Compositional Mediation Analysis for Microbiome Studies

Michael B. Sohn and Hongzhe Li

Department of Biostatistics and Epidemiology, University of Pennsylvania

Abstract

Motivated by recent advances in causal inference on mediation analysis and problems in the analysis of metagenomic data, we consider the effect of a treatment on an outcome transmitted through microbes, or compositional mediators. Compositional and high dimensional natures of such mediators make the standard mediation analysis not directly applicable. In this paper, we propose a method for estimating the expected causal direct and indirect (or mediation) effects for a sparse high-dimensional compositional mediation model utilizing the algebra for compositions in the simplex space and the nature of compositional data. The test of the total and component-wise mediation effects are also proposed. We conduct extensive simulation studies to assess the performance of the proposed method and apply the method to a real metagenomic dataset to investigate the effect of fat intake on body mass index (BMI) mediated through the gut microbiome composition.

1 Introduction

It has been shown that fat intake is associated with BMI (Bray and Popkin, 1998) and obesity is associated with the gut microbiome (Ley *et al.*, 2006; Turnbaugh *et al.*, 2006). From this information, a very natural question to ask is whether fat intake has some effects on BMI mediated through the perturbation of the gut microbiome. The approach to answering this type of questions is known as "mediation analysis". The mediation analysis is a statistical method of studying the effect of a treatment or exposure on an outcome transmitted through intermediate variables, referred to as "mediators" or "intervening variables". It has been widely applied in various disciplines such as sociology, psychology and epidemiology and become increasingly popular due to recent advances in causal inference (Pearl, 2000; Rubin, 2005; Imai *et al.*, 2010), which clarifies the assumptions needed for a causal interpretation. Until recently, mediation analysis has been restricted to a single mediator as depicted in Figure 1, and the effect of a treatment on an outcome transmitted through a mediator is often formulated and implemented within the framework of linear structural equation models (LSEMs).

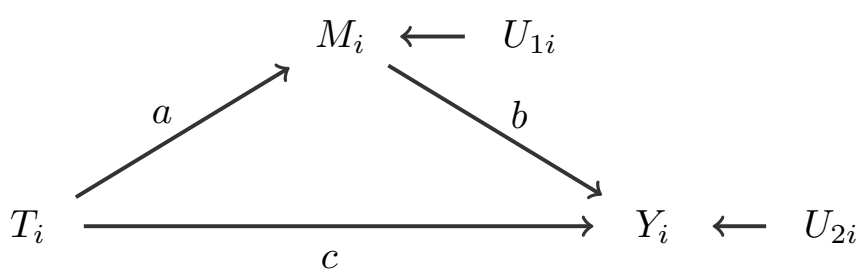

Figure 1: A single-mediator model: a, b and c are path coefficients; U_{1i} and U_{2i} are the disturbance or unmeasured variables for the mediator M_i and the outcome Y_i, respectively.

For instance, an LSEM for the path diagram in Figure 1 can be formulated as

$$M_i = a_0 + aT_i + U_{1i},$$
$$Y_i = c_0 + cT_i + bM_i + U_{2i},$$

where T_i, M_i and Y_i are the treatment, mediator and outcome variable for unit i; a, b and c are path coefficients; U_{1i} and U_{2i} are disturbance variables for M_i and Y_i, respectively. Under this model, the expected effect of T_i on Y_i transmitted through M_i, often called the *indirect or mediation effect*, is generally defined by the product of the path coefficients a and b; the expected effect of T_i on Y_i not transmitted through M_i, the *direct effect*, is defined by the path coefficient c. The model formulation for Figure 1 shows that the expected total effect of T_i on Y_i is the sum of the expected direct and indirect effects, $c + ab$.

In recent years, numerous studies have extended the applicability of mediation analysis: incorporating nonlinearity and interaction between the treatment and mediator (Pearl, 2001; Imai *et al.*, 2010; VanderWeele and Vansteelandt, 2010); multiple mediators (Preacher and Hayes, 2008; VanderWeele and Vansteelandt, 2014). A few studies have also proposed methods for high-dimensional mediators. Chén *et al.* (2015) proposed a method to estimate path coefficients of an LSEM by finding the linear combinations of mediators that maximize the likelihood of linear structural equations, which is similar to the principal components. Huang and Pan (2016) introduced a transformation model using the principal components and included the interaction in their model. Zhao and Luo (2016) proposed a sparse mediation model using a regularized LSEM approach.

In this paper, we contribute to extending the applicability of mediation analysis further by proposing an estimating method for the expected direct and indirect effects when mediators are compositional. Compositional data refer to proportions or percentages of a whole and arise frequently in a wide range of disciplines such as mineral components of a rock in geology and vote shares of an election in psephology. In microbiome and metagenomic studies, to account for different sizes of sequencing libraries for 16S rRNA or shotgun metagenomic sequencing, the sequencing reads (i.e., count data) are often normalized into proportions. This normalization introduces the unit-sum constraint (i.e., proportions sum to unity), which transforms a k dimensional Euclidian space $\mathbb{R}^k$ into a $k-1$ dimensional simplex space $\mathbb{S}^{k-1}$, thus making the statistical models for unconstrained data inappropriate for compositional data. To deal with the nature of compositional data, Aitchison (1982) introduced an axiomatic approach with various operations under a logratio transformation, which provides a one-to-one mapping between $\mathbb{R}^k$ and $\mathbb{S}^{k-1}$, and various researchers including himself have formalized and extended this approach. Aitchison and Bacon-Shone (1984) proposed a linear and quadratic log contrast model for compositional covariates. Billheimer *et al.* (2001) formulated the algebra for compositions in the simplex space. Lin *et al.* (2014) developed a linear log contrast regression model for compositional covariates in a high-dimensional setting, and Shi *et al.* (2016) generalized it.

We propose a framework for mediation analysis when mediators are high dimensional and compositional. Our framework utilizes two components: the compositional algebra of Billheimer *et al.* (2001) and the linear log contrast regression of Lin *et al.* (2014); Shi *et al.* (2016). We employ the first component (i.e., compositional algebra) to jointly estimate the effect of a randomly assigned treatment on all compositional mediators. To this end, we propose to minimize the difference between observed and estimated compositions using the compositional norm. We use the second component to quantify the effect of a treatment and compositional mediators on an outcome. Under this compositional mediation framework, we show that the expected direct and indirect effects are identifiable and have a causal interpretation under some assumptions.

Section 2 introduces the compositional mediation model for continuous outcomes and discusses model

assumptions and identifiability conditions. Section 3 describes the methods of estimating parameters and their covariance matrices. We also discuss null hypotheses for the total and component-wise mediation effects. Section 4 compares the performance of our method in extensive simulation studies with two methods that can be applied to compositional data. Section 4 also presents an application of the proposed method to a gut microbiome dataset. Finally, Section 5 presents a brief discussion of the method and results.

2 Compositional Mediation Model and Causal Interpretation

2.1 Notation

For unit i, we let T_i be the observed treatment; $\boldsymbol{M}_i$ be a vector of the observed k compositional mediators; Y_i be the observed outcome; $\boldsymbol{X}_i$ be a vector of the observed n_x covariates that are nondescendants of T_i or $\boldsymbol{M}_i$; $\boldsymbol{U}_{1i}$ be a vector of the disturbance variable for $\boldsymbol{M}_i$; U_{2i} be the disturbance variable for Y_i. We use $\mathcal{T}$, $\boldsymbol{\mathcal{M}}$ and $\boldsymbol{\mathcal{X}}$ to denote the support of the distribution of T_i, $\boldsymbol{M}_i$ and $\boldsymbol{X}_i$, respectively.

In the model assumptions and the identification for the expected direct and indirect effects, we adopt the potential outcomes framework: $\boldsymbol{M}_i(t)$ to denote the potential outcome for unit i would have under $T_i = t$ and $Y_i(t, \boldsymbol{m})$ to denote the potential outcome for unit i would have under $T_i = t$ and $\boldsymbol{M}_i = \boldsymbol{m}$. Thus, we can express the observed variable $\boldsymbol{M}_i = \boldsymbol{M}_i(T_i)$. Similarly, $Y_i = Y_i(T_i, \boldsymbol{M}_i(T_i))$.

2.2 Compositional Mediation Model

Suppose we have a random sample of size n from a population where for each unit i we observe Y_i, T_i, and $\boldsymbol{M}_i$. For clarity, we exclude $\boldsymbol{X}_i$ in the model formulation. See Appendix B for the model with $\boldsymbol{X}_i$. Note that $\boldsymbol{M}_i \in \mathbb{S}^{k-1}$ for all i, that is, $M_i = \{(M_{i1}, \ldots, M_{ik}) : M_{ij} > 0,\ j = 1, \ldots, k,\ \sum_{j=1}^{k} M_{ij} = 1\}$. Figure 2 shows the effect of T_i on Y_i mediated through $\boldsymbol{M}_i$.

