Supporting Information for "Gene-gene interaction analysis incorporating network information via a structured Bayesian approach"

by Xing Qin, Shuangge Ma, Mengyun Wu

The Supporting Information contains the specification of the posterior distribution and mathematical details of the variational Bayesian expectation-maximization algorithm for the proposed method in Section S1, and additional simulation and data analysis results in Section S2.

S1 The variational Bayesian expectation-maximization algorithm

In order to describe the entire computation process more clearly, we add the corresponding random variable to the symbolic representation of each distribution. For example, we re-express $p\left(\boldsymbol{y}|\boldsymbol{\beta}\right) = \mathcal{N}\left(\tilde{\boldsymbol{X}}\boldsymbol{\beta}, \tau^{-1}\boldsymbol{I}\right)$ as $p\left(\boldsymbol{y}|\boldsymbol{\beta}\right) = \mathcal{N}\left(\boldsymbol{y}|\tilde{\boldsymbol{X}}\boldsymbol{\beta}, \tau^{-1}\boldsymbol{I}\right)$.

In the following, we denote $\tilde{\boldsymbol{L}}_k = \boldsymbol{L}_k + \xi\boldsymbol{I} \triangleq \begin{pmatrix} \tilde{\boldsymbol{L}}_k^{(\mathbf{1})} & \mathbf{0}_{p_k\times\bar{p}_k} \\ \mathbf{0}_{\bar{p}_k\times p_k} & \tilde{\boldsymbol{L}}_k^{(\mathbf{2})} \end{pmatrix}$, and $\boldsymbol{\beta}_k^{(\mathbf{1})} = \left(\beta_j^{(\mathbf{1})}\right)_{\{j\in V_k\}}$, $\boldsymbol{\beta}_k^{(\mathbf{2})} = \left(\beta_{l_1 l_2}^{(\mathbf{2})}\right)_{\{l_1,l_2\in V_k, l_1<l_2\}}$, i.e., $\boldsymbol{\beta}_k^{(\mathbf{1})}$ and $\boldsymbol{\beta}_k^{(\mathbf{2})}$ consist of the coefficients for the main effects and interactions in the kth network, respectively. Here $\bar{p}_k = (p_k - 1)\,p_k/2$.

S1.1 Computation of the posterior distribution

We rewrite the following priors as the generative models with observation vector $\mathbf{0}$, i.e.,

$$\begin{aligned}
p\left(\bar{\beta}_{l_1l_2}^{(\mathbf{2})}|\gamma_{l_1}^{(\mathbf{1})}\gamma_{l_2}^{(\mathbf{1})}\right) &= \mathcal{N}\left(\bar{\beta}_{l_1l_2}^{(\mathbf{2})}|0, s_1\right)^{\gamma_{l_1}^{(\mathbf{1})}\gamma_{l_2}^{(\mathbf{1})}} \mathcal{N}\left(\bar{\beta}_{l_1l_2}^{(\mathbf{2})}|0, s_2\right)^{1-\gamma_{l_1}^{(\mathbf{1})}\gamma_{l_2}^{(\mathbf{1})}} \\
&= \mathcal{N}\left(0|\bar{\beta}_{l_1l_2}^{(\mathbf{2})}, s_1\right)^{\gamma_{l_1}^{(\mathbf{1})}\gamma_{l_2}^{(\mathbf{1})}} \mathcal{N}\left(0|\bar{\beta}_{l_1l_2}^{(\mathbf{2})}, s_2\right)^{1-\gamma_{l_1}^{(\mathbf{1})}\gamma_{l_2}^{(\mathbf{1})}} \\
&= p\left(0|\bar{\beta}_{l_1l_2}^{(\mathbf{2})}, \gamma_{l_1}^{(\mathbf{1})}\gamma_{l_2}^{(\mathbf{1})}\right), \\
p\left(\tilde{\boldsymbol{\beta}}_k|\alpha_k\right) &= \mathcal{N}\left(\tilde{\boldsymbol{\beta}}_k|\mathbf{0}, s_1\tilde{\boldsymbol{L}}_k^{-1}\right)^{\alpha_k} \mathcal{N}\left(\tilde{\boldsymbol{\beta}}_k|\mathbf{0}, s_2\boldsymbol{I}\right)^{1-\alpha_k} \\
&= \mathcal{N}\left(\mathbf{0}|\tilde{\boldsymbol{\beta}}_k, s_1\tilde{\boldsymbol{L}}_k^{-1}\right)^{\alpha_k} \mathcal{N}\left(\mathbf{0}|\tilde{\boldsymbol{\beta}}_k, s_2\boldsymbol{I}\right)^{1-\alpha_k} \\
&= p\left(\mathbf{0}|\tilde{\boldsymbol{\beta}}_k, \alpha_k\right).
\end{aligned}$$

As shown above, $p\left(\bar{\beta}_{l_1l_2}^{(\mathbf{2})}|\gamma_{l_1}^{(\mathbf{1})}\gamma_{l_2}^{(\mathbf{1})}\right)$ has the same form as $p\left(0|\bar{\beta}_{l_1l_2}^{(\mathbf{2})}, \gamma_{l_1}^{(\mathbf{1})}\gamma_{l_2}^{(\mathbf{1})}\right)$, and $p\left(\tilde{\boldsymbol{\beta}}_k|\alpha_k\right)$ has the same form as $p\left(\mathbf{0}|\tilde{\boldsymbol{\beta}}_k, \alpha_k\right)$, where we can view the observation vector $\mathbf{0}$ as samples drawn from

$\bar{\beta}^{(\mathbf{2})}_{l_1 l_2}$ and $\tilde{\boldsymbol{\beta}}_k$. Thus we obtain

$$\begin{aligned}
& p\left(\boldsymbol{y}, \boldsymbol{\beta}, \tilde{\boldsymbol{\beta}}, \bar{\boldsymbol{\beta}}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\zeta} | \boldsymbol{X}; \tau, \theta\right) \\
=& \mathcal{N}\left(\boldsymbol{y} | \tilde{\boldsymbol{X}} \boldsymbol{\beta}, \tau^{-1} \boldsymbol{I}\right) \prod_{j=1}^{p(p+1)/2} \mathcal{N}\left(\beta_j | 0, s_1\right)^{\gamma_j} \mathcal{N}\left(\beta_j | 0, s_2\right)^{1-\gamma_j} \operatorname{Bern}\left(\gamma_j | \zeta_j\right) \operatorname{Beta}\left(\zeta_j | a, b\right) \\
& \prod_{k=1}^{K} \left\{ p\left(\tilde{\boldsymbol{\beta}}_k | \boldsymbol{\beta}_k\right) \mathcal{N}\left(\mathbf{0} | \tilde{\boldsymbol{\beta}}_k, s_1 \tilde{\boldsymbol{L}}_k^{-1}\right)^{\alpha_k} \mathcal{N}\left(\mathbf{0} | \tilde{\boldsymbol{\beta}}_k, s_2 \boldsymbol{I}\right)^{1-\alpha_k} \operatorname{Bern}\left(\alpha_k | \theta\right) \right\} \prod_{l_1=1}^{p} \prod_{l_2 > l_1}^{p} \left[p\left(\bar{\beta}^{(\mathbf{2})}_{l_1 l_2} | \beta^{(\mathbf{2})}_{l_1 l_2}\right) \right. \\
& \left. p\left(0 | \bar{\beta}^{(\mathbf{2})}_{l_1 l_2}, \gamma^{(\mathbf{1})}_{l_1} \gamma^{(\mathbf{1})}_{l_2}\right) \right].
\end{aligned}$$

With the assumption of the same prior distribution for $\boldsymbol{\beta}$, $\bar{\boldsymbol{\beta}}$, and $\tilde{\boldsymbol{\beta}}$, i.e., $p\left(\tilde{\boldsymbol{\beta}}_k | \boldsymbol{\beta}_k\right) = \mathbf{1}_{\{\tilde{\boldsymbol{\beta}}_k = \boldsymbol{\beta}_k\}}$ and $p\left(\bar{\beta}^{(\mathbf{2})}_{l_1 l_2} | \beta^{(\mathbf{2})}_{l_1 l_2}\right) = \mathbf{1}_{\left\{\bar{\beta}^{(\mathbf{2})}_{l_1 l_2} = \beta^{(\mathbf{2})}_{l_1 l_2}\right\}}$, we can replace $\bar{\beta}^{(\mathbf{2})}_{l_1 l_2}$ (or $\tilde{\boldsymbol{\beta}}_k$) with $\beta^{(\mathbf{2})}_{l_1 l_2}$ (or $\boldsymbol{\beta}_k$). So the parameters' posterior distribution can be written as

$$\begin{aligned}
& p\left(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\zeta} | \boldsymbol{y}, \boldsymbol{X}; \tau, \theta\right) \\
=& \frac{1}{C} \mathcal{N}\left(\boldsymbol{y} | \tilde{\boldsymbol{X}} \boldsymbol{\beta}, \tau^{-1} \boldsymbol{I}\right) \prod_{j=1}^{p(p+1)/2} \mathcal{N}\left(\beta_j | 0, s_1\right)^{\gamma_j} \mathcal{N}\left(\beta_j | 0, s_2\right)^{1-\gamma_j} \operatorname{Bern}\left(\gamma_j | \zeta_j\right) \operatorname{Beta}\left(\zeta_j | a, b\right) \\
& \prod_{k=1}^{K} \left\{ \mathcal{N}\left(\boldsymbol{\beta}_k | \mathbf{0}, s_1 \tilde{\boldsymbol{L}}_k^{-1}\right)^{\alpha_k} \mathcal{N}\left(\boldsymbol{\beta}_k | \mathbf{0}, s_2 \boldsymbol{I}\right)^{1-\alpha_k} \operatorname{Bern}\left(\alpha_k | \theta\right) \right\} \\
& \prod_{l_1=1}^{p} \prod_{l_2 > l_1}^{p} \mathcal{N}\left(\beta_{l_1 l_2} | 0, s_1\right)^{\gamma_{l_1} \gamma_{l_2}} \mathcal{N}\left(\beta_{l_1 l_2} | 0, s_2\right)^{1-\gamma_{l_1} \gamma_{l_2}},
\end{aligned}$$

where C is the normalization constant.

S1.2 Graphical representation of the proposed hybrid Bayesian approach

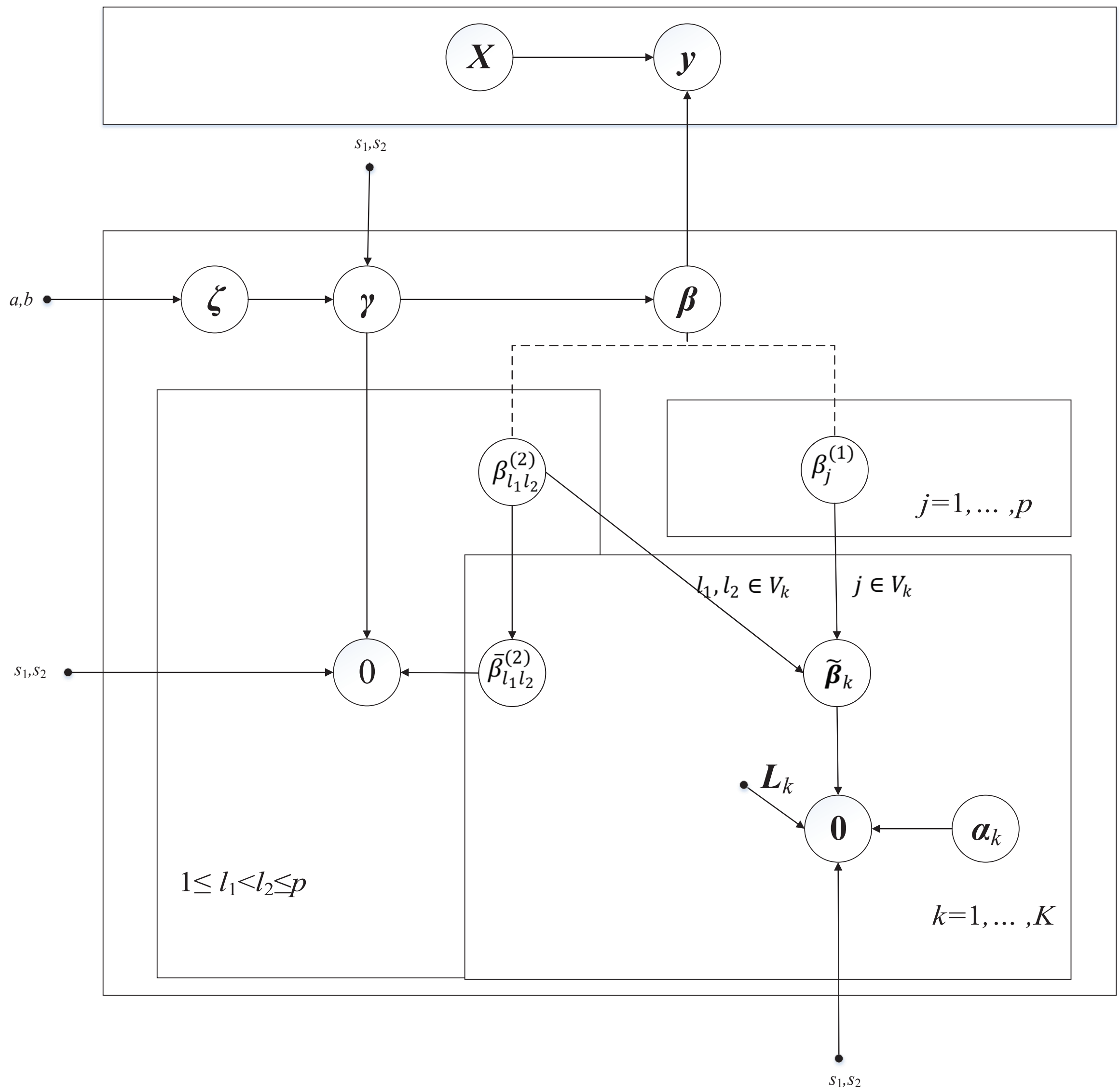

Figure S1: Graphical representation of the proposed hybrid Bayesian approach.

S1.3 Derivation of the optimal variational distribution

As stated in the main text, we denote $q(\Omega)$ as a candidate approximating distribution of the true posterior distribution. For simplicity, denote $\mathrm{E}(\cdot)$ as the expectation taken with respect to $q(\Omega)$, and the logarithm of the marginal likelihood can be decomposed as

$$\log p(\boldsymbol{y}|\boldsymbol{X};\tau,\theta) = \mathcal{L}_q + \mathrm{KL}\left(q(\Omega)\|p(\Omega|\boldsymbol{y},\boldsymbol{X};\tau,\theta)\right),$$

where

$$\mathcal{L}_q = \mathrm{E}\log\left[\frac{p(\boldsymbol{y},\Omega|\boldsymbol{X};\tau,\theta)}{q(\Omega)}\right],$$

$$\mathrm{KL}\left(q(\Omega)\|p(\Omega|\boldsymbol{y},\boldsymbol{X};\tau,\theta)\right) = \mathrm{E}\log\left[\frac{q(\Omega)}{p(\Omega|\boldsymbol{y},\boldsymbol{X};\tau,\theta)}\right] = \int q(\Omega)\log\left[\frac{q(\Omega)}{p(\Omega|\boldsymbol{y},\boldsymbol{X};\tau,\theta)}\right]d\Omega.$$

Minimizing the KL divergence with respect to the approximating distribution $q(\Omega)$ is equivalent to maximizing the evidence lower bound $\mathcal{L}_q$. In the expectation (E) step, we optimize the KL divergence ($\mathcal{L}_q$) with respect to the variational parameters Ω while holding the model parameters τ and θ fixed. In the maximization (M) step, we optimize the KL divergence with respect to the model parameters while keeping the variational parameters fixed.

S1.3.1 E-step

Notice that the logarithm of the joint probability function for $\Omega, \boldsymbol{y}$ can be written as

$$\begin{aligned}
&\log p(\boldsymbol{y},\Omega|\boldsymbol{X};\tau,\theta)\\
=&\log\Bigg\{\mathcal{N}\left(\boldsymbol{y}|\tilde{\boldsymbol{X}}\boldsymbol{\beta},\tau^{-1}\boldsymbol{I}\right)\prod_{j=1}^{p(p+1)/2}\mathcal{N}(\beta_j|0,s_1)^{\gamma_j}\mathcal{N}(\beta_j|0,s_2)^{1-\gamma_j}\,\mathrm{Bern}(\gamma_j|\zeta_j)\,\mathrm{Beta}(\zeta_j|a,b)\\
&\prod_{k=1}^{K}\left\{\mathcal{N}\left(\boldsymbol{\beta}_k|\boldsymbol{0},s_1\tilde{\boldsymbol{L}}_k^{-1}\right)^{\alpha_k}\mathcal{N}(\boldsymbol{\beta}_k|\boldsymbol{0},s_2\boldsymbol{I})^{1-\alpha_k}\,\mathrm{Bern}(\alpha_k|\theta)\right\}\\
&\prod_{l_1=1}^{p}\prod_{l_2>l_1}^{p}\mathcal{N}(\beta_{l_1l_2}|0,s_1)^{\gamma_{l_1}\gamma_{l_2}}\mathcal{N}(\beta_{l_1l_2}|0,s_2)^{1-\gamma_{l_1}\gamma_{l_2}}\Bigg\}.
\end{aligned}$$

We consider the variational distribution with the form

$$q(\Omega) = q(\boldsymbol{\beta})\,q(\boldsymbol{\gamma})\,q(\boldsymbol{\alpha})\,q(\boldsymbol{\zeta}).$$

According to the general formula of the variational method, the optimal solution of $q(\Omega)$ is computed as follows.
1. For $q(\boldsymbol{\beta})$, to avoid high-dimensional covariance matrix operations, we assume $q(\boldsymbol{\beta}) = \prod_{j=1}^{p} q\left(\beta_j^{(\mathbf{1})}\right)$ $\prod_{l_1=1}^{p}\prod_{l_2>l_1}^{p} q\left(\beta_{l_1l_2}^{(\mathbf{2})}\right)$, where $q(\cdot)$ denotes the variational distribution of the corresponding parameter.

