Supplemental Materials to "Scalable Bayesian Variable Selection Using Nonlocal Prior Densities in Ultrahigh-dimensional Settings"

1 Preliminary Results

Lemma 1. *For $Q_{\mathbf{k}}$ defined in* (6), $\prod_{j=1}^{k} Q_{\mathbf{k},j}^{L} \leq Q_{\mathbf{k}} \leq \prod_{j=1}^{k} Q_{\mathbf{k},j}^{U}$,

where

$$\begin{aligned} Q_{\mathbf{k},j}^{L} &= c_1(\sigma^2)^{1/2}(n\nu_{\mathbf{k}}^{*} + 1/\tau_{n,p})^{-1/2}\exp\{-\tau_{n,p}/\widetilde{\beta}_{\mathbf{k},j}^{*2}\}, \\ Q_{\mathbf{k},j}^{U} &= c_2(\sigma^2)^{1/2}(n\nu_{\mathbf{k}*} + 1/\tau_{n,p})^{-1/2}\exp\{-\tau_{n,p}/(|\widetilde{\beta}_{\mathbf{k},j}| + \widetilde{\epsilon}_n)^2\}, \end{aligned}$$

and $\widetilde{\epsilon}_n \asymp (n\nu_{\mathbf{k}*}/\tau_{n,p})^{-1/4}$, *with* $\widetilde{\beta}_{\mathbf{k},j}^{*} \in [\widetilde{\beta}_{\mathbf{k},j} - \widetilde{\epsilon}_n, \widetilde{\beta}_{\mathbf{k},j} + \widetilde{\epsilon}_n] \setminus (-\widetilde{\epsilon}_n, \widetilde{\epsilon}_n)^c$ *for some positive constants* c_1 *and* c_2.

Proof. Recall $\widetilde{\Sigma}_{\mathbf{k}} = (X_{\mathbf{k}}^T X_{\mathbf{k}} + 1/\tau_{n,p}\mathrm{I}_{\mathbf{k}})^{-1}$. From (8), all eigenvalues of $(\widetilde{\Sigma}_{\mathbf{k}})^{-1}$ are bounded between $n\nu_{\mathbf{k}*} + 1/\tau_{n,p}$ and $n\nu_{\mathbf{k}}^{*} + 1/\tau_{n,p}$, which implies for all $x \in \mathbb{R}^{|\mathbf{k}|}$, $(n\nu_{\mathbf{k}*} + 1/\tau_{n,p})x^T x \leq x^T(\widetilde{\Sigma}_{\mathbf{k}})^{-1}x \leq (n\nu_{\mathbf{k}}^{*} + 1/\tau_{n,p})x^T x$. Let $T_{1n} = \{(n\nu_{\mathbf{k}}^{*} + 1/\tau_{n,p})/\sigma^2\}^{1/2}$ and $T_{2n} = \{(n\nu_{\mathbf{k}*} + 1/\tau_{n,p})/\sigma^2\}^{1/2}$. Substituting the above inequality in the expression for $Q_{\mathbf{k}}$, we have

$$\prod_{j=1}^{|\mathbf{k}|} g_1(\widetilde{\beta}_{\mathbf{k},j}) \leq Q_{\mathbf{k}} \leq \prod_{j=1}^{|\mathbf{k}|} g_2(\widetilde{\beta}_{\mathbf{k},j}), \tag{S1}$$

where

$$g_i(\widetilde{\beta}_{\mathbf{k},j}) = \int_{-\infty}^{\infty} \exp\{-T_{in}^2(\beta_{\mathbf{k},j} - \widetilde{\beta}_{\mathbf{k},j})^2/2 - \tau_{n,p}/\beta_{\mathbf{k},j}^2\}d\beta_{\mathbf{k},j}, \tag{S2}$$

for $i = 1, 2$. We establish the lower bound first by showing that $g_1(\widetilde{\beta}_{\mathbf{k},j}) \geq Q_{\mathbf{k},j}^{L}$ for all $j =$

$1, \ldots, |\mathbf{k}|$. Recall $\widetilde{\epsilon}_n \asymp (n\nu_{\mathbf{k}*}/\tau_{n,p})^{-1/4}$ from the statement of the Lemma. We have

$$\begin{aligned} g_1(\widetilde{\beta}_{\mathbf{k},j}) &\geq \int_{[\widetilde{\beta}_{\mathbf{k},j}-\widetilde{\epsilon}_n, \widetilde{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n]\setminus(-\widetilde{\epsilon}_n,\widetilde{\epsilon}_n)^c} \exp\{-T_{1n}^2(\beta_{\mathbf{k},j}-\widetilde{\beta}_{\mathbf{k},j})^2/2 - \tau_{n,p}/\beta_{\mathbf{k},j}^2\} d\beta_{\mathbf{k},j} \\ &\geq \exp\{-\tau_{n,p}/\widetilde{\beta}^{*2}_{\mathbf{k},j}\} \int_{[\widetilde{\beta}_{\mathbf{k},j}-\widetilde{\epsilon}_n, \widetilde{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n]\setminus(-\widetilde{\epsilon}_n,\widetilde{\epsilon}_n)^c} \exp\{-T_{1n}^2(\beta_{\mathbf{k},j}-\widetilde{\beta}_{\mathbf{k},j})^2/2\} d\beta_{\mathbf{k},j}, \end{aligned}$$

for some $\widetilde{\beta}^*_{\mathbf{k},j} \in [\widetilde{\beta}_{\mathbf{k},j} - \widetilde{\epsilon}_n, \widetilde{\beta}_{\mathbf{k},j} + \widetilde{\epsilon}_n] \setminus (-\widetilde{\epsilon}_n, \widetilde{\epsilon}_n)^c$. Then, the integral in the last line of the above display is equivalent to

$$\int_{[-\widetilde{\epsilon}_n,\widetilde{\epsilon}_n]\setminus(-\widetilde{\beta}_{\mathbf{k},j}-\widetilde{\epsilon}_n, -\widetilde{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n)^c} e^{-T_{1n}^2 t^2/2} dt \geq c_1 T_{1n}^{-1} \int_0^{T_{1n}\widetilde{\epsilon}_n} e^{-z^2/2} dz \geq c_2 T_{1n}^{-1},$$

where c_1 and c_2 are some positive constants and the last inequality in the above display follows since $T_{1n}\widetilde{\epsilon}_n \geq 1$ for large n. Substituting back in the previous display, $g_1(\widetilde{\beta}_{\mathbf{k,j}}) \geq c_1 T_{1n}^{-1} \exp\{-\tau_{n,p}/\widetilde{\beta}^{*2}_{\mathbf{k,j}}\}$ for some constant $c_1 > 0$, completing the proof of the lower bound.

We now establish the upper bound by showing that $g_2(\widetilde{\beta}_{\mathbf{k},j}) \leq Q^U_{\mathbf{k},j}$ for all $j = 1, \ldots, |\mathbf{k}|$. It is straightforward to see that g_2 is a symmetric function (i.e, $g_2(\widetilde{\beta}_{\mathbf{k},j}) = g_2(|\widetilde{\beta}_{\mathbf{k},j}|)$), so that it is enough to establish the bound for $\widetilde{\beta}_{\mathbf{k},j} > 0$; without loss of generality we assume that $\widetilde{\beta}_{\mathbf{k},j} > 0$. We have

$$\begin{aligned} & \int_{-\infty}^{\infty} \exp\{-T_{2n}^2(\beta_{\mathbf{k},j}-\widetilde{\beta}_{\mathbf{k},j})^2/2 - \tau_{n,p}\beta_{\mathbf{k},j}^2\} d\beta_{\mathbf{k},j} \\ = & \int_{-\infty}^{0} \exp\{-T_{2n}^2(\beta_{\mathbf{k},j}-\widetilde{\beta}_{\mathbf{k},j})^2/2 - \tau_{n,p}/\beta_{\mathbf{k},j}^2\} d\beta_{\mathbf{k},j} \\ + & \int_{0}^{\widetilde{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n} \exp\{-T_{2n}^2(\beta_{\mathbf{k},j}-\widetilde{\beta}_{\mathbf{k},j})^2/2 - \tau_{n,p}/\beta_{\mathbf{k},j}^2\} d\beta_{\mathbf{k},j} \\ + & \int_{\widetilde{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n}^{\infty} \exp\{-T_{2n}^2(\beta_{\mathbf{k},j}-\widetilde{\beta}_{\mathbf{k},j})^2/2 - \tau_{n,p}/\beta_{\mathbf{k},j}^2\} d\beta_{\mathbf{k},j}. \end{aligned}$$

Define the first term of the above as W_1, the second as W_2, and the third term as W_3. First, we shall show that $W_1 \leq c T_{2n}^{-1} \exp\{-T_{2n}(2\tau_{n,p})^{1/2}\}$ for some positive constant c. By transforming the variable $t = \beta_{\mathbf{k},j} - \widetilde{\beta}_{\mathbf{k},j}$,

$$
\begin{aligned}
W_1 &= \int_{-\infty}^{0} \exp\{-T_{2n}^2 t^2/2 + T_{2n}^2 t \widetilde{\beta}_{\mathbf{k},j} - T_{2n}^2 \widetilde{\beta}_{\mathbf{k},j}^2/2 - \tau_{n,p}/t^2\} dt \\
&\leq \int_{-\infty}^{0} \exp\{-T_{2n}^2 t^2/2 - \tau_{n,p}/t^2\} dt \\
&\leq c_3 T_{2n}^{-1} \exp\{-T_{2n}(2\tau_{n,p})^{1/2}\},
\end{aligned}
$$

for some constant c_3, since $\int \exp\{-\mu/t^2 - \zeta t^2\} dt = (\pi/\zeta)^{-1/2} \exp\{-2(\mu\zeta)^{1/2}\}$ for $\mu > 0$ and $\zeta > 0$.

