Semiparametric Inference for the Functional Cox Model

Meiling Hao[a], Kin-yat Liu[b], Wei Xu[c] and Xingqiu Zhao[d]

[a]School of Statistics, University of International Business and Economics, Beijing, China

[b]Department of Mathematics and Statistics, The Hang Seng University of Hong Kong, Hong Kong

[c]Dalla Lana School of Public Health, University of Toronto, Toronto, Canada

[d]Department of Applied Mathematics, The Hong Kong Polytechnic University, Hong Kong

Supplementary Materials

Appendix: Proofs

To show the theorems, we need some notations and lemmas. In the following, we will use c to denote different positive constants in different places, while $a \lesssim b$ means $a \leq cb$, and $a \gtrsim b$ means $a \geq cb$.

A Notation

For ease of presentation, we introduce more notation related to the Fréchet derivatives. Let $\mathcal{S}_n(\alpha)$ and $\mathcal{S}_{n,\lambda}(\alpha)$ be the Fréchet derivatives of $l_n(\alpha)$ and $l_{n,\lambda}(\alpha)$, respectively. Denote $l(\alpha)$ as the asymptotic value of $l_n(\alpha)$, and $l_\lambda(\alpha) = l(\alpha) - \lambda J(\beta, \beta)/2$. Similarly, let $\mathcal{S}(\alpha)$ and $\mathcal{S}_\lambda(\alpha)$ be the Fréchet derivatives of $l(\alpha)$ and $l_\lambda(\alpha)$, respectively. Let D be the Fréchet derivative operator and $\alpha_i = (\theta_i^\top, \beta_i(\cdot)), i = 1, 2, 3 \in \mathcal{H}$ be any direction. Then, we have

$$\mathcal{S}_{n,\lambda}(\alpha)\alpha_1 = \frac{1}{n}\sum_{i=1}^{n}\Delta_i\left[\eta_{\alpha_1}(W_i) - \frac{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_1}(W_j)}{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}}\right] - \lambda J(\beta,\beta_1)$$
$$\equiv \mathcal{S}_n(\alpha)\alpha_1 - \lambda J(\beta,\beta_1),$$

$$\begin{aligned}
&D\mathcal{S}_{n,\lambda}(\alpha)\alpha_1\alpha_2\\
&= -\frac{1}{n}\sum_{i=1}^{n}\Delta_i\left[\frac{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_1}(W_j)\eta_{\alpha_2}(W_j)}{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}}\right.\\
&\left.-\frac{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_1}(W_j)\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_2}(W_j)}{[\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}]^2}\right] - \lambda J(\beta_1,\beta_2)\\
&\equiv D\mathcal{S}_n(\alpha)\alpha_1\alpha_2 - \lambda J(\beta_1,\beta_2),
\end{aligned}$$

and

$$\begin{aligned}
&D^2\mathcal{S}_{n,\lambda}(\alpha)\alpha_1\alpha_2\alpha_3 = -\frac{1}{n}\sum_{i=1}^{n}\Delta_i\Big[\frac{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_1}(W_j)\eta_{\alpha_2}(W_j)\eta_{\alpha_3}(W_j)}{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}}\\
&-\frac{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_1}(W_j)\eta_{\alpha_2}(W_j)\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_3}(W_j)}{[\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}]^2}\\
&-\frac{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_1}(W_j)\eta_{\alpha_3}(W_j)\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_2}(W_j)}{[\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}]^2}\\
&-\frac{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_1}(W_j)\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_2}(W_j)\eta_{\alpha_3}(W_j)}{[\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}]^2}\\
&+2\frac{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_1}(W_j)\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_2}(W_j)}{[\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}]^3}\\
&\times\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_\alpha(W_j)\}\eta_{\alpha_3}(W_j)\Big]\\
&\equiv D^2\mathcal{S}_n(\alpha)\alpha_1\alpha_2\alpha_3.
\end{aligned}$$

There exists a sequence of functions $\omega_k, k = 1, 2, \ldots, p$, such that $\langle\omega_k, \beta\rangle_m = V(G_k, \beta)$. Direct calculations yield that $\omega_k(\cdot) = \sum_{j=1}^{\infty} G_{jk}h_j(\cdot)/(1 + \lambda\rho_j)$. Denote $\omega = (\omega_1, \omega_2, \ldots, \omega_p)^\top$. Thus,

$\omega = (id - W_\lambda)G$. Furthermore, by the Riesz representation theorem, there exists an element in $\mathcal{H}^{(m)}$, denoted as π_x, such that $\langle \pi_x, \beta \rangle_m = \int_{\mathbb{I}} x(t)\beta(t)\,dt$. Through direct calculations, we have $\pi_x = \sum_{j=1}^{\infty} \int_{\mathbb{I}} x(t)h_j(t)\,dt/(1+\lambda\rho_j)h_j(\cdot)$. Define $\mathcal{R}_w = (H_w, T_w)$ with $w = (z, x(\cdot))$, where

$$H_w = \{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1}\{z - V(G, \pi_x)\}, \quad \text{and}$$

$$T_w = \pi_x - \omega^\top \{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1}\{z - V(G, \pi_x)\}.$$

Then we have $\langle \mathcal{R}_w, \alpha \rangle_\lambda = \theta^\top z + \int_{\mathbb{I}} x(t)\beta(t)\,dt$.

Also define $\tilde{\mathcal{R}}_u$ as $\tilde{\mathcal{R}}_u : u \to (\tilde{H}_u, \tilde{T}_u) \in \mathcal{H}$, where $u = (z^\top, t)$,

$$\tilde{H}_u = \{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1}(z - \omega(t)), \quad \text{and}$$

$$\tilde{T}_u = K_t - \omega^\top \{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1}(z - \omega(t)).$$

Then we have that $\left\langle \tilde{\mathcal{R}}_u, \alpha \right\rangle_\lambda = \theta^\top z + \beta(t)$.

Define $\mathcal{P}_\lambda \alpha = (\tilde{H}_\alpha, \tilde{T}_\alpha)$, where

$$\tilde{H}_\alpha = -\{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1} V(G, W_\lambda \beta), \quad \text{and}$$

$$\tilde{T}_\alpha = W_\lambda \beta - \{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1} V(G, W_\lambda \beta).$$

Then $\mathcal{P}_\lambda \alpha \in \mathcal{H}$ and $\langle \mathcal{P}_\lambda \alpha, \alpha_1 \rangle_\lambda = \langle W_\lambda \beta, \beta_1 \rangle_m$ for any $\alpha_1 = (\theta_1^\top, \beta_1) \in \mathcal{H}$. It follows from the Cauchy-Schwarz inequality that $\|\mathcal{P}_\lambda\|_\lambda \le 1$ and $\mathcal{P}_\lambda$ is self-adjoint.

B Lemmas

Lemma B.1 *Under Condition (C1), we have* $D\mathcal{S}_\lambda(\alpha_0) = -id$*, where id is the identity operator.*

This result directly follows from the definitions of the inner product and $D\mathcal{S}_\lambda(\alpha_0)$.

Denote $\|\alpha\|_e = \|\theta\|_2 + \|\beta\|_{L_2}$. The following lemma provides the relationship between the general Euclidean norm $\|\cdot\|_e$ and the norm $\|\cdot\|_\lambda$.

Lemma B.2 *There exists a constant* $\kappa > 0$ *such that for any* $\alpha \in \mathcal{H}$, $\|\alpha\|_e \leq \kappa h^{-(2a+1)/2}\|\alpha\|_\lambda$.

Proof of Lemma B.2. For any $u = (z^\top, t)$, direct calculations yield

$$\left\langle \tilde{\mathcal{R}}_u, \tilde{\mathcal{R}}_u \right\rangle_\lambda = \{z - \omega(t)\}^\top \{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1}\{z - \omega(t)\} + K_t(t).$$

It follows from Condition (C3) that $\{\Sigma - V(G,G)\}^{-1}$ is positive definite and $V(G, W_\lambda G) \to 0$. Let c denote the minimum eigenvalue of $\{\Sigma - V(G,G) + V(G, W_\lambda G)\}$. Then we have $\{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1} \leq c^{-1}\mathbf{1}$ with $\mathbf{1}$ being the identity matrix. Thus, we have $\{z - \omega(t)\}^\top \{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1}\{z - \omega(t)\} \lesssim \|z - \omega(t)\|_2^2$. Then we have

$$\left\langle \tilde{\mathcal{R}}_u, \tilde{\mathcal{R}}_u \right\rangle_\lambda \lesssim \|z - \omega(t)\|_2^2 + ||K_t||_\lambda^2.$$

By the arguments in Cheng and Shang (2015, page 1379, lines 5-11), we hace $\sup_{t \in \mathbb{I}} |\omega(t)| = O(1)$, and it follows from the definition of K_t that $||K_t||_\lambda \lesssim h^{-(a+1/2)}$. This implies that

$||\tilde{\mathcal{R}}_u||_\lambda \lesssim h^{-(a+1/2)}$. Then, we have

$$\begin{aligned}||\alpha||_e =&||\theta||_2 + ||\beta||_{L_2} \leq ||\theta||_2 + ||\beta||_{\sup}\\ =&\sup_{||z||_2=1, t\in\mathbb{I}} |\beta(t) + \theta^\top z| = \sup_{||z||_2=1, t\in\mathbb{I}} \left\langle \tilde{\mathcal{R}}_u, \alpha \right\rangle_\lambda\\ \leq&||\alpha||_\lambda \sup_{|z|_2=1, t\in\mathbb{I}} \mid \tilde{\mathcal{R}}_u \mid_\lambda\\ \lesssim&h^{-(a+1/2)}||\alpha||_\lambda.\end{aligned}$$

Lemma B.3 *Suppose that Conditions (C1)–(C4) hold. Then for any $\alpha \in \mathcal{H}$,*

$$E(\left|\langle \mathcal{R}_\mathcal{W}, \alpha\rangle_\lambda^4\right|) \lesssim \|\alpha\|_\lambda^4.$$

Proof of Lemma B.3. By Conditions (C1) and (C4), we have

$$\begin{aligned}&E_W(\langle R_W, \alpha\rangle_\lambda)^4 = E_W\left\{\theta^\top Z + \int_\mathbb{I} X(t)\beta(t)\,dt\right\}^4\\ &\leq M_0\left\{E_W|\theta^\top Z + \int_\mathbb{I} X(t)\beta(t)\,dt|^2\right\}^2\\ &\lesssim \left\{\int_0^\tau Var[\eta_\alpha(W)|T = v, \Delta = 1]E\,dN(v)\right\}^2 \lesssim \|\alpha\|_\lambda^4.\end{aligned}$$

Lemma B.4 *Suppose that Conditions (C1)–(C3) hold. Then for any $x \in L_2([0,1])$, there exists a universal positive constant c_r such that*

$$\langle \mathcal{R}_{Z,x}, \mathcal{R}_{Z,x}\rangle_\lambda \leq c_r(\|Z\|_2^2 + \|x\|_{L_2}^2 h^{-2a-1}) \quad \textit{and} \quad E\{\|\mathcal{R}_W\|_\lambda^2\} \leq c_r h^{-1}.$$

Proof of Lemma B.4. Direct calculations yield that

$$\begin{aligned}&\langle \mathcal{R}_w, \mathcal{R}_w\rangle_\lambda = z^\top\{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1}z + \langle \pi_x, \pi_x\rangle_m - 2z^\top\{\Sigma - V(G,G)\\ &+ V(G, W_\lambda G)\}^{-1}V(G,\pi_x) + V(G,\pi_x)^\top\{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1}V(G,\pi_x). \quad \text{(B.1)}\end{aligned}$$

It follows from Condition (C3) that $\{\Sigma - V(G,G)\}^{-1}$ is positive definite and $V(G, W_\lambda G) \to 0$. Let c denote the minimum eigenvalue of $\{\Sigma - V(G,G) + V(G, W_\lambda G)\}$. Then we have $\{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1} \leq c^{-1}\mathbf{1}$ with $\mathbf{1}$ being the identity matrix. Thus, we have $z^\top\{\Sigma - V(G,G) + V(G, W_\lambda G)\}^{-1} z \lesssim \|z\|_2^2$. Direct calculations yield that $\langle \pi_X, \pi_X\rangle_m \lesssim \|X\|_{L^2}^2 h^{-2a-1}$ and $V(G, \pi_X) \lesssim \|X\|_{L^2} h^{-a-1/2}$. Thus, there exists a constant $c_r > 0$ such that

$$\langle R_W, R_W\rangle_\lambda \leq c_r(\|Z\|_2^2 + \|X\|_{L^2}^2 h^{-2a-1}).$$

Besides, it follows from (B.1) and the proof of Lemma S.4 in Shang and Cheng (2015) that

$$E\,\langle R_W, R_W\rangle_\lambda \leq c_r h^{-1}.$$

The proof is completed.

