

A multimodal approach for determining brain networks by jointly modeling functional and structural connectivity

Wenqiong Xue^{1*}, F. DuBois Bowman², Anthony V. Pileggi³ and Andrew R. Mayer^{4,5}

¹ Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT, USA

² Department of Biostatistics, Mailman School of Public Health, Columbia University, New York, NY, USA

³ Department of Biostatistics and Bioinformatics, Rollins School of Public Health, Emory University, Atlanta, GA, USA

⁴ The Mind Research Network/Lovelace Biomedical and Environmental Research Institute, Albuquerque, NM, USA

⁵ Neurology Department, University of New Mexico School of Medicine, Albuquerque, NM, USA

Edited by:

Qawi K. Telesford, Wake Forest University School of Medicine, USA

Reviewed by:

Ani Eloyan, Johns Hopkins University, USA

Mary Beth Nebel, Johns Hopkins School of Medicine, USA

*Correspondence:

Wenqiong Xue, Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, 900 Ridgebury Road, Ridgefield, CT 06877, USA
 e-mail: nora.xue@gmail.com

Recent innovations in neuroimaging technology have provided opportunities for researchers to investigate connectivity in the human brain by examining the anatomical circuitry as well as functional relationships between brain regions. Existing statistical approaches for connectivity generally examine resting-state or task-related functional connectivity (FC) between brain regions or separately examine structural linkages. As a means to determine brain networks, we present a unified Bayesian framework for analyzing FC utilizing the knowledge of associated structural connections, which extends an approach by Patel et al. (2006a) that considers only functional data. We introduce an FC measure that rests upon assessments of functional coherence between regional brain activity identified from functional magnetic resonance imaging (fMRI) data. Our structural connectivity (SC) information is drawn from diffusion tensor imaging (DTI) data, which is used to quantify probabilities of SC between brain regions. We formulate a prior distribution for FC that depends upon the probability of SC between brain regions, with this dependence adhering to structural-functional links revealed by our fMRI and DTI data. We further characterize the functional hierarchy of functionally connected brain regions by defining an ascendancy measure that compares the marginal probabilities of elevated activity between regions. In addition, we describe topological properties of the network, which is composed of connected region pairs, by performing graph theoretic analyses. We demonstrate the use of our Bayesian model using fMRI and DTI data from a study of auditory processing. We further illustrate the advantages of our method by comparisons to methods that only incorporate functional information.

Keywords: functional connectivity, structural connectivity, Bayesian analysis, small-world network, MCMC, fMRI, DTI, neuroimaging

1. INTRODUCTION

There has been a recent emergence of complex functional brain network analyses. A typical neuroimaging network analysis involves defining brain regions (nodes) representing the whole brain, quantifying a measure of association between all pairs of brain regions to produce a connectivity matrix, thresholding these associations to obtain a more sparse connectivity matrix, and then calculating summary statistics that reflect properties of the network. Network summaries include metrics that reflect local or global communication ability (e.g., clustering coefficient, path length, and efficiency), centrality metrics (e.g., betweenness, closeness, and eigenvector centrality), and community structure (e.g., small-worldness).

A critical step in network analyses is that of quantifying associations, or functional connectivity (FC), between every pair of nodes because it fundamentally impacts the integrity, reliability, and interpretability of the resulting networks and their associated properties (Seghier and Friston, 2013; Ramsey et al., 2014). Friston et al. (1993) define FC as the “temporal correlations

between spatially remote neurophysiological events.” This non-directional association may pertain to resting-state brain activity or to neural activity stemming from cognitive, emotional, visual, and behavioral tasks. In addition, estimation of directed connectivity is of great interest to describe graph theoretic properties in larger networks (Seghier and Friston, 2013). In contrast, structural connectivity (SC) refers to the structural white-matter fiber tracts (bundles of axons) linking different brain regions, which may be considered to identify structural brain networks. SC provides a mechanism for functional relationships in neural activity by enabling the transmission of electrical signals that pass along axons (Hendelman, 2000). Yet, FC and SC properties typically are evaluated separately.

The Pearson correlation coefficient and partial correlation are popular statistics used to describe nodal FC associations (Hampson et al., 2002; Marrelec et al., 2006). Some researchers use alternative measures of association, e.g., mutual information (Cover and Thomas, 1991), or consider transformations of the time series data, e.g., to examine associations in the frequency

domain (Fiecas et al., 2010) or in the wavelet domain (Patel et al., 2006b). Patel et al. (2006a) develop a Bayesian model that assesses the FC based on the joint activation in pairs of brain regions. SC is based on diffusion tensor imaging (DTI), which is a magnetic resonance imaging (MRI) technique for characterizing the integrity of axonal fibers by measuring molecular diffusion (Kollias, 2009). Probabilistic diffusion tensor tractography (DTT) uses DTI data to empirically reconstruct fiber tracts by quantifying the likelihood of white-matter SC (Behrens et al., 2007).

A notable limitation of the aforementioned procedures for determining FC is that they do not consider any information about the associated SC, which may adversely impact the integrity of the networks used for complex whole-brain network analyses. Previous studies have suggested that structural fiber properties (degree of myelination and conduction times) contribute to FC amongst homolog brain regions (Stark et al., 2008; Zuo et al., 2010; Gee et al., 2011), but do not solely determine FC (Honey et al., 2009). There have been a few attempts to examine both FC and SC. These studies generally either examine the correspondence between SC and localized (voxel-level) analysis of functional MRI (fMRI) data or assess SC and FC sequentially and use SC to guide region selection for FC evaluation (or vice versa) (Rykhlevskaia et al., 2008). From such analyses, Morgan et al. (2009) suggest that FC is supported by SC along the language pathways. Also, Greicius et al. (2009) and van den Heuvel et al. (2009) indicate that resting-state FC reflects SC to a large degree. Bowman et al. (2012) present a framework that simultaneously considers fMRI and DTI data to determine FC, and they demonstrate that the supplemental SC information is particularly beneficial in the presence of fMRI noise. As the association between brain structure and function is revealed, an important next step is to develop unified, model-based statistical frameworks that incorporate both sources of information simultaneously. We feel that such methods will in turn improve the validity of complex network analyses.

We present a novel multimodal approach, using FC and supplementary SC information, to determine FC as the basis for defining and evaluating whole-brain networks. We extend the previously developed model by Patel et al. (2006a) that determines FC by examining the concurrence of elevated activity in pairs of brain locations. Our Bayesian model utilizes DTT information as a supplement to fMRI, and here we apply our model to evaluate FC globally between all pairs of defined regions of interest. We determine the hierarchy among functionally connected pairs of brain regions based on the associated probabilities of elevated activity for each node, giving rise to directed networks. We develop formal inference frameworks regarding task-related functional coherence for both our undirected and directed measures of association. Using the inference frameworks as a basis for thresholding, we build a functional network based on region pairs that are connected with high probability, and we explore the associated topological properties from the graphical network. We perform estimation using Markov Chain Monte Carlo (MCMC) techniques. We apply our method to an auditory spatial-cueing task data set and conduct simulation studies to evaluate the performance of our approach.

2. DATA

2.1. EXPERIMENTAL DATA

Combined structural and functional data were collected from 32 right-handed adults (15 males, 17 females) on a 3 Tesla Siemens scanner. The functional data were collected during an auditory spatial-cueing task, which has been described in previous publications (Mayer et al., 2009a,b). The auditory stimuli consist of a series of two pure frequency tone pips per trial. The first auditory tone (2000 Hz) served as a spatial cue, and occurred with equal frequency in the left or the right ear. The second tone (1000 Hz) served as a target, and participants were instructed to press either their right index or middle finger for targets appearing in their left or right ear, respectively. Cues correctly predicted (i.e., valid trials) the location of the targets on 50% of the trials, with invalid trials (i.e., cue and target in opposite ears) occurring during the other 50% of trials. The stimulus onset asynchrony (SOA) between the cues and the targets was either 200 or 700 ms. Prior to the start of the experiment, participants were informed that the cues did not contain any useful information to predict the location of the target. Therefore, the cue validity ratio (50%) and task instructions were designed to evoke exogenous orienting conditions, with SOA designed to respectively measure facilitation (200 ms SOA) or inhibition of return (700 ms SOA). Participants were also required to maintain fixation on a centrally presented cross-hair to reduce the likelihood of eye movements. A total of 84 valid and 84 invalid trials were presented in pseudo-random order across three imaging runs. The main objectives considered here are to identify the functionally connected brain regions associated with the neural processing stemming from an auditory task (both valid and invalid trials), to determine underlying auditory networks, and to assess the topological properties of the networks, although a broader set of objectives has been previously considered (Mayer et al., 2009a,b).

