

A novel multi-exposure-to-multi-mediator mediation model for imaging genetic study of brain disorders

Neng Wang^{a,c}, Eric V. Slud ^{b,*}, Tianzhou Ma ^{c,**}

^a*Department of Mathematics, University of Maryland, College Park, MD 20742, USA*

^b*Department of Mathematics, University of Maryland, College Park, MD 20742, USA*

^c*Department of Epidemiology and Biostatistics, University of Maryland, College Park, MD 20742, USA*

Abstract

Common psychiatric and brain disorders are highly heritable and affected by a number of genetic risk factors, yet the mechanism by which these genetic factors contribute to the disorders through alterations in brain structure and function remain poorly understood. Contemporary imaging genetic studies integrate genetic and neuroimaging data to investigate how genetic variation contributes to brain disorders via intermediate neuroimaging endophenotypes. However, the large number of potential exposures (genes) and mediators (neuroimaging features) pose new challenges to the traditional mediation analysis. In this paper, we propose a novel multi-exposure-to-multi-mediator mediation model that integrates genetic, neuroimaging and phenotypic data to investigate the “gene-neuroimaging-brain disorder” mediation pathway. Our method jointly reduces the dimensions of exposures and mediators into low-dimensional aggregators where the mediation effect is maximized. We further introduce sparsity into the loadings to improve the interpretability. To target the bi-convex optimization problem, we implement an efficient alternating direction method of multipliers algorithm with block coordinate updates. We provide theoretical guarantees for the convergence of our algorithm and establish the asymptotic properties of the resulting estimators. Through extensive simulations, we demonstrate that our method outperforms other competing methods in recovering true loadings and true mediation proportions across a wide range of signal strengths, noise levels, and correlation structures. We further illustrate the utility of the method through a mediation analysis that integrates genetic, brain functional connectivity and smoking behavior data from UK Biobank, and identifies critical genes that impact nicotine dependence via changing the functional connectivity in specific brain regions.

1. Introduction

Common psychiatric and brain disorders are highly heritable and polygenic, influenced by numerous genetic risk factors [1, 2]. Previous genome-wide association studies (GWAS) and transcriptome-wide association studies (TWAS) have identified critical genetic risk factors at both SNP and gene levels and improved our understanding of the genetic influence on these

*Corresponding author. Email: slud@umd.edu

**Corresponding author. Email: tma0929@umd.edu

disorders [1, 3, 4, 5]. Yet, how these genes impact the disorders via altering the brain structure or function are not well understood. Studies using the magnetic resonance imaging (MRI) data have found heritable neuroimaging features including both structural and functional signatures of these brain disorders [6], laying out the basis of imaging genetic studies of brain disorders that examine how genetic variation influences brain disorders through intermediate neuroimaging endophenotypes [7].

Mediation analysis is a statistical framework used to delineate how exposures affect outcomes through intermediate variables known as mediators, providing key insights into the causal mechanism [8]. Traditional mediation analysis deals with one single exposure and one single mediator at a time. With the advancement of technology, contemporary studies frequently involve a large number of exposures and mediators, e.g. genetic and neuroimaging features in imaging genetic studies (Fig 1), which present new methodological challenges that classical mediation approaches are ill-equipped to address. A number of methods have been developed to address these challenges in mediation analysis. One major class of methods considers performing screening or regularization procedures to reduce the number of mediators and/or exposures first before the mediation analysis [9, 10, 11, 12]. While effective for large mediator sets, these methods usually separate the mediator selection from mediation effect estimation, potentially overlooking joint mediator effects and introducing false positives. Another class of methods applies dimension reduction on high-dimensional mediators, including principal component analysis and its sparse variants [13, 14, 15]. However, the dimension reduction step is independent of mediation analysis thus the mapped low-dimensional space is not directly related to the mediation pathway. [16] considered seeking linear mediator combinations maximizing mediated likelihood within structural equation models using the Directions of Mediation (DMs) approach. [17] proposed a multi-mediator method that mapped the high-dimensional mediators to a space that maximizes the mediation proportion, and [18] introduced the Mediation Multiple Pathways (MMP) framework, which aggregates evidence across multiple linear mediation directions by enforcing orthogonality and controlling false discovery in both marginal and conditional indirect effects. Few methods have effectively balanced between model flexibility, model interpretability and computational efficiency in these new mediation analysis problems.

In imaging genetic study of brain disorders, both genetic risk factors and neuroimaging features are potentially high-dimensional. Genes are correlated and working together to carry similar functions [19]. At the same time, the neuroimaging features (e.g. MRI features from different brain regions) are inherently spatially-correlated [20] (Fig 1). We need to consider the simultaneous selection of exposures and mediators from a large number of correlated candidates. In addition, methods creating latent mediator composites often sacrifice interpretability, complicating the translation of findings into actionable biological insights. Lastly, heavily penalized models may exhibit instability, with sparse solutions highly sensitive to data perturbations and tuning parameter choices, thus limiting reproducibility [13, 15].

To overcome these challenges, in this paper, we introduced a novel multi-exposure-to-multi-mediator mediation model that integrates genetic, neuroimaging and phenotypic data to investigate the “gene-neuroimaging-brain disorder” pathway (Figure 1). In our model, we jointly reduced the dimensions of exposures and mediators into a lower-dimensional space that captures the strongest mediating effect of brain imaging in the mediation pathway. In addition, we imposed structured sparsity constraints on loadings to select a small subset of

exposures and mediators, maintaining interpretability and stability. To efficiently estimate the proposed model, we developed a tailored optimization algorithm based on alternating direction method of multipliers (ADMM). We also theoretically showed the convergence of our algorithm and asymptotic properties of our estimates. Our method consistently outperformed competing methods in simulation studies, showing improved power, enhanced mediator selection stability, and better interpretability. Additionally, we illustrated the utility of our proposed method through an integrative mediation analysis of gene expression, brain functional connectivity, and smoking behavior data, demonstrating how critical genes (e.g. *CHRNA5-CHRNA3-CHRN*_{B4} gene cluster) impact nicotine dependence via changing the brain functional connectivity between middle frontal gyrus, precentral gyrus, thalamus and superior parietal lobule regions.

The remainder of this paper is structured as follows. Section 2 details our methodological framework, estimation procedure and theoretical guarantee. Section 3 presents extensive simulations comparing our method against existing approaches. Section 4 applies our method to a real data example investigating the role of brain functional connectivity features that mediate the genetic effect on nicotine addiction in UK Biobank. Section 5 concludes with a discussion of the findings, potential limitations, and avenues for future research.

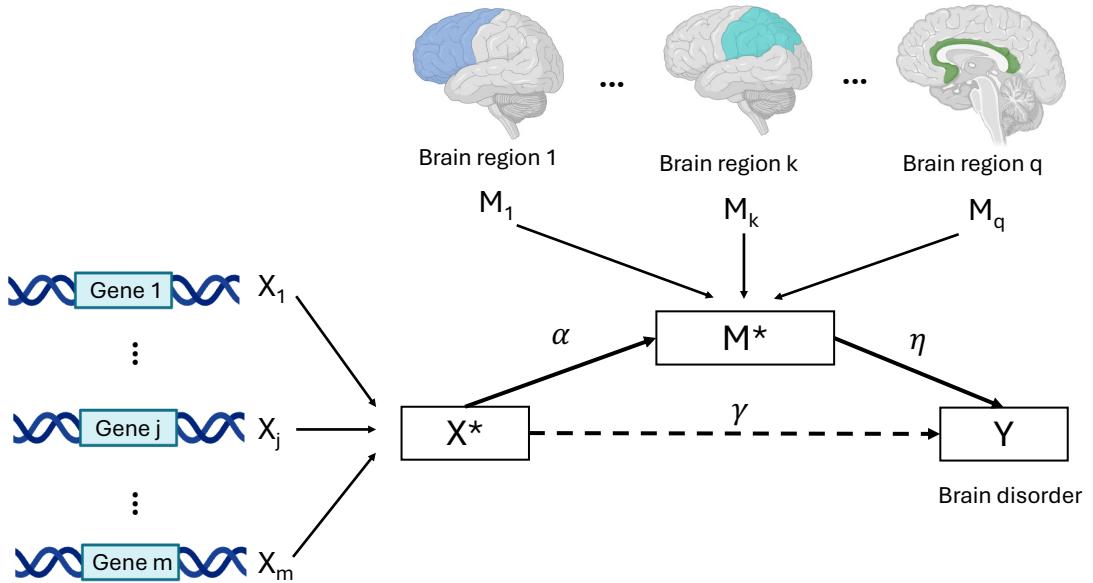


Figure 1: The proposed multi-exposure-to-multi-mediator framework to investigate how multiple genetic risk factors (e.g. gene expression from TWAS) impact brain disorders via changing multiple neuroimaging features (e.g. MRI features from different brain regions). In this model, genes act as exposures, neuroimaging features act as mediators, and brain disorder is the outcome.

2. Method

2.1. A novel multi-exposure-to-multi-mediator model

We consider a multi-exposure-to-multi-mediator model to investigate the gene-neuroimaging-brain disorder mediation pathway (Fig 1). We are specifically interested in knowing how a group of functionally related genetic risk factors impact the brain disorder via changing the function and structure of the brain. Denote by $\mathbf{X} = (X_1, \dots, X_m)$ the $n \times m$ matrix of

exposures of m related genetic risk factors of a brain disorder for n subjects. These genetic risk factors could refer to the genotypes of a set of related SNPs (e.g., in the same linkage disequilibrium (LD) region) associated with the outcome from a GWAS, or the predicted expression levels (also known as genetically regulated expression (GReX) [21]) of a set of genes associated with the outcome from a TWAS. Denote by $\mathbf{M} = (M_1, \dots, M_q)$ the $n \times q$ matrix of q neuroimaging mediators for n subjects, where each M_k ($1 \leq k \leq q$) may represent an imaging-derived phenotype (IDP) such as functional connectivity or cortical thickness measured across q brain regions from structural or functional MRI. Denote by $\mathbf{Y} = (Y_1, \dots, Y_n)^\top$ the vector of phenotypic outcome. Here we assume a linear mediation model so Y needs to be a quantitative measure (e.g. cigarettes per day to capture nicotine dependence in real data example). When the outcome is categorical, it is readily extensible to a logistic or multinomial outcome model. In real data application, we adjust for known covariates (e.g., age, sex, and ancestry principal components) by regressing them out from each column of \mathbf{X} , \mathbf{M} , and \mathbf{Y} , i.e. each is replaced by its residuals from a linear model that includes the covariates. While this adjustment is standard practice, it introduces dependencies among the entries of \mathbf{Y} that must be acknowledged; the residualized outcomes are not strictly independent. Nonetheless, this preprocessing step ensures that subsequent mediation analysis targets variation in \mathbf{Y} , \mathbf{X} , and \mathbf{M} that is not explained by nuisance variables.

Given the large sets of correlated exposures and correlated mediators, we adopt a reduced-rank modeling approach by assuming that the mediation effect is primarily conveyed through a few dominant directions in the exposure and mediator spaces. This assumption is supported by the empirical success of dimension reduction techniques such as principal component analysis (PCA) and canonical correlation analysis (CCA), which often reveal low-dimensional structures capturing most of the variability or association. Specifically, we posit that the mediation effect is concentrated along one principal direction in the exposure space and one in the mediator space. These directions serve as linear combinations that summarize the primary pathway through which the exposures influence the outcome via the mediators [15, 22]. Specifically, we define low-dimensional aggregators:

$$\mathbf{X}^* = \mathbf{X} \mathbf{a}, \quad \mathbf{M}^* = \mathbf{M} \mathbf{b}, \quad (1)$$

where $\mathbf{a} \in \mathbb{R}^m$ and $\mathbf{b} \in \mathbb{R}^q$ are weight vectors to be estimated. In this model, we consider that only one latent variable (e.g. one-dimensional aggregator) exists for exposures and mediators (i.e. $\mathbf{X}^* = (X_1^*, \dots, X_n^*)$ and $\mathbf{M}^* = (M_1^*, \dots, M_n^*)$ for n subjects), respectively. In a real application, there might exist more than one latent variable and we may consider extending our model to a scenario with multiple latent variables. As we will see next, the general idea of capturing the primary mediation pathway still applies. To ensure identifiability, these aggregators are subject to the unit-norm constraints $\|\mathbf{X} \mathbf{a}\|_2 = 1$ and $\|\mathbf{M} \mathbf{b}\|_2 = 1$, which fix scale, removing ambiguity due to rescaling.

We assume a linear mediation structure on the aggregated variables \mathbf{X}^* and \mathbf{M}^* , following is a standard formulation in the mediation literature [23, 24, 25]:

$$\mathbf{M}^* = \alpha \mathbf{X}^* + \boldsymbol{\varepsilon}_M, \quad (2)$$

$$\mathbf{Y} = \gamma \mathbf{X}^* + \eta \mathbf{M}^* + \boldsymbol{\varepsilon}_Y, \quad (3)$$

where $\varepsilon_M \sim \mathcal{N}(0, \sigma_M^2 I_{n \times n})$, $\varepsilon_Y \sim \mathcal{N}(0, \sigma_Y^2 I_{n \times n})$, and $\theta = (\mathbf{a}, \mathbf{b}, \alpha, \gamma, \eta, \sigma_M^2, \sigma_Y^2)$ is the set of unknown parameters. The error terms are assumed to be independent with exposures and mediators, i.e. $\varepsilon_M \perp \mathbf{X}$, $\varepsilon_Y \perp (\mathbf{X}, \mathbf{M})$ (see section 2.3 for details). Without loss of generality, we assume $\alpha > 0$ and $\text{sign}(\eta) = \text{sign}(\gamma)$, ensuring the mediation proportion $\text{MP} = \alpha\eta/\tau$ where $\tau = \gamma + \alpha\eta$ is well defined and bounded by 1.

