SUPPLEMENT TO "POLYNOMIAL SPLINE CONFIDENCE BANDS FOR REGRESSION CURVES"

Jing Wang and Lijian Yang

University of Illinois at Chicago, Michigan State University

Appendix A: Proof of Theorem 1

A. 1. Preliminaries

Throughout Appendices A and B, we denote by the same letters c, C, any positive constants, without distinction in each case. Detailed proof is given in Wang and Yang (2006).

Lemma A.1 *Under Assumptions (A3) and (A4), there exists $\alpha_0 > 0$ such that the sequence $\{D_n\} = \{n^{\alpha_0}\}$ satisfies*

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

For such a sequence $\{D_n\}$, $P\{\omega \,|\exists N(\omega), \ni |\varepsilon_i| \le D_n, 1 \le i \le n, n > N(\omega)\} = 1$.

Denote the theoretical norms of the basis $c_{j,n} = \|b_{j,1}\|_2^2$ and $d_{j,n} = \|b_{j,2}\|_2^2$ by

$$c_{j,n} = \int_a^b I_j(x) f(x) dx, d_{j,n} = \int_a^b K^2\left(\frac{x - t_{j+1}}{h}\right) f(x) dx.$$

Lemma A.2 *Under Assumptions (A2) and (A3), as $n \to \infty$,*

$$c_{j,n} = f(t_j) h (1 + r_{j,n,1}), \langle b_{j,1}, b_{j',1}\rangle \equiv 0, j \ne j' \tag{A.2}$$

$$d_{j,n} = \frac{2}{3} f(t_{j+1}) h \times \begin{cases} 1 + r_{j,n,2} & j = 0, ..., N-1, \\ 1/2 + r_{j,n,2} & j = -1, N, \end{cases} \tag{A.3}$$

$$\langle b_{j,2}, b_{j',2}\rangle = \frac{1}{6} f(t_{j+1}) h \times \begin{cases} 1 + \tilde{r}_{j,n,2} & |j' - j| = 1, \\ 0 & |j' - j| > 1, \end{cases} \tag{A.4}$$

where

$$\max_{0\le j\le N}|r_{j,n,1}|+\max_{-1\le j\le N}\{|r_{j,n,2}|+|\tilde r_{j,n,2}|\}\le C\omega(f,h). \tag{A.5}$$

In particular,

$$\frac{1}{3}f(t_{j+1})h\{1-C\omega(f,h)\}\le d_{j,n}\le\frac{2}{3}f(t_{j+1})h\{1+C\omega(f,h)\}. \tag{A.6}$$

Proof of Lemma 3.1. For brevity, we give only the proof of (3.1) for $A_{n,1}$. Take any $j=0,1,...,N$

$$\left|\|B_{j,1}\|_{2,n}^2-1\right|=\left|\sum\nolimits_{i=1}^n\xi_i\right|,\xi_i=\left\{B_{j,1}^2(X_i)-1\right\}n^{-1}$$

with $E\xi_i=0$ and for any $k\ge2$. Minkowski's inequality implies that

$$E|\xi_i|^k=n^{-k}E\left|B_{j,1}^2(X_i)-1\right|^k\le2^{k-1}n^{-k}E\left[B_{j,1}^{2k}(X_i)+1\right]\le\left\{\frac{2}{nh}\right\}^kC_0h,$$

while (A.2) implies that $E\xi_i^2\ge n^{-2}E\left[\frac{1}{2}B_{j,1}^4(X_i)-1\right]\ge\{2/(nh)\}^2C_1h$. One can then find a constant $c>0$ such that for $k>2$, $E|\xi_i|^k\le(cn^{-1}h^{-1})^{k-2}k!E|\xi_i|^2$. Applying Bernstein's inequality, we conclude that $P\left\{|\sum_{i=1}^n\xi_i|\ge\eta_0\log^{1/2}(n)(nh)^{-1/2}\right\}\le2n^{-3}$ for large enough $\eta_0>0$. Thus,

$$\sum\nolimits_{n=1}^\infty P\left\{\sup_{0\le j\le N}\left|\|B_{j,1}\|_{2,n}^2-1\right|\ge\eta\log^{1/2}(n)(nh)^{-1/2}\right\}<\infty$$

for such $\eta_0>0$, so that (3.1) follows. □

A. 2. Proof of Theorem 1

In this section, we investigate the behavior of $\tilde\varepsilon_1(x)$ defined in (3.4). Since $\langle\mathbf{B}_{j',1}(\mathbf{X}),\mathbf{B}_{j,1}(\mathbf{X})\rangle_n=0$ unless $j=j'$, $\tilde\varepsilon_1(x)$ can be written as $\tilde\varepsilon_1(x)=\sum_{j=0}^N\varepsilon_j^*B_{j,1}(x)\|B_{j,1}\|_{2,n}^{-2}$, in which $\varepsilon_j^*=\left\langle\mathbf{E},\mathbf{B}_{j,1}(\mathbf{X})\right\rangle_n=n^{-1}\sum_{i=1}^nB_{j,1}(X_i)\sigma(X_i)\varepsilon_i$.

Lemma A.3 *Let $\hat\varepsilon_1(x)=\sum_{j=0}^N\varepsilon_j^*B_{j,1}(x),x\in[a,b]$, for $A_{n,1}$ defined in (3.1)*

$$|\tilde\varepsilon_1(x)-\hat\varepsilon_1(x)|\le A_{n,1}(1-A_{n,1})^{-1}|\hat\varepsilon_1(x)|,x\in[a,b].$$

Thus, $\sup_{x\in[a,b]}|\tilde\varepsilon_1(x)|$ and $\sup_{x\in[a,b]}|\hat\varepsilon_1(x)|$ have the same asymptotic behavior.

