Bayesian Methods in High-Dimensional Sparse Mediation Analysis

by

Yanyi Song

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biostatistics)
in The University of Michigan
2020

Doctoral Committee:

Professor Bhramar Mukherjee, Co-Chair
Associate Professor Xiang Zhou, Co-Chair
Professor Jian Kang
Associate Professor Jennifer A. Smith
Professor Min Zhang

Yanyi Song

yanys@umich.edu

ORCID iD: 0000-0002-7652-6864

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF APPENDICES

Appendix

ABSTRACT

Causal mediation analysis aims to examine the role of a mediator or a group of mediators that lie in the pathway between an exposure and an outcome. Recent biomedical studies often involve a large number of potential mediators, typically a large ensemble of biomarkers that are measured via high-throughput technologies. The goal of my dissertation is to develop novel statistical methods that can accommodate and leverage high-dimensional mediators in mediation analysis. We provide an overview of mediation analysis and an outline of our work in Chapter I. We elaborate our methodological developments in the following chapters.

In Chapter II, we develop a Bayesian inference method using continuous shrinkage priors to simultaneously analyze high-dimensional mediators. Simulations demonstrate that our method improves the power of global mediation analysis compared to simpler alternatives and has decent performance to identify true non-null contributions to the mediation effects of the pathway. The Bayesian method also helps us to understand the structure of the composite null cases for inactive mediators in the pathway. We applied our method to Multi-Ethnic Study of Atherosclerosis (MESA) and identified DNA methylation regions that may actively mediate the effect of socioeconomic status (SES) on cardiometabolic outcomes.

In Chapter III, we develop methods to directly perform targeted penalization of the natural indirect effect (NIE) in a Bayesian paradigm. Specifically, we develop two novel prior models for identification of the NIEs in high-dimensional mediation analysis, both with a joint distribution on the coefficients of the exposure-mediator and mediator-outcome models: (a) four-component Gaussian mixture prior, and (b) product threshold Gaussian prior. By jointly modeling the two parameters that contribute to the NIE, the proposed methods

enable penalization on their product in a targeted way. Resultant inference can take into account the four-component composite structure underlying the indirect effect. We show through extensive simulations that the proposed methods improve both selection and estimation accuracy compared to other existing or alternative shrinkage/penalization based methods. We applied our methods to two ongoing epidemiological studies: the MESA and the LIFECODES birth cohort. The identified active mediators reveal important biological pathways that may be useful for understanding disease mechanism.

In Chapter IV, we further extend the Gaussian mixture method in Chapter III to explicitly incorporate the useful correlation structural information among mediators in the model building process. Instead of assuming independent prior for each mediator as in our previous methods, we propose to (a) jointly model the mixing probabilities for correlated mediator selection, or (b) jointly model the group indicators by a Potts distribution, both adding the possible grouping effect across mediators through another layer in the Bayesian hierarchy. We develop efficient sampling algorithms under non-conjugate priors and large state space. Various simulations demonstrate that our methods enable effective identification of active mediators with high correlations, which could be missed using independent priors. The proposed methods also suggest new mediation findings in the LIFECODES and MESA data applications.

CHAPTER I

Introduction

1.1 Mediation Analysis

Mediation analysis has received great attention and are broadly applied across various disciplines (*MacKinnon et al.*, 2007a; *Imai et al.*, 2010a; *Albert*, 2008; *Jo*, 2008; *Ten Have et al.*, 2007). It attempts to decompose the exposure's effect into indirect effect that acts through an intermediate variable, called mediator, and direct effect on the outcome that is unexplained by the mediator. The exposure influences the mediator, which in turn influences the outcome. This entire mediating relationship can be visually represented in the diagram 1.1. For example, in the field of educational psychology, researchers investigated the extent to which the academic achievement could be improved by encouraging mastery achievement goals, compared with fostering strategic learning (*Diseth and Kobbeltvedt*, 2010). The former concerns the direct effect of establishing achievement goals, while the latter concerns the indirect effect mediated by strategic approaches of deep learning.

The importance of mediating variables has long been recognized in the area of psychology, mainly owing to the historic popularity of stimulus organism response model (*Woodworth*, 1930). In this model, an active organism transmits messages from a stimulus to elicit a response. Similar mediating frameworks also underlie many other psychological theories (*Fishbein and Ajzen*, 1977). Later, *Baron and Kenny* (1986) outlined the classical causal

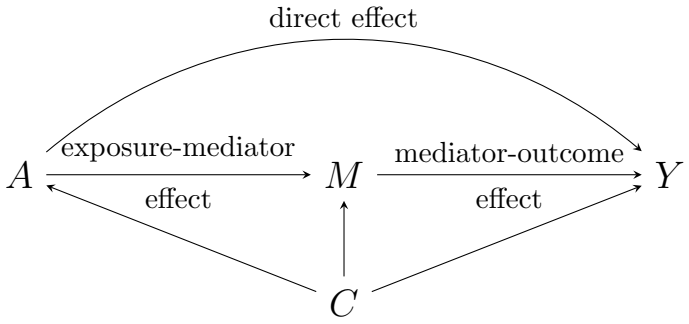

Figure 1.1: The visual illustration of the mediating relationship in a simple mediation model, with an exposure A, a mediator M, an outcome Y and the potential confounders C.

steps approach to assess mediation. This approach involves a single mediator and three regression equations, i.e. relating the exposure to the outcome, relating the exposure to the mediator, and relating the mediator to the outcome adjusted for the exposure. This influential work has built the basic foundation for the subsequent development on mediation testing (*MacKinnon et al.*, 1995, 2002, 2007b; *Preacher and Hayes*, 2008). Those approaches treat difference of coefficients or product of coefficients as statistics of interest and use different approximation formulas in obtaining the standard error of its sampling distribution. Some literature also expressed caution or criticism about Baron and Kenny's prerequisite on a significant exposure-outcome relationship to establish mediation (*Hayes*, 2009; *Rucker et al.*, 2011). They have argued that it was possible for an indirect effect to be significantly away from zero even if the total effect of exposure on the outcome is not. This single-mediator model has then been extended to a wide range of applications in psychology and psychiatry, epidemiology, prevention and treatment research (*Ditlevsen et al.*, 2005; *MacKinnon*, 2008). It has also been extended to simple settings with a moderate number of mediators by including multiple mediators in the outcome regression and one exposure-mediator regression for each mediator (*MacKinnon*, 2000). The moderate number of mediators considered was usually no more than 10.

The above classical mediation analysis has been formulated and conducted within the statistical linear regressions, but a general definition of casual mediation effects beyond a par-

ticular model is still missing. At the same time, the approach of causal modeling based on counterfactual/potential outcome framework (*Rubin*, 1974) has received great attention and interest (*Schafer and Kang*, 2008; *Mulaik*, 2009). Consequently, causal mediation analysis, a general, unified approach built on the counterfactual framework has been proposed, with the key assumptions for identification and causal interpretation being specified (*Pearl*, 2001; *Imai et al.*, 2010a,b). This framework allows for formal definition of natural/controlled direct and indirect effects, and further gave rise to other possible extensions in mediation analysis, such as categorical mediator/outcome (*VanderWeele and Vansteelandt*, 2010; *Albert and Nelson*, 2011), survival data (*VanderWeele*, 2011a), exposure-mediator interaction (*Valeri and VanderWeele*, 2013), etc.

1.2 Motivation

With the rapid advances in high-throughput biological technologies, recent studies often involve a large number of potential mediators, such as human molecular-level traits or brain image features. How to perform mediation analysis in a high-dimensional setting or leverage information from the high-dimensional mediators has become an emerging topics. Most of the previous literature focus on univariate mediator analysis or settings with a moderate number of mediators. In this dissertation, building on the potential outcome framework for causal inference, we attempt to develop novel Bayesian methods for mediator selection and causal effect estimation in high-dimensional sparse mediation analysis. The sparsity assumption implies only a small proportion of the mediators exhibit individually large mediation effects, which we refer to as active mediators, while the majority would contribute small background effects in mediating the exposure-outcome relationship. The sparse setup has become a standard scheme in genetic and epigenetic studies (*Zhou et al.*, 2013; *Shen et al.*, 2013) and been proven to yield better performance in detecting important features. Those Bayesian models will enable us to jointly analyze a large ensemble of correlated mediators without

making any causal ordering assumptions on the mediators, and provide a natural way to induce sparsity and characterize the uncertainty in the parameters.

As pointed out in *Hayes* (2009), many of the previous methods for mediation analysis is not based on a quantification of the very thing it is attempting to test: the mediation (indirect) effect. Instead, they infer mediation from two partial associations (exposure-mediator association and mediator-outcome association). The requirement of both associations being significant to declare an indirect effect was shown to be stringent at the cost of possibly missing some important mediation discoveries. The most well-known inferential technique for quantifying indirect effects rather than their constituent effects separately is probably the Sobel test (*Sobel*, 1982). But one limitation of the Sobel test is that it assumes the sampling distribution of the indirect effect is normal. Given that the sampling distribution is essentially a product of normal, the normality assumption may not be valid in many cases and can lead to over-conservative results (*MacKinnon et al.*, 2007b). As proposed in the first chapter, a Bayesian approach in high-dimensional mediation analysis can provide an efficient way for the identification of active mediators and the embedding of mediation structure. However, as far as we know, none of the prior distributions is designed for shrinkage on the indirect effect in a targeted way. In addition, genome-wide mediation analysis has been an emerging area of research that will continue to grow as more data are collected. Testing the indirect effect for each single mediator is challenging due to the complexity in its null distribution, which involves composite structure arising from the product-type test statistics. Fortunately, a large number of mediators make it possible to characterize that composite space underlying mediation mechanisms. Recent literature began to recognize and leverage the composite structure in the null hypothesis of no indirect effect, and have proposed effective testing procedure for one-at-a-time single mediator analysis (*Huang et al.*, 2019). Motivated by the underlying composite structure of indirect effect and also the goal of directly targeting the non-null indirect effect, we are interested in seeking the Bayesian parallel with a joint prior on the exposure-mediator and mediator-outcome coefficients, which

is so far lacking in the literature.

The above methods will enable a joint analysis of high-dimensional mediators and efficient identification of active mediators in a sparse setting. However, they do not explicitly take into account the correlation structure among the mediators. Treating mediators independent *a priori*, the subsequent inference may fail to distinguish between highly correlated mediators and result in loss of power for mediator selection. Correlated mediators presumably exhibit similar association pattern with exposure and outcome, and therefore can be grouped together to borrow strength for Bayesian learning. Incorporating correlated structural information in mediation analysis enjoys great support from previous literature on Bayesian structural variable selection, and is promising to boost the overall performance in the presence of high-dimensional correlated mediators.

1.3 Summary of Objectives

With an emphasis on the problems described above, in this dissertation, I present methodologies that aim to achieve the following objectives:
(1) To develop Bayesian shrinkage models that can simultaneously accommodate high-dimensional mediators;
(2) To develop Bayesian sparse mediation methods that can penalize the indirect effect in a target way;
(3) To further extend the previous methods to explicitly incorporate the correlated structural information from mediators.

The above three objectives are addressed in Chapter II, III and IV, respectively. More details on the background, pertinent literature review, motivation and methodology development can be found in the introduction sections of each chapters.

1.4 Data Acknowledgement

The proposed methods in this dissertation are generally applicable to many settings, and here we focus on the following epidemiological and environmental studies:

Multi-Ethnic Study of Atherosclerosis (MESA) is a population-based longitudinal study designed to identify risk factors for the progression of subclinical cardiovascular disease (CVD) (*Bild et al.*, 2002). Between July 2000 and August 2002, a total of 6,814 participants without clinically apparent CVD were recruited from six regions in the U.S. All the participants were measured for exposure variables that include various aspects of childhood, adulthood and neighborhood socioeconomic position. They were also measured for clinical health outcomes, including body mass index/obesity, and diabetes/glucose/insulin/HbA1c. Between April 2010 and February 2012 (corresponding to MESA Exam 5), DNA methylation and gene expression data were collected on a random subsample of 1,264 participants. We hypothesize that those molecular-level omics traits are part of the mediating mechanism through which socioeconomic and neighborhood characteristics affect physical health.

LIFECODES Birth Cohort is a cross-sectional study of approximately 1,600 pregnant women (recruited < 15 weeks) between 2006 and 2008 at the Brigham and Womens Hospital in Boston, MA. At the initial study visit, questionnaires were administered to collect demographic and health-related information of the participants. At each of the following four study visits, participants urine and plasma samples were collected. Environmental exposure analytes, including phthalates, phenols and parabens, trace metals and polycyclic aromatic hydrocarbons, were measured from urine samples. A large group of endogenous biomarkers of lipid metabolism, inflammation, and oxidative stress were measured in plasma samples. Among participants recruited in the LIFECODES cohort, 1,181 participants were followed until delivery and had live singleton infants, with the gestational age recorded. We hypothesize that the endogenous biomarkers may mediate the effects of prenatal exposure to environmental contamination on adverse pregnancy outcomes. The integration of

molecular/biological data with epidemiologic data in the mediation framework will provide interesting and important insights into underlying disease mechanisms.

CHAPTER II

Bayesian Shrinkage Estimation of High Dimensional Causal Mediation Effects in Omics Studies

2.1 Introduction

Causal mediation analysis has been of great interest across many disciplines (*VanderWeele*, 2016b; *Ten Have and Joffe*, 2012). It investigates how an intermediate variable, referred to as mediator, explains the mechanism through which the exposure variable affects the outcome. Under certain regularity conditions, mediation analysis allows us to disentangle the exposure's effect into two parts: effect that acts through the mediator of interest (indirect/mediation effect) and effect that is unexplained by the mediator (direct effect). The state-of-the-art causal mediation analysis (*Ten Have and Joffe*, 2012), which is built upon the counterfactual framework (*Robins and Greenland*, 1992; *Imai et al.*, 2010a), establishes rigorous assumptions regarding the exposure-outcome, exposure-mediator and mediator-outcome relationships to justify appropriate use of the classical formulas from Baron and Kenny in the linear regression setting (*Baron and Kenny*, 1986; *MacKinnon*, 2008) and creates a framework for other general extensions. Many of the existing methods focus on univariate mediator analysis that analyzes one mediator at a time in the causal inference framework, and are applicable to both continuous (*Imai et al.*, 2010b) and binary outcomes (*VanderWeele and Vansteelandt*, 2010). These methods have been widely applied in areas of social,

economic, epidemiological and genetic studies (*VanderWeele*, 2016b; *MacKinnon*, 2008), including recent extensions to multiple exposure variables that lead to more powerful single nucleotide polymorphism (SNP) set tests in presence of gene expression data (*Huang et al.*, 2014). Several studies have recently extended mediation analysis models to jointly account for multiple mediators. However, most of the literature considered settings with two or three mediators, where each mediator is ordered along a priori known mediation pathways and the path-specific effects are estimated (*Daniel et al.*, 2015). In the presence of multiple unordered mediators, one often has to rely on an *ad hoc* approach to fit a series of mediation models with one mediator and one exposure (*VanderWeele and Vansteelandt*, 2014; *Taguri et al.*, 2015)/outcome (*Huang and Pan*, 2016) at a time and then summarize the mediation effects across all the mediators. Such approach ignores correlation among mediators and the estimated mediation effect does not necessarily have an intuitive causal interpretation, particularly when the dimension of the potential mediators is truly large.

In this chapter, building on the potential outcome framework for causal inference, we develop a Bayesian mediation analysis method to characterize the indirect effect through an entire set of high-dimensional mediators. Note that Bayesian methods for mediation have also been proposed in a principal stratification framework (*Elliott et al.*, 2010), though there are subsequent discussions on whether the principal stratification framework is a plausible framework to estimate indirect effects (*VanderWeele*, 2011b). In addition, for estimating natural direct and indirect effects, recent work applied Bayesian non-parametric models, especially Dirichlet process mixture models (*Kim et al.*, 2017, 2019) in both univariate and multiple mediators analysis. In contrast, here, we rely on Bayesian variable selection models to simultaneously analyze a relatively large number of mediators in a pathway with potentially a small number being truly active. With sparsity inducing priors on active coefficients, we assume only a small proportion of mediators in the whole set may mediate the exposure effect on the outcome. This sparsity assumption allows us to extend previous univariate mediator analysis to a high-dimensional setting by framing the identification of active mediators in the

whole set as a variable selection problem and applying Bayesian methods with continuous shrinkage priors on the effects. Unlike previous methods developed for multiple mediators, ours can jointly analyze much larger number of mediators without making any path-specific or causal ordering assumptions on mediators. Our method enables us to identify the joint indirect effects of all the mediators and the subset of active ones in the set, and propagates uncertainty in inference in a principled way. Recently, there has been emerging interest in high-dimensional mediation analysis, and our method adds to the burgeoning literature for high-dimensional mediators (*Chén et al.*, 2017; *Derkach et al.*, 2019).

While our method is generally applicable to many settings, we examine the performance of our method in the setting of genomics studies. Due to fast advances in high-throughput biological technologies, genomics studies can nowadays measure a large number of molecular-level traits such as gene expression and DNA methylation (DNAm) levels. Recent studies have proposed these molecular traits may act as a mechanism through which various aspects of socioeconomic status (SES) and neighborhood disadvantages affect physical health. For example, childhood SES, adult SES, social mobility, and neighborhood crime rates have recently been shown to influence DNAm in several genes related to stress and inflammation (*Needham et al.*, 2015; *Smith et al.*, 2017). DNAm of inflammatory markers have also been associated with the status of cardiovascular disease (CVD) and type 2 diabetes (T2D) (*Zhong et al.*, 2016). Here, we show through simulations and data analysis that our high-dimensional mediation analysis framework can increase power of a joint analysis and facilitate the identification of active mediators in the set.

2.2 Notation, Definitions and Assumptions

In this chapter, we focus on causal mediation analysis for the setting where there is a single exposure of interest but there exists a high-dimensional set of candidate mediators that may mediate the effect of exposure on an outcome. Suppose our analysis is based on a study

of n subjects and for subject i, $i = 1, ..., n$, we collect data on exposure A_i, p candidate mediators $\boldsymbol{M_i} = (M_i^{(1)}, M_i^{(2)}, ..., M_i^{(p)})^T$, outcome Y_i, and q covariates $\boldsymbol{C_i} = (C_i^{(1)}, ..., C_i^{(q)})^T$. In particular, we focus on the case where Y_i and $\boldsymbol{M_i}$ are all continuous variables.

We adopt the counterfactual (or potential outcomes) framework to formally define mediators and their causal effects. Let $M_i^{(j)}(a)$ denote the potential (or counterfactual) value of the jth mediator, $j = 1, \ldots, p$, for subject i under exposure level at a. Suppose the exposure has K levels, then $K \times p$ potential counterfactual random variables for mediators are defined, i.e. $M^{(1)}(1), M^{(2)}(1), ..., M^{(p)}(1), M^{(1)}(2), M^{(2)}(2), ..., M^{(p)}(2), M^{(1)}(K), M^{(2)}(K), ..., M^{(p)}(K)$. Let $Y_i(a, \boldsymbol{m}) = Y_i(a, m^{(1)}, ..., m^{(p)})$ denote the ith subject's potential outcome if the subject's exposure were a and mediators were $\boldsymbol{m} = (m^{(1)}, ..., m^{(p)})$. As this chapter focuses on the joint effects of the whole set of mediators, for simplicity, we define $\boldsymbol{M_i}(a) = (M_i^{(1)}(a), M_i^{(2)}(a), \ldots, M_i^{(p)}(a))$. These potential (or counterfactual) variables are hypothetical variables and may not be observed in real data. To connect potential variables to observed data, we make the Stable Unit Treatment Value Assumption (SUTVA) (*Rubin*, 1980), which is a commonly made assumption for performing causal inference. Specifically, the SUTVA assumes there is no interference between subjects and the consistency assumption, which states that the observed variables are the same as the potential variables corresponding to the actually observed treatment level, i.e., $\boldsymbol{M_i} = \sum_a \boldsymbol{M_i}(a) I(A_i = a)$, and $Y_i = \sum_a \sum_{\boldsymbol{m}} Y_i(a, \boldsymbol{m}) I(A_i = a, \boldsymbol{M_i} = \boldsymbol{m})$, where $I(\cdot)$ is the indicator function. For simplicity in notation, we define $Y_i(a) = Y_i(a, \boldsymbol{M_i}(a))$, i.e., the potential outcome had the exposure been a and the whole set of mediators been the value that would have been observed under exposure a. Although potential or counterfactual variables are useful concepts in order to formally define causal effects, they are hypothetical and actually most of them are not observed in real data. For example, if $A_i \neq a$, then $Y_i(a)$ or $\boldsymbol{M_i}(a)$ are not observed. Also $Y_i(a)$ and $Y_i(a^\star)$ are never simultaneously observed for a subject.

We may decompose the effect of an exposure into its direct effect and effect mediated through the whole set of mediators (*VanderWeele and Vansteelandt*, 2014). The controlled

direct effect (CDE) of the exposure on the outcome is defined as $Y_i(a, \boldsymbol{m}) - Y_i(a^\star, \boldsymbol{m})$, which is the effect of changing exposure from level $a^\star$ (the reference level) to a while hypothetically controlling mediators at level $\boldsymbol{m}$. The natural direct effect (NDE) is defined as $Y_i(a, \boldsymbol{M_i}(a^\star)) - Y_i(a^\star, \boldsymbol{M_i}(a^\star))$, which is the CDE when mediators are controlled at the level that would have naturally been had the exposure been $a^\star$. The natural indirect effect (NIE) is defined by $Y_i(a, \boldsymbol{M_i}(a)) - Y_i(a, \boldsymbol{M_i}(a^\star))$, capturing the effect mediated through the whole set of mediators, i.e., the change in potential outcomes when mediators change from $\boldsymbol{M_i}(a^\star)$ to $\boldsymbol{M_i}(a)$ while fixing exposure at a. The total effect (TE), $Y_i(a) - Y_i(a^\star)$, can then be decomposed into natural direct effect and natural indirect effect, written as $Y_i(a) - Y_i(a^\star) = Y_i(a, \boldsymbol{M_i}(a)) - Y_i(a^\star, \boldsymbol{M_i}(a^\star)) = Y_i(a, \boldsymbol{M_i}(a)) - Y_i(a, \boldsymbol{M_i}(a^\star)) + Y_i(a, \boldsymbol{M_i}(a^\star)) - Y_i(a^\star, \boldsymbol{M_i}(a^\star))$ =NIE+NDE.

Causal effects are formally defined in terms of potential variables which are not necessarily observed, but the identification of causal effects must be based on observed data. Therefore, similar to missing data problems, further assumptions regarding the confounders are required for the identification of causal effects in mediation analysis (*VanderWeele and Vansteelandt*, 2014). We will use $A \perp\!\!\!\perp B|C$ to denote that A is independent of B conditional on C. For estimating the average CDE, two assumptions on confounding are needed: (1) $Y_i(a, \boldsymbol{m}) \perp\!\!\!\perp A_i|\boldsymbol{C_i}$, namely, there is no unmeasured confounding for the exposure effect on the outcome; (2) $Y_i(a, \boldsymbol{m}) \perp\!\!\!\perp \boldsymbol{M_i}|\{\boldsymbol{C_i}, A_i\}$, namely, there is no unmeasured confounding for any of mediator-outcome relationship after controlling for the exposure. The two assumptions are illustrated in the left panel of Figure 2.1, and controlling for exposure-outcome and mediator-outcome confounding corresponds to controlling for C_1, C_2 in the figure. In practice, both sets of covariates C_1 and C_2 need not to be distinguished from one another and can simply be included in the overall set of C that we adjust for. The identification of the average NDE and NIE requires assumption (1) and (2), along with two additional assumptions: (3) $\boldsymbol{M_i}(a) \perp\!\!\!\perp A_i|\boldsymbol{C_i}$, namely, there is no unmeasured confounding for the exposure effect on all the mediators; (4) $Y_i(a, \boldsymbol{m}) \perp\!\!\!\perp \boldsymbol{M_i}(a^\star)|\boldsymbol{C_i}$, which can be interpreted as there is no downstream effect of

the exposure that confounds the mediator-outcome relationship for any of the mediators. Graphically, assumption (4) implies that there should be no arrow going from exposure A to mediator-outcome confounder C_2 in Figure 2.1(a). It is thus violated in Figure 2.1(b) since the mediator-outcome confounder L is itself affected by the exposure. The four assumptions are required to hold with respect to the whole set of mediators $\boldsymbol{M_i}(a)$. Finally, as in all mediation analysis, in order for associations to represent causal effects, the temporal ordering assumption also needs to be satisfied, i.e., the exposure precedes the mediators and the mediators precede the outcome.

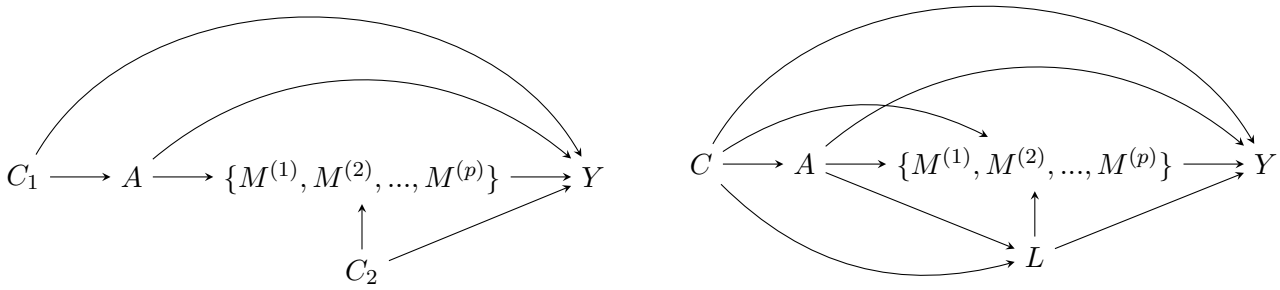

Figure 2.1: Left (a): High-dimensional mediators ($(M^{(1)}, M^{(2)}, ..., M^{(p)})$) between exposure ($A$) and outcome ($Y$) with exposure-outcome confounders C_1 and mediator-outcome confounders C_2; Right (b): An example of mediator-outcome confounder L that is affected by the exposure A.

Now we show that if the above assumptions hold, then the average natural direct and indirect effects can be identified from the observed data. We first notice that $E[Y_i(a, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}]$ can be expressed as below (see the Supporting Information for details),

$$E[Y_i(a, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}] = \int_{\boldsymbol{m}} E(Y_i|a, \boldsymbol{m}, \boldsymbol{C_i}) P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a^\star) \mathrm{d}\boldsymbol{m} \tag{2.1}$$

If we replace a with $a^\star$ in $E[Y_i(a, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}]$, then we get $E[Y_i(a^\star, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}] = \int_{\boldsymbol{m}} E(Y_i|a^\star, \boldsymbol{m}, \boldsymbol{C_i}) \times P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a^\star) \mathrm{d}\boldsymbol{m}$. Therefore, we can express the average natural

direct effect conditional on C as,

$$\begin{aligned} & E[Y_i(a, \boldsymbol{M_i}(a^\star)) - Y_i(a^\star, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}] \\ = & \int_{\boldsymbol{m}} \{E(Y_i|a, \boldsymbol{m}, \boldsymbol{C_i}) - E(Y_i|a^\star, \boldsymbol{m}, \boldsymbol{C_i})\} P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a^\star) \mathrm{d}\boldsymbol{m}. \end{aligned} \tag{2.2}$$

If we replace $a^\star$ with a in $E[Y_i(a, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}]$, then we get $E[Y_i(a, \boldsymbol{M_i}(a))|\boldsymbol{C_i}] = \int_{\boldsymbol{m}} E(Y_i|a, \boldsymbol{m}, \boldsymbol{C_i}) \times P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a) \mathrm{d}\boldsymbol{m}$, and thus the average indirect effect conditional on C is given by,

$$\begin{aligned} & E[Y_i(a, \boldsymbol{M_i}(a)) - Y_i(a, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}] \\ = & \int_{\boldsymbol{m}} E(Y_i|a, \boldsymbol{m}, \boldsymbol{C_i}) \{P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a) - P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a^\star)\} \mathrm{d}\boldsymbol{m}. \end{aligned} \tag{2.3}$$

Finally, one can get the average NDE and NIE by taking expectation over C of the two conditional effects defined in (2.2) and (2.3). Importantly, Equations (2.2), (2.3) show that, under the assumptions we made, the average NDE and the average NIE can be identified by modeling $Y_i|A_i, \boldsymbol{M_i}, \boldsymbol{C_i}$ and $\boldsymbol{M_i}|A_i, \boldsymbol{C_i}$ using observed data.

We note that as the main interest of this chapter lies in the joint effect of the whole set of mediators,thus the definition of NIE and NDE only involve the counterfactuals of the form $\boldsymbol{M_i}(a) = (M_i^{(1)}(a), M_i^{(2)}(a), ..., M_i^{(p)}(a))$. If one is interested in estimating the effect of a specific mediator, then one needs to consider the K^p counterfactuals $(M_i^{(1)}(a_1), M_i^{(2)}(a_2), \ldots, M_i^{(p)}(a_p))$, $a_1, a_2, ..., a_p \in \{1, 2, ..., K\}$. Characterizing mediator-specific NIE is a much more challenging task and requires stronger assumptions, in particular when the multiple mediators influence and interact with one another.

2.3 Models and Estimands

As discussed in Section 2.2, effects of mediators (average NDE and NIE) defined in terms of potential outcomes can be deduced from two conditional models for $Y_i|A_i, \boldsymbol{M_i}, \boldsymbol{C_i}$ and $\boldsymbol{M_i}|A_i, \boldsymbol{C_i}$ using observed data. Therefore, we propose two regression models for the two conditional relationships and subsequently deduce the causal effects of mediators. For modeling $Y_i|A_i, \boldsymbol{M_i}, \boldsymbol{C_i}$, we assume for subject i $(i = 1, ..., n)$, a continuous outcome of interest Y_i is associated with exposure A_i, p potential mediators $\boldsymbol{M_i} = (M_i^{(1)}, M_i^{(2)}, ..., M_i^{(p)})^T$ that may be on the pathway from A_i to Y_i, and q covariates $\boldsymbol{C_i}$ with the first element being the scalar 1 for the intercept:

$$Y_i = \boldsymbol{M_i^T}\boldsymbol{\beta_m} + A_i\beta_a + \boldsymbol{C_i^T}\boldsymbol{\beta_c} + \epsilon_{Yi} \tag{2.4}$$

where $\boldsymbol{\beta_m} = (\beta_{m1}, ..., \beta_{mp})^T$, $\boldsymbol{\beta_c} = (\beta_{c1}, ..., \beta_{cq})^T$, $\epsilon_{Yi} \sim N(0, \sigma_e^2)$. Here we assume there is no interaction between A_i and $\boldsymbol{M_i}$. Next for modeling $\boldsymbol{M_i}|A_i, \boldsymbol{C_i}$ we consider a multivariate regression model that jointly analyzes the p potential mediators:

$$\boldsymbol{M_i} = A_i\boldsymbol{\alpha_a} + \boldsymbol{\alpha_c}\boldsymbol{C_i} + \boldsymbol{\epsilon}_{Mi} \tag{2.5}$$

where $\boldsymbol{\alpha_a} = (\alpha_{a1}, ..., \alpha_{ap})^T$, $\boldsymbol{\alpha_c} = (\boldsymbol{\alpha_{c1}^T}, ..., \boldsymbol{\alpha_{cp}^T})^T$, $\boldsymbol{\alpha_{c1}}, ..., \boldsymbol{\alpha_{cp}}$ are q-by-1 vectors, $\boldsymbol{\epsilon}_{Mi} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma})$, $\boldsymbol{\Sigma}$ captures the correlation among the mediators. ϵ_{Yi} and $\boldsymbol{\epsilon}_{Mi}$ are assumed independent of A_i, $\boldsymbol{C_i}$ and each other.

With assumptions made in Section 2.2 and under the regression models specified for the outcome $E(Y_i|A_i, \boldsymbol{M_i}, \boldsymbol{C_i})$ and for the mediators $P(\boldsymbol{M_i}|A_i, \boldsymbol{C_i})$, we can analytically calculate the right-hand side of Equations (2.2), (2.3). We show in Supporting Information that the average NDE, NIE and TE can then be computed as below, and in the rest of the chapter,

we refer to NDE as direct effect and NIE as indirect/mediation effect.

$$\text{NDE:} \quad E[Y_i(a, \boldsymbol{M_i}(a^\star)) - Y_i(a^\star, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}] = \beta_a(a - a^\star). \tag{2.6}$$

$$\text{NIE:} \quad E[Y_i(a, \boldsymbol{M_i}(a)) - Y_i(a, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}] = (a - a^\star) \sum_{j=1}^{p} (\boldsymbol{\alpha_a})_j (\boldsymbol{\beta_m})_j. \tag{2.7}$$

$$\text{TE:} \quad E[Y_i(a) - Y_i(a^\star)|\boldsymbol{C_i}] = (\beta_a + \boldsymbol{\alpha_a^T \beta_m})(a - a^\star). \tag{2.8}$$

As noted in Equation (2.7), under the assumptions of model (2.4) the NIE through the whole set of mediators turns out to be the sum of the product of $(\boldsymbol{\alpha_a})_j$ and $(\boldsymbol{\beta_m})_j$ over the entire set. Those individual product terms do not correspond to the NIE of a specific (say j-th) mediator. We define active mediators as the ones with non-null contribution to the global NIE, i.e. $(\boldsymbol{\alpha_a})_j(\boldsymbol{\beta_m})_j$ being non-zero. The proposed Bayesian shrinkage and selection methods are used to identify and estimate these active components. Any inactive mediator will naturally fall into one of the following three categories: $(\boldsymbol{\beta_m})_j$ is non-zero while $(\boldsymbol{\alpha_a})_j$ is zero; $(\boldsymbol{\alpha_a})_j$ is non-zero while $(\boldsymbol{\beta_m})_j$ is zero; both are zero. Such a refined partition for the high-dimensional set of mediators provides useful and insightful interpretations for the structure of the composite null.

Regarding a global measure of the indirect effects, we note that the quantity in Equation (2.7), summation of each mediator's marginal mediation effect, is a good summary of the global mediation effects when the marginal mediation effect for each mediator is of the same direction. However, when marginal mediation effects have opposite directions, their effects may cancel out and result in a small or zero indirect effect. Considering this, we propose to use the L_2 norm of the vector of marginal mediation effects (*Huang and Pan*, 2016) as a

global measure of mediation effects, i.e.,

$$\begin{aligned}\tau &= ||((\boldsymbol{\alpha_a})_1(\boldsymbol{\beta_m})_1, (\boldsymbol{\alpha_a})_1(\boldsymbol{\beta_m})_2, ..., (\boldsymbol{\alpha_a})_p(\boldsymbol{\beta_m})_p)||^2 \\ &= \sum_{j=1}^{p}\{(\boldsymbol{\alpha_a})_j(\boldsymbol{\beta_m})_j\}^2. \end{aligned} \tag{2.9}$$

2.4 Bayesian Method for Estimation

2.4.1 Prior Specification

In order to conduct high-dimensional mediation analysis, we need to make certain model assumptions on the effect sizes. In genome-wide association studies, Bayesian sparse regression models, such as Bayesian variable selection regression models (BVSR), have been proven to yield better power in detecting relevant covariates (*Guan and Stephens*, 2011). For high-dimensional mediation analysis, we also make the reasonable sparsity assumption, which implies that only a small proportion of mediators mediate the exposure effects on the outcome. Linear mixed models (LMM), on the other hand, assume that every mediator transmits certain effects from exposure to outcome, with the effect sizes normally distributed. Here, we first assume that all the potential mediators contribute small, non-zero effects in mediating the exposure-outcome relationship, which is aligned with the main idea of polygenic (*Zhou et al.*, 2013) and omnigenic (*Boyle et al.*, 2017) models. In genetics settings, it has been widely argued that all genetic markers exhibit non-zero effect on disease related complex traits. The sum of those small effects in our models capture the overall contribution from the whole set of mediators. Besides these small effects, the sparsity assumption indicates that there is a small proportion of mediators exhibiting additional/large effects. We refer to these mediators with additional effects as active mediators, which is consistent with the concept of core genes defined in the omnigenic model. Therefore, in this chapter, we use the Baysian Sparse Linear Mixed Model (BSLMM), which imposes continuous shrinkage on the effects (*Zhou et al.*, 2013) and assumes the presence of small and additional effects,

for high-dimensional mediation analysis. The BSLMM is capable of learning the underlying mediation architecture from the data, producing good performances across a wide range of scenarios. Our model assumptions are also akin to the notion of quasi-sparsity that has become popular with continuous shrinkage priors (*Ge et al.*, 2019). Specifically, we assume a mixture of two normal components a priori for the jth mediator, $j = 1, 2, ...p$,

$$\begin{aligned}(\boldsymbol{\beta_m})_j &\sim \pi_m N(0, \sigma^2_{m1}) + (1 - \pi_m) N(0, \sigma^2_{m0}) \\ (\boldsymbol{\alpha_a})_j &\sim \pi_a N(0, \sigma^2_{ma1}) + (1 - \pi_a) N(0, \sigma^2_{ma0})\end{aligned}$$

where $\sigma^2_{m1} > \sigma^2_{m0}$, $\sigma^2_{ma1} > \sigma^2_{ma0}$, and π_m, π_a denote the proportion of coefficients that belong to the normal distribution with a larger variance.

For the other coefficients, we assume,

$$\beta_a \sim N(0, \sigma^2_a) \quad \text{and} \quad \boldsymbol{\beta_C}, \boldsymbol{\alpha_C} \sim MVN(\mathbf{0}, \sigma^2_c \boldsymbol{I}), \sigma^2_c \to \infty$$

Here we use a limiting normal prior for $\boldsymbol{\beta_C}, \boldsymbol{\alpha_C}$ with its variance going to infinity, since we often have insufficient information from the data to overwhelm any prior assumptions. For the convenience of modeling, we set the correlation structure among mediators $\boldsymbol{\Sigma}$ as $\sigma^2_g \boldsymbol{I}$.

For the hyper-parameters of variances in the model, we use the standard conjugate priors,

$$\begin{aligned}\sigma^2_{ms} &\sim \text{inverse-gamma}(k_{ms}, l_{ms}), s = 0, 1 \\ \sigma^2_a &\sim \text{inverse-gamma}(k_a, l_a) \\ \sigma^2_{mas} &\sim \text{inverse-gamma}(k_{mas}, l_{mas}), s = 0, 1 \\ \sigma^2_e &\sim \text{inverse-gamma}(k_e, l_e) \\ \sigma^2_g &\sim \text{inverse-gamma}(k_g, l_g)\end{aligned}$$

We set $k_{m0} = k_{m1} = k_a = k_{ma0} = k_{ma1} = k_e = k_g = 2.0$, and $l_{m0} = l_{ma0} = 10^{-4}, l_a = l_{m1} =$

$l_{ma1} = l_e = l_g = 1.0$. The prior inclusion probabilities π_m, π_a encode the prior information about the sparsity of the coefficients. We place a uniform prior on $\log(\pi_m), \log(\pi_a)$,:

$$\log(\pi_m), \log(\pi_a) \sim U(\log(1/p), \log(1))$$

where p is the number of mediators. The priors were chosen so that π_m and π_a range from $1/p$ to 1, and the lower and upper bounds correspond to an expectation of 1 and p covariates in each model. A uniform prior on $\log(\pi_m)$ and $\log(\pi_a)$ reflects the fact that the uncertainty in π_m, π_a spans orders of magnitude due to the sparsity of the models. We do not choose a uniform prior on π_m, π_a since that would put appreciable prior probability on large numbers of covariates (*Guan and Stephens*, 2011).

2.4.2 Posterior Sampling Algorithm

We develop a Markov chain Monte Carlo (MCMC) sampling algorithm to obtain the posterior samples from our Bayesian method. To facilitate MCMC, we introduce indicator variables $\boldsymbol{r_m}, \boldsymbol{r_a} \in \{0,1\}^p$ to indicate which normal component $(\boldsymbol{\beta_m})_j$ and $(\boldsymbol{\alpha_a})_j$ come from, and for the jth mediator, $r_{mj} = \mathrm{I}((\boldsymbol{\beta_m})_j \sim N(0, \sigma^2_{m1})), r_{aj} = \mathrm{I}((\boldsymbol{\alpha_a})_j \sim N(0, \sigma^2_{ma1}))$. Let $\boldsymbol{\theta_1} = (\boldsymbol{\beta_m}, \beta_a, \pi_m, \boldsymbol{r_m}, \sigma^2_{m1}, \sigma^2_{m0}, \sigma^2_a, \sigma^2_e)$ denote all the unknown parameters in model (2.4), and $\boldsymbol{\theta_2} = (\boldsymbol{\alpha_a}, \pi_a, \boldsymbol{r_a}, \sigma^2_{ma1}, \sigma^2_{ma0}, \sigma^2_g)$ for model (2.5). The joint log posterior distribution is,

$$\begin{aligned} & \log P(\boldsymbol{\theta_1}, \boldsymbol{\theta_2} | (Y_i, \boldsymbol{M_i}, A_i)_{i=1}^n) \\ \propto & \sum_{i=1}^n \log P(Y_i | \boldsymbol{\theta_1}, A_i, \boldsymbol{M_i}) + \sum_{i=1}^n \log P(\boldsymbol{M_i} | \boldsymbol{\theta_2}, A_i) + \log P(\boldsymbol{\theta_1}) + \log P(\boldsymbol{\theta_2}) \end{aligned}$$

We use a Hastings-within-Gibbs algorithm to obtain posterior samples, and full details of the sampling algorithm appear in Supporting Information.

For the jth mediator, we can estimate the posterior probability of both $(\boldsymbol{\beta_m})_j$ and $(\boldsymbol{\alpha_a})_j$ being in the normal components with larger variances as the posterior inclusion probability

	$(\boldsymbol{\beta_M})_j$	Larger component	Smaller component
$(\boldsymbol{\alpha_A})_j$			
Larger component		$r_{mj} * r_{aj} = 1$ (Group 1)	$r_{mj} = 0, r_{aj} = 1$ (Group 2)
Smaller component		$r_{mj} = 1, r_{aj} = 0$ (Group 3)	$r_{mj} = r_{aj} = 0$ (Group 4)

Table 2.1: Mediators are categorized into four groups based on their relationships with exposure and outcome; Group 1: Both $(\boldsymbol{\beta_M})_j$ and $(\boldsymbol{\alpha_A})_j$ come from larger normal components; Group 2: $(\boldsymbol{\alpha_A})_j$ from larger normal component while $(\boldsymbol{\beta_M})_j$ from smaller normal component; Group 3: $(\boldsymbol{\beta_M})_j$ from larger normal component while $(\boldsymbol{\alpha_A})_j$ from smaller normal component; Group 4: Both $(\boldsymbol{\beta_M})_j$ and $(\boldsymbol{\alpha_A})_j$ come from smaller normal components.

(PIP), defined as $P(r_{mj} = 1, r_{aj} = 1|\text{Data})$ in our model. Mediators with larger $(\boldsymbol{\beta_M})_j$ and $(\boldsymbol{\alpha_A})_j$ tend to be categorized into larger variance normals, and such tendency can be quantified by the mediator's PIP. PIP provides non-null evidence for both $(\boldsymbol{\beta_M})_j$ and $(\boldsymbol{\alpha_A})_j$, and therefore, we select mediators with the highest PIP as potentially active mediators.

2.4.3 Mediator Categorization

Under the above Bayesian mediation framework, active mediators are the ones whose $(\boldsymbol{\beta_M})_j$ and $(\boldsymbol{\alpha_A})_j$ both come from larger normal components. The three categories for the inactive mediators are: $(\boldsymbol{\beta_M})_j$ from larger normal component while $(\boldsymbol{\alpha_A})_j$ from smaller normal component; $(\boldsymbol{\alpha_A})_j$ from larger normal component while $(\boldsymbol{\beta_M})_j$ from smaller normal component; both from smaller components. In addition to identifying true mediators, our method automatically classifies all the mediators into four groups based on their relationship with exposure and outcome. In practice, we have the indicator variables r_{mj} and r_{aj} to denote which component the coefficients $(\boldsymbol{\beta_M})_j$, $(\boldsymbol{\alpha_A})_j$ belong to and can easily obtain the posterior probabilities for each group. The four groups are illustrated in Table 2.1,

2.5 Simulations

We evaluate the performance of the proposed Bayesian mediation method and compare it with other existing mediation methods in simulations. The three existing frequentist meth-

ods include single mediation analysis, multivariate mediation analysis and high-dimensional multivariate mediation (HDMM) methodology of *Chén et al.* (2017). Single mediation analysis tests one mediator at a time for its mediation effect. We use the R package **mediation** to run single mediation analysis with the nonparametric bootstrap option for standard error estimation. Multivariate mediation analysis (*VanderWeele and Vansteelandt*, 2014), on the other hand, jointly analyzes all the mediators in both model (2.4) and (2.5) and tests the product term $(\boldsymbol{\beta_m})_j(\boldsymbol{\alpha_a})_j$ for each j at a time while controlling for all other variables. This method can only be fit when a multivariate ordinary least squares regression model can be fit for the outcome model (2.4). We implement the multivariate mediation analysis and compute the standard error based on the delta method as $se((\hat{\boldsymbol{\beta}}_{\boldsymbol{m}})_j(\hat{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j) = \sqrt{(\hat{\boldsymbol{\beta}}_{\boldsymbol{m}})_j^2 Var((\hat{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j) + (\hat{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j^2 Var((\hat{\boldsymbol{\beta}}_{\boldsymbol{m}})_j)}$. Afterwards, we obtain a z-statistics for the jth mediator by dividing $(\hat{\boldsymbol{\beta}}_{\boldsymbol{m}})_j(\hat{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j$ with its standard error and compute the corresponding p-value based on asymptotic normality. The HDMM is a novel method recently developed for high-dimensional mediation analysis and aims to identify active mediators through dimension reduction techniques. We use p-values for univariate and multivariate mediation analysis, estimated indirect effect for HDMM and PIP for our Bayesian method as measures of the evidence for mediation. We compare the power to identify active mediators based on either 5% or 10% false discovery rate (FDR).

