HHS Public Access

Author manuscript

Biometrics. Author manuscript; available in PMC 2025 October 09.

Published in final edited form as:

Biometrics. 2025 July 03; 81(3): . doi:10.1093/biomtc/ujaf097.

Covariance-on-covariance regression

Yi Zhao[1,*], **Yize Zhao**[2]

[1]Department of Biostatistics and Health Data Science, Indiana University School of Medicine, 410 West 10th Street, Indianapolis, IN 46202, United States

[2]Department of Biostatistics, Yale School of Public Health, New Haven, CT 06511, United States

Abstract

A covariance-on-covariance regression model is introduced in this manuscript. It is assumed that there exists (at least) a pair of linear projections on outcome covariance matrices and predictor covariance matrices such that a log-linear model links the variances in the projection spaces, as well as additional covariates of interest. An ordinary least square type of estimator is proposed to simultaneously identify the projections and estimate model coefficients. Under regularity conditions, the proposed estimator is asymptotically consistent. The superior performance of the proposed approach over existing methods is demonstrated via simulation studies. Applying to data collected in the Human Connectome Project Aging study, the proposed approach identifies 3 pairs of brain networks, where functional connectivity within the resting-state network predicts functional connectivity within the corresponding task-state network. The 3 networks correspond to a global signal network, a task-related network, and a task-unrelated network. The findings are consistent with existing knowledge about brain function.

Keywords

common diagonalization; generalized linear model; linear projection; ordinary least squares

1 INTRODUCTION

In this manuscript, a covariance-on-covariance regression (CoCReg) problem is studied. This is motivated by functional magnetic resonance imaging (fMRI) experiments. Two types of fMRI experiments are widely studied, resting-state (rs-fMRI) and task-based (tb-fMRI). In an rs-fMRI experiment, participants lie in the scanner at rest with eyes open. The study interest is to characterize the coactivation pattern between brain regions, the so-called brain functional connectivity (FC), captured by the correlation between fMRI time courses. In a tb-fMRI experiment, participants are instructed to complete a certain task during the scan. Such type of fMRI experiments attempts to identify activated brain regions in response to task stimuli and depict task-state functional networks. Building upon the assumption that

[*]Corresponding author: Yi Zhao, Department of Biostatistics and Health Data Science, Indiana University School of Medicine, 410 West 10th Street, Indianapolis, IN 46202, United States (yz125@iu.edu).

individual variation in brain responses to a task is intrinsic in the brain and predictable, existing research demonstrates that one can use task-free measurements to predict task-related activations (Tavor et al., 2016; Ngo et al., 2022). In tb-fMRI experiments, it is essential to study FC modulated by task demands. When performing cognitive tasks, small alterations from the intrinsic network organization occur and strongly contribute to task performance (Cole et al., 2021), where the intrinsic network organization can be captured by the FC map using rs-fMRI. Considering the fact that cognitive task activations emerge through network interactions, predicting task-state FC from resting-state is feasible and promising. In the motivating Lifespan Human Connectome Project (HCP) Aging study, both rs-fMRI and tb-fMRI are collected from the same individual. One tb-fMRI is the FACENAME task aiming to locate encoding-related and retrieval-related neural processes. With the availability of rs-fMRI scans, it is of interest to study how informative the intrinsic network organization is to the task-related and extrinsic network organization, especially among brain regions related to working memory and visual processing. Putting it into a statistical framework, it is a regression problem with a covariance matrix (tb-connectivity) as the outcome and a covariance matrix (rs-connectivity) as one of the predictors.

Popular approaches of using the rs-connectivity features to predict tb-connectivity include the general linear model and statistical/machine learning techniques enriching the collection of literature on “connectome fingerprinting” (Tavor et al., 2016). After choosing a brain parcellation, rs-connectivity and tb-connectivity are calculated for each pair of brain regions. One straightforward approach is to fit models to each pair of tb-connectivity and loop over all possible pairs, where in each model, all pairs of rs-connectivity can be entered as predictors. This approach, however, ignores the structure of the connectivity matrices (both tb-connectivity and rs-connectivity), represented by the covariance matrix of the standardized fMRI time series, such as the positive definiteness and topological architecture, and suffers from a deficient power due to multiplicity. Here, the topological architecture of a covariance matrix infers the relationship between brain regions through the structure of its eigenvalues and eigenvectors. Considering covariance matrices as the outcome, a type of regression model, called covariance regression, was introduced. Examples include modeling the covariance matrix as a quadratic function of the covariates (Hoff and Niu, 2012; Seiler and Holmes, 2017) or a linear combination of similarity matrices of the covariates (Zou et al., 2017), nonparametric covariance regression utilizing low-rank approximation (Fox and Dunson, 2015), a Bayesian hierarchical modeling of covariance matrices accounting for individual heterogeneity (Acharyya et al., 2023), common diagonalization based on eigendecomposition (Flury, 1984; Franks and Hoff, 2019; Boik, 2002; Hoff, 2009) and Cholesky decomposition (Pourahmadi et al., 2007), and a recent approach of covariate assisted principal regression for covariance matrix outcomes (Zhao et al., 2021b). In Zhao et al. (2021b), it is assumed that there exists a common diagonalization on the covariance matrices and the corresponding diagonal elements satisfy a log-linear model on the covariates. The advantage is to preserve the positive definiteness of the covariance matrices and offer high flexibility in parsimonious modeling. In this study, building upon this idea of common diagonalization, an approach to perform CoCReg is introduced. Common linear projections are assumed for outcome covariance matrices and predictor covariance matrices and a log-linear model is assumed for the projected data and scalar

covariates. The objective is to identify the projections and simultaneously estimate model coefficients. With proper thresholding or sparsifying on the loading profiles, the model offers a network-level interpretation, that is the resting-state FC within the network can predict the connectivity between regions within the corresponding task-based network. For tb-fMRI signals, normality cannot be assumed. An ordinary least squares (OLS) type of estimator is proposed to relax distributional assumption and asymptotic consistency is achieved.

Another group of related work is the image-on-image regression. Here, an image means a vector of scalar outcomes with spatial information. Compared with image-on-scalar and scalar-on-image regression, image-on-image has been less explored but received growing attention along the need in modern neuroscientific research. Guo et al. (2022) developed a spatial Bayesian latent factor model to predict individual task-evoked images using the corresponding task-independent images. Data dimension was significantly reduced by choosing a small number of latent factors, and simultaneously taking spatial dependence into consideration. Some earlier attempts include Sweeney et al. (2013), which considered voxel-wise regression for imaging prediction, but ignored the spatial correlations. Hazra et al. (2019) proposed to perform voxel-wise models but to include effects from the neighboring voxels. The proposed framework can be viewed as an image network-on-image network regression, which is sharply distinguished from all the existing works. Another group of related approaches is the graphical regression. Examples include Ni et al. (2019, 2022) and Zhang and Li (2023). In these approaches, regressions were considered for both mean and precision matrix of graphical models. Our proposed approach is distinct from the following aspects. First, our study focuses on the covariance matrix rather than the precision matrix. Though considering eigendecompsition, with the logarithmic link, modeling the 2 is equivalent. However, the interpretation differs. Second, the proposal considers the scenario that a covariance matrix is acquired from each unit rather than to estimate one from units. Third, the current study considers the case that the predictors are a mixture of a covariance matrix and scalars.

