SUPPLEMENTAL MATERIALS OF "AN EMPIRICAL BAYES REGRESSION FOR MULTI-TISSUE GENE EXPRESSION PREDICTION"

BY FEI XUE[1] AND HONGZHE LI[2]

[1]Purdue University, feixue@purdue.edu

[2]University of Pennsylvania, hongzhe@pennmedicine.upenn.edu

SECTION A: EM ALGORITHM FOR MISSING DATA

In this section, we introduce the EM algorithm with missing data. The marginal distribution of $\boldsymbol{Y}_{\text{obs}}^{(t)}$ and $\boldsymbol{W}^{(t)}$ is

$$p(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2) = \tau_1 g_1(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \boldsymbol{\beta}, \eta, \sigma^2) + (1-\tau_1) g_0(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \sigma^2),$$

where $g_0(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \sigma^2) = p(\boldsymbol{W}^{(t)})\psi(\boldsymbol{Y}_{\text{obs}}^{(t)}; \boldsymbol{0}, \sigma^2 \boldsymbol{I}_{n_t})$, $g_1(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \boldsymbol{\beta}, \eta, \sigma^2) = p(\boldsymbol{W}^{(t)})\psi(\boldsymbol{Y}_{\text{obs}}^{(t)}; \boldsymbol{X}_t\boldsymbol{\beta}, \sigma^2\boldsymbol{I}_{n_t} + \eta \boldsymbol{H}_t)$, and $p(\boldsymbol{W}^{(t)})$ is the probability mass function of $\boldsymbol{W}^{(t)}$. Then the marginal distribution of all $\boldsymbol{Y}_{\text{obs}}^{(1)}, \dots, \boldsymbol{Y}_{\text{obs}}^{(m)}, \boldsymbol{W}^{(1)}, \dots, \boldsymbol{W}^{(m)}$ is

$$p(\boldsymbol{Y}_{\text{obs}}^{(1)}, \boldsymbol{W}^{(1)}, \dots, \boldsymbol{Y}_{\text{obs}}^{(m)}, \boldsymbol{W}^{(m)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2) = \prod_{t=1}^{m} p(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2), \tag{11}$$

and the complete-data likelihood is

$$\begin{aligned} & p(\boldsymbol{Y}_{\text{obs}}^{(1)}, \boldsymbol{W}^{(1)}, \dots, \boldsymbol{Y}_{\text{obs}}^{(m)}, \boldsymbol{W}^{(m)}, I^{(1)}, \dots, I^{(m)}; \tau_1, \tau_0, \boldsymbol{\beta}, \eta, \sigma^2) \\ = & \prod_{t=1}^{m}\prod_{s=0}^{1} \left\{ \tau_s g_s(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \boldsymbol{\beta}, \eta, \sigma^2) \right\}^{\mathbb{I}(I^{(t)}=s)}, \end{aligned}$$

where $g_0(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \boldsymbol{\beta}, \eta, \sigma^2) = g_0(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \sigma^2)$.

We use the expectation–maximization (EM) algorithm similarly as in Section 3. In the expectation step, given our current estimate of the parameters $\boldsymbol{\theta}_{(k)}$ at the k-th iteration, we first calculate the posterior distribution of $I^{(t)}$:

$$\begin{aligned} T_{s,(k)}^{(t)} &= P(I^{(t)} = s \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}, \boldsymbol{\theta}_{(k)}) \\ &= \frac{\tau_{s,(k)} g_s(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \boldsymbol{\beta}_{(k)}, \eta_{(k)}, \sigma_{(k)})}{\tau_{1,(k)} g_1 \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \boldsymbol{\beta}_{(k)}, \eta_{(k)}, \sigma_{(k)}) + \tau_{0,(k)} g_0(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \sigma_{(k)})}, \end{aligned}$$

and then calculate the expectation of the complete-data log-likelihood under the posterior distribution of $\{I^{(1)}, \dots, I^{(m)}\}$:

$$\begin{aligned} & Q(\boldsymbol{\theta} \mid \boldsymbol{\theta}_{(k)}) \\ = & E_{I^{(1)},\dots,I^{(m)} \mid \boldsymbol{Y}_{\text{obs}}^{(1)}, \boldsymbol{W}^{(1)}, \dots, \boldsymbol{Y}_{\text{obs}}^{(m)}, \boldsymbol{W}^{(m)}, \boldsymbol{\theta}_{(k)}} \Big[\log p(\boldsymbol{Y}_{\text{obs}}^{(1)}, \boldsymbol{W}^{(1)}, \dots, \boldsymbol{Y}_{\text{obs}}^{(m)}, \boldsymbol{W}^{(m)}, I^{(1)}, \dots, I^{(m)}; \\ & \tau_1, \tau_0, \boldsymbol{\beta}, \eta, \sigma^2) \Big] \\ = & \sum_{t=1}^{m}\sum_{s=0}^{1} P(I^{(t)} = s \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}, \boldsymbol{\theta}_{(k)}) \left\{ \log \tau_s + \log g_s(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \boldsymbol{\beta}, \eta, \sigma^2) \right\} \end{aligned}$$

$$= \sum_{t=1}^{m} \sum_{s=0}^{1} T_{s,(k)}^{(t)} \left\{ \log \tau_s + \log g_s(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \boldsymbol{\beta}, \eta, \sigma^2) \right\}.$$

In the maximization step, we maximize $Q(\boldsymbol{\theta} \mid \boldsymbol{\theta}_{(k)})$ to determine the next estimate for all the parameters. The maximizer consists of

$$\tau_{0,(k+1)} = \frac{\sum_{t=1}^{m} T_{0,(k)}^{(t)}}{\sum_{t=1}^{m} \left\{ T_{0,(k)}^{(t)} + T_{1,(k)}^{(t)} \right\}}, \quad \tau_{1,(k+1)} = \frac{\sum_{t=1}^{m} T_{1,(k)}^{(t)}}{\sum_{t=1}^{m} \left\{ T_{0,(k)}^{(t)} + T_{1,(k)}^{(t)} \right\}},$$

$$\boldsymbol{\beta}_{(k+1)} = \left[\sum_{t=1}^{m} T_{1,(k)}^{(t)} \left\{ \boldsymbol{X}_t^T (\sigma^2 \boldsymbol{I} + \eta \boldsymbol{H}_t)^{-1} \boldsymbol{X}_t \right\} \right]^{-1} \sum_{t=1}^{m} T_{1,(k)}^{(t)} \left\{ \boldsymbol{X}_t^T (\sigma^2 \boldsymbol{I} + \eta \boldsymbol{H}_t)^{-1} \boldsymbol{Y}_{\text{obs}}^{(t)} \right\},$$

$$\eta_{(k+1)} = \eta_{(k)} \left[\sum_{t=1}^{m} T_{1,(k)}^{(t)} \text{tr} \left\{ (\sigma^2 \boldsymbol{I} + \eta_{(k)} \boldsymbol{H}_t)^{-1} \boldsymbol{H}_t \right\} \right]^{-1}$$

$$\cdot \sum_{t=1}^{m} T_{1,(k)}^{(t)} \left[(\boldsymbol{Y}_{\text{obs}}^{(t)} - \boldsymbol{X}_t \boldsymbol{\beta})^T (\sigma^2 \boldsymbol{I} + \eta_{(k)} \boldsymbol{H}_t)^{-1} \boldsymbol{H}_t (\sigma^2 \boldsymbol{I} + \eta_{(k)} \boldsymbol{H}_t)^{-1} (\boldsymbol{Y}_{\text{obs}}^{(t)} - \boldsymbol{X}_t \boldsymbol{\beta}) \right],$$

and

$$\sigma_{(k+1)}^2 = \frac{\sigma_{(k)}^4 \sum_{t=1}^{m} T_{1,(k)}^{(t)} (\boldsymbol{Y}_{\text{obs}}^{(t)} - \boldsymbol{X}_t \boldsymbol{\beta}_{(k)})^T (\sigma_{(k)}^2 \boldsymbol{I} + \eta_{(k)} \boldsymbol{H}_t)^{-2} (\boldsymbol{Y}_{\text{obs}}^{(t)} - \boldsymbol{X}_t \boldsymbol{\beta}_{(k)})}{\sum_{t=1}^{m} T_{0,(k)}^{(t)} n_t + \sigma_{(k)}^2 \sum_{t=1}^{m} T_{1,(k)}^{(t)} \text{tr} \left\{ (\sigma_{(k)}^2 \boldsymbol{I} + \eta_{(k)} \boldsymbol{H}_t)^{-1} \right\}}$$

$$+ \frac{\sum_{t=1}^{m} T_{0,(k)}^{(t)} \left\{ \left(\boldsymbol{Y}_{\text{obs}}^{(t)} \right)^T \left(\boldsymbol{Y}_{\text{obs}}^{(t)} \right) \right\}}{\sum_{t=1}^{m} T_{0,(k)}^{(t)} n_t + \sigma_{(k)}^2 \sum_{t=1}^{m} T_{1,(k)}^{(t)} \text{tr} \left\{ (\sigma_{(k)}^2 \boldsymbol{I} + \eta_{(k)} \boldsymbol{H}_t)^{-1} \right\}}.$$

