Published in final edited form as:

Comput Stat Data Anal. 2020 February ; 142: . doi:10.1016/j.csda.2019.106835.

Sparse Principal Component based High-Dimensional Mediation Analysis

Yi Zhao^{a,*}, Martin A. Lindquist^a, Brian S. Caffo^a

^aDepartment of Biostatistics, Johns Hopkins Bloomberg School of Public Health

Abstract

Causal mediation analysis aims to quantify the intermediate effect of a mediator on the causal pathway from treatment to outcome. When dealing with multiple mediators, which are potentially causally dependent, the possible decomposition of pathway effects grows exponentially with the number of mediators. An existing approach incorporated the principal component analysis (PCA) to address this challenge based on the fact that the transformed mediators are conditionally independent given the orthogonality of the principal components (PCs). However, the transformed mediator PCs, which are linear combinations of original mediators, can be difficult to interpret. A sparse high-dimensional mediation analysis approach is proposed which adopts the sparse PCA method to the mediation setting. The proposed approach is applied to a task-based functional magnetic resonance imaging study, illustrating its ability to detect biologically meaningful results related to an identified mediator.

Keywords

Functional magnetic resonance imaging; Mediation analysis; Regularized regression; Structural equation model

2010 MSC:

00-01; 99-00

1. Introduction

Causal mediation analysis has been widely applied in social, psychological, and biological studies to evaluate the intermediate effect of a variable (called a mediator) on the causal pathway from an exposure/treatment to a target outcome for the purpose of exploring hypothesized causal mechanisms. The single mediator setting has been extensively studied [1, 2, 3, 4, 5, 6, 7, 8, 9, 10]. During the past decade, methods for dealing with multiple mediators have attracted increasing attention [11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22,

*Correspondence to: 615 N Wolfe St, Baltimore MD, 21205, USA, yzhao99@jhmi.edu (Yi Zhao).

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

23, 24, 25, 26]. Most of these methods are designed for dealing with relatively low-dimensional data. With the emergence of modern technologies (e.g., high-throughput technologies in omics studies and neuroimaging technologies), data sets with a large number of variables are being collected. However, methods for conducting mediation analysis with high-dimensional mediators are limited. Motivated by a genetics study, Huang and Pan (2016) [27] proposed a principal component analysis (PCA) based approach for reducing the high-dimensional gene expression mediators to a series of marginal mediation problems. Incorporating a regularized regression for the outcome model, Zhang et al. (2016) [28] introduced an independent screening approach for high-dimensional mediation analysis.

In the field of neuroimaging, studies on the impact of brain mediators on cognitive behavior are becoming increasingly popular [29, 30, 31, 32, 33, 34, 35]. Caffo et al. (2007) [29] presented an early attempt at addressing neurological images as mediators, though the analysis was conducted on univariate summaries extracted from the multivariate images. Recently, Chén et al. (2017) [36] proposed a mediation analysis approach that transforms the high-dimensional mediator candidates into independent directions of mediation (DMs). Geuter et al. [37] extended this approach by transforming components based on the amount of the indirect (mediating) effect they explained. Under a linear structural equation modeling (LSEM) framework, the directions are ranked based on the proportion of the likelihood that they account for. Zhao and Luo (2016) [38] recently proposed a general mediation model under the LSEM framework to account for the causal dependencies between the mediators and introduced a lasso-type penalty to regularize the mediation pathway effects to achieve simultaneous mediator selection and mediation effect estimation.

In both the PCA and the directions of mediation analysis, the acquired PCs and DMs are linear combinations of the original mediating variables. With nonzero loadings, their interpretation is not always straightforward. An informal way of reducing the number of variables is to set a hard threshold, and force loadings with absolute value below the threshold to be zero. However, this can be potentially misleading, as it is not only the loadings but also the variance of the variable that governs its importance [39]. Jolliffe et al. (2003) [40] introduced a modified PCA approach based on the lasso [41]. Built on the fact that sparsifying the PC loadings can be expressed as a regression-type optimization problem, Zou et al. (2006) [42] introduced the sparse principal component analysis (SPCA) approach. The same idea was later implemented in canonical correlation analysis (CCA) [43, 44, 45].

In this study, we propose a sparse principal component of mediation (SPCM) approach to perform high-dimensional mediation analysis. This approach has two stages: the first performs a PCA for high-dimensional mediation using the method proposed in Huang and Pan (2016) [27]; and the second sparsifies the loading vector using a (structured) regularization. In neuroimaging studies, spatial smoothing regularization is generally imposed to enforce spatial smoothness and yield meaningful biological interpretations [46, 47]. In this study, the fused lasso [48], as a special case of the generalized lasso [49], will be employed to impose local smoothness and constancy. We apply the proposed approach to a task-based functional magnetic resonance imaging (fMRI) study, where participants were instructed to perform a probabilistic classification learning task. Our method enables the

identification of biologically interpretable brain pathways that are responsive to the task stimulus and also have an impact on the reaction time to the task.

This paper is organized as follows. Section 2 introduces the PCA based mediation approach for multiple mediators proposed in Huang and Pan (2016) [27]. In Section 3, we present the sparse principal component of mediation approach. Section 4 summarizes the simulation results. We apply our proposed method to a task-based fMRI study in Section 5. Section 6 summarizes this paper with discussions.

2. Causal Mediation Analysis with Multiple Mediators

Mediation analysis aims to quantify the causal effect of a treatment/exposure (X) on the outcome (Y) mediated by a third variable, called the mediator (M). This causal relationship can be represented using a causal diagram as in Figure 1a. Linear structural equation modeling (LSEM) is a popular approach for performing mediation analysis. Let $M(x)$ and $Y(x, M(x))$ denote the potential outcome of the mediator and the outcome under treatment assignment x [50, 51], the mediation models are written as

$$M(x) = \alpha_0 + \alpha x + \epsilon, \quad \text{and} \quad (1)$$

$$Y(x, M(x)) = \beta_0 + \gamma x + \beta M(x) + \eta, \quad (2)$$

where ϵ and η are independent model errors with mean zero. The average total treatment effect is decomposed as

$$\begin{aligned} \text{ATE}(x, x^*) &= \mathbb{E}[Y(x, M(x))] - \mathbb{E}[Y(x^*, M(x^*))] \\ &= \mathbb{E}[Y(x, M(x)) - Y(x, M(x^*))] + \mathbb{E}[Y(x, M(x^*)) - Y(x^*, M(x^*))] \\ &= \text{AIE}(x, x^*) + \text{ADE}(x, x^*), \end{aligned} \quad (3)$$

where x and x^* are two distinct treatment assignments. Under models (1) and (2), $\text{ADE}(x, x^*) = \mathbb{E}[Y(x, M(x^*)) - Y(x^*, M(x^*))] = \gamma(x - x^*)$ is the average (controlled) direct effect of the treatment on the outcome, and $\text{AIE}(x, x^*) = \mathbb{E}[Y(x, M(x)) - Y(x, M(x^*))] = \alpha\beta(x - x^*)$ is the average indirect effect. Under assumptions (see Section A of the supplementary material), these causal estimands can be identified from the observed data [9, 10].