Since $\boldsymbol{M}_i \in \mathbb{S}^{k-1}$, the multivariate regression, which assumes $\boldsymbol{M}_i \in \mathbb{R}^k$, is inappropriate to estimate the expected effect of T_i on $\boldsymbol{M}_i$. For the same reason, the multiple regression is not appropriate to estimate the expected effect of T_i and $\boldsymbol{M}_i$ on Y_i. Refer to Aitchison and Bacon-Shone (1984); Lin *et al.* (2014) for details. Before introducing the compositional mediation model that accounts for the nature of compositional data, we define two compositional operators. For two compositions $\boldsymbol{\eta}, \boldsymbol{\zeta} \in \mathbb{S}^{k-1}$, the perturbation operator is defined by

$$\boldsymbol{\eta} \oplus \boldsymbol{\zeta} = \left(\frac{\eta_1 \zeta_1}{\sum_{j=1}^{k} \eta_j \zeta_j}, \frac{\eta_2 \zeta_2}{\sum_{j=1}^{k} \eta_j \zeta_j}, \ldots, \frac{\eta_k \zeta_k}{\sum_{j=1}^{k} \eta_j \zeta_j} \right)^{\top},$$

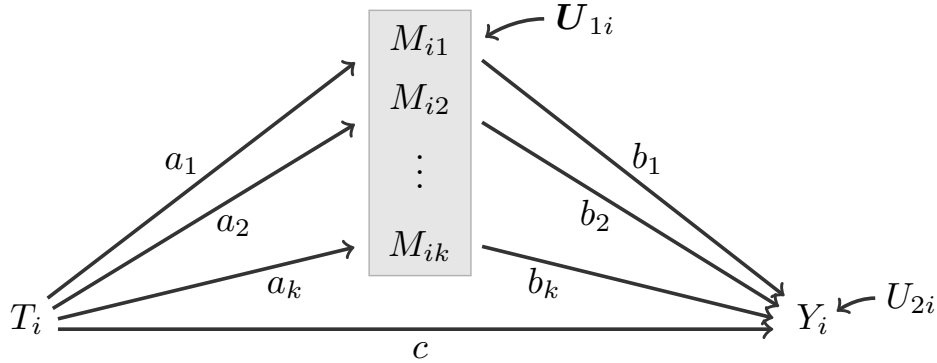

Figure 2: A compositional mediation model: a_j, b_j and c are path coefficients, $j = 1, \ldots, k$; $\boldsymbol{U}_{1i}$ and U_{2i} are the disturbance variables for the k compositional mediators $\boldsymbol{M}_i$ and the outcome Y_i, respectively. $\sum_{j=1}^{k} M_{ij} = 1$.

and the power transformation for a composition $\boldsymbol{\eta}$ by a scalar α by

$$\boldsymbol{\eta}^{\alpha} = \left(\frac{\eta_1^{\alpha}}{\sum_{j=1}^{k} \eta_j^{\alpha}}, \frac{\eta_2^{\alpha}}{\sum_{j=1}^{k} \eta_j^{\alpha}}, \ldots, \frac{\eta_k^{\alpha}}{\sum_{j=1}^{k} \eta_j^{\alpha}} \right)^{\top}.$$

With these operators, we propose the following compositional mediation model:

$$\boldsymbol{M}_i = \left(\boldsymbol{m}_0 \oplus \boldsymbol{a}^{T_i} \right) \oplus \boldsymbol{U}_{1i} \tag{1}$$

$$Y_i = c_0 + c\, T_i + (\log \boldsymbol{M}_i)^{\top} \boldsymbol{b} + U_{2i}, \quad \text{subject to } \boldsymbol{b}^{\top} \mathbf{1}_k = 0, \tag{2}$$

where $\boldsymbol{m}_0$ is the baseline composition (i.e., when $T_i = \mathbb{E}(T_i)$); similarly, c_0 is the baseline for Y_i; $\boldsymbol{a}, \boldsymbol{b}$ and c are path coefficients; $\mathbf{1}_k$ is a vector of k ones. We do not specify the distribution of $\boldsymbol{U}_{2i}$, but we assume $U_{2i} \sim N(0, \sigma^2)$. The model (1) formulates how a treatment perturbs a composition from the baseline composition, which is measured by the compositional parameter $\boldsymbol{a}$. The model (2) links the treatment and composition to an outcome. In order to account for the compositional nature of $\boldsymbol{M}_i$, a linear constraint, $\boldsymbol{b}^{\top} \mathbf{1}_k = 0$, is imposed, which warrants the subcompositional invariance (Lin *et al.*, 2014; Shi *et al.*, 2016).

Under this model, the expected total effect of T_i on Y_i can be decomposed into the expected direct effect, c, and the expected total indirect effect, $(\log \boldsymbol{a})^{\top} \boldsymbol{b}$, which is the sum of the expected component-wise indirect effects. Note that under the model (1), $\boldsymbol{a}$ is interpreted as the expected change in $\boldsymbol{M}_i$ due to T_i from the identity element $\mathcal{J}_{k-1}$ of $\mathbb{S}^{k-1}$, which is defined by $\mathcal{J}_{k-1} = (1/k, 1/k, \ldots, 1/k)^{\top}$. Thus, to estimate the expected component-wise indirect effects, $\boldsymbol{a}$ should be divided by $\mathcal{J}_{k-1}$, or multiplied by k, so that $\log(k a_j)$ represents the expected change in $\log(M_{ij})$ from 0 by a one unit change in T_i.

2.3 Model Assumptions and Identification

The identification of the expected causal direct and indirect effects requires several assumptions. To define our identifying assumptions, we use the potential outcomes notation described in Section 2.1. Combined with the stable unit treatment value assumption (SUTVA) and the positivity assumption (i.e., $0 < P(T_i = t|\boldsymbol{X}_i = \boldsymbol{x})$ and $0 < P(\boldsymbol{M}_i(t) = \boldsymbol{m}|T_i = t, \boldsymbol{X}_i = \boldsymbol{x})$), the assumptions for the compositional mediation model are given by

$$\{Y_i(t', \log(\boldsymbol{m})), \log \boldsymbol{M}_i(t)\} \perp\!\!\!\perp T_i|\boldsymbol{X}_i = \boldsymbol{x} \tag{3}$$

$$Y_i(t', \log(\boldsymbol{m})) \perp\!\!\!\perp \log \boldsymbol{M}_i(t)|T_i = t, \boldsymbol{X}_i = \boldsymbol{x} \tag{4}$$

$$\text{No interaction between } T_i \text{ and } \boldsymbol{M}_i \tag{5}$$

for $t, t' \in \mathcal{T}$, $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ and $\boldsymbol{m} \in \boldsymbol{\mathcal{M}}$. The SUTVA consists of two components: no interference (i.e., no effect of a treatment applied to one unit on the outcome for other units) and no hidden variation of treatment (i.e., consistent treatment levels) (Imbens and Rubin, 2015). Note that the second component is irrelevant for continuous outcomes. The assumptions (3) - (4) with the positivity assumption are an extension of the sequential ignorability assumptions for the single mediator model (Imai *et al.*, 2010). Basically, they state that there is no unmeasured confounding variable after controlling covariates, $\boldsymbol{X}_i$.

These assumptions appear like those for the multiple mediators model (Imai and Yamamoto, 2013; VanderWeele and Vansteelandt, 2014). However, they are fundamentally very different. The assumption (3) implies that a treatment is conditionally independent of all possible potential outcomes. Thus, the k mediators model requires 2^{k+1} cases of conditional independence. For instance, the two mediators model requires eight cases: $Y_i(t', M_{i1}(t), M_{i2}(t'')) \perp\!\!\!\perp T_i|\boldsymbol{X}_i = \boldsymbol{x}$, where $t'' = t$ or t'. However, for the compositional mediation model, all the components of a composition must be under the same treatment (i.e., $\log \boldsymbol{M}_i(t)$ or $\log \boldsymbol{M}_i(t')$); thus, there are, regardless of k, only four possible cases of conditional independence like the single mediator model. Similarly, for the assumption (4).

Under the assumptions of the compositional mediation model, we have the following theorem to show that the model (1) and (2) lead to quantification of the expected causal direct and total indirect effects.

Theorem 1. (Identification for the compositional mediation model) *Under the assumptions (3) - (5) combined with the SUTVA and the positivity assumptions, the expected causal direct effect $\zeta(t)$ and the expected*

causal total indirect effect $\delta(t)$ for the compositional mediation model are identifiable and given by

$$\begin{aligned}\zeta(\tau) &\equiv \mathbb{E}\left[Y_i(t, \log \boldsymbol{M}_i(\tau)) - Y_i(t_0, \log \boldsymbol{M}_i(\tau))\right] \\ &= c(t - t_0),\end{aligned} \tag{6}$$

$$\begin{aligned}\delta(\tau) &\equiv \mathbb{E}\left[Y_i(\tau, \log \boldsymbol{M}_i(t)) - Y_i(\tau, \log \boldsymbol{M}_i(t_0))\right] \\ &= (\log \boldsymbol{a})^\top \boldsymbol{b}(t - t_0),\end{aligned} \tag{7}$$

where t is the observed treatment for unit i, t_0 a reference value for the treatment and $\tau = t$ or t_0.

