Local Polynomial Modelling for Varying-Coefficient Informative Survival Models

Wenyang Zhang[1], Yan Sun[2], Jin-Ting Zhang[3] and Duolao Wang[4]

[1] *University of Bath*, [2] *Shanghai University of Finance and Economics*,

[3] *National University of Singapore and* [4] *London School of Hygiene and Tropical Medicine*

Supplementary Material

This note contains conditions and proofs for the asymptotic normality of the local log-likelihood estimator.

S1. Notation

Let $N_{1i}(s) = I(t_i \le s,\ \delta_i = 1),\ N_{2i}(s) = I(t_i \le s,\ \delta_i = 0),\ T_i(s) = I(t_i \ge s)$, and

$$\boldsymbol{\vartheta} = (\mathbf{a}_0^{\mathrm{T}},\ \mathbf{a}_1^{\mathrm{T}},\ \mathbf{a}_2^{\mathrm{T}},\ \mathbf{b}_0^{\mathrm{T}},\ \mathbf{b}_1^{\mathrm{T}},\ \mathbf{b}_2^{\mathrm{T}})^{\mathrm{T}},$$

$$X_{i1}^* = (X_{i1}^{\mathrm{T}},\ X_{i2}^{\mathrm{T}},\ \mathbf{0}_{1\times(p-p_1)},\ X_{i1}^{\mathrm{T}}(U_i-u),\ X_{i2}^{\mathrm{T}}(U_i-u),\ \mathbf{0}_{1\times(p-p_1)})^{\mathrm{T}},$$

$$X_{i2}^* = (X_{i1}^{\mathrm{T}},\ \mathbf{0}_{1\times(p-p_1)},\ X_{i2}^{\mathrm{T}},\ X_{i1}^{\mathrm{T}}(U_i-u),\ \mathbf{0}_{1\times(p-p_1)},\ X_{i2}^{\mathrm{T}}(U_i-u))^{\mathrm{T}}.$$

The local log-likelihood function ℓ_2 can be expressed in terms of counting process as

$$\ell_3(\boldsymbol{\vartheta},\ \tau) = \sum_{l=1}^{2}\sum_{i=1}^{n}\int_0^{\tau} K_h(U_i-u)\log\frac{g_l(\boldsymbol{\vartheta}^{\mathrm{T}}X_{il}^*)}{\sum_{j=1}^{n} T_j(v)g_l(\boldsymbol{\vartheta}^{\mathrm{T}}X_{jl}^*)K_h(U_j-u)}dN_{li}(v) \qquad \text{(S1.1)}$$

where we omitted D as it is independent of $\boldsymbol{\vartheta}$.

Let $\Big(\Omega,\ \mathcal{F}, P_{(\ \boldsymbol{\eta},\ h_{0,y},\ h_{0,c})}\Big)$ be the sample space equipped with a right-continuous nondecreasing family of σ- algebras $(\mathcal{F}_s\ :\ s\in[0,\ \tau])$, where $\mathcal{F}_s = \sigma\Big\{t_i \le v,\ X_i,\ U_i,\ T_i(v),\ i =$

$1, 2, \cdots, n,\ 0 \le v \le s\Big\}$. All stochastic processes in this paper are assumed to be $\mathcal{F}_s$ measurable. Define $M_{li}(s) = N_{li}(s) - \int_0^s \lambda_{li}(v)dv,\ l = 1, 2,\ i = 1, 2, \cdots, n$, where λ_{li} is the intensity process of N_{li}. Obviously, $M_{li}(s)$ is a $\mathcal{F}_s$ martingale and $M_{li}(s)$ are orthogonal with predictable variation process

$$< M_{li} > (s) =< M_{li},\ M_{li} > (s) = \int_0^s \lambda_{li}(v)dv.$$

Let $\boldsymbol{\zeta} = H(\boldsymbol{\vartheta} - \boldsymbol{\xi})$, that is $\boldsymbol{\vartheta} = H^{-1}\boldsymbol{\zeta} + \boldsymbol{\xi}$, then (S1.1) can be reparametrized to