We consider various simulation settings with $n = 1,000$ samples and p mediators ($p = 100$ or 2,000). Since the multivariate mediation analysis can only be applied to settings where the number of mediators is smaller than the number of observations (i.e. $p < n$), we first examine the settings of $p = 100$ in order to include the multivariate mediation analysis for comparison. We will later consider the high-dimensional setting of $p = 2,000$. For each simulation setting, we first simulate a set of continuous exposure variables $\{A_i, i = 1, ..., 1000\}$ independently from a standard normal distribution. We then generate a p-vector of mediators for the ith individual from $\boldsymbol{M_i} = A_i\boldsymbol{\alpha_a} + \boldsymbol{\epsilon}_{Mi}$. Each element of $\boldsymbol{\alpha_a}$, $(\boldsymbol{\alpha_a})_j$ $(j = 1, ..., p)$, is simulated from a point-normal prior: $\pi_a N(0, 1) + (1 - \pi_a)\delta_0$, where δ_0 is

a point mass at zero. The residual errors $\boldsymbol{\epsilon}_{Mi}$ are simulated from a multivariate normal distribution with mean zero and a covariance $\boldsymbol{\Sigma}$. $\boldsymbol{\Sigma}$ accounts for the correlation among mediators commonly seen in real data, and we use the sample covariance estimated from the Multi-Ethnic Study of Atherosclerosis (MESA) data to serve as $\boldsymbol{\Sigma}$. Because our Bayesian mediation model does not explicitly account for the correlation structure of mediators in the model between mediators and exposure, the simulations with correlated mediators allow us to examine the robustness of our modeling assumption regarding independence. We scale the two terms $A_i\boldsymbol{\alpha_a}$ and $\boldsymbol{\epsilon}_{Mi}$ further so that the former explains a fixed proportion of variance: $PVE_A = Var(A_i\boldsymbol{\alpha_a})/Var(\boldsymbol{M_i})$, where Var denotes the sample variance.

Given the exposure and mediators, we then generate the outcome Y_i from the linear model: $Y_i = \boldsymbol{M_i^T}\boldsymbol{\beta_m} + A_i\beta_a + \epsilon_{Yi}$. Here, each element of $\boldsymbol{\beta_m}$, $(\boldsymbol{\beta_m})_j$ $(j = 1, ..., p)$, is simulated from $\pi_m N(0, 1) + (1-\pi_m)\delta_0$, and β_a from a standard normal distribution. The residual error ϵ_{Yi} is simulated independently from a standard normal distribution. We assume that only 10% of the mediators are truly mediating the exposure effects on the outcome (i.e. active mediators), whose $(\boldsymbol{\beta_m})_j$ and $(\boldsymbol{\alpha_a})_j$ are both sampled from the large variance normal distribution. After simulating $\boldsymbol{M_i^T}\boldsymbol{\beta_m}$, $A_i\beta_a$ and ϵ_{Yi}, we scale these three terms further to achieve two desirable PVEs: $PVE_{IE} = Var(\boldsymbol{\alpha_a^T}\boldsymbol{\beta_m}A_i)/Var(Y_i)$ and $PVE_{DE} = Var(A_i\beta_a)/Var(Y_i)$.

To explore a variety of simulation scenarios, we first examine a baseline scenario where we set $PVE_A = 0.5, PVE_{IE} = 0.4, PVE_{DE} = 0.1, \pi_a = 0.3, \pi_m = 0.2$. We then vary each of the four parameters $(PVE_A, PVE_{IE}, \pi_a, \pi_m)$ one at a time to investigate their individual influences on the results. We perform 200 replicates for each scenario to do the power comparison.

We first examine the settings for $p = 100$ and display the comparative results in Figure 2.2. The results show that our Bayesian multivariate mediation method outperforms the other three methods in all scenarios. For example, in the baseline setting, at 10% FDR, Bayesian mediation method achieves a power of 0.725, while the univariate and multivariate methods,

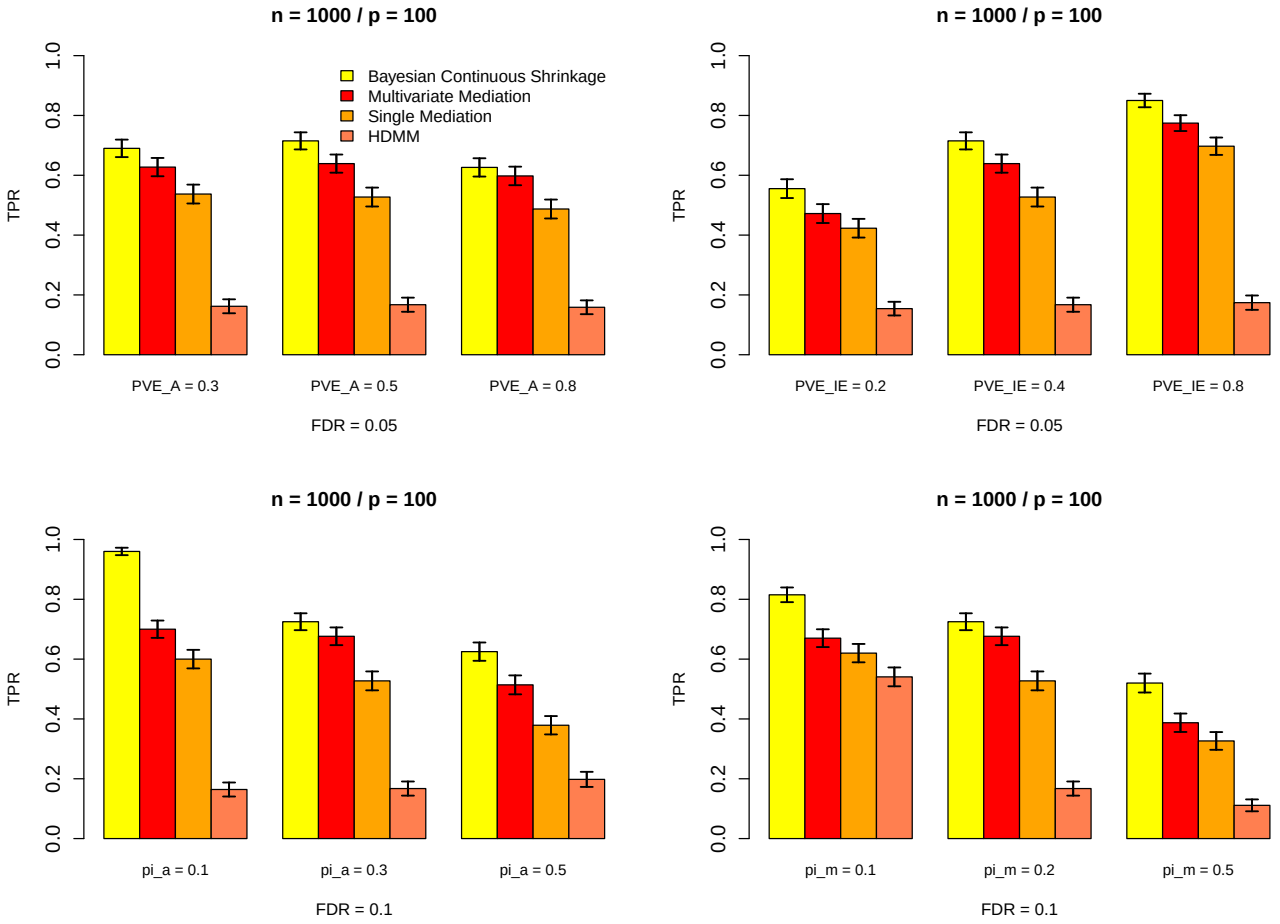

Figure 2.2: Power comparison among our Bayesian mediation method (yellow), multivariate mediation method (red), single mediation method (orange) and HDMM (coral) when the number of mediators is 100 and sample size 1,000. The x-axis marks the one parameter we change at a time from the baseline setting. The average TPR at FDR = 0.05/0.1 and its error bar based on ± 2 standard errors are calculated across 200 replicates. The standard error of the proportions are computed from a binomial distribution.

and HDMM achieve a power of 0.527, 0.676 and 0.167, respectively. The power of the four approaches increases with increasing PVE_{IE}, which increases the effect sizes of $\boldsymbol{\beta_m}$. In addition, the power of most approaches reduces with increased π_a or π_m, which reduces the effect sizes of either $\boldsymbol{\alpha_a}$ or $\boldsymbol{\beta_m}$, respectively. As one would expect, the advantage of our Bayesian method over the univariate and multivariate methods is more apparent in sparse settings with smaller values of π_a and π_m. In terms of PVE_A, which determines the effect size of $\boldsymbol{\alpha_a}$, we found that the power of different methods first increases slightly when PVE_A changes from 0.3 to 0.5 and then decreases slightly as PVE_A changes further to 0.8. The later decrease in power in the setting of $PVE_A = 0.8$ is presumably due to the increased correlation between the exposure and mediators, which makes it difficult for all the methods to distinguish between direct and indirect effects in model (2.4). The performance of HDMM is relatively stable to PVE_A, PVE_{IE} and π_a, and improves slightly with increased π_a. There is no sparsity assumption on mediation effects in HDMM and thus it may not fare well in relatively sparse simulations. Between the two competing methods of single and multivariate mediation analysis method, the latter yields better power than the former in all scenarios, as the multivariate mediation analysis properly controls for the correlation among mediators.

Next, we examine the settings for $p = 2,000$. Now we select 1% of the mediators to be active and set $\pi_m = 2\%$, $\pi_a = 3\%$ as the baseline setting with all other configurations being same as in the baseline setting of $p = 100$. Since the multivariate mediation analysis is unfeasible when $p > n$, we compare our method with single mediation analysis and HDMM. We use a threshold of 1% false positive rate (FPR) instead of false discovery rate due to low power in the $p = 2,000$ settings. The comparisons are shown in Figure 2.3. The Bayesian mediation method yields more power than the single mediator analysis and HDMM in all the scenarios. For example, in the baseline setting, at 1% FPR, Bayesian mediation method achieves a power of 0.470, while the univariate method and HDMM have a power of 0.357 and 0.248, respectively. The power of our method and the univariate approach again increases with increasing PVE_{IE} and decreases with increasing π_a or π_m. Increasing PVE_A decreases the

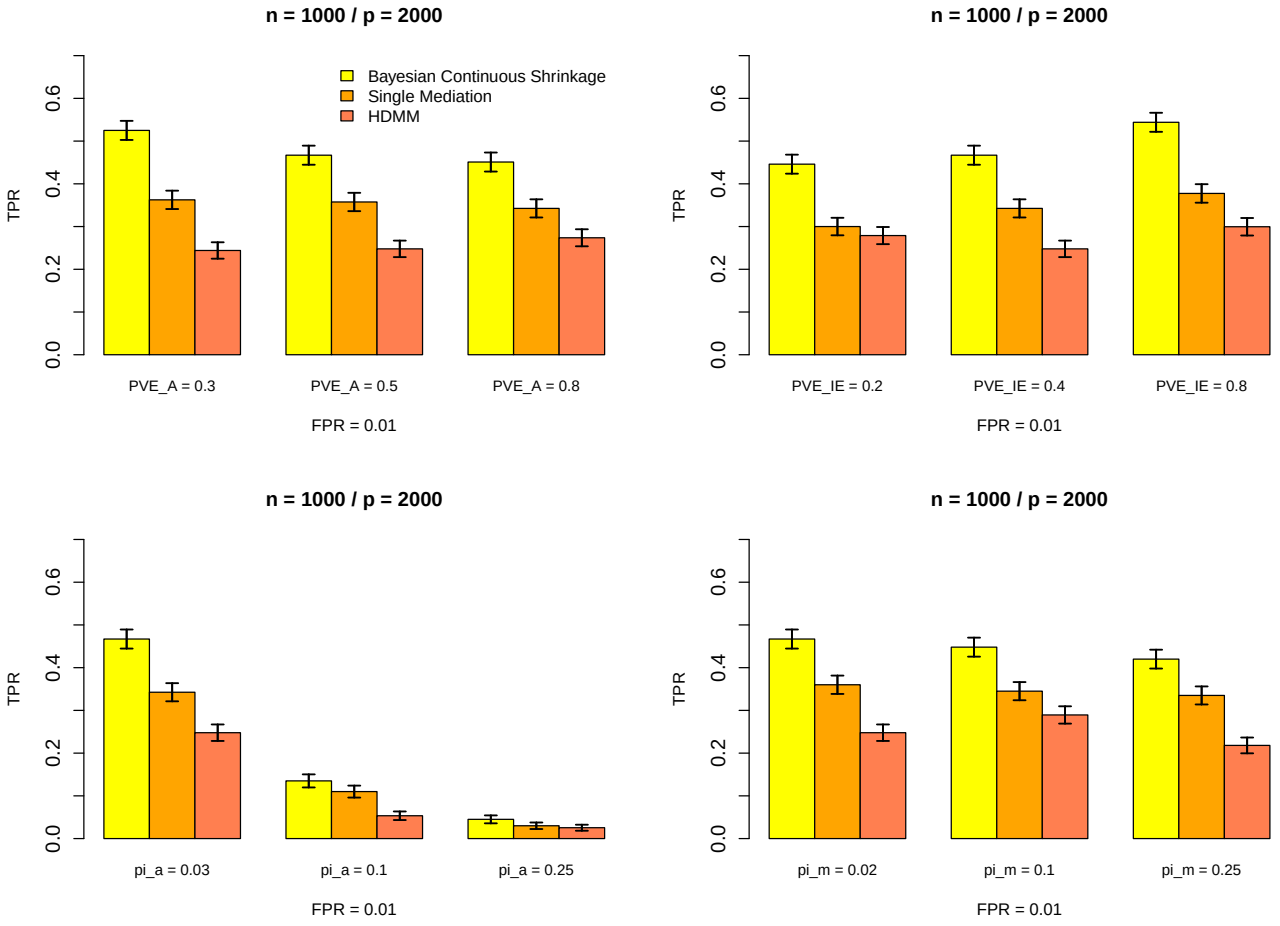

Figure 2.3: Power comparison between our Bayesian mediation method (yellow), single mediation method (orange) and HDMM (coral) when the number of mediators is 2,000 and sample size 1,000. The x-axis marks the one parameter we change at a time from the baseline setting. The average TPR at FPR = 0.01 and its error bar based on ± 2 standard errors are calculated across 200 replicates. The standard error of the proportions are computed from a binomial distribution.

power of the Bayesian method, while tends to improve the performance of HDMM possibly due to the dimension reduction applied on the high-dimensional mediators. Comparing the settings with varied π_a, we note that the major power gain of our method lies in the joint analysis of mediators in the outcome model and appropriate shrinkage on the vector of $\boldsymbol{\beta_m}$. For the mediator model, we are essentially fitting a series of regression models for each mediator and exposure. Therefore, shrinkage on the vector of $\boldsymbol{\alpha_a}$ does not help much in mediator selection, especially if π_a is relatively large, e.g. 0.1 or 0.25. Also, when $\pi_m = 0.1$, the true number of non-zero $\boldsymbol{\beta_m}$ gets closest to the reduced dimension of $\boldsymbol{M}$ after generalized population value decomposition (PVD) in HDMM, which leads to its power improvement.

In the above high-dimensional settings, besides a mixture of normals prior, horseshoe prior and spike-and-slab prior (*Mitchell and Beauchamp*, 1988; *Carvalho et al.*, 2010) are also commonly used prior specifications for Bayesian shrinkage, and it is natural to apply them separately to the two regression models in high-dimensional mediation analysis. The comparative results are presented in Table A.1 and A.2 in Supporting Information. Our Bayesian mediation method performs well in identifying active mediators in a variety of scenarios, and remains decent power when the effects are not polygenic as we assume. Our method can distinguish the large effects from the small effects in the polygenic model, and is also more practically appealing than the horseshoe prior since it can directly categorize mediators into four possible groups without the need of specifying a thresholding rule.

Finally, we examine the ability of our method to estimate the overall mediation effects and the proportion of mediators in the four different categories as shown in Table 2.1. We use π_{g1}, π_{g2}, π_{g3}, π_{g4} to represent the proportion of mediators in Group 1, Group 2, Group 3 and Group 4, respectively. We examine eight different simulation scenarios based on different combinations of π_{g1}, π_{g2}, π_{g3} and π_{g4}, which include four null scenarios with $\pi_{g1} = 0$ and four alternative scenarios with $\pi_{g1} \neq 0$. In these simulations, we set PVEs to be the same as in the baseline setting ($PVE_A = 0.5, PVE_{IE} = 0.4, PVE_{DE} = 0.1$; except when $\pi_{g4} = 1$ where PVE_A and PVE_{IE} are zero). We provide the estimated global mediation effects (τ) and

p	π_{g1}	π_{g2}	π_{g3}	π_{g4}	$\widehat{\pi_{g1}}$	(95% CI)	τ	$\widehat{\tau}$	(95% CI)
100	0	0.2	0.1	0.7	0.003	(0.000, 0.010)	0	0.025	(0.006, 0.066)
	0	0.1	0.2	0.7	0.003	(0.000, 0.010)	0	0.035	(0.006, 0.117)
	0	0.1	0.1	0.8	0.001	(0.000, 0.012)	0	0.009	(0.000, 0.012)
	0	0	0	1	0.001	(0.000, 0.008)	0	0.000	(0.000, 0.000)
100	0.1	0.2	0.1	0.6	0.064	(0.010, 0.092)	0.128	0.194	(0.100, 0.324)
	0.1	0.1	0.2	0.6	0.058	(0.030, 0.080)	0.249	0.263	(0.159, 0.400)
	0.1	0.1	0.1	0.7	0.078	(0.006, 0.110)	0.110	0.172	(0.014, 0.960)
	0.1	0	0	0.9	0.051	(0.040, 0.063)	0.961	0.458	(0.215, 0.792)
2,000	0	0.03	0.02	0.95	0.000	(0.000, 0.000)	0	0.230	(0.074, 0.565)
	0	0.1	0.02	0.88	0.000	(0.000, 0.000)	0	0.315	(0.149, 0.644)
	0	0.1	0.1	0.8	0.000	(0.000, 0.000)	0	0.211	(0.075, 0.509)
	0	0	0	1	0.000	(0.000, 0.000)	0	0.000	(0.000, 0.000)
2,000	0.01	0.02	0.01	0.96	0.001	(0.000, 0.003)	1.392	0.642	(0.408, 0.941)
	0.01	0.04	0.01	0.94	0.001	(0.000, 0.004)	1.273	0.436	(0.156, 0.891)
	0.01	0.09	0.09	0.81	0.001	(0.000, 0.003)	0.347	0.544	(0.217, 1.074)
	0.01	0	0	0.99	0.010	(0.002, 0.015)	0.103	0.113	(0.018, 0.334)

Table 2.2: Estimation of the global mediation effects τ under different compositions. We report the posterior mean ($\widehat{\tau}$) of τ and its 95% credible intervals when $p = 100/2,000$. We denote π_{g1}, π_{g2}, π_{g3} and π_{g4} to represent the proportion of mediators in Group 1, Group 2, Group 3 and Group 4 as defined in Table 2.1, and $\hat{\pi}_{g1}$ is the estimated proportion of active mediators from our Bayesian method. We also provide the 95% credible intervals (CI) for $\widehat{\tau}$ and $\hat{\pi}_{g1}$.

proportion of active mediators (π_{g1}), as well as their 95% credible intervals from posterior samples in Table 2.2.

From the above table, we find that our method provides decent estimates for π_{g1} and τ across different scenarios, especially when $p = 100$. Note that our estimates for π_{g1} are slightly conservative due to the fact that our model does not have full power to detect all the mediators. The 95% credible intervals of τ also shows that the posterior distribution of τ is asymmetric and depends on the composition of the four groups. We also show a distribution graph from the posterior samples of τ in four different scenarios with $n = 1000, p = 100$ in Figure A.2 of Supporting Information.

2.6 Data Analysis

We applied the proposed Bayesian method to investigate the mediation mechanism of DNAm in the pathway from adult socioeconomic status (SES) to glycated hemoglobin (HbA1c) in the Multi-Ethnic Study of Atherosclerosis (MESA) (*Bild et al.*, 2002). The exposure, adult SES, is indicated by adult educational attainment and is an important risk factor for cardiovascular diseases. The outcome, HbA1c, is a surrogate measurement of average blood glucose levels and a critical variable for various diseases including T2D and CVD (*Selvin et al.*, 2010). Thus, understanding how methylation at different CpG sites mediates the effects of adult SES on HbA1c can shed light on the molecular mechanisms of CVD. We provide our detailed processing steps for MESA data in the SI. Briefly, we selected 1,231 individuals with both adult SES and HbA1c measurements as well as DNA methylation profiles measured from purified monocytes. Due to computational reasons, we focused on a final set of 2,000 CpG sites that have the strongest marginal associations with adult SES for the following mediation analysis.

We applied both univariate mediation analysis and our Bayesian multivariate mediation analysis to analyze the selected 2,000 CpG sites. For the multivariate analysis, we consider

$$Y_i = \boldsymbol{M_i^T \beta_m} + A_i \beta_a + \boldsymbol{C_{2i}^T \beta_c} + \epsilon_{Yi} \tag{2.10}$$

$$\boldsymbol{M_i} = A_i \boldsymbol{\alpha_a} + \boldsymbol{\alpha_c C_{1i}} + \boldsymbol{\epsilon}_{Mi} \tag{2.11}$$

where Y_i represnts HbA1c levels; A_i represents adult SES values; and $\boldsymbol{M_i}$ represnts methylation level for 2,000 CpG sites. In Equation (2.10), the model controls for age, gender and race/ethnicity, and in Equation (2.11), we adjust for age, gender, race/ethnicity and enrichment scores for 4 major blood cell types (neutrophils, B cells, T cells and natural killer cells). All the continuous variables are standardized to have zero mean and unit variance. The univariate analysis is applied in a similar fashion except that it is used to analyze one

site at a time.

We discussed the identifiability assumptions required for causal inference in high-dimensional mediation analysis in Section 2.2. For Assumption 1 (no unmeasured exposure-outcome confounding), we believe that the available covariates of age, gender and race/ethnicity in model 10 are sufficient to control for confounders that are associated both to adult SES and HbA1c (*Burke et al.*, 2008). For Assumption 2 (no unmeasured mediator-outcome confounding given exposure), within each of the two exposure groups (low adult SES and high adult SES), age, gender and race/ethnicity are also sufficient to control for mediator-outcome confounding. For Assumption 3 (no unmeasured exposure-mediator confounding), we included all the important potential exposure-mediator confounders (age, gender, race/ethnicity and enrichment scores for 4 major blood cell types) in model (2.11) as in *Needham et al.* (2015). Assumption 4 (no mediator-outcome confounding affected by the exposure) is very difficult and may not hold, but it is only required for identifying the natural effects so we don't need to worry too much about it. The influence of violating the above identifiability assumptions can be assessed using sensitivity analysis, which has been developed for the single mediator setting (*Imai et al.*, 2010b), and additional work is required to extend that to the high-dimensional setting. Regarding the temporal assumptions, in MESA, adult SES (exposure) was collected in Exam 1 between July 2000 and August 2002, and DNAm (mediators) and HbA1c (outcome) were assessed in Exam 5, and all of them are one-time measurements. While it is hard to disentangle the temporality between DNAm and HbA1c measurements, we note that DNAm levels are relatively stable over time and thus a DNAm measurement at Exam 5 could be highly correlated with a DNAm measurement at an earlier stage.

We display PIP values for each of the 2,000 CpG sites from the Bayesian multivariate analysis in Figure 2.4. Two CpG sites were identified with strong evidence (PIP > 0.5) for mediating the adult SES effects on HbA1c. They are also among the top ten sites with the smallest p-values obtained from univariate mediation analysis. In addition, these two CpG sites are close to genes *CCDC54* and *CCND2*, both of which are known candidates associated with

HbA1c. Specifically, the expression of *CCND2* has been shown to be associated with risk of T2D and the related glycemic traits of glucose, HbA1c, and insulin (*Yaghootkar et al.*, 2015). The gene *CCDC54* interacts with valproic acid and acrylamide, both of which are associated with diabetes and blood insulin (*Lin et al.*, 2009). Therefore, strong evidence from our method suggests that adult SES may act through these two genes to affect HbA1c. We also apply the HDMM and Bayesian methods with spike-and-slab and horseshoe priors to the data. For the HDMM, the weights for the first direction of mediation do not suggest obvious signal or pattern, and a plot of the estimated weights is provided in Figure A.3. We also list the top 2 sites and their nearby genes from the other competing methods in Table A.3. We note that there is a lack of biological evidence to support a mediating role of those genes picked out by the other methods, except for one gene (CLU).

In addition, we estimate the global mediation effects $\hat{\tau}$ as 0.0084 and its 95% credible interval from the posterior as (0.0063, 0.0115). The $\widehat{PVE}_{IE}$ is 0.096, indicating that approximately 10% of the outcome variance is indirectly explained by DNAm after controlling for covariates. In addition, we estimate the proportion of CpG sites in each of the four categories as defined in Section 2.4.3: $\hat{\pi}_{g1} = 0.002$, $\hat{\pi}_{g2} = 0.031$, $\hat{\pi}_{g3} = 0.001$, $\hat{\pi}_{g4} = 0.966$. We find that a small proportion of DNAm has large effects on the HbA1c level, and a small proportion of DNAm is notably associated with adult SES. The results also suggest that adult SES acts through certain important DNAm sites to influence HbA1c. Finally, we perform a posterior predictive check on the outcome model, and the results are shown in Figure A.4. The Bayesian predictive p-values (*Neelon et al.*, 2010) are 0.5 and 0.45 for the sample mean and variance, respectively, suggesting adequate fit of the outcome model.

2.7 Discussion

In this chapter, we develop a Bayesian sparse linear mixed model for high-dimensional mediation analysis. The advantage of a Bayesian method is to propagate uncertainty for functions

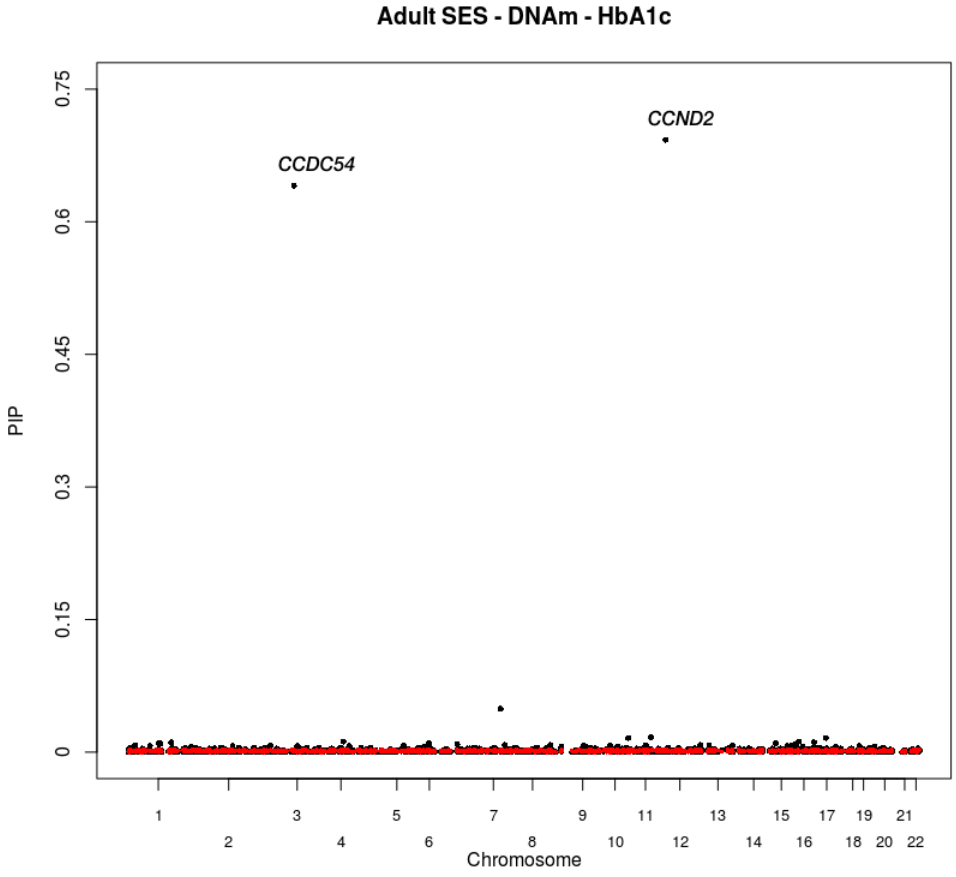

Figure 2.4: Consider the trio: Adult SES → DNAm → HbA1c. The black dots are the estimated posterior inclusion probability (PIP) for each CpG site from the Bayesian mediation method and the red dots are the estimated PIPs when we permute the outcome once and fit the Bayesian mediation method.

of parameters in a natural way instead of resorting to Delta methods or two-step approaches. Our method can jointly analyze a large number of unordered mediators and characterize their global mediation effect without making any assumptions on their joint distribution. By imposing continuous shrinkage priors on the key regression coefficients for the mediation analysis, our method achieves up to 30% power gain in identifying true non-null mediators compared with univariate mediation method and approximately 10% power gain compared with multivariate method based on simulations. The Bayesian method also provides better interpretations of the way in which a mediator links or does not link exposure to outcome, and automatically categorize mediators into four components based on exposure-mediator and mediator-outcome relationship. Implementing our method to MESA, we have identified two genes, *CCDC54* and *CCND2*, with strong evidence for actively mediating the adult SES effects on HbA1c. Both of them are candidate genes associated with diabetes and blood insulin.

Although our proposed method can simultaneously analyze high-dimensional mediators, like other posterior sampling based methods, the computation speed is not fast due to the large number of sampling iterations required for reasonable convergence. Also, throughout the chapter, we focus on one continuous outcome of interest. For binary outcome, we can treat it as a quantitative trait, which is justified by recognizing the linear model as a first order Taylor approximation to a generalized linear model (*Zhou et al.*, 2013). One may hope to adapt our method to directly model the nonlinear outcome through link functions within a generalized linear model framework, but such approach will substantially increase the computational cost and may sometimes not bring much power gain, as shown in *Zhou et al.* (2013). Future development of new algorithms or methods will likely be required to scale our method to handle thousands of individuals and millions of mediators in generalized regression models.

Recent literature proposes a convex penalty on the product term of indirect effect (*Zhao and Luo*, 2016), which improves power of pathway selection and reduces estimation bias in

the indirect effects. Under the Bayesian framework, direct shrinkage on the product term may be a more appropriate choice, as it takes into account the correlation between the two models in the mediation analysis and is more straightforward when the goal is to identify non-null mediators. Directly incorporating the correlation between the mediators will be another avenue to pursue. In addition, the biological annotations like pathways can be important predictors for the underlying mediation mechanism, and integrating them into high-dimensional mediation analysis would be promising to facilitate the identification of active mediators. Possible extensions include linking the functional annotation information for mediators either to the mediator-specific group probabilities, e.g. π_{mj}, π_{aj} for the j-th mediator through a logistic regression model (*Carbonetto and Stephens*, 2013), or to the effect sizes, e.g. $\sigma^2_{m1j}, \sigma^2_{ma1j}$ for the j-th mediator (*Hao et al.*, 2018). We leave these interesting extensions for future work.

CHAPTER III

Bayesian Sparse Mediation Analysis with Targeted Penalization of Natural Indirect Effects

3.1 Introduction

This chapter describes an extension to the Bayesian shrinkage approach for high-dimensional mediation analysis in Chapter III. Recall in Chapter III, two Bayesian Sparse Linear Mixed Models (BSLMM) are proposed to impose continuous shrinkage on the outcome-mediator and mediator-exposure effect, separately. In this chapter, we develop alternative approaches in order to directly target the non-null indirect effects to identify active mediators within the same mediation framework, and we compare the proposed methods with a different set of competing methods from the previous chapter.

With the rapid development of high-throughput technologies and more availability of larger-scale omics data, there is expanding interest in mediation analysis with a large number of mediators. For example, *Huang and Pan* (2016) and *Chén et al.* (2017) transform the high-dimensional unordered set of mediators into lower-dimensional orthogonal components using dimension reduction techniques. The extracted low-dimensional components are then analyzed through single mediation analysis. However, it is often not straightforward to interpret the low-dimensional components in these approaches. Shrinkage methods via regularization have also been explored to tackle this high-dimensional regression problem involving two

models, the exposure-mediator model and the outcome-exposure model. The Lasso (*Tibshirani*, 1996) penalty can be naturally applied to the two models in mediation analysis. *Zhang et al.* (2016) also proposed a regularized regression with minimax concave penalty for the outcome model after a sure independence screening on mediators. The above methods penalize the mediator-outcome and exposure-mediator coefficients separately without taking into account the structure of the indirect effect. To directly target the mediators with strong indirect effects, *Zhao and Luo* (2016) recently developed a new convex, Lasso-type penalty on the indirect effect, which is the product of the two path coefficients. This direct penalization on the pathway effects is shown to improve power for mediator selection and reduce the estimation bias of indirect effects. In addition to frequentist approaches, Bayesian non-parametric models (*Kim et al.*, 2017, 2019) have been applied in the analysis with single or a moderate number of mediators. *Song et al.* (2018) handles high-dimensional mediators through a Bayesian variable selection method and specifies separate shrinkage priors on both the exposure-mediator effects and mediator-outcome effects. However, not modeling the indirect effects in a targeted way may lead to loss of power for selection of active mediators.

The indirect effect of a mediator is known to be proportional to the product of the exposure-mediator and mediator-outcome effects under certain assumptions (*MacKinnon*, 2008). Testing for this product term is not easy due to the complexity in its null distribution. Recent literature began to recognize and leverage the composite structure in the null hypothesis of no indirect effect in the genome-wide mediation analysis setting, where a one-at-a-time single mediator analysis is performed across the entire set of mediators (*Huang et al.*, 2019). Given a large number of mediators, we can characterize the composite space and learn about the structure of mediation through the four components arising from the product of the two effects, i.e. one component of mediators with non-zero indirect effects (active mediators), and three components with zero indirect effects.

Motivated by the goal of directly targeting the non-null indirect effects to identify active me-

diators, we are interested in seeking the Bayesian parallel with a joint prior on the exposure-mediator and mediator-outcome coefficients, which is so far lacking in the literature. One common choice of the bivariate prior would be a Gaussian prior, and it is natural to assume a four-component Gaussian mixture structure on the two effects, corresponding to the composite structure underlying their product. On the other hand, a direct thresholding prior on the indirect effects would also achieve the same goal, and we can extend the hard-thresholding priors (*Ni et al.*, 2019; *Cai et al.*, 2020) to product thresholding for mediation analysis. Therefore, in this chapter, building on the potential outcome framework for causal inference, we develop two novel prior models for high-dimensional mediation analysis: (a) four-component Gaussian mixture prior, and (b) product threshold Gaussian prior. Both models can simultaneously analyze a large number of mediators without making any path-specific or causal ordering assumptions on mediators. The mediator categorization into four groups provides useful interpretations on the way in which a mediator links or does not link exposure to outcome. More importantly, by jointly modeling the exposure-mediator and mediator-outcome coefficients via either bivariate Gaussian distributions or thresholding functions, we place direct shrinkage on the product of the two coefficients, i.e. indirect effect, in a targeted way. Hence, our methods are expected to outperform other penalization methods that apply separate shrinkage in the two regression models independently, for identifying active mediators with non-zero indirect effects.

The proposed methods are generally applicable to many settings, and we examine their performance for both large-scale genomic and environmental data. Due to fast advances in high-throughput biological technologies, genomic studies can nowadays measure large-scale molecular-level traits such as gene expression and DNA methylation (DNAm) levels. Recent studies have proposed these molecular traits may act as a mechanism through which neighborhood disadvantages affect physical health (*Smith et al.*, 2017). Our methods are implemented for a high-dimensional mediation analysis with DNAm as mediators in the Multi-Ethnic Study of Atherosclerosis (MESA) (*Bild et al.*, 2002), focusing on the relation-

ship between neighborhood disadvantage and body mass index (BMI). BMI is a critical risk factor for various diseases like type 2 diabetes (T2D) and cardiovascular disease (CVD) (*Hjellvik et al.*, 2012), and the important scientific discoveries will advance our biological understanding of disease etiology. Another study we examine is the LIFECODES prospective birth cohort, and our primary aim is to evaluate the mediating role of endogenous biomarkers of lipid metabolism, inflammation, and oxidative stress in the association between prenatal exposure to environmental contamination and pregnancy outcomes. Our methods are applied to identify multiple active biomarker mediators and reveal important biological pathways between toxicant exposure and preterm birth, which are useful for early detection and prevention of disease in pregnancy. Besides the data analysis, we also perform extensive simulation studies under different structures of effects. We show through both simulations and data analysis that our proposed methods can increase power of a joint analysis and enable efficient identification of individual mediators.

3.2 Method

We adopt the same notations and counterfactual framework as in Chapter II. We also make the same set of identifiability assumptions as discussed in Section 2.2, including the consistency assumption and four non-unmeasured confounding assumptions (*VanderWeele and Vansteelandt*, 2014). It has been shown that under the required assumptions, the average NDE and NIE can be identified by modeling $Y_i|A_i, \boldsymbol{M}_i, \boldsymbol{C_i}$ and $\boldsymbol{M}_i|A_i, \boldsymbol{C_i}$ using observed data (*Song et al.*, 2018). Therefore we still work with the two conditional models for $Y_i|A_i, \boldsymbol{M}_i, \boldsymbol{C_i}$ and $\boldsymbol{M}_i|A_i, \boldsymbol{C_i}$, and subsequently deduce the causal effects of mediators. For the outcome model, we assume

$$Y_i = \boldsymbol{M}_i^\top \boldsymbol{\beta_m} + A_i\beta_a + \boldsymbol{C_i}^\top \boldsymbol{\beta_c} + \epsilon_{Yi}, \tag{3.1}$$

where $\boldsymbol{\beta_m} = (\beta_{m1}, \ldots, \beta_{mp})^\top$; $\boldsymbol{\beta_c} = (\beta_{c1}, \ldots, \beta_{cq})^\top$; and $\epsilon_{Yi} \sim \mathrm{N}(0, \sigma_e^2)$. For the mediator model, we consider a multivariate regression model that jointly analyzes all p potential mediators together as dependent variables:

$$\boldsymbol{M}_i = A_i \boldsymbol{\alpha_a} + \boldsymbol{\alpha_C} \boldsymbol{C_i} + \boldsymbol{\epsilon}_{Mi}, \tag{3.2}$$

where $\boldsymbol{\alpha_a} = (\alpha_{a1}, \ldots, \alpha_{ap})^\top$; $\boldsymbol{\alpha_c} = (\boldsymbol{\alpha_{c1}}^\top, \ldots, \boldsymbol{\alpha_{cp}}^\top)^\top$; $\boldsymbol{\alpha_{c1}}, \ldots, \boldsymbol{\alpha_{cp}}$ are q-by-1 vectors; $\boldsymbol{\epsilon}_{Mi} \sim \mathrm{MVN}(\mathbf{0}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma}$ capturing potential residual error covariance. ϵ_{Yi} and $\boldsymbol{\epsilon}_{Mi}$ are assumed to be independent of each other and independent of A_i and $\boldsymbol{C_i}$. With the identifiability assumptions and the modeling assumptions (linearity, no interaction in the outcome and mediator model) in (3.1)-(3.2), we can compute the average NDE, NIE and TE as in (2.6)-(2.8). Particularly, the NIE is proportional to $\sum_{j=1}^{p} \alpha_{aj}\beta_{mj}$. Therefore, the marginal indirect contribution from the j-th mediator is the product of α_{aj} and β_{mj}. We propose to jointly model β_{mj} and α_{aj} and perform targeted shrinkage on the NIE using two prior models described in the following sections.

3.2.1 Gaussian Mixture Model (GMM)

The first model we develop to characterize the composite structure of the exposure-mediator and mediator-outcome effects in mediation analysis and induce targeted shrinkage on NIE is the four-component Gaussian mixture model. Mixture models have been studied vastly for classifying subjects into different categories and inferring their association patterns or category-specific properties (*Zeng et al.*, 2018; *Cai et al.*, 2019). In the context of mediation analysis, previous mixture model approaches have primarily been proposed in the form of a principal stratification model (*Gallop et al.*, 2009), or focused on grouping individuals by their covariate profiles for estimating the heterogeneous effects (*Kim et al.*, 2018). Here, we introduce a Gaussian mixture model for the joint modeling of β_{mj} and α_{aj} and the subsequent inference of the composite association patterns. Specifically, we consider four components

in the Gaussian mixture model: a component representing $\beta_{mj}\alpha_{aj} \neq 0$, that both β_{mj} and α_{aj} are non-zero; a component representing $\beta_{mj} \neq 0$ and $\alpha_{aj} = 0$; a component representing $\beta_{mj} = 0$ and $\alpha_{aj} \neq 0$; and a component representing $\beta_{mj} = 0$ and $\alpha_{aj} = 0$. To characterize the composite structure underlying the product $\beta_{mj}\alpha_{aj}$, we assume that the effects for each mediator follow a four-component Gaussian mixture distribution as below,

$$[\beta_{mj}, \alpha_{aj}]^\top | \{\boldsymbol{V_k}\}_{k=1}^3 \sim \pi_1 \text{MVN}_2(\mathbf{0}, \boldsymbol{V_1}) + \pi_2 \text{MVN}_2(\mathbf{0}, \boldsymbol{V_2}) + \pi_3 \text{MVN}_2(\mathbf{0}, \boldsymbol{V_3}) + \pi_4 \boldsymbol{\delta_0}$$

with prior probabilities π_k $(k = 1, 2, 3, 4)$ summing to one and MVN_2 denoting a bivariate normal distribution. Here, π_1 represents the prior probability of being an active mediator, with non-zero marginal mediation effect $\beta_{mj}\alpha_{aj}$; and $\boldsymbol{V_1}$ models the covariance of $[\beta_{mj}, \alpha_{aj}]^\top$ in model (3.1) and (3.2) when both effects are non-zero. Any inactive mediator will fall into one of the remaining three components. π_2 is the prior probability of having non-zero mediator-outcome effect but zero exposure-mediator effect; and $\boldsymbol{V_2} = \begin{bmatrix} \sigma_2^2 & 0 \\ 0 & 0 \end{bmatrix}$ is a low-rank covariance matrix restricting that only the effect of mediator on outcome β_{mj} is non-zero. π_3 is the prior probability of having non-zero exposure-mediator effect but zero mediator-outcome effect; and $\boldsymbol{V_3} = \begin{bmatrix} 0 & 0 \\ 0 & \sigma_3^2 \end{bmatrix}$ is a low-rank covariance matrix restricting that only the effect of exposure on mediator α_{aj} is non-zero. Lastly, π_4 denotes the prior probability of zero mediator-outcome effect and zero exposure-mediator effect; and $\boldsymbol{\delta_0}$ is a point mass at zero. Our method automatically classifies all the mediators into four groups based on their relationship with exposure and outcome. We note that the recently developed Bayesian mediation analysis method (BAMA, *Song et al.* (2018)) can be viewed as a two-component version of GMM: in BAMA, the mediator-outcome effect is non-zero and follows a normal distribution with probability $\pi_1 + \pi_2$; while the exposure-mediator effect is non-zero and follows another normal distribution with probability $\pi_1 + \pi_3$. Consequently, the active mediator in BAMA has *a priori* probability $(\pi_1+\pi_2)(\pi_1+\pi_3)$, which is determined by

the non-zero exposure-mediator effect probability and the non-zero mediator-outcome effect probability.

In GMM, we specify a conjugate inverse-Wishart prior on $\boldsymbol{V_1}$, $\boldsymbol{V_1} \sim$ Inv-Wishart($\boldsymbol{\Psi_0}, \nu$), where $\boldsymbol{\Psi_0} = \text{diag}\{\psi_{01}, \psi_{02}\}$ is a diagonal matrix, and ν is the degrees of freedom, and inverse-gamma priors on σ_2^2, σ_3^2, $\sigma_2^2 \sim$ Inv-Gamma($\nu/2, \psi_{01}/2$), $\sigma_3^2 \sim$ Inv-Gamma($\nu/2, \psi_{02}/2$). We also assume $\{\pi_1, \pi_2, \pi_3, \pi_4\} \sim$ Dirichlet(a_1, a_2, a_3, a_4) with a_1, a_2 and a_3 set to be smaller than a_4 to encourage sparsity of the first three components. For the coefficients of the other covariates, we assume $\beta_a \sim \text{N}(0, \sigma_a^2)$ and $\boldsymbol{\beta_C}, \boldsymbol{\alpha_{C1}}, ..., \boldsymbol{\alpha_{Cp}} \sim \text{MVN}(\boldsymbol{0}, \sigma_c^2 \boldsymbol{I})$. Since we often have inadequate information from the data to infer $\boldsymbol{\beta_C}$ and $\boldsymbol{\alpha_C}$, we simply use a limiting prior by setting $\sigma_c^2 \to \infty$. For the convenience of modeling, we also set the correlation structure among mediators $\boldsymbol{\Sigma}$ as $\sigma_g^2 \boldsymbol{I}$. We use weakly informative inverse-gamma priors on the variance hyper-parameters (σ_a^2, σ_e^2 and σ_g^2) in the models.

To facilitate computation, for the jth mediator, we create a membership indicator variable γ_j, where $\gamma_j = k$ if $[\beta_{mj}, \alpha_{aj}]^\top$ is from normal component $k, k = 1, 2, 3, 4$. Since the priors used here are all conjugate, we implement a standard Gibbs sampling algorithm and iterate each mediator one at a time to obtain posterior samples. The full details of the algorithm appear in the Appendix. With the Gibbs sampling, for the j-th mediator, we can estimate its indirect effect as the product of the posterior mean of β_{mj} and α_{aj}. We also calculate the posterior probability of both β_{mj} and α_{aj} being non-zero as the posterior inclusion probability (PIP), which is $P(\gamma_j = 1|$ Data). The PIP provides evidence for a non-zero indirect effect, and therefore, we identify mediators with the highest PIP as potentially active mediators.

