ORACLE-EFFICIENT GLOBAL INFERENCE FOR VARIANCE FUNCTION IN NONPARAMETRIC REGRESSION WITH MISSING COVARIATES

Li Cai[1] and Suojin Wang[2]

[1]*Zhejiang Gongshang University and* [2]*Texas A&M University*

Supplementary Material

This supplement provides all the proofs for the main results. Throughout this supplement, we denote by $\|\xi\|$ the Euclidean norm and by $\|\xi\|_\infty$ the largest absolute value of the elements of any vector ξ. For any $l \times k$ matrix $M = \{m_{ij}\}_{i=1,j=1}^{l,k}$, denote $\|M\|_\infty = \max_{\zeta\in\mathbb{R}^k,\zeta\neq 0} \|M\zeta\|_\infty / \|\zeta\|_\infty$ which is easily seen to be equivalent to $\|M\|_\infty = \max\limits_{1\leq i\leq l} \sum_{j=1}^{k} |m_{ij}|$. For any function $\psi(x) \in L_2[a,b]$, let $\|\psi(x)\|_\infty = \sup_{x\in[a,b]} |\psi(x)|$.

S1 Preliminaries

Lemma S.1. *(Theorem 1.2 of Bosq (1998)) Let $\xi_1, \ldots, \xi_n$ be independent random variables with mean* 0*. If there exists a constant $r > 0$ such that*

(Cramér's Conditions)

$$\mathrm{E}\left|\xi_i\right|^k \le r^{k-2}k!\,\mathrm{E}\,\xi_i^2 < +\infty \text{ for } 1 \le i \le n, k \ge 3,$$

then for any $t > 0$

$$P\left\{\left|\sum_{i=1}^{n}\xi_i\right| > t\right\} \le 2\exp\left\{-\frac{t^2}{4\sum_{i=1}^{n}\mathrm{E}\,\xi_i^2 + 2rt}\right\}.$$

The next lemma is an important result from de Boor (2001, p.149).

Lemma S.2. *For any* $\phi(x) \in C^{(p)}[a,b]$*, there exist a constant* $C_p > 0$ *and a spline function* $m_p(x) \in G_N^{(p-2)}$ *such that* $\|\phi(x) - m_p(x)\|_\infty \le C_p\left\|\phi^{(p)}(x)\right\|_\infty N^{-p}$.

For any function $\psi(\cdot), \phi(\cdot) \in L_2[a,b]$, define the theoretical and empirical inner products as

$$\langle\psi,\phi\rangle_2 = \int_a^b \psi(x)\,\phi(x)\,f_X(x)\,dx,$$

and

$$\langle\psi,\phi\rangle_{2,n} = n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}\psi(X_i)\,\phi(X_i).$$

The corresponding norms are defined as $\|\phi\|_2^2 = \int_a^b \phi^2(x)\,f(x)\,dx$ and $\|\phi\|_{2,n}^2 = n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}\phi^2(X_i)$. Meanwhile, define the following theoretical and empirical inner product matrices of $\{B_{J,p}(\cdot)\}_{J=1-p}^{N}$:

$$V_p = \left(\langle B_{J,p}, B_{J',p}\rangle_2\right)_{J,J'=1-p}^{N}, \hat{V}_p^* = \left(\langle B_{J,p}, B_{J',p}\rangle_{2,n}\right)_{J,J'=1-p}^{N}.$$

It is clear that $n^{-1}\mathbf{B}^T\mathbf{\Delta}\mathbf{B} = \hat{V}_p^*$. Moreover, denote the matrix

$$\hat{V}_p = \left(n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\hat{\pi}_i}B_{J,p}(X_i)B_{J',p}(X_i)\right)_{J,J'=1-p}^{N},$$

and hence $n^{-1}\mathbf{B}^T\hat{\mathbf{\Delta}}\mathbf{B} = \hat{V}_p$. We next give some properties about the inner product matrices of V_p, $\hat{V}_p$ and $\hat{V}_p^*$ which are needed in the study of the uniform convergence of $\hat{\sigma}^2(x)$.

Lemma S.3. *For any positive p, there exists a constant K_{p1} depending only on p such that $\left\|V_p^{-1}\right\|_\infty \le K_{p1} N$.*

It is a direct conclusion of Lemma A.3 in Cao et al. (2012).

Lemma S.4. *Under Assumption (A6), one has that*

(a) $\left\|\hat{V}_p^* - V_p\right\|_\infty = O_p(n^{-1/2} N^{-1/2} \log^{1/2} n)$;

(b) $\left\|\hat{V}_p^* - \hat{V}_p\right\|_\infty = O_p(n^{-1/2} N^{-1})$;

(c) there exist constants K_{p2}, K_{p3} such that $\left\|\hat{V}_p^{*-1}\right\|_\infty \le K_{p2} N$ and $\left\|\hat{V}_p^{-1}\right\|_\infty \le K_{p3} N$ in probability.

Proof of Lemma S.4(a). Denote

$$\xi_{J,J'}(X_i) = n^{-1} \left\{ \frac{\delta_i}{\pi_i} B_{J,p}(X_i) B_{J',p}(X_i) - \mathrm{E}\{B_{J,p}(X_i) B_{J',p}(X_i)\} \right\}.$$

Since $B_{J,p}(x) B_{J',p}(x) \neq 0$ only if $x \in \left[\chi_J, \chi_{J+p}\right] \cap \left[\chi_{J'}, \chi_{J'+p}\right]$ and $|J - J'| < p$, one has that

$$\begin{aligned}
&\left\|\hat{V}_p^* - V_p\right\|_\infty \\
&= \max_{1-p \le J \le N} \sum_{J' \in \{J' : |J'-J| < p\}} \left| n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi_i} B_{J,p}(X_i) B_{J',p}(X_i) - \mathrm{E}\{B_{J,p}(X_i) B_{J',p}(X_i)\} \right| \\
&= \max_{1-p \le J \le N} \sum_{J' \in \{J' : |J'-J| < p\}} \left| \sum_{i=1}^n \xi_{J,J'}(X_i) \right|.
\end{aligned}$$

It is clear that $\mathrm{E}\,\xi_{J,J'}(X_i) = 0$ and

$$\begin{aligned}
\mathrm{E}\,\xi_{J,J'}^2(X_i) &\le n^{-2}\, \mathrm{E} \left\{ \frac{\delta_i}{\pi_i} B_{J,p}(X_i) B_{J',p}(X_i) \right\}^2 \\
&\le n^{-2} c_\pi^{-1}\, \mathrm{E}\{B_{J,p}(X_i) B_{J',p}(X_i)\}^2 \\
&\le c_\pi^{-1} C n^{-2} N^{-1}, \qquad \text{(S1.1)}
\end{aligned}$$

for some constant $C > 0$ since $\max_{1-p \le J \le N} |B_{J,p}(x)| = O(1)$. On the other hand,

$$\begin{aligned}
\mathrm{E}\, \xi^2_{J,J'}(X_i) &= n^{-2} \left(\mathrm{E} \left\{ \frac{1}{\pi_i} B^2_{J,p}(X_i) B^2_{J',p}(X_i) \right\} - [\mathrm{E}\{B_{J,p}(X_i)B_{J',p}(X_i)\}]^2 \right) \\
&\ge n^{-2} \left(\mathrm{E} \left\{ B^2_{J,p}(X_i) B^2_{J',p}(X_i) \right\} - [\mathrm{E} \left\{ B_{J,p}(X_i)B_{J',p}(X_i) \right\}]^2 \right) \\
&\ge cn^{-2} N^{-1},
\end{aligned}$$

for some $c > 0$ which holds since

$$\mathrm{E} \left\{ B^2_{J,p}(X_i) B^2_{J',p}(X_i) \right\} \gg [\mathrm{E} \left\{ B_{J,p}(X_i)B_{J',p}(X_i) \right\}]^2$$

