Supporting Information for (Structured gene-environment interaction analysis) by (Mengyun Wu, Qingzhao Zhang and Shuangge Ma)

1 Web Appendix A

1.1 Estimation under the AFT model

For subject i, denote T_i as the survival time of interest. Use notations similar to those in the main text. For T_i, consider the accelerated failure time (AFT) model

$$\log(T_i) = \alpha_0 + \sum_{k=1}^{q} Z_{ik}\alpha_k + \sum_{j=1}^{p} X_{ij}\beta_j + \sum_{k=1}^{q}\sum_{j=1}^{p} Z_{ik}X_{ij}\eta_{kj} + \varepsilon_i,$$

where α_0 is the intercept. In practice, right censoring is usually present. Denote C_i as the censoring time for subject i, then we observe $Y_i = \log(\min(T_i, C_i))$ and $\tilde{\delta}_i = I(T_i \leq C_i)$. Assume that data $\{(\boldsymbol{Z}_{i\cdot}, \boldsymbol{X}_{i\cdot}, Y_i, \tilde{\delta}_i), i = 1, \ldots, n\}$ have been sorted according to Y_i's from the smallest to the largest. For estimation, the following weighted least squared loss function is adopted,

$$\frac{1}{2n}\sum_{i=1}^{n} w_i \left[Y_i - \left(\alpha_0 + \sum_{k=1}^{q} Z_{ik}\alpha_k + \sum_{j=1}^{p} X_{ij}\beta_j + \sum_{k=1}^{q}\sum_{j=1}^{p} Z_{ik}X_{ij}\eta_{kj} \right) \right]^2, \tag{1}$$

where w_i's are the Kaplan-Meier weights defined as

$$w_1 = \frac{\tilde{\delta}_1}{n},\ w_i = \frac{\tilde{\delta}_i}{n-i+1}\prod_{l=1}^{i-1}\left(\frac{n-l}{n-l+1}\right)^{\tilde{\delta}_l}, i = 2, \cdots, n.$$

We center Y_i, $\boldsymbol{Z}_{i\cdot}$, $\boldsymbol{X}_{i\cdot}$, and $\boldsymbol{W}_{i\cdot}^{(k)} = (Z_{ik}X_{i1}, \cdots, Z_{ik}X_{ip})$ using their weighted means. Specifically,

$$Y_i = \sqrt{w_i}(Y_i - \overline{Y}),\ \boldsymbol{Z}_{i\cdot} = \sqrt{w_i}(\boldsymbol{Z}_{i\cdot} - \overline{\boldsymbol{Z}}),\ \boldsymbol{X}_{i\cdot} = \sqrt{w_i}(\boldsymbol{X}_{i\cdot} - \overline{\boldsymbol{X}}),\ \boldsymbol{W}_{i\cdot}^{(k)} = \sqrt{w_i}\left(\boldsymbol{W}_{i\cdot}^{(k)} - \overline{\boldsymbol{W}}^{(k)}\right),$$

where $\overline{Y} = \sum_{i=1}^{n} w_i Y_i / \sum_{i=1}^{n} w_i$, $\overline{\boldsymbol{Z}} = \sum_{i=1}^{n} w_i \boldsymbol{Z}_{i\cdot} / \sum_{i=1}^{n} w_i$, $\overline{\boldsymbol{X}} = \sum_{i=1}^{n} w_i \boldsymbol{X}_{i\cdot} / \sum_{i=1}^{n} w_i$, and $\overline{\boldsymbol{W}}^{(k)} = \sum_{i=1}^{n} w_i \boldsymbol{W}_{i\cdot}^{(k)} / \sum_{i=1}^{n} w_i$. Then, loss function (1) can be rewritten as

$$\frac{1}{2n}\left\| \boldsymbol{Y} - \boldsymbol{Z}\boldsymbol{\alpha} - \boldsymbol{X}\boldsymbol{\beta} - \sum_{k=1}^{q} \boldsymbol{W}^{(k)}\boldsymbol{\eta}_k \right\|_2^2.$$

1.2 Computation

Consider the objective function

$$\begin{aligned} Q_n(\boldsymbol{\theta}) &= \frac{1}{2n}\left\|\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{\alpha}-\boldsymbol{X}\boldsymbol{\beta}-\sum_{k=1}^{q}\boldsymbol{W}^{(k)}(\boldsymbol{\beta}\odot\boldsymbol{\gamma}_k)\right\|_2^2+\sum_{j=1}^{p}\rho(|\beta_j|;\lambda_1,r)+\sum_{j=1}^{p}\sum_{k=1}^{q}\rho(|\gamma_{kj}|;\lambda_1,r) \\ &+ \frac{1}{2}\lambda_2\boldsymbol{\beta}'\boldsymbol{J}\boldsymbol{\beta}+\frac{1}{2}\lambda_2\sum_{k=1}^{q}\boldsymbol{\gamma}_k'\boldsymbol{J}\boldsymbol{\gamma}_k. \end{aligned} \tag{2}$$

With fixed tuning parameters, optimization of (2) can be conducted using an iterative coordinate descent (CD) algorithm, which optimizes the objective function with respect to one of the three vectors, $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, and $\boldsymbol{\gamma}$, at a time and iteratively cycles through all parameters until convergence. The proposed algorithm proceeds as follows:

Step 1 Initialize $t=0$, $\boldsymbol{\beta}^{(t)}=\mathbf{0}$, $\boldsymbol{\gamma}^{(t)}=\mathbf{0}$, $\boldsymbol{\alpha}^{(t)}=(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{Y}$, and $\boldsymbol{res}^{(t)}=\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{\alpha}^{(t)}-\boldsymbol{X}\boldsymbol{\beta}^{(t)}-\sum_{k=1}^{q}\boldsymbol{W}^{(k)}\left(\boldsymbol{\beta}^{(t)}\odot\boldsymbol{\gamma}_k^{(t)}\right)$, where $\boldsymbol{\alpha}^{(t)}$, $\boldsymbol{\beta}^{(t)}$, $\boldsymbol{\gamma}^{(t)}$ and $\boldsymbol{res}^{(t)}$ denote the estimates of $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$, $\boldsymbol{\gamma}$ and residual vector at iteration t, respectively.

Step 2 Update $t=t+1$. With $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}$ fixed at $\boldsymbol{\gamma}^{(t-1)}$ and $\boldsymbol{\alpha}^{(t-1)}$, optimize (2) with respect to $\boldsymbol{\beta}$. Let $\tilde{\boldsymbol{Y}}^{(t)}=\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{\alpha}^{(t-1)}$ and $\tilde{\boldsymbol{X}}^{(t)}=\boldsymbol{X}+\sum_{k=1}^{q}\boldsymbol{W}^{(k)}\odot\left(\mathbf{1}_{n\times 1}\left(\boldsymbol{\gamma}_k^{(t-1)}\right)'\right)$ with $\mathbf{1}_{n\times 1}=(1,\cdots,1)_{n\times 1}$. Then

$$\boldsymbol{\beta}^{(t)}=\text{argmin}\frac{1}{2n}\left\|\tilde{\boldsymbol{Y}}^{(t)}-\tilde{\boldsymbol{X}}^{(t)}\boldsymbol{\beta}\right\|_2^2+\sum_{j=1}^{p}\rho(|\beta_j|;\lambda_1,r)+\frac{1}{2}\lambda_2\boldsymbol{\beta}'\boldsymbol{J}\boldsymbol{\beta}. \tag{3}$$

For $j=1,\cdots,p$, carry out the following steps sequentially.

Step 2.1 Compute

$$\boldsymbol{res}_{-j}^{(t)}=\boldsymbol{res}^{(t-1)}+\tilde{\boldsymbol{X}}_j^{(t)}\beta_j^{(t-1)},\ \chi_j^{(t)}=\frac{1}{n}\left(\tilde{\boldsymbol{X}}_j^{(t)}\right)'\tilde{\boldsymbol{X}}_j^{(t)},$$

$$\varphi_j^{(t)}=\frac{1}{n}\left(\tilde{\boldsymbol{X}}_j^{(t)}\right)'\boldsymbol{res}_{-j}^{(t)},\ \Delta_j^{(t)}=\sum_{l=1}^{j-1}\beta_l^{(t)}J_{jl}+\sum_{l=j+1}^{p}\beta_l^{(t-1)}J_{jl}.$$

Step 2.2 Update the estimate of β_j as

$$\beta_j^{(t)} = \begin{cases} \frac{\mathrm{ST}\left(\varphi_j^{(t)} - \lambda_2 \Delta_j^{(t)}, \lambda_1\right)}{\chi_j^{(t)} + \lambda_2 J_{jj} - \frac{1}{r}}, & \left|\varphi_j^{(t)} - \lambda_2 \Delta_j^{(t)}\right| \leq \lambda_1 r \left(\chi_j^{(t)} + \lambda_2 J_{jj}\right) \\ \frac{\varphi_j^{(t)} - \lambda_2 \Delta_j}{\chi_j^{(t)} + \lambda_2 J_{jj}}, & \left|\varphi_j^{(t)} - \lambda_2 \Delta_j^{(t)}\right| > \lambda_1 r \left(\chi_j^{(t)} + \lambda_2 J_{jj}\right) \end{cases}, \tag{4}$$

where $\mathrm{ST}(\nu, \lambda_1) = sgn(\nu)(|\nu| - \lambda_1)_+$ is the soft-thresholding operator.

Step 2.3 Update $\boldsymbol{res}^{(t-1)} = \boldsymbol{res}^{(t-1)} + \tilde{\boldsymbol{X}}_j^{(t)} \beta_j^{(t-1)} - \tilde{\boldsymbol{X}}_j^{(t)} \beta_j^{(t)}$.

Step 3 With $\boldsymbol{\beta}$ and $\boldsymbol{\alpha}$ fixed at $\boldsymbol{\beta}^{(t)}$ and $\boldsymbol{\alpha}^{(t-1)}$, optimize (2) with respect to $\boldsymbol{\gamma}$. Let $\breve{\boldsymbol{Y}}^{(t)} = \boldsymbol{Y} - \boldsymbol{Z}\boldsymbol{\alpha}^{(t-1)} - \boldsymbol{X}\boldsymbol{\beta}^{(t)}$ and $\left(\tilde{\boldsymbol{W}}^{(k)}\right)^{(t)} = \boldsymbol{W}^{(k)} \odot \left(\mathbf{1}_{n\times 1}\left(\boldsymbol{\beta}^{(t)}\right)'\right)$. Then

$$\left(\boldsymbol{\gamma}_1^{(t)}, \cdots, \boldsymbol{\gamma}_q^{(t)}\right) = \mathrm{argmin} \frac{1}{2n} \left\| \breve{\boldsymbol{Y}}^{(t)} - \sum_{k=1}^{q} \left(\tilde{\boldsymbol{W}}^{(k)}\right)^{(t)} \boldsymbol{\gamma}_k \right\|_2^2 + \sum_{k=1}^{q} \sum_{j=1}^{p} \rho(|\gamma_{kj}|; \lambda_1, r) + \frac{1}{2}\lambda_2 \sum_{k=1}^{q} \boldsymbol{\gamma}_k' \boldsymbol{J} \boldsymbol{\gamma}_k.$$

For $k = 1, \cdots, q$ and $j \in \left\{j : \beta_j^{(t)} \neq 0, j = 1, \cdots, p\right\}$, conduct estimation similar to Steps 2.1, 2.2 and 2.3.

Step 4 Compute $\boldsymbol{\alpha}^{(t)} = (\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\left(\boldsymbol{res}^{(t-1)} + \boldsymbol{Z}\boldsymbol{\alpha}^{(t-1)}\right)$ and $\boldsymbol{res}^{(t)} = \boldsymbol{res}^{(t-1)} + \boldsymbol{Z}\boldsymbol{\alpha}^{(t-1)} - \boldsymbol{Z}\boldsymbol{\alpha}^{(t)}$.

Step 5 Repeat Steps 2-4 until convergence. In our numerical study, convergence is concluded if $\frac{|Q_n(\boldsymbol{\theta}^{(t)}) - Q_n(\boldsymbol{\theta}^{(t-1)})|}{|Q_n(\boldsymbol{\theta}^{(t-1)})|} < 10^{-4}$.

1.2.1 Details for Steps 2.1 and 2.2 of the proposed algorithm

Consider objective function (3). For $j = 1, \ldots, p$, the CD algorithm optimizes the objective function with respect to β_j while fixing the other parameters $\beta_l (l \neq j)$ at their current estimates $\beta_l^{(t)}$ for $l < j$ or $\beta_l^{(t-1)}$ for $l > j$. Specifically, consider the following simplified objective function

$$Q_s(\beta_j) = \frac{1}{2n} \left\| \boldsymbol{res}_{-j}^{(t)} - \tilde{\boldsymbol{X}}_j^{(t)} \beta_j \right\|_2^2 + \rho(|\beta_j|; \lambda_1, r) + \frac{1}{2}\lambda_2 \left(J_{jj}\beta_j^2 + 2\sum_{l=1}^{j-1} \beta_l^{(t)} J_{jl} \beta_j + 2 \sum_{l=j+1}^{p} \beta_l^{(t-1)} J_{jl} \beta_j \right), \tag{5}$$

where $\boldsymbol{res}_{-j}^{(t)} = \tilde{\boldsymbol{Y}}^{(t)} - \sum_{l=1}^{j-1} \tilde{\boldsymbol{X}}_l^{(t)} \beta_l^{(t)} - \sum_{l=j+1}^{p} \tilde{\boldsymbol{X}}_l^{(t)} \beta_l^{(t-1)} = \boldsymbol{res}^{(t-1)} + \tilde{\boldsymbol{X}}_j^{(t)} \beta_j^{(t-1)}$ with $\boldsymbol{res}^{(t-1)} = \tilde{\boldsymbol{Y}}^{(t)} - \sum_{l=1}^{j-1} \tilde{\boldsymbol{X}}_l^{(t)} \beta_l^{(t)} - \sum_{l=j}^{p} \tilde{\boldsymbol{X}}_l^{(t)} \beta_l^{(t-1)}$. The first order derivative of (5) is

$$\begin{aligned}
\frac{\partial Q_s(\beta_j)}{\partial \beta_j} &= -\frac{1}{n}\left(\tilde{\boldsymbol{X}}_j^{(t)}\right)' \boldsymbol{res}_{-j}^{(t)} + \frac{1}{n}\left(\tilde{\boldsymbol{X}}_j^{(t)}\right)' \tilde{\boldsymbol{X}}_j^{(t)} \beta_j + \lambda_1 sgn(\beta_j) \begin{cases} 1 - \frac{|\beta_j|}{\lambda_1 r} & |\beta_j| \le \lambda_1 r \\ 0 & |\beta_j| > \lambda_1 r, \end{cases} + \lambda_2 J_{jj} \beta_j + \lambda_2 \Delta_j^{(t)}, \\
&\triangleq -\varphi_j^{(t)} + \chi_j^{(t)} \beta_j + \lambda_1 sgn(\beta_j) \begin{cases} 1 - \frac{|\beta_j|}{\lambda_1 r} & |\beta_j| \le \lambda_1 r \\ 0 & |\beta_j| > \lambda_1 r, \end{cases} + \lambda_2 J_{jj} \beta_j + \lambda_2 \Delta_j^{(t)},
\end{aligned}$$

where

$$\varphi_j^{(t)} = \frac{1}{n}\left(\tilde{\boldsymbol{X}}_j^{(t)}\right)' \boldsymbol{res}_{-j}^{(t)},\ \chi_j^{(t)} = \frac{1}{n}\left(\tilde{\boldsymbol{X}}_j^{(t)}\right)' \tilde{\boldsymbol{X}}_j^{(t)},\ \Delta_j^{(t)} = \sum_{l=1}^{j-1} \beta_l^{(t)} J_{jl} + \sum_{l=j+1}^{p} \beta_l^{(t-1)} J_{jl}.$$

By setting the first order derivative equal to zero, we have

$$\beta_j^{(t)} = \begin{cases} \frac{\mathrm{ST}\left(\varphi_j^{(t)} - \lambda_2 \Delta_j^{(t)}, \lambda_1\right)}{\chi_j^{(t)} + \lambda_2 J_{jj} - \frac{1}{r}} & \left|\varphi_j^{(t)} - \lambda_2 \Delta_j^{(t)}\right| \le \lambda_1 r (\chi_j^{(t)} + \lambda_2 J_{jj}) \\ \frac{\varphi_j^{(t)} - \lambda_2 \Delta_j}{\chi_j^{(t)} + \lambda_2 J_{jj}} & \left|\varphi_j^{(t)} - \lambda_2 \Delta_j^{(t)}\right| > \lambda_1 r (\chi_j^{(t)} + \lambda_2 J_{jj}) \end{cases},$$

where $\mathrm{ST}(\nu, \lambda_1) = sgn(\nu)(|\nu| - \lambda_1)_+$ is the soft-thresholding operator.

1.2.2 Convergence of the proposed algorithm

We now consider convergence of the proposed coordinate descent-based algorithm. A similar algorithm has been adopted in Choi et al. (2010), which has demonstrated satisfactory convergence and numerical performance. Convergence of the CD algorithm for MCP has been studied in Breheny and Huang (2011). Specifically, with $\boldsymbol{\gamma}$ and $\boldsymbol{\alpha}$ fixed at $\boldsymbol{\gamma}^{(t-1)}$ and $\boldsymbol{\alpha}^{(t-1)}$, the objective function with respect to $\boldsymbol{\beta}$ can be rewritten as

$$Q(\boldsymbol{\beta}) = \frac{1}{2n}\left\|\tilde{\boldsymbol{Y}}^{(t)} - \tilde{\boldsymbol{X}}^{(t)}\boldsymbol{\beta}\right\|_2^2 + \sum_{j=1}^{p} \rho(|\beta_j|; \lambda_1, r) + \frac{1}{2}\lambda_2 \boldsymbol{\beta}' \boldsymbol{J} \boldsymbol{\beta}, \tag{6}$$

where $\tilde{\boldsymbol{Y}}^{(t)} = \boldsymbol{Y} - \boldsymbol{Z}\boldsymbol{\alpha}^{(t-1)}$ and $\tilde{\boldsymbol{X}}^{(t)} = \boldsymbol{X} + \sum_{k=1}^{q} \boldsymbol{W}^{(k)} \odot \left(\mathbf{1}_{n\times 1}\left(\boldsymbol{\gamma}_k^{(t-1)}\right)'\right)$ with $\mathbf{1}_{n\times 1} = (1, \cdots, 1)_{n\times 1}$. Then, we have the following proposition.

Proposition 1: Let $c_*(\boldsymbol{\gamma})$ denote the minimum eigenvalue of $\frac{1}{n}\left(\tilde{\boldsymbol{X}}^{(t)}\right)' \tilde{\boldsymbol{X}}^{(t)} + \lambda_2 \boldsymbol{J}$. Then the objective function defined in (3) is a convex function of $\boldsymbol{\beta}$ when $r > 1/c_*(\boldsymbol{\gamma})$.

Proof: The objective function possesses directional derivatives and second directional derivatives at all points $\boldsymbol{\beta}$ and in all directions $\boldsymbol{u}_p$ for $\boldsymbol{\beta}$, although it is not differentiable. Denote $d_{\boldsymbol{u}}Q$ and $d^2_{\boldsymbol{u}}Q$ as the derivative and second derivative of Q in the direction $\boldsymbol{u}$. Then, for all $\boldsymbol{\beta}$, we have

$$\min_{\boldsymbol{u}}\{d^2_{\boldsymbol{u}}Q(\boldsymbol{\beta})\} \geq \frac{1}{n}\left(\tilde{\boldsymbol{X}}^{(t)}\right)'\tilde{\boldsymbol{X}}^{(t)} + \lambda_2\boldsymbol{J} - \frac{1}{r}\boldsymbol{I}.$$

This completes the proof of Proposition 1.

According to Lemma 1, Propositions 1 and 2 in Breheny and Huang (2011), and Theorems 4.1 and 5.1 in Tseng (2001), with the convexity, the value of the objective function decreases at each step and converges to a coordinate-wise minimum when $r > 1/c_*(\boldsymbol{\gamma})$. Furthermore, because all directional derivatives exist, every coordinate-wise minimum is also a local minimum. The same conclusion can be drawn for $\boldsymbol{\gamma}$. Therefore, with a large enough value of r, the proposed algorithm is guaranteed to converge.

