Causal mediation analysis for time-varying heritable risk factors with Mendelian Randomization

Zixuan Wu[1], Ethan Lewis[1], Qingyuan Zhao[2], Jingshu Wang[1,*]

Abstract

Understanding the causal pathogenic mechanisms of diseases is crucial in clinical research. When randomized controlled experiments are not available, Mendelian Randomization (MR) offers an alternative, leveraging genetic mutations as a natural "experiment" to mitigate environmental confoundings. However, most MR analyses treat the risk factors as static variables, potentially oversimplifying dynamic risk factor effects. The framework of life-course MR has been introduced to address this issue. However, current methods face challenges especially when the age-specific GWAS datasets have limited cohort sizes and there are substantial correlations between time points for a single trait. This study proposes a novel approach, estimating a unified system of structural equations for a sequence of temporally ordered heritable traits, requiring only GWAS summary statistics. The method facilitates statistical inference on direct, indirect, and path-wise causal effects and demonstrates superior efficiency and reliability, particularly with noisy GWAS data. By incorporating a spike-and-slab prior for genetic effects, the approach can address extreme polygenicity and weak instrument bias. Through this methodology, we uncovered a protective effect of BMI on breast cancer during a confined period of childhood development. We also analyzed how BMI, systolic blood pressure (SBP), and low-density cholesterol levels influence stroke risk across childhood and adulthood, and identified the intriguing relationships between these risk factors.

[1]Department of Statistics, The University of Chicago, Chicago, IL 60615, USA
[2]Statistical Laboratory, Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, CB2 0WB, UK

*Corresponding author, jingshuw@uchicago.edu

1 Introduction

Understanding the pathogenic mechanisms of diseases is a foundational challenge in clinical research. Given the limitations of conducting randomized controlled experiments in certain cases, there has been a growing reliance on Mendelian Randomization (MR) as an alternative approach. MR leverages genetic mutations and inheritance as a natural "experiment," effectively mitigating unmeasured environmental confounding in epidemiological studies [1, 2]. However, most MR analyses use a cross-sectional design and treat the risk factor as a static variable, ignoring the fact that inheritable risk factors often change over time and may have a time-varying effect on diseases [3, 4]. As shown in earlier studies, this can lead to oversimplified and even misleading conclusions from MR studies [5]. For instance, observational studies have shown that vitamin D levels in childhood, but not adulthood, are associated with the risk of multiple sclerosis. However, standard MR approaches based on the vitamin D level in adult individuals suggest its causal effect in the etiology of multiple sclerosis [6].

To address this challenge, life-course MR has recently emerged as a framework to consider how risk factors that are measured throughout an individual's lifetime may influence later life outcomes [7]. One commonly employed approach is to apply multivariable MR (MVMR) methods such as IVW-MVMR [8] to estimate the *direct* causal effects of the risk factor at each time point under linear and additive causal effect models. However, the efficiency of multivariable MR may be compromised due to the limited cohort size in Genome-wide Association Studies (GWAS) of earlier life traits and high auto-correlations of the risk factor across time. It is also challenging to evaluate the indirect causal effect of the risk factor at earlier time points mediated by later time points using multivariable MR. Alternative approaches involve g-estimations of structural mean models [4, 9], or functional principal component analysis to aggregate the effects of a risk factor across time points [10]. However, both methods require access to individual-level GWAS data that is typically not readily available on a large scale. Additionally, the method from [10] is unable to distinguish between the direct and indirect effects across different time points.

Using genetic variations as instrumental variables throughout the life course, we propose a new approach to estimate a unified system of structural equations for a sequence of heritable traits that are ordered temporally, only requiring GWAS summary statistics for each trait. Our estimated model gives a "full picture" of the time-varying risk factors and enables causal mediation analysis by allowing researchers to quantify various aspects of causal effects, including direct effects, indirect effects, and the proportion of each path-wise effect relative to the total effect. While our model shares similarities with simultaneous equations models in econometrics (Davidson, 1993), we can deal with weak and invalid instruments and allow arbitrary interactions among genetic variants and environmental factors. Furthermore, our model can simultaneously analyze multiple traits at any given time point, accounting for known confounding exposures.

In comparison to alternative approaches, our method shows superior efficiency and reliability, particularly when dealing with noisy GWAS data from small cohorts for age-specific traits. By incorporating a spike-and-slab prior for genetic effects, we account for the extreme polygenicity of complex traits and avoid weak instrument bias, a common issue in MR. Additionally, we discussed sufficient conditions on the identifiability of any direct and indirect causal effects, accommodating pervasive horizontal pleiotropy and arbitrary mediator-outcome confoundings. Applying our new method, we uncover a protective effect of Body Mass Index (BMI) on breast cancer which is confined to a specific period during

childhood development. The method also allows us to unravel the intricate relationships among BMI, systolic blood pressure (SBP), and low-density cholesterol levels across both adulthood and childhood, and their effects on stroke. Our analyses suggest that, among these traits, adulthood SBP may be the only factor with a direct causal effect on stroke. Furthermore, the previously identified causal effect of BMI on SBP in adulthood [11, 12] might be explained by the confounding effect of childhood SBP.

2 Material and methods

2.1 Structural equations on individual-level data

Consider a sequence of risk factors $(X_1, X_2, \ldots, X_{K-1})$ in temporal order, where X_i temporally precedes X_j when $i < j$. We are interested in the causal relationship between these traits and their causal effect on the outcome, Y which will be denoted as X_K and is typically a specific disease status in adulthood. Denote $\mathbf{Z}$ as the vector of all genetic variants. Figure 1 illustrates the causal directed acyclic graph (DAG) showing the causal relationships among the genetic variants, risk factors, outcome, and unmeasured environmental confounders $\mathbf{U}$. Causal effects between the traits must follow the temporal order (only an earlier trait can causally affect a later trait), but we allow the genetic variants to have direct causal effects on any trait at any time. Based on this, we assume the following additive structural equations for each individual:

$$\begin{aligned} X_K := Y &= \sum_{l=1}^{K-1} \beta_{Kl} X_l + f_K(\boldsymbol{U}, \boldsymbol{Z}, \boldsymbol{E}_Y), \\ X_k &= \sum_{l=1}^{k-1} \beta_{kl} X_l + f_k(\boldsymbol{U}, \boldsymbol{Z}, \boldsymbol{E}_{X_k}), \quad k = 2, \ldots, K-1 \\ X_1 &= f_1(\boldsymbol{U}, \boldsymbol{Z}, \boldsymbol{E}_{X_1}). \end{aligned} \tag{1}$$

In these equations, the functions $f_1(\cdot), \cdots, f_K(\cdot)$ are the direct effects of genetics and environmental factors on traits, which can be non-linear and involve arbitrary interactions. Our main assumption is linearity and homogeneity of the causal effect β_{kl} for any earlier trait X_l $(1 \le l \le K-1)$ on any later trait X_k $(l+1 \le k \le K)$. As discussed in the next section, this assumption of linearity and homogeneity becomes crucial, especially when our data only includes GWAS summary statistics, a condition also shared by previous summary-data MR methods.

2.2 Model on summary statistics

Two-sample MR is a popular implementation of MR that uses two independent GWAS samples, one for the exposure and one for the outcome [13]. Similar to many existing methods for two-sample MR, we only require GWAS summary statistics for each trait, so one may call this design "K-sample MR". The fact that summary statistics are sufficient is due to the linearity and homogeneity of the causal effect in (1). This implies a linear model with measurement error on the GWAS summary statistics, as we will explain next.

Specifically, let $\gamma_{kj} \equiv \text{argmin}_\gamma \text{Var}[X_k - \gamma Z_j]$ denote the coefficient corresponding to the least squares projection of the trait X_k on the genetic variant Z_j. The GWAS summary statistics provide estimates of these marginal associations along with their standard errors. In particular, we assume we observe

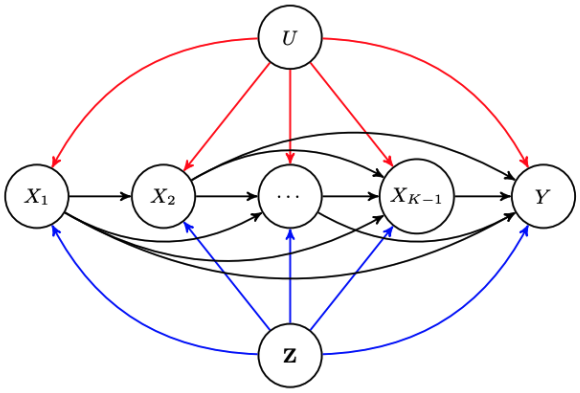

Figure 1: Model overview. The causal directed graph associated with structural equations (1). $X_1, ..., X_{K-1}$ are exposure traits in temporal increasing order and Y is the outcome trait. The blue arrows represent the causal effects of the genetic variants $\boldsymbol{Z}$ on the traits. The red arrows represent the effects of the unmeasured non-heritable confounders $\boldsymbol{U}$.

$\hat{\gamma}_{kj} \sim \mathcal{N}(\gamma_{kj}, \delta_{kj}^2)$ for any SNP Z_j and trait X_k, where the variances δ_{kj}^2 are known. Further, let $\alpha_{kj} \equiv \text{argmin}_\alpha \text{Var}[f_k(U, \boldsymbol{Z}, \boldsymbol{E}_{X_k}) - \alpha Z_j]$ denote the projection of the genetic and environmental direct effects of trait X_k onto SNP Z_j. If a SNP Z_j is used as an instrumental variable for trait X_k, then for any $k' \neq k$ the value $\alpha_{k'j}$ can be viewed as the pleiotropic effects of SNP Z_j on another trait $X_{k'}$. Notice that, unlike the marginal associations γ_{kj}, these direct genetic effects α_{kj} are generally not identifiable from GWAS data. We will make a "balanced pleiotropy" assumption in later sections that is akin to the Instrument Strength Independent of Direct Effect (InSIDE) assumption in the literature [12, 14].

By projecting the structural equations (1) onto the SNP Z_j, we obtain the following linear equations (Supplementary Text):

$$\gamma_{kj} = \sum_{l=1}^{k-1} \beta_{kl}\gamma_{lj} + \alpha_{kj}, \quad k = 1, \ldots, K.$$

These equations can be more conveniently expressed in matrix form as:

$$\boldsymbol{\Gamma} = \tilde{\boldsymbol{B}} \cdot \boldsymbol{\Gamma} + \boldsymbol{A}, \tag{2}$$

where P is the number of SNPs and the matrices are defined as

$$\boldsymbol{\Gamma} \equiv \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1P} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2P} \\ \vdots & \vdots & \ldots & \vdots \\ \gamma_{K1} & \gamma_{K2} & \cdots & \gamma_{KP} \end{bmatrix}, \ \boldsymbol{A} \equiv \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1P} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2P} \\ \vdots & \vdots & \ldots & \vdots \\ \alpha_{K1} & \alpha_{K2} & \cdots & \alpha_{KP} \end{bmatrix}, \ \tilde{\boldsymbol{B}} \equiv \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \beta_{21} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \ldots & \vdots & \vdots \\ \beta_{K1} & \beta_{K2} & \cdots & \beta_{K(K-1)} & 0 \end{bmatrix}.$$

Equation (2) can be alternatively represented as

$$\boldsymbol{\Gamma} = (\boldsymbol{I} + \boldsymbol{B})\boldsymbol{A}, \tag{3}$$

where $\boldsymbol{I}$ is the $K \times K$ identity matrix, and $\boldsymbol{B}$ is the lower-triangular matrix that solves $(\boldsymbol{I} - \tilde{\boldsymbol{B}})^{-1} = \boldsymbol{I} + \boldsymbol{B}$. Compared to (2), the parametrization in (3) avoids matrix inversion, making it more amenable for statistical estimation. It is easy to show that $\tilde{\boldsymbol{B}}$ can be written as a Neumann series: $\tilde{\boldsymbol{B}} = \boldsymbol{B} + \boldsymbol{B}^2 + \ldots$.

Intuitively, the (k, l) entry of the matrix $\boldsymbol{B}$ denotes the total causal effects of X_l on X_k through all directed pathways [15].

We assume that we can use separate GWAS datasets to obtain p-values for SNP selection, thereby mitigating potential selection biases [16, 17]. The selection p-value for each SNP is defined as the Bonferroni-corrected p-value, computed from $K-1$ GWAS summary datasets corresponding to traits X_1 to X_{K-1}. SNPs are subsequently chosen based on the selection p-values using LD clumping [18], ensuring that selected SNPs are approximately independent. Compared to a stringent p-value threshold (such as 10^{-8}) for SNP selection, our method allows for a higher threshold (such as 10^{-4} or 0.01) and uses SNPs that are weakly associated with the exposure traits. This can generally increase the power of the MR analysis.

Upon selecting the SNPs, another set of GWAS datasets is used to obtain the summary statistics for the exposure and outcome traits. For each SNP Z_j, the summary statistics $\hat{\boldsymbol{\Gamma}}j$ follow a normal distribution $\hat{\boldsymbol{\Gamma}}_j \sim \mathcal{N}(\boldsymbol{\Gamma}_j, \boldsymbol{\Sigma}_j)$, where $\boldsymbol{\Gamma}_j := (\gamma_{1j}, \gamma_{2j}, \ldots, \gamma_{Kj})^T$ is a vector of marginal associations and $\boldsymbol{\Sigma}_j$ is a covariance matrix obtained from the GWAS standard errors and a correlation matrix that depends on the extent of sample overlap and the correlation between the traits. This correlation matrix is shared across the SNPs and can be estimated from the non-statistically significant GWAS summary statistics using the method described in [17] (Supplementary Text).

2.3 Identifiability of direct and indirect causal effects

As the MR design uses genetic variants as instrumental variables, a crucial assumption is that these instrumental variables are valid. In particular, MR relies on the assumption that the genetic variants have no horizontal pleiotropic effects [19]. Selecting suitable SNPs for this purpose is particularly challenging due to the complex polygenic nature of complex traits [20, 21]. Many recent MR studies have proposed methods to conduct MR in the presence of horizontal pleiotropy, with various additional assumptions on the pleiotropic effects [12, 14, 22, 23].

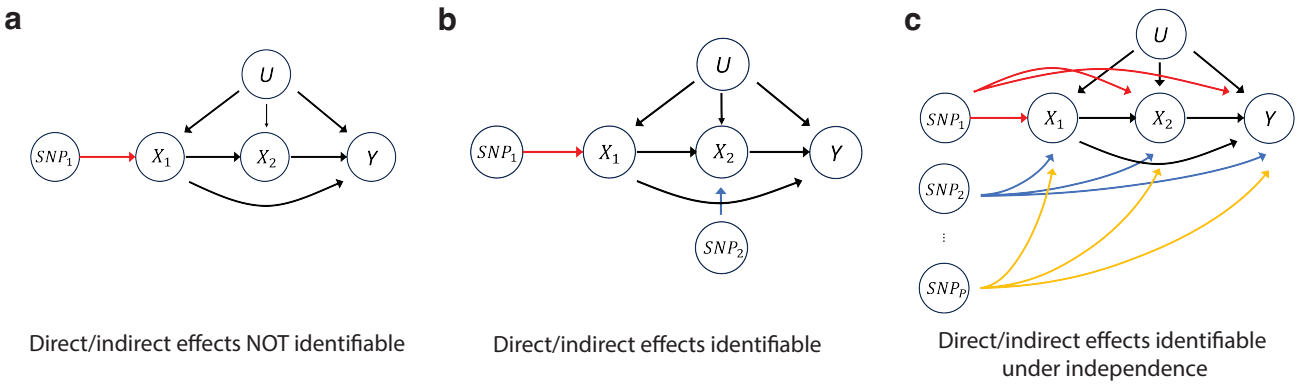

Figure 2: Model identifiability under three scenarios

In the context of life-course MR, we confront an additional challenge that arises from unmeasured mediator-outcome confounding, which can incur biases in the estimation of direct and indirect causal effects even when the treatment is completely randomized [24]. This is illustrated in Figure 2 with two exposures and a outcome ($K = 3$). If we only use genetic instrumental variables for the first risk factor X_1 (Figure 2a), even if all SNPs are valid IV, it is not possible to separate the direct and indirect effects of X_1 due to unmeasured mediator-outcome confounding between X_2 and Y (see Supplementary

Text for a counter-example). In contrast, if each risk factor has its own set of instrumental variables (Figure 2b), it is possible to identify all direct and indirect causal effects under the linear structural equation model in (1). However, given that the sequence of risk factors is often the same risk factor measured at different time points, it may be difficult to find SNPs that only exert their effects at a specific time point.

We propose an alternative and more realistic assumption, which generalizes the InSIDE assumption [14] and is illustrated in Figure 2c. We allow the SNPs to exert effects on all risk factors continuously, and they may all have pleiotropic effects on the outcome and are thus not strictly valid instrumental variables. Nonetheless, we show that the causal coefficient matrix $\tilde{\boldsymbol{B}}$ can still be identified provided two key assumptions are satisfied. First, we require the direct effects $\alpha_{1j}, \ldots, \alpha_{Kj}$ are independent across the traits. Second, the number of SNPs P needs to converge to infinity; in practice, this means that P needs to be sufficiently large. Notice that we only assume independence among α_{kj}, and the marginal associations γ_{kj} can still be correlated across the traits indexed by k. Additionally, we do not need to assume that the genetic associations across SNPs within each trait are identically distributed and allow for heterogeneity across SNPs. These new assumptions allow us to address both pleiotropic effects and mediator-outcome confounding. A formal mathematical statement of this new identifiability result and its proof can be found in the Supplementary Text.

2.4 Model estimation and inference

We use a hierarchical Bayesian framework and Gibbs sampler to infer the direct and indirect causal effects, which can all be expressed in terms of the matrix $\boldsymbol{B}$. The Gibbs sampler allows us to conveniently use posterior samples to construct credible intervals of any function of $\boldsymbol{B}$, including any direct and indirect effects, and proportions of these effects out of the total effects.