With multiple mediators, a challenge is to delineate the causal structure among the mediators. When the mediators maintain an ordered structure, one can directly penalize each causal connection in a directed acyclic graph [52]. In certain applications, for example the process of inhibitory motor control [53], causality between brain activation has been established via transcranial magnetic stimulation. This type of causal relationship between brain regions is referred to as the effective connectivity, which captures the influence that one neural system exerts over another [54, 55]. This presumes an ordering of neural systems. Therefore, when considering brain activity from different brain regions as mediators, the assumption that there are p -ordered mediators is consistent with a working model of

connectivity. However, functional magnetic resonance imaging (fMRI) is not sufficiently informative to determine the causal ordering of the brain regions, given its low temporal resolution and high noise. To avoid sequentially modeling multiple mediators, Huang and Pan (2016) [27] introduced the concept of applying principal components to the mediators, which can be used to linearly combine the candidate mediators. Imposing orthogonality, the mediation principal components are conditionally independent given the treatment. The complex causal structure in Figure 1b can be transformed into a problem with parallel causal mediation pathways, as shown in Figure 1c. As discussed in Imai et al. (2013) [13] and VanderWeele (2015) [10], with causally independent multiple mediators, it is equivalent to performing a series of marginal mediation analyses.

Let X denote the treatment assignment, $\mathbf{M}(x) = (M_1(x), \dots, M_p(x))^T \in \mathbb{R}^p$ the p -dimensional potential outcome of mediator given the treatment assignment at level x . Let $\widetilde{\mathbf{M}}^{(j)}(x) = \mathbf{M}(x)^T \phi_j (j = 1, \dots, p)$ be a linear projection of the potential outcome $\mathbf{M}(x)$, such that

$$\widetilde{\mathbf{M}}^{(j)}(x) \perp \widetilde{\mathbf{M}}^{(k)}(x) \mid X = x, \quad \text{for } j \neq k; \quad (4)$$

that is, the mediators in the projection space are causally independent under the definition in Imai and Yamamoto (2013) [13]. In this setting, the problem is equivalent to conducting a series of marginal mediation analyses. For subject $i (i = 1, \dots, n)$, under the LSEM framework, for each $j = 1, \dots, p$,

$$\widetilde{M}_i^{(j)} = \alpha_{0j} + \alpha_j X_i + \xi_{ij}, \quad \text{and} \quad (5)$$

$$Y_i = \beta_{0j} + \gamma_j X_i + \beta_j \widetilde{M}_i^{(j)} + \eta_{ij}, \quad (6)$$

where $\{\alpha_{0j}, \alpha_j, \beta_{0j}, \gamma_j, \beta_j\}$ is the model parameter, and ξ_{ij} and η_{ij} are independent model errors normally distributed with mean zero. Here $\alpha_j \beta_j (x - x^*)$ is the average indirect effect of the projected mediator $\widetilde{M}^{(j)}$ comparing treatment x and x^* , and $\sum_{j=1}^p \alpha_j \beta_j (x - x^*)$ is the total average indirect effect. Under this marginal LSEM, $\gamma_j (x - x^*)$ is interpreted as the treatment effect not mediated through the mediator $\widetilde{M}^{(j)}$.

As proposed in Huang and Pan (2016) [27], obtaining these causally independent mediators is achieved through principal component analysis. Consider

$$M_{ij} = \tau_{0j} + \tau_j X_i + \epsilon_{ij}, \quad \text{for } i = 1, \dots, n \text{ and } j = 1, \dots, p, \quad (7)$$

where $\{\tau_{0j}, \tau_j\}$ are model coefficients, and ϵ_{ij} is the normally distributed random error with mean zero. Let $\boldsymbol{\epsilon}_i = (\epsilon_{i1}, \dots, \epsilon_{ip})^T$, assume

$$\mathbb{E}(\epsilon_i \epsilon_i^T) = \boldsymbol{\Sigma} = \boldsymbol{\Phi} \boldsymbol{\Lambda} \boldsymbol{\Phi}^T, \quad (8)$$

where $\boldsymbol{\Phi} = (\phi_1, \dots, \phi_p) \in \mathbb{R}^{p \times p}$ is an orthonormal matrix such that $\boldsymbol{\Phi}^T \boldsymbol{\Phi} = \mathbf{I}$, and $\boldsymbol{\Lambda} = \text{diag}\{\lambda_1, \dots, \lambda_p\}$ is a p -dimensional diagonal matrix, and thus $\boldsymbol{\Phi} \boldsymbol{\Lambda} \boldsymbol{\Phi}^T$ is the spectral

decomposition of the positive-definite matrix Σ . The columns of $\tilde{\mathbf{M}} = \mathbf{M}\Phi$ are conditionally independent given X , where $\mathbf{M} = (\mathbf{M}_1, \dots, \mathbf{M}_p)^\top$ and $\mathbf{M}_i = (M_{i1}, \dots, M_{ip})^\top$ for $i = 1, \dots, n$. The method proposed in Huang and Pan (2016) [27] is summarized as follows:

Step 1. For $j = 1, \dots, p$, fit model (7) and denote the residuals as $\{e_{j1}, \dots, e_{jp}\}$ for $i = 1, \dots, n$.

Step 2. Conduct PCA on the residuals $\{e_{i1}, \dots, e_{ip}\}_{i=1}^n$ to obtain $\hat{\Phi} = (\hat{\phi}_1, \dots, \hat{\phi}_p)$ and $\hat{\Lambda}$.

