

Bayesian regularized quantile mixed models for longitudinal studies

by

Kun Fan

B.S., University of Illinois at Urbana-Champaign, 2015

M.A., Columbia University in the City of New York, 2017

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Abstract

In longitudinal studies, the same subjects are measured repeatedly over time, leading to correlations among the repeated measurements. Properly accounting for the intra-cluster correlations in the presence of data heterogeneity and long tailed distributions of the disease phenotype is challenging, especially in the context of high dimensional regressions. Here, we aim at developing novel Bayesian regularized quantile mixed effect models to tackle these challenges.

In the first project, we have proposed a Bayesian variable selection method in the mixed effect models for longitudinal lipidomics studies. To dissect important lipid-environment interactions, our model can simultaneously identify important main and interaction effects on the individual and group level, which have been facilitated by imposing the spike-and-slab priors through Laplacian shrinkage in the Bayesian quantile hierarchical models. The within- subject dependence among data can be accommodated by incorporating the random effects. The Gibbs sampler has been developed along with the Markov Chain Monte Carlo (MCMC). We have established the advantage of the proposed method over multiple competing methods in extensive simulation studies and a high-dimensional lipidomics study with repeated measurements.

In the second project, we further extend the sparse Bayesian quantile mixed models to nonlinear longitudinal interactions. Specifically, the proposed Bayesian quantile semiparametric model is robust not only to outliers and heavy-tailed distributions of the response variable, but also to the misspecification of interaction effect in the forms other than nonlinear interactions. We have developed the Gibbs sampler with the spike-and-slab priors to promote sparse identification of appropriate forms of main and interaction effects. Simulation results reveal superior performance in identification, estimation and statistical inference. In particular, the proposed method that incorporates the spike-and-slab priors can enable

exact statistical inference by yielding Bayesian credible intervals with nominal coverage probabilities on parametric and nonparametric fixed effects simultaneously. Application of the proposed method on high dimensional longitudinal biomedical studies shed novel insight on disease etiology.

The Bayesian regularized quantile mixed models proposed in this dissertation aim to tackle challenges arising from longitudinal gene environment interaction studies under the linear and nonlinear interaction assumption in chapter 2 and chapter 3, respectively. In a regression analysis framework, gene-environment interactions can be divided into linear and non-linear types, based on whether the effect of genetic factors on disease traits can be represented by linear or non-linear functions of these genetic factors. The two models proposed in this dissertation fill a significant technical gap since robust identification and inference of the two types of interactions has rarely been explored in published longitudinal studies.

We have also developed C++ based R packages `mixedBayes` on CRAN to facilitate reproducible and fast computations using all methods under comparison in this dissertation.

Bayesian regularized quantile mixed models for longitudinal studies

by

Kun Fan

B.S., University of Illinois at Urbana-Champaign, 2015

M.A., Columbia University in the City of New York, 2017

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Statistics
College of Arts and Sciences

KANSAS STATE UNIVERSITY
Manhattan, Kansas

2024

Approved by:

Major Professor
Cen Wu

Copyright

© Kun Fan 2024.

Abstract

In longitudinal studies, the same subjects are measured repeatedly over time, leading to correlations among the repeated measurements. Properly accounting for the intra-cluster correlations in the presence of data heterogeneity and long tailed distributions of the disease phenotype is challenging, especially in the context of high dimensional regressions. Here, we aim at developing novel Bayesian regularized quantile mixed effect models to tackle these challenges.

In the first project, we have proposed a Bayesian variable selection method in the mixed effect models for longitudinal lipidomics studies. To dissect important lipid-environment interactions, our model can simultaneously identify important main and interaction effects on the individual and group level, which have been facilitated by imposing the spike-and-slab priors through Laplacian shrinkage in the Bayesian quantile hierarchical models. The within- subject dependence among data can be accommodated by incorporating the random effects. The Gibbs sampler has been developed along with the Markov Chain Monte Carlo (MCMC). We have established the advantage of the proposed method over multiple competing methods in extensive simulation studies and a high-dimensional lipidomics study with repeated measurements.

In the second project, we further extend the sparse Bayesian quantile mixed models to nonlinear longitudinal interactions. Specifically, the proposed Bayesian quantile semiparametric model is robust not only to outliers and heavy-tailed distributions of the response variable, but also to the misspecification of interaction effect in the forms other than nonlinear interactions. We have developed the Gibbs sampler with the spike-and-slab priors to promote sparse identification of appropriate forms of main and interaction effects. Simulation results reveal superior performance in identification, estimation and statistical inference. In particular, the proposed method that incorporates the spike-and-slab priors can enable

exact statistical inference by yielding Bayesian credible intervals with nominal coverage probabilities on parametric and nonparametric fixed effects simultaneously. Application of the proposed method on high dimensional longitudinal biomedical studies shed novel insight on disease etiology.

The Bayesian regularized quantile mixed models proposed in this dissertation aim to tackle challenges arising from longitudinal gene environment interaction studies under the linear and nonlinear interaction assumption in chapter 2 and chapter 3, respectively. In a regression analysis framework, gene-environment interactions can be divided into linear and non-linear types, based on whether the effect of genetic factors on disease traits can be represented by linear or non-linear functions of these genetic factors. The two models proposed in this dissertation fill a significant technical gap since robust identification and inference of the two types of interactions has rarely been explored in published longitudinal studies.

We have also developed C++ based R packages `mixedBayes` on CRAN to facilitate reproducible and fast computations using all methods under comparison in this dissertation.

Table of Contents

List of Figures	x
List of Tables	xii
Acknowledgements	xviii
1 Introduction	1
1.1 Penalized Variable Selection	2
1.2 Bayesian Variable Selection	5
1.3 Longitudinal Data with Linear Mixed Effects Model	8
1.3.1 Random Intercept Model	9
1.3.2 Random Intercept and Slope Model	10
1.4 Works in this dissertation	11
2 Robust Sparse Bayesian Regression for Longitudinal Gene-Environment Interac- tions	13
2.1 Introduction	13
2.2 Materials and Methods	18
2.2.1 Data and Model Settings	18
2.2.2 The robust Bayesian interaction analysis with longitudinal traits . . .	19
2.2.3 Robust Bayesian variable selection for longitudinal interaction studies	21
2.2.4 Gibbs Sampler	26
2.3 Simulation	31
2.4 Real Data Analysis	40

2.5	Discussion	43
3	Bayesian Regularized Semiparametric Quantile Mixed Models for Longitudinal Studies.	45
3.1	Introduction	45
3.2	Data and Model settings	48
3.2.1	Partially Linear Varying Coefficient Model	48
3.2.2	The reparameterization on random effect	49
3.2.3	Basis expansion for structure identification of additive and dominant effects	50
3.2.4	Basis expansion for varying intercept and clinical effect	51
3.2.5	The quantile mixed effects varying coefficient model	52
3.2.6	The Bayesian formulation of the quantile varying coefficient model with longitudinal traits	52
3.2.7	The Bayesian regularized quantile mixed effects model for longitudinal study	54
3.2.8	Gibbs Sampler	61
3.3	Summary of all methods	69
3.4	Simulation	69
3.5	Real Data Analysis	77
	Bibliography	90
A	Appendices for chapter 2	100
A.1	Summary of all models under comparison	100
A.2	Additional simulation results	102
A.2.1	Identification results for $\theta = 0.5$ in simulation	102
A.2.2	Identification results for $\theta = 0.3$ in simulation	108
A.2.3	Identification results for $\theta = 0.7$ in simulation	116

A.2.4	Figures corresponding to Tables of identification and estimation results	124
A.2.5	Computational cost	129
A.2.6	Hyper-parameters sensitivity analysis	130
A.2.7	Credibility of identification results	133
A.2.8	Comparison between RBVSS-BI and interep	134
A.3	Assessment of the convergence of MCMC chains	138
A.3.1	The convergence of MCMC chains under random intercept - slope model	138
A.3.2	The convergence of MCMC chains under random intercept model . .	139
A.4	Estimation results for data analysis	140
A.5	Posterior inference	143
A.5.1	RBVSS-BI	143
A.5.2	RBVSS	148
A.5.3	RBV-BI	153
A.5.4	RBV	157
A.5.5	BVSS-BI	161
A.5.6	BVSS	166
A.5.7	BV-BI	171
A.5.8	BV	175
B	Appendices for chapter 3	179
B.1	Additional simulation results	179
B.1.1	Identification results for $\theta = 0.5$ in simulation	179
B.1.2	Identification results for $\theta = 0.3$ in simulation	182
B.1.3	Identification results for $\theta = 0.7$ in simulation	186
B.1.4	The estimated quantile varying coefficient functions	192
B.2	Posterior inference	193
B.2.1	RBVSS-SI	193
B.2.2	RBVSS	198

B.2.3	RBV-SI	202
B.2.4	RBV	205
B.2.5	BVSS-SI	208
B.2.6	BVSS	213
B.2.7	BV-SI	216
B.2.8	BV	220

List of Figures

2.1	(a) The trajectory plot for body weights of 60 CD-1 mice, (b) The trajectory plot for mean body weight of 60 CD-1 mice , (c) Bar plot for the mean body weight of 60 CD-1 mice at begin (week 1) and end (week 11)stage , (d) Histogram of body weights for 60 CD-1 mice.	16
3.1	Distribution of the FEV1 (left) and non-linear interaction effect of SNP rs10167668 (right) from the CAMP data. The blue dashed lines denote the 95% credible interval.	46
3.2	Time-varying total heritability	81
3.3	Real data analysis for the proposed method.	87
3.4	Real data analysis for the proposed method.	88
3.5	Real data analysis for the proposed method.	89
A.1	Identification results for $n = 200, m = 30, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 30th, 50th and 70th quantile levels.	125
A.2	Identification results for $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 30th, 50th and 70th quantile levels.	126
A.3	Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept - slope model based on 100 replicates at the 30th, 50th and 70th quantile levels.	127
A.4	Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept - slope model based on 100 replicates at the 30th, 50th and 70th quantile levels.	128
A.5	Real data analysis for the proposed method.	138

A.6	Real data analysis for the proposed method.	139
B.1	Estimation of non-zero varying coefficients under the $t(2)$ error (Error 2) for the proposed method (RBVSS-SI) at 50% quantile level. Red line: true varying coefficients. Black line: posterior median estimates of varying coefficients from RBVSS-SI. Blue lines: 95% credible intervals for the estimated varying coefficients.	192

List of Tables

2.1	Identification results for $n = 200, m = 30, p = 3$. Mean (sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 50th quantile.	36
2.2	Identification results for $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 50th quantile.	37
2.3	Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept - slope model based on 100 replicates at the 50th quantile.	38
2.4	Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept - slope model based on 100 replicates at 50th quantile.	39
2.5	Real data analysis result from RBVSS-BI under random intercept - slope model	43
2.6	Real data analysis result from RBVSS-BI under random intercept model . .	43
3.1	Summary of comparisons among all Bayesian regularized mixed-effects models.	69
3.2	Simulation results in example 1 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at 50th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.	75
3.3	Estimation accuracy in example 1 for $n = 400, m = 200$ under random intercept - slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 50th quantile with i.i.d. errors. .	76
3.4	Identification results for varying and constant effects for the additive effects.	82
3.5	Identification results for varying and constant effects for dominant effects. . .	83

3.6	Overlaps in CAMP.	84
3.7	Heritability of single additive effects SNPs by RBVSS-SI	85
3.8	Heritability of single dominant effects SNPs by RBVSS-SI	86
A.1	Summary of the proposed and alternative methods.	101
A.2	Summary of comparisons among all Bayesian regularized mixed-effects models.	101
A.3	Identification results for $n = 200, m = 30, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 50th quantile.	104
A.4	Identification results for $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 50th quantile.	105
A.5	Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept model based on 100 replicates at the 50th quantile.	106
A.6	Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept model based on 100 replicates at the 50th quantile.	107
A.7	Identification results for $n = 200, m = 30, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 30th quantile.	108
A.8	Identification results for $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 30th quantile.	109
A.9	Identification results for $n = 200, m = 30, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 30th quantile.	110
A.10	Identification results for $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 30th quantile.	111

A.11 Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept - slope model based on 100 replicates at the 30th quantile.	112
A.12 Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept - slope model based on 100 replicates at the 30th quantile.	113
A.13 Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept model based on 100 replicates at the 30th quantile.	114
A.14 Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept model based on 100 replicates at the 30th quantile.	115
A.15 Identification results for $n = 200, m = 30, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 70th quantile.	116
A.16 Identification results for $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 70th quantile.	117
A.17 Identification results for $n = 200, m = 30, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 70th quantile.	118
A.18 Identification results for $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 70th quantile.	119
A.19 Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept - slope model based on 100 replicates at the 70th quantile.	120
A.20 Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept - slope model based on 100 replicates at the 70th quantile.	121
A.21 Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept model based on 100 replicates at the 70th quantile.	122

A.22	Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept model based on 100 replicates at the 70th quantile.	123
A.23	Correspondence between Figures 2–5 and associated Tables with $n=200$ and $p=3$. m is the number of lipids and θ is quantile level.	124
A.24	Computational time for RBVSS-BI with $k = 5$ repeated measurements under random intercept - slope model (Model I) and random intercept model (Model II) under $t(2)$ error (Error 2) at the 50th quantile. m : number of lipids. Time: CPU time (in minutes) for 10,000 MCMC iterations. The total number of regression coefficients to be estimated is $pm + m + 1$, where p is the number of environmental factors. In this study, $p = 3$	129
A.25	Sensitivity analysis of RBVSS - BI under random intercept - slope model (Model I) and random intercept model (Model II) with $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP), false positives (FP) and total mean absolute deviation (TMAD) based on 100 replicates.	131
A.26	Sensitivity analysis of RBVSS - BI under random intercept - slope model (Model I) and random intercept model (Model II) with $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP), false positives (FP) and total mean absolute deviation (TMAD) based on 100 replicates.	132
A.27	Credibility of identification results with $n = 200, m = 60, p = 3$ for RBVSS - BI by adopting both MPM and 95% CI. Mean (sd) of true positives (TP) and false positives (FP) based on 100 replicates.	133
A.28	Identification performance of RBVSS-BI and interep under random intercept - slope model with $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) based on 100 replicates.	136
A.29	Estimation accuracy of RBVSS-BI and interep under random intercept - slope model with $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) based on 100 replicates.	137

A.30	Real data analysis result from RBVSS under random intercept - slope model	140
A.31	Real data analysis result from RBVSS under model random intercept model	140
A.32	Real data analysis result from BVSS-BI under random intercept - slope model	141
A.33	Real data analysis result from BVSS-BI under random intercept model . . .	141
A.34	Real data analysis result from BVSS under random intercept - slope model .	142
A.35	Real data analysis result from BVSS under random intercept model	142
B.1	Simulation results in example 2 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at 50th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.	180
B.2	Estimation accuracy in example 2 for $n = 400, m = 200$ under random intercept - slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 50th quantile with i.i.d. errors. .	181
B.3	Simulation results in example 1 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at 30th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.	182
B.4	Estimation accuracy in example 1 for $n = 400, m = 200$ under random intercept - slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 30th quantile with i.i.d. errors. .	183
B.5	Simulation results in example 2 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at 30th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.	184
B.6	Estimation accuracy in example 2 for $n = 400, m = 200$ under random intercept-slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 30th quantile with i.i.d. errors. .	185

B.7	Simulation results in example 1 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at 70th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.	186
B.8	Estimation accuracy in example 1 for $n = 400, m = 200$ under random intercept - slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 70th quantile with i.i.d. errors. .	187
B.9	Simulation results in example 2 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept -slope model based on 100 replicates at 70th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.	188
B.10	Estimation accuracy in example 2 for $n = 400, m = 200$ under random intercept-slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 70th quantile with i.i.d. errors. .	189
B.11	Empirical 95% coverage probabilities under simulated SNP data and random intercept - slope model with i.i.d. $t(2)$ error based on 200 replicates.	190
B.12	Empirical 95% coverage probabilities under simulated SNP data and random intercept - slope model with i.i.d. $N(0,1)$ error based on 200 replicates.	191

Acknowledgments

I would like to begin by expressing my heartfelt appreciation to my advisor, Dr. Cen Wu, for his invaluable guidance and support throughout my Ph.D. journey. Dr. Wu has been an inspiring mentor, full of innovative ideas. Our conversations have provided me with countless insights that significantly shaped my research. His deep knowledge, expertise, and passion for research have set a standard that I aim to follow in my future endeavors.

I am also profoundly grateful to my Ph.D. committee members, Dr. Wexing Song, Dr. Juan Du, and Dr. Weiqun Wang, for their unwavering support, thoughtful advice, and insightful feedback.

My sincere thanks go to the Department of Statistics and Dr. Lynn Ying-Shiang Lin for granting me assistantships and scholarships, which provided essential financial support during my studies. I am also thankful to the department's professors for delivering outstanding courses, and to my friends for their continuous help and encouragement.

Finally, I owe my deepest gratitude to my family, particularly my parents, for their unconditional love and support throughout this journey.

Chapter 1

Introduction

Nowadays, high-dimensional data analysis plays a crucial role in many areas, including biomedical studies. For instance, in the multiomics study, gene-environment interactions shed fresh insight in elucidating the genetic metabolism and phenotype variations. Previously, the ordinary least square regression (OLS) can be used to predict disease phenotype with low dimensional predictors. However, due to the nature of the genetics data, “large p , small n ”, which means the sample size is much smaller than the number of genetic features, challenges statistician to develop novel statistical methods to handle the high dimensionality of the high-throughput data. Since only a part of genetic features is related to the phenotype of interest, the identification of important features becomes an important task. To address this issue, a novel variable selection method needs to be accommodated. One commonly used variable selection method is the least absolute shrinkage and selection operator (LASSO) ([Tibshirani \(1996\)](#)). LASSO can be used to select important variables by shrinking all regression coefficients towards 0. With LASSO, for the high-dimensional data analysis, the shrinkage estimation and variable selection can be done simultaneously so that the set of predictors will be dramatically reduced, leading to a more accurate outcome prediction and model interpretation. In the past two decades, extensive amount of variable selection methods have been developed for analyzing high-dimensional data as well. These variable selection methods may have advantages over LASSO under certain situations. In the remaining sec-

tions, a detailed overview of those variable selection methods for high-dimensional data will be provided from both frequentist and Bayesian perspectives.

1.1 Penalized Variable Selection

To overcome the limitation of ordinary least square regression, penalized regression becomes a widely adopted method to select important genetics features in genomics studies. Consider Y as a vector of the disease outcome, which can be a continuous disease phenotype. Denote G as the m -dimensional matrix of genetics factors, the penalized regression model can be formulated as :

$$L(\beta; Y, G) + \sum_{j=1}^m p_{\lambda}(|\beta_j|).$$

where $\beta = (\beta_1, \dots, \beta_m)^T$ is the coefficient vector, $L(\beta; Y, G)$ represents the loss function and $p_{\lambda}(|\beta_j|)$ is the penalty function of β_j with a tuning parameter λ . The tuning parameter λ determines the amount of shrinkage imposed. Larger λ value will lead to a decreasing model complexity with less variable being selected. On the other hand, the model becomes more complicated when more variable have been selected. Finding the optimal tuning parameter is crucial for variable selection. In order to determine the best tuning parameter λ , cross validation method can be adopted. The estimation of the regression coefficients and variable selection can be conducted by minimizing the penalized loss function(1.1). LASSO is a special case of penalized least square regression with penalty function $p_{\lambda}(|\beta_j|) = \lambda|\beta_j|$. In this case, the penalized loss function can be expressed as below

$$||Y - G\beta||_2^2 + \lambda \sum_{j=1}^m |\beta_j|.$$

where $\beta = (\beta_1, \dots, \beta_m)^T$ is the coefficient vector and G is the m -dimensional matrix of genetics factors. Although LASSO is a popular variable selection method for high-dimensional data analysis, it does not meet the one of the three properties of a good penalty function(Fan and Li (2001)). LASSO only satisfies the sparsity and continuity criteria, however, it introduces

the biased estimation which disobeys the unbiasedness criteria. In this situation, new penalties need to be developed to resolve the biased estimation issue. In the past two decades, the smoothly clipped absolute deviation (SCAD)([Fan and Li \(2001\)](#)), minimax concave penalty (MCP) ([Zhang \(2010\)](#)) and adaptive LASSO ([Zou \(2006\)](#)) have been developed to achieve all three criterias. Their penalties can be represented as following: respectively,

SCAD:

$$p_{\lambda}(\beta_j) = \begin{cases} \lambda|\beta_j| & \text{if } |\beta_j| \leq \lambda, \\ -\frac{\beta_j^2 - 2a\lambda|\beta_j| + \lambda^2}{2(a-1)} & \text{if } \lambda < |\beta_j| \leq a\lambda, \\ \frac{1}{2}(a+1)\lambda^2 & \text{if } |\beta_j| > a\lambda, \end{cases}$$

where $a > 2$.

MCP:

$$p_{\lambda}(\beta_j) = \begin{cases} \lambda|\beta_j| - (2a)^{-1}\beta_j^2 & \text{if } |\beta_j| \leq a\lambda, \\ \frac{1}{2}a\lambda^2 & \text{if } |\beta_j| > a\lambda, \end{cases}$$

where $a > 1$.

Adaptive LASSO:

$$p_{\lambda}(\beta_j) = \lambda(|\beta_j^{(0)}|^{-a})|\beta_j|,$$

where $a > 0$ and $\beta_j^{(0)}$ is the initial estimation of the j th element in the coefficient vector $\hat{\beta}$. Besides, handling complex data structure such as the group structure has motivated new penalty functions. To select grouped variables, group LASSO method has been developed by [Yuan and Lin \(2006a\)](#) with the penalty function as :

$$p_{\lambda}(\beta_j) = \lambda \sum_{j=1}^J \sqrt{m_j} \|\beta_j\|_2,$$

where $\beta_j = (\beta_{j1}, \dots, \beta_{jm_j})^T$ is a coefficient vector of length m_j and $\|\beta_j\|_2 = \sqrt{\beta_j^T \beta_j}$. The group size in the penalty function can be adjusted.

In addition, the SCAD and MCP have also been extended to the group setting. For example,

the penalty function of the group SCAD(Huang et al. (2012))can be defined as :

$$p_{\lambda}(\beta_j) = \begin{cases} \sqrt{m_j}\lambda\|\beta_j\|_2 & \text{if } \|\beta_j\|_2 \leq \sqrt{m_j}\lambda, \\ -\frac{\beta_j^T \beta_j - 2\sqrt{m_j}a\lambda\|\beta_j\|_2 + m_j\lambda^2}{2(a-1)} & \text{if } \sqrt{m_j}\lambda < \|\beta_j\|_2 \leq \sqrt{m_j}a\lambda, \\ \frac{\sqrt{m_j}}{2}(a+1)\lambda^2 & \text{if } \|\beta_j\|_2 > \sqrt{m_j}a\lambda, \end{cases}$$

where $a > 2$.

and the group MCP(Huang et al. (2012)) can be expressed as :

$$p_{\lambda}(\beta_j) = \begin{cases} \sqrt{m_j}\lambda\|\beta_j\|_2 - (2a)^{-1}\beta_j^T \beta_j & \text{if } \|\beta_j\|_2 \leq \sqrt{m_j}a\lambda, \\ \frac{m_j}{2}a\lambda^2 & \text{if } \|\beta_j\|_2 > \sqrt{m_j}a\lambda, \end{cases}$$

where $a > 1$.

When encountering with heavy-tailed errors and outliers in response variables, traditional variable selection methods are not robust to accomodate data heterogeneity. Novel robust variable selection methods need to be developed to handle data contamination and avoid model misspecification. Other than the penalized least square regression, the penalized robust regression, in the form of, “robust loss function + penalty function”, can be considered to deal with the robustness of the data. Several loss functions have been adopted to deal with the robustness, for instance, check loss function and the least absolute deviation (LAD) function(Wu and Ma (2015a)). The LAD loss fuction is a special case of the quantile-based loss functions. Since the form of the θ th ($0 < \theta < 1$) linear quantile regression (Li et al. (2010a))can be expressed as:

$$y_i = g_i^T \beta + u_i, i = 1, \dots, n,$$

where $\beta = (\beta_1, \dots, \beta_m)^T$ and u_i 's are independent with their θ th quantiles equal to 0. The estimation of the coefficients vector β can be performed by minimizing the following equation:

$$\min_{\beta} \sum_{i=1}^n \rho_{\theta}(y_i - g_i^T \beta).$$

where $\rho_{\theta}(\cdot)$ is the check loss function

$$\rho_{\theta}(t) = \begin{cases} \theta t & \text{if } t \geq 0, \\ -(1 - \theta)t & \text{if } t < 0, \end{cases}$$

and g_i^T is the i th row of the m -dimensional matrix G . When $\theta = 0.5$, the quantile based loss function becomes LAD loss function:

$$\sum_{i=1}^n |y_i - g_i^T \beta|.$$

A major limitation of robust methods when analyzing the high-dimesional data is the high computational cost, due to the nonsmoothness of the loss functions. To accomodate the computational convenience, the LAD loss fuction is commonly adopted for omics studies (Wu and Ma (2015a);Huang et al. (2011);Wu et al. (2018a)).

Similar to the penalized least square regression, the LAD penalized regression can be expressed as

$$\sum_{i=1}^n |y_i - g_i^T \beta| + \sum_{j=1}^m p_{\lambda}(|\beta_j|).$$

1.2 Bayesian Variable Selection

In addition to the frequentist methods for variable selection, Bayesian variable selection has also been widely used method to handle high-dimensional data. Traditionally, the frequentist methods utilize the minimizer of the “unregularized loss+ penalty” function to obtain the optimal estimated coefficients. From the Bayesian perspective, the coefficient estimation

can be achieved by computing the posterior mode through assigning certain prior onto the regression parameters. The LASSO estimates can be achieved by computing the posterior mode from imposing independent and identical Laplace priors on regression coefficients, i.e. $\pi(\beta) = \prod_{j=1}^m \frac{\lambda}{2} \exp\{-\lambda|\beta_j|\}$ (Tibshirani (1996)). Specially, with the independent and noninformative scale-invariant marginal prior $\pi(\sigma^2) = \frac{1}{\sigma^2}$ (Park and Casella (2008a)) on σ^2 , the posterior distribution can be represented as :

$$\pi(\beta, \sigma^2 | y) \propto (\sigma^2)^{-\frac{n}{2}-1} \exp\left\{-\frac{1}{2\sigma^2} \|Y - G\beta\|_2^2 - \lambda \sum_{j=1}^m |\beta_j|\right\}. \quad (1.1)$$

The fully Bayesian LASSO has been developed (Park and Casella (2008a),) i.e. the Bayesian LASSO, through replacing the Laplace prior with the conditional Laplace prior

$$\pi(\beta_j | \sigma^2) = \frac{\lambda}{2\sqrt{\sigma^2}} e^{-\lambda|\beta_j|/\sqrt{\sigma^2}}.$$

The importance of conditioning on σ^2 is that a unimodal posterior can be guaranteed. Otherwise, the posterior may not be unimodal (Park and Casella (2008a)). One major limitation of Bayesian LASSO is that it cannot shrink regression coefficient to zero exactly. However, the beauty of Bayesian LASSO is to quantify the uncertainty in estimation and readily make statistical inference on each coefficient. In Bayesian framework, selecting grouped predictors is very important. As an extension of Bayesian LASSO, Kyung et. al (Kyung et al. (2010)) has developed the Bayesian group LASSO to handle grouping structure. The prior of the Bayesian group LASSO can be expressed by adopting a multivariate Laplace prior

$$\pi(\beta_j | \sigma^2) \propto \exp\left\{-\frac{\sqrt{m_j}\lambda}{\sigma} \|\beta_j\|_2\right\} \quad (1.2)$$

where β_j is a coefficient vector with size m_j and $\frac{\sqrt{m_j}\lambda}{\sigma}$ is the scale parameter for the multi-

variate Laplace distribution.

The major drawback of the Bayesian LASSO is that it cannot shrink each estimated coefficient to zero exactly, i. e. lack of model selection property. The spike and slab priors (Mitchell and Beauchamp (1988)) can achieve the model selection property throughout imposing the spike and slab prior on each regression coefficient. The prior for β_j becomes:

$$\beta_j | \phi_j \sim (1 - \phi_j) \pi_0(\beta_j) + \phi_j \pi_1(\beta_j), \quad j = 1, \dots, m,$$

where $\pi_1(\beta_j)$ is the flat “slab distribution” for nonzero large coefficients and $\pi_0(\beta_j)$ is the “spike distribution” for either zero or reasonably small values close to zero, and ϕ_j is an auxiliary indicator variable. When $\phi_j = 1$, β_j comes from the slab proportion that represents the existence of the j th effect. When $\phi_j = 0$, the j th effect is not associated with y . Most commonly, by fixing ϕ_j at $\frac{1}{2}$, it will provide equal probabilities to all submodels and represents no prior information on the true model (Xu and Ghosh (2015)). For the Bayesian LASSO with spike and slab prior, $\pi_1(\beta_j)$ is set as a normal distribution with large variance and $\pi_0(\beta_j)$ is set to $\delta_0(\beta_j)$, which denotes a point mass at 0. Therefore, the unimportant coefficients can be set to zero in the spike part. The prior can be expressed as

$$\beta_j | \phi_j, \sigma^2, \tau_j^2 \sim (1 - \phi_j) \delta_0(\beta_j) + \phi_j N(0, \sigma^2 \tau_j^2), \quad j = 1, \dots, m.$$

Similarly, for the Bayesian group LASSO with spike and slab prior, $\pi_1(\beta_j)$ is set to be a multivariate normal distribution with large variance and $\pi_0(\beta_j)$ is set to be $\delta_0(\beta_j)$, which denotes a point mass at $\mathbf{0}$. Thus, the groupwise unimportant coefficients, can be set to zero in the spike part. The prior can be expressed as:

$$\beta_j | \phi_j, \sigma^2, \tau_j^2 \sim (1 - \phi_j) \delta_0(\beta_j) + \phi_j N_{m_j}(0, \sigma^2 \tau_j^2 \mathbf{I}_{m_j}), \quad j = 1, \dots, m.$$

To accomodate the heavy-tailed distributed data, under Bayesian framework, Li et al. (Li et al. (2010a)) has proposed the Bayesian quantile LASSO by imposing the Laplace prior $\pi(\beta|\tau, \lambda) = (\frac{\tau\lambda}{2})^m \exp(-\tau\lambda \sum_{j=1}^m |\beta_j|)$ to induce shrinkage and assuming that the error term follows a skewed Laplace distribution. The posterior distribution of β can be expressed as:

$$\pi(\beta|\tau, \lambda) \propto \exp\{-\tau \sum_{i=1}^n \rho_{\theta}(y_i - g_i^T \beta) - \tau\lambda \sum_{j=1}^m |\beta_j|\}. \quad (1.3)$$

Li et al. (2010a) has proposed the Bayesian quantile group LASSO regression by imposing the Laplace prior $\pi(\beta_j|\eta) = C_{m_j} \eta^{m_j} \exp(-\eta \|\beta_j\|_2)$ where $C_{m_j} = 2^{-(m_j+1)/2} (2\pi)^{-(m_j-1)/2} / \Gamma((m_j+1)/2)$, $\eta = \tau\lambda$, and $\Gamma(\cdot)$ is the gamma function. The posterior distribution of β_j can be represented as:

$$\pi(\beta|\tau, \lambda) \propto \exp\{-\tau \sum_{i=1}^n \rho_{\theta}(y_i - g_i^T \beta) - \tau\lambda \sum_{j=1}^m \|\beta_j\|_2\}. \quad (1.4)$$

Spike and slab prior can be incorporated on both Bayesian quantile LASSO and Bayesian quantile group LASSO.

1.3 Longitudinal Data with Linear Mixed Effects Model

In longitudinal studies, the same subjects are measured repeatedly across time. One prominent feature of the repeated measurement data is that they are intra-correlated within the same subject, while independence is assumed for measurements taken on different subjects. A diversity of statistical methods that can properly account for the intra-cluster correlations among longitudinal data have been proposed (Verbeke et al. (2014a), Bandyopadhyay et al. (2011a)). Among those methods, the generalized estimating equation (GEE) bypasses the challenge of specifying the full likelihood function especially under non-normal assumptions (Liang and Zeger (1986a)). Instead, it utilizes the first two marginal moments of the response, as well as a working correlation matrix involving only a small number of nuisance

parameters. Although the GEE estimator is consistent even when the working correlation is (partially) misspecified, it is not robust to outlying observations and data heterogeneity. In this dissertation, we consider developing the Bayesian quantile mixed effect models to robustly analyze the longitudinal data for the following reasons. First, the asymmetric Laplace distribution can be readily adopted to specify robust likelihood function in mixed effect models. Second, the random effects in the mixed effect model can be well characterized by a properly chosen prior and corresponding posterior distribution. In addition, with the fully Bayesian analysis, exact inference on model parameters can be interpreted even when sample size is limited. Below we briefly review two widely used mixed effect models.

1.3.1 Random Intercept Model

Consider the following model:

$$Y_{ij} = X_{ij}^T \beta + b_i + \epsilon_{ij}, i = 1, \dots, n, j = 1, \dots, k, N = nk,$$

where Y_{ij} is the j th measurement of a continuous random variable on the i th subject, X_{ij}^T are the row vectors of the design matrix with dimension $N \times p$ and β is a coefficient vector with length p and ϵ_{ij} represents the within-subject measurement error. The assumption of the fixed effects and the measurement error are random with mean zero and with variances $Var(b_i) = \sigma_b^2$ and $Var(\epsilon_{ij}) = \sigma^2$, respectively. In addition, the b_i and ϵ_{ij} are independent with each other. Then the marginal mean response over time can be expressed as

$$E(Y_{ij}) = X_{ij}^T \beta.$$

and marginal variance can be represented as

$$Var(Y_{ij}) = \sigma_b^2 + \sigma^2$$

Consider the marginal covariance between two response variables, Y_{ij} and Y_{ik} is given by

$$Cov(Y_{ij}, Y_{ik}) = \sigma_b^2$$

then, the marginal covariance matrix can be represented as

$$Cov(Y_i) = \begin{pmatrix} \sigma_b^2 + \sigma^2 & \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \sigma^2 & \sigma_b^2 & \dots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 + \sigma^2 & \dots & \sigma_b^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \sigma_b^2 & \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 + \sigma^2 \end{pmatrix}$$

and the correlation between two response variable is

$$Corr(Y_{ij}, Y_{ik}) = \frac{\sigma_b^2}{\sigma_b^2 + \sigma^2}$$

1.3.2 Random Intercept and Slope Model

Consider the following model:

$$Y_{ij} = X_{ij}^T \beta + Z_{ij}^T b_i + \epsilon_{ij}, i = 1, \dots, n, j = 1, \dots, k, N = nk,$$

where Y_{ij} is the j th measurement of a continuous random variable on the i th subject, X_{ij}^T are the row vectors of the design matrix with dimension $N \times p$ and β is a coefficient vector with length p and ϵ_{ij} represents the within-subject measurement error. $Z_{ij} = (1, j)^T$ is a 2×1 covariate associated with random effects and b_i is a 2×1 vector of random effects. The marginal mean response over time can be represented as

$$E(Y_{ij}) = X_{ij}^T \beta.$$

The marginal covariance of Y_i by averaging over the distribution of b_i is

$$Cov(Y_i) = Z_i G Z_i^T + \sigma^2 I_k$$

where $G = Cov(b_i)$, Y_i is a $k \times 1$ vector with each element Y_{ij} and Z_i is a $k \times 2$ matrix which has the form:

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ \vdots & \vdots \\ 1 & k \end{pmatrix}$$

Variable selection methods can also be applied on the linear mixed effects model for longitudinal study. For instance, Bayesian LASSO. We can assign the same prior on the coefficients parameter and other hyperparameters as the penalized linear regression model under the univariate case.

1.4 Works in this dissertation

In this dissertation, we have proposed sparse Bayesian quantile mixed models for interaction analysis in longitudinal studies. In a broader context, interaction structure in high-dimensional scenarios is a typical type of structured sparsity that also includes strong inter-relatedness ((Zou and Hastie, 2005; Tibshirani et al., 2005)), graphs and networks ((Wu et al., 2018b; Ren et al., 2017, 2019)), pathways and gene sets ((Wu and Cui, 2014; Jiang et al., 2017; Wu et al., 2012)) among others. From the regression perspective, gene-environment interactions can be categorized as linear interactions and non-linear interactions, depending on whether the contribution of genetic factors to disease phenotypes can be modeled as linear or nonlinear functions of genetic factors ((Zhou et al., 2021)). Besides structured sparsity, heavy-tailed distributions and outliers in disease phenotypes are widely encountered in genomics studies, which demands robust variable selection methods ((Wu and Ma, 2015b)).

In the first project, we have developed Bayesian regularized quantile mixed models for interaction structure arising from repeated measures in one-way ANOVA with high-dimensional lipidomics features. In a one-way ANOVA setting, we model the interaction between treatment (in terms of a group of binary indicators corresponding to treatment) and high-dimensional lipid features, which is essentially a linear interaction problem. Identification of important main and interaction effects can be achieved by utilizing structured Laplacian shrinkage and spike-and-slab priors within Bayesian hierarchical models. In the proposed quantile mixed effect models, we adopt the asymmetric Laplace distribution to specify the likelihood function under longitudinal traits, which has a nice correspondence to the frequentist quantile regression models resistant to outliers in the response variable. Besides, intra-cluster correlations from repeated measures are accommodated by random effects from the mixed model. Gibbs samplers of proposed and alternative methods have been developed to generate posterior samples for estimation and prediction. We demonstrated the superiority of the proposed method over alternatives in both simulation and real data analysis.

In the second project, we have proposed a sparse Bayesian quantile semiparametric model to capture time-varying non-linear interactions in longitudinal Genome-wide association studies (GWAS). In GWAS, nonlinear gene-environment interactions were first studied by (Wu and Cui (2013)) in a hypothesis-testing framework. Regularized varying coefficient models have also been developed for the analysis goal from both frequentist and Bayesian frameworks ((Wang et al., 2008; Wu et al., 2018c,d; Tang et al., 2013; Ren et al., 2020)). The proposed Bayesian model is robust to heterogeneous repeated measures and can properly handle correlations among longitudinal observations. Furthermore, it can conduct structure identification by detecting whether high-dimensional SNPs have nonlinear interactions, just main effects, or no genetic effects at all. Our proposed fully Bayesian analysis can also enable exact statistical inference by providing Bayesian credible intervals on parametric and non-parametric fixed effects at the same time. We demonstrated the advantage of the proposed method in extensive simulations and a case study using CAMP longitudinal GWAS data.

Chapter 2

Robust Sparse Bayesian Regression for Longitudinal Gene-Environment Interactions

2.1 Introduction

The importance of longitudinal omics data analysis has been increasingly recognized. Compared to cross-sectional omics studies, the repeated measurements of the disease phenotype and/or omics features have provided additional insight to elucidate the etiology of complex diseases. As the longitudinal measures have been taken over a series of time points with respect to the same subjects, they are intrinsically interconnected ([Verbeke et al. \(2014b\)](#), [Bandyopadhyay et al. \(2011b\)](#)). Therefore, it is critical to appropriately model the intra-correlations among the longitudinal observations in the presence of high-dimensional genetic and genomics data, and other types of omics data such as lipidomics, metabolomics and proteomics data. Meanwhile, published studies have revealed that the genetic factors (and omics features) alone cannot fully unravel the genetic underpinnings of complex diseases. Their interactions with the environmental factors, or gene-environment ($G \times E$) interactions, needs to be dissected on a fine scale to better characterize the genetic control of complex

disease traits (Hunter (2005); Kraft et al. (2007)). With the longitudinal genetic design, the interaction analysis leverages the repeated measurements of phenotypic traits to further decipher the dynamic landscape of $G \times E$ interactions.

In this study, we are interested in conducting a longitudinal $G \times E$ analysis, which has been partially motivated by a published work using repeated measure ANOVA to examine changes in the lipidomics profiles from weight controlled CD-1 mice (King et al. (2015)). The trait of interest is the repeatedly measured body weight of 60 female CD-1 mice, as shown in Figure 2.1 (a) 60 female CD-1 mice have been randomly assigned with 4 different treatment groups, which are control (ad libitum feeding and without treadmill exercise), AE (ad libitum feeding and with treadmill exercise), PE (pair feeding and with treadmill exercise), and DCR (20% dietary calorie restriction and without treadmill exercise). 20% dietary restriction means the mice are fed a diet with 20% less total calories from carbohydrates and fat in comparison with the basal AIN-93 diet, while the levels of protein and essential micro-nutrients have been kept the same as the basal diet. Figure 2.1 (b) demonstrates the trajectory plots of average weight across 11 weeks for each group. Figure 2.1 (c) shows that without dietary control and exercise, mice in the control group have the largest increase in body weight. A histogram of all the pooled measurements of bodyweight in Figure 2.1 (d) reveals outliers and deviation from normality. In particular, the bar on the rightmost position is due to an outlying weight in control group measured at the 5th week.

In addition, the dietary lipids of CD-1 mice have also been generated and treated as omics features in the $G \times E$ study. The triacylglycerol (TG) and diacylglycerol (DG) lipid species have been profiled from plasma samples using the electrospray ionization MS/MS (King et al. (2015)). We focus on the DG profiles and treat them as the G factors. The environmental factors adopted for the $G \times E$ analysis is the factor corresponding to the four treatment groups in ANOVA, which has been converted to a group of three binary indicators with control level as the baseline. The G factors are the high-dimensional lipidomics factors. Identifying important main lipid factors and their interactions with the environment factors will shed fresh insight in elucidating the lipid metabolism and phenotypic variations in a longitudinal setting.

In a broader context, the proposed $G \times E$ analysis stems from the ANOVA problem in the longitudinal design where selecting important main and interaction effects is of paramount interest (Yuan and Lin (2006b)). Analyzing such a longitudinal lipidomics data in an interaction study is challenging in the following aspects. First, as shown in Figure 2.1 (a) and 2.1 (d), the repeated measurements of body weight of the CD-1 mice have skewed distributions and outliers, requiring robustness of analysis methods (Zhou et al. (2021); Wu and Ma (2015b)). Second, accommodating intra-cluster correlations of longitudinal observations under the robustness constraints is not a trivial task. Many widely used longitudinal modeling methods, such as those based on the generalized estimating equation (GEE) (Liang and Zeger (1986b)), are not robust and vulnerable to outlying observations. Third, due to the formulation of the E factors in a group of binary indicators representing the four treatment of control, DCR, AE and PE, the $G \times E$ interactions have a group structure. Detecting important lipid by environment interaction rising from the repeated measure ANOVA eventually amounts to group level variable selection (Yuan and Lin (2006b)), and pinpointing the important lipid main effects leads to individual level selection, which together demand a tailored interaction analysis.

Despite success, published high-dimensional analysis of repeated measure data cannot tackle the above challenges simultaneously. For instance, penalization methods proposed within the major frameworks for longitudinal data, such as the penalized GEE (Wang et al. (2012)) and regularized linear mixed model (Fan and Li (2012)) among others, have focused on main effects and are vulnerable to heavy-tailed errors. Although robust sparse longitudinal methods have been developed, including (Fan et al. (2012a)) and (Lv et al. (2015)), they cannot be adopted for performing gene-environment interaction studies.

To date, high-dimensional longitudinal methods have been underdeveloped for interaction studies (Fan et al. (2012b)). Our limited literature mining suggests that, in the presence of the structured sparsity of high-dimensional data, there appears a tendency to concentrate more on feature selection and ignore serial correlations by assuming independence among repeated measurements, as shown by a plethora of work in non-/semi-parametric variable selection for longitudinal data (Wang et al. (2008); Tang et al. (2013); Ma et al. (2013)) with

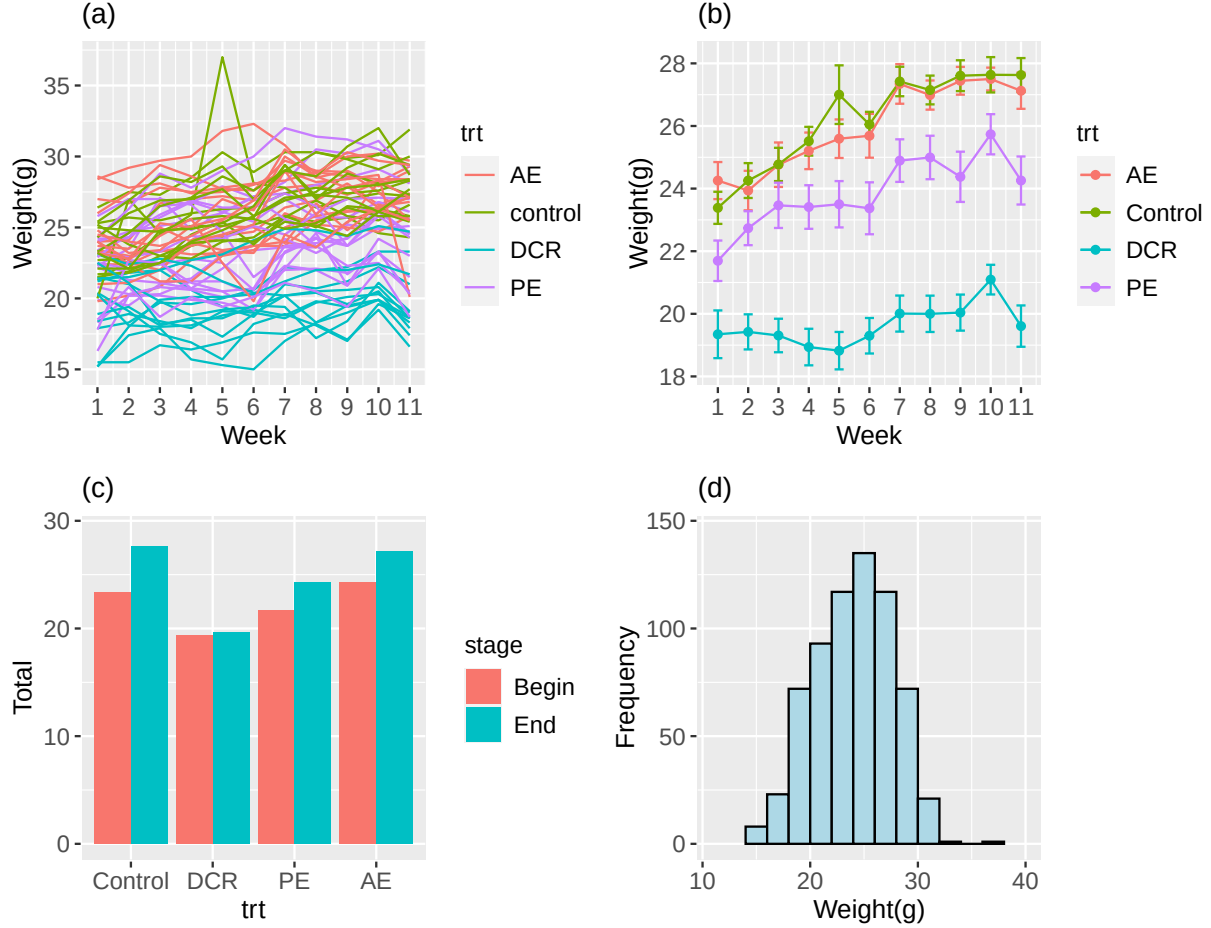


Figure 2.1: (a) The trajectory plot for body weights of 60 CD-1 mice, (b) The trajectory plot for mean body weight of 60 CD-1 mice, (c) Bar plot for the mean body weight of 60 CD-1 mice at begin (week 1) and end (week 11) stage, (d) Histogram of body weights for 60 CD-1 mice.

nonlinear interaction assumptions (Zhou et al. (2021)). The $G \times E$ analysis of the motivating lipidomics data models the gene-environment interactions as the product between the lipids and group of dummy variables representing the E variables, which corresponds to the more general linear gene-environment interactions (Zhou et al. (2021)). Only a limited number of work have been published with the linear $G \times E$ interaction assumption, including Zhou et al. 2019 and 2022 (Zhou et al. (2019, 2022a)). However, these studies adopts GEE and thus are not robust to outliers and heavy-tailed distributions in the repeated measurements.

Besides, both studies are focused on penalization without providing inference procedures for parameter estimation.

The challenges of analyzing the longitudinal lipidomics data and the limitations of existing methods have motivated us to consider a robust sparse Bayesian mixed model for $G \times E$ interactions with repeated measurements. Compared to the frequentist-based shrinkage methods, fully Bayesian analyses enables exact statistical inference even under small sample sizes. Specifically, to incorporate robustness, we propose to adopt the asymmetric Laplace distribution for setting up the robust likelihood function so the proposed model is not sensitive to outlying and heterogeneous repeated measurements (Yu and Moyeed (2001)). In addition, the structure of mixed effect model readily allows us to account for correlations among longitudinal measurements by specifying random effects. Therefore, accommodating robustness and correlations at the same time does not incur extra cost compared to studies robustifying the generalized estimating equations (Fan and Li (2012); Fan et al. (2012a)). Furthermore, within the Bayesian mixed model framework, we can impose structured sparse priors to select the main and interaction effects jointly to conduct a tailored $G \times E$ analysis for the lipidomics data.

With the high-dimensional genetic factors, selecting important main and interactions in interaction studies that root in the ANOVA problem are no longer corresponding to the group LASSO (Yuan and Lin (2006b)), but requires a more refined bi-level structure to select main effects on the individual level and interaction on the group level concurrently. In literature, Bayesian LASSO and Bayesian group LASSO have been widely adopted to conduct Bayesian shrinkage on the individual and group level, correspondingly (Casella et al. (2010)). A major disadvantage of not being able to yield exact 0 posterior estimates in fully Bayesian analysis based on these priors have been widely recognized. Here, we incorporate univariate and multivariate spike-and-slab priors in Laplacian shrinkage to promote exact sparsity (George and McCulloch (1993), Ishwaran and Rao (2005), Ročková and George (2018), Ren et al. (2020)). In our study, we assign individual and group level sparse priors to select main and interaction effects, correspondingly within the quantile mixed effect model framework. Although Bayesian LASSO and Bayesian group LASSO can be adopted to

achieve such an analysis goal (Casella et al. (2010)), they conduct Bayesian shrinkage based on the conditional (group) Laplace prior which cannot lead to exact 0 posterior estimates in fully Bayesian analysis. Through extensive simulation studies and the case study, we have demonstrated the superiority of the proposed model over multiple competing alternatives. For fast computation and reproducible research, we implement the proposed and alternative methods in C++ and encapsulate them in a publicly available R package `mixedBayes` (Fan and Wu).

2.2 Materials and Methods

2.2.1 Data and Model Settings

Consider a longitudinal study with n subjects with k repeated measurement for each subject. Let Y_{ij} be phenotype for the i th subject at the j th time point ($1 \leq i \leq n, 1 \leq j \leq k$). We use the m -dimensional vector $\mathbf{L}_{ij}$ to denote the measurement of lipid factors for the i th subject at time point j , where $\mathbf{L}_{ij} = (L_{ij1}, \dots, L_{ijm})^\top$. Also, let the p -dimensional vector $\mathbf{T}_{ij}$ denote the measurements of factors, where $\mathbf{T}_{ij} = (T_{ij1}, \dots, T_{ijp})^\top$. In a typical repeated measure ANOVA analysis, the factors do not vary across time, which has been incorporated as a special case under this formulation. To model the main effects of lipid and treatment, as well as their interactions that are collectively associated with the longitudinal phenotype having heavy-tailed distributions, we propose the following quantile mixed-effects model at a given quantile level θ , ($0 < \theta < 1$):

$$Y_{ij} = \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_{0,\theta} + \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_{1,\theta} + \mathbf{L}_{ij}^\top \boldsymbol{\gamma}_{2,\theta} + (\mathbf{L}_{ij} \otimes \mathbf{T}_{ij})^\top \boldsymbol{\gamma}_{3,\theta} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_{i,\theta} + \epsilon_{ij,\theta}. \quad (2.1)$$

where the fixed effects correspond to (a) the quadratic polynomials of time effects, $\mathbf{X}_{ij} = (1, j, j^2)^\top$, which are determined from linear, quadratic and cubic polynomials on the repeated measure data (King et al. (2015)) through model selections using AIC/BIC values; (b) the lipid and environmental main effects that are represented by $\mathbf{L}_{ij}$ and $\mathbf{T}_{ij}$ respec-

tively; and (c) the lipid–environment interactions that can be expressed as a Kronecker product between the two types of main effects as a mp -dimensional vector:

$$\mathbf{L}_{ij} \otimes \mathbf{T}_{ij} = [L_{ij1}T_{ij1}, L_{ij1}T_{ij2}, \dots, L_{ij1}T_{ijp}, L_{ij2}T_{ij1}, \dots, L_{ijm}T_{ijp}]^\top.$$

The fixed effects are $\boldsymbol{\gamma}_{0,\theta}$, the 3-dimensional coefficient vector for $\mathbf{X}_{ij}$, and $\boldsymbol{\gamma}_{1,\theta}, \boldsymbol{\gamma}_{2,\theta}, \boldsymbol{\gamma}_{3,\theta}$, the p -, m - and mp - dimensional regression coefficient vectors that represent the treatment and lipid main effects and their interactions, correspondingly. The above model also includes $\mathbf{Z}_{ij}$ with random effects $\boldsymbol{\alpha}_{i,\theta}$ to account for the intra-correlation among repeated measurements. The $h \times 1$ vector $\mathbf{Z}_{ij}$ corresponds to the random intercept - slope model and random intercept model under $h = 2$ and 1 , accordingly. The model error $\epsilon_{ij,\theta}$'s are independent with the θ th quantile being 0. Without loss of generality, we suppress the subscript “ θ ” of the regression coefficient vectors for both fixed and random effects from now on for simplicity of notation. For instance, the regression coefficients for the main and interaction effects can be further written as $\boldsymbol{\gamma}_1 = (\gamma_{11}, \dots, \gamma_{1p})^\top$, $\boldsymbol{\gamma}_2 = (\gamma_{21}, \dots, \gamma_{2m})^\top$, $\boldsymbol{\gamma}_3 = (\boldsymbol{\gamma}_{31}, \dots, \boldsymbol{\gamma}_{3m})^\top$ where $\boldsymbol{\gamma}_{3l} = (\gamma_{3l1}, \dots, \gamma_{3lp})^\top$ for $l = 1, \dots, m$.

The quantile mixed effect model (2.1) provides a general formulation for modeling longitudinal gene–environment interactions where both the main and interaction effects, as well as the phenotypic traits, are repeatedly measured. Such a general formulation is applicable to the motivating lipidomics data where the G and E factors are cross-sectional, and only the response variable is longitudinal.

2.2.2 The robust Bayesian interaction analysis with longitudinal traits

The histogram on repeatedly measured bodyweight of the CD-1 mice in Figure 2.1 shows heavy-tailed distribution with outliers, demanding robust longitudinal methods for interaction studies. Recently, robust Bayesian analysis has emerged as a powerful tool to dissect G×E interactions in the presence data heterogeneity arising from complex disease traits (Ren

et al. (2023)). The key to incorporate robustness in a Bayesian framework is to specify the likelihood based on a heavy-tailed distribution. We consider assuming that the error term ϵ_{ij} 's follow i.i.d asymmetric Laplace distribution with density function

$$f(\epsilon_{ij}|\theta, \tau) = \theta(1 - \theta)\tau \exp\{-\tau\rho_\theta(\epsilon_{ij})\}, \quad i = 1, \dots, n, \quad j = 1, \dots, k,$$

where τ^{-1} is the scale parameter indicating the skewness of the asymmetric Laplace distribution (ALD), and $\rho_\theta(\cdot)$ is the check loss function with quantile level $\theta \in (0, 1)$:

$$\rho_\theta(t) = \begin{cases} \theta t, & \text{if } t \geq 0, \\ -(1 - \theta)t, & \text{if } t < 0, \end{cases}$$

which is essential for the Bayesian formulation of quantile regression. The heavy-tailed pattern of ALD can be readily manifested when it reduces to the well-known double exponential (or Laplace) distribution at the 50% quantile level.

With the above probabilistic assumption on model error, by conditioning on the fixed and random effects as well as the associated regression coefficients defined in model (2.1), the distribution of Y_{ij} can be expressed as:

$$f(Y_{ij}|\mu_{ij}, \boldsymbol{\alpha}_i, \mathbf{Z}_{ij}, \tau) = \theta(1 - \theta)\tau \exp\{-\tau\rho_\theta(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i)\}.$$

where the mean function μ_{ij} is defined as $\mu_{ij} = \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0 + \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1 + \mathbf{L}_{ij}^\top \boldsymbol{\gamma}_2 + \mathbf{W}_{ij}^\top \boldsymbol{\gamma}_3$ including the main genetic and environmental effects, and their interactions.

Kozumi and Kobayashi (Kozumi and Kobayashi (2011)) have provided an equivalent representation of the skewed Laplace distribution as a mixture of an exponential and a scaled normal distribution, which facilitates the setup of a Bayesian hierarchical model for quantile regression. Specifically, with $\xi_1 = \frac{1-2\theta}{\theta(1-\theta)}$ and $\xi_2 = \sqrt{\frac{2}{\theta(1-\theta)}}$ at quantile level $0 < \theta < 1$, the error term ϵ_{ij} follows the ALD with a scale parameter τ^{-1} as

$$\epsilon_{ij} = \tau^{-1}\xi_1 v_{ij} + \tau^{-1}\xi_2 \sqrt{v_{ij}} z_{ij},$$

where v_{ij} is the standard exponential distribution, $\text{Exp}(1)$, and z_{ij} is a standard normal random variable. Subsequently, model (2.1) can be rewritten as:

$$Y_{ij} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i + \tau^{-1} \xi_1 v_{ij} + \tau^{-1} \xi_2 \sqrt{v_{ij}} z_{ij},$$

Let $\tilde{v}_{ij} = \tau^{-1} v_{ij}$, then $\tilde{v}_{ij} \stackrel{\text{ind}}{\sim} \text{Exp}(\tau^{-1})$ for $i = 1, \dots, n$, and $j = 1, \dots, k$. Then the Bayesian hierarchal model has the following representation :

$$\begin{aligned} Y_{ij} &= \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i + \xi_1 \tilde{v}_{ij} + \tau^{-1/2} \xi_2 \sqrt{\tilde{v}_{ij}} z_{ij}, \\ \tilde{v}_{ij} | \tau &\stackrel{\text{ind}}{\sim} \tau \exp(-\tau \tilde{v}_{ij}), \\ z_{ij} &\stackrel{\text{ind}}{\sim} \text{N}(0, 1). \end{aligned}$$

2.2.3 Robust Bayesian variable selection for longitudinal interaction studies

The connection between ANOVA and grouped variable selection of main and interaction effects has been pointed out in Yuan and Lin 2006 (Yuan and Lin (2006b)). In the repeated measure ANOVA analysis conducted in King et al. (King et al. (2015)), considering the high-dimensional lipidomics features eventually leads to the proposed longitudinal $G \times E$ study that demands tailored variable selection of main and interaction effects. Recall that in our study, there are 4 treatment groups of exercise and/or dietary restriction, which are sedentary and ad libitum feeding (control), 20% dietary calorie restriction and sedentary (DCR), Pair feeding and exercise (PE) and Ad libitum feeding and exercise (AE). With the control as the baseline level, the treatment group can be represented by 3 binary indicators, or dummy variables, for DCR, PE and AE, respectively. In model (2.1), the E factors are modeled as $\mathbf{T}_{ij}$, and the $G \times E$ interactions are represented as the Kronecker product $\mathbf{L}_{ij} \otimes \mathbf{T}_{ij}$. The $\mathbf{T}_{ij}$ has a group structure, that is, the treatment groups can be correctly specified only when all the three binary indicators are present. Although $\mathbf{T}_{ij}$ is low dimensional and not subject to selection, a direct consequence is that the interaction between $\mathbf{T}_{ij}$ and any lipidomics

features also have the group structure, and identifying G×E interactions needs to take the grouping pattern into account.

Besides the G×E interactions, we also perform selection of the lipidomics main effect $\mathbf{L}_{ij}$ individually. Therefore, detection of important main and interaction effects amounts to variable selection accommodating the individual and group level selection simultaneously. The regularized estimates of the main and interaction effects can be obtained by minimizing the following loss function:

$$\sum_{i=1}^n \sum_{j=1}^k \rho_{\theta}(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^{\top} \boldsymbol{\alpha}_i) + \lambda_1 \sum_{l=1}^m |\gamma_{2l}| + \lambda_2 \sum_{l=1}^m \|\boldsymbol{\gamma}_{3l}\|_2. \quad (2.2)$$

where the L_2 norm of $\boldsymbol{\gamma}_{3l}$ is defined as $\|\boldsymbol{\gamma}_{3l}\|_2 = (\boldsymbol{\gamma}_{3l}^{\top} \boldsymbol{\gamma}_{3l})^{1/2}$, and λ_1 and λ_2 are two positive tuning parameters. Here, γ_{2l} is the coefficient for the main effect of the l th lipid feature, and $\boldsymbol{\gamma}_{3l}$ correspond to the interaction between the l th lipid feature and environmental factors, i.e., the group of dummy variables. If L_2 norm of the vector $\boldsymbol{\gamma}_{3l}$ is estimated as 0, i.e., $\boldsymbol{\gamma}_{3l} = \mathbf{0}$, there is no corresponding interaction effects. In addition, if $\gamma_{2l}=0$, then the l th lipid does not have main effect and is thus not associated with the repeatedly measured bodyweight of CD1 mice.

With Bayesian analysis, selection of important main and interaction effects can be achieved by assigning individual and group level sparse priors on γ_{2l} and $\boldsymbol{\gamma}_{3l}$, ($1 \leq l \leq m$), accordingly. In published studies, the conditional Laplace prior and its multivariate version have been developed for Bayesian LASSO and Bayesian group LASSO, respectively (Casella et al. (2010); Park and Casella (2008b)). Their posterior modes correspond to the regularized LASSO and group LASSO estimates in the frequentist framework, respectively. Further more, under the ALD based likelihood, (Li et al. (2010b)) have shown that the Laplacian shrinkage priors work as well in regularized Bayesian quantile regression. A major limitation of the priors is that the corresponding posterior estimates do not lead to exact 0 coefficients, which can be overcome by incorporating the spike-and-slab priors in the hierarchical model (Ren et al. (2023); Zhou et al. (2023)). Next, we impose the two types of sparse priors hierarchically

to formulate the robust Bayesian hierarchical model and identify the main and interaction effects under the longitudinal trait.

Selecting main and interaction effects with Laplacian shrinkage To identify main lipidomics effects that are associated with the phenotype, we assign the independent and identical Laplace prior on each component of coefficient vector $\boldsymbol{\gamma}_2$ as follows:

$$\pi(\boldsymbol{\gamma}_2|\tau, \lambda_1) = \left(\frac{\tau\lambda_1}{2}\right)^m \exp(-\tau\lambda_1 \sum_{l=1}^m |\gamma_{2l}|).$$

Meanwhile, to pinpoint important lipid-environment interactions, we impose the group-level shrinkage by adopting multivariate Laplace prior on $\boldsymbol{\gamma}_{3l}$:

$$\pi(\boldsymbol{\gamma}_{3l}|\tau, \lambda_2) = C_p(\tau\lambda_2)^p \exp(-\tau\lambda_2 \|\boldsymbol{\gamma}_{3l}\|_2),$$

where $C_p = 2^{-(p+1)/2} (2\pi)^{-(p-1)/2} / \Gamma((p+1)/2)$ with gamma function $\Gamma(\cdot)$, and p is the group size.

Then the joint posterior distribution of $\boldsymbol{\gamma}_2$ and $\boldsymbol{\gamma}_3$ is:

$$f(\boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3|Y_{ij}, \mu_{ij}, \mathbf{Z}_{ij}^\top, \alpha_i, \tau, \lambda_1, \lambda_2) \propto \exp\left\{-\tau \sum_{i=1}^n \sum_{j=1}^k \rho_\theta(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \alpha_i) - \tau\lambda_1 \sum_{l=1}^m |\gamma_{2l}| - \tau\lambda_2 \sum_{l=1}^m \|\boldsymbol{\gamma}_{3l}\|_2\right\},$$

where the power of the exponential kernel is corresponding to the objective function in eq. (2.2). To facilitate the set up the Bayesian hierarchical model, we denote $\eta_1 = \tau\lambda_1$. Then the *i.i.d* Laplace priors on $\boldsymbol{\gamma}_2$ representing the main genetic effects can be written as

$$\pi(\boldsymbol{\gamma}_2|\tau, \lambda_1) = \prod_{l=1}^m \frac{\eta_1}{2} \exp\{-\eta_1 |\gamma_{2l}|\} = \prod_{l=1}^m \int_0^\infty \frac{1}{\sqrt{2\pi s_{1l}}} \exp\left(-\frac{\gamma_{2l}^2}{2s_{1l}}\right) \frac{\eta_1^2}{2} \exp\left(-\frac{\eta_1^2}{2} s_{1l}\right) ds_{1l}.$$

The above Laplacian priors can be hierarchically expressed as scale mixture of normals after s_{1l} is integrated out:

$$\begin{aligned} \gamma_{2l}|s_{1l} &\overset{ind}{\sim} N(0, s_{1l}), \\ s_{1l} &\overset{ind}{\sim} \frac{\eta_1^2}{2} \exp\left\{-\frac{\eta_1^2}{2} s_{1l}\right\}. \end{aligned} \tag{2.3}$$

Besides, with the reparametrization $\eta_2 = \tau\lambda_2$, the multivariate Laplace prior on $\boldsymbol{\gamma}_{3l}$ can be written as:

$$\begin{aligned}\pi(\boldsymbol{\gamma}_{3l}|\eta_2) &\triangleq \text{M-Laplace}(\boldsymbol{\gamma}_{3l}|\eta_2) \\ &= C_p \eta_2^p \exp(-\eta_2 \|\boldsymbol{\gamma}_{3l}\|_2) \\ &= \int_0^\infty (2\pi s_{2l})^{-p/2} \exp\left\{-\frac{1}{2s_{2l}} \boldsymbol{\gamma}_{3l}^\top \boldsymbol{\gamma}_{3l}\right\} \times \frac{(\eta_2^2/2)^{(p+1)/2}}{\Gamma(\frac{p+1}{2})} s_{2l}^{(p-1)/2} \exp\left(-\frac{\eta_2^2}{2} s_{2l}\right) ds_{2l}.\end{aligned}$$

Then, the prior of $\boldsymbol{\gamma}_{3l}$ can be expressed as the hierarchical representation below:

$$\begin{aligned}\pi(\boldsymbol{\gamma}_{3l}|s_{2l}) &\stackrel{\text{ind}}{\sim} N_p(0, s_{2l} \mathbf{I}_p), \\ s_{2l} &\stackrel{\text{ind}}{\sim} \text{Gamma}\left(\frac{p+1}{2}, \frac{\eta_2^2}{2}\right),\end{aligned}\tag{2.4}$$

where $\frac{p+1}{2}$ is the shape parameter and $\frac{\eta_2^2}{2}$ is the rate parameter of the gamma distribution.

Selection of main and interaction effects with structured spike-and-slab priors

In the Bayesian hierarchical model, the prior formulations of univariate and multivariate Laplace priors in (2.3) and (3.7) will lead to the shrinkage estimation for the identification of main and interaction effects, respectively. However, a well-acknowledged drawback of Laplace priors is that they do not produce exact 0 estimates. To promote exact sparsity, we refine the prior formulation by further incorporating the spike-and-slab priors.

To detect the gene-environment interactions, we have

$$\begin{aligned}\boldsymbol{\gamma}_{3l}|\phi_{2l}, s_{2l} &\stackrel{\text{ind}}{\sim} \phi_{2l} N_p(\mathbf{0}, \text{diag}(s_{2l}, \dots, s_{2l})) + (1 - \phi_{2l}) \delta_0(\boldsymbol{\gamma}_{3l}), \\ \phi_{2l} &\stackrel{\text{ind}}{\sim} \text{Bernoulli}(\pi_2), \\ s_{2l} &\stackrel{\text{ind}}{\sim} \text{Gamma}\left(\frac{p+1}{2}, \frac{\eta_2^2}{2}\right).\end{aligned}\tag{2.5}$$

where the slab component $N_p(0, s_{2l} \mathbf{I}_p)$ models the large and diffuse signals, and the spike distribution $\delta_0(\boldsymbol{\gamma}_{3l})$ denotes the point mass at $\mathbf{0}_{p \times 1}$, with parameter $\pi_2 \in [0, 1]$. The ϕ_{2l} , ($1 \leq l \leq m$), is a latent binary indicator variable for the interactions between the l th lipid factor and the environmental factors, which indicates whether or not the corresponding interaction

is selected. When $\phi_{2l} = 0$, the coefficient vector γ_{3l} has point mass density at zero, which means that the l th lipid factor does not have interaction effect with the treatment groups. On the other hand, if $\phi_{2l} = 1$, the spike-and-slab prior in (2.5) reduces to the multivariate Laplace prior in (3.7) and results in the Bayesian quantile group LASSO, indicating that the corresponding interactions are associated with the longitudinal phenotypic traits. Moreover, the prior formulation in (2.5) can be treated as the following multivariate spike-and-slab priors after ϕ_{2l} and s_{2l} are integrated out:

$$\gamma_{3l}|\eta_2 \stackrel{ind}{\sim} \phi_{2l}\text{M-Laplace}(\gamma_{3l}|\eta_2) + (1 - \phi_{2l})\delta_0(\gamma_{3l}),$$

where the parameter $\pi_2 \in [0, 1]$.

For identification of lipidomics main effect, we impose the univariate spike-and-slab priors for γ_{2l} , $l = (1, \dots, m)$ as follows:

$$\begin{aligned} \gamma_{2l}|\phi_{1l}, s_{1l} &\stackrel{ind}{\sim} \phi_{1l}\text{N}(0, s_{1l}) + (1 - \phi_{1l})\delta_0(\gamma_{2l}), \\ \phi_{1l} &\stackrel{ind}{\sim} \text{Bernoulli}(\pi_1), \\ s_{1l} &\stackrel{ind}{\sim} \text{Gamma}(1, \frac{\eta_1^2}{2}). \end{aligned} \tag{2.6}$$

where the spike component $\delta_0(\gamma_{2l})$ is a point mass at 0 and ϕ_{1l} determines whether the spike or slab component is in action, with $\pi_1 \in [0, 1]$. The prior specified in (2.6) is equivalent to the univariate Laplace prior in (2.3) when $\phi_{1l} = 1$, indicating that the l th lipid has a main effect. Otherwise, when $\phi_{1l} = 0$, the l th lipid does not have a main effect.

We also assign conjugate beta priors $\pi_1 \sim \text{Beta}(r_1, w_1)$ and $\pi_2 \sim \text{Beta}(r_2, w_2)$ Gamma priors are imposed on $\eta_1^2 \sim \text{Gamma}(a_1, b_1)$ and $\eta_2^2 \sim \text{Gamma}(a_2, b_2)$, $\tau \sim \text{Gamma}(c, d)$, also inverse-gamma prior $\phi^2 \sim \text{Inverse-Gamma}(e, f)$. In particular, we set the hyperparameters $a_1 = b_1 = a_2 = b_2 = c = d = r_1 = w_1 = r_2 = w_2 = e = f = 1$. For the random effect α_i , normal prior has been imposed on it. For the random intercept - slope model: $\alpha_i \sim N(\mathbf{0}, \phi^2 \mathbf{I})$. For the random intercept model: $\alpha_i \sim N(0, \phi^2)$.

2.2.4 Gibbs Sampler

The joint posterior distribution can be expressed as

$$\begin{aligned}
& \pi(\boldsymbol{\gamma}_0, \boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3, s_{1l}, s_{2l}, \eta_1^2, \eta_2^2, \pi_1, \pi_2, \tau | Y_{ij}) \\
& \propto (\tau^{-1})^{-\frac{nk}{2}} \left(\prod_{i=1}^n \prod_{j=1}^k \tilde{v}_{ij} \right)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\tau^{-1}\xi_2^2} \sum_{i=1}^n \sum_{j=1}^k \left(\frac{Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}}{\sqrt{\tilde{v}_{ij}}} \right)^2 \right\} \\
& \times \exp \left(-\frac{1}{2} \boldsymbol{\gamma}_0^\top \boldsymbol{\Sigma}_0^{-1} \boldsymbol{\gamma}_0 \right) \\
& \times \exp \left(-\frac{1}{2} \boldsymbol{\gamma}_1^\top \boldsymbol{\Sigma}_1^{-1} \boldsymbol{\gamma}_1 \right) \\
& \times \prod_{l=1}^m \left((1 - \pi_1)(2\pi s_{1l})^{-\frac{1}{2}} \exp \left(-\frac{1}{2s_{1l}} \gamma_{2l}^2 \right) \mathbf{I}_{\{\gamma_{2l} \neq 0\}} + \pi_1 \delta_0(\gamma_{2l}) \right) \\
& \times \tau^{nk} \exp \left(-\tau \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij} \right) \\
& \times \tau^{c-1} \exp(-d\tau) \\
& \times (\eta_1^2)^{a_1-1} \exp(-b_1 \eta_1^2) \prod_{l=1}^p \frac{\eta_1^2}{2} \exp \left(-\frac{\eta_1^2}{2} s_{1l} \right) \\
& \times \pi_1^{r_1-1} (1 - \pi_1)^{w_1-1} \\
& \times \prod_{l=1}^m \left((1 - \pi_2)(2\pi s_{2l})^{-\frac{p}{2}} \exp \left(-\frac{1}{2s_{2l}} \boldsymbol{\gamma}_{3l}^\top \boldsymbol{\gamma}_{3l} \right) \mathbf{I}_{\{\boldsymbol{\gamma}_{3l} \neq \mathbf{0}\}} + \pi_2 \delta_0(\boldsymbol{\gamma}_{3l}) \right) \\
& \times (\eta_2^2)^{a_2-1} \exp(-b_2 \eta_2^2) \prod_{l=1}^m \left(\frac{p\eta_2^2}{2} \right)^{p+1} (\eta_2^2)^{\frac{p+1}{2}-1} \exp \left(-\frac{\eta_2^2}{2} s_{2l} \right) \\
& \times \pi_2^{r_2-1} (1 - \pi_2)^{w_2-1} \\
& \times (\phi^2)^{-\frac{nh}{2}} \prod_{i=1}^n \prod_{t=1}^h \exp \left(-\frac{1}{2\phi^2} \alpha_{it}^2 \right) \\
& \times (\phi^2)^{e-1} \exp(-f\phi^2).
\end{aligned}$$

- Let $\mu_{ij(-\boldsymbol{\gamma}_0)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of $\boldsymbol{\gamma}_0$.

The full conditional distribution of $\boldsymbol{\gamma}_0$ is a multivariate normal distribution $N_3(\boldsymbol{\mu}_{\boldsymbol{\gamma}_0}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0})$ with

mean

$$\boldsymbol{\mu}_{\gamma_0} = \boldsymbol{\Sigma}_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{X}_{ij} (Y_{ij} - \mu_{ij(-\gamma_0)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\boldsymbol{\Sigma}_{\gamma_0} = \left(\boldsymbol{\Sigma}_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{X}_{ij} \mathbf{X}_{ij}^\top \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_1)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1$, ($i = 1, \dots, n, j = 1, \dots, k$), then the full conditional distribution of $\boldsymbol{\gamma}_1$ is multivariate normal distribution $N_p(\boldsymbol{\mu}_{\gamma_1}, \boldsymbol{\Sigma}_{\gamma_1})$ with mean

$$\boldsymbol{\mu}_{\gamma_1} = \boldsymbol{\Sigma}_{\gamma_1} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{T}_{ij} (Y_{ij} - \mu_{ij(-\gamma_1)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\boldsymbol{\Sigma}_{\gamma_1} = \left(\boldsymbol{\Sigma}_1^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{T}_{ij} \mathbf{T}_{ij}^\top \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{2l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - L_{ijl} \gamma_{2l}$ and $l_{1l} = \pi(\gamma_{2l} = 0 | \text{rest})$, the conditional posterior distribution of the coefficient of lipid factor γ_{2l} is a spike-and-slab distribution:

$$\gamma_{2l} | \text{rest} \sim (1 - l_{1l}) N(\mu_{\gamma_{2l}}, \Sigma_{\gamma_{2l}}) + l_{1l} \delta_0(\gamma_{2l}),$$

where

$$\begin{aligned} \mu_{\gamma_{2l}} &= \Sigma_{\gamma_{2l}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\gamma_{2l}} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl}^2 + \frac{1}{s_{1l}} \right)^{-1}. \end{aligned}$$

and

$$l_{1l} = \frac{\pi_1}{\pi_1 + (1 - \pi_1)(s_{1l})^{-\frac{1}{2}}(\Sigma_{\gamma_{2l}})^{\frac{1}{2}} \exp\{\frac{1}{2} \Sigma_{\gamma_{2l}} (\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl}(\text{res}))^2\}}.$$

$$\text{res} = Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}.$$

The posterior distribution of γ_{2l} is a mixture of a normal distribution and a point mass at 0. At each iteration of MCMC, γ_{2l} is drawn from $N(\mu_{\gamma_{2l}}, \Sigma_{\gamma_{2l}})$ with probability $(1 - l_{1l})$ and is set to 0 with probability l_{1l} . If γ_{2l} is set to 0, then we have $\phi_{1l} = 0$. Otherwise, $\phi_{1l} = 1$.