Based on earlier investigations [17], both the genetic associations and pleiotropic effects can be highly polygenic, indicating that most elements in $\boldsymbol{A}$ are likely nonzero. Although most SNPs only have weak effects on complex traits, a small set of SNPs may be responsible for the core biological process and have a strong effect on the traits. Thus, as the most critical component of our Bayesian hierarchical model, we assume the following spike-and-slab distributional assumptions on α_{kj}:

$$\boldsymbol{\alpha}_{kj} \mid p_k, \sigma^2_{k0}, \sigma^2_{k1} \overset{\text{ind}}{\sim} (1-p_k)\mathcal{N}(0, \sigma^2_{k0}) + p_k \mathcal{N}(0, \sigma^2_{k1}), \quad k = 1, 2, \ldots, K. \tag{4}$$

This model assumes a two-component Gaussian mixture on the direct genetic effects, wherein all SNPs are permitted to have non-zero genetic effects on each trait and a subset of SNPs is allowed to exhibit larger effects. In addition, we put Gaussian priors on elements of the matrix $\boldsymbol{B}$ and conjugate priors on the hyperparameters:

$$\begin{aligned}
\beta_{kl} \mid \sigma^2 &\overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2), \quad 1 \le l < k \le K \\
\sigma^2 &\sim \text{InvGamma}(\alpha, \beta) \\
p_k &\overset{\text{ind}}{\sim} \text{Beta}(a_k, b_k), \quad k = 1, 2, \ldots, K \\
\sigma^2_{k0} &\overset{\text{ind}}{\sim} \text{InvGamma}(\alpha^0_k, \beta^0_k) \quad k = 1, 2, \ldots, K \\
\sigma^2_{k1} &\overset{\text{ind}}{\sim} \text{InvGamma}(\alpha^1_k, \beta^1_k) \quad k = 1, 2, \ldots, K
\end{aligned}$$

To estimate the above Bayesian hierarchical model, we employ a Gibbs sampler algorithm which can efficiently generate posterior samples when K is small. Posterior samples of $\boldsymbol{B}$ directly give us posterior samples of $\tilde{\boldsymbol{B}}$, which can then be used to construct credible intervals for any direct and indirect causal effects, or their functions. Additionally, we use a simplified empirical Bayes approach to choose the hyper-parameters $(\alpha_k^0, \beta_k^0), (\alpha_k^1, \beta_k^1)$ and (a_k, b_k), recognizing their significant impact on the posterior distributions of $\boldsymbol{B}$. For further computational and mathematical details, see the Supplementary Text.

2.5 Extension to multiple traits at each time point

To account for known confounders, we further expand our model to accommodate multiple traits at any given time point (Figure 3). Specifically, at each time point k, we assume there are $n_k \geq 1$ exposures of interest, which are denoted as $X_{k1}, \ldots, X_{kn_k}$. To facilitate the MR analysis, we require two key assumptions. Firstly, we assume that there is not causal effect between the traits at the same time point. This assumption avoids the need to identify the causal directions of any two traits that are at the same time point. If there is a known causal direction between two traits measured at time point k, we can add a "pseudo-time point" (or stage) $k+1$ after k where we move the outcome trait of the two traits to stage $k+1$. Our second assumption is a generalization of the independence assumption on direct genetic effects described in Section 2.3. In the more general setting considered here, we require that the direct genetic effects α_{kj} of any SNP j must be independent across all traits, including those at the same and different time points. With these assumptions in place, we can adapt our model and Gibbs sampler to infer the direct and indirect causal effects. For more details, see Supplementary Text.

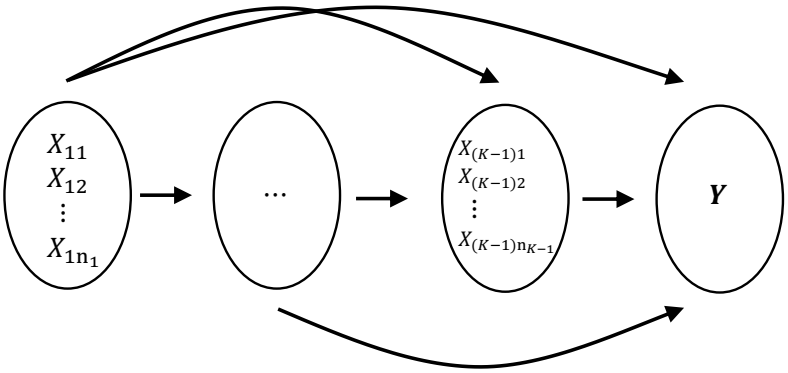

Figure 3: Illustration of causal relationships across traits allowing multiple traits at each time point.

2.6 Data sources

GWAS summary statistics are downloaded from public sources. For adult BMI, data is obtained from the GIANT consortium website (`https://portals.broadinstitute.org/collaboration/giant/index.php/GIANT_consortium_data_files`) and the UK Biobank Neale's lab (`http://www.nealelab.is/uk-biobank/`) using phenotype code 21001. Adult lipid traits (LDL-C, HDL-C, and triglycerides) summary statistics are acquired from the GLGC cohort via the GLGC website (`https://csg.sph.umich.edu/willer/public/lipids2013/`) and from the GERA cohort through GWAS Catalog with study accession numbers GCST007141, GCST007140, and GCST007142. Breast cancer GWAS summary statistics come from GWAS Catalog with study accession number GCST004988, while those for T2D are downloaded from the DIAGRAM website (`https://diagram-consortium.org/downloads.html`), and for stroke from GWAS Catalog with study accession number GCST006906.

For childhood traits, GWAS summary statistics on 1-year-old BMI and 8-year-old BMI are contributed by the Centre For Diabetes Research, University of Bergen, Norway, and the Norwegian Mother, Father and Child study, downloadable from their website (`https://www.fhi.no/en/studies/moba/for-forskere-artikler/gwas-data-from-moba/`). Childhood BMI GWAS data from the EGG Consortium can be found at `http://egg-consortium.org/childhood-bmi.html`. The ALSPAC datasets for childhood lipid traits, SBP, and BMI are obtained from GWAS Catalog with study accession numbers GCST90104679, GCST90104678, GCST90104680, GCST90104683, GCST90104677.

3 Results

3.1 Simulation studies: benchmarking with multivariable MR

We compare our Bayesian framework with the alternative approach using MVMR to perform causal mediation analysis [6, 25]. Specifically, we benchmark our method with IVW-MVMR [8], which is widely used in practice, and GRAPPLE [17], a frequentist approach designed to account for independent pleiotropic effects and weak instruments. At each step k, we apply the chosen multivariable MR method to estimate the direct causal effects of $X_1, \cdots, X_{k-1}$ on X_k. A major drawback of this multi-step approach to life-course MR is that is is difficult to provide reliable statistical inference on indirect and pathwise effects, as the estimates across the different steps are correlated.

We generate synthetic GWAS summary datasets by soft-thresholding real GWAS summary statistics, allowing the "true" genetic effects to be exactly 0 for some SNPs. Specifically, if we observe $\hat{\gamma}_{kj}^{\text{real}}$ with variance $\hat{\delta}_{kj}^2$ for a particular SNP j in a real GWAS dataset, we set $\alpha_{kj} = \text{sign}(\hat{\gamma}_{kj}^{\text{real}})(|\hat{\gamma}_{kj}^{\text{real}}| - 0.01)_+$ and also randomly shuffle α_{kj} across j within each trait k to ensure independence across traits. Once the matrix $\boldsymbol{A}$ is generated, we further generate $\boldsymbol{\Gamma}$ following our linear model (2) with a pre-specified matrix $\boldsymbol{B}$. To specify $\boldsymbol{B}$, we simulate three scenarios with $K = 3$, $K = 4$, and a multivariate case where there are multiple exposures at each time point (Figure 4). Finally, we simulate our synthetic summary statistics $\hat{\gamma}_{kj}^{\text{simu}} \sim \mathcal{N}(\gamma_{kj}, \hat{\delta}_{kj}^2)$ independently across all SNPs and traits.

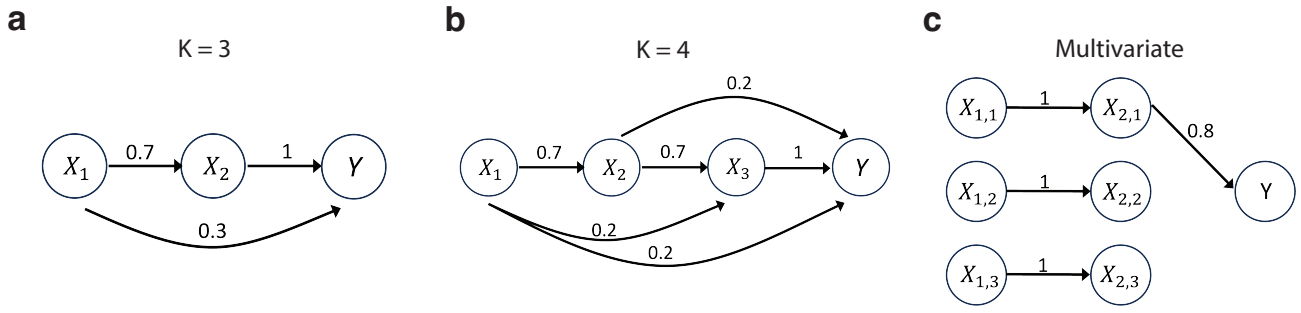

Figure 4: Simulation scenarios

To mimic real MR mediation analysis, where earlier childhood traits typically have a smaller sample size than adulthood traits, we generate synthetic data for earlier traits based on childhood GWAS data. Specifically, we use childhood BMI data from the Norwegian Mother, Father, and Child Cohort Study (MoBa) [26] and childhood lipid traits from the Avon Longitudinal Study of Parents and Children (ALSPAC) cohort [27]. Adult GWAS datasets are used to generate summary statistics for the remaining traits. See SI test for additional details of the simulation setup.

Figure 5 shows the coverage and average lengths of the confidence and credible intervals of the direct

and indirect causal effects when $K = 3$ obtained using different methods. SNPs are selected based on cutoffs of the p-values calculated from $\gamma_{kj}/\hat{\delta}_{kj}$, indicating the strength of the true association between SNP j and trait k. Genetic variants that are weakly associated with the traits are selected under a large p-value threshold. As shown in the figure, for the direct causal effects, though the IVW-MVMR method has the shortest confidence intervals, they are also severely under-covered, possibly due to the presence of pleiotropic effects and weak instruments. In contrast, both our Bayesian approach and GRAPPLE do not suffer from the weak instrument bias, and have good coverage irrespective of the strength of the SNPs. Compared to GRAPPLE, our Bayesian approach offers more efficient and powerful inference. Additionally, regarding the indirect causal effects of X_1, only our Bayesian approach provides reliable inference, and our credible intervals demonstrate good coverage and reasonable power. We also observe similar advantages of our Bayesian approach in the other two simulation scenarios: $K = 4$ and the multivariate case (Figure S1-S2).

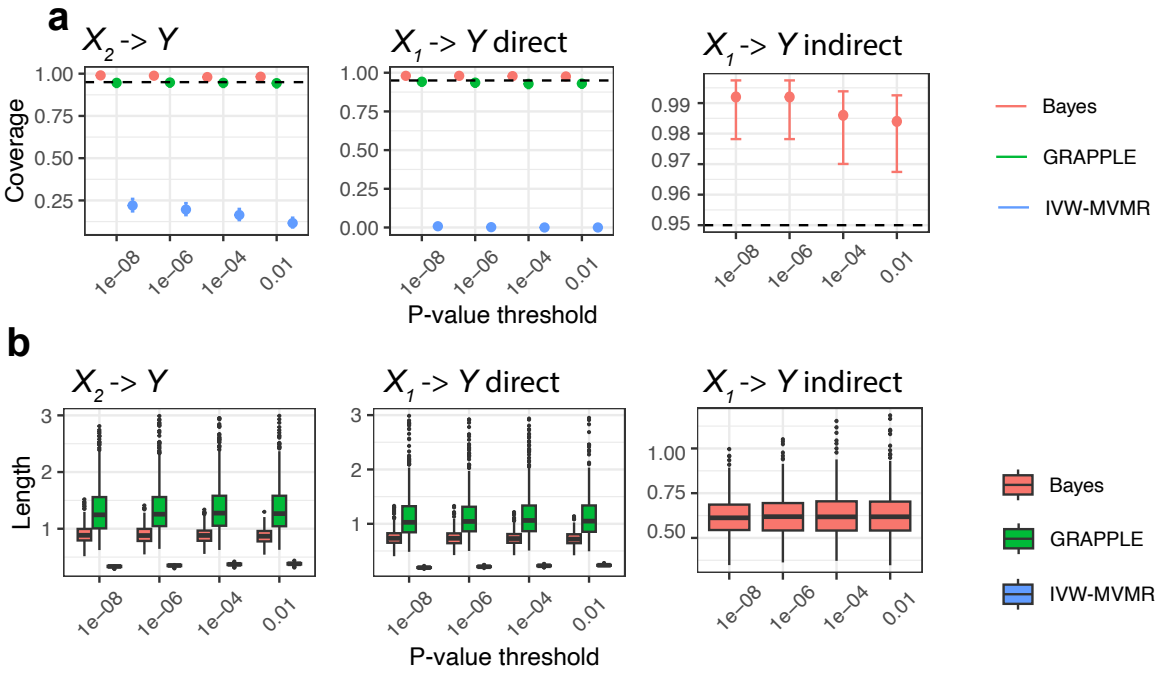

Figure 5: Simulation results for $K = 3$. The first row illustrates empirical coverage of 95% confidence intervals on direct effects of X_1 and X_2 on Y and indirect effects of X_1 over 500 repeated simulations. The second row displays the boxplots of lengths of confidence intervals over repeated simulations.

3.2 Effect of early life body size on breast cancer

Several recent studies have indicated that early-life body size serves as a protective factor against the risk of breast cancer [28, 29, 30]. Using IVW-MVMR for the mediation analysis, [3] observed that the protective effect of early-life body size on breast cancer is not mediated by adult body size, while adult body size itself does not causally influence breast cancer. However, in their analysis, though the adult body size was quantified by the adult BMI trait, the representation of early-life body size relied on the early-life body score from UK Biobank, which is a questionnaire recall trait asking adults to recall whether they were thinner or plumper than average at the age of 10. As discussed by the original authors, the use of such an imprecise measure for childhood body size may raise concerns about the credibility of the scientific conclusions when compared to using direct measurements of childhood BMI.

To address this, we revisit the analysis, comparing the earlier results with a new approach employing our Bayesian method and directly utilizing childhood BMI GWAS summary statistics to represent childhood body size. Similar to [3], for the adult body size we use the adult BMI trait from UK Biobank and use GWAS summary statistics on breast cancer from [31] for the outcome trait. To avoid instrument selection biases, we select SNPs based on their significance in separate GIANT Adult BMI [32] and EGG childhood BMI [33] GWAS datasets. As shown in Figure 6a, when early-life body score from the UK Biobank is used as the exposure, our Bayesian method, along with IVW-MVMR and GRAPPLE, replicates the findings from [3], indicating no causal effect of adult BMI and a direct protective effect of early-life body size on breast cancer.

However, replacing the early-life body size trait with the GWAS trait of 8-year-old BMI from MoBa yields surprising results. IVW-MVMR suggests a protective causal effect of adult BMI on breast cancer, contradicting earlier conclusions, while childhood BMI shows no direct effect (Figure 6b). In contrast, GRAPPLE loses its power to detect any causal effects. Only our Bayesian approach provides a similar conclusion as the earlier analysis, affirming the direct protective effect of childhood BMI. Despite the 8-year-old BMI trait offering a more precise measurement of childhood BMI, its use in MR is challenging due to its small sample size (3K samples compared to half a million in the UK Biobank). To assess the change in instrument strength when replacing the early-life body size trait from UK Biobank with the 8-year-old BMI GWAS trait from MoBa, we compute the conditional F-statistics [34] of the exposure traits. We observe a substantial decrease in the conditional F-statistics in the childhood dataset (Figure S3), indicating a loss of instrument strength due to the limited cohort size of the MoBa study. Similar to our simulations, the analysis demonstrates that IVW-MVMR can produce unreliable confidence intervals, while GRAPPLE lacks power when the GWAS dataset sample size for any of the exposure traits is small. Our new Bayesian method demonstrates both efficiency and robustness, especially in the mediation analysis where GWAS studies for early-life traits always have small sample sizes.

To further enhance our understanding of the impact of early-life body size on breast cancer, we include infant BMI, specifically the 1-year-old BMI GWAS trait from MoBa, in our mediation analysis. As the 1-year-old BMI and 8-year-old BMI traits share the same cohort, we also incorporate the estimated noise correlation matrix in our Bayesian approach (Figure S4). As shown in Figure 6c, we observe a causal effect of 1-year-old BMI on 8-year-old BMI, but no direct causal effects of 1-year-old BMI on adult BMI or breast cancer. This suggests that the protective effect of body size on breast cancer is confined to a specific period during childhood development.

3.3 high blood pressure, BMI and stroke

High blood pressure is widely acknowledged as the primary modifiable risk factor for stroke [35]. Through univariate MR analyses, a recent study has also revealed additional potential causal risks for stroke outcomes including the adult BMI [36]. Moreover, using univariate MR, recent studies have suggested that BMI has a significant positive causal effect on high blood pressure in adulthood [11, 12]. We aim to unravel the intertwined impacts of high blood pressure and BMI on stroke, delving into the dynamic evolution of these causal relationships throughout the life course. To ensure robust findings, we will also account for potential confounding risk factors, such as lipid traits.