Lemma A.4 *The pointwise variance of $\hat\varepsilon_1(x)$ is the function $\sigma_{n,1}^2(x)$ defined in (2.6) which satisfies for $\sup_{x\in[a,b]}|r_{n,1}(x)|\to0$*

$$E\{\hat\varepsilon_1(x)\}^2\equiv\sigma_{n,1}^2(x)=\frac{\sigma^2(x)}{f(x)nh}\{1+r_{n,1}(x)\},x\in[a,b]. \tag{A.7}$$

Lemma A.5 *Let the sequence* $\{D_n\}$ *satisfy (A.1), then as* $n \to \infty$

$$\left\|\hat{\varepsilon}_{n,1}(x) - \hat{\varepsilon}_{n,1}^{D}(x)\right\|_{\infty} = O\left(D_n^{-(1+\eta)}\sqrt{nh}\right) = o(1), \ \ w.\ p.\ 1,$$

where, for $x \in [a,b]$,

$$\hat{\varepsilon}_{n,1}(x) = \sigma_{n,1}(x)^{-1}\sum_{j=0}^{N} B_{j,1}(x)\,\varepsilon_j^* = \sigma_{n,1}(x)^{-1}\sum_{j=0}^{N} B_{j,1}(x)\left(\varepsilon_j^* - E\varepsilon_j^*\right),$$

$$\hat{\varepsilon}_{n,1}^{D}(x) = \sigma_{n,1}(x)^{-1}\sum_{j=0}^{N} B_{j,1}(x)\left(\varepsilon_j^* - E\varepsilon_j^*\right) I_{\{|\varepsilon|<D_n\}}. \tag{A.8}$$

Proof. Notice that $E\varepsilon_j^* = E\{n^{-1}\sum_{i=1}^{n} B_{j,1}(X_i)\,\sigma(X_i)\,\varepsilon_i\} = 0$, so that

$$\hat{\varepsilon}_{n,1}(x) = \left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int\int I_{j(x)}(v)\,\sigma(v)\,\varepsilon dZ_n(v,\varepsilon)$$

according to the definition of $Z_n(v,\varepsilon)$ in (3.9). The truncated part $\hat{\varepsilon}_{n,1}^{D}(x)$ is defined in (A.8). The tail part $\hat{\varepsilon}_{n,1}(x) - \hat{\varepsilon}_{n,1}^{D}(x)$ is bounded uniformly over $[a,b]$ by

$$\sup_{x\in[a,b]}\left|\left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int\int I_{j(x)}(v)\,\sigma(v)\,\varepsilon I_{\{|\varepsilon|\geq D_n\}} dZ_n(v,\varepsilon)\right|$$

$$\leq \sup_{x\in[a,b]}\left|\left\{\sigma_{n,1}(x)\,c_{j(x),n}\right\}^{-1}\frac{1}{n}\sum_{i=1}^{n} I_{j(x)}(X_i)\,\sigma(X_i)\,\varepsilon_i I_{\{|\varepsilon_i|\geq D_n\}}\right| \tag{A.9}$$

$$+\sup_{x\in[a,b]}\left|\left\{\sigma_{n,1}(x)\,c_{j(x),n}\right\}^{-1}\int\int I_{j(x)}(v)\,\sigma(v)\,\varepsilon I_{\{|\varepsilon|\geq D_n\}}) dF(v,\varepsilon)\right|. \tag{A.10}$$

By Lemma A.1, the term (A.9) is 0 almost surely. The term (A.10) is bounded by

$$\sup_{x\in[a,b]}\left\{\sigma_{n,1}(x)\,c_{j(x),n}\right\}^{-1}\int I_{j(x)}(v)\,\sigma(v)\,f(v)\left[\int|\varepsilon|\,I_{\{|\varepsilon|\geq D_n\}} dF(\varepsilon\,|v)\right]dv$$

$$\leq \sup_{x\in[a,b]}\left\{\sigma_{n,1}(x)\,c_{j(x),n}\right\}^{-1}\int I_{j(x)}(v)\,\sigma(v)\,f(v)\,dv\frac{M_\eta}{D_n^{1+\eta}} \leq C\frac{\sqrt{nh}}{D_n^{1+\eta}}.$$

The lemma follows immediately by the third condition in (A.1). □

Lemma A.6 *Define for* $x \in [a,b]$

$$\hat{\varepsilon}_{n,1}^{(0)}(x) = \left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int\int I_{j(x)}(v)\,\sigma(v)\,\varepsilon I_{\{|\varepsilon|<D_n\}} dB\{M(v,\varepsilon)\}$$

then as $n \to \infty$

$$\sup_{x\in[a,b]}\left|\hat{\varepsilon}_{n,1}^{(0)}(x) - \hat{\varepsilon}_{n,1}^{D}(x)\right| = O\left(h^{-1/2}n^{-1/2}D_n\log^2 n\right) = o(1), \ \ w.\ p.\ 1.$$

Proof. First, $\left|\hat{\varepsilon}_{n,1}^{(0)}(x)-\hat{\varepsilon}_{n,1}^{D}(x)\right|$ can be written as

$$\left|\left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int\int I_{j(x)}(v)\,\sigma(v)\,\varepsilon I_{\{|\varepsilon|<D_n\}}d\left[Z_n(v,\varepsilon)-B\left\{M(v,\varepsilon)\right\}\right]\right|,$$

which becomes the following via integration by parts

$$\left|\left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int\int\left[Z_n(v,\varepsilon)-B\left\{M(v,\varepsilon)\right\}\right]d\left\{I_{j(x)}(v)\,\sigma(v)\,\varepsilon I_{\{|\varepsilon|<D_n\}}\right\}\right|$$

$$\leq\left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int\int\left|Z_n(v,\varepsilon)-B\left\{M(v,\varepsilon)\right\}\right|d\left\{\varepsilon I_{\{|\varepsilon|<D_n\}}\right\}d\left\{I_{j(x)}(v)\,\sigma(v)\right\}.$$