- The posterior distribution of the interaction between the l th lipid and treatment factors $\boldsymbol{\gamma}_{3l}$ ($l = 1, \dots, m$) is also a spike-and-slab distribution. Denote $\mu_{ij(-\gamma_{3l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{W}_{ijl}^\top \boldsymbol{\gamma}_{3l}$ and $l_{2l} = \pi(\boldsymbol{\gamma}_{3l} = \mathbf{0} | \text{rest})$, the conditional posterior distribution of $\boldsymbol{\gamma}_{3l}$ is a multivariate spike-and-slab distribution :

$$\boldsymbol{\gamma}_{3l} | \text{rest} \sim (1 - l_{2l}) N_p(\boldsymbol{\mu}_{\gamma_{3l}}, \Sigma_{\gamma_{3l}}) + l_{2l} \delta_0(\boldsymbol{\gamma}_{3l}),$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\gamma_{3l}} &= \Sigma_{\gamma_{3l}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{W}_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{3l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\gamma_{3l}} &= (\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{W}_{ijl} \mathbf{W}_{ijl}^\top + \frac{1}{s_{2l}} \mathbf{I}_p)^{-1}. \end{aligned}$$

and

$$l_{2l} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)(s_{2l})^{-\frac{p}{2}} |\Sigma_{\gamma_{3l}}|^{\frac{1}{2}} \exp\{\frac{1}{2} \|\Sigma_{\gamma_{3l}}^{-\frac{1}{2}} \sum_{i=1}^n \sum_{j=1}^k \tau \xi_2^{-2} \tilde{v}_{ij}^{-1} \mathbf{W}_{ijl}(\text{res})\|_2^2\}}.$$

$$\text{res} = Y_{ij} - \mu_{ij(-\gamma_{3l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}.$$

The posterior distribution of $\boldsymbol{\gamma}_{3l}$ is a mixture of a multivariate normal distribution and a point mass at $\mathbf{0}$. At each iteration of MCMC, $\boldsymbol{\gamma}_{3l}$ is drawn from $N_p(\boldsymbol{\mu}_{\gamma_{3l}}, \Sigma_{\gamma_{3l}})$ with probability $(1 - l_{2l})$ and is set to 0 with probability l_{2l} . If $\boldsymbol{\gamma}_{3l}$ is set to $\mathbf{0}$, then we have $\phi_{2l} = 0$. Otherwise, $\phi_{2l} = 1$.

- The full conditional posterior distribution of $(s_{1l})^{-1}$ is:

$$(s_{1l})^{-1}|\text{rest} \sim \begin{cases} \text{Inverse-Gamma}(1, \frac{\eta_1^2}{2}) & \text{if } \gamma_{2l} = 0 \\ \text{Inverse-Gaussian}(\eta_1^2, \sqrt{\frac{\eta_1^2}{\gamma_{2l}^2}}) & \text{if } \gamma_{2l} \neq 0 \end{cases}$$

- Similarly, The full conditional posterior distribution of $(s_{2l})^{-1}$ is:

$$(s_{2l})^{-1}|\text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{p+1}{2}, \frac{\eta_2^2}{2}) & \text{if } \gamma_{3l} = 0 \\ \text{Inverse-Gaussian}(\eta_2^2, \sqrt{\frac{\eta_2^2}{\|\gamma_{3l}\|_2^2}}) & \text{if } \gamma_{3l} \neq 0 \end{cases}$$

- The full conditional posterior distribution of η_1^2 :

$$\eta_1^2|\text{rest} \sim \text{Gamma}(a_1 + p, b_1 + \frac{\sum_{l=1}^m s_{1l}}{2}).$$

- Likewise, the full conditional posterior distribution of η_2^2 :

$$\eta_2^2|\text{rest} \sim \text{Gamma}(a_2 + \frac{m(p+1)}{2}, b_2 + \frac{\sum_{l=1}^m s_{2l}}{2}).$$

- The full conditional posterior distribution of π_1 :

$$\pi_1|\text{rest} \sim \text{Beta}(r_1 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l}=0\}}, w_1 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l} \neq 0\}}).$$

- Similarly, the full conditional posterior distribution of π_2 :

$$\pi_2|\text{rest} \sim \text{Beta}(r_2 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{3l}=0\}}, w_2 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{3l} \neq 0\}}).$$

- The full conditional posterior distribution of $\tilde{v}_{ij}$:

$$\tilde{v}_{ij}|\text{rest} \sim \text{Inverse-Gaussian}\left(\sqrt{\frac{\xi_1^2 + 2\xi_2^2}{(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i)^2}}, \frac{\tau \xi_1^2}{\xi_2^2} + 2\tau\right).$$

- The full conditional posterior distribution of τ :

$$\tau|\text{rest} \sim \text{Gamma}\left(c + \frac{3nk}{2}, d + \sum_{i=1}^n \sum_{j=1}^k (\tilde{v}_{ij} + (2\xi_2^2)^{-1} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij})^2)\right).$$

- The full conditional posterior distribution of $\boldsymbol{\alpha}_i$:

$$\boldsymbol{\alpha}_i|\text{rest} \sim \text{N}_2(\boldsymbol{\mu}_{\boldsymbol{\alpha}_i}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i})$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\alpha}_i} &= \frac{\tau}{\xi_2^2} \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} \sum_{j=1}^k \frac{\mathbf{Z}_{ij}(Y_{ij} - \mu_{ij} - \xi_1 \tilde{v}_{ij})}{\tilde{v}_{ij}} \\ \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} &= \left(\frac{\tau}{\xi_2^2} \sum_{j=1}^k \frac{\mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top}{\tilde{v}_{ij}} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

- Lastly, the full conditional posterior distribution of ϕ^2 :

$$\phi^2|\text{rest} \sim \text{Inverse-Gamma}\left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_i\right)$$

With the spike-and-slab priors, the proposed robust Bayesian variable selection method can simultaneously identify main and interaction effects on the individual and group level, or bi-level, respectively. Therefore, we term it as RBVSS-BI. An immediate competitor of RBVSS-BI ignores the group structure of the environmental factors and select all the main and interaction effects individually, which is thus denoted as RBVSS. Their counterpart without incorporating the spike and slab priors are termed as RBV-BI and RBV, correspondingly. These 4 methods are robust under the ALD based likelihood. Their non-robust counterpart using the normal likelihood in the hierarchical models are denoted as BVSS-BI,

BVSS, BV-BI and BV, respectively. The C++ implementations of all the 8 methods under comparison have been provided in R package `mixedBayes` (Fan and Wu) on CRAN. We have shown the detailed derivations of the Gibbs sampler of the other 7 alternative methods in the Appendix. Besides, Table A.1 and A.2 in the Appendix have provided detailed summaries of and comparisons between all the methods under comparison.

2.3 Simulation

The aforementioned 8 methods under comparison can be classified by three criteria: with or without robustness, with or without spike-and-slab priors and with or without respecting the bi-level structure for main and interaction effects in variable selection under the longitudinal traits. Detailed summaries of these methods are available in Table A.1 and A.2 in the Appendix. In simulation study, we comprehensively assess the performance of all the methods and demonstrate the utility of proposed RBVSS-BI due to accommodating robustness, spike-and-slab priors and structured variable selection simultaneously.

The longitudinal response variable is generated from model (2.1) with sample size $n = 200$. Specifically, we consider two types of random effects: (1) the random intercept model, and (2) the random intercept-slope model, which are corresponding to $h = 1$ and 2 in model (2.1), respectively. The number of time points k is set to 5. We set the lipid factors $\mathbf{L}_{ij}$ with dimensions $m = 30$ and 60. The dimension of the factors $\mathbf{T}_{ij}$ is $p = 3$. Therefore the dimension of lipid by treatment interaction is pm . In ANOVA, the factors can be represented by a group of dummy variables. Therefore, we generate a group of three dummy variables by dichotomizing the vector at the 30th and 70th percentiles. In addition, the lipid factors are simulated from a multivariate normal distribution with mean $\mathbf{0}$ and the variance covariance matrix following a AR(1) structure with marginal variance 1 and auto-correlation coefficient $\rho = 0.5$.

The fixed effects in model (2.1) have been generated from a uniform distribution $U(0.4, 0.8)$ for a total of 26 important effects, consisting of 1 for the intercept, 3 for treatment dummy variables, 4 for lipid main effects, and 18 for 6 groups of lipid-treatment interactions. The

nonzero coefficients for the main effects are corresponding to the 2nd, 4th, 7th and 12th lipids in the design matrix and the nonzero coefficients for the interactions effects are corresponding to groups (or interactions) 2, 4, 6, 7, 12 and 15 with group size 3. Therefore, 18 nonzero coefficients need to be identified through regularized Bayesian estimation. The intercept and treatment dummy variables are not subjected to selection. Each component in the h by 1 vector of random effect α_i , where $h=1$ and 2 for the random intercept and random intercept-slope models, has been generated from a uniform distribution $U(0, 1)$. The remaining coefficients were all set to zero. The model errors have been simulated from $N(0,1)$ (Error 1), $t(2)$ (Error 2), $\text{Laplace}(0, 2)$ (Error 3), $20\%\text{Laplace}(0, 1)+80\%\text{Laplace}(0, \sqrt{5})$ (Error 4), and $\text{LogNormal}(0, 1)$ (Error 5). All the distributions have heavy-tails except $N(0,1)$.

Identification and estimation criteria are adopted to assess the performance using posterior samples from the Gibbs sampler. In terms of identification, we use the median probability model (MPM) to identify predictors that are associated with the response variable for methods with spike-and-slab priors ([Barbieri and Berger \(2004\)](#)). Suppose we have collected H posterior samples from MCMC after dropping burn-ins. At the h th MCMC iteration, the l th predictor is included in the model if the corresponding indicator $\phi_l^{(h)} = 1$. Therefore, the posterior probability of including the l th predictor in the final model is defined as:

$$p_l = \hat{\pi}(\phi_l = 1|y) = \frac{1}{H} \sum_{h=1}^H \phi_l^{(h)}, l = 1, \dots, m.$$

In general, higher posterior inclusion probability p_l suggests stronger empirical evidence that the l th predictor has a non-zero coefficient and thus a stronger association with the response variable. The MPM model is defined as the predictors that have been selected in the model when the posterior inclusion probability is at least 0.5 ([Barbieri and Berger \(2004\)](#); [Barbieri et al. \(2021\)](#)). For the non spike-and-slab methods, the 95% credible interval (95% CI) has been adopted to select important main and interaction effects. Performance in identification is assessed in terms of the number of true positives (TP) and false positives (FP).

The estimation accuracy is evaluated by using the L1 difference between the estimated

and true regression coefficients over multiple simulation replicates. For convenience, we define $\mathbf{\Gamma} = (\boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3)^\top$, for the coefficients corresponding to the genetic main effects and G×E interactions that are subject to selection in model (2.1). We calculate the total mean absolute deviations, or TMAD, for the coefficient vector $\mathbf{\Gamma}$, as:

$$\text{TMAD} = \frac{1}{M} \sum_{i=1}^M |\hat{\mathbf{\Gamma}}^{(i)} - \mathbf{\Gamma}|/p_{\mathbf{\Gamma}}$$

where $p_{\mathbf{\Gamma}}$ is the dimension of $\mathbf{\Gamma}$ and $\hat{\mathbf{\Gamma}}^{(i)}$ is the estimated value of $\mathbf{\Gamma}$ in the i th simulated dataset, $i = 1, \dots, M$. M is the total number of simulation replications. In addition, we also compute the mean absolute deviations corresponding to the true nonzero and zero coefficients, denoted as MAD and NMAD, respectively.

We collect the posterior samples via Gibbs sampler with 10,000 MCMC iterations where the first 5,000 are burn-ins, and compute the Bayesian estimates using posterior medians. All methods under comparison have been evaluated under both the random intercept-slope model and random intercept model. We have listed identification and estimation results in tables and highlighted the best method in boldface. Table 3.3 and Table 3.2 present the identification results under random intercept-slope model for $m=30$ and 60 respectively, at the 50% quantile level ($\theta=0.5$), which clearly shows the advantage of incorporating (1) robustness, (2) the spike-and-slab prior, and (3) bi-level selection of main and interaction effects. Under heavy tailed errors and data contamination, the robust Bayesian variable selection methods outperform the non-robust alternatives in general. For instance, under $t(2)$ (Error 2), while the proposed RBVSS-BI and its non-robust counterpart BVSS-BI have similar TP counts, the FP count reduces dramatically from 7.59 to 0.08 across 100 replicates, indicating the strength of robust methods in the presence of data heterogeneity. Besides, the methods with spike-and-slab priors have outperformed the ones without adopting the priors to impose exact sparsity. For instance, in Table 3.3, under $N(0,1)$ (Error 1), the proposed RBVSS-BI and its non-spike-and-slab counterpart RBV-BI both have a small fraction of FP counts, but TP count of main and interaction effects improves from 14.00 to 21.25. Similarly, under

the Laplace(0,2) (Error 3), still between the two robust methods, the TP counts increases dramatically from 11.80 to 21.35 given similar number of FP counts. Furthermore, we can also observe the superior performance of the best method highlighted with boldface due to appropriately accommodate the bi-level sparsity. For example, with $t(2)$ (Error 2), RBV and RBVSS, the two robust methods that ignore the structured sparsity, both lead to inferior identification performance with smaller TP counts and higher FP counts, compared to those identified by RBVSS-BI. Same patterns have been observed under the high-dimensional setting as well in Table 3.2.

Table 2.3 and Table 2.4 have reported the estimation accuracy of all methods under the random intercept-slope model with $(n, m, p) = (200, 30, 3)$ and $(200, 60, 3)$ respectively. The mean absolute deviations between the estimated and true coefficients under the nonzero true coefficients (MAD), zero true coefficients (NMAD) and all coefficients (TMAD) have been provided for all the error distributions at 50% quantile level ($\theta = 0.5$). In Table 2.3, under Error 2, RBVSS-BI leads to a MAD of 0.045(sd0.017), NMAD of 0.002(sd0.003), and TMAD of 0.009(sd0.004), all of which are the smallest among all methods under comparison. Such a pattern indicates the superiority of the proposed RBVSS-BI over the other 7 competing methods due to account for robustness, structured variable selection and exact sparsity. The trend also applies under the high dimensional setting as well in Table 2.4.

In addition to the simulation results on the random intercept-slope model as shown in Table 3.3 to Table 2.4, we have also conducted simulations under the random intercept model, and tabulate the identification results in Table B.2 and Table B.3, and the estimation results in Table B.7 and Table A.6 respectively in the appendix. Besides, the performance has been assessed at 30% and 70% quantile levels. In the appendix, when $\theta = 0.3$ under random intercept-slope model and random intercept model, the identification results are provided in Table A.7 to Table A.10, while the estimation accuracy can be found in Table A.11 to Table A.14. With the 70% quantile level, Table A.15 to Table A.18 report the identification results, and Table A.19 to Table A.22 show the parameter estimation performance.

In addition, we have provided graphical display of tabulated results under random intercept-slope model to assist quick observations of the superior performance of RBVSS-BI. In the

Appendix, Table A.23 demonstrates the correspondence between Figures A.1, A.2, A.3, A.4 and their associated Tables. For example, Figure A.1 shows the TP and FP counts of all methods across 3 quantile levels and 5 error distributions with $m=30$, corresponding to Tables 3.3, A.7, A.15, respectively. Under the $N(0,1)$ error (Error 1), the proposed RBVSS-BI and its non-robust cousin BVSS-BI both lead to the largest number of TPs compared to others while maintaining very small FP counts, which manifests the importance of incorporating spike-and-slab priors and accommodating the interaction effects due to the ANOVA design. In the presence of heavy-tailed errors, although the non-robust counterpart still yields high TP counts, its FP number also increases remarkably. Meanwhile, other alternatives have detected a much smaller number of true signals. Bar-plots of TP and FP counts in Figure A.2 which corresponds to Table 3.2, A.8, A.16 with $m=60$, also shows a similar pattern.

Besides, the estimation performance of all methods under evaluation has been visualized in Figures A.3 and A.4 in the Appendix. Specifically, Figure A.3 presents the bar-plots of estimation errors from Tables 2.3, A.11 and A.19 under the random intercept model with $m=30$. The Figure clearly shows that RBVSS-BI has the best estimation performance, especially under heavy-tailed errors. On the other hand, BV has the worst estimation accuracy, which is reasonable since this approach is non-robust, ignores spike-and-slab priors and cannot accommodate the interaction structure. We can observe the same trend in A.4 that corresponds to Table 2.4, A.12 and A.20. Overall, Figures A.1, A.2, A.3 and A.4 in the Appendix shows the advantage of RBVSS-BI compared to the alternatives.

Table 2.1: Identification results for $n = 200, m = 30, p = 3$. Mean (sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 50th quantile.

$n = 200 \quad m = 30$		Main		Interaction		Total	
$\theta = 0.5$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.75(0.71)	0.02(0.82)	17.50(0.12)	0.02(0.32)	21.25(1.41)	0.04(0.83)
N(0,1)	RBVSS	3.75(0.71)	0.02(0.82)	14.30(0.71)	1.50(0.71)	18.05(1.41)	1.52(0.71)
	RBV-BI	2.50(2.12)	0.50(0.54)	11.50(1.31)	0.60(0.71)	14.00(2.34)	0.60(0.89)
	RBV	2.50(1.41)	0.60(0.71)	8.50(0.55)	1.50(0.81)	11.00(1.41)	2.10(1.41)
	BVSS-BI	3.55(0.82)	0.04(0.51)	17.30(0.23)	3.80(1.64)	20.85(0.65)	3.84(1.64)
	BVSS	3.50(2.12)	0.05(0.21)	14.12(0.33)	0.20(0.45)	17.62(2.02)	0.25(0.45)
	BV-BI	2.05(1.41)	0.10(0.31)	10.14(2.42)	2.40(0.55)	12.19(2.31)	2.50(0.55)
	BV	2.09(0.81)	0.25(0.45)	8.55(0.45)	0.85(0.56)	10.64(0.45)	1.10(0.58)
Error 2	RBVSS – BI	3.85(1.41)	0.03(0.41)	17.90(0.21)	0.05(0.21)	21.75(1.41)	0.08(0.51)
$t(2)$	RBVSS	3.84(1.52)	0.05(0.53)	15.30(1.41)	0.50(0.71)	19.14(1.52)	0.55(0.72)
	RBV-BI	2.02(0.81)	0.02(0.13)	10.17(0.44)	0.70(0.90)	12.19(0.85)	0.72(0.86)
	RBV	1.80(1.41)	0.03(0.25)	8.50(1.83)	0.20(0.45)	10.30(1.30)	0.23(0.45)
	BVSS-BI	3.20(0.71)	0.04(0.25)	17.89(0.23)	7.55(0.21)	21.09(0.56)	7.59(0.31)
	BVSS	3.02(0.45)	0.07(0.89)	12.40(2.41)	0.60(0.89)	15.42(0.59)	0.60(0.89)
	BV-BI	2.30(0.55)	0.50(0.45)	9.50(1.23)	0.45(0.56)	11.80(1.52)	0.95(0.55)
	BV	2.20(0.57)	0.60(0.55)	7.50(0.56)	0.57(0.45)	9.70(1.48)	1.17(0.58)
Error 3	RBVSS – BI	3.85(0.71)	0.09(0.71)	17.50(2.12)	0.07(0.17)	21.35(1.14)	0.16(0.41)
Laplace(0,2)	RBVSS	3.50(0.21)	0.50(0.71)	14.40(0.89)	0.16(0.31)	17.90(0.89)	0.66(0.43)
	RBV-BI	2.30(0.55)	0.40(0.21)	9.50(1.30)	0.67(0.46)	11.80(1.30)	1.07(0.48)
	RBV	2.20(0.45)	0.70(0.91)	7.50(1.52)	0.40(0.32)	9.70(1.82)	1.10(0.29)
	BVSS-BI	3.20(1.30)	0.21(0.25)	17.89(0.31)	7.35(0.32)	21.09(1.30)	7.35(0.42)
	BVSS	3.04(1.25)	0.02(0.21)	10.40(3.36)	0.80(1.30)	13.44(4.04)	0.82(1.30)
	BV-BI	2.60(1.14)	0.40(0.31)	8.50(2.55)	0.90(0.34)	11.10(2.07)	0.94(0.53)
	BV	2.40(1.10)	0.55(0.76)	7.20(1.84)	0.75(0.54)	9.60(1.79)	1.30(0.45)
Error 4	RBVSS – BI	3.55(0.71)	0.03(0.11)	17.95(0.05)	0.15(0.12)	21.53(0.25)	0.15(0.31)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.50(0.73)	0.02(0.41)	15.30(1.41)	0.30(0.65)	18.80(0.67)	0.32(0.85)
	RBV-BI	2.50(0.75)	0.40(0.31)	12.50(0.71)	0.85(0.91)	15.00(1.82)	1.25(0.86)
	RBV	2.50(0.71)	0.55(0.43)	11.40(1.45)	0.95(0.76)	13.90(1.30)	1.50(0.80)
	BVSS-BI	2.80(1.41)	0.02(0.38)	17.23(0.31)	8.27(0.34)	20.03(0.68)	8.29(0.33)
	BVSS	2.80(0.35)	0.02(0.54)	13.67(2.52)	0.10(0.31)	16.47(2.62)	0.12(0.31)
	BV-BI	2.50(0.77)	0.07(0.25)	11.57(2.64)	0.13(0.35)	14.07(2.90)	0.20(0.48)
	BV	2.30(0.84)	0.08(0.25)	9.25(2.26)	0.17(0.38)	11.55(2.76)	0.25(0.50)
Error 5	RBVSS – BI	3.50(0.72)	0.04(1.83)	17.50(2.12)	0.45(1.43)	21.00(1.32)	0.49(1.85)
Lognormal(0,1)	RBVSS	3.50(0.73)	0.03(0.65)	15.23(0.25)	0.75(0.21)	18.73(0.83)	0.78(0.76)
	RBV-BI	2.04(1.41)	0.45(0.56)	10.50(0.55)	1.55(0.76)	12.50(0.55)	3.59(0.46)
	RBV	2.50(0.71)	0.50(0.71)	7.50(0.75)	0.55(0.72)	10.00(0.57)	1.05(0.78)
	BVSS-BI	3.05(0.84)	0.02(0.11)	17.96(0.32)	6.56(0.21)	21.01(0.84)	6.58(0.51)
	BVSS	3.20(0.45)	0.04(0.22)	13.60(0.89)	1.20(1.30)	16.80(1.14)	1.24(1.30)
	BV-BI	2.60(0.89)	0.05(0.35)	11.80(0.84)	1.09(0.71)	14.40(1.14)	1.14(0.71)
	BV	2.50(0.74)	0.07(0.45)	9.70(0.84)	1.11(0.71)	12.20(1.21)	1.18(0.71)

Table 2.2: Identification results for $n = 200, m = 60, p = 3$. *mean(sd)* of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 50th quantile.

$n = 200 \quad m = 60$		Main		Interaction		Total	
$\theta = 0.5$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.97(0.52)	0.04(0.22)	17.70(0.67)	0.04(0.17)	21.67(0.54)	0.08(0.11)
N(0,1)	RBVSS	3.85(0.43)	0.03(0.21)	15.22(0.27)	0.40(0.14)	19.07(0.76)	0.43(0.25)
	RBV-BI	3.40(0.55)	0.12(0.31)	11.80(0.84)	1.50(0.51)	15.20(1.10)	1.62(0.51)
	RBV	3.30(0.71)	0.06(0.32)	9.78(1.41)	0.80(0.89)	13.08(2.00)	0.86(0.89)
	BVSS-BI	3.75(0.73)	0.03(0.32)	17.40(0.35)	0.03(0.65)	21.15(0.75)	0.06(0.52)
	BVSS	3.65(0.85)	0.05(0.21)	15.21(0.31)	0.55(0.24)	18.86(0.34)	0.60(0.74)
	BV-BI	3.30(0.45)	0.02(0.15)	12.75(1.64)	1.55(0.54)	16.05(2.00)	1.57(0.76)
	BV	3.25(0.55)	0.04(0.25)	11.20(2.78)	0.56(0.67)	14.45(3.12)	0.60(0.87)
Error 2	RBVSS – BI	3.87(0.47)	0.02(0.21)	17.75(0.42)	0.03(0.35)	21.62(0.34)	0.05(0.45)
$t(2)$	RBVSS	3.76(0.51)	0.03(0.35)	15.14(0.21)	0.07(0.51)	18.90(0.56)	0.10(0.25)
	RBV-BI	2.80(1.79)	0.09(0.33)	11.50(3.24)	1.75(0.45)	14.30(2.44)	1.84(0.35)
	RBV	2.75(1.41)	0.08(0.34)	8.80(3.11)	0.76(0.45)	11.55(3.96)	0.84(0.67)
	BVSS-BI	3.55(0.55)	0.03(0.16)	17.86(0.31)	20.51(0.41)	21.41(0.55)	20.54(0.56)
	BVSS	3.50(1.14)	0.20(0.45)	13.55(4.16)	0.80(1.79)	17.05(4.64)	1.00(1.73)
	BV-BI	2.40(0.55)	0.35(0.65)	10.25(1.67)	1.27(0.45)	12.65(2.10)	1.62(0.45)
	BV	2.40(0.55)	0.27(0.76)	8.55(1.64)	0.20(0.45)	10.95(1.95)	0.47(0.58)
Error 3	RBVSS – BI	3.78(0.58)	0.07(0.47)	17.90(0.45)	0.03(0.33)	21.68(0.56)	0.10(0.47)
Laplace(0,2)	RBVSS	3.70(0.35)	0.08(0.52)	15.20(0.84)	0.05(0.29)	18.90(0.84)	0.13(0.34)
	RBV-BI	2.75(0.89)	0.15(0.35)	11.54(1.95)	0.57(0.26)	14.29(1.79)	0.72(0.58)
	RBV	2.65(1.14)	0.07(0.25)	9.75(1.82)	0.43(0.15)	12.40(1.48)	0.50(0.25)
	BVSS-BI	3.55(0.55)	0.07(0.43)	17.95(0.44)	19.31(0.22)	21.50(0.55)	19.38(0.23)
	BVSS	3.40(0.89)	0.05(0.21)	12.20(2.49)	0.32(0.45)	15.60(3.21)	0.37(0.55)
	BV-BI	2.65(0.84)	0.15(0.32)	10.60(1.34)	0.56(0.34)	13.25(1.95)	0.61(0.89)
	BV	2.55(0.84)	0.08(0.26)	8.85(1.10)	0.09(0.22)	11.40(1.14)	0.17(0.45)
Error 4	RBVSS – BI	3.95(0.21)	0.04(0.23)	17.95(0.31)	0.02(0.21)	21.95(0.52)	0.06(0.23)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.85(0.33)	0.03(0.27)	15.82(0.55)	0.03(0.24)	19.67(0.55)	0.06(0.14)
	RBV-BI	3.03(0.71)	0.10(0.45)	12.45(0.84)	1.57(0.67)	15.48(1.48)	1.67(0.67)
	RBV	2.80(1.10)	0.08(0.25)	9.57(1.23)	0.08(0.45)	12.37(1.64)	0.16(0.56)
	BVSS-BI	3.75(0.89)	0.02(0.45)	17.77(0.21)	20.15(0.33)	21.52(0.89)	20.17(0.45)
	BVSS	3.55(0.84)	0.03(0.25)	12.55(2.88)	0.03(0.00)	16.10(3.36)	0.06(0.15)
	BV-BI	2.55(0.71)	0.04(0.31)	11.25(1.14)	0.05(0.26)	13.80(1.14)	0.09(0.25)
	BV	2.45(0.55)	0.06(0.15)	9.50(1.14)	0.04(0.25)	11.95(1.10)	0.10(0.45)
Error 5	RBVSS – BI	3.90(0.42)	0.03(0.21)	17.95(0.21)	0.04(0.22)	21.85(0.45)	0.07(0.33)
Lognormal(0,1)	RBVSS	3.85(0.31)	0.04(0.35)	14.50(0.24)	0.05(0.28)	18.35(0.56)	0.09(0.23)
	RBV-BI	2.75(0.84)	0.05(0.31)	10.40(1.52)	0.45(0.23)	13.15(2.30)	0.50(0.56)
	RBV	2.70(0.45)	0.04(0.45)	8.50(2.31)	0.08(0.37)	11.20(2.68)	0.12(0.59)
	BVSS-BI	3.55(1.52)	0.05(0.45)	17.92(0.33)	18.24(0.31)	21.47(1.52)	18.29(0.52)
	BVSS	3.50(1.30)	0.05(0.45)	13.40(2.61)	0.60(1.34)	16.90(2.91)	0.65(1.35)
	BV-BI	2.80(0.84)	0.06(0.53)	10.25(2.30)	0.40(0.55)	13.05(2.17)	0.46(0.58)
	BV	2.57(0.71)	0.07(0.56)	8.55(2.07)	0.60(0.89)	11.12(2.30)	0.67(0.88)

Table 2.3: *Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept - slope model based on 100 replicates at the 50th quantile.*

$n = 200$	$m = 30$	$\theta = 0.5$	MAD	NMAD	TMAD
Error 1		RBVSS – BI	0.025(0.010)	0.001(0.003)	0.005(0.003)
N(0,1)		RBVSS	0.052(0.026)	0.003(0.001)	0.007(0.008)
		RBV-BI	0.085(0.030)	0.035(0.010)	0.046(0.013)
		RBV	0.158(0.048)	0.051(0.005)	0.062(0.012)
		BVSS-BI	0.045(0.005)	0.005(0.004)	0.012(0.003)
		BVSS	0.053(0.013)	0.006(0.005)	0.017(0.007)
		BV-BI	0.121(0.008)	0.058(0.001)	0.074(0.003)
		BV	0.165(0.009)	0.063(0.001)	0.081(0.008)
Error 2		RBVSS – BI	0.045(0.017)	0.002(0.003)	0.009(0.004)
$t(2)$		RBVSS	0.087(0.014)	0.004(0.005)	0.013(0.003)
		RBV-BI	0.123(0.027)	0.035(0.009)	0.045(0.015)
		RBV	0.136(0.032)	0.059(0.007)	0.057(0.015)
		BVSS-BI	0.071(0.024)	0.025(0.005)	0.045(0.004)
		BVSS	0.092(0.018)	0.035(0.001)	0.053(0.003)
		BV-BI	0.195(0.020)	0.136(0.009)	0.152(0.052)
		BV	0.223(0.019)	0.142(0.009)	0.163(0.011)
Error 3		RBVSS – BI	0.086(0.037)	0.001(0.011)	0.014(0.007)
Laplace(0,2)		RBVSS	0.126(0.035)	0.002(0.005)	0.034(0.005)
		RBV-BI	0.144(0.041)	0.047(0.012)	0.085(0.025)
		RBV	0.185(0.033)	0.057(0.007)	0.095(0.009)
		BVSS-BI	0.164(0.045)	0.024(0.015)	0.045(0.015)
		BVSS	0.182(0.098)	0.032(0.003)	0.052(0.019)
		BV-BI	0.195(0.011)	0.117(0.005)	0.128(0.004)
		BV	0.258(0.045)	0.145(0.004)	0.156(0.008)
Error 4		RBVSS – BI	0.075(0.024)	0.002(0.005)	0.025(0.005)
20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)		RBVSS	0.125(0.075)	0.002(0.003)	0.034(0.015)
		RBV-BI	0.142(0.027)	0.053(0.006)	0.057(0.007)
		RBV	0.157(0.041)	0.065(0.005)	0.077(0.008)
		BVSS-BI	0.145(0.036)	0.002(0.017)	0.056(0.019)
		BVSS	0.167(0.123)	0.004(0.003)	0.076(0.020)
		BV-BI	0.184(0.010)	0.108(0.003)	0.145(0.003)
		BV	0.195(0.011)	0.115(0.004)	0.154(0.004)
Error 5		RBVSS – BI	0.062(0.071)	0.003(0.005)	0.019(0.013)
Lognormal(0,1)		RBVSS	0.156(0.045)	0.005(0.011)	0.037(0.008)
		RBV-BI	0.164(0.036)	0.053(0.014)	0.087(0.017)
		RBV	0.176(0.057)	0.086(0.008)	0.095(0.011)
		BVSS-BI	0.145(0.057)	0.009(0.012)	0.027(0.036)
		BVSS	0.153(0.108)	0.012(0.024)	0.036(0.025)
		BV-BI	0.174(0.015)	0.117(0.035)	0.108(0.013)
		BV	0.189(0.023)	0.214(0.045)	0.125(0.023)

Table 2.4: *Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept - slope model based on 100 replicates at 50th quantile.*

$n = 200$	$p = 60$	$\theta = 0.5$	MAD	NMAD	TMAD
Error 1		RBVSS – BI	0.004(0.001)	0.001(0.003)	0.003(0.004)
N(0,1)		RBVSS	0.015(0.003)	0.004(0.001)	0.012(0.001)
		RBV-BI	0.185(0.023)	0.024(0.017)	0.038(0.027)
		RBV	0.238(0.033)	0.032(0.003)	0.044(0.005)
		BVSS-BI	0.008(0.004)	0.002(0.001)	0.007(0.001)
		BVSS	0.013(0.003)	0.003(0.002)	0.015(0.002)
		BV-BI	0.157(0.043)	0.023(0.008)	0.039(0.011)
		BV	0.174(0.031)	0.078(0.017)	0.047(0.015)
Error 2		RBVSS – BI	0.015(0.027)	0.005(0.002)	0.012(0.006)
$t(2)$		RBVSS	0.143(0.016)	0.024(0.002)	0.037(0.003)
		RBV-BI	0.154(0.025)	0.041(0.005)	0.047(0.008)
		RBV	0.198(0.041)	0.054(0.002)	0.058(0.004)
		BVSS-BI	0.045(0.038)	0.013(0.030)	0.025(0.014)
		BVSS	0.075(0.035)	0.029(0.014)	0.036(0.007)
		BV-BI	0.254(0.123)	0.115(0.123)	0.124(0.134)
		BV	0.298(0.131)	0.125(0.145)	0.135(0.132)
Error 3		RBVSS – BI	0.012(0.006)	0.002(0.001)	0.007(0.001)
Laplace(0,2)		RBVSS	0.023(0.005)	0.007(0.004)	0.012(0.004)
		RBV-BI	0.207(0.017)	0.028(0.001)	0.045(0.002)
		RBV	0.253(0.027)	0.043(0.003)	0.062(0.005)
		BVSS-BI	0.031(0.023)	0.012(0.007)	0.038(0.005)
		BVSS	0.054(0.025)	0.017(0.012)	0.019(0.012)
		BV-BI	0.245(0.024)	0.025(0.003)	0.056(0.004)
		BV	0.257(0.018)	0.034(0.003)	0.075(0.004)
Error 4		RBVSS – BI	0.013(0.009)	0.005(0.002)	0.006(0.001)
20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)		RBVSS	0.019(0.002)	0.007(0.006)	0.008(0.005)
		RBV-BI	0.198(0.032)	0.027(0.005)	0.048(0.004)
		RBV	0.232(0.036)	0.034(0.003)	0.053(0.006)
		BVSS-BI	0.043(0.029)	0.011(0.012)	0.027(0.013)
		BVSS	0.054(0.028)	0.013(0.006)	0.035(0.007)
		BV-BI	0.184(0.015)	0.048(0.004)	0.135(0.031)
		BV	0.231(0.036)	0.057(0.005)	0.175(0.007)
Error 5		RBVSS – BI	0.015(0.004)	0.004(0.002)	0.006(0.004)
Lognormal(0,1)		RBVSS	0.022(0.021)	0.005(0.003)	0.008(0.003)
		RBV-BI	0.197(0.074)	0.046(0.013)	0.059(0.016)
		RBV	0.281(0.077)	0.143(0.239)	0.082(0.224)
		BVSS-BI	0.023(0.037)	0.021(0.007)	0.028(0.010)
		BVSS	0.039(0.038)	0.035(0.010)	0.035(0.012)
		BV-BI	0.216(0.040)	0.052(0.008)	0.066(0.011)
		BV	0.314(0.056)	0.065(0.009)	0.156(0.031)

According to (Li et al. (2015)), the convergence of the MCMC is evaluated by the potential scale reduction factor (PSRF) (Gelman and Rubin (1992) Brooks and Gelman (1998)). The chains converge to the stationary distribution if the PSRF values close to 1. We use 1.1 as the benchmark of the convergence in our study according to (Gelman et al. (2004)). We compute the PSRF for the nonzero parameters and find the convergence of all chains after the burn-ins. Figure A.5 and Figure A.6 in the Appendix show the convergence of the proposed method for both models under Error 2. The Figures A.1, A.2, A.3 and A.4 represent the corresponding tables 3.3, 3.2, 2.3, 2.4, A.7, A.8, A.11, A.12, A.15, A.16, A.19 and A.20. The computational cost was presented in Table A.24. The sensitivity analysis represents in Table A.25 and Table A.26. The credibility of identification results can be found in Table A.27. The identification performance and estimation accuracy between the proposed method and intep are represented in Table A.28 and Table A.29.

2.4 Real Data Analysis

We analyzed the longitudinal lipidomics data generated from Dr. Wang’s lab. Details of the data collection and cleaning process been described in (King et al. (2015)). In the study, 60 female CD-1 mice have been randomly assigned with 4 different treatment, which are control (ad libitum feeding and sedentary), AE (exercise and ad libitum feeding), PE (exercise and pair feeding), and DCR (sedentary and 20% dietary calorie restriction). The mice’s body weight is the phenotype of interest. We measure the mice’s body weight every week for 11 weeks. In the control group, mice are sedentary and fed with ad libitum feeding as much as they want without doing treadmill exercises. In the AE group, mice are given ad libitum feeding with treadmill exercise every day at speed of 0.5 mph, 1 hour per day, and 5 days a week. In the PE group, mice are on treadmill exercise in the same way as AE group but with pair fed, which means they are given the same amount of food as the control mice. In the DCR group, mice are sedentary and fed with ad libitum with 20% less calorie than the control group but with the same level of protein, vitamins, and minerals. We have measured the composition of 176 plasma neutral lipid species and focus on diacylglycerols. We have

excluded diacylglycerol lipid species with most of samples below the detection limits so that the dataset contains 31 lipid features, leading to 93 lipid-treatment interactions.

In total, there are 8 methods under comparison. Our simulation suggests that methods with spike-and-slab priors generally outperform their counterpart without incorporating the priors. Therefore, we adopt these 4 methods, which are RBVSS-BI, RBVSS, BVSS-BI and BVSS to analyze the real data at the 50% quantile level. Table 2.5 shows the results for the analysis of the association of body weight with weight control treatment and lipids using the proposed RBVSS-BI under random intercept - slope model. The left column shows the phospholipid species, where the leading C denotes as carbon. For example, C18:2/16:1 represents that the phospholipid specie has two chains. One chain has 18 carbons with 2 unsaturated bonds and the other chain has 16 carbons with 1 unsaturated bond. In the analysis, control mice is selected as the reference group. Column AE, PE, DCR in the table represent their interaction effects. The proposed method identifies 4 lipid species that are significantly associated with body weight of the mice while having no interaction with treatments. The 4 lipid species are: C18:2/16:1, C16:1/16:0, C20:6/C16:0, C18:1/C20:4. The proposed method identifies 7 lipid species that have different association with the body weight of mice in exercise and/or caloric restricted mice comparing to that of control mice. Among them, C20:0/16:1 has a positive association with body weight in control mice, however it has negative interactions in all the other three groups. The results indicate that mice with an increased amount of C20:0/16:1 tend to have elevated body weights in control group, but lower body weights in AE, DCR or PE group. For AE mice, the largest positive interaction is observed in lipid specie C20:0/18:3, with regression coefficient 1.272. For DCR mice, all the interactions are negative except for C18:2/20:4. The interaction for C18:2/20:4 in PE mice is also positive, which is 0.982. It implies that C18:2/20:4 is positively associated with body weight in DCR and PE group. Lipid species that have positive interactions in PE mice also include C18:1/16:0, and C18:2/18:1. Species have negative interactions include C20:0/16:1, C20:0/18:3, C18:2/18:2, C18:1/18:2 for PE mice. We also conducted similar analysis using proposed method under the random intercept model. The selected lipids and their regression coefficient for main effects and interactions are summarized in

Table 2.6. Compared to random intercept - slope model, random intercept model selected lipids with more unsaturated bonds. For example, C20:3/16:1 has the strongest positive interactions for all three groups. C20:5/16:1 has a positive interaction in AE and PE group, but negative interaction in DCR group. Under both the random intercept - slope model and random intercept models, all the lipid we found have at least one chain of poly unsaturated fatty acids. Overall, the current findings confirm that weight control treatments via exercise and/or controlled dietary intake have interactions with phospholipid profile in the association with body weight. Especially, poly unsaturated fatty acids may play a critical role in this process as what we have observed previously (Ouyang et al. (2010)).

In the Appendix, Table A.30 – Table A.35 show the identification results from RBVSS (Table A.30–A.31), BVSS-BI (Table A.32–A.33, and BVSS (Table A.34–A.35), respectively. Selection of individual instead of groups of lipid–environment interactions by RBVSS and BVSS is not consistent with the group nature of E factors, and thus of much less important implications. Furthermore, it can be observed that in general, the magnitude of coefficients has more heterogeneity in non-robust methods. For example, in Table A.32, under the random intercept-slope model of BVSS-BI, the top 2 largest magnitude are 8.083 and 6.523, while the smallest one is merely 0.053.

The advantage of RBVSS-BI can be partially justified through evaluating prediction performance of all methods under comparison. Using the posterior median estimates, we have computed the mean absolute deviation (MAD) across all samples and time points according to model (2.1). Under random intercept - slope model , the proposed method has the smallest MAD of 2.67, compared to 3.84,18.27,19.75, which have been yielded from RBVSS, BVSS-BI and BVSS, accordingly. Besides, under the random intercept model, RBVSS-BI again has the smallest MAD of 2.74. The MAD calculated from the other three alternatives are 3.92,19.78 and 21.21, respectively.

Table 2.5: *Real data analysis result from RBVSS-BI under random intercept - slope model*

	Lipid	AE	DCR	PE
C18:2/16:1	-1.409	0.000	0.000	0.000
C20:0/16:1	0.138	-0.849	-0.265	-0.239
C16:1/16:0	-0.116	0.000	0.000	0.000
C18:1/16:0	0.000	0.578	-0.980	0.191
C20:6/16:0	-0.705	0.000	0.000	0.000
C20:0/18:3	-0.727	1.272	-0.156	-0.094
C18:2/18:2	0.000	-0.282	-0.762	-0.747
C18:1/18:2	-0.163	0.117	-0.220	-0.396
C18:2/18:1	0.000	0.561	-0.256	0.417
C18:2/20:4	0.000	-0.432	0.122	0.982
C18:1/20:4	0.845	0.000	0.000	0.000

Table 2.6: *Real data analysis result from RBVSS-BI under random intercept model*

	Lipid	AE	DCR	PE
C20:0/16:1	-1.029	0.000	0.000	0.000
C18:3/16:1	0.000	-1.698	-1.060	-5.429
C20:5/16:1	-0.568	1.214	-0.075	0.616
C20:6/16:0	0.060	0.000	0.000	0.000
C20:0/18:3	1.245	-1.174	-0.289	0.412
C18:2/18:2	0.000	-0.635	-0.220	-0.051
C18:1/18:2	-1.760	0.000	0.000	0.000
C20:3/16:1	0.305	3.259	1.816	0.527
C22:6/18:1	0.094	0.000	0.000	0.000
C18:2/20:4	0.000	0.134	-0.607	-0.074
C22:3/16:1	0.352	0.000	0.000	0.000

2.5 Discussion

(King et al. (2015)) has generated the longitudinal lipidomics data to conduct a classical one-way repeated measure ANOVA analysis. With high-dimensional lipid measurements, the goal of selecting important main and interaction effects in ANOVA problems (Yuan and Lin (2006b)) leads to the proposed Bayesian longitudinal $G \times E$ interaction study that accommodates robustness for heterogeneous repeated measures, intra-cluster correlations among longitudinal observations and structured sparsity of gene-environment interactions. Therefore, although the motivating data comes from a cancer prevention study using animal

models, the sparse Bayesian quantile mixed effect model developed in our study is broadly applicable to the longitudinal ANOVA with high-dimensional omics features.

The robustness of the proposed longitudinal Bayesian variable selection methods lies in the robust likelihood function based on the asymmetric Laplace distributions, which corresponds to the quantile check loss within the frequentist framework. The merit of the proposed method, as well as other Bayesian quantile variable selection methods over their non-robust counterpart have been fully revealed in the simulation study where four heavy-tailed errors have been used in the data generating model. The case study also suggests that the proposed method can outperform all the alternatives.

Interaction studies with a complex high-dimensional structure has been extensively studied, such as gene-gene ($G \times G$) and gene-environment ($G \times E$) interactions. Our study mainly focuses on the identification of the lipid-treatment interactions arising from the repeated measure ANOVA with high-dimensional genetic factors. In literature, longitudinal gene-gene interactions has not been widely examined ([Fan et al. \(2012b\)](#)). The proposed model can be further extended to handle the $G \times G$ interactions, although it is computationally more challenging due to the high dimensionality of the high order interactions among the main genetic effects.

Chapter 3

Bayesian Regularized Semiparametric Quantile Mixed Models for Longitudinal Studies.

3.1 Introduction

In genome-wide association studies (GWAS), millions of single nucleotide polymorphisms (SNPs) have been explored to identify important associations with complex disease traits such as case-control disease status, blood pressure, body mass index (BMI), among others. In classical GWAS, disease phenotypes are cross-sectional, which emphasizes phenotype measured at a single time point. The significance of associations can be established through hypothesis testing ([Wu et al., 2012](#); [Cantor et al., 2010](#)). The proposed study extends beyond this realm in the following aspects. First, we consider repeatedly measured phenotypes that are contaminated with outliers due to the heterogeneity of complex diseases. The heavy-tailed nature of the longitudinal disease phenotype demands models that are robust against outliers and can handle intra-cluster correlations among repeated measures. Second, appropriate modeling of nonlinear time-varying effects and structure identification when genetic contributions of SNPs to disease phenotypes are not time-varying or there are no genetic

effects at all. Third, valid statistical inferences on different genetic effects in robust models. These extensions allow us to gain a much deeper understanding of the genetic architecture under complex longitudinal disease traits.

Specifically, our work has been partially motivated by the Childhood Asthma Management Program (CAMP), which involves children aged 5 to 12 years who were diagnosed with chronic asthma (Group, 1999; Covar et al., 2012a). These children were monitored through follow-up visits over a period of 4 years. The disease trait of interest is the forced expiratory volume in one second (FEV1), which measures the amount of air one can expel in one second. Our analysis focuses on FEV1 measurements recorded through the 12 follow-up visits after applying treatment of nedocromil, budesonide, and Control. The left panel of Figure 3.1 shows that pooled repeated measures of FEV1 over all time points have a heavy-tailed distribution. Besides, we can observe a significant nonlinear time-varying effect of SNP XXX on the longitudinal trait, which should be properly accommodated. It is worth noting that not all SNPs have non-linear time-varying effects. They may have a main effect only or no genetic effect at all.

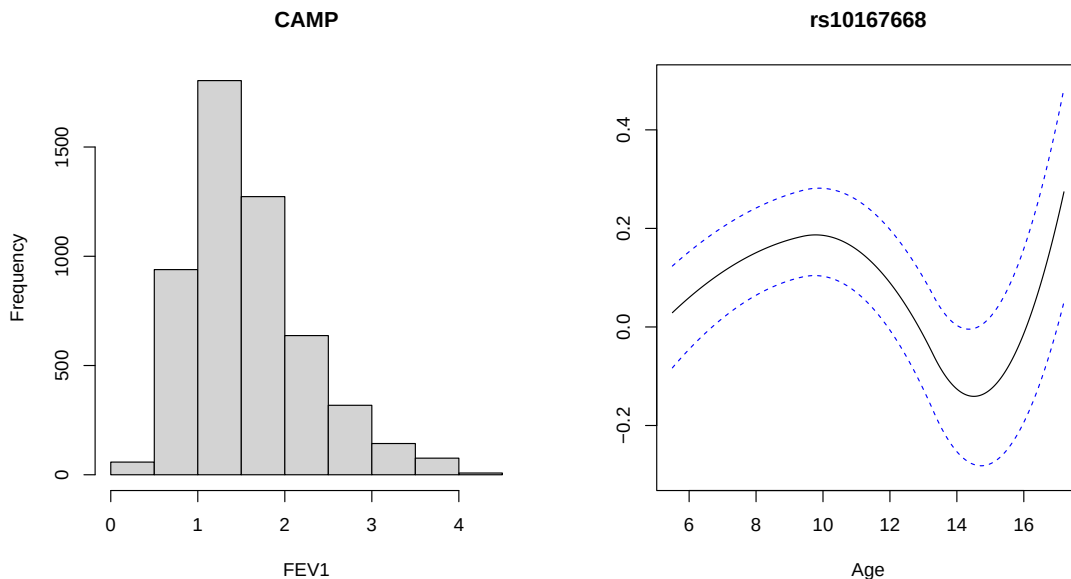


Figure 3.1: *Distribution of the FEV1 (left) and non-linear interaction effect of SNP rs10167668 (right) from the CAMP data. The blue dashed lines denote the 95% credible interval.*

In the literature, a broad spectrum of longitudinal variable selection models have been proposed, mostly from the generalized estimation equation (GEE) or mixed-effects model framework (Wang et al., 2012; Fan and Li, 2012; Zhou et al., 2019, 2022a). While the GEE is well known to be vulnerable to outliers in response variables, the mixed effect model is also not robust to data heterogeneity under the least square loss (Fan and Li, 2012). Although robust Bayesian mixed models have also been developed for longitudinal studies, they are usually low-dimensional and cannot achieve the selection of omics features. With nonparametric longitudinal models for nonlinear interactions, the regularized varying coefficient model proposed in Wang et al. (2008) and the corresponding quantile VC model in Tang et al. (2013) cannot achieve the separation of interaction and main effects. Moreover, both studies (Wang et al., 2008; Tang et al., 2013) assume independence among repeated measures, leaving the intra-cluster correlation unmodeled. Therefore, no published models can comprehensively take care of all the defining characteristics from CAMP data.

Here, we have developed Bayesian regularized semiparametric quantile mixed models for longitudinal GWAS data. We adopt the asymmetric Laplace distribution to specify the likelihood function, which has a natural correspondence with the frequentist quantile regression robust to heterogeneity in the response variable. To account for non-linear time-varying effects, we utilize varying coefficient modeling in the fixed effects. In particular, through basis expansion, we achieve structured modeling by separating the main and interaction effects. The Laplacian shrinkage and structured spike-and-slab priors have been imposed to select parametric and nonparametric fixed effects. To capture the intra-cluster correlations among repeated measures, we adopt a random intercept-slope model as random effects. With the Bayesian hierarchical model, we have developed Gibbs samplers and corresponding MCMC algorithms. The advantage of our proposed model has been convincingly demonstrated in extensive simulations in terms of identification, estimation, and statistical inference. In particular, our numerical results suggest that the incorporation of spike-and-slab priors facilitates valid statistical inference by producing Bayesian credible intervals with nominal coverage probabilities on parametric main effects and nonparametric nonlinear time-varying effects simultaneously. Applications of the proposed method on longitudinal CAMP data

shed novel insight on asthma etiology.

3.2 Data and Model settings

3.2.1 Partially Linear Varying Coefficient Model

Consider a longitudinal study on n subjects with J_i repeat measurement for each subject. let Y_{ij} be the phenotype for the i th subject at the j th time point t_{ij} . ($1 \leq i \leq n, 1 \leq j \leq J_i$). We use a m -dimensional indicator vectors $\mathbf{A}_{ij}$ and $\mathbf{D}_{ij}$ to denote the additive and dominant effects of SNPs for the i th subject at time point j , where $\mathbf{A}_{ij} = (A_{ij1}, \dots, A_{ijm})^\top$ and $\mathbf{D}_{ij} = (D_{ij1}, \dots, D_{ijm})^\top$. Also, $\mathbf{C}_{ij}$ represents low-dimensional clinical factor, where $\mathbf{C}_{ij} = (C_{ij1}, \dots, C_{ijq})^\top$. We include age for t_{ij} . Gender serves as clinical covariates, making $q = 1$. To model the effects of the additive and dominant effects of SNPs that are collectively associated with the longitudinal phenotype exhibiting heavy-tailed distributions, we propose the following quantile mixed-effects model at a given quantile level θ (where $0 < \theta < 1$):

$$Y_{ij} = \alpha_0(t_{ij}) + \sum_{c=1}^q \alpha_{1t,\theta}(t_{ij})C_{ijc} + \sum_{k=1}^m \beta_{1k,\theta}(t_{ij})A_{ijk} + \sum_{k=1}^m \beta_{2k,\theta}(t_{ij})D_{ijk} + \mathbf{Z}_{ij}^\top \boldsymbol{\zeta}_{i,\theta} + \epsilon_{ij,\theta}. \quad (3.1)$$

$$A_{ijk} = \begin{cases} 1, & \text{if the genotype of SNP } k \text{ is } AA, \\ 0, & \text{if the genotype of SNP } k \text{ is } Aa, \\ -1, & \text{if the genotype of SNP } k \text{ is } aa, \end{cases}$$

$$D_{ijk} = \begin{cases} 0.5, & \text{if the genotype of SNP } k \text{ is } Aa, \\ -0.5, & \text{if the genotype of SNP } k \text{ is } AA \text{ or } aa, \end{cases}$$

where the fixed effects include (a) the varying intercept $\alpha_0(t_{ij})$ and the varying coefficients $\alpha_{1t,\theta}(t_{ij})$ for the clinical factors $\mathbf{C}_{ij}$; (b) The coefficients $\beta_{1,\theta}(t_{ij})$ and $\beta_{2,\theta}(t_{ij})$ represent the varying effects of the additive and dominant components associated the design matrices

of additive and dominant effects, denoted by $\mathbf{A}_{ij}$ and $\mathbf{D}_{ij}$, respectively. The above model also incorporates the random effect $\boldsymbol{\zeta}_{i,\theta}$ with the corresponding design matrix $\mathbf{Z}_{ij}$. The $s \times 1$ vector $\mathbf{Z}_{ij}$ corresponds to the random intercept-slope model and the random intercept model when $s = 2$ and 1, respectively. The model errors $\epsilon_{ij,\theta}$ are independent with the θ th quantile being 0. For simplicity, we will omit the subscript “ θ ” from the regression coefficient vectors for both fixed and random effects moving forward. For example, the varying intercept, the varying coefficients for the clinical factors, the varying additive and dominant effects can be further expressed as $\alpha_0(t_{ij})$, $\boldsymbol{\alpha}_1(\mathbf{t}_{ij}) = (\alpha_{1,1}(t_{ij}), \dots, \alpha_{1,q}(t_{ij}))^\top$, $\boldsymbol{\beta}_1(\mathbf{t}_{ij}) = (\beta_{1,1}(t_{ij}), \dots, \beta_{1,m}(t_{ij}))^\top$ and $\boldsymbol{\beta}_2(\mathbf{t}_{ij}) = (\beta_{2,1}(t_{ij}), \dots, \beta_{2,m}(t_{ij}))^\top$.

3.2.2 The reparameterization on random effect

Consider the quantile linear mixed effect model is as follows:

$$Y_{ij} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\zeta}_i + \epsilon_{ij}, \boldsymbol{\zeta}_i \sim N(0, \boldsymbol{\Lambda}). \quad (3.2)$$

where $\mu_{ij} = \alpha_0(t_{ij}) + \sum_{c=1}^q \alpha_{1c,\theta}(t_{ij})C_{ijc} + \sum_{k=1}^m \beta_{1k,\theta}(t_{ij})A_{ijk} + \sum_{k=1}^m \beta_{2k,\theta}(t_{ij})D_{ijk}$ is the mean vector of the fixed effects, and $\boldsymbol{\zeta}_i$ is a $s \times 1$ vector of unobservable random effects with covariance matrix $\boldsymbol{\Lambda}$.

To ensure that $\boldsymbol{\Lambda}$ is positive semidefinite, we use the Cholesky decomposition, resulting in

$$\boldsymbol{\Lambda} = \boldsymbol{\Delta} \boldsymbol{\Delta}^\top,$$

where $\boldsymbol{\Delta}$ is a $s \times s$ lower triangular matrix with nonnegative diagonal entries. Similar to (Ibrahim et al. (2011)), the model (3.2) can be re-expressed as follows:

$$Y_{ij} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\Delta} \mathbf{b}_i + \epsilon_{ij} = \mu_{ij} + (\mathbf{b}_i^\top \bigotimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} + \epsilon_{ij}, \mathbf{b}_i \sim N(0, \mathbf{I}_s).$$

where $\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top$ is a Kronecker product between $\mathbf{b}_i^\top$ and $\mathbf{Z}_{ij}^\top$:

$$\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top = [b_{i1}\mathbf{Z}_{i1}^\top, b_{i1}\mathbf{Z}_{i2}^\top, \dots, b_{i1}\mathbf{Z}_{iJ_i}^\top, b_{i2}\mathbf{Z}_{i1}^\top, \dots, b_{is}\mathbf{Z}_{iJ_i}^\top]^\top.$$

Here $\boldsymbol{\delta}^\top = (\boldsymbol{\delta}_1^\top, \dots, \boldsymbol{\delta}_s^\top)$ is a $1 \times s(s+1)/2$ vector, where $\boldsymbol{\delta}_l$ consists of the non-zero lower triangular elements of the l th row of $\boldsymbol{\Delta}$, i.e., $\boldsymbol{\delta}_l = (\delta_{l1}, \dots, \delta_{ll})^\top, l = 1, \dots, s$, $\mathbf{J}_s$ is a transformation matrix with size $s^2 \times s(s+1)/2$ such that $\text{vec}(\boldsymbol{\Delta}) = \mathbf{J}_s \boldsymbol{\delta}$, $\mathbf{I}_s$ denotes the $s \times s$ identity matrix. By reparameterizing the linear predictor in this manner, selecting mixed effects becomes analogous to the issue of grouped variable selection in regression models with missing covariates. In this context, the design matrix of the random effects $(\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s$ can be viewed as missing covariates. In our case, $s = 2$, then

$$\boldsymbol{\Delta} = \begin{pmatrix} \delta_{11} & 0 \\ \delta_{21} & \delta_{22} \end{pmatrix}$$

the transformation matrix $\mathbf{J}_2$ would have the following form

$$\mathbf{J}_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

3.2.3 Basis expansion for structure identification of additive and dominant effects

As we discussed, it is crucial to differentiate between main-effect-only and nonlinear interaction, as incorrect specification of these effects can lead to overfitting. The following basis expansion is essential for distinguishing the various types of effects. Let N_n represent the number of uniform interior knots, and O denote the polynomial degree, then $L = N_n + O + 1$ (Schumaker (2007)) is the number of basis functions. The varying coefficient function $\beta_{1,k}(t_{ij})$

and $\beta_{2,k}(t_{ij})$ are approximated by the basis expansion, respectively.

$$\beta_{1,k}(t_{ij}) \approx \sum_{u=1}^L \tilde{B}_{1,ku}(t_{ij}) \tilde{\gamma}_{1,ku} = \tilde{\mathbf{B}}_{1,k}(\mathbf{t}_{ij})^\top \tilde{\boldsymbol{\gamma}}_{1,k}.$$

$$\beta_{2,k}(t_{ij}) \approx \sum_{u=1}^L \tilde{B}_{2,ku}(t_{ij}) \tilde{\gamma}_{2,ku} = \tilde{\mathbf{B}}_{2,k}(\mathbf{t}_{ij})^\top \tilde{\boldsymbol{\gamma}}_{2,k}.$$

where $\tilde{\mathbf{B}}_{1,k}(\mathbf{t}_{ij}) = (\tilde{B}_{1,k1}(t_{ij}), \dots, \tilde{B}_{1,kL}(t_{ij}))^\top$ and $\tilde{\mathbf{B}}_{2,k}(\mathbf{t}_{ij}) = (\tilde{B}_{2,k1}(t_{ij}), \dots, \tilde{B}_{2,kL}(t_{ij}))^\top$ are sets of normalized B spline basis, and $\tilde{\boldsymbol{\gamma}}_{1,k} = (\tilde{\gamma}_{1,k1}, \dots, \tilde{\gamma}_{1,kL})^\top$ and $\tilde{\boldsymbol{\gamma}}_{2,k} = (\tilde{\gamma}_{2,k1}, \dots, \tilde{\gamma}_{2,kL})^\top$ are the coefficient vectors. By changing of basis, the previously mentioned basis expansions are equivalent to

$$\beta_{1,k}(t_{ij}) \approx \sum_{u=1}^L \tilde{B}_{1,ku}(t_{ij}) \tilde{\gamma}_{1,ku} \doteq \gamma_{1,k1} + \tilde{\mathbf{B}}_{1,k*}(\mathbf{t}_{ij})^\top \boldsymbol{\gamma}_{1,k*}.$$

$$\beta_{2,k}(t_{ij}) \approx \sum_{u=1}^L \tilde{B}_{2,ku}(t_{ij}) \tilde{\gamma}_{2,ku} \doteq \gamma_{2,k1} + \tilde{\mathbf{B}}_{2,k*}(\mathbf{t}_{ij})^\top \boldsymbol{\gamma}_{2,k*}.$$

where $\tilde{\mathbf{B}}_{1,k*}(\mathbf{t}_{ij}) = (\tilde{B}_{1,k2}(t_{ij}), \dots, \tilde{B}_{1,kL}(t_{ij}))^\top$ and $\tilde{\mathbf{B}}_{2,k*}(\mathbf{t}_{ij}) = (\tilde{B}_{2,k2}(t_{ij}), \dots, \tilde{B}_{2,kL}(t_{ij}))^\top$. $\gamma_{1,k1}$ and $\boldsymbol{\gamma}_{1,k*} = (\gamma_{1,k2}, \dots, \gamma_{1,kL})^\top$ correspond to the constant and varying components of $\beta_{1,k}(t_{ij})$, respectively. In addition, $\gamma_{2,k1}$ and $\boldsymbol{\gamma}_{2,k*} = (\gamma_{2,k2}, \dots, \gamma_{2,kL})^\top$ correspond to the constant and varying components of $\beta_{2,k}(t_{ij})$, respectively.

3.2.4 Basis expansion for varying intercept and clinical effect

The varying coefficient function $\alpha_0(t_{ij})$ and $\alpha_{1,c}(t_{ij})$ are approximated by the basis expansion. Consider L be the number of basis functions, then the basis expansion can be expressed as:

$$\alpha_0(t_{ij}) \approx \sum_{u=1}^L \tilde{B}_{0u}(t_{ij}) \tilde{\gamma}_{0u} = \tilde{\mathbf{B}}_0(\mathbf{t}_{ij})^\top \tilde{\boldsymbol{\gamma}}_0.$$

$$\alpha_{1,c}(t_{ij}) \approx \sum_{u=1}^L \tilde{B}_{3,cu}(t_{ij}) \tilde{\gamma}_{3,cu} = \tilde{\mathbf{B}}_{3,c}(\mathbf{t}_{ij})^\top \tilde{\boldsymbol{\gamma}}_{3,c}.$$

where $\tilde{\mathbf{B}}_{0,u}(\mathbf{t}_{ij}) = (\tilde{B}_{0,1}(\mathbf{t}_{ij}), \dots, \tilde{B}_{0,L}(\mathbf{t}_{ij}))^\top$ and $\tilde{\mathbf{B}}_{3,c}(\mathbf{t}_{ij}) = (\tilde{B}_{3,c1}(\mathbf{t}_{ij}), \dots, \tilde{B}_{3,cL}(\mathbf{t}_{ij}))^\top$ are sets of normalized B spline basis. $\tilde{\boldsymbol{\gamma}}_0 = (\tilde{\gamma}_{01}, \dots, \tilde{\gamma}_{0L})^\top$ and $\tilde{\boldsymbol{\gamma}}_{3,c} = (\tilde{\gamma}_{3,c1}, \dots, \tilde{\gamma}_{3,cL})^\top$ are the coefficient vectors correspond to the varying components of $\alpha_0(\mathbf{t}_{ij})$ and $\alpha_{1,c}(\mathbf{t}_{ij})$.

3.2.5 The quantile mixed effects varying coefficient model

Define $\boldsymbol{\gamma}_{1,k} = (\gamma_{1,k1}, \boldsymbol{\gamma}_{1,k*}^\top)^\top$, $\boldsymbol{\gamma}_{2,k} = (\gamma_{2,k1}, \boldsymbol{\gamma}_{2,k*}^\top)^\top$, $\mathbf{B}_{1,k}(\mathbf{t}_{ij}) = (1, \tilde{\mathbf{B}}_{1,k*}(\mathbf{t}_{ij})^\top)^\top \doteq (B_{1,k1}(\mathbf{t}_{ij}), \dots, B_{1,kL}(\mathbf{t}_{ij}))^\top$, and $\mathbf{B}_{2,k}(\mathbf{t}_{ij}) = (1, \tilde{\mathbf{B}}_{2,k*}(\mathbf{t}_{ij})^\top)^\top \doteq (B_{2,k1}(\mathbf{t}_{ij}), \dots, B_{2,kL}(\mathbf{t}_{ij}))^\top$.

Then, model (3.1) can be represented as

$$\begin{aligned} Y_{ij} &= \tilde{\mathbf{B}}_0(\mathbf{t}_{ij})^\top \tilde{\boldsymbol{\gamma}}_0 + \sum_{c=1}^q \tilde{\mathbf{B}}_{3,c}(\mathbf{t}_{ij})^\top \tilde{\boldsymbol{\gamma}}_{3,c} C_{ijc} + \sum_{k=1}^m \mathbf{B}_{1,k}(\mathbf{t}_{ij})^\top \boldsymbol{\gamma}_{1,k} A_{ijk} + \sum_{k=1}^m \mathbf{B}_{2,k}(\mathbf{t}_{ij})^\top \boldsymbol{\gamma}_{2,k} D_{ijk} \\ &\quad + (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} + \epsilon_{ij} \\ &= \tilde{\mathbf{B}}_0(\mathbf{t}_{ij})^\top \tilde{\boldsymbol{\gamma}}_0 + \sum_{c=1}^q \tilde{\mathbf{B}}_{3,c}(\mathbf{t}_{ij})^\top \tilde{\boldsymbol{\gamma}}_{3,c} C_{ijc} + \sum_{k=1}^m (A_{ijk} \gamma_{1,k1} + \mathbf{W}_{ijk}^\top \boldsymbol{\gamma}_{1,k*}) + \sum_{k=1}^m (D_{ijk} \gamma_{2,k1} + \mathbf{V}_{ijk}^\top \boldsymbol{\gamma}_{2,k*}) \\ &\quad + (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} + \epsilon_{ij} \end{aligned} \quad (3.3)$$

where $\mathbf{W}_{ijk} = (B_{1,k2}(\mathbf{t}_{ij}) A_{ijk}, \dots, B_{1,kL}(\mathbf{t}_{ij}) A_{ijk})^\top$ and $\mathbf{V}_{ijk} = (B_{2,k2}(\mathbf{t}_{ij}) D_{ijk}, \dots, B_{2,kL}(\mathbf{t}_{ij}) D_{ijk})^\top$.

3.2.6 The Bayesian formulation of the quantile varying coefficient model with longitudinal traits

The histogram of repeatedly measured FEV1, shown in Figure 3.1, reveals a heavy-tailed distribution with outliers. This underscores the need for robust longitudinal methods. Recently, robust Bayesian analysis has proven to be a powerful approach for investigating G $\times$ E interactions in the presence of data heterogeneity associated with complex disease traits (Ren et al. (2023)). To achieve robustness in a Bayesian framework, it is essential to define the likelihood using a heavy-tailed distribution. In our approach, we assume that the error terms ϵ_{ij} are independently and identically distributed (i.i.d.) according to an asymmetric

Laplace distribution with the following density function

$$f(\epsilon_{ij}|\theta, \tau) = \theta(1 - \theta)\tau \exp\{-\tau\rho_\theta(\epsilon_{ij})\}, \quad i = 1, \dots, n, \quad j = 1, \dots, J_i,$$

where τ^{-1} is the scale parameter indicating the skewness of the asymmetric Laplace distribution (ALD), and $\rho_\theta(\cdot)$ is the check loss function with quantile level $\theta \in (0, 1)$: (Yu and Moyeed (2001), Yu and Zhang (2005))

$$\rho_\theta(t) = \begin{cases} \theta t, & \text{if } t \geq 0, \\ -(1 - \theta)t, & \text{if } t < 0, \end{cases}$$

which is crucial for the Bayesian formulation of quantile regression. The heavy-tailed nature of the Asymmetric Laplace Distribution (ALD) becomes evident when it simplifies to the well-known double exponential (or Laplace) distribution at the 50% quantile level.

Given the above probabilistic assumption on model error, by conditioning on the fixed and random effects, as well as the associated regression coefficients defined in model (3.2), the distribution of $Y_i(t_{ij})$ can be described as:

$$f(Y_{ij}|\mu_{ij}, \boldsymbol{\zeta}_i, \mathbf{Z}_{ij}, \tau) = \theta(1 - \theta)\tau \exp\{-\tau\rho_\theta(Y_{ij} - \mu_{ij} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta})\}.$$

where the mean function μ_{ij} is defined as $\mu_{ij} = \tilde{\mathbf{B}}_0(\mathbf{t}_{ij})^\top \tilde{\boldsymbol{\gamma}}_0 + \sum_{c=1}^q \tilde{\mathbf{B}}_{3,c}(\mathbf{t}_{ij})^\top \tilde{\boldsymbol{\gamma}}_{3,c} C_{ijc} + \sum_{k=1}^m (A_{ijk} \gamma_{1,k1} + \mathbf{W}_{ijk}^\top \boldsymbol{\gamma}_{1,k*}) + \sum_{k=1}^m (D_{ijk} \gamma_{2,k1} + \mathbf{V}_{ijk}^\top \boldsymbol{\gamma}_{2,k*})$.

(Kozumi and Kobayashi (2011)) presented an equivalent formulation of the skewed Laplace distribution as a combination of an exponential distribution and a scaled normal distribution. This formulation aids in establishing a Bayesian hierarchical model for quantile regression. Specifically, with $\xi_1 = \frac{1-2\theta}{\theta(1-\theta)}$ and $\xi_2 = \sqrt{\frac{2}{\theta(1-\theta)}}$ at quantile level $0 < \theta < 1$, the error term ϵ_{ij} follows the ALD with a scale parameter τ^{-1} as

$$\epsilon_{ij} = \tau^{-1} \xi_1 v_{ij} + \tau^{-1} \xi_2 \sqrt{v_{ij}} z_{ij},$$

where v_{ij} is the standard exponential distribution, $\text{Exp}(1)$, and z_{ij} is a standard normal random variable. Subsequently, model (3.2) can be rewritten as:

$$Y_{ij} = \mu_{ij} + (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} + \tau^{-1} \xi_1 v_{ij} + \tau^{-1} \xi_2 \sqrt{v_{ij}} z_{ij},$$

Let $\tilde{v}_{ij} = \tau^{-1} v_{ij}$, then $\tilde{v}_{ij} \stackrel{\text{ind}}{\sim} \text{Exp}(\tau^{-1})$ for $i = 1, \dots, n$ and $j = 1, \dots, J_i$. The Bayesian hierarchal model has the following representation :

$$\begin{aligned} Y_{ij} &= \mu_{ij} + (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} + \xi_1 \tilde{v}_{ij} + \tau^{-1/2} \xi_2 \sqrt{\tilde{v}_{ij}} z_{ij}, \\ \tilde{v}_{ij} | \tau &\stackrel{\text{ind}}{\sim} \tau \exp(-\tau \tilde{v}_{ij}), \\ z_{ij} &\stackrel{\text{ind}}{\sim} \text{N}(0, 1). \end{aligned}$$

3.2.7 The Bayesian regularized quantile mixed effects model for longitudinal study

Because of the “large p , small n ” setting for the proposed semi-parametric model, variable selection method needs to be adopted. Also, due to the heavy-tailed distribution for the response variables, quantile Bayesian variable selection method needs to be accomodated. The loss function can be defined as following:

$$L(\mu_{ij}, \boldsymbol{\delta}) = \sum_{i=1}^n \sum_{j=1}^{J_i} \rho_\theta(Y_{ij}) - \mu_{ij} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta}$$

The coresponding penalized loss function can be written as

$$L(\mu_{ij}, \boldsymbol{\delta}) + \lambda_1 \sum_{k=1}^m |\gamma_{1,k1}| + \lambda_2 \sum_{k=1}^m \|\gamma_{1,k*}\|_2 + \lambda_3 \sum_{k=1}^m |\gamma_{2,k1}| + \lambda_4 \sum_{k=1}^m \|\gamma_{2,k*}\|_2 + \lambda_5 \sum_{l=1}^s \|\boldsymbol{\delta}_l\|_2. \quad (3.4)$$

where $\gamma_{1,k1}$ is the main additive effect for the k th SNP and the l_2 norm of the spline coefficients $\|\gamma_{1,k*}\|_2$ is corresponding to the varying parts of $\beta_{1,k}(t_{ij})$. If $\|\gamma_{1,k*}\|_2 = 0$, then the time - varying additive effect of the k th SNP is identically zero. Furthermore, if $\gamma_{1,k1} = 0$, then A_{ijk}

has no main additive effect and is not associated with the phenotype. Likewise, $\gamma_{2,k1}$ is the main effect for the dominant effect for the k th SNP and the l_2 norm of the spline coefficients $\|\gamma_{2,k*}\|_2$ is corresponding to the varying parts of $\beta_{2,k}(t_{ij})$. If $\|\gamma_{2,k*}\|_2 = 0$, then the time-varying dominant effect of the k th SNP is identically zero. Furthermore, if $\gamma_{2,k1} = 0$, then D_{ijk} has no main dominant effect and is not associated with the phenotype. The l_2 norm of $\|\delta_l\|_2$ corresponding to the group level regularization on the random effect.

Because of the low dimensionality of the clinical factors, we are only interested in applying robust Bayesian variable selection on both the individual level of $\gamma_{1,k1}$, $\gamma_{2,k1}$ and the group level of $\gamma_{1,k*}, \gamma_{2,k*} (k = 1, \dots, m)$, and δ_l .

Selecting the fixed effects and random effects with Laplacian shrinkage

To determine the main additive effects linked with the phenotype, we apply an independent and identical Laplace prior to $\gamma_{1,k1}$ as follows:

$$\pi(\gamma_{1,11}, \dots, \gamma_{1,m1} | \tau, \lambda_1) = \prod_{k=1}^m \frac{\tau \lambda_1}{2} \exp(-\tau \lambda_1 |\gamma_{1,k1}|).$$

Meanwhile, to identify the varying additive effects, we employ group-level shrinkage by applying a multivariate Laplace prior to $\gamma_{1,k*}$:

$$\pi(\gamma_{1,k*} | \tau, \lambda_2) = C_p (\tau \lambda_2)^p \exp(-\tau \lambda_2 \|\gamma_{1,k*}\|_2).$$

where $p = L - 1$, $C_p = 2^{-(p+1)/2} (2\pi)^{-(p-1)/2} / \Gamma((p+1)/2)$ with gamma function $\Gamma(\cdot)$, and p is the group size.

Similarly, to identify the main dominant effects associated with the phenotype, we impose an independent and identical Laplace prior on $\gamma_{2,k1}$ as follows:

$$\pi(\gamma_{2,11}, \dots, \gamma_{2,m1} | \tau, \lambda_3) = \prod_{k=1}^m \frac{\tau \lambda_3}{2} \exp(-\tau \lambda_3 |\gamma_{2,k1}|).$$

Meanwhile, to identify the varying additive effects, we use group-level shrinkage by imposing

a multivariate Laplace prior on $\gamma_{2,k*}$:

$$\pi(\gamma_{2,k*}|\tau, \lambda_4) = C_p(\tau\lambda_4)^p \exp(-\tau\lambda_4 \|\gamma_{2,k*}\|_2).$$

where $p = L - 1$, $C_p = 2^{-(p+1)/2} (2\pi)^{-(p-1)/2} / \Gamma((p+1)/2)$ with gamma function $\Gamma(\cdot)$, and p is the group size.