and by the Mean Value Theorem,

$$\begin{aligned}
\mathrm{E}\left\{ B^2_{J,p}(X_i) B^2_{J',p}(X_i) \right\} &= \int_{x \in [\chi_J, \chi_{J+p}] \cap [\chi_{J'}, \chi_{J'+p}], |J-J'|<p} B^2_{J,p}(x) B^2_{J',p}(x) f_X(x) dx \\
&= B^2_{J,p}(\xi) B^2_{J',p}(\xi) f_X(\xi) \int_{x \in [\chi_J, \chi_{J+p}] \cap [\chi_{J'}, \chi_{J'+p}], |J-J'|<p} 1 dx \sim N^{-1}
\end{aligned}$$

for some $\xi \in (\chi_J, \chi_{J+p}) \cap (\chi_{J'}, \chi_{J'+p})$, $|J - J'| < p$. When $k \ge 3$, the k-th moment $\mathrm{E} \left| \xi_{J,J'}(X_i) \right|^k$ is bounded as follows:

$$\begin{aligned}
\mathrm{E} \left| \xi_{J,J'}(X_i) \right|^k &= n^{-k}\, \mathrm{E} \left| \frac{\delta_i}{\pi_i} B_{J,p}(X_i) B_{J',p}(X_i) - \mathrm{E}\{B_{J,p}(X_i)B_{J',p}(X_i)\} \right|^k \\
&\le n^{-k} 2^{k-1} \left[\mathrm{E} \left\{ \frac{1}{\pi_i^{k-1}} \left| B_{J,p}(X_i) B_{J',p}(X_i) \right|^k \right\} + \left| \mathrm{E}\{B_{J,p}(X_i)B_{J',p}(X_i)\} \right|^k \right] \\
&\le n^{-k} 2^k\, \mathrm{E} \left\{ \frac{1}{\pi_i^{k-1}} \left| B_{J,p}(X_i) B_{J',p}(X_i) \right|^k \right\} \\
&\le n^{-k} 2^k c_\pi^{1-k} C^* N^{-1} = n^{-(k-2)} 2^k c_\pi^{1-k} C^* c^{-1} c n^{-2} N^{-1} \\
&\le n^{-(k-2)} 2^k c_\pi^{1-k} C^* c^{-1}\, \mathrm{E}\, \xi^2_{J,J'}(X_i) \le \left(\frac{2c_0}{n}\right)^{k-2} k!\, \mathrm{E}\, \xi^2_{J,J'}(X_i),
\end{aligned}$$

for some constants $C^*, c_0 > 0$. The first inequality is easy to see by mathematical induction and Young's Inequality. Thus, $\xi_{J,J'}(X_i), 1 \le i \le n$, satisfy Cramér's Conditions with $r = 2c_0/n$ in Lemma S.1. Then one obtains that for any given $\rho > 0$,

$$
\begin{aligned}
P&\left\{\left|\sum_{i=1}^{n}\xi_{J,J'}(X_i)\right| > \rho n^{-1/2}N^{-1/2}\log^{1/2} n\right\}\\
&\le 2\exp\left\{-\frac{\rho^2 n^{-1}N^{-1}\log n}{4\sum_{i=1}^{n}\mathrm{E}\,\xi_{J,J'}^2(X_i)+4\rho c_0 n^{-3/2}N^{-1/2}\log^{1/2} n}\right\}\\
&= 2\exp\left\{-\frac{\rho^2\log n}{4n^2N\,\mathrm{E}\,\xi_{J,J'}^2(X_1)+4\rho c_0 n^{-1/2}N^{1/2}\log^{1/2} n}\right\}\\
&\le n^{-t},
\end{aligned}
$$

for some $t > 2$ by choosing a large enough ρ, which holds since $4n^2N\times \mathrm{E}\,\xi_{J,J'}^2(X_1)$ is bounded by (S1.1) and $n^{-1/2}N^{1/2}\log^{1/2} n \to 0$ by Assumption (A6). Therefore,

$$
\begin{aligned}
P&\left\{\max_{1-p\le J,J'\le N}\left|\sum_{i=1}^{n}\xi_{J,J'}(X_i)\right| > \rho n^{-1/2}N^{-1/2}\log^{1/2} n\right\}\\
&\le \sum_{J,J'=1-p}^{N} P\left\{\left|\sum_{i=1}^{n}\xi_{J,J'}(X_i)\right| > \rho n^{-1/2}N^{-1/2}\log^{1/2} n\right\}\\
&\le (N+p)^2 n^{-t},
\end{aligned}
$$

and hence

$$
\begin{aligned}
\sum_{n=1}^{\infty}P&\left\{\max_{1-p\le J,J'\le N}\left|\sum_{i=1}^{n}\xi_{J,J'}(X_i)\right| > \rho n^{-1/2}N^{-1/2}\log^{1/2} n\right\}\\
&\le \sum_{n=1}^{\infty}(N+p)^2 n^{-t} < \infty.
\end{aligned}
$$

By Borel-Cantelli's Lemma, one immediately obtains that

$$
\max_{1-p\le J,J'\le N}\left|\sum_{i=1}^{n}\xi_{J,J'}(X_i)\right| = O_p\left(n^{-1/2}N^{-1/2}\log^{1/2} n\right), \tag{S1.2}
$$

which concludes that

$$
\left\|\hat{V}_p^* - V_p\right\|_\infty = \max_{1-p\le J\le N}\sum_{J'\in\{J':\,|J'-J|<p\}}\left|\sum_{i=1}^{n}\xi_{J,J'}(X_i)\right| = O_p(n^{-1/2}N^{-1/2}\log^{1/2} n),
$$

completing the proof.

Proof of Lemma S.4(b). According to (S1.2), one has that

$$\begin{aligned}
&\max_{1-p\leq J,J'\leq N}\left|n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}B_{J,p}(X_i)B_{J',p}(X_i)\right| \\
&\quad\leq \max_{1-p\leq J,J'\leq N}\left|\mathrm{E}\{B_{J,p}(X_i)B_{J',p}(X_i)\right| + O_p(n^{-1/2}N^{-1/2}\log^{1/2} n) \\
&\quad= O_p\left(N^{-1}\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
&\left\|\hat{V}_p - \hat{V}_p^*\right\|_\infty \\
&= \max_{1-p\leq J\leq N}\sum_{J'\in\{J':\,|J'-J|<p\}}\left|n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}\left(\frac{\pi_i-\hat{\pi}_i}{\hat{\pi}_i}\right)B_{J,p}(X_i)B_{J',p}(X_i)\right| \\
&= O_p\left(n^{-1/2}\right)(2p-1)\max_{1-p\leq J,J'\leq N}\left|n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}B_{J,p}(X_i)B_{J',p}(X_i)\right| \\
&= O_p\left(n^{-1/2}N^{-1}\right),
\end{aligned}$$

completing the proof.