A proof of Theorem 1 is given in Appendix A. Pearl (2001) calls $\zeta(\tau)$ the *average natural direct effect* and $\delta(\tau)$ the *average natural indirect effect*. Pearl (2001) also defines the *average controlled direct effect* that is defined in terms of a specific value of the mediator, rather than its potential values. Note that $\zeta(t) = \zeta(t_0)$ and $\delta(t) = \delta(t_0)$ under the no interaction assumption. Thus, the controlled and the natural direct effects coincide and are equal to c for a one unit change in t, and similarly, the natural total indirect effect is $(\log \boldsymbol{a})^\top \boldsymbol{b}$ for a one unit change in t.

3 Parameter Estimation, Variance Estimation and Tests of Mediation Effects

3.1 Estimation of Composition Parameters and Covariance Matrix

With the compositional operators, Billheimer *et al.* (2001) show that $\mathbb{S}^{k-1}$ constitutes a complete inner product space, allowing the definition of a norm for a composition $\boldsymbol{\eta}$:

$$\|\boldsymbol{\eta}\| = \left(\boldsymbol{\eta}^T \boldsymbol{\eta}\right)^{1/2} = \left(\boldsymbol{\theta}^T \boldsymbol{\mathcal{N}}^{-1} \boldsymbol{\theta}\right)^{1/2},$$

where $\boldsymbol{\theta}$ is the additive logratio transformation of $\boldsymbol{\eta}$, that is,

$$\boldsymbol{\theta} = \mathrm{alt}(\boldsymbol{\eta}) = \left(\log \frac{\eta_1}{\eta_k}, \log \frac{\eta_2}{\eta_k}, \ldots, \log \frac{\eta_{k-1}}{\eta_k}\right)^\top,$$

and $\boldsymbol{\mathcal{N}}^{-1}$ is the inverse matrix of a $(k-1) \times (k-1)$ matrix $\boldsymbol{\mathcal{N}}$ defined by

$$\boldsymbol{\mathcal{N}} = \boldsymbol{\mathcal{I}}_{k-1} + \mathbf{1}_{k-1} \mathbf{1}_{k-1}^T,$$

where $\boldsymbol{\mathcal{I}}_{k-1}$ is the $(k-1)\times(k-1)$ identity matrix. Lastly, the inverse of the perturbation operator is defined by

$$\boldsymbol{\eta} \ominus \boldsymbol{\zeta} = \left(\frac{\eta_1 \zeta_1^{-1}}{\sum_{j=1}^k \eta_j \zeta_j^{-1}}, \frac{\eta_2 \zeta_2^{-1}}{\sum_{j=1}^k \eta_j \zeta_j^{-1}}, \ldots, \frac{\eta_k \zeta_k^{-1}}{\sum_{j=1}^k \eta_j \zeta_j^{-1}} \right)^\top .$$

To estimate the parameters in the model (1), we propose the following objective function, which minimizes the compositional norm of the difference between the observed and estimated compositions,

$$\begin{aligned} \hat{\boldsymbol{a}} &= \underset{\boldsymbol{a}, \boldsymbol{m}_0 \in \mathbb{S}^{k-1}}{\operatorname{argmin}} \sum_{i=1}^n \left\| \boldsymbol{M}_i \ominus (\boldsymbol{m}_0 \oplus \boldsymbol{a}^{T_i}) \right\|^2 \\ &= \underset{\boldsymbol{a}, \boldsymbol{m}_0 \in \mathbb{S}^{k-1}}{\operatorname{argmin}} \sum_{i=1}^n \sum_{j=1}^{k-1} \left\{ (k-1) \left[\log \left(\frac{M_{ij} m_{0k} a_k^{T_i}}{M_{ik} m_{0j} a_j^{T_i}} \right) \right]^2 \right. \\ &\qquad \left. - \log \left(\frac{M_{ij} m_{0k} a_k^{T_i}}{M_{ik} m_{0j} a_j^{T_i}} \right) \sum_{\ell \neq j}^{k-1} \log \left(\frac{M_{i\ell} m_{0k} a_k^{T_i}}{M_{ik} m_{0\ell} a_\ell^{T_i}} \right) \right\} . \end{aligned} \tag{8}$$

The objective function (8) is not convex in terms of a_j, but is in terms of $\operatorname{alt}(\boldsymbol{a})_j$ for $j = 1, \ldots, k-1$. Thus, the optimal solution can be obtained by solving the following system of linear equations,

$$\begin{bmatrix} (k-1)f & -f & \ldots & -f \\ -f & (k-1)f & \ldots & -f \\ \vdots & \vdots & \ddots & \vdots \\ -f & -f & \ldots & (k-1)f \end{bmatrix} \begin{bmatrix} \operatorname{alt}(\boldsymbol{a})_1 \\ \operatorname{alt}(\boldsymbol{a})_2 \\ \vdots \\ \operatorname{alt}(\boldsymbol{a})_{k-1} \end{bmatrix} = \begin{bmatrix} g_1 \\ g_2 \\ \vdots \\ g_{k-1} \end{bmatrix}$$

with the constraint $\boldsymbol{a}^\top \mathbf{1}_k = 1$, where $f = \sum_{i=1}^n (T_i - \mathbb{E}(T_i))^2$ and $g_j = k \sum_{i=1}^n (T_i - \mathbb{E}(T_i)) \log M_{ij} - \sum_{i=1}^n \sum_{\ell=1}^k (T_i - \mathbb{E}(T_i)) \log M_{i\ell}$. The matrix in the above equation is always invertible by Sherman-Morrison formula (Sherman and Morrison, 1950) since it can be expressed as the sum of a diagonal matrix of kf and a constant matrix of $-f$. Note that for brevity, we assume no confounding variable and exclude $\boldsymbol{X}_i$. However, the presence of $\boldsymbol{X}_i$ does not fundamentally change the underlying method in estimating the compositional parameters. For instance, if n_x covariates are present, the system of $n_x(k-1)$ linear equations needs to be solved; see Appendix B for details. To estimate its covariance matrix $\hat{\Sigma}_{\boldsymbol{a}}$, we use a bootstrap distribution of $\hat{\boldsymbol{a}}$ obtained by the percentile method of Machado and Parente (2005). See Appendix C for details.

3.2 Estimation of Compositional Regression Parameters and Covariance Matrix

The log-contrast model has been the most general solution to incorporate the unit-sum constraint in the linear regression model for compositional covariates (Aitchison and Bacon-Shone, 1984; Lin *et al.*, 2014). Shi *et al.* (2016) developed a debias procedure for the ℓ_1 regularized estimates of high dimensional compositional

covariates. We use the linear log-contrast model and the debias procedure to estimate regression parameters, $\boldsymbol{b}$ and c. Specifically, we first solve the following objective function,

$$\hat{\boldsymbol{\beta}} = \underset{\beta}{\operatorname{argmin}} \left(\frac{1}{2n} \left\| \boldsymbol{Y} - \boldsymbol{\mathcal{Q}}\boldsymbol{\beta} \right\|_2^2 + \lambda \left\| \boldsymbol{\beta} \right\|_1 \right), \text{ subject to } \sum_{j=1}^{k} \beta_j = 0, \tag{9}$$

where λ is a turning parameter, $\boldsymbol{\beta} = (b_1, \ldots, b_k, c)^\top$ and

$$\boldsymbol{\mathcal{Q}} = (\log \boldsymbol{M}_1, \ldots, \log \boldsymbol{M}_k, \boldsymbol{T}) \left(\boldsymbol{\mathcal{I}}_{k+1} - \boldsymbol{\mathcal{P}} \right),$$

where $\log \boldsymbol{M}_j = (\log M_{1j}, \ldots, \log M_{nj})^\top$ for each component $j = 1, \ldots, k$, $\boldsymbol{T} = (T_i, \ldots, T_n)^\top$ and $\boldsymbol{\mathcal{P}} = (\mathbf{1}_k^\top, 0)^\top (\mathbf{1}_k^\top, 0)/k$. For simplicity, the intercept is excluded in the model, which can be eliminated by centering all the variables in the model. We then apply the debias procedure to the solution of the objective function (9) to obtain unbiased estimates and their covariance matrix.