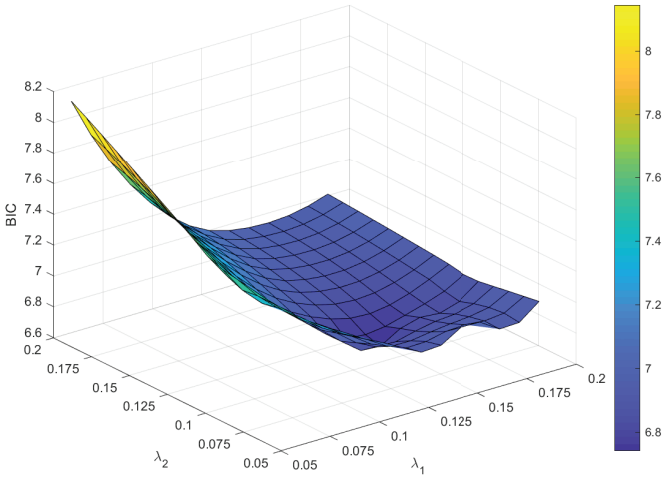
1.2.3 Computational cost

In the proposed algorithm, the most time-consuming are Steps 2 and 3. Here, following Friedman et al. (2010), $\boldsymbol{res}^{(t)}$ is introduced to improve efficiency. Suppose that there are $\tilde{s}_\beta$ nonzero G effects in the model. With the hierarchial constraint, updating for each main G factor or interaction costs at most $O(\max(n, \tilde{s}_\beta))$ operations. Thus, a complete cycle through all $p+q+\tilde{s}_\beta q$ variables costs at most $O(\max(n, \tilde{s}_\beta)(p+q+\tilde{s}_\beta q))$ operations. The space consumption is mainly for storing matrices $\boldsymbol{Z}, \boldsymbol{X}, \boldsymbol{W}^{(1)}, \cdots, \boldsymbol{W}^{(q)}$ and initializing $\boldsymbol{\theta}$, leading to complexity $O((n+1)(p+q+pq))$. For example, consider $q = 5$ and $p = 5000$. With fixed tuning parameters, for a simulated dataset with $n = 250$ and a continuous outcome, the average computer time is 23.15 seconds. For a simulated dataset with $n = 350$ and a censored survival outcome with a censoring rate 20%, the proposed analysis takes about 28.54 seconds.

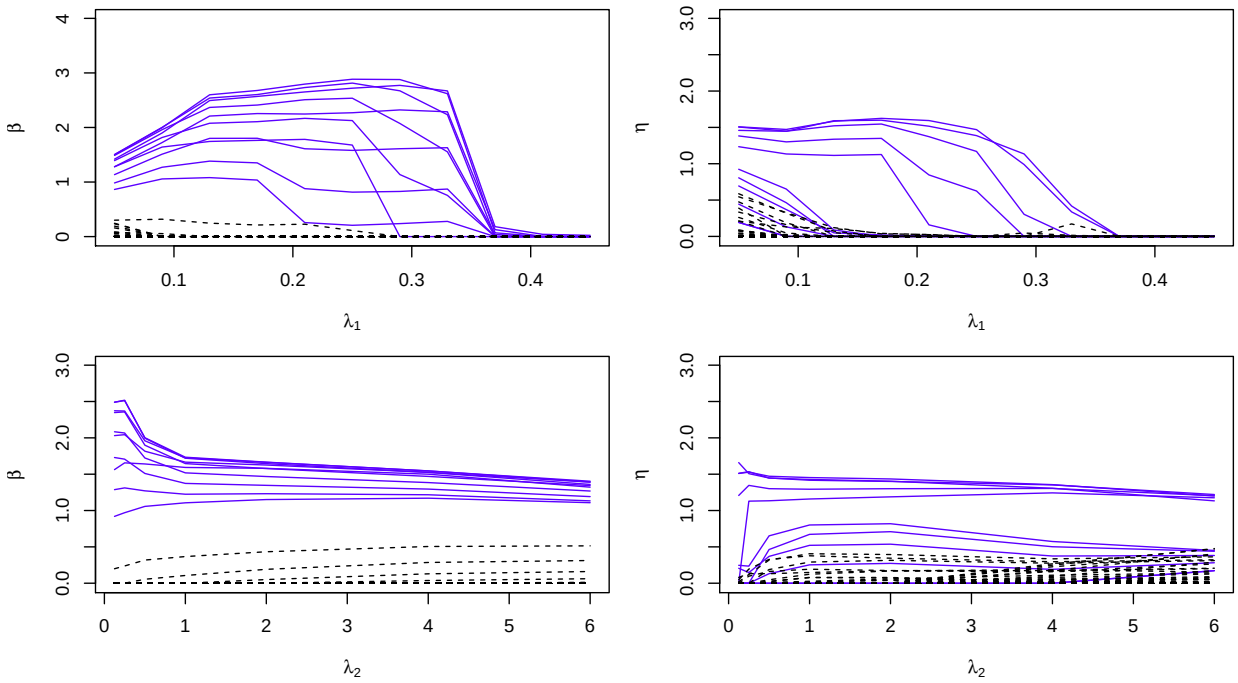
1.2.4 Parameter path

To produce a parameter path, for each value of λ_2 in a sensible range, we compute solutions for a decreasing sequence of values for λ_1, starting at the smallest $\lambda_{1,\max}$ at which all penalized coefficients are zero. With the hierarchial constraint, it suffices to examine $\lambda_{1,\max}$ for $\boldsymbol{\beta}$. As shown in (4), when $\boldsymbol{\beta} = 0$, β_j will stay zero if $\frac{1}{n}\left|\left(\tilde{\boldsymbol{X}}_j\right)' \tilde{\boldsymbol{Y}}^{(t)}\right| < \lambda_1$. With Step 1 and the definitions of $\tilde{\boldsymbol{X}}_j$ and $\tilde{\boldsymbol{Y}}^{(t)}$, $\lambda_{1,\max}$ can be calculated as $\max_j \frac{1}{n}\left|\boldsymbol{X}_j'(\boldsymbol{Y} - \boldsymbol{Z}(\boldsymbol{Z}'\boldsymbol{Z})^{-1}\boldsymbol{Z}'\boldsymbol{Y})\right|$. Following the literature (Friedman et al., 2010), we select a minimum value $\lambda_{1,\min} = 0.001\lambda_{1,\max}$ and construct a decreasing sequence of 50 λ_1-values on the log-scale. The estimates from the previous value of λ_1 are used as warm starts for the next one to speed up computation as well as improve stability.

We simulate one replicate under the linear model with MAF setting M1 and correlation structure AR(0.3). Details on the data settings are described in Section 3 (main text). With the proposed approach, we first examine the values of BIC as a function of λ_1 and λ_2 in Web Figure 1. The optimal point with $(\lambda_1, \lambda_2) = (0.135, 0.095)$ is clearly identified. We further examine the parameter paths in Web Figure 2. The proposed approach is observed to have parameter paths similar to those of other penalized estimations. The model is sparser with larger λ_1 and smoother with larger λ_2. For this simulated dataset, with the optimal tuning parameters, the proposed approach can correctly identify the majority of true positives while having a small number of false positives.



Web Figure 1: Simulation: BIC as a function of λ_1 and λ_2



Web Figure 2: Simulation: parameter paths for one replicate under the linear model with MAF setting M1 and correlation structure AR(0.3). The blue solid lines represent the first ten true positives, and the black dashed lines represent the true negatives.

1.3 Conditions for Theorems 1 and 2

(C1) Components of the residual $\boldsymbol{\varepsilon}$ are i.i.d and sub-Gaussian with noise level σ. That is, for any vector $\boldsymbol{\nu}$ with $||\boldsymbol{\nu}||_2 = 1$ and any constant $\epsilon > 0$, $P(|\boldsymbol{\nu}'\boldsymbol{\varepsilon}| \geq \epsilon) \leq 2\exp\left(-\frac{\epsilon^2}{2\sigma^2}\right)$.

(C2) Let $b_0 = \min\left\{\left\{|\beta_j^0| : j \in \mathcal{A}_1\right\}, \left\{|\gamma_{kj}^0| : j \in \mathcal{A}_2^k, k = 1, \cdots, q\right\}\right\}$. Then, $b_0\sqrt{n/s} \to \infty$.

(C3) Denote $\lambda_{\min}(\boldsymbol{M})$ and $\lambda_{\max}(\boldsymbol{M})$ as the smallest and largest eigenvalues of $\boldsymbol{M}$. Then,

$$\max_{\boldsymbol{\theta}_{\mathcal{A}} \in \mathcal{N}_0} \lambda_{\max}\left(\frac{1}{n}\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}, \boldsymbol{\gamma}_{\mathcal{A}_1})'\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}, \boldsymbol{\gamma}_{\mathcal{A}_1})\right) \leq s\bar{c},$$

$$\min_{\boldsymbol{\theta}_{\mathcal{A}} \in \mathcal{N}_0} \lambda_{\min}\left(\frac{1}{n}\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}, \boldsymbol{\gamma}_{\mathcal{A}_1})'\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}, \boldsymbol{\gamma}_{\mathcal{A}_1}) + \frac{1}{n}\boldsymbol{F}(\boldsymbol{\theta}_{\mathcal{A}}\right) \geq \underline{c},$$

where $\boldsymbol{\gamma}_{\mathcal{A}_1} = \left(\boldsymbol{\gamma}_{1,\mathcal{A}_1}', \cdots, \boldsymbol{\gamma}_{q,\mathcal{A}_1}'\right)'$ with $\gamma_{kj} = 0$, if $j \in \mathcal{A}_1$ but $j \notin \mathcal{A}_2^k$, $\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}, \boldsymbol{\gamma}_{\mathcal{A}_1}) = \left(\boldsymbol{Z}, \boldsymbol{U}(\boldsymbol{\gamma}_{\mathcal{A}_1}), \boldsymbol{V}^{(1)}(\boldsymbol{\beta}_{\mathcal{A}_2^1}), \boldsymbol{V}^{(2)}(\boldsymbol{\beta}_{\mathcal{A}_2^2}), \cdots, \boldsymbol{V}^{(q)}(\boldsymbol{\beta}_{\mathcal{A}_2^q})\right)_{n\times(q+s)}$ with $\boldsymbol{U}(\boldsymbol{\gamma}_{\mathcal{A}_1}) = \boldsymbol{X}_{\mathcal{A}_1} + \sum_{k=1}^q \boldsymbol{W}_{\mathcal{A}_1}^{(k)} \odot \left(\boldsymbol{1}_{n\times 1}\left(\boldsymbol{\gamma}_{k,\mathcal{A}_1}\right)'\right)$ and $\boldsymbol{V}^{(k)}(\boldsymbol{\beta}_{\mathcal{A}_2^k}) = \boldsymbol{W}_{\mathcal{A}_2^k}^{(k)} \odot \left(\boldsymbol{1}_{n\times 1}\left(\boldsymbol{\beta}_{\mathcal{A}_2^k}\right)'\right)$, $\boldsymbol{F}(\boldsymbol{\theta}_{\mathcal{A}}) = \left(f_{jl}(\boldsymbol{\theta}_{\mathcal{A}})\right)_{(q+s)\times(q+s)}$ with $f_{jl}(\boldsymbol{\theta}_{\mathcal{A}}) = -\left(\boldsymbol{W}_{\varsigma}^{(k)}\right)'\left(\boldsymbol{Y} - \boldsymbol{Z}\boldsymbol{\alpha} - \boldsymbol{X}_{\mathcal{A}_1}\boldsymbol{\beta}_{\mathcal{A}_1} - \sum_{g=1}^q \boldsymbol{W}_{\mathcal{A}_2^g}^{(g)}(\boldsymbol{\beta}_{\mathcal{A}_2^g} \odot \boldsymbol{\gamma}_{g,\mathcal{A}_2^g})\right)$ if both j and l correspond to the ςth element of $\mathcal{A}_2^k$, and 0 otherwise, $\mathcal{N}_0 = \{\boldsymbol{\theta}_{\mathcal{A}} : ||\boldsymbol{\theta}_{\mathcal{A}} - \boldsymbol{\theta}_{\mathcal{A}}^0||_\infty \leq \frac{b_0}{2}\}$, and $\bar{c}$ and $\underline{c}$ are two positive constants.

(C4) $\lambda_2 = O(\sqrt{1/n})$.

(C5) $\lambda_{\min}\left(\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\right) \geq 0$ and $||\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\boldsymbol{\theta}_{\mathcal{A}}^0||_2 = O(\sqrt{s})$, where $\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}} = \text{diag}\left(\boldsymbol{0}_{q\times q}, \boldsymbol{J}_{\mathcal{A}_1,\mathcal{A}_1}, \cdots, \boldsymbol{J}_{\mathcal{A}_2^q,\mathcal{A}_2^q}\right)$ is a block diagonal matrix with the diagonal blocks being $\boldsymbol{0}_{q\times q}, \boldsymbol{J}_{\mathcal{A}_1,\mathcal{A}_1}, \cdots$, and $\boldsymbol{J}_{\mathcal{A}_2^q,\mathcal{A}_2^q}$.

(C6) $||\boldsymbol{U}(\boldsymbol{\gamma}_{\mathcal{A}_1^c}^0)'\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}^0, \boldsymbol{\gamma}_{\mathcal{A}_1}^0)||_{2,\infty} = O(n)$, $\left|\left|\boldsymbol{V}^{(k)}\left(\boldsymbol{\beta}_{\left(\tilde{\mathcal{A}}_2^k\right)^c}^0\right)'\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}^0, \boldsymbol{\gamma}_{\mathcal{A}_1}^0)\right|\right|_{2,\infty} = O(n)$, $||\boldsymbol{U}(\boldsymbol{\gamma}_j^0)||_2 = O(\sqrt{n}), ||\boldsymbol{V}^{(k)}(\beta_j^0)||_2 = O(\sqrt{n}), j = 1, \cdots, p$, where for $\boldsymbol{M}$, $||\boldsymbol{M}||_{2,\infty} = \max_{||\boldsymbol{\nu}||_2=1}||\boldsymbol{M}\boldsymbol{\nu}||_\infty$. $\max_{\boldsymbol{\theta}_{\mathcal{A}} \in \mathcal{N}_0}\max_j \lambda_{\max}\left(\boldsymbol{T}_1^{(j)}(\boldsymbol{\gamma}_j)\right) = O(n)$, where $\boldsymbol{T}_1^{(j)}(\boldsymbol{\gamma}_j) = \left(t_{lh}^{(j)}(\boldsymbol{\gamma}_j)\right)_{(q+s)\times(q+s)}$ with $t_{lh}^{(j)}(\boldsymbol{\gamma}_j) = \left(\boldsymbol{X}_j + \sum_{g=1}^q \boldsymbol{W}_j^{(g)}\gamma_{gj}\right)'\boldsymbol{W}_{\varsigma}^{(k)}$ if both l and h correspond to the ςth element of $\mathcal{A}_2^k$, and 0 otherwise. $\max_{\boldsymbol{\theta}_{\mathcal{A}} \in \mathcal{N}_0}\max_j \lambda_{\max}\left(\boldsymbol{T}_2^{(j)}(\beta_j)\right) = O(n)$, where $\boldsymbol{T}_2^{(j)}(\beta_j) = \left(t_{lh}^{(j)}(\beta_j)\right)_{(q+s)\times(q+s)}$ with

$t_{lh}^{(j)}(\beta_j) = \left(\boldsymbol{W}_j^{(k)}\beta_j\right)' \boldsymbol{W}_\varsigma^{(k)}$ if both l and h correspond to the ςth element of $\mathcal{A}_2^k$, and 0 otherwise.

(C7) $\log(p) = n^a, a \in (0, \frac{1}{2})$.

(C8) $\frac{\lambda_1}{\sqrt{s/n}} \to \infty$, $\frac{\lambda_1}{n^{a/2-1/2}\sqrt{\log n}} \to \infty$.

(C9) $b_0\lambda_1^{-1} \to \infty$.

Condition (C1) is the sub-Gaussian condition which is assumed in Fan and Lv (2011), Guo et al. (2016), Huang et al. (2017), and others. Condition (C2) puts a lower bound on the size of the smallest signal, and allows the nonzero effects to vanish asymptotically but at a rate that is not faster than $\sqrt{n/s}$. Condition (C3) assumes that the predictor matrix is "well behaved". Compared to the common predictor matrix-based assumptions, Condition (C3) may be slightly more complicated due to the decomposition strategy for accommodating the hierarchy. If model (1) (Main Text) is applied directly without the decomposition of η_{kj}, Condition (C3) goes back to nearly the same as Condition (A1) in Zou and Zhang (2009), where $\boldsymbol{G} = \left(\boldsymbol{Z}, \boldsymbol{X}_{\mathcal{A}_1}, \boldsymbol{W}_{\mathcal{A}_2^1}^{(1)}, \cdots, \boldsymbol{W}_{\mathcal{A}_2^q}^{(q)}\right)$ and $\boldsymbol{F}(\boldsymbol{\theta}_{\mathcal{A}}) = 0$. Similar conditions have also been assumed in Fan and Lv (2011) and others. Condition (C4) restricts the rate of λ_2. Condition (C5) makes a weak constraint on $\boldsymbol{J}$. It needs to be checked on a case-by-case basis, as $\boldsymbol{J}$ may vary across data. Conditions (C2) and (C5) are not difficult to be satisfied simultaneously. Take the spline type penalty for SNP data as an example. As shown in Main Text, $\boldsymbol{J}$ is symmetric and has $\max|J_{jl}| = 6$ and $|\{j : J_{jl} \neq 0\}| \leq 5, l = 1, \cdots, p$. Denote $b_1 = \max\left\{\left\{|\beta_j^0| : j \in \mathcal{A}_1\right\}, \left\{|\gamma_{kj}^0| : j \in \mathcal{A}_2^k, k = 1, \cdots, q\right\}\right\}$, then we have $||\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\boldsymbol{\theta}_{\mathcal{A}}^0||_2 \leq 30\sqrt{s}b_1$. With Condition (C5), we have $b_1 = O(1)$ which is clearly able to be satisfied together with $b_0\sqrt{n/s} \to \infty$. In addition, if matrix $\boldsymbol{J}$ is an identity matrix, the proposed structure-based penalty goes back to ridge. Two similar conditions have been assumed in Condition (A6) in Zou and Zhang (2009) for adaptive elastic net. For the Laplacian type penalty, it is also satisfied for example when the network is sparse. Condition (C6) is similar to Condition 4 in Fan and Lv (2011), where the

first two equations control the "correlations" between the unimportant variables (those in $\mathcal{A}_1^c$ and $\left(\tilde{\mathcal{A}}_2^k\right)^c$) and important variables (those in $\mathcal{A}_1$ and $\mathcal{A}_2$). Condition (C7) allows the number of G factors to increase as the sample size increases. Condition (C8) has also been assumed in Fan and Lv (2011) and others. Condition (C9) provides the rate at which the nonzero coefficients can be distinguished from zero (Huang et al., 2017).

In the objective function, λ_1 controls the sparsity of coefficients. Both Conditions (C7) and (C8) concern the order of λ_1, indicating that it is related to $\log(p)$. For the minimum signal b_0, with additional Conditions (C2) and (C9), we also have that b_0 depends on $\log(p)$. Similar assumptions have been made in Condition 5 in Fan and Lv (2011). λ_2 plays a different role, and controls the smoothness of coefficients based on the structure of G factors, with a larger value leading to smoother estimates. Condition (C4) is partly motivated by those in Huang et al. (2011), Guo et al. (2016), and others. Guo et al. (2016) suggests that the optimal level of λ_2 can be obtained with assumptions on the smoothness of coefficients. We note that the order of $\sqrt{1/n}$ is slightly stronger than those in Huang et al. (2011) and Guo et al. (2016). However, as interaction analysis with hierarchy is more challenging than the aforementioned main effect analysis, we assume $O(\sqrt{1/n})$ to facilitate a less complicated proof, and the rigorous proof in the following sections can demonstrate its rationality.