Proof of Lemma S.4(c). According to Lemma S.3, for any $(N+p)$ length vector $\boldsymbol{\eta}$, $\left\|V_p^{-1}\boldsymbol{\eta}\right\|_\infty \leq K_{p1}N\left\|\boldsymbol{\eta}\right\|_\infty$. Thus one has $\left\|V_p\boldsymbol{\eta}\right\|_\infty \geq K_{p1}^{-1}N^{-1}\left\|\boldsymbol{\eta}\right\|_\infty$. Since $n^{-1/2}N^{1/2}\log^{1/2} n \to 0$ by Assumption (A6) and Lemma S.4(a), one has that

$$\begin{aligned}
\left\|\hat{V}_p^*\boldsymbol{\eta}\right\|_\infty &\geq \left\|V_p\boldsymbol{\eta}\right\|_\infty - \left\|\left(\hat{V}_p^*-V_p\right)\boldsymbol{\eta}\right\|_\infty \\
&\geq K_{p1}^{-1}N^{-1}\left\|\boldsymbol{\eta}\right\|_\infty - O_p(n^{-1/2}N^{-1/2}\log^{1/2} n)\left\|\boldsymbol{\eta}\right\|_\infty \\
&= K_{p1}^{-1}N^{-1}\left\|\boldsymbol{\eta}\right\|_\infty\left(1-O_p(n^{-1/2}N^{1/2}\log^{1/2} n)\right) \\
&\geq K_{p2}^{-1}N^{-1}\left\|\boldsymbol{\eta}\right\|_\infty
\end{aligned}$$

in probability for some constant $K_{p2}>0$. Therefore, $\left\|\hat{V}_p^{*-1}\boldsymbol{\eta}\right\|_\infty \leq K_{p2}N\left\|\boldsymbol{\eta}\right\|_\infty$ which together with Lemma S.4(b) concludes that, for any $(N+p)$ length vector $\boldsymbol{\eta}$,

$$\begin{aligned}\left\|\hat{V}_p\boldsymbol{\eta}\right\|_\infty &\geq \left\|\hat{V}_p^*\boldsymbol{\eta}\right\|_\infty - \left\|\left(\hat{V}_p - \hat{V}_p^*\right)\boldsymbol{\eta}\right\|_\infty \\ &\geq K_{p2}^{-1}N^{-1}\left\|\boldsymbol{\eta}\right\|_\infty - O_p\left(n^{-1/2}N^{-1}\right)\left\|\boldsymbol{\eta}\right\|_\infty \\ &\geq K_{p2}^{-1}N^{-1}\left\|\boldsymbol{\eta}\right\|_\infty(1 - O_p(n^{-1/2})) \\ &\geq K_{p3}^{-1}N^{-1}\left\|\boldsymbol{\eta}\right\|_\infty\end{aligned}$$

for some constant $K_{p3} > 0$ in probability. Hence $\left\|\hat{V}_p^{-1}\boldsymbol{\eta}\right\|_\infty \leq K_{p3}N\left\|\boldsymbol{\eta}\right\|_\infty$ in probability, completing the proof.

For any function $\varphi(x) \in C^{(p)}[a,b]$, let $\boldsymbol{\varphi} = (\varphi(X_1), \ldots, \varphi(X_n))^T$ and denote

$$\tilde{\varphi}_p^*(x) = (B_{1-p,p}(x), \ldots, B_{N,p}(x))\,\hat{V}_p^{*-1}n^{-1}\mathbf{B}^T\boldsymbol{\Delta}\boldsymbol{\varphi}. \tag{S1.3}$$

Lemma S.5. *Under Assumption (A6), there exist constants M_p and $M_{\varphi,p}$ such that any function $\tilde{\varphi}_p^*(x)$ given in (S1.3) satisfies*

$$\left\|\tilde{\varphi}_p^*(x)\right\|_\infty \leq M_p \times \left\|\varphi(x)\right\|_\infty \quad \textit{and} \quad \left\|\tilde{\varphi}_p^*(x) - \varphi(x)\right\|_\infty \leq M_{\varphi,p}N^{-p}$$

in probability.

Proof of Lemma S.5. Let the vector $I_N = (1, \ldots, 1)^T$ with length N. Then by Lemma S.1, similar to the proof of (S1.2) it is easy to show that

$$\begin{aligned}\left\|n^{-1}\mathbf{B}^T\boldsymbol{\Delta}I_N\right\|_\infty &= \max_{1-p\leq J\leq N}\left|n^{-1}\sum_{i=1}^n\frac{\delta_i}{\pi_i}B_{J,p}(X_i)\right| \\ &\leq \max_{1-p\leq J\leq N}\left|n^{-1}\sum_{i=1}^n\frac{\delta_i}{\pi_i}B_{J,p}(X_i) - \mathrm{E}\,B_{J,p}(X_1)\right| \\ &\quad + \max_{1-p\leq J\leq N}\left|\mathrm{E}\,B_{J,p}(X_1)\right| \\ &= O_p(n^{-1/2}N^{-1/2}\log^{-1/2}n) + O(N^{-1}) = O_p(N^{-1}).\end{aligned}$$

This together with Lemma S.4(c) and the fact that at most $(p+1)$ of the numbers $B_{1-p,p}(x), \ldots, B_{N,p}(x)$ are between 0 and 1, others being 0 implies that

$$\begin{aligned}\left\|\tilde{\varphi}_p^*(x)\right\|_\infty &= \left\|\left(B_{1-p,p}(x),\ldots,B_{N,p}(x)\right)\hat{V}_p^{*-1}n^{-1}\mathbf{B}^T\boldsymbol{\Delta}\boldsymbol{\varphi}\right\|_\infty\\ &\leq (p+1)\left\|\hat{V}_p^{*-1}n^{-1}\mathbf{B}^T\boldsymbol{\Delta}\boldsymbol{\varphi}\right\|_\infty\\ &\leq (p+1)\left\|\hat{V}_p^{*-1}\right\|_\infty\left\|n^{-1}\mathbf{B}^T\boldsymbol{\Delta}\boldsymbol{\varphi}\right\|_\infty\\ &\leq (p+1)\left\|\hat{V}_p^{*-1}\right\|_\infty\left\|n^{-1}\mathbf{B}^T\boldsymbol{\Delta}I_N\right\|_\infty\left\|\varphi(x)\right\|_\infty\\ &\leq (p+1)K_{p2}N\left\|n^{-1}\mathbf{B}^T\boldsymbol{\Delta}I_N\right\|_\infty\left\|\varphi(x)\right\|_\infty\\ &\leq (p+1)K_{p2}NCN^{-1}\left\|\varphi(x)\right\|_\infty\equiv M_p\left\|\varphi(x)\right\|_\infty \qquad \text{(S1.4)}\end{aligned}$$

in probability, where C is some positive constant and $M_p=(p+1)K_{p2}C$.

Moreover, according to Lemma S.2, there exists a spline function $m_p(x) \in G_N^{(p-2)}$ such that $m_p(x)\equiv\tilde{m}_p^*(x)$ and $\left\|\varphi(x)-m_p(x)\right\|_\infty\leq C_p\left\|\varphi^{(p)}\right\|_\infty N^{-p}$. Then $\varphi(x)-m_p(x)\in C^{(p)}[a,b]$ and $\tilde{\varphi}_p^*(x)-\tilde{m}_p^*(x)$ can be expressed as in (S1.3) with $\varphi(x)$ replaced by $\varphi(x)-m_p(x)$. Therefore, with this substitution in (S1.4) one has

$$\begin{aligned}\left\|\tilde{\varphi}_p^*(x)-\varphi(x)\right\|_\infty &\leq \left\|\tilde{\varphi}_p^*(x)-\tilde{m}_p^*(x)\right\|_\infty+\left\|m_p(x)-\varphi(x)\right\|_\infty\\ &\leq M_p\left\|\varphi(x)-m_p(x)\right\|_\infty+\left\|m_p(x)-\varphi(x)\right\|_\infty\\ &= (M_p+1)\left\|\varphi(x)-m_p(x)\right\|_\infty\\ &\leq C_p(M_p+1)\left\|\varphi^{(p)}\right\|_\infty N^{-p}\equiv M_{\varphi,p}N^{-p}\end{aligned}$$

in probability, where $M_{\varphi,p}=C_p(M_p+1)\left\|\varphi^{(p)}\right\|_\infty$. The proof is completed.

