BIOMETRICS 000, 000–000 DOI: 000
000 0000

Web-based Supplementary Materials for "Estimating subject-specific dependent competing risk profile with censored event time observations"

by Li, Tian and Wei

Appendix A: The Asymptotical Properties of $\hat{\eta}(v)$

Recalling that $\hat{\beta}$ is a solution from the estimating equation

$$n^{-1}\sum_{i=1}^{n}\frac{w_i}{\hat{G}(T_i \wedge t_0)}X_i\{I(T_i \leqslant t_0, \epsilon = 1) - g(\beta' X_i)\} = 0,$$

it follows from the similar arguments used in Tian et al.(2007) that $\hat{\beta}$ converges to a deterministic limit β_0 and

$$\hat{\beta} - \beta_0 = n^{-1}\xi_i + o_p(n^{-1})$$

where β_0 is the solution of $r(\beta) = 0$,

$$\xi_i = [E\{\dot{g}(\beta_0' X_i)X_i^{\otimes 2}\}]^{-1}\Bigg(X_i\{I(\tilde{T}_i \leqslant t_0, \epsilon = 1) - g(\beta' X_i)\}$$
$$- \int_0^{t_0}\mathbb{K}(X_i\{I(\tilde{T}_i \leqslant t_0, \epsilon = 1) - g(\beta' X_i)\}, u)\frac{dM_i^C(u)}{G(T_i \wedge t_0)}\Bigg),$$

$\mathbb{K}(W, u) = W - E\{WI(\tilde{T} \wedge t_0 \geqslant u)\}/\mathrm{pr}(\tilde{T} \wedge t_0 \geqslant u)$ for any random vector W and $M_i^C(u)$ is the martingale process associated with the censoring time C_i. Let $V_i = \beta_0' X_i$ and $\hat{V}_i = \hat{\beta}' X_i$. With slightly abuse of notation, we let $\{\hat{a}(v)', \hat{b}(v)'\}'$ be the maximizer of

$$\sum_{i=1}^{n}\frac{w_i K_h(\hat{V}_i - v)}{\hat{G}(T_i \wedge t_0)}\left(\sum_{k=1}^{K-1} Y_{ik}\{a_k + b_k(\hat{V}_i - v)\} - \log\left[1 + \sum_{k=1}^{K-1}\exp\{a_k + b_k(\hat{V}_i - v)\}\right]\right),$$

and then it is the solution to the estimating equation

$$\hat{S}(a, b; v) = \{\hat{S}_1'(a, b; v), \cdots, \hat{S}_{K-1}'(a, b; v)\}' = 0$$

where

$$\hat{S}_k(a, b; v) = n^{-1}\sum_{i=1}^{n}\frac{w_i K_h(\hat{V}_i - v)}{\hat{G}(T_i \wedge t_0)}\begin{pmatrix} 1 \\ \hat{V}_i - v \end{pmatrix}\left\{Y_{ik} - \frac{\exp\{a_k + b_k(\hat{V}_i - v)\}}{1 + \sum_{k=1}^{K-1}\exp\{a_k + b_k(\hat{V}_i - v)\}}\right\}.$$

To study the asymptotical properties of $\hat{a}(v)$, we let $\hat{\Delta}_a(v) = \{\hat{a}_1(v) - m_1(v), \cdots, \hat{a}_{K-1}(v) - m_{K-1}(v)\}'$ and $\hat{\Delta}_b(v) = h\{\hat{b}_1(v) - \dot{m}_1(v), \cdots, \hat{b}_{K-1}(v) - \dot{m}_{K-1}(v)\}$, where $m_j(v) = \log\{\eta_j(v)/\eta_K(v)\}$ and $\dot{m}_j(v) = dm_j(v)/dv, j = 1, \cdots, K-1$. Therefore, $\{\hat{\Delta}_a(v)', \hat{\Delta}_b(v)'\}'$ is the solution to the estimating equation

