Identifying Gene-environment interactions with robust marginal Bayesian variable selection

Xi Lu[1], Kun Fan[1], Jie Ren[2] and Cen Wu[1]

[1] Department of Statistics, Kansas State University, Manhattan, KS

[2] Department of Biostatistics, Indiana University School of Medicine, Indianapolis, IN

Abstract

In high-throughput cancer studies, an important aim is to identify gene-environment interactions associated with the clinical outcomes. Recently, multiple marginal penalization methods have been developed and shown to be effective in G×E studies. However, within the Bayesian framework, marginal variable selection has not received much attention. In this study, we propose a novel marginal Bayesian variable selection method for G×E studies. In particular, our marginal Bayesian method is robust to data contamination and outliers in the outcome variables. With the incorporation of spike-and-slab priors, we have implemented the Gibbs sampler based on MCMC. The proposed method outperforms a number of alternatives in extensive simulation studies. The utility of the marginal robust Bayesian variable selection method has been further demonstrated in the case studies using data from the Nurse Health Study (NHS). Some of the identified main and interaction effects from the real data analysis have important biological implications.

Keywords: Gene-environment interaction; marginal analysis; MCMC; robust Bayesian variable selection; spike-and-slab priors.

1 Introduction

The risk and progression of complex diseases including cancer, asthma and type 2 diabetes, are associated with the coordinated functioning of genetic factors, the environmental (and clinical) factors, as well as their interactions([1, 2, 3, 4]). The identification of important Gene-environment(G×E) interactions leads to novel insight in dissecting the genetic basis of complex diseases in addition to the main effects of genetic and environmental factors. In the last two decades, searching for the important G×E interactions has been extensively conducted based on genetic association studies([5, 6]). One representative example is the genome wide association study (GWAS), where the statistical significance of interaction between the environmental exposure and the genetic variant has been marginally assessed one at a time across the whole genome. Important findings are evidenced by genome wide significant p-values after adjusting for multiple comparisons.

Recently, substantial efforts have been devoted to novel penalized variable selection methods for G×E studies([7]). In particular, marginal penalization has achieved very competitive performances with the aforementioned significance based G×E analysis([8, 9, 10]). For exam-

ple, within the framework of maximum rank correlation, Shi et al.(2014)([8]) has developed a penalization method robust to outliers and model misspecification in determining important G×E interactions one at a time. Zhang et al. (2019)([10]) has imposed hierarchical structure between the main effects and interactions in marginal identification of G×E interactions using regularization. Despite success, these studies have limitations. First, as a common tuning parameter is demanded for all the marginal models, its selection requires pooling all genes together to conduct a joint model based cross validation. While such a strategy is not rare, it seems not in favor of the marginal nature of the proposed G×E studies. Second, a rigorous measure to quantify uncertainty is not available. Zhang et al. (2019)([10]) has constructed 95% confidence intervals based on the observed occurrence index (OOI) values([11]), nevertheless, this measure has been used to demonstrate stability of identified effects rather than quantifying uncertainty of penalized estimates.

These limitations have motivated us to consider Bayesian analyses. In literature, Bayesian variable selection methods have been developed for G×E analysis in multiple studies([7]). For example, with indicator model selection, Liu et al. (2015) ([12]) has imposed hierarchical Bayesian variable selection for linear G×E interactions. Li et al. (2015)([13]) has proposed a Bayesian group LASSO to identify non-linear interactions in nonparametric varying coefficient models. Ren et al. (2020)([14]) has further incorporated selection of linear and nonlinear G×E interactions simultaneously while accounting for structured identification in the Bayesian adaptive shrinkage framework. All these fully Bayesian methods can efficiently provide uncertainty quantification based on the posterior samples from MCMC. Nevertheless, our limited literature mining shows that none of the marginal Bayesian variable selection methods have been proposed for interaction studies so far.

Historically, marginal analysis has prevailed in G×E interaction studies within the framework of genetic association studies. Although recent studies have confirmed the utility of regularized variable selection in joint G×E analysis, more efforts are needed for marginal penalizations especially through the Bayesian point of view. The step towards marginal Bayesian variable selection is of particular significance in developing a coherent framework of analyzing G×E interactions.

Here, we propose a novel marginal Bayesian variable selection method for the robust identification of G×E interactions. As heavy-tailed distributions and outliers in the response variable have been widely observed, robust modelling is essential for yielding reliable results. Specifically, the robustness of the proposed method is facilitated by the Bayesian formulation of the least absolute deviation (LAD) regression which has been a popular choice in frequentist G×E studies but seldom investigated in a similar context from the Bayesian perspective. We consider the Bayesian LAD LASSO for regularized identification of interaction effects. As Bayesian LAD LASSO does not lead to zero coefficients, the spike-and-slab priors([15, 16]) has been incorporated to impose exact sparsity in the adaptive shrinkage framework. The corresponding MCMC algorithm has been developed to accommodate fast computations. We have demonstrated the advantage of the proposed robust Bayesian marginal analysis in simulation. The findings from the case study of the Nurses' Health Study (NHS) with SNP measurements have important biological implications.

2 Method

We use Y to denote a continuous response variable representing the the cancer outcome or disease phenotype. Let $X = (X_1, \ldots, X_p)$ be the p genetic variants, $E = (E_1, \ldots, E_q)$ be the q environmental factors and $C = (C_1, \ldots, C_m)$ be the m clinical factors. We denote the ith subject with i. Let (Y_i, E_i, C_i, X_i) $(i = 1, \ldots, n)$ be independent and identically distributed random vectors. For X_{ij} $(j = 1, \ldots, p)$, the measurement of the jth genetic factor on the ith subject, consider the following marginal model:

$$\begin{aligned} Y_i &= \sum_{k=1}^{q} \alpha_k E_{ik} + \sum_{t=1}^{m} \gamma_t C_{it} + \beta_j X_{ij} + \sum_{k=1}^{q} \eta_{jk} X_{ij} E_{ik} + \epsilon_i \\ &= \sum_{k=1}^{q} \alpha_k E_{ik} + \sum_{t=1}^{m} \gamma_t C_{it} + \beta_j X_{ij} + \eta_j \tilde{W}_i + \epsilon_i, \end{aligned} \tag{1}$$

where α_k's and γ_t's are the regression coefficients corresponding to effects of environmental and clinical factors, respectively. For the jth gene X_j $(j = 1, \ldots, p)$, the G×E interactions effects are defined with $W_j = (X_j E_1, \ldots, X_j E_q)$, $\eta_j = (\eta_{j1}, \ldots, \eta_{jq})^T$. With a slight abuse of notation, denote $\tilde{W} = W_j$. The β_j's and η_{jk}'s are the regression coefficients of the genetic variants and G×E interactions effects, correspondingly. Denote $\alpha = (\alpha_1, \ldots, \alpha_q)^T$ and $\gamma = (\gamma_1, \ldots, \gamma_m)^T$. Then model (1) can be written as

$$Y_i = E_i \alpha + C_i \gamma + X_{ij} \beta_j + \tilde{W}_i \eta_j + \epsilon_i. \tag{2}$$

2.1 Bayesian formulation of the LAD regerssion

The necessity of accounting for robustness in interaction studies has been increasingly recognized ([7]). Within the frequentist framework, it is essentially dependent on adopting a robust loss function to quantify lack of fit ([17]). Among a variety of popular robust losses, the least absolute deviation (LAD) loss function is well known for its advantages in dealing with heavy-tailed error distributions or outliers in response. The estimation of regression coefficients amounts to the following minimization problem

$$\min_{\alpha, \gamma, \beta_j, \eta_j} \sum_{i=1}^{n} |Y_i - E_i \alpha - C_i \gamma - X_{ij} \beta_j - \tilde{W}_i \eta_j|.$$

Here, we propose the robust marginal Bayesian variable selection based on LAD. As the Laplace distribution is equivalent to the mixture of an exponential distribution and a scaled normal distribution([18]), for a Bayesian formulation of LAD regression, we assume that $\epsilon_i (i = 1, \ldots, n)$ are i.i.d. random variables following the Laplace distribution with density

$$f(\epsilon_i | \tau) = \frac{\tau}{2} \exp(-\tau |\epsilon_i|),$$

where τ is the inverse of the scale parameters from the Laplace density. Then the likelihood function of our marginal G×E model can be expressed as:

$$f(Y | \alpha, \gamma, \beta_j, \eta_j) = \prod_{i=1}^{n} \frac{\tau}{2} \exp(-\tau |Y_i - E_i \alpha - C_i \gamma - X_{ij} \beta_j - \tilde{W}_i \eta_j|).$$

The above formulation using Laplace distrubution is a special case of the asymmetric Laplace distribution, which has been widely adopted in Baysian quantile regression([19, 20]). In Baysian quantile regression, ϵ_i's are assumed to follow the skewed Laplace distribution with density

$$f(\epsilon|\tau) = \theta(1-\theta)\tau\exp(-\tau\rho_\theta(\epsilon)).$$

The random errors can be written as

$$\epsilon_i = \xi_1 v_i + \tau^{-1/2}\xi_2\sqrt{v_i}z_i,$$

where

$$\xi_1 = \frac{1-2\theta}{\theta(1-\theta)} \quad and \quad \xi_2 = \sqrt{\frac{2}{\theta(1-\theta)}}$$

with quantile level $\theta \in (0,1)$, $v_i \sim \exp(\tau^{-1})$, and $z_i \sim \mathrm{N}(0,1)$.

The Bayesian LAD regression is a special case of Bayesian quantile regression([21]) with θ=0.5, resulting in that $\xi_1 = 0$ and $\xi_2 = \sqrt{8}$. Therefore, the response Y_i can be written as:

$$\begin{aligned} Y_i &= \mu_i + \tau^{-1/2}\xi_2\sqrt{v_i}z_i, \\ v_i|\tau &\overset{iid}{\sim} \tau\exp(-\tau v_i), \\ z_i &\overset{iid}{\sim} \mathrm{N}(0,1), \end{aligned} \tag{3}$$

where $\mu_i = E_i\alpha + C_i\gamma + X_{ij}\beta_j + \tilde{W}_i\eta_j$.

2.2 Bayesian LAD LASSO with spike-and-slab priors

In model (1), the coefficients β_j and η_j corresponds to the main and interaction effects with respect to the jth genentic variant, respectively. When $\beta_j = 0$ and $\eta_j = 0$, the genetic variant has no effect on the phenotype. A non-zero β_j suggests the presence of main genetic effect. For η_j, if at least one of its component is not zero, then the G×E interaction effect exist. In the literature, Bayesian quantile LASSO, with Bayesian LAD LASSO as its special case, has been proposed to conduct variable selection([21]). However, a major limitation is that Bayesian quantile LASSO cannot shrink regression coefficients to 0 exactly, resulting in inaccurate identification and biased estimation. To overcome such an limitation, we incorporate spike-and-slab priors to impose sparsity within Bayesian LAD LASSO framework as follows.

For the jth gene ($j = 1, \ldots, p$), the marginal LAD LASSO model is given by:

$$\sum_{i=1}^{n}|Y_i - E_i\alpha - C_i\gamma - X_{ij}\beta_j - \tilde{W}_i\eta_j| + \lambda_1|\beta_j| + \lambda_2\sum_{k=1}^{q}|\eta_{jk}|.$$

Let $\varphi_1 = \tau\lambda_1$ and $\varphi_2 = \tau\lambda_2$. Then the conditional Laplace prior on the coefficient of main effect β_j can be expressed as scale mixtures of normals:

$$\begin{aligned} \pi(\beta_j|\tau,\lambda_1) &= \frac{\varphi_1}{2}\exp\{-\varphi_1|\beta_j|\} \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi s_1}}\exp(-\frac{\beta_j^2}{2s_1})\frac{\varphi_1^2}{2}\exp(\frac{-\varphi_1^2}{2}s_1)ds_1. \end{aligned}$$

The conditional Laplace prior on the coefficients of interaction effect η_j can be written as:

$$\begin{aligned}\pi(\eta_j|\tau,\lambda_2) &= \prod_{k=1}^{q}\frac{\varphi_2}{2}\exp\{-\varphi_2|\eta_{jk}|\}\\ &= \prod_{k=1}^{q}\int_0^\infty \frac{1}{\sqrt{2\pi s_2}}\exp(-\frac{\eta_{jk}^2}{2s_2})\frac{\varphi_2^2}{2}\exp(\frac{-\varphi_2^2}{2}s_2)ds_2.\end{aligned}$$

Therefore, we consider the following hierarchical formulation for the marginal G×E model:

$$\begin{aligned}\beta_j|s_1,\pi_1 &\sim (1-\pi_1)\mathrm{N}(0,s_1)+\pi_1\delta_0(\beta_j),\\ s_1|\varphi_1^2 &\sim \frac{\varphi_1^2}{2}\exp(-\frac{\varphi_1^2}{2}s_1),\\ \eta_{jk}|s_{2k},\pi_2 &\overset{iid}{\sim} (1-\pi_2)\mathrm{N}(0,s_{2k})+\pi_2\delta_0(\eta_{jk})(k=1,\ldots,q),\\ s_{2k}|\varphi_2^2 &\overset{iid}{\sim} \frac{\varphi_2^2}{2}\exp(-\frac{\varphi_2^2}{2}s_{2k})(k=1,\ldots,q),\end{aligned} \tag{4}$$

where $\delta_0(\beta_j)$ and $\delta_0(\eta_{jk})$ denote the spike at 0, respectively, and the slab distributions are represented by two normal distributions, $\mathrm{N}(0,s_1)$ and $\mathrm{N}(0,s_2k)$. Here, $\pi_1\in[0,1]$ and $\pi_2\in[0,1]$. The mixture of the spike and slab components facilitate the selection of main and interaction effects. Instead of setting π_1 and π_2 to a fixed value such as 0.5, we assign conjugate beta priors on them as $\pi_1\sim$Beta(r_1,u_1) and $\pi_2\sim$Beta(r_2,u_2) which account for the uncertainty in π_1 and π_2. In this paper, we choose $r_1=u_1=r_2=u_2=1$ as it gives a prior mean with 0.5 and it also allows a prior to spread out.

