Technical supplement to Nonparametric estimation of conditional value-at-risk and expected shortfall based on extreme value theory

Carlos Martins-Filho

Department of Economics
University of Colorado
Boulder, CO 80309-0256, USA
email: carlos.martins@colorado.edu
Voice: + 1 303 492 4599

&

IFPRI
2033 K Street NW
Washington, DC 20006-1002, USA
email: c.martins-filho@cgiar.org
Voice: + 1 202 862 8144

Feng Yao

Department of Economics
West Virginia University
Morgantown, WV 26505, USA
email: feng.yao@mail.wvu.edu
Voice: +1 304 293 7867

Maximo Torero

IFPRI
2033 K Street NW
Washington, DC 20006-1002, USA
email: m.torero@cgiar.org
Voice: + 1 202 862 5662

December, 2013

Abstract. This supplement contains statements and proofs for Lemmas 1, 2, 3 and 4 in Martins-Filho et al. (2013). In addition, there are statements for Corollary 1, Lemmas 5 and 6.

Keywords and phrases. Value-at-risk, expected shortfall, extreme value theory, nonparametric location-scale models, strong mixing.

JEL classifications. C14, C15, C22, G10.

AMS-MS classifications. Primary: 62G32, 62G07, 62G08, 62G20.

1 Lemmas and proofs

Lemma 1. *Let* $w(\mathbf{X}_t - \mathbf{x};\mathbf{x}) : \mathbb{R}^d \to \mathbb{R}$ *and* $g(\varepsilon) : \mathbb{R} \to \mathbb{R}$ *be measurable functions and define*

$$s(\mathbf{x}) = \frac{1}{nh_n^d}\sum_{t=1}^{n} K\left(\frac{\mathbf{X}_t - \mathbf{x}}{h_n}\right)\left(\frac{X_{ti} - x_i}{h_n}\right)^{p_1}\left(\frac{X_{tj} - x_j}{h_n}\right)^{p_2}\left(\frac{X_{tl} - x_l}{h_n}\right)^{p_3} w(\mathbf{X}_t - \mathbf{x};\mathbf{x})g(\varepsilon_t) \tag{1}$$

where K *is a multivariate kernel given by* $K(\mathbf{x}) = \prod_{j=1}^{d}\mathcal{K}(x_j)$, $h_n > 0$ *is a bandwidth, for* $i, j = 1, \cdots, d$ *and* $p_1, p_2, p_3 = 0, 1$. *Assume that A1 and A2 are holding and that:*

a) $E\left(|g(\varepsilon_t)|^{\zeta}\right) < \infty$ *for some* $\zeta > 2$*;*

b) $w(\mathbf{X}_t - \mathbf{x};\mathbf{x})$ *satisfies a Lipschitz condition of order 1, i.e.,* $|w(\mathbf{X}_t - \mathbf{x};\mathbf{x}) - w(\mathbf{X}_t - \mathbf{x}^k;\mathbf{x}^k)| \leq C\|\mathbf{x} - \mathbf{x}^k\|_E$ *for some* $C > 0$ *and* $\mathbf{x} \neq \mathbf{x}^k$ *in* $\mathbb{R}^d$ *and* $|w(\mathbf{X}_t - \mathbf{x};\mathbf{x})| < C$ *for all* $x \in \mathbb{R}^d$*;*

c) The joint density of $\mathbf{X}_i$ *and* $\mathbf{X}_j$ *conditional on* ε_i *and* ε_j *denoted by* $f_{\mathbf{X}_i\mathbf{X}_j|\varepsilon_i\varepsilon_j}(\mathbf{X}_i, \mathbf{X}_j) < C$.

