Supplementary Material for "Variable selection with group structure in competing risks quantile regression" by Kwang Woo Ahn and Soyoung Kim

December 5, 2017

1 Proof of Lemma 2.1

We have the following assumptions: We assume as follows:

(a) $\int_0^L \lambda_0^G(t)dt < \infty$ and $\mathrm{P}\{Y_i(t) = 1\} > 0$ for $t \in [0, L]$, $i = 1, \ldots, n$, and $d_n^4/n \to 0$ as $n \to \infty$.

(b) Z_{ij} is bounded almost surely for all i, j and $\boldsymbol{\alpha}^T\tilde{\mathbf{Z}}$ is bounded almost surely for any $\tilde{\mathbf{Z}}$ and $\boldsymbol{\alpha} \in \mathcal{B}$, where $\mathcal{B}$ is a neighborhood $\boldsymbol{\alpha}_0$.

(c) For $d = 0, 1, 2$, there exists a neighborhood $\mathcal{B}$ of $\boldsymbol{\alpha}_0$ such that $s_G^{(d)}(\boldsymbol{\alpha}, t)$ are continuous functions and $\sup_{t\in(0,L),\boldsymbol{\alpha}\in\mathcal{B}} \|\mathbb{S}_G^{(d)}(\boldsymbol{\alpha}, t) - s_G^{(d)}(\boldsymbol{\alpha}, t)\| \to 0$ in probability.

(d) The matrix $A(\boldsymbol{\alpha}_0) = \int_0^L v_G(\boldsymbol{\alpha}_0, t)s_G^{(0)}(\boldsymbol{\alpha}_0, t)\lambda_0^G(t)dt$ is positive definite, where $v_G(\boldsymbol{\alpha}, t) = s_G^{(2)}(\boldsymbol{\alpha}, t)/s_G^{(0)}(\boldsymbol{\alpha}, t) - e_G(\boldsymbol{\alpha}, t)^{\otimes 2}$ and $e_G(\boldsymbol{\alpha}, t) = s_G^{(1)}(\boldsymbol{\alpha}, t)/s_G^{(0)}(\boldsymbol{\alpha}, t)$.

(e) For all $\boldsymbol{\alpha} \in \mathcal{B}, t \in [0, L]$, $\mathbb{S}_G^{(1)}(\boldsymbol{\alpha}, t) = \partial\mathbb{S}_G^{(0)}(\boldsymbol{\alpha}, t)/\partial\boldsymbol{\alpha}$, and $\mathbb{S}_G^{(2)}(\boldsymbol{\alpha}, t) = \partial^2\mathbb{S}_G^{(0)}(\boldsymbol{\alpha}, t)/(\partial\boldsymbol{\alpha}\partial\boldsymbol{\alpha}^T)$, where $\mathbb{S}_G^{(d)}(\boldsymbol{\alpha}, t)$, $d = 0, 1, 2$ are continuous functions of $\boldsymbol{\alpha} \in \mathcal{B}$ uniformly in $t \in [0, L]$ and are bounded on $\mathcal{B} \times [0, L]$, and $s_G^{(0)}$ is bounded away from zero on $\mathcal{B} \times [0, L]$.

Define $M_i^G(t) = N_i^G(t) - \int_0^t Y_i(u)e^{\boldsymbol{\alpha}_0^T\tilde{\mathbf{Z}}}d\Lambda_0^G(u)$. Then,

$$\hat{\Lambda}_0^G(t : \boldsymbol{\alpha}_0) - \Lambda_0^G(t) = \int_0^t \frac{\sum_{i=1}^n dM_i^G(u)}{n\mathbb{S}_G^{(0)}(\boldsymbol{\alpha}_0, u)}.$$

We have

$$\hat{\Lambda}_0^G(t:\hat{\boldsymbol{\alpha}}) - \Lambda_0^G(t) = \int_0^t \left\{ \frac{1}{n\mathbb{S}_G^{(0)}(\hat{\boldsymbol{\alpha}},u)} - \frac{1}{n\mathbb{S}_G^{(0)}(\boldsymbol{\alpha}_0,u)} \right\} d\sum_{i=1}^n M_i^G(u)$$
$$+ \int_0^t \left\{ \frac{1}{\mathbb{S}_G^{(0)}(\hat{\boldsymbol{\alpha}},u)} - \frac{1}{\mathbb{S}_G^{(0)}(\boldsymbol{\alpha}_0,u)} \right\} \mathbb{S}_G^{(0)}(\boldsymbol{\alpha}_0,u) d\Lambda_0^G(u) \quad (1)$$
$$+ \int_0^t \frac{1}{n\mathbb{S}_G^{(0)}(\boldsymbol{\alpha}_0,u)} d\sum_{i=1}^n M_i^G(u).$$

Let $var\{M_i^G(t)\} = \int_0^t s_G^{(0)}(\boldsymbol{\alpha}_0,u)d\Lambda_0^G(u)$ be $\sigma^2(\boldsymbol{\alpha}_0,t)$. Then, by the proof of Lemma 5 of Ni et al. [1], $n^{-1/2}\sum_{i=1}^n M_i^G(t)/\sigma(\boldsymbol{\alpha}_0,t)$ converges weakly to a tight zero mean Gaussian process with continuous sample paths. In addition, $\sigma(\boldsymbol{\alpha}_0,t)$ is bounded and bounded away from 0 due to Conditions (a), (b) and (e). Similarly to Theorem 1 of Ni et al. [1], using Conditions (a) to (e) we can show $\|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0\| = O_p(\sqrt{d_n/n})$. Then, by Taylor expansion, Conditions (c) and (e), Lemma 1 of Lin [2], the first term of (1) $\times\sqrt{n/d_n}$ is