In addition, the normal prior has been placed on the coefficients of environmental factor $\alpha_k(k=1,\ldots,q)$ and clinical factor $\gamma_t(t=1,\ldots,m)$ as:

$$\begin{aligned}\alpha_k &\overset{iid}{\sim} \frac{1}{\sqrt{(2\pi\alpha_0)}}\exp(-\frac{\alpha_k^2}{2\alpha_0})(k=1,\ldots,q)\\ \gamma_t &\overset{iid}{\sim} \frac{1}{\sqrt{(2\pi\gamma_0)}}\exp(-\frac{\gamma_t^2}{2\gamma_0})(t=1,\ldots,m),\end{aligned}$$

We also assume conjugate Gamma priors on τ, φ_1^2 and φ_2^2 with

$$\begin{aligned}\tau &\sim \mathrm{Gamma}(a,b),\\ \varphi_1^2 &\sim \mathrm{Gamma}(c_1,d_1),\\ \varphi_2^2 &\sim \mathrm{Gamma}(c_2,d_2).\end{aligned}$$

In typical G×E studies, the environmental and clinical factors are of low dimensionality and the selection of them is not of interest. Therefore, the sparsity-inducing priors have not been adopted for these factors. We consider the Bayesian LAD LASSO type of regularization in the proposed study as published studies have demonstrated that baseline penalty such as MCP and LASSO work well for marginal variable selection ([8, 9]).

It is noted that Zhang et al. (2020) ([10]) has proposed a marginal sparse group MCP to respect the strong hierarchy between main and interaction effects. Their results are promising

when long tailed distributions and outliers are not present in the response variable. Although sparse group (or, bi–level) variable selection has been demonstrated as being very effective in multiple G×E studies based on joint models ([7]), in our study, there is only one group per each marginal model. The sparse group no longer has significant advantages over individual level selection. Therefore, it has not been considered here.

Our model respects the weak hierarchy of "main effects, interactions". If imposing the strong hierarchy is needed, the genetic factor, once it is not selected given the presence of corresponding interaction effects, can be added back to the identified marginal model for a refit to impose strong hierarchy ([9]). While such a practice is not uncommon in marginal interaction studies , Shi et al. (2014) ([8]) has also revealed satisfactory performance when strong hierarchy has not been pursued.

2.3 The Gibbs sampler for robust marginal G×E analysis

For the jth genetic factor, the joint posterior distribution of all the unknown parameters conditional on data can be expressed as

$$
\begin{aligned}
&\pi(\alpha,\gamma,\beta_j,\eta_j,v,s_1,s_2,\tau,\varphi_1,\varphi_2,\pi_1,\pi_2,z_i|Y)\\
&\qquad\propto\prod_{i=1}^{n}\frac{1}{\sqrt{2\pi\tau^{-1}\xi_2^2 v_i}}\exp\Big\{-\frac{(y_i-E_i\alpha-C_i\gamma-X_{ij}\beta_j-\tilde{W}_i\eta_j)^2}{2\tau^{-1}\xi_2^2 v_i}\Big\}\\
&\qquad\times\prod_{i=1}^{n}\tau\exp(-\tau v_i)\tau^{a-1}\exp(-b\tau)\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}z_i^2)\\
&\qquad\times\prod_{k=1}^{q}\frac{1}{\sqrt{(2\pi\alpha_0)}}\exp(-\frac{\alpha_k^2}{2\alpha_0})\\
&\qquad\times\prod_{t=1}^{m}\frac{1}{\sqrt{(2\pi\gamma_0)}}\exp(-\frac{\gamma_t^2}{2\gamma_0})\\
&\qquad\times\Big((1-\pi_1)(2\pi s_1)^{-1/2}\exp(-\frac{\beta_j^2}{2s_1})\mathbf{I}_{\{\beta_j\neq 0\}}+\pi_1\delta_0(\beta_j)\Big)\\
&\qquad\times\prod_{k=1}^{q}\Big((1-\pi_2)(2\pi s_{2k})^{-1/2}\exp(-\frac{\eta_{jk}^2}{2s_{2k}})\mathbf{I}_{\{\eta_{jk}\neq 0\}}+\pi_2\delta_0(\eta_{jk})\Big)\\
&\qquad\times\frac{\varphi_1^2}{2}\exp(-\frac{\varphi_1^2}{2}s_1)\\
&\qquad\times\prod_{k=1}^{q}\frac{\varphi_2^2}{2}\exp(-\frac{\varphi_2^2}{2}s_{2k})\\
&\qquad\times(\varphi_1^2)^{c_1-1}\exp(-d_1\varphi_1^2)\\
&\qquad\times(\varphi_2^2)^{c_2-1}\exp(-d_2\varphi_2^2)\\
&\qquad\times\pi_1^{r_1-1}(1-\pi_1)^{u_1-1}\\
&\qquad\times\pi_2^{r_2-1}(1-\pi_2)^{u_2-1}
\end{aligned}
$$

Let $\mu_{(-\alpha_k)}=E(y_i)-E_{ik}\alpha_k,(i=1,\ldots,n),(k=1,\ldots,q)$, representing the mean effect

without the contribution of $E_{ik}\alpha_k$. The posterior distribution of the coefficient of environmental factor α_k conditional on all other parmeters can be expressed as

$$\begin{aligned}
&\pi(\alpha_k|\text{rest}) \\
&\quad\propto \pi(\alpha_k)\pi(Y|\cdot) \\
&\quad\propto \exp\Big\{-\sum_{i=1}^{n}\frac{(y_i-E_i\alpha-C_i\gamma-X_{ij}\beta_j-\tilde{W}_i\eta_j)^2}{2\tau^{-1}\xi_2^2 v_i}\Big\}\times\exp(-\frac{\alpha_k^2}{2\alpha_0}) \\
&\quad\propto \exp\Big\{-\frac{1}{2}\big[(\sum_{i=1}^{n}\frac{\tau E_{ik}^2}{\xi_2^2 v_i}+\frac{1}{\alpha_0})\alpha_k^2-2\sum_{i=1}^{n}\frac{\tau(y_i-\mu_{(-\alpha_k)})E_{ik}}{\xi_2^2 v_i}\alpha_k\big]\Big\}.
\end{aligned}$$

Hence, the full conditional distribution of α_k is normal distribution $N(\mu_{\alpha_k},\sigma^2_{\alpha_k})$with mean

$$\mu_{\alpha_k}=\big(\sum_{i=1}^{n}\frac{\tau(y_i-\mu_{(-\alpha_k)})E_{ik}}{\xi_2^2 v_i}\big)\sigma^2_{\alpha_k},$$

and variance

$$\sigma^2_{\alpha_k}=\big(\sum_{i=1}^{n}\frac{\tau E_{ik}^2}{\xi_2^2 v_i}+\frac{1}{\alpha_0}\big)^{-1}.$$

The posterior distribution of the coefficient of clinical factor $\gamma_t(t=1,\ldots,m)$ conditional on all other parameters can be obtained in similiar way. Let $\mu_{(-\gamma_t)}=E(y_i)-C_{it}\gamma_t$, $i=1,\ldots,n$, then

$$\gamma_t|\text{rest}\sim N(\mu_{\gamma_k},\sigma^2_{\gamma_t}),$$

where

$$\begin{aligned}
\mu_{\gamma_t}&=\big(\sum_{i=1}^{n}\frac{\tau(y_i-\mu_{(-\gamma_t)})C_{it}}{\xi_2^2 v_i}\big)\sigma^2_{\gamma_t},\\
\sigma^2_{\gamma_t}&=\big(\sum_{i=1}^{n}\frac{\tau C_{it}^2}{\xi_2^2 v_i}+\frac{1}{\gamma_0}\big)^{-1}.
\end{aligned}$$

Let $\mu_{(-\beta_j)}=E(y_i)-X_{ij}\beta_j$ and $l_1=\pi(\beta_j=0|\text{rest})$, the conditional posterior distribution of the coefficient of genetic factor β_j is a spike-and-slab distribution:

$$\beta_j|\text{rest}\sim(1-l_1)N(\mu_{\beta_j},\sigma^2_{\beta_j})+l_1\delta_0(\beta_j), \tag{5}$$

where

$$\begin{aligned}
\mu_{\beta_j}&=\big(\sum_{i=1}^{n}\frac{\tau(y_i-\mu_{(-\beta_j)})X_{ij}}{\xi_2^2 v_i}\big)\sigma^2_{\beta_j},\\
\sigma^2_{\beta_j}&=\big(\sum_{i=1}^{n}\frac{\tau X_{ij}^2}{\xi_2^2 v_i}+\frac{1}{s_1}\big)^{-1}.
\end{aligned}$$

We can show that

$$l_1=\frac{\pi_1}{\pi_1+(1-\pi_1)s_1^{-1/2}(\sigma^2_{\beta_j})^{1/2}\exp\{\frac{1}{2}(\sum_{i=1}^{n}\frac{\tau(y_i-\mu_{(-\beta_j)})X_{ij}}{\xi_2^2 v_i})^2\sigma^2_{\beta_j}\}}.$$

The posterior distribution of β_j is a mixture of a normal distribution and a point mass at 0. That is, at each iteractio of MCMC, β_j is drawn from $N(\mu_{\beta_j}, \sigma^2_{\beta_j})$ with probability $(1 - l_1)$ and is set to 0 with probability l_1.

Similiarly, the posterior distribution of the interaction of the jth gene and environmental factors $\eta_{jk}(k = 1, \ldots, q)$ is also a spike-and-slab distribution. Denote $\mu_{(-\eta_{jk})} = E(y_i) - W_{ik}\eta_{jk}$ and $l_{2k} = \pi(\eta_{jk} = 0|\text{rest})$, η_{jk} follows this distribution:

$$\eta_{jk}|\text{rest} \sim (1 - l_{2k})N(\mu_{\eta_{jk}}, \sigma^2_{\eta_{jk}}) + l_{2k}\delta_0(\eta_{jk}), \tag{6}$$

where

$$\mu_{\eta_{jk}} = \Big(\sum_{i=1}^{n} \frac{\tau(y_i - \mu_{(-\eta_{jk})})\tilde{W}_{ik}}{\xi_2^2 v_i}\Big)\sigma^2_{\eta_{jk}},$$
$$\sigma^2_{\beta_j} = \Big(\sum_{i=1}^{n} \frac{\tau\tilde{W}^2_{ik}}{\xi_2^2 v_i} + \frac{1}{s_{2k}}\Big)^{-1}.$$

And

$$l_{2k} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)s_{2k}^{-1/2}(\sigma^2_{\eta_{jk}})^{1/2}\exp\{\frac{1}{2}(\sum_{i=1}^{n} \frac{\tau(y_i - \mu_{(-\eta_{jk})})\tilde{W}_{ik}}{\xi_2^2 v_i})^2\sigma^2_{\eta_{jk}}\}}. \tag{7}$$

The full conditional posterior distribution of s_1 is:

$$\begin{aligned} &s_1|\text{rest} \\ &\quad \propto \pi(\beta_j|s_1, \pi_1)\pi(s_1|\varphi_1^2) \\ &\quad \propto \Big((1 - \pi_1)(2\pi s_1)^{-1/2}\exp(-\frac{\beta_j^2}{2s_1})\mathbf{I}_{\{\beta_j \neq 0\}} + \pi_1\delta_0(\beta_j)\Big)\exp(-\frac{\varphi_1^2}{2}s_1). \end{aligned} \tag{8}$$

When $\beta_j = 0$, equation(8) is proportional to $\exp(-\frac{\varphi_1^2}{2}s_1)$. Therefore, the posterior distribution of s_1 is $\exp(\frac{\varphi_1^2}{2})$.

When $\beta_j \neq 0$, equation(8) is proportional to

$$\begin{aligned} &\frac{1}{\sqrt{s_1}}\exp(-\frac{\varphi_1^2}{2}s_1)\exp(-\frac{\beta_j^2}{2s_1}) \\ &\qquad \propto \frac{1}{\sqrt{s_1}}\exp\Big\{-\frac{1}{2}[\varphi_1^2 s_1 + \frac{\beta_j^2}{s_1}]\Big\}. \end{aligned}$$

Therefore, when $\beta_j \neq 0$, the posterior distribution for s_1^{-1} is Inverse-Gaussian$(\sqrt{\frac{\varphi_1^2}{\beta_j^2}}, \varphi_1^2)$.