2.2. IMAGE ACQUISITION AND DATA PREPROCESSING

High resolution 5-echo multi-echo MPRAGE T1 [TR (repetition time) = 2.53 s, 7° flip angle, number of excitations (NEX) = 1, slice thickness = 1 mm, FOV (field of view) = 256 mm, resolution = 256 × 256] and T2 [echo time = 77.0 ms, TR = 1.55 s, flip angle 152°, NEX = 1, slice thickness = 1.5 mm, FOV = 220 mm, matrix = 192 × 192, voxel size = 1.15 × 1.15 × 1.5 mm³] sequences were collected on a 3 Tesla Siemens Trio scanner. Echo-planar images (EPI) were collected using a single-shot, gradient-echo echoplanar pulse sequence [TR = 2000 ms; TE = 29 ms; flip angle = 75°; FOV = 240 mm; matrix size = 64 × 64]. A total of $T^* = 483$ fMRI scans were collected for each of the $N = 32$ subjects, 161 for each of three runs, after eliminating the first scan of each run to account for equilibrium effects. Two DTI scans with $b = 800$ s/mm² and 30 diffusion gradients were acquired using a twice-refocused spin echo sequence to reduce the effects of eddy currents and artifacts associated with head movement and to allow increased time for diffusion sensitizing gradients. Additional DTI scanning parameters include the following: 72 interleaved slices, TE = 84 ms, TR = 9 s, flip angle = 90, slice thickness = 2.0 mm, FOV = 256 × 256 mm, matrix size = 128 × 128, voxel resolution = 2 × 2 × 2 mm³ (Ling et al., 2012).

We performed several standard preprocessing steps to the functional images using FMRIB (Functional Magnetic Resonance Imaging of the Brain) Software Library (FSL) (Smith et al., 2004). These steps included slice-timing correction, 3D motion correction and spatial normalization to Montreal Neurological Institute (MNI) space using a 12 degrees of freedom affine linear transformation with trilinear interpolation. Pre-whitening was conducted to remove the temporal correlations between scans from the same subject by iteratively estimating the autocorrelation matrix of the residuals to achieve independence through the whole time series (Woolrich et al., 2001).

2.3. DETERMINING REGIONAL ACTIVITY

We consider region-to-region connectivity in our network analysis. To define the brain regions, we begin with 90 regions from the automated anatomic labeling (AAL) system (Tzourio-Mazoyer et al., 2002) excluding the cerebellum. We further refine the regions by applying parcellation to obtain subregions within each AAL region as described in Appendix 1, which yields a total of 205 subregions. Each subregion that we consider contains more than 50 voxels. We identify the voxel within each subregion that is most involved with the auditory task, determined on the basis of a standardized statistic calculated by subtracting the mean and dividing by the standard deviation of each voxel time course. We summarize the neural activity for each defined subregion by selecting the roughly 150 closest voxels (in Euclidean distance) to the voxel most involved with the task, ensuring that these closest voxels all fall within the subregion. If there are fewer than 150 voxels in the subregion, we use the entire subregion. The area of the 150 voxels generally corresponds to a roughly spherical shape with a 6 mm radius, although we do not strictly require a spherical shape, e.g., to address boundary constraints. To obtain a single fMRI temporal profile representing each subregion for each subject, we perform a singular value decomposition (SVD) in the time domain to a $T^* \times 150$ matrix. We extract the first right singular vector, yielding a single temporal profile reflecting the most dominant temporal trend in that subregion. Since singular vectors are unique up to multiplication by a unit phase factor, we compare the singular vector to the subregion mean signal to ensure that the selected signal represents the subregion correctly, and we apply a sign change to our extracted signal, if necessary. The resulting 205 temporal profiles reflect the neural activity representing each node in our network analysis.

2.4. DETERMINING STRUCTURAL CONNECTIVITY

We employ the widely used approach of Behrens et al. (2007), implemented in FSL, to perform probabilistic DTT. We define subregions within the AAL system for DTT, which are centered in white matter proximal to the fMRI-based subregions. Probabilistic tractography successively initiates streams, which are intended to follow or trace the paths of white matter tracts in the brain. A given number of streams (5000 in our analysis) are sent from the seed voxel, and each stream chooses a path based on the principle diffusion direction of the underlying white matter at each voxel and ceases according to a stopping rule. The probabilistic DTT for each pair of regions initially yields voxel-level counts (out of 5000 trials) indicating the likelihood of a fiber tract extending from the voxel in the seed region to (or through) the

target region. We use the 90th percentile of the voxel-level counts connecting voxels in the seed region to voxels in the target region to reflect the strongest anatomical connectivity between pairs of regions. The voxel-level counts connecting two regions are usually asymmetric; yet for our purpose, we regard FC as a symmetric measure. Therefore, we impose symmetry of SC between two regions by calculating the maximum of the two directional counts for each region pair. To reduce the noise, we analyze both DTI scans and average the resulting SC counts. We adjust the SC counts by the corresponding geometric distances between regions by fitting a zero-inflated Poisson regression model on voxel-level SC counts adjusted for the minimum geometric distance between regions.

3. METHODS

We develop a framework for network analysis, which jointly considers FC and SC. We introduce a statistic κ to capture the functional coherence between region pairs and an associated ascendancy measure τ to quantify the hierarchy of identified coherent regions. The way that we describe functionally coherent brain regions is not a typical definition of FC. The joint activation framework is defined conceptually by Patel et al. (2006a). This method performs well-relative to other methods for FC and comparisons conducted by Smith et al. (2010) and by Ramsey et al. (2014), which considers performance after removing and replacing the high-pass filter by a less stringent filter. A preliminary look at our data reveals that higher levels of SC counts tend to have associated larger values of functional coherence (see Figure 1). Specifically, we examine the distribution of functional coherence at lower and higher levels of voxel-level SC for each subject across 20910 region pairs based on 205 brain regions. Figure 1 illustrates the results for selected subjects, but the results across all the subjects reveal similar patterns. Therefore, we build our Bayesian model based on the observation that increased SC is generally associated with higher functional coherence. SC is a static property, whereas FC is a transient characteristic that may vary with the performance of different tasks. Therefore, we do not make the link too strong in our model since high SC may exist without corresponding elevated FC during the auditory task.

3.1. JOINT ACTIVATION AND STRUCTURAL CONNECTIVITY

Define $A_{gnt} = I(R_{gnt} > c \times \sigma_{gn}^2)$, where $R_{gnt} = Y_{gnt} - \hat{\mu}_{gn}$ is the mean-adjusted level of neural activity for region g , subject n , and scan t ; c is a constant; $\hat{\mu}_{gn}$ is the mean of the subregion level fMRI profile Y_{gnt} across time, and σ_{gn}^2 is the variance of Y_{gnt} , with $n = 1, \dots, N$ and $t = 1, \dots, T^*$. Thus, A_{gnt} serves as an indicator of elevated regional brain activity at time t . We choose $c = 0.01$ when analyzing the auditory spatial-cueing task.

The joint activation between two regions a and b for subject n can be expressed as:

$$\begin{aligned} Z_1^* &= \sum_{t=1}^{T^*} I(A_{ant} = 1, A_{bnt} = 1), & Z_2^* &= \sum_{t=1}^{T^*} I(A_{ant} = 1, A_{bnt} = 0) \\ Z_3^* &= \sum_{t=1}^{T^*} I(A_{ant} = 0, A_{bnt} = 1), & Z_4^* &= \sum_{t=1}^{T^*} I(A_{ant} = 0, A_{bnt} = 0). \end{aligned} \quad (1)$$