Next, by Lemma A.4, the bounded variation of the function $\sigma(x)$ in Assumption (A2), the strong approximation result (3.10), and the first condition in (A.1), $\sup_{x\in[a,b]}\left|\hat{\varepsilon}_{n,1}^{(0)}(x)-\hat{\varepsilon}_{n,1}^{D}(x)\right|$ is bounded as

$$O\left\{(nh)^{1/2}n^{-1/2}h^{-1}\left(n^{-1/2}\log^2 n\right)D_n\right\}=O\left(n^{-1/2}h^{-1/2}D_n\log^2 n\right)=o(1)$$

with probability 1, thus completing the proof of the lemma. □

The next lemma finds a process $\hat{\varepsilon}_{n,1}^{(1)}(x)$ defined in terms of the 2-dimensional Brownian motion to approximate $\hat{\varepsilon}_{n,1}^{(0)}(x)$.

Lemma A.7 *Define for $x\in[a,b]$*

$$\hat{\varepsilon}_{n,1}^{(1)}(x)=\left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int\int I_{j(x)}(v)\,\sigma(v)\,\varepsilon I_{\{|\varepsilon|<D_n\}}dW\left\{M(v,\varepsilon)\right\}$$

then as $n\to\infty$, $\left\|\hat{\varepsilon}_{n,1}^{(1)}(x)-\hat{\varepsilon}_{n,1}^{(0)}(x)\right\|_\infty=O\left(h^{1/2}D_n^{-(1+\eta)}\right)=o(1)$ w. p. 1.

Proof. Based on the Rosenblatt transformation $M(x,\varepsilon)$ defined in (3.8), and $\frac{\partial M(x,\varepsilon)}{\partial(x,\varepsilon)}=f(x,\varepsilon)$, the term $\left\|\hat{\varepsilon}_{n,1}^{(1)}(x)-\hat{\varepsilon}_{n,1}^{(0)}(x)\right\|_\infty$ is bounded by

$$\sup_{x\in[a,b]}\left|\left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int\int I_{j(x)}(v)\,\sigma(v)\,|\varepsilon|\,I_{\{|\varepsilon|<D_n\}}dM(v,\varepsilon)\,W(1,1)\right|$$

$$\leq\sup_{x\in[a,b]}\left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int I_{j(x)}(v)\,\sigma(v)\,f(v)\,dv$$

$$\times\left\{\int|\varepsilon|\,I_{\{|\varepsilon|<D_n\}}dF(\varepsilon\,|v)\right\}|W(1,1)|$$

$$\leq C\left(\frac{\sqrt{nh}}{\sqrt{n}h}\right)h\frac{M_\eta}{D_n^{1+\eta}}|W(1,1)|=O\left(h^{1/2}D_n^{-(1+\eta)}\right)=o(1)\quad\text{w. p. 1.}$$

The last step is obtained by applying the third condition in (A.1). □

The next lemma expresses the distribution of $\hat{\varepsilon}_{n,1}^{(1)}(x)$ in terms of 1-dimensional Brownian motion.

Lemma A.8 *The process* $\hat{\varepsilon}_{n,1}^{(1)}(x)$ *has the same distribution as*

$$\hat{\varepsilon}_{n,1}^{(2)}(x) = \left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int I_{j(x)}(v)\,\sigma(v)\,s_n(v)\,f^{\frac{1}{2}}(v)\,dW(v)\,, x \in [a,b]$$

where

$$s_n^2(v) = \int \varepsilon^2 I_{\{|\varepsilon|<D_n\}} dF(\varepsilon\,|v)\,. \tag{A.11}$$

Proof. According to Itô's Isometry Theorem, var $\left\{\hat{\varepsilon}_{n,1}^{(1)}(x)\right\}$ and var $\left\{\hat{\varepsilon}_{n,1}^{(2)}(x)\right\}$ are exactly the same for any $x \in [a,b]$. Hence, the two Gaussian processes $\hat{\varepsilon}_{n}^{(1)}(x)$ and $\hat{\varepsilon}_{n}^{(2)}(x)$ have the same distribution. □

Lemma A.9 *Define for any* $x \in [a,b]$

$$\hat{\varepsilon}_{n,1}^{(3)}(x) = \left\{\sigma_{n,1}(x)\sqrt{n}c_{j(x),n}\right\}^{-1}\int I_{j(x)}(v)\,\sigma(v)\,f^{\frac{1}{2}}(v)\,dW(v)$$

then as $n \to \infty$, $\left\|\hat{\varepsilon}_{n,1}^{(2)}(x) - \hat{\varepsilon}_{n,1}^{(3)}(x)\right\|_{\infty} = O\left(D_n^{-\eta}h^{-1/2}\right) = o(1)$ *w. p.* 1.

Proof. By the fourth condition in (A.1) , $\sup_{x\in[a,b]}\left|\hat{\varepsilon}_{n,1}^{(2)}(x) - \hat{\varepsilon}_{n,1}^{(3)}(x)\right|$ is almost surely bounded by

$$\sup_{v\in[a,b]}\left|s_n^2(v) - 1\right| \sup_{x\in[a,b]}\left|\sigma_{n,1}^{-1}(x)\,c_{j(x),n}^{-1}n^{-1/2}\int I_{j(x)}(v)\,\sigma(v)\,f^{\frac{1}{2}}(v)\,dW(v)\right|$$
$$= O\left(D_n^{-\eta}h^{-1/2}\right) = o(1)\,.\ \text{w. p. } 1$$ □

Lemma A.10 *The process* $\hat{\varepsilon}_{n,1}^{(3)}(x)$ *is a Gaussian process with mean* 0, *variance* 1, *and covariance* cov $\left\{\hat{\varepsilon}_{n,1}^{(3)}(x)\,, \hat{\varepsilon}_{n,1}^{(3)}(y)\right\} = \delta_{j(x),j(y)}, \forall x, y \in [a,b]$.