3.2.2 Product Threshold Gaussian (PTG) Prior

Although the GMM model is flexible for a range of applications, the method does not directly impose sparsity on $\beta_{mj}\alpha_{aj}$ for mediator selection. To address this issue, we develop a product threshold Gaussian (PTG) prior for the indirect effects of the $j$the mediator. Threshold priors

have been recently proposed for Bayesian variable selection. For example, *Ni et al.* (2019) introduced a hard-thresholding mechanism in edge selection for sparse graphical structure; *Cai et al.* (2020) performed a feature selection over networks using the threshold graph Laplacian prior; and *Kang et al.* (2018) developed a soft-thresholding Gaussian process for scalar-on-image regression. As compelling alternatives to shrinkage priors, the threshold priors are equivalent to the non-local priors (*Rossell and Telesca*, 2017) which enjoy appealing theoretical properties and excellent performance in variable selection for high-dimensional regression, especially when the predictors are strongly correlated(*Kang et al.*, 2018; *Cai et al.*, 2020). In this work, we extend the threshold priors to the product threshold priors for mediation analysis. In particular, for the bivariate vector $(\beta_{mj}, \alpha_{aj})$, $j = 1, ..., p$,

$$\begin{aligned} \beta_{mj} &= \tilde{\beta}_{mj} \max\left\{ I\left(|\tilde{\beta}_{mj}| > \lambda_1\right), I\left(|\tilde{\beta}_{mj}\tilde{\alpha}_{aj}| > \lambda_0\right)\right\} \\ \alpha_{aj} &= \tilde{\alpha}_{aj} \max\left\{ I\left(|\tilde{\alpha}_{aj}| > \lambda_2\right), I\left(|\tilde{\beta}_{mj}\tilde{\alpha}_{aj}| > \lambda_0\right)\right\} \end{aligned}$$

where the underlying un-thresholded effects $(\tilde{\beta}_{mj}, \tilde{\alpha}_{aj})^\top \sim \text{MVN}_2(0, \boldsymbol{\Sigma}_{\boldsymbol{u}})$ and $I(\mathcal{A})$ is the indicator function with $I(\mathcal{A}) = 1$ if $\mathcal{A}$ occurs and $I(\mathcal{A}) = 0$ otherwise. We denote $(\beta_{mj}, \alpha_{aj}) \sim \text{PTG}(\boldsymbol{\Sigma}_{\boldsymbol{u}}, \lambda)$ with $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ being thresholding parameters.

As one may note, a mediator would escape thresholding and have non-zero indirect effect $\beta_{mj}\alpha_{aj}$ only when (i) both the absolute values of the marginal effects $\tilde{\beta}_{mj}$ and $\tilde{\alpha}_{aj}$ are larger than the threshold values, or (ii) the absolute value of the un-thresholded indirect effect $\tilde{\beta}_{mj}\tilde{\alpha}_{aj}$ is larger than the threshold value. In practice, condition (ii) does not necessarily indicate condition (i). The product threshold prior will facilitate the selection of active mediators by thresholding on the indirect effects in addition to the marginal effects, and shrinking insignificant effects to zero. Similar to GMM, one group of active mediators and three groups of inactive ones are naturally formed. The thresholding on the product term also adds dependency between β_{mj} and α_{aj}, and we impose no more dependency on the un-thresholded values, namely setting $\boldsymbol{\Sigma}_{\boldsymbol{u}} = \text{diag}\left\{\tau_\beta^2, \tau_\alpha^2\right\}$ in the rest of this chapter.

The threshold parameters $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ control *a priori* the sparsity of the non-zero effects, and larger values tend to produce a smaller subset of active mediators. Previous literature (*Ni et al.*, 2019; *Cai et al.*, 2020) have considered uniform priors on those threshold parameters, e.g. $\lambda_0 \sim U[0, \lambda_{0h}]$, $\lambda_1 \sim U[0, \lambda_{1h}]$, $\lambda_2 \sim U[0, \lambda_{2h}]$, with the upper bounds $\lambda_{0h}, \lambda_{1h}, \lambda_{2h}$ being some pre-defined large values. This approach is straightforward and requires little prior knowledge, however, the control of false positives is a concern due to the common under-estimation of λ. In this chapter, we instead determine the threshold parameters from the un-thresholded distributions and the desired number of declared positives, and fix them *a priori*. For example, if we set $\lambda_0 = 0.36, \lambda_1 = \lambda_2 = 0.6$ under $\tau_\beta^2 = 0.1, \tau_\alpha^2 = 0.1$, then the Monte Carlo estimate of the prior proportion of active mediators is approximately 0.01, which could also be tuned to match with π_1 in the Gaussian mixture model. In practice, we can grid search the three hyper-parameters together with priors on τ_β^2 and τ_α^2, and find the values that achieve desired prior proportions. The thresholds λ can also be interpreted as the minimal detectable signal, and determined based on their practical meaning. Although the resulting selection may be conservative and heavily informed by the pre-defined thresholds, our specification is helpful in guarding against false positive findings. As in the GMM model described in 3.2.1, conjugate inverse-gamma priors are used for the variance terms ($\tau_\beta^2, \tau_\alpha^2, \sigma_e^2$ and σ_g^2) in the model. The full conditional distributions for β_{mj} and α_{aj} are mixtures of truncated normals and can be sampled from Gibbs sampling. The full algorithm appears in the Appendix. Similar to GMM, we can calculate the posterior mean of β_{mj} and α_{aj}, and the posterior probability of both β_{mj} and α_{aj} being non-zero as PIP, and use the PIP to rank and select active mediators.

The proposed GMM relies on small values of π_1, π_2, π_3 to admit sparsity on the effects. The Gaussian priors shrink the effects continuously toward zero, and help the model achieve better estimation and prediction performance, but not necessarily mediator selection by the indirect effects. On the other hand, the PTG utilizes a hard threshold function to directly select on the product term $\beta_{mj}\alpha_{aj}$ and map near zero effects to zero. Instead of centering

around zero, the effects produced from PTG will be similar to truncated normals away from zero. As a practical procedure, we suggest median inclusion probabilities (PIP = 0.5) as the significance threshold for mediator selection.

3.2.3 Other Approaches for High-dimensional Mediation Analysis

Besides GMM and PTG, we also explore a few other approaches. Many of them place simple penalty functions or shrinkage priors on the natural indirect effects.

Univariate Mediation Analysis is perhaps the simplest approach to perform mediation analysis. In univariate mediation analysis, we examine one mediator at a time and test whether the mediator has non-zero indirect effect. We extract P-values for testing the indirect effects using the R package **mediation**.

Bi-Lasso The least absolute shrinkage and selection operator (Lasso) introduced by *Tibshirani* (1996) is a widely used penalty function to perform both variable regularization and selection. Here, we consider placing Lasso regularization on the mediator-outcome effects and the exposure-mediator effects separately. For the mediator-outcome effects, we attempt to minimize the following loss function based on the outcome model (3.1): $f(\boldsymbol{\beta_m}, \beta_a, \boldsymbol{\beta_c}) = \frac{1}{2}\sum_{i=1}^{n}(Y_i - \boldsymbol{M}_i^\top \boldsymbol{\beta_m} - A_i\beta_a - \boldsymbol{C_i}^\top \boldsymbol{\beta_c})^2 + \lambda_1 \sum_{j=1}^{p}|\beta_{mj}|$. For the exposure-mediator effects, we attempt to minimize the following loss function based on the mediator model (3.2): $f(\boldsymbol{\alpha_a}, \boldsymbol{\alpha_c}) = \frac{1}{2}\sum_{i=1}^{n}(\boldsymbol{M}_i - A_i\boldsymbol{\alpha_a} - \boldsymbol{\alpha_c}\boldsymbol{C_i})^\top(\boldsymbol{M}_i - A_i\boldsymbol{\alpha_a} - \boldsymbol{\alpha_c}\boldsymbol{C_i}) + \lambda_2 \sum_{j=1}^{p}|\alpha_{aj}|$. We perform optimization in the first function using the R package **glmnet** and perform optimization in the second function using soft-thresholding. We choose the two tuning parameters $\lambda_1 > 0$ and $\lambda_2 > 0$ through 10-fold cross validation in the two functions separately. We refer this approach of applying Lasso separately to the outcome and mediator models as Bi-Lasso.

Bi-Bayesian Lasso is effectively the Bayesian version of Bi-Lasso. It is equivalent to placing a Bayesian Lasso prior (*Park and Casella*, 2008) on the mediator-outcome effects $\boldsymbol{\beta_m}$

and a separate Bayesian Lasso prior on the exposure-mediator effects $\boldsymbol{\alpha_a}$. Here, we specify the Bayesian Lasso prior for the j-th element of $\boldsymbol{\beta_m}$ or $\boldsymbol{\alpha_a}$ as a scale mixture of normal distributions $\mathrm{N}(0, z_j\sigma_z^2)$, where the scale parameter z_j follows an exponential distribution $\exp(s^2/2)$ and $1/s^2$ is given a diffuse inverse-gamma prior. We implement the Bi-Bayesian Lasso using a Gibbs sampler following *Park and Casella* (2008) and obtain posterior samples for $\boldsymbol{\beta_m}$ and $\boldsymbol{\alpha_a}$.

Pathway Lasso is a method developed by *Zhao and Luo* (2016) for high-dimensional mediation analysis under the linear structural equation modeling (LSEM) framework. To see how Pathway Lasso works, we first define the squared-error loss in the joint model from Equations (3.1) and (3.2) as $l(\boldsymbol{\beta_m}, \boldsymbol{\alpha_a}, \beta_a, \boldsymbol{\beta_c}, \boldsymbol{\alpha_c}) = \sum_{i=1}^n (Y_i - \boldsymbol{M}_i^\top \boldsymbol{\beta_m} - A_i\beta_a - \boldsymbol{C_i}^\top \boldsymbol{\beta_c})^2 + \sum_{i=1}^n (\boldsymbol{M}_i - A_i\boldsymbol{\alpha_a} - \boldsymbol{\alpha_c}\boldsymbol{C_i})^\top (\boldsymbol{M}_i - A_i\boldsymbol{\alpha_a} - \boldsymbol{\alpha_c}\boldsymbol{C_i})$. The Pathway Lasso then aims to minimize the following penalized function,

$$\begin{aligned} f(\boldsymbol{\beta_m}, \boldsymbol{\alpha_a}, \beta_a, \boldsymbol{\beta_c}, \boldsymbol{\alpha_c}) &= \frac{1}{2} l(\boldsymbol{\beta_m}, \boldsymbol{\alpha_a}, \beta_a, \boldsymbol{\beta_c}, \boldsymbol{\alpha_c}) + \lambda[\sum_{j=1}^p (|\beta_{mj}\alpha_{aj}| + \phi(\beta_{mj}^2 + \alpha_{aj}^2)) + |\beta_a|] \\ &\quad + \omega[\sum_{j=1}^p (|\beta_{mj}| + |\alpha_{aj}|)], \phi \geq 1/2 \\ &= \frac{1}{2} l(\boldsymbol{\beta_m}, \boldsymbol{\alpha_a}, \beta_a, \boldsymbol{\beta_c}, \boldsymbol{\alpha_c}) + \lambda P_1(\boldsymbol{\beta_m}, \boldsymbol{\alpha_a}, \beta_a) + \omega P_2(\boldsymbol{\beta_m}, \boldsymbol{\alpha_a}), \end{aligned} \tag{3.3}$$

In Equation (3.3), the first penalty term P_1 stabilizes and shrinks the estimates for the indirect effects $\beta_{mj}\alpha_{aj}$. The second penalty term P_2 provides additional shrinkage on $\boldsymbol{\beta_m}$ and $\boldsymbol{\alpha_a}$ through a common Lasso penalty placed on both of them. We use the algorithm from *Zhao and Luo* (2016) to fit Pathway Lasso. We choose the three tuning parameters (ϕ, ω, and λ): $\phi = 2$, $\omega = 0.1\lambda$, and choose λ through 10-fold cross-validation as in the original paper.

HIMA is another frequentist method developed for high-dimensional mediation analysis (*Zhang et al.*, 2016). HIMA first applies a sure independence screening to the outcome

model to select a small set of potential mediators. With the selected mediators, HIMA then places a minimax concave penalty on the mediator-outcome effects in the outcome model (3.1) to obtain effect estimates. The method finally performs a joint significance test and rejects the null hypothesis of no indirect effect with the j-th mediator if both β_{mj} and α_{aj} are significant. Using the HIMA software, we obtain the Bonferroni corrected P-values for testing the indirect effects.

In addition to the aforementioned methods, we note that several other approaches exist. For example, the methods developed by *Huang and Pan* (2016) and *Chén et al.* (2017) first perform dimension reduction on the mediators to extract low dimensional factors on the reduced dimensional space, and then carry out mediation analysis by treating the low dimensional factors as new mediators. Because these approaches analyze the latent factors instead of the original mediators, we do not compare our methods with them in the present study.

3.3 Simulations

3.3.1 Simulation Design

We consider one small sample scenario with $n = 100$, $p = 200$, and one large sample scenario with $n = 1000$, $p = 2000$. In both scenarios, we set the proportions of the four different mediator groups to be $\pi_1 = 0.05, \pi_2 = 0.05, \pi_3 = 0.10, \pi_4 = 0.80$. In each scenario, we further explore two different settings. In Setting (I), we fix the non-zero effects of both β_{mj} and α_{aj} to be 0.5, with their signs randomly assigned as positive or negative. In Setting (II), we fix 40% of the non-zero β_{mj} (or α_{aj}) to be 0.3, 30% of them to be 0.5, and 30% of them to be 0.7, with their signs randomly assigned as positive or negative. In both settings, we simulate the continuous exposure $\{A_i, i = 1, ..., n\}$ independently from a standard normal distribution N(0, 1). We simulate the residual error ϵ_{Yi} in the outcome model independently from N(0,

1), and simulate the residual errors $\boldsymbol{\epsilon}_{Mi}$ in the mediator model from MVN($\mathbf{0}$, $\boldsymbol{\Sigma}$). Here, we use the sample covariance estimated from MESA data to serve as $\boldsymbol{\Sigma}$ in the simulations. Afterwards, we generate a p-vector of mediators for the ith individual from $\boldsymbol{M}_i = A_i\boldsymbol{\alpha_a}+\boldsymbol{\epsilon}_{Mi}$. We also generate the outcome Y_i for the ith individual from $Y_i = \boldsymbol{M}_i^\top\boldsymbol{\beta_m} + A_i\beta_a + \epsilon_{Yi}$, with $\beta_a = 0.5$.

In the above settings, we have fixed the effect sizes to specific values across replicates. To further examine the performance of our methods over a wide range of effect sizes, we perform additional simulations where we simulate $[\beta_{mj}, \alpha_{aj}]^\top$ differently in each simulation replicate. Specifically, we generate these two effects from three different joint distributions detailed below (Figure B.1): the first two correspond to the prior distributions assumed in PTG and GMM, respectively, while the last one is a horseshoe distribution, i.e.

(A) Simulate effects under the PTG model: $[\beta_{mj}, \alpha_{aj}]^\top \sim \text{PTG}(\text{diag}\{\sigma_u^2, \sigma_u^2\}, \lambda)$, where $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ are set to satisfy the desired proportions of the four groups (π_1, π_2, π_3, π_4). We set $\sigma_u^2 = 0.3$ for $p = 200$, and $\sigma_u^2 = 0.1$ for $p = 2000$.

(B) Simulate effects under the GMM model: $[\beta_{mj}, \alpha_{aj}]^\top \sim \pi_1\text{MVN}(\mathbf{0}, \begin{bmatrix} \sigma^2 & \sigma^2/3 \\ \sigma^2/3 & \sigma^2 \end{bmatrix}) + \pi_2\text{MVN}(\mathbf{0}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix}) + \pi_3\text{MVN}(\mathbf{0}, \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2 \end{bmatrix}) + \pi_4\boldsymbol{\delta}_0$. We set $\sigma^2 = 0.3$ for $p = 200$, and $\sigma^2 = 0.1$ for $p = 2000$.

(C) Simulate effects from a mixture of bivariate horseshoe distributions, which can be generated from a scale mixture of normals: $[\beta_{mj}, \alpha_{aj}]^\top \sim \pi_1\text{MVN}(\mathbf{0}, Z_j^2 \begin{bmatrix} \sigma^2 & \sigma^2/3 \\ \sigma^2/3 & \sigma^2 \end{bmatrix}) + \pi_2\text{MVN}(\mathbf{0}, Z_j^2 \begin{bmatrix} \sigma^2 & 0 \\ 0 & 0 \end{bmatrix}) + \pi_3\text{MVN}(\mathbf{0}, Z_j^2 \begin{bmatrix} 0 & 0 \\ 0 & \sigma^2 \end{bmatrix}) + \pi_4\boldsymbol{\delta}_0$. Here, $Z_j \sim \text{halfCauchy}(0, 1)$, but truncated at a value of b to avoid impractically large values. We set $\sigma^2 = 0.5$ for $p = 200$, and $\sigma^2 = 0.3$ for $p = 2000$, and $b = 3$. Note that the effect size distribution assumed here is different from either of our proposed models, thus allowing us to study the robustness of

our methods. With the effect size distributions, we follow the same procedure described as in the fixed effects settings and perform 200 simulation replicates for each scenario. The effect sizes used in the simulations here are chosen to ensure a reasonable level of selection power for different methods. In real data, the variance explained by the exposure in the mediator model (PVE_A) and the variance explained by the mediators in the outcome model (PVE_{IE}) is likely to be lower than what we set in the simulations. For example, in Setting (A), PVE_{IE} is around 0.85 and PVE_A around 0.01, while in the LIFECODES data analysis, the estimated PVE_{IE} is 0.75 and the estimated PVE_A is 0.005.

We apply different methods to fit the simulated data. In GMM, we set the Dirichlet parameters $a_1 = 0.01p$, $a_2 = a_3 = 0.05p$, $a_4 = 0.89p$. We use an empirical Bayesian approach to set the diagonal entries of $\boldsymbol{\Psi}_0$, which are set to be the sample variance of the non-zero $\boldsymbol{\beta_m}$ and $\boldsymbol{\alpha_a}$ fitted through Lasso. We set the degree of freedom ν in the inverse-Wishart distribution to be two, which makes the distribution reasonably non-informative while still well-defined. In PTG, we set the pre-defined minimal detectable effect sizes $(\lambda_0, \lambda_1, \lambda_2)$ to be the 90% quantiles of the estimated $|\boldsymbol{\beta_m}|$ and $|\boldsymbol{\alpha_a}|$ fitted through Lasso. To be consistent with the GMM, we choose the parameter $\hat{\tau}^2$ in the priors $\tau_\beta^2 \sim \text{IG}(1.1, \hat{\tau}^2), \tau_\alpha^2 \sim \text{IG}(1.1, \hat{\tau}^2)$ to ensure that the prior inclusion probability is around 0.01. For example, in Setting (I) and (II), we set $\lambda_0 = 0.15, \lambda_1 = \lambda_2 = 0.4$, and $\hat{\tau}^2 = 0.01$. Note that although the prior means of τ_β^2 and τ_α^2 are small, their prior variances do not exist, indicating that such less informative priors would not affect the posterior inference on effect sizes much. For the Bayesian methods, we perform 150,000 iterations and discard the first 100,000 iterations as burn-in. We check the MCMC convergence by running five chains with random initial values and calculating the potential scale reduction factor (PSRF) for the PIPs. All the PSRFs fall within (1.0, 1.2), indicating the convergence of our algorithms.

3.3.2 Evaluation Metrics

We evaluate the performance of the two proposed methods (GMM and PTG) and compare them with existing methods in different simulation scenarios. As described in Section 3.2, we consider a total of eight methods: one univariate method and seven multivariate methods that include four Bayesian methods (GMM, PTG, BAMA and Bi-Bayesian Lasso) and three frequentist methods (Bi-Lasso, Pathway Lasso, and HIMA). We examine the power of different methods to detect true mediators in the simulations. To do so, we rely on PIP to prioritize mediators in PTG, GMM and BAMA; rely on P-value to rank mediators in the univariate method and HIMA; and rely on the estimated indirect effects as an measure of evidence for mediation for the remaining methods. We calculate the true positive rate (TPR) based on a fixed false discovery rate (FDR) of 10% for the selection accuracy. For estimation, we measure the mean square error (MSE) for the indirect effects of the truly active mediators ($\text{MSE}_{\text{non-null}}$), and MSE for the indirect effects of the truly inactive mediators (MSE_{null}).

Under Setting (A)-(C), in each replicate, we will have a different set of underlying true effects and truly active mediators, therefore the reported average metrics in those settings, such as the average true positive rate, are expectations taken over the distribution of corresponding random variables. Specifically, we calculate the MSE of the indirect effect for active mediators as

$$\text{MSE}_{\text{non-null}} = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{|\mathcal{C}(k)|}\sum_{j\in\mathcal{C}(k)}((\beta_{mj})^{(k)}(\alpha_{aj})^{(k)} - (\hat{\beta}_{mj})^{(k)}(\hat{\alpha}_{aj})^{(k)})^2 \tag{3.4}$$

where K is the number of replicates. For the k-th replicate, $\mathcal{C}(k)$ is the set of active mediators, $(\beta_{mj})^{(k)}$ and $(\alpha_{aj})^{(k)}$ are the true effects of the j-th mediator, $(\hat{\beta}_{mj})^{(k)}$ and $(\hat{\alpha}_{aj})^{(k)}$ are the estimated effects. This is the expected value over the distribution of $(\boldsymbol{\beta_m}, \boldsymbol{\alpha_a})$ and $\mathcal{C}(k)$.

Similarly, the MSE of the indirect effect for inactive mediators is calculated as

$$\text{MSE}_{\text{null}} = \frac{1}{K}\sum_{k=1}^{K}\frac{1}{|\mathcal{C}_0(k)|}\sum_{j\in\mathcal{C}_0(k)}((\beta_{mj})^{(k)}(\alpha_{aj})^{(k)} - (\hat{\beta}_{mj})^{(k)}(\hat{\alpha}_{aj})^{(k)})^2 \tag{3.5}$$

where $\mathcal{C}_0(k)$ is the set of inactive mediators for the k-th replicate.

3.3.3 Simulation Results: Setting (I)-(II)

Table 3.1 and 3.2 shows the results under the fixed effects for the small sample scenario $n = 100, p = 200$ and the large sample scenario $n = 1000, p = 2000$, respectively. Overall, our proposed methods, GMM and PTG, outperform the other methods. These two methods achieve the highest area under the ROC curve (AUC) and are up to $\sim 30\%$ more powerful than the other methods in identifying active mediators, with performance gain more apparent in the large sample scenario. Under Setting (I) where the mediation effects are large, the PTG method has the highest average TPR for both small and large sample settings. The performance of PTG is followed by GMM and BAMA. In contrast, under Setting (II) where the mediation effects are uneven, PTG may fail to identify some of the active mediators with small effects due to the thresholding set by the pre-defined parameter λ. Instead, GMM performs the best and its performance is followed by PTG and BAMA. Importantly, median inclusion probabilities (PIP = 0.5) in both GMM and PTG can be used as a criterion to declare active mediators (details in the Appendix), producing decent empirical estimates for FDR in simulations (Table B.1, B.2). Among the frequentist methods, the Bi-Lasso performs best over the others and is also competitive in the small sample setting. HIMA and the univariate method are among the worst methods for mediator selection, presumably because neither models the entire set of mediators jointly in the outcome model.

In terms of the effects estimation, GMM has the lowest $\text{MSE}_{\text{non-null}}$ across most simulation scenarios. Due to hard thresholding, PTG tends to provide a conservative list of the active mediators. Consequently, the non-zero indirect effects of some active mediators are shrunk

to zero in PTG, leading to relatively high $\text{MSE}_{\text{non-null}}$ but small MSE_{null} by PTG. Meanwhile, we find that the Pathway Lasso does not appear to exhibit much advantage over the simple alternative Bi-Lasso. Indeed, Pathway Lasso requires multiple tuning parameters for inducing the penalty term on the indirect effects, and those tuning parameters may benefit from more careful specifications than the algorithm default setting. The univariate method in particular has a quite high MSE_{null} as it does not apply any shrinkage on the effects. Overall, by jointly analyzing multiple mediators in a coherent statistical framework, both PTG and GMM outperform the other methods in simulations.

3.3.4 Simulation Results: Setting (A)-(C)

Table 3.3 shows the results in the small sample scenario and Table 3.4 shows the results in the large sample scenario. In all the settings, our proposed methods, PTG and GMM, outperform the other methods with an approximately 10% power gain in identifying active mediators. Between PTG and GMM, we find that both methods work preferably well in the setting where their corresponding effect size distribution is used. Specifically, in Setting (A) with $p = 2000$, the PTG method has the highest AUC (0.98) and TPR (0.40) at FDR = 10%. The performance of PTG is followed by GMM (AUC = 0.98, TPR = 0.37). In Setting (B) with $p = 2000$, the GMM method has the highest AUC (0.95) and TPR (0.51). The performance of GMM is followed by PTG (AUC = 0.92, TPR = 0.42). In Setting (C) where the effects are simulated with a horseshoe distribution, we find that GMM performs the best and its performance is followed by PTG and BAMA. The horseshoe distribution has a tall spike near zero and heavy tails, and therefore leads to a particularly challenging setting for most methods. The good performance of GMM in Setting (C) thus supports the robustness of the method. In addition, as before, both PTG and GMM provide reasonable empirical estimates of FDR and TPR (Table B.1, B.2 in the Appendix) based on a PIP = 0.5 cutoff. The accuracy gain in indirect effects estimation basically follows the same pattern as the power gain in mediator selection. The computing time of the proposed methods is reported

$n = 100, p = 200, p_{11} = 10$, *fixed effects (I)*				
Method	AUC	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
PTG (0.15,0.4,0.4)	0.99(0.001)	**0.54**(0.025)	0.040	0.486
GMM	0.98(0.002)	**0.42**(0.023)	0.054	1.336
BAMA	0.98(0.001)	0.40(0.022)	0.054	1.870
Bi-BLasso	0.90(0.005)	0.27(0.015)	0.092	21.879
PathLasso	0.75(0.004)	0.35(0.023)	0.098	17.220
Bi-Lasso	0.81(0.008)	0.38(0.019)	0.079	12.436
HIMA	0.61(0.005)	0.23(0.010)	0.081	1.832
Univariate	0.83(0.007)	0.25(0.014)	0.088	26.220
$n = 100, p = 200, p_{11} = 10$, *fixed effects (II)*				
Method	AUC	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
PTG (0.15,0.4,0.4)	0.96(0.003)	**0.34**(0.017)	0.074	0.549
GMM	0.96(0.003)	**0.39**(0.020)	0.029	0.791
BAMA	0.96(0.003)	0.31(0.015)	0.038	1.502
Bi-BLasso	0.90(0.005)	0.25(0.013)	0.044	11.040
PathLasso	0.72(0.007)	0.23(0.011)	0.072	3.225
Bi-Lasso	0.72(0.006)	0.30(0.014)	0.041	0.445
HIMA	0.56(0.005)	0.17(0.009)	0.056	2.526
Univariate	0.81(0.006)	0.19(0.013)	0.030	46.764

Table 3.1: Simulation results for fixed effects under $n = 100, p = 200$, p_{11} is the number of true active mediators. TPR: true positive rate at false discovery rate (FDR) = 0.10. $\text{MSE}_{\text{non-null}}$: mean squared error for the indirect effects of active mediators. MSE_{null}: mean squared error for the indirect effects of inactive mediators. The results are based on 200 replicates for each setting, and the standard errors are shown within parentheses. For PTG, we include the pre-defined thresholds $(\lambda_0, \lambda_1, \lambda_2)$ under each setting. Bolded TPRs indicate the top two performers.

$n = 1000, p = 2000, p_{11} = 100$, *fixed effects (I)*				
Method	AUC	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
PTG (0.15,0.4,0.4)	0.98(0.001)	**0.64**(0.008)	0.028	0.070
GMM	0.99(0.001)	**0.61**(0.009)	0.023	0.170
BAMA	0.98(0.001)	0.57(0.010)	0.038	0.141
Bi-BLasso	0.90(0.002)	0.23(0.004)	0.063	5.711
PathLasso	0.70(0.002)	0.20(0.005)	0.057	3.982
Bi-Lasso	0.76(0.001)	0.25(0.003)	0.051	0.290
HIMA	0.57(0.001)	0.16(0.003)	0.077	1.891
Univariate	0.93(0.001)	0.10(0.005)	0.092	225.056
$n = 1000, p = 2000, p_{11} = 100$, *fixed effects (II)*				
Method	AUC	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-6}$
PTG (0.15,0.4,0.4)	0.90(0.002)	**0.40**(0.008)	0.008	0.164
GMM	0.97(0.001)	**0.48**(0.006)	0.003	3.257
BAMA	0.96(0.001)	0.36(0.007)	0.005	7.346
Bi-BLasso	0.85(0.001)	0.18(0.004)	0.011	184.761
PathLasso	0.67(0.002)	0.19(0.003)	0.017	19.540
Bi-Lasso	0.70(0.001)	0.23(0.005)	0.007	4.925
HIMA	0.56(0.002)	0.09(0.004)	0.013	23.048
Univariate	0.90(0.002)	0.12(0.003)	0.075	208.660

Table 3.2: Simulation results for fixed effects under $n = 1000, p = 2000$, p_{11} is the number of true active mediators. TPR: true positive rate at false discovery rate (FDR) = 0.10. $\text{MSE}_{\text{non-null}}$: mean squared error for the indirect effects of active mediators. MSE_{null}: mean squared error for the indirect effects of inactive mediators. The results are based on 200 replicates for each setting, and the standard errors are shown within parentheses. For PTG, we include the pre-defined thresholds $(\lambda_0, \lambda_1, \lambda_2)$ under each setting. Bolded TPRs indicate the top two performers.

in Table C.1 of the Appendix, and the algorithms are relatively efficient for both $p = 200$ and $p = 2000$.

Finally, among the three frequentist methods, the bi-Lasso yields higher power as compared to the other two in all the scenarios and has smaller MSE in almost all the settings except for the horseshoe setting. Between bi-Lasso and bi-Bayesian Lasso, we find that the former outperforms the latter by higher TPR and smaller MSE_{null}. This comparison between bi-Lasso and bi-Bayesian Lasso suggests that the estimated indirect effects in bi-Bayesian Lasso may not be ideal for classifying mediators as compared to that used in bi-Lasso.

In summary, the simulations demonstrate the great advantage and robustness of GMM regarding selection and estimation accuracy, while PTG is more powerful under potentially large non-zero effects in mediator selection.

3.4 Data Application

3.4.1 Analysis of DNA Methylation in the MESA Cohort

We applied the proposed GMM and PTG to investigate the mediation mechanism of DNAm in the pathway from neighborhood socioeconomic disadvantage to BMI in the MESA data. Neighborhood SES is the exposure variable and is created based on a principal components analysis of 16 census-tract level variables reflecting dimensions of education, occupation, income, poverty, housing, etc. Higher values on the scale indicate greater neighborhood socioeconomic disadvantage. BMI is the outcome variable and also a critical risk factor for various diseases including T2D and CVD (*Hjellvik et al.*, 2012). Understanding how methylation at different CpG sites mediates the effects of neighborhood SES on BMI can shed light on the molecular mechanisms of complex diseases, thus leading to potential therapeutic strategies. The detailed processing steps for MESA data are provided in the Appendix. Briefly, we selected 1,225 individuals with non-missing data. Due to computational reasons,

$n = 100, p = 200, p_{11} = 10, PTG, \sigma_u^2 = 0.3$				
Method	AUC	TPR	$\text{MSE}_{\text{non-null}}$	MSE_{null} $\times 10^{-4}$
PTG (0.15, 0.4, 0.4)	0.98(0.002)	**0.45**(0.020)	0.05	1.59
GMM	0.98(0.001)	**0.43**(0.015)	0.03	4.25
BAMA	0.98(0.001)	0.41(0.019)	0.04	2.64
Bi-BLasso	0.89(0.006)	0.35(0.017)	0.05	6.83
PathLasso	0.65(0.013)	0.31(0.015)	0.06	2.43
Bi-Lasso	0.78(0.009)	0.40(0.020)	0.05	1.12
HIMA	0.60(0.007)	0.29(0.012)	0.07	5.46
Univariate	0.85(0.008)	0.29(0.023)	0.15	76.25
$n = 100, p = 200, p_{11} = 10, Gaussian, \sigma^2 = 0.3$				
Method	AUC	TPR	$\text{MSE}_{\text{non-null}}$ $\times 10^{-3}$	MSE_{null} $\times 10^{-5}$
PTG (0.04, 0.2, 0.2)	0.92(0.002)	**0.38**(0.008)	6.24	4.05
GMM	0.94(0.003)	**0.41**(0.006)	3.92	3.56
BAMA	0.95(0.003)	**0.38**(0.011)	5.06	3.39
Bi-BLasso	0.83(0.006)	0.28(0.014)	23.31	14.38
PathLasso	0.75(0.008)	0.30(0.011)	11.57	3.09
Bi-Lasso	0.75(0.003)	0.36(0.011)	7.50	1.52
HIMA	0.65(0.005)	0.21(0.009)	14.98	7.93
Univariate	0.75(0.006)	0.26(0.025)	62.46	234.30
$n = 100, p = 200, p_{11} = 10, Horseshoe, \sigma^2 = 0.5, b = 3$				
Method	AUC	TPR	$\text{MSE}_{\text{non-null}}$	MSE_{null} $\times 10^{-4}$
PTG (0.15, 0.5, 0.3)	0.80(0.009)	**0.30**(0.015)	0.42	7.16
GMM	0.83(0.006)	**0.33**(0.011)	0.03	5.21
BAMA	0.80(0.008)	0.28(0.017)	0.11	6.28
Bi-BLasso	0.76(0.011)	0.23(0.010)	0.45	42.36
PathLasso	0.65(0.019)	0.25(0.026)	0.51	6.04
Bi-Lasso	0.68(0.009)	0.27(0.017)	0.46	5.41
HIMA	0.60(0.006)	0.20(0.010)	0.41	26.51
Univariate	0.72(0.009)	0.20(0.020)	0.44	512.33

Table 3.3: Simulation results for $n = 100, p = 200$, p_{11} is the number of true active mediators. TPR: true positive rate at false discovery rate (FDR) = 0.10. $\text{MSE}_{\text{non-null}}$: mean squared error for the indirect effects of active mediators. MSE_{null}: mean squared error for the indirect effects of inactive mediators. The results are based on 200 replicates for each setting, and the standard errors are shown within parentheses. For PTG, we include the pre-defined thresholds $(\lambda_0, \lambda_1, \lambda_2)$ under each setting. Bolded TPRs indicate the top two performers.

$n = 1000, p = 2000, p_{11} = 100, PTG, \sigma_u^2 = 0.1$				
Method	AUC	TPR	$\text{MSE}_{\text{non-null}} \times 10^{-4}$	$\text{MSE}_{\text{null}} \times 10^{-6}$
PTG (0.05,0.15,0.15)	0.98(0.001)	**0.40**(0.008)	5.28	2.46
GMM	0.98(0.001)	**0.37**(0.010)	3.86	4.26
BAMA	0.98(0.001)	0.30(0.012)	4.84	3.62
Bi-BLasso	0.92(0.003)	0.29(0.018)	7.92	11.38
PathLasso	0.77(0.009)	0.22(0.007)	7.02	1.74
Bi-Lasso	0.83(0.003)	0.28(0.014)	5.60	1.81
HIMA	0.53(0.002)	0.14(0.004)	9.96	4.96
Univariate	0.85(0.003)	0.11(0.023)	60.24	214.57
$n = 1000, p = 2000, p_{11} = 100, Gaussian, \sigma^2 = 0.1$				
Method	AUC	TPR	$\text{MSE}_{\text{non-null}} \times 10^{-3}$	$\text{MSE}_{\text{null}} \times 10^{-5}$
PTG (0.02,0.2,0.1)	0.92(0.002)	**0.42**(0.006)	4.76	0.874
GMM	0.95(0.001)	**0.51**(0.007)	2.09	0.712
BAMA	0.90(0.003)	0.41(0.018)	2.85	0.722
Bi-BLasso	0.88(0.002)	0.32(0.007)	4.85	1.632
PathLasso	0.78(0.011)	0.25(0.003)	4.88	1.256
Bi-Lasso	0.81(0.002)	0.38(0.010)	2.53	0.368
HIMA	0.55(0.002)	0.19(0.004)	8.41	1.544
Univariate	0.82(0.003)	0.19(0.017)	34.08	20.05
$n = 1000, p = 2000, p_{11} = 100, Horseshoe, \sigma^2 = 0.3, b = 3$				
Method	AUC	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
PTG (0.03,0.3,0.1)	0.74(0.002)	**0.29**(0.008)	0.18	10.04
GMM	0.80(0.001)	**0.38**(0.007)	0.14	2.94
BAMA	0.75(0.002)	0.27(0.006)	0.25	3.88
Bi-BLasso	0.71(0.002)	0.09(0.003)	0.26	127.55
PathLasso	0.66(0.008)	0.05(0.002)	0.41	2.03
Bi-Lasso	0.72(0.003)	0.24(0.007)	0.24	1.57
HIMA	0.55(0.002)	0.09(0.004)	0.39	1.56
Univariate	0.77(0.003)	0.09(0.015)	0.59	644.07

Table 3.4: Simulation results for $n = 1000, p = 2000$, p_{11} is the number of true active mediators. TPR: true positive rate at false discovery rate (FDR) = 0.10, $\text{MSE}_{\text{non-null}}$: mean squared error for the indirect effects of active mediators. MSE_{null}: mean squared error for the indirect effects of inactive mediators. The results are based on 200 replicates for each setting, and the standard errors are shown within parentheses. For PTG, we include the pre-defined thresholds $(\lambda_0, \lambda_1, \lambda_2)$ under each setting. Bolded TPRs indicate the top two performers.

we focused on a final set of 2,000 CpG sites that have the strongest marginal associations with neighborhood SES. We applied various methods for the mediation analysis. In the outcome model, we adjust for age, gender, race/ethnicity, childhood socioeconomic status (SES) and adult SES. In the mediator model, we control for age, gender, race/ethnicity, childhood SES, adult SES, and enrichment scores for 4 major blood cell types (neutrophils, B cells, T cells and natural killer cells). All the continuous variables are standardized to have zero mean and unit variance.

We display the PIP values for each of the 2,000 CpG sites from PTG and GMM in Figure 3.1. GMM identified nine CpG sites with significance evidence for mediating the neighborhood SES effects on BMI based on 0.5 cutoff of PIPs. In contrast, PTG identified twelve significant CpG sites at the same threshold, which include all the nine sites selected by GMM method. The top five CpG sites identified by the two methods are identical. The rank correlation for the mediator rank lists obtained from both methods is 0.87, supporting the high consistency between the two methods. We carefully examine the nearby genes of the detected methylation sites by GMM and PTG. Among them, the protein-coding gene *PTK2* has been previously discovered as BMI risk loci (*Zeller et al.*, 2018); *PCID2* and *NFE2L1* have been shown to be associated with obesity, glucose, diabetes and related metabolic diseases (*Zheng et al.*, 2015; *Erdmann et al.*, 2018); *COX6A1P2* was robustly recognized to link with obesity development in multiple epigenome-wide studies (*Kvaløy et al.*, 2018) and *EVI2B* was reported as one of the regulatory genes related to obesity (*Kogelman et al.*, 2014). Therefore, the genes nearby the detected CpG sites may play an important role in transmitting the effects of neighborhood SES to BMI. For the other competing methods, BAMA, HIMA and the univariate methods do not have sufficient power to identify any significant CpG sites at 0.10 FDR. Bi-Lasso and Pathway Lasso tend to produce a large number of false positives in simulations, and thus it is hard to verify their findings in the real data application.

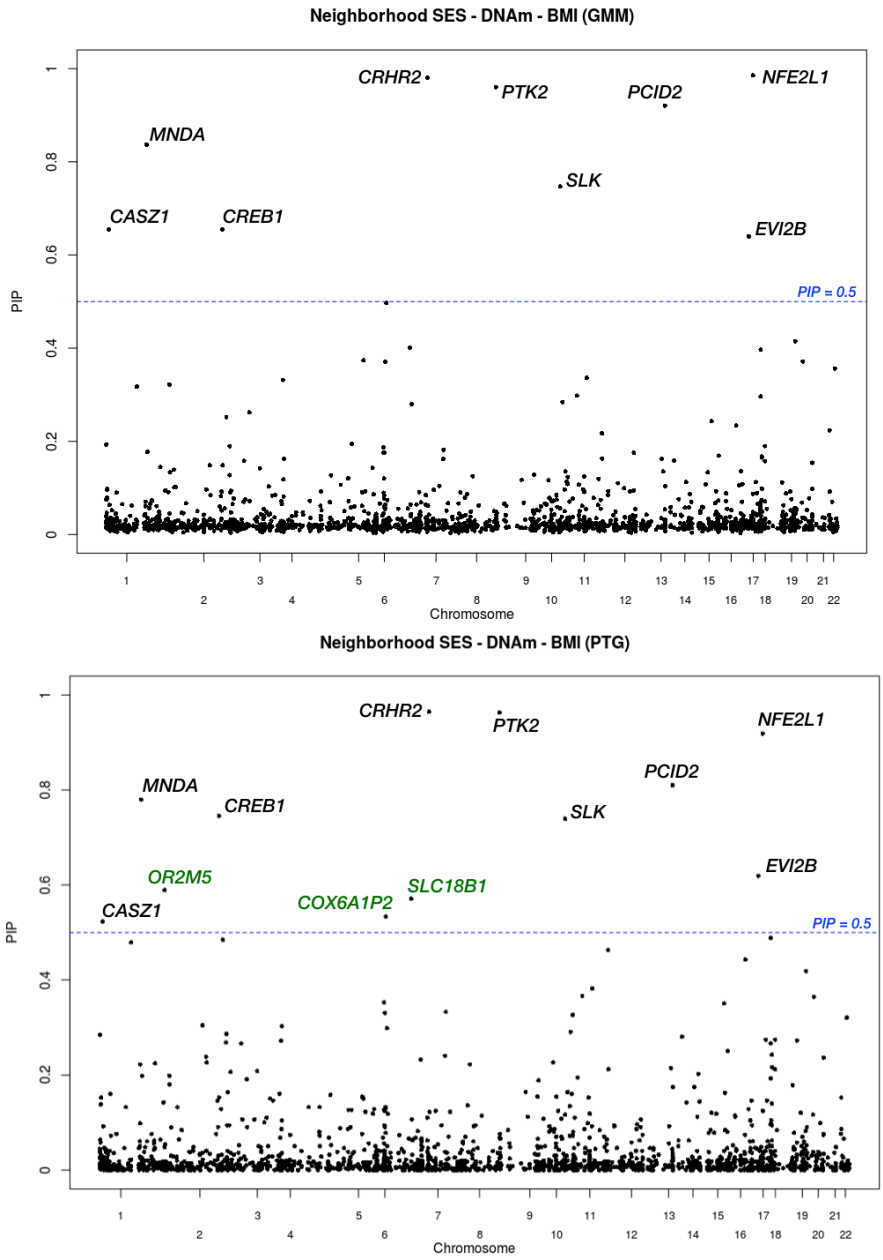

Figure 3.1: Data analysis results for the trio Neighborhood SES → DNAm → BMI in MESA data. The upper panel shows the PIPs obtained from the GMM method, and the lower panels shows the PIPs obtained from the PTG method. The blue lines mark the PIP = 0.5 threshold, and we include the nearby genes of the selected CpG sites. Most of the sites are identified by both methods, and the three genes in green are additional findings from PTG.

3.4.2 Analysis of Endogenous Biomarkers and Environmental Data in the LIFECODES Birth Cohort

As another data example, we study the collective impact of endogenous signaling molecules derived from lipids, peptides, and DNA in mediating prenatal exposure to environmental contaminants on the risk of preterm birth in the LIFECODES birth cohort. Detailed description of the study is provided in the Appendix. Briefly, we consider $n = 161$ pregnant women registered at the Brigham and Women's Hospital in Boston, MA between 2006 and 2008. Subjects' urine and plasma specimens were collected at one study visit occurring between 23.1 and 28.9 weeks gestation. Four classes of environmental contaminants, including phthalates, phenols, polycyclic aromatic hydrocarbons, and trace metals, were measured in each urine sample. Among them, phthalates are the high-production volume chemicals commonly used as plasticizers in numerous consumer products. Previous studies have shown that everyday exposure to phthalates during pregnancy would increase risk of delivering preterm (*Ferguson et al.*, 2014b). Recent studies have also uncovered associations between multiple lipid biomarkers and preterm birth (*Aung et al.*, 2019). Based on these previous literature, we aim to understand the molecular mechanism underlying the effects of phthalates on preterm. To do so, we follow *Aung et al.* (2020) to create an environmental risk score for the phthalate class and treat such risk score as the exposure variable. We recorded the gestational age at delivery as the continuous birth outcome. In terms of mediators, we obtained 61 endogenous biomarkers from urine and plasma that included 51 eicosanoids, five oxidative stress biomarkers and five immunological biomarkers. With these variables, we examine if any of these 61 available biomarkers mediates the effects of phthalate exposure on gestational age at delivery. In the analysis, we perform log-transformation on all measurements of the exposure metabolites and endogenous biomarkers. We adjust for age and maternal BMI from the initial visit, race, and urinary specific gravity levels inside both models of the mediation analysis. Since the cohort is oversampled for preterm cases (< 37 weeks gestation), we multiply the data by the case-control sampling weights to adjust for

that.