3.3 Hypothesis Test of Mediation Effect

The null hypothesis of no total compositional mediation effect is given by

$$H_0 : (\log \boldsymbol{a})^\top \boldsymbol{b} = 0, \tag{10}$$

and the null hypothesis of no component-wise mediation effect is given by

$$H_0 : \log(k a_j)\, b_j = 0, \ \forall j \in \{1, 2, \ldots, k\}. \tag{11}$$

The null hypothesis (10) reflects the total mediation effect on an outcome; however, it can disguise the actual mediation effect, which may be captured by the null hypothesis (11). If for instance, the mediation effects of two mediators are equal, but their directions are opposite, then the total mediation effect is zero. In other words, we cannot reject the null hypothesis (10) but might reject the null hypothesis (11). Therefore, we need to test both hypotheses to avoid a misleading conclusion about the mediation effect.

To test the null hypothesis (10) and (11), we propose two approaches: an extension of the Sobel test (Sobel, 1982) and a bootstrap approach. In testing the null hypothesis (10) with the former, the square root of the first order asymptotic variance of a total indirect effect is computed with the estimated covariance matrices of $\log(k\hat{\boldsymbol{a}})$ and $\hat{\boldsymbol{b}}$ by the method described in Bollen (1987), and then it is used as a standard error of the total indirect effect in the Z-test. The expressions for the first order asymptotic variances of a total

indirect effect and component-wise indirect effects are given in Appendix D.

The distribution of a composition $\boldsymbol{M}_i$ is not known, but $\mathrm{alt}(\boldsymbol{M}_i)_j$ is well approximated by a normal distribution (Aitchison, 1986), that is, the distribution of $\log(a_j/a_k)$ is well approximated by a normal distribution. Therefore, $\delta(\tau)$ can also be approximated by a normal distribution assuming the product of two normal variables (i.e., $\log(a_j/a_k)$ and b) follows a normal distribution. Recall that $\sum_{j=1}^{k-1}\log(a_j/a_k)b_j = (\log \boldsymbol{a})^\top \boldsymbol{b}$ because $\sum_{j=1}^{k} b_j = 0$. In general, the product of two normal variables does not follow a normal distribution. However, the misspecified distribution will just reduce the power but not affect the type I error rate when the null hypothesis of no indirect effect is false (MacKinnon *et al.*, 2002; Shrout and Bolger, 2002).

To avoid the assumption of normality for the product of two normal variables, we can use a bootstrap approach (Shrout and Bolger, 2002; VanderWeele and Vansteelandt, 2014). To this end, we use a non-parametric bootstrap for $\log(k\hat{\boldsymbol{a}})$ and a parametric bootstrap for $\hat{\boldsymbol{b}}$ using a multivariate normal distribution to approximate the sampling distribution of $\delta(\tau)$. The p-value for $\delta(\tau)$ is then approximated by utilizing the fact that any bootstrap replicate $\delta(\tau)_b - \delta(\tau)$ should have a distribution close to that of $\delta(\tau)$ when the null hypothesis is true, where $\delta(\tau)_b$ denotes an estimated total indirect effect derived from a resampled dataset (Efron and Tibshirani, 1993). That is, the p-value can be approximated by $2\sum_{b=1}^{n_b}\mathrm{I}(\delta(\tau)_b - \delta(\tau) \geq \delta(\tau))/n_b$ when $\delta(\tau) \geq 0$ and $2\sum_{b=1}^{n_b}\mathrm{I}(\delta(\tau)_b - \delta(\tau) < \delta(\tau))/n_b$ when $\delta(\tau) < 0$, where $\mathrm{I}(\cdot)$ is the indicator function and n_b is a number of bootstrap samples. Similarly, the null hypothesis (11) can be tested. In the simulation, we observed no significant difference in performance between the two methods in terms of power and type I error, concurring with Huang and Pan (2016).

4 Simulation Studies

Mediation analysis for multiple or high dimensional mediators typically assumes independence between mediators to establish a causal interpretation so principal components of mediators are often used. This approach of using principal component is also applicable for compositional mediators to estimate the expected direct and total indirect effects. Another possible approach for compositional mediators is to utilize an ℓ_1 regularization, which has a tendency of dropping correlated variables. We used these two approaches to evaluate the performance of our compositional mediation model. For the hypothesis test, we used the extension of the Sobel test for fair comparison.

In data generation, we randomly generated the treatment T_i from the standard normal distribution and the baseline composition $\boldsymbol{m}_0$ from the standard uniform distribution under the unit-sum constraint. The path coefficients $\boldsymbol{a}$, $\boldsymbol{b}$ and c were selected such that the direct effect is 1.00 and the total indirect effect is approximately 0.00, 0.50, 0.75 or 1.00. For the compositional disturbance, we used a multivariate

logistic normal distribution with mean $\mathbf{0}_{k-1}$ and covariance $2\boldsymbol{\mathcal{N}}$, that is, $\boldsymbol{U}_{1i} \sim \mathrm{LN}(\mathbf{0}_{k-1}, 2\boldsymbol{\mathcal{N}})$, where $\boldsymbol{\mathcal{N}} = \boldsymbol{\mathcal{I}}_{k-1} + \mathbf{1}_{k-1}\mathbf{1}_{k-1}^T$; for the compositional regression disturbance, we used a normal distribution with mean 0 and variance 2, or $U_{2i} \sim \mathrm{N}(0, 2)$. Recall that our model does not specify the distribution of $\boldsymbol{U}_{1i}$. The composition $\boldsymbol{M}_i$ and the outcome Y_i were then generated according to the model (1) and (2), respectively.

We first estimated the power and type I error rate in testing the direct effect (DE) and the total indirect effect (TIDE) for three methods: the principal component regression (PCR), the two stage adaptive lasso (TSAL) and our compositional mediation model (CMM). TSAL, as an ℓ_1 regularization approach, uses the standard lasso in the first stage to screen irrelevant variables and the adaptive lasso in the second stage to select consistent variables (Bühlmann and van de Geer, 2011). We used 1500 simulations for each k mediators with a sample size $n = 100$ at various probability thresholds, where $k = 5, 49, 99$: 250 simulations with each of TIDE $= 1.00, 0.75, 0.50$; 250 simulations with no effect of T_i on $\boldsymbol{M}_i$ (i.e., $a_j = 0, \forall j$); 250 simulations with no effect of $\boldsymbol{M}_i$ on Y_i (i.e., $b_j = 0, \forall j$); 250 simulations with *inconsistent* TIDEs (i.e. the sum of the component-wise indirect effects is zero). For the PCR, we included only the first k_{pc} principal components that explain 90% of the total variance. As shown in Table 1, while all three methods roughly control the type 1 errors, CMM outperforms both the PCR and TSAL approach in power, especially when k is large. We then examined the estimated DE and TIDE at various sample sizes to see the bias and variance of the three methods. Figure 3 shows the results for $k = 49$ based on 100 simulations. The estimates obtained by all three methods converge to the true values, but CMM has the smallest variance.

Table 1: Power and type 1 error rate in testing TIDEs: number of mediators $k = 5, 49, 99$; sample size $n = 100$; significance level $\alpha = 0.001, 0.01, 0.05$. A total of 1500 simulations for each k were used: 750 simulations with non-zero TIDEs and 750 simulations for zero TIDEs. CMM: proposed compositional mediation model; PCR: principal component regression; TSAL: two-stage adaptive lasso.

		Power			Type 1 error		
α		0.001	0.01	0.05	0.001	0.01	0.05
$k = 5$	CMM	0.452	0.623	0.752	0.001	0.011	0.035
	PCR	0.381	0.604	0.732	0.000	0.007	0.040
	TSAL	0.413	0.621	0.740	0.001	0.008	0.037
$k = 49$	CMM	0.476	0.675	0.820	0.001	0.015	0.043
	PCR	0.051	0.207	0.423	0.000	0.005	0.029
	TSAL	0.277	0.487	0.631	0.001	0.012	0.067
$k = 99$	CMM	0.397	0.645	0.791	0.001	0.007	0.040
	PCR	0.023	0.105	0.273	0.000	0.004	0.028
	TSAL	0.304	0.495	0.628	0.003	0.016	0.064

PCR is not capable of testing the component-wise indirect effects (IDEs); therefore, we compared the performance of CMM on the component-wise IDEs only with TSAL. In data generation, we selected the parameters $\boldsymbol{a}$ and $\boldsymbol{b}$ such that the first 7 IDEs are at a nonzero constant level but different combinations of a_j and b_j; the remaining 43 IDEs are at zero. The level of nonzero IDEs was increased by 0.1 from 0.0.