To derive the rate of convergence, we still need the following concentration inequality as a preliminary step. Define $\mathcal{F}_{p_n} = \{\alpha = (\theta^\top, \beta(\cdot)) \in \mathcal{H} : \|\theta\|_2 \leq 1, \|\beta\|_{L_2} \leq 1, J(\beta,\beta) \leq p_n\}$,

$$H_n(\alpha) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} [\phi_n(Y_i, W_i; \alpha) R_{W_i} - E\phi_n(Y_i, W_i; \alpha) R_{W_i}],$$

where $\phi_n(Y_i, W_i; \alpha)$ is a function of the observation and parameter, which may depend on n.

Lemma B.5 *Suppose that Conditions (C1)–(C4) hold. If $\phi_n(Y_i, W_i; \boldsymbol{0}) = 0$ a.s., and there exists a constant $C_\phi > 0$ such that*

$$|\phi_n(Y_i, W_i; \alpha_1) - \phi_n(Y_i, W_i; \alpha_2)| \leq C_\phi \|\alpha_1 - \alpha_2\|_e, \quad \textit{for any} \quad \alpha_1, \alpha_2 \in \mathcal{H},$$

then we have

$$\lim_n P\Big(\sup_{\alpha \in \mathcal{F}_{p_n}} \frac{\|H_n(\alpha)\|_\lambda}{p_n^{1/(4m)} \|\alpha\|_e^\gamma + n^{-1/2}} \leq \{5h^{-1} \log\log(n)\}^{1/2} \Big) = 1,$$

where $\gamma = 1 - 1/(2m)$.

Proof of Lemma B.5. Denote $N(\delta, \mathcal{F}_{p_n}, \|\cdot\|_2)$ as the $\delta-$covering number of the class $\mathcal{F}_{p_n}$, in terms of $\|\cdot\|_2$ norm. Then it follows from Theorem 9.20 of Kosorok et al. (2008) that

$$\begin{aligned}\log N(\delta, \mathcal{F}_{p_n}, \|\cdot\|_2) &\leq N(\delta, p_n^{1/2}\mathcal{F}_1, \|\cdot\|_2)\\ &\leq N(p_n^{-1/2}\delta, \mathcal{F}_1, \|\cdot\|_2)\\ &\lesssim \max\{(p_n^{-1/2}\delta)^{-1/m}, (p_n^{-1/2}\delta)^{-1/p}\},\end{aligned}$$

where p is the dimension of θ. Thus the conclusion of the lemma directly follows from

$$\exp(-c\max\{(p_n^{-1/2}\delta)^{-1/m}, (p_n^{-1/2}\delta)^{-1/p}\}) \leq \exp\{-c(p_n^{-1/2}\delta)^{-1/m}\}$$

and the proof of Lemma 3.4 in Shang and Cheng (2015).

C Proofs of Theorems

C.1 Proof of Theorem 3.1

First, we show that there exists a unique α_λ such that $\mathcal{S}_\lambda(\alpha_\lambda) = 0$. Let $r_{1n} = 2\{J(\beta_0, \beta_0) + 1\}^{1/2}h^k$, and define the operator: $T_{1h}(\alpha) = \alpha + \mathcal{S}_\lambda(\alpha_0 + \alpha), \alpha \in \mathcal{H}$. It is easy to see that

$$\|T_{1h}(\alpha)\|_\lambda = \|\alpha + \mathcal{S}_\lambda(\alpha + \alpha_0)\|_\lambda \leq \|\alpha + \mathcal{S}_\lambda(\alpha + \alpha_0) - \mathcal{S}_\lambda(\alpha_0)\|_\lambda + \|\mathcal{S}_\lambda(\alpha_0)\|_\lambda.$$

Let $\mathbb{B}(\epsilon) = \{\alpha \in \mathcal{H}, \|\alpha\|_\lambda \leq \epsilon\}$ be the ball of radius ϵ in $\mathcal{H}$. Note that $\mathcal{S}(\alpha_0) = 0$, which implies that $\mathcal{S}_\lambda(\alpha_0) = -\mathcal{P}_\lambda\alpha_0$. It follows from the Cauchy-Schwarz inequality that

$$\|\mathcal{S}_\lambda(\alpha_0)\|_\lambda = \|\mathcal{P}_\lambda\alpha_0\|_\lambda \leq \{\lambda J(\beta_0, \beta_0)\}^{1/2} \leq \{J(\beta_0, \beta_0) + 1\}^{1/2}h^k = \frac{r_{1n}}{2}. \tag{C.1}$$

By Lemma B.1, we have

$$\begin{aligned}\|\alpha+\mathcal{S}_\lambda(\alpha+\alpha_0)-\mathcal{S}_\lambda(\alpha_0)\|_\lambda =&\|\alpha+D\mathcal{S}_\lambda(\alpha_0)\alpha+\int_0^1\int_0^1 sD^2\mathcal{S}_\lambda(\alpha_0+ss'\alpha)\alpha\alpha\,ds\,ds'\|_\lambda\\ =&\|\int_0^1\int_0^1 sD^2\mathcal{S}_\lambda(\alpha_0+ss'\alpha)\alpha\alpha\,ds\,ds'\|_\lambda\\ \le&\int_0^1\int_0^1 s\|D^2\mathcal{S}_\lambda(\alpha_0+ss'\alpha)\alpha\alpha\|_\lambda\,ds\,ds'. \end{aligned}\tag{C.2}$$

From the definition of $D^2\mathcal{S}_\lambda(\alpha)$, Lemmas B.2 and B.4, and Condition (C1), we have

$$\|D^2\mathcal{S}_\lambda(\alpha_0+ss'\alpha)\alpha\alpha\|_\lambda\lesssim\{E\,\langle\mathcal{R}_W,\alpha\rangle_\lambda^4\}^{1/2}\{E\|\mathcal{R}_W\|_\lambda^2\}^{1/2}\lesssim\|\alpha\|_\lambda^2c_r^{1/2}h^{-1/2}.\tag{C.3}$$

From inequalities (C.1)–(C.3), we have

$$\|T_{1h}\|_\lambda\le c\|\alpha\|_\lambda^2c_r^{1/2}h^{-1/2}+\frac{r_{1n}}{2}.\tag{C.4}$$

Since $h=o(1)$ and $k>a+1/2\ge1/2$, we have $r_{1n}h^{-1/2}=o(1)$. Then for any $\alpha\in\mathbb{B}(r_{1n})$, $\|T_{1h}\|_\lambda<r_{1n}$ for large enough n. This implies $T_{1h}(\mathbb{B}(r_{1h}))\subset\mathbb{B}(r_{1h})$. Next, we show that T_{1h} is a contraction mapping. For any $\alpha_j=(\theta_j^\top,\beta_j(\cdot))\in\mathcal{H}, j=1,2$, we have

$$\begin{aligned}&T_{1h}(\alpha_1)-T_{1h}(\alpha_2)=\alpha_1-\alpha_2+\mathcal{S}_\lambda(\alpha_0+\alpha_1)-\mathcal{S}_\lambda(\alpha_0+\alpha_2)\\ &=\int_0^1[D\mathcal{S}_\lambda\{\alpha_0+\alpha_2+s(\alpha_1-\alpha_2)\}-D\mathcal{S}_\lambda(\alpha_0)](\alpha_1-\alpha_2)\,ds\\ &=\int_0^1\int_0^1 s'D^2\mathcal{S}_\lambda[\alpha_0+s'\{\alpha_2+s(\alpha_1-\alpha_2)\}](\alpha_1-\alpha_2)\{\alpha_2+s(\alpha_1-\alpha_2)\}\,ds\,ds'.\end{aligned}$$

By applying the similar arguments used in proving inequality (C.4), we have

$$
\begin{aligned}
&\|T_{1h}(\alpha_1) - T_{1h}(\alpha_2)\|_\lambda \\
&\leq \int_0^1 \int_0^1 s' \|D^2\mathcal{S}_\lambda[\alpha_0 + s'\{\alpha_2 + s(\alpha_1 - \alpha_2)\}](\alpha_1 - \alpha_2)\{\alpha_2 + s(\alpha_1 - \alpha_2)\}\|_\lambda ds ds' \\
&\lesssim \int_0^1 \int_0^1 s'\{E\left\langle \mathcal{R}_W, \alpha_1 - \alpha_2\right\rangle_\lambda^4\}^{1/4}\{E\|\mathcal{R}_W\|_\lambda^2\}^{1/2}\{E\left\langle \mathcal{R}_W, \alpha_2 + s(\alpha_1 - \alpha_2)\right\rangle_\lambda^4\}^{1/4}\, ds\, ds' \\
&\lesssim \|\alpha_1 - \alpha_2\|_\lambda c_r^{1/2} h^{-1/2}(\|\alpha_1 - \alpha_2\|_\lambda + \|\alpha_2\|_\lambda) \\
&\lesssim r_{1n}\|\alpha_1 - \alpha_2\|_\lambda c_r^{1/2} h^{-1/2} \\
&\leq 1/2\|\alpha_1 - \alpha_2\|_\lambda.
\end{aligned}
$$

The last inequality follows from the fact that $r_{1n}h^{-1/2} = o(1)$. Then $T_{1h}(\alpha)$ is a contraction mapping on $\mathbb{B}(r_{1n})$. By the Banach fixed-point theorem, there exists a unique $\alpha'_\lambda \in \mathbb{B}(r_{1n})$ such that $T_{1h}(\alpha'_\lambda) = \alpha'_\lambda$. Define $\alpha_\lambda = \alpha'_\lambda + \alpha_0$. Then $\mathcal{S}_\lambda(\alpha_\lambda) = 0$ and $\|\alpha_\lambda - \alpha_0\|_\lambda \leq r_{1n}$.

Second, we show that there exists a unique $\hat{\alpha}_{n,\lambda}$ such that $\mathcal{S}_{n,\lambda}(\hat{\alpha}_{n,\lambda}) = 0$. Since $\|\alpha_\lambda - \alpha_0\|_\lambda = O(r_{1n}) = o(1)$ and $D\mathcal{S}_\lambda(\alpha_0) = -id$, it follows from the Taylor expansion and inequality (C.3) that $D\mathcal{S}_\lambda(\alpha_\lambda)$ is invertible. By the similar arguments used in Shang and Cheng (2015), we can get that $\|D\mathcal{S}_\lambda(\alpha_\lambda)\|_\lambda \in (1/2, 3/2)$. Now define the operator

$$
\begin{aligned}
&T_{2h}(\alpha) = \alpha - \{D\mathcal{S}_\lambda(\alpha_\lambda)\}^{-1}\mathcal{S}_{n,\lambda}(\alpha_\lambda + \alpha) \\
&= -\{D\mathcal{S}_\lambda(\alpha_\lambda)\}^{-1}\{D\mathcal{S}_{n,\lambda}(\alpha_\lambda)\alpha - D\mathcal{S}_\lambda(\alpha_\lambda)\alpha\} \\
&\quad - \{D\mathcal{S}_\lambda(\alpha_\lambda)\}^{-1}\{\mathcal{S}_{n,\lambda}(\alpha_\lambda + \alpha) - \mathcal{S}_{n,\lambda}(\alpha_\lambda) - D\mathcal{S}_{n,\lambda}(\alpha_\lambda)\alpha\} - \{D\mathcal{S}_\lambda(\alpha_\lambda)\}^{-1}\mathcal{S}_{n,\lambda}(\alpha_\lambda) \\
&\equiv I_1 + I_2 + I_3.
\end{aligned}
$$

It follows from the functional central limit theorem that uniformly in $t \in \mathbb{I}$

$$\Big\|\frac{1}{n}\sum_{j=1}^{n}\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\} - s_1^{(0)}(t,\alpha_\lambda)\Big\|_\infty = O_p(n^{-1/2}). \tag{C.5}$$

By Lemma B.3 and the functional central limit theorem, we have

$$\begin{aligned}
&\Big\|\frac{1}{n}\sum_{j=1}^{n}\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\mathcal{R}_{W_j} - E[\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\mathcal{R}_{W_j}]\Big\|_\lambda \\
&= \sup_{\|\alpha_1\|_\lambda=1}\left\langle \frac{1}{n}\sum_{j=1}^{n}\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\mathcal{R}_{W_j} - E[\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\mathcal{R}_{W_j}], \alpha_1\right\rangle_\lambda \\
&= O_p(n^{-1/2}h^{-a-1/2}).
\end{aligned} \tag{C.6}$$