1.4 Proof of Theorem 1

To prove Theorem 1, it suffices to show that under conditions (C1)-(C5), for a given ξ,

$$P\left\{\inf_{\boldsymbol{\theta}_{\mathcal{A}}\in\mathcal{N}_1}\tilde{Q}_n(\boldsymbol{\theta}_{\mathcal{A}})>\tilde{Q}_n(\boldsymbol{\theta}^0_{\mathcal{A}})\right\}\geq 1-\xi,$$

where $\mathcal{N}_1=\{\boldsymbol{\theta}_{\mathcal{A}}:||\boldsymbol{\theta}_{\mathcal{A}}-\boldsymbol{\theta}^0_{\mathcal{A}}||_2=\delta_n\}$.

Let $\boldsymbol{w}=\left(\boldsymbol{g}'_{q\times 1},\boldsymbol{u}'_{|\mathcal{A}_1|\times 1},\boldsymbol{v}'_{1_{|\mathcal{A}_2^1|\times 1}},\cdots,\boldsymbol{v}'_{q_{|\mathcal{A}_2^q|\times 1}}\right)'$ with $||\boldsymbol{w}||_2=1$ and $\boldsymbol{\theta}_{\mathcal{A}}=\boldsymbol{\theta}^0_{\mathcal{A}}+\delta_n\boldsymbol{w}$. Let $L_n(\boldsymbol{\theta}_{\mathcal{A}})=\left\|\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{\alpha}-\boldsymbol{X}_{\mathcal{A}_1}\boldsymbol{\beta}_{\mathcal{A}_1}-\sum_{k=1}^q\boldsymbol{W}^{(k)}_{\mathcal{A}_2^k}(\boldsymbol{\beta}_{\mathcal{A}_2^k}\odot\boldsymbol{\gamma}_{k,\mathcal{A}_2^k})\right\|_2^2$, then

$$\begin{aligned}
D_n(\boldsymbol{w}) &= \tilde{Q}_n(\boldsymbol{\theta}^0_{\mathcal{A}}+\delta_n\boldsymbol{w})-\tilde{Q}_n(\boldsymbol{\theta}^0_{\mathcal{A}})\\
&= \frac{1}{2n}L_n(\boldsymbol{\theta}^0_{\mathcal{A}}+\delta_n\boldsymbol{w})-\frac{1}{2n}L_n(\boldsymbol{\theta}^0_{\mathcal{A}})\\
&+ \frac{1}{2}\lambda_2(\boldsymbol{\beta}^0_{\mathcal{A}_1}+\delta_n\boldsymbol{u})'\boldsymbol{J}_{\mathcal{A}_1,\mathcal{A}_1}(\boldsymbol{\beta}^0_{\mathcal{A}_1}+\delta_n\boldsymbol{u})-\frac{1}{2}\lambda_2(\boldsymbol{\beta}^0_{\mathcal{A}_1})'\boldsymbol{J}_{\mathcal{A}_1,\mathcal{A}_1}\boldsymbol{\beta}^0_{\mathcal{A}_1}\\
&+ \frac{1}{2}\lambda_2\sum_{k=1}^q(\boldsymbol{\gamma}^0_{\mathcal{A}_2^k}+\delta_n\boldsymbol{v}_k)'\boldsymbol{J}_{\mathcal{A}_2^k,\mathcal{A}_2^k}(\boldsymbol{\gamma}^0_{\mathcal{A}_2^k}+\delta_n\boldsymbol{v}_k)-\frac{1}{2}\lambda_2\sum_{k=1}^q(\boldsymbol{\gamma}^0_{\mathcal{A}_2^k})'\boldsymbol{J}_{\mathcal{A}_2^k,\mathcal{A}_2^k}\boldsymbol{\gamma}^0_{\mathcal{A}_2^k}.
\end{aligned}$$

We have

$$\begin{aligned}
I &\triangleq \frac{1}{2n}L_n(\boldsymbol{\theta}^0_{\mathcal{A}}+\delta_n\boldsymbol{w})-\frac{1}{2n}L_n(\boldsymbol{\theta}^0_{\mathcal{A}})\\
&= \frac{1}{2n}\delta_n\boldsymbol{w}'(\nabla L_n(\boldsymbol{\theta}_{\mathcal{A}})|_{\boldsymbol{\theta}^0_{\mathcal{A}}})+\frac{1}{4n}\delta_n^2\boldsymbol{w}'(\nabla^2 L_n(\boldsymbol{\theta}_{\mathcal{A}})\big|_{\tilde{\boldsymbol{\theta}}_{\mathcal{A}}})\boldsymbol{w}\\
&= \delta_n\boldsymbol{w}'\left[-\frac{1}{n}\boldsymbol{G}(\boldsymbol{\beta}^0_{\mathcal{A}_2},\boldsymbol{\gamma}^0_{\mathcal{A}_1})'\boldsymbol{\varepsilon}\right]\\
&+ \frac{1}{2}\delta_n^2\boldsymbol{w}'\left[\frac{1}{n}\boldsymbol{G}(\tilde{\boldsymbol{\beta}}_{\mathcal{A}_2},\tilde{\boldsymbol{\gamma}}_{\mathcal{A}_1})'\boldsymbol{G}(\tilde{\boldsymbol{\beta}}_{\mathcal{A}_2},\tilde{\boldsymbol{\gamma}}_{\mathcal{A}_1})+\frac{1}{n}\boldsymbol{F}(\tilde{\boldsymbol{\theta}}_{\mathcal{A}})\right]\boldsymbol{w}\\
&\triangleq I_1+I_2,
\end{aligned}$$

where $\boldsymbol{\varepsilon}=\boldsymbol{Y}-\boldsymbol{Z}\boldsymbol{\alpha}^0-\boldsymbol{X}_{\mathcal{A}_1}\boldsymbol{\beta}^0_{\mathcal{A}_1}-\sum_{k=1}^q\boldsymbol{W}^{(k)}_{\mathcal{A}_2^k}(\boldsymbol{\beta}^0_{\mathcal{A}_2^k}\odot\boldsymbol{\gamma}^0_{k,\mathcal{A}_2^k})$, $\boldsymbol{\gamma}_{\mathcal{A}_1}=\left(\boldsymbol{\gamma}'_{1,\mathcal{A}_1},\cdots,\boldsymbol{\gamma}'_{q,\mathcal{A}_1}\right)'$ with $\gamma_{kj}=0$, if $j\in\mathcal{A}_1$ but $j\notin\mathcal{A}_2^k$,

$$\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2},\boldsymbol{\gamma}_{\mathcal{A}_1})=\left(\boldsymbol{Z},\boldsymbol{U}(\boldsymbol{\gamma}_{\mathcal{A}_1}),\boldsymbol{V}^{(1)}(\boldsymbol{\beta}_{\mathcal{A}_2^1}),\boldsymbol{V}^{(2)}(\boldsymbol{\beta}_{\mathcal{A}_2^2}),\cdots,\boldsymbol{V}^{(q)}(\boldsymbol{\beta}_{\mathcal{A}_2^q})\right)_{n\times(q+s)},$$

with

$$\boldsymbol{U}(\boldsymbol{\gamma}_{\mathcal{A}_1})=\boldsymbol{X}_{\mathcal{A}_1}+\sum_{k=1}^q\boldsymbol{W}^{(k)}_{\mathcal{A}_1}\odot\left(\boldsymbol{1}_{n\times 1}\left(\boldsymbol{\gamma}_{k,\mathcal{A}_1}\right)'\right),\ \boldsymbol{V}^{(k)}(\boldsymbol{\beta}_{\mathcal{A}_2^k})=\boldsymbol{W}^{(k)}_{\mathcal{A}_2^k}\odot\left(\boldsymbol{1}_{n\times 1}\left(\boldsymbol{\beta}_{\mathcal{A}_2^k}\right)'\right),$$

$\boldsymbol{F}(\boldsymbol{\theta}_{\mathcal{A}}) = (f_{jl}(\boldsymbol{\theta}_{\mathcal{A}}))_{(q+s)\times(q+s)}$ with $f_{jl}(\boldsymbol{\theta}_{\mathcal{A}}) = -\left(\boldsymbol{W}_{\varsigma}^{(k)}\right)'(\boldsymbol{Y} - \boldsymbol{Z}\boldsymbol{\alpha} - \boldsymbol{X}_{\mathcal{A}_1}\boldsymbol{\beta}_{\mathcal{A}_1} - \sum_{g=1}^{q}\boldsymbol{W}_{\mathcal{A}_2^g}^{(g)}(\boldsymbol{\beta}_{\mathcal{A}_2^g} \odot \boldsymbol{\gamma}_{g,\mathcal{A}_2^g}))$ if both j and l correspond to the ςth element of $\mathcal{A}_2^k$, and 0 otherwise, and $\tilde{\boldsymbol{\theta}}_{\mathcal{A}}$ lies on the line segment connecting $\boldsymbol{\theta}_{\mathcal{A}}$ and $\boldsymbol{\theta}_{\mathcal{A}}^0$. Moreover,

$$\begin{aligned} II &\triangleq \frac{1}{2}\lambda_2(\boldsymbol{\beta}_{\mathcal{A}_1}^0 + \delta_n\boldsymbol{u})'\boldsymbol{J}_{\mathcal{A}_1,\mathcal{A}_1}(\boldsymbol{\beta}_{\mathcal{A}_1}^0 + \delta_n\boldsymbol{u}) - \frac{1}{2}\lambda_2(\boldsymbol{\beta}_{\mathcal{A}_1}^0)'\boldsymbol{J}_{\mathcal{A}_1,\mathcal{A}_1}\boldsymbol{\beta}_{\mathcal{A}_1}^0 \\ &+ \frac{1}{2}\lambda_2\sum_{k=1}^{q}(\boldsymbol{\gamma}_{\mathcal{A}_2^k}^0 + \delta_n\boldsymbol{v}_k)'\boldsymbol{J}_{\mathcal{A}_2^k,\mathcal{A}_2^k}(\boldsymbol{\gamma}_{\mathcal{A}_2^k}^0 + \delta_n\boldsymbol{v}_k) - \frac{1}{2}\lambda_2\sum_{k=1}^{q}(\boldsymbol{\gamma}_{\mathcal{A}_2^k}^0)'\boldsymbol{J}_{\mathcal{A}_2^k,\mathcal{A}_2^k}\boldsymbol{\gamma}_{\mathcal{A}_2^k}^0 \\ &= \delta_n\lambda_2\boldsymbol{w}'\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\boldsymbol{\theta}_{\mathcal{A}}^0 + \frac{1}{2}\delta_n^2\lambda_2\boldsymbol{w}'\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\boldsymbol{w} \\ &\geq -\delta_n\lambda_2||\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\boldsymbol{\theta}_{\mathcal{A}}^0||_2, \end{aligned}$$

where $\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}} = \text{diag}\left(\boldsymbol{0}_{q\times q}, \boldsymbol{J}_{\mathcal{A}_1,\mathcal{A}_1}, \cdots, \boldsymbol{J}_{\mathcal{A}_2^q,\mathcal{A}_2^q}\right)$ is a block diagonal matrix with the diagonal blocks being $\boldsymbol{0}_{q\times q}, \boldsymbol{J}_{\mathcal{A}_1,\mathcal{A}_1}, \cdots,$ and $\boldsymbol{J}_{\mathcal{A}_2^q,\mathcal{A}_2^q}$, and $\frac{1}{2}\delta_n^2\lambda_2\boldsymbol{w}'\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\boldsymbol{w} \geq \frac{1}{2}\delta_n^2\lambda_2\lambda_{\min}\left(\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\right) \geq 0$ with condition (C5).

With $\delta_n = \frac{4\lambda_2||\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\boldsymbol{\theta}_{\mathcal{A}}^0||_2}{\underline{c}} + E\sqrt{s/n}$, and Conditions (C2), (C4) and (C5), we have

$$||\tilde{\boldsymbol{\theta}}_{\mathcal{A}} - \boldsymbol{\theta}_{\mathcal{A}}^0||_\infty \leq ||\boldsymbol{\theta}_{\mathcal{A}} - \boldsymbol{\theta}_{\mathcal{A}}^0||_\infty \leq \delta_n < b_0/2.$$

Then, with Condition (C3), we have

$$I_2 \geq \frac{1}{2}\delta_n^2\underline{c} > 0.$$

For I_1, with Conditions (C1) and (C3), we have

$$\begin{aligned} & P\left(\delta_n\boldsymbol{w}'\left[-\frac{1}{n}\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}^0, \boldsymbol{\gamma}_{\mathcal{A}_1}^0)'\boldsymbol{\varepsilon}\right] \leq -\delta_n\epsilon\right) \\ = & P\left(\frac{\boldsymbol{w}'\left[-\frac{1}{n}\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}^0, \boldsymbol{\gamma}_{\mathcal{A}_1}^0)'\boldsymbol{\varepsilon}\right]}{\left\|\boldsymbol{w}'\left[-\frac{1}{n}\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}^0, \boldsymbol{\gamma}_{\mathcal{A}_1}^0)'\right]\right\|_2} \leq -\frac{\epsilon}{\left\|\boldsymbol{w}'\left[-\frac{1}{n}\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}^0, \boldsymbol{\gamma}_{\mathcal{A}_1}^0)'\right]\right\|_2}\right) \\ \leq & \exp\left(-\frac{n\epsilon^2}{2\sigma^2\bar{c}s}\right). \end{aligned}$$

Setting $\epsilon = \frac{1}{4}\underline{c}\delta_n$, we have

$$P\left(\delta_n\boldsymbol{w}'\left[-\frac{1}{n}\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}^0, \boldsymbol{\gamma}_{\mathcal{A}_1}^0)'\boldsymbol{\varepsilon}\right] \geq -\frac{1}{4}\underline{c}\delta_n^2\right) \geq 1 - \exp\left(-\frac{n\underline{c}^2\delta_n^2}{32\sigma^2\bar{c}s}\right).$$

Thus, with $\delta_n = \frac{4\lambda_2 ||\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\boldsymbol{\theta}^0_{\mathcal{A}}||_2}{\underline{c}} + E\sqrt{s/n}$, we have

$$\begin{aligned}
P\left\{\inf_{\hat{\theta}\in\mathcal{N}_1} Q_n(\hat{\boldsymbol{\theta}}) > Q_n(\boldsymbol{\theta}^0)\right\} &\geq P\left\{D_n(\boldsymbol{w}) > 0\right\} \\
&\geq P\left\{\delta_n \boldsymbol{w}'\left[-\frac{1}{n}\boldsymbol{G}(\boldsymbol{\beta}^0_{\mathcal{A}_2}, \boldsymbol{\gamma}^0_{\mathcal{A}_1})'\boldsymbol{\varepsilon}\right] + \frac{1}{2}\delta_n^2\underline{c} - \delta_n\lambda_2||\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\boldsymbol{\theta}^0_{\mathcal{A}}||_2 > 0\right\} \\
&\geq P\left(\delta_n \boldsymbol{w}'\left[-\frac{1}{n}\boldsymbol{G}(\boldsymbol{\beta}^0_{\mathcal{A}_2}, \boldsymbol{\gamma}^0_{\mathcal{A}_1})'\boldsymbol{\varepsilon}\right] \geq -\frac{1}{4}\underline{c}\delta_n^2\right) \\
&\geq 1 - \exp\left(-\frac{n\underline{c}^2\delta_n^2}{32\sigma^2\bar{c}s}\right) \\
&= 1 - \exp\left(-\frac{\left[4\sqrt{n/s}\lambda_2||\widetilde{\boldsymbol{J}}_{\mathcal{A},\mathcal{A}}\boldsymbol{\theta}^0_{\mathcal{A}}||_2 + E\underline{c}\right]^2}{32\sigma^2\bar{c}}\right).
\end{aligned}$$

This completes the proof of Theorem 1.

1.5 Proof of Theorem 2

First, consider $\hat{\boldsymbol{\beta}}_{\mathcal{A}_1^c}$. Following Theorem 1 in Fan and Lv (2011), with Condition (C9) and Theorem 1, it suffices to check condition (8) in Fan and Lv (2011). Let

$$h_1 = (n\lambda_1)^{-1}\left[\frac{1}{2}\,\nabla_{\boldsymbol{\beta}_{\mathcal{A}_1^c}} L_n(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}} + \lambda_2 n \boldsymbol{J}_{\mathcal{A}_1^c\cdot}\hat{\boldsymbol{\beta}}\right].$$

Since $\hat{\boldsymbol{\beta}}_{\mathcal{A}_1^c} = 0$, with a Taylor expansion, we have

$$\begin{aligned} h_1 &= (n\lambda_1)^{-1}\left[-\boldsymbol{U}(\boldsymbol{\gamma}_{\mathcal{A}_1^c})'\left(\boldsymbol{Y} - \boldsymbol{Z}\hat{\boldsymbol{\alpha}} - \boldsymbol{X}\hat{\boldsymbol{\beta}} - \sum_{k=1}^{q}\boldsymbol{W}^{(k)}(\hat{\boldsymbol{\beta}}\odot\hat{\boldsymbol{\gamma}}_k)\right) + \lambda_2 n \boldsymbol{J}_{\mathcal{A}_1^c,\mathcal{A}_1}\hat{\boldsymbol{\beta}}_{\mathcal{A}_1}\right] \\ &= (n\lambda_1)^{-1}\left[-\boldsymbol{U}(\boldsymbol{\gamma}^0_{\mathcal{A}_1^c})'\boldsymbol{\varepsilon} + \boldsymbol{U}(\boldsymbol{\gamma}^0_{\mathcal{A}_1^c})'\boldsymbol{G}(\boldsymbol{\beta}^0_{\mathcal{A}_2}, \boldsymbol{\gamma}^0_{\mathcal{A}_1})'(\hat{\boldsymbol{\theta}}_{\mathcal{A}} - \boldsymbol{\theta}^0_{\mathcal{A}}) + \boldsymbol{\kappa} + \lambda_2 n \boldsymbol{J}_{\mathcal{A}_1^c,\mathcal{A}_1}\hat{\boldsymbol{\beta}}_{\mathcal{A}_1}\right] \\ &= (n\lambda_1)^{-1}\left[-\boldsymbol{U}(\boldsymbol{\gamma}^0_{\mathcal{A}_1^c})'\boldsymbol{\varepsilon} + III + \lambda_2 n \boldsymbol{J}_{\mathcal{A}_1^c,\mathcal{A}_1}\hat{\boldsymbol{\beta}}_{\mathcal{A}_1}\right]. \end{aligned}$$