$$\begin{aligned}\ell_4(\boldsymbol{\zeta},\ \tau) &= \ell_3(H^{-1}\boldsymbol{\zeta} + \boldsymbol{\xi},\ \tau)\\ &= \sum_{l=1}^{2}\sum_{i=1}^{n}\int_0^\tau K_h(U_i - u)\log\frac{g_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^* + \boldsymbol{\xi}^{\mathrm{T}}X_{il}^*)}{\sum\limits_{j=1}^{n} T_j(v)g_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{jl}^* + \boldsymbol{\xi}^{\mathrm{T}}X_{jl}^*)K_h(U_j - u)}dN_{li}(v),\end{aligned} \tag{S1.2}$$

where $W_{il}^* = H^{-1}X_{il}^*$.

For easy description, we write $h_1(v) = h_{0,y}(v),\ h_2(v) = h_{0,c}(v)$, $Z_1^* = (X^{\mathrm{T}}, \mathbf{0}_{1\times(p-p_1)})^{\mathrm{T}},\ Z_2^* = (X_{01}^{\mathrm{T}}, \mathbf{0}_{1\times(p-p_1)}, X_{02}^{T})^{\mathrm{T}},\ Z_l(s) = (Z_l^{*\mathrm{T}}, Z_l^{*\mathrm{T}}s)^{\mathrm{T}},\ X^{\otimes 0} = 1,\ X^{\otimes 1} = X,\ X^{\otimes 2} = XX^{\mathrm{T}}$. We use $\mathbf{0}$ without subscript to denote $\mathbf{0}_{2q\times 1}$, and $g_l^{(k)}(\cdot)$ to denote the kth derivative of $g_l(\cdot)$, $k = 0,\ 1,\ 2$, $l = 1,\ 2$. For any matrix A and vector $\mathbf{a}$, let $\|A\| = \sup_{i,\ j}|a_{ij}|,\ \|\mathbf{a}\| = \sup_i |a_i|$ and $|\mathbf{a}| = (\mathbf{a}^{\mathrm{T}}\mathbf{a})^{1/2}$. For $l = 1,\ 2,\ k = 0,\ 1,\ 2$, let

$$A_{nlk}(\boldsymbol{\zeta},\ v,\ u) = \frac{1}{n}\sum_{i=1}^{n} K_h(U_i - u)T_i(v)g_l^{(k)}(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^* + \boldsymbol{\xi}^{\mathrm{T}}X_{il}^*)(W_{il}^*)^{\otimes k},$$

$$A_{nl0}^*(v,\ u) = \frac{1}{n}\sum_{i=1}^{n} K_h(U_i - u)T_i(v)g_l(\theta_{0l}(U_i)^{\mathrm{T}}X_i),$$

$$a_{lk}(\boldsymbol{\zeta},\ v,\ u) = f(u)\int E\Big[P(v,\ X,\ u)g_l^{(k)}(\boldsymbol{\zeta}^{\mathrm{T}}Z_l(s) + \theta_{0l}(u)^{\mathrm{T}}X)Z_l(s)^{\otimes k}|U = u\Big]K(s)ds,$$

$$\alpha_{lk}(v,\ u) = f(u)E\Big[P(v,\ X,\ u)g_l^{(k)}(\theta_{0l}(u)^{\mathrm{T}}X)Z_l^{*\otimes k}|U = u\Big],$$

$$\rho_l(v,\ u) = f(u)E\Big\{P(v,\ X,\ u)[g_l^{'}(\theta_{0l}(u)^{\mathrm{T}}X)]^2/g_l(\theta_{0l}(u)^{\mathrm{T}}X)Z_l^{*\otimes 2}|U = u\Big\},$$

$$\rho_l(u) = \int_0^\tau \rho_l(v,\ u)h_l(v)dv, \quad \Gamma_l(u) = \rho_l(u) - \int_0^\tau \alpha_{l1}(v,\ u)^{\otimes 2}\alpha_{l0}(v,\ u)^{-1}h_l(v)dv.$$

S2. Condition

The following conditions are imposed to establish the asymptotic normality of the proposed estimator.