The Bayesian regularization on the random effects can be done by imposing multivariate Laplace prior on δ_l :

$$\pi(\delta_s|\tau, \lambda_5) = C_s(\tau\lambda_5)^s \exp(-\tau\lambda_5 \|\delta_s\|_2).$$

The joint posterior distribution of fixed effects and random effects is given by:

$$\begin{aligned} & f(\gamma_{1,k1}, \gamma_{2,k1}, \gamma_{1,k*}, \gamma_{2,k*}, \mathbf{Z}_{ij}^\top, \delta|Y_{ij}, \mu_{ij}, \tau, \lambda_1, \lambda_2, \lambda_3, \lambda_4, \lambda_5) \\ & \propto \exp \left\{ -\tau \sum_{i=1}^n \sum_{j=1}^{J_i} \rho_\theta(Y_{ij} - \mu_{ij} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \delta) - \tau\lambda_1 \sum_{k=1}^m |\gamma_{1,k1}| \right. \\ & \quad \left. - \tau\lambda_3 \sum_{k=1}^m |\gamma_{2,k1}| - \tau\lambda_2 \sum_{k=1}^m \|\gamma_{1,k*}\|_2 - \tau\lambda_4 \sum_{k=1}^m \|\gamma_{2,k*}\|_2 - \tau\lambda_5 \sum_{l=1}^s \|\delta_l\|_2 \right\}. \end{aligned}$$

where the power of the exponential kernel aligns with the objective function in Equation 3.4. We set up the Bayesian hierarchical model by defining η_1 as $\tau\lambda_1$. Consequently, the i.i.d. Laplace priors on $\gamma_{1,k1}$, which represent the main additive effects, can be expressed as follows.

$$\pi(\gamma_{1,11}, \dots, \gamma_{1,m1}|\tau, \lambda_1) = \prod_{k=1}^m \frac{\eta_1}{2} \exp\{-\eta_1 |\gamma_{1,k1}|\} = \prod_{k=1}^m \int_0^\infty \frac{1}{\sqrt{2\pi s_{1,k}}} \exp(-\frac{\gamma_{1,k1}^2}{2s_{1,k}}) \frac{\eta_1^2}{2} \exp(-\frac{\eta_1^2}{2} s_{1,k}) ds_{1,k}.$$

The aforementioned Laplacian priors can be hierarchically represented as a scale mixture of normals once $s_{1,k}$ is integrated out:

$$\begin{aligned} \pi(\gamma_{1,k1}|s_{1,k}) & \stackrel{ind}{\sim} N(0, s_{1,k}), \\ s_{1,k} & \stackrel{ind}{\sim} \frac{\eta_1^2}{2} \exp\{-\frac{\eta_1^2}{2} s_{1,k}\}. \end{aligned} \tag{3.5}$$

Similarly, we set up the Bayesian hierarchical model by defining η_3 as $\tau\lambda_3$. Consequently, the i.i.d. Laplace priors on $\gamma_{2,k1}$, which represent the main dominant effects, can be expressed as follows.

$$\pi(\gamma_{2,11}, \dots, \gamma_{2,m1} | \tau, \lambda_3) = \prod_{k=1}^m \frac{\eta_3}{2} \exp\{-\eta_3 |\gamma_{2,k1}|\} = \prod_{k=1}^m \int_0^\infty \frac{1}{\sqrt{2\pi s_{3,k}}} \exp\left(-\frac{\gamma_{2,k1}^2}{2s_{3,k}}\right) \frac{\eta_3^2}{2} \exp\left(-\frac{\eta_3^2}{2} s_{3,k}\right) ds_{3,k}.$$

Then the above Laplacian priors can be hierarchically expressed as scale mixture of normals after $s_{3,k}$ is integrated out:

$$\begin{aligned} \pi(\gamma_{2,k1} | s_{3,k}) &\stackrel{ind}{\sim} N(0, s_{3,k}), \\ s_{3,k} &\stackrel{ind}{\sim} \frac{\eta_3^2}{2} \exp\left\{-\frac{\eta_3^2}{2} s_{3,k}\right\}. \end{aligned} \tag{3.6}$$

Additionally, using the reparameterization $\eta_2 = \tau\lambda_2$, the multivariate Laplace prior on $\gamma_{1,k*}$ can be expressed as:

$$\begin{aligned} \pi(\gamma_{1,k*} | \eta_2) &\triangleq \text{M-Laplace}(\gamma_{1,k*} | \eta_2) \\ &= C_p \eta_2^p \exp(-\eta_2 \|\gamma_{1,k*}\|_2) \\ &= \int_0^\infty (2\pi s_{2,k})^{-p/2} \exp\left\{-\frac{1}{2s_{2,k}} \gamma_{1,k*}^\top \gamma_{1,k*}\right\} \times \frac{(\eta_2^2/2)^{(p+1)/2}}{\Gamma(\frac{p+1}{2})} s_{2,k}^{(p-1)/2} \exp\left(-\frac{\eta_2^2}{2} s_{2,k}\right) ds_{2,k} \end{aligned}$$

Then, the prior of $\gamma_{1,k*}$ can be expressed as the hierarchical representation below:

$$\begin{aligned} \pi(\gamma_{1,k*} | s_{2,k}) &\stackrel{ind}{\sim} N_p(0, s_{2,k} \mathbf{I}_p), \\ s_{2,k} &\stackrel{ind}{\sim} \text{Gamma}\left(\frac{p+1}{2}, \frac{\eta_2^2}{2}\right), \end{aligned} \tag{3.7}$$

where $\frac{p+1}{2}$ is the shape parameter and $\frac{\eta_2^2}{2}$ is the rate parameter of the gamma distribution.

Besides, with the reparametrization $\eta_4 = \tau\lambda_4$, the multivariate Laplace prior on $\gamma_{2,k*}$

can be written as:

$$\begin{aligned}
\pi(\boldsymbol{\gamma}_{2,k*}|\eta_4) &\triangleq \text{M-Laplace}(\boldsymbol{\gamma}_{2,k*}|\eta_4) \\
&= C_p \eta_4^p \exp(-\eta_4 \|\boldsymbol{\gamma}_{2,k*}\|_2) \\
&= \int_0^\infty (2\pi s_{4,k})^{-p/2} \exp\left\{-\frac{1}{2s_{4,k}} \boldsymbol{\gamma}_{2,k*}^\top \boldsymbol{\gamma}_{2,k*}\right\} \times \frac{(\eta_4^2/2)^{(p+1)/2}}{\Gamma(\frac{p+1}{2})} s_{4,k}^{(p-1)/2} \exp\left(-\frac{\eta_4^2}{2} s_{4,k}\right) ds_{4,k}.
\end{aligned}$$

Then, the prior of $\boldsymbol{\gamma}_{2,k*}$ can be expressed as the hierarchical representation below:

$$\begin{aligned}
\pi(\boldsymbol{\gamma}_{2,k*}|s_{4,k}) &\stackrel{ind}{\sim} N_p(0, s_{4,k} \mathbf{I}_p), \\
s_{4,k} &\stackrel{ind}{\sim} \text{Gamma}\left(\frac{p+1}{2}, \frac{\eta_4^2}{2}\right),
\end{aligned} \tag{3.8}$$

where $\frac{p+1}{2}$ is the shape parameter and $\frac{\eta_4^2}{2}$ is the rate parameter of the gamma distribution.

Similarly, using the reparameterization $\eta_5 = \tau \lambda_5$, the multivariate Laplace prior on $\boldsymbol{\delta}_l$ can be expressed as:

$$\begin{aligned}
\pi(\boldsymbol{\delta}_l|\eta_5) &\triangleq \text{M-Laplace}(\boldsymbol{\delta}_l|\eta_5) \\
&= C_s \eta_5^s \exp(-\eta_5 \|\boldsymbol{\delta}_l\|_2) \\
&= \int_0^\infty (2\pi s_{5l})^{-s/2} \exp\left\{-\frac{1}{2s_{5l}} \boldsymbol{\delta}_l^\top \boldsymbol{\delta}_l\right\} \times \frac{(\eta_5^2/2)^{(s+1)/2}}{\Gamma(\frac{s+1}{2})} s_{5l}^{(s-1)/2} \exp\left(-\frac{\eta_5^2}{2} s_{5l}\right) ds_{5l}.
\end{aligned}$$

Then, the prior of $\boldsymbol{\delta}_l$ can be expressed as the hierarchical representation below:

$$\begin{aligned}
\pi(\boldsymbol{\delta}_l|s_{5l}) &\stackrel{ind}{\sim} N_s(0, s_{5l} \mathbf{I}_s), \\
s_{5l} &\stackrel{ind}{\sim} \text{Gamma}\left(\frac{s+1}{2}, \frac{\eta_5^2}{2}\right),
\end{aligned} \tag{3.9}$$

where $\frac{s+1}{2}$ is the shape parameter and $\frac{\eta_5^2}{2}$ is the rate parameter of the gamma distribution.

Selection of fixed and random effects with structured spike-and-slab priors

In the Bayesian hierarchical model, the prior formulations of univariate and multivariate Laplace priors in (3.5), (3.6), (3.7), (3.8) and (3.9) result in shrinkage estimation for identifying fixed effects and random effects, respectively. However, a known limitation of Laplace

priors is that they do not yield exact zero estimates. To achieve precise sparsity, we enhance the prior formulation by integrating spike-and-slab priors.

To identify varying additive effects, we have

$$\begin{aligned}\boldsymbol{\gamma}_{1,k*}|\phi_{2,k}, s_{2,k} &\stackrel{ind}{\sim} \phi_{2,k}\mathbf{N}_p(\mathbf{0}, \text{diag}(s_{2,k}, \dots, s_{2,k})) + (1 - \phi_{2,k})\delta_0(\boldsymbol{\gamma}_{1,k*}), \\ \phi_{2,k} &\stackrel{ind}{\sim} \text{Bernoulli}(\pi_2), \\ s_{2,k} &\stackrel{ind}{\sim} \text{Gamma}\left(\frac{p+1}{2}, \frac{\eta_2^2}{2}\right).\end{aligned}\tag{3.10}$$

To identify varying dominant effects, we have

$$\begin{aligned}\boldsymbol{\gamma}_{2,k*}|\phi_{4,k}, s_{4,k} &\stackrel{ind}{\sim} \phi_{4,k}\mathbf{N}_p(\mathbf{0}, \text{diag}(s_{4,k}, \dots, s_{4,k})) + (1 - \phi_{4,k})\delta_0(\boldsymbol{\gamma}_{2,k*}), \\ \phi_{4,k} &\stackrel{ind}{\sim} \text{Bernoulli}(\pi_4), \\ s_{4,k} &\stackrel{ind}{\sim} \text{Gamma}\left(\frac{p+1}{2}, \frac{\eta_4^2}{2}\right).\end{aligned}\tag{3.11}$$

To identify the random effects $\boldsymbol{\delta}_l$, we have

$$\begin{aligned}\boldsymbol{\delta}_l|\phi_{5l}, s_{5l} &\stackrel{ind}{\sim} \phi_{5l}\mathbf{N}(\mathbf{0}, \text{diag}(s_{5l}, \dots, s_{5l})) + (1 - \phi_{5l})\delta_0(\boldsymbol{\delta}_l), \\ \phi_{5l} &\stackrel{ind}{\sim} \text{Bernoulli}(\pi_5), \\ s_{5l} &\stackrel{ind}{\sim} \text{Gamma}\left(\frac{s+1}{2}, \frac{\eta_5^2}{2}\right).\end{aligned}\tag{3.12}$$

The slab component $\mathbf{N}_p(\mathbf{0}, s_{2,k}\mathbf{I}_p)$, $\mathbf{N}_p(\mathbf{0}, s_{4,k}\mathbf{I}_p)$ and $\mathbf{N}_s(\mathbf{0}, s_{5l}\mathbf{I}_s)$ model the large and diffuse signals, while the spike distribution $\delta_0(\boldsymbol{\gamma}_{1,k*})$, $\delta_0(\boldsymbol{\gamma}_{2,k*})$ and $\delta_0(\boldsymbol{\delta}_l)$ represent the point mass at $\mathbf{0}_{p \times 1}$ and $\mathbf{0}_{s \times 1}$, with parameter $\pi_2 \in [0, 1]$, $\pi_4 \in [0, 1]$ and $\pi_5 \in [0, 1]$. The $\phi_{2,k}$, $\phi_{4,k}$, for $1 \leq k \leq m$, and ϕ_{5l} , for $1 \leq l \leq s$ are latent binary indicator variables for fixed effects and random effects, indicating whether the corresponding fixed and random effects are selected. When $\phi_{2,k} = 0$, $\phi_{4,k} = 0$ and $\phi_{5l} = 0$, the coefficient vector $\boldsymbol{\gamma}_{1,k*}$, $\boldsymbol{\gamma}_{2,k*}$ and $\boldsymbol{\delta}_l$ have a point mass density at zero, implying that there is fixed and random effect. Conversely, if $\phi_{2,k} = 1$, $\phi_{4,k} = 1$ and $\phi_{5l} = 1$, the spike-and-slab prior in (3.10), (3.11) and (3.12) reduce to the multivariate Laplace prior in (3.7), (3.8), and (3.9) resulting in the Bayesian quantile group

LASSO, indicating that the fixed and random effects are associated with the longitudinal phenotypic traits. Furthermore, the prior formulation in (3.10), (3.11) and (3.12) can be interpreted as the following multivariate spike-and-slab priors after integrating out $\phi_{2,k}$, $s_{2,k}$, $\phi_{4,k}$, $s_{4,k}$, ϕ_{5l} and s_{5l} :

$$\gamma_{1,k*}|\eta_2 \stackrel{ind}{\sim} \phi_{2,k} \text{M-Laplace}(\gamma_{1,k*}|\eta_2) + (1 - \phi_{2,k})\delta_0(\gamma_{1,k*}),$$

$$\gamma_{2,k*}|\eta_4 \stackrel{ind}{\sim} \phi_{4,k} \text{M-Laplace}(\gamma_{2,k*}|\eta_4) + (1 - \phi_{4,k})\delta_0(\gamma_{2,k*}),$$

$$\delta_l|\eta_5 \stackrel{ind}{\sim} \phi_{5l} \text{M-Laplace}(\delta_l|\eta_5) + (1 - \phi_{5l})\delta_0(\delta_l),$$

where the parameter $\pi_2 \in [0, 1]$, $\pi_4 \in [0, 1]$ and $\pi_5 \in [0, 1]$.

For identification of main additive and dominant effect, we impose the univariate spike-and-slab priors for $\gamma_{1,k1}$ and $\gamma_{2,k1}$ $k = (1, \dots, m)$ as follows:

$$\begin{aligned} \gamma_{1,k1}|\phi_{1,k}, s_{1,k} &\stackrel{ind}{\sim} \phi_{1,k} \text{N}(0, s_{1,k}) + (1 - \phi_{1,k})\delta_0(\gamma_{1,k1}), \\ \phi_{1,k} &\stackrel{ind}{\sim} \text{Bernoulli}(\pi_1), \\ s_{1,k} &\stackrel{ind}{\sim} \text{Gamma}(1, \frac{\eta_1^2}{2}). \end{aligned} \tag{3.13}$$

$$\begin{aligned} \gamma_{2,k1}|\phi_{3,k}, s_{3,k} &\stackrel{ind}{\sim} \phi_{3,k} \text{N}(0, s_{3,k}) + (1 - \phi_{3,k})\delta_0(\gamma_{2,k1}), \\ \phi_{3,k} &\stackrel{ind}{\sim} \text{Bernoulli}(\pi_3), \\ s_{3,k} &\stackrel{ind}{\sim} \text{Gamma}(1, \frac{\eta_3^2}{2}). \end{aligned} \tag{3.14}$$

where the spike component $\delta_0(\gamma_{1,k1})$ and $\delta_0(\gamma_{2,k1})$ are point mass at 0. $\phi_{1,k}$ and $\phi_{3,k}$ determine whether the spike or slab component is in action, with $\pi_1 \in [0, 1]$ and $\pi_3 \in [0, 1]$. The prior specified in (3.13) and (3.14) are equivalent to the univariate Laplace prior in (3.5) and (3.6) when $\phi_{1,k} = 1$ and $\phi_{3,k} = 1$, indicating that the k th SNP has a main additive or dominant effect. Otherwise, when $\phi_{1,k} = 0$ and $\phi_{3,k} = 0$, the k th SNP does not have a main additive or dominant effect.

We assign conjugate beta priors $\pi_1 \sim \text{Beta}(r_1, w_1)$ and $\pi_2 \sim \text{Beta}(r_2, w_2)$, $\pi_3 \sim \text{Beta}(r_3, w_3)$,

$\pi_4 \sim \text{Beta}(r_4, w_4)$ and $\pi_5 \sim \text{Beta}(r_5, w_5)$ and gamma priors $\eta_1^2 \sim \text{Gamma}(a_1, b_1)$ and $\eta_2^2 \sim \text{Gamma}(a_2, b_2)$, gamma priors $\eta_3^2 \sim \text{Gamma}(a_3, b_3)$, $\eta_4^2 \sim \text{Gamma}(a_4, b_4)$ and $\eta_5^2 \sim \text{Gamma}(a_5, b_5)$. $\tau \sim \text{Gamma}(c, d)$. In particular, we set all the hyperparameters to 1.

3.2.8 Gibbs Sampler

The joint posterior distribution can be expressed as

$$\begin{aligned}
& \pi(\tilde{\gamma}_0, \tilde{\gamma}_3, \gamma_{1,k1}, \gamma_{1,k*}, \gamma_{2,k1}, \gamma_{2,k*}, \boldsymbol{\delta}, s_{1,k}, s_{2,k}, s_{3,k}, s_{4,k}, s_{5,l}, \eta_1^2, \eta_2^2, \eta_3^2, \eta_4^2, \eta_5^2, \pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \tau | Y_{ij}) \\
& \propto (\tau^{-1})^{-\frac{\sum_{i=1}^n J_i}{2}} \left(\prod_{i=1}^n \prod_{j=1}^{J_i} \tilde{v}_{ij} \right)^{-\frac{1}{2}} \\
& \exp \left\{ -\frac{1}{2\tau^{-1}\xi_2^2} \sum_{i=1}^n \sum_{j=1}^{J_i} \left(\frac{Y_{ij} - \mu_{ij} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}}{\sqrt{\tilde{v}_{ij}}} \right)^2 \right\} \\
& \times \exp \left(-\frac{1}{2} \tilde{\gamma}_0^\top \boldsymbol{\Sigma}_0^{-1} \tilde{\gamma}_0 \right) \times \prod_{c=1}^q \exp \left(-\frac{1}{2} \tilde{\gamma}_3^\top \boldsymbol{\Sigma}^{-1} \tilde{\gamma}_3 \right) \\
& \times \prod_{k=1}^m \left((1 - \pi_1) (2\pi s_{1,k})^{-\frac{1}{2}} \exp \left(-\frac{1}{2s_{1,k}} \gamma_{1,k1}^2 \right) \mathbf{I}_{\{\gamma_{1,k1} \neq 0\}} + \pi_1 \delta_0(\gamma_{1,k1}) \right) \\
& \times \prod_{k=1}^m \left((1 - \pi_3) (2\pi s_{3,k})^{-\frac{1}{2}} \exp \left(-\frac{1}{2s_{3,k}} \gamma_{2,k1}^2 \right) \mathbf{I}_{\{\gamma_{2,k1} \neq 0\}} + \pi_3 \delta_0(\gamma_{2,k1}) \right) \\
& \times \tau^{\sum_{i=1}^n J_i} \exp \left(-\tau \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij} \right) \\
& \times \tau^{c-1} \exp(-d\tau) \\
& \times (\eta_1^2)^{a_1-1} \exp(-b_1 \eta_1^2) \prod_{k=1}^m \frac{\eta_1^2}{2} \exp \left(-\frac{\eta_1^2}{2} s_{1k} \right) \\
& \times (\eta_2^2)^{a_2-1} \exp(-b_2 \eta_2^2) \prod_{k=1}^m \frac{\eta_2^2}{2} \exp \left(-\frac{\eta_2^2}{2} s_{2k} \right) \\
& \times \pi_1^{r_1-1} (1 - \pi_1)^{w_1-1} \\
& \times \pi_2^{r_2-1} (1 - \pi_2)^{w_2-1} \\
& \times \prod_{k=1}^m \left((1 - \pi_2) (2\pi s_{2,k})^{-\frac{p}{2}} \exp \left(-\frac{1}{2s_{2,k}} \gamma_{1,k*}^\top \gamma_{1,k*} \right) \mathbf{I}_{\{\gamma_{1,k*} \neq \mathbf{0}\}} + \pi_2 \delta_0(\gamma_{1,k*}) \right)
\end{aligned}$$

$$\begin{aligned}
& \times \prod_{k=1}^m \left((1 - \pi_4)(2\pi s_{4,k})^{-\frac{p}{2}} \exp\left(-\frac{1}{2s_{4,k}} \boldsymbol{\gamma}_{2,k*}^\top \boldsymbol{\gamma}_{2,k*}\right) \mathbf{I}_{\{\boldsymbol{\gamma}_{2,k*} \neq \mathbf{0}\}} + \pi_4 \delta_0(\boldsymbol{\gamma}_{2,k*}) \right) \\
& \times \prod_{l=1}^s \left((1 - \pi_5)(2\pi s_{5l})^{-\frac{s}{2}} \exp\left(-\frac{1}{2s_{5l}} \boldsymbol{\delta}_l^\top \boldsymbol{\delta}_l\right) \mathbf{I}_{\{\boldsymbol{\delta}_l \neq \mathbf{0}\}} + \pi_5 \delta_0(\boldsymbol{\delta}_l) \right) \\
& \times (\eta_3^2)^{a_3-1} \exp(-b_3 \eta_3^2) \prod_{k=1}^m \left(\frac{p\eta_3^2}{2} \right)^{p+1} (\eta_3^2)^{\frac{p+1}{2}-1} \exp\left(-\frac{\eta_3^2}{2} s_{3,k}\right) \\
& \times (\eta_4^2)^{a_4-1} \exp(-b_4 \eta_4^2) \prod_{k=1}^m \left(\frac{p\eta_4^2}{2} \right)^{p+1} (\eta_4^2)^{\frac{p+1}{2}-1} \exp\left(-\frac{\eta_4^2}{2} s_{4,k}\right) \\
& \times (\eta_5^2)^{a_5-1} \exp(-b_5 \eta_5^2) \prod_{l=1}^s \left(\frac{s\eta_5^2}{2} \right)^{s+1} (\eta_5^2)^{\frac{s+1}{2}-1} \exp\left(-\frac{\eta_5^2}{2} s_{5l}\right) \\
& \times \pi_3^{r_3-1} (1 - \pi_3)^{w_3-1} \\
& \times \pi_4^{r_4-1} (1 - \pi_4)^{w_4-1} \\
& \times \pi_5^{r_5-1} (1 - \pi_5)^{w_5-1}.
\end{aligned}$$

Let $\mu_{ij(-\tilde{\gamma}_0)} = E(Y_{ij}) - \tilde{\mathbf{B}}_0(\mathbf{t}_{ij})^\top \tilde{\gamma}_0$, ($i = 1, \dots, n, j = 1, \dots, J_i$), representing the mean effect without the contribution of $\tilde{\gamma}_0$.

The full conditional distribution of $\tilde{\gamma}_0$ is a multivariate normal distribution $N(\mu_{\tilde{\gamma}_0}, \Sigma_{\tilde{\gamma}_0})$ with mean

$$\mu_{\tilde{\gamma}_0} = \Sigma_{\tilde{\gamma}_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\tilde{\gamma}_0)} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\tilde{\gamma}_0} = \left(\Sigma_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} \tilde{\mathbf{B}}_0(\mathbf{t}_{ij}) \tilde{\mathbf{B}}_0(\mathbf{t}_{ij})^\top \right)^{-1}.$$

Let $\mu_{ij(-\tilde{\gamma}_{3,c})} = E(Y_{ij}) - \tilde{\mathbf{B}}_{3,c}(\mathbf{t}_{ij})^\top \tilde{\gamma}_{3,c}$, ($i = 1, \dots, n, j = 1, \dots, J_i$), representing the mean effect without the contribution of $\tilde{\gamma}_{3,c}$.

The full conditional distribution of $\tilde{\gamma}_{3,c}$ is a multivariate normal distribution $N(\mu_{\tilde{\gamma}_{3,c}}, \Sigma_{\tilde{\gamma}_{3,c}})$ with mean

$$\mu_{\tilde{\gamma}_{3,c}} = \Sigma_{\tilde{\gamma}_{3,c}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\tilde{\gamma}_{3,c})} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\tilde{\gamma}_{3,c}} = \left(\Sigma^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} \tilde{\mathbf{B}}_{3,c}(\mathbf{t}_{ij}) \tilde{\mathbf{B}}_{3,c}(\mathbf{t}_{ij})^\top \right)^{-1}.$$

Let $\mu_{ij(-\gamma_{1,k1})} = E(Y_{ij}) - A_{ijk}\gamma_{1,k1}$ and $l_{1,k} = \pi(\gamma_{1,k1} = 0|\text{rest})$, the conditional posterior distribution of the coefficient of constant additive effect $\gamma_{1,k1}$ is a spike-and-slab distribution:

$$\gamma_{1,k1}|\text{rest} \sim (1 - l_{1,k})N(\mu_{\gamma_{1,k1}}, \Sigma_{\gamma_{1,k1}}) + l_{1,k}\delta_0(\gamma_{1,k1}),$$

where

$$\begin{aligned} \mu_{\gamma_{1,k1}} &= \Sigma_{\gamma_{1,k1}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} A_{ijk} (Y_{ij} - \mu_{ij(-\gamma_{1,k1})} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\gamma_{1,k1}} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} A_{ijk}^2 + \frac{1}{s_{1,k}} \right)^{-1}. \end{aligned}$$

And

$$l_{1,k} = \frac{\pi_1}{\pi_1 + (1 - \pi_1)(s_{1,k})^{-\frac{1}{2}} (\Sigma_{\gamma_{1,k1}})^{\frac{1}{2}} \exp\left\{ \frac{1}{2} \Sigma_{\gamma_{1,k1}} (\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} A_{ijk} \tilde{Y}_{ij})^2 \right\}}.$$

where $\tilde{Y}_{ij} = Y_{ij} - \mu_{ij(-\gamma_{1,k1})} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}$.

The posterior distribution of $\gamma_{1,k1}$ is a mixture of a normal distribution and a point mass at 0. At each iteration of MCMC, $\gamma_{1,k1}$ is drawn from $N(\mu_{\gamma_{1,k1}}, \Sigma_{\gamma_{1,k1}})$ with probability $(1 - l_{1,k})$ and is set to 0 with probability $l_{1,k}$. If $\gamma_{1,k1}$ is set to 0, then we have $\phi_{1,k} = 0$. Otherwise, $\phi_{1,k} = 1$.

Let $\mu_{i(-\gamma_{2,k1})}(\mathbf{t}_{ij}) = E(Y_i(\mathbf{t}_{ij})) - D_{ik}\gamma_{2,k1}$ and $l_{2,k} = \pi(\gamma_{2,k1} = 0|\text{rest})$, the conditional posterior distribution of the coefficient of constant dominant effect $\gamma_{2,k1}$ is a spike-and-slab distribution:

$$\gamma_{2,k1}|\text{rest} \sim (1 - l_{2,k})N(\mu_{\gamma_{2,k1}}, \Sigma_{\gamma_{2,k1}}) + l_{2,k}\delta_0(\gamma_{2,k1}),$$

where

$$\begin{aligned}\mu_{\gamma_{2,k1}} &= \Sigma_{\gamma_{2,k1}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} D_{ijk} (Y_{ij} - \mu_{ij(-\gamma_{2,k1})} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\gamma_{2,k1}} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} D_{ijk}^2 + \frac{1}{s_{2,k}} \right)^{-1}.\end{aligned}$$

And

$$l_{2,k} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)(s_{2,k})^{-\frac{1}{2}} (\Sigma_{\gamma_{2,k1}})^{\frac{1}{2}} \exp\left\{\frac{1}{2} \Sigma_{\gamma_{2,k1}} (\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} D_{ijk} \tilde{y}_{ij})^2\right\}}.$$

where $\tilde{y}_{ij} = Y_{ij} - \mu_{ij(-\gamma_{2,k1})} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}$.

The posterior distribution of $\gamma_{2,k1}$ is a mixture of a normal distribution and a point mass at 0. At each iteration of MCMC, $\gamma_{2,k1}$ is drawn from $N(\mu_{\gamma_{2,k1}}, \Sigma_{\gamma_{2,k1}})$ with probability $(1 - l_{2,k})$ and is set to 0 with probability $l_{2,k}$. If $\gamma_{2,k1}$ is set to 0, then we have $\phi_{3,k} = 0$. Otherwise, $\phi_{3,k} = 1$.

The posterior distribution of the varying coefficient of additive effects $\gamma_{1,k*}(k = 1, \dots, m)$ is also a spike-and-slab distribution. Denote $\mu_{ij(-\gamma_{1,k*})} = E(Y_{ij}) - \mathbf{W}_{ijk}^\top \gamma_{1,k*}$ and $l_{3,k} = \pi(\gamma_{1,k*} = \mathbf{0} | \text{rest})$, the conditional posterior distribution of $\gamma_{1,k*}$ is a multivariate spike-and-slab distribution :

$$\gamma_{1,k*} | \text{rest} \sim (1 - l_{3,k}) N(\mu_{\gamma_{1,k*}}, \Sigma_{\gamma_{1,k*}}) + l_{3,k} \delta_0(\gamma_{1,k*}),$$

where

$$\begin{aligned}\mu_{\gamma_{1,k*}} &= \Sigma_{\gamma_{1,k*}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} \mathbf{W}_{ijk} (Y_{ij} - \mu_{ij(-\gamma_{1,k*})} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\gamma_{1,k*}} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} \mathbf{W}_{ijk} \mathbf{W}_{ijk}^\top + \frac{1}{s_{2,k}} \mathbf{I}_p \right)^{-1}.\end{aligned}$$

And

$$l_{3,k} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)(s_{2,k})^{-\frac{p}{2}} |\Sigma_{\gamma_{1,k*}}|^{\frac{1}{2}} \exp\{\frac{1}{2} \|\Sigma_{\gamma_{1,k*}}^{-\frac{1}{2}} \sum_{i=1}^n \sum_{j=1}^{J_i} \tau \xi_2^{-2} \tilde{v}_{ij}^{-1} \mathbf{W}_{ijk} \tilde{R}_{ijk}\|_2^2\}}.$$

where $\tilde{R}_{ij} = Y_{ij} - \mu_{ij(-\gamma_{1,k*})} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}$.

The posterior distribution of $\gamma_{1,k*}$ is a mixture of a multivariate normal distribution and a point mass at 0. At each iteration of MCMC, $\gamma_{1,k*}$ is drawn from $N(\mu_{\gamma_{1,k*}}, \Sigma_{\gamma_{1,k*}})$ with probability $(1 - l_{3,k})$ and is set to 0 with probability $l_{3,k}$. If $\gamma_{1,k*}$ is set to $\mathbf{0}$, then we have $\phi_{2,k} = 0$. Otherwise, $\phi_{2,k} = 1$.

The posterior distribution of the varying coefficient of dominant effects $\gamma_{2,k*}$ ($k = 1, \dots, m$) is also a spike-and-slab distribution. Denote $\mu_{ij(-\gamma_{2,k*})} = E(Y_i(t_{ij})) - \mathbf{V}_{ijk}^\top \gamma_{2,k*}$ and $l_{4,k} = \pi(\gamma_{2,k*} = \mathbf{0} | \text{rest})$, the conditional posterior distribution of $\gamma_{2,k*}$ is a multivariate spike-and-slab distribution :

$$\gamma_{2,k*} | \text{rest} \sim (1 - l_{4,k}) N(\mu_{\gamma_{2,k*}}, \Sigma_{\gamma_{2,k*}}) + l_{4,k} \delta_0(\gamma_{2,k*}),$$

where

$$\begin{aligned} \mu_{\gamma_{2,k*}} &= \Sigma_{\gamma_{2,k*}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} \mathbf{V}_{ijk} (Y_{ij} - \mu_{ij(-\gamma_{2,k*})} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\gamma_{2,k*}} &= (\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} \mathbf{V}_{ijk} \mathbf{V}_{ijk}^\top + \frac{1}{s_{4,k}} \mathbf{I}_p)^{-1}. \end{aligned}$$

And

$$l_{4,k} = \frac{\pi_4}{\pi_4 + (1 - \pi_4)(s_{4,k})^{-\frac{p}{2}} |\Sigma_{\gamma_{2,k*}}|^{\frac{1}{2}} \exp\{\frac{1}{2} \|\Sigma_{\gamma_{2,k*}}^{-\frac{1}{2}} \sum_{i=1}^n \sum_{j=1}^{J_i} \tau \xi_2^{-2} \tilde{v}_{ij}^{-1} \mathbf{V}_{ijk} \tilde{r}_{ij}\|_2^2\}}.$$

where $\tilde{r}_{ij} = Y_{ij} - \mu_{ij(-\gamma_{2,k*})} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij}$.

The posterior distribution of $\gamma_{2,k*}$ is a mixture of a multivariate normal distribution and a point mass at 0. At each iteration of MCMC, $\gamma_{2,k*}$ is drawn from $N(\mu_{\gamma_{2,k*}}, \Sigma_{\gamma_{2,k*}})$ with

probability $(1 - l_{4,k})$ and is set to 0 with probability $l_{4,k}$. If $\gamma_{2,k*}$ is set to $\mathbf{0}$, then we have $\phi_{4,k} = 0$. Otherwise, $\phi_{4,k} = 1$.

The full conditional posterior distribution of $(s_{1,k})^{-1}$ is:

$$(s_{1,k})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(1, \frac{\eta_1^2}{2}) & \text{if } \gamma_{1,k1} = 0 \\ \text{Inverse-Gaussian}(\eta_1^2, \sqrt{\frac{\eta_1^2}{\gamma_{1,k1}}}) & \text{if } \gamma_{1,k1} \neq 0 \end{cases}$$

The full conditional posterior distribution of $(s_{3,k})^{-1}$ is:

$$(s_{3,k})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(1, \frac{\eta_2^2}{2}) & \text{if } \gamma_{2,k1} = 0 \\ \text{Inverse-Gaussian}(\eta_2^2, \sqrt{\frac{\eta_2^2}{\gamma_{2,k1}}}) & \text{if } \gamma_{2,k1} \neq 0 \end{cases}$$

Similarly, The full conditional posterior distribution of $(s_{2,k})^{-1}$ is:

$$(s_{2,k})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{p+1}{2}, \frac{\eta_3^2}{2}) & \text{if } \gamma_{1,k*} = \mathbf{0} \\ \text{Inverse-Gaussian}(\eta_3^2, \sqrt{\frac{\eta_3^2}{\|\gamma_{1,k*}\|_2^2}}) & \text{if } \gamma_{1,k*} \neq \mathbf{0} \end{cases}$$

Similarly, The full conditional posterior distribution of $(s_{4,k})^{-1}$ is:

$$(s_{4,k})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{p+1}{2}, \frac{\eta_4^2}{2}) & \text{if } \gamma_{2,k*} = \mathbf{0} \\ \text{Inverse-Gaussian}(\eta_4^2, \sqrt{\frac{\eta_4^2}{\|\gamma_{2,k*}\|_2^2}}) & \text{if } \gamma_{2,k*} \neq \mathbf{0} \end{cases}$$

The full conditional posterior distribution of η_1^2 :

$$\eta_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{k=1}^m s_{1,k}}{2}).$$

The full conditional posterior distribution of η_3^2 :

$$\eta_3^2 | \text{rest} \sim \text{Gamma}(a_3 + m, b_3 + \frac{\sum_{k=1}^m s_{3,k}}{2}).$$

Likewise, the full conditional posterior distribution of η_2^2 :

$$\eta_2^2 | \text{rest} \sim \text{Gamma}(a_2 + \frac{m(p+1)}{2}, b_2 + \frac{\sum_{k=1}^m s_{2,k}}{2}).$$

Likewise, the full conditional posterior distribution of η_4^2 :

$$\eta_4^2 | \text{rest} \sim \text{Gamma}(a_4 + \frac{m(p+1)}{2}, b_4 + \frac{\sum_{k=1}^m s_{4,k}}{2}).$$

The full conditional posterior distribution of π_1 :

$$\pi_1 | \text{rest} \sim \text{Beta}(r_1 + \sum_{k=1}^m \mathbf{I}_{\{\gamma_{1,k1}=0\}}, w_1 + \sum_{k=1}^m \mathbf{I}_{\{\gamma_{1,k1} \neq 0\}}).$$

The full conditional posterior distribution of π_3 :

$$\pi_3 | \text{rest} \sim \text{Beta}(r_3 + \sum_{k=1}^m \mathbf{I}_{\{\gamma_{2,k1}=0\}}, w_3 + \sum_{k=1}^m \mathbf{I}_{\{\gamma_{2,k1} \neq 0\}}).$$

Similarly, the full conditional posterior distribution of π_2 :

$$\pi_2 | \text{rest} \sim \text{Beta}(r_2 + \sum_{k=1}^m \mathbf{I}_{\{\gamma_{1,k*}=0\}}, w_2 + \sum_{k=1}^m \mathbf{I}_{\{\gamma_{1,k*} \neq 0\}}).$$

Similarly, the full conditional posterior distribution of π_4 :

$$\pi_4 | \text{rest} \sim \text{Beta}(r_4 + \sum_{k=1}^m \mathbf{I}_{\{\gamma_{2,k*}=0\}}, w_4 + \sum_{k=1}^m \mathbf{I}_{\{\gamma_{2,k*} \neq 0\}}).$$

The full conditional posterior distribution of $\tilde{v}_{ij}$:

$$\tilde{v}_{ij} | \text{rest} \sim \text{Inverse-Gaussian}(a_{\tilde{v}_{ij}}, \frac{\tau \xi_1^2}{\xi_2^2} + 2\tau).$$

where

$$a_{\tilde{v}_{ij}} = \sqrt{\frac{\xi_1^2 + 2\xi_2^2}{(Y_{ij} - \mu_{ij} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij})^2}}.$$

The full conditional posterior distribution of τ :

$$\tau|\text{rest} \sim \text{Gamma}(c + \frac{3 \sum_{i=1}^n J_i}{2}, b_\tau).$$

where

$$b_\tau = d + \sum_{i=1}^n \sum_{j=1}^{J_i} (\tilde{v}_{ij} + (2\xi_2^2)^{-1} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s \boldsymbol{\delta} - \xi_1 \tilde{v}_{ij})^2)$$

Let $\mathbf{F}_{ij} = (\mathbf{b}_i^\top \otimes \mathbf{Z}_{ij}^\top) \mathbf{J}_s$. The posterior distribution of the random effect coefficient $\boldsymbol{\delta}_l$ ($l = 1, \dots, s$) is also a spike-and-slab distribution. Denote $\mu_{ij(-\boldsymbol{\delta}_l)} = E(Y_{ij}) - \mathbf{F}_{ijl}^\top \boldsymbol{\delta}_l$ and $l_{5,l} = \pi(\boldsymbol{\delta}_l = \mathbf{0}|\text{rest})$, the conditional posterior distribution of $\boldsymbol{\delta}_l$ is a multivariate spike-and-slab distribution :

$$\boldsymbol{\delta}_l|\text{rest} \sim (1 - l_{5,l}) N(\mu_{\boldsymbol{\delta}_l}, \Sigma_{\boldsymbol{\delta}_l}) + l_{5,l} \delta_0(\boldsymbol{\delta}_l),$$

where

$$\begin{aligned} \mu_{\boldsymbol{\delta}_l} &= \Sigma_{\boldsymbol{\delta}_l} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} \mathbf{F}_{ijl} (Y_{ij} - \mu_{ij(-\boldsymbol{\delta}_l)} - \tilde{F}_{ijl}), \\ \Sigma_{\boldsymbol{\delta}_l} &= (\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^{J_i} \tilde{v}_{ij}^{-1} \mathbf{F}_{ijl} \mathbf{F}_{ijl}^\top + \frac{1}{s_{5,l}} \mathbf{I}_s)^{-1}. \end{aligned}$$

where $\tilde{F}_{ij} = \mathbf{F}_{ijl} \boldsymbol{\delta}_l - \xi_1 \tilde{v}_{ij}$ And

$$l_{5,l} = \frac{\pi_5}{\pi_5 + (1 - \pi_5)(s_{5,l})^{-\frac{s}{2}} |\Sigma_{\boldsymbol{\delta}_l}|^{\frac{1}{2}} \exp\{\frac{1}{2} \|\Sigma_{\boldsymbol{\delta}_l}^{\frac{1}{2}} \sum_{i=1}^n \sum_{j=1}^{J_i} \tau \xi_2^{-2} \tilde{v}_{ij}^{-1} \mathbf{F}_{ijl} R_{ij}\|_2^2\}}.$$

where $R_{ij} = Y_{ij} - \mu_{ij(-\boldsymbol{\delta}_l)} - \xi_1 \tilde{v}_{ij}$.

The full conditional posterior distribution of $(s_{5,l})^{-1}$ is:

$$(s_{5,l})^{-1}|\text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{s+1}{2}, \frac{\eta_5^2}{2}) & \text{if } \boldsymbol{\delta}_l = \mathbf{0} \\ \text{Inverse-Gaussian}(\eta_5^2, \sqrt{\frac{\eta_5^2}{\|\boldsymbol{\delta}_l\|_2^2}}) & \text{if } \boldsymbol{\delta}_l \neq \mathbf{0} \end{cases}$$

The full conditional posterior distribution of η_5^2 :

$$\eta_5^2 | \text{rest} \sim \text{Gamma}(a_5 + \frac{s(s+1)}{2}, b_5 + \frac{\sum_{l=1}^s s_{5,l}}{2}).$$

The full conditional posterior distribution of π_5 :

$$\pi_5 | \text{rest} \sim \text{Beta}(r_5 + \sum_{l=1}^s \mathbf{I}_{\{\delta_l=0\}}, w_5 + \sum_{l=1}^s \mathbf{I}_{\{\delta_l \neq 0\}}).$$

3.3 Summary of all methods

Table 3.1: *Summary of comparisons among all Bayesian regularized mixed-effects models.*

Methods	Robustness	Likelihood	Variable Selection	Structure	Identification
RBVSS-SI	Yes	Asymmetric Laplacian	Spike-and-slab prior	Bi-level	MPM
RBVSS	Yes	Asymmetric Laplacian	Spike-and-slab prior	Group level	MPM
RBV-SI	Yes	Asymmetric Laplacian	Laplacian shrinkage	Bi-level	Credible interval
RBV	Yes	Asymmetric Laplacian	Laplacian shrinkage	Group level	Credible interval
BVSS-SI	No	Gaussian	Spike-and-slab prior	Bi-level	MPM
BVSS	No	Gaussian	Spike-and-slab prior	Group level	MPM
BV-SI	No	Gaussian	Laplacian shrinkage	Bi-level	Credible interval
BV	No	Gaussian	Laplacian shrinkage	Group level	Credible interval

Bi-level: both the group and individual level.

MPM: Median Probability Model.

3.4 Simulation

The eight methods being compared can be categorized based on three criteria: whether they incorporate robustness, whether they use spike-and-slab priors, and whether they consider structure identification in variable selection for longitudinal traits. In our simulation study, we thoroughly evaluate the performance of all the methods and show the advantages of the proposed RBVSS-SI, as it simultaneously incorporates robustness, spike-and-slab priors,

and structured identification. We consider RBVSS as another proposed method, but it does not include structural identification. It does not differentiate between the nonzero constant effect and the nonlinear effect. In RBVSS, without implementing structural identification, the coefficients of L basis functions of β_{1k} and β_{2k} are selected together as one group at a certain group level. We illustrate the benefit of structural identification in detecting interaction effects by comparing RBVSS-SI with RBVSS. Similarly, the methods being compared without spike-and-slab priors are RBV-SI and RBV, where RBV-SI distinguishes between constant and varying effects, but RBV does not. The remaining four methods—BVSS-SI, BVSS, BV-SI, and BV—do not incorporate robustness. The goal of the proposed method is to highlight the benefits of using the spike-and-slab prior and to demonstrate the importance of including structural identification.

In our simulation setup, we consider two examples. The response variables are generated from model (3.1) under all settings with $n = 400$ and $m = 200$. After basis expansion, the total dimension of the regression coefficients to be estimated is significantly higher than the sample size $n = 400$. For instance, when the number of basis function $L = 5$, the effective dimension of the regression coefficient is 505. In each example, we evaluate performance in terms of both identification and estimation. We use the integrated mean absolute deviation (IMAD) and the integrated mean squared error (IMSE) to assess the accuracy of the estimation for the nonlinear effects. The time points of observation t_{ij} were generated using a methodology similar to that described by (Huang et al. (2002)). Each subject was assigned a series of "scheduled" time points $\{1, \dots, 30\}$, with each scheduled time having a 60% chance being skipped. Subsequently, the actual observation time was derived by adding a random perturbation from Uniform $(-0.5, 0.5)$ to the scheduled times that were not skipped. Let $\hat{\beta}_{1k}(t_{ij})$ and $\hat{\beta}_{2k}(t_{ij})$ denote the estimates of the nonparametric functions $\beta_{1k}(t_{ij})$ and $\beta_{2k}(t_{ij})$, respectively, and let $\{t_g\}_{g=1}^{n_{grid}}$ represent the grid points at which β_{1k} and β_{2k} are evaluated. For the methods that incorporate robustness, the integrated mean absolute deviation (IMAD) of $\hat{\beta}_{1k}(t_{ij})$ and $\hat{\beta}_{2k}(t_{ij})$ are defined as $\text{IMAD}(\hat{\beta}_{1k}(t_{ij})) = \frac{1}{n_{grid}} \sum_{g=1}^{n_{grid}} |\hat{\beta}_{1k}(t_g) - \beta_{1k}(t_g)|$ and $\text{IMAD}(\hat{\beta}_{2k}(t)) = \frac{1}{n_{grid}} \sum_{g=1}^{n_{grid}} |\hat{\beta}_{2k}(t_g) - \beta_{2k}(t_g)|$. Note that $\text{IMAD}(\hat{\beta}_{1k})$ and $\text{IMAD}(\hat{\beta}_{2k})$ reduce to $\text{MAD}(\hat{\beta}_{1k})$ and $\text{MAD}(\hat{\beta}_{2k})$ respectively when β_{1k} and β_{2k} are constant. For the

methods that do not incorporate robustness, the integrated mean squared error (IMSE) of $\hat{\beta}_{1k}(t_{ij})$ and $\hat{\beta}_{2k}(t_{ij})$ are defined as $\text{IMSE}(\hat{\beta}_{1k}(t_{ij})) = \frac{1}{n_{grid}} \sum_{g=1}^{n_{grid}} \{\hat{\beta}_{1k}(t_g) - \beta_{1k}(t_g)\}^2$ and $\text{IMSE}(\hat{\beta}_{2k}(t_{ij})) = \frac{1}{n_{grid}} \sum_{g=1}^{n_{grid}} \{\hat{\beta}_{2k}(t_g) - \beta_{2k}(t_g)\}^2$. Note that $\text{IMSE}(\hat{\beta}_{1k})$ and $\text{IMSE}(\hat{\beta}_{2k})$ reduce to $\text{MSE}(\hat{\beta}_{1k})$ and $\text{MSE}(\hat{\beta}_{2k})$ respectively when β_{1k} and β_{2k} are constant. Identification accuracy is evaluated based on the number of true positives and false positives.

Example 1

First, we generate an $n \times m$ matrix of gene expressions, where $n = 400$ and $m = 200$, from a multivariate normal distribution with zero mean vector. We use an autoregressive (AR) correlation structure for gene expression data, where the correlation coefficient between genes j and k is $\rho^{|j-k|}$, with $\rho = 0.5$. We evaluate whether the proposed method performs better than the alternatives on simulated single-nucleotide polymorphism (SNP) data. The SNP genotype data $\mathbf{X}_i$ are simulated by dichotomizing gene expression values at the 1st and 3rd quartiles, assigning a 3-level genotype scheme (2, 1, 0) for the genotypes (AA, Aa, aa), respectively. The genetic variables $\mathbf{A}_i$ and $\mathbf{D}_i$ were coded as (1, 0, -1) and (-0.5, 0.5, -0.5) for the genotypes (AA, Aa, aa), respectively, following an orthogonal quantitative genetic model (Cockerham (1954)). The longitudinal response variable is generated from model (3.1) with a sample size of $n = 400$. Specifically, we consider two types of random effects: (1) the random intercept model, and (2) the random intercept-slope model, corresponding to $s = 1$ and 2 in model (3.1), respectively. The model errors were simulated as follows: $N(\mu, 1)$ (Error 1), $t(2)$ with mean $= \mu$ (Error 2), $\text{LogNormal}(\mu, 1)$ (Error 3), $80\%N(\mu, 1) + 20\%N(\mu, 3)$ (Error 4), and $\text{Laplace}(\mu, b)$ with the scale parameter $b = 1$ (Error 5). All distributions except $N(0, 1)$ have heavy tails. The parameter μ is chosen so that the θ th quantile of each distribution is equal to 0.

The coefficients are generated as follows: $\alpha_0(t_{ij}) = 2 + 2\sin(2\pi t_{ij})$, $\beta_{11}(t_{ij}) = 2.5\exp(2.5t_{ij} - 1)$, $\beta_{21}(t_{ij}) = 3t_{ij}^2 - 2t_{ij} + 2$, $\beta_{12}(t_{ij}) = -4t_{ij}^3 + 3$, $\beta_{13}(t_{ij}) = 3 - 2t_{ij}$, $\beta_{22} = 0.8$, $\beta_{23} = -1.2$, $\beta_{14} = 0.7$, $\beta_{24} = -1.1$. We set all the remaining coefficients to 0.

Example 2

In this example, we generate a $n \times m$ matrix of SNP genotype data using a pairwise linkage disequilibrium (LD) structure. For the two minor alleles A and B of two adjacent SNPs, let

q_1 and q_2 represent the minor allele frequencies (MAFs) for these alleles, respectively. The frequencies of four haplotypes are calculated 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 represents the linkage disequilibrium (LD). Under Hardy-Weinberg equilibrium, the SNP genotype (AA, Aa, aa) at locus 1 can be generated from a multinomial distribution with frequencies $(q_1^2, 2q_1(1 - q_1), (1 - q_1)^2)$. Using the conditional genotype probability matrix, we can simulate the genotypes for locus 2. With MAFs 0.3 and pairwise correlation $r = 0.5$, we have $\delta = r\sqrt{q_1(1 - q_1)q_2(1 - q_2)}$.

Identification and estimation criteria are employed to evaluate performance using the posterior samples obtained from the Gibbs sampler. For identification, we use the median probability model (MPM) to identify predictors associated with the response variable for methods that incorporate spike-and-slab priors (Barbieri and Berger (2004)). Assume we have collected H posterior samples from MCMC procedure after discarding the burn-in period.. At the h th MCMC iteration, the k th predictor is included in the model if the corresponding indicator $\phi_k^{(h)} = 1$. Therefore, the posterior probability of including the k th predictor in the final model is defined as:

$$p_k = \hat{\pi}(\phi_k = 1|y) = \frac{1}{H} \sum_{h=1}^H \phi_k^{(h)}, k = 1, \dots, m.$$

In general, higher posterior inclusion probability p_k indicates stronger empirical evidence that the k th predictor has a non-zero coefficient and therefore a stronger association with the response variable. The MPM model is defined as the set of predictors selected when the posterior inclusion probability is at least 0.5 (Barbieri and Berger (2004)). For the non spike-and-slab methods, the 95% credible interval (95% CI) is used to identify important main and interaction effects. Performance in identification is evaluated based on the number of true positives (TP) and false positives (FP).

Tables 3.2 presents the identification results for the random intercept-slope model for $m=200$, at the 50% quantile level ($\theta=0.5$). These tables clearly demonstrate the advantages of incorporating (1) robustness, (2) the spike-and-slab prior, and (3) structured identification. Under heavy-tailed errors and data contamination, the robust Bayesian variable selection

methods generally outperform the non-robust alternatives. In particular, we can see that the proposed method performs better than RBVSS. RBVSS fails to identify any non-zero constant effects and has a high rate of false positives for detecting varying effects because it lacks structural identification to distinguish between main-effect-only cases and varying effects. In contrast, RBVSS-SI identifies most of the true effects with very few false positives. For example, in Example 1 under $t(2)$ (Error 2), the proposed RBVSS-SI detected almost all of the varying effects, achieving a true positive rate of 3.98 (sd0.69) and a small number of false positives, with a rate of 1.83 (sd 0.33). It also identifies 3.97 (sd0.24) of the 4 true constant effects, with a smaller false positive rate of 1.54 (sd0.68). Its non-robust counterpart, BVSS-SI, achieved a comparable true positive rate of 3.74 (sd0.56) for varying effects but had a much higher false positive rate of 3.26 (sd0.75). Additionally, it had a smaller true positive rate of 2.54 (sd0.53) for constant effects and a higher false positive rate of 2.88 (sd0.95). This highlights the strength of robust methods in the presence of data heterogeneity. Additionally, methods with spike-and-slab priors have outperformed those without such priors in imposing exact sparsity. For instance, RBVSS-SI detected a true positive rate of 3.98 (sd0.69) and a false positive rate of 1.83(sd0.33) for the true varying effects, which is better compared to the true positive rate of 3.92 (sd0.57) and false positive rate of 3.89 (sd0.51) achieved by RBV-BI. Besides, it identifies 3.97 (sd0.24) out of the 4 true constant effects with a smaller false positive rate of 1.54 (sd0.68), which is better than the true positive rate of 2.88 (sd 0.56) and false positive rate of 3.96 (sd 0.29) achieved by RBV-BI.

The estimation performance also been evaluated by the MADs and MSEs. The results were listed in Table 3.3 for Example 1. The IMAD and IMSE for all true varying effects, MAD and MSE for constant effects, as well as the total absolute deviation and the total squared errors for all coefficient estimates are provided in the Table. Overall, the proposed method has the smallest IMAD, MAD and the total absolute deviation among all the alternative methods. For example, in Table 3.3 under $t(2)$ (Error 2), the RBVSS-SI has the smallest total absolute deviation 0.292(sd 0.104) among all the approaches. The key of the superior performance lies in (1) robustness to deal with heavy-tailed error, (2) the spike and slab

priors for achieving sparsity and (3) structural identification for the constant and varying effects. Compared with RBV-SI which has (1) but without spike and slab prior, RBVSS-SI performs better when estimating both varying and constant coefficients. For instance, under $t(2)$ (Error 2), the IMAD and MAD on $\alpha_0(t_{ij})$ and β_{22} are 0.102(sd 0.430) and 0.097(sd 0.047), respectively. While RBV-SI yields 0.291(sd 0.104) and 0.103(sd 0.063), correspondingly. Besides, compared with RBVSS which adopts the spike and slab priors without considering structured Bayesian variable selection, RBVSS-SI has comparable estimation performance on coefficients. Overall, the methods incorporating robustness outperform the methods without robustness under heavy tailed distributed errors.

After fitting the proposed method to the data generated for each replicate, we obtained the posterior median estimates and the 95% credible intervals. The median estimates, along with the lower and upper bounds of the credible intervals, were averaged across 100 replicates to produce the estimated varying coefficients and their corresponding 95% credible intervals, as displayed in Figure B.1. Additionally, we assessed the empirical 95% coverage probabilities of 8 Bayesian methods using their pointwise 95% credible intervals over the $N = \sum_{i=1}^{J_i}$ grid points for the varying coefficients, as long as the 95% coverage probabilities for the constant coefficients. For methods with structural identification, the empirical 95% coverage probabilities were applied on the varying coefficients and 95% coverage probabilities for the constant effects, while the methods without structural identification, the empirical 95% coverage probabilities were applied on the varying coefficients and constant coefficients. Table B.11 in the Appendix presents the 95% coverage probabilities for 5 varying coefficient functions and 4 constant coefficients under simulated gene expression data with i.i.d. $t(2)$ errors.

In addition to the simulation results of example 1 on the random intercept-slope model as shown in Table 3.2 to Table 3.3, we have also conducted simulations of example 2 under the random intercept - slope model, and tabulate the identification results in Table B.1 and the estimation results in Table B.2 respectively in the appendix. Besides, the performance has been assessed at 30% and 70% quantile levels. In the appendix, when $\theta = 0.3$ under random intercept-slope model, the identification results are provided in Table B.3 and Table

B.5, while the estimation accuracy can be found in Table B.4 and Table B.6. With the 70% quantile level, Table B.7 and Table B.9 report the identification results, and Table B.8 and Table B.10 show the parameter estimation performance.

Table 3.2: *Simulation results in example 1 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at 50th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.*

$n = 400$ $m = 200$		Varying		Constant	
		TP	FP	TP	FP
Error 1	RBVSS-SI	3.98(0.69)	1.83(0.33)	3.97(0.27)	1.54(0.68)
	RBV-SI	3.96(0.10)	4.20(0.83)	3.70(0.46)	3.49(0.76)
	BVSS-SI	3.99(0.15)	1.75(0.21)	3.94(0.31)	1.65(0.24)
	BV-SI	3.95(0.12)	4.34(0.25)	3.75(0.78)	3.53(0.67)
	RBVSS	3.96(0.18)	5.13(0.37)	0.00(0.00)	0.00(0.00)
	RBV	3.93(0.15)	7.62(0.66)	0.00(0.00)	0.00(0.00)
	BVSS	3.95(0.32)	5.15(0.43)	0.00(0.00)	0.00(0.00)
	BV	3.91(0.19)	7.71(0.27)	0.00(0.00)	0.00(0.00)
Error 2	RBVSS-SI	3.96(0.69)	1.27(0.74)	3.95(0.24)	1.26(0.57)
	RBV-SI	3.92(0.57)	3.89(0.51)	2.88(0.56)	3.96(0.29)
	BVSS-SI	3.74(0.56)	3.26(0.75)	2.54(0.53)	2.88(0.95)
	BV-SI	3.06(0.81)	3.95(0.76)	2.07(0.64)	5.52(1.31)
	RBVSS	3.93(0.45)	5.16(0.51)	0.00(0.00)	0.00(0.00)
	RBV	3.89(0.34)	6.78(1.11)	0.00(0.00)	0.00(0.00)
	BVSS	3.73(0.45)	7.79(1.95)	0.00(0.00)	0.00(0.00)
	BV	3.11(0.26)	8.68(1.27)	0.00(0.00)	0.00(0.00)
Error 3	RBVSS-SI	3.97(0.48)	0.88(0.35)	3.93(0.54)	1.08(0.76)
	RBV-SI	3.96(0.54)	2.56(0.25)	2.47(1.82)	3.07(0.15)
	BVSS-SI	3.68(0.89)	2.96(2.78)	1.98(0.27)	2.15(0.45)
	BV-SI	2.98(0.84)	4.48(0.26)	1.84(1.10)	4.97(0.22)
	RBVSS	3.92(0.38)	5.13(0.37)	0.00(0.00)	0.00(0.00)
	RBV	3.87(0.49)	7.62(1.66)	0.00(0.00)	0.00(0.00)
	BVSS	3.69(0.25)	7.28(0.35)	0.00(0.00)	0.00(0.00)
	BV	3.09(0.37)	8.87(0.29)	0.00(0.00)	0.00(0.00)
Error 4	RBVSS-SI	3.96(0.33)	0.74(0.27)	3.89(0.55)	1.98(0.24)
	RBV-SI	3.89(1.10)	1.85(0.25)	2.19(1.23)	2.98(0.45)
	BVSS-SI	3.76(0.84)	2.07(0.25)	1.79(2.88)	2.18(1.06)
	BV-SI	3.05(0.23)	4.79(0.15)	1.81(1.14)	3.19(0.25)
	RBVSS	3.91(0.44)	5.24(0.37)	0.00(0.00)	0.00(0.00)
	RBV	3.88(0.29)	7.27(1.66)	0.00(0.00)	0.00(0.00)
	BVSS	3.73(0.38)	6.37(0.58)	0.00(0.00)	0.00(0.00)
	BV	3.06(0.19)	7.65(1.68)	0.00(0.00)	0.00(0.00)
Error 5	RBVSS-SI	3.92(0.31)	0.94(0.35)	3.88(0.24)	2.01(0.28)
	RBV-SI	3.90(0.45)	2.14(0.45)	2.77(1.31)	2.76(0.37)
	BVSS-SI	3.67(1.30)	2.87(0.45)	1.96(2.61)	2.65(1.34)
	BV-SI	3.12(0.71)	5.11(0.56)	1.89(1.07)	3.98(0.89)
	RBVSS	3.94(0.26)	4.69(0.60)	0.00(0.00)	0.00(0.00)
	RBV	3.93(0.39)	6.48(0.84)	0.00(0.00)	0.00(0.00)
	BVSS	3.71(0.26)	5.77(0.95)	0.00(0.00)	0.00(0.00)
	BV	3.10(0.29)	7.67(0.95)	0.00(0.00)	0.00(0.00)

Table 3.3: Estimation accuracy in example 1 for $n = 400, m = 200$ under random intercept - slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 50th quantile with i.i.d. errors.

	RBVSS-SI	RBV-SI	BVSS-SI	BV-SI	RBVSS	RBV	BVSS	BV
Error 1								
IMSE								
$\mu(t)$	0.024(0.007)	0.039(0.088)	0.025(0.024)	0.039(0.097)	0.025(0.005)	0.040(0.009)	0.025(0.004)	0.041(0.006)
$\beta_{11}(t)$	0.032(0.015)	0.088(0.035)	0.031(0.026)	0.089(0.065)	0.030(0.006)	0.101(0.013)	0.085(0.006)	0.090(0.014)
$\beta_{21}(t)$	0.051(0.017)	0.102(0.045)	0.050(0.018)	0.103(0.056)	0.043(0.008)	0.124(0.011)	0.047(0.006)	0.124(0.012)
$\beta_{12}(t)$	0.017(0.012)	0.098(0.041)	0.019(0.026)	0.110(0.036)	0.015(0.008)	0.119(0.020)	0.017(0.007)	0.120(0.015)
$\beta_{13}(t)$	0.035(0.018)	0.111(0.051)	0.035(0.026)	0.112(0.036)	0.033(0.008)	0.131(0.022)	0.035(0.009)	0.132(0.033)
MSE								
β_{22}	0.009(0.018)	0.035(0.038)	0.011(0.021)	0.032(0.032)	0.015(0.023)	0.039(0.029)	0.012(0.019)	0.038(0.021)
β_{23}	0.003(0.005)	0.021(0.022)	0.005(0.022)	0.029(0.028)	0.004(0.009)	0.019(0.009)	0.007(0.018)	0.028(0.028)
β_{14}	0.006(0.008)	0.032(0.030)	0.009(0.025)	0.041(0.032)	0.008(0.023)	0.038(0.009)	0.011(0.038)	0.039(0.029)
β_{24}	0.008(0.009)	0.018(0.022)	0.012(0.035)	0.022(0.024)	0.007(0.028)	0.022(0.048)	0.014(0.059)	0.025(0.054)
Total	0.134(0.040)	0.752(0.196)	0.138(0.037)	0.764(0.265)	0.149(0.026)	0.821(0.018)	0.153(0.020)	0.835(0.062)
Error 2								
IMSE								
$\mu(t)$	0.096(0.430)	0.157(0.104)	0.218(0.287)	0.376(0.841)	0.093(0.007)	0.160(0.103)	0.213(1.431)	0.372(0.675)
$\beta_{11}(t)$	0.031(0.011)	0.229(0.112)	0.328(0.844)	0.557(0.988)	0.029(0.015)	0.228(0.144)	0.326(0.720)	0.551(0.421)
$\beta_{21}(t)$	0.083(0.041)	0.147(0.099)	0.289(0.397)	0.385(1.666)	0.084(0.020)	0.147(0.107)	0.283(0.444)	0.379(0.223)
$\beta_{12}(t)$	0.075(0.056)	0.186(0.121)	0.333(0.378)	0.485(0.718)	0.076(0.033)	0.184(0.146)	0.328(0.385)	0.483(0.321)
$\beta_{13}(t)$	0.068(0.061)	0.168(0.138)	0.397(0.561)	0.471(0.555)	0.071(0.045)	0.164(0.151)	0.401(0.375)	0.468(0.329)
MSE								
β_{22}	0.097(0.047)	0.126(0.063)	0.196(0.079)	0.240(0.537)	0.112(0.045)	0.128(0.045)	0.192(0.038)	0.251(0.085)
β_{23}	0.058(0.551)	0.095(0.009)	0.121(0.285)	0.265(0.672)	0.067(0.089)	0.092(0.048)	0.129(0.088)	0.271(0.048)
β_{14}	0.011(0.008)	0.068(0.061)	0.134(0.577)	0.316(0.673)	0.013(0.045)	0.062(0.079)	0.139(0.078)	0.323(0.057)
β_{24}	0.014(0.003)	0.034(0.037)	0.107(0.213)	0.193(0.167)	0.018(0.089)	0.039(0.058)	0.114(0.045)	0.189(0.067)
Total	0.292(0.104)	1.382(0.340)	1.372(0.456)	2.091(0.556)	0.316(0.022)	1.672(0.518)	1.563(0.889)	3.021(0.574)
Error 3								
IMSE								
$\mu(t)$	0.071(0.040)	0.137(0.237)	0.145(0.598)	0.435(0.589)	0.072(0.036)	0.135(0.153)	0.142(0.587)	0.434(0.897)
$\beta_{11}(t)$	0.282(0.113)	0.510(0.256)	0.657(0.462)	0.859(0.319)	0.281(0.046)	0.506(0.284)	0.652(0.239)	0.861(0.495)
$\beta_{21}(t)$	0.193(0.073)	0.302(0.161)	0.342(0.441)	0.626(0.285)	0.190(0.130)	0.305(0.202)	0.339(0.520)	0.623(0.246)
$\beta_{12}(t)$	0.163(0.188)	0.278(0.262)	0.312(0.497)	0.512(0.381)	0.164(0.112)	0.277(0.270)	0.309(0.456)	0.509(0.321)
$\beta_{13}(t)$	0.081(0.098)	0.209(0.111)	0.341(0.287)	0.499(0.311)	0.080(0.109)	0.208(0.281)	0.339(0.212)	0.494(0.211)
MSE								
β_{22}	0.139(0.107)	0.196(0.099)	0.205(0.030)	0.321(0.541)	0.141(0.098)	0.192(0.111)	0.211(0.041)	0.329(0.418)
β_{23}	0.050(0.130)	0.142(0.117)	0.182(0.240)	0.460(0.270)	0.042(0.128)	0.145(0.121)	0.179(0.224)	0.472(0.117)
β_{14}	0.034(0.057)	0.131(0.146)	0.148(0.616)	0.332(0.671)	0.038(0.049)	0.137(0.132)	0.151(0.099)	0.299(0.511)
β_{24}	0.016(0.025)	0.078(0.075)	0.088(0.194)	0.214(0.378)	0.014(0.031)	0.084(0.087)	0.092(0.071)	0.234(0.211)
Total	0.416(0.295)	1.455(0.623)	1.619(0.323)	2.127(0.289)	0.892(0.456)	1.715(0.720)	1.929(0.456)	3.127(0.321)
Error 4								
IMSE								
$\mu(t)$	0.025(0.006)	0.079(0.009)	0.113(0.214)	0.724(0.943)	0.026(0.006)	0.077(0.009)	0.112(0.093)	0.721(0.967)
$\beta_{11}(t)$	0.018(0.026)	0.045(0.093)	0.068(0.235)	0.118(0.239)	0.017(0.007)	0.043(0.128)	0.068(1.415)	0.116(0.352)
$\beta_{21}(t)$	0.019(0.021)	0.054(0.023)	0.079(0.808)	0.211(0.891)	0.017(0.007)	0.054(0.046)	0.077(0.494)	0.209(0.331)
$\beta_{12}(t)$	0.016(0.047)	0.045(0.075)	0.058(0.351)	0.262(0.892)	0.015(0.008)	0.043(0.088)	0.056(0.962)	0.259(0.047)
$\beta_{13}(t)$	0.014(0.039)	0.039(0.046)	0.051(0.22)	0.209(0.822)	0.013(0.011)	0.037(0.038)	0.048(0.711)	0.207(0.089)
MSE								
β_{22}	0.006(0.026)	0.011(0.009)	0.035(0.032)	0.097(0.218)	0.008(0.045)	0.014(0.011)	0.041(0.067)	0.091(0.043)
β_{23}	0.003(0.004)	0.007(0.006)	0.042(0.290)	0.088(0.672)	0.005(0.012)	0.009(0.013)	0.038(0.211)	0.094(0.051)
β_{14}	0.005(0.003)	0.013(0.025)	0.065(0.632)	0.079(0.349)	0.004(0.008)	0.015(0.031)	0.059(0.041)	0.084(0.067)
β_{24}	0.003(0.005)	0.012(0.013)	0.045(0.581)	0.088(0.786)	0.005(0.028)	0.015(0.022)	0.048(0.078)	0.093(0.038)
Total	0.081(0.092)	0.123(0.182)	0.213(0.425)	0.444(0.521)	0.127(0.052)	0.267(0.416)	0.495(1.780)	1.578(0.952)
Error 5								
IMSE								
$\mu(t)$	0.054(0.023)	0.178(0.013)	0.224(0.019)	0.415(0.146)	0.051(0.010)	0.176(0.036)	0.226(0.101)	0.413(0.122)
$\beta_{11}(t)$	0.027(0.032)	0.123(0.217)	0.213(0.018)	0.327(0.513)	0.027(0.006)	0.124(0.018)	0.210(0.027)	0.325(0.061)
$\beta_{21}(t)$	0.014(0.021)	0.114(0.120)	0.157(0.019)	0.215(0.765)	0.012(0.007)	0.115(0.021)	0.156(0.030)	0.214(0.044)
$\beta_{12}(t)$	0.032(0.024)	0.206(0.341)	0.217(0.015)	0.412(0.654)	0.033(0.008)	0.204(0.025)	0.216(0.033)	0.410(0.065)
$\beta_{13}(t)$	0.019(0.038)	0.109(0.222)	0.199(0.039)	0.277(0.312)	0.017(0.011)	0.108(0.043)	0.200(0.073)	0.275(0.077)
MSE								
β_{22}	0.013(0.023)	0.028(0.051)	0.076(0.045)	0.134(0.123)	0.011(0.029)	0.035(0.049)	0.083(0.054)	0.139(0.111)
β_{23}	0.048(0.021)	0.068(0.025)	0.116(0.058)	0.189(0.325)	0.054(0.039)	0.074(0.043)	0.123(0.071)	0.197(0.048)
β_{14}	0.018(0.024)	0.041(0.029)	0.098(0.034)	0.187(0.256)	0.022(0.045)	0.051(0.045)	0.111(0.178)	0.195(0.123)
β_{24}	0.021(0.013)	0.099(0.034)	0.117(0.067)	0.210(0.111)	0.025(0.061)	0.111(0.243)	0.126(0.078)	0.208(0.321)
Total	0.185(0.028)	0.214(0.056)	0.312(0.124)	0.459(0.135)	0.265(0.025)	0.541(0.114)	0.823(0.268)	1.084(0.480)

3.5 Real Data Analysis

Asthma is a chronic respiratory disease characterized by lung inflammation and reversible airflow obstruction, leading to breathing difficulties. According to the Centers for Disease Control and Prevention (CDC), over 25 million Americans suffer from asthma, including 7.7 percent of adults and 8.4 percent of children ([for Disease Control et al. Akinbami \(2006\)](#)). Asthma is the most common chronic disease among children. Our case study analyzes data from the Childhood Asthma Management Program (CAMP) ([Group et al. \(1999\)](#), [Sze-fler et al. \(2000\)](#), [Covar et al. \(2012b\)](#)), utilizing SNP and phenotype datasets (accession pht000701.v1.p1) that have been downloaded and pre-processed. We selected subjects aged 5 to 12 years diagnosed with chronic asthma and monitored them over four years. Before treatment, there were three visits spaced 1-2 weeks apart. After treatment, thirteen visits were conducted: the first two visits two months apart, followed by twelve visits four months apart, which we focused on in our study. Two types of treatments, Budesonide and Nedocromil, were administered to 30% of the subjects, while the remainder received a placebo. We considered treatment, age, and gender as environmental factors. The phenotype of interest, forced expiratory volume in one second (FEV1), which measures the total volume of air expelled from the lungs in one second, was repeatedly measured during each visit. Each subject’s genotype data included over nine hundred thousand SNPs. We matched genotypes with phenotypes using subject IDs, removed SNPs with a minor allele frequency (MAF) less than 0.05 or those deviating from Hardy-Weinberg equilibrium, and obtained a working dataset of 438 Caucasian subjects and 447,850 SNPs.

For computational ease in studies involving extremely high-dimensional data, like Genome Wide Association Studies (GWAS) and multi-omics integration studies, it is necessary to conduct marginal feature prescreening first. This allows for the application of regularization on datasets of a more manageable size. ([Fan and Lv \(2010\)](#), [Wu et al. \(2019\)](#)) For example ([Li et al. \(2015\)](#), [Jiang et al. \(2015\)](#), [Wu et al. \(2014\)](#)) have utilized single SNP analysis for prescreening prior to implementing their proposed variable selection methods in longitudinal and multivariate GWAS studies, respectively. In this case, we utilize a marginal time varying

coefficient model with FEV1 as the response variable to filter SNPs. The predictors in the marginal model include a time-varying intercept and the time-varying effects of both the additive and dominant effects for a single SNP. The proposed RBVSS-SI method was utilized to identify the significant additive and dominant effects. Additive or dominant effects with a standard deviation of the estimated varying coefficients greater than 0.2 are retained. This screening process resulted in 400 additive and dominant effects passing the filter.

We analyzed the data using the proposed method, along with RBVSS, BVSS-SI, and BVSS. The RBVSS method lacks structural identification, while BVSS-SI and BVSS do not account for robustness. Methods RBV-SI, RBV, BV-SI and BV demonstrated inferior performance in simulations and were therefore excluded from the real data analysis. The proposed method identified 4 SNPs with constant additive effects and 5 SNPs with constant dominant effects, along with 23 SNPs displaying varying additive effects and 22 SNPs exhibiting varying dominant effects. In contrast, RBVSS identified 25 SNPs with varying additive effects and 30 SNPs with varying dominant effects. BVSS-SI identified 3 SNPs with constant additive effects and 3 SNPs with constant dominant effects, as well as 16 SNPs with varying additive effects and 15 SNPs with varying dominant effects. BVSS identified 19 SNPs with varying additive effects and 20 SNPs with varying dominant effects. The identification results for varying and constant effects are summarized in Table 3.4 and Table 3.5. In Table 3.4, it is clear that the four SNPs (rs10085668, rs12474254, and rs17247848) identified as having constant additive effects only by RBVSS-SI are also selected by RBVSS. Similarly, Table 3.5 shows that the five SNPs (rs2736100, rs10085668, rs7294610, rs11852876, and rs17247848) recognized as having constant dominant effects exclusively by RBVSS-SI are also selected by RBVSS. However, due to the absence of structural identification, RBVSS categorizes them as SNPs with varying effects. Additionally, due to a lack of robustness, BVSS-SI and BVSS identify fewer constant and varying effects, or only varying effects, for both additive and dominant effects. Out of all the additive effects selected by the proposed RBVSS-SI method, 3 were uniquely identified, not detected by other methods. These SNPs are associated with the genes CYP1A1, FCER1A and ZNRF3. In addition, out of all the dominant effects selected by the proposed RBVSS-SI method, 1 was uniquely identified, not

detected by other methods. The SNP is associated with the gene TBX21. Figure 3.3 presents the estimated coefficient functions for the varying intercept and the gender effect. The intercept shows an increasing trend between ages 6 and 10 and 14 to 16, while it decreases between ages 10 and 14, reflecting the increase in FEV1 as lung development and treatments take effect over time. Regarding the gender effect, the decline before age 10 is attributed to young females growing faster than young males prior to age 12. The rise between ages 10 and 14 suggests that boys begin to develop more lung capacity along with other body functions. There is also a decreasing trend observed between ages 14 and 16. The 23 additive and 22 dominant varying coefficients of age identified by RBVSS-SI are presented in Figure 3.4 and Figure 3.5.

We wrap up this section by analyzing the heritability identified by the models and the heritability attributed to individual SNPs. Heritability refers to the proportion of trait variation that can be attributed to genetic factors. Evaluating heritability is crucial for comprehending the genetic structure of complex diseases and is a common practice in genome-wide association studies (GWAS). Since we are selecting a relatively small subset of SNPs, the heritability discussed here represents only the heritability attributable to our submodel, rather than the overall heritability of the disease. This is particularly important because it allows us to more directly observe how heritability changes with age. We assess both the overall heritability of all selected SNPs and the heritability of individual SNPs.

For the methods with robustness, the total heritability of all SNPs is calculated using the formula:

$$H(\text{FEV}) = \frac{\text{RSAD}(\text{FEV}|\text{Gender})}{\text{RSAD}(\text{FEV}|\text{Gender})} - \frac{\text{RSAD}(\text{FEV}|\text{Gender}, \text{SNP}_1, \dots, \text{SNP}_p)}{\text{RSAD}(\text{FEV}|\text{Gender})}$$

Here, the RSAD (Residual Sum of Absolute Deviation) is defined as the sum of the absolute deviated residuals, given by:

$$\text{RSAD} = \sum_{i=1}^n \sum_{j=1}^{J_i} |y_i(t_{ij}) - \hat{y}_i(t_{ij})|$$

For the methods without robustness, the total heritability of all SNPs is calculated using the formula:

$$H(\text{FEV}) = \frac{\text{RSS}(\text{FEV}|\text{Gender})}{\text{RSS}(\text{FEV}|\text{Gender})} - \frac{\text{RSS}(\text{FEV}|\text{Gender}, \text{SNP}_1, \dots, \text{SNP}_p)}{\text{RSS}(\text{FEV}|\text{Gender})}$$

Here, the RSS (Residual Sum of Squares) is defined as the sum of the squared residuals, given by:

$$\text{RSS} = \sum_{i=1}^n \sum_{j=1}^{J_i} (y_i(t_{ij}) - \hat{y}_i(t_{ij}))^2$$

where $\hat{y}_i(t_{ij})$ represents the fitted value obtained from the model using the posterior median estimates, which takes into account the fixed effects and the random effects. The total heritability for RBVSS-SI is 41.246%, which is noticeably higher than that of RBVSS (36.786%), BVSS-SI (33.542%), and BVSS (29.318%). Additionally, we estimate the time-varying heritability for all SNPs.

Finally, we calculate the heritability of individual SNPs. This is determined by the order in which each SNP is selected into the model during the stepwise selection procedure. Let $\text{SNP}(k)$ be the k -th SNP to be selected into the model; its heritability is then calculated by:

$$H(\text{SNP}(k)) = \frac{\text{RSS}(\text{FEV}|\text{Gender}, \text{SNP}(1), \dots, \text{SNP}(k-1))}{\text{RSS}(\text{FEV}|\text{Gender})} - \frac{\text{RSS}(\text{FEV}|\text{Gender}, \text{SNP}(1), \dots, \text{SNP}(k-1), \text{SNP}(k))}{\text{RSS}(\text{FEV}|\text{Gender})}$$

Tables 3.7 and 3.8 display the heritability of individual SNPs for both additive and dominant effects using the proposed method. The heritability of the SNPs ranges quite evenly between 2 and 8 percent.