For III, let $m_j(\boldsymbol{\theta}_{\mathcal{A}}) = \left(\boldsymbol{X}_j + \sum_{k=1}^{q}\boldsymbol{W}_j^{(k)}\gamma_{kj}\right)'\left(\boldsymbol{Z}\boldsymbol{\alpha} + \boldsymbol{X}_{\mathcal{A}_1}\boldsymbol{\beta}_{\mathcal{A}_1} + \sum_{k=1}^{q}\boldsymbol{W}^{(k)}_{\mathcal{A}_2^k}(\boldsymbol{\beta}_{\mathcal{A}_2^k}\odot\boldsymbol{\gamma}_{k,\mathcal{A}_2^k})\right)$. Then $\boldsymbol{\kappa} = (\kappa_j, j\in\mathcal{A}_1^c)'$ with

$$\begin{aligned} \kappa_j &= \frac{1}{2}(\hat{\boldsymbol{\theta}}_{\mathcal{A}} - \boldsymbol{\theta}^0_{\mathcal{A}})\left(\nabla^2_{\theta_{\mathcal{A}}} m_j(\boldsymbol{\theta}_{\mathcal{A}})\big|_{\tilde{\boldsymbol{\theta}}_{\mathcal{A}}}\right)(\hat{\boldsymbol{\theta}}_{\mathcal{A}} - \boldsymbol{\theta}^0_{\mathcal{A}}), \\ &\le \max_j \frac{1}{2}\lambda_{\max}\left(\boldsymbol{T}_1^{(j)}(\tilde{\boldsymbol{\gamma}}_j)\right)||\boldsymbol{\theta}^*_{\mathcal{A}} - \boldsymbol{\theta}^0_{\mathcal{A}}||_2, \end{aligned}$$

where $\tilde{\boldsymbol{\theta}}_{\mathcal{A}}$ lies on the line segment connecting $\theta^*_{\mathcal{A}}$ and $\theta^0_{\mathcal{A}}$. Here $\boldsymbol{T}_1^{(j)}(\boldsymbol{\gamma}_j) = \left(t_{lh}^{(j)}(\boldsymbol{\gamma}_j)\right)_{(q+s)\times(q+s)}$ with $t_{lh}^{(j)}(\boldsymbol{\gamma}_j) = \left(\boldsymbol{X}_j + \sum_{g=1}^{q}\boldsymbol{W}_j^{(g)}\gamma_{gj}\right)'\boldsymbol{W}_{\varsigma}^{(k)}$, if both l and h correspond to the ςth element of $\mathcal{A}_2^k$, and 0 otherwise. Consider the event

$$\Omega_1 = \left\{||\boldsymbol{U}(\boldsymbol{\gamma}^0_{\mathcal{A}_1^c})'\boldsymbol{\varepsilon}||_\infty \le \zeta_n\sqrt{n}\right\},$$

with $\zeta_n = n^a(\log(n))^{1/2}$. With Conditions (C6) and (C7), we have

$$\begin{aligned} P(\Omega_1) &= 1 - P\left\{||\boldsymbol{U}(\boldsymbol{\gamma}^0_{\mathcal{A}_1^c})'\boldsymbol{\varepsilon}||_\infty > \zeta_n\sqrt{n}\right\} \\ &\ge 1 - \sum_{j\in\mathcal{A}_1^c} P\left\{||\boldsymbol{U}(\boldsymbol{\gamma}^0_j)'\boldsymbol{\varepsilon}|| > \zeta_n\sqrt{n}\right\} \\ &\ge 1 - 2(p-s_0)\exp\left(-\frac{\zeta_n^2 n}{2\sigma^2\max_{j\in\mathcal{A}_1^c}||\boldsymbol{U}(\boldsymbol{\gamma}^0_j)||_2^2}\right) \end{aligned}$$

$$\geq \quad 1 - 2p \exp\left(-\frac{\zeta_n^2 n}{2\sigma^2 \max_{j \in \mathcal{A}_1^c} ||\boldsymbol{U}(\boldsymbol{\gamma}_j^0)||_2^2} \right) \to 1,$$

as $\log(p) = O(n^a)$ and $||\boldsymbol{U}(\boldsymbol{\gamma}_j^0)||_2 = O(\sqrt{n})$. Thus, with probability approaching 1,

$$||\boldsymbol{U}(\boldsymbol{\gamma}_{\mathcal{A}_1^c}^0)'\boldsymbol{\varepsilon}||_\infty = O(n^{a/2+1/2}\sqrt{\log n}).$$

Then, Condition (C8) gives

$$(n\lambda_1)^{-1}||\boldsymbol{U}(\boldsymbol{\gamma}_{\mathcal{A}_1^c}^0)'\boldsymbol{\varepsilon}||_\infty = o(1).$$

For III, with Conditions (C6) and (C8),

$$\begin{aligned}
(n\lambda_1)^{-1}||III||_\infty &= (n\lambda_1)^{-1}\left[||\boldsymbol{U}(\boldsymbol{\gamma}_{\mathcal{A}_1^c}^0)'\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}^0, \boldsymbol{\gamma}_{\mathcal{A}_1}^0)'(\hat{\boldsymbol{\theta}}_{\mathcal{A}} - \boldsymbol{\theta}_{\mathcal{A}}^0)||_\infty + ||\boldsymbol{\kappa}||_\infty\right] \\
&= (n\lambda_1)^{-1}\left[O(n)||\boldsymbol{\theta}_{\mathcal{A}}^* - \boldsymbol{\theta}_{\mathcal{A}}^0||_2 + O(n)||\boldsymbol{\theta}_{\mathcal{A}}^* - \boldsymbol{\theta}_{\mathcal{A}}^0||_2^2\right] \\
&= O(\lambda_1^{-1}\sqrt{s/n}) = o(1).
\end{aligned}$$

With Conditions (C4), (C5) and (C8),

$$\begin{aligned}
(n\lambda_1)^{-1}||\lambda_2 n\boldsymbol{J}_{\mathcal{A}_1^c,\mathcal{A}_1}\hat{\boldsymbol{\beta}}_{\mathcal{A}_1}||_\infty &= (\lambda_1)^{-1}||\lambda_2\boldsymbol{J}_{\mathcal{A}_1^c,\mathcal{A}_1}\boldsymbol{\beta}_{\mathcal{A}_1}^0 - \lambda_2\boldsymbol{J}_{\mathcal{A}_1^c,\mathcal{A}_1}(\hat{\boldsymbol{\beta}}_{\mathcal{A}_1} - \boldsymbol{\beta}_{\mathcal{A}_1}^0)||_\infty \\
&\leq (\lambda_1)^{-1}||\lambda_2\boldsymbol{J}_{\mathcal{A}_1^c,\mathcal{A}_1}\boldsymbol{\beta}_{\mathcal{A}_1}^0||_\infty + (\lambda_1)^{-1}||\lambda_2\boldsymbol{J}_{\mathcal{A}_1^c,\mathcal{A}_1}(\hat{\boldsymbol{\beta}}_{\mathcal{A}_1} - \boldsymbol{\beta}_{\mathcal{A}_1}^0)||_\infty \\
&= O(\lambda_1^{-1}\sqrt{s/n}) = o(1).
\end{aligned}$$

Next, consider $\hat{\boldsymbol{\gamma}}_{k,(\tilde{\mathcal{A}}_2^k)^c}$. A similar process is adopted to check condition (8) in Fan and Lv (2011). Let

$$h_2 = (n\lambda_1)^{-1}\left[\frac{1}{2}\nabla_{(\tilde{\mathcal{A}}_2^k)^c}L_n(\boldsymbol{\theta})\Big|_{\hat{\boldsymbol{\theta}}} + \lambda_2 n\boldsymbol{J}_{(\tilde{\mathcal{A}}_2^k)^c\cdot}\hat{\boldsymbol{\gamma}}_k\right].$$

Since $\hat{\boldsymbol{\gamma}}_{(\tilde{\mathcal{A}}_2^k)^c} = 0$ and $\hat{\boldsymbol{\beta}}_{(\tilde{\mathcal{A}}_2^k)^c} \neq 0$, with a Taylor expansion, we have

$$\begin{aligned}
h_2 &= (n\lambda_1)^{-1}\left[-\boldsymbol{V}^{(k)}(\boldsymbol{\beta}_{(\tilde{\mathcal{A}}_2^k)^c})'\left(\boldsymbol{Y} - \boldsymbol{Z}\hat{\boldsymbol{\alpha}} - \boldsymbol{X}\hat{\boldsymbol{\beta}} - \sum_{k=1}^{q}\boldsymbol{W}^{(k)}(\hat{\boldsymbol{\beta}} \odot \hat{\boldsymbol{\gamma}}_k)\right) + \lambda_2 n\boldsymbol{J}_{(\tilde{\mathcal{A}}_2^k)^c\cdot}\hat{\boldsymbol{\gamma}}_k\right] \\
&= (n\lambda_1)^{-1}\left[-\boldsymbol{V}^{(k)}(\boldsymbol{\beta}_{(\tilde{\mathcal{A}}_2^k)^c}^0)'\boldsymbol{\varepsilon} + \boldsymbol{V}^{(k)}(\boldsymbol{\beta}_{(\tilde{\mathcal{A}}_2^k)^c}^0)'\boldsymbol{G}(\boldsymbol{\beta}_{\mathcal{A}_2}^0, \boldsymbol{\gamma}_{\mathcal{A}_1}^0)'(\hat{\boldsymbol{\theta}}_{\mathcal{A}} - \boldsymbol{\theta}_{\mathcal{A}}^0) + \tilde{\boldsymbol{\kappa}} + \lambda_2 n\boldsymbol{J}_{(\tilde{\mathcal{A}}_2^k)^c,\cdot}\hat{\boldsymbol{\gamma}}_k\right] \\
&= (n\lambda_1)^{-1}\left[-\boldsymbol{V}^{(k)}(\beta_{(\tilde{\mathcal{A}}_2^k)^c}^0)'\boldsymbol{\varepsilon} + IV + \lambda_2 n\boldsymbol{J}_{(\tilde{\mathcal{A}}_2^k)^c,\mathcal{A}_2^k}\hat{\boldsymbol{\gamma}}_{k,\mathcal{A}_2^k}\right].
\end{aligned}$$

For IV, let $\tilde{m}_j(\boldsymbol{\theta}_{\mathcal{A}}) = \left(\boldsymbol{W}_j^{(k)}\beta_j\right)'\left(\boldsymbol{Z}\boldsymbol{\alpha} + \boldsymbol{X}_{\mathcal{A}_1}\boldsymbol{\beta}_{\mathcal{A}_1} + \sum_{k=1}^{q}\boldsymbol{W}^{(k)}_{\mathcal{A}_2^k}(\boldsymbol{\beta}_{\mathcal{A}_2^k}\odot\boldsymbol{\gamma}_{k,\mathcal{A}_2^k})\right)$, then $\tilde{\boldsymbol{\kappa}} = (\tilde{\kappa}_j, j \in (\tilde{\mathcal{A}}_2^k)^c)'$ with

$$\begin{aligned}\tilde{\kappa}_j &= \frac{1}{2}(\hat{\boldsymbol{\theta}}_{\mathcal{A}} - \boldsymbol{\theta}^0_{\mathcal{A}})\left(\nabla^2_{\theta_{\mathcal{A}}}\tilde{m}_j(\boldsymbol{\theta}_{\mathcal{A}})\big|_{\tilde{\boldsymbol{\theta}}_{\mathcal{A}}}\right)(\hat{\boldsymbol{\theta}}_{\mathcal{A}} - \boldsymbol{\theta}^0_{\mathcal{A}}),\\ &\leq \max_j \frac{1}{2}\lambda_{\max}\left(\boldsymbol{T}_2^{(j)}(\tilde{\beta}_j)\right)||\boldsymbol{\theta}^*_{\mathcal{A}} - \boldsymbol{\theta}^0_{\mathcal{A}}||_2,\end{aligned}$$

where $\tilde{\boldsymbol{\theta}}_{\mathcal{A}}$ lies on the line segment connecting $\boldsymbol{\theta}^*_{\mathcal{A}}$ and $\boldsymbol{\theta}^0_{\mathcal{A}}$. Here $\boldsymbol{T}_2^{(j)}(\beta_j) = \left(t_{lh}^{(j)}(\beta_j)\right)_{(q+s)\times(q+s)}$ with $t_{lh}^{(j)}(\beta_j) = \left(\boldsymbol{W}_j^{(k)}\beta_j\right)'\boldsymbol{W}_\varsigma^{(k)}$ if both l and h correspond to the ςth element of $\mathcal{A}_2^k$, and 0 otherwise.

Consider the event

$$\Omega_2 = \left\{||\boldsymbol{V}^{(k)}(\boldsymbol{\beta}^0_{(\tilde{\mathcal{A}}_2^k)^c})'\boldsymbol{\varepsilon}||_\infty \leq \zeta_n\sqrt{n}\right\},$$

with $\zeta_n = n^a(\log(n))^{1/2}$. We have

$$\begin{aligned}P(\Omega_2) &= 1 - P\left\{||\boldsymbol{V}^{(k)}(\boldsymbol{\beta}^0_{(\tilde{\mathcal{A}}_2^k)^c})'\boldsymbol{\varepsilon}||_\infty > \zeta_n\sqrt{n}\right\}\\ &\geq 1 - \sum_{j\in(\tilde{\mathcal{A}}_2^k)^c} P\left\{||\boldsymbol{V}^{(k)}(\beta_j^0)'\boldsymbol{\varepsilon}|| > \zeta_n\sqrt{n}\right\}\\ &\geq 1 - 2p\exp\left(-\frac{\zeta_n^2 n}{2\sigma^2\max_{j\in(\tilde{\mathcal{A}}_2^k)^c}||\boldsymbol{V}^{(k)}(\beta_j^0)||_2^2}\right)\to 1,\end{aligned}$$

as $\log(p) = O(n^a)$ and $||\boldsymbol{V}^{(k)}(\beta_j^0)||_2 = O(\sqrt{n})$. Thus, we have, with probability approaching 1,

$$||\boldsymbol{V}^{(k)}(\boldsymbol{\beta}^0_{(\tilde{\mathcal{A}}_2^k)^c})'\boldsymbol{\varepsilon}||_\infty = O(n^{a/2+1/2}\sqrt{\log n}).$$

Condition (C8) gives

$$(n\lambda_1)^{-1}||\boldsymbol{V}^{(k)}(\boldsymbol{\beta}^0_{(\tilde{\mathcal{A}}_2^k)^c})'\boldsymbol{\varepsilon}||_\infty = o(1).$$

For IV, with Conditions (C6) and (C8),

$$\begin{aligned}(n\lambda_1)^{-1}||IV||_\infty &= (n\lambda_1)^{-1}\left[||\boldsymbol{V}^{(k)}(\boldsymbol{\beta}^0_{(\tilde{\mathcal{A}}_2^k)^c})'\boldsymbol{G}(\boldsymbol{\beta}^0_{\mathcal{A}_2},\boldsymbol{\gamma}^0_{\mathcal{A}_1})'(\hat{\boldsymbol{\theta}}_{\mathcal{A}} - \boldsymbol{\theta}^0_{\mathcal{A}})||_\infty + ||\tilde{\boldsymbol{\kappa}}||_\infty\right]\\ &= (n\lambda_1)^{-1}\left[O(n)||\boldsymbol{\theta}^*_{\mathcal{A}} - \boldsymbol{\theta}^0_{\mathcal{A}}||_2 + O(n)||\boldsymbol{\theta}^*_{\mathcal{A}} - \boldsymbol{\theta}^0_{\mathcal{A}}||_2^2\right]\\ &= O(\lambda_1^{-1}\sqrt{s/n}) = o(1).\end{aligned}$$

With Conditions (C4), (C5) and (C8),

$$\begin{aligned}(\lambda_1)^{-1}||\lambda_2 \boldsymbol{J}_{(\tilde{\mathcal{A}}_2^k)^c,\mathcal{A}_2^k}\hat{\boldsymbol{\gamma}}_{k,\mathcal{A}_2^k}||_\infty &= (\lambda_1)^{-1}||\lambda_2 \boldsymbol{J}_{(\tilde{\mathcal{A}}_2^k)^c,\mathcal{A}_2^k}\boldsymbol{\gamma}^0_{k,\mathcal{A}_2^k} - \lambda_2 \boldsymbol{J}_{(\tilde{\mathcal{A}}_2^k)^c,\mathcal{A}_2^k}(\hat{\boldsymbol{\gamma}}_{k,\mathcal{A}_2^k} - \boldsymbol{\gamma}^0_{k,\mathcal{A}_2^k})||_\infty \\ &\leq (\lambda_1)^{-1}||\lambda_2 \boldsymbol{J}_{(\tilde{\mathcal{A}}_2^k)^c,\mathcal{A}_2^k}\boldsymbol{\gamma}^0_{k,\mathcal{A}_2^k}||_\infty + (\lambda_1)^{-1}||\lambda_2 \boldsymbol{J}_{(\tilde{\mathcal{A}}_2^k)^c,\mathcal{A}_2^k}(\hat{\boldsymbol{\gamma}}_{k,\mathcal{A}_2^k} - \boldsymbol{\gamma}^0_{k,\mathcal{A}_2^k})||_\infty \\ &= O(\lambda_1^{-1}\sqrt{s/n}) = o(1).\end{aligned}$$

This completes the proof.

2 Web Appendix B

This section includes additional results for parameter selection, simulation, and data analysis.

2.1 Analysis of c^{Pcorr} value for G factors with a network structure

To construct the adjacency matrix $\boldsymbol{A}$ for gene expression data with a network structure, the cutoff c^{Pcorr} is calculated from the Fisher transformation following the literature (Huang et al., 2011). Specifically, with r_{jl}^{Pcorr}, the Fisher transformation is conducted as

$$u_{jl} = 0.5 \log\left((1 + r_{jl}^{Pcorr})/(1 - r_{jl}^{Pcorr})\right).$$

If the correlation between gene expressions j and l is zero, then $\sqrt{n-3}u_{jl}$ is approximately distributed as $N(0,1)$. With this, a threshold c can be determined for $\sqrt{n-3}u_{jl}$, such as the 95th percentile. The corresponding threshold c^{Pcorr} for r_{jl}^{Pcorr} is then defined as $c^{Pcorr} = (\exp(2c/\sqrt{n-3}) - 1)/(\exp(2c/\sqrt{n-3}) + 1)$.

We conduct analysis on simulated data with a network structure to examine performance of the proposed approach with various values of c^{Pcorr}. Specifically, we set $n = 350$, $q = 5$ and $p = 5,000$. Thus, there are a total of 5,005 main effects and 25,000 interactions. The settings for E factors are the same as those in Section 3 (main text). Among the 5,000 G factors, there are 10 highly correlated clusters, each with size 5. G factors in different clusters are independent and generated from a multivariate Normal distribution with mean $\mathbf{0}$ and covariance matrix AR(0.9) (we refer to Section 3 for details). The remaining 4,950 G factors are generated from a multivariate Normal distribution with mean $\mathbf{0}$ and covariance matrix AR(0.3). There are 20 main G effects and 40 G-E interactions in the first four clusters with nonzero coefficients generated from Uniform(0.8,1.2). Both the network structure and the "main effects, interactions" hierarchial structure are satisfied. We consider a continuous response under the linear model with the random error following a standard Normal distribution. The evaluation measures are the same as those in Section 3, except

for the root structured error (RSE). As the true network structure can not be objectively measured, the estimated RSE is not available. The means and standard deviations (sd) of the evaluation measures as well as the means of average degree $\frac{1}{p}\sum_{j=1}^{p}\sum_{l=1}^{p}|a_{jl}|$ over 500 replicates are provided in Web Table 1. It is observed that with a larger c^{Pcorr}, the adjacency matrix is sparser with a lower average degree. When $c^{Pcorr} > 0.70$, the adjacency matrix is almost an identity matrix, and the proposed approach simplifies to that without accommodating any structure. For sensible values in the range of $[0.15, 0.25]$, results are similar. The average values of c^{Pcorr} calculated from the Fisher transformation is 0.21, which leads to satisfactory results. It is noted that the sensible range and value of c^{Pcorr} calculated from the Fisher transformation are data-dependent.