Given the latent aggregates $\mathbf{X}^* = \mathbf{X}\mathbf{a}$ and $\mathbf{M}^* = \mathbf{M}\mathbf{b}$, the ordinary least squares (OLS) coefficients of α , γ , η and τ are defined as the unique minimizers of $\|\mathbf{M}^* - \alpha \mathbf{X}^*\|_2^2$, $\|\mathbf{Y} - \gamma \mathbf{X}^* - \eta \mathbf{M}^*\|_2^2$ and $\|\mathbf{Y} - \tau \mathbf{X}\mathbf{a}\|_2^2$. Under the constraints $\|\mathbf{X}^*\|_2 = \|\mathbf{M}^*\|_2 = 1$, the normal equations yield

$$\begin{aligned}\hat{\alpha} &= (\mathbf{X}^*)^\top \mathbf{M}^*, \quad \hat{\gamma} = \frac{(\mathbf{X}^*)^\top \mathbf{Y} - \hat{\alpha} (\mathbf{M}^*)^\top \mathbf{Y}}{1 - \hat{\alpha}^2}, \\ \hat{\eta} &= \frac{(\mathbf{M}^*)^\top \mathbf{Y} - \hat{\alpha} (\mathbf{X}^*)^\top \mathbf{Y}}{1 - \hat{\alpha}^2}, \quad \hat{\tau} = (\mathbf{X}^*)^\top \mathbf{Y} = \mathbf{a}^\top \mathbf{X}^\top \mathbf{Y}.\end{aligned}$$

However, the latent aggregates $\mathbf{X}^* = \mathbf{X}\mathbf{a}$ and $\mathbf{M}^* = \mathbf{M}\mathbf{b}$ depend on unknown weight vectors $\mathbf{a} \in \mathbb{R}^m$ and $\mathbf{b} \in \mathbb{R}^q$. Out of the possible weight vectors, we wish to capture the primary gene-neuroimaging-brain disorder mediation pathway, thus targeting at a large mediated effect (via the term involving $\hat{\alpha}\hat{\eta}/\hat{\tau}$) while encouraging sparsity in \mathbf{a}, \mathbf{b} for the best interpretability of the resulting latent aggregates. To balance three competing aims—(i) accurate estimation of the linear mediation model, (ii) a large mediation proportion, and (iii) sparsity for interpretability, we maximize a penalized (Lagrangian) sum of squared error objective with both a penalty on mediation proportions and sparsity penalties on weight vectors. Formally, for observed data $(\mathbf{Y}, \mathbf{X}, \mathbf{M})$ and tuning parameters $\lambda_n, \lambda_a, \lambda_b > 0$, we define the following loss function:

$$\begin{aligned}\mathcal{L}(\mathbf{a}, \mathbf{b}; \mathbf{Y}, \mathbf{X}, \mathbf{M}, \lambda_n, \lambda_a, \lambda_b) &= \frac{1}{2n} \left\{ SSR_{\mathbf{Y} \sim \mathbf{X}^*} + SSR_{\mathbf{M}^* \sim \mathbf{X}^*} + \frac{SSR_{\mathbf{Y} \sim \mathbf{X}^*, \mathbf{M}^*}}{1 - \hat{\alpha}^2} \right\} \\ &\quad - \lambda_n \left(\frac{\hat{\alpha} \hat{\eta}}{\hat{\tau}} \right) + \lambda_a \|\mathbf{a}\|_1 + \lambda_b \|\mathbf{b}\|_1,\end{aligned}\tag{4}$$

where the first three sum of squared residual (SSR) terms correspond to the estimation error from outcome on latent exposure ($SSR_{\mathbf{Y} \sim \mathbf{X}^*} = \|\mathbf{Y} - \hat{\tau} \mathbf{X}\mathbf{a}\|_2^2$), latent mediator on latent exposure ($SSR_{\mathbf{M}^* \sim \mathbf{X}^*} = \|\mathbf{M}\mathbf{b} - \hat{\alpha} \mathbf{X}\mathbf{a}\|_2^2$), and outcome on latent exposure and mediator ($SSR_{\mathbf{Y} \sim \mathbf{X}^*, \mathbf{M}^*} = \|\mathbf{Y} - \hat{\gamma} \mathbf{X}\mathbf{a} - \hat{\eta} \mathbf{M}\mathbf{b}\|_2^2$), respectively. The denominator $1 - \hat{\alpha}^2$ in the third term compensates for the correlation between \mathbf{X}^* and \mathbf{M}^* , partially compensating variance inflation due to multicollinearity. The term involving $\hat{\alpha}\hat{\eta}/\hat{\tau}$ aims to maximize the estimated mediation proportion, while the ℓ_1 penalties encourage sparse selections of manifest exposures and mediators in the latent variables, yielding interpretable models with a small number of active features [26]. This objective can, under certain regularity conditions in Section 2.3, recover the aggregation directions up to scaling and sign.

Remark 1. *The parameters $(\mathbf{a}, \mathbf{b}, \alpha, \gamma, \eta)$ are identifiable only under the assumptions that $\tau = \alpha + \gamma\eta \neq 0$ and that, conditional on the exposures, no mediator direction orthogonal to $\mathbf{M}\mathbf{b}$ is associated with \mathbf{X} (Assumption 4 in Section 2.3). Further justification and a proof of identifiability of the parameters (Prop. 1) can be found in Section 2.3.*

2.2. Optimization using alternating direction method of multipliers (ADMM) algorithm

The objective function in (4) results in a bi-convex optimization problem in \mathbf{a} and \mathbf{b} , we propose to solve using an alternating direction method of multipliers (ADMM) algorithm [27]. Each subproblem is convex and admits a closed-form soft-thresholding update, enabling efficient computation. The method thereby produces interpretable weight vectors that reveal which exposures and mediators contribute most significantly to the mediation pathway. To facilitate efficient updates under ℓ_1 -regularization, we introduce auxiliary variables $\mathbf{z}_a \in \mathbb{R}^m$ and $\mathbf{z}_b \in \mathbb{R}^q$ that decouple the sparsity-inducing penalties from the smooth residual terms. The augmented Lagrangian for this problem introduces scaled dual variables \mathbf{u}_a and \mathbf{u}_b , and includes quadratic penalties for constraint violations:

$$\begin{aligned}\mathcal{L}_\rho(\mathbf{a}, \mathbf{b}, \mathbf{z}_a, \mathbf{z}_b, \mathbf{u}_a, \mathbf{u}_b) = & \frac{1}{2n} \|\mathbf{Y} - \hat{\tau} \mathbf{X}\mathbf{a}\|^2 + \frac{1}{2n} \|\mathbf{Mb} - \hat{\alpha} \mathbf{X}\mathbf{a}\|^2 + \frac{1}{2n} \frac{\|\mathbf{Y} - \hat{\gamma} \mathbf{X}\mathbf{a} - \hat{\eta} \mathbf{Mb}\|^2}{1 - \hat{\alpha}^2} \\ & - \lambda_n \frac{\hat{\alpha}\hat{\eta}}{\hat{\tau}} + \lambda_a \|\mathbf{z}_a\|_1 + \lambda_b \|\mathbf{z}_b\|_1 \\ & + \frac{\rho}{2} \|\mathbf{a} - \mathbf{z}_a + \mathbf{u}_a\|_2^2 + \frac{\rho}{2} \|\mathbf{b} - \mathbf{z}_b + \mathbf{u}_b\|_2^2,\end{aligned}$$

where ρ is the penalty parameter in the augmented Lagrangian that controls the convergence and balance between primal and dual residuals in ADMM algorithm. These quadratic terms enforce the consensus constraints $\mathbf{a} = \mathbf{z}_a$ and $\mathbf{b} = \mathbf{z}_b$ in the scaled-dual ADMM form.

ADMM proceeds by iteratively updating (1) the primal variables (\mathbf{a}, \mathbf{b}) , (2) the auxiliary variables $(\mathbf{z}_a, \mathbf{z}_b)$, and (3) the dual variables $(\mathbf{u}_a, \mathbf{u}_b)$. The scalar quantities $(\tau, \alpha, \gamma, \eta)$ are not updated by ADMM steps; instead, at each iteration, they are profiled in closed form from the current (\mathbf{a}, \mathbf{b}) (see Remark 2). Convergence is assessed via the primal and dual residuals

$$\mathbf{r}^{(k+1)} = \begin{bmatrix} \mathbf{a}^{(k+1)} - \mathbf{z}_a^{(k+1)} \\ \mathbf{b}^{(k+1)} - \mathbf{z}_b^{(k+1)} \end{bmatrix}, \quad \mathbf{s}^{(k+1)} = \rho \begin{bmatrix} \mathbf{z}_a^{(k+1)} - \mathbf{z}_a^{(k)} \\ \mathbf{z}_b^{(k+1)} - \mathbf{z}_b^{(k)} \end{bmatrix},$$

and we terminate when $\|\mathbf{r}^{(k+1)}\|_\infty \leq \varepsilon_{\text{pri}}$ and $\|\mathbf{s}^{(k+1)}\|_\infty \leq \varepsilon_{\text{dual}}$; in practice we also stop when the relative change in the augmented Lagrangian falls below 10^{-6} for 20 consecutive iterations, or after 500 iterations, whichever occurs first. Detailed optimization steps can be found in Algorithm 1.

To mitigate non-convexity, we initialize (\mathbf{a}, \mathbf{b}) using singular vectors of $X^\top Y$ and residualized $M^\top Y$, and perform multiple random restarts. In practice, this helps avoid poor local minima in the bi-convex optimization. The SSR terms in (4) enforce adequate prediction in the outcome model, while the LASSO penalty parameters λ_a and λ_b emphasize sparse parameterization (i.e., the smallest possible number of nonzero a, b entries) and the tuning parameter λ_n emphasizes maximization of the mediation proportion. The best set of tuning parameters was selected via K -fold cross-validation (e.g. $K=5$) that minimizes MSE over a grid of candidate values. Careful tuning is essential to our method (see Remark 4), so we conducted additional sensitivity analysis by scaling the LASSO penalty parameters λ_a and λ_b by a constant factor C_λ relative to the value chosen by the cross-validation and evaluated the impact of this modification on the performance of our method in simulations. (See details in the Simulation section and Table S5).

Algorithm 1 ADMM-based optimization for penalized mediation estimation in our multi-exposure-to-multi-mediator model

Initialize: $\mathbf{a}^{(0)}, \mathbf{b}^{(0)}, \mathbf{z}_a^{(0)} = \mathbf{a}^{(0)}, \mathbf{z}_b^{(0)} = \mathbf{b}^{(0)}, \mathbf{u}_a^{(0)} = \mathbf{0}, \mathbf{u}_b^{(0)} = \mathbf{0}$

1: **for** $k = 0, 1, 2, \dots$ **until convergence do**

2: **Update a by blockwise coordinate descent:**

$$\mathbf{a}^{(k+1)} = \arg \min_{\mathbf{a}} \mathcal{L}_\rho(\mathbf{a}, \mathbf{b}^{(k)}, \mathbf{z}_a^{(k)}, \mathbf{z}_b^{(k)}, \mathbf{u}_a^{(k)}, \mathbf{u}_b^{(k)})$$

3: **Update b by blockwise coordinate descent:**

$$\mathbf{b}^{(k+1)} = \arg \min_{\mathbf{b}} \mathcal{L}_\rho(\mathbf{a}^{(k+1)}, \mathbf{b}, \mathbf{z}_a^{(k)}, \mathbf{z}_b^{(k)}, \mathbf{u}_a^{(k)}, \mathbf{u}_b^{(k)})$$

4: **Update \mathbf{z}_a (soft-thresholding):**

$$\mathbf{z}_a^{(k+1)} = \arg \min_{\mathbf{z}_a} \lambda_a \|\mathbf{z}_a\|_1 + \frac{\rho}{2} \|\mathbf{z}_a - (\mathbf{a}^{(k+1)} + \rho^{-1} \mathbf{u}_a^{(k)})\|_2^2$$

$$\implies (z_a^{(k+1)})_i = S_{\lambda_a/\rho}((a^{(k+1)})_i + \rho^{-1}(u_a^{(k)})_i).$$

5: **Update \mathbf{z}_b (soft-thresholding):**

$$\mathbf{z}_b^{(k+1)} = \arg \min_{\mathbf{z}_b} \lambda_b \|\mathbf{z}_b\|_1 + \frac{\rho}{2} \|\mathbf{z}_b - (\mathbf{b}^{(k+1)} + \rho^{-1} \mathbf{u}_b^{(k)})\|_2^2$$

$$\implies (z_b^{(k+1)})_j = S_{\lambda_b/\rho}((b^{(k+1)})_j + \rho^{-1}(u_b^{(k)})_j).$$

6: **Dual updates:**

$$\mathbf{u}_a^{(k+1)} = \mathbf{u}_a^{(k)} + \rho(\mathbf{a}^{(k+1)} - \mathbf{z}_a^{(k+1)}), \quad \mathbf{u}_b^{(k+1)} = \mathbf{u}_b^{(k)} + \rho(\mathbf{b}^{(k+1)} - \mathbf{z}_b^{(k+1)}).$$