Second, by changing the variable $z = t - \widetilde{\epsilon}$,

$$
\begin{aligned}
W_2 &= \int_{-\widetilde{\epsilon}_n}^{\widetilde{\beta}_{\mathbf{k},j}} \exp\{-T_{2n}^2 (z - \widetilde{\beta}_{\mathbf{k},j} + \widetilde{\epsilon}_n)^2/2 - \tau_{n,p}/(z + \widetilde{\epsilon}_n)^2\} dz \\
&\leq \exp\{-\tau_{n,p}/(\widetilde{\beta}_{\mathbf{k},j} + \widetilde{\epsilon}_n)^2\} \int_{-\infty}^{\infty} \exp\{-T_{2n}^2 (z - \widetilde{\beta}_{\mathbf{k},j} + \widetilde{\epsilon}_n)^2/2\} \\
&\leq c_4 T_{2n}^{-1} \exp\{-\tau_{n,p}/(\widetilde{\beta}_{\mathbf{k},j} + \widetilde{\epsilon}_n)^2\},
\end{aligned}
$$

for some positive constant c_4.

Third, by changing the variable $z = t - \widetilde{\beta}_{\mathbf{k},j}$, there exists some positive constant c such that

$$
\begin{aligned}
W_3 &= \int_{\widetilde{\epsilon}_n}^{\infty} \exp\{-T_{2n}^2 z^2/2 - \tau_{n,p}/(z + \widetilde{\beta}_{\mathbf{k},j})^2\} dz \\
&\leq \exp\{-T_{2n}^2 \widetilde{\epsilon}_n^2/4\} \int_{-\infty}^{\infty} \exp\{-T_{2n}^2 z^2/4\} dz \\
&\leq c_5 T_{2n}^{-1} \exp\{-c_6 T_{2n} \tau_{n,p}^{1/2}\},
\end{aligned}
$$

for some constants $c_5$and c_6, since $\widetilde{\epsilon}_n \asymp (n\nu_{\mathbf{k}*}/\tau_{n,p})^{-1/4}$. Then,

$$
\begin{aligned}
g_2(\widetilde{\beta}_{\mathbf{k},j}) \leq c_3 T_{2n}^{-1} \exp\{-T_{2n}(2\tau_{n,p})^{1/2}\} + c_4 T_{2n}^{-1} \exp\{-\tau_{n,p}/(\widetilde{\beta}_{\mathbf{k},j} + \widetilde{\epsilon}_n)^2\} \\
+ c_5 T_{2n}^{-1} \exp\{-c_6 T_{2n} \tau_{n,p}^{1/2}\}.
\end{aligned}
$$

Since $\widetilde{\epsilon}_n \asymp (n\nu_{\mathbf{k}*}/\tau_{n,p})^{-1/4}$, when $\widetilde{\beta}_{\mathbf{k},j} < \widetilde{\epsilon}_n$, $\tau_{n,p}/(\widetilde{\beta}_{\mathbf{k},j} + \widetilde{\epsilon}_n)^2 < \tau_{n,p}/(4\widetilde{\epsilon}_n^2) \asymp T_{2n}\tau_{n,p}^{1/2}$, and when $\widetilde{\beta}_{\mathbf{k},j} \geq \widetilde{\epsilon}_n$, $\tau_{n,p}/(\widetilde{\beta}_{\mathbf{k},j} + \widetilde{\epsilon}_n)^2 \leq \tau_{n,p}/(4\widetilde{\beta}^2_{\mathbf{k},j}) < T_{2n}\tau_{n,p}^{1/2}$. In overall, the right-hand side of the above display would be dominated by the second term, which shows that $g_2(\widetilde{\beta}_{\mathbf{k},j}) \leq cT_{2n}^{-1}\exp\{-\tau_{n,p}/(\widetilde{\beta}_{\mathbf{k},j} + \widetilde{\epsilon}_n)^2\}$ for some constant c. When $\widetilde{\beta}_{\mathbf{k},j} < 0$, we can show the same result by following exactly the same steps explained above. □

We now present some auxiliary results that are used to prove Theorems 1 and 2. We make use of the following simple union bound multiple times: for non-negative random variables $V_1, \ldots, V_m$ and $a > 0$,

$$P(\sum_{l=1}^{m} V_l > a) \leq \sum_{l=1}^{m} P(V_l > a/m) \leq m \max_{1\leq l\leq m} P(V_l > a/m). \tag{S3}$$

We define some notations that are used in the subsequent proofs. Let $\mathbf{t}$ denote the true data generating model, and let $\beta^0_{\mathbf{t}}$ denote the true regression coefficient corresponding to $\mathbf{t}$. Let $\mathbf{c_t} = \mathbf{t} \setminus \mathbf{k}$, $\mathbf{c_k} = \mathbf{k} \setminus \mathbf{t}$, and $\mathbf{u} = \mathbf{k} \cup \mathbf{t}$. Also, we define the cardinality of a model $\mathbf{k}$ as k and in the same spirit, denote $c_k = |\mathbf{c_k}|$, $c_t = |\mathbf{c_t}|$, and $t = |\mathbf{t}|$. $\{x\}_j$ denotes the *j*-th element of the vector x, and $diag\{A\}_j$ refers to the *j*-th diagonal element in the square matrix A. We denote $\chi^2_m(\lambda)$ a non-central chi-square distribution with the degrees of freedom m and non-centrality parameter λ; a central chi-square distribution is simply denoted by χ^2_m.

An important property that is used in the subsequent proofs concerns the distribution of the marginal ridge estimator. Let $\widetilde{\beta}_{\mathbf{k}} = (X_{\mathbf{k}}^{\mathrm{T}}X + 1/\tau_{n,p}I_{\mathbf{k}})^{-1}X_{\mathbf{k}}^{\mathrm{T}}y$ and $\widetilde{\beta}_{\mathbf{k},j} = \{\widetilde{\beta}_{\mathbf{k}}\}_j$. Then,

$$\widetilde{\beta}_{\mathbf{k},j} \sim N(\beta^*_{\mathbf{k},j}, \sigma^{2*}_{\mathbf{k},j}), \tag{S4}$$

where $\beta^*_{\mathbf{k},j} = \{(X_{\mathbf{k}}^{\mathrm{T}}X + 1/\tau_{n,p}I_{\mathbf{k}})^{-1}X_{\mathbf{k}}^{\mathrm{T}}X_{\mathbf{t}}\beta^*_{\mathbf{t}}\}_j$ and $\sigma^{2*}_{\mathbf{k},j} = \sigma^2 diag\{(X_{\mathbf{k}}^{\mathrm{T}}X_{\mathbf{k}} + 1/\tau_{n,p}I_{\mathbf{k}})^{-1}\}_j$. It is also evident that $(\widetilde{\beta}_{\mathbf{k},j} - \beta^*_{\mathbf{k},j})^2/\sigma^{2*}_{\mathbf{k},j} \sim \chi^2_1$.

A set of technical results follow that are used in the proof of the main results. Define

$$H_{1n} = \sum_{\substack{\mathbf{k}:\mathbf{t}\subsetneq\mathbf{k},\\ |\mathbf{k}|\le q_n}} \frac{m_{\mathbf{k}}(y)\pi(\mathbf{k})}{m_{\mathbf{t}}(y)\pi(\mathbf{t})} = \sum_{\substack{\mathbf{k}:\mathbf{t}\subsetneq\mathbf{k},\\ |\mathbf{k}|\le q_n}} \frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)}, \quad H_{2n} = \sum_{\substack{\mathbf{k}:\mathbf{t}\not\subseteq\mathbf{k},\\ |\mathbf{k}|\le q_n}} \frac{m_{\mathbf{k}}(y)\pi(\mathbf{k})}{m_{\mathbf{t}}(y)\pi(\mathbf{t})} = \sum_{\substack{\mathbf{k}:\mathbf{t}\not\subseteq\mathbf{k},\\ |\mathbf{k}|\le q_n}} \frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)}. \tag{S5}$$

Lemma 2. *Fix* $\epsilon > 0$. *Let* $\Gamma_d = \{\mathbf{k} : |\mathbf{k}| \le q_n, \mathbf{t} \subsetneq \mathbf{k}, |\mathbf{k}| - |\mathbf{t}| = d\}$ *for* $d = 1, \dots, q_n - |\mathbf{t}|$. *Suppose there exist constants* $c, \delta > 0$ *such that* $\max_{\mathbf{k}\in\Gamma_d} P\{\pi(\mathbf{k} \mid y)/\pi(\mathbf{t} \mid y) > \epsilon p^{-d}/q_n\} \le cp^{-d(1+\delta)}$ *for* $d = 1, \dots, q_n - |\mathbf{t}|$. *Then,* H_{1n} *converges to zero in probability as* n *tends to* ∞, *where* H_{1n} *is as in* (S5).

Proof. Clearly, $|\Gamma_d| = \binom{p-|\mathbf{t}|}{d}$. Using (S3), we bound

$$\begin{aligned} P\Big\{\sum_{\mathbf{k}:\mathbf{t}\subsetneq\mathbf{k}} \frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)} > \epsilon\Big\} &= P\Big\{\sum_{d=1}^{q_n-|\mathbf{t}|}\sum_{\mathbf{k}\in\Gamma_d} \frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)} > \epsilon\Big\} \\ &\le \sum_{d=1}^{q_n-|\mathbf{t}|} P\Big\{\sum_{\mathbf{k}\in\Gamma_d} \frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)} > \epsilon/q_n\Big\} \\ &\le \sum_{d=1}^{q_n-|\mathbf{t}|} \binom{p-|\mathbf{t}|}{d} \max_{\mathbf{k}\in\Gamma_d} P\Big\{\frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)} > \epsilon p^{-d}/q_n\Big\} \le \sum_{d=1}^{q_n-|\mathbf{t}|} cp^{-d\delta}. \end{aligned}$$

Finally, $\sum_{d=1}^{q_n-|\mathbf{t}|} cp^{-d\delta} \le cq_n p^{-\delta} \to 0$ as $n \to \infty$. □

Lemma 3. *Fix* $\epsilon > 0$ *and let* $t = |\mathbf{t}|$. *Define* $\Gamma_{k,c_k,c_t} = \{\mathbf{k} : |\mathbf{k}| \le q_n, |\mathbf{k}| = k, |\mathbf{k}\backslash\mathbf{t}| = c_k, |\mathbf{t}\backslash\mathbf{k}| = c_t\}$ *for* $k = 0, \dots, q_n;\ c_k = 0, \dots, k;\ c_t = 1, \dots, t$. *Suppose*

$$\max_{\mathbf{k}\in\Gamma_{k,c_k,c_t}} P\Big[\frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)} > \epsilon n^{-3} p^{-k} n^{-c_k} t^{-t}\Big] \le cp^{-k(1+\delta)},$$

with some postive constants c *and* δ. *Then,* H_{2n} *converges to zero as* n *tends to* ∞, *where* H_{2n} *is as in* (S5).