$$\sqrt{n/d_n}\int_0^t \Big\{ \frac{1}{n\mathbb{S}_G^{(0)}(\hat{\boldsymbol{\alpha}},u)} - \frac{1}{n\mathbb{S}_G^{(0)}(\boldsymbol{\alpha}_0,u)} \Big\} d\sum_{i=1}^n M_i^G(u)$$
$$= \frac{1}{\sqrt{d_n}}\int_0^t \Big\{ \frac{1}{\mathbb{S}_G^{(0)}(\hat{\boldsymbol{\alpha}},u)} - \frac{1}{\mathbb{S}_G^{(0)}(\boldsymbol{\alpha}_0,u)} \Big\} \sigma(\boldsymbol{\alpha}_0,t) d\frac{\sum_{i=1}^n M_i^G(u)}{\sqrt{n}\sigma(\boldsymbol{\alpha}_0,t)}$$
$$= \frac{1}{\sqrt{d_n}}\int_0^t \Big\{ -\frac{\mathbb{S}_G^{(1)}(\boldsymbol{\alpha}_*,u)}{\mathbb{S}_G^{(0)}(\boldsymbol{\alpha}_*,u)^2}(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0) \Big\} \sigma(\boldsymbol{\alpha}_0,t) d\frac{\sum_{i=1}^n M_i^G(u)}{\sqrt{n}\sigma(\boldsymbol{\alpha}_0,t)}$$
$$= O_p(\frac{1}{\sqrt{n}}) = o_p(1),$$

where $\boldsymbol{\alpha}_*$ lies between $\hat{\boldsymbol{\alpha}}$ and $\boldsymbol{\alpha}_0$, which holds uniformly in t. This shows the first term of (1) is $o_P(\sqrt{d_n/n})$ uniformly in t. The second term of (1) $\times\sqrt{n/d_n}$ converges to $-\sqrt{n/d_n}\int_0^t \{e(\boldsymbol{\alpha}_0,u)d\Lambda_0^G(u)\}^T(\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0)$ in probability by Taylor expansion and Conditions (a), (c), and (e). Because of $\|\hat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0\| = O_p(\sqrt{d_n/n})$ and the boundedness of $\Lambda_0^G(t)$ and $e(\boldsymbol{\alpha}_0,t)$, the second term of (1) is $O_p(\sqrt{d_n/n})$ uniformly in t. The third term of (1) $\times\sqrt{n/d_n}$ converges to

$$\frac{1}{\sqrt{d_n}}\int_0^t \frac{1}{s_G^{(0)}(\boldsymbol{\alpha}_0,u)} \sigma(\boldsymbol{\alpha}_0,t) d\frac{\sum_{i=1}^n M_i^G(u)}{\sqrt{n}\sigma(\boldsymbol{\alpha}_0,t)}$$

in probability by Condition (c). Because $s_G^{(0)}(\boldsymbol{\alpha}_0,u)$ and $\sigma(\boldsymbol{\alpha}_0,t)$ are bounded, and bounded away from 0, the third term of (1) $\times\sqrt{n/d_n}$ is $O_p(1/\sqrt{d_n}) = o_p(1)$ uniformly in t and thus the third term of (1) is $o_P(\sqrt{d_n/n})$ uniformly in t. Therefore, $\sup_t|\hat{\Lambda}_0^G(t:\hat{\boldsymbol{\alpha}}) - \Lambda_0^G(t)| = O_p(\sqrt{d_n/n})$. By Taylor expansion and the consistency of $\hat{\Lambda}_0^G(t:\hat{\boldsymbol{\alpha}})$, $\sqrt{n/d_n}\{\hat{G}(t|\tilde{\mathbf{Z}}) - G(t|\tilde{\mathbf{Z}})\}$ is asymptotically equivalent to

$$
\begin{aligned}
&-\sqrt{n/d_n}G(t|\tilde{\mathbf{Z}})\{\hat{\Lambda}_0^G(t:\hat{\boldsymbol{\alpha}})\exp(\hat{\boldsymbol{\alpha}}^T\tilde{\mathbf{Z}})-\Lambda_0^G(t)\exp(\boldsymbol{\alpha}_0^T\tilde{\mathbf{Z}})\}\\
&=-\sqrt{n/d_n}G(t|\tilde{\mathbf{Z}})[\{\hat{\Lambda}_0^G(t:\hat{\boldsymbol{\alpha}})\exp(\hat{\boldsymbol{\alpha}}^T\tilde{\mathbf{Z}})-\hat{\Lambda}_0^G(t:\hat{\boldsymbol{\alpha}})\exp(\boldsymbol{\alpha}_0^T\tilde{\mathbf{Z}})\}\\
&\quad+\{\hat{\Lambda}_0^G(t:\hat{\boldsymbol{\alpha}})\exp(\boldsymbol{\alpha}_0^T\tilde{\mathbf{Z}})-\Lambda_0^G(t)\exp(\boldsymbol{\alpha}_0^T\tilde{\mathbf{Z}})\}]\\
&=-\sqrt{n/d_n}G(t|\tilde{\mathbf{Z}})[\hat{\Lambda}_0^G(t:\hat{\boldsymbol{\alpha}})\exp(\boldsymbol{\alpha}_*^T\tilde{\mathbf{Z}})\tilde{\mathbf{Z}}^T(\hat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}_0)+\exp(\boldsymbol{\alpha}_0^T\tilde{\mathbf{Z}})\{\hat{\Lambda}_0^G(t:\hat{\boldsymbol{\alpha}})-\Lambda_0^G(t)\}],
\end{aligned}
$$

where $\boldsymbol{\alpha}_*$ lies between $\boldsymbol{\alpha}_0$ and $\hat{\boldsymbol{\alpha}}$. Because of $\sup_t|\hat{\Lambda}_0^G(t:\hat{\boldsymbol{\alpha}})-\Lambda_0^G(t)|=O_p(\sqrt{d_n/n})$, $\|\hat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}_0\|=O_p(\sqrt{d_n/n})$, and Conditions (a) & (b), we have $\sup_t|\hat{G}(t|\tilde{\mathbf{Z}})-G(t|\tilde{\mathbf{Z}})|=O_p(\sqrt{d_n/n})$.