SECTION B: BAYES FACTOR

For a given tissue t, to determine whether the cis-SNPs is relevant to the gene expression, we can calculate the posterior probability $p(I^{(t)} = 0 | \boldsymbol{Y})$ or the Bayesian factor. Specifically, we access the plausibility of the cis-SNP and gene expression association via the Bayes factor (BF)

$$K_t(\boldsymbol{Y}; \boldsymbol{\beta}, \eta, \sigma^2) = \frac{p(\boldsymbol{Y} \mid H_0^{(t)})}{p(\boldsymbol{Y} \mid H_1^{(t)})} = \frac{p(\boldsymbol{Y}^{(t)} \mid I^{(t)} = 0)}{p(\boldsymbol{Y}^{(t)} \mid I^{(t)} = 1)} = \frac{\psi(\boldsymbol{Y}^{(t)}; \boldsymbol{0}, \sigma^2 \boldsymbol{I})}{\psi(\boldsymbol{Y}^{(t)}; \boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{I} + \eta \boldsymbol{H})}, \tag{12}$$

based on the prior distribution in Equations (2.3) and (2.4), where $H_0^{(t)}$ represents the hypothesis $\boldsymbol{\beta}^{(t)} = 0$ and $H_0^{(t)}$ represents the hypothesis $\boldsymbol{\beta}^{(t)} \neq 0$. The second equality in (12) is due to that columns in $\boldsymbol{Y}$ other than $\boldsymbol{Y}^{(t)}$ do not depend on $I^{(t)}$, which implies that $K_t(\boldsymbol{Y}; \boldsymbol{\beta}, \eta, \sigma^2) = K_t(\boldsymbol{Y}^{(t)}; \boldsymbol{\beta}, \eta, \sigma^2)$. In addition, the posterior odds ratio is

$$O_t(\boldsymbol{Y}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2) = \frac{P(I^{(t)} = 0 \mid \boldsymbol{Y}^{(t)})}{P(I^{(t)} = 1 \mid \boldsymbol{Y}^{(t)})} = K_t(\boldsymbol{Y}; \boldsymbol{\beta}, \eta, \sigma^2) \cdot \frac{\tau_0}{\tau_1}. \tag{13}$$

From equation (13), we see that the BF can indicate whether the observed data provides evidence for or against the cis-SNP gene expression association. If BF > 1 then the posterior odds are greater than the prior odds τ_0/τ_1, indicating that the observed data provides evidence for the no cis-SNP gene expression association. If BF < 1 then the data provides evidence for cis-SNP gene expression association.

B.1. Bayes factor under the setting with missing data. The Bayes factor with missing data is

$$K_t^{(obs)}(\boldsymbol{Y}_{\text{obs}}, \boldsymbol{W}; \boldsymbol{\beta}, \eta, \sigma^2) = \frac{p(\boldsymbol{Y}_{\text{obs}}, \boldsymbol{W} \mid H_0^{(t)})}{p(\boldsymbol{Y}_{\text{obs}}, \boldsymbol{W} \mid H_1^{(t)})} = \frac{\psi(\boldsymbol{Y}_{\text{obs}}^{(t)}; \boldsymbol{0}, \sigma^2 \boldsymbol{I})}{\psi(\boldsymbol{Y}_{\text{obs}}^{(t)}; \boldsymbol{X}_t \boldsymbol{\beta}, \sigma^2 \boldsymbol{I} + \boldsymbol{H}_t)},$$

and the posterior odds ratio is

$$O_t^{(obs)}(\boldsymbol{Y}_{\text{obs}}, \boldsymbol{W}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2) = \frac{P(I^{(t)} = 0 \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})}{P(I^{(t)} = 1 \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})} = K_t^{(obs)}(\boldsymbol{Y}_{\text{obs}}, \boldsymbol{W}; \boldsymbol{\beta}, \eta, \sigma^2) \cdot \frac{\tau_0}{\tau_1}.$$

Similarly as in Section B, the Bayes factor $K_t^{(obs)} > 1$ implies that the observed data support the hypothesis $H_0^{(t)}$, while $K_t^{(obs)} < 1$ represents that the observed data are against $H_0^{(t)}$.

SECTION C: PROOFS

PROOF OF PROPOSITION 1. The posterior density of $\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)}, I^{(t)}$ is

$$\begin{aligned} p(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)}, I^{(t)}) &= \frac{p(\boldsymbol{Y}^{(t)} \mid \boldsymbol{\beta}^{(t)}, I^{(t)}) p(\boldsymbol{\beta}^{(t)} \mid I^{(t)})}{\int p(\boldsymbol{Y}^{(t)} \mid \boldsymbol{\beta}^{(t)}, I^{(t)}) p(\boldsymbol{\beta}^{(t)} \mid I^{(t)}) d\boldsymbol{\beta}^{(t)}} \\ &= \frac{\psi(\boldsymbol{Y}^{(t)}; \boldsymbol{X}\boldsymbol{\beta}^{(t)}, \sigma^2 \boldsymbol{I}) p(\boldsymbol{\beta}^{(t)} \mid I^{(t)})}{\int \psi(\boldsymbol{Y}^{(t)}; \boldsymbol{X}\boldsymbol{\beta}^{(t)}, \sigma^2 \boldsymbol{I}) p(\boldsymbol{\beta}^{(t)} \mid I^{(t)}) d\boldsymbol{\beta}^{(t)}}. \end{aligned}$$

Since $p(\boldsymbol{\beta}^{(t)} \mid I^{(t)} = 1) = \psi(\boldsymbol{\beta}^{(t)}; \boldsymbol{\beta}, \eta(\boldsymbol{X}^T \boldsymbol{X})^{-1})$, we have

$$\begin{aligned} & p(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)}, I^{(t)} = 1) \\ &= \psi\left(\boldsymbol{\beta}^{(t)}; (\boldsymbol{X}^T\boldsymbol{X}/\eta + \boldsymbol{X}^T\boldsymbol{X}/\sigma^2)^{-1}(\boldsymbol{X}^T\boldsymbol{X}\boldsymbol{\beta}/\eta + \boldsymbol{X}^T\boldsymbol{Y}^{(t)}/\sigma^2), (\boldsymbol{X}^T\boldsymbol{X}/\eta + \boldsymbol{X}^T\boldsymbol{X}/\sigma^2)^{-1}\right) \\ &= \psi\left(\boldsymbol{\beta}^{(t)}; (1/\eta + 1/\sigma^2)^{-1}(\boldsymbol{\beta}/\eta + \hat{\boldsymbol{\beta}}^{(t)}/\sigma^2), (1/\eta + 1/\sigma^2)^{-1}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\right). \end{aligned}$$

Thus, $E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)}, I^{(t)} = 1) = (1/\eta + 1/\sigma^2)^{-1}(\boldsymbol{\beta}/\eta + \hat{\boldsymbol{\beta}}^{(t)}/\sigma^2)$.

Since

$$\begin{aligned} \int_{\boldsymbol{Y}^{(t)} \in \Omega_Y} P(I^{(t)} = 1 \mid \boldsymbol{Y}^{(t)}) p(\boldsymbol{Y}^{(t)}) d\boldsymbol{Y}^{(t)} &= P(I^{(t)} = 1, \boldsymbol{Y}^{(t)} \in \Omega_Y) \\ &= P(\boldsymbol{Y}^{(t)} \in \Omega_Y \mid I^{(t)} = 1) P(I^{(t)} = 1), \end{aligned}$$

we have

$$\begin{aligned} & P(I^{(t)} = 1 \mid \boldsymbol{Y}^{(t)}) \\ &= \frac{\tau_1 p(\boldsymbol{Y}^{(t)} \mid I^{(t)} = 1)}{p(\boldsymbol{Y}^{(t)})} \\ &= \frac{\tau_1 p(\boldsymbol{Y}^{(t)} \mid I^{(t)} = 1)}{\tau_1 p(\boldsymbol{Y}^{(t)} \mid I^{(t)} = 1) + (1 - \tau_1) p(\boldsymbol{Y}^{(t)} \mid I^{(t)} = 0)} \\ &= \frac{\tau_1 \int p(\boldsymbol{Y}^{(t)}, \boldsymbol{\beta}^{(t)} \mid I^{(t)} = 1) d\boldsymbol{\beta}^{(t)}}{\tau_1 \int p(\boldsymbol{Y}^{(t)}, \boldsymbol{\beta}^{(t)} \mid I^{(t)} = 1) d\boldsymbol{\beta}^{(t)} + (1 - \tau_1) \int p(\boldsymbol{Y}^{(t)}, \boldsymbol{\beta}^{(t)} \mid I^{(t)} = 0) d\boldsymbol{\beta}^{(t)}} \\ &= \left\{\tau_1 \int p(\boldsymbol{Y}^{(t)} \mid \boldsymbol{\beta}^{(t)}, I^{(t)} = 1) p(\boldsymbol{\beta}^{(t)} \mid I^{(t)} = 1) d\boldsymbol{\beta}^{(t)}\right\} \end{aligned}$$

$$\cdot\left\{\tau_1 \int p(\boldsymbol{Y}^{(t)} \mid \boldsymbol{\beta}^{(t)}, I^{(t)}=1) p(\boldsymbol{\beta}^{(t)} \mid I^{(t)}=1) d\boldsymbol{\beta}^{(t)}\right.$$
$$\left.+(1-\tau_1) \int p(\boldsymbol{Y}^{(t)} \mid \boldsymbol{\beta}^{(t)}, I^{(t)}=0) p(\boldsymbol{\beta}^{(t)} \mid I^{(t)}=0) d\boldsymbol{\beta}^{(t)}\right\}^{-1}.$$