Web Table 1: Simulation results for data with a network structure and various values of c^{Pcorr}. In each cell, mean (sd) based on 500 replicates.

c^{Pcorr}	M:TP	M:FP	I:TP	I:FP	EMSE	PMSE	Average degree
0.05	18.9(1.5)	0.1(0.0)	37.1(1.5)	29.9(3.0)	5.54(0.42)	6.91(2.00)	136.0325(0.0512)
0.07	19.0(1.5)	0.1(0.0)	37.5(1.5)	29.9(3.0)	5.48(0.41)	6.75(1.94)	74.7153(0.0544)
0.09	19.2(0.7)	0.1(0.0)	37.8(1.5)	25.2(3.0)	5.42(0.43)	6.61(1.99)	53.3912(0.0508)
0.11	19.8(0.0)	0.0(0.0)	39.2(0.0)	22.0(3.0)	5.15(0.36)	5.71(1.67)	19.8843(0.0295)
0.13	19.8(0.0)	0.0(0.0)	39.2(0.0)	22.0(3.0)	5.15(0.36)	5.71(1.67)	19.8843(0.0295)
0.15	19.9(0.0)	0.0(0.0)	39.8(0.0)	16.7(3.0)	4.88(0.29)	4.89(1.37)	10.5375(0.0236)
0.17	20.0(0.0)	0.0(0.0)	40.0(0.0)	13.0(3.0)	4.38(0.31)	3.75(1.02)	5.4117(0.0135)
0.19	20.0(0.0)	0.0(0.0)	40.0(0.0)	10.9(4.4)	3.54(0.22)	2.46(0.67)	2.8851(0.0137)
0.21	20.0(0.0)	0.8(1.5)	40.0(0.0)	7.7(2.2)	3.07(0.38)	1.64(0.29)	1.7502(0.0083)
0.23	20.0(0.0)	0.8(1.5)	40.0(0.0)	7.7(2.2)	3.07(0.38)	1.64(0.29)	1.7502(0.0083)
0.25	20.0(0.0)	5.4(5.2)	38.5(1.5)	10.9(5.2)	4.44(1.05)	2.81(0.52)	1.2859(0.0057)
0.27	20.0(0.0)	10.4(3.7)	36.0(1.5)	11.4(4.4)	6.12(2.28)	3.02(0.94)	1.1093(0.0029)
0.29	20.0(0.0)	17.4(3.7)	36.0(1.5)	11.4(4.4)	7.12(2.28)	6.02(0.94)	1.0493(0.0017)
0.31	20.0(0.0)	19.2(3.7)	35.3(1.5)	11.2(3.0)	7.20(2.55)	6.43(0.83)	1.0303(0.0012)
0.33	20.0(0.0)	19.2(3.7)	35.3(1.5)	11.2(3.0)	7.20(2.55)	6.43(0.83)	1.0245(0.0007)
0.35	20.0(0.0)	20.6(5.9)	35.0(3.0)	10.2(5.9)	7.67(2.27)	6.53(1.08)	1.0245(0.0007)
0.37	20.0(0.0)	20.6(5.9)	35.0(3.0)	10.2(5.9)	7.67(2.27)	6.53(1.08)	1.0228(0.0009)
0.39	19.9(0.0)	21.3(3.7)	34.0(2.2)	9.9(4.4)	7.37(1.88)	6.68(0.94)	1.0220(0.0010)
0.41	20.0(0.0)	21.3(1.5)	35.4(0.0)	10.1(2.2)	6.95(0.16)	6.61(0.39)	1.0213(0.0011)
0.43	20.0(0.0)	23.3(3.0)	35.0(0.0)	10.3(5.9)	6.93(1.32)	7.92(0.50)	1.0207(0.0011)
0.45	20.0(0.0)	23.3(3.7)	35.0(1.5)	11.0(5.2)	7.66(2.28)	8.60(0.74)	1.0206(0.0011)
0.47	20.0(0.0)	23.9(5.2)	34.8(1.5)	11.1(6.7)	7.72(2.89)	7.81(1.33)	1.0198(0.0013)
0.49	20.0(0.0)	25.9(4.4)	33.8(1.5)	10.3(3.7)	7.83(1.77)	8.44(0.70)	1.0188(0.0013)
0.51	20.0(0.0)	31.4(9.6)	33.5(3.7)	12.0(4.4)	8.85(3.09)	9.11(1.18)	1.0175(0.0013)
0.53	19.4(0.0)	33.0(6.7)	32.0(3.0)	11.6(3.7)	8.67(2.16)	10.92(1.34)	1.0163(0.0015)
0.55	20.0(0.0)	35.7(3.0)	33.7(0.0)	11.5(5.2)	7.40(2.34)	11.51(0.82)	1.0154(0.0015)
0.57	20.0(0.0)	37.8(8.2)	32.1(2.2)	11.7(4.4)	9.28(2.37)	11.11(1.10)	1.0148(0.0013)
0.59	18.4(0.0)	46.9(37.1)	27.6(11.1)	13.2(9.6)	11.38(6.99)	17.43(5.92)	1.0134(0.0012)
0.61	20.0(0.0)	50.2(14.1)	30.6(3.0)	13.3(3.7)	10.76(3.99)	14.77(1.89)	1.0121(0.0012)
0.63	18.1(3.0)	61.9(66.7)	22.3(15.6)	13.2(6.7)	13.40(8.51)	20.78(15.22)	1.0104(0.0013)
0.65	18.1(3.0)	61.9(66.7)	22.3(15.6)	13.2(6.7)	13.40(8.51)	20.78(15.22)	1.0104(0.0013)
0.67	17.1(1.5)	75.3(49.7)	18.6(9.6)	14.4(4.4)	15.41(4.78)	24.67(18.40)	1.0086(0.0013)
0.69	16.1(1.5)	87.6(58.6)	15.7(5.2)	13.8(3.0)	16.95(5.36)	30.66(24.67)	1.0065(0.0012)
0.71	14.1(1.5)	89.8(33.4)	11.8(7.4)	15.1(6.7)	18.51(4.81)	31.63(22.03)	1.0041(0.0013)
0.73	13.8(1.5)	98.4(25.2)	10.1(5.2)	17.0(5.2)	19.78(3.55)	34.28(16.49)	1.0022(0.0009)
0.75	14.1(1.5)	101.5(26.7)	9.8(3.7)	16.0(5.9)	19.99(2.52)	36.20(16.26)	1.0022(0.0009)
0.77	14.1(1.5)	101.5(26.7)	9.8(3.7)	16.0(5.9)	19.99(2.52)	36.20(16.26)	1.0011(0.0006)
0.79	13.0(1.5)	102.2(33.4)	9.6(3.7)	16.1(4.4)	20.10(3.13)	37.81(15.93)	1.0004(0.0005)
0.81	13.0(1.5)	105.7(43.7)	9.3(4.4)	15.9(5.2)	20.72(3.30)	38.16(14.86)	1.0001(0.0006)
0.83	13.0(1.5)	105.7(43.7)	9.3(4.4)	15.9(5.2)	20.72(3.30)	38.16(14.86)	1.0000(0.0001)
0.85	12.9(1.5)	105.2(47.4)	9.4(4.4)	16.1(5.2)	20.71(3.30)	38.02(14.86)	1.0000(0.0000)
0.87	12.9(1.5)	105.2(47.4)	9.4(4.4)	16.1(5.2)	20.71(3.30)	38.02(14.86)	1.0000(0.0000)
0.89	12.9(1.5)	105.2(47.4)	9.4(4.4)	16.1(5.2)	20.71(3.30)	38.02(14.86)	1.0000(0.0000)

2.2 Analysis of r value in MCP

For r, published studies suggest setting it as fixed or examining a small number of values. Following the literature (Liu et al., 2013; Zhang, 2011), we examine four values of r, including 1.8, 3, 6, and 10, with simulation under the linear model. Scenarios with two auto-regressive (AR) correlation structures and two minor allele frequency (MAF) values have been considered. We refer to Section 3 (main text) for the detailed data generation, settings, and evaluation measures. Summary results are provided in Web Table 2. It is observed that compared to $r = 1.8$ or 10, models with $r = 3$ or 6 are less complex with fewer false positives and also fewer true positives. Overall, the proposed approach is not very sensitive to the choice of r when it is in a sensible range.

Web Table 2: Simulation results under the linear model with various values of r in MCP. In each cell, mean (sd) based on 500 replicates.

r	M:TP	M:FP	I:TP	I:FP	EMSE	SMSE	PMSE
				M1: AR(0.3)			
1.8	19.8(0.6)	0.1(0.3)	34.3(4.2)	7.3(3.4)	3.08(0.33)	2.76(0.20)	1.92(0.35)
3	19.2(1.0)	0.0(0.0)	32.6(2.9)	6.6(2.9)	3.57(0.87)	3.34(0.58)	1.91(0.64)
6	19.4(0.9)	0.0(0.0)	33.5(3.3)	9.1(3.6)	3.60(0.93)	3.22(0.56)	1.80(0.44)
10	19.7(0.6)	0.1(0.4)	35.3(2.3)	13.0(3.4)	3.53(0.87)	3.23(0.45)	1.79(0.45)
				M1: AR(0.5)			
1.8	19.6(0.7)	0.1(0.4)	33.5(5.2)	8.5(3.7)	3.36(0.41)	2.93(0.26)	2.07(0.29)
3	19.2(0.8)	0.0(0.2)	33.6(2.5)	5.1(3.1)	3.14(0.56)	3.47(0.34)	1.85(0.56)
6	19.7(0.6)	0.0(0.0)	35.2(2.7)	7.1(2.2)	3.19(0.85)	3.33(0.39)	1.76(0.44)
10	19.8(0.4)	0.1(0.3)	36.5(2.0)	11.5(3.4)	3.12(0.81)	3.35(0.33)	1.72(0.39)
				M2: AR(0.3)			
1.8	19.7(0.7)	0.5(1.9)	33.2(6.1)	7.5(3.8)	3.11(0.33)	2.75(0.20)	1.91(0.39)
3	19.1(1.2)	0.0(0.0)	32.2(3.6)	5.9(2.5)	3.57(0.93)	3.36(0.64)	1.95(0.73)
6	19.3(1.3)	0.6(1.2)	32.9(4.0)	8.9(3.1)	3.78(0.98)	3.29(0.64)	1.92(0.84)
10	19.3(1.5)	2.6(4.1)	34.1(3.4)	11.3(4.2)	3.71(0.92)	3.42(0.47)	2.08(0.98)
				M2: AR(0.5)			
1.8	19.7(0.5)	0.2(1.0)	33.9(4.7)	8.4(3.7)	3.12(0.40)	2.89(0.23)	2.01(0.32)
3	19.1(0.9)	0.0(0.2)	33.6(2.6)	5.4(3.4)	3.20(0.65)	3.51(0.34)	1.86(0.54)
6	19.6(0.8)	0.3(0.7)	34.5(3.0)	7.0(2.7)	3.33(0.82)	3.42(0.42)	1.83(0.43)
10	19.4(0.9)	1.8(2.1)	35.4(2.7)	9.8(3.2)	3.51(0.85)	3.64(0.45)	2.00(0.62)

2.3 Detailed simulation settings

Two approaches, A1 and A2, are adopted to simulate G factors which mimic SNP data coded with three categories (0, 1, 2) for genotypes (aa, Aa, AA).

Approach A1 includes two steps, under which we first generate p continuous variables from a multivariate Normal distribution with mean $\mathbf{0}$ and covariance matrix $\mathbf{\Sigma} = (\sigma_{jl})_{p\times p}$, and then dichotomize the continuous variables at the q_1 and q_2 percentiles to generate 3-level G measurements (0, 1, 2). In the first step, two correlation structures are considered with different parameters. The first is the auto-regressive (AR) structure with $\sigma_{jl} = \rho^{|j-l|}$. We consider two levels of correlation with $\rho = 0.3$ and 0.5 (referred to as AR(0.3) and AR(0.5)). The second is the banded correlation structure where two specific scenarios are considered. The first one (Band1) has $\sigma_{jl} = 1$ if $j = l$, 0.3 if $|j - l| = 1$, and 0 otherwise. The second one (Band2) has $\sigma_{jl} = 1$ if $j = l$, 0.5 if $|j - l| = 1$, 0.3 if $|j - l| = 2$, and 0 otherwise. In the second step, q_1 and q_2 are adjusted to generate G factors with different minor allele frequency (MAF) values. Consider two specific scenarios. Under the first scenario (M1), all of the G factors have MAF=0.05 with $q_1 = 0.91$ and $q_2 = 0.99$. Under the second one (M2), a half of the G factors have MAF=0.05, and the other half have MAF=0.15 with $q_1 = 0.73$ and $q_2 = 0.97$.

Under A2, we simulate G factors with the pairwise LD structure. Specifically, denote p_A and p_B as the MAFs of alleles A and B for two adjacent SNPs. The LD is defined as $\phi = r_{LD}\sqrt{p_A(1-p_A)p_B(1-p_B)}$ with pairwise correlation r_{LD}. Then, the four haplotypes ab, aB, Ab, AB have frequencies $(1-p_A)(1-p_B)+\phi$, $(1-p_A)p_B-\phi$, $p_A(1-p_B)-\phi$, and $p_Ap_B+\phi$, respectively. With the Hardy-Weinberg equilibrium assumption, we simulate the SNP genotype (AA, Aa, aa) at locus 1 from a multinomial distribution given corresponding frequencies $(p_A^2, 2p_A^2(1-p_A), (1-p_A)^2)$ and that at locus 2 accordingly from the conditional probability defined in Cui et al. (2008). Two pairwise correlations are considered with $r_{LD} = 0.3$ and $r_{LD} = 0.5$ (referred to as LD(0.3) and LD(0.5)). For MAF, two scenarios similar to those in Step 2 of A1 are considered.

The true coefficient values for the main G effects and interactions are set as follows: $\beta_j = \sin(0.2j + 0.9) + 0.2$ for $j = 1, \cdots, 10$, $\beta_j = 0.5(j - 10)$ for $j = 11, \cdots, 15$, $\beta_j = 0.5(21 - j)$ for $j = 16, \cdots, 20$, $\eta_{1j} = 0.2j + 0.2$ for $j = 1, \cdots, 5$, $\eta_{1j} = 0.2(11 - j) + 0.2$ for $j = 6, \cdots, 11$, $\eta_{2j} = 0.2\sqrt{3j - 32}$ for $j = 11, \cdots, 15$, $\eta_{2j} = 0.2\sqrt{63 - 3j}$ for $j = 16, \cdots, 20$, $\eta_{3j} = -(0.2j - 0.9)^2 + 1.5$ for $j = 1, \cdots, 10$, and $\eta_{3j} = -(0.2j - 3.2)^2 + 1.6$ for $j = 11, \cdots, 20$. The rest of the effects are zero.

2.4 Simulation results under the AFT model

Web Table 3: Simulation results under the AFT model with MAF setting M1. In each cell, mean (sd) based on 500 replicates.

	M:TP	M:FP	I:TP	I:FP	RSSE	RSE	Cstat
				AR(0.3)			
MA	0.8(1.7)	40.3(38.0)	2.8(2.8)	95.5(77.6)	14.05(4.77)	26.37(19.93)	0.74(0.05)
HierMCP	13.5(1.9)	38.3(6.3)	0.8(0.9)	0.5(0.8)	9.72(0.57)	15.84(1.81)	0.81(0.03)
SMCP	4.8(4.8)	6.5(13.7)	9.1(6.6)	31.9(15.7)	8.54(0.56)	3.24(0.80)	0.85(0.03)
Proposed	19.2(1.2)	1.2(6.7)	33.1(6.1)	6.1(4.3)	2.99(1.09)	2.29(0.72)	0.93(0.02)
				AR(0.5)			
MA	1.8(2.7)	55.8(38.9)	4.7(4.0)	127.3(75.0)	71.88(62.71)	173.12(166.62)	0.56(0.10)
HierMCP	12.9(1.8)	40.6(6.9)	0.9(1.0)	0.9(0.9)	10.13(0.69)	17.32(2.16)	0.80(0.03)
SMCP	6.5(4.9)	7.9(12.9)	8.9(7.2)	34.4(16.3)	8.36(0.72)	3.50(0.91)	0.85(0.03)
Proposed	19.3(1.3)	0.4(2.3)	34.2(4.9)	5.4(4.0)	2.82(1.02)	2.18(0.58)	0.93(0.01)
				Band1			
MA	1.1(1.9)	41.6(42.5)	2.5(2.6)	92.7(80.5)	13.71(4.68)	25.49(20.59)	0.72(0.08)
HierMCP	13.5(1.7)	37.1(5.7)	0.7(0.7)	0.4(0.8)	9.66(0.51)	15.60(1.67)	0.81(0.03)
SMCP	4.6(4.7)	5.1(12.0)	9.2(6.5)	28.5(15.6)	8.51(0.65)	3.20(0.92)	0.86(0.03)
Proposed	19.3(1.2)	0.4(1.4)	33.4(5.4)	6.1(4.9)	2.92(0.90)	2.25(0.76)	0.93(0.01)
				Band2			
MA	1.8(2.3)	59.6(41.1)	5.5(3.9)	131.5(73.7)	100.86(90.94)	245.78(233.22)	0.54(0.07)
HierMCP	12.8(1.9)	42.0(8.2)	1.2(1.0)	0.6(0.8)	10.19(0.76)	17.70(2.15)	0.80(0.03)
SMCP	9.4(5.2)	17.8(18.5)	9.3(6.5)	35.8(14.1)	8.10(0.83)	4.00(1.01)	0.85(0.03)
Proposed	19.0(2.3)	1.6(7.1)	33.2(7.5)	5.5(4.9)	2.97(1.46)	2.24(0.90)	0.93(0.02)
				LD(0.3)			
MA	1.5(2.6)	48.4(42.2)	4.1(3.5)	103.0(78.6)	17.58(8.62)	37.39(34.40)	0.70(0.09)
HierMCP	13.4(2.0)	50.4(8.2)	0.6(0.9)	0.2(0.6)	10.58(0.81)	18.96(2.47)	0.80(0.03)
SMCP	4.4(4.5)	4.7(11.7)	11.7(8.1)	21.3(11.2)	8.33(0.88)	2.98(0.67)	0.86(0.03)
Proposed	19.2(1.6)	0.3(1.5)	34.0(5.7)	4.6(3.2)	2.79(1.04)	2.11(0.68)	0.93(0.01)
				LD(0.5)			
MA	2.1(3.1)	57.5(39.7)	7.3(4.9)	121.8(70.1)	49.40(40.45)	125.38(118.24)	0.58(0.10)
HierMCP	12.6(2.1)	55.4(9.6)	0.8(0.8)	0.3(0.6)	11.18(0.90)	21.26(2.79)	0.78(0.07)
SMCP	6.8(5.7)	8.0(13.8)	13.8(8.6)	22.7(12.9)	7.87(1.02)	3.16(1.04)	0.85(0.08)
Proposed	19.4(1.1)	0.2(1.8)	34.1(4.7)	4.8(3.9)	2.78(0.89)	2.10(0.60)	0.93(0.02)

Web Table 4: Simulation results under the AFT model with MAF setting M2. In each cell, mean (sd) based on 500 replicates.

	M:TP	M:FP	I:TP	I:FP	RSSE	RSE	Cstat
				AR(0.3)			
MA	2.1(2.7)	60.5(28.0)	5.4(3.8)	157.2(67.9)	145.73(130.23)	360.20(327.86)	0.55(0.04)
HierMCP	13.6(1.9)	34.5(6.7)	1.2(1.1)	0.7(0.9)	9.56(0.60)	15.53(2.01)	0.82(0.03)
SMCP	5.8(4.8)	9.4(15.1)	4.1(2.7)	81.9(20.0)	8.68(0.37)	3.47(0.57)	0.82(0.05)
Proposed	18.8(2.0)	8.1(17.1)	30.1(9.9)	6.5(4.1)	3.52(1.74)	2.66(1.07)	0.92(0.03)
				AR(0.5)			
MA	3.2(3.5)	68.6(21.8)	7.8(4.9)	172.3(47.2)	184.51(107.73)	451.07(267.52)	0.53(0.04)
HierMCP	12.8(1.8)	37.9(7.4)	1.2(1.2)	0.9(0.9)	10.05(0.75)	17.06(2.41)	0.80(0.03)
SMCP	7.3(5.0)	12.4(18.1)	3.8(3.3)	80.6(20.1)	8.56(0.53)	3.60(0.55)	0.82(0.03)
Proposed	18.9(2.2)	6.5(19.5)	31.9(9.4)	5.6(4.9)	3.22(1.80)	2.37(0.97)	0.92(0.03)
				Band1			
MA	2.2(2.8)	61.6(33.5)	5.1(3.8)	151.4(71.4)	135.71(124.35)	330.96(314.37)	0.52(0.05)
HierMCP	13.8(1.7)	33.5(6.0)	1.3(1.1)	0.7(0.7)	9.45(0.55)	15.20(1.91)	0.82(0.03)
SMCP	6.3(4.7)	8.0(13.0)	4.0(3.2)	79.8(19.9)	8.62(0.42)	3.40(0.51)	0.84(0.02)
Proposed	18.9(1.7)	14.8(28.5)	28.3(11.3)	6.3(4.6)	3.87(1.99)	2.82(1.22)	0.91(0.03)
				Band2			
MA	3.8(3.6)	64.9(24.8)	8.8(4.7)	167.1(51.5)	207.33(179.97)	519.71(455.48)	0.53(0.04)
HierMCP	12.6(1.8)	38.8(8.1)	1.6(1.3)	1.0(1.0)	10.11(0.75)	17.74(2.17)	0.80(0.03)
SMCP	8.0(4.5)	12.8(17.4)	3.6(3.2)	79.5(21.2)	8.56(0.49)	3.71(0.62)	0.82(0.03)
Proposed	18.5(2.8)	11.3(24.1)	29.4(11.5)	5.5(4.0)	3.67(2.08)	2.65(1.26)	0.91(0.03)
				LD(0.3)			
MA	2.7(3.6)	68.5(32.7)	6.7(4.6)	163.2(69.2)	144.42(131.19)	357.04(335.25)	0.50(0.07)
HierMCP	13.6(2.0)	46.7(8.5)	1.1(1.2)	0.4(0.6)	10.39(0.85)	18.61(2.69)	0.81(0.03)
SMCP	5.8(4.5)	7.6(12.9)	6.6(4.5)	72.0(15.5)	8.46(0.67)	3.19(0.35)	0.84(0.04)
Proposed	18.8(2.2)	11.0(20.2)	30.0(10.6)	4.9(2.9)	3.53(1.91)	2.58(1.08)	0.92(0.02)
				LD(0.5)			
MA	3.8(4.2)	66.5(23.0)	10.5(5.8)	162.1(45.1)	158.95(107.48)	401.32(279.82)	0.56(0.06)
HierMCP	12.4(2.1)	52.4(9.8)	1.2(1.1)	0.5(0.7)	11.10(0.92)	21.15(2.92)	0.78(0.07)
SMCP	6.5(5.4)	7.8(15.9)	8.4(5.0)	67.7(18.4)	8.20(0.76)	3.07(0.66)	0.82(0.08)
Proposed	18.9(1.8)	8.1(18.9)	31.0(8.9)	4.9(3.9)	3.33(1.59)	2.40(0.98)	0.92(0.03)

2.5 Simulation results under the linear model with highly correlated predictors

For the linear model with MAF setting M1, we simulate three additional scenarios with highly correlated predictors. Specifically, the first and second scenarios are based on approach A1. The first (AR(0.9)) has an AR correlation structure with $\rho = 0.9$. The second (Band3) has $\sigma_{jl} = 1$ if $j = l$, 0.7 if $|j - l| = 1$, 0.4 if $|j - l| = 2$, 0.2 if $|j - l| = 3$, and 0 otherwise. The third scenario (LD(0.9)) is based on approach A2 with the pairwise correlation $r_{LD} = 0.9$. Summary results are provided in Web Table 5.