Step 3. Let $\tilde{M}_i^{(j)} = \mathbf{M}_i \hat{\phi}_j$. Using the transformed mediators, perform marginal mediation analysis under models (5) and (6), for $j = 1, \dots, q$, where analogous to PCA, q is determined by the designated proportion (for example 80%) of variance explained.

Though the focus of Huang and Pan (2016) [27] is on hypothesis testing for the direct and total indirect effects, one can adapt their approach to make inference about individual pathway effects.

It is well-known that PC loadings are sign nonidentifiable. As a following step, parameters in the mediation models (5) and (6) are sign nonidentifiable as well. Here, we show that though the estimate of α and β are not sign identifiable, the estimate of the indirect and direct effects are sign consistent. Since in mediation analysis, the indirect and direct effects are the effects of interests, with Proposition 1, one can directly interpret the findings.

Proposition 1. Let $\tilde{\mathbf{M}}^{(1j)} = \mathbf{M}\phi_j$ and $\tilde{\mathbf{M}}^{(2j)} = \mathbf{M}(-\phi_j) = -\mathbf{M}\phi_j$. Let $(\hat{\alpha}_j^{(s)}, \hat{\beta}_j^{(s)}, \hat{\gamma}_j^{(s)})$ denote the estimate from the transformed mediator $\tilde{\mathbf{M}}^{(sj)}$ using models (5) and (6), for $s = 1, 2$. Then

$$\hat{\alpha}_j^{(1)} = -\hat{\alpha}_j^{(2)}, \quad \hat{\beta}_j^{(1)} = -\hat{\beta}_j^{(2)}, \quad \text{and} \quad \hat{\gamma}_j^{(1)} = \hat{\gamma}_j^{(2)}.$$

Thus the estimated direct and indirect (estimated by the product $\alpha\beta$) effects are sign invariant.

3. Sparse High-Dimensional Mediation Analysis

3.1. Motivation

As discussed in Huang and Pan (2016) [27], the estimated causal effects “do not necessarily have an intuitive interpretation”, since the transformed mediators are linear combinations of the original mediators. This drawback commonly occurs in PCA-based studies. An informal way to reduce the number of variables is to set a hard threshold and force the loadings with absolute value below the threshold to be zero. However, this can be potentially misleading; for example, see [39]. Jolliffe et al. (2003) [40] introduced the modified PCA based on the lasso [41] to yield possible zero loadings. This sparse PCA framework was then further studied by Zou et al. (2006) [42] based on the fact that sparsifying the PC loadings is equivalent to a regression-type optimization problem. In this study, we propose a sparse PCA based mediation analysis approach to estimate the mediator PCs with sparse loadings.

3.2. The lasso, the generalized lasso and the elastic net

The least absolute shrinkage and selection operator (lasso) was introduced by Tibshirani (1996) [41] to perform simultaneous variable selection and estimation in linear regression. Let $Y = (Y_1, \dots, Y_n)^\top$ denote the dependent variable, and $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_n)^\top$ where $\mathbf{X}_i = (X_{i1}, \dots, X_{ip})^\top$ ($i = 1, \dots, n$) the design matrix with p predictors. The lasso solution minimizes the squared-error loss under ℓ_1 regularization. That is,

$$\hat{\beta}_{\text{lasso}} = \arg \min_{\beta \in \mathbb{R}^p} \|Y - \mathbf{X}\beta\|_2^2 + \lambda \|\beta\|_1, \quad (9)$$

where $\beta \in \mathbb{R}^p$ is the model coefficient, $\lambda \geq 0$ is a tuning parameter, and $\|\mathbf{x}\|_1 = \sum_{j=1}^p |x_j|$ is the ℓ_1 -norm of a p -dimensional vector $\mathbf{x} \in \mathbb{R}^p$. When the tuning parameter λ is large enough, some coefficients will be shrunk to exactly zero. Under regularity conditions, the lasso estimator has been shown to be both consistent and sparsistent [see 56, 57, 58].

Tibshirani and Taylor (2011) [49] considered the problem of the generalized lasso to enforce structured constraints instead of pure sparsity. The problem can be formalized as

$$\hat{\beta}_{\text{glasso}} = \arg \min_{\beta \in \mathbb{R}^p} \|Y - \mathbf{X}\beta\|_2^2 + \lambda \|\mathbf{D}\beta\|_1, \quad (10)$$

where $\mathbf{D} \in \mathbb{R}^{m \times p}$ is a prespecified penalty matrix. The asymptotic properties of the solution was studied in She (2010) [59].

The lasso has several limitations. As discussed in Zou and Hastie (2005) [60], one limitation is that for predictors with high collinearity, the lasso tends to randomly select one of them; and second, when $p > n$, the lasso selects at most n variables. In the mediation analysis setting, the mediators are potentially causally dependent, which violates the incoherence assumption for the lasso when taking the mediators as predictors in the outcome model. Considering brain voxels as mediators, where $p \sim 100,000$, with limited number of trials $n < 100$, it is not desirable to be constrained to select at most n voxels. Zou and Hastie (2005) [60] introduced the elastic net to address these drawbacks by introducing a convex combination of ℓ_1 and ℓ_2 penalties. The elastic net solution is written as

$$\hat{\beta}_{\text{en}} = (1 + \lambda_2) \left\{ \arg \min_{\beta \in \mathbb{R}^p} \|Y - \mathbf{X}\beta\|_2^2 + \lambda_1 \|\beta\|_1 + \lambda_2 \|\beta\|_2^2 \right\}, \quad (11)$$

where $\lambda_1, \lambda_2 \geq 0$. When λ_2 is positive, the elastic net approach can potentially choose all the variables and overcomes the drawbacks with the ℓ_1 penalty only. In this study, we consider the following generalized elastic net solution, i.e.,

$$\hat{\beta}_{\text{gen}} = (1 + \lambda_2) \left\{ \arg \min_{\beta \in \mathbb{R}^p} \|Y - \mathbf{X}\beta\|_2^2 + \lambda_1 \|\mathbf{D}\beta\|_1 + \lambda_2 \|\beta\|_2^2 \right\}, \quad (12)$$

to impose a structured regularization, for example the fused lasso [48]. Motivated by a protein mass spectroscopy study, the fused lasso was introduced to encourage sparsity in the coefficients as well as in their successive differences. This local constancy property is also desirable in fMRI data, as a small-world topology supports both segregated/specialized and distributed/integrated information processing in the human brain [61].