It follows from $\mathcal{S}_\lambda(\alpha_\lambda) = 0$, Lemma B.2, and equations (C.5) and (C.6) that $E\|[D\mathcal{S}_\lambda(\alpha_\lambda)]I_3\|_\lambda^2 = O((hn)^{-1})$. This implies that $\|\mathcal{S}_{n,\lambda}(\alpha_\lambda)\|_\lambda = O_p((nh)^{-1/2})$. Let c be a positive constant satisfying $P(\|\mathcal{S}_{n,\lambda}(\alpha_\lambda)\|_\lambda \le c(nh)^{-1/2}) \to 1$. Define $r_{2n} = 2c(nh)^{-1/2}$ and $\mathbb{B}(r_{2n}) = \{\alpha \in \mathcal{H} : \|\alpha\|_\lambda \le r_{2n}\}$. Then we have $P(\|\mathcal{S}_{n,\lambda}(\alpha_\lambda)\|_\lambda \le r_{2n}/2) \to 1$. Define $\Gamma = \cap_{i=1}^n A_{ni}$, where

$$A_{ni} = \{\|Z_i\|_2 \le c\log(n), \|X_i\|_{L^2} \le c\log(n), \exp\{\eta_{\alpha_\lambda}(W_i)\} \le c\log(n)\}$$

for a constant c. From Condition (C4), we can choose c large enough such that $P(\Gamma) \to 1$ and $P(A_{ni}^c) = O(n^{-1})$. To handle I_1, we have

$$\begin{aligned}
\|[D\mathcal{S}_\lambda(\alpha_\lambda)]I_1\|_\lambda \le\ & \Big\|\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{n^{-1}\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\mathcal{R}_{W_j}}{S_0^{(1)}(Y_i,\alpha_\lambda)} \\
&- \int_0^\tau s_1^{(0)}(t,\alpha_0)\frac{E[\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\mathcal{R}_{W_j}]}{s_1^{(0)}(t,\alpha_\lambda)}h_0(t)\,dt\Big\|_\lambda \\
&+ \Big\|\frac{1}{n}\sum_{i=1}^{n}\Delta_i\frac{\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\sum_{j=1}^{n}\mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_\lambda}(W_j)\}\mathcal{R}_{W_j}}{[nS_0^{(1)}(Y_i,\alpha_\lambda)]^2} \\
&- \int_0^\tau s_1^{(0)}(t,\alpha_0)\frac{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_\lambda}(W)\}\eta_\alpha(W)]E[\mathcal{Y}(t)\exp\{\eta_{\alpha_\lambda}(W)\}\mathcal{R}_W]}{[s_1^{(0)}(t,\alpha_\lambda)]^2}h_0(t)\,dt\Big\|_\lambda \\
\equiv\ & I_{11} + I_{12}.
\end{aligned} \tag{C.7}$$

For I_{11}, we have

$$
\begin{aligned}
\|I_{11}\|_\lambda \le \Big\| \frac{1}{n}\sum_{i=1}^n \int_0^\tau & \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\mathcal{R}_{W_j}}{s_1^{(0)}(t,\alpha_\lambda)} \\
& - \frac{E[\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\mathcal{R}_{W_j}]}{s_1^{(0)}(t,\alpha_\lambda)}\, dN_i(t)\Big\|_\lambda \\
& + \frac{1}{n}\Big\| \sum_{i=1}^n \int_0^\tau n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\mathcal{R}_{W_j}\{\frac{1}{s_1^{(0)}(t,\alpha_\lambda)} - \frac{1}{S_1^{(0)}(t,\alpha_\lambda)}\}\, dN_i(t)\Big\|_\lambda \\
& + \Big\|\frac{1}{n}\sum_{i=1}^n \int_0^\tau \frac{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\mathcal{R}_{W_j}}{s_1^{(0)}(t,\alpha_\lambda)}\, dN_i(t) \\
& - \int_0^\tau s_1^{(0)}(t,\alpha_0)\frac{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\mathcal{R}_{W_j}}{s_1^{(0)}(t,\alpha_\lambda)} h_0(t)\, dt\Big\|_\lambda \\
\equiv I_{111} + I_{112} + I_{113}. &
\end{aligned}
$$

For I_{113}, we have

$$
\begin{aligned}
I_{113} &= \Big\| \int_0^\tau \frac{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\mathcal{R}_{W_j}}{s_1^{(0)}(t,\alpha_\lambda)} \frac{1}{n}\sum_{i=1}^n \{ dN_i(t) - E\, dN_i(t)\} \Big\|_\lambda \\
&= O_p((nh)^{-1/2})\|\alpha\|_\lambda.
\end{aligned}
$$

To infer I_{111}, define

$$
\phi(Y_j, W_j;\alpha) = \frac{\mathcal{Y}_j(t_0)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)}{s_1^{(0)}(t_0,\alpha_\lambda)} I_{A_{nj}}.
$$

Then for any $\alpha_1, \alpha_2 \in \mathcal{H}$, we have

$$
\begin{aligned}
|\phi(Y_j, W_j;\alpha_1) - \phi(Y_j, W_j;\alpha_2)| &= \frac{1}{s_1^{(0)}(t_0,\alpha_\lambda)} \mathcal{Y}_j(t_0)\exp\{\eta_{\alpha_\lambda}(W_j)\}|\{\eta_{\alpha_1}(W_j) - \eta_{\alpha_2}(W_j)\}| I_{A_{nj}} \\
&\le \frac{c\log(n)}{s_1^{(0)}(t_0,\alpha_\lambda)} |\left\langle \mathcal{R}_{W_j}, \alpha_1 - \alpha_2 \right\rangle_\lambda | I_{A_{nj}} \\
&\le \frac{\{c\log(n)\}^2}{s_1^{(0)}(t_0,\alpha_\lambda)} \|\alpha_1 - \alpha_2\|_e.
\end{aligned}
$$

Define $\phi_n(Y_j, W_j; \alpha) = s_1^{(0)}(t_0, \alpha_\lambda)c^{-2}\{\log(n)\}^{-2}\phi(Y_j, W_j; \alpha_1)$. Then

$$|\phi_n(Y_j, W_j; \alpha_1) - \phi_n(Y_j, W_j; \alpha_1)| \leq \|\alpha_1 - \alpha_2\|_e.$$

For any $\alpha \neq 0 \in \mathcal{H}$, let $\tilde{\alpha} = \alpha/(d_n\|\alpha\|_\lambda)$, where $d_n = \kappa h^{-(2a+1)/2}$. It follows from Lemma B.3 that $\|\tilde{\alpha}\|_e \leq d_n\|\tilde{\alpha}\|_\lambda = 1$. Then we have $\|\tilde{\theta}\|_2 + \|\tilde{\beta}\|_{L_2} \leq 1$. Meanwhile, we have $\lambda J(\tilde{\beta}, \tilde{\beta}) \leq \|\tilde{\alpha}\|_\lambda^2 = d_n^{-2}$. Then $J(\tilde{\beta}, \tilde{\beta}) \leq \lambda^{-1}d_n^{-2} \equiv p_n$. By Lemma B.5, we obtain that for any $\alpha \in \mathbb{B}(r_{2n})$,

$$\lim_n P\Big(\Big\|\sum_{j=1}^n [\phi_n(Y_j, W_j; \tilde{\alpha})\mathcal{R}_{W_j} - E\{\phi_n(Y_j, W_j; \tilde{\alpha})\mathcal{R}_{W_j}\}]\Big\|_\lambda \lesssim (n^{1/2}p_n^{1/(4m)} + 1)\{h^{-1}\log\log(n)\}^{1/2}\Big) = 1.$$

Therefore, we have

$$\lim_n P\Big(\Big\|\sum_{j=1}^n [\phi(Y_j, W_j; \alpha)\mathcal{R}_{W_j} - E\{\phi(Y_j, W_j; \alpha)\mathcal{R}_{W_j}\}]\Big\|_\lambda \lesssim d_n\{\log(n)\}^2\|\alpha\|_\lambda(n^{1/2}p_n^{1/(4m)} + 1)\{h^{-1}\log\log(n)\}^{1/2}\Big) = 1.$$

It follows from the definition of A_{ni} that

$$\Big\|\frac{E[\mathcal{Y}_j(t_0)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\mathcal{R}_{W_j}I_{A_{nj}^c}]}{s_1^{(0)}(t, \alpha_\lambda)}\Big\|_\lambda \leq E\|\left\langle \mathcal{R}_{W_j}, \alpha\right\rangle_\lambda \mathcal{R}_{W_j}I_{A_{nj}^c}\Big\|_\lambda = O(P(A_{ni}^c)^{1/2}h^{-1/2})\|\alpha\|_\lambda = o(1)\|\alpha\|_\lambda.$$

Thus, we have

$$I_{111} = O_p(n^{-1/2}h^{-(a+1)-\frac{2k-2a-1}{4m}}\{\log(n)\}^2\{\log\log(n)\}^{1/2})\|\alpha\|_\lambda + o_p(1)\|\alpha\|_\lambda = o_p(1)\|\alpha\|_\lambda.$$

From Lemma B.4, we can get that

$$\frac{1}{n}\Big\|\sum_{i=1}^n\int_0^\tau E\big[\frac{1}{n}\sum_{j=1}^n\mathcal{Y}_j(t)\exp\{\eta_{\alpha_\lambda}(W_j)\}\eta_\alpha(W_j)\mathcal{R}_{W_j}\big]\,dN_i(t)\Big\|_\lambda = O_p(h^{-1/2}\|\alpha\|_\lambda).$$

Then by equation (C.5) and $(nh)^{-1} = o(1)$, we have $I_{112} = o_p(1)\|\alpha\|_\lambda$.

Similar to I_{11}, we can get that $I_{12} = o_p(1)\|\alpha\|_\lambda$. Therefore, for any $\alpha \in \mathbb{B}(r_{2n})$, $\|[D\mathcal{S}_\lambda(\alpha_\lambda)]I_1\| \leq r_{2n}/18$. For $\|[D\mathcal{S}_\lambda(\alpha_\lambda)]I_2\|_\lambda$, we have

$$\begin{aligned}
\|[D\mathcal{S}_\lambda(\alpha_\lambda)]I_2\|_\lambda &= \|\{\mathcal{S}_{n,\lambda}(\alpha_\lambda + \alpha) - \mathcal{S}_{n,\lambda}(\alpha_\lambda) - D\mathcal{S}_{n,\lambda}(\alpha_\lambda)\alpha\}\|_\lambda \\
&= \|\int_0^1 \int_0^1 sD^2\mathcal{S}_{n,\lambda}(\alpha_\lambda + ss'\alpha)\alpha\alpha \, ds \, ds'\|_\lambda.
\end{aligned}$$

It follows from inequality (C.3) that

$$\begin{aligned}
&\|D^2\mathcal{S}_{n,\lambda}(\alpha_\lambda + ss'\alpha)\alpha\alpha\|_\lambda \\
&\leq \|D^2\mathcal{S}_{n,\lambda}(\alpha_\lambda + ss'\alpha)\alpha\alpha - D^2\mathcal{S}_\lambda(\alpha_\lambda + ss'\alpha)\alpha\alpha\|_\lambda + \|D^2\mathcal{S}_\lambda(\alpha_\lambda + ss'\alpha)\alpha\alpha\|_\lambda \\
&= \|D^2\mathcal{S}_{n,\lambda}(\alpha_\lambda + ss'\alpha)\alpha\alpha - D^2\mathcal{S}_\lambda(\alpha_\lambda + ss'\alpha)\alpha\alpha\|_\lambda + O(h^{-1/2})\|\alpha\|_\lambda^2.
\end{aligned}$$

By using the similar arguments applied to I_{111}, we can get that

$$\begin{aligned}
&\|D^2\mathcal{S}_{n,\lambda}(\alpha_\lambda + ss'\alpha)\alpha\alpha - D^2\mathcal{S}_\lambda(\alpha_\lambda + ss'\alpha)\alpha\alpha\|_\lambda \\
&= O_p\Big(n^{-1}h^{-(2a+1)-\frac{2k-2a-1}{4m}}\log(n)^3\{\log\log(n)\}^{1/2}\{1+n^{-1/2}\}\Big)\|\alpha\|_\lambda \\
&\quad + O_p\Big(\frac{1}{h^{1/2}} + \frac{1}{\sqrt{n}h^{(a+1)+\frac{2k-2a-1}{4m}}}\{\log(n)\}^2\{\log\log(n)\}^{1/2}h^{-1/2}(1+\{nh^2\}^{-1/2}) \\
&\quad + n^{-1/2}\log(n)h^{-(2a+1)/2} + n^{-1/2}h^{-(a+1)-\frac{2k-2a-1}{4m}}\{\log(n)\}^2\{\log\log(n)\}^{1/2} + n^{-1/2}h^{-1}\Big)\|\alpha\|_\lambda^2.
\end{aligned}$$

It follows from $\alpha \in \mathbb{B}(r_{2n})$ and the conditions in the theorem that

$$\|D^2\mathcal{S}_{n,\lambda}(\alpha_\lambda + ss'\alpha)\alpha\alpha - D^2\mathcal{S}_\lambda(\alpha_\lambda + ss'\alpha)\alpha\alpha\|_\lambda = o_p(1)\|\alpha\|_\lambda.$$

Then we have $\|[D\mathcal{S}_\lambda(\alpha_\lambda)]I_2\|_\lambda \leq 11\|\alpha\|_\lambda/18$. Therefore, for any $\alpha \in \mathbb{B}(r_{2n})$,

$$\|T_{2h}(\alpha)\|_\lambda \leq \|I_1\|_\lambda + \|I_2\|_\lambda + \|I_3\|_\lambda \leq 11r_{2n}/12.$$