Then, for an arbitrary compact set $\mathcal{G} \subseteq \mathbb{R}^d$*, we have*

$$\sup_{\mathbf{x}\in\mathcal{G}}|s(\mathbf{x}) - E(s(\mathbf{x}))| = O_p\left(\left(\frac{log\, n}{nh_n^d}\right)^{1/2}\right) \tag{2}$$

provided that for $\zeta, B > 2$, $\theta > 0$*, we have*

$$n^{1-\frac{2}{\zeta}-2\theta}h_n^d \to \infty \tag{3}$$

and

$$n^{(B+1.5)(\frac{1}{\zeta}+\theta)-\frac{B}{2}+0.75+\frac{d}{2}}h_n^{-1.75d-\frac{d}{2}(d+B)}(log\, n)^{0.25+0.5(B-d)} \to 0. \tag{4}$$

Proof. We first establish the result for the case where $p_1 = p_2 = p_3 = 0$. The proof follows Martins-Filho and Yao (2009) and has three steps: (1) we show that $\sup_{\mathbf{x}\in\mathcal{G}}|s(\mathbf{x}) - E(s(\mathbf{x}))| \leq \max_{1\leq k\leq l_n}|s_{\mathbf{0}}(\mathbf{x}^k) - E(s_{\mathbf{0}}(\mathbf{x}^k))| + 2\,C\left(\frac{log\, n}{nh_n^d}\right)^{1/2}$ for a suitably defined sequence l_n and $\mathbf{x}^k \in \mathcal{G}$; (2) we show that $\sup_{\mathbf{x}\in\mathcal{G}}|s(\mathbf{x}) - s^{\tau}(\mathbf{x}) - E(s(\mathbf{x}) - s^{\tau}(\mathbf{x}))| = O_{as}(B_n^{1-\zeta})$, where

$$s^{\tau}(\mathbf{x}) = (nh_n^d)^{-1}\sum_{t=1}^{n} K\left(\frac{\mathbf{X}_t - \mathbf{x}}{h_n}\right) w(\mathbf{X}_t)g(\varepsilon_t)\chi_{\{|g(\varepsilon_t)|\leq B_n\}}$$

with $B_1 \leq B_2 \leq ...$ such that $\sum_{t=1}^{\infty} B_t^{-\zeta} < \infty$ for some $\zeta > 0$; (3) we show that for $0 < \Delta < \infty$, $B > 2$ and $\varepsilon_n = \left(\frac{nh_n^d}{log\, n}\right)^{-1/2}\Delta$, $P\left(\max_{1\leq k\leq l_n}\left|s^{\tau}(\mathbf{x}^k) - E(s^{\tau}(\mathbf{x}^k))\right| \geq \varepsilon_n\right) = O\left(d_n\right)$ where

$$d_n = n^{(B+1.5)(\frac{1}{\zeta}+\theta)-B/2+0.75+\frac{d}{2}}h_n^{-1.75d-\frac{d}{2}(d+B)}(log\, n)^{0.25+0.5(B-d)}.$$

Let $B(\mathbf{x}_0, r) = \{\mathbf{x} \in \mathbb{R}^d : \|\mathbf{x} - \mathbf{x}_0\|_E < r\}$ for $r \in \mathbb{R}^+$. $\mathcal{G}$ compact implies that there exists $\mathbf{x}_0 \in \mathbb{R}^d$ such that $\mathcal{G} \subseteq B(\mathbf{x}_0, r)$. Therefore, for all $\mathbf{x}, \mathbf{z} \in \mathcal{G}$, $\|\mathbf{x} - \mathbf{z}\|_E < 2r$. Let $h_n > 0$ be such that $h_n \to 0$ as $n \to \infty$ where $n \in \{1, 2, 3 \cdots\}$. For any n, by the Heine-Borel Theorem, every infinite cover for $\mathcal{G}$ contains a finite subcover $\left\{B\left(\mathbf{x}^k, C\left(\frac{n}{h_n^{d+2} log\, n}\right)^{-1/2}\right)\right\}_{k=1}^{l_n}$ with $\mathbf{x}^k \in \mathcal{G}$ and $l_n \leq \left(\frac{n}{h_n^{d+2} log\, n}\right)^{d/2}$. Step (1): For $\mathbf{x} \in B\left(\mathbf{x}^k, C\left(\frac{n}{h_n^{d+2} log\, n}\right)^{-1/2}\right)$,

