Multi-object Data Integration in the Study of Primary Progressive Aphasia

Rene Gutierrez*
Department of Statistics, Texas A & M University,
College Station, TX
Aaron Scheffler
Department of Epidemiology and Biostatistics,
UC San Francisco, San Francisco, CA
Rajarshi Guhaniyogi
Department of Statistics, Texas A & M University,
College Station, TX
Maria Luisa Gorno-Tempini
Memory and Aging Center, Department of Neurology,
University of California, San Francisco, CA
Maria Luisa Mandelli
Memory and Aging Center, Department of Neurology,
University of California, San Francisco, CA
Giovanni Battistella
Department of Otolaryngology–Head and Neck Surgery,
Harvard Medical School, Boston, MA

February 10, 2023

Abstract

This article focuses on a multi-modal imaging data application where structural/anatomical information from grey matter (GM) and brain connectivity information in the form of a brain connectome network from functional magnetic resonance imaging (fMRI) are available for a number of subjects with different degrees of primary progressive aphasia (PPA), a neurodegenerative disorder (ND) measured

*The authors gratefully acknowledges *acknowledges funding from National Science Foundation Grant DMS-2220840 and DMS-2210672. We thank the Memory and Aging Center at UC San Francisco for providing the MR images and the patients and their families for the time and effort they dedicated to the research. The study was supported by Grants from the National Institutes of Health (NINDS R01NS050915, NIDCD K24DC015544, NIA P50AG023501).*

through a speech rate measure on motor speech loss. The clinical/scientific goal in this study becomes the identification of brain regions of interest significantly related to the speech rate measure to gain insight into ND pathways. Viewing the brain connectome network and GM images as objects, we develop a flexible joint object response regression framework of network and GM images on the speech rate measure. A novel joint prior formulation is proposed on network and structural image coefficients in order to exploit network information of the brain connectome, while leveraging the topological linkages among connectome network and anatomical information from GM to draw inference on brain regions significantly related to the speech rate measure. The principled Bayesian framework allows precise characterization of the uncertainty in ascertaining a region being actively related to the speech rate measure. Our framework yields new insights into the relationship of brain regions with PPA, offering deeper understanding of neuro-degeneration pathways for PPA. Supplementary file adds details about posterior computation and additional empirical results.

Keywords: Bayesian inference; brain connectome; functional magnetic resonance imaging; grey matter; multi-modal imaging; primary progressive aphasia.

1 Introduction

This article is motivated by a clinical application on patients suffering from brain loss due to Primary Progressive Aphasia (PPA), an ND characterized by loss of language ability which shares pathological signatures with Alzheimer's disease and frontotemporal dementia. Specifically, we consider the nonfluent/aggramatic variant of PPA (nfvPPA) characterized by motor speech and grammar loss and left inferior frontal atrophy (Gorno-Tempini *et al.*, 2008). Multi-modal imaging data are available for each of these PPA patients which include: (a) *brain network information* and (b) *brain structural information*. Brain network envisions each region of interest (ROI) as a network node and quantifies connectivity between the pairs of network nodes using functional magnetic resonance imaging (fMRI). Brain structural information is obtained using structural MRI (sMRI), e.g., grey matter (GM) images over brain volumetric pixels (voxels). Both images are collected on a common brain atlas which segments a human brain into different regions of interest (ROI), with each ROI containing a number of voxels. A scalar-valued speech rate score is available for these patients to measure the degree of their motor speech loss due to PPA.

The major scientific goals of the study is two fold, (a) to estimate regression relationships between the speech rate and the two imaging modalities; and (b) to identify ROIs and voxels significantly related to the speech rate, and thus infer on the ND pathway for PPA. To this end, the identification of influential ROIs via principled modeling framework is necessary to yield results that are scientifically interpretable in the lingua franca of well characterized brain anatomy and function. While many ROI based inferential techniques are constrained by considering the ROI as the most basic unit of analysis, the PPA data application allows for additional granular inference at the level of the voxel which will provide neuroscientists with the dual ability to frame results at the macro level of the ROI and probe the micro

voxel-level signals that drive these results. To our knowledge, this article proposes the first statistical framework that allows for flexibility of macro and micro inference on multi-object imaging data which is urgently needed in the neuroscience community (Calhoun and Sui, 2016). Uncertainty quantification regarding the inference on influential regions becomes crucial since the analysis only involves a limited number of PPA patients.

When inference on influential brain regions is of interest, one may fit a regression model with the speech rate as the response and brain network and structural images as predictors, extending the popular literature of scalar-on-object regressions (Guhaniyogi *et al.*, 2017; Fan *et al.*, 2019; Guha and Rodriguez, 2021; Feng *et al.*, 2019) to scalar-on-multi-object regressions. However, limited sample size in our application precludes drawing satisfactory inference from these models. As an alternative, we address the inferential objective by formulating a multi-modal regression framework with the brain network (defined over brain ROIs) and structural images (defined over brain voxels) as two sets of responses and the speech rate as the predictor. A hierarchical Bayesian approach is adopted wherein a joint prior structure on coefficients of the speech rate corresponding to the structural and network objects (hereon referred to as the structural and network coefficients, respectively) is developed to account for topology in multi-object data and to allow the information in separate image objects to complement and reinforce each other in their relation to the scalar predictor. To elaborate on it, we begin with a common brain atlas for both image objects to ensure that it provides an organizing principle that links together structural and network information via a shared set of ROIs and a group of voxels within each ROI. The joint prior on network and structural coefficients are constructed respecting the *hierarchical* constraint in multi-object topology which ensures all voxels within an ROI is uninfluential if the ROI is un-influential (see Section 3.2). While this article does not explicitly make

use of the structural information in the GM images by careful spatial modeling of GM image coefficient, it partially exploits the structural information by imposing the *hierarchical* constraint in the prior construction step. The problem of identifying influential ROIs is cast under a nonlinear variable selection framework, wherein latent activation indicator corresponding to all ROIs are shared among both sets of coefficients to enforce that all voxels from a particular ROI and all network edges connected to that particular ROI have no relationship with the predictor when the activation indicator corresponding to the ROI is zero. As a byproduct of our construction, *symmetry* and *transitivity* property (Hoff, 2005) of the un-directed network object is preserved, as discussed in Section 3.2. The prior construction achieves efficient computation, identifies influential ROIs and voxels within an influential ROI which are key to study neuronal atrophy, and produces well-calibrated interval estimates for the multi-modal regression coefficients. Moreover, our framework attaches uncertainty in identifying these ROIs and offers improved inference over regression methods with a single imaging modality.

There is a dearth of principled Bayesian literature addressing the inferential objectives of the motivating application, and our proposal is arguably the first Bayesian multi-object regression approach to answer the inferential questions stated before. We now provide a brief overview of the available literature to contrast them with our proposal. In the course to determine the association between network or structural objects and the speech rate score, the most popular approach estimates the association between each network edge or GM voxel and the speech rate independently, providing a p-value "map" (Friston, 2003; Lazar, 2016), after adjusting for multiple comparisons to identify "significant" ROIs and voxels. However, these methods do not take into account the joint impact of all network edges and voxels of the GM images simultaneously. A more principled approach

is to vectorize both objects and regress them jointly on the speech rate, leading to a high dimensional vector response regression problem. This approach can take advantage of the recent developments in high dimensional multivariate reduced rank sparse regression literature, consisting of both penalized optimization (Yuan *et al.*, 2007; Chen and Huang, 2012) and Bayesian shrinkage (Goh *et al.*, 2017). However, vectorization of both sets of object responses during analysis ignores their individual topology (e.g., the symmetry and transitivity in the network) and linkage between the topology through the hierarchical constraint, and apparently does not allow identification of influential ROIs with uncertainty.

There are some recent efforts to build regression models with a single-object response, mainly for the functional and structural neuro-imaging data (Guhaniyogi and Spencer, 2021; Guha and Guhaniyogi, 2021), though there is a dearth of literature on exploiting principled linkages among image objects (e.g., through the hierarchical constraint). Failure to consider the structure and cross information from multiple images (as in a uni-modal analysis) have generally a negative impact on ND research in terms of lower detection power (Li *et al.*, 2018), bias in estimated effects (Dai and Li, 2021), and sensitivity of results to noise (Calhoun and Sui, 2016). Additionally, the existing approaches involving network objects consider low-dimensional structure for the network coefficient, while our approach does not rely on such restrictive assumptions.

Our proposed approach is considerably different from the existing statistical literature on multi-modal data integration. In particular, there have been a class of unsupervised multi-modal analysis built on matrix or tensor factorization (Lock *et al.*, 2013), or methods exploiting structural connectivity information in the prior construction for the functional connectivity analysis from fMRI data (Xue *et al.*, 2015). Xue *et al.* (2018) proposes regression on disease status on low-frequency fluctuation (fALFF) from resting-state fMRI scans,

voxel based morphometry (VBM) from T1-weighted MRI scans, and fractional anisotropy (FA) from DTI scans. In the same vein Li and Li (2021); Dai and Li (2021) develop frameworks to account for non-linear association between a scalar response and multi-modal predictors. While these approaches do form linkages among image modalities, they do not properly model within image correlations and thus are not able to address our inferential goals of jointly modeling information across images. Moreover, all these approaches do not naturally offer identification of influential ROIs with uncertainty.

