Analysis on Censored Quantile Residual Life Model via Spline Smoothing (Web Appendix)

YANYUAN MA

Texas A&M University, Department of Statistics College Stage, TX, U.S.A.

YING WEI

Columbia University, Department of Biostatistics New York, NY, U.S.A.

Appendix

Define

$$\begin{aligned}
M(\alpha, G) &= \left(\frac{\partial E\left[s_i^{\mathrm{T}}\{t_1, \boldsymbol{\alpha}\mathbf{b}(t_1), G\}\right]}{\partial \alpha}, \ldots, \frac{\partial E\left[s_i^{\mathrm{T}}\{t_J, \boldsymbol{\alpha}\mathbf{b}(t_J), G\}\right]}{\partial \alpha} \right), \\
\mathcal{A}(\alpha, G) &= M(\alpha, G) M^{\mathrm{T}}(\alpha, G), \\
\mu_i(t_j, \alpha, G) &= s_i\{t_j, \boldsymbol{\alpha}\mathbf{b}(t_j), G\} - \mathbf{q}_2(\alpha, t_j) \int_{-\infty}^{t_j} h^{-1}(s) \left\{dI(Y_i \le s, D_i = 0) - I(Y_i \ge s) d\Lambda_G(s)\right\} \\
&\quad + \int_{-\infty}^{\infty} G^{-1}(s) \int_{-\infty}^{s} h^{-1}(v) \left\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v) d\Lambda_G(v)\right\} d\mathbf{q}_{1j}(\alpha, s), \\
\mathbf{q}_{1j}(\alpha, s) &= E\left(\frac{\partial m\left\{X_i, \boldsymbol{\alpha}\mathbf{b}(t_j)\right\}}{\partial \boldsymbol{\alpha}\mathbf{b}(t_j)} I\left[t_j + m\left\{X_i, \boldsymbol{\alpha}\mathbf{b}(t_j)\right\} \le \min(s, Y_i)\right] \right), \\
\mathbf{q}_2(\alpha, t_j) &= (1-\tau) G(t_j)^{-1} E\left[I(Y_i \ge t_j) \frac{\partial m\left\{X_i, \boldsymbol{\alpha}\mathbf{b}(t_j)\right\}}{\partial \boldsymbol{\alpha}\mathbf{b}(t_j)} \right], \\
\mu_i(\alpha, G) &= \left\{\mu_i^{\mathrm{T}}(t_1, \alpha, G), \ldots, \mu_i^{\mathrm{T}}(t_J, \alpha, G)\right\}^{\mathrm{T}},
\end{aligned}$$

where $h(s) = E(Y_1 \ge s)$, Λ_G is the cumulative hazard function of the censoring process.

Proof of Theorem 1: Without loss of generosity, we assume in this proof that $p = 1$. The results hold for any finite p. Let $\mathbf{t} = (t_1, t_2, ..., t_J)$ be a set of times used to construct the estimating equations, then the estimating equations can be written as

$$S_n(\boldsymbol{\alpha}\mathbf{b}(\mathbf{t}), \widehat{G}) = \sum_{i=1}^{n} f_i(\boldsymbol{\alpha}\mathbf{b}(\mathbf{t}), \widehat{G}) = 0.$$

We first consider the situation where G is known. Let

$$u_i(\boldsymbol{\alpha}, \boldsymbol{\alpha}_0) = f_i(\boldsymbol{\alpha}\mathbf{b}(\mathbf{t}), G) - f_i(\boldsymbol{\alpha}_0\mathbf{b}(\mathbf{t}), G) - E f_i(\boldsymbol{\alpha}\mathbf{b}(\mathbf{t}), G).$$

To show the consistency defined in (4), we first establish the following uniform consistency,

$$\sup_{\|\boldsymbol{\alpha}-\boldsymbol{\alpha}_0\|<B(k_n/n)^{1/2}} \|\sum_{i=1}^{n}\eta^{\mathrm{T}}u_i(\boldsymbol{\alpha},\boldsymbol{\alpha}_0)\| = o(n^{1/2}k_n^{1/2}), \tag{A.1}$$

for any $B>0$, and any $\|\eta\|=1$. Based on Lemmas 2.1 and 3.3 of He and Shao (2000), the sufficient conditions for (A.1) are

(C1) $\max_i \sup_{\|\boldsymbol{\alpha}-\boldsymbol{\alpha}_0\|<B(k_n/n)^{1/2}} \|u_i(\boldsymbol{\alpha},\boldsymbol{\alpha}_0)\|^2 = O(k_n^2/n)$

(C2) There exist $0<c\le 2, 0<s\le 2$ such that $\max_{i\le n} E\sup_{\|\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2\|<d}\|f_i(\boldsymbol{\alpha}_1\mathbf{b}(\mathbf{t}),G) - f_i(\boldsymbol{\alpha}_2\mathbf{b}(\mathbf{t}),G) - Ef_i(\boldsymbol{\alpha}_1\mathbf{b}(\mathbf{t}),G) + Ef_i(\boldsymbol{\alpha}_2\mathbf{b}(\mathbf{t}),G)\| \le n^c d^s$, for all $0<d\le 1$.

In what follows, we show that Conditions (C1) and (C2) are satisfied under Assumptions **A1**, **A2** and **A4**. To show condition (C1), we note that under Assumption **A2**, the quantile function $m(x,\boldsymbol{\beta})$ and its first derivative with respect to $\boldsymbol{\beta}$, denoted as $\dot{m}(x,\boldsymbol{\beta})$, satisfy the Lipschitz conditions. That is, there exist constants K_1 and K_2, such that

$$\begin{aligned}
&\max_i \sup_t |m(x_i,\beta_1(t),G) - m(x_i,\beta_2(t),G)| < K_1|\beta_1(t)-\beta_2(t)|.\\
&\max_i \sup_t |\dot{m}(x_i,\beta_1(t),G) - \dot{m}(x_i,\beta_2(t),G)| < K_2|\beta_1(t)-\beta_2(t)|.
\end{aligned}$$