$$
\begin{aligned}
|s(\mathbf{x}) - s(\mathbf{x}^k)| = & \left| \frac{1}{nh_n^d} \sum_{t=1}^{n} \left(K\left(\frac{\mathbf{X}_t - \mathbf{x}}{h_n}\right) w(\mathbf{X}_t - \mathbf{x}; \mathbf{x}) - K\left(\frac{\mathbf{X}_t - \mathbf{x}^k}{h_n}\right) w(\mathbf{X}_t - \mathbf{x}; \mathbf{x}) \right.\right. \\
& + \left.\left. K\left(\frac{\mathbf{X}_t - \mathbf{x}^k}{h_n}\right) w(\mathbf{X}_t - \mathbf{x}; \mathbf{x}) - K\left(\frac{\mathbf{X}_t - \mathbf{x}^k}{h_n}\right) w(\mathbf{X}_t - \mathbf{x}^k; \mathbf{x}^k) \right) g(\varepsilon_t) \right| \\
& \leq \frac{1}{nh_n^d} \sum_{t=1}^{n} \left(\left| K\left(\frac{\mathbf{X}_t - \mathbf{x}}{h_n}\right) - K\left(\frac{\mathbf{X}_t - \mathbf{x}^k}{h_n}\right) \right| |w(\mathbf{X}_t - \mathbf{x}; \mathbf{x})| \right. \\
& + \left. \left| K\left(\frac{\mathbf{X}_t - \mathbf{x}^k}{h_n}\right) \right| |w(\mathbf{X}_t - \mathbf{x}; \mathbf{x}) - w(\mathbf{X}_t - \mathbf{x}^k; \mathbf{x}^k)| \right) |g(\varepsilon_t)| \\
& \leq \frac{C}{h_n^{d+1}} \left\|\mathbf{x}^k - \mathbf{x}\right\|_E \frac{1}{n} \sum_{t=1}^{n} |g(\varepsilon_t)| \text{ by A1 1) and 4) and condition b)} \qquad (5) \\
& \leq \frac{C}{nh_n^{d+1}} \left(\frac{n}{h_n^{d+2} log\, n}\right)^{-1/2} \sum_{t=1}^{n} |g(\varepsilon_t)| = C \left(\frac{log\, n}{nh_n^d}\right)^{1/2} \frac{1}{n} \sum_{t=1}^{n} |g(\varepsilon_t)|
\end{aligned}
$$

By the measurability of g and assumption A2 1) we have that $\{|g(\varepsilon_t)|\}_{t=1,2,\ldots}$ is α-mixing of size -2. By condition a) and McLeish's LLN (White (2001), p. 49) $\frac{1}{n}\sum_{t=1}^{n}(|g(\varepsilon_t)| - E(|g(\varepsilon_t)|) = o_p(1)$ and since $E(|g(\varepsilon_t)|) < C$ we have $|s(\mathbf{x}) - s(\mathbf{x}^k)| \leq C\left(\frac{log\, n}{nh_n^d}\right)^{1/2}$. Following similar arguments it is easily verified that $E(|s(\mathbf{x}) - s(\mathbf{x}^k)|) \leq C\left(\frac{log\, n}{nh_n^d}\right)^{1/2}$. Combining these two bounds, $\sup_{\mathbf{x}\in\mathcal{G}}|s_{\mathbf{0}}(\mathbf{x}) - E(s_{\mathbf{0}}(\mathbf{x}))| \leq \max_{1\leq k\leq l_n}|s_{\mathbf{0}}(\mathbf{x}^k) - E(s_{\mathbf{0}}(\mathbf{x}^k))| + 2\, C\left(\frac{log\, n}{nh_n^d}\right)^{1/2}$. Step (2): $\sup_{\mathbf{x}\in\mathcal{G}}|s(\mathbf{x}) - s^\tau(\mathbf{x}) - E(s(\mathbf{x}) - s^\tau(\mathbf{x}))| \leq T_1 + T_2$, where $T_1 = \sup_{\mathbf{x}\in\mathcal{G}}|s(\mathbf{x}) - s^\tau(\mathbf{x})|$ and $T_2 = \sup_{\mathbf{x}\in\mathcal{G}}|E(s(\mathbf{x}) - s^\tau(\mathbf{x}))|$. We show that $T_1 = o_{as}(1)$ and $T_2 = O(B_n^{1-\zeta})$ for $\zeta > 1$. Note that $T_1 = \sup_{\mathbf{x}\in\mathcal{G}}\left|(nh_n^d)^{-1}\sum_{t=1}^{n} K\left(\frac{\mathbf{X}_t - \mathbf{x}}{h_n}\right) g(\varepsilon_t) w(\mathbf{X}_t - \mathbf{x}; \mathbf{x}) \chi_{\{|g(\varepsilon_t)| > B_n\}}\right|$. By the Borel-Cantelli Lemma for any $\epsilon > 0$ and for all m satisfying $m' < m < n$ we have $P(|g(\varepsilon_m)| \leq B_n) > 1 - \epsilon$, since $\{B_t\}_{t=1,2,\ldots}$ is an increasing sequence. By Chebyshev's Inequality, for $t = 1, \cdots, m'$ and $\zeta > 0$, $P((|g(\varepsilon_t)| > B_n) < \frac{E(|g(\varepsilon_t)|^\zeta)}{B_n^\zeta} < \frac{C}{B_n^\zeta}$ by a). Consequently, for all $\epsilon > 0$ and sufficiently large n, we have $P(|g(\varepsilon_t)| < B_n) > 1 - \epsilon$. Hence, for $n > max\{N, m\}$ we have that for all $t \leq n$, $P(|g(\varepsilon_t)| < B_n) > 1 - \epsilon$ and therefore $\chi_{\{|g(\varepsilon_t)| > B_n\}} = 0$ with probability 1, which gives $T_1 = o_{as}(1)$. For T_2, note that by A2 1) and A2 2)