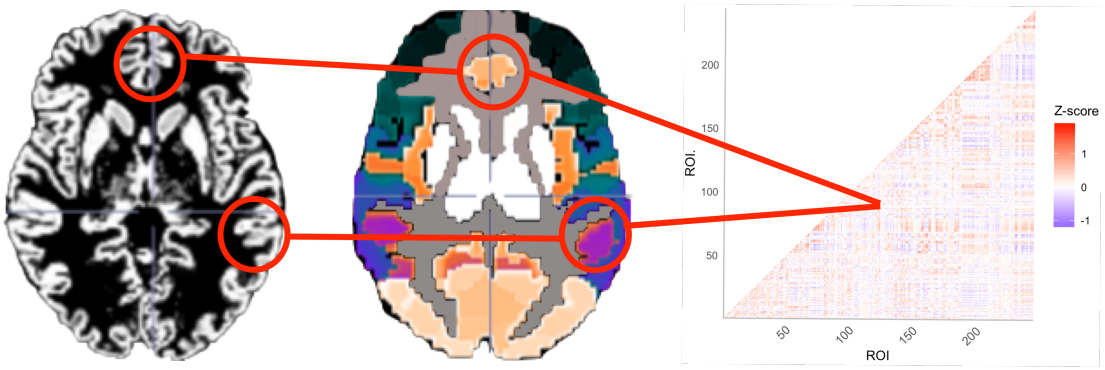
The specific application of the proposed framework to our motivating PPA imaging data will substantially improve characterization of dementia related language loss in contrast to previous analyses. Up to this point, the foundational work in the cognitive neuroscience of PPA (Mandelli *et al.*, 2016; Gorno-Tempini *et al.*, 2008) has not directly combined structural information relating to neuronal atrophy with network information on brain connectivity. At best, brain networks that underlie language have been identified via an informal fusion of information from multi-object image data via seed based methods (Battistella *et al.*, 2020). Specifically, several ROIs are identified which display the highest degree of atrophy in structural images (e.g. GM degradation) of PPA patients. These ROIs are then used to define "seed regions" which are used to construct speech networks that quantify the nodes highly connected to the seed regions. While the speech networks identified via this approach are likely important to study PPA, the ad-hoc way of selecting seed regions and the lack of a joint model that simultaneously consider the influence of structural and network images on language dysfunction means that the influence of atrophy in other ROIs in conjunction with speech networks have been overlooked. Thus, the application of our proposed approach to the PPA data will provide a new range of opportunities anchored in multi-modal image integration to study cognitive language models.

The rest of the manuscript proceeds as follows. Section 2 provides a description of the structure of multi-modal data we analyze in this article. Section 3 describes the model development and prior framework to draw inference from multi-modal images and Section 4 discusses posterior computation of the proposed model. Empirical investigations with data generated under various simulation settings are reported in Section 5. Section 6 analyzes the multi-modal dataset, offering scientific findings on influential ROIs. Finally, Section 7 summarizes the idea and highlights some of the extensions to be explored in the near future.

2 Clinical Case Study on Nonfluent Primary Progressive Aphasia (nfvPPA) Neurodegenerative Disorder

We focus on a clinical application derived from multi-modal image studies conducted on patients with nonfluent/aggramatic variant of PPA (nfvPPA), a form of ND characterized by motor speech and grammar loss and left inferior frontal atrophy (Gorno-Tempini *et al.*, 2008). To investigate the neural underpinnings of disruption to motor speech/fluency in nfvPPA patients, clinical images from multiple modalities was collected as detailed below.

Clinical images and language evaluation: Imaging data is acquired on 24 nfvPPA patients during the course of clinical research activity. Data is collected from the following imaging modalities: sMRI derived gray matter (GM) **(Figure 1a)** which measures the likelihood a voxel containing neuronal cell bodies; and task-free resting state functional magnetic resonance imaging (fMRI) to measure brain activation via neuronal oxygen consumption in subjects at rest. All images are registered to the Montreal Neurological Institute (MNI) template space with voxels parcellated into 245 ROIs using the Brainnetome atlas such that images across modalities and subjects can be directly compared and each



(a) Structural image (b) Atlas (c) Network image

Figure 1: Schematic of the multi-object brain imaging data structure for a PPA patient. (a) Structural image encoding voxel-level gray matter (GM) probability, (b) Brainnetome atlas parcellation of the brain into anatomical ROIs, (c) Network image obtained by calculating the pairwise Pearson correlation Z-score for the average fMRI signal in each ROI. Red circles and lines connect (a) structural and (c) network information from images via the (b) parcellated atlas.

voxel is nested in an anatomically defined ROI **(Figure 1b)** (Fan *et al.*, 2016). For each subject, a 'brain network' represented by a symmetric adjacency matrix is obtained from the fMRI image by considering rows and columns of this matrix corresponding to different ROIs and entries corresponding to the Z-scores obtained by transforming the Pearson correlation between average fMRI data of two ROIs **(Figure 1c)**. We focus our analysis of language loss on speech rate, the number of words spoken per minute, a measure of motor speech via a subject's articulation rate. This speech rate measure is automatically extracted from recorded speech from the Grandfather passage, a 129-word block of text meant to elicit a comprehensive set of phonemes in English (Ogar *et al.*, 2007), via SALT software (https://www.saltsoftware.com/), a software platform used to automatically extract language features from recorded speech.

Scientific question of interest: Language loss in nfvPPA patients is driven by neurodegeneration in the *left inferior frontal* region but the dual role of structural damage

and brain connectivity in language loss is not well characterized (Mandelli *et al.*, 2016). Prior clinical studies have identified 38 regions of interest (ROIs) which are related to motor speech/fluency, known as the motor speech/fluency speech production network (SPN) (Mandelli *et al.*, 2016). One of our primary goals is to extend the multi-modal study of motor speech/fluency beyond this established SPN to the whole brain. To do so, we apply our proposed multi-modal framework to regress GM maps (which capture focal neurodegeneration), and, fMRI brain connectivity networks (which capture disruptions of brain connectivity), on the speech rate score. The next section describes the novel regression framework to achieve these scientific goals.

3 Bayesian Regression with Multiple Imaging Responses and Speech Rate as a Predictor

3.1 Model Framework

For the ith subject, let $y_i \in \mathbb{R}$ denote the speech rate measure, $\boldsymbol{x}_i = (x_{i,1}, ..., x_{i,H})'$ represent the biological and demographic variables (e.g., age, gender) and $\boldsymbol{A}_i$ represent the weighted brain network object. We assume that the network of all subjects are defined on a common set of nodes, with elements of $\boldsymbol{A}_i$ encoding the strength of network connections between different nodes for the i-th subject. In particular, the network object $\boldsymbol{A}_i$ is expressed in the form of a $P \times P$ matrix with the (p, p')-th entry of the matrix $a_{i,(p,p')}$ signifying the strength of association between the pth and p'th node, where $p, p' = 1, ..., P$ and P is the number of network nodes. This paper specifically focuses on networks that contain no self relationship, i.e., $a_{i,(p,p)} \equiv 0$, and are un-directed ($a_{i,(p,p')} = a_{i,(p',p)}$). Such assumptions hold for the data application pertaining to Section 2, where $\boldsymbol{A}_i$ represents the brain connectome

network matrix obtained from the fMRI scan, with each node representing a specific brain region of interest (ROI) and edges signify correlations between fMRI signals in two regions. Let $\boldsymbol{g}_{i,1}$,...,$\boldsymbol{g}_{i,P}$ denote the $V_1, ..., V_P$ dimensional structural objects in regions $\mathcal{R}_1$,...,$\mathcal{R}_P$, respectively. In the context of our data application, they represent volumetric elements (voxels) of the GM image from the P ROIs. The nested structure of voxels within ROIs provides biologically plausible organization and is instrumental for variable selection and computation as detailed in the upcoming methodological development.

For $i = 1, \ldots, n$, we assume that the relationship between the speech rate measure y_i varies with every network edge and every GM voxel after accounting for the edge and voxel specific effect of the biological covariates, and propose a set of conditionally independent generalized linear models for every network edge and GM voxel, given by

$$E[a_{i,(p,p')}] = H_1^{-1}\Big(\sum_{h=1}^{H} \psi^a_{p,p',h} x_{i,h} + \theta_{p,p'}\, y_i\Big), \;\; E[g_{i,v,p}] = H_2^{-1}\Big(\sum_{h=1}^{H} \psi^g_{p,h} x_{i,h} + \beta_{v,p}\, y_i\Big), \quad (1)$$

for $v = 1, \ldots, V_p$; $p = 1, \ldots, P$, where $H_1(\cdot)$ and $H_2(\cdot)$ are the link functions, $\theta_{p,p'}$ is the (p, p')th element of the $P \times P$ matrix $\boldsymbol{\Theta}$, $\beta_{v,p}$ is the vth element of the V_p dimensional vector of coefficients $\boldsymbol{\beta}_p$, $\psi^a_{p,p',h}$ denotes the effect of $x_{i,h}$ on the edge connecting pth and p'th nodes of the network matrix and $\psi^g_{p,h}$ determines the effect of $x_{i,h}$ on the GM image at a voxel within $\mathcal{R}_p$. Considering the symmetry and zero diagonal constraint in the network object $\boldsymbol{A}_i$, we set $\theta_{p,p'} = \theta_{p',p}$ and $\theta_{p,p} = 0$, for all $1 \leq p < p' \leq P$. Additionally, we assume a bilinear effect of $\psi^a_{p,p',h} = \psi^a_{p,h}\psi^a_{p',h}$ to achieve parsimony following the popular literature on bilinear regression analysis (Gabriel, 1998; Von Rosen, 2018). When both sets of responses follow a normal linear model with an identity link function, (1) becomes

$$a_{i,(p,p')} = \sum_{h=1}^{H} \psi^a_{p,h}\psi^a_{p',h} x_{i,h} + \theta_{p,p'}\, y_i + e_i^{(p,p')}, \;\; g_{i,v,p} = \sum_{h=1}^{H} \psi^g_{p,h} x_{i,h} + \beta_{v,p}\, y_i + w_i^{(v,p)}, \quad (2)$$

for $v = 1, \ldots, V_p$; $p = 1, \ldots, P$, where $e_i^{(p,p')}$ and $w_i^{(v,p)}$ represent the errors in the two regression models. While carefully constructed spatial covariance structure can be imposed

on $e_i^{(p,p')}$ and $w_i^{(j,p)}$, limited sample size in our application is not conducive to adding sufficiently expressive structures for covariance matrices of the errors. We relegate it as a future work and offer discussion in Section 7. Instead, this article focuses on joint learning of the mean structure for two sets of models and assumes $e_i^{(p,p')}, w_i^{(v,p)} \overset{ind.}{\sim} N(0, \tau^2)$ (for $1 \leq p < p' \leq P$ and $v = 1, \ldots, V_p$) for simplicity, following multivariate linear response regression literature (Goh *et al.*, 2017). Consistent with the symmetry and zero diagonal entry in $\boldsymbol{A}_i$, we assume $e_i^{(p,p')} = e_i^{(p',p)}$ and $e_i^{(p,p)} = 0$. Therefore, stacking over elements of the network matrix and elements of the GM voxels over each region, (2) can be written as

$$\boldsymbol{A}_i = \sum_{h=1}^{H} \boldsymbol{\psi}_h^a (\boldsymbol{\psi}_h^a)' x_{i,h} + \boldsymbol{\Theta}\, y_i + \boldsymbol{E}_i, \;\; \boldsymbol{g}_{i,p} = \sum_{h=1}^{H} \mathbf{1}_{V_p} \psi_{p,h}^g x_{i,h} + \boldsymbol{\beta}_p\, y_i + \boldsymbol{w}_i^{(p)}, \tag{3}$$