Proof. This follows from Itô's Isometry Theorem and (A.7). □

Proof of Proposition 3.1. The proof follows immediately from Lemmas A.3, A.5, A.6, A.7, A.8, A.9 and A.10. □

Proof of Theorem 1. It is clear from Proposition 3.1 that the Gaussian process $U(x)$ consists of $(N+1)$ i.i.d. standard normal variables $U(t_0)\,,...,U(t_N)$. Hence Theorem 3.4 implies that as $n \to \infty$

$$P\left\{\sup_{x\in[a,b]}|U(x)| \le \tau/a_{N+1} + b_{N+1}\right\} \to \exp\left(-2e^{-\tau}\right).$$

By letting $\tau = -\log\left\{-\frac{1}{2}\log(1-\alpha)\right\}$, and using the definition of a_{N+1} and b_{N+1}, we obtain

$$\lim_{n\to\infty} P\left[\sup_{x\in[a,b]}|U(x)| \le -\log\left\{-\frac{1}{2}\log(1-\alpha)\right\}\{2\log(N+1)\}^{-1/2}\right.$$
$$\left.+\{2\log(N+1)\}^{1/2} - \frac{1}{2}\{2\log(N+1)\}^{-1/2}\{\log\log(N+1) + \log 4\pi\}\right] = 1-\alpha.$$

Replacing $U(x)$ with $\sigma_{n,1}(x)^{-1}\tilde{\varepsilon}_1(x)$ (Proposition 3.1), and the definition of d_n in (2.9) implies that

$$\lim_{n\to\infty} P\left[\sup_{x\in[a,b]}\left|\sigma_{n,1}(x)^{-1}\tilde{\varepsilon}_1(x)\right| \le \{2\log(N+1)\}^{1/2} d_n\right] = 1-\alpha.$$

As (3.5) implies that $\sqrt{nh/\log(N+1)}\,\|\tilde{m}_1(x)-m(x)\|_\infty = o_p(1)$. According to (3.3),

$$\begin{aligned}
& \lim_{n\to\infty} P\left[m(x)\in\hat{m}_1(x)\pm\sigma_{n,1}(x)\{2\log(N+1)\}^{1/2} d_n, \forall x\in[a,b]\right] \\
= & \lim_{n\to\infty} P\left[\{2\log(N+1)\}^{-1/2} d_n^{-1}\sup_{x\in[a,b]}\sigma_{n,1}^{-1}(x)\left|\tilde{\varepsilon}_1(x)+\tilde{m}_1(x)-m(x)\right|\le 1\right] \\
= & \lim_{n\to\infty} P\left[\{2\log(N+1)\}^{-1/2} d_n^{-1}\sup_{x\in[a,b]}\sigma_{n,1}^{-1}(x)\left|\tilde{\varepsilon}_1(x)\right|\le 1\right] = 1-\alpha. \quad \square
\end{aligned}$$

Appendix B: Proof of Theorem 2

B. 1. Preliminaries

In this subsection we examine matrices used in (2.10) of Theorem 2. In what follows, we use $|T|$ to denote the maximal absolute value of any matrix T, and M_{N+2} is the tridiagonal matrix as defined in (4.9).

Lemma B.1 *The inner product matrix V of the B-spline basis $\{B_{j,2}(x)\}_{j=-1}^{N}$ defined as $V = (v_{j'j})_{j,j'=-1}^{N} = (\langle B_{j',2}, B_{j,2}\rangle)_{j,j'=-1}^{N}$, has the following decomposition*

$$V = M_{N+2} + (\tilde{v}_{j'j})_{j,j'=-1}^{N} = M_{N+2} + \tilde{V}$$

where $\tilde{v}_{j'j}\equiv 0$ if $|j-j'|\ge 1$, and $\left|\tilde{V}\right| \le C\omega(f,h)$.

Proof. By (A.3), (A.4) and (A.5), the inner product of $\langle b_{j',2}, b_{j,2}\rangle$ can be replaced by $\frac{1}{6}f(t_{j+1})h$ if $|j'-j|=1$, and $\frac{1}{3}f(t_{j+1})h$ or $\frac{2}{3}f(t_{j+1})h$ when $j'=j$, plus some uniformly infinitesimal differences dominated by $\omega(f,h)$. Based on the definition of $B_{j,2}(x)$, the lemma follows immediately. $\square$

The next lemma shows that multiplication by M_{N+2} behaves similarly to multiplication by a constant.

Lemma B.2 *Given the matrix $\Omega = M_{N+2}+\Gamma$, for which $\Gamma = \left(\gamma_{jj'}\right)_{j,j'=-1}^{N}$ satisfies $\gamma_{jj'}\equiv 0$ if $|j-j'|\ge 1$ and $|\Gamma| \xrightarrow{p} 0$, there exist constants $c, C>0$ independent of n and Γ, such that in probability*

$$c|\boldsymbol{\xi}| \le |\Omega\boldsymbol{\xi}| \le C|\boldsymbol{\xi}|, C^{-1}|\boldsymbol{\xi}| \le \left|\Omega^{-1}\boldsymbol{\xi}\right| \le c^{-1}|\boldsymbol{\xi}|, \forall\boldsymbol{\xi}\in R^{N+2}. \tag{B.1}$$