(a) For $q\left(\beta_j^{(\mathbf{1})}\right)$, for simplicity, assume $j \in V_{k_j}$, i.e., the jth main effect belongs to the k_jth network. We have

$$
\begin{aligned}
\log q\left(\beta_j^{(\mathbf{1})}\right) &= \mathrm{E}_{\left(-\beta_j^{(\mathbf{1})}\right)}\left[\log p\left(\boldsymbol{y}, \Omega | \boldsymbol{X}; \tau, \theta\right)\right] + \text{const} \\
&= \mathrm{E}_{\left(-\beta_j^{(\mathbf{1})}\right)}\left\{\log \mathcal{N}\left(\boldsymbol{y} | \tilde{\boldsymbol{X}}\boldsymbol{\beta}, \tau^{-1}\boldsymbol{I}\right) + \log p\left(\beta_j^{(\mathbf{1})} | \gamma_j^{(\mathbf{1})}\right) + \log p\left(\boldsymbol{\beta}_{k_j} | \alpha_{k_j}\right)\right\} + \text{const} \\
&= \mathrm{E}_{\left(-\beta_j^{(\mathbf{1})}\right)}\left\{-\frac{\tau}{2}\left(\boldsymbol{y} - \tilde{\boldsymbol{X}}\boldsymbol{\beta}\right)^{\mathrm{T}}\left(\boldsymbol{y} - \tilde{\boldsymbol{X}}\boldsymbol{\beta}\right) - \frac{\left(\beta_j^{(\mathbf{1})}\right)^2}{2s_1}\gamma_j^{(\mathbf{1})} - \frac{\left(\beta_j^{(\mathbf{1})}\right)^2}{2s_2}\left(1 - \gamma_j^{(\mathbf{1})}\right)\right. \\
&\quad \left. - \left(\boldsymbol{\beta}_{k_j}^{(\mathbf{1})}\right)^{\mathrm{T}}\left(\frac{\alpha_{k_j}}{2s_1}\tilde{\boldsymbol{L}}_{k_j}^{(\mathbf{1})} + \frac{1 - \alpha_{k_j}}{2s_2}\boldsymbol{I}\right)\boldsymbol{\beta}_{k_j}^{(\mathbf{1})}\right\} + \text{const} \\
&= -\frac{\tau}{2}\sum_i\left(\left\langle res_i^{(-j)}\right\rangle - x_{ij}\beta_j^{(\mathbf{1})}\right)^2 - \frac{\left(\beta_j^{(\mathbf{1})}\right)^2}{2}\left[\frac{\left\langle\gamma_j^{(\mathbf{1})}\right\rangle}{s_1} + \frac{1 - \left\langle\gamma_j^{(\mathbf{1})}\right\rangle}{s_2}\right] \\
&\quad - \mathrm{E}_{\left(-\beta_j^{(\mathbf{1})}\right)}\left[\left(\boldsymbol{\beta}_{k_j}^{(\mathbf{1})}\right)^{\mathrm{T}}\left(\frac{\alpha_{k_j}}{2s_1}\tilde{\boldsymbol{L}}_{k_j}^{(\mathbf{1})} + \frac{1 - \alpha_{k_j}}{2s_2}\boldsymbol{I}\right)\boldsymbol{\beta}_{k_j}^{(\mathbf{1})}\right] + \text{const},
\end{aligned}
\tag{S1}
$$

where $\mathrm{E}_{\left(-\beta_j^{(\mathbf{1})}\right)}(\cdot)$ denotes the expectation taken with respect to the approximate posterior distribution over Ω except $\beta_j^{(\mathbf{1})}$, $\langle\cdot\rangle$ denotes the expectation over the corresponding variational distribution, $\tilde{\boldsymbol{x}}_i$ is the ith row of $\tilde{\boldsymbol{X}}$, and $res_i^{(-j)} = y_i - \tilde{\boldsymbol{x}}_i\boldsymbol{\beta} + x_{ij}\beta_j^{(\mathbf{1})}$. For convenience, we rewrite

$$
\underline{\boldsymbol{B}}_{k_j}^{(\mathbf{1})} =: \frac{\alpha_{k_j}}{s_1}\tilde{\boldsymbol{L}}_{k_j}^{(\mathbf{1})} + \frac{1 - \alpha_{k_j}}{s_2}\boldsymbol{I} \triangleq \left(\underline{b}_{lj}^{(k_j)}\right)_{p_{k_j} \times p_{k_j}}.
$$

Note that $\underline{\boldsymbol{B}}_{k_j}^{(\mathbf{1})}$ is a symmetric matrix, and the right side of (S1) can be written as

$$
\begin{aligned}
\log q\left(\beta_j^{(\mathbf{1})}\right) &= -\frac{\tau}{2}\sum_i \left(\left\langle res_i^{(-j)}\right\rangle - x_{ij}\beta_j^{(\mathbf{1})}\right)^2 - \frac{1}{2}\mathrm{E}_{\left(-\beta_j^{(\mathbf{1})}\right)}\left(\left(\boldsymbol{\beta}_{k_j}^{(\mathbf{1})}\right)^{\mathrm{T}} \underline{\boldsymbol{B}}_{k_j}^{(\mathbf{1})}\boldsymbol{\beta}_{k_j}^{(\mathbf{1})}\right) \\
&\quad - \frac{\left(\beta_j^{(\mathbf{1})}\right)^2}{2}\left[\frac{\left\langle \gamma_j^{(\mathbf{1})}\right\rangle}{s_1} + \frac{1-\left\langle \gamma_j^{(\mathbf{1})}\right\rangle}{s_2}\right] + \mathrm{const} \\
&= -\frac{\tau}{2}\sum_i \left(\left\langle res_i^{(-j)}\right\rangle - x_{ij}\beta_j^{(\mathbf{1})}\right)^2 - \frac{1}{2}\left\langle \underline{b}_{jj}^{(k_j)}\right\rangle \left(\beta_j^{(\mathbf{1})}\right)^2 - \beta_j^{(\mathbf{1})}\sum_{l\neq j, k_l=k_j}\left\langle \underline{b}_{lj}^{(k_j)}\right\rangle\left\langle \beta_l^{(\mathbf{1})}\right\rangle \\
&\quad - \frac{\left(\beta_j^{(\mathbf{1})}\right)^2}{2}\left[\frac{\left\langle \gamma_j^{(\mathbf{1})}\right\rangle}{s_1} + \frac{1-\left\langle \gamma_j^{(\mathbf{1})}\right\rangle}{s_2}\right] + \mathrm{const} \\
&= \tau\sum_i\left[\left\langle res_i^{(-j)}\right\rangle x_{ij}\beta_j^{(\mathbf{1})} - \frac{\left(x_{ij}\beta_j^{(\mathbf{1})}\right)^2}{2}\right] - \frac{1}{2}\left\langle \underline{b}_{jj}^{(k_j)}\right\rangle \left(\beta_j^{(\mathbf{1})}\right)^2 - \beta_j^{(\mathbf{1})}\sum_{l\neq j, k_l=k_j}\left\langle \underline{b}_{lj}^{(k_j)}\right\rangle\left\langle \beta_l^{(\mathbf{1})}\right\rangle \\
&\quad - \frac{\left(\beta_j^{(\mathbf{1})}\right)^2}{2}\left[\frac{\left\langle \gamma_j^{(\mathbf{1})}\right\rangle}{s_1} + \frac{1-\left\langle \gamma_j^{(\mathbf{1})}\right\rangle}{s_2}\right] + \mathrm{const} \\
&= -\frac{\frac{\left\langle \gamma_j^{(\mathbf{1})}\right\rangle}{s_1} + \frac{1-\left\langle \gamma_j^{(\mathbf{1})}\right\rangle}{s_2} + \left\langle \underline{b}_{jj}^{(k_j)}\right\rangle + \tau\sum_i (x_{ij})^2}{2}\left(\beta_j^{(\mathbf{1})}\right)^2 + \beta_j^{(\mathbf{1})}\left[\tau\sum_i\left\langle res_i^{(-j)}\right\rangle x_{ij}\right. \\
&\quad \left. - \sum_{l\neq j, k_l=k_j}\left\langle \underline{b}_{lj}^{(k_j)}\right\rangle\left\langle \beta_l^{(\mathbf{1})}\right\rangle\right] + \mathrm{const}.
\end{aligned}
$$

Thus, we have $q\left(\beta_j^{(\mathbf{1})}\right) = \mathcal{N}\left(\beta_j^{(\mathbf{1})}|m_j^{(\mathbf{1})}, \left(\sigma_j^{(\mathbf{1})}\right)^2\right)$ with the parameters

$$
\begin{cases}
\left(\sigma_j^{(\mathbf{1})}\right)^2 = \left(\frac{\left\langle \gamma_j^{(\mathbf{1})}\right\rangle}{s_1} + \frac{1-\left\langle \gamma_j^{(\mathbf{1})}\right\rangle}{s_2} + \left\langle \underline{b}_{jj}^{(k_j)}\right\rangle + \tau\sum_i (x_{ij})^2\right)^{-1}, \\
m_j^{(\mathbf{1})} = \left(\sigma_j^{(\mathbf{1})}\right)^2\left(\tau\sum_i x_{ij}\left\langle res_i^{(-j)}\right\rangle - \sum_{l\neq j, k_l=k_j}\left\langle \underline{b}_{lj}^{(k_j)}\right\rangle\left\langle \beta_l^{(\mathbf{1})}\right\rangle\right).
\end{cases}
$$

(b) For $q\left(\beta^{(\mathbf{2})}_{l_1l_2}\right), k_{l_1}=k_{l_2}\triangleq k$, i.e., for coefficients of the interactions within the networks, we have

$$\begin{aligned}
\log q\left(\beta^{(\mathbf{2})}_{l_1l_2}\right) &= \mathrm{E}_{\left(-\beta^{(\mathbf{2})}_{l_1l_2}\right)}\left[\log p\left(\boldsymbol{y},\Omega|\boldsymbol{X};\tau,\theta\right)\right]+\text{const}\\
&= \mathrm{E}_{\left(-\beta^{(\mathbf{2})}_{l_1l_2}\right)}\left\{\log\mathcal{N}\left(\boldsymbol{y}|\tilde{\boldsymbol{X}}\boldsymbol{\beta},\tau^{-1}\boldsymbol{I}\right)+\log p\left(\beta^{(\mathbf{2})}_{l_1l_2}|\gamma^{(\mathbf{2})}_{l_1l_2}\right)+\log p\left(\boldsymbol{\beta}_k|\alpha_k\right)\right.\\
&\quad\left.+\log p\left(\beta^{(\mathbf{2})}_{l_1l_2}|\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right)\right\}+\text{const}\\
&= \mathrm{E}_{\left(-\beta^{(\mathbf{2})}_{l_1l_2}\right)}\left\{-\frac{\tau}{2}\left(\boldsymbol{y}-\tilde{\boldsymbol{X}}\boldsymbol{\beta}\right)^{\mathrm{T}}\left(\boldsymbol{y}-\tilde{\boldsymbol{X}}\boldsymbol{\beta}\right)-\frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_1}\gamma^{(\mathbf{2})}_{l_1l_2}\right.\\
&\quad-\frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_2}\left(1-\gamma^{(\mathbf{2})}_{l_1l_2}\right)-\frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_1}\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}-\frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_2}\left(1-\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right)\\
&\quad\left.-\left(\boldsymbol{\beta}^{(\mathbf{2})}_k\right)^{\mathrm{T}}\left(\frac{\alpha_k}{2s_1}\tilde{\boldsymbol{L}}^{(\mathbf{2})}_k+\frac{1-\alpha_k}{2s_2}\boldsymbol{I}\right)\boldsymbol{\beta}^{(\mathbf{2})}_k\right\}+\text{const}\\
&= -\frac{\tau}{2}\sum_i\left(\left\langle res_i^{(-l_1l_2)}\right\rangle-x_{il_1}x_{il_2}\beta^{(\mathbf{2})}_{l_1l_2}\right)^2-\frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2}\left[\frac{\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle}{s_1}\right.\\
&\quad\left.+\frac{1-\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle}{s_2}+\frac{\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_1}+\frac{1-\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_2}\right]\\
&\quad-\mathrm{E}_{\left(-\beta^{(\mathbf{2})}_{l_1l_2}\right)}\left[\left(\boldsymbol{\beta}^{(\mathbf{2})}_k\right)^{\mathrm{T}}\left(\frac{\alpha_k}{2s_1}\tilde{\boldsymbol{L}}^{(\mathbf{2})}_k+\frac{1-\alpha_k}{2s_2}\boldsymbol{I}\right)\boldsymbol{\beta}^{(\mathbf{2})}_k\right]+\text{const},
\end{aligned}\tag{S2}$$

where $res_i^{(-l_1l_2)}=y_i-\tilde{\boldsymbol{x}}_i\boldsymbol{\beta}+x_{il_1}x_{il_2}\beta^{(\mathbf{2})}_{l_1l_2}$. Rewrite

$$\underline{\boldsymbol{B}}^{(\mathbf{2})}_k=:\frac{\alpha_k}{s_1}\tilde{\boldsymbol{L}}^{(\mathbf{2})}_k+\frac{1-\alpha_k}{s_2}\boldsymbol{I}\triangleq\left(\hat{b}^{(k)}_{l_1\sim j_1,l_2\sim j_2}\right)_{\bar{p}_k\times\bar{p}_k},$$

where $\hat{b}^{(k)}_{l_1\sim j_1,l_2\sim j_2}$ denotes the corresponding element for the interactions indexed by l_1j_1 and l_2j_2.

The right side of (S2) can be written as

$$
\begin{aligned}
&-\frac{\tau}{2}\sum_i \left(\left\langle res_i^{(-l_1l_2)}\right\rangle - x_{il_1}x_{il_2}\beta^{(\mathbf{2})}_{l_1l_2}\right)^2 - \frac{1}{2}\mathrm{E}_{\left(-\beta^{(\mathbf{2})}_{l_1l_2}\right)}\left(\left(\boldsymbol{\beta}^{(\mathbf{2})}_k\right)^{\mathrm{T}} \underline{\boldsymbol{B}}^{(\mathbf{2})}_k \boldsymbol{\beta}^{(\mathbf{2})}_k\right) \\
&-\frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2}\left[\frac{\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle}{s_1} + \frac{1-\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle}{s_2} + \frac{\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_1} + \frac{1-\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_2}\right] + \text{const} \\
=&-\frac{\tau}{2}\sum_i \left(\left\langle res_i^{(-l_1l_2)}\right\rangle - x_{il_1}x_{il_2}\beta^{(\mathbf{2})}_{l_1l_2}\right)^2 - \beta^{(\mathbf{2})}_{l_1l_2}\sum_{\substack{(j',l')\neq(l_1,l_2)\\ k_{j'}=k_{l'}=k}}\left\langle \hat{b}^{(k)}_{j'\sim l',l_1\sim l_2}\right\rangle\left\langle\beta^{(\mathbf{2})}_{j'l'}\right\rangle \\
&-\frac{1}{2}\left\langle \hat{b}^{(k)}_{l_1\sim l_2,l_1\sim l_2}\right\rangle\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2 - \frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2}\left[\frac{2-\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle-\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_2} + \frac{\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle+\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle}{s_1}\right] \\
&+\text{const} \\
=&\,\tau\sum_i\left[\left\langle res_i^{(-l_1l_2)}\right\rangle x_{il_1}x_{il_2}\beta^{(\mathbf{2})}_{l_1l_2} - \frac{\left(x_{il_1}x_{il_2}\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2}\right] - \beta^{(\mathbf{2})}_{l_1l_2}\sum_{\substack{(j',l')\neq(l_1,l_2)\\ k_{j'}=k_{l'}=k}}\left\langle \hat{b}^{(k)}_{j'\sim l',l_1\sim l_2}\right\rangle\left\langle\beta^{(\mathbf{2})}_{j'l'}\right\rangle \\
&-\frac{1}{2}\left\langle \hat{b}^{(k)}_{l_1\sim l_2,l_1\sim l_2}\right\rangle\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2 - \frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2}\left[\frac{2-\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle-\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_2} + \frac{\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle+\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle}{s_1}\right] \\
&+\text{const} \\
=&-\frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2}\left[\left(2-\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle-\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle\right)s_2^{-1} + \left(\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle+\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle\right)s_1^{-1} + \left\langle \hat{b}^{(k)}_{l_1\sim l_2,l_1\sim l_2}\right\rangle\right. \\
&\left.+\tau\sum_i\left(x_{il_1}x_{il_2}\right)^2\right] + \beta^{(\mathbf{2})}_{l_1l_2}\left[\tau\sum_i\left\langle res_i^{(-l_1l_2)}\right\rangle x_{il_1}x_{il_2} - \sum_{\substack{(j',l')\neq(l_1,l_2)\\ k_{j'}=k_{l'}=k}}\left\langle \hat{b}^{(k)}_{j'\sim l',l_1\sim l_2}\right\rangle\left\langle\beta^{(\mathbf{2})}_{j'l'}\right\rangle\right] \\
&+\text{const}.
\end{aligned}
$$

Thus, for $k_{l_1} = k_{l_2} = k$, we have $q\left(\beta^{(\mathbf{2})}_{l_1l_2}\right) = \mathcal{N}\left(\beta^{(\mathbf{2})}_{l_1l_2}|m^{(\mathbf{2})}_{l_1l_2}, \left(\sigma^{(\mathbf{2})}_{l_1l_2}\right)^2\right)$ with the parameters

$$
\begin{cases}
\left(\sigma^{(\mathbf{2})}_{l_1l_2}\right)^2 = \left(\frac{2-\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle-\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_2} + \frac{\left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle+\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle}{s_1} + \left\langle \hat{b}^{(k)}_{l_1\sim l_2,l_1\sim l_2}\right\rangle + \tau\sum_i\left(x_{il_1}x_{il_2}\right)^2\right)^{-1}, \\
m^{(\mathbf{2})}_{l_1l_2} = \left(\sigma^{(\mathbf{2})}_{l_1l_2}\right)^2\left(\tau\sum_i x_{il_1}x_{il_2}\left\langle res_i^{(-l_1l_2)}\right\rangle - \sum_{\substack{(j',l')\neq(l_1,l_2)\\ k_{j'}=k_{l'}=k}}\left\langle \hat{b}^{(k)}_{j'\sim l',l_1\sim l_2}\right\rangle\left\langle\beta^{(\mathbf{2})}_{j'l'}\right\rangle\right).
\end{cases}
$$

(c) For $q\left(\beta^{(\mathbf{2})}_{l_1l_2}\right), k_{l_1} \neq k_{l_2}$, i.e., for coefficients of the interactions accross the networks, we have

$$
\begin{aligned}
\log q\left(\beta^{(\mathbf{2})}_{l_1l_2}\right) &= \mathrm{E}_{\left(-\beta^{(\mathbf{2})}_{l_1l_2}\right)}\left[\log p\left(\boldsymbol{y}, \Omega | \boldsymbol{X}; \tau, \theta\right)\right] + \text{const} \\
&= \mathrm{E}_{\left(-\beta^{(\mathbf{2})}_{l_1l_2}\right)}\left\{\mathcal{N}\left(\boldsymbol{y}|\tilde{\boldsymbol{X}}\boldsymbol{\beta}, \tau^{-1}\boldsymbol{I}\right) + \log p\left(\beta^{(\mathbf{2})}_{l_1l_2}|\gamma^{(\mathbf{2})}_{l_1l_2}\right)\right. \\
&\quad \left. + \log p\left(\beta^{(\mathbf{2})}_{l_1l_2}|\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right)\right\} + \text{const} \\
&= \mathrm{E}_{\left(-\beta^{(\mathbf{2})}_{l_1l_2}\right)}\left\{-\frac{\tau}{2}\left(\boldsymbol{y} - \tilde{\boldsymbol{X}}\boldsymbol{\beta}\right)^{\mathrm{T}}\left(\boldsymbol{y} - \tilde{\boldsymbol{X}}\boldsymbol{\beta}\right) - \frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_1}\gamma^{(\mathbf{2})}_{l_1l_2}\right. \\
&\quad \left. -\frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_2}\left(1 - \gamma^{(\mathbf{2})}_{l_1l_2}\right) - \frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_1}\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2} - \frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_2}\left(1 - \gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right)\right\} \\
&\quad + \text{const} \\
&= \mathrm{E}_{\left(-\beta^{(\mathbf{2})}_{l_1l_2}\right)}\left[-\frac{\tau}{2}\sum_i\left(res_i^{(-l_1l_2)} - x_{il_1}x_{il_2}\beta^{(\mathbf{2})}_{l_1l_2}\right)^2 - \frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_1}\gamma^{(\mathbf{2})}_{l_1l_2}\right. \\
&\quad \left. -\frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_1}\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2} - \frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2s_2}\left(2 - \gamma^{(\mathbf{2})}_{l_1l_2} - \gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right)\right] + \text{const} \\
&= \tau\beta^{(\mathbf{2})}_{l_1l_2}\sum_i\left\langle res_i^{(-l_1l_2)}\right\rangle x_{il_1}x_{il_2} - \frac{\left(\beta^{(\mathbf{2})}_{l_1l_2}\right)^2}{2}\left[\frac{2 - \left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle - \left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_2}\right. \\
&\quad \left. +\frac{\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle + \left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_1} + \tau\sum_i\left(x_{il_1}x_{il_2}\right)^2\right] + \text{const},
\end{aligned}
$$

where $res_i^{(-l_1l_2)} = y_i - \tilde{\boldsymbol{x}}_i\boldsymbol{\beta} + x_{il_1}x_{il_2}\beta^{(\mathbf{2})}_{l_1l_2}$. Thus, $q\left(\beta^{(\mathbf{2})}_{l_1l_2}\right) = \mathcal{N}\left(\beta^{(\mathbf{2})}_{l_1l_2}|m^{(\mathbf{2})}_{l_1l_2}, \left(\sigma^{(\mathbf{2})}_{l_1l_2}\right)^2\right)$ with the parameters

$$
\begin{cases}
\left(\sigma^{(\mathbf{2})}_{l_1l_2}\right)^2 = \left(\tau\sum_i\left(x_{il_1}x_{il_2}\right)^2 + \frac{\left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle + \left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_1} + \frac{2 - \left\langle\gamma^{(\mathbf{2})}_{l_1l_2}\right\rangle - \left\langle\gamma^{(\mathbf{1})}_{l_1}\gamma^{(\mathbf{1})}_{l_2}\right\rangle}{s_2}\right)^{-1}, \\
m^{(\mathbf{2})}_{l_1l_2} = \left(\sigma^{(\mathbf{2})}_{l_1l_2}\right)^2\tau\sum_i x_{il_1}x_{il_2}\left\langle res_i^{(-l_1l_2)}\right\rangle.
\end{cases}
$$

According to (a), (b) and (c), we conclude that

$$
Q(\boldsymbol{\beta}) = \prod_{j=1}^{p}\mathcal{N}\left(\beta^{(\mathbf{1})}_j|m^{(\mathbf{1})}_j, \left(\sigma^{(\mathbf{1})}_j\right)^2\right)\prod_{l_1=1}^{p}\prod_{l_2>l_1}^{p}\mathcal{N}\left(\beta^{(\mathbf{2})}_{l_1l_2}|m^{(\mathbf{2})}_{l_1l_2}, \left(\sigma^{(\mathbf{2})}_{l_1l_2}\right)^2\right) \triangleq \prod_{j=1}^{p(p+1)/2}\mathcal{N}\left(\beta_j|m_j, \sigma_j^2\right).
$$

2. For $q(\boldsymbol{\gamma})$, assume $q(\boldsymbol{\gamma}) = \prod_{j=1}^{p(p+1)/2} q(\gamma_j)$. Consider the following two parts.