Similiarly, for $s_{2k}(k = 1, \ldots, q)$, when $\eta_{jk} = 0$, the posterior distribution of s_{2k} is $\exp(\frac{\varphi_2^2}{2})$. When $\eta_{jk} \neq 0$, the posterior distribution for s_{2k}^{-1} is Inverse-Gaussian$(\sqrt{\frac{\varphi_2^2}{\eta_{jk}^2}}, \varphi_2^2)$.

The full conditional posterior distribution of φ_1^2:

$$\begin{aligned} &\varphi_1^2|\text{rest} \\ &\quad \propto \pi(s_1|\varphi_1^2)\pi(\varphi_1^2) \\ &\quad \propto \frac{\varphi_1^2}{2}\exp(-\frac{\varphi_1^2 s_1}{2})(\varphi_1^2)^{c_1 - 1}\exp(-d_1\varphi_1^2) \\ &\quad \propto (\varphi_1^2)^{c_1}\exp\Big(-\varphi_1^2(s_1/2 + d_1)\Big). \end{aligned}$$

Therefore, the posterior distribution for φ_1^2 is Gamma$(c_1+1, s_1/2+d_1)$. Similiarly, the posterior distribution for φ_2^2 is Gamma$(c_2+q, \sum_{k=1}^{q} s_{2k}/2+d_2)$.

The full conditional posterior distribution of π_1:

$$\begin{aligned}
&\pi_1|\text{rest} \\
&\quad\propto \pi(s_1|\varphi_1^2)\pi(\varphi_1^2) \\
&\quad\propto \pi_1^{r_1-1}(1-\pi_1)^{u_1-1} \\
&\quad\times \Big((1-\pi_1)(2\pi s_1)^{-1/2}\exp(-\frac{\beta_j^2}{2s_1})\mathbf{I}_{\{\beta_j\neq 0\}}+\pi_1\delta_0(\beta_j)\Big).
\end{aligned}$$

Then, the posterior distribution for π_1 is Beta $(1+r_1-\mathbf{I}(\beta_j\neq 0), u_1+\mathbf{I}(\beta_j\neq 0))$.

The full conditional posterior distribution of π_2:

$$\begin{aligned}
&\pi_2|\text{rest} \\
&\quad\propto \pi(s_2|\varphi_2^2)\pi(\varphi_2^2) \\
&\quad\propto \pi_2^{r_2-1}(1-\pi_2)^{u_2-1} \\
&\quad\times \prod_{k=1}^{q}\Big((1-\pi_2)(2\pi s_{2k})^{-1/2}\exp(-\frac{\eta_{jk}^2}{2s_{2k}})\mathbf{I}_{\{\eta_{jk}\neq 0\}}+\pi_2\delta_0(\eta_{jk})\Big).
\end{aligned}$$

So, the posterior distribution for π_2 is Beta $(1+r_1-\sum_{k=1}^{q}\mathbf{I}(\eta_{jk}\neq 0), u_1+\sum_{k=1}^{q}\mathbf{I}(\eta_{jk}\neq 0))$.

The full conditional posterior distribution of τ:

$$\begin{aligned}
&\tau|\text{rest} \\
&\quad\propto \pi(v|\tau)\pi(\tau)\pi(Y|\cdot) \\
&\quad\propto \tau^{n/2}\exp\Big\{-\sum_{i=1}^{n}\frac{(y_i-E_i\alpha-C_i\gamma-X_{ij}\beta_j-\tilde{W}_i\eta_j)^2}{2\tau^{-1}\xi_2^2 v_i}\Big\} \\
&\quad\times \tau^n\exp(-\tau\sum_{i=1}^{n}v_i)\tau^{a-1}\exp(-b\tau) \\
&\quad\propto \tau^{a+\frac{3}{2}n-1}\exp\Big\{-\tau\big[\sum_{i=1}^{n}(\frac{(y_i-E_i\alpha-C_i\gamma-X_{ij}\beta_j-\tilde{W}_i\eta_j)^2}{2\xi_2^2 v_i}+v_i)+b\big]\Big\}.
\end{aligned}$$

Therefore, the posterior distribution for τ is Gamma$(a+\frac{3}{2}n, \big[\sum_{i=1}^{n}(\frac{(y_i-E_i\alpha-C_i\gamma-X_{ij}\beta_j-\tilde{W}_i\eta_j)^2}{2\xi_2^2 v_i}+v_i)+b\big])$.

Last, we have the full conditional posterior distribution of v_i:

$$\begin{aligned}
&v_i|\text{rest} \\
&\quad\propto \pi(v|\tau)\pi(Y|\cdot) \\
&\quad\propto \frac{1}{\sqrt{v_i}}\exp\Big\{-\frac{(y_i-E_i\alpha-C_i\gamma-X_{ij}\beta_j-\tilde{W}_i\eta_j)^2}{2\tau^{-1}\xi_2^2 v_i}\Big\}\times\exp(-\tau v_i) \\
&\quad\propto \frac{1}{\sqrt{v_i}}\exp\Big\{-\frac{1}{2}\big[(2\tau)v_i+\frac{\tau(y_i-E_i\alpha-C_i\gamma-X_{ij}\beta_j-\tilde{W}_i\eta_j)^2}{\xi_2^2 v_i}\big]\Big\}.
\end{aligned}$$

It is easy to show that

$$\frac{1}{v_i}|\text{rest} \sim \text{Inverse-Gaussian}(\sqrt{\frac{2\xi_2^2}{(y_i - E_i\alpha - C_i\gamma - X_{ij}\beta_j - \tilde{W}_i\eta_j)^2}},\ 2\tau).$$

The spirit of marginal penalization for G×E interactions lies in the usage of a common sparsity cutoff to determine a list of important main and interaction effects. Instead of focusing on a fixed cutoff, varying the cutoff can generate different lists, resulting in a comprehensive view of important findings. The tuning parameter in penalized estimation serves as the cutoff. Therefore, the same tuning parameter has to be adopted for all the sub models ([8, 9, 10]). To further justify such a common tuning parameter, Zhang et al. (2020) ([10]) has attempted using the joint model to select the common tuning through cross validation. However, this seems not coherent with the nature of marginal analysis.

Ideally, the tuning parameter should be determined by each model itself to allow for flexibility in controlling sparsity individually, and a common cutoff is still available to examine different lists of important effects. With the Bayesian formulation, we can avoid such a limitation of frequentist marginal penalization methods. In particular, the priors have been placed on regularization parameters to determine the sparsity in a data-driven manner for each sub model. With the spike-and-slab priors, the posterior distributions on the coefficients of main and interaction effects naturally lead to the usage of inclusion probability as a common cutoff to pin down the list of important effects, which is described in detail in the next section.

3 Simulation

To demonstrate the utility of the proposed approach, we evaluate the performance through simulation study. In particular, we compare the performance of the proposed method, LAD Bayesian Lasso with spike-and-slab priors (denoted as LADBLSS) with three alternatives, LAD Bayesian Lasso (denoted as LADBL), Bayesian Lasso with spike-and-slab priors (denoted as BLSS) and Bayesian Lasso (denoted as BL). LADBL is similar to the proposed method, except that it does not adopt the spike-and-slab prior. The details of posterior inference are available from the Appendix.

Under all settings, the sample size is set as $n = 200$, and the number of G factors is $p = 500$ with $q = 4$, $m = 3$. For environmental factors, we simulate four continuous variables from multivariate normal distributions with marginal mean 0, marginal variance 1 and AR1 correlation structure with $\rho = 0.5$. In addition, three clinical factors are generated from a multivariate normal distribution with margianl mean 0 and marginal variance 1 and AR1 structure with $\rho = 0.5$. Among the p main G effects and pq G×E interactions, 8 and 12 effects are set as being associated with the response, respectively. All the environmental and clinical factors are important with nonzero coefficients, which are randomly generated from a uniform distribution Unif$[0.1, 0.5]$. The random error are generated from: (1) N$(0, 1)$(Error 1), (2) t-distribution with 2 degress of freedom ($t(2)$) (Error2), (3) LogNormal(0,2)(Error3), (4) 90%N(0,1)+10%Cauchy(0,1)(Error4), (5) 80%N(0,1)+20%Cauchy(0,1)(Error5). All of them are heavy-tailed distribution except the first one.

In addition, the genetic factors are simulated in the following four settings.

Setting 1. In simulating continuous genetic variants, we generate multivariate normal distributions with marginal mean 0 and variance 1. The AR structure is considered in computing the correlation of G factors, under which gene j and k have correlation $\rho^{|j-k|}$ with $\rho = 0.5$.

Setting 2. We assess the performance under single-nucleotide polymorphism (SNP) data. The SNPs are obtained by dichotomizing the gene expression values at the 1st and 3rd quartiles, with the 3–level (0,1,2) for genotypes (aa,Aa,AA) respectively. Here, the gene expressions are generated from the first setting.

Setting 3. Consider simulating the SNP data under a pairwise linkage disequilibrium (LD) structure. For the two minor alleles A and B of two adjacent SNPs, let q_1 and q_2 be the minor allele frequencies (MAFs), respectively. The frequencies of four haplotypes are as $p_{AB} = q_1q_2+\delta$, $p_{ab} = (1-q_1)(1-q_2)+\delta$, $p_{Ab} = q_1(1-q_2)-\delta$, and $p_{aB} = (1-q_1)q_2-\delta$, where δ denotes the LD. Assuming Hardy-Weinberg equilibrium and given the allele frequency for A at locus 1, we can generate the SNP genotype (AA, Aa, aa) from a multinomial distribution with frequencies $(q_1^2, 2q_1(1-q_1), (1-q_1)^2)$. Based on the conditional genotype probability matrix, we can simulate the genotypes for locus 2. With MAFs 0.3 and pairwise correlation $r = 0.6$, we have $\delta = r\sqrt{q_1(1-q_1)q_2(1-q_2)}$.

We collect the posterior samples from the Gibbs Sampler with 10,000 interations and discard the first 5,000 samples as burn-ins. The posterior medians are used to estimate the coefficients. For approaches incorporating spike-and-slab priors, we consider computing the inclusion probability to indicate the importance of predictors. Here we use a binary indicator ϕ to denote that the membership of the non-spike distribution. Take the main effect of the jth genetic factor, X_j, as an example. Suppose we have collected H posterior samples from MCMC after burn-ins. The jth G factor is included in the marginal G×E model at the hth MCMC iteration if the corresponding indicator is 1, i.e., $\phi_j^{(g)} = 1$. Subsequently, the posterior probability of retaining the jth genetic main effect in the final marginal model is defined as the average of all the indicators for the jth G factor among the H posterior samples. That is,

$$p_j = \hat{\pi}(\phi_j = 1|y) = \frac{1}{H}\sum_{h=1}^{H}\phi_j^{(h)},\ j = 1, \ldots, p.$$

A larger posterior inclusion probability p_j indicates a stronger empirical evidence that the jth genetic main effect has a non-zero coefficient, i.e., a stronger association with the phenotypic trait.

To comprehensively assess the performance of the proposed and alternative methods, we consider a sequence of probabilities as cutting-offs in inclusion probability for methods with spike-and-slab priors. Given a cutoff probability, the main or interaction is included in the final marginal model if its posterior inclusion probability is larger than the cutoff, and is excluded otherwise. Provided with a sequence of cutting-off probabilities from small to large, we can investigate the set of identified effects and calculate the true/false positive rates (T/FPR) as the ground truth is known in simulation. For the sequence of cut-offs, we are able to compute the area under curve (AUC) as a comprehensive measure. Besides, for methods without spike-and-slab priors, the confidence level of the credible intervals can be adopted as the cut-off to compute TPR and FPRs. Therefore, all the methods under

comparison can be evaluated on the same ground.

In addition, we also consider Top100, which is defined as the number of true signals when 100 important main effects (or interactions) are identified. For methods with spik-and-slab priors, 100 main effects or interactions are chosen with the highest inclusion probabilities. For methods without spike-and-slab priors, the indicators of all effects are computed for a sequence of credible levels. The top 100 main effects or interactions are chosen in terms of the highest average identification values.