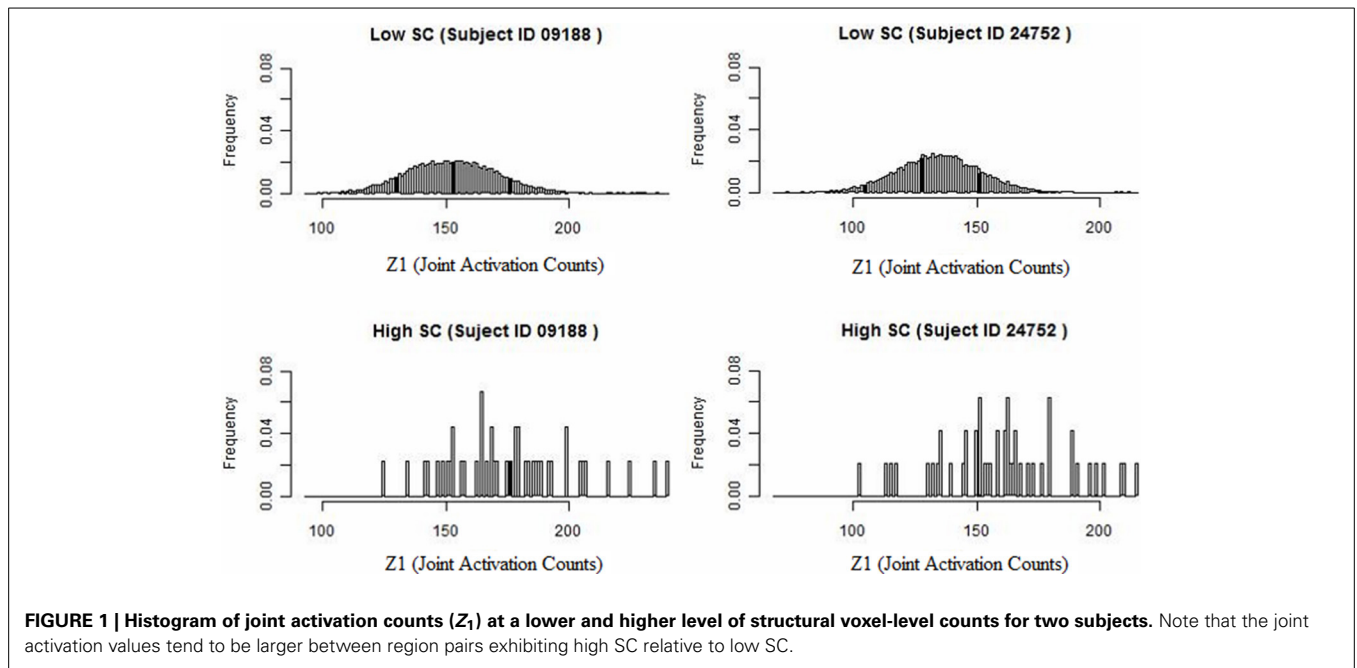


FIGURE 1 | Histogram of joint activation counts (Z_1) at a lower and higher level of structural voxel-level counts for two subjects. Note that the joint activation values tend to be larger between region pairs exhibiting high SC relative to low SC.

Z_1^* is interpreted as the number of times that both regions a and b experience an elevated fMRI signal (for subject n). Note that we omit subscript indexing subjects to simply our notation. We assume $Z^* = (Z_1^*, \dots, Z_4^*)'$ follows a multinomial distribution with parameters T^* and $\theta = (\theta_1, \dots, \theta_4)'$, where

$$\begin{aligned} \theta_1 &= P(A_{ant} = 1, A_{bnt} = 1), \quad \theta_2 = P(A_{ant} = 1, A_{bnt} = 0) \\ \theta_3 &= P(A_{ant} = 0, A_{bnt} = 1), \quad \theta_4 = P(A_{ant} = 0, A_{bnt} = 0). \end{aligned} \tag{2}$$

To facilitate interpretations across different analyses, we standardize Z_i^* ($i = 1, \dots, 4$) by scaling it with a specified number of scans. We set the scaling unit to $T = 100$ for our data, so our standardized measure, Z_i , is the average number of times that a and b are coherent per one hundred scans. Z also follows a multinomial distribution with parameters T and θ .

For the anatomical data, we denote the DTT counts between regions a and b for subject n by S^* (omitting the subject and region subscripts for simplicity). Let M^* be the number of trials for probabilistic DTT fiber tracking from the voxels in the seed region. We assume that S^* follows a binomial distribution with parameters M^* and π , where π is the probability of SC between regions a and b for any subject. Using similar scaling applied to Z_i^* , we generate scaled counts S , which follow a binomial distribution with parameters M and π , where here we choose $M = 1000$.

To estimate the model parameters, we perform 10,000 iterations preceded by 2000 burn-in iterations. The programming is implemented in Matlab, and the computation is performed on a Linux cluster with 16 GB of RAM. Execution time is approximately 3–4 h. Additional gains in computation time may be achievable by parallel programming. In simulation studies, 7000 iterations preceded by 2000 burn-in iterations are conducted for each simulated dataset. For both our data analysis and simulations, we implement a thinning procedure by retaining every tenth iteration from the MCMC chain

Table 1 | Joint activation probabilities for regions a and b .

		Region a		
		Active	Inactive	
Region b	Active	θ_1	θ_3	$\theta_1 + \theta_3$
	Inactive	θ_2	θ_4	$\theta_2 + \theta_4$
		$\theta_1 + \theta_2$	$\theta_3 + \theta_4$	1

for easy storage and for promoting independence between the samples. A total of 10,000 datasets are generated for each combination of hyperparameters and each job is completed within 1 day.

3.2. FUNCTIONAL COHERENCE AND ASCENDANCY

We extend the agreement measure of Patel et al. (2006a), which evaluates joint activation only (see Appendix 2) based on Cohen’s Kappa (Cohen, 1960), to describe functional coherence between pairs of brain regions. Considering Table 1, our functional coherence measure κ is defined as:

$$\kappa = \begin{cases} \frac{\theta_1 + \theta_4 - E}{1 - E} & \text{if } \theta_1\theta_4 > \theta_2\theta_3 \\ 0 & \text{otherwise,} \end{cases} \tag{3}$$

where $E = (\theta_1 + \theta_2)(\theta_1 + \theta_3) + (\theta_3 + \theta_4)(\theta_2 + \theta_4)$. The numerator of κ measures the difference between the probability of coherence and the expected probability of coherence under independence. We restrict our attention to non-negative values of κ , so our measure of agreement ranges from 0 to 1. κ equals 1 when the probability of joint activation and deactivation θ_1 and θ_4 sums to 1, and hence θ_2 and θ_3 are 0, which indicates complete coherence. If there is no agreement between regions a and b , $\kappa = 0$.

Given that a and b are functionally connected, i.e., κ exceeds a specified threshold (say e_κ) with high probability, we define a measure of ascendancy to determine the hierarchical relationship between

the regions. Unlike the definition in Patel et al. (2006a), in which the definition of ascendancy is based on the ratio of $P(A_a = 1)$ and $P(A_b = 1)$ (see Appendix 2), our ascendancy measure, τ_{ab} , is based on the ratio of $P(A_a = 1)/(1 - P(A_a = 1))$ and $P(A_b = 1)/(1 - P(A_b = 1))$ and takes the following form:

$$\tau_{ab} = \left(\frac{\theta_1 + \theta_2}{\theta_3 + \theta_4} \right) / \left(\frac{\theta_1 + \theta_3}{\theta_2 + \theta_4} \right). \quad (4)$$

τ_{ab} ranges from 0 to ∞ and is interpreted as the odds of region a being active relative to the odds of region b being active. If two regions a and b become active together and inactive together, we consider them as functionally connected; however, an ascendancy relationship does not necessarily exist between them. Given that region a and region b are functionally connected, we say that a is ascendant to b whenever the marginal odds of activation of a are larger than that of b . Given $\kappa > e_\kappa$, $\tau_{ab} > 1$ indicates that a is ascendant to b , while $\tau_{ab} < 1$ indicates that b is ascendant to a . **Figure 2** presents a hierarchical relationship among four regions. We use shading to denote a region exhibiting elevated activity. As shown in **Figure 2**, while region a is functionally connected to regions b , c , and d (represented by solid lines in the graph), regions c and d exhibit elevated activity for a strict subset of the stimuli which a exhibits elevated activity, suggesting that a is ascendant to c and d .

3.3. BAYESIAN STATISTICAL MODEL

For any pair of regions a and b , the likelihood function takes the form:

$$p(\mathbf{Z}, S|\theta, \pi) \propto \prod_{i=1}^4 \theta_i^{\sum_{n=1}^N Z_{in}} \pi^{\sum_{n=1}^N S_n} (1 - \pi)^{N \times M - \sum_{n=1}^N S_n}, \quad (5)$$

where the proportional sign (\propto) defines an equivalent relation up to a multiplicative constant. Following the approach by Patel et al. (2006a), we assume that each repeated measure on the same region pair is independent across subjects. We also assume that each repeated measure on the same region pair is independent over time since we have performed pre-whitening in our pre-processing to remove the temporal correlations between scans from the same subject. In addition, given both probability measurements θ and π , S is independent of \mathbf{Z} because we build structure-function dependence in the distribution of $[\theta|\pi]$. This is a conditional independence assumption between S and \mathbf{Z} , but marginally our model still captures the dependence between S and \mathbf{Z} through the corresponding parameters θ and π .