Proof. Clearly, $|\Gamma_{k,c_k,c_t}| = \binom{p}{k}\binom{k}{c_k}\binom{t}{c_t}$.

$$
\begin{aligned}
P\Big\{\sum_{\mathbf{k}:\mathbf{t}\not\subseteq\mathbf{k}} \frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)} > \epsilon\Big\} &\le P\Big\{\sum_{k=1}^{q_n}\sum_{c_k=0}^{k}\sum_{c_t=1}^{t}\sum_{\mathbf{k}\in\Gamma_{k,c_k,c_t}} \frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)} > \epsilon\Big\} \\
&\le P\Big\{\sum_{k=1}^{q_n}\sum_{c_k=0}^{k}\sum_{c_t=1}^{t}\sum_{\mathbf{k}\in\Gamma_{k,c_k,c_t}} \frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)} > \epsilon\Big\} \\
&\le \sum_{k=1}^{q_n}\sum_{c_k=0}^{k}\sum_{c_t=1}^{t} P\Big\{\sum_{\mathbf{k}\in\Gamma_{k,c_k,c_t}} \frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)} > \epsilon n^{-3}\Big\} \\
&\le \sum_{k=1}^{q_n}\sum_{c_k=0}^{k}\sum_{c_t=1}^{t} p^k n^{c_k} t^t \max_{\mathbf{k}\in\Gamma_{k,c_k,c_t}} P\Big\{\frac{\pi(\mathbf{k}\mid y)}{\pi(\mathbf{t}\mid y)} > \epsilon n^{-3} p^{-k} n^{-c_k} t^{-t}\Big\} \\
&\le \sum_{k=1}^{q_n}\sum_{c_k=0}^{k}\sum_{c_t=1}^{t} p^k n^{c_k} t^t p^{-k(1+\delta)} \to 0,
\end{aligned}
$$

as $n \to \infty$. □

Lemma 4. *Suppose W follows a non-central chi-square distribution with the degree of freedom m_n that is a positive integer and the non-central parameter $\lambda_n \ge 0$, i.e, $W \sim \chi^2_{m_n}(\lambda_n)$. Also, consider w_n and t_n such that $w_n \to 0$ and $t_n \to \infty$ as n tends to ∞. Also, assume that $m_n \prec t_n$. Then,*

$$
P(W \le \lambda_n w_n) \le c_1 \lambda_n^{-1} \exp\{-\lambda_n(1-w_n)^2\}, \tag{S6}
$$

And

$$
P(W > \lambda_n + t_n) \le c_2 \left(\frac{t_n}{2m_n}\right)^{m_n/2} \exp\{m_n/2 - t_n/2\} + c_3 \lambda_n^{1/2} t_n^{-1} \exp\left\{-\frac{t_n^2}{32\lambda_n}\right\}, \tag{S7}
$$

where c_1, c_2, and c_3 are some positive constants.

Proof. W can be expressed as $W = \sum_{i=1}^{m_n}\{Z_i + (\lambda_n/m_n)^{1/2}\}^2$, where $Z_i \overset{i.i.d}{\sim} N(0,1)$ for $i = 1,\dots,m$. Then, by the fact that $P(Z > a) \le (2\pi)^{-1/2} a^{-1} \exp\{-a^2/2\}$ for any $a > 0$, we can show that there exist some positive constants c_1 such that

$$\begin{aligned} P\left(W \le \lambda_n w_n\right) &= P\Big\{\sum_{i=1}^{m_n} Z_i^2 + 2(\lambda_n/m_n)^{1/2}\sum_{i=1}^{m_n} Z_i + \lambda_n \le \lambda_n w_n\Big\} \\ &\le P\Big\{m_n^{-1/2}\sum_{i=1}^{m_n} Z_i \le -\lambda_n^{1/2}(1-w_n)/2\Big\} \\ &= P\Big\{|Z_1| \ge \lambda_n^{1/2}(1-w_n)/2\Big\}/2 \\ &\le c_1\lambda_n^{-1}\exp\{-\lambda_n(1-w_n)^2/2\}, \end{aligned}$$

since Z_1 follows a standard normal distribution.

Also, by using Chernoffs's bound and the fact that $P(Z > a) \le (2\pi)^{-1/2}a^{-1}\exp\{-a^2/2\}$ for any $a > 0$, one can show that

$$\begin{aligned} &P(W > \lambda_n + t_n) = P\left\{\sum_{i=1}^{m_n} Z_i^2 + 2(\lambda_n/m_n)^{1/2}\sum_{i=1}^{m_n} Z_i > t_n\right\} \\ \le\ & P\left(\sum_{i=1}^{m_n} Z_i^2 > t_n/2\right) + P\left\{m_n^{-1/2}\sum_{i=1}^{m_n} Z_i > \lambda_n^{-1/2}t_n/4\right\} \\ \le\ & c_2\left(\frac{t_n}{2m_n}\right)^{m_n/2}\exp\left\{m_n/2 - t_n/2\right\} + c_3\lambda_n^{1/2}t_n^{-1}\exp\left\{-\frac{t_n^2}{32\lambda_n}\right\}, \end{aligned}$$

where c_2 and c_3 are some positive constants. □

Lemma 5. *Consider $Q_{\mathbf{k}}$ defined in (6) for an arbitrary model $\mathbf{k}$. Fix any $\delta > 0$. For any $\mathbf{k}$ with $\mathbf{t} \subsetneq \mathbf{k}$,*

$$P\left[Q_{\mathbf{k}}/Q_{\mathbf{t}} > \exp\left\{-|\mathbf{k}\setminus\mathbf{t}|\tau_{n,p}^{2/3}(n\nu_{\mathbf{k}*})^{1/3} + |\mathbf{t}|\tau_{n,p}^{1-\delta/8}(n\nu_{\mathbf{k}*})^{\delta/8}\right\}\right] \le p^{-|\mathbf{k}\setminus\mathbf{t}|(1+\delta)}, \tag{S8}$$

and for $\mathbf{k}$ such that $\mathbf{t} \nsubseteq \mathbf{k}$,

$$P\left[Q_{\mathbf{k}}/Q_{\mathbf{t}} > \exp\left\{\|\beta_{\mathbf{t}}^0\|_2^2 n\nu_{\mathbf{u}*}/\{2\log(\tau_{n,p}/\log p)\}\right\}\right] \le p^{-|\mathbf{k}|(1+\delta)}. \tag{S9}$$

Proof. By Lemma 1, it is sufficient to show that

$$P\left[\prod_{j\in\mathbf{t}}(Q^U_{\mathbf{k},j}/Q^L_{\mathbf{t},j}) > \exp\{|\mathbf{t}|\tau_{n,p}^{1-\delta/8}(n\nu_{\mathbf{k}*})^{\delta/8}\}\right] + P\left[\prod_{j\in\mathbf{k}\setminus\mathbf{t}} Q^U_{\mathbf{k},j} > \exp\{-|\mathbf{k}\setminus\mathbf{t}|\tau_{n,p}^{2/3}(n\nu_{\mathbf{k}*})^{1/3}\}\right]$$
$$\le \; p^{-|\mathbf{k}\setminus\mathbf{t}|(1+\delta)}. \tag{S10}$$

We first shall show that the first term in the left-hand side of (S10) is bounded above by $\exp\{-cn\nu_{\mathbf{k}*}\}$ for some constant c.

$$\begin{aligned} & P\left[\prod_{j\in\mathbf{t}}\frac{Q^U_{\mathbf{k},j}}{Q^L_{\mathbf{t},j}} > \exp\left\{|\mathbf{t}|\tau_{n,p}^{1-\delta/8}(n\nu_{\mathbf{k}*})^{\delta/8}\right\}\right] \le \sum_{j\in\mathbf{t}} P\left[\frac{Q^U_{\mathbf{k},j}}{Q^L_{\mathbf{t},j}} > \exp\left\{\tau_{n,p}^{1-\delta/8}(n\nu_{\mathbf{k}*})^{\delta/8}\right\}\right] \\ = \;& \sum_{j\in\mathbf{t}} P\left[c'\left(\frac{n\nu_{\mathbf{k}*}+1/\tau_{n,p}}{n\nu^*_{\mathbf{t}}+1/\tau_{n,p}}\right)^{-1/2}\exp\left\{-\tau_{n,p}\left(1/(|\widetilde{\beta}_{\mathbf{k},j}|+\widetilde{\epsilon}_n)^2 - 1/\widetilde{\beta}^{*2}_{\mathbf{k},j}\right)\right\} > \exp\left\{\tau_{n,p}^{1-\delta/8}(n\nu_{\mathbf{k}*})^{\delta/8}\right\}\right] \\ \le \;& \sum_{j\in\mathbf{t}} P[|\widetilde{\beta}_{\mathbf{k},j}-\beta^*_{\mathbf{k},j}| > \epsilon'] + \sum_{j\in\mathbf{t}} P[|\widetilde{\beta}_{\mathbf{t},j}-\beta^*_{\mathbf{t},j}| > \epsilon'], \end{aligned} \tag{S11}$$

for some small enough $\epsilon' > 0$ and some positive constant c' and $\widetilde{\beta}^*_{\mathbf{k},j} \in [\widetilde{\beta}_{\mathbf{k},j}-\widetilde{\epsilon}_n, \widetilde{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n] \setminus (-\widetilde{\epsilon}_n, \widetilde{\epsilon}_n)^c$ as defined in Lemma 1, and $\widetilde{\beta}_{\mathbf{k},j}$ and $\beta^*_{\mathbf{k},j}$ defined in (S4). The last inequality in the above display asymptotically holds, since

$$\tau_{n,p}^{1-\delta/8}(n\nu_{\mathbf{k}*})^{\delta/8} \succ \tau_{n,p}/(|\beta^*_{\mathbf{k},j}|-\epsilon'-\widetilde{\epsilon}_n)^2,$$

for any $\delta > 0$.