Specifically, we analyze the causal effects of three risk factors: systolic blood pressure (SBP), BMI and low-density lipoprotein cholesterol (LDL-C) in both childhood and adulthood on stroke, using our multivariate mediation analysis framework (Figure 3). As our multivariate framework does not

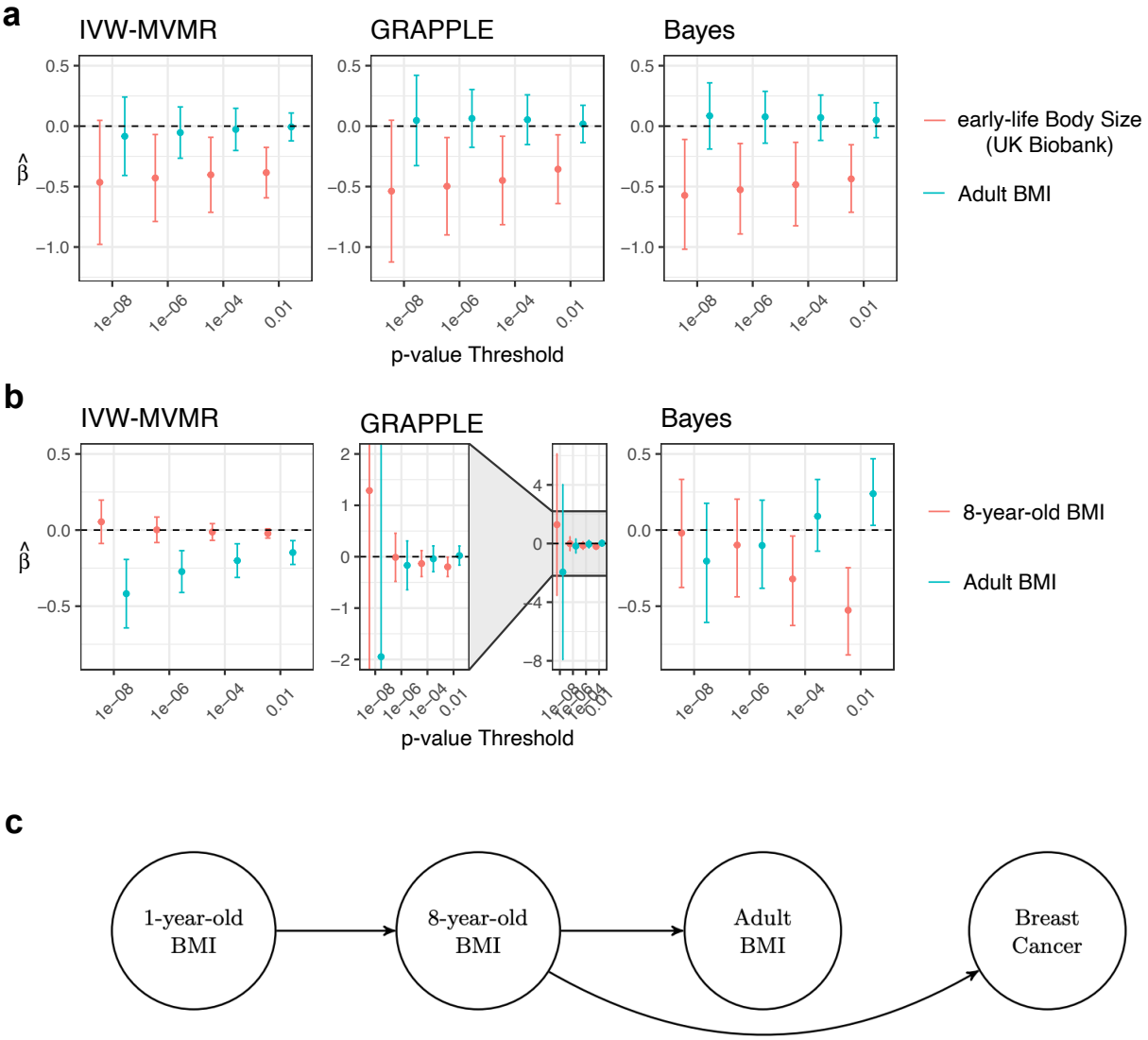

Figure 6: Evaluation of the effect of body size on breast cancer at different ages. a) Estimated effects of childhood body size (from UK Biobank) and adult BMI on breast cancer risk as estimated by MV-IVW, GRAPPLE, and our Bayesian approach. b) Estimated effects of childhood BMI and adult BMI on breast cancer risk from different methods. c) Estimated causal DAG from our Bayesian approach with selection p-value threshold at 10^{-2}. The black arrows indicate significant direct effects.

allow any causal relationships among the risk factors at the same time point, to accommodate the potential causal effects of BMI on SBP, we create two additional "time points" (stages) for the SBP traits (Figure 7a). For the childhood BMI, LDL-C and SBP traits we use GWAS summary statistics from the ALSPAC cohort [27]. We use GWAS summary statistics from UK Biobank for adult BMI and SBP, and the summary statistics from Global Lipids Genetics Consortium [37] for adult LDL-C. For stroke, the summary statistics are from [38]. SNP selections are based on the p-values in the GERA GWAS for SBP and LDL-C [39, 40], and the p-values in GIANT adult BMI dataset and EGG childhood BMI dataset.

Figure 7b illustrates the estimated causal directed graph at various selection p-value thresholds using our Bayesian method. More significant arrows emerge with milder p-value thresholds, suggesting increased power with the inclusion of weak instruments. As anticipated, all childhood exposures exhibit

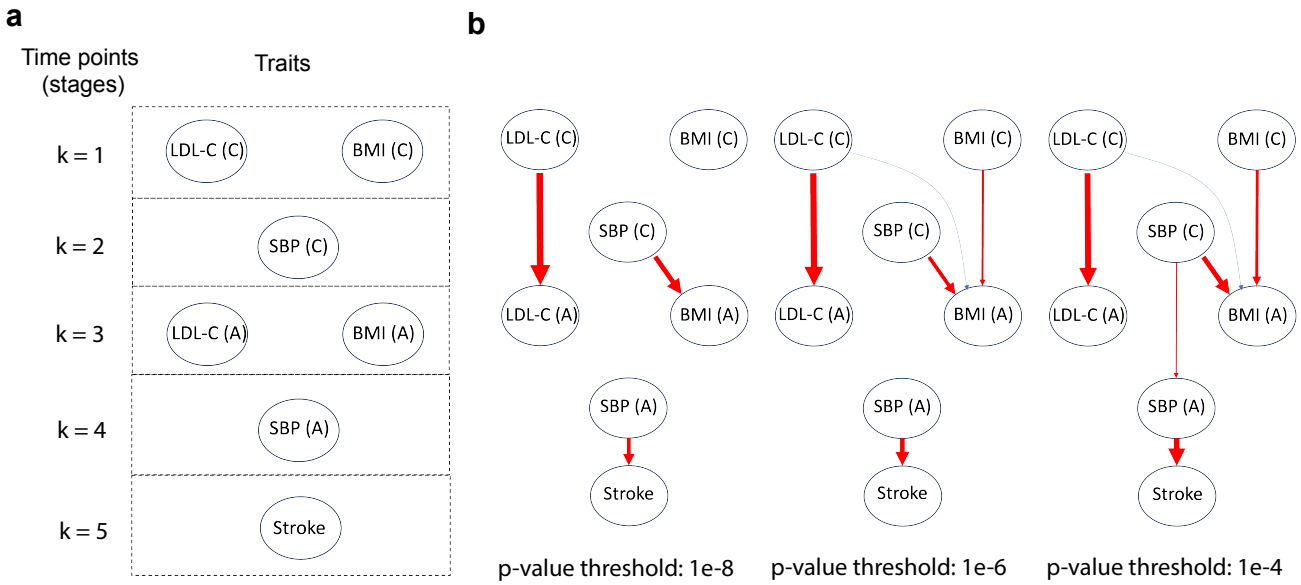

Figure 7: Multivariate mediation analysis on LDL-C, BMI, SBP and stroke. a) Design of the stages. Childhood traits are marked with label (C) and adulthood traits are marked with label (A). b) Estimated causal effects with different selection p-value thresholds. Arrows correspond to the significant effects, with thicker arrows indicating larger effects. Positive effects are red and negative effects are blue.

significant causal impacts on their corresponding adulthood exposures. Our findings indicate a positive causal direct effect of adulthood SBP on stroke, while no compelling evidence supports direct effects from other traits on stroke, including childhood traits and adult BMI.

A surprising result from our analysis is a lack of support for a causal effect of adult BMI on adult SBP, which is contrary to earlier univariate Mendelian Randomization (MR) findings [11, 12]. To take a closer examination, we also perform multivariable MR, with childhood SBP and adult BMI as exposures and adult SBP as the outcome. Both IVW-MVMR and GRAPPLE indicate that adult BMI no longer exerts a positive causal effect on adult SBP after accounting for childhood SBP (Figure S5a). Additionally, we observe a stronger genetic correlation between childhood SBP and adult BMI compared to that between adult SBP and BMI (Figure S5b). Collectively, these results suggest that the confounding effect of childhood SBP may have contributed the association of adult BMI on SBP identified in previous univariable MR analysis.

4 Discussion

Based on a unified model across all traits, we propose a Bayesian approach with GWAS summary data for life-course Mendelian Randomization. This proposed method allows the assessment of time-varying causal effects of heritable risk factors, distinguishing between direct and indirect causal effects of traits in temporal order. Addressing a key challenge in life-course MR—specifically, the high genetic association of a trait across different ages and the limited cohort size of age-specific GWAS—our method exhibits superior performance. Our method enjoys robustness to bias in using weakly associated SNPs as instruments and can efficiently integrate information across traits.

A key assumption in our causal structural equation system is linearity and homogeneity of the risk factors' causal effects. Concerns may arise about the use of linear structural equations when the outcome trait or the exposure trait is binary. In the former case, MR methods based on linear models

can still estimate meaningful but attenuated causal effects [12]. In the latter case, caveats of MR are discussed in [41]. Interpreting the estimated linear structural equations may be challenging in practice, particularly when there is a sequence of binary exposures across time.

For the identification of causal effects in life-course MR, a similar argument was discussed in [7]. They show that in order for the causal direct effects to be identified, genetic variants must exert different effects on each exposure in the model, and these effects must be linearly independent, or equivalently, the true SNP-trait association matrix has full rank. This is necessary but not sufficient for identifying all direct and indirect effects. For instance, in univariable MR, pleiotropic effects must be restricted (such as by the InSIDE assumption) for causal identification. Merely assuming that the pleiotropic effects are not perfectly correlated with the genetic associations with the exposure does not suffice to ensure the identification of the causal effect. An additional assumption, such as our independence assumption of direct associations across traits, is needed to ensure the separation of direct and indirect causal effects. While our independence assumption permits pervasive pleiotropy, it assumes no correlated pleiotropy for any traits. In practice, if confounding pathways exist between traits leading to correlated pleiotropy, we recommend collecting additional GWAS data for confounding traits and explicitly adjusting them using the multivariate extension of our approach.

A limitation of life-course MR is missing time points, where exposures at additional time points play a role in causal relationships but are not considered in the MR analysis [6, 42]. If the exposure trait at the missing time point X_k is a confounder of any two later-time traits X_{k_1} and X_{k_2} in our model, it may invalidate the assumptions in our model. Specifically, if a selected SNP Z_j is associated with X_k, its direct associations with X_{k_1} and X_{k_2} become correlated through the paths $Z_j \to X_k \to X_{k_1}$ and $Z_j \to X_k \to X_{k_2}$, thereby violating our independence assumption. However, our independence assumption remains true if X_k is not a hidden confounder for any pair of subsequent traits. Examples include when X_k solely influences the outcome or the exposure at the next t time point.

Another concern about life-course MR is the possibility of reverse causality, where the outcome trait may exert causal effects on the exposure at the latest time point [6]. Typically, life-course MR benefits from clear causal directions owing to the temporal order of traits. However, in many applications a notable challenge arises when the last exposure and the outcome are collected at the same time, such as in adulthood. In cases where the selected SNPs primarily associate with the outcome through their connections with the exposures, MR generally remains resilient to bias caused by reverse causation [6, 17, 43]. Our framework shares this advantageous property of MR.

5 Declaration of Interests

The authors have declared that no competing interests exist.

6 Acknowledgments

J.W. is partly supported by the National Science Foundation under grants DMS-2113646 and DMS-2238656. Q.Z. is partly supported by EPSRC grant EP/V049968/1.

7 Code availability

The R package MrMediation for conducting our Beyesian mediation MR analysis is publicly available for installation at (`https://github.com/ZixuanWu1/MrMediation`).

References

[1] Davey Smith, G. and Ebrahim, S. (2003). 'mendelian randomization': can genetic epidemiology contribute to understanding environmental determinants of disease? International journal of epidemiology *32*, 1–22.

[2] Davey Smith, G. and Hemani, G. (2014). Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Human molecular genetics *23*, R89–R98.

[3] Richardson, T.G., Sanderson, E., Elsworth, B., Tilling, K., and Smith, G.D. (2020). Use of genetic variation to separate the effects of early and later life adiposity on disease risk: mendelian randomisation study. bmj *369*.

[4] Shi, J., Swanson, S.A., Kraft, P., Rosner, B., De Vivo, I., and Hernán, M.A. (2022). Mendelian randomization with repeated measures of a time-varying exposure: an application of structural mean models. Epidemiology *33*, 84–94.

[5] Labrecque, J.A. and Swanson, S.A. (2019). Interpretation and potential biases of mendelian randomization estimates with time-varying exposures. American journal of epidemiology *188*, 231–238.

[6] Sanderson, E., Richardson, T.G., Morris, T.T., Tilling, K., and Davey Smith, G. (2022). Estimation of causal effects of a time-varying exposure at multiple time points through multivariable mendelian randomization. PLoS Genetics *18*, e1010290.

[7] Power, G.M., Sanderson, E., Pagoni, P., Fraser, A., Morris, T., Prince, C., Frayling, T.M., Heron, J., Richardson, T.G., Richmond, R., et al. (2023). Methodological approaches, challenges, and opportunities in the application of mendelian randomisation to lifecourse epidemiology: A systematic literature review. European Journal of Epidemiology , 1–20.

[8] Burgess, S., Dudbridge, F., and Thompson, S.G. (2015). Re:"multivariable mendelian randomization: the use of pleiotropic genetic variants to estimate causal effects". American journal of epidemiology *181*, 290–291.

[9] Shi, J., Swanson, S.A., Kraft, P., Rosner, B., De Vivo, I., and Hernán, M.A. (2021). Instrumental variable estimation for a time-varying treatment and a time-to-event outcome via structural nested cumulative failure time models. BMC medical research methodology *21*, 1–12.

[10] Cao, Y., Rajan, S.S., and Wei, P. (2016). Mendelian randomization analysis of a time-varying exposure for binary disease outcomes using functional data analysis methods. Genetic epidemiology *40*, 744–755.

[11] Lee, M.R., Lim, Y.H., and Hong, Y.C. (2018). Causal association of body mass index with hypertension using a mendelian randomization design. Medicine *97*.

[12] Zhao, Q., Wang, J., Bowden, J., and Small, D. (2018). Statistical inference in two-sample summary-data mendelian randomization using robust adjusted profile score. Annals of Statistics *48*.

[13] Hartwig, F.P., Davies, N.M., Hemani, G., and Davey Smith, G. (2016). Two-sample mendelian randomization: avoiding the downsides of a powerful, widely applicable but potentially fallible technique.

[14] Bowden, J., Davey Smith, G., and Burgess, S. (2015). Mendelian randomization with invalid instruments: effect estimation and bias detection through egger regression. International journal of epidemiology *44*, 512–525.

[15] Sullivant, S., Talaska, K., and Draisma, J. (2008). Trek separation for gaussian graphical models. Annals of Statistics *38*.

[16] Zhao, Q., Chen, Y., Wang, J., and Small, D.S. (2019). Powerful three-sample genome-wide design and robust statistical inference in summary-data mendelian randomization. International journal of epidemiology *48*, 1478–1492.

[17] Wang, J., Zhao, Q., Bowden, J., Hemani, G., Davey Smith, G., Small, D.S., and Zhang, N.R. (2021). Causal inference for heritable phenotypic risk factors using heterogeneous genetic instruments. PLoS genetics *17*, e1009575.

[18] Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M.A., Bender, D., Maller, J., Sklar, P., De Bakker, P.I., Daly, M.J., et al. (2007). Plink: a tool set for whole-genome association and population-based linkage analyses. The American journal of human genetics *81*, 559–575.

[19] Ebrahim, S. and Davey Smith, G. (2008). Mendelian randomization: can genetic epidemiology help redress the failures of observational epidemiology? Human genetics *123*, 15–33.

[20] Timpson, N.J., Greenwood, C.M., Soranzo, N., Lawson, D.J., and Richards, J.B. (2018). Genetic architecture: the shape of the genetic contribution to human traits and disease. Nature Reviews Genetics *19*, 110–124.

[21] O'Connor, L.J., Schoech, A.P., Hormozdiari, F., Gazal, S., Patterson, N., and Price, A.L. (2019). Extreme polygenicity of complex traits is explained by negative selection. The American Journal of Human Genetics *105*, 456–476.

[22] Verbanck, M., Chen, C.Y., Neale, B., and Do, R. (2018). Detection of widespread horizontal pleiotropy in causal relationships inferred from mendelian randomization between complex traits and diseases. Nature genetics *50*, 693–698.

[23] Morrison, J., Knoblauch, N., Marcus, J.H., Stephens, M., and He, X. (2020). Mendelian randomization accounting for correlated and uncorrelated pleiotropic effects using genome-wide summary statistics. Nature genetics *52*, 740–747.

[24] VanderWeele, T.J., Vansteelandt, S., and Robins, J.M. (2014). Effect decomposition in the presence of an exposure-induced mediator-outcome confounder. Epidemiology (Cambridge, Mass.) *25*, 300.

[25] Carter, A.R., Sanderson, E., Hammerton, G., Richmond, R.C., Davey Smith, G., Heron, J., Taylor,

A.E., Davies, N.M., and Howe, L.D. (2021). Mendelian randomisation for mediation analysis: current methods and challenges for implementation. European journal of epidemiology *36*, 465–478.

[26] Helgeland, Ø., Vaudel, M., Juliusson, P.B., Lingaas Holmen, O., Juodakis, J., Bacelis, J., Jacobsson, B., Lindekleiv, H., Hveem, K., Lie, R.T., et al. (2019). Genome-wide association study reveals dynamic role of genetic variation in infant and early childhood growth. Nature communications *10*, 4448.

[27] O'Nunain, K., Sanderson, E., Holmes, M.V., Smith, G.D., and Richardson, T.G. (2023). A genome-wide association study of childhood adiposity and blood lipids. Wellcome Open Research *6*, 303.