Proof. In (4.9), M_{N+2} has diagonal elements 1, and the sum of the absolute values of off-diagonal elements in each row does not exceed $1/\sqrt{2}$. Hence it follows that $\left(1 - 1/\sqrt{2} - 3\left|\Gamma\right|\right)\left|\boldsymbol{\xi}\right| \leq \left|\Omega\boldsymbol{\xi}\right| \leq 3\left(1+\left|\Gamma\right|\right)\left|\boldsymbol{\xi}\right|$, which implies the first inequality of (B.1), and the second one follows by switching the roles of $\boldsymbol{\xi}$ and $\Omega\boldsymbol{\xi}$. □

As an application of Lemma B.2, consider the matrix $S = V^{-1}$ defined in (2.5). Let $\tilde{\boldsymbol{\xi}}_{j'} = \left\{\operatorname{sgn}\left(s_{j'j}\right)\right\}_{j=-1}^{N}$, then there exists a positive C_s such that

$$\sum\nolimits_{j=-1}^{N}\left|s_{j'j}\right| \leq \left|S\tilde{\boldsymbol{\xi}}_{j'}\right| \leq C_s\left|\tilde{\boldsymbol{\xi}}_{j'}\right| = C_s, \forall j' = -1, 0, ..., N. \tag{B.2}$$

The matrix S in the construction of the confidence band can not be computed exactly as it involves the unknown density $f(x)$. We approximate S by the inverse of M_{N+2}, with a simpler, distribution-free form in (4.9). This approximation is uniform for S_j in (2.5) and Ξ_j in (4.8) as well.

Lemma B.3 *As* $n \to \infty$, $\left|M_{N+2}^{-1} - S\right| \to 0$ *and* $\max_{0\leq j\leq N}\left|\Xi_j - S_j\right| \to 0$.

Proof. By definition, $M_{N+2}M_{N+2}^{-1} = I = VS = \left(M_{N+2} + \tilde{V}\right)S$. Denote by e_i the unit vector with i-th element 1. Applying Lemma B.2 with $\Omega = M_{N+2}$,

$$\begin{aligned} & c\left|M_{N+2}^{-1} - S\right| = c\max\nolimits_{i=1}^{N+2}\left|\left(M_{N+2}^{-1} - S\right)e_i\right| \\ \leq\ & \max\nolimits_{i=1}^{N+2}\left|M_{N+2}\left(M_{N+2}^{-1} - S\right)e_i\right| \leq c\left|\tilde{V}\right|\left(\left|M_{N+2}^{-1} - S\right| + \left|M_{N+2}^{-1}\right|\right) \end{aligned}$$

Since Lemma B.1 implies $\left|\tilde{V}\right| \leq C\omega(f,h)$, as $n \to \infty$, $\left|M_{N+2}^{-1} - S\right| = O\left\{\omega(f,h)\right\} \to 0$. By definition of submatrices S_j and Ξ_j, $\max_{0\leq j\leq N}\left|\Xi_j - S_j\right| \leq \left|M_{N+2}^{-1} - S\right|$. The lemma follows. □

B. 2. Variance calculation

We now examine the asymptotic behavior of

$$\tilde{\varepsilon}_2(x) = \operatorname{Proj}_{G_n^{(0)}}\mathbf{E} = \sum\nolimits_{j=-1}^{N}\tilde{a}_j B_{j,2}(x), x \in [a,b] \tag{B.3}$$

where the coefficient vector $\tilde{\mathbf{a}} = \left(\tilde{a}_{-1}, ..., \tilde{a}_N\right)^T$ is the solution to the normal equations

$$\left(\left\langle B_{j,2}, B_{j',2}\right\rangle_n\right)_{j,j'=-1}^{N}\left(\tilde{a}_j\right)_{j=-1}^{N} = \left(n^{-1}\sum\nolimits_{i=1}^{n} B_{j,2}(X_i)\sigma(X_i)\varepsilon_i\right)_{j=-1}^{N}.$$

In other words

$$\tilde{\mathbf{a}} = \left(\tilde{a}_j\right)_{j=-1}^{N} = \left(V + \tilde{B}\right)^{-1}\left(n^{-1}\sum\nolimits_{i=1}^{n} B_{j,2}(X_i)\sigma(X_i)\varepsilon_i\right)_{j=-1}^{N}, \tag{B.4}$$

where $\left|\tilde{B}\right| \leq A_{n,2} = O_p\left(n^{-1/2}h^{-1/2}\log^{1/2}(n)\right)$ by (3.2).

Now define the $\hat{a}_j$'s by replacing $\left(V+\tilde{B}\right)^{-1}$ with $V^{-1}=S$ in above formula, i.e.

$$\hat{\mathbf{a}} = \left(\hat{a}_j\right)_{j=-1}^{N} = \left(\sum\nolimits_{j=-1}^{N} s_{j'j} n^{-1} \sum\nolimits_{i=1}^{n} B_{j,2}\left(X_i\right)\sigma\left(X_i\right)\varepsilon_i\right)_{j=-1}^{N} \tag{B.5}$$

and define for $x \in [a,b]$

$$\hat{\varepsilon}_2(x) = \sum_{j'=-1}^{N} \hat{a}_{j'} B_{j',2}(x) = \sum_{j,j'=-1}^{N} s_{j'j} n^{-1} \sum\nolimits_{i=1}^{n} B_{j,2}\left(X_i\right)\sigma\left(X_i\right)\varepsilon_i B_{j',2}(x). \tag{B.6}$$

The next lemma is a special case of the unconditional version of (6.2) in Huang (2003).