That is, $T_{2h}(\mathbb{B}(r_{2n})) \subset \mathbb{B}(r_{2n})$. Using the same arguments as the above, we can get that T_{2h} is a contraction mapping in $\mathbb{B}(r_{2n})$. Therefore, there exists a unique $\alpha' \in \mathbb{B}(r_{2n})$ such that $T_{2h}(\alpha') = \alpha'$, implying $\mathcal{S}_{n,\lambda}(\alpha_\lambda + \alpha') = 0$. Let $\hat{\alpha}_{n,\lambda} = \alpha_\lambda + \alpha'$. Then $\mathcal{S}_{n,\lambda}(\hat{\alpha}_{n,\lambda}) = 0$. Therefore, with probability approaching to 1, we have

$$\|\hat{\alpha}_{n,\lambda} - \alpha_0\|_\lambda \leq r_{1n} + r_{2n} = O_P((nh)^{-1/2} + h^k).$$

C.2 Proof of Theorem 3.2

It follows from Theorem 3.1 that there exists a constant $M > 0$ such that, with probability approaching to one, $\|\hat{\alpha}_{n,\lambda} - \alpha_0\|_\lambda \leq Mr_n$. For simplicity, denote $\alpha = \hat{\alpha}_{n,\lambda} - \alpha_0$. We can assume that $\|\alpha\|_\lambda \leq Mr_n$ since its complement is negligible in terms of probability. Let $d_n = \kappa M h^{-(2a+1)/2} r_n$, $\tilde{\alpha} = d_n^{-1}\alpha$, and $p_n = \kappa^{-2} h^{1-2k}$, where κ is a constant given in Lemma B.3. Clearly, $p_n \geq 1$ when n is large enough since $h \to 0$ with $n \to \infty$ and $1 - 2k < 0$. It can be shown that $\|\alpha\|_\lambda \leq Mr_n$ implies $\tilde{\alpha} \in \mathcal{F}_{p_n}$. To see this, write $\tilde{\alpha} = (\tilde{\theta}^\top, \tilde{\beta}(\cdot))$. Then $\|\tilde{\alpha}\|_e = d_n^{-1}\|\alpha\|_e \leq d_n^{-1}\kappa h^{-(2a+1)/2}\|\alpha\|_\lambda \leq d_n^{-1}\kappa h^{-(2a+1)/2} M r_n = 1$. Thus, we can get

$$J(\tilde{\beta}, \tilde{\beta}) = d_n^{-2}\lambda^{-1}\{\lambda J(\beta, \beta)\} \leq d_n^{-2}\lambda^{-1}\|\alpha\|_\lambda^2 \leq d_n^{-2}\lambda^{-1}(Mr_n)^2 = \kappa^{-2}h^{1-2k} = p_n.$$

Besides, we have

$$\begin{aligned}
&\|\mathcal{S}_{n,\lambda}(\alpha + \alpha_0) - \mathcal{S}_{n,\lambda}(\alpha_0) - \{\mathcal{S}_\lambda(\alpha + \alpha_0) - \mathcal{S}_\lambda(\alpha_0)\}\|_\lambda \\
=&\|\mathcal{S}_n(\alpha + \alpha_0) - \mathcal{S}_n(\alpha_0) - \{\mathcal{S}(\alpha + \alpha_0) - \mathcal{S}(\alpha_0)\}\|_\lambda. \qquad \text{(C.8)}
\end{aligned}$$

For the right hand side of equation (C.8), we have

$$
\begin{aligned}
&\|\mathcal{S}_n(\alpha+\alpha_0)-\mathcal{S}_n(\alpha_0)-\{\mathcal{S}(\alpha+\alpha_0)-\mathcal{S}(\alpha_0)\}\|_\lambda \\
&=\Big\|\frac{1}{n}\sum_{i=1}^n \Delta_i\left[\mathcal{R}_{W_i}-\frac{n^{-1}\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha+\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{S_1^{(0)}(Y_i,\alpha+\alpha_0)}\right] \\
&\quad -\frac{1}{n}\sum_{i=1}^n \Delta_i\left[\mathcal{R}_{W_i}-\frac{n^{-1}\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{S_1^{(0)}(Y_i,\alpha_0)}\right]-\{\mathcal{S}(\alpha+\alpha_0)-\mathcal{S}(\alpha_0)\}\Big\|_\lambda \\
&\leq \Big\|\frac{1}{n}\sum_{i=1}^n \int_0^\tau \left[\left\{\frac{\sum_j \mathcal{Y}_j(t)\exp\{\eta_{\alpha+\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{nS_1^{(0)}(t,\alpha_0+\alpha)}-\frac{\sum_j \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{nS_1^{(0)}(t,\alpha_0)}\right\}\right. \\
&\qquad\qquad \left.+\left\{\frac{E\mathcal{Y}(t)\exp\{\eta_{\alpha+\alpha_0}(W)\}\mathcal{R}_W}{s_1^{(0)}(t,\alpha_0+\alpha)}-\frac{E\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}\mathcal{R}_W}{s_1^{(0)}(t,\alpha_0)}\right\}\right] dN_i(t)\Big\|_\lambda+(nh)^{-1/2} \\
&\leq \Big\|\frac{1}{n}\sum_{i=1}^n \int_0^\tau \left[\left\{\frac{\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha+\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{ns_1^{(0)}(t,\alpha_0+\alpha)}-\frac{\sum_j \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{ns_1^{(0)}(t,\alpha_0)}\right\}\right. \\
&\qquad\qquad\quad \left.+\left\{\frac{E\mathcal{Y}(t)\exp\{\eta_{\alpha+\alpha_0}(W)\}\mathcal{R}_W}{s_1^{(0)}(t,\alpha_0+\alpha)}-\frac{E\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}\mathcal{R}_W}{s_1^{(0)}(t,\alpha_0)}\right\}\right] dN_i(t)\Big\|_\lambda \\
&\quad +O_p\left(\frac{1}{(nh)^{1/2}}+\frac{1}{h^{1/2+a}n^{1/2}}\right).
\end{aligned}
$$

Define $\Gamma=\cap_{i=1}^n A_{ni}$, where

$$A_{ni}=\{\|Z_i\|_2\leq c\log(n),\|X_i\|_{L^2}\leq c\log(n),\exp\{\eta_{\alpha_\lambda}(W_i)\}\leq c\log(n)\}.$$

For any t_0, define $\varphi(Y_j;\alpha)=[\mathcal{Y}_j(t_0)\exp\{\eta_{\alpha+\alpha_0}(W_j)\}-\mathcal{Y}_j(t_0)\exp\{\eta_{\alpha_0}(W_j)\}]$, $D_n=\{c\log(n)\}^2 d_n^{-1}$, and $\varphi_n(Y_j;\tilde{\alpha})=D_n\varphi(Y_j;d_n\tilde{\alpha})\mathbf{1}_{A_{ni}}$. Then $|\varphi_n(Y_j;\tilde{\alpha}_1)-\varphi_n(Y_j;\tilde{\alpha}_2)|\leq\|\tilde{\alpha}_1-\tilde{\alpha}_2\|_e$. Since $\|\alpha\|_\lambda\leq Mr_n$, $\tilde{\alpha}\in\mathcal{F}_{p_n}$, it follows from Lemma B.5 that with probability approaching to one,

$$
\begin{aligned}
&n^{-1/2}\|\sum_{j=1}^n \varphi_n(Y_j;\tilde{\alpha})\mathcal{R}_{W_j}-E\varphi_n(Y_j;\tilde{\alpha})\mathcal{R}_{W_j}\|_\lambda \lesssim (p_n^{1/(4m)}\|\tilde{\alpha}\|_e^\gamma+n^{-1/2})\{h^{-1}\log\log(n)\}^{1/2} \\
&\lesssim (p_n^{1/(4m)}+n^{-1/2})\{h^{-1}\log\log(n)\}^{1/2},
\end{aligned}
$$

where $\gamma = 1 - 1/(2m)$. On the other hand, by the Taylor expansion, the Cauchy-Schwarz inequality, Lemma B.2 and Theorem 3.1, we have

$$\begin{aligned}
&\|E\{\varphi(Y_j; d_n\tilde{\alpha})\mathcal{R}_{W_j} I^c_{A_{n_j}}\}\|_\lambda \leq \{E|\varphi(Y_j; d_n\tilde{\alpha}) I^c_{A_{nj}}|^2\}^{1/2}\{E\|\mathcal{R}_{W_j}\|^2_\lambda\}^{1/2} \\
\lesssim &\big(E[\mathcal{Y}_j(t_0)\exp\{\eta_{\alpha_0}(W_j)\}\left\langle \mathcal{R}_{W_j}, \tilde{\alpha}\right\rangle_\lambda]^2\big)^{1/2} c_r^{1/2} h^{-1/2} \\
\lesssim &\big(E[\mathcal{Y}_j(t_0)\exp\{\eta_{\alpha_0}(W_j)\}]^4\big)^{1/4} P(A^c_{n_j})^{1/4}[E\{(\|Z\|_2 + \|X_j\|_{L_2})\}^4]^{1/4} c_r h^{-1/2}.
\end{aligned}$$

From Condition (C4), we can choose c large enough such that

$$n^{1/2} h^{-1/2} P(A^c_{n_i})^{1/4} = o(p_n^{1/(4m)}\{h^{-1}\log\log(n)\}^{1/2}).$$

Then

$$n^{1/2}\|E\{\varphi(Y_j; d_n\tilde{\alpha})\mathcal{R}_{W_j} I^c_{A_{n_j}}\}\|_\lambda \lesssim p_n^{1/(4m)}\{h^{-1}\log\log(n)\}^{1/2}.$$

Thus, on Γ_n, as $n \to \infty$, we have

$$n^{-1/2} D_n \|\mathcal{S}_n(\alpha + \alpha_0) - \mathcal{S}_n(\alpha_0) - \{\mathcal{S}(\alpha + \alpha_0) - \mathcal{S}(\alpha_0)\}\|_\lambda \lesssim p_n^{1/(4m)}\{h^{-1}\log\log(n)\}^{1/2}. \tag{C.9}$$

For the left hand side of equation (C.8), we have

$$\begin{aligned}
&\|\mathcal{S}_{n,\lambda}(\alpha + \alpha_0) - \mathcal{S}_{n,\lambda}(\alpha_0) - \{\mathcal{S}_\lambda(\alpha + \alpha_0) - \mathcal{S}_\lambda(\alpha_0)\}\|_\lambda \\
=&\| - \mathcal{S}_{n,\lambda}(\alpha_0) - D\mathcal{S}_\lambda(\alpha_0)\alpha - \int_0^1\int_0^1 sD^2\mathcal{S}_\lambda(\alpha_0 + ss'\alpha)\alpha\alpha\, ds\, ds'\|_\lambda \\
=&\Big\|\alpha - \mathcal{S}_{n,\lambda}(\alpha_0) - \int_0^1\int_0^1 sD^2\mathcal{S}_\lambda(\alpha_0 + ss'\alpha)\alpha\alpha\, ds\, ds'\Big\|_\lambda \\
\geq&\Big\|\alpha - \mathcal{S}_{n,\lambda}(\alpha_0)\|_\lambda - \|\int_0^1\int_0^1 sD^2\mathcal{S}_\lambda(\alpha_0 + ss'\alpha)\alpha\alpha\, ds\, ds'\Big\|_\lambda.
\end{aligned} \tag{C.10}$$

It follows from inequality (B.3) that

$$\Big\| \int_0^1 \int_0^1 sD^2\mathcal{S}_\lambda(\alpha_0 + ss'\alpha)\alpha\alpha\, ds\, ds' \Big\|_\lambda \le \int_0^1 \int_0^1 s\|D^2\mathcal{S}_\lambda(\alpha_0 + ss'\alpha)\alpha\alpha\|_\lambda\, ds\, ds'$$
$$\lesssim \|\alpha\|_\lambda^2 c_r^{1/2} h^{-1/2} \lesssim h^{-1/2} r_n^2. \tag{C.11}$$

Therefore, it follows from (C.8)–(C.11) that

$$\|\alpha - \mathcal{S}_{n,\lambda}(\alpha_0)\|_\lambda \le O_p(a_n).$$

The proof of Theorem 3.2 is completed.

