SIMULTANEOUS CONFIDENCE BANDS IN NONLINEAR REGRESSION MODELS WITH NONSTATIONARITY

Degui Li[1], Weidong Liu[2], Qiying Wang[3] and Wei Biao Wu[4]

[1] *The University of York,* [2] *Shanghai Jiao Tong University,*
[3] *The University of Sydney and* [4] *The University of Chicago*

Supplementary Material

In this supplemental material, we provide the proofs of Propositions 2–4 stated in Section 5 of the main document and some technical lemmas with the proofs; see Appendices A and B. In Appendix C we provide a discussion on how to construct the SCBs when the Nadaraya-Watson kernel smoothing method is used to estimate the regression function.

Appendix A: Some technical lemmas

We start with some technical lemmas, which are very useful in the proofs of Propositions 2–4. The first lemma is a well-known exponential inequality for the martingale differences, see, for example, de la Peña (1999).

Lemma A.1 *Let* $(d_t, \mathcal{F}_t)_{t\geq 1}$ *be a sequence of martingale differences and* $\overline{\sigma}_n^2 = \sum_{t=1}^n \mathsf{E}(d_t^2|\mathcal{F}_{t-1})$*. Suppose there exists a constant* $a_1 > 0$ *such that* $\mathsf{P}(|d_t| \leq a_1|\mathcal{F}_{t-1}) = 1$ *for all* $t \geq 2$*. Then,*

$$\mathsf{P}\left(\sum_{t=1}^n d_t > x,\, \overline{\sigma}_n^2 \leq y\right) \;\leq\; \exp\left\{-\frac{x^2}{2(y+a_1x)}\right\}$$

for all $x, y > 0$.

Lemma A.2 *Let* $f(x)$ *be a real function on a compact support* $[-A_1, A_1]$ *satisfying* $|f(x)-f(y)| \leq C|x-y|$*. Under the conditions (C1) and (C3)(ii), for any* $B_n \leq M_0\sqrt{n}$ *with* M_0 *being a positive constant, we have*

$$\sup_{|x|\leq B_n}\left|\sum_{t=1}^n \{f[(X_t+x)/h] - f(X_t/h)\}\right| = O_P\left(\overline{\eta}_n \log n\right), \tag{A.1}$$

where

$$\overline{\eta}_n = \begin{cases} (nB_n \log n)^{1/3}\, h, & \textit{if } B_n\, h \log n \geq 1, \\ [nh/(B_n \log n)]^{1/3}, & \textit{if } B_n\, h \log n < 1. \end{cases}$$

In particular, by letting $B_n = M_0\sqrt{n}/\log^{a_2} n$ *with* $a_2 \geq 0$, *we get*

$$\sup_{|x| \leq M_0\sqrt{n}/\log^{a_2} n} \left| \sum_{t=1}^{n} \{f[(X_t + x)/h] - f(X_t/h)\} \right| = O_P(\sqrt{n}h \log^{(4-a_2)/3} n). \qquad \text{(A.2)}$$

Proof. The proof is similar to Theorem 2.3 of Chan and Wang (2014). Due to the condition that $\sum_{k=0}^{\infty} \phi_k \neq 0$ and $\sum_{k=0}^{\infty} |\phi_k| < \infty$, there exists a positive integer q_0 such that $\sum_{k=0}^{q} \phi_k \neq 0$ for all $q \geq q_0$. Without loss of generality we assume $q_0 = 0$ (otherwise it only requires a routine modification). We start with several facts which can facilitate the proof.

F1. *For any* $t > s$, $(X_t - X_s)/\sqrt{t-s}$ *has a uniformly bounded density* $d_{s,t}(x)$, *satisfying* $\sup_{x \in R} |d_{s,t}(x+u) - d_{s,t}(x)| \leq C \min\{|u|, 1\}$.

F2. *There exists a positive constant* H_0 *which is independent of* k_1, k_2, k_3 *and* m *such that*

$$\sup_x E\left(\Big| \sum_{t=k_2}^{k_3} f[(X_t + x)/h] \Big|^m \mid \mathcal{F}_{k_1} \right)$$
$$\leq H_0^m (m+1)! (k_3 - k_1)^{1/2} h \left[1 + \left\{ (k_3 - k_2)^{1/2} h \right\}^{m-1} \right], \qquad \text{(A.3)}$$

for all $0 \leq k_1 < k_2 < k_3 \leq n$ *and integer* $m \geq 1$, *where* $\mathcal{F}_s = \sigma\{\eta_s, \eta_{s-1}, \cdots\}$.

F3. $\sup_x |\sum_{t=1}^{n} f[(X_t + x)/h]| = O\left(\max\{\sqrt{n}h, 1\} \log n\right)$ *a.s.*

The fact F1 is proved in Example 2.2 of Chan and Wang (2014). Using F1 and the conditional arguments, fact F2 follows easily and the details can be found in Lemma 5.1 of Chan and Wang (2014). Using F2 as a main tool (taking $m = \log n$), fact F3 can be proved by Markov's inequality and standard arguments (see Theorem 2.1 of Chan and Wang, 2014, for more details).

We are now ready to prove Lemma A.2. Let

$$b_n = \begin{cases} (nB_n/\log^2 n)^{2/3}, & \text{if } B_n h \log n \geq 1, \\ (nhB_n^2/\log n)^{2/3}, & \text{if } B_n h \log n < 1. \end{cases}$$

Let T_n be the largest integer s such that $sb_n \leq n$. Furthermore write $y_j = -B_n + j/m_n', j = 0, 1, 2, \cdots, m_n$, where $m_n' = [n/(h\overline{\eta}_n \log n)]$ and $m_n = [2B_n m_n'] + 1$, where $\overline{\eta}_n$ is defined in the lemma. It is easy to find that $T_n b_n \leq n$ and

$$m_n \leq 2B_n m_n' + 1 \leq Cn^2, \tag{A.4}$$

due to $nh^2 \to \infty$. Using the fact $|f(x) - f(y)| \leq C|x - y|$ and by the standard arguments, we can show that

$$\begin{aligned} &\sup_{|x| \leq B_n} \left| \sum_{t=1}^{n} \left\{ f[(X_t + x)/h] - f(X_t/h) \right\} \right| \\ \leq\ & \max_{1 \leq j \leq m_n} \left| \sum_{t=1}^{n} \left\{ f[(X_t + y_j)/h] - f(X_t/h) \right\} \right| + O_{a.s.}(\overline{\eta}_n \log n) \\ \leq\ & \max_{1 \leq j \leq m_n} \left| \sum_{s=2}^{T_n - 1} \Delta_{ns}(y_j) \right| + \max_{1 \leq j \leq m_n} \Delta_n(y_j) + O_{a.s.}(\overline{\eta}_n \log n), \end{aligned} \tag{A.5}$$

where

$$\begin{aligned} \Delta_{ns}(x) &= \sum_{t = sb_n + 1}^{(s+1)b_n} \left\{ f[(X_t + x)/h] - f(X_t/h) \right\}, \quad \text{for } s = 2, \ldots, T_n - 1, \\ \Delta_n(x) &\leq \left(\sum_{t=1}^{2b_n} + \sum_{t = T_n b_n + 1}^{n} \right) \left| f[(X_t + x)/h] - f(X_t/h) \right|. \end{aligned}$$

With the fact F3, it is readily seen that

$$\begin{aligned} \max_{1 \leq j \leq m_n} \Delta_n(y_j) &\leq C \max\{h\sqrt{b_n + |n - T_n b_n|}, 1\} \log n \\ &\leq C\overline{\eta}_n \log n \quad a.s. \end{aligned}$$

This, together with (A.5), implies that (A.1) will follow if we prove

$$\max_{1 \leq j \leq m_n} \left(\Big| \sum_{\substack{s=2 \\ s \text{ is even}}}^{T_n} \Delta_{ns}(y_j) \Big| + \Big| \sum_{\substack{s=3 \\ s \text{ is odd}}}^{T_n} \Delta_{ns}(y_j) \Big| \right) = O_P(\overline{\eta}_n \log n). \tag{A.6}$$