We first bound $\|u_i(\boldsymbol{\alpha},\boldsymbol{\alpha}_0)\|^2$ by

$$\|u_i(\boldsymbol{\alpha},\boldsymbol{\alpha}_0)\|^2 \le \|\left(f_i(\boldsymbol{\alpha}\mathbf{b}(\mathbf{t}),G) - f_i(\boldsymbol{\alpha}_0\mathbf{b}(\mathbf{t}),G)\right)\|^2 + \|Ef_i(\boldsymbol{\alpha}\mathbf{b}(\mathbf{t}),G)\|^2.$$

Let $M_1 = \max_i \sup_{\boldsymbol{\beta}(t)\in\Omega}\left|\frac{\partial m(x_i,\boldsymbol{\beta}(t))}{\partial\boldsymbol{\beta}(t)}\right|$, then, for any two coefficient functions, $\boldsymbol{\beta}_1(t)$ and $\boldsymbol{\beta}_2(t)$, we have

$$\begin{aligned}
&\Big|\frac{\partial m(x_i,\boldsymbol{\beta}_1(t))}{\partial\boldsymbol{\beta}_1(t)}\frac{I[Y_i\ge t+m(x_i,\boldsymbol{\beta}_1(t))]}{G[t+m(x_i,\boldsymbol{\beta}_1(t))]} - \frac{\partial m(x_i,\boldsymbol{\beta}_2(t))}{\partial\boldsymbol{\beta}_2(t)}\frac{I[Y_i\ge t+m(x_i,\boldsymbol{\beta}_2(t))]}{G[t+m(x_i,\boldsymbol{\beta}_2(t))]}\Big|\\
=\;& \Big|\frac{\partial m(x_i,\boldsymbol{\beta}_1(t))}{\partial\boldsymbol{\beta}_1(t)}\left\{\frac{I[Y_i\ge t+m(x_i,\boldsymbol{\beta}_1(t))]}{G[t+m(x_i,\boldsymbol{\beta}_1(t))]} - \frac{I[Y_i\ge t+m(x_i,\boldsymbol{\beta}_2(t))]}{G[t+m(x_i,\boldsymbol{\beta}_2(t))]}\right\}\\
&+\left\{\frac{\partial m(x_i,\boldsymbol{\beta}_1(t))}{\partial\boldsymbol{\beta}_1(t)} - \frac{\partial m(x_i,\boldsymbol{\beta}_2(t))}{\partial\boldsymbol{\beta}_2(t)}\right\}\frac{I[Y_i\ge t+m(x_i,\boldsymbol{\beta}_2(t))]}{G[t+m(x_i,\boldsymbol{\beta}_2(t))]}\Big|\\
\le\;& M_1\left\{\frac{I[|Y_i-t-m(x_i,\boldsymbol{\beta}_2(t))|<|m(x_i,\boldsymbol{\beta}_1(t))-m(x_i,\boldsymbol{\beta}_2(t))|]}{G(T)}\right.\\
&\left.+\left|G^{-1}(t+m(x_i,\boldsymbol{\beta}_1(t))) - G^{-1}(t+m(x_i,\boldsymbol{\beta}_2(t)))\right|\right\}\\
&+K_2\|\boldsymbol{\beta}_1(t)-\boldsymbol{\beta}_2(t)\|G^{-1}(T)\\
\le\;& M_1\left\{\frac{I[|T_i-t-m(x_i,\boldsymbol{\beta}_2(t))|<|m(x_i,\boldsymbol{\beta}_1(t))-m(x_i,\boldsymbol{\beta}_2(t))|]}{G(T)}\right.\\
&+\frac{I[|C_i-t-m(x_i,\boldsymbol{\beta}_2(t))|<|m(x_i,\boldsymbol{\beta}_1(t))-m(x_i,\boldsymbol{\beta}_2(t))|]}{G(T)}\\
&\left.+\left|G^{-1}(t+m(x_i,\boldsymbol{\beta}_1(t))) - G^{-1}(t+m(x_i,\boldsymbol{\beta}_2(t)))\right|\right\}\\
&+K_2\|\boldsymbol{\beta}_1(t)-\boldsymbol{\beta}_2(t)\|G^{-1}(T)\\
=\;& O_p(\|\boldsymbol{\beta}_1(t)-\boldsymbol{\beta}_2(t)\|)
\end{aligned} \tag{A.2}$$

In the above derivation, we used the Lipschitz conditions to obtain the two inequality, and used condition (**A4**) to bound the probability of the indicator function being 1. Similarly, we can show that

$$\left\{\frac{\partial m(x_i,\boldsymbol{\beta}_1(t))}{\partial\boldsymbol{\beta}_1(t)}-\frac{\partial m(x_i,\boldsymbol{\beta}_2(t))}{\partial\boldsymbol{\beta}_2(t)}\right\}\frac{I[Y_i\geq t]}{G[t]} \quad = \quad O_p(\|\boldsymbol{\beta}_1(t)-\boldsymbol{\beta}_2(t)\|) \tag{A.3}$$

Combining (A.2) and (A.3), we have

$$\max_i\|s_i(\boldsymbol{\beta}_1(t),G)-s_i(\boldsymbol{\beta}_2(t),G)\|=O_p(\|\boldsymbol{\beta}_1(t)-\boldsymbol{\beta}_2(t)\|) \tag{A.4}$$

Following the similar arguments, we could show that

$$\max_i\|\frac{\partial Es_i(\boldsymbol{\beta}_1(t),G)}{\partial\boldsymbol{\beta}_1(t)}-\frac{\partial Es_i(\boldsymbol{\beta}_2(t),G)}{\partial\boldsymbol{\beta}_2(t)}\|=O(\|\boldsymbol{\beta}_1(t)-\boldsymbol{\beta}_2(t)\|) \tag{A.5}$$