Simulation results for the gene expression data in the first setting are tabulated in Tables 1and 2. We can observe that the proposed method has the best performance among all approaches, especially when the response variable has heavy-tailed distributions. First, the performance of methods with spike-and-slab priors is consistently better than methods without spike-and-slab priors. For example, in Table 1, under error 3, the AUC of LADBLSS is 0.9558(sd 0.0161), which is much larger than that of the robust method without spike-and-slab priors, i.e., 0.8432(sd 0.0115) from LADBL. Also, the AUC of robust methods is much larger than that of non-robust methods, especially in the presence of heavy-tailed errors. For instance, in the first setting under error3, the AUC of LADBLSS is 0.9558 and the AUC of LADBL is 0.8432 while that of BLSS and BL is around 0.5. Similar advantageous performance can also be observed from the identification results with Top100. In Table 2 under error 5, LADBLSS identifies 7.80(sd 0.55) out of the 8 main effects and 10.53(sd 1.36) out of the 12 interaction effects. This is higher than the results of LADBL with 7.57(sd 0.57) of main effects and 6.83(sd 1.07) of interaction effects. Second, among all the methods with spike-and-slab priors, Bayesian LAD method with spike-and-slab priors has the best performance in all identification results. Under error 3, in Table 1, the AUC of LADBLSS is 0.9558(sd 0.0161) while the AUC of BLSS is 0.5473(sd 0.0576). Under error 4 in Table 2, LADBLSS identifies 7.77(sd 0.57) main effects and 10.67(sd 1.50) interaction effects while BLSS identifies 6.2(sd 2.62) main effects and 8.3(sd 3.98) interaction effects, respectively.

Table 1: Simulation results of the first setting. AUC (mean of AUC), SD (sd of AUC) based on 100 replicates. n=200, p=500, q=4 and $m = 3$.

		BL	BLSS	LADBL	LADBLSS
Error 1	AUC	0.9182	0.9901	0.9258	0.9887
N(0,1)	SD	0.0052	0.0021	0.0076	0.0026
Error 2	AUC	0.8332	0.9420	0.9004	0.9841
$t(2)$	SD	0.0107	0.0235	0.0078	0.0031
Error 3	AUC	0.5343	0.5473	0.8432	0.9558
Lognormal(0,2)	SD	0.0144	0.0576	0.0115	0.0161
Error 4	AUC	0.8221	0.9124	0.9222	0.9895
90%N(0,1)+10%Cauchy(0,1)	SD	0.0212	0.0410	0.0071	0.0024
Error 5	AUC	0.7507	0.8431	0.9192	0.9904
80%N(0,1)+20%Cauchy(0,1)	SD	0.0217	0.0633	0.0059	0.0018

Table 2: Identification results of the first setting with Top100 method. mean(sd) based on 100 replicates. n=200, p=500, q=4 and $m = 3$.

		Main	Interaction	Total
Error 1	BL	7.60(0.49)	6.80(1.6)	14.40(1.73)
N(0,1)	BLSS	7.80(0.41)	10.80(0.92)	18.60(1.13)
	LADBL	7.67(0.55)	6.53(1.85)	14.20(1.81)
	LADBLSS	7.76(0.5)	10.53(1.36)	18.30(1.49)
Error 2	BL	6.37(1.90)	3.90(2.07)	10.27(3.19)
$t(2)$	BLSS	6.33(1.63)	8.53(2.46)	14.87(3.71)
	LADBL	7.43(0.94)	5.80(1.71)	13.23(2.01)
	LADBLSS	7.53(0.51)	9.90(1.56)	17.43(1.76)
Error 3	BL	0.90(1.21)	0.50(0.97)	1.40(1.45)
Lognormal(0,2)	BLSS	0.73(0.94)	0.47(0.68)	1.20(1.35)
	LADBL	6.27(1.55)	3.67(1.94)	9.93(2.75)
	LADBLSS	6.10(1.37)	8.93(2.02)	15.03(3.09)
Error 4	BL	5.57(2.99)	3.63(2.53)	9.20(5.05)
90%N(0,1)	BLSS	6.20(2.62)	8.30(3.98)	14.50(6.39)
+10%Cauchy(0,1)	LADBL	7.77(0.43)	7.00(1.93)	14.77(1.81)
	LADBLSS	7.77(0.57)	10.67(1.50)	18.23(1.67)
Error 5	BL	5.07(2.89)	3.00(2.49)	8.07(5.01)
80%N(0,1)	BLSS	4.60(3.25)	5.70(4.23)	10.30(7.27)
+20%Cauchy(0,1)	LADBL	7.57(0.57)	6.83(1.07)	14.40(1.83)
	LADBLSS	7.80(0.55)	10.53(1.36)	18.33(1.69)

Similar patterns can be observed in Table 4, 5 for the second setting, and Table 6, 7 for the third setting in Appendix. Overall, the advantages of conducting robust Bayesian G×E analysis using the proposed approach can be justified based on the results of comprehensive simulation studies. The convergence of the MCMC chains with the potential scale reduction factor (PSRF) ([22]) has been conducted. In this study, we use PSRF ≤ 1.1 ([23]) as the cut-off point which indicates that chains converge to a stationary distribution. The convergence of chains after burn-ins has been checked for all parameters with the value of PSRF less than 1.1. Figure 1 shows the convergence pattern of PSRF for the main and interaction coefficients of the first genetic factors in Example 1 under Error 3.

4 Real Data Analysis

In this study, we analyze the type 2 diabetes (T2D) data from Nurses' Health Study (NHS), which is a well-characterized cohort study of women with high dimensional SNP data, as well as measurements on lifestyle and dietary factors. We consider SNPs on chronmosome 10 to identify main and gene-environment interactions associated with weight, which is an important phenotypic trait related to type 2 diabetes. Here, weight is used as response

and five environment factors, age (age), total physical activity (act), trans fat intake (trans), cereal fiber intake (ceraf) and reported high blood cholesterol (chol) are considered. Data are available on 3391 subjects and 17016 gene expressions after cleaning the raw data through matching phenotypes and genotypes and removing SNPs with minor allele frequency (MAF) less than 0.05. A prescreening is done before downstream analysis. We use a marginal linear model with weight as response and age, act, trans, ceraf, chol as environment factors. 10,000 SNPs which have at least two main or interaction effects with p-value less than 0.05 are kept.

We use Top 100 method to identify 100 most important main and interaction effects. The proposed method LADBLSS identifies 20 main SNP effects and 80 gene-environment interactions, which are listed in Table 8. Our study provides crucial implications in identifying the important main and interactions of SNPs and its associations with weight. For example, three SNPs, rs17011106, rs4838643 and rs17011115, located within gene WDFY4 are identified. WDFY4 has been observed as an influential factor related to weight and obesity ([24, 25]). In addition, SNPs rs10994364, rs10821773 and rs10994308, located within gene ANK3, are identified with interacting environment factors age and chol. There are findings showing an association between ANK3 and higher systolic blood pressure ([26]). Published studies have also shown that ANK3 is linked to plumonary and renal hypertension ([26]). Allele risk variants have been identified in ANK3, and these variants explain a proportion of the heritability of BD (bipolar disorder), which is associated with higher body mass index (BMI) and increased metabolic comorbidity and the genetic risk for BD relates to common genetic risk with T2D ([27]). Our proposed method identifies its interaction with chol, the high blood cholesterol.Data from several sources suggest that islet cholesterol metabolism contributes to the pathogenesis of T2D ([28]).

Analysis with alternatives BL, BLSS and LADBL has also been conducted. To compare the alternative methods with the proposed method, we provide the numbers of main effects and interactions identified by these methods with pairwise overlaps in Table 3. It clearly shows that the proposed one results in a very different set of effects compared to alternatives. We refit the regularized marginal models by LADBL and LADBLSS using robust Bayesian Lasso, and those identified by BL and BLSS using Bayesian Lasso. In addition, the inclusion probabilities of the selected main and interaction effects using LADBLSS are provided in Table 9. Results from the alternative methods are available from the Supplementary files. The proposed method selects the 100 most important effects with the inclusion probability larger than 0.9, which shows advantage in quantifying uncertain compared to marginal penalization methods ([8, 9, 10]).

Table 3: The numbers of main G effects and interactions identified by different approaches and their overlaps.

T2D	Main				Interaction			
	BL	BLSS	LADBL	LADBLSS	BL	BLSS	LADBL	LADBLSS
BL	86	5	6	8	14	14	4	8
BLSS		24	3	6		76	20	23
LADBL			20	12			80	50
LADBLSS				20				80

5 Discussion

In the past, G×E interaction studies have been mainly conducted through marginal hypothesis testing, based on a diversity of study designs utilizing parametric, nonparametric and semiparametric models ([29, 30, 31]), which later have been extended to joint analyses driven primarily by the pathway or gene set based association studies ([32, 33, 34]). In addition, published literature has also reported the success of marginal screening studies, including those based on partial correlations ([36, 35]). Recently, the effectiveness of regularized variable selection in G×E interaction studies has been increasingly recognized, and a large number of regularization methods have been proposed for joint interaction studies ([7]). Marginal penalization has also been demonstrated as promising competitors, although they have only been investigated in a limited number of frequentist studies ([8, 10, 9]).

Therefore, the proposed marginal robust Bayesian variable selection is of particular importance, since joint and marginal analysis cannot replace each other and marginal Bayesian penalization has not been examined for G×E studies so far. In particular, with the robustness and incorporation of spike- and-slab priors in the adaptive Bayesian shrinkage, the LADBLSS has an analysis framework more coherent with that of the joint robust analysis ([37]), which significantly facilitates methodological developments for interaction studies.

The marginal Bayesian regularization can be extended to different types of response, for example, under binary, categorical, prognostic and multivariate outcomes. Nevertheless, considering robustness in the generalized models with the Bayesian framework is not trivial, especially under the multivariate responses ([38, 39]). We postpone the investigations to the future studies.The interaction between genetic and environmental factors in this study has been modeled as the product of the two corresponding variables, which amounts to "linear" interactions. In practice, the linear interaction assumption has been frequently violated([40, 41, 42]), which demands accommodation of these nonlinear effects through nonparametic and semiparametric models ([13, 14, 43, 44]). It is of great interest and importance to migrate the nonlinear G×E studies to marginal cases in the near future.

References

[1] Hunter, D. J. (2005). Gene–environment interactions in human diseases. *Nature Reviews Genetics*, 6(4), 287-298.

[2] Simonds, N. I., Ghazarian, A. A., Pimentel, C. B., Schully, S. D., Ellison, G. L., Gillanders, E. M. and Mechanic, L. E. (2016). Review of the gene-environment interaction literature in cancer: what do we know?. *Genetic epidemiology*, 40(5), 356-365.

[3] Von Mutius, E. (2009). Gene-environment interactions in asthma. *Journal of Allergy and Clinical Immunology*, 123(1), 3-11.

[4] Cornelis, M. C. and Hu, F. B. (2012). Gene-environment interactions in the development of type 2 diabetes: recent progress and continuing challenges. *Annual review of nutrition*, 32, 245-259.

[5] Cordell, H. J. and Clayton, D. G. (2005). Genetic association studies. *The Lancet*, 366(9491), 1121-1131.

[6] Wu, C., Li, S. and Cui, Y. (2012). Genetic association studies: an information content perspective. *Current genomics*, 13(7), 566-573.

[7] Zhou, F., Ren, J., Lu, X., Ma, S. and Wu, C. (2021). Gene–Environment Interaction: a Variable Selection Perspective. Epistasis. *Methods in Molecular Biology*. Humana Press.

[8] Shi, X., Liu, J., Huang, J., Zhou, Y., Xie, Y. and Ma, S. (2014). A penalized robust method for identifying gene–environment interactions. *Genetic epidemiology*, 38(3), 220-230.

[9] Chai, H., Zhang, Q., Jiang, Y., Wang, G., Zhang, S., Ahmed, S. E. and Ma, S. (2017). Identifying gene-environment interactions for prognosis using a robust approach. *Econometrics and statistics*, 4, 105-120.

[10] Zhang, S., Xue, Y., Zhang, Q., Ma, C., Wu, M. and Ma, S. (2020). Identification of gene–environment interactions with marginal penalization. *Genetic epidemiology*, 44(2), 159-196.

[11] Huang, J. and Ma, S. (2010). Variable selection in the accelerated failure time model via the bridge method. *Lifetime data analysis*, 16(2), 176-195.

[12] Liu, C., Ma, J. and Amos, C. I. (2015). Bayesian variable selection for hierarchical gene–environment and gene–gene interactions. *Human genetics*, 134(1), 23-36.

[13] Li, J., Wang, Z., Li, R. and Wu, R. (2015). Bayesian group lasso for nonparametric varying-coefficient models with application to functional genome-wide association studies. *The annals of applied statistics*, 9(2), 640.

[14] Ren, J., Zhou, F., Li, X., Chen, Q., Zhang, H., Ma, S., Jiang, Y. and Wu, C. (2020). Semiparametric Bayesian variable selection for gene-environment interactions. *Statistics in Medicine*, 39(5), 617-638.

[15] George, E. I. and McCulloch, R. E. (1993). Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, 88(423), 881-889.

[16] Ishwaran, H. and Rao, J. S. (2005). Spike and slab variable selection: frequentist and Bayesian strategies. *The Annals of Statistics*, 33(2), 730-773.

[17] Wu, C. and Ma, S. (2015). A selective review of robust variable selection with applications in bioinformatics, *Briefings in Bioinformatics*, 16(5): 873-–883,

[18] Kozumi, H. and Kobayashi, G. (2011). Gibbs sampling methods for bayesian quantile regression. *Journal of Statistical Computation and Simulation*, 81(11):1565–1578, 2011. doi: 10.1080/00949655.2010.496117.