7: **Check convergence:**

$$\|\mathbf{a}^{(k+1)} - \mathbf{z}_a^{(k+1)}\|, \quad \|\mathbf{b}^{(k+1)} - \mathbf{z}_b^{(k+1)}\|, \quad \rho \|\mathbf{z}_a^{(k+1)} - \mathbf{z}_a^{(k)}\|, \quad \rho \|\mathbf{z}_b^{(k+1)} - \mathbf{z}_b^{(k)}\|.$$

where i, j refer to blocks, using notations $(t)_+ := \max\{t, 0\}$ and $\text{sgn}(0) := 0$,

$$S_{\lambda_a/\rho}((a^{(k+1)})_i + \rho^{-1}(u_a^{(k)})_i) = \text{sgn}((a^{(k+1)})_i + \rho^{-1}(u_a^{(k)})_i) \left(\left| (a^{(k+1)})_i + \rho^{-1}(u_a^{(k)})_i \right| - \frac{\lambda_a}{\rho} \right)_+$$

$$S_{\lambda_b/\rho}((b^{(k+1)})_j + \rho^{-1}(u_b^{(k)})_j) = \text{sgn}((b^{(k+1)})_j + \rho^{-1}(u_b^{(k)})_j) \left(\left| (b^{(k+1)})_j + \rho^{-1}(u_b^{(k)})_j \right| - \frac{\lambda_b}{\rho} \right)_+$$

Remark 2. One could in principle apply ADMM to the “natural” joint objective that treats $(\alpha, \gamma, \eta, \tau)$ and (\mathbf{a}, \mathbf{b}) as free variables. In practice this leads to coupled, nonconvex subproblems: the mediation term $-\lambda_n \alpha \eta / \tau$ (and the factor $1/(1 - \alpha^2)$, if used) prevents closed-form proximal steps for \mathbf{a} and \mathbf{b} . Each outer iteration would require an inner constrained solve for $(\alpha, \gamma, \eta, \tau)$, and the \mathbf{a}, \mathbf{b} updates cease to be standard Lasso steps.

By contrast, profiling the least-squares part yields explicit maps $(\mathbf{a}, \mathbf{b}) \mapsto (\hat{\alpha}, \hat{\gamma}, \hat{\eta}, \hat{\tau})$, after which each block update reduces to a smooth quadratic plus an ℓ_1 penalty. This admits simple two-block ADMM with soft-thresholding, fewer hyperparameters, and better numerical stability (especially when we enforce $\hat{\tau} \geq r_0$ and $|\hat{\alpha}| \leq \rho < 1$). Empirically, this formulation decreases the objective monotonically and converges rapidly.

Remark 3 (Scaling). Algorithm 1 enforces unit length on the aggregates, $\|X\mathbf{a}\|_2 = \|M\mathbf{b}\|_2 = 1$, which is numerically convenient once columns of X, M are standardized. For the asymptotic theory in the Supplement we instead fix the projection weights \mathbf{a}, \mathbf{b} to have unit Euclidean norm, $\|\mathbf{a}\|_2 = \|\mathbf{b}\|_2 = 1$. The entries a_i and b_j in the Algorithm are therefore smaller by respective factors $h_n^{(a)} \equiv \|\mathbf{X}\mathbf{a}\|_2 / \|\mathbf{a}\|_2$ and $h_n^{(b)} \equiv \|\mathbf{M}\mathbf{b}\|_2 / \|\mathbf{b}\|_2$ than those in the Supplement, and these factors $h_n^{(a)}, h_n^{(b)}$ are uniformly asymptotically of order \sqrt{n} under the sub-Gaussian designs with independent data (X_i, M_i, Y_i) assumed in Assumption 4 of Section 2.3. The identifiability proof of Proposition 1 shows explicitly how scaling differences affect the parameters α, γ, η and thereby change the objective function terms and the meaning of the penalty parameters. The statistical consistency theory in the Supplement works with unit vectors \mathbf{a}, \mathbf{b} and parameters α, γ, η that are constants, not affected by sample size n . As n grows, the scaling of the Algorithm (with $\|\mathbf{X}\mathbf{a}\|_2 = \|\mathbf{M}\mathbf{b}\|_2 = 1$) also scales α (when nonzero) up by a factor $h_n^{(a)}$ and $\eta, \gamma/\alpha$ up by a factor $h_n^{(b)}$, leaving the mediation proportion $MP = |\alpha\eta|/(\gamma + \alpha\eta)|$ invariant. The SSR norm-squared terms for $Y \sim X^*$ and $Y \sim M^*, X^*$ in the objective function \mathcal{L}_p therefore retain the same meaning in this Section as in the Supplement, while the second SSR term in \mathcal{L}_p is smaller by a factor of order $1/n$ than the other SSR terms in \mathcal{L}_p and is therefore dropped in the Supplement. Because of the invariance of MP under rescaling, the penalty coefficient λ_n plays the same role in \mathcal{L}_p and in the Supplement. However, the LASSO penalty coefficients λ_a, λ_b which are taken of order $\sqrt{\log n}/\sqrt{n}$ in the consistency results of the Supplement are scaled up by factors $h_n^{(a)}, h_n^{(b)}$ in the numerical calculations with the Algorithm, so that they should be – and are – of order $\sqrt{\log n}$ in the simulations and real-data computations. In simulations, we conducted additional sensitivity analysis to show the impact of scaling λ_a, λ_b on the performance of our method and justified our selection (see details in Simulation section and results in Table S5).

Remark 4 (Estimation paradigm and tuning). Our method follows the penalized-least-squares convention in linear mediation [23]. Alternatively, one may use likelihood-based estimation (MLE/MPLE). We prioritize discovery of a strong mediating pathway via the term involving $\hat{\alpha}\hat{\eta}/\hat{\tau}$, while the ℓ_1 penalties provide interpretability. Overly aggressive penalization with large λ_n can distort (\mathbf{a}, \mathbf{b}) , so careful tuning is essential.

2.3. Convergence Theorem and Consistency

To rigorously justify the theoretical performance of the proposed method, we established convergence guarantees for the ADMM-based optimization algorithm described in

Section 2.2. In addition, in the Supplement, we also showed the statistical consistency of the estimators.

For fixed data matrices $(\mathbf{X}, \mathbf{M}, \mathbf{Y})$ let

$$F_n(\mathbf{a}, \mathbf{b}) := \frac{1}{2n} \left\{ SSR_{\mathbf{Y} \sim \mathbf{X}^*} + SSR_{\mathbf{M}^* \sim \mathbf{X}^*} + \frac{SSR_{\mathbf{Y} \sim \mathbf{X}^*, \mathbf{M}^*}}{1 - \hat{\alpha}(\mathbf{a}, \mathbf{b})^2} \right\} - \lambda_n \frac{\hat{\alpha}(\mathbf{a}, \mathbf{b}) \hat{\eta}(\mathbf{a}, \mathbf{b})}{\hat{\tau}(\mathbf{a})},$$

where $\hat{\tau}, \hat{\alpha}, \hat{\eta}$ are the closed-form expressions in Section 2.1. We allow λ_n to depend on n ; the finite-sample rates below use $\lambda_n = C_n n^{-1/2}$ with $C_n = O(1)$. The penalised objective can thus be defined as:

$$\Phi(\mathbf{a}, \mathbf{b}) := F_n(\mathbf{a}, \mathbf{b}) + \lambda_a \|\mathbf{a}\|_1 + \lambda_b \|\mathbf{b}\|_1.$$

Orders of magnitude under the two normalizations.. With the algorithmic normalization $\|\mathbf{X}\mathbf{a}\|_2 = \|\mathbf{M}\mathbf{b}\|_2 = 1$, the retained SSR terms in $F_n/(2n)$ are $O(1)$ at the optimum, while the l_1 -norm terms scale like $O(\lambda_a \sqrt{s}/\sqrt{n} + \lambda_b \sqrt{s}/\sqrt{n})$. Under the theoretical normalization $\|\mathbf{a}\|_2 = \|\mathbf{b}\|_2 = 1$ used in the Supplement, the l_1 -norm terms are $O(\sqrt{s \log d/n})$ for $\lambda_a, \lambda_b \asymp \sqrt{\log d/n}$, matching the standard high-dimensional regime. Our simulation therefore includes the multiplicative factor C_λ to align the penalty magnitudes with the algorithmic scaling; see also Table S5 for performance as C_λ increases.

Assumption 1 (Near-global solution / oracle dominance). *With probability tending to 1, the penalized objective Φ has a (possibly unique) global minimizer $\theta^* = (\mathbf{a}^*, \mathbf{b}^*)$ in the admissible set, and the algorithm returns $\hat{\theta} = (\hat{\mathbf{a}}, \hat{\mathbf{b}})$ such that*

$$\Phi(\hat{\theta}) - \Phi(\theta^*) = o_p(1) \quad \text{and in particular} \quad \Phi(\hat{\theta}) \leq \Phi(\theta_0) + o_p(1),$$

where $\theta_0 = (a_0, b_0)$ denotes the true aggregation directions.

We emphasize that our ADMM result guarantees convergence to a first-order stationary point; consistency then invokes this assumption to connect the attained objective value to the global optimum. We make the following assumptions to show our convergence theorem:

Assumption 2 (Kurdyka–Łojasiewicz property). *Φ is semi-algebraic, hence satisfies the KL property at every critical point (with possibly sample-dependent neighbourhoods and constants); see [28].*

Assumption 3. *Let $\mathcal{C} \subset \mathbb{R}^m \times \mathbb{R}^q$ denote the feasible set for (\mathbf{a}, \mathbf{b}) (including any scaling/box/sparsity constraints used in Algorithm 1). Fix constants $r_0 > 0$ and $\delta \in (0, 1)$ and define the admissible region*

$$\mathcal{C}_{r_0, \delta} = \left\{ (\mathbf{a}, \mathbf{b}) \in \mathcal{C} : \hat{\tau}(\mathbf{a}) \geq r_0, 1 - \hat{\alpha}(\mathbf{a}, \mathbf{b})^2 \geq \delta \right\},$$

where $\hat{\tau}(\mathbf{a}) = \mathbf{a}^\top \mathbf{X}^\top \mathbf{Y}$ and $\hat{\alpha}(\mathbf{a}, \mathbf{b}) = \mathbf{a}^\top \mathbf{X}^\top \mathbf{M} \mathbf{b}$. Assume Algorithm 1 is run within $\mathcal{C}_{r_0, \delta}$ and that the mediation objective Φ uses a fixed tuning level $\lambda_n \in (0, \infty)$. Then $\nabla_{\mathbf{a}} \Phi$ and $\nabla_{\mathbf{b}} \Phi$ are Lipschitz on $\mathcal{C}_{r_0, \delta}$, and each block subproblem in Algorithm 1 is convex with a unique global minimiser.

Assumption 4 (Data-generating model and underlying aggregation structure). *We assume the following data-generating mechanism for the mediation model involving multiple exposures and mediators:*

Latent aggregators and mediation structure. *There exist unknown sparse weight vectors $\mathbf{a}_0 \in \mathbb{R}^m$ and $\mathbf{b}_0 \in \mathbb{R}^q$ such that the primary linear projections*

$$X^* := \mathbf{X}\mathbf{a}_0, \quad M^* := \mathbf{M}\mathbf{b}_0$$

satisfy the structural mediation model

$$M^* = \alpha_0 X^* + \varepsilon_M, \quad Y = \gamma_0 X^* + \eta_0 M^* + \varepsilon_Y, \quad (5)$$

where $\alpha_0, \gamma_0, \eta_0 \in \mathbb{R}$ are fixed constants such that $\tau_0 = \gamma_0 + \alpha_0 \eta_0 \neq 0$, and $(\varepsilon_M, \varepsilon_Y)$ are independent errors with zero mean and finite variance, and with $E(\varepsilon_M^\top \varepsilon_M) > 0$.

Independence/regularity. *We assume:*

(C1) Sub-Gaussian covariates. *Each row $(\mathbf{X}_i, \mathbf{M}_i) \in \mathbb{R}^{m+q}$ is independent across $i = 1, \dots, n$, mean-zero, and σ -sub-Gaussian:*

$$\mathbb{E}[\exp\{\langle \mathbf{v}, (\mathbf{X}_i, \mathbf{M}_i) \rangle\}] \leq \exp\left\{\frac{\sigma^2}{2} \|\mathbf{v}\|_2^2\right\}, \quad \forall \mathbf{v} \in \mathbb{R}^{m+q}.$$

(C2) Population nonsingularity. *Let $\Sigma_X := \mathbb{E}[\mathbf{X}_i \mathbf{X}_i^\top] \in \mathbb{R}^{m \times m}$, $\Sigma_M := \mathbb{E}[\mathbf{M}_i \mathbf{M}_i^\top] \in \mathbb{R}^{q \times q}$, and $\Sigma_{XM} := \mathbb{E}[\mathbf{X}_i \mathbf{M}_i^\top] \in \mathbb{R}^{m \times q}$. Assume $\Sigma_X \succ 0$, $\Sigma_M \succ 0$, and the joint covariance*

$$\Sigma = \begin{pmatrix} \Sigma_X & \Sigma_{XM} \\ \Sigma_{XM}^\top & \Sigma_M \end{pmatrix}$$

is nonsingular.