C.3 Proof of Theorem 3.3

Define $\hat{\alpha}_{n,\lambda}^h = (\hat{\theta}_{n,\lambda}, h^{a+1/2}\hat{\beta}_{n,\lambda})$, $\alpha_0^* = (id - \mathcal{P}_\lambda)\alpha_0$, $\alpha_0^{*h} = (\theta_0^*, h^{a+1/2}\beta_0^*)$, $\tilde{\mathcal{R}}_u^h = (\tilde{H}_u, h^{a+1/2}\tilde{T}_u)$, and

$$Rem_n = \hat{\alpha}_{n,\lambda} - \alpha_0^* - \frac{1}{n}\sum_{i=1}^n \int_0^\tau \left[\mathcal{R}_{W_i} - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}} \right] dN_i(t).$$

It follows from Theorem 3.2 that $\|Rem_n\|_\lambda = O_p(a_n)$. Thus, we have

$$\|\hat{\theta}_{n,\lambda} - \theta_0^* - \frac{1}{n}\sum_{i=1}^n \int_0^\tau \left[\mathcal{H}_{W_i} - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{H}_{W_j}}{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}} \right] dN_i(t)\|_2 = O_p(a_n).$$

Define

$$Rem_n^h = \hat{\alpha}_{n,\lambda}^h - \alpha_0^{*h} - \frac{1}{n}\sum_{i=1}^n \int_0^\tau \left[\mathcal{R}_{W_i}^h - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}^h}{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}} \right] dN_i(t).$$

Then it is easy to show that $\|Rem_n^h - h^{a+1/2}Rem_n\|_\lambda = O_p(a_n)$. It follows from $a_n = o(n^{-1/2})$ that

$$\|Rem_n^h\|_\lambda \le \|Rem_n^h - h^{a+1/2}Rem_n\|_\lambda + h^{a+1/2}\|Rem_n\|_\lambda = o_p(n^{-1/2}).$$

Next, we will use Rem_n^h to obtain the joint limiting distribution. The idea is to employ the Cramér-Wold device. For any $u = (z^\top, t) \in \mathbb{R}^p \times \mathbb{I}$, we will obtain the limiting distribution of $n^{1/2} z^\top(\hat{\theta}_{n,\lambda} - \theta_0^*) + n^{1/2} h^{a+1/2}\{\hat{\beta}_{n,\lambda}(t) - \beta_0^*(t)\}$. Note that this is equivalent to getting the asymptotic results about $n^{1/2}\left\langle \tilde{\mathcal{R}}_u, \hat{\alpha}_{n,\lambda}^h - \alpha_0^{*h}\right\rangle_\lambda$. It follows from Theorem 3.2 that $n^{1/2}|\left\langle \tilde{\mathcal{R}}_u, Rem_n^h\right\rangle_\lambda| = O_p(n^{1/2}h^{-(a+1/2)}a_n)$. Thus, we only need to get the limiting distribution of

$$n^{1/2}\left\langle \tilde{\mathcal{R}}_u, \frac{1}{n}\sum_{i=1}^n \int_0^\tau \left[\mathcal{R}_{W_i}^h - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}^h}{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}}\right] dN_i(t)\right\rangle_\lambda .$$

Direct calculations yield that

$$\begin{aligned}
& n^{1/2}\left\langle \tilde{\mathcal{R}}_u, \frac{1}{n}\sum_{i=1}^n \int_0^\tau \left[\mathcal{R}_{W_i}^h - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}^h}{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}}\right] dN_i(s)\right\rangle_\lambda \\
=& n^{-1/2}\sum_{i=1}^n \int_0^\tau \Big[z^\top H_{W_i} + h^{a+1/2}T_{W_i}(t) \\
&\qquad\qquad - \frac{\sum_{j=1}^n \mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}\{z^\top H_{W_j} + h^{a+1/2}T_{W_j}(t)\}}{\sum_{j=1}^n \mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}}\Big] dM_i(s) \\
\equiv& \mathcal{U}_n.
\end{aligned}$$

Define $\mathcal{K}_i(u) \equiv z^\top H_{W_i} + h^{a+1/2}T_{W_i}(t)$. Therefore, we have

$$\begin{aligned}
\mathcal{U}_n =& n^{-1/2}\sum_{i=1}^n \int_0^\tau \{\mathcal{K}_i(u) - \frac{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}[\mathcal{K}_j(u)]}{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}}\} dM_i(s) + O_p(n^{-1/2}h^{-a-1/2}) \\
\equiv& n^{-1/2}\sum_{i=1}^n \mathcal{U}_i + o_p(1).
\end{aligned}$$

Direct calculations yield that

$$
\begin{aligned}
Var(\mathcal{U}_i) &= E\int_0^\tau \left[\mathcal{K}_i(u) - \frac{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{K}_j(u)}{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}}\right]^2 dN_i(s) \\
&= h^{2a+1}E\int_0^\tau \left[\pi_{X_i}(t) - \frac{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}\pi_{X_j}(t)}{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}}\right]^2 dN_i(s) \\
&\quad +2h^{a+1/2}(z - h^{a+1/2}\omega(t))^\top\{\Sigma - V(G,G) + V(G,W_\lambda G)\}^{-1} \\
&\qquad \times E\int_0^\tau \left[\pi_{X_i}(t) - \frac{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}\pi_{X_j}(t)}{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}}\right] \\
&\qquad \times \left[\{Z_i - V(G,\pi_{X_i})\} - \frac{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}\{Z_j - V(G,\pi_{X_j})\}}{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}}\right] dN_i(s) \\
&\quad +(z - h^{a+1/2}\omega(t))^\top\{\Sigma - V(G,G) + V(G,W_\lambda G)\}^{-1} \\
&\qquad \times E\int_0^\tau \left[\{Z_i - V(G,\pi_{X_i})\} - \frac{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}\{Z_j - V(G,\pi_{X_j})\}}{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}}\right]^{\otimes 2} dN_i(s) \\
&\qquad \times\{\Sigma - V(G,G) + V(G,W_\lambda G)\}^{-1}(z - h^{a+1/2}\omega(t)),
\end{aligned}
$$

and

$$
\begin{aligned}
& E\int_0^\tau \left[\{Z_i - V(G,\pi_{X_i})\} - \frac{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}\{Z_j - V(G,\pi_{X_j})\}}{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}}\right] \\
& \qquad \times \left\{\pi_{X_i}(t) - \frac{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}\pi_{X_j}(t)}{E\mathcal{Y}_j(s)\exp\{\eta_{\alpha_0}(W_j)\}}\right\} dN_i(s) \\
&= \sum_{j=1}^\infty \left\{\frac{G_j}{1+\lambda\rho_j}h_j(t) - \frac{G_j}{(1+\lambda\rho_j)^2}h_j(t)\right\} \\
&= \sum_{j=1}^\infty \frac{\lambda\rho_j G_j}{(1+\lambda\rho_j)^2}h_j(t) = W_\lambda\omega(t).
\end{aligned}
$$

Using the arguments similar to the proof of Theorem 3.1 in Cheng and Shang (2015), we have $h^{a+1/2} \to 0$, $\omega(t) = O(1)$, $h^{a+1/2}W_\lambda\omega(t) \to 0$, $\sqrt{n}\{\theta_0^* - \theta_0\} \to 0$, and $\sqrt{n}h^{a+1/2}\{\beta_0^*(t) -$

$\beta_0(t) + \{W_\lambda(\beta_0)\}(t)\} \to 0$. Then, as $\lambda \to 0$, we have

$$
\begin{aligned}
Var(\mathcal{U}_i) \to & \sigma_t^2 + 2(z+\gamma_0)^\top \{\Sigma - V(G,G)\}^{-1}\xi_0 + (z+\gamma_0)^\top \{\Sigma - V(G,G)\}^{-1}(z+\gamma_0) \\
\equiv & (z^\top, 1)\Phi(z^\top, 1)^\top.
\end{aligned}
$$

It follows from the Lindeberg's central limit theorem that

$$
\begin{bmatrix} \sqrt{n}(\hat{\theta}_{n,\lambda} - \theta_0) \\ \sqrt{nh}h^a\{\hat{\beta}_{n,\lambda}(t) - \beta_0(t) + (W_\lambda\beta_0)(t_0)\} \end{bmatrix} \to N(0, \Phi).
$$

Since $n^{1/2}h^{k(1+b)} = o(1)$, we can get that $nh^{4k} = o(1)$. Then, we have

$$
\begin{aligned}
|(W_\lambda\beta_0)(t_0)| &= \Big| \sum_{j=1}^n \frac{b_j\lambda\rho_j}{1+\lambda\rho_j} h_j(t_0) \Big| \le c_h\lambda \left\{ \sum_{j=1}^\infty b_j^2\rho_j^2 \right\}^{1/2} \left\{ \sum_{j=1} \frac{j^{2a}}{(1+\lambda\rho_j)^2} \right\}^{1/2} \\
&= O(\lambda h^{-a-1/2}) = o(1).
\end{aligned}
$$

Hence, it leads to $\sqrt{n}h^{a+1/2}\{W_\lambda(\beta_0)\}(t) = o(1)$. Thus, the conclusion follows directly.

C.4 Proof of Theorem 3.4

By Theorem 3.2 and the proof of Theorem 3.3, we have

$$
n^{1/2}h^{a+1/2} \sup_{s\in\mathbb{I}} |\hat{\beta}_{n,\lambda}(s) - \beta_0(s) - \tilde{\mathcal{S}}_n(\alpha_0)(s)| = o_p(1),
$$

where

$$
\begin{aligned}
\tilde{\mathcal{S}}_n(\alpha_0)(s) \equiv & \frac{1}{n}\int_0^\tau \sum_{i=1}^n \Big(T_{W_i}(s) - \frac{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}T_W(s)]}{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}]}\Big)\, dM_i(t) \\
= & \frac{1}{n}\int_{\mathbb{I}} K(s,u)\int_0^\tau \sum_{i=1}^n \Big(X_i(u) - \frac{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}X(u)]}{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}]}\Big)\, dM_i(t)\, du \\
& - \omega(s)\{\Sigma - \Omega + V(G, W_\lambda G^\top)\}^{-1}\frac{1}{n}\sum_{i=1}^n \int_0^\tau \{Z_i - V(G, \pi_{X_i})\}\, dM_i(t) \\
\equiv & \tilde{\mathcal{S}}_{n1}(\alpha_0)(s) - \tilde{\mathcal{S}}_{n2}(\alpha_0)(s).
\end{aligned}
$$

We will show the theorem through the following two steps.

(i) Denote $H_n(s) = \sqrt{nh}h^a\tilde{\mathcal{S}}_{n1}(\alpha_0)(s)$. The first step is to show that $H_n(s)$ converges to the Gaussian process $\mathcal{G}(t)$ in the Hilbert Space $\mathcal{H}^{(m)}$ with the inner product $V(\cdot,\cdot)$, where $h_j, j = 1, 2, \ldots,$ are the orthonormal bases. Direct calculations yield

$$
\begin{aligned}
H_n(s) = & \frac{h^{1/2+a}}{\sqrt{n}}\int_{\mathbb{I}} K(s,u)\int_0^\tau \sum_{i=1}^n \Big(X_i(u) - \frac{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}X(u)]}{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}]}\Big)\, dM_i(t)\, du \\
= & \frac{1}{\sqrt{n}}\sum_{j=1}^\infty \int_{\mathbb{I}} \frac{h_j(u)h_j(s)h^{1/2+a}}{1+\lambda\rho_j}\int_0^\tau \sum_{i=1}^n \Big(X_i(u) - \frac{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}X(u)]}{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}]}\Big)\, dM_i(t)\, du.
\end{aligned}
$$

From Theorem 1.8.4 in van der Vaart and Wellner (1996), to prove that $H_n(t)$ converges to the Gaussian process $\mathcal{G}(t)$ in the Hilbert Space $\mathcal{H}^m$, we only need to prove that $H_n(\cdot)$ is asymptotically finite-dimensional and $V(H_n, h_j)$ converges in distribution to $V(\mathcal{G}, h_j)$. It follows from the definition of $H_n(\cdot)$ that

$$
\begin{aligned}
\sum_{j=1}^\infty V(H_n, h_j)^2 = & \sum_{j=1}^\infty \frac{h^{1+2a}}{(1+\lambda\rho_j)^2}\Bigg\{\frac{1}{\sqrt{n}}\sum_{i=1}^n \int_{\mathbb{I}} h_j(u)\int_0^\tau \Big(X_i(u) \\
& - \frac{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}X(u)]}{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}]}\Big)\, dM_i(t)\, du\Bigg\}^2.
\end{aligned}
$$

It is easy to verify that

$$\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_{\mathbb{I}} h_j(u)\int_0^{\tau}\left(X_i(u)-\frac{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}X(u)]}{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}]}\right)dM_i(t)\,du$$

is asymptotically tight and bounded by cj^a. Besides, we have $\sum_j \frac{h^{1+2a}j^{2a}}{(1+\lambda\rho_j)^2} \asymp \int_0^\infty \frac{x^{2a}}{(1+x^{2k})^2}\,dx < \infty$. Then for every $\varepsilon > 0$, there exists J_0 such that $\sum_{j\geq J_0} \frac{h^{1+2a}j^{2a}}{(1+\lambda\rho_j)^2} < \varepsilon$. Thus for any $\varepsilon > 0$ and $\delta > 0$, we have

$$\limsup_n P\left(\sum_{j\geq J_0} V(H_n, h_j)^2 > \varepsilon\right) < \delta.$$

Namely, H_n is asymptotically finite-dimensional.

Furthermore, it follows from the definitions of h_j and $V(\cdot,\cdot)$ that

$$\begin{aligned} V(H_n,h_j) &= \frac{h^{1/2+a}}{(1+\lambda\rho_j)}\left\{\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\int_{\mathbb{I}} h_j(u)\int_0^{\tau}\left(X_i(u)-\frac{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}X(u)]}{E[\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}]}\right)dM_i(t)\,du\right\} \\ &\xrightarrow{d} N\left(0, \frac{h^{1+2a}}{(1+\lambda\rho_j)^2}\right). \qquad \text{(C.12)} \end{aligned}$$

Following the lines of the proof of Karhunen-Loéve theorem (Alexanderian, 2015), we can get that

$$\mathcal{G}(t) = \sum_j \frac{h^{1/2+a}}{(1+\lambda\rho_j)}\eta_j h_j(t),$$

where $\eta_j, j = 1, 2, \cdots$, are i.i.d standard normal random variables. Thus, $V(\mathcal{G}, h_j)$ follows $N(0, \frac{h^{1+2a}}{(1+\lambda\rho_j)^2})$. From equation (C.12), we have that $V(H_n, h_j)$ converges in distribution to $V(\mathcal{G}, h_j)$.