(1) The kernel function $K(t)$ is a bounded and symmetric density function with compact support.

(2) The functions $\boldsymbol{\beta}_l(\cdot), l = 0, 1, 2,\ g_1(\cdot),\ g_2(\cdot)$ have continuous second derivatives around the point u and $g_1(\cdot),\ g_2(\cdot)$ are strictly positive.

(3) The marginal density $f(\cdot)$ of U has a continuous derivative in some neighborhood of u, and $f(u) \neq 0$

(4) $\int_0^\tau h_{0,y}(v)dv < \infty,\ \int_0^\tau h_{0,c}(v)dv < \infty.$

(5) The conditional probability $P(v,\ x,\ u)$ is continuous with respect to u.

(6) $n \to \infty,\ h \to 0,\ nh/\log n \to \infty,\ nh^5$ is bounded.

(7) (Asymptotic regularity conditions) The matrix $\Delta(u)$ is positive definite at u, $\Omega(u)$ is nonsingular at u.

(8) (Lindeberg condition)

$$(nh)^{-\frac{1}{2}} \sup_{l,i,s} \left| \frac{{g_l}'(\boldsymbol{\beta}_0(U_i)^{\mathrm{T}} X_{i1} + \boldsymbol{\beta}_l(U_i)^{\mathrm{T}} X_{i2})}{g_l(\boldsymbol{\beta}_0(U_i)^{\mathrm{T}} X_{i1} + \boldsymbol{\beta}_l(U_i)^{\mathrm{T}} X_{i2})} X_i \right| T_i(s) \xrightarrow{P} 0.$$

S3. Proof of asymptotic normality

We first present some useful lemmas.

Lemma S3.1. *Under conditions (1)-(6), we have*

$$A_{nlk}(\boldsymbol{\zeta},\ v,\ u) = a_{lk}(\boldsymbol{\zeta},\ v,\ u) + o_p(1), \quad A^*_{nl0}(v,\ u) = \alpha_{l0}(v,\ u) + o_p(1),$$

$l = 1,\ 2,\ k = 0,\ 1,\ 2,$ *where* $\boldsymbol{\zeta}$ *lies in a neighborhood of* $\mathbf{0}$ *for fixed* u.

This Lemma follows immediately from the same argument for Lemma A.1 in Fan, *et al.*, 2006.

Lemma S3.2. *Suppose the* k*-variate counting process* $\mathbf{N}$ *has intensity process* $\boldsymbol{\lambda}$. *Let* $M_i(s) = N(s) - \int_0^s \lambda(v)dv,\ i = 1, \cdots, k,\ 0 < s \le \tau,\ \mathbf{M}(s) = (M_1(s), \cdots, M_k(s))^T,$ *and* $\mathbf{H}(s)$ *be a* $p \times k$ *matrix of locally bounded and predictable. Then* $\mathbf{M}(s)$ *and* $\int_0^s \mathbf{H}(v)d\mathbf{M}(v)$ *are local square integrable martingales with*

$$< \mathbf{M} > (s) = \Big(diag\int_0^s \boldsymbol{\lambda}(v)dv\Big), \qquad < \int \mathbf{H}(v)d\mathbf{M}(v) > (s) = \int_0^s \mathbf{H}(v)\ \Big(diag\boldsymbol{\lambda}(v)\Big)\ \mathbf{H}(v)^T dv,$$

where $\int_0^s \mathbf{H}(v)d\mathbf{M}(v)$ *is the* p *dimensional vector whose* j*th component,* $j = 1, \cdots, p$, *is the sum of integrals, with respect to the* k *components of* $\mathbf{M}(v)$*, of all entries on* j*th row of* $\mathbf{H}(v)$.