$$\hat{Q}(\Delta_a, \Delta_b; v) = \{\hat{Q}_1'(\Delta_a, \Delta_b; v), \cdots, \hat{Q}_{K-1}'(\Delta_a, \Delta_b; v)\}' = 0$$

where $\hat{Q}_k(\Delta_a, \Delta_b; v)$ is

$$n^{-1}\sum_{i=1}^{n}\frac{w_iK_h(\hat{V}_i-v)}{\hat{G}(T_i\wedge t_0)}\begin{pmatrix}1\\(\hat{V}_i-v)/h\end{pmatrix}\left\{Y_{ik}-\frac{e^{\Delta_{a_k}+\Delta_{b_k}(\hat{V}_i-v)/h+\bar{m}_k(\hat{V}_i,v)}}{1+\sum_{k=1}^{K-1}e^{\Delta_{a_k}+\Delta_{b_k}(\hat{V}_i-v)/h+\bar{m}_k(\hat{V}_i,v)}}\right\}$$

$\Delta_a = (\Delta_{a_1}, \cdots, \Delta_{a_{K-1}})'$, $\Delta_b = (\Delta_{b_1}, \cdots, \Delta_{b_{K-1}})'$ and $\bar{m}_k(u, v) = m_k(v) + \dot{m}_k(v)(u - v)$. Following the similar arguments used in Cai et al. (2008), one may show that the changes in $\hat{Q}_k(\Delta_a, \Delta_b; v)$ by replacing $\hat{G}(\cdot)$ and $\hat{V}_i$ by $G(\cdot)$ and V_i, respectively, are asymptotically negligible. Let $Q_k(\Delta_a, \Delta_b; h, v)$ be

$$n^{-1}\sum_{i=1}^{n}\frac{w_iK_h(V_i-v)}{G(T_i\wedge t_0)}\begin{pmatrix}1\\(V_i-v)/h\end{pmatrix}\left\{Y_{ik}-\frac{e^{\Delta_{a_k}+\Delta_{b_k}(V_i-v)/h+\bar{m}_k(V_i,v)}}{1+\sum_{k=1}^{K-1}e^{\Delta_{a_k}+\Delta_{b_k}(V_i-v)/h+\bar{m}_k(V_i,v)}}\right\},$$

and write difference $\hat{Q}_k(\Delta_a, \Delta_b; v) - Q_k(\Delta_a, \Delta_b; v)$ as

$$-n^{-1}\sum_{i=1}^{n}\{\hat{G}(T_i\wedge t_0)-G(T_i\wedge t_0)\}\frac{w_iK_h(\hat{V}_i-v)}{\hat{G}(T_i\wedge t_0)G(T_i\wedge t_0)}\begin{pmatrix}1\\(\hat{V}_i-v)/h\end{pmatrix}$$
$$\left\{Y_{ik}-\frac{\exp\{\Delta_{a_k}+\Delta_{b_k}(\hat{V}_i-v)/h+\bar{m}_k(\hat{V}_i,v)\}}{1+\sum_{k=1}^{K-1}\exp\{\Delta_{a_k}+\Delta_{b_k}(\hat{V}_i-v)/h+\bar{m}_k(\hat{V}_i,v)\}}\right\}+\int_{v-h}^{v+h}K_h(s-v)$$
$$d\mathbf{P}_n\left[I(s<\hat{\beta}'X)\begin{pmatrix}1\\(\hat{\beta}'X-v)/h\end{pmatrix}\left\{Y_k-\frac{e^{\Delta_{a_k}+\Delta_{b_k}(\hat{\beta}'X-v)/h+\bar{m}_k(\hat{\beta}'X,v)}}{1+\sum_{k=1}^{K-1}e^{\Delta_{a_k}+\Delta_{b_k}(\hat{\beta}'X-v)/h+\bar{m}_k(\hat{\beta}'X,v)}}\right\}\right.$$
$$\left.-I(s<\beta_0'X)\begin{pmatrix}1\\(\beta_0'X-v)/h\end{pmatrix}\left\{Y_k-\frac{e^{\Delta_{a_k}+\Delta_{b_k}(\beta_0'X-v)/h+\bar{m}_k(\beta_0'X,v)}}{1+\sum_{k=1}^{K-1}e^{\Delta_{a_k}+\Delta_{b_k}(\beta_0'X-v)/h+\bar{m}_k(\beta_0'X,v)}}\right\}\right]$$
$$\times\frac{w}{G(T\wedge t_0)},$$