(ii) The second step is to show that $\sqrt{n}h^{1/2+a}\tilde{\mathcal{S}}_{n2}(\alpha_0)(s)$ converges to zero in probability uniformly in s. It follows from the arguments in Cheng and Shang (2015, page 1379, lines 5–11) that $\omega(s) = O(1)$. Besides, since $h \to 0$, $\{\Sigma - \Omega + V(G, W_\lambda G^\top)\}^{-1} \to \{\Sigma - \Omega\}^{-1}$, and $n^{-1/2}\sum_{i=1}^{n}\int_0^\tau\{Z_i - V(G,\pi_{X_i})\}\,dM_i(t)$ converges to a normal distribution, we have $h^{a+1/2}\sqrt{n}\tilde{\mathcal{S}}_{n2}(\alpha_0)(s) \xrightarrow{p} 0$ uniformly in s.

Combining (i) and (ii) yields the conclusion of the theorem.

C.5 Proof of Theorem 4.1

Define $\alpha = \hat{\alpha}_{n,\lambda} - \alpha_0$. It follows from Theorem 3.1 that for some $M > 0$, $\|\alpha\|_\lambda \leq Mr_n$ with probability approaching to one. Therefore, we can assume $\|\alpha\|_\lambda \leq Mr_n$. By the Taylor expansion, we can get that

$$\begin{aligned}
l_{n,\lambda}(\alpha_0) - l_{n,\lambda}(\hat{\alpha}_{n,\lambda}) &= -\mathcal{S}_{n,\lambda}(\hat{\alpha}_{n,\lambda})\alpha + \int_0^1\int_0^1 sD\mathcal{S}_{n,\lambda}(\hat{\alpha}_{n,\lambda} - ss'\alpha)\alpha\alpha\,ds\,ds' \\
&= \int_0^1\int_0^1 s\{D\mathcal{S}_{n,\lambda}(\hat{\alpha}_{n,\lambda} - ss'\alpha) - D\mathcal{S}_{n,\lambda}(\alpha_0)\}\alpha\alpha\,ds\,ds' + \frac{1}{2}D\mathcal{S}_{n,\lambda}(\alpha_0)\alpha\alpha. \qquad \text{(C.13)}
\end{aligned}$$

It follows from Lemma B.1 that

$$\begin{aligned}
l_{n,\lambda}(\alpha_0) - l_{n,\lambda}(\hat{\alpha}_{n,\lambda}) = &\int_0^1\int_0^1 s[D\mathcal{S}_{n,\lambda}(\hat{\alpha}_{n,\lambda} - ss'\alpha) - D\mathcal{S}_{n,\lambda}(\alpha_0)]\alpha\alpha\,ds\,ds' \\
&+ \frac{1}{2}[D\mathcal{S}_{n,\lambda}(\alpha_0) - D\mathcal{S}_{\lambda}(\alpha_0)]\alpha\alpha - \frac{1}{2}\|\alpha\|_\lambda \\
\equiv &I_1 + I_2 - \frac{1}{2}\|\alpha\|_\lambda.
\end{aligned}$$

To get the order of I_1, we define $\alpha' = \hat{\alpha}_{n,\lambda} - ss'\alpha - \alpha_0 = (1-ss')\alpha$, where $0 \le s, s' \le 1$. Direct calculations yield that

$$
\begin{aligned}
&|[D\mathcal{S}_{n,\lambda}(\hat{\alpha}_{n,\lambda} - ss'\alpha) - D\mathcal{S}_{n,\lambda}(\alpha_0)]\alpha\alpha| = |D^2\mathcal{S}_{n,\lambda}(\alpha_0 + \delta\alpha')\alpha\alpha\alpha'| \\
&\asymp \Big|\frac{1}{n}\sum_{i=1}^{n}\Delta_i\Big[\frac{\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\eta_\alpha(W_j)\eta_{\alpha'}(W_j)}{nS_1^{(0)}(Y_i,\alpha_0)} \\
&\qquad - \frac{\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\eta_\alpha(W_j)\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_0}(W_j)\}\eta_{\alpha'}(W_j)}{[nS_1^{(0)}(Y_i,\alpha_0)]^2} \\
&\qquad - 2\frac{\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\eta_{\alpha'}(W_j)\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)}{[nS_1^{(0)}(Y_i,\alpha_0)]^2} \\
&\qquad + 2\frac{\{\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\}^2\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_0}(W_j)\}\eta_{\alpha'}(W_j)}{[nS_1^{(0)}(Y_i,\alpha_0)]^3}\Big]\Big| \\
&\lesssim \sup_{t\in\mathbb{I}}\Big|\frac{\sum_j \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\eta_\alpha(W_j)\eta_{\alpha'}(W_j)}{ns_1^{(0)}(t,\alpha_0)}\Big| \\
&+ \sup_{t\in\mathbb{I}}\Big|\frac{\sum_j \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\eta_\alpha(W_j)\sum_j \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_0}(W_j)\}\eta_{\alpha'}(W_j)}{\{ns_1^{(0)}(t,\alpha_0)\}^2}\Big| \\
&+ 2\sup_{t\in\mathbb{I}}\Big|\frac{\sum_j \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\eta_{\alpha'}(W_j)\sum_j \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)}{\{ns_1^{(0)}(t,\alpha_0)\}^2}\Big| \\
&+ 2\sup_{t\in\mathbb{I}}\Big|\frac{\{\sum_j \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\}^2\sum_j \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\eta_{\alpha'}(W_j)}{[ns_1^{(0)}(t,\alpha_0)]^3}\Big| \\
&\equiv I_{11} + I_{12} + I_{13} + I_{14},
\end{aligned}
$$

where $0 \le \delta \le 1$. Define $\Gamma = \cap_{i=1}^n A_{ni}$, where

$$A_{ni} = \{\|Z_i\|_2 \le c\log(n), \|X_i\|_{L^2} \le c\log(n), \exp\{\eta_{\alpha_0}(W_i)\} \le c\log(n)\}$$

for a constant c. For I_{11}, we have

$$\begin{aligned}
I_{11} &= \sup_{t\in\mathbb{I}} \Big| \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\eta_\alpha(W_j)\eta_{\alpha'}(W_j)}{s_1^{(0)}(t,\alpha_0)} \Big| \\
&\leq \Big| \frac{n^{-1}\sum_{j=1}^n \exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\eta_\alpha(W_j)\eta_{\alpha'}(W_j)}{s_1^{(0)}(\tau,\alpha_0)} \Big| \\
&\leq \frac{Mr_n}{ns_1^{(0)}(\tau,\alpha_0)} \Big| \sum_{j=1}^n \exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\eta_\alpha(W_j)\|\mathcal{R}_{W_j}\|_\lambda \Big| \\
&\leq \frac{Mr_n}{ns_1^{(0)}(\tau,\alpha_0)} \Big| \Big\langle \sum_{j=1}^n \exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\|\mathcal{R}_{W_j}\|_\lambda \mathcal{R}_{W_j}, \alpha \Big\rangle_\lambda \Big|.
\end{aligned}$$

Let $d_n = \kappa M h^{-(2a+1)} r_n$, and $\tilde{\alpha} = d_n^{-1}\alpha$, where κ is given in Lemma B.3. Note that $\tilde{\alpha} \in \mathcal{F}_{p_n}$, where $p_n = \kappa^{-2}h^{2a+1-2k} > 1$ for large enough n. Denote

$$\phi_n(Y_i,\Delta_i,W_i;\tilde{\alpha}) = \frac{\exp\{\eta_{\alpha_0}(W_i)\}\eta_{\tilde{\alpha}}(W_i)\|\mathcal{R}_{W_i}\|_\lambda \mathcal{R}_{W_i}}{\sqrt{2c_r}\{c\log(n)\}^3 h^{-(a+1/2)}} I_{A_{ni}}.$$

Then it can be shown that

$$|\phi_n(Y_i,\Delta_i,W_i;\tilde{\alpha}_1) - \phi_n(Y_i,\Delta_i,W_i;\tilde{\alpha}_2)| \leq \|\tilde{\alpha}_1 - \tilde{\alpha}_2\|_e.$$

It follows from Lemma B.5 that with probability approaching to one,

$$\Big\| n^{-1/2}\sum_{j=1}^n \left[\exp\{\eta_{\alpha_0}(W_j)\}\eta_{\tilde{\alpha}}(W_j)\|\mathcal{R}_{W_j}\|_\lambda \mathcal{R}_{W_j} I_{A_{nj}} - E\exp\{\eta_{\alpha_0}(W_j)\}\eta_{\tilde{\alpha}}(W_j)\|\mathcal{R}_{W_j}\|_\lambda \mathcal{R}_{W_j} I_{A_{nj}}\right] \Big\|_\lambda$$

$$\lesssim p_n^{1/(4m)}\{h^{-1}\log\log(n)\}^{1/2}\{c\log(n)\}^3 h^{-(a+1/2)}.$$

Therefore, we have

$$\Big\| \sum_{j=1}^n [\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\|\mathcal{R}_{W_j}\|_\lambda \mathcal{R}_{W_j} I_{A_{nj}} - E\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\|\mathcal{R}_{W_j}\|_\lambda \mathcal{R}_{W_j} I_{A_{nj}}] \Big\|_\lambda$$

$$\begin{aligned}
&\lesssim n^{1/2}p_n^{1/(4m)}\{h^{-1}\log\log(n)\}^{1/2}\{c\log(n)\}^3 h^{-(a+1/2)} d_n \\
&= c^3 M r_n n^{1/2}\kappa^{1-1/(2m)} h^{-(2a+3/2)+(2a+1-2k)/(4m)}\{\log(n)\}^3\{\log\log(n)\}^{1/2}.
\end{aligned}$$

It follows from the Cauchy-Schwarz inequality, Lemma B.2, and Lemma B.4 that

$$
\begin{aligned}
&|E[\exp\{\eta_{\alpha_0}(W_j)\}\eta_\alpha(W_j)\|\mathcal{R}_{W_j}\|_\lambda \left\langle \mathcal{R}_{W_j}, \alpha\right\rangle_\lambda]| \\
&\leq \left(E[\exp\{\eta_{\alpha_0}(W_j)\}]^2\right)^{1/2}\{E\|\mathcal{R}_W\|_\lambda^2\}\{E\left\langle \mathcal{R}_{W_j}, \alpha\right\rangle_\lambda^4\}^{1/2} \\
&\lesssim \left(E[\exp\{\eta_{\alpha_0}(W_j)\}]^2\right)^{1/2}\sqrt{c_r}h^{-1/2}\sqrt{\|\alpha\|_\lambda^4} \\
&\lesssim \left(E[\exp\{\eta_{\alpha_0}(W_j)\}]^2\right)^{1/2}h^{-1/2}M^2 r_n^2.
\end{aligned}
$$

Thus, with probability approaching to one,

$$
|I_{11}| \lesssim (r_n^3 n^{-1/2} h^{-(2a+3/2)+(2a+1-2k)/(4m)}\{\log(n)\}^3\{\log\log(n)\}^{1/2} + h^{-1/2}r_n^3).
$$

Similarly, we can prove that

$I_{12} = O_p(r_n^3 n^{-1/2} h^{-(2a+3/2)+(2a+1-2k)/(4m)}\{\log(n)\}^3\{\log\log(n)\}^{1/2} + h^{-1/2}r_n^3)$,

$I_{13} = O_p(r_n^3 n^{-1/2} h^{-(2a+3/2)+(2a+1-2k)/(4m)}\{\log(n)\}^3\{\log\log(n)\}^{1/2} + h^{-1/2}r_n^3)$,

$I_{14} = O_p(r_n^3 n^{-1/2} h^{-(2a+3/2)+(2a+1-2k)/(4m)}\{\log(n)\}^3\{\log\log(n)\}^{1/2} + h^{-1/2}r_n^3)$. Therefore, we have

$$
I_1 = O_p(r_n^3 n^{-1/2} h^{-(2a+3/2)+(2a+1-2k)/(4m)}\{\log(n)\}^3\{\log\log(n)\}^{1/2} + h^{-1/2}r_n^3) = o_p(n^{-1}h^{-1/2}).
$$

It follows from equation (C.7) that

$$
\begin{aligned}
2|I_2| &= |[D\mathcal{S}_{n,\lambda}(\alpha_0) - D\mathcal{S}_\lambda(\alpha_0)]\alpha\alpha| \\
&= O_p(n^{-1/2}h^{-(a+1)-\frac{2k-2a-1}{4m}}\{\log(n)\}^2\{\log\log(n)\}^{1/2}r_n^2) = o_p(n^{-1}h^{-1/2}).
\end{aligned}
$$

Therefore, we have

$$
-2n\mathrm{PLRT}_{n,\lambda} = n\|\hat{\alpha}_{n,\lambda} - \alpha_0\|_\lambda^2 + o_p(h^{-1/2}).
$$