The rest of the manuscript is organized as the following. Section 2 introduces the proposed CoCReg model. A least-square type of estimator is introduced to simultaneously estimate linear projections and model coefficients. Asymptotic consistency is discussed under regularity conditions. The performance is evaluated via simulation studies in Supplementary Materials Section B and the task/resting-state fMRI data collected in the HCP Aging Study in Section 3. Section 4 summarizes the manuscript with discussions.

2 MODEL AND METHOD

Assume data are collected from n subjects. Let $\mathbf{x}_{is} \in \mathbb{R}^p$ denote the p-dimensional predictor of the sth observation from subject i, where $s = 1, \ldots, u_i$ and u_i is the number of observations. Let $\mathbf{y}_{it} \in \mathbb{R}^q$ denote the q-dimensional outcome of the tth observation from subject i, where $t = 1, \ldots, v_i$ and v_i is the number of observations. Denote $\mathbf{\Delta}_i \in \mathbb{R}^{p \times p}$ as the covariance matrix of $\mathbf{x}_{is}(s = 1, \ldots, u_i)$ and $\mathbf{\Sigma}_i \in \mathbb{R}^{q \times q}$ as the covariance matrix of $\mathbf{y}_{it}(t = 1, \ldots, v_i)$, for $i = 1, \ldots, n$. The objective is to generalize the concept of regression to study the association between the

2 sets of covariance matrices. In the data application, $\mathbf{x}_{is}$'s are the rs-fMRI signals and $\mathbf{y}_{it}$'s are the tb-fMRI signals. Assuming the signals are centralized to mean zero and standardized to unit variance, the covariance matrices represent the resting-state FC ($\boldsymbol{\Delta}_i$'s) and brain connectivity in task ($\boldsymbol{\Sigma}_i$'s), respectively. The study interest is to investigate if the resting-state FC can predict brain connectivity under an on-going in-scanner task. Let $\mathbf{w}_i \in \mathbb{R}^r$ denote the r-dimensional other covariates of interest (with the first element of one for the intercept term). It is assumed that there exists a linear projection on the $\mathbf{x}_{is}$'s, denoted as $\boldsymbol{\theta} \in \mathbb{R}^p$, and a linear projection on the $\mathbf{y}_{it}$'s, denoted as $\boldsymbol{\gamma} \in \mathbb{R}^q$, such that the following regression model holds:

$$\log\left(\boldsymbol{\gamma}^\top \boldsymbol{\Sigma}_i \boldsymbol{\gamma}\right) = \alpha \log\left(\boldsymbol{\theta}^\top \boldsymbol{\Delta}_i \boldsymbol{\theta}\right) + \mathbf{w}_i^\top \boldsymbol{\beta}, \tag{1}$$

where $\alpha \in \mathbb{R}$ and $\boldsymbol{\beta} \in \mathbb{R}^r$ are model coefficients. Denote $\zeta_{it} = \boldsymbol{\gamma}^\top \mathbf{y}_{it}$. $\boldsymbol{\gamma}$ projects $\mathbf{y}_{it}$ into $\mathbb{R}$ and $\mathrm{Var}(\zeta_{it}) = \boldsymbol{\gamma}^\top \boldsymbol{\Sigma}_i \boldsymbol{\gamma}$. Analogously, denote $\kappa_{is} = \boldsymbol{\theta}^\top \mathbf{x}_{is}$ and $\mathrm{Var}(\kappa_{is}) = \boldsymbol{\theta}^\top \boldsymbol{\Delta}_i \boldsymbol{\theta}$. Model (1) is thus a type of generalized regression model with a logarithmic link on variance components in the projected spaces and to study the association between the 2 variances. α in the model quantifies the association between the 2 variances on the logarithmic scale. $\boldsymbol{\beta}$ measures the association between the outcome data variation in the projection space and the covariates of interest on the original scale. Expanding the quadratic term, $\boldsymbol{\gamma}^\top \boldsymbol{\Sigma}_i \boldsymbol{\gamma} = \sum_j \gamma_j^2 \sigma_{ijj} + \sum_{j \neq k} \gamma_j \gamma_k \sigma_{ijk}$, it represents a linear combination of the correlations if assuming each dimension is scaled to standard deviation one ($\sigma_{ijj} = 1$). In the fMRI application, $\boldsymbol{\gamma}$ defines a brain network where each element in $\boldsymbol{\gamma}$ denotes the weight of the corresponding brain region. $\boldsymbol{\gamma}^\top \boldsymbol{\Sigma}_i \boldsymbol{\gamma}$ can be then considered as a summary of brain FC within the network. $\boldsymbol{\theta}^\top \boldsymbol{\Delta}_i \boldsymbol{\theta}$ can be interpreted in the same manner. Via Model (1), α reveals the association of the connectivity of a task-related network with the connectivity of a resting-state network and $\boldsymbol{\beta}$ uncovers the association with the covariates.

When $\mathbf{w}_i$ contains the intercept term only ($r = 1$), the model can be written as

$$\log\left(\boldsymbol{\gamma}^\top \boldsymbol{\Sigma}_i \boldsymbol{\gamma}\right) = \beta_0 + \alpha \log\left(\boldsymbol{\theta}^\top \boldsymbol{\Delta}_i \boldsymbol{\theta}\right). \tag{2}$$

This can be viewed as a generalization of the canonical correlation analysis (CCA), but to characterize the association between the second-order moments, that is the covariance matrices of $\mathbf{y}$ and $\mathbf{x}$. When $\boldsymbol{\theta}$ is a prespecified projection vector, for example, a subgroup of brain regions with equal weight, plugging in an estimate of $\boldsymbol{\Delta}_i$ (denoted as $\hat{\boldsymbol{\Delta}}_i$) and treating $\log\left(\boldsymbol{\theta}^\top \hat{\boldsymbol{\Delta}}_i \boldsymbol{\theta}\right)$ as a covariate, Model (1) reduces to the covariate assisted principal regression model proposed in Zhao et al. (2021b). However, since distributional assumptions on $\mathbf{y}_{it}$ are not imposed, instead of a likelihood-based estimator introduced by Zhao et al. (2021b), a distribution-free estimator is introduced in the next section. When $\boldsymbol{\gamma}$ is prespecified, Model

(1) should be distinguished from the principal component regression as the study interest is on the association to the variation of the components, that is, $\mathrm{Var}\left(\boldsymbol{\theta}^\top \mathbf{x}_{is}\right)$, rather than the principal components themselves.