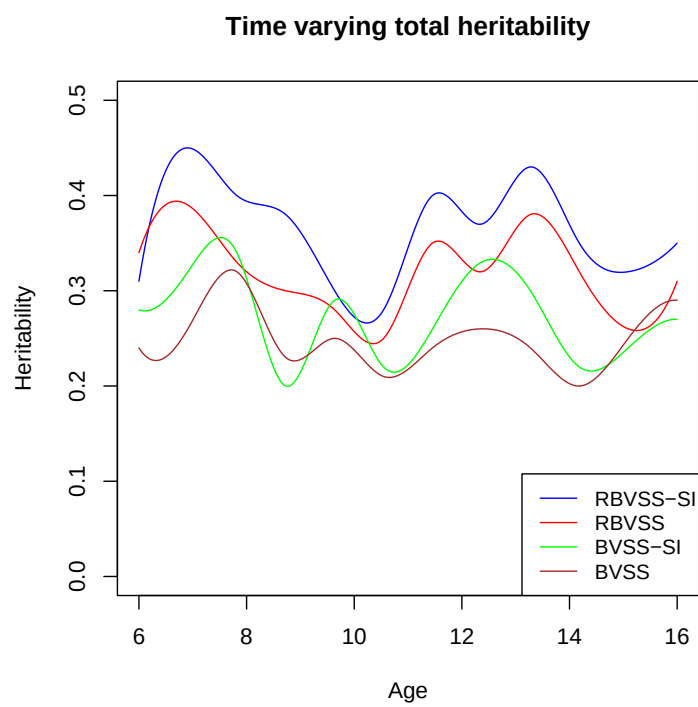


Figure 3.2: *Time-varying total heritability*

Table 3.4: *Identification results for varying and constant effects for the additive effects.*

SNP	Gene	RBVSS-SI		RBVSS	BVSS-SI		BVSS
		V(Age)	C	V(Age)	V(Age)	C	V(Age)
rs7750817	CDKN2B		2.488	Varying			
rs10085668	CILP2		-1.058	Varying			Varying
rs2078323	CD36	Varying		Varying		-1.156	
rs1986953	INHBB	Varying		Varying			Varying
rs10208874	KCNJ2	Varying		Varying	Varying		Varying
rs12472552	RPH3A	Varying		Varying			Varying
rs2108622	CYP4F3	Varying		Varying	Varying		Varying
rs9303277	IL33	Varying		Varying			
rs13103603	LEPR	Varying			Varying		
rs11943130	ESR1	Varying		Varying	Varying		Varying
rs1048943	CYP1A1	Varying					
rs12539562	CDKAL1	Varying		Varying		2.761	Varying
rs10108740	GLI2	Varying		Varying	Varying		Varying
rs2427828	FCER1A	Varying					
rs12474254	IL18R1		2.973	Varying		1.989	Varying
rs16923238	ZNRF3	Varying					
rs4837929	NCAM1	Varying		Varying	Varying		Varying
rs10828439	TRIB1	Varying		Varying			
rs10783445	PTGER4	Varying		Varying	Varying		Varying
rs1109760	FTO	Varying		Varying	Varying		Varying
rs1557075	IRF1	Varying		Varying			
rs4770659	LCORL	Varying		Varying			
rs4842910	TRIB1	Varying		Varying	Varying		Varying
rs10937405	TP63	Varying		Varying			
rs2825653	GALNT10	Varying		Varying	Varying		Varying
rs12170951	SLC22A5	Varying		Varying	Varying		Varying
rs11852876	TBX5			Varying	Varying		Varying
rs17247848	TCF7L2		2.345	Varying	Varying		Varying
rs135058	CYP2A6			Varying	Varying		Varying

Table 3.5: *Identification results for varying and constant effects for dominant effects.*

SNP	Gene	RBVSS-SI		RBVSS	BVSS-SI		BVSS
		V(Age)	C	V(Age)	V(Age)	C	V(Age)
rs1986953	INHBB			Varying	Varying		Varying
rs4645225	CYP1A2			Varying			
rs2736100	TERT		2.463	Varying			
rs10085668	CILP2		-0.940	Varying			Varying
rs6679930	ANK3	Varying		Varying	Varying		Varying
rs11207487	KLF7	Varying		Varying	Varying		Varying
rs2427828	FCER1A	Varying		Varying			
rs2108622	CYP4F3	Varying		Varying	Varying		Varying
rs9303277	IL33	Varying			Varying		Varying
rs10167668	LPA	Varying		Varying		-1.235	Varying
rs12474254	IL18R1	Varying		Varying			
rs7640124	SORT1	Varying		Varying			
rs2321761	PPARG	Varying		Varying	Varying		Varying
rs13103603	FAM78B	Varying		Varying			
rs2427837	TBX21	Varying					
rs10108740	GLI2	Varying		Varying	Varying		Varying
rs7845059	CPED1			Varying			
rs12675273	BCAR1	Varying		Varying			
rs10937405	TP63	Varying		Varying			
rs1339760	DISP1	Varying		Varying	Varying		Varying
rs16923238	ZNRF3	Varying		Varying	Varying		Varying
rs4837929	NCAM1	Varying		Varying	Varying		Varying
rs10122210	TRIB1			Varying	Varying		Varying
rs10828439	TRIB1	Varying		Varying			Varying
rs2135829	CFH	Varying		Varying	Varying		Varying
rs17066875	ZHX2	Varying		Varying			
rs4842910	TRIB1	Varying		Varying		4.081	Varying
rs12170951	SLC22A5	Varying		Varying	Varying		Varying
rs7294610	MYO1C		3.027	Varying		2.991	Varying
rs1718375	PTPRJ			Varying	Varying		
rs11852876	TBX5		3.555	Varying			
rs17247848	TCF7L2		3.395	Varying			Varying

Table 3.6: *The numbers of varying effects and constant effects identified by different approaches and their overlaps.*

Additive	Varying				Constant			
	RBVSS-SI	RBVSS	BVSS-SI	BVSS	RBVSS-SI	RBVSS	BVSS-SI	BVSS
RBVSS-SI	23	19	11	13	4	0	1	0
RBVSS		25	13	18		0	0	0
BVSS-SI			14	13			3	0
BVSS				18				0
Dominant	Varying				Constant			
	RBVSS-SI	RBVSS	BVSS-SI	BVSS	RBVSS-SI	RBVSS	BVSS-SI	BVSS
RBVSS-SI	22	20	11	15	5	0	1	0
RBVSS		30	13	18		0	0	0
BVSS-SI			14	12			3	0
BVSS				19				0

Table 3.7: *Heritability of single additive effects SNPs by RBVSS-SI*

SNP	Gene	Heritability
rs7750817	CDKN2B	5.475%
rs10085668	CILP2	6.128%
rs2078323	CD36	4.189%
rs1986953	INHBB	7.137%
rs10208874	KCNJ2	3.912%
rs12472552	RPH3A	2.998%
rs2108622	CYP4F3	5.321%
rs9303277	IL33	4.921%
rs13103603	LEPR	2.768%
rs11943130	ESR1	2.589%
rs1048943	CYP1A1	3.058%
rs12539562	CDKAL1	2.194%
rs10108740	GLI2	3.139%
rs2427828	FCER1A	8.159%
rs12474254	IL18R1	6.987%
rs16923238	ZNRF3	4.555%
rs4837929	NCAM1	1.987%
rs10828439	TRIB1	4.324%
rs10783445	PTGER4	2.005%
rs1109760	FTO	3.167%
rs1557075	IRF1	4.196%
rs4770659	LCORL	1.897%
rs4842910	TRIB1	2.578%
rs10937405	TP63	2.957%
rs2825653	GALNT10	7.773%
rs12170951	SLC22A5	5.578%
rs17247848	TCF7L2	3.112%

Table 3.8: *Heritability of single dominant effects SNPs by RBVSS-SI*

SNP	Gene	Heritability
rs2736100	TERT	6.147%
rs10085668	CILP2	5.569%
rs6679930	ANK3	3.776%
rs11207487	KLF7	3.895%
rs2427828	FCER1A	2.176%
rs2108622	CYP4F3	2.097%
rs9303277	IL33	3.326%
rs10167668	LPA	1.868%
rs1247254	IL18R1	1.999%
rs7640124	SORT1	7.436%
rs2321761	PPARG	6.914%
rs13103603	FAM78B	5.487%
rs2427837	TBX21	4.335%
rs10108740	GLI2	6.113%
rs12675273	BCAR1	1.898%
rs10937405	TP63	3.788%
rs1339760	DISP1	4.234%
rs16923238	ZNRF3	2.789%
rs4837929	NCAM1	2.559%
rs10828439	TRIB1	3.437%
rs2135829	CFH	3.662%
rs17066875	ZHX2	8.289%
rs4842910	TRIB1	5.987%
rs12170951	SLC22A5	2.786%
rs7294610	MYO1C	4.685%
rs11852876	TBX5	3.472%
rs17247848	TCF7L2	5.011%

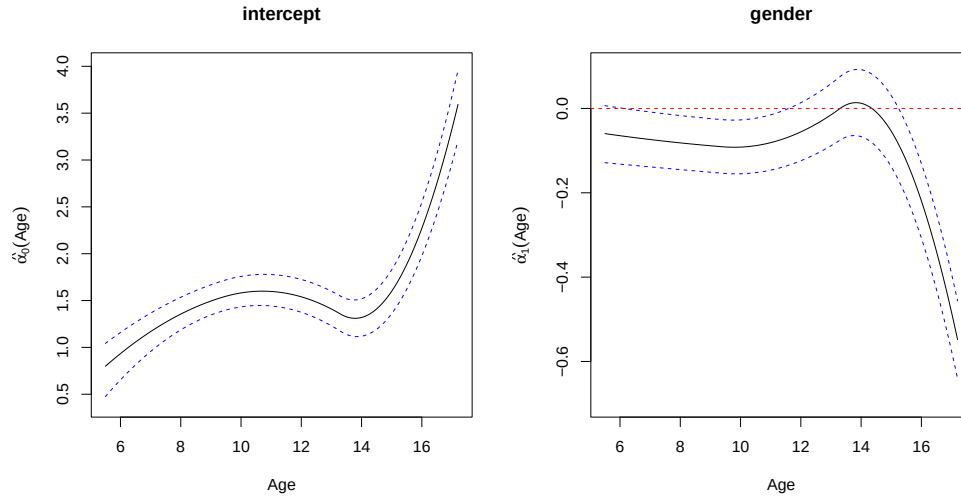


Figure 3.3: Real data analysis for the proposed method (RBVSS-SI). Black line: median estimates of varying intercept and varying coefficients for gender for RBVSS-SI. Blue dashed lines: 95% credible intervals for the estimated varying intercept and varying coefficients for gender.

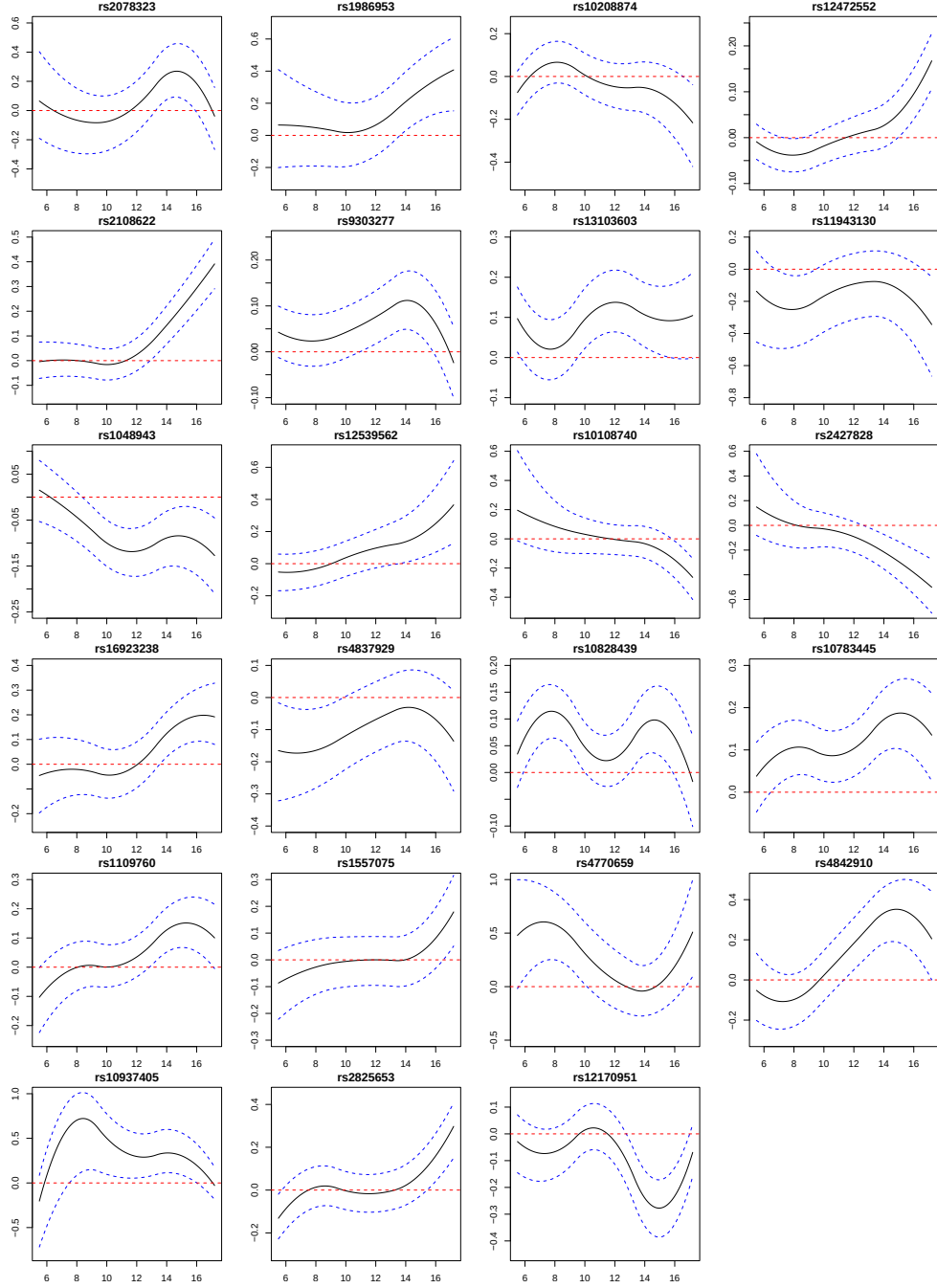


Figure 3.4: Real data analysis for the proposed method (RBVSS-SI). Black line: median estimates of varying additive effects coefficients for RBVSS-SI. Blue dashed lines: 95% credible intervals for the estimated varying additive effects coefficients.

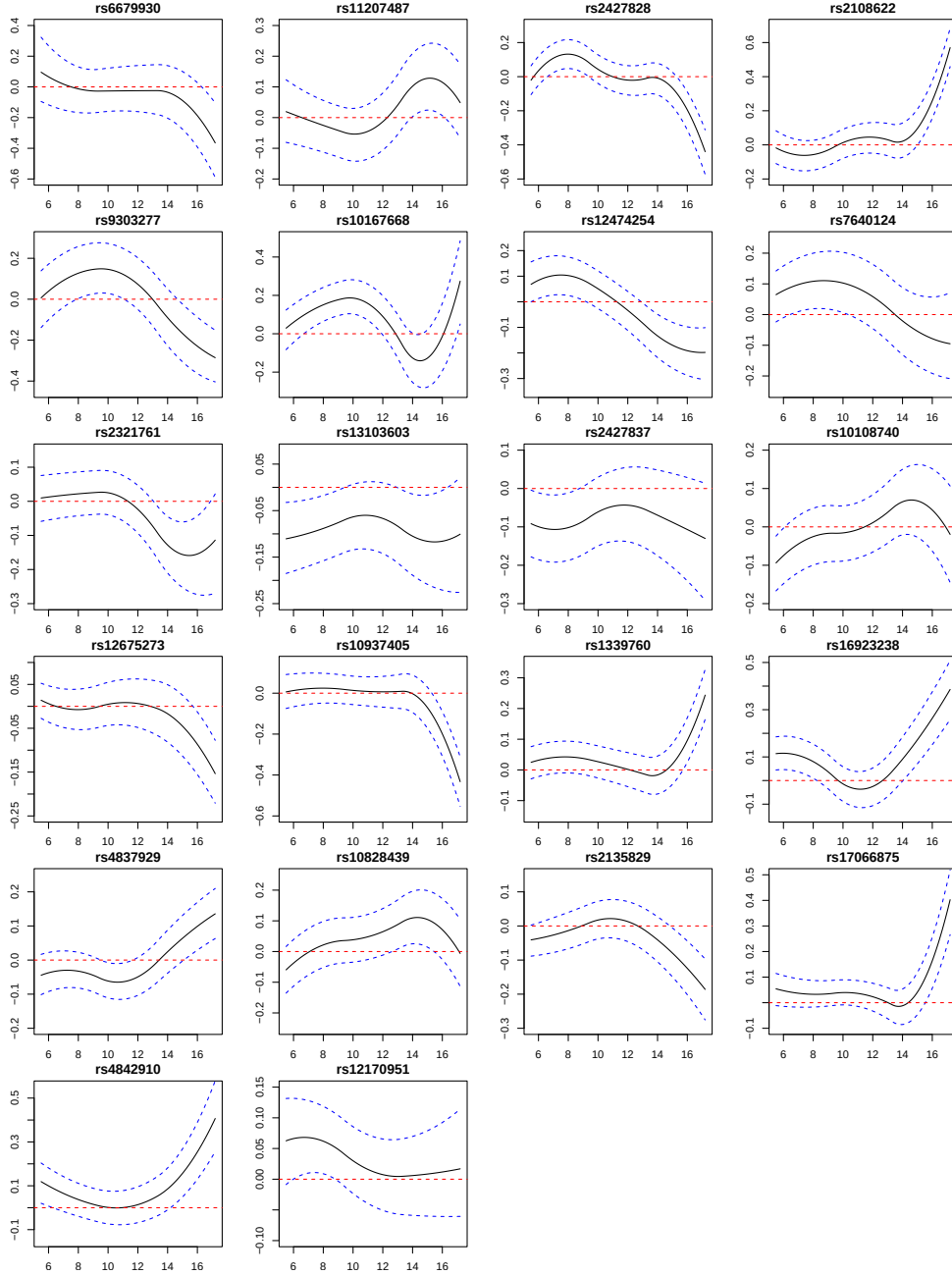


Figure 3.5: Real data analysis for the proposed method (RBVSS-SI). Black line: median estimates of varying dominant effects coefficients for RBVSS-SI. Blue dashed lines: 95% credible intervals for the estimated varying dominant effects coefficients.

Bibliography

- Robert Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, 58(1):267–288, 1996.
- Jianqing Fan and Runze Li. Variable selection via nonconcave penalized likelihood and its oracle properties. *Journal of the American Statistical Association*, 94(456):1348–1360, 2001. doi: 10.1198/016214501753382273.
- CunHui Zhang. Nearly unbiased variable selection under minimax concave penalty. *The Annals of Statistics*, 38(2):894–942, 2010. doi: 10.1214/09-AOS729.
- Hui Zou. The adaptive lasso and its oracle properties. *Journal of the American statistical association*, 101:1418–1429, 2006.
- Ming Yuan and Yi Lin. Model selection and estimation in regression with grouped variables. *J. R. Statist. Soc. B*, pages 49–67, 2006a.
- Jian Huang, Patrick Breheny, and Shuangge Ma. A selective review of group selection in high-dimensional models. *Statistical Science*, 27:481–499, 2012.
- Cen Wu and Shuangge Ma. A selective review of robust variable selection with applications in bioinformatics. *Briefings in Bioinformatics*, 16(5):873–883, 2015a. doi: 10.1093/bib/bbu046.
- Qing Li, Ruibin Xi, and Nan Lin. Bayesian regularized quantile regression. *Bayesian Analysis*, 5(3):533–556, 2010a. doi: 10.1214/10-BA521.
- Jian Huang, Shuangge Ma, Hongzhe Li, and Cun-Hui Zhang. The sparse laplacian shrinkage estimator for high-dimensional regression. *Annals of statistics*, 39:2021, 2011.

- Cen Wu, Yu Jiang, Jie Ren, Yuehua Cui, and Shuangge Ma. Dissecting gene-environment interactions: A penalized robust approach accounting for hierarchical structures. *Statistics in medicine*, 37:437–456, 2018a.
- Trevor Park and George Casella. The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686, 2008a. doi: 10.1198/016214508000000337.
- Minjung Kyung, Jeff Gill, Malay Ghosh, and George Casella. Penalized regression, standard errors, and bayesian lassos. *Bayesian Anal.*, 5(2):369–411, 2010. doi: 10.1214/10-BA607.
- T. J. Mitchell and J. J. Beauchamp. Bayesian variable selection in linear regression. *Journal of the American Statistical Association*, 83(404):1023–1032, 1988.
- Xiaofan Xu and Malay Ghosh. Bayesian variable selection and estimation for group lasso. *Bayesian Anal.*, 10(4):909–936, 2015. doi: 10.1214/14-BA929.
- Geert Verbeke, Steffen Fieuws, Geert Molenberghs, and Marie Davidian. The analysis of multivariate longitudinal data: a review. *Statistical methods in medical research*, 23:42–59, 2014a.
- S Bandyopadhyay, B Ganguli, and A Chatterjee. A review of multivariate longitudinal data analysis. *Statistical methods in medical research*, 20:299–330, 2011a.
- Kung-Yee Liang and Scott L Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, pages 13–22, 1986a.
- Hui Zou and Trevor Hastie. Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(2):301–320, 2005.
- Robert Tibshirani, Michael Saunders, Saharon Rosset, Ji Zhu, and Keith Knight. Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(1):91–108, 2005.

- Cen Wu, Qingzhao Zhang, Yu Jiang, and Shuangge Ma. Robust network-based analysis of the associations between (epi) genetic measurements. *Journal of multivariate analysis*, 168:119–130, 2018b.
- Jie Ren, Tao He, Ye Li, Sai Liu, Yinhao Du, Yu Jiang, and Cen Wu. Network-based regularization for high dimensional snp data in the case–control study of type 2 diabetes. *BMC Genetics*, 18:1–12, 2017.
- Jie Ren, Yinhao Du, Shaoyu Li, Shuangge Ma, Yu Jiang, and Cen Wu. Robust network-based regularization and variable selection for high-dimensional genomic data in cancer prognosis. *Genetic epidemiology*, 43(3):276–291, 2019.
- Cen Wu and Yuehua Cui. Boosting signals in gene-based association studies via efficient snp selection. *Briefings in Bioinformatics*, 15(2):279–291, 2014.
- Yu Jiang, Yuan Huang, Yinhao Du, Yinjun Zhao, Jie Ren, Shuangge Ma, and Cen Wu. Identification of prognostic genes and pathways in lung adenocarcinoma using a bayesian approach. *Cancer informatics*, 16:1176935116684825, 2017.
- Cen Wu, Shaoyu Li, and Yuehua Cui. Genetic association studies: an information content perspective. *Current genomics*, 13(7):566–573, 2012.
- Fei Zhou, Jie Ren, Xi Lu, Shuangge Ma, and Cen Wu. Gene–environment interaction: A variable selection perspective. *Epistasis: Methods and Protocols*, pages 191–223, 2021.
- Cen Wu and Shuangge Ma. A selective review of robust variable selection with applications in bioinformatics. *Briefings in bioinformatics*, 16(5):873–883, 2015b.
- Cen Wu and Yuehua Cui. A novel method for identifying nonlinear gene–environment interactions in case–control association studies. *Human genetics*, 132:1413–1425, 2013.
- Lifeng Wang, Hongzhe Li, and Jianhua Z Huang. Variable selection in nonparametric varying-coefficient models for analysis of repeated measurements. *Journal of the American Statistical Association*, 103(484):1556–1569, 2008.

- Cen Wu, Yu Jiang, Jie Ren, Yuehua Cui, and Shuangge Ma. Dissecting gene-environment interactions: A penalized robust approach accounting for hierarchical structures. *Statistics in medicine*, 37(3):437–456, 2018c.
- Cen Wu, Ping-Shou Zhong, and Yuehua Cui. Additive varying-coefficient model for nonlinear gene-environment interactions. *Statistical applications in genetics and molecular biology*, 17(2):20170008, 2018d.
- Yanlin Tang, Huixia Judy Wang, and Zhongyi Zhu. Variable selection in quantile varying coefficient models with longitudinal data. *Computational Statistics & Data Analysis*, 57(1):435–449, 2013.
- Jie Ren, Fei Zhou, Xiaoxi Li, Qi Chen, Hongmei Zhang, Shuangge Ma, Yu Jiang, and Cen Wu. Semiparametric bayesian variable selection for gene-environment interactions. *Statistics in medicine*, 39(5):617–638, 2020.
- Geert Verbeke, Steffen Fieuws, Geert Molenberghs, and Marie Davidian. The analysis of multivariate longitudinal data: a review. *Statistical methods in medical research*, 23(1):42–59, 2014b.
- S Bandyopadhyay, B Ganguli, and A Chatterjee. A review of multivariate longitudinal data analysis. *Statistical methods in medical research*, 20(4):299–330, 2011b.
- David J Hunter. Gene–environment interactions in human diseases. *Nature reviews genetics*, 6(4):287–298, 2005.
- Peter Kraft, Yu-Chun Yen, Daniel O Stram, John Morrison, and W James Gauderman. Exploiting gene-environment interaction to detect genetic associations. *Human heredity*, 63(2):111–119, 2007.
- Brenee S King, Lizhi Lu, Miao Yu, Yu Jiang, Joseph Standard, Xiaoyu Su, Zhihui Zhao, and Weiqun Wang. Lipidomic profiling of di-and tri-acylglycerol species in weight-controlled mice. *Plos one*, 10(2):e0116398, 2015.

- Ming Yuan and Yi Lin. Model selection and estimation in regression with grouped variables. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 68(1):49–67, 2006b.
- Kung-Yee Liang and Scott L Zeger. Longitudinal data analysis using generalized linear models. *Biometrika*, 73(1):13–22, 1986b.
- Lan Wang, Jianhui Zhou, and Annie Qu. Penalized generalized estimating equations for high-dimensional longitudinal data analysis. *Biometrics*, 68(2):353–360, 2012.
- Yingying Fan and Runze Li. Variable selection in linear mixed effects models. *Annals of statistics*, 40(4):2043, 2012.
- Yali Fan, Guoyou Qin, and Zhongyi Zhu. Variable selection in robust regression models for longitudinal data. *Journal of Multivariate Analysis*, 109:156–167, 2012a.
- Jing Lv, Hu Yang, and Chaohui Guo. An efficient and robust variable selection method for longitudinal generalized linear models. *Computational Statistics & Data Analysis*, 82: 74–88, 2015.
- Ruzong Fan, Paul S Albert, and Enrique F Schisterman. A discussion of gene-gene and gene-environment interactions and longitudinal genetic analysis of complex traits. *Statistics in medicine*, 31(22):2565, 2012b.
- Shujie Ma, Qiongxia Song, and Li Wang. Simultaneous variable selection and estimation in semiparametric modeling of longitudinal/clustered data. *Bernoulli*, pages 252–274, 2013.
- Fei Zhou, Jie Ren, Gengxin Li, Yu Jiang, Xiaoxi Li, Weiqun Wang, and Cen Wu. Penalized variable selection for lipid–environment interactions in a longitudinal lipidomics study. *Genes*, 10(12):1002, 2019.
- Fei Zhou, Xi Lu, Jie Ren, Kun Fan, Shuangge Ma, and Cen Wu. Sparse group variable selection for gene–environment interactions in the longitudinal study. *Genetic Epidemiology*, 46(5-6):317–340, 2022a.

- Keming Yu and Rana A Moyeed. Bayesian quantile regression. *Statistics & Probability Letters*, 54(4):437–447, 2001.
- George Casella, Malay Ghosh, Jeff Gill, and Minjung Kyung. Penalized regression, standard errors, and bayesian lassos. *Bayesian Analysis*, 06(2):369–411, 2010.
- Edward I George and Robert E McCulloch. Variable selection via gibbs sampling. *Journal of the American Statistical Association*, 88(423):881–889, 1993.
- Hemant Ishwaran and J Sunil Rao. Spike and slab variable selection: frequentist and bayesian strategies. *The Annals of Statistics*, 33(2):730 – 773, 2005.
- Veronika Ročková and Edward I George. The spike-and-slab lasso. *Journal of the American Statistical Association*, 113(521):431–444, 2018.
- Kun Fan and Cen Wu. *mixedBayes: Bayesian Longitudinal Regularized Quantile Mixed Model*. URL <https://CRAN.R-project.org/package=mixedBayes>. R package version 0.1.0.
- Jie Ren, Fei Zhou, Xiaoxi Li, Shuangge Ma, Yu Jiang, and Cen Wu. Robust bayesian variable selection for gene–environment interactions. *Biometrics*, 79(2):684–694, 2023.
- Hideo Kozumi and Genya Kobayashi. Gibbs sampling methods for bayesian quantile regression. *Journal of statistical computation and simulation*, 81(11):1565–1578, 2011.
- Trevor Park and George Casella. The bayesian lasso. *Journal of the American Statistical Association*, 103(482):681–686, 2008b.
- Qing Li, Nan Lin, and Ruibin Xi. Bayesian regularized quantile regression. *Bayesian Analysis*, 5(3):533–556, 2010b.
- Fei Zhou, Jie Ren, Shuangge Ma, and Cen Wu. The bayesian regularized quantile varying coefficient model. *Computational Statistics & Data Analysis*, page 107808, 2023.

- Maria Maddalena Barbieri and James O Berger. Optimal predictive model selection. *The Annals of Statistics*, 32(3):870–897, 2004.
- Maria M Barbieri, James O Berger, Edward I George, and Veronika Ročková. The median probability model and correlated variables. *Bayesian Analysis*, 16(4):1085–1112, 2021.
- Jiahua Li, Zhong Wang, Runze Li, and Rongling Wu. Bayesian group lasso for nonparametric varying-coefficient models with application to functional genome-wide association studies. *The annals of applied statistics*, 9(2):640, 2015.
- Andrew Gelman and Donald B Rubin. Inference from iterative simulation using multiple sequences. *Statistical science*, pages 457–472, 1992.
- Stephen P Brooks and Andrew Gelman. General methods for monitoring convergence of iterative simulations. *Journal of computational and graphical statistics*, 7(4):434–455, 1998.
- Andrew Gelman, John B Carlin, Hal S Stern, David B Dunson, and Aki Vehtari. Rubin db. bayesian data analysis. *Chapman & Hall*, 2004.
- Ping Ouyang, Yu Jiang, Hieu M Doan, Linglin Xie, David Vasquez, Ruth Welti, Xiaoyu Su, Nanyan Lu, Betty Herndon, Shie-Shien Yang, et al. Weight loss via exercise with controlled dietary intake may affect phospholipid profile for cancer prevention in murine skin tissues. *Cancer Prevention Research*, 3(4):466–477, 2010.
- R. M. Cantor, K. Lange, and J. S. Sinsheimer. Prioritizing gwas results: a review of statistical methods and recommendations for their application. *The American Journal of Human Genetics*, 86(1):6–22, 2010.
- Childhood Asthma Management Program Research Group. The childhood asthma management program (camp): design, rationale, and methods. *Controlled Clinical Trials*, 20(1): 91–120, 1999.
- Ronald A. Covar, Anne L. Fuhlbrigge, Paula Williams, Henry W. Kelly, and Childhood Asthma Management Program Research Group. The childhood asthma management pro-

- gram (camp): contributions to the understanding of therapy and the natural history of childhood asthma. *Current Respiratory Care Reports*, 1:243–250, 2012a.
- Joseph G Ibrahim, Hongtu Zhu, Ramon I Garcia, and Ruixin Guo. Fixed and random effects selection in mixed effects models. *Biometrics*, 67(2):495–503, 2011.
- Larry Schumaker. *Spline functions: basic theory*. Cambridge university press, 2007.
- Keming Yu and Jin Zhang. A three-parameter asymmetric laplace distribution and its extension. *Communications in Statistics—Theory and Methods*, 34(9-10):1867–1879, 2005.
- Jianhua Z Huang, Colin O Wu, and Lan Zhou. Varying-coefficient models and basis function approximations for the analysis of repeated measurements. *Biometrika*, 89(1):111–128, 2002.
- C Clark Cockerham. An extension of the concept of partitioning hereditary variance for analysis of covariances among relatives when epistasis is present. *Genetics*, 39(6):859, 1954.
- Centers for Disease Control, Prevention, et al. Data, statistics, and surveillance: asthma surveillance data [internet]. atlanta (ga): Centers for disease control and prevention; 2020 [cited 2020 jan 23].
- Omolara Jean Akinbami. *The state of childhood asthma: United states, 1980-2005*. Number 381. US Department of Health and Human Services, Centers for Disease Control and . . . , 2006.
- Childhood Asthma Management Program Research Group et al. The childhood asthma management program (camp): design, rationale, and methods. *Controlled clinical trials*, 20(1):91–120, 1999.
- Stanley Szeffler, Scott Weiss, James Tonascia, N Franklin Adkinson, Bruce Bender, Reuben Cherniack, Michele Donithan, H William Kelly, Joseph Reisman, Gail G Shapiro, et al.

- Long-term effects of budesonide or nedocromil in children with asthma. *The New England journal of medicine*, 343(15):1054–1063, 2000.
- Ronina A Covar, Anne L Fuhlbrigge, Paul Williams, H William Kelly, and Childhood Asthma Management Program Research Group. The childhood asthma management program (camp): contributions to the understanding of therapy and the natural history of childhood asthma. *Current respiratory care reports*, 1:243–250, 2012b.
- Jianqing Fan and Jinchi Lv. A selective overview of variable selection in high dimensional feature space. *Statistica Sinica*, 20(1):101, 2010.
- Cen Wu, Fei Zhou, Jie Ren, Xiaoxi Li, Yu Jiang, and Shuangge Ma. A selective review of multi-level omics data integration using variable selection. *High-throughput*, 8(1):4, 2019.
- Libo Jiang, Jingyuan Liu, Xuli Zhu, Meixia Ye, Lidan Sun, Xavier Lacaze, and Rongling Wu. 2higwas: a unifying high-dimensional platform to infer the global genetic architecture of trait development. *Briefings in bioinformatics*, 16(6):905–911, 2015.
- Cen Wu, Yuehua Cui, and Shuangge Ma. Integrative analysis of gene–environment interactions under a multi-response partially linear varying coefficient model. *Statistics in medicine*, 33(28):4988–4998, 2014.
- Fei Zhou, Jie Ren, Yuwen Liu, Xiaoxi Li, Weiqun Wang, and Cen Wu. Interep: An r package for high-dimensional interaction analysis of the repeated measurement data. *Genes*, 13(3):544, 2022b.
- Fei Zhou, Jie Ren, Yuwen Liu, Xiaoxi Li, Cen Wu, and Yu Jiang. *interep: Interaction Analysis of Repeated Measure Data*. URL <https://CRAN.R-project.org/package=interep>. R package version 0.4.1.
- Annie Qu and Peter X-K Song. Assessing robustness of generalised estimating equations and quadratic inference functions. *Biometrika*, 91(2):447–459, 2004.

Garrett M Fitzmaurice, Nan M Laird, and James H Ware. *Applied longitudinal analysis*.
John Wiley & Sons, 2012.

Appendix A

Appendices for chapter 2

A.1 Summary of all models under comparison

We have provided a summary of all methods under comparison in Table [A.1](#) and [A.2](#). All the 8 mixed-effects model based Bayesian methods have been proposed for the first time and can accommodate intra-cluster(or intra-subject) correlations through random effects. Unlike the GEE (generalized estimating equation) based methods that focus on population-level estimates, the mixed-effects models are subject-level methods which capture the correlations among repeated measures through modeling the random effects. As shown in Table [A.2](#), these methods can be categorized based on three criteria: whether they include robustness, whether they utilize spike-and-slab priors, and whether they account for bi-level main and interaction effects in the repeated measures ANOVA design. The “bi-level” in Table [A.2](#) refers to both the individual level (main lipid effects) and group level (lipid by treatment interactions). Therefore, all 4 methods– RBVSS-BI, RBV-BI, BVSS-BI and BV-BI –can effectively accommodate the interaction structures that arise from the ANOVA design. The rest 4 methods ignore the interaction structure and only conduct individual level selections.

Table A.1: *Summary of all Bayesian regularized mixed models under comparison.*

Type	Method	Fixed effects	Random effects	Reference
Robust	RBVSS-BI	Robust Bayesian variable selection on both individual and group level with spike-and-slab priors	Random intercept-slope and random intercept model	Proposed for the first time
	RBVSS	Robust Bayesian variable selection on individual level with spike-and-slab priors	Random intercept-slope and random intercept model	Proposed for the first time
	RBV-BI	Robust Bayesian variable selection on both individual and group level	Random intercept-slope and random intercept model	Proposed for the first time
	RBV	Robust Bayesian variable selection on individual level	Random intercept-slope and random intercept model	Proposed for the first time
Non-robust	BVSS-BI	Bayesian variable selection on both individual and group level with spike-and-slab priors	Random intercept-slope and random intercept model	Proposed for the first time
	BVSS	Bayesian variable selection on individual level with spike-and-slab priors	Random intercept-slope and random intercept model	Proposed for the first time
	BV-BI	Bayesian variable selection on both individual and group level	Random intercept-slope and random intercept model	Proposed for the first time
	BV	Bayesian variable selection on individual level	Random intercept-slope and random intercept model	Proposed for the first time

Table A.2: *Summary of comparisons among all Bayesian regularized mixed-effects models.*

Methods	Robustness	Likelihood	Variable Selection	Structure	Identification
RBVSS-BI	Yes	Asymmetric Laplacian	Spike-and-slab prior	Bi-level	MPM
RBVSS	Yes	Asymmetric Laplacian	Spike-and-slab prior	Individual level	MPM
RBV-BI	Yes	Asymmetric Laplacian	Laplacian shrinkage	Bi-level	Credible interval
RBV	Yes	Asymmetric Laplacian	Laplacian shrinkage	Individual level	Credible interval
BVSS-BI	No	Gaussian	Spike-and-slab prior	Bi-level	MPM
BVSS	No	Gaussian	Spike-and-slab prior	Individual level	MPM
BV-BI	No	Gaussian	Laplacian shrinkage	Bi-level	Credible interval
BV	No	Gaussian	Laplacian shrinkage	Individual level	Credible interval

Bi-level: both the group and individual level.

MPM: Median Probability Model.

A.2 Additional simulation results

A.2.1 Identification results for $\theta = 0.5$ in simulation

These results also hold under the random intercept model, which can be found in Table B.2, Table B.3, Table B.7, and Table A.6 in the Appendix. In general, the methods incorporating the spike - and - slab priors have better performance than the ones without spike - and - slab priors across all the error settings. For instance, under Error 1, RBVSS - BI has higher identification than RBV - BI, with the total TP is 21.97(0.14) and the total FP is 0.02(0.23). For the main effect, the TP is 3.99(0.21) and the FP is 0.01(0.15). For the interaction effects, the TP is 17.98(0.34) and the FP is 0.01(0.31). Also, the performance of the methods with bi - level selection over the methods without bi - level selection. For instance, under Error 2, RBVSS - BI has higher performance than RBVSS, with the total TP is 21.90(0.45) and the total FP is 0.10(0.67). For the main effect, the TP is 3.95(0.42) and the FP is 0.02(0.56). For the interaction effects, the TP is 17.95(0.45) and the FP is 0.08(0.15). Under heavy tailed errors and data contamination, the robust Bayesian variable selection methods beat the nonrobust Bayesian variable selection methods. For instance, under Error 2, RBVSS - BI has higher total TP 21.90(0.45) and lower FP is 0.10(0.67) than BVSS - BI. For the main effect, the TP is 3.95(0.42) and the FP is 0.02(0.56). For the interaction effects, the TP is 17.95(0.45) and the FP is 0.08(0.15). Although the TPs for the interaction effects are comparable between RBVSS - BI and BVSS - BI, the FP for the RBVSS - BI is significant lower than the BVSS - BI. Within the heavy tailed errors and data contamination, the proposed method RBVSS - BI outperforms the other methods. Among all the methods with spike - and - slab priors, the proposed method RBVSS-BI has better performance among the alternatives for the identification under heavy tailed errors and data contamination. For instance, under the mixture Laplace error (Error 3), the total TP is 21.77(0.77) and the total FP is 0.12(0.55). For the main effect, the TP is 3.87(0.35) and the FP is 0.02(0.33). For the interaction effects, the TP is 17.90(0.55) and the FP is 0.10(0.55). Same patterns have been observed under the high dimensional setting as well in Table B.3. In terms of the estimation

accuracy, for random intercept and slope model of the 8 methods under $\theta = 0.5$ were shown in Table [B.7](#) and Table [A.6](#), In general, the proposed method has the best estimation accuracy than the other methods across all the error settings. This trend also applies under the high dimensional setting as well in Table [A.6](#).

Table A.3: Identification results for $n = 200, m = 30, p = 3$. $mean(sd)$ of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 50th quantile.

$n = 200 \quad m = 30$		Main		Interaction		Total	
$\theta = 0.5$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.99(0.21)	0.01(0.15)	17.98(0.34)	0.01(0.31)	21.97(0.14)	0.02(0.23)
N(0,1)	RBVSS	3.97(0.45)	0.07(0.25)	15.11(0.35)	0.07(0.25)	19.08(0.33)	0.14(0.35)
	RBV-BI	3.25(0.54)	0.45(0.82)	12.55(0.37)	1.95(1.77)	15.80(0.39)	2.40(1.73)
	RBV	3.20(0.26)	0.55(1.12)	10.25(0.26)	2.45(1.81)	13.45(0.28)	3.00(1.89)
	BVSS-BI	3.57(1.01)	0.03(0.18)	17.85(0.32)	0.08(0.92)	21.42(1.24)	0.11(0.92)
	BVSS	3.50(1.45)	0.05(0.24)	14.45(0.26)	0.27(0.31)	17.95(0.18)	0.32(0.25)
	BV-BI	3.02(1.15)	0.25(0.38)	11.54(0.32)	0.57(0.68)	14.56(0.45)	0.82(0.86)
	BV	3.09(0.67)	0.37(0.43)	9.75(0.47)	0.89(0.53)	12.84(0.27)	1.26(0.95)
Error 2	RBVSS – BI	3.95(0.42)	0.02(0.56)	17.95(0.45)	0.08(0.15)	21.90(0.45)	0.10(0.67)
$t(2)$	RBVSS	3.85(0.25)	0.04(0.45)	15.45(0.56)	0.45(0.82)	19.30(0.54)	0.49(0.81)
	RBV-BI	3.14(0.45)	0.23(0.85)	11.89(0.24)	2.15(0.98)	15.04(0.37)	2.38(0.97)
	RBV	3.03(0.58)	0.43(0.86)	9.95(0.37)	1.15(0.96)	12.98(0.48)	1.58(0.57)
	BVSS-BI	3.31(1.01)	0.05(0.28)	17.78(0.31)	7.24(0.33)	21.09(1.01)	7.29(0.18)
	BVSS	3.25(1.08)	0.08(0.25)	13.74(2.56)	0.31(0.47)	16.99(2.69)	0.39(0.86)
	BV-BI	2.97(1.07)	0.47(0.73)	10.43(3.82)	1.54(1.07)	13.40(4.49)	2.01(1.28)
	BV	2.75(1.05)	0.57(0.73)	8.75(3.66)	0.96(1.03)	11.50(2.12)	1.53(1.45)
Error 3	RBVSS – BI	3.87(0.35)	0.02(0.33)	17.90(0.55)	0.10(0.55)	21.77(0.77)	0.12(0.55)
Laplace(0,2)	RBVSS	3.65(0.45)	0.09(0.35)	14.77(0.68)	0.80(1.03)	18.32(0.85)	0.89(1.26)
	RBV-BI	3.23(0.54)	0.45(0.58)	12.21(1.78)	2.56(1.00)	15.44(1.46)	3.01(1.17)
	RBV	3.15(0.48)	0.57(0.86)	10.25(1.41)	1.23(1.03)	13.40(1.75)	1.80(1.48)
	BVSS-BI	3.55(0.71)	0.03(0.18)	17.59(0.33)	8.28(0.33)	21.14(0.71)	8.31(0.18)
	BVSS	3.50(0.54)	0.10(0.31)	12.43(2.08)	0.27(0.45)	15.93(1.17)	0.37(0.62)
	BV-BI	2.75(0.68)	0.35(0.54)	9.77(1.72)	0.50(0.57)	12.52(2.08)	0.85(0.91)
	BV	2.50(0.76)	0.47(0.78)	7.55(1.81)	0.60(0.72)	10.05(2.16)	1.07(1.05)
Error 4	RBVSS – BI	3.97(0.45)	0.07(0.25)	17.58(0.32)	0.05(0.34)	21.65(0.35)	0.12(0.25)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.85(0.31)	0.08(0.25)	14.83(0.38)	0.77(0.97)	18.68(0.38)	0.85(0.95)
	RBV-BI	3.30(0.25)	0.58(0.98)	11.35(1.27)	1.76(1.38)	15.65(1.26)	2.34(1.84)
	RBV	3.13(0.25)	0.75(1.02)	10.75(1.48)	0.97(0.89)	13.88(1.48)	1.72(1.25)
	BVSS-BI	3.34(0.68)	0.12(0.18)	17.88(0.33)	6.17(0.23)	21.22(0.68)	6.29(0.18)
	BVSS	3.12(0.55)	0.15(0.25)	13.03(1.73)	0.33(0.55)	16.15(1.71)	0.48(0.62)
	BV-BI	2.65(0.47)	0.35(0.73)	9.85(1.75)	0.67(0.84)	12.50(1.83)	1.02(1.03)
	BV	2.55(0.56)	0.47(0.63)	7.75(1.65)	0.75(0.92)	10.30(1.78)	1.22(0.99)
Error 5	RBVSS – BI	3.96(0.31)	0.02(0.21)	17.55(0.35)	0.02(0.42)	21.51(0.31)	0.22(0.41)
Lognormal(0,1)	RBVSS	3.97(0.15)	0.03(0.18)	15.10(0.31)	0.33(0.76)	19.07(0.31)	0.36(0.77)
	RBV-BI	3.24(0.51)	0.47(0.73)	12.45(0.56)	2.30(1.37)	15.69(0.43)	2.77(1.52)
	RBV	3.12(0.25)	0.53(0.57)	10.27(0.40)	0.97(0.96)	13.39(0.43)	1.50(1.25)
	BVSS-BI	3.75(0.54)	0.07(0.25)	17.79(0.33)	8.32(0.37)	21.54(0.70)	8.39(0.25)
	BVSS	3.45(0.25)	0.12(0.56)	14.50(0.94)	0.63(0.96)	17.95(1.14)	0.75(0.96)
	BV-BI	2.58(0.63)	0.30(0.60)	10.75(1.52)	1.13(1.33)	13.33(1.76)	1.43(1.48)
	BV	2.27(0.67)	0.43(0.66)	8.56(1.68)	0.85(1.46)	10.83(2.06)	1.28(1.63)

Table A.4: Identification results for $n = 200, m = 60, p = 3$. $mean(sd)$ of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 50th quantile.

$n = 200 \quad m = 60$		Main		Interaction		Total	
$\theta = 0.5$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.95(0.11)	0.08(0.45)	17.95(0.31)	0.02(0.14)	21.90(0.35)	0.10(0.45)
N(0,1)	RBVSS	3.95(0.23)	0.12(2.13)	15.23(0.58)	0.15(0.58)	19.18(0.45)	0.27(1.58)
	RBV-BI	3.37(1.34)	1.80(1.92)	10.56(3.27)	1.80(2.05)	13.93(2.21)	3.60(3.91)
	RBV	3.25(0.84)	2.75(2.22)	8.54(1.87)	0.86(0.78)	11.79(2.59)	3.61(1.97)
	BVSS-BI	3.97(0.23)	0.14(0.45)	17.75(0.21)	0.03(0.33)	21.72(0.35)	0.17(0.52)
	BVSS	3.84(0.33)	0.22(0.51)	15.15(0.15)	0.05(0.12)	18.99(0.34)	0.27(0.34)
	BV-BI	3.34(0.34)	0.34(0.34)	12.37(1.14)	1.56(0.37)	15.71(1.14)	1.90(0.67)
	BV	3.12(0.16)	0.46(0.56)	10.80(0.84)	0.76(0.37)	13.92(1.23)	1.22(0.56)
Error 2	RBVSS – BI	3.97(0.34)	0.05(0.23)	17.85(0.24)	0.04(0.28)	21.82(0.33)	0.09(0.32)
$t(2)$	RBVSS	3.86(0.42)	0.07(0.31)	15.55(0.56)	0.08(0.76)	19.41(0.15)	0.15(0.34)
	RBV-BI	2.85(0.45)	0.13(0.42)	11.40(1.82)	0.26(0.46)	14.25(0.89)	0.39(0.54)
	RBV	2.57(0.32)	0.23(0.56)	9.56(1.06)	0.14(0.32)	12.13(1.10)	0.37(0.35)
	BVSS-BI	3.25(1.06)	0.08(0.32)	17.77(0.35)	20.14(0.22)	21.02(1.67)	20.22(0.47)
	BVSS	3.20(1.10)	0.11(0.35)	14.56(2.60)	0.28(0.45)	17.76(0.87)	0.38(0.58)
	BV-BI	2.25(0.58)	0.21(0.67)	10.23(2.97)	1.20(0.32)	12.48(2.06)	1.41(0.57)
	BV	2.14(0.71)	0.25(0.45)	9.25(1.12)	0.06(0.45)	11.39(1.74)	0.27(0.45)
Error 3	RBVSS – BI	3.88(0.54)	0.02(0.75)	17.54(0.45)	0.07(0.31)	21.42(0.78)	0.09(0.47)
Laplace(0,2)	RBVSS	3.75(0.32)	0.05(0.18)	15.26(0.31)	0.09(0.35)	19.01(0.89)	0.14(0.34)
	RBV-BI	2.85(0.89)	0.13(0.56)	10.56(0.45)	0.45(0.32)	13.41(0.76)	0.58(0.43)
	RBV	2.55(1.23)	0.15(0.58)	9.40(0.58)	0.05(0.13)	11.95(1.14)	0.20(0.57)
	BVSS-BI	3.15(1.48)	0.08(0.34)	17.76(0.33)	19.05(0.31)	21.91(1.48)	19.13(0.18)
	BVSS	3.10(0.78)	0.11(0.35)	13.20(3.49)	0.60(0.55)	16.30(2.49)	0.71(0.43)
	BV-BI	2.10(0.84)	0.23(0.45)	10.27(1.82)	0.95(0.21)	12.37(2.17)	1.18(0.32)
	BV	2.07(0.05)	0.27(0.51)	8.45(1.23)	0.15(0.23)	10.52(1.82)	0.42(0.34)
Error 4	RBVSS – BI	3.79(0.32)	0.02(0.21)	17.56(0.36)	0.01(0.12)	21.35(0.41)	0.03(0.15)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.75(0.13)	0.04(0.17)	15.23(0.38)	0.07(0.14)	18.98(0.45)	0.11(0.18)
	RBV-BI	3.40(0.89)	0.18(0.31)	11.20(1.92)	0.23(0.34)	14.60(1.95)	0.41(0.54)
	RBV	3.25(0.89)	0.19(0.23)	9.57(1.52)	0.15(0.23)	12.82(1.58)	0.34(0.52)
	BVSS-BI	3.34(0.84)	0.09(0.31)	17.79(0.32)	20.15(0.33)	21.13(0.84)	20.24(0.31)
	BVSS	3.21(0.45)	0.20(0.45)	12.40(1.82)	0.15(0.32)	15.61(1.92)	0.35(0.45)
	BV-BI	2.24(0.84)	0.23(0.46)	10.50(1.30)	0.95(0.15)	12.74(1.95)	1.18(0.56)
	BV	2.17(0.55)	0.31(0.51)	8.55(1.10)	0.18(0.23)	10.72(1.52)	0.49(0.42)
Error 5	RBVSS – BI	3.89(0.15)	0.03(0.45)	17.95(2.68)	0.02(0.31)	21.84(0.32)	0.05(0.18)
Lognormal(0,1)	RBVSS	3.76(0.18)	0.05(2.15)	15.60(1.34)	0.06(1.56)	19.36(1.34)	0.11(1.67)
	RBV-BI	3.15(0.45)	0.20(0.45)	12.51(2.13)	0.28(0.45)	15.66(1.47)	0.48(0.89)
	RBV	3.05(1.30)	0.25(0.23)	9.80(1.96)	0.08(0.23)	12.85(1.64)	0.33(0.25)
	BVSS-BI	3.25(1.52)	0.09(0.15)	17.59(0.31)	18.07(0.12)	21.84(1.52)	18.16(0.15)
	BVSS	3.05(0.98)	0.12(0.25)	14.60(0.89)	0.08(0.26)	17.65(0.83)	0.20(0.15)
	BV-BI	2.40(1.82)	0.25(0.15)	11.15(1.10)	0.25(0.45)	13.55(1.92)	0.50(0.45)
	BV	2.09(1.30)	0.31(0.45)	8.77(1.58)	0.07(0.45)	10.86(1.30)	0.38(0.45)

Table A.5: *Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept model based on 100 replicates at the 50th quantile.*

$n = 200$	$m = 30$	$\theta = 0.5$	MAD	NMAD	TMAD
Error 1 N(0,1)	RBVSS – BI	0.011(0.005)	0.004(0.012)	0.002(0.001)	
	RBVSS	0.026(0.003)	0.008(0.004)	0.005(0.001)	
	RBV-BI	0.066(0.013)	0.050(0.006)	0.053(0.006)	
	RBV	0.088(0.014)	0.067(0.006)	0.059(0.006)	
	BVSS-BI	0.017(0.005)	0.002(0.005)	0.004(0.001)	
	BVSS	0.028(0.003)	0.003(0.007)	0.005(0.001)	
	BV-BI	0.071(0.002)	0.045(0.001)	0.055(0.001)	
	BV	0.095(0.005)	0.069(0.001)	0.058(0.001)	
Error 2 $t(2)$	RBVSS – BI	0.015(0.115)	0.002(0.008)	0.012(0.021)	
	RBVSS	0.035(0.037)	0.004(0.012)	0.052(0.007)	
	RBV-BI	0.062(0.017)	0.034(0.006)	0.062(0.006)	
	RBV	0.084(0.018)	0.053(0.006)	0.078(0.007)	
	BVSS-BI	0.123(0.133)	0.024(0.015)	0.050(0.034)	
	BVSS	0.247(0.109)	0.035(0.002)	0.062(0.020)	
	BV-BI	0.345(0.020)	0.138(0.012)	0.148(0.013)	
	BV	0.456(0.021)	0.145(0.014)	0.155(0.014)	
Error 3 Laplace(0,2)	RBVSS – BI	0.024(0.016)	0.005(0.008)	0.011(0.003)	
	RBVSS	0.048(0.012)	0.007(0.012)	0.013(0.002)	
	RBV-BI	0.116(0.026)	0.024(0.034)	0.056(0.014)	
	RBV	0.123(0.034)	0.046(0.009)	0.085(0.016)	
	BVSS-BI	0.037(0.014)	0.016(0.004)	0.024(0.003)	
	BVSS	0.064(0.042)	0.024(0.006)	0.035(0.008)	
	BV-BI	0.179(0.011)	0.127(0.004)	0.138(0.004)	
	BV	0.184(0.011)	0.134(0.005)	0.145(0.005)	
Error 4 20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)	RBVSS – BI	0.035(0.002)	0.001(0.008)	0.009(0.005)	
	RBVSS	0.053(0.005)	0.005(0.012)	0.012(0.001)	
	RBV-BI	0.121(0.028)	0.082(0.009)	0.075(0.015)	
	RBV	0.146(0.029)	0.084(0.009)	0.083(0.015)	
	BVSS-BI	0.058(0.031)	0.017(0.005)	0.018(0.009)	
	BVSS	0.073(0.005)	0.025(0.007)	0.024(0.015)	
	BV-BI	0.184(0.013)	0.130(0.004)	0.141(0.004)	
	BV	0.192(0.014)	0.144(0.004)	0.148(0.005)	
Error 5 Lognormal(0,1)	RBVSS – BI	0.026(0.009)	0.007(0.004)	0.010(0.002)	
	RBVSS	0.067(0.008)	0.012(0.005)	0.030(0.001)	
	RBV-BI	0.085(0.015)	0.043(0.005)	0.054(0.005)	
	RBV	0.117(0.017)	0.054(0.005)	0.067(0.006)	
	BVSS-BI	0.093(0.009)	0.014(0.002)	0.028(0.002)	
	BVSS	0.123(0.008)	0.023(0.008)	0.035(0.001)	
	BV-BI	0.138(0.009)	0.100(0.004)	0.105(0.005)	
	BV	0.141(0.009)	0.105(0.005)	0.110(0.005)	

Table A.6: *Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept model based on 100 replicates at the 50th quantile.*

$n = 200$	$m = 60$	$\theta = 0.5$	MAD	NMAD	TMAD
Error 1 N(0,1)	RBVSS – BI	0.006(0.002)	0.002(0.001)	0.002(0.001)	
	RBVSS	0.012(0.021)	0.016(0.010)	0.018(0.011)	
	RBV-BI	0.171(0.056)	0.074(0.030)	0.083(0.031)	
	RBV	0.252(0.066)	0.143(0.088)	0.135(0.085)	
	BVSS-BI	0.011(0.002)	0.004(0.001)	0.003(0.001)	
	BVSS	0.018(0.003)	0.006(0.000)	0.004(0.001)	
	BV-BI	0.114(0.020)	0.027(0.002)	0.035(0.004)	
	BV	0.129(0.032)	0.029(0.003)	0.038(0.006)	
Error 2 $t(2)$	RBVSS – BI	0.034(0.021)	0.009(0.006)	0.012(0.007)	
	RBVSS	0.145(0.022)	0.032(0.002)	0.043(0.001)	
	RBV-BI	0.117(0.030)	0.043(0.003)	0.038(0.005)	
	RBV	0.121(0.017)	0.057(0.003)	0.057(0.003)	
	BVSS-BI	0.111(0.058)	0.033(0.013)	0.040(0.015)	
	BVSS	0.124(0.055)	0.046(0.007)	0.053(0.010)	
	BV-BI	0.228(0.057)	0.053(0.046)	0.105(0.044)	
	BV	0.238(0.045)	0.093(0.043)	0.116(0.043)	
Error 3 Laplace(0,2)	RBVSS – BI	0.015(0.004)	0.005(0.002)	0.006(0.002)	
	RBVSS	0.021(0.004)	0.008(0.001)	0.008(0.002)	
	RBV-BI	0.136(0.015)	0.025(0.002)	0.040(0.002)	
	RBV	0.156(0.019)	0.034(0.002)	0.055(0.003)	
	BVSS-BI	0.065(0.030)	0.026(0.008)	0.030(0.009)	
	BVSS	0.085(0.005)	0.036(0.005)	0.046(0.005)	
	BV-BI	0.211(0.025)	0.049(0.004)	0.053(0.005)	
	BV	0.219(0.027)	0.059(0.005)	0.074(0.005)	
Error 4 20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)	RBVSS – BI	0.017(0.003)	0.005(0.001)	0.006(0.005)	
	RBVSS	0.028(0.004)	0.008(0.001)	0.008(0.001)	
	RBV-BI	0.133(0.011)	0.030(0.001)	0.039(0.002)	
	RBV	0.145(0.020)	0.047(0.003)	0.048(0.004)	
	BVSS-BI	0.075(0.022)	0.026(0.006)	0.030(0.008)	
	BVSS	0.089(0.014)	0.035(0.002)	0.070(0.003)	
	BV-BI	0.186(0.016)	0.060(0.004)	0.081(0.004)	
	BV	0.205(0.017)	0.075(0.004)	0.094(0.004)	
Error 5 Lognormal(0,1)	RBVSS – BI	0.018(0.006)	0.005(0.002)	0.006(0.002)	
	RBVSS	0.023(0.004)	0.007(0.001)	0.008(0.001)	
	RBV-BI	0.119(0.030)	0.039(0.013)	0.046(0.014)	
	RBV	0.144(0.023)	0.045(0.012)	0.054(0.012)	
	BVSS-BI	0.041(0.021)	0.023(0.009)	0.025(0.010)	
	BVSS	0.052(0.005)	0.029(0.003)	0.045(0.003)	
	BV-BI	0.165(0.019)	0.055(0.007)	0.065(0.007)	
	BV	0.178(0.027)	0.064(0.007)	0.076(0.007)	

A.2.2 Identification results for $\theta = 0.3$ in simulation

Table A.7: Identification results for $n = 200, m = 30, p = 3$. *mean(sd)* of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 30th quantile.

$n = 200 \quad m = 30$		Main		Interaction		Total	
$\theta = 0.3$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.97(0.25)	0.03(0.04)	17.97(0.11)	0.04(0.21)	21.94(0.25)	0.07(0.21)
N(0,1)	RBVSS	3.86(0.41)	0.04(0.05)	15.25(0.25)	0.08(0.25)	19.11(0.31)	0.12(0.56)
	RBV-BI	3.15(0.55)	0.14(0.32)	11.65(0.89)	0.40(0.82)	14.80(1.30)	0.54(0.89)
	RBV	3.10(0.52)	0.17(0.24)	9.25(2.55)	0.32(0.34)	12.35(3.05)	0.49(0.37)
	BVSS-BI	3.85(0.24)	0.05(0.12)	17.25(0.23)	0.05(0.68)	21.10(0.56)	0.10(0.67)
	BVSS	3.75(0.23)	0.08(0.23)	15.07(0.12)	0.09(0.23)	18.82(0.24)	0.17(0.23)
	BV-BI	3.05(0.15)	0.15(0.41)	10.87(1.30)	0.55(0.21)	13.92(1.30)	0.70(0.41)
	BV	3.01(0.14)	0.16(0.52)	8.79(2.30)	0.12(0.32)	11.80(2.30)	0.28(0.32)
Error 2	RBVSS – BI	3.87(0.25)	0.01(0.45)	17.56(0.23)	0.02(0.28)	21.43(0.43)	0.03(0.26)
$t(2)$	RBVSS	3.75(0.43)	0.04(0.51)	14.80(0.45)	0.20(0.45)	18.55(0.45)	0.24(0.45)
	RBV-BI	3.08(0.46)	0.17(0.25)	10.75(1.79)	0.85(0.89)	13.83(1.79)	1.02(0.89)
	RBV	2.97(0.45)	0.19(0.35)	8.25(1.67)	0.24(0.43)	11.22(0.78)	0.43(0.37)
	BVSS-BI	3.15(0.79)	0.04(0.35)	17.59(0.21)	8.08(0.22)	20.74(1.48)	8.12(0.08)
	BVSS	3.02(1.73)	0.08(0.23)	13.78(1.48)	0.12(0.33)	16.80(2.07)	0.20(0.08)
	BV-BI	2.30(1.79)	0.18(0.44)	9.25(1.32)	1.55(0.45)	11.55(1.78)	1.73(0.42)
	BV	2.10(1.73)	0.23(0.32)	7.55(1.85)	0.54(0.32)	9.65(1.36)	0.77(0.51)
Error 3	RBVSS – BI	3.95(0.35)	0.02(0.47)	17.92(0.43)	0.01(0.25)	21.87(0.18)	0.03(0.57)
Laplace(0,2)	RBVSS	3.85(0.46)	0.05(0.45)	14.40(0.89)	0.05(0.17)	18.25(0.89)	0.10(0.45)
	RBV-BI	3.13(0.55)	0.12(0.34)	10.40(2.41)	0.95(0.25)	13.53(2.45)	1.07(0.34)
	RBV	2.97(0.45)	0.14(0.45)	8.80(3.56)	0.24(0.35)	11.77(3.94)	0.38(0.86)
	BVSS-BI	3.24(0.55)	0.08(0.56)	17.77(0.15)	7.12(0.18)	21.01(0.55)	7.20(0.02)
	BVSS	3.15(0.23)	0.12(0.78)	13.20(1.10)	0.06(0.18)	16.35(1.10)	0.18(0.43)
	BV-BI	2.53(0.41)	0.21(0.85)	10.15(1.30)	0.95(0.27)	12.68(1.30)	1.16(0.56)
	BV	2.35(0.83)	0.28(0.40)	7.70(2.41)	0.17(0.38)	10.05(2.47)	0.45(0.52)
Error 4	RBVSS – BI	3.95(0.17)	0.01(0.16)	17.95(0.25)	0.05(0.15)	21.90(0.23)	0.05(0.25)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.95(0.23)	0.04(0.21)	15.14(0.21)	0.09(0.14)	19.09(2.00)	0.13(0.17)
	RBV-BI	3.55(0.31)	0.18(0.22)	11.25(0.37)	0.96(0.25)	14.80(0.43)	1.14(0.25)
	RBV	3.32(1.06)	0.25(0.31)	9.25(2.62)	0.03(0.18)	12.57(1.27)	0.28(0.18)
	BVSS-BI	3.22(0.68)	0.10(0.25)	17.87(0.21)	8.06(0.22)	21.09(0.68)	8.16(0.35)
	BVSS	3.13(0.35)	0.15(0.34)	13.67(2.52)	0.10(0.31)	16.80(2.62)	0.25(0.31)
	BV-BI	2.45(0.77)	0.24(0.25)	9.85(2.64)	0.35(0.25)	12.30(2.90)	0.59(0.48)
	BV	2.16(0.84)	0.34(0.28)	7.85(2.72)	0.17(0.38)	10.01(3.19)	0.51(0.54)
Error 5	RBVSS – BI	3.93(0.25)	0.10(0.31)	17.80(0.76)	0.10(0.55)	21.73(0.79)	0.20(0.61)
Lognormal(0,1)	RBVSS	3.75(0.35)	0.15(0.25)	14.27(1.31)	0.13(0.43)	18.02(1.46)	0.28(0.55)
	RBV-BI	3.14(1.04)	0.24(0.25)	11.35(2.14)	0.17(0.38)	14.49(2.76)	0.41(0.50)
	RBV	3.05(1.22)	0.32(0.31)	9.33(2.60)	0.21(0.45)	12.38(3.39)	0.53(0.31)
	BVSS-BI	3.15(0.81)	0.15(0.18)	17.97(0.15)	7.11(0.17)	21.15(0.81)	7.26(0.18)
	BVSS	3.10(0.71)	0.21(0.24)	12.37(2.54)	0.23(0.57)	15.47(2.60)	0.44(0.57)
	BV-BI	2.32(0.89)	0.43(0.18)	9.25(2.15)	0.43(0.31)	11.57(2.41)	0.86(0.43)
	BV	2.21(0.95)	0.52(0.31)	7.33(2.26)	0.27(0.25)	9.54(2.70)	0.79(0.46)

Table A.8: Identification results for $n = 200, m = 60, p = 3$. $mean(sd)$ of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 30th quantile.

$n = 200 \quad m = 60$		Main		Interaction		Total	
$\theta = 0.3$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.89(0.24)	0.08(0.25)	17.95(0.25)	0.06(0.24)	21.84(0.34)	0.14(0.25)
N(0,1)	RBVSS	3.72(0.25)	0.12(0.35)	15.00(0.00)	0.12(0.16)	18.72(0.35)	0.24(0.25)
	RBV-BI	3.26(0.84)	0.20(0.45)	10.00(3.67)	0.20(0.45)	13.26(3.77)	0.40(0.55)
	RBV	3.08(1.00)	0.24(0.32)	8.57(2.07)	0.07(0.23)	11.65(3.05)	0.31(0.54)
	BVSS-BI	3.35(0.43)	0.13(0.89)	15.25(0.34)	0.15(2.80)	18.60(0.43)	0.28(1.80)
	BVSS	3.31(0.18)	0.15(0.25)	13.15(0.57)	0.23(0.35)	16.46(0.69)	0.38(0.35)
	BV-BI	2.97(0.87)	0.24(0.35)	10.53(2.42)	1.35(0.18)	13.50(3.02)	1.59(0.18)
	BV	2.27(1.04)	0.33(0.24)	8.25(2.48)	0.89(0.25)	10.52(3.17)	1.22(0.25)
Error 2	RBVSS – BI	3.97(0.35)	0.01(0.48)	17.95(0.31)	0.02(0.18)	21.92(0.27)	0.03(0.25)
$t(2)$	RBVSS	3.85(0.68)	0.03(0.32)	15.11(0.28)	0.08(0.34)	18.96(0.45)	0.11(0.35)
	RBV-BI	3.35(0.84)	0.16(0.67)	12.34(2.41)	0.59(0.42)	15.69(2.78)	0.75(0.36)
	RBV	3.15(1.14)	0.19(0.35)	10.57(2.51)	0.23(0.26)	13.72(2.92)	0.42(0.51)
	BVSS-BI	3.23(1.01)	0.07(0.25)	17.94(0.11)	20.08(0.00)	21.17(1.01)	20.15(0.25)
	BVSS	3.03(1.05)	0.09(0.25)	13.45(1.99)	0.30(0.60)	16.48(1.85)	0.39(0.67)
	BV-BI	2.75(0.83)	0.15(0.34)	9.86(1.76)	0.95(0.23)	12.61(2.27)	1.10(0.24)
	BV	2.25(0.72)	0.24(0.47)	7.26(1.60)	0.45(0.17)	9.51(2.02)	0.69(0.35)
Error 3	RBVSS – BI	3.94(0.21)	0.01(0.22)	17.94(0.21)	0.03(0.15)	21.88(0.19)	0.04(0.16)
Laplace(0,2)	RBVSS	3.80(0.45)	0.02(0.15)	15.32(3.74)	0.05(0.25)	19.12(3.96)	0.07(0.42)
	RBV-BI	3.32(0.84)	0.14(0.35)	11.45(1.00)	0.55(0.35)	14.77(1.64)	0.69(0.53)
	RBV	3.08(0.89)	0.18(0.52)	9.54(1.14)	0.15(0.24)	12.62(1.92)	0.33(0.46)
	BVSS-BI	3.31(0.45)	0.08(0.41)	17.56(0.19)	18.14(0.23)	20.87(0.45)	18.22(0.43)
	BVSS	3.06(0.89)	0.12(0.56)	12.54(2.59)	0.60(0.55)	15.60(2.49)	0.72(0.55)
	BV-BI	2.25(0.71)	0.24(0.32)	10.08(1.10)	0.72(0.33)	12.33(1.48)	0.96(0.57)
	BV	2.14(0.45)	0.35(0.76)	8.07(1.00)	0.34(0.37)	10.21(1.30)	0.69(0.15)
Error 4	RBVSS – BI	3.85(0.45)	0.04(0.56)	17.40(1.34)	0.05(0.23)	21.25(1.30)	0.09(0.32)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.60(0.55)	0.06(0.34)	15.52(1.92)	0.08(0.24)	19.12(3.36)	0.14(0.45)
	RBV-BI	3.21(0.71)	0.15(0.32)	11.85(1.79)	0.98(0.21)	15.06(1.16)	1.13(0.25)
	RBV	3.15(0.71)	0.23(0.34)	9.45(1.82)	0.25(0.34)	12.60(2.07)	0.48(0.43)
	BVSS-BI	3.43(0.55)	0.07(0.31)	17.45(0.21)	20.19(0.21)	20.88(0.55)	20.26(0.31)
	BVSS	3.20(0.84)	0.09(0.22)	12.23(2.88)	0.15(0.36)	15.43(2.88)	0.24(0.35)
	BV-BI	2.85(0.55)	0.16(0.35)	10.16(1.52)	0.54(0.34)	13.01(1.41)	0.70(0.24)
	BV	2.54(0.45)	0.19(0.21)	8.47(1.34)	0.35(0.42)	11.01(1.64)	0.54(0.35)
Error 5	RBVSS – BI	3.80(0.45)	0.02(0.35)	17.20(0.68)	0.03(0.45)	21.00(0.00)	0.05(0.38)
Lognormal(0,1)	RBVSS	3.60(0.55)	0.06(0.45)	14.60(1.95)	0.05(0.45)	18.20(2.49)	0.11(0.55)
	RBV-BI	3.33(1.34)	0.23(0.32)	10.54(4.69)	0.20(0.45)	13.87(5.90)	0.43(0.45)
	RBV	3.15(1.10)	0.34(0.25)	8.66(2.35)	0.23(0.42)	11.81(3.12)	0.57(0.52)
	BVSS-BI	3.35(0.89)	0.07(0.21)	17.89(0.21)	19.07(0.24)	21.24(0.89)	19.14(0.16)
	BVSS	3.23(0.89)	0.09(0.17)	12.76(2.97)	0.16(0.15)	15.99(3.81)	0.25(0.23)
	BV-BI	2.22(0.45)	0.45(0.25)	9.87(2.24)	0.57(0.24)	12.09(2.59)	1.02(0.45)
	BV	2.02(0.25)	0.55(0.15)	7.57(2.19)	0.20(0.45)	9.59(2.19)	0.75(0.45)

Table A.9: Identification results for $n = 200, m = 30, p = 3$. *mean(sd)* of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 30th quantile.

$n = 200 \quad m = 30$		Main		Interaction		Total	
$\theta = 0.3$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.96(0.13)	0.04(0.28)	17.96(0.22)	0.04(0.13)	21.92(0.33)	0.08(0.27)
N(0,1)	RBVSS	3.87(0.35)	0.08(0.45)	15.27(0.16)	0.08(0.17)	19.14(0.34)	0.16(0.23)
	RBV-BI	3.35(0.27)	0.18(0.37)	11.34(0.56)	0.98(0.29)	14.69(0.45)	1.16(0.67)
	RBV	3.28(0.89)	0.24(0.45)	9.54(0.35)	0.45(0.32)	12.82(0.31)	0.69(0.32)
	BVSS-BI	3.75(0.84)	0.06(0.45)	16.75(0.34)	0.08(0.92)	20.50(0.84)	0.14(0.45)
	BVSS	3.53(0.78)	0.08(0.34)	14.89(0.18)	0.10(0.31)	18.42(0.18)	0.18(0.31)
	BV-BI	3.28(0.34)	0.22(0.38)	10.89(0.32)	1.39(0.63)	14.17(0.32)	1.61(0.86)
	BV	3.18(0.42)	0.33(0.43)	8.98(0.37)	0.57(0.73)	12.16(0.37)	0.90(0.89)
Error 2	RBVSS – BI	3.98(0.25)	0.03(0.79)	17.85(0.24)	0.05(0.34)	21.83(0.45)	0.08(0.25)
$t(2)$	RBVSS	3.75(0.34)	0.05(0.81)	15.21(0.23)	0.15(0.25)	18.96(0.37)	0.20(0.31)
	RBV-BI	3.22(0.56)	0.15(0.37)	11.25(0.37)	0.59(0.33)	14.47(0.52)	0.74(0.53)
	RBV	3.12(0.47)	0.23(0.47)	9.29(0.43)	0.23(0.45)	12.41(0.34)	0.46(0.48)
	BVSS-BI	3.43(1.01)	0.09(0.18)	17.85(0.23)	8.11(0.20)	21.28(1.01)	8.20(0.18)
	BVSS	3.37(1.00)	0.17(0.25)	13.03(4.89)	0.27(0.64)	14.40(5.69)	0.33(0.76)
	BV-BI	2.97(1.07)	0.47(0.73)	10.43(3.82)	0.97(1.07)	13.40(4.49)	1.43(1.28)
	BV	2.35(1.05)	0.56(0.73)	8.90(3.66)	0.90(1.03)	11.25(4.32)	1.46(1.33)
Error 3	RBVSS – BI	3.87(0.35)	0.03(0.11)	17.90(0.55)	0.06(0.55)	21.77(0.77)	0.09(0.55)
Laplace(0,2)	RBVSS	3.58(0.35)	0.06(0.35)	14.77(0.68)	0.08(1.03)	18.35(0.85)	0.14(1.26)
	RBV-BI	3.27(0.41)	0.17(0.66)	11.74(1.39)	1.10(1.00)	15.01(1.46)	1.27(1.10)
	RBV	3.13(0.48)	0.32(0.86)	9.52(1.51)	0.97(1.03)	12.65(1.64)	1.29(1.36)
	BVSS-BI	3.33(0.71)	0.06(0.18)	17.96(0.17)	7.06(0.17)	21.29(0.71)	7.12(0.18)
	BVSS	3.28(0.54)	0.10(0.31)	13.43(2.08)	0.27(0.45)	16.71(2.11)	0.37(0.62)
	BV-BI	2.75(0.68)	0.43(0.68)	9.77(1.72)	0.50(0.57)	12.52(2.08)	0.93(0.91)
	BV	2.25(0.76)	0.56(0.78)	7.97(1.81)	0.60(0.72)	10.22(2.16)	1.16(1.05)
Error 4	RBVSS – BI	3.98(0.45)	0.07(0.25)	17.97(0.18)	0.02(0.15)	21.95(0.27)	0.09(0.25)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.89(0.25)	0.09(0.25)	15.83(0.38)	0.09(0.97)	19.72(0.38)	0.18(0.95)
	RBV-BI	3.35(0.18)	0.33(0.98)	11.25(1.27)	1.53(1.38)	14.60(1.26)	1.86(1.84)
	RBV	3.24(0.25)	0.83(1.02)	9.27(1.48)	0.97(0.89)	12.51(1.48)	1.80(1.22)
	BVSS-BI	3.53(0.68)	0.09(0.18)	17.88(0.21)	8.08(0.23)	21.41(0.68)	8.17(0.18)
	BVSS	3.34(0.52)	0.13(0.25)	12.56(1.73)	0.85(0.55)	15.90(1.71)	0.40(0.62)
	BV-BI	2.87(0.47)	0.43(0.73)	10.80(1.75)	0.95(0.84)	13.67(1.83)	1.38(1.03)
	BV	2.75(0.56)	0.47(0.63)	8.97(1.65)	0.70(0.92)	11.72(1.78)	1.17(0.99)
Error 5	RBVSS – BI	3.99(0.26)	0.04(0.13)	17.87(0.34)	0.02(0.16)	21.86(0.34)	0.06(0.15)
Lognormal(0,1)	RBVSS	3.87(0.45)	0.07(0.18)	15.10(0.31)	0.33(0.76)	18.97(0.31)	0.40(0.77)
	RBV-BI	3.38(0.45)	0.27(0.73)	11.27(0.43)	1.30(1.37)	14.65(0.43)	1.57(1.52)
	RBV	3.29(0.31)	0.35(0.57)	9.10(0.40)	0.97(0.96)	12.39(0.40)	1.32(1.25)
	BVSS-BI	3.53(0.70)	0.07(0.25)	17.56(0.12)	7.11(0.12)	21.09(0.70)	7.18(0.25)
	BVSS	3.34(0.25)	0.12(0.35)	13.50(0.94)	0.63(0.96)	16.84(1.14)	0.75(0.96)
	BV-BI	2.67(0.61)	0.20(0.60)	11.36(1.48)	1.13(1.33)	14.03(1.76)	1.33(1.48)
	BV	2.45(0.67)	0.33(0.66)	8.67(1.68)	1.23(1.46)	11.12(2.06)	1.56(1.63)

Table A.10: Identification results for $n = 200, m = 60, p = 3$. *mean(sd)* of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 30th quantile.