Lemma B.4 *The pointwise variance of $\hat{\varepsilon}_2(x)$ is the function $\sigma_{n,2}^2(x)$ defined in (2.6), which satisfies*

$$E\left\{\hat{\varepsilon}_2^2(x)\right\} \equiv \sigma_{n,2}^2(x) = \frac{3\sigma^2(x)}{2f(x)\,nh}\mathbf{\Delta}^T(x) S_{j(x)} \mathbf{\Delta}(x)\left\{1+r_{n,2}(x)\right\}$$

with $\sup_{x\in[a,b]}\left|r_{n,2}(x)\right| \to 0$, *$j(x)$ in (2.3), $\mathbf{\Delta}(x)$ in (4.7) and matrix S_j in (2.5). Consequently, there exist $0 < c_\sigma < C_\sigma$ such that for n large enough*

$$c_\sigma (nh)^{-1/2} \le \sigma_{n,2}(x) \le C_\sigma (nh)^{-1/2}, \forall x \in [a,b]. \tag{B.7}$$

Proof. See Wang and Yang (2006). □

B. 3. Proof of Theorem 2

The next several lemmas are needed for the proof of Proposition 3.2.

Lemma B.5 *Define for $x \in [a,b]$*

$$\begin{aligned} \hat{\varepsilon}_{n,2}(x) &= \sigma_{n,2}^{-1}(x)\,\hat{\varepsilon}_2(x) = \sigma_{n,2}^{-1}(x)\sum\nolimits_{j'=-1}^{N} \hat{a}_{j'} B_{j',2}(x), \\ \hat{\varepsilon}_{n,2}^{D}(x) &= \sigma_{n,2}^{-1}(x)\sum\nolimits_{j'=-1}^{N} \hat{a}_{j'} B_{j',2}(x)\, I_{\{|\varepsilon|<D_n\}}. \end{aligned}$$

where D_n satisfies (A.1). Then with probability 1

$$\left\|\hat{\varepsilon}_{n,2}(x) - \hat{\varepsilon}_{n,2}^{D}(x)\right\|_\infty = O\left(n^{1/2}h^{1/2}D_n^{-(1+\eta)}\right) = o(1).$$

Proof. Since obviously $E\hat{\varepsilon}_{n,2}(x) = 0$, $\forall x \in [a,b]$,

$$\hat{\varepsilon}_{n,2}(x) = \sigma_{n,2}^{-1}(x)\, n^{-1/2} \sum_{j'=j(x)-1}^{j(x)} B_{j',2}(x) \sum_{j=-1}^{N} s_{j'j} \int\int B_{j,2}(v)\,\sigma(v)\,\varepsilon\, dZ_n(v,\varepsilon)$$

where $Z_n(x,\varepsilon)$ is defined in (3.9). The technical proof is very similar to Lemma A.5, except that we employ (B.2) to deal with $\sum_{j=-1}^{N} s_{j'j}$. The same order is also achieved. □

Lemma B.6 *Let M be the Rosenblatt transformation given in (3.8) and define*

$$\hat{\varepsilon}_{n,2}^{(0)}(x) = \frac{1}{\sqrt{n}\sigma_{n,2}(x)} \sum_{j',j=-1}^{N} B_{j',2}(x) s_{j'j} \int\int B_{j,2}(v)\sigma(v)\varepsilon I_{\{|\varepsilon|<D_n\}} dB\{M(v,\varepsilon)\}$$

for $x \in [a,b]$. Then as $n \to \infty$

$$\sup_{x\in[a,b]} \left|\hat{\varepsilon}_{n,2}^{(0)}(x) - \hat{\varepsilon}_{n,2}^{D}(x)\right| = O\left(n^{-1/2}h^{-1/2}D_n \log^2 n\right) = o(1) \quad w.\ p.\ 1.$$

Proof. See Lemma A.6. □

Lemma B.7 *Define for $x \in [a,b]$*

$$\hat{\varepsilon}_{n,2}^{(1)}(x) = \frac{\sigma_{n,2}^{-1}(x)}{\sqrt{n}} \sum_{j',j=-1}^{N} B_{j',2}(x) s_{j'j} \int\int B_{j,2}(v)\sigma(v)\varepsilon I_{\{|\varepsilon|<D_n\}} dW\{M(v,\varepsilon)\},$$

then as $n \to \infty$

$$\sup_{x\in[a,b]} \left|\hat{\varepsilon}_{n,2}^{(1)}(x) - \hat{\varepsilon}_{n,2}^{(0)}(x)\right| = O\left(h^{1/2}D_n^{-(1+\eta)}\right) = o(1) \quad w.\ p.\ 1.$$

Lemma B.8 *The process $\hat{\varepsilon}_{n,2}^{(1)}(x), x \in [a,b]$ has the same distribution as*

$$\hat{\varepsilon}_{n,2}^{(2)}(x) = \sigma_{n,2}^{-1}(x) n^{-1/2} \sum\nolimits_{j',j=-1}^{N} B_{j',2}(x) s_{j'j} \int\int B_{j,2}(v)\sigma(v) s_n(v) f^{\frac{1}{2}}(v) dW(v)$$

for $x \in [a,b]$, where $s_n^2(v)$ is as defined in (A.11).