We applied the proposed methods to the data and summarize the results in Table 3.5. Both PTG and GMM identified significant mediators that mediate the effects of the phthalate exposure on gestational age at delivery based on PIP = 0.5 cutoff (Figure 3.2), with rank lists of mediators positively correlated with each other (rank correlation = 0.48). Specifically, GMM identified two significant biomarkers (9-oxooctadeca-dienoic acid [9-oxoODE], 12,13-epoxy-octadecenoic acid [12(13)-EpoME]). PTG identified three significant biomarkers (8-hydroxydeoxyguanosine [8-OHdG], 12(13)-EpoME, leukotriene D4 [LTD4]), one of which (12(13)-EpoME) overlaps with those identified by GMM. Among the identified biomarkers, 8-OHdG is commonly utilized as a marker of oxidative stress generated upon repair of oxidative DNA damage and has been found strongly associated with decreased gestational length and increased risk of preterm (*Ferguson et al.*, 2015; *Hsieh et al.*, 2012); while LTD4 has been shown to exhibit significant associations with preterm birth, and 9-oxoODE and 12(13)-EpoME had an important protective effect on preterm birth (*Aung et al.*, 2019). Therefore, our results help improve the understanding of the molecular mechanisms underlying the effects of environmental exposure on preterm, and could further lead to improvement of treatment and prevention strategies.

3.5 Discussion

In this chapter, we present two novel joint modeling methods, PTG and GMM, for high-dimensional mediation analysis. Our methods can jointly model a large number of mediators and enable penalization on the indirect effects in a targeted way. Our methods effectively characterize the high-dimensional set of potential mediators into four groups based on the exposure-mediator and mediator-outcome effects: the active mediating group and three non-mediating groups. These group categorizations are in consonance with the composite structure for testing the indirect effect recently proposed in genome-wide mediation analyses

Method	Selected Mediators
MESA: Neighborhood SES → DNAm → BMI	
GMM	CRHR2, NFE2L1, PTK2, PCID2, MNDA, SLK, CREB1, CASZ1, EVI2B
PTG (0.01,0.05,0.1)	CRHR2, NFE2L1, PTK2, PCID2, MNDA, CREB1, SLK, EVI2B, OR2M5, SLC18B1, COX6A1P2, CASZ1
LIFECODES: Phthalates → Biomarkers → Gestational age	
GMM	12(13)-EpoME, 9-oxoODE
PTG (3.0,2.0,1.5)	12(13)-EpoME, 8-OHdG, LTD4

Table 3.5: Summary of the identified active mediators from the data application on MESA and LIFECODES study. For PTG, we include the pre-defined thresholds $(\lambda_0, \lambda_1, \lambda_2)$ for the two real datasets.

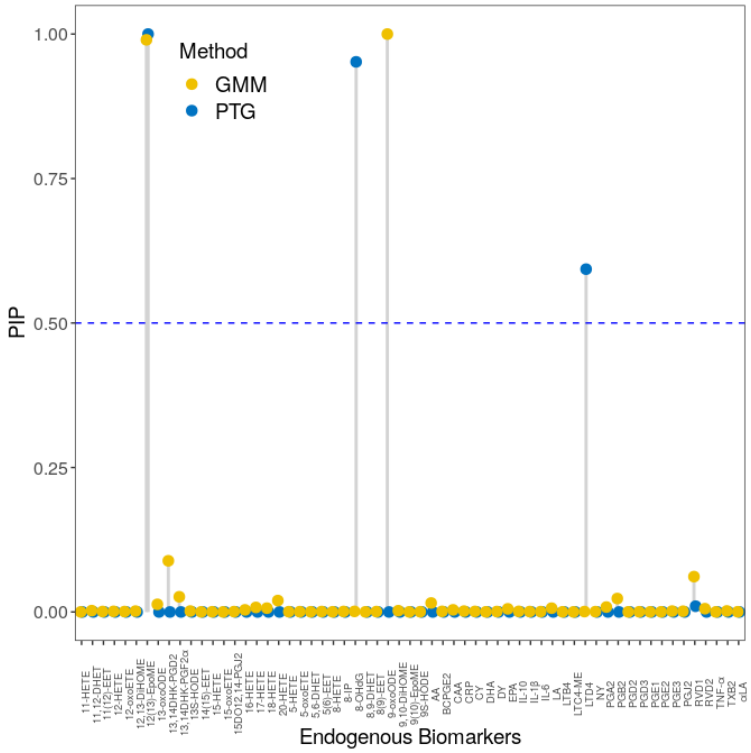

Figure 3.2: Data analysis results for the LIFECODES cohort. The panel shows the PIPs obtained from GMM and PTG methods for the trio Phthalates → Biomarkers → Gestational Age. The blue line marks the PIP = 0.5 threshold.

(*Huang et al.*, 2019). With extensive simulations, we show that our methods achieve up to 30% power gain in identifying true non-null mediators compared with other alternatives, including several recently developed penalized and Bayesian methods for mediation analysis. We have demonstrated the benefits of our methods in the MESA and LIFECODES cohorts. For example, in the MESA cohort, we identify several methylation sites and their nearby genes, such as *NFE2L1* and *PTK2*, with strong evidence for mediating the neighborhood SES effects on BMI.

Bayesian FDR control is of great importance to safeguard false positives in the scientific discovery. For PTG and GMM, we rely on median inclusion probabilities (PIP = 0.5) to identify active mediators while maintaining a reasonable FDR. For bi-Bayesian Lasso and other continuous shrinkage methods, such as the scale mixture of normals prior (*Carvalho et al.*, 2009), we have also attempted to estimate the Bayesian FDR using shrinkage factors following *Carvalho et al.* (2010). However, we find it challenging to adapt the shrinkage factors in these methods as an optimal strategy for ranking correlated mediators and to identify a significance threshold for declaring signals while bounding the false discoveries. Therefore, coming up with an analog of PIP in the global-local shrinkage framework as the selection criterion and developing a thresholding rule for determining significance levels in mediation analysis remains a topic of future investigation.

One limitation of this current work is that the proposed methods do not explicitly incorporate the correlation structure among mediators in the model building process. Treating mediators independent *a priori*, the models may fail to distinguish among highly correlated mediators and lose power in mediator selection when active mediators tend to correlate with each other. Correlations among mediators are commonly seen in modern data analysis; such examples include genomic data that measure hundreds of thousands of gene expressions/single nucleotide polymorphisms (SNPs), and brain image data that contain a large number of voxels/regions. Incorporating mediator correlation information into our Bayesian paradigm could be a promising direction to pursue, and we explore more on that in the next chapter.

CHAPTER IV

Bayesian Hierarchical Modeling of High-Dimensional Mediation Analysis for Coordinated Selection of Correlated Mediators

4.1 Introduction

In the last two chapters, we have introduced multiple Bayesian methods developed in high-dimensional sparse mediation analysis. They enable a joint analysis of high-dimensional correlated mediators and exhibit great advantages and flexibility in a wide range of scenarios. However, one common issue with the previous described methods is that mediators are assumed to be independent *a priori* and the specified priors ignore the possible correlation structure among the mediators. In modern data applications, thanks to the technology advances, the substantial correlation has been frequently seen in the increasingly high-dimensional mediators. For example, in large-scale genomic studies, hundreds of thousands of single nucleotide polymorphisms (SNPs) were measured and shown highly correlated within linkage disequilibrium (LD)-based blocks; and in functional MRI (fMRI) studies, the brain images are composed of a large number of voxels/regions and true signals usually represent connected regions. As far as we know, none of the existing methods for mediation analysis has incorporated the useful correlation structural information in the model build-

ing process. Intuitively, positively correlated mediators tend to transmit similar mediation effects and should be grouped together when modeled simultaneously. Incorporating prior knowledge on the structure of the mediators is expected to improve active mediator detection and strengthen Bayesian learning.

Bayesian variable selection with covariate structural information has received much attention over the years. Bayesian group Lasso (*Raman et al.*, 2009) and Bayesian sparse group selection method (*Chen et al.*, 2016) allow for the inclusion of grouping effects and lead to more parsimonious models with reduced estimation error compared with standard Lasso. *Yuan and Lin* (2005) also develop a correlation prior on the binary selection indicators to distinguish models with the same size. Bayesian graphical models represent another stream of work on structural variable selection. Both *Liu et al.* (2014) and *Cai et al.* (2018) utilize the graph Laplacian matrix to encode the network information into the regression coefficients. *Stingo et al.* (2011) and *Peng et al.* (2013) propose the simultaneous selection of pathways and genes, using the pathway summaries of the group behavior and structure dependency within pathways to inform the selection. *Chang et al.* (2018) assume a multivariate normal prior on the shrinkage parameters in the Laplace priors for a informative network based prior. Along with the above methods, emerging literature considers the extension of the "spike-and-slab" type of mixture prior (*Mitchell and Beauchamp*, 1988) in combination with Markov random field (MRF) prior to incorporate graph information. Ising prior, a binary spatial MRF, and its variations have been effectively applied to induce sparsity and accommodate selection dependency. *Li and Zhang* (2010) and *Chekouo et al.* (2016) show that the structural information through Ising priors can greatly improve selection and prediction accuracy over the independent priors. In addition to smoothing over the latent selection indicators, recent studies deploy different types of "slab distribution", such as the Dirichlet Process (*Li et al.*, 2015), the Gaussian MRF (*Goldsmith et al.*, 2014), the group fused Lasso prior (*Zhang et al.*, 2014), etc., to include the grouping and smoothing effect in the non-zero regression coefficients due to local dependence or high correlation.

In the last chapter, we have introduced the Gaussian mixture model (GMM) for a joint prior on the exposure-mediator and mediator-outcome effects and consequently a targeted shrinkage on their products, i.e. the indirect effect. GMM assumes that each mediator can be independently categorized into one of the four components based on association pattern, and its group indicator follows the same multinomial distribution as the other mediators. In this chapter, with the goal of utilizing the correlation structure among mediators in the modeling process, we may replace the independent priors on the mediators' group indicators with a joint Potts prior. The Potts distribution (*Potts*, 1952), a generalization from the Ising distribution, allows for more than two groups and complex dependency between neighbors, which can be spatially or statistically correlated. On the other hand, we can also jointly model the mediator-specific mixing probabilities via a logistic normal distribution (*Atchison and Shen*, 1980), with the group probabilities reflecting the underlying correlation structure. To sum up, we develop two methods: Potts mixture model and correlated multinomial model for high-dimensional mediation analysis, adding the possible grouping effect across mediators through another layer in the Bayesian heirachy. Both methods are built off the GMM introduced in Chapter 3, and thus inherit the merits of the GMM method. Furthermore, the proposed methods incorporate the structural information into a prior favoring selection of correlated mediators, and are expected to allow the identification of mediators that could be missed otherwise. The proposed method will also facilitate the interpretation of the results, particularly for the selected mediators with high correlations.

We illustrate the advantage of proposed method on the LIFECODES prospective birth cohort. Preterm birth remains the leading cause of infant mortality in the U.S. (*Callaghan et al.*, 2006), and one suspected risk factor for preterm delivery is exposure to environmental contaminants (*Vrijheid et al.*, 2016). Those toxicants (e.g. phthalates, toxic heavy metals) were also shown to disturb receptor activity and induce their responses, which could affect the signaling molecules related to inflammation and metabolism (*Kiyama and Wada-Kiyama*, 2015; *Milnerowicz et al.*, 2015). Meanwhile, recent studies revealed associations

between preterm birth and certain biomarkers of inflammation and oxidative stress (*Ferguson et al.*, 2014a, 2015). In light of those multiple lines of evidence, we hypothesize that the relationship between toxicant exposure during pregnancy and preterm may be mediated by endogenous biomarkers of lipid metabolism, inflammation, and oxidative stress. In the present study, moderate to strong correlations across biomarkers are observed and not only occur within the same biological pathways. The whole correlation structure will be utilized to inform biomarker selection and interpretation. Another application on the MESA data will demonstrate our methods under a larger number of mediators and relatively weak correlation structure. In this application, DNA methylations are hypothesized to mediate the effect of neighborhood factors on clinical health outcomes. Previous studies have provided evidence of the associations between neighborhood conditions and changes in DNA methylation (*Smith et al.*, 2017), as well as associations between changes in DNA methylation and diabetes/CVD risks (*Rakyan et al.*, 2011; *Dayeh et al.*, 2014; *Abi Khalil*, 2014; *Zhong et al.*, 2016). Understanding the molecular basis of those complex diseases through the integrative mediation analysis will facilitate the development of prevention and treatment strategies, where the interventions can be designed to target the mediating DNAm or nearby genes.

4.2 Method

We start with a brief review of our high-dimensional mediation models. For the outcome model, $i = 1, 2, ..., n$,

$$Y_i = \boldsymbol{M}_i^\top \boldsymbol{\beta_m} + A_i \beta_a + \boldsymbol{C_i}^\top \boldsymbol{\beta_c} + \epsilon_{Yi}, \tag{4.1}$$

where $\boldsymbol{\beta_m} = (\beta_{m1}, \ldots, \beta_{mp})^\top$ is the mediator-outcome effect for the p mediators; $\boldsymbol{\beta_c} = (\beta_{c1}, \ldots, \beta_{cq})^\top$ is the coefficient for the q covariates; and $\epsilon_{Yi} \sim \mathrm{N}(0, \sigma_e^2)$.

For the mediator model,

$$\boldsymbol{M}_i = A_i \boldsymbol{\alpha_a} + \boldsymbol{\alpha_c} \boldsymbol{C_i} + \boldsymbol{\epsilon}_{Mi}, \tag{4.2}$$

where $\boldsymbol{\alpha_a} = (\alpha_{a1}, \ldots, \alpha_{ap})^\top$ is the exposure-mediator effect for the p mediators; $\boldsymbol{\alpha_c} = (\boldsymbol{\alpha_{c1}^\top}, \ldots, \boldsymbol{\alpha_{cp}^\top})^\top$; $\boldsymbol{\alpha_{c1}}, \ldots, \boldsymbol{\alpha_{cp}}$ are q-by-1 vectors of coefficients for the q covariates; $\boldsymbol{\epsilon}_{Mi} \sim \text{MVN}(\mathbf{0}, \boldsymbol{\Sigma})$, with $\boldsymbol{\Sigma}$ capturing potential residual error covariance. ϵ_{Yi} and $\boldsymbol{\epsilon}_{Mi}$ are assumed to be independent of each other and independent of A_i and $\boldsymbol{C_i}$.

With the same rationale as discussed in Chapter III, we consider a four-component Gaussian mixture model for the effects of the j-th mediator,

$$[\beta_{mj}, \alpha_{aj}]^\top \sim \pi_{1j}\text{MVN}_2(\mathbf{0}, \boldsymbol{V}_1) + \pi_{2j}\text{MVN}_2(\mathbf{0}, \boldsymbol{V}_2) + \pi_{3j}\text{MVN}_2(\mathbf{0}, \boldsymbol{V}_3) + \pi_{4j}\boldsymbol{\delta}_0 \tag{4.3}$$

with a prior probabilities π_{kj} ($k \in \Omega, \Omega = \{1, 2, 3, 4\}$) summing to one for the jth mediator. We also introduce a membership indicator variable γ_j for the j-th mediator, where $\gamma_j = k$ if $[\beta_{mj}, \alpha_{aj}]^\top$ is from normal component k, $k \in \{1, 2, 3, 4\}$.The other parameters are defined similarly as in Chapter III. If we assume independent prior distributions across $\pi_{k1}, \pi_{k2}, ..., \pi_{kp}$ (or $\gamma_1, \gamma_2, ..., \gamma_p$), then each mediator is independent *a priori* and the prior distribution on $\left[\boldsymbol{\beta_m}, \boldsymbol{\alpha_a}\right]^\top$ after integrating out π_{kj} (or γ_j) is essentially a separable product of distributions of $[\beta_{mj}, \alpha_{aj}]^\top$. This is also akin to the concept of "separable prior" in *Ročková and George* (2018). Recall that in the previous developed GMM method, we have a common set of $\pi_1, \pi_2, \pi_3, \pi_4$ for all the mediators. This specification enables the information borrowing across mediators and tie mediators together through the mixing probabilities, making the prior "non-separable". However, since this approach assumes the same mixing probabilities for all the mediators *a priori*, it does not differentiate highly correlated mediators from uncorrelated ones to inform mediator categorization. For example, if the j-th mediator and the $(j+1)$-th mediator are highly correlated, then presumably γ_j and γ_{j+1} are more likely to be the same. To improve the flexibility of the previous GMM, we turn to mediator-specific $\{\pi_{kj}, k = 1, 2, 3, 4\}$ and γ_j, and instead of independent, separable prior, we consider embedding the correlated information to $\{\pi_{kj}, k = 1, 2, 3, 4\}$'s or γ_j's for mediator selection. In the following sections, we describe the proposed methods with more details.

4.2.1 Hierarchical Potts Mixture Model: GMM-Potts

The Potts model (*Potts*, 1952) was initially developed as a generalization of the Ising model in statistical physics (*Ising*, 1925; *Winkler*, 2012). However, it has enjoyed great success as a prior model for the spatial modeling in image analysis (*Feng et al.*, 2012; *Li et al.*, 2019), disease mapping (*Best et al.*, 2005), genetics studies (*Thomas et al.*, 2003; *Yu et al.*, 2012), etc. In those applications, Potts models incorporate spatial Markovian dependency by assigning homogeneous relationships for the "neighboring" regions. In the context of mediation analysis, we can allocate the high-dimensional mediators into four clusters based on their exposure-mediator and mediator-outcome effects. To further draw the connection, we may think of highly correlated mediators as neighbors, and they tend to be assigned to the same cluster (Gaussian component) through Potts model.

To specifically formulate our Potts mixture model, we assume $\boldsymbol{\gamma} = (\gamma_1, \gamma_2, ..., \gamma_p)$ follows a Potts distribution,

$$p(\boldsymbol{\gamma}|\boldsymbol{\theta}_0, \boldsymbol{\theta}_1) = c(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1)^{-1} \exp\Big\{ \sum_{i=1}^{p} \theta_{0k} I[\gamma_i = k] \Big\} \times \exp\Big\{ \sum_{i=1}^{p} \sum_{i \sim j} \sum_{k=1}^{4} \theta_{1k} I[\gamma_i = \gamma_j = k] \Big\} \quad (4.4)$$

where $i \sim j$ indicates neighboring pairs and $I(\cdot)$ is the indicator function. The neighboring relationship can be defined in terms of domain knowledge or in our case, the correlated information. $\boldsymbol{\theta}_0 = (\theta_{01}, \theta_{02}, \theta_{03}, \theta_{04})$ represents the relative group distributions *a priori* without any neighboring information and controls the overall sparsity. $\boldsymbol{\theta}_1 = (\theta_{11}, \theta_{12}, \theta_{13}, \theta_{14})$ represents the prior belief on the strength of association between neighboring pairs. For $\theta_{1k} > 0$, the Potts distribution encourages configurations where "neighboring mediators" belong to the same cluster, and the larger θ_{1k}, the tighter this coupling. When $\boldsymbol{\theta}_1 = \mathbf{0}$, cluster membership of one mediator is independent of its neighbors. The exact calculation of the normalizing constant $c(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1)$ in Potts distribution requires the summation over the entire space of $\boldsymbol{\gamma}$, which consists of 4^p states. Even for a moderate number of mediators, $c(\boldsymbol{\theta}_0, \boldsymbol{\theta}_1)$ is computationally intractable, and this complicates the Bayesian inference unless

$\boldsymbol{\theta}_0, \boldsymbol{\theta}_1$ are assumed as fixed hyperparameters. Based on the full probability distribution in Equation 4.4, the probability for the j-th mediator belonging to cluster k conditional on its neighbors is,

$$p(\gamma_j = k|\{\gamma_i\}_{i\neq j}, \boldsymbol{\theta}_0, \boldsymbol{\theta}_1) = \frac{\exp\{\theta_{0k}\} \times \exp\{\sum_{i\sim j} \theta_{1k} I[\gamma_i = \gamma_j = k]\}}{\sum_{k=1}^{4} \exp\{\theta_{0k}\} \times \exp\{\sum_{i\sim j} \theta_{1k} I[\gamma_i = \gamma_j = k]\}} \tag{4.5}$$

This conditional probability depends on the neighbors of the j-th mediator and demonstrates the Markov property of the Potts distribution.

We develop a Markov chain Monte Carlo (MCMC) sampling strategy for the proposed model. Due to the intractable normalizing constant in Potts distribution, the update of $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1$ cannot be handled by the standard Metropolis Hastings (MH) algorithm. To address this issue, we employ the double MH sampler (*Liang*, 2010) to generate auxiliary variables via the MH transition kernels and eliminate the normalizing constants. For $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1$, we consider normal priors, and the prior means of $\{\theta_{0k}\}$ are set to have the desired inclusion probability while the prior means of $\{\theta_{1k}\}$ are set to be the same positive number. This prior information favors the grouping of correlated mediators. According to Equation 4.5, the updating of $\boldsymbol{\gamma}$ can be realized through single site Gibbs sampling. Since the sampling space of $\boldsymbol{\gamma}$ is huge and discrete, the efficiency of the standard Gibbs updates can be improved by the Swendsen-Wang (SW) algorithm (*Higdon*, 1998). The SW algorithm partitions the whole set of mediators into blocks within which the mediators belong to the same normal component, and then updates each block independently. Following the strategy in *Higdon* (1998), we alternate between the single site Gibbs updates of $\boldsymbol{\gamma}$ and SW updates to ensure movement in large patches and fast mixing of the algorithm. The detailed algorithm is given in the Appendix.

In our Potts mixture model, the "neighboring" mediators are predefined to capture the correlation structure among mediators. Based on our experience, including too many neighbors into the model will cause irrelevant noises to the group probabilities and blur the cluster boundary; while including too few neighbors will certainly lose some of the important struc-

tural information. In this work, we apply the common clustering method on the $p(p-1)/2$ pairwise correlations among mediators to divide them into two groups: high correlation and background noise. This procedure essentially sets a correlation threshold for neighbors and non-neighbors in a data dependent way, and we define the i-th mediator and j-the mediator as neighbors if their pairwise correlation is above this threshold. The threshold should reflect the prior knowledge on the neighborhood structure and relationships across mediators.

The Potts mixture model translates the correlation structure and incorporates the local dependency through mediators' predefined neighbors. For each mediator, its four-component group probabilities will only smooth over the most correlated mediators with this one, and ignore the non-neighboring ones. This local property does not incur much additional computational burden compared to the previous GMM, and more importantly, filters out the uncorrelated information. However, if the neighbors and non-neighbors are hard to distinguish through a clustering method, e.g. within a weak correlation structure, the inaccurately specified neighbors may adversely affect the performance of the method. To avoid the need of neighborhood pre-specification and also for a more direct and general position of the overall correlation structure, we consider a different approach as described in the next section.

4.2.2 Hierarchical GMM with Correlated Selection: GMM-CorrS

In this section, we take another approach to accounting for the correlation among mediators in mediator selection. For each mediator, the selection/group indicator γ_j follows a multinomial distribution with parameters π_{1j}, π_{2j}, π_{3j}, π_{4j}, and $\sum_{k=1}^{4} \pi_{kj} = 1$. We propose to jointly model all the mediators' mixing probabilities and their continuous dependence structure via latent logistic normal distributions. The logistic normal (*Atchison and Shen*, 1980) has been studied in the context of analyzing compositional data, such as bacterial composition in human microbiome data (*Xia et al.*, 2013) and topics proportions associated with document collections in correlated topics model (*Chen et al.*, 2013a). In mediation analysis,

it will allow for a flexible covariance structure among the mediators and give a more realistic model where correlated mediators will have similar group probabilities *a priori*, including the prior inclusion probability. However, adding this Gaussian correlation structure among multinomial parameters breaks the Dirichlet-multinomial conjugacy as used in the previous chapters, and poses a challenge in Bayesian computation. Approximation techniques, such as variational inference are feasible, but they do not always come with the theoretical guarantees as MCMC and may require additional modeling assumptions (*Blei et al.*, 2007).

On the other hand, Bayesian inference with binomial likelihood or Bayesian logistic regression has long been explored given its inconvenient analytic form of the likelihood and the non-existence of a conjugate prior for parameters of interest. To solve the computational issue, *Holmes et al.* (2006) develops auxiliary variable approaches using normal-scale mixture for the noise process as an extension to the probit model; and *Frühwirth-Schnatter and Frühwirth* (2010) approximates the noise process with a discrete mixture of normals. More recently, *Polson et al.* (2013) constructs a new data-augmentation strategy based on the novel class of Pólya-Gamma (PG) distributions, and the method is notably simpler and more efficient than the previous schemes for Bayesian hierarchical models with binomial likelihoods. To achieve a multinomial generalization, we leverage a logistic stick-breaking representation in the Pólya-Gamma augmentation (*Linderman et al.*, 2015) to formulate the multinomial distribution in terms of latent variables with jointly Gaussian likelihoods.

We rewrite 4-dimensional multinomial in terms of 3 binomial densities $\widetilde{\pi}_{j1}$, $\widetilde{\pi}_{j2}$ and $\widetilde{\pi}_{j3}$,

$$\begin{aligned}
p(\gamma_j = 1) &= \widetilde{\pi}_{j1} = \pi_{j1} \\
p(\gamma_j = 2 | \gamma_j \neq 1) &= \widetilde{\pi}_{j2} = \pi_{j2}/(1-\pi_{j1}) \\
p(\gamma_j = 3 | \gamma_j \neq 1 \text{ or } 2) &= \widetilde{\pi}_{j3} = \pi_{j3}/(1-\pi_{j1}-\pi_{j2}) \\
p(\gamma_j = 4 | \gamma_j \neq 1 \text{ or } 2 \text{ or } 3) &= \widetilde{\pi}_{j4} = \pi_{j4}/(1-\pi_{j1}-\pi_{j2}-\pi_{j3}) = 1 \\
\text{Multinomial}(\gamma_j | 1, \{\pi_{j1}, \pi_{j2}, \pi_{j3}, \pi_{j4}\}) &= \prod_{k=1}^{3} \text{Binomial}(I(\gamma_j = k) | n_{jk}, \widetilde{\pi}_{jk})
\end{aligned}$$

where $n_{jk} = 1 - \sum_{k'<k} I(\gamma_j = k')$, $n_{j1} = 1$. The multinomial distribution is now expressed with three binomial distributions and each $\widetilde{\pi}_{jk}$ describes the faction of the remaining probability for the k-th group (details in the Appendix). To better aid the interpretation of the above stick-breaking representation, consider a testing strategy for the indirect effect $\beta_{mj}\alpha_{aj}$ implemented on each mediator, and we get the subset of active mediators with $\gamma_j = 1$. For the remaining mediators, we further look into their marginal effects: $p(\gamma_j = 2|\gamma_j \neq 1)$ is the conditional probability of having non-zero β_{mj} effect given it is not an active mediator; and $p(\gamma_j = 3|\gamma_j \neq 1 \text{ or } 2)$ is the conditional probability of having non-zero α_{aj} effect given that $\beta_{mj} = 0$. The rest of the mediators will surely have $\gamma_j = 4$. We note that under the sparsity assumption, for most of the mediators, $\widetilde{\pi}_{j2} \approx \pi_{j2}$, $\widetilde{\pi}_{j3} \approx \pi_{j3}$ due to the small values of π_{j1} and π_{j2}.

We define $b_{jk} = \text{logit}(\widetilde{\pi}_{jk})$ for $k = 1, 2, 3$ and $j = 1, 2, \ldots, p$. We then stack the $3 \times p$ b_{jk}'s as one random vector, and assume a multivariate normal prior on it, that is,

$$\boldsymbol{b} := \{b_{jk}\}_{j=1,\ldots,p;k=1,2,3}$$

$$\boldsymbol{b} \sim \text{MVN}(\boldsymbol{a}, \text{diag}\{\sigma^2_{d1}, \sigma^2_{d2}, \sigma^2_{d3}\} \otimes \boldsymbol{D}) \tag{4.6}$$

where $\otimes$ denotes the Kronecker product. The logistic transformation maps the transformed multinomial parameters to the $3p$-dimensional open real space. The prior mean $\boldsymbol{a} = \{a_{jk}\}_{j=1,\ldots,p;k=1,2,3}$, and it is chosen such that $a_{jk} = a_{j'k}$ for $k = 1, 2, 3$ and $1 \leq j < j' \leq p$. It reflects our prior belief on the overall group proportions and induces sparsity. The $\boldsymbol{D}$ is a p-by-p covariance matrix and will incorporate the mediator-wise correlation/structure dependency to the transformed mixing probabilities. In our setting, we estimate the correlation matrix among mediators from data and replace the negative correlations with their absolute values. We then find the nearest positive definite matrix to the absolute correlation matrix, and use that as $\boldsymbol{D}$ matrix in the model fitting. Since the variation may be different for $\text{logit}(\widetilde{\pi}_{j1})$, $\text{logit}(\widetilde{\pi}_{j2})$ and $\text{logit}(\widetilde{\pi}_{j3})$, we introduce the group-wise $\sigma^2_{dk}, k = 1, 2, 3$ for a more

general pattern. This correlation embedded GMM exploits the whole correlation information from all the mediators and does not require the predefined neighbors as in the Potts mixture model.

Following the idea of data augmentation (*Polson et al.*, 2013), we introduce the Pólya-Gamma variables for an effective fully Bayesian approach. The augmented posterior leads to conditional distributions from which we can easily draw samples and the entire vector $\boldsymbol{b}$ can be sampled as a block in a single Gibbs update. The detailed derivation and algorithm can be found in the Appendix.

4.3 Simulations

We evaluate the performance of the proposed models compared with existing methods under different scenarios through simulations.

4.3.1 Small Sample Scenarios: $n = 100, p = 200$

Simulation Design

Following settings in the previous chapters, we adopt the four-component structure to generate the exposure-mediator and mediator-outcome effects, i.e. simulate $[\beta_{mj}, \alpha_{aj}]^\top$ from

$$[\beta_{mj}, \alpha_{aj}]^\top \sim \pi_1 \text{MVN}(\mathbf{0}, \begin{bmatrix} 0.5 & 0.2 \\ 0.2 & 0.5 \end{bmatrix}) + \pi_2 \text{MVN}(\mathbf{0}, \begin{bmatrix} 0.5 & 0 \\ 0 & 0 \end{bmatrix}) + \pi_3 \text{MVN}(\mathbf{0}, \begin{bmatrix} 0 & 0 \\ 0 & 0.5 \end{bmatrix}) + \pi_4 \boldsymbol{\delta}_0$$

To introduce sparsity, we assume the proportion of active mediators $\pi_1 = 0.05$, and the other three null components $\pi_2 = 0.05, \pi_3 = 0.10, \pi_4 = 0.80$. We generate a p-vector of correlated mediators for the ith individual from $\boldsymbol{M}_i = A_i \boldsymbol{\alpha_a} + \boldsymbol{\epsilon}_{M_i}$, where the continuous exposure $\{A_i, i = 1, ..., n\}$ are independently from a standard normal distribution. The residual errors $\boldsymbol{\epsilon}_{M_i} \sim \text{MVN}(\mathbf{0}, \boldsymbol{\Sigma})$ and $\boldsymbol{\Sigma}$ models the correlation structure across mediators.

For the outcome, we simulate it from the linear model: $Y_i = \boldsymbol{M}_i^\top \boldsymbol{\beta_m} + A_i\beta_a + \epsilon_{Y_i}$, with $\beta_a = 0.5$, and the residual error $\epsilon_{Y_i} \sim \mathrm{N}(0, 1)$.

For the correlation structure, we assume 10 highly-correlated blocks of size 10×10, within which the pairwise correlation of mediators is ρ_1, e.g. $\rho_1 = 0.5 - 0.03|i-j|$ or $0.9 - 0.05|i-j|$, and the correlation between blocks (ρ_2) is relatively weak (e.g. $\rho_2 = 0$ or 0.1). Such correlation structure mimics the local dependency due to physical adjacency or biologically functional pathway of biomarkers, which is commonly seen in the high-dimensional mediators. There are 10 active mediators, and they are assumed to cluster within one block or scatter over a few blocks, while the other blocks contain no active mediators. We also consider settings where there is no correlation or such structural information underlying active mediators, that is, setting $\boldsymbol{\Sigma}$ to be identical matrix or estimated covariance based on a random subset of DNAm from MESA.

The Potts mixture model needs the input of a reliable neighborhood matrix. In practice, we may not be able to specify a completely precise neighborhood structure, but instead a deviated version of that. To examine how sensitive our Potts mixture model is to the incorrect neighborhood relationship, we randomly convert a proportion of r neighboring mediator pairs to be non-neighboring, and randomly convert the same amount of non-neighboring pairs to be neighbors. The other configurations are same as in the previous simulations. We vary the perturbation rate r from 0.05 to 0.5 to mimic different degrees of bias. In addition, for the Gaussian mixture model, since it directly takes the correlation matrix as an input, we examine its sensitivity to the observed correlation matrix by adding mild changes from $\mathrm{N}(0, \sigma^2)$ to the estimated matrix. We vary σ from 0.1 to 0.3 for different levels of noise.

Evaluation Metrics

To examine the mediator selection accuracy, for the proposed GMM-based methods, we use PIP to prioritize mediators as we did in the last two chapters. We calculate the true positive rate (TPR) for active mediators based on the fixed 10% false discovery rate (FDR). For

the estimation accuracy, we calculate the mean square error (MSE) of the indirect effects for both non-null and null mediators, denoted as $\text{MSE}_{\text{non-null}}$ and MSE_{null}. We perform 200 replicates for each scenario and report the means of those metrics in the result tables.

Competing Methods

In addition to the proposed methods, we consider the following existing methods: GMM with no correlated information included, Bi-Lasso (apply two separate Lasso (*Tibshirani*, 1996) regression to the outcome and mediator model, respectively), Bi-Ridge (apply two separate ridge (*Hoerl and Kennard*, 1988) regressions to the outcome and mediator model, respectively), and Pathway Lasso (*Zhao and Luo*, 2016). In Bi-Lasso and Bi-Ridge, we adopt 10-fold cross validation to choose the tuning parameter in each regression separately. The three frequentist methods provide optimized solutions of $\boldsymbol{\beta_m}$, $\boldsymbol{\alpha_a}$ to the three different penalized likelihoods, and the marginal indirect contribution from each mediator, i.e. $\beta_{mj}\alpha_{aj}$ is used to rank mediators for these methods.

Simulation Results

Table 4.1 shows the results under the small sample scenarios with $n = 100, p = 200$. Overall, by leveraging mediators' correlation structure, the two proposed approaches, GMM-Potts and GMM-CorrS, substantially improve the selection accuracy over the other methods. When the active mediators are concentrated within one block, the GMM-Potts achieves the highest TPR (> 0.90) at a fixed 10% FDR for identifying this whole block, followed by GMM-CorrS ($\sim$0.80 TPR). The advantage of the proposed methods grows with stronger correlations. Without such "group selection" ability, the GMM under independent priors tends to lose half of the power for detecting correlated mediators. On the other hand, if the active ones are evenly distributed into two blocks, then highly correlated mediators within the same block may not be concurrently active. This could happen if their correlation does not mainly link with mediation as we assume, and therefore may disturb mediator selection. Under those settings, we do observe power decrease for the proposed methods. Particularly,

the GMM-Potts model becomes less preferable as it smoothes over non-mediating neighbors to infer active mediators; while GMM-CorrS uses a more flexible Gaussian distribution for dependent group probabilities and thus has the best TPR. In the settings where there is no systematic correlation structure underlying mediators, we find that GMM-CorrS behaves quite similarly to the GMM, and outperforms the others. GMM-Potts is less robust presumably due to the inclusion of irrelevant neighbors, but still better than the frequentist methods. The three frequentist methods have relatively poor selection performance with highly correlated mediators, and Bi-Lasso is most competitive under zero or weak correlation. In terms of the effects estimation, the proposed methods mostly achieve the smallest $\text{MSE}_{\text{non-null}}$ and a reasonable level of MSE_{null}. Among the three frequentist methods, since in general Lasso tends to select less correlated variables than the elastic net type penalty, Bi-Lasso has a relatively larger $\text{MSE}_{\text{non-null}}$ but noticeably smaller MSE_{null} than the pathway Lasso. Given the sparse setup in the above simulations, Bi-Ridge does not exhibit much advantage over the other methods.

Tables 4.2 and 4.3 summarize the sensitivity analysis for GMM-Potts and GMM-CorrS, respectively, regarding the input correlation structure. As expected, with increasing noise added to the correlation structure, the overall accuracy of GMM-Potts and GMM-CorrS gets reduced. However, the power of our methods remains 75% of the original level for reasonable r and σ ($r < 0.3$, $\sigma < 0.3$). Even with large $r = 0.5$ and $\sigma = 0.3$, GMM-CorrS still has better performance (TPR, $\text{MSE}_{\text{non-null}}$) over methods with no structural information in all the settings, and GMM-Potts does for most of the settings. Generally speaking, the proposed methods are not sensitive to small alteration of the input correlation structure.

4.3.2 Large Sample Scenarios: $n = 1000, p = 2000$

Simulation Design

Next, we examine the settings for $n = 1000, p = 2000$. We simulate the exposure, exposure-

$\rho_1 = 0.5 - 0.03\|i-j\|, \rho_2 = 0$						
	(A) Signals in one block			(B) Signals in two blocks		
Method	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
GMM-CorrS	**0.78**	0.029	1.360	**0.62**	0.039	1.919
GMM-Potts	**0.93**	0.035	2.251	**0.49**	0.040	2.112
GMM	0.45	0.042	1.211	0.46	0.047	1.203
Bi-Lasso	0.26	0.238	0.520	0.23	0.238	0.584
Bi-Ridge	0.22	0.283	2.639	0.21	0.286	2.642
Pathway Lasso	0.24	0.233	2.598	0.23	0.180	6.405
$\rho_1 = 0.9 - 0.05\|i-j\|, \rho_2 = 0.1$						
	(A) Signals in one block			(B) Signals in two blocks		
Method	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
GMM-CorrS	**0.81**	0.208	1.146	**0.49**	0.182	4.080
GMM-Potts	**0.92**	0.171	3.515	**0.41**	0.233	1.651
GMM	0.33	0.206	2.158	0.22	0.201	3.112
Bi-Lasso	0.11	0.342	0.173	0.13	0.343	0.179
Bi-Ridge	0.15	0.322	2.170	0.16	0.326	1.690
Pathway Lasso	0.21	0.237	5.495	0.19	0.264	3.457
No systematic correlation structure (signals in two blocks)						
	(A) $\rho_1 = 0$			(B) Weak correlation from MESA		
Method	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
GMM-CorrS	**0.52**	0.020	1.042	**0.44**	0.023	1.780
GMM-Potts	0.46	0.043	1.970	0.40	0.030	3.041
GMM	**0.52**	0.021	0.805	**0.45**	0.023	1.642
Bi-Lasso	0.45	0.081	0.542	0.35	0.139	0.740
Bi-Ridge	0.35	0.238	3.645	0.28	0.247	4.003
Pathway Lasso	0.35	0.164	0.314	0.32	0.177	0.400
Each block contains only one signal						
	(A) $\rho_1 = 0.5 - 0.03\|i-j\|, \rho_2 = 0$			(B) $\rho_1 = 0.9 - 0.05\|i-j\|, \rho_2 = 0.1$		
Method	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
GMM-CorrS	0.46	0.022	1.488	**0.39**	0.025	2.284
GMM-Potts	**0.48**	0.027	2.336	**0.39**	0.028	3.035
GMM	**0.47**	0.021	1.291	**0.41**	0.025	2.089
Bi-Lasso	0.36	0.101	0.753	0.26	0.195	0.763
Bi-Ridge	0.27	0.251	3.298	0.22	0.277	2.363
Pathway Lasso	0.29	0.160	0.344	0.27	0.194	0.286

Table 4.1: Simulation results of $n = 100, p = 200$ under different correlation structures. TPR: true positive rate at false discovery rate (FDR) = 0.10. $\text{MSE}_{\text{non-null}}$: mean squared error for the indirect effects of active mediators. MSE_{null}: mean squared error for the indirect effects of inactive mediators. The results are based on 200 replicates for each setting. Bolded TPRs indicate the top two performers.

mediator and mediator-outcome effects using the same distribution as above. For the correlation structure, we now consider 50 blocks of size 20×20, with relatively high within-block

$\rho_1 = 0.5 - 0.03\|i-j\|, \rho_2 = 0$						
	(A) Signals in one block			(B) Signals in two blocks		
Perturbation rate	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
0	0.93	0.035	2.251	0.49	0.040	2.112
0.05	0.78	0.076	1.496	0.44	0.091	1.733
0.1	0.72	0.077	1.578	0.43	0.091	1.827
0.2	0.69	0.087	1.568	0.42	0.086	1.822
0.3	0.61	0.097	1.736	0.41	0.088	2.019
0.4	0.53	0.102	1.525	0.40	0.085	1.952
0.5	0.49	0.094	2.082	0.41	0.081	1.847
$\rho_1 = 0.9 - 0.05\|i-j\|, \rho_2 = 0.1$						
	(A) Signals in one block			(B) Signals in two blocks		
Perturbation rate	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
0	0.92	0.171	3.515	0.41	0.233	1.651
0.05	0.91	0.180	0.819	0.33	0.191	1.876
0.1	0.91	0.181	1.203	0.35	0.183	2.156
0.2	0.91	0.175	1.393	0.32	0.201	1.815
0.3	0.89	0.174	1.129	0.32	0.177	2.081
0.4	0.88	0.173	1.395	0.32	0.200	1.492
0.5	0.83	0.166	2.046	0.30	0.188	1.884

Table 4.2: Sensitivity analysis for Potts mixture model (GMM-Potts) for $n = 100, p = 200$.

$\rho_1 = 0.5 - 0.03\|i-j\|, \rho_2 = 0$						
	(A) Signals in one block			(B) Signals in two blocks		
Noise level	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$	TPR	$\text{MSE}_{\text{non-null}}$	MSE_{null}
0	0.78	0.029	1.360	0.62	0.039	1.919
0.1	0.71	0.029	2.481	0.56	0.036	2.246
0.2	0.60	0.031	2.575	0.50	0.037	2.043
0.3	0.53	0.033	2.235	0.47	0.037	1.910
$\rho_1 = 0.9 - 0.05\|i-j\|, \rho_2 = 0.1$						
	(A) Signals in one block			(B) Signals in two blocks		
Noise level	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
0	0.81	0.208	1.146	0.49	0.182	4.080
0.1	0.72	0.168	4.017	0.40	0.127	3.288
0.2	0.63	0.170	3.442	0.37	0.130	3.370
0.3	0.54	0.176	3.413	0.34	0.133	3.283

Table 4.3: Sensitivity analysis for the Gaussian mixture model with correlated selection (GMM-CorrS) for $n = 100, p = 200$.

mediator correlation ρ_1 and zero between-block correlation. We first set the four group proportions same as in the small sample scenarios, and the resultant 100 active mediators are

assumed to evenly distribute over five blocks. The other blocks contain no active mediators. In one of the settings, we use the covariance matrix estimated from a random subset of DNAm in MESA as $\mathbf{\Sigma}$ to simulate mediators with no underlying systematic correlation structure.

Then we study a much sparser setting with only 10 active mediators to better reflect the situation we observe in the MESA application. The 10 active mediators exist in two blocks, each of which contains five active ones and 15 inactive ones. Furthermore, we consider another worse-case scenario for GMM-Potts model by reducing ρ_1 to 0.2 and remaining the high sparsity. The weak correlation makes it hard for GMM-Potts model to identify the true neighboring relationship via the clustering method, and the performance of the Potts model is quite dependent on the smoothing effects from neighbors.

Simulation Results

Table 4.4 shows the results under the large sample scenarios with $n = 1000, p = 2000$. Our methods enjoy up to 30% power gain on mediator selection utilizing the correlation structure compared to the other methods. In the first setting, both methods identify almost all the active blocks, and GMM-Potts has a slightly higher TPR (0.97) at 10% FDR than GMM-CorrS (TPR = 0.92). When the mediator correlation has no implication for mediation effects in the second setting, the overall performance of GMM-CorrS is similar to that of GMM, and better than GMM-Potts. Those patterns are consistent with what we have observed in the small sample scenarios. Under the much sparser settings with only 10 active mediators and varied correlation ρ_1, the GMM-CorrS maintains good and stable performance with TPR around 0.80. By contrast, the performance of GMM-Potts is dependent on how obvious the correlation patterns are and subsequently how well the clustering method does in defining neighbors and non-neighbors. For example, with $\rho_1 = 0.5 - 0.02|i - j|$, the GMM-Potts models can accurately identify the underlying correlation structure and achieve the highest TPR (0.85), smallest MSE ($\text{MSE}_{\text{non-null}} = 0.002$, $\text{MSE}_{\text{null}} = 7.607 \times 10^{-7}$). However, as the

within-block correlation ρ_1 reduces to 0.25, it becomes challenging for the clustering method to separate true correlation versus noise, and we do observe many noisy pairs in the neighborhood matrix. As a consequence, the results of GMM-Potts model get compromised by the inclusion of those irrelevant neighbors. This setting is actually in agreement with our observation of the ambiguous correlation structure and sparse signals in the MESA application, which may not fare well for GMM-Potts model. Among the other three frequentist methods, Bi-Lasso performs best regarding to the selection and estimation accuracy.