which is bounded by

$$\sup_t|\hat{G}(t)-G(t)|\times n^{-1}\sum_{i=1}^{n}\frac{w_iK_h(\hat{V}_i-v)}{\hat{G}(T_i\wedge t_0)G(T_i\wedge t_0)}\begin{pmatrix}1\\(\hat{V}_i-v)/h\end{pmatrix}\left\{Y_{ik}\right.$$
$$\left.-\frac{\exp\{\Delta_{a_k}+\Delta_{b_k}(\hat{V}_i-v)/h+\bar{m}_k(\hat{V}_i,v)\}}{1+\sum_{k=1}^{K-1}\exp\{\Delta_{a_k}+\Delta_{b_k}(\hat{V}_i-v)/h+\bar{m}_k(\hat{V}_i,v)\}}\right\}+O_p(h^{-1}n^{-1/2})\times$$
$$\left(\mathbf{G}_n\left[I(s<\hat{\beta}'X)\begin{pmatrix}1\\(\hat{\beta}'X-v)/h\end{pmatrix}\left\{Y_k-\frac{e^{\Delta_{a_k}+\Delta_{b_k}(\hat{\beta}'X-v)/h+\bar{m}_k(\hat{\beta}'X,v)}}{1+\sum_{k=1}^{K-1}e^{\Delta_{a_k}+\Delta_{b_k}(\hat{\beta}'X-v)/h+\bar{m}_k(\hat{\beta}'X,v)}}\right\}\right.\right.$$
$$\left.-I(s<\beta_0'X)\begin{pmatrix}1\\(\beta_0'X-v)/h\end{pmatrix}\left\{Y_k-\frac{e^{\Delta_{a_k}+\Delta_{b_k}(\beta_0'X-v)/h+\bar{m}_k(\beta_0'X,v)}}{1+\sum_{k=1}^{K-1}e^{\Delta_{a_k}+\Delta_{b_k}(\beta_0'X-v)/h+\bar{m}_k(\beta_0'X,v)}}\right\}\right]\times$$
$$\left.\frac{w}{G(T\wedge t_0)}+O_p(h^2+n^{-1/2})\right)=O_p\{n^{-1/2}+(nh^2)^{-3/4+\delta_0}+(nh)^{-1}\},$$

for some small $\delta_0 > 0$, where $\mathbf{P}_n$ and $\mathbf{P}$ are the expectation operator with respect to the empirical distribution of $\{(T_i, \Delta_i, \epsilon_i, U_i), i = 1, \cdots, n\}$ and the distribution of (T, Δ, ϵ, U), respectively, and $\mathbf{G}_n = n^{1/2}(\mathbf{P}_n - \mathbf{P})$. Furthermore, since $Q(\Delta_a, \Delta_b; v) = \{Q_1(\Delta_a, \Delta_b; v), \cdots, Q_{K-1}(\Delta_a, \Delta_b; v)\}'$ can be written as sum of n identically distributed independent random functions, it follows from the standard arguments that it uniformly converges to $q(\Delta_a, \Delta_b; v) = \{q_1(\Delta_a, \Delta_b; v), \cdots, q_{K-1}(\Delta_a, \Delta_b; v)\}'$, where