The above equations (A.4) and (A.5) further imply that

$$\begin{aligned}
&\max_i \sup_{\|\boldsymbol{\alpha}-\boldsymbol{\alpha}_0\|<B(k_n/n)^{1/2}} \|\left(f_i(\boldsymbol{\alpha}\mathbf{b}(\mathbf{t}),G)-f_i(\boldsymbol{\alpha}_0\mathbf{b}(\mathbf{t}),G)\right)\|^2\\
\leq\ & \max_i \sup_{\|\boldsymbol{\alpha}-\boldsymbol{\alpha}_0\|<B(k_n/n)^{1/2}} \|\sum_{j=1}^{J}\frac{\partial Es_i(t_j,\boldsymbol{\alpha}\mathbf{b}(t_j),G)}{\partial\boldsymbol{\alpha}\mathbf{b}(t_j)}\left\{s_i(t_j,\boldsymbol{\alpha}\mathbf{b}(t_j),G)-s_i(t_j,\boldsymbol{\alpha}_0\mathbf{b}(t_j),G)\right\}\mathbf{b}(t_j)\|^2\\
& +\max_i \sup_{\|\boldsymbol{\alpha}-\boldsymbol{\alpha}_0\|<B(k_n/n)^{1/2}} \|\sum_{j=1}^{J}\left\{\frac{\partial Es_i(t_j,\boldsymbol{\alpha}\mathbf{b}(t_j),G)}{\partial\boldsymbol{\alpha}\mathbf{b}(t_j)}-\frac{\partial Es_i(t_j,\boldsymbol{\alpha}_0\mathbf{b}(t_j),G)}{\partial\boldsymbol{\alpha}\mathbf{b}(t_j)}\right\}\\
& \times s_i(t_j,\boldsymbol{\alpha}_0\mathbf{b}(t_j),G)\mathbf{b}(t_j)\|^2\\
=\ & O_p(k_n^2/n).
\end{aligned}$$

The last equality holds due to the fact that $\|\mathbf{b}(t)\|^2=O(1)$ for all t by construction, and $J=O(k_n)$. We denote

$$f_i(\boldsymbol{\beta}_0(\mathbf{t}),G)=\sum_{j=1}^{J}\left(\frac{\partial E[s_i\{t_j,\boldsymbol{\beta}_0(\mathbf{t}),G\}]}{\partial\boldsymbol{\beta}_0(\mathbf{t})}\right)s_i\{t_j,\boldsymbol{\beta}_0(\mathbf{t}),G\}\mathbf{b}(t_j)$$

as the estimating function evaluated at true coefficient $\boldsymbol{\beta}_0(t)$. Then, for any t, $Ef_i(\boldsymbol{\beta}_0(\mathbf{t}),G)=0$. We have

$$\begin{aligned}
&\max_i \sup_{\|\boldsymbol{\alpha}-\boldsymbol{\alpha}_0\|<B(k_n/n)^{1/2}} \|Ef_i(\boldsymbol{\alpha}\mathbf{b}(\mathbf{t}),G)\|^2\\
=\ &\max_i \sup_{\|\boldsymbol{\alpha}-\boldsymbol{\alpha}_0\|<B(k_n/n)^{1/2}} \|Ef_i(\boldsymbol{\alpha}\mathbf{b}(\mathbf{t}),G)-Ef_i(\boldsymbol{\beta}_0(\mathbf{t}),G)\|^2=O(k_n^2/n)+O(k_n^{-2r+1})
\end{aligned}$$

Combining the equations above, condition (C1) is satisfied provided that $k_n^{-2r+1}=O(k_n^2/n)$.

Following the similar arguments for (A.4) and (A.5), with some derivations, we can show that,

$$\begin{aligned}
&\max_i \sup_{\|\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2\|\leq d} \|Es_i(t,\boldsymbol{\alpha}_1\mathbf{b}(t),G)-Es_i(t,\boldsymbol{\alpha}_2\mathbf{b}(t),G)\|\\
&\leq \max_i E \sup_{\|\boldsymbol{\alpha}_1-\boldsymbol{\alpha}_2\|\leq d} \|s_i(t,\boldsymbol{\alpha}_1\mathbf{b}(t),G)-s_i(t,\boldsymbol{\alpha}_2\mathbf{b}(t),G)\|\\
&\leq \mathcal{C}d\|\mathbf{b}(t)\|,
\end{aligned} \tag{A.6}$$

where $\mathcal{C} = M_1K_1\{\sup_{t\in[0,T]} \partial G^{-1}(t)/\partial t + G^{-1}(T)\sup_y f_{Y_i}(y)\} + (2-\tau)G^{-1}(T)K_2$. Combining (A.6) with the facts that $J = O(k_n)$, and $s_i(\cdot)$ and its first derivative are bounded away from infinity, Condition (C2) is satisfied. The uniform expansion (A.1) is hence proved.

Let

$$D(t) = \frac{\partial \left\{ \frac{Es_i(t,\boldsymbol{\beta}_0(t),G)}{\partial \boldsymbol{\beta}_0(t)} Es_i(t,\boldsymbol{\beta}_0(t),G) \right\}}{\partial \boldsymbol{\beta}_0(t)},$$

we further note that

$$D_0 = \inf_t D(t) > 0.$$

This is because by Assumption $\mathbf{A_3}$, $\boldsymbol{\beta}_0(t)$ is the unique solution to $E(S_n(\boldsymbol{\beta}(t))) = 0$ for all t, hence the objective function is convex. Note that

$$\begin{aligned}
& Ef_i(\boldsymbol{\alpha}\mathbf{b}(t), G) \\
= & \sum_{j=1}^{J} \frac{Es_i(t_j, \boldsymbol{\alpha}\mathbf{b}(t_j), G)}{\partial \boldsymbol{\alpha}\mathbf{b}(t_j)} Es_i(t, \boldsymbol{\alpha}\mathbf{b}(t), G)\mathbf{b}(t_j) \\
= & \sum_{j=1}^{J} \left\{ \frac{\partial \left\{ \frac{Es_i(t_j,\boldsymbol{\beta}_0(t_j),G)}{\partial \boldsymbol{\beta}_0(t_j)} Es_i(t_j,\boldsymbol{\beta}_0(t_j),G) \right\}}{\partial \boldsymbol{\beta}_0(t_j)} (\boldsymbol{\alpha}\mathbf{b}(t_j) - \boldsymbol{\beta}_0(t)) + O(\|\boldsymbol{\alpha}\mathbf{b}(t_j) - \boldsymbol{\beta}_0(t)\|^2) \right\} \mathbf{b}(t_j).
\end{aligned}$$