(a) For $q\left(\gamma_j^{(\mathbf{1})}\right)$, we have

$$
\begin{aligned}
\log q\left(\gamma_j^{(\mathbf{1})}\right) &= \mathrm{E}_{\left(-\gamma_j^{(\mathbf{1})}\right)}\left[\log p\left(\boldsymbol{y}, \Omega | \boldsymbol{X}; \tau, \theta\right)\right] + \text{const} \\
&= \mathrm{E}_{\left(-\gamma_j^{(\mathbf{1})}\right)}\Bigg[\log p\left(\beta_j^{(\mathbf{1})} | \gamma_j^{(\mathbf{1})}\right) + \log \mathrm{Bern}\left(\gamma_j^{(\mathbf{1})} | \zeta_j^{(\mathbf{1})}\right) + \sum_{l \neq j} \log p\left(\beta_{jl}^{(\mathbf{2})} | \gamma_j^{(\mathbf{1})} \gamma_l^{(\mathbf{1})}\right) \\
&\quad + \text{const} \\
&= -\gamma_j^{(\mathbf{1})}\left\{\frac{1}{2}\log\frac{s_1}{s_2} + \frac{1}{2}\left(\frac{1}{s_1} - \frac{1}{s_2}\right)\left\langle\left(\beta_j^{(\mathbf{1})}\right)^2\right\rangle + \left\langle\log\left(1 - \zeta_j^{(\mathbf{1})}\right)\right\rangle - \left\langle\log\zeta_j^{(\mathbf{1})}\right\rangle\right. \\
&\quad \left. + \frac{1}{2}\sum_{l \neq j}\left\langle\gamma_l^{(\mathbf{1})}\right\rangle\left[\log\frac{s_1}{s_2} + \left\langle\left(\beta_{jl}^{(\mathbf{2})}\right)^2\right\rangle\left(\frac{1}{s_1} - \frac{1}{s_2}\right)\right]\right. + \text{const.}
\end{aligned}
$$

Here we simplify the notation where $\beta_{jl}^{(\mathbf{2})} = \beta_{lj}^{(\mathbf{2})}$. With the form of $q(\boldsymbol{\beta})$, we have $\left\langle\left(\beta_{jl}^{(\mathbf{2})}\right)^2\right\rangle = \left(m_{jl}^{(\mathbf{2})}\right)^2 + \left(\sigma_{jl}^{(\mathbf{2})}\right)^2$, $\left\langle\left(\beta_j^{(\mathbf{1})}\right)^2\right\rangle = \left(m_j^{(\mathbf{1})}\right)^2 + \left(\sigma_j^{(\mathbf{1})}\right)^2$. Hence, $q\left(\gamma_j^{(\mathbf{1})}\right) = \mathrm{Bern}\left(\gamma_j^{(\mathbf{1})} | \eta_j^{(\mathbf{1})}\right)$, with the parameter

$$
\begin{aligned}
\eta_j^{(\mathbf{1})} = &\left(1 + \exp\left\{\frac{1}{2}\log\frac{s_1}{s_2} + \frac{1}{2}\left(\frac{1}{s_1} - \frac{1}{s_2}\right)\left(\left(m_j^{(\mathbf{1})}\right)^2 + \left(\sigma_j^{(\mathbf{1})}\right)^2\right) + \left\langle\log\left(1 - \zeta_j^{(\mathbf{1})}\right)\right\rangle - \left\langle\log\zeta_j^{(\mathbf{1})}\right\rangle\right.\right. \\
&\left.\left. + \frac{1}{2}\sum_{l \neq j}\left\langle\gamma_l^{(\mathbf{1})}\right\rangle\left[\log\frac{s_1}{s_2} + \left(\left(m_{jl}^{(\mathbf{2})}\right)^2 + \left(\sigma_{jl}^{(\mathbf{2})}\right)^2\right)\left(\frac{1}{s_1} - \frac{1}{s_2}\right)\right]\right\}\right)^{-1}.
\end{aligned}
$$

(b) For $\gamma_{l_1 l_2}^{(\mathbf{2})}$, we have

$$
\begin{aligned}
\log q\left(\gamma_{l_1 l_2}^{(\mathbf{2})}\right) &= \mathrm{E}_{\left(-\gamma_{l_1 l_2}^{(\mathbf{2})}\right)}\left[\log p\left(\boldsymbol{y}, \Omega | \boldsymbol{X}; \tau, \theta\right)\right] + \text{const} \\
&= \mathrm{E}_{\left(-\gamma_{l_1 l_2}^{(\mathbf{2})}\right)}\left[\log p\left(\beta_{l_1 l_2}^{(\mathbf{2})} | \gamma_{l_1 l_2}^{(\mathbf{2})}\right) + \log \mathrm{Bern}\left(\gamma_{l_1 l_2}^{(\mathbf{2})} | \zeta_{l_1 l_2}^{(\mathbf{2})}\right)\right] + \text{const} \\
&= -\gamma_{l_1 l_2}^{(\mathbf{2})}\left[\frac{1}{2}\log\frac{s_1}{s_2} + \frac{1}{2}\left(\frac{1}{s_1} - \frac{1}{s_2}\right)\left\langle\left(\beta_{l_1 l_2}^{(\mathbf{2})}\right)^2\right\rangle + \left\langle\log\left(1 - \zeta_{l_1 l_2}^{(\mathbf{2})}\right)\right\rangle - \left\langle\log\zeta_{l_1 l_2}^{(\mathbf{2})}\right\rangle\right] \\
&\quad + \text{const} \\
&= -\gamma_{l_1 l_2}^{(\mathbf{2})}\left[\frac{1}{2}\log\frac{s_1}{s_2} + \frac{1}{2}\left(\frac{1}{s_1} - \frac{1}{s_2}\right)\left(\left(m_{l_1 l_2}^{(\mathbf{2})}\right)^2 + \left(\sigma_{l_1 l_2}^{(\mathbf{2})}\right)^2\right) + \left\langle\log\left(1 - \zeta_{l_1 l_2}^{(\mathbf{2})}\right)\right\rangle\right. \\
&\quad \left. - \left\langle\log\zeta_{l_1 l_2}^{(\mathbf{2})}\right\rangle\right] + \text{const.}
\end{aligned}
$$

Hence $q\left(\gamma^{(\mathbf{2})}_{l_1 l_2}\right) = \mathrm{Bern}\left(\gamma^{(\mathbf{2})}_{l_1 l_2}|\eta^{(\mathbf{2})}_{l_1 l_2}\right)$ with the parameter

$$\begin{aligned}\eta^{(\mathbf{2})}_{l_1 l_2} =&1/\left(1+\exp\left(\frac{1}{2}\log\frac{s_1}{s_2}+\frac{1}{2}\left(\frac{1}{s_1}-\frac{1}{s_2}\right)\left(\left(m^{(\mathbf{2})}_{l_1 l_2}\right)^2+\left(\sigma^{(\mathbf{2})}_{l_1 l_2}\right)^2\right)\right.\right.\\ &\left.\left.+\left\langle\log\left(1-\zeta^{(\mathbf{2})}_{l_1 l_2}\right)\right\rangle-\left\langle\log\zeta^{(\mathbf{2})}_{l_1 l_2}\right\rangle\right)\right).\end{aligned}$$

According to (a), (b), and the structure of $\boldsymbol{\gamma}$, we conclude that

$$\begin{aligned}Q(\boldsymbol{\gamma}) =& \prod_{j=1}^{p}\left(\eta^{(\mathbf{1})}_j\right)^{\gamma^{(\mathbf{1})}_j}\left(1-\eta^{(\mathbf{1})}_j\right)^{1-\gamma^{(\mathbf{1})}_j}\prod_{l_1=1}^{p}\prod_{l_2>l_1}^{p}\left(\eta^{(\mathbf{2})}_{l_1 l_2}\right)^{\gamma^{(\mathbf{2})}_{l_1 l_2}}\left(1-\eta^{(\mathbf{2})}_{l_1 l_2}\right)^{1-\gamma^{(\mathbf{2})}_{l_1 l_2}}\\ \triangleq& \prod_{j=1}^{p(p+1)/2}\left(\eta_j\right)^{\gamma_j}\left(1-\eta_j\right)^{1-\gamma_j}.\end{aligned}$$

3. For $q(\alpha_k)$, we have

$$\begin{aligned}\log q(\alpha_k) &= \mathrm{E}_{(-\alpha_k)}\left[\log p\left(\boldsymbol{y},\Omega|\boldsymbol{X};\tau,\theta\right)\right]+\mathrm{const}\\ &= \mathrm{E}_{(-\alpha_k)}\left[\log p\left(\boldsymbol{\beta_k}|\alpha_k\right)+\log\mathrm{Bern}\left(\alpha_k|\theta\right)\right]+\mathrm{const}\\ &= -\alpha_k\left[\frac{\hat{p}_k}{2}\log\frac{s_1}{s_2}-\frac{1}{2}\log\left|\tilde{\boldsymbol{L}}_k\right|+\frac{1}{2}\left\langle\boldsymbol{\beta}_k^{\mathrm{T}}\left(\frac{1}{s_1}\tilde{\boldsymbol{L}}_k-\frac{1}{s_2}\boldsymbol{I}\right)\boldsymbol{\beta}_k\right\rangle\right]\\ &\quad+(1-\alpha_k)\log(1-\theta)+\alpha_k\log\theta+\mathrm{const}\\ &= -\alpha_k\left[\frac{1}{2}\mathrm{tr}\left(\left\langle\boldsymbol{\beta}_k\boldsymbol{\beta}_k^{\mathrm{T}}\right\rangle\left(\frac{1}{s_1}\tilde{\boldsymbol{L}}_k-\frac{1}{s_2}\boldsymbol{I}\right)\right)+\frac{\hat{p}_k}{2}\log\frac{s_1}{s_2}-\frac{1}{2}\log\left|\tilde{\boldsymbol{L}}_k\right|\right.\\ &\quad\left.+\log(1-\theta)-\log\theta\right]+\mathrm{const}.\end{aligned}$$

Hence, $q(\alpha_k) = \mathrm{Bern}(\alpha_k|r_k)$ with the parameter

$$\begin{aligned}r_k =& \left[1+\exp\left(\frac{1}{2}\mathrm{tr}\left(\left\langle\boldsymbol{\beta}_k\boldsymbol{\beta}_k^{\mathrm{T}}\right\rangle\left(\frac{1}{s_1}\tilde{\boldsymbol{L}}_k-\frac{1}{s_2}\boldsymbol{I}\right)\right)+\frac{\hat{p}_k}{2}\log\frac{s_1}{s_2}\right.\right.\\ &\left.\left.-\frac{1}{2}\log\left|\tilde{\boldsymbol{L}}_k\right|+\log(1-\theta)-\log\theta\right)\right]^{-1}.\end{aligned}$$

4. For $q(\boldsymbol{\zeta})$, with the separable structure of $\boldsymbol{\zeta}$ in the joint probability function, we have $q(\boldsymbol{\zeta}) = \prod_{j=1}^{p(p+1)/2} q(\zeta_j)$ with

$$\begin{aligned}\log q(\zeta_j) &= \mathrm{E}_{(-\zeta_j)}\left[\log p\left(\boldsymbol{y},\Omega|\boldsymbol{X};\tau,\theta\right)\right]+\mathrm{const}\\ &= \mathrm{E}_{(-\zeta_j)}\left[\log\mathrm{Bern}\left(\gamma_j|\zeta_j\right)+\log\mathrm{Beta}\left(\zeta_j|a,b\right)\right]+\mathrm{const}\\ &= \left(\langle\gamma_j\rangle+a-1\right)\log\zeta_j+\left(b-\langle\gamma_j\rangle\right)\log\left(1-\zeta_j\right)+\mathrm{const}.\end{aligned}$$

With the form of $q(\boldsymbol{\gamma})$, we have $\langle\gamma_j\rangle = \eta_j$. Hence, $q(\zeta_j) = \mathrm{Beta}\left(\zeta_j|\tilde{a}_j,\tilde{b}_j\right)$ with

$$\begin{cases}\tilde{a}_j = \eta_j + a,\\ \tilde{b}_j = 1-\eta_j+b.\end{cases}$$

Then, we have

$$Q(\boldsymbol{\zeta}) \propto \prod_{j=1}^{p(p+1)/2} (\zeta_j)^{\tilde{a}_j - 1} (1-\zeta_j)^{\tilde{b}_j - 1}.$$

S1.3.2 M-step

We update model parameters τ and θ. First, we evaluate the lower bound $\mathcal{L}_q = -\mathrm{E}_q(\log q(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\zeta})) + \mathrm{E}_q(\log p(\boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\zeta} | \boldsymbol{X}; \tau, \theta))$. The first part of the lower bound $\mathcal{L}_q$:

$$\begin{aligned} & -\mathrm{E}_q[\log q(\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\zeta})] \\ = & -\mathrm{E}_q\left\{\sum_{k=1}^{K} \log \mathrm{Bern}(\alpha_k | r_k) + \sum_{j=1}^{p(p+1)/2} \left[\log \mathcal{N}\left(\beta_j | m_j, \sigma_j^2\right) + \log \mathrm{Bern}(\gamma_j | \eta_j)\right.\right. \\ & \left.\left. + \log \mathrm{Beta}\left(\zeta_j | \tilde{a}_j, \tilde{b}_j\right)\right]\right\}. \end{aligned}$$

The second part of $\mathcal{L}_q$:

$$\begin{aligned} & \mathrm{E}_q(\log p(\boldsymbol{y}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\alpha}, \boldsymbol{\zeta} | \boldsymbol{X}; \tau, \theta)) \\ = & \mathrm{E}_q\left\{\log \mathcal{N}\left(\boldsymbol{y} | \tilde{\boldsymbol{X}}\boldsymbol{\beta}, \tau^{-1}\boldsymbol{I}\right) + \sum_{j=1}^{p(p+1)/2} [\gamma_j \log \mathcal{N}(\beta_j | 0, s_1) + (1-\gamma_j) \log \mathcal{N}(\beta_j | 0, s_2)\right. \\ & + \log \mathrm{Bern}(\gamma_j | \zeta_j) + \log \mathrm{Beta}(\zeta_j | a, b)] + \sum_{k=1}^{K} \left[\alpha_k \log \mathcal{N}\left(\boldsymbol{\beta}_k | \boldsymbol{0}, s_1 \tilde{\boldsymbol{L}}_k^{-1}\right) + (1-\alpha_k) \mathcal{N}(\boldsymbol{\beta}_k | \boldsymbol{0}, s_2 \boldsymbol{I})\right. \\ & \left. + \log \mathrm{Bern}(\alpha_k | \theta)] + \sum_{l_1=1}^{p} \sum_{l_2 > l_1}^{p} \log \mathcal{N}(\beta_{l_1 l_2} | 0, s_1)^{\gamma_{l_1}\gamma_{l_2}} \mathcal{N}(\beta_{l_1 l_2} | 0, s_2)^{1-\gamma_{l_1}\gamma_{l_2}}\right\}. \end{aligned}$$

By taking derivatives with respect to model parameters and setting them to zero,

$$\begin{aligned} \frac{\partial \mathcal{L}_q}{\partial \tau} = & \frac{n}{2\tau} - \frac{1}{2}\left(\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} - 2\boldsymbol{y}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{m} + \mathrm{tr}\left(\tilde{\boldsymbol{X}}\left(\boldsymbol{m}\boldsymbol{m}^{\mathrm{T}} + \mathrm{diag}\left(\boldsymbol{\sigma}^2\right)\right)\tilde{\boldsymbol{X}}^{\mathrm{T}}\right)\right) = 0, \\ \frac{\partial \mathcal{L}_q}{\partial \theta} = & -\sum_{k=1}^{K}\left(\frac{1-r_k}{1-\theta} - \frac{r_k}{\theta}\right) = 0, \end{aligned}$$

we can obtain

$$\begin{aligned} \tau = & \frac{n}{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} - 2\boldsymbol{y}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{m} + \mathrm{tr}\left(\tilde{\boldsymbol{X}}\left(\boldsymbol{m}\boldsymbol{m}^{\mathrm{T}} + \mathrm{diag}(\boldsymbol{\sigma}^2)\right)\tilde{\boldsymbol{X}}^{\mathrm{T}}\right)}, \\ \theta = & \sum_{k=1}^{K} \frac{r_k}{K}. \end{aligned} \tag{S3}$$

S1.4 Variational Bayesian expectation-maximization algorithm for the proposed approach

Algorithm 1 Variational Bayesian expectation-maximization algorithm for the proposed approach

Input: $\boldsymbol{X}$: design matrix; $\boldsymbol{y}$: response; G: network information; p_0: initial value of vector $\boldsymbol{w} = \left(r_k\text{'s}, \eta_j^{(\mathbf{1})}\text{'s}, \eta_{l_1 l_2}^{(\mathbf{2})}\text{'s}\right)$; s_2: tuning parameter; $s_1 = 1$; $a = b = 1$;

Output: optimal variational parameters for $q(\Omega)$;

Initialize $\boldsymbol{m} = \left(m_j^{(\mathbf{1})}\text{'s}, m_{l_1 l_2}^{(\mathbf{2})}\text{'s}\right) = \boldsymbol{\sigma} = \left(\sigma_j^{(\mathbf{1})}\text{'s}, \sigma_{l_1 l_2}^{(\mathbf{2})}\text{'s}\right) = \mathbf{0}$, $\tilde{\boldsymbol{a}} = (\tilde{a}_1, \cdots, \tilde{a}_p) = \left(\tilde{a}_j^{(\mathbf{1})}\text{'s}, \tilde{a}_{l_1 l_2}^{(\mathbf{2})}\text{'s}\right) = \mathbf{1}$, $\tilde{\boldsymbol{b}} = \left(\tilde{b}_1, \cdots, \tilde{b}_p\right) = \left(\tilde{b}_j^{(\mathbf{1})}\text{'s}, \tilde{b}_{l_1 l_2}^{(\mathbf{2})}\text{'s}\right) = \mathbf{1}$, for $1 \le j \le p$ and $1 \le l_1 < l_2 \le p$. $\tau^{-1} = \mathrm{SD}(\boldsymbol{y})$, and $\theta = \frac{\sum r_k}{K}$.

repeat

E-step:

$\mathscr{P} = \{k : r_k > 0.5\}$;

for $k \in \mathscr{P}$ **do**

$$r_k = 1/\left(1 + \exp\left(\frac{1}{2}\operatorname{tr}\left(\left(\boldsymbol{m}_k \boldsymbol{m}_k^{\mathrm{T}} + \operatorname{diag}\left(\boldsymbol{\sigma}_k^2\right)\right)\left(\frac{1}{s_1}\tilde{\boldsymbol{L}}_k - \frac{1}{s_2}\boldsymbol{I}\right)\right) + \frac{\hat{p}_k}{2}\log\frac{s_1}{s_2} - \frac{1}{2}\log\left|\tilde{\boldsymbol{L}}_k\right| + \log(1-\theta) - \log\theta\right)\right);$$

for $j \in V_k$ **do**

$$\left(\sigma_j^{(\mathbf{1})}\right)^2 = \left(\frac{\eta_j^{(\mathbf{1})}}{s_1} + \frac{1-\eta_j^{(\mathbf{1})}}{s_2} + \left\langle \underline{b}_{jj}^{(k)} \right\rangle + \tau \sum_i (x_{ij})^2\right)^{-1};$$

$$m_j^{(\mathbf{1})} = \left(\sigma_j^{(\mathbf{1})}\right)^2 \left(\tau \sum_i x_{ij} \left\langle res_i^{(-j)} \right\rangle - \sum_{l \ne j, k_l = k_j} \left\langle \underline{b}_{lj}^{(k)} \right\rangle m_l^{(\mathbf{1})}\right);$$

$$\eta_j^{(\mathbf{1})} = 1/\left(1 + \exp\left\{\psi\left(\tilde{b}_j^{(\mathbf{1})}\right) - \psi\left(\tilde{a}_j^{(\mathbf{1})}\right) + \frac{1}{2}\log\frac{s_1}{s_2} + \frac{1}{2}\left(\left(m_j^{(\mathbf{1})}\right)^2 + \left(\sigma_j^{(\mathbf{1})}\right)^2\right)\left(\frac{1}{s_1} - \frac{1}{s_2}\right) + \frac{1}{2}\sum_{l \ne j} \eta_l^{(\mathbf{1})}\left(\log\frac{s_1}{s_2} + \left(\left(m_{jl}^{(\mathbf{2})}\right)^2 + \left(\sigma_{jl}^{(\mathbf{2})}\right)^2\right)\left(\frac{1}{s_1} - \frac{1}{s_2}\right)\right)\right\}\right);$$

$\tilde{a}_j^{(\mathbf{1})} = \eta_j^{(\mathbf{1})} + a$;

$\tilde{b}_j^{(\mathbf{1})} = 1 - \eta_j^{(\mathbf{1})} + b$;

for $l > j$ and $l \in V_k$ **do**

$$\left(\sigma_{jl}^{(\mathbf{2})}\right)^2 = \left(\frac{2-\eta_{jl}^{(\mathbf{2})} - \eta_j^{(\mathbf{1})}\eta_l^{(\mathbf{1})}}{s_2} + \frac{\eta_j^{(\mathbf{1})}\eta_l^{(\mathbf{1})} + \eta_{jl}^{(\mathbf{2})}}{s_1} + \left\langle \hat{b}_{j\sim l, j\sim l}^{(k)} \right\rangle + \tau \sum_i (x_{ij}x_{il})^2\right)^{-1};$$

$$m_{jl}^{(\mathbf{2})} = \left(\sigma_{jl}^{(\mathbf{2})}\right)^2 \left(\tau \sum_i x_{ij}x_{il} \left\langle res_i^{(-jl)} \right\rangle - \sum_{\substack{(j',l') \ne (j,l) \\ k_{j'} = k_{l'} = k}} \left\langle \hat{b}_{j'\sim l', j\sim l}^{(k)} \right\rangle m_{j'l'}^{(\mathbf{2})}\right);$$

$$\eta_{jl}^{(\mathbf{2})} = 1/\left(1 + \exp\left(\frac{1}{2}\log\frac{s_1}{s_2} + \frac{1}{2}\left(\left(m_{jl}^{(\mathbf{2})}\right)^2 + \left(\sigma_{jl}^{(\mathbf{2})}\right)^2\right)\left(\frac{1}{s_1} - \frac{1}{s_2}\right) + \psi\left(\tilde{b}_{jl}^{(\mathbf{2})}\right) - \psi\left(\tilde{a}_{jl}^{(\mathbf{2})}\right)\right)\right);$$

$\tilde{a}_{jl}^{(\mathbf{2})} = \eta_{jl}^{(\mathbf{2})} + a$;

$\tilde{b}_{jl}^{(\mathbf{2})} = 1 - \eta_{jl}^{(\mathbf{2})} + b$;

end for

end for

for $l_1, l_2 \in \left\{(j', l') | k_{j'}, k_{l'} \in \mathscr{P} \text{ and } k_{j'} \ne k_{l'}\right\}$ **do**

$\left(\sigma^{(\mathbf{2})}_{l_1l_2}\right)^2 = 1/\left(\tau \sum_i \left(x_{il_1}x_{il_2}\right)^2 + \frac{1}{s_1}\left(\eta^{(\mathbf{2})}_{l_1l_2} + \eta^{(\mathbf{1})}_{l_1}\eta^{(\mathbf{1})}_{l_2}\right) + \frac{1}{s_2}\left(2 - \eta^{(\mathbf{2})}_{l_1l_2} - \eta^{(\mathbf{1})}_{l_1}\eta^{(\mathbf{1})}_{l_2}\right)\right);$

$m^{(\mathbf{2})}_{l_1l_2} = \tau \sum_i x_{il_1}x_{il_2}\left\langle res_i^{(-l_1l_2)}\right\rangle\left(\sigma^{(\mathbf{2})}_{l_1l_2}\right)^2;$

$$\eta^{(\mathbf{2})}_{l_1l_2} =1/\left(1+\exp\left(\frac{1}{2}\log\frac{s_1}{s_2} + \frac{1}{2}\left(\left(m^{(\mathbf{2})}_{l_1l_2}\right)^2 + \left(\sigma^{(\mathbf{2})}_{l_1l_2}\right)^2\right)\left(\frac{1}{s_1}-\frac{1}{s_2}\right) + \psi\left(\tilde{b}^{(\mathbf{2})}_{l_1l_2}\right) - \psi\left(\tilde{a}^{(\mathbf{2})}_{l_1l_2}\right)\right)\right);$$

$\tilde{a}^{(\mathbf{2})}_{l_1l_2} = \eta^{(\mathbf{2})}_{l_1l_2} + a;$

$\tilde{b}^{(\mathbf{2})}_{l_1l_2} = 1 - \eta^{(\mathbf{2})}_{l_1l_2} + b;$

end for

end for

M-step:

$\tau = \frac{n}{\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y} - 2\boldsymbol{y}^{\mathrm{T}}\tilde{\boldsymbol{X}}\boldsymbol{m} + \mathrm{tr}\left(\tilde{\boldsymbol{X}}(\boldsymbol{m}\boldsymbol{m}^{\mathrm{T}} + \mathrm{diag}(\boldsymbol{\sigma}^2))\tilde{\boldsymbol{X}}^{\mathrm{T}}\right)};$

$\theta = \sum_{k=1}^{K}\frac{r_k}{K};$

until the change of Ω is smaller than a threshold 10^{-3}.

In the algorithm above,

- $\left\langle \underline{b}^{(k)}_{lj}\right\rangle$ is the element of matrix $\boldsymbol{B}^{(\mathbf{1})}_k = \frac{1}{s_1}\left(r_k\tilde{\boldsymbol{L}}^{(\mathbf{1})}_k\right) + \frac{1}{s_2}\left((1-r_k)\,\boldsymbol{I}\right);$
- $\left\langle res_i^{(-j)}\right\rangle = y_i - \tilde{\boldsymbol{x}}_i\boldsymbol{m} + x_{ij}m^{(\mathbf{1})}_j$, $\left\langle res_i^{(-l_1l_2)}\right\rangle = y_i - \tilde{\boldsymbol{x}}_i\boldsymbol{m} + x_{il_1}x_{il_2}m^{(\mathbf{2})}_{l_1l_2};$
- $\left\langle \hat{b}^{(k)}_{lj}\right\rangle$ is the element of matrix $\boldsymbol{B}^{(\mathbf{2})}_k = \frac{1}{s_1}\left(r_k\tilde{\boldsymbol{L}}^{(\mathbf{2})}_k\right) + \frac{1}{s_2}\left((1-r_k)\,\boldsymbol{I}\right);$
- $\psi(x) = \frac{\mathrm{d}}{\mathrm{dx}}\ln\Gamma(x)$, where $\Gamma(x)$ is the gamma function.

S2 Results for simulation and data analysis

This section includes additional results for simulation and data analysis.

S2.1 Sensitivity analysis for parameter s_1

Table S1: Simulation results under the scenarios with $\rho = 0.4, K = 100$, and $r = 1/\sqrt{5}$, and various values of s_1. In each cell, mean (SD) based on 100 replicates.

s_1	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
				S1			
0.5	17.90(0.36)	4.16(1.35)	0.46(0.07)	16.86(0.40)	15.40(5.46)	0.62(0.08)	0.89(0.20)
1	17.86(0.35)	2.06(1.06)	0.35(0.10)	16.90(0.30)	9.44(3.64)	0.45(0.09)	0.55(0.10)
2	17.86(0.35)	2.04(1.16)	0.32(0.11)	16.90(0.30)	8.70(3.47)	0.41(0.09)	0.91(0.19)
3	17.80(0.40)	1.76(1.27)	0.34(0.12)	16.82(0.39)	7.88(3.61)	0.41(0.09)	0.95(0.22)
5	17.74(0.49)	1.52(1.22)	0.36(0.14)	16.84(0.37)	7.48(3.86)	0.40(0.09)	0.99(0.25)
8	17.60(0.57)	1.38(1.23)	0.39(0.15)	16.76(0.48)	7.20(4.04)	0.41(0.10)	1.01(0.23)
10	17.56(0.61)	1.30(1.11)	0.40(0.15)	16.70(0.51)	7.12(3.94)	0.42(0.11)	1.03(0.23)
20	17.32(0.74)	1.06(1.11)	0.46(0.17)	16.62(0.70)	7.28(4.53)	0.44(0.13)	1.09(0.24)
30	17.14(0.78)	1.12(1.12)	0.50(0.17)	16.50(0.74)	7.36(4.63)	0.47(0.14)	1.11(0.27)
50	16.76(0.96)	0.96(1.11)	0.57(0.18)	16.18(0.87)	7.24(4.49)	0.53(0.14)	1.16(0.29)
100	16.04(1.12)	0.66(0.72)	0.69(0.18)	15.62(1.31)	5.72(3.61)	0.60(0.17)	1.20(0.27)
				S2			
0.5	17.64(0.66)	2.76(1.65)	0.45(0.11)	16.62(0.64)	12.66(5.27)	0.66(0.10)	1.05(0.24)
1	17.30(0.79)	1.77(1.41)	0.43(0.13)	16.57(0.68)	8.17(3.56)	0.52(0.12)	0.69(0.18)
2	17.30(0.86)	1.46(1.30)	0.40(0.15)	16.58(0.70)	7.66(3.98)	0.47(0.14)	1.05(0.25)
3	17.14(0.95)	1.24(1.24)	0.43(0.16)	16.56(0.76)	7.62(4.18)	0.46(0.15)	1.07(0.25)
5	16.66(1.17)	0.70(0.86)	0.50(0.19)	16.36(0.90)	5.56(3.23)	0.48(0.16)	1.13(0.27)
8	16.54(1.15)	0.76(1.00)	0.53(0.18)	16.36(0.90)	5.86(4.03)	0.49(0.16)	1.14(0.27)
10	16.40(1.25)	0.54(0.86)	0.55(0.19)	16.26(1.07)	5.68(4.10)	0.49(0.19)	1.17(0.32)
20	16.10(1.23)	0.60(0.93)	0.60(0.19)	15.94(1.13)	5.94(3.63)	0.56(0.20)	1.21(0.34)
30	15.80(1.32)	0.54(0.84)	0.65(0.20)	15.76(1.10)	5.92(3.94)	0.60(0.18)	1.27(0.35)
50	15.24(1.53)	0.34(0.59)	0.72(0.21)	15.26(1.23)	4.82(3.38)	0.67(0.19)	1.35(0.45)
100	14.52(1.53)	0.24(0.48)	0.82(0.19)	14.72(1.51)	4.38(2.81)	0.73(0.22)	1.45(0.45)
				S3			
0.5	17.80(6.91)	2.58(1.73)	0.76(0.31)	14.93(5.90)	13.18(7.05)	0.89(0.36)	1.00(0.45)
1	17.63(0.56)	1.19(1.10)	0.57(0.10)	15.04(1.37)	8.87(4.35)	0.77(0.13)	0.70(0.21)
2	17.65(6.87)	1.23(1.22)	0.48(0.21)	14.93(5.94)	8.57(5.27)	0.72(0.32)	1.02(0.44)
3	17.38(6.78)	0.98(1.03)	0.49(0.23)	14.60(5.85)	7.38(4.63)	0.74(0.33)	1.09(0.47)
5	17.02(6.66)	1.05(1.26)	0.53(0.26)	14.07(5.67)	7.82(5.14)	0.79(0.36)	1.13(0.49)
8	16.68(6.54)	0.95(1.08)	0.58(0.28)	13.88(5.62)	7.70(5.25)	0.82(0.36)	1.15(0.48)
10	16.38(6.43)	0.90(1.05)	0.62(0.30)	13.47(5.47)	7.62(5.28)	0.86(0.37)	1.21(0.52)
20	15.43(6.05)	0.55(0.81)	0.76(0.33)	12.32(5.01)	6.03(4.22)	0.96(0.39)	1.34(0.60)
30	14.80(5.83)	0.38(0.68)	0.83(0.35)	11.62(4.78)	4.70(3.63)	1.01(0.41)	1.37(0.63)
50	14.32(5.63)	0.17(0.37)	0.89(0.36)	11.05(4.55)	4.08(3.03)	1.05(0.42)	1.40(0.62)
100	13.43(5.32)	0.17(0.37)	0.98(0.40)	10.05(4.20)	3.70(2.83)	1.12(0.45)	1.45(0.67)
				S4			
0.5	17.85(6.93)	3.45(1.99)	0.44(0.19)	14.45(5.78)	20.68(9.03)	0.89(0.37)	1.16(0.51)
1	17.94(0.24)	2.10(1.12)	0.37(0.07)	13.37(2.17)	15.10(3.98)	0.92(0.20)	0.77(0.20)
2	17.90(6.95)	2.38(1.47)	0.36(0.16)	14.18(5.72)	16.60(7.41)	0.84(0.37)	1.16(0.57)
3	17.82(6.93)	2.05(1.33)	0.37(0.17)	13.80(5.61)	15.60(7.03)	0.87(0.39)	1.24(0.62)
5	17.80(6.92)	1.77(1.26)	0.39(0.18)	13.43(5.46)	15.18(6.65)	0.92(0.40)	1.24(0.61)
8	17.55(6.83)	1.50(1.20)	0.44(0.21)	12.97(5.37)	14.25(6.47)	0.96(0.42)	1.28(0.63)
10	17.58(6.83)	1.52(1.22)	0.44(0.21)	12.78(5.29)	14.22(6.47)	0.98(0.43)	1.31(0.65)
20	17.02(6.65)	1.02(0.91)	0.53(0.25)	11.78(5.01)	12.70(5.66)	1.08(0.47)	1.40(0.65)
30	16.75(6.55)	0.80(0.82)	0.57(0.27)	11.43(4.88)	11.95(5.61)	1.12(0.48)	1.46(0.70)
50	16.08(6.34)	0.60(0.79)	0.67(0.31)	10.25(4.50)	10.62(5.26)	1.22(0.51)	1.62(0.78)
100	14.93(5.86)	0.35(0.50)	0.83(0.34)	9.05(4.00)	9.18(4.27)	1.32(0.54)	1.78(0.84)

S2.2 Computer time

Table S2: Computer time (minutes) of the proposed analysis with fixed tunning parameters. In each cell, mean based on 100 replicates.

	$n = 150$	$n = 200$	$n = 250$	$n = 300$	$n = 400$	$n = 500$
$p = 200$	0.052	0.052	0.052	0.056	0.057	0.063
$p = 300$	0.082	0.085	0.083	0.085	0.090	0.097
$p = 400$	0.110	0.109	0.109	0.123	0.126	0.155
$p = 500$	0.139	0.128	0.149	0.152	0.146	0.187
$p = 600$	0.171	0.174	0.165	0.222	0.325	0.459
$p = 800$	0.225	0.241	0.244	0.347	0.614	0.770
$p = 1000$	0.289	0.312	0.338	0.487	0.850	1.036
$p = 1200$	0.443	0.524	0.779	0.971	1.412	1.736
$p = 1400$	0.666	0.848	1.150	1.540	1.941	3.102

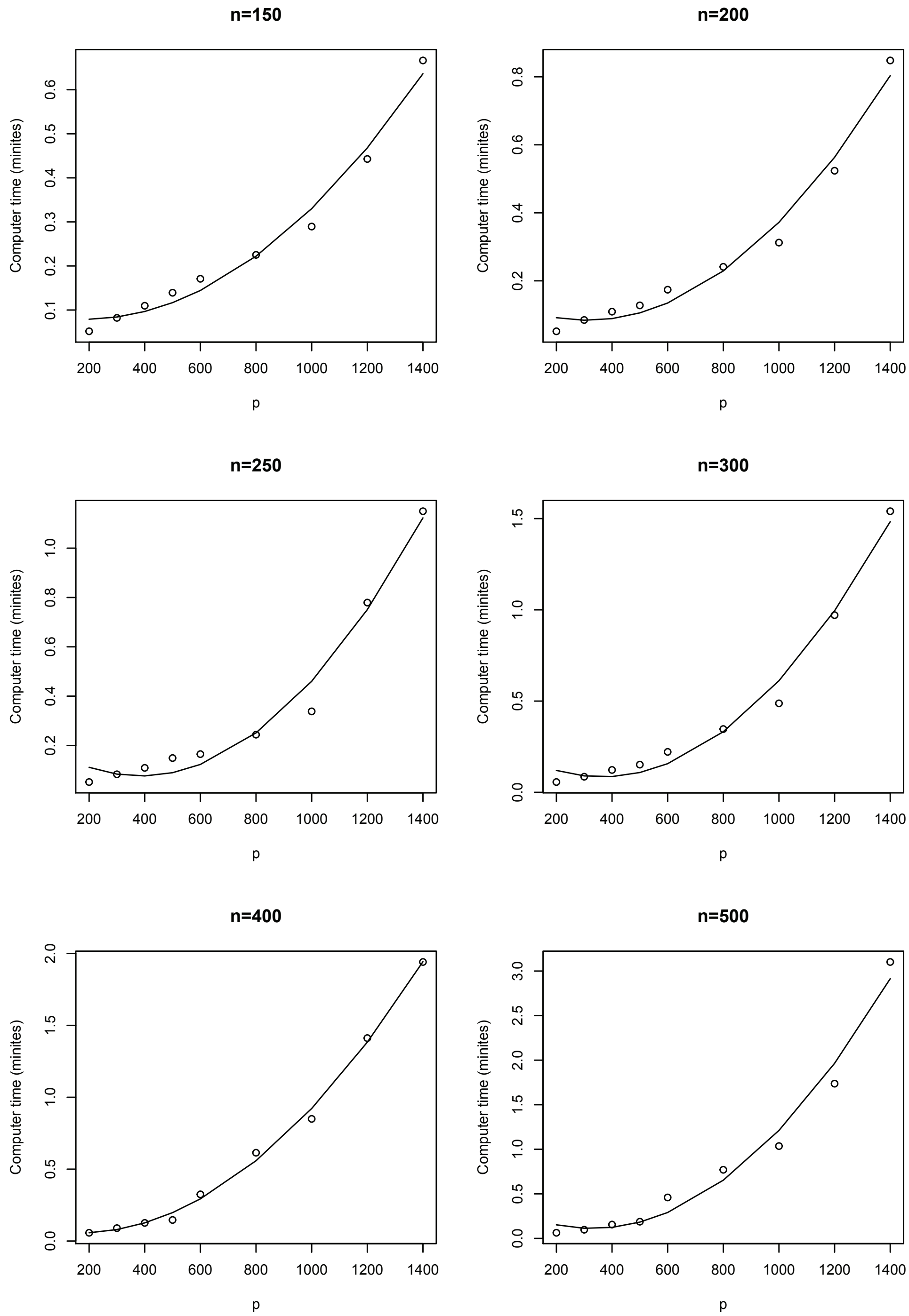

Figure S2: Computer time of the proposed analysis with various p values with a given sample size n.