See Andersen, *et al.*,1993, Proposition II.4.1.

Lemma S3.3. (consistency of $\hat{\boldsymbol{\xi}}$) *Under conditions (1)-(7) , we have*

$$H(\hat{\boldsymbol{\xi}} - \boldsymbol{\xi}) \xrightarrow{P} \mathbf{0}.$$

Proof: By (S1.2), it follows that

$$\begin{aligned}
& n^{-1}\ell_4(\boldsymbol{\zeta},\ s) - n^{-1}\ell_4(\mathbf{0},\ s) \\
= & \ \frac{1}{n}\sum_{l=1}^{2}\sum_{i=1}^{n}\int_0^s K_h(U_i - u)\Big[\log\frac{g_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^* + \boldsymbol{\xi}^{\mathrm{T}}X_{il}^*)}{g_l(\boldsymbol{\xi}^{\mathrm{T}}X_{il}^*)} - \log\frac{A_{nl0}(\boldsymbol{\zeta},\ v,\ u)}{A_{nl0}(\mathbf{0},\ v,\ u)}\Big]dM_{li}(v) \\
& + \sum_{l=1}^{2}\Big\{\int_0^s B_{nl0}(\boldsymbol{\zeta},\ v,\ u)h_l(v)dv - \int_0^s \log\frac{A_{nl0}(\boldsymbol{\zeta},\ v,\ u)}{A_{nl0}(\mathbf{0},\ v,\ u)}A_{nl0}^*(v,\ u)h_l(v)dv\Big\} \\
\equiv & \ \sum_{l=1}^{2} J_{nl}^{(1)}(\boldsymbol{\zeta},\ s) + \sum_{l=1}^{2} J_{nl}^{(2)}(\boldsymbol{\zeta},\ s),
\end{aligned}$$

where $B_{nl0}(\boldsymbol{\zeta},\ v,\ u)=\frac{1}{n}\sum_{i=1}^{n}K_h(U_i-u)\log\frac{g_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^{*}+\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})}{g_l(\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})}T_i(v)g_l(\theta_{0l}(U_i)^{\mathrm{T}}X_i)h_l(v)dv$. For each $\boldsymbol{\zeta}$, $J_{nl}^{(1)}(\boldsymbol{\zeta},\ s)$ is a local square integrable martingale with

$$<J_{nl}^{(1)}(\boldsymbol{\zeta},\ \cdot)>(s)=n^{-2}\sum_{i=1}^{n}\int_0^s K_h^2(U_i-u)\Big[\log\frac{g_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^{*}+\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})}{g_l(\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})}-\log\frac{A_{nl0}(\boldsymbol{\zeta},\ v,\ u)}{A_{nl0}(\mathbf{0},\ v,\ u)}\Big]^2\lambda_{li}(v)dv.$$

By using the same argument as Lemma S3.1, it can be shown that $nh<J_{nl}^{(1)}(\boldsymbol{\zeta},\cdot)>(\tau)$ converges in probability to some finite quantity (depending on $\boldsymbol{\zeta}$). By the inequality of Lengart, we have $J_{nl}^{(1)}(\boldsymbol{\zeta},\ s)=O_p((nh)^{-\frac{1}{2}}),\ 0\le s\le\tau$.

Using the same argument as that for Lemma S3.1, we obtain

$$\begin{aligned}J_{nl}^{(2)}(\boldsymbol{\zeta},\ s) &= \int_0^s b_{l0}(\boldsymbol{\zeta},\ v,\ u)h_l(v)dv-\int_0^s\log\frac{a_{l0}(\boldsymbol{\zeta},\ v,\ u)}{a_{l0}(\mathbf{0},\ v,\ u)}\alpha_{l0}(v,\ u)h_l(v)dv+o_p(1)\\ &\equiv I_l(\boldsymbol{\zeta},\ s)+o_p(1),\end{aligned}$$