2.1 Estimation

In this section, an OLS type of estimator is introduced to relax the distributional assumptions on the data. For resting-state fMRI data, one can assume that the signal ($\mathbf{x}_{is}$) follows a normal distribution with covariance matrix $\boldsymbol{\Delta}_i$. While for task-based fMRI data ($\mathbf{y}_{it}$), this normality assumption does not hold as the signals are a convolution of the task onsite and the hemodynamic response function (HRF), where a widely accepted theoretical distribution for the HRF is a gamma distribution (Lindquist, 2008). Thus, the solution to the following optimization problem is introduced as the estimator without imposing any distributional assumption:

$$\begin{aligned} &\underset{(\boldsymbol{\gamma},\boldsymbol{\theta},\alpha,\boldsymbol{\beta})}{\text{minimize}} \quad \ell = \frac{1}{n}\sum_{i=1}^{n}\left\{\log\left(\boldsymbol{\gamma}^\top \widehat{\boldsymbol{\Sigma}}_i \boldsymbol{\gamma}\right) - \alpha \log\left(\boldsymbol{\theta}^\top \widehat{\boldsymbol{\Delta}}_i \boldsymbol{\theta}\right) - \mathbf{w}_i^\top \boldsymbol{\beta}\right\}^2, \\ &\text{such that } \boldsymbol{\gamma}^\top \mathbf{H}_y \boldsymbol{\gamma} = 1, \quad \boldsymbol{\theta}^\top \mathbf{H}_x \boldsymbol{\theta} = 1. \end{aligned} \tag{3}$$

$\hat{\boldsymbol{\Sigma}}_i$ and $\hat{\boldsymbol{\Delta}}_i$ are estimators of $\boldsymbol{\Sigma}_i$ and $\boldsymbol{\Delta}_i$, respectively. A natural choice is the sample covariance matrices, $\hat{\boldsymbol{\Sigma}}_i = \mathbf{S}_i^y = v_i^{-1}\sum_{t=1}^{v_i}(\mathbf{y}_{it} - \bar{\mathbf{y}}_i)(\mathbf{y}_{it} - \bar{\mathbf{y}}_i)^\top$ and $\hat{\boldsymbol{\Delta}}_i = \mathbf{S}_i^x = u_i^{-1}\sum_{s=1}^{u_i}(\mathbf{x}_{is} - \bar{\mathbf{x}}_i)(\mathbf{x}_{is} - \bar{\mathbf{x}}_i)^\top$ (where $\bar{\mathbf{y}}_i = v_i^{-1}\sum_{t=1}^{v_i}\mathbf{y}_{it}$ and $\bar{\mathbf{x}}_i = u_i^{-1}\sum_{s=1}^{u_i}\mathbf{x}_{is}$), when the data dimensions are not too large and both $\mathbf{S}_i^y$ and $\mathbf{S}_i^x$ are positive definite. It is also discussed in Section 2.3 that to yield consistent estimates, it is required that the data dimensions are fixed and less than the number of corresponding data observations from each individual (Assumption A1). In the optimization problem (3), constraints are imposed on $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$ as the solutions are 0 vectors otherwise, where $\mathbf{H}_y \in \mathbb{R}^{q\times q}$ and $\mathbf{H}_x \in \mathbb{R}^{p\times p}$ are both positive definite. Examples of such matrices include the identity matrix and the overall sample covariance matrices, $\mathbf{H}_y = \bar{\mathbf{S}}_y = \sum_{i=1}^{n} v_i \mathbf{S}_i^y / \sum_{i=1}^{n} v_i$ and $\mathbf{H}_x = \bar{\mathbf{S}}_x = \sum_{i=1}^{n} u_i \mathbf{S}_i^x / \sum_{i=1}^{n} u_i$. For a likelihood-based estimator, $\mathbf{H}_y = \bar{\mathbf{S}}_y$ and $\mathbf{H}_x = \bar{\mathbf{S}}_x$ are considered to incorporate sample information and avoid undesired component estimate (see Zhao et al., 2021b, for a discussion). For the proposed OLS estimator (or moment estimator), identity matrices, that is, $\mathbf{H}_y = \mathbf{I}_q$ and $\mathbf{H}_x = \mathbf{I}_p$, are considered, which relaxes the distributional constraints, analogous to the Principal Component Analysis (PCA) and CCA. In Section B.3 of the Supplementary Materials, we use a simulation study to demonstrate that compared to $\bar{\mathbf{S}}_y$ and $\bar{\mathbf{S}}_x$, the identity matrices are better choices for the proposed approach.

Algorithm 1 summarizes the estimation procedure of optimizing (3). One can show that the objective function in (3) is biconvex over $(\boldsymbol{\gamma}, \boldsymbol{\theta}, \alpha, \boldsymbol{\beta})$. Thus, a coordinate-descent algorithm is considered. For model coefficients α and $\boldsymbol{\beta}$, given the rest parameters, the updates are provided in the following:

$$\hat{\alpha} = \left\{ \frac{1}{n} \sum_{i=1}^{n} \log^2\left(\boldsymbol{\theta}^\top \widehat{\boldsymbol{\Delta}}_i \boldsymbol{\theta}\right) \right\}^{-1} \left[\frac{1}{n} \sum_{i=1}^{n} \left\{ \log\left(\boldsymbol{\gamma}^\top \widehat{\boldsymbol{\Sigma}}_i \boldsymbol{\gamma}\right) - \mathbf{w}_i^\top \boldsymbol{\beta} \right\} \log\left(\boldsymbol{\theta}^\top \widehat{\boldsymbol{\Delta}}_i \boldsymbol{\theta}\right) \right], \tag{4}$$

$$\hat{\boldsymbol{\beta}} = \left(\frac{1}{n} \sum_{i=1}^{n} \mathbf{w}_i \mathbf{w}_i^\top \right)^{-1} \left[\frac{1}{n} \sum_{i=1}^{n} \left\{ \log\left(\boldsymbol{\gamma}^\top \widehat{\boldsymbol{\Sigma}}_i \boldsymbol{\gamma}\right) - \alpha \log\left(\boldsymbol{\theta}^\top \widehat{\boldsymbol{\Delta}}_i \boldsymbol{\theta}\right) \right\} \mathbf{w}_i \right]. \tag{5}$$

For $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$, with quadratic constraints, the method of Lagrange multiplier is employed. Using $\boldsymbol{\gamma}$ as an example, let $U_i = \alpha \log\left(\boldsymbol{\theta}^\top \hat{\boldsymbol{\Delta}}_i \boldsymbol{\theta}\right) + \mathbf{w}_i^\top \boldsymbol{\beta}$, the Lagrangian form is

$$\mathscr{L}(\boldsymbol{\gamma}) = \frac{1}{n} \sum_{i=1}^{n} \left\{ \log\left(\boldsymbol{\gamma}^\top \hat{\boldsymbol{\Sigma}}_i \boldsymbol{\gamma}\right) - U_i \right\}^2 - \lambda_1 \left(\boldsymbol{\gamma}^\top \mathbf{H}_y \boldsymbol{\gamma} - 1\right), \tag{6}$$

where λ_1 is the Lagrange multiplier. Details of optimizing (6), as well as optimizing for $\boldsymbol{\theta}$, are provided in Supplementary Materials Section A.1.

Algorithm 1

An algorithm of solving (3).