$p = 1, \ldots, P$, where $\boldsymbol{\psi}_h^a = (\psi_{1,h}^a, ..., \psi_{P,h}^a)'$ and $\mathbf{1}_{V_p}$ represents a V_p-dimensional vector of ones. The error $\boldsymbol{E}_i \in \mathbb{R}^{P \times P}$ is the symmetric matrix with zero diagonal entries corresponding to the network object and $\boldsymbol{w}_i^{(p)}$ is the V_p-dimensional error vector corresponding to the GM image at the pth ROI. The key to joint learning of the multi-modal data lies in the development of a joint prior structure on $\boldsymbol{\Theta}$ and $\boldsymbol{\beta}_p$'s, as described in the next section.

3.2 Prior Distribution on Multi-Modal Coefficients

Our joint prior construction on coefficients $\boldsymbol{\beta}_p$'s and $\{\theta_{p,p'} : p < p'\}$ for multi-modal predictors is fundamental to exploiting topology of the image objects and cross-information among them by forming principled linkages among images. The prior construction is aimed at: (a) identification of influential ROIs with uncertainty; (b) shrinkage of unimportant voxel coefficients to zero within an influential ROI; and (c) guaranteeing efficient computation of the posterior for the proposed prior. We cast the problem of identifying influential ROIs from the multi-modal images as a high-dimensional variable selection problem and formulate prior distributions on multi-modal object coefficients building upon the existing

literature on prior constructions for high-dimensional regression coefficients (George and McCulloch, 1993; Scott and Berger, 2010; Carvalho *et al.*, 2010).

The direct application of existing variable selection prior on multi-modal coefficients is unappealing for multiple reasons. First, an ordinary variable selection prior on coefficients $\boldsymbol{\Theta}$ and $\boldsymbol{\beta}_p$'s identifies cells in $\boldsymbol{A}_i$ and $\boldsymbol{g}_{i,p}$ (which in our application refer to the network edges and GM voxels) significantly related to the predictor, rather than identifying influential ROIs. Second, we seek to impose two additional restrictions on the prior construction of $\boldsymbol{\Theta}$ and $\boldsymbol{\beta}_p$ motivated by the neuro-scientific application. First, if at least one of the pth and p'th ROIs is not related to the speech rate, the edge coefficient $\theta_{p,p'}$ corresponding to the edge between pth and p'th nodes is unimportant. Second, if the pth ROI is deemed uninfluential then all voxels within the pth ROI are unrelated to the speech rate, i.e., $\boldsymbol{\beta}_p = \mathbf{0}$. These restrictions are relevant due to the hierarchical arrangement of voxels and ROIs and are jointly referred to as the *hierarchical constraint*. Finally, we expect the matrix of coefficients $\boldsymbol{\Theta}$ (which itself can be regarded as a weighted network) to exhibit *transitivity effects*, i.e., we expect that if the interactions between regions p and p' and between regions p' and p'' are both related to the speech rate, the interaction between regions p and p'' will likely associated with the speech rate (see, e.g., Li *et al.* (2013)). An ordinary variable selection prior on multi-modal coefficients does not necessarily conform to all these requirements.

Let $\xi_1, ..., \xi_P$ denote the binary inclusion indicators corresponding to the P ROIs taking values in $\{0, 1\}$, with $\xi_p = 0$ determining no effect of the pth ROI on the covariate of interest. In our construction, the network edge coefficient $\theta_{p,p'}$ is endowed with a variable selection prior given by

$$\theta_{p,p'}|\lambda_{p,p'}, \tau^2, \sigma_\theta, \xi_p, \xi_{p'} \overset{ind.}{\sim} \xi_p\xi_{p'}N(0, \tau^2\sigma_\theta^2\lambda_{p,p'}^2) + (1 - \xi_p\xi_{p'})\delta_0, \;\; p < p', \tag{4}$$

where δ_0 corresponds to the Dirac-delta function, $\lambda_{p,p'}$ is a local shrinking parameter cor-

responding to the (p, p')th edge and σ_θ is the global shrinking parameter for the network coefficient. The prior closely mimics the spike-and-slab variable selection structure (George and McCulloch, 1993) with an important difference. While an ordinary spike-and-slab prior introduces a binary inclusion indicator corresponding to each variable, (4) enforces $\theta_{p,p'} = 0$ when either $\xi_p = 0$ or $\xi_{p'} = 0$. Such a formulation is sensible from a network perspective as it implies that the edge connecting two network nodes is unrelated to the predictor when at least one of the network nodes is not influential, i.e., it satisfies *hierarchical constraint*. Additionally, the formulation naturally incorporates transitivity effects in the network coefficient $\mathbf{\Theta}$. We further assign $\sigma_\theta \sim C^+(0,1)$ and $\lambda_{p,p'} \overset{ind.}{\sim} C^+(0,1)$ to complete prior specification on network coefficient. Integrating out σ_θ and $\lambda_{p,p'}$ in (4), $\theta_{p,p'}|\tau^2, \xi_p = 1, \xi_{p'} = 1$ follows the popular horseshoe prior (Carvalho *et al.*, 2010) which offers a flexible prior structure for precise estimation of nonzero network edge coefficients.

The GM coefficient $\beta_{v,p}$ in the pth ROI is modeled using $\beta_{v,p} = \xi_p \gamma_{v,p}$, to ensure all voxel coefficients in the pth ROI become unrelated to the predictor if the pth ROI is uninfluential (i.e., $\xi_p = 0$), thus satisfying the *hierarchical constraint*. To estimate voxel level effects in the pth ROI on the predictor, each $\gamma_{v,p}$ is assigned a horseshoe shrinkage prior which takes the following scale-mixture representation,

$$\gamma_{v,p}|\phi_{v,p}, \eta_p^2, \tau^2 \sim N(0, \tau^2\eta_p^2\phi_{v,p}^2), \;\; \phi_{v,p} \overset{i.i.d.}{\sim} C^+(0,1), \;\; \eta_p \overset{i.i.d.}{\sim} C^+(0,1), \tag{5}$$

for $v = 1, ..., V_p$; $p = 1, ..., P$. The prior structure (5) induces approximate sparsity in voxel-level GM coefficients $\gamma_{v,p}$ by shrinking the components which are less influential toward zero while retaining the true signals. Finally, the binary inclusion indicators are assigned Bernoulli prior distribution $\xi_p \overset{i.i.d.}{\sim} Ber(\nu)$ with $\nu \sim Beta(a_\nu, b_\nu)$ to account for multiplicity correction (Scott and Berger, 2010). Notably, an estimate of the posterior probability of the event $\{\xi_p = 1\}$ shows the uncertainty in identifying the pth ROI to be influential. Thus,

$P(\xi_p = 1|\text{Data})$ close to 1 or 0 signifies strong evidence in favor of identifying the pth ROI to be active or inactive, repectively. The prior specification on multi-modal coefficients is completed by assigning an inverse Gamma $IG(a_\tau, b_\tau)$ prior on the error variance τ^2 and an $IG(a_\theta, b_\theta)$ prior on the error variance σ_θ. We assign $N(0,1)$ prior distribution on $\psi^a_{p,h}, \psi^g_{p,h}$ for all $p = 1, ..., P$ and $h = 1, ..., H$.

4 Posterior Computation

Full conditional distributions for all the parameters are available and mostly correspond to standard families (available in Section 1 of the supplementary file). Thus, posterior computation can proceed through a Markov chain Monte Carlo algorithm. The MCMC sampler is run for 5000 iterations, with the first 1000 discarded as burn-in. All posterior inference is based on post-convergence samples suitably thinned. The effective sample size averaged over all $\boldsymbol{\Theta}$ and $\boldsymbol{\beta}_p$'s for the post-convergence iterations is 3143.

We have implemented our code in R (without using any C++, Fortran, or Python interface) on a cluster computing environment with three interactive analysis servers, 56 cores each with the Dell PE R820: 4x Intel Xeon Sandy Bridge E5-4640 processor, 16GB RAM and 1TB SATA hard drive. Different replications of the model are implemented under a parallel architecture by making use of the packages `doparallel` and `foreach` within R. The computation times of running 5000 MCMC iterations with $P = 100$ and $V_1 = ... = V_P = 50$ is given by 142 min on average across all simulations.

L (suitably thinned) post-convergence MCMC samples $\xi_p^{(1)}, ..., \xi_P^{(L)}$ of the indicator ξ_p are used to empirically assess if the pth ROI is significantly associated with the predictor. In particular, the pth ROI $\mathcal{R}_p$ is related to y_i if $\sum_{l=1}^{L} \xi_p^{(l)}/L > t$, for $0 < t < 1$. The ensuing sections set $t = 0.5$ to decide which ROIs influentially related to y_i.