We next only prove (A.6) for the case that s is even. The other case can be dealt with similarly, and the details are thus omitted. To this end, let $\mathcal{F}^*_{n,\nu} = \mathcal{F}_{n,(2\nu+1)b_n}$ for $\nu \geq 0$, and

$$\begin{aligned} \Delta'_{ns}(x) &= \Delta_{n,2s}(x) I\left\{|\Delta_{n,2s}(x)| \leq C_*\,\overline{\eta}_n\right\}, \\ \Delta^*_{ns}(x) &= \Delta'_{ns}(x) - \mathsf{E}\big(\Delta'_{ns}(x) \mid \mathcal{F}^*_{n,s-1}\big), \end{aligned}$$

where C_* is a positive constant which will be specified later. With these notation, to prove (A.6) when s is even, it suffices to show

$$\lambda_{1n} \equiv \max_{1\leq j\leq m_n} \Big|\sum_{s=1}^{T_n/2} \Delta^*_{ns}(y_j)\Big| = O_P\left(\overline{\eta}_n \log n\right), \tag{A.7}$$

$$\lambda_{2n} \equiv \max_{1\leq j\leq m_n} \left|\sum_{s=1}^{T_n/2} \mathsf{E}\big(\Delta_{n,2s}(y_j) \mid \mathcal{F}^*_{n,s-1}\big)\right| = O_P\left(\overline{\eta}_n \log n\right), \tag{A.8}$$

$$\begin{aligned} \lambda_{3n} &\equiv \max_{1\leq j\leq m_n} \left|\sum_{s=1}^{T_n/2} \Delta_{n,2s}(y_j) I\left\{|\Delta_{n,2s}(y_j)| > C_*\overline{\eta}_n\right\}\right| + \\ &\quad \max_{1\leq j\leq m_n} \left|\sum_{s=1}^{T_n/2} \mathsf{E}\big[\Delta_{n,2s}(y_j) I\left\{|\Delta_{n,2s}(y_j)| > C_*\overline{\eta}_n\right)\}\big]\right| \\ &= O_P\left(\overline{\eta}_n \log n\right). \end{aligned} \tag{A.9}$$

We first give the proof of (A.8). Letting $s_n = (2s-1)b_n$, note that, for any $2sb_n < t \leq (2s+1)b_n$,

$$\begin{aligned} & \left|\mathsf{E}\left\{f[(X_t+x)/h] - f(X_t/h)\right\}|\mathcal{F}^*_{n,s-1}]\right| \\ = & \left|\mathsf{E}\Big\{f[(X_t+x)/h] - f(X_t/h)\Big\}\Big|\mathcal{F}_{n,s_n}]\right| \\ = & \left|\int_{-\infty}^{\infty} \left\{f[(X_{s_n} + y\sqrt{t-s_n} + x)/h] - f[(X_{s_n} + y\sqrt{t-s_n})/h]\right\} d_{s_n,k}(y) dy\right| \\ \leq & \frac{h}{\sqrt{t-s_n}} \int_{-\infty}^{\infty} f(X_{s_n}/h + y) \left|d_{s_n,k}\left[(hy-x)/\sqrt{t-s_n}\right] - d_{s_n,k}(hy/\sqrt{t-s_n})\right| dy \\ \leq & C|x|h/(t-s_n), \end{aligned} \tag{A.10}$$

due to the fact F1. It is readily seen that

$$\begin{aligned}\lambda_{2n} &\leq \sum_{s=1}^{T_n/2} \max_{1\leq j\leq m_n} \left|\mathsf{E}\big(\Delta_{n,2s}(y_j) \mid \mathcal{F}^*_{n,s-1}\big)\right| \\ &\leq CT_n b_n h b_n^{-1} \max_{1\leq j\leq m_n} |y_j| \\ &\leq Cnhb_n^{-1} B_n \leq C\overline{\eta}_n \log n, \qquad \text{(A.11)}\end{aligned}$$

which yields (A.8).

We next consider the proof of (A.9). Using (A.3) in the fact F2 with $k_1 = 0, k_2 = 2sb_n + 1$ and $k_3 = (2s+1)b_n$, for any integer $m \geq 1$,

$$\sup_x \mathsf{E}|\Delta_{n,2s}(x)|^m \leq H_0^m (m+1)! b_n^{1/2} h \left[1 + (b_n^{1/2}h)^{m-1}\right].$$

Using this fact, we have

$$\begin{aligned}\mathsf{E}[\lambda_{3n}] &\leq 2\sum_{j=1}^{m_n}\sum_{s=1}^{T_n/2} \mathsf{E}\left[|\Delta_{n,2s}(y_j)| I\left\{|\Delta_{n,2s}(y_j)| > C_*\overline{\eta}_n\right\}\right] \\ &\leq 2(C_*\overline{\eta}_n)^{(1-m)} \sum_{j=1}^{m_n}\sum_{s=1}^{T_n/2} \mathsf{E}|\Delta_{n,2s}(y_j)|^m \\ &\leq C\overline{\eta}_n^{(1-m)} (H_0/C_*)^m (m+1)! m_n T_n \max\{1, (b_n^{1/2}h)^m\} \\ &\leq Cn^3 (\log n)^{1-m} (H_0/C_*)^m (m+1)!,\end{aligned}$$

due to (A.4) and the definitions of $\overline{\eta}_n$ and b_n. By taking $m = \log n$ and $C_* \geq H_0\, e^5$, it follows from the Stirling approximation of $(m+1)!$ that

$$\begin{aligned}\mathsf{E}[\lambda_{3n}] &\leq Cn^3 e^{-5\log n} \sqrt{2\pi(m+1)} \Big(\frac{m+1}{e}\Big)^{m+1} (\log n)^{1-m} \\ &\leq Cn^{-2} \log^{5/2} n = O(\overline{\eta}_n \log n), \qquad \text{(A.12)}\end{aligned}$$

which implies (A.9).

Finally, we consider the proof of (A.7). Note that, similarly to the proof of (A.10),

$$\begin{aligned}I_{k,j} &\equiv \left|\mathsf{E}\left(\{f[(X_j + x)/h] - f(X_j/h)\} \mid \mathcal{F}_{n,k}\right)\right| \\ &\leq \frac{h}{\sqrt{k-j}} \int_{-\infty}^{\infty} f(X_j/h + y) \left| d_{k,j}[(hy - x)/\sqrt{j-k}] - d_{k,j}(hy/\sqrt{j-k}) \right| dy \\ &\leq C|x|h/(j-k),\end{aligned}$$

for any $k < j$. This, together with (A.3) with $k_1 = (2s-1)b_n, k_2 = 2sb_n + 1$ and $k_3 = (2s+1)b_n$, implies that, for any $x \in \mathsf{R}$, that

$$\begin{aligned} & \mathsf{E}\left[\Delta_{ns}^{*2}(x) \mid \mathcal{F}_{n,s-1}^*\right] \le 2\mathsf{E}\left[\Delta_{n,2s}^2(x) \mid \mathcal{F}_{n,(2s-1)b_n}\right] \\ \le\ & 2 \sum_{t=2sb_n+1}^{(2s+1)b_n} \mathsf{E}\Big(\{f[(X_t+x)/h] - f(X_t/h)\}^2 \mid \mathcal{F}_{n,(2s-1)b_n}\Big) + 4 \sum_{2sb_n+1 \le t_1 < t_2 \le (2s+1)b_n} \\ & \Big|\mathsf{E}\Big(\{f[(X_{t_1}+x)/h] - f(X_{t_1}/h)\}\{f[(X_{t_2}+x)/h] - f(X_{t_2}/h)\} \mid \mathcal{F}_{n,(2s-1)b_n}\Big)\Big| \\ \le\ & C\, b_n^{1/2} h + 4 \sum_{2sb_n+1 \le t_1 < t_2 \le (2s+1)b_n} \mathsf{E}\Big(\big|f[(X_{t_1}+x)/h] - f(X_{t_2}/h)\big| I_{t_1,t_2} \mid \mathcal{F}_{n,(2s-1)b_n}\Big) \\ \le\ & C b_n^{1/2} h + C b_n^{-1/2} |x| h^2 \sum_{2sb_n+1 \le t_1 < t_2 \le (2s+1)b_n} (t_2 - t_1)^{-1} \\ \le\ & C b_n^{1/2} h (1 + |x| h \log b_n). \end{aligned}$$

Hence, we have

$$\begin{aligned} & \max_{0 \le j \le m_n} \sum_{s=1}^{T_n/2} \mathsf{E}\left[\Delta_{ns}^{*2}(y_j) \mid \mathcal{F}_{n,s-1}^*\right] \\ \le\ & C T_n b_n^{1/2} h \left(1 + h \log b_n \max_{0 \le j \le m_n} |y_j|\right) \\ \le\ & C n b_n^{-1/2} h (1 + B_n h \log n) \le C \overline{\eta}_n^2 \log n. \end{aligned}$$

Note that $|\Delta_{ns}^*(y_j)| \le 2\overline{\eta}_n$ and for each j, $\{\Delta_{ns}^*(y_j), \mathcal{F}_{n,s}^*\}$ forms a sequence of martingale differences. It follows from the martingale exponential inequality in Lemma A.1 that, there exists an $M_* > 0$ sufficiently large such that, as $n \to \infty$,

$$\begin{aligned} & \mathsf{P}(\lambda_{1n} \ge M_* \overline{\eta}_n \log n) \\ \le\ & \mathsf{P}\left(\lambda_{1n} \ge M_* \overline{\eta}_n \log n,\ \ \max_{0 \le j \le m_n} \sum_{s=1}^{T_n/2} \mathsf{E}[\Delta_{ns}^{*2}(y_j) \mid \mathcal{F}_{n,s-1}^*] \le C \overline{\eta}_n^2 \log n\right) \\ \le\ & \sum_{j=0}^{m_n} \mathsf{P}\left(\sum_{s=1}^{T_n/2} \Delta_{ns}^*(y_j) \ge M_* \overline{\eta}_n \log n,\ \ \sum_{s=1}^{T_n/2} \mathsf{E}[\Delta_{ns}^{*2}(y_j) \mid \mathcal{F}_{n,s-1}^*] \le C \overline{\eta}_n^2 \log n\right) \\ \le\ & m_n \exp\left\{-\frac{M_*^2 \log^2 n}{2C \log n + 2M_* \log n}\right\} \\ \le\ & C n^2 \exp\{-M_* \log n\} \to 0, \end{aligned} \tag{A.13}$$

where the last inequality follows from (A.4). This yields $\lambda_{1n} = O_P(\overline{\eta}_n \log n)$. Combining (A.11)–(A.13), we prove (A.6). □

Lemma A.3 *Let $f(x)$ be a real function on a compact support $[-A_1, A_1]$ satisfying $|f(x)-f(y)| \leq C|x-y|$. Under the conditions (C1) and (C3)(ii), we have*

$$\sup_{|x|\leq M_0\sqrt{n}/\log^4 n} \left| \sum_{t=1}^{n} f[(X_t+x)/h] \right| = O_P\left(\sqrt{n}h\right), \tag{A.14}$$

where M_0 is a positive constant.