Input: $\left\{ \left(\mathbf{y}_{i1}, \ldots, \mathbf{y}_{iv_i}\right), \left(\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iu_i}\right), \mathbf{w}_i \right\}$

1: For $i = 1, \ldots, n$, estimate $\boldsymbol{\Sigma}_i$ and $\boldsymbol{\Delta}_i$, de noted as $\widehat{\boldsymbol{\Sigma}}_i$ and $\widehat{\boldsymbol{\Delta}}_i$ respectively.

2: **Initialization:** $\left(\boldsymbol{\gamma}^{(0)}, \boldsymbol{\theta}^{(0)}, \alpha^{(0)}, \boldsymbol{\beta}^{(0)}\right)$

3: **repeat** for iteration $h = 0, 1, 2, \ldots,$

4: update α as $\alpha^{(h+1)}$ using (4) with $\left(\boldsymbol{\gamma}^{(h)}, \boldsymbol{\theta}^{(h)}, \boldsymbol{\beta}^{(h)}\right)$;

5: update $\boldsymbol{\beta}$ as $\boldsymbol{\beta}^{(h+1)}$ using (5) with $\left(\boldsymbol{\gamma}^{(h)}, \boldsymbol{\theta}^{(h)}, \alpha^{(h+1)}\right)$;

6: update $\boldsymbol{\theta}$ as $\boldsymbol{\theta}^{(h+1)}$ following the approach described in Supplementary Materials Section A.1 with $\left(\boldsymbol{\gamma}^{(h)}, \alpha^{(h+1)}, \boldsymbol{\beta}^{(h+1)}\right)$;

7: update $\boldsymbol{\gamma}$ as $\boldsymbol{\gamma}^{(h+1)}$ following the approach described in Supplementary Materials Section A.1 with $\left(\boldsymbol{\theta}^{(h+1)}, \alpha^{(h+1)}, \boldsymbol{\beta}^{(h+1)}\right)$;

8: **until** the objective function in (3) converges.

9: Consider a random series of initializations, repeat Steps 2–8, and choose the solution with the minimum objective value.

Output: $(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\theta}}, \hat{\alpha}, \hat{\boldsymbol{\beta}})$

2.2 Inference

In this section, we focus on introducing a bootstrap procedure to perform inference on α and $\boldsymbol{\beta}$, the model coefficients. Bootstrap inference on $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$ is beyond the scope of the current study as it may require a procedure of matching on the projections across bootstrap samples. In addition, the performance of this matching procedure highly depends the chosen similarity metric. Thus, a more rigorous and thorough theoretical and numerical investigation is left for future research. The following is a procedure to conduct bootstrap inference on α and $\boldsymbol{\beta}$.

Step 0. Obtain an estimate of $(\boldsymbol{\gamma}, \boldsymbol{\theta})$, denoted as $\left(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\theta}}\right)$, using the full dataset.

Step 1. Generate a bootstrap sample, $\left\{\{\mathbf{y}_{it}\}_t^*, \{\mathbf{x}_{is}\}_s^*, \mathbf{w}_i^*\right\}$, of size n by sampling with replacement.

Step 2. Estimate $(\alpha, \boldsymbol{\beta})$ using Algorithm 1 with $\left(\hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\theta}}\right)$ known.

Step 3. Repeat Steps 1–2 for B times.

Step 4. Construct bootstrap confidence intervals for $(\alpha, \boldsymbol{\beta})$ under a prespecified significance level.

In Step 1, the resampling procedure is conducted at the subject level. All observations within a subject will be used for estimation if being sampled. For datasets with a hierarchically nested structure, resampling on the highest level has been demonstrated to better preserve the original sample information and yield better performance (Ren et al., 2010).

2.3 Asymptotic properties

This section discusses the asymptotic properties of the proposed estimator under regularity conditions. For $i = 1, \ldots, n$, it is assumed that $\boldsymbol{\Sigma}_i$ has the eigendecomposition of $\boldsymbol{\Sigma}_i = \boldsymbol{\Pi}_i \boldsymbol{\Lambda}_i \boldsymbol{\Pi}_i^\top$ and $\boldsymbol{\Delta}_i$ has the eigendecomposition of $\boldsymbol{\Delta}_i = \boldsymbol{\Upsilon}_i \boldsymbol{\Omega}_i \boldsymbol{\Upsilon}_i^\top$, where $\boldsymbol{\Pi}_i = (\boldsymbol{\pi}_{i1}, \ldots, \boldsymbol{\pi}_{iq}) \in \mathbb{R}^{q \times q}$ and $\boldsymbol{\Upsilon}_i = (\boldsymbol{\upsilon}_{i1}, \ldots, \boldsymbol{\upsilon}_{ip}) \in \mathbb{R}^{p \times p}$ are orthonormal eigenmatrices, $\boldsymbol{\Lambda}_i = \mathrm{diag}\{\lambda_{i1}, \ldots, \lambda_{iq}\} \in \mathbb{R}^{q \times q}$ and $\boldsymbol{\Omega}_i = \mathrm{diag}\{\omega_{i1}, \ldots, \omega_{ip}\} \in \mathbb{R}^{p \times p}$ are diagonal matrices of corresponding eigenvalues. Let $\boldsymbol{\zeta}_{it} = \boldsymbol{\Pi}_i^\top \mathbf{y}_{it} = (\boldsymbol{\zeta}_{itk})_k \in \mathbb{R}^q$ and $\boldsymbol{\kappa}_{is} = \boldsymbol{\Upsilon}_i^\top \mathbf{x}_{is} = (\kappa_{isj})_j \in \mathbb{R}^p$. Then, $\mathrm{Cov}(\boldsymbol{\zeta}_{it}) = \boldsymbol{\Lambda}_i$ and $\mathrm{Cov}(\boldsymbol{\kappa}_{is}) = \boldsymbol{\Omega}_i$. The elements in ζ_{it} are uncorrelated and so as the elements in $\boldsymbol{\kappa}_{\mathrm{is}}$. The following assumptions are imposed:

Assumption A1. Let $u = \min_i u_i$ and $v = \min_i v_i$. $p \ll u$ and $q \ll v$ are fixed.

Assumption A2. There exist constants C_1 independent of u and C_2 independent of v, such that for $\forall j = 1, \ldots, p$, $\mathbb{E}\left(\kappa_{i1j}^4\right) \leq C_1$, and for $\forall k = 1, \ldots, q$, $\mathbb{E}\left(\zeta_{i1k}^4\right) \leq C_2$, for $\forall i = 1, \ldots, n$.

Assumption A3 $\boldsymbol{\Sigma}_i$'s share the same set of eigenvectors, i.e., $\boldsymbol{\Pi}_i = \boldsymbol{\Pi}$, for $i = 1, \ldots, n$; and $\boldsymbol{\Delta}_i$'s share the same set of eigenvectors, i.e., $\boldsymbol{\Upsilon}_i = \boldsymbol{\Upsilon}$, for $i = 1, \ldots, n$.