$n = 200 \quad m = 60$		Main		Interaction		Total	
$\theta = 0.3$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.87(0.27)	0.07(0.35)	17.77(0.42)	0.07(0.45)	21.64(0.35)	0.14(0.35)
N(0,1)	RBVSS	3.73(0.25)	0.14(1.42)	15.23(0.45)	0.12(1.99)	18.96(0.45)	0.26(1.06)
	RBV-BI	3.34(1.52)	0.26(4.21)	11.78(2.70)	1.68(1.21)	15.12(2.92)	1.94(1.36)
	RBV	3.24(1.14)	0.40(0.58)	9.80(1.10)	0.98(1.65)	13.04(1.95)	1.38(1.85)
	BVSS-BI	3.75(0.35)	0.09(0.25)	17.25(0.35)	0.09(0.34)	21.00(0.25)	0.18(0.52)
	BVSS	3.55(0.25)	0.13(0.35)	14.98(0.45)	0.13(0.45)	18.53(0.76)	0.26(0.48)
	BV-BI	3.12(0.45)	0.24(0.25)	11.60(1.14)	0.96(0.25)	14.72(1.14)	1.20(0.56)
	BV	3.05(0.45)	0.35(0.31)	8.89(0.84)	0.54(0.47)	11.94(1.14)	0.89(0.33)
Error 2	RBVSS – BI	3.98(0.26)	0.03(0.45)	17.99(0.25)	0.04(0.36)	21.97(0.26)	0.07(0.38)
$t(2)$	RBVSS	3.87(0.43)	0.05(0.57)	15.04(0.11)	0.08(0.25)	18.91(0.39)	0.13(0.59)
	RBV-BI	3.35(0.57)	0.14(0.36)	11.22(0.71)	0.13(0.23)	14.57(0.71)	0.27(0.45)
	RBV	3.24(0.22)	0.23(0.29)	9.45(1.10)	0.23(0.45)	12.69(1.10)	0.46(0.45)
	BVSS-BI	3.33(1.67)	0.09(0.45)	17.59(0.18)	20.12(0.15)	20.92(1.67)	20.21(0.34)
	BVSS	3.20(1.10)	0.12(0.00)	13.25(1.60)	0.20(0.45)	16.45(1.66)	0.32(0.45)
	BV-BI	2.45(1.14)	0.34(0.34)	10.45(2.97)	0.96(0.32)	12.90(1.06)	1.30(0.45)
	BV	2.23(0.71)	0.45(0.42)	8.99(2.12)	0.45(0.28)	11.22(2.74)	0.90(0.45)
Error 3	RBVSS – BI	3.95(0.21)	0.02(0.13)	17.95(0.33)	0.02(0.24)	21.90(0.35)	0.04(0.32)
Laplace(0,2)	RBVSS	3.89(0.45)	0.03(0.28)	15.25(0.23)	0.13(0.45)	19.14(0.89)	0.16(0.23)
	RBV-BI	3.40(0.89)	0.15(0.45)	11.12(0.21)	1.45(0.35)	14.52(0.89)	1.60(0.67)
	RBV	3.30(1.00)	0.24(0.57)	9.25(0.55)	0.56(0.67)	12.55(1.14)	0.80(0.55)
	BVSS-BI	3.34(1.48)	0.08(0.35)	17.77(0.21)	19.07(0.13)	21.11(1.48)	19.15(0.35)
	BVSS	3.27(0.78)	0.12(0.34)	13.20(1.49)	0.60(0.55)	16.47(1.49)	0.72(0.56)
	BV-BI	2.98(0.84)	0.35(0.45)	10.26(1.82)	0.95(0.35)	13.24(2.17)	1.30(0.43)
	BV	2.87(1.14)	0.45(0.56)	8.99(1.23)	0.56(0.32)	11.86(1.82)	1.01(0.30)
Error 4	RBVSS – BI	3.95(0.25)	0.03(0.35)	17.98(0.41)	0.02(0.54)	21.53(0.32)	0.05(0.35)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.87(0.38)	0.05(0.48)	15.32(0.45)	0.08(0.34)	19.19(0.45)	0.13(0.57)
	RBV-BI	3.36(0.89)	0.13(0.47)	11.20(1.92)	0.98(0.32)	14.56(1.95)	1.12(0.29)
	RBV	3.28(0.79)	0.24(0.21)	9.26(1.52)	0.27(0.37)	12.54(1.58)	0.51(0.35)
	BVSS-BI	3.43(0.84)	0.08(0.32)	17.85(0.16)	20.15(0.23)	21.28(0.84)	20.23(0.15)
	BVSS	3.32(0.45)	0.12(0.45)	12.40(1.82)	0.13(0.42)	15.72(1.92)	0.25(0.45)
	BV-BI	2.76(0.84)	0.20(0.38)	10.51(1.30)	1.25(0.38)	13.27(1.95)	1.45(0.58)
	BV	2.15(0.55)	0.34(0.45)	8.25(1.10)	0.34(0.43)	10.40(1.52)	0.68(0.14)
Error 5	RBVSS – BI	3.97(0.78)	0.06(0.31)	17.55(1.68)	0.08(0.43)	21.52(0.34)	0.14(0.12)
Lognormal(0,1)	RBVSS	3.85(0.35)	0.09(1.03)	15.60(1.34)	0.12(0.42)	19.45(1.34)	0.21(1.45)
	RBV-BI	3.32(0.45)	0.20(0.45)	12.60(3.13)	0.20(0.45)	15.92(1.58)	0.40(0.89)
	RBV	3.20(1.30)	0.26(0.32)	10.80(1.96)	0.27(0.31)	14.00(1.64)	0.53(0.47)
	BVSS-BI	3.25(1.52)	0.13(0.32)	17.55(0.23)	18.12(0.23)	20.80(1.52)	18.25(0.25)
	BVSS	3.20(0.58)	0.18(0.28)	13.60(0.89)	0.15(0.35)	16.80(0.89)	0.33(0.15)
	BV-BI	2.40(1.82)	0.34(0.41)	10.27(1.10)	1.25(0.45)	12.67(1.92)	1.59(0.45)
	BV	2.23(1.30)	0.42(0.37)	8.24(1.58)	0.82(0.45)	10.47(1.30)	1.24(0.45)

Table A.11: *Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept - slope model based on 100 replicates at the 30th quantile.*

$n = 200$	$m = 30$	$\theta = 0.3$	MAD	NMAD	TMAD
Error 1 N(0,1)	RBVSS – BI	0.023(0.015)	0.001(0.035)	0.007(0.003)	
	RBVSS	0.056(0.026)	0.004(0.001)	0.009(0.008)	
	RBV-BI	0.090(0.038)	0.029(0.010)	0.038(0.013)	
	RBV	0.162(0.048)	0.048(0.005)	0.057(0.012)	
	BVSS-BI	0.039(0.009)	0.005(0.004)	0.012(0.003)	
	BVSS	0.064(0.013)	0.007(0.005)	0.021(0.007)	
	BV-BI	0.135(0.008)	0.059(0.001)	0.067(0.003)	
	BV	0.158(0.009)	0.068(0.001)	0.079(0.008)	
Error 2 $t(2)$	RBVSS – BI	0.032(0.017)	0.004(0.015)	0.011(0.008)	
	RBVSS	0.074(0.014)	0.008(0.023)	0.015(0.003)	
	RBV-BI	0.135(0.027)	0.027(0.009)	0.042(0.015)	
	RBV	0.146(0.032)	0.042(0.007)	0.059(0.015)	
	BVSS-BI	0.065(0.024)	0.033(0.005)	0.057(0.003)	
	BVSS	0.085(0.018)	0.045(0.006)	0.068(0.007)	
	BV-BI	0.176(0.025)	0.121(0.009)	0.143(0.058)	
	BV	0.235(0.019)	0.157(0.014)	0.176(0.015)	
Error 3 Laplace(0,2)	RBVSS – BI	0.046(0.054)	0.007(0.026)	0.018(0.007)	
	RBVSS	0.117(0.025)	0.015(0.005)	0.025(0.008)	
	RBV-BI	0.156(0.045)	0.053(0.014)	0.075(0.035)	
	RBV	0.192(0.035)	0.064(0.007)	0.084(0.012)	
	BVSS-BI	0.176(0.045)	0.032(0.015)	0.037(0.015)	
	BVSS	0.185(0.098)	0.045(0.003)	0.057(0.019)	
	BV-BI	0.205(0.015)	0.108(0.008)	0.112(0.007)	
	BV	0.247(0.045)	0.149(0.004)	0.135(0.008)	
Error 4 20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)	RBVSS – BI	0.055(0.027)	0.006(0.009)	0.022(0.005)	
	RBVSS	0.108(0.075)	0.011(0.003)	0.041(0.015)	
	RBV-BI	0.139(0.027)	0.053(0.006)	0.057(0.007)	
	RBV	0.163(0.038)	0.070(0.012)	0.082(0.011)	
	BVSS-BI	0.156(0.036)	0.012(0.023)	0.062(0.019)	
	BVSS	0.172(0.138)	0.024(0.011)	0.084(0.023)	
	BV-BI	0.196(0.023)	0.103(0.024)	0.133(0.015)	
	BV	0.221(0.045)	0.123(0.027)	0.165(0.033)	
Error 5 Lognormal(0,1)	RBVSS – BI	0.044(0.076)	0.005(0.045)	0.014(0.016)	
	RBVSS	0.096(0.055)	0.007(0.032)	0.027(0.012)	
	RBV-BI	0.144(0.023)	0.032(0.025)	0.039(0.022)	
	RBV	0.176(0.057)	0.086(0.008)	0.095(0.011)	
	BVSS-BI	0.158(0.063)	0.009(0.012)	0.038(0.056)	
	BVSS	0.163(0.128)	0.018(0.035)	0.045(0.056)	
	BV-BI	0.154(0.015)	0.104(0.054)	0.102(0.016)	
	BV	0.182(0.023)	0.192(0.048)	0.136(0.027)	

Table A.12: *Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept - slope model based on 100 replicates at the 30th quantile.*

$n = 200$	$m = 60$	$\theta = 0.3$	MAD	NMAD	TMAD
Error 1 N(0,1)	RBVSS – BI	0.021(0.018)	0.003(0.011)	0.012(0.005)	
	RBVSS	0.039(0.035)	0.004(0.024)	0.018(0.008)	
	RBV-BI	0.076(0.042)	0.029(0.010)	0.043(0.015)	
	RBV	0.163(0.028)	0.053(0.009)	0.069(0.025)	
	BVSS-BI	0.043(0.013)	0.008(0.023)	0.024(0.008)	
	BVSS	0.054(0.025)	0.015(0.026)	0.037(0.023)	
	BV-BI	0.154(0.011)	0.034(0.017)	0.057(0.018)	
	BV	0.164(0.012)	0.054(0.022)	0.068(0.011)	
Error 2 $t(2)$	RBVSS – BI	0.026(0.024)	0.004(0.015)	0.011(0.008)	
	RBVSS	0.055(0.021)	0.008(0.037)	0.019(0.012)	
	RBV-BI	0.143(0.027)	0.013(0.019)	0.032(0.027)	
	RBV	0.152(0.043)	0.022(0.011)	0.045(0.015)	
	BVSS-BI	0.075(0.024)	0.028(0.005)	0.043(0.011)	
	BVSS	0.088(0.026)	0.035(0.008)	0.052(0.009)	
	BV-BI	0.182(0.034)	0.117(0.011)	0.089(0.082)	
	BV	0.207(0.027)	0.133(0.043)	0.143(0.068)	
Error 3 Laplace(0,2)	RBVSS – BI	0.033(0.084)	0.005(0.016)	0.015(0.011)	
	RBVSS	0.098(0.053)	0.011(0.005)	0.021(0.017)	
	RBV-BI	0.128(0.057)	0.047(0.029)	0.043(0.035)	
	RBV	0.153(0.045)	0.070(0.032)	0.054(0.012)	
	BVSS-BI	0.168(0.053)	0.041(0.038)	0.046(0.021)	
	BVSS	0.190(0.081)	0.053(0.009)	0.052(0.029)	
	BV-BI	0.201(0.037)	0.098(0.010)	0.104(0.007)	
	BV	0.221(0.045)	0.112(0.013)	0.141(0.018)	
Error 4 20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)	RBVSS – BI	0.029(0.023)	0.006(0.009)	0.027(0.009)	
	RBVSS	0.095(0.075)	0.013(0.009)	0.047(0.027)	
	RBV-BI	0.119(0.033)	0.047(0.011)	0.062(0.011)	
	RBV	0.145(0.039)	0.061(0.021)	0.075(0.031)	
	BVSS-BI	0.173(0.029)	0.023(0.039)	0.069(0.025)	
	BVSS	0.184(0.138)	0.035(0.034)	0.081(0.032)	
	BV-BI	0.201(0.035)	0.112(0.054)	0.109(0.021)	
	BV	0.232(0.062)	0.127(0.015)	0.143(0.047)	
Error 5 Lognormal(0,1)	RBVSS – BI	0.039(0.064)	0.003(0.065)	0.026(0.021)	
	RBVSS	0.087(0.072)	0.009(0.056)	0.032(0.028)	
	RBV-BI	0.138(0.039)	0.041(0.039)	0.042(0.031)	
	RBV	0.165(0.084)	0.072(0.019)	0.085(0.026)	
	BVSS-BI	0.143(0.032)	0.011(0.027)	0.031(0.036)	
	BVSS	0.165(0.128)	0.021(0.073)	0.051(0.086)	
	BV-BI	0.178(0.025)	0.098(0.064)	0.088(0.025)	
	BV	0.192(0.036)	0.175(0.123)	0.153(0.067)	

Table A.13: *Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept model based on 100 replicates at the 30th quantile.*

$n = 200$	$m = 30$	$\theta = 0.3$	MAD	NMAD	TMAD
Error 1 N(0,1)	RBVSS – BI	0.019(0.024)	0.002(0.011)	0.004(0.015)	
	RBVSS	0.023(0.013)	0.006(0.018)	0.011(0.021)	
	RBV-BI	0.055(0.026)	0.033(0.011)	0.043(0.011)	
	RBV	0.071(0.045)	0.043(0.012)	0.061(0.021)	
	BVSS-BI	0.023(0.012)	0.004(0.014)	0.008(0.011)	
	BVSS	0.028(0.021)	0.007(0.017)	0.012(0.014)	
	BV-BI	0.063(0.018)	0.023(0.031)	0.035(0.021)	
	BV	0.081(0.035)	0.049(0.025)	0.052(0.015)	
Error 2 $t(2)$	RBVSS – BI	0.021(0.121)	0.004(0.011)	0.009(0.021)	
	RBVSS	0.041(0.027)	0.006(0.019)	0.023(0.007)	
	RBV-BI	0.071(0.021)	0.028(0.009)	0.052(0.011)	
	RBV	0.080(0.023)	0.047(0.006)	0.081(0.011)	
	BVSS-BI	0.129(0.138)	0.028(0.015)	0.048(0.034)	
	BVSS	0.236(0.117)	0.042(0.008)	0.068(0.026)	
	BV-BI	0.333(0.043)	0.114(0.021)	0.124(0.035)	
	BV	0.411(0.043)	0.151(0.038)	0.149(0.014)	
Error 3 Laplace(0,2)	RBVSS – BI	0.023(0.016)	0.003(0.008)	0.016(0.012)	
	RBVSS	0.037(0.025)	0.008(0.021)	0.022(0.006)	
	RBV-BI	0.105(0.026)	0.019(0.034)	0.056(0.014)	
	RBV	0.129(0.034)	0.035(0.019)	0.079(0.016)	
	BVSS-BI	0.041(0.024)	0.018(0.004)	0.031(0.007)	
	BVSS	0.058(0.033)	0.029(0.006)	0.040(0.008)	
	BV-BI	0.164(0.015)	0.128(0.004)	0.119(0.004)	
	BV	0.178(0.019)	0.141(0.008)	0.135(0.005)	
Error 4 20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)	RBVSS – BI	0.028(0.007)	0.003(0.009)	0.011(0.005)	
	RBVSS	0.048(0.008)	0.007(0.024)	0.021(0.010)	
	RBV-BI	0.117(0.033)	0.022(0.011)	0.055(0.034)	
	RBV	0.139(0.029)	0.034(0.012)	0.073(0.026)	
	BVSS-BI	0.047(0.043)	0.022(0.019)	0.021(0.009)	
	BVSS	0.091(0.011)	0.033(0.012)	0.029(0.015)	
	BV-BI	0.164(0.023)	0.121(0.009)	0.139(0.012)	
	BV	0.188(0.034)	0.152(0.009)	0.158(0.005)	
Error 5 Lognormal(0,1)	RBVSS – BI	0.033(0.011)	0.007(0.032)	0.018(0.002)	
	RBVSS	0.058(0.008)	0.012(0.000)	0.030(0.001)	
	RBV-BI	0.076(0.019)	0.033(0.005)	0.061(0.005)	
	RBV	0.123(0.018)	0.058(0.015)	0.078(0.006)	
	BVSS-BI	0.075(0.009)	0.022(0.004)	0.025(0.002)	
	BVSS	0.119(0.008)	0.037(0.008)	0.035(0.001)	
	BV-BI	0.141(0.009)	0.098(0.004)	0.112(0.005)	
	BV	0.157(0.011)	0.114(0.005)	0.125(0.005)	

Table A.14: *Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept model based on 100 replicates at the 30th quantile.*

$n = 200$	$m = 60$	$\theta = 0.3$	MAD	NMAD	TMAD
Error 1 N(0,1)	RBVSS – BI	0.017(0.002)	0.002(0.001)	0.009(0.001)	
	RBVSS	0.023(0.021)	0.011(0.014)	0.021(0.016)	
	RBV-BI	0.181(0.039)	0.064(0.036)	0.079(0.031)	
	RBV	0.242(0.066)	0.138(0.098)	0.146(0.085)	
	BVSS-BI	0.021(0.012)	0.004(0.005)	0.011(0.011)	
	BVSS	0.028(0.011)	0.006(0.006)	0.026(0.001)	
	BV-BI	0.198(0.026)	0.087(0.002)	0.087(0.004)	
	BV	0.253(0.032)	0.156(0.003)	0.154(0.006)	
Error 2 $t(2)$	RBVSS – BI	0.025(0.021)	0.003(0.006)	0.012(0.007)	
	RBVSS	0.056(0.022)	0.012(0.002)	0.031(0.001)	
	RBV-BI	0.109(0.030)	0.033(0.003)	0.040(0.005)	
	RBV	0.124(0.017)	0.059(0.006)	0.060(0.003)	
	BVSS-BI	0.121(0.058)	0.031(0.013)	0.033(0.015)	
	BVSS	0.135(0.065)	0.046(0.012)	0.059(0.018)	
	BV-BI	0.207(0.029)	0.058(0.033)	0.099(0.044)	
	BV	0.241(0.035)	0.073(0.043)	0.116(0.043)	
Error 3 Laplace(0,2)	RBVSS – BI	0.020(0.014)	0.004(0.010)	0.010(0.002)	
	RBVSS	0.033(0.004)	0.008(0.001)	0.015(0.002)	
	RBV-BI	0.126(0.013)	0.015(0.002)	0.039(0.012)	
	RBV	0.145(0.029)	0.028(0.005)	0.052(0.003)	
	BVSS-BI	0.054(0.035)	0.021(0.008)	0.025(0.009)	
	BVSS	0.075(0.012)	0.033(0.011)	0.036(0.008)	
	BV-BI	0.202(0.035)	0.042(0.015)	0.049(0.005)	
	BV	0.223(0.027)	0.062(0.013)	0.065(0.012)	
Error 4 20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)	RBVSS – BI	0.018(0.003)	0.005(0.001)	0.011(0.000)	
	RBVSS	0.025(0.011)	0.008(0.001)	0.015(0.006)	
	RBV-BI	0.129(0.016)	0.027(0.011)	0.035(0.002)	
	RBV	0.141(0.020)	0.038(0.010)	0.051(0.004)	
	BVSS-BI	0.058(0.022)	0.032(0.009)	0.036(0.012)	
	BVSS	0.078(0.018)	0.056(0.002)	0.059(0.003)	
	BV-BI	0.172(0.016)	0.069(0.004)	0.076(0.004)	
	BV	0.198(0.017)	0.082(0.004)	0.089(0.008)	
Error 5 Lognormal(0,1)	RBVSS – BI	0.022(0.009)	0.005(0.002)	0.011(0.002)	
	RBVSS	0.034(0.004)	0.007(0.001)	0.019(0.001)	
	RBV-BI	0.113(0.030)	0.029(0.013)	0.036(0.014)	
	RBV	0.138(0.023)	0.045(0.012)	0.054(0.012)	
	BVSS-BI	0.048(0.021)	0.027(0.011)	0.029(0.010)	
	BVSS	0.056(0.005)	0.037(0.003)	0.049(0.011)	
	BV-BI	0.155(0.019)	0.059(0.007)	0.069(0.007)	
	BV	0.181(0.017)	0.069(0.011)	0.079(0.011)	

A.2.3 Identification results for $\theta = 0.7$ in simulation

Table A.15: *Identification results for $n = 200, m = 30, p = 3$. $mean(sd)$ of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 70th quantile.*

$n = 200 \quad m = 30$		Main		Interaction		Total	
$\theta = 0.7$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.92(0.31)	0.01(0.12)	17.24(0.34)	0.03(0.25)	21.16(0.25)	0.04(0.21)
N(0,1)	RBVSS	3.79(0.23)	0.06(0.09)	15.48(0.38)	0.12(0.35)	19.27(0.31)	0.18(0.56)
	RBV-BI	3.21(0.55)	0.18(0.22)	12.12(0.84)	0.39(0.72)	15.33(0.56)	0.57(0.79)
	RBV	3.08(0.34)	0.20(0.35)	10.17(1.23)	0.43(0.74)	13.25(1.79)	0.63(0.17)
	BVSS-BI	3.86(0.37)	0.03(0.21)	16.97(0.23)	0.05(0.68)	20.83(0.56)	0.08(0.67)
	BVSS	3.69(0.27)	0.08(0.23)	14.97(0.17)	0.14(0.23)	18.66(0.24)	0.22(0.63)
	BV-BI	3.12(0.39)	0.23(0.54)	11.28(1.25)	0.55(0.21)	14.40(1.30)	0.78(0.41)
	BV	3.03(0.14)	0.31(0.52)	9.26(1.30)	0.46(0.32)	12.29(1.82)	0.77(0.32)
Error 2	RBVSS – BI	3.85(0.38)	0.04(0.52)	17.97(0.23)	0.01(0.38)	21.82(0.67)	0.05(0.36)
$t(2)$	RBVSS	3.70(0.43)	0.08(0.62)	14.85(0.55)	0.18(0.25)	18.55(0.29)	0.26(0.45)
	RBV-BI	3.10(0.39)	0.15(0.25)	10.33(1.79)	0.78(0.89)	13.43(1.79)	0.93(0.49)
	RBV	2.87(0.48)	0.20(0.45)	8.29(0.67)	0.58(0.43)	11.16(0.78)	0.78(0.37)
	BVSS-BI	3.13(0.59)	0.08(0.35)	17.58(0.33)	8.06(0.17)	21.13(1.48)	8.08(0.08)
	BVSS	3.00(1.54)	0.18(0.28)	13.14(1.59)	0.25(0.43)	16.14(2.07)	0.43(0.08)
	BV-BI	2.45(1.89)	0.24(0.65)	10.12(1.45)	0.98(0.52)	12.57(1.78)	1.22(0.42)
	BV	2.17(1.49)	0.38(0.46)	7.23(1.19)	0.68(0.54)	9.40(1.36)	1.06(0.51)
Error 3	RBVSS – BI	3.88(0.54)	0.02(0.47)	17.75(0.46)	0.04(0.49)	21.63(0.38)	0.06(0.57)
Laplace(0,2)	RBVSS	3.75(0.46)	0.07(0.55)	14.78(0.84)	0.11(0.23)	18.53(0.59)	0.18(0.45)
	RBV-BI	3.08(0.38)	0.17(0.48)	10.21(1.41)	0.95(0.25)	13.29(0.45)	1.12(0.34)
	RBV	2.77(0.69)	0.23(0.62)	8.85(1.28)	0.68(0.52)	11.62(1.94)	0.91(0.86)
	BVSS-BI	3.18(0.55)	0.16(0.56)	17.82(0.23)	8.05(0.20)	21.00(0.55)	8.21(0.22)
	BVSS	3.10(0.23)	0.21(0.78)	13.12(1.15)	0.11(0.38)	16.22(1.15)	0.32(0.43)
	BV-BI	2.49(0.41)	0.29(0.85)	9.97(1.30)	0.76(0.27)	12.46(1.33)	1.05(0.56)
	BV	2.28(0.36)	0.41(0.56)	7.28(1.38)	0.54(0.88)	9.56(1.47)	0.95(0.52)
Error 4	RBVSS – BI	3.92(0.43)	0.01(0.39)	17.95(0.25)	0.05(0.15)	21.97(0.23)	0.06(0.35)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.85(0.23)	0.04(0.21)	15.36(0.38)	0.12(0.43)	19.21(2.00)	0.16(0.37)
	RBV-BI	3.31(0.36)	0.18(0.22)	12.12(0.37)	0.90(0.35)	15.43(0.43)	1.08(0.25)
	RBV	3.22(1.06)	0.25(0.31)	9.73(1.62)	0.56(0.14)	12.95(1.27)	0.81(0.38)
	BVSS-BI	3.35(0.68)	0.16(0.25)	17.89(0.11)	8.11(0.14)	21.14(0.68)	8.27(0.35)
	BVSS	3.13(0.55)	0.25(0.34)	13.26(1.52)	0.25(0.31)	16.39(0.62)	0.50(0.31)
	BV-BI	2.38(0.77)	0.34(0.25)	10.12(1.64)	0.52(0.55)	12.50(0.90)	0.86(0.48)
	BV	2.15(0.87)	0.49(0.36)	7.54(1.72)	0.42(0.38)	9.69(1.19)	0.91(0.54)
Error 5	RBVSS – BI	3.93(0.25)	0.07(0.41)	17.80(0.56)	0.03(0.35)	21.73(0.79)	0.10(0.61)
Lognormal(0,1)	RBVSS	3.70(0.35)	0.15(0.25)	15.12(0.31)	0.13(0.43)	18.82(1.46)	0.28(0.55)
	RBV-BI	3.26(1.04)	0.23(0.25)	12.12(1.14)	0.15(0.38)	15.38(0.76)	0.38(0.50)
	RBV	3.14(1.22)	0.38(0.35)	10.11(1.60)	0.21(0.45)	13.25(1.39)	0.59(0.25)
	BVSS-BI	3.28(0.81)	0.19(0.28)	17.56(0.14)	8.05(0.11)	20.84(0.81)	8.24(0.18)
	BVSS	3.12(0.71)	0.28(0.33)	12.17(1.92)	0.29(0.37)	15.29(2.60)	0.57(0.27)
	BV-BI	2.75(0.49)	0.35(0.18)	9.99(2.03)	0.58(0.31)	12.74(0.41)	0.93(0.43)
	BV	2.33(0.55)	0.52(0.42)	7.19(0.89)	0.42(0.41)	9.52(2.70)	0.94(0.46)

Table A.16: Identification results for $n = 200, m = 60, p = 3$. $mean(sd)$ of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 70th quantile.

$n = 200 \quad m = 60$		Main		Interaction		Total	
$\theta = 0.7$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.93(0.54)	0.04(0.55)	17.29(0.35)	0.03(0.44)	21.22(0.34)	0.07(0.25)
N(0,1)	RBVSS	3.82(0.47)	0.09(0.35)	15.18(0.37)	0.09(0.21)	19.00(0.35)	0.18(0.25)
	RBV-BI	3.31(0.84)	0.18(0.85)	11.18(1.72)	0.28(0.45)	14.49(3.77)	0.46(0.55)
	RBV	3.19(1.00)	0.24(0.32)	9.21(1.07)	0.19(0.23)	12.40(1.05)	0.43(0.54)
	BVSS-BI	3.71(0.53)	0.08(0.86)	16.13(0.34)	0.08(1.80)	19.84(0.43)	0.16(1.80)
	BVSS	3.61(0.18)	0.10(0.25)	14.12(0.57)	0.12(0.35)	17.73(0.69)	0.22(0.35)
	BV-BI	3.15(0.87)	0.24(0.35)	10.53(2.42)	0.53(0.18)	13.68(2.02)	0.77(0.18)
	BV	3.04(1.04)	0.33(0.24)	8.35(0.48)	0.52(0.25)	11.39(1.17)	0.85(0.25)
Error 2	RBVSS – BI	3.92(0.35)	0.01(0.48)	17.95(0.36)	0.01(0.38)	21.87(0.37)	0.02(0.35)
$t(2)$	RBVSS	3.79(0.68)	0.07(0.32)	15.09(0.21)	0.07(0.34)	18.88(0.45)	0.14(0.35)
	RBV-BI	3.29(0.84)	0.21(0.67)	10.89(1.41)	0.72(0.42)	14.18(0.78)	0.93(0.36)
	RBV	3.11(1.14)	0.31(0.35)	8.87(0.51)	0.51(0.26)	11.98(0.92)	0.82(0.51)
	BVSS-BI	3.21(1.01)	0.12(0.25)	17.45(0.13)	20.11(0.14)	20.66(1.01)	20.23(0.25)
	BVSS	3.08(1.05)	0.19(0.31)	13.12(1.19)	0.21(0.60)	16.20(1.85)	0.40(0.67)
	BV-BI	2.82(0.83)	0.27(0.34)	9.72(0.76)	1.27(0.23)	12.54(2.27)	1.54(0.24)
	BV	2.62(0.72)	0.39(0.47)	7.12(1.60)	0.98(0.17)	9.74(1.02)	1.37(0.55)
Error 3	RBVSS – BI	3.89(0.21)	0.02(0.32)	17.86(0.21)	0.03(0.12)	21.75(0.39)	0.05(0.28)
Laplace(0,2)	RBVSS	3.72(0.45)	0.08(0.28)	15.12(0.74)	0.09(0.25)	18.84(1.96)	0.17(0.42)
	RBV-BI	3.21(0.84)	0.19(0.35)	11.21(1.08)	0.89(0.35)	14.42(1.64)	1.08(0.53)
	RBV	3.15(0.89)	0.27(0.52)	9.12(1.14)	0.76(0.32)	12.27(1.92)	1.03(0.46)
	BVSS-BI	3.28(0.45)	0.11(0.41)	17.26(0.16)	20.12(0.13)	20.54(0.45)	20.23(0.43)
	BVSS	3.19(0.89)	0.12(0.56)	12.23(0.59)	0.28(0.55)	15.42(0.49)	0.40(0.55)
	BV-BI	2.31(0.81)	0.28(0.45)	10.27(1.10)	0.85(0.43)	12.58(1.48)	1.13(0.57)
	BV	2.21(0.45)	0.38(0.31)	8.28(1.00)	0.41(0.37)	10.49(1.30)	0.79(0.15)
Error 4	RBVSS – BI	3.91(0.57)	0.02(0.56)	17.92(1.34)	0.01(0.31)	21.83(1.30)	0.03(0.32)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.73(0.32)	0.06(0.34)	15.17(1.92)	0.10(0.24)	18.90(0.36)	0.16(0.45)
	RBV-BI	3.19(0.52)	0.19(0.32)	11.17(1.79)	0.88(0.21)	14.36(1.16)	1.07(0.25)
	RBV	3.09(0.71)	0.29(0.31)	10.12(1.82)	0.68(0.34)	13.21(1.07)	0.97(0.43)
	BVSS-BI	3.34(0.47)	0.12(0.31)	17.25(0.15)	20.06(0.12)	20.59(0.59)	20.18(0.41)
	BVSS	3.18(0.62)	0.19(0.22)	13.10(1.92)	0.11(0.28)	16.28(1.28)	0.30(0.35)
	BV-BI	2.91(0.55)	0.26(0.35)	10.17(1.52)	0.82(0.34)	13.08(1.41)	1.08(0.24)
	BV	2.61(0.37)	0.33(0.31)	8.59(1.34)	0.47(0.42)	11.20(1.64)	0.80(0.35)
Error 5	RBVSS – BI	3.86(0.45)	0.02(0.35)	17.43(0.68)	0.03(0.45)	21.29(1.08)	0.05(0.38)
Lognormal(0,1)	RBVSS	3.71(0.55)	0.07(0.45)	15.11(1.75)	0.09(0.45)	18.82(0.49)	0.16(0.55)
	RBV-BI	3.41(1.34)	0.21(0.28)	10.29(0.69)	0.30(0.45)	13.70(2.70)	0.51(0.45)
	RBV	3.26(1.10)	0.34(0.25)	8.35(1.35)	0.37(0.42)	11.61(1.12)	0.71(0.32)
	BVSS-BI	3.48(0.89)	0.09(0.21)	17.56(0.15)	20.14(0.23)	21.04(0.74)	20.23(0.38)
	BVSS	3.27(0.39)	0.12(0.22)	12.98(1.97)	0.16(0.15)	16.25(1.81)	0.28(0.23)
	BV-BI	2.45(0.32)	0.35(0.54)	10.07(1.24)	0.62(0.24)	12.45(0.59)	0.87(0.45)
	BV	2.32(0.25)	0.55(0.15)	8.01(1.12)	0.42(0.45)	10.33(2.19)	0.97(0.45)

Table A.17: Identification results for $n = 200, m = 30, p = 3$. *mean(sd)* of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 70th quantile.

$n = 200 \quad m = 30$		Main		Interaction		Total	
$\theta = 0.7$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.95(0.21)	0.02(0.17)	17.96(0.22)	0.03(0.22)	21.91(0.13)	0.05(0.12)
N(0,1)	RBVSS	3.89(0.35)	0.03(0.55)	16.12(0.16)	0.10(0.17)	20.01(0.34)	0.13(0.23)
	RBV-BI	3.43(0.37)	0.15(0.57)	12.05(0.56)	0.85(0.29)	15.48(0.65)	1.00(0.67)
	RBV	3.32(0.89)	0.20(0.45)	10.12(0.45)	0.58(0.62)	13.44(0.31)	0.78(0.32)
	BVSS-BI	3.81(0.84)	0.02(0.85)	17.13(0.28)	0.03(0.56)	20.94(0.84)	0.05(0.65)
	BVSS	3.61(0.59)	0.11(0.34)	14.33(0.28)	0.10(0.31)	17.94(0.34)	0.21(0.31)
	BV-BI	3.37(0.47)	0.21(0.38)	11.12(0.45)	1.57(0.38)	14.49(0.51)	1.78(0.57)
	BV	3.20(0.38)	0.31(0.43)	9.05(0.48)	0.88(0.67)	12.25(0.56)	1.19(0.79)
Error 2	RBVSS – BI	3.89(0.38)	0.02(0.79)	17.89(0.24)	0.03(0.34)	21.78(0.55)	0.05(0.45)
$t(2)$	RBVSS	3.74(0.34)	0.09(0.71)	15.32(0.37)	0.15(0.25)	19.06(0.37)	0.24(0.31)
	RBV-BI	3.19(0.56)	0.18(0.37)	11.15(0.49)	0.98(0.35)	14.34(0.52)	1.16(0.53)
	RBV	3.09(0.47)	0.25(0.47)	8.98(0.43)	0.45(0.75)	12.07(0.34)	0.70(0.38)
	BVSS-BI	3.35(1.14)	0.08(0.22)	17.76(0.16)	8.04(0.15)	21.11(1.01)	8.12(0.18)
	BVSS	3.29(0.78)	0.12(0.45)	12.97(0.89)	0.33(0.64)	16.26(1.69)	0.45(0.76)
	BV-BI	2.75(1.07)	0.35(0.53)	9.98(0.82)	1.55(1.07)	12.73(1.49)	1.90(1.28)
	BV	2.48(1.05)	0.48(0.73)	8.90(2.12)	0.97(1.03)	11.38(1.32)	1.45(1.33)
Error 3	RBVSS – BI	3.91(0.35)	0.02(0.33)	17.96(0.55)	0.02(0.55)	21.87(0.77)	0.04(0.55)
Laplace(0,2)	RBVSS	3.81(0.35)	0.08(0.35)	14.89(0.68)	0.08(1.03)	18.70(0.55)	0.16(1.46)
	RBV-BI	3.39(0.47)	0.21(0.64)	11.85(1.39)	1.15(1.00)	15.24(1.46)	1.36(1.10)
	RBV	3.30(0.48)	0.37(0.86)	10.12(0.51)	0.82(1.03)	13.42(1.64)	1.19(1.36)
	BVSS-BI	3.42(0.71)	0.06(0.18)	17.56(0.23)	8.11(0.15)	20.98(0.71)	8.17(0.18)
	BVSS	3.34(0.54)	0.10(0.31)	12.87(1.08)	0.22(0.45)	16.21(0.11)	0.32(0.62)
	BV-BI	2.95(0.68)	0.41(0.38)	10.12(1.72)	1.23(0.57)	13.07(1.08)	1.64(0.91)
	BV	2.82(0.76)	0.51(0.78)	8.05(1.81)	0.73(0.32)	10.87(1.16)	1.24(1.05)
Error 4	RBVSS – BI	3.92(0.45)	0.03(0.25)	17.77(0.37)	0.02(0.32)	21.69(0.27)	0.05(0.25)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.82(0.25)	0.09(0.25)	15.14(0.38)	0.13(0.97)	18.96(0.38)	0.22(0.95)
	RBV-BI	3.47(0.18)	0.37(0.98)	11.17(1.27)	1.42(1.22)	14.64(1.26)	1.79(0.84)
	RBV	3.39(0.25)	0.52(1.02)	9.54(1.48)	0.97(0.39)	12.93(1.48)	1.49(1.22)
	BVSS-BI	3.62(0.68)	0.09(0.18)	17.35(0.12)	7.79(0.17)	20.97(0.65)	7.88(0.18)
	BVSS	3.51(0.52)	0.12(0.32)	11.95(1.73)	0.71(0.45)	15.46(1.71)	0.83(0.62)
	BV-BI	3.02(0.47)	0.38(0.42)	10.22(1.75)	0.85(0.84)	13.24(1.83)	1.24(1.03)
	BV	2.82(0.56)	0.47(0.63)	8.97(1.65)	0.65(0.52)	11.79(1.78)	1.12(0.99)
Error 5	RBVSS – BI	3.88(0.26)	0.02(0.13)	17.99(0.44)	0.01(0.36)	21.87(0.34)	0.03(0.15)
Lognormal(0,1)	RBVSS	3.79(0.45)	0.06(0.18)	16.12(0.31)	0.28(0.76)	19.91(0.31)	0.34(0.77)
	RBV-BI	3.41(0.51)	0.21(0.73)	12.21(0.43)	1.54(1.37)	15.62(0.43)	1.75(1.52)
	RBV	3.33(0.31)	0.30(0.37)	10.15(0.40)	0.97(0.96)	13.48(0.40)	1.27(1.25)
	BVSS-BI	3.47(0.64)	0.07(0.25)	17.47(0.21)	8.04(0.33)	20.94(0.70)	8.11(0.25)
	BVSS	3.39(0.25)	0.15(0.35)	13.22(0.94)	0.52(0.86)	16.61(1.14)	0.67(0.96)
	BV-BI	2.98(0.61)	0.28(0.60)	10.29(1.48)	1.15(1.33)	13.27(1.76)	1.43(1.48)
	BV	2.79(0.67)	0.33(0.66)	8.89(1.68)	1.11(1.33)	11.68(0.79)	1.44(1.43)

Table A.18: Identification results for $n = 200, m = 60, p = 3$. $mean(sd)$ of true positives (TP) and false positives (FP) under random intercept model based on 100 replicates at the 70th quantile.

$n = 200 \quad m = 60$		Main		Interaction		Total	
$\theta = 0.7$		TP	FP	TP	FP	TP	FP
Error 1	RBVSS – BI	3.89(0.33)	0.03(0.35)	17.88(0.32)	0.03(0.33)	21.77(0.35)	0.06(0.35)
N(0,1)	RBVSS	3.76(0.34)	0.10(1.22)	15.09(0.75)	0.10(0.99)	18.85(0.45)	0.20(1.06)
	RBV-BI	3.41(1.52)	0.19(4.21)	11.11(0.70)	1.54(1.21)	14.52(0.92)	1.63(1.36)
	RBV	3.33(1.14)	0.32(0.58)	9.85(1.10)	0.95(1.65)	13.18(0.95)	1.27(1.85)
	BVSS-BI	3.79(0.39)	0.05(0.25)	17.10(0.35)	0.07(0.34)	20.89(0.25)	0.12(0.52)
	BVSS	3.52(0.25)	0.15(0.44)	14.19(0.45)	0.18(0.25)	17.71(0.76)	0.33(0.48)
	BV-BI	3.16(0.45)	0.24(0.25)	11.25(1.14)	0.95(0.31)	14.41(1.14)	1.19(0.56)
	BV	3.07(0.39)	0.39(0.32)	8.53(0.84)	0.68(0.47)	11.60(1.14)	1.07(0.33)
Error 2	RBVSS – BI	3.92(0.32)	0.04(0.38)	17.75(0.43)	0.02(0.36)	21.67(0.36)	0.06(0.41)
$t(2)$	RBVSS	3.81(0.38)	0.09(0.38)	15.21(0.33)	0.11(0.38)	19.02(0.39)	0.20(0.59)
	RBV-BI	3.41(0.38)	0.21(0.29)	10.28(0.81)	0.89(0.41)	13.69(0.71)	1.10(0.45)
	RBV	3.35(0.41)	0.33(0.39)	8.37(1.15)	0.41(0.45)	11.72(1.10)	0.74(0.45)
	BVSS-BI	3.42(0.67)	0.11(0.29)	17.69(0.12)	20.06(0.31)	21.11(1.67)	20.17(0.34)
	BVSS	3.31(1.08)	0.22(0.33)	13.47(1.60)	0.23(0.45)	16.78(1.66)	0.45(0.45)
	BV-BI	2.86(1.21)	0.38(0.42)	11.12(1.97)	1.12(0.32)	13.98(1.06)	1.50(0.45)
	BV	2.41(0.86)	0.51(0.29)	9.10(0.12)	0.86(0.28)	11.51(1.74)	1.37(0.45)
Error 3	RBVSS – BI	3.96(0.15)	0.02(0.17)	17.96(0.21)	0.03(0.15)	21.92(0.13)	0.05(0.30)
Laplace(0,2)	RBVSS	3.78(0.32)	0.07(0.18)	15.07(0.18)	0.08(0.45)	18.85(0.89)	0.15(0.23)
	RBV-BI	3.47(0.59)	0.19(0.35)	10.89(0.31)	1.75(0.35)	14.36(0.89)	1.94(0.67)
	RBV	3.37(1.21)	0.29(0.57)	8.87(0.34)	0.62(0.67)	12.19(1.14)	0.91(0.55)
	BVSS-BI	3.35(1.48)	0.11(0.35)	17.54(0.18)	17.79(0.15)	20.89(1.48)	17.90(0.35)
	BVSS	3.26(0.75)	0.18(0.34)	13.56(0.49)	0.55(0.45)	16.82(1.49)	0.73(0.56)
	BV-BI	3.03(0.86)	0.25(0.45)	10.74(1.82)	1.08(0.35)	13.77(0.17)	1.33(0.43)
	BV	2.92(1.14)	0.43(0.56)	8.34(1.23)	0.68(0.32)	11.26(1.82)	1.11(0.30)
Error 4	RBVSS – BI	3.88(0.25)	0.02(0.35)	17.78(0.41)	0.04(0.54)	21.66(0.32)	0.06(0.35)
20%Laplace(0,1) +80%Laplace(0, $\sqrt{5}$)	RBVSS	3.73(0.38)	0.08(0.48)	14.98(0.51)	0.10(0.34)	18.61(0.45)	0.18(0.57)
	RBV-BI	3.32(0.69)	0.17(0.47)	11.12(0.13)	0.82(0.32)	14.44(1.95)	0.99(0.29)
	RBV	3.20(0.79)	0.26(0.31)	9.75(1.52)	0.33(0.37)	12.95(1.58)	0.59(0.35)
	BVSS-BI	3.51(0.76)	0.09(0.32)	17.59(0.17)	20.05(0.33)	21.10(0.84)	20.14(0.15)
	BVSS	3.37(0.45)	0.18(0.55)	12.21(0.82)	0.18(0.42)	15.58(1.92)	0.36(0.45)
	BV-BI	2.74(0.75)	0.25(0.38)	9.98(1.30)	1.45(0.68)	12.72(1.95)	1.70(0.58)
	BV	2.45(0.55)	0.38(0.45)	8.25(1.10)	0.47(0.43)	10.70(1.52)	0.85(0.14)
Error 5	RBVSS – BI	3.95(0.56)	0.02(0.31)	17.77(0.68)	0.05(0.43)	21.72(0.34)	0.07(0.12)
Lognormal(0,1)	RBVSS	3.80(0.35)	0.07(1.12)	15.07(1.34)	0.10(0.42)	18.87(1.34)	0.17(1.45)
	RBV-BI	3.47(0.45)	0.25(0.35)	13.01(0.13)	0.95(0.45)	16.48(1.58)	1.10(0.89)
	RBV	3.33(1.35)	0.38(0.32)	10.27(0.96)	0.52(0.31)	13.60(1.64)	0.90(0.47)
	BVSS-BI	3.48(1.52)	0.19(0.22)	17.49(0.15)	19.56(0.18)	20.97(0.52)	19.65(0.25)
	BVSS	3.27(0.58)	0.25(0.28)	12.79(0.59)	0.19(0.35)	16.06(0.59)	0.44(0.15)
	BV-BI	2.78(0.82)	0.38(0.38)	9.98(1.15)	1.58(0.57)	12.76(1.92)	1.96(0.45)
	BV	2.45(1.38)	0.47(0.33)	8.12(1.58)	0.92(0.45)	10.57(1.30)	1.39(0.45)

Table A.19: *Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept - slope model based on 100 replicates at the 70th quantile.*

$n = 200$	$m = 30$	$\theta = 0.7$	MAD	NMAD	TMAD
Error 1 N(0,1)	RBVSS – BI	0.021(0.045)	0.002(0.015)	0.005(0.011)	
	RBVSS	0.048(0.056)	0.006(0.006)	0.011(0.008)	
	RBV-BI	0.085(0.018)	0.025(0.013)	0.041(0.013)	
	RBV	0.143(0.048)	0.051(0.013)	0.061(0.012)	
	BVSS-BI	0.029(0.011)	0.005(0.004)	0.009(0.003)	
	BVSS	0.064(0.013)	0.011(0.013)	0.026(0.009)	
	BV-BI	0.141(0.007)	0.049(0.001)	0.055(0.003)	
	BV	0.152(0.009)	0.063(0.001)	0.069(0.008)	
Error 2 $t(2)$	RBVSS – BI	0.028(0.017)	0.003(0.022)	0.015(0.008)	
	RBVSS	0.067(0.014)	0.009(0.023)	0.019(0.003)	
	RBV-BI	0.122(0.047)	0.017(0.005)	0.032(0.015)	
	RBV	0.149(0.032)	0.033(0.007)	0.059(0.015)	
	BVSS-BI	0.057(0.024)	0.041(0.012)	0.038(0.003)	
	BVSS	0.071(0.018)	0.053(0.012)	0.055(0.007)	
	BV-BI	0.132(0.025)	0.118(0.009)	0.143(0.048)	
	BV	0.219(0.019)	0.143(0.014)	0.167(0.015)	
Error 3 Laplace(0,2)	RBVSS – BI	0.031(0.024)	0.004(0.026)	0.011(0.007)	
	RBVSS	0.109(0.025)	0.018(0.005)	0.022(0.008)	
	RBV-BI	0.147(0.025)	0.043(0.017)	0.062(0.035)	
	RBV	0.181(0.035)	0.068(0.009)	0.089(0.012)	
	BVSS-BI	0.165(0.055)	0.036(0.027)	0.041(0.015)	
	BVSS	0.178(0.098)	0.055(0.003)	0.049(0.019)	
	BV-BI	0.221(0.015)	0.102(0.011)	0.109(0.007)	
	BV	0.251(0.035)	0.141(0.008)	0.147(0.008)	
Error 4 20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)	RBVSS – BI	0.035(0.027)	0.004(0.011)	0.025(0.007)	
	RBVSS	0.098(0.055)	0.008(0.005)	0.033(0.015)	
	RBV-BI	0.122(0.031)	0.031(0.006)	0.043(0.007)	
	RBV	0.151(0.038)	0.049(0.015)	0.062(0.015)	
	BVSS-BI	0.144(0.036)	0.017(0.033)	0.032(0.019)	
	BVSS	0.166(0.121)	0.028(0.017)	0.065(0.023)	
	BV-BI	0.188(0.023)	0.112(0.024)	0.124(0.015)	
	BV	0.203(0.075)	0.131(0.029)	0.169(0.033)	
Error 5 Lognormal(0,1)	RBVSS – BI	0.033(0.056)	0.005(0.045)	0.019(0.016)	
	RBVSS	0.055(0.015)	0.007(0.032)	0.023(0.012)	
	RBV-BI	0.149(0.023)	0.027(0.025)	0.031(0.022)	
	RBV	0.169(0.057)	0.046(0.008)	0.055(0.011)	
	BVSS-BI	0.147(0.043)	0.011(0.012)	0.031(0.056)	
	BVSS	0.154(0.128)	0.020(0.035)	0.052(0.086)	
	BV-BI	0.165(0.015)	0.111(0.054)	0.109(0.016)	
	BV	0.187(0.023)	0.176(0.048)	0.146(0.027)	

Table A.20: *Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept - slope model based on 100 replicates at the 70th quantile.*

$n = 200$	$m = 60$	$\theta = 0.7$	MAD	NMAD	TMAD
Error 1 N(0,1)	RBVSS – BI	0.026(0.018)	0.002(0.012)	0.015(0.005)	
	RBVSS	0.036(0.032)	0.007(0.024)	0.023(0.008)	
	RBV-BI	0.079(0.042)	0.021(0.018)	0.049(0.015)	
	RBV	0.159(0.028)	0.044(0.009)	0.069(0.025)	
	BVSS-SI	0.033(0.013)	0.008(0.023)	0.024(0.008)	
	BVSS	0.047(0.025)	0.015(0.026)	0.037(0.023)	
	BV-BI	0.161(0.019)	0.041(0.017)	0.052(0.018)	
	BV	0.171(0.021)	0.059(0.033)	0.077(0.018)	
Error 2 $t(2)$	RBVSS – BI	0.031(0.023)	0.004(0.015)	0.011(0.008)	
	RBVSS	0.044(0.021)	0.008(0.037)	0.020(0.012)	
	RBV-BI	0.139(0.028)	0.013(0.019)	0.039(0.027)	
	RBV	0.159(0.043)	0.026(0.021)	0.051(0.021)	
	BVSS-BI	0.067(0.024)	0.028(0.005)	0.043(0.011)	
	BVSS	0.085(0.026)	0.039(0.008)	0.057(0.011)	
	BV-BI	0.172(0.034)	0.121(0.013)	0.085(0.082)	
	BV	0.211(0.029)	0.141(0.043)	0.139(0.048)	
Error 3 Laplace(0,2)	RBVSS – BI	0.027(0.084)	0.003(0.016)	0.014(0.011)	
	RBVSS	0.059(0.053)	0.014(0.005)	0.028(0.017)	
	RBV-BI	0.127(0.057)	0.038(0.029)	0.047(0.035)	
	RBV	0.148(0.045)	0.058(0.032)	0.062(0.012)	
	BVSS-BI	0.159(0.053)	0.045(0.038)	0.051(0.021)	
	BVSS	0.190(0.081)	0.059(0.009)	0.062(0.029)	
	BV-BI	0.268(0.037)	0.087(0.010)	0.113(0.007)	
	BV	0.291(0.045)	0.113(0.013)	0.157(0.018)	
Error 4 20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)	RBVSS – BI	0.026(0.023)	0.003(0.009)	0.022(0.019)	
	RBVSS	0.063(0.075)	0.015(0.011)	0.049(0.027)	
	RBV-BI	0.121(0.033)	0.041(0.017)	0.067(0.019)	
	RBV	0.152(0.039)	0.067(0.021)	0.079(0.031)	
	BVSS-BI	0.165(0.031)	0.027(0.039)	0.054(0.025)	
	BVSS	0.179(0.112)	0.036(0.034)	0.076(0.032)	
	BV-BI	0.213(0.039)	0.114(0.054)	0.111(0.021)	
	BV	0.241(0.062)	0.131(0.025)	0.151(0.047)	
Error 5 Lognormal(0,1)	RBVSS – BI	0.019(0.064)	0.003(0.065)	0.014(0.021)	
	RBVSS	0.048(0.072)	0.009(0.056)	0.029(0.028)	
	RBV-BI	0.129(0.019)	0.038(0.035)	0.047(0.031)	
	RBV	0.154(0.084)	0.077(0.021)	0.062(0.026)	
	BVSS-BI	0.139(0.032)	0.012(0.027)	0.038(0.026)	
	BVSS	0.159(0.128)	0.026(0.073)	0.054(0.086)	
	BV-BI	0.181(0.031)	0.057(0.064)	0.076(0.025)	
	BV	0.191(0.036)	0.089(0.123)	0.164(0.067)	

Table A.21: *Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept model based on 100 replicates at the 70th quantile.*

$n = 200$	$m = 30$	$\theta = 0.7$	MAD	NMAD	TMAD
Error 1 N(0,1)	RBVSS – BI	0.018(0.044)	0.001(0.011)	0.012(0.015)	
	RBVSS	0.031(0.013)	0.008(0.018)	0.019(0.021)	
	RBV-BI	0.061(0.026)	0.036(0.018)	0.033(0.012)	
	RBV	0.075(0.045)	0.049(0.014)	0.068(0.023)	
	BVSS-BI	0.031(0.012)	0.005(0.014)	0.019(0.011)	
	BVSS	0.042(0.028)	0.009(0.037)	0.027(0.032)	
	BV-BI	0.071(0.018)	0.019(0.033)	0.039(0.027)	
	BV	0.085(0.035)	0.037(0.025)	0.057(0.015)	
Error 2 $t(2)$	RBVSS – BI	0.023(0.134)	0.003(0.016)	0.007(0.025)	
	RBVSS	0.045(0.027)	0.009(0.023)	0.019(0.007)	
	RBV-BI	0.067(0.023)	0.021(0.009)	0.043(0.011)	
	RBV	0.084(0.043)	0.037(0.006)	0.063(0.011)	
	BVSS-BI	0.132(0.138)	0.021(0.026)	0.050(0.034)	
	BVSS	0.217(0.117)	0.039(0.008)	0.072(0.026)	
	BV-BI	0.312(0.035)	0.121(0.021)	0.124(0.035)	
	BV	0.424(0.043)	0.158(0.038)	0.143(0.014)	
Error 3 Laplace(0,2)	RBVSS – BI	0.028(0.016)	0.004(0.008)	0.014(0.012)	
	RBVSS	0.039(0.035)	0.009(0.025)	0.028(0.006)	
	RBV-BI	0.108(0.026)	0.022(0.034)	0.046(0.014)	
	RBV	0.125(0.034)	0.039(0.019)	0.068(0.016)	
	BVSS-BI	0.046(0.024)	0.017(0.004)	0.034(0.007)	
	BVSS	0.055(0.036)	0.025(0.011)	0.045(0.008)	
	BV-BI	0.159(0.015)	0.125(0.012)	0.113(0.004)	
	BV	0.172(0.019)	0.141(0.008)	0.142(0.005)	
Error 4 20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)	RBVSS – BI	0.028(0.007)	0.003(0.009)	0.011(0.005)	
	RBVSS	0.042(0.008)	0.008(0.014)	0.026(0.010)	
	RBV-BI	0.119(0.033)	0.019(0.011)	0.042(0.036)	
	RBV	0.134(0.029)	0.031(0.012)	0.062(0.026)	
	BVSS-BI	0.058(0.043)	0.026(0.019)	0.027(0.019)	
	BVSS	0.074(0.011)	0.036(0.018)	0.036(0.015)	
	BV-BI	0.143(0.023)	0.115(0.009)	0.122(0.012)	
	BV	0.163(0.034)	0.142(0.012)	0.153(0.005)	
Error 5 Lognormal(0,1)	RBVSS – BI	0.035(0.031)	0.005(0.032)	0.011(0.002)	
	RBVSS	0.047(0.012)	0.012(0.000)	0.034(0.001)	
	RBV-BI	0.072(0.019)	0.035(0.005)	0.045(0.005)	
	RBV	0.112(0.018)	0.057(0.015)	0.058(0.006)	
	BVSS-BI	0.069(0.011)	0.024(0.004)	0.023(0.002)	
	BVSS	0.107(0.009)	0.041(0.008)	0.038(0.001)	
	BV-BI	0.138(0.011)	0.072(0.012)	0.117(0.005)	
	BV	0.169(0.015)	0.112(0.005)	0.128(0.006)	

Table A.22: Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept model based on 100 replicates at the 70th quantile.

$n = 200$	$m = 60$	$\theta = 0.7$	MAD	NMAD	TMAD
Error 1 N(0,1)	RBVSS – BI	0.015(0.008)	0.002(0.001)	0.011(0.001)	
	RBVSS	0.025(0.027)	0.015(0.014)	0.026(0.016)	
	RBV-BI	0.163(0.039)	0.036(0.039)	0.068(0.031)	
	RBV	0.221(0.066)	0.118(0.098)	0.125(0.085)	
	BVSS-BI	0.023(0.012)	0.004(0.005)	0.021(0.011)	
	BVSS	0.031(0.015)	0.012(0.006)	0.031(0.001)	
	BV-BI	0.159(0.026)	0.069(0.002)	0.054(0.004)	
	BV	0.228(0.032)	0.132(0.003)	0.164(0.006)	
Error 2 $t(2)$	RBVSS – BI	0.020(0.021)	0.001(0.006)	0.009(0.007)	
	RBVSS	0.041(0.025)	0.009(0.002)	0.021(0.001)	
	RBV-BI	0.106(0.033)	0.022(0.003)	0.046(0.005)	
	RBV	0.132(0.017)	0.061(0.006)	0.062(0.003)	
	BVSS-BI	0.127(0.058)	0.025(0.013)	0.024(0.015)	
	BVSS	0.141(0.035)	0.036(0.012)	0.038(0.018)	
	BV-BI	0.211(0.029)	0.059(0.033)	0.076(0.044)	
	BV	0.252(0.035)	0.075(0.043)	0.118(0.043)	
Error 3 Laplace(0,2)	RBVSS – BI	0.031(0.014)	0.004(0.016)	0.012(0.002)	
	RBVSS	0.042(0.004)	0.011(0.009)	0.023(0.002)	
	RBV-BI	0.136(0.015)	0.023(0.007)	0.037(0.012)	
	RBV	0.152(0.029)	0.037(0.005)	0.055(0.008)	
	BVSS-BI	0.061(0.037)	0.021(0.008)	0.028(0.009)	
	BVSS	0.078(0.015)	0.036(0.019)	0.038(0.008)	
	BV-BI	0.214(0.035)	0.049(0.015)	0.050(0.005)	
	BV	0.234(0.029)	0.068(0.017)	0.067(0.012)	
Error 4 20%Laplace(0,1) +80% Laplace(0, $\sqrt{5}$)	RBVSS – BI	0.016(0.003)	0.003(0.001)	0.017(0.003)	
	RBVSS	0.034(0.016)	0.012(0.001)	0.028(0.006)	
	RBV-BI	0.117(0.016)	0.024(0.011)	0.039(0.002)	
	RBV	0.133(0.027)	0.043(0.017)	0.062(0.004)	
	BVSS-BI	0.061(0.022)	0.037(0.009)	0.025(0.012)	
	BVSS	0.081(0.019)	0.063(0.012)	0.052(0.003)	
	BV-BI	0.153(0.023)	0.075(0.011)	0.078(0.004)	
	BV	0.185(0.026)	0.088(0.004)	0.092(0.008)	
Error 5 Lognormal(0,1)	RBVSS – BI	0.027(0.016)	0.005(0.002)	0.019(0.002)	
	RBVSS	0.039(0.004)	0.007(0.001)	0.028(0.012)	
	RBV-BI	0.124(0.036)	0.035(0.013)	0.041(0.014)	
	RBV	0.140(0.029)	0.056(0.012)	0.058(0.012)	
	BVSS-BI	0.052(0.021)	0.038(0.011)	0.031(0.017)	
	BVSS	0.065(0.042)	0.052(0.006)	0.053(0.011)	
	BV-BI	0.168(0.017)	0.069(0.007)	0.071(0.009)	
	BV	0.189(0.024)	0.078(0.011)	0.088(0.011)	

A.2.4 Figures corresponding to Tables of identification and estimation results

To facilitate better extraction of key information from Tables listing the identification and estimation results, we have provided Figures A.1–A.4 to display the results using bar-plots. Table A.23 shows the correspondence between Figures and their associated Tables.

Table A.23: *Correspondence between Figures 2–5 and associated Tables with $n=200$ and $p=3$. m is the number of lipids and θ is quantile level.*

Figure	Table	θ	m	Random Effects	Results
A.1	3.3	50	30	Random intercept-slope model	Identification
A.1	A.7	30	30	Random intercept-slope model	Identification
A.1	A.15	70	30	Random intercept-slope model	Identification
A.2	3.2	50	60	Random intercept-slope model	Identification
A.2	A.8	30	60	Random intercept-slope model	Identification
A.2	A.16	70	60	Random intercept-slope model	Identification
A.3	2.3	50	30	Random intercept-slope model	Estimation
A.3	A.11	30	30	Random intercept-slope model	Estimation
A.3	A.19	70	30	Random intercept-slope model	Estimation
A.4	2.4	50	60	Random intercept-slope model	Estimation
A.4	A.12	30	60	Random intercept-slope model	Estimation
A.4	A.20	70	60	Random intercept-slope model	Estimation

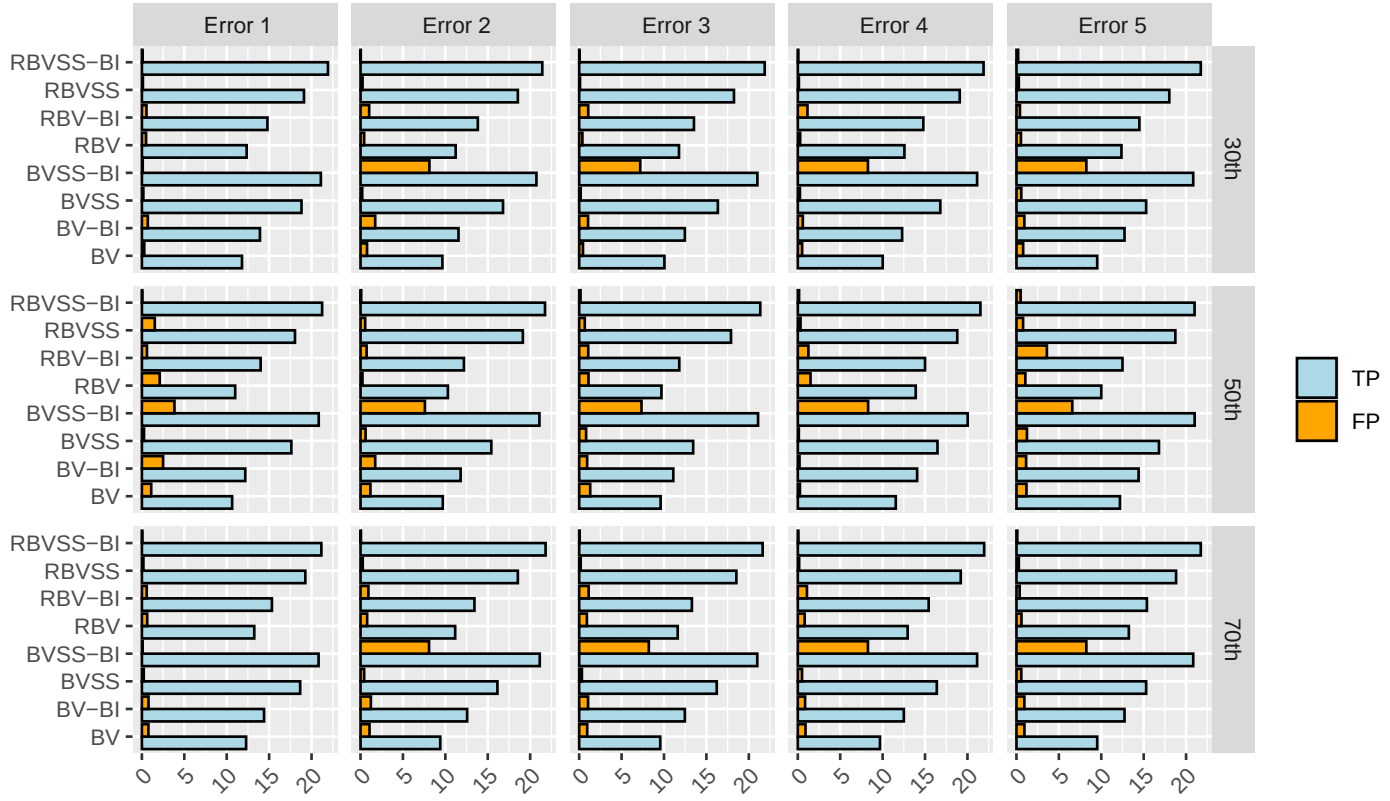


Figure A.1: Identification results for $n = 200, m = 30, p = 3$. $\text{mean}(\text{sd})$ of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 30th, 50th and 70th quantile levels.

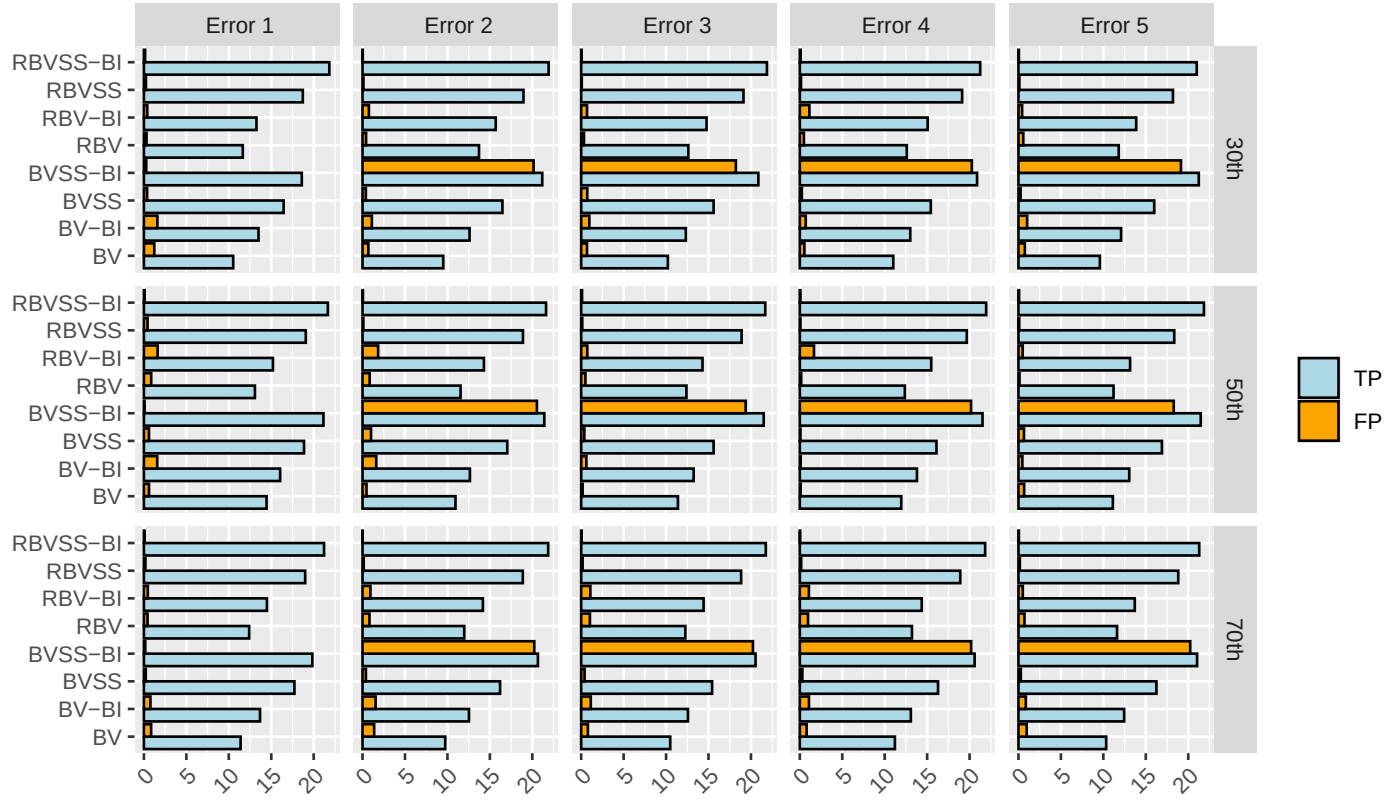


Figure A.2: Identification results for $n = 200, m = 60, p = 3$. $\text{mean}(\text{sd})$ of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at the 30th, 50th and 70th quantile levels.

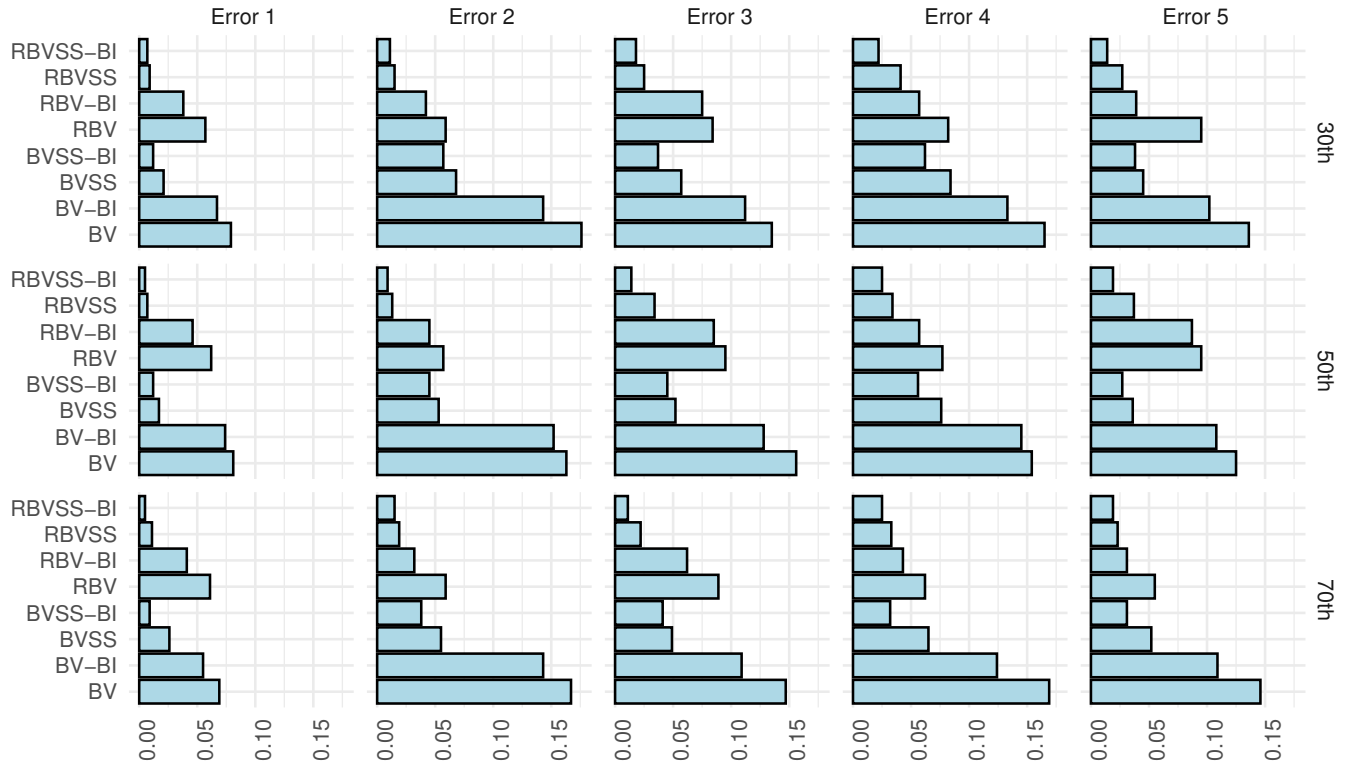


Figure A.3: Estimation accuracy for $n = 200, m = 30, p = 3$ under random intercept - slope model based on 100 replicates at the 30th, 50th and 70th quantile levels.

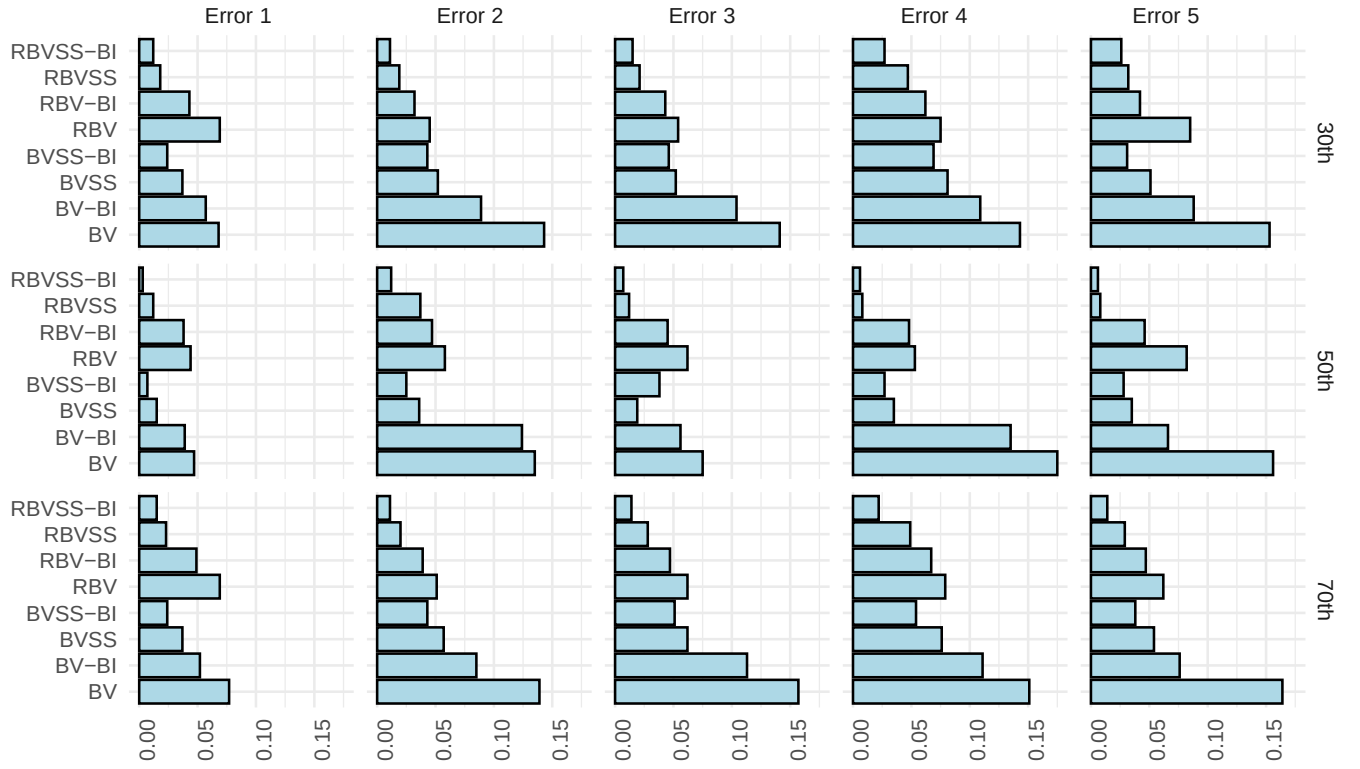


Figure A.4: Estimation accuracy for $n = 200, m = 60, p = 3$ under random intercept - slope model based on 100 replicates at the 30th, 50th and 70th quantile levels.

A.2.5 Computational cost

We have examined computational costs of the proposed RBVSS-BI under the random intercept-slope model (Model I) and the random intercept model (Model II) in the presence of $t(2)$ error (Error 2). Table A.24 shows the computational time under different combinations of sample size and dimensionality given $k=5$ repeated measures and $p = 3$ environmental factors. The coefficient vector of regression coefficients corresponding to intercept, main and interaction effects combined is $pm + m + 1$. For instance, if the number of lipids is $m=1000$, then the coefficient vector is of dimension $1000 \times 3 + 1000 + 3 + 1 = 4,004$. We can observe from the table that even if under such a large study size with $n=800$ and a total dimension of 4,004, the computational time is still reasonable being around 50 minutes to run 10,000 MCMC iterations.

Table A.24: *Computational time for RBVSS-BI with $k = 5$ repeated measurements under random intercept - slope model (Model I) and random intercept model (Model II) under $t(2)$ error (Error 2) at the 50th quantile. m : number of lipids. Time: CPU time (in minutes) for 10,000 MCMC iterations. The total number of regression coefficients to be estimated is $pm + m + 1$, where p is the number of environmental factors. In this study, $p = 3$.*

Model I					
$n = 200$		$n = 500$		$n = 800$	
# of lipids	Time	# of lipids	Time	# of lipids	Time
$m = 100$	1.15	$m = 400$	21.88	$m = 600$	35.06
$m = 200$	1.89	$m = 500$	23.12	$m = 800$	41.60
$m = 300$	7.64	$m = 600$	30.71	$m = 1,000$	50.84
Model II					
$n = 200$		$n = 500$		$n = 800$	
# of lipids	Time	# of lipids	Time	# of lipids	Time
$m = 100$	1.09	$m = 400$	21.55	$m = 600$	34.88
$m = 200$	1.79	$m = 500$	22.97	$m = 800$	40.87
$m = 300$	7.58	$m = 600$	29.95	$m = 1,000$	50.28

A.2.6 Hyper-parameters sensitivity analysis

We illustrate the sensitivity of RBVSS-BI in variable selection to the choice of hyperparameters for π_1 and π_2 . We consider five different Beta priors: (1) Beta(0.5, 0.5), which represents a U-shaped curve between (0, 1); (2) Beta(1, 1), which is essentially a uniform prior; (3) Beta(2, 2), which forms a quadratic curve; (4) Beta(1, 5), which is highly right-skewed; and (5) Beta(5, 1), which is highly left-skewed. As an example, we use the random intercept-slope model (Model I) and the random intercept model (Model II) with $n = 200, m = 60$ to generate data under $t(2)$ error (Error 2). Table [A.25](#) displays the identification performance of the median probability model (MPM) with different Beta priors. For all choices of Beta priors, the MPM model demonstrates high stability. Additionally, RBVSS-BI accurately identifies most true effects with low false positives across all cases. Consequently, we use Beta(1, 1) as the prior for π_1 and π_2 in this study.