$$E(s(\mathbf{x}) - s^\tau(\mathbf{x})) = \frac{1}{nh_n^d}\sum_{t=1}^{n}\int\int_{|g(\varepsilon_t)|>B_n} K\left(\frac{\mathbf{X}_t - \mathbf{x}}{h_n}\right) w(\mathbf{X}_t - \mathbf{x};\mathbf{x}) g(\varepsilon_t) f_X(\mathbf{X}_t) f(\varepsilon_t) d\mathbf{X}_t d\varepsilon_t$$

$$\leq \int_{\substack{|v_l|\leq C\\ l=1,\cdots,d}} K(\mathbf{v}) w(\mathbf{x} + h_n\mathbf{v};\mathbf{x}) f_{\mathbf{X}}(\mathbf{x} + h_n\mathbf{v}) d\mathbf{v} \int |g(\varepsilon)| f(\varepsilon) \chi_{\{|g(\varepsilon_t)|>B_n\}} d\varepsilon \tag{6}$$

where v_l is the l^{th} element of $\mathbf{v}$

$$\leq C\int |g(\varepsilon)| f(\varepsilon)\chi_{\{|g(\varepsilon_t)|>B_n\}} d\varepsilon.$$

The last inequality follows from A1 1), 2) and 4), condition b) and the the fact that w and f_X are continuous functions. By Hölder's inequality, for $\zeta > 1$,

$$\int \chi_{|g(\varepsilon_t)|>B_n} |g(\varepsilon_t)| f(\varepsilon_t) d\varepsilon_t \leq \left(\int |g(\varepsilon_t)|^\zeta f(\varepsilon_t) d\varepsilon_t\right)^{1/\zeta} \left(\int \chi_{\{|g(\varepsilon_t)|>B_n\}} f(\varepsilon_t) d\varepsilon_t\right)^{1-1/\zeta}$$

where the first integral after the inequality is uniformly bounded by a) and by Chebyshev's Inequality $\left(\int \chi_{\{|g(\varepsilon_t)|>B_n\}} f(\varepsilon_t) d\varepsilon_t\right)^{1-1/\zeta} \leq C(P(|g(\varepsilon_t)| > B_n))^{1-1/\zeta} \leq CB_n^{1-\zeta}$. Hence, $T_2 = O(B_n^{1-\zeta})$. As in Step (1) and given the orders for T_1 and T_2,

$$\sup_{\mathbf{x}\in\mathcal{G}}|s(\mathbf{x}) - E(s(\mathbf{x}))| \leq \sup_{\mathbf{x}\in\mathcal{G}}|s^\tau(\mathbf{x}) - E(s^\tau(\mathbf{x}))| + \sup_{\mathbf{x}\in\mathcal{G}}|s(\mathbf{x}) - s^\tau(\mathbf{x}) - (E(s(\mathbf{x})) - E(s^\tau(\mathbf{x})))|$$

$$\leq \max_{1\leq k\leq l_n} |s^\tau(\mathbf{x}) - E(s^\tau(\mathbf{x}))| + O\left(\left(\frac{log\, n}{nh_n^d}\right)^{1/2}\right) + O(B_n^{1-\zeta}) \text{ for } \zeta > 1.$$