Since $(\widetilde{\beta}_{\mathbf{k},j}-\beta^*_{\mathbf{k},j})^2/\sigma^{*2}_{\mathbf{k},j} \sim \chi^2_1$ and $\sigma^{*2}_{\mathbf{k},j} \ge (n\nu_{\mathbf{k}*}+1/\tau_{n,p})^{-1}$, by using Lemma 4, one can show that the first term in (S11) bounded above by $\exp\{-c_1\epsilon'^2 n\nu_{\mathbf{k}*}\}$ for some constant c_1. Similarly, the second term in (S11) is bounded above by $\exp\{-c_2\epsilon'^2 n\}$ for some constant c_2, since *Assumption 5* states that $X_{\mathbf{t}}^{\mathrm{T}}X_{\mathbf{t}}/n$ is asymptically isotropic. Therefore, (S11) is asymptotically bounded by $p^{-q_n(1+\delta)}$ by *Assumption 3*.

Next, we shall show that the second term in the left-hand side of (S10) is bounded above by

$\exp\{-c\tau_{n,p}^{1/3}(n\nu_{\mathbf{k}*})^{2/3}\}$ for some positive constant c. Since when $j \in \mathbf{k} \setminus \mathbf{t}$ and $\mathbf{t} \subsetneq \mathbf{k}$, $\beta^*_{\mathbf{k},j} \asymp n^{-1}$,

$$
\begin{aligned}
& P\left[\prod_{j\in\mathbf{k}\setminus\mathbf{t}} Q^U_{\mathbf{k},j} > \exp\{-|\mathbf{k}\setminus\mathbf{t}|\tau_{n,p}^{2/3}(n\nu_{\mathbf{k}*})^{1/3}\}\right] \\
\leq\ & \sum_{j\in\mathbf{k}\setminus\mathbf{t}} P\left[c'(n\nu_{\mathbf{k},j}+1/\tau_{n,p})^{-1/2}\exp\left\{-\frac{\tau_{n,p}}{(|\widetilde{\beta}_{\mathbf{k},j}|+\widetilde{\epsilon}_n)^2}\right\} > \exp\{-\tau_{n,p}^{2/3}(n\nu_{\mathbf{k}*})^{1/3}\}\right] \\
=\ & \sum_{j\in\mathbf{k}\setminus\mathbf{t}} P\left[\widetilde{\beta}^2_{\mathbf{k},j} > \left\{\tau_{n,p}^{1/2}\left((n\nu_{\mathbf{k}*})^{1/3}\tau_{n.p}^{2/3} - \log(n\nu_{\mathbf{k}*}+1/\tau_{n,p})/2+\log c'\right)^{-1/2} - \widetilde{\epsilon}_n\right\}^2\right] \\
\leq\ & \sum_{j\in\mathbf{k}\setminus\mathbf{t}} P\left[(\widetilde{\beta}_{\mathbf{k},j}-\beta^*_{\mathbf{k},j})^2/\sigma^*_{\mathbf{k},j} > c''\left(\frac{\tau_{n,p}}{n\nu_{\mathbf{k}*}}\right)^{1/3}(n\nu_{\mathbf{k}*}+1/\tau_{n,p})/\sigma^2\right],
\end{aligned}
$$

for some positive contant c' and c''. Since $(\widetilde{\beta}_{\mathbf{k},j}-\beta^*_{\mathbf{k},j})^2/\sigma^*_{\mathbf{k},j} \sim \chi^2_1$, by Lemma 4 the last quantity in the above display can be bounded by $\exp\{-c\tau_{n,p}^{1/3}(n\nu_{\mathbf{k}*})^{2/3}\}$ for some contant c. By *Assumption 3*, $\exp\{-c\tau_{n,p}^{1/3}(n\nu_{\mathbf{k}*})^{2/3}\} \prec p^{-q_n(1+\delta)} \leq p^{|\mathbf{k}\setminus\mathbf{t}|(1+\delta)|}$, which proves the statement (S10).

We now shall show that the equation (S9) holds for any $\delta > 0$. The left-hand side of (S9) can be bounded above by

$$
\begin{aligned}
& P\left[\prod_{j\in\mathbf{k}} Q^U_{\mathbf{k},j}\Big(\prod_{j\in\mathbf{t}} Q^L_{\mathbf{t},j}\Big)^{-1} > \exp\left\{\|\beta^0_{\mathbf{t}}\|_2^2 n\nu_{\mathbf{u}*}/\{2\log(\tau_{n,p}/\log p)\}\right\}\right] \\
\leq\ & \sum_{j\in\mathbf{k}} P\left[c(n\nu_{\mathbf{k}*}+1/\tau_{n,p})^{-1/2}\exp\left\{-\tau_{n,p}/(|\widetilde{\beta}_{\mathbf{k},j}|+\widetilde{\epsilon}_n)^2\right\} > \exp\left\{\|\beta^0_{\mathbf{t}}\|_2^2 n\nu_{\mathbf{u}*}/\{4|\mathbf{k}|\log(\tau_{n,p}/\log p)\}\right\}\right] \\
& + \sum_{j\in\mathbf{t}} P\left[c'(n\nu_{\mathbf{t}*}+1/\tau_{n,p})^{1/2}\exp\left\{\tau_{n,p}/(\widetilde{\beta}^{*2}_{\mathbf{t},j})\right\} > \exp\left\{\|\beta^0_{\mathbf{t}}\|_2^2 n\nu_{\mathbf{u}*}/\{4|\mathbf{t}|\log(\tau_{n,p}/\log p)\}\right\}\right] \\
\leq\ & \sum_{j\in\mathbf{k}} P\left[-\frac{\tau_{n,p}}{(|\widetilde{\beta}_{\mathbf{k},j}|+\widetilde{\epsilon}_n)^2} > \|\beta^0_{\mathbf{t}}\|_2^2 n\nu_{\mathbf{u}*}/\{4|\mathbf{k}|\log(\tau_{n,p}/\log p)\} + \log c\right] \qquad \text{(S12)} \\
& + \sum_{j\in\mathbf{t}} P\left[|\widetilde{\beta}^*_{\mathbf{t},j}| < c''\|\beta^0_{\mathbf{t}}\|_2^{-1}(n\nu_{\mathbf{u}*})^{-1/2}\{4|\mathbf{t}|\log(\tau_{n,p}/\log p)\}^{1/2}\tau_{n,p}^{1/2}\right], \qquad \text{(S13)}
\end{aligned}
$$

where c, c', and c'' are some positive constants.

(S12) is always zero since the left-hand side in the probability is always negative and the right-hand side in the probability operator is always positive. So, we focus on (S13) as below:

Since $\widetilde{\beta}_{\mathbf{t},j} - \widetilde{\epsilon}_n \leq \widetilde{\beta}^*_{\mathbf{t},j} \leq \widetilde{\beta}_{\mathbf{t},j} + \widetilde{\epsilon}_n$ implies $|\widetilde{\beta}_{\mathbf{t},j}| - \widetilde{\epsilon}_n \leq |\widetilde{\beta}^*_{\mathbf{t},j}| \leq |\widetilde{\beta}_{\mathbf{t},j}| + \widetilde{\epsilon}_n$, (S13) can be bounded above by

$$\begin{aligned} & \sum_{j\in\mathbf{t}} P\left[|\widetilde{\beta}^*_{\mathbf{t},j}| < c''\|\beta^0_{\mathbf{t}}\|_2^{-1}(n\nu_{\mathbf{u}*})^{-1/2}\{4|\mathbf{t}|\log(\tau_{n,p}/\log p)\}^{1/2}\tau_{n,p}^{1/2}\right] \\ \leq & \sum_{j\in\mathbf{t}} P\left[|\widetilde{\beta}_{\mathbf{t},j}| < c''\|\beta^0_{\mathbf{t}}\|_2^{-1}(n\nu_{\mathbf{u}*})^{-1/2}\{4|\mathbf{t}|\log(\tau_{n,p}/\log p)\}^{1/2}\tau_{n,p}^{1/2} + \widetilde{\epsilon}_n\right], \end{aligned}$$

where $\beta^*_{\mathbf{t},j}$ is defined in (S4). Since $\widetilde{\beta}^2_{\mathbf{t},j}/\sigma^2_{\mathbf{t},j} \sim \chi^2_1(\beta^{*2}_{\mathbf{t},j}/\sigma^2_{\mathbf{t},j})$ and $\sigma^2_{\mathbf{t},j} \asymp \sigma^2/n$ for $j \in \mathbf{t}$, by using Lemma 4 and *Assumption 5*, one can show that the probability is bounded by $\exp\{-cn\}$ for some constant c, and it is evident that $\exp\{-cn\} \prec p^{-|\mathbf{k}|(1+\delta)}$, which completes the proof of the Lemma. □

2 Proofs of Main Results

Proof of Theorem 1. We have $\pi(\mathbf{t} \mid \mathbf{y}) = \mathbf{m_t}(\mathbf{y})\pi(\mathbf{t})/\{\sum_{\mathbf{k}:|\mathbf{k}|\leq \mathbf{q_n}} \mathbf{m_k}(\mathbf{y})\pi(\mathbf{k})\}$, since $\pi(\mathbf{k}) = 0$ for any $\mathbf{k}$ with $|\mathbf{k}| > q_n$. Recall H_{1n} and H_{2n} from (S5) and note that $\pi(\mathbf{t} \mid y) = (1+H_{1n}+H_{2n})^{-1}$. Hence to show that $\pi(\mathbf{t} \mid y)$ converges to one in probability, it is sufficient to establish that H_{1n} and H_{2n} both converge in probability to zero as n tends to ∞. We shall prove the Theorem by showing:

For any $\delta \in (0, 8/3)$ and any model $\mathbf{k} \in \Gamma_d$ (defined in Lemma 2),

$$P\left[\frac{\pi(\mathbf{k} \mid y)}{\pi(\mathbf{t} \mid y)} > \epsilon p^{-d} q_n^{-1}\right] \leq p^{-d(1+\delta)}, \tag{S14}$$

and for any model $\mathbf{k} \in \Gamma_{k,c_k,c_t}$ (defined in Lemma 3),

$$P\left[\frac{\pi(\mathbf{k} \mid y)}{\pi(\mathbf{t} \mid y)} > \epsilon n^{-3} p^{-k} n^{-c_k} t^{-t}\right] \leq c p^{-k(1+\delta)}. \tag{S15}$$

Then, it is evident that H_{1n} and H_{2n} both converge to zero in probability by Lemma 2 and 3 respectively.