(C3) Error independence. *The error terms satisfy (with \perp denoting independence):*

$$\varepsilon_M \perp \mathbf{X}, \quad \varepsilon_Y \perp (\mathbf{X}, \mathbf{M}), \quad \text{and} \quad \varepsilon_M, \varepsilon_Y \text{ are each } \sigma\text{-sub-Gaussian.}$$

(C4) Identifiability condition. *To ensure uniqueness of the projected directions, we assume the conditional expectation $\mathbb{E}[\mathbf{M} | \mathbf{X}]$ is linear:*

$$\mathbb{E}[\mathbf{M} | \mathbf{X}] = \mathbf{X}A^\top,$$

for some matrix $A \in \mathbb{R}^{q \times m}$ satisfying

$$A^\top \mathbf{b}_0 = \alpha_0 \mathbf{a}_0, \quad \text{and} \quad A^\top v = 0 \quad \text{for all } v \perp \mathbf{b}_0.$$

$$\text{Equivalently, } \text{col}(A^\top) = \text{span}\{\mathbf{a}_0\} \quad \text{and} \quad A^\top = (\alpha_0 / \|\mathbf{b}_0\|_2^2) \mathbf{a}_0 \mathbf{b}_0^\top.$$

The latent projections (X^*, M^*) and coefficients $(\alpha_0, \gamma_0, \eta_0)$ are uniquely determined (up to scale-factors on a, b respectively defined by either the pair X^*, M^* or the pair a, b being unit

vectors), by the joint distribution of $(\mathbf{X}, \mathbf{M}, Y)$ under Assumption 4. For this Proposition, we adopt the notation for any vector $\mathbf{v} \neq \mathbf{0}$ that $\hat{\mathbf{v}} = \mathbf{v}/\|\mathbf{v}\|_2$ is the rescaled unit vector.

Proposition 1 (Identifiability). *Let $(\mathbf{X}, \mathbf{M}, Y)$ satisfy Assumption 4, with fixed constants $(\alpha_0, \gamma_0, \eta_0)$ such that $\tau_0 = \gamma_0 + \alpha_0\eta_0 \neq 0$. Then the coefficients $(\alpha_0, \gamma_0, \eta_0)$ and unit vectors $\hat{\mathbf{a}}_0$ and $\hat{\mathbf{b}}_0$ are uniquely determined by $E(\mathbf{X}^\top \mathbf{M})$, $E(\mathbf{X}^\top \mathbf{Y})$, $E(\mathbf{M}^\top \mathbf{Y})$.*

Proof. The proof is slightly different according to whether $\alpha_0 = 0$. First, by Assumption 4(C4), the condition $\alpha_0 = 0$ is equivalent to $E(\mathbf{X}^\top \mathbf{M}) = \mathbf{0}$. If this holds, then the assumption $\tau_0 \neq 0$ implies both $\gamma_0 \neq 0$, $\eta_0 \neq 0$. In this case, $\eta_0 \mathbf{b} = (E(\mathbf{M}^\top \mathbf{M}))^{-1} E(\mathbf{M}^\top \mathbf{Y})$ uniquely identifies $\eta_0 \|\mathbf{b}_0\|_2$ and $\hat{\mathbf{b}}_0$; and then $\gamma_0 \mathbf{a}_0 = (E(\mathbf{X}^\top \mathbf{X}))^{-1} E(\mathbf{X}^\top \mathbf{Y})$ uniquely identifies $\gamma_0 \|\mathbf{a}_0\|_2$ and $\hat{\mathbf{a}}_0$.

Next consider the case where $\alpha_0 \neq 0$. Again we apply the expectation of the structural equations (5) premultiplied by \mathbf{X}^\top or \mathbf{M}^\top . Then $\tau_0 \mathbf{a}_0 = (E(\mathbf{X}^\top \mathbf{X}))^{-1} E(\mathbf{X}^\top \mathbf{Y})$ uniquely identifies $\tau_0 \|\mathbf{a}_0\|_2$ and $\hat{\mathbf{a}}_0$. Also, $(\alpha_0/\|\mathbf{b}_0\|_2^2) \mathbf{a}_0 \mathbf{b}_0^\top = A = (E(\mathbf{X}^\top \mathbf{X}))^{-1} E(\mathbf{X}^\top \mathbf{M})$ uniquely identifies the sum of squares of A^\top entries as $(\alpha_0 \|\mathbf{a}_0\|_2 / \|\mathbf{b}_0\|_2)^2$. Then

$$A^\top \tau_0 \mathbf{a}_0 = \frac{\tau_0}{\alpha_0} \left(\frac{\alpha_0 \|\mathbf{a}_0\|_2}{\|\mathbf{b}_0\|_2} \right)^2 \mathbf{b}_0$$

uniquely identifies $(\tau_0/\alpha_0) \mathbf{b}_0$, $(\tau_0/\alpha_0) \|\mathbf{b}_0\|_2$ and $\hat{\mathbf{b}}_0$. We next write

$$\begin{aligned} \hat{\mathbf{b}}_0^\top E(\mathbf{M}^\top \mathbf{Y}) &= \gamma_0 \hat{\mathbf{b}}_0^\top E(\mathbf{X}^\top \mathbf{M})^\top \mathbf{a}_0 + \eta_0 \hat{\mathbf{b}}_0^\top E(\mathbf{M}^\top \mathbf{M}) \mathbf{b}_0 \\ &= \frac{\gamma_0}{\alpha_0} \|\mathbf{b}_0\|_2 \left(\frac{\alpha_0 \|\mathbf{a}_0\|_2}{\|\mathbf{b}_0\|_2} \right)^2 \hat{\mathbf{a}}_0^\top E(\mathbf{X}^\top \mathbf{X}) \hat{\mathbf{a}}_0 + \eta_0 \hat{\mathbf{b}}_0^\top E(\mathbf{M}^\top \mathbf{M}) \mathbf{b}_0 \end{aligned}$$

so that

$$\begin{aligned} \hat{\mathbf{b}}_0^\top E(\mathbf{M}^\top \mathbf{Y}) - \frac{\tau_0}{\alpha_0} \|\mathbf{b}_0\|_2 \left(\frac{\alpha_0 \|\mathbf{a}_0\|_2}{\|\mathbf{b}_0\|_2} \right)^2 \hat{\mathbf{a}}_0^\top E(\mathbf{X}^\top \mathbf{X}) \hat{\mathbf{a}}_0 \\ = \left\{ \hat{\mathbf{b}}_0^\top E(\mathbf{M}^\top \mathbf{M}) \hat{\mathbf{b}}_0 - \left(\frac{\alpha_0 \|\mathbf{a}_0\|_2}{\|\mathbf{b}_0\|_2} \right)^2 \hat{\mathbf{a}}_0^\top E(\mathbf{X}^\top \mathbf{X}) \hat{\mathbf{a}}_0 \right\} \eta_0 \|\mathbf{b}_0\|_2 \end{aligned}$$

The left-hand side of this last expression has already been shown to be identified, as has the curly-bracketed expression on the right-hand side. Therefore, the right-hand side uniquely identifies $\eta_0 \|\mathbf{b}_0\|_2$, as long as we can show the curly-bracketed expression on the right-hand side is a nonzero number. But the curly-bracketed number times $\|\mathbf{b}_0\|_2^2$ equals

$$\mathbf{b}_0^\top E(\mathbf{M}^\top \mathbf{M}) \mathbf{b}_0 - \alpha_0^2 \mathbf{a}_0^\top E(\mathbf{X}^\top \mathbf{X}) \mathbf{a}_0 = E((\mathbf{M}\mathbf{b}_0 - \alpha_0 \mathbf{X}\mathbf{a}_0)^\top (\mathbf{M}\mathbf{b}_0 - \alpha_0 \mathbf{X}\mathbf{a}_0)) = E(\epsilon_M^\top \epsilon_M) > 0$$

The Proposition has been proved. More precisely, taking general scaling into account, the proof shows that when $E(\mathbf{X}^\top \mathbf{M}) \neq \mathbf{0}$ the expectations $E(\mathbf{X}^\top \mathbf{M})$, $E(\mathbf{X}^\top \mathbf{Y})$, $E(\mathbf{M}^\top \mathbf{Y})$ jointly identify $\hat{\mathbf{a}}_0$, $\hat{\mathbf{b}}_0$, $\tau_0 \|\mathbf{a}_0\|_2$, $\eta_0 \|\mathbf{b}_0\|_2$ and $(\gamma_0/\alpha_0) \|\mathbf{b}_0\|_2$; and when $E(\mathbf{X}^\top \mathbf{M}) = \mathbf{0}$, they identify $\alpha_0 = 0$ along with $\hat{\mathbf{a}}_0$, $\hat{\mathbf{b}}_0$, $\gamma_0 \|\mathbf{a}_0\|_2$ and $\eta_0 \|\mathbf{b}_0\|_2$. \square

Theorem 1 (Global sublinear and local linear convergence). *Suppose Assumptions 2 and 3 hold. Let $\{(\mathbf{a}^k, \mathbf{b}^k)\}_{k \geq 0}$ be the ADMM iterates from Algorithm 1. Then:*

- (i) **Cluster-point stationarity and isolation.** *The sequence is bounded; every cluster point is a first-order stationary point of Φ on $\mathcal{C}_{r_0, \delta}$, and cluster points are isolated.*
- (ii) **Global sublinear rate.** *There exists $C > 0$ such that for all $K \geq 1$,*

$$\min_{0 \leq t \leq K} \|\nabla_{\mathbf{a}} \Phi(\mathbf{a}^t, \mathbf{b}^t)\|_2^2 + \|\nabla_{\mathbf{b}} \Phi(\mathbf{a}^t, \mathbf{b}^t)\|_2^2 \leq \frac{C}{K},$$

and the primal/dual residuals are square-summable.

- (iii) **Local linear rate.** *If, in a neighbourhood of a cluster point $(\mathbf{a}^*, \mathbf{b}^*)$, either block map is μ -strongly convex (with the other block fixed), then there exists $\theta \in (0, 1)$ and a neighbourhood \mathcal{N} such that*

$$\Phi(\mathbf{a}^{k+1}, \mathbf{b}^{k+1}) - \Phi(\mathbf{a}^*, \mathbf{b}^*) \leq \theta [\Phi(\mathbf{a}^k, \mathbf{b}^k) - \Phi(\mathbf{a}^*, \mathbf{b}^*)], \quad \text{whenever } (\mathbf{a}^k, \mathbf{b}^k) \in \mathcal{N}.$$

Proof of Theorem 1. See Appendix A. □

In addition to establishing the convergence behavior and local optimality diagnostics for the optimization algorithm, we also showed the statistical consistency of the estimators. Detailed assumptions, proof of the main theorem (Theorem S1) and relevant lemmas can be found in the Supplement.

Remark 5 (Numerical optimization strategy and convergence in practice). *Theorem 1 guarantees convergence only to a first-order stationary point of the non-convex objective Φ . To mitigate suboptimal convergence in practice, we employ two safeguards: (i) a local-minimum check—after convergence, compute the Hessian of Φ restricted to the constraint tangent space at $(\hat{\mathbf{a}}, \hat{\mathbf{b}})$; if the smallest eigenvalue is negative, deem the point a strict saddle and discard the run; (ii) a multi-start screening—run from R random initializations (typically $R = 10$) and retain the admissible solution with the lowest objective. In our numerical experiments, this strategy returned the lowest (or numerically indistinguishable) objective value in $> 95\%$ of data sets.*

3. Simulation

3.1. Simulation settings

In this section, we evaluated the performance of our method across scenarios involving multiple exposures, multiple mediators, and an outcome of interest. We first simulated the data of multiple exposures $\mathbf{X}_{n \times m}$ (e.g. gene expression) for n subjects, where each subject's exposure vector $X_i = (x_{i1}, x_{i2}, \dots, x_{im})$ was independently drawn from a multivariate normal distribution $\mathcal{N}_m(\mathbf{0}, \Sigma_X)$, with $\Sigma_X = (1 - \rho_X)\mathbf{I}_m + \rho_X \mathbf{1}_m \mathbf{1}_m^\top$, where \mathbf{I}_m is the $m \times m$ identity matrix, $\mathbf{1}_m$ is an m -element vector of one's, and ρ_X controls the correlation among the m genes. Next, we generated the data of multiple mediators $\mathbf{M}_{n \times q}$ (e.g. regional neuroimaging

features), where each subject's mediator vector $M_i = (m_{i1}, m_{i2}, \dots, m_{iq})^\top$ was drawn from a multivariate normal distribution $\mathcal{N}_q(L^T X_i, \Sigma_M)$, where $L_{m \times q}$ is the fixed coefficient matrix with $L_{j,k} = -0.5$ for $j = 1, \dots, m$ and $k = 1, \dots, 5$, and zero otherwise, $\Sigma_M = (1 - \rho_M) \mathbf{I}_q + \rho_M \mathbf{1}_q \mathbf{1}_q^\top$, where ρ_M introduces correlation among q mediators. We then aggregated the exposures and mediators $X_i^* = \mathbf{a}^T X_i$ and $M_i^* = \mathbf{b}^T M_i$ for $i = 1, 2, \dots, n$, where the true weight vectors are sparse with only the first five elements being non-zero: i.e. $\mathbf{a}_{m \times 1} = (c, c, c, c, c, 0, \dots, 0)^T$ and $\mathbf{b}_{q \times 1} = (c, c, c, c, c, 0, \dots, 0)^T$. Finally, we simulated the outcome \mathbf{Y} as $Y_i = \gamma X_i^* + \eta M_i^* + \varepsilon_Y$, for $i = 1, 2, \dots, n$, where $\varepsilon_Y \sim \mathcal{N}(0, \sigma_Y^2)$.