Since

$$\begin{aligned}
&\int p(\boldsymbol{Y}^{(t)} \mid \boldsymbol{\beta}^{(t)}, I^{(t)}=1) p(\boldsymbol{\beta}^{(t)} \mid I^{(t)}=1) d\boldsymbol{\beta}^{(t)} \\
=&\int \psi(\boldsymbol{Y}^{(t)}; \boldsymbol{X}\boldsymbol{\beta}^{(t)}, \sigma^2 \boldsymbol{I}) \psi(\boldsymbol{\beta}^{(t)}; \boldsymbol{\beta}, \eta(\boldsymbol{X}^T\boldsymbol{X})^{-1}) d\boldsymbol{\beta}^{(t)} \\
=&\frac{\psi(\boldsymbol{Y}^{(t)}; \boldsymbol{X}\boldsymbol{\beta}^{(t)}, \sigma^2 \boldsymbol{I}) \psi(\boldsymbol{\beta}^{(t)}; \boldsymbol{\beta}, \eta(\boldsymbol{X}^T\boldsymbol{X})^{-1})}{\psi\left(\boldsymbol{\beta}^{(t)}; (1/\eta+1/\sigma^2)^{-1}(\boldsymbol{\beta}/\eta+\hat{\boldsymbol{\beta}}^{(t)}/\sigma^2), (1/\eta+1/\sigma^2)^{-1}(\boldsymbol{X}^T\boldsymbol{X})^{-1}\right)} \\
=&\psi(\boldsymbol{Y}^{(t)}; \boldsymbol{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{I}+\eta\boldsymbol{H})
\end{aligned}$$

and

$$\int p(\boldsymbol{Y}^{(t)} \mid \boldsymbol{\beta}^{(t)}, I^{(t)}=0) p(\boldsymbol{\beta}^{(t)} \mid I^{(t)}=0) d\boldsymbol{\beta}^{(t)} = p(\boldsymbol{Y}^{(t)} \mid \boldsymbol{\beta}^{(t)}=\boldsymbol{0}) = \psi(\boldsymbol{Y}^{(t)}; \boldsymbol{0}, \sigma^2\boldsymbol{I}),$$

we have

$$P(I^{(t)}=1 \mid \boldsymbol{Y}^{(t)}) = \frac{\tau_1 \psi(\boldsymbol{Y}^{(t)}; \boldsymbol{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{I}+\eta\boldsymbol{H})}{\tau_1 \psi(\boldsymbol{Y}^{(t)}; \boldsymbol{X}\boldsymbol{\beta}, \sigma^2\boldsymbol{I}+\eta\boldsymbol{H}) + (1-\tau_1)\psi(\boldsymbol{Y}^{(t)}; \boldsymbol{0}, \sigma^2\boldsymbol{I})}.$$

Note that

$$\int_{\boldsymbol{\beta}^{(t)} \in [a,b]^p} p(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)}, I^{(t)}=0) d\boldsymbol{\beta}^{(t)} = P(\boldsymbol{\beta}^{(t)} \in [a,b]^p \mid \boldsymbol{Y}^{(t)}, I^{(t)}=0),$$

and

$$\begin{aligned}
&\int_{\boldsymbol{Y}^{(t)} \in \Omega_Y} P(\boldsymbol{\beta}^{(t)} \in [a,b]^p \mid \boldsymbol{Y}^{(t)}, I^{(t)}=0) p(\boldsymbol{Y}^{(t)} \mid I^{(t)}=0) d\boldsymbol{Y}^{(t)} \\
=&P(\boldsymbol{\beta}^{(t)} \in [a,b]^p, \boldsymbol{Y}^{(t)} \in \Omega_Y \mid I^{(t)}=0) \\
=&P(\boldsymbol{\beta}^{(t)} \in [a,b]^p, \boldsymbol{Y}^{(t)} \in \Omega_Y, I^{(t)}=0)/P(I^{(t)}=0) \\
=&P(\boldsymbol{Y}^{(t)} \in \Omega_Y \mid \boldsymbol{\beta}^{(t)} \in [a,b]^p, I^{(t)}=0) P(\boldsymbol{\beta}^{(t)} \in [a,b]^p \mid I^{(t)}=0).
\end{aligned}$$

If $\boldsymbol{0} \in [a,b]^p$, $P(\boldsymbol{\beta}^{(t)} \in [a,b]^p \mid I^{(t)}=0)=1$. If $\boldsymbol{0} \notin [a,b]^p$, $P(\boldsymbol{\beta}^{(t)} \in [a,b]^p \mid I^{(t)}=0)=0$. Thus, $P(\boldsymbol{\beta}^{(t)} \neq \boldsymbol{0} \mid \boldsymbol{Y}^{(t)}, I^{(t)}=0)=0$ and

$$P(\boldsymbol{\beta}^{(t)}=\boldsymbol{0} \mid \boldsymbol{Y}^{(t)}, I^{(t)}=0) = \frac{p(\boldsymbol{Y}^{(t)} \mid \boldsymbol{\beta}^{(t)}=\boldsymbol{0}, I^{(t)}=0)}{p(\boldsymbol{Y}^{(t)} \mid I^{(t)}=0)} = \frac{p(\boldsymbol{Y}^{(t)} \mid \boldsymbol{\beta}^{(t)}=\boldsymbol{0})}{p(\boldsymbol{Y}^{(t)} \mid I^{(t)}=0)}.$$

Thus, $E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)}, I^{(t)}=0)=0$.

Therefore, the posterior mean is

$$E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)}) = h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)\left(\frac{1}{\eta}+\frac{1}{\sigma^2}\right)^{-1}\left(\frac{\boldsymbol{\beta}}{\eta}+\frac{\hat{\boldsymbol{\beta}}^{(t)}}{\sigma^2}\right).$$

□

PROOF OF THEOREM 1. The Bayes risk function of an estimator $\boldsymbol{\delta}_m$ is

$$R_m(\boldsymbol{\delta}_m) = \int \left\{\boldsymbol{\delta}_m(\boldsymbol{Y}) - \boldsymbol{\beta}^{(t)}\right\}^T \Delta \left\{\boldsymbol{\delta}_m(\boldsymbol{Y}) - \boldsymbol{\beta}^{(t)}\right\} \prod_{i=1}^m \left\{p(\boldsymbol{Y}^{(i)} \mid \boldsymbol{\beta}^{(i)}) p(\boldsymbol{\beta}^{(i)}) d\boldsymbol{Y}^{(i)} d\boldsymbol{\beta}^{(i)}\right\}$$
$$= C + \int \left\{\boldsymbol{\delta}_m(\boldsymbol{Y}) - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y})\right\}^T \Delta \left\{\boldsymbol{\delta}_m(\boldsymbol{Y}) - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y})\right\}$$
$$\cdot \prod_{i=1}^m \left\{p(\boldsymbol{Y}^{(i)}) d\boldsymbol{Y}^{(i)}\right\},$$

where C is a constant independent of $\boldsymbol{\delta}_m$. Thus, $R_m(\boldsymbol{\delta}_m)$ is minimized at $\boldsymbol{\delta}_m = E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}) = E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)}) = \bar{\boldsymbol{\beta}}^{(t)}(\boldsymbol{Y}^{(t)})$, where the second equality follows from the independence among $\boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(m)}$, and independence among $\boldsymbol{\varepsilon}^{(1)}, \ldots, \boldsymbol{\varepsilon}^{(m)}$. Then,

$$\inf_{\boldsymbol{\delta}_m \in \mathcal{E}_m^*} R_m(\boldsymbol{\delta}_m) = R_m(\bar{\boldsymbol{\beta}}^{(t)}) = C. \tag{14}$$

In addition, we have

$$R_m(\hat{\boldsymbol{\beta}}^{(t)}) - R_m(\bar{\boldsymbol{\beta}}^{(t)}) = \int \left\{\hat{\boldsymbol{\beta}}^{(t)} - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)})\right\}^T \Delta \left\{\hat{\boldsymbol{\beta}}^{(t)} - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)})\right\}$$
$$\cdot \prod_{i=1}^m \left\{p(\boldsymbol{Y}^{(i)}) d\boldsymbol{Y}^{(i)}\right\}$$
$$\geq \lambda_{\min}(\Delta) \int \left\|\hat{\boldsymbol{\beta}}^{(t)}(\boldsymbol{Y}^{(t)}) - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)})\right\|_2^2 \cdot \prod_{i=1}^m \left\{p(\boldsymbol{Y}^{(i)}) d\boldsymbol{Y}^{(i)}\right\}$$
$$= \lambda_{\min}(\Delta) \int \left\|\hat{\boldsymbol{\beta}}^{(t)}(\boldsymbol{Y}^{(t)}) - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)})\right\|_2^2 \cdot p(\boldsymbol{Y}^{(t)}) d\boldsymbol{Y}^{(t)}.$$