Lemma S.6. *Under Assumptions (A1)–(A5), for any sequence of measurable functions $s_n(x)$ with $\left\|s_n(x)\right\|_\infty=u_n>0$,*

$$\sup_{x\in[a,b]}\left|n^{-1}\sum_{i=1}^n\frac{\delta_i}{\pi_i}K_h(X_i-x)\,s_n(X_i)\,\varepsilon_i\right|=O_p\left(u_nn^{-1/2}h^{-1/2}\log^{1/2}n\right).$$

Proof of Lemma S.6. Denote $D_n=n^\lambda$ for $1/(2+\eta)<\lambda<1/3$, which together with Assumption (A5) implies

$$D_n^{-(\eta+1)} n^{1/2} h^{1/2} \to 0, \quad \sum_{n=1}^{\infty} D_n^{-(2+\eta)} < \infty, \quad D_n n^{-1/2} h^{-1/2} \log^{1/2} n \to 0.$$

We decompose the noise ε_i as

$$\varepsilon_i = \varepsilon_{i,1} + \varepsilon_{i,2} + \mu_i, \tag{S1.5}$$

where $\varepsilon_{i,1} = \varepsilon_i I\left(|\varepsilon_i| > D_n\right)$, $\mu_i = \mathrm{E}\left\{\varepsilon_i I\left(|\varepsilon_i| \le D_n\right) | X_i\right\}$ and $\varepsilon_{i,2} = \varepsilon_i \times I\left(|\varepsilon_i| \le D_n\right) - \mu_i$, in which $I\left(\cdot\right)$ is the indicator function.

Firstly, note that

$$\begin{aligned} |\mu_i| &= \left|-\mathrm{E}\left\{\varepsilon_i I\left(|\varepsilon_i| > D_n\right) | X_i\right\}\right| \\ &\le D_n^{-(\eta+1)} \mathrm{E}\left\{|\varepsilon_i|^{2+\eta} | X_i\right\} \\ &\le D_n^{-(\eta+1)} C_\eta \end{aligned}$$

in probability, where C_η is the upper bound of $\mathrm{E}(|\varepsilon_i|^{2+\eta} | X_i)$ by Assumption (A2). Meanwhile, according to Lemma 3 in Cai et al. (2021), one has that

$$\sup_{x\in[a,b]} \left| n^{-1} \sum_{i=1}^{n} \frac{\delta_i}{\pi_i} K_h\left(X_i - x\right) \right| = \sup_{x\in[a,b]} \left| \tilde{f}_X(x) \right| = O_p\left(1\right). \tag{S1.6}$$

Thus,

$$\begin{aligned} &\sup_{x\in[a,b]} \left| n^{-1} \sum_{i=1}^{n} \frac{\delta_i}{\pi_i} K_h\left(X_i - x\right) s_n\left(X_i\right) \mu_i \right| \\ &\le u_n D_n^{-(\eta+1)} C_\eta \sup_{x\in[a,b]} \left| \tilde{f}_X(x) \right| \\ &= O_p\left(u_n D_n^{-(\eta+1)}\right) = o_p\left(u_n n^{-1/2} h^{-1/2}\right). \end{aligned} \tag{S1.7}$$

Next, since

$$\sum_{n=1}^{\infty} P\left\{I\left(|\varepsilon_n| > D_n\right)\right\} \le C_\eta \sum_{n=1}^{\infty} D_n^{-(2+\eta)} < \infty,$$

Borel-Cantelli's Lemma implies that

$$P\{\omega\colon \text{there exists } N_0(\omega)>0 \text{ such that } |\varepsilon_n(\omega)|\le D_n \text{ for } n>N_0(\omega)\}=1.$$

Therefore,

$$P\{\omega\colon \text{there exists } N(\omega) \text{ such that } |\varepsilon_i(\omega)|\le D_n \text{ for } 1\le i\le n,\ n>N(\omega)\}=1,$$

which concludes that

$$P\{\omega\colon \text{there exists } N(\omega) \text{ such that } \varepsilon_{i,1}=0 \text{ for } 1\le i\le n,\ n>N(\omega)\}=1.$$

Hence, for any $\gamma>0$

$$\sup_{x\in[a,b]}|n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h(X_i-x)\,s_n(X_i)\,\varepsilon_{i,1}|=O_{a.s.}(n^{-\gamma}), \tag{S1.8}$$

where a.s. stands for almost surely.

Finally, we deal with the truncated part $\varepsilon_{i,2}$. It is easy to see that $\mathrm{E}(\varepsilon_{i,2}|X_i)=0$ and

$$\begin{aligned}\mathrm{E}\left(\varepsilon_{i,2}^2|X_i\right) &= \mathrm{E}\left\{\varepsilon_i^2 I(|\varepsilon_i|\le D_n)|X_i\right\}-\mu_i^2\\ &= \sigma^2(X_i)-\mathrm{E}\left\{\varepsilon_i^2 I(|\varepsilon_i|>D_n)|X_i\right\}-\mu_i^2\\ &= \sigma^2(X_i)+O_p\left(D_n^{-\eta}+D_n^{-2(\eta+1)}\right).\end{aligned}$$

For convenience, denote $\xi_{in}(x)=n^{-1}\frac{\delta_i}{\pi_i}K_h(X_i-x)\,s_n(X_i)\,\varepsilon_{i,2}$. One then gets that $\mathrm{E}\,\xi_{in}(x)=0$ and

$$\begin{aligned}\mathrm{E}\,\xi_{in}^2(x) &= n^{-2}\,\mathrm{E}\left\{\frac{\delta_i^2}{\pi_i^2}K_h^2(X_i-x)\,s_n^2(X_i)\,\varepsilon_{i,2}^2\right\}\\ &= n^{-2}\,\mathrm{E}\left\{\frac{1}{\pi_i}K_h^2(X_i-x)\,s_n^2(X_i)\,\varepsilon_{i,2}^2\right\}\\ &\le c_\pi^{-1}n^{-2}\,\mathrm{E}\left\{K_h^2(X_i-x)\,s_n^2(X_i)\,\varepsilon_{i,2}^2\right\}\\ &\le c_\pi^{-1}u_n^2n^{-2}\,\mathrm{E}\left\{K_h^2(X_i-x)\,\varepsilon_{i,2}^2\right\}\\ &= c_\pi^{-1}u_n^2n^{-2}h^{-1}\sigma^2(x)\,f_X(x)\int K^2(v)\,dv\,\{1+u(1)\}.\end{aligned}$$

The k-th moment $\mathrm{E}\left|\xi_{in}(x)\right|^k$ for $k \geq 3$ is bounded as follows:

$$\begin{aligned}
\mathrm{E}\left|\xi_{in}(x)\right|^k &= \mathrm{E}\left\{\left|\xi_{in}(x)\right|^{k-2}\left(\xi_{in}(x)\right)^2\right\} \\
&\leq n^{-(k-2)} c_\pi^{-(k-2)} h^{-(k-2)} \left\|K\right\|_\infty^{k-2} u_n^{k-2} \left(2D_n\right)^{k-2} \mathrm{E}\left(\xi_{in}(x)\right)^2 \\
&= \left(q u_n n^{-1} h^{-1} D_n\right)^{k-2} \mathrm{E}\left(\xi_{in}(x)\right)^2 \leq \left(q u_n n^{-1} h^{-1} D_n\right)^{k-2} k!\, \mathrm{E}\left(\xi_{in}(x)\right)^2,
\end{aligned}$$

where $q = 2c_\pi^{-1}\left\|K\right\|_\infty$. Thus $\xi_{in}(x), 1 \leq i \leq n$, fulfill Cramér's Conditions in Lemma S.1 with $r = q u_n n^{-1} h^{-1} D_n$. Then for large n, one has that