5 Simulation Studies

5.1 Data Generation

In all our simulations, we assume no biological and demographic covariates, i.e. $\boldsymbol{\psi}^a_{p,h} = \boldsymbol{\psi}^g_{p,h} = \mathbf{0}$ and simulate $y_1, \ldots, y_n$ from N(0,1). We then generate responses $\boldsymbol{A}_i$ and $\boldsymbol{g}_{i,p}$ from model (3) with the true network coefficient $\boldsymbol{\Theta}_t$ and true structural coefficient $\boldsymbol{\beta}_{p,t}$ for the pth ROI. The subscript t indicates the true data generating parameters. In all simulations we set a limited sample size of $n = 16$ to resemble our simulation settings with the PPA data application and assume equal number of voxels per ROI, i.e., $V_1 = \cdots = V_P = V$.

Simulating true coefficients $\boldsymbol{\Theta}_t$ and $\boldsymbol{\beta}_{p,t}$. To simulate the true coefficients $\boldsymbol{\Theta}_t$ and $\boldsymbol{\beta}_{p,t}$, we first generate binary variables $\xi_{1,t}, ..., \xi_{P,t} \overset{i.i.d.}{\sim} Ber(\nu_t)$ with $\xi_{p,t} = 1$ sets the p-th region influentially related to the scalar predictor. Since $(1 - \nu_t)$ is the probability of a region not being "influential," it is referred to as the node sparsity parameter. For our simulation studies we consider two sparsity levels, $(1-\nu_t) = 0.85$ and $(1-\nu_t) = 0.70$. For each sparsity level, the coefficient corresponding to the edge connecting the p-th and p'-th region is drawn from the following mixture distribution,

$$\theta_{(p,p'),t} | \xi_{p,t}, \xi_{p',t} \sim \xi_{p,t}\xi_{p',t} N(\mu_{\theta,(p,p')}, \sigma^2_\theta) + (1 - \xi_{p,t}\xi_{p',t})\delta_0, \quad \theta_{(p,p'),t} = \theta_{(p',p),t}; \;\; p < p'. \tag{6}$$

(6) ensures that any edge connecting to the p-th region in the network response is unrelated to the predictor if the p-th region is un-influential, i.e., $\xi_{p,t} = 0 \Rightarrow \theta_{(p,p'),t} = 0$ for all $p' \in \{1, .., P\}$. Similarly, corresponding to each un-influential region $\mathcal{R}_p$, the $V \times 1$ dimensional GM coefficient $\boldsymbol{\beta}_{p,t}$ is set at $\mathbf{0}$. When $\xi_{p,t} = 1$, i.e., the p-th region is influential, we randomly choose $\upsilon_t = 0.4$ proportion of cell coefficients in the p-th region to be nonzero and rest are set at zero. These nonzero coefficients within $\boldsymbol{\beta}_{p,t}$ are simulated from $N(\mu_{\beta,(p,p')}, \sigma^2_\beta)$, where the values of $\mu_{\theta,(p,p')}$ and $\mu_{\beta,(p,p')}$ are drawn from $Unif(0.25, 1)$. All simulations fix

$\sigma_\theta^2 = 1$, $\sigma_\beta^2 = 1$ and the error variance $\tau_t^2 = 1$. For each of the two node sparsity levels, different number of ROIs and voxels within an ROI are considered. Specifically, we consider two cases, (a) "**small dimensional example:**" $P = 20$ and $V = 10$, and (b) "**high dimensional example:**" $P = 100$ and $V = 50$. In both these cases we have approximately similar number of parameters to estimate from the network and the structural coefficients.

5.2 Competitors and Metrics of Comparison

The simulated data will be used to assess the performance of: (A) identifying influential regions; (B) estimating the true network coefficient $\boldsymbol{\Theta}_t$ and structural coefficients $\boldsymbol{\beta}_{1,t}, ..., \boldsymbol{\beta}_{P,t}$; and (C) quantifying uncertainty in the point estimation of network and structural coefficients. We construct a series of competitors of BMRR to assess (A)-(C) as described below.

5.2.1 Frequentist competitors

As competitors, we implement popularly used approaches wherein each network edge and each cell of the structural image is independently regressed on the predictor to obtain p-values corresponding to the point estimates of $\theta_{p,p'}$ and $\beta_{v,p}$, denoted by p-value$(\hat{\theta}_{p,p'})$ and p-value$(\hat{\beta}_{v,p})$, respectively, for $v = 1, ..., V_p$ and $p = 1, ..., P$. These p-values will be compared to a threshold to declare if a region is influential, after correcting for the multiple comparison issue. Given the structured nature of our problem, we will consider several implementations of this correction as described below. We set $\alpha_0 = 0.05$ throughout.

Global and Regional Bonferroni's correction on structural images. The first competitor, referred to as the **Global Bonferroni** on Grey Matter, focuses on $\{\hat{\boldsymbol{\beta}}_1, ..., \hat{\boldsymbol{\beta}}_P\}$ together and declares a region p to be influential if p-value$(\hat{\beta}_{v,p}) \leq \frac{\alpha_0}{\sum_{p=1}^{P} V_p}$ for at least one $v \in \{1, ..., V_p\}$. As discussed in Genovese *et al.* (2002), applying Bonferroni's correc-

tion at a regional level may improve performance. Thus we consider a second competitor called **Regional Bonferroni** on Grey Matter, which selects a region p as influential if p-value$(\hat{\beta}_{v,p}) \leq \frac{\alpha_0}{V_p}$ for at least one $v \in \{1, ..., V_p\}$.

Global and Regional Bonferroni's correction on the network object. Similar to the structural images, the **Global Bonferroni** on the network object identifies region p as influential if at least one edge connecting to this region is influential, i.e., if p-value$(\hat{\theta}_{p,p'}) \leq \frac{\alpha_0}{P(P-1)/2}$ for at least one $p' \neq p$. In contrast, the **Regional Bonferroni** identifies region p to be influential if p-value$(\hat{\theta}_{p,p'}) \leq \frac{\alpha_0}{P-1}$ for at least one $p' \neq p$.

Global and Regional Bonferroni's correction on all objects jointly. **Global Bonferroni** on structural and network objects jointly determines region p to be influential if at least one edge connecting to the region or at least one voxel coefficient in region p is significant, i.e., p-value$(\hat{\theta}_{p,p'}) \leq \frac{\alpha_0}{P(P-1)/2+\sum_{p=1}^{P} V_p}$ for at least one $p' \neq p$ or p-value$(\hat{\beta}_{p,v}) \leq \frac{\alpha_0}{P(P-1)/2+\sum_{p=1}^{P} V_p}$ for at least one $v \in \{1, .., V_p\}$. The **Regional Bonferroni** for objects jointly identifies pth region as influential if p-value$(\hat{\theta}_{p,p'}) \leq \frac{\alpha_0}{(P-1)+V_p}$ for at least one $p' \in \{1, ..., P\}$ or p-value$(\hat{\beta}_{p,v}) \leq \frac{\alpha_0}{(P-1)+V_p}$ for at least one $v \in \{1, .., V_p\}$.

The second group of frequentist competitors are based on the mass univariate analysis (MUA) approach that controls the False Discovery Rate (FDR) (Genovese *et al.*, 2002).

Global and Regional MUA on the network and structural objects separately. The global implementation of MUA on the Network Matrix proceeds by ordering p-values corresponding to the $P(P-1)/2$ edge coefficients, denoted by p-value$_{(1)}, ...,$ p-value$_{\left(\frac{P(P-1)}{2}\right)}$. We consider p-value$_{(i^*)}$ to be the threshold where $i^* \in \{1, .., \frac{P(P-1)}{2}\}$ is the largest index such that p-value$_{(i)} \leq \frac{i}{\frac{P(P-1)}{2}}\alpha_0$. A region p is considered influential if p-value$(\hat{\theta}_{p,p'}) \leq$ p-value$_{(i^*)}$ for at least one $p' \neq p$. This approach is referred to as the **Global MUA** approach for the network object only. The **Regional MUA** for only the network object focuses on p-

values for coefficients of edges attached to a node/region. Let p-value$_{p,(1)}$, ..., p-value$_{p,(P-1)}$ be the ordering of such p-values for region p and i_p^* be the largest index in region p such that p-value$_{p,(i_p^*)} \leq \frac{i_p^*}{P-1}\alpha_0$. **Regional MUA** of the network object selects a region p if p-value$(\hat{\theta}_{p,p'}) \leq$ max(p-value$_{(i_p^*)}$, p-value$_{(i_{p'}^*)}$) for at least one $p' \neq p$. **Global MUA** and **Regional MUA** for structural objects are similarly defined.

Global and Regional MUA on the network and structural objects jointly. The **Global MUA** on structural and network objects jointly proceeds by ordering p-value$(\hat{\beta}_{v,p})$ and p-value$(\hat{\theta}_{p,p'})$ for all $v = 1, ..., V_p$, $1 \leq p < p' \leq P$ in ascending order. Let p-value$_{(1)}$, ..., p-value$_{\left(\frac{P(P-1)}{2}+\sum_{p=1}^{P} V_p\right)}$ be the ordered p-values and let i^* be the largest index such that p-value$_{(i^*)} \leq \frac{i^*\alpha_0}{\frac{P(P-1)}{2}+\sum_{p=1}^{P} V_p}$. A region p is identified as influential if p-value$(\hat{\theta}_{p,p'}) \leq$ p-value$_{(i^*)}$ for at least one $p' \neq p$, or p-value$(\hat{\beta}_{v,p}) \leq$ p-value$_{(i^*)}$ for at least one $v \in \{1, ..., V_p\}$. To implement **Regional MUA** jointly for both objects, the p-values corresponding to all cell coefficients for the pth region and edge coefficients connecting to the pth node/region are ordered in an ascending manner, p-value$_{p,(1)} \leq \ldots \leq$ p-value$_{p,(P-1+V_p)}$. Let i_p^* be the largest index such that p-value$_{p,(i_p^*)} \leq \frac{i_p^*\alpha_0}{P-1+V_p}$. The region p is considered influential if p-value$(\hat{\theta}_{p,p'}) \leq$ max(p-value$_{(i_p^*)}$, p-value$_{(i_{p'}^*)}$) for at least one $p' \neq p$, or p-value$(\hat{\beta}_{v,p}) \leq$ p-value$_{(i_p^*)}$ for at least one cell $v \in \{1, ..., V_p\}$ in the pth region.