We set $n = 200$, $m, q \in \{20, 50\}$, $\rho_X, \rho_M \in \{0, 0.3\}$, noise variances $\sigma_Y^2 \in (1, 3)$ and the weights $c = 0.5$, to evaluate our method. We considered both complete mediation ($\gamma = 0$ and $MP = 1$) and partial mediation ($\gamma = \alpha\eta = 0.5$ and $MP = 0.5$) cases by varying the parameters γ and η . Each scenario was repeated for 1000 times. For each simulation, we selected the best set of tuning parameters $(\lambda_a, \lambda_b, \lambda_n)$ by five-fold cross-validation over a grid of candidate values ranging from 0.02 to 0.5. To evaluate the selection performance of manifest exposures and mediators in the latent variables, i.e. identifying the true nonzero and zero weights in $\hat{\mathbf{a}}$ and $\hat{\mathbf{b}}$, we adopted standard classification metrics: $\text{Precision} = \frac{TP}{TP+FP}$; $\text{Recall} = \frac{TP}{TP+FN}$; $\text{F1-score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}$ and $\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}$, where TP is the number of true nonzero weights in $\hat{\mathbf{a}}$ or $\hat{\mathbf{b}}$, FP is the number of false nonzero weights, TN is the number of true zero weights and FN is the number of false zero weights. To evaluate the estimation of mediation proportion, we reported the empirical absolute bias and standard deviation (SD) of the estimated mediation proportion across simulation replicates.

To benchmark, we compared the proposed multi-exposure-to-multi-mediator (MEMM) method against seven competitor approaches from three different classes:

1. Regularization-based methods: the screening-based two-stage procedure with MCP penalty (SIS-MCP; [29]), the Pathway Lasso estimator (Path and its two-stage Lasso special case TS-Path; [30]), with default parameters $\phi = 2$ and $\psi = 0$.
2. Dimension reduction-based methods: the directions-of-mediation approaches with one and two directions (DM1 and DM2; [31]) and the sparse PCA mediation algorithm (SPCMA; [32]).
3. Hybrid dimension-reduction method targeting the mediation pathway (MMP; 17).

In addition to the above parameter settings, to further assess the robustness of our method, we also performed a series of sensitivity analyses with: (i) misspecified \mathbf{a} or \mathbf{b} ($\vec{a}_{mis} = \vec{0}$ or $\vec{b}_{mis} = \vec{0}$), (ii) smaller mediation proportion $MP = 0.25$ and (iii) weaker signal strength $c = 0.1$. Lastly, as a sensitivity check and to ensure consistency with our theoretical guidance (Theorem S1), we also scaled the LASSO penalty parameters by a constant factor C_λ and evaluated its impact on the performance of our method. These results were summarized in Supplementary Tables S2–S5.

3.2. Simulation Results

Across all simulation scenarios, the proposed MEMM method consistently achieves the lowest absolute bias in mediation proportion (MP) estimation and exhibits the best or near-best weight selection performance (Table 1, 2). Under complete mediation, MEMM accurately recovers the true MP with negligible bias, especially when both exposure and mediator

Table 1: Simulation results for *complete* mediation (MP=1) with $\sigma_Y^2 = 1$, varying m , q , $\rho_{\mathbf{X}}$ and $\rho_{\mathbf{M}}$.

m, q	$\rho_{\mathbf{X}}$	$\rho_{\mathbf{M}}$	Method	MP (SD)	Abs. bias	Precision	Recall	F1	Accuracy	
m=20, q=20	0.0	0.0	MEMM	0.8746 (0.0505)	0.1254	0.5885	1.0000	0.7360	0.8130	
			SIS-MCP	0.2699 (0.0429)	0.7330	0.9088	0.9572	0.9253	0.9588	
			Path	0.1079 (0.0168)	0.8900	1.0000	0.4317	0.6013	0.8581	
			TS-Path	0.2112 (0.0307)	0.7900	0.7468	0.6195	0.6732	0.8486	
			DM1	0.8295 (0.0739)	0.1630	0.3034	1.0000	0.4650	0.4206	
			DM2	0.3194 (0.0217)	0.6780	0.2488	0.3177	0.2787	0.5895	
			SPCMA	0.8555 (0.0828)	0.1310	0.2944	1.0000	0.4545	0.3945	
			MMP	0.3147 (0.0468)	0.6910	0.4086	0.5184	0.4547	0.6944	
	0.3	0.0	MEMM	0.9371 (0.0487)	0.0629	0.8660	1.0000	0.9227	0.9540	
			SIS-MCP	0.4008 (0.0930)	0.6180	0.5991	0.9127	0.7096	0.8056	
			Path	0.2500 (0.0001)	0.7500	1.0000	0.4263	0.6173	0.8463	
			TS-Path	0.2990 (0.0304)	0.7070	0.8485	1.0000	0.9148	0.9512	
			DM1	0.8550 (0.0340)	0.1520	0.2929	1.0000	0.4529	0.3950	
			DM2	0.3186 (0.0228)	0.6840	0.2472	0.3149	0.2767	0.5890	
			SPCMA	0.8765 (0.0760)	0.1410	0.2872	1.0000	0.4459	0.3735	
			MMP	0.2509 (0.0055)	0.7490	0.9873	0.9904	0.9887	0.9945	
	0.0	0.3	MEMM	0.9486 (0.0375)	0.0514	0.8685	1.0000	0.9239	0.9545	
			SIS-MCP	0.3215 (0.0894)	0.6820	0.7605	0.9282	0.8220	0.8926	
			Path	0.1080 (0.0157)	0.8920	1.0000	0.4321	0.6012	0.8582	
			TS-Path	0.2070 (0.0298)	0.7940	0.7684	0.6279	0.6876	0.8571	
			DM1	0.8310 (0.0716)	0.1720	0.3024	1.0000	0.4639	0.4189	
			DM2	0.3194 (0.0247)	0.6840	0.2442	0.3119	0.2735	0.5868	
			SPCMA	0.8530 (0.0788)	0.1500	0.2949	1.0000	0.4553	0.3970	
			MMP	0.3133 (0.0418)	0.6820	0.4153	0.5221	0.4610	0.6978	
	0.3	0.3	MEMM	0.9679 (0.0338)	0.0321	0.9867	1.0000	0.9927	0.9960	
			SIS-MCP	0.3976 (0.0838)	0.5860	0.5643	0.8400	0.6536	0.7724	
			Path	0.2500 (0.0001)	0.7500	1.0000	0.4523	0.6134	0.8321	
			TS-Path	0.2916 (0.0274)	0.7090	0.8657	1.0000	0.9253	0.9581	
			DM1	0.8493 (0.0317)	0.1510	0.2949	1.0000	0.4552	0.4008	
			DM2	0.3162 (0.0241)	0.6830	0.2540	0.3208	0.2831	0.5944	
			SPCMA	0.8585 (0.0774)	0.1370	0.2935	1.0000	0.4534	0.3915	
			MMP	0.2523 (0.0061)	0.7477	0.9878	0.9969	0.9922	0.9963	
	m=50, q=50	0.0	0.0	MEMM	0.9236 (0.0356)	0.0764	0.7124	0.9725	0.8193	0.8954
				SIS-MCP	0.1960 (0.0264)	0.8040	1.0000	0.8170	0.8990	0.8040
				Path	0.0509 (0.0164)	0.9491	1.0000	0.2120	0.3500	0.9490
				TS-Path	0.0854 (0.0223)	0.9146	0.8710	0.3090	0.4550	0.9150
				DM1	0.7240 (0.0495)	0.2760	0.3330	1.0000	0.4990	0.2760
				DM2	0.3170 (0.0132)	0.6830	0.2400	0.3170	0.2730	0.6830
				SPCMA	0.7420 (0.0534)	0.2580	0.3260	1.0000	0.4910	0.2580
				MMP	0.1660 (0.0263)	0.8340	0.0016	0.0012	0.0014	0.8340
	0.3	0.0	MEMM	0.9110 (0.0326)	0.0890	1.0000	1.0000	1.0000	1.0000	
			SIS-MCP	0.1910 (0.0634)	0.8090	1.0000	0.7960	0.8850	0.8090	
			Path	0.0497 (0.0008)	0.9503	1.0000	0.2070	0.3430	0.9500	
			TS-Path	0.0861 (0.0231)	0.9139	0.8570	0.3050	0.4490	0.9140	
			DM1	0.7180 (0.0246)	0.2820	0.3370	1.0000	0.5030	0.2820	
			DM2	0.3170 (0.0153)	0.6830	0.2410	0.3180	0.2740	0.6830	
			SPCMA	0.7380 (0.0634)	0.2620	0.3280	1.0000	0.4930	0.2620	
			MMP	0.1600 (0.0275)	0.8400	0.0008	0.0005	0.0006	0.8400	
	0.0	0.3	MEMM	0.9296 (0.0265)	0.0704	0.9984	0.9983	0.9983	0.9992	
			SIS-MCP	0.1440 (0.0532)	0.8560	1.0000	0.6000	0.7440	0.8560	
			Path	0.2400 (0.0070)	0.7600	1.0000	1.0000	1.0000	0.7600	
			TS-Path	0.1960 (0.0172)	0.8040	1.0000	0.8160	0.8990	0.8040	
			DM1	0.7590 (0.0312)	0.2410	0.3160	1.0000	0.4810	0.2410	
			DM2	0.3170 (0.0165)	0.6830	0.2410	0.3180	0.2740	0.6830	
			SPCMA	0.7390 (0.0385)	0.2610	0.3270	1.0000	0.4920	0.2610	
			MMP	0.0119 (0.0262)	0.9881	0.0000	0.0000	0.0000	0.9880	
	0.3	0.3	MEMM	0.9872 (0.0204)	0.0128	1.0000	1.0000	1.0000	1.0000	
			SIS-MCP	0.1150 (0.0701)	0.8850	0.9980	0.4750	0.6380	0.8850	
			Path	0.2400 (0.0063)	0.7600	1.0000	1.0000	1.0000	0.7600	
			TS-Path	0.1880 (0.0154)	0.8120	1.0000	0.7810	0.8770	0.8120	
			DM1	0.7550 (0.0223)	0.2450	0.3180	1.0000	0.4830	0.2450	
			DM2	0.3160 (0.0123)	0.6840	0.2390	0.3150	0.2720	0.6840	
			SPCMA	0.7390 (0.0462)	0.2610	0.3270	1.0000	0.4930	0.2610	
			MMP	0.0114 (0.0343)	0.9886	0.0000	0.0000	0.0000	0.9890	

Table 2: Simulation results for *partial* mediation (MP=0.5) with $\sigma_Y^2 = 1$, varying m , q , ρ_X and ρ_M .