Since

$$E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)}) - \hat{\boldsymbol{\beta}}^{(t)}$$
$$= \frac{\eta\sigma^2 \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)}{\eta + \sigma^2} \left(\frac{\boldsymbol{\beta}}{\eta} + \frac{\hat{\boldsymbol{\beta}}^{(t)}}{\sigma^2}\right) - \hat{\boldsymbol{\beta}}^{(t)}$$
$$= \frac{\sigma^2 \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)}{\eta + \sigma^2} \boldsymbol{\beta} - \left(1 - \frac{\eta \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)}{\eta + \sigma^2}\right) \hat{\boldsymbol{\beta}}^{(t)}$$
$$= \left(1 - \frac{\eta \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)}{\eta + \sigma^2}\right) \left(\frac{\sigma^2 \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)}{\eta + \sigma^2 - \eta \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)} \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}^{(t)}\right),$$

we have

$$\left\|E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}^{(t)}) - \hat{\boldsymbol{\beta}}^{(t)}\right\|_2$$
$$= \left|1 - \frac{\eta \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)}{\eta + \sigma^2}\right| \left\|\frac{\sigma^2 \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)}{\eta + \sigma^2 - \eta \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)} \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}^{(t)}\right\|_2$$
$$\geq \left|1 - \frac{\eta}{\eta + \sigma^2}\right| \left\|\frac{\sigma^2 \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)}{\eta + \sigma^2 - \eta \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)} \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}^{(t)}\right\|_2.$$

Since

$$0 \leq \frac{\sigma^2 \cdot h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)}{\sigma^2 + \eta\left\{1 - h_1(\boldsymbol{Y}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma^2)\right\}} \leq 1,$$

$$\int \left\| \hat{\boldsymbol\beta}^{(t)} - E(\boldsymbol\beta^{(t)} \mid \boldsymbol Y^{(t)}) \right\|_2^2 p(\boldsymbol Y^{(t)}; \tau_1, \boldsymbol\beta, \eta, \sigma^2) d\boldsymbol Y^{(t)}$$
$$\geq \frac{\sigma^2}{\eta+\sigma^2} \int_{\|\hat{\boldsymbol\beta}^{(t)}\|_2 > \|\boldsymbol\beta\|_2 + \alpha} \alpha^2 p(\boldsymbol Y^{(t)}; \tau_1, \boldsymbol\beta, \eta, \sigma^2) d\boldsymbol Y^{(t)}$$
$$= \frac{\sigma^2}{\eta+\sigma^2} \int_{\|\hat{\boldsymbol\beta}^{(t)}\|_2 > \|\boldsymbol\beta\|_2 + \alpha} \alpha^2 \left\{ \tau_1 g_1(\boldsymbol Y^{(t)}; \boldsymbol\beta, \eta, \sigma^2) + (1-\tau_1) g_0(\boldsymbol Y^{(t)}; \sigma^2) \right\} d\boldsymbol Y^{(t)}$$
$$\geq \frac{\sigma^2\alpha^2\tau_1}{\eta+\sigma^2} \int_{\boldsymbol A\hat{\boldsymbol\beta}^{(t)} \notin \mathcal{B}(\boldsymbol A\boldsymbol\beta, \lambda_{\max}((\boldsymbol X^T\boldsymbol X)^{1/2})(2\|\boldsymbol\beta\|_2+\alpha)/(\sigma^2+\eta)^{1/2})} g_1(\boldsymbol Y^{(t)}; \boldsymbol\beta, \eta, \sigma^2) d\boldsymbol Y^{(t)}$$
$$+ \frac{\sigma^2\alpha^2(1-\tau_1)}{\eta+\sigma^2} \int_{(\boldsymbol X^T\boldsymbol X)^{1/2}\hat{\boldsymbol\beta}^{(t)}/\sigma \notin \mathcal{B}(\boldsymbol 0, \lambda_{\max}((\boldsymbol X^T\boldsymbol X)^{1/2})(\|\boldsymbol\beta\|_2+\alpha)/\sigma)} g_0(\boldsymbol Y^{(t)}; \sigma^2) d\boldsymbol Y^{(t)}$$
$$= \frac{\sigma^2\alpha^2}{\eta+\sigma^2} \Big(\tau_1 \varphi \left[\lambda_{\max}\{(\boldsymbol X^T\boldsymbol X)^{1/2}\}(2\|\boldsymbol\beta\|_2+\alpha)/(\sigma^2+\eta)^{1/2} \right]$$
$$+ (1-\tau_1)\varphi \left[\lambda_{\max}\{(\boldsymbol X^T\boldsymbol X)^{1/2}\}(\|\boldsymbol\beta\|_2+\alpha)/\sigma \right] \Big),$$

where $\boldsymbol A = (\boldsymbol X^T\boldsymbol X)^{1/2}/(\sigma^2+\eta)^{1/2}$. Thus, for any $1 \leq t \leq m$, we have

$$R_m(\hat{\boldsymbol\beta}^{(t)}) - R_m(\bar{\boldsymbol\beta}^{(t)}) \geq \frac{\sigma^2\alpha^2\lambda_{\min}(\Delta)}{\eta+\sigma^2} \Big(\tau_1 \varphi \left[\lambda_{\max}\{(\boldsymbol X^T\boldsymbol X)^{1/2}\}(2\|\boldsymbol\beta\|_2+\alpha)/(\sigma^2+\eta)^{1/2} \right]$$
$$+ (1-\tau_1)\varphi \left[\lambda_{\max}\{(\boldsymbol X^T\boldsymbol X)^{1/2}\}(\|\boldsymbol\beta\|_2+\alpha)/\sigma \right] \Big).$$

□

PROOF OF THEOREM 2. Let $\widetilde{\boldsymbol\gamma} = (\widetilde\tau_1, \widetilde\eta, \widetilde\sigma^2, \widetilde{\boldsymbol\beta}^T)^T$ and $\boldsymbol\gamma = (\tau_1, \eta, \sigma^2, \boldsymbol\beta^T)^T$. Then $\widetilde{\boldsymbol\gamma}$ is MLE of $\boldsymbol\gamma$ based on i.i.d samples $\boldsymbol Y^{(1)}, \ldots, \boldsymbol Y^{(m)}$. Thus, we have $\widetilde{\boldsymbol\gamma} \xrightarrow{p} \boldsymbol\gamma$.

Let

$$\boldsymbol\phi(\boldsymbol\gamma, \boldsymbol Y^{(t)}) = h_1(\boldsymbol Y^{(t)}; \tau_1, \boldsymbol\beta, \eta, \sigma) \left(\frac{1}{\eta} + \frac{1}{\sigma^2} \right)^{-1} \left(\frac{1}{\eta}\boldsymbol\beta + \frac{1}{\sigma^2}\hat{\boldsymbol\beta} \right).$$

Then, $\bar{\boldsymbol\beta}^{(t)}(\boldsymbol Y^{(t)}) = \boldsymbol\phi(\boldsymbol\gamma, \boldsymbol Y^{(t)})$ and $\widetilde{\boldsymbol\beta}^{(t)}(\boldsymbol Y) = \boldsymbol\phi(\widetilde{\boldsymbol\gamma}, \boldsymbol Y^{(t)})$. For any $\epsilon > 0$ and $\delta > 0$, there exists a constant $c_2(n)$ such that $P(\|\boldsymbol Y^{(t)}\|_2 > c_2(n)) < \delta/3$. Since $\widetilde{\boldsymbol\gamma} \xrightarrow{p} \boldsymbol\gamma$, there exist a constant c_3 and an integer M_1 such that for $m > M_1$, $P(\|\widetilde{\boldsymbol\gamma} - \boldsymbol\gamma\|_2 > c_3) < \delta/3$.

$$P\left\{ \left\| \widetilde{\boldsymbol\beta}^{(t)}(\boldsymbol Y) - \bar{\boldsymbol\beta}^{(t)}(\boldsymbol Y^{(t)}) \right\|_2 > \epsilon \right\}$$
$$< 2\delta/3 + P\left\{ \left\| \boldsymbol\phi(\widetilde{\boldsymbol\gamma}, \boldsymbol Y^{(t)}) - \boldsymbol\phi(\boldsymbol\gamma, \boldsymbol Y^{(t)}) \right\|_2 > \epsilon, \|\widetilde{\boldsymbol\gamma} - \boldsymbol\gamma\|_2 \leq c_3, \|\hat{\boldsymbol\beta}^{(t)}\|_2 \leq c_2(n) \right\}.$$

Since the continuous function $\boldsymbol\phi$ is uniformly continuous on a compact set, there exists a positive constant ϵ' such that, if $\|\widetilde{\boldsymbol\gamma} - \boldsymbol\gamma\|_2 < \epsilon'$, then $\|\boldsymbol\phi(\widetilde{\boldsymbol\gamma}, \boldsymbol Y^{(t)}) - \boldsymbol\phi(\boldsymbol\gamma, \boldsymbol Y^{(t)})\|_2 < \epsilon$. Then,

$$P\left\{ \left\| \boldsymbol\phi(\widetilde{\boldsymbol\gamma}, \boldsymbol Y) - \boldsymbol\phi(\boldsymbol\gamma, \boldsymbol Y^{(t)}) \right\|_2 > \epsilon, \|\widetilde{\boldsymbol\gamma} - \boldsymbol\gamma\|_2 \leq c_3, \|\hat{\boldsymbol\beta}^{(t)}\|_2 \leq c_2(n) \right\}$$
$$= P\left\{ \left\| \boldsymbol\phi(\widetilde{\boldsymbol\gamma}, \boldsymbol Y) - \boldsymbol\phi(\boldsymbol\gamma, \boldsymbol Y^{(t)}) \right\|_2 > \epsilon, \epsilon' \leq \|\widetilde{\boldsymbol\gamma} - \boldsymbol\gamma\|_2 \leq c_3, \|\hat{\boldsymbol\beta}^{(t)}\|_2 \leq c_2(n) \right\}$$
$$\leq P(\|\widetilde{\boldsymbol\gamma} - \boldsymbol\gamma\|_2 \geq \epsilon').$$