Using a Bayesian formulation, we express our prior belief about structural connection probabilities π by defining a beta prior which takes the form:

$$p(\pi) \propto \pi^{\alpha_0 - 1} (1 - \pi)^{\beta_0 - 1}. \quad (6)$$

We choose the beta distribution as the prior both for mathematical convenience, as it is the conjugate prior for the binomial distribution, and for its flexibility to implement priors with different shapes of the density using various hyper prior parameter specifications.

We specify a flat prior for our DTT data by setting $\alpha_0 = \beta_0 = 1$ for each region pair, suggesting no available prior information regarding the SC of any region pair. In our simulation studies, we evaluate the performance of our method under different combinations of the hyperparameters α_0 and β_0 .

The prior for θ is taken to follow a Dirichlet distribution with parameters $(\alpha(\pi) + \alpha_1, \alpha_2, \alpha_3, \alpha_4)'$, where $\alpha(\pi)$ is a function of π and reflects the assumed relationship between FC and SC. Specifically, we assume that

$$p(\theta|\pi) \propto \frac{\Gamma(\alpha(\pi) + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)}{\Gamma(\alpha(\pi) + \alpha_1)} \times \theta_1^{\alpha(\pi) + \alpha_1 - 1} \theta_2^{\alpha_2 - 1} \theta_3^{\alpha_3 - 1} \theta_4^{\alpha_4 - 1}. \quad (7)$$

We set $\alpha_1 = 5, \alpha_2 = \alpha_3 = \alpha_4 = 10$, and $\alpha(\pi) = 10 / \left(\frac{9}{\ln(10)} - 1 \right) \times 10^\pi - 10 / \left(\frac{9}{\ln(10)} - 1 \right)$, so the average value of $\alpha(\pi)$ on $\pi \in [0, 1]$ is 10. In this case, $\theta_2, \dots, \theta_4$ have the same expected values and moderate variances. Our prior is based on the observation that weak SC corresponds to relatively few joint functional activations, and extremely strong SC is assumed to yield an expected value of θ_1 to be around 0.5, which is approximately the highest maximum likelihood estimate of θ_1 . When $\alpha(\pi)$ is an increasing function, the expected value of θ_1 , which takes the form of $(\alpha(\pi) + \alpha_1) / (\alpha(\pi) + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)$ is also an increasing function with respect to π ; thus, matching our observation from the data. Later, we present results from a sensitivity analysis of our choice of $\alpha(\pi)$, which shows that our results do not change much with respect to different functions of $\alpha(\pi)$.

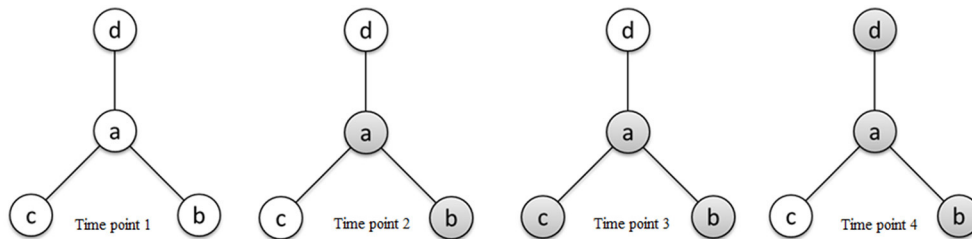


FIGURE 2 | Schematic diagram of activity states of four regions a, b, c , and d at different time points and connectivity relationships between the regions reflecting functional connectivity and ascendancy. Shading for a given region indicates elevated activity at the corresponding time point.

The line segments connecting regions define functionally connected region pairs, illustrating that region a is functionally connected to regions b, c , and d . Region a is ascendant to regions c and d based on the relative levels of elevated activity between functionally connected region pairs.

3.4. SAMPLING THE JOINT POSTERIOR PROBABILITY DISTRIBUTION

We express the posterior probability distribution from our Bayesian model as:

$$p(\theta, \pi | \mathbf{Z}, S) \propto \frac{\Gamma(\alpha(\pi) + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)}{\Gamma(\alpha(\pi) + \alpha_1)} \quad (8)$$

$$\times \pi^{\sum_{n=1}^N S_n + \alpha_0 - 1} (1 - \pi)^{N \times M - \sum_{n=1}^N S_n + \beta_0 - 1}$$

$$\times \theta_1^{\alpha(\pi)} \prod_{i=1}^4 \theta_i^{\sum_{n=1}^N Z_{in} + \alpha_i - 1}.$$

The full conditional posterior distributions take the form:

$$p(\pi | \theta, \mathbf{Z}, S) \propto \frac{\Gamma(\alpha(\pi) + \alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)}{\Gamma(\alpha(\pi) + \alpha_1)}$$

$$\times \pi^{\sum_{n=1}^N S_n + \alpha_0 - 1} (1 - \pi)^{N \times M - \sum_{n=1}^N S_n + \beta_0 - 1} \theta_1^{\alpha(\pi)} \quad (9)$$

and

$$p(\theta | \pi, \mathbf{Z}, S) \sim \text{Dirichlet} \left(\sum_{n=1}^N Z_{1n} + \alpha(\pi) + \alpha_1, \sum_{n=1}^N Z_{2n} + \alpha_2, \right.$$

$$\left. \sum_{n=1}^N Z_{3n} + \alpha_3, \sum_{n=1}^N Z_{4n} + \alpha_4 \right). \quad (10)$$

Estimation of κ and τ are based on the posterior distribution, $p(\theta | \mathbf{Z}, S)$, as we are able to estimate $p(\kappa | \mathbf{Z}, S)$ and $p(\tau | \mathbf{Z}, S)$ by sampling $p(\theta, \pi | \mathbf{Z}, S)$. We set effect sizes for κ , denoted e_κ , and τ , denoted e_τ , in our analysis to compute exceedance probabilities $P(\kappa > e_\kappa | \mathbf{Z}, S)$ and $P(\tau > e_\tau | \mathbf{Z}, S)$ from our modeling framework. For κ , we choose $e_\kappa = 0.4$, which reflects moderate agreement or above (Landis and Koch, 1977). As no standard has been established for hierarchy, we explore the ascendancy relationship when τ is above 50th percentile across all region pairs. We conduct inference on κ and τ by estimating $P(\kappa > e_\kappa | \mathbf{Z}, S) > p_\kappa$ and $P(\tau > e_\tau | \mathbf{Z}, S) > p_\tau$, respectively. We set both p_κ and p_τ to 0.5 to capture the most information. The choices of effect sizes and probability thresholds are made to reflect characteristics of functional coherence and ascendancy, rather than on the basis of statistical properties. Ideally, these thresholds should be determined before performing the data analysis. However, due to the complexity and variability of the neuroimaging data, the user has the flexibility to investigate connections between regions at different levels of these thresholds.

As revealed by the conditional posterior distribution of θ , our prior belief has an impact on the posterior through $\alpha(\pi)$, but does not drive the direction of results. As the SC π becomes stronger, the expected value of θ_1 increases, and the expected values of $\theta_2, \dots, \theta_4$ decrease accordingly, but at a slower rate.

We estimate our Bayesian hierarchical model using MCMC methods, implemented via the Gibbs sampler with an embedded Metropolis step. The parameter θ is updated from a Dirichlet distribution with π specified from the previous step. The parameter π is updated by π^* , which is sampled from a Normal jumping distribution $J_t(\pi^* | \pi^{t-1}) = N(\pi^* | \pi^{t-1}, \tau^2)$ at time t , with probability $\min(r, 1)$, where r is the ratio of the conditional densities from

time $t - 1$ to the conditional densities of the proposed value with respect to θ ,

$$r = \frac{p(\pi^* | \theta, \mathbf{Z}, S)}{p(\pi^{t-1} | \theta, \mathbf{Z}, S)}. \quad (11)$$

The variance τ^2 in the Normal jumping distribution is adjusted based on the data, which ensures the acceptance ratio close to 25% to achieve the optimal efficiency of Metropolis algorithm Gelman et al. (1995).

3.5. ASSESSING BRAIN NETWORKS

We build both a undirected and a directed brain network based on the coherence and ascendancy relationships between regions, respectively, and perform graph theoretic analyses to demonstrate topological properties of the brain networks. Suppose there are s brain regions and t functionally connected region pairs in the network, which are represented by s nodes and t edges, respectively. In a directed network, if region a is ascendant to region b , the edge between them is directed from region a to region b .