3.3. Sparse approximation

For PCA, Zou et al. (2006) [42] studied a simple regression approach to recover the PC loadings and showed that with a ridge penalty, the normalized solution to the regression problem, by regressing the loadings on the variables, is independent of λ . With this property, the inclusion of ridge penalty is not meant to penalize the regression coefficients but to ensure the reconstruction of principal components. As described in Section 2, in mediation analysis, PCA is conducted on the residuals of the mediation models. Therefore, we propose to sparsify the loadings using these model residuals, i.e., considering the following optimization problem, for $k = 1, \dots, q$,

$$\hat{\mathbf{v}}_k = (1 + \lambda_2) \left\{ \arg \min_{\mathbf{v} \in \mathbb{R}^p} \|\tilde{\mathbf{e}}^{(k)} - \mathbf{E}\mathbf{v}\|_2^2 + \lambda_1 \|\mathbf{D}\mathbf{v}\|_1 + \lambda_2 \|\mathbf{v}\|_2^2 \right\}, \quad (13)$$

where $\mathbf{v} = (v_1, \dots, v_p)^\top \in \mathbb{R}^p$; $\mathbf{E} = (\mathbf{e}_1, \dots, \mathbf{e}_n)^\top$ with $\mathbf{e}_i = (e_{i1}, \dots, e_{ip})^\top$ is the residual matrix, and $\tilde{\mathbf{e}}^{(k)} = \mathbf{E}\hat{\phi}_k$ is the k th principal component. Then $\hat{\mathbf{w}}_k = \hat{\mathbf{v}}_k / \|\hat{\mathbf{v}}_k\|_2$ is a sparse approximation of $\hat{\phi}_k$.

In neuroimaging studies, a spatial smoothness constraint is commonly applied (for example, see [46, 47]). In this study, we consider a fused lasso penalty [48] to impose a local constancy of the PC profile, i.e., neighbor brain voxels/regions are assumed to have the same loading weights:

$$\hat{\mathbf{v}}_k = (1 + \lambda_2) \left\{ \arg \min_{\mathbf{v} \in \mathbb{R}^p} \|\tilde{\mathbf{e}}^{(k)} - \mathbf{E}\mathbf{v}\|_2^2 + \lambda_{11} \|\mathbf{v}\|_1 + \lambda_{12} \sum_{(j, j') \in \mathcal{E}} |v_j - v_{j'}| + \lambda_2 \|\mathbf{v}\|_2^2 \right\}, \quad (14)$$

where $\mathcal{E}$ is an edge set such that $(j, j') \in \mathcal{E}$ if M_j and $M_{j'}$ are neighboring brain voxels/regions as defined by their spatial locations. For other areas of application, the edge set can be defined based on domain knowledge. For example in genetic studies, it can be defined based on locus information. Formulation (14) is a special case of the generalized elastic net (12) with $\mathbf{D}$ the fused lasso matrix corresponding to the underlying graph with edge set $\mathcal{E}$.

3.3.1. Tuning parameter selection—Zou et al. (2006) [42] showed that the inclusion of a ridge penalty does not penalize the regression coefficients. Thus, when $n \gg p$, we can set the tuning parameter for ridge penalty λ_2 to zero; and when $n \ll p$, we can in principle use any positive λ_2 . The objective of sparse PCA is to sparsify the loadings while preserving the proportion of variance explained. Zou et al. (2006) [42] proposed a data-driven approach

to choose λ_1 (the ℓ_1 -penalty tuning parameter) by examining the trace plot of the percentage of total variance explained calculated from the adjusted total variance. We apply the same strategy to choose tuning parameters λ_{11} and λ_{12} in (14).

3.4. Mediation analysis with sparse principal components

In this section, we discuss analysis with the sparse PC of the mediators. Let

$\tilde{M}_i^{(k)} = \mathbf{M}_i^T \mathbf{w}_k = \sum_{j=1}^p M_{ij} w_{kj}$, and for $k = 1, \dots, q$, fit models

$$\begin{aligned}\tilde{M}_i^{(k)} &= a_0 + a_k X_i + \epsilon_i^{(k)}, \quad \text{and} \\ Y_i &= b_0 + c_k X_i + b_k \tilde{M}_i^{(k)} + \eta_i^{(k)},\end{aligned}\tag{15}$$

where $\epsilon_i^{(k)}$ and $\eta_i^{(k)}$ are independent random errors normally distributed with mean zero. One advantage of the PCA mediation analysis is that the transformed mediator PCs are conditionally independent, and fitting the LSEM with multiple mediators is equivalent to using marginal LSEMs for each individual mediator. By sparsifying the loading vector, the orthogonality constraint is not explicitly imposed. To achieve the conditional independence, we include a regression projection step to remove the conditional linear dependence between the transformed mediators, analogous to the procedure proposed in Zou et al. (2006) [42].

Let $\tilde{M}_i^{(k \cdot 1, \dots, k-1)}$ denote the residual after adjusting for $\tilde{M}_i^{(1)}, \dots, \tilde{M}_i^{(k-1)}$ when controlling X_i (for $i = 1, \dots, n$). That is,

$$\tilde{M}_i^{(k \cdot 1, \dots, k-1)} = \tilde{M}_i^{(k)} - \tilde{\mathbf{M}}_i^{(1, \dots, k-1)T} \hat{\boldsymbol{\Pi}}_{1, \dots, k-1},\tag{16}$$

where $\tilde{\mathbf{M}}_i^{(1, \dots, k-1)} = (\tilde{M}_i^{(1)}, \dots, \tilde{M}_i^{(k-1)})^T$, and $\hat{\boldsymbol{\Pi}}_{1, \dots, k-1}$ is the estimated coefficient in model

$$\tilde{M}_i^{(k)} = \pi_0 + \pi_1 X_i + \tilde{\mathbf{M}}_i^{(1, \dots, k-1)T} \boldsymbol{\Pi}_{1, \dots, k-1} + \tau_i^{(k)},\tag{17}$$

where $\tau_i^{(k)}$ is normally distributed model error with mean zero. The new mediators

$\tilde{M}_i^{(k \cdot 1, \dots, k-1)}$ are uncorrelated given the treatment X . Thus, we can use model (15) to estimate the indirect effect of each individual mediation pathway.

We summarize the steps of mediation analysis with sparse principal components in Algorithm 1. To perform inference on model parameters, we propose a bootstrap procedure.