Since $\widetilde{\boldsymbol{\gamma}} \xrightarrow{p} \boldsymbol{\gamma}$, there exists an integer M_2 such that for $m > M_2$, $P(\|\widetilde{\boldsymbol{\gamma}} - \boldsymbol{\gamma}\|_2 \geq \epsilon') < \delta/3$. Thus, we have

$$P\left\{\left\|\widetilde{\boldsymbol{\beta}}^{(t)}(\boldsymbol{Y}) - \bar{\boldsymbol{\beta}}^{(t)}(\boldsymbol{Y}^{(t)})\right\|_2 > \epsilon\right\} < \delta$$

for $m > \max\{M_1, M_2\}$.

□

PROOF OF THEOREM 3. The Bayes risk function of an estimator $\boldsymbol{\delta}_m$ with missing data is

$$\begin{aligned}
R_m^{(obs)}(\boldsymbol{\delta}_m) &= \int \left\{\boldsymbol{\delta}_m(\boldsymbol{Y}_{\text{obs}}, \boldsymbol{W}) - \boldsymbol{\beta}^{(t)}\right\}^T \Delta \left\{\boldsymbol{\delta}_m(\boldsymbol{Y}_{\text{obs}}, \boldsymbol{W}) - \boldsymbol{\beta}^{(t)}\right\} \\
&\quad \prod_{i=1}^m \left\{p(\boldsymbol{W}^{(i)}, \boldsymbol{Y}_{\text{obs}}^{(i)} \mid \boldsymbol{\beta}^{(i)}) p(\boldsymbol{\beta}^{(i)}) d\boldsymbol{Y}_{\text{obs}}^{(i)} d\boldsymbol{W}^{(i)} d\boldsymbol{\beta}^{(i)}\right\} \\
&= C + \int \left\{\boldsymbol{\delta}_m(\boldsymbol{Y}_{\text{obs}}, \boldsymbol{W}) - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}, \boldsymbol{W})\right\}^T \Delta \left\{\boldsymbol{\delta}_m(\boldsymbol{Y}_{\text{obs}}, \boldsymbol{W}) - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}, \boldsymbol{W})\right\} \\
&\quad \cdot \prod_{i=1}^m \left\{p(\boldsymbol{W}^{(i)}, \boldsymbol{Y}_{\text{obs}}^{(i)}) d\boldsymbol{W}^{(i)} d\boldsymbol{Y}_{\text{obs}}^{(i)}\right\},
\end{aligned}$$

where C is a constant independent of $\boldsymbol{\delta}_m$. Thus, $R_m^{(obs)}(\boldsymbol{\delta}_m)$ is minimized at $\boldsymbol{\delta}_m = E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}, \boldsymbol{W}) = E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) = \bar{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})$, where the second equality follows from the independence among $\boldsymbol{\beta}^{(1)}, \ldots, \boldsymbol{\beta}^{(m)}$, and independence among $\boldsymbol{\varepsilon}^{(1)}, \ldots, \boldsymbol{\varepsilon}^{(m)}$. Then,

$$\inf_{\boldsymbol{\delta}_m \in \bar{\mathcal{E}}_m^*} R_m^{(obs)}(\boldsymbol{\delta}_m) = R_m^{(obs)}(\bar{\boldsymbol{\beta}}_{obs}^{(t)}) = C. \tag{15}$$

In addition, we have

$$\begin{aligned}
R_m^{(obs)}(\hat{\boldsymbol{\beta}}_{obs}^{(t)}) - R_m^{(obs)}(\bar{\boldsymbol{\beta}}_{obs}^{(t)}) &= \int \left\{\hat{\boldsymbol{\beta}}_{obs}^{(t)} - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})\right\}^T \Delta \left\{\hat{\boldsymbol{\beta}}_{obs}^{(t)} - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})\right\} \\
&\quad \cdot \prod_{i=1}^m \left\{p(\boldsymbol{W}^{(i)}, \boldsymbol{Y}_{\text{obs}}^{(i)}) d\boldsymbol{W}^{(i)} d\boldsymbol{Y}_{\text{obs}}^{(i)}\right\} \\
&\geq \lambda_{\min}(\Delta) \int \left\|\hat{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})\right\|_2^2 \\
&\quad \cdot \prod_{i=1}^m \left\{p(\boldsymbol{W}^{(i)}, \boldsymbol{Y}_{\text{obs}}^{(i)}) d\boldsymbol{W}^{(i)} d\boldsymbol{Y}_{\text{obs}}^{(i)}\right\} \\
&= \lambda_{\min}(\Delta) \int \left\|\hat{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})\right\|_2^2 \\
&\quad \cdot p(\boldsymbol{W}^{(t)}, \boldsymbol{Y}_{\text{obs}}^{(t)}) d\boldsymbol{W}^{(t)} d\boldsymbol{Y}_{\text{obs}}^{(t)}.
\end{aligned}$$

Let $\boldsymbol{A}(\boldsymbol{X}, \eta) = \boldsymbol{X}^T\boldsymbol{X}/\eta$ and $\boldsymbol{B}(\boldsymbol{X}, \boldsymbol{W}^{(t)}, \sigma^2) = \boldsymbol{X}^T\boldsymbol{W}^{(t)}\boldsymbol{X}/\sigma^2$. Note that

$$\begin{aligned}
&E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) - \hat{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) \\
&= h_2(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma) \left(\frac{\boldsymbol{X}^T\boldsymbol{X}}{\eta} + \frac{\boldsymbol{X}^T\boldsymbol{W}^{(t)}\boldsymbol{X}}{\sigma^2}\right)^{-1} \left(\frac{\boldsymbol{X}^T\boldsymbol{X}}{\eta}\boldsymbol{\beta} + \frac{\boldsymbol{X}^T\boldsymbol{W}^{(t)}\boldsymbol{Y}^{(t)}}{\sigma^2}\right) - \hat{\boldsymbol{\beta}}_{obs}^{(t)} \\
&= h_2(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma) \left(\frac{\boldsymbol{X}^T\boldsymbol{X}}{\eta} + \frac{\boldsymbol{X}^T\boldsymbol{W}^{(t)}\boldsymbol{X}}{\sigma^2}\right)^{-1} \left(\frac{\boldsymbol{X}^T\boldsymbol{X}}{\eta}\boldsymbol{\beta} + \frac{\boldsymbol{X}^T\boldsymbol{W}^{(t)}\boldsymbol{X}}{\sigma^2}\hat{\boldsymbol{\beta}}_{obs}^{(t)}\right) - \hat{\boldsymbol{\beta}}_{obs}^{(t)}
\end{aligned}$$

$$= h_2(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma) \left(\frac{\boldsymbol{X}^T\boldsymbol{X}}{\eta} + \frac{\boldsymbol{X}^T\boldsymbol{W}^{(t)}\boldsymbol{X}}{\sigma^2} \right)^{-1} \frac{\boldsymbol{X}^T\boldsymbol{X}}{\eta}\boldsymbol{\beta}$$

$$- \left\{ \boldsymbol{I} - h_2(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \tau_1, \boldsymbol{\beta}, \eta, \sigma) \left(\frac{\boldsymbol{X}^T\boldsymbol{X}}{\eta} + \frac{\boldsymbol{X}^T\boldsymbol{W}^{(t)}\boldsymbol{X}}{\sigma^2} \right)^{-1} \frac{\boldsymbol{X}^T\boldsymbol{W}^{(t)}\boldsymbol{X}}{\sigma^2} \right\} \hat{\boldsymbol{\beta}}_{obs}^{(t)}$$

$$= h_2 \cdot (\boldsymbol{A} + \boldsymbol{B})^{-1} \boldsymbol{A}\boldsymbol{\beta} - \left\{ \boldsymbol{I} - h_2 \cdot (\boldsymbol{A} + \boldsymbol{B})^{-1} \boldsymbol{B} \right\} \hat{\boldsymbol{\beta}}_{obs}^{(t)}$$

$$= \left\{ \boldsymbol{I} - h_2 \cdot (\boldsymbol{A} + \boldsymbol{B})^{-1} \boldsymbol{B} \right\} \left\{ \left\{ \boldsymbol{I} - h_2 \cdot (\boldsymbol{A} + \boldsymbol{B})^{-1} \boldsymbol{B} \right\}^{-1} h_2 \cdot (\boldsymbol{A} + \boldsymbol{B})^{-1} \boldsymbol{A}\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{obs}^{(t)} \right\}$$

$$= \left\{ \boldsymbol{I} - h_2 \cdot (\boldsymbol{A} + \boldsymbol{B})^{-1} \boldsymbol{B} \right\} \left[\{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\}^{-1} h_2 \cdot \boldsymbol{A}\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{obs}^{(t)} \right]$$

$$= (\boldsymbol{A} + \boldsymbol{B})^{-1} \{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\} \left[\{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\}^{-1} h_2 \cdot \boldsymbol{A}\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{obs}^{(t)} \right].$$