Let $\alpha = \alpha_0 + B(k_n/n)^{1/2}\eta^{\mathrm{T}}$, $||\eta|| = 1$, and $k_n^{-2r+1} = o(k_n^2/n)$, we obtain

$$\begin{aligned}
& Ef_i((\alpha_0 + B(k_n/n)^{1/2}\eta^{\mathrm{T}})\mathbf{b}(t), G) \\
= & \sum_{j=1}^{J} \left\{ \frac{\partial \left\{ \frac{Es_i(t_j,\boldsymbol{\beta}_0(t_j),G)}{\partial \boldsymbol{\beta}_0(t_j)} Es_i(t_j,\boldsymbol{\beta}_0(t_j),G) \right\}}{\partial \boldsymbol{\beta}_0(t_j)} B(k_n/n)^{1/2}\eta^{\mathrm{T}}\mathbf{b}(t_j) + o((k_n/n)^{1/2}) \right\} \mathbf{b}(t_j),
\end{aligned}$$

hence

$$\eta^{\mathrm{T}} Ef_i((\alpha_0 + B(k_n/n)^{1/2}\eta^{\mathrm{T}})\mathbf{b}(t), G) \geq \sum_{j=1}^{J} D_0 B(k_n/n)^{1/2}(\eta^{\mathrm{T}}\mathbf{b}(t_j))^2 + k_n o((k_n/n)^{1/2}). \quad \text{(A.7)}$$

The uniform expansion (A.1), together with (A.7), imply that

$$\begin{aligned}
& \sum_{i=1}^{n} \eta^{\mathrm{T}} f_i(\{\boldsymbol{\alpha}_0 + B(k_n/n)^{1/2}\eta^{\mathrm{T}}\}\mathbf{b}(\mathbf{t}), G) \\
= & \sum_{i=1}^{n} \eta^{\mathrm{T}} f_i(\boldsymbol{\alpha}_0\mathbf{b}(\mathbf{t}), G) + \sum_{i=1}^{n} \eta^{\mathrm{T}} Ef_i(\{\boldsymbol{\alpha}_0 + B(k_n/n)^{1/2}\eta^{\mathrm{T}})\}\mathbf{b}(\mathbf{t}), G) + o(n^{1/2}k_n^{1/2}) \\
\geq & \sum_{i=1}^{n} \eta^{\mathrm{T}} f_i(\boldsymbol{\alpha}_0\mathbf{b}(\mathbf{t}), G) + n\sum_{j=1}^{J} D_0 B(k_n/n)^{1/2}(\eta^{\mathrm{T}}\mathbf{b}(t_j))^2 + nk_n o((k_n/n)^{1/2}) + o(n^{1/2}k_n^{1/2}) \\
= & \sum_{i=1}^{n} \eta^{\mathrm{T}} f_i(\boldsymbol{\alpha}_0\mathbf{b}(\mathbf{t}), G) + D_0 B n^{1/2} k_n^{1/2} \sum_{j=1}^{J} (\eta^{\mathrm{T}}\mathbf{b}(t_j))^2 + o(n^{1/2}k_n^{3/2}). \quad \text{(A.8)}
\end{aligned}$$

We now show that the dominant term in (A.8) is the second term. We only need to compare the first two terms. Let $M_2 = \max_i \sup_t \partial ES_i(t, \boldsymbol{\beta}(t), G)/\partial \boldsymbol{\beta}(t)$. Under Assumption **A5**, we have

$$\begin{aligned}
\sum_{i=1}^{n} E\|f_i(\boldsymbol{\alpha}_0 \mathbf{b}(\mathbf{t}), G)\|^2 &\leq \sum_{i=1}^{n} E \sum_{j=1}^{J} \left\| \frac{\partial E s_i(t_j, \boldsymbol{\alpha}_0 \mathbf{b}(t_j), G)}{\partial \boldsymbol{\alpha}_0 \mathbf{b}(t_j)} s_i(t_j, \boldsymbol{\alpha}_0 \mathbf{b}(t_j), G) \mathbf{b}(t_j) \right\|^2 \\
&= \sum_{i=1}^{n} \sum_{j=1}^{J} \left\{ \frac{\partial E s_i(t_j, \boldsymbol{\alpha}_0 \mathbf{b}(t_j), G)}{\partial \boldsymbol{\alpha}_0 \mathbf{b}(t_j)} \right\}^2 E s_i^2(t_j, \boldsymbol{\alpha}_0 \mathbf{b}(t_j), G) \|\mathbf{b}(t_j)\|^2 \\
&= O(n k_n^3).
\end{aligned}$$

The inequality above implies that $\|\sum_{i=1}^{n} f_i(\boldsymbol{\alpha}_0 \mathbf{b}(\mathbf{t}), G)\| = O_p(n^{1/2} k_n^{3/2})$, which is much smaller than the second term in (A.8). The second term in (A.8) is positive, therefore, the probability for the left side of (A.8) larger than 0 tends to 1, i.e.

$$Prob(\inf_{\|\eta\|=1} \sum_{i=1}^{n} \eta^{\mathrm{T}} f_i(\{\boldsymbol{\alpha}_0 + B(k_n/n)^{1/2}\eta\} \mathbf{b}(\mathbf{t}), G) > 0) \to 1.$$

Following Gsorgo and Horvath (1983), for all $\epsilon > 0$, the Kaplan-Meier estimates is uniformly consistent with $\sup_t |\widehat{G}(t) - G(t)| = o(n^{-1/2+\epsilon}), a.s.$ Under assumption **A6**, using G^* to denote a quantity between G and $\widehat{G}$, we have

$$\begin{aligned}
& \sum_{i=1}^{n} \eta^{\mathrm{T}} f_i(\{\boldsymbol{\alpha}_0 + B(k_n/n)^{1/2}\eta\} \mathbf{b}(\mathbf{t}), \widehat{G}) \\
= & \sum_{i=1}^{n} \eta^{\mathrm{T}} f_i(\{\boldsymbol{\alpha}_0 + B(k_n/n)^{1/2}\eta\} \mathbf{b}(\mathbf{t}), G) + \sum_{i=1}^{n} \eta^{\mathrm{T}} \frac{\partial f_i(\{\boldsymbol{\alpha}_0 + B(k_n/n)^{1/2}\eta\} \mathbf{b}(\mathbf{t}), G^*)}{\partial G} (\widehat{G} - G) \\
= & \sum_{i=1}^{n} \eta^{\mathrm{T}} f_i(\{\boldsymbol{\alpha}_0 + B(k_n/n)^{1/2}\eta\} \mathbf{b}(\mathbf{t}), G) + o_p(n k_n n^{-1/2+\epsilon}) \\
= & \sum_{i=1}^{n} \eta^{\mathrm{T}} f_i(\{\boldsymbol{\alpha}_0 + B(k_n/n)^{1/2}\eta\} \mathbf{b}(\mathbf{t}), G) + o_p(n^{1/2} k_n n^{\epsilon}).
\end{aligned}$$