- i. Generate a bootstrap sample $(X_i^*, \tilde{M}_i^{(k \cdot 1, \dots, k-1)*}, Y_i^*)$ of size n by resampling the data with replacement, where $\tilde{M}_i^{(k \cdot 1, \dots, k-1)}$ is the modified mediator PC obtained in Step 4 of Algorithm 1.
- ii. Estimate the model parameters in (15) using the bootstrap sample.
- iii. Repeat steps (i)-(ii) B times.

Bootstrap confidence intervals can be then calculated using either the percentile or bias-corrected approach [62] under a pre-specified significance level.

The motivation of sparsifying the loading vectors is to facilitate intuitive interpretations of the findings. The estimated causal effects under the new mediators are desired to be as close as possible to the ones estimated from the original mediator PCs (the $\tilde{M}^{(j)}$'s). In Section B.2 of the supplementary material, we provide a sketch of the asymptotic convergence of the estimators from Algorithm 1 and those from $\tilde{M}^{(j)}$'s in models (5) and (6).

Algorithm 1 Mediation analysis with sparse principal components.

Step 1. For $j = 1, \dots, p$, fit model (7) and denote the residuals as $\{e_{i1}, \dots, e_{ip}\}$ for $i = 1, \dots, n$.

Step 2. Conduct PCA on the residuals $\{e_{i1}, \dots, e_{ip}\}_{i=1}^n$ and estimate the loading matrix as $\hat{\Phi} = (\hat{\phi}_1, \dots, \hat{\phi}_p)$.

Step 3. For $k = 1, \dots, q$, let $\tilde{\mathbf{e}}^{(k)} = \mathbf{E}\hat{\phi}_k$, where $\mathbf{E} = (\mathbf{e}_1, \dots, \mathbf{e}_n)^\top$ and $\mathbf{e}_i = (e_{i1}, \dots, e_{ip})^\top$ ($i = 1, \dots, n$). Perform regularized regression using the generalized elastic net penalty (12) and attain the estimator $\hat{\mathbf{v}}_k(\hat{\lambda}_k)$ and its normalization $\hat{\mathbf{w}}_k(\hat{\lambda}_k) = \hat{\mathbf{v}}_k(\hat{\lambda}_k) / \|\hat{\mathbf{v}}_k(\hat{\lambda}_k)\|_2$, for $k = 1, \dots, q$, where $\hat{\lambda}_k$ is chosen based on the method discussed in Section 3.3.1.

Step 4. Let $\tilde{M}_i^{(k)} = \mathbf{M}_i^\top \hat{\mathbf{w}}_k$ for $i = 1, \dots, n$. For $k = 2, \dots, q$ obtain the modified $\tilde{M}_i^{(k \cdot 1, \dots, k-1)}$.

Step 5. Fit model (15) using the causally independent $\{\tilde{M}_i^{(1)}, \tilde{M}_i^{(2 \cdot 1)}, \dots, \tilde{M}_i^{(q \cdot 1, \dots, q-1)}\}$ to yield estimates of the causal effects.

4. Simulation Study

A simulation study was conducted to examine the performance of the proposed sparse PC based mediation analysis approach. A description of the study and the results are presented in Section D of the supplementary material available online. To summarize, the proposed sparse PC based method achieves comparable estimates of the indirect effect to the PCA based approach introduced in Huang and Pan (2016) [27] under various values of p and n . Using bootstrap confidence intervals following the procedure described in Section 3.4, we find that for those PCs with a nonzero indirect effect, the coverage probability can be slightly lower than the nominal level, but the power is reasonably high and converges to one as n increases. In addition, for those PCs with a null indirect effect, the coverage probability and Type I error are well controlled at the designated level.

5. A Task-Based Functional MRI Study

We analyze a task-based fMRI study using the proposed sparse principal component of mediation approach. The data set is downloaded from the OpenfMRI database (accession number ds000002). In the experiment, participants were instructed to perform a probabilistic classification learning (PCL) task for weather prediction [63]. To avoid inter-subject heterogeneity, we use the data from a single healthy right-handed English-speaking subject aged between 21 to 26. The experiment consisted of $n = 80$ trials with ten cycles, and within each there are five PCL trials intermixed with three baseline trials. Under the weather prediction trial, a visual stimulus was presented at a randomized location. The participant

would respond by pressing either the left button for a “sun” prediction or the right button for “rain”. The experiment included baseline trials to control for visual stimulation, button press and computer response to button press. A fundamental question that PCL task-fMRI study aims to address is whether and how the memory systems interact. In this study, we take a step further considering reaction time of button pressing (Y) as the outcome of interest, and the goal is to discover the brain networks that have an intermediate effect on the reaction time when comparing PCL ($X = 1$) and baseline ($X = 0$) trials.

One hundred and eighty functional T2*-weighted echoplanar images (EPI) (4 mm slice thickness, 33 slices, TR = 2 s, TE = 30 ms, flip angle = 90° , matrix 64×64 , field of view 200) were acquired from a 3 T Siemens Allegra MRI scanner. For registration purposes, a matched-bandwidth High-Resolution scan (same slice prescription as EPI) and MPRAGE (TR = 2.3, TE = 2.1, FOV = 256, matrix = 192×192 , sagittal plane, slice thickness = 1 mm, 160 slices) were acquired for each participant. Preprocessing for both anatomical and functional images was conducted using Statistical Parametric Mapping version 5 (SPM5) (Wellcome Department of Imaging Neuroscience, University College London, London, UK), including slice timing correction, realignment, coregistration, normalization, and smoothing. The blood-oxygen-level dependent (BOLD) time courses are extracted from $p = 264$ putative brain functional regions defined in Power et al. (2011) [64]. These brain regions are grouped into eleven (ten functional and one uncertain) modules. We employed the general linear model approach to acquire a summary measure of brain activity for each trial at each region [32, 65, 66]. These single-trial brain activities were used as the mediators (M).