$$q_k(\Delta_a, \Delta_b; v) = \begin{pmatrix} g_0(v) \int K(x) \left[\eta_k(v) - \frac{\exp\{\Delta_{a_k} + \Delta_{b_k} x + m_k(v)\}}{1+\sum_{k=1}^{K-1} \exp\{\Delta_{a_k} + \Delta_{b_k} x + m_k(v)\}}\right] dx \\ g_0(v) \int x K(x) \left[\eta_k(v) - \frac{\exp\{\Delta_{a_k} + \Delta_{b_k} x + m_k(v)\}}{1+\sum_{k=1}^{K-1} \exp\{\Delta_{a_k} + \Delta_{b_k} x + m_k(v)\}}\right] dx \end{pmatrix},$$

and $g_0(\cdot)$ is the density function of the random variable $\beta_0' X$. Since $(\Delta_a', \Delta_b')' = (0', 0')'$ is the unique solution of $q(\Delta_a, \Delta_b; v) = 0$. $\hat{\Delta}_a(v)$ and $\hat{\Delta}_b(v)$ converge to zero uniformly in v, assuming that the "slope" matrix of $q(\Delta_a, \Delta_b; v)$ is nonsingular. Coupled with the consistency of $\hat{\Delta}_a$ and $\hat{\Delta}_b$, the Taylor series expansion can be used to show that

$$\hat{\Delta}_a(v) = \mathbf{A}(u)^{-1} n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(\hat{V}_i - v)}{\hat{G}(T_i \wedge t_0)} \begin{pmatrix} Y_{i1} - \frac{\exp\{\bar{m}_1(\hat{V}_i, v)\}}{1+\sum_{k=1}^{K-1} \exp\{\bar{m}_k(\hat{V}_i, v)\}} \\ \cdots \\ Y_{i(K-1)} - \frac{\exp\{\bar{m}_{K-1}(\hat{V}_i, v)\}}{1+\sum_{k=1}^{K-1} \exp\{\bar{m}_k(\hat{V}_i, v)\}} \end{pmatrix} + o_p\{(nh)^{-1/2}\},$$

where

$$\mathbf{A}(u) = g_0(v) \begin{pmatrix} \eta_1(u)\{1 - \eta_1(u)\} & -\eta_1(u)\eta_2(u) & \cdots & -\eta_1(u)\eta_{K-1}(u) \\ -\eta_2(u)\eta_1(u) & \eta_2(u)\{1 - \eta_2(u)\} & \cdots & -\eta_2(u)\eta_{K-1}(u) \\ \cdots & \cdots & \cdots & \cdots \\ -\eta_{K-1}(u)\eta_1(u) & -\eta_{K-1}(u)\eta_2(u) & \cdots & \eta_{K-1}(u)\{1 - \eta_{K-1}(u)\} \end{pmatrix}.$$

Again, following the similar arguments for estimating the asymptotical order of (B.2) in Appendix B of Cai et al. (2008), one may show that

$$n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(\hat{V}_i - v)}{\hat{G}(T_i \wedge t_0)} \left\{ Y_{ik} - \frac{\exp\{\bar{m}_k(\hat{V}_i, v)\}}{1 + \sum_{k=1}^{K-1} \exp\{\bar{m}_k(\hat{V}_i, v)\}} \right\}$$

can be approximated by

$$n^{-1} \sum_{i=1}^{n} \frac{w_i K_h(V_i - v)}{G(T_i \wedge t_0)} \left\{ Y_{ik} - \frac{\exp\{\bar{m}_k(V_i, v)\}}{1 + \sum_{k=1}^{K-1} \exp\{\bar{m}_k(V_i, v)\}} \right\}$$

uniformly in v up to an order of $o_p\{(nh)^{-1/2}\}$ for $h = n^{-\nu}, \nu \in (1/5, 1/2)$. Noting that the consistent estimator for $\eta_k(v)$ is

$$\hat{\eta}_k(v) = \frac{\exp\{\hat{a}_k(v)\}}{1 + \sum_{j=1}^{K-1} \exp\{\hat{a}_j(v)\}},$$