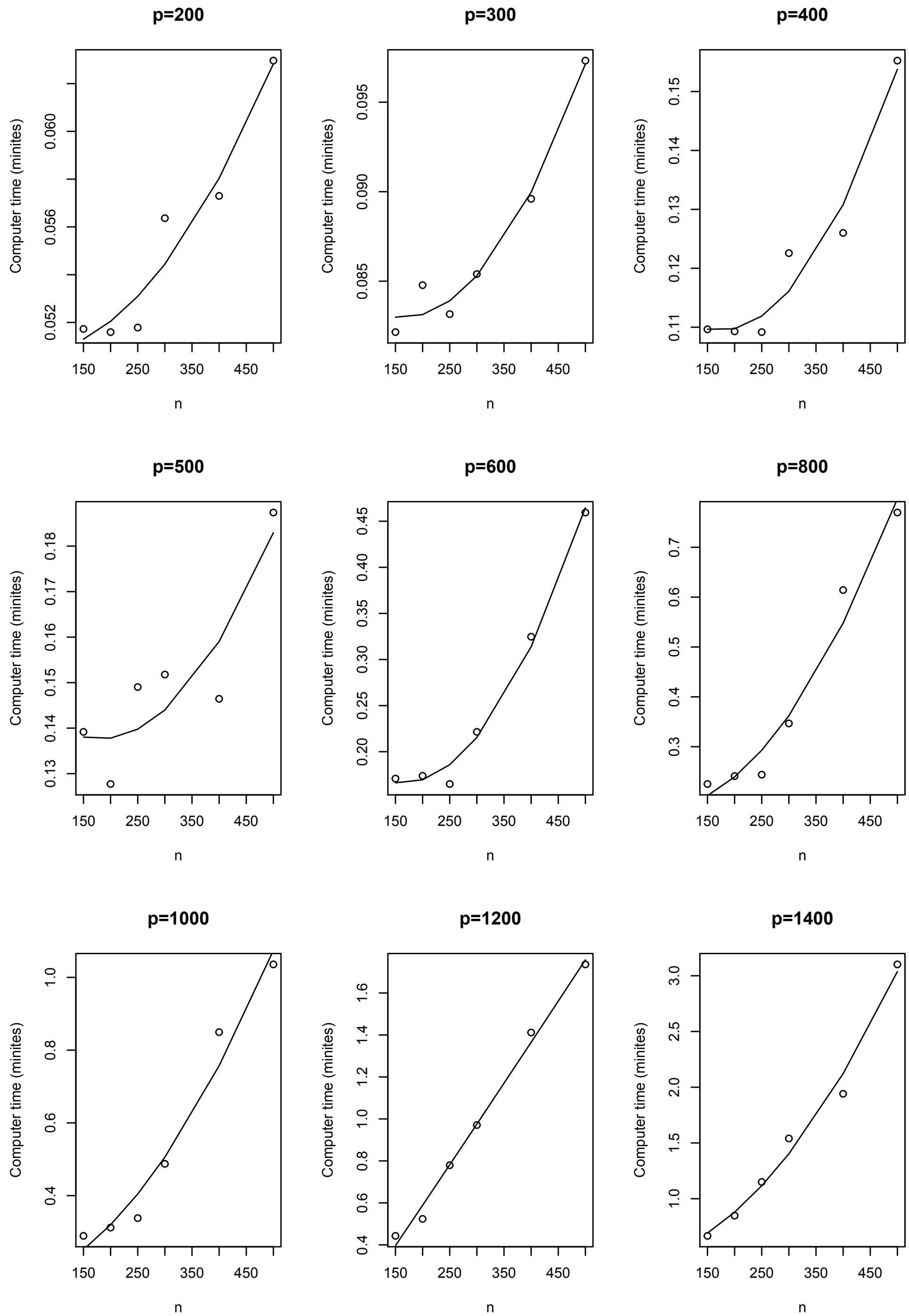

Figure S3: Computer time of the proposed analysis with various n values with a given number of genetic factors p.

S2.3 Detailed simulation settings

We set that only the first three networks have effects on the response. Specifically, taking z_i generated from Uniform$(0.8, 1.2), i = 1, 2, 3, 4$, and $(n_1, n_2, n_3) = (5, 5, 2)$, we consider

1. S1: all signals are positive.

$$\boldsymbol{\beta}_k = \left(z_k, \underbrace{rz_k, \ldots, rz_k}_{5}, \underbrace{0, \ldots, 0}_{p_k-6}, \underbrace{rz_k, \ldots, rz_k}_{n_k}, 0, \ldots, 0 \right)^{\mathrm{T}}, \ k = 1, 2, 3,$$

$$\boldsymbol{\beta}_{12}^{(\mathbf{2})} = \left(\underbrace{rz_4, \ldots, rz_4}_{5}, 0, \ldots, 0 \right)^{\mathrm{T}};$$

2. S2 is the same as S1 except that the signals in the second network and those of interactions between the first and second networks are negative, i.e.,

$$\boldsymbol{\beta}_2 = \left(-z_2, \underbrace{-rz_2, \ldots, -rz_2}_{5}, \underbrace{0, \ldots, 0}_{p_2-6}, \underbrace{-rz_2, \ldots, -rz_2}_{n_2}, 0, \ldots, 0 \right)^{\mathrm{T}},$$

$$\boldsymbol{\beta}_k = \left(z_k, \underbrace{rz_k, \ldots, rz_k}_{5}, \underbrace{0, \ldots, 0}_{p_k-6}, \underbrace{rz_k, \ldots, rz_k}_{n_k}, 0, \ldots, 0 \right)^{\mathrm{T}}, \ k = 1, 3,$$

$$\boldsymbol{\beta}_{12}^{(\mathbf{2})} = \left(\underbrace{-rz_4, \ldots, -rz_4}_{5}, 0, \ldots, 0 \right)^{\mathrm{T}};$$

3. S3: Within each network, the signals can be either positive or negative.

$$\boldsymbol{\beta}_k = \left(z_k, -rz_k, -rz_k, rz_k, rz_k, rz_k, \underbrace{0, \ldots, 0}_{p_k-6}, -rz_k, \underbrace{rz_k, \ldots, rz_k}_{4}, 0, \ldots, 0 \right)^{\mathrm{T}}, \text{ for } k = 1, 2,$$

$$\boldsymbol{\beta}_3 = \left(z_3, -rz_3, \underbrace{rz_3, \ldots, rz_3}_{4}, \underbrace{0, \ldots, 0}_{p_k-6}, rz_3, rz_3, 0, \ldots, 0 \right)^{\mathrm{T}},$$

$$\boldsymbol{\beta}_{12}^{(\mathbf{2})} = \left(-rz_4, \underbrace{rz_4, \ldots, rz_4}_{4}, 0, \ldots, 0 \right)^{\mathrm{T}};$$

4. S4: the important interactions only involve none-TF main effects with weaker signals.

$$\boldsymbol{\beta}_k = \left(z_k, \underbrace{rz_k, \ldots, rz_k}_{5}, \underbrace{0, \ldots, 0}_{2p_k-7}, \underbrace{rz_k, \ldots, rz_k}_{n_k}, 0, \ldots, 0 \right)^{\mathrm{T}}, k = 1, 2, 3,$$

$$\boldsymbol{\beta}_{12}^{(\mathbf{2})} = \left(\underbrace{0, \ldots, 0}_{p_k}, \underbrace{rz_4, \ldots, rz_4}_{5}, 0, \ldots, 0 \right)^{\mathrm{T}}.$$

S2.4 Additional Simulation Results

Table S3: Simulation results under the scenarios with $\rho = 0.4, K = 50$, and $r = 1/\sqrt{5}$. In each cell, mean (SD) based on 100 replicates.

Approach	M:TP	M:FP	M:RSSE	I:TP	I:FP	I:RSSE	PMSE
				S1			
proposed	16.77(1.07)	3.70(2.94)	0.63(0.13)	15.83(1.02)	10.07(5.49)	0.62(0.15)	0.73(0.16)
triBayes	15.86(1.44)	14.60(4.57)	1.31(0.09)	6.66(1.26)	7.80(2.09)	1.36(0.07)	2.17(0.55)
glinternet	17.38(0.95)	6.70(3.21)	0.97(0.10)	14.44(1.72)	6.18(2.80)	1.05(0.11)	1.57(0.42)
Lasso	7.02(2.06)	0.04(0.20)	1.51(0.08)	10.56(2.50)	15.66(7.01)	1.26(0.12)	2.41(0.77)
iFORM	15.90(2.31)	37.68(4.51)	1.43(0.38)	11.70(4.39)	31.32(3.99)	1.37(0.37)	2.39(1.18)
HierNet	13.28(3.43)	1.36(1.59)	1.27(0.17)	7.98(2.74)	9.04(6.24)	1.38(0.13)	2.48(0.77)
Grace	8.88(2.24)	0.84(0.96)	1.50(0.12)	11.70(1.96)	10.90(7.27)	1.26(0.10)	2.35(0.63)
GEL	10.98(2.24)	25.62(6.15)	2.98(0.89)	7.18(2.05)	340.74(77.28)	9.44(3.82)	31.38(28.41)
				S2			
proposed	16.02(1.09)	2.90(2.26)	0.64(0.14)	16.33(1.10)	7.48(4.48)	0.57(0.15)	0.80(0.24)
triBayes	16.38(1.47)	12.16(4.40)	1.38(0.30)	5.02(3.15)	1.46(1.70)	1.44(0.14)	2.60(1.20)
glinternet	17.36(0.94)	6.18(3.81)	0.94(0.11)	14.78(1.37)	5.26(3.13)	1.02(0.12)	1.44(0.40)
Lasso	7.24(3.07)	0.16(0.42)	1.46(0.11)	8.26(3.64)	42.76(92.48)	1.41(0.14)	2.67(0.85)
iFORM	15.58(2.37)	37.76(3.59)	1.48(0.39)	12.16(4.10)	31.32(4.57)	1.42(0.41)	2.52(1.20)
HierNet	13.24(3.10)	1.52(1.69)	1.23(0.17)	9.40(2.77)	8.54(5.36)	1.35(0.15)	2.00(0.55)
Grace	5.42(1.67)	0.36(0.56)	1.73(0.05)	5.04(1.29)	8.86(8.27)	1.49(0.06)	3.37(0.95)
GEL	11.16(2.10)	26.18(5.68)	2.76(0.76)	8.56(1.97)	346.44(71.58)	8.61(2.97)	25.87(24.97)
				S3			
proposed	15.63(1.70)	3.77(3.80)	0.90(0.13)	11.58(2.10)	8.44(6.28)	1.04(0.12)	0.99(0.28)
triBayes	12.32(2.07)	6.58(4.76)	1.94(0.33)	0.82(1.79)	0.78(1.79)	1.54(0.23)	3.35(1.19)
glinternet	13.12(2.68)	3.64(3.35)	1.40(0.16)	6.96(3.00)	2.42(2.33)	1.40(0.10)	2.16(0.64)
Lasso	5.94(2.28)	0.08(0.27)	1.74(0.14)	6.80(1.91)	9.16(5.15)	1.39(0.08)	2.54(0.68)
iFORM	15.88(2.42)	37.74(4.54)	1.46(0.43)	10.82(4.13)	32.62(4.11)	1.42(0.36)	2.31(1.29)
HierNet	9.50(2.46)	0.94(1.13)	1.53(0.14)	5.20(2.19)	5.32(4.28)	1.44(0.08)	2.30(0.64)
Grace	8.72(1.51)	0.94(1.15)	1.58(0.11)	10.20(1.60)	18.64(13.34)	1.31(0.07)	2.03(0.52)
GEL	9.02(3.17)	23.26(9.22)	2.51(0.50)	4.78(2.61)	303.18(109.17)	7.13(2.18)	17.17(9.91)
				S4			
proposed	16.69(1.24)	2.56(2.40)	0.62(0.15)	10.56(2.48)	11.38(3.71)	1.14(0.19)	0.97(0.22)
triBayes	16.62(1.43)	11.90(7.24)	1.55(0.34)	0.06(0.24)	2.08(2.12)	1.65(0.05)	3.13(1.54)
glinternet	15.66(1.84)	4.24(2.76)	1.09(0.18)	6.90(2.87)	5.52(2.67)	1.47(0.09)	1.79(0.46)
Lasso	6.96(2.15)	0.10(0.36)	1.51(0.08)	3.82(2.24)	9.58(5.87)	1.56(0.06)	2.55(0.70)
iFORM	11.78(3.03)	44.44(4.66)	2.34(0.53)	2.96(3.54)	39.06(3.87)	2.25(0.33)	5.13(2.26)
HierNet	12.16(3.22)	0.80(1.03)	1.29(0.18)	4.96(2.52)	8.00(4.72)	1.69(0.13)	2.04(0.67)
Grace	8.46(1.98)	0.48(0.68)	1.48(0.08)	4.16(2.54)	11.76(7.45)	1.63(0.08)	2.42(0.61)
GEL	11.88(0.87)	26.66(5.55)	2.80(0.66)	7.30(1.75)	354.38(75.68)	8.96(2.97)	28.46(21.82)

Table S4: Simulation results under the scenarios with $\rho = 0.4, K = 50$, and $r = 1/\sqrt{12}$. In each cell, mean (SD) based on 100 replicates.

Approach	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
				S1			
proposed	15.24(1.78)	2.42(2.46)	0.60(0.09)	13.46(2.06)	4.66(3.54)	0.57(0.13)	0.72(0.19)
triBayes	14.94(2.25)	11.28(6.59)	1.32(0.26)	3.74(2.54)	1.62(1.59)	0.93(0.07)	1.71(0.68)
glinternet	14.80(2.56)	3.54(2.57)	0.77(0.09)	10.72(2.63)	3.26(2.10)	0.80(0.06)	1.23(0.38)
Lasso	5.30(1.66)	0.12(0.39)	1.02(0.06)	8.34(2.06)	11.84(5.53)	0.88(0.07)	1.47(0.42)
iFORM	11.30(2.04)	45.18(4.34)	1.61(0.19)	4.84(2.03)	35.34(2.98)	1.39(0.12)	2.61(0.84)
HierNet	9.38(2.68)	0.46(0.76)	0.93(0.06)	5.32(2.28)	5.22(4.34)	0.95(0.06)	1.52(0.39)
Grace	6.46(1.42)	0.60(0.76)	1.03(0.05)	9.52(1.73)	8.94(5.53)	0.86(0.06)	1.52(0.36)
GEL	11.76(1.19)	28.24(1.19)	2.21(0.33)	7.70(1.63)	372.30(1.63)	6.81(1.21)	15.48(6.96)
				S2			
proposed	14.56(1.58)	2.80(2.39)	0.60(0.10)	14.05(1.41)	4.34(3.31)	0.54(0.11)	0.72(0.19)
triBayes	13.22(3.07)	6.02(6.71)	1.46(0.25)	1.46(2.41)	0.10(0.36)	1.02(0.07)	2.47(1.01)
glinternet	14.96(2.51)	4.40(2.65)	0.71(0.10)	12.16(2.88)	3.08(2.20)	0.77(0.09)	1.13(0.33)
Lasso	5.46(2.28)	0.20(0.57)	0.98(0.06)	6.12(3.20)	50.92(106.17)	0.98(0.08)	1.65(0.45)
iFORM	10.96(2.63)	45.04(3.90)	1.59(0.21)	5.76(2.58)	34.68(4.06)	1.41(0.14)	2.39(0.68)
HierNet	9.80(2.59)	0.86(1.05)	0.88(0.08)	7.50(2.43)	6.72(4.29)	0.93(0.08)	1.31(0.34)
Grace	4.16(1.23)	0.32(0.55)	1.34(0.03)	3.84(1.38)	7.24(6.53)	0.99(0.04)	2.12(0.55)
GEL	11.50(1.72)	26.74(5.83)	2.14(0.50)	8.78(1.83)	352.86(75.75)	6.39(1.83)	13.56(6.71)
				S3			
proposed	13.40(2.35)	3.46(3.42)	0.82(0.07)	8.23(2.31)	4.33(3.46)	0.81(0.08)	0.88(0.19)
triBayes	10.02(1.77)	1.22(1.00)	1.73(0.02)	0.00(0.00)	0.00(0.00)	1.03(0.00)	2.45(0.62)
glinternet	9.70(2.98)	2.14(2.08)	1.02(0.10)	3.92(2.52)	1.20(1.47)	0.96(0.05)	1.36(0.33)
Lasso	5.10(1.56)	0.14(0.45)	1.21(0.09)	5.40(1.80)	7.94(5.28)	0.93(0.05)	1.52(0.36)
iFORM	11.96(2.45)	44.62(3.87)	1.54(0.20)	4.40(2.34)	35.80(4.32)	1.38(0.15)	2.34(0.65)
HierNet	7.40(2.20)	0.58(0.70)	1.08(0.09)	3.58(1.74)	4.22(3.89)	0.96(0.04)	1.38(0.33)
Grace	6.56(1.58)	0.68(0.91)	1.15(0.09)	8.08(1.78)	12.04(9.38)	0.88(0.04)	1.35(0.31)
GEL	10.22(3.01)	23.94(10.64)	1.96(0.57)	5.56(2.58)	310.62(138.18)	5.64(3.00)	12.98(12.07)
				S4			
proposed	14.58(2.34)	2.79(3.29)	0.64(0.10)	6.50(3.07)	7.38(2.80)	0.93(0.14)	0.91(0.23)
triBayes	13.56(2.67)	5.82(7.92)	1.55(0.23)	0.00(0.00)	0.36(0.78)	1.04(0.02)	2.74(1.11)
glinternet	11.62(3.28)	1.66(1.89)	0.85(0.10)	3.12(2.33)	2.66(1.86)	1.00(0.04)	1.28(0.31)
Lasso	5.20(1.58)	0.06(0.24)	1.01(0.04)	2.22(1.75)	6.40(3.73)	1.02(0.03)	1.51(0.38)
iFORM	9.42(2.17)	48.10(4.54)	1.91(0.22)	0.96(1.37)	39.58(3.64)	1.67(0.14)	3.43(0.98)
HierNet	8.90(2.59)	0.50(0.74)	0.93(0.08)	2.48(1.85)	5.14(3.33)	1.10(0.09)	1.33(0.34)
Grace	6.18(1.60)	0.54(0.68)	1.01(0.04)	2.26(1.86)	7.78(5.47)	1.05(0.04)	1.54(0.37)
GEL	11.92(0.57)	27.44(3.96)	2.26(0.55)	7.50(1.45)	364.96(52.54)	6.88(1.96)	16.05(9.81)

Table S5: Simulation results under the scenarios with $\rho = 0.6, K = 100$, and $r = 1/\sqrt{5}$. In each cell, mean (SD) based on 100 replicates.