Network topology is described as a small-world network if compared to a similar random network, the small-world index $\sigma = (C/C_{\text{random}})/(L/L_{\text{random}}) > 1$ (Watts and Strogatz, 1998; Humphries et al., 2006). Here, a similar random network is defined as a network with the same number of nodes, the same number of edges, and the same degree distribution (Simpson et al., 2013). The clustering coefficient C measures the average likelihood of connecting neighbors. The clustering coefficient is defined as $C_i = 2E_i/k_i(k_i - 1)$ for undirected network, and $C_i = E_i/k_i(k_i - 1)$ for directed network for each node i , where k_i is the number of neighbors of node i and E_i is the number of direct links connecting neighbors of node i . And the path length L is the average minimum number of connections to link two nodes.

We also examine the hubs in the network as they play a central role in a network since they serve as the common connections to other nodes. We define hubs as the nodes with degree (D), out-degree (D_{out}) or in-degree (D_{in}) at least one standard deviation above the mean degree of all the nodes in the network (Sporns et al., 2007). For directed networks, driving hubs satisfy $D_{\text{out}} > \text{mean} + \text{SD}$ and driven hubs satisfy $D_{\text{in}} > \text{mean} + \text{SD}$. Roughly speaking, hubs are regions that exhibit numerous connections to other regions, with driving hubs, being relatively more active than the regions to which they are connected; and driven hubs, being relatively less active than the regions to which they are connected.

4. RESULTS

4.1. AUDITORY DATA RESULTS

We apply our Bayesian model to the auditory processing fMRI data to determine functionally connected regions and also to examine the corresponding coherence and ascendancy relationship to illustrate the neural integration underlying auditory and related processing. We find strongly connected regions within the auditory cortex with associated hierarchical relationships between these functionally connected pairs of regions.

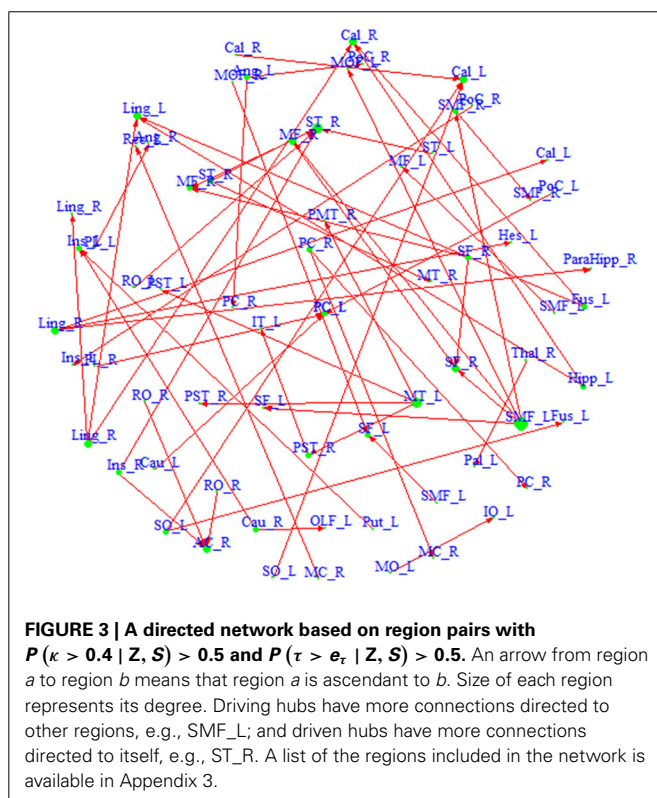
We identify an undirected network based on the region pairs with $P(\kappa > e_\kappa | \mathbf{Z}, S) > p_\kappa$. We determine the hubs in the network using the criteria described in Section 3.5. A total of 27 hubs are found, including brain regions within the auditory cortex, visual cortex, motor cortex, and sensory regions. For example, as an important region in the auditory cortex, the left superior temporal gyrus plays a central role in the undirected network. Also, the bilateral Rolandic

operculum, which is one of the auditory-speech encoding regions, shares many connections with other regions, which may reflect the use by many people of language encoding areas during auditory processing involving perception of the spatial location of a tone (Klatzky et al., 2002). The design of the spatial-cueing auditory experiment calls for subjects to momentarily remember the location of the target tone before they press the button. Many people make use of visualization to aid the memory process, especially spatial memory, which may in turn activate the visual cortex (Ungerleider, 1995). Our findings support this relationship by revealing the prominent roles of superior parietal gyrus, the left precuneus, and other regions within the visual cortex in the network. In addition, our results reflect the neural processing related to button presses in response to the target cue by identifying the bilateral supplementary motor area, the right precentral gyrus, and the left postcentral gyrus as hubs, all of which are included in motor and sensory systems. We further examine other graph theoretic properties of the brain network associated with the auditory task. The clustering coefficient C of this undirected network is 0.02, which reflects a low likelihood of connected neighbors in the network. However, the path length L is 3.24, which indicates that it takes an average of only three intermediate connections to link any pair of regions in the network. The identified undirected network has the small-world property with the small-world index $\hat{\sigma} = 35.16$ compared to the average of 1000 similar random networks ($C_{random} = 8.04 \times 10^{-4}$, $L_{random} = 4.58$).

We also build a directed network based on the region pairs with $P(\kappa > e_\kappa | \mathbf{Z}, \mathbf{S}) > p_\kappa$ and $P(\tau > e_\tau | \mathbf{Z}, \mathbf{S}) > p_\tau$ (Figure 3). Similarly, both driving hubs and driven hubs are identified. Here, we focus on the driving hubs, since they are the regions that play more central roles when connecting to others. With the constraint

of ascendancy, substantially fewer hubs are found compared to the undirected network. We identify a total of seven driving hubs, including the right Rolandic Operculum, the left olfactory gyrus, the left supplementary motor area, the right middle cingulate gyrus, the left superior occipital gyrus, the right middle temporal gyrus and the left superior medial frontal gyrus (component figures for each driven hub are shown in Appendix 4). These driving hubs are more active than the other regions to which they are functionally connected. The clustering coefficient C of the directed network is 0.26, and the path length L is 4.17. The directed network has higher probability of connecting neighbors comparing to undirected network, and it takes about four steps to connect any pair of regions within the network. Compared to the average of 1000 similar random networks ($C_{random} = 0.02$, $L_{random} = 4.71$), the identified directed network has the small-world index $\hat{\sigma} = 12.93$, which indicates that it demonstrates small-worldness.

For comparison, we apply the approach of Patel et al. (2006a) to fMRI data from the auditory spatial cueing task. Their method only utilizes functional imaging data and is limited to using joint activation to define connectivity and ascendancy. The Patel method fails to identify the left superior temporal gyrus as a hub in the undirected network, which is an important region within the auditory cortex. Another major difference is that more region pairs within the visual cortex are detected by the original approach. Although visualization is an essential component of the experimental design, we do not expect to see the high degree of connectivity within the visual cortex because the task only involves fixation on a cross hair, with more engaging auditory processing. This leads us to conjecture that—without structural information, more false positives are generated while the major findings may go undetected.



4.2. SIMULATION RESULTS

We conduct two simulation studies to compare our approach to alternative methods that solely consider FC, without regard for underlying SC. Specifically, we compare our method to the approach of Patel et al. (2006a) and to a traditional correlation-based analysis. In addition, we conduct an analysis to evaluate the sensitivity of our approach to the impact of various choices of the function $\alpha(\pi)$, which links FC and SC.

We first discuss simulation results from the comparison our combined fMRI and DTI approach (labeled as the *FC with SC* method) with the approach by Patel et al. (2006a) utilizing fMRI data only (labeled as the *FC only* method). We generate data \mathbf{Z} and \mathbf{S} from our model with different settings of hyperparameters α_0 and β_0 in the prior distribution of π . These prior parameters control the shape of the density function and provide a flexible range of possible prior distributions. We evaluate the methods by comparing the bias of the corresponding posterior means of θ , κ and τ . For each simulation setting, 10π 's are generated, and 10θ 's are simulated for each π , 100 data sets are generated from each set of θ . Therefore, a total of 10,000 data sets are simulated to compute the mean biases. The results indicate that our method performs better in all settings with smaller bias. Table 2 shows that the mean bias of θ_i 's from our method is smaller than that from the *FC only* method in every case, which indicates that incorporating structural information improves the estimation of FC. Since our definitions of κ and τ are both functions of θ_i 's, it follows that our model also outperforms the *FC only* method for estimating these measures as defined in (3) and (4). Alternatively, we compare estimation performance of κ and τ from the *FC only* method

Table 2 | Comparison of mean of bias between two Bayesian methods.