m, q	ρ_X	ρ_M	Method	MP (SD)	Abs. bias	Precision	Recall	F1	Accuracy
m=20, q=20	0.0	0.0	MEMM	0.3089 (0.0610)	0.1911	0.5028	0.9880	0.6626	0.7405
			SIS-MCP	0.2580 (0.0238)	0.2420	0.9690	0.9950	0.9810	0.2420
			Path	0.1120 (0.0153)	0.3880	1.0000	0.4490	0.6170	0.3880
			TS-Path	0.2110 (0.0287)	0.2890	0.7600	0.6330	0.6880	0.2890
			DM1	0.8330 (0.0214)	0.3330	0.3020	1.0000	0.4630	0.3330
			DM2	0.3180 (0.0401)	0.1820	0.2490	0.3170	0.2790	0.1820
			SPCMA	0.8620 (0.0763)	0.3620	0.2920	1.0000	0.4520	0.3620
			MMP	0.3080 (0.0639)	0.1920	0.4140	0.5140	0.4570	0.1920
0.3	0.0	0.0	MEMM	0.4427 (0.0522)	0.0573	0.7939	1.0000	0.8791	0.9260
			SIS-MCP	0.3400 (0.0959)	0.1600	0.7440	0.9460	0.8160	0.1600
			Path	0.2500 (0.0000)	0.2500	1.0000	1.0000	1.0000	0.2500
			TS-Path	0.2800 (0.0273)	0.2200	0.9010	1.0000	0.9460	0.2200
			DM1	0.8510 (0.0233)	0.3510	0.2940	1.0000	0.4540	0.3510
			DM2	0.3130 (0.0063)	0.1870	0.2470	0.3090	0.2740	0.1870
			SPCMA	0.8610 (0.0771)	0.3610	0.2930	1.0000	0.4520	0.3610
			MMP	0.2500 (0.0321)	0.2500	0.9900	0.9910	0.9910	0.2500
0.0	0.3	0.0	MEMM	0.4159 (0.0856)	0.0841	0.7636	1.0000	0.8586	0.9105
			SIS-MCP	0.3130 (0.0745)	0.1870	0.8070	0.9680	0.8690	0.1870
			Path	0.1110 (0.0157)	0.3890	1.0000	0.4460	0.6140	0.3890
			TS-Path	0.2130 (0.0298)	0.2870	0.7540	0.6340	0.6850	0.2870
			DM1	0.8350 (0.0225)	0.3350	0.3010	1.0000	0.4620	0.3350
			DM2	0.3160 (0.0466)	0.1840	0.2510	0.3170	0.2800	0.1840
			SPCMA	0.8620 (0.0755)	0.3620	0.2920	1.0000	0.4520	0.3620
			MMP	0.3200 (0.0570)	0.1800	0.4180	0.5380	0.4680	0.1800
0.3	0.3	0.0	MEMM	0.4933 (0.0361)	0.0067	0.9493	1.0000	0.9721	0.9845
			SIS-MCP	0.4260 (0.0800)	0.0740	0.5630	0.9310	0.6920	0.0745
			Path	0.2500 (0.0001)	0.2500	1.0000	1.0000	1.0000	0.2500
			TS-Path	0.2870 (0.0274)	0.2130	0.8790	1.0000	0.9330	0.2130
			DM1	0.8420 (0.0235)	0.3420	0.2970	1.0000	0.4580	0.3420
			DM2	0.3170 (0.0053)	0.1830	0.2520	0.3190	0.2810	0.1830
			SPCMA	0.8380 (0.0759)	0.3380	0.3010	1.0000	0.4620	0.3380
			MMP	0.2530 (0.0385)	0.2470	0.9840	0.9950	0.9900	0.2470
m=50, q=50	0.0	0.0	MEMM	0.3691 (0.0425)	0.1309	0.6299	0.9625	0.7579	0.8496
			SIS-MCP	0.1980 (0.0204)	0.3020	1.0000	0.8260	0.9050	0.3020
			Path	0.0506 (0.0134)	0.4494	1.0000	0.2110	0.3480	0.4490
			TS-Path	0.0869 (0.0233)	0.4131	0.8500	0.3070	0.4500	0.4130
			DM1	0.7370 (0.0204)	0.2370	0.3270	1.0000	0.4930	0.2370
			DM2	0.3180 (0.0321)	0.1820	0.2420	0.3200	0.2750	0.1820
			SPCMA	0.7560 (0.0620)	0.2560	0.3190	1.0000	0.4830	0.2560
			MMP	0.1640 (0.0570)	0.3360	0.0015	0.0011	0.0013	0.3360
0.3	0.0	0.0	MEMM	0.4591 (0.0456)	0.0409	0.9962	1.0000	0.9980	0.9990
			SIS-MCP	0.1700 (0.0564)	0.3300	1.0000	0.7090	0.8260	0.3300
			Path	0.2400 (0.0001)	0.2600	1.0000	1.0000	1.0000	0.2600
			TS-Path	0.1990 (0.0215)	0.3010	1.0000	0.8270	0.9050	0.3010
			DM1	0.7500 (0.0218)	0.2500	0.3200	1.0000	0.4850	0.2500
			DM2	0.3170 (0.0329)	0.1830	0.2410	0.3180	0.2740	0.1830
			SPCMA	0.7330 (0.0661)	0.2670	0.3300	1.0000	0.4950	0.2330
			MMP	0.0126 (0.0302)	0.4874	0.0000	0.0000	0.0000	0.4870
0.0	0.3	0.0	MEMM	0.4586 (0.0447)	0.0414	0.9954	1.0000	0.9976	0.9988
			SIS-MCP	0.1950 (0.0533)	0.3050	1.0000	0.8120	0.8960	0.3050
			Path	0.0502 (0.0127)	0.4498	1.0000	0.2090	0.3450	0.4500
			TS-Path	0.0869 (0.0285)	0.4131	0.8650	0.3120	0.4580	0.4130
			DM1	0.7310 (0.0216)	0.2310	0.3300	1.0000	0.4960	0.2310
			DM2	0.3170 (0.0419)	0.1830	0.2380	0.3140	0.2710	0.1830
			SPCMA	0.7490 (0.0642)	0.2510	0.3230	1.0000	0.4870	0.2490
			MMP	0.1600 (0.0494)	0.3400	0.0005	0.0003	0.0004	0.3400
0.3	0.3	0.0	MEMM	0.4950 (0.0341)	0.0050	1.0000	1.0000	1.0000	1.0000
			SIS-MCP	0.1350 (0.0463)	0.3650	1.0000	0.5640	0.7170	0.3650
			Path	0.2400 (0.0001)	0.2600	1.0000	1.0000	1.0000	0.2600
			TS-Path	0.1950 (0.0209)	0.3050	1.0000	0.8130	0.8970	0.3050
			DM1	0.7590 (0.0194)	0.2410	0.3170	1.0000	0.4810	0.2590
			DM2	0.3170 (0.0358)	0.1830	0.2400	0.3170	0.2730	0.1830
			SPCMA	0.7430 (0.0497)	0.2430	0.3250	1.0000	0.4900	0.2430
			MMP	0.0128 (0.0420)	0.4872	0.0000	0.0000	0.0000	0.4870

correlations are high ($\rho_X, \rho_M = 0.3$), suggesting that the aggregation structure is particularly effective in the presence of correlated features. The advantage of MEMM remained even with larger number of exposures and mediators ($m = 50, q = 50$). In contrast, regularization based methods (SIS-MCP, Path and TS-Path) accurately selected the weights but persistently displayed large bias in MP estimation with even worse performance when the correlation and number of features increase. The dimension reduction methods (DM1/DM2 and SPCMA), on the other hand, had smaller bias for MP estimation but poorer performance in weight selection (e.g. selecting many false positive weights thus low precision), making it harder to interpret the final aggregators. The hybrid method MMP, originally designed for one exposure only, performed poorly in this setting especially when the number of features and between-feature correlation increased. Our method MEMM well balanced the accurate estimation of MP that captures the predominant mediation pathway and accurate selection of weights for the best interpretability of the aggregated exposures and mediators, well suited to the multi-exposure-to-multi-mediator problem.

Our advantage over other competing methods remained with partial mediation (MP=0.5). Specifically, MEMM maintained a low absolute bias in MP estimation for partial mediation without over-estimating MP, showing the strength of our objective function that balanced the likelihood, mediation proportion and weight estimation. By contrast, DM1/DM2 and SPCMA tend to over-estimate the mediation proportion. The weight selection performance also dropped dramatically for the regularization based methods with partial mediation. As the noise level increased ($\sigma_Y^2 = 3$; Table S1), all methods had worsened performance but MEMM remained as the best performer especially when the between-feature correlation increased.

When a was misspecified, the MP estimation and the overall selection of true zero and nonzero weights remained accurate (Table S2). Misspecification of b had a stronger negative impact on the MP estimation, though the precision of weight selection by our method remained high, and the overall accuracy of selection increased as the correlation increased (Table S2). When the true MP was smaller (MP=0.25), or the signal strength was weakened ($c = 0.1$), our method performed consistently well with close to true MP estimation and high accuracy in selecting the true zero and nonzero weights (Table S3-S4). Table S5 examines the impact of scaling up the LASSO penalty parameters λ_a and λ_b on the performance of our method under complete mediation (MP=1) when $n=200$, $m=q=20$ and $(\rho_X, \rho_M) = (0.3, 0.3)$. The new set of penalty parameters $\lambda_a = C_\lambda \lambda_a^{\text{base}}$, $\lambda_b = C_\lambda \lambda_b^{\text{base}}$, where $\lambda_a^{\text{base}} = \lambda_b^{\text{base}} = 0.15$ falls in the original range of grid search, $C_\lambda \in \{1, 2, 5, 10, 15\}$ is the scaling factor, λ_n is fixed to be 0.1. The MP estimation and the selection of true zero and nonzero weights remained highly accurate with modest penalization ($C_\lambda = 1, 2, 5$, which is of order $\sqrt{\log n}$; Table S5). With too strong penalization ($C_\lambda = 10, 15$), the model is subject to over-shrinkage so the recall and F1 score dropped rapidly.

4. Application to Real Data

Nicotine dependence is a multifaceted brain disorder with a significant heritable component, and a leading modifiable risk factor for cancer, pulmonary, cardiovascular and cerebrovascular disorders [33]. Large-scale GWAS and TWAS on different measures of tobacco smoking behaviors [34, 35, 36, 4] have reported multiple replicable genetic loci and clusters

of genes potentially associated with nicotine dependence. In the same time, neuroimaging studies using magnetic resonance imaging (MRI) data have revealed alterations in both gray and white matter structural and functional characteristics associated with nicotine addiction [37, 38], yet the pathways from genes to brain alternations to nicotine addiction remained unclear. In this section, we applied our MEMM method that integrates genetically regulated expression (GReX), resting-state functional MRI (rs-fMRI), and smoking behavior data from UK Biobank (UKB) to evaluate the potential “gene-brain functional connectivity-nicotine addiction” mediation pathway, exploring how genetic predispositions influence brain circuits, which in turn make people more addicted to nicotine.

As in most other studies, we treated cigarettes per day (CPD) as the main outcome of interest to assess nicotine dependence, and also considered the former and current smoking status in selecting the samples for analysis following the validated diagnostic criteria for nicotine dependence [39]. We first obtained the GReX data from the genotype data in UKB and the GTEx [40] reference panel data in three addiction-relevant brain tissues—Nucleus Accumbens (NAcc), Caudate (Ca), and Putamen (Pu), all subregions of the basal ganglia. Following our previous works of TWAS on cigarettes per day (CPD) using MultiXcan [41, 4, 42], we selected a group of $m = 21$ genes most significantly associated with CPD on chromosome 15 ($p < 1e - 8$; see Fig S1 for the TWAS Manhattan plot) and used their GReX values aggregated over three brain tissues as exposures in our study. Region of Interest (ROI)-to-ROI functional connectivity (FC) features were computed from preprocessed rs-fMRI data. We restricted the analysis to basal ganglia and subcortical lobe regions most relevant to nicotine addiction and kept $q = 635$ ROI-to-ROI FC measures as potential mediators for our analysis (Fig S2). Our final analytical sample includes $n = 710$ participants with complete exposure, mediator and outcome data. As we can see, many genes and many FC measures in these samples are highly correlated with each other (Fig 2A and 2B), laying out the basis for our reduced rank based modeling of both exposures and mediators.

We applied a five-fold cross-validation over a coarse-to-fine grid to select the best λ_a , λ_b and λ_n , where $\lambda_a, \lambda_b \in \{0.02, 0.05, 0.10, 0.20, 0.30, 0.50\}$, $\lambda_n \in \{0.02, 0.05, 0.08, 0.10, 0.15\}$. The final set of penalty parameters selected was $(\lambda_a, \lambda_b, \lambda_n) = (0.30, 0.10, 0.08)$. The ADMM algorithm converged in 9 iterations, and the full run (including cross-validation) completed in about 20 minutes. The estimated MP is approximately 0.40, indicating that roughly half of the genetic effects on CPD on chromosome 15 was mediated through the brain FC between ROIs retained in the model. This result suggested that the functional connectome between specific brain regions, the brain’s dynamic wiring diagram, may serve as a key conduit translating genetic influences into brain disorders like nicotine addiction. At the same time, the non-mediated portion of the effect suggested alternative biological pathways (e.g., hormonal, metabolic, or peripheral nervous system routes). For the exposures, we identified a coherent subset of six genes with nonzero weights (Table 3). All the three genes in the *CHRNA5-CHRNA3-CHRN4* nicotinic receptor subunit gene cluster [43] were selected, affirming their role in nicotine response and smoking behavior. These genes encoded subunits of nicotinic acetylcholine receptors (nAChRs), and their altered expression is known to influence nicotine intake and receptor sensitivity [44]. *PSMA4* and *CTSH*, both encoded for protease involved in protein degradation pathways, may influence brain signaling pathways, neuroadaptations and neuroinflammation linked to nicotine addiction [45, 46].

For the mediators, we identified a sparse subset of 32 FC measures spanning several

addiction-relevant regions (Table S6). Among these were functional connectivity between middle frontal gyrus (MFG), precentral gyrus (PrG), thalamus and superior parietal lobule (SPL) regions. MFG has played a significant role in inhibitory control and reward processing of nicotine [47]. The functional connectivity between SPL and the other brain regions (e.g. the insula for reward processing and craving) has been found correlated with the severity of nicotine dependence [48], and the thalamus is enriched in nicotinic receptors [49]. Our findings suggest that genetic variation affecting expression of nicotinic receptors may exert influence on CPD via alterations in MFG, SPL and thalamus subregions.