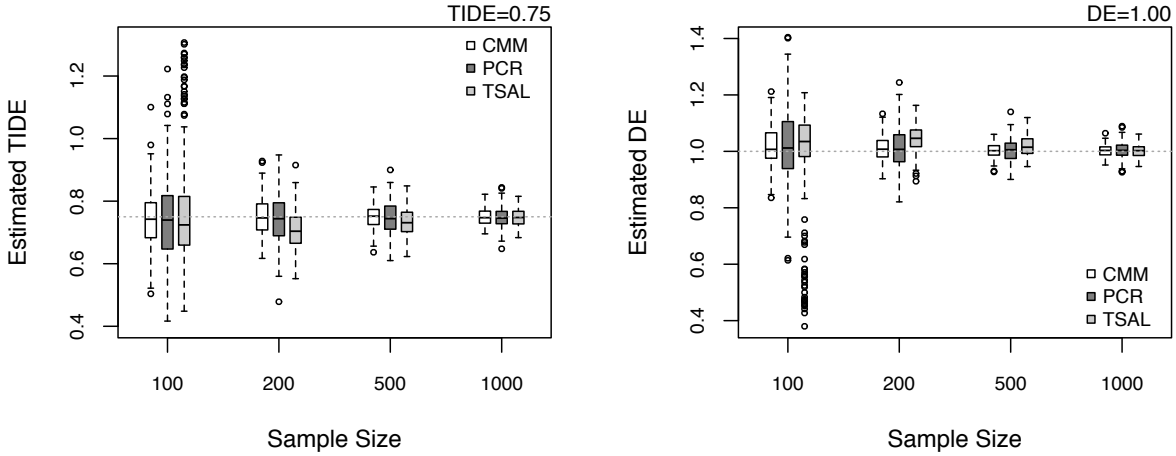

Figure 3: Estimated DEs and TIDEs at various sample sizes for $k = 49$ with the true DE of 1.00 and the true TIDE of 0.75, which are indicated by dotted lines. The results are based on 100 simulations. CMM: proposed compositional mediation model; PCR: principal component regression; TSAL: two-stage adaptive lasso.

We used two sample sizes $n = 100$ and 200. For multiple testing corrections, we used the Benjamini and Yekutieli (2001) false discovery rate (FDR) at 0.05. As a comparison measure, the F_1 score, which is the harmonic mean of precision and recall, was used. Figure 4 shows results: better performance of CMM over TSAL.

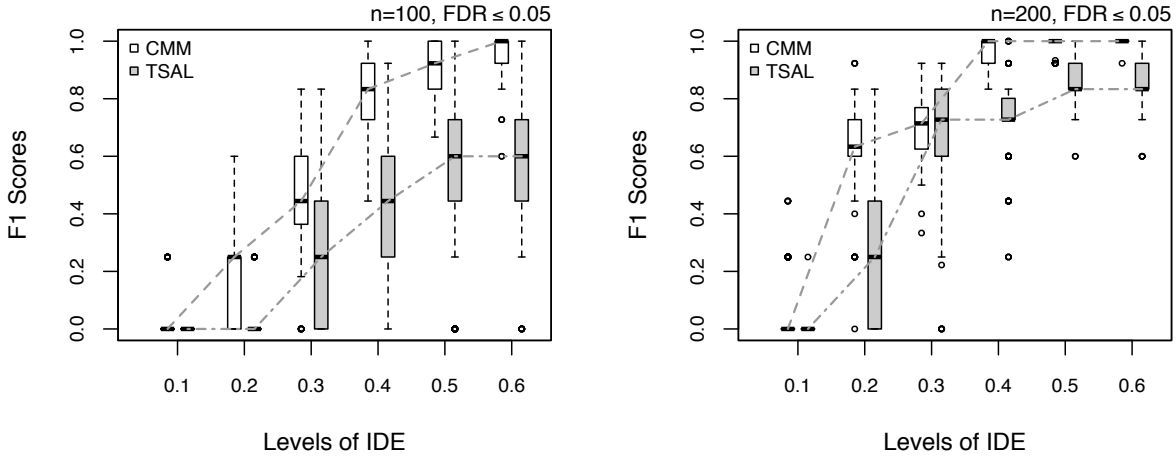

Figure 4: F_1 score versus levels of IDEs at FDR ≤ 0.05. P-values are adjusted by the Benjamini-Yekutieli FDR procedure. CMM: proposed compositional mediation model; TSAL: two-stage adaptive lasso.

5 Real Data Application

We applied CMM to the dataset reported in Wu *et al.* (2011), which is the source of motivation for this study. The dataset consists of 16S rRNA sequences from fecal samples of 98 healthy volunteers from the University of Pennsylvania. It also contains demographic and clinical information including fat intakes and BMI. We summarized OTUs at the genus level and then filtered out the genera that appear less than 10% of the samples, leaving 45 genera in 96 samples. Due to different total counts throughout the samples, the OTU counts assigned to these genera were transformed into proportions after replacing zero counts by the maximum rounding error 0.5, which is commonly used in compositional data analysis (Aitchison, 1986).

The gut microbiota can influence host adiposity through energy extraction from the diet, with variable efficiency depending on community composition; furthermore, the microbiota can also affect host adiposity by influencing metabolism throughout the body. It is therefore highly likely that gut microbiome can potentially mediate the effect of diets such as fat on host adiposity and BMI. Since the 98 samples were roughly randomly sampled, it was reasonable to assume that fat was randomly sampled. CMM was applied to the dataset with BMI as the outcome, fat intake as the treatment, and the 45 genera as the compositional mediators. The estimated DE and TIDE are reported in Table 2 with their 95% confidence intervals (CIs). The estimated component-wise IDEs are shown in Figure 5.

Table 2: The estimated DE and TIDE of fat intakes on BMI through the gut microbiome using the proposed compositional mediation model (CMM). The values in parenthesis represent 95% CIs. The second row presents the percentages of DE and TIDE in the total effect of fat intakes on BMI.

DE	TIDE
0.949 (0.003, 1.901)	0.732 (-0.331, 2.114)
56.5%	43.5%

The estimated DE and TIDE are 0.949 and 0.732, respectively, which correspond to 56.5% and 43.5% of the total effect of fat intakes on BMI. Note that under LSEM the total effect of treatment can be decomposed into direct and indirect effects. The estimated DE is statistically significant at the significance level of 0.05, but the estimated TIDE, as well as all component-wise IDEs, are not. It is most likely due to an insufficient sample size. In the simulation, we observed that TIDE of 0.75 has a significant probability of its CI containing zero, and component-wise IDEs at around 0.2 can be rarely detected with a sample size of 100.

As shown in Figure 5, the potential genera that would have statistically significant non-zero component-wise IDEs with a sufficient sample size are *Alistipes, Oscillibacter, Acidaminococcus*, and *Allisonella*. All these genera have positive values for IDEs, but their responses to fat intakes are quite different. The

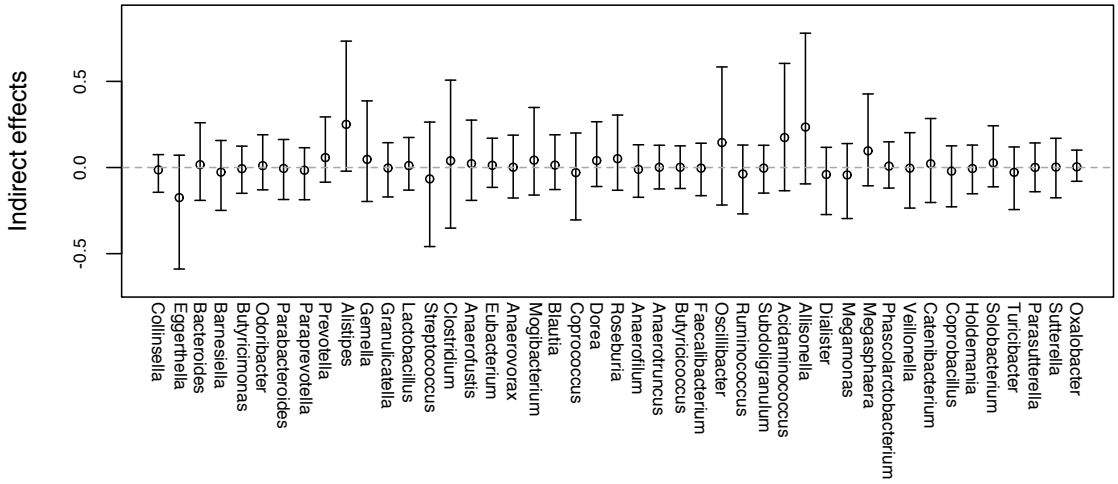

Figure 5: The estimated component-wise IDEs of fat intakes on BMI through the gut microbiome.

abundance of *Alistipes* and *Oscillibacter* is negatively correlated with fat intake and BMI, whereas that of *Acidaminococcus* and *Allisonella* is positively correlated with fat intake and BMI. Based on this information, we can hypothesize that the increase in fat intake causes the decrease in the abundance of *Alistipes* and *Oscillibacter* and the increase in the abundance of *Acidaminococcus* and *Allisonella* which in turn cause the increase in BMI. Lam *et al.* (2012) identified Oscillibacter-like organisms as a potentially important gut microbe that mediates high fat-induced gut dysfunction and gut permeability and showed that decrease of *Oscillibacter* led to increased gut permeability, which was shown to be associated with obesity (Teixeira *et al.*, 2012). These observations were largely consistent with what were observed in mice fed with high-fat diet (Daniel *et al.*, 2014).