α_0	β_0	$E(\pi)$	FC with SC($\times 10^{-3}$)						FC only($\times 10^{-3}$)					
			θ_1	θ_2	θ_3	θ_4	κ	τ	θ_1	θ_2	θ_3	θ_4	κ	τ
N = 15														
1	100	0.01	6.882	8.906	9.006	9.078	2.874	44.578	6.951	9.011	9.154	9.183	221.313	106.883
2	18	0.1	7.39	9.073	8.851	8.822	4.442	42.553	7.461	9.162	8.993	8.914	204.43	122.885
2	5	0.3	8.565	8.861	8.544	8.58	8.545	41.804	8.667	9.038	8.636	8.689	120.509	126.061
2	2	0.5	9.239	8.333	8.484	8.499	13.308	41.421	9.297	8.44	8.63	8.644	71.363	128.054
5	2	0.7	9.798	8.146	8.086	8.196	16.415	40.797	9.862	8.301	8.174	8.316	59.667	136.606
18	2	0.9	10.056	7.379	7.52	7.313	19.624	41.791	10.233	7.503	7.68	7.457	88.121	136.543
N = 30														
1	100	0.01	4.925	6.550	6.394	6.391	2.185	30.759	4.963	6.587	6.432	6.416	233.875	116.591
2	18	0.1	5.127	6.361	6.282	6.469	2.791	29.577	5.173	6.427	6.307	6.508	188.989	117.824
2	5	0.3	5.979	6.292	6.206	6.211	5.702	29.458	6.017	6.330	6.247	6.247	116.983	127.239
2	2	0.5	6.271	6.044	6.045	5.952	7.556	30.224	6.321	6.106	6.077	5.998	112.727	119.231
5	2	0.7	7.007	5.777	5.653	5.692	12.085	29.761	7.078	5.829	5.711	5.756	57.719	123.946
18	2	0.9	7.150	5.387	5.330	5.390	13.531	28.887	7.242	5.434	5.376	5.442	80.766	139.150
N = 100														
1	100	0.01	2.704	3.536	3.597	3.534	1.120	17.098	2.710	3.541	3.597	3.545	232.029	115.003
2	18	0.1	2.996	3.483	3.481	3.532	2.103	16.705	3.006	3.492	3.485	3.538	181.038	114.898
2	5	0.3	3.139	3.435	3.377	3.380	2.717	16.048	3.150	3.446	3.389	3.385	153.390	126.220
2	2	0.5	3.540	3.362	3.266	3.315	4.950	16.250	3.541	3.362	3.273	3.327	93.892	127.077
5	2	0.7	3.748	3.172	3.122	3.066	6.080	15.228	3.761	3.177	3.131	3.077	78.071	161.000
18	2	0.9	3.939	2.938	3.024	2.920	7.495	16.320	3.959	2.949	3.037	2.931	84.696	144.013

The table reveals the improvements of FC with SC from FC only in terms of the mean of bias.

based on the original definitions from Patel et al. (2006a) with estimation of our extended definitions of κ and τ under our FC with SC approach. We contrast how these two methods address functional coherence/association and ascendancy. The standard deviation of the bias also yields similar conclusions favoring our combined FC with SC approach over the FC only method (see Table 4 in Appendix 5). We examine the performance of our method using samples sizes of 15, 30, and 100 subjects. Although the difference in bias between the two approaches is relatively small, our FC with SC approach outperforms the FC only approach in every case that we consider.

Our second simulation study compares our method to a traditional correlation analysis. We use the same simulated θ from the previous simulation study to generate the neural activity profiles Y_{ant} and Y_{bnt} , for regions a and b , respectively, from a bivariate normal distribution with variances σ_a^2 and σ_b^2 and correlation ρ . Thus, the mean adjusted level of neural activity profiles R_{ant} and R_{bnt} also follow a bivariate normal distribution. We derive the expectation of Z_i 's as follows:

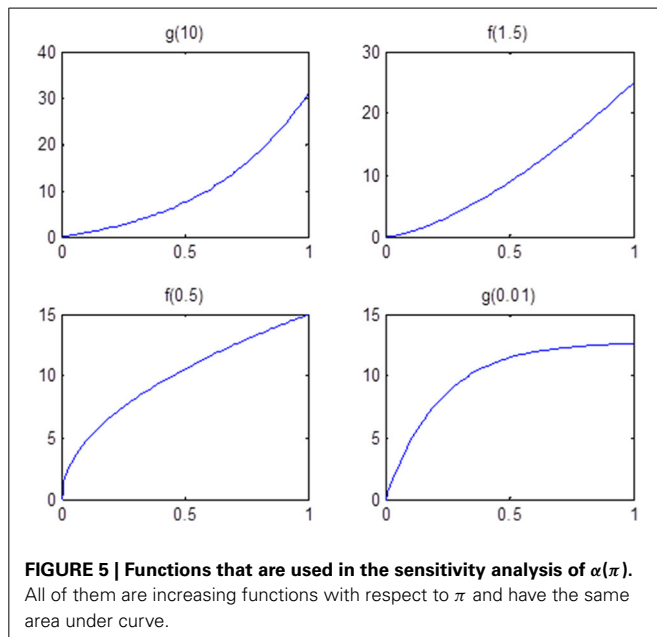
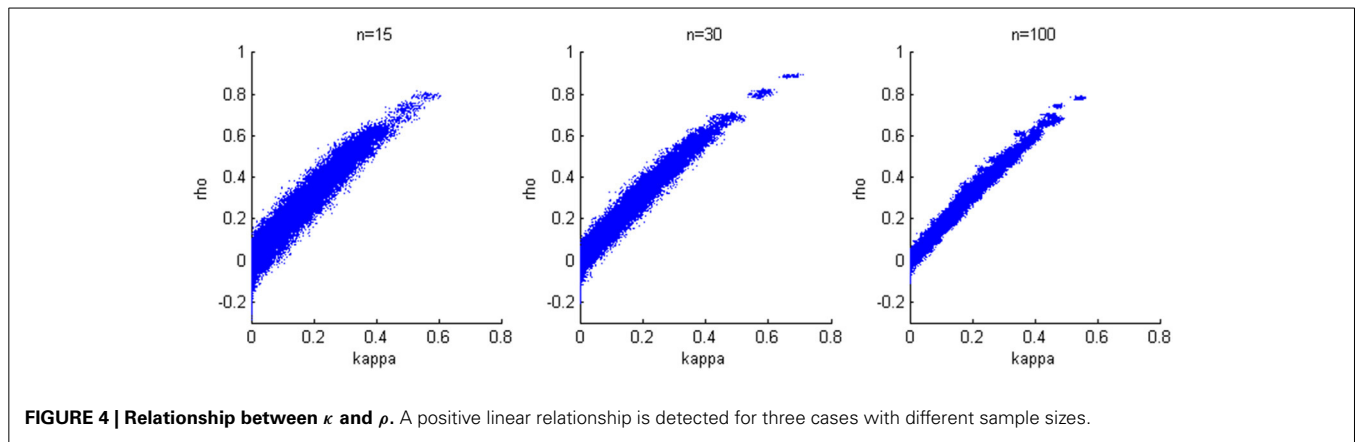
$$\begin{aligned}
 E(Z_1) &= P(Z_1 = 1) = P(R_{ant} > c_a, R_{bnt} > c_b) = \theta_1 \\
 E(Z_2) &= P(Z_2 = 1) = P(R_{ant} > c_a, R_{bnt} < c_b) = \theta_2 \\
 E(Z_3) &= P(Z_3 = 1) = P(R_{ant} < c_a, R_{bnt} > c_b) = \theta_3 \\
 E(Z_4) &= P(Z_4 = 1) = P(R_{ant} < c_a, R_{bnt} < c_b) = \theta_4.
 \end{aligned}
 \tag{12}$$

We solve for c_a and c_b using the marginal probabilities of R_{ant} and R_{bnt} , which are functions of θ , and we subsequently solve for ρ

using any of the above equations. We estimate the Pearson correlation coefficient ρ from our simulated data and compare it to the functional coherence κ . We expect to see substantial correspondence between these two measures since they both capture the functional associations between two regions.