$$\begin{aligned}
& P\left\{\left|\sum_{i=1}^n \xi_{in}(x)\right| > \rho u_n n^{-1/2} h^{-1/2} \log^{1/2} n\right\} \\
&\leq 2\exp\left\{-\frac{\rho^2 u_n^2 n^{-1} h^{-1} \log n}{4\sum_{i=1}^n \mathrm{E}\,\xi_{in}^2(x) + 2\rho q n^{-1} h^{-1} \mu_n D_n n^{-1/2} h^{-1/2} u_n \log^{1/2} n}\right\} \\
&= 2\exp\left\{-\frac{\rho^2 \log n}{4u_n^{-2} n^2 h\, \mathrm{E}\,\xi_{1n}^2(x) + 2\rho q D_n n^{-1/2} h^{-1/2} \log^{1/2} n}\right\} \\
&\leq 2n^{-4}
\end{aligned}$$

by choosing large ρ which holds since $4nhu_n^{-2}\,\mathrm{E}\,\xi_{1n}^2(x)$ is bounded and $D_n n^{-1/2} \times\ h^{-1/2} \log^{1/2} n \to 0$. To bound the truncated part uniformly for all $x \in [a, b]$, we discretize $[a, b]$ by equally spaced points $a = x_0 < x_1 < \cdots < x_{M_n} = b$ with $M_n = n^2$. One then gets

$$\begin{aligned}
& P\left\{\left|\max_{j=0}^{M_n} \sum_{i=1}^n \xi_{in}(x_j)\right| > \rho u_n n^{-1/2} h^{-1/2} \log^{1/2} n\right\} \\
&\leq \sum_{j=0}^{M_n} P\left\{\left|\sum_{i=1}^n \xi_{in}(x_j)\right| > \rho u_n n^{-1/2} h^{-1/2} \log^{1/2} n\right\} \\
&\leq 2n^{-4}(M_n + 1).
\end{aligned}$$

Therefore,

$$\begin{aligned}
\sum_{n=1}^\infty P\left\{\left|\max_{j=0}^{M_n} \sum_{i=1}^n \xi_{in}(x_j)\right| > \rho u_n n^{-1/2} h^{-1/2} \log^{1/2} n\right\} &\leq \sum_{n=1}^\infty 2n^{-4}(M_n + 1) \\
&< \infty.
\end{aligned}$$

Borel-Cantelli's Lemma implies

$$\max_{j=0}^{M_n}\left|\sum_{i=1}^{n}\xi_{in}\left(x_j\right)\right| = O_{a.s.}\left(u_n n^{-1/2}h^{-1/2}\log^{1/2} n\right).$$

Notice that

$$\begin{aligned}
&\max_{x\in[a,b]}\left|\sum_{i=1}^{n}\xi_{in}\left(x\right)\right| \\
&\leq \max_{j=0}^{M_n}\left|\sum_{i=1}^{n}\xi_{in}\left(x_j\right)\right| + \max_{j=0}^{(M_n-1)}\sup_{x\in[x_j,x_{j+1}]}\left|\sum_{i=1}^{n}\xi_{in}\left(x\right)-\sum_{i=1}^{n}\xi_{in}\left(x_j\right)\right| \\
&= \max_{j=0}^{M_n}\left|\sum_{i=1}^{n}\xi_{in}\left(x_j\right)\right| + \\
&\max_{j=0}^{(M_n-1)}\sup_{x\in[x_j,x_{j+1}]}\left|\sum_{i=1}^{n}n^{-1}\frac{\delta_i}{\pi_i}\left\{K_h\left(X_i-x\right)-K_h\left(X_i-x_j\right)\right\}s_n\left(X_i\right)\varepsilon_{i,2}\right| \\
&\leq \max_{j=0}^{M_n}\left|\sum_{i=1}^{n}\xi_{in}\left(x_j\right)\right| + 2c_\pi^{-1}\left\|K^{(1)}\right\|_\infty h^{-2}u_nD_n\max_{j=0}^{(M_n-1)}\sup_{x\in[x_j,x_{j+1}]}\left|x-x_j\right| \\
&\leq O_{a.s.}\left(u_n n^{-1/2}h^{-1/2}\log^{1/2} n\right) + 2c_\pi^{-1}\left\|K^{(1)}\right\|_\infty h^{-2}u_nD_n(b-a)M_n^{-1} \\
&= O_{a.s.}\left(u_n n^{-1/2}h^{-1/2}\log^{1/2} n\right). \qquad \text{(S1.9)}
\end{aligned}$$

Thus, (S1.5), (S1.7), (S1.8), and (S1.9) imply the result.

Lemma S.7. *There exist positive constants c and C independent of n such that*

$$cN^{-1}\sum_{J=1-p}^{N}\alpha_J^2 \leq \left\|\sum_{J=1-p}^{N}\alpha_J B_{J,p}(x)\right\|_2^2 \leq CN^{-1}\sum_{J=1-p}^{N}\alpha_J^2.$$

This lemma is adapted from Lemma A.5 of Wang and Yang (2007).

Lemma S.8. *Under Assumptions (A1) and (A6), as $n\to\infty$,*

$$\Upsilon_n = \sup_{m_1(\cdot),m_2(\cdot)\in G_N^{(p-2)}}\left|\frac{\langle m_1,m_2\rangle_{2,n}-\langle m_1,m_2\rangle_2}{\|m_1\|_2\|m_2\|_2}\right| = O_p(n^{-1/2}N^{1/2}\log^{1/2} n).$$

The proof is similar to that of Lemma A.2 in Song and Yang (2009) by applying Lemma S.1 and is omitted here.

S2 Proofs of Proposition 1 and Theorem 2

Proof of Proposition 1(a). Note that by the definition of $\tilde{\varepsilon}_p^*(x)$ given in (3.4), one has

$$\tilde{\varepsilon}_p^*(x) = (B_{1-p,p}(x), \ldots, B_{N,p}(x))\, \hat{V}_p^{*-1} n^{-1} \mathbf{B}^T \mathbf{\Delta E}.$$

Applying Lemma S.1, Borel-Cantelli's Lemma and the truncation techniques again as in the proof of Lemma S.6, one has that

$$\begin{aligned} \left\| n^{-1}\mathbf{B}^T \mathbf{\Delta E} \right\|_\infty &= \left\| \left(n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi_i} B_{J,p}(X_i)\varepsilon_i \right)_{J=1-p}^N \right\|_\infty \\ &= O_p(n^{-1/2} N^{-1/2} \log^{1/2} n). \end{aligned}$$

Thus,

$$\begin{aligned} \left\| \tilde{\varepsilon}_p^*(x) \right\|_\infty &= \left\| (B_{1-p,p}(x), \ldots, B_{N,p}(x))\, \hat{V}_p^{*-1} n^{-1} \mathbf{B}^T \mathbf{\Delta E} \right\|_\infty \\ &\leq (p+1) \left\| \hat{V}_p^{*-1} \right\|_\infty \left\| n^{-1}\mathbf{B}^T \mathbf{\Delta E} \right\|_\infty \\ &\leq (p+1) K_{p2} N \left\| n^{-1}\mathbf{B}^T \mathbf{\Delta E} \right\|_\infty \\ &= O_p(n^{-1/2} N^{1/2} \log^{1/2} n). \end{aligned} \tag{S2.1}$$

Moreover, according to Lemma S.5, there exists a constant $M_{g,p}$ such that $\tilde{g}_p^*(x)$ given in (3.4) satisfies

$$\left\| \tilde{g}_p^*(x) - g(x) \right\|_\infty \leq M_{g,p} N^{-p}. \tag{S2.2}$$

Therefore,

$$\begin{aligned} \sup_{x\in[a,b]} |I_1(x)| &\leq \sup_{x\in[a,b]} \left| 2\tilde{f}_X^{-1}(x)\, n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi_i} K_h(X_i - x) \left(g(X_i) - \tilde{g}_p^*(X_i) \right)^2 \right| \\ &\quad + \sup_{x\in[a,b]} \left| 2\tilde{f}_X^{-1}(x)\, n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi_i} K_h(X_i - x)\, \tilde{\varepsilon}_p^{*2}(X_i) \right| \\ &\leq 2 \left\| g(x) - \tilde{g}_p^*(x) \right\|_\infty^2 + 2 \left\| \tilde{\varepsilon}_p^*(x) \right\|_\infty^2 \\ &= O_p(N^{-2p} + n^{-1} N \log n). \end{aligned}$$