[19] Yu, K. and Moyeed, R. A. (2001). Bayesian quantile regression. *Statistics and Probability Letters*, 54(4): 437–447, 2001.

[20] Yu, K. and Zhang, J. (2005). A three-parameter asymmetric laplace distribution and its extension. *Communications in Statistics - Theory and Methods*, 34(9-10):1867—1879

[21] Li, Q., Xi, R., and Lin, N. (2010). Bayesian regularized quantile regression. *Bayesian Analysis*, 5(3):533—556.

[22] Brooks, S. P. and Gelman, A. (1998). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434–455, 1998. doi: 10.1080/10618600.1998.10474787.

[23] Gelman, A., Carlin, J. B., Stern, H.S., Dunson, D. B., Vehtari, A. and Rubin, D. B. *Bayesian Data Analysis*. Chapman and Hall/CRC, 2004.

[24] Barclay, S. F., Rand, C. M., Borch, L. A., Nguyen, L., Gray, P. A., Gibson, W. T., Wilson, R. J. A., Gordon, P. M. K., Aung, Z., Berry-Kravis, E. M., Ize-Ludlow, D., Weese-Mayer, D. E. and Bech-Hansen, N. T. (2015). Rapid-Onset Obesity with Hypothalamic Dysfunction, Hypoventilation, and Autonomic Dysregulation (ROHHAD): exome sequencing of trios, monozygotic twins and tumours. *Orphanet Journal of Rare Diseases*, 10:103, 2015. doi:10.1186/s13023-015-0314-x.

[25] Martin, C. L., Jima, D., Sharp, G. C., McCullough, L. E., Park, S. S., Gowdy, K. M., Skaar, D., Cowley, M., Maguire, R. L., Fuemmeler, B., Collier, D., Relton, C. L., Murphy, S. K. and Hoyo, C. (2019). Maternal pre-pregnancy obesity, offspring cord blood DNA methylation, and offspring cardiometabolic health in early childhood: an epigenome-wide association study. *Epigenetics*, 14:4, 325-340,2019. doi: 10.1080/15592294.2019.1581594.

[26] Ghanbari, M., Vries, P. S., Looper, H., Peters, M. J., Schurmann, C.,Yaghootkar, H., Dörr, M., Frayling, T. M., Uitterlinden, A. G.,Hofman, A.,Meurs, J. B. J., Erkeland, S. J.,Franco, O. H. and Dehghan, A. (2014). A Genetic Variant in the Seed Region of miR-4513 Shows Pleiotropic Effects on Lipid and Glucose Homeostasis, Blood Pressure, and Coronary Artery Disease. *Hum Mutat*, 35:1524–1531, 2014, doi: 10.1002/humu.22706.

[27] Winham, S. J., Cuellar-Barboza, A. B., Oliveros, A., McElroy, S. L., Crow, S. Colby, C., Choi, D. S., Chauhan, M., Frye, M. and Biernacka, J. M. (2014). Genome-wide association study of bipolar disorder accounting for effect of body mass index identifies a new risk allele in TCF7L2.*Molecular Psychiatry*, 19, 1010–1016; doi:10.1038/mp.2013.159.

[28] Brunham, L. R., Kruit, J. K., Verchere, C. B. and Hayden, M. R. (2008). Cholesterol in islet dysfunction and type 2 diabetes. *J. Clin. Invest*, 118:403–408 (2008). doi:10.1172/JCI33296.

[29] Murcray, C. E., Lewinger, J. P. and Gauderman, W. J. (2009). Gene-environment interaction in genome-wide association studies. *American journal of epidemiology*, 169(2), 219-226.

[30] Thomas, D. (2010). Methods for investigating gene-environment interactions in candidate pathway and genome-wide association studies. *Annual review of public health*, 31, 21-36.

[31] Mukherjee, B., Ahn, J., Gruber, S. B. and Chatterjee, N. (2012). Testing gene-environment interaction in large-scale case-control association studies: possible choices and comparisons. *American journal of epidemiology*, 175(3), 177-190.

[32] Wu, C. and Cui, Y. (2013). Boosting signals in gene-based association studies via efficient SNP selection. *Briefings in Bioinformatics*, 15(2), 279-291.

[33] Jin, L., Zuo, X., Su, W., Zhao, X., Yuan, M., Han, L.,Zhao, X., Chen, Y. and Rao, S.(2014). Pathway-based analysis tools for complex diseases: a review. *Genomics, Proteomics and Bioinformatics*, 12(5), 210-220.

[34] Jiang, Y., Huang, Y., Du, Y., Zhao, Y., Ren, J., Ma, S. and Wu, C. (2017). Identification of prognostic genes and pathways in lung adenocarcinoma using a Bayesian approach. *Cancer Informatics*, 1(7).

[35] Niu, Y. S., Hao, N., and Zhang, H. H. (2018). Interaction screening by partial correlation. *Statistics and Its Interface*, 11(2), 317–325.

[36] Xu, Y., Wu, M., Zhang, Q. and Ma, S. (2019). Robust identification of gene-environment interactions for prognosis using a quantile partial correlation approach. *Genomics*, 111(5), 1115–1123.

[37] Ren, J., Zhou, F., Li, X., Ma, S., Jiang, Y. and Wu, C. (2021). Robust Bayesian variable selection for gene-environment interactions. (under revision)

[38] Wu, C., Cui, Y., and Ma, S. (2014). Integrative analysis of gene–environment interactions under a multi-response partially linear varying coefficient model. *Statistics in medicine*, 33(28), 4988–4998.

[39] Zhou, F., Ren, J., Li, G., Jiang, Y., Li, X., Wang, W. and Wu, C. (2019). Penalized Variable Selection for Lipid–Environment interactions in a longitudinal lipidomics study. *Genes*, 10(12), 1002.

[40] Ma, S., Yang, L., Romero, R. and Cui, Y. (2011). Varying coefficient model for gene–environment interaction: a non-linear look. *Bioinformatics*, 27(15), 2119-2126.

[41] Wu, C. and Cui, Y. (2013). A novel method for identifying nonlinear gene–environment interactions in case–control association studies. *Human genetics*, 132(12), 1413-1425.

[42] Zhao, N., Zhang, H., Clark, J. J., Maity, A., and Wu, M. C. (2019). Composite kernel machine regression based on likelihood ratio test for joint testing of genetic and gene–environment interaction effect. *Biometrics*, 75(2), 625-637.

[43] Wu, C., Zhong, P. S. and Cui, Y. (2018). Additive varying-coefficient model for nonlinear gene-environment interactions. *Statistical applications in genetics and molecular biology*, 17(2).

[44] Wu, C., Shi, X., Cui, Y. and Ma, S. (2015). A penalized robust semiparametric approach for gene–environment interactions. *Statistics in medicine*, 34(30), 4016-4030.

A Additional simulation results

A.1 Identification results in simulation

Table 4: Simulation results of the second setting. AUC (mean of AUC), SD (sd of AUC) based on 100 replicates. n=200, p=500, q=4 and $m = 3$.

		BL	BLSS	LADBL	LADBLSS
Error 1	AUC	0.9089	0.9881	0.9148	0.9888
N(0,1)	SD	0.0059	0.0019	0.0051	0.0037
Error 2	AUC	0.8187	0.9255	0.8877	0.9769
$t(2)$	SD	0.0142	0.0524	0.0057	0.0048
Error 3	AUC	0.5333	0.5533	0.8239	0.9459
Lognormal(0,2)	SD	0.0096	0.0656	0.1045	0.0162
Error 4	AUC	0.8113	0.9122	0.9111	0.9849
90%N(0,1)+10%Cauchy(0,1)	SD	0.0166	0.0502	0.0083	0.0033
Error 5	AUC	0.7425	0.8086	0.9076	0.9856
80%N(0,1)+20%Cauchy(0,1)	SD	0.0241	0.0746	0.0065	0.0024

Table 5: Identification results of the second setting with Top100 method. mean(sd) based on 100 replicates. n=200, p=500, q=4 and $m = 3$.

		Main	Interaction	Total
Error 1	BL	7.50(0.86)	6.70(1.49)	14.20(1.83)
N(0,1)	BLSS	7.60(0.67)	10.20(0.09)	17.80(1.32)
	LADBL	7.67(0.66)	6.83(1.82)	14.5(1.96)
	LADBLSS	7.63(0.56)	9.97(1.54)	17.6(1.67)
Error 2	BL	5.83(2.21)	3.47(1.57)	9.30(2.98)
$t(2)$	BLSS	6.33(2.09)	7.57(3.15)	13.90(4.73)
	LADBL	7.07(0.94)	5.97(1.61)	13.03(1.96)
	LADBLSS	7.40(0.62)	9.20(1.94)	16.60(2.11)
Error 3	BL	0.77(0.86)	0.73(0.94)	1.50(1.11)
Lognormal(0,2)	BLSS	0.57(1.01)	0.67(1.06)	1.23(1.77)
	LADBL	5.90(1.65)	3.50(1.96)	9.40(2.43)
	LADBLSS	5.67(1.73)	9.00(2.35)	14.67(3.73)
Error 4	BL	6.03(2.19)	4.40(2.44)	10.43(4.17)
90%N(0,1)	BLSS	6.03(2.57)	8.00(3.33)	14.03(5.76)
+10%Cauchy(0,1)	LADBL	7.27(0.91)	6.87(1.48)	14.13(1.74)
	LADBLSS	7.53(0.63)	10.00(1.43)	17.53(1.57)
Error 5	BL	5.53(2.45)	3.63(2.19)	9.16(4.13)
80%N(0,1)	BLSS	5.07(2.57)	6.73(3.37)	11.80(5.65)
+20%Cauchy(0,1)	LADBL	7.47(0.97)	5.43(1.77)	12.90(2.04)
	LADBLSS	7.37(0.85)	10.47(1.46)	17.83(1.91)

Table 6: Simulation results of the third setting. AUC (mean of AUC), SD (sd of AUC) based on 100 replicates. n=200, p=500, q=4 and $m = 3$.

		BL	BLSS	LADBL	LADBLSS
Error 1	AUC	0.9158	0.9895	0.9251	0.9878
N(0,1)	SD	0.0041	0.0022	0.0054	0.0028
Error 2	AUC	0.8323	0.9461	0.8972	0.9833
$t(2)$	SD	0.0117	0.0342	0.0062	0.0028
Error 3	AUC	0.5268	0.5531	0.8415	0.9595
Lognormal(0,2)	SD	0.0127	0.0590	0.0107	0.0156
Error 4	AUC	0.8261	0.9323	0.9245	0.9889
90%N(0,1)+10%Cauchy(0,1)	SD	0.0191	0.0352	0.0056	0.0034
Error 5	AUC	0.7533	0.8591	0.9204	0.9862
80%N(0,1)+20%Cauchy(0,1)	SD	0.0201	0.0657	0.0067	0.0114

Table 7: Identification results of the third setting with Top100 method. mean(sd) based on 100 replicates. n=200, p=500, q=4 and $m = 3$.

		Main	Interaction	Total
Error 1	BL	7.70(0.47)	6.80(1.63)	14.50(1.79)
N(0,1)	BLSS	7.63(0.72)	10.93(0.98)	18.57(1.22)
	LADBL	7.70(0.75)	7.33(1.95)	15.03(2.14)
	LADBLSS	7.87(0.35)	10.33(1.35)	18.20(1.45)
Error 2	BL	6.57(1.87)	4.47(1.69)	11.03(2.88)
$t(2)$	BLSS	6.60(1.57)	8.40(2.51)	15.00(3.68)
	LADBL	7.57(0.62)	5.77(1.50)	13.33(1.77)
	LADBLSS	7.43(0.68)	9.30(2.15)	16.73(2.43)
Error 3	BL	0.50(0.73)	0.83(1.02)	1.33(1.47)
Lognormal(0,2)	BLSS	0.70(0.99)	0.40(0.86)	1.10(1.54)
	LADBL	6.13(2.05)	3.80(1.39)	9.93(1.32)
	LADBLSS	6.63(1.16)	10.10(1.73)	16.73(2.52)
Error 4	BL	5.73(2.82)	4.30(2.64)	10.03(5.11)
90%N(0,1)	BLSS	5.73(3.02)	7.67(4.19)	13.40(7.05)
+10%Cauchy(0,1)	LADBL	7.80(0.48)	6.87(1.61)	14.67(1.54)
	LADBLSS	7.83(0.38)	10.50(1.25)	18.33(1.39)
Error 5	BL	5.60(2.61)	2.93(2.23)	8.53(4.27)
80%N(0,1)	BLSS	5.27(2.27)	6.90(3.64)	12.17(5.66)
+20%Cauchy(0,1)	LADBL	7.87(0.35)	6.87(1.45)	14.73(1.46)
	LADBLSS	7.70(0.53)	10.70(1.12)	18.40(1.28)