First, we shall show that (S14) holds. For any $\mathbf{k} \in \Gamma_d$, recall that

$$P\Big[\frac{\pi(\mathbf{k} \mid y)}{\pi(\mathbf{t} \mid y)} > \epsilon p^{-d} q_n^{-1}\Big] \leq P\Big[C_{n,p}^{-d}\frac{Q_{\mathbf{k}}}{Q_{\mathbf{t}}}\exp\Big\{-\frac{1}{2\sigma^2}\big(\widetilde{R}_{\mathbf{k}} - \widetilde{R}_{\mathbf{t}}\big)\Big\} > \epsilon p^{-d}/q_n\Big].$$

Since $\widetilde{R}_{\mathbf{k}} > R^*_{\mathbf{k}}$ and $\widetilde{R}_{\mathbf{t}} < R^*_{\mathbf{t}} + \eta$, where $\eta = d_1\widehat{\beta}_{\mathbf{t}}^T\widehat{\beta}_{\mathbf{t}}/\tau_{n,p}$ for some constant d_1 and $\widehat{\beta}_{\mathbf{t}}$ is the ordinary least square estimator of $\beta_{\mathbf{t}}$ in the true model **t**, by using (S3), the term in the last display can be bounded above by

$$\begin{aligned}
& P\Big[C_{n,p}^{-d}\frac{Q_{\mathbf{k}}}{Q_{\mathbf{t}}}\exp\big\{-\big(R^*_{\mathbf{k}} - R^*_{\mathbf{t}}\big)/(2\sigma^2) + \eta/(2\sigma^2)\big\} > \epsilon p^{-d}/q_n\Big] \\
\leq\ & P\Big[C_{n,p}^{-d}\frac{Q_{\mathbf{k}}}{Q_{\mathbf{t}}}p^{d(1+\delta)+\delta} > \epsilon p^{-d}/q_n\Big] && \text{(S16)} \\
& +P\left[R^*_{\mathbf{t}} - R^*_{\mathbf{k}} > 2\sigma^2 d(1+\delta)\log p\right] && \text{(S17)} \\
& +P\left[\exp\{\eta/(2\sigma^2)\} > \epsilon p^{\delta}\right]. && \text{(S18)}
\end{aligned}$$

By using Lemma 5, (S16) is less than $p^{-d(1+\delta)}$ when $\delta < 8/3$. Since $(R^*_{\mathbf{t}} - R^*_{\mathbf{k}})/\sigma^2 \sim \chi^2_{|\mathbf{k}\setminus\mathbf{t}|}$, by using (S6) in Lemma 4, we can show that (S17) is bounded by $cp^{-d(1+\delta)}$ for some positive constant c. Since $\tau_{n,p}n\nu_{\mathbf{t}*}\eta/d_1\sigma^2 \leq \widehat{\beta}_{\mathbf{t}}^T X_{\mathbf{t}}^T X_{\mathbf{t}}\widehat{\beta}_{\mathbf{t}}/\sigma^2 \sim \chi^2_{|\mathbf{t}|}\left(\beta_{\mathbf{t}}^{0T}X_{\mathbf{t}}^T X_{\mathbf{t}}\beta_{\mathbf{t}}^0\right)$, by using the inequality (S7) in Lemma 4, (S18) can be expressed as

$$\begin{aligned}
P\left[\exp\left\{\eta/2\sigma^2\right\} > \epsilon p^{\delta}\right] &\leq P\left[\tau_{n,p}n\nu_{\mathbf{t}*}\eta/d_1\sigma^2 > 2\tau_{n,p}n\nu_{\mathbf{t}*}(\log\epsilon + \delta\log p)/d_1\right] \\
&\leq P\big[\widehat{\beta}_{\mathbf{t}}^T X_{\mathbf{t}}^T X_{\mathbf{t}}\widehat{\beta}_{\mathbf{t}}/\sigma^2 > 2\tau_{n,p}n\nu_{\mathbf{t}*}(\log\epsilon + \delta\log p)/d_1\big] \\
&\leq (n\delta\log p)^{|\mathbf{t}|/2}\exp\{-c_1\delta(n\log p)\} + n^{-1/2}(\delta\log p)^{-1}\exp\{-c_2(n\log p)^2/n\} \\
&\leq c_3 p^{-|\mathbf{k}|(1+\delta)}, && \text{(S19)}
\end{aligned}$$

for some positive constant c_1, c_2, and c_3, which proves that (S14) holds.

Next, we consider (S15). Recall that $\mathbf{u} = \mathbf{k} \cup \mathbf{t}$. By using (S3), it can be shown that

$$
\begin{aligned}
& P\Big[\frac{\pi(\mathbf{k} \mid y)}{\pi(\mathbf{t} \mid y)} > \epsilon n^{-3} p^{-|\mathbf{k}|} n^{-|\mathbf{k}\setminus\mathbf{t}|} |\mathbf{t}|^{-|\mathbf{t}|}\Big] \\
\leq\ & P\Big[C_{n,p}^{-(|\mathbf{k}|-|\mathbf{t}|)} \frac{Q_{\mathbf{k}}}{Q_{\mathbf{t}}} \exp\Big\{-(\widetilde{R}_{\mathbf{k}} - \widetilde{R}_{\mathbf{t}})/(2\sigma^2)\Big\} > \epsilon n^{-3} p^{-|\mathbf{k}|} n^{-|\mathbf{k}\setminus\mathbf{t}|} |\mathbf{t}|^{-|\mathbf{t}|}\Big] \\
\leq\ & P\Big[C_{n,p}^{-|\mathbf{k}|-|\mathbf{t}|)} \frac{Q_{\mathbf{k}}}{Q_{\mathbf{t}}} \exp\big\{-(R^*_{\mathbf{k}} - R^*_{\mathbf{u}})/(2\sigma^2)\big\} > n^{-3-|\mathbf{k}\setminus\mathbf{t}|} |\mathbf{t}|^{-|\mathbf{t}|} p^{-|\mathbf{k}|(2+\delta)+\delta}\Big] \\
& + P\big[\exp\big\{\big(R^*_{\mathbf{t}} - R^*_{\mathbf{u}}\big)/(2\sigma^2)\big\} \geq \epsilon p^{|\mathbf{k}|(1+\delta)}\big] + P\big[\exp\big(\eta/(2\sigma^2)\big) > p^{\delta}\big] \\
\leq\ & P\big[\exp\big\{\big(R^*_{\mathbf{t}} - R^*_{\mathbf{u}}\big)/2\sigma^2\big\} > \epsilon p^{|\mathbf{k}|(1+\delta)}\big] && \text{(S20)} \\
& + P\big[\exp\big(\eta/2\sigma^2\big) > p^{\delta}\big] && \text{(S21)} \\
& + P\left[R^*_{\mathbf{k}} - R^*_{\mathbf{u}} < 2\sigma^2 \|\beta^0_{\mathbf{t}}\|_2^2 n\nu_{\mathbf{u}*}/\log(\tau_{n,p}/\log p)\right] && \text{(S22)} \\
& + P\left[Q_{\mathbf{k}}/Q_{\mathbf{t}} > \exp\left\{\|\beta^0_{\mathbf{t}}\|_2^2 n\nu_{\mathbf{u}*}/\{2\log(\tau_{n,p}/\log p)\}\right\}\right]. && \text{(S23)}
\end{aligned}
$$

Since $(R^*_{\mathbf{t}} - R^*_{\mathbf{u}})/\sigma^2$ follows a $\chi^2_{|\mathbf{u}\setminus\mathbf{t}|}$ distribution, (S20) is also bounded by $c_1 p^{-|\mathbf{k}|(1+\delta)}$ with some constant c_1. By following the same steps regarding (S19), one can show that (S21) is bounded by $c_2 p^{-|\mathbf{k}|(1+\delta)}$ for some constant c_2. We note that $(R^*_{\mathbf{k}} - R^*_{\mathbf{u}})/\sigma^2 \sim \chi^2_{|\mathbf{u}\setminus\mathbf{k}|}(\lambda_n)$ with $\lambda_n = \beta^{0T}_{\mathbf{t}} X^T_{\mathbf{t}} (P_{\mathbf{u}} - P_{\mathbf{k}}) X_{\mathbf{t}} \beta^0_{\mathbf{t}}$, where $P_{\mathbf{k}}$ is the projection matrix of $X_{\mathbf{k}}$. As discussed in Narisetty and He (2014), $\lambda_n \geq n\nu_{\mathbf{u}*}\|\beta^0_{\mathbf{t}}\|_2^2$. Hence, by using Lemma 4, one can show that (S22) is bounded by $\exp\{-c_3\|\beta^0_{\mathbf{t}}\|_2^2 n\nu_{\mathbf{u}*}/\log(\tau_{n,p}/\log p)\}$ for some constant c_3. Lemma 5 states that (S23) is bounded by $p^{-|\mathbf{k}|(1+\delta)}$. In summary, since $q_n \prec \tau_{n,p}/\log p$ by *Assumption 3*, there exists some positive constant c_4 such that $P[\pi(\mathbf{k} \mid y)/\pi(\mathbf{t} \mid y) > \epsilon n^{-3} p^{-|\mathbf{k}|} n^{-|\mathbf{k}\setminus\mathbf{t}|} |\mathbf{t}|^{-|\mathbf{t}|}] \leq c_4 p^{-|\mathbf{k}|(1+\delta)}$. which completes the proof of Theorem 1. □