(i) If, in addition, $\int_{-A_1}^{A_1} f(x)dx \neq 0$, then

$$\left(\inf_{|x|\leq M_0\sqrt{n}/\log^4 n} \left| \sum_{t=1}^{n} f[(X_t+x)/h] \right| \right)^{-1} = O_P\left[(\sqrt{n}h)^{-1}\right]. \tag{A.15}$$

(ii) If, in addition, $\int_{-A_1}^{A_1} f(x)dx = 0$, then

$$\sup_{|x|\leq B_n} \left| \sum_{t=1}^{n} f[(X_t+x)/h] \right| = O_P\left[\overline{\eta}_n \log n + (nh^2)^{1/4}\right] \tag{A.16}$$

for any $B_n \leq M_0\sqrt{n}$, where $\overline{\eta}_n$ is defined as in Lemma A.2. In particular, by letting $B_n = M_0\sqrt{n}/\log^{a_2} n$, we have

$$\sup_{|x|\leq M_0\sqrt{n}/\log^{a_2} n} \left| \sum_{t=1}^{n} f[(X_t+x)/h] \right| = O_P\left[\sqrt{n}h\log^{(4-a_2)/3} n + (nh^2)^{1/4}\right], \tag{A.17}$$

for any $a_2 \geq 0$.

Proof. Under the condition (C1) and the condition on $f(x)$, it follows from Corollary 2.2 of Wang and Phillips (2009) that

$$\frac{1}{\sqrt{n}h} \sum_{t=1}^{n} f(X_t/h) \to_D \int_{-A_1}^{A_1} f(x)dx L_W(1,0),$$

where $L_W(s,t)$ is the local time of the standard Brownian motion $W(x)$, defined by

$$L_W(s,t) = \lim_{\epsilon\to 0} \frac{1}{2\epsilon} \int_0^t I(|W(x)-s| \leq \epsilon)dx.$$

The results (A.14) and (A.15) follow easily from (A.2) and the fact that $P(L_W(1,0) > 0) = 1$. Similarly, the results (A.16) and (A.17) follows from

(A.1) and the fact that

$$\mathsf{E}\left|\sum_{t=1}^{n} f(X_t/h)\right|^2 \le C\sqrt{n}h,$$

whenever $\int_{-A_1}^{A_1} f(x)dx = 0$ (see, for example, Proposition 3.1 of Wang and Phillips, 2011). We have thus completed the proof of Lemma A.3. □

Lemma A.4 *Suppose that (i) $\{f_n(x,y)\}$ is a sequence of real functions satisfying $\sup_{x,y,n}|f_n(x,y)| < \infty$ and there exists an $\alpha > 0$ such that, whenever $|y-y_1|$ is sufficiently small,*

$$\sup_{x,y,n}|f_n(x,y) - f_n(x,y_1)| \le Cn^{\alpha}|y-y_1|; \tag{A.18}$$

(ii) there exist positive constant sequences $\gamma_n \to \infty$ and $B_n^* = O(n^k)$ for some $k > 0$ such that

$$\sup_{|y|\le B_n^*}\sum_{t=1}^{n} f_n^2(X_t,y) = O_P(\gamma_n). \tag{A.19}$$

Furthermore, suppose that the condition (C4) is satisfied. Then, for any $n\gamma_n^{-p}\log^{p-1} n = O(1)$ with p defined in (C4)(i), we have

$$\sup_{|y|\le B_n^*}\left|\sum_{t=1}^{n} u_t f_n(X_t,y)\right| = O_P\left[(\gamma_n \log n)^{1/2}\right]. \tag{A.20}$$

where $u_t = e_t$ or $|e_t| - \mathsf{E}[|e_t|]$ or $e_t^2 - 1$. Consequently, under the conditions of Theorem 1, we have, for $j = 0, 1, \cdots$,

$$\sup_{|x|\le M_0\sqrt{n}/\log^4 n}\left|\sum_{t=1}^{n} K_j[(X_t-x)/h]u_t\right| = O_P\left[(nh^2)^{1/4}\log^{1/2} n\right], \tag{A.21}$$

$$\sup_{|x|\le M_0\sqrt{n}/\log^4 n}\left|\sum_{t=1}^{n} K_j[(X_t-x)/h]\frac{\sigma(X_t)-\sigma(x)}{\sigma(x)}u_t\right| = O_P\left[h(nh^2)^{1/4}\log^{1/2} n\right], \tag{A.22}$$

where M_0 is a positive constant and $K_j(x) = x^j K(x)$.

Proof. The proof of (A.20) is similar to Theorem 2.1 of Wang and Chan (2014), and we thus omit the details. In order to prove (A.22), let

$$f_n(x,y) = \frac{\sigma(x)-\sigma(y)}{\sigma(x)}K[(x-y)/h].$$

Due to $K(x) = 0$ for $|x| \geq A$ and (2.2) in the condition (C4), it is easy to verify that $f_n(x,y)$ satisfies (A.18) and (A.19) with $\gamma_n = h^2(nh^2)^{1/2}$ and then the result (A.22) follows from (A.20). The proof of (A.21) is similar and hence the details are omitted. □

Lemma A.5 *Under the conditions (C1), (C3)(ii) and (C5), we have*

$$\begin{aligned} \Upsilon_{1n} &\equiv \sup_{|x|\leq M_0\sqrt{n}/(h\log^{c_0} n)} \left| \frac{1}{S_n^2(xh)} \sum_{t=1}^n K(X_t/h - x) - \frac{\lambda_1}{\lambda_2} \right| \\ &= O_P[\log^{-33} n], \end{aligned} \tag{A.23}$$

and for any $|s| \leq 2A$,

$$\begin{aligned} \Upsilon_{2n} &\equiv \sup_{|x|\leq M_0\sqrt{n}/(h\log^{c_0} n)} \left| \sum_{t=1}^n Z_t(s+x)Z_t(x) - r(s) \right| \\ &= O_P[\log^{-33} n], \end{aligned} \tag{A.24}$$

where $r(s) = \int K(x)K(x+s)dx/\lambda_2$ defined as in (2.3) of the main document, $c_0 > 103$, λ_1 and λ_2 are defined in Section 2 of the main document, and M_0 is a positive constant.

Proof. We only prove (A.24) as the proof of (A.23) is similar but simpler. Let

$$g_2(y) = \lambda_2\big[K(y)K(y-s) - r(s)\,K^2(y)\big]$$

and $B_n^\diamond = M_0\sqrt{n}/(h\log^{c_0} n)$. By the condition (C5), $g_2(y)$ has a compact support on $[-A, A]$ with $|g_2(x) - g_2(y)| \leq C|x-y|$ and $\int_{-A}^A g_2(y)dy = 0$. Letting

$$A_n(x, s_\diamond, s) = \frac{1}{S_n^2(xh + s_\diamond h)} \cdot \sum_{t=1}^n K(X_t/h - x - s)K(X_t/h - x)$$

with $s_\diamond = s$ or 0, it follows from Lemma A.3 with $f(x) = g_2(x)$ and $a_2 = c_0 > 103$ that, for any $|s| \leq 2A$,

$$\begin{aligned} &\sup_{|x|\leq B_n^\diamond} |A_n(x, s_\diamond, s) - r(s)| \\ &= \left[\lambda_2 \inf_{|x|\leq B_n^\diamond} \sum_{t=1}^n K^2(X_t/h - x - s_\diamond)\right]^{-1} \sup_{|x|\leq B_n^\diamond} \left|\sum_{t=1}^n g_2(X_t/h - x)\right| \\ &= O_P[\log^{-33} n]. \end{aligned} \tag{A.25}$$