by $\delta-$method we have

$$\begin{aligned}
&f\{\eta^*(v)\} - f\{\eta(v)\} \\
&=\mathbf{D}(v)\mathbf{A}(u)\hat{\Delta}_a(v)/g_0(v) + o_p\{(nh)^{-1/2}\} \\
&=\mathbf{D}(v)n^{-1}\sum_{i=1}^{n} \frac{w_i K_h(V_i - v)}{g_0(v)G(T_i \wedge t_0)}\varsigma_i(v) + o_p\{(nh)^{-1/2}\}
\end{aligned}$$

where $\mathbf{D}(v) = \mathrm{diag}[\dot{f}\{\eta_1(v)\}, \cdots, \dot{f}\{\eta_{K-1}(v)\}]$, $\dot{f}(\cdot)$ is the derivative of $f(\cdot)$, and

$$\varsigma_i(v) = \begin{pmatrix} Y_{i1} - \frac{\exp\{\bar{m}_1(V_i,v)\}}{1+\sum_{k=1}^{K-1}\exp\{\bar{m}_k(V_i,v)\}} \\ \cdots \\ Y_{i(K-1)} - \frac{\exp\{\bar{m}_{K-1}(V_i,v)\}}{1+\sum_{k=1}^{K-1}\exp\{\bar{m}_k(V_i,v)\}} \end{pmatrix}.$$

Therefore by the central limit theorem

$$(nh)^{1/2}[f\{\hat{\eta}(v)\} - f\{\eta(v)\}] \rightarrow N\{0, \Sigma(v)\},$$

in distribution as $n \rightarrow \infty$.

To justify the consistency of the variance-covariance matrix estimator $\hat{\Sigma}(v)$ based on the resampling method, we first note that following the same arguments above, we have

$$f\{\eta^*(v)\} - f\{\eta(v)\} = \mathbf{D}(v)n^{-1}\sum_{i=1}^{n} \frac{w_i K_h(V_i - v)}{g_0(v)G(T_i \wedge t_0)}\varsigma_i(v)B_i + o_{P^*}\{(nh)^{-1/2}\},$$

where the probability measure P^* is on the joint product spaces of the random data and $\{B_i\}$. Therefore $(nh)^{1/2}[f\{\eta^*(v)\} - f\{\eta(v)\}]$ is asymptotically equivalent to

$$\mathbf{D}(v)n^{-1}\sum_{i=1}^{n} \frac{w_i K_h(V_i - v)}{g_0(v)G(T_i \wedge t_0)}\varsigma_i(v)(B_i - 1)$$

whose conditional variance is

$$\mathbf{D}(v)n^{-1}\sum_{i=1}^{n} h\left\{\frac{w_i K_h(V_i - v)}{g_0(v)G(T_i \wedge t_0)}\right\}^2 \varsigma_i(v)^{\otimes 2}\mathbf{D}(v)$$

which converges to $\Sigma(v)$ in probability as $n \rightarrow \infty$. Therefore, we have shown that $P(|\hat{\Sigma}(v) - \Sigma(v)| > \epsilon \mid \text{ data })$ converges to 0 for any $\epsilon > 0$.

Appendix B: The Justification of the Resampling Methods

To justify the resampling-based variance estimator, note that the variance estimator $\tilde{\sigma}_k^2(v)$ can be approximated by

$$\frac{\dot{f}^2\{\hat{\eta}_k(v)\}}{\hat{g}_0^2(v)} n^{-1} \sum_{i=1}^n \mathrm{E}_{B_i} \left[\frac{w_i K_h(V_i^* - v)}{G^*(T_i \wedge t_0)} \{Y_{ik} - \frac{e^{\hat{a}_k(v)+\hat{b}_k(V_i^*-v)}}{1+\sum_{k=1}^{K-1} e^{\hat{a}_k(v)+\hat{b}_k(V_i^*-v)}}\} B_i \right]^2$$
$$= \frac{\dot{f}^2\{\hat{\eta}_k(v)\}}{\hat{g}_0^2(v)} n^{-1} \sum_{i=1}^n \left[\frac{w_i K_h(\hat{V}_i - v)}{\hat{G}(T_i \wedge t_0)} \{Y_{ik} - \frac{e^{\hat{a}_k(v)+\hat{b}_k(\hat{V}_i-v)}}{1+\sum_{k=1}^{K-1} e^{\hat{a}_k(v)+\hat{b}_k(\hat{V}_i-v)}}\} \right]^2 + o_p(1),$$