Table A.25: *Sensitivity analysis of RBVSS - BI under random intercept - slope model (Model I) and random intercept model (Model II) with $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP), false positives (FP) and total mean absolute deviation (TMAD) based on 100 replicates.*

$n = 200$	$m = 60$	$p = 3$	Total		
			TP	FP	TMAD
Model I					
$\theta = 0.3$	Beta(0.5,0.5)	21.94(0.27)	0.04(0.25)	0.012(0.008)	
	Beta(1,1)	21.92(0.27)	0.03(0.25)	0.011(0.008)	
	Beta(2,2)	21.91(0.29)	0.05(0.45)	0.012(0.011)	
	Beta(1,5)	21.92(0.18)	0.03(0.28)	0.011(0.065)	
	Beta(5,1)	21.92(0.15)	0.04(0.71)	0.010(0.031)	
$\theta = 0.5$	Beta(0.5,0.5)	21.66(0.28)	0.04(0.15)	0.011(0.003)	
	Beta(1,1)	21.62(0.54)	0.05(0.45)	0.012(0.006)	
	Beta(2,2)	21.62(0.54)	0.06(0.49)	0.012(0.011)	
	Beta(1,5)	21.63(0.18)	0.05(0.38)	0.011(0.085)	
	Beta(5,1)	21.62(0.14)	0.04(0.35)	0.011(0.061)	
$\theta = 0.7$	Beta(0.5,0.5)	21.88(0.37)	0.04(0.35)	0.010(0.008)	
	Beta(1,1)	21.87(0.37)	0.02(0.35)	0.011(0.008)	
	Beta(2,2)	21.87(0.17)	0.02(0.39)	0.011(0.011)	
	Beta(1,5)	21.86(0.25)	0.03(0.48)	0.012(0.055)	
	Beta(5,1)	21.88(0.27)	0.04(0.75)	0.012(0.035)	
Model II					
$\theta = 0.3$	Beta(0.5,0.5)	21.96(0.26)	0.07(0.38)	0.012(0.007)	
	Beta(1,1)	21.97(0.26)	0.07(0.38)	0.012(0.007)	
	Beta(2,2)	21.98(0.39)	0.06(0.39)	0.014(0.010)	
	Beta(1,5)	21.97(0.24)	0.05(0.18)	0.013(0.075)	
	Beta(5,1)	21.95(0.54)	0.04(0.25)	0.012(0.029)	
$\theta = 0.5$	Beta(0.5,0.5)	21.79(0.35)	0.10(0.45)	0.012(0.007)	
	Beta(1,1)	21.82(0.33)	0.09(0.32)	0.012(0.007)	
	Beta(2,2)	21.81(0.39)	0.10(0.29)	0.011(0.016)	
	Beta(1,5)	21.83(0.18)	0.09(0.35)	0.013(0.035)	
	Beta(5,1)	21.82(0.17)	0.11(0.25)	0.012(0.061)	
$\theta = 0.7$	Beta(0.5,0.5)	21.70(0.36)	0.06(0.41)	0.010(0.007)	
	Beta(1,1)	21.67(0.36)	0.06(0.41)	0.009(0.007)	
	Beta(2,2)	21.66(0.28)	0.07(0.39)	0.010(0.010)	
	Beta(1,5)	21.67(0.21)	0.06(0.68)	0.010(0.025)	
	Beta(5,1)	21.65(0.19)	0.06(0.35)	0.011(0.031)	

We also assess the sensitivity of RBVSS-BI to the selection of Gamma hyperpriors on η_1^2 , η_2^2 , and τ . Specifically, we test the shape parameter of the Gamma prior with five different values: $\{0.1, 0.5, 1, 2, 5\}$, covering a range from highly skewed exponential shapes to highly diffuse unimodal shapes. The rate parameter is fixed at $\{1, 2, 5\}$, and we explore different combinations of shape and rate parameters on a two-dimensional grid. Table A.26 presents the simulation results for some representative cases under Error 2. The RBVSS-BI consistently shows stable performance with high true positives (TP) and low false positives (FP) across various Gamma priors. Similar patterns were observed for all other cases. For

this study, we use $\text{Gamma}(1, 1)$ for η_1^2 , η_2^2 , and τ across all scenarios.

Table A.26: *Sensitivity analysis of RBVSS - BI under random intercept - slope model (Model I) and random intercept model (Model II) with $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP), false positives (FP) and total mean absolute deviation (TMAD) based on 100 replicates.*

$n = 200$ $m = 60$ $p = 3$		Total		
		TP	FP	TMAD
Model I				
$\theta = 0.3$	Gamma (0.1,1)	21.90(0.32)	0.02(0.15)	0.013(0.003)
	Gamma (0.5,2)	21.93(0.28)	0.04(0.45)	0.012(0.006)
	Gamma (1,1)	21.92(0.27)	0.03(0.25)	0.011(0.008)
	Gamma (1,5)	21.90(0.18)	0.03(0.38)	0.013(0.085)
	Gamma (2,5)	21.89(0.14)	0.02(0.35)	0.014(0.061)
	Gamma (5,1)	21.91(0.14)	0.03(0.35)	0.011(0.061)
$\theta = 0.5$	Gamma (0.1,1)	21.64(0.26)	0.05(0.17)	0.011(0.003)
	Gamma (0.5,2)	21.63(0.54)	0.04(0.25)	0.011(0.006)
	Gamma (1,1)	21.62(0.54)	0.05(0.45)	0.012(0.006)
	Gamma (1,5)	21.58(0.16)	0.05(0.38)	0.013(0.077)
	Gamma (2,5)	21.63(0.13)	0.04(0.37)	0.011(0.061)
	Gamma (5,1)	21.62(0.14)	0.03(0.35)	0.012(0.061)
$\theta = 0.7$	Gamma (0.1,1)	21.88(0.28)	0.03(0.15)	0.011(0.003)
	Gamma (0.5,2)	21.85(0.54)	0.05(0.45)	0.012(0.006)
	Gamma (1,1)	21.87(0.37)	0.02(0.35)	0.011(0.008)
	Gamma (1,5)	21.89(0.18)	0.03(0.38)	0.010(0.085)
	Gamma (2,5)	21.87(0.14)	0.04(0.35)	0.011(0.061)
	Gamma (5,1)	21.88(0.14)	0.04(0.35)	0.012(0.061)
Model II				
$\theta = 0.3$	Gamma (0.1,1)	21.96(0.28)	0.06(0.15)	0.011(0.003)
	Gamma (0.5,2)	21.97(0.54)	0.05(0.45)	0.012(0.006)
	Gamma (1,1)	21.97(0.26)	0.07(0.38)	0.012(0.007)
	Gamma (1,5)	21.99(0.18)	0.03(0.38)	0.010(0.085)
	Gamma (2,5)	21.96(0.14)	0.04(0.35)	0.012(0.061)
	Gamma (5,1)	21.99(0.14)	0.04(0.35)	0.011(0.061)
$\theta = 0.5$	Gamma (0.1,1)	21.83(0.28)	0.08(0.15)	0.011(0.003)
	Gamma (0.5,2)	21.80(0.54)	0.06(0.45)	0.012(0.006)
	Gamma (1,1)	21.82(0.33)	0.09(0.32)	0.012(0.007)
	Gamma (1,5)	21.84(0.18)	0.03(0.38)	0.013(0.085)
	Gamma (2,5)	21.79(0.14)	0.04(0.35)	0.014(0.061)
	Gamma (5,1)	21.81(0.14)	0.04(0.35)	0.012(0.061)
$\theta = 0.7$	Gamma (0.1,1)	21.67(0.28)	0.04(0.15)	0.010(0.003)
	Gamma (0.5,2)	21.66(0.54)	0.05(0.45)	0.012(0.006)
	Gamma (1,1)	21.67(0.36)	0.06(0.41)	0.009(0.007)
	Gamma (1,5)	21.71(0.18)	0.03(0.38)	0.013(0.085)
	Gamma (2,5)	21.66(0.14)	0.04(0.35)	0.012(0.061)
	Gamma (5,1)	21.67(0.14)	0.04(0.35)	0.009(0.061)

A.2.7 Credibility of identification results

The theoretical underpinnings and advantage of adopting MPM (median probability model) for variable selection with spike-and-slab priors have been justified in [Barbieri and Berger \(2004\)](#); [Barbieri et al. \(2021\)](#) among many others. Note that MPM can only be adopted if spike-and-slab priors have been incorporated. For demonstration purpose, we show the identification performance of RBVSS-BI using both the MPM model and 95% credible interval in Table A.27, under the random intercept - slope model (Model I) and the random intercept model Model II). The data has been generated with $n = 200, m = 60, p = 3$ and $t(2)$ error (Error 2). The total number of true positives is 25. Table A.27 indicates that for RBVSS-BI that incorporates spike-and-slab priors, the 95% credible interval yields smaller number of true positives compared to those obtained under the MPM model. The 95% credible interval generally yields inferior identification results for methods incorporating spike-and-slab priors.

Table A.27: *Credibility of identification results with $n = 200, m = 60, p = 3$ for RBVSS - BI by adopting both MPM and 95% CI. Mean (sd) of true positives (TP) and false positives (FP) based on 100 replicates.*

$n = 200 \quad m = 60 \quad p = 3$			Main		Interaction		Total	
			TP	FP	TP	FP	TP	FP
Model I								
$\theta = 0.3$	MPM		3.97 (0.35)	0.01 (0.48)	17.95 (0.31)	0.02 (0.18)	21.92 (0.27)	0.03 (0.25)
	95% CI		2.89 (0.24)	0.01 (0.27)	16.52 (0.19)	0.02 (0.13)	19.41 (0.31)	0.03 (0.15)
$\theta = 0.5$	MPM		3.87 (0.47)	0.02 (0.21)	17.75 (0.42)	0.03 (0.35)	21.62 (0.54)	0.05 (0.45)
	95% CI		2.85 (0.21)	0.01 (0.13)	16.92 (0.32)	0.02 (0.17)	19.77 (0.34)	0.03 (0.13)
$\theta = 0.7$	MPM		3.92 (0.35)	0.01 (0.48)	17.95 (0.36)	0.01 (0.38)	21.87 (0.37)	0.02 (0.35)
	95% CI		2.95 (0.21)	0.02 (0.31)	16.02 (0.25)	0.01 (0.28)	18.97 (0.14)	0.03 (0.46)
Model II								
$\theta = 0.3$	MPM		3.98 (0.26)	0.03 (0.45)	17.99 (0.25)	0.04 (0.36)	21.97 (0.26)	0.07 (0.38)
	95% CI		2.86 (0.31)	0.02 (0.37)	16.88 (0.32)	0.04 (0.53)	19.74 (0.31)	0.06 (0.27)
$\theta = 0.5$	MPM		3.97 (0.34)	0.05 (0.23)	17.85 (0.24)	0.04 (0.28)	21.82 (0.33)	0.09 (0.32)
	95% CI		2.91 (0.31)	0.03 (0.13)	16.36 (0.28)	0.03 (0.33)	19.27 (0.34)	0.06 (0.43)
$\theta = 0.7$	MPM		3.92 (0.32)	0.04 (0.38)	17.75 (0.43)	0.02 (0.36)	21.67 (0.36)	0.06 (0.41)
	95% CI		2.91 (0.28)	0.04 (0.31)	16.21 (0.47)	0.03 (0.34)	19.12 (0.34)	0.07 (0.36)

A.2.8 Comparison between RBVSS-BI and interep

We further assess the performance of the proposed Bayesian method against *interep*, a regularized variable selection method based on penalized GEE (Generalized Estimating Equation) for interaction analysis with repeated measure data [Zhou et al. \(2019, 2022b\)](#). The frequentist *interep* has also been proposed to conduct gene-environment interaction analysis under repeated measure ANOVA design with high-dimensional omics features [King et al. \(2015\)](#); [Zhou et al. \(2019\)](#). Simulations in our study indicates that RBVSS-BI has superior performance over the rest of Bayesian methods under comparison, therefore we conduct a direct comparison between RBVSS-BI and *interep* in terms of identification results and estimation accuracy. The method *interep* has been implemented via R package *interep* [Zhou et al.](#). The tuning parameters have been chosen by using the cross-validation function provided by package *interep*. Besides, we have specified “AR-1” correlation in the numerical experiments. The package has also implemented two other correlation structures, “independence” and “exchangeable”. Both correlations lead to similar results which have been omitted. More details of using the package are available in [Zhou et al. \(2022b\)](#).

The data has been generated under random intercept-slope model with $n = 200, m = 60, p = 3$ according to Section 2.3. We simulate data based on exactly the same data generating model described in Section 2.3. So the total number of true positives is still 25 with fixed effects generated from the same uniform distribution $U(0.4, 0.8)$. Overall, RBVSS-BI has outperformed *interep* in terms of both identification and estimation, especially under heavy-tailed errors. Table A.28 shows that under all the 3 quantile levels, the number of TPs of RBVSS-BI and *interep* are very close, with *interep* having slightly larger FPs. When we examine the performance under heavy-tailed errors, we can observe a much higher number of FPs generated by *interep*. For example, with $\theta=0.5$ and Error 2, the *interep* has resulted a FP of 19.49 (sd 0.25), which is much larger than the number of FPs produced by RBVSS-BI. Similarly, we observe the same pattern for estimation errors shown in Table A.29 which indicates *interep* has much larger estimation errors under the heavy-tailed errors.

Overall, numerical comparison shows inferior performance of *interep* under heavy-tailed

errors, which is consistent with published studies regarding non-robustness of GEE based methods [Qu and Song \(2004\)](#); [Fitzmaurice et al. \(2012\)](#). Although GEE can provide consistent parameter estimates even when correlation structure is mis-specified, it is not robust to data contamination and outliers in the responses [Qu and Song \(2004\)](#); [Fitzmaurice et al. \(2012\)](#). In addition, the poor performance may also be due to the data generating model. We have simulated data under the mixed-effects model described in [Section 2.3](#), which is an individual-level model and utilizes random effects to capture correlations among repeated measurements. On the other hand, GEE is a population-level method that provides population-averaged parameter estimates [Fitzmaurice et al. \(2012\)](#).

Table A.28: Identification performance of RBVSS-BI and interep under random intercept - slope model with $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) based on 100 replicates.

$n = 200 \quad m = 60 \quad p = 3$			Main		Interaction		Total	
			TP	FP	TP	FP	TP	FP
$\theta = 0.3$								
Error 1	RBVSS – BI		3.92(0.24)	0.07(0.25)	17.97(0.25)	0.04(0.24)	21.89(0.34)	0.11(0.25)
	interep		3.87(0.56)	0.53(0.21)	16.89(0.27)	1.37(0.14)	20.76(0.76)	1.90(0.25)
Error 2	RBVSS – BI		3.96(0.35)	0.02(0.48)	17.96(0.31)	0.03(0.18)	21.92(0.27)	0.05(0.25)
	interep		3.89(0.68)	9.09(0.32)	16.98(0.28)	9.89(0.34)	20.87(0.45)	18.98(0.35)
Error 3	RBVSS – BI		3.95(0.21)	0.01(0.22)	17.95(0.21)	0.03(0.15)	21.90(0.19)	0.04(0.16)
	interep		3.85(0.45)	9.14(0.15)	17.02(0.74)	9.59(0.25)	20.87(0.96)	18.73(0.42)
Error 4	RBVSS – BI		3.90(0.45)	0.04(0.56)	17.61(1.34)	0.04(0.23)	21.51(0.30)	0.08(0.32)
	interep		3.87(0.55)	9.74(0.34)	17.21(0.92)	10.01(0.24)	21.08(0.36)	19.75(0.45)
Error 5	RBVSS – BI		3.89(0.45)	0.02(0.35)	17.51(0.68)	0.05(0.45)	21.40(0.47)	0.07(0.38)
	interep		3.84(0.55)	9.38(0.45)	17.18(0.95)	9.39(0.45)	21.02(0.49)	18.77(0.55)
$\theta = 0.5$								
Error 1	RBVSS – BI		3.93(0.24)	0.04(0.22)	17.90(0.67)	0.03(0.17)	21.83(0.54)	0.07(0.11)
	interep		3.89(0.76)	0.49(0.21)	16.93(0.27)	1.29(0.14)	20.82(0.76)	1.78(0.25)
Error 2	RBVSS – BI		3.91(0.47)	0.02(0.21)	17.88(0.42)	0.02(0.35)	21.79(0.34)	0.04(0.45)
	interep		3.88(0.51)	9.48(0.35)	17.29(0.21)	10.01(0.51)	21.17(0.56)	19.49(0.25)
Error 3	RBVSS – BI		3.86(0.58)	0.05(0.47)	17.94(0.45)	0.03(0.33)	21.80(0.56)	0.08(0.47)
	interep		3.83(0.35)	9.28(0.52)	17.31(0.84)	9.99(0.29)	21.14(0.84)	19.27(0.34)
Error 4	RBVSS – BI		3.94(0.21)	0.03(0.23)	17.94(0.31)	0.02(0.21)	21.88(0.52)	0.05(0.23)
	interep		3.89(0.33)	9.72(0.27)	16.96(0.55)	10.18(0.24)	20.85(0.55)	19.90(0.14)
Error 5	RBVSS – BI		3.93(0.42)	0.02(0.21)	17.95(0.21)	0.03(0.22)	21.88(0.45)	0.05(0.33)
	interep		3.87(0.31)	9.79(0.35)	16.85(0.24)	10.25(0.28)	20.72(0.56)	20.04(0.23)
$\theta = 0.7$								
Error 1	RBVSS – BI		3.93(0.54)	0.03(0.55)	17.58(0.35)	0.03(0.44)	21.51(0.34)	0.06(0.25)
	interep		3.85(0.47)	0.64(0.35)	17.08(0.37)	1.19(0.21)	20.93(0.35)	1.83(0.25)
Error 2	RBVSS – BI		3.94(0.35)	0.01(0.48)	17.93(0.36)	0.02(0.38)	21.87(0.37)	0.03(0.35)
	interep		3.84(0.68)	9.82(0.32)	17.53(0.21)	9.89(0.34)	21.37(0.45)	19.71(0.35)
Error 3	RBVSS – BI		3.92(0.21)	0.02(0.32)	17.90(0.21)	0.03(0.12)	21.82(0.39)	0.05(0.28)
	interep		3.89(0.45)	9.48(0.28)	17.37(0.74)	10.04(0.25)	21.26(0.96)	19.52(0.42)
Error 4	RBVSS – BI		3.92(0.57)	0.03(0.56)	17.88(1.34)	0.02(0.31)	21.80(0.30)	0.05(0.32)
	interep		3.85(0.32)	9.19(0.34)	16.99(0.92)	9.79(0.24)	20.84(0.36)	18.98(0.45)
Error 5	RBVSS – BI		3.89(0.45)	0.02(0.35)	17.57(0.68)	0.03(0.45)	21.46(0.28)	0.05(0.38)
	interep		3.84(0.55)	9.53(0.45)	17.46(0.75)	10.02(0.45)	21.30(0.49)	19.55(0.55)

Table A.29: *Estimation accuracy of RBVSS-BI and interep under random intercept - slope model with $n = 200, m = 60, p = 3$. mean(sd) of true positives (TP) and false positives (FP) based on 100 replicates.*

$n = 200$	$m = 60$	$p = 3$	MAD	NMAD	TMAD
$\theta = 0.3$					
Error 1	RBVSS – BI	0.020(0.018)	0.002(0.011)	0.011(0.005)	
	interep	0.031(0.035)	0.009(0.024)	0.016(0.008)	
Error 2	RBVSS – BI	0.024(0.024)	0.003(0.015)	0.013(0.008)	
	interep	0.124(0.021)	0.093(0.037)	0.109(0.012)	
Error 3	RBVSS – BI	0.031(0.084)	0.005(0.016)	0.014(0.011)	
	interep	0.132(0.053)	0.102(0.005)	0.114(0.017)	
Error 4	RBVSS – BI	0.028(0.023)	0.003(0.009)	0.021(0.009)	
	interep	0.137(0.075)	0.106(0.009)	0.112(0.027)	
Error 5	RBVSS – BI	0.036(0.064)	0.004(0.065)	0.023(0.021)	
	interep	0.126(0.072)	0.101(0.056)	0.111(0.028)	
$\theta = 0.5$					
Error 1	RBVSS – BI	0.023(0.018)	0.002(0.012)	0.013(0.005)	
	interep	0.029(0.032)	0.004(0.024)	0.015(0.008)	
Error 2	RBVSS – BI	0.028(0.023)	0.003(0.015)	0.011(0.008)	
	interep	0.118(0.021)	0.101(0.037)	0.111(0.012)	
Error 3	RBVSS – BI	0.026(0.084)	0.003(0.016)	0.012(0.011)	
	interep	0.122(0.053)	0.099(0.005)	0.118(0.017)	
Error 4	RBVSS – BI	0.027(0.023)	0.004(0.009)	0.024(0.019)	
	interep	0.119(0.075)	0.100(0.011)	0.116(0.027)	
Error 5	RBVSS – BI	0.016(0.064)	0.003(0.065)	0.011(0.021)	
	interep	0.118(0.072)	0.104(0.056)	0.115(0.028)	
$\theta = 0.7$					
Error 1	RBVSS – BI	0.024(0.018)	0.002(0.012)	0.013(0.005)	
	interep	0.030(0.032)	0.005(0.024)	0.017(0.008)	
Error 2	RBVSS – BI	0.027(0.023)	0.003(0.015)	0.010(0.008)	
	interep	0.125(0.021)	0.106(0.037)	0.114(0.012)	
Error 3	RBVSS – BI	0.025(0.084)	0.002(0.016)	0.011(0.011)	
	interep	0.131(0.053)	0.100(0.005)	0.120(0.017)	
Error 4	RBVSS – BI	0.024(0.023)	0.003(0.009)	0.019(0.019)	
	interep	0.128(0.075)	0.103(0.011)	0.112(0.027)	
Error 5	RBVSS – BI	0.020(0.064)	0.002(0.065)	0.013(0.021)	
	interep	0.118(0.072)	0.099(0.056)	0.100(0.028)	

A.3 Assessment of the convergence of MCMC chains

A.3.1 The convergence of MCMC chains under random intercept - slope model

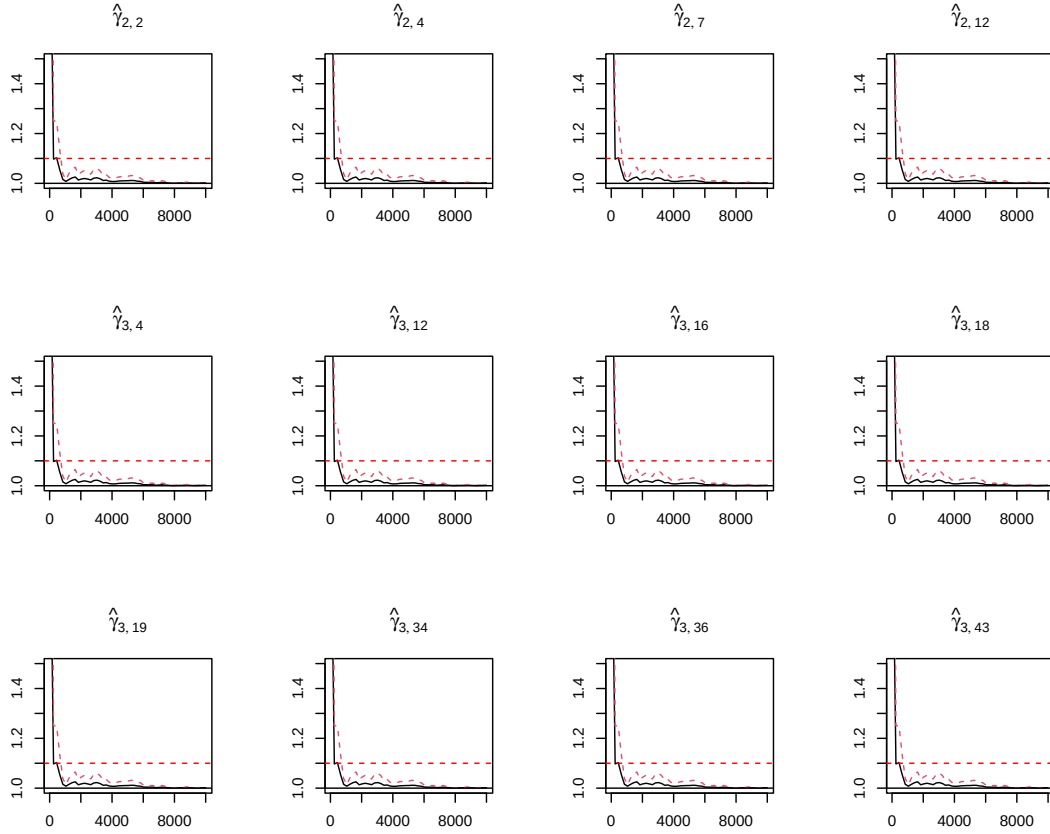


Figure A.5: *Potential scale reduction factor (PSRF) against iterations for nonzero coefficients in simulation with $n = 200, m = 60$ and $p = 3$ under Error 2 at 50th quantile. Black line: the PSRF. Red line: the threshold of 1.1.*

A.3.2 The convergence of MCMC chains under random intercept model

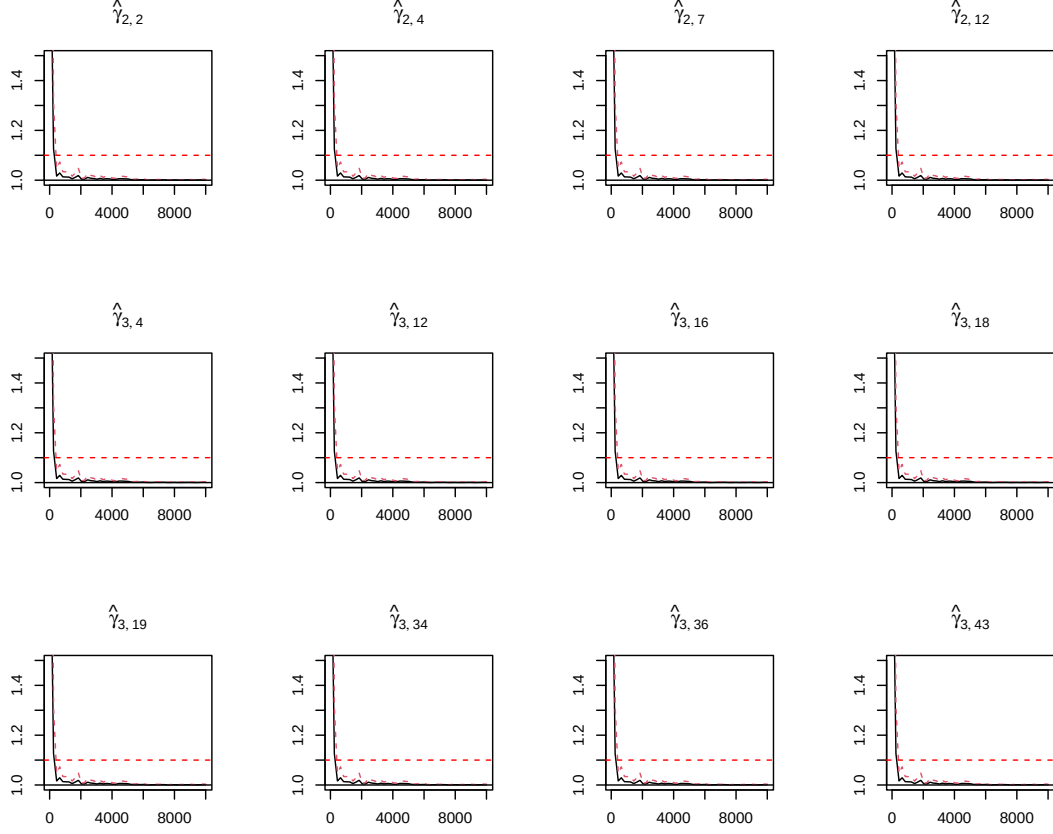


Figure A.6: *Potential scale reduction factor (PSRF) against iterations for nonzero coefficients in simulation with $n = 200, m = 60$ and $p = 3$ under Error 2 at 50th quantile. Black line: the PSRF. Red line: the threshold of 1.1.*

A.4 Estimation results for data analysis

Table A.30: *Real data analysis result from RBVSS under random intercept - slope model*

	Lipid	AE	DCR	PE
C18:3/16:1	-0.546	1.376	0.046	0.871
C20:0/16:1	-0.969	0.606	0.000	0.918
C18:1/16:1	-1.232	-0.111	0.000	-1.340
C20:4/16:0	0.000	0.000	0.000	1.232
C16:1/18:2	0.000	0.000	1.501	0.000
C18:4/16:0	1.329	0.000	0.376	0.000
C18:3/16:0	0.000	1.014	1.221	0.000
C18:2/18:1	0.010	-0.943	0.000	0.000
C20:2/18:1	0.172	0.466	0.000	1.139
C18:2/20:4	0.000	0.365	0.000	0.391
C18:3/20:4	0.000	0.000	0.000	-0.500

Table A.31: *Real data analysis result from RBVSS under model random intercept model*

	Lipid	AE	DCR	PE
C18:2/16:1	0.000	0.000	0.119	1.325
C20:0/16:1	-1.570	-0.746	0.000	0.000
C18:1/16:0	0.000	0.000	0.000	-4.460
C20:6/16:0	1.378	0.000	0.410	0.000
C20:0/18:3	0.000	1.486	0.000	0.783
C16:1/18:2	-0.198	0.000	0.000	0.000
C18:2/18:2	-0.155	0.000	1.215	0.000
C18:1/18:2	0.000	1.953	0.000	0.000
C22:6/18:1	0.000	0.895	0.000	1.320
C18:2/20:4	0.000	0.000	0.000	0.658
C18:1/20:4	1.083	0.000	0.000	-0.877

Table A.32: *Real data analysis result from BVSS-BI under random intercept - slope model*

	Lipid	AE	DCR	PE
C18:2/16:1	0.000	0.716	-0.773	0.433
C16:1/16:0	-1.988	0.000	0.000	0.000
C20:6/16:0	0.585	0.667	-3.315	-0.537
C20:0/18:3	6.523	0.000	0.000	0.000
C16:1/18:2	0.000	1.861	2.331	1.175
C18:1/18:2	8.083	0.000	0.000	0.000
C18:2/18:1	0.263	0.000	0.000	0.000
C22:6/18:1	0.000	3.049	0.361	1.120
C18:2/20:4	0.000	-0.396	0.053	1.729
C18:1/20:4	3.881	0.000	0.000	0.000

Table A.33: *Real data analysis result from BVSS-BI under random intercept model*

	Lipid	AE	DCR	PE
C16:1/16:0	0.000	0.301	-0.868	-1.489
C18:1/16:0	0.000	1.789	-0.710	-0.285
C20:6/16:0	-5.132	0.799	-0.849	-1.765
C20:0/18:3	-1.005	0.000	0.000	0.000
C16:1/18:2	-0.971	0.000	0.000	0.000
C18:2/18:2	0.000	0.217	-0.213	-1.839
C18:1/18:2	0.000	-0.511	-1.144	-0.364
C18:2/20:4	0.000	-0.032	0.001	-1.037
C18:1/20:4	2.625	0.000	0.000	0.000

Table A.34: *Real data analysis result from BVSS under random intercept - slope model*

	Lipid	AE	DCR	PE
C18:2/16:1	-4.946	0.000	0.000	0.000
C20:0/16:1	-1.179	-1.956	0.000	0.000
C18:1/16:0	0.000	0.000	0.000	-4.453
C20:6/16:0	1.280	-0.940	0.000	0.761
C16:1/18:2	0.000	0.000	-0.611	1.697
C18:2/18:2	-0.104	0.000	0.000	2.893
C18:1/18:2	1.975	0.000	0.000	0.000
C18:2/18:1	0.000	0.000	0.000	1.659
C18:1/20:4	4.561	0.000	0.000	0.000

Table A.35: *Real data analysis result from BVSS under random intercept model*

	Lipid	AE	DCR	PE
C18:2/16:1	0.000	0.186	0.000	0.000
C20:0/16:1	-3.741	0.000	0.000	-0.446
C16:1/16:0	-2.165	0.000	0.000	0.052
C18:1/16:0	0.000	0.000	0.000	0.261
C20:6/16:0	0.544	0.000	-0.447	0.000
C20:0/18:3	-0.858	0.000	0.000	0.000
C16:1/18:2	0.000	0.000	0.000	1.584
C18:1/18:2	0.000	0.000	0.000	-1.902
C18:2/20:4	0.312	0.000	0.000	0.776
C18:1/20:4	0.648	0.000	-0.045	-1.239

A.5 Posterior inference

A.5.1 RBVSS-BI

Priors

$$Y_{ij} = \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0 + \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1 + \mathbf{L}_{ij}^\top \boldsymbol{\gamma}_2 + \mathbf{W}_{ij}^\top \boldsymbol{\gamma}_3 + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i + \xi_1 \tilde{v}_{ij} + \tau^{-1/2} \xi_2 \sqrt{\tilde{v}_{ij}} z_{ij} \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$z_{ij} \stackrel{ind}{\sim} \text{N}(0, 1) \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\tilde{v}_{ij} | \tau \stackrel{ind}{\sim} \tau \exp(\tau \tilde{v}_{ij}) \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\boldsymbol{\gamma}_0 \sim \text{N}_3(0, \boldsymbol{\Sigma}_0)$$

$$\boldsymbol{\gamma}_1 \sim \text{N}_p(0, \boldsymbol{\Sigma}_1)$$

$$\gamma_{2l} | \phi_{1l}, s_{1l} \stackrel{ind}{\sim} \phi_{1l} \text{N}(0, s_{1l}) + (1 - \phi_{1l}) \delta_0(\gamma_{2l}) \quad l = 1, \dots, m$$

$$\phi_{1l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_1) \quad l = 1, \dots, m$$

$$\boldsymbol{\gamma}_{3l} | \phi_{2l}, s_{2l} \stackrel{ind}{\sim} \phi_{2l} \text{MVN}(\mathbf{0}, \text{diag}(s_{2l}, \dots, s_{2l})) + (1 - \phi_{2l}) \delta_0(\boldsymbol{\gamma}_{3l}) \quad l = 1, \dots, m$$

$$\phi_{2l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_2) \quad l = 1, \dots, m$$

$$s_{1l} | \eta_1^2 \stackrel{iid}{\sim} \frac{\eta_1^2}{2} \exp\left(-\frac{\eta_1^2}{2} s_{1l}\right) \quad l = 1, \dots, m$$

$$s_{2l} | \eta_2^2 \stackrel{ind}{\sim} \text{Gamma}\left(\frac{p+1}{2}, \frac{\eta_2^2}{2}\right) \quad l = 1, \dots, m$$

$$\tau \sim \text{Gamma}(c, d)$$

$$\eta_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\eta_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\boldsymbol{\alpha}_i \sim \text{MVN}(\mathbf{0}, \phi^2 \mathbf{I}) \quad i = 1, \dots, n$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\pi_1 \sim \text{Beta}(r_1, w_1)$$

$$\pi_2 \sim \text{Beta}(r_2, w_2)$$

Gibbs Sampler

- Let $\mu_{ij(-\gamma_0)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of $\boldsymbol{\gamma}_0$.

The full conditional distribution of $\boldsymbol{\gamma}_0$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\gamma}_0}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0})$ with mean

$$\boldsymbol{\mu}_{\boldsymbol{\gamma}_0} = \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{X}_{ij} (Y_{ij} - \mu_{ij(-\gamma_0)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0} = \left(\boldsymbol{\Sigma}_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{X}_{ij} \mathbf{X}_{ij}^\top \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_1)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1$, then the full conditional distribution of $\boldsymbol{\gamma}_1$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\gamma}_1}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_1})$ with mean

$$\boldsymbol{\mu}_{\boldsymbol{\gamma}_1} = \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_1} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{T}_{ij} (Y_{ij} - \mu_{ij(-\gamma_1)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_1} = \left(\boldsymbol{\Sigma}_1^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{T}_{ij} \mathbf{T}_{ij}^\top \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{2l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - L_{ijl} \gamma_{2l}$ and $l_{1l} = \pi(\gamma_{2l} = 0 | \text{rest})$, the conditional posterior distribution of the coefficient of lipid factor γ_{2l} is a spike-and-slab distribution:

$$\gamma_{2l} | \text{rest} \sim (1 - l_{1l}) \text{N}(\mu_{\gamma_{2l}}, \Sigma_{\gamma_{2l}}) + l_{1l} \delta_0(\gamma_{2l}),$$

where

$$\begin{aligned}\mu_{\gamma_{2l}} &= \Sigma_{\gamma_{2l}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\gamma_{2l}} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl}^2 + \frac{1}{s_{1l}} \right)^{-1}.\end{aligned}$$

Where

$$l_{1l} = \frac{\pi_1}{\pi_1 + (1 - \pi_1)(s_{1l})^{-\frac{1}{2}} (\Sigma_{\gamma_{2l}})^{\frac{1}{2}} \exp\left\{\frac{1}{2} \Sigma_{\gamma_{2l}} (\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ijl}^{-1} L_{ijl}(\text{res}))^2\right\}}.$$

where $\text{res} = Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}$.

The posterior distribution of γ_{2l} is a mixture of a normal distribution and a point mass at 0. At each iteration of MCMC, γ_{2l} is drawn from $N(\mu_{\gamma_{2l}}, \Sigma_{\gamma_{2l}})$ with probability $(1 - l_{1l})$ and is set to 0 with probability l_{1l} . If γ_{2l} is set to 0, then we have $\phi_{1l} = 0$. Otherwise, $\phi_{1l} = 1$.

- The posterior distribution of the interaction between the l th lipid and treatment factors $\boldsymbol{\gamma}_{3l} (l = 1, \dots, m)$ is also a spike-and-slab distribution. Denote $\mu_{ij(-\boldsymbol{\gamma}_{3l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{W}_{ijl}^\top \boldsymbol{\gamma}_{3l}$ and $l_{2l} = \pi(\boldsymbol{\gamma}_{3l} = \mathbf{0} | \text{rest})$, the conditional posterior distribution of $\boldsymbol{\gamma}_{3l}$ is a multivariate spike-and-slab distribution:

$$\boldsymbol{\gamma}_{3l} | \text{rest} \sim (1 - l_{2l}) \text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\gamma}_{3l}}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_{3l}}) + l_{2l} \delta_0(\boldsymbol{\gamma}_{3l}),$$

where

$$\begin{aligned}\boldsymbol{\mu}_{\boldsymbol{\gamma}_{3l}} &= \Sigma_{\boldsymbol{\gamma}_{3l}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{W}_{ijl} (Y_{ij} - \mu_{ij(-\boldsymbol{\gamma}_{3l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\boldsymbol{\gamma}_{3l}} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{W}_{ijl} \mathbf{W}_{ijl}^\top + \frac{1}{s_{2l}} \mathbf{I}_p \right)^{-1}.\end{aligned}$$

And

$$l_{2l} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)(s_{2l})^{-\frac{p}{2}} |\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_{3l}}|^{\frac{1}{2}} \exp\{\frac{1}{2} \|\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_{3l}}^{-\frac{1}{2}} \sum_{i=1}^n \sum_{j=1}^k \tau \xi_2^{-2} \tilde{v}_{ij}^{-1} \mathbf{W}_{ijl}(\text{res})\|_2^2\}}.$$

where $\text{res} = Y_{ij} - \mu_{ij(-\boldsymbol{\gamma}_{3l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}$.

The posterior distribution of $\boldsymbol{\gamma}_{3l}$ is a mixture of a multivariate normal distribution and a point mass at $\mathbf{0}$. At each iteration of MCMC, $\boldsymbol{\gamma}_{3l}$ is drawn from $\text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\gamma}_{3l}}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_{3l}})$ with probability $(1 - l_{2l})$ and is set to 0 with probability l_{2l} . If $\boldsymbol{\gamma}_{3l}$ is set to $\mathbf{0}$, then we have $\phi_{2l} = 0$. Otherwise, $\phi_{2l} = 1$.

- The full conditional posterior distribution of $(s_{1l})^{-1}$ is:

$$(s_{1l})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(1, \frac{\eta_1^2}{2}) & \text{if } \gamma_{2l} = 0 \\ \text{Inverse-Gaussian}(\eta_1^2, \sqrt{\frac{\eta_1^2}{\gamma_{2l}^2}}) & \text{if } \gamma_{2l} \neq 0 \end{cases}$$

- Similarly, The full conditional posterior distribution of $(s_{2l})^{-1}$ is:

$$(s_{2l})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{p+1}{2}, \frac{\eta_2^2}{2}) & \text{if } \boldsymbol{\gamma}_{3l} = \mathbf{0} \\ \text{Inverse-Gaussian}(\eta_2^2, \sqrt{\frac{\eta_2^2}{\|\boldsymbol{\gamma}_{3l}\|_2^2}}) & \text{if } \boldsymbol{\gamma}_{3l} \neq \mathbf{0} \end{cases}$$

- The full conditional posterior distribution of η_1^2 :

$$\eta_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m s_{1l}}{2}).$$

- Likewise, the full conditional posterior distribution of η_2^2 :

$$\eta_2^2 | \text{rest} \sim \text{Gamma}(a_2 + \frac{m(p+1)}{2}, b_2 + \frac{\sum_{l=1}^m s_{2l}}{2}).$$

- The full conditional posterior distribution of π_1 :

$$\pi_1 | \text{rest} \sim \text{Beta}(r_1 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l}=0\}}, w_1 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l} \neq 0\}}).$$

- Similarly, the full conditional posterior distribution of π_2 :

$$\pi_2 | \text{rest} \sim \text{Beta}(r_2 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{3l}=0\}}, w_2 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{3l} \neq 0\}}).$$

- The full conditional posterior distribution of $\tilde{v}_{ij}$:

$$\tilde{v}_{ij} | \text{rest} \sim \text{Inverse-Gaussian}(\sqrt{\frac{\xi_1^2 + 2\xi_2^2}{(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i)^2}}, \frac{\tau \xi_1^2}{\xi_2^2} + 2\tau).$$

- The full conditional posterior distribution of τ :

$$\tau | \text{rest} \sim \text{Gamma}(c + \frac{3nk}{2}, d + \sum_{i=1}^n \sum_{j=1}^k (\tilde{v}_{ij} + (2\xi_2^2)^{-1} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij})^2).$$

- The full conditional posterior distribution of $\boldsymbol{\alpha}_i$:

$$\boldsymbol{\alpha}_i | \text{rest} \sim \text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\alpha}_i}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i})$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\alpha}_i} &= \frac{\tau}{\xi_2^2} \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} \sum_{j=1}^k \frac{\mathbf{Z}_{ij} (Y_{ij} - \mu_{ij} - \xi_1 \tilde{v}_{ij})}{\tilde{v}_{ij}} \\ \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} &= \left(\frac{\tau}{\xi_2^2} \sum_{j=1}^k \frac{\mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top}{\tilde{v}_{ij}} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

- Lastly, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma}(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_i)$$

A.5.2 RBVSS

Priors

$$Y_{ij} = \mathbf{X}_{ij}^\top \gamma_0 + \mathbf{T}_{ij}^\top \gamma_1 + \mathbf{L}_{ij}^\top \gamma_2 + \mathbf{W}_{ij}^\top \gamma_3 + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i + \xi_1 \tilde{v}_{ij} + \tau^{-1/2} \xi_2 \sqrt{\tilde{v}_{ij}} z_{ij} \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$z_{ij} \stackrel{ind}{\sim} \text{N}(0, 1) \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\tilde{v}_{ij} | \tau \stackrel{ind}{\sim} \tau \exp(\tau \tilde{v}_{ij}) \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\gamma_0 \sim \text{N}_3(\mathbf{0}, \boldsymbol{\Sigma}_0)$$

$$\gamma_1 \sim \text{N}_p(\mathbf{0}, \boldsymbol{\Sigma}_1)$$

$$\gamma_{2l} | \phi_{1l}, s_{1l} \stackrel{ind}{\sim} \phi_{1l} \text{N}(0, s_{1l}) + (1 - \phi_{1l}) \delta_0(\gamma_{2l}) \quad l = 1, \dots, m$$

$$\phi_{1l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_1)$$

$$\gamma_{3q} | \phi_{2q}, s_{2q} \stackrel{ind}{\sim} \phi_{2q} \text{N}(0, s_{2q}) + (1 - \phi_{2q}) \delta_0(\gamma_{3q}) \quad q = 1, \dots, pm$$

$$\phi_{2l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_2)$$

$$s_{1l} | \eta_1^2 \stackrel{iid}{\sim} \frac{\eta_1^2}{2} \exp\left(-\frac{\eta_1^2}{2} s_{1l}\right) \quad l = 1, \dots, m$$

$$s_{2q} | \eta_2^2 \stackrel{iid}{\sim} \frac{\eta_2^2}{2} \exp\left(-\frac{\eta_2^2}{2} s_{2q}\right) \quad q = 1, \dots, pm$$

$$\tau \sim \text{Gamma}(c, d)$$

$$\eta_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\eta_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\boldsymbol{\alpha}_i \sim \text{N}(\mathbf{0}, \phi^2 \mathbf{I}) \quad i = 1, \dots, n$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\pi_1 \sim \text{Beta}(r_1, w_1)$$

$$\pi_2 \sim \text{Beta}(r_2, w_2)$$

Gibbs Sampler

- Let $\mu_{ij(-\gamma_0)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of $\boldsymbol{\gamma}_0$.

The full conditional distribution of $\boldsymbol{\gamma}_0$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\gamma}_0}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0})$ with mean

$$\boldsymbol{\mu}_{\boldsymbol{\gamma}_0} = \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{X}_{ij} (Y_{ij} - \mu_{ij(-\gamma_0)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0} = \left(\boldsymbol{\Sigma}_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{X}_{ij} \mathbf{X}_{ij}^\top \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_1)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1$, then the full conditional distribution of $\boldsymbol{\gamma}_1$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\gamma}_1}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_1})$ with mean

$$\boldsymbol{\mu}_{\boldsymbol{\gamma}_1} = \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_1} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{T}_{ij} (Y_{ij} - \mu_{ij(-\gamma_1)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\boldsymbol{\Sigma}_{\boldsymbol{\gamma}_1} = \left(\boldsymbol{\Sigma}_1^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{T}_{ij} \mathbf{T}_{ij}^\top \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{2l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - L_{ijl} \gamma_{2l}$ and $l_{1l} = \pi(\gamma_{2l} = 0 | \text{rest})$, the conditional posterior distribution of the coefficient of lipid factor γ_{2l} is a spike-and-slab distribution:

$$\gamma_{2l} | \text{rest} \sim (1 - l_{1l}) \text{N}(\mu_{\gamma_{2l}}, \Sigma_{\gamma_{2l}}) + l_{1l} \delta_0(\gamma_{2l}),$$

where

$$\begin{aligned}\mu_{\gamma_{2l}} &= \Sigma_{\gamma_{2l}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\gamma_{2l}} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl}^2 + \frac{1}{s_{1l}} \right)^{-1}.\end{aligned}$$

and

$$l_{1l} = \frac{\pi_1}{\pi_1 + (1 - \pi_1)(s_{1l})^{-\frac{1}{2}} (\Sigma_{\gamma_{2l}})^{\frac{1}{2}} \exp\left\{\frac{1}{2} \Sigma_{\gamma_{2l}} (\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl}(\text{res}))^2\right\}}.$$

$$\text{res} = Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}.$$

The posterior distribution of γ_{2l} is a mixture of a normal distribution and a point mass at 0. At each iteration of MCMC, γ_{2l} is drawn from $N(\mu_{\gamma_{2l}}, \Sigma_{\gamma_{2l}})$ with probability $(1 - l_{1l})$ and is set to 0 with probability l_{1l} . If γ_{2l} is set to 0, then we have $\phi_{1l} = 0$. Otherwise, $\phi_{1l} = 1$.

- The posterior distribution of the interaction between the q th lipid and treatment γ_{3q} ($q = 1, \dots, pm$) is also a spike-and-slab distribution. Denote $\mu_{ij(-\gamma_{3q})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - W_{ijq} \gamma_{3q}$ and $l_{2q} = \pi(\gamma_{3q} = 0 | \text{rest})$, the conditional posterior distribution of γ_{3q} is a spike-and-slab distribution:

$$\gamma_{3q} | \text{rest} \sim (1 - l_{2q}) N(\mu_{\gamma_{3q}}, \Sigma_{\gamma_{3q}}) + l_{2q} \delta_0(\gamma_{3q}),$$

where

$$\begin{aligned}\mu_{\gamma_{3q}} &= \Sigma_{\gamma_{3q}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijq} (Y_{ij} - \mu_{ij(-\gamma_{3q})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\gamma_{3q}} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k W_{ijq}^2 + \frac{1}{s_{2q}} \right)^{-1}.\end{aligned}$$

and

$$l_{2q} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)(s_{2q})^{-\frac{1}{2}}(\Sigma_{\gamma_{3q}})^{\frac{1}{2}} \exp\{\frac{1}{2}\Sigma_{\gamma_{3q}}(\sum_{i=1}^n \sum_{j=1}^k \tau \xi_2^{-2} \tilde{v}_{ij}^{-1} W_{ijq}(\text{res}))^2\}}.$$

$$\text{res} = Y_{ij} - \mu_{ij(-\gamma_{3q})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}.$$

The posterior distribution of γ_{3q} is a mixture of a normal distribution and a point mass at 0. At each iteration of MCMC, γ_{3q} is drawn from $N(\mu_{\gamma_{3q}}, \Sigma_{\gamma_{3q}})$ with probability $(1 - l_{2q})$ and is set to 0 with probability l_{2q} . If γ_{3q} is set to 0, then we have $\phi_{2q} = 0$. Otherwise, $\phi_{2q} = 1$.

- The full conditional posterior distribution of $(s_{1l})^{-1}$ is:

$$(s_{1l})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(1, \frac{\eta_1^2}{2}) & \text{if } \gamma_{2l} = 0 \\ \text{Inverse-Gaussian}(\eta_1^2, \sqrt{\frac{\eta_1^2}{\gamma_{2l}^2}}) & \text{if } \gamma_{2l} \neq 0 \end{cases}$$

- Similarly, The full conditional posterior distribution of $(s_{2q})^{-1}$ is:

$$(s_{2q})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{m+1}{2}, \frac{\eta_2^2}{2}) & \text{if } \gamma_{3q} = 0 \\ \text{Inverse-Gaussian}(\eta_2^2, \sqrt{\frac{\eta_2^2}{\gamma_{3q}^2}}) & \text{if } \gamma_{3q} \neq 0 \end{cases}$$

- The full conditional posterior distribution of η_1^2 :

$$\eta_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m s_{1l}}{2}).$$

- Likewise, the full conditional posterior distribution of η_2^2 :

$$\eta_2^2 | \text{rest} \sim \text{Gamma}(a_2 + pm, b_2 + \frac{\sum_{q=1}^{pm} s_{2q}}{2}).$$

- The full conditional posterior distribution of π_1 :

$$\pi_1|\text{rest} \sim \text{Beta}(r_1 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l}=0\}}, w_1 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l} \neq 0\}}).$$

- Similarly, the full conditional posterior distribution of π_2 :

$$\pi_2|\text{rest} \sim \text{Beta}(r_2 + \sum_{q=1}^{pm} \mathbf{I}_{\{\gamma_{3q}=0\}}, w_2 + \sum_{q=1}^{pm} \mathbf{I}_{\{\gamma_{3q} \neq 0\}}).$$

- The full conditional posterior distribution of $\tilde{v}_{ij}$:

$$\tilde{v}_{ij}|\text{rest} \sim \text{Inverse-Gaussian}(\sqrt{\frac{\xi_1^2 + 2\xi_2^2}{(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i)^2}}, \frac{\tau \xi_1^2}{\xi_2^2} + 2\tau).$$

- The full conditional posterior distribution of τ :

$$\tau|\text{rest} \sim \text{Gamma}(c + \frac{3nk}{2}, d + \sum_{i=1}^n \sum_{j=1}^k (\tilde{v}_{ij} + (2\xi_2^2)^{-1} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij})^2).$$

- The full conditional posterior distribution of $\boldsymbol{\alpha}_i$:

$$\boldsymbol{\alpha}_i|\text{rest} \sim \text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\alpha}_i}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i})$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\alpha}_i} &= \frac{\tau}{\xi_2^2} \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} \sum_{j=1}^k \frac{\mathbf{Z}_{ij}(Y_{ij} - \mu_{ij} - \xi_1 \tilde{v}_{ij})}{\tilde{v}_{ij}} \\ \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} &= \left(\frac{\tau}{\xi_2^2} \sum_{j=1}^k \frac{\mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top}{\tilde{v}_{ij}} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

- Lastly, the full conditional posterior distribution of ϕ^2 :

$$\phi^2|\text{rest} \sim \text{Inverse-Gamma}(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_i)$$

A.5.3 RBV-BI

Priors

$$Y_{ij} = \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0 + \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1 + \mathbf{L}_{ij}^\top \boldsymbol{\gamma}_2 + \mathbf{W}_{ij}^\top \boldsymbol{\gamma}_3 + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i + \xi_1 \tilde{v}_{ij} + \tau^{-1/2} \xi_2 \sqrt{\tilde{v}_{ij}} z_{ij} \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$z_{ij} \stackrel{ind}{\sim} \text{N}(0, 1) \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\tilde{v}_{ij} | \tau \stackrel{ind}{\sim} \tau \exp(\tau \tilde{v}_{ij}) \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\boldsymbol{\gamma}_0 \sim \text{N}_3(\mathbf{0}, \boldsymbol{\Sigma}_0)$$

$$\boldsymbol{\gamma}_1 \sim \text{N}_p(\mathbf{0}, \boldsymbol{\Sigma}_1)$$

$$\gamma_{2l} | s_{1l} \stackrel{ind}{\sim} \text{N}(0, s_{1l}) \quad l = 1, \dots, m$$

$$\boldsymbol{\gamma}_{3l} | s_{2l} \stackrel{ind}{\sim} \text{MVN}(\mathbf{0}, \text{diag}(s_{2l}, \dots, s_{2l})) \quad l = 1, \dots, m$$

$$s_{1l} | \eta_1^2 \stackrel{iid}{\sim} \frac{\eta_1^2}{2} \exp\left(-\frac{\eta_1^2}{2} s_{1l}\right) \quad l = 1, \dots, m$$

$$s_{2l} | \eta_2^2 \stackrel{ind}{\sim} \text{Gamma}\left(\frac{p+1}{2}, \frac{\eta_2^2}{2}\right) \quad l = 1, \dots, m$$

$$\tau \sim \text{Gamma}(c, d)$$

$$\eta_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\eta_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\boldsymbol{\alpha}_i \sim \text{N}(\mathbf{0}, \phi^2 \mathbf{I}) \quad i = 1, \dots, n$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

Gibbs Sampler

- Let $\mu_{ij(-\boldsymbol{\gamma}_0)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of $\boldsymbol{\gamma}_0$.

The full conditional distribution of $\boldsymbol{\gamma}_0$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\gamma}_0}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0})$

with mean

$$\boldsymbol{\mu}_{\gamma_0} = \boldsymbol{\Sigma}_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{X}_{ij} (Y_{ij} - \mu_{ij(-\gamma_0)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\boldsymbol{\Sigma}_{\gamma_0} = \left(\boldsymbol{\Sigma}_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{X}_{ij} \mathbf{X}_{ij}^\top \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_1)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1$, then the full conditional distribution of $\boldsymbol{\gamma}_1$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\gamma_1}, \boldsymbol{\Sigma}_{\gamma_1})$ with mean

$$\boldsymbol{\mu}_{\gamma_1} = \boldsymbol{\Sigma}_{\gamma_1} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{T}_{ij} (Y_{ij} - \mu_{ij(-\gamma_1)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\boldsymbol{\Sigma}_{\gamma_1} = \left(\boldsymbol{\Sigma}_1^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{T}_{ij} \mathbf{T}_{ij}^\top \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{2l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - L_{ijl} \gamma_{2l}$, the full conditional distribution of γ_{2l} is a normal distribution with mean

$$\mu_{\gamma_{2l}} = \boldsymbol{\Sigma}_{\gamma_{2l}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

variance

$$\boldsymbol{\Sigma}_{\gamma_{2l}} = \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl}^2 + \frac{1}{s_{1l}} \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_{3l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{W}_{ijl}^\top \boldsymbol{\gamma}_{3l}$, then the full conditional distribution of $\boldsymbol{\gamma}_{3l}$ is a multivariate normal distribution with mean

$$\boldsymbol{\mu}_{\gamma_{3l}} = \boldsymbol{\Sigma}_{\gamma_{3l}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{W}_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{3l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\gamma_{3l}} = \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{W}_{ijl} \mathbf{W}_{ijl}^\top + \frac{1}{s_{2l}} \mathbf{I}_m \right)^{-1}.$$

- The full conditional posterior distribution of $(s_{1l})^{-1}$ is:

$$(s_{1l})^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\eta_1^2, \sqrt{\frac{\eta_1^2}{\gamma_{2l}^2}})$$

- Similarly, The full conditional posterior distribution of $(s_{2l})^{-1}$ is:

$$(s_{2l})^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\eta_2^2, \sqrt{\frac{\eta_2^2}{\|\gamma_{3l}\|_2^2}})$$

- The full conditional posterior distribution of η_1^2 :

$$\eta_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m s_{1l}}{2}).$$

- Likewise, the full conditional posterior distribution of η_2^2 :

$$\eta_2^2 | \text{rest} \sim \text{Gamma}(a_2 + \frac{m(p+1)}{2}, b_2 + \frac{\sum_{l=1}^m s_{2l}}{2}).$$

- The full conditional posterior distribution of $\tilde{v}_{ij}$:

$$\tilde{v}_{ij} | \text{rest} \sim \text{Inverse-Gaussian}(\sqrt{\frac{\xi_1^2 + 2\xi_2^2}{(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i)^2}}, \frac{\tau \xi_1^2}{\xi_2^2} + 2\tau).$$

- The full conditional posterior distribution of τ :

$$\tau | \text{rest} \sim \text{Gamma}(c + \frac{3nk}{2}, d + \sum_{i=1}^n \sum_{j=1}^k (\tilde{v}_{ij} + (2\xi_2^2)^{-1} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij})^2).$$

- The full conditional posterior distribution of α_i :

$$\alpha_i | \text{rest} \sim \text{MVN}(\mu_{\alpha_i}, \Sigma_{\alpha_i})$$

where

$$\begin{aligned} \mu_{\alpha_i} &= \frac{\tau}{\xi_2^2} \Sigma_{\alpha_i} \sum_{j=1}^k \frac{\mathbf{Z}_{ij}(Y_{ij} - \mu_{ij} - \xi_1 \tilde{v}_{ij})}{\tilde{v}_{ij}} \\ \Sigma_{\alpha_i} &= \left(\frac{\tau}{\xi_2^2} \sum_{j=1}^k \frac{\mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top}{\tilde{v}_{ij}} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

- Lastly, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma}\left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \alpha_i^\top \alpha_i\right)$$

A.5.4 RBV

Priors

$$Y_{ij} = \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0 + \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1 + \mathbf{L}_{ij}^\top \boldsymbol{\gamma}_2 + \mathbf{W}_{ij}^\top \boldsymbol{\gamma}_3 + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i + \xi_1 \tilde{v}_{ij} + \tau^{-1/2} \xi_2 \sqrt{\tilde{v}_{ij}} z_{ij} \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$z_{ij} \stackrel{ind}{\sim} \text{N}(0, 1) \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\tilde{v}_{ij} | \tau \stackrel{ind}{\sim} \tau \exp(\tau \tilde{v}_{ij}) \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\boldsymbol{\gamma}_0 \sim \text{N}_3(\mathbf{0}, \boldsymbol{\Sigma}_0)$$

$$\boldsymbol{\gamma}_1 \sim \text{N}_p(\mathbf{0}, \boldsymbol{\Sigma}_1)$$

$$\gamma_{2l} | s_{1l} \stackrel{ind}{\sim} \text{N}(0, s_{1l}) \quad l = 1, \dots, m$$

$$\gamma_{3q} | s_{2q} \stackrel{ind}{\sim} \text{N}(0, s_{2q}) \quad q = 1, \dots, pm$$

$$s_{1l} | \eta_1^2 \stackrel{iid}{\sim} \frac{\eta_1^2}{2} \exp\left(-\frac{\eta_1^2}{2} s_{1l}\right) \quad l = 1, \dots, m$$

$$s_{2q} | \eta_2^2 \stackrel{iid}{\sim} \frac{\eta_2^2}{2} \exp\left(-\frac{\eta_2^2}{2} s_{2q}\right) \quad q = 1, \dots, pm$$

$$\tau \sim \text{Gamma}(c, d)$$

$$\eta_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\eta_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\boldsymbol{\alpha}_i \sim \text{N}(\mathbf{0}, \phi^2 \mathbf{I}) \quad i = 1, \dots, n$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

Gibbs Sampler

- Let $\mu_{ij(-\boldsymbol{\gamma}_0)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of $\boldsymbol{\gamma}_0$.

The full conditional distribution of $\boldsymbol{\gamma}_0$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\gamma}_0}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0})$ with

mean

$$\boldsymbol{\mu}_{\gamma_0} = \boldsymbol{\Sigma}_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{X}_{ij} (Y_{ij} - \mu_{ij(-\gamma_0)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\boldsymbol{\Sigma}_{\gamma_0} = \left(\boldsymbol{\Sigma}_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{X}_{ij} \mathbf{X}_{ij}^\top \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_1)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1$, then the full conditional distribution of $\boldsymbol{\gamma}_1$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\gamma_1}, \boldsymbol{\Sigma}_{\gamma_1})$ with mean

$$\boldsymbol{\mu}_{\gamma_1} = \boldsymbol{\Sigma}_{\gamma_1} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{T}_{ij} (Y_{ij} - \mu_{ij(-\gamma_1)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\boldsymbol{\Sigma}_{\gamma_1} = \left(\boldsymbol{\Sigma}_1^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \mathbf{T}_{ij} \mathbf{T}_{ij}^\top \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{2l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - L_{ijl} \gamma_{2l}$, the full conditional distribution of γ_{2l} is a normal distribution with mean

$$\mu_{\gamma_{2l}} = \boldsymbol{\Sigma}_{\gamma_{2l}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

variance

$$\boldsymbol{\Sigma}_{\gamma_{2l}} = \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl}^2 + \frac{1}{s_{1l}} \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{3q})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - W_{ijq} \gamma_{3q}$, the full conditional distribution of γ_{3q} is a normal distribution with mean

$$\mu_{\gamma_{3q}} = \boldsymbol{\Sigma}_{\gamma_{3q}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijq} (Y_{ij} - \mu_{ij(-\gamma_{3q})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\gamma_{3q}} = \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijq}^2 + \frac{1}{s_{2q}} \right)^{-1}.$$

- The full conditional posterior distribution of $(s_{1l})^{-1}$ is:

$$(s_{1l})^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\eta_1^2, \sqrt{\frac{\eta_1^2}{\gamma_{2l}^2}})$$

- Similarly, The full conditional posterior distribution of $(s_{2q})^{-1}$ is:

$$(s_{2q})^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\eta_2^2, \sqrt{\frac{\eta_2^2}{\gamma_{3q}^2}})$$

- The full conditional posterior distribution of η_1^2 :

$$\eta_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m s_{1l}}{2}).$$

- Likewise, the full conditional posterior distribution of η_2^2 :

$$\eta_2^2 | \text{rest} \sim \text{Gamma}(a_2 + \frac{m(p+1)}{2}, b_2 + \frac{\sum_{q=1}^{pm} s_{2q}}{2}).$$

- The full conditional posterior distribution of $\tilde{v}_{ij}$:

$$\tilde{v}_{ij} | \text{rest} \sim \text{Inverse-Gaussian}(\sqrt{\frac{\xi_1^2 + 2\xi_2^2}{(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i)^2}}, \frac{\tau \xi_1^2}{\xi_2^2} + 2\tau).$$

- The full conditional posterior distribution of τ :

$$\tau | \text{rest} \sim \text{Gamma}(c + \frac{3nk}{2}, d + \sum_{i=1}^n \sum_{j=1}^k (\tilde{v}_{ij} + (2\xi_2^2)^{-1} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \xi_1 \tilde{v}_{ij})^2).$$

- The full conditional posterior distribution of $\boldsymbol{\alpha}_i$:

$$\boldsymbol{\alpha}_i | \text{rest} \sim \text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\alpha}_i}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i})$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\alpha}_i} &= \frac{\tau}{\xi_2^2} \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} \sum_{j=1}^k \frac{\mathbf{Z}_{ij}(Y_{ij} - \mu_{ij} - \xi_1 \tilde{v}_{ij})}{\tilde{v}_{ij}} \\ \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} &= \left(\frac{\tau}{\xi_2^2} \sum_{j=1}^k \frac{\mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top}{\tilde{v}_{ij}} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

- Lastly, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma}\left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_i\right)$$

A.5.5 BVSS-BI

Priors

$$\mathbf{Y}|\gamma_0, \gamma_1, \gamma_2, \gamma_3, \boldsymbol{\alpha}_i, \sigma^2 \propto (\sigma^2)^{-\frac{nk}{2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^k \left(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i\right)^2\right\} \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\boldsymbol{\gamma}_0 \sim \text{N}_3(\mathbf{0}, \boldsymbol{\Sigma}_0)$$

$$\boldsymbol{\gamma}_1 \sim \text{N}_p(\mathbf{0}, \boldsymbol{\Sigma}_1)$$

$$\gamma_{2l}|\phi_{1l}, \tau_{1l}^2, \sigma^2 \stackrel{\text{ind}}{\sim} \phi_{1l} \text{N}(0, \sigma^2 \tau_{1l}^2) + (1 - \phi_{1l}) \delta_0(\gamma_{2l}) \quad l = 1, \dots, m$$

$$\phi_{1l} \stackrel{\text{ind}}{\sim} \text{Bernoulli}(\pi_1)$$

$$\boldsymbol{\gamma}_{3l}|\phi_{2l}, \tau_{2l}^2, \sigma^2 \stackrel{\text{ind}}{\sim} \phi_{2l} \text{MVN}(\mathbf{0}, \text{diag}(\sigma^2 \tau_{2l}^2, \dots, \sigma^2 \tau_{2l}^2)) + (1 - \phi_{2l}) \delta_0(\boldsymbol{\gamma}_{3l}) \quad l = 1, \dots, m$$

$$\phi_{2l} \stackrel{\text{ind}}{\sim} \text{Bernoulli}(\pi_2)$$

$$\tau_{1l}^2|\lambda_1^2 \stackrel{\text{iid}}{\sim} \frac{\lambda_1^2}{2} \exp\left(-\frac{\lambda_1^2}{2} \tau_{1l}^2\right) \quad l = 1, \dots, m$$

$$\tau_{2l}^2|\lambda_2^2 \stackrel{\text{ind}}{\sim} \text{Gamma}\left(\frac{p+1}{2}, \frac{p\lambda_2^2}{2}\right) \quad l = 1, \dots, m$$

$$\lambda_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\lambda_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\boldsymbol{\alpha}_i \sim \text{N}(\mathbf{0}, \phi^2 \mathbf{I}) \quad i = 1, \dots, n$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\pi_1 \sim \text{Beta}(r_1, w_1)$$

$$\pi_2 \sim \text{Beta}(r_2, w_2)$$

$$\sigma^2 \sim \frac{1}{\sigma^2}$$

Gibbs Sampler

- Let $\mu_{ij}(-\gamma_0) = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of $\boldsymbol{\gamma}_0$.

The full conditional distribution of $\boldsymbol{\gamma}_0$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\gamma}_0}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0})$ with

mean

$$\boldsymbol{\mu}_{\gamma_0} = \boldsymbol{\Sigma}_{\gamma_0} \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{X}_{ij} (Y_{ij} - \mu_{ij(-\gamma_0)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\boldsymbol{\Sigma}_{\gamma_0} = \left(\boldsymbol{\Sigma}_0^{-1} + \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{X}_{ij} \mathbf{X}_{ij}^\top \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_1)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1$, then the full conditional distribution of $\boldsymbol{\gamma}_1$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\gamma_1}, \boldsymbol{\Sigma}_{\gamma_1})$ with mean

$$\boldsymbol{\mu}_{\gamma_1} = \boldsymbol{\Sigma}_{\gamma_1} \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{T}_{ij} (Y_{ij} - \mu_{ij(-\gamma_1)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\boldsymbol{\Sigma}_{\gamma_1} = \left(\boldsymbol{\Sigma}_1^{-1} + \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{T}_{ij} \mathbf{T}_{ij}^\top \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{2l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - L_{ijl} \gamma_{2l}$ and $l_{1l} = \pi(\gamma_{2l} = 0 | \text{rest})$, the conditional posterior distribution of the coefficient of lipid factor γ_{2l} is a spike-and-slab distribution:

$$\gamma_{2l} | \text{rest} \sim (1 - l_{1l}) \text{N}(\mu_{\gamma_{2l}}, \Sigma_{\gamma_{2l}}) + l_{1l} \delta_0(\gamma_{2l}),$$

where

$$\begin{aligned} \mu_{\gamma_{2l}} &= \Sigma_{\gamma_{2l}} (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k L_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i), \\ \Sigma_{\gamma_{2l}} &= \sigma^2 \left(\sum_{i=1}^n \sum_{j=1}^k L_{ijl}^2 + \frac{1}{\tau_{1l}^2} \right)^{-1}. \end{aligned}$$

Where

$$l_{1l} = \frac{\pi_1}{\pi_1 + (1 - \pi_1)(\tau_{1l}^2)^{-\frac{1}{2}}(\Sigma_{\gamma_{2l}})^{\frac{1}{2}} \exp\{\frac{1}{2\sigma^2}\Sigma_{\gamma_{2l}}(\sum_{i=1}^n \sum_{j=1}^k L_{ijl}(Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i))^2\}}.$$

The posterior distribution of γ_{2l} is a mixture of a normal distribution and a point mass at 0. At each iteration of MCMC, γ_{2l} is drawn from $N(\mu_{\gamma_{2l}}, \Sigma_{\gamma_{2l}})$ with probability $(1 - l_{1l})$ and is set to 0 with probability l_{1l} . If γ_{2l} is set to 0, then we have $\phi_{1l} = 0$. Otherwise, $\phi_{1l} = 1$.

- The posterior distribution of the interaction between the l th lipid and treatment γ_{3l} ($l = 1, \dots, m$) is also a spike-and-slab distribution. Denote $\mu_{ij(-\gamma_{3l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{W}_{ijl}^\top \gamma_{3l}$ and $l_{2l} = \pi(\gamma_{3l} = \mathbf{0} | \text{rest})$, the conditional posterior distribution of γ_{3l} is a multivariate spike-and-slab distribution :

$$\gamma_{3l} | \text{rest} \sim (1 - l_{2l}) \text{MVN}(\boldsymbol{\mu}_{\gamma_{3l}}, \sigma^2 \Sigma_{\gamma_{3l}}) + l_{2l} \delta_0(\gamma_{3l}),$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\gamma_{3l}} &= \Sigma_{\gamma_{3l}} \sum_{i=1}^n \sum_{j=1}^k \mathbf{W}_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{3l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i), \\ \Sigma_{\gamma_{3l}} &= \left(\sum_{i=1}^n \sum_{j=1}^k \mathbf{W}_{ijl} \mathbf{W}_{ijl}^\top + \frac{1}{\tau_{2l}^2} \mathbf{I}_m \right)^{-1}. \end{aligned}$$

And

$$l_{2l} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)(\tau_{2l}^2)^{-\frac{m}{2}} |\Sigma_{\gamma_{3l}}|^{\frac{1}{2}} \exp\{\frac{1}{2\sigma^2} \|\Sigma_{\gamma_{3l}}^{-\frac{1}{2}} \sum_{i=1}^n \sum_{j=1}^k \mathbf{W}_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{3l})} - \boldsymbol{\alpha}_i)\|_2^2\}}. \quad (\text{A.1})$$

The posterior distribution of γ_{3l} is a mixture of a multivariate normal distribution and a point mass at $\mathbf{0}$. At each iteration of MCMC, γ_{3l} is drawn from $\text{MVN}(\boldsymbol{\mu}_{\gamma_{3l}}, \sigma^2 \Sigma_{\gamma_{3l}})$ with probability $(1 - l_{2l})$ and is set to 0 with probability l_{2l} . If γ_{3l} is set to $\mathbf{0}$, then we have $\phi_{2l} = 0$. Otherwise, $\phi_{2l} = 1$.

- The full conditional posterior distribution of $(\tau_{1l}^2)^{-1}$ is:

$$(\tau_{1l}^2)^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(1, \frac{\lambda_1^2}{2}) & \text{if } \gamma_{2l} = 0 \\ \text{Inverse-Gaussian}(\lambda_1^2, \sqrt{\frac{\sigma^2 \lambda_1^2}{\gamma_{2l}^2}}) & \text{if } \gamma_{2l} \neq 0 \end{cases}$$

- Similarly, The full conditional posterior distribution of $(\tau_{2l}^2)^{-1}$ is:

$$(\tau_{2l}^2)^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{m+1}{2}, \frac{m\lambda_2^2}{2}) & \text{if } \gamma_{3l} = 0 \\ \text{Inverse-Gaussian}(\lambda_2^2, \sqrt{\frac{\sigma^2 \lambda_2^2}{\|\gamma_{3l}\|_2^2}}) & \text{if } \gamma_{3l} \neq 0 \end{cases}$$

- The full conditional posterior distribution of λ_1^2 :

$$\lambda_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m \tau_{1l}^2}{2}).$$

- Likewise, the full conditional posterior distribution of λ_2^2 :

$$\lambda_2^2 | \text{rest} \sim \text{Gamma}(a_2 + \frac{m(p+1)}{2}, b_2 + \frac{m \sum_{l=1}^m \tau_{2l}^2}{2}).$$

- The full conditional posterior distribution of π_1 :

$$\pi_1 | \text{rest} \sim \text{Beta}(r_1 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l}=0\}}, w_1 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l} \neq 0\}}).$$

- Similarly, the full conditional posterior distribution of π_2 :

$$\pi_2 | \text{rest} \sim \text{Beta}(r_2 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{3l}=0\}}, w_2 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{3l} \neq 0\}}).$$

- The full conditional posterior distribution of σ^2 :

$$\sigma^2 | \text{rest} \sim \text{Inv-Gamma}(\mu_{\sigma^2}, \Sigma_{\sigma^2})$$

with shape

$$\mu_{\sigma^2} = \frac{nk + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l} \neq 0\}} + m \sum_{l=1}^m \mathbf{I}_{\{\gamma_{3l} \neq 0\}}}{2}$$

and scale

$$\Sigma_{\sigma^2} = \frac{\sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i)^2 + \sum_{l=1}^m (\tau_{1l}^2)^{-1} \gamma_{2l}^2 + \sum_{l=1}^m (\tau_{2l}^2)^{-1} \gamma_{3l}^\top \gamma_{3l}}{2}$$

- The full conditional posterior distribution of $\boldsymbol{\alpha}_i$:

$$\boldsymbol{\alpha}_i | \text{rest} \sim \text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\alpha}_i}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i})$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\alpha}_i} &= (\sigma^2)^{-1} \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} \sum_{j=1}^k \mathbf{Z}_{ij} (Y_{ij} - \mu_{ij}) \\ \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} &= \left(\sum_{j=1}^k \frac{\mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top}{\sigma^2} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

- Lastly, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma}\left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_i\right)$$

A.5.6 BVSS

Priors

$$\mathbf{Y}|\gamma_0, \gamma_1, \gamma_2, \gamma_3, \boldsymbol{\alpha}_i, \sigma^2 \propto (\sigma^2)^{-\frac{nk}{2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^k \left(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i\right)^2\right\} \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\gamma_0 \sim \text{N}_3(\mathbf{0}, \boldsymbol{\Sigma}_0)$$

$$\gamma_1 \sim \text{N}_p(\mathbf{0}, \boldsymbol{\Sigma}_1)$$

$$\gamma_{2l}|\phi_{1l}, \tau_{1l}^2, \sigma^2 \stackrel{ind}{\sim} \phi_{1l} \text{N}(0, \sigma^2 \tau_{1l}^2) + (1 - \phi_{1l})\delta_0(\gamma_{2l}) \quad l = 1, \dots, m$$

$$\phi_{1l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_1)$$

$$\gamma_{3q}|\phi_{2q}, \tau_{2q}^2, \sigma^2 \stackrel{ind}{\sim} \phi_{2q} \text{N}(0, \sigma^2 \tau_{2q}^2) + (1 - \phi_{2q})\delta_0(\gamma_{3q}) \quad q = 1, \dots, pm$$

$$\phi_{2q} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_2)$$

$$\tau_{1l}^2|\lambda_1^2 \stackrel{iid}{\sim} \frac{\lambda_1^2}{2} \exp\left(-\frac{\lambda_1^2}{2} \tau_{1l}^2\right) \quad l = 1, \dots, m$$

$$\tau_{2q}^2|\lambda_2^2 \stackrel{iid}{\sim} \frac{\lambda_2^2}{2} \exp\left(-\frac{\lambda_2^2}{2} \tau_{2q}^2\right) \quad q = 1, \dots, pm$$

$$\lambda_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\lambda_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\boldsymbol{\alpha}_i \sim \text{N}(\mathbf{0}, \phi^2 \mathbf{I}) \quad i = 1, \dots, n$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\pi_1 \sim \text{Beta}(r_1, w_1)$$

$$\pi_2 \sim \text{Beta}(r_2, w_2)$$

$$\sigma^2 \sim \frac{1}{\sigma^2}$$

Gibbs Sampler

- Let $\mu_{ij(-\gamma_0)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{X}_{ij}^\top \gamma_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of γ_0 .