[28] Ooi, B.N.S., Loh, H., Ho, P.J., Milne, R.L., Giles, G., Gao, C., Kraft, P., John, E.M., Swerdlow, A., Brenner, H., et al. (2019). The genetic interplay between body mass index, breast size and breast cancer risk: a mendelian randomization analysis. International journal of epidemiology *48*, 781–794.

[29] Vatten, L.J. and Kvinnsland, S. (1990). Body mass index and risk of breast cancer. a prospective study of 23,826 norwegian women. International journal of cancer *45*, 440–444.

[30] Guo, Y., Warren Andersen, S., Shu, X.O., Michailidou, K., Bolla, M.K., Wang, Q., Garcia-Closas, M., Milne, R.L., Schmidt, M.K., Chang-Claude, J., et al. (2016). Genetically predicted body mass index and breast cancer risk: Mendelian randomization analyses of data from 145,000 women of european descent. PLoS medicine *13*, e1002105.

[31] Michailidou, K., Lindström, S., Dennis, J., Beesley, J., Hui, S., Kar, S., Lemaçon, A., Soucy, P., Glubb, D., Rostamianfar, A., et al. (2017). Association analysis identifies 65 new breast cancer risk loci. Nature *551*, 92–94.

[32] Justice, A.E., Winkler, T.W., Feitosa, M.F., Graff, M., Fisher, V.A., Young, K., Barata, L., Deng, X., Czajkowski, J., Hadley, D., et al. (2017). Genome-wide meta-analysis of 241,258 adults accounting for smoking behaviour identifies novel loci for obesity traits. Nature communications *8*, 14977.

[33] Felix, J.F., Bradfield, J.P., Monnereau, C., Van Der Valk, R.J., Stergiakouli, E., Chesi, A., Gaillard, R., Feenstra, B., Thiering, E., Kreiner-Møller, E., et al. (2016). Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. Human molecular genetics *25*, 389–403.

[34] Sanderson, E., Spiller, W., and Bowden, J. (2021). Testing and correcting for weak and pleiotropic instruments in two-sample multivariable mendelian randomization. Statistics in medicine *40*, 5434–5452.

[35] Willmot, M., Leonardi-Bee, J., and Bath, P.M. (2004). High blood pressure in acute stroke and subsequent outcome: a systematic review. Hypertension *43*, 18–24.

[36] Chen, L., Peters, J.E., Prins, B., Persyn, E., Traylor, M., Surendran, P., Karthikeyan, S., Yonova-Doing, E., Di Angelantonio, E., Roberts, D.J., et al. (2022). Systematic mendelian randomization

using the human plasma proteome to discover potential therapeutic targets for stroke. Nature communications *13*, 6143.

[37] Willer, C., Schmidt, E., Sengupta, S., Peloso, G., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M., Mora, S., et al. (2013). Discovery and refinement of loci associated with lipid levels. Nature genetics *45*, 1274–1283.

[38] Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., Rutten-Jacobs, L., Giese, A.K., Van Der Laan, S.W., Gretarsdottir, S., et al. (2018). Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nature genetics *50*, 524–537.

[39] Hoffmann, T.J., Ehret, G.B., Nandakumar, P., Ranatunga, D., Schaefer, C., Kwok, P.Y., Iribarren, C., Chakravarti, A., and Risch, N. (2017). Genome-wide association analyses using electronic health records identify new loci influencing blood pressure variation. Nature genetics *49*, 54–64.

[40] Hoffmann, T.J., Theusch, E., Haldar, T., Ranatunga, D.K., Jorgenson, E., Medina, M.W., Kvale, M.N., Kwok, P.Y., Schaefer, C., Krauss, R.M., et al. (2018). A large electronic-health-record-based genome-wide study of serum lipids. Nature genetics *50*, 401–413.

[41] Burgess, S. and Labrecque, J.A. (2018). Mendelian randomization with a binary exposure variable: interpretation and presentation of causal estimates. European journal of epidemiology *33*, 947–952.

[42] Tian, H. and Burgess, S. (2023). Estimation of time-varying causal effects with multivariable mendelian randomization: some cautionary notes. International journal of epidemiology *52*, 846–857.

[43] Burgess, S., Swanson, S.A., and Labrecque, J.A. (2021). Are mendelian randomization investigations immune from bias due to reverse causation? European Journal of Epidemiology *36*, 253–257.

Supplemental materials for "Causal mediation analysis for time-varying heritable risk factors with Mendelian Randomization"

S1 Additional mathematical details

S1.1 Derivation of summary-data linear models from individual structural equations

Recall that we define

$$\gamma_{kj} \equiv \text{argmin}_{\gamma}\text{Var}[X_k - \gamma Z_j]$$
$$\alpha_{kj} \equiv \text{argmin}_{\alpha}\text{Var}[f_k(\boldsymbol{U}, \boldsymbol{Z}, \boldsymbol{E}_{X_k}) - \alpha Z_j]$$

Projecting X_k onto Z_j, we have

$$X_k = \gamma_{kj} Z_j + (X_k - \gamma_{kj} Z_j) = \gamma_{kj} Z_j + \epsilon_{kj} \tag{1}$$

where $\text{corr}(Z_j, \epsilon_{kj}) = 0$. On the other hand, we have

$$X_k = \sum_{l=1}^{k-1} \beta_{kl} X_l + \alpha_{kj} Z_j + (f_k(\boldsymbol{U}, \boldsymbol{Z}, \boldsymbol{E}_{\boldsymbol{X_k}}) - \alpha_{kj} Z_j) = \sum_{l=1}^{k-1} \beta_{kl} X_l + \alpha_{kj} Z_j + \tilde{e}_{kj}, \tag{2}$$

where $\text{corr}(Z_j, \tilde{e}_{kj}) = 0$. Substitute (1) into (2) for $l = 1, \ldots, k-1$, we further have

$$X_k = \left(\sum_{l=1}^{k-1} \beta_{kl}\gamma_{lj}\right) Z_j + \alpha_{kj} Z_j + \left(\sum_{l=1}^{k-1} \beta_{kl}\epsilon_{lj}\right) + \tilde{e}_{kj} = \left(\sum_{l=1}^{k-1} \beta_{kl}\gamma_{lj} + \alpha_{kj}\right) Z_j + \left(\sum_{l=1}^{k-1} \beta_{kl}\epsilon_{lj} + \tilde{e}_{kj}\right). \tag{3}$$

Since $\text{corr}(Z_j, (\sum_{l=1}^{k-1} \beta_{kl}\epsilon_{lj} + \tilde{e}_{kj})) = 0$, both (1) and (3) are linear projections of X_k on Z_j, thus

$$\gamma_{kj} = \sum_{l=1}^{k-1} \beta_{kl}\gamma_{lj} + \alpha_{kj}, \quad k = 1, \ldots, K \tag{4}$$

S1.2 Noise correlation estimation

In section 2.3 we claim that when traits have overlapping cohorts, the estimates $\hat{\gamma}_{kj}$ are not independent across k, but approximately, all SNPs share the same correlation matrix which can be effectively estimated from the GWAS summary statistics themselves As shown in [1], for any risk factor k and l we have

$$\text{Corr}(\hat{\gamma}_{kj}, \hat{\gamma}_{lj}) \approx \frac{N_{kl}}{\sqrt{N_k N_l}} \text{Corr}(X_{ks}, X_{ls}), \tag{5}$$

where N_k and N_l are the sample sizes of the risk factor X_k and X_l, N_{kl} is the size of overlapping samples and X_{ks} denotes the measure of X_k for individuals s. The correlation of X_k and X_l of any shared sample is $\text{Corr}[X_{ks}, X_{ls}]$. As a consequence, we assume

$$\begin{pmatrix} \hat{\gamma}_{1j} \\ \hat{\gamma}_{2j} \\ \vdots \\ \hat{\gamma}_{Kj} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \gamma_{1j} \\ \gamma_{2j} \\ \vdots \\ \gamma_{Kj} \end{pmatrix}, \begin{pmatrix} \delta_{1j} & & & \\ & \delta_{2j} & & \\ & & \ddots & \\ & & & \delta_{Kj} \end{pmatrix} \boldsymbol{R} \begin{pmatrix} \delta_{1j} & & & \\ & \delta_{2j} & & \\ & & \ddots & \\ & & & \delta_{Kj} \end{pmatrix} \right),$$

where $\boldsymbol{R}$ is the unknown shared correlation matrix.

To estimate $\boldsymbol{R}$, we choose SNPs where $\gamma_{kj} = 0$ for all k so that we can estimate the shared correlations using the sample correlation of the chosen SNPs. To do this we select SNPs with p-values $p_j \geq 0.5$ in all selection files.

S1.3 Identifiability

In this section, we provide a formal mathematical statement for identifying the causal effect matrix $\tilde{\boldsymbol{B}}$ in the linear model:

$$\boldsymbol{\Gamma} = \tilde{\boldsymbol{B}} \cdot \boldsymbol{\Gamma} + \boldsymbol{A} \tag{6}$$

where

$$\boldsymbol{\Gamma} \equiv \begin{bmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1P} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2P} \\ \vdots & \vdots & \cdots & \vdots \\ \gamma_{K1} & \gamma_{K2} & \cdots & \gamma_{KP} \end{bmatrix}, \ \boldsymbol{A} \equiv \begin{bmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1P} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2P} \\ \vdots & \vdots & \cdots & \vdots \\ \alpha_{K1} & \alpha_{K2} & \cdots & \alpha_{KP} \end{bmatrix}, \ \tilde{\boldsymbol{B}} \equiv \begin{bmatrix} 0 & 0 & \cdots & 0 & 0 \\ \beta_{21} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ \beta_{K1} & \beta_{K2} & \cdots & \beta_{K(K-1)} & 0 \end{bmatrix}.$$

In the main text, it was highlighted that the causal effect matrix $\tilde{\boldsymbol{B}}$ is not identifiable when only genetic instrumental variables for the first risk factor X_1 is involved in the multivariable MR analysis. Here we give a simple example to illustrate how the direct and indirect effects can be not inseparable in these cases. Suppose

$$\begin{aligned} X_1 &= \sum_{j=1}^{P} \alpha_j Z_j, \\ X_2 &= \beta_1 X_1 = \sum_{j=1}^{P} \beta_1 \alpha_j Z_j, \\ Y &= \beta_2 X_1 + \beta_3 X_2 = \sum_{j=1}^{P} (\beta_2 + \beta_1 \beta_3) \alpha_j Z_j, \end{aligned}$$

where $\boldsymbol{Z} = (Z_1, Z_2, \ldots, Z_P)$ are independent SNPs. Suppose we have infinite sample size, thus $(\alpha_1, \alpha_2, \ldots, \alpha_P)$, $\beta_1(\alpha_1, \alpha_2, \cdots, \alpha_P)$ and $(\beta_2 + \beta_1\beta_3)(\alpha_1, \alpha_2, \cdots, \alpha_P)$ are directly observed. In this case β_1 can be simply identified by looking at the ratio of the first two marginal effect. On the other hand, for any $\beta_3' \in \mathbb{R}$, one can define $\beta_2' = \beta_2 - \beta_1(\beta_3' - \beta_3)$, so that $(\beta_2' + \beta_1\beta_3') = (\beta_2 + \beta_1\beta_3)$. Hence in this case the direct effects and indirect effects of X_1 on Y are not identifiable.

Now suppose for each single exposure trait we have some instrumental strength. As discussed in the main text, we do not assume that the SNPs are valid IVs for any single exposure trait but allow a large number of SNPs P to be used. Unlike the identifiability problems in standard linear models, we do not assume identical distributions across SNPs to allow for arbitrary heterogeneity. We can identify $\tilde{\boldsymbol{B}}$ in the following sense:

Theorem 1. *Under model* (6)*, we further assume*

- *Infinite GWAS data: the sample sizes n_k of the GWAS data for each trait k are large enough that the marginal associations γ_{kj} can be uniformly consistently estimated across j and k by the summary statistics when $P \to \infty$ and $\min_k n_k \to \infty$.*
- *Independence: α_{kj}'s are mutually independent across k and j*
- *Bounded moment: $\sup_{k,j} \mathbb{E}[\alpha_{kj}^4] < M$ for some constant M.*
- *Well-behaved limiting average moments: For any $k, l \in [K]$, we have*

1. $\mu_k := \lim_{P\to\infty} \frac{1}{P}\sum_{j=1}^{P} \mathbb{E}[\alpha_{kj}] < \infty$ *exists and* $\lim_{P\to\infty} \frac{1}{P}\sum_{j=1}^{P} |\mathbb{E}[\alpha_{kj}] - \mu_k| = 0$
2. $\lim_{P\to\infty} \frac{1}{P}\sum_{j=1}^{P} \mathbb{E}[\alpha_{kj}\alpha_{lj}] < \infty$ *exists*

- *Non-zero instrumental effect: For any* $k \in [K-1]$*, we have*

$$\lim_{P\to\infty} \frac{1}{P}\sum_{j=1}^{P} Var(\alpha_{kj}) > 0$$

Then elements of $\tilde{\boldsymbol{B}}$ *in* (6) *can be consistently estimated when* $P \to \infty$ *and* $\min_k n_k \to \infty$..

Proof. Let $\boldsymbol{B} = \left(\boldsymbol{I} - \tilde{\boldsymbol{B}}\right)^{-1} - \boldsymbol{I}$. Then alternatively we can write

$$\boldsymbol{\Gamma} = \boldsymbol{B}\cdot\boldsymbol{A} + \boldsymbol{A} \tag{7}$$

We only need to show that elements in $\boldsymbol{B}$ can be consistently estimated as by definition, $\tilde{\boldsymbol{B}} = \boldsymbol{I} - (\boldsymbol{I} + \boldsymbol{B})^{-1}$. Under the infinite GWAS data assumption, without loss of generality, we treat all γ_{kj} as directly observed.

First, we prove the following lemma.

Lemma 2. *Let* $(\boldsymbol{X}_j, y_j)$ *be independent random variables. Suppose*

$$y_j = \boldsymbol{\beta}^T\boldsymbol{X}_j + \epsilon_j,$$

where $\boldsymbol{X}_j \perp\!\!\!\perp \epsilon_j$*. Assume the following quantities exist:*

$$\mu_y := \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[y_j],$$

$$\boldsymbol{\mu}_X := \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\boldsymbol{X}_j],$$

$$\boldsymbol{S}_X := \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\boldsymbol{X}_j\boldsymbol{X}_j^T],$$

$$\boldsymbol{c} := \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\boldsymbol{X}_j y_j].$$

In addition, assume that

1. $\sup_j \mathbb{E}[\|\boldsymbol{X}_j\|_2^2 + \epsilon_j^2] < M$ *for some constant* M
2. $\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} |\mathbb{E}[\epsilon_j] - \mu_\epsilon| = 0$ *where* $\mu_\epsilon = \lim_{n\to\infty}\sum_{j=1}^{n}\mathbb{E}[\epsilon_j]/n$
3. $\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T$ *is full rank.*

Then

$$\boldsymbol{\beta} = -(\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}\boldsymbol{\mu}_X\mu_y + (\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}\boldsymbol{c}$$

is identifiable from the data when $n \to \infty$*.*

Proof of Lemma 2. First note we have

$$\begin{aligned}
\mu_y &= \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[y_j] \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\boldsymbol{X}_j]^T\boldsymbol{\beta} + \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\epsilon_j] \\
&= \boldsymbol{\mu}_X^T\boldsymbol{\beta} + \mu_\epsilon.
\end{aligned}$$

Additionally, since

$$\begin{aligned}
\boldsymbol{c} &= \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\boldsymbol{X}_j y_j] \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\boldsymbol{X}_j(\boldsymbol{X}_j^T\boldsymbol{\beta} + \epsilon_j)] \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \left(\mathbb{E}[\boldsymbol{X}_j\boldsymbol{X}_j^T]\boldsymbol{\beta} + \mathbb{E}[\boldsymbol{X}_j\epsilon_j]\right) \\
&= \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \left(\mathbb{E}[\boldsymbol{X}_j\boldsymbol{X}_j^T]\boldsymbol{\beta} + \mathbb{E}[\boldsymbol{X}_j]\mu_\epsilon + \mathbb{E}[\boldsymbol{X}_j]\mathbb{E}[\epsilon_j - \mu_\epsilon]\right) \\
&= \boldsymbol{S}_x\boldsymbol{\beta} + \boldsymbol{\mu}_X\mu_\epsilon + \lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\boldsymbol{X}_j]\mathbb{E}[(\epsilon_j - \mu_\epsilon)]
\end{aligned}$$

we obtain that

$$\begin{aligned}
\|\boldsymbol{c} - \boldsymbol{S}_x\boldsymbol{\beta} - \boldsymbol{\mu}_X\mu_\epsilon\|_1 &= \left\| \left(\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}[\boldsymbol{X}_j]\mathbb{E}[(\epsilon_j - \mu_\epsilon)] \right) \right\|_1 \\
&\le \left(\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} \mathbb{E}\left[\|\boldsymbol{X}_j\|_1\right] |\mathbb{E}[(\epsilon_j - \mu_\epsilon)]| \right) \\
&\le (1+M)\left(\lim_{n\to\infty} \frac{1}{n}\sum_{j=1}^{n} |\mathbb{E}[\epsilon_j] - \mu_\epsilon| \right) \\
&= 0,
\end{aligned}$$

where the inequality in the third line is due to $\mathbb{E}[\|X\|_1] \le 1 + \mathbb{E}[\|X\|_2^2] \le 1 + M$. Thus we have

$$\begin{aligned}
&-(\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}\boldsymbol{\mu}_X\mu_y + (\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}\boldsymbol{c} \\
=&-(\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}\boldsymbol{\mu}_X(\boldsymbol{\mu}_X^T\boldsymbol{\beta} + \mu_\epsilon) + (\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}(\boldsymbol{S}_X\boldsymbol{\beta} + \boldsymbol{\mu}_X\mu_\epsilon) \\
=&(\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}(\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)\boldsymbol{\beta} - (\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}\boldsymbol{\mu}_X\mu_\epsilon + (\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}\boldsymbol{\mu}_X\mu_\epsilon \\
=&\boldsymbol{\beta}
\end{aligned}$$

□

Now we prove Theorem 1 by induction. We first define the following notations. For any matrix $\boldsymbol{M}$, define

$$\boldsymbol{M}_{<k,j} := \left(M_{1j}, M_{2j}, \cdots, M_{(k-1)j}\right)^T \tag{8}$$

$$\boldsymbol{M}_{k,<j} := \left(M_{k1}, M_{k2}, \cdots, M_{k(j-1)}\right) \tag{9}$$

(similarly for $\boldsymbol{M}_{\le k,j}$ and $\boldsymbol{M}_{k,\le j}$). We shall show by induction that for $k = 1, 2, \ldots, K-1$, we have

1. $\boldsymbol{B}_{k,<k}$ can be consistently estimated

2.

$$\mu_k := \lim_{P\to\infty} \frac{1}{P}\sum_{j=1}^{P} \mathbb{E}[\alpha_{kj}],$$
$$v_k := \lim_{P\to\infty} \frac{1}{P}\sum_{j=1}^{P} \mathbb{E}[\alpha_{kj}^2]$$
$$\boldsymbol{c}_k := \lim_{P\to\infty} \frac{1}{P}\sum_{j=1}^{P} \mathbb{E}[\gamma_{(k+1)j}\boldsymbol{A}_{\leq k,j}] \in \mathbb{R}^k$$

exist and can be consistently estimated when $P \to \infty$.