Approach	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
				S1			
proposed	17.96(0.20)	1.94(1.12)	0.42(0.08)	16.90(0.36)	6.58(2.78)	0.44(0.10)	0.54(0.12)
triBayes	17.46(0.95)	12.60(2.32)	1.21(0.07)	7.94(1.32)	4.96(1.28)	1.27(0.06)	2.37(0.65)
glinternet	17.98(0.14)	6.56(3.60)	0.85(0.12)	16.02(1.00)	8.34(3.84)	0.84(0.12)	1.15(0.36)
Lasso	5.88(1.60)	0.02(0.14)	1.58(0.06)	10.64(1.88)	6.92(3.71)	1.26(0.11)	2.77(0.71)
iFORM	15.20(1.96)	38.54(3.86)	1.59(0.31)	8.18(2.98)	34.08(3.39)	1.72(0.24)	3.35(1.45)
HierNet	16.30(1.80)	1.72(1.85)	1.18(0.21)	9.62(2.32)	10.86(4.60)	1.35(0.13)	3.06(1.23)
Grace	10.10(3.37)	1.18(2.25)	1.49(0.27)	12.22(1.94)	7.10(7.84)	1.54(0.15)	2.57(0.93)
GEL	17.88(0.33)	3.38(2.73)	0.68(0.17)	12.10(1.04)	28.54(28.02)	1.25(0.42)	1.15(0.38)
				S2			
proposed	17.57(0.63)	1.43(1.04)	0.41(0.10)	16.87(0.43)	6.10(2.26)	0.47(0.08)	0.73(0.21)
triBayes	17.18(1.52)	10.12(2.39)	1.25(0.26)	6.86(2.64)	2.52(1.78)	1.38(0.12)	2.86(1.54)
glinternet	17.74(0.63)	5.44(2.90)	0.90(0.12)	15.78(1.17)	7.02(3.20)	0.87(0.11)	1.07(0.26)
Lasso	12.02(2.54)	0.30(0.54)	1.42(0.14)	13.18(1.97)	33.44(61.65)	1.15(0.13)	1.97(0.60)
iFORM	14.38(2.74)	38.68(4.62)	1.73(0.47)	8.02(3.23)	34.78(3.66)	1.78(0.27)	2.91(1.28)
HierNet	15.80(2.04)	1.26(1.27)	1.25(0.18)	8.90(1.91)	10.18(4.74)	1.49(0.12)	2.12(0.60)
Grace	5.02(1.45)	0.30(0.51)	1.78(0.06)	4.70(1.25)	2.72(3.02)	1.52(0.07)	4.71(1.57)
GEL	17.36(1.27)	5.74(4.56)	0.83(0.24)	12.52(1.05)	57.38(54.29)	1.73(0.88)	1.39(0.50)
				S3			
proposed	17.92(0.27)	0.92(1.01)	0.64(0.07)	16.33(0.76)	6.92(2.85)	0.69(0.11)	0.61(0.16)
triBayes	12.90(2.25)	9.64(3.41)	1.89(0.22)	2.68(3.08)	1.12(1.59)	1.54(0.11)	3.55(1.31)
glinternet	15.00(1.54)	4.84(2.95)	1.28(0.10)	9.66(1.78)	5.64(2.99)	1.29(0.08)	1.70(0.43)
Lasso	8.84(1.78)	0.10(0.30)	1.62(0.11)	8.90(1.75)	11.34(4.86)	1.29(0.07)	1.90(0.44)
iFORM	14.00(3.08)	39.68(4.98)	1.79(0.46)	6.38(3.76)	35.78(4.24)	1.86(0.30)	3.23(1.55)
HierNet	11.50(1.28)	0.84(0.98)	1.39(0.11)	6.00(1.76)	6.06(3.76)	1.44(0.08)	2.07(0.55)
Grace	8.48(1.71)	0.34(0.48)	1.59(0.13)	11.46(1.68)	7.40(8.39)	1.66(0.10)	2.17(0.61)
GEL	16.44(2.31)	8.42(4.87)	0.98(0.23)	11.18(2.50)	93.32(59.19)	2.20(0.69)	1.55(0.44)
				S4			
proposed	18.00(0.00)	2.65(1.98)	0.41(0.07)	12.67(2.29)	20.60(4.94)	1.09(0.22)	0.73(0.21)
triBayes	17.36(1.35)	10.14(3.02)	1.30(0.27)	0.70(1.46)	7.60(3.29)	1.69(0.09)	2.81(1.49)
glinternet	17.20(1.37)	5.68(3.35)	0.87(0.20)	11.14(3.04)	9.74(3.74)	1.43(0.16)	1.42(0.43)
Lasso	11.54(2.42)	0.14(0.35)	1.44(0.15)	8.68(2.58)	16.90(6.45)	1.56(0.15)	1.82(0.52)
iFORM	10.28(3.15)	44.20(4.81)	2.71(0.63)	1.82(3.01)	40.28(3.66)	2.58(0.36)	6.21(2.63)
HierNet	15.82(1.86)	1.50(1.18)	1.12(0.17)	9.54(2.28)	10.06(4.09)	1.85(0.19)	1.53(0.39)
Grace	10.46(2.83)	0.50(0.74)	1.53(0.16)	4.24(2.49)	14.76(9.31)	1.98(0.10)	2.23(0.54)
GEL	17.74(0.49)	3.62(2.93)	0.72(0.17)	11.54(0.91)	24.92(33.28)	1.30(0.55)	1.12(0.40)

Table S6: Simulation results under the scenarios with $\rho = 0.6, K = 100$, and $r = 1/\sqrt{12}$. In each cell, mean (SD) based on 100 replicates.

Approach	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
			S1				
proposed	17.36(0.62)	0.46(0.58)	0.47(0.08)	16.29(0.81)	2.07(1.39)	0.35(0.10)	0.54(0.11)
triBayes	16.76(2.31)	9.32(2.78)	1.22(0.21)	5.64(2.55)	2.82(1.52)	0.89(0.07)	1.85(1.15)
glinternet	17.02(1.17)	4.44(2.71)	0.64(0.09)	14.04(1.70)	5.14(3.30)	0.67(0.06)	0.99(0.27)
Lasso	6.68(2.01)	0.02(0.14)	1.01(0.06)	10.58(1.60)	10.20(4.95)	0.82(0.07)	1.40(0.33)
iFORM	10.92(1.93)	44.24(3.84)	1.68(0.17)	3.58(1.43)	37.24(3.15)	1.49(0.10)	3.14(0.90)
HierNet	13.44(2.07)	0.64(0.90)	0.88(0.11)	8.08(1.99)	7.42(3.90)	0.95(0.08)	1.69(0.52)
Grace	8.08(2.20)	0.42(0.67)	1.01(0.09)	10.62(1.96)	5.22(5.86)	1.00(0.09)	1.64(0.50)
GEL	17.08(1.07)	3.82(3.77)	0.64(0.15)	10.62(1.74)	36.90(44.59)	1.12(0.64)	1.03(0.40)
			S2				
proposed	16.58(1.14)	0.54(0.91)	0.48(0.07)	16.14(1.05)	2.22(1.72)	0.37(0.10)	0.56(0.13)
triBayes	14.40(3.31)	4.86(5.07)	1.39(0.25)	2.64(3.29)	0.46(0.81)	0.99(0.09)	2.82(1.38)
glinternet	16.80(1.14)	4.16(2.35)	0.62(0.08)	14.00(1.53)	4.64(2.20)	0.69(0.07)	0.89(0.21)
Lasso	7.66(2.45)	0.28(0.54)	0.99(0.08)	9.52(2.45)	21.32(53.10)	0.87(0.08)	1.47(0.36)
iFORM	10.30(2.72)	44.66(4.16)	1.72(0.23)	3.62(1.58)	37.10(3.07)	1.49(0.09)	2.62(0.79)
HierNet	12.86(2.76)	0.72(1.05)	0.88(0.12)	8.00(1.88)	9.26(5.38)	1.00(0.10)	1.32(0.33)
Grace	4.68(1.54)	0.24(0.43)	1.34(0.04)	4.36(1.26)	2.82(3.04)	0.99(0.04)	2.71(0.75)
GEL	16.14(2.14)	5.18(4.72)	0.73(0.20)	11.24(1.86)	52.62(56.06)	1.38(0.84)	1.21(0.42)
			S3				
proposed	17.24(0.87)	0.59(0.80)	0.59(0.09)	14.41(1.67)	2.65(1.98)	0.56(0.11)	0.59(0.15)
triBayes	9.22(1.81)	1.38(1.88)	1.72(0.05)	0.14(0.99)	0.00(0.00)	1.03(0.02)	2.66(0.57)
glinternet	12.86(1.97)	3.14(2.30)	0.99(0.07)	7.64(2.05)	2.84(2.22)	0.88(0.05)	1.17(0.26)
Lasso	7.56(1.66)	0.10(0.30)	1.12(0.10)	7.58(1.58)	12.34(5.19)	0.87(0.05)	1.23(0.30)
iFORM	10.82(1.76)	44.36(4.01)	1.63(0.19)	3.16(1.61)	37.78(3.72)	1.50(0.13)	2.69(0.76)
HierNet	10.22(1.52)	0.72(0.95)	1.00(0.07)	5.78(1.80)	5.70(4.09)	0.93(0.05)	1.18(0.29)
Grace	7.32(1.50)	0.28(0.50)	1.14(0.11)	9.98(1.41)	6.06(5.88)	1.04(0.07)	1.33(0.37)
GEL	15.42(2.91)	7.16(5.16)	0.88(0.15)	9.88(2.84)	78.68(63.44)	1.77(0.85)	1.31(0.55)
			S4				
proposed	16.73(1.39)	0.92(1.03)	0.53(0.09)	6.96(3.05)	12.98(3.19)	1.03(0.16)	0.71(0.21)
triBayes	13.84(2.48)	3.32(4.47)	1.53(0.22)	0.04(0.28)	1.72(2.74)	1.06(0.05)	3.23(1.35)
glinternet	15.64(1.90)	3.20(2.79)	0.72(0.10)	6.42(2.42)	7.40(2.85)	1.02(0.07)	1.09(0.25)
Lasso	8.54(1.79)	0.08(0.27)	0.98(0.07)	5.60(2.46)	13.80(5.64)	1.05(0.09)	1.23(0.32)
iFORM	8.80(1.80)	46.26(3.24)	2.04(0.29)	0.70(1.05)	41.16(2.90)	1.83(0.14)	3.75(1.17)
HierNet	12.64(2.73)	0.72(0.86)	0.86(0.12)	5.74(2.11)	8.34(4.09)	1.37(0.17)	1.15(0.28)
Grace	8.30(2.35)	0.46(0.65)	1.01(0.07)	2.90(2.15)	12.70(8.77)	1.25(0.08)	1.45(0.36)
GEL	16.64(1.41)	2.82(2.46)	0.62(0.14)	10.26(1.40)	17.02(25.53)	0.89(0.39)	0.89(0.30)

Table S7: Simulation results under the scenarios with $\rho = 0.6, K = 50$, and $r = 1/\sqrt{5}$. In each cell, mean (SD) based on 100 replicates.

Approach	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
				S1			
proposed	17.76(0.43)	2.89(1.72)	0.49(0.10)	16.73(0.50)	5.76(2.89)	0.46(0.11)	0.57(0.12)
triBayes	17.28(1.11)	28.18(5.68)	1.51(0.07)	7.94(1.36)	11.22(2.43)	1.33(0.07)	2.41(0.61)
glinternet	17.90(0.30)	9.30(4.00)	0.88(0.11)	15.70(0.99)	11.52(4.00)	0.89(0.11)	1.24(0.30)
Lasso	5.66(1.61)	0.14(0.40)	1.58(0.06)	9.40(1.95)	11.04(4.41)	1.36(0.14)	2.84(0.66)
iFORM	15.24(2.21)	38.10(4.06)	1.64(0.36)	8.04(2.95)	34.62(3.65)	1.71(0.27)	3.55(1.76)
HierNet	15.18(2.66)	2.40(2.04)	1.27(0.21)	8.00(2.31)	11.22(5.67)	1.45(0.14)	3.21(1.02)
Grace	8.92(2.90)	1.64(1.63)	1.58(0.19)	11.44(1.75)	9.90(7.25)	1.61(0.13)	2.76(0.84)
GEL	12.00(1.12)	26.82(5.46)	4.15(1.26)	8.26(1.80)	357.26(70.10)	15.05(4.62)	39.11(25.71)
				S2			
proposed	16.94(0.89)	1.73(1.57)	0.53(0.14)	16.88(0.32)	5.65(2.83)	0.45(0.10)	0.66(0.14)
triBayes	17.86(0.40)	25.76(4.61)	1.42(0.08)	7.98(1.39)	3.82(2.30)	1.34(0.06)	2.16(0.62)
glinternet	17.84(0.42)	8.54(3.69)	0.87(0.11)	15.72(1.09)	10.50(3.78)	0.87(0.10)	1.07(0.30)
Lasso	11.76(2.83)	0.80(0.90)	1.37(0.16)	12.76(1.92)	32.12(47.66)	1.20(0.12)	1.99(0.75)
iFORM	14.66(2.34)	38.70(4.07)	1.71(0.36)	8.82(2.77)	33.46(3.19)	1.76(0.26)	2.91(1.06)
HierNet	15.94(1.52)	2.38(1.84)	1.20(0.19)	9.42(1.76)	13.34(5.21)	1.45(0.13)	2.02(0.51)
Grace	5.12(1.41)	0.48(0.71)	1.78(0.08)	5.14(1.28)	4.72(4.01)	1.51(0.06)	4.30(1.56)
GEL	11.48(1.67)	25.48(7.33)	3.74(1.19)	8.92(1.94)	337.92(98.12)	12.85(4.33)	28.55(13.46)
				S3			
proposed	17.40(0.87)	1.85(1.96)	0.76(0.10)	15.27(1.78)	6.08(3.11)	0.79(0.18)	0.69(0.19)
triBayes	13.12(2.56)	16.14(5.26)	1.73(0.29)	4.86(2.16)	3.70(2.10)	1.44(0.22)	2.35(1.05)
glinternet	13.96(1.59)	5.18(3.01)	1.37(0.11)	8.44(1.91)	5.98(3.42)	1.35(0.08)	1.94(0.57)
Lasso	7.98(2.09)	0.40(0.61)	1.65(0.11)	8.20(1.69)	15.64(6.53)	1.35(0.08)	2.02(0.61)
iFORM	14.12(2.68)	39.72(4.66)	1.81(0.44)	6.26(3.47)	35.92(3.72)	1.87(0.28)	3.30(1.52)
HierNet	11.18(1.83)	1.88(1.33)	1.42(0.12)	5.62(1.89)	7.82(4.58)	1.46(0.07)	2.11(0.72)
Grace	8.14(1.85)	1.24(1.19)	1.63(0.12)	11.08(1.43)	12.38(7.87)	1.69(0.13)	2.11(0.47)
GEL	10.84(3.16)	23.42(10.74)	2.75(0.85)	6.26(2.28)	307.46(149.32)	8.74(4.04)	15.47(11.49)
				S4			
proposed	17.37(0.86)	2.52(1.82)	0.57(0.11)	9.85(2.48)	17.60(3.54)	1.33(0.20)	0.92(0.27)
triBayes	17.58(0.84)	28.44(4.95)	1.50(0.06)	0.28(0.57)	9.70(2.22)	1.74(0.04)	2.37(0.55)
glinternet	17.40(1.25)	8.04(3.95)	0.90(0.19)	10.74(3.19)	11.56(3.99)	1.47(0.17)	1.39(0.39)
Lasso	10.68(2.86)	0.42(0.70)	1.48(0.17)	7.90(3.10)	18.48(7.25)	1.59(0.17)	1.87(0.56)
iFORM	10.38(3.08)	43.18(4.42)	2.70(0.62)	1.82(2.63)	41.38(3.67)	2.53(0.30)	5.44(2.13)
HierNet	14.88(2.65)	2.52(1.89)	1.18(0.24)	8.16(2.57)	9.58(3.89)	1.90(0.20)	1.64(0.49)
Grace	8.74(2.13)	1.16(0.96)	1.60(0.10)	2.78(2.05)	14.18(6.47)	2.02(0.11)	2.44(0.61)
GEL	12.02(0.14)	27.96(0.20)	3.99(0.88)	7.80(1.32)	371.96(2.18)	13.91(2.58)	33.54(16.15)

Table S8: Simulation results under the scenarios with $\rho = 0.6, K = 50$, and $r = 1/\sqrt{12}$. In each cell, mean (SD) based on 100 replicates.

Approach	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
				S1			
proposed	16.54(1.16)	1.34(1.33)	0.55(0.09)	15.08(1.52)	2.28(1.98)	0.44(0.13)	0.60(0.14)
triBayes	17.28(1.03)	23.82(4.87)	1.35(0.05)	6.36(1.14)	4.80(1.43)	0.88(0.04)	1.47(0.31)
glinternet	17.04(1.09)	7.66(3.19)	0.67(0.09)	13.52(1.66)	8.60(3.53)	0.71(0.07)	1.04(0.26)
Lasso	5.86(1.53)	0.32(0.74)	1.01(0.05)	8.96(1.85)	13.98(5.19)	0.89(0.09)	1.47(0.35)
iFORM	10.78(2.13)	44.02(2.69)	1.70(0.21)	3.94(1.39)	37.24(2.95)	1.47(0.10)	3.18(0.83)
HierNet	12.58(2.94)	1.74(1.37)	0.90(0.11)	6.84(1.96)	9.84(5.69)	0.99(0.09)	1.76(0.56)
Grace	7.12(2.21)	1.46(1.27)	1.04(0.10)	10.28(1.67)	8.94(5.44)	1.04(0.09)	1.63(0.38)
GEL	12.02(0.71)	26.92(5.35)	3.06(0.83)	8.20(1.54)	357.86(68.66)	11.06(3.63)	21.03(12.65)
				S2			
proposed	15.82(1.17)	0.88(1.48)	0.55(0.08)	16.08(1.18)	2.26(1.69)	0.38(0.12)	0.59(0.16)
triBayes	17.86(0.35)	21.96(4.31)	1.28(0.05)	5.80(1.76)	0.52(0.65)	0.90(0.05)	1.36(0.41)
glinternet	16.78(1.11)	7.36(3.12)	0.63(0.07)	14.18(1.19)	7.64(3.24)	0.69(0.05)	0.91(0.22)
Lasso	8.24(2.44)	0.62(0.83)	0.95(0.07)	9.78(2.26)	31.14(65.03)	0.89(0.08)	1.41(0.46)
iFORM	10.50(2.53)	44.44(3.73)	1.74(0.24)	4.52(1.96)	36.62(3.53)	1.48(0.13)	2.58(0.74)
HierNet	11.98(2.93)	1.64(1.86)	0.88(0.11)	7.68(2.33)	9.36(5.59)	0.98(0.10)	1.30(0.37)
Grace	4.38(1.41)	0.56(0.76)	1.34(0.04)	4.52(1.27)	4.14(3.18)	0.99(0.04)	2.59(0.77)
GEL	11.90(1.07)	27.62(2.16)	2.85(0.63)	9.24(1.45)	367.82(28.47)	10.51(2.27)	17.78(9.47)
				S3			
proposed	16.35(1.53)	1.52(1.64)	0.72(0.09)	12.92(2.56)	2.69(1.86)	0.63(0.11)	0.66(0.20)
triBayes	12.32(3.58)	10.44(8.30)	1.58(0.14)	2.76(2.49)	0.46(0.93)	0.97(0.06)	1.83(0.81)
glinternet	11.78(2.42)	4.92(2.91)	1.03(0.09)	6.46(2.13)	4.52(3.03)	0.91(0.05)	1.24(0.36)
Lasso	6.72(2.07)	0.46(0.73)	1.14(0.10)	6.90(1.49)	15.72(5.21)	0.92(0.05)	1.27(0.34)
iFORM	11.48(2.11)	43.56(3.27)	1.64(0.18)	4.02(1.73)	36.56(3.28)	1.43(0.13)	2.51(0.79)
HierNet	9.22(2.22)	1.46(1.28)	1.04(0.09)	4.88(1.96)	6.50(3.87)	0.96(0.04)	1.26(0.38)
Grace	6.90(1.42)	1.44(1.20)	1.16(0.11)	9.90(1.22)	10.14(8.05)	1.06(0.09)	1.31(0.31)
GEL	11.64(1.72)	25.40(8.49)	2.22(0.70)	6.68(1.50)	333.34(122.33)	7.27(2.79)	9.85(5.97)
				S4			
proposed	15.02(2.21)	0.79(1.35)	0.63(0.11)	4.65(2.76)	10.75(3.07)	1.15(0.16)	0.86(0.19)
triBayes	17.68(0.62)	24.12(4.30)	1.34(0.04)	0.00(0.00)	5.14(1.46)	1.14(0.03)	1.41(0.34)
glinternet	15.02(1.95)	5.70(3.49)	0.74(0.12)	5.84(2.52)	8.58(3.15)	1.04(0.08)	1.13(0.28)
Lasso	8.20(2.33)	0.46(0.76)	0.99(0.09)	5.48(2.34)	15.06(4.86)	1.07(0.08)	1.29(0.35)
iFORM	8.18(2.34)	47.06(3.44)	2.20(0.32)	0.38(0.75)	41.08(2.57)	1.84(0.13)	3.71(1.09)
HierNet	12.22(3.07)	1.96(1.67)	0.87(0.12)	5.00(2.39)	8.52(3.42)	1.41(0.14)	1.18(0.31)
Grace	7.02(1.91)	1.12(1.06)	1.02(0.06)	1.94(1.62)	12.12(6.70)	1.27(0.08)	1.58(0.41)
GEL	12.00(0.00)	28.00(0.00)	3.19(0.71)	7.90(1.39)	372.10(1.39)	11.53(2.54)	23.31(13.88)

Table S9: Simulation results under the scenarios with $\rho = 0.4, K = 100$, and $r = 1/\sqrt{5}$, where some genetic factors are involved in multiple networks. In each cell, mean (SD) based on 100 replicates.