6 Discussion

In this study, we propose a compositional mediation model for a continuous outcome. Our method takes the characteristics of compositional data into account and treats the whole compositional mediators as a unit, that is, it estimates the effect of treatment on compositional mediators simultaneously instead of each mediator separately. In our simulation studies, we have shown better performance of our method over the two potential methods for compositional mediators. Our method also provides a clear interpretation of component-wise indirect effects. Its application to gut microbiome and BMI dataset has clearly indicated a potential mediation effect of the gut microbiome in linking the association between fat intake and BMI. Although the results were not statistically significant, several bacterial genera have been shown to be directly associated with gut permeability and therefore BMI. It would be interesting and important to replicate these

results in larger datasets.

In many clinical microbiome studies, the outcome variable is binary such as whether a subject is diseased or not. In these cases, Model (2) can be rewritten for logistic or probit regression, assuming the outcome variable is a latent continuous variable indicated by an observed dichotomous variable. Then, Model (1) and the modified Model (2) will provide the identifiability of direct and indirect effects (Winship and Mare, 1983). Another interesting extension of our method is for longitudinal compositional data, which is also very common in microbial studies.

Even though we used only a continuous treatment variable in the simulation study, a binary treatment variable can be used without any modification in our method. However, a treatment variable with more than two categories cannot be used since it requires $k-1$ parameters to fully represent the effect of a categorical variable with k mutually exclusive categories on a dependent variable. A compositional mediation model with a general categorical treatment variable is another interesting extension of our method.

Appendix A Proof of Theorem 1

Rewrite equations (1) and (2) using the potential outcome notation:

$$\boldsymbol{M}_i(T_i) = \left(\boldsymbol{m}_0 \oplus \boldsymbol{a}^{T_i}\right) \oplus \boldsymbol{U}_{1i}(T_i) \tag{12}$$

$$Y_i(T_i, \log \boldsymbol{M}_i(T_i)) = c_0 + c\,T_i + (\log \boldsymbol{M}_i(T_i))^\top \boldsymbol{b} + U_i(T_i, \log \boldsymbol{M}_i(T_i)), \tag{13}$$

where $\boldsymbol{b}^\top \mathbf{1}_k = 0$. Note the equation (12) implicitly assumes $\mathbb{E}[\mathrm{alt}(\boldsymbol{U}_{1i}(\tau))] = \mathbf{0}_{k-1}$ for $\tau \in \mathcal{T}$, which assures consistency of our estimators, particularly $\hat{\boldsymbol{a}}$; as a model assumption for the equation (13), $\mathbb{E}[U_i(\tau, \log \boldsymbol{m})] = 0$ for $\tau \in \mathcal{T}$ and $\boldsymbol{m} \in \boldsymbol{\mathcal{M}}$. The assumption (3) implies $\log(\boldsymbol{U}_{1i}(\tau)) \perp\!\!\!\perp T_i$, which entails $\mathrm{alt}(\boldsymbol{U}_{1i}(\tau)) \perp\!\!\!\perp T_i$, thus yielding $\mathbb{E}[\mathrm{alt}(\boldsymbol{U}_{1i}(T_i))|T_i = \tau] = \mathbb{E}[\mathrm{alt}(\boldsymbol{U}_{1i}(\tau))] = 0$ for any τ. Therefore,

$$\mathbb{E}[\mathrm{alt}(\boldsymbol{M}_i)|T_i] = \mathrm{alt}(\boldsymbol{m}_0) + T_i\,\mathrm{alt}(\boldsymbol{a}). \tag{14}$$

Similarly, the assumption (4) implies $U_i(\tau, \log \boldsymbol{m}) \perp\!\!\!\perp \log \boldsymbol{M}_i | T_i = \tau$ for all $\tau \in \mathcal{T}$ and $\boldsymbol{m} \in \boldsymbol{\mathcal{M}}$. Thus, $\mathbb{E}[U_i(T_i, \log \boldsymbol{M}_i(T_i))|T_i = \tau, \log \boldsymbol{M}_i = \log \boldsymbol{m}] = \mathbb{E}[U_i(T_i, \log \boldsymbol{m})|T_i = \tau] = \mathbb{E}[U_i(\tau, \log \boldsymbol{m})] = 0$, where the second equality follows from the assumption (3). Thus,

$$\begin{aligned} \mathbb{E}[Y_i | \log \boldsymbol{M}_i, T_i] &= c_0 + c\,T_i + (\log \boldsymbol{M}_i)^\top \boldsymbol{b} \\ &= c_0 + c\,T_i + (\mathrm{alt}(\boldsymbol{M}_i))^\top \boldsymbol{b}_{-k}, \end{aligned} \tag{15}$$

where $\boldsymbol{b}_{-k} = (b_1, b_2, \ldots, b_{k-1})^\top$. The second equality holds because of the constraint, $\boldsymbol{b}^\top \mathbf{1}_k = 0$. Finally, under the causal assumption for the compositional mediation model, the expected causal direct effect $\zeta(\tau)$ is

$$
\begin{aligned}
\zeta(\tau) &= \mathbb{E}\left[Y_i(t, \log \boldsymbol{M}_i(\tau)) - Y_i(t_0, \log \boldsymbol{M}_i(\tau))\right] \\
&= \int \cdots \int_k \Big[\mathbb{E}(Y_i(t, \log \boldsymbol{m}) | \log \boldsymbol{M}_i(\tau) = \log \boldsymbol{m}) \\
&\qquad - \mathbb{E}(Y_i(t_0, \log \boldsymbol{m}) | \log \boldsymbol{M}_i(\tau) = \log \boldsymbol{m})\Big] dF_{\log \boldsymbol{M}_i(\tau)}(\log \boldsymbol{m}) \\
&= \int \cdots \int_k \Big[\mathbb{E}(Y_i(t, \log \boldsymbol{m}) | \log \boldsymbol{M}_i(\tau) = \log \boldsymbol{m}, T_i = \tau) \\
&\qquad - \mathbb{E}(Y_i(t_0, \log \boldsymbol{m}) | \log \boldsymbol{M}_i(\tau) = \log \boldsymbol{m}, T_i = \tau)\Big] dF_{\log \boldsymbol{M}_i(\tau)}(\log \boldsymbol{m}) \\
&= \int \cdots \int_k \Big[\mathbb{E}(Y_i(t, \log \boldsymbol{m}) | T_i = \tau) \\
&\qquad - \mathbb{E}(Y_i(t_0, \log \boldsymbol{m}) | T_i = \tau)\Big] dF_{\log \boldsymbol{M}_i(\tau)}(\log \boldsymbol{m}) \\
&= \int \cdots \int_k \Big[\mathbb{E}(Y_i(t, \log \boldsymbol{m}) | T_i = t) \\
&\qquad - \mathbb{E}(Y_i(t_0, \log \boldsymbol{m}) | T_i = t_0)\Big] dF_{\log \boldsymbol{M}_i | T_i = \tau}(\log \boldsymbol{m}) \\
&= \int \cdots \int_k \Big[\mathbb{E}(Y_i(t, \log \boldsymbol{m}) | \log \boldsymbol{M}_i(t) = \log \boldsymbol{m}, T_i = t) \\
&\quad - \mathbb{E}(Y_i(t_0, \log \boldsymbol{m}) | \log \boldsymbol{M}_i(t_0) = \log \boldsymbol{m}, T_i = t_0)\Big] dF_{\log \boldsymbol{M}_i | T_i = \tau}(\log \boldsymbol{m}) \\
&= \int \cdots \int_k \Big[\mathbb{E}(Y_i | \log \boldsymbol{M}_i = \log \boldsymbol{m}, T_i = t) \\
&\qquad - \mathbb{E}(Y_i | \log \boldsymbol{M}_i = \log \boldsymbol{m}, T_i = t_0)\Big] dF_{\log \boldsymbol{M}_i | T_i = \tau}(\log \boldsymbol{m}) \\
&= \int \cdots \int_{k-1} \Big[\mathbb{E}(Y_i | \mathrm{alt}(\boldsymbol{M}_i) = \mathrm{alt}(\boldsymbol{m}), T_i = t) \\
&\qquad - \mathbb{E}(Y_i | \mathrm{alt}(\boldsymbol{M}_i) = \mathrm{alt}(\boldsymbol{m}), T_i = t_0)\Big] dF_{\mathrm{alt}(\boldsymbol{M}_i) | T_i = \tau}(\mathrm{alt}(\boldsymbol{m})) \\
&= \int \cdots \int_{k-1} c\,(t - t_0) dF_{\mathrm{alt}(\boldsymbol{M}_i) | T_i = \tau}(\mathrm{alt}(\boldsymbol{m})) \\
&= c\,(t - t_0),
\end{aligned}
$$