2 Proofs of Lemma 2.3 and Theorem 2.5

Let $\mathbf{S}_n^G(\mathbf{b},\tau)=n^{-1/2}\sum_{i=1}^n\mathbf{Z}_i[I\{X_i\le g(X_i^T\mathbf{b})\}I(\delta_i=1)/G(X_i|\tilde{\mathbf{Z}})-\tau]$. Because of (C1), we have $\sup_{t<\omega}|\hat{G}(t|\tilde{\mathbf{Z}})-G(t|\tilde{\mathbf{Z}})|=o_p(n^{-1/2+q}d_n^{1/2})$ for any $q>0$. Consider

$$
\begin{aligned}
&n^{-1/2}\mathbf{S}_n(\mathbf{b},\tau)-n^{-1/2}\mathbf{S}_n^G(\mathbf{b},\tau)\\
&=n^{-1}\Big(\sum_{i=1}^n\mathbf{Z}_i\Big[\frac{I\{X_i\le g(\mathbf{Z}_i^T\mathbf{b})\}I(\delta_i=1)}{\hat{G}(X_i|\tilde{\mathbf{Z}}_i)}-\tau\Big]-\sum_{i=1}^n\mathbf{Z}_i\Big[\frac{I\{X_i\le g(\mathbf{Z}_i^T\mathbf{b})\}I(\delta_i=1)}{G(X_i|\tilde{\mathbf{Z}}_i)}-\tau\Big]\Big)\\
&=n^{-1}\sum_{i=1}^n\mathbf{Z}_i\Big[\frac{I\{X_i\le g(\mathbf{Z}_i^T\mathbf{b})\}I(\delta_i=1)\{G(X_i|\tilde{\mathbf{Z}}_i)-\hat{G}(X_i|\tilde{\mathbf{Z}}_i)\}}{\hat{G}(X_i|\tilde{\mathbf{Z}}_i)G(X_i|\tilde{\mathbf{Z}}_i)}\Big].
\end{aligned}
$$

Because of the boundedness of the d_n elements of $\mathbf{Z}_i$, (C1), and Lemma 2.1, for $0<q\le 1/8$ we have

$$
\sup_{\mathbf{b}}\|n^{-1/2}\mathbf{S}_n(\mathbf{b},\tau)-n^{-1/2}\mathbf{S}_n^G(\mathbf{b},\tau)\|=o_p(n^{-1/2+q}d_n^{1/2}d_n)=o_p(1). \tag{2}
$$

By Chebyshev's inequality, (C2), and (C3) for any ϵ,

$$
\begin{aligned}
&P(\|n^{-1/2}\mathbf{S}_n^G(\mathbf{b},\tau)-E\{n^{-1/2}\tilde{\mathbf{S}}_n(\mathbf{b},\tau)\}\|\ge\epsilon)\\
&\le\frac{1}{\epsilon^2}E\|n^{-1/2}\mathbf{S}_n^G(\mathbf{b},\tau)-E\{n^{-1/2}\tilde{\mathbf{S}}_n(\mathbf{b},\tau)\}\|^2\\
&=O(\frac{d_n}{n\epsilon^2})=o(1),
\end{aligned}
$$

which holds uniformly in $\mathbf{b}\in\mathcal{B}(\rho_0)$. Similarly, by Chebyshev's inequality, (C3), (C4), (C6), and (C7), for any ϵ,

$$
P(\|n^{-1/2}\bigtriangledown\tilde{\mathbf{S}}_n(\mathbf{b},\tau)-\mathbf{H}(\mathbf{b})\|\ge\frac{\epsilon}{d_n})\le O(\frac{d_n^4}{n\epsilon^2})=o(1), \tag{3}
$$

which holds uniformly in $\mathbf{b}\in\mathcal{B}(\rho_0)$.

For $0 < q \le 1/8$, by (2) and (C2) we have

$$\begin{aligned}\frac{1}{n}U_n(\mathbf{b},\tau) &= \frac{1}{n}\Big\{\sum_{i=1}^{n} I(\delta_i=1)\Big|\frac{g^{-1}(X_i)-\mathbf{b}^T\mathbf{Z}_i}{\hat{G}(X_i|\tilde{\mathbf{Z}})}\Big| \\ &\quad + \Big|M-\mathbf{b}^T\sum_{i=1}^{n}\frac{-\mathbf{Z}_i I(\delta_i=1)}{\hat{G}(X_i|\tilde{\mathbf{Z}})}\Big| + \Big|M-\mathbf{b}^T\sum_{i=1}^{n}2\mathbf{Z}_i\tau\Big|\Big\} \\ &= \frac{1}{n}\Big\{\sum_{i=1}^{n} I(\delta_i=1)\Big|\frac{g^{-1}(X_i)-\mathbf{b}^T\mathbf{Z}_i}{G(X_i|\tilde{\mathbf{Z}})}\Big| + \Big|M-\mathbf{b}^T\sum_{i=1}^{n}\frac{-\mathbf{Z}_i I(\delta_i=1)}{G(X_i|\tilde{\mathbf{Z}})}\Big| \\ &\quad + \Big|M-\mathbf{b}^T\sum_{i=1}^{n}2\mathbf{Z}_i\tau\Big|\Big\} + o_p(n^{-1/2+q}d_n^{3/2}),\end{aligned}$$

where $o_p(n^{-1/2+q}d_n^{3/2}) = o_p(1)$. Let

$$\begin{aligned}\tilde{U}_n(\mathbf{b},\tau) &= \sum_{i=1}^{n} I(\delta_i=1)\Big|\frac{g^{-1}(X_i)-\mathbf{b}^T\mathbf{Z}_i}{G(X_i|\tilde{\mathbf{Z}})}\Big| + \Big|M-\mathbf{b}^T\sum_{i=1}^{n}\frac{-\mathbf{Z}_i I(\delta_i=1)}{G(X_i|\tilde{\mathbf{Z}})}\Big| \\ &\quad + \Big|M-\mathbf{b}^T\sum_{i=1}^{n}2\mathbf{Z}_i\tau\Big|.\end{aligned}$$