Proof. Similar to that of Lemma A.8, see Wang and Yang (2006) for details. □

Lemma B.9 *Define for any $x \in [a,b]$*

$$\hat{\varepsilon}_{n,2}^{(3)}(x) = \frac{1}{\sqrt{n}\sigma_{n,2}(x)} \sum_{j',j=-1}^{N} B_{j',2}(x) s_{j'j} \int B_{j,2}(v)\sigma(v) f^{\frac{1}{2}}(v) dW(v)$$

then $\operatorname{var}\left\{\hat{\varepsilon}_{n,2}^{(3)}(x)\right\} \equiv 1, \forall x \in [a,b]$, *and as $n \to \infty$*

$$\left\|\hat{\varepsilon}_{n,2}^{(2)}(x) - \hat{\varepsilon}_{n,2}^{(3)}(x)\right\|_\infty = O\left(h^{-1/2}D_n^{-\eta}\right) = o(1) \quad w.\ p.\ 1.$$

Proof. Using (A.1) in the last step, the term $\sup_{x\in[a,b]} \left|\hat{\varepsilon}_{n,2}^{(2)}(x) - \hat{\varepsilon}_{n,2}^{(3)}(x)\right|$ is bounded by

$$\sup_{x\in[a,b]} \left|1 - s_n^2(x)\right| \sup_{x\in[a,b]} \left\{\frac{\sigma_{n,2}^{-1}(x)}{\sqrt{n}} \sum_{j',j=-1}^{N} B_{j',2}(x) |s_{j'j}| \int B_{j,2}(v)\sigma(v) f^{\frac{1}{2}}(v) dW(v)\right\}$$

$$\le M_\eta D_n^{-\eta} h^{1/2} C \left| \int \sigma(v) f^{\frac{1}{2}}(v)\, dW(v) \right| = O\left(h^{-1/2} D_n^{-\eta}\right) = o(1) \text{ w. p. 1.}$$

Meanwhile, directly from (2.7) and (2.6), for any $x \in [a, b]$

$$\text{var}\left\{\hat{\varepsilon}_{n,2}^{(3)}(x)\right\} = E\left\{\frac{\sigma_{n,2}^{-1}(x)}{\sqrt{n}} \sum_{j',j=-1}^{N} B_{j',2}(x)\, s_{j'j} \int B_{j,2}(v)\, \sigma(v)\, f^{\frac{1}{2}}(v)\, dW(v)\right\}^2 = 1.$$

□

Now define for any $j' = -1, ..., N$ and $x \in [a, b]$, the functions

$$\zeta_{j'}(x) = n^{-1/2} \sigma_{n,2}^{-1}(x)\, B_{j',2}(x), \tilde{\boldsymbol{\zeta}}(x) = \left(\zeta_{j(x)-1}(x), \zeta_{j(x)}(x)\right)^T$$

and the random vector $\boldsymbol{\Lambda} = (\Lambda_{-1}, \Lambda_0, \ldots, \Lambda_N)^T$ where

$$\Lambda_{j'} = \sum_{j=-1}^{N} s_{j'j} \int B_{j,2}(v)\, \sigma(v)\, f^{\frac{1}{2}}(v)\, dW(v).$$

Then $\boldsymbol{\Lambda} \sim \mathbf{N}(\mathbf{0}, S \sum S)$ as $E\Lambda_{j'} = 0, \forall j' = -1, .., N$, and the covariance is $E\Lambda_{j'}\Lambda_{l'} = \sum_{j,l=-1}^{N} s_{j'j}\sigma_{jl}s_{ll'}$, for any $j', l' = -1, .., N$, and σ_{jl} is defined in (2.7). Notice that

$$\hat{\varepsilon}_{n,2}^{(3)}(x) \equiv \sum_{j'=j(x)-1,j(x)} \zeta_{j'}(x)\, \Lambda_{j'} = \tilde{\boldsymbol{\zeta}}(x)^T \boldsymbol{\Lambda}_{j(x)}, \boldsymbol{\Lambda}_j = (\Lambda_{j-1}, \Lambda_j)^T, j = 0, ..., N.$$

Since Lemma B.9 states that the variance of $\hat{\varepsilon}_{n,2}^{(3)}(x) \equiv 1$, it follows that

$$\hat{\varepsilon}_{n,2}^{(3)}(x) = \frac{\tilde{\boldsymbol{\zeta}}(x)^T \boldsymbol{\Lambda}_{j(x)}}{\sqrt{\tilde{\boldsymbol{\zeta}}(x)^T \left\{\text{cov}\left(\boldsymbol{\Lambda}_{j(x)}\right)\right\} \tilde{\boldsymbol{\zeta}}(x)}}. \tag{B.8}$$

Lemma B.10 *For any given* $0 < \alpha < 1$,

$$\liminf_{n\to\infty} P\left(\sup_{x\in[a,b]} |\hat{\varepsilon}_{n,2}(x)| \le [2\{\log(N+1) - \log\alpha\}]^{1/2}\right) \ge 1 - \alpha. \tag{B.9}$$

PROOF. Define for any $0 \le j \le N, Q_j = \boldsymbol{\Lambda}_j^T \{\text{cov}(\boldsymbol{\Lambda}_j)\}^{-1} \boldsymbol{\Lambda}_j$. Result 4.7 (a), page 140 of Johnson and Wichern (1992) ensures that Q_j is distributed as χ_2^2, hence

$$P[Q_j > 2\{\log(N+1) - \log\alpha\}] = \frac{\alpha}{N+1}, \forall\, 0 \le j \le N.$$

Then (B.8) and the Maximization Lemma of Johnson and Wichern (1992), page 66, ensures that

$$\left\{\hat{\varepsilon}_{n,2}^{(3)}(x)\right\}^2 = \frac{\left|\tilde{\boldsymbol{\zeta}}(x)^T \boldsymbol{\Lambda}_{j(x)}\right|^2}{\tilde{\boldsymbol{\zeta}}(x)^T \left\{\text{cov}\left(\boldsymbol{\Lambda}_{j(x)}\right)\right\} \tilde{\boldsymbol{\zeta}}(x)} \le \boldsymbol{\Lambda}_{j(x)}^T \left\{\text{cov}\left(\boldsymbol{\Lambda}_{j(x)}\right)\right\}^{-1} \boldsymbol{\Lambda}_{j(x)} = Q_{j(x)},$$

for any $x \in [a,b]$. Therefore $\sup_{x\in[a,b]} \left|\hat{\varepsilon}_{n,2}^{(3)}(x)\right|^2 \leq \max_{0\leq j\leq N}\{Q_j\}$ and

$$P\left[\sup_{x\in[a,b]}\left|\hat{\varepsilon}_{n,2}^{(3)}(x)\right|^2 \leq 2\{\log(N+1)-\log\alpha\}\right]$$
$$\geq P\left[\max_{0\leq j\leq N}\{Q_j\} \leq 2\{\log(N+1)-\log\alpha\}\right] \geq 1-\alpha.$$