Web Table 5: Simulation results under the linear model with MAF setting M1 and highly correlated predictors. In each cell, mean (sd) based on 500 replicates.

	M:TP	M:FP	I:TP	I:FP	EMSE	SMSE	PMSE
				AR(0.9)			
MA	4.8(5.0)	33.3(21.4)	20.0(5.2)	169.0(36.7)	69.96(11.00)	167.22(123.35)	89.07(44.88)
HierMCP	10.9(1.8)	64.3(86.3)	11.0(4.4)	41.6(13.6)	21.91(4.77)	52.49(10.10)	34.18(33.73)
SMCP	12.8(3.2)	0.0(0.0)	31.1(3.6)	67.0(16.7)	6.39(1.59)	3.32(0.28)	3.41(1.58)
Proposed	19.3(0.8)	0.1(0.3)	34.3(2.5)	5.5(3.9)	3.27(0.67)	3.66(0.50)	2.03(0.67)
				Band3			
MA	0.7(1.9)	15.8(14.4)	5.2(4.2)	127.3(77.9)	27.16(13.65)	63.85(39.38)	65.07(31.68)
HierMCP	15.3(1.9)	211.7(71.7)	6.2(3.7)	23.8(10.8)	16.70(2.05)	34.76(4.55)	35.78(9.47)
SMCP	8.4(5.6)	0.0(0.0)	31.6(5.7)	69.5(20.6)	5.85(1.70)	3.21(0.33)	3.21(2.57)
Proposed	19.8(0.5)	0.0(0.2)	36.6(1.6)	5.8(2.9)	2.68(0.46)	3.16(0.35)	1.63(0.33)
				LD(0.9)			
MA	0.0(0.2)	8.2(11.4)	10.3(6.1)	88.8(84.3)	60.01(46.61)	175.64(152.29)	85.87(50.86)
HierMCP	8.9(1.7)	90.5(100.4)	8.8(4.1)	28.0(12.1)	19.95(3.97)	47.84(7.93)	26.69(23.07)
SMCP	10.9(5.4)	0.0(0.0)	32.9(3.0)	54.3(16.1)	5.39(1.36)	3.29(0.20)	2.21(0.55)
Proposed	18.4(1.1)	0.0(0.0)	31.9(2.7)	2.6(2.8)	4.05(0.73)	4.32(0.42)	1.82(0.47)

2.6 Estimation results for data analysis with the proposed approach

Web Table 6: Analysis of the GENEVA diabetes data (NHS/HPFS) using the proposed approach: identified main effects and interactions.

SNP	Location	Gene*		age	famdb	act	trans	ceraf	heme
				-0.3331	0.1711	-0.2659	0.2185	-0.3332	0.5615
rs17090278	61679934	RP11-593F5.2	-0.0016						
rs10019557	61684734	RP11-593F5.2	-0.0019						
rs17090285	61695580	RP11-593F5.2	-0.0019						
rs17090286	61695978	RP11-593F5.2	-0.0016						
rs11731112	65554491	RP11-63H19.1	0.0015						
rs4355422	65557183	RP11-63H19.1	0.0015						
rs1430504	65681190	RP11-707A18.1	-0.0035						
rs6551878	65690589	RP11-707A18.1	-0.0056						
rs6823601	65691562	RP11-707A18.1	-0.0043						
rs13151560	67160442	MIR1269A	0.0025						
rs1858306	67161812	MIR1269A	0.0083	0.0021	0.002	-0.0013	0.0027		
rs10016795	67167064	MIR1269A	0.0174	0.0047	0.0074	-0.0054	0.0096	-0.0026	-0.0034
rs17087008	67188696	MIR1269A	0.0282	0.0049	0.0169	-0.0145	0.0213	-0.0046	-0.0119
rs12331987	67188980	MIR1269A	0.0373		0.0261	-0.0265	0.0308	-0.0056	-0.0256
rs10000219	67200024	MIR1269A	0.0405	-0.0105	0.0274	-0.0336	0.0304	-0.0028	-0.0357
rs4860208	67201368	MIR1269A	0.035	-0.0142	0.02	-0.0277	0.0197		-0.0301
rs1511286	67213473	MIR1269A	0.0215	-0.0077	0.0082	-0.0126	0.0072		-0.0139
rs1033095	67232011	RPS23P3	0.0064	-0.0012	0.0011	-0.0019			-0.002
rs11936928	67489994	RPS23P3	0.0014						
rs6838523	67494918	RPS23P3	0.0014						
rs10033058	69177408	YTHDC1	0.0055						
rs2293595	69178920	YTHDC1	0.0097						0.0031
rs12649108	69181942	YTHDC1	0.0095						0.0036
rs17089267	69183791	YTHDC1	0.0067						0.002
rs1730872	69189048	YTHDC1	0.0018						
rs1399247	70973970	CSN1S2AP	-0.0012						
rs1717600	70974315	CSN1S2AP	-0.0013						
rs11936367	72884978	NPFFR2	-0.0013						
rs7699403	72893324	NPFFR2	-0.0042						
rs6856651	72896457	NPFFR2	-0.0068	0.0013					-0.001
rs7654531	72900621	NPFFR2	-0.0079	0.0018					-0.0013
rs6824342	72903182	NPFFR2	-0.0074	0.0016					
rs6824703	72903318	NPFFR2	-0.0051						
rs7687603	72915996	NPFFR2	-0.002						
rs12649753	74940765	CXCL2	0.0092			-0.0023			0.0026
rs546829	74956372	CXCL2	0.0199	-0.0012	-0.0028	-0.012	0.0024	0.0027	0.0129
rs1837559	74959093	CXCL2	0.0257	-0.003	-0.0061	-0.0217	0.0033	0.0034	0.0226

Continued on the next page

Web Table 6: Continued from the previous page

SNP	Position	Gene		age	famdb	act	trans	ceraf	heme
rs9131	74963049	CXCL2	0.0232	-0.0038	-0.0066	-0.0196	0.0021	0.0011	0.0199
rs1866755	74978340	MTHFD2L	0.0156	-0.0024	-0.0038	-0.0098			0.0096
rs7686861	74998484	MTHFD2L	0.0064			-0.0019			0.0018
rs11737437	80262521	NAA11	-0.0019						
rs10004440	80272792	NAA11	-0.0043						
rs2903619	80281513	NAA11	-0.0056						
rs11731223	80290084	GK2	-0.0051						
rs6534350	80305179	GK2	-0.0025						
rs17003746	80314643	GK2	-0.0015						
rs11930550	80317724	GK2	-0.0019						
rs17003749	80317772	GK2	-0.0014						
rs7680648	82666782	RP11-689K5.3	-0.0018						
rs17561568	82667783	RP11-689K5.3	-0.0068	-0.0011					
rs35036928	82671170	RP11-689K5.3	-0.0137	-0.0036		-0.0026			-0.0032
rs4693369	82671234	RP11-689K5.3	-0.0156	-0.0046		-0.0027			-0.0049
rs12508164	82671299	RP11-689K5.3	-0.012	-0.0029		-0.001			-0.0038
rs7672440	82671938	RP11-689K5.3	-0.0079	-0.0013					-0.0018
rs1353661	82672523	RP11-689K5.3	-0.0025						
rs676592	82733530	RP11-689K5.3	-0.0012						
rs1993798	82762741	RP11-689K5.3	0.0047						
rs2868257	82762839	RP11-689K5.3	0.0072						
rs6535281	82763010	RP11-689K5.3	0.0048						
rs6535291	82926694	RP11-689K5.3	-0.0011						
rs434193	86253489	ARHGAP24	-0.0032						
rs6842681	86253994	ARHGAP24	-0.0084	0.0019		-0.0036	-0.0033		-0.0031
rs425196	86255297	ARHGAP24	-0.0144	0.0054		-0.0083	-0.0073		-0.0076
rs416035	86255366	ARHGAP24	-0.0203	0.0111	-0.0014	-0.014	-0.0117		-0.0136
rs432755	86255399	ARHGAP24	-0.0252	0.0174	-0.003	-0.0196	-0.016		-0.0196
rs375432	86255845	ARHGAP24	-0.0276	0.0214	-0.0041	-0.0224	-0.0185		-0.023
rs425642	86255997	ARHGAP24	-0.0257	0.0207	-0.0037	-0.0194	-0.016		-0.021
rs407430	86256356	ARHGAP24	-0.0204	0.0151	-0.0023	-0.0128	-0.0109		-0.0148
rs400023	86256538	ARHGAP24	-0.0131	0.0077		-0.0062	-0.0056		-0.0077
rs585787	86257453	ARHGAP24	-0.006	0.0023		-0.0019	-0.0019		-0.0025
rs380632	86264123	ARHGAP24	-0.0012						
rs2726516	106346206	PPA2	0.0022						
rs2636739	106352105	PPA2	0.0022						

* Genes that SNPs belong to or are the closest to.

Web Table 7: Analysis of the TCGA SKCM data using the proposed approach: identified main effects and interactions

Gene		Age	PN	Gender	Breslow's depth	Clark level
		-0.1381	-0.3077	0.0536	-0.2158	-0.1590
ACTL6B	-0.0067					
BLOC1S5	0.0188	-0.0012				
C3ORF67	0.0420	-0.0014	0.0064		0.0169	0.0016
CLEC2L	-0.0069					
CLPB	-0.0048					
CREG1	0.0281			-0.0018		
CRYBA1	-0.0037					
ENDOD1	0.0160		-0.0013	-0.0014		-0.0020
ETV3	-0.0019					
FAM131B	0.0041					
GOLPH3L	-0.0024					
IFNA7	-0.0018					
IL17A	0.0046					
IL17F	0.0143			-0.0014		
IL34	0.0030					
INPP5K	0.0093					
INTS4	-0.0055					
ISL2	-0.0022					
KCNE1	0.0281	-0.0026		0.0024		-0.0055
LAMTOR1	-0.0078					
LANCL2	0.0149					0.0012
LYNX1	-0.0261	0.0012	0.0021	-0.0011		-0.0026
MEPE	0.0144	-0.0012				-0.0014
METTL21C	0.0087					
NKAIN2	-0.0239	0.0019	0.0027		0.0013	0.0033
NKAIN3	-0.0019					
NOV	0.0422		-0.0070	-0.0085	-0.0060	0.0047
OR5L2	0.0452	-0.0103	-0.0064	-0.0103	-0.0037	0.0053
PRSS3	-0.0100					
PXDNL	-0.0106					
RAC1	-0.0177					0.0013
RAET1L	0.0093					
RIMS2	0.0076					
RPTN	-0.0023					
SERPINB13	-0.0079					
SERPINB3	-0.0018					

Continued on the next page

Web Table 7: Continued from the previous page

Gene		Age	PN	Gender	Breslow's depth	Clark level
SETD3	0.0139					
SKIDA1	-0.0075					
SLFN13	0.0212	-0.0023		-0.0023		
SPINK4	0.0012					
SPRR2B	0.0015					
STMN4	0.0035					
STPG4	-0.0055					
SYT12	0.0027					
TAS2R1	-0.0056					
TRIM46	-0.0117					
UBE2V1	-0.0317	-0.0018	0.0045			0.0019
UGT1A7	-0.0056					
WDPCP	-0.0843	-0.0194	-0.0322	-0.0018	0.0206	-0.0192
WDR77	-0.0137					

2.7 Implications of the markers identified with the proposed approach

2.7.1 GENEVA diabetes data (NHS/HPFS)

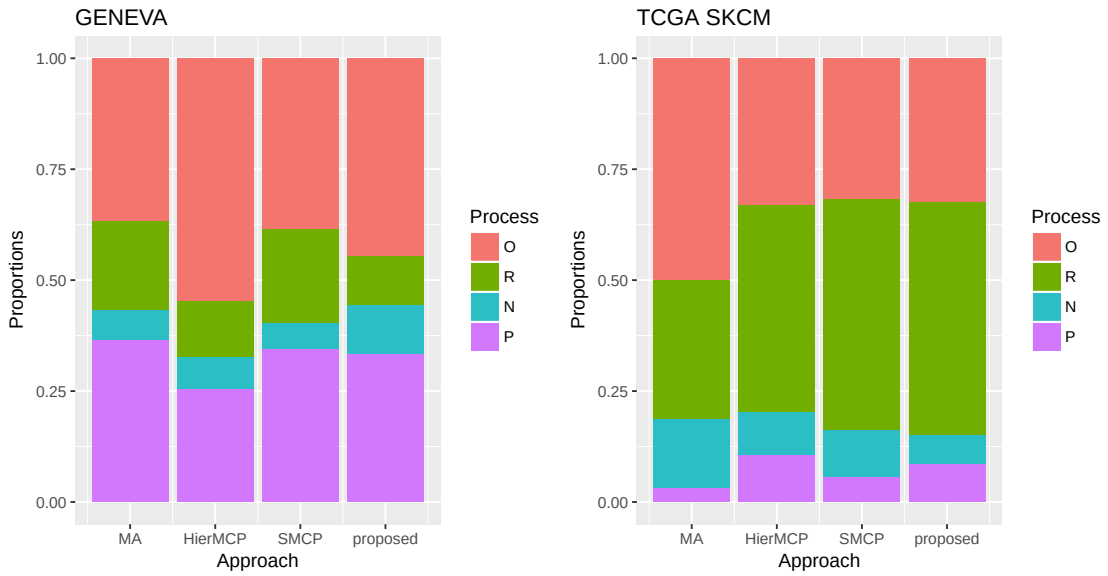
Gene NPFFR2 has been found to play an important role in obesity predisposition, and some NPFFR2 haplotypes have been suggested to be strongly protective against obesity (Hunt et al., 2011). Gene CXCL2 has been shown to be up-regulated in obese subjects and contribute to the chemotaxis of neutrophils which are one type of circulating cells greatly activated in obese subjects (Rouault et al., 2013). Published analysis has also found that the enzyme encoded by gene GK2 plays a key role in the regulation of glycerol uptake and metabolism (Song et al., 2011), and its activity in human adipose tissue is related to obesity (Chakrabarty et al., 1984).

2.7.2 TCGA skin cutaneous melanoma data

ACTL6A (BAF53) is a subunit of the SWI/SNF complex which has been found to be critical for the expression of microphthalmia-associated transcription factor in melanoma cells (Vachtenheim et al., 2010). FAM131B-BRAF fusion has been observed to comprise an alternative mechanism of MAPK pathway activation (Cin et al., 2011), and MAPK pathway plays important roles in melanoma etiology, prognosis, and treatment (Fecher et al., 2008). Gene GOLPH3 has been shown to regulate cell size and enhance growth-factor-induced mTOR signaling in melanoma cells (Liu et al., 2018), and suggested as a new oncogene that is commonly targeted for amplification in melanoma (Scott et al., 2009). Gene IL17A has been found to have tumorigenic effects in melanoma cell lines, which are related to the signal transducer and activator of transcription pathway signaling (Wang et al., 2009). It has been demonstrated that mutations in RAC1 are potentially biologically associated with cutaneous melanoma, and the pharmacological inhibition of downstream effectors of RAC1 signaling can be of therapeutic benefit (Krauthammer et al., 2012). In addition, gene SERPINB3 has been reported to be up-regulated in benign hyperplasia in melanoma (Jung et al., 2011).

2.8 Biological similarity analysis based on Gene Ontology

To gain deeper insight into differences of the markers identified by different approaches, we conduct a closer examination of the Gene Ontology (GO) biological processes. For GENEVA data, the proposed approach together with three alternatives identify a total of 110 unique genes (that the identified SNPs belong to or are the closest to). These genes represent a total of 51 GO biological processes, of which the p-values computed from the GO enrichment analysis are smaller than 0.05. Here the GO enrichment analysis is realized using the R package *GOSim*. It is observed that the majority of these processes are related to "regulation". We further separate the 51 processes into four categories: positive regulation (P), negative regulation (N), regulation (R, without a well-defined "direction"), and other (O). In Web Figure 3, we provide the proportions of genes that have the four categories of processes with different approaches. It is observed that the four approaches have different distributions with a moderate level of similarity. Similar analysis is conducted on the TCGA SKCM data, where a total of 238 unique genes are identified. These genes represent a total of 124 GO biological processes, which are also separated into four categories. The proportions of genes that have the four categories of processes with different approaches are provided in Web Figure 3. The three joint analysis approaches, HierMCP, SMCP, and the proposed one, have a higher level of similarity compared to MA. These three approaches have slightly different proportions of positive regulation, negative regulation, and regulation.



Web Figure 3: Data analysis: proportions of genes that have the four categories of processes with different approaches. Left: GENEVA data. Right: TCGA SKCM data.

2.9 Estimation results for data analysis with the three alternatives

Web Table 8: Analysis of the GENEVA diabetes data (NHS/HPFS) using MA: identified main effects and interactions.