which uniformly converges to σ_k^2, the asymptotical variance of $n^{1/2}[f\{\hat{\eta}_k(v)\} - f\{\eta_k(v)\}]$, in probability as $n \to \infty$, where E_{B_i} is the expectation with respect to the random weights $\{B_1, \cdots, B_n\}$, which are independent of the observed data. The first approximation follows from the fact that $|\hat{\beta}^* - \hat{\beta}| + \sup_t |G^*(t) - \hat{G}(t)|$ is in the order of $O_p(n^{-1/2})$ and similar arguments used to bound the difference between $\hat{Q}_k(\Delta_a, \Delta_b; v)$ and $Q_k(\Delta_a, \Delta_b; v)$.

To justify the proposed procedure for constructing the simultaneous confidence band of $\eta_k(v), v \in \mathcal{I}$, first note that we have already established that uniformly in v,

$$(nh)^{1/2}[f\{\hat{\eta}_k(v)\} - f\{\eta_k(v)\}]$$
$$=(nh)^{1/2} \frac{\dot{f}\{\eta_k(v)\}}{g_0(v)} n^{-1} \sum_{i=1}^n \frac{w_i K_h(V_i - v)}{G(T_i \wedge t_0)} \left\{ Y_{i1} - \frac{e^{\bar{m}_k(V_i,v)}}{1+\sum_{k=1}^{K-1} e^{\bar{m}_k(V_i,v)}} \right\} + o_p(n^{-\delta_0})$$
$$=(nh)^{-1/2} \sum_{i=1}^n K\left(\frac{V_i - v}{h}\right) \xi_{ki} + o_p(n^{-\delta_0})$$
$$=h^{-1/2} \int K\left(\frac{x-v}{h}\right) y dZ_{nk}(x,y) + o_p(n^{-\delta_0})$$

for some $\delta_0 > 0$, where

$$\xi_{ki} = \frac{\dot{f}\{\eta_k(V_i)\} w_i}{g_0(V_i) G(T_i \wedge t_0)} \{Y_{ik} - \eta_k(V_i)\},$$

$Z_{kn}(x,y) = n^{1/2}\{n^{-1}\sum_{i=1}^n I(V_i, \xi_{ki}) - F_k(x,y)\}$, and $F_k(x,y)$ is the CDF of (V, ξ_k). From the strong approximation theorem (Tusnady, 1977), one may construct a sequence of standard bivariate Brownian bridge processes $B_n(x,y)$ such that

$$\sup_{x,y} |B_n\{M(x,y)\} - Z_{nk}(x,y)| = O_p(n^{-1/2}\{\log(n)\}^2),$$

where $M(x,y)$ is the Rosenblatt transformation such that $M(V_i, \xi_i)$ is uniformly distributed on the unit square. Note that the similar strong approximation result for empirical process with more than 2 dimensions is not established yet. Therefore, by integration by part, $(nh)^{1/2}[f\{\hat{\eta}_k(v)\} - f\{\eta_k(v)\}]/\tilde{\sigma}_k(v)$ can be further approximately uniformly by

$$\mathcal{Y}_{1,n}(v) = \frac{1}{h^{1/2}\sigma_k(v)} \int K\left(\frac{x-v}{h}\right) y dB_n\{M(x,y)\}.$$