We next deal with the second term in Proposition 1(a). Notice that

$$\sup_{x\in[a,b]}|J_1(x)| \leq \sup_{x\in[a,b]}\left|n^{-1}\hat{f}_X^{-1}(x)\sum_{i=1}^{n}\frac{\delta_i}{\hat{\pi}_i}K_h\left(X_i-x\right)\left(\hat{g}_p^*(X_i)-\hat{g}_p(X_i)\right)^2\right|$$
$$+\sup_{x\in[a,b]}\left|2n^{-1}\hat{f}_X^{-1}(x)\sum_{i=1}^{n}\frac{\delta_i}{\hat{\pi}_i}K_h\left(X_i-x\right)\left(\hat{g}_p^*(X_i)-\hat{g}_p(X_i)\right)\left(Y_i-\hat{g}_p^*(X_i)\right)\right|. \tag{S2.3}$$

On the one hand, by (2.5) and (3.3), one has that

$$\begin{aligned}\hat{g}_p^*\left(x\right) &= \left(B_{1-p,p}\left(x\right),\ldots,B_{N,p}\left(x\right)\right)\hat{V}_p^{*-1}n^{-1}\mathbf{B}^T\mathbf{\Delta}\mathbf{Y},\\ \hat{g}_p\left(x\right) &= \left(B_{1-p,p}\left(x\right),\ldots,B_{N,p}\left(x\right)\right)\hat{V}_p^{-1}n^{-1}\mathbf{B}^T\hat{\mathbf{\Delta}}\mathbf{Y}.\end{aligned}$$

Applying Lemma S.1 again, similar to the proof of Lemma S.6, one has that

$$\max_{1-p\leq J\leq N}n^{-1}\sum_{i=1}^{n}B_{J,p}(X_i)\left|g\left(X_i\right)\right| = O_p\left(N^{-1}\right),$$

and

$$\max_{1-p\leq J\leq N}n^{-1}\sum_{i=1}^{n}B_{J,p}(X_i)\left|\varepsilon_i\right| = O_p\left(N^{-1}\right),$$

which imply that

$$\begin{aligned}\max_{1-p\leq J\leq N}n^{-1}\sum_{i=1}^{n}\left|B_{J,p}(X_i)Y_i\right| &\leq \max_{1-p\leq J\leq N}n^{-1}\sum_{i=1}^{n}B_{J,p}(X_i)\left|g\left(X_i\right)\right|\\ &\quad+\max_{1-p\leq J\leq N}n^{-1}\sum_{i=1}^{n}B_{J,p}(X_i)\left|\varepsilon_i\right|\\ &= O_p\left(N^{-1}\right).\end{aligned}$$

Then one has that

$$\left\|n^{-1}\mathbf{B}^T\mathbf{\Delta}\mathbf{Y}\right\|_\infty = \max_{1-p\leq J\leq N}\left|n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}B_{J,p}(X_i)Y_i\right|$$
$$\leq \left\|\pi^{-1}(y)\right\|_\infty \max_{1-p\leq J\leq N}n^{-1}\sum_{i=1}^{n}\left|B_{J,p}(X_i)Y_i\right| = O_p\left(N^{-1}\right), \tag{S2.4}$$

and

$$\left\| n^{-1}\mathbf{B}^T\mathbf{\Delta}\mathbf{Y} - n^{-1}\mathbf{B}^T\hat{\mathbf{\Delta}}\mathbf{Y} \right\|_\infty = \max_{1-p\le J\le N} \left| n^{-1}\sum_{i=1}^n \delta_i \left(\frac{\hat{\pi}_i - \pi_i}{\hat{\pi}_i\pi_i}\right) B_{J,p}(X_i)Y_i \right|$$
$$\le O_p\left(n^{-1/2}\right) \max_{1-p\le J\le N} n^{-1}\sum_{i=1}^n \left|B_{J,p}(X_i)Y_i\right| = O_p\left(n^{-1/2}N^{-1}\right). \qquad \text{(S2.5)}$$

By (S2.4), (S2.5), and Lemma S.4(b), one has that

$$\begin{aligned}
&\left\|\hat{g}_p^*(x) - \hat{g}_p(x)\right\|_\infty \\
&= \Big\| \left(B_{1-p,p}(x), \ldots, B_{N,p}(x)\right)\left(\hat{V}_p^{*-1} - \hat{V}_p^{-1}\right) n^{-1}\mathbf{B}^T\mathbf{\Delta}\mathbf{Y} \\
&\qquad + \left(B_{1-p,p}(x), \ldots, B_{N,p}(x)\right)\hat{V}_p^{-1}\left(n^{-1}\mathbf{B}^T\mathbf{\Delta}\mathbf{Y} - n^{-1}\mathbf{B}^T\hat{\mathbf{\Delta}}\mathbf{Y}\right)\Big\|_\infty \\
&= \Big\| -\left(B_{1-p,p}(x), \ldots, B_{N,p}(x)\right)\hat{V}_p^{*^{-1}}\left(\hat{V}_p^* - \hat{V}_p\right)\hat{V}_p^{-1} n^{-1}\mathbf{B}^T\mathbf{\Delta}\mathbf{Y} \\
&\qquad + \left(B_{1-p,p}(x), \ldots, B_{N,p}(x)\right)\hat{V}_p^{-1}\left(n^{-1}\mathbf{B}^T\mathbf{\Delta}\mathbf{Y} - n^{-1}\mathbf{B}^T\hat{\mathbf{\Delta}}\mathbf{Y}\right)\Big\|_\infty \\
&\le (p+1)\left\|\hat{V}_p^{*^{-1}}\right\|_\infty \left\|\hat{V}_p^* - \hat{V}_p\right\|_\infty \left\|\hat{V}_p^{-1}\right\|_\infty \left\|n^{-1}\mathbf{B}^T\mathbf{\Delta}\mathbf{Y}\right\|_\infty \\
&\qquad + (p+1)\left\|\hat{V}_p^{-1}\right\|_\infty \left\|n^{-1}\mathbf{B}^T\mathbf{\Delta}\mathbf{Y} - n^{-1}\mathbf{B}^T\hat{\mathbf{\Delta}}\mathbf{Y}\right\|_\infty \\
&= O_p\left(n^{-1/2}\right),
\end{aligned}$$

which concludes that

$$\sup_{x\in[a,b]} \left| n^{-1}\hat{f}_X^{-1}(x)\sum_{i=1}^n \frac{\delta_i}{\hat{\pi}_i} K_h\left(X_i - x\right)\left(\hat{g}_p^*(X_i) - \hat{g}_p(X_i)\right)^2 \right|$$
$$\le \sup_{x\in[a,b]} \left\|\hat{g}_p^*(x) - \hat{g}_p(x)\right\|_\infty^2 = O_p(n^{-1}). \qquad \text{(S2.6)}$$

On the other hand, applying Lemma S.1 again similar to the proof of Lemma S.6, it is easy to show that

$$\sup_{x\in[a,b]} \left| n^{-1}\sum_{i=1}^n \frac{\delta_i}{\pi_i} K_h\left(X_i - x\right)\left|\varepsilon_i\right| \right| = O_p(1).$$