SNP	Location	Gene		age	famdb	act	trans	ceraf	heme
				0.18	0.1547	-0.3866	0.0691	-0.6648	0.5573
rs10021002	61129136	AC095061.1		-0.2414					
rs10019682	61228770	AC095061.1		0.1505					
rs11946495	61238161	AC095061.1		0.4238					
rs4318680	61249881	AC095061.1		-0.5856					
rs1511104	61287186	AC095061.1		0.0450					
rs1546511	61327179	AC095061.1		-0.4675					
rs10021608	61352529	AC095061.1		0.0969					
rs17090278	61679934	RP11-593F5.2	0.2229						
rs17090286	61695978	RP11-593F5.2	-0.6965						
rs13122165	61762061	RP11-593F5.2	0.0407						
rs17828144	61803167	RP11-593F5.2	0.0200						
rs17239101	62495735	LPHN3							0.3395
rs7671984	63243647	RP11-30P21.2						0.1845	
rs1430504	65681190	RP11-707A18.1	-0.2050						
rs6551878	65690589	RP11-707A18.1	-0.1019						
rs13435819	65736743	RP11-707A18.1							-0.0715
rs10002424	65744523	RP11-707A18.1							0.1317
rs10028673	65748015	RP11-707A18.1							-0.1254
rs13139468	65763976	RP11-707A18.1							0.1179
rs10005825	65778213	RP11-707A18.1							-0.4144
rs7667410	66486986	EPHA5					1.3617		
rs7659227	66496170	EPHA5					-1.2760		
rs4370201	66500571	EPHA5					0.2358		
rs13107026	67139034	MIR1269A	-0.3099						
rs1397755	67140834	MIR1269A	0.6150						
rs13151560	67160442	MIR1269A	0.2824						
rs1858306	67161812	MIR1269A	-0.5931						
rs12331987	67188980	MIR1269A	0.2053						
rs10000219	67200024	MIR1269A	-0.3886						
rs4860208	67201368	MIR1269A	0.2988						
rs1511286	67213473	MIR1269A	0.2312						
rs2136822	67482764	RPS23P3	0.1948						
rs11936928	67489994	RPS23P3	0.0932						
rs6838523	67494918	RPS23P3	0.0932						

Continued on the next page

Web Table 8: Continued from the previous page

SNP	Position	Gene	age	famdb	act	trans	ceraf	heme
rs17088752	68747036	UBA6-AS1	0.6826					
rs2293595	69178920	YTHDC1	0.5827					
rs17089267	69183791	YTHDC1	-0.2146					
rs11249477	70934425	CSN1S2AP	-0.2276					
rs1399247	70973970	CSN1S2AP	0.1586					
rs1717600	70974315	CSN1S2AP	-0.3458					
rs1842478	71227802	SMR3A						-0.3286
rs10003790	71957713	DCK	0.5628					
rs12649753	74940765	RP11-629B11.4	0.2633					
rs13148163	77504400	SHROOM3						-0.2491
rs7810	77968346	CCNI				0.5323		
rs4272042	78053975	RNU6-1187P		0.4129				
rs2903455	79095407	FRAS1						1.0048
rs17003019	79095443	FRAS1						-1.3661
rs143371	79257710	FRAS1			-0.4509			
rs7681755	80184659	LINC01088	-0.4519					
rs11731223	80290084	GK2	-0.1388					
rs17003746	80314643	GK2	-0.1966					
rs11937407	81413618	C4orf22		-2.7727				
rs6858262	81441388	C4orf22		0.2385				
rs17004924	81445307	C4orf22		-2.7994				
rs1843563	81454827	C4orf22		4.931				
rs10004901	81897005	C4orf22	-0.441					
rs733241	82116677	RP11-100N20.1					0.2034	
rs1391262	82663506	RP11-689K5.3	-0.0371					
rs35036928	82671170	RP11-689K5.3	-0.9472					
rs4693369	82671234	RP11-689K5.3	0.8823					
rs7672440	82671938	RP11-689K5.3	-0.0615					
rs676592	82733530	RP11-689K5.3	-0.1434					
rs1993798	82762741	RP11-689K5.3	1.6689					
rs2868257	82762839	RP11-689K5.3	0.2131					
rs6535281	82763010	RP11-689K5.3	-0.461					
rs612318	82764165	RP11-689K5.3	-0.2764					
rs1824657	82783136	RP11-689K5.3	-1.4809					

Continued on the next page

Web Table 8: Continued from the previous page

SNP	Position	Gene		age	famdb	act	trans	ceraf	heme
rs11722328	82818968	RP11-689K5.3	0.5812						
rs2199487	82820008	RP11-689K5.3	-0.3459						
rs7436836	84288014	HPSE							0.2181
rs392112	86244117	RP11-218C23.1	-2.8736						
rs434193	86253489	RP11-218C23.1	-0.8468						
rs416035	86255366	RP11-218C23.1	-4.5803						
rs432755	86255399	RP11-218C23.1	2.7866						
rs375432	86255845	RP11-218C23.1	-0.2408						
rs407430	86256356	RP11-218C23.1	8.9916						
rs400023	86256538	RP11-218C23.1	-3.514						
rs7656367	89196821	PPM1K							0.2831
rs3775373	89743821	FAM13A	-0.3285						
rs10031177	91941942	CCSER1					-0.7535		
rs1898905	94650978	GRID2			0.1739				
rs17021080	94674019	GRID2			0.2608				
rs10022681	97114273	RNU6-34P						-0.3874	
rs6532594	97114706	RNU6-34P						0.1882	
rs1395199	97177286	RP11-145G20.1						-0.5635	
rs6817234	97205150	RP11-145G20.1						0.5911	
rs6817554	97293102	RP11-145G20.1						0.8506	
rs6840466	97311763	RP11-145G20.1						-0.3888	
rs9307172	97395506	RP11-145G20.1						-1.1833	
rs28526823	97395638	RP11-145G20.1						1.67	
rs17025552	97409408	RP11-145G20.1						-0.7591	
rs11724479	97419924	RP11-145G20.1						-0.3917	
rs6532609	97421141	RP11-145G20.1						1.0698	
rs10022608	97444956	RP11-145G20.1						-0.098	
rs6832719	97458376	RP11-145G20.1						-9.5649	
rs6532632	97459243	RP11-145G20.1						9.3413	
rs2004316	99381148	TSPAN5			-0.4301				
rs11935423	100477462	TRMT10A		-0.4326					
rs13126505	102865304	BANK1		-0.5155					
rs2686293	107511296	RP13-612N21.1	-0.3627						
rs1004472	109215088	LEF1-AS1							-2.58
rs2078126	109215175	LEF1-AS1							0.5862
rs6831553	109218208	LEF1-AS1							1.8794
rs220615	109239451	LEF1-AS1							0.1964
rs3932216	109242863	LEF1-AS1							-3.1306
rs11097979	109248691	LEF1-AS1							3.3899

Web Table 9: Analysis of the GENEVA diabetes data (NHS/HPFS) using HierMCP: identified main effects and interactions.

SNP	Location	Gene		age	famdb	act	trans	ceraf	heme
				-0.4322	0.3991	-0.2762	0.3159	-0.4147	0.33
rs12645408	60879952	Y_RNA	-0.4366					0.2041	-0.1355
rs17090278	61679934	RP11-593F5.2	-0.4349	0.1885	0.0229			0.0841	
rs17828330	62226863	LPHN3	0.4944			-0.0854	0.0088	0.1576	
rs6828446	65304222	TECRL	-0.6452			-0.3828			0.0124
rs12233864	65608452	RP11-63H19.1	-0.4627	0.2458			-0.1327		0.1765
rs13130005	66100747	RP11-498E11.2	0.8005	0.0363			0.0066	-0.0442	
rs13124187	66208911	EPHA5	-0.6015		-0.1894				
rs13107026	67139034	MIR1269A	0.433				0.1242	-0.0644	
rs17088752	68747036	UBA6-AS1	0.8609	-0.1508	-0.1282	0.0236	0.1898	-0.0605	0.051
rs10033058	69177408	YTHDC1	0.2919			-0.0128			0.061
rs11249453	70488786	UGT2A1	-0.4048	0.18	-0.2025	0.286		0.0446	0.0205
rs11249477	70934425	CSN1S2AP	-0.4281		-0.0995		-0.1076	0.3327	0.0075
rs11249478	70934473	CSN1S2AP	0.6009	-0.2061	-0.0859	0.2153	0.0246	-0.0719	0.1284
rs13102142	70951325	CSN1S2AP	0.6451			-0.4601		0.2998	0.3027
rs4563469	71656101	RUFY3	-0.6335	0.1512	0.0161	0.036			
rs13119998	71781186	MOB1B	0.6384	-0.3633					-0.5033
rs6857348	72388965	SLC4A4	0.0165						
rs2366715	73854103	RNU4ATAC9P	0.0749						
rs11938646	74537705	AC112518.3	-0.5411	-0.1758			0.018	-0.2986	
rs16850160	74758993	CXCL1	-0.5147			0.1367			-0.1084
rs12649753	74940765	RP11-629B11.4	0.3374			-0.0187		0.0226	0.1663
rs4610413	75555658	AC142293.3	-0.4844			0.2083			0.0554
rs10022187	75688708	BTC	0.5642		0.3142	-0.0995			
rs6817948	75762800	BTC	0.5838			-0.2409	0.1108		0.0773
rs17000199	75997311	RP11-44F21.5	-0.4003						
rs7677150	76644109	G3BP2	-0.4625					-0.0381	
rs10028707	77671439	SHROOM3	-0.3461		0.1419				0.0152
rs12500486	77754619	AC104687.1	0.5001	-0.1349				0.0073	0.1066
rs4150060	78080963	CCNG2	0.5117		0.0686				
rs7697415	78249882	CCNG2	-0.5616	-0.0257	0.3309	0.0751			
rs6827052	79949076	LINC01088	0.5551		0.0218	-0.3137		-0.3029	
rs17003679	80155510	LINC01088	-0.4822						
rs10857104	80575728	RP11-452C8.1	0.2978		-0.0242	-0.2286			0.1105
rs2867702	81060836	RP11-162K6.1	0.4913	-0.0498	0.0047	-0.0787			

Continued on the next page

Web Table 9: Continued from the previous page

SNP	Position	Gene		age	famdb	act	trans	ceraf	heme
rs10004901	81897005	C4orf22	-0.5762	0.0417	-0.115	0.0029		0.2165	
rs35036928	82671170	RP11-689K5.3	-0.3807	-0.0929		-0.1247			-0.0184
rs1993798	82762741	RP11-689K5.3	0.2292	-0.0373				-0.1176	
rs10516697	84618260	RP11-767N15.1	-0.4875				0.078		
rs17007216	84636196	RP11-767N15.1	-0.5071						
rs17007783	84881829	RP11-8L2.1	0.4986	0.0088	0.1229	0.1573	0.397		
rs392112	86244117	RP11-218C23.1	-0.3109	0.0191	-0.0036		-0.0963		
rs4334746	87072464	RP11-778J15.1	-0.4234	-0.0254	-0.0671		-0.0963	0.0895	-0.055
rs17409687	87250436	MAPK10	-0.7202	0.0177	0.0177		0.2061		-0.1671
rs13147739	87759673	SLC10A6	0.6486	0.3326	0.1686	0.0854		-0.2256	
rs7668684	88159940	KLHL8	0.656	0.102	0.0596		-0.4326	-0.0908	
rs9995093	89220944	RP11-10L7.1	0.5144		0.0881			0.0188	0.0252
rs1552972	90473540	RP11-115D19.1	0.1628						
rs10516839	90508340	RP11-115D19.1	0.4203			-0.079			
rs6841431	91116156	CCSER1	-0.5755	0.1587	-0.2974		-0.2365	-0.0252	
rs7667572	92053582	CCSER1	-0.5232			0.2386	-0.0139	0.0545	
rs17019545	93468433	GRID2	0.439				0.0379		
rs10856909	95268748	HPGDS	0.0401						
rs997464	96273480	UNC5C	-0.0134						
rs16996519	96759448	PDHA2	-0.4671	0.0639		-0.0949	-0.0433	0.0017	
rs10016725	98363606	RP11-681L8.1	-0.5328	0.0703		0.0998	-0.0511		
rs13149070	99479684	TSPAN5	-0.0431						
rs9998528	99692618	BTF3P13	0.0283						
rs7375429	100145719	RP11-696N14.1	0.4196						
rs4306962	101228493	EMCN	-0.3764	0.0453	-0.0462	0.0352	0.1218	0.0035	-0.1781
rs12645499	102611757	BANK1	0.6012				-0.1083	-0.2208	0.0816
rs2726516	106346206	PPA2	0.0142						
rs17261094	106557429	ARHGEF38	-0.4555				-0.2091	0.0644	
rs7690115	107448046	RP13-612N21.1	0.0788						
rs17357756	107455176	RP13-612N21.1	0.5251	-0.0414	-0.2133	-0.0481		0.0566	
rs17036882	107541582	RP13-612N21.1	0.5604		-0.0895			-0.246	
rs7669708	108202333	RP11-713M6.2	0.6216		-0.0686				
rs7661968	110866582	EGF	-0.7588			-0.2884	-0.069	-0.1115	

Web Table 10: Analysis of the GENEVA diabetes data (NHS/HPFS) using SMCP: identified main effects and interactions.

SNP	Location	Gene		age	famdb	act	trans	ceraf	heme
				-0.2785	0.2019	-0.4544	0.2361	-0.3411	0.4440
rs1507131	60564866	RP11-525J21.1			-0.0005				
rs2045634	60847314	Y_RNA							0.0004
rs1460346	60852284	Y_RNA							0.0003
rs9996745	61115698	Y_RNA		-0.0003					
rs10021002	61129136	AC095061.1		-0.0009					
rs41318750	61219473	AC095061.1		-0.0003					
rs10019682	61228770	AC095061.1		-0.0029					
rs11946495	61238161	AC095061.1		-0.0059					
rs4318680	61249881	AC095061.1		-0.0045					
rs10517495	61270978	AC095061.1		-0.0003					
rs1511104	61287186	AC095061.1		-0.0008					
rs1546511	61327179	AC095061.1		-0.0007					
rs10021608	61352529	AC095061.1		-0.0008					
rs10780045	61365594	AC095061.1		0.0006					
rs17090278	61679934	RP11-593F5.2	-0.0002						
rs17090286	61695978	RP11-593F5.2	-0.0002						
rs13122165	61762061	RP11-593F5.2	-0.0004						
rs1378365	61775645	RP11-593F5.2		0.0001					
rs12642037	62503837	LPHN3						0.0002	
rs4860422	62524423	LPHN3						0.0008	
rs11131340	62535059	LPHN3						0.0015	
rs2036199	62540857	LPHN3						0.0009	
rs996208	62559531	LPHN3						0.0001	
rs17226412	62907023	LPHN3					0.0002		
rs7681041	62982567	RP11-84A1.3						0.0006	
rs7681041	62982567	RP11-84A1.3						0.0007	
rs950313	63025823	RP11-84A1.3						0.0001	
rs1124974	63037324	RP11-84A1.3						0.0003	
rs9995712	63062142	RP11-84A1.3						0.0003	
rs778937	63242346	RP11-30P21.2						0.0006	
rs2604592	63312469	HMGN1P11						0.0005	
rs778567	63312966	HMGN1P11						0.0004	
rs6818094	64726856	RP11-12K22.1			-0.0002				
rs1878564	64791788	TECRL			-0.0003				

Continued on the next page

Web Table 10: Continued from the previous page

SNP	Position	Gene	age	famdb	act	trans	ceraf	heme
rs10002424	65744523	RP11-707A18.1						-0.0003
rs10028673	65748015	RP11-707A18.1						-0.0005
rs11721709	65775277	RP11-707A18.1						-0.0004
rs10005825	65778213	RP11-707A18.1						-0.0005
rs1376412	66267015	EPHA5		-0.0006				
rs17086181	66267300	EPHA5		-0.0007				
rs6551926	66292123	EPHA5		0.0013				
rs7659865	66292306	EPHA5		0.0012				
rs7667410	66486986	EPHA5				0.0009		
rs4422467	66494645	EPHA5				0.001		
rs7659227	66496170	EPHA5				0.0008		
rs4370201	66500571	EPHA5				0.0003		
rs11131612	66525024	EPHA5				0.0002		
rs13107026	67139034	MIR1269A	0.0001					
rs1397755	67140834	MIR1269A	0.0003					
rs12331987	67188980	MIR1269A	0.002					
rs10000219	67200024	MIR1269A	0.005					
rs4860208	67201368	MIR1269A	0.0063					
rs1511286	67213473	MIR1269A	0.0041					
rs11936928	67489994	RPS23P3	0.0014					
rs6838523	67494918	RPS23P3	0.0014					
rs920482	68241081	RP11-584P21.2			0.0001			
rs6419917	68241545	RP11-584P21.2			0.0003			
rs1348079	68718552	UBA6-AS1		-0.0004				
rs9312190	68718911	UBA6-AS1		-0.0007				
rs10033058	69177408	YTHDC1	0.0011					
rs2293595	69178920	YTHDC1	0.0011					
rs17089267	69183791	YTHDC1	0.0002					
rs2013562	70354575	UGT2B4	-0.0002					
rs2046911	70848044	STATH		0.0004				
rs776822	70861285	STATH		0.0005				
rs1842478	71227802	SMR3A						-0.0013
rs6446910	73833081	RNU4ATAC9P	0.0002					
rs1957659	74472081	RASSF6		0.0004				
rs17805665	74472268	RASSF6		0.0005				

Continued on the next page

Web Table 10: Continued from the previous page

SNP	Position	Gene	age	famdb	act	trans	ceraf	heme
rs2010989	74472896	RASSF6		0.0003				
rs12649753	74940765	RP11-629B11.4	0.0003					
rs12641287	75259691	EREG		0.0001				
rs16996019	76644744	G3BP2					0.0002	
rs17000786	76667709	USO1					0.0002	
rs11938067	76669538	USO1					0.0001	
rs13110602	76704860	USO1					0.0001	
rs324735	76705014	USO1					0.0002	
rs12506745	77220837	FAM47E-STBD1	-0.0005					
rs1376412	66267015	EPHA5		-0.0006				
rs17086181	66267300	EPHA5		-0.0007				
rs6551926	66292123	EPHA5		0.0013				
rs7659865	66292306	EPHA5		0.0012				
rs7667410	66486986	EPHA5				0.0009		
rs4422467	66494645	EPHA5				0.001		
rs7659227	66496170	EPHA5				0.0008		
rs4370201	66500571	EPHA5				0.0003		
rs11131612	66525024	EPHA5				0.0002		
rs13107026	67139034	MIR1269A	0.0001					
rs1397755	67140834	MIR1269A	0.0003					
rs12331987	67188980	MIR1269A	0.002					
rs10000219	67200024	MIR1269A	0.005					
rs4860208	67201368	MIR1269A	0.0063					
rs1511286	67213473	MIR1269A	0.0041					
rs11936928	67489994	RPS23P3	0.0014					
rs6838523	67494918	RPS23P3	0.0014					
rs920482	68241081	RP11-584P21.2			0.0001			
rs6419917	68241545	RP11-584P21.2			0.0003			
rs1348079	68718552	UBA6-AS1		-0.0004				
rs9312190	68718911	UBA6-AS1		-0.0007				
rs10033058	69177408	YTHDC1	0.0011					
rs2293595	69178920	YTHDC1	0.0011					
rs17089267	69183791	YTHDC1	0.0002					
rs2013562	70354575	UGT2B4	-0.0002					
rs2046911	70848044	STATH		0.0004				
rs776822	70861285	STATH		0.0005				

Continued on the next page

Web Table 10: Continued from the previous page

SNP	Position	Gene		age	famdb	act	trans	ceraf	heme
rs1842478	71227802	SMR3A							-0.0013
rs6446910	73833081	RNU4ATAC9P	0.0002						
rs1957659	74472081	RASSF6			0.0004				
rs17805665	74472268	RASSF6			0.0005				
rs2010989	74472896	RASSF6			0.0003				
rs12649753	74940765	RP11-629B11.4	0.0003						
rs12641287	75259691	EREG			0.0001				
rs16996019	76644744	G3BP2						0.0002	
rs17000786	76667709	USO1						0.0002	
rs11938067	76669538	USO1						0.0001	
rs13110602	76704860	USO1						0.0001	
rs324735	76705014	USO1						0.0002	
rs12506745	77220837	FAM47E-STBD1		-0.0005					
rs907446	77254804	CCDC158		-0.0002					
rs6857452	77317124	CCDC158		-0.0002					
rs6853053	77489645	SHROOM3							-0.0007
rs6824297	77492155	SHROOM3							-0.0009
rs13148163	77504400	SHROOM3							-0.0012
rs2645645	77859067	11-Sep					-0.0004		
rs2703141	77860806	11-Sep					-0.0003		
rs4272042	78053975	RNU6-1187P			0.0007				
rs6831175	78389293	RP11-625I7.1					0.0001		
rs1051225463	79458537	FRAS1		-0.0004					
rs7694246	80686562	PCAT4							-0.0005
rs4285108	80688257	PCAT4							-0.0006
rs6534560	80688396	PCAT4							-0.0005
rs7682543	81154975	FGF5					0.0003		
rs11937407	81413618	C4orf22			-0.0005				
rs6858262	81441388	C4orf22			-0.001				
rs6534992	81441764	C4orf22			-0.0009				
rs17004924	81445307	C4orf22			-0.0004				
rs1843563	81454827	C4orf22			-0.0004				
rs1391262	82663506	RP11-689K5.3	-0.0004						
rs35036928	82671170	RP11-689K5.3	-0.0023						
rs4693369	82671234	RP11-689K5.3	-0.0022						