where $b_{l0}(\boldsymbol{\zeta},\ v,\ u)=f(u)\int E\Big[P(v,\ X,\ u)\log\frac{g_l(\boldsymbol{\zeta}^{\mathrm{T}}Z_l(s)+\theta_{0l}(u)^{\mathrm{T}}X)}{g_l(\theta_{0l}(u)^{\mathrm{T}}X)}g_l(\theta_{0l}(u)^{\mathrm{T}}X)|U=u\Big]K(s)ds$. Hence we have

$$n^{-1}\ell_4(\boldsymbol{\zeta},\ s)-n^{-1}\ell_4(\boldsymbol{\zeta},\ s)=I_1(\boldsymbol{\zeta},\ s)+I_2(\boldsymbol{\zeta},\ s)+o_p(1).$$

By simple calculation, we can see the first derivative of $\Big(I_1(\boldsymbol{\zeta},\ s)+I_2(\boldsymbol{\zeta},\ s)\Big)$ is zero at $\boldsymbol{\zeta}=\mathbf{0}$ and its second derivative is a negative definite matrix at $\boldsymbol{\zeta}=\mathbf{0}$ by condition (7).

Finally we will show that the local log-likelihood function is concave. Since the second derivative of $J_{nl}^{(2)}(\boldsymbol{\zeta},\ s)$ with respect to $\boldsymbol{\zeta}$ is

$$\begin{aligned}&\int_0^s\Bigg\{\frac{1}{n}\sum_{i=1}^{n}K_h(U_i-u)T_i(v)\gamma_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^{*}+\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})W_{il}^{*\otimes 2}\Big[g_l(\theta_{0l}(U_i)^{\mathrm{T}}X_i)-g_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^{*}+\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})\Big]\\ &\quad\times A_{nl0}(\boldsymbol{\zeta},\ v,\ u)-\frac{1}{n}\sum_{i=1}^{n}K_h(U_i-u)T_i(v)\gamma_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^{*}+\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})g_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^{*}+\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})W_{il}^{*\otimes 2}\\ &\quad\times\frac{1}{n}\sum_{i=1}^{n}K_h(U_i-u)T_i(v)\Big[g_l(\theta_{0l}(U_i)^{\mathrm{T}}X_i)-g_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^{*}+\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})\Big]\Bigg\}\frac{1}{A_{nl0}(\boldsymbol{\zeta},\ v,\ u)}\\ &\quad-\frac{1}{n}\sum_{i=1}^{n}K_h(U_i-u)T_i(v)g_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^{*}+\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})\Big[\frac{g_l^{'}(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^{*}+\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})}{g_l(\boldsymbol{\zeta}^{\mathrm{T}}W_{il}^{*}+\boldsymbol{\xi}^{\mathrm{T}}X_{il}^{*})}W_{il}^{*}-\frac{A_{nl1}(\boldsymbol{\zeta},\ v,\ u)}{A_{nl0}(\boldsymbol{\zeta},\ v,\ u)}\Big]^{\otimes 2}\\ &\quad\times\frac{A_{nl0}^{*}(v,\ u)}{A_{nl0}(\boldsymbol{\zeta},\ v,\ u)}h_l(v)dv,\end{aligned}$$