Due to (A.25), we have

$$\begin{aligned}\sup_{|x|\le B_n^\diamond}\left|A_n^{1/2}(x,s_\diamond,s)-r^{1/2}(s)\right| &\le \sup_{|x|\le B_n^\diamond}\frac{\left|A_n(x,s_\diamond,s)-r(s)\right|}{\left|A_n^{1/2}(x,s_\diamond,s)+r^{1/2}(s)\right|}\\ &= O_P\left[\log^{-33} n\right]\end{aligned}$$

and thus

$$\begin{aligned}\Upsilon_{2n} &= \sup_{|x|\le B_n^\diamond}\left|A_n^{1/2}(x,s,s)A_n^{1/2}(x,0,s)-r(s)\right|\\ &\le \sup_{|x|\le B_n^\diamond}\left|\left[A_n^{1/2}(x,s,s)-r^{1/2}(s)\right]A_n^{1/2}(x,0,s)\right|+\\ &\quad \sup_{|x|\le B_n^\diamond}\left|r^{1/2}(s)\left[A_n^{1/2}(x,0,s)-r^{1/2}(s)\right]\right|\\ &= O_P[\log^{-33} n],\end{aligned}$$

as required. □

To introduce Lemma A.6, let $H(a)=H_2(a)$ be defined as in Lemma A3 of Bickel and Rosenblatt (1973). It follows that

$$\lim_{a\downarrow 0} a^{-1}H(a)=1/\sqrt{\pi}.$$

Further denote by P_ξ the conditional probability given $\xi=(\eta_k;-\infty<k<\infty)$ and $\mathsf{P}_{\mathbf{A}}(\mathbf{B})=\mathsf{P}(\mathbf{B}\cap\mathbf{A})$ for any event $\mathbf{B}$. Set

$$\begin{aligned}a_j^{(1)} &= j/(\log n)^8, \qquad 1\le j\le[(\log n)^8 w],\\ a_j^{(2)} &= ja/x_n, \qquad 1\le j\le wx_n/a,\end{aligned}$$

where $a>0$ is a constant, $x_n=d_n+z/(2\log\bar{h}^{-1})^{1/2}$ for $z\in\mathsf{R}$ and $w>0$ is a constant such that

$$\inf\{s^{-2}(1-r(s)):0\le s\le w\}>0. \tag{A.26}$$

It is clear that such w exists, and is finite due to (2.3) in the main document. Furthermore, write

$$\overline{m}=\max\left\{[(\log n)^6 w], wx_n/a\right\}$$

and

$$\mathbf{A} = \Big\{ \max_{i=1,2} \max_{1\le j\le \overline{m}} \sup_{x\in\mathcal{I}_n} \Big| \sum_{t=1}^{n} Z_t(x+a_j^{(i)}) Z_k(x) - r(a_j^{(i)}) \Big| \le (\log n)^{-33},$$
$$\inf_{x\in\mathcal{I}_n} \sum_{t=1}^{n} K(X_t/h - x) \ge \sqrt{n} h \log^{-1/2} n,$$
$$\sup_{x\in\mathcal{I}_n} \sum_{t=1}^{n} K(X_t/h - x) \le \sqrt{n} h \log^{1/2} n \Big\}. \qquad \text{(A.27)}$$

It follows from Lemmas A.3 and A.5 that $\mathsf{P}(\mathbf{A}) \to 1$ as $n \to \infty$.

Lemma A.6 *For any $a > 0$ and $0 < t \le w$ with w defined as that in (A.26), under the conditions of Theorem 1, we have*

$$P_{\mathbf{A}}\left(\bigcup_{j=1}^{[wx_n/a]} \left\{ \widetilde{M}_n(v + a_j^{(2)}) \ge x_n \right\} \right) = x_n\psi(x_n)\frac{H(a)}{a} C_0^{1/2} w + o(x_n\psi(x_n)), \qquad \text{(A.28)}$$

$$P_{\mathbf{A}}\left(\sup_{0\le s\le w} \widetilde{M}_n(v+s) > x_n \right) = \frac{1}{\sqrt{\pi}} x_n\psi(x_n) C_0^{1/2} w + o(x_n\psi(x_n)) \qquad \text{(A.29)}$$

and

$$P_{\mathbf{A}}\left(\bigcup_{j=1}^{[wx_n/a]} \left\{ \widetilde{M}_n(v + a_j^{(2)}) \ge x_n \right\}, \bigcup_{j=1}^{[wx_n/a]} \left\{ \widetilde{M}_n(v + a_j^{(2)}) < -x_n \right\} \right) = o(x_n\psi(x_n)) \qquad \text{(A.30)}$$

uniformly over $|v| \le \bar{h}^{-1} - w$, where $\psi(x) = e^{-x^2/2}/(\sqrt{2\pi})$. On the other hand, we also have

$$P_{\mathbf{A}}\left(\bigcup_{j=1}^{[wx_n/a]} \left\{ \widetilde{M}_n(v + a_j^{(2)}) < -x_n \right\} \right) = x_n\psi(x_n)\frac{H(a)}{a} C_0^{1/2} w + o(x_n\psi(x_n)) \qquad \text{(A.31)}$$

and

$$P_{\mathbf{A}}\left(\inf_{0\le s\le w} \widetilde{M}_n(v+s) < -x_n \right) = \frac{1}{\sqrt{\pi}} x_n\psi(x_n) C_0^{1/2} w + o(x_n\psi(x_n)). \qquad \text{(A.32)}$$

Proof. We only prove (A.29) as the proofs of the other results are similar. Throughout the proof, we let $t_n = [(\log n)^8 w] + 1$, $s_j = j/(\log n)^8$, $1 \le j \le t_n - 1$ and $s_{t_n} = w$. Then

$$\mathsf{P}_{\mathbf{A}}\left(\sup_{0\le s\le w} \widetilde{M}_n(v+s) > x_n \right)$$

$$\begin{aligned}
&\leq \ \mathsf{P}_{\mathbf{A}}\left(\max_{1\leq j\leq t_n} \widetilde{M}_n(v+s_j) > x_n - (\log n)^{-2}\right) + \\
&\quad\ \mathsf{P}_{\mathbf{A}}\left(\max_{1\leq j\leq t_n} \sup_{s_{j-1}<s\leq s_j} |\widetilde{M}_n(v+s) - \widetilde{M}_n(v+s_{j-1})| > (\log n)^{-2}\right) \\
&\equiv \ \Xi_{n1} + \Xi_{n2}.
\end{aligned}$$

Using the notation in Section 5 of the main document, we may write

$$\widetilde{M}_n(v+s) - \widetilde{M}_n(v+s_{j-1}) = \sum_{t=1}^{n} W_t(v,j)e_t',$$

where $W_t(v,j) = Z_t(v+s) - Z_t(v+s_{j-1})$. Note that, on $\mathbf{A}$, for any $|v| \leq \bar{h}^{-1}$ and $s_{j-1} \leq s \leq s_j$,

$$|W_t(v,j)e_t'| \leq 2(nh^2)^{-1/4}(\log n)^{3/2}$$

and

$$\begin{aligned}
\sum_{t=1}^{n} W_t(v,j)^2\, \mathsf{E}(e_t')^2 \ &\leq \ S_n^{-2}(vh+sh)\sum_{t=1}^{n}\left[K\big(X_t/h - v - s\big) - K\big(X_t/h - v - s_{j-1}\big)\right]^2 + \\
&\quad\ \sum_{t=1}^{n} K^2\big(X_t/h - v - s_{j-1}\big)\left[1/S_n(vh+sh) - 1/S_n(vh+s_{j-1}h)\right]^2 \\
&\leq \ C\,\log^{-32} n.
\end{aligned}$$

Then, as in the proof of Proposition 3 (see Appendix B below), we have $\Xi_{n2} \to 0$ as $n \to \infty$.

For Ξ_{n1}, by Theorem 1.1 in Zaïtsev (1987) and letting $\vartheta_n = x_n - 2(\log n)^{-2}$, we have

$$\begin{aligned}
&\mathsf{P}_{\xi}\left(\max_{1\leq j\leq t_n} \widetilde{M}_n(v+s_j) > x_n - (\log n)^{-2}\right) I\{\mathbf{A}\} \\
\leq \ &\mathsf{P}_{\xi}\left(\max_{1\leq j\leq t_n} Y_n(v+s_j) > \vartheta_n\right) I\{\mathbf{A}\} + C t_n^{5/2}\exp\left(-\frac{C(nh^2)^{1/4}}{t_n^{5/2}(\log n)^4}\right),
\end{aligned}$$

where $\xi = (\eta_k; k = 0, \pm 1, \pm 2, \cdots)$ and $Y_n(\cdot)$ is separable Gaussian processes with mean zero and covariance function

$$\mathsf{Cov}_{\xi}(Y_n(s_1), Y_n(s_2)) = \sum_{t=1}^{n} Z_t(s_1)Z_t(s_2).$$

Set

$$\sigma_{ij} = \frac{1}{S_n(vh + s_i h)S_n(vh + s_j h)} \sum_{t=1}^{n} K\Big(\frac{X_t}{h} - v - s_j\Big) K\Big(\frac{X_t}{h} - v - s_i\Big).$$