For sufficiently small ϵ, the dominant term of the above expression is the first term. Hence we have

$$Prob(\inf_{\|\eta\|=1} \sum_{i=1}^{n} \eta^{\mathrm{T}} f_i(\{\boldsymbol{\alpha}_0 + B(k_n/n)^{1/2}\eta\} \mathbf{b}(\mathbf{t}), \widehat{G}) > 0) \to 1. \tag{A.9}$$

Since $\widehat{\boldsymbol{\alpha}}$ is the minimizer of (3), (A.9) further implies that there exists a local minimizer $\widehat{\boldsymbol{\alpha}}$, such that

$$\|\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}_0\|^2 = O(k_n/n).$$

□

Proof of Theorem 2: The uniform consistency established in Theorem 1 and the uniform consistency of $\widehat{G}$ allow us to expand the estimating equation at α_0, G.

$$
\begin{aligned}
& 0 \\
= \; & n^{-1/2} S_n\{\widehat{\boldsymbol\alpha}\mathbf{b}(\mathbf{t}), \widehat{G}(\mathbf{t})\} \\
= \; & n^{-1/2} \frac{\partial E S_n\{\boldsymbol\alpha_0\mathbf{b}(\mathbf{t}), \widehat{G}(\mathbf{t})\}}{\partial \alpha_0}(\widehat{\alpha}-\alpha_0) + o_p(k_n) \\
& + n^{-1/2}\sum_{i=1}^n\sum_{j=1}^J \left(\frac{\partial E\left[s_i^{\mathrm{T}}\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\}\right]}{\partial \alpha} + o_p(n^{-1/2+\epsilon}) \right) s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\} \\
& + n^{-1/2}\sum_{i=1}^n\sum_{j=1}^J \left(\frac{\partial E\left[s_i^{\mathrm{T}}\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), \widehat{G}\}\right]}{\partial \alpha} + o_p(n^{-1/2+\epsilon}) \right) \\
& \times \left[s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), \widehat{G}\} - s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\} \right] \\
& + n^{-1/2}\sum_{i=1}^n\sum_{j=1}^J \left(\frac{\partial E\left[s_i^{\mathrm{T}}\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), \widehat{G}\}\right]}{\partial \alpha} - \frac{\partial E\left[s_i^{\mathrm{T}}\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\}\right]}{\partial \alpha} + o_p(n^{-1/2+\epsilon}) \right) \\
& \times s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\} \\
= \; & n^{-1/2} \frac{\partial E S_n\{\boldsymbol\alpha_0\mathbf{b}(\mathbf{t}), \widehat{G}(\mathbf{t})\}}{\partial \alpha}(\widehat{\alpha}-\alpha_0) + n^{-1/2}\sum_{i=1}^n\sum_{j=1}^J \frac{\partial E\left[s_i^{\mathrm{T}}\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\}\right]}{\partial \alpha} s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\} \\
& + n^{-1/2}\sum_{i=1}^n\sum_{j=1}^J \frac{\partial E\left[s_i^{\mathrm{T}}\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), \widehat{G}\}\right]}{\partial \alpha} \left[s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), \widehat{G}\} - s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\} \right] + o_p(k_n) \\
= \; & n^{-1/2} \frac{\partial E S_n\{\boldsymbol\alpha_0\mathbf{b}(\mathbf{t}), G(\mathbf{t})\}}{\partial \alpha}(\widehat{\alpha}-\alpha_0) + n^{-1/2}\sum_{i=1}^n\sum_{j=1}^J \frac{\partial E\left[s_i^{\mathrm{T}}\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\}\right]}{\partial \alpha} s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\} \\
& + n^{-1/2}\sum_{i=1}^n\sum_{j=1}^J \frac{\partial E\left[s_i^{\mathrm{T}}\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\}\right]}{\partial \alpha} \left[s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), \widehat{G}\} - s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\} \right] + o_p(k_n)
\end{aligned}
\tag{A.10}
$$

as long as $J = O(k_n) = o(n^{1/2-\epsilon})$. We can expand the following term

$$
\begin{aligned}
& n^{-1/2}\sum_{i=1}^n \left[s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), \widehat{G}\} - s_i\{t_j, \boldsymbol\alpha_0\mathbf{b}(t_j), G\} \right] \\
= \; & n^{-1}\sum_{i=1}^n \frac{\partial m\{X_i, \boldsymbol\alpha_0\mathbf{b}(t_j)\}}{\partial\{\boldsymbol\alpha_0\mathbf{b}(t_j)\}} I\left[Y_i \ge t_j + m\{X_i, \boldsymbol\alpha_0\mathbf{b}(t_j)\}\right] \\
& \times n^{1/2}\left(\frac{1}{\widehat{G}\left[t_j + m\{X_i, \boldsymbol\alpha_0\mathbf{b}(t_j)\}\right]} - \frac{1}{G\left[t_j + m\{X_i, \boldsymbol\alpha_0\mathbf{b}(t_j)\}\right]} \right) \\
& -(1-\tau)n^{-1}\sum_{i=1}^n \frac{\partial m\{X_i, \boldsymbol\alpha_0\mathbf{b}(t_j)\}}{\partial\{\boldsymbol\alpha_0\mathbf{b}(t_j)\}} I(Y_i \ge t_j) n^{1/2}\left\{ \frac{1}{\widehat{G}(t_j)} - \frac{1}{G(t_j)} \right\}.
\end{aligned}
\tag{A.11}
$$