where $\gamma_l(\boldsymbol{\zeta}^{\mathrm{T}} W_{il}^* + \boldsymbol{\xi}^{\mathrm{T}} X_{il}^*)$ is the second derivative of $\log g_l(\boldsymbol{\zeta}^{\mathrm{T}} W_{il}^* + \boldsymbol{\xi}^{\mathrm{T}} X_{il}^*)$ with respect to $\boldsymbol{\zeta}$. Obviously, the last term is negatively semidefinite for any $\boldsymbol{\zeta}$ and with $\boldsymbol{\zeta}$ lies in a neighborhood of $\mathbf{0}$ for fixed u. Using the same argument as that for Lemma S3.1, we have that each component of the first two terms inside the bracket converges in probability to zero. Thus $J_{nl}^{(2)}(\boldsymbol{\zeta},\ s)$ is negatively semidefinite matrix for any $\boldsymbol{\zeta}$. Therefore, $n^{-1}\Big(\ell_4(\boldsymbol{\zeta},\ \tau) - \ell_4(\mathbf{0},\ \tau)\Big)$ is concave with maximiser being $\boldsymbol{\zeta} = \hat{\boldsymbol{\zeta}}$. Using the convex Theorem II.1 of Andersen & Gill (1982, Appendix II), we have $\hat{\boldsymbol{\zeta}}$ converges in probability to the maximiser of $\Big(I_1(\boldsymbol{\zeta},\ \tau) + I_2(\boldsymbol{\zeta},\ \tau)\Big)\ \mathbf{0}$.

Proof of Theorem: It is easy to see that

$$
\begin{aligned}
\tilde{\ell}_{n1}(\mathbf{0},\ \tau) &= n^{-1}\frac{\partial \ell_4(\boldsymbol{\zeta},\ \tau)}{\partial \boldsymbol{\zeta}}\Big|_{\boldsymbol{\zeta}=\mathbf{0}} \\
&= \frac{1}{n}\sum_{l=1}^{2}\sum_{i=1}^{n}\int_0^{\tau} K_h(U_i-u)\Big[\frac{g_l^{'}(\boldsymbol{\xi}^{\mathrm{T}} X_{il}^*)}{g_l(\boldsymbol{\xi}^{\mathrm{T}} X_{il}^*)}W_{il}^* - \frac{A_{nl1}(\mathbf{0},\ v,\ u)}{A_{nl0}(\mathbf{0},\ v,\ u)}\Big]dM_{li}(v) \\
&+ \frac{1}{n}\sum_{l=1}^{2}\sum_{i=1}^{n}\int_0^{\tau} K_h(U_i-u)\Big[\frac{g_l^{'}(\boldsymbol{\xi}^{\mathrm{T}} X_{il}^*)}{g_l(\boldsymbol{\xi}^{\mathrm{T}} X_{il}^*)}W_{il}^* - \frac{A_{nl1}(\mathbf{0},\ v,\ u)}{A_{nl0}(\mathbf{0},\ v,\ u)}\Big]T_i(v) g_l(\theta_{0l}(U_i)^{\mathrm{T}} X_i) h_l(v) dv \\
&\equiv \sum_{l=1}^{2} R_{l1}(\mathbf{0},\tau) + \sum_{l=1}^{2} R_{l2}(\mathbf{0},\tau).
\end{aligned}
$$

We first deal with $R_{l2}(\mathbf{0},\tau)$. When U_i in the small neighborhood of u, by Taylor's expansion, it can be shown that

$$
\begin{aligned}
\sum_{l=1}^{2} R_{l2}(\mathbf{0},\ \tau) &= \frac{1}{2}h^2\mu_2\Big\{\begin{pmatrix}\Gamma_1(u)\\ \mathbf{0}_{q\times q}\end{pmatrix}\begin{pmatrix}\boldsymbol{\beta}_0^{''}(u)\\ \boldsymbol{\beta}_1^{''}(u)\\ \mathbf{0}_{(p-p_1)\times 1}\end{pmatrix} + \begin{pmatrix}\Gamma_2(u)\\ \mathbf{0}_{q\times q}\end{pmatrix}\begin{pmatrix}\boldsymbol{\beta}_0^{''}(u)\\ \mathbf{0}_{(p-p_1)\times 1}\\ \boldsymbol{\beta}_2^{''}(u)\end{pmatrix}\Big\}(1+o_p(1)) \\
&\equiv r_n(u)(1+o_p(1)).
\end{aligned}
$$