Note the existence of μ_k and v_k are guaranteed by assumptions. So we will only show the existence of $\boldsymbol{c}_k$ in our induction.

Base case. When $k = 1$, we need to show $\mu_1, v_1, \boldsymbol{c}_1$ exist and can be consistently estimated. Here the existence of $\boldsymbol{c}_1$ is implied by the existence of v_1 and μ_2 once we observe that

$$\lim_{P\to\infty} \frac{1}{P}\sum_{i=1}^{P} \mathbb{E}[\gamma_{2j}\alpha_{1j}] = \lim_{P\to\infty} \frac{1}{P}\sum_{i=1}^{P} \mathbb{E}[\beta_{21}\alpha_{1j}\alpha_{1j} + \alpha_{2j}] = \beta_{21} \lim_{P\to\infty} \frac{1}{P}\sum_{i=1}^{P} \mathbb{E}[\alpha_{1j}^2] + \lim_{P\to\infty} \frac{1}{P}\sum_{i=1}^{P} \mathbb{E}[\alpha_{2j}]$$

Note here we have $\alpha_{1j} = \gamma_{1j}$ are directly observed. Therefore by law of large numbers, under the bounded moment assumption, we have

$$\frac{1}{P}\sum_{i=1}^{P} \alpha_{1j} \xrightarrow{p} \lim_{P\to\infty} \frac{1}{P}\sum_{i=1}^{P} \mathbb{E}[\alpha_{1j}] = \mu_1$$
$$\frac{1}{P}\sum_{i=1}^{P} \alpha_{1j}^2 \xrightarrow{p} \lim_{P\to\infty} \frac{1}{P}\sum_{i=1}^{P} \mathbb{E}[\alpha_{1j}^2] = v_1$$
$$\frac{1}{P}\sum_{i=1}^{P} \gamma_{2j}\alpha_{1j} \xrightarrow{p} \lim_{P\to\infty} \frac{1}{P}\sum_{i=1}^{P} \mathbb{E}[\gamma_{2j}\alpha_{1j}] = \boldsymbol{c}_1$$

Induction Step. Suppose 1, 2 hold for all $k \leq k_0 - 1$. Then we use Lemma 2 to show $\boldsymbol{B}_{k_0,<k_0}$ can be consistently estimated. Denote

$$\boldsymbol{\beta} = \boldsymbol{B}_{k_0,<k_0}^T$$
$$\boldsymbol{X}_j = (\alpha_{1j}, \alpha_{2j}, \cdots, \alpha_{(k_0-1)j})^T,$$
$$\epsilon_j = \alpha_{k_0 j},$$
$$y_j = \gamma_{k_0 j}.$$

To apply Lemma 2, we first check assumptions. Note

$$\boldsymbol{\mu}_X = \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\boldsymbol{X}_j] = \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}(\mathbb{E}[\alpha_{1j}],\mathbb{E}[\alpha_{2j}],\cdots,\mathbb{E}[\alpha_{(k_0-1)j}])^T = (\mu_1,\mu_2,\cdots,\mu_{k_0-1})^T$$

$$\mu_y = \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\gamma_{k_0j}] = \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\gamma_{k_0j}] = \boldsymbol{B}_{k_0,<k_0}^T\boldsymbol{\mu}_X + \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\alpha_{k_0j}] = \boldsymbol{\beta}^T\boldsymbol{\mu}_X + \mu_{k_0}$$

$$(\boldsymbol{S}_x)_{kl} = \left(\lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\boldsymbol{X}_j\boldsymbol{X}_j^T]\right)_{kl} = \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\alpha_{kj}\alpha_{lj}]$$

$$\boldsymbol{c} = \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\boldsymbol{X}_j y_j] = \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\boldsymbol{X}_j(\boldsymbol{X}_j^T\boldsymbol{\beta}+\alpha_{k_0j})]$$

$$= \boldsymbol{S}_x\boldsymbol{\beta} + \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}(\mathbb{E}[\alpha_{1j}\alpha_{k_0j}],\cdots,\mathbb{E}[\alpha_{(k_0-1)j}\alpha_{k_0j}])^T$$

By the induction hypothesis and our assumptions of existence of limits, all of these quantities exist. Moreover, the *bounded second moment assumption* in Lemma 2 is satisfied by our *bounded fourth moment assumption*. The condition $\lim_{n\to\infty}\frac{1}{n}\sum_{j=1}^{n}|\mathbb{E}[\epsilon_j]-\mu_\epsilon| = 0$ is directly assumed in the statement of the theorem. The full-rank assumption of $\boldsymbol{S}_x - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T$ is implied by the *non-zero instrumental effect* because

$$\boldsymbol{S}_x - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T = diag\left(\lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathrm{Var}(\alpha_{1j}),\cdots,\lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathrm{Var}(\alpha_{(k_0-1)j})\right)$$

To see this, note for any k, we have

$$\begin{aligned}\frac{1}{P}\sum_{j=1}^{P}\mathrm{Var}(\alpha_{kj}) &= \frac{1}{P}\sum_{j=1}^{P}\left(\mathbb{E}[\alpha_{kj}^2]-\mathbb{E}[\alpha_{kj}]^2\right)\\ &= \left(\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\alpha_{kj}^2]\right) - \left(\frac{1}{P}\sum_{j=1}^{P}(\mu_k+\mathbb{E}[\alpha_{kj}]-\mu_k)^2\right)\\ &= v_k - \mu_k^2 - 2\mu_k\left(\frac{1}{P}\sum_{j=1}^{P}(\mathbb{E}[\alpha_{kj}]-\mu_k)\right) - \left(\frac{1}{P}\sum_{j=1}^{P}(\mathbb{E}[\alpha_{kj}]-\mu_k)^2\right)\end{aligned}$$

Therefore

$$\begin{aligned}\left|\lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathrm{Var}(\alpha_{kj}) - v_k - \mu_k^2\right| &\le \left|\lim_{P\to\infty}2\mu_k\left(\frac{1}{P}\sum_{j=1}^{P}(\mathbb{E}[\alpha_{kj}]-\mu_k)\right)\right| + \left|\lim_{P\to\infty}\left(\frac{1}{P}\sum_{j=1}^{P}(\mathbb{E}[\alpha_{kj}]-\mu_k)^2\right)\right|\\ &\le 0 + \left(\lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}|\mathbb{E}[\alpha_{kj}]-\mu_k|\right)^2\\ &= 0\end{aligned}$$

Hence $\left(\boldsymbol{S}_x - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T\right)_{kk} = \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathrm{Var}(\alpha_{kj})$ for any k. One can use similar arguments to show the off-diagonals of $\left(\boldsymbol{S}_x - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T\right)$ are zeros.

Combining all the results we have, by Lemma 2,

$$\boldsymbol{B}_{k_0,<k_0}^T = \boldsymbol{\beta} = -(\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}\boldsymbol{\mu}_X\mu_y + (\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)^{-1}\boldsymbol{c}.$$

We can consistently estimate $\boldsymbol{B}_{k_0,<k_0}^T$ as long as $\boldsymbol{\mu}_X$, $\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T$, μ_y and $\boldsymbol{c}$ can be consistently estimated. Given their definition and the induction hypotheses, we can consistently estimate $\boldsymbol{\mu}_X, (\boldsymbol{S}_X - \boldsymbol{\mu}_X\boldsymbol{\mu}_X^T)$ and $\boldsymbol{c}$. In addition, by the law of large numbers, under the bounded moments condition, we have

$$\frac{1}{P}\sum_{i=1}^{P}\gamma_{k_0 j} \xrightarrow{p} \lim_{P\to\infty}\frac{1}{P}\sum_{i=1}^{P}\mathbb{E}[\gamma_{k_0 j}] = \mu_y$$

This shows that $\boldsymbol{B}_{k_0,<k_0}$ can be consistently estimated.

It remains to show $\mu_{k_0}, v_{k_0}, \boldsymbol{c}_{k_0}$ exist and can be consistently estimated. Here $\boldsymbol{c}_{k_0}$ exists as it is weights sum of moments of α's:

$$\begin{aligned}\boldsymbol{c}_{k_0} = \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\gamma_{(k_0+1)j}\boldsymbol{A}_{\le k_0 j}] &= \lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[(\boldsymbol{B}_{(k_0+1),\le k_0}\boldsymbol{A}_{\le k_0 j} + \alpha_{(k_0+1)j})\boldsymbol{A}_{\le k_0 j}] \\ &= \boldsymbol{B}_{(k_0+1),\le k_0}\lim_{P\to\infty}\frac{1}{P}\sum_{j=1}^{P}\mathbb{E}[\boldsymbol{A}_{\le k_0 j}\boldsymbol{A}_{\le k_0 j}^T] + \lim_{P\to\infty}\frac{1}{P}\mathbb{E}[\boldsymbol{A}_{\le k_0 j}\alpha_{(k_0+1)j}]\end{aligned}$$

Since $\boldsymbol{B}_{k_0,<k_0}$ can be consistently estimated, we know $\tilde{\boldsymbol{B}}_{k_0,<k_0}$ can also be consistently estimated. Let $\hat{\boldsymbol{B}}_{k_0,<k_0}$ be our consistent estimator of $\tilde{\boldsymbol{B}}_{k_0,<k_0}$. We know

$$\begin{aligned}\frac{1}{P}\sum_{i=1}^{P}\left(\gamma_{k_0 j} - \hat{\boldsymbol{B}}_{k_0,<k_0}^T\boldsymbol{\Gamma}_{<k_0,j}\right) &= \frac{1}{P}\sum_{i=1}^{P}\left(\gamma_{k_0 j} - \tilde{\boldsymbol{B}}_{k_0,<k_0}^T\boldsymbol{\Gamma}_{<k_0,j}\right) + o(1) \\ &\xrightarrow{p} \lim_{P\to\infty}\frac{1}{P}\sum_{i=1}^{P}\mathbb{E}[\alpha_{k_0 j}] = \mu_{k_0}\end{aligned}$$

$$\begin{aligned}\frac{1}{P}\sum_{i=1}^{P}\left(\gamma_{k_0 j} - \hat{\boldsymbol{B}}_{k_0,<k_0}^T\boldsymbol{\Gamma}_{<k_0,j}\right)^2 &= \frac{1}{P}\sum_{i=1}^{P}\left(\gamma_{k_0 j} - \tilde{\boldsymbol{B}}_{k_0,<k_0}^T\boldsymbol{\Gamma}_{<k_0,j}\right)^2 + o(1) \\ &\xrightarrow{p} \lim_{P\to\infty}\frac{1}{P}\sum_{i=1}^{P}\mathbb{E}[\alpha_{k_0 j}^2] = v_{k_0}\end{aligned}$$

And for any $k \le k_0$

$$\begin{aligned}\frac{1}{P}\sum_{i=1}^{P}\gamma_{k_0+1,j}\left(\gamma_{k,j} - \hat{\boldsymbol{B}}_{k,<k}^T\boldsymbol{\Gamma}_{<k,j}\right) &= \frac{1}{P}\sum_{i=1}^{P}\gamma_{k_0+1,j}\left(\gamma_{k,j} - \tilde{\boldsymbol{B}}_{k,<k}^T\boldsymbol{\Gamma}_{<k,j}\right) + o(1) \\ &\xrightarrow{p} \lim_{P\to\infty}\frac{1}{P}\sum_{i=1}^{P}\mathbb{E}[\gamma_{(k_0+1)j}\alpha_{kj}] = (\boldsymbol{c}_{k_0})_k\end{aligned}$$

This finishes our proof of induction. Now we have for any $k \in [K-1]$, $\boldsymbol{B}_{k,<k}, \mu_k, v_k, \boldsymbol{c}_k$ can be consistently estimated. Following the same proof in the induction step, by Lemma 2, we know $B_{K,<K}$ can be consistently estimated. This finishes the proof. □

S1.4 Details of the Gibbs Sampler

We observe GWAS summary statistics, $\hat{\gamma}_{kj}$, of marginal associations between SNPs and traits, with standard errors δ_{kj}. Denote $\boldsymbol{\Delta}_j = \mathrm{diag}(\delta_{1j}, \ldots \delta_{Kj})$. Denote the correlation matrix betwen trait-association for all SNPs by $\boldsymbol{R}$.

The model is

$$\hat{\boldsymbol{\Gamma}} = (\boldsymbol{I} + \boldsymbol{B})\boldsymbol{A} + \boldsymbol{\epsilon}$$

$$\boldsymbol{\epsilon_{*,j}} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_j := \boldsymbol{\Delta}_j \boldsymbol{R} \boldsymbol{\Delta_j}), \text{ with } \boldsymbol{\epsilon_{*,j}} \perp\!\!\!\perp \{\boldsymbol{\epsilon_{*,j'}} : j' \neq j\}$$

$$\begin{aligned}
\beta_{kl} &\overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2) \text{ for } k < l \\
\sigma^2 &\overset{\text{iid}}{\sim} \text{InvGamma}(\alpha^B, \beta^B)
\end{aligned}$$

$$\begin{aligned}
\alpha_{kj} &\overset{\text{iid}}{\sim} \begin{cases} \mathcal{N}(0, \sigma^2_{k0}), & \text{if } z_{kj} = 0 \\ \mathcal{N}(0, \sigma^2_{k1}), & \text{if } z_{kj} = 1 \end{cases} \\
\sigma_{k0} &\overset{\text{iid}}{\sim} \text{InvGamma}(\alpha^0_k, \beta^0_k) \\
\sigma_{k1} &\overset{\text{iid}}{\sim} \text{InvGamma}(\alpha^1_k, \beta^1_k) \\
z_{kj} &\overset{\text{iid}}{\sim} \text{Bernoulli}(p_k) \\
p_k &\overset{\text{iid}}{\sim} \text{Beta}(a_k, b_k)
\end{aligned}$$

We have K traits and P genes. The prior parameters are α^B, β^B, α^0_k, β^0_k, α^1_k, β^1_k, a_k and b_k. Also let

$$\begin{aligned}
\boldsymbol{Z} &= (z_{kj})_{K \times P} \\
\boldsymbol{p} &= (p_1, p_2, \ldots, p_K)^T \\
\boldsymbol{\sigma_0} &= (\sigma_{10}, \sigma_{20}, \ldots, \sigma_{K0})^T \\
\boldsymbol{\sigma_1} &= (\sigma_{11}, \sigma_{21}, \ldots, \sigma_{K1})^T
\end{aligned}$$

S1.4.1 Notation

For a $K \times P$ matrix $\boldsymbol{M}$, let

- $\boldsymbol{M}_{i,<k}$ denotes a row vector of all elements of $\boldsymbol{M}$ in the i^{th} row and in columns less than k
- $\boldsymbol{M}_{i,*}$ denotes a whole row, and $\boldsymbol{M}_{*,j}$ a column.
- $\text{vec}(\boldsymbol{M})$ denotes the vectorization of $\boldsymbol{M}$, i.e.,

$$\text{vec}(\boldsymbol{M}) = \begin{pmatrix} \boldsymbol{M}_{*,1} \\ \boldsymbol{M}_{*,2} \\ \vdots \\ \boldsymbol{M}_{*,P} \end{pmatrix}.$$

- For $K = P$, define $\text{LTvec}(\boldsymbol{M})$ to be the *lower-triangular vectorization* of $\boldsymbol{M}$ (excluding the diagonal):

$$\text{LTvec}(\boldsymbol{M}) := \begin{pmatrix} M_{2,1} \\ \boldsymbol{M}^T_{3,<3} \\ \vdots \\ \boldsymbol{M}^T_{K-1,<K-1} \\ \boldsymbol{M}^T_{K,<K} \end{pmatrix},$$

  which is a $\binom{K}{2}$-dimensional vector. Let $\boldsymbol{v}$ be a K dimensional vector:

- Define

$$\text{LTstack}(\boldsymbol{v}) := \begin{pmatrix} 0 & v_1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & v_1 & 0 & & \\ 0 & 0 & v_2 & 0 & & \\ 0 & 0 & 0 & v_1 & & \\ 0 & 0 & 0 & v_2 & & \\ 0 & 0 & 0 & v_3 & & \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & 0 & v_1 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & 0 & & v_{k-2} \\ 0 & 0 & 0 & 0 & \cdots & v_{k-1} \end{pmatrix} = \begin{pmatrix} 0 & \boldsymbol{v}_{<2} & 0 & & \cdots & 0 \\ \mathbf{0}_2 & \mathbf{0}_2 & \boldsymbol{v}_{<3} & & \cdots & \vdots \\ \vdots & \vdots & & \ddots & & \\ \mathbf{0}_{K-2} & \mathbf{0}_{K-2} & & & \boldsymbol{v}_{<K-1} & \mathbf{0}_{K-2} \\ \mathbf{0}_{K-1} & \mathbf{0}_{K-1} & \cdots & & \mathbf{0}_{K-1} & \boldsymbol{v}_{<K} \end{pmatrix},$$

  which has dimensions $\binom{K}{2} \times K$. Here $\mathbf{0}_k$ denotes the zero vector of dimension k. Hence, if $\boldsymbol{M}$ is square, lower-triangular with zero diagonal,

$$\boldsymbol{M}\boldsymbol{v} = \left[\text{LTvec}(\boldsymbol{M})^T \text{LTstack}(\boldsymbol{v})\right]^T.$$

  Also, for random variables U and V, $p(U|V)$ denotes the conditional density of U given V.