We first check the first term in (A.11). Following Fleming and Harrington (1991, Corollary 3.2.1), We can use a martingale integral representation to obtain

$$
\begin{aligned}
& -n^{1/2}\{\widehat{G}(s)-G(s)\}/G(s) \\
= & \int_{-\infty}^{s} \frac{n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\Lambda_G(v)\}}{n^{-1}\sum_{i=1}^{n} I(Y_i \ge v)} \\
= & \int_{-\infty}^{s} h^{-1}(v) n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\Lambda_G(v)\} + o_p(1),
\end{aligned}
$$

where $\Lambda_G(\cdot)$ is the cumulative hazard function for the censoring process. Thus, we can write the first term in (A.11) as

$$
\int_{-\infty}^{\infty} G(s)^{-1}\int_{-\infty}^{s} h^{-1}(v)n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\Lambda_G(v)\}d\mathbf{q}_{1j}(\alpha_0, s) + o_p(1).
$$

Let

$$
\mathbf{q}_{2j}(\alpha, s) = E\left[\frac{\partial m\left\{X_i, \boldsymbol{\alpha}\mathbf{b}(t_j)\right\}}{\partial \boldsymbol{\alpha}\mathbf{b}(t_j)} I\left\{t_j \le \min(s, Y_i)\right\}\right] = E\left[\frac{\partial m\left\{X_i, \boldsymbol{\alpha}\mathbf{b}(t_j)\right\}}{\partial \boldsymbol{\alpha}\mathbf{b}(t_j)} I(t_j \le Y_i)\right] I(t_j \le s),
$$

similar derivation can show that the second term in (A.11) can be written as

$$
\begin{aligned}
& -(1-\tau)\int_{-\infty}^{\infty} G(s)^{-1}\int_{-\infty}^{s} h^{-1}(v)n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\Lambda_G(v)\}d\mathbf{q}_{2j}(\alpha_0, s) \\
& +o_p(1) \\
= & -(1-\tau)E\left[\frac{\partial m\left\{X_i, \boldsymbol{\alpha}_0\mathbf{b}(t_j)\right\}}{\partial \alpha} I(t_j \le Y_i)\right] G(t_j)^{-1} \\
& \times \int_{-\infty}^{t_j} h^{-1}(v)n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\Lambda_G(v)\} + o_p(1) \\
= & -\mathbf{q}_2(\alpha_0, t_j)\int_{-\infty}^{t_j} h^{-1}(v)n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\Lambda_G(v)\} + o_p(1).
\end{aligned}
$$

Thus, continue from (A.10), we have

$$
\begin{aligned}
0 = & \; n^{-1/2}\frac{\partial ES_n\{\boldsymbol{\alpha}_0\mathbf{b}(\mathbf{t}), G(\mathbf{t})\}}{\partial \alpha}(\widehat{\alpha}-\alpha_0) + \sum_{j=1}^{J}\frac{\partial E\left[s_i^{\mathrm{T}}\{t_j, \boldsymbol{\alpha}_0\mathbf{b}(t_j), G\}\right]}{\partial \alpha} n^{-1/2}\sum_{i=1}^{n} \\
& \Bigg[s_i\{t_j, \boldsymbol{\alpha}_0\mathbf{b}(t_j), G\} + \int_{-\infty}^{\infty} G(s)^{-1}\int_{-\infty}^{s} h^{-1}(v)\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\Lambda_G(v)\}d\mathbf{q}_{1j}(\alpha_0, s) \\
& -\mathbf{q}_2(\alpha_0, t_j)\int_{-\infty}^{t_j} h^{-1}(v)\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\Lambda_G(v)\}\Bigg] + r_n \\
= & \; n^{-1/2}\frac{\partial ES_n\{\boldsymbol{\alpha}_0\mathbf{b}(\mathbf{t}), G(\mathbf{t})\}}{\partial \alpha}(\widehat{\alpha}-\alpha_0) + n^{-1/2}\sum_{i=1}^{n}\sum_{j=1}^{J}\frac{\partial E\left[s_i^{\mathrm{T}}\{t_j, \boldsymbol{\alpha}_0\mathbf{b}(t_j), G\}\right]}{\partial \alpha}\mu_i(t_j, \alpha_0, G) + r_n \\
= & \; \mathcal{A}(\alpha_0, G)n^{1/2}(\widehat{\alpha}-\alpha_0) + M(\alpha_0, G)n^{-1/2}\sum_{i=1}^{n}\mu_i(\alpha_0, G) + r_n,
\end{aligned}
$$

where $\|r_n\| = o_p(k_n)$. Thus, we have

$$n^{1/2}\eta^{\mathrm{T}}(\widehat{\alpha}-\alpha_0) = -\eta^{\mathrm{T}}\mathcal{A}(\alpha_0, G)^{-1}M(\alpha_0, G)n^{-1/2}\sum_{i=1}^{n}\mu_i(\alpha_0, G) + o_p(1)$$

for any $\eta \in R^{k_n}, ||\eta|| = 1$, where $o_p(1)$ is a scalar that goes to zero in probability when $n \to \infty$. The results thus follow. □

<u>Proof of Theorem 3:</u> In the scope of this proof, we define

$$\begin{aligned}
\mathbf{q}_{1j}(\boldsymbol{\beta}_j, s) &= E\left[\frac{\partial m\left(X_i, \boldsymbol{\beta}_j\right)}{\partial \boldsymbol{\beta}_j^{\mathrm{T}}} I\left\{t_j + m\left(X_i, \boldsymbol{\beta}_j\right) \le \min(s, Y_i)\right\}\right], \\
\mathbf{q}_2(\boldsymbol{\beta}_j, t_j) &= (1-\tau)G(t_j)^{-1}E\left\{I(Y_i \ge t_j)\frac{\partial m\left(X_i, \boldsymbol{\beta}_j\right)}{\partial \boldsymbol{\beta}_j^{\mathrm{T}}}\right\}.
\end{aligned}$$