Define $R_l^*(s) = \sqrt{nh}\ R_{l1}(\mathbf{0},\ s),\ l = 1,\ 2.$ Using same argument as that for Lemmas S3.1 and S3.2, we have that

$$
< (R_1^* + R_2^*) > (\tau) \xrightarrow{P} \begin{pmatrix}(\Gamma_1(u)+\Gamma_2(u))\nu_0 & \mathbf{0}_{q\times q}\\ \mathbf{0}_{q\times q} & (\rho_1(u)+\rho_2(u))\nu_2\end{pmatrix} \equiv \Sigma_1(u).
$$

It remains to be proved that for all $\epsilon > 0$,

$$\sum_{l=1}^{2}\sum_{i=1}^{n}\int_0^\tau \{S_{lij}(v)\}^2\lambda_{li}(v)I(|S_{lij}(v)|>\epsilon)dv \xrightarrow{P} 0,$$

where $S_{lij}(v) = \sqrt{\frac{h}{n}}\ K_h(U_i-u)\Big[\frac{g_l^{'}(\boldsymbol{\xi}^{\rm T}X_{il}^*)}{g_l(\boldsymbol{\xi}^{\rm T}X_{il}^*)}W_{il}^* - \frac{A_{nl1}(\mathbf{0},\ v,\ u)}{A_{nl0}(\mathbf{0},\ v,\ u)}\Big]_j,\ j=1,2,\cdots,2q.$

By using the elementary inequality

$$|\mathbf{a}-\mathbf{b}|^2 I(|\mathbf{a}-\mathbf{b}|>\epsilon) \le 4|\mathbf{a}|^2 I(|\mathbf{a}|>\frac{\epsilon}{2}) + 4|\mathbf{b}|^2 I(|\mathbf{b}|>\frac{\epsilon}{2})$$

and Taylor's expansion of $\boldsymbol{\beta}_j(U_i),\ j=0,\ 1,\ 2,$ at u and the continuity of $g_l(\cdot), {g_l}'(\cdot),\ l=1,2,$ together with condition (8), we obtain the above result.

Appealing Rebolledos's martingale central limit theorem, we have

$$\sqrt{nh}(\tilde{\ell}_{n1}(\mathbf{0},\ \tau) - r_n(u)) \xrightarrow{D} \mathbf{N}(\mathbf{0},\ \Sigma_1(u)). \tag{S3.1}$$

We are now going to show that $\tilde{\ell}_{n2}(\boldsymbol{\zeta}^*,\tau) = \frac{1}{n}\frac{\partial^2\ell_4(\boldsymbol{\zeta},\ \tau)}{\partial\boldsymbol{\zeta}\partial\boldsymbol{\zeta}^{\rm T}}\Big|_{\boldsymbol{\zeta}=\boldsymbol{\zeta}^*}$ converges in probability to a finite constant matrix for any random $\boldsymbol{\zeta}^*$ between $\mathbf{0}$ and $\hat{\boldsymbol{\zeta}}$. Since $\boldsymbol{\zeta}^* \xrightarrow{P} \mathbf{0}$, by the mean-value theorem we have that

$$\tilde{\ell}_{n2}(\boldsymbol{\zeta}^*,\ \tau) = \tilde{\ell}_{n2}(\mathbf{0},\ \tau) + o_p(1),$$