It is easy to see that, on $\mathbf{A}$, uniformly in v and i, j,

$$\Big|\sigma_{i,j} - r(s_j - s_i)\Big| \le C(\log n)^{-33}. \tag{A.33}$$

Therefore, using Lemma A4 in Bickel and Rosenblatt (1973), we have on $\mathbf{A}$

$$\begin{aligned} \mathsf{P}_\xi\left(\max_{1\le j\le t_n} Y_n(j) > \vartheta_n\right) &\le \mathsf{P}\left(\max_{1\le j\le t_n} \widetilde{Y}_n(s_j) > \vartheta_n\right) + \frac{Ct_n^2 e^{-\vartheta_n^2/2}}{(\log n)^{33/2}} \\ &\le \mathsf{P}\left(\max_{1\le j\le t_n} \widetilde{Y}_n(s_j) > \vartheta_n\right) + Ch(\log n)^{-1/2}, \end{aligned}$$

where $\widetilde{Y}_n(\cdot)$ is a separable stationary Gaussian processes with mean zero and covariance function $r(\cdot)$. Furthermore, by Lemma A3 in Bickel and Rosenblatt (1973), we have

$$\begin{aligned} \mathsf{P}\left(\max_{1\le j\le t_n} \widetilde{Y}_n(s_j) > \vartheta_n\right) &\le \mathsf{P}\left(\sup_{0\le s\le w} \widetilde{Y}_n(s) > \vartheta_n\right) \\ &= \frac{1}{\sqrt{\pi}} x_n \psi(x_n) C_0^{1/2} w + o(x_n\psi(x_n)). \end{aligned}$$

This implies the upper bound in (A.29) [Lemma A3 in Bickel and Rosenblatt (1973) assumes that $C_0 = 1$. For general $C_0 > 0$, one only needs to use a simple scale transform]. The lower bound in (A.29) can be obtained similarly. □

Appendix B: Proofs of Propositions 2–4

We next give the proofs of Propositions 2–4 stated in Section 5 of the main document.

Proof of Proposition 2. By the condition (C5), $B_n = \sqrt{n}/(\log^{c_0} n)$ with $c_0 > 103$, and using Lemma A.3, we have

$$\inf_{|x|\le B_n} S_n^{-2}(x) = O_P\left[(nh^2)^{-1/2}\right], \tag{B.1}$$

$$\inf_{|x|\le B_n} V_{n2}^{-1}(x) = O_P\left[(nh^2)^{-1/2}\right], \tag{B.2}$$

$$\sup_{|x|\le B_n} |V_{n1}(x)| = O_P\left[(nh^2)^{1/2}\log^{-33} n\right]. \tag{B.3}$$

Consequently, by (A.21) of Lemma A.4, we have

$$\sup_{|x|\le B_n} |\Gamma_{2n}(x)| = O_P(\log^{-32} n). \tag{B.4}$$

Similarly, it follows from (A.22) of Lemma A.4 that

$$\begin{aligned}\sup_{|x|\le B_n} |\Gamma_{3n}(x)| &\le \sup_{|x|\le B_n} \frac{1}{S_n(x)} \cdot \left|\sum_{t=1}^n K[(X_t-x)/h][\sigma(X_t)-\sigma(x)]e_t/\sigma(x)\right| + \\ &\quad \sup_{|x|\le B_n} \frac{|V_{n1}(x)|}{S_n(x)V_{n2}(x)} \cdot \left|\sum_{t=1}^n K_1[(X_t-x)/h]\frac{\sigma(X_t)-\sigma(x)}{\sigma(x)} e_t\right| \\ &= O_P(h\, \log^{1/2} n) = O_P(\log^{-2} n), \end{aligned}\tag{B.5}$$

due to the fact that $nh^{10}\log^8 n = O(1)$. For $\Gamma_{1n}(x)$, by noting

$$\sum_{t=1}^n w_t(x)(X_t - x) = 0,$$

the conditions (C2) and (C3)(iii), (B.1)–(B.3) as well as Lemma A.3, we have

$$\begin{aligned}\sup_{|x|\le B_n} |\Gamma_{1n}(x)| &= \sup_{|x|\le B_n} \frac{1}{S_n(x)V_{n2}(x)} \cdot \left|\sum_{t=1}^n w_t(x)\left[g(X_t)-g(x)-g'(x)(X_t-x)\right]\right| \\ &\le O_P\left[(nh^2)^{-1/4}\right] \sup_{|x|\le B_n} |g_0(x)| \sup_{|x|\le B_n} \frac{1}{V_{n2}(x)} \cdot \sum_{t=1}^n |w_t(x)|\cdot|X_t-x|^2 \\ &\le O_P\left[h^2(nh^2)^{-1/4}\right] \sup_{|x|\le B_n} |g_0(x)| \left[\sup_{|x|\le B_n} \sum_{t=1}^n K_2[(X_t-x)/h] + \right. \\ &\quad \left. \sup_{|x|\le B_n} \frac{|V_{n1}(x)|}{V_{n2}(x)} \cdot \sum_{t=1}^n |K_3[(X_t-x)/h]|\right] \\ &= O_P\left(\left[nh^{10} \sup_{|x|\le B_n} g_0^4(x)\right]^{1/4}\right) = O_P(\log^{-2} n). \end{aligned}\tag{B.6}$$

By (B.4)–(B.6), we complete the proof of Proposition 2. □

Proof of Proposition 3. Letting $\breve{e}_t = e_t - \hat{e}_t - \tilde{e}_t$ with $\tilde{e}_t$ defined in (5.2) of the main document and

$$\hat{e}_t = e_t I\{|e_t| \ge (nh^2)^{1/4}(\log n)^{-4}\} - \mathsf{E}\left[e_t I\{|e_t| \ge (nh^2)^{1/4}(\log n)^{-4}\}\right],$$

we have

$$M_n(x) - \widetilde{M}_n(x) = \frac{1}{S_n(xh)} \sum_{t=1}^n K(X_t/h - x)\left\{\hat{e}_t + \breve{e}_t + \tilde{e}_t\left[1 - \frac{1}{\mathsf{E}(\tilde{e}_t^2)^{1/2}}\right]\right\}.$$

Note that, by the conditions (C3)(ii), (C4)(i) and (C5),

$$\begin{aligned}
& \mathsf{E}\big[\sup_{x\in\mathcal{I}_n}|\sum_{t=1}^n K(X_t/h-x)\hat{e}_t|\big] \\
\le\ & Cn\mathsf{E}|e_1|I\{|e_1|\ge (nh^2)^{1/4}(\log n)^{-4}\} \\
\le\ & C(nh^2)^{1/4}(n^{1/2-\delta_0}h)^{-p}\, n^{1-\delta_0\, p}(\log n)^{4(2p-1)}\mathsf{E}|e_1|^{2p} \\
=\ & o\big((nh^2)^{1/4}n^{1-\delta_0\, p}(\log n)^{4(2p-1)}\big).
\end{aligned}$$

As $p > 1+[1/\delta_0]$ assumed in the condition (C4)(i), it follows that

$$\sup_{x\in\mathcal{I}_n}\left|\sum_{t=1}^n K(X_t/h-x)\hat{e}_t\right| = o_P\left((nh^2)^{1/4}\log^{-2} n\right).$$

This, together with the fact that $\mathsf{P}(\mathbf{A})\to 1$ where $\mathbf{A}$ is defined in (A.27), implies that Proposition 3 can be proved if we show that

$$\mathsf{P}_{\mathbf{A}}\left\{\sup_{x\in\mathcal{I}_n}\left|\sum_{t=1}^n K(X_t/h-x)e_t^*\right| \ge \frac{1}{2}(nh^2)^{1/4}(\log n)^{-2}\right\} \;=\; o(1), \tag{B.7}$$

where $e_t^* = \breve{e}_t + \tilde{e}_t\left[1-1/\mathsf{E}(\tilde{e}_t^2)^{1/2}\right]$.

The proof of (B.7) is similar to (A.21) by making use of the exponential inequality in Lemma A.1. Hence, we next only sketch the proof to save space. Let $\omega_j = -a_n + j\, m_n^{-1}$, $j = 1,2,\cdots,2a_n m_n$, where $a_n = h^{-1}\sqrt{n}/\log^{c_0} n$, $m_n = n^5$. Note that

$$\begin{aligned}
\mathsf{E}\big[(e_1^*)^2\big] &\le 2\mathsf{E}\big[\breve{e}_1^2\big] + 2\mathsf{E}\left\{\tilde{e}_1^2\left[1-\frac{1}{\mathsf{E}(\tilde{e}_1^2)^{1/2}}\right]\right\} \\
&\le C\mathsf{E}\left[e_1^2 I\left\{|e_1|\ge \log n\right\}\right] \le C\log^{-7} n
\end{aligned}$$

due to the condition (C4)(i) with $p > 1+[1/\delta_0] > 5$ ($\delta_0 < 1/4$), and for given $\eta_1,\eta_2,\cdots$ and conditional on $\mathbf{A}$, (e_t^*) forms a martingale difference with the conditional variance

$$\begin{aligned}
\overline{V}_n^2 &= \sum_{t=1}^n K^2(X_t/h-\omega_j)\mathsf{E}[(e_t^*)^2] \le C\sqrt{n}h\log n\mathsf{E}[(e_1^*)^2] \\
&\le C\sqrt{n}h\log n/\log^7 n = C\sqrt{n}h/\log^6 n.
\end{aligned}$$