First of all, we show consistency of $\hat{\boldsymbol{\beta}}(\tau)$. It is sufficient to show that for any $\epsilon > 0$, there exists a large constant C such that

$$P\Big[\inf_{\|\mathbf{u}\|=1} W_n\{\boldsymbol{\beta}_0(\tau)+C\alpha_n\mathbf{u},\tau\} > W_n\{\boldsymbol{\beta}_0(\tau),\tau\}\Big] > 1-\epsilon,$$

where $\|\mathbf{u}\| = \mathbf{1}$. Let $\alpha_n = \sqrt{d_n/n}$. We have

$$\begin{aligned}&\frac{1}{n}[W_n\{\boldsymbol{\beta}_0(\tau)+C\alpha_n\mathbf{u},\tau\} - W_n\{\boldsymbol{\beta}_0(\tau),\tau\}] \\ &= \frac{1}{n}[U_n\{\boldsymbol{\beta}_0(\tau)+C\alpha_n\mathbf{u},\tau\} - U_n\{\boldsymbol{\beta}_0(\tau),\tau\}] + D_n,\end{aligned}$$

where

$$D_n = \frac{\lambda_n}{n}\sum_{k=1}^{K} c_k\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^{\nu}}\Big)^{\gamma} - \frac{\lambda_n}{n}\sum_{k=1}^{K} c_k\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^{\nu}}\Big)^{\gamma}.$$

By Taylor expansion, for $0 < q \le 1/4$ we have

$$\begin{aligned}&\frac{1}{n}[U_n\{\boldsymbol{\beta}_0(\tau)+C\alpha_n\mathbf{u},\tau\} - U_n\{\boldsymbol{\beta}_0(\tau),\tau\}] \\ &= \frac{1}{n}C\alpha_n\mathbf{u}^T \bigtriangledown \tilde{U}_n\{\boldsymbol{\beta}_0(\tau),\tau\} + \frac{C^2}{2n}\alpha_n^2\mathbf{u}^T \bigtriangledown^2 \tilde{U}_n\{\boldsymbol{\beta}^*(\tau),\tau\}\mathbf{u} + o_p(n^{-1/2+q}d_n). \\ &= R_1 + R_2 + o_p(1),\end{aligned}$$

where $\boldsymbol{\beta}^*(\tau)$ is between $\boldsymbol{\beta}_0(\tau) + \alpha_n \mathbf{u}$ and $\boldsymbol{\beta}_0(\tau)$. By the Cauchy-Schwartz inequality, (C1), and (C2),

$$R_1 = \frac{1}{n} C\alpha_n \mathbf{u}^T \bigtriangledown \tilde{U}_n\{\boldsymbol{\beta}_0(\tau), \tau\} = \frac{1}{n} C\alpha_n \|\mathbf{u}\| \| \bigtriangledown \tilde{U}_n\{\boldsymbol{\beta}_0(\tau), \tau\}\| = CO_p(\alpha_n^2)\|\mathbf{u}\|.$$

By $\bigtriangledown \tilde{U}_n\{\mathbf{b}, \tau\} = \mathbf{S}_n^G\{\mathbf{b}, \tau\} = \tilde{\mathbf{S}}_n\{\mathbf{b}, \tau\} + o_p(1)$, (3), and (C4), $R_2 = C^2\alpha_n^2 \mathbf{u}^T \bigtriangledown^2 \tilde{U}\{\boldsymbol{\beta}_0(\tau), \tau\}\mathbf{u}\{1 + o_p(1)\}$. Therefore, R_1 is of order $C\alpha_n^2$ and R_2 is of order $C^2\alpha_n^2$. If we choose sufficiently large C, $R_2 > 0$ dominates R_1, which shows Lemma 2.3.

Next, D_n can be written as:

$$\begin{aligned}
D_n &= \frac{\lambda_n}{n} \sum_{k=1}^{K} c_k \Big(\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau) + C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^\gamma - \frac{\lambda_n}{n} \sum_{k=1}^{K} c_k \Big(\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^\gamma \\
&= \Big\{ \frac{\lambda_n}{n} \sum_{k=1}^{K_1} c_k \Big(\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau) + C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^\gamma - \frac{\lambda_n}{n} \sum_{k=1}^{K_1} c_k \Big(\sum_{j \in A_k \cap B_1} \frac{|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^\gamma \Big\} \\
&\quad + \frac{\lambda_n}{n} \sum_{k=K_1+1}^{K} c_k \Big(\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau) + C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^\gamma \\
&= I_1 + I_2.
\end{aligned}$$

Consider I_1 first. We consider 2 cases: Case i) when $\beta_{j,0}(\tau) \neq 0$ for all $j \in A_k$ for all $k \in \{1, 2, \ldots, K_1\}$; and Case ii) there is at least one $\beta_{j,0}(\tau)$ such that $\beta_{j,0}(\tau) = 0$ for some $j \in A_k$ for some $k \in \{1, 2, \ldots, K_1\}$.