5.2.2 Bayesian competitors

We implement two Bayesian competitors, both capturing joint effects of all network edges and cells of the structural images on the predictor, but do not acknowledge the connection between the two sets of objects. The first Bayesian competitor, referred to as the Spike & Slab, applies an ordinary spike & slab prior (George and McCulloch, 1993) on each $\theta_{p,p'}$ and $\beta_{v,p}$. A voxel or an edge will be identified as influential if the posterior probability of corresponding coefficient equals to zero exceeds 0.5 (Barbieri and Berger, 2004). To assess

joint modeling of objects vs. modeling them individually, we fit only the network on scalar regression (i.e., the first equation in (3)), or the structural images on scalar regression (i.e., the second equation in (3), or both of them jointly. For the network on scalar regression Spike & Slab competitor identifies a region to be influential if at least an edge connecting to that region is influential. Spike & Slab competitor for structural images on scalar regression identifies a region as influential if it has at least one influential voxel in that region. When both objects are used jointly, Spike & Slab identifies an influential region by observing if any voxel in that region or any edge connected to that region is influential.

Our second Bayesian competitor applies a Horseshoe shrinkage prior (Carvalho *et al.*, 2010) on each $\theta_{p,p'}$ and $\beta_{v,p}$. Since the Horseshoe prior does not result in exact zeros for the iterates of coefficients, we perform a post-processing approach (Guha and Rodriguez, 2021) to determine which voxel coefficients and network edges are related to the predictor. We identify influential regions using these estimates either on both sets of objects separately, or jointly, following the similar strategy outlined in the Spike & Slab competitor.

5.2.3 Metrics of comparison

We will present True Positive Rate (TPR) and True Negative Rate (TNR) for correctly identifying important regions, corresponding to all simulation cases. Since there is a natural trade-off between TPR and TNR, we will present a single measure F_1 score to balance the performance between true positives and true negatives for all competitors.

The point estimation of every competitor is assessed using mean squared errors (MSE) of estimating the network coefficient $\boldsymbol{\Theta}_t$ and the structural coefficients $\boldsymbol{\beta}_{1,t}, ..., \boldsymbol{\beta}_{P,t}$. The MSE for network and structural coefficients are given by $2\sum_{p<p'}(\theta_{p,p',t} - \widehat{\theta}_{p,p'})^2/P(P-1)$ and $\sum_{p=1}^{P} ||\boldsymbol{\beta}_{p,t} - \widehat{\boldsymbol{\beta}}_p||^2/VP$, respectively, where $\widehat{\theta}_{p,p'}$ is the point estimate of $\theta_{p,p'}$, and $\widehat{\boldsymbol{\beta}}_p$ represents the point estimate of $\boldsymbol{\beta}_p$. MSE for both sets of coefficients jointly is given by

$[\sum_{p<p',t}(\theta_{p,p',t} - \widehat{\theta}_{p,p'})^2 + \sum_{p=1}^{P} ||\boldsymbol{\beta}_{p,t} - \widehat{\boldsymbol{\beta}}_p||^2]/(P(P-1)/2 + VP)$. The point estimates are taken to be the posterior median for the Bayesian competitors. For Bayesian competitors, we evaluate the length and coverage of 95% credible intervals averaged across coefficients in $\boldsymbol{\Theta}$ and present them for all cases. Similar quantities for $\boldsymbol{\beta}_p$'s are also presented.

5.3 Results

The Table 1 shows the results for different competitors in terms of influential region identification. All frequentist competitors identify influential nodes using a two-stage approach. Hence, their performances are evaluated applying the second stage on the structural and network objects jointly or separately, as described in Section 5.2. Bayesian competitors Horseshoe and Spike & Slab are fitted on structural and network objects separately as well as jointly. Hence TPR and TNR for these methods are recorded when they are applied on individual objects or they are applied jointly. In contrast, identification of influential regions is obtained as part of the inference from BMRR. Hence, we just show the TPR and TNR results for BMRR under the "joint object" column in Table 1. A few interesting patterns emerge from the table. First, both Regional MUA and Regional Bonferroni perform considerably better than Global MUA and Global bonferroni, respectively, in terms of TPR. In contrast, the global implementations of these methods perform marginally better than regional implementation in terms of TNR. This is due to the fact that the global implementations are more conservative than the corresponding regional implementations. Second, all frequentist and Bayesian competitors show significantly better TPR when they consider both objects jointly as opposed to inferring on them separately. No significant differences are found for TNR in this aspect. Third, the simulation cases under the smaller dimensional example show significantly better performance for all competitors than the

higher dimensional examples. Fourth, Horseshoe prior applied on joint modeling of objects yield significantly better results in small dimensional example than applying Horseshoe prior in the scenario where the two sets of objects are fitted separately. This also serves as a highlighting point of modeling objects jointly to draw better inference. Finally, our proposed approach BMRR demonstrates better TPR and TNR than Horseshoe. Spike & Slab approach performs similar to BMRR in terms of TPR, but massively underperforms in terms of TNR. In contrast, BMRR significantly outperforms frequentist competitors in terms of TPR and shows marginally inferior performance to them in terms of TNR. Since there is a trade-off between TPR and TNR performance, we show results for a combining measure, F_1-score, for all competitors in the supplementary file. Figure 1 in the supplementary file demonstrates superior performance of BMRR w.r.t. its competitors.

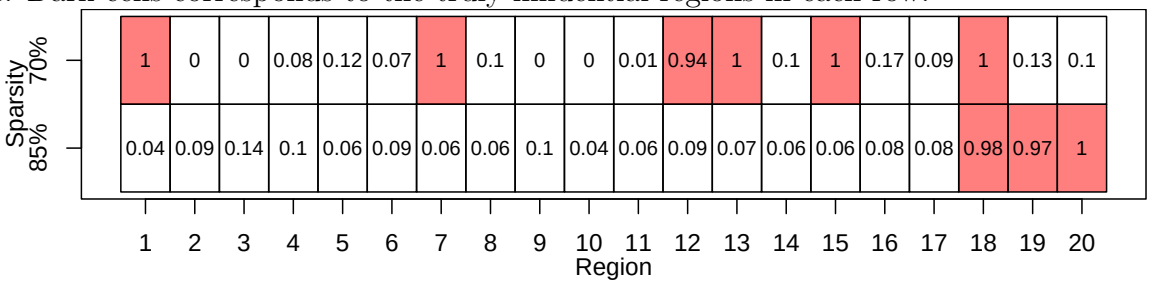
One notable advantage of BMRR over its competitors is that it is able to provide characterization of uncertainty for each region being influential by offering posterior probability of $\{\xi_p = 1\}$. Figure 2 shows the posterior probability of the pth region being influential for small dimensional simulation example under two different node sparsities, with dark cells corresponding to the regions which are truly influential. We present the plot for the large dimensional example in Figure 2 of the supplementary file due to space constraint. The results suggest that the posterior probability of a region being influential is close to 1 or 0 if the region is truly influential and un-influential, respectively.

In terms of estimating the regression coefficients, Figure 3 shows significantly better performance by BMRR over its competitors. The significant improvement of point estimation by BMRR demonstrates the importance of exploiting multi-object topology to offer better point estimation when sample size is limited. We also observe 95% confidence/credible intervals from BMRR attend close to the nominal level in all simulations,

Table 1: True Positive Rate (TPR) and True Negative Rate (TNR) for competing methods under smaller dimensional and high dimensional examples. Both smaller and higher dimensional examples include two different node sparsity levels. We compute TPR and TNR separately for grey matter and network matrix, as well as for them together, as explained in Section 5.2. The best performing method in each simulation case is boldfaced.

	Sparsity 85%						Sparsity 70%					
	TPR			TNR			TPR			TNR		
Method	Grey Matter	Network Matrix	Joint Objects	Grey Matter	Network Matrix	Joint Objects	Grey Matter	Network Matrix	Joint Objects	Grey Matter	Network Matrix	Joint Objects
High dimensional example												
Global Bonferroni	0.0947	0.1580	0.1633	**0.9985**	**0.9920**	**0.9951**	0.2443	0.3757	0.3920	**0.9950**	**0.9940**	**0.9960**
Regional Bonferroni	0.3667	0.4167	0.5180	0.9359	0.9528	0.9411	0.6743	0.7523	0.8430	0.9389	0.9610	0.9533
Global MUA	0.1313	0.1973	0.2213	0.9693	0.9547	0.9564	0.4030	0.5420	0.6777	0.9557	0.9493	0.9419
Regional MUA	0.3700	0.4447	0.5413	0.9355	0.9355	0.9274	0.6913	0.7810	0.8737	0.9373	0.9440	0.9400
Horseshoe	0.5320	**0.7400**	0.6960	0.7701	0.6956	0.7981	0.7170	0.8613	0.8593	0.8177	0.7249	0.9153
BMRR	–	–	**0.6967**	–	–	0.8666	–	–	**0.9703**	–	–	0.9460
Spike and Slab	0.3040	0.4007	0.5120	0.8866	0.8576	0.8524	0.7267	0.8323	0.9527	0.7727	0.7041	0.6161
Small dimensional example												
Global Bonferroni	0.1339	0.1853	0.1965	**0.9990**	**0.9988**	**0.9991**	0.2517	0.3484	0.3468	0.9991	0.9989	0.9994
Regional Bonferroni	0.8637	0.8393	0.9352	0.9244	0.9452	0.9345	0.9788	0.9759	0.9949	0.9258	0.9362	0.9261
Global MUA	0.4496	0.4285	0.6400	0.9667	0.9814	0.9648	0.9847	0.9787	0.9997	0.8875	0.9082	0.8186
Regional MUA	0.8997	0.8781	0.9663	0.9216	0.9276	0.9244	0.9935	0.9917	0.9995	0.9238	0.9241	0.9184
Horseshoe	0.9947	0.9875	0.9852	0.3739	0.5598	0.9187	1.0000	1.0000	1.0000	0.2855	0.3577	0.7693
BMRR	–	–	**0.9997**	–	–	0.9454	–	–	**1.0000**	–	–	**0.9996**
Spike and Slab	0.9996	0.9960	0.9997	0.0769	0.3616	0.0020	1.0000	1.0000	1.0000	0.0285	0.0960	0.0020

Figure 2: The figure presents posterior probability of each region being influential under the two simulation cases in smaller dimensional example. Each row corresponds to a simulation case. Dark cells corresponds to the truly influential regions in each row.