It is readily from Lemma A.1 that

$$\mathsf{P}_{\mathbf{A}}\left\{\max_{1\le j\le m_n}|\sum_{t=1}^n K(X_t/h-\omega_j)e_t^*| \ge \frac{1}{2}(nh^2)^{1/4}(\log n)^{-2}\right\}$$

$$
\begin{aligned}
&\le\ 2\sum_{j=1}^{m_n}\mathsf{P}_{\mathbf{A}}\left\{\sum_{t=1}^{n}K(X_t/h-\omega_j)e_t^*\ge (nh^2)^{1/4}(\log n)^{-2}/2,\ \ \overline{V}_n^2\le C\sqrt{n}h/\log^6 n\right\}\\
&\le\ 2m_n\exp\big\{-C(\log n)^2\big\}=o(1),
\end{aligned}
$$

where we have used the fact that $|K(X_t/h-\omega_j)e_t^*|\le C(nh^2)^{1/4}/\log^4 n$. By recalling $|K(x)-K(y)|\le C|x-y|$, the above result and standard Taylor's expansion yield (B.7). The proof of Proposition 3 is thus completed. □

Proof of Proposition 4. Recall $x_n=d_n+z/(2\log\bar{h}^{-1})^{1/2}$ for $z\in\mathsf{R}$ and $\mathsf{P}(\mathbf{A})\to 1$, as $n\to\infty$. It suffices to show that, for any $z\in\mathsf{R}$,

$$
\left|\mathsf{P}_{\mathbf{A}}\left(\widetilde{M}_n\ge x_n\right)-(1-e^{-2e^{-z}})\right|\to 0. \tag{B.8}
$$

As in Bickel and Rosenblatt (1973), we split the interval $\mathcal{I}_n$ into $2N$ subintervals:

$$
\mathcal{W}_1,\mathcal{V}_1,\cdots,\mathcal{W}_N,\mathcal{V}_N
$$

with the length of $\mathcal{W}_i$ being $w>0$, length of $\mathcal{V}_i$ being $v>0$, and $N=\lfloor|\mathcal{I}_n|/(w+v)\rfloor$, where w is defined as in (A.26) and v is sufficiently small. We can ignore the last two incomplete intervals in view of Lemma A.6, which implies that

$$
\sup_{x\in\mathcal{I}_n}\widetilde{M}_n(x)=\max_{1\le k\le N}\sup_{x\in\mathcal{W}_k\cup\mathcal{V}_k}\widetilde{M}_n(x),
$$

and

$$
\inf_{x\in\mathcal{I}_n}\widetilde{M}_n(x)=\min_{1\le k\le N}\inf_{x\in\mathcal{W}_k\cup\mathcal{V}_k}\widetilde{M}_n(x).
$$

It can be verified that

$$
\begin{aligned}
&\Bigg|\mathsf{P}_{\mathbf{A}}\left(\sup_{x\in\mathcal{I}_n}\widetilde{M}_n(x)\ge x_n\quad\text{or}\quad\inf_{x\in\mathcal{I}_n}\widetilde{M}_n(x)<-x_n\right)-\\
&\quad\mathsf{P}_{\mathbf{A}}\left(\max_{1\le k\le N}\sup_{x\in\mathcal{W}_k}\widetilde{M}_n(x)\ge x_n\quad\text{or}\quad\min_{1\le k\le N}\inf_{x\in\mathcal{W}_k}\widetilde{M}_n(x)<-x_n\right)\Bigg|\\
\le\ &\mathsf{P}_{\mathbf{A}}\left(\max_{1\le k\le N}\sup_{x\in\mathcal{V}_k}\widetilde{M}_n(x)\ge x_n\right)+\mathsf{P}_{\mathbf{A}}\left(\max_{1\le k\le N}\inf_{x\in\mathcal{V}_k}\widetilde{M}_n(x)<-x_n\right)\\
\equiv\ &\mathcal{R}_{n1}+\mathcal{R}_{n2}.
\end{aligned}
$$

Without loss of generality, assume that $\mathcal{W}_k=[a_k,a_k+w)$. We further have

$$
\Bigg|\mathsf{P}_{\mathbf{A}}\left(\max_{1\le k\le N}\sup_{x\in\mathcal{W}_k}\widetilde{M}_n(x)\ge x_n\quad\text{or}\quad\min_{1\le k\le N}\inf_{x\in\mathcal{W}_k}\widetilde{M}_n(x)<-x_n\right)-
$$

$$
\begin{aligned}
& \mathsf{P}_{\mathbf{A}}\left(\bigcup_{k=1}^{N} \bigcup_{j=1}^{[wx_n/a]} \left\{ \widetilde{M}_n(a_k + a_j^{(2)}) \geq x_n \right\} \quad \text{or} \quad \bigcup_{k=1}^{N} \bigcup_{j=1}^{[wx_n/a]} \left\{ \widetilde{M}_n(a_k + a_j^{(2)}) < -x_n \right\} \right) \Bigg| \\
\leq \quad & \sum_{k=1}^{N} \left| \mathsf{P}_{\mathbf{A}}\left(\sup_{x \in \mathcal{W}_k} \widetilde{M}_n(x) \geq x_n \right) - \mathsf{P}_{\mathbf{A}}\left(\bigcup_{j=1}^{[wx_n/a]} \left\{ \widetilde{M}_n(a_k + a_j^{(2)}) \geq x_n \right\} \right) \right| + \\
& \sum_{k=1}^{N} \left| \mathsf{P}_{\mathbf{A}}\left(\inf_{x \in \mathcal{W}_k} \widetilde{M}_n(x) < -x_n \right) - \mathsf{P}_{\mathbf{A}}\left(\bigcup_{j=1}^{[wx_n/a]} \left\{ \widetilde{M}_n(a_k + a_j^{(2)}) < -x_n \right\} \right) \right| \\
\equiv \quad & \mathcal{R}_{n3} + \mathcal{R}_{n4}.
\end{aligned}
$$

Recalling $\lim_{a\downarrow 0} a^{-1}H(a) = 1/\sqrt{\pi}$, $N \leq C_1\bar{h}^{-1}$ and $x_n\psi(x_n) \leq C_2\bar{h}$, where C_1 and C_2 are two positive constants, it follows easily from Lemma A.6 that

$$
\begin{aligned}
\lim_{v\to 0} \limsup_{n\to\infty} (\mathcal{R}_{n1} + \mathcal{R}_{n2}) &= 0, \\
\lim_{a\to 0} \limsup_{v\to 0} \limsup_{n\to\infty} (\mathcal{R}_{n3} + \mathcal{R}_{n4}) &= 0.
\end{aligned}
$$

Combining the above facts, the result (B.8) follows if we prove, for any $z \in \mathsf{R}$,

$$
\lim_{a\to 0} \limsup_{v\to 0} \limsup_{n\to\infty} \left| \mathsf{P}_{\mathbf{A}}\left(\bigcup_{k=1}^{N} \mathbf{A}_k \right) - (1 - e^{-2e^{-z}}) \right| = 0, \tag{B.9}
$$

where $\mathbf{A}_k = \bigcup_{j=1}^{[wx_n/a]} \mathbf{B}_{k,j}$ and

$$
\mathbf{B}_{k,j} = \left\{ \widetilde{M}_n(a_k + a_j^{(2)}) \geq x_n \right\} \cup \left\{ \widetilde{M}_n(a_k + a_j^{(2)}) \leq -x_n \right\}.
$$

We next give the proof of (B.9), which is similar to the relevant argument in Liu and Wu (2010). Let

$$
\begin{aligned}
\mathbf{D}_{k,j} &= \left\{ Y_n(a_k + a_j^{(2)}) \geq x_n \right\} \cup \left\{ Y_n(a_k + a_j^{(2)}) \leq -x_n \right\}, \\
\mathbf{D}_{k,j}^{\pm} &= \left\{ Y_n(a_k + a_j^{(2)}) \geq x_n \pm (\log n)^{-2} \right\} \cup \left\{ Y_n(a_k + a_j^{(2)}) \leq -x_n - (\pm(\log n)^{-2}) \right\}, \\
\widetilde{\mathbf{D}}_{k,j} &= \left\{ \widetilde{Y}_n(a_k + a_j^{(2)}) \geq x_n \right\} \cup \left\{ \widetilde{Y}_n(a_k + a_j^{(2)}) \leq -x_n \right\}, \\
\widetilde{\mathbf{D}}_{k,j}^{\pm} &= \left\{ \widetilde{Y}_n(a_k + a_j^{(2)}) \geq x_n \pm (\log n)^{-2} \right\} \cup \left\{ \widetilde{Y}_n(a_k + a_j^{(2)}) \leq -x_n - (\pm(\log n)^{-2}) \right\},
\end{aligned}
$$

where, conditioning on $\xi = (\eta_k; k \in Z)$, $Y_n(\cdot)$ is separable Gaussian processes with mean zero and covariance function

$$
\mathsf{Cov}_{\xi}(Y_n(s_1), Y_n(s_2)) = \sum_{t=1}^{n} Z_t(s_1)Z_t(s_2),
$$

and $\widetilde{Y}_n(\cdot)$ is separable stationary Gaussian processes with mean zero and covariance function $r(s) = \int K(x)K(x+s)dx/\lambda(K^2)$.