Therefore,

$$
\begin{aligned}
&\sup_{x\in[a,b]}\left|2n^{-1}\hat{f}_X^{-1}(x)\sum_{i=1}^{n}\frac{\delta_i}{\hat{\pi}_i}K_h\left(X_i-x\right)\left(\hat{g}_p^*(X_i)-\hat{g}_p(X_i)\right)\left(Y_i-\hat{g}_p^*(X_i)\right)\right| \\
&\leq \sup_{x\in[a,b]} 2\left\|\hat{g}_p^*(x)-\hat{g}_p(x)\right\|_\infty \times \\
&\quad \sup_{x\in[a,b]}\left|n^{-1}\hat{f}_X^{-1}(x)\sum_{i=1}^{n}\frac{\delta_i}{\hat{\pi}_i}K_h\left(X_i-x\right)\left\{\left|g(X_i)-\tilde{g}_p^*(X_i)\right|+\left|\varepsilon_i-\tilde{\varepsilon}_p^*(X_i)\right|\right\}\right| \\
&\leq 2\left\|\hat{g}_p^*(x)-\hat{g}_p(x)\right\|_\infty\left\|g(x)-\tilde{g}_p^*(x)\right\|_\infty \\
&\quad + 2\left\|\hat{g}_p^*(x)-\hat{g}_p(x)\right\|_\infty \sup_{x\in[a,b]}\left|\hat{f}_X^{-1}(x)n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\hat{\pi}_i}K_h\left(X_i-x\right)\left|\varepsilon_i\right|\right| \\
&\quad + 2\left\|\hat{g}_p^*(x)-\hat{g}_p(x)\right\|_\infty\left\|\tilde{\varepsilon}_p^*(x)\right\|_\infty \\
&= O_p(n^{-1/2}N^{-p}+n^{-1/2}+n^{-1/2}n^{-1/2}N^{1/2}\log^{1/2}n) = O_p(n^{-1/2}). \qquad \text{(S2.7)}
\end{aligned}
$$

By (S2.3), (S2.6), and (S2.7), one obtains that

$$\sup_{x\in[a,b]}|J_1(x)| = O_p(n^{-1/2}).$$

Proof of Proposition 1(b). According to Lemma A.5 in Cai et al. (2021), one has that

$$\left\|\tilde{f}_X(x)-\hat{f}_X(x)\right\|_\infty = O_p(n^{-1/2}),$$

which with (S1.6) and Assumption (A1) concludes that

$$\left\|\tilde{f}_X^{-1}(x)\right\|_\infty = O_p(1),\ \left\|\hat{f}_X^{-1}(x)\right\|_\infty = O_p(1),\ \left\|\tilde{f}_X^{-1}(x)-\hat{f}_X^{-1}(x)\right\|_\infty = O_p(n^{-1/2}). \qquad \text{(S2.8)}$$

Moreover, by Lemma S.2, for $g\left(x\right) \in C^{(p)}\left[a,b\right]$, there exist a constant C_p and a spline function $m_p(x) \in G_N^{(p-2)}$ such that $\left\|m_p(x)-g\left(x\right)\right\|_\infty \leq$

$C_p \left\| g^{(p)}(x) \right\|_\infty N^{-p}$. One then has that

$$\begin{aligned}
&\sup_{x\in[a,b]} |I_2(x)| \\
&\quad = \sup_{x\in[a,b]} \left| 2\tilde{f}_X^{-1}(x) n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi_i} K_h(X_i - x) \left(g(X_i) - \tilde{g}_p^*(X_i) \right) \varepsilon_i \right| \\
&\quad \le \sup_{x\in[a,b]} \left| 2\tilde{f}_X^{-1}(x) n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi_i} K_h(X_i - x) \left(g(X_i) - m_p(X_i) \right) \varepsilon_i \right| \\
&\qquad + \sup_{x\in[a,b]} \left| 2\tilde{f}_X^{-1}(x) n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi_i} K_h(X_i - x) \left(m_p(X_i) - \tilde{g}_p^*(X_i) \right) \varepsilon_i \right|.
\end{aligned} \tag{S2.9}$$

Applying Lemma S.6 with $s_n(x) = g(x) - m_p(x)$ and $u_n = O_p(N^{-p})$, one has that

$$\begin{aligned}
&\sup_{x\in[a,b]} \left| 2\tilde{f}_X^{-1}(x) n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi_i} K_h(X_i - x) \left(g(X_i) - m_p(X_i) \right) \varepsilon_i \right| \\
&\qquad = O_p(n^{-1/2} h^{-1/2} N^{-p} \log^{1/2} n).
\end{aligned} \tag{S2.10}$$

Meanwhile, since both $m_p(x)$ and $\tilde{g}_p^*(x)$ belong to the spline space G_N^{p-2}, one can write $m_p(x) - \tilde{g}_p^*(x) = \sum_{J=1-p}^N \theta_{J,p} B_{J,p}(x)$. By Lemmas S.7, S.8 and (S2.2), there exists a constant $c > 0$ such that

$$\begin{aligned}
& cN^{-1} \sum_{J=1-p}^N \theta_{J,p}^2 \le \| m_p - \tilde{g}^* \|_2^2 \le (1 - \Upsilon_n)^{-1} \| m_p - \tilde{g}^* \|_{2,n}^2 \\
\le\ & (1 - \Upsilon_n)^{-1} \left\| \pi^{-1}(y) \right\|_\infty \| m_p(x) - \tilde{g}^*(x) \|_\infty^2 \\
\le\ & (1 - \Upsilon_n)^{-1} \left\| \pi^{-1}(y) \right\|_\infty \left\{ \| m_p(x) - g(x) \|_\infty^2 + \| g(x) - \tilde{g}^*(x) \|_\infty^2 \right\} \\
=\ & O_p(N^{-2p}).
\end{aligned}$$

Moreover, applying Lemma S.6 again with $s_n(x) = B_{J,p}(x)$ and $u_n = O_p(1)$, one has

$$\sup_{x\in[a,b]} \left| n^{-1} \sum_{i=1}^n \frac{\delta_i}{\pi_i} K_h(X_i - x) B_{J,p}(X_i) \varepsilon_i \right| = O_p(n^{-1/2} h^{-1/2} \log^{1/2} n). \tag{S2.11}$$

Therefore,

$$\begin{aligned}
&\sup_{x\in[a,b]}\left|2\tilde{f}_X^{-1}(x)\,n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h(X_i-x)\left(m_p(X_i)-\tilde{g}_p^*(X_i)\right)\varepsilon_i\right| \\
&=\sup_{x\in[a,b]}\left|2\tilde{f}_X^{-1}(x)\,n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h(X_i-x)\sum_{J=1-p}^{N}\theta_{J,p}B_{J,p}(X_i)\varepsilon_i\right| \\
&\le\sup_{x\in[a,b]}\left|2\tilde{f}_X^{-1}(x)\left\{\sum_{J=1-p}^{N}\theta_{J,p}^2\right\}^{1/2}\times\right. \\
&\qquad\left.\left[\sum_{J=1-p}^{N}\left\{n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h(X_i-x)B_{J,p}(X_i)\varepsilon_i\right\}^2\right]^{1/2}\right| \\
&=O_p(n^{-1/2}h^{-1/2}N^{1-p}\log^{1/2}n). \qquad\qquad \text{(S2.12)}
\end{aligned}$$

Putting (S2.9), (S2.10), and (S2.12) together, one concludes that

$$\sup_{x\in[a,b]}|I_2(x)| \;=\; O_p(n^{-1/2}h^{-1/2}N^{1-p}\log^{1/2}n).$$

Next, by Lemma S.1 it is easy to show that

$$\sup_{x\in[a,b]}\left|n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h(X_i-x)\,\varepsilon_i^2\right|=O_p(1).$$