and

$$\begin{aligned}
\tilde{\ell}_{n2}(\mathbf{0},\ \tau) \ =\ & \frac{1}{n}\sum_{l=1}^{2}\sum_{i=1}^{n}\int_0^\tau K_h(U_i-u)\Big[\frac{g_l^{''}(\boldsymbol{\xi}^{\rm T}X_{il}^*)g_l(\boldsymbol{\xi}^{\rm T}X_{il}^*) - [{g_l}'(\boldsymbol{\xi}^{\rm T}X_{il}^*)]^2}{g_l^2(\boldsymbol{\xi}^{\rm T}X_{il}^*)}W_{il}^{*\otimes 2}\Big]dM_{li}(v) \\
& -\frac{1}{n}\sum_{l=1}^{2}\sum_{i=1}^{n}\int_0^\tau K_h(U_i-u)\frac{A_{nl2}(\mathbf{0},\ v,\ u)A_{nl0}(\mathbf{0},\ v,\ u) - A_{nl1}(\mathbf{0},\ v,\ u)^{\otimes 2}}{A_{nl0}^2(\mathbf{0},\ v,\ u)}dM_{li}(v) \\
& +\frac{1}{n}\sum_{l=1}^{2}\sum_{i=1}^{n}\int_0^\tau K_h(U_i-u)\Big[\frac{g_l^{''}(\boldsymbol{\xi}^{\rm T}X_{il}^*)g_l(\boldsymbol{\xi}^{\rm T}X_{il}^*) - [{g_l}'(\boldsymbol{\xi}^{\rm T}X_{il}^*)]^2}{g_l^2(\boldsymbol{\xi}^{\rm T}X_{il}^*)}W_{il}^{*\otimes 2} \\
& \quad - \frac{A_{nl2}(\mathbf{0},\ v,\ u)A_{nl0}(\mathbf{0},\ v,\ u) - A_{nl1}(\mathbf{0},\ v,\ u)^{\otimes 2}}{A_{nl0}^2(\mathbf{0},\ v,\ u)}\Big]\lambda_{li}(v)dv.
\end{aligned}$$

Similar to the proof of $J_{nl}^{(1)}(\boldsymbol{\zeta},\ s)$ in Lemma S3.3, it can be shown that each component in the first and second terms equal $O_p((nh)^{-\frac{1}{2}})$. By simple calculation, we can see the third term converges in probability to

$$-\begin{pmatrix} \Gamma_1(u)+\Gamma_2(u) & \mathbf{0}_{q\times q} \\ \mathbf{0}_{q\times q} & (\rho_1(u)+\rho_2(u))\mu_2 \end{pmatrix} \equiv -\Sigma_2(u).$$

Hence $\tilde{\ell}_{n2}(\boldsymbol{\zeta}^*,\ \tau) \stackrel{P}{\longrightarrow} -\Sigma_2(u)$.

As $\hat{\boldsymbol{\zeta}}$ maximizes $n^{-1}\ell_4(\boldsymbol{\zeta},\ \tau)$ and $\hat{\boldsymbol{\zeta}} \stackrel{P}{\longrightarrow} \mathbf{0}$, by Taylor's expansion at $\mathbf{0}$ and the above result, we have

$$\hat{\boldsymbol{\zeta}} - \Sigma_2(u)^{-1}r_n(u) = -\left(\tilde{\ell}_{n2}(\boldsymbol{\zeta}^*,\ \tau)\right)^{-1}\left(\tilde{\ell}_{n1}(\mathbf{0},\ \tau) - r_n(u)\right) + o_p(1).$$

This together with (S3.1) lead to

$$\sqrt{nh}\left(\hat{\boldsymbol{\zeta}} - \Sigma_2(u)^{-1}r_n(u)\right) \stackrel{D}{\longrightarrow} \mathbf{N}\left(\mathbf{0}, \Sigma_2(u)^{-1}\Sigma_1(u)\Sigma_2(u)^{-1}\right).$$

By simple and straightforward calculation, it follows that

$$r_n(u) = \frac{1}{2}h^2\mu_2 \begin{pmatrix} \Delta(u) \\ \mathbf{0}_{q\times q} \end{pmatrix} \boldsymbol{\eta}''(u), \quad \Sigma_2(u)^{-1}r_n(u) = \frac{1}{2}h^2\mu_2\mathbf{e}_{2q\times q}\boldsymbol{\eta}''(u),$$

and

$$\Sigma_2(u)^{-1}\Sigma_1(u)\Sigma_2(u)^{-1} = \begin{pmatrix} \Delta(u)^{-1}\nu_0 & \mathbf{0}_{q\times q} \\ \mathbf{0}_{q\times q} & \Omega(u)^{-1}\mu_2^{-2}\nu_2 \end{pmatrix}.$$