Step (3): For $\varepsilon_n = \left(\frac{nh_n^d}{log\, n}\right)^{-1/2}\Delta$ with $0 < \Delta < \infty$, note that $P\left(\max_{1\leq k\leq l_n}\left|s^\tau(\mathbf{x}^k) - E(s^\tau(\mathbf{x}^k))\right| \geq \varepsilon_n\right) \leq \sum_{k=1}^{l_n} P\left(\left|s^\tau(\mathbf{x}^k) - E(s^\tau(\mathbf{x}^k))\right| \geq \varepsilon_n\right)$. Let $s^\tau(\mathbf{x}^k) - E(s^\tau(\mathbf{x}^k)) = \frac{1}{n}\sum_{t=1}^n Z_t$ with

$$Z_t = \frac{1}{h_n^d}K\left(\frac{\mathbf{X}_t - \mathbf{x}^k}{h_n}\right) w(\mathbf{X}_t - \mathbf{x}^k;\mathbf{x}^k) g(\varepsilon_t)\chi_{\{|g(\varepsilon_t)|\leq B_n\}} - E\left(\frac{1}{h_n^d}K\left(\frac{\mathbf{X}_t - \mathbf{x}^k}{h_n}\right) w(\mathbf{X}_t - \mathbf{x}^k;\mathbf{x}^k) g(\varepsilon_t)\chi_{\{|g(\varepsilon_t)|\leq B_n\}}\right).$$

By A1 1), b) and $|g(\varepsilon_t)|\chi_{\{|g(\varepsilon_t)|\leq B_n)\}} \leq B_n$ we have that

$$|Z_t| \leq Ch_n^{-d}B_n. \tag{7}$$

Let $||Z_t||_\infty = inf\{a : P(Z_t > a) = 0\}$, then $\sup_{1\leq t\leq n} ||Z_t||_\infty \leq C\frac{B_n}{h_n^d}$. Then, from Theorem 1.3 in Bosq (1996) we have that for each $q = 1, 2, ..., [n/2]$

$$P\left(\frac{1}{n}\left|\sum_{t=1}^n Z_t\right| > \varepsilon_n\right) \leq 4exp\left(\frac{-\varepsilon_n^2 q}{8v^2(q)}\right) + 22\left(1 + \frac{4CB_n}{\varepsilon_n h_n^d}\right)^{1/2} q\alpha\left(\left[\frac{n}{2q}\right]\right)$$

where $p = n/2q$, $v^2(q) = \frac{2}{p^2}\sigma^2(q) + \frac{CB_n\varepsilon_n}{2h_n^d}$ and $\sigma^2(q) = \max_{0\leq j\leq 2q-1} E\left(\left(([jp]+1-jp)Z_{[jp]+1} + Z_{[jp]+2} + \cdots + Z_{[(j+1)p]} + ((j+1)p - [(j+1)p])Z_{[(j+1)p+1]}\right)^2\right)$. We first show that $\frac{h_n^d}{p}\sigma^2(q) = O(1)$. To see this, note that

$$\sigma^2(q) \leq \max_{0\leq j\leq 2q-1}\left(\sum_{[jp]<i\leq[(j+1)p+1]} E(Z_i^2) + 2\sum_{[jp]+1\leq l\leq[(j+1)p]}\ \sum_{\substack{[jp]+1<i\leq[(j+1)p+1]\\ i>l}} |E(Z_l Z_i)|\right).$$