	$p_{11} = 100$, Signals in five blocks					
	(A) $\rho_1 = 0.5 - 0.02\|i-j\|$			(B) Weak correlation from MESA		
Method	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
GMM-CorrS	**0.92**	0.031	0.440	**0.83**	0.002	0.240
GMM-Potts	**0.97**	0.030	0.018	0.76	0.004	1.013
GMM	0.76	0.077	0.630	**0.84**	0.002	0.176
Bi-Lasso	0.73	0.031	0.199	0.65	0.042	0.446
Bi-Ridge	0.32	0.244	2.680	0.36	0.202	3.795
Pathway Lasso	0.44	0.112	1.162	0.42	0.107	1.427
	$p_{11} = 10$, Signals in two blocks					
	(A) $\rho_1 = 0.5 - 0.02\|i-j\|$			(B) $\rho_1 = 0.25$		
Method	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$	TPR	$\text{MSE}_{\text{non-null}}$	$\text{MSE}_{\text{null}} \times 10^{-4}$
GMM-CorrS	**0.83**	0.003	0.015	**0.82**	0.002	0.017
GMM-Potts	**0.85**	0.002	0.008	0.61	0.018	0.228
GMM	0.80	0.003	0.013	**0.81**	0.002	0.016
Bi-Lasso	0.73	0.013	0.036	0.76	0.010	0.035
Bi-Ridge	0.41	0.061	1.508	0.39	0.063	1.517
Pathway Lasso	0.55	0.046	0.133	0.56	0.047	0.141
	$p_{11} = 10$, Each block contains only one signal					
	(A) $\rho_1 = 0.5 - 0.02\|i-j\|, \rho_2 = 0$			(B) $\rho_1 = 0.9 - 0.03\|i-j\|, \rho_2 = 0.1$		
Method	TPR	$\text{MSE}_{\text{non-null}} \times 10^{-3}$	$\text{MSE}_{\text{null}} \times 10^{-6}$	TPR	$\text{MSE}_{\text{non-null}} \times 10^{-3}$	$\text{MSE}_{\text{null}} \times 10^{-6}$
GMM-CorrS	**0.81**	2.317	1.484	**0.78**	4.501	4.920
GMM-Potts	**0.81**	2.892	0.724	0.73	7.257	4.076
GMM	**0.81**	2.396	1.256	**0.78**	4.511	4.970
Bi-Lasso	0.72	10.425	3.761	0.64	17.676	5.053
Bi-Ridge	0.50	14.084	15.273	0.41	28.471	14.049
Pathway Lasso	0.56	13.507	14.001	0.50	25.940	14.609

Table 4.4: Simulation results of $n = 1000, p = 2000$ under different correlation structures, p_{11} is the number of true active mediators. TPR: true positive rate at false discovery rate (FDR) = 0.10. $\text{MSE}_{\text{non-null}}$: mean squared error for the indirect effects of active mediators. MSE_{null}: mean squared error for the indirect effects of inactive mediators. The results are based on 200 replicates for each setting. Bolded TPRs indicate the top two performers.

Finally, we examine the empirical FDR estimates using the local FDR approach (*Efron et al.*,

2007) , median PIP cutoff, and 0.90 PIP cutoff for a targeted 10% FDR. For the local FDR approach, we compute the local false discovery rate for each mediator following *Efron et al.* (2007). We define the local false discovery rate for the j-th mediator being in the active group as locfdr_{j1}, and it can be expressed as $1 - P(\gamma_j = 1|\text{ Data})$. We first sort locfdr_{j1} from the smallest to the largest, where the jth ordered value is $\text{locfdr}_1^{(j)}, j = 1, ..., p$. Then the cutoff value c_1 for locfdr_{j1} to guarantee a 10% FDR can be identified from,

$$\underset{c_1}{\text{argmax}} \frac{1}{\sum_{j=1}^{p} I(\text{locfdr}_1^{(j)} < c_1)} \sum_{j=1}^{p} I(\text{locfdr}_1^{(j)} < c_1)\text{locfdr}_1^{(j)} < 0.1$$

where I is an indicator function. Following *Newton et al.* (2004), we declare mediators with an locfdr_{j1} smaller than the threshold c_1 as active mediators.

To evaluate the performance of those significance rules, we report the empirical FDR and TPR in Table 4.5 and 4.6 under all the simulation scenarios. Under the small sample scenarios (Table 4.5), the local FDR approach provides decent and well-controlled empirical FDR for both of the proposed methods, while the estimates by median PIP cutoff and 0.90 PIP cutoff tend to be either slightly overestimated or very conservative. Under the large sample scenarios (Table 4.6), the local FDR approach and median PIP cutoff still produces reasonable FDR estimates for GMM-CorrS across different settings and for GMM-Potts when neighbors reflect connected signals. However, including irrelevant neighbors in GMM-Potts could lead to increased false discoveries, and instead a more stringent 0.90 PIP cutoff may be used if one seeks a lower limit on the false discovery. Without accounting for correlation structure, the GMM method does not tend to select correlated mediators and therefore mostly has lower empirical FDR estimates, accompanied with lower TPR estimates when active mediators are correlated. Therefore in practice, we would recommend the local FDR and median PIP cutoff for reasonable FDR estimates and control, and we recognize the potential caveat concerning inflated FDR for GMM-Potts.

To summarize our findings from the simulations, GMM-CorrS takes the overall correlation

Method	TPR	TPR(locfdr)	FDR(locfdr)	TPR(PIP>0.5)	FDR(PIP>0.5)	TPR(PIP>0.9)	FDR(PIP>0.9)
$\rho_1 = 0.5 - 0.03\|i-j\|, \rho_2 = 0$, Signals in one block							
GMM-CorrS	0.78	0.69(0.021)	0.05(0.008)	0.82(0.019)	0.12(0.012)	0.49(0.016)	0.02(0.006)
GMM-Potts	0.93	0.79(0.019)	0.05(0.007)	0.86(0.014)	0.07(0.010)	0.61(0.017)	0.01(0.002)
GMM	0.45	0.35(0.011)	0.03(0.009)	0.41(0.012)	0.08(0.013)	0.29(0.009)	0.02(0.007)
$\rho_1 = 0.5 - 0.03\|i-j\|, \rho_2 = 0$, Signals in two blocks							
GMM-CorrS	0.62	0.52(0.018)	0.07(0.010)	0.67(0.021)	0.14(0.012)	0.40(0.015)	0.01(0.005)
GMM-Potts	0.49	0.34(0.041)	0.06(0.025)	0.66(0.023)	0.22(0.022)	0.24(0.032)	0.02(0.017)
GMM	0.46	0.36(0.011)	0.02(0.007)	0.44(0.013)	0.07(0.011)	0.29(0.009)	0.01(0.004)
$\rho_1 = 0.9 - 0.05\|i-j\|, \rho_2 = 0.1$, Signals in one block							
GMM-CorrS	0.81	0.49(0.020)	0.06(0.013)	0.83(0.014)	0.17(0.007)	0.36(0.018)	0.02(0.015)
GMM-Potts	0.92	0.51(0.043)	0.05(0.015)	0.83(0.049)	0.08(0.014)	0.23(0.014)	0.01(0.012)
GMM	0.33	0.19(0.014)	0.03(0.014)	0.25(0.014)	0.09(0.018)	0.017(0.010)	0.01(0.008)
$\rho_1 = 0.9 - 0.05\|i-j\|, \rho_2 = 0.1$, Signals in two blocks							
GMM-CorrS	0.49	0.31(0.032)	0.09(0.023)	0.55(0.032)	0.23(0.023)	0.22(0.021)	0.05(0.020)
GMM-Potts	0.40	0.27(0.006)	0.06(0.005)	0.43(0.038)	0.17(0.022)	0.17(0.006)	0.04(0.008)
GMM	0.22	0.19(0.012)	0.06(0.030)	0.28(0.012)	0.15(0.022)	0.14(0.009)	0.03(0.032)
$\rho_1 = 0$, Signals in two blocks							
GMM-CorrS	0.52	0.44(0.015)	0.03(0.008)	0.50(0.015)	0.07(0.012)	0.35(0.012)	0.01(0.004)
GMM-Potts	0.46	0.42(0.022)	0.06(0.016)	0.50(0.016)	0.19(0.019)	0.33(0.016)	0.02(0.011)
GMM	0.52	0.42(0.014)	0.02(0.006)	0.48(0.014)	0.05(0.008)	0.33(0.010)	0.01(0.003)
Weak correlation from MESA, Signals in two blocks							
GMM-CorrS	0.44	0.32(0.009)	0.03(0.009)	0.39(0.011)	0.08(0.013)	0.27(0.007)	0.01(0.006)
GMM-Potts	0.40	0.35(0.014)	0.07(0.015)	0.45(0.017)	0.23(0.022)	0.27(0.010)	0.03(0.011)
GMM	0.45	0.33(0.010)	0.03(0.009)	0.39(0.011)	0.06(0.011)	0.27(0.007)	0.01(0.005)
Each block contains only one signal, $\rho_1 = 0.5 - 0.03\|i-j\|, \rho_2 = 0$							
GMM-CorrS	0.46	0.36(0.012)	0.06(0.011)	0.44(0.016)	0.10(0.013)	0.28(0.009)	0.01(0.005)
GMM-Potts	0.48	0.38(0.016)	0.04(0.010)	0.53(0.015)	0.17(0.017)	0.29(0.012)	0.02(0.007)
GMM	0.47	0.37(0.013)	0.04(0.009)	0.43(0.014)	0.08(0.012)	0.29(0.009)	0.01(0.005)
Each block contains only one signal, $\rho_1 = 0.9 - 0.05\|i-j\|, \rho_2 = 0.1$							
GMM-CorrS	0.39	0.29(0.008)	0.04(0.010)	0.35(0.011)	0.08(0.013)	0.24(0.006)	0.01(0.005)
GMM-Potts	0.39	0.31(0.013)	0.06(0.013)	0.41(0.015)	0.17(0.023)	0.24(0.010)	0.02(0.008)
GMM	0.41	0.29(0.008)	0.04(0.009)	0.36(0.011)	0.07(0.012)	0.24(0.006)	0.01(0.005)

Table 4.5: Empirical estimates of TPR and FDR in simulations of $n = 100, p = 200$. The results are based on 200 replicates for each setting, and the standard errors are shown within parentheses. TPR is the true positive rate controlled at a fixed FDR of 10%; TPR(locfdr) and FDR(locfdr) are the empirical estimates based on the local FDR approach; TPR(PIP>0.9) and FDR(PIP>0.9) are the empirical estimates when the PIP threshold for identifying active mediators is 0.9; TPR(PIP>0.5) and FDR(PIP>0.5) are the empirical estimates when the PIP threshold for identifying active mediators is 0.5.

structure among the mediators directly into modeling process, and shows excellent performance and robustness under different correlation structures. On the other hand, the performance of GMM-Potts is related to how well the specified neighborhood matrix reflects the underlying connection of active mediators. When the correlation-based neighboring has good implication on the mediation effects, then GMM-Potts usually achieves the best selection and estimation accuracy. Its performance will likely get compromised by the inclusion

Method	TPR	TPR(locfdr)	FDR(locfdr)	TPR(PIP>0.5)	FDR(PIP>0.5)	TPR(PIP>0.9)	FDR(PIP>0.9)
$\rho_1 = 0.5 - 0.02\|i-j\|, p_{11} = 100$, Signals in five block							
GMM-CorrS	0.92	0.90(0.001)	0.08(0.002)	0.88(0.002)	0.02(0.012)	0.80(0.003)	0.00(0.001)
GMM-Potts	0.97	0.96(0.002)	0.09(0.002)	0.96(0.002)	0.01(0.002)	0.93(0.002)	0.00(0.002)
GMM	0.76	0.49(0.004)	0.04(0.002)	0.48(0.003)	0.03(0.002)	0.36(0.004)	0.01(0.001)
Weak correlation from MESA, $p_{11} = 100$, Signals in five blocks							
GMM-CorrS	0.83	0.83(0.002)	0.12(0.003)	0.81(0.003)	0.04(0.003)	0.77(0.003)	0.00(0.001)
GMM-Potts	0.76	0.86(0.017)	0.33(0.022)	0.88(0.010)	0.35(0.024)	0.82(0.011)	0.16(0.017)
GMM	0.84	0.84(0.002)	0.10(0.002)	0.81(0.002)	0.03(0.002)	0.77(0.002)	0.00(0.001)
$\rho_1 = 0.5 - 0.02\|i-j\|, p_{11} = 10$, Signals in two blocks							
GMM-CorrS	0.83	0.82(0.007)	0.09(0.006)	0.81(0.007)	0.05(0.007)	0.74(0.009)	0.01(0.003)
GMM-Potts	0.85	0.88(0.007)	0.34(0.012)	0.95(0.007)	0.65(0.008)	0.83(0.008)	0.04(0.012)
GMM	0.80	0.80(0.006)	0.08(0.007)	0.78(0.007)	0.03(0.006)	0.74(0.008)	0.00(0.002)
$\rho_1 = 0.25, p_{11} = 10$, Signals in two blocks							
GMM-CorrS	0.82	0.80(0.006)	0.09(0.007)	0.80(0.006)	0.06(0.007)	0.74(0.007)	0.01(0.003)
GMM-Potts	0.61	0.79(0.018)	0.35(0.043)	0.81(0.011)	0.56(0.037)	0.75(0.010)	0.15(0.047)
GMM	0.81	0.81(0.006)	0.08(0.006)	0.80(0.006)	0.05(0.007)	0.77(0.007)	0.01(0.003)
Each block contains only one signal, $\rho_1 = 0.5 - 0.02\|i-j\|, \rho_2 = 0, p_{11} = 10$							
GMM-CorrS	0.81	0.81(0.007)	0.10(0.007)	0.80(0.006)	0.05(0.008)	0.76(0.007)	0.00(0.002)
GMM-Potts	0.81	0.80(0.006)	0.06(0.007)	0.79(0.006)	0.04(0.007)	0.74(0.007)	0.00(0.000)
GMM	0.81	0.81(0.006)	0.08(0.006)	0.80(0.006)	0.04(0.007)	0.76(0.007)	0.00(0.002)
Each block contains only one signal, $\rho_1 = 0.9 - 0.03\|i-j\|, \rho_2 = 0.1, p_{11} = 10$							
GMM-CorrS	0.78	0.78(0.008)	0.07(0.007)	0.77(0.008)	0.04(0.007)	0.69(0.009)	0.01(0.003)
GMM-Potts	0.73	0.73(0.009)	0.10(0.019)	0.72(0.010)	0.09(0.023)	0.66(0.009)	0.04(0.016)
GMM	0.78	0.78(0.008)	0.08(0.007)	0.77(0.008)	0.05(0.007)	0.70(0.009)	0.01(0.003)

Table 4.6: Empirical estimates of TPR and FDR in simulations of $n = 1000, p = 2000$, p_{11} is the number of true active mediators. The results are based on 200 replicates for each setting, and the standard errors are shown within parentheses. TPR is the true positive rate controlled at a fixed FDR of 10%; TPR(locfdr) and FDR(locfdr) are the empirical estimates based on our PIP approach; TPR(PIP>0.9) and FDR(PIP>0.9) are the empirical estimates when the PIP threshold for identifying active mediators is 0.9; TPR(PIP>0.5) and FDR(PIP>0.5) are the empirical estimates when the PIP threshold for identifying active mediators is 0.5.

of irrelevant neighbors.

4.4 Data Application

In this section, we study two real data applications of the proposed methods: the LIFECODES birth cohort and the Multi-Ethnic Study of Atherosclerosis (MESA).

4.4.1 The LIFECODES Birth Cohort

In this application, we consider a subset set of $n = 161$ pregnant women registered at the Brigham and Women's Hospital in Boston, MA between 2006 and 2008. We are interested

in the mediation mechanism linking environmental contaminant exposure during pregnancy to preterm birth through endogenous signaling molecules. Those endogenous biomarkers are derived from lipids, peptides, and DNA, and were measured with subjects' urine and plasma specimens collected at one study visit between 23.1 and 28.9 weeks gestation. We focus on $p = 61$ available endogenous biomarkers as potential mediators, including 51 eicosanoids, five oxidative stress biomarkers and five immunological biomarkers. The correlation structure across mediators are shown in Figure 4.1, and clear pattern with moderate to strong correlations can be observed. For the prenatal exposure to environmental toxicants, we focus the attention of this present study on one class of environmental contaminants, polycyclic aromatic hydrocarbons (PAHs). PAHs are a group of organic contaminants that form due to the incomplete combustion of hydrocarbons, and commonly present in tobacco smoke, smoked and grilled food products, polluted water and soil, vehicle exhaust gas (*Alegbeleye et al.*, 2017). Previous studies have suggested association between PAH exposure and adverse birth outcomes (*Padula et al.*, 2014). Since the PAH class contains multiple chemical analytes in our study, we follow *Aung et al.* (2020) to construct an environmental risk score for the PAH class and use that risk score as the exposure variable. The continuous birth outcome, gestational age, was recorded at delivery for each participant, and preterm is defined as delivery prior to 37 weeks gestation. Since the cohort is oversampled for preterm cases, we multiply the data by the case-control sampling weights to adjust for that. We log-transform all measurements of the exposure metabolites and endogenous biomarkers. We apply the proposed methods with the aforementioned exposure, mediator and outcome variables, controlling for age and maternal BMI from the initial visit, race, and urinary specific gravity levels in both regressions of the mediation analysis.

The results are summarized in Table 4.7. Based on 10% FDR using the local FDR approach, GMM-Potts identifies four biomarkers for actively mediating the impact of PAH exposure on gestational age at delivery, 8,9-epoxy-eicosatrienoic acid (8(9)-EET), 9,10-dihydroxy-octadecenoic acid (9,10-DiHOME), 12,13-epoxy-octadecenoic acid (12(13)-EpoME), 9-oxooctadeca-

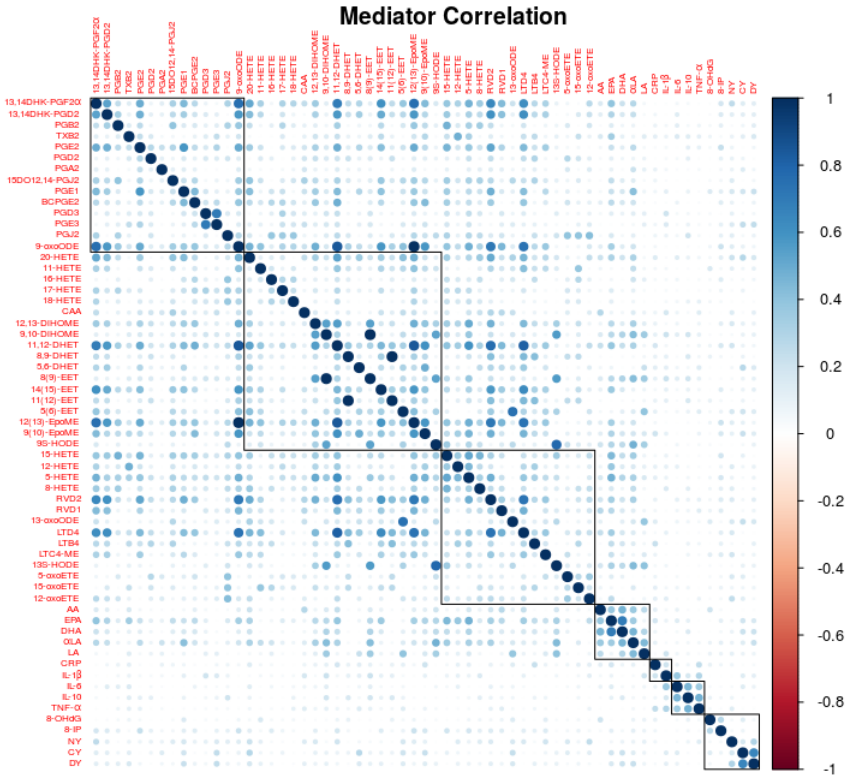

Figure 4.1: Correlations among biomarkers in LIFECODES birth cohort. The negative correlations (∼37% of all the pairwise correlations) were replaced with their absolute values. The 61 biomarkers were grouped by biological pathways (black lines).

dienoic acid (9-oxoODE); while both GMM-CorrS and GMM only identifies two of them, 8(9)-EET and 9,10-DiHOME. Among the four biomarkers, 8(9)-EET, 9,10-DiHOME and 12(13)-EpoME belong to the same Cytochrome p450 (CYP450) Pathway; while 9-oxoODE is within Cyclooxygenase (COX) Pathway. CYP450 is a family of enzymes that function to metabolize environmental toxicants, drugs, and endogenous compounds (*Sadler et al.*, 2016), and thus the PAH exposure may cause perturbations in the functions of these enzymes. It has also been suggested that the group of CYP450 metabolites as well as the related genes may play a role in the etiology of preterm delivery (*Banerjee et al.*, 2014), and the underlying mechanisms involve increased maternal oxidative stress and inflammation (*Ferguson and Chin*, 2017). Those evidence helps explain the potential mediating mechanism of CYP450 metabolites from PAH exposure to preterm delivery. Additionally, single biomarker analysis also demonstrated the protective effect of 12(13)-EpoME on preterm (*Aung et al.*, 2019). We also performed the posterior predictive checks on the outcome model for the three methods, in which the data generated from the posterior predictive distribution are compared with the observed outcome. We find the Bayesian predictive P-values (*Neelon et al.*, 2010) of the GMM-Potts model are 0.72 and 0.48 for sample first and second moments, respectively, which are closest to 0.5 among the three methods and indicate the most adequate fit of the outcome model.

Besides the estimated correlation structure, we also consider the input of biological pathway based structural information. That is, only mediators within the same biological pathway are neighbors in GMM-Potts and has non-zero pairwise correlation in GMM-CorrS. The findings are shown in the last two rows in Table 4.7. GMM-Potts identifies a subset of the above four biomarkers: 8(9)-EET, 9,10-DiHOME, and GMM-CorrS declares the other two biomarkers as active mediators: 12(13)-EpoME, 9-oxoODE. The overlapping lists of active mediators add confidence to our findings, and also reveal the fact that only adjusting for biological pathways may lose the correlated information between different pathways. In addition to PIP, we also report the indirect effect estimates and their 95% credible intervals for selected

mediators. We note that the direction of effects are consistent among different methods.

Method	Selected Mediators	PIP	$\hat{\beta}_{mj}\hat{\alpha}_{aj}$ (95% CI)
Polycyclic aromatic hydrocarbons $\rightarrow$ *Biomarkers* $\rightarrow$ *Gestational Age*			
GMM-Potts^0	12(13)-EpoME	0.99	0.419(0.295, 0.579)
	8(9)-EET	0.98	0.368(0.179, 0.567)
	9-oxoODE	0.97	-0.296(-0.441, 0.000)
	9,10-DiHOME	0.87	-0.185(-0.383, 0.000)
GMM-CorrS^0	8(9)-EET	1.00	0.698(0.391, 1.005)
	9,10-DiHOME	0.85	-0.345(-0.688, 0.000)
GMM-Potts^1	8(9)-EET	0.99	0.698(0.391, 1.005)
	9,10-DiHOME	0.91	-0.359(-0.671, 0.000)
GMM-CorrS^1	12(13)-EpoME	1.00	1.132(0.867, 1.426)
	9-oxoODE	0.99	-0.834(-1.119, -0.535)
GMM	8(9)-EET	1.00	0.698(0.407, 0.990)
	9,10-DiHOME	0.99	-0.394(-0.693, -0.091)

Table 4.7: Summary of the identified active mediators from the data application on LIFECODES study based on 10% FDR with the local FDR approach. GMM-Potts^0: define neighboring based on biomarker correlation clustering; GMM-Potts^1: define neighboring based on biological pathways; GMM-CorrS^0: use the estimated correlation matrix from data; GMM-CorrS^1: use the neighborhood structure based on biological pathways; GMM-CorrS^2: use the neighborhood structure based on biomarker correlation clustering. Besides the PIP, we also report the effect estimation $\hat{\beta}_{mj}\hat{\alpha}_{aj}$ and its 95% credible interval.

4.4.2 The MESA Cohort

In this application, we study the mediation mechanism of DNAm in the pathway from neighborhood socioeconomic disadvantage to glucose. We focus on $n = 1226$ participants with no missing data, and a subset of $p = 2000$ CpG sites that have the strongest marginal associations with neighborhood disadvantage for computational reasons. As the exposure, neighborhood socioeconomic disadvantage evaluates the neighborhood physical and social conditions from dimensions of education, occupation, income and wealth, poverty, employment, and housing. Previous literature has demonstrated the relationship between DNA methylation patterns and socially patterned stressors including low adult socioeconomic status (SES) (*Needham et al.*, 2015), unfavorable neighborhood conditions (*Smith et al.*, 2017), and neighborhood crime (*Lei et al.*, 2015). It has also been long known that disadvantaged

neighborhood conditions can lead to a variety of health problems, such as chronic psychological distress (*Ross and Mirowsky*, 2009), obesity (*Moore et al.*, 2013) and increased risk of cardiovascular disease (*Kaplan and Keil*, 1993). The outcome, glucose, is one of the most important blood parameters and should be kept within a safe range in order to support vital body functions and reduce the risk of diabetes and heart disease (*Sasso et al.*, 2004). Multiple evidence has supported the association between glucose metabolism and differential DNAm patterns (*Zheng et al.*, 2014; *Kriebel et al.*, 2016). However, the underlying molecular mechanisms that link neighborhood conditions to physical health profiles are not fully elucidated. To take a step forward, we apply the proposed methods for high-dimensional mediation analysis. In the outcome model, we adjust for age, gender, race/ethnicity, childhood SES and adult SES. In the mediator model, we control for age, gender, race/ethnicity, childhood SES, adult SES, and enrichment scores for 4 major blood cell types (neutrophils, B cells, T cells and natural killer cells). All the continuous variables are standardized to have zero mean and unit variance. In general, the correlation among DNAm is relatively weak, and only 3% of DNAm pairs have correlation larger than 0.2.

The results can be found in Table 4.8. Because of the relatively ambiguous correlation structure observed across mediators in MESA, we do not expect big improvement from our methods. Indeed, the GMM-CorrS identifies one more CpG site as active mediators compared to GMM, and three other CpG sites are detected by both GMM-CorrS and GMM. The rank correlation for the mediator rank lists obtained from the two methods is 0.74, indicating the high consistency between them. The indirect effect estimates from the GMM-CorrS are also close to those from the GMM model. The one additional finding of CpG site by GMM-CorrS, cg27090988, is close to the gene *OGG1*. This gene, which is involved in the repair of oxidative DNA damage, has been shown up-regulated in type 2 diabetic islet cell mitochondria, and studies have suggested a crucial role of oxidative DNA damage in the pathogenesis of type 2 diabetes (T2D) (*Tyrberg et al.*, 2002; *Pan et al.*, 2007). We also examine the nearby genes to the other three jointly selected CpG sites. Among them,

MYBPC3 is a known cardiomyopathy gene (*Dhandapany et al.*, 2009), and the increased risk of cardiac hypertrophy and heart failure is likely to alter the glucose metabolism (*Tran and Wang*, 2019); the expression level of *CD101* was found associated with T2D in a Mendelian randomization analysis (*Xue et al.*, 2018). As shown in the simulations, GMM-Potts is not quite suitable for a weak correlation structure as in the MESA data, and the method does not identify any active mediators based on 10% FDR.

Method	Selected Mediators	Nearby Genes	PIP	$\hat{\beta}_{mj}\hat{\alpha}_{aj}$ (95% CI)
Neighborhood SES $\rightarrow$ *Biomarkers* $\rightarrow$ *Glucose*				
GMM-CorrS	cg19515398	EIF2C2	0.97	-0.013(-0.026, 0.000)
	cg04000940	MYBPC3	0.96	0.016(0.000, 0.029)
	cg17907003	CD101	0.88	0.016(0.000, 0.034)
	cg27090988	OGG1	0.84	-0.011(-0.024, 0.000)
GMM	cg04000940	MYBPC3	0.97	0.015(0.000, 0.028)
	cg19515398	EIF2C2	0.94	-0.012(-0.024, 0.000)
	cg17907003	CD101	0.85	0.015(0.000, 0.032)

Table 4.8: Summary of the identified active mediators from the data application on MESA study based on 10% FDR using the local FDR approach. We include the nearby gene and PIP for each selected CpG site. The GMM-Potts does not identify any active mediators based on 10% FDR. Besides the PIP, we also report the effect estimation $\hat{\beta}_{mj}\hat{\alpha}_{aj}$ and its 95% credible interval.

We note that the validity of identifiability assumptions cannot be verified empirically from the observed data (*Little and Rubin*, 2019), and we can only justify our selection of covariates based on scientific knowledge. The influence of violating those identifiability assumptions can be assessed using sensitivity analysis, which has been well-developed for the single mediator setting (*Imai et al.*, 2010b; *Smith and VanderWeele*, 2019). For example, if we just focused on the single mediator of 9-oxoODE, then a correlation of 0.1 between the residuals of the mediator and outcome models would explain away the indirect effect. Additional work is required to extend that approach to the high-dimensional mediator setting.

4.5 Discussion

In this chapter, we present two hierarchical Bayesian approaches to incorporating the correlation structure across mediators in high-dimensional mediation analysis: (1) through a logistic normal for mixing probabilities (GMM-CorrS), or (2) through a Potts distribution on the group indicators (GMM-Potts). The consequent "non-separable" priors of both methods inform the grouping and selection of correlated mediators under the composite structure of mediation. The simulation studies show that utilizing the correlation pattern in mediators, the proposed methods greatly enhance the selection and estimation accuracy over the methods that do not account for such correlation, and maintain decent and comparable performance under no obvious or mis-specified correlation structure. In addition, the analysis on the LIFECODES birth cohort and MESA cohort indicates that our methods can promote the detection of new active mediators, which may have important implications on future research in targeted interventions and treatment for preterm birth and diabetes.

There are several limitations of the proposed methods. First, for GMM-Corr, it requires the inversion of a $p \times p$ matrix in each iteration of the sampling algorithm, and as p increases to the scale of hundreds of thousands, that step could become the computational bottleneck of the method. Techniques on matrix approximation or fast parallel matrix inversion will be required to speed up the computing time and reduce the memory footprint. Second, for GMM-Potts, smoothing over arbitrary or inaccurately specified neighbors may have a negative effect on its performance, and this can be further improved by imposing adaptive weight for each neighbor to reflect their relative importance. Moreover, the method can be extended to allow for simultaneous inference of both the active mediators and the neighborhood/network structure linking them. In that way, the neighborhood/network structure among mediators does not need to be known *a priori*.

As promising directions for future work, we note that there may be other ways to incorporate mediators' correlation into the modeling process. Recently, testing the multivariate media-

tion effects from groups of potential mediators has received growing attention (*Djordjilović et al.*, 2019), and the variance component tests developed by *Huang* (2019) can naturally take into account the correlation within groups. Also, *Bobb et al.* (2015) develops a Bayesian kernel machine regression to incorporate the structure of the multi-pollutant mixtures into the hierarchical model. Those methodologies may provide insightful perspectives to applying correlation kernels under the global testing setup in the context of high-dimensional mediation analysis.

CHAPTER V

Conclusion

The use of rigorous causal mediation analysis has been limited for studying the indirect effects of social/neighborhood/environmental factors on health outcomes, mediated through markers of biological pathways. Mediation analysis can help understand mechanisms underlying the etiology of chronic diseases. Most of the existing methods for causal mediation analysis can only analyze one or a moderate number of multiple mediators. In this dissertation, our main goal is to develop Bayesian methods that can handle high-dimensional mediators and perform efficient selection of active mediators. We first described the counterfactual framework for multiple-mediator analysis, and justified estimands with causal interpretation and the conditions that are needed for such interpretation. Estimation under a causal mediation analysis involves regression coefficients from two models: one with the outcome-mediator-exposure and the other with mediator-exposure. Since the natural indirect effect involves sum of products of coefficients from these models (summed over the mediator set), any shrinkage method for identifying active mediators needs to target this estimand. Due to the nature of the composite null hypothesis underlying a mediation analysis, one has to be thoughtful while specifying appropriate priors for selection and shrinkage. Motivated by this objective, in each chapter we proposed Bayesian shrinkage methods with a sparsity assumption to jointly analyze high-dimensional mediators and identify individual non-null indirect effects. Since our proposed methods rely on exact posterior sampling, we can char-

acterize uncertainty in estimation and provide estimates of other quantities of interest (e.g. proportion mediated) without relying on large-sample approximations.

The method proposed in Chapter II introduces Bayesian variable selection models using continuous mixture priors to high-dimensional mediation analysis, and is shown to yield excellent performance in mediator selection. The flexibility of a mixture prior specification in a Bayesian framework allows for penalization of regression coefficients from the two key models. Since our goal is identification of important mediators, the desired target is the contribution of each mediator to the composite NIE. Therefore, in Chapter III, we improve upon existing methods by proposing novel joint priors on exposure-mediator effects and mediator-outcome effects to enable targeted penalization on the indirect effect. In Chapter IV, we build on the work in Chapters II and III to explicitly incorporate the correlation structure among mediators to inform active mediator selection. The proposed methods are shown to have superior group selection ability when active mediators are correlated and robustness under different correlation structures. In all of the above Bayesian methods, the mediator-specific posterior inclusion probability provides a valid selection criterion for active mediators that contribute to the NIE. Our methodologies developed in this dissertation bridge an important gap in the literature for high-dimensional mediation analysis. The advantage of our methods over state-of-the-art existing methods has been illustrated in the applications of both genetic and environmental data. The proposed methods identified multiple genes/biomarkers for mediating the social/neighborhood/environmental factors on health outcomes. Those are important findings as researchers try to characterize how the insults from our external/social environment impact the internal cellular environment, and finally manifest into development of chronic diseases. To reach a broader audience, we have been working on easy-to-use R packages for our methods, and you can find the latest updates in `https://github.com/umich-cphds`.

The work presented in this dissertation points to many areas of potential future research. With the increasing availability of multi-platform data, it would be desirable to extend our

methods to accommodate discrete outcome/mediators in a principled way and place the methods into the generalized linear model framework. That will lead to wider applications of our methods. We ignored the possible presence of exposure-mediator interactions. How to interpret the causal mediation effects and perform efficient inference to characterize each component of a four-way decomposition (*VanderWeele*, 2016a) would be another direction of future work. Due to the large number of sampling iterations required for reasonable convergence, the current software for Bayesian mediation analysis is not quite computationally efficient in ultra-high dimensional settings. Approximation techniques such as variational Bayes or fast matrix computation could be explored to further improve the efficiency of our algorithms as the dimension increases to the scale of hundreds of thousands of mediators. Additionally, when individual-level data is not readily available, we will need to adapt our high-dimensional mediation analysis methods to construct models based on summary statistics only. We can take the approach in *Zhu and Stephens* (2017) for inference and extend that to regression with multivariate dependent variables for the mediator model. How to reduce the dimension of multiple exposures is also something we have been grappling with (*Aung et al.*, 2020). Another concern from a methodological perspective is that we do not have a good way to control the false discovery rate beyond computationally intensive permutation-based approach. This is a problem that prevails in all chapters, and a valid procedure to assess the false discovery rate is important for the error rate control and an objective comparison across different methods. The approach adopted in this dissertation work (*Newton et al.*, 2004) tends to be sensitive to prior specification, while the ordering of the PIP's is relatively robust. Ranking-based selection rules may be developed, and efficient methods for empirical null construction may also be possible directions to address this problem.

The proposed methods in this dissertation can be applied to a wide spectrum of mediation problems. We believe these methods will facilitate our understanding of the underlying biological mechanisms of complex diseases/traits and advance discovery of intervention and prevention strategies. We hope that this dissertation contributes to statistical methods for

mediation analysis and provides important insights to future research on high-dimensional mediation models.

APPENDICES

APPENDIX A

Supplement for Chapter II

A.1 Detailed Proofs

A.1.1 *Proof of Equation (2.1)*

Given the assumptions in 2.2, we can express $E[Y_i(a, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}]$ as below,

$$
\begin{aligned}
& E[Y_i(a, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}] \\
= & \int_{\boldsymbol{m}} E[Y_i(a, \boldsymbol{m})|\boldsymbol{C_i}, \boldsymbol{M_i}(a^\star) = \boldsymbol{m}] P(\boldsymbol{M_i}(a^\star) = \boldsymbol{m}|\boldsymbol{C_i}) \mathrm{d}\boldsymbol{m} \\
= & \int_{\boldsymbol{m}} E[Y_i(a, \boldsymbol{m})|\boldsymbol{C_i}] P(\boldsymbol{M_i}(a^\star) = \boldsymbol{m}|\boldsymbol{C_i}, A_i = a^\star) \mathrm{d}\boldsymbol{m} \\
& \text{(by assumption (4) \& (3))} \\
= & \int_{\boldsymbol{m}} E[Y_i(a, \boldsymbol{m})|A_i = a, \boldsymbol{C_i}] P(\boldsymbol{M_i}(a^\star) = \boldsymbol{m}|\boldsymbol{C_i}, A_i = a^\star) \mathrm{d}\boldsymbol{m} \\
& \text{(by assumption (1))} \\
= & \int_{\boldsymbol{m}} E[Y_i(a, \boldsymbol{m})|A_i = a, \boldsymbol{M_i} = \boldsymbol{m}, \boldsymbol{C_i}] P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, A_i = a^\star) \mathrm{d}\boldsymbol{m} \\
& \text{(by assumption (2) and consistency)} \\
= & \int_{\boldsymbol{m}} E(Y_i|a, \boldsymbol{m}, \boldsymbol{C_i}) P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a^\star) \mathrm{d}\boldsymbol{m} \\
& \text{(by consistency)}
\end{aligned}
$$

A.1.2 *Proof of Equation (2.6), (2.7)*

$$
\begin{aligned}
\text{NDE:} \quad & E[Y_i(a, \boldsymbol{M_i}(a^\star)) - Y_i(a^\star, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}] \\
= \; & \int_{\boldsymbol{m}} \{E(Y_i|a, \boldsymbol{m}, \boldsymbol{C_i}) - E(Y_i|a^\star, \boldsymbol{m}, \boldsymbol{C_i})\} P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a^\star) \mathrm{d}\boldsymbol{m} \\
= \; & \int_{\boldsymbol{m}} (a\beta_a - a^\star\beta_a) P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a^\star) \mathrm{d}\boldsymbol{m} \\
= \; & \beta_a(a - a^\star) \\
\text{NIE:} \quad & E[Y_i(a, \boldsymbol{M_i}(a)) - Y_i(a, \boldsymbol{M_i}(a^\star))|\boldsymbol{C_i}] \\
= \; & \int_{\boldsymbol{m}} E(Y_i|a, \boldsymbol{m}, \boldsymbol{C_i}) \{P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a) - P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a^\star)\} \mathrm{d}\boldsymbol{m} \\
= \; & \int_{\boldsymbol{m}} (\boldsymbol{m}^T\boldsymbol{\beta_m} + a\beta_a + \boldsymbol{C_i}^T\boldsymbol{\beta_c}) \{P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a) \\
& - P(\boldsymbol{M_i} = \boldsymbol{m}|\boldsymbol{C_i}, a^\star)\} \mathrm{d}\boldsymbol{m} \\
= \; & \{E(\boldsymbol{M_i}|\boldsymbol{C_i}, a) - E(\boldsymbol{M_i}|\boldsymbol{C_i}, a^\star)\}^T \boldsymbol{\beta_m} \\
= \; & \boldsymbol{\alpha_a}^T \boldsymbol{\beta_m} (a - a^\star) \\
= \; & (a - a^\star) \sum_{j=1}^{p} (\boldsymbol{\alpha_a})_j (\boldsymbol{\beta_m})_j
\end{aligned}
$$

A.2 Posterior Sampling Algorithm Details for Bayesian Mediation Analysis

Sampling β_{mj} and r_{mj}

$$
\log p(\beta_{mj}|r_{mj} = 1, .) = -\frac{\beta_{mj}^2}{2\sigma_{m1}^2} - \sum_{i=1}^{n} \{\frac{(M_i^{(j)}\beta_{mj})^2}{2\sigma_e^2} + \sigma_e^{-2} M_i^{(j)} (Y_i - A_i\beta_a - \sum_{s \neq j} M_i^{(s)}\beta_{ms} - \boldsymbol{C_i}^T\boldsymbol{\beta_c})\beta_{mj}\}
$$

$$
p(\beta_{mj}|r_{mj} = z, .) = N(\mu_{mjz}, s_{mjz}^2), z = 0, 1
$$

$$
\mu_{mjz} = \frac{\sum_{i=1}^{n} M_i^{(j)} (Y_i - A_i\beta_a - \sum_{s \neq j} M_i^{(s)}\beta_{ms} - \boldsymbol{C_i}^T\boldsymbol{\beta_c})}{\sigma_e^2/\sigma_{mz}^2 + \sum_{i=1}^{n} (M_i^{(j)})^2}, s_{mjz}^2 = \frac{1}{1/\sigma_{mz}^2 + \sum_{i=1}^{n} (M_i^{(j)})^2/\sigma_e^2}
$$

$$p(r_{mj}=z|.) \propto \exp(\mu_{mjz}^2/2s_{mjz}^2 + \log(s_{mjz}) - \log(\sigma_{mz}) + \log(p(r_{mj}=z)))$$

Sampling β_a

$$\log p(\beta_a|.) = -\frac{\beta_a^2}{2\sigma_a^2} - \sum_{i=1}^{n}\{\frac{(A_i\beta_a)^2}{2\sigma_e^2} + \sigma_e^{-2}A_i(Y_i - \boldsymbol{M_i^T\beta_m} - \boldsymbol{C_i^T\beta_c})\beta_a\}$$

$$p(\beta_a|.) = N(\mu_a, s_a^2)$$

$$\mu_a = \frac{\sum_{i=1}^{n} A_i(Y_i - \boldsymbol{M_i^T\beta_m} - \boldsymbol{C_i^T\beta_c})}{\sigma_e^2/\sigma_a^2 + \sum_{i=1}^{n} A_i^2}, s_a^2 = \frac{1}{1/\sigma_a^2 + \sum_{i=1}^{n} A_i^2/\sigma_e^2}$$

Sampling α_{aj} *and* r_{aj}

$$\log p(\alpha_{aj}|r_{aj}=1,.) = -\frac{\alpha_{aj}^2}{2\sigma_{ma1}^2} - \sum_{i=1}^{n}\{\frac{(A_i\alpha_{aj})^2}{2\sigma_g^2} + \sigma_g^{-2}A_i(M_i^{(j)} - (\boldsymbol{\alpha_c C_i})_j)\alpha_{aj}\}$$

$$p(\alpha_{aj}|r_{aj}=z,.) = N(\frac{\sum_{i=1}^{n} A_i(M_i^{(j)} - (\boldsymbol{\alpha_c C_i})_j)}{\sigma_g^2/\sigma_{maz}^2 + \sum_{i=1}^{n} A_i^2}, \frac{1}{1/\sigma_{maz}^2 + \sum_{i=1}^{n} A_i^2/\sigma_g^2})$$

$$p(r_{aj}=z|.) \propto \exp(\mu_{\alpha jz}^2/2s_{\alpha jz}^2 + \log(s_{\alpha jz}) - \log(\sigma_{maz}) + \log(p(r_{aj}=z))), z=0,1$$

Sampling $\sigma_{m1}^2, \sigma_{m0}^2$

$$\log p(\sigma_{m1}^2|.) = -(\frac{\sum_{j=1}^{q} r_{mj}}{2} + k_m + 1)\log(\sigma_{m1}^2) - (\frac{\sum_{j=1}^{q} r_{mj}\beta_{mj}^2}{2} + l_m)\sigma_{m1}^{-2}$$

$$\log p(\sigma_{m0}^2|.) = -(\frac{\sum_{j=1}^{q}(1-r_{mj})}{2} + k_m + 1)\log(\sigma_{m0}^2) - (\frac{\sum_{j=1}^{q}(1-r_{mj})\beta_{mj}^2}{2} + l_m)\sigma_{m0}^{-2}$$

$$p(\sigma_{m1}^2|.) \sim \text{inverse-gamma}(\frac{\sum_{j=1}^{q} r_{mj}}{2} + k_m, \frac{\sum_{j=1}^{q} r_{mj}\beta_{mj}^2}{2} + l_m)$$

$$p(\sigma_{m0}^2|.) \sim \text{inverse-gamma}(\frac{\sum_{j=1}^{q}(1-r_{mj})}{2} + k_m, \frac{\sum_{j=1}^{q}(1-r_{mj})\beta_{mj}^2}{2} + l_m)$$

Sampling σ_a^2

$$\log p(\sigma_a^2|.) = -(\frac{1}{2} + k_a + 1)\log(\sigma_a^2) - (\frac{\beta_a^2}{2} + l_a)\sigma_a^{-2}$$

$$p(\sigma_a^2|.) \sim \text{inverse-gamma}(\frac{1}{2} + k_a, \frac{\beta_a^2}{2} + l_a)$$

Sampling $\sigma_{ma1}^2, \sigma_{ma0}^2$

$$\log p(\sigma_{ma1}^2|.) = -(\frac{\sum_j r_{aj}}{2} + k_{ma} + 1)\log(\sigma_{ma1}^2) - (\frac{\sum_j r_{aj}\alpha_{aj}^2}{2} + l_{ma})\sigma_{ma1}^{-2}$$

$$\log p(\sigma_{ma0}^2|.) = -(\frac{\sum_j (1 - r_{aj})}{2} + k_{ma} + 1)\log(\sigma_{ma0}^2) - (\frac{\sum_j (1 - r_{aj})\alpha_{aj}^2}{2} + l_{ma})\sigma_{ma0}^{-2}$$

$$p(\sigma_{ma1}^2|.) \sim \text{inverse-gamma}(\frac{\sum_j r_{aj}}{2} + k_{ma}, \frac{\sum_j r_{aj}\alpha_{aj}^2}{2} + l_{ma})$$

$$p(\sigma_{ma0}^2|.) \sim \text{inverse-gamma}(\frac{\sum_j (1 - r_{aj})}{2} + k_{ma}, \frac{\sum_j (1 - r_{aj})\alpha_{aj}^2}{2} + l_{ma})$$

Sampling σ_e^2

$$\log p(\sigma_e^2|.) = -(\frac{n}{2} + k_e + 1)\log(\sigma_e^2) - (\frac{\sum_{i=1}^n (Y_i - \boldsymbol{M_i^T}\boldsymbol{\beta_m} - A_i\beta_a - \boldsymbol{C_i^T}\boldsymbol{\beta_c})^2}{2} + l_e)\sigma_e^{-2}$$

$$p(\sigma_e^2|.) \sim \text{inverse-gamma}(\frac{n}{2} + k_e, \frac{\sum_{i=1}^n (Y_i - \boldsymbol{M_i^T}\boldsymbol{\beta_m} - A_i\beta_a - \boldsymbol{C_i^T}\boldsymbol{\beta_c})^2}{2} + l_e)$$

Sampling σ_g^2

$$\log p(\sigma_g^2|.) = -(\frac{qn}{2} + k_g + 1)\log(\sigma_g^2) - (\frac{\sum_{i=1}^n \sum_q (M_i^{(q)} - A_i\alpha_{aq} - (\boldsymbol{\alpha_c}\boldsymbol{C_i})_q)^2}{2} + l_g)\sigma_g^{-2}$$

$$p(\sigma_g^2|.) \sim \text{inverse-gamma}(\frac{qn}{2} + k_g, (\frac{\sum_{i=1}^n \sum_q (M_i^{(q)} - A_i\alpha_{aq} - (\boldsymbol{\alpha_c}\boldsymbol{C_i})_q)^2}{2} + l_g))$$

Sampling β_{cw}

$$\log p(\beta_{cw}|.) = -\sum_{i=1}^{n}\{\frac{(C_{iw}\beta_{cw})^2}{2\sigma_e^2} + \sigma_e^{-2}C_{iw}(Y_i - \boldsymbol{M_i^T\beta_m} - A_i\beta_a - \sum_{s\neq w} C_{iw}\beta_{cw})\beta_{cw}\}$$

$$p(\beta_{cw}|.) = N(\frac{\sum_{i=1}^{n} C_{iw}(Y_i - A_i\beta_a - \boldsymbol{M_i^T\beta_m} - \sum_{s\neq w} C_{iw}\beta_{cw})}{\sum_{i=1}^{n} C_{iw}^2}, \frac{\sigma_e^2}{\sum_{i=1}^{n} C_{iw}^2})$$

Sampling $(\boldsymbol{\alpha_{cw}})_j$

$$\log p((\boldsymbol{\alpha_{cw}})_j|.) = -\sum_{i=1}^{n}\{\frac{(C_{iw}(\boldsymbol{\alpha_{cw}})_j)^2}{2\sigma_g^2} + \sigma_g^{-2}C_{iw}(M_i^{(j)} - A_i\alpha_{aj} - \sum_{s\neq w} C_{is}(\boldsymbol{\alpha_{cs}})_j)(\boldsymbol{\alpha_{cw}})_j\}$$

$$p((\boldsymbol{\alpha_{cw}})_j|.) = N(\frac{\sum_{i=1}^{n} C_{iw}(M_i^{(j)} - A_i\alpha_{aj} - \sum_{s\neq w} C_{is}(\boldsymbol{\alpha_{cs}})_j)}{\sum_{i=1}^{n} C_{iw}^2}, \frac{\sigma_g^2}{\sum_{i=1}^{n} C_{iw}^2})$$

Sampling π_m, π_a

For π_m, π_a, their conditional distributions don't appear to be of any known form, so we use a random-walk standard Metropolis-Hastings algorithm to draw posterior samples of them.