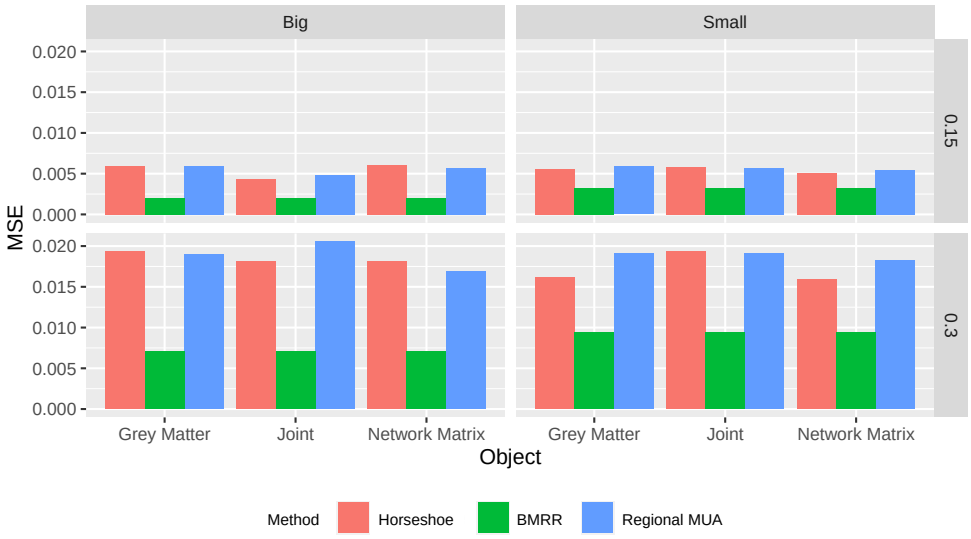
see Figure 4. Regional MUA also enjoys close to nominal coverage. While Horseshoe fitted with both objects jointly yields coverage similar to its competitors, severe under-coverage is observed when Horseshoe is applied on structural and network objects separately. Importantly, BMRR shows similar coverage with much narrower credible intervals than its competitors, as demonstrated in Figure 5. The 95% CIs estimated from Regional MUA are much wider than both its competitors, which is explained by the fact that it estimates every cell of the structural and network coefficient independently of each other.

6 Application to the multi-modal PPA data

6.1 Data structure and methods

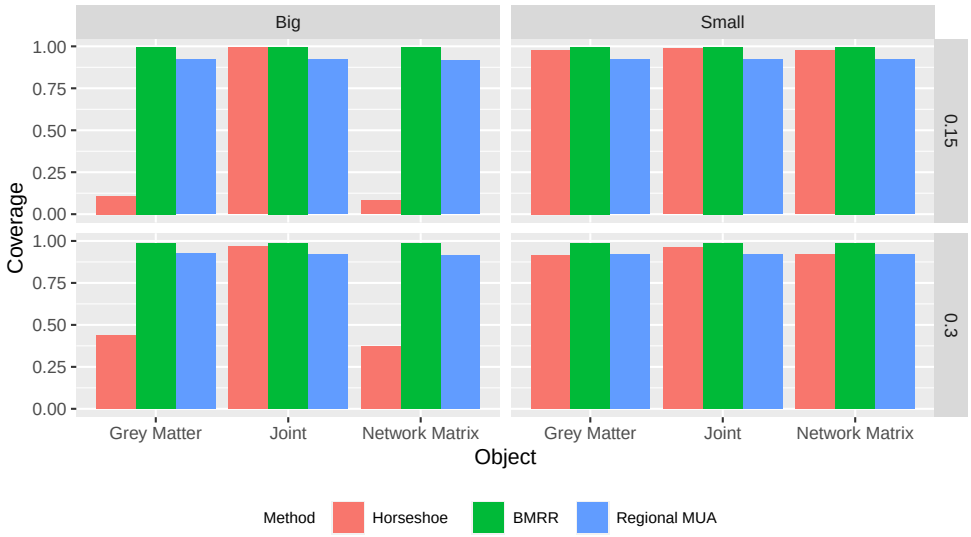
In our motivating study, sMRI GM, resting state fMRI connectivity images, and speech rates (described in Section 2) were collected at the University of California, San Francisco Memory and Aging Center on 24 subjects with nfvPPA (10 males and 14 females) aged between 57 and 81 years (68.9 years old on average). sMRI GM maps were preprocessed as in Canu *et al.* (2020). For the fMRI network images, the first five volumes of the acquisition were discarded to allow T1 equilibrium to be established. The remaining volumes were

Figure 3: MSE for the Horseshoe, the selected MUA approach and BMRR for different simulation scenarios. Here "Big" and "Small" refer to high-dimensional and smaller dimensional examples. The top and bottom row correspond to the true sparsity level of 85% and 70%, respectively.



slice-time corrected, realigned to the mean functional image, and assessed for rotational and translational head motion. The functional volumes were then linearly registered to an Echo Planar Imaging (EPI) template created from the mean functional images of the participants enrolled in this study. The EPI template was normalized to the MNI space using a combination of linear and non-linear warping, and the transformations were subsequently applied to the task-free fMRI data. These volumes were spatially smoothed with a 5 mm full width at half maximum (FWHM) Gaussian kernel. CSF and white matter (WM) tissue probability maps were then used to compute the mean time-series used as regressors. Functional data were then bandpass filtered (0.008 Hz $< f <$ 0.15 Hz), and the nuisance variables were regressed out from the data.We adjust our analysis for biological sex, age, and cognitive function measured via the mini-mental state examination (MMSE) scale as a measure of dementia severity as in García *et al.* (2022).

Figure 4: Coverage of 95% confidence/credible intervals for the selected MUA approach and BMRR for different simulation scenarios. Here "Big" and "Small" refer to high-dimensional and smaller dimensional examples. The top and bottom row correspond to the true sparsity level of 85% and 70%, respectively.

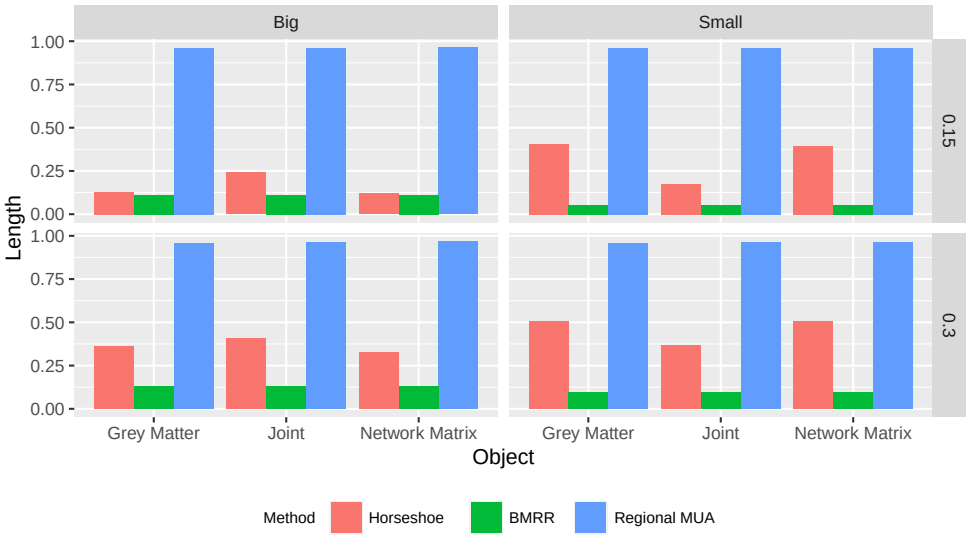


The fMRI network connectivity images and sMRI Grey Matter images were centered in such a way that the average across subjects for each transformed Z-score and grey matter probability is 0. The biological sex indicator is encoded as 1 for male and 0 for female. The MMSE is centered. Finally the speech rate variable has been normalized. Since simulation experiments demonstrate Regional MUA and Horseshoe as the two best performing competitors, we compare the results from the application of the proposed BMRR to the MUA with regional correction (Regional MUA) and Horseshoe.

6.2 Data analysis results

We present the results from our application of the proposed BMRR to the multi-modal PPA imaging data. In total, 35 ROIs out of 245 ROIs have a posterior probability $P(\xi_p = 1|\text{Data}) > 0.5$ of being associated with speech rate in the multi-modal data. The set of

Figure 5: Length of 95% confidence/credible intervals for the selected MUA approach and BMRR for different simulation scenarios. Here "Big" and "Small" refer to high-dimensional and smaller dimensional examples. The top and bottom row correspond to the true sparsity level of 85% and 70%, respectively.



influential ROIs and their posterior probability of being influential (which is a measure of uncertainty) are recorded in Table 2. In general, the ROIs lie in areas of the brain previously associated with language, motor speech, and neurodegeneration in PPA, specifically the frontal lobe including the superior frontal gyrus (SFG) (Landin-Romero *et al.*, 2021), the middle frontal gyrus (MFG) (Landin-Romero *et al.*, 2021), the inferior frontal gyrus (IFG) (García *et al.*, 2022), and the superior parietal lobe (SPL) (Alahmadi, 2021). The 35 ROIs display a bilateral distribution which is consistent with previous findings that indicate that articulation can be impacted by disruptions in the motor speech network in either the left or right hemisphere of the motor speech network (Landin-Romero *et al.*, 2021).