Then we have

$$\left\| E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) - \hat{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) \right\|_2$$

$$\geq \left\| \left[(\boldsymbol{A} + \boldsymbol{B})^{-1} \{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\} \right]^{-1} \right\|_2^{-1} \left\| \{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\}^{-1} h_2 \cdot \boldsymbol{A}\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{obs}^{(t)} \right\|_2$$

$$= \left\| \{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\}^{-1} (\boldsymbol{A} + \boldsymbol{B}) \right\|_2^{-1} \left\| \{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\}^{-1} h_2 \cdot \boldsymbol{A}\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{obs}^{(t)} \right\|_2$$

$$\geq \frac{\lambda_{\min}\left(\boldsymbol{X}^T\boldsymbol{X}\right)}{\lambda_{\max}\left(\boldsymbol{X}^T\boldsymbol{X}\right)} \frac{\sigma^2}{\eta + \sigma^2} \left\| \{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\}^{-1} h_2 \cdot \boldsymbol{A}\boldsymbol{\beta} - \hat{\boldsymbol{\beta}}_{obs}^{(t)} \right\|_2$$

$$\geq \frac{\lambda_{\min}\left(\boldsymbol{X}^T\boldsymbol{X}\right)}{\lambda_{\max}\left(\boldsymbol{X}^T\boldsymbol{X}\right)} \frac{\sigma^2}{\eta + \sigma^2} \left[\left\| \hat{\boldsymbol{\beta}}_{obs}^{(t)} \right\|_2 - h_2 \left\| \{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\}^{-1} \boldsymbol{A}\boldsymbol{\beta} \right\|_2 \right].$$

Note that

$$\begin{aligned} h_2 \left\| \{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\}^{-1} \boldsymbol{A}\boldsymbol{\beta} \right\|_2 &\leq h_2 \left\| \{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\}^{-1} \right\|_2 \|\boldsymbol{A}\boldsymbol{\beta}\|_2 \\ &\leq h_2 \lambda_{\min} \{\boldsymbol{A} + (1 - h_2)\boldsymbol{B}\}^{-1} \|\boldsymbol{A}\boldsymbol{\beta}\|_2 \\ &\leq \lambda_{\min}(\boldsymbol{A})^{-1} \|\boldsymbol{A}\boldsymbol{\beta}\|_2 \\ &\leq \frac{\lambda_{\min}(\boldsymbol{A})}{\lambda_{\max}(\boldsymbol{A})} \|\boldsymbol{\beta}\|_2. \end{aligned}$$

Let $\kappa(\boldsymbol{X}^T\boldsymbol{X}) = \lambda_{\min}(\boldsymbol{A}) / \lambda_{\max}(\boldsymbol{A})$. Then we have

$$\begin{aligned} &\left\| E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) - \hat{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) \right\|_2 \\ &\geq \kappa \frac{\sigma^2}{\eta + \sigma^2} \left(\|\hat{\boldsymbol{\beta}}_{obs}^{(t)}\|_2 - \kappa \|\boldsymbol{\beta}\|_2 \right). \end{aligned}$$

Note that

$$\int \left\| \hat{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) - E(\boldsymbol{\beta}^{(t)} \mid \boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}) \right\|_2^2 p(\boldsymbol{W}^{(t)}, \boldsymbol{Y}_{\text{obs}}^{(t)}) d\boldsymbol{W}^{(t)} d\boldsymbol{Y}_{\text{obs}}^{(t)}$$

$$\geq \kappa \frac{\sigma^2}{\eta + \sigma^2} \int_{\|\hat{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})\|_2 > \kappa \|\boldsymbol{\beta}\|_2 + \alpha} \alpha^2 p(\boldsymbol{W}^{(t)}, \boldsymbol{Y}_{\text{obs}}^{(t)}) d\boldsymbol{W}^{(t)} d\boldsymbol{Y}_{\text{obs}}^{(t)}$$

$$
\begin{aligned}
&= \kappa \frac{\sigma^2}{\eta+\sigma^2} \int_{\|\hat{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})\|_2 > \kappa\|\boldsymbol{\beta}\|_2+\alpha} \alpha^2 \Big\{ \tau_1 g_1(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \boldsymbol{\beta}, \eta, \sigma^2) \\
&\quad + (1-\tau_1) g_0(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)}; \sigma^2) \Big\} d\boldsymbol{W}^{(t)} d\boldsymbol{Y}_{\text{obs}}^{(t)} \\
&\geq \kappa \frac{\sigma^2\alpha^2\tau_1}{\eta+\sigma^2} \int_{\boldsymbol{B}^{1/2}\hat{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})/(\sigma^2+\eta)^{1/2} \notin \mathcal{B}(\boldsymbol{\beta}, \lambda_{\max}(\boldsymbol{B}^{1/2})((1+\kappa)\|\boldsymbol{\beta}\|_2+\alpha)/(\sigma^2+\eta)^{1/2})} \\
&\quad \cdot \psi(\boldsymbol{Y}_{\text{obs}}^{(t)}; \boldsymbol{X}_t\boldsymbol{\beta}, \sigma^2\boldsymbol{I}_{n_t} + \eta\boldsymbol{H}_t) d\boldsymbol{Y}_{\text{obs}}^{(t)} p(\boldsymbol{W}^{(t)}) d\boldsymbol{W}^{(t)} \\
&\quad + \kappa \frac{\sigma^2\alpha^2(1-\tau_1)}{\eta+\sigma^2} \int_{\boldsymbol{B}^{1/2}\hat{\boldsymbol{\beta}}_{obs}^{(t)}(\boldsymbol{Y}_{\text{obs}}^{(t)}, \boldsymbol{W}^{(t)})/\sigma \notin \mathcal{B}(\boldsymbol{0}, \lambda_{\max}(\boldsymbol{B}^{1/2})(\kappa\|\boldsymbol{\beta}\|_2+\alpha)/\sigma)} \psi(\boldsymbol{Y}_{\text{obs}}^{(t)}; \boldsymbol{0}, \sigma^2\boldsymbol{I}_{n_t}) d\boldsymbol{Y}_{\text{obs}}^{(t)} \\
&\quad \cdot p(\boldsymbol{W}^{(t)}) d\boldsymbol{W}^{(t)} \\
&= \kappa \frac{\sigma^2\alpha^2}{\eta+\sigma^2} \int \Big(\tau_1 \varphi \Big[\lambda_{\max}\{\boldsymbol{B}^{1/2}\}((1+\kappa)\|\boldsymbol{\beta}\|_2+\alpha)/(\sigma^2+\eta)^{1/2} \Big] \\
&\quad + (1-\tau_1)\varphi \Big[\lambda_{\max}\{\boldsymbol{B}^{1/2}\}(\kappa\|\boldsymbol{\beta}\|_2+\alpha)/\sigma \Big] \Big) p(\boldsymbol{W}^{(t)}) d\boldsymbol{W}^{(t)} \\
&\geq \kappa \frac{\sigma^2\alpha^2}{\eta+\sigma^2} \int \Big(\tau_1 \varphi \Big[\lambda_{\max}\{\boldsymbol{B}^{1/2}\}((1+\kappa)\|\boldsymbol{\beta}\|_2+\alpha)/(\sigma^2+\eta)^{1/2} \Big] \\
&\quad + (1-\tau_1)\varphi \Big[\lambda_{\max}\{\boldsymbol{B}^{1/2}\}(\kappa\|\boldsymbol{\beta}\|_2+\alpha)/\sigma \Big] \Big) p(\boldsymbol{W}^{(t)}) d\boldsymbol{W}^{(t)} \\
&\geq \kappa \frac{\sigma^2\alpha^2}{\eta+\sigma^2} \Big(\tau_1 \varphi \Big[\lambda_{\max}\{\boldsymbol{A}^{1/2}\}((1+\kappa)\|\boldsymbol{\beta}\|_2+\alpha)/(\sigma^2+\eta)^{1/2} \Big] \\
&\quad + (1-\tau_1)\varphi \Big[\lambda_{\max}\{\boldsymbol{A}^{1/2}\}(\kappa\|\boldsymbol{\beta}\|_2+\alpha)/\sigma \Big] \Big).
\end{aligned}
$$

Thus, for any $1 \leq t \leq m$,

$$
\begin{aligned}
R_m^{(obs)}(\hat{\boldsymbol{\beta}}_{obs}^{(t)}) - R_m^{(obs)}(\bar{\boldsymbol{\beta}}_{obs}^{(t)}) \geq \kappa \frac{\sigma^2\alpha^2\lambda_{\min}(\Delta)}{\eta+\sigma^2} \Big(&\tau_1 \varphi \Big[\lambda_{\max}\{(\boldsymbol{X}^T\boldsymbol{X})^{1/2}\}((1+\kappa)\|\boldsymbol{\beta}\|_2 \\
&+\alpha)/(\sigma^2+\eta)^{1/2} \Big] \\
&+(1-\tau_1)\varphi \Big[\lambda_{\max}\{(\boldsymbol{X}^T\boldsymbol{X})^{1/2}\}(\kappa\|\boldsymbol{\beta}\|_2+\alpha)/\sigma \Big] \Big).
\end{aligned}
$$

□

SECTION D: ADDITIONAL SIMULATION RESULTS

D.1. Simulation results in Setting 3. Table D.1 shows the simulation results under the Setting 3.