We applied the PCA based high-dimensional mediation analysis proposed in Huang and Pan (2016) [27] and our proposed sparse PCA based method. When conducting sparse approximation, we consider a fused lasso penalty (14), where edge set $\mathcal{E}$ is defined based on the spatial location of the brain region, as well as module information given the modular processing nature of human brains [61]. If brain regions j and k are from the same functional module, then $(j, k) \in \mathcal{E}$; if region k is the spatially nearest neighbor of region j , then $(j, k) \in \mathcal{E}$. In the PCA based analysis, the first 18 PCs, which account for 76.2% of the total variation, were tested for mediation effect (see Figure C.1 in the supplementary material). PC3 shows significant positive indirect effect, where the linear combination of all brain regions shows deactivation in PCL compared to baseline (α estimate -0.376 with 95% confidence interval $(-0.592, -0.160)$), and this deactivation further increases the reaction time (β estimate -0.235 with 95% confidence interval $(-0.365, -0.105)$). The loadings of PC3 are shown in Figure C.2. In order to achieve the goal of region selection and make the interpretation feasible, a generally applied approach is to choose a threshold and only regions with magnitude above the threshold are considered to contribute to the PC. However, the threshold is usually chosen arbitrarily. Furthermore, including a thresholding step may lead to deceptive interpretations (see the discussion in Section C of the supplementary materials). Here, we utilize a principled procedure through lasso regularization. Figure 2 shows the sparse approximation under the fused lasso penalty. The estimated model coefficients, as well as the indirect effect are very close to those in the PCA based analysis (Table 1 and Figure C.1 in the supplementary material). From the figure, the whole

cerebellum module is regularized to zero. All the visual modules, including lateral, medial and occipital pole visual, yield negative loadings; the positive loadings are mainly from the auditory, default mode network, executive control and frontoparietal modules (see Figure C.3 in the supplementary material). Figure 3 shows the regions with positive and negative loadings in a brain map. The medial frontal and parietal cortex and the auditory cortex are often identified to be deactivated (positive loadings with negative α estimate) when the task involves visual stimuli [67, 68]. The positive loading map (Figure 3a) also includes the medial temporal lobe (MTL) regions, which is in line with the findings in the existing literature [67]. The negative loading map (Figure 3b) consists of the visual cortex and the basal ganglia (caudate nucleus). This classification learning task is a nondeclarative memory procedure. The opposite sign of the loadings in MTL and striatum verifies the competing role of these two memory system regions during learning [67]. Positive correlations between reaction time and activation in visual cortex, fusiform gyrus and thalamus have been discovered, which are presumably involved in processing sensory feedback related to the motor response [69]. Through our mediation analysis, our finding connects the brain pathways that are responsive to the classification learning task and contribute to the longer reaction time.

6. Discussion

In this study, we introduced a sparse principal component analysis (PCA) based approach for performing mediation analysis with high-dimensional mediators. As an extension of the method introduced in Huang and Pan (2016) [27], the proposed approach increases the interpretability of the transformed mediators by employing the sparse PCA method proposed in Zou et al. (2006) [42]. Though our proposed sparse PCA based method introduces a slight estimation bias in model coefficients compared to the PCA based method, the estimate of the mediation effect converges and the findings are more interpretable. Thus, this can be seen as a trade-off between estimation bias and interpretability. In the task-based fMRI application, we consider the fused lasso penalty based on spatial information to enforce local smoothness and constancy. Using sparse loadings, the activation patterns of the brain regions in the PC with significant mediation effect are consistent with those found in the existing literature. The proposed framework is based on the assumption that there are p underlying ordered mediators, where the knowledge of the specific ordering is not necessarily known. When considering brain activation as the mediators of cognitive behaviors, the causal relationship between brain regions is referred to as effective connectivity [54]. Thus, the assumption of p -ordered mediators is consistent with a working model of connectivity; though we acknowledge potential limitations should this prove deficient.

Though our proposed high-dimensional mediation analysis approach is motivated by neuroimaging studies, the fundamental principle can be generalized to other areas, for example genetics studies. Given the spatial information of the brain, we consider a special type of the generalized lasso, namely the fused lasso. Other lasso-type structured regularization can be adopted based on data characteristics, including group lasso [70] and its variations [71].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institutes of Health [P41 EB015909, R01 EB016061 and R01 EB026549 from the National Institute of Biomedical Imaging and Bioengineering].

References

- [1]. Baron RM, Kenny DA, The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations., *Journal of personality and social psychology* 51 (6) (1986) 1173. [PubMed: 3806354]
- [2]. Holland PW, Causal inference, path analysis, and recursive structural equations models, *Sociological methodology* 18 (1) (1988) 449–484.
- [3]. Robins JM, Greenland S, Identifiability and exchangeability for direct and indirect effects, *Epidemiology* (1992) 143–155. [PubMed: 1576220]
- [4]. Pearl J, Direct and indirect effects, in: *Proceedings of the seventeenth conference on uncertainty in artificial intelligence*, Morgan Kaufmann Publishers Inc., 2001, pp. 411–420.
- [5]. Ten Have TR, Joffe MM, Lynch KG, Brown GK, Maisto SA, Beck AT, Causal mediation analyses with rank preserving models, *Biometrics* 63 (3) (2007) 926–934. [PubMed: 17825022]
- [6]. MacKinnon DP, *Introduction to statistical mediation analysis*, Routledge, 2008.
- [7]. Sobel ME, Identification of causal parameters in randomized studies with mediating variables, *Journal of Educational and Behavioral Statistics* 33 (2) (2008) 230–251.
- [8]. VanderWeele TJ, Vansteelandt S, Conceptual issues concerning mediation, interventions and composition, *Statistics and its Interface* 2 (2009) 457–468.
- [9]. Imai K, Keele L, Yamamoto T, Identification, inference and sensitivity analysis for causal mediation effects, *Statistical Science* (2010) 51–71.
- [10]. VanderWeele TJ, *Explanation in Causal Inference: Methods for Mediation and Interaction*, Oxford University Press, 2015.
- [11]. MacKinnon DP, Contrasts in multiple mediator models.
- [12]. Preacher KJ, Hayes AF, Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models, *Behavior research methods* 40 (3) (2008) 879–891. [PubMed: 18697684]
- [13]. Imai K, Yamamoto T, Identification and sensitivity analysis for multiple causal mechanisms: Revisiting evidence from framing experiments, *Political Analysis* 21 (2) (2013) 141–171.
- [14]. Wang W, Nelson S, Albert JM, Estimation of causal mediation effects for a dichotomous outcome in multiple-mediator models using the mediation formula, *Statistics in medicine* 32 (24) (2013) 4211–4228. [PubMed: 23650048]
- [15]. VanderWeele TJ, Vansteelandt S, Robins JM, Effect decomposition in the presence of an exposure-induced mediator-outcome confounder, *Epidemiology* 25 (2) (2014) 300–306. [PubMed: 24487213]
- [16]. VanderWeele TJ, Vansteelandt S, Mediation analysis with multiple mediators, *Epidemiologic methods* 2 (1) (2014) 95–115. [PubMed: 25580377]
- [17]. Zhao SD, Cai TT, Li H, More powerful genetic association testing via a new statistical framework for integrative genomics, *Biometrics* 70 (4) (2014) 881–890. [PubMed: 24975802]
- [18]. Boca SM, Sinha R, Cross AJ, Moore SC, Sampson JN, Testing multiple biological mediators simultaneously, *Bioinformatics* 30 (2) (2014) 214–220. [PubMed: 24202540]
- [19]. Daniel R, De Stavola B, Cousens S, Vansteelandt S, Causal mediation analysis with multiple mediators, *Biometrics* 71 (1) (2015) 1–14. [PubMed: 25351114]
- [20]. Taguri M, Featherstone J, Cheng J, Causal mediation analysis with multiple causally non-ordered mediators, *Statistical methods in medical research* (2015) 0962280215615899.