It follows from Theorem 3.2 and $n^{1/2}a_n = o(1)$ that

$$-2n\mathrm{PLRT}_{n,\lambda} = n\|\mathcal{S}_{n,\lambda}(\alpha_0)\|_\lambda^2(1+o_p(1)) + o_p(h^{-1/2}).$$

Direct calculations yield that

$$\begin{aligned}
& n\|\mathcal{S}_{n,\lambda}(\alpha_0)\|_\lambda^2 \\
=& \frac{1}{n}\Big\|\sum_{i=1}^n \int_0^\tau \Big[\mathcal{R}_{W_i} - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{S_1^{(0)}(t,\alpha_0)}\Big]\, dM_i(t) - \mathcal{P}_\lambda\alpha_0\Big\|_\lambda^2 \\
=& \frac{1}{n}\Big\|\sum_{i=1}^n \int_0^\tau \Big[\mathcal{R}_{W_i} - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{S_1^{(0)}(t,\alpha_0)}\Big]\, dM_i(t)\Big\|_\lambda^2 \\
& - 2\Big\langle \frac{1}{n}\sum_{i=1}^n \int_0^\tau \Big[\mathcal{R}_{W_i} - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{S_1^{(0)}(t,\alpha_0)}\Big]\, dM_i(t), \mathcal{P}_\lambda\alpha_0\Big\rangle_\lambda + n\|\mathcal{P}_\lambda\alpha_0\|_\lambda^2 \\
\equiv& J_1 + J_2 + J_3.
\end{aligned}$$

For J_1 and J_2, it follows from Condition (C1) that

$$\begin{aligned}
J_1 =& \frac{1}{n}\Big\|\sum_{i=1}^n \int_0^\tau \Big[\mathcal{R}_{W_i} - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{S_1^{(0)}(t,\alpha_0)}\Big]\, dM_i(t)\Big\|_\lambda^2 \\
=& \frac{1}{n}\Big\|\sum_{i=1}^n \int_0^\tau \Big[\mathcal{R}_{W_i} - \frac{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{s_1^{(0)}(t,\alpha_0)}\Big]\, dM_i(t)\Big\|_\lambda^2 + O_p(n^{-1}h^{-1-2a}) \\
=& \frac{1}{n}\Big\|\sum_{i=1}^n \int_0^\tau \Big[\mathcal{R}_{W_i} - \frac{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{s_1^{(0)}(t,\alpha_0)}\Big]\, dM_i(t)\Big\|_\lambda^2 + o_p(1),
\end{aligned}$$

and

$$\begin{aligned}
& \Big|\frac{J_2}{2}\Big| = \Big|\Big\langle \frac{1}{n}\sum_{i=1}^n \int_0^\tau \Big[\mathcal{R}_{W_i} - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{S_1^{(0)}(t,\alpha_0)}\Big]\, dM_i(t), \mathcal{P}_\lambda\alpha_0\Big\rangle_\lambda\Big| \\
\le& \Big|\Big\langle \frac{1}{n}\sum_{i=1}^n \int_0^\tau \Big[\mathcal{R}_{W_i} - \frac{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{s_1^{(0)}(t,\alpha_0)}\Big]\, dM_i(t), \mathcal{P}_\lambda\alpha_0\Big\rangle_\lambda\Big| \\
+& \Big|\Big\langle \frac{1}{n}\sum_{i=1}^n \int_0^\tau \Big[\frac{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{s_1^{(0)}(t,\alpha_0)} - \frac{n^{-1}\sum_{j=1}^n \mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{S_1^{(0)}(t,\alpha_0)}\Big]\, dM_i(t), \mathcal{P}_\lambda\alpha_0\Big\rangle_\lambda\Big| \\
=& \Big|\Big\langle \frac{1}{n}\sum_{i=1}^n \int_0^\tau \Big[\mathcal{R}_{W_i} - \frac{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0}(W_j)\}\mathcal{R}_{W_j}}{s_1^{(0)}(t,\alpha_0)}\Big]\, dM_i(t), \mathcal{P}_\lambda\alpha_0\Big\rangle_\lambda\Big| + O_p(n^{-1/2}h^{-1/2-a})\|\mathcal{P}_\lambda\alpha_0\|_\lambda.
\end{aligned}$$

Denote $\beta_0 = \sum_j b_j h_j$. Since $J(\beta_0, \beta_0) = \sum_j b_j^2 \rho_j < \infty$, and $\lambda = o(1)$, it follows from the Lebesgue dominated convergence theorem that

$$\begin{aligned}
&E\left[\left\langle \sum_{i=1}^{n} \int_0^{\tau} \left\{ \mathcal{R}_{W_i} - \frac{E[\mathcal{Y}_j(t) \exp\{\eta_{\alpha_0}(W_j)\} \mathcal{R}_{W_j}]}{s_1^{(0)}(t,\alpha_0)} \right\} dM_i(t), \mathcal{P}_\lambda \alpha_0 \right\rangle_\lambda^2\right] \\
=& nE\left[\int_0^{\tau} \left\{ \int_{\mathbb{I}} [X(t) - E\{X(t)|T=v, \Delta=1\}] W_\lambda(\beta_0)(t)\, dt \right\}^2 \mathcal{Y}(v) \exp\{\eta_{\alpha_0}(W)\} h_0(v)\, dv\right] \\
=& nV(W_\lambda(\beta_0), W_\lambda(\beta_0)) \le n\|W_\lambda(\beta_0)\|_m = n\lambda \sum_j b_j^2 \rho_j \frac{\lambda \rho_j}{(1+\lambda\rho_j)} = o_p(n\lambda).
\end{aligned}$$

Therefore, we have $J_2 = o_p((n\lambda)^{1/2})(1+(nh)^{-1/2}) = o_p((n\lambda)^{1/2})$. Note that $J_3 = n\|\mathcal{P}_\lambda \alpha_0\|_\lambda^2 = n\|W_\lambda(\beta_0)\|_m^2 = o(n\lambda)$. Therefore, we have

$$\begin{aligned}
&-2n\mathrm{PLRT}_{n,\lambda} \\
=& \frac{1}{n} \left\| \sum_{i=1}^{n} \int_0^{\tau} \left(\mathcal{R}_{W_i} - \frac{E[\mathcal{Y}_j(t) \exp\{\eta_{\alpha_0}(W_j)\} \mathcal{R}_{W_j}]}{s_1^{(0)}(t,\alpha_0)} \right) dM_i(t) \right\|_\lambda^2 + n\|W_\lambda(\beta_0)\|_m^2 + o_p(h^{-1/2}).
\end{aligned}$$

Define $R_i(t) = \mathcal{R}_{W_i} - E[\mathcal{Y}_j(t) \exp\{\eta_{\alpha_0}(W_j)\} \mathcal{R}_{W_j}]/s_1^{(0)}(t,\alpha_0)$. To get the asymptotic results of $-2n\mathrm{PLRT}_{n,\lambda}$, we only need to figure out the properties of

$$\frac{1}{n} \left\| \sum_{i=1}^{n} \int_0^{\tau} R_i(t)\, dM_i(t) \right\|_\lambda^2 = \frac{1}{n} \sum_{i=1}^{n} \int_0^{\tau} \int_0^{\tau} \langle R_i(t), R_i(s) \rangle_\lambda \, dM_i(t)\, dM_i(s) + \frac{1}{n} \sum_{1 \le i < j \le n} W_{ij},$$

where $W_{ij} = 2\int_0^{\tau} \int_0^{\tau} \langle R_i(t), R_j(s) \rangle_\lambda \, dM_i(t)\, dM_j(s)$. Write $W_n = \sum_{1 \le i < j \le n} W_{ij}$. So, W_n is clean (Jong, 1987). Next, we aim to derive the limiting distribution of W_n. Let $\sigma_n^2 = Var(W_n)$. Then

$$\begin{aligned}
\sigma_n^2 &= \frac{n(n-1)}{2} E(W_{ij}^2) = 2n(n-1) E\left\{ \int_0^{\tau} \int_0^{\tau} \langle R_i(t), R_j(s) \rangle_\lambda \, dM_i(t)\, dM_j(s) \right\}^2 \\
&\asymp 2n(n-1) \left\{ \sum_{l=1}^{\infty} \frac{1}{(1+\lambda\rho_l)^2} + 1 \right\} \asymp 2n^2 h^{-1} \sigma_\lambda^4 / \rho_\lambda^2.
\end{aligned}$$

Define F_1, F_2, and F_3 as follows:

$$F_1 \equiv \sum_{i<j} E(W_{ij}^4), \quad F_2 \equiv \sum_{i<j<k} \left\{E(W_{ij}^2 W_{ik}^2) + E(W_{ji}^2 W_{jk}^2) + E(W_{ki}^2 W_{kj}^2)\right\}, \quad \text{and}$$

$$F_3 \equiv \sum_{i<j<k<l} \left\{E(W_{ij}W_{ik}W_{lj}W_{lk}) + E(W_{ij}W_{il}W_{kj}W_{kl}) + E(W_{ik}W_{il}W_{jk}W_{jl})\right\}.$$

By Proposition 3.2 of Jong (1987), if F_1, F_2, F_3 are all of lower order than σ_n^4, then $\sigma_n^{-1}W_n$ converges weakly to the standard normal distribution. Now, we study the order of each $F_i, i = 1, 2, 3$. First, observe that

$$\begin{aligned}
&E\left(W_{ij}^4\right) = 16E\left\{\int_0^\tau \int_0^\tau \langle R_i(t), R_j(s)\rangle_\lambda \, dM_i(t)\, dM_j(s)\right\}^4 \\
&= 16 \int_0^\tau \int_0^\tau \int_0^\tau \int_0^\tau \int_0^\tau \int_0^\tau \int_0^\tau \int_0^\tau E\, \langle R_i(t_1), R_j(s_1)\rangle_\lambda \langle R_i(t_2), R_j(s_2)\rangle_\lambda \langle R_i(t_3), R_j(s_3)\rangle_\lambda \\
&\langle R_i(t_4), R_j(s_4)\rangle_\lambda \left\{ dM_i(t_1)\, dM_j(s_1)\, dM_i(t_2)\, dM_j(s_2)\, dM_i(t_3)\, dM_j(s_3)\, dM_i(t_4)\, dM_j(s_4)\right\} \\
&= O(h^{-4}),
\end{aligned}$$

which implies $F_1 = O(n^2h^{-4})$. Next, by the Cauchy-Schwarz inequity,

$$E(W_{ij}^2 W_{ik}^2) \le \{E(W_{ij}^4)\}^{1/2}\{E(W_{ik}^4)\}^{1/2} = O(h^{-4}),$$

which yields $F_2 = O(n^3h^{-4})$. A straightforward calculation yields that

$$E(W_{ij}W_{ik}W_{lj}W_{lk}) \sim 16\sum_{j=1}^{\infty} \frac{1}{(1+\lambda\rho_j)^4} = O(h^{-1}).$$

Therefore, $F_3 = O(n^4h^{-1})$. Combining the fact that $\sigma_n^4 = (\sigma_n^2)^2 = O(n^4h^{-2})$ with the assumptions that $nh^2 \to \infty$ and $h = o(1)$, F_1, F_2, F_3 are of lower order than that of σ_n^4. Hence, by Jong (1987), $\sigma_n^{-1}W_n \xrightarrow{d} N(0,1)$ as $n \to \infty$. Recall that $\rho_\lambda^2 = \sum_{j=1}^{\infty} h/(1+\lambda\rho_j)^2$. We have

$$\frac{1}{\sqrt{2h^{-1}}n\rho_\lambda} W_n \xrightarrow{d} N(0,1). \tag{C.14}$$

Lastly, we consider $n^{-1}\sum_{i=1}^{n}\int_0^\tau\int_0^\tau \langle R_i(t), R_i(s)\rangle_\lambda \, dM_i(t)\, dM_i(s)$. Through direct calculations, we can obtain that

$$E\left\{\int_0^\tau\int_0^\tau \langle R_i(t), R_i(s)\rangle_\lambda \, dM_i(t)\, dM_i(s)\right\}^2 = O(\{E\|\mathcal{R}_{W_i}\|_\lambda^2\}^2) = O(h^{-2}).$$

Then,

$$\begin{aligned} & E\left\{\sum_{i=1}^{n}\int_0^\tau\int_0^\tau \langle R_i(t), R_i(s)\rangle_\lambda \, dM_i(t)\, dM_i(s) - h^{-1}\sigma_\lambda^2 - 1\right\}^2 \\ \leq \; & nE\left\{\int_0^\tau\int_0^\tau \langle R_i(t), R_i(s)\rangle_\lambda \, dM_i(t)\, dM_i(s)\right\}^2 = O(nh^{-2}), \end{aligned}$$

where $\sigma_\lambda^2 = \sum_{j=0}^{\infty} h/(1+\lambda\rho_j)$. Combining these gives

$$\frac{1}{n}\sum_{i=1}^{n}\int_0^\tau\int_0^\tau \langle R_i(t), R_i(s)\rangle_\lambda \, dM_i(t)\, dM_i(s) = 1 + h^{-1}\sigma_\lambda^2 + O_p\{(n^{1/2}h)^{-1}\}. \tag{C.15}$$

By (C.14) and (C.15), we have $n\|\mathcal{S}_{n,\lambda}(\alpha_0)\|_\lambda^2 = O_p(h^{-1})$ and therefore $n^{1/2}\|\mathcal{S}_{n,\lambda}(\alpha_0)\|_\lambda = O_p(h^{-1/2})$. As a result,

$$-2n\mathrm{PLRT}_{n,\lambda} = \{n^{1/2}\|\mathcal{S}_{n,\lambda}(\alpha_0)\|_\lambda + o_p(1)\}^2 + o_p(h^{-1/2}) = n\|\mathcal{S}_{n,\lambda}(\alpha_0)\|_\lambda^2 + o_p(h^{-1/2}). \tag{C.16}$$

In view of (C.14), (C.15) and (C.16), we conclude that as $n \to \infty$,

$$(2h^{-1}\sigma_\lambda^4/\rho_\lambda^2)^{-1/2}\left\{-2n\gamma_\lambda PLRT_{n,\lambda} - n\gamma_\lambda\|W_\lambda\beta_0(t)\|_\lambda^2 - h^{-1}\sigma_\lambda^4/\rho_\lambda^2\right\} \xrightarrow{d} N(0,1).$$

The proof of Theorem 4.1 is completed.