Assumption A4. For $\forall i = 1, \ldots, n$, there exists (at least) a pair of columns in $\boldsymbol{\Pi}_i$ and $\boldsymbol{\Upsilon}_i$, indexed by k_i and j_i), respectively, such that $\boldsymbol{\gamma} = \boldsymbol{\pi}_{ik_i}$, $\boldsymbol{\theta} = \boldsymbol{\upsilon}_{ij_i}$, and Model (1) is satisfied.

Assumption A1 assumes a low-dimensional scenario for both **x** and **y**. Under this assumption, the sample covariance matrices are of full rank and are consistent estimators of the covariance matrices. Though distributional assumptions are not imposed, Assumption A2 regulates the higher-order moments of $\boldsymbol{\kappa}$ and $\boldsymbol{\zeta}$, which is equivalent to regulating the higher-order moments of **x** and **y**, respectively. Assumption A3 assumes that $\boldsymbol{\Sigma}_i$'s share the same set of eigenvectors and so as $\boldsymbol{\Delta}_i$'s. In the fMRI data application, this assumption assumes that human brain can be divided into subnetworks and this division is the same across individuals. This type of assumption is also imposed for brain parcellation and the identification of functional modulation (see an example in Power et al., 2011). Assumption A4 assumes the loglinear model to be correctly specified. With Assumptions A1–A4 satisfied, one can consider setting the eigenvectors of $\overline{\mathbf{S}}_y$ and $\overline{\mathbf{S}}_x$ as the initial value of $\boldsymbol{\gamma}$ and $\boldsymbol{\theta}$, respectively, in Algorithm 1. The following proposition suggests the consistency of the proposed estimator under such initialization. To achieve consistency, it is required that the number of subjects (n), as well as the number of observations in **x** (denoted as u) and **y** (denoted as v) from each subject, increase. In fMRI experiments, longer scans increase the value of u and v.

Proposition 1: *Assume Assumptions A1–A4 hold. As* $n, u, v \to \infty$, *the proposed estimator of* $(\boldsymbol{\gamma}, \boldsymbol{\theta}, \alpha, \boldsymbol{\beta})$ *is consistent.*

2.4 Higher-order components and choosing the number of components

Section 2.1 introduces an algorithm to identify the first component. This section introduces an approach to identify higher-order components. Here, the projections are assumed to be one-to-one, that is, 1 linear projection on **y** corresponds to 1 linear projection on **x**. Assume $\hat{\boldsymbol{\Gamma}}^{(k-1)} = (\hat{\boldsymbol{\gamma}}_1, \ldots, \hat{\boldsymbol{\gamma}}_{(k-1)}) \in \mathbb{R}^{q \times (k-1)}$ and $\hat{\boldsymbol{\Theta}}^{(k-1)} = \left(\hat{\boldsymbol{\theta}}_1, \ldots, \hat{\boldsymbol{\theta}}_{(k-1)}\right) \in \mathbb{R}^{p \times (k-1)}$ are the first $(k-1)$ identified component pairs, for $k = 2, \ldots, \min(p, q)$. It is proposed to remove these identified components from the data to identify the next pair. For $i = 1, \ldots, n$, let

$$\begin{aligned} \widehat{\mathbf{Y}}_i^{(k)} &= \mathbf{Y}_i - \mathbf{Y}_i \widehat{\boldsymbol{\Gamma}}^{(k-1)} \widehat{\boldsymbol{\Gamma}}^{(k-1)\top} \quad \text{and} \quad \widehat{\mathbf{X}}_i^{(k)} \\ &= \mathbf{X}_i - \mathbf{X}_i \widehat{\boldsymbol{\Theta}}^{(k-1)} \widehat{\boldsymbol{\Theta}}^{(k-1)\top}, \end{aligned} \tag{7}$$

where $\mathbf{Y}_i = (\mathbf{y}_{i1}, \ldots, \mathbf{y}_{iv_i})^\top \in \mathbb{R}^{v_i \times q}$ and $\mathbf{X}_i = (\mathbf{x}_{i1}, \ldots, \mathbf{x}_{iu_i})^\top \in \mathbb{R}^{u_i \times p}$ denote the original data. Consider $\left\{\hat{\mathbf{Y}}_i^{(k)}, \hat{\mathbf{X}}_i^{(k)}, \mathbf{w}_i\right\}$ as the new data and apply Algorithm 1 to identify the kth component. It should be noted that by removing the identified components using (7), the effect from the covariates associated with these components is removed simultaneously.

To choose the number of components, Zhao et al. (2021b) introduced a criterion named average deviation from diagonality (DfD). It was introduced based on the nature that there exists a simultaneous diagonalization of the covariance matrices. Extending to the CoCReg introduced in this manuscript, the following metric is considered. Let $\hat{\boldsymbol{\Gamma}}^{(k)} \in \mathbb{R}^{q \times k}$ and $\hat{\boldsymbol{\Theta}}^{(k)} \in \mathbb{R}^{p \times k}$ be the estimated first k components, define

$$\mathrm{DfD}(k) = \max\left\{\mathrm{DfD}\left(\widehat{\mathbf{\Gamma}}^{(k)}\right), \mathrm{DfD}\left(\widehat{\mathbf{\Theta}}^{(k)}\right)\right\}, \tag{8}$$

where

$$\mathrm{DfD}\left(\widehat{\mathbf{\Gamma}}^{(k)}\right) = \prod_{i=1}^{n} v\left(\widehat{\mathbf{\Gamma}}^{(k)\top}\widehat{\mathbf{\Sigma}}_i\widehat{\mathbf{\Gamma}}^{(k)}\right)^{v_i/\sum_i v_i}, \text{ and } \mathrm{DfD}\left(\widehat{\mathbf{\Theta}}^{(k)}\right) = \prod_{i=1}^{n} v\left(\widehat{\mathbf{\Theta}}^{(k)\top}\widehat{\mathbf{\Delta}}_i\widehat{\mathbf{\Theta}}^{(k)}\right)^{u_i/\sum_i u_i}. \tag{9}$$

For a square matrix $\mathbf{A}$, $v(\mathbf{A}) = \det\{\mathrm{diag}(\mathbf{A})\}/\det(\mathbf{A})$, where $\mathrm{diag}(\mathbf{A})$ is a diagonal matrix with the same diagonal elements as in $\mathbf{A}$ and $\det(\mathbf{A})$ is the determinant. $v(\mathbf{A}) \geq 1$ and the equality holds if and only if $\mathbf{A}$ is a diagonal matrix. Considering a threshold ϱ, for example, $\varrho = 2$ as recommended in Zhao et al. (2021b), the number of components is chosen as

$$\hat{k} = \max\{k : \mathrm{DfD}(k) \leq \varrho\}. \tag{10}$$

3 THE LIFESPAN HUMAN CONNECTOME PROJECT AGING STUDY

We apply the proposed approach to the Lifespan HCP Aging Study. The goal is to utilize the technological advances developed by the HCP study of healthy young adults to explore typical aging trajectory and how connectome changes in the brain among mature and older adults. In this study, it is considered to use resting-state FC to predict FC during a FACENAME task, a paired-associates memory task adapted from Zeineh et al. (2003). During the task-based scanning, participants were instructed to encode the names of different face identities in some blocks and mentally recall them in other blocks. The task was designed as a localizer for encoding and retrieval related neural processes. A total of $n = 551$ subjects aged between 36 and 90 (245 Male and 306 Female) having both the resting-state and task-based fMRI available without any quality control issue are included. The fMRI data were minimally preprocessed (Glasser et al., 2013). Signals are extracted from $p = q = 75$ brain regions, including 60 cortical and 15 subcortical regions grouped into 10 functional modules, using the Harvard-Oxford Atlas in FSL (Smith et al., 2004). Both resting-state and task-based signals are motion corrected. In the resting-state data ($\mathbf{x}_{is}$), there are $u_i = u = 478$ observations of each subject and in the task-based data ($\mathbf{y}_{it}$), the number is $v_i = v = 335$. Other covariates include gender at birth and age ($\mathbf{w}_i$) with $r = 3$ including the intercept. The validity of model assumptions is examined in Supplementary Materials Section C.1.