Proof of Corollary 2. Recall the penalty term of a model $\mathbf{k}$, $Q^*_{\mathbf{k}}$, based on the piMoM priors is

$$
Q^*_{\mathbf{k}} = \int \exp\Big\{ - (\beta_{\mathbf{k}} - \widehat{\beta}_{\mathbf{k}})^T \Sigma^{*-1}_{\mathbf{k}} (\beta_{\mathbf{k}} - \widehat{\beta}_{\mathbf{k}})/(2\sigma^2) - \sum_{j=1}^{|\mathbf{k}|} \tau_{n,p}/\beta^2_{\mathbf{k},j} - r \sum_{j=1}^{|\mathbf{k}|} \log(\beta^2_{\mathbf{k},j}) \Big\} d\beta_{\mathbf{k}},
$$

in (7). Since, for any $\epsilon > 0$, $\exp\big[-\sum_{j=1}^{|\mathbf{k}|}\{\epsilon\tau_{n,p}/\beta^2_{\mathbf{k},j} + r\log(\beta^2_{\mathbf{k},j})\}\big]$ is bounded above with respect to $\beta_{\mathbf{k},j}$, $Q^*_{\mathbf{k}} \leq C \int \exp\{-(\beta_{\mathbf{k}} - \widehat{\beta}_{\mathbf{k}})^T \Sigma^{*-1}_{\mathbf{k}} (\beta_{\mathbf{k}} - \widehat{\beta}_{\mathbf{k}})/(2\sigma^2) - \sum_{j=1}^{|\mathbf{k}|}(1-\epsilon)\tau_{n,p}/\beta^2_{\mathbf{k},j}\} d\beta_{\mathbf{k}}$ for some constant C. Following the exactly same steps in Lemma 1, $Q^*_{\mathbf{k}} \leq C'(n\nu^*_{\mathbf{k}})^{-1/2}\prod_{j=1}^{|\mathbf{k}|}\exp\{-(1-\epsilon)\tau_{n,p}/(|\widehat{\beta}_{\mathbf{k},j}|+\widetilde{\epsilon}_n)^2\}$ for some constant $C' > 0$.

We shall show that the model selection procedure based on piMoM priors as in (4) assures consistency by proving that $Q^*_{\mathbf{k}}$ and $Q_{\mathbf{k}}$ are asymptotically equivalent.

Next, we shall show that $Q^*_{\mathbf{k}}$ is bounded below by $C(n\nu^*_{\mathbf{k}})^{-1/2}\prod_{j=1}^{|\mathbf{k}|}\exp\{-(1-\epsilon)\tau_{n,p}/\widehat{\beta}^{*2}_{\mathbf{k},j}\}$ for some constant $C > 0$ and $\widehat{\beta}^*_{\mathbf{k},j} \in [\widehat{\beta}_{\mathbf{k},j}-\widetilde{\epsilon}_n, \widehat{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n]$. Since $\exp\big\{-\epsilon\tau_{n,p}/\beta^2_{\mathbf{k},j}+r\log(\beta^2_{\mathbf{k},j})\big\}$ can be minimized in $[\widehat{\beta}_{\mathbf{k},j}-\widetilde{\epsilon}_n, \widehat{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n]$, by following the proof of Lemma 1,

$$\begin{aligned}
&\int_{-\infty}^{\infty}\exp\{-n\nu^*_{\mathbf{k}}(\beta-\widehat{\beta}_{\mathbf{k},j})^2/(2\sigma^2)-\tau_{n,p}/\beta^2-r\log(\beta^2)\}d\beta\\
\geq\ &\int_{\widehat{\beta}_{\mathbf{k},j}-\widetilde{\epsilon}_n}^{\widehat{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n}\exp\{-n\nu^*_{\mathbf{k}}(\beta-\widehat{\beta}_{\mathbf{k},j})^2/(2\sigma^2)-(1-\epsilon)\tau_{n,p}/\beta^2\}\exp\{-\epsilon\tau_{n,p}/\beta^2-r\log(\beta^2)\}d\beta\\
\geq\ &C(n\nu^*_{\mathbf{k}})^{-1/2}\exp\left\{-(1-\epsilon)\tau_{n,p}/\widehat{\beta}^{*2}_{\mathbf{k},j}\right\},
\end{aligned}$$

where C is some constant and $\widehat{\beta}^*_{\mathbf{k},j} \in [\widehat{\beta}_{\mathbf{k},j}-\widetilde{\epsilon}_n, \widehat{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n] \setminus (-\widetilde{\epsilon}_n, \widetilde{\epsilon}_n)^c$.

Therefore, due to the asymptotic similarity between the ridge estimator and the least square estimator, the lower and upper bounds of $Q^*_{\mathbf{k}}$ are asymptotically equivalent to those of $Q_{\mathbf{k}}$ with the penalty parameter $(1-\epsilon)\tau_{n,p}$, which assures the strong consistency of the model selection based on the piMoM priors. □

Proof of Theorem 3. Under a situation where σ^2 is unknown, it is clear that

$$m_{\mathbf{k}}(y) = \tau_{n,p}^{-\frac{|\mathbf{k}|}{2}}\int(2\pi\sigma^2)^{-\frac{n+|\mathbf{k}|}{2}}\int\exp\left\{|\mathbf{k}|\left(\frac{2}{\sigma^2}\right)^{1/2}-\frac{(\beta_{\mathbf{k}}-\widetilde{\beta}_{\mathbf{k}})^T\widetilde{\Sigma}_{\mathbf{k}}^{-1}(\beta_{\mathbf{k}}-\widetilde{\beta}_{\mathbf{k}})}{2\sigma^2}-\sum_{j=1}^{|\mathbf{k}|}\frac{\tau_{n,p}}{\beta^2_{\mathbf{k},j}}\right\}\pi(\sigma^2)d\beta_{\mathbf{k}}d\sigma^2,$$

where $\pi(\sigma^2)$ is the prior for σ^2 (Inverse-gamma density with hyperparameters a_0 and b_0).

First, we shall show that the ratio between marginal likelihoods of a model $\mathbf{k}$ and the true model

$\mathbf{t}$ can be bounded as

$$\frac{m_{\mathbf{k}}(y)}{m_{\mathbf{t}}(y)} \le c^{\frac{|\mathbf{k}|-|\mathbf{t}|}{2}} \left(\frac{\widetilde{R}_{\mathbf{k}}+2b_0}{\widetilde{R}_{\mathbf{t}}+2b_0}\right)^{-n/2-a_0} \exp\left\{-\sum_{j=1}^{|\mathbf{k}|}\frac{\tau_{n,p}}{(|\widetilde{\beta}_{\mathbf{k},j}|+\widetilde{\epsilon}_n)^2}+\sum_{j=1}^{|\mathbf{t}|}\frac{\tau_{n,p}}{\widetilde{\beta}^{*2}_{\mathbf{t},j}}\right\}\frac{(n\nu_{\mathbf{k}*}\tau_{n,p}+1)^{-|\mathbf{k}|/2}}{(n\nu^*_{\mathbf{t}}\tau_{n,p}+1)^{-|\mathbf{t}|/2}}, \tag{S24}$$

where $\widetilde{\beta}^*_{\mathbf{t},j} \in [\widetilde{\beta}_{\mathbf{t},j}-\widetilde{\epsilon}_n, \widetilde{\beta}_{\mathbf{t},j}+\widetilde{\epsilon}_n] \setminus (-\widetilde{\epsilon}_n, \widetilde{\epsilon}_n)^c$ for $j \in 1,\ldots,|\mathbf{t}|$ and c is some constant. Next, we shall show that $\{(\widetilde{R}_{\mathbf{k}}+2b_0)/(\widetilde{R}_{\mathbf{t}}+2b_0)\}^{-n/2-a_0} \le \exp\{-(\widetilde{R}_{\mathbf{k}}-\widetilde{R}_{\mathbf{t}})/(2\sigma_0^2(1+u_n))\}$, where σ_0^2 is the true regression variance that involves in the data-generating process, and u_n is some random variable that is concentrated around a finite value with at least probability $1-\exp\{-cn\}$ for some constant c. Then, by following the same steps in the proof of Theorem 1, the proof of Corollary 2 is completed.

By Lemma 1, the marginal likelihood of a model $\mathbf{k}$ can be bounded by

$$\begin{aligned} m_{\mathbf{k}}(y) &\le \{c_1(n\nu_{\mathbf{k}*}\tau_{n,p}+1)\}^{-\frac{|\mathbf{k}|}{2}} \int (\sigma^2)^{-\frac{n+2a_0}{2}-1} \exp\left\{|\mathbf{k}|\left(\frac{2}{\sigma^2}\right)^{1/2} - \sum_{j=1}^{|\mathbf{k}|}\frac{\tau_{n,p}}{(|\widetilde{\beta}_{\mathbf{k},j}|+\widetilde{\epsilon}_n)^2} - \frac{\widetilde{R}_{\mathbf{k}}+2b_0}{2\sigma^2}\right\} d\sigma^2 \\ &\le \{c_1(n\nu_{\mathbf{k}*}\tau_{n,p}+1)\}^{-\frac{|\mathbf{k}|}{2}} \exp\left\{-\sum_{j=1}^{|\mathbf{k}|}\frac{\tau_{n,p}}{(|\widetilde{\beta}_{\mathbf{k},j}|+\widetilde{\epsilon}_n)^2}\right\}(1+\exp\{2|\mathbf{k}|\})\left(\widetilde{R}_{\mathbf{k}}+2b_0\right)^{-\frac{n+2a_0}{2}}, \end{aligned}$$

for some constant c_1.