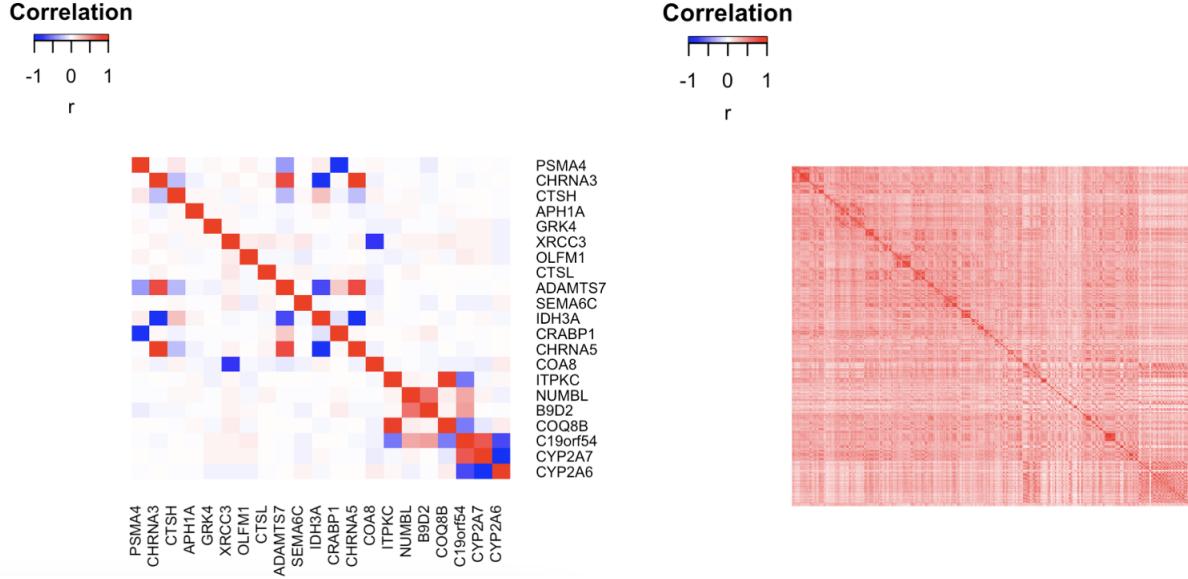


Figure 2: (A) Pairwise correlations among the expression of 21 genes used as exposures in our analysis. (B) Pairwise correlations among the 635 ROI-to-ROI functional connectivity (FC) measures used as mediators in our analysis.

Table 3: List of genes selected by MEMM for the “gene-brain FC-nicotine addiction” mediation pathway

Gene symbols	Gene function
<i>CHRNA3</i>	encodes a subunit of the neuronal nicotinic acetylcholine receptor (nAChR) called alpha-3 subunit.
<i>CHRNA5</i>	encodes a subunit of the neuronal nicotinic acetylcholine receptor (nAChR) called alpha-5 subunit.
<i>PSMA4</i>	encodes a subunit of the 20S proteasome, part of the ubiquitin-proteasome system responsible for protein degradation within cells.
<i>GRK4</i>	encodes G protein-coupled receptor kinase 4, which plays a key role in regulating blood pressure.
<i>CRABP1</i>	encodes a carrier protein binding to retinoic acid and facilitating its transport within the cell.
<i>CTSH</i>	encodes cathepsin H, a lysosomal cysteine protease crucial for protein degradation within cells.

5. Discussion

In this paper, we proposed a new multi-exposure-to-multi-mediator mediation model for imaging genetic study of brain disorders to understand how the multivariate genetic risk factors impact the brain disorders via changing the multivariate neuroimaging features, including both structural and functional changes. We constructed sparse aggregated variables that captured much of the indirect effect in the mediation pathway, balancing the estimation of mediation proportion and feature selection for interpretability in a mediation model with multiple exposures and mediators. We proposed a fast algorithm based on ADMM to solve the optimization and theoretically justified the convergence guarantees of the algorithm and consistency of the estimator. An efficient R package called “MEMM” was developed (<https://github.com/nwang123/MEMM>) to implement the method.

Our model is among the very few that perform mediation analysis with multiple exposures and multiple mediators. In imaging genetic study specifically, it is widely known that structural or functional changes in the brain can partially explain the genetic effect on brain disorders. However, with the large number of relevant genes and neuroimaging features, identifying the predominant mediation pathway through which primary genetic determinants of brain disorders exert their effects via functional or structural alterations in specific brain regions remains challenging. Our method proposed one solution that identified the best mediation pathway with large mediation proportion while in the same time selecting the most critical genes and regional neuroimaging features involved in the pathway for better interpretability. In the examples we showed, we assumed the number of both exposures and mediators are of moderate size (at most n) as both need to be related to the outcome of interest. When both exposures and mediators are high-dimensional ($\gg n$), the algorithm can be less stable and computationally intensive. In this case, we encourage a two-stage approach by first screening out noisy exposures and mediators simultaneously using advanced high-dimensional-to-high-dimensional sure screening techniques [50] and then performing our method.

The current method is built on the basis of a linear mediation model. For genetic and neuroimaging features, it is common that more complex nonlinear relationship and mediation effect may exist. Future work could focus on incorporating nonlinearity into our current mediation model [51] and adopting nonlinear dimension reduction in the reduce rank aggregation [52]. Lastly, establishing the true causal mediation relies on major identification assumptions [53] that are usually not directly testable for the large-scale genetic and imaging data in cross-sectional studies like UKB. Future research could focus on longitudinal cohorts such as Adolescent Brain Cognitive Development (ABCD) study [54] that includes longitudinal imaging mediators and outcome with possible temporal ordering to validate the identified mediation pathways.

Author contributions

Neng Wang: methodology, theoretical proof, simulation, data collection and analysis, software, original manuscript preparation and editing. **Eric V Slud:** supervision of the project, theoretical proof, manuscript review and editing. **Tianzhou Ma:** conceptualization, supervision of the project, manuscript review and editing.

Acknowledgments

This work was supported in part by the National Institutes of Health (NIDA 1K01DA059603) and Grand Challenge Grant by the University of Maryland.

Appendix A.

Proof of Theorem 1. Define the augmented Lagrangian

$$\mathcal{L}_\rho(\mathbf{a}, \mathbf{b}, \mathbf{z}_a, \mathbf{z}_b, \mathbf{u}_a, \mathbf{u}_b) := F(\mathbf{a}, \mathbf{b}) + G_a(\mathbf{z}_a) + G_b(\mathbf{z}_b) + \frac{\rho}{2} \|\mathbf{a} - \mathbf{z}_a + \mathbf{u}_a\|_2^2 + \frac{\rho}{2} \|\mathbf{b} - \mathbf{z}_b + \mathbf{u}_b\|_2^2,$$

where $G_a(\mathbf{z}_a) = \lambda_a \|\mathbf{z}_a\|_1$ and $G_b(\mathbf{z}_b) = \lambda_b \|\mathbf{z}_b\|_1$ and abbreviate

$$\mathcal{L}_\rho^k := \mathcal{L}_\rho(\mathbf{a}^k, \mathbf{b}^k, \mathbf{z}_a^k, \mathbf{z}_b^k, \mathbf{u}_a^k, \mathbf{u}_b^k), \quad \Delta_k := \|\mathbf{a}^{k+1} - \mathbf{z}_a^{k+1}\|_2^2 + \|\mathbf{b}^{k+1} - \mathbf{z}_b^{k+1}\|_2^2.$$

Step 1: Descent under the \mathbf{a} - and \mathbf{b} -updates.. By Assumption 3 (convexity and unique minimisers of the block subproblems), the exact minimisation in each of the \mathbf{a} and \mathbf{b} blocks yields

$$\mathcal{L}_\rho(\mathbf{a}^{k+1}, \mathbf{b}^{k+1}, \mathbf{z}_a^k, \mathbf{z}_b^k, \mathbf{u}_a^k, \mathbf{u}_b^k) \leq \mathcal{L}_\rho^k. \quad (1)$$

Step 2: Descent under the \mathbf{z} -updates.. The proximal updates $\mathbf{z}_a^{k+1} = S_{\lambda_a/\rho}(\mathbf{a}^{k+1} + \mathbf{u}_a^k)$ and $\mathbf{z}_b^{k+1} = S_{\lambda_b/\rho}(\mathbf{b}^{k+1} + \mathbf{u}_b^k)$ satisfy

$$\mathcal{L}_\rho(\mathbf{a}^{k+1}, \mathbf{b}^{k+1}, \mathbf{z}_a^{k+1}, \mathbf{z}_b^{k+1}, \mathbf{u}_a^k, \mathbf{u}_b^k) \leq \mathcal{L}_\rho(\mathbf{a}^{k+1}, \mathbf{b}^{k+1}, \mathbf{z}_a^k, \mathbf{z}_b^k, \mathbf{u}_a^k, \mathbf{u}_b^k) - \frac{\rho}{2} \Delta_k. \quad (2)$$

Step 3: Summability of the primal residuals.. The dual updates $\mathbf{u}_a^{k+1} = \mathbf{u}_a^k + \mathbf{a}^{k+1} - \mathbf{z}_a^{k+1}$, $\mathbf{u}_b^{k+1} = \mathbf{u}_b^k + \mathbf{b}^{k+1} - \mathbf{z}_b^{k+1}$ leave \mathcal{L}_ρ unchanged. Combining (1) and (2) gives

$$\mathcal{L}_\rho^{k+1} \leq \mathcal{L}_\rho^k - \frac{\rho}{2} \Delta_k.$$

Summing from $k = 0$ to $K - 1$ yields

$$\mathcal{L}_\rho^K + \frac{\rho}{2} \sum_{k=0}^{K-1} \Delta_k \leq \mathcal{L}_\rho^0.$$

By Assumption 3 and the fact that the domain of the iterates is restricted to the admissible region $\mathcal{C}_{r_0, \delta}$, the augmented Lagrangian is bounded below; hence $\sum_{k=0}^{\infty} \Delta_k < \infty$ and, in particular, $\Delta_k \rightarrow 0$.

Step 4: Boundedness and stationarity.. Since $(\mathbf{a}^k, \mathbf{b}^k) \in \mathcal{C}$ for all k and \mathcal{C} is compact, the primal iterates are bounded. Dual iterates remain bounded because their increments $\mathbf{a}^{k+1} - \mathbf{z}_a^{k+1}$ and $\mathbf{b}^{k+1} - \mathbf{z}_b^{k+1}$ vanish. Hence cluster points exist.

Exact optimality in the \mathbf{a} block gives

$$0 = \nabla_{\mathbf{a}} F(\mathbf{a}^{k+1}, \mathbf{b}^k) + \rho(\mathbf{a}^{k+1} - \mathbf{z}_a^k + \mathbf{u}_a^k),$$

and after the dual update one shows $\nabla_{\mathbf{a}}\Phi(\mathbf{a}^{k+1}, \mathbf{b}^{k+1}) + \rho(\mathbf{a}^{k+1} - \mathbf{z}_a^{k+1}) = 0$. Similarly for \mathbf{b} . Taking norms and using $\Delta_k \rightarrow 0$ yields $\|\nabla\Phi(\mathbf{a}^k, \mathbf{b}^k)\| \rightarrow 0$. Thus every cluster point is a first-order stationary point of Φ on \mathcal{C} , proving part (i).

Step 5: Convergence rates via the KL inequality.. From Steps 1–3 we have a descent $\Phi(\mathbf{a}^k, \mathbf{b}^k) - \Phi(\mathbf{a}^{k+1}, \mathbf{b}^{k+1}) \geq \frac{\rho}{2} \Delta_k$ and $\sum \Delta_k < \infty$. By the Kurdyka–Łojasiewicz (KL) property assumed in Assumption 2, there exist constants $c > 0$ and $\theta \in [0, 1)$ such that, in a neighbourhood of any cluster point,

$$\|\nabla\Phi(\mathbf{a}, \mathbf{b})\| \geq c [\Phi(\mathbf{a}, \mathbf{b}) - \Phi^*]^\theta.$$

Standard arguments (cf. [55, Thm. 2]) then yield the global $\mathcal{O}(1/K)$ rate in part (ii). If additionally F is strongly convex in one block (with the other block fixed), the KL exponent satisfies $\theta = \frac{1}{2}$, giving the local linear rate in part (iii). \square

References

- [1] Brainstorm Consortium, V. Anttila, B. Bulik-Sullivan, H. K. Finucane, R. K. Walters, J. Bras, L. Duncan, V. Escott-Price, G. J. Falcone, P. Gormley, et al., Analysis of shared heritability in common disorders of the brain, *Science* 360 (6395) (2018) eaap8757.
- [2] T. Kaufmann, D. van der Meer, N. T. Doan, et al., Common brain disorders are associated with heritable patterns of apparent aging of the brain, *Nature Neuroscience* 22 (10) (2019) 1617–1623. [doi:10.1038/s41593-019-0471-7](https://doi.org/10.1038/s41593-019-0471-7).
- [3] T. Horwitz, K. Lam, Y. Chen, Y. Xia, C. Liu, A decade in psychiatric gwas research, *Molecular psychiatry* 24 (3) (2019) 378–389.
- [4] Z. Ye, C. Mo, H. Ke, Q. Yan, C. Chen, P. Kochunov, L. E. Hong, B. D. Mitchell, S. Chen, T. Ma, Meta-analysis of transcriptome-wide association studies across 13 brain tissues identified novel clusters of genes associated with nicotine addiction, *Genes* 13 (1) (2021) 37.
- [5] A. S. Hatoum, S. M. Colbert, E. C. Johnson, S. B. Huggett, J. D. Deak, G. A. Pathak, M. V. Jennings, S. E. Paul, N. R. Karcher, I. Hansen, et al., Multivariate genome-wide association meta-analysis of over 1 million subjects identifies loci underlying multiple substance use disorders, *Nature Mental Health* 1 (3) (2023) 210–223.
- [6] L. T. Elliott, K. Sharp, F. Alfaro-Almagro, et al., Genome-wide association studies of brain imaging phenotypes in UK biobank, *Nature* 562 (7726) (2018) 210–216. [doi:10.1038/s41586-018-0571-7](https://doi.org/10.1038/s41586-018-0571-7).
- [7] X. Bi, L. Yang, T. Li, B. Wang, H. Zhu, H. Zhang, Genome-wide mediation analysis of psychiatric and cognitive traits through imaging phenotypes, *Human brain mapping* 38 (8) (2017) 4088–4097.
- [8] D. P. MacKinnon, J. L. Krull, C. M. Lockwood, Equivalence of the mediation, confounding and suppression effect, *Prevention Science* 1 (4) (2000) 173–181. [doi:10.1023/A:1026595011371](https://doi.org/10.1023/A:1026595011371).