where $F_V()$ and $F_{V|W}()$ represent the distribution function of a random variable V and the conditional distribution function of V given W. The third and fifth equalities hold due to the assumption (3); the fourth and sixth due to the assumption (4); the eighth and ninth because of the equation (15). Similarly, the causal

total indirect effect $\delta(\tau)$ is

$$
\begin{aligned}
\delta(\tau) &= \mathbb{E}\left[Y_i(\tau, \log \boldsymbol{M}_i(t)) - Y_i(\tau, \log \boldsymbol{M}_i(t_0))\right] \\
&= \int \cdots \int_{k-1} \mathbb{E}(Y_i | \mathrm{alt}(\boldsymbol{M}_i) = \mathrm{alt}(\boldsymbol{m}), T_i = \tau) \\
&\qquad \left[dF_{\mathrm{alt}(\boldsymbol{M}_i)|T_i=t}(\mathrm{alt}(\boldsymbol{m})) - dF_{\mathrm{alt}(\boldsymbol{M}_i)|T_i=t_0}(\mathrm{alt}(\boldsymbol{m}))\right] \\
&= \int \cdots \int_{k-1} (c_0 + c\,\tau + (\mathrm{alt}(\boldsymbol{m}))^\top \boldsymbol{b}_{-k}) \\
&\qquad \left[dF_{\mathrm{alt}(\boldsymbol{M}_i)|T_i=t}(\mathrm{alt}(\boldsymbol{m})) - dF_{\mathrm{alt}(\boldsymbol{M}_i)|T_i=t_0}(\mathrm{alt}(\boldsymbol{m}))\right] \\
&= (\mathrm{alt}(\boldsymbol{a}))^\top \boldsymbol{b}_{-k}(t - t_0) \\
&= (\log \boldsymbol{a})^\top \boldsymbol{b}\,(t - t_0).
\end{aligned}
$$

The fourth equality holds because of the equation (14) and the last one due to $\boldsymbol{b}^\top \mathbf{1}_k = 0$.

Appendix B Model with Covariates

Let X_i be a covariate and $\boldsymbol{\psi}$ be its compositional parameters. Then, the parameters $\boldsymbol{a}$ and $\boldsymbol{\psi}$ can be estimated by solving

$$
\begin{aligned}
(\hat{\boldsymbol{a}}, \hat{\boldsymbol{\psi}}) &= \underset{\boldsymbol{a}, \boldsymbol{\psi}, \boldsymbol{m}_0 \in \mathbb{S}^{k-1}}{\operatorname{argmin}} \sum_{i=1}^{n} \left\| M_i \ominus (\boldsymbol{m}_0 \oplus \boldsymbol{a}^{T_i} \oplus \boldsymbol{\psi}^{X_i}) \right\|^2 \\
&= \underset{\boldsymbol{a}, \boldsymbol{\psi}, \boldsymbol{m}_0 \in \mathbb{S}^{k-1}}{\operatorname{argmin}} \sum_{i=1}^{n} \sum_{j=1}^{k-1} \left\{ (k-1) \left[\log \left(\frac{M_{ij} m_{0k} a_k^{T_i} \psi_k^{X_i}}{M_{ik} m_{0j} a_j^{T_i} \psi_j^{X_i}} \right) \right]^2 \right. \\
&\qquad \left. - \log \left(\frac{M_{ij} m_{0k} a_k^{T_i} \psi_k^{X_i}}{M_{ik} m_{0j} a_j^{T_i} \psi_j^{X_i}} \right) \sum_{\ell \neq j}^{k-1} \log \left(\frac{M_{i\ell} m_{0k} a_k^{T_i} \psi_k^{X_i}}{M_{ik} m_{0\ell} a_\ell^{T_i} \psi_\ell^{X_i}} \right) \right\},
\end{aligned}
$$

or solving the system of $2(k-1)$ linear equations:

$$
\begin{bmatrix} D(T^{*2}) & D(X^* T^*) \\ D(X^* T^*) & D(X^{*2}) \end{bmatrix} \begin{bmatrix} \mathrm{alt}(\boldsymbol{a}) \\ \mathrm{alt}(\boldsymbol{\psi}) \end{bmatrix} = \begin{bmatrix} m(T^*) \\ m(X^*) \end{bmatrix}
$$

with a constraint $\boldsymbol{a}^\top \mathbf{1}_k = 1$ and $\boldsymbol{\psi}^\top \mathbf{1}_k = 1$, where $T^* = T - \mathbb{E}(T)$; $X^* = X - \mathbb{E}(X)$; $m(\xi)_j = k \sum_{i=1}^n \xi_i \log M_{ij} - \sum_{i=1}^n \sum_{\ell=1}^k \xi_i \log M_{i\ell},\ \forall j \in \{1, \dots k-1\}$;

$$D(\xi) = \begin{bmatrix} (k-1)\sum_{i=1}^n \xi_i & -\sum_{i=1}^n \xi_i & \dots & -\sum_{i=1}^n \xi_i \\ -\sum_{i=1}^n \xi_i & (k-1)\sum_{i=1}^n \xi_i & \dots & -\sum_{i=1}^n \xi_i \\ \vdots & \vdots & \ddots & \vdots \\ -\sum_{i=1}^n \xi_i & -\sum_{i=1}^n \xi_i & \dots & (k-1)\sum_{i=1}^n \xi_i \end{bmatrix}.$$

Similarly, we can estimate the compositional parameters for n_x covariates by solving the system of $(k-1)(n_x+1)$ linear equations.

Appendix C Covariance Matrix of Composition Parameters $\hat{\boldsymbol{a}}$

In Model (1), we do not specify the distribution of $\boldsymbol{U}_i$. Therefore, we use a bootstrap distribution of $\hat{\boldsymbol{a}}$ to estimate its covariance matrix $\hat{\Sigma}_{\boldsymbol{a}}$, specifically using the percentile method of Machado and Parente (2005). We define a matrix $\mathcal{H}$ as

$$\mathcal{H} = \left[\begin{array}{c|c} \mathcal{I}_k & \mathcal{O} \\ \hline \mathcal{C} & \mathcal{D}_{-2} \end{array}\right],$$

where $\mathcal{I}_k$ is a $k \times k$ identity matrix, $\mathcal{O}$ is a $k \times k(k-1)/2$ matrix of zeros, $\mathcal{C}$ is a $k(k-1)/2 \times k$ binary matrix whose rows are all possible combinations of 2 ones and $k-2$ zeros, and $\mathcal{D}_{-2}$ is a $k(k-1)/2 \times k(k-1)/2$ scalar matrix with -2. Then, $\hat{\Sigma}_{\boldsymbol{a}}$ can be obtained by solving the system of linear equations,

$$\mathcal{H}\boldsymbol{u} = \boldsymbol{v},$$

where $\boldsymbol{u} = (\hat{\sigma}_{11}, \dots, \hat{\sigma}_{kk}, \hat{\sigma}_{12}, \dots, \hat{\sigma}_{(k-1)k})^\top$ is a vector of all the unique elements in $\hat{\Sigma}_{\boldsymbol{a}}$ and $\boldsymbol{v} = (s^2_{h_1}, \dots, s^2_{h_{k(k+1)/2}})^T$ is a vector of the bootstrap estimators of $\hat{\sigma}_h$. For instance, from the first, second, and $(k+1)$th rows, we can obtain $\hat{\sigma}_{11} = s^2_{h_1}$, $\hat{\sigma}_{22} = s^2_{h_2}$, and $\hat{\sigma}_{11} + \hat{\sigma}_{22} - 2\hat{\sigma}_{12} = s^2_{h_{k+1}}$ or $\hat{\sigma}_{12} = (s^2_{h_1} + s^2_{h_2} - s^2_{h_{k+1}})/2$. For the estimators of $\hat{\sigma}_H$, we use functionals censored at the top and bottom 5% of the bootstrap distribution (Machado and Parente, 2005),

$$s_{h_i} = \frac{1}{0.847 n_b} \sum_{j=s}^{\ell} \Phi^{-1}(j/n_b) q^*_{h_i}(j/n_b) + 0.103 \left[q^*_{h_i}(\ell/n_b) - q^*_{h_i}(s/n_b)\right]$$

with $s = \lfloor 0.05 n_b \rfloor$ and $\ell = n_b - \lfloor 0.05 n_b \rfloor$, where $\lfloor x \rfloor$ denotes the integer part of x, n_b is a bootstrap sample size, $\Phi^{-1}(\alpha)$ is the α-percentile of the standard normal distribution, and $q^*_{h_i}(\alpha)$ is the α-percentile of the

bootstrap distribution of $\left\{\sqrt{n}(\boldsymbol{h}_i^\top \hat{\boldsymbol{a}}_b^* - \boldsymbol{h}_i^\top \hat{\boldsymbol{a}}), b = 1, \ldots, n_b\right\}$. $\hat{\boldsymbol{a}}_b^*$ is a vector of bootstrap estimations and $\boldsymbol{h}_i$ is a vector satisfying $\boldsymbol{h}_i^\top \hat{\Sigma}_{\boldsymbol{a}} \boldsymbol{h}_i = \hat{\sigma}_{h_i}$. For the $(k+1)$th row, as an example, $\boldsymbol{h}_{k+1} = (1, -1, 0, \ldots, 0)^\top$.