B Assessment of the convergence of MCMC chains

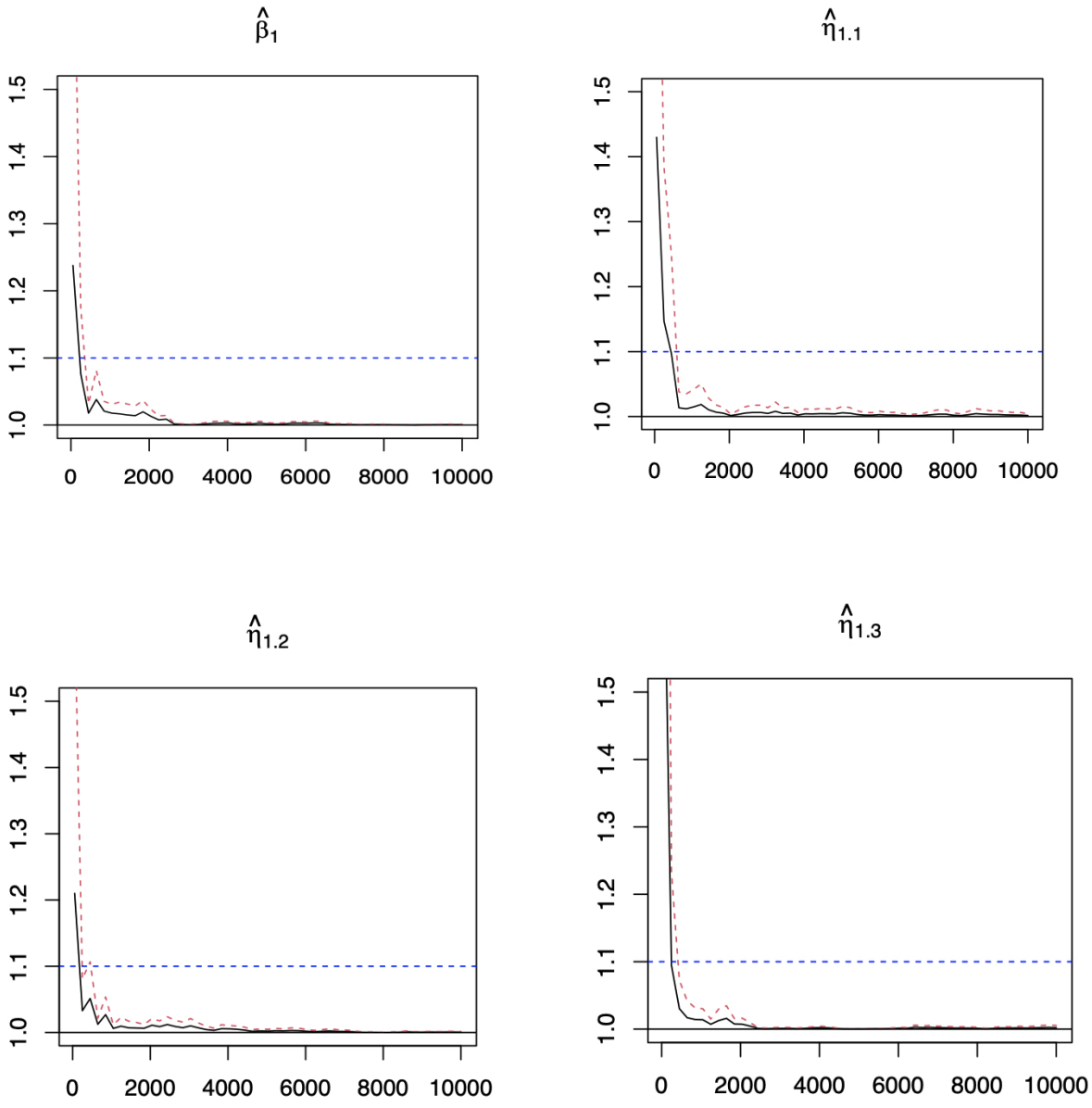

Figure 1: Potential scale reduction factor (PSRF) against iterations for the coefficients of the first genetic factors and its interaction with environmental factors in Example 1 under Error 3. Black line: the PSRF. Red dotted line: the upper limits of the 95% confidence interval for the PSRF. Blue dotted line: the threshold of 1.1. The $\hat{\beta}_1$ represents the estimated coefficients of the main effects for the first genetic factor. The $\hat{\eta}_{11}$ to $\hat{\eta}_{13}$ represent the estimated coefficients of the first three interaction effects for the first genetic factor.

C Estimation results for data analysis

Table 8: Analysis of the NHS T2D data using LADBLSS.

			Interactions				
SNP	Gene	Main Effects	age	act	trans	ceraf	chol
rs17011106	WDFY4	-0.024					
rs7077294	KIAA1217						-0.0491
rs7093682	RP11-170M17.1				-0.1239		
rs17011106	WDFY4	-0.0953					
rs10826028	MIR3924					-0.0524	
rs4748996	THNSL1						0.0064
rs2646392	KRT8P37	0.0148					
rs7904629	RP11-170M17.1				-0.0592		
rs1244416	ATP5C1						0.0851
rs4838643	WDFY4	-0.0051					
rs1916458	RP11-170M17.1				-0.0264		
rs1537615	RP11-526P5.2	0.0477					
rs2765398	KRT8P37	-0.0157					
rs4317891	CELF2		0.0647				
rs7922793	LINC00845	-0.0345					
rs1916412	RP11-170M17.1				-0.0614		
rs1916411	RP11-170M17.1				-0.0448		
rs4747800	KRT8P37	0.0036					
rs11258040	CAMK1D					-0.0983	
rs1984275	RP11-319F12.2					0.0065	
rs17432763	MIR5100						-0.0677
rs10796113	FRMD4A	-0.0931					
rs224765	RP11-490O24.2		-0.0521				
rs6482387	KIAA1217						0.011
rs1492608	ENKUR						-0.0287
rs11257323	ECHDC3						0.0084
rs4434904	KIAA1217						-0.0374
rs10994364	ANK3		0.1086				
rs12220246	KIAA1462				0.0371		
rs11010390	RP11-309N24.1		0.0271				
rs10828584	KIAA1217	0.087					
rs10857590	ARHGAP22				-0.1379		
rs1537616	RP11-526P5.2	0.086					
rs17295031	KIAA1462				-0.0468		
rs10905778	RP11-271F18.4						0.0055
rs7093161	SNRPEP8		-0.0577				0.0107
rs2377872	CHAT					-0.0213	
rs1916409	RP11-170M17.1				-0.0642		

Continued on the next page

Table 8: Continued from the previous page.

			Interactions				
SNP	Gene	Main Effects	age	act	trans	ceraf	chol
rs2245456	MALRD1		-0.0042				
rs787116	RP11-478H13.1		0.0259				
rs2817825	RP11-492M23.2		0.0278				
rs11255338	KIN						-0.0401
rs17011115	WDFY4	-0.0045					
rs11010821	Y-RNA	-0.025					
rs2532760	RP11-492M23.2		0.0272				
rs10821773	ANK3		-0.0107				
rs17454012	CELF2		-0.1076				
rs4372368	RP11-478B11.2		-0.0449				
rs1916420	RP11-170M17.1				0.0885		
rs2446588	FRMD4A		-0.0142				
rs10995687	RP11-170M17.1				0.1065		
rs161279	RP11-192P3.5				0.0333		
rs161279	ZEB1				0.0333		
rs161258	ZEB1				0.0362		
rs10509149	TMEM26						-0.0428
rs3740000	LINC00837	-0.106					
rs17314489	ZNF365		0.1543				
rs17453876	CELF2		0.0518				
rs10793451	ZNF485				-0.1028		
rs4749527	KIAA1462				0.0093		
rs12570207	SEPHS1		-0.0329				
rs902904	THNSL1						-0.0885
rs7921813	CAMK1D		0.0063				
rs10218945	SNRPEP8						-0.0351
rs2804551	RP11-492M23.2	0.0616					
rs12266433	CELF2		-0.0301				
rs16919385	PLXDC2						0.0112
rs4750039	CELF2		0.0333				
rs12249964	KIAA1217						0.0607
rs4745829	RP11-170M17.1				-0.0252		
rs11257932	CAMK1D		0.0256				
rs10827602	RP11-810B23.1			0.0167			
rs7081466	RP11-526P5.2					0.0267	
rs12256642	THNSL1						-0.0258
rs2796304	RP11-492M23.2		0.0875				
rs10826964	ZEB1				0.0316		
rs11257933	CAMK1D		0.0547				
rs17432532	MIR5100						0.017

Continued on the next page

Table 8: Continued from the previous page.

			Interactions				
SNP	Gene	Main Effects	age	act	trans	ceraf	chol
rs10826899	UBE2V2P1					-0.0234	
rs11592473	UBE2V2P1					-0.0642	
rs12764778	OR13A1		0.0263				
rs12762462	GPR158				-0.0153		
rs1011763	MIR3924					0.1871	
rs1916450	RP11-170M17.1				-0.1767		
rs1917814	CHAT					-0.0194	
rs6602809	DCLRE1CP1	0.0043					
rs923757	THNSL1						0.0226
rs7092368	RP11-526P5.2					-0.0531	
rs6602806	DCLRE1CP1	0.0527					
rs6602806	ACBD7	0.0527					
rs10994308	ANK3						-0.0124
rs224699	RP11-490O24.2		-0.0351				
rs7083349	KIAA1217	-0.0651					
rs10828905	RNU6-632P		-0.0799				

Table 9: Inclusion probability of the NHS T2D data using LADBLSS.

SNP	Gene	Main Effects	age	act	trans	ceraf	chol
rs17011106	WDFY4	0.9930					
rs7077294	KIAA1217						0.9736
rs7093682	RP11-170M17.1				0.9938		
rs17011106	WDFY4	0.9900					
rs10826028	MIR3924					0.9612	
rs4748996	THNSL1						0.9834
rs2646392	KRT8P37	0.9818					
rs7904629	RP11-170M17.1				0.9646		
rs1244416	ATP5C1						0.9656
rs4838643	WDFY4	0.9768					
rs1916458	RP11-170M17.1				0.9832		
rs1537615	RP11-526P5.2	0.9956					
rs2765398	KRT8P37	0.9756					
rs4317891	CELF2		1.000				
rs7922793	LINC00845	0.9744					
rs1916412	RP11-170M17.1				0.9774		
rs1916411	RP11-170M17.1				0.9700		
rs4747800	KRT8P37	0.9840					
rs11258040	CAMK1D					0.9738	

Continued on the next page

Table 9: Continued from the previous page.

SNP	Gene	Main Effects	age	act	trans	ceraf	chol
rs1984275	RP11-319F12.2					0.9952	
rs17432763	MIR5100						0.9862
rs10796113	FRMD4A	0.9636					
rs224765	RP11-490O24.2		0.9710				
rs6482387	KIAA1217						0.9638
rs1492608	ENKUR						0.9680
rs11257323	ECHDC3						0.9892
rs4434904	KIAA1217						0.9716
rs10994364	ANK3		0.9942				
rs12220246	KIAA1462				0.9610		
rs11010390	RP11-309N24.1		0.9820				
rs10828584	KIAA1217	0.9752					
rs10857590	ARHGAP22				0.9848		
rs1537616	RP11-526P5.2	0.9944					
rs17295031	KIAA1462				0.9816		
rs10905778	RP11-271F18.4						0.9988
rs7093161	SNRPEP8		0.9542				0.9902
rs2377872	CHAT					0.9728	
rs1916409	RP11-170M17.1				0.9612		
rs2245456	MALRD1		0.9630				
rs787116	RP11-478H13.1		0.9638				
rs2817825	RP11-492M23.2		0.9550				
rs11255338	KIN						0.9964
rs17011115	WDFY4	0.9712					
rs11010821	Y-RNA	0.9916					
rs2532760	RP11-492M23.2		0.9720				
rs10821773	ANK3		0.9586				
rs17454012	CELF2		0.9998				
rs4372368	RP11-478B11.2		0.9618				
rs1916420	RP11-170M17.1				0.9672		
rs2446588	FRMD4A		0.9724				
rs10995687	RP11-170M17.1				0.9588		
rs161279	RP11-192P3.5				0.9770		
rs161279	ZEB1				0.9770		
rs161258	ZEB1				0.9876		
rs10509149	TMEM26						0.9726
rs3740000	LINC00837	0.9964					
rs17314489	ZNF365		0.9866				
rs17453876	CELF2		0.9952				
rs10793451	ZNF485				0.9604		
rs4749527	KIAA1462				0.9794		

Continued on the next page

Table 9: Continued from the previous page.