One then obtains that

$$\begin{aligned}
&\sup_{x\in[a,b]}\left|n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h(X_i-x)\,\hat{R}_i^*\right| \\
&\le\sup_{x\in[a,b]}\left|2n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h(X_i-x)\,(g(X_i)-\tilde{g}_p^*(X_i))^2\right| \\
&\quad+\sup_{x\in[a,b]}\left|2n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h(X_i-x)\,(\varepsilon_i-\tilde{\varepsilon}_p^*(X_i))^2\right| \\
&\le 2\left\|g(x)-\tilde{g}_p^*(x)\right\|_\infty^2\sup_{x\in[a,b]}\left|\tilde{f}_X^{-1}(x)\right|
\end{aligned}$$

$$+\sup_{x\in[a,b]}\left|4n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h\left(X_i-x\right)\varepsilon_i^2\right|+4\left\|\tilde{\varepsilon}_p^*(x)\right\|_\infty^2\sup_{x\in[a,b]}\left|\tilde{f}_X^{-1}\left(x\right)\right|$$
$$=O_p(1), \tag{S2.13}$$

which concludes that

$$\begin{aligned}
&\sup_{x\in[a,b]}|J_2(x)| \\
&\quad=\sup_{x\in[a,b]}\left|n^{-1}\hat{f}_X^{-1}(x)\sum_{i=1}^{n}\left(\frac{\delta_i(\pi_i-\hat{\pi}_i)}{\hat{\pi}_i\pi_i}\right)K_h\left(X_i-x\right)\hat{R}_i^*\right| \\
&\quad\le\left\|\pi(y)-\hat{\pi}(y)\right\|_\infty\left\|\hat{\pi}^{-1}(y)\right\|_\infty\sup_{x\in[a,b]}\left|n^{-1}\hat{f}_X^{-1}(x)\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h\left(X_i-x\right)\hat{R}_i^*\right| \\
&\quad=O_p(n^{-1/2}).
\end{aligned}$$

Proof of Proposition 1(c). By (3.4), one has $\tilde{\varepsilon}_p^*(x)=\sum_{J=1-p}^{N}\tilde{\beta}_{J,p}^*B_{J,p}(x)\in G_N^{(p-2)}$. Lemmas S.7, S.8 and (S2.1) imply that there exists a constant $t>0$ such that

$$\begin{aligned}
tN^{-1}\sum\nolimits_{J=1-p}^{N}\tilde{\beta}_{J,p}^{*2} &\le \left\|\tilde{\varepsilon}_p^*\right\|_2^2\le(1-\Upsilon_n)^{-1}\left\|\tilde{\varepsilon}_p^*\right\|_{2,n}^2 \\
&\le (1-\Upsilon_n)^{-1}\left\|\pi^{-1}(y)\right\|_\infty\left\|\tilde{\varepsilon}_p^*(x)\right\|_\infty^2=O_p(n^{-1}N\log n).
\end{aligned}$$

Hence, $\sum_{J=1-p}^{N}\tilde{\beta}_{J,p}^{*2}=O_p(n^{-1}N^2\log n)$, which together with (S2.11) implies that

$$\begin{aligned}
\sup_{x\in[a,b]}|I_3(x)| &= \sup_{x\in[a,b]}\left|-2\tilde{f}_X^{-1}(x)n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h\left(X_i-x\right)\varepsilon_i\sum\nolimits_{J=1-p}^{N}\tilde{\beta}_{J,p}^*B_{J,p}(X_i)\right| \\
&\le\sup_{x\in[a,b]}\left|2\tilde{f}_X^{-1}(x)\right|\left\{\sum_{J=1-p}^{N}\tilde{\beta}_{J,p}^{*2}\right\}^{1/2}\times \\
&\qquad\sup_{x\in[a,b]}\left[\sum_{J=1-p}^{N}\left(n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h\left(X_i-x\right)B_{J,p}(X_i)\varepsilon_i\right)^2\right]^{1/2}
\end{aligned}$$

$$= O_p(n^{-1/2}N\log^{1/2}n)\times(N+p)^{1/2}\times O_p(n^{-1/2}h^{-1/2}\log^{1/2}n)$$
$$= O_p(n^{-1}h^{-1/2}N^{3/2}\log n).$$

Next, by (S2.8) and (S2.13), one has

$$\sup_{x\in[a,b]}|J_3(x)| \leq \sup_{x\in[a,b]}\left|\hat{f}_X^{-1}(x)-\tilde{f}_X^{-1}(x)\right|\sup_{x\in[a,b]}\left|n^{-1}\sum_{i=1}^{n}\frac{\delta_i}{\pi_i}K_h\left(X_i-x\right)\hat{R}_i^*\right|$$
$$= O_p(n^{-1/2}).$$

The proof is completed.

Proof of Theorem 2. According to (3.5) and Proposition 1, one has that

$$\sup_{x\in[a,b]}\left|\hat{\sigma}^{*2}_{\mathrm{SNW}}(x)-\tilde{\sigma}^2_{\mathrm{NW}}(x)\right| \leq \sup_{x\in[a,b]}|I_1(x)|+\sup_{x\in[a,b]}|I_2(x)|+\sup_{x\in[a,b]}|I_3(x)|$$
$$= O_p(N^{-2p}+n^{-1}N\log n+n^{-1/2}h^{-1/2}N^{1-p}\log^{1/2}n+n^{-1}h^{-1/2}N^{3/2}\log n).$$

Since $n^{1/(4p)} \ll N$ and $N \ll n^{1/2}\log^{-1}n$ in Assumption (A6), one gets that

$$N^{-2p} \ll n^{-1/2}\text{and } n^{-1}N\log n \ll n^{-1/2}.$$

Furthermore, by $h^{-1/2(p-1)}\log^{1/2(p-1)}n \ll N$ and $N \ll n^{1/3}h^{1/3}\log^{-2/3}n$ in Assumption (A6), one obtains that

$$n^{-1/2}h^{-1/2}N^{1-p}\log^{1/2}n \ll n^{-1/2}, n^{-1}h^{-1/2}N^{3/2}\log n \ll n^{-1/2}.$$

Therefore,

$$\sup_{x\in[a,b]}\left|\hat{\sigma}^{*2}_{\mathrm{SNW}}(x)-\tilde{\sigma}^2_{\mathrm{NW}}(x)\right| = O_p(n^{-1/2}).$$

Next, according to (3.6) and Proposition 1, one obtains that

$$\sup_{x\in[a,b]}\left|\hat{\sigma}^{2}_{\mathrm{SNW}}(x)-\hat{\sigma}^{*2}_{\mathrm{SNW}}(x)\right| \leq \sup_{x\in[a,b]}|J_1(x)|+\sup_{x\in[a,b]}|J_2(x)|+\sup_{x\in[a,b]}|J_3(x)|$$
$$= O_p(n^{-1/2}).$$

Therefore,

$$\sup_{x\in[a,b]}\left|\hat{\sigma}^{2}_{\mathrm{SNW}}(x)-\tilde{\sigma}^2_{\mathrm{NW}}(x)\right| = O_p(n^{-1/2}),$$

completing the proof.

References

Bosq, D. (1998). *Nonparametric Statistics for Stochastic Processes.* Springer-Verlag, New York.

Cai, L., Gu, L., Wang, Q. and Wang, S. (2021). Simultaneous confidence bands for nonparametric regression with missing covariate data. *Annals of the Institute of Statistical Mathematics* accepted.

Cao, G., Yang, L. and Todem, D. (2012). Simultaneous inference for the mean function based on dense functional data. *Journal of Nonparametric Statistics* **24**, 359–377.

de Boor, C. (2001). *A Practical Guide to Splines.* Springer, New York.

Song, Q. and Yang, L. (2009). Spline confidence bands for variance functions. *Journal of Nonparametric Statistics* **5**, 589–609.

Wang, L. and Yang, L. (2007). Spline-backfitted kernel smoothing of nonlinear additive autoregression model. *Annals of Statistics* **35**, 2474–2503.