S1.4.2 Updating $\boldsymbol{B}$

With prior:

$$\text{LTvec}(\boldsymbol{B}) \sim N(\mathbf{0}, \sigma^2 \boldsymbol{I}), \tag{10}$$

likelihood (noise distribution):

$$\boldsymbol{\epsilon}_{*,j} \overset{\text{iid}}{\sim} N(\mathbf{0}, \boldsymbol{\Sigma}_j),$$

so that

$$\text{vec}(\boldsymbol{\epsilon}) \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma} := \begin{pmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 & \cdots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Sigma}_P \end{pmatrix}\right) \tag{11}$$

We have that

$$\begin{aligned} & \hat{\boldsymbol{\Gamma}} = \boldsymbol{A} + \boldsymbol{B}\boldsymbol{A} + \boldsymbol{\epsilon} \\ \implies & \hat{\boldsymbol{\Gamma}} - \boldsymbol{A} = \boldsymbol{B}\boldsymbol{A} + \boldsymbol{\epsilon} \\ \implies & \text{vec}(\hat{\boldsymbol{\Gamma}} - \boldsymbol{A})^T = \text{LTvec}(\boldsymbol{B})^T \tilde{\boldsymbol{A}} + \text{vec}(\boldsymbol{\epsilon})^T \\ \implies & \text{vec}(\hat{\boldsymbol{\Gamma}} - \boldsymbol{A}) = \tilde{\boldsymbol{A}}^T \text{LTvec}(\boldsymbol{B}) + \text{vec}(\boldsymbol{\epsilon}) \end{aligned}$$

Where

$$\tilde{\boldsymbol{A}} := \begin{pmatrix} \text{LTstack}(\boldsymbol{A}_{*,1}) & \text{LTstack}(\boldsymbol{A}_{*,2}) & \cdots & \text{LTstack}(\boldsymbol{A}_{*,P}) \end{pmatrix}$$

From Theorem 2.2 in [2], we know the posterior is

$$\begin{aligned} \text{LTvec}(\boldsymbol{B}) &\sim N(\boldsymbol{m}, \boldsymbol{C}), \\ \boldsymbol{C} &= (\tilde{\boldsymbol{A}}\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{A}}^T + \sigma^{-2}\boldsymbol{I})^{-1}, \\ \boldsymbol{m} &= \boldsymbol{C}\tilde{\boldsymbol{A}}\boldsymbol{\Sigma}^{-1}\,\text{vec}(\hat{\boldsymbol{\Gamma}} - \tilde{\boldsymbol{A}}). \end{aligned}$$

S1.4.3 Updating $\sigma^2, \boldsymbol{\sigma}_0, \boldsymbol{\sigma}_1$

First observe that conditioning on $\boldsymbol{B}$, we have

$$\sigma \perp\!\!\!\perp (\boldsymbol{A}, \boldsymbol{Z}, \boldsymbol{p}, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}, \hat{\boldsymbol{\Gamma}}) \mid \boldsymbol{B}$$

(as σ is independent with $\boldsymbol{A}, \boldsymbol{Z}, \boldsymbol{p}, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}$ and depends on $\hat{\boldsymbol{\Gamma}}$ only through $\boldsymbol{B}$)
If follows that

$$\begin{aligned} p(\sigma|\boldsymbol{B}, \boldsymbol{A}, \boldsymbol{Z}, \boldsymbol{p}, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}, \hat{\boldsymbol{\Gamma}}) &= \frac{p(\sigma, \boldsymbol{A}, \boldsymbol{Z}, \boldsymbol{p}, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}, \hat{\boldsymbol{\Gamma}}|\boldsymbol{B})}{p(\boldsymbol{A}, \boldsymbol{Z}, \boldsymbol{p}, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}, \hat{\boldsymbol{\Gamma}}|\boldsymbol{B})} \\ &= \frac{p(\boldsymbol{A}, \boldsymbol{Z}, \boldsymbol{p}, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}, \hat{\boldsymbol{\Gamma}}|\boldsymbol{B})p(\sigma|\boldsymbol{B})}{p(\boldsymbol{A}, \boldsymbol{Z}, \boldsymbol{p}, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}, \hat{\boldsymbol{\Gamma}}|\boldsymbol{B})} \\ &= p(\sigma|\boldsymbol{B}) \end{aligned}$$

So if suffices to consider $p(\sigma|\boldsymbol{B})$.

We have prior:

$$\sigma^2 \sim \text{InvGamma}(\alpha^B, \beta^B),$$

and likelihood:

$$\beta_{kl} \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2),$$

so we have posterior:

$$\sigma^2 \sim \text{InvGamma}\left(\alpha^B + \frac{n}{2},\ \beta^B + \frac{1}{2}\|\boldsymbol{B}\|_F^2\right),$$

where $n = K(K-1)/2$ is the number of degrees of freedom in $\boldsymbol{B}$.

Similarly for $\boldsymbol{\sigma_0}$ and $\boldsymbol{\sigma_1}$, we have

$$\boldsymbol{\sigma_0} \perp\!\!\!\perp (\boldsymbol{B}, \boldsymbol{p}, \sigma, \boldsymbol{\sigma_1}, \hat{\boldsymbol{\Gamma}}) \mid (\boldsymbol{A}, \boldsymbol{Z})$$

$$\boldsymbol{\sigma_1} \perp\!\!\!\perp (\boldsymbol{B}, \boldsymbol{p}, \sigma, \boldsymbol{\sigma_0}, \hat{\boldsymbol{\Gamma}}) \mid (\boldsymbol{A}, \boldsymbol{Z})$$

Thus

$$\begin{aligned} p(\boldsymbol{\sigma_0}|\boldsymbol{A}, \boldsymbol{Z}) &= p(\boldsymbol{\sigma_0}|\boldsymbol{A}, \boldsymbol{Z}, \boldsymbol{B}, \boldsymbol{p}, \sigma, \boldsymbol{\sigma_1}, \hat{\boldsymbol{\Gamma}}) \\ p(\boldsymbol{\sigma_1}|\boldsymbol{A}, \boldsymbol{Z}) &= p(\boldsymbol{\sigma_1}|\boldsymbol{A}, \boldsymbol{Z}, \boldsymbol{B}, \boldsymbol{p}, \sigma, \boldsymbol{\sigma_0}, \hat{\boldsymbol{\Gamma}}) \end{aligned}$$

Also for any $k \in [K]$, since σ_{k0} is independent of $(\boldsymbol{A}_{k',*}, \boldsymbol{Z}_{k',*})$ for $k' \neq k$, we have

$$p(\sigma_{k0}|\boldsymbol{A}, \boldsymbol{Z}) = p(\sigma_{k0}|\alpha_{k1}, \ldots, \alpha_{kp}, z_{k1}, \ldots, z_{kp}).$$

Similarly

$$p(\sigma_{k1}|\boldsymbol{A}, \boldsymbol{Z}) = p(\sigma_{k1}|\alpha_{k1}, \ldots, \alpha_{kp}, z_{k1}, \ldots, z_{kp}).$$

Note

$$\alpha_{kj} \sim \begin{cases} N(0, \sigma_{k0}^2) & \text{if } z_{kj} = 0, \\ N(0, \sigma_{k1}^2) & \text{if } z_{kj} = 1 \end{cases},$$

with prior,

$$\sigma_{kq}^2 \sim \text{InvGamma}(\alpha_k^q, \beta_k^q),$$

where $q \in \{0, 1\}$.

Then the posterior is,

$$\sigma_{kq}^2 \sim \text{InvGamma}\left(\alpha_k^q + \frac{1}{2}n_{kq},\ \beta_k^q + \frac{1}{2}\sum_{j=1}^{P} \alpha_{kj}^2 \mathbb{1}\{z_{kj} = q\}\right),$$

where $n_{iq} := \sum_{j=1}^{P} \mathbb{1}\{z_{kj} = q\}$.

S1.4.4 Updating p_k

Note

$$\boldsymbol{p} \perp\!\!\!\perp (\boldsymbol{B}, \sigma, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}, \boldsymbol{A}, \hat{\boldsymbol{\Gamma}}) \mid \boldsymbol{Z}$$

Therefore

$$p(\boldsymbol{p}|\boldsymbol{Z}) = p(\boldsymbol{p}|\boldsymbol{B}, \sigma, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}, \boldsymbol{A}, \hat{\boldsymbol{\Gamma}}, \boldsymbol{Z})$$

Since p_k is independent with $\boldsymbol{Z}_{k',*}$ for $k' \neq k$, we have

$$p(p_k|\boldsymbol{Z}) = p(p_k|z_{k1}, \ldots, z_{kp})$$

With prior

$$p_k \sim \text{Beta}(a_k, b_k),$$

and likelihood

$$z_{kj} \sim \text{Bernoulli}(p_k),$$

we have Posterior:

$$p_k \sim \text{Beta}(a_k + n_k,\ b_k + P - n_k),$$

where $n_k = \sum_{j=1}^{P} z_{kj}$.

S1.4.5 Updating $(\boldsymbol{A}, \boldsymbol{Z})$

To sample $(\boldsymbol{A}, \boldsymbol{Z})$ we first sample marginal of $\boldsymbol{Z}$ given everything but $\boldsymbol{A}$. Observe that

$$\begin{aligned} & p(\boldsymbol{Z}_{*,j}, \hat{\boldsymbol{\Gamma}}|\boldsymbol{B}, \boldsymbol{p}, \sigma, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}) \\ =& p(\boldsymbol{Z}_{*,j}, \hat{\boldsymbol{\Gamma}}_{*,j}|\boldsymbol{B}, \boldsymbol{p}, \sigma, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}) \cdot \left(\prod_{j' \neq j} p(\hat{\boldsymbol{\Gamma}}_{j'}|\boldsymbol{B}, \boldsymbol{p}, \sigma, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1} \right). \end{aligned}$$

Divide both sides by $p(\hat{\boldsymbol{\Gamma}})$. By independence among columns of $\hat{\boldsymbol{\Gamma}}$, we have

$$p(\boldsymbol{Z}_{*,j}|\hat{\boldsymbol{\Gamma}}, \boldsymbol{B}, \boldsymbol{p}, \sigma, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1}) = p(\boldsymbol{Z}_{*,j}|\hat{\boldsymbol{\Gamma}}_{*,j}, \boldsymbol{B}, \sigma, \boldsymbol{p}, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1})$$

Therefore we can just inspect each $\boldsymbol{Z}_{*,j}$ on its own. We have a likelihood given by:

$$\begin{aligned} \hat{\boldsymbol{\Gamma}}_{*,j}|\boldsymbol{Z}_{*,j}, \boldsymbol{\sigma_0}, \boldsymbol{\sigma_1} &= (\boldsymbol{I} + \boldsymbol{B})\boldsymbol{A}_{*,j} + \boldsymbol{\epsilon}_{*,j} \\ &\sim \mathcal{N}(\boldsymbol{0},\ (\boldsymbol{I} + \boldsymbol{B})\boldsymbol{\Sigma}_{Z_{*,j}}(\boldsymbol{I} + \boldsymbol{B})^T + \boldsymbol{\Sigma}_j), \end{aligned} \tag{12}$$

where

$$\boldsymbol{\Sigma}_{Z_{*,j}} := \text{diag}(\sigma_{1z_{1j}}, \ldots, \sigma_{Kz_{Kj}}),$$

and we have prior

$$z_{kj} \overset{\text{iid}}{\sim} \text{Bernoulli}(p_k).$$

In cases where K is small enough, we can easily compute the un-normalized probability weights for all possibilities of Z_j, and sample from a discrete distribution.

Now that we have a sample from the marginal of $\boldsymbol{Z}$ we can sample from the conditional of $\boldsymbol{A}$, which has prior of

$$\boldsymbol{A}_{*,j} \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_{Z_{*,j}}). \tag{13}$$

Recall that we have the linear Guassian model,

$$\hat{\boldsymbol{\Gamma}}_{*,j} = (\boldsymbol{I} + \boldsymbol{B})\boldsymbol{A}_{*,j} + \boldsymbol{\epsilon}_{*,j}$$

with noise distribution,

$$\boldsymbol{\epsilon}_{*,j}^T \sim \mathcal{N}(\boldsymbol{0}, \boldsymbol{\Sigma}_j),$$

so we get a posterior similar to before of

$$\begin{aligned}\boldsymbol{A}_{*,j}|\hat{\boldsymbol{\Gamma}}_{*,j}, \boldsymbol{B}, \boldsymbol{\sigma}_0, \boldsymbol{\sigma}_1 &\sim \mathcal{N}(\boldsymbol{m}, \boldsymbol{c}),\\ \boldsymbol{C} &= [(\boldsymbol{I}+\boldsymbol{B})^T \boldsymbol{\Sigma}_j^{-1}(\boldsymbol{I}+\boldsymbol{B}) + \boldsymbol{\Sigma}_{\boldsymbol{Z}_{*,j}}^{-1}]^{-1},\\ \boldsymbol{m} &= \boldsymbol{C}(\boldsymbol{I}+\boldsymbol{B})^T \boldsymbol{\Sigma}_j^{-1} \hat{\boldsymbol{\Gamma}}_{*,j}.\end{aligned}$$

S1.5 Estimation for hyper-parameters

In our empirical studies, we find that the posterior distributions of B can be sensitive to the choice of the hyper-parameters $(\alpha_k^0, \beta_k^0), (\alpha_k^1, \beta_k^1)$ and (a_k, b_k). Here we take an empirical based approach to choose the hyper-parameters. Specifically, these hyperparameters will be set to the corresponding maximum likelihood estimators in the marginalized model, which can be obtained by the expectation-maximization (EM) algorithm. Recall that our model assumes for $k \in \{1, 2, \ldots, K\}$,

$$\alpha_{kj} \sim (1-p_k)\mathcal{N}(0, \sigma_{k0}^2) + p_k \mathcal{N}(0, \sigma_{k1}^2) \tag{14}$$

In particular,

$$\hat{\gamma}_{1j} = \alpha_{1j} + \epsilon_{1j} \sim (1-p_1)N(0, \sigma_{10}^2 + \delta_{1j}^2) + p_1 \mathcal{N}(0, \sigma_{11}^2 + \delta_{1j}^2) \tag{15}$$

We shall first estimate the hyper-parameters for the outcome pleiotropies. In the E step, we compute the posterior distribution of the latent variables given our current estimates of parameters. In the M step we solve the maximization problem of the expectation of the log likelihood over the latent variables. To get rid of the identifiability issue, we initialize σ_{10}^2 and σ_{11}^2 to be far from each other. When solving for σ_{10}^2 and σ_{11}^2 in the M-step, we put the restriction that they can be at most 10 times larger than the their values from last iteration. Once we have obtained the point estimates of $(p_1, \sigma_{10}, \sigma_{11})$, we choose the hyper-parameters for their priors such that the prior means are equal to the points estimates.

For $k \in \{2, 3, \ldots, K\}$, however, this is theoretically unachievable because $\hat{\alpha}_{kj} \equiv \alpha_{kj} + \epsilon_{kj}$ are no longer accessible. In practice we set $\hat{\alpha}_{kj}$ to $\hat{\gamma}_{kj}$ to determine the hyper-parameters. The priors will be biased especially when the genetic effects of a trait is largely mediated by an upstream trait that is already in the model. Nonetheless, our numerical simulations suggest that this bias may be small in most reasonable scenarios. We point out that one possible alternative approach is to use this as a first degree approximation and then use the Gibbs sampler to estimate $\boldsymbol{B}$ and then use that to estimate the direct genetic association, and repeat the expectation maximization. In the optimal settings, if the estimates of $\boldsymbol{B}$ in the first stage is close to the true value, this two-stage procedure will improve the precision of the algorithm. Indeed it turns out that this second stage estimation can usually be quite helpful in univariate Mendelian Randomization. On the other hand, the step of estimating the direct genetic association aggregates the noise from multiple exposures and makes it harder for EM algorithm to find meaning and reliable estimates for the hyperparameters. Therefore the two-state procedure is only recommended when the number of traits are small and the noise levels are fair.

S1.5.1 EM details

Recall that our model assumes for $k \in \{1, 2, \ldots, K\}$,

$$\alpha_{kj} \sim (1-p_k)\mathcal{N}(0, \sigma_{k0}^2) + p_k \mathcal{N}(0, \sigma_{k1}^2)$$

In particular,

$$\hat{\gamma}_{1j} = \alpha_{1j} + \epsilon_{1j} \sim (1-p_1)N(0, \sigma_{10}^2 + \delta_{1j}^2) + p_1 \mathcal{N}(0, \sigma_{11}^2 + \delta_{1j}^2)$$

For simplicity, we will drop the subscript 1 for the time being.

E-step. Denote $\theta^{(t)} = (p^{(t)}, \sigma_1^{(t)}, \sigma_0^{(t)})$ .With latent variable $z_j \sim \text{Bernoulli}(p)$, where $\hat{\gamma}_j$ comes from the first distribution when $z_j = 1$, the membership probabilities are

$$w_j \equiv \mathbb{P}(z_j = 1|\theta^{(t)}) = \frac{p^{(t)}\phi(\hat{\gamma}_j; 0, (\sigma_1^{(t)})^2 + \delta_j^2)}{p^{(t)}\phi(\hat{\gamma}_j; 0, (\sigma_1^{(t)})^2 + \delta_j^2) + (1-p^{(t)})\phi(\hat{\gamma}_j; 0, (\sigma_0^{(t)})^2 + \delta_j^2)},$$

where $\phi(;\mu,\sigma)$ is the density of normal distribution with mean μ and standard deviation σ. Then we can compute the expectation of the data-log-likelihood with respect to the distribution of $\boldsymbol{Z}$ given the data and $\theta^{(t)}$:

$$\begin{aligned}
\mathbb{E}_{\boldsymbol{Z}}\left[\log L(\theta)|\hat{\gamma}_j, \theta^{(t)}\right] &= \textstyle\sum_{j=1}^{P}\Big(w_j \log\left[p\phi(\hat{\gamma}_j; 0, \sigma_1^2 + \delta_j^2)\right] \\
&\quad + (1-w_j)\log\left[(1-p)\phi(\hat{\gamma}_j; 0, \sigma_0^2 + \delta_j^2)\right]\Big) \\
&= \textstyle\sum_{j=1}^{P}\Big(w_j\left[\log(p) - \frac{1}{2}\log(\delta_j^2 + \sigma_1^2) - \frac{\hat{\gamma}_j^2}{2(\delta_j^2+\sigma_1^2)}\right] \\
&\quad + (1-w_j)\left[\log(1-p) - \frac{1}{2}\log(\delta_j^2 + \sigma_0^2) - \frac{\hat{\gamma}_j^2}{2(\delta_j^2+\sigma_0^2)}\right]\Big) + C,
\end{aligned}$$

where C is a constant that does not depend on θ.