SNP	Gene	Main Effects	age	act	trans	ceraf	chol
rs12570207	SEPHS1		0.9698				
rs902904	THNSL1						0.9884
rs7921813	CAMK1D		0.9998				
rs10218945	SNRPEP8						0.9612
rs2804551	RP11-492M23.2	0.9848					
rs12266433	CELF2		0.9618				
rs16919385	PLXDC2						0.9806
rs4750039	CELF2		0.9910				
rs12249964	KIAA1217						0.9558
rs4745829	RP11-170M17.1				0.9940		
rs11257932	CAMK1D		0.9826				
rs10827602	RP11-810B23.1			0.9728			
rs7081466	RP11-526P5.2					0.9714	
rs12256642	THNSL1						0.9616
rs2796304	RP11-492M23.2		0.9928				
rs10826964	ZEB1				0.9592		
rs11257933	CAMK1D		0.9726				
rs17432532	MIR5100						0.9834
rs10826899	UBE2V2P1					0.9784	
rs11592473	UBE2V2P1					0.9864	
rs12764778	OR13A1		0.9894				
rs12762462	GPR158				0.9636		
rs1011763	MIR3924					0.9954	
rs1916450	RP11-170M17.1				0.9820		
rs1917814	CHAT					0.9670	
rs6602809	DCLRE1CP1	0.9614					
rs923757	THNSL1						0.9964
rs7092368	RP11-526P5.2					0.9868	
rs6602806	DCLRE1CP1	0.9912					
rs6602806	ACBD7	0.9912					
rs10994308	ANK3						0.9542
rs224699	RP11-490O24.2		0.9768				
rs7083349	KIAA1217	0.9986					
rs10828905	RNU6-632P		0.9626				

D Posterior inference

D.1 LADBL

D.1.1 Hierarchical model specification

$$
\begin{aligned}
Y_i &= E_i\alpha + C_i\gamma + X_{ij}\beta_j + \tilde{W}_i\eta_j + \tau^{-1/2}\xi_2\sqrt{v_i}z_i \quad i = 1,\ldots,n \\
v_i|\tau &\overset{iid}{\sim} \tau\exp(-\tau v_i) \quad i = 1,\ldots,n \\
z_i &\overset{iid}{\sim} \mathrm{N}(0,1) \quad i = 1,\ldots,n \\
\beta_j|s_1 &\sim \frac{1}{\sqrt{2\pi s_1}}\exp(-\frac{\beta_j^2}{2s_1}) \\
s_1|\varphi_1^2 &\sim \frac{\varphi_1^2}{2}\exp(-\frac{\varphi_1^2}{2}s_1) \\
\eta_{jk}|s_{2k} &\overset{iid}{\sim} \frac{1}{\sqrt{2\pi s_{2k}}}\exp(-\frac{\eta_{jk}^2}{2s_{2k}}) \quad k = 1,\ldots,q \\
s_{2k}|\varphi_2^2 &\overset{iid}{\sim} \frac{\varphi_2^2}{2}\exp(-\frac{\varphi_2^2}{2}s_{2k}) \quad k = 1,\ldots,q \\
\alpha_k &\overset{iid}{\sim} \frac{1}{\sqrt{(2\pi\alpha_0)}}\exp(-\frac{\alpha_k^2}{2\alpha_0}) \quad k = 1,\ldots,q \\
\gamma_t &\overset{iid}{\sim} \frac{1}{\sqrt{(2\pi\gamma_0)}}\exp(-\frac{\gamma_t^2}{2\gamma_0}) \quad t = 1,\ldots,m \\
\tau &\sim \mathrm{Gamma}(a,b) \\
\varphi_1^2 &\sim \mathrm{Gamma}(c_1,d_1) \\
\varphi_2^2 &\sim \mathrm{Gamma}(c_2,d_2)
\end{aligned}
$$

D.1.2 Gibbs Sampler

Let $\mu_{(-\alpha_k)} = E(y_i) - E_{ik}\alpha_k$, then

$$
\begin{aligned}
&\pi(\alpha_k|\mathrm{rest}) \\
&\quad\propto \pi(Y|\cdot)\pi(\alpha_k) \\
&\quad\propto \exp\Big\{-\sum_{i=1}^{n}\frac{(y_i - E_i\alpha - C_i\gamma - X_{ij}\beta_j - \tilde{W}_i\eta_j)^2}{2\tau^{-1}\xi_2^2 v_i}\Big\} \times \exp(-\frac{\alpha_k^2}{2\alpha_0}) \\
&\quad\propto \exp\Big\{-\frac{1}{2}\big[(\sum_{i=1}^{n}\frac{\tau E_{ik}^2}{\xi_2^2 v_i} + \frac{1}{\alpha_0})\alpha_k^2 - 2\sum_{i=1}^{n}\frac{\tau(y_i - \mu_{(-\alpha_k)})E_{ik}}{\xi_2^2 v_i}\alpha_k\big]\Big\}.
\end{aligned}
$$

Hence, $\alpha_k|\text{rest}\sim N(\mu_{\alpha_k}, \sigma^2_{\alpha_k})$, where

$$\mu_{\alpha_k} = \big(\sum_{i=1}^{n} \frac{\tau(y_i - \mu_{(-\alpha_k)})E_{ik}}{\xi_2^2 v_i}\big)\sigma^2_{\alpha_k},$$
$$\sigma^2_{\alpha_k} = \big(\sum_{i=1}^{n} \frac{\tau E_{ik}^2}{\xi_2^2 v_i} + \frac{1}{\alpha_0}\big)^{-1}.$$

Let $\mu_{(-\gamma_t)} = E(y_i) - C_{it}\gamma_t$, So $\gamma_t|\text{rest}\sim N(\mu_{\gamma_k}, \sigma^2_{\gamma_t})$, where

$$\mu_{\gamma_t} = \big(\sum_{i=1}^{n} \frac{\tau(y_i - \mu_{(-\gamma_t)})C_{it}}{\xi_2^2 v_i}\big)\sigma^2_{\gamma_t},$$
$$\sigma^2_{\gamma_t} = \big(\sum_{i=1}^{n} \frac{\tau C_{it}^2}{\xi_2^2 v_i} + \frac{1}{\gamma_0}\big)^{-1}.$$

Let $\mu_{(-\beta_j)} = E(y_i) - X_{ij}\beta_j$, then

$$\begin{aligned}
&\pi(\beta_j|\text{rest}) \\
&\quad\propto \pi(y|\cdot)\pi(\beta_j|s_1) \\
&\quad\propto \exp\Big\{-\sum_{i=1}^{n} \frac{(y_i - E_i\alpha - C_i\gamma - X_{ij}\beta_j - \tilde{W}_i\eta_j)^2}{2\tau^{-1}\xi_2^2 v_i}\Big\} \times \exp(-\frac{\beta_j^2}{2s_1}) \\
&\quad\propto \exp\Big\{-\frac{1}{2}\big[(\sum_{i=1}^{n} \frac{\tau X_{ij}^2}{\xi_2^2 v_i} + \frac{1}{s_1})\beta_j^2 - 2\sum_{i=1}^{n} \frac{\tau(y_i - \mu_{(-\beta_j)})X_{ij}}{\xi_2^2 v_i}\beta_j\big]\Big\}.
\end{aligned}$$

So, $\beta_j|\text{rest}\sim N(\mu_{\beta_j}, \sigma^2_{\beta_j})$ with

$$\mu_{\beta_j} = \big(\sum_{i=1}^{n} \frac{\tau(y_i - \mu_{(-\beta_j)})X_{ij}}{\xi_2^2 v_i}\big)\sigma^2_{\beta_j},$$
$$\sigma^2_{\beta_j} = \big(\sum_{i=1}^{n} \frac{\tau X_{ij}^2}{\xi_2^2 v_i} + \frac{1}{s_1}\big)^{-1}.$$

Let $\mu_{(-\eta_{jk})} = E(y_i) - W_{ik}\eta_{jk}$, then $\eta_{jk}|\text{rest}\sim N(\mu_{\eta_{jk}}, \sigma^2_{\eta_{jk}})$, where

$$\mu_{\eta_{jk}} = \big(\sum_{i=1}^{n} \frac{\tau(y_i - \mu_{(-\eta_{jk})})\tilde{W}_{ik}}{\xi_2^2 v_i}\big)\sigma^2_{\eta_{jk}},$$
$$\sigma^2_{\beta_j} = \big(\sum_{i=1}^{n} \frac{\tau \tilde{W}_{ik}^2}{\xi_2^2 v_i} + \frac{1}{s_{2k}}\big)^{-1}.$$

The full conditional posterior distribution of s_1 is:

$$\begin{aligned} & s_1|\text{rest} \\ & \quad \propto \pi(\beta_j|s_1)\pi(s_1|\varphi_1^2) \\ & \quad \propto \frac{1}{\sqrt{s_1}}\exp(-\frac{\varphi_1^2}{2}s_1)\exp(-\frac{\beta_j^2}{2s_1}) \\ & \quad \propto \frac{1}{\sqrt{s_1}}\exp\Big\{-\frac{1}{2}[\varphi_1^2 s_1+\frac{\beta_j^2}{s_1}]\Big\}. \end{aligned}$$

Therefore, $s_1^{-1}|\text{rest}\sim$ Inverse-Gaussian$(\sqrt{\frac{\varphi_1^2}{\beta_j^2}},\varphi_1^2)$.

Similiarly, for $s_{2k}(k=1,\ldots,q)$, the posterior distribution for is $s_{2k}^{-1}|\text{rest}\sim$ Inverse-Gaussian$(\sqrt{\frac{\varphi_2^2}{\eta_{jk}^2}},\varphi_2^2)$.

The full conditional posterior distribution of φ_1^2 is:

$$\begin{aligned} & \varphi_1^2|\text{rest} \\ & \quad \propto \pi(s_1|\varphi_1^2)\pi(\varphi_1^2) \\ & \quad \propto \frac{\varphi_1^2}{2}\exp(-\frac{\varphi_1^2 s_1}{2})(\varphi_1^2)^{c_1-1}\exp(-d_1\varphi_1^2) \\ & \quad \propto (\varphi_1^2)^{c_1}\exp\Big(-\varphi_1^2(s_1/2+d_1)\Big). \end{aligned}$$

Therefore, the posterior distribution for φ_1^2 is Gamma$(c_1+1, s_1/2+d_1)$.

The full conditional posterior distribution of φ_2^2 is:

$$\begin{aligned} & \varphi_2^2|\text{rest} \\ & \quad \propto \pi(s_2|\varphi_2^2)\pi(\varphi_2^2) \\ & \quad \propto \prod_{k=1}^{q}\frac{\varphi_2^2}{2}\exp(-\frac{\varphi_2^2 s_{2k}}{2})(\varphi_2^2)^{c_2-1}\exp(-d_2\varphi_2^2) \\ & \quad \propto (\varphi_2^2)^{q+c_2-1}\exp\Big(-\varphi_2^2(\sum_{k=1}^{q}\frac{s_{2k}}{2}+d_2)\Big). \end{aligned}$$

The posterior distribution for φ_2^2 is Gamma$(c_2+q, \sum_{k=1}^{q} s_{2k}/2+d_2)$.

The full conditional posterior distribution of τ:

$$\begin{aligned} & \tau|\text{rest} \\ & \quad \propto \pi(v|\tau)\pi(\tau)\pi(Y|\cdot) \\ & \quad \propto \tau^{n/2}\exp\Big\{-\sum_{i=1}^{n}\frac{(y_i-E_i\alpha-C_i\gamma-X_{ij}\beta_j-\tilde{W}_i\eta_j)^2}{2\tau^{-1}\xi_2^2 v_i}\Big\} \\ & \quad \times \tau^n\exp(-\tau\sum_{i=1}^{n}v_i)\tau^{a-1}\exp(-b\tau) \\ & \quad \propto \tau^{a+\frac{3}{2}n-1}\exp\Big\{-\tau\big[\sum_{i=1}^{n}(\frac{(y_i-E_i\alpha-C_i\gamma-X_{ij}\beta_j-\tilde{W}_i\eta_j)^2}{2\xi_2^2 v_i}+v_i)+b\big]\Big\}. \end{aligned}$$

Therefore, $\tau|\text{rest} \sim \text{Gamma}(a + \frac{3}{2}n,\ \big[\sum_{i=1}^{n}(\frac{(y_i - E_i\alpha - C_i\gamma - X_{ij}\beta_j - \tilde{W}_i\eta_j)^2}{2\xi_2^2 v_i} + v_i) + b\big])$.