- [9] H. Zhang, Y. Zheng, Z. Zhang, T. Gao, B. Joyce, G. Yoon, W. Zhang, J. Schwartz, A. Just, E. Colicino, et al., Estimating and testing high-dimensional mediation effects in epigenetic studies, *Bioinformatics* 32 (20) (2016) 3150–3154.
- [10] Y. Zhao, X. Luo, Pathway Lasso: pathway estimation and selection with high-dimensional mediators, *Statistics and Its Interface* 15 (1) (2022) 39–50. [doi:10.4310/21-sii673](https://doi.org/10.4310/21-sii673).
- [11] C. Perera, H. Zhang, Y. Zheng, L. Hou, A. Qu, C. Zheng, K. Xie, L. Liu, Hima2: high-dimensional mediation analysis and its application in epigenome-wide dna methylation data, *BMC bioinformatics* 23 (1) (2022) 296.
- [12] Q. Zhang, High-dimensional mediation analysis with applications to causal gene identification, *Statistics in biosciences* 14 (3) (2022) 432–451.
- [13] Y.-T. Huang, W.-C. Pan, Hypothesis test of mediation effect in causal mediation model with high-dimensional continuous mediators, *Biometrics* 72 (2) (2016) 402–413. [doi:10.1111/biom.12421](https://doi.org/10.1111/biom.12421).
- [14] Y. Gao, H. Yang, R. Fang, Y. Zhang, E. L. Goode, Y. Cui, Testing mediation effects in high-dimensional epigenetic studies, *Frontiers in genetics* 10 (2019) 1195.
- [15] Y. Zhao, Y. Fan, A. Qiu, et al., Multimodal mediation analysis of brain imaging data, *IEEE Transactions on Medical Imaging* 39 (9) (2020) 2862–2873. [doi:10.1109/TMI.2020.2993485](https://doi.org/10.1109/TMI.2020.2993485).
- [16] O. Y. Chén, C. M. Crainiceanu, E. L. Ogburn, et al., High-dimensional multivariate mediation with application to neuroimaging data, *Biostatistics* 19 (2) (2018) 121–136. [doi:10.1093/biostatistics/kxx027](https://doi.org/10.1093/biostatistics/kxx027).
- [17] H. Lee, C. Chen, P. Kochunov, E. Hong, L. S. Chen, A new multiple-mediator model maximally uncovering the mediation pathway: Evaluating the role of neuroimaging measures in age-related cognitive decline, *Annals of Applied Statistics* 18 (4) (2024) to appear. [doi:10.1214/24-aoas1905](https://doi.org/10.1214/24-aoas1905).
- [18] Y. Dai, X. Li, H. Liu, Integrative mediation analysis of multiple pathways in high-dimensional data, arXiv preprintReplace with journal details & DOI when published (2024). [arXiv:arXiv:2406.XXXX](https://arxiv.org/abs/2406.XXXX).
- [19] P. F. Sullivan, A. Agrawal, C. M. Bulik, O. A. Andreassen, A. D. Børglum, G. Breen, S. Cichon, H. J. Edenberg, S. V. Faraone, J. Gelernter, et al., Psychiatric genomics: an update and an agenda, *American Journal of Psychiatry* 175 (1) (2018) 15–27.
- [20] S. M. Smith, T. E. Nichols, D. Vidaurre, et al., A positive-negative mode of population covariation links brain connectivity, demographics and behavior, *Nature Neuroscience* 18 (11) (2015) 1565–1567. [doi:10.1038/nn.4125](https://doi.org/10.1038/nn.4125).

- [21] E. R. Gamazon, H. E. Wheeler, K. P. Shah, et al., A gene-based association method for mapping traits using reference transcriptome data, *Nature Genetics* 47 (9) (2015) 1091–1098. [doi:10.1038/ng.3367](https://doi.org/10.1038/ng.3367).
- [22] H. Zhou, W. Pan, C. Hu, Sparse pca-based mediation analysis for high-dimensional data, *Biostatistics* 21 (4) (2020) e576–e592. [doi:10.1093/biostatistics/kxy065](https://doi.org/10.1093/biostatistics/kxy065).
- [23] R. M. Baron, D. A. Kenny, The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations, *Journal of Personality and Social Psychology* 51 (6) (1986) 1173–1182. [doi:10.1037/0022-3514.51.6.1173](https://doi.org/10.1037/0022-3514.51.6.1173).
- [24] D. P. MacKinnon, *Introduction to Statistical Mediation Analysis*, Erlbaum, New York, 2007.
- [25] K. Imai, L. Keele, D. Tingley, A general approach to causal mediation analysis, *Psychological Methods* 15 (4) (2010) 309–334. [doi:10.1037/a0020761](https://doi.org/10.1037/a0020761).
- [26] R. Tibshirani, Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society, Series B* 58 (1) (1996) 267–288. [doi:10.1111/j.2517-6161.1996.tb02080.x](https://doi.org/10.1111/j.2517-6161.1996.tb02080.x).
- [27] S. Boyd, N. Parikh, E. Chu, B. Peleato, J. Eckstein, Distributed optimization and statistical learning via the alternating direction method of multipliers, *Foundations and Trends in Machine Learning* 3 (1) (2011) 1–122. [doi:10.1561/2200000016](https://doi.org/10.1561/2200000016).
- [28] J. Bolte, S. Sabach, M. Teboulle, Proximal alternating linearized minimization for non-convex and nonsmooth problems, *Mathematical Programming* 146 (1–2) (2014) 459–494. [doi:10.1007/s10107-013-0701-9](https://doi.org/10.1007/s10107-013-0701-9).
- [29] S. Zhang, L. Zhang, L. Yin, S. Ma, High-dimensional mediation analysis with applications to causal gene identification, *Biometrics* 72 (4) (2016) 1291–1302. [doi:10.1111/biom.12513](https://doi.org/10.1111/biom.12513).
- [30] Y. Zhao, X. Luo, Pathway lasso: Pathway estimation and selection with high-dimensional mediators, *Statistics and Its Interface* 15 (1) (2022) 39–50. [doi:10.4310/SII.2021-SII673](https://doi.org/10.4310/SII.2021-SII673).
- [31] S. Chen, C. Crainiceanu, E. L. Ogburn, B. S. Caffo, M. A. Lindquist, High-dimensional multivariate mediation with application to neuroimaging data, *Biostatistics* 19 (2) (2018) 121–136. [doi:10.1093/biostatistics/kxx023](https://doi.org/10.1093/biostatistics/kxx023).
- [32] Y. Zhao, M. A. Lindquist, B. S. Caffo, Sparse principal component based high-dimensional mediation analysis, *Computational Statistics & Data Analysis* 142 (2020) 106835. [doi:10.1016/j.csda.2019.106835](https://doi.org/10.1016/j.csda.2019.106835).
- [33] M. Bartal, Health effects of tobacco use and exposure, *Monaldi Archives for Chest Disease* 56 (6) (2001) 545–554. [doi:10.4081/monaldi.2001.549](https://doi.org/10.4081/monaldi.2001.549).

- [34] A. M. Erzurumluoglu, M. Liu, V. E. Jackson, D. R. Barnes, G. Datta, C. A. Melbourne, R. Young, C. Batini, P. Surendran, T. Jiang, et al., Meta-analysis of up to 622,409 individuals identifies 40 novel smoking behaviour associated genetic loci, *Molecular psychiatry* 25 (10) (2020) 2392–2409.
- [35] K. Xu, J. W. J. Li, J. Yang, Genome-wide association study of cigarettes per day in > 300,000 european ancestry individuals, *Addiction Biology* 25 (5) (2020) e12812. [doi:10.1111/adb.12812](https://doi.org/10.1111/adb.12812).
- [36] D. P. Buchwald, B. M. Neale, J. Gelernter, Genome-wide association meta-analysis of smoking quantity identifies multiple novel loci, *Molecular Psychiatry* 26 (2021) 2906–2918. [doi:10.1038/s41380-020-00913-3](https://doi.org/10.1038/s41380-020-00913-3).
- [37] S. Kuhn, J. Gallinat, Brain structure adaptations in smokers: Voxel-based morphometry and meta-analysis, *Human Brain Mapping* 33 (12) (2012) 292–308. [doi:10.1002/hbm.21219](https://doi.org/10.1002/hbm.21219).
- [38] M. A. Gray, G. D. Smith, H. M. Grant, Associations between cigarette smoking and structural brain mri measures in uk biobank, *Addiction Biology* 25 (2) (2020) e12770. [doi:10.1111/adb.12770](https://doi.org/10.1111/adb.12770).
- [39] T. F. Heatherton, L. T. Kozlowski, R. C. Frecker, W. L. Fagerström, The fagerström test for nicotine dependence: A revision of the fagerström tolerance questionnaire, *British Journal of Addiction* 86 (9) (1991) 1119–1127. [doi:10.1111/j.1360-0443.1991.tb01879.x](https://doi.org/10.1111/j.1360-0443.1991.tb01879.x).
- [40] G. Consortium, The gtex consortium atlas of genetic regulatory effects across human tissues, *Science* 369 (6509) (2020) 1318–1330.
- [41] A. N. Barbeira, M. Pividori, J. Zheng, H. E. Wheeler, D. L. Nicolae, H. K. Im, Integrating predicted transcriptome from multiple tissues improves association detection, *PLoS genetics* 15 (1) (2019) e1007889.
- [42] N. Wang, Z. Ye, T. Ma, TIPS: a novel pathway-guided joint model for transcriptome-wide association studies, *Briefings in Bioinformatics* 25 (6) (2024) bbae587. [doi:10.1093/bib/bbae587](https://doi.org/10.1093/bib/bbae587).
- [43] G. Lassi, A. E. Taylor, N. J. Timpson, P. J. Kenny, R. J. Mather, T. Eisen, M. R. Munafò, The chrna5–a3–b4 gene cluster and smoking: from discovery to therapeutics, *Trends in neurosciences* 39 (12) (2016) 851–861.
- [44] C. D. Fowler, N. Arends, P. M. Kenny, The $\alpha 5$ nicotinic receptor subunit plays a key role in nicotine reward and withdrawal, *Current Biology* 21 (22) (2011) 221–227. [doi:10.1016/j.cub.2011.10.041](https://doi.org/10.1016/j.cub.2011.10.041).
- [45] K. Rezvani, Y. Teng, D. Shim, M. De Biasi, Nicotine regulates multiple synaptic proteins by inhibiting proteasomal activity, *Journal of Neuroscience* 27 (39) (2007) 10508–10519.

- [46] R. Sherva, A. B. Smoller, R. Walters, Genetic associations of the *chrna5–chrna3–chrnb4* gene cluster with smoking behaviors and lung cancer risk, *Nature Genetics* 48 (1) (2016) 59–65. [doi:10.1038/ng.3452](https://doi.org/10.1038/ng.3452).
- [47] L. E. Hong, H. Gu, Y. Yang, T. J. Ross, B. J. Salmeron, B. Buchholz, G. K. Thaker, E. A. Stein, Association of nicotine addiction and nicotine’s actions with separate cingulate cortex functional circuits, *Archives of general psychiatry* 66 (4) (2009) 431–441.
- [48] D. G. Ghahremani, J.-B. F. Pochon, M. P. Diaz, R. F. Tyndale, A. C. Dean, E. D. London, Nicotine dependence and insula subregions: functional connectivity and cue-induced activation, *Neuropsychopharmacology* 48 (6) (2023) 936–945.
- [49] D. Spurden, J. Court, S. Lloyd, A. Oakley, R. Perry, C. Pearson, R. Pullen, E. Perry, Nicotinic receptor distribution in the human thalamus: autoradiographical localization of [3h] nicotine and [125i] α -bungarotoxin binding, *Journal of Chemical Neuroanatomy* 13 (2) (1997) 105–113.
- [50] H. Ke, Z. Ren, J. Qi, S. Chen, G. C. Tseng, Z. Ye, T. Ma, High-dimension to high-dimension screening for detecting genome-wide epigenetic and noncoding rna regulators of gene expression, *Bioinformatics* 38 (17) (2022) 4078–4087.
- [51] W. W. Loh, B. Moerkerke, T. Loeys, S. Vansteelandt, Nonlinear mediation analysis with high-dimensional mediators whose causal structure is unknown, *Biometrics* 78 (1) (2022) 46–59.
- [52] J. A. Lee, M. Verleysen, Nonlinear dimensionality reduction, Springer Science & Business Media, 2007.
- [53] K. Imai, L. Keele, D. Tingley, A general approach to causal mediation analysis., *Psychological methods* 15 (4) (2010) 309.
- [54] B. J. Casey, T. Cannonier, M. I. Conley, A. O. Cohen, D. M. Barch, M. M. Heitzeg, M. E. Soules, T. Teslovich, D. V. Dellarco, H. Garavan, et al., The adolescent brain cognitive development (abcd) study: imaging acquisition across 21 sites, *Developmental cognitive neuroscience* 32 (2018) 43–54.
- [55] M. Hong, Z.-Q. Luo, M. Razaviyayn, J.-S. Pang, Convergence of alternating direction method of multipliers for nonconvex and nonsmooth optimization, *SIAM Journal on Optimization* 26 (1) (2016) 337–364. [doi:10.1137/140961791](https://doi.org/10.1137/140961791).