Using the DfD criterion, the proposed approach identifies 3 orthogonal components, denoted as C1, C2, and C3. Table 1 presents the estimated model coefficients and 95% bootstrap

confidence intervals from 500 samples. In all 3 components, FC within the identified task-based network is significantly associated with FC within the identified resting-state network, positive in C1 and C2 and negative in C3. To better interpret the components, γ and θ are sparsified following an *ad hoc* procedure using a fused lasso regression (Tibshirani et al., 2005), where local smoothness and constancy are imposed within each brain functional module (Grosenick et al., 2013). Figure C.1 in Supplementary Materials Section C.3 presents the sparsified loading profile colored by functional modules and Figure 1 presents the nonzero loaded regions in a brain map and the river plot of the loading configuration by modules.

The resting-state network of C1 is primarily the default mode network (DMN), which is best known for being active during wakeful rest and consistently captured in rs-fMRI studies (Shen, 2015). The related task-state network is primarily cortical areas of the brain and the cerebellum with similar positive loading values suggesting a global signal pattern during the task. Emerging evidence suggests that there is a potential extrinsic mode network (EMN) during task that can be predicted by DMN at the resting state (Hugdahl et al., 2015). Our finding confirms this with the identified task-state network signatures contributing to the establishment of EMN. The 2 networks in C2 yields a high similarity of 0.663 and the module configuration is consistent in the DMN, fronto-parietal network, limbic system, and visual network. This high similarity between the 2 networks is in line with the emerging theory of functional connectome fingerprinting (Finn et al., 2015). Leading consistent areas include the anterior cingulate cortex (ACC), the rostral middle frontal cortex, the medial orbitofrontal cortex, the temporal pole, and the ventromedial prefrontal cortex with a positive loading and the inferior parietal lobule, the lateral occipital cortex, the cuneus, the lingual gyrus, and the fusiform gyrus with a negative loading. Regions in the prefrontal cortex and frontoparietal network are well known to be involved in working memory functions (Barch et al., 2013; Funahashi, 2006). Hypothetically, FC between these regions is modulated by working memory loads during tasks and potentially a third brain region at rest. A recent study found that the ACC showed different modulatory interactions with regions in the network both in the resting-state and working memory tasks (Di et al., 2020). The cuneus, the lingual gyrus, and the fusiform gyrus (also known as the lateral occipitotemporal gyrus) are in the visual network involved in visual processing, visual memory, and face recognition. Coactivation in these regions has been observed in the processing of visual information and working memory tasks (Bogousslavsky et al., 1987; McCarthy et al., 1999; Palejwala et al., 2021; Sellal, 2022). C3 is a component with a low similarity between the 2 networks (similarity = −0.095). The rs-network consists of the DMN, including the middle temporal gyrus, and the superior frontal gyrus (Left), and the caudal middle frontal gyrus, and subcortical regions, including the parahippocampal (Left) and the caudate. The tb-network mainly consists of the visual network, the somato-motor network, the frontal-parietal network, and the limbic system. Regions include the lateral occipital cortex, the fusiform, the precentral and postcentral gyrus, the superior temporal gyrus, the posterior insula, the temporal gyrus and pole, the inferior temporal gyrus, the medial orbitofrontal cortex, and the rostral middle frontal cortex. Primarily consisting of the visual and motor networks, the tb-network of C3 is a sensory binding and motor control component potentially activated to the distraction period between task blocks (Dörfel et al.,

2014; Li et al., 2021). In summary, the proposed approach identifies 3 orthogonal associated resting-state and task-state brain networks, one related to global signaling (C1), one related to the working memory task (C2), and one unrelated to the task (C3).

4 DISCUSSION

In this study, a CoCReg model is introduced. It is assumed that there exists a pair of linear projections on the outcome covariance matrices and the predictor covariance matrices, such that in the projection spaces, a log-linear model is satisfied to associate the variances. An OLSs type of estimator is introduced for simultaneous projection identification and model coefficient estimation. Under regularity conditions, the proposed estimator is asymptotically consistent. Simulation studies demonstrate the superior performance of the proposed approach over a modified existing method. Applying to data collected in the HCP Aging study, the proposed approach identifies 3 pairs of networks, where FC within the resting-state network can predict FC within the corresponding task-state network. The 3 networks correspond to a global signal network, a task-related network, and a task-unrelated network. The findings are consistent with existing knowledge about brain function and activation at rest and during a memory task.

The asymptotic consistency of the proposed estimator is achieved under the assumption of complete common diagonalization for the outcome covariance matrices and the predictor covariance matrices. In Zhao et al. (2021b), a complete common diagonalization was also assumed. Via simulation studies, it is demonstrated that this assumption can be relaxed to partial common diagonalization. In Section B of Supplementary Materials, a partial common diagonalization scenario is also considered and the numerical result suggests that the proposed approach is robust to this relaxation. Thus, a theoretical study of the asymptotic consistency under this relaxation is one future direction. The proposed framework considers the low-dimensional scenario, where the dimensions of the outcome covariance matrix and the predictor covariance matrix are both lower than the number of observations acquired from the subject. Asymptotic properties are investigated under this scenario. One extension is to consider the case of high-dimensional data. In Zhao et al. (2021a), a shrinkage estimator of the outcome covariance matrices was introduced, where the shrinkage parameters were assumed to be shared across subjects to yield optimal performance. Shrinkage estimators of the outcome covariance matrices and the predictor covariance matrices can be then considered analogously and we leave it as our future research. For covariance matrices in a high dimension, one may consider imposing sparsity on the projection vectors, which serves better interpretation. The current study considers an *ad hoc* procedure of sparsifying the projection vectors following the strategy of sparse PCA (Zou et al., 2006). Another future direction is to directly impose sparsity in the estimation procedure. A bootstrap procedure is introduced for inference on the model coefficients, not on the linear projections. As for each bootstrap sample, the identified projections and the order of identifying these projections may differ, performing inference on these projection vectors requires a matching procedure. The inference will then highly depend on the performance of matching and the metric used for matching. Thus, a more thorough theoretical and numerical investigation is necessary. As this is beyond the scope of the current manuscript, it will be studied in the future.