M-step. Taking derivative with respect to p and setting it to zero yields

$$\sum_{j=1}^{P}\left(\frac{w_j}{p} - \frac{(1-w_j)}{1-p}\right) = 0 \implies p^{(t+1)} = \frac{\sum_{j=1}^{P} w_j}{P}$$

For σ_1^2, it is equivalent to minimize

$$\sum_{j=1}^{P}\left(\log(\delta_j^2 + \sigma_1^2) + \frac{\hat{\Gamma}_j^2}{(\delta_j^2 + \sigma_1^2)}\right),$$

and similarly for σ_0^2 .The objectives can be maximized with a numerical optimization method.

S1.6 Extension to multiple traits at each time point

In this section we propose an extension of the Bayesian framework which allows multiple traits at a single time stage. Specifically, suppose there are K time stages and N traits in total. At each time stage k, there are n_k exposures/outcomes of interests $X_{k1}, \ldots, X_{kn_k}$, with no internal interactions. Define $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_K$ as

$$\boldsymbol{X}_1 = \begin{pmatrix} X_{11} \\ X_{12} \\ \vdots \\ X_{1n_1} \end{pmatrix}, \boldsymbol{X}_2 = \begin{pmatrix} X_{21} \\ X_{22} \\ \vdots \\ X_{2n_2} \end{pmatrix}, \boldsymbol{X}_3 = \begin{pmatrix} X_{31} \\ X_{32} \\ \vdots \\ X_{3n_3} \end{pmatrix}, \cdots, \boldsymbol{X}_K = \begin{pmatrix} X_{K1} \\ X_{K2} \\ \vdots \\ X_{Kn_K} \end{pmatrix}.$$

Let

$$\gamma_{ki,j} = \text{argmin}_\gamma \text{Var}\left[X_{ki} - \gamma Z_j\right]$$

be the marginal association between X_{ki} and Z_j. Define $\boldsymbol{\gamma}_{1,j}, \ldots, \boldsymbol{\gamma}_{K,j}$ as

$$\boldsymbol{\gamma}_{1,j} = \begin{pmatrix} \gamma_{11,j} \\ \gamma_{12,j} \\ \vdots \\ \gamma_{1n_1,j} \end{pmatrix}, \boldsymbol{\gamma}_{2,j} = \begin{pmatrix} \gamma_{21,j} \\ \gamma_{22,j} \\ \vdots \\ \gamma_{2n_2,j} \end{pmatrix}, \boldsymbol{\gamma}_{3,j} = \begin{pmatrix} \gamma_{31,j} \\ \gamma_{32,j} \\ \vdots \\ \gamma_{3n_3,j} \end{pmatrix}, \cdots, \boldsymbol{\gamma}_{K,j} = \begin{pmatrix} \gamma_{K1,j} \\ \gamma_{K2,j} \\ \vdots \\ \gamma_{Kn_K,j} \end{pmatrix}$$

Similarly define $\boldsymbol{\alpha}_{1,j}, \ldots, \boldsymbol{\alpha}_{K,j}$ as

$$\boldsymbol{\alpha}_{1,j} = \begin{pmatrix} \alpha_{11,j} \\ \alpha_{12,j} \\ \vdots \\ \alpha_{1n_1,j} \end{pmatrix}, \boldsymbol{\alpha}_{2,j} = \begin{pmatrix} \alpha_{21,j} \\ \alpha_{22,j} \\ \vdots \\ \alpha_{2n_2,j} \end{pmatrix}, \boldsymbol{\alpha}_{3,j} = \begin{pmatrix} \alpha_{31,j} \\ \alpha_{32,j} \\ \vdots \\ \alpha_{3n_3,j} \end{pmatrix}, \cdots, \boldsymbol{\alpha}_{K,j} = \begin{pmatrix} \alpha_{K1,j} \\ \alpha_{K2,j} \\ \vdots \\ \alpha_{Kn_K,j} \end{pmatrix}$$

Let

$$\boldsymbol{\Gamma} := \begin{bmatrix} \gamma_{1,1} & \gamma_{1,2} & \cdots & \gamma_{1,P} \\ \vdots & \vdots & \vdots & \vdots \\ \gamma_{K,1} & \gamma_{K,2} & \cdots & \gamma_{K,P} \end{bmatrix} \in \mathbb{R}^{N \times P}$$

$$\boldsymbol{A} := \begin{bmatrix} \boldsymbol{\alpha}_{1,1} & \boldsymbol{\alpha}_{1,2} & \cdots & \boldsymbol{\alpha}_{1,P} \\ \vdots & \vdots & \vdots & \vdots \\ \boldsymbol{\alpha}_{K,1} & \boldsymbol{\alpha}_{K,2} & \cdots & \boldsymbol{\alpha}_{K,P} \end{bmatrix} \in \mathbb{R}^{N \times P}$$

Similar to the previous setting, from the individual level equations we have

$$\boldsymbol{\Gamma} = \tilde{\boldsymbol{B}}\boldsymbol{\Gamma} + \boldsymbol{A}, \tag{16}$$

where $\tilde{\boldsymbol{B}}$ is the matrix of coefficients. Furthermore, since there is no interactions between X_{ki}'s within each time stage k, we the coefficient matrix can be written as

$$\tilde{\boldsymbol{B}} = \begin{bmatrix} \mathbf{0}_{\boldsymbol{n_1}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \tilde{\boldsymbol{B}}_2 & \mathbf{0}_{\boldsymbol{n_2}} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ & \tilde{\boldsymbol{B}}_3 & \mathbf{0}_{\boldsymbol{n_3}} & \mathbf{0} & \ldots & \mathbf{0} \\ & & \tilde{\boldsymbol{B}}_4 & \mathbf{0}_{\boldsymbol{n_4}} & \ldots & \mathbf{0} \\ & & \ldots & & \ddots & \mathbf{0} \\ & & & \tilde{\boldsymbol{B}}_K & & \mathbf{0}_{\boldsymbol{n_K}} \end{bmatrix},$$

where $\mathbf{0}_{\boldsymbol{n_k}} \in \mathbb{R}^{n_k \times n_k}$ represents the square matrix of zeros, $\tilde{\boldsymbol{B}}_k$ represents the matrix of marginal associations between the $\boldsymbol{X}_k$ and $(\boldsymbol{X}_l)_{1 \leq l < k}$.

Equivalently by defining $\boldsymbol{B} = (\boldsymbol{I} - \tilde{\boldsymbol{B}})^{-1} - \boldsymbol{I}$ we can write

$$\boldsymbol{\Gamma} = (\boldsymbol{I} + \boldsymbol{B})\boldsymbol{A} \tag{17}$$

$$\hat{\boldsymbol{\Gamma}} = (\boldsymbol{I} + \boldsymbol{B})\boldsymbol{A} + \boldsymbol{\epsilon}, \tag{18}$$

where $\boldsymbol{B}$ is of the form

$$\boldsymbol{B} = \begin{bmatrix} \mathbf{0}_{\boldsymbol{n_1}} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ \boldsymbol{B}_2 & \mathbf{0}_{\boldsymbol{n_2}} & \mathbf{0} & \mathbf{0} & \ldots & \mathbf{0} \\ & \boldsymbol{B}_3 & \mathbf{0}_{\boldsymbol{n_3}} & \mathbf{0} & \ldots & \mathbf{0} \\ & & \boldsymbol{B}_4 & \mathbf{0}_{\boldsymbol{n_4}} & \ldots & \mathbf{0} \\ & & \ldots & & \ddots & \mathbf{0} \\ & & & \boldsymbol{B}_K & & \mathbf{0}_{\boldsymbol{n_K}} \end{bmatrix}.$$

This is a direct conclusion from the fact that

$$\left(\boldsymbol{I} - \tilde{\boldsymbol{B}}\right)^{-1} - \boldsymbol{I} = \boldsymbol{I} + \sum_{k=1}^{\infty} \tilde{\boldsymbol{B}}^k - \boldsymbol{I} = \sum_{k=1}^{\infty} \tilde{\boldsymbol{B}}^k$$

S1.6.1 Adaption of Gibbs Sampler

The implementation of the Gibbs sampler for the multivariate version is identical to the previous version except for the updates of $\boldsymbol{B}$, since we impose addition structure on $\boldsymbol{B}$. To preceed, we make the following definitions:

- Define $N_k = \sum_{i=1}^{k} n_i$ for $k = 1, 2, \ldots, K$.
- For any $i \in \{n_1 + 1, 2, \ldots, N_k\}$, define $f(i)$ to be the unique integer k in $\{1, 2, \ldots, K\}$ such that $N_{k-1} < i \leq N_k$. Define $g(i) = N_{f(i)}$.
- For any matrix $\boldsymbol{B} \in \mathbb{R}^{N \times N}$, given $n_1, \ldots, n_k$, define

$$\text{BLvec}(\boldsymbol{B}) := \begin{pmatrix} B_{N_1+1,1} \\ B_{N_1+1,2} \\ \vdots \\ B_{N_1+1,g(N_1+1)} \\ B_{N_1+2,1} \\ B_{N_1+2,2} \\ \vdots \\ B_{N_1+2,g(N_1+2)} \\ \vdots \\ B_{N,1} \\ B_{N,2} \\ \vdots \\ B_{N,g(N)} \end{pmatrix}$$

- For any $\boldsymbol{v} \in \mathbb{R}^N$, define

$$\text{BLstack}(\boldsymbol{v}) = \begin{pmatrix} \mathbf{0}_{n_1}^T & v_1 & 0 & \cdots & 0 \\ \mathbf{0}_{n_1}^T & v_2 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & 0 \\ \mathbf{0}_{n_1}^T & v_{g(N_1+1)} & 0 & \cdots & 0 \\ \mathbf{0}_{n_1}^T & 0 & v_1 & \cdots & 0 \\ \mathbf{0}_{n_1}^T & 0 & v_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0}_{n_1}^T & 0 & v_{g(N_1+2)} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0}_{n_1}^T & 0 & 0 & \cdots & v_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{0}_{n_1}^T & 0 & 0 & \cdots & v_{g(N)} \end{pmatrix}$$

Hence, we have

$$\boldsymbol{B}\boldsymbol{v} = [\text{BLvec}(\boldsymbol{B})^T \text{BLstack}(\boldsymbol{v})]^T$$

Then we show how to update $\boldsymbol{B}$. With prior

$$\text{BLvec}(\boldsymbol{B}) \sim \mathcal{N}(\mathbf{0}, \sigma^2 \boldsymbol{I})$$

likelihood

$$\boldsymbol{\epsilon}_{*,j} \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_j)$$

and

$$\mathrm{vec}(\boldsymbol{\epsilon}) \sim \mathcal{N}\left(0, \begin{pmatrix} \boldsymbol{\Sigma}_1 & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_2 & \cdots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \boldsymbol{\Sigma}_P \end{pmatrix}\right),$$

we have

$$\begin{aligned} \hat{\boldsymbol{\Gamma}} &= (\boldsymbol{I} + \boldsymbol{B})\boldsymbol{A} + \boldsymbol{\epsilon} \\ \hat{\boldsymbol{\Gamma}} - \boldsymbol{A} &= \boldsymbol{B}\boldsymbol{A} + \boldsymbol{\epsilon} \\ \mathrm{vec}(\hat{\boldsymbol{\Gamma}} - \boldsymbol{A})^T &= \mathrm{BLvec}(\boldsymbol{B})^T \tilde{\boldsymbol{A}} + \mathrm{vec}(\boldsymbol{\epsilon})^T \\ \mathrm{vec}(\hat{\boldsymbol{\Gamma}} - \boldsymbol{A}) &= \tilde{\boldsymbol{A}}^T \mathrm{BLvec}(\boldsymbol{B}) + \mathrm{vec}(\boldsymbol{\epsilon}), \end{aligned}$$

where

$$\tilde{\boldsymbol{A}} = \begin{pmatrix} \mathrm{BLstack}(\boldsymbol{A}_{*,1}) & \mathrm{BLstack}(\boldsymbol{A}_{*,2}) & \cdots & \mathrm{BLstack}(\boldsymbol{A}_{*,P}) \end{pmatrix}$$

So

$$\begin{aligned} \mathrm{BLvec}(\boldsymbol{B}) &\sim \mathcal{N}(\boldsymbol{m}, \boldsymbol{C}) \\ \boldsymbol{C} &= (\tilde{\boldsymbol{A}}\boldsymbol{\Sigma}^{-1}\tilde{\boldsymbol{A}}^T + \sigma^{-2}\boldsymbol{I})^{-1} \\ \boldsymbol{m} &= \boldsymbol{c}\tilde{\boldsymbol{A}}\boldsymbol{\Sigma}^{-1}\mathrm{vec}(\hat{\boldsymbol{\Gamma}} - \boldsymbol{A}) \end{aligned}$$

The other parts are identical to the previous case.

S2 Additional details of the simulation setup and real data analysis

S2.1 Simulation setup

Our simulation is based on multiple real GWAS summary statistics datasets. For the $K = 3$ scenario, we select SNPs using data from GIANT Adult BMI and EGG childhood BMI GWAS datasets [3]. The chosen SNPs are then associated with traits including 8-year-old BMI from MoBa, adult BMI from UK Biobank, and Type II diabetes (T2D) from DIAGRAM to generate "true" direct associations with the exposure and outcome traits. Extending to $K = 4$, the same SNPs and traits are retained, with an additional inclusion of the 1-year-old BMI GWAS trait from MoBa as the exposure trait. In the simulation scenario involving multivariate traits at each time point, for SNP selection we emply summary statistics from low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), and triglycerides in Genetic Epidemiology Research on Adult Health and Aging (GERA) [4]. The simulation comprises three time points: a first stage with childhood traits with LDL-C, HDL-C, and TG summary statistics from [5]; a second stage with adulthood traits with LDL-C, HDL-C, and TG summary statistics from [6]; and a third stage with the stroke trait as the outcome, utilizing summary statistics from [7].

To generate GWAS summary data for simulation, we first select SNPs based on the selection files with LD clumping and a p-value cutoff at 0.01. For each selected SNP j and each exposure k, we record the estimated effects $\hat{\gamma}_{kj}^{real}$ along with their corresponding standard errors $\hat{\delta}_{kj}$ in the exposure and outcome files. Subsequently, we set "true" direct associations $\alpha_{kj} = \mathrm{sign}(\hat{\gamma}_{kj}^{real})(|\hat{\gamma}_{kj}^{real}| - 0.01)_+$, and we perform a random shuffle of α_{kj} across j within each trait k to ensure independence across traits. Once the matrix $\boldsymbol{A}$ is generated, we create $\boldsymbol{\Gamma}$ following our linear model with a pre-specified matrix $\boldsymbol{B}$. In the simulation, the matrix $\boldsymbol{B}$ is typically designed as an approximation of the true relationships among the traits in the real data. Finally, we simulate our synthetic summary statistics $\hat{\gamma}_{kj}^{simu} \sim \mathcal{N}(\gamma_{kj}, \hat{\delta}_{kj}^2)$ independently across all SNPs and traits. The results of experiments with $K = 3$ and $K = 4$ are based on 500 replications, while for $K = 7$, we conduct 100 replications.

At each replication, in order to mimic the real case, we would like to run the algorithms with p-values cutoffs for the selection of SNPs. Unlike the real cases, where independent selection files are usually utilized to compute the instrumental strength without selection bias, we directly set p-values for each SNP by computing p-values from $\gamma_{kj}/\hat{\delta}_{kj}$ and taking the minimum of p-values across k. This can be heuristically considered as if we have independent selection files with no measurement errors. Specially, the p-values for each SNP j is computed by

$$\min_{k<K}\left(2\cdot\Phi^{-1}(-|\gamma_{kj}/\hat{\delta}_{kj}|)\right),$$

where Φ is the Gaussian cumulative distribution function. Bonferroni correction is then applied to select the significant SNPs under different p-values cutoffs.

S2.2 Calculation of genetic correlation

We apply the LDSC package developed in [8] to compute the genetic correlation between each pair of phenotypes. Here we briefly review the calculation of genetic correlation via LD Score regression [9]. Let X_k and X_l be two phenotypes of interests and S be a set of P SNPs $\boldsymbol{Z}=(Z_j)_{i=1}^P$. Define

$$\boldsymbol{\beta}_k^{cor}=\text{argmax}_{\boldsymbol{\alpha}\in\mathbb{R}^P}\text{Cor}(X_k,\boldsymbol{\alpha}^T\boldsymbol{Z}).$$

Define $h_S^2(X_k)$, the heritability explained by SNPs in S, as

$$h_S^2(X_k)=\|\boldsymbol{\beta}_k^{cor}\|_2^2,$$

and $e_S(X_k,X_l)$, the genetic covariance among SNPs in S, as

$$e_S(X_k,X_l)=(\boldsymbol{\beta}_k^{cor})^T\boldsymbol{\beta}_l^{cor}.$$

The genetic correlation between X_k and X_l is then defined as

$$r_S(X_k,X_l)=\frac{e_S(X_k,X_l)}{\sqrt{h_S^2(X_k)h_S^2(X_l)}}.$$

The estimation of genetic correlation involves the estimation of both heritability and genetic correlation. The heritability can be estimated via the single-trait LD Score regressions from [8]. The single-trait LD Score regression equation of phenotype X_k is

$$\mathbb{E}[\chi_j^2\mid l_j]=N_kh^2l_j/P+N_ka+1,$$

where χ_j^2 is the expected χ^2-statistic of variant j, N_k is the sample size, P is the number of SNPs, l_j is the LD-score, a measures the contribution of confounding biases, and h^2 is the average heritability. The cross-trait LD regression can be applied to estimate the genetic covariance. The cross-trait LD Score regression equation is

$$\mathbb{E}[\mathtt{z}_{kj}\mathtt{z}_{lj}\mid l_j]=\frac{\sqrt{N_kN_l}e_g}{P}l_j+\frac{eN_{kl}}{\sqrt{N_kN_l}},$$

where $\mathtt{z}_{kj}$ denotes the z-score for study k and SNP j, N_k is the sample size for study k, e_g is the genetic covariance, l_j is the LD Score, N_{kl} is the number of individuals included in both studies and e is the phenotypic correlation among the overlapping samples. Then the genetic covariance can be estimated by regressing $\mathtt{z}_{kj}\mathtt{z}_{lj}$ against $\frac{\sqrt{N_kN_l}}{P}l_j$.