The full conditional posterior distribution of v_i is:

$$
\begin{aligned}
& v_i|\text{rest} \\
& \quad \propto \pi(v|\tau)\pi(y|\cdot) \\
& \quad \propto \frac{1}{\sqrt{v_i}} \exp\Big\{ - \frac{(y_i - E_i\alpha - C_i\gamma - X_{ij}\beta_j - \tilde{W}_i\eta_j)^2}{2\tau^{-1}\xi_2^2 v_i} \Big\} \times \exp(-\tau v_i) \\
& \quad \propto \frac{1}{\sqrt{v_i}} \exp\Big\{ - \frac{1}{2}\big[(2\tau)v_i + \frac{\tau(y_i - E_i\alpha - C_i\gamma - X_{ij}\beta_j - \tilde{W}_i\eta_j)^2}{\xi_2^2 v_i}\big] \Big\}.
\end{aligned}
$$

Therefore,

$$
\frac{1}{v_i}|\text{rest} \sim \text{ Inverse-Gaussian}(\sqrt{\frac{2\xi_2^2}{(y_i - E_i\alpha - C_i\gamma - X_{ij}\beta_j - \tilde{W}_i\eta_j)^2}},\ 2\tau).
$$

D.2 BLSS

D.2.1 Hierarchical model specification

$$
\begin{gathered}
Y \propto (\sigma^2)^{-\frac{n}{2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - E_i\alpha - C_i\gamma - X_{ij}\beta_j - \tilde{W}_i\eta_j)^2 \right\} \\
\alpha \sim \mathrm{N}_q(0,\ \Sigma_{\alpha 0}) \\
\gamma \sim \mathrm{N}_m(0,\ \Sigma_{\gamma 0}) \\
\beta_j|\pi_c, \tau_c^2, \sigma^2 \sim (1 - \pi_c)\,\mathrm{N}\left(0,\ \sigma^2\tau_c^2\right) + \pi_c\,\delta_0(\beta_j) \quad j = 1, \ldots, p \\
\eta_{jk}|\pi_e, \tau_{ek}^2, \sigma^2 \overset{iid}{\sim} (1 - \pi_e)\,\mathrm{N}\left(0,\ \sigma^2\tau_{ek}^2\right) + \pi_e\,\delta_0(\eta_{jk}) \quad j = 1, \ldots, p,\ k = 1, \ldots, q \\
\tau_c^2|\lambda_c^2 \sim \text{Gamma}(1, \frac{\lambda_c^2}{2}) \\
\tau_{ek}^2|\lambda_e^2 \overset{iid}{\sim} \text{Gamma}(1, \frac{\lambda_e^2}{2}) \quad k = 1, \ldots, q \\
\pi_c \sim \text{Beta}\,(r_c,\ u_c) \\
\pi_e \sim \text{Beta}\,(r_e,\ u_e) \\
\lambda_c^2 \sim \text{Gamma}\,(a_c,\ b_c) \\
\lambda_e^2 \sim \text{Gamma}\,(a_e,\ b_e) \\
\sigma^2 \sim \text{Inverse-Gamma}\,(s,\ h)
\end{gathered}
$$

D.2.2 Gibbs Sampler

Denote $\mu_{(-\alpha)} = \mathrm{E}(Y) - E\alpha$, then $\alpha|\text{rest} \sim \mathrm{N}(\mu_\alpha,\ \Sigma_\alpha)$, where

$$
\begin{aligned}
\mu_\alpha &= \Sigma_\alpha(\frac{1}{\sigma^2}(Y - \mu_{(-\alpha)})^\top E)^\top, \\
\Sigma_\alpha &= \left(\frac{1}{\sigma^2}E^\top E + \Sigma_{\alpha 0}^{-1}\right)^{-1}.
\end{aligned}
$$

Denote $\mu_{(-\gamma)} = \mathrm{E}(Y) - C\gamma$, then $\gamma|\text{rest} \sim \mathrm{N}(\mu_\gamma,\, \Sigma_\gamma)$, where

$$\mu_\gamma = \Sigma_\gamma(\frac{1}{\sigma^2}(Y - \mu_{(-\gamma)})^\top C)^\top,$$
$$\Sigma_\gamma = \left(\frac{1}{\sigma^2}C^\top C + \Sigma_{\gamma 0}^{-1}\right)^{-1}.$$

Denote $\mu_{(-\beta_j)} = \mathrm{E}(Y) - X_j\beta_j$, then $\beta_j|\text{rest} \sim (1-l_c)\mathrm{N}(\mu_{\beta_j},\, \sigma^2\Sigma_{\beta_j}) + l_c\delta_0(\beta_j)$, where

$$\mu_{\beta_j} = \Sigma_{\beta_j}X_j^\top(Y - \mu_{(-\beta_j)}),$$
$$\Sigma_{\beta_j} = \left(X_j^\top X_j + \frac{1}{\tau_c^2}\right)^{-1},$$
$$l_c = \frac{\pi_c}{\pi_c + (1-\pi_c)(\tau_c^2)^{-1/2}|\Sigma_{\beta_j}|^{1/2}\exp\left\{\frac{1}{2\sigma^2}\Sigma_{\beta_j}||X_j^\top(Y - \mu_{(-\beta_j)}||_2^2\right\}}.$$

Denote $\mu_{(-\eta_{jk})} = \mathrm{E}(Y) - \tilde{W}_k\eta_{jk}$, then $\eta_{jk}|\text{rest} \sim (1-l_{ek})\mathrm{N}(\mu_{\eta_{jk}},\, \sigma^2\Sigma_{\eta_{jk}}) + l_e\delta_0(\eta_{jk})$, where

$$\mu_{\eta_{jk}} = \Sigma_{\eta_{jk}}\tilde{W}_k^\top(Y - \mu_{(-\eta_{jk})}),$$
$$\Sigma_{\eta_{jk}} = \left(\tilde{W}_k^\top\tilde{W}_k + \frac{1}{\tau_{ek}^2}\right)^{-1},$$
$$l_e = \frac{\pi_e}{\pi_e + (1-\pi_e)(\tau_{ek}^2)^{-1/2}|\Sigma_{\eta_{jk}}|^{1/2}\exp\left\{\frac{1}{2\sigma^2}\Sigma_{\eta_{jk}}||\tilde{W}_k^\top(Y - \mu_{(-\eta_{jk})})||_2^2\right\}}.$$

The posterior of τ_c^2 is:

$$\frac{1}{\tau_c^2}|\text{rest} \sim \begin{cases}\text{Inverse-Gamma}(1, \frac{\lambda_c^2}{2}) & \text{if } \beta_j = 0\\ \text{Inverse-Gaussian}(\sqrt{\frac{\sigma^2}{\beta_j^2}\lambda_c^2},\, \lambda_c^2) & \text{if } \beta_j \neq 0\end{cases}.$$

The posterior of τ_{ek}^2 is:

$$\frac{1}{\tau_{ek}^2}|\text{rest} \sim \begin{cases}\text{Inverse-Gamma}(1, \frac{\lambda_e^2}{2}) & \text{if } \eta_{jk} = 0\\ \text{Inverse-Gaussian}(\sqrt{\frac{\sigma^2}{\eta_{jk}^2}\lambda_e^2},\, \lambda_e^2) & \text{if } \eta_{jk} \neq 0\end{cases}.$$

λ_c^2 and λ_e^2 have Gamma posterior distributions:

$$\lambda_c^2|\text{rest} \sim \text{Gamma}(a_c + 1,\ \frac{\tau_c^2}{2} + b_c),$$
$$\lambda_e^2|\text{rest} \sim \text{Gamma}(a_e + q,\ \sum_{k=1}^{q}\frac{\tau_{ek}^2}{2} + b_e).$$

π_c and π_e have Gamma posterior distributions:

$$\pi_c|\text{rest} \sim \text{Beta}(r_c - \mathbf{I}_{\{\beta_j \neq 0\}} + 1,\ u_c + \mathbf{I}_{\{\beta_j \neq 0\}}),$$
$$\pi_e|\text{rest} \sim \text{Beta}(r_e - \sum_{k=1}^{q}\mathbf{I}_{\{\eta_{jk} \neq 0\}} + q,\ u_e + \sum_{k=1}^{q}\mathbf{I}_{\{\eta_{jk} \neq 0\}}).$$

$\sigma^2 \sim$ Inverse-Gamma($\mu_{\sigma^2},\ \Sigma_{\sigma^2}$), where

$$\mu_{\sigma^2} = s + \frac{n + \mathbf{I}_{\{\beta_j \neq 0\}} + \sum_{k=1}^{q} \mathbf{I}_{\{\eta_{jk} \neq 0\}}}{2},$$
$$\Sigma_{\sigma^2} = h + \frac{(Y-\mu)^\top (Y-\mu) + (\tau_c^2)^{-1}\beta_j^2 + \sum_{k=1}^{q} (\tau_{ek}^2)^{-1} \eta_j^\top \eta_j}{2}.$$

D.3 BL

D.3.1 Hierarchical model specification

$$Y \propto (\sigma^2)^{-\frac{n}{2}} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (y_i - E_i\alpha - C_i\gamma - X_{ij}\beta_j - \tilde{W}_i\eta_j)^2 \right\}$$
$$\alpha \sim \mathrm{N}_q(0,\ \Sigma_{\alpha 0})$$
$$\gamma \sim \mathrm{N}_m(0,\ \Sigma_{\gamma 0})$$
$$\beta_j | \tau_c^2, \sigma^2 \sim \mathrm{N}\left(0,\ \sigma^2\tau_c^2\right) \quad j = 1, \ldots, p$$
$$\eta_{jk} | \tau_{ek}^2, \sigma^2 \overset{iid}{\sim} \mathrm{N}\left(0,\ \sigma^2\tau_{ek}^2\right) \quad j = 1, \ldots, p,\ k = 1, \ldots, q$$
$$\tau_c^2 | \lambda_c^2 \sim \exp(\frac{\lambda_c^2}{2})$$
$$\tau_{ek}^2 | \lambda_e^2 \overset{iid}{\sim} \exp(\frac{\lambda_e^2}{2}) \quad k = 1, \ldots, q$$
$$\lambda_c^2 \sim \mathrm{Gamma}\,(a_c,\ b_c)$$
$$\lambda_e^2 \sim \mathrm{Gamma}\,(a_e,\ b_e)$$
$$\sigma^2 \propto \frac{1}{\sigma^2}$$

D.3.2 Gibbs Sampler

Denote $\mu_{(-\alpha)} = \mathrm{E}(Y) - E\alpha$, then $\alpha|\text{rest} \sim \mathrm{N}(\mu_\alpha,\ \Sigma_\alpha)$, where

$$\mu_\alpha = \Sigma_\alpha (\frac{1}{\sigma^2}(Y - \mu_{(-\alpha)})^\top E)^\top,$$
$$\Sigma_\alpha = \left(\frac{1}{\sigma^2} E^\top E + \Sigma_{\alpha 0}^{-1}\right)^{-1}.$$

Denote $\mu_{(-\gamma)} = \mathrm{E}(Y) - C\gamma$, then $\gamma|\text{rest} \sim \mathrm{N}(\mu_\gamma,\ \Sigma_\gamma)$, where

$$\mu_\gamma = \Sigma_\gamma (\frac{1}{\sigma^2}(Y - \mu_{(-\gamma)})^\top C)^\top,$$
$$\Sigma_\gamma = \left(\frac{1}{\sigma^2} C^\top C + \Sigma_{\gamma 0}^{-1}\right)^{-1}.$$

Denote $\mu_{(-\beta_j)} = \mathrm{E}(Y) - X_j\beta_j$, then $\beta_j|\text{rest} \sim \mathrm{N}(\mu_{\beta_j},\, \sigma^2\Sigma_{\beta_j})$, where

$$\begin{aligned}\mu_{\beta_j} &= \Sigma_{\beta_j} X_j^\top (Y - \mu_{(-\beta_j)}),\\ \Sigma_{\beta_j} &= \left(X_j^\top X_j + \frac{1}{\tau_c^2}\right)^{-1}.\end{aligned}$$

Denote $\mu_{(-\eta_{jk})} = \mathrm{E}(Y) - \tilde{W}_k\eta_{jk}$, then $\eta_{jk}|\text{rest} \sim \mathrm{N}(\mu_{\eta_{jk}},\, \sigma^2\Sigma_{\eta_{jk}})$, where

$$\begin{aligned}\mu_{\eta_{jk}} &= \Sigma_{\eta_{jk}} \tilde{W}_k^\top (Y - \mu_{(-\eta_{jk})}),\\ \Sigma_{\eta_{jk}} &= \left(\tilde{W}_k^\top \tilde{W}_k + \frac{1}{\tau_{ek}^2}\right)^{-1}.\end{aligned}$$

The posterior of τ_c^2 is:

$$\frac{1}{\tau_c^2}|\text{rest} \sim \text{Inverse-Gaussian}(\sqrt{\frac{\sigma^2}{\beta_j^2}\lambda_c^2}, \lambda_c^2).$$

The posterior of τ_{ek}^2 is:

$$\frac{1}{\tau_{ek}^2}|\text{rest} \sim \text{Inverse-Gaussian}(\sqrt{\frac{\sigma^2}{\eta_{jk}^2}\lambda_e^2}, \lambda_e^2).$$

λ_c^2 and λ_e^2 have Gamma posterior distributions:

$$\begin{aligned}\lambda_c^2|\text{rest} &\sim \text{Gamma}(a_c + 1,\ \frac{\tau_c^2}{2} + b_c),\\ \lambda_e^2|\text{rest} &\sim \text{Gamma}(a_e + q,\ \sum_{k=1}^{q} \frac{\tau_{ek}^2}{2} + b_e).\end{aligned}$$

$\sigma^2 \sim \text{Inverse-Gamma}(\mu_{\sigma^2},\ \Sigma_{\sigma^2})$, where

$$\begin{aligned}\mu_{\sigma^2} &= \frac{n + 1 + q}{2},\\ \Sigma_{\sigma^2} &= \frac{(Y-\mu)^\top(Y-\mu) + (\tau_c^2)^{-1}\beta_j^2 + \sum_{k=1}^{q}(\tau_{ek}^2)^{-1}\eta_j^\top\eta_j}{2}.\end{aligned}$$