One objective of fitting a regression model is to make prediction. For a new subject with acquired rs-fMRI data, the proposed approach provides a low-rank representation of the task-state connectivity matrix. This low-rank representation consists of latent components that are related to resting-state brain connectivity and covariates considered in the model. A prediction of the full-rank task-state connectivity matrix or even tb-fMRI signals is beyond the capability of the current framework.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

FUNDING

Yi Zhao was supported in part by NIH grants R01MH126970, P30AG072976, and U54AG065181. Yize Zhao was supported in part by National Institutes of Health grants R01EB034720, RF1AG068191, and RF1AG081413.

DATA AVAILABILITY

Data used in this paper were downloaded from the Human Connectome Project Aging Study and are publicly available in the National Institute of Mental Health (NIMH) Data Archive, at https://nda.nih.gov/.

REFERENCES

Acharyya S, Zhang Z, Bhattacharya A and Pati D (2023). Bayesian hierarchical modeling on covariance valued data. Stat, 12, e534.

Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M et al. (2013). Function in the human connectome: taskfMRI and individual differences in behavior. Neuroimage, 80, 169–189. [PubMed: 23684877]

Bogousslavsky J, Miklossy J, Deruaz J-P, Assal G and Regli F (1987). Lingual and fusiform gyri in visual processing: a clinico-pathologic study of superior altitudinal hemianopia. Journal of Neurology, Neurosurgery and Psychiatry, 50, 607–614. [PubMed: 3585386]

Boik RJ. (2002). Spectral models for covariance matrices. Biometrika, 89, 159–182.

Cole MW, Ito T, Cocuzza C and Sanchez-Romero R (2021). The functional relevance of task-state functional connectivity. Journal of Neuroscience, 41, 2684–2702. [PubMed: 33542083]

Di X, Zhang H and Biswal BB (2020). Anterior cingulate cortex differently modulates frontoparietal functional connectivity between resting-state and working memory tasks. Human Brain Mapping, 41, 1797–1805. [PubMed: 31904907]

Dörfel D, Lamke J-P, Hummel F, Wagner U, Erk S and Walter H (2014). Common and differential neural networks of emotion regulation by detachment, reinterpretation, distraction, and expressive suppression: a comparative fMRI investigation. NeuroImage, 101, 298–309. [PubMed: 24993897]

Finn ES, Shen X, Scheinost D, Rosenberg MD, Huang J, Chun MM et al. (2015). Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. Nature Neuroscience, 18, 1664–1671. [PubMed: 26457551]

Flury BN. (1984). Common principal components in k groups. Journal of the American Statistical Association, 79, 892–898.

Fox EB and Dunson DB (2015). Bayesian nonparametric covariance regression. Journal of Machine Learning Research, 16, 2501–2542.

Franks AM and Hoff P (2019). Shared subspace models for multi-group covariance estimation. Journal of Machine Learning Research, 20, 1–37.

Funahashi S. (2006). Prefrontal cortex and working memory processes. Neuroscience, 139, 251–261. [PubMed: 16325345]

Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL et al. (2013). The minimal preprocessing pipelines for the Human Connectome Project. NeuroImage, 80, 105–124. [PubMed: 23668970]

Grosenick L, Klingenberg B, Katovich K, Knutson B and Taylor JE (2013). Interpretable whole-brain prediction analysis with GraphNet. NeuroImage, 72, 304–321. [PubMed: 23298747]

Guo C, Kang J and Johnson TD (2022). A spatial Bayesian latent factor model for image-on-image regression. Biometrics, 78, 72–84. [PubMed: 33368210]

Hazra A, Reich BJ, Reich DS, Shinohara RT and Staicu A-M (2019). A spatio-temporal model for longitudinal image-on-image regression. Statistics in Biosciences, 11, 22–46. [PubMed: 31156722]

Hoff PD. (2009). A hierarchical eigenmodel for pooled covariance estimation. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 71, 971–992.

Hoff PD and Niu X (2012). A covariance regression model. Statistica Sinica, 22, 729–753.

Hugdahl K, Raichle ME, Mitra A and Specht K (2015). On the existence of a generalized non-specific task-dependent network. Frontiers in Human Neuroscience, 9, 430. [PubMed: 26300757]

Li W, Yang P, Ngetich RK, Zhang J, Jin Z and Li L (2021). Differential involvement of frontoparietal network and insula cortex in emotion regulation. Neuropsychologia, 161, 107991. [PubMed: 34391808]

Lindquist MA. (2008). The statistical analysis of fMRI data. Statistical Science, 23, 439–464.

McCarthy G, Puce A, Belger A and Allison T (1999). Electrophysiological studies of human face perception II: Response properties of face-specific potentials generated in occipitotemporal cortex. Cerebral Cortex, 9, 431–444. [PubMed: 10450889]

Ngo GH, Khosla M, Jamison K, Kuceyeski A and Sabuncu MR (2022). Predicting individual task contrasts from resting-state functional connectivity using a surface-based convolutional network. NeuroImage, 248, 118849. [PubMed: 34965456]

Ni Y, Stingo FC and Baladandayuthapani V (2019). Bayesian graphical regression. Journal of the American Statistical Association, 114, 184–197. [PubMed: 36937091]

Ni Y, Stingo FC and Baladandayuthapani V (2022). Bayesian covariate-dependent Gaussian graphical models with varying structure. The Journal of Machine Learning Research, 23, 11049–11077.

Palejwala AH, Dadario NB, Young IM, O'Connor K, Briggs RG, Conner AK et al. (2021). Anatomy and white matter connections of the lingual gyrus and cuneus. World Neurosurgery, 151, e426–e437. [PubMed: 33894399]

Pourahmadi M, Daniels MJ and Park T (2007). Simultaneous modelling of the Cholesky decomposition of several covariance matrices. Journal of Multivariate Analysis, 98, 568–587.

Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA et al. (2011). Functional network organization of the human brain. Neuron, 72, 665–678. [PubMed: 22099467]

Ren S, Lai H, Tong W, Aminzadeh M, Hou X and Lai S (2010). Nonparametric bootstrapping for hierarchical data. Journal of Applied Statistics, 37, 1487–1498.

Seiler C and Holmes S (2017). Multivariate heteroscedasticity models for functional brain connectivity. Frontiers in Neuroscience, 11, 696. [PubMed: 29311777]

Sellal F. (2022). Anatomical and neurophysiological basis of face recognition. Revista de Neurología (Paris), 178, 649–653.

Shen HH. (2015). Resting-state connectivity. Proceedings of the National Academy of Sciences, 112, 14115–14116.

Smith SM, Jenkinson M, Woolrich MW, Beckmann CF, Behrens TE, Johansen-Berg H et al. (2004). Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage, 23, S208–S219. [PubMed: 15501092]

Sweeney E, Shinohara R, Shea C, Reich D and Crainiceanu CM (2013). Automatic lesion incidence estimation and detection in multiple sclerosis using multisequence longitudinal MRI. American Journal of Neuroradiology, 34, 68–73. [PubMed: 22766673]

Tavor I, Jones OP, Mars R, Smith S, Behrens T and Jbabdi S (2016). Task-free MRI predicts individual differences in brain activity during task performance. Science, 352, 216–220. [PubMed: 27124457]

Tibshirani R, Saunders M, Rosset S, Zhu J and Knight K (2005). Sparsity and smoothness via the fused lasso. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 67, 91–108.