C.6 Proof of Theorem 4.2

Throughout this proof, we only consider $\alpha_{n_0} = \alpha_0 + \alpha_n$ for $\alpha_n \in \mathcal{A}$ in H_1. To prove the theorem, we write

$$-2n \cdot \mathrm{PLRT}_{n,\lambda} = -2n\{l_{n,\lambda}(\alpha_0) - l_{n,\lambda}(\alpha_{n_0})\} - 2n\{l_{n,\lambda}(\alpha_{n_0}) - l_{n,\lambda}(\hat{\alpha}_{n,\lambda})\} \equiv I_1 + I_2. \tag{C.17}$$

We first consider I_1. For simplicity, we denote

$$
\begin{aligned}
R_i &= \Delta_i \left[\eta_{\alpha_0}(W_i) - \log \frac{1}{n}\sum_{j=1}^{n} \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_0}(W_j)\}\right] \\
&- \Delta_i \left[\eta_{\alpha_{n_0}}(W_i) - \log \frac{1}{n}\sum_{j=1}^{n} \mathcal{Y}_j(Y_i)\exp\{\eta_{\alpha_{n_0}}(W_j)\}\right] \\
&= -\int_0^\tau \left[\eta_{\alpha_n}(W_i) - \frac{\sum_{j=1}^{n}\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0+s'\alpha_n}(W_j)\}\eta_{\alpha_n}(W_j)}{\sum_{j=1}^{n}\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0+s'\alpha_n}(W_j)\}}\right] dN_i(t) \\
&= -\int_0^\tau \left[\eta_{\alpha_n}(W_i) - \frac{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0+s'\alpha_n}(W_j)\}\eta_{\alpha_n}(W_j)}{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0+s'\alpha_n}(W_j)\}}\right] dN_i(t) \\
&\quad + \int_0^\tau \left[\frac{\sum_{j=1}^{n}\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0+s'\alpha_n}(W_j)\}\eta_{\alpha_n}(W_i)}{\sum_{j=1}^{n}\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0+s'\alpha_n}(W_j)\}} - \frac{E[\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0+s'\alpha_n}(W_j)\}\eta_{\alpha_n}(W_j)]}{E[\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0+s'\alpha_n}(W_j)\}]}\right] dN_i(t) \\
&= -\int_0^\tau \left[\eta_{\alpha_n}(W_i) - \frac{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0+s'\alpha_n}(W_j)\}\eta_{\alpha_n}(W_j)}{E\mathcal{Y}_j(t)\exp\{\eta_{\alpha_0+s'\alpha_n}(W_j)\}}\right] dN_i(t) + o_p(\|\alpha_n\|_\lambda),
\end{aligned}
$$

where $0 \le s' \le 1$. Then

$$
E\{R_i^2\} \asymp E\int_0^\tau Var\{\eta_{\alpha_n}(W)|T=t,\Delta=1\}\mathcal{Y}(t)\exp\{\eta_{\alpha_0}(W)\}h_0(t)\,dt = O(\|\alpha_n\|_\lambda^2).
$$

Therefore, we can get

$$
E\left\{\left|\sum_{i=1}^{n}(R_i - ER_i)\right|^2\right\} \le nE\{R_i^2\} = O(n\|\alpha_n\|_\lambda^2).
$$

Combining these gives

$$
n\left[l_{n,\lambda}(\alpha_0) - l_{n,\lambda}(\alpha_{n_0}) - E\{l_{n,\lambda}(\alpha_0) - l_{n,\lambda}(\alpha_{n_0})\}\right] = O_p(n^{1/2}\|\alpha_n\|_\lambda).
$$

On the other hand, since $D\mathcal{S}_\lambda(\alpha)\alpha_n\alpha_n < 0$ for any $\alpha \in \mathcal{H}$, there exists a constant $c > 0$ such that $\{D\mathcal{S}_\lambda(\alpha^*_{n_0})\alpha_n\alpha_n\} \le c\{D\mathcal{S}_\lambda(\alpha_{n_0})\alpha_n\alpha_n\} = -c\|\alpha_n\|^2_\lambda$. Then, we have

$$
\begin{aligned}
&E\{l_{n,\lambda}(\alpha_0) - l_{n,\lambda}(\alpha_{n_0})\} = E\left\{\mathcal{S}_{n,\lambda}(\alpha_{n_0})(-\alpha_n) + \frac{1}{2}D\mathcal{S}_{n,\lambda}(\alpha^*_{n_0})\alpha_n\alpha_n\right\}\\
&\le \lambda J(\alpha_{n_0}, \alpha_n) - \frac{c\|\alpha_n\|^2_\lambda}{2} \le \{J(\alpha_n, \alpha_n) + J(\alpha_0, \alpha_n)\} - \frac{c\|\alpha_n\|^2_\lambda}{2}\\
&\le \{J(\alpha_n, \alpha_n) + J(\alpha_0, \alpha_0)^{1/2}J(\alpha_n, \alpha_n)^{1/2}\} - \frac{c\|\alpha_n\|^2_\lambda}{2}\\
&= O(\lambda) - \frac{c\|\alpha_n\|^2_\lambda}{2}.
\end{aligned}
$$

It then follows that

$$
I_1 \ge n\|\alpha_n\|^2_\lambda + O_p(n\lambda + n^{1/2}\|\alpha_n\|_\lambda) = n\|\alpha_n\|^2_\lambda\{1 + O_p(\lambda\|\alpha_n\|^{-2}_\lambda + n^{-1/2}\|\alpha_n\|^{-1}_\lambda)\}. \quad \text{(C.18)}
$$

Second, we consider I_2. Under the alternative hypothesis, $\|\hat{\alpha}_{n,\lambda} - \alpha_{n_0}\| = O_p\{(nh)^{-1/2} + h^k\}$. By the joint functional Bahadur representation given in Theorem 3.2, we have

$$
\inf_{n\ge N}\inf_{\alpha_n\in\mathcal{A}} P_{\alpha_{n_0}}\left(\|\hat{\alpha}_{n,\lambda} - \alpha_{n_0} - S_{n,\lambda}(\alpha_{n_0})\|_\lambda \le Mr_n\right) \to 1, \quad \text{(C.19)}
$$

where $r_n = (nh)^{-1/2} + h^k$, and $P_{\alpha_{n_0}}$ means that the probability relies on α_{n_0}. Note that under the alternative hypothesis H_{1n}, I_2 is the same as (C.13) except one constant term $-2n$. Along the lines of Theorem 4.1, we can show that I_2 has the same limiting distribution as given in Theorem 4.1, uniformly for any $\alpha_n \in \mathcal{A}$. In other words, uniformly over all $\alpha_n \in \mathcal{A}$,

$$
(2\nu_{n_0})^{-1/2}(I_2 - n\|W_\lambda\beta_{n_0}\|^2_m - h^{-1}\sigma^2_{n_0,\lambda}) = O_p(1), \quad \text{(C.20)}
$$

where $\nu_{n_0} = h^{-1}\sigma^4_{n_0,\lambda}/\rho^2_{n_0,\lambda}$, $\sigma^2_{n_0,\lambda}$ and $\rho^2_{n_0,\lambda}$ are defined as the same as σ^2_λ and ρ^2_λ but with eigenvalues and eigenfunctions obtained under α_{n_0}. Next, let $V_{n_0}(f, g) = \int_{\mathbb{I}}\int_{\mathbb{I}} F_{\alpha_{n_0}}(s,t)f(t)g(s)\,dt\,ds$

and $V_0(f,g) = \int_{\mathbb{I}}\int_{\mathbb{I}} F_{\alpha_0}(s,t)f(t)g(s)\,dt\,ds$, where $F_{\alpha_0}(s,t) = F(s,t)$, while $F_{\alpha_{n_0}}(s,t)$ has the same formula as $F_{\alpha_0}(s,t)$ but replacing α_0 with α_{n_0}. Thus, for any $f \in \mathcal{H}^{(m)}$, there exists a constant c such that

$$|V_{n_0}(f,f) - V_0(f,f)| = \left| \int_{\mathbb{I}}\int_{\mathbb{I}} \big[F_{\alpha_{n_0}}(s,t) - F_{\alpha_0}(s,t)\big] f(t)f(s)\,dt\,ds \right|$$
$$\leq E\|\exp\{\alpha_n(W)\}\|_\infty V_0(f,f)\|\alpha_n\|_\infty = cV_0(f,f)\|\alpha_n\|_\infty.$$

It follows from the Supplementary Material (page 56) of Shang and Cheng (2013) that

$$\sigma^2_{n_0,\lambda} - \sigma^2_\lambda = O(h^{-(a+1)/2}\|\alpha_n\|_\lambda). \tag{C.21}$$

Combining (C.18), (C.20) and (C.21) gives

$$(2\nu_n)^{-1/2}(-2nr_\lambda \mathrm{PLRT}_{n,\lambda} - \nu_n) = (2\nu_n)^{-1/2}\{-r_\lambda(I_1 + I_2) - \nu_n\}$$
$$=(2\nu_n)^{-1/2} r_\lambda (I_2 - n\|\mathcal{P}_\lambda \alpha n_0\|^2_\lambda - h^{-1}\sigma^2_{n_0,\lambda}) + (2\nu_n)^{-1/2} r_\lambda n \|\mathcal{P}_\lambda \alpha n_0\|^2_\lambda$$
$$+ (2\nu_n)^{-1/2} r_\lambda I_1 + (2\nu_n)^{-1/2} r_\lambda h^{-1}(\sigma^2_{n_0,\lambda} - \sigma^2_\lambda)$$
$$\geq O_p(1) + (2\nu_n)^{-1/2} r_\lambda n \|\alpha_n\|^2_\lambda \{1 + O_p(\lambda\|\alpha_n\|^{-2}_\lambda + n^{-1/2}\|\alpha_n\|^{-1}_\lambda)\}$$
$$+ O(h^{-3/2-a/2}\|\alpha_n\|_\lambda),$$

where $O_p(1)$ holds uniformly in $\mathcal{A}$, $\nu_n = h^{-1}\sigma^4_\lambda/\rho^2_\lambda$, and r_λ is defined in Theorem 4.1. Let $\lambda\|\alpha_n\|^{-2}_\lambda \leq 1/c$, $n^{-1/2}\|\alpha_n\|^{-1}_\lambda \leq 1/c$, $ch^{-3/2-a/2}\|\alpha_n\|_\lambda \leq n\|\alpha_n\|^2_\lambda$, and $\|\alpha_n\|^2_\lambda \geq c(nh^{1/2})^{-1}$ for a sufficiently small constant c. In other words,

$$|(2\nu_n)^{-1/2}(-2nr_\lambda \mathrm{PLRT}_{n,\lambda} - \nu_n)| \geq c_\alpha,$$

where c_α is the critical value (based on $N(0,1)$) to H_0^{global} at nominal level α. This leads to

$$\|\alpha_n\|^2_\lambda \gtrsim \{h^{2k} + (nh^{1/2})^{-1}\}. \tag{C.22}$$

Combining (C.19) and (C.22), we complete the proof of Theorem 4.2.

References

Alexanderian A. (2015). A brief note on the Karhunen-Loéve expansion. *http://users.ices.utexas.edu/∼alen/articles/KL.pdf.*

Jong, D. P. (1987). A central limit theorem for generalized quadratic forms. *Probability Theory and Related Fields* **75**, 261-277.

Kosorok, M. R. (2008). *Introduction to Empirical Processes and Semiparametric Inference.* Springer Science and Business Media.

Shang, Z., and Cheng, G. (2013). Local and global asymptotic inference in smoothing spline models. *The Annals of Statistics* **41**, 2608-2638.

Shang, Z. and Cheng, G. (2015). Nonparametric inference in generalized functional linear models. *The Annals of Statistics* **43**, 1742-1773.

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes with Applications to Statistics.* Springer-Verlag, New York.