Figure 6 displays horizontal slices (higher z coordinate refers to top of the brain) of the GM coefficient map organized by ROI where the colored shading indicates the magnitude and direction of association of the speech rate with probability of GM. Due to the

Table 2: Influential ROIs identified by BMRR in the PPA data analysis along with their posterior probabilities of being influential.

ROI	Prob.	ROI	Prob.	ROI	Prob.	ROI	Prob.
MFG_L_7_1	1.000	MFG_R_7_7	0.944	SFG_L_7_1	0.972	IPL_R_6_5	1.000
MFG_R_7_1	1.000	IFG_R_6_1	0.676	SFG_L_7_3	0.636	PCun_L_4_1	0.78
MFG_L_7_3	1.000	IFG_L_6_2	0.840	SFG_R_7_3	0.992	PCun_R_4_1	1.000
MFG_R_7_3	0.972	IFG_R_6_2	1.000	OrG_R_6_6	1.000	PCun_R_4_2	0.620
MFG_L_7_4	1.000	IFG_L_6_3	1.000	PrG_L_6_2	0.976	PCun_R_4_4	0.620
MFG_L_7_5	0.836	IFG_R_6_3	1.000	ITG_R_7_2	0.914	CG_L_7_2	0.862
MFG_R_7_5	1.000	IFG_L_6_4	0.850	SPL_R_5_2	1.000	CG_R_7_3	1.000
MFG_L_7_6	1.000	IFG_R_6_5	0.798	SPL_R_5_3	0.870	CG_R_7_5	1.000
MFG_R_7_6	0.640	IFG_L_6_6	0.624	SPL_L_5_5	1.000		

hierarchical nature of the BMRR model, the distribution of influential GM voxels (as captured by their associated coefficients) is determined by the ROI selection and thus reflects the distribution of the selected 35 ROIs discussed above, with concentrations in the SFG, MFG, IFG, and SPL in a bilateral distribution. Overall, speech rate is positively associated with increases in voxel level GM which affirms that a decrease in GM in regions in the frontal gyrus (Landin-Romero *et al.*, 2021; García *et al.*, 2022) is correlated with lower articulation rates. But, there is heterogeneity in the voxel level signal both within and across ROIs with a substantial number of voxels displaying a negative association of GM with speech rate. The effect sizes range from -0.241 to 0.200 across voxels which indicates that a standard deviation shift in speech rate (25 words/minute) can produce a change in GM probability from between -0.24 to 0.24. These results reveal a dynamic portrait of how GM and neurodegeneration are associated with motor speech and highlights heterogeneity

in the direction of the association as well as varying levels of sparsity in influential voxels.

Figure 6: Point estimates (median) of the regression coefficients for grey matter probability for selected voxels across horizontal slices using BMRR with speech rate as outcome.

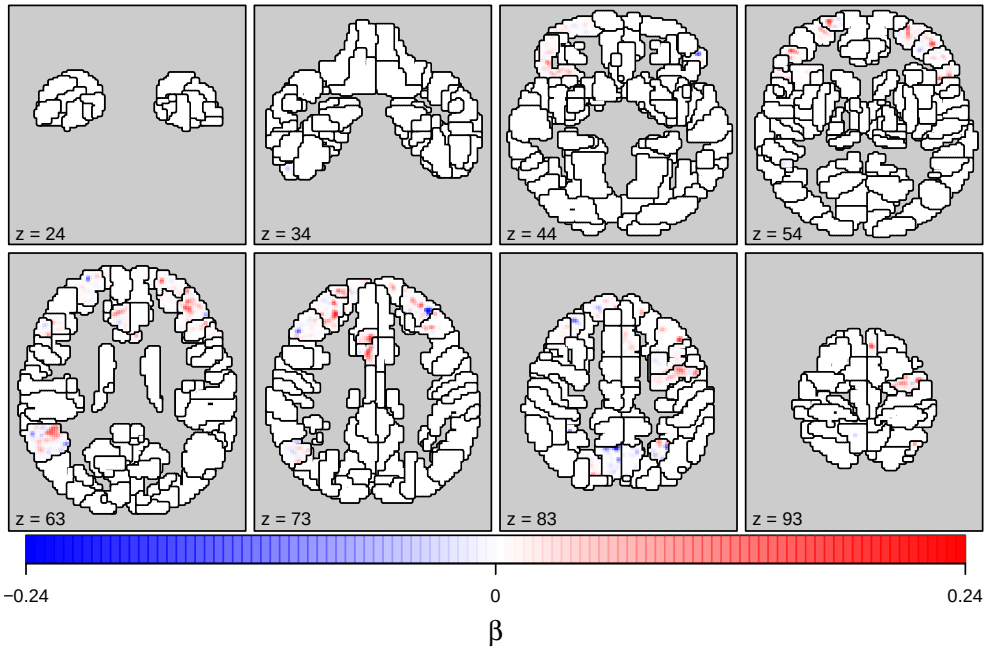
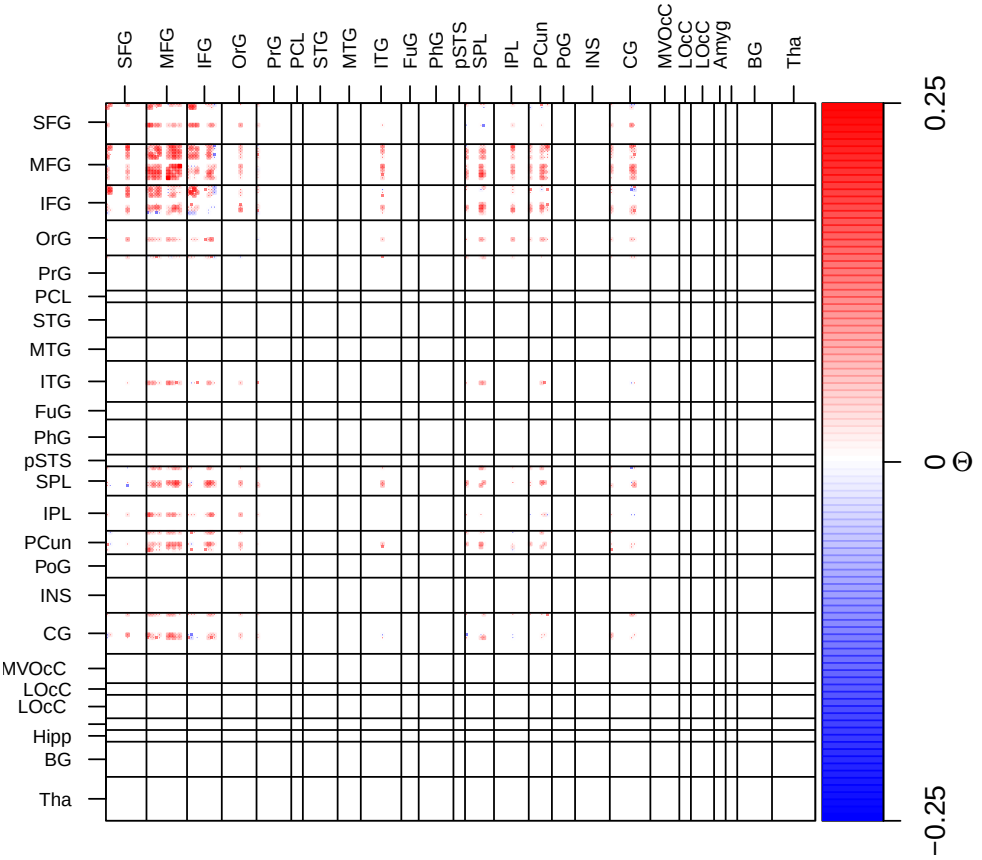


Figure 7 displays the ROI level network coefficient maps organized by anatomical brain region (e.g. superior frontal gyrus = SFG, see Fan *et al.* (2016) for full details) where the top and bottom half of each anatomical region indicated by a horizontal dash corresponds to the left hemisphere and right hemispheres, respectively. The colored shading indicates the magnitude and direction of association of the speech rate with resting state functional connectivity. It appears that the network edges associated with the speech rate are sparse and there is almost always a positive association between functional connectivity and speech rate. There is a high degree of bilateral functional connectivity associated with speech rate both among and between the SFG, MFG, and IFG brain regions in the frontal lobe. Connections extend outside the frontal lobe and higher speech rates are also associated with higher functional connectivity between the frontal lobe and the parietal lobe which has been indicated in higher order language processing (Coslett and Schwartz, 2018), including the

SPL, inferior parietal lobe (IPL), and the precuneus (PCun). The effect sizes range from -0.047 to 0.251 across network edges which indicates that a standard deviation shift in speech rate can produce a change in Fisher Z-transformed correlation between -0.25 to 0.25. Supplementary Figures 3-5 show the GM coefficient maps for the covariates gender, age, and cognitive decline, and Figures 6-8 show the same for network coefficient maps. Interpretation of these maps is also available in Section 2.2 of the supplementary file.

Figure 7: Point estimates (median) of the regression coefficients for Network Matrix using BMRR. Each cell displays the ROI level network coefficient maps organized by anatomical brain region where the top and bottom half of each anatomical region indicated by a horizontal dash corresponds to the left hemisphere and right hemispheres, respectively.