We generate results for all combinations of hyperparameters specified in the first simulation study and again consider sample sizes of 15, 30, and 100. The accuracy of the estimation is not heavily influenced by the variance of the bivariate normal distribution; therefore, we present results when $\sigma_a^2 = \sigma_b^2 = 0.2$, which corresponds to estimates from our experimental data. We find a positive linear relationship between κ and ρ (Figure 4), while larger sample size yields smaller variability in the estimates.

Finally, we examine the impact of $\alpha(\pi)$ on the estimation of θ . As the data suggests, the functional coherence tends to increase as the SC increases. Therefore, we use four different increasing functions for the parameter $\alpha(\pi)$ as a parameter of the prior distribution of θ , based on power functions and exponential functions. Figure 5 shows the functions that we consider in the posterior simulations. As each function has the same integration over the interval $[0, 1]$, the expected values of all the θ_i 's simulated from each function are the same. We choose specific forms of functions f and g to satisfy the above criterion. We consider both convex functions, i.e., $g(10)$ and $f(1.5)$, and concave functions, i.e., $f(0.5)$ and $g(0.01)$. Table 3 summarizes the biases of θ estimated using different functions, where the true values are generated from the two most extreme cases $g(10)$ and $g(0.01)$. Here, the bias is calculated from the sum of the biases in all θ_i 's. We also vary the probability of SC, π , from weak to strong in the simulation study. The results indicate that the biases of θ across all



tested functions are comparable. Thus, we conclude that our method is not very sensitive to the choice of $\alpha(\pi)$, among those considered.

5. DISCUSSION

We build a unified Bayesian framework that provides a novel approach to combine functional and structural brain imaging for an integrated assessment of FC to evaluate the whole-brain networks. Joint analysis of both imaging modalities is an important tool to gain a better understanding of sensory and cognitive functions in the human brain as well as pathophysiology associated with psychiatric or neurologic disorders. Different from other methods that examine both FC and SC in a descriptive way, our method incorporates SC information into the model and allows for making statistical inferences, which acts as a basis of building a functional network. Our Bayesian model utilizes SC information to estimate the functional coherence between pairs of regions, yet our model does not allow the structural information to unduly drive the functional results. Our method is a purely data-driven, hypothesis-unconstrained approach, which can efficiently search across all pairs of defined brain regions of

interest. We develop two measures, κ and τ , to capture the functional coherence and degree of ascendancy, between the brain regions. κ is based on the probability of joint activation and deactivation, leading to an undirected network; while τ assesses the ascendancy between functionally connected regions, enabling us to construct a directed brain network. By incorporating both FC and SC, we more accurately identify the hubs and reduce the noise in the brain network.

Our method permits analyses examining all 20910 possible brain region pairs, between 205 subregions, to construct FC networks. We conduct a whole-brain analysis rather than requiring a pre-defined network or regions of interest. Our method is based on subregions centered on the most active voxel, yielding neural activity profiles that are representative of the brain activity within the small spherical subregions. Unlike other methods that average data across an entire region, we generate neural activity summaries for each subregion to allow heterogeneous patterns across the whole region. The mean of the time course puts equal weight on each voxel, which may not be able to explain the variation across the subregion. Instead, we use the first singular vector from SVD to represent the subregion, which is a weighted combination of neural activity profiles, and provides additional flexibility to summarize the information. We may allow multiple subregions from each AAL region, which provides more complete coverage in our whole-brain analyses.

We dichotomize the time series data to define functional joint activations, from which we evaluate the functional connectivity using our proposed κ metric that differs from the traditional correlation approach. There is no scale for the fMRI signal in the human brain that lends itself to natural interpretations of the level of neural activity. We define high and low neural activity from the fMRI signal based on a selected threshold. A possible extension to our current method is to use finer categories such as ordinal or even continuous to define joint functional activations.

In the context of dichotomizing the data, our modeling framework depends on c to declare elevated and inactive states of neural activity. Drastic changes to c will have a direct impact on the data that are input to our model, e.g., by setting c sufficiently high (or low), all the regions will become inactive (or active). We choose an arbitrary value of c reflecting the threshold for a standard score of fMRI BOLD activity. A reasonable default threshold is $c = 0$ to designate increased or decreased activity relative to the mean level. We use a slightly elevated value of $c = 0.01$ when determining the indicator of elevated brain activity for regional fMRI profiles from our data, which results in about 45% of the time points being regarded as active states. We

Table 3 | The bias of θ from estimations of different functions of $\alpha(\pi)$.

π	Generate Data from $g(10)$				π	Generate Data from $g(0.01)$			
	$g(10)$	$f(1.5)$	$f(0.5)$	$g(0.01)$		$g(10)$	$f(1.5)$	$f(0.5)$	$g(0.01)$
0.2045	0.0024	0.0026	0.0047	0.0051	0.0325	0.0018	0.0017	0.0023	0.0021
0.4626	0.0034	0.0035	0.004	0.0042	0.164	0.0035	0.0034	0.0026	0.0027
0.6894	0.0041	0.0038	0.0043	0.0043	0.4724	0.0038	0.0037	0.0035	0.0034
0.8987	0.0066	0.0073	0.0086	0.009	0.7031	0.0059	0.0058	0.0061	0.0059

Note that not much difference is detected across different functions with the same structural connectivity. The bias is calculated from the sum of the bias in all θ_i 's. Here, $f(a) = 10 \times (a + 1) \times \pi^a$, $g(a) = 10 / \left(\frac{a-1}{\ln(a)} - 1 \right) \times a^\pi - 10 / \left(\frac{a-1}{\ln(a)} - 1 \right)$

conduct sensitivity analysis of the choice of c , and find that when c fluctuates within a small range, i.e., for c up to 0.1, the major findings in our data application do not change.

We propose a functional coherence measurement that builds on Cohen's κ -statistic, which evaluates the levels of agreement adjusted for chance. Here, the chance agreement is defined as $(\theta_1 + \theta_2)(\theta_1 + \theta_3) + (\theta_3 + \theta_4)(\theta_2 + \theta_4)$. In addition to Cohen's κ , we may consider other agreement measures, e.g., Scott's π -statistic (Scott, 1955), in which the chance agreement is obtained by $[\frac{(\theta_1 + \theta_2) + (\theta_1 + \theta_3)}{2}]^2 + [\frac{(\theta_3 + \theta_4) + (\theta_2 + \theta_4)}{2}]^2$; Fleiss' κ -statistic (Fleiss, 1971), which is a generalization of Scott's π ; and other alternative chance-corrected statistics (Gwet, 2002). The major difference among these statistics is the way they calculate the chance agreement. Some researchers (Gwet, 2002) argue that the conditions that Cohen's κ requires, e.g., the chance-agreement probability is less than 0.5, are not always met in practice. In our case, however, the sum of the marginal probabilities $P(A_a = 1) + P(A_b = 1)$ is close to 1, which ensures that the chance-agreement probability does not exceed 0.5. In addition, other statisticians (Strijbos et al., 2006) believe that when fewer categories are included, Cohen's κ is a more conservative measurement of agreement. Therefore, we use this more strict measurement in our case.

The study of functional networks in the human brain is important to understand basic cognition, mental and neurological disorders, and response to treatments for these disorders. Moreover, the structural circuitry underlying functional connections may offer additional insights. We develop a Bayesian model that combines both functional and structural information to help characterize FC networks. Leveraging SC to quantify FC, our model yields more accurate and more informative results than considering solely functional data.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <http://www.frontiersin.org/journal/10.3389/fncom.2015.00022/abstract>

REFERENCES

- Behrens, T. E. J., Berg, H. J., Jbabdi, S., Rushworth, M. F. S., and Woolrich, M. W. (2007). Probabilistic diffusion tractography with multiple fibre orientations: what can we gain? *Neuroimage* 34, 144–155. doi: 10.1016/j.neuroimage.2006.09.018
- Bowman, F. D., Zhang, L., Derado, G., and Chen, S. (2012). Determining functional connectivity using fMRI data with diffusion-based anatomical weighting. *Neuroimage* 62, 1769–1779. doi: 10.1016/j.neuroimage.2012.05.032
- Cohen, J. A. (1960). A coefficient of agreement for nominal scales. *Educ. Psychol. Meas.* 41, 687–699.