For Case i), we assume that $\beta_{j,0}(\tau) \neq 0$ for all $j \in A_k$ for all $k \in \{1, 2, \ldots, K_1\}$. Since $b^\gamma - a^\gamma \leq 2(b-a)b^{\gamma-1}$ for $0 < a \leq b$ and $|\beta_{j,0}(\tau)| = O_p\{(d_n/n)^{\nu_1/2}\}$ with $0 < \nu_1 < 1$ as in (C8b) for $j \in B_1$ and any $\tau \in [\tau_L, \tau_U]$, we have that for sufficiently large n such that $C\alpha_n \leq |\beta_{j,0}(\tau)|$ for all $j \in A_k$ and any $\tau \in [\tau_L, \tau_U]$,

$$\begin{aligned}
&\frac{\lambda_n}{n} \sum_{k=1}^{K_1} c_k \Big(\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau) + C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^\gamma - \frac{\lambda_n}{n} \sum_{k=1}^{K_1} c_k \Big(\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^\gamma \\
&\leq \frac{\lambda_n}{n} \sum_{k=1}^{K_1} c_k \Big(\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau)| + C|u_j|\alpha_n}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^\gamma - \frac{\lambda_n}{n} \sum_{k=1}^{K_1} c_k \Big(\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^\gamma \\
&\leq 2\frac{\lambda_n}{n} \sum_{k=1}^{K_1} c_k \Big\{ \Big(\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau)| + C|u_j|\alpha_n}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^{\gamma-1} \sum_{j \in A_k} \frac{C|u_j|\alpha_n}{|\tilde{\beta}_j(\tau)|^\nu} \Big\} \qquad (4) \\
&\leq 2\frac{\lambda_n}{n} \sum_{k=1}^{K_1} c_k \Big\{ \Big(\sum_{j \in A_k} \frac{2|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^{\gamma-1} \sum_{j \in A_k} \frac{C|u_j|\alpha_n}{|\tilde{\beta}_j(\tau)|^\nu} \Big\} \\
&\leq 2\alpha_n^2 \lambda_n (d_n n)^{-1/2} \sum_{k=1}^{K_1} c_k \Big\{ \Big(\sum_{j \in A_k} \frac{2|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu} \Big)^{\gamma-1} \sum_{j \in A_k} \frac{C|u_j|}{|\tilde{\beta}_j(\tau)|^\nu} \Big\}.
\end{aligned}$$

Using $|\tilde{\beta}_j(\tau)|^\nu \to_p |\beta_{j,0}(\tau)|^\nu \neq 0$ for $j \in B_1$, (C8b) and (C9b), this term is dominated by $R_2 > 0$, where $\to_p$ indicates convergence in probability.

For Case ii), assume that there are $\beta_{j,0}(\tau)$'s such that $\beta_{j,0}(\tau) = 0$ for some $j \in A_k$ for some $k \in \{1, 2, \dots, K_1\}$. Consider

$$\begin{aligned}
&\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau) + C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu} - \sum_{j \in A_k \cap B_1} \frac{|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu} \\
&= \Big(\sum_{j \in A_k \cap B_1} \frac{|\beta_{j,0}(\tau) + C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu} - \sum_{j \in A_k \cap B_1} \frac{|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu} \Big) + \sum_{j \in A_k \cap B_2} \frac{|\beta_{j,0}(\tau) + C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu} \\
&= \Big(\sum_{j \in A_k \cap B_1} \frac{|\beta_{j,0}(\tau) + C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu} - \sum_{j \in A_k \cap B_1} \frac{|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu} \Big) + \sum_{j \in A_k \cap B_2} \frac{C\alpha_n |u_j|}{|\tilde{\beta}_j(\tau)|^\nu} \\
&= I_{11} + I_{12}.
\end{aligned}$$

We have

$$\begin{aligned}
|I_{11}| &\leq \sum_{j \in A_k \cap B_1} \frac{C\alpha_n |u_j|}{|\tilde{\beta}_j(\tau)|^\nu}, \\
|I_{12}| &= \sum_{j \in A_k \cap B_2} \frac{C\alpha_n^{1-\nu} |u_j|}{|\sqrt{n/d_n}\tilde{\beta}_j(\tau)|^\nu}.
\end{aligned} \tag{5}$$

Since $\tilde{\beta}_j(\tau)$ converges in probability to non-zero $\beta_{j,0}(\tau) = O_p\{(d_n/n)^{\nu_1/2}\}$ for $j \in B_1$ and $\max_k |A_k \cap B_1| = O\{(n/d_n)^{\nu_2/2}\}$, $|I_{11}| \leq O_p(\alpha_n^{1-\nu\nu_1-\nu_2}) = o_p(\alpha_n^{1-\nu})$ because $\nu\nu_1 + \nu_2 < \nu$ as in (C8b). On the other hand, because $\sqrt{n/d_n}\tilde{\beta}_j(\tau) = O_p(1)$ for $j \in B_2$, $|I_{12}|$ is at least $O_p(\alpha_n^{1-\nu})$. Thus, $I_{12} > 0$ dominates I_{11} as $n \to \infty$. Therefore, for sufficiently large n,

$$\sum_{j \in A_k} \frac{|\beta_{j,0}(\tau) + C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu} > \sum_{j \in A_k \cap B_1} \frac{|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu}.$$