The full conditional distribution of γ_0 is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\gamma_0}, \boldsymbol{\Sigma}_{\gamma_0})$

with mean

$$\boldsymbol{\mu}_{\gamma_0} = \boldsymbol{\Sigma}_{\gamma_0} \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k j = 1^k \mathbf{X}_{ij} (Y_{ij} - \mu_{ij(-\gamma_0)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\boldsymbol{\Sigma}_{\gamma_0} = \left(\boldsymbol{\Sigma}_0^{-1} + \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{X}_{ij} \mathbf{X}_{ij}^\top \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_1)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1$, then the full conditional distribution of $\boldsymbol{\gamma}_1$ is multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\gamma_1}, \boldsymbol{\Sigma}_{\gamma_1})$ with mean

$$\boldsymbol{\mu}_{\gamma_1} = \boldsymbol{\Sigma}_{\gamma_1} \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{T}_{ij} (Y_{ij} - \mu_{ij(-\gamma_1)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\boldsymbol{\Sigma}_{\gamma_1} = \left(\boldsymbol{\Sigma}_1^{-1} + \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{T}_{ij} \mathbf{T}_{ij}^\top \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{2l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - L_{ijl} \gamma_{2l}$ and $l_{1l} = \pi(\gamma_{2l} = 0 | \text{rest})$, the conditional posterior distribution of the coefficient of lipid factor γ_{2l} is a spike-and-slab distribution:

$$\gamma_{2l} | \text{rest} \sim (1 - l_{1l}) \text{N}(\mu_{\gamma_{2l}}, \boldsymbol{\Sigma}_{\gamma_{2l}}) + l_{1l} \delta_0(\gamma_{2l}),$$

where

$$\begin{aligned} \mu_{\gamma_{2l}} &= \boldsymbol{\Sigma}_{\gamma_{2l}} (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k L_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i), \\ \boldsymbol{\Sigma}_{\gamma_{2l}} &= \sigma^2 \left(\sum_{i=1}^n \sum_{j=1}^k L_{ijl}^2 + \frac{1}{\tau_{1l}^2} \right)^{-1}. \end{aligned}$$

Where

$$l_{1l} = \frac{\pi_1}{\pi_1 + (1 - \pi_1)(\tau_{1l}^2)^{-\frac{1}{2}}(\Sigma_{\gamma_{2l}})^{\frac{1}{2}} \exp\{\frac{1}{2\sigma^2} \Sigma_{\gamma_{2l}} (\sum_{i=1}^n \sum_{j=1}^k L_{ijl}(Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i))^2\}}.$$

The posterior distribution of γ_{2l} is a mixture of a normal distribution and a point mass at 0. At each iteration of MCMC, γ_{2l} is drawn from $N(\mu_{\gamma_{2l}}, \Sigma_{\gamma_{2l}})$ with probability $(1 - l_{1l})$ and is set to 0 with probability l_{1l} . If γ_{2l} is set to 0, then we have $\phi_{1l} = 0$. Otherwise, $\phi_{1l} = 1$.

- Let $\mu_{ij(-\gamma_{3q})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - W_{ijq} \gamma_{3q}$ and $l_{2q} = \pi(\gamma_{3q} = 0 | \text{rest})$, the conditional posterior distribution of the coefficient of interaction factor γ_{3q} is a spike-and-slab distribution:

$$\gamma_{3q} | \text{rest} \sim (1 - l_{2q})N(\mu_{\gamma_{3q}}, \Sigma_{\gamma_{3q}}) + l_{2q}\delta_0(\gamma_{3q}),$$

where

$$\begin{aligned} \mu_{\gamma_{3q}} &= \Sigma_{\gamma_{3q}}(\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k W_{ijq}(Y_{ij} - \mu_{ij(-\gamma_{3q})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i), \\ \Sigma_{\gamma_{3q}} &= \sigma^2 \left(\sum_{i=1}^n \sum_{j=1}^k W_{ijq}^2 + \frac{1}{\tau_{2q}^2} \right)^{-1}. \end{aligned}$$

Where

$$l_{2q} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)(\tau_{2q}^2)^{-\frac{1}{2}}(\Sigma_{\gamma_{3q}})^{\frac{1}{2}} \exp\{\frac{1}{2\sigma^2} \Sigma_{\gamma_{3q}} (\sum_{i=1}^n \sum_{j=1}^k W_{ijq}(Y_{ij} - \mu_{ij(-\gamma_{3q})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i))^2\}}.$$

- The full conditional posterior distribution of $(\tau_{1l}^2)^{-1}$ is:

$$(\tau_{1l}^2)^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(1, \frac{\lambda_1^2}{2}) & \text{if } \gamma_{2l} = 0 \\ \text{Inverse-Gaussian}(\lambda_1^2, \sqrt{\frac{\sigma^2 \lambda_1^2}{\gamma_{2l}^2}}) & \text{if } \gamma_{2l} \neq 0 \end{cases}$$

- Similarly, The full conditional posterior distribution of $(\tau_{2q}^2)^{-1}$ is:

$$(\tau_{2q}^2)^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(1, \frac{\lambda_2^2}{2}) & \text{if } \gamma_{3q} = 0 \\ \text{Inverse-Gaussian}(\lambda_2^2, \sqrt{\frac{\sigma^2 \lambda_2^2}{\gamma_{3q}^2}}) & \text{if } \gamma_{3q} \neq 0 \end{cases}$$

- The full conditional posterior distribution of λ_1^2 :

$$\lambda_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m \tau_{1l}^2}{2}).$$

- Likewise, the full conditional posterior distribution of λ_2^2 :

$$\lambda_2^2 | \text{rest} \sim \text{Gamma}(a_2 + pm, b_2 + \frac{\sum_{q=1}^{pm} \tau_{2q}^2}{2}).$$

- The full conditional posterior distribution of π_1 :

$$\pi_1 | \text{rest} \sim \text{Beta}(r_1 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l}=0\}}, w_1 + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l} \neq 0\}}).$$

- Similarly, the full conditional posterior distribution of π_2 :

$$\pi_2 | \text{rest} \sim \text{Beta}(r_2 + \sum_{q=1}^{pm} \mathbf{I}_{\{\gamma_{3q}=0\}}, w_2 + \sum_{q=1}^{pm} \mathbf{I}_{\{\gamma_{3q} \neq 0\}}).$$

- The full conditional posterior distribution of σ^2 :

$$\sigma^2 | \text{rest} \sim \text{Inv-Gamma}(\mu_{\sigma^2}, \Sigma_{\sigma^2})$$

with shape

$$\mu_{\sigma^2} = \frac{nk + \sum_{l=1}^m \mathbf{I}_{\{\gamma_{2l} \neq 0\}} + \sum_{q=1}^{pm} \mathbf{I}_{\{\gamma_{3q} \neq 0\}}}{2}$$

and scale

$$\Sigma_{\sigma^2} = \frac{\sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i)^2 + \sum_{l=1}^m (\tau_{1l}^2)^{-1} \gamma_{2l}^2 + \sum_{q=1}^{pm} (\tau_{2q}^2)^{-1} \gamma_{3q}^2}{2}$$

- The full conditional posterior distribution of $\boldsymbol{\alpha}_i$:

$$\boldsymbol{\alpha}_i | \text{rest} \sim \text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\alpha}_i}, \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i})$$

where

$$\begin{aligned} \boldsymbol{\mu}_{\boldsymbol{\alpha}_i} &= (\sigma^2)^{-1} \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} \sum_{j=1}^k \mathbf{Z}_{ij} (Y_{ij} - \mu_{ij}) \\ \boldsymbol{\Sigma}_{\boldsymbol{\alpha}_i} &= \left(\sum_{j=1}^k \frac{\mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top}{\sigma^2} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

- Lastly, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma} \left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \boldsymbol{\alpha}_i^\top \boldsymbol{\alpha}_i \right)$$

A.5.7 BV-BI

Priors

$$\mathbf{Y}|\gamma_0, \gamma_1, \gamma_2, \gamma_3, \boldsymbol{\alpha}_i, \sigma^2 \propto (\sigma^2)^{-\frac{nk}{2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^k \left(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i\right)^2\right\} \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\gamma_0 \sim \text{N}_3(\mathbf{0}, \boldsymbol{\Sigma}_0)$$

$$\gamma_1 \sim \text{N}_p(\mathbf{0}, \boldsymbol{\Sigma}_1)$$

$$\gamma_{2l}|\tau_{1l}^2, \sigma^2 \stackrel{\text{ind}}{\sim} \text{N}(0, \sigma^2 \tau_{1l}^2) \quad l = 1, \dots, m$$

$$\gamma_{3l}|\tau_{2l}^2, \sigma^2 \stackrel{\text{ind}}{\sim} \text{MVN}(\mathbf{0}, \text{diag}(\sigma^2 \tau_{2l}^2, \dots, \sigma^2 \tau_{2l}^2)) \quad l = 1, \dots, m$$

$$\tau_{1l}^2|\lambda_1^2 \stackrel{\text{iid}}{\sim} \frac{\lambda_1^2}{2} \exp\left(-\frac{\lambda_1^2}{2} \tau_{1l}^2\right) \quad l = 1, \dots, m$$

$$\tau_{2l}^2|\lambda_2^2 \stackrel{\text{ind}}{\sim} \text{Gamma}\left(\frac{p+1}{2}, \frac{p\lambda_2^2}{2}\right) \quad l = 1, \dots, m$$

$$\lambda_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\lambda_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\boldsymbol{\alpha}_i \sim \text{N}(\mathbf{0}, \phi^2 \mathbf{I}) \quad i = 1, \dots, n$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\sigma^2 \sim \frac{1}{\sigma^2}$$

Gibbs Sampler

- Let $\mu_{ij(-\gamma_0)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{X}_{ij}^\top \gamma_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of γ_0 .

The full conditional distribution of γ_0 is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\gamma_0}, \boldsymbol{\Sigma}_{\gamma_0})$ with mean

$$\boldsymbol{\mu}_{\gamma_0} = \boldsymbol{\Sigma}_{\gamma_0} \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{X}_{ij} (Y_{ij} - \mu_{ij(-\gamma_0)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\Sigma_{\gamma_0} = \left(\Sigma_0^{-1} + \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{1}^k \mathbf{X}_{ij} \mathbf{X}_{ij}^\top \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_1)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{T}_{ij}^\top \boldsymbol{\gamma}_1$, then the full conditional distribution of $\boldsymbol{\gamma}_1$ is multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\gamma_1}, \Sigma_{\gamma_1})$ with mean

$$\boldsymbol{\mu}_{\gamma_1} = \Sigma_{\gamma_1} \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{T}_{ij} (Y_{ij} - \mu_{ij(-\gamma_1)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\Sigma_{\gamma_1} = \left(\Sigma_1^{-1} + \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{T}_{ij} \mathbf{T}_{ij}^\top \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{2l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - L_{ijl} \gamma_{2l}$, the conditional posterior distribution of the coefficient of lipid factor γ_{2l} is a normal distribution:

$$\mu_{\gamma_{2l}} = \Sigma_{\gamma_{2l}} (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k L_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\Sigma_{\gamma_{2l}} = \sigma^2 \left(\sum_{i=1}^n \sum_{j=1}^k L_{ijl}^2 + \frac{1}{\tau_{1l}^2} \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_{3l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{W}_{ijl}^\top \boldsymbol{\gamma}_{3l}$. The posterior distribution of the interaction between the l th lipid and treatment factors $\boldsymbol{\gamma}_{3l} (l = 1, \dots, m)$ is a multivariate normal distribution:

$$\boldsymbol{\mu}_{\gamma_{3l}} = \Sigma_{\gamma_{3l}} \sum_{i=1}^n \sum_{j=1}^k \mathbf{W}_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{3l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

And variance

$$\Sigma_{\gamma_{3l}} = \left(\sum_{i=1}^n \sum_{j=1}^k \mathbf{W}_{ijl} \mathbf{W}_{ijl}^\top + \frac{1}{\tau_{2l}^2} \mathbf{I}_p \right)^{-1}.$$

- The full conditional posterior distribution of $(\tau_{1l}^2)^{-1}$ is:

$$(\tau_{1l}^2)^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\lambda_1^2, \sqrt{\frac{\sigma^2 \lambda_1^2}{\gamma_{2l}^2}})$$

- Similarly, The full conditional posterior distribution of $(\tau_{2l}^2)^{-1}$ is:

$$(\tau_{2l}^2)^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\lambda_2^2, \sqrt{\frac{\sigma^2 \lambda_2^2}{\|\gamma_{3l}\|_2^2}})$$

- The full conditional posterior distribution of λ_1^2 :

$$\lambda_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m \tau_{1l}^2}{2}).$$

- Likewise, the full conditional posterior distribution of λ_2^2 :

$$\lambda_2^2 | \text{rest} \sim \text{Gamma}(a_2 + \frac{m(p+1)}{2}, b_2 + \frac{p \sum_{l=1}^m \tau_{2l}^2}{2}).$$

- The full conditional posterior distribution of σ^2 :

$$\sigma^2 | \text{rest} \sim \text{Inv-Gamma}(\mu_{\sigma^2}, \Sigma_{\sigma^2})$$

with shape

$$\mu_{\sigma^2} = \frac{nk + m + pm}{2}$$

and scale

$$\Sigma_{\sigma^2} = \frac{\sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i)^2 + \sum_{l=1}^m (\tau_{1l}^2)^{-1} \gamma_{2l}^2 + \sum_{l=1}^m (\tau_{2l}^2)^{-1} \gamma_{3l}^\top \gamma_{3l}}{2}$$

- The full conditional posterior distribution of α_i :

$$\alpha_i | \text{rest} \sim \text{MVN}(\mu_{\alpha_i}, \Sigma_{\alpha_i})$$

where

$$\begin{aligned} \mu_{\alpha_i} &= (\sigma^2)^{-1} \Sigma_{\alpha_i} \sum_{j=1}^k \mathbf{Z}_{ij} (Y_{ij} - \mu_{ij}) \\ \Sigma_{\alpha_i} &= \left(\sum_{j=1}^k \frac{\mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top}{\sigma^2} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

- Lastly, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma} \left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \alpha_i^\top \alpha_i \right)$$

A.5.8 BV

Priors

$$\mathbf{Y}|\gamma_0, \gamma_1, \gamma_2, \gamma_3, \boldsymbol{\alpha}_i, \sigma^2 \propto (\sigma^2)^{-\frac{nk}{2}} \exp\left\{-\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^k \left(Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i\right)^2\right\} \quad i = 1, \dots, n \quad j = 1, \dots, k$$

$$\boldsymbol{\gamma}_0 \sim \text{N}_3(\mathbf{0}, \boldsymbol{\Sigma}_0)$$

$$\boldsymbol{\gamma}_1 \sim \text{N}_p(\mathbf{0}, \boldsymbol{\Sigma}_1)$$

$$\gamma_{2l}|\tau_{1l}^2, \sigma^2 \stackrel{\text{ind}}{\sim} \text{N}(0, \sigma^2 \tau_{1l}^2) \quad l = 1, \dots, m$$

$$\gamma_{3q}|\tau_{2q}^2, \sigma^2 \stackrel{\text{ind}}{\sim} \text{N}(0, \sigma^2 \tau_{2q}^2) \quad q = 1, \dots, pm$$

$$\tau_{1l}^2|\lambda_1^2 \stackrel{\text{iid}}{\sim} \frac{\lambda_1^2}{2} \exp\left(-\frac{\lambda_1^2}{2} \tau_{1l}^2\right) \quad l = 1, \dots, m$$

$$\tau_{2q}^2|\lambda_2^2 \stackrel{\text{iid}}{\sim} \frac{\lambda_2^2}{2} \exp\left(-\frac{\lambda_2^2}{2} \tau_{2q}^2\right) \quad q = 1, \dots, pm$$

$$\lambda_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\lambda_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\boldsymbol{\alpha}_i \sim \text{N}(\mathbf{0}, \phi^2 \mathbf{I}) \quad i = 1, \dots, n$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\sigma^2 \sim \frac{1}{\sigma^2}$$

Gibbs Sampler

- Let $\mu_{ij(-\gamma_0)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{X}_{ij}^\top \boldsymbol{\gamma}_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of $\boldsymbol{\gamma}_0$.

The full conditional distribution of $\boldsymbol{\gamma}_0$ is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\boldsymbol{\gamma}_0}, \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0})$ with mean

$$\boldsymbol{\mu}_{\boldsymbol{\gamma}_0} = \boldsymbol{\Sigma}_{\boldsymbol{\gamma}_0} \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{X}_{ij} (Y_{ij} - \mu_{ij(-\gamma_0)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\Sigma_{\gamma_0} = \left(\Sigma_0^{-1} + \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{X}_{ij} \mathbf{X}_{ij}^\top \right)^{-1}.$$

- Denote $\mu_{ij(-\gamma_1)} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - \mathbf{T}_{ij}^\top \gamma_1$, then the full conditional distribution of γ_1 is a multivariate normal distribution $\text{MVN}(\boldsymbol{\mu}_{\gamma_1}, \Sigma_{\gamma_1})$ with mean

$$\boldsymbol{\mu}_{\gamma_1} = \Sigma_{\gamma_1} \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{T}_{ij} (Y_{ij} - \mu_{ij(-\gamma_1)} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\Sigma_{\gamma_1} = \left(\Sigma_1^{-1} + \sigma^{-2} \sum_{i=1}^n \sum_{j=1}^k \mathbf{T}_{ij} \mathbf{T}_{ij}^\top \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{2l})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - L_{ijl} \gamma_{2l}$, the conditional posterior distribution of the coefficient of lipid factor γ_{2l} is a normal distribution: with mean

$$\mu_{\gamma_{2l}} = \Sigma_{\gamma_{2l}} (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k L_{ijl} (Y_{ij} - \mu_{ij(-\gamma_{2l})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\Sigma_{\gamma_{2l}} = \sigma^2 \left(\sum_{i=1}^n \sum_{j=1}^k L_{ijl}^2 + \frac{1}{\tau_{1l}^2} \right)^{-1}.$$

- Let $\mu_{ij(-\gamma_{3q})} = \mu_{ij} + \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i - W_{ijq} \gamma_{3q}$, the conditional posterior distribution of the coefficient of interaction factor γ_{3q} is a normal distribution: with mean

$$\mu_{\gamma_{3q}} = \Sigma_{\gamma_{3q}} (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k j = 1^k W_{ijq} (Y_{ij} - \mu_{ij(-\gamma_{3q})} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i),$$

and variance

$$\Sigma_{\gamma_{3q}} = \sigma^2 \left(\sum_{i=1}^n \sum_{j=1}^k W_{ijq}^2 + \frac{1}{\tau_{2q}^2} \right)^{-1}.$$

- The full conditional posterior distribution of $(\tau_{1l}^2)^{-1}$ is:

$$(\tau_{1l}^2)^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\lambda_1^2, \sqrt{\frac{\sigma^2 \lambda_1^2}{\gamma_{2l}^2}})$$

- Similarly, the full conditional posterior distribution of $(\tau_{2q}^2)^{-1}$ is:

$$(\tau_{2q}^2)^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\lambda_2^2, \sqrt{\frac{\sigma^2 \lambda_2^2}{\gamma_{3q}^2}})$$

- The full conditional posterior distribution of λ_1^2 :

$$\lambda_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m \tau_{1l}^2}{2}).$$

- Likewise, the full conditional posterior distribution of λ_2^2 :

$$\lambda_2^2 | \text{rest} \sim \text{Gamma}(a_2 + pm, b_2 + \frac{\sum_{q=1}^{pm} \tau_{2q}^2}{2}).$$

- The full conditional posterior distribution of σ^2 :

$$\sigma^2 | \text{rest} \sim \text{Inv-Gamma}(\mu_{\sigma^2}, \Sigma_{\sigma^2})$$

with shape

$$\mu_{\sigma^2} = \frac{nk + m + pm}{2}$$

and scale

$$\Sigma_{\sigma^2} = \frac{\sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij} - \mathbf{Z}_{ij}^\top \boldsymbol{\alpha}_i)^2 + \sum_{l=1}^m (\tau_{1l}^2)^{-1} \gamma_{2l}^2 + \sum_{q=1}^{pm} (\tau_{2q}^2)^{-1} \gamma_{3q}^2}{2}$$

- The full conditional posterior distribution of α_i :

$$\alpha_i | \text{rest} \sim \text{MVN}(\mu_{\alpha_i}, \Sigma_{\alpha_i})$$

where

$$\begin{aligned} \mu_{\alpha_i} &= (\sigma^2)^{-1} \Sigma_{\alpha_i} \sum_{j=1}^k \mathbf{Z}_{ij} (Y_{ij} - \mu_{ij}) \\ \Sigma_{\alpha_i} &= \left(\sum_{j=1}^k \frac{\mathbf{Z}_{ij} \mathbf{Z}_{ij}^\top}{\sigma^2} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

- Lastly, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma} \left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \alpha_i^\top \alpha_i \right)$$

Appendix B

Appendices for chapter 3

B.1 Additional simulation results

B.1.1 Identification results for $\theta = 0.5$ in simulation

Table B.1: *Simulation results in example 2 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at 50th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.*

$n = 400 \quad m = 200$		Varying		Constant	
		TP	FP	TP	FP
Error 1 $N(0,1)$	RBVSS - SI	3.93(0.27)	1.76(0.21)	3.98(0.27)	1.61(0.14)
	RBV-SI	3.91(0.21)	4.15(0.32)	3.82(1.41)	3.58(0.89)
	BVSS -SI	3.94(0.13)	1.73(0.21)	3.95(0.31)	1.68(0.24)
	BV - SI	3.91(0.55)	4.28(0.25)	3.87(0.78)	3.71(0.67)
		TP		FP	
	RBVSS	3.94(0.32)		5.45(0.16)	
	RBV	3.91(0.28)		7.29(1.22)	
	BVSS	3.93(0.19)		5.54(0.20)	
Error 2 $t(2)$	BV	3.90(0.19)		7.27(1.17)	
	RBVSS - SI	3.98(0.14)	1.29(0.74)	3.88(0.24)	1.33(0.57)
	RBV - SI	3.94(1.79)	3.58(0.33)	3.04(0.24)	3.76(0.45)
	BVSS - SI	3.80(0.55)	3.17(0.16)	2.65(0.31)	2.77(0.41)
	BV - SI	3.12(0.55)	4.06(0.76)	2.21(1.64)	5.24(0.45)
		TP		FP	
	RBVSS	3.98(0.14)		5.04(1.47)	
	RBV	3.93(0.10)		6.64(0.99)	
Error 3 Lognormal(0,1)	BVSS	3.82(0.44)		7.51(0.20)	
	BV	3.13(0.41)		8.29(1.17)	
	RBVSS - SI	3.91(0.35)	0.93(0.52)	3.95(0.84)	1.21(0.29)
	RBV -SI	3.89(0.14)	2.71(0.25)	3.02(0.82)	3.76(0.15)
	BVSS - SI	3.71(0.89)	3.15(0.21)	2.08(0.49)	2.91(0.45)
	BV - SI	3.07(0.84)	4.59(0.26)	2.04(1.10)	4.79(0.22)
		TP		FP	
	RBVSS	3.91(0.10)		5.51(1.05)	
Error 4 80% $N(0,1)$ +20% $N(0,3)$	RBV	3.87(0.47)		6.86(1.22)	
	BVSS	3.69(0.59)		7.27(1.29)	
	BV	3.06(0.63)		8.98(0.94)	
	RBVSS - SI	3.92(0.23)	0.88(0.27)	3.92(0.55)	1.95(0.24)
	RBV - SI	3.86(0.12)	1.77(0.25)	3.43(0.53)	3.10(0.45)
	BVSS - SI	3.72(0.84)	2.33(0.25)	2.57(2.88)	2.11(0.15)
	BV - SI	3.07(0.55)	5.14(0.15)	2.05(1.14)	3.98(0.25)
		TP		FP	
Error 5 Laplace (0,1)	RBVSS	3.90(0.34)		4.93(0.67)	
	RBV	3.85(0.44)		6.95(0.98)	
	BVSS	3.72(1.08)		5.85(0.20)	
	BV	3.07(0.99)		7.70(1.97)	
	RBVSS - SI	3.95(0.31)	1.21(0.35)	3.94(0.24)	1.87(0.28)
	RBV - SI	3.84(0.45)	2.25(0.45)	3.05(0.31)	3.05(0.37)
	BVSS - SI	3.69(1.30)	2.75(0.45)	2.67(0.61)	2.79(1.34)
	BV - SI	3.12(0.71)	4.89(0.56)	2.32(0.47)	4.05(0.89)
		TP		FP	
	RBVSS	3.95(0.10)		4.66(0.90)	
	RBV	3.88(0.30)		5.98(0.51)	
	BVSS	3.69(0.37)		5.29(0.53)	
	BV	3.13(0.41)		7.29(1.67)	

Table B.2: Estimation accuracy in example 2 for $n = 400, m = 200$ under random intercept - slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 50th quantile with i.i.d. errors.

	RBVSS-SI	RBV-SI	BVSS-SI	BV-SI	RBVSS	RBV	BVSS	BV
Error 1 N(0,1)								
IMSE								
$\beta_0(t)$	0.032(0.009)	0.123(0.021)	0.032(0.027)	0.124(0.014)	0.031(0.012)	0.121(0.011)	0.033(0.008)	0.123(0.011)
$\beta_1(t)$	0.028(0.018)	0.131(0.013)	0.031(0.084)	0.134(0.051)	0.027(0.017)	0.129(0.024)	0.031(0.008)	0.135(0.026)
$\beta_2(t)$	0.027(0.024)	0.118(0.032)	0.031(0.041)	0.118(0.089)	0.025(0.015)	0.119(0.019)	0.032(0.011)	0.115(0.023)
$\beta_3(t)$	0.021(0.017)	0.110(0.021)	0.022(0.035)	0.111(0.065)	0.019(0.052)	0.111(0.031)	0.023(0.010)	0.112(0.029)
$\beta_4(t)$	0.018(0.017)	0.115(0.021)	0.020(0.035)	0.116(0.065)	0.016(0.052)	0.113(0.031)	0.019(0.010)	0.114(0.029)
MSE								
ζ_1	0.007(0.014)	0.013(0.015)	0.011(0.064)	0.015(0.054)	0.008(0.015)	0.015(0.123)	0.010(0.045)	0.018(0.017)
ζ_2	0.005(0.007)	0.017(0.025)	0.006(0.078)	0.021(0.067)	0.006(0.012)	0.018(0.031)	0.005(0.047)	0.022(0.071)
ζ_3	0.003(0.005)	0.018(0.016)	0.004(0.088)	0.023(0.065)	0.004(0.015)	0.019(0.021)	0.006(0.071)	0.025(0.031)
ζ_4	0.006(0.006)	0.021(0.015)	0.015(0.014)	0.034(0.045)	0.007(0.008)	0.023(0.016)	0.016(0.021)	0.033(0.071)
Total	0.123(0.041)	0.287(0.024)	0.125(0.034)	0.293(0.015)	0.158(1.167)	1.090(0.142)	0.163(0.038)	1.112(0.029)
Error 2 t(2)								
IMSE								
$\beta_0(t)$	0.111(0.104)	0.162(0.077)	0.222(0.305)	0.342(1.841)	0.110(0.066)	0.163(0.120)	0.219(0.901)	0.339(0.272)
$\beta_1(t)$	0.026(0.011)	0.231(0.112)	0.343(0.841)	0.553(0.988)	0.024(0.074)	0.229(0.202)	0.341(0.366)	0.550(0.559)
$\beta_2(t)$	0.065(0.041)	0.171(0.099)	0.419(0.397)	0.585(1.666)	0.066(0.171)	0.169(0.143)	0.421(0.491)	0.584(0.301)
$\beta_3(t)$	0.082(0.056)	0.197(0.121)	0.398(0.358)	0.585(0.718)	0.080(0.140)	0.195(0.200)	0.400(0.368)	0.584(0.346)
$\beta_4(t)$	0.036(0.056)	0.152(0.121)	0.311(0.358)	0.498(0.718)	0.034(0.140)	0.149(0.200)	0.309(0.368)	0.496(0.346)
MSE								
ζ_1	0.088(0.047)	0.124(0.063)	0.220(0.079)	0.253(0.537)	0.091(0.056)	0.128(0.071)	0.228(0.073)	0.261(0.025)
ζ_2	0.062(0.009)	0.092(0.061)	0.142(0.285)	0.273(0.067)	0.071(0.011)	0.098(0.081)	0.148(0.038)	0.284(0.321)
ζ_3	0.014(0.008)	0.074(0.063)	0.146(0.577)	0.298(0.067)	0.018(0.023)	0.081(0.034)	0.151(0.023)	0.314(0.035)
ζ_4	0.011(0.003)	0.024(0.025)	0.104(0.213)	0.203(0.045)	0.016(0.023)	0.031(0.039)	0.113(0.029)	0.214(0.038)
Total	0.285(0.104)	1.271(0.340)	1.317(0.302)	1.991(0.256)	0.402(1.318)	1.721(0.658)	1.643(0.941)	3.577(0.232)
Error 3 Lognormal(0,1)								
IMSE								
$\beta_0(t)$	0.085(0.088)	0.142(0.373)	0.139(0.592)	0.385(0.927)	0.083(0.060)	0.141(0.173)	0.141(0.504)	0.381(0.044)
$\beta_1(t)$	0.271(0.160)	0.529(0.288)	0.710(0.579)	0.842(0.441)	0.268(0.079)	0.526(0.340)	0.708(0.301)	0.839(0.611)
$\beta_2(t)$	0.184(0.194)	0.298(0.320)	0.335(0.460)	0.595(0.819)	0.185(0.189)	0.296(0.198)	0.333(0.497)	0.600(0.250)
$\beta_3(t)$	0.154(0.308)	0.275(0.508)	0.343(0.508)	0.605(0.165)	0.155(0.108)	0.274(0.279)	0.339(0.573)	0.602(0.363)
$\beta_4(t)$	0.177(0.308)	0.238(0.508)	0.319(0.508)	0.533(0.165)	0.178(0.108)	0.236(0.279)	0.316(0.573)	0.531(0.363)
MSE								
ζ_1	0.145(0.071)	0.205(0.072)	0.227(0.031)	0.334(0.233)	0.153(0.088)	0.217(0.045)	0.233(0.067)	0.341(0.067)
ζ_2	0.071(0.237)	0.158(0.123)	0.217(0.579)	0.455(0.617)	0.079(0.038)	0.163(0.056)	0.225(0.038)	0.464(0.079)
ζ_3	0.051(0.171)	0.183(0.160)	0.193(0.599)	0.368(0.627)	0.061(0.078)	0.189(0.076)	0.201(0.045)	0.389(0.047)
ζ_4	0.021(0.143)	0.098(0.086)	0.111(0.173)	0.372(0.158)	0.029(0.035)	0.095(0.039)	0.119(0.038)	0.381(0.048)
Total	0.459(0.471)	1.328(0.637)	1.457(1.248)	2.282(0.438)	0.786(0.857)	1.515(0.773)	1.862(1.381)	3.231(0.008)
Error 4 80%N(0,1)+ 20%N(0,3)								
IMSE								
$\beta_0(t)$	0.035(0.033)	0.115(0.060)	0.134(1.112)	0.743(1.387)	0.036(0.009)	0.113(0.013)	0.135(0.967)	0.741(0.230)
$\beta_1(t)$	0.022(0.024)	0.061(0.039)	0.087(0.198)	0.122(0.527)	0.020(0.012)	0.059(0.075)	0.090(0.438)	0.119(0.554)
$\beta_2(t)$	0.019(0.034)	0.046(0.027)	0.088(0.606)	0.223(0.921)	0.018(0.011)	0.047(0.032)	0.086(0.877)	0.222(0.183)
$\beta_3(t)$	0.014(0.025)	0.033(0.043)	0.051(0.830)	0.234(0.843)	0.012(0.025)	0.031(0.085)	0.048(0.890)	0.235(0.592)
$\beta_4(t)$	0.031(0.025)	0.051(0.043)	0.069(0.830)	0.199(0.843)	0.030(0.025)	0.049(0.085)	0.068(0.890)	0.200(0.592)
MSE								
ζ_1	0.013(0.030)	0.015(0.013)	0.033(0.265)	0.099(0.919)	0.019(0.035)	0.018(0.038)	0.041(0.038)	0.102(0.048)
ζ_2	0.007(0.011)	0.018(0.020)	0.047(0.321)	0.111(0.692)	0.009(0.041)	0.020(0.029)	0.051(0.045)	0.114(0.038)
ζ_3	0.008(0.027)	0.021(0.038)	0.061(0.927)	0.123(0.131)	0.010(0.054)	0.028(0.039)	0.069(0.041)	0.125(0.046)
ζ_4	0.007(0.011)	0.027(0.025)	0.081(0.987)	0.095(0.278)	0.011(0.038)	0.031(0.089)	0.083(0.056)	0.096(0.047)
Total	0.115(0.048)	0.154(0.153)	0.233(0.278)	0.487(0.688)	0.211(0.443)	0.282(0.271)	0.554(0.533)	1.677(0.592)
Error 5 Laplace (0,1)								
IMSE								
$\beta_0(t)$	0.043(0.026)	0.154(0.033)	0.236(0.036)	0.388(0.049)	0.042(0.031)	0.152(0.055)	0.234(0.119)	0.388(0.143)
$\beta_1(t)$	0.033(0.045)	0.148(0.082)	0.221(0.059)	0.335(0.077)	0.031(0.012)	0.149(0.048)	0.220(0.032)	0.334(0.067)
$\beta_2(t)$	0.012(0.066)	0.117(0.049)	0.148(0.077)	0.199(0.063)	0.011(0.018)	0.116(0.032)	0.147(0.053)	0.197(0.178)
$\beta_3(t)$	0.035(0.064)	0.196(0.088)	0.222(0.035)	0.389(0.092)	0.034(0.015)	0.195(0.050)	0.221(0.038)	0.387(0.075)
$\beta_4(t)$	0.052(0.064)	0.185(0.088)	0.207(0.035)	0.325(0.092)	0.051(0.015)	0.184(0.050)	0.206(0.038)	0.323(0.075)
MSE								
ζ_1	0.012(0.077)	0.031(0.029)	0.066(0.109)	0.125(0.077)	0.018(0.034)	0.039(0.047)	0.071(0.035)	0.131(0.045)
ζ_2	0.038(0.048)	0.057(0.035)	0.098(0.070)	0.176(0.044)	0.040(0.051)	0.058(0.051)	0.111(0.048)	0.181(0.059)
ζ_3	0.015(0.014)	0.054(0.057)	0.082(0.063)	0.191(0.049)	0.019(0.045)	0.061(0.059)	0.086(0.047)	0.195(0.039)
ζ_4	0.025(0.029)	0.086(0.052)	0.121(0.106)	0.235(0.054)	0.028(0.091)	0.088(0.058)	0.128(0.045)	0.241(0.048)
Total	0.191(0.124)	0.233(0.234)	0.449(0.139)	0.726(0.579)	0.277(0.450)	0.430(0.179)	0.950(0.472)	1.372(0.556)

B.1.2 Identification results for $\theta = 0.3$ in simulation

Table B.3: *Simulation results in example 1 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at 30th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.*

$n = 400 \quad m = 200$		Varying		Constant	
		TP	FP	TP	FP
Error 1 N(0,1)	RBVSS-SI	3.91(0.22)	1.66 (0.33)	3.95(0.27)	1.62(1.78)
	RBV-SI	3.88(0.10)	3.99(0.83)	3.64(0.46)	3.32(0.76)
	BVSS-SI	3.92(0.31)	1.72(0.21)	3.92(0.31)	1.71(0.24)
	BV-SI	3.82(0.12)	4.58(0.25)	3.59(1.78)	3.61(0.67)
		TP		FP	
	RBVSS	3.91(0.33)		5.36(0.37)	
	RBV	3.89(0.21)		7.38(1.66)	
	BVSS	3.92(0.43)		5.31(0.24)	
	BV	3.83(0.21)		7.55(1.27)	
Error 2 $t(2)$	RBVSS-SI	3.98(0.14)	0.57(1.74)	3.94(0.24)	1.89(0.57)
	RBV-SI	3.94(0.27)	2.97(0.51)	2.95(0.50)	3.88(0.29)
	BVSS-SI	3.83(0.56)	3.34(0.75)	2.63(1.53)	3.17(2.05)
	BV-SI	3.14(0.81)	3.82(0.76)	1.97(1.64)	5.74(1.31)
		TP		FP	
	RBVSS	3.97(0.45)		5.08(0.51)	
	RBV	3.92(0.37)		6.55(1.11)	
	BVSS	3.89(0.45)		7.44(1.05)	
	BV	3.14(0.30)		8.24(1.27)	
Error 3 Lognormal(0,1)	RBVSS-SI	3.94(0.15)	0.79(0.35)	3.86(0.54)	1.07(0.76)
	RBV-SI	3.91(0.54)	1.98(0.25)	2.99(1.82)	3.25(0.15)
	BVSS-SI	3.74(0.89)	2.75(0.78)	2.52(0.27)	3.15(0.45)
	BV-SI	3.11(0.84)	4.31(0.26)	1.99(1.12)	4.67(0.22)
		TP		FP	
	RBVSS	3.93(0.18)		5.27(0.37)	
	RBV	3.89(0.29)		7.12(1.66)	
	BVSS	3.87(0.33)		7.05(0.44)	
	BV	3.11(0.52)		8.55(1.27)	
Error 4 80%N(0,1) +20%N(0,3)	RBVSS -SI	3.94(0.33)	0.84(0.27)	3.75(0.55)	1.26(0.24)
	RBV-SI	3.91(1.10)	1.48(0.25)	3.02(1.23)	2.54(0.45)
	BVSS-SI	3.82(0.84)	2.86(0.25)	2.85(1.88)	2.22(1.06)
	BV-SI	3.15(0.18)	4.21(0.15)	2.02(1.14)	3.21(0.25)
		TP		FP	
	RBVSS	3.94(0.11)		4.74(0.26)	
	RBV	3.91(0.32)		6.27(1.66)	
	BVSS	3.82(0.29)		5.77(0.58)	
	BV	3.15(0.18)		6.55(1.68)	
Error 5 Laplace (0,1)	RBVSS-SI	3.95(0.31)	1.14(0.35)	3.91(0.24)	1.92(0.28)
	RBV-SI	3.70(0.45)	2.18(0.45)	3.15(1.31)	2.64(0.37)
	BVSS-SI	3.83(1.30)	4.88(0.45)	3.02(1.61)	2.71(1.34)
	BV-SI	3.05(0.71)	5.03(0.56)	2.12(1.07)	3.53(0.89)
		TP		FP	
	RBVSS	3.95(0.18)		4.71(0.60)	
	RBV	3.71(0.22)		5.88(0.84)	
	BVSS	3.81(0.26)		5.47(0.95)	
	BV	3.06(0.29)		7.24(1.95)	

Table B.4: Estimation accuracy in example 1 for $n = 400, m = 200$ under random intercept - slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 30th quantile with i.i.d. errors.

	RBVSS-SI	RBV-SI	BVSS-SI	BV-SI	RBVSS	RBV	BVSS	BV
Error 1 N(0,1)								
IMSE								
$\beta_0(t)$	0.028(0.007)	0.115(0.088)	0.031(0.024)	0.113(0.097)	0.027(0.005)	0.114(0.009)	0.029(0.004)	0.115(0.006)
$\beta_1(t)$	0.025(0.015)	0.110(0.035)	0.031(0.026)	0.121(0.065)	0.023(0.006)	0.108(0.013)	0.032(0.006)	0.120(0.014)
$\beta_2(t)$	0.036(0.017)	0.135(0.045)	0.041(0.018)	0.139(0.056)	0.035(0.008)	0.134(0.011)	0.040(0.006)	0.138(0.012)
$\beta_3(t)$	0.022(0.012)	0.077(0.041)	0.021(0.023)	0.082(0.036)	0.021(0.008)	0.074(0.020)	0.020(0.007)	0.079(0.015)
$\beta_4(t)$	0.041(0.012)	0.086(0.041)	0.053(0.023)	0.079(0.036)	0.040(0.008)	0.086(0.020)	0.052(0.007)	0.077(0.015)
MSE								
ζ_1	0.011(0.018)	0.028(0.038)	0.012(0.021)	0.033(0.032)	0.017(0.035)	0.031(0.061)	0.015(0.025)	0.041(0.045)
ζ_2	0.005(0.005)	0.023(0.022)	0.007(0.022)	0.026(0.028)	0.009(0.012)	0.029(0.037)	0.008(0.021)	0.027(0.031)
ζ_3	0.011(0.008)	0.032(0.030)	0.014(0.025)	0.039(0.032)	0.016(0.031)	0.038(0.052)	0.019(0.037)	0.041(0.052)
ζ_4	0.006(0.009)	0.016(0.022)	0.005(0.035)	0.018(0.024)	0.009(0.025)	0.019(0.038)	0.008(0.041)	0.021(0.079)
Total	0.129(0.040)	1.097(0.196)	0.132(0.037)	1.111(0.265)	0.151(0.026)	1.099(0.018)	0.134(0.020)	1.114(0.062)
Error 2 t(2)								
IMSE								
$\beta_0(t)$	0.102(0.430)	0.191(0.104)	0.279(0.287)	0.412(0.841)	0.101(0.007)	0.190(0.103)	0.278(1.431)	0.411(0.675)
$\beta_1(t)$	0.027(0.011)	0.214(0.112)	0.331(0.844)	0.612(0.988)	0.026(0.015)	0.215(0.144)	0.330(0.720)	0.611(0.421)
$\beta_2(t)$	0.086(0.041)	0.157(0.099)	0.311(0.397)	0.415(1.666)	0.087(0.020)	0.156(0.107)	0.310(0.444)	0.412(0.223)
$\beta_3(t)$	0.045(0.056)	0.207(0.121)	0.312(0.378)	0.435(0.718)	0.042(0.033)	0.206(0.146)	0.310(0.385)	0.433(0.321)
$\beta_4(t)$	0.068(0.056)	0.233(0.121)	0.342(0.378)	0.501(0.718)	0.066(0.033)	0.230(0.146)	0.341(0.385)	0.499(0.321)
MSE								
ζ_1	0.083(0.047)	0.103(0.063)	0.220(0.079)	0.251(0.537)	0.085(0.032)	0.104(0.051)	0.231(0.045)	0.263(0.071)
ζ_2	0.047(0.551)	0.081(0.009)	0.132(0.285)	0.271(0.672)	0.051(0.042)	0.083(0.038)	0.140(0.052)	0.279(0.038)
ζ_3	0.017(0.008)	0.048(0.061)	0.138(0.577)	0.406(0.673)	0.021(0.028)	0.052(0.021)	0.139(0.041)	0.410(0.077)
ζ_4	0.022(0.003)	0.044(0.037)	0.111(0.213)	0.205(0.167)	0.024(0.018)	0.051(0.045)	0.115(0.032)	0.207(0.045)
Total	0.275(0.104)	1.228(0.340)	1.237(1.456)	2.111(0.556)	0.378(0.022)	1.557(0.518)	1.525(0.889)	3.424(0.574)
Error 3 Lognormal(0,1)								
IMSE								
$\beta_0(t)$	0.064(0.040)	0.124(0.237)	0.151(0.598)	0.536(0.589)	0.062(0.036)	0.122(0.153)	0.150(0.587)	0.534(0.897)
$\beta_1(t)$	0.252(0.113)	0.433(0.256)	0.578(0.462)	0.843(0.319)	0.251(0.046)	0.432(0.084)	0.572(0.239)	0.841(0.495)
$\beta_2(t)$	0.175(0.073)	0.288(0.161)	0.311(0.441)	0.588(0.285)	0.173(0.130)	0.286(0.202)	0.310(0.520)	0.586(0.246)
$\beta_3(t)$	0.222(0.188)	0.265(0.262)	0.333(0.497)	0.517(0.381)	0.221(0.112)	0.264(0.270)	0.331(0.456)	0.513(0.321)
$\beta_4(t)$	0.243(0.188)	0.321(0.262)	0.428(0.497)	0.603(0.381)	0.240(0.112)	0.319(0.270)	0.426(0.456)	0.600(0.321)
MSE								
ζ_1	0.126(0.107)	0.186(0.099)	0.212(0.030)	0.331(0.541)	0.133(0.038)	0.192(0.041)	0.221(0.025)	0.341(0.025)
ζ_2	0.042(0.130)	0.132(0.117)	0.172(0.240)	0.430(0.270)	0.051(0.047)	0.141(0.049)	0.181(0.049)	0.441(0.041)
ζ_3	0.029(0.057)	0.121(0.146)	0.187(0.616)	0.322(0.671)	0.035(0.051)	0.126(0.035)	0.189(0.028)	0.323(0.051)
ζ_4	0.021(0.025)	0.068(0.075)	0.111(0.194)	0.229(0.378)	0.027(0.049)	0.072(0.045)	0.114(0.041)	0.231(0.056)
Total	0.515(0.295)	1.255(0.623)	1.419(1.323)	2.017(1.289)	0.792(0.456)	1.412(0.720)	1.594(0.456)	2.982(0.321)
Error 4 80%N(0,1)+ 20%N(0,3)								
IMSE								
$\beta_0(t)$	0.031(0.006)	0.065(0.009)	0.123(0.214)	0.803(0.943)	0.029(0.006)	0.064(0.009)	0.121(1.093)	0.802(2.967)
$\beta_1(t)$	0.022(0.026)	0.075(0.093)	0.114(0.235)	0.528(0.239)	0.020(0.007)	0.071(0.128)	0.113(1.415)	0.526(0.352)
$\beta_2(t)$	0.021(0.021)	0.065(0.023)	0.099(0.808)	0.213(0.891)	0.019(0.007)	0.063(0.046)	0.096(0.494)	0.211(0.331)
$\beta_3(t)$	0.017(0.047)	0.042(0.075)	0.066(0.351)	0.362(0.892)	0.015(0.008)	0.039(0.088)	0.064(0.962)	0.360(0.047)
$\beta_4(t)$	0.029(0.047)	0.056(0.075)	0.079(0.351)	0.308(0.892)	0.028(0.008)	0.054(0.088)	0.077(0.962)	0.305(0.047)
MSE								
ζ_1	0.011(0.026)	0.019(0.009)	0.047(0.032)	0.111(0.218)	0.018(0.031)	0.025(0.011)	0.058(0.067)	0.118(0.045)
ζ_2	0.007(0.004)	0.013(0.006)	0.052(0.290)	0.075(0.672)	0.011(0.045)	0.019(0.061)	0.061(0.048)	0.088(0.041)
ζ_3	0.005(0.003)	0.016(0.025)	0.081(0.632)	0.117(0.349)	0.009(0.031)	0.021(0.054)	0.089(0.076)	0.129(0.122)
ζ_4	0.001(0.005)	0.009(0.013)	0.061(0.581)	0.082(0.786)	0.006(0.041)	0.017(0.027)	0.069(0.041)	0.091(0.023)
Total	0.092(0.092)	0.118(0.182)	0.229(0.425)	0.413(0.521)	0.131(0.052)	0.258(0.416)	0.475(0.781)	0.524(0.951)
Error 5 Laplace(0,1)								
IMSE								
$\beta_0(t)$	0.044(0.023)	0.162(0.013)	0.312(0.019)	0.467(0.146)	0.042(0.010)	0.160(0.036)	0.309(0.101)	0.464(0.122)
$\beta_1(t)$	0.022(0.032)	0.114(0.217)	0.203(0.018)	0.333(0.513)	0.020(0.006)	0.112(0.018)	0.200(0.027)	0.334(0.061)
$\beta_2(t)$	0.023(0.021)	0.122(0.120)	0.161(0.019)	0.223(0.765)	0.021(0.007)	0.121(0.021)	0.159(0.030)	0.224(0.044)
$\beta_3(t)$	0.041(0.024)	0.313(0.341)	0.326(0.015)	0.521(0.654)	0.039(0.008)	0.310(0.025)	0.322(0.033)	0.519(0.065)
$\beta_4(t)$	0.039(0.023)	0.213(0.341)	0.298(0.015)	0.444(0.654)	0.038(0.008)	0.211(0.025)	0.299(0.033)	0.443(0.065)
MSE								
ζ_1	0.022(0.023)	0.031(0.051)	0.086(0.045)	0.144(0.123)	0.028(0.041)	0.039(0.038)	0.091(0.028)	0.149(0.047)
ζ_2	0.053(0.021)	0.078(0.025)	0.122(0.058)	0.196(0.325)	0.061(0.045)	0.084(0.035)	0.129(0.045)	0.200(0.049)
ζ_3	0.028(0.024)	0.051(0.029)	0.186(0.034)	0.287(0.256)	0.034(0.038)	0.059(0.074)	0.193(0.056)	0.296(0.059)
ζ_4	0.019(0.013)	0.088(0.034)	0.111(0.067)	0.222(0.111)	0.028(0.051)	0.091(0.033)	0.119(0.045)	0.229(0.045)
Total	0.174(0.028)	0.222(0.056)	0.287(0.124)	0.389(0.135)	0.254(0.025)	0.441(0.114)	0.756(0.268)	0.999(0.480)

Table B.5: *Simulation results in example 2 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at 30th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.*

$n = 400 \quad m = 200$		Varying		Constant	
		TP	FP	TP	FP
Error 1 $N(0,1)$	RBVSS - SI	3.95(0.17)	1.88(0.21)	3.97(0.27)	1.54(0.14)
	RBV-SI	3.91(0.21)	3.69(0.29)	3.89(1.41)	3.75(0.89)
	BVSS -SI	3.94(0.13)	1.67(0.21)	3.93(0.31)	1.61(0.24)
	BV - SI	3.93(0.55)	3.77(0.25)	3.90(0.78)	3.91(0.67)
		TP		FP	
	RBVSS	3.95(0.41)		5.21(0.16)	
	RBV	3.92(0.31)		6.31(1.22)	
	BVSS	3.94(0.21)		5.33(0.20)	
	BV	3.91(0.18)		6.37(1.17)	
Error 2 $t(2)$	RBVSS - SI	3.91(0.14)	1.35(0.74)	3.86(0.24)	1.46(0.57)
	RBV - SI	3.88(1.79)	3.10(0.33)	3.25(0.24)	2.59(0.45)
	BVSS - SI	3.85(0.55)	2.87(0.16)	2.85(0.31)	2.88(0.41)
	BV - SI	3.27(0.55)	4.55(0.76)	2.02(1.64)	6.24(0.45)
		TP		FP	
	RBVSS	3.91(0.14)		5.27(1.47)	
	RBV	3.88(0.10)		6.63(0.99)	
	BVSS	3.85(0.44)		7.11(0.20)	
	BV	3.31(0.41)		8.12(1.17)	
Error 3 Lognormal(0,1)	RBVSS - SI	3.86(0.35)	1.11(0.52)	3.75(0.84)	1.51(0.29)
	RBV -SI	3.78(0.14)	2.44(0.25)	3.11(0.82)	3.88(0.15)
	BVSS - SI	3.81(0.89)	3.27(0.21)	2.79(0.49)	3.11(0.45)
	BV - SI	3.08(0.84)	4.71(0.26)	2.17(1.10)	3.88(0.22)
		TP		FP	
	RBVSS	3.86(0.10)		5.78(1.05)	
	RBV	3.74(0.47)		6.89(1.22)	
	BVSS	3.81(0.59)		7.36(1.29)	
	BV	3.06(0.63)		8.19(0.94)	
Error 4 80% $N(0,1)$ +20% $N(0,3)$	RBVSS - SI	3.89(0.18)	0.99(0.27)	3.95(0.55)	1.84(0.24)
	RBV - SI	3.84(0.12)	1.99(0.25)	3.03(0.53)	2.99(0.45)
	BVSS - SI	3.72(0.84)	2.71(0.25)	2.12(2.88)	2.21(0.15)
	BV - SI	3.11(0.55)	4.82(0.15)	2.08(1.14)	3.22(0.25)
		TP		FP	
	RBVSS	3.90(0.11)		4.83(0.67)	
	RBV	3.83(0.56)		6.85(0.98)	
	BVSS	3.71(1.08)		5.35(0.20)	
	BV	3.13(0.99)		7.11(1.97)	
Error 5 Laplace (0,1)	RBVSS - SI	3.91(0.31)	1.09(0.35)	3.91(0.24)	1.37(0.28)
	RBV - SI	3.78(0.45)	2.12(0.45)	2.97(0.31)	2.99(0.37)
	BVSS - SI	3.89(1.30)	2.69(0.45)	2.36(0.61)	2.77(1.34)
	BV - SI	3.31(0.71)	4.57(0.56)	2.12(0.47)	4.11(0.89)
		TP		FP	
	RBVSS	3.90(0.10)		4.54(0.90)	
	RBV	3.78(0.43)		5.69(0.51)	
	BVSS	3.88(0.37)		5.34(0.53)	
	BV	3.33(0.41)		7.23(1.67)	

Table B.6: *Estimation accuracy in example 2 for $n = 400, m = 200$ under random intercept-slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 30th quantile with i.i.d. errors.*

	RBVSS-SI	RBV-SI	BVSS-SI	BV-SI	RBVSS	RBV	BVSS	BV
Error 1 N(0,1)								
IMSE								
$\beta_0(t)$	0.026(0.009)	0.116(0.021)	0.029(0.027)	0.120(0.014)	0.024(0.012)	0.114(0.011)	0.028(0.008)	0.122(0.011)
$\beta_1(t)$	0.031(0.018)	0.122(0.013)	0.034(0.084)	0.125(0.051)	0.029(0.017)	0.122(0.024)	0.033(0.008)	0.124(0.026)
$\beta_2(t)$	0.032(0.024)	0.121(0.032)	0.033(0.041)	0.123(0.089)	0.030(0.015)	0.119(0.019)	0.032(0.011)	0.121(0.023)
$\beta_3(t)$	0.025(0.017)	0.131(0.021)	0.028(0.035)	0.134(0.065)	0.023(0.052)	0.130(0.031)	0.027(0.010)	0.132(0.029)
$\beta_4(t)$	0.041(0.017)	0.138(0.021)	0.038(0.035)	0.139(0.065)	0.040(0.052)	0.137(0.031)	0.037(0.010)	0.138(0.029)
MSE								
ζ_1	0.009(0.014)	0.015(0.015)	0.008(0.064)	0.017(0.054)	0.013(0.026)	0.019(0.045)	0.010(0.041)	0.021(0.031)
ζ_2	0.006(0.007)	0.019(0.025)	0.007(0.078)	0.021(0.067)	0.009(0.034)	0.023(0.045)	0.009(0.032)	0.028(0.041)
ζ_3	0.004(0.005)	0.021(0.016)	0.005(0.088)	0.023(0.065)	0.008(0.021)	0.028(0.032)	0.008(0.056)	0.030(0.035)
ζ_4	0.005(0.006)	0.017(0.015)	0.005(0.014)	0.022(0.045)	0.009(0.014)	0.023(0.041)	0.009(0.034)	0.029(0.025)
Total	0.127(0.041)	0.265(0.024)	0.131(0.034)	0.271(0.015)	0.192(1.167)	1.112(0.142)	0.201(0.038)	1.118(0.029)
Error 2 t(2)								
IMSE								
$\beta_0(t)$	0.132(0.104)	0.171(0.077)	0.245(0.305)	0.356(1.841)	0.131(0.066)	0.170(0.120)	0.243(0.901)	0.354(0.272)
$\beta_1(t)$	0.031(0.011)	0.243(0.112)	0.332(0.841)	0.497(0.988)	0.029(0.074)	0.244(0.202)	0.331(0.366)	0.496(0.559)
$\beta_2(t)$	0.071(0.041)	0.182(0.099)	0.421(0.397)	0.591(0.666)	0.070(0.171)	0.180(0.143)	0.420(0.491)	0.590(0.301)
$\beta_3(t)$	0.075(0.056)	0.203(0.121)	0.376(0.358)	0.591(0.718)	0.074(0.140)	0.202(0.200)	0.374(0.368)	0.589(0.346)
$\beta_4(t)$	0.061(0.056)	0.221(0.121)	0.389(0.358)	0.543(0.718)	0.060(0.140)	0.222(0.200)	0.387(0.368)	0.541(0.346)
MSE								
ζ_1	0.097(0.047)	0.133(0.063)	0.234(0.079)	0.262(0.537)	0.111(0.052)	0.140(0.037)	0.239(0.081)	0.271(0.045)
ζ_2	0.071(0.009)	0.097(0.061)	0.161(0.285)	0.282(0.067)	0.079(0.035)	0.111(0.025)	0.171(0.032)	0.291(0.032)
ζ_3	0.011(0.008)	0.068(0.063)	0.138(0.577)	0.276(0.067)	0.016(0.041)	0.075(0.031)	0.143(0.041)	0.284(0.039)
ζ_4	0.014(0.003)	0.034(0.025)	0.113(0.213)	0.224(0.045)	0.019(0.035)	0.043(0.045)	0.118(0.071)	0.229(0.078)
Total	0.292(0.104)	1.283(0.340)	1.346(3.302)	1.888(0.256)	0.389(1.318)	1.657(0.658)	1.548(1.941)	3.476(0.232)
Error 3 Lognormal(0,1)								
IMSE								
$\beta_0(t)$	0.091(0.088)	0.151(0.373)	0.311(0.592)	0.325(1.927)	0.090(0.060)	0.149(0.173)	0.310(0.504)	0.324(1.044)
$\beta_1(t)$	0.330(0.160)	0.539(0.288)	0.776(0.579)	0.886(0.441)	0.328(0.079)	0.536(0.340)	0.775(0.301)	0.884(0.611)
$\beta_2(t)$	0.176(0.194)	0.308(0.320)	0.367(0.460)	0.635(0.819)	0.175(0.189)	0.306(0.198)	0.365(0.497)	0.633(0.250)
$\beta_3(t)$	0.162(0.308)	0.270(0.508)	0.351(0.508)	0.626(0.165)	0.160(0.108)	0.268(0.279)	0.349(0.573)	0.624(0.363)
$\beta_4(t)$	0.207(0.308)	0.312(0.508)	0.401(0.508)	0.687(0.165)	0.206(0.108)	0.311(0.279)	0.399(0.573)	0.685(0.363)
MSE								
ζ_1	0.155(0.071)	0.215(0.072)	0.225(0.031)	0.341(0.233)	0.166(0.091)	0.226(0.057)	0.231(0.041)	0.349(0.031)
ζ_2	0.083(0.237)	0.173(0.123)	0.232(0.579)	0.501(0.617)	0.088(0.041)	0.181(0.045)	0.235(0.057)	0.503(0.047)
ζ_3	0.063(0.171)	0.168(0.160)	0.202(0.599)	0.388(0.627)	0.067(0.043)	0.173(0.045)	0.211(0.048)	0.392(0.071)
ζ_4	0.024(0.143)	0.101(0.086)	0.124(0.173)	0.412(0.158)	0.028(0.039)	0.108(0.043)	0.129(0.048)	0.419(0.088)
Total	0.444(0.471)	1.128(0.637)	1.472(1.248)	2.322(0.438)	0.764(0.857)	1.613(0.773)	1.725(1.381)	3.179(4.008)
Error 4 80%N(0,1)+ 20%N(0,3)								
IMSE								
$\beta_0(t)$	0.042(0.033)	0.125(0.060)	0.141(1.112)	0.688(1.387)	0.042(0.009)	0.123(0.013)	0.139(0.967)	0.686(0.230)
$\beta_1(t)$	0.029(0.024)	0.067(0.039)	0.091(0.198)	0.131(0.527)	0.028(0.012)	0.065(0.075)	0.090(0.438)	0.129(0.554)
$\beta_2(t)$	0.022(0.034)	0.058(0.027)	0.098(0.606)	0.239(1.921)	0.021(0.011)	0.057(0.032)	0.096(0.877)	0.237(1.183)
$\beta_3(t)$	0.018(0.025)	0.035(0.043)	0.064(0.830)	0.243(0.843)	0.016(0.025)	0.034(0.085)	0.062(0.890)	0.241(0.592)
$\beta_4(t)$	0.031(0.025)	0.042(0.043)	0.055(0.830)	0.207(0.843)	0.029(0.025)	0.040(0.085)	0.053(0.890)	0.206(0.592)
MSE								
ζ_1	0.017(0.030)	0.022(0.013)	0.042(0.265)	0.111(0.919)	0.019(0.045)	0.026(0.078)	0.049(0.071)	0.119(0.048)
ζ_2	0.009(0.011)	0.021(0.020)	0.059(0.321)	0.145(0.692)	0.011(0.034)	0.029(0.064)	0.068(0.034)	0.146(0.075)
ζ_3	0.011(0.027)	0.031(0.038)	0.071(0.927)	0.133(0.131)	0.016(0.029)	0.035(0.078)	0.080(0.075)	0.137(0.049)
ζ_4	0.005(0.011)	0.023(0.025)	0.091(0.987)	0.103(0.278)	0.008(0.057)	0.029(0.045)	0.092(0.048)	0.106(0.018)
Total	0.113(0.048)	0.145(0.153)	0.241(0.278)	0.511(0.688)	0.311(0.443)	0.379(0.271)	0.666(0.533)	0.765(0.592)
Error 5 Laplace (0,1)								
IMSE								
$\beta_0(t)$	0.051(0.023)	0.168(0.033)	0.276(0.036)	0.396(0.049)	0.049(0.031)	0.167(0.055)	0.275(0.119)	0.395(0.143)
$\beta_1(t)$	0.042(0.045)	0.152(0.082)	0.254(0.059)	0.355(0.077)	0.042(0.012)	0.151(0.048)	0.254(0.032)	0.353(0.067)
$\beta_2(t)$	0.017(0.066)	0.125(0.049)	0.157(0.077)	0.188(0.063)	0.015(0.018)	0.124(0.032)	0.157(0.053)	0.186(0.178)
$\beta_3(t)$	0.038(0.064)	0.179(0.088)	0.218(0.035)	0.376(0.092)	0.036(0.015)	0.177(0.050)	0.218(0.038)	0.377(0.075)
$\beta_4(t)$	0.022(0.064)	0.136(0.088)	0.171(0.035)	0.246(0.092)	0.021(0.015)	0.135(0.050)	0.170(0.038)	0.244(0.075)
MSE								
ζ_1	0.013(0.077)	0.041(0.029)	0.077(0.109)	0.135(0.077)	0.018(0.038)	0.049(0.032)	0.083(0.028)	0.143(0.079)
ζ_2	0.031(0.048)	0.062(0.035)	0.092(0.070)	0.166(0.044)	0.039(0.028)	0.071(0.048)	0.111(0.049)	0.174(0.038)
ζ_3	0.013(0.014)	0.064(0.057)	0.092(0.063)	0.210(0.049)	0.019(0.057)	0.071(0.032)	0.096(0.032)	0.218(0.038)
ζ_4	0.022(0.029)	0.079(0.052)	0.124(0.106)	0.244(0.054)	0.029(0.049)	0.088(0.038)	0.129(0.071)	0.261(0.078)
Total	0.211(0.124)	0.287(0.234)	0.469(0.139)	0.876(0.579)	0.377(0.450)	0.589(0.179)	0.988(0.472)	1.415(0.556)

B.1.3 Identification results for $\theta = 0.7$ in simulation

Table B.7: *Simulation results in example 1 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept - slope model based on 100 replicates at 70th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.*

$n = 400 \quad m = 200$		Varying		Constant	
		TP	FP	TP	FP
Error 1 $N(0,1)$	RBVSS-SI	3.89(0.22)	1.58(0.33)	3.98(0.27)	1.59(1.78)
	RBV-SI	3.88(0.10)	3.71(0.83)	3.72(0.46)	3.65(0.76)
	BVSS-SI	3.90(0.31)	1.62(0.21)	3.96(0.31)	1.64(0.24)
	BV-SI	3.85(0.12)	3.82(0.25)	3.68(1.78)	3.71(0.67)
		TP		FP	
	RBVSS	3.90(0.33)		5.59(0.37)	
	RBV	3.89(0.21)		7.21(1.66)	
	BVSS	3.91(0.43)		5.61(0.24)	
Error 2 $t(2)$	BV	3.88(0.21)		7.33(1.27)	
	RBVSS-SI	3.98(0.14)	0.46(1.74)	3.94(0.24)	1.76(0.57)
	RBV-SI	3.91(0.27)	2.89(0.51)	3.17(0.50)	3.91(0.29)
	BVSS-SI	3.79(0.56)	3.47(0.75)	2.91(1.53)	3.28(2.05)
	BV-SI	3.18(0.81)	3.79(0.76)	2.01(1.64)	5.33(1.31)
		TP		FP	
	RBVSS	3.97(0.45)		5.28(0.51)	
	RBV	3.92(0.37)		6.77(1.11)	
Error 3 Lognormal(0,1)	BVSS	3.78(0.45)		7.55(1.05)	
	BV	3.17(0.30)		8.11(1.27)	
	RBVSS-SI	3.96(0.15)	0.83(0.35)	3.89(0.54)	1.11(0.76)
	RBV-SI	3.90(0.54)	2.09(0.25)	3.18(1.82)	3.88(0.15)
	BVSS-SI	3.74(0.89)	2.89(0.78)	2.88(0.27)	3.73(0.45)
	BV-SI	3.04(0.84)	4.22(0.26)	2.11(1.12)	4.99(0.22)
		TP		FP	
	RBVSS	3.95(0.18)		5.25(0.37)	
Error 4 80%N(0,1) +20%N(0,3)	RBV	3.89(0.29)		7.29(1.66)	
	BVSS	3.75(0.33)		7.11(0.44)	
	BV	3.03(0.52)		8.42(1.27)	
	RBVSS -SI	3.94(0.33)	0.92(0.27)	3.85(0.55)	1.31(0.24)
	RBV-SI	3.91(1.10)	1.89(0.25)	3.33(1.23)	2.49(0.45)
	BVSS-SI	3.81(0.84)	2.92(0.25)	2.89(1.88)	3.12(1.06)
	BV-SI	3.01(0.18)	4.33(0.15)	2.22(1.14)	4.78(0.25)
		TP		FP	
Error 5 Laplace (0,1)	RBVSS	3.94(0.11)		4.89(0.26)	
	RBV	3.91(0.32)		6.12(1.66)	
	BVSS	3.82(0.29)		5.81(0.58)	
	BV	3.02(0.18)		6.65(1.68)	
	RBVSS-SI	3.95(0.31)	1.22(0.35)	3.95(0.24)	1.85(0.28)
	RBV-SI	3.81(0.45)	2.25(0.45)	3.41(1.31)	2.53(0.37)
	BVSS-SI	3.79(1.30)	4.76(0.45)	3.07(1.61)	2.83(1.34)
	BV-SI	3.11(0.71)	5.49(0.56)	2.25(1.07)	3.89(0.89)
		TP		FP	
	RBVSS	3.95(0.18)		4.61(0.60)	
	RBV	3.81(0.22)		5.75(0.84)	
	BVSS	3.79(0.26)		5.62(0.95)	
	BV	3.10(0.29)		7.14(1.95)	

Table B.8: Estimation accuracy in example 1 for $n = 400, m = 200$ under random intercept - slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 70th quantile with i.i.d. errors.