TABLE D.1

MSEs and PMSEs of OLS, UTMOST, and the proposed method under Setting 3. "MSE_OLS", "MSE_UTMOST", "MSE_mEBmix" represent MSEs of the OLS, UTMOST, and the proposed method, respectively. "PMSE_OLS", "PMSE_UTMOST", "PMSE_mEBmix" represent PMSEs of the OLS, UTMOST, and the proposed method mEBmix, respectively.

Correlation ρ	0			0.6			0.8		
Signal level β_S	0.5	1	2	0.5	1	2	0.5	1	2
MSE_OLS	0.112	0.110	0.110	0.272	0.274	0.267	0.528	0.541	0.538
MSE_UTMOST	0.099	0.124	0.223	0.077	0.105	0.164	0.091	0.119	0.218
MSE_mEBmix	0.010	0.010	0.010	0.027	0.030	0.025	0.052	0.053	0.054
PMSE_OLS	4.370	4.297	4.318	4.365	4.407	4.308	4.286	4.341	4.311
PMSE_UTMOST	3.989	4.699	7.742	8.051	11.046	16.202	8.838	10.380	17.867
PMSE_mEBmix	1.310	1.284	1.288	1.331	1.367	1.306	1.330	1.330	1.330

D.2. Simulation results in Setting 4 and comparisons with Mr.Mash. Table D.2 shows the simulation results under the Setting 4.

TABLE D.2

MSEs and PMSEs of OLS, UTMOST, Mr.Mash, and the proposed method under Setting 4. "MSE_OLS", "MSE_UTMOST", "MSE_Mash", "MSE_mEBmix" represent MSEs of the OLS, UTMOST, Mr.Mash, and the proposed method, respectively. "PMSE_OLS", "PMSE_UTMOST", "PMSE_Mash", "PMSE_mEBmix" represent PMSEs of the OLS, UTMOST, Mr.Mash, and the proposed method mEBmix, respectively. "SD" represents standard deviation.

Noise SD	$\sigma = 1$								
Correlation ρ	0			0.6			0.8		
Signal level β_S	0.5	1	2	0.5	1	2	0.5	1	2
MSE_OLS	0.047	0.048	0.046	0.048	0.048	0.047	0.047	0.047	0.048
MSE_UTMOST	0.027	0.064	0.140	0.051	0.063	0.142	0.062	0.082	0.122
MSE_Mash	0.041	0.105	0.272	0.096	0.115	0.237	0.103	0.127	0.225
MSE_mEBmix	0.003	0.005	0.004	0.005	0.006	0.005	0.005	0.005	0.005
PMSE_OLS	1.150	1.141	1.138	1.140	1.145	1.141	1.141	1.139	1.147
PMSE_UTMOST	1.220	2.266	5.854	3.689	4.028	8.649	5.084	6.093	7.710
PMSE_Mash	1.158	1.503	2.559	1.506	1.641	2.628	1.584	1.793	2.622
PMSE_mEBmix	1.010	1.012	1.008	1.012	1.016	1.013	1.015	1.012	1.011
Noise SD	$\sigma = 2$								
Correlation ρ	0			0.6			0.8		
Signal level β_S	0.5	1	2	0.5	1	2	0.5	1	2
MSE_OLS	0.196	0.191	0.194	0.185	0.188	0.196	0.184	0.189	0.193
MSE_UTMOST	0.065	0.076	0.106	0.115	0.088	0.131	0.085	0.102	0.127
MSE_Mash	0.085	0.097	0.140	0.175	0.155	0.199	0.114	0.147	0.192
MSE_mEBmix	0.019	0.016	0.014	0.025	0.023	0.018	0.020	0.022	0.015
PMSE_OLS	4.598	4.562	4.555	4.593	4.589	4.556	4.570	4.554	4.608
PMSE_UTMOST	4.858	5.494	7.696	11.203	8.556	10.767	9.528	10.341	12.011
PMSE_Mash	4.380	4.422	4.735	4.942	4.844	5.076	4.611	4.739	5.037
PMSE_mEBmix	4.050	4.042	4.035	4.086	4.086	4.007	4.094	4.052	4.040

SECTION E: ADDITIONAL REAL DATA RESULTS

E.1. Comparisons of PMSEs and correlations. Table E.1 shows the PMSEs of different methods for each issue.

Table E.2 shows the correlations of observed and predicted expressions by different methods for each issue.

TABLE E.1

PMSEs of the OLS, UTMOST, Mr.Mash, and mEBmix for each tissue. "OLS", "UTMOST", "Mash", and "mEBmix" represent the OLS, UTMOST, Mr.Mash, and the proposed method mEBmix, respectively.

Tissue	OLS	UTMOST	Mash	**mEBmix**
Adipose_Subcutaneous	0.962	0.963	0.953	0.949
Adipose_Visceral_Omentum	0.968	0.963	0.956	0.950
Adrenal_Gland	0.965	0.944	0.965	0.930
Artery_Aorta	0.960	0.955	0.949	0.939
Artery_Coronary	0.974	0.945	0.945	0.928
Artery_Tibial	0.963	0.963	0.955	0.951
Brain_Cerebellum	0.957	0.935	0.946	0.924
Brain_Cortex	0.970	0.941	0.950	0.930
Brain_Nucleus_accumbens_basal_ganglia	0.979	0.944	0.954	0.935
Breast_Mammary_Tissue	0.970	0.961	0.954	0.949
Cells_Cultured_fibroblasts	0.967	0.961	0.958	0.955
Colon_Sigmoid	0.965	0.953	0.954	0.937
Colon_Transverse	0.973	0.961	0.979	0.950
Esophagus_Gastroesophageal_Junction	0.961	0.952	0.956	0.935
Esophagus_Mucosa	0.966	0.963	0.961	0.952
Esophagus_Muscularis	0.960	0.958	0.964	0.943
Heart_Atrial_Appendage	0.964	0.956	0.951	0.942
Heart_Left_Ventricle	0.969	0.960	0.954	0.948
Liver	0.978	0.946	0.952	0.934
Lung	0.969	0.964	0.958	0.954
Muscle_Skeletal	0.970	0.969	0.963	0.961
Nerve_Tibial	0.957	0.958	0.948	0.944
Pancreas	0.958	0.948	0.969	0.934
Pituitary	0.970	0.946	0.964	0.934
Prostate	0.975	0.947	0.945	0.932
Skin_Not_Sun_Exposed_Suprapubic	0.962	0.961	0.955	0.951
Skin_Sun_Exposed_Lower_leg	0.965	0.965	0.958	0.955
Spleen	0.959	0.940	0.970	0.925
Stomach	0.973	0.958	0.979	0.947
Testis	0.971	0.957	0.957	0.950
Thyroid	0.959	0.961	0.951	0.947
Whole_Blood	0.972	0.969	0.966	0.964

E.2. Robustness to normality assumptions. Table E.3 shows average of PMSEs across genes whose expression values are not normally distributed in at least 20 tissues.

E.3. Comparison of computation time. Table E.4 shows the computation time (in seconds) of OLS, UTMOST, Mr.Mash, and mEBmix for one gene in GTEx data.

E.4. Posterior probability of cis-SNP gene expression association. In this subsection, we apply the proposed method to the whole dataset, and calculate the posterior probability of $I^{(t)} = 1$ for $t = 1, \ldots, m$ and all the genes. This posterior probability indicates whether the cis-SNPs of a given gene have effects on the corresponding gene expression in a particular tissue based on the observed data. We use the "gplots" R package (`https://cran.r-project.org/web/packages/gplots/index.html`) to generate a heat map for all the posterior probabilities, where each column represents a tissue and each row represents a gene. We first note that for many genes, the cis-SNP and gene expression associations are observed all the tissues (top blue rows). There are only a few genes that do not have their corresponding cis-SNPs in all the tissues (bottom red rows). For many other genes, we observe tissue-specific cis-SNP gene expression associations, but for most of these genes, the cis-SNP associations are observed in most of the tissues.

As shown in Figure E.1, the heat map clusters similar genes and similar tissues together based on the posterior probability of observing cis-SNP and gene expression association. We observe that the tissues that are clustered together are indeed For example, the "Esimilar. sophagus_Gastroesophageal_Junction" tissue and the "Esophagus_Muscularis" tissue are clustered together in Figure E.1. They are both related to the esophagus. Similarly, the

TABLE E.2

Correlations of the OLS, UTMOST, Mr.Mash, and mEBmix for each tissue. "OLS", "UTMOST", "Mash", and "Proposed" represent the OLS, UTMOST, Mr.Mash, and the proposed method mEBmix, respectively.