Then, due to $\gamma b^{\gamma-1}(b-a) \le b^\gamma - a^\gamma$ for $0 \le a \le b$, for sufficiently large n we have

$$
\begin{aligned}
&\frac{\lambda_n}{n}\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu}\Big)^\gamma - \frac{\lambda_n}{n}\Big(\sum_{j\in A_k\cap B_1}\frac{|\beta_{j,0}(\tau)|}{|\tilde\beta_j(\tau)|^\nu}\Big)^\gamma \\
&\ge \gamma\frac{\lambda_n}{n}\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu}\Big)^{\gamma-1}\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu} - \sum_{j\in A_k\cap B_1}\frac{|\beta_{j,0}(\tau)|}{|\tilde\beta_j(\tau)|^\nu}\Big) \\
&= \gamma\frac{\lambda_n}{n}\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu}\Big)^{\gamma-1}\Big(\sum_{j\in A_k\cap B_1}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j| - |\beta_{j,0}(\tau)|}{|\tilde\beta_j(\tau)|^\nu} \\
&\quad + \sum_{j\in A_k\cap B_2}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu}\Big) \\
&= \gamma\frac{\lambda_n}{n}\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu}\Big)^{\gamma-1}\sum_{j\in A_k\cap B_1}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j| - |\beta_{j,0}(\tau)|}{|\tilde\beta_j(\tau)|^\nu} \\
&\quad + \gamma\frac{\lambda_n}{n}\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu}\Big)^{\gamma-1}\sum_{j\in A_k\cap B_2}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu} \\
&= I_{13} + I_{14}.
\end{aligned}
$$

By similar argument to (5), I_{14} dominates I_{13}. Consider I_{14}:

$$
\begin{aligned}
&\gamma\frac{\lambda_n}{n}\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu}\Big)^{\gamma-1}\sum_{j\in A_k\cap B_2}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu} \\
&= \gamma\frac{\lambda_n}{n}\Big(\frac{n}{d_n}\Big)^{\nu/2}\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu}\Big)^{\gamma-1}\sum_{j\in A_k\cap B_2}\frac{C\alpha_n|u_j|}{|\sqrt{n/d_n}\tilde\beta_j(\tau)|^\nu} \\
&= \gamma\frac{\lambda_n}{n}\Big(\frac{n}{d_n}\Big)^{\nu/2}\Big(\frac{d_n}{n}\Big)^{1/2}\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu}\Big)^{\gamma-1}\sum_{j\in A_k\cap B_2}\frac{C|u_j|}{|\sqrt{n/d_n}\tilde\beta_j(\tau)|^\nu} \\
&= \gamma\alpha_n^2\lambda_n n^{(\nu-1)/2}d_n^{-(\nu+1)/2}\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde\beta_j(\tau)|^\nu}\Big)^{\gamma-1}\sum_{j\in A_k\cap B_2}\frac{C|u_j|}{|\sqrt{n/d_n}\tilde\beta_j(\tau)|^\nu},
\end{aligned}
\tag{6}
$$

where $\sqrt{n/d_n}\tilde\beta_j(\tau) = O_p(1)$. Because $\lambda_n n^{(\nu-1)/2}d_n^{-(1+\nu)/2} \to \infty$ by (C9b), $I_{14} > 0$ dominates R_1 and R_2. Hence, by (4) and (6), if there exists at least one $\beta_{j,0}(\tau)$ equal to 0, $I_1 > 0$ dominates R_1 and R_2 for sufficiently large n.

Next, consider I_2:

$$\begin{aligned}
&\frac{\lambda_n}{n}\sum_{k=K_1+1}^{K} c_k\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^{\nu}}\Big)^{\gamma}\\
&=\frac{\lambda_n}{n}\Big(\frac{n}{d_n}\Big)^{\nu\gamma/2}\sum_{k=K_1+1}^{K} c_k\Big(\sum_{j\in A_k\cap B_2}\frac{C\alpha_n|u_j|}{(\sqrt{n/d_n}|\tilde{\beta}_j(\tau)|)^{\nu}}\Big)^{\gamma} \qquad (7)\\
&=\alpha^2\lambda_n n^{\gamma(\nu-1)/2}d_n^{-1+\gamma(1-\nu)/2}\sum_{k=K_1+1}^{K} c_k\Big(\sum_{j\in A_k\cap B_2}\frac{C|u_j|}{(\sqrt{n/d_n}|\tilde{\beta}_j(\tau)|)^{\nu}}\Big)^{\gamma}.
\end{aligned}$$

Because $\lambda_n n^{\gamma(\nu-1)/2}d_n^{-1+\gamma(1-\nu)/2}\to\infty$ by (C9b), $I_2>0$ dominates R_1 and R_2.

Therefore, by (4), (6), and (7), for sufficiently large n,

$$W_n\{\boldsymbol{\beta}_0(\tau)+C\alpha_n\mathbf{u},\tau\}-W_n\{\boldsymbol{\beta}_0(\tau),\tau\}>0$$

which proves (1) of Theorem 2.4.

Next, we show variable selection consistency. Let $\boldsymbol{\beta}(\tau)=\boldsymbol{\beta}_0(\tau)+C\alpha_n\mathbf{u}$. Then, we have

$$\begin{aligned}
&\frac{1}{n}W_n[\{\boldsymbol{\beta}_{B_1}(\tau)^T,\mathbf{0}^T\}^T,\tau]-\frac{1}{n}W_n[\{\boldsymbol{\beta}_{B_1}(\tau)^T,\boldsymbol{\beta}_{B_2}(\tau)^T\}^T,\tau]\\
&=\frac{1}{n}\Big(W_n[\{\boldsymbol{\beta}_{B_1}(\tau)^T,\mathbf{0}^T\}^T,\tau]-W_n[\{\boldsymbol{\beta}_{B_1,0}(\tau)^T,\mathbf{0}^T\}^T,\tau]\Big)\\
&\quad-\frac{1}{n}\Big(W_n[\{\boldsymbol{\beta}_{B_1}(\tau)^T,\boldsymbol{\beta}_{B_2}(\tau)^T\}^T,\tau]-W_n[\{\boldsymbol{\beta}_{B_1,0}(\tau)^T,\mathbf{0}^T\}^T,\tau]\Big)\\
&=\frac{1}{n}\Big(U_n[\{\boldsymbol{\beta}_{B_1}(\tau)^T,\mathbf{0}^T\}^T,\tau]-U_n[\{\boldsymbol{\beta}_{B_1,0}(\tau)^T,\mathbf{0}^T\}^T,\tau]\Big)\\
&\quad-\frac{1}{n}\Big(U_n[\{\boldsymbol{\beta}_{B_1}(\tau)^T,\boldsymbol{\beta}_{B_2}(\tau)^T\}^T,\tau]-U_n[\{\boldsymbol{\beta}_{B_1,0}(\tau)^T,\mathbf{0}^T\}^T,\tau]\Big)\\
&\quad+\Big\{\frac{\lambda_n}{n}\sum_{k=1}^{K_1}c_k\Big(\sum_{j\in A_k\cap B_1}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^{\nu}}\Big)^{\gamma}-\frac{\lambda_n}{n}\sum_{k=1}^{K}c_k\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^{\nu}}\Big)^{\gamma}\Big\}\\
&=J_1-J_2+J_3.
\end{aligned}$$