The choices of weights of LD score regression are discussed in the details in [9]. The assessment of statistical significance is then conducted by block jackknife.

S3 Supplemental figures

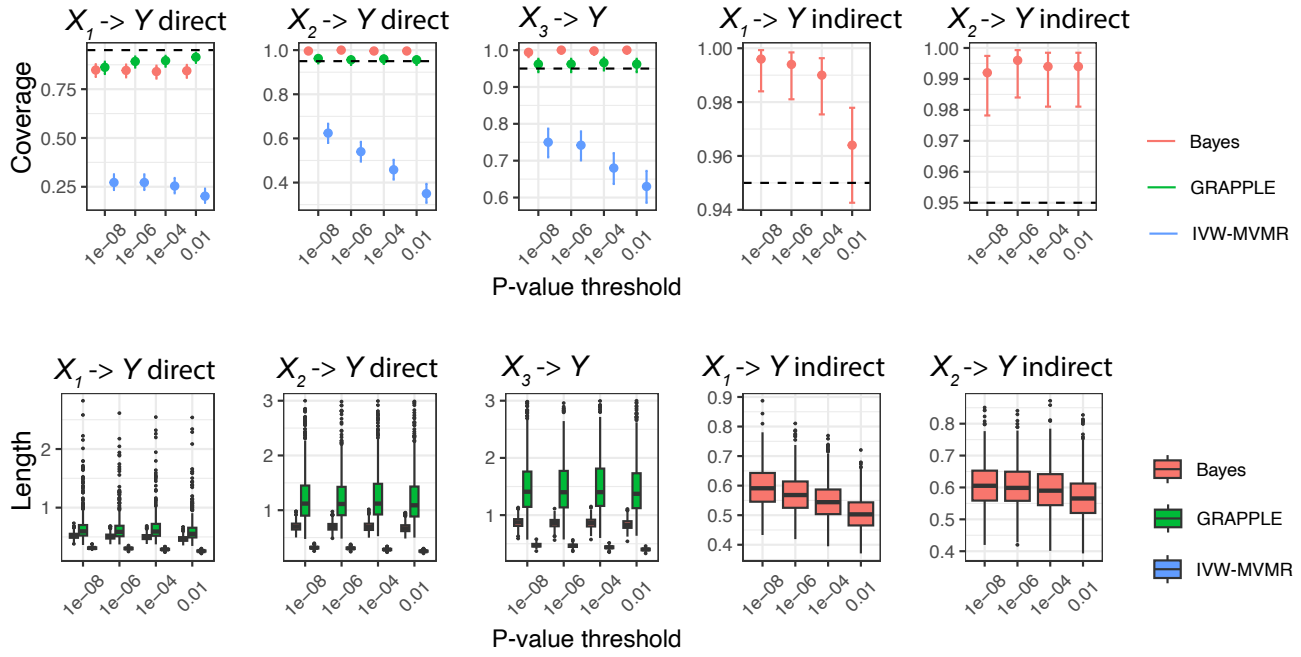

Figure S1: Simulation results for $K = 4$. The first row illustrates empirical coverage of 95% confidence intervals of all direct/indirect effects of the exposures on Y over 500 repeated simulations. The second row displays the boxplots of lengths of confidence intervals over repeated simulations.

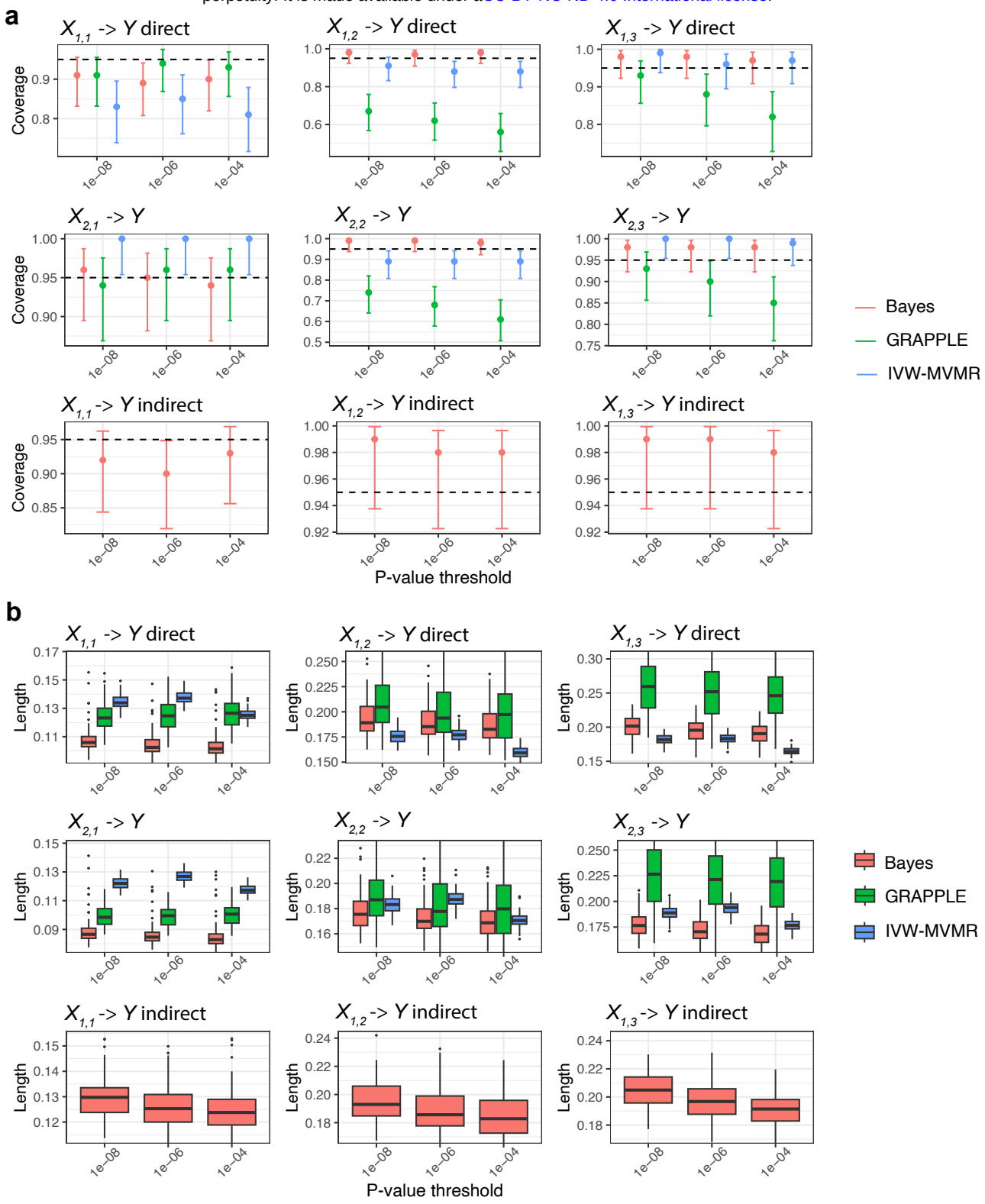

Figure S2: Simulation results for the multivariate case. a) Empirical coverage of 95% confidence intervals of all direct/indirect effects of the exposures on Y over 500 repeated simulations. b) Boxplots of lengths of confidence intervals over repeated simulations.

Conditional F statistics

a

selection p-values	1e-8	1e-6	1e-4	1e-2
early-life Body Size (UK Biobank)	8.53	7.74	4.98	3.22
Adult BMI	8.86	8.83	5.79	3.92

b

selection p-values	1e-8	1e-6	1e-4	1e-2
8-year-old BMI	0.99	1.21	1.22	1.07
Adult BMI	1.50	3.17	5.36	9.07

Figure S3: Conditional F statistics. a) Conditional F statistics of the exposure traits when we use childhood body size from UK Biobank and adult BMI from GIANT consortium as exposure GWAS datasets b) Conditional F statistics of the exposure traits when we use 8-year-old BMI from MoBa and adult BMI from GIANT consortium as exposure GWAS datasets. Notice that the selection files are the same, thus both scenarios have the same set of SNPs at any given selection p-value threshold.

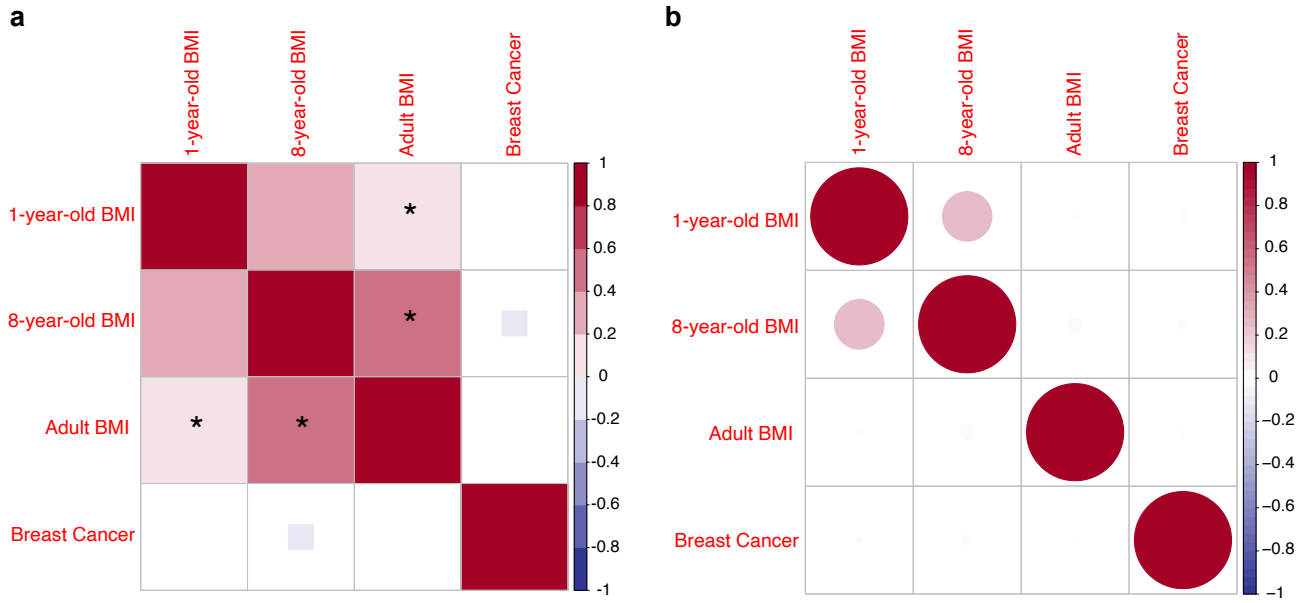

Figure S4: Additional results for the breast cancer case study. a) Estimated pairwise genetic correlations across traits. A '*' indicates significantly correlated pairs. b) Estimated noise correlation across traits.

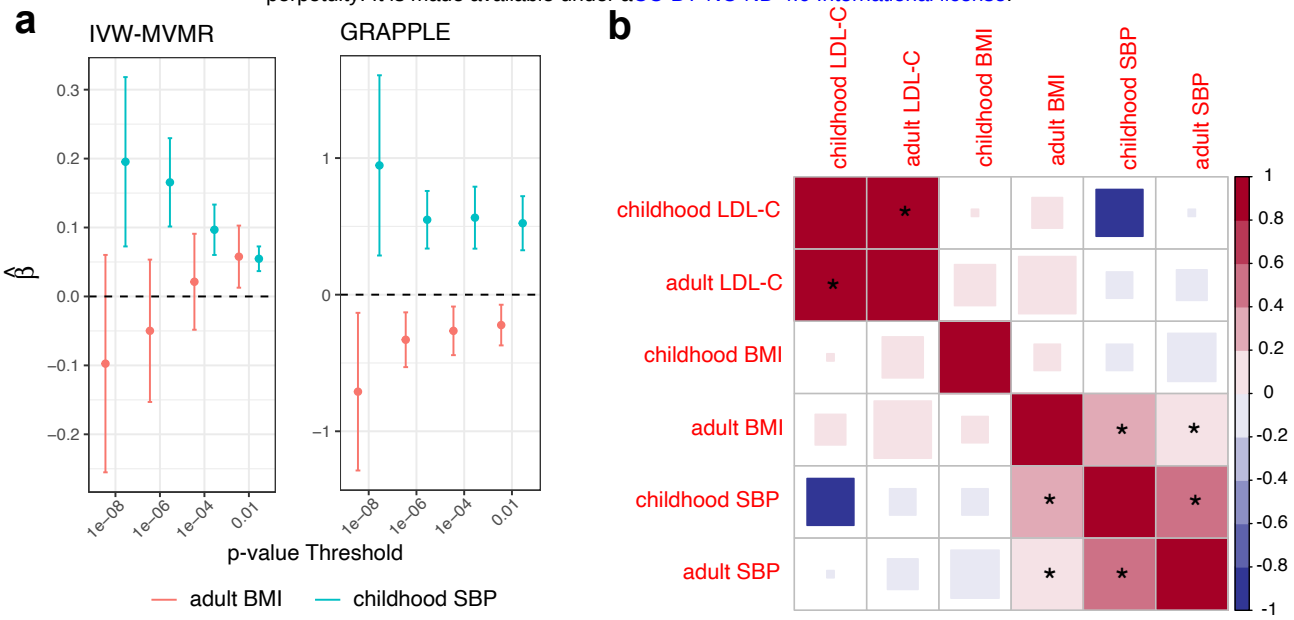

Figure S5: Additional results for the stroke case study. a) MVMR results estimating the joint causal effects of adult BMI and childhood SBP on adult SBP. b) Estimated pairwise genetic correlations across traits. A '*' indicates significantly correlated pairs.

S4 Supplemental tables

Selection p-value thresholds	10^{-8}	10^{-6}	10^{-4}	10^{-2}
Breast cancer (K=3, body size)	56	119	305	811
Breast cancer (K=3, BMI)	56	118	304	810
Breast cancer (K=4)	55	117	303	739
Stroke (K=7)	116	204	464	847
Simulation (K=3)	250	274	312	362
Simulation (K=4)	156	175	207	248
Simulation (K=7)	356	420	489	NA

Table S1: Number of SNPs used in simulations and real data analysis at any given selection p-value threshold

References

[1] Wang, J., Zhao, Q., Bowden, J., Hemani, G., Davey Smith, G., Small, D.S., and Zhang, N.R. (2021). Causal inference for heritable phenotypic risk factors using heterogeneous genetic instruments. PLoS genetics *17*, e1009575.

[2] Sanz-Alonso, D., Stuart, A.M., and Taeb, A. (2018). Inverse problems and data assimilation. arXiv e-prints , arXiv–1810.

[3] Felix, J.F., Bradfield, J.P., Monnereau, C., Van Der Valk, R.J., Stergiakouli, E., Chesi, A., Gaillard, R., Feenstra, B., Thiering, E., Kreiner-Møller, E., et al. (2016). Genome-wide association analysis identifies three new susceptibility loci for childhood body mass index. Human molecular genetics *25*, 389–403.

[4] Hoffmann, T.J., Theusch, E., Haldar, T., Ranatunga, D.K., Jorgenson, E., Medina, M.W., Kvale, M.N., Kwok, P.Y., Schaefer, C., Krauss, R.M., et al. (2018). A large electronic-health-record-based genome-wide study of serum lipids. Nature genetics *50*, 401–413.

[5] O'Nunain, K., Sanderson, E., Holmes, M.V., Smith, G.D., and Richardson, T.G. (2023). A genome-wide association study of childhood adiposity and blood lipids. Wellcome Open Research *6*, 303.

[6] Willer, C., Schmidt, E., Sengupta, S., Peloso, G., Gustafsson, S., Kanoni, S., Ganna, A., Chen, J., Buchkovich, M., Mora, S., et al. (2013). Discovery and refinement of loci associated with lipid levels. Nature genetics *45*, 1274–1283.

[7] Malik, R., Chauhan, G., Traylor, M., Sargurupremraj, M., Okada, Y., Mishra, A., Rutten-Jacobs, L., Giese, A.K., Van Der Laan, S.W., Gretarsdottir, S., et al. (2018). Multiancestry genome-wide association study of 520,000 subjects identifies 32 loci associated with stroke and stroke subtypes. Nature genetics *50*, 524–537.

[8] Bulik-Sullivan, B.K., Loh, P.R., Finucane, H.K., Ripke, S., Yang, J., of the Psychiatric Genomics Consortium, S.W.G., Patterson, N., Daly, M.J., Price, A.L., and Neale, B.M. (2015). Ld score regression distinguishes confounding from polygenicity in genome-wide association studies. Nature genetics *47*, 291–295.

[9] Bulik-Sullivan, B., Finucane, H.K., Anttila, V., Gusev, A., Day, F.R., Loh, P.R., Consortium, R., Consortium, P.G., for Anorexia Nervosa of the Wellcome Trust Case Control Consortium 3, G.C., Duncan, L., et al. (2015). An atlas of genetic correlations across human diseases and traits. Nature genetics *47*, 1236–1241.