The uniform consistency established in Theorem 1 certainly also applies to a single $\check{\boldsymbol{\beta}}_j$ by treating $\boldsymbol{\beta}_j$ as a special case of $\boldsymbol{\alpha}$ where we take $k_n = 1$ basis function $b(t) \equiv 1$. The uniform consistency of $\widehat{G}$ further allows us to expand the estimating equation at $\boldsymbol{\beta}_j, G$.

$$\begin{aligned}
0 &= n^{-1/2}\sum_{i=1}^{n} s_i(t_j, \check{\boldsymbol{\beta}}_j, \widehat{G}) \\
&= n^{1/2}\frac{\partial E s_i(t_j, \boldsymbol{\beta}_j, \widehat{G})}{\partial \boldsymbol{\beta}_j^{\mathrm{T}}}(\check{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j) + n^{-1/2}\sum_{i=1}^{n} s_i(t_j, \boldsymbol{\beta}_j, G) \\
&\quad + n^{-1/2}\sum_{i=1}^{n}\left\{s_i(t_j, \boldsymbol{\beta}_j, \widehat{G}) - s_i(t_j, \boldsymbol{\beta}_j, G)\right\} + o_p(1) \\
&= n^{1/2}\frac{\partial E s_i(t_j, \boldsymbol{\beta}_j, G)}{\partial \boldsymbol{\beta}_j^{\mathrm{T}}}(\check{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j) + n^{-1/2}\sum_{i=1}^{n} s_i(t_j, \boldsymbol{\beta}_j, G) \\
&\quad + n^{-1/2}\sum_{i=1}^{n}\left\{s_i(t_j, \boldsymbol{\beta}_j, \widehat{G}) - s_i(t_j, \boldsymbol{\beta}_j, G)\right\} + o_p(1).
\end{aligned} \tag{A.12}$$

We can expand the following term

$$\begin{aligned}
& n^{-1/2}\sum_{i=1}^{n}\left\{s_i(t_j, \boldsymbol{\beta}_j, \widehat{G}) - s_i(t_j, \boldsymbol{\beta}_j, G)\right\} \\
= \; & n^{-1}\sum_{i=1}^{n}\frac{\partial m(X_i, \boldsymbol{\beta}_j)}{\partial \boldsymbol{\beta}_j^{\mathrm{T}}} I\left\{Y_i \ge t_j + m(X_i, \boldsymbol{\beta}_j)\right\} n^{1/2}\left[\frac{1}{\widehat{G}\left\{t_j + m(X_i, \boldsymbol{\beta}_j)\right\}} - \frac{1}{G\left\{t_j + m(X_i, \boldsymbol{\beta}_j)\right\}}\right] \\
& -(1-\tau)n^{-1}\sum_{i=1}^{n}\frac{\partial m(X_i, \boldsymbol{\beta}_j)}{\partial \boldsymbol{\beta}_j^{\mathrm{T}}} I(Y_i \ge t_j) n^{1/2}\left\{\frac{1}{\widehat{G}(t_j)} - \frac{1}{G(t_j)}\right\}.
\end{aligned} \tag{A.13}$$

We first check the first term in (A.13). Following Fleming and Harrington (1991, Corollary 3.2.1), We can use a martingale integral representation to obtain

$$
\begin{aligned}
& -n^{1/2}\{\widehat{G}(s)-G(s)\}/G(s) \\
= & \int_{-\infty}^{s} \frac{n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i=0)-I(Y_i \ge v)d\Lambda_G(v)\}}{n^{-1}\sum_{i=1}^{n} I(Y_i \ge v)} \\
= & \int_{-\infty}^{s} h^{-1}(v) n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i=0)-I(Y_i \ge v)d\Lambda_G(v)\}+o_p(1),
\end{aligned}
$$

where $\Lambda_G(\cdot)$ is the cumulative hazard function for the censoring process. Thus, we can write the first term in (A.13) as

$$
\int_{-\infty}^{\infty} G(s)^{-1}\int_{-\infty}^{s} h^{-1}(v)n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i=0)-I(Y_i\ge v)d\Lambda_G(v)\}d\mathbf{q}_{1j}(\boldsymbol{\beta}_j, s)+o_p(1).
$$

Let

$$
\mathbf{q}_{2j}(\boldsymbol{\beta}_j, s) = E\left[\frac{\partial m\left(X_i, \boldsymbol{\beta}_j\right)}{\partial \boldsymbol{\beta}_j^{\mathrm{T}}} I\left\{t_j \le \min(s, Y_i)\right\}\right] = E\left[\frac{\partial m\left\{X_i, \boldsymbol{\beta}_j\right\}}{\partial \boldsymbol{\beta}_j^{\mathrm{T}}} I(t_j \le Y_i)\right] I(t_j \le s),
$$

similar derivation can show that the second term in (A.13) can be written as

$$
\begin{aligned}
& -(1-\tau)\int_{-\infty}^{\infty} G(s)^{-1}\int_{-\infty}^{s} h^{-1}(v)n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i=0)-I(Y_i \ge v)d\Lambda_G(v)\}d\mathbf{q}_{2j}(\boldsymbol{\beta}_j, s) \\
& +o_p(1) \\
= & -(1-\tau)E\left[\frac{\partial m(X_i, \boldsymbol{\beta}_j)}{\partial \boldsymbol{\beta}_j^{\mathrm{T}}} I(t_j \le Y_i)\right] G(t_j)^{-1} \\
& \times \int_{-\infty}^{t_j} h^{-1}(v)n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i=0)-I(Y_i\ge v)d\Lambda_G(v)\}+o_p(1) \\
= & -\mathbf{q}_2(\boldsymbol{\beta}_j, t_j)\int_{-\infty}^{t_j} h^{-1}(v)n^{-1/2}\sum_{i=1}^{n}\{dI(Y_i \le v, D_i=0)-I(Y_i \ge v)d\Lambda_G(v)\}+o_p(1).
\end{aligned}
$$