Continued on the next page

Web Table 10: Continued from the previous page

SNP	Position	Gene		age	famdb	act	trans	ceraf heme
rs7672440	82671938	RP11-689K5.3	-0.0004					
rs676592	82733530	RP11-689K5.3	-0.0014					
rs1993798	82762741	RP11-689K5.3	0.0012					
rs2868257	82762839	RP11-689K5.3	0.002					
rs6535281	82763010	RP11-689K5.3	0.0013					
rs11722328	82818968	RP11-689K5.3	0.0003					
rs2199487	82820008	RP11-689K5.3	0.0003					
rs2035911	83115185	RNU6-499P				-0.0003		
rs17359809	84554429	Y_RNA		0.0002				
rs4693655	84820059	RP11-8L2.1					0.0003	
rs4693656	84820334	RP11-8L2.1					0.0005	
rs4272003	84822518	RP11-8L2.1					0.0004	
rs1838039	85625701	WDFY3					0.0004	
rs6857037	86069378	RP11-218C23.1					-0.0002	
rs4485812	86101088	RP11-218C23.1					-0.0004	
rs6822829	86101771	RP11-218C23.1					0.0002	
rs7656038	86143981	RP11-218C23.1					-0.0002	
rs6531805	86153851	RP11-218C23.1					-0.0003	
rs13115311	86156826	RP11-218C23.1					-0.0002	
rs392112	86244117	RP11-218C23.1	-0.0004					
rs433657	86252435	RP11-218C23.1	-0.0007					
rs434193	86253489	RP11-218C23.1	-0.001					
rs6842681	86253994	RP11-218C23.1	-0.0011					
rs425196	86255297	RP11-218C23.1	-0.0012					
rs416035	86255366	RP11-218C23.1	-0.0014					
rs432755	86255399	RP11-218C23.1	-0.0014					
rs375432	86255845	RP11-218C23.1	-0.0011					
rs425642	86255997	RP11-218C23.1	-0.0008					
rs407430	86256356	RP11-218C23.1	-0.0008					
rs400023	86256538	RP11-218C23.1	-0.0008					
rs585787	86257453	RP11-218C23.1	-0.0004					
rs340199	86350683	ARHGAP24				-0.0009		
rs340200	86351156	ARHGAP24				-0.0017		
rs340202	86353796	ARHGAP24				-0.0014		
rs340203	86353867	ARHGAP24				-0.001		
rs10222732	88295008	HSD17B11		-0.0003				
rs10017282	88347368	NUDT9		0.0002				

Continued on the next page

Web Table 10: Continued from the previous page

SNP	Position	Gene	age	famdb	act	trans	ceraf	heme
rs7656367	89196821	PPM1K						0.0008
rs893971	89203670	RP11-10L7.1						0.0005
rs9994576	91822311	CCSER1						0.0003
rs13119689	94719474	RNA5SP164				-0.0009		
rs10031584	94719872	RNA5SP164				-0.001		
rs2632401	95147055	SMARCAD1				-0.0001		
rs3775051	96126206	UNC5C			-0.0001			
rs1369985	96545816	RPL30P6	0.0001					
rs1816849	97122531	RNU6-34P					-0.0006	
rs1823757	97122591	RNU6-34P					-0.0012	
rs1395199	97177286	RP11-145G20.1					0.0008	
rs6532609	97421141	RP11-145G20.1					0.0007	
rs10022608	97444956	RP11-145G20.1					0.001	
rs6832719	97458376	RP11-145G20.1					0.0003	
rs4596245	97797275	COX7A2P2			-0.0001			
rs6841643	97806589	COX7A2P2			-0.0001			
rs7685402	99660764	BTF3P13		0.0007				
rs6532833	100747972	DAPP1					-0.0007	
rs6532833	100747972	DAPP1					-0.0007	
rs3822103	100754665	DAPP1				-0.0004		
rs1348161	102252696	MIR1255A		0.0002				
rs2850329	102275264	MIR1255A		0.0008				
rs6845263	102276682	MIR1255A		0.0013				
rs10433982	102283898	MIR1255A		0.0007				
rs6845368	102285899	MIR1255A		0.0002				
rs6846097	102286050	MIR1255A		0.0002				
rs7377083	102708997	BANK1	-0.0003					
rs4615176	102737936	BANK1				-0.0009		
rs4411998	102738147	BANK1				-0.0015		
rs4276281	102746780	BANK1				-0.0011		
rs11735227	103125383	SLC39A8				-0.0005		
rs151414	103137879	SLC39A8						0.0004
rs151413	103137941	SLC39A8						0.0006

Continued on the next page

Web Table 10: Continued from the previous page

SNP	Position	Gene		age	famdb	act	trans ceraf	heme
rs1540052	103137977	SLC39A8						0.0004
rs9997118	105898577	RNU6-351P		0.0006				
rs902444	105923047	RP11-556I14.1		0.0004				
rs2636696	106297308	PPA2			-0.0004			
rs2713834	106310520	PPA2			-0.0003			
rs2636732	106336855	PPA2			-0.0002			
rs2726516	106346206	PPA2	0.0008		-0.0011			
rs2636739	106352105	PPA2	0.0008		-0.0011			
rs2636743	106356581	PPA2			-0.0004			
rs2686293	107511296	RP13-612N21.1	-0.0007					
rs7671347	108792735	RP11-286E11.1				-0.0001		
rs17509519	109192790	LEF1-AS1				0.0001		
rs3851415	109194387	LEF1-AS1				0.0002		
rs1027681	109195380	LEF1-AS1				0.0002		
rs1004472	109215088	LEF1-AS1						0.004
rs2078126	109215175	LEF1-AS1						0.0073
rs6831553	109218208	LEF1-AS1						0.0075
rs7699097	109227508	LEF1-AS1						0.0045
rs16996944	109232275	LEF1-AS1						0.0009
rs220630	109232304	LEF1-AS1						0.0009
rs11726135	109232492	LEF1-AS1						0.0009
rs220615	109239451	LEF1-AS1						0.002
rs3932216	109242863	LEF1-AS1						0.002
rs11097979	109248691	LEF1-AS1						0.0006
rs11938222	110555678	CCDC109B						0.0002

Web Table 11: Analysis of the TCGA SKCM data using MA: identified main effects and interactions

Gene		Age	PN	Gender	Breslow's depth	Clark level
		0.0019	-0.0331	3.3675	0.0744	-0.0098
ACTA1	-0.0167					
ANKRD30A	-0.1554					
BCYRN1	0.0163		0.0247			
C2ORF83	-0.0251		0.0307	-0.0059		
C3ORF79	-0.9345					
CFC1B	-0.2176					
CT47A6	1.5287					
DMRTB1	-0.8216					
DPPA3	0.0363					
FABP12	-0.0525			-0.0243		
FANCD2OS	-0.1159					
FTHL17	-0.14					
GSX1	0.2476	0.4562		-0.1929		0.0096
HES3	-10.2916			12.7309		
KCNK18	0.0573					
OR4X2	-0.0459		0.0152			
OR5B2			0.0648			0.0562
OR8B12	-0.0085					
PIAS3	-0.0258					
RGS21	-0.1029				-0.1823	
RNASE11	0.3474	-0.2649	0.2106	0.1886	0.8826	
SLC17A6	0.0685					
SPATA31D3	-0.0026		0.0089			
TAS2R7	0.0481	-0.037				-0.0356
TMEM30CP	0.0446					
TMIGD1	-0.0422		-0.0021	0.037		
TMPRSS11BNL	0.2116					
VTA1	0.0431					

Web Table 12: Analysis of the TCGA SKCM data using HierMCP: identified main effects and interactions

Gene		Age	PN	Gender	Breslow's depth	Clark level
		-0.3087	-0.4998	-0.1149	-0.1555	-0.067
ABCC13	-0.4337					
ADGRF2	0.532					
APBA2	-0.0472					
APC	-0.0702					
AQP10	-3.2088			0.0446		0.2722
ARHGAP5	0.0358					
ATF2	0.0206					
ATP6V0A4	0.1658					
AUNIP	-0.0846					
BMI1	-0.0546					
BRCC3	-0.0459					
BRINP3	0.3568					0.0364
C1QL4	-0.0327					
CES3	0.1063	0.0046				
CFC1B	0.0567					
CGB7	0.0412					
CHD9	-0.0406					
CLCA1	1.2615					
CLDN20	-0.4075	-0.057				
CLIC4	0.0471					
CT47A6	-0.644					
CTXN3	-0.7147					
DGLUCY	0.0498					
DLGAP3	0.2987					
DMRTB1	-0.3941					
EHMT1	-0.1286	-0.0028				
ENTPD3	0.603					
ENTPD5	-0.0919	-0.0162	-0.0019			
EVA1B	-0.1816					
FDX1	0.2654		-0.0055	0.0084		
FERMT2	0.0896		-0.0066			
FGF2	-0.1431					
FHL5	-0.0491					
FNDC3A	-0.0964					

Continued on the next page

Web Table 12: Continued from the previous page

Gene		Age	PN	Gender	Breslow's depth	Clark level
GLB1L3	-0.4256					0.001
GPR15	-0.1709					
GTF2A1	0.0756					
GTF3C5	0.0442					
GUCA1A	0.0364					
HADHA	-0.0758					
INSR	0.074					
ISM2	-0.308					
KRBA1	0.0307					
KTN1	0.107	0.0215				
LGI3	-0.0563					
LINC00265	-0.0597		-0.0023			
LRFN2	0.2679					
LRP12	0.0397					
LYPLA2	-0.107	0.0053	-0.0195	0.0021		
MAJIN	0.3281					
MAP2K4	-0.057			-0.0074		
MARF1	0.0588					
MIA3	0.066					
MTG1	-0.1014					
MTO1	0.0831					
MTR	0.0853		-0.0036	0.0222		
MYL2	0.3822					
NEK7	-0.037					
NFAT5	-0.0695					0.017
NR3C1	-0.0935			-0.0032		
OR2T6	-0.1199					
OR5A1	-0.1817		0.0021		-0.0012	
OTX1	0.2702			0.0027		-0.0026
PAFAH1B2	-0.0497					
PCDHB16	0.0398					
PCDHB5	-0.1094					0.0087
PCGEM1	-1.0071		-0.7877			
PEX14	0.0982					
PHF3	0.0952		-0.0015	-0.0012		-0.001
PHOX2A	0.2075					
PJA2	0.0886					
PKD1L3	0.2358			-0.0198		
PMPCA	0.0374					
POGLUT1	-0.0382					

Continued on the next page

Web Table 12: Continued from the previous page

Gene		Age	PN	Gender	Breslow's depth	Clark level
PPIAL4G	-0.104			0.0061		-0.0126
PPM1A	0.0605					
PPY	0.1489					
PTPN11	-0.1561	-0.0281				
RAB18	-0.1036				0.0033	
RBPMS2	0.2032	-0.006	-0.035			0.0148
REG1A	0.111					
RFPL1S	-0.1125	0.0303			-0.0169	0.0019
RGS14	0.0724					
RIC8B	0.1271		-0.0079			0.0249
RNASE7	-0.0012					
ROCK1	-0.071	0.0078				
RORB	-0.0434					
RSRC1	-0.1321	0.0458		-0.0038		
RWDD2A	0.0929		0.0212			
SEC61B	0.1049					
SEC62	-0.0759			-0.0072		
SERPINB10	-0.1255					
SERPINB2	-0.1047					
SH3TC2	0.0919	-0.0192				
SLC22A1	1.151					
SLC22A13	0.6331					
SLC2A13	0.0385					
SLC37A4	0.1188					
SLU7	-0.1255		-0.0027			
SNAPC1	-0.0694					
SPATS1	-0.1707	-0.0022	-0.0085			-0.0042
SPRYD4	-0.0799					
TBC1D19	0.0532					
TCEA3	-0.1186			-0.0014		-0.0025
TCTA	-0.116			0.0143		-0.0092
TENT2	-0.095					0.0453
TEX13B	0.2882					
TFDP3	-0.4634			0.0683		
TLX1	-0.4839					
TLX1NB	1.2187					
TMEM184C	-0.1015	0.0029		-0.0057		-0.0024

Continued on the next page

Web Table 12: Continued from the previous page

Gene		Age	PN	Gender	Breslow's depth	Clark level
TMEM223	-0.6452					
TMPRSS3	-0.4446					
TMPRSS9	-0.5598					
TMX1	0.0931					
TRIM68	0.0481					
TSC22D2	-0.0711					
TULP1	0.108					
UBE2V1	-0.0379					
UCP1	-1.6798					
UPRT	0.1142		0.032	-0.0438	0.0166	
USP12	0.0891	0.0272				
VWA2	2.4251	-0.1801				0.02
WBP1	-0.125					
WDR7	-0.1281	0.0117		-0.0031	-0.0072	
WFDC2	-0.0802					
WNT6	-0.2608	0.0707		-0.0213		
ZBTB1	0.0563					
ZNF25	0.3841	-0.0495	-0.0181	0.0469	0.0421	
ZNF335	-0.0394					

Web Table 13: Analysis of the TCGA SKCM data using SMCP: identified main effects and interactions

Gene		Age	PN	Gender	Breslow's depth	Clark level
		-0.1093	-0.3384	-0.0212	-0.0644	-0.1736
ACTL6B		0.0027			-0.0011	
BARHL2		-0.001				
C10ORF99		0.0013				
CHRNA4	-0.0014					
CLCA2	-0.0015					
COX8C		-0.001				
CSPG4	0.001					
CSTF1	-0.0013					
CYP26C1				-0.001		
DHX35	-0.0012					
DSG3	-0.0021					
ERGIC3	-0.001					
FAM131B				0.0011		
FAM19A4						0.0011
FBXO22	0.0011					
GNGT1			0.0012			
HIST2H2BA				0.0011		
HIST2H2BF				0.0013		
IL17A	0.0024		-0.0015	-0.0011		
IL17F			-0.0015	-0.0019		
INTS4	-0.0013					
JARID2	-0.0015					
KCTD21	-0.0011					
KDELR2	-0.0011					
KLHL12	-0.0015					
LYNX1	-0.0012					
METTL21C	0.0013		-0.0012	0.0016		
MRM2	-0.001					
NKAIN2	-0.0015	0.001	0.0013	0.0011		
NOV	0.0011			-0.0012		
OR5C1		-0.001				
OR5L2				-0.001		
PRSS3	-0.0011					0.0018
RAC1	-0.0014					
RORB		-0.0011				

Continued on the next page

Web Table 13: Continued from the previous page

Gene		Age	PN	Gender	Breslow's depth	Clark level
SERPINB13	-0.0029					
SERPINB3	-0.0014					
SETDB2	-0.0015	0.001	0.0012	0.0018		
SKIDA1					-0.0011	
SLC18A3	-0.0014	0.0013		0.0014		
SMG7	-0.001					
SMIM21		-0.0011				
STMN4	0.0015					
SYT12						-0.0011
TAS2R1	-0.0011				-0.0021	
TMEM266	0.0013					
TPD52L2	-0.0011					
TRIM46	-0.0012					
TRPC4AP	-0.0011					
UBAP2L	-0.0011					
UBE2Q1	-0.0011					
UBE2V1	-0.0013					
UFC1	-0.0011					
WDR77	-0.001					
YWHAB	-0.0011					
ZC3HC1	-0.0012					

References

Breheny, P. and Huang, J. (2009). Penalized methods for bi-level variable selection. *Statistics and its Interface* **2**, 369-380.

Breheny, P. and Huang, J. (2011). Coordinate descent algorithms for nonconvex penalized regression, with applications to biological feature selection. *The Annals of Applied Statistics* **5**, 232.

Chakrabarty, K., Tauber, J.W., Sigel, B., Bombeck, C.T., and Jeffay, H. (1984). Glycerokinase activity in human adipose tissue as related to obesity. *International Journal of Obesity* **8**, 609-622.

Choi, N. H., Li, W., and Zhu, J. (2010). Variable selection with the strong heredity constraint and its oracle property. *Journal of the American Statistical Association* **105**, 354-364.

Cin, H., Meyer, C., Herr, R., Janzarik, W.G., Lambert, S., Jones, et al. (2011). Oncogenic FAM131B-CBRAF fusion resulting from 7q34 deletion comprises an alternative mechanism of MAPK pathway activation in pilocytic astrocytoma. *Acta Neuropathologica* **121**, 763-774.

Cui, Y., Kang, G., Sun, K., Qian, M., Romero, R., and Fu, W. (2008). Gene-centric genomewide association study via entropy. *Genetics* **179**, 637-650.

Fan, J and Lv, J. (2011). Nonconcave penalized likelihood with NP-dimensionality. *IEEE Transactions on Information Theory* **57**, 5467-5484.

Fecher, L.A., Amaravadi, R.K., and Flaherty, K.T. (2008). The MAPK pathway in melanoma. *Current Opinion in Oncology* **20**, 183-189.

Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software* **33**, 1-22.

Guo, J., Hu, J., Jing, B. Y., and Zhang, Z. (2016). Spline-lasso in high-dimensional linear regression. *Journal of the American Statistical Association* **111**, 288-297.

Huang, J., Ma, S., Li, H., and Zhang, C. H. (2011). The sparse Laplacian shrinkage estimator for high-dimensional regression. *Annals of Statistics* **39**, 2021-2046.

Huang, Y., Zhang, Q., Zhang, S., Huang, J., and Ma, S. (2017). Promoting similarity of sparsity structures in integrative analysis with penalization. *Journal of the American Statistical Association* **112**, 342-350.

Hunt, S.C., Hasstedt, S.J., Xin, Y., Dalley, B. K., Milash, B. A., Yakobson, E., et al. (2011). Polymorphisms in the NPY2R gene show significant associations with BMI that are additive to FTO, MC4R, and NPFFR2 gene effects. *Obesity* **19**, 2241-2247.

Jung, I. K., Lee, J. S., Choi, C. H., and Kim, D. S. (2011). Inferring relative activity between pathway and downstream genes to classify melanoma cancer progression. *Interdisciplinary Bio Central* **3**, 1-5.

Krauthammer, M., Kong, Y., Ha, B. H., Evans, P., Bacchiocchi, A., McCusker, J. P., et al. (2012). Exome sequencing identifies recurrent somatic RAC1 mutations in melanoma. *Nature Genetics* **44**, 1006-1014.

Liu, J., Huang, J., Zhang, Y., Lan, Q., Rothman, N., Zheng, T., et al. (2013). Identification of gene-environment interactions in cancer studies using penalization. *Genomics* **102**, 189-194.

Liu, H., Wang, X., Feng, B., Tang, L., Li, W., Zheng, X., et al. (2018). Golgi phosphoprotein 3 (GOLPH3) promotes hepatocellular carcinoma progression by activating mTOR signaling pathway. *BMC Cancer* **18**, 661.

Rouault, C., Pellegrinelli, V., Schilch, R., Cotillard, A., Poitou, C., Tordjman, J., et al. (2013).

Roles of chemokine ligand-2 (CXCL2) and neutrophils in influencing endothelial cell function and inflammation of human adipose tissue. *Endocrinology* **154**, 1069-1079.

Scott, K. L., Kabbarah, O., Liang, M. C., Ivanova, E., Anagnostou, V., Wu, J., et al. (2009). GOLPH3 modulates mTOR signalling and rapamycin sensitivity in cancer. *Nature 459*, 1085-1090.

Song, M.H., Ha, J.C., Lee, S.M., Park, Y.M., and Lee, S. Y. (2011). Identification of BCP-20 (FBXO39) as a cancer/testis antigen from colon cancer patients by SEREX. *Biochemical and Biophysical Research Communications* **408**, 195-201.

Tseng, P. (2001). Convergence of a block coordinate descent method for nondifferentiable minimization. *Journal of Optimization Theory and Applications* **109**, 475-494.

Vachtenheim, J., Ondrusova, L., and Borovansky, J. (2010). SWI/SNF chromatin remodeling complex is critical for the expression of microphthalmia-associated transcription factor in melanoma cells. *Biochemical and Biophysical Research Communications* **392**, 454-459.

Wang, L., Yi, T., Kortylewski, M., Pardoll, D. M., Zeng, D., and Yu, H. (2009). IL-17 can promote tumor growth through an IL-6-CStat3 signaling pathway. *Journal of Experimental Medicine* **206**, 1457-1464.

Zhang, C. (2011). Nearly unbiased variable selection under minimax concave penalty. *Annals of Statistics* **38**, 894-942.

Zou, H., and Zhang, H. H. (2009). On the adaptive elastic-net with a diverging number of parameters. *Annals of Statistics* **37**, 1733-1751.