- [21]. Nguyen TQ, Webb-Vargas Y, Koning IM, Stuart EA, Causal mediation analysis with a binary outcome and multiple continuous or ordinal mediators: Simulations and application to an alcohol intervention, *Structural equation modeling: a multidisciplinary journal* 23 (3) (2016) 368–383. [PubMed: 27158217]
- [22]. Lin S-H, VanderWeele TJ, Interventional approach for path-specific effects, *Journal of Causal Inference* 5 (1).
- [23]. Vansteelandt S, Daniel RM, Interventional effects for mediation analysis with multiple mediators, *Epidemiology (Cambridge, Mass.)* 28 (2) (2017) 258.
- [24]. Steen J, Loeys T, Moerkerke B, Vansteelandt S, Medflex: An r package for flexible mediation analysis using natural effect models, *Journal of Statistical Software* 76 (11).
- [25]. Calcagni A, Lombardi L, Avanzi L, Pascali E, Multiple mediation analysis for interval-valued data, *Statistical Papers* (2017) 1–23.
- [26]. Park S, Kürüm E, Causal mediation analysis with multiple mediators in the presence of treatment noncompliance, *Statistics in Medicine*.
- [27]. Huang Y-T, Pan W-C, Hypothesis test of mediation effect in causal mediation model with high-dimensional continuous mediators, *Biometrics* 72 (2) (2016) 402–413. [PubMed: 26414245]
- [28]. Zhang H, Zheng Y, Zhang Z, Gao T, Joyce B, Yoon G, Zhang W, Schwartz J, Just A, Colicino E, Estimating and testing high-dimensional mediation effects in epigenetic studies, *Bioinformatics* (2016) btw351.
- [29]. Caffo B, Chen S, Stewart W, Bolla K, Yousem D, Davatzikos C, Schwartz BS, Are brain volumes based on magnetic resonance imaging mediators of the associations of cumulative lead dose with cognitive function?, *American journal of epidemiology* 167 (4) (2007) 429–437. [PubMed: 18079133]
- [30]. Wager TD, Davidson ML, Hughes BL, Lindquist MA, Ochsner KN, Prefrontal-subcortical pathways mediating successful emotion regulation, *Neuron* 59 (6) (2008) 1037–1050. [PubMed: 18817740]
- [31]. Wager TD, Waugh CE, Lindquist M, Noll DC, Fredrickson BL, Taylor SF, Brain mediators of cardiovascular responses to social threat: part i: Reciprocal dorsal and ventral sub-regions of the medial prefrontal cortex and heart-rate reactivity, *Neuroimage* 47 (3) (2009) 821–835. [PubMed: 19465137]
- [32]. Atlas LY, Bolger N, Lindquist MA, Wager TD, Brain mediators of predictive cue effects on perceived pain, *The Journal of neuroscience* 30 (39) (2010) 12964–12977. [PubMed: 20881115]
- [33]. Lindquist MA, Functional causal mediation analysis with an application to brain connectivity, *Journal of the American Statistical Association* 107 (500) (2012) 1297–1309. [PubMed: 25076802]
- [34]. Atlas LY, Lindquist MA, Bolger N, Wager TD, Brain mediators of the effects of noxious heat on pain, *PAIN®* 155 (8) (2014) 1632–1648. [PubMed: 24845572]
- [35]. Woo C-W, Roy M, Buhle JT, Wager TD, Distinct brain systems mediate the effects of nociceptive input and self-regulation on pain, *PLoS biology* 13 (1) (2015) e1002036. [PubMed: 25562688]
- [36]. Chén OY, Crainiceanu C, Ogburn EL, Caffo BS, Wager TD, Lindquist MA, High-dimensional multivariate mediation with application to neuroimaging data, *Biostatistics* 19 (2) (2017) 121–136.
- [37]. Geuter S, Losin EA, Roy M, Atlas LY, Schmidt L, Krishnan A, Koban L, Wager TD, Lindquist MA, Multiple brain networks mediating stimulus-pain relationships in humans, *bioRxiv* (2018) 298927.
- [38]. Zhao Y, Luo X, Pathway lasso: Estimate and select sparse mediation pathways with high dimensional mediators, *arXiv preprint arXiv:1603.07749*.
- [39]. Cadima J, Jolliffe IT, Loading and correlations in the interpretation of principle compenents, *Journal of Applied Statistics* 22 (2) (1995) 203–214.
- [40]. Jolliffe IT, Trendafilov NT, Uddin M, A modified principal component technique based on the lasso, *Journal of computational and Graphical Statistics* 12 (3) (2003) 531–547.
- [41]. Tibshirani R, Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society. Series B (Methodological)* (1996) 267–288.