It suffices to show that for sufficiently large n,

$$\frac{1}{n}W_n[\{\boldsymbol{\beta}_{B_1}(\tau)^T,\mathbf{0}^T\}^T,\tau]-\frac{1}{n}W_n[\{\boldsymbol{\beta}_{B_1}(\tau)^T,\boldsymbol{\beta}_{B_2}(\tau)^T\}^T,\tau]<0.$$

As in the proof of consistency of the adaptive group bridge estimator, J_1 and J_2 are of order $C^2\alpha_n^2$.

Consider J_3:

$$\begin{aligned}
&\frac{\lambda_n}{n}\sum_{k=1}^{K_1} c_k\Big(\sum_{j\in A_k\cap B_1}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^\gamma - \frac{\lambda_n}{n}\sum_{k=1}^{K} c_k\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^\gamma \\
&= \frac{\lambda_n}{n}\sum_{k=1}^{K_1} c_k\Big(\sum_{j\in A_k\cap B_1}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^\gamma - \frac{\lambda_n}{n}\sum_{k=1}^{K_1} c_k\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^\gamma \\
&\quad - \frac{\lambda_n}{n}\sum_{k=K_1+1}^{K} c_k\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^\gamma \\
&= J_{31} - J_{32} - J_{33}
\end{aligned} \tag{8}$$

Because $\gamma b^{\gamma-1}(b-a) \le b^\gamma - a^\gamma$ for $0 \le a \le b$, similarly to (6) we have

$$\begin{aligned}
&J_{32} - J_{31} \\
&= \frac{\lambda_n}{n}\sum_{k=1}^{K_1} c_k\Big(\sum_{j\in A_k}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^\gamma - \frac{\lambda_n}{n}\sum_{k=1}^{K_1} c_k\Big(\sum_{j\in A_k\cap B_1}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^\gamma \\
&\ge \frac{\lambda_n}{n}\sum_{k=1}^{K_1} \gamma c_k\Big(\sum_{j\in A_k\cap B_1}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^{\gamma-1}\Big(\sum_{j\in A_k\cap B_2}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big) \\
&= \alpha^2\lambda_n n^{(\nu-1)/2} d_n^{-(\nu+1)/2}\sum_{k=1}^{K_1} \gamma c_k\Big(\sum_{j\in A_k\cap B_1}\frac{|\beta_{j,0}(\tau)+C\alpha_n u_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^{\gamma-1} \\
&\quad \times \sum_{j\in A_k\cap B_2}\frac{C|u_j|}{|\sqrt{n/d_n}\tilde{\beta}_j(\tau)|^\nu}.
\end{aligned}$$

Because $\sqrt{n/d_n}\tilde{\beta}_j(\tau) = O_p(1)$ for $j \in B_2$ and $\lambda_n n^{(\nu-1)/2} d_n^{-(\nu+1)/2} \to \infty$ by (C9b), $J_{31} - J_{32} < 0$ dominates J_1 and J_2. Similarly to (7), $-J_{33} < 0$ also dominates J_1 and J_2. Therefore, for sufficiently large n,

$$\frac{1}{n}W_n[\{\boldsymbol{\beta}_{B_1}(\tau)^T, \mathbf{0}^T\}^T, \tau] - \frac{1}{n}W_n[\{\boldsymbol{\beta}_{B_1}(\tau)^T, \boldsymbol{\beta}_{B_2}(\tau)^T\}^T, \tau] < 0,$$

which shows individual variable selection consistency.

To show the asymptotics $\hat{\boldsymbol{\beta}}_{B_1}(\tau)$, we study the asymptotic normality of $\tilde{\boldsymbol{\beta}}(\tau)$, which is the solution of $\mathbf{S}_n(\mathbf{b},\tau) = \mathbf{0}$, when d_n is fixed. Using (5.1) and (5.2) of He et al. [3], we have

$$\begin{aligned}
\frac{1}{\hat{G}(t|\mathbf{Z}_i)} - \frac{1}{G(t|\mathbf{Z}_i)} = \frac{1}{nG(t|\mathbf{Z}_i)}\sum_{i=1}^n \int_0^L \Big[&\mathbf{h}^T(t,0,\mathbf{Z}_i)\mathbf{A}(\boldsymbol{\alpha}_0)^{-1}\{\mathbf{Z}_i - e_G(\boldsymbol{\alpha}_0, t)\} \\
&+ \frac{e^{\boldsymbol{\alpha}_0^T\mathbf{Z}_i} I(u \le t)}{s_G^{(0)}(\boldsymbol{\alpha}_0, u)}\Big] dM_i^G(u) + o_p(n^{-1/2}).
\end{aligned} \tag{9}$$

Using (9), similarly to Peng and Fine [4], we can show $\mathbf{S}_n(\boldsymbol{\beta}_0(\tau), \tau)$ is asymptotically equivalent to $n^{-1/2}\sum_{i=1}^n \eta_i(\tau)$ and $n^{1/2}\{\tilde{\boldsymbol{\beta}}_0(\tau) - \boldsymbol{\beta}_0(\tau)\}$ is asymptotically equivalent to $n^{-1/2}\sum_{i=1}^n \mathbf{H}\{\boldsymbol{\beta}_0(\tau)\}^{-1}\eta_i(\tau)$. Therefore, by the multivariate central limit theorem, $n^{1/2}\{\tilde{\boldsymbol{\beta}}_0(\tau) - \boldsymbol{\beta}_0(\tau)\}$ converges to $N(\mathbf{0}, \mathbf{H}\{\boldsymbol{\beta}_0(\tau)\}^{-1}E\{\eta_1(\tau)\eta_1(\tau)^T\}\mathbf{H}\{\boldsymbol{\beta}_0(\tau)\}^{-1})$ in distribution for fixed d_n.