- Cover, T. M., and Thomas, J. A. (1991). *Elements of Information Theory*. New York, NY: John Wiley & Sons. doi: 10.1002/0471200611
- Fiecas, M., Ombao, H., Linkletter, C., Thompson, W., and Sanes, J. (2010). Functional connectivity: shrinkage estimation and randomization test. *Neuroimage* 49, 3005–3014. doi: 10.1016/j.neuroimage.2009.12.022
- Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychol. Bull.* 76, 378–382. doi: 10.1037/h0031619
- Friston, K. J., Frith, C. D., Liddle, P. F., and Frackowiak, R. S. (1993). Functional connectivity: the principal component analysis of large (PET) data sets. *J. Cereb. Blood Flow Metab.* 13, 5–14. doi: 10.1038/jcbfm.1993.4
- Gee, M. S., Nimkin, K., Hsu, M., Israel, E. J., Biller, J. A., Katz, A. J., et al. (2011). Prospective evaluation of MR enterography as the primary imaging modality for pediatric Crohn disease assessment. *Am. J. Roentgenol.* 197, 224–231. doi: 10.2214/AJR.10.5970
- Gelman, A., Rubin, D. B., Carlin, J., and Stern, H. (1995). *Bayesian Data Analysis*. London: Chapman & Hall.
- Greicius, M. D., Supekar, K., Menon, V., and Dougherty, R. F. (2009). Resting-state functional connectivity reflects structural connectivity in the default mode network. *Cereb. Cortex* 19, 72–78. doi: 10.1093/cercor/bhn059
- Gwet, K. (2002). Kappa statistic is not satisfactory for assessing the extent of agreement between raters. *Stat. Methods Inter-Rater Reliabil. Assess.* 1, 1–6.
- Hampson, M., Peterson, B., Skudlarski, P., Gatenby, C., and Gore, J. (2002). Detection of functional connectivity using temporal correlations in MRI images. *Hum. Brain Mapp.* 15, 247–262. doi: 10.1002/hbm.10022
- Hendelman, W. (2000). *Atlas of Functional Neuroanatomy*. Boca Raton, FL: CRC Press. doi: 10.1201/9781420048209
- Honey, C. J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J. P., Meuli, R., et al. (2009). Predicting human resting-state functional connectivity from structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 106, 2035–2040. doi: 10.1073/pnas.0811168106
- Humphries, M. D., Gurney, K., and Prescott, T. J. (2006). The brainstem reticular formation is a small-world, not scale-free, network. *Proc. Biol. Sci.* 273, 503–511. doi: 10.1098/rspb.2005.3354
- Klatzky, R. L., Lippa, Y., Loomis, J. M., and Golledge, R. G. (2002). Learning directions of objects specified by vision, spatial audition, or auditory spatial language. *Learn. Mem.* 9, 364–367. doi: 10.1101/lm.51702
- Kollias, S. (2009). Parcelation of the white matter using DTI: insights into the functional connectivity of the brain. *Neuroradiol. J.* 22(Suppl. 1), 74–84.
- Landis, J. R., and Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics* 33, 159–174. doi: 10.2307/2529310
- Ling, J., Merideth, F., Caprihan, A., Pena, A., Teshiba, T., and Mayer, A. R. (2012). Head injury or head motion? Assessment and quantification of motion artifacts in diffusion tensor imaging studies. *Hum. Brain Mapp.* 33, 50–62. doi: 10.1002/hbm.21192
- Marrelec, G., Krainik, A., Duffau, H., Pelegrini-Issac, M., Lehericy, S., Doyon, J., et al. (2006). Partial correlation for functional brain interactivity investigation in functional MRI. *Neuroimage* 32, 228–237. doi: 10.1016/j.neuroimage.2005.12.057
- Mayer, A. R., Franco, A. R., and Harrington, D. L. (2009a). Neuronal modulation of auditory attention by informative and uninformative spatial cues. *Hum. Brain Mapp.* 30, 1652–1666. doi: 10.1002/hbm.20631

- Mayer, A. R., Mannell, M. V., Ling, J., Elgie, R., Gasparovic, C., Phillips, J. P., et al. (2009b). Auditory orienting and inhibition of return in mild traumatic brain injury: a fMRI study. *Hum. Brain Mapp.* 30, 4152–4166. doi: 10.1002/hbm.20836
- Morgan, V. L., Mishra, A., Newton, A. T., Gore, J. C., and Ding, Z. (2009). Integrating functional and diffusion magnetic resonance imaging for analysis of structure-function relationship in the human language network. *PLoS ONE* 4:e6660. doi: 10.1371/journal.pone.0006660
- Patel, R., Bowman, F. D., and Rilling, J. K. (2006a). A Bayesian approach to determining connectivity of the human. *Hum. Brain Mapp.* 27, 267–276. doi: 10.1002/hbm.20182
- Patel, R., Van De Ville, D., and Bowman, F. D. (2006b). Determining Significant Connectivity by 4D spatiotemporal wavelet packet resampling of functional neuroimaging data. *Neuroimage* 31, 1142–1155. doi: 10.1016/j.neuroimage.2006.01.012
- Ramsey, J. D., Sanchez-Romero, R., and Glymour, C. (2014). Non-Gaussian methods and high-pass filters in the estimation of effective connections. *Neuroimage* 84, 986–1006. doi: 10.1016/j.neuroimage.2013.09.062
- Rykhlevskaia, E., Gratton, G., and Fabiani, M. (2008). Combining structural and functional neuroimaging data for studying brain connectivity: a review. *Psychophysiology* 45, 173–187. doi: 10.1111/j.1469-8986.2007.00621.x
- Scott, W. (1955). Reliability of content analysis: the case of nominal scale coding. *Public Opin. Q.* 19, 321–325. doi: 10.1086/266577
- Seghier, M. L., and Friston, K. J. (2013). Network discovery with large DCMs. *Neuroimage* 68, 181–191. doi: 10.1016/j.neuroimage.2012.12.005
- Simpson, S., Laurienti, P., and Bowman, F. D. (2013). Analyzing complex functional brain networks: fusing statistics and network science to understand the brain. *Stat. Surv.* 7, 1–36. doi: 10.1214/13-SS103
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, 208–219. doi: 10.1016/j.neuroimage.2004.07.051
- Smith, S. M., Miller, K. L., Salimi-Khorshidi, G., Webster, M., Beckmann, C. F., Nichols, T. E., et al. (2010). Network modelling methods for FMRI. *Neuroimage* 54, 875–891. doi: 10.1016/j.neuroimage.2010.08.063
- Sporns, O., Honey, C. J., and Kötter, R. (2007). Identification and classification of hubs in brain networks. *PLoS ONE* 10:e1049. doi: 10.1371/journal.pone.0001049
- Stark, D. E., Margulies, D. S., Shehzad, Z. E., Reiss, P., Kelly, A. M. C., Uddin, L. Q., et al. (2008). Regional variation in interhemispheric coordination of intrinsic hemodynamic fluctuations. *J. Neurosci.* 28, 13754–13764. doi: 10.1523/JNEUROSCI.4544-08.2008
- Strijbos, J. W., Martens, R. L., Prins, F. J., and Jochems, W. M. G. (2006). Content analysis: what are they talking about? *Comput. Educ.* 46, 29–48. doi: 10.1016/j.compedu.2005.04.002
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi: 10.1006/nimg.2001.0978
- Ungerleider, L. G. (1995). Functional brain imaging studies of cortical mechanisms for memory. *Science* 270, 769–774. doi: 10.1126/science.270.5237.769
- van den Heuvel, M. P., Mandl, R. C. W., Kahn, R. S., and Hulshoff Pol, H. E. (2009). Functionally linked resting-state networks reflect the underlying structural connectivity architecture of the human brain. *Hum. Brain Mapp.* 30, 3127–3141. doi: 10.1002/hbm.20737
- Watts, D. J., and Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *Nature* 393, 440–442. doi: 10.1038/30918
- Woolrich, M. W., Ripley, B. D., Brady, M., and Smith, S. M. (2001). Temporal autocorrelation in univariate linear modeling of FMRI data. *Neuroimage* 14, 1370–1386. doi: 10.1006/nimg.2001.0931
- Zuo, X. N., Di Martino, A., Kelly, C., Shehzad, Z. E., Gee, D. G., Klein, D. F., et al. (2010). The oscillating brain: complex and reliable. *Neuroimage* 49, 1432–1445. doi: 10.1016/j.neuroimage.2009.09.037

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 01 October 2013; accepted: 04 February 2015; published online: 20 February 2015.

Citation: Xue W, Bowman FD, Pileggi AV and Mayer AR (2015) A multimodal approach for determining brain networks by jointly modeling functional and structural connectivity. *Front. Comput. Neurosci.* 9:22. doi: 10.3389/fncom.2015.00022

This article was submitted to the journal *Frontiers in Computational Neuroscience*. Copyright © 2015 Xue, Bowman, Pileggi and Mayer. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.