Equation (B.9) follows from Lemmas B.5, B.6 B.7, B.8, B.9. □

Lemma B.11

$$\left|\sup_{x\in[a,b]}\left|\frac{\hat{\varepsilon}_2(x)}{\sigma_{n,2}(x)}\right| - \sup_{x\in[a,b]}\left|\frac{\tilde{\varepsilon}_2(x)}{\sigma_{n,2}(x)}\right|\right| = O_p\left(\sqrt{\frac{\log n}{nh}}\right) = o_p(1).$$

Proof. Recall the definition for $\tilde{\mathbf{a}} = (\tilde{a}_{-1}, \tilde{a}_0, ..., \tilde{a}_N)^T$ and $\hat{\mathbf{a}} = (\hat{a}_{-1}, \hat{a}_0, \ldots, \hat{a}_N)^T$ in (B.4) and (B.5). Then $\left(V+\tilde{B}\right)\tilde{\mathbf{a}} = V\hat{\mathbf{a}}$. Based on Lemma B.2 and (3.2), there exists a constant c such that $c\left|\hat{\mathbf{a}}-\tilde{\mathbf{a}}\right| \leq \left|V\left(\hat{\mathbf{a}}-\tilde{\mathbf{a}}\right)\right| = \left|\tilde{B}\tilde{\mathbf{a}}\right| \leq A_{n,2}\left(\left|\hat{\mathbf{a}}-\tilde{\mathbf{a}}\right| + \left|\hat{\mathbf{a}}\right|\right)$, it implies that $\left|\hat{\mathbf{a}}-\tilde{\mathbf{a}}\right| \leq \frac{A_{n,2}}{c-A_{n,2}}\left|\hat{\mathbf{a}}\right|$. From the definitions of $\tilde{\varepsilon}_2(x)$ in (B.3) and $\hat{\varepsilon}_2(x)$ in (B.6), plus (B.7) and (A.6), as $n\to\infty$

$$\sup_{x\in[a,b]}\left|\frac{\hat{\varepsilon}_2(x)}{\sigma_{n,2}(x)} - \frac{\tilde{\varepsilon}_2(x)}{\sigma_{n,2}(x)}\right| \leq \sup_{x\in[a,b]}\left|\sum\nolimits_{j=-1}^{N}\frac{\left|\widehat{\mathbf{a}}-\tilde{\mathbf{a}}\right| B_{j,2}(x)}{\sigma_{n,2}(x)}\right| \leq Cn^{1/2}\frac{A_{n,2}}{c-A_{n,2}}\left|\hat{\mathbf{a}}\right|.$$

Using (A.6) again, we conclude that as $n\to\infty$

$$\sup_{x\in[a,b]}\left|\frac{\hat{\varepsilon}_2(x)}{\sigma_{n,2}(x)}\right| \geq \frac{\sqrt{nh}}{C_\sigma}\sup_{x\in[a,b]}\left|\hat{\mathbf{a}}\mathbf{B}_2^T(x)\right| \geq C\sqrt{n}\left|\hat{\mathbf{a}}\right|$$

where $\mathbf{B}_2(x) = \{B_{-1,2}(x), ..., B_{N,2}(x)\}^T$, $\mathbf{b}_2(x) = \{b_{-1,2}(x), ..., b_{N,2}(x)\}^T$.

The desired result follows, i.e.

$$\sup_{x\in[a,b]}\left|\frac{\hat{\varepsilon}_2(x)}{\sigma_{n,2}(x)} - \frac{\tilde{\varepsilon}_2(x)}{\sigma_{n,2}(x)}\right| \leq C\frac{A_{n,2}}{c-A_{n,2}}\sup_{x\in[a,b]}\left|\frac{\hat{\varepsilon}_2(x)}{\sigma_{n,2}(x)}\right| = O_p\left(\sqrt{\frac{\log n}{nh}}\right). \qquad \square$$

Proof of Proposition 3.2. This follows from Lemma B.10 and Lemma B.11. □

Proof of Theorem 2. Now (3.5) implies that

$$\sqrt{nh/\log(N+1)}\left\|\tilde{m}_2(x)-m(x)\right\|_\infty = O_p\left\{\sqrt{nh^5/\log(N+1)}\right\} = o_p(1).$$

Applying (3.6) in Proposition 3.2

$$\liminf_{n\to\infty} P\left[m(x) \in \hat{m}_2(x) \pm \sigma_{n,2}(x)\{2\log(N+1)-2\log\alpha\}^{1/2}, \forall x\in[a,b]\right]$$
$$= \liminf_{n\to\infty} P\left[\sup_{x\in[a,b]}\sigma_{n,2}^{-1}(x)\left|\tilde{\varepsilon}_2(x)+\tilde{m}_2(x)-m(x)\right| \leq \{2\log(N+1)-2\log\alpha\}^{1/2}\right]$$
$$= \liminf_{n\to\infty} P\left[\sup_{x\in[a,b]}\left|\frac{\tilde{\varepsilon}_2(x)}{\sigma_{n,2}(x)}\right| \leq \{2\log(N+1)-2\log\alpha\}^{1/2}\right] \geq 1-\alpha. \qquad \square$$