Thus, continue from (A.12), we have

$$
\begin{aligned}
0 = & n^{1/2}\frac{\partial E s_i(t_j, \boldsymbol{\beta}_j, G)}{\partial \boldsymbol{\beta}_j^{\mathrm{T}}}(\check{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j) + n^{-1/2}\sum_{i=1}^{n} \\
& \Bigg[s_i(t_j, \boldsymbol{\beta}_j, G) + \int_{-\infty}^{\infty} G(s)^{-1}\int_{-\infty}^{s} h^{-1}(v)\{dI(Y_i \le v, D_i=0)-I(Y_i \ge v)d\Lambda_G(v)\}d\mathbf{q}_{1j}(\boldsymbol{\beta}_j, s) \\
& -\mathbf{q}_2(\boldsymbol{\beta}_j, t_j)\int_{-\infty}^{t_j} h^{-1}(v)\{dI(Y_i \le v, D_i=0)-I(Y_i\ge v)d\Lambda_G(v)\}\Bigg]+o_p(1) \\
= & \mathcal{M}_j n^{1/2}(\check{\boldsymbol{\beta}}_j - \boldsymbol{\beta}_j) + n^{-1/2}\sum_{i=1}^{n}\nu_i(t_j, \boldsymbol{\beta}_j, G)+o_p(1),
\end{aligned}
$$

where $\mathcal{M}_j$ is the jth block diagonal of $\mathcal{M}$. Ensembling the above equation for $j = 1, \ldots, J$, multiple $\eta^{\mathrm{T}}\mathcal{C}\mathcal{M}$ from the left yields

$$
\begin{aligned}
0 &= \eta^{\mathrm{T}}\mathcal{C}n^{1/2}\{(\check{\boldsymbol{\beta}}(t_1)^{\mathrm{T}}, \ldots, \check{\boldsymbol{\beta}}(t_J)^{\mathrm{T}}) - (\boldsymbol{\beta}(t_1)^{\mathrm{T}}, \ldots, \boldsymbol{\beta}(t_J)^{\mathrm{T}})\}^{\mathrm{T}} \\
&\quad + n^{-1/2}\eta^{\mathrm{T}}\mathcal{C}\mathcal{M}^{-1}\sum_{i=1}^{n}(\nu_i(t_1, \boldsymbol{\beta}_1, G)^{\mathrm{T}}, \ldots, \nu_i(t_1, \boldsymbol{\beta}_1, G)^{\mathrm{T}})^{\mathrm{T}} \\
&= \eta^{\mathrm{T}}n^{1/2}(\widetilde{\alpha} - \alpha) + n^{-1/2}\eta^{\mathrm{T}}\mathcal{C}\mathcal{M}^{-1}\sum_{i=1}^{n}(\nu_i(t_1, \boldsymbol{\beta}_1, G)^{\mathrm{T}}, \ldots, \nu_i(t_1, \boldsymbol{\beta}_1, G)^{\mathrm{T}})^{\mathrm{T}},
\end{aligned}
$$

which directly yields the result. □

Estimation of cov$\{\mu(\alpha_0, G)\}$

For any value t, we make the approximation

$$
\begin{aligned}
d\widehat{\Lambda}_G(t) &= \left\{\sum_{i=1}^{n} I(Y_i \ge t)\right\}^{-1} d\left\{\sum_{i=1}^{n} I(Y_i \le t, D_i = 0)\right\}, \\
\widehat{\mathbf{q}}_{1j}(\alpha, t) &= n^{-1}\sum_{i=1}^{n}\frac{\partial m\left\{X_i, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j)\right\}}{\partial\alpha} I\left[t_j + m\left\{X_i, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j)\right\} \le \min(t, Y_i)\right].
\end{aligned}
$$

We then obtain

$$
\begin{aligned}
&\int_{-\infty}^{t}\widehat{h}^{-1}(v)\left\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\widehat{\Lambda}_G(v)\right\} \\
= \; & h^{-1}(Y_i)(1 - D_i)I(Y_i \le t) - n^{-1}\sum_{k=1}^{n}\widehat{h}^{-2}(Y_k)(1 - D_k)I\{Y_k \le \min(t, Y_i)\}, \\
& d\widehat{\mathbf{q}}_{1j}(\alpha, t) \\
= \; & n^{-1}\sum_{i=1}^{n}\frac{\partial m\left\{X_i, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j)\right\}}{\partial\alpha} I\left[t_j + m\left\{X_i, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j)\right\} \le Y_i\right] dI\left[t_j + m\left\{X_i, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j)\right\} \le t\right].
\end{aligned}
$$

Combine the above two results, we have

$$
\begin{aligned}
&\int_{-\infty}^{\infty}\widehat{G}^{-1}(s)\int_{-\infty}^{s}\widehat{h}^{-1}(v)\left\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\widehat{\Lambda}_G(v)\right\} d\widehat{\mathbf{q}}_{1j}(\alpha, s) \\
= \; & \widehat{G}^{-1}[t_j + m\{X_i, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j)\}]\Big(h^{-1}(Y_i)(1 - D_i)I[Y_i \le t_j + m\{X_i, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j)\}] \\
& - n^{-1}\sum_{k=1}^{n}\widehat{h}^{-2}(Y_k)(1 - D_k)I\left(Y_k \le \min\left[t_j + m\{X_i, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j)\}, Y_i\right]\right)\Big) \\
& \left(n^{-1}\sum_{i=1}^{n}\frac{\partial m\left\{X_i, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j)\right\}}{\partial\alpha} I\left[t_j + m\left\{X_i, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j)\right\} \le Y_i\right]\right).
\end{aligned}
$$

Since

$$
\begin{aligned}
\widehat{\mu}_i(t_j, \widehat{\alpha}, \widehat{G}) &= s_i\{t_j, \widehat{\boldsymbol{\alpha}}\mathbf{b}(t_j), \widehat{G}\} - \widehat{\mathbf{q}}_2(\alpha, t_j)\int_{-\infty}^{t_j}\widehat{h}^{-1}(s)\left\{dI(Y_i \le s, D_i = 0) - I(Y_i \ge s)d\widehat{\Lambda}_G(s)\right\} \\
&\quad + \int_{-\infty}^{\infty}\widehat{G}^{-1}(s)\int_{-\infty}^{s}\widehat{h}^{-1}(v)\left\{dI(Y_i \le v, D_i = 0) - I(Y_i \ge v)d\widehat{\Lambda}_G(v)\right\} d\widehat{\mathbf{q}}_{1j}(\alpha, s),
\end{aligned}
$$

inserting the above results, we obtain (9). □