Also, by using Lemma 1, one can show that

$$\begin{aligned} m_{\mathbf{k}}(y) &\ge \{c_2(n\nu_{\mathbf{k}*}\tau_{n,p}+1)\}^{-\frac{|\mathbf{k}|}{2}} \int (\sigma^2)^{-\frac{n+2a_0}{2}-1} \exp\left\{|\mathbf{k}|\left(\frac{2}{\sigma^2}\right)^{1/2} - \sum_{j=1}^{|\mathbf{k}|}\frac{\tau_{n,p}}{\widetilde{\beta}^{*2}_{\mathbf{k},j}} - \frac{\widetilde{R}_{\mathbf{k}}+2b_0}{2\sigma^2}\right\} d\sigma^2 \\ &\ge \{c_2(n\nu_{\mathbf{k}*}\tau_{n,p}+1)\}^{-\frac{|\mathbf{k}|}{2}} \exp\left\{-\sum_{j=1}^{|\mathbf{k}|}\frac{\tau_{n,p}}{\widetilde{\beta}^{*2}_{\mathbf{k},j}}\right\}\left(\widetilde{R}_{\mathbf{k}}+2b_0\right)^{-\frac{n+2a_0}{2}}, \end{aligned}$$

where c_2 is some constant and $\widetilde{\beta}^*_{\mathbf{k},j} \in [\widetilde{\beta}_{\mathbf{k},j}-\widetilde{\epsilon}_n, \widetilde{\beta}_{\mathbf{k},j}+\widetilde{\epsilon}_n] \setminus (-\widetilde{\epsilon}_n, \widetilde{\epsilon}_n)^c$ for $j \in 1,\ldots,|\mathbf{k}|$. These results shows that (S24) holds.

Next, we consider the asymptotic behavior of $\{(\widetilde{R}_{\mathbf{k}}+2b_0)/(\widetilde{R}_{\mathbf{t}}+2b_0)\}^{-n/2-a_0}$ in (S24). Define

ρ_n as the follows:

$$\rho_n = (\widetilde{R}_{\mathbf{t}} + 2b_0)/(n\sigma_0^2) - 1.$$

Since $-\log(1-u) < u/(1-u)$ for $u \in \mathbb{R}$,

$$\begin{aligned} -\log\{(\widetilde{R}_{\mathbf{k}} + 2b_0)/(\widetilde{R}_{\mathbf{t}} + 2b_0)\} &= -\log[1 + (\widetilde{R}_{\mathbf{k}} - \widetilde{R}_{\mathbf{t}})/\{n(1+\rho_n)\sigma_0^2\}] \\ &\leq (\widetilde{R}_{\mathbf{t}} - \widetilde{R}_{\mathbf{k}})/\{n\sigma_0^2(1+u_n)\}, \end{aligned}$$

where $u_n = \rho_n + (\widetilde{R}_{\mathbf{k}} - \widetilde{R}_{\mathbf{t}})/(n\sigma_0^2)$.

Since $(R^*_{\mathbf{k}} - R^*_{\mathbf{u}})/\sigma_0^2 \sim \chi_{|\mathbf{u}\setminus\mathbf{k}|}(\lambda_n)$ with $\lambda_n = \beta_{\mathbf{t}}^{0T} X_{\mathbf{t}}^T (P_{\mathbf{u}} - P_{\mathbf{k}}) X_{\mathbf{t}} \beta_{\mathbf{t}}^0/\sigma_0^2$, by using Lemma 4 one can show that

$$\begin{aligned} P\left(|u_n - \lambda_n/n| > \epsilon\right) &\leq P\left(|\rho_n| > \epsilon/4\right) + P\left\{(R^*_{\mathbf{t}} - R^*_{\mathbf{u}})/(n\sigma_0^2) > \epsilon/4\right\} \\ &\quad + P\left\{\left|(R^*_{\mathbf{k}} - R^*_{\mathbf{u}})/(n\sigma_0^2) - \lambda_n/n\right| > \epsilon/4\right\} + P\left(\eta/2n\sigma_0^2 > \epsilon/4\right) \\ &\leq \exp\{-c'n\} + P\left\{\left|(R^*_{\mathbf{k}} - R^*_{\mathbf{u}})/(n\sigma_0^2) - \lambda_n/n\right| > \epsilon/4\right\} \\ &\leq \exp\{-c''n\}, \end{aligned}$$

for some constant c' and c'', and η is defined in the proof of Theorem 1. Also, by Assumption 5, λ_n/n will be bounded below and above. □

Proof of Corollary 4. Since we showed that the asymptotic equivalence between $Q_{\mathbf{k}}$ and $Q^*_{\mathbf{k}}$ in the proof of Corollary 2, by following exactly same steps in the proof of Theorem 3 we can prove the model selection consistency under piMoM prior densities. □

Proof of Proposition 5. We shall show that for any $\alpha_{\mathbf{k}} = \widehat{\beta}_{\mathbf{k}} + \boldsymbol{\epsilon}_n$ with $\boldsymbol{\epsilon}_n = \{\epsilon_{n,j}\}_{j=1,\ldots,|\mathbf{k}|}$ and $|\epsilon_{n,j}| \succ \epsilon^*_n$ for at least one $j \in \{1,\ldots,|\mathbf{k}|\}$, $P\{g(\alpha_{\mathbf{k}};\mathbf{k}) < g(\widetilde{\beta}^*_{\mathbf{k}};\mathbf{k})\} \to 0$ as n tends to ∞, where $\widetilde{\beta}^*_{\mathbf{k}} \in B(\widehat{\beta}_{\mathbf{k}};\epsilon^*_n)$ with $\epsilon^*_n \asymp (\tau_{n,p}/n)^{1/3}$. More specifically, we set $\widetilde{\beta}^*_{\mathbf{k},j} = \widehat{\beta}_{\mathbf{k},j} + \epsilon^*_n$ for $j \in \mathbf{t}$ and $\widetilde{\beta}^*_{\mathbf{k},j} = \widehat{\beta}_{\mathbf{k},j}$ for $j \in \mathbf{t}^c$. Without loss of generality, we assume that $X_j^{\mathrm{T}} X_j = n$ for $j = 1,\ldots,p$.

Note that

$$\begin{aligned} g(\alpha_{\mathbf{k}};\mathbf{k}) &= ||X_{\mathbf{k}}\alpha_{\mathbf{k}} - X_{\mathbf{k}}\widehat{\beta}_{\mathbf{k}}||_2^2 + \sum_{j=1}^{|\mathbf{k}|}\tau_{n,p}/|\alpha_{\mathbf{k},j}| + D_n \\ &= \sum_{j=1}^{|\mathbf{k}|}\{c_j n\epsilon_{n,j}^2 + \tau_{n,p}/|\widehat{\beta}_{\mathbf{k},j} + \epsilon_{n,j}|\} + D_n, \end{aligned}$$

for some constants c_j such that $C_L < c_j < C_U$ for $j = 1, \ldots, |\mathbf{k}|$, and some randome variable D_n that are not relevant to $\alpha_{\mathbf{k}}$. Then,

$$\begin{aligned} & P\{g(\alpha_{\mathbf{k}};\mathbf{k}) < g(\widetilde{\beta}^*_{\mathbf{k}};\mathbf{k})\} \\ \leq\ & P\left[\sum_{j=1}^{|\mathbf{k}|}\left\{c_j n\epsilon_{n,j}^2 + \frac{\tau_{n,p}}{|\widehat{\beta}_{\mathbf{k},j} + \epsilon_{n,j}|}\right\} < \sum_{j=1}^{|\mathbf{k}|}\left\{c_j n\epsilon_n^{*2} + \frac{\tau_{n,p}}{|\widetilde{\beta}^*_{\mathbf{k},j}|}\right\}\right] \\ \leq\ & P\left[\sum_{j\in S^*\cap S_{\mathbf{k},n}}\left\{c_j n\epsilon_{n,j}^2 + \frac{\tau_{n,p}}{|\widehat{\beta}_{\mathbf{k},j}| + |\epsilon_{n,j}|} - t_{n,j}\right\} < \sum_{j\in S^*\cap S_{\mathbf{k},n}}\left\{c_j n\epsilon_n^{*2} + \frac{\tau_{n,p}}{|\widetilde{\beta}^*_{\mathbf{k},j}|}\right\}\right] \quad \text{(S25)} \\ & + P\left[\sum_{j\in S^*\cap S^c_{\mathbf{k},n}}\left\{c_j n\epsilon_{n,j}^2 + \frac{\tau_{n,p}}{|\widehat{\beta}_{\mathbf{k},j}| + |\epsilon_{n,j}|} - t_{n,j}\right\} < \sum_{j\in S^*\cap S^c_{\mathbf{k},n}}\left\{c_j n\epsilon_n^{*2} + \frac{\tau_{n,p}}{|\widetilde{\beta}^*_{\mathbf{k},j}|}\right\}\right] \quad \text{(S26)} \\ & + P\left[\sum_{j\in S^{*c}}\left\{c_j n\epsilon_{n,j}^2 + \frac{\tau_{n,p}}{|\widehat{\beta}_{\mathbf{k},j}| + |\epsilon_{n,j}|} + \sum_{j\in S^*}\frac{t_{n,j}}{|S^{*c}|}\right\} < \sum_{j\in S^{*c}}\left\{c_j n\epsilon_n^{*2} + \frac{\tau_{n,p}}{|\widetilde{\beta}^*_{\mathbf{k},j}|}\right\}\right], \quad \text{(S27)} \end{aligned}$$

where t_n is an arbitrary sequence such that $t_{n,j} = n^{2/3}\tau_{n,p}^{1/3}\epsilon_{n,j}$, and $S^* = \{j \in \{1, \ldots, p\} : |\epsilon_{n,j}| \succ \epsilon_n^*\}$, and $S_{\mathbf{k},n} = \{j \in \mathbf{k} : |\widehat{\beta}_{\mathbf{k},j}| < \epsilon_n^*\}$. Then, to complete the proof, it is sufficient to show that each of (S25), (S26), and (S27) converges to zero.