Next, we show the asymptotics of $\hat{\boldsymbol{\beta}}_{B_1}(\tau)$, where $\boldsymbol{\beta}_{B_1}(\tau) = \{\beta_{j,0}(\tau) + C\alpha_n u_j; j \in B_1\}^T$. Let b_i be the dimension of B_i for $i = 1, 2$. Note that Theorem 2.5 works on restricted β_j's and $Z_{i,j}$ for $j \in B_1$. Thus, we have

$$\|\hat{\boldsymbol{\beta}}_{B_1}(\tau) - \boldsymbol{\beta}_{B_1,0}(\tau)\| = O_p(n^{-1}),\ \|\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}_0(\tau)\| = O_p(n^{-1}).$$

Consider $V_{1n} = W_n\{\boldsymbol{\beta}_0(\tau) + n^{-1/2}(\mathbf{a}^T, \mathbf{0}^T)^T, \tau\} - W_n\{\boldsymbol{\beta}_0(\tau), \tau\}$, where $\mathbf{a} = (a_1, \ldots, a_{b_1})^T$ is a b_1–dimensional vector and $\mathbf{0}$ is a b_2–dimensional zero vector. From the consistency of $\hat{\boldsymbol{\beta}}(\tau)$, we have $\hat{\boldsymbol{\beta}}(\tau) - \boldsymbol{\beta}(\tau) = n^{-1/2}(\hat{\mathbf{a}}^T, \mathbf{0}^T)^T$ with large probability, where $\hat{\mathbf{a}} = \text{argmin}\{V_{1n}(\mathbf{a})\}$. Similarly to Huang et al. [5], V_{1n} can be written as

$$\begin{aligned} V_{1n} &= n^{-1/2}(\mathbf{a}^T, \mathbf{0}^T)^T\mathbf{S}_n\{\boldsymbol{\beta}_0(\tau), \tau\} + \frac{1}{2}\mathbf{a}^T\mathbf{H}_{11}\{\boldsymbol{\beta}_0, \tau\}\mathbf{a} + \mathbf{a}^T o_p(1)\mathbf{a} \\ &\quad + \lambda_n \sum_{k=1}^{K_1} c_k \Big\{\Big(\sum_{j\in A_k\cap B_1} \frac{|\beta_{j,0}(\tau) + n^{-1/2}a_j|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^\gamma - \Big(\sum_{j\in A_k\cap B_1} \frac{|\beta_{j,0}(\tau)|}{|\tilde{\beta}_j(\tau)|^\nu}\Big)^\gamma\Big\} \\ &= T_{1n}(\mathbf{a}) + T_{2n}(\mathbf{a}). \end{aligned}$$

By Peng and Fine [4], $n^{-1/2}(\mathbf{1}^T, \mathbf{0}^T)^T\mathbf{S}_n\{\boldsymbol{\beta}_0(\tau), \tau\}$ is asymptotically equal to $n^{-1}(\mathbf{1}^T, \mathbf{0}^T)^T\sum_{i=1}^n \eta_i(\tau)$, where $\mathbf{1}$ is a b_1–dimensional vector and $\mathbf{0}$ is a b_2–dimensional zero vector. Thus, $T_{1n}(\mathbf{a})$ converges in distribution to $\mathbf{a}^T N\{\mathbf{0}, \Sigma_{11}(\tau)\} + \mathbf{a}^T\mathbf{H}_{11}\{\boldsymbol{\beta}_0, \tau\}\mathbf{a}/2$. Similarly to Huang et al. [6], by (C9b) we have $T_{2n}(\mathbf{a}) \to 0$. Thus, V_{1n} converges in distribution to $V_1(\mathbf{a})$. By the argmin continuous mapping theorem of Kim and Pollard [7], $\sqrt{n}(\hat{\boldsymbol{\beta}}_{B_1} - \boldsymbol{\beta}_{B_1,0})$ converges to $\text{argmin}\{V_1(\mathbf{a})\}$. This completes the proof of Theorem 2.5.

References

[1] Ni A, Cai J, Zeng D. Variable selection for case-cohort studies with failure time outcome. *Biometrika* 2016; **103**:547–562.

[2] Lin DY. On fitting cox's proportional hazards models to survey data. *Biometrika* 2000; **87**:37–47.

[3] He P, Eriksson F, Scheike TH, Zhang MJ. A proportional hazards regression model for the subdistribution with covariates-adjusted censoring weight for competing risks data. *Scandinavian Journal of Statistics* 2016; **43**:103122.

[4] Peng L, Fine JP. Competing risks quantile regression. *Journal of the American Statistical Association* 2009; **104**:1440–1453.

[5] Huang J, Li L, Liu Y, Zhao X. Group selection in the Cox model with a diverging number of covariates. *Statistica Sinica* 2014; **24**:1787–1810.

[6] Huang J, Ma S, Xie H, Zhang CH. A group bridge approach for variable selection. *Biometrika* 2009; **96**:339–355.

[7] Kim J, Pollard D. Cube root asymptotics. *Annals of Statistics* 1990; **18**:191–219.