Zeineh MM, Engel SA, Thompson PM and Bookheimer SY (2003). Dynamics of the hippocampus during encoding and retrieval of face-name pairs. Science, 299, 577–580. [PubMed: 12543980]

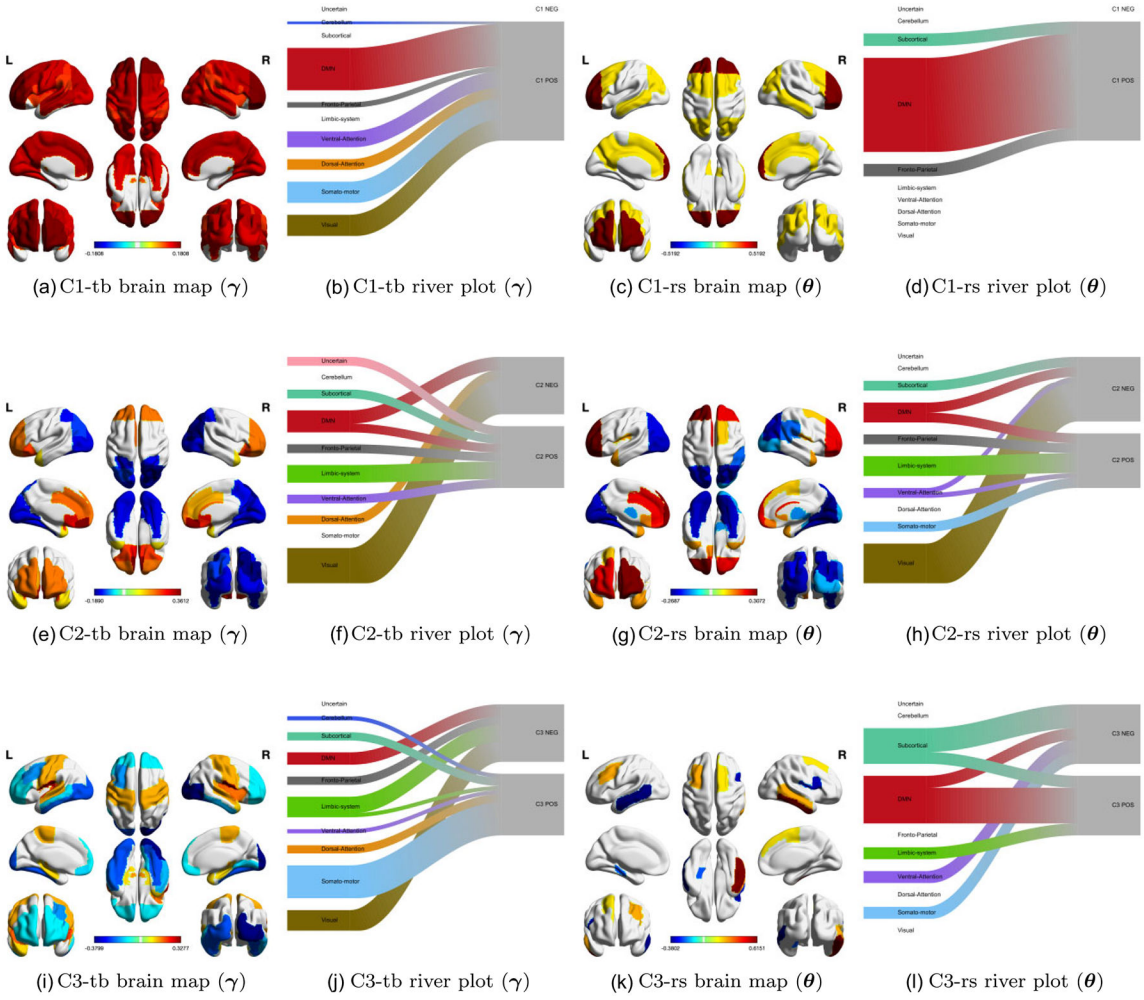
Zhang J and Li Y (2023). High-dimensional Gaussian graphical regression models with covariates. Journal of the American Statistical Association, 118, 2088–2100. [PubMed: 38143787]

Zhao Y, Caffo BS and Luo X (2021a). Principal regression for high dimensional covariance matrices. Electronic Journal of Statistics, 15, 4192–4235. [PubMed: 35782590]

Zhao Y, Wang B, Mostofsky SH, Caffo BS and Luo X (2021b). Covariate assisted principal regression for covariance matrix outcomes. Biostatistics, 22, 629–645. [PubMed: 31851318]

Zou H, Hastie T and Tibshirani R (2006). Sparse principal component analysis. Journal of Computational and Graphical Statistics, 15, 265–286.

Zou T, Lan W, Wang H and Tsai C-L (2017). Covariance regression analysis. Journal of the American Statistical Association, 112, 266–281.



(a) C1-tb brain map (γ) (b) C1-tb river plot (γ) (c) C1-rs brain map ($\boldsymbol{\theta}$) (d) C1-rs river plot ($\boldsymbol{\theta}$)

(e) C2-tb brain map (γ) (f) C2-tb river plot (γ) (g) C2-rs brain map ($\boldsymbol{\theta}$) (h) C2-rs river plot ($\boldsymbol{\theta}$)

(i) C3-tb brain map (γ) (j) C3-tb river plot (γ) (k) C3-rs brain map ($\boldsymbol{\theta}$) (l) C3-rs river plot ($\boldsymbol{\theta}$)

FIGURE 1.

Regions with a nonzero loading in a brain map and the river plot of module configuration of the 3 identified components (C1, C2, and C3), using the proposed CoCReg approach in the HCP Aging study. tb, results of task-based fMRI; rs, results of resting-state fMRI.

TABLE 1

Estimated model coefficients, 95% confidence intervals from 500 bootstrap samples, and the similarity between $\hat{\boldsymbol{\gamma}}$ and $\hat{\boldsymbol{\theta}}$ (denoted as $\langle \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\theta}} \rangle$) of the 3 identified components (C1, C2, and C3) using the proposed CoCReg approach in the HCP Aging study.

	rs-fMRI		Male-Female		Age		$\langle \hat{\gamma}, \hat{\theta} \rangle$
	Estimate	**95% CI**	**Estimate**	**95% CI**	**Estimate**	**95% CI**	
C1	0.669	(−0.572, −0.766)	0.028	(−0.009, −0.065)	0.002	(−0.017, −0.020)	0.490
C2	0.217	(−0.142, −0.290)	−0.067	(−0.118, −0.015)	0.043	(−0.018, −0.069)	0.663
C3	−0.301	(−0.407, −0.195)	0.002	(−0.052, −0.056)	−0.053	(−0.080, −0.027)	−0.095