By Bonferroni's inequality, we have for fixed $1 \le l \le N/2$,

$$\begin{aligned}\sum_{d=1}^{2l}(-1)^{d-1} \sum_{1\le i_1<\cdots<i_d\le N} \mathsf{P}_{\mathbf{A}}\left(\bigcap_{j=1}^{d} \mathbf{A}_{i_j}\right) &\le \mathsf{P}_{\mathbf{A}}\left(\bigcup_{k=1}^{N} \mathbf{A}_k\right) \\ \le \sum_{d=1}^{2l-1}(-1)^{d-1} \sum_{1\le i_1<\cdots<i_d\le N} \mathsf{P}_{\mathbf{A}}\left(\bigcap_{j=1}^{d} \mathbf{A}_{i_j}\right).\end{aligned}$$

To estimate the probability $\mathsf{P}_{\mathbf{A}}\left(\bigcap_{j=1}^{d} \mathbf{A}_{i_j}\right)$, recall $\mathcal{W}_k = [a_k, a_k + w)$. Let $q_j = i_{j+1} - i_j$ for $1 \le j \le d-1$ and define

$$\Sigma' = \left\{1 \le i_1 < \cdots < i_d \le N : \min_{1\le j\le d-1} q_j \le \lfloor 2/w + 2\rfloor\right\}. \tag{B.10}$$

Let $0 \le d_0 \le d-2$ and

$$\Sigma'_{d_0} = \left\{1 \le i_1 < \cdots < i_d \le N : \text{the number of } j \text{ such that } q_j > \lfloor 2/w+2\rfloor \text{ is } d_0\right\}.$$

We have $\Sigma' = \cup_{d_0=0}^{d-2}\Sigma'_{d_0}$ and the number of elements in the sum $\sum_{\Sigma'_{d_0}} \mathsf{P}_{\mathbf{A}}\left(\bigcap_{j=1}^{d} \mathbf{A}_{i_j}\right)$ is bounded by $CN^{d_0+1} = O(\bar{h}^{-d_0-1})$, where C is a positive constant independent of N. Suppose now $i_1, \cdots, i_d$ are in Σ'_{d_0}. Note that

$$\bigcap_{j=1}^{d} \mathbf{A}_{i_j} = \bigcup_{j_1=1}^{[wx_n/a]} \cdots \bigcup_{j_d=1}^{[wx_n/a]} \left\{\mathbf{B}_{i_1,j_1} \cap \cdots \cap \mathbf{B}_{i_d,j_d}\right\}.$$

Without loss of generality, assume that $q_1 \le \lfloor 2/w+2\rfloor$, $q_2 > \lfloor 2/w + 2\rfloor, \cdots, q_{d_0+1} > \lfloor 2/w+2\rfloor$. By Theorem 1.1 in Zaïtsev (1987), on $\mathbf{A}$, we have

$$\mathsf{P}_\xi\left(\mathbf{B}_{i_1,j_1} \cap \cdots \cap \mathbf{B}_{i_d,j_d}\right) \le \mathsf{P}_\xi\left(\mathbf{D}^-_{i_1,j_1} \cap \cdots \cap \mathbf{D}^-_{i_d,j_d}\right) + C\exp\left\{-(\log \bar{h}^{-1})^2\right\}. \tag{B.11}$$

Set $\sigma_{lk} = \sum_{t=1}^{n} Z_t(a_{i_l} + a^{(2)}_{j_l})Z_t(a_{i_k} + a^{(2)}_{j_k})$. Recalling the definition of $Z_t(s)$, we have, on $\mathbf{A}$, for $3 \le k \le d_0+1$, $l = 1, 2$, $\sigma_{lk} = 0$; for $3 \le k \ne s \le d_0+1$, $\sigma_{sk} = 0$; for $1 \le k \le d_0+1$, $\sigma_{kk} = 1$; and

$$\left|\sigma_{12} - r\left(a_{i_2} - a_{i_1} + (j_2 - j_1)ax_n^{-1}\right)\right| \le C(\log n)^{-33}.$$

For notational simplicity, we set

$$\mathbf{Y}_k = Y_n(a_{i_k} + a^{(2)}_{j_k}), \quad 1 \le k \le d, \ \ 1 \le j_k \le [wx_n/a].$$

Using the estimations of the covariance above, we get the covariance matrix $\mathbf{V}_n$ of $(\mathbf{Y}_k)_{1\le k\le d_0+1}$ satisfying

$$\Big|\mathbf{V}_n - \mathbf{V}\Big| \le C(\log n)^{-33}, \tag{B.12}$$

where

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_1 & 0 \\ 0 & \mathbf{I}_{d_0-1} \end{pmatrix}, \quad \mathbf{V}_1 = \begin{pmatrix} 1 & \mu \\ \mu & 1 \end{pmatrix},$$

$\mu = r\left(a_{i_2} - a_{i_1} + (j_2 - j_1)ax_n^{-1}\right)$. By (B.12), we have

$$\left|\mathbf{V}_n^{-1} - \mathbf{V}^{-1}\right| \le C(\log n)^{-33}, \quad \left|\sqrt{\det(\mathbf{V})} - \sqrt{\det(\mathbf{V}_n)}\right| \le C(\log n)^{-33}. \tag{B.13}$$

Let $p_n(y)$ denote the density function of $(\mathbf{Y}_k)_{1\le k\le d_0+1}$, and $p(y)$ denote the density function of the Gaussian random vector with covariance matrix $\mathbf{V}$. Then, by (B.13), we have

$$|p_n(y) - p(y)| \le C\left[(\log n)^{-2}p(y) + \exp\left(-y\mathbf{V}^{-1}y'/2\right)\left|\exp\left\{C(\log n)^{-2}|y|^2\right\} - 1\right|\right]. \tag{B.14}$$

Note that $|j_2 - j_1|ax_n^{-1} \le w$ and $a_{j_2} - a_{j_1} \ge w + v$. It is readily seen that $|\mu| \le \sup_{t\ge v}|r(t)| < 1$. Then it follows from Lemma 2 in Berman (1962) that, for some $\delta > 0$, we have

$$\begin{aligned} & \mathsf{P}\left(\mathbf{D}^-_{i_1,j_1} \cap \cdots \cap \mathbf{D}^-_{i_d,j_d}\right) \\ \le\ & C(1 + (\log n)^{-2}) \int_{\Xi^-} p(y)dy + C\exp\left\{-(\log \bar{h}^{-1})^2\right\} \\ \le\ & C\bar{h}^{d_0+1+\delta}, \end{aligned} \tag{B.15}$$

where $y = (y_1, \cdots y_{d_0+1})$ and

$$\Xi^{\pm} = \bigcap_{j=1}^{d_0+1}\left[\left\{y_j \ge x_n + (\pm\log n)^{-2d}\right\} \cup \left\{y_j \le -x_n - (\pm\log n)^{-2d}\right\}\right].$$

Noting that $([wx_n/a])^d = O((\log n)^{2d})$ and by (B.11), we have for some $\delta > 0$,

$$\sum_{d_0=0}^{d-2}\sum_{\Sigma'_{d_0}} \mathsf{P}_{\mathbf{A}}\left(\bigcap_{j=1}^{d}\mathbf{A}_{i_j}\right) \le C\bar{h}^{\delta}.$$

We now estimate

$$\left(\sum_{1\le i_1<\cdots<i_d\le N} - \sum_{\Sigma'}\right)\mathsf{P}_{\mathbf{A}}\left(\bigcap_{j=1}^{d}\mathbf{A}_{i_j}\right). \tag{B.16}$$

Since $i_{j+1}-i_j \ge \lfloor 2/w+2\rfloor$, we have $a_{i_{j+1}}-a_{i_j} \ge (w+v)(\lfloor 2/w+2\rfloor) > 2+w+v$. Then,

$$\sigma_{sk}=0 \ \text{ for } \ 1\le s\ne k\le d,\ 1\le j_s, j_k \le [wx_n/a]. \tag{B.17}$$

By Theorem 1.1 in Zaïtsev (1987), on $\mathbf{A}$, we have

$$\mathsf{P}_\xi\left(\bigcap_{j=1}^{d}\mathbf{A}_{i_j}\right) \le \mathsf{P}_\xi\left(\bigcap_{j=1}^{d}\mathbf{D}_{i_j}^{-}\right) + C\exp\left\{-(\log \bar{h}^{-1})^2\right\} \tag{B.18}$$

and

$$\mathsf{P}_\xi\left(\bigcap_{j=1}^{d}\mathbf{A}_{i_j}\right) \ge \mathsf{P}_\xi\left(\bigcap_{j=1}^{d}\mathbf{D}_{i_j}^{+}\right) - C\exp\left\{-(\log \bar{h}^{-1})^2\right\}. \tag{B.19}$$

By (B.17),

$$\mathsf{P}_\xi\left(\bigcap_{j=1}^{d}\mathbf{D}_{i_j}^{\pm}\right) = \prod_{j=1}^{d}\mathsf{P}_\xi(\mathbf{D}_{i_j}^{\pm})$$