Approach	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
				S1			
proposed	17.82(0.39)	2.02(1.35)	0.37(0.11)	16.60(0.57)	9.14(4.05)	0.49(0.10)	0.59(0.14)
triBayes	16.30(1.26)	8.95(1.79)	1.32(0.40)	5.50(3.35)	2.75(2.15)	1.36(0.15)	2.50(1.21)
glinternet	17.60(0.68)	5.55(2.98)	0.92(0.12)	14.70(1.53)	4.35(2.21)	1.01(0.13)	1.43(0.43)
Lasso	8.20(3.17)	0.00(0.00)	1.48(0.12)	11.50(1.96)	15.60(7.58)	1.19(0.11)	2.29(0.62)
iFORM	16.90(1.59)	36.35(3.57)	1.30(0.30)	13.20(3.58)	30.40(3.55)	1.29(0.33)	2.18(1.81)
HierNet	14.35(2.37)	1.20(1.77)	1.19(0.17)	9.05(2.91)	9.05(7.10)	1.31(0.14)	2.18(0.78)
Grace	9.90(2.34)	0.35(0.93)	1.45(0.14)	12.00(1.72)	8.55(8.53)	1.20(0.11)	2.14(0.57)
GEL	18.00(0.00)	10.90(3.39)	0.72(0.13)	13.15(0.88)	123.45(37.83)	2.06(0.42)	1.96(0.96)
				S2			
proposed	16.98(1.10)	1.48(1.20)	0.46(0.17)	16.12(1.00)	7.44(3.70)	0.57(0.14)	0.68(0.17)
triBayes	16.00(1.34)	4.05(3.20)	1.81(0.25)	0.75(2.31)	0.15(0.49)	1.63(0.10)	4.33(1.06)
glinternet	17.65(0.81)	5.05(3.43)	0.96(0.12)	14.55(1.64)	4.25(2.75)	1.07(0.10)	1.36(0.51)
Lasso	6.20(2.50)	0.15(0.37)	1.50(0.07)	7.65(2.78)	49.35(106.84)	1.47(0.07)	2.66(0.65)
iFORM	15.55(2.70)	38.70(4.54)	1.53(0.44)	11.45(4.63)	31.60(4.48)	1.49(0.45)	2.47(1.23)
HierNet	13.00(4.13)	1.15(1.27)	1.27(0.20)	9.05(2.80)	8.75(6.37)	1.38(0.13)	1.92(0.69)
Grace	5.40(1.39)	0.25(0.55)	1.75(0.05)	4.75(1.41)	6.75(5.18)	1.51(0.05)	3.36(1.03)
GEL	18.00(0.00)	11.20(2.48)	0.72(0.12)	13.20(0.62)	124.70(37.91)	2.05(0.41)	1.68(0.43)
				S3			
proposed	17.58(0.73)	1.30(1.22)	0.61(0.09)	14.16(1.36)	8.78(3.58)	0.84(0.11)	0.71(0.15)
triBayes	12.25(3.06)	4.30(2.41)	1.94(0.46)	0.00(0.00)	0.00(0.00)	1.52(0.36)	3.30(1.11)
glinternet	13.70(2.25)	3.25(2.83)	1.36(0.16)	7.25(2.59)	2.00(1.78)	1.38(0.11)	1.96(0.51)
Lasso	6.75(1.89)	0.05(0.22)	1.73(0.11)	7.50(1.64)	7.30(3.81)	1.36(0.08)	2.28(0.62)
iFORM	16.85(1.60)	37.10(4.17)	1.34(0.28)	11.50(3.75)	31.70(3.01)	1.39(0.31)	2.05(0.80)
HierNet	10.65(1.93)	0.95(0.89)	1.46(0.13)	5.50(1.91)	6.25(3.39)	1.44(0.08)	2.07(0.68)
Grace	9.20(1.36)	0.15(0.37)	1.54(0.11)	9.50(1.28)	11.90(10.48)	1.29(0.08)	1.93(0.51)
GEL	17.70(0.98)	11.45(4.07)	0.79(0.16)	12.35(0.99)	127.90(40.54)	2.11(0.39)	1.88(0.67)
				S4			
proposed	17.88(0.39)	2.14(1.26)	0.38(0.09)	12.76(2.07)	15.18(3.54)	1.00(0.17)	0.78(0.19)
triBayes	15.90(1.41)	2.25(1.86)	1.95(0.03)	0.00(0.00)	0.00(0.00)	1.60(0.00)	4.52(1.47)
glinternet	15.95(2.72)	5.10(3.60)	1.02(0.18)	8.25(3.32)	4.90(2.75)	1.42(0.08)	1.78(0.62)
Lasso	6.70(1.89)	0.00(0.00)	1.53(0.08)	3.85(2.37)	8.55(6.33)	1.54(0.07)	2.77(0.99)
iFORM	13.10(3.78)	42.80(6.44)	2.07(0.70)	5.25(5.66)	37.20(4.03)	2.04(0.50)	4.44(2.51)
HierNet	13.30(3.34)	1.00(1.38)	1.21(0.16)	6.10(2.86)	9.10(5.63)	1.62(0.10)	1.94(0.53)
Grace	10.00(2.27)	0.10(0.31)	1.44(0.11)	4.85(2.39)	11.80(7.51)	1.59(0.07)	2.25(0.81)
GEL	17.90(0.45)	10.25(3.65)	0.71(0.14)	12.65(1.69)	112.30(50.97)	1.96(0.49)	1.73(0.64)

Table S10: Simulation results under the scenarios with $\rho = 0.4, K = 100$, and $r = 1/\sqrt{12}$, where some genetic factors are involved in multiple networks. In each cell, mean (SD) based on 100 replicates.

Approach	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
				S1			
proposed	17.00(0.97)	1.00(1.12)	0.43(0.08)	15.32(1.19)	4.48(2.99)	0.43(0.10)	0.59(0.15)
triBayes	13.40(2.01)	2.45(2.42)	1.61(0.22)	0.65(2.01)	0.30(0.92)	1.01(0.05)	2.72(0.91)
glinternet	15.70(2.94)	4.30(3.11)	0.72(0.09)	11.70(2.74)	2.55(1.70)	0.77(0.08)	1.12(0.31)
Lasso	5.90(2.05)	0.00(0.00)	1.02(0.07)	9.25(2.24)	12.55(6.04)	0.83(0.07)	1.42(0.37)
iFORM	11.60(2.28)	45.05(2.61)	1.56(0.19)	5.30(2.11)	34.55(3.69)	1.38(0.13)	2.57(0.94)
HierNet	10.60(2.84)	0.40(0.50)	0.91(0.08)	5.80(2.14)	4.80(3.52)	0.93(0.06)	1.42(0.33)
Grace	6.85(1.09)	0.10(0.31)	1.01(0.04)	10.10(1.68)	6.15(4.66)	0.82(0.06)	1.44(0.43)
GEL	17.75(0.64)	9.50(4.52)	0.68(0.12)	12.30(1.26)	105.70(56.25)	1.68(0.56)	1.50(0.58)
				S2			
proposed	16.04(1.58)	0.88(0.80)	0.45(0.10)	14.12(1.60)	3.32(2.14)	0.53(0.12)	0.65(0.18)
triBayes	12.35(1.53)	0.40(0.60)	1.62(0.02)	0.00(0.00)	0.00(0.00)	1.07(0.00)	2.99(0.63)
glinternet	14.45(3.86)	3.10(2.55)	0.75(0.13)	10.60(3.79)	2.55(2.24)	0.83(0.10)	1.10(0.42)
Lasso	4.90(1.68)	0.15(0.37)	0.99(0.06)	5.25(2.24)	47.75(106.05)	1.00(0.06)	1.53(0.37)
iFORM	11.15(2.58)	43.95(5.12)	1.64(0.27)	4.60(2.06)	37.35(4.08)	1.46(0.14)	2.63(0.95)
HierNet	8.25(2.47)	0.20(0.52)	0.94(0.05)	5.30(2.05)	3.75(2.02)	0.98(0.06)	1.36(0.39)
Grace	4.05(1.10)	0.20(0.52)	1.35(0.04)	3.45(1.39)	4.85(4.97)	1.00(0.03)	2.18(0.65)
GEL	17.50(1.28)	10.25(4.28)	0.70(0.10)	12.50(1.32)	119.20(47.25)	1.81(0.49)	1.62(0.55)
				S3			
proposed	15.96(1.46)	1.18(1.40)	0.61(0.08)	11.32(1.77)	5.00(3.55)	0.70(0.08)	0.68(0.16)
triBayes	10.80(1.28)	0.75(0.64)	1.72(0.02)	0.00(0.00)	0.00(0.00)	1.03(0.00)	2.18(0.50)
glinternet	10.50(2.50)	1.70(1.81)	1.01(0.08)	4.55(2.61)	0.85(1.42)	0.95(0.06)	1.28(0.27)
Lasso	5.20(1.58)	0.00(0.00)	1.21(0.08)	5.55(1.61)	6.20(3.87)	0.92(0.05)	1.41(0.42)
iFORM	11.85(2.37)	45.00(4.28)	1.53(0.18)	3.95(2.54)	36.60(3.59)	1.39(0.10)	2.31(0.71)
HierNet	8.80(1.94)	0.70(0.92)	1.05(0.08)	4.30(2.27)	5.65(4.00)	0.96(0.06)	1.30(0.35)
Grace	7.40(1.50)	0.10(0.31)	1.14(0.08)	7.95(1.32)	8.65(8.51)	0.87(0.04)	1.25(0.37)
GEL	17.40(1.96)	10.85(3.54)	0.71(0.11)	11.65(1.76)	126.15(40.81)	1.84(0.37)	1.52(0.48)
				S4			
proposed	16.74(1.31)	1.14(1.18)	0.46(0.09)	8.56(2.33)	9.26(2.82)	0.86(0.12)	0.77(0.17)
triBayes	13.75(1.68)	0.45(0.60)	1.67(0.02)	0.00(0.00)	0.00(0.00)	1.03(0.00)	3.13(1.02)
glinternet	12.11(2.49)	1.26(1.88)	0.84(0.09)	3.42(1.84)	2.58(2.09)	0.99(0.03)	1.18(0.14)
Lasso	5.26(1.41)	0.00(0.00)	1.01(0.05)	2.37(1.57)	5.84(3.25)	1.01(0.03)	1.43(0.21)
iFORM	9.84(1.68)	48.05(3.08)	1.88(0.15)	1.16(1.12)	38.58(3.17)	1.63(0.12)	3.05(0.92)
HierNet	8.68(2.00)	0.32(0.58)	0.93(0.08)	2.74(1.66)	5.74(4.16)	1.08(0.07)	1.31(0.27)
Grace	7.16(2.43)	0.21(0.71)	0.99(0.08)	2.89(2.47)	7.95(10.37)	1.04(0.03)	1.45(0.28)
GEL	15.79(2.70)	4.95(5.58)	0.74(0.14)	10.05(2.74)	55.68(65.32)	1.26(0.60)	1.29(0.44)

Table S11: Simulation results under the scenarios with $\rho = 0.6, K = 100$, and $r = 1/\sqrt{5}$, where some genetic factors are involved in multiple networks. In each cell, mean (SD) based on 100 replicates.

Approach	M:TP	M:FP	M:RSSE	I:TP	I:FP	I:RSSE	PMSE
				S1			
proposed	17.90(0.30)	3.72(1.63)	0.43(0.08)	16.16(1.08)	12.74(5.64)	0.57(0.16)	0.61(0.15)
triBayes	17.45(0.89)	12.85(2.21)	1.22(0.08)	7.80(1.54)	4.80(1.06)	1.28(0.07)	2.30(0.63)
glinternet	18.00(0.00)	7.50(4.50)	0.81(0.12)	16.10(0.91)	8.40(4.57)	0.82(0.11)	1.11(0.33)
Lasso	5.80(1.11)	0.05(0.22)	1.58(0.06)	10.90(1.59)	6.85(3.54)	1.23(0.09)	2.70(0.68)
iFORM	15.40(2.09)	39.20(4.27)	1.56(0.34)	8.75(3.11)	32.85(3.60)	1.66(0.27)	3.05(1.37)
HierNet	16.45(1.67)	2.25(1.74)	1.13(0.20)	9.60(2.52)	11.00(4.26)	1.37(0.15)	2.89(1.25)
Grace	9.95(3.10)	1.00(2.13)	1.50(0.24)	11.20(1.82)	5.85(4.70)	1.54(0.16)	2.67(0.96)
GEL	17.90(0.31)	3.75(2.81)	0.65(0.14)	12.30(1.08)	25.95(26.68)	1.23(0.40)	1.14(0.39)
				S2			
proposed	17.34(0.77)	2.88(1.55)	0.46(0.11)	16.14(1.05)	10.28(4.78)	0.57(0.15)	0.63(0.17)
triBayes	17.30(1.42)	10.00(2.13)	1.21(0.24)	7.15(2.35)	2.80(1.77)	1.38(0.11)	2.60(1.48)
glinternet	17.90(0.45)	6.15(3.20)	0.87(0.13)	15.95(1.23)	7.40(3.47)	0.86(0.11)	0.96(0.24)
Lasso	12.85(2.11)	0.35(0.59)	1.42(0.15)	13.45(1.73)	57.25(93.46)	1.15(0.11)	1.71(0.40)
iFORM	14.30(2.85)	38.70(4.74)	1.75(0.55)	7.05(2.84)	35.60(3.02)	1.87(0.26)	2.94(1.56)
HierNet	15.65(2.13)	0.80(1.01)	1.27(0.19)	8.80(2.12)	9.45(3.97)	1.52(0.14)	1.88(0.57)
Grace	5.30(1.38)	0.15(0.37)	1.81(0.06)	4.85(1.14)	2.55(2.26)	1.50(0.05)	4.07(1.31)
GEL	17.70(0.92)	5.00(4.39)	0.77(0.21)	12.70(0.80)	48.45(51.33)	1.62(0.99)	1.22(0.46)
				S3			
proposed	17.78(0.46)	1.50(1.39)	0.69(0.10)	14.50(1.47)	8.02(4.05)	0.84(0.14)	0.65(0.17)
triBayes	13.30(2.08)	10.50(3.33)	1.84(0.23)	3.05(3.09)	1.50(1.79)	1.52(0.11)	3.10(1.16)
glinternet	14.90(1.52)	4.95(3.39)	1.30(0.10)	9.70(1.81)	5.60(2.50)	1.30(0.07)	1.72(0.41)
Lasso	8.85(1.66)	0.15(0.37)	1.62(0.09)	8.90(1.48)	12.00(5.15)	1.28(0.08)	1.68(0.41)
iFORM	14.20(3.21)	38.95(5.05)	1.77(0.52)	6.35(4.27)	36.15(3.25)	1.85(0.32)	3.25(1.65)
HierNet	11.45(1.36)	0.85(1.31)	1.39(0.12)	5.30(1.59)	6.50(4.37)	1.48(0.07)	2.08(0.63)
Grace	8.60(1.43)	0.35(0.49)	1.59(0.11)	10.75(1.48)	8.55(9.88)	1.67(0.12)	2.03(0.66)
GEL	16.20(2.42)	7.90(5.23)	0.97(0.22)	11.10(2.47)	89.45(60.16)	2.14(0.68)	1.56(0.50)
				S4			
proposed	17.92(0.34)	3.08(1.90)	0.43(0.10)	11.90(2.54)	19.26(4.96)	1.15(0.22)	0.75(0.19)
triBayes	17.00(1.62)	10.15(3.39)	1.34(0.30)	0.70(1.49)	7.25(3.63)	1.69(0.09)	2.85(1.46)
glinternet	17.70(0.66)	6.30(3.40)	0.79(0.19)	12.00(2.41)	9.85(3.77)	1.40(0.18)	1.38(0.46)
Lasso	11.80(2.59)	0.30(0.47)	1.41(0.15)	8.50(2.74)	17.10(7.18)	1.57(0.17)	1.92(0.59)
iFORM	11.05(3.65)	43.55(5.74)	2.64(0.77)	2.55(4.02)	39.60(4.01)	2.48(0.42)	5.76(2.78)
HierNet	16.00(1.97)	1.55(1.23)	1.10(0.19)	9.15(2.30)	9.90(4.48)	1.86(0.15)	1.54(0.45)
Grace	10.65(2.68)	0.35(0.59)	1.51(0.13)	4.00(2.03)	13.05(6.49)	2.00(0.09)	2.15(0.43)
GEL	17.80(0.52)	3.70(2.54)	0.67(0.15)	11.55(0.76)	21.05(27.94)	1.23(0.44)	1.05(0.39)

Table S12: Simulation results under the scenarios with $\rho = 0.6, K = 100$, and $r = 1/\sqrt{12}$, where some genetic factors are involved in multiple networks. In each cell, mean (SD) based on 100 replicates.

Approach	M:TP	M:FP	M:RSSE	I:TP	I:FP	I:RSSE	PMSE
			S1				
proposed	17.22(0.74)	1.50(1.20)	0.49(0.08)	14.48(1.37)	4.44(3.56)	0.51(0.13)	0.56(0.14)
triBayes	16.85(2.18)	9.65(2.96)	1.21(0.20)	5.80(2.46)	2.80(1.47)	0.88(0.07)	1.72(1.15)
glinternet	17.45(1.10)	4.90(2.75)	0.61(0.07)	14.45(1.54)	5.40(3.12)	0.65(0.06)	0.96(0.25)
Lasso	6.60(1.50)	0.05(0.22)	1.00(0.06)	10.65(1.42)	9.95(5.17)	0.79(0.07)	1.37(0.28)
iFORM	11.40(2.21)	43.35(3.91)	1.67(0.20)	3.60(1.47)	37.75(3.45)	1.50(0.11)	3.20(0.92)
HierNet	13.90(1.68)	0.80(1.15)	0.86(0.12)	7.75(1.80)	6.65(3.27)	0.95(0.09)	1.69(0.48)
Grace	8.45(2.58)	0.50(0.76)	0.99(0.10)	10.25(1.97)	5.55(4.38)	1.00(0.10)	1.58(0.56)
GEL	17.20(0.89)	5.00(4.28)	0.64(0.14)	11.00(1.69)	46.15(54.14)	1.26(0.76)	1.10(0.40)
			S2				
proposed	16.28(1.34)	0.90(0.79)	0.50(0.08)	14.14(1.46)	2.42(1.95)	0.53(0.12)	0.59(0.16)
triBayes	14.00(3.18)	3.70(4.58)	1.43(0.25)	2.00(3.18)	0.40(0.68)	1.01(0.09)	2.85(1.19)
glinternet	17.20(0.83)	4.05(2.16)	0.62(0.08)	14.15(1.27)	4.50(2.16)	0.69(0.06)	0.81(0.20)
Lasso	8.20(2.75)	0.30(0.57)	0.99(0.08)	9.85(1.81)	37.45(82.22)	0.87(0.07)	1.32(0.30)
iFORM	10.60(2.85)	43.80(4.34)	1.69(0.24)	3.65(1.50)	37.60(3.52)	1.48(0.10)	2.41(0.78)
HierNet	13.15(3.15)	0.40(0.82)	0.90(0.11)	8.15(2.11)	9.00(4.92)	1.01(0.11)	1.22(0.33)
Grace	4.60(1.70)	0.20(0.41)	1.35(0.06)	4.40(1.31)	2.35(2.37)	0.98(0.03)	2.44(0.62)
GEL	16.35(1.79)	4.05(4.43)	0.69(0.14)	11.20(1.91)	41.25(49.40)	1.23(0.83)	1.11(0.49)
			S3				
proposed	16.92(1.05)	0.68(0.98)	0.63(0.10)	12.44(1.67)	3.26(2.51)	0.65(0.09)	0.62(0.15)
triBayes	9.05(1.57)	1.30(1.17)	1.72(0.02)	0.00(0.00)	0.00(0.00)	1.03(0.00)	2.44(0.48)
glinternet	13.20(1.94)	3.75(2.24)	0.99(0.08)	8.00(1.92)	3.25(2.10)	0.89(0.06)	1.18(0.29)
Lasso	7.45(1.57)	0.15(0.37)	1.12(0.09)	7.85(1.46)	12.50(5.33)	0.87(0.05)	1.15(0.26)
iFORM	11.35(1.46)	43.30(3.54)	1.57(0.15)	3.55(1.43)	37.55(3.73)	1.50(0.13)	2.63(0.73)
HierNet	10.35(1.53)	1.00(1.08)	1.00(0.07)	5.10(1.62)	6.45(3.82)	0.94(0.06)	1.19(0.31)
Grace	7.50(1.28)	0.20(0.41)	1.14(0.09)	9.65(1.46)	6.80(6.99)	1.05(0.09)	1.23(0.39)
GEL	16.05(2.82)	8.55(4.87)	0.90(0.15)	10.55(2.31)	95.20(60.42)	2.03(0.87)	1.44(0.58)
			S4				
proposed	16.54(1.40)	0.80(0.95)	0.52(0.10)	6.08(2.51)	11.64(2.66)	1.07(0.13)	0.77(0.18)
triBayes	13.75(2.47)	3.10(4.15)	1.54(0.22)	0.00(0.00)	1.70(2.90)	1.05(0.05)	3.29(1.44)
glinternet	16.00(1.65)	3.35(2.46)	0.70(0.11)	6.35(2.48)	6.80(3.11)	1.03(0.08)	1.14(0.33)
Lasso	8.65(1.27)	0.15(0.37)	0.98(0.06)	5.20(2.12)	13.80(5.80)	1.06(0.10)	1.26(0.35)
iFORM	8.85(1.73)	44.90(2.99)	2.04(0.31)	0.70(0.86)	42.60(2.72)	1.87(0.14)	3.80(1.25)
HierNet	12.85(3.03)	0.75(0.79)	0.84(0.10)	5.60(2.33)	8.00(4.08)	1.37(0.15)	1.16(0.29)
Grace	8.25(2.24)	0.45(0.51)	1.00(0.05)	2.55(1.39)	11.10(5.75)	1.26(0.07)	1.42(0.34)
GEL	17.00(1.17)	3.20(2.75)	0.58(0.12)	10.75(1.16)	18.35(28.39)	0.90(0.45)	0.84(0.31)

Table S13: Simulation results under the scenarios with $n = 300, \rho = 0.4, K = 100$, $r = 1/\sqrt{5}$, coefficient type S1, and various values of p. In each cell, mean (SD) based on 100 replicates. Note that since HierNet is not applicable when the dimension of predictors is too high, results are not available for $p = 1200$ and 1400.