As for the proposal distribution, we update the parameters by adding a random variable from U(-0.07, 0.07) to the current value. New values that lie outside the boundary [0,1] are reflected back.

A.3 Convergence Diagnosis

We used the potential scale reduction factor (PSRF) (*Gelman and Rubin*, 1992) to quantify the mixing property of the proposed MCMC algorithm. With multiple MCMC chains, PSRF for a parameter is essentially the ratio between the overall-chain variance and the average within-chain variance. A PSRF value in the range of (0.9, 1.2) suggests that the MCMC algorithm has good mixing property and the posterior samples converge well. As an example, in Figure A.1, we present the PSRFs for the PIPs of 60 top significant mediators identified from univariate analysis in the baseline simulation setting with the number of mediators $p = 2,000$. We find that all the PSRFs from our MCMC algorithm fall within (0.9, 1.2),

which indicates the good mixing property of our algorithm.

A.4 Power Comparison with Spike-and-slab Priors and Horseshoe Priors

Both horseshoe priors and spike-and-slab priors (*Mitchell and Beauchamp*, 1988; *Carvalho et al.*, 2010) are widely used methods for Bayesian shrinkage, and it is natural to apply them to the two regression models in high-dimensional mediation analysis. The horseshoe prior can be represented as a scale mixture of normals, with the mixing distribution being a standard half-Cauchy distribution. The horseshoe prior is not a discrete mixture prior and therefore it does not directly categorize mediators into one actively mediating group and three inactive (null) groups. To achieve categorization and selection of mediators, one can use the shrinkage factors (*Carvalho et al.*, 2010) for the coefficients and develop a thresholding rule on the continuous values to determine inclusion or not. For the spike-and-slab prior, we can directly use the posterior inclusion probability to perform mediator selection.

We implemented both horseshoe priors and spike-and-slab priors for high-dimensional mediation analysis and compared them with our method in simulations. We first simulated $(\boldsymbol{\beta_m})_j$ $(j = 1, ..., p)$ from a mixture of two normals: $\pi_m N(0, 1) + (1 - \pi_m)N(0, 0.001)$, and $(\boldsymbol{\alpha_a})_j$ $(j = 1, ..., p)$ from $\pi_a N(0, 1) + (1 - \pi_a)N(0, 0.001)$. The other configurations are same as the baseline setting for $p = 2,000$. We find that our Bayesian method with normal-normal priors outperforms the other two methods. For example, when $PVE_{IE} = 0.8$, at 0.01 FPR, our method achieves a power of 0.528, while the methods with point-normal priors and horseshoe priors have a power of 0.484 and 0.467, respectively. The full results are shown in Table A.1.

In addition, we also performed simulations in which $(\boldsymbol{\beta_m})_j$ $(j = 1, ..., p)$ from point-normal priors: $\pi_m N(0, 1) + (1 - \pi_m)\delta_0$, and $(\boldsymbol{\alpha_a})_j$ $(j = 1, ..., p)$ from $\pi_a N(0, 1) + (1 - \pi_a)\delta_0$, where δ_0 is a point mass at zero. The other configurations are same as the baseline setting for

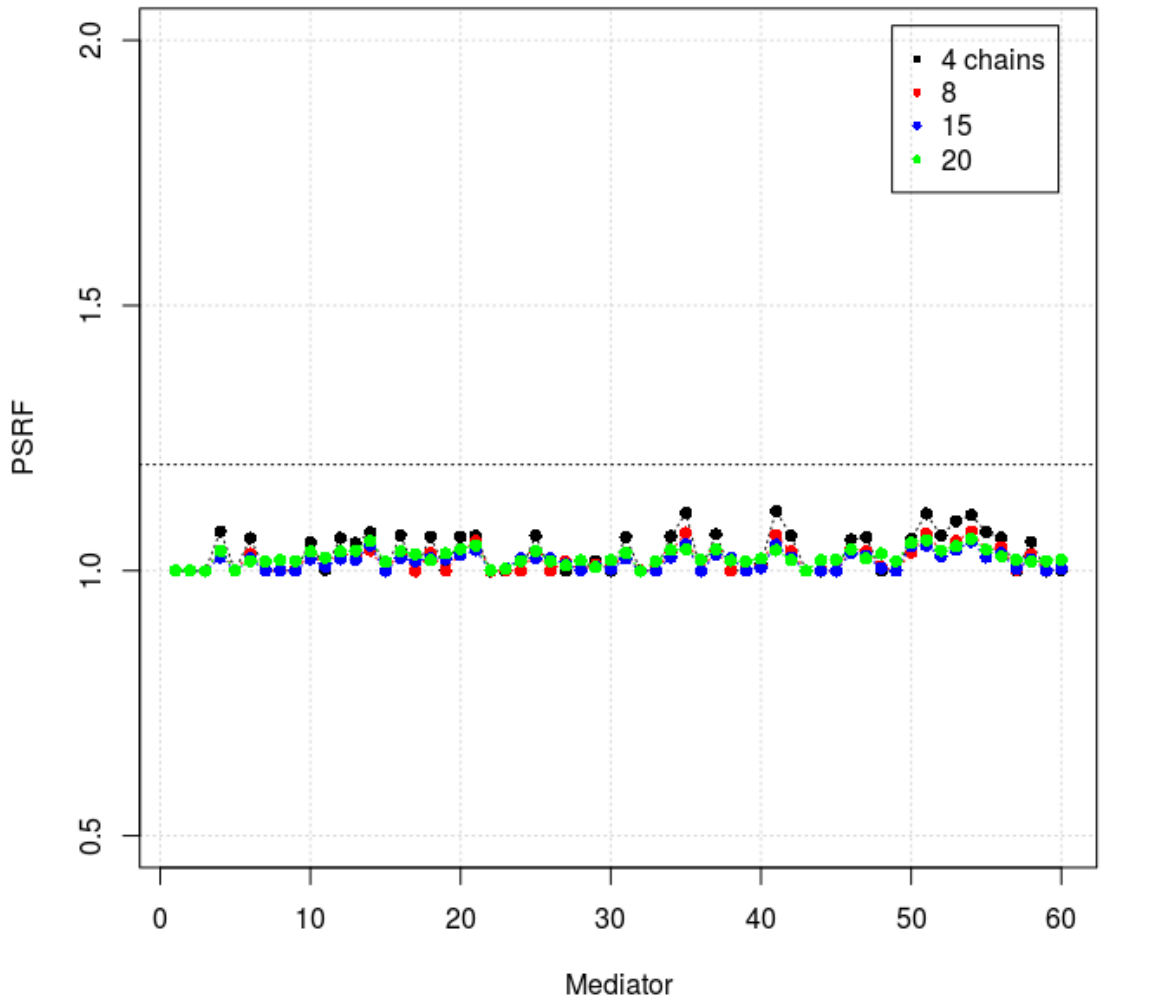

Figure A.1: Potential scale reduction factors (PSRF) of the Bayesian posterior inclusion probabilities of 60 top marginally significant mediators with 3, 8, 15, and 20 MCMC chains, where PSRF within (0.9, 1.2) suggests good mixing property.

p	Setting	Normal-normal priors	Point-normal priors	Horseshoe priors
2,000	$PVE_A = 0.3$	0.509	0.461	0.437
	$PVE_A = 0.5$	0.474	0.424	0.461
	$PVE_A = 0.8$	0.512	0.413	0.479
	$PVE_{IE} = 0.2$	0.473	0.415	0.453
	$PVE_{IE} = 0.4$	0.474	0.424	0.461
	$PVE_{IE} = 0.8$	0.528	0.484	0.467
	$\pi_a = 0.03$	0.474	0.424	0.461
	$\pi_a = 0.1$	0.146	0.131	0.092
	$\pi_a = 0.25$	0.072	0.062	0.042
	$\pi_m = 0.02$	0.474	0.424	0.461
	$\pi_m = 0.1$	0.471	0.420	0.454
	$\pi_m = 0.25$	0.462	0.401	0.440

Table A.1: Power comparison among our normal-normal priors, the point-normal priors and horseshoe priors when $p = 2,000, n = 1,000$, and the effect sizes are sampled from a mixture of two normals. In each setting, we change one parameter at a time from the baseline setting. The average TPR at FPR = 0.01 is calculated across 200 replicates.

$p = 2,000$. Now the effect size distribution is not a mixture of two normals, and favors the spike-and-slab priors. The results (in Table A.2) show that all the three methods have similar performance in most scenarios, and our method remains decent power when the effects are not polygenic.

Therefore, our Bayesian mediation method with mixture of normals prior performs well in identifying active mediators in a wide range of scenarios, and is relatively robust to the effect size distribution.

A.5 Posterior Distribution of the Global Mediation Effects τ

We examine the posterior distribution of τ, which is bounded at zero and not symmetric. The distribution also depends on the composition of the four groups. In below we show a distribution graph (Figure A.2) based on the posterior samples of τ in four different scenarios

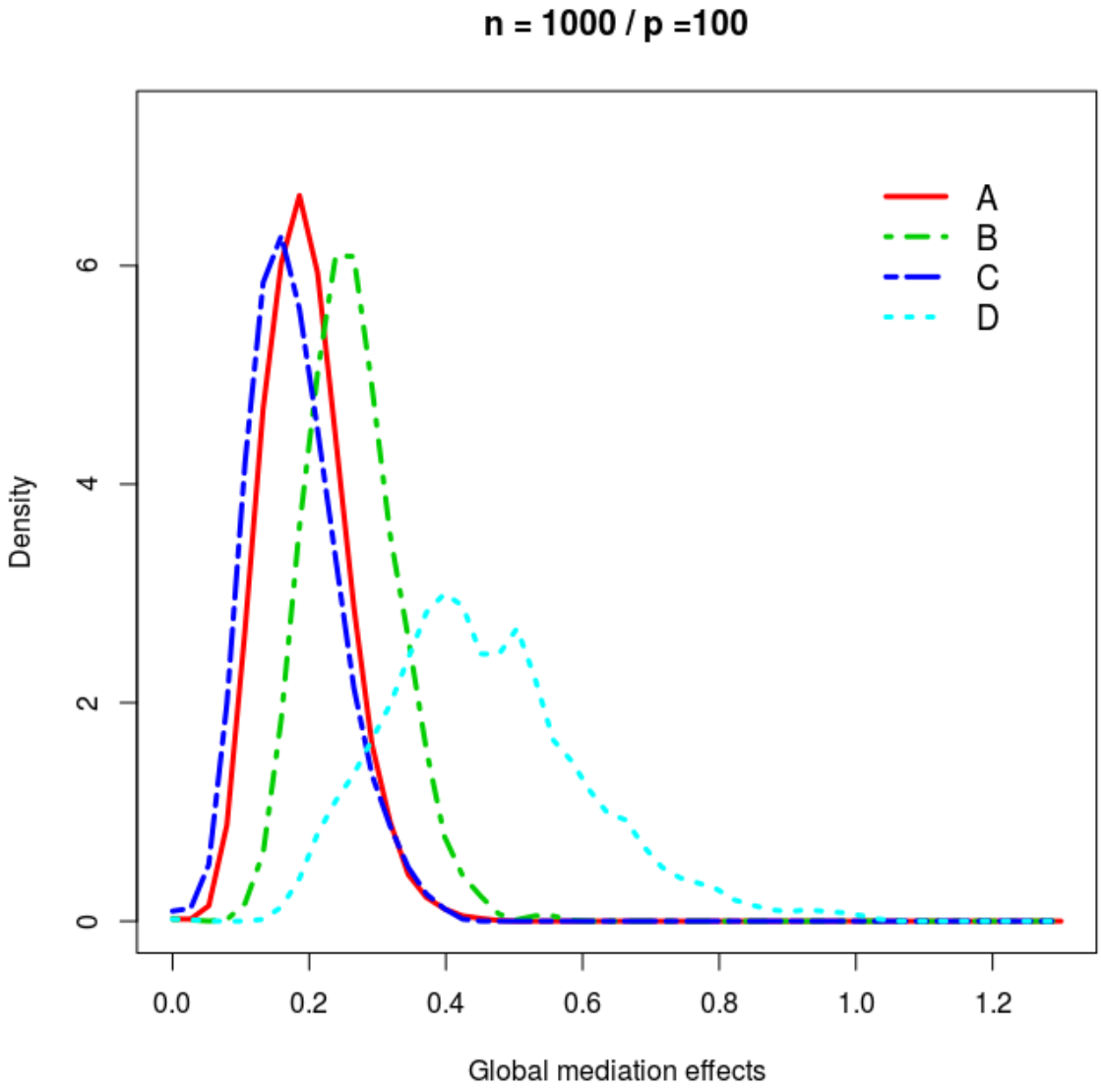

Figure A.2: The distribution from the posterior samples of τ in four different scenarios with $n = 1000, p = 100$. We denote π_{g1}, π_{g2}, π_{g3} and π_{g4} to represent the proportion of mediators in Group 1, Group 2, Group 3 and Group 4 as defined in Table 1 in the main Chapter II. The four settings are: A: $\pi_{g1} = 0.1, \pi_{g2} = 0.2, \pi_{g3} = 0.1, \pi_{g4} = 0.6$; B: $\pi_{g1} = 0.1, \pi_{g2} = 0.1, \pi_{g3} = 0.2, \pi_{g4} = 0.6$; C: $\pi_{g1} = 0.1, \pi_{g2} = 0.1, \pi_{g3} = 0.1, \pi_{g4} = 0.7$; D: $\pi_{g1} = 0.1, \pi_{g2} = 0, \pi_{g3} = 0, \pi_{g4} = 0.9$;

p	Setting	Normal-normal priors	Point-normal priors	Horseshoe priors
2,000	$PVE_A = 0.3$	0.525	0.526	0.465
	$PVE_A = 0.5$	0.483	0.503	0.490
	$PVE_A = 0.8$	0.470	0.513	0.493
	$PVE_{IE} = 0.2$	0.456	0.488	0.476
	$PVE_{IE} = 0.4$	0.483	0.503	0.490
	$PVE_{IE} = 0.8$	0.510	0.543	0.491
	$\pi_a = 0.03$	0.483	0.503	0.490
	$\pi_a = 0.1$	0.135	0.145	0.106
	$\pi_a = 0.25$	0.047	0.094	0.052
	$\pi_m = 0.02$	0.483	0.503	0.490
	$\pi_m = 0.1$	0.468	0.488	0.486
	$\pi_m = 0.25$	0.450	0.465	0.470

Table A.2: Power comparison among our normal-normal priors, the point-normal priors and horseshoe priors when $p = 2,000, n = 1,000$, and the effect sizes are sampled from point-normal priors. In each setting, we change one parameter at a time from the baseline setting. The average TPR at FPR = 0.01 is calculated across 200 replicates.

with $n = 1000, p = 100$ as in Table 2 in the main manuscript.

A.6 Detailed Description of MESA Data

MESA is a population-based longitudinal study designed to identify risk factors for the progression of subclinical cardiovascular disease (CVD) (*Bild et al.*, 2002). A total of 6,814 non-Hispanic white, African-American, Hispanic, and Chinese-American women and men aged 45−84 without clinically apparent CVD were recruited between July 2000 and August 2002 from the following 6 regions in the US: Forsyth County, NC; Northern Manhattan and the Bronx, NY; Baltimore City and Baltimore County, MD; St. Paul, MN; Chicago, IL; and Los Angeles County, CA. Each field center recruited from locally available sources, which included lists of residents, lists of dwellings, and telephone exchanges. At Exam 1, respondents reported the highest level of education they completed. We created a dichotomous measure of respondent's educational attainment (less than college, low adult SES = 1; college degree or more = 0). The descriptive statistics for the exposure and outcome can be found in Table C.2.

In the MESA data, between April 2010 and February 2012 (corresponding to MESA Exam 5), DNAm were assessed on a random subsample of 1,264 non-Hispanic white, African-American, and Hispanic MESA participants aged 55−94 from the Baltimore, Forsyth County, New York, and St. Paul field centers. After excluding respondents with missing data on one or more variables, we had phenotype and DNAm data from purified monocytes on a total of 1,231 individuals and we focused on this set of individuals for analysis. The detailed description of DNAm data collection, quantitation and data processing procedures can be found in Liu et al (*Liu et al.*, 2013). Briefly, the Illumina HumanMethylation450 BeadChip was used to measure DNAm, and bead-level data were summarized in GenomeStudio. Quantile normalization was performed using the *lumi* package with default settings (*Du et al.*, 2008). Quality control (QC) measures included checks for sex and race/ethnicity mismatches and outlier identification by multidimensional scaling plots. Further probe filtering criteria included: "detected" DNAm levels in <90% of MESA samples (detection p-value cut-off = 0.05), existence of a SNP within 10 base pairs of the target CpG site, overlap with a non-unique region, and suggestions by DMRcate (*Chen et al.*, 2013b) (mostly cross-reactive probes). Those procedures leave us 403,713 autosomal probes for analysis.

For computational reasons, we first selected a subset of CpG sites to be used in the final multivariate mediation analysis model. In particular, for each single CpG site in turn, we fit the following linear mixed model to test the marginal association between the CpG site and the exposure variable:

$$M_i = A_i\psi_a + \boldsymbol{C_{1i}^T\psi_c} + \boldsymbol{Z_i^T\psi_u} + \epsilon_i, i = 1, ..., n \tag{A.1}$$

where A_i represents adult SES value for the i'th individual and ψ_a is its coefficient; $\boldsymbol{C_{1i}}$ is a vector of covariates that include age, gender, race/ethnicity and enrichment scores for each of 4 major blood cell types (neutrophils, B cells, T cells and natural killer cells) to account for potential contamination by non-monocyte cell types; $\boldsymbol{Z_i^T\psi_u}$ represent methylation chip

and position random effects and are used to control for possible batch effects. The error term $\epsilon_i \sim MVN(0, \sigma^2 I_n)$ and is independent of the random effects. We obtained p-values for testing the null hypothesis $\psi_a = 0$ from the above model. We further applied the R/Bioconductor package BACON (*van Iterson et al.*, 2017) to these p-values to further adjust for possible inflation using an empirical null distribution. Based on these marginal p-values, we selected top 2,000 CpG sites with the smallest p-values for our Bayesian multivariate analysis.

We implemented our proposed methods as well as methods with different prior specifications and HDMM on the MESA data. The current HDMM cannot handle covariates, so we apply it to the residuals after regressing the Y and $\boldsymbol{M}$ on the covariates. There may exist certain numerical stability issue with the HDMM on the MESA data, and the resulting weights of the first direction of mediation do not suggest obvious signal or pattern. The estimated first direction of mediation across the selected 2,000 sites is presented in Figure A.3.

We also listed the top 2 sites and nearby genes identified by the four methods in the Table A.3,

Method	Top 2 sites	Nearby genes
Our method	cg19582614	CCND2
	cg04514392	CCDC54
Spike-and-Slab Priors	cg19582614	CCND2
	cg26610247	RP11-10J21.3
Horseshoe Priors	cg15531249	C16orf74
	cg15149205	TRIB1
HDMM	cg13488078	CLU
	cg12880602	MAP3K7

Table A.3: The top 2 sites and their nearby genes identified by our proposed method as well as methods with different prior specifications and HDMM for mediation analysis on Adult SES $\rightarrow$ DNAm $\rightarrow$ HbA1c.

We do not know the truth in the real data so it is hard evaluate the effectiveness of different methods here. In addition to the genes *CCDC54* and *CCND2*, which are associated with

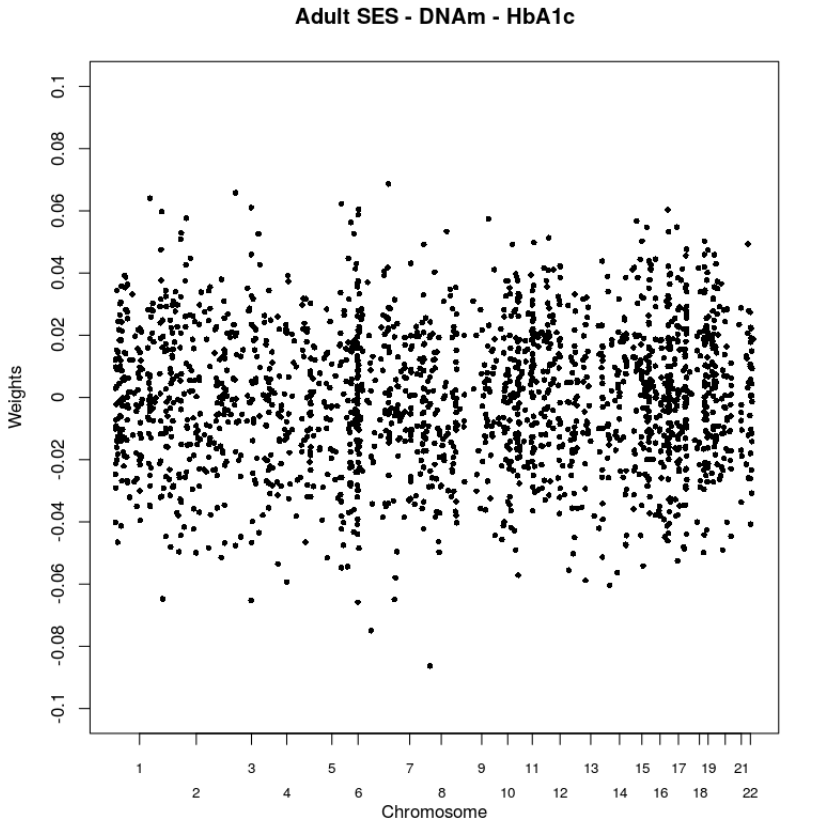

Figure A.3: Consider the trio: Adult SES → DNAm → HbA1c. The black dots are the weights for the first direction of mediation for the selected 2,000 sites across the genome from HDMM.

the outcome HbA1c as discussed in Chapter II, *CLU* is associated with diabetes, probably through an increase in insulin resistance (*Daimon et al.*, 2011). There is a lack of biological evidence to support a mediating role of the other genes.

After fitting the Bayesian mediation models, we then empirically check whether models 11 and 12 in the main manuscript are reasonable. We perform the posterior predictive checks on the outcome model, and create the following graphical displays comparing the observed outcome to the replications drawn from the posterior predictive distribution.

In Figure A.4, we compare the distribution of the observed outcomes HbA1c (y) and the kernel density estimates of replications of the outcome, y_{rep} from the posterior predictive distribution. Those plots makes it easy to see that the distributions of the simulated outcomes do not deviate much from the true distribution of HbA1c in the data and our model is relatively reasonable. The Bayesian predictive p-values (*Neelon et al.*, 2010) are 0.5 and 0.45 for the sample mean and variance, respectively, also suggesting adequate fit of the outcome model.

	Full Sample (n, %)	**Low Adult SES %**	**HbA1c Mean (SD)**
Full sample	1231 (100)	67	5.99 (0.92)
Age			
55−65 years	466 (38)	64	5.92 (0.97)
66−75 years	398 (32)	68	6.08 (0.98)
76−85 years	301 (24)	67	5.98 (0.80)
86−95 years	66 (5)	74	5.95 (0.72)
Race			
Non-Hispanic white	582 (47)	51	5.76 (0.65)
African-American	263 (22)	72	6.23 (1.03)
Hispanic	386 (31)	86	6.16 (1.11)
Gender			
Female	633 (51)	73	5.99 (0.88)
Male	598 (49)	60	5.99 (0.97)

Table A.4: Descriptive statistics for adult SES measures and HbA1c. n: number of subjects. %: proportion in the corresponding category. SD: standard deviation.

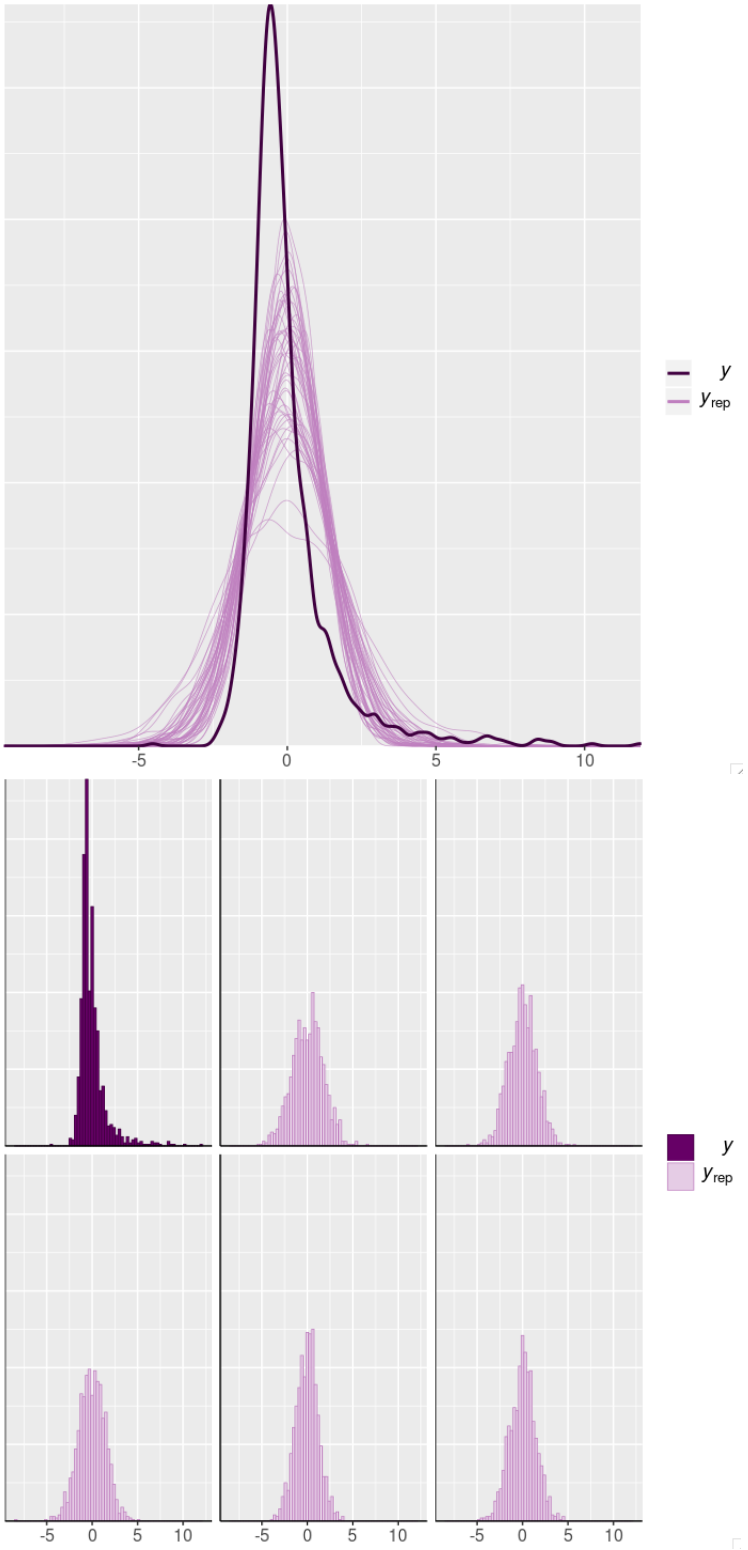

Figure A.4: The top panel shows the distribution of the observed outcomes HbA1c (y, the dark line) and each of the 100 lighter lines is the kernel density estimate of one of the replications of y, y_{rep} from the posterior predictive distribution. The bottom panel displays the separate histograms of y and five of the y_{rep} datasets.

APPENDIX B

Supplement for Chapter III

B.1 Identifiability Assumptions for Causal Mediation Analysis

We use the same counterfactual notation as in the first chapter. To connect potential variables to observed data, we make the Stable Unit Treatment Value Assumption (SUTVA) (*Rubin*, 1980, 1986). Specifically, the SUTVA assumes there is no interference between subjects and the consistency assumption, which states that the observed variables are the same as the potential variables corresponding to the actually observed treatment level, i.e., $\boldsymbol{M_i} = \sum_a \boldsymbol{M_i}(a)I(A_i = a)$, and $Y_i = \sum_a \sum_{\boldsymbol{m}} Y_i(a, \boldsymbol{m})I(A_i = a, \boldsymbol{M_i} = \boldsymbol{m})$, where $I(\cdot)$ is the indicator function.

Causal effects are formally defined in terms of potential variables which are not necessarily observed, but the identification of causal effects must be based on observed data. Therefore further assumptions regarding the confounders are required for the identification of causal effects in mediation analysis (*VanderWeele and Vansteelandt*, 2014). We will use $A \perp\!\!\!\perp B|C$ to denote that A is independent of B conditional on C. To estimate the average NDE and NIE from observed data, the following assumptions are needed: (1) $Y_i(a, \boldsymbol{m}) \perp\!\!\!\perp A_i|\boldsymbol{C_i}$, no unmeasured confounding for exposure-outcome relationship; (2) $Y_i(a, \boldsymbol{m}) \perp\!\!\!\perp \boldsymbol{M}_i|\{\boldsymbol{C_i}, A_i\}$, no

unmeasured confounding for any of mediator-outcome relationship after controlling for the exposure; (3) $\boldsymbol{M}_i(a) \perp\!\!\!\perp A_i | \boldsymbol{C_i}$, no unmeasured confounding for the exposure effect on all the mediators; (4) $Y_i(a, \boldsymbol{m}) \perp\!\!\!\perp \boldsymbol{M}_i(a^\star) | \boldsymbol{C_i}$, no downstream effect of the exposure that confounds any mediator-outcome relationship. The four assumptions are required to hold with respect to the whole set of mediators. Finally, as in all mediation analysis, the temporal ordering assumption also needs to be satisfied, i.e., the exposure precedes the mediators, and the mediators precede the outcome.

B.2 Posterior Sampling Algorithm Details for Gaussian Mixture Model (GMM)

Let $\boldsymbol{\Theta}_{GMM} = (\boldsymbol{\beta_m}, \boldsymbol{\alpha_a}, \boldsymbol{V_k}, \beta_a, \boldsymbol{\beta_c}, \boldsymbol{\alpha_c}, \{\gamma_j\}_{j=1}^p, \pi_k, k = 1, 2, 3, 4, \sigma_e^2, \boldsymbol{\Sigma}, \sigma_a^2)$ denote all the unknown parameters in our Gaussian mixture model. The joint likelihood of $\{Y_i, \boldsymbol{M_i}\}_{i=1}^n$ given $\boldsymbol{\Theta}_{GMM}$ is,

$$
\begin{aligned}
& \log P(\{Y_i, \boldsymbol{M_i}\}_{i=1}^n | \boldsymbol{\Theta}_{GMM}, \{A_i, \boldsymbol{C_i}\}_{i=1}^n) \\
= & \sum_{i=1}^n \log P(Y_i, \boldsymbol{M_i} | \boldsymbol{\Theta}_{GMM}, A_i, \boldsymbol{C_i}) \\
= & \sum_{i=1}^n \log P(Y_i | \boldsymbol{M_i}, \boldsymbol{\beta_m}, \sigma_e^2, \beta_a, \boldsymbol{\beta_c}, A_i, \boldsymbol{C_i}) + \log P(\boldsymbol{M_i} | \boldsymbol{\alpha_a}, \boldsymbol{\alpha_c}, \boldsymbol{\Sigma}, A_i, \boldsymbol{C_i}) \\
= & \sum_{i=1}^n -\frac{1}{2}\log\sigma_e^2 - \frac{1}{2\sigma_e^2}(Y_i - \boldsymbol{M_i}^T\boldsymbol{\beta_m} - A_i\beta_a - \boldsymbol{C_i}^T\boldsymbol{\beta_c})^2 \\
& -\frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{M_i} - A_i\boldsymbol{\alpha_a} - \boldsymbol{\alpha_c}\boldsymbol{C_i})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{M_i} - A_i\boldsymbol{\alpha_a} - \boldsymbol{\alpha_c}\boldsymbol{C_i})
\end{aligned}
$$

The joint log posterior distribution is,

$$
\begin{aligned}
& \log P(\boldsymbol{\Theta}_{GMM} | \{Y_i, \boldsymbol{M_i}, A_i, \boldsymbol{C_i}\}_{i=1}^n) \\
\propto & \sum_{i=1}^n \log P(Y_i | \boldsymbol{M_i}, \boldsymbol{\beta_m}, \sigma_e^2, \beta_a, \boldsymbol{\beta_c}, A_i, \boldsymbol{C_i}) + \log P(\boldsymbol{M_i} | \boldsymbol{\alpha_a}, \boldsymbol{\alpha_c}, \boldsymbol{\Sigma}, A_i, \boldsymbol{C_i}) \\
& + \log P(\boldsymbol{\Theta}_{GMM})
\end{aligned}
$$

$$
\begin{aligned}
= & \sum_{i=1}^{n} -\frac{1}{2}\log\sigma_e^2 - \frac{1}{2\sigma_e^2}(Y_i - \boldsymbol{M_i^T\beta_m} - A_i\beta_a - \boldsymbol{C_i^T\beta_c})^2 \\
& -\frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{M_i} - A_i\boldsymbol{\alpha_a} - \boldsymbol{\alpha_c C_i})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{M_i} - A_i\boldsymbol{\alpha_a} - \boldsymbol{\alpha_c C_i}) \\
& + \sum_{j=1}^{p}\sum_{k=1}^{4} I(\gamma_j = k)(-\frac{d}{2}\log 2\pi - \frac{1}{2}\log|\boldsymbol{V_k}| - \frac{1}{2}\begin{bmatrix}(\boldsymbol{\beta_m})_j \\ (\boldsymbol{\alpha_a})_j\end{bmatrix}^T \boldsymbol{V_k^{-1}} \begin{bmatrix}(\boldsymbol{\beta_m})_j \\ (\boldsymbol{\alpha_a})_j\end{bmatrix}) \\
& -\frac{q}{2}\log 2\pi\sigma_c^2 - \frac{\boldsymbol{\beta_c^T\beta_c}}{2\sigma_c^2} - \frac{pq}{2}\log 2\pi\sigma_c^2 - \sum_{j=1}^{p}\frac{\boldsymbol{\alpha_{cj}^T\alpha_{cj}}}{2\sigma_c^2} \\
& + \sum_{j=1}^{p}\sum_{k=1}^{4} I(\gamma_j = k)\log(\pi_k) \\
& + \sum_{k=1}^{4} a_k\log(\pi_k) + \sum_{k=1}^{4}(-\frac{\nu + d + 1}{2}\log|\boldsymbol{V_k}| + \frac{1}{2}tr(\boldsymbol{\Psi_0 V_k^{-1}})))
\end{aligned}
$$

Sampling $\begin{bmatrix}(\boldsymbol{\beta_m})_j \\ (\boldsymbol{\alpha_a})_j\end{bmatrix}$ *and* γ_j

$$
\log p(\begin{bmatrix}(\boldsymbol{\beta_m})_j \\ (\boldsymbol{\alpha_a})_j\end{bmatrix} | \gamma_j = k, \cdot) \propto -\frac{1}{2}\begin{bmatrix}(\boldsymbol{\beta_m})_j \\ (\boldsymbol{\alpha_a})_j\end{bmatrix}^T (\boldsymbol{W_j} + \boldsymbol{V_k^{-1}})\begin{bmatrix}(\boldsymbol{\beta_m})_j \\ (\boldsymbol{\alpha_a})_j\end{bmatrix} + \boldsymbol{w_j^T}\begin{bmatrix}(\boldsymbol{\beta_m})_j \\ (\boldsymbol{\alpha_a})_j\end{bmatrix}
$$

where $\boldsymbol{W_j} = \begin{bmatrix}\sum_{i=1}^{n}(\sigma_e^2)^{-1}M_{ij}^2 & 0 \\ 0 & \sum_{i=1}^{n}\boldsymbol{\Sigma}^{-1}A_i^2\end{bmatrix}$ ($\boldsymbol{\Sigma}$ need to be diagonal, and can be replaced as σ_g^2), and $\boldsymbol{w_j} = (\sum_{i=1}^{n}(\sigma_e^2)^{-1}(Y_i - A_i\beta_a - \sum_{j'\neq j}M_{ij'}(\boldsymbol{\beta_m})_{j'})M_{ij}, \sum_{i=1}^{n}\boldsymbol{\Sigma}^{-1}M_{ij}A_i)^T$

$$
p(\begin{bmatrix}(\boldsymbol{\beta_m})_j \\ (\boldsymbol{\alpha_a})_j\end{bmatrix} | \gamma_j = k, \cdot) \sim MVN_2((\boldsymbol{W_j} + \boldsymbol{V_k^{-1}})^{-1}\boldsymbol{w_j}, (\boldsymbol{W_j} + \boldsymbol{V_k^{-1}})^{-1})
$$

$$
\log p(\gamma_j = k|.) \propto -\frac{1}{2}\log|\boldsymbol{W_j V_k} + \boldsymbol{I_2}| + \frac{1}{2}\boldsymbol{w_j^T}(\boldsymbol{W_j} + \boldsymbol{V_k^{-1}})^{-1}\boldsymbol{w_j} + \log(\pi_k)
$$

Sampling π_k

$$
\{\pi_1, \pi_2, \pi_3, \pi_4\}
$$

$$\propto \quad \text{Dirichlet}(a_1 + \sum_{j=1}^{p} I(\gamma_j = 1), a_2 + \sum_{j=1}^{p} I(\gamma_j = 2), a_3 + \sum_{j=1}^{p} I(\gamma_j = 3), a_4 + \sum_{j=1}^{p} I(\gamma_j = k))$$

Sampling $\boldsymbol{V_k}$

$$\begin{aligned} \log p(\boldsymbol{V_k}|.) \propto \quad & -\frac{1}{2}(\sum_{j=1}^{p} I(\gamma_j = k) + \nu + d + 1)\log|\boldsymbol{V_k}| - \frac{1}{2} tr(\boldsymbol{\Psi_0 V_k^{-1}}) \\ & + \sum_{j=1}^{p} I(\gamma_j = k)(-\frac{1}{2} \begin{bmatrix} (\boldsymbol{\beta m})_j \\ (\boldsymbol{\alpha a})_j \end{bmatrix}^T \boldsymbol{V_k^{-1}} \begin{bmatrix} (\boldsymbol{\beta m})_j \\ (\boldsymbol{\alpha a})_j \end{bmatrix}) \end{aligned}$$

$$p(\boldsymbol{V_k}|.) \sim \text{Inv-Wishart}(\boldsymbol{\Psi_0} + \sum_{j=1}^{p} I(\gamma_j = k) \begin{bmatrix} (\boldsymbol{\beta m})_j \\ (\boldsymbol{\alpha a})_j \end{bmatrix} \begin{bmatrix} (\boldsymbol{\beta m})_j \\ (\boldsymbol{\alpha a})_j \end{bmatrix}^T, \sum_{j=1}^{p} I(\gamma_j = k) + \nu)$$

Sampling β_a

$$\log p(\beta_a|.) \propto -\frac{\beta_a^2}{2\sigma_a^2} - \sum_{i=1}^{n} \{\frac{(A_i\beta_a)^2}{2\sigma_e^2} - \sigma_1^{-2} A_i(Y_i - \boldsymbol{M_i^T \beta_m} - \boldsymbol{C_i^T \beta_c})\beta_a\}$$

$$p(\beta_a|.) \sim N(\frac{\sum_{i=1}^{n} A_i(Y_i - \boldsymbol{M_i^T \beta_m} - \boldsymbol{C_i^T \beta_c})}{\sigma_e^2/\sigma_a^2 + \sum_{i=1}^{n} A_i^2}, \frac{1}{1/\sigma_a^2 + \sum_{i=1}^{n} A_i^2/\sigma_e^2})$$

Sampling σ_a^2

$$\log p(\sigma_a^2|.) \propto -(\frac{1}{2} + h_a + 1)\log(\sigma_a^2) - (\frac{\beta_a^2}{2} + l_a)\sigma_a^{-2}$$

$$p(\sigma_a^2|.) \sim \text{inverse-gamma}(\frac{1}{2} + h_a, \frac{\beta_a^2}{2} + l_a)$$

Sampling σ_e^2

$$\log p(\sigma_e^2|.) = -(\frac{n}{2} + h_1 + 1)\log(\sigma_e^2) - (\frac{\sum_{i=1}^{n}(Y_i - \boldsymbol{M_i^T \beta_m} - A_i\beta_a - \boldsymbol{C_i^T \beta_c})^2}{2} + l_1)\sigma_1^{-2}$$

$$p(\sigma_e^2|.) \sim \text{inverse-gamma}(\frac{n}{2} + h_1, \frac{\sum_{i=1}^{n}(Y_i - \boldsymbol{M_i^T \beta_m} - A_i\beta_a - \boldsymbol{C_i^T \beta_c})^2}{2} + l_1)$$

Sampling σ_g^2

$$\begin{aligned}\log p(\sigma_g^2|.) &= -(\frac{pn}{2}+h_2+1)\log(\sigma_g^2) \\ &\quad -(\frac{\sum_{i=1}^n(\boldsymbol{M_i}^T - A_i\boldsymbol{\alpha_a} - \boldsymbol{C_i}^T\boldsymbol{\alpha_c})(\boldsymbol{M_i}^T - A_i\boldsymbol{\alpha_a} - \boldsymbol{C_i}^T\boldsymbol{\alpha_c})^T}{2}+l_2)\sigma_2^{-2}\end{aligned}$$

$$p(\sigma_g^2|.) \sim \text{inverse-gamma}(\frac{pn}{2}+h_2, \frac{\sum_{i=1}^n(\boldsymbol{M_i}^T - A_i\boldsymbol{\alpha_a} - \boldsymbol{C_i}^T\boldsymbol{\alpha_c})(\boldsymbol{M_i}^T - A_i\boldsymbol{\alpha_a} - \boldsymbol{C_i}^T\boldsymbol{\alpha_c})^T}{2}+l_2)$$

Sampling β_{cw}

$$\log p(\beta_{cw}|.) = -\sum_{i=1}^n\{\frac{(C_{iw}\beta_{cw})^2}{2\sigma_e^2} + \sigma_e^{-2}C_{iw}(Y_i - \boldsymbol{M_i}^T\boldsymbol{\beta_m} - A_i\beta_a - \sum_{s\neq w}C_{iw}\beta_{cw})\beta_{cw}\}$$

$$p(\beta_{cw}|.) = N(\frac{\sum_{i=1}^n C_{iw}(Y_i - A_i\beta_a - \boldsymbol{M_i}^T\boldsymbol{\beta_m} - \sum_{s\neq w}C_{iw}\beta_{cw})}{\sum_{i=1}^n C_{iw}^2}, \frac{\sigma_e^2}{\sum_{i=1}^n C_{iw}^2})$$