Appendix D Variance of Indirect Effects

To approximate the variances of the total indirect effect and component-wise indirect effects, the first order asymptotic method described in Bollen (1987) is used. That is, the variance of each component-wise indirect effect is approximated by

$$\mathrm{Var}(\log(k\hat{a}_j)\hat{b}_j) \approx [\mathbb{E}(\log(k\hat{a}_j))]^2 \mathrm{Var}(\hat{b}_j) + [\mathbb{E}(\hat{b}_j)]^2 \mathrm{Var}(\log(k\hat{a}_j))$$

and the variance of the total indirect effect is approximated by

$$\begin{aligned}\mathrm{Var}((\log \hat{a})^T \hat{b}) \approx & \sum_{j=1}^{k} [\mathbb{E}(\log \hat{a}_j)]^2 \mathrm{Var}(\hat{b}_j) + \sum_{j=1}^{k} [\mathbb{E}(\hat{b}_j)]^2 \mathrm{Var}(\log \hat{a}_j) \\ & + 2 \sum_{j<\ell}^{k} \mathrm{Cov}(\log \hat{a}_j, \log \hat{a}_\ell) \mathbb{E}(\hat{b}_j) \mathbb{E}(\hat{b}_\ell) \\ & + 2 \sum_{j<\ell}^{k} \mathrm{Cov}(\hat{b}_j, \hat{b}_\ell) \mathbb{E}(\log \hat{a}_j) \mathbb{E}(\log \hat{a}_\ell).\end{aligned}$$

References

Aitchison, J. (1982). The statistical analysis of compositional data. *J. R. Stat. Soc., Ser. B (Stat. Methodol.)*, **44**(2),139-177.

Aitchison, J. and Bacon-Shone, J. (1984). Log contrast models for experiments with mixtures. *Biometrika*, **71**(2), 323-330.

Aitchison, J. (1986). *The Statistical Analysis of Compositional Data*. New York: Chapman & Hall.

Benjamini, Y. and Yekutieli, D. (2001). The Control of the False Discovery Rate in Multiple Testing Under Dependency. *The Annals of Statistics*, **29**(4), 1165-1188.

Billheimer, D., Guttorp, P. and Fagan, W. F. (2001). Statistical Interpretation of Species Composition. *Journal of the American Statistical Association*, **96**(456), 1205-1214.

Bollen, K. A. (1987). Total, direct, and indirect effects in structural equation models. *Sociological Methodology*. **17**, 37-69.

Bray, G. A. and Popkin, B. M. (1998). Dietary fat intake does affect obesity! *Am J Clin Nutr*. **68**(6), 1157-1173.

Bühlmann, P. and van de Geer, S. (2011). *Statistics for High-Dimensional Data: Method, Theory and Applications*. Berlin: Springer.

Chén, O. Y., Crainiceanu, C. M., Ogburn, E. L., Caffo, B. S., Wager, T. D. and Lindquist, M. A. (2015). High-dimensional Multivariate Mediation with Application to Neuroimaging Data. arXiv:1511.09354

Daniel, H., Gholami, A. M., Berry, D., Desmarchelier, C., Hahne, H., Loh, G., Mondot, S., Lepage, P., Rothballer, M., Walker, A., Böhm C., Wenning, M., Wagner, M., Blaut, M., Schmitt-Kopplin, P., Kuster, B., Haller, D. and Clavel, T. (2014). High-fat diet alters gut microbiota physiology in mice. *ISEM J*, **8**(2), 295-308.

Efron, B. and Tibshirani, R. (1993). *An Introduction to the Bootstrap*. Chapman & Hall.

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

Lam, Y. Y., Ha, C. W., Campbell, C. R., Mitchell, A. J., Dinudom, A., Oscarsson, J., Cook, D. I., Hunt, N. H., Caterson, I. D., Holmes, A. J. and Storlien, L. H. (2012). Increased Gut Permeability and Microbiota Change Associate with Mesenteric Fat Inflammation and Metabolic Dysfunction in Diet-Induced Obese Mice. *PLoS ONE*, **7**(3): e34233.

Imai, K., Keele, L. and Yamamoto, T. (2010). Identification, Inference and Sensitivity Analysis for Causal Mediation Effects. *Statistical Science*, **25**(1), 51-71.

Imai, K., Keele, L. and Tingley, D. (2010). A General Approach to Causal Mediation Analysis. *Psychol Methods*, **15**(4), 309-334.

Imai, K. and Yamamoto, T. (2013). Identification and Sensitivity Analysis for Multiple Causal Mechanisms: Revisiting Evidence from Framing Experiments. *Political Analysis*, **21**(2), 141-171.

Imbens, G. and Rubin, D. (2015). *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge: Cambridge University Press.

Machado, J. A. F. and Parente, P. (2005). Bootstrap estimation of covariance matrices via the percentile method. *The Econometrics Journal*, **8**(1), 70-78.

MacKinnon, D. P., Lockwood, C. M., Hoffman, J. M., West, S. G. and Sheets, V. (2002). A comparison of methods to test mediation and other intervening variable effects. *Psychological Methods*, **7**(1), 83-104.

Ley, R. E., Turnbaugh, P. J., Klein, S. and Gordon, J. I. (2006). Human gut microbes associated with obesity. *Nature*, **444**, 1022-1023.

Lin, W., Shi, P., Feng, R. and Li, H. (2014). Variable selection in regression with compositional covariates. *Biometrika*, **101**(4), 785-797.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*. New York: Cambridge University Press.

Pearl, J. (2001). Direct and indirect effects. *In Proceedings of the Seventeenth Conference on Uncertainty and Artificial Intelligence*, 411-420. San Francisco, CA: Morgan Kaufmann.

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

Robins, J. M. (2003). Semantics of causal DAG models and the identification of direct and indirect effects. In *Highly Structured Stochastic Systems* (P. J. Green, N. L. Hjort and S. Richardson, eds.) 70-81. Oxford: Oxford Univ. Press.

Rubin D. B. (2005). Causal inference using potential outcomes. *Journal of the American Statistical Association*. **100**(469), 322-331.

Sherman, J. and Morrison, W. J. (1950). Adjustment of an Inverse Matrix Corresponding to a Change in One Element of a Given Matrix. *Annals of Mathematical Statistics*. **21**(1), 124-127.

Shi, P., Zhang, A. and Li, H. (2016). Regression Analysis for Microbiome Compositional Data. *Annals of Applied Statistics*, **10(2)**, 1019-1040.

Shrout, P. E. and Bolger, N. (2002). Mediation in Experimental and Nonexperimental Studies: New Procedures and Recommendations. *Psychological Methods*, **7**(4), 422-445.

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

Teixeira, T. F., Collado, M. C., Ferreira, C. L., Bressan, J. and Peluzio, M. C. (2012). Potential mechanisms for the emerging link between obesity and increased intestinal permeability. *Nutr Res.*, **32**(9), 637-47.

Turnbaugh, P. J., Ley, R. E., Mahowald, M. A., Magrini, V., Mardis, E. R. and Gordon, J. I. (2006). An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature*, **444**, 1027-1031.

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

VanderWeele, T. J. and Vansteelandt, S. (2014). Mediation Analysis with Multiple Mediators. *Epidemiol Method*. **2**(1): 95-115.

Winship, C. and Mare, R. D. (1983). Structural Equations and Path Analysis for Discrete Data. *The American Journal of Sociology*. **89**(1):54-110.

Wu, G., Chen, J., Hoffmann, C., Bittinger, K., Chen, Y. Y., Keilbaugh, S. A., Bewtra, M., Knights, D., Walters, W. A., Knight, R., Sinha, R., Gilroy, E., Gupta, K., Baldassano, R., Nessel, L., Li, H., Bushman, F. D. and Lewis J. D. (2011). Linking long-term dietary patterns with gut microbial enterotypes. *Science*. **334**(6052) 105-108.

Zhao, Y. and Luo, X. (2016). Pathway Lasso: Estimate and Select Sparse Mediation Pathways with High Dimensional Mediators. arXiv:1603.07749.