Let

$$\mathcal{Y}_{2,n}(v) = \frac{1}{h^{1/2}\sigma_k(v)} \int K\left(\frac{x-v}{h}\right) y dW_n\{M(x,y)\},$$

$$\mathcal{Y}_{3,n}(v) = \frac{1}{h^{1/2}\sigma_k(v)} \int \sigma_k(x) K(\frac{x-v}{h}) dW(x) \quad \text{and} \quad \mathcal{Y}_{4,n}(v) = \frac{1}{h^{1/2}} \int K(\frac{x-v}{h}) dW(x),$$

where $W_n(\cdot,\cdot)$ be a sequence of bivariate Wiener processes satisfying that

$$B_n(x,y) = W_n(x,y) - xyW_n(1,1)$$

and $W(\cdot)$ is the one-dimensional Wiener process. Following the similar arguments in Bickel and Rosenblatt (1973), we have the approximation $\sup_{\mathcal{I}} |\mathcal{Y}_{1,n}(v) - \mathcal{Y}_{2,n}(v)| = O_p(h^{1/2})$ and $\sup_{\mathcal{I}} |\mathcal{Y}_{3,n}(v) - \mathcal{Y}_{4,n}(v)| = O_p(h^{1/2})$. This, coupled with the fact that $\mathcal{Y}_{2,n}(v)$ and $\mathcal{Y}_{3,n}(v)$ are mean zero Gaussian processes with the identical variance-covariance function, implies that

$$\mathcal{S} = \sup_{\mathcal{I}} (nh)^{1/2} \frac{|f\{\hat{\eta}_k(v)\} - f\{\eta_k(v)\}|}{\tilde{\sigma}_k(v)} = \sup_{\mathcal{I}} \frac{1}{h^{1/2}} \int K(\frac{x-v}{h}) dW(x) + o_p(n^{-\delta_0}),$$

for some $\delta_0 > 0$. Therefore, it follows from Bickel and Rosenblatt (1973) that

$$\text{pr}(\{-2\log(h)\}^{1/2}(\mathcal{S} - d_h) \leqslant s) = \exp(-2e^{-x}) + o(1),$$

as $n \to \infty$, where

$$d_h = \{-2\log(h)\}^{1/2} + \frac{1}{\{-2\log(h)\}^{1/2}} \log\left\{\frac{\int \dot{K}(t)^2 dt}{4\pi \int K(t)^2 dt}\right\}.$$

Unlike the supremum value of tight processes, $\mathcal{S}$ itself does not converge in distribution, since $d_h \to \infty$ as $n \to \infty$. In parallel arguments $\mathcal{S}^*$, the resampling counterpart of $\mathcal{S}$, is equivalent to

$$\begin{aligned}
&\sup_{\mathcal{I}} \left| \frac{1}{(nh)^{1/2}\sigma_k(v)} \sum_{i=1}^{n} K\left(\frac{V_i - v}{h}\right) \hat{\xi}_{ki} B_i \right| + o_p(n^{-\delta_0}) \\
=& \sup_{\mathcal{I}} \left| \frac{1}{(nh)^{1/2}\sigma_k(v)} \int K\left(\frac{x-v}{h}\right) y dW_n^*\{M^*(x,y)\} \right| + o_p(n^{-\delta_0})
\end{aligned}$$

for some $\delta_0 > 0$ where

$$\hat{\xi}_{ki} = \frac{\dot{f}\{\hat{\eta}_k(\hat{V}_i)\}w_i}{\hat{g}_0(\hat{V}_i)\hat{G}(T_i \wedge t_0)}\{Y_{ik} - \hat{\eta}_k(\hat{V}_i)\}$$

and $W_n^*\{M^*(x,y)\}$ is a sequence of mean zero Gaussian processes, whose covariance function is identical to that of $W_n\{M(x,y)\}$ conditional on the observed data. Let $\mathcal{T}^* = \{-2\log(h)\}^{1/2}(\mathcal{S}^* - d_h)$ and $\mathcal{T} = \{-2\log(h)\}^{1/2}(\mathcal{S} - d_h)$. It follows that

$$|\mathrm{pr}_B(\mathcal{T}^* \leqslant s) - \mathrm{pr}(\mathcal{T} \leqslant s)| = o_p(n^{-\delta_0}),$$

which implies that we can use the conditional distribution of $\mathcal{S}^*$ to approximate that of $\mathcal{S}$, where pr_B is conditional on the observed data.