Approach	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
				$p = 800$			
proposed	18.00(0.00)	2.50(1.17)	1.85(0.02)	16.90(0.31)	10.37(3.78)	0.45(0.10)	0.56(0.14)
triBayes	15.63(1.71)	9.23(2.69)	2.10(0.17)	0.00(0.00)	9.07(5.51)	0.78(0.44)	2.27(1.04)
glinternet	16.53(1.14)	6.83(3.67)	2.04(0.08)	0.00(0.00)	20.07(4.78)	0.90(0.14)	1.51(0.51)
Lasso	13.60(2.37)	1.17(0.87)	2.17(0.09)	0.00(0.00)	274.03(68.84)	1.26(0.10)	1.99(0.51)
iFORM	15.97(2.04)	35.77(4.02)	2.36(0.25)	0.00(0.00)	44.30(2.45)	1.88(0.19)	2.29(1.15)
HierNet	13.77(2.86)	2.00(1.76)	2.16(0.11)	0.00(0.00)	20.03(8.63)	1.01(0.25)	2.21(0.56)
Grace	8.97(3.00)	0.33(0.71)	2.28(0.07)	11.90(2.14)	7.40(8.33)	1.23(0.11)	2.25(0.62)
GEL	17.87(0.73)	10.03(4.09)	2.02(0.09)	13.00(1.02)	112.77(48.19)	1.96(0.51)	1.79(0.65)
				$p = 1200$			
proposed	17.90(0.31)	2.33(1.21)	1.85(0.02)	16.93(0.25)	9.43(4.11)	0.45(0.08)	0.60(0.13)
triBayes	15.77(1.14)	9.80(2.61)	2.23(0.23)	0.00(0.00)	6.20(5.25)	0.51(0.43)	3.24(1.68)
glinternet	16.40(1.25)	6.67(3.87)	2.07(0.07)	0.00(0.00)	18.07(4.34)	0.82(0.14)	1.64(0.55)
Lasso	12.53(2.10)	1.20(0.85)	2.21(0.07)	0.00(0.00)	296.87(10.75)	1.24(0.10)	2.16(0.71)
iFORM	15.50(2.13)	40.87(3.70)	2.40(0.24)	0.00(0.00)	40.47(2.90)	1.76(0.17)	2.75(1.78)
HierNet	-	-	-	-	-	-	-
Grace	10.10(3.17)	0.57(1.25)	2.26(0.10)	12.40(1.83)	12.00(12.35)	1.21(0.13)	2.30(0.78)
GEL	18.00(0.00)	11.37(2.41)	1.99(0.05)	13.20(0.61)	128.10(31.33)	2.06(0.35)	1.93(0.52)
				$p = 1400$			
proposed	17.90(0.31)	3.47(1.36)	1.85(0.02)	16.73(0.52)	11.73(3.55)	0.53(0.10)	0.66(0.18)
triBayes	15.63(1.38)	10.30(3.42)	2.22(0.22)	0.00(0.00)	6.87(5.13)	0.52(0.38)	3.33(1.67)
glinternet	16.17(2.07)	6.10(3.49)	2.08(0.07)	0.00(0.00)	18.53(4.76)	0.83(0.19)	1.64(0.55)
Lasso	12.37(2.74)	0.83(0.91)	2.22(0.07)	0.00(0.00)	288.77(49.55)	1.25(0.11)	2.14(0.71)
iFORM	14.23(2.66)	44.13(4.68)	2.57(0.37)	0.00(0.00)	38.53(3.34)	1.74(0.15)	3.09(1.35)
HierNet	-	-	-	-	-	-	-
Grace	8.23(2.27)	0.13(0.35)	2.29(0.06)	11.33(1.54)	6.27(6.20)	1.22(0.13)	2.56(0.76)
GEL	18.00(0.00)	10.10(3.91)	2.00(0.06)	12.83(0.95)	111.93(49.25)	1.96(0.49)	1.76(0.54)

Table S14: Simulation results under the scenarios with $n = 300, \rho = 0.4, K = 100$, $r = 1/\sqrt{5}$, coefficient type S2, and various values of p. In each cell, mean (SD) based on 100 replicates. Note that since HierNet is not applicable when the dimension of predictors is too high, results are not available for $p = 1200$ and 1400.

Approach	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
			$p = 800$				
proposed	17.70(0.47)	1.17(0.99)	3.52(0.04)	16.67(0.66)	7.17(2.68)	0.49(0.14)	0.58(0.15)
triBayes	15.40(1.35)	5.33(2.80)	2.77(0.19)	0.00(0.00)	1.83(3.95)	0.17(0.35)	3.45(1.14)
glinternet	16.30(1.12)	6.80(3.54)	3.36(0.10)	0.00(0.00)	19.47(5.20)	0.95(0.12)	1.40(0.48)
Lasso	12.20(2.71)	1.47(1.01)	3.29(0.11)	0.00(0.00)	292.93(52.37)	1.20(0.10)	2.25(0.64)
iFORM	15.03(2.76)	36.90(3.88)	3.88(0.20)	0.00(0.00)	44.67(2.86)	1.87(0.16)	2.49(1.29)
HierNet	11.70(3.29)	1.20(1.06)	3.29(0.16)	0.00(0.00)	14.40(6.90)	1.09(0.25)	2.14(0.70)
Grace	5.40(1.81)	0.03(0.18)	2.77(0.11)	4.73(0.98)	4.87(5.02)	1.51(0.04)	3.31(0.70)
GEL	17.77(1.28)	11.47(2.45)	3.66(0.16)	13.27(1.44)	128.63(29.06)	2.06(0.28)	1.88(0.35)
			$p = 1200$				
proposed	17.60(0.56)	1.80(1.37)	3.49(0.05)	16.70(0.47)	7.43(3.63)	0.52(0.11)	0.67(0.20)
triBayes	15.50(1.48)	4.10(2.28)	2.64(0.08)	0.00(0.00)	0.33(1.83)	0.02(0.13)	4.69(1.45)
glinternet	16.20(1.16)	7.17(3.94)	3.37(0.12)	0.00(0.00)	18.90(4.15)	0.89(0.12)	1.41(0.51)
Lasso	10.83(1.98)	1.13(0.63)	3.26(0.13)	0.00(0.00)	308.93(8.77)	1.15(0.08)	2.64(0.80)
iFORM	14.33(3.20)	42.87(6.41)	3.96(0.19)	0.00(0.00)	39.83(4.28)	1.77(0.17)	2.92(1.52)
HierNet	-	-	-	-	-	-	-
Grace	5.23(1.63)	0.13(0.43)	2.81(0.11)	4.77(1.25)	8.70(7.03)	1.49(0.04)	3.32(0.93)
GEL	18.00(0.00)	11.63(2.20)	3.65(0.12)	13.20(0.55)	132.40(21.54)	2.11(0.29)	1.88(0.52)
			$p = 1400$				
proposed	17.03(3.26)	2.20(1.83)	3.36(0.64)	16.13(3.08)	10.03(5.33)	0.56(0.15)	0.70(0.21)
triBayes	14.23(4.14)	3.70(1.93)	2.44(0.66)	0.00(0.00)	0.00(0.00)	0.00(0.00)	4.37(1.58)
glinternet	15.17(3.21)	5.23(2.50)	3.25(0.62)	0.00(0.00)	16.77(4.43)	0.86(0.20)	1.48(0.52)
Lasso	9.77(3.09)	1.10(0.80)	3.13(0.60)	0.00(0.00)	296.40(56.89)	1.11(0.23)	2.69(0.86)
iFORM	13.00(3.49)	43.77(9.37)	3.92(0.77)	0.00(0.00)	37.00(7.67)	1.71(0.37)	3.10(1.42)
HierNet	-	-	-	-	-	-	-
Grace	5.30(2.22)	0.10(0.31)	2.62(0.72)	4.43(1.72)	8.97(9.22)	1.40(0.38)	3.13(1.15)
GEL	16.80(4.57)	11.20(4.02)	3.44(0.94)	12.17(3.32)	125.20(41.22)	2.02(0.65)	1.94(0.79)

Table S15: Simulation results under the scenarios with $p = 1000, \rho = 0.4, K = 100$, $r = 1/\sqrt{5}$, coefficient type S1, and various sample sizes. In each cell, mean (SD) based on 100 replicates.

Approach	M:TP	M:FP	M:RSSE	I:TP	I:FP	I:RSSE	PMSE
			$n = 200$				
proposed	17.83(0.46)	4.93(1.64)	0.54(0.08)	16.13(1.20)	16.73(4.14)	0.75(0.12)	0.83(0.18)
triBayes	15.43(1.10)	8.73(3.33)	2.01(0.16)	0.27(1.46)	0.13(0.73)	1.59(0.05)	4.28(1.43)
glinternet	10.87(3.67)	2.27(2.12)	1.40(0.18)	6.13(3.50)	1.90(1.94)	1.41(0.12)	2.68(0.70)
Lasso	7.97(1.77)	0.23(0.43)	1.53(0.12)	8.67(2.47)	193.17(9.06)	1.65(0.11)	3.07(0.88)
iFORM	10.17(2.36)	47.43(4.27)	2.44(0.29)	3.50(1.63)	35.87(3.32)	2.02(0.19)	5.59(1.88)
HierNet	8.20(2.28)	0.63(0.72)	1.50(0.12)	3.57(1.79)	4.70(3.15)	1.53(0.08)	3.09(0.63)
Grace	8.83(1.90)	0.27(0.52)	1.51(0.12)	10.43(2.05)	17.03(11.33)	1.29(0.14)	2.58(0.81)
GEL	16.90(3.00)	11.60(5.10)	2.02(0.55)	11.17(2.12)	116.10(39.24)	4.86(1.77)	8.64(6.24)
			$n = 400$				
proposed	17.97(0.18)	1.27(0.98)	0.26(0.07)	17.00(0.00)	7.40(3.79)	0.33(0.07)	0.55(0.14)
triBayes	17.67(0.55)	12.60(2.69)	1.02(0.08)	8.13(1.31)	4.20(1.27)	1.22(0.07)	2.02(0.72)
glinternet	18.00(0.00)	6.30(3.16)	0.71(0.10)	16.57(0.77)	6.07(3.17)	0.78(0.10)	1.14(0.36)
Lasso	15.33(2.34)	0.00(0.00)	1.16(0.16)	15.33(1.35)	27.47(11.18)	0.90(0.12)	1.62(0.57)
iFORM	18.00(0.00)	33.27(2.61)	0.93(0.07)	16.90(0.31)	28.43(2.54)	0.88(0.08)	1.17(0.25)
HierNet	17.57(0.94)	2.13(1.72)	0.88(0.18)	13.37(1.88)	12.73(5.30)	1.15(0.15)	1.75(0.87)
Grace	12.73(4.09)	1.83(3.00)	1.20(0.36)	14.20(1.88)	11.53(14.54)	1.12(0.15)	2.00(0.81)
GEL	18.00(0.00)	8.27(5.09)	0.54(0.09)	13.90(1.30)	101.60(65.30)	1.33(0.31)	1.11(0.30)
			$n = 500$				
proposed	18.00(0.00)	0.43(0.73)	0.21(0.05)	17.00(0.00)	4.30(2.18)	0.24(0.04)	0.52(0.13)
triBayes	17.93(0.25)	15.77(3.22)	0.93(0.08)	8.90(1.03)	4.67(1.18)	1.17(0.05)	1.85(0.41)
glinternet	18.00(0.00)	5.13(3.05)	0.60(0.07)	17.00(0.00)	5.00(2.60)	0.62(0.07)	0.92(0.27)
Lasso	17.67(0.96)	0.10(0.31)	0.86(0.12)	16.87(0.35)	25.97(8.44)	0.64(0.11)	1.05(0.31)
iFORM	18.00(0.00)	33.10(2.81)	0.83(0.06)	17.00(0.00)	28.67(2.87)	0.74(0.05)	1.04(0.25)
HierNet	17.97(0.18)	2.00(1.82)	0.71(0.12)	14.93(1.23)	10.43(4.15)	1.01(0.11)	1.36(0.36)
Grace	16.83(2.74)	4.40(3.24)	0.81(0.28)	16.37(1.50)	27.13(16.19)	0.95(0.15)	1.18(0.53)
GEL	18.00(0.00)	5.90(5.29)	0.44(0.06)	14.23(1.68)	78.33(72.50)	1.01(0.15)	0.95(0.18)

Table S16: Simulation results under the scenarios with $p = 1000, \rho = 0.4, K = 100$, $r = 1/\sqrt{5}$, coefficient type S2, and various sample sizes. In each cell, mean (SD) based on 100 replicates.

Approach	**M:TP**	**M:FP**	**M:RSSE**	**I:TP**	**I:FP**	**I:RSSE**	**PMSE**
			$n = 200$				
proposed	17.37(0.81)	3.97(1.33)	0.57(0.11)	15.17(1.58)	11.27(3.61)	0.84(0.16)	0.95(0.29)
triBayes	15.13(1.50)	4.80(2.51)	1.97(0.03)	0.00(0.00)	0.00(0.00)	1.66(0.00)	4.60(1.21)
glinternet	9.53(3.01)	1.47(1.55)	1.40(0.14)	5.70(2.96)	0.83(1.02)	1.48(0.11)	2.60(0.79)
Lasso	6.07(1.76)	0.53(0.73)	1.55(0.09)	5.50(1.57)	205.40(7.39)	1.84(0.08)	3.53(1.17)
iFORM	8.67(2.07)	50.50(3.33)	2.60(0.30)	3.30(1.73)	34.30(3.49)	2.10(0.16)	5.27(1.41)
HierNet	6.33(2.37)	0.53(0.78)	1.51(0.09)	3.43(1.98)	3.80(2.61)	1.59(0.08)	2.83(0.83)
Grace	5.40(1.67)	0.37(0.56)	1.76(0.07)	4.03(1.65)	21.73(11.38)	1.55(0.07)	3.37(0.91)
GEL	17.47(2.18)	11.23(2.92)	1.87(0.43)	11.53(1.94)	118.03(30.49)	4.66(1.21)	9.16(4.51)
			$n = 400$				
proposed	17.87(0.35)	0.60(0.77)	0.28(0.09)	16.90(0.31)	6.73(3.14)	0.37(0.11)	0.56(0.13)
triBayes	17.03(1.75)	8.43(3.73)	1.19(0.40)	5.70(3.51)	1.17(1.39)	1.40(0.16)	2.47(1.56)
glinternet	17.97(0.18)	5.67(3.59)	0.76(0.08)	16.23(0.94)	4.70(3.31)	0.84(0.12)	1.11(0.29)
Lasso	12.47(3.96)	0.27(0.52)	1.27(0.22)	12.93(2.57)	96.87(156.68)	1.18(0.16)	1.88(0.54)
iFORM	17.97(0.18)	34.17(2.28)	0.96(0.08)	16.63(0.76)	28.43(2.43)	0.90(0.12)	1.08(0.29)
HierNet	17.17(1.26)	2.60(2.01)	0.95(0.17)	13.33(1.52)	12.60(5.98)	1.21(0.14)	1.42(0.41)
Grace	5.30(1.39)	0.10(0.31)	1.75(0.06)	5.27(1.23)	1.63(1.59)	1.48(0.05)	3.51(1.10)
GEL	17.83(0.91)	11.07(3.46)	0.55(0.13)	14.13(1.48)	137.00(41.09)	1.56(0.24)	1.22(0.37)
			$n = 500$				
proposed	17.97(0.18)	0.30(0.65)	0.22(0.06)	16.97(0.18)	6.17(3.85)	0.31(0.07)	0.55(0.13)
triBayes	17.93(0.25)	10.97(2.76)	0.86(0.07)	8.57(0.90)	1.73(1.39)	1.27(0.05)	1.64(0.33)
glinternet	18.00(0.00)	5.47(3.09)	0.64(0.10)	16.80(0.48)	4.77(2.34)	0.68(0.09)	0.89(0.19)
Lasso	17.10(1.49)	0.20(0.41)	0.99(0.17)	16.03(1.16)	26.43(10.47)	0.88(0.12)	1.29(0.20)
iFORM	18.00(0.00)	33.37(2.89)	0.83(0.06)	17.00(0.00)	28.30(2.91)	0.74(0.05)	0.95(0.18)
HierNet	13.13(8.06)	1.33(1.83)	0.58(0.38)	11.17(6.92)	7.20(5.45)	0.79(0.49)	0.85(0.58)
Grace	6.63(2.40)	0.13(0.43)	1.71(0.09)	5.67(1.21)	1.33(2.12)	1.45(0.07)	3.14(0.68)
GEL	18.00(0.00)	8.43(4.51)	0.44(0.08)	14.63(1.79)	106.70(72.23)	1.08(0.16)	0.94(0.22)

S2.5 Additional data analysis results

Table S17: Analysis of the SKCM data using the proposed approach: identified main effects and interactions.

Gene	Main	HK3	HK2	ADH5	DLAT	FBP1	CS	MDH2	GUSB	GMPPA	PMM1	PFKFB4	PMM2	ACSL5
HK3	-0.2589													
HK2	-0.1789													-0.3450
ADH5	-0.1107													
DLAT	0.0996	0.0527												-0.0252
FBP1	0.0180				-0.1857		0.0848	0.2843			-0.2398	0.1033	-0.1339	0.0493
PCK2	-0.1593						0.1013					-0.0556		0.0036
CS	0.0713													-0.0574
MDH2	-0.0767	-0.4624												0.1532
DLST	0.0715					0.0490		0.0500			-0.0126			
GUSB	0.1671	0.2827												
GMPPA	0.1709													
PMM1	0.1908	0.2184	-0.0410						-0.0297					0.0970
PFKFB4	-0.0041		-0.1000	-0.0342						0.0364	-0.0067			0.4628
PMM2	-0.2602	0.1326										0.2020		
GALE	0.0204													-0.2027
ACSL5	0.0919								0.0267	0.0943				

Table S18: Analysis of the LUAD data using the proposed approach: identified main effects and interactions.

Gene	Main	ACSS2	ALDH2	PGAM4	PGM2	PGAM1	HK2	AKR1A1	LDHA	PGK1	ACLY	CS
ACSS2	0.0423											
ALDH2	0.0617	0.0063										
PGAM4	0.0101	-0.0604	0.0733								0.1912	
PGM2	0.0253		0.0527									
PGAM1	-0.0153	0.1354	0.0680									
HK2	-0.0268	0.0120	0.0320	0.2139	-0.0817	-0.1755						
AKR1A1	-0.0009		0.0410	-0.0440		0.0240	0.0408					
LDHA	-0.0426	-0.0978	-0.0151					0.0249			-0.1009	
PGK1	0.0100		-0.0828								-0.0766	
ALDH3A2	-0.0216			0.0486		-0.0730	0.0535	-0.0620	-0.0563	-0.0366	-0.0579	0.0512
ACLY	-0.0616											
CS	-0.0096		-0.0658	-0.1090			-0.0067		-0.0333	0.1001		
DLST	0.0515			-0.0376		0.0526			-0.0481			