Tissue	OLS	UTMOST	Mash	**mEBmix**
Adipose_Subcutaneous	0.101	0.101	0.117	0.135
Adipose_Visceral_Omentum	0.081	0.086	0.101	0.123
Adrenal_Gland	0.065	0.083	0.082	0.125
Artery_Aorta	0.094	0.098	0.111	0.140
Artery_Coronary	0.046	0.066	0.078	0.126
Artery_Tibial	0.098	0.098	0.111	0.129
Brain_Cerebellum	0.070	0.092	0.077	0.114
Brain_Cortex	0.046	0.070	0.061	0.096
Brain_Nucleus_accumbens_basal_ganglia	0.024	0.051	0.039	0.078
Breast_Mammary_Tissue	0.067	0.075	0.089	0.116
Cells_Cultured_fibroblasts	0.062	0.075	0.065	0.078
Colon_Sigmoid	0.079	0.087	0.103	0.137
Colon_Transverse	0.052	0.064	0.076	0.103
Esophagus_Gastroesophageal_Junction	0.091	0.095	0.115	0.146
Esophagus_Mucosa	0.083	0.087	0.094	0.114
Esophagus_Muscularis	0.104	0.104	0.123	0.145
Heart_Atrial_Appendage	0.086	0.090	0.103	0.131
Heart_Left_Ventricle	0.071	0.078	0.091	0.116
Liver	0.032	0.056	0.056	0.099
Lung	0.075	0.080	0.092	0.112
Muscle_Skeletal	0.079	0.081	0.090	0.100
Nerve_Tibial	0.110	0.110	0.124	0.144
Pancreas	0.084	0.096	0.096	0.130
Pituitary	0.048	0.069	0.069	0.106
Prostate	0.040	0.062	0.069	0.118
Skin_Not_Sun_Exposed_Suprapubic	0.095	0.097	0.101	0.117
Skin_Sun_Exposed_Lower_leg	0.092	0.094	0.100	0.113
Spleen	0.074	0.092	0.088	0.133
Stomach	0.050	0.064	0.072	0.103
Testis	0.051	0.066	0.054	0.078
Thyroid	0.107	0.106	0.118	0.138
Whole_Blood	0.061	0.067	0.066	0.071

"Artery_Tibial", "Artery_Coronary", and "Artery_Aorta" tissues are all related to the artery and are clustered together in the heat map. Thus, the posterior probabilities based on the proposed method indeed capture the similarity between tissues in terms of the relationship between gene expression and cis-SNPs.

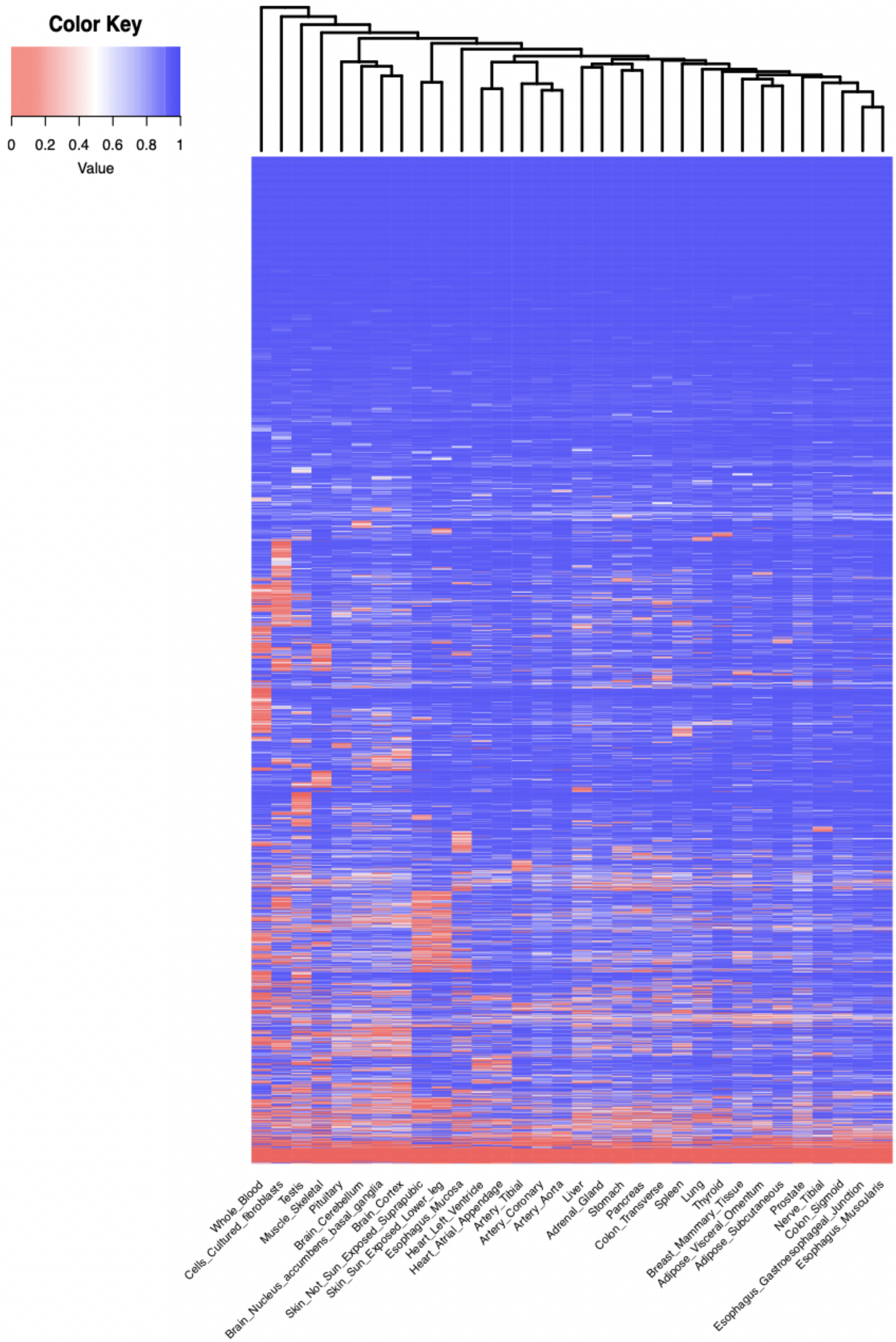

FIG E.1. *Posterior probabilities of $I^{(t)} = 1$ for all the genes and tissues.*

TABLE E.3

Average PMSEs across genes whose expression values are not normally distributed in at least 20 *tissues. "OLS", "UTMOST", "Mash", and "mEBmix" represent the OLS, UTMOST, Mr.Mash, and the proposed method mEBmix, respectively.*

Tissue	OLS	UTMOST	Mash	**mEBmix**
Adipose_Subcutaneous	0.773	0.804	0.762	0.756
Adipose_Visceral_Omentum	0.832	0.849	0.813	0.808
Adrenal_Gland	0.870	0.855	0.837	0.824
Artery_Aorta	0.784	0.806	0.764	0.758
Artery_Coronary	0.819	0.790	0.775	0.765
Artery_Tibial	0.741	0.760	0.725	0.722
Brain_Cerebellum	0.799	0.814	0.767	0.760
Brain_Cortex	0.745	0.751	0.719	0.702
Brain_Nucleus_accumbens_basal_ganglia	0.724	0.728	0.688	0.678
Breast_Mammary_Tissue	0.834	0.869	0.809	0.802
Cells_Cultured_fibroblasts	0.750	0.768	0.738	0.735
Colon_Sigmoid	0.800	0.810	0.770	0.762
Colon_Transverse	0.822	0.822	0.803	0.791
Esophagus_Gastroesophageal_Junction	0.785	0.787	0.762	0.751
Esophagus_Mucosa	0.831	0.854	0.815	0.812
Esophagus_Muscularis	0.787	0.807	0.774	0.763
Heart_Atrial_Appendage	0.798	0.804	0.774	0.771
Heart_Left_Ventricle	0.761	0.771	0.739	0.734
Liver	0.900	0.886	0.850	0.837
Lung	0.839	0.857	0.825	0.822
Muscle_Skeletal	0.757	0.787	0.746	0.745
Nerve_Tibial	0.737	0.770	0.728	0.723
Pancreas	0.870	0.874	0.840	0.827
Pituitary	0.833	0.877	0.810	0.790
Prostate	0.838	0.825	0.789	0.780
Skin_Not_Sun_Exposed_Suprapubic	0.774	0.804	0.761	0.753
Skin_Sun_Exposed_Lower_leg	0.746	0.778	0.735	0.731
Spleen	0.873	0.862	0.837	0.828
Stomach	0.860	0.864	0.839	0.830
Testis	0.926	0.930	0.905	0.890
Thyroid	0.797	0.821	0.783	0.780
Whole_Blood	0.793	0.819	0.779	0.777

TABLE E.4

Computation time (in seconds) of OLS, UTMOST, Mr.Mash, and mEBmix for one gene in GTEx data. "OLS", "UTMOST", "Mash", and "mEBmix" represent the OLS, UTMOST, Mr.Mash, and the proposed method mEBmix, respectively.

Number of SNPs	OLS	UTMOST	Mash	**mEBmix**
$p = 30$	0.035	31.773	146.986	40.741
$p = 50$	0.062	63.093	366.848	58.693
$p = 80$	0.126	94.922	404.225	61.245
$p = 105$	0.134	181.168	1186.156	79.368
$p = 125$	0.193	230.753	770.069	98.004