Sampling $(\boldsymbol{\alpha_{cw}})_j$

$$\log p((\boldsymbol{\alpha_{cw}})_j|.) = -\sum_{i=1}^n\{\frac{(C_{iw}(\boldsymbol{\alpha_{cw}})_j)^2}{2\sigma_g^2} + \sigma_g^{-2}C_{iw}(M_i^{(j)} - A_i\alpha_{aj} - \sum_{s\neq w}C_{is}(\boldsymbol{\alpha_{cs}})_j)(\boldsymbol{\alpha_{cw}})_j\}$$

$$p((\boldsymbol{\alpha_{cw}})_j|.) = N(\frac{\sum_{i=1}^n C_{iw}(M_i^{(j)} - A_i\alpha_{aj} - \sum_{s\neq w}C_{is}(\boldsymbol{\alpha_{cs}})_j)}{\sum_{i=1}^n C_{iw}^2}, \frac{\sigma_g^2}{\sum_{i=1}^n C_{iw}^2})$$

B.3 Posterior Sampling Algorithm Details for Product Threshold Gaussian (PTG) Prior

Let $\boldsymbol{\Theta}_{PTG} = (\boldsymbol{\beta_m}, \boldsymbol{\alpha_a}, \tilde{\boldsymbol{\beta}}_{\boldsymbol{m}}, \tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}}, \tau_\beta^2, \tau_\alpha^2, \beta_a, \boldsymbol{\beta_c}, \boldsymbol{\alpha_c}, \sigma_e^2, \boldsymbol{\Sigma})$ denote all the unknown parameters in the model. Under the PTG prior, the joint log posterior distribution is,

$$\begin{aligned}&\log P(\boldsymbol{\Theta}_{PTG}|\{Y_i, \boldsymbol{M_i}, A_i, \boldsymbol{C_i}\}_{i=1}^n) \\ \propto\ &\sum_{i=1}^n \log P(Y_i|\boldsymbol{M_i}, \boldsymbol{\beta_m}, \sigma_e^2, \beta_a, \boldsymbol{\beta_c}, A_i, \boldsymbol{C_i}) + \log P(\boldsymbol{M_i}|\boldsymbol{\alpha_a}, \boldsymbol{\alpha_c}, \boldsymbol{\Sigma}, A_i, \boldsymbol{C_i})\end{aligned}$$

$$
\begin{aligned}
& + \log P(\boldsymbol{\Theta}_{PTG}) \\
= & \sum_{i=1}^{n} -\frac{1}{2}\log\sigma_e^2 - \frac{1}{2\sigma_e^2}(Y_i - \boldsymbol{M_i^T}\boldsymbol{\beta_m} - A_i\beta_a - \boldsymbol{C_i^T}\boldsymbol{\beta_c})^2 \\
& -\frac{1}{2}\log|\boldsymbol{\Sigma}| - \frac{1}{2}(\boldsymbol{M_i} - A_i\boldsymbol{\alpha_a} - \boldsymbol{\alpha_c}\boldsymbol{C_i})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{M_i} - A_i\boldsymbol{\alpha_a} - \boldsymbol{\alpha_c}\boldsymbol{C_i}) \\
& + \sum_{i=1}^{p} -\frac{1}{2}\log\tau_\beta^2 - \frac{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j^2}{2\tau_\beta^2} + \sum_{i=1}^{p} -\frac{1}{2}\log\tau_\alpha^2 - \frac{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j^2}{2\tau_\alpha^2} \\
& -\frac{q}{2}\log 2\pi\sigma_c^2 - \frac{\boldsymbol{\beta_c^T}\boldsymbol{\beta_c}}{2\sigma_c^2} - \frac{pq}{2}\log 2\pi\sigma_c^2 - \sum_{j=1}^{p}\frac{\boldsymbol{\alpha_{cj}^T}\boldsymbol{\alpha_{cj}}}{2\sigma_c^2}
\end{aligned}
$$

Sampling $(\boldsymbol{\beta_m})_j$

For $(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j$, we denote its threshold conditional on the other parameters as

$$
u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j} = \begin{cases} \min(\lambda_1, \lambda_0/|(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j|), & \text{for } (\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j \neq 0 \\ \lambda_1, & \text{for } (\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j = 0 \end{cases}
$$

$$
\log p((\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j || (\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j| < u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}) \propto -(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j^2/(2\tau_\beta^2)
$$

$$
(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j || (\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j| < u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j} \sim TN(0, \tau_\beta^2, -u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}, u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j})
$$

where $TN(\mu, \sigma^2, a, b)$ denotes a truncated normal distribution with mean μ, variance σ^2 truncated between $[a, b]$.

$$
\begin{aligned}
& \log p((\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j || (\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j| >= u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}) \\
& \propto -\frac{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j^2}{2\tau_\beta^2} - \sum_{i=1}^{n}\{\frac{(M_i^{(j)}(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j)^2}{2\sigma_e^2} + \sigma_e^{-2}M_i^{(j)}(Y_i - A_i\beta_a - \sum_{s\neq j} M_i^{(s)}(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_s - \boldsymbol{C_i^T}\boldsymbol{\beta_c})(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j\}
\end{aligned}
$$

$$
(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j | (\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j >= u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j} \sim TN(\mu_{mj}, s_{mj}^2, u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}, \infty)
$$

$$
(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j | (\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j <= -u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j} \sim TN(\mu_{mj}, s_{mj}^2, -\infty, -u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j})
$$

$$\mu_{mj} = \frac{\sum_{i=1}^{n} M_i^{(j)}(Y_i - A_i\beta_a - \sum_{s\neq j} M_i^{(s)}(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_s - \boldsymbol{C}_{\boldsymbol{i}}^T\boldsymbol{\beta}_{\boldsymbol{c}})}{\sigma_e^2/\tau_\beta^2 + \sum_{i=1}^{n}(M_i^{(j)})^2}, s_{mj}^2 = \frac{1}{1/\tau_\beta^2 + \sum_{i=1}^{n}(M_i^{(j)})^2/\sigma_e^2}$$

And,

$$\begin{aligned} p(|(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j| < u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}) &= \frac{B_1}{B_1+B_2+B_3} \\ p((\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j >= u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}) &= \frac{B_2}{B_1+B_2+B_3} \\ p((\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j <= -u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}) &= \frac{B_3}{B_1+B_2+B_3} \end{aligned}$$

where $B_1 = \int_{-u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}}^{u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}} \frac{1}{\sqrt{2\pi\tau_\beta^2}}\exp(-\frac{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j^2}{2\tau_\beta^2}) = 1 - 2\Phi(-\frac{u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}}{\tau_\beta^2})$, $\Phi(x)$ is the CDF for standard normal distribution, $B_2 = \exp(\mu_{mj}^2/2s_{mj}^2 + \log(s_{mj}) - \log(\tau_\beta))(1 - \Phi(\frac{u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}}{\tau_\beta^2}))$, $B_3 = \exp(\mu_{mj}^2/2s_{mj}^2 + \log(s_{mj}) - \log(\tau_\beta))\Phi(-\frac{u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j}}{\tau_\beta^2})$.

$$(\boldsymbol{\beta}_{\boldsymbol{m}})_j = \begin{cases} (\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j, & \text{for } |(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j| >= u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j} \\ 0, & \text{for } |(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j| < u_{(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j} \end{cases}$$

Sampling $(\boldsymbol{\alpha}_{\boldsymbol{a}})_j$

For $(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j$, we denote its threshold conditional on the other parameters as

$$u_{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j} = \begin{cases} \min(\lambda_2, \lambda_0/|(\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j|), & \text{for } (\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j \neq 0 \\ \lambda_2, & \text{for } (\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}})_j = 0 \end{cases}$$

$$\log p((\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j || |(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j| < u_{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j}) \propto -(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j^2/(2\tau_\alpha^2)$$

$$(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j || |(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j| < u_{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j} \sim TN(0, \tau_\alpha^2, -u_{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j}, u_{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j})$$

$$\log p((\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j || |(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j| >= u_{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j}) \propto -\frac{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j^2}{2\tau_\alpha^2} - \sum_{i=1}^{n}\{\frac{(A_i(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j)^2}{2\sigma_g^2} + \sigma_g^{-2}A_i(M_i^{(j)} - (\boldsymbol{\alpha}_{\boldsymbol{c}}\boldsymbol{C}_{\boldsymbol{i}})_j)(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j\}$$

$$(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j | (\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j >= u_{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j} \sim TN(\mu_{aj}, s_{aj}^2, u_{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j}, \infty)$$

$$(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j | (\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j <= -u_{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j} \sim TN(\mu_{aj}, s_{aj}^2, -\infty, -u_{(\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}})_j})$$

$$\mu_{aj} = \frac{\sum_{i=1}^{n} A_i(M_i^{(j)} - (\boldsymbol{\alpha_C C_i})_j)}{\sigma_g^2/\tau_\alpha^2 + \sum_{i=1}^{n} A_i^2}, s_{aj}^2 = \frac{1}{1/\tau_\alpha^2 + \sum_{i=1}^{n} A_i^2/\sigma_g^2}$$

And,

$$p(|(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j| < u_{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j}) = \frac{A_1}{A_1 + A_2 + A_3}$$
$$p((\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j >= u_{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j}) = \frac{A_2}{A_1 + A_2 + A_3}$$
$$p((\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j <= -u_{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j}) = \frac{A_3}{A_1 + A_2 + A_3}$$

where $A_1 = \int_{-u_{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j}}^{u_{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j}} \frac{1}{\sqrt{2\pi\tau_\alpha^2}} \exp(-\frac{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j^2}{2\tau_\alpha^2}) = 1 - 2\Phi(-\frac{u_{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j}}{\tau_\alpha^2})$, $\Phi(x)$ is the CDF for standard normal distribution, $A_2 = \exp(\mu_{aj}^2/2s_{aj}^2 + \log(s_{aj}) - \log(\tau_\alpha))(1 - \Phi(\frac{u_{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j}}{\tau_\alpha^2}))$, $A_3 = \exp(\mu_{aj}^2/2s_{aj}^2 + \log(s_{aj}) - \log(\tau_\alpha))\Phi(-\frac{u_{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j}}{\tau_\alpha^2})$.

$$(\boldsymbol{\alpha a})_j = \begin{cases} (\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j, & \text{for } |(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j| >= u_{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j} \\ 0, & \text{for } |(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j| < u_{(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j} \end{cases}$$

Sampling β_a

$$\log p(\beta_a|.) \propto -\frac{\beta_a^2}{2\sigma_a^2} - \sum_{i=1}^{n}\{\frac{(A_i\beta_a)^2}{2\sigma_1^2} - \sigma_1^{-2}A_i(Y_i - \boldsymbol{M_i^T\beta_m} - \boldsymbol{C_i^T\beta_c})\beta_a\}$$

$$p(\beta_a|.) \sim N(\frac{\sum_{i=1}^{n} A_i(Y_i - \boldsymbol{M_i^T\beta_m} - \boldsymbol{C_i^T\beta_c})}{\sigma_1^2/\sigma_a^2 + \sum_{i=1}^{n} A_i^2}, \frac{1}{1/\sigma_a^2 + \sum_{i=1}^{n} A_i^2/\sigma_1^2})$$

Sampling σ_a^2

$$\log p(\sigma_a^2|.) \propto -(\frac{1}{2} + h_a + 1)\log(\sigma_a^2) - (\frac{\beta_a^2}{2} + l_a)\sigma_a^{-2}$$

$$p(\sigma_a^2|.) \sim \text{inverse-gamma}(\frac{1}{2} + h_a, \frac{\beta_a^2}{2} + l_a)$$

Sampling σ_e^2

$$\log p(\sigma_e^2|.) = -(\frac{n}{2} + h_1 + 1)\log(\sigma_e^2) - (\frac{\sum_{i=1}^{n}(Y_i - \boldsymbol{M_i^T}\boldsymbol{\beta_m} - A_i\beta_a - \boldsymbol{C_i^T}\boldsymbol{\beta_c})^2}{2} + l_1)\sigma_e^{-2}$$

$$p(\sigma_e^2|.) \sim \text{inverse-gamma}(\frac{n}{2} + h_1, \frac{\sum_{i=1}^{n}(Y_i - \boldsymbol{M_i^T}\boldsymbol{\beta_m} - A_i\beta_a - \boldsymbol{C_i^T}\boldsymbol{\beta_c})^2}{2} + l_1)$$

Sampling σ_g^2

$$\begin{aligned}\log p(\sigma_g^2|.) &= -(\frac{pn}{2} + h_2 + 1)\log(\sigma_g^2) \\ &\quad -(\frac{\sum_{i=1}^{n}(\boldsymbol{M_i^T} - A_i\boldsymbol{\alpha_a} - \boldsymbol{C_i^T}\boldsymbol{\alpha_c})(\boldsymbol{M_i^T} - A_i\boldsymbol{\alpha_a} - \boldsymbol{C_i^T}\boldsymbol{\alpha_c})^T}{2} + l_2)\sigma_g^{-2}\end{aligned}$$

$$p(\sigma_g^2|.) \sim \text{inverse-gamma}(\frac{pn}{2}+h_2, \frac{\sum_{i=1}^{n}(\boldsymbol{M_i^T} - A_i\boldsymbol{\alpha_a} - \boldsymbol{C_i^T}\boldsymbol{\alpha_c})(\boldsymbol{M_i^T} - A_i\boldsymbol{\alpha_a} - \boldsymbol{C_i^T}\boldsymbol{\alpha_c})^T}{2}+l_2)$$

Sampling τ_β^2

$$\log p(\tau_\beta^2|.) = -(\frac{q}{2} + k_m + 1)\log(\tau_\beta^2) - (\frac{\sum_{j=1}^{q}(\tilde{\boldsymbol{\beta}}\boldsymbol{m})_j^2}{2} + l_m)\tau_\beta^{-2}$$

$$p(\tau_\beta^2|.) \sim \text{inverse-gamma}(\frac{q}{2} + k_m, \frac{\sum_{j=1}^{q}(\tilde{\boldsymbol{\beta}}\boldsymbol{m})_j^2}{2} + l_m)$$

Sampling τ_α^2

$$\log p(\tau_\alpha^2|.) = -(\frac{q}{2} + k_{ma} + 1)\log(\tau_\alpha^2) - (\frac{\sum_{j=1}^{q}(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j^2}{2} + l_{ma})\tau_\alpha^{-2}$$

$$p(\tau_\alpha^2|.) \sim \text{inverse-gamma}(\frac{q}{2} + k_{ma}, \frac{\sum_{j=1}^{q}(\tilde{\boldsymbol{\alpha}}\boldsymbol{a})_j^2}{2} + l_{ma})$$

Sampling β_{cw}

$$\log p(\beta_{cw}|.) = -\sum_{i=1}^{n}\{\frac{(C_{iw}\beta_{cw})^2}{2\sigma_e^2} + \sigma_e^{-2}C_{iw}(Y_i - \boldsymbol{M_i^T}\boldsymbol{\beta_m} - A_i\beta_a - \sum_{s\neq w}C_{iw}\beta_{cw})\beta_{cw}\}$$

$$p(\beta_{cw}|.) = N(\frac{\sum_{i=1}^{n} C_{iw}(Y_i - A_i\beta_a - \boldsymbol{M_i^T}\boldsymbol{\beta_m} - \sum_{s\neq w} C_{iw}\beta_{cw})}{\sum_{i=1}^{n} C_{iw}^2}, \frac{\sigma_e^2}{\sum_{i=1}^{n} C_{iw}^2})$$

Sampling $(\boldsymbol{\alpha_{cw}})_j$

$$\log p((\boldsymbol{\alpha_{cw}})_j|.) = -\sum_{i=1}^{n}\{\frac{(C_{iw}(\boldsymbol{\alpha_{cw}})_j)^2}{2\sigma_g^2} + \sigma_g^{-2}C_{iw}(M_i^{(j)} - A_i\alpha_{aj} - \sum_{s\neq w} C_{is}(\boldsymbol{\alpha_{cs}})_j)(\boldsymbol{\alpha_{cw}})_j\}$$

$$p((\boldsymbol{\alpha_{cw}})_j|.) = N(\frac{\sum_{i=1}^{n} C_{iw}(M_i^{(j)} - A_i\alpha_{aj} - \sum_{s\neq w} C_{is}(\boldsymbol{\alpha_{cs}})_j)}{\sum_{i=1}^{n} C_{iw}^2}, \frac{\sigma_g^2}{\sum_{i=1}^{n} C_{iw}^2})$$

B.4 Effects Distribution, Empirical FDR Results and Computing Time in Simulations

To better understand the generated effects under the three different data generating mechanism in the simulation Setting (A)-(C), we examine the corresponding distributions of the simulated non-zero marginal effects, $(\boldsymbol{\beta_m})_j$ (or $(\boldsymbol{\alpha_a})_j$) and indirect effects, $(\boldsymbol{\beta_m})_j(\boldsymbol{\alpha_a})_j$ in Figure B.1.

The PTG prior model essentially produces effects truncated away from zero, where the thresholding parameter $\lambda = (\lambda_0, \lambda_1, \lambda_2)$ is determined by the proportion of non-zero effects (Setting (A)). For example, choosing $\lambda_0 = |\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}}\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}}|^{(95)}, \lambda_1 = |\tilde{\boldsymbol{\beta}}_{\boldsymbol{m}}|^{(85)}, \lambda_2 = |\tilde{\boldsymbol{\alpha}}_{\boldsymbol{a}}|^{(93)}$ approximately makes $\pi_1 = 0.05, \pi_2 = 0.10, \pi_3 = 0.05, \pi_4 = 0.80$. The relatively small non-zero marginal effects are picked up by its indirect effects exceeding the product threshold. The Setting (B) with four components of bivariate Gaussian mixture is straightforward, and the resulting indirect effects distribute as a product of two normal distributions. Under the Setting (C), we can see that the horseshoe distribution has a tall spike near zero and heavy tails on large effects, and this generates uneven effects different from either PTG or GMM prior model. The distribution of the corresponding indirect effects show a stronger contrast between small and large effects.

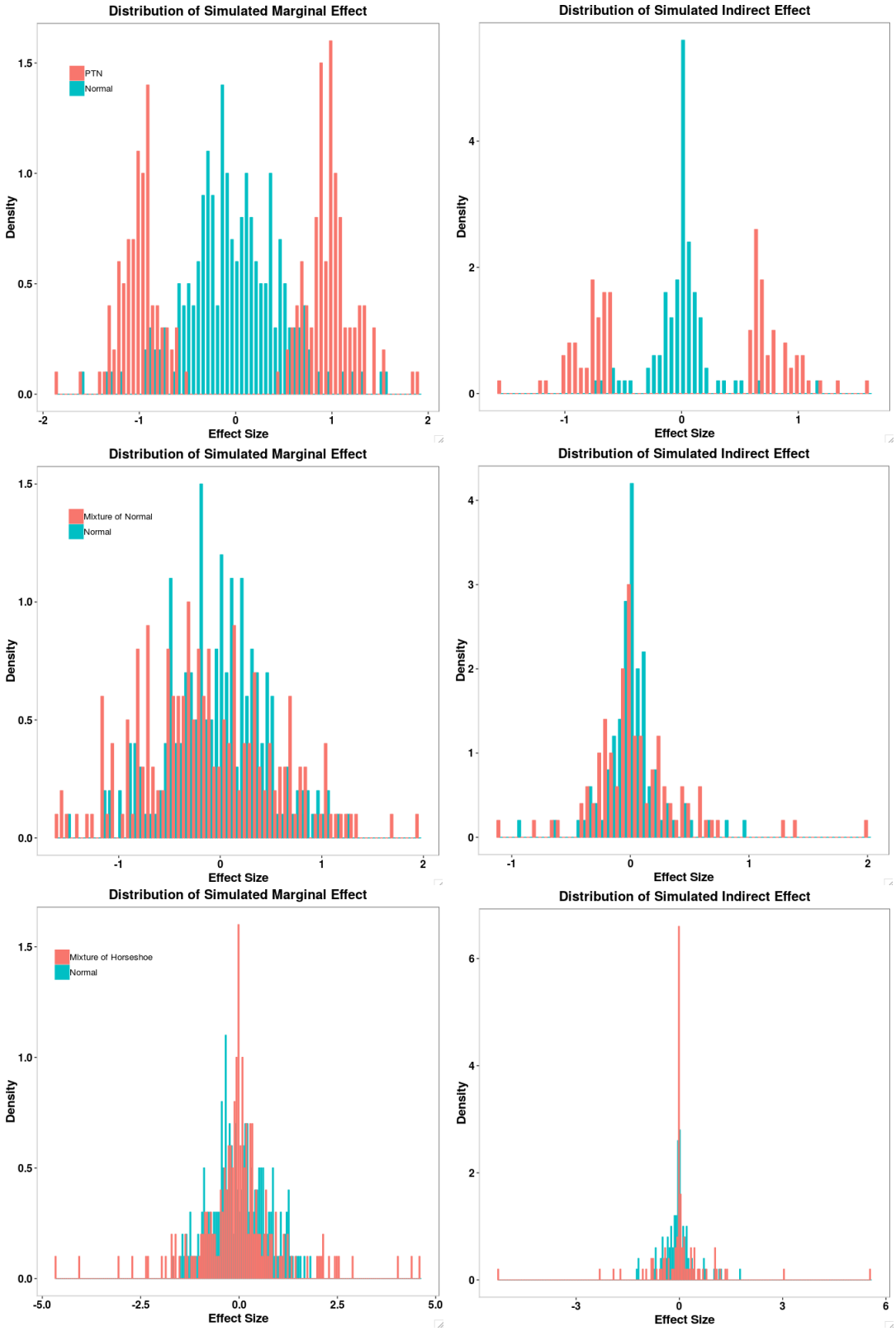

Figure B.1: The distributions of the simulated non-zero marginal effects, $(\boldsymbol{\beta_m})_j$ (or $(\boldsymbol{\alpha_a})_j$) and indirect effects, $(\boldsymbol{\beta_m})_j(\boldsymbol{\alpha_a})_j$ under the three simulation settings when $n = 100, p = 200$. Each row represents one scenario, i.e. effects under prior model PTG, GMM and Mixture of Horseshoe. We include marginal effects from normals with the same variances as the simulation distributions and the corresponding indirect effects as a comparison.

As a practical procedure, we suggest a cutoff on the posterior inclusion probabilities (PIP) to identify a significance threshold for declaring active mediators. To evaluate the performance of this significance rule, we report the empirical FDR and TPR in Table B.1 and B.2 under all the simulation scenarios. We find that at PIP = 0.5 cutoff, the two proposed methods, PTG and GMM, exhibit good selection performance while maintaining a reasonable FDR in most scenarios. At PIP = 0.9 cutoff, the two methods provide over conservative estimates of FDR, leading to reduced power in mediator selection. Therefore, we will use the 0.5 cutoff on the PIPs as a selection criterion in the following applications.

Method	TPR(FDR=0.1)	TPR(PIP>0.9)	FDR(PIP>0.9)	TPR(PIP>0.5)	FDR(PIP>0.5)
$n = 100, p = 200, p_{11} = 10,$ *fixed effects (I)*					
PTG	0.54(0.025)	0.27(0.017)	0.03(0.014)	0.55(0.017)	0.13(0.016)
GMM	0.42(0.023)	0.17(0.022)	0.03(0.021)	0.44(0.023)	0.16(0.017)
$n = 100, p = 200, p_{11} = 10,$ *fixed effects (II)*					
PTG	0.34(0.017)	0.27(0.008)	0.04(0.019)	0.37(0.013)	0.14(0.019)
GMM	0.39(0.020)	0.21(0.010)	0.03(0.016)	0.39(0.013)	0.11(0.017)
$n = 100, p = 200, p_{11} = 10, PTG$					
PTG	0.45(0.020)	0.19(0.014)	0.01(0.007)	0.49(0.018)	0.18(0.015)
GMM	0.43(0.015)	0.26(0.011)	0.03(0.012)	0.45(0.014)	0.11(0.012)
$n = 100, p = 200, p_{11} = 10, Gaussian$					
PTG	0.38(0.008)	0.26(0.008)	0.01(0.006)	0.56(0.010)	0.39(0.011)
GMM	0.41(0.006)	0.27(0.005)	0.01(0.002)	0.35(0.006)	0.06(0.008)
$n = 100, p = 200, p_{11} = 10, Horseshoe$					
PTG	0.30(0.015)	0.24(0.014)	0.08(0.016)	0.37(0.016)	0.38(0.019)
GMM	0.33(0.011)	0.26(0.011)	0.03(0.008)	0.35(0.012)	0.16(0.014)

Table B.1: Empirical estimates of TPR and FDR in simulations under $n = 100, p = 200$, p_{11} is the number of true active mediators. The results are based on 200 replicates for each setting, and the standard errors are shown within parentheses. TPR(FDR=0.1) is the true positive rate controlled at a fixed FDR of 10%; TPR(PIP>0.9) and FDR(PIP>0.9) are the empirical estimates when the PIP threshold for declaring active mediators is 0.9; TPR(PIP>0.5) and FDR(PIP>0.5) are the empirical estimates when the PIP threshold for declaring active mediators is 0.5.

We performed simulations on a single core of Intel(R) Xeon(R) Platinum 8176 CPU @ 2.10GHz, and the runtime comparison of the proposed methods is shown in Table C.1. For both the small sample scenario with $n = 100$, $p = 200$, and the large sample scenario with $n = 1000$, $p = 2000$, the proposed algorithms can be finished in a reasonable amount of time. We still acknowledge that future development of new algorithms and/or new methods

Method	TPR(FDR=0.1)	TPR(PIP>0.9)	FDR(PIP>0.9)	TPR(PIP>0.5)	FDR(PIP>0.5)
$n = 1000, p = 2000, p_{11} = 100,$ *fixed effects (I)*					
PTG	0.64(0.008)	0.49(0.017)	0.01(0.002)	0.55(0.017)	0.06(0.016)
GMM	0.61(0.009)	0.40(0.004)	0.01(0.003)	0.55(0.005)	0.07(0.010)
$n = 1000, p = 2000, p_{11} = 100,$ *fixed effects (II)*					
PTG	0.40(0.008)	0.20(0.004)	0.01(0.003)	0.37(0.012)	0.07(0.010)
GMM	0.48(0.006)	0.29(0.003)	0.01(0.002)	0.43(0.004)	0.06(0.007)
$n = 1000, p = 2000, p_{11} = 100, PTG$					
PTG	0.40(0.008)	0.19(0.004)	0.01(0.011)	0.44(0.007)	0.13(0.006)
GMM	0.37(0.010)	0.10(0.004)	0.05(0.008)	0.47(0.006)	0.17(0.007)
$n = 1000, p = 2000, p_{11} = 100, Gaussian$					
PTG	0.42(0.006)	0.20(0.005)	0.03(0.002)	0.51(0.005)	0.17(0.004)
GMM	0.51(0.007)	0.36(0.005)	0.01(0.002)	0.49(0.006)	0.10(0.004)
$n = 1000, p = 2000, p_{11} = 100, Horseshoe$					
PTG	0.29(0.008)	0.30(0.004)	0.05(0.006)	0.39(0.008)	0.24(0.004)
GMM	0.38(0.007)	0.35(0.004)	0.03(0.003)	0.45(0.004)	0.18(0.015)

Table B.2: Empirical estimates of TPR and FDR in simulations under $n = 1000, p = 2000$, p_{11} is the number of true active mediators. The results are based on 200 replicates for each setting, and the standard errors are shown within parentheses. TPR(FDR=0.1) is the true positive rate controlled at a fixed FDR of 10%; TPR(PIP>0.9) and FDR(PIP>0.9) are the empirical estimates when the PIP threshold for declaring active mediators is 0.9; TPR(PIP>0.5) and FDR(PIP>0.5) are the empirical estimates when the PIP threshold for declaring active mediators is 0.5.

will likely be required to scale our method to handle thousands of individuals and millions of mediators.

B.5 Detailed Description of MESA Data

MESA is a population-based longitudinal study designed to identify risk factors for the progression of subclinical cardiovascular disease (CVD) *Bild et al.* (2002). A total of 6,814

Method	$n = 100$, $p = 200$	$n = 1000$, $p = 2000$
PTG	30.5sec	23.0min
GMM	88.8sec	29.8min

Table B.3: The average runtime of the proposed methods for $n = 100$, $p = 200$ and $n = 1000$, $p = 2000$ in the simulations. Comparison was carried out on a single core of Intel(R) Xeon(R) Platinum 8176 CPU @ 2.10GHz. For the proposed methods, we in total ran 150,000 iterations.

non-Hispanic white, African-American, Hispanic, and Chinese-American women and men aged 45−84 without clinically apparent CVD were recruited between July 2000 and August 2002 from the following 6 regions in the US: Forsyth County, NC; Northern Manhattan and the Bronx, NY; Baltimore City and Baltimore County, MD; St. Paul, MN; Chicago, IL; and Los Angeles County, CA. Each field center recruited from locally available sources, which included lists of residents, lists of dwellings, and telephone exchanges. Neighborhood socioeconomic disadvantage scores for each neighborhood were created based on a principal components analysis of 16 census-tract level variables from the 2000 US Census. These variables reflect dimensions of education, occupation, income and wealth, poverty, employment, and housing. For the neighborhood measures, we use the cumulative average of the measure across all available MESA examinations. The descriptive statistics for the exposure and outcome can be found in Table C.2.

In the MESA data, between April 2010 and February 2012 (corresponding to MESA Exam 5), DNAm were assessed on a random subsample of 1,264 non-Hispanic white, African-American, and Hispanic MESA participants aged 55−94 from the Baltimore, Forsyth County, New York, and St. Paul field centers. After excluding respondents with missing data on one or more variables, we had phenotype and DNAm data from purified monocytes on a total of 1,225 individuals and we focused on this set of individuals for analysis. The detailed description of DNAm data collection, quantitation and data processing procedures can be found in Liu et al *Liu et al.* (2013). Briefly, the Illumina HumanMethylation450 BeadChip was used to measure DNAm, and bead-level data were summarized in GenomeStudio. Quantile normalization was performed using the *lumi* package with default settings *Du et al.* (2008). Quality control (QC) measures included checks for sex and race/ethnicity mismatches and outlier identification by multidimensional scaling plots. Further probe filtering criteria included: "detected" DNAm levels in <90% of MESA samples (detection p-value cut-off = 0.05), existence of a SNP within 10 base pairs of the target CpG site, overlap with a non-unique region, and suggestions by DMRcate *Chen et al.* (2013b) (mostly cross-reactive probes).

Those procedures leave us 403,713 autosomal probes for analysis.

For computational reasons, we first selected a subset of CpG sites to be used in the final multivariate mediation analysis model. In particular, for each single CpG site in turn, we fit the following linear mixed model to test the marginal association between the CpG site and the exposure variable:

$$M_i = A_i\psi_a + \boldsymbol{C_{1i}^T\psi_c} + \boldsymbol{Z_i^T\psi_u} + \epsilon_i, i = 1, ..., n \tag{B.1}$$

where A_i represents neighborhood SES value for the i'th individual and ψ_a is its coefficient; $\boldsymbol{C_{1i}}$ is a vector of covariates that include age, gender, race/ethnicity, childhood socioeconomic status, adult socioeconomic status and enrichment scores for each of 4 major blood cell types (neutrophils, B cells, T cells and natural killer cells) to account for potential contamination by non-monocyte cell types; $\boldsymbol{Z_i^T\psi_u}$ represent methylation chip and position random effects and are used to control for possible batch effects. The error term $\epsilon_i \sim MVN(0, \sigma^2 I_n)$ and is independent of the random effects. We obtained p-values for testing the null hypothesis $\psi_a = 0$ from the above model. We further applied the R/Bioconductor package BACON *van Iterson et al.* (2017) to these p-values to further adjust for possible inflation using an empirical null distribution. Based on these marginal p-values, we selected top 2,000 CpG sites with the smallest p-values for our Bayesian multivariate analysis.

B.6 Detailed Description of LIFECODES Data

The LIFECODES prospective birth cohort enrolled approximately 1,600 pregnant women between 2006 and 2008 at the Brigham and Womens Hospital in Boston, MA. Participants between 20 and 46 years of age were all at less than 15 weeks gestation at the initial study visit, and followed up to four visits (targeted at median 10, 18, 26, and 35 weeks gestation). At the initial study visit, questionnaires were administered to collect demographic and health-related information. Subjects' urine and plasma samples were collected at each

	Full Sample (n, %)	Neighborhood Socioeconomic Disadvantage Mean (SD)	Body Mass Index (BMI) Mean (SD)
Full sample	1225 (100)	-0.32 (1.11)	29.5 (5.49)
Age			
55−65 years	462 (38)	-0.18 (0.96)	30.3 (6.02)
66−75 years	397 (32)	-0.30 (1.16)	30.1 (5.21)
76−85 years	300 (24)	-0.47 (1.15)	28.2 (4.65)
86−95 years	66 (5)	-0.67 (1.46)	26.6 (4.66)
Race/ethnic group			
Non-Hispanic white	580 (47)	-0.56 (1.18)	28.7 (5.40)
African-American	263 (22)	-0.16 (0.98)	30.5 (5.69)
Hispanic	382 (31)	-0.05 (1.00)	30.0 (5.32)
Gender			
Female	633 (52)	-0.24 (1.09)	30.1 (6.20)
Male	592 (48)	-0.40 (1.12)	28.9 (4.54)

Table B.4: Characteristics of 1225 participants from MESA. %: proportion in the corresponding category. SD: standard deviation.

study visit. Among participants recruited in the LIFECODES cohort, 1,181 participants were followed until delivery and had live singleton infants. The birth outcome, gestational age, was also recorded at delivery, and preterm birth was defined as delivery prior to 37 weeks gestation. This study received institutional review board (IRB) approval from the Brigham and Womens Hospital and all participants provided written informed consent. All of the methods within this study were performed in accordance with the relevant guidelines and regulations approved by the IRB. Additional details regarding recruitment and study design can be found in *McElrath et al.* (2012); *Ferguson et al.* (2014b).

In this study, we focused on a subset of $n = 161$ individuals with their urine and plasma samples collected at one study visit occurring between 23.1 and 28.9 weeks gestation (median = 26 weeks). Characteristics of the subset sample is described in Table B.5. Subjects' urine samples were refrigerated (4°C) for a maximum of 2 hours before being processed and stored at −80°C. Approximately 10mL of blood was collected using ethylenediaminetetraacetic acid plasma tubes and temporarily stored at 4°C for less than 4 hours. Afterwards, blood was

centrifuged for 20 minutes and stored at −80°C. Environmental exposure analytes were measured from urine samples by NSF International in Ann Arbor, MI, following the methods developed by the Centers for Disease Control (CDC) (*Silva et al.*, 2007). Those exposure analytes includes phthalates, phenols and parabens, trace metals and polycyclic aromatic hydrocarbons. To adjust for urinary dilution, specific gravity (SG) levels were measured in each urine sample using a digital handheld refractometer (ATAGO Company Ltd., Tokyo, Japan), and was included as a covariate in regression models. Urine and plasma were subsequently analyzed for endogenous biomarkers, including 51 eicosanoids, five oxidative stress biomarkers and five immunological biomarkers in the present study. For a detailed description of the biomarkers that we analyzed and the media (urine or plasma) in which they were measured, please refer to *Aung et al.* (2019).

	Full Sample (n = 161)	**Preterm (<37 weeks gestation, n = 52)**	**Control (n = 109)**
Age[a]	32.7 (4.4)	32.1 (5.0)	33.0 (4.2)
BMI at Initial Visit[a]	26.7 (6.4)	28.5 (7.6)	25.8 (5.6)
Race/ethnic group[b]			
White	102 (63%)	29 (56%)	73 (67%)
African-American	18 (11%)	7 (13%)	11 (10%)
Other	41 (26%)	16 (31%)	25 (23%)
Gestational weeks[a]	37.5 (3.1)	34.1 (3.2)	39.1 (1.1)

Table B.5: Characteristics of all participants in the subset sample from the LIFECODES prospective birth cohort(n = 161). [a]Continuous variables presented as: mean (standard deviation). [b]Categorical variables presented as: count (proportion).

APPENDIX C

Supplement for Chapter IV

C.1 Posterior Sampling Algorithm Details for GMM-Potts

Sampling $\begin{bmatrix}(\boldsymbol{\beta_m})_j\\(\boldsymbol{\alpha_a})_j\end{bmatrix}$ *and* γ_j

$$\log p(\begin{bmatrix}(\boldsymbol{\beta_m})_j\\(\boldsymbol{\alpha_a})_j\end{bmatrix}|\gamma_j=k,.)\propto-\frac{1}{2}\begin{bmatrix}(\boldsymbol{\beta_m})_j\\(\boldsymbol{\alpha_a})_j\end{bmatrix}^\top(\boldsymbol{W_j}+\boldsymbol{V_k}^{-1})\begin{bmatrix}(\boldsymbol{\beta_m})_j\\(\boldsymbol{\alpha_a})_j\end{bmatrix}+\boldsymbol{w_j}^\top\begin{bmatrix}(\boldsymbol{\beta_m})_j\\(\boldsymbol{\alpha_a})_j\end{bmatrix}$$

where $\boldsymbol{W_j}=\text{diag}\{\sum_{i=1}^n(\sigma_e^2)^{-1}M_{ij}^2,\sum_{i=1}^n(\sigma_g^2)^{-1}A_i^2\}$, and

$\boldsymbol{w_j}=(\sum_{i=1}^n(\sigma_e^2)^{-1}(Y_i-A_i\beta_a-\sum_{j'\neq j}M_{ij'}(\boldsymbol{\beta_m})_{j'})M_{ij},\sum_{i=1}^n\boldsymbol{\Sigma}^{-1}M_{ij}A_i)^\top$

$$p(\begin{bmatrix}(\boldsymbol{\beta_m})_j\\(\boldsymbol{\alpha_a})_j\end{bmatrix}|\gamma_j=k,.)\sim MVN_2((\boldsymbol{W_j}+\boldsymbol{V_k}^{-1})^{-1}\boldsymbol{w_j},(\boldsymbol{W_j}+\boldsymbol{V_k}^{-1})^{-1})$$

$$\log p(\gamma_j=k|.)\propto-\frac{1}{2}\log|\boldsymbol{W_jV_k}+\boldsymbol{I_2}|+\frac{1}{2}\boldsymbol{w_j}^\top(\boldsymbol{W_j}+\boldsymbol{V_k}^{-1})^{-1}\boldsymbol{w_j}+\theta_{0k}+\sum_{i\sim j}\theta_{1k}I[\gamma_i=\gamma_j=k]$$

Sampling $\boldsymbol{V_k}$

$$\log p(\boldsymbol{V_k}|.)\quad\propto\quad-\frac{1}{2}(\sum_{j=1}^p I[\gamma_j=k]+df+d+1)\log|\boldsymbol{V_k}|-\frac{1}{2}\text{tr}(\boldsymbol{\Psi_0V_k}^{-1})$$

$$+\sum_{j=1}^{p} I[\gamma_j = k](-\frac{1}{2}\begin{bmatrix}(\boldsymbol{\beta m})_j\\(\boldsymbol{\alpha a})_j\end{bmatrix}^\top \boldsymbol{V_k^{-1}}\begin{bmatrix}(\boldsymbol{\beta m})_j\\(\boldsymbol{\alpha a})_j\end{bmatrix})$$

$$p(\boldsymbol{V_k}|.) \sim \text{Inv-Wishart}(\boldsymbol{\Psi_0} + \sum_{j=1}^{p} I[\gamma_j = k]\begin{bmatrix}(\boldsymbol{\beta m})_j\\(\boldsymbol{\alpha a})_j\end{bmatrix}\begin{bmatrix}(\boldsymbol{\beta m})_j\\(\boldsymbol{\alpha a})_j\end{bmatrix}^\top, \sum_{j=1}^{p} I[\gamma_j = k] + df)$$

Sampling β_a

$$\log p(\beta_a|.) \propto -\frac{\beta_a^2}{2\sigma_a^2} - \sum_{i=1}^{n}\{\frac{(A_i\beta_a)^2}{2\sigma_e^2} - \sigma_e^{-2}A_i(Y_i - \boldsymbol{M_i^\top}\boldsymbol{\beta m} - \boldsymbol{C_i^\top}\boldsymbol{\beta c})\beta_a\}$$

$$p(\beta_a|.) \sim N(\frac{\sum_{i=1}^{n} A_i(Y_i - \boldsymbol{M_i^\top}\boldsymbol{\beta m} - \boldsymbol{C_i^\top}\boldsymbol{\beta c})}{\sigma_e^2/\sigma_a^2 + \sum_{i=1}^{n} A_i^2}, \frac{1}{1/\sigma_a^2 + \sum_{i=1}^{n} A_i^2/\sigma_e^2})$$

Sampling σ_a^2

$$\log p(\sigma_a^2|.) \propto -(\frac{1}{2} + h_a + 1)\log(\sigma_a^2) - (\frac{\beta_a^2}{2} + l_a)\sigma_a^{-2}$$

$$p(\sigma_a^2|.) \sim \text{inverse-gamma}(\frac{1}{2} + h_a, \frac{\beta_a^2}{2} + l_a)$$

Sampling σ_e^2

$$\log p(\sigma_e^2|.) = -(\frac{n}{2} + h_1 + 1)\log(\sigma_e^2) - (\frac{\sum_{i=1}^{n}(Y_i - \boldsymbol{M_i^\top}\boldsymbol{\beta m} - A_i\beta_a - \boldsymbol{C_i^\top}\boldsymbol{\beta c})^2}{2} + l_1)\sigma_e^{-2}$$

$$p(\sigma_e^2|.) \sim \text{inverse-gamma}(\frac{n}{2} + h_1, \frac{\sum_{i=1}^{n}(Y_i - \boldsymbol{M_i^\top}\boldsymbol{\beta m} - A_i\beta_a - \boldsymbol{C_i^\top}\boldsymbol{\beta c})^2}{2} + l_1)$$

Sampling σ_g^2

$$\begin{aligned}\log p(\sigma_g^2|.) &= -(\frac{pn}{2} + h_2 + 1)\log(\sigma_g^2)\\ &\quad -(\frac{\sum_{i=1}^{n}(\boldsymbol{M_i^\top} - A_i\boldsymbol{\alpha a} - \boldsymbol{C_i^\top}\boldsymbol{\alpha c})(\boldsymbol{M_i^\top} - A_i\boldsymbol{\alpha a} - \boldsymbol{C_i^\top}\boldsymbol{\alpha c})^\top}{2} + l_2)\sigma_g^{-2}\end{aligned}$$

$$p(\sigma_g^2|.) \sim \text{inverse-gamma}(\frac{pn}{2}+h_2, \frac{\sum_{i=1}^{n}(\boldsymbol{M_i^\top} - A_i\boldsymbol{\alpha_a} - \boldsymbol{C_i^\top}\boldsymbol{\alpha_c})(\boldsymbol{M_i^\top} - A_i\boldsymbol{\alpha_a} - \boldsymbol{C_i^\top}\boldsymbol{\alpha_c})^\top}{2}+l_2)$$

Sampling $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1$

We update each of the $\theta_{0k}, \theta_{1k}, k \in \{1,2,3,4\}$ using a double Metropolis-Hastings (DMH) algorithm (*Liang*, 2010). For example, for updating θ_{0k}, we first propose a new $\theta_{0k}^\star$ from $\text{N}(\theta_{0k}, \tau_\theta^2)$ and then simulate an auxiliary variable $\boldsymbol{\gamma}^\star$ starting from $\boldsymbol{\gamma}$ based on the new $\boldsymbol{\theta}_0^\star, \boldsymbol{\theta}_1$ where all the elements are the same as $\boldsymbol{\theta}_0, \boldsymbol{\theta}_1$, excluding θ_{0k}. The proposed value $\theta_{0k}^\star$ will be accepted with probability $\min(1, r_\theta)$ and the Hastings ratio is,

$$r_\theta = \frac{\phi(\theta_{0k}^\star; \mu_{0k}, \sigma_{0k}^2)p(\boldsymbol{\gamma}^\star|\boldsymbol{\theta}_0, \boldsymbol{\theta}_1)p(\boldsymbol{\gamma}|\theta_0^\star, \boldsymbol{\theta}_1)}{\phi(\theta_{0k}; \mu_{0k}, \sigma_{0k}^2)p(\boldsymbol{\gamma}|\boldsymbol{\theta}_0, \boldsymbol{\theta}_1)p(\boldsymbol{\gamma}^\star|\theta_0^\star, \boldsymbol{\theta}_1)}$$

where $\phi(\theta_{0k}; \mu_{0k}, \sigma_{0k}^2)$ is the pdf for the normal distribution $\text{N}(\mu_{0k}, \sigma_{0k}^2)$. The form of $p(\boldsymbol{\gamma}|\boldsymbol{\theta}_0, \boldsymbol{\theta}_1)$ is given by Equation (4.5) as in the main text and the normalizing constants are canceled in the ratio.