While comparing BMRR with regional MUA, Figures 9 and 10 in the supplementary file show that very few GM voxels and Network edges are selected by regional MUA. This

Table 3: Posterior Predictive Loss Criterion (PPLC) for BMRR and Horseshoe.

Method	G	P	PPLC
BMRR	1557.506	35104.18	36661.686
Horseshoe	1796.926	35102.32	36899.246

is expected as we observed in simulations that regional MUA tends to have a lower TPR in contrast to BMMR. We also contrast BMRR with the horseshoe implementation in terms of the Posterior Predictive Loss Criterion (PPLC) (Gelfand and Ghosh, 1998) that shows better fit by BMRR (see Table 3). Overall, BMRR develops novel insight on the neuro-degeneration pathway for PPA and outperforms competitors due to accounting for science driven constraints in the prior structure for coefficients.

7 Conclusion and Future Work

This article develops a regression approach with structural and network-valued objects on a scalar predictor. A novel prior structure is developed jointly on coefficients corresponding to different object responses which can exploit linked topology of these objects to draw inference on network nodes significantly related to the scalar predictor with uncertainty. The proposed approach is arguably the first statistical multi-modal response regression approach that allows for flexibility of drawing inference at the ROI-level and at the voxel-level simultaneously, and equipped with identifying brain regions significantly related to speech rate measuring neuro-degeneration due to PPA with uncertainty. The analysis of PPA data leads to important understanding of neuro-degeneration pathway for PPA.

As an immediate future work, we will extend our approach to incorporate the spatial correlation in the GM image. We also plan to capture more complex non-linear dependence

between nodes in the network object and cells in the structural objects while identifying their join relationship with the speech rate in our upcoming articles.

8 Acknowledgement

.

References

Alahmadi, A. A. S. (2021). Investigating the sub-regions of the superior parietal cortex using functional magnetic resonance imaging connectivity. *Insights Imaging*, **12**(1), 47.

Barbieri, M. M. and Berger, J. O. (2004). Optimal predictive model selection. *The annals of statistics*, **32**(3), 870–897.

Battistella, G., Borghesani, V., Henry, M., Shwe, W., Lauricella, M., Miller, Z., Deleon, J., Miller, B. L., Dronkers, N., Brambati, S. M., Seeley, W. W., Mandelli, M. L., and Gorno-Tempini, M. L. (2020). Task-Free functional language networks: Reproducibility and clinical application. *J. Neurosci.*, **40**(6), 1311–1320.

Calhoun, V. D. and Sui, J. (2016). Multimodal fusion of brain imaging data: A key to finding the missing link (s) in complex mental illness. *Biological psychiatry: cognitive neuroscience and neuroimaging*, **1**(3), 230–244.

Canu, E., Agosta, F., Battistella, G., Spinelli, E. G., DeLeon, J., Welch, A. E., Mandelli, M. L., Hubbard, H. I., Moro, A., Magnani, G., Cappa, S. F., Miller, B. L., Filippi, M., and Gorno-Tempini, M. L. (2020). Speech production differences in english and italian speakers with nonfluent variant PPA. *Neurology*, **94**(10), e1062–e1072.

Carvalho, C. M., Polson, N. G., and Scott, J. G. (2010). The horseshoe estimator for sparse signals. *Biometrika*, **97**(2), 465–480.

Chen, L. and Huang, J. Z. (2012). Sparse reduced-rank regression for simultaneous dimension reduction and variable selection. *Journal of the American Statistical Association*, **107**(500), 1533–1545.

Coslett, H. B. and Schwartz, M. F. (2018). The parietal lobe and language. *Handb. Clin. Neurol.*, **151**, 365–375.

Dai, X. and Li, L. (2021). Orthogonal statistical inference for multimodal data analysis. *arXiv preprint arXiv:2103.07088*.

Fan, J., Gong, W., and Zhu, Z. (2019). Generalized high-dimensional trace regression via nuclear norm regularization. *Journal of econometrics*, **212**(1), 177–202.

Fan, L., Li, H., Zhuo, J., Zhang, Y., Wang, J., Chen, L., Yang, Z., Chu, C., Xie, S., Laird, A. R., Fox, P. T., Eickhoff, S. B., Yu, C., and Jiang, T. (2016). The human brainnetome atlas: A new brain atlas based on connectional architecture. *Cereb Cortex*, **26**(8), 3508–3526.

Feng, X., Li, T., Song, X., and Zhu, H. (2019). Bayesian scalar on image regression with nonignorable nonresponse. *Journal of the American Statistical Association*, pages 1–24.

Friston, K. J. (2003). Statistical parametric mapping. In *Neuroscience databases*, pages 237–250. Springer.

Gabriel, K. R. (1998). Generalised bilinear regression. *Biometrika*, **85**(3), 689–700.

García, A. M., Welch, A. E., Mandelli, M. L., Henry, M. L., Lukic, S., Torres-Prioris, M. J., Deleon, J., Ratnasiri, B. M., Lorca Puls, D. L., Miller, B. L., Seeley, W. W., Vogel, A. P.,

and Gorno-Tempini, M. L. (2022). Automated detection of speech timing alterations in autopsy-confirmed non-fluent/agrammatic variant primary progressive aphasia.

Gelfand, A. E. and Ghosh, S. K. (1998). Model choice: a minimum posterior predictive loss approach. *Biometrika*, **85**(1), 1–11.

Genovese, C. R., Lazar, N. A., and Nichols, T. (2002). Thresholding of Statistical Maps in Functional Neuroimaging Using the False Discovery Rate. *NeuroImage*, **15**(4), 870–878.

George, E. I. and McCulloch, R. E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, **88**(423), 881–889.

Goh, G., Dey, D. K., and Chen, K. (2017). Bayesian sparse reduced rank multivariate regression. *Journal of multivariate analysis*, **157**, 14–28.

Gorno-Tempini, M. L., Brambati, S. M., Ginex, V., Ogar, J., Dronkers, N. F., Marcone, A., Perani, D., Garibotto, V., Cappa, S. F., and Miller, B. L. (2008). The logopenic/phonological variant of primary progressive aphasia. *Neurology*, **71**(16), 1227–1234.

Guha, S. and Guhaniyogi, R. (2021). Bayesian generalized sparse symmetric tensor-on-vector regression. *Technometrics*, **63**(2), 160–170.

Guha, S. and Rodriguez, A. (2021). Bayesian regression with undirected Network predictors with an application to brain connectome data. *Journal of the American Statistical Association*, **116**(534), 581–593.

Guhaniyogi, R. and Spencer, D. (2021). Bayesian tensor response regression with an application to brain activation studies. *Bayesian Analysis*, **16**(4), 1221–1249.

Guhaniyogi, R., Qamar, S., and Dunson, D. B. (2017). Bayesian tensor regression. *The Journal of Machine Learning Research*, **18**(1), 2733–2763.

Hoff, P. D. (2005). Bilinear mixed-effects models for dyadic data. *Journal of the American Statistical Association*, **100**(469), 286–295.

Landin-Romero, R., Liang, C. T., Monroe, P. A., Higashiyama, Y., Leyton, C. E., Hodges, J. R., Piguet, O., and Ballard, K. J. (2021). Brain changes underlying progression of speech motor programming impairment. *Brain Commun*, **3**(3), fcab205.

Lazar, N. A. (2016). Corrections for multiplicity in functional neuroimaging data. *Handbook of Neuroimaging Data Analysis*, page 355.

Li, L., Kang, J., Lockhart, S. N., Adams, J., and Jagust, W. J. (2018). Spatially adaptive varying correlation analysis for multimodal neuroimaging data. *IEEE transactions on medical imaging*, **38**(1), 113–123.

Li, Q. and Li, L. (2021). Integrative factor regression and its inference for multimodal data analysis. *Journal of the American Statistical Association*, pages 1–15.

Li, Y., Qin, Y., Chen, X., and Li, W. (2013). Exploring the functional brain network of alzheimer's disease: based on the computational experiment. *PloS one*, **8**(9), e73186.

Lock, E. F., Hoadley, K. A., Marron, J. S., and Nobel, A. B. (2013). Joint and individual variation explained (jive) for integrated analysis of multiple data types. *The annals of applied statistics*, **7**(1), 523.

Mandelli, M. L., Vilaplana, E., Brown, J. A., Hubbard, H. I., Binney, R. J., Attygalle, S., Santos-Santos, M. A., Miller, Z. A., Pakvasa, M., Henry, M. L., Rosen, H. J., Henry, R. G., Rabinovici, G. D., Miller, B. L., Seeley, W. W., and Gorno-Tempini, M. L. (2016).

Healthy brain connectivity predicts atrophy progression in non-fluent variant of primary progressive aphasia. *Brain*, **139**(Pt 10), 2778–2791.

Ogar, J. M., Dronkers, N. F., Brambati, S. M., Miller, B. L., and Gorno-Tempini, M. L. (2007). Progressive nonfluent aphasia and its characteristic motor speech deficits. *Alzheimer Dis. Assoc. Disord.*, **21**(4), S23–30.

Scott, J. G. and Berger, J. O. (2010). Bayes and empirical-Bayes multiplicity adjustment in the variable-selection problem. *The Annals of Statistics*, **38**(5), 2587–2619.

Von Rosen, D. (2018). Bilinear regression analysis. *Lecture notes in statistics*, **220**.

Xue, W., Bowman, F. D., Pileggi, A. V., and Mayer, A. R. (2015). A multimodal approach for determining brain networks by jointly modeling functional and structural connectivity. *Frontiers in computational neuroscience*, **9**, 22.

Xue, W., Bowman, F. D., and Kang, J. (2018). A Bayesian spatial model to predict disease status using imaging data from various modalities. *Frontiers in neuroscience*, **12**, 184.

Yuan, M., Ekici, A., Lu, Z., and Monteiro, R. (2007). Dimension reduction and coefficient estimation in multivariate linear regression. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, **69**(3), 329–346.