	RBVSS-SI	RBV-SI	BVSS-SI	BV-SI	RBVSS	RBV	BVSS	BV
Error 1 N(0,1)								
IMSE								
$\beta_0(t)$	0.021(0.007)	0.121(0.088)	0.026(0.024)	0.123(0.097)	0.024(0.005)	0.120(0.009)	0.022(0.004)	0.120(0.006)
$\beta_1(t)$	0.028(0.015)	0.118(0.035)	0.030(0.026)	0.120(0.065)	0.025(0.006)	0.116(0.013)	0.025(0.006)	0.118(0.014)
$\beta_2(t)$	0.031(0.017)	0.129(0.045)	0.038(0.018)	0.131(0.056)	0.029(0.008)	0.128(0.011)	0.040(0.006)	0.134(0.012)
$\beta_3(t)$	0.019(0.012)	0.065(0.041)	0.021(0.023)	0.071(0.036)	0.020(0.008)	0.068(0.020)	0.019(0.007)	0.074(0.015)
$\beta_4(t)$	0.026(0.012)	0.105(0.041)	0.024(0.023)	0.113(0.036)	0.024(0.008)	0.107(0.020)	0.021(0.007)	0.112(0.015)
MSE								
ζ_1	0.015(0.018)	0.031(0.038)	0.017(0.021)	0.033(0.032)	0.014(0.023)	0.029(0.041)	0.019(0.027)	0.031(0.031)
ζ_2	0.006(0.005)	0.028(0.022)	0.008(0.022)	0.031(0.028)	0.005(0.011)	0.027(0.032)	0.006(0.021)	0.029(0.041)
ζ_3	0.015(0.008)	0.034(0.030)	0.017(0.025)	0.036(0.032)	0.013(0.025)	0.035(0.019)	0.016(0.038)	0.034(0.045)
ζ_4	0.008(0.009)	0.022(0.022)	0.010(0.035)	0.024(0.024)	0.009(0.038)	0.021(0.038)	0.011(0.038)	0.022(0.076)
Total	0.131(0.040)	0.473(0.196)	0.132(0.037)	0.481(0.265)	0.154(0.026)	0.722(0.018)	0.157(0.020)	0.725(0.062)
Error 2 t(2)								
IMSE								
$\beta_0(t)$	0.104(0.430)	0.184(0.104)	0.262(0.287)	0.442(0.841)	0.103(0.007)	0.190(0.103)	0.258(1.431)	0.437(0.675)
$\beta_1(t)$	0.023(0.011)	0.203(0.112)	0.334(0.844)	0.631(0.988)	0.031(0.015)	0.201(0.144)	0.329(0.720)	0.629(0.421)
$\beta_2(t)$	0.091(0.041)	0.162(0.099)	0.321(0.397)	0.465(1.666)	0.093(0.020)	0.158(0.107)	0.319(0.444)	0.460(0.223)
$\beta_3(t)$	0.051(0.056)	0.211(0.121)	0.328(0.378)	0.454(0.718)	0.053(0.033)	0.206(0.146)	0.323(0.385)	0.450(0.321)
$\beta_4(t)$	0.038(0.056)	0.187(0.121)	0.211(0.378)	0.387(0.718)	0.032(0.033)	0.185(0.146)	0.209(0.385)	0.383(0.321)
MSE								
ζ_1	0.079(0.047)	0.112(0.063)	0.226(0.089)	0.261(0.537)	0.081(0.051)	0.113(0.051)	0.223(0.071)	0.259(0.189)
ζ_2	0.042(0.551)	0.093(0.009)	0.152(0.285)	0.259(0.672)	0.041(0.099)	0.091(0.011)	0.149(0.111)	0.257(0.231)
ζ_3	0.021(0.008)	0.051(0.061)	0.141(0.577)	0.391(0.673)	0.019(0.113)	0.049(0.098)	0.139(0.477)	0.389(0.067)
ζ_4	0.019(0.003)	0.048(0.037)	0.122(0.213)	0.287(0.167)	0.017(0.018)	0.045(0.111)	0.120(0.099)	0.282(0.123)
Total	0.268(0.104)	1.215(0.340)	1.325(1.456)	2.101(0.556)	0.389(0.022)	1.457(0.518)	1.528(0.889)	2.028(0.574)
Error 3 Lognormal(0,1)								
IMSE								
$\beta_0(t)$	0.071(0.040)	0.131(0.237)	0.161(0.598)	0.612(0.589)	0.069(0.036)	0.129(0.153)	0.159(0.587)	0.610(0.897)
$\beta_1(t)$	0.232(0.113)	0.383(0.256)	0.478(0.462)	0.793(0.319)	0.230(0.046)	0.381(0.284)	0.475(0.239)	0.800(0.495)
$\beta_2(t)$	0.181(0.073)	0.291(0.161)	0.338(0.441)	0.677(0.285)	0.179(0.130)	0.289(0.202)	0.336(0.520)	0.675(0.246)
$\beta_3(t)$	0.231(0.188)	0.273(0.262)	0.341(0.497)	0.621(0.381)	0.229(0.112)	0.270(0.270)	0.339(0.456)	0.619(0.321)
$\beta_4(t)$	0.158(0.188)	0.231(0.262)	0.299(0.497)	0.418(0.381)	0.156(0.112)	0.229(0.270)	0.297(0.456)	0.416(0.321)
MSE								
ζ_1	0.123(0.107)	0.191(0.099)	0.207(0.030)	0.328(0.541)	0.121(0.041)	0.189(0.048)	0.205(0.048)	0.325(0.041)
ζ_2	0.051(0.130)	0.141(0.117)	0.181(0.240)	0.403(0.270)	0.049(0.056)	0.139(0.056)	0.179(0.058)	0.400(0.039)
ζ_3	0.032(0.057)	0.129(0.146)	0.191(0.616)	0.343(0.671)	0.029(0.048)	0.127(0.133)	0.189(0.041)	0.341(0.541)
ζ_4	0.029(0.025)	0.075(0.075)	0.117(0.194)	0.238(0.378)	0.028(0.067)	0.073(0.111)	0.115(0.098)	0.235(0.126)
Total	0.413(0.295)	1.212(0.623)	1.403(0.323)	1.997(0.289)	0.727(0.456)	1.433(0.720)	1.528(0.456)	2.099(0.321)
Error 4 80%N(0,1)+ 20%N(0,3)								
IMSE								
$\beta_0(t)$	0.029(0.006)	0.055(0.009)	0.115(0.214)	0.779(0.943)	0.027(0.006)	0.052(0.009)	0.113(0.093)	0.776(2.967)
$\beta_1(t)$	0.023(0.026)	0.085(0.093)	0.121(0.235)	0.569(0.239)	0.021(0.007)	0.082(0.128)	0.117(0.415)	0.566(0.352)
$\beta_2(t)$	0.019(0.021)	0.060(0.023)	0.102(0.808)	0.224(0.891)	0.017(0.007)	0.057(0.046)	0.100(0.494)	0.222(0.331)
$\beta_3(t)$	0.021(0.047)	0.051(0.075)	0.077(0.351)	0.382(0.892)	0.018(0.008)	0.049(0.088)	0.080(0.962)	0.379(0.047)
$\beta_4(t)$	0.017(0.047)	0.044(0.075)	0.061(0.351)	0.291(0.892)	0.015(0.008)	0.041(0.088)	0.059(0.962)	0.289(0.147)
MSE								
ζ_1	0.013(0.026)	0.023(0.009)	0.051(0.032)	0.118(0.218)	0.012(0.031)	0.021(0.078)	0.049(0.054)	0.116(0.199)
ζ_2	0.009(0.004)	0.018(0.006)	0.057(0.290)	0.086(0.672)	0.007(0.011)	0.016(0.014)	0.057(0.117)	0.084(0.555)
ζ_3	0.008(0.003)	0.021(0.025)	0.079(0.632)	0.121(0.349)	0.006(0.034)	0.019(0.111)	0.071(0.113)	0.119(0.444)
ζ_4	0.002(0.005)	0.011(0.013)	0.058(0.581)	0.088(0.786)	0.002(0.123)	0.010(0.121)	0.060(0.444)	0.086(0.111)
Total	0.087(0.092)	0.116(0.182)	1.232(0.425)	1.446(0.521)	0.134(0.052)	1.248(0.416)	1.428(0.781)	1.899(0.951)
Error 5 Laplace(0,1)								
IMSE								
$\beta_0(t)$	0.039(0.023)	0.151(0.013)	0.232(0.019)	0.489(0.146)	0.038(0.010)	0.149(0.036)	0.230(0.101)	0.487(0.122)
$\beta_1(t)$	0.025(0.032)	0.122(0.217)	0.211(0.018)	0.356(0.513)	0.022(0.006)	0.121(0.018)	0.209(0.027)	0.354(0.061)
$\beta_2(t)$	0.021(0.021)	0.119(0.120)	0.151(0.019)	0.213(0.765)	0.044(0.007)	0.139(0.021)	0.195(0.030)	0.291(0.044)
$\beta_3(t)$	0.039(0.024)	0.299(0.341)	0.329(0.015)	0.499(0.654)	0.037(0.008)	0.297(0.025)	0.327(0.033)	0.496(0.065)
$\beta_4(t)$	0.018(0.024)	0.141(0.341)	0.199(0.015)	0.307(0.654)	0.016(0.008)	0.139(0.025)	0.195(0.033)	0.304(0.065)
MSE								
ζ_1	0.024(0.023)	0.038(0.051)	0.092(0.045)	0.152(0.123)	0.022(0.034)	0.036(0.045)	0.090(0.051)	0.150(0.115)
ζ_2	0.049(0.021)	0.081(0.025)	0.116(0.058)	0.189(0.325)	0.047(0.045)	0.079(0.056)	0.115(0.111)	0.190(0.216)
ζ_3	0.027(0.024)	0.061(0.029)	0.184(0.034)	0.283(0.256)	0.025(0.067)	0.059(0.047)	0.183(0.068)	0.280(0.113)
ζ_4	0.021(0.013)	0.091(0.034)	0.115(0.067)	0.231(0.111)	0.019(0.095)	0.089(0.051)	0.113(0.067)	0.229(0.117)
Total	0.165(0.028)	0.216(0.056)	1.291(0.124)	1.396(0.135)	0.261(0.025)	1.154(0.114)	1.446(0.268)	1.982(0.480)

Table B.9: *Simulation results in example 2 for $n = 400, m = 200$. mean(sd) of true positives (TP) and false positives (FP) under random intercept -slope model based on 100 replicates at 70th quantile with i.i.d. errors. True effects that are subject to selection: 4 varying effects, 4 constant effects.*

$n = 400 \quad m = 200$		Varying		Constant	
		TP	FP	TP	FP
Error 1 $N(0,1)$	RBVSS - SI	3.93(0.32)	1.36(0.21)	3.95(0.27)	1.82(0.14)
	RBV-SI	3.91(0.13)	3.13(0.32)	3.91(0.41)	3.83(0.89)
	BVSS -SI	3.92(0.13)	1.41(0.21)	3.94(0.31)	1.88(0.24)
	BV - SI	3.90(0.55)	3.18(0.25)	3.89(0.78)	3.89(0.67)
		TP		FP	
	RBVSS	3.93(0.32)		5.31(0.16)	
	RBV	3.91(0.28)		6.29(1.22)	
	BVSS	3.92(0.19)		5.35(0.20)	
Error 2 $t(2)$	BV	3.90(0.44)		6.33(1.17)	
	RBVSS - SI	3.93(0.14)	1.41(0.74)	3.89(0.24)	1.37(0.57)
	RBV - SI	3.79(1.79)	3.24(0.33)	3.22(0.24)	2.76(0.45)
	BVSS - SI	3.75(0.55)	3.03(0.16)	2.87(0.31)	3.02(0.41)
	BV - SI	3.24(0.55)	4.61(0.76)	2.13(1.64)	5.65(0.45)
		TP		FP	
	RBVSS	3.93(0.14)		5.36(1.47)	
	RBV	3.79(0.10)		6.54(0.99)	
Error 3 Lognormal(0,1)	BVSS	3.75(0.44)		6.99(0.20)	
	BV	3.24(0.41)		7.96(1.17)	
	RBVSS - SI	3.89(0.35)	1.15(0.52)	3.83(0.84)	1.62(0.29)
	RBV -SI	3.75(0.14)	2.55(0.25)	3.17(0.82)	3.93(0.15)
	BVSS - SI	3.80(0.89)	3.35(0.21)	2.95(0.49)	3.21(0.45)
	BV - SI	3.05(0.84)	4.69(0.26)	2.35(1.10)	3.86(0.22)
		TP		FP	
	RBVSS	3.89(0.10)		5.59(1.05)	
Error 4 80% $N(0,1)$ +20% $N(0,3)$	RBV	3.75(0.47)		6.86(1.22)	
	BVSS	3.80(0.59)		7.27(1.29)	
	BV	3.05(0.63)		8.06(0.94)	
	RBVSS - SI	3.92(0.32)	1.16(0.27)	3.98(0.55)	1.88(0.24)
	RBV - SI	3.81(0.12)	2.04(0.25)	3.11(0.53)	3.10(0.45)
	BVSS - SI	3.77(0.84)	2.89(0.25)	2.88(0.88)	2.34(0.15)
	BV - SI	3.13(0.55)	4.96(0.15)	2.22(1.14)	3.31(0.25)
		TP		FP	
Error 5 Laplace (0,1)	RBVSS	3.92(0.32)		4.98(0.67)	
	RBV	3.81(0.12)		6.51(0.98)	
	BVSS	3.77(1.08)		5.41(0.20)	
	BV	3.13(0.99)		6.99(1.97)	
	RBVSS - SI	3.93(0.31)	1.14(0.35)	3.94(0.24)	1.44(0.28)
	RBV - SI	3.81(0.45)	2.23(0.45)	3.31(0.31)	3.05(0.37)
	BVSS - SI	3.90(1.30)	2.81(0.45)	2.99(0.61)	2.81(1.34)
	BV - SI	3.29(0.71)	4.48(0.56)	2.38(0.47)	4.05(0.89)
		TP		FP	
	RBVSS	3.93(0.10)		4.66(0.90)	
	RBV	3.81(0.32)		5.72(0.51)	
	BVSS	3.90(0.37)		5.29(0.53)	
	BV	3.29(0.41)		7.09(1.67)	

Table B.10: *Estimation accuracy in example 2 for $n = 400, m = 200$ under random intercept-slope model, mean(sd) of IMSE, MSE, total squared errors for all estimates (Total) based on 100 replicates at 70th quantile with i.i.d. errors.*

	RBVSS-SI	RBV-SI	BVSS-SI	BV-SI	RBVSS	RBV	BVSS	BV
Error 1 N(0,1)								
IMSE								
$\beta_0(t)$	0.028(0.009)	0.119(0.021)	0.029(0.027)	0.121(0.014)	0.026(0.012)	0.118(0.011)	0.027(0.008)	0.119(0.011)
$\beta_1(t)$	0.023(0.018)	0.121(0.013)	0.025(0.084)	0.123(0.051)	0.021(0.017)	0.119(0.024)	0.022(0.008)	0.120(0.026)
$\beta_2(t)$	0.032(0.024)	0.127(0.032)	0.034(0.041)	0.129(0.089)	0.030(0.015)	0.125(0.019)	0.032(0.011)	0.128(0.023)
$\beta_3(t)$	0.021(0.017)	0.095(0.021)	0.022(0.035)	0.110(0.065)	0.019(0.052)	0.094(0.031)	0.020(0.010)	0.108(0.029)
$\beta_4(t)$	0.041(0.017)	0.117(0.021)	0.041(0.035)	0.116(0.065)	0.039(0.052)	0.115(0.031)	0.039(0.010)	0.114(0.029)
MSE								
ζ_1	0.012(0.014)	0.028(0.015)	0.015(0.064)	0.031(0.054)	0.011(0.067)	0.026(0.111)	0.014(0.051)	0.029(0.067)
ζ_2	0.005(0.007)	0.027(0.025)	0.006(0.078)	0.029(0.067)	0.008(0.061)	0.026(0.061)	0.004(0.048)	0.031(0.051)
ζ_3	0.011(0.005)	0.031(0.016)	0.013(0.088)	0.034(0.065)	0.009(0.078)	0.029(0.065)	0.011(0.078)	0.032(0.111)
ζ_4	0.005(0.006)	0.017(0.015)	0.006(0.014)	0.020(0.045)	0.004(0.075)	0.015(0.111)	0.004(0.057)	0.019(0.057)
Total	0.123(0.041)	0.387(0.024)	0.125(0.034)	0.390(0.015)	0.163(1.167)	0.822(0.142)	0.165(0.038)	0.825(0.029)
Error 2 t(2)								
IMSE								
$\beta_0(t)$	0.111(0.104)	0.177(0.077)	0.271(0.305)	0.461(0.841)	0.110(0.066)	0.175(0.120)	0.269(0.901)	0.459(0.272)
$\beta_1(t)$	0.027(0.011)	0.209(0.112)	0.358(0.841)	0.641(0.988)	0.025(0.074)	0.207(0.202)	0.354(0.366)	0.639(0.559)
$\beta_2(t)$	0.096(0.041)	0.158(0.099)	0.338(0.397)	0.486(0.666)	0.093(0.171)	0.154(0.143)	0.331(0.491)	0.483(0.301)
$\beta_3(t)$	0.063(0.056)	0.227(0.121)	0.416(0.358)	0.489(0.718)	0.061(0.140)	0.223(0.200)	0.414(0.368)	0.485(0.346)
$\beta_4(t)$	0.088(0.056)	0.188(0.121)	0.299(0.358)	0.401(0.718)	0.087(0.140)	0.186(0.200)	0.296(0.368)	0.398(0.346)
MSE								
ζ_1	0.097(0.047)	0.121(0.063)	0.233(0.079)	0.259(0.537)	0.095(0.054)	0.120(0.078)	0.231(0.078)	0.257(0.111)
ζ_2	0.051(0.009)	0.092(0.061)	0.165(0.285)	0.261(0.067)	0.048(0.115)	0.089(0.056)	0.163(0.156)	0.259(0.108)
ζ_3	0.025(0.008)	0.061(0.063)	0.151(0.577)	0.404(0.067)	0.023(0.127)	0.059(0.125)	0.149(0.211)	0.400(0.104)
ζ_4	0.017(0.003)	0.052(0.025)	0.133(0.213)	0.291(0.045)	0.015(0.045)	0.049(0.011)	0.130(0.191)	0.289(0.129)
Total	0.272(0.104)	1.231(0.340)	1.378(0.302)	2.137(0.256)	0.412(0.318)	1.557(0.658)	1.657(0.941)	2.215(0.232)
Error 3 Lognormal(0,1)								
IMSE								
$\beta_0(t)$	0.088(0.088)	0.142(0.373)	0.164(0.592)	0.603(0.927)	0.087(0.060)	0.140(0.173)	0.162(0.504)	0.602(0.044)
$\beta_1(t)$	0.241(0.160)	0.391(0.288)	0.486(0.579)	0.812(0.441)	0.240(0.079)	0.389(0.340)	0.484(0.301)	0.810(0.611)
$\beta_2(t)$	0.173(0.194)	0.303(0.320)	0.341(0.460)	0.655(0.819)	0.170(0.189)	0.300(0.198)	0.342(0.497)	0.651(0.250)
$\beta_3(t)$	0.238(0.308)	0.281(0.508)	0.339(0.508)	0.609(0.165)	0.234(0.108)	0.279(0.279)	0.337(0.573)	0.607(0.363)
$\beta_4(t)$	0.238(0.308)	0.281(0.508)	0.339(0.508)	0.609(0.165)	0.236(0.108)	0.279(0.279)	0.336(0.573)	0.603(0.363)
MSE								
ζ_1	0.125(0.071)	0.187(0.072)	0.205(0.031)	0.319(0.233)	0.123(0.091)	0.186(0.117)	0.202(0.041)	0.315(0.078)
ζ_2	0.057(0.237)	0.143(0.123)	0.192(0.579)	0.411(0.617)	0.054(0.123)	0.141(0.118)	0.190(0.378)	0.408(0.517)
ζ_3	0.034(0.171)	0.131(0.160)	0.220(0.599)	0.351(0.627)	0.031(0.168)	0.129(0.037)	0.218(0.341)	0.349(0.478)
ζ_4	0.030(0.143)	0.081(0.086)	0.116(0.173)	0.241(0.158)	0.029(0.123)	0.079(0.038)	0.114(0.154)	0.239(0.146)
Total	0.459(0.471)	1.228(0.637)	1.513(0.248)	2.001(0.438)	0.764(0.857)	1.399(0.773)	1.531(0.381)	2.735(0.218)
Error 4 80%N(0,1)+ 20%N(0,3)								
IMSE								
$\beta_0(t)$	0.041(0.033)	0.061(0.060)	0.123(0.112)	0.781(0.387)	0.042(0.009)	0.059(0.013)	0.121(0.967)	0.782(0.230)
$\beta_1(t)$	0.021(0.024)	0.075(0.039)	0.125(0.198)	0.589(2.527)	0.020(0.012)	0.073(0.075)	0.123(0.438)	0.586(0.554)
$\beta_2(t)$	0.017(0.034)	0.058(0.027)	0.105(0.606)	0.213(1.921)	0.015(0.011)	0.054(0.032)	0.103(0.877)	0.210(0.183)
$\beta_3(t)$	0.024(0.025)	0.058(0.043)	0.082(0.830)	0.394(0.843)	0.022(0.025)	0.058(0.085)	0.081(0.890)	0.390(0.592)
$\beta_4(t)$	0.019(0.025)	0.042(0.043)	0.071(0.830)	0.305(0.843)	0.021(0.025)	0.040(0.085)	0.069(0.890)	0.303(0.592)
MSE								
ζ_1	0.015(0.030)	0.025(0.013)	0.061(0.265)	0.123(0.919)	0.013(0.021)	0.023(0.011)	0.058(0.218)	0.121(0.717)
ζ_2	0.007(0.011)	0.016(0.020)	0.052(0.321)	0.078(0.692)	0.006(0.013)	0.015(0.031)	0.049(0.245)	0.075(0.451)
ζ_3	0.009(0.027)	0.024(0.038)	0.081(0.927)	0.123(0.131)	0.007(0.031)	0.022(0.039)	0.079(0.824)	0.124(0.319)
ζ_4	0.003(0.011)	0.014(0.025)	0.061(0.987)	0.094(0.278)	0.002(0.028)	0.012(0.034)	0.058(0.781)	0.092(0.312)
Total	0.091(0.048)	1.122(0.153)	1.282(0.278)	1.451(0.688)	0.128(0.443)	1.241(0.271)	1.438(0.533)	1.581(0.592)
Error 5 Laplace (0,1)								
IMSE								
$\beta_0(t)$	0.041(0.023)	0.157(0.033)	0.233(0.036)	0.493(0.049)	0.038(0.031)	0.155(0.055)	0.231(0.119)	0.490(0.143)
$\beta_1(t)$	0.028(0.045)	0.133(0.082)	0.222(0.059)	0.367(0.077)	0.026(0.012)	0.131(0.048)	0.220(0.032)	0.365(0.067)
$\beta_2(t)$	0.018(0.066)	0.117(0.049)	0.148(0.077)	0.209(0.063)	0.015(0.018)	0.114(0.032)	0.147(0.053)	0.207(0.178)
$\beta_3(t)$	0.037(0.064)	0.287(0.088)	0.318(0.035)	0.488(0.092)	0.034(0.015)	0.288(0.050)	0.316(0.038)	0.484(0.075)
$\beta_4(t)$	0.037(0.064)	0.287(0.088)	0.318(0.035)	0.488(0.092)	0.035(0.015)	0.285(0.050)	0.315(0.038)	0.486(0.075)
MSE								
ζ_1	0.022(0.077)	0.037(0.029)	0.089(0.109)	0.165(0.077)	0.021(0.074)	0.036(0.031)	0.092(0.104)	0.163(0.081)
ζ_2	0.051(0.048)	0.086(0.035)	0.121(0.070)	0.193(0.044)	0.049(0.049)	0.085(0.041)	0.119(0.067)	0.191(0.055)
ζ_3	0.025(0.014)	0.054(0.057)	0.182(0.063)	0.291(0.049)	0.023(0.019)	0.051(0.067)	0.179(0.078)	0.289(0.078)
ζ_4	0.019(0.029)	0.087(0.052)	0.114(0.106)	0.228(0.054)	0.018(0.056)	0.084(0.109)	0.112(0.112)	0.223(0.471)
Total	0.154(0.124)	1.222(0.234)	1.349(0.139)	1.426(0.579)	0.271(0.450)	1.430(0.179)	1.750(0.472)	1.907(0.556)

Table B.11: Empirical 95% coverage probabilities under simulated SNP data and random intercept - slope model with i.i.d. $t(2)$ error based on 200 replicates.

$t(2)$ error		RBVSS-SI	RBV-SI	RBVSS	RBV	BVSS-SI	BV-SI	BVSS	BV
$\theta = 0.3$	$\beta_0(t)$	0.885	0.905	0.876	0.899	0.865	0.886	0.878	0.871
	$\beta_1(t)$	0.815	0.130	0.823	0.141	0.920	0.151	0.915	0.163
	$\beta_2(t)$	0.945	0.920	0.952	0.916	0.845	0.899	0.853	0.905
	$\beta_3(t)$	0.925	0.775	0.919	0.800	0.625	0.791	0.675	0.638
	$\beta_4(t)$	0.930	0.875	0.934	0.887	0.555	0.865	0.600	0.605
	ζ_1	0.825	0.847	0.675	0.793	0.600	0.754	0.460	0.526
	ζ_2	0.825	0.795	0.705	0.635	0.610	0.615	0.555	0.426
	ζ_3	0.680	0.721	0.545	0.715	0.440	0.486	0.220	0.427
	ζ_4	0.885	0.905	0.875	0.840	0.765	0.785	0.730	0.610
$\theta = 0.5$	$\beta_0(t)$	0.912	0.945	0.900	0.939	0.925	0.898	0.900	0.903
	$\beta_1(t)$	0.871	0.136	0.860	0.128	0.853	0.143	0.861	0.152
	$\beta_2(t)$	0.958	0.905	0.965	0.913	0.850	0.861	0.849	0.863
	$\beta_3(t)$	0.955	0.785	0.950	0.791	0.855	0.779	0.863	0.784
	$\beta_4(t)$	0.950	0.875	0.950	0.881	0.489	0.504	0.492	0.495
	ζ_1	0.910	0.885	0.815	0.748	0.815	0.821	0.560	0.404
	ζ_2	0.860	0.640	0.835	0.512	0.835	0.567	0.545	0.430
	ζ_3	0.770	0.770	0.620	0.498	0.620	0.654	0.405	0.415
	ζ_4	0.910	0.850	0.935	0.814	0.875	0.819	0.760	0.624
$\theta = 0.7$	$\beta_0(t)$	0.890	0.910	0.902	0.917	0.871	0.931	0.912	0.937
	$\beta_1(t)$	0.880	0.110	0.879	0.117	0.931	0.095	0.896	0.111
	$\beta_2(t)$	0.935	0.930	0.941	0.928	0.851	0.835	0.877	0.841
	$\beta_3(t)$	0.910	0.755	0.923	0.768	0.711	0.745	0.842	0.759
	$\beta_4(t)$	0.935	0.910	0.928	0.933	0.589	0.888	0.863	0.897
	ζ_1	0.860	0.882	0.710	0.895	0.665	0.855	0.475	0.525
	ζ_2	0.810	0.765	0.765	0.665	0.680	0.592	0.584	0.413
	ζ_3	0.799	0.844	0.690	0.795	0.585	0.676	0.335	0.412
	ζ_4	0.905	0.955	0.885	0.830	0.775	0.835	0.753	0.539

Table B.12: Empirical 95% coverage probabilities under simulated SNP data and random intercept - slope model with i.i.d. $N(0,1)$ error based on 200 replicates.

$t(2)$ error		RBVSS-SI	RBV-SI	RBVSS	RBV	BVSS-SI	BV-SI	BVSS	BV
$\theta = 0.3$	$\beta_0(t)$	0.885	0.905	0.876	0.899	0.865	0.886	0.878	0.871
	$\beta_1(t)$	0.815	0.130	0.823	0.141	0.920	0.151	0.915	0.163
	$\beta_2(t)$	0.945	0.920	0.952	0.916	0.845	0.899	0.853	0.905
	$\beta_3(t)$	0.925	0.775	0.919	0.800	0.625	0.791	0.675	0.638
	$\beta_4(t)$	0.930	0.875	0.934	0.887	0.555	0.865	0.600	0.605
	ζ_1	0.825	0.847	0.675	0.793	0.600	0.754	0.460	0.526
	ζ_2	0.825	0.795	0.705	0.635	0.610	0.615	0.555	0.426
	ζ_3	0.680	0.721	0.545	0.715	0.440	0.486	0.220	0.427
	ζ_4	0.885	0.905	0.875	0.840	0.765	0.785	0.730	0.610
$\theta = 0.5$	$\beta_0(t)$	0.912	0.945	0.900	0.939	0.925	0.898	0.900	0.903
	$\beta_1(t)$	0.871	0.136	0.860	0.128	0.853	0.143	0.861	0.152
	$\beta_2(t)$	0.958	0.905	0.965	0.913	0.850	0.861	0.849	0.863
	$\beta_3(t)$	0.955	0.785	0.950	0.791	0.855	0.779	0.863	0.784
	$\beta_4(t)$	0.950	0.875	0.950	0.881	0.489	0.504	0.492	0.495
	ζ_1	0.910	0.885	0.815	0.748	0.815	0.821	0.560	0.404
	ζ_2	0.860	0.640	0.835	0.512	0.835	0.567	0.545	0.430
	ζ_3	0.770	0.770	0.620	0.498	0.620	0.654	0.405	0.415
	ζ_4	0.910	0.850	0.935	0.814	0.875	0.819	0.760	0.624
$\theta = 0.7$	$\beta_0(t)$	0.890	0.910	0.902	0.917	0.871	0.931	0.912	0.937
	$\beta_1(t)$	0.880	0.110	0.879	0.117	0.931	0.095	0.896	0.111
	$\beta_2(t)$	0.935	0.930	0.941	0.928	0.851	0.835	0.877	0.841
	$\beta_3(t)$	0.910	0.755	0.923	0.768	0.711	0.745	0.842	0.759
	$\beta_4(t)$	0.935	0.910	0.928	0.933	0.589	0.888	0.863	0.897
	ζ_1	0.860	0.882	0.710	0.895	0.665	0.855	0.475	0.525
	ζ_2	0.810	0.765	0.765	0.665	0.680	0.592	0.584	0.413
	ζ_3	0.799	0.844	0.690	0.795	0.585	0.676	0.335	0.412
	ζ_4	0.905	0.955	0.885	0.830	0.775	0.835	0.753	0.539

B.1.4 The estimated quantile varying coefficient functions

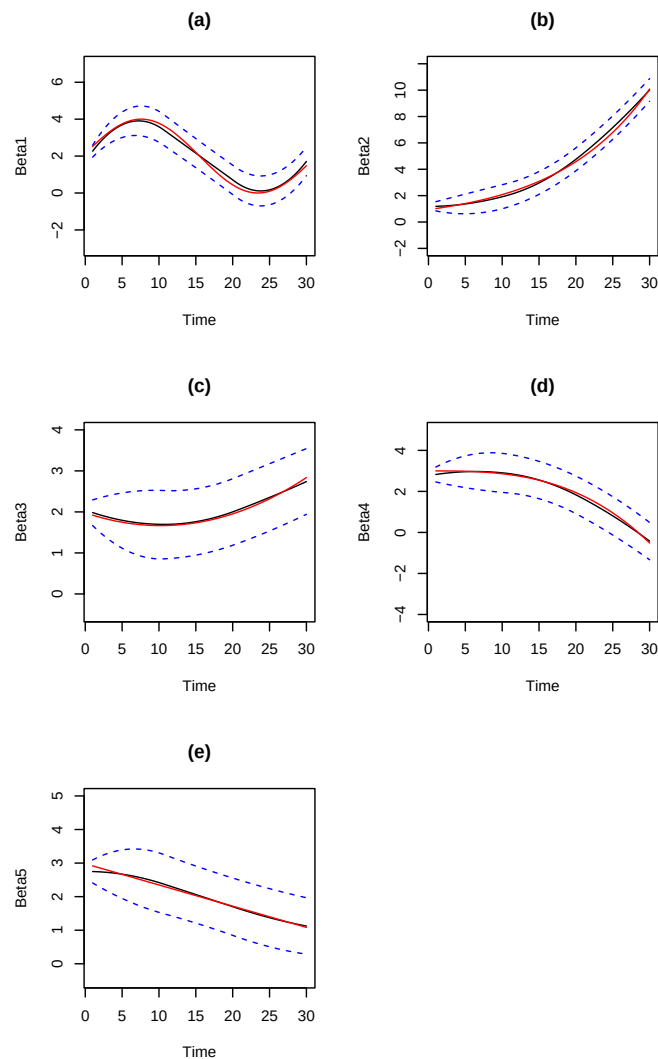


Figure B.1: Estimation of non-zero varying coefficients under the $t(2)$ error (Error 2) for the proposed method (RBVSS-SI) at 50% quantile level. Red line: true varying coefficients. Black line: posterior median estimates of varying coefficients from RBVSS-SI. Blue lines: 95% credible intervals for the estimated varying coefficients.

B.2 Posterior inference

B.2.1 RBVSS-SI

Priors

$$Y \propto (\tau^{-1})^{-\frac{nk}{2}} \left(\prod_{i=1}^n \prod_{j=1}^k \tilde{v}_{it_j} \right)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\tau^{-1}\xi_2^2} \sum_{i=1}^n \sum_{j=1}^k \left(\frac{Y_{it_j} - \mu_{it_j} - (\mathbf{b}_i^\top \otimes \mathbf{Z}_{it_j}^\top) \mathbf{J}_q \boldsymbol{\gamma} - \xi_1 \tilde{v}_{ij}}{\sqrt{\tilde{v}_{ij}}} \right)^2 \right\}$$

$$\tilde{v}_{it_j} | \tau \stackrel{ind}{\sim} \tau \exp(\tau \tilde{v}_{it_j})$$

$$\boldsymbol{\gamma}_0 \sim N_3(0, \Sigma_{\boldsymbol{\gamma}_0})$$

$$\boldsymbol{\alpha} \sim N_L(0, \Sigma_{\boldsymbol{\alpha}_0})$$

$$\beta_{l1} | \phi_{1l}, s_{1l} \stackrel{ind}{\sim} \phi_{1l} N(0, s_{1l}) + (1 - \phi_{1l}) \delta_0(\beta_{l1}) \quad l = 1, \dots, m$$

$$\phi_{1l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_1)$$

$$\beta_{l*} | \phi_{2l}, s_{2l} \stackrel{ind}{\sim} \phi_{2l} N_q(0, \text{diag}(s_{2l}, \dots, s_{2l})) + (1 - \phi_{2l}) \delta_0(\beta_{l*})$$

$$\phi_{2l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_2)$$

$$s_{11} | \eta_1^2 \stackrel{iid}{\sim} \exp\left(\frac{\eta_1^2}{2}\right) \quad l = 1, \dots, m$$

$$s_{2l} | \eta_2^2 \stackrel{ind}{\sim} \text{Gamma}\left(\frac{q+1}{2}, \frac{\eta_2^2}{2}\right)$$

$$\tau \sim \text{Gamma}(c, d)$$

$$\eta_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\eta_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\boldsymbol{\zeta}_i \sim N(0, \phi^2 \mathbf{I})$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\pi_1 \sim \text{Beta}(r_1, w_1)$$

$$\pi_2 \sim \text{Beta}(r_2, w_2)$$

Gibbs Sampler

Let $\mu_{ij(-\gamma_0)} = E(Y_{ij}) - X_{ij}^T \gamma_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of γ_0 .

The full conditional distribution of γ_0 is a multivariate normal distribution $N(\mu_{\gamma_0}, \Sigma_{\gamma_0})$ with mean

$$\mu_{\gamma_0} = \Sigma_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\gamma_0)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\gamma_0} = \left(\Sigma_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \right)^{-1}.$$

Let $\mu_{ij(-\alpha)} = E(Y_{ij}) - G_0(T_{ij})^T \alpha$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of α .

The full conditional distribution of α is normal distribution $N(\mu_{\alpha}, \Sigma_{\alpha})$ with mean

$$\mu_{\alpha} = \Sigma_{\alpha} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\alpha)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\alpha} = \left(\Sigma_{\alpha_0}^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} G_0(T_{ij}) G_0(T_{ij})^T \right)^{-1}.$$

Let $\mu_{ij(-\beta_{l1})} = E(Y_{ij}) - L_{ijl} \gamma_{l1}$ and $l_{1l} = \pi(\beta_{l1} = 0 | \text{rest})$, the conditional posterior distribution of the coefficient of lipid factor β_{l1} is a spike-and-slab distribution:

$$\beta_{l1} | \text{rest} \sim (1 - l_{1l}) N(\mu_{\beta_{l1}}, \Sigma_{\beta_{l1}}) + l_{1l} \delta_0(\beta_{l1}),$$

where

$$\begin{aligned}\mu_{\beta_{l1}} &= \Sigma_{\beta_{l1}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l1})} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\beta_{l1}} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl}^2 + \frac{1}{s_{1l}} \right)^{-1}.\end{aligned}$$

And

$$l_{1l} = \frac{\pi_1}{\pi_1 + (1 - \pi_1)(s_{1l})^{-\frac{1}{2}} (\Sigma_{\beta_{l1}})^{\frac{1}{2}} \exp\{\frac{1}{2} \Sigma_{\beta_{l1}} (\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l1})} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}))^2\}}.$$

The posterior distribution of β_{l1} is a mixture of a normal distribution and a point mass at 0. At each iteration of MCMC, β_{l1} is drawn from $N(\mu_{\beta_{l1}}, \Sigma_{\beta_{l1}})$ with probability $(1 - l_{1l})$ and is set to 0 with probability l_{1l} . If β_{l1} is set to 0, then we have $\phi_{1l} = 0$. Otherwise, $\phi_{1l} = 1$.

The posterior distribution of the interaction between the l th lipid and treatment factors $\beta_{l*}(l = 1, \dots, m)$ is also a spike-and-slab distribution. Denote $\mu_{ij(-\beta_{l*})} = E(Y_{ij}) - W_{ijl}^T \beta_{l*}$ and $l_{2l} = \pi(\beta_{l*} = \mathbf{0} | \text{rest})$, the conditional posterior distribution of β_{l*} is a multivariate spike-and-slab distribution :

$$\beta_{l*} | \text{rest} \sim (1 - l_{2l}) N(\mu_{\beta_{l*}}, \Sigma_{\beta_{l*}}) + l_{2l} \delta_0(\beta_{l*}),$$

where

$$\begin{aligned}\mu_{\beta_{l*}} &= \Sigma_{\beta_{l*}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l*})} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\beta_{l*}} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijl} W_{ijl}^T + \frac{1}{s_{2l}} \mathbf{I}_q \right)^{-1}.\end{aligned}$$

And

$$l_{2l} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)(s_{2l})^{-\frac{q}{2}} |\Sigma_{\beta_{l*}}|^{\frac{1}{2}} \exp\{\frac{1}{2} \|\Sigma_{\beta_{l*}}^{-\frac{1}{2}} \sum_{i=1}^n \sum_{j=1}^k \tau \xi_2^{-2} \tilde{v}_{ij}^{-1} W_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l*})} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij})\|_2^2\}}.$$

The posterior distribution of β_{l*} is a mixture of a multivariate normal distribution and a point mass at 0. At each iteration of MCMC, β_{l*} is drawn from $N(\mu_{\beta_{l*}}, \Sigma_{\beta_{l*}})$ with probability

$(1 - l_{2l})$ and is set to 0 with probability l_{2l} . If β_{l*} is set to $\mathbf{0}$, then we have $\phi_{2l} = 0$. Otherwise, $\phi_{2l} = 1$.

The full conditional posterior distribution of $(s_{1l})^{-1}$ is:

$$(s_{1l})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(1, \frac{\eta_1^2}{2}) & \text{if } \beta_{l1} = 0 \\ \text{Inverse-Gaussian}(\eta_1^2, \sqrt{\frac{\eta_1^2}{\beta_{l1}^2}}) & \text{if } \beta_{l1} \neq 0 \end{cases}$$

Similarly, The full conditional posterior distribution of $(s_{2l})^{-1}$ is:

$$(s_{2l})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{q+1}{2}, \frac{\eta_2^2}{2}) & \text{if } \beta_{l*} = \mathbf{0} \\ \text{Inverse-Gaussian}(\eta_2^2, \sqrt{\frac{\eta_2^2}{\|\beta_{l*}\|_2^2}}) & \text{if } \beta_{l*} \neq \mathbf{0} \end{cases}$$

The full conditional posterior distribution of η_1^2 :

$$\eta_1^2 | \text{rest} \sim \text{Gamma}(a_1 + p, b_1 + \frac{\sum_{l=1}^m s_{1l}}{2}).$$

Likewise, the full conditional posterior distribution of η_2^2 :

$$\eta_2^2 | \text{rest} \sim \text{Gamma}(a_2 + \frac{m(q+1)}{2}, b_2 + \frac{\sum_{l=1}^m s_{2l}}{2}).$$

The full conditional posterior distribution of π_1 :

$$\pi_1 | \text{rest} \sim \text{Beta}(r_1 + \sum_{l=1}^m \mathbf{I}_{\{\beta_{l1}=0\}}, w_1 + \sum_{l=1}^m \mathbf{I}_{\{\beta_{l1} \neq 0\}}).$$

Similarly, the full conditional posterior distribution of π_2 :

$$\pi_2 | \text{rest} \sim \text{Beta}(r_2 + \sum_{l=1}^m \mathbf{I}_{\{\beta_{l*}=\mathbf{0}\}}, w_2 + \sum_{l=1}^m \mathbf{I}_{\{\beta_{l*} \neq \mathbf{0}\}}).$$

The full conditional posterior distribution of $\tilde{v}_{ij}$:

$$\tilde{v}_{ij}|\text{rest} \sim \text{Inverse-Gaussian}\left(\sqrt{\frac{\xi_1^2 + 2\xi_2^2}{(Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i)^2}}, \frac{\tau \xi_1^2}{\xi_2^2} + 2\tau\right).$$

The full conditional posterior distribution of τ :

$$\tau|\text{rest} \sim \text{Gamma}\left(c + \frac{3nk}{2}, d + \sum_{i=1}^n \sum_{j=1}^k (\tilde{v}_{ij} + (2\xi_2^2)^{-1} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij})^2)\right).$$

The full conditional posterior distribution of ζ_i :

$$\zeta_i|\text{rest} \sim N(\mu_{\zeta_i}, \Sigma_{\zeta_i})$$

where

$$\begin{aligned} \mu_{\zeta_i} &= \frac{\tau}{\xi_2^2} \Sigma_{\zeta_i} \sum_{j=1}^k \frac{Z_{ij}(Y_{ij} - \mu_{ij} - \xi_1 \tilde{v}_{ij})}{\tilde{v}_{ij}} \\ \Sigma_{\zeta_i} &= \left(\frac{\tau}{\xi_2^2} \sum_{j=1}^k \frac{Z_{ij} Z_{ij}^T}{\tilde{v}_{ij}} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

Last, the full conditional posterior distribution of ϕ^2 :

$$\phi^2|\text{rest} \sim \text{Inverse-Gamma}\left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \zeta_i^T \zeta_i\right)$$

B.2.2 RBVSS

Priors

$$Y \propto (\tau^{-1})^{-\frac{nk}{2}} \left(\prod_{i=1}^n \prod_{j=1}^k \tilde{v}_{ij} \right)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\tau^{-1}\xi_2^2} \sum_{i=1}^n \sum_{j=1}^k \left(\frac{Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}}{\sqrt{\tilde{v}_{ij}}} \right)^2 \right\}$$

$$\tilde{v}_{ij} | \tau \stackrel{ind}{\sim} \tau \exp(\tau \tilde{v}_{ij})$$

$$\gamma_0 \sim N_3(0, \Sigma_{\gamma_0})$$

$$\alpha \sim N_L(0, \Sigma_{\alpha_0})$$

$$\beta_l | \phi_{1l}, s_{1l} \stackrel{ind}{\sim} \phi_{1l} N_L(0, \text{diag}(s_{1l}, \dots, s_{1l})) + (1 - \phi_{1l}) \delta_0(\beta_l)$$

$$\phi_{1l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_1)$$

$$s_{1l} | \eta_1^2 \stackrel{ind}{\sim} \text{Gamma}\left(\frac{L+1}{2}, \frac{\eta_1^2}{2}\right)$$

$$\tau \sim \text{Gamma}(c, d)$$

$$\eta_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\zeta_i \sim N(0, \phi^2 \mathbf{I})$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\pi_1 \sim \text{Beta}(r_1, w_1)$$

Gibbs Sampler

Let $\mu_{ij(-\gamma_0)} = E(Y_{ij}) - X_{ij}^T \gamma_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of γ_0 .

The full conditional distribution of γ_0 is a multivariate normal distribution $N(\mu_{\gamma_0}, \Sigma_{\gamma_0})$ with mean

$$\mu_{\gamma_0} = \Sigma_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\gamma_0)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\gamma_0} = \left(\Sigma_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \right)^{-1}.$$

Let $\mu_{ij(-\alpha)} = E(Y_{ij}) - G_0(T_{ij})^T \alpha$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of α .

The full conditional distribution of α is normal distribution $N(\mu_\alpha, \Sigma_\alpha)$ with mean

$$\mu_\alpha = \Sigma_\alpha \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\alpha)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_\alpha = \left(\Sigma_{\alpha_0}^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} G_0(T_{ij}) G_0(T_{ij})^T \right)^{-1}.$$

β_l ($l = 1, \dots, m$) is also a spike-and-slab distribution. Denote $\mu_{ij(-\beta_l)} = E(Y_{ij}) - W_{ijl}^T \beta_l$ and $l_{1l} = \pi(\beta_l = \mathbf{0} | \text{rest})$, the conditional posterior distribution of β_l is a multivariate spike-and-slab distribution :

$$\beta_l | \text{rest} \sim (1 - l_{1l}) N(\mu_{\beta_l}, \Sigma_{\beta_l}) + l_{1l} \delta_0(\beta_l),$$

where

$$\begin{aligned} \mu_{\beta_l} &= \Sigma_{\beta_l} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijl} (Y_{ij} - \mu_{ij(-\beta_l)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}), \\ \Sigma_{\beta_l} &= \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijl} W_{ijl}^T + \frac{1}{s_{1l}} \mathbf{I}_L \right)^{-1}. \end{aligned}$$

And

$$l_{1l} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)(s_{1l})^{-\frac{L}{2}} |\Sigma_{\beta_l}|^{\frac{1}{2}} \exp\left\{\frac{1}{2} \left\| \Sigma_{\beta_l}^{-\frac{1}{2}} \sum_{i=1}^n \sum_{j=1}^k \tau \xi_2^{-2} \tilde{v}_{ij}^{-1} W_{ijl} (Y_{ij} - \mu_{ij(-\beta_l)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}) \right\|_2^2\right\}}.$$

The posterior distribution of β_l is a mixture of a multivariate normal distribution and a point mass at 0. At each iteration of MCMC, β_l is drawn from $N(\mu_{\beta_l}, \Sigma_{\beta_l})$ with probability $(1 - l_{1l})$ and is set to 0 with probability l_{1l} . If β_l is set to $\mathbf{0}$, then we have $\phi_{1l} = 0$. Otherwise,

$$\phi_{1l} = 1.$$

The full conditional posterior distribution of $(s_{2l})^{-1}$ is:

$$(s_{1l})^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{L+1}{2}, \frac{\eta_1^2}{2}) & \text{if } \beta_l = \mathbf{0} \\ \text{Inverse-Gaussian}(\eta_1^2, \sqrt{\frac{\eta_1^2}{\|\beta_l\|_2^2}}) & \text{if } \beta_{l*} \neq \mathbf{0} \end{cases}$$

The full conditional posterior distribution of η_2^2 :

$$\eta_1^2 | \text{rest} \sim \text{Gamma}(a_1 + \frac{m(L+1)}{2}, b_1 + \frac{\sum_{l=1}^m s_{1l}}{2}).$$

The full conditional posterior distribution of π_2 :

$$\pi_1 | \text{rest} \sim \text{Beta}(r_1 + \sum_{l=1}^m \mathbf{I}_{\{\beta_l = \mathbf{0}\}}, w_1 + \sum_{l=1}^m \mathbf{I}_{\{\beta_l \neq \mathbf{0}\}}).$$

The full conditional posterior distribution of $\tilde{v}_{ij}$:

$$\tilde{v}_{ij} | \text{rest} \sim \text{Inverse-Gaussian}(\sqrt{\frac{\xi_1^2 + 2\xi_2^2}{(Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i)^2}}, \frac{\tau \xi_1^2}{\xi_2^2} + 2\tau).$$

The full conditional posterior distribution of τ :

$$\tau | \text{rest} \sim \text{Gamma}(c + \frac{3nk}{2}, d + \sum_{i=1}^n \sum_{j=1}^k (\tilde{v}_{ij} + (2\xi_2^2)^{-1} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij})^2).$$

The full conditional posterior distribution of ζ_i :

$$\zeta_i | \text{rest} \sim N(\mu_{\zeta_i}, \Sigma_{\zeta_i})$$

where

$$\mu_{\zeta_i} = \frac{\tau}{\xi_2^2} \Sigma_{\zeta_i} \sum_{j=1}^k \frac{Z_{ij}(Y_{ij} - \mu_{ij} - \xi_1 \tilde{v}_{ij})}{\tilde{v}_{ij}}$$

$$\Sigma_{\zeta_i} = \left(\frac{\tau}{\xi_2^2} \sum_{j=1}^k \frac{Z_{ij} Z_{ij}^T}{\tilde{v}_{ij}} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1}$$

Last, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma} \left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \zeta_i^T \zeta_i \right)$$

B.2.3 RBV-SI

Priors

$$Y \propto (\tau^{-1})^{-\frac{nk}{2}} \left(\prod_{i=1}^n \prod_{j=1}^k \tilde{v}_{ij} \right)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\tau^{-1}\xi_2^2} \sum_{i=1}^n \sum_{j=1}^k \left(\frac{Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}}{\sqrt{\tilde{v}_{ij}}} \right)^2 \right\}$$

$$\tilde{v}_{ij} | \tau \stackrel{ind}{\sim} \tau \exp(\tau \tilde{v}_{ij})$$

$$\gamma_0 \sim N_3(0, \Sigma_{\gamma_0})$$

$$\alpha \sim N_L(0, \Sigma_{\alpha_0})$$

$$\beta_{l1} | s_{1l} \stackrel{ind}{\sim} N(0, s_{1l}) \quad l = 1, \dots, m$$

$$\beta_{l*} | s_{2l} \stackrel{ind}{\sim} N(0, \text{diag}(s_{2l}, \dots, s_{2l}))$$

$$s_{1l} | \eta_1^2 \stackrel{iid}{\sim} \exp\left(\frac{\eta_1^2}{2}\right) \quad l = 1, \dots, m$$

$$s_{2l} | \eta_2^2 \stackrel{ind}{\sim} \text{Gamma}\left(\frac{q+1}{2}, \frac{\eta_2^2}{2}\right)$$

$$\tau \sim \text{Gamma}(c, d)$$

$$\eta_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\eta_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\zeta_i \sim N(0, \phi^2 \mathbf{I})$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

Gibbs Sampler

Let $\mu_{ij(-\gamma_0)} = E(Y_{ij}) - X_{ij}^T \gamma_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of γ_0 .

The full conditional distribution of γ_0 is a multivariate normal distribution $N(\mu_{\gamma_0}, \Sigma_{\gamma_0})$ with mean

$$\mu_{\gamma_0} = \Sigma_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\gamma_0)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\gamma_0} = \left(\Sigma_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \right)^{-1}.$$

Let $\mu_{ij(-\alpha)} = E(Y_{ij}) - G_0(T_{ij})^T \alpha$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of α .

The full conditional distribution of α is normal distribution $N(\mu_\alpha, \Sigma_\alpha)$ with mean

$$\mu_\alpha = \Sigma_\alpha \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\alpha)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_\alpha = \left(\Sigma_{\alpha_0}^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} G_0(T_{ij}) G_0(T_{ij})^T \right)^{-1}.$$

Let $\mu_{ij(-\beta_{l1})} = E(Y_{ij}) - L_{ijl} \beta_{l1}$, then the full conditional distribution of β_{l1} is a normal distribution with mean

$$\mu_{\beta_{l1}} = \Sigma_{\beta_{l1}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l1})} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

variance

$$\Sigma_{\beta_{l1}} = \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} L_{ijl}^2 + \frac{1}{s_{1l}} \right)^{-1}.$$

Denote $\mu_{ij(-\beta_{l*})} = E(Y_{ij}) - W_{ijl}^T \beta_{l*}$, then the full conditional distribution of β_{l*} is a multivariate normal distribution with mean

$$\mu_{\beta_{l*}} = \Sigma_{\beta_{l*}} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l*})} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\beta_{l*}} = \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijl} W_{ijl}^T + \frac{1}{s_{2l}} \mathbf{I}_q \right)^{-1}.$$

The full conditional posterior distribution of $(s_{1l})^{-1}$ is:

$$(s_{1l})^{-1}|\text{rest} \sim \text{Inverse-Gaussian}(\eta_1^2, \sqrt{\frac{\eta_1^2}{\beta_{l1}^2}})$$

Similarly, The full conditional posterior distribution of $(s_{2l})^{-1}$ is:

$$(s_{2l})^{-1}|\text{rest} \sim \text{Inverse-Gaussian}(\eta_2^2, \sqrt{\frac{\eta_2^2}{\|\beta_{l*}\|_2^2}})$$

The full conditional posterior distribution of η_1^2 :

$$\eta_1^2|\text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m s_{1l}}{2}).$$

Likewise, the full conditional posterior distribution of η_2^2 :

$$\eta_2^2|\text{rest} \sim \text{Gamma}(a_2 + \frac{m(q+1)}{2}, b_2 + \frac{\sum_{l=1}^m s_{2l}}{2}).$$

The full conditional posterior distribution of $\tilde{v}_{ij}$:

$$\tilde{v}_{ij}|\text{rest} \sim \text{Inverse-Gaussian}(\sqrt{\frac{\xi_1^2 + 2\xi_2^2}{(Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i)^2}}, \frac{\tau \xi_1^2}{\xi_2^2} + 2\tau).$$

The full conditional posterior distribution of τ :

$$\tau|\text{rest} \sim \text{Gamma}(c + \frac{3nk}{2}, d + \sum_{i=1}^n \sum_{j=1}^k (\tilde{v}_{ij} + (2\xi_2^2)^{-1} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij})^2).$$

The full conditional posterior distribution of ζ_i :

$$\zeta_i|\text{rest} \sim N(\mu_{\zeta_i}, \Sigma_{\zeta_i})$$

where

$$\mu_{\zeta_i} = \frac{\tau}{\xi_2^2} \Sigma_{\zeta_i} \sum_{j=1}^k \frac{Z_{ij}(Y_{ij} - \mu_{ij} - \xi_1 \tilde{v}_{ij})}{\tilde{v}_{ij}}$$

$$\Sigma_{\zeta_i} = \left(\frac{\tau}{\xi_2^2} \sum_{j=1}^k \frac{Z_{ij} Z_{ij}^T}{\tilde{v}_{ij}} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1}$$

Last, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma} \left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \zeta_i^T \zeta_i \right)$$

B.2.4 RBV

Priors

$$Y \propto (\tau^{-1})^{-\frac{nk}{2}} \left(\prod_{i=1}^n \prod_{j=1}^k \tilde{v}_{ij} \right)^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2\tau^{-1}\xi_2^2} \sum_{i=1}^n \sum_{j=1}^k \left(\frac{Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}}{\sqrt{\tilde{v}_{ij}}} \right)^2 \right\}$$

$$\tilde{v}_{ij} | \tau \stackrel{\text{ind}}{\sim} \tau \exp(\tau \tilde{v}_{ij})$$

$$\gamma_0 \sim \text{N}_3(0, \Sigma_{\gamma_0})$$

$$\alpha \sim \text{N}_L(0, \Sigma_{\alpha_0})$$

$$\beta_l | s_{1l} \stackrel{\text{ind}}{\sim} \text{N}(0, \text{diag}(s_{1l}, \dots, s_{1l})) \quad l = 1, \dots, m$$

$$s_{1l} | \eta_1^2 \stackrel{\text{ind}}{\sim} \text{Gamma} \left(\frac{L+1}{2}, \frac{\eta_1^2}{2} \right) \quad l = 1, \dots, m$$

$$\tau \sim \text{Gamma}(c, d)$$

$$\eta_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\zeta_i \sim \text{N}(0, \phi^2 \mathbf{I})$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

Gibbs Sampler

Let $\mu_{ij(-\gamma_0)} = E(Y_{ij}) - X_{ij}^T \gamma_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of γ_0 .

The full conditional distribution of γ_0 is a multivariate normal distribution $N(\mu_{\gamma_0}, \Sigma_{\gamma_0})$ with mean

$$\mu_{\gamma_0} = \Sigma_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\gamma_0)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\gamma_0} = \left(\Sigma_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \right)^{-1}.$$

Let $\mu_{ij(-\alpha)} = E(Y_{ij}) - G_0(T_{ij})^T \alpha$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of α .

The full conditional distribution of α is normal distribution $N(\mu_{\alpha}, \Sigma_{\alpha})$ with mean

$$\mu_{\alpha} = \Sigma_{\alpha} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\alpha)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\alpha} = \left(\Sigma_{\alpha_0}^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} G_0(T_{ij}) G_0(T_{ij})^T \right)^{-1}.$$

Denote $\mu_{ij(-\beta_l)} = E(Y_{ij}) - W_{ijl}^T \beta_l$, then the full conditional distribution of β_l is a multivariate normal distribution with mean

$$\mu_{\beta_l} = \Sigma_{\beta_l} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijl} (Y_{ij} - \mu_{ij(-\beta_l)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\beta_l} = \left(\tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} W_{ijl} W_{ijl}^T + \frac{1}{s_{1l}} \mathbf{I}_L \right)^{-1}.$$

The full conditional posterior distribution of $(s_{1l})^{-1}$ is:

$$(s_{1l})^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\eta_1^2, \sqrt{\frac{\eta_1^2}{\|\beta_l\|_2^2}})$$

The full conditional posterior distribution of η_1^2 :

$$\eta_1^2 | \text{rest} \sim \text{Gamma}(a_1 + \frac{m(L+1)}{2}, b_1 + \frac{\sum_{l=1}^m s_{1l}}{2}).$$

The full conditional posterior distribution of $\tilde{v}_{ij}$:

$$\tilde{v}_{ij} | \text{rest} \sim \text{Inverse-Gaussian}(\sqrt{\frac{\xi_1^2 + 2\xi_2^2}{(Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i)^2}}, \frac{\tau \xi_1^2}{\xi_2^2} + 2\tau).$$

The full conditional posterior distribution of τ :

$$\tau | \text{rest} \sim \text{Gamma}(c + \frac{3nk}{2}, d + \sum_{i=1}^n \sum_{j=1}^k (\tilde{v}_{ij} + (2\xi_2^2)^{-1} \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij})^2).$$

The full conditional posterior distribution of ζ_i :

$$\zeta_i | \text{rest} \sim N(\mu_{\zeta_i}, \Sigma_{\zeta_i})$$

where

$$\begin{aligned} \mu_{\zeta_i} &= \frac{\tau}{\xi_2^2} \Sigma_{\zeta_i} \sum_{j=1}^k \frac{Z_{ij}(Y_{ij} - \mu_{ij} - \xi_1 \tilde{v}_{ij})}{\tilde{v}_{ij}} \\ \Sigma_{\zeta_i} &= \left(\frac{\tau}{\xi_2^2} \sum_{j=1}^k \frac{Z_{ij} Z_{ij}^T}{\tilde{v}_{ij}} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

Last, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma}(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \zeta_i^T \zeta_i)$$

B.2.5 BVSS-SI

Priors

$$Y \propto (\sigma^2)^{-\frac{nk}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^k \left(Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i \right)^2 \right\}$$

$$\gamma_0 \sim N_3(0, \Sigma_{\gamma_0})$$

$$\alpha \sim N_L(0, \Sigma_{\alpha_0})$$

$$\beta_{l1} | \phi_{1l}, \tau_{1l}^2, \sigma^2 \stackrel{ind}{\sim} \phi_{1l} N(0, \sigma^2 \tau_{1l}^2) + (1 - \phi_{1l}) \delta_0(\beta_{l1}) \quad l = 1, \dots, m$$

$$\phi_{1l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_1)$$

$$\beta_{l*} | \phi_{2l}, \tau_{2l}^2, \sigma^2 \stackrel{ind}{\sim} \phi_{2l} N(0, \text{diag}(\sigma^2 \tau_{2l}^2, \dots, \sigma^2 \tau_{2l}^2)) + (1 - \phi_{2l}) \delta_0(\beta_{l*})$$

$$\phi_{2l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_2)$$

$$\tau_{1l}^2 | \lambda_1^2 \stackrel{iid}{\sim} \exp\left(\frac{\lambda_1^2}{2}\right) \quad l = 1, \dots, m$$

$$\tau_{2l}^2 | \lambda_2^2 \stackrel{ind}{\sim} \text{Gamma}\left(\frac{q+1}{2}, \frac{q\lambda_2^2}{2}\right)$$

$$\lambda_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\lambda_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\zeta_i \sim N(0, \phi^2 \mathbf{I})$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\pi_1 \sim \text{Beta}(r_1, w_1)$$

$$\pi_2 \sim \text{Beta}(r_2, w_2)$$

$$\sigma^2 \propto \frac{1}{\sigma^2}$$

Gibbs Sampler

Let $\mu_{ij(-\gamma_0)} = E(Y_{ij}) - X_{ij}^T \gamma_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of γ_0 .

The full conditional distribution of γ_0 is a multivariate normal distribution $N(\mu_{\gamma_0}, \Sigma_{\gamma_0})$ with

mean

$$\mu_{\gamma_0} = \Sigma_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\gamma_0)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\gamma_0} = \left(\Sigma_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \right)^{-1}.$$

Denote $\mu_{ij(-\alpha)} = E(Y_{ij}) - G_0(T_{ij})^T \alpha$, ($i = 1, \dots, n, j = 1, \dots, k$), then the full conditional distribution of α is multivariate normal distribution $MVN(\mu_\alpha, \Sigma_\alpha)$ with mean

$$\mu_\alpha = \Sigma_\alpha (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij(-\alpha)} - Z_{ij}^T \zeta_i) G_0(T_{ij}),$$

and variance

$$\Sigma_\alpha = \left(\Sigma_{\alpha_0}^{-1} + (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k G_0(T_{ij}) G_0(T_{ij})^T \right)^{-1}.$$

Let $\mu_{ij(-\beta_{l1})} = E(Y_{ij}) - L_{ijl} \beta_{l1}$ and $l_{1l} = \pi(\beta_{l1} = 0 | \text{rest})$, the conditional posterior distribution of the coefficient of lipid factor β_{l1} is a spike-and-slab distribution:

$$\beta_{l1} | \text{rest} \sim (1 - l_{1l}) N(\mu_{\beta_{l1}}, \Sigma_{\beta_{l1}}) + l_{1l} \delta_0(\beta_{l1}),$$

where

$$\begin{aligned} \mu_{\beta_{l1}} &= \Sigma_{\beta_{l1}} (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k L_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l1})} - Z_{ij}^T \zeta_i), \\ \Sigma_{\beta_{l1}} &= \sigma^2 \left(\sum_{i=1}^n \sum_{j=1}^k L_{ijl}^2 + \frac{1}{\tau_{1l}^2} \right)^{-1}. \end{aligned}$$

Where

$$l_{1l} = \frac{\pi_1}{\pi_1 + (1 - \pi_1) (\tau_{1l}^2)^{-\frac{1}{2}} (\Sigma_{\beta_{l1}})^{\frac{1}{2}} \exp\left\{ \frac{1}{2\sigma^2} \Sigma_{\beta_{l1}} \left(\sum_{i=1}^n \sum_{j=1}^k L_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l1})} - Z_{ij}^T \zeta_i) \right)^2 \right\}}.$$

The posterior distribution of β_{l1} is a mixture of a normal distribution and a point mass at 0. At each iteration of MCMC, β_{l1} is drawn from $N(\mu_{\beta_{l1}}, \Sigma_{\beta_{l1}})$ with probability $1 - l_{1l}$ and

is set to 0 with probability l_{1l} . If β_{l1} is set to 0, then we have $\phi_{1l} = 0$. Otherwise, $\phi_{1l} = 1$.

The posterior distribution of the interaction between the l th lipid and treatment factors $\beta_{l*}(l = 1, \dots, m)$ is also a spike-and-slab distribution. Denote $\mu_{ij(-\beta_{l*})} = E(Y_{ij}) - W_{ijl}^T \beta_{l*}$ and $l_{2l} = \pi(\beta_{l*} = \mathbf{0} | \text{rest})$, the conditional posterior distribution of β_{l*} is a multivariate spike-and-slab distribution :

$$\beta_{l*} | \text{rest} \sim (1 - l_{2l}) N(\mu_{\beta_{l*}}, \sigma^2 \Sigma_{\beta_{l*}}) + l_{2l} \delta_0(\beta_{l*}),$$

where

$$\begin{aligned} \mu_{\beta_{l*}} &= \Sigma_{\beta_{l*}} \sum_{i=1}^n \sum_{j=1}^k W_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l*})} - Z_{ij}^T \zeta_i), \\ \Sigma_{\beta_{l*}} &= \left(\sum_{i=1}^n \sum_{j=1}^k W_{ijl} W_{ijl}^T + \frac{1}{\tau_{2l}^2} \mathbf{I}_q \right)^{-1}. \end{aligned}$$

And

$$l_{2l} = \frac{\pi_2}{\pi_2 + (1 - \pi_2)(\tau_{2l}^2)^{-\frac{q}{2}} |\Sigma_{\beta_{l*}}|^{\frac{1}{2}} \exp\left\{ \frac{1}{2\sigma^2} \left\| \Sigma_{\beta_{l*}}^{\frac{1}{2}} \sum_{i=1}^n \sum_{j=1}^k W_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l*})} - Z_{ij}^T \zeta_i) \right\|_2^2 \right\}}.$$

The posterior distribution of β_{l*} is a mixture of a multivariate normal distribution and a point mass at 0. At each iteration of MCMC, β_{l*} is drawn from $N(\mu_{\beta_{l*}}, \sigma^2 \Sigma_{\beta_{l*}})$ with probability $(1 - l_{2l})$ and is set to 0 with probability l_{2l} . If β_{l*} is set to $\mathbf{0}$, then we have $\phi_{2l} = 0$. Otherwise, $\phi_{2l} = 1$.

The full conditional posterior distribution of $(\tau_{1l}^2)^{-1}$ is:

$$(\tau_{1l}^2)^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(1, \frac{\lambda_1^2}{2}) & \text{if } \beta_{l1} = 0 \\ \text{Inverse-Gaussian}(\lambda_1^2, \sqrt{\frac{\sigma^2 \lambda_1^2}{\beta_{l1}^2}}) & \text{if } \beta_{l1} \neq 0 \end{cases}$$

Similarly, the full conditional posterior distribution of $(\tau_{2l}^2)^{-1}$ is:

$$(\tau_{2l}^2)^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{q+1}{2}, \frac{q\lambda_2^2}{2}) & \text{if } \beta_{l*} = \mathbf{0} \\ \text{Inverse-Gaussian}(\lambda_2^2, \sqrt{\frac{\sigma^2 \lambda_2^2}{\|\beta_{l*}\|_2^2}}) & \text{if } \beta_{l*} \neq \mathbf{0} \end{cases}$$

The full conditional posterior distribution of λ_1^2 :

$$\lambda_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m \tau_{1l}^2}{2}).$$

Likewise, the full conditional posterior distribution of λ_2^2 :

$$\lambda_2^2 | \text{rest} \sim \text{Gamma}(a_2 + \frac{m(q+1)}{2}, b_2 + \frac{q \sum_{l=1}^m \tau_{2l}^2}{2}).$$

The full conditional posterior distribution of π_1 :

$$\pi_1 | \text{rest} \sim \text{Beta}(r_1 + \sum_{l=1}^m \mathbf{I}_{\{\beta_{l1}=0\}}, w_1 + \sum_{l=1}^m \mathbf{I}_{\{\beta_{l1} \neq 0\}}).$$

Similarly, the full conditional posterior distribution of π_2 :

$$\pi_2 | \text{rest} \sim \text{Beta}(r_2 + \sum_{l=1}^m \mathbf{I}_{\{\beta_{l*}=\mathbf{0}\}}, w_2 + \sum_{l=1}^m \mathbf{I}_{\{\beta_{l*} \neq \mathbf{0}\}}).$$

The full conditional posterior distribution of σ^2 :

$$\sigma^2 | \text{rest} \sim \text{Inv-Gamma}(\mu_{\sigma^2}, \Sigma_{\sigma^2})$$

with shape

$$\mu_{\sigma^2} = \frac{nk + \sum_{l=1}^m \mathbf{I}_{\{\beta_{l1} \neq 0\}} + q \sum_{l=1}^m \mathbf{I}_{\{\beta_{l*} \neq \mathbf{0}\}}}{2}$$

and scale

$$\Sigma_{\sigma^2} = \frac{\sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i)^2 + \sum_{l=1}^m (\tau_{2l}^2)^{-1} \beta_{l1}^2 + \sum_{l=1}^m (\tau_{2l}^2)^{-1} \beta_{l*}^T \beta_{l*}}{2}$$

The full conditional posterior distribution of ζ_i :

$$\zeta_i | \text{rest} \sim N(\mu_{\zeta_i}, \Sigma_{\zeta_i})$$

where

$$\begin{aligned} \mu_{\zeta_i} &= (\sigma^2)^{-1} \Sigma_{\zeta_i} \sum_{j=1}^k Z_{ij} (Y_{ij} - \mu_{ij}) \\ \Sigma_{\zeta_i} &= \left(\sum_{j=1}^k \frac{Z_{ij} Z_{ij}^T}{\sigma^2} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

Last, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma} \left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \zeta_i^T \zeta_i \right)$$

B.2.6 BVSS

Priors

$$Y \propto (\sigma^2)^{-\frac{nk}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^k \left(Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i \right)^2 \right\}$$

$$\gamma_0 \sim N_3(0, \Sigma_{\gamma_0})$$

$$\alpha \sim N_L(0, \Sigma_{\alpha_0})$$

$$\beta_l | \phi_{1l}, \tau_{1l}^2, \sigma^2 \stackrel{ind}{\sim} \phi_{1l} N(0, \text{diag}(\sigma^2 \tau_{1l}^2, \dots, \sigma^2 \tau_{1l}^2)) + (1 - \phi_{1l}) \delta_0(\beta_l)$$

$$\phi_{1l} \stackrel{ind}{\sim} \text{Bernoulli}(\pi_2)$$

$$\tau_{1l}^2 | \lambda_1^2 \stackrel{ind}{\sim} \text{Gamma}\left(\frac{L+1}{2}, \frac{L\lambda_1^2}{2}\right)$$

$$\lambda_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\zeta_i \sim N(0, \phi^2 \mathbf{I})$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\pi_1 \sim \text{Beta}(r_1, w_1)$$

$$\sigma^2 \propto \frac{1}{\sigma^2}$$

Gibbs Sampler

Let $\mu_{ij(-\gamma_0)} = E(Y_{ij}) - X_{ij}^T \gamma_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of γ_0 .

The full conditional distribution of γ_0 is a multivariate normal distribution $N(\mu_{\gamma_0}, \Sigma_{\gamma_0})$ with mean

$$\mu_{\gamma_0} = \Sigma_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\gamma_0)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\gamma_0} = \left(\Sigma_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \right)^{-1}.$$

Denote $\mu_{ij(-\alpha)} = E(Y_{ij}) - G_0(T_{ij})^T \alpha$, ($i = 1, \dots, n, j = 1, \dots, k$), then the full conditional distribution of α is multivariate normal distribution $MVN(\mu_\alpha, \Sigma_\alpha)$ with mean

$$\mu_\alpha = \Sigma_\alpha (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij(-\alpha)} - Z_{ij}^T \zeta_i) G_0(T_{ij}),$$

and variance

$$\Sigma_\alpha = \left(\Sigma_{\alpha_0}^{-1} + (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k G_0(T_{ij}) G_0(T_{ij})^T \right)^{-1}.$$

The posterior distribution of β_l ($l = 1, \dots, m$) is also a spike-and-slab distribution. Denote $\mu_{ij(-\beta_l)} = E(Y_{ij}) - W_{ijl}^T \beta_l$ and $l_{1l} = \pi(\beta_l = \mathbf{0} | \text{rest})$, the conditional posterior distribution of β_l is a multivariate spike-and-slab distribution :

$$\beta_l | \text{rest} \sim (1 - l_{1l}) N(\mu_{\beta_l}, \sigma^2 \Sigma_{\beta_l}) + l_{1l} \delta_0(\beta_l),$$

where

$$\begin{aligned} \mu_{\beta_l} &= \Sigma_{\beta_l} \sum_{i=1}^n \sum_{j=1}^k W_{ijl} (Y_{ij} - \mu_{ij(-\beta_l)} - Z_{ij}^T \zeta_i), \\ \Sigma_{\beta_l} &= \left(\sum_{i=1}^n \sum_{j=1}^k W_{ijl} W_{ijl}^T + \frac{1}{\tau_{1l}^2} \mathbf{I}_q \right)^{-1}. \end{aligned}$$

And

$$l_{1l} = \frac{\pi_1}{\pi_1 + (1 - \pi_1) (\tau_{1l}^2)^{-\frac{L}{2}} |\Sigma_{\beta_l}|^{\frac{1}{2}} \exp\left\{ \frac{1}{2\sigma^2} \left\| \Sigma_{\beta_l}^{-\frac{1}{2}} \sum_{i=1}^n \sum_{j=1}^k W_{ijl} (Y_{ij} - \mu_{ij(-\beta_l)} - Z_{ij}^T \zeta_i) \right\|_2^2 \right\}}. \quad (\text{B.1})$$

The posterior distribution of β_l is a mixture of a multivariate normal distribution and a point mass at 0. At each iteration of MCMC, β_l is drawn from $N(\mu_{\beta_l}, \sigma^2 \Sigma_{\beta_l})$ with probability $(1 - l_{1l})$ and is set to 0 with probability l_{1l} . If β_l is set to $\mathbf{0}$, then we have $\phi_{1l} = 0$. Otherwise,

$\phi_{1l} = 1$. The full conditional posterior distribution of $(\tau_{1l}^2)^{-1}$ is:

$$(\tau_{1l}^2)^{-1} | \text{rest} \sim \begin{cases} \text{Inverse-Gamma}(\frac{L+1}{2}, \frac{L\lambda_1^2}{2}) & \text{if } \beta_l = \mathbf{0} \\ \text{Inverse-Gaussian}(\lambda_1^2, \sqrt{\frac{\sigma^2 \lambda_1^2}{\|\beta_l\|_2^2}}) & \text{if } \beta_l \neq \mathbf{0} \end{cases}$$

The full conditional posterior distribution of λ_1^2 :

$$\lambda_1^2 | \text{rest} \sim \text{Gamma}(a_1 + \frac{m(L+1)}{2}, b_1 + \frac{L \sum_{l=1}^m \tau_{1l}^2}{2}).$$

The full conditional posterior distribution of π_1 :

$$\pi_1 | \text{rest} \sim \text{Beta}(r_1 + \sum_{l=1}^m \mathbf{I}_{\{\beta_l = \mathbf{0}\}}, w_1 + \sum_{l=1}^m \mathbf{I}_{\{\beta_l \neq \mathbf{0}\}}).$$

The full conditional posterior distribution of σ^2 :

$$\sigma^2 | \text{rest} \sim \text{Inv-Gamma}(\mu_{\sigma^2}, \Sigma_{\sigma^2})$$

with shape

$$\mu_{\sigma^2} = \frac{nk + L \sum_{l=1}^m \mathbf{I}_{\{\beta_l \neq \mathbf{0}\}}}{2}$$

and scale

$$\Sigma_{\sigma^2} = \frac{\sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i)^2 + \sum_{l=1}^m (\tau_{1l}^2)^{-1} \beta_l^T \beta_l}{2}$$

The full conditional posterior distribution of ζ_i :

$$\zeta_i | \text{rest} \sim N(\mu_{\zeta_i}, \Sigma_{\zeta_i})$$

where

$$\begin{aligned}\mu_{\zeta_i} &= (\sigma^2)^{-1} \Sigma_{\zeta_i} \sum_{j=1}^k Z_{ij} (Y_{ij} - \mu_{ij}) \\ \Sigma_{\zeta_i} &= \left(\sum_{j=1}^k \frac{Z_{ij} Z_{ij}^T}{\sigma^2} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1}\end{aligned}$$

Last, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma} \left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \zeta_i^T \zeta_i \right)$$

B.2.7 BV-SI

Priors

$$Y \propto (\sigma^2)^{-\frac{nk}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^k \left(Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i \right)^2 \right\}$$

$$\gamma_0 \sim \text{N}_3(0, \Sigma_{\gamma_0})$$

$$\alpha \sim \text{N}_L(0, \Sigma_{\alpha_0})$$

$$\beta_{1l} | \tau_{1l}^2, \sigma^2 \stackrel{\text{ind}}{\sim} \text{N}(0, \sigma^2 \tau_{1l}^2) \quad l = 1, \dots, m$$

$$\beta_{l*} | \tau_{2l}^2, \sigma^2 \stackrel{\text{ind}}{\sim} \text{N}(0, \text{diag}(\sigma^2 \tau_{2l}^2, \dots, \sigma^2 \tau_{2l}^2))$$

$$\tau_{1l}^2 | \lambda_1^2 \stackrel{\text{iid}}{\sim} \exp\left(\frac{\lambda_1^2}{2}\right) \quad l = 1, \dots, m$$

$$\tau_{2l}^2 | \lambda_2^2 \stackrel{\text{ind}}{\sim} \text{Gamma}\left(\frac{q+1}{2}, \frac{L\lambda_2^2}{2}\right)$$

$$\lambda_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\lambda_2^2 \sim \text{Gamma}(a_2, b_2)$$

$$\zeta_i \sim \text{N}(0, \phi^2 \mathbf{I})$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\sigma^2 \propto \frac{1}{\sigma^2}$$

Gibbs Sampler

Let $\mu_{ij(-\gamma_0)} = E(Y_{ij}) - X_{ij}^T \gamma_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of γ_0 .

The full conditional distribution of γ_0 is a multivariate normal distribution $N(\mu_{\gamma_0}, \Sigma_{\gamma_0})$ with mean

$$\mu_{\gamma_0} = \Sigma_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\gamma_0)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\gamma_0} = \left(\Sigma_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \right)^{-1}.$$

Denote $\mu_{ij(-\alpha)} = E(Y_{ij}) - G_0(T_{ij})^T \alpha$, ($i = 1, \dots, n, j = 1, \dots, k$), then the full conditional distribution of α is multivariate normal distribution $MVN(\mu_{\alpha}, \Sigma_{\alpha})$ with mean

$$\mu_{\alpha} = \Sigma_{\alpha} (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij(-\alpha)} - Z_{ij}^T \zeta_i) G_0(T_{ij}),$$

and variance

$$\Sigma_{\alpha} = \left(\Sigma_{\alpha_0}^{-1} + (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k G_0(T_{ij}) G_0(T_{ij})^T \right)^{-1}.$$

Let $\mu_{ij(-\beta_{l1})} = E(Y_{ij}) - L_{ijl} \beta_{l1}$ and $l_{1l} = \pi(\beta_{l1} = 0 | \text{rest})$, the conditional posterior distribution of the coefficient of lipid factor β_{l1} is a normal distribution: with mean

$$\mu_{\beta_{l1}} = \Sigma_{\beta_{l1}} (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k L_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l1})} - Z_{ij}^T \zeta_i),$$

and variance

$$\Sigma_{\beta_{l1}} = \sigma^2 \left(\sum_{i=1}^n \sum_{j=1}^k L_{ijl}^2 + \frac{1}{\tau_{1l}^2} \right)^{-1}.$$

Denote $\mu_{ij(-\beta_{l*})} = E(Y_{ij}) - W_{ijl}^T \beta_{l*}$. The posterior distribution of the interaction between the l th lipid and treatment factors $\beta_{l*} (l = 1, \dots, m)$ is a multivariate normal distribution with

mean

$$\mu_{\beta_{l*}} = \Sigma_{\beta_{l*}} \sum_{i=1}^n \sum_{j=1}^k W_{ijl} (Y_{ij} - \mu_{ij(-\beta_{l*})} - Z_{ij}^T \zeta_i),$$

And variance

$$\Sigma_{\beta_{l*}} = \left(\sum_{i=1}^n \sum_{j=1}^k W_{ijl} W_{ijl}^T + \frac{1}{\tau_{2l}^2} \mathbf{I}_q \right)^{-1}.$$

The full conditional posterior distribution of $(\tau_{1l}^2)^{-1}$ is:

$$(\tau_{1l}^2)^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\lambda_1^2, \sqrt{\frac{\sigma^2 \lambda_1^2}{\beta_{1l}^2}})$$

Similarly, The full conditional posterior distribution of $(\tau_{2l}^2)^{-1}$ is:

$$(\tau_{2l}^2)^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\lambda_2^2, \sqrt{\frac{\sigma^2 \lambda_2^2}{\|\beta_{l*}\|_2^2}})$$

The full conditional posterior distribution of λ_1^2 :

$$\lambda_1^2 | \text{rest} \sim \text{Gamma}(a_1 + m, b_1 + \frac{\sum_{l=1}^m \tau_{1l}^2}{2}).$$

Likewise, the full conditional posterior distribution of λ_2^2 :

$$\lambda_2^2 | \text{rest} \sim \text{Gamma}(a_2 + \frac{m(q+1)}{2}, b_2 + \frac{q \sum_{l=1}^m \tau_{2l}^2}{2}).$$

The full conditional posterior distribution of σ^2 :

$$\sigma^2 | \text{rest} \sim \text{Inv-Gamma}(\mu_{\sigma^2}, \Sigma_{\sigma^2})$$

with shape

$$\mu_{\sigma^2} = \frac{nk + m + qm}{2}$$

and scale

$$\Sigma_{\sigma^2} = \frac{\sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i)^2 + \sum_{l=1}^m (\tau_{1l}^2)^{-1} \beta_{1l}^2 + \sum_{l=1}^m (\tau_{2l}^2)^{-1} \beta_{l*}^T \beta_{l*}}{2}$$

The full conditional posterior distribution of ζ_i :

$$\zeta_i | \text{rest} \sim N(\mu_{\zeta_i}, \Sigma_{\zeta_i})$$

where

$$\begin{aligned} \mu_{\zeta_i} &= (\sigma^2)^{-1} \Sigma_{\zeta_i} \sum_{j=1}^k Z_{ij} (Y_{ij} - \mu_{ij}) \\ \Sigma_{\zeta_i} &= \left(\sum_{j=1}^k \frac{Z_{ij} Z_{ij}^T}{\sigma^2} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

Last, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma} \left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \zeta_i^T \zeta_i \right)$$

B.2.8 BV

Priors

$$Y \propto (\sigma^2)^{-\frac{nk}{2}} \exp \left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^n \sum_{j=1}^k \left(Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i \right)^2 \right\}$$

$$\gamma_0 \sim N_3(0, \Sigma_{\gamma_0})$$

$$\alpha \sim N_L(0, \Sigma_{\alpha_0})$$

$$\beta_l | \tau_{1l}^2, \sigma^2 \stackrel{ind}{\sim} N(0, \text{diag}(\sigma^2 \tau_{1l}^2, \dots, \sigma^2 \tau_{1l}^2))$$

$$\tau_{1l}^2 | \lambda_1^2 \stackrel{ind}{\sim} \text{Gamma}\left(\frac{L+1}{2}, \frac{L\lambda_1^2}{2}\right)$$

$$\lambda_1^2 \sim \text{Gamma}(a_1, b_1)$$

$$\zeta_i \sim N(0, \phi^2 \mathbf{I})$$

$$\phi^2 \sim \text{Inverse-Gamma}(e, f)$$

$$\sigma^2 \propto \frac{1}{\sigma^2}$$

Gibbs Sampler

Let $\mu_{ij(-\gamma_0)} = E(Y_{ij}) - X_{ij}^T \gamma_0$, ($i = 1, \dots, n, j = 1, \dots, k$), representing the mean effect without the contribution of γ_0 .

The full conditional distribution of γ_0 is a multivariate normal distribution $N(\mu_{\gamma_0}, \Sigma_{\gamma_0})$ with mean

$$\mu_{\gamma_0} = \Sigma_{\gamma_0} \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} (Y_{ij} - \mu_{ij(-\gamma_0)} - Z_{ij}^T \zeta_i - \xi_1 \tilde{v}_{ij}),$$

and variance

$$\Sigma_{\gamma_0} = \left(\Sigma_0^{-1} + \tau \xi_2^{-2} \sum_{i=1}^n \sum_{j=1}^k \tilde{v}_{ij}^{-1} \right)^{-1}.$$

Denote $\mu_{ij(-\alpha)} = E(Y_{ij}) - G_0(T_{ij})^T \alpha$, ($i = 1, \dots, n, j = 1, \dots, k$), then the full conditional

distribution of α is multivariate normal distribution $\text{MVN}(\mu_\alpha, \Sigma_\alpha)$ with mean

$$\mu_\alpha = \Sigma_\alpha (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij(-\alpha)} - Z_{ij}^T \zeta_i) G_0(T_{ij}),$$

and variance

$$\Sigma_\alpha = \left(\Sigma_{\alpha_0}^{-1} + (\sigma^2)^{-1} \sum_{i=1}^n \sum_{j=1}^k G_0(T_{ij}) G_0(T_{ij})^T \right)^{-1}.$$

Denote $\mu_{ij(-\beta_l)} = E(Y_{ij}) - W_{ijl}^T \beta_l$. The posterior distribution of $\beta_l (l = 1, \dots, m)$ is a multivariate normal distribution with mean

$$\mu_{\beta_l} = \Sigma_{\beta_l} \sum_{i=1}^n \sum_{j=1}^k W_{ijl} (Y_{ij} - \mu_{ij(-\beta_l)} - Z_{ij}^T \zeta_i),$$

And variance

$$\Sigma_{\beta_l} = \left(\sum_{i=1}^n \sum_{j=1}^k W_{ijl} W_{ijl}^T + \frac{1}{\tau_{1l}^2} \mathbf{I}_L \right)^{-1}.$$

The full conditional posterior distribution of $(\tau_{1l}^2)^{-1}$ is:

$$(\tau_{1l}^2)^{-1} | \text{rest} \sim \text{Inverse-Gaussian}(\lambda_1^2, \sqrt{\frac{\sigma^2 \lambda_1^2}{\|\beta_l\|_2^2}})$$

The full conditional posterior distribution of λ_1^2 :

$$\lambda_1^2 | \text{rest} \sim \text{Gamma}(a_1 + \frac{m(L+1)}{2}, b_1 + \frac{L \sum_{l=1}^m \tau_{1l}^2}{2}).$$

The full conditional posterior distribution of σ^2 :

$$\sigma^2 | \text{rest} \sim \text{Inv-Gamma}(\mu_{\sigma^2}, \Sigma_{\sigma^2})$$

with shape

$$\mu_{\sigma^2} = \frac{nk + m + Lm}{2}$$

and scale

$$\Sigma_{\sigma^2} = \frac{\sum_{i=1}^n \sum_{j=1}^k (Y_{ij} - \mu_{ij} - Z_{ij}^T \zeta_i)^2 + \sum_{l=1}^m (\tau_{1l}^2)^{-1} \beta_l^T \beta_l}{2}$$

The full conditional posterior distribution of ζ_i :

$$\zeta_i | \text{rest} \sim N(\mu_{\zeta_i}, \Sigma_{\zeta_i})$$

where

$$\begin{aligned} \mu_{\zeta_i} &= (\sigma^2)^{-1} \Sigma_{\zeta_i} \sum_{j=1}^k Z_{ij} (Y_{ij} - \mu_{ij}) \\ \Sigma_{\zeta_i} &= \left(\sum_{j=1}^k \frac{Z_{ij} Z_{ij}^T}{\sigma^2} + \frac{1}{\phi^2} \mathbf{I} \right)^{-1} \end{aligned}$$

Last, the full conditional posterior distribution of ϕ^2 :

$$\phi^2 | \text{rest} \sim \text{Inverse-Gamma} \left(\frac{nh}{2} + e, f + \frac{1}{2} \sum_{i=1}^n \zeta_i^T \zeta_i \right)$$