By Lemma A4 in Bickel and Rosenblatt (1973), on $\mathbf{A}$ we have

$$\begin{aligned}\left|\mathsf{P}_\xi(\mathbf{D}_{i_j}^{\pm}) - \mathsf{P}(\widetilde{\mathbf{D}}_{i_j}^{\pm})\right| &\le Ct_n^2(\log n)^{-33/2}\exp\left\{-(x_n \pm (\log n)^{-2})^2/2\right\}\\ &\le Cht_n^2(\log n)^{-33/2}, \end{aligned} \tag{B.20}$$

where $\widetilde{\mathbf{D}}_{i_j}^{\pm}$ is defined as

$$\left[\bigcup_{k=1}^{[wx_n/a]}\left\{\widetilde{Y}_n(a_{i_j}+a_k^{(2)}) > x_n \pm (\log n)^{-2}\right\}\right]\bigcup\left[\bigcup_{k=1}^{[wx_n/a]}\left\{\widetilde{Y}_n(a_{i_j}+a_k^{(2)}) < -x_n - (\pm(\log n)^{-2})\right\}\right].$$

By Lemma A3 in Bickel and Rosenblatt (1973), we have

$$\mathsf{P}(\widetilde{\mathbf{D}}_{i_j}^{\pm}) = (2+o(1))x_n\psi(x_n)\frac{H_\alpha(a)}{a}C_0^{1/2}w. \tag{B.21}$$

Hence, by (B.20) and (B.21),

$$\mathsf{P}_\xi(\mathbf{D}_{i_j}^{\pm})I\{\mathbf{A}\} = (2+o(1))x_n\psi(x_n)\frac{H_2(a)}{a}C_0^{1/2}w\cdot I\{\mathbf{A}\}$$

$$= \quad (2+o(1))\sqrt{\pi} h w e^{-z} \frac{H_2(a)}{a} \cdot I\{\mathbf{A}\},$$

where $o(1)$ is bounded by an non-random number which tends to zero. Hence, we have

$$\begin{aligned}
& \left(\sum_{1\le i_1<\cdots<i_d\le N} - \sum_{\Sigma'} \right) \mathsf{P}_{\mathbf{A}} \left(\bigcap_{j=1}^{d} \mathbf{A}_{i_j} \right) \\
= \quad & \left(\sum_{1\le i_1<\cdots<i_d\le N} - \sum_{\Sigma'} \right) \mathsf{E} \left[\mathsf{P}_{\xi} \left(\bigcap_{j=1}^{d} \mathbf{A}_{i_j} \right) I\{\mathbf{A}\} \right] \\
= \quad & \left[(2+o(1)) N \sqrt{\pi} h w e^{-z} \frac{H_2(a)}{a} \right]^d \mathsf{P}(\mathbf{A})/d! \\
= \quad & \left[(2+o(1))(w+v)^{-1} w e^{-z} \sqrt{\pi} \frac{H_2(a)}{a} \right]^d \mathsf{P}(\mathbf{A})/d!,
\end{aligned}$$

which implies the result (B.9), by letting $n \to \infty$, $l \to \infty$, $v \to 0$ and $a \to 0$. $\square$

Appendix C: Asymptotic Gumbel distribution for the Nadaraya-Watson estimator

We next consider using the Nadaraya-Watson kernel smoothing method to estimate the regression function g in model (1.2), develop the asymptotic Gumbel distribution for the maximum deviation of the estimator with appropriate centering and scaling, and then give a brief discussion on how to construct the SCBs for the function g.

The Nadaraya-Watson estimation is defined as

$$\widetilde{g}_n(x) = \sum_{t=1}^{n} K\left[(X_t - x)/h\right] Y_t / \sum_{t=1}^{n} K\left[(X_t - x)/h\right], \tag{C.1}$$

where K is a non-negative kernel function and h is a bandwidth. To construct the SCBs, we have to obtain the asymptotic distribution for the normalized maximum absolute deviation

$$\widetilde{\Delta}_n = \sup_{|x|\le B_n} \left| \widetilde{V}_n(x) \left[\widetilde{g}_n(x) - g(x) - b_g(x)\right] / \sigma(x) \right|, \tag{C.2}$$

where B_n is a sequence of constants which may diverge to infinity,

$$b_g(x) = \sum_{t=1}^{n} g(X_t) K\left[(X_t - x)/h\right] / \sum_{t=1}^{n} K\left[(X_t - x)/h\right] - g(x) \qquad \text{(C.3)}$$

is the bias term for $\widetilde{g}_n(\cdot)$ and the normalizing term

$$\widetilde{V}_n(x) = \left\{\sum_{t=1}^{n} K\left[(X_t - x)/h\right]\right\} / \left\{\sum_{t=1}^{n} K^2\left[(X_t - x)/h\right]\right\}^{1/2}. \qquad \text{(C.4)}$$

Similarly to the proofs in Section 5 of the main document, we may derive the following asymptotic Gumbel distribution.

Theorem C.1. *Let Conditions (C1) (C3)(i)(ii), (C4) and (C5) be satisfied. Then, for* $z \in R$,

$$P\left\{(2\log \bar{h}^{-1})^{1/2}\left(\widetilde{\Delta}_n - d_n\right) \le z\right\} \to e^{-2e^{-z}}, \qquad \text{(C.5)}$$

where $\bar{h}$ *and* d_n *are defined as in Theorem 1.*

In order to use the above theorem to construct the SCBs of the unknown regression function g, we need to consistently estimate the variance function $\sigma^2(x)$ and the asymptotic bias term $b_g(x)$. The function $\sigma^2(x)$ can be estimated by using (3.1) in the main document with $\widehat{g}_n(\cdot)$ replaced by $\widetilde{g}_n(\cdot)$, and we denote the resulting estimate by $\widetilde{\sigma}_n^2(x)$. Furthermore, under some smoothness condition on g, we may further show that

$$\sup_{x \le B_n} \left|b_g(x) - b_g^{\diamond}(x)\right| = o_P(h^2), \qquad \text{(C.6)}$$

where $b_g^{\diamond}(x) = g'(x)\frac{V_{n1}(x)}{V_{n0}(x)}h + \frac{1}{2}g''(x)\frac{V_{n2}(x)}{V_{n0}(x)}h^2$, g' and g'' are the first and second-order derivatives of g, and $V_{nj}(x)$ is defined in Section 1 of the main document. Note that g' and g'' can be consistently estimated by using the local cubic smoothing method, and then we can obtain the estimation of $b_g^{\diamond}(x)$, which is denoted by $\widetilde{b}_g(x)$. Using Theorem C.1 and (C.6), for given α, the $(1-\alpha)$-SCB for g over the set $\{x : |x| \le B_n\}$ can be constructed by

$$\left[\widetilde{g}_n(x) - \widetilde{b}_g(x) - \widetilde{l}_\alpha(x),\ \ \widetilde{g}_n(x) - \widetilde{b}_g(x) + \widetilde{l}_\alpha(x)\right], \qquad \text{(C.7)}$$

where

$$\widetilde{l}_\alpha(x) = \left[z_\alpha(2\log \bar{h}^{-1})^{-1/2} + d_n\right] \widetilde{\sigma}_n(x)\, \widetilde{V}_n^{-1}(x).$$

and z_α is defined as in Section 3 of the main document. However, Chan and Wang (2014) showed that the performance of the local linear estimation is superior to that of the conventional Nadaraya-Watson estimator in uniform asymptotics for nontationary time series. Therefore, in the present paper, we concentrate on the SCBs using the local linear smoothing method for the regression function g.

REFERENCES

Bickel, P. J. and Rosenblatt, M. (1973). On some global measures of the deviations of density function estimates. *The Annals of Statistics* **1**, 1071–1095.

Berman, S. (1962). A law of large numbers for the maximum of a stationary Gaussian sequence. *The Annals of Mathematical Statistics* **33** 93–97.

Chan, N. and Wang, Q. (2014). Uniform convergence for nonparametric estimators with non-stationary data. *Econometric Theory* **30**, 1110–1133.

Liu, W. and Wu, W. B. (2010). Simultaneous nonparametric inference of time series. *The Annals of Statistics* **38**, 2388–2421.

de la Peña, V. H. (1999). A general class of exponential inequalities for martingales and ratios. *The Annals of Probability* **27**, 537–564.

Wang, Q. and Chan, N. (2014). Uniform convergence for a class of martingales with applications in non-linear cointegrating regression. *Bernoulli* **20**, 207–230.

Wang, Q. and Phillips, P. C. B. (2009). Asymptotic theory for local time density estimation and nonparametric cointegrating regression. *Econometric Theory* **25**, 710–738.

Wang, Q. and Phillips, P. C. B. (2011). Asymptotic theory for zero energy functionals with nonparametric regression applications. *Econometric Theory* **27**, 235–259.

Zaïtsev, A. Yu. (1987). On the Gaussian approximation of convolutions under multidimensional analogues of S.N. Bernstein's inequality conditions. *Probability Theory and Related Fields* **74**, 535–566.