- [42]. Zou H, Hastie T, Tibshirani R, Sparse principal component analysis, *Journal of computational and graphical statistics* 15 (2) (2006) 265–286.
- [43]. Zhou J, He X, et al., Dimension reduction based on constrained canonical correlation and variable filtering, *The Annals of Statistics* 36 (4) (2008) 1649–1668.
- [44]. Witten DM, Tibshirani R, Hastie T, A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis, *Biostatistics* 10 (3) (2009) 515–534. [PubMed: 19377034]
- [45]. Witten DM, Tibshirani RJ, Extensions of sparse canonical correlation analysis with applications to genomic data, *Statistical applications in genetics and molecular biology* 8 (1) (2009) 1–27.
- [46]. Grosenick L, Klingenberg B, Katovich K, Knutson B, Taylor JE, Interpretable whole-brain prediction analysis with graphnet, *Neuroimage* 72 (2013) 304–321. [PubMed: 23298747]
- [47]. Liu LY-F, Liu Y, Zhu H, Initiative ADN, Smac: Spatial multi-category angle-based classifier for high-dimensional neuroimaging data, *NeuroImage*.
- [48]. Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K, Sparsity and smoothness via the fused lasso, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67 (1) (2005) 91–108.
- [49]. Tibshirani RJ, Taylor J, The solution path of the generalized lasso, *The Annals of Statistics* 39 (3) (2011) 1335–1371. doi:10.1214/11-AOS878. URL 10.1214/11-AOS878
- [50]. Rubin DB, Bayesian inference for causal effects: The role of randomization, *The Annals of Statistics* (1978) 34–58.
- [51]. Rubin DB, Causal inference using potential outcomes, *Journal of the American Statistical Association* 100 (469).
- [52]. Shojaie A, Michailidis G, Penalized likelihood methods for estimation of sparse high-dimensional directed acyclic graphs, *Biometrika* 97 (3) (2010) 519–538. [PubMed: 22434937]
- [53]. Obeso I, Cho SS, Antonelli F, Houle S, Jahanshahi M, Ko JH, Strafella AP, Stimulation of the pre-sma influences cerebral blood flow in frontal areas involved with inhibitory control of action, *Brain stimulation* 6 (5) (2013) 769–776. [PubMed: 23545472]
- [54]. Friston KJ, Functional and effective connectivity in neuroimaging: a synthesis, *Human brain mapping* 2 (1–2) (1994) 56–78.
- [55]. Lindquist MA, Sobel ME, Effective connectivity and causal inference in neuroimaging, *Handbook of Neuroimaging Data Analysis* (2016) 419.
- [56]. Meinshausen N, Bühlmann P, High-dimensional graphs and variable selection with the lasso, *The Annals of Statistics* (2006) 1436–1462.
- [57]. Wainwright M, Sharp thresholds for noisy and high-dimensional recovery of sparsity using ℓ_1 -constrained quadratic programming (lasso), *IEEE Transactions on Information Theory* 55 (5) (2009) 2183–2202.
- [58]. Zhao P, Yu B, On model selection consistency of lasso, *The Journal of Machine Learning Research* 7 (2006) 2541–2563.
- [59]. She Y, Sparse regression with exact clustering, *Electronic Journal of Statistics* 4 (2010) 1055–1096.
- [60]. Zou H, Hastie T, Regularization and variable selection via the elastic net, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67 (2) (2005) 301–320.
- [61]. Bassett DS, Bullmore E, Small-world brain networks, *The neuroscientist* 12 (6) (2006) 512–523. [PubMed: 17079517]
- [62]. Efron B, Better bootstrap confidence intervals, *Journal of the American statistical Association* 82 (397) (1987) 171–185.
- [63]. Aron AR, Gluck MA, Poldrack RA, Long-term test-retest reliability of functional mri in a classification learning task, *Neuroimage* 29 (3) (2006) 1000–1006. [PubMed: 16139527]
- [64]. Power JD, Cohen AL, Nelson SM, Wig GS, Barnes KA, Church JA, Vogel AC, Laumann TO, Miezin FM, Schlaggar BL, Functional network organization of the human brain, *Neuron* 72 (4) (2011) 665–678. [PubMed: 22099467]
- [65]. Lindquist MA, The statistical analysis of fmri data, *Statistical Science* 23 (4) (2008) 439–464.

- [66]. Rissman J, Gazzaley A, D'Esposito M, Measuring functional connectivity during distinct stages of a cognitive task, *Neuroimage* 23 (2) (2004) 752–763. [PubMed: 15488425]
- [67]. Poldrack RA, Clark J, Pare-Blagoev E, Shohamy D, Moyano JC, Myers C, Gluck MA, Interactive memory systems in the human brain, *Nature* 414 (6863) (2001) 546–550. [PubMed: 11734855]
- [68]. Aron AR, Shohamy D, Clark J, Myers C, Gluck MA, Poldrack RA, Human midbrain sensitivity to cognitive feedback and uncertainty during classification learning, *Journal of neurophysiology* 92 (2) (2004) 1144–1152. [PubMed: 15014103]
- [69]. Yarkoni T, Barch DM, Gray JR, Conturo TE, Braver TS, Bold correlates of trial-by-trial reaction time variability in gray and white matter: a multi-study fmri analysis, *PLoS One* 4 (1) (2009) e4257. [PubMed: 19165335]
- [70]. Yuan M, Lin Y, Model selection and estimation in regression with grouped variables, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 68 (1) (2006) 49–67.
- [71]. Yuan L, Liu J, Ye J, Efficient methods for overlapping group lasso, in: *Advances in Neural Information Processing Systems*, 2011, pp. 352–360.

Highlights

- We propose an approach that performs mediation analysis on a large number of causally dependent mediators.
- Our proposed method facilitates intuitive interpretations through sparse principal component analysis.
- Software, in the form of R code, together with a sample data set and complete documentation is available on Github at <https://github.com/zhaoyi1026/spcma>.

Figure 1:

Causal diagram of (a) single mediator, (b) p dependent ordered mediators and (c) q causally independent transformed mediators.

Figure 2:
The sparse approximation of the loadings of PC3. The x -axis shows the index of the 264 brain regions which are colored by functional modules.

(a) Brain regions with positive loadings.

(b) Brain regions with negative loadings

Figure 3: Brain regions with (a) positive and (b) negative loadings in sparsified PC3. The module color code is the same as in Figure 2. The size of the node is proportion to the absolute value of the loading.

Table 1:

The estimate (Est.) of model parameters α , β , and the indirect effect (IE), as well as the 95% bootstrap confidence interval (CI) of PC3, which yields significant IE.

	PC3			
	PCA		SPCA	
	Est.	95% CI	Est.	95% CI
α	-0.376	(-0.592, -0.160)	-0.343	(-0.534, -0.151)
β	-0.235	(-0.365, -0.105)	-0.262	(-0.410, -0.115)
IE ($\alpha\beta$)	0.089	(0.020, 0.176)	0.090	(0.025, 0.178)