Since $n(\widehat{\beta}_{\mathbf{k},j} - \beta^0_{\mathbf{t},j})^2/\sigma^2 \sim \chi_1^2$ for $j = 1, \ldots, |\mathbf{k}|$,

$$P(|\widehat{\beta}_{\mathbf{t},j} - \beta^0_{\mathbf{t},j}| > \zeta_n) \leq (\pi n\zeta_n^2/2)^{-1/2}\exp\{-n\zeta_n^2/(2\sigma^2)\},$$

for any $\zeta_n > 0$. This implies that $S_{\mathbf{k},n} = \mathbf{t}$ at least probability $1 - |\mathbf{t}^c|(\pi n\epsilon_n^{*2}/2)^{-1/2}\exp\{-n\epsilon_n^{*2}/(2\sigma^2)\}$.

Therefore, the equation (S25) can be asymptotically bounded by

$$\begin{aligned} &\sum_{j\in S^*\cap \mathbf{t}} P\left[c_j n\epsilon_{n,j}^2 + \frac{\tau_{n,p}}{2|\epsilon_{n,j}|} - t_{n,j} < c_j n\epsilon_n^{*2} + \frac{\tau_{n,p}}{|\widehat{\beta}_{\mathbf{k},j} + \epsilon_n^*|}\right] \\ \leq \quad &\sum_{j\in S^*\cap \mathbf{t}} P\left[|\widehat{\beta}_{\mathbf{k},j} + \epsilon_n^*| < c\tau_{n,p}(n\epsilon_{n,j}^2 - t_{n,j} + \tau_{n,p}/|\epsilon_{n,j}|)^{-1}\right], \end{aligned}$$

for some positive constant c. Consider Lemma 4 with $\lambda_n = n\epsilon_n^{*2}/\sigma^2$ and $w_n = c^2\tau_{n,p}^2/\{\epsilon_n^{*2}(n\epsilon_{n,j}^2 - t_{n,j} + \tau_{n,p}/|\epsilon_{n,j}|)^2\}$ for $j \in S^* \cap \mathbf{t}$. Since $n\epsilon_{n,j}^2 \succ n^{1/3}\tau_{n,p}^{2/3}$ for $j \in S^*$ implies $w_n \to 0$, Lemma 4 guarantees that the last display is bounded by $c'|S^* \cap \mathbf{t}|\lambda_n^{-1}\exp\{-\lambda_n(1-w_n)^2\}$ for some constant c', which means that (S25) converges to zero as n tends to 0. By following the same steps, one can show that (S26) converges to zero.

Also, (S27) can be asymptotically bounded by

$$\begin{aligned} &\sum_{j\in S^{*c}\cap \mathbf{t}} P\left[c_j n\epsilon_{n,j}^2 + \frac{\tau_{n,p}}{2|\epsilon_{n,j}|} + c\min_{j\in S^*} t_{n,j} < c_j n\epsilon^{*2} + \frac{\tau_{n,p}}{|\widehat{\beta}_{\mathbf{k},j} + \epsilon_n^*|}\right] \\ &+ \sum_{j\in S^{*c}\cap \mathbf{t}^c} P\left[c_j n\epsilon_{n,j}^2 + \frac{\tau_{n,p}}{2|\widehat{\beta}_{\mathbf{k},j} + \epsilon_n^*|} + c\min_{j\in S^*} t_{n,j} < c_j n\epsilon^{*2} + \frac{\tau_{n,p}}{|\widehat{\beta}_{\mathbf{k},j} + \epsilon_n^*|}\right] \\ \leq \quad &\sum_{j\in S^{*c}\cap \mathbf{t}} P\left[|\widehat{\beta}_{\mathbf{k},j} + \epsilon_n^*| < c'\tau_{n,p}(n\epsilon_{n,j}^2 - n\epsilon_n^{*2} + c\min_{j\in S^*} t_{n,j} + \tau_{n,p}/|\epsilon_{n,j}|)^{-1}\right] \\ &+ \sum_{j\in S^{*c}\cap \mathbf{t}^c} P\left[|\widehat{\beta}_{\mathbf{k},j} + \epsilon_n^*| < c''\tau_{n,p}(n\epsilon_{n,j}^2 - n\epsilon_n^{*2} + c\min_{j\in S^*} t_{n,j} + \tau_{n,p}/|\epsilon_{n,j}|)^{-1}/2\right], \end{aligned}$$

where c, c', and c'' are some positive constants. For the first term in the last line of the above display, by setting $\lambda_n = n\epsilon^{*2}/\sigma^2$ and $w_n = c^2\tau_{n,p}^2/\{\epsilon_n^{*2}(n\epsilon_{n,j}^2 - n\epsilon_n^* + c\min_{j\in S^*} t_{n,j} + \tau_{n,p}/|\epsilon_{n,j}|)^2\}$, we can apply Lemma 4. Since $w_n \prec \tau_{n,p}^2(\epsilon_n^* \min_{j\in S^*} t_{n,j})^{-2}$ implies $w_n \to 0$, the first term in the above display converges to zero by Lemma 4. Similarly, the second term also converges to zero. □

3 Laplace Approximations of Marginal Likelihoods

In this section, we provide the Laplace approximation of the marginal likelihoods based on the nonlocal priors. Because closed form expressions for posterior model probabilities based on modified peMoM priors and modified piMoM priors are not available, we estimate the posterior model probabilities using Laplace approximations. For posterior probabilities based on the peMoM priors, an inverse-Gamma density with parameters (a_0, b_0) on σ^2 the Laplace approximation to the marginal density of the data for model $\mathbf{k}$ can be expressed as

$$\pi(\mathbf{k} \mid y) \propto (2\pi)^{|\mathbf{k}|/2} \left|V(\beta_{\mathbf{k}}^*, \sigma^{2*})\right|^{-1/2} \exp\{f(\beta_{\mathbf{k}}^*, \sigma^{2*})\}p(\mathbf{k}), \tag{S28}$$

where

$$\begin{aligned}
(\beta_{\mathbf{k}}^*, \sigma^{2*}) &= \underset{(\beta_{\mathbf{k}}, \sigma^2)}{\operatorname{argmax}} f(\beta_{\mathbf{k}}, \sigma^2) \\
f(\beta_{\mathbf{k}}, \sigma^2) &= -(n/2 + |\mathbf{k}|/2 + a_0 + 1)\log\sigma^2 - (y - X_{\mathbf{k}}\beta_{\mathbf{k}})^T(y - X_{\mathbf{k}}\beta_{\mathbf{k}})/(2\sigma^2) - \beta_{\mathbf{k}}^T\beta_{\mathbf{k}}/(2\sigma^2\tau_{n,p}) \\
&\quad - \sum_{j=1}^{|\mathbf{k}|} \tau_{n,p}/\beta_{\mathbf{k},j}^2 + |\mathbf{k}|(2/\sigma^2)^{1/2} - b_0/\sigma^2 + |\mathbf{k}|(\log\tau_{n,p})/2,
\end{aligned}$$

and $V(\beta_{\mathbf{k}}, \sigma^2)$ is a $(|\mathbf{k}| + 1) \times (|\mathbf{k}| + 1)$ matrix with the following blocks:

$$\begin{aligned}
V_{11} &= X_{\mathbf{k}}^T X_{\mathbf{k}}/\sigma^2 + I_{\mathbf{k}}/\sigma^2\tau_{n,p} + diag\left\{6\tau_{n,p}/\beta_{\mathbf{k},j}^4\right\}_{j=1,\ldots,|\mathbf{k}|} \\
V_{12} &= X_{\mathbf{k}}^T(y - X_{\mathbf{k}}\beta_{\mathbf{k}})/\sigma^4 - \beta_{\mathbf{k}}/\{\sigma^4\tau_{n,p}\} \\
V_{22} &= -(n/2 + |\mathbf{k}|/2 + a_0 + 1)/\sigma^4 + (y - X_{\mathbf{k}}\beta_{\mathbf{k}})^T(y - X_{\mathbf{k}}\beta_{\mathbf{k}})/\sigma^6 - \beta_{\mathbf{k}}^T\beta_{\mathbf{k}}/\tau_{n,p} \\
&\quad -3|\mathbf{k}|2^{1/2}\sigma^{-5}/4 + 2b_0/\sigma^6.
\end{aligned}$$

For the piMoM priors on $\beta_{\mathbf{k}}$, the Laplace approximation of the posterior model probability can be expressed as in (S28), but with

$$
\begin{aligned}
f(\beta_{\mathbf{k}}, \sigma^2) &= -(n/2 + a_0 + 1)\log\sigma^2 - (y - X_{\mathbf{k}}\beta_{\mathbf{k}})^T(y - X_{\mathbf{k}}\beta_{\mathbf{k}})/(2\sigma^2) \\
&\quad - \sum_{j=1}^{|\mathbf{k}|}\left\{r\log(\beta_{\mathbf{k},j}^2) + \tau_{n,p}/\beta_{\mathbf{k},j}^2\right\} + |\mathbf{k}|\left\{(r - 1/2)\log\tau_{n,p} - \log\Gamma(r - 1/2)\right\} - b_0/\sigma^2,
\end{aligned}
$$

and $V(\beta_{\mathbf{k}}, \sigma^2)$ a $(|\mathbf{k}| + 1) \times (|\mathbf{k}| + 1)$ matrix with the following blocks:

$$
\begin{aligned}
V_{11} &= X_{\mathbf{k}}^T X_{\mathbf{k}}/\sigma^2 + diag\left\{6\tau_{n,p}/\beta_{\mathbf{k},j}^4 - 2r/\beta_{\mathbf{k},j}^2\right\}_{j=1,\ldots,|\mathbf{k}|} \\
V_{12} &= X_{\mathbf{k}}^T(y - X_{\mathbf{k}}\beta_{\mathbf{k}})/\sigma^4 \\
V_{22} &= -(n/2 + a_0 + 1)/\sigma^4 + (y - X_{\mathbf{k}}\beta_{\mathbf{k}})^T(y - X_{\mathbf{k}}\beta_{\mathbf{k}})/\sigma^6 + 2b_0/\sigma^6.
\end{aligned}
$$