Sampling β_{cw}

$$\log p(\beta_{cw}|.) = -\sum_{i=1}^{n}\{\frac{(C_{iw}\beta_{cw})^2}{2\sigma_e^2} + \sigma_e^{-2}C_{iw}(Y_i - \boldsymbol{M_i^\top}\boldsymbol{\beta_m} - A_i\beta_a - \sum_{s\neq w}C_{iw}\beta_{cw})\beta_{cw}\}$$

$$p(\beta_{cw}|.) \sim N(\frac{\sum_{i=1}^{n}C_{iw}(Y_i - A_i\beta_a - \boldsymbol{M_i^\top}\boldsymbol{\beta_m} - \sum_{s\neq w}C_{iw}\beta_{cw})}{\sum_{i=1}^{n}C_{iw}^2}, \frac{\sigma_e^2}{\sum_{i=1}^{n}C_{iw}^2})$$

Sampling $(\boldsymbol{\alpha_{cw}})_j$

$$\log p((\boldsymbol{\alpha_{cw}})_j|.) = -\sum_{i=1}^{n}\{\frac{(C_{iw}(\boldsymbol{\alpha_{cw}})_j)^2}{2\sigma_g^2} + \sigma_g^{-2}C_{iw}(M_i^{(j)} - A_i\alpha_{aj} - \sum_{s\neq w}C_{is}(\boldsymbol{\alpha_{cs}})_j)(\boldsymbol{\alpha_{cw}})_j\}$$

$$p((\boldsymbol{\alpha_{cw}})_j|.) \sim N(\frac{\sum_{i=1}^{n}C_{iw}(M_i^{(j)} - A_i\alpha_{aj} - \sum_{s\neq w}C_{is}(\boldsymbol{\alpha_{cs}})_j)}{\sum_{i=1}^{n}C_{iw}^2}, \frac{\sigma_g^2}{\sum_{i=1}^{n}C_{iw}^2})$$

Swendsen-Wang algorithm

We propose to use Swendsen-Wang algorithm (*Higdon*, 1998) to update the Markov random field, $\boldsymbol{\gamma}$. It is a particular case of auxiliary variable methods. In applying SW, we introduce "bond variables", $\boldsymbol{u} = \{u_{ij}, i \sim j\}$, for each neighbor pair $i \sim j$. Given $\boldsymbol{\gamma}$, the non-negative random variable u_{ij} are assumed to be independent and uniformly distributed as below,

$$p(u_{ij}|\boldsymbol{\gamma}) = \exp\{-\sum_{k=1}^{4} \theta_{1k} I[\gamma_i = \gamma_j = k]\} \times I[0 \leq u_{ij} \leq \exp\{\sum_{k=1}^{4} \theta_{1k} I[\gamma_i = \gamma_j = k]\}]$$

$$p(\boldsymbol{u}|\boldsymbol{\gamma}) = \prod_{i \sim j} \exp\{-\sum_{k=1}^{4} \theta_{1k} I[\gamma_i = \gamma_j = k]\} \times I[0 \leq u_{ij} \leq \exp\{\sum_{k=1}^{4} \theta_{1k} I[\gamma_i = \gamma_j = k]\}] \quad \text{(C.1)}$$

Furthermore,

$$p(\boldsymbol{\gamma}|\boldsymbol{u}, \cdot) \propto p(\boldsymbol{\beta_m}, \boldsymbol{\alpha_a}|\boldsymbol{\gamma}) \exp\{\sum_{i=1}^{p} \sum_{k=1}^{4} \theta_{0k} I[\gamma_i = k]\} \times \prod_{i \sim j} I[0 \leq u_{ij} \leq \exp\{\sum_{k=1}^{4} \theta_{1k} I[\gamma_i = \gamma_j = k]\}] \quad \text{(C.2)}$$

To sample from the joint posterior of $\boldsymbol{\gamma}$ and $\boldsymbol{u} = \{u_{ij}, i \sim j\}$, we can iteratively sample from Equation (C.1) and (C.2). To sample from (C.2), we note that $u_{ij} > 1$ implies that $\gamma_i = \gamma_j$, so that the bond variable $\boldsymbol{u}$ partitions mediators into same-labeled clusters, and this happens with a probability of $1 - \exp\{-\sum_{k=1}^{4} \theta_{1k} I[\gamma_i = \gamma_j = k]\}$. For a particular cluster, C, the probability of belonging to component k is $\propto \prod_{i \in C} p((\boldsymbol{\beta_m})_i, (\boldsymbol{\alpha_a})_i|\gamma_i) \exp\{\theta_{0k}\gamma_i\}$, and each cluster can be updated independently in turn according to its conditional distribution. The SW implementation can be described as below:

1. Update each bond variable according to a uniform distribution:

$$u_{ij}|\boldsymbol{\gamma} \sim U[0, \exp\{\sum_{k=1}^{4} \theta_{1k} I[\gamma_i = \gamma_j = k]\}]$$

   Bonds are forbidden from forming wherever the two neighbors are in different groups.

2. Form the same-labeled clusters (connected components) induced by u_{ij}

   i. The Union-Find algorithm

ii. Simplifies in the 1-D case

3. For each cluster C, update its label according to its conditional distribution,

$$p(\gamma_C = k|\cdot) \propto \prod_{i \in C} p((\boldsymbol{\beta m})_i, (\boldsymbol{\alpha a})_i|\gamma_i)\exp\{\theta_{0k}\gamma_i\}, k = 1, 2, 3, 4$$

We alternate between Swendsen-Wang updates of $\boldsymbol{\gamma}$ and single site Gibbs updates to ensure movement in large patches.

C.2 Posterior Sampling Algorithm Details for GMM-CorrS

To sample from the posterior distribution using the Pólya-Gamma method, simply iterate two steps:

Sampling w_{jk} for each j and k

$$w_{jk}|\cdot \sim \text{Pólya-Gamma}(n_{jk}, b_{kj})$$

where $n_{jk} = 1 - \sum_{k'<k} I(\gamma_j = k')$, $n_{j1} = 1$. The samples from Pólya-Gamma distribution can be generated using the algorithm and software in *Polson et al.* (2013).

Sampling $\boldsymbol{b_k}$

We can rewrite 4-dimensional multinomial in terms of 3 binomial densities $\widetilde{\pi}_{j1}$, $\widetilde{\pi}_{j2}$ and $\widetilde{\pi}_{j3}$. Specifically,

$$\begin{aligned} p(\boldsymbol{b_k}) \quad &\propto \quad \prod_j \widetilde{\pi}_{j1}^{I(\gamma_j=1)}((1-\widetilde{\pi}_{j1})\widetilde{\pi}_{j2})^{I(\gamma_j=2)}((1-\widetilde{\pi}_{j1})(1-\widetilde{\pi}_{j2})\widetilde{\pi}_{j3})^{I(\gamma_j=3)} \\ &\quad \times((1-\widetilde{\pi}_{j1})(1-\widetilde{\pi}_{j2})(1-\widetilde{\pi}_{j3})\widetilde{\pi}_{j4})^{I(\gamma_j=4)}\text{MVN}(\boldsymbol{a_k}, \sigma_{dk}^2\boldsymbol{D}) \\ &\propto \quad \prod_j \widetilde{\pi}_{j1}^{I(\gamma_j=1)}(1-\widetilde{\pi}_{j1})^{1-I(\gamma_j=1)}\widetilde{\pi}_{j2}^{I(\gamma_j=2)}(1-\widetilde{\pi}_{j2})^{I(\gamma_j=3)+I(\gamma_j=4)}\widetilde{\pi}_{j3}^{I(\gamma_j=3)}(1-\widetilde{\pi}_{j3})^{I(\gamma_j=4)} \\ &\quad \text{MVN}(\boldsymbol{a_k}, \sigma_{dk}^2\boldsymbol{D}) \end{aligned}$$

$$
\begin{aligned}
I(\gamma_j = 1) &\sim \text{Binom}(1, \text{expit}(b_{1j})), \\
I(\gamma_j = 2) &\sim \text{Binom}(n_{j2}, \text{expit}(b_{2j})), \\
I(\gamma_j = 3) &\sim \text{Binom}(n_{j3}, \text{expit}(b_{3j}))
\end{aligned}
$$

The multinomial distribution is now expressed with three binomial distributions involving b_{kj}, k = 1,2,3. Following the derivation in *Polson et al.* (2013), we will have,

$$
\boldsymbol{b_k}|\cdot \sim \text{MVN}(\boldsymbol{\mu_{bk}}, \boldsymbol{V_{bk}})
$$

$$
\begin{aligned}
\boldsymbol{V_{bk}} &= (\boldsymbol{\Omega} + (\sigma^2_{dk})^{-1}\boldsymbol{D}^{-1})^{-1} \\
\boldsymbol{\mu_{bk}} &= \boldsymbol{V_{bk}}(\boldsymbol{\kappa_k} + (\sigma^2_{dk})^{-1}\boldsymbol{D}^{-1}\boldsymbol{a_k})
\end{aligned}
$$

where $\boldsymbol{\Omega}$ is the diagnol matrix of w_{jk}'s, and $\boldsymbol{\kappa_k} = (I(\gamma_1 = k) - n_{1k}/2, I(\gamma_2 = k) - n_{2k}/2, ..., I(\gamma_p = k) - n_{pk}/2)$. Then we can update $\boldsymbol{\pi_j}$ accordingly.

Sampling σ^2_{dk}

$$
\sigma^2_{dk}|\cdot \sim \text{IG}(u + \frac{p}{2}, v + \frac{(\boldsymbol{b_k} - \boldsymbol{a_k})^T\boldsymbol{D}^{-1}(\boldsymbol{b_k} - \boldsymbol{a_k})}{2})
$$

The other parameters can be sampled in a similar way as in the GMM-Potts, with details described in the previous section.

C.3 Computing Time in Simulations

We performed simulations on a single core of Intel(R) Xeon(R) Platinum 8176 CPU @ 2.10GHz, and the runtime comparison of the proposed methods is shown in Table C.1. For both the small sample scenario with $n = 100$, $p = 200$, and the large sample scenario with $n = 1000$, $p = 2000$, the proposed algorithms can be finished in a reasonable amount of time. We still acknowledge that future development of new algorithms and/or new methods will likely be required to scale our method to handle thousands of individuals and millions

Method	$n = 100$, $p = 200$	$n = 1000$, $p = 2000$
GMM-CorrS	3.5 min	9.8 hr
GMM-Potts	2.2 min	4.0 hr

Table C.1: The average runtime of the proposed methods for $n = 100$, $p = 200$ and $n = 1000$, $p = 2000$ in the simulations. Comparison was carried out on a single core of Intel(R) Xeon(R) Platinum 8176 CPU @ 2.10GHz. For the proposed methods, we in total ran 150,000 iterations.

of mediators.

C.4 Detailed Description of MESA Data

MESA is a population-based longitudinal study designed to identify risk factors for the progression of subclinical cardiovascular disease (CVD) *Bild et al.* (2002). A total of 6,814 non-Hispanic white, African-American, Hispanic, and Chinese-American women and men aged 45−84 without clinically apparent CVD were recruited between July 2000 and August 2002 from the following 6 regions in the US: Forsyth County, NC; Northern Manhattan and the Bronx, NY; Baltimore City and Baltimore County, MD; St. Paul, MN; Chicago, IL; and Los Angeles County, CA. Each field center recruited from locally available sources, which included lists of residents, lists of dwellings, and telephone exchanges. Neighborhood socioeconomic disadvantage scores for each neighborhood were created based on a principal components analysis of 16 census-tract level variables from the 2000 US Census. These variables reflect dimensions of education, occupation, income and wealth, poverty, employment, and housing. For the neighborhood measures, we use the cumulative average of the measure across all available MESA examinations. The descriptive statistics for the exposure and outcome can be found in Table C.2.

In the MESA data, between April 2010 and February 2012 (corresponding to MESA Exam 5), DNAm were assessed on a random subsample of 1,264 non-Hispanic white, African-American, and Hispanic MESA participants aged 55−94 from the Baltimore, Forsyth County, New York, and St. Paul field centers. After excluding respondents with missing data on one or more

variables, we had phenotype and DNAm data from purified monocytes on a total of 1,225 individuals and we focused on this set of individuals for analysis. The detailed description of DNAm data collection, quantitation and data processing procedures can be found in Liu et al *Liu et al.* (2013). Briefly, the Illumina HumanMethylation450 BeadChip was used to measure DNAm, and bead-level data were summarized in GenomeStudio. Quantile normalization was performed using the *lumi* package with default settings *Du et al.* (2008). Quality control (QC) measures included checks for sex and race/ethnicity mismatches and outlier identification by multidimensional scaling plots. Further probe filtering criteria included: "detected" DNAm levels in <90% of MESA samples (detection p-value cut-off = 0.05), existence of a SNP within 10 base pairs of the target CpG site, overlap with a non-unique region, and suggestions by DMRcate *Chen et al.* (2013b) (mostly cross-reactive probes). Those procedures leave us 403,713 autosomal probes for analysis.

For computational reasons, we first selected a subset of CpG sites to be used in the final multivariate mediation analysis model. In particular, for each single CpG site in turn, we fit the following linear mixed model to test the marginal association between the CpG site and the exposure variable:

$$M_i = A_i\psi_a + \boldsymbol{C_{1i}^\top}\boldsymbol{\psi_c} + \boldsymbol{Z_i^\top}\boldsymbol{\psi_u} + \epsilon_i, i = 1, ..., n \tag{C.3}$$

where A_i represents neighborhood SES value for the i'th individual and ψ_a is its coefficient; $\boldsymbol{C_{1i}}$ is a vector of covariates that include age, gender, race/ethnicity, childhood socioeconomic status, adult socioeconomic status and enrichment scores for each of 4 major blood cell types (neutrophils, B cells, T cells and natural killer cells) to account for potential contamination by non-monocyte cell types; $\boldsymbol{Z_i^\top}\boldsymbol{\psi_u}$ represent methylation chip and position random effects and are used to control for possible batch effects. The error term $\epsilon_i \sim MVN(0, \sigma^2 I_n)$ and is independent of the random effects. We obtained p-values for testing the null hypothesis $\psi_a = 0$ from the above model. We further applied the R/Bioconductor package BACON

van Iterson et al. (2017) to these p-values to further adjust for possible inflation using an empirical null distribution. Based on these marginal p-values, we selected top 2,000 CpG sites with the smallest p-values for our Bayesian multivariate analysis.

	Full Sample (n, %)	**Neighborhood Socioeconomic Disadvantage Mean (SD)**	**Glucose Mean (SD)**
Full sample	1225 (100)	-0.32 (1.11)	29.5 (5.49)
Age			
55−65 years	462 (38)	-0.18 (0.96)	30.3 (6.02)
66−75 years	397 (32)	-0.30 (1.16)	30.1 (5.21)
76−85 years	300 (24)	-0.47 (1.15)	28.2 (4.65)
86−95 years	66 (5)	-0.67 (1.46)	26.6 (4.66)
Race/ethnic group			
Non-Hispanic white	580 (47)	-0.56 (1.18)	28.7 (5.40)
African-American	263 (22)	-0.16 (0.98)	30.5 (5.69)
Hispanic	382 (31)	-0.05 (1.00)	30.0 (5.32)
Gender			
Female	633 (52)	-0.24 (1.09)	30.1 (6.20)
Male	592 (48)	-0.40 (1.12)	28.9 (4.54)

Table C.2: Characteristics of 1225 participants from MESA. %: proportion in the corresponding category. SD: standard deviation.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abi Khalil, C. (2014), The emerging role of epigenetics in cardiovascular disease, *Therapeutic advances in chronic disease*, *5*(4), 178–187.

Albert, J. M. (2008), Mediation analysis via potential outcomes models., *Statistics in medicine*, *27*(8), 1282–1304.

Albert, J. M., and S. Nelson (2011), Generalized causal mediation analysis, *Biometrics*, *67*(3), 1028–1038.

Alegbeleye, O. O., B. O. Opeolu, and V. A. Jackson (2017), Polycyclic aromatic hydrocarbons: a critical review of environmental occurrence and bioremediation, *Environmental management*, *60*(4), 758–783.

Atchison, J., and S. M. Shen (1980), Logistic-normal distributions: Some properties and uses, *Biometrika*, *67*(2), 261–272.

Aung, M. T., Y. Yu, K. K. Ferguson, D. E. Cantonwine, L. Zeng, T. F. McElrath, S. Pennathur, B. Mukherjee, and J. D. Meeker (2019), Prediction and associations of preterm birth and its subtypes with eicosanoid enzymatic pathways and inflammatory markers, *Scientific reports*, *9*(1), 1–17.

Aung, M. T., Y. Song, K. K. Ferguson, D. E. Cantonwine, L. Zeng, T. F. McElrath, S. Pennathur, J. D. Meeker, and B. Mukherjee (2020), Application of a novel analytical pipeline for high-dimensional multivariate mediation analysis of environmental data, *medRxiv*, doi: 10.1101/2020.05.30.20117655.

Banerjee, B., M. Mustafa, T. Sharma, V. Tyagi, R. Ahmed, A. Tripathi, and K. Guleria (2014), Assessment of toxicogenomic risk factors in etiology of preterm delivery, *Reprod Syst Sex Disord*, *3*(2), 1–10.

Baron, R. M., and D. A. Kenny (1986), The moderator–mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations., *Journal of personality and social psychology*, *51*(6), 1173.

Best, N., S. Richardson, and A. Thomson (2005), A comparison of bayesian spatial models for disease mapping, *Statistical methods in medical research*, *14*(1), 35–59.

Bild, D. E., D. A. Bluemke, G. L. Burke, R. Detrano, A. V. Diez Roux, A. R. Folsom, et al. (2002), Multi-ethnic study of atherosclerosis: objectives and design, *American journal of epidemiology*, *156*(9), 871–881.

Blei, D. M., J. D. Lafferty, et al. (2007), A correlated topic model of science, *The Annals of Applied Statistics*, *1*(1), 17–35.

Bobb, J. F., L. Valeri, B. Claus Henn, D. C. Christiani, R. O. Wright, M. Mazumdar, J. J. Godleski, and B. A. Coull (2015), Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures, *Biostatistics*, *16*(3), 493–508.

Boyle, E. A., Y. I. Li, and J. K. Pritchard (2017), An expanded view of complex traits: from polygenic to omnigenic, *Cell*, *169*(7), 1177–1186.

Burke, G. L., A. G. Bertoni, S. Shea, R. Tracy, K. E. Watson, R. S. Blumenthal, H. Chung, and M. R. Carnethon (2008), The impact of obesity on cardiovascular disease risk factors and subclinical vascular disease: the multi-ethnic study of atherosclerosis, *Archives of internal medicine*, *168*(9), 928–935.

Cai, Q., J. Kang, T. Yu, et al. (2018), Bayesian network marker selection via the thresholded graph laplacian gaussian prior, *Bayesian Analysis*.

Cai, Q., J. Kang, and T. Yu (2020), Bayesian network marker selection via the thresholded graph laplacian gaussian prior, *Bayesian Analysis*, *15*(1), 79–102.

Cai, T. T., J. Ma, and L. Zhang (2019), Chime: Clustering of high-dimensional gaussian mixtures with em algorithm and its optimality, *The Annals of Statistics*, *47*(3), 1234–1267.

Callaghan, W. M., M. F. MacDorman, S. A. Rasmussen, C. Qin, and E. M. Lackritz (2006), The contribution of preterm birth to infant mortality rates in the united states, *Pediatrics*, *118*(4), 1566–1573.

Carbonetto, P., and M. Stephens (2013), Integrated enrichment analysis of variants and pathways in genome-wide association studies indicates central role for il-2 signaling genes in type 1 diabetes, and cytokine signaling genes in crohn's disease, *PLoS genetics*, *9*(10), e1003,770.

Carvalho, C. M., N. G. Polson, and J. G. Scott (2009), Handling sparsity via the horseshoe, in *Artificial Intelligence and Statistics*, pp. 73–80.

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

Chang, C., S. Kundu, and Q. Long (2018), Scalable bayesian variable selection for structured high-dimensional data, *Biometrics*, *74*(4), 1372–1382.

Chekouo, T., F. C. Stingo, M. Guindani, K.-A. Do, et al. (2016), A bayesian predictive model for imaging genetics with application to schizophrenia, *The Annals of Applied Statistics*, *10*(3), 1547–1571.

Chen, J., J. Zhu, Z. Wang, X. Zheng, and B. Zhang (2013a), Scalable inference for logistic-normal topic models, in *Advances in Neural Information Processing Systems*, pp. 2445–2453.

Chén, O. Y., C. Crainiceanu, E. L. Ogburn, B. S. Caffo, T. D. Wager, and M. A. Lindquist (2017), High-dimensional multivariate mediation with application to neuroimaging data, *Biostatistics*, *19*(2), 121–136.

Chen, R.-B., C.-H. Chu, S. Yuan, and Y. N. Wu (2016), Bayesian sparse group selection, *Journal of Computational and Graphical Statistics*, *25*(3), 665–683.

Chen, Y.-a., M. Lemire, S. Choufani, D. T. Butcher, D. Grafodatskaya, B. W. Zanke, et al. (2013b), Discovery of cross-reactive probes and polymorphic cpgs in the illumina infinium humanmethylation450 microarray, *Epigenetics*, *8*(2), 203–209.

Daimon, M., et al. (2011), Association of the clusterin gene polymorphisms with type 2 diabetes mellitus, *Metabolism*, *60*(6), 815–822.

Daniel, R., B. De Stavola, S. Cousens, and S. Vansteelandt (2015), Causal mediation analysis with multiple mediators, *Biometrics*, *71*(1), 1–14.

Dayeh, T., et al. (2014), Genome-wide dna methylation analysis of human pancreatic islets from type 2 diabetic and non-diabetic donors identifies candidate genes that influence insulin secretion, *PLoS Genet*, *10*(3), e1004,160.

Derkach, A., R. M. Pfeiffer, T. H. Chen, and J. N. Sampson (2019), High dimensional mediation analysis with latent variables, *Biometrics*.

Dhandapany, P. S., et al. (2009), A common mybpc3 (cardiac myosin binding protein c) variant associated with cardiomyopathies in south asia, *Nature genetics*, *41*(2), 187–191.

Diseth, Å., and T. Kobbeltvedt (2010), A mediation analysis of achievement motives, goals, learning strategies, and academic achievement, *British Journal of Educational Psychology*, *80*(4), 671–687.

Ditlevsen, S., U. Christensen, J. Lynch, M. T. Damsgaard, and N. Keiding (2005), The mediation proportion: a structural equation approach for estimating the proportion of exposure effect on outcome explained by an intermediate variable, *Epidemiology*, pp. 114–120.

Djordjilović, V., C. M. Page, J. M. Gran, T. H. Nøst, T. M. Sandanger, M. B. Veierød, and M. Thoresen (2019), Global test for high-dimensional mediation: Testing groups of potential mediators, *Statistics in medicine*, *38*(18), 3346–3360.

Du, P., W. A. Kibbe, and S. M. Lin (2008), lumi: a pipeline for processing illumina microarray, *Bioinformatics*, *24*(13), 1547–1548.

Efron, B., et al. (2007), Size, power and false discovery rates, *The Annals of Statistics*, *35*(4), 1351–1377.

Elliott, M. R., T. E. Raghunathan, and Y. Li (2010), Bayesian inference for causal mediation effects using principal stratification with dichotomous mediators and outcomes, *Biostatistics*, *11*(2), 353–372.

Erdmann, J., T. Kessler, L. Munoz Venegas, and H. Schunkert (2018), A decade of genome-wide association studies for coronary artery disease: the challenges ahead, *Cardiovascular research*, *114*(9), 1241–1257.

Feng, D., L. Tierney, and V. Magnotta (2012), Mri tissue classification using high-resolution bayesian hidden markov normal mixture models, *Journal of the American Statistical Association*, *107*(497), 102–119.

Ferguson, K. K., and H. B. Chin (2017), Environmental chemicals and preterm birth: biological mechanisms and the state of the science, *Current epidemiology reports*, *4*(1), 56–71.

Ferguson, K. K., T. F. McElrath, Y.-H. Chen, B. Mukherjee, and J. D. Meeker (2014a), Longitudinal profiling of inflammatory cytokines and c-reactive protein during uncomplicated and preterm pregnancy, *American journal of reproductive immunology*, *72*(3), 326–336.

Ferguson, K. K., T. F. McElrath, Y.-A. Ko, B. Mukherjee, and J. D. Meeker (2014b), Variability in urinary phthalate metabolite levels across pregnancy and sensitive windows of exposure for the risk of preterm birth, *Environment international*, *70*, 118–124.

Ferguson, K. K., T. F. McElrath, Y.-H. Chen, R. Loch-Caruso, B. Mukherjee, and J. D. Meeker (2015), Repeated measures of urinary oxidative stress biomarkers during pregnancy and preterm birth, *American journal of obstetrics and gynecology*, *212*(2), 208–e1.

Fishbein, M., and I. Ajzen (1977), Belief, attitude, intention, and behavior: An introduction to theory and research.

Frühwirth-Schnatter, S., and R. Frühwirth (2010), Data augmentation and mcmc for binary and multinomial logit models, in *Statistical modelling and regression structures*, pp. 111–132, Springer.

Gallop, R., D. S. Small, J. Y. Lin, M. R. Elliott, M. Joffe, and T. R. Ten Have (2009), Mediation analysis with principal stratification, *Statistics in medicine*, *28*(7), 1108–1130.

Ge, T., C.-Y. Chen, Y. Ni, Y.-C. A. Feng, and J. W. Smoller (2019), Polygenic prediction via bayesian regression and continuous shrinkage priors, *Nature communications*, *10*(1), 1776.

Gelman, A., and D. B. Rubin (1992), Inference from iterative simulation using multiple sequences, *Statistical science*, pp. 457–472.

Goldsmith, J., L. Huang, and C. M. Crainiceanu (2014), Smooth scalar-on-image regression via spatial bayesian variable selection, *Journal of Computational and Graphical Statistics*, *23*(1), 46–64.

Guan, Y., and M. Stephens (2011), Bayesian variable selection regression for genome-wide association studies and other large-scale problems, *The Annals of Applied Statistics*, pp. 1780–1815.

Hao, X., P. Zeng, S. Zhang, and X. Zhou (2018), Identifying and exploiting trait-relevant tissues with multiple functional annotations in genome-wide association studies, *PLoS genetics*, *14*(1), e1007,186.

Hayes, A. F. (2009), Beyond baron and kenny: Statistical mediation analysis in the new millennium, *Communication monographs*, *76*(4), 408–420.

Higdon, D. M. (1998), Auxiliary variable methods for markov chain monte carlo with applications, *Journal of the American statistical Association*, *93*(442), 585–595.

Hjellvik, V., S. Sakshaug, and H. Strøm (2012), Body mass index, triglycerides, glucose, and blood pressure as predictors of type 2 diabetes in a middle-aged norwegian cohort of men and women, *Clinical epidemiology*, *4*, 213.

Hoerl, A., and R. Kennard (1988), Ridge regression, in encyclopedia of statistical sciences, vol. 8.

Holmes, C. C., L. Held, et al. (2006), Bayesian auxiliary variable models for binary and multinomial regression, *Bayesian analysis*, *1*(1), 145–168.

Hsieh, T.-T., S.-F. Chen, L.-M. Lo, M.-J. Li, Y.-L. Yeh, and T.-H. Hung (2012), The association between maternal oxidative stress at mid-gestation and subsequent pregnancy complications, *Reproductive sciences*, *19*(5), 505–512.

Huang, Y.-T. (2019), Variance component tests of multivariate mediation effects under composite null hypotheses, *Biometrics*.

Huang, Y.-T., and W.-C. Pan (2016), Hypothesis test of mediation effect in causal mediation model with high-dimensional continuous mediators, *Biometrics*, *72*(2), 402–413.

Huang, Y.-T., T. J. VanderWeele, and X. Lin (2014), Joint analysis of snp and gene expression data in genetic association studies of complex diseases, *The annals of applied statistics*, *8*(1), 352.

Huang, Y.-T., et al. (2019), Genome-wide analyses of sparse mediation effects under composite null hypotheses, *The Annals of Applied Statistics*, *13*(1), 60–84.

Imai, K., L. Keele, and D. Tingley (2010a), A general approach to causal mediation analysis., *Psychological methods*, *15*(4), 309.

Imai, K., L. Keele, T. Yamamoto, et al. (2010b), Identification, inference and sensitivity analysis for causal mediation effects, *Statistical science*, *25*(1), 51–71.

Ising, E. (1925), Beitrag zur theorie des ferromagnetismus, *Zeitschrift für Physik*, *31*(1), 253–258.

Jo, B. (2008), Causal inference in randomized experiments with mediational processes., *Psychological Methods*, *13*(4), 314.

Kang, J., B. J. Reich, and A.-M. Staicu (2018), Scalar-on-image regression via the soft-thresholded gaussian process, *Biometrika*, *105*(1), 165–184.

Kaplan, G. A., and J. E. Keil (1993), Socioeconomic factors and cardiovascular disease: a review of the literature., *Circulation*, *88*(4), 1973–1998.

Kim, C., M. J. Daniels, B. H. Marcus, and J. A. Roy (2017), A framework for bayesian nonparametric inference for causal effects of mediation, *Biometrics*, *73*(2), 401–409.

Kim, C., M. Daniels, Y. Li, K. Milbury, and L. Cohen (2018), A bayesian semiparametric latent variable approach to causal mediation, *Statistics in medicine*, *37*(7), 1149–1161.

Kim, C., M. Daniels, J. Hogan, C. Choirat, and C. Zigler (2019), Bayesian methods for multiple mediators: Relating principal stratification and causal mediation in the analysis of power plant emission controls, *arXiv preprint arXiv:1902.06194*.

Kiyama, R., and Y. Wada-Kiyama (2015), Estrogenic endocrine disruptors: Molecular mechanisms of action, *Environment International*, *83*, 11–40.

Kogelman, L. J., S. Cirera, D. V. Zhernakova, M. Fredholm, L. Franke, and H. N. Kadarmideen (2014), Identification of co-expression gene networks, regulatory genes and pathways for obesity based on adipose tissue rna sequencing in a porcine model, *BMC medical genomics*, *7*(1), 57.

Kriebel, J., et al. (2016), Association between dna methylation in whole blood and measures of glucose metabolism: Kora f4 study, *PloS one*, *11*(3), e0152,314.

Kvaløy, K., C. M. Page, and T. L. Holmen (2018), Epigenome-wide methylation differences in a group of lean and obese women–a hunt study, *Scientific reports*, *8*(1), 1–9.

Lei, M.-K., S. R. Beach, R. L. Simons, and R. A. Philibert (2015), Neighborhood crime and depressive symptoms among african american women: Genetic moderation and epigenetic mediation of effects, *Social Science & Medicine*, *146*, 120–128.

Li, F., and N. R. Zhang (2010), Bayesian variable selection in structured high-dimensional covariate spaces with applications in genomics, *Journal of the American statistical association*, *105*(491), 1202–1214.

Li, F., T. Zhang, Q. Wang, M. Z. Gonzalez, E. L. Maresh, J. A. Coan, et al. (2015), Spatial bayesian variable selection and grouping for high-dimensional scalar-on-image regression, *The Annals of Applied Statistics*, *9*(2), 687–713.

Li, Q., X. Wang, F. Liang, F. Yi, Y. Xie, A. Gazdar, and G. Xiao (2019), A bayesian hidden potts mixture model for analyzing lung cancer pathology images, *Biostatistics*, *20*(4), 565–581.

Liang, F. (2010), A double metropolis–hastings sampler for spatial models with intractable normalizing constants, *Journal of Statistical Computation and Simulation*, *80*(9), 1007–1022.

Lin, C.-Y., Y.-C. Lin, H.-K. Kuo, J.-J. Hwang, J.-L. Lin, P.-C. Chen, and L.-Y. Lin (2009), Association between acrylamide, blood insulin and insulin resistance in adults, *Diabetes care*.

Linderman, S., M. J. Johnson, and R. P. Adams (2015), Dependent multinomial models made easy: Stick-breaking with the pólya-gamma augmentation, in *Advances in Neural Information Processing Systems*, pp. 3456–3464.

Little, R. J., and D. B. Rubin (2019), *Statistical analysis with missing data*, vol. 793, Wiley.

Liu, F., S. Chakraborty, F. Li, Y. Liu, A. C. Lozano, et al. (2014), Bayesian regularization via graph laplacian, *Bayesian Analysis*, *9*(2), 449–474.

Liu, Y., J. Ding, L. M. Reynolds, K. Lohman, T. C. Register, A. De La Fuente, et al. (2013), Methylomics of gene expression in human monocytes, *Human molecular genetics*, *22*(24), 5065–5074.

MacKinnon, D. P. (2000), Contrasts in multiple mediator models, *Multivariate applications in substance use research: New methods for new questions*, pp. 141–160.

MacKinnon, D. P. (2008), *Introduction to statistical mediation analysis*, Routledge.

MacKinnon, D. P., G. Warsi, and J. H. Dwyer (1995), A simulation study of mediated effect measures, *Multivariate behavioral research*, *30*(1), 41–62.

MacKinnon, D. P., C. M. Lockwood, J. M. Hoffman, S. G. West, and V. Sheets (2002), A comparison of methods to test mediation and other intervening variable effects., *Psychological methods*, *7*(1), 83.

MacKinnon, D. P., A. J. Fairchild, and M. S. Fritz (2007a), Mediation analysis, *Annu. Rev. Psychol.*, *58*, 593–614.

MacKinnon, D. P., M. S. Fritz, J. Williams, and C. M. Lockwood (2007b), Distribution of the product confidence limits for the indirect effect: Program prodclin, *Behavior research methods*, *39*(3), 384–389.

McElrath, T. F., K.-H. Lim, E. Pare, J. Rich-Edwards, D. Pucci, R. Troisi, and S. Parry (2012), Longitudinal evaluation of predictive value for preeclampsia of circulating angiogenic factors through pregnancy, *American journal of obstetrics and gynecology*, *207*(5), 407–e1.

Milnerowicz, H., M. Ściskalska, and M. Dul (2015), Pro-inflammatory effects of metals in persons and animals exposed to tobacco smoke, *Journal of Trace Elements in Medicine and Biology*, *29*, 1–10.

Mitchell, T. J., and J. J. Beauchamp (1988), Bayesian variable selection in linear regression, *Journal of the American Statistical Association*, *83*(404), 1023–1032.

Moore, K., A. V. D. Roux, A. Auchincloss, K. R. Evenson, J. Kaufman, M. Mujahid, and K. Williams (2013), Home and work neighbourhood environments in relation to body mass index: the multi-ethnic study of atherosclerosis (mesa), *J Epidemiol Community Health*, *67*(10), 846–853.

Mulaik, S. A. (2009), *Linear causal modeling with structural equations*, CRC press.

Needham, B. L., J. A. Smith, W. Zhao, X. Wang, B. Mukherjee, S. L. Kardia, et al. (2015), Life course socioeconomic status and dna methylation in genes related to stress reactivity and inflammation: the multi-ethnic study of atherosclerosis, *Epigenetics*, *10*(10), 958–969.

Neelon, B. H., A. J. O′Malley, and S.-L. T. Normand (2010), A bayesian model for repeated measures zero-inflated count data with application to outpatient psychiatric service use, *Statistical modelling*, *10*(4), 421–439.

Newton, M. A., A. Noueiry, D. Sarkar, and P. Ahlquist (2004), Detecting differential gene expression with a semiparametric hierarchical mixture method, *Biostatistics*, *5*(2), 155–176.

Ni, Y., F. C. Stingo, and V. Baladandayuthapani (2019), Bayesian graphical regression, *Journal of the American Statistical Association*, *114*(525), 184–197.

Padula, A. M., E. M. Noth, S. K. Hammond, F. W. Lurmann, W. Yang, I. B. Tager, and G. M. Shaw (2014), Exposure to airborne polycyclic aromatic hydrocarbons during pregnancy and risk of preterm birth, *Environmental research*, *135*, 221–226.

Pan, H.-Z., D. Chang, L. G. Feng, F.-J. Xu, H.-Y. Kuang, and M.-J. Lu (2007), Oxidative damage to dna and its relationship with diabetic complications., *Biomedical and environmental sciences: BES*, *20*(2), 160.

Park, T., and G. Casella (2008), The bayesian lasso, *Journal of the American Statistical Association*, *103*(482), 681–686.

Pearl, J. (2001), Direct and indirect effects, in *Proceedings of the seventeenth conference on uncertainty in artificial intelligence*, pp. 411–420, Morgan Kaufmann Publishers Inc.

Peng, B., D. Zhu, B. P. Ander, X. Zhang, F. Xue, F. R. Sharp, and X. Yang (2013), An integrative framework for bayesian variable selection with informative priors for identifying genes and pathways, *PloS one*, *8*(7).

Polson, N. G., J. G. Scott, and J. Windle (2013), Bayesian inference for logistic models using pólya–gamma latent variables, *Journal of the American statistical Association*, *108*(504), 1339–1349.

Potts, R. B. (1952), Some generalized order-disorder transformations, in *Mathematical proceedings of the cambridge philosophical society*, vol. 48, pp. 106–109, Cambridge University Press.

Preacher, K. J., and A. F. Hayes (2008), Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models, *Behavior research methods*, *40*(3), 879–891.

Rakyan, V. K., et al. (2011), Identification of type 1 diabetes–associated dna methylation variable positions that precede disease diagnosis, *PLoS Genet*, *7*(9), e1002,300.

Raman, S., T. J. Fuchs, P. J. Wild, E. Dahl, and V. Roth (2009), The bayesian group-lasso for analyzing contingency tables, in *Proceedings of the 26th Annual International Conference on Machine Learning*, pp. 881–888.

Robins, J. M., and S. Greenland (1992), Identifiability and exchangeability for direct and indirect effects, *Epidemiology*, pp. 143–155.

Ročková, V., and E. I. George (2018), The spike-and-slab lasso, *Journal of the American Statistical Association*, *113*(521), 431–444.

Ross, C. E., and J. Mirowsky (2009), Neighborhood disorder, subjective alienation, and distress, *Journal of health and social behavior*, *50*(1), 49–64.

Rossell, D., and D. Telesca (2017), Nonlocal priors for high-dimensional estimation, *Journal of the American Statistical Association*, *112*(517), 254–265.

Rubin, D. B. (1974), Estimating causal effects of treatments in randomized and nonrandomized studies., *Journal of educational Psychology*, *66*(5), 688.

Rubin, D. B. (1980), Randomization analysis of experimental data: The fisher randomization test comment, *Journal of the American Statistical Association*, *75*(371), 591–593.

Rubin, D. B. (1986), Comment: Which ifs have causal answers, *Journal of the American Statistical Association*, *81*(396), 961–962.

Rucker, D. D., K. J. Preacher, Z. L. Tormala, and R. E. Petty (2011), Mediation analysis in social psychology: Current practices and new recommendations, *Social and Personality Psychology Compass*, *5*(6), 359–371.

Sadler, N. C., P. Nandhikonda, B.-J. Webb-Robertson, C. Ansong, L. N. Anderson, J. N. Smith, R. A. Corley, and A. T. Wright (2016), Hepatic cytochrome p450 activity, abundance, and expression throughout human development, *Drug Metabolism and Disposition*, *44*(7), 984–991.

Sasso, F. C., O. Carbonara, R. Nasti, B. Campana, R. Marfella, M. Torella, G. Nappi, R. Torella, and D. Cozzolino (2004), Glucose metabolism and coronary heart disease in patients with normal glucose tolerance, *Jama*, *291*(15), 1857–1863.

Schafer, J. L., and J. Kang (2008), Average causal effects from nonrandomized studies: a practical guide and simulated example., *Psychological methods*, *13*(4), 279.

Selvin, E., M. W. Steffes, H. Zhu, K. Matsushita, L. Wagenknecht, J. Pankow, et al. (2010), Glycated hemoglobin, diabetes, and cardiovascular risk in nondiabetic adults, *New England Journal of Medicine*, *362*(9), 800–811.

Shen, R., S. Wang, and Q. Mo (2013), Sparse integrative clustering of multiple omics data sets, *The annals of applied statistics*, *7*(1), 269.

Silva, M. J., E. Samandar, J. L. Preau Jr, J. A. Reidy, L. L. Needham, and A. M. Calafat (2007), Quantification of 22 phthalate metabolites in human urine, *Journal of Chromatography B*, *860*(1), 106–112.

Smith, J. A., W. Zhao, X. Wang, S. M. Ratliff, B. Mukherjee, S. L. Kardia, et al. (2017), Neighborhood characteristics influence dna methylation of genes involved in stress response and inflammation: The multi-ethnic study of atherosclerosis, *Epigenetics*, *12*(8), 662–673.

Smith, L. H., and T. J. VanderWeele (2019), Mediational e-values: approximate sensitivity analysis for unmeasured mediator–outcome confounding, *Epidemiology*, *30*(6), 835–837.

Sobel, M. E. (1982), Asymptotic confidence intervals for indirect effects in structural equation models, *Sociological methodology*, *13*, 290–312.

Song, Y., et al. (2018), Bayesian shrinkage estimation of high dimensional causal mediation effects in omics studies, *Biometrics*.

Stingo, F. C., Y. A. Chen, M. G. Tadesse, and M. Vannucci (2011), Incorporating biological information into linear models: A bayesian approach to the selection of pathways and genes, *The annals of applied statistics*, *5*(3).

Taguri, M., J. Featherstone, and J. Cheng (2015), Causal mediation analysis with multiple causally non-ordered mediators, *Statistical methods in medical research*, p. 0962280215615899.

Ten Have, T. R., and M. M. Joffe (2012), A review of causal estimation of effects in mediation analyses, *Statistical Methods in Medical Research*, *21*(1), 77–107.

Ten Have, T. R., M. M. Joffe, K. G. Lynch, G. K. Brown, S. A. Maisto, and A. T. Beck (2007), Causal mediation analyses with rank preserving models, *Biometrics*, *63*(3), 926–934.

Thomas, D. C., D. O. Stram, D. Conti, J. Molitor, and P. Marjoram (2003), Bayesian spatial modeling of haplotype associations, *Human heredity*, *56*(1-3), 32–40.

Tibshirani, R. (1996), Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288.

Tran, D. H., and Z. V. Wang (2019), Glucose metabolism in cardiac hypertrophy and heart failure, *Journal of the American Heart Association*, *8*(12), e012,673.

Tyrberg, B., K. A. Anachkov, S. A. Dib, J. Wang-Rodriguez, K.-H. Yoon, and F. Levine (2002), Islet expression of the dna repair enzyme 8-oxoguanosine dna glycosylase (ogg1) in human type 2 diabetes, *BMC endocrine disorders*, *2*(1), 1–10.

Valeri, L., and T. J. VanderWeele (2013), Mediation analysis allowing for exposure–mediator interactions and causal interpretation: Theoretical assumptions and implementation with sas and spss macros., *Psychological methods*, *18*(2), 137.

van Iterson, M., E. W. van Zwet, and B. T. Heijmans (2017), Controlling bias and inflation in epigenome-and transcriptome-wide association studies using the empirical null distribution, *Genome biology*, *18*(1), 19.

VanderWeele, T. (2016a), A unification of mediation and interaction: A 4-way decomposition (vol 25, pg 749, 2014), *Epidemiology*, *27*(5), E36–E36.

VanderWeele, T., and S. Vansteelandt (2014), Mediation analysis with multiple mediators, *Epidemiologic methods*, *2*(1), 95–115.

VanderWeele, T. J. (2011a), Causal mediation analysis with survival data, *Epidemiology (Cambridge, Mass.)*, *22*(4), 582.

VanderWeele, T. J. (2011b), Principal stratification–uses and limitations, *The international journal of biostatistics*, *7*(1), 1–14.

VanderWeele, T. J. (2016b), Mediation analysis: a practitioner's guide, *Annual review of public health*, *37*, 17–32.

VanderWeele, T. J., and S. Vansteelandt (2010), Odds ratios for mediation analysis for a dichotomous outcome, *American journal of epidemiology*, *172*(12), 1339–1348.

Vrijheid, M., M. Casas, M. Gascon, D. Valvi, and M. Nieuwenhuijsen (2016), Environmental pollutants and child healtha review of recent concerns, *International journal of hygiene and environmental health*, *219*(4-5), 331–342.

Winkler, G. (2012), *Image analysis, random fields and Markov chain Monte Carlo methods: a mathematical introduction*, vol. 27, Springer Science & Business Media.

Woodworth, R. S. (1930), Dynamic psychology.

Xia, F., J. Chen, W. K. Fung, and H. Li (2013), A logistic normal multinomial regression model for microbiome compositional data analysis, *Biometrics*, *69*(4), 1053–1063.

Xue, A., et al. (2018), Genome-wide association analyses identify 143 risk variants and putative regulatory mechanisms for type 2 diabetes, *Nature communications*, *9*(1), 1–14.

Yaghootkar, H., A. Stancáková, R. M. Freathy, J. Vangipurapu, M. N. Weedon, W. Xie, et al. (2015), Association analysis of 29,956 individuals confirms that a low frequency variant at ccnd2 halves the risk of type 2 diabetes by enhancing insulin secretion., *Diabetes*, p. db141456.

Yu, K., S. Wacholder, W. Wheeler, Z. Wang, N. Caporaso, M. T. Landi, and F. Liang (2012), A flexible bayesian model for studying gene–environment interaction, *PLoS genetics*, *8*(1).

Yuan, M., and Y. Lin (2005), Efficient empirical bayes variable selection and estimation in linear models, *Journal of the American Statistical Association*, *100*(472), 1215–1225.

Zeller, Z., et al. (2018), Impact of polymorphisms in ptk2 on intrinsic muscle strength.

Zeng, P., X. Hao, and X. Zhou (2018), Pleiotropic mapping and annotation selection in genome-wide association studies with penalized gaussian mixture models, *Bioinformatics*, *34*(16), 2797–2807.

Zhang, H., et al. (2016), Estimating and testing high-dimensional mediation effects in epigenetic studies, *Bioinformatics*, *32*(20), 3150–3154.

Zhang, L., V. Baladandayuthapani, B. K. Mallick, G. C. Manyam, P. A. Thompson, M. L. Bondy, and K.-A. Do (2014), Bayesian hierarchical structured variable selection methods with application to molecular inversion probe studies in breast cancer, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *63*(4), 595–620.

Zhao, Y., and X. Luo (2016), Pathway lasso: estimate and select sparse mediation pathways with high dimensional mediators, *arXiv preprint arXiv:1603.07749*.

Zheng, H., et al. (2015), Cnc-bzip protein nrf1-dependent regulation of glucose-stimulated insulin secretion, *Antioxidants & redox signaling*, *22*(10), 819–831.

Zheng, J., X. Xiao, Q. Zhang, and M. Yu (2014), Dna methylation: the pivotal interaction between early-life nutrition and glucose metabolism in later life, *British journal of nutrition*, *112*(11), 1850–1857.

Zhong, J., G. Agha, and A. A. Baccarelli (2016), The role of dna methylation in cardiovascular risk and disease, *Circulation research*, *118*(1), 119–131.

Zhou, X., P. Carbonetto, and M. Stephens (2013), Polygenic modeling with bayesian sparse linear mixed models, *PLoS genetics*, *9*(2), e1003,264.

Zhu, X., and M. Stephens (2017), Bayesian large-scale multiple regression with summary statistics from genome-wide association studies, *The annals of applied statistics*, *11*(3), 1561.