Given a), b), c) and A2 we have $\sum_{[jp]<i\leq[(j+1)p+1]} E(Z_i^2) \leq O(ph_n^{-d})$. Since $E(|Z_t|^\delta) = O(h_n^{d(1-\delta)})$ for $\delta > 2$, by Theorem 3 (1) in Doukhan (1994) $|E(Z_i Z_l)| \leq Ch_n^{2d\left(\frac{1}{\delta}-1\right)}\alpha(i-l)^{1-\frac{2}{\delta}}$. Now, for any l such that $[jp]+1 \leq l \leq [(j+1)p]$ we have that $\sum_{[jp]+1<i\leq[(j+1)p+1]} |E(Z_l Z_i)| \leq \sum_{i=1}^{p^*-1} |E(Z_l Z_{l+i})| + \sum_{i=1}^{p^*-1} |E(Z_l Z_{l-i})|$ where $p^* = [(j+1)p+1] - [jp]$. Letting $d_n = h_n^{-\frac{d}{a_1}\left(1-\frac{2}{\delta}\right)}$ be a sequence of integers, we have that $d_n h_n^d \to 0$ whenever $a_1 > 1 - \frac{2}{\delta}$. Hence, we can write

$$\sum_{i=1}^{p^*-1} |E(Z_l Z_{l+i})| = \sum_{i=1}^{d_n-1} |E(Z_l Z_{l+i})| + \sum_{i=d_n}^{p^*-1} |E(Z_l Z_{l+i})| = J_1 + J_2$$

and easily show that $J_1 = o(h_n^{-d})$ and $J_2 = O(h_n^{-d})$. Similarly we obtain $\sum_{i=1}^{p^*-1} |E(Z_l Z_{l-i})| = O(h_n^{-d})$. Combining the results on the variance and covariances we have that $\frac{h_n^d}{p}\sigma^2(q) \leq C$ for n sufficiently large. Hence, we have that $ph_n^d v^2(q) \leq C + CpB_n\varepsilon_n$ and choosing $p = (B_n\varepsilon_n)^{-1}$ (with $B_n\varepsilon_n \to 0$) we have that for n sufficiently large $ph_n^d v^2(q) \leq C$. Then, $4exp\left(\frac{-\varepsilon_n^2 q}{8v^2(q)}\right) \leq 4exp\left(\frac{-\varepsilon_n^2 nh_n^d}{16C}\right) \leq 4n^{-\frac{\Delta^2}{16C}}$. Now, given A2, for $B > 2$ and n sufficiently large

$$\begin{aligned} 22\left(1 + \frac{4CB_n}{\varepsilon_n h_n^d}\right)^{1/2} q\alpha\left(\left[\frac{n}{2q}\right]\right) &\leq C\left(\frac{B_n}{\varepsilon_n}\right)^{1/2} h_n^{-d/2}\frac{n}{2p}[p]^{-B} \\ &\leq Cnh_n^{-d/2}B_n^{B+1.5}\varepsilon_n^{B+0.5}. \end{aligned}$$

Thus, $P\left(\max_{1\leq k\leq l_n}\left|s^\tau(\mathbf{x}^k) - E(s^\tau(\mathbf{x}^k))\right| \geq \varepsilon_n\right) < \left(\frac{n}{h_n^{d+2}log\, n}\right)^{d/2}\left(4n^{-\frac{\Delta^2}{16C}} + Cnh_n^{-d/2}B_n^{B+1.5}\varepsilon_n^{B+0.5}\right)$ and if Δ is chosen such that $\frac{\Delta^2}{16C} > 1$ the first term in the sum to the right of the inequality is negligible and we have that

$$P\left(\max_{1\leq k\leq l_n}\left|s^\tau(\mathbf{x}^k) - E(s^\tau(\mathbf{x}^k))\right| \geq \varepsilon_n\right) < CB_n^{B+1.5}(log\, n)^{0.25+0.5(B-d)}n^{0.75+0.5(d-B)}h_n^{-1.75d-0.5d(d+B)}. \quad (8)$$

Choosing $B_n \propto n^{1/\zeta+\theta}$ for $\zeta > 2, \theta > 0$ we have $B_n^{1-\zeta} < n^{-0.5-\theta}$ and $B_n^{1-\zeta} = o(n^{-1/2})$. Furthermore, the restriction $B_n \varepsilon_n \to 0$ requires that $n^{1-\frac{2}{\zeta}-2\theta} h_n^d \to \infty$. Lastly, for our choice of B_n

$$P\left(\max_{1\le k\le l_n} \left|s^\tau(\mathbf{x}_k) - E(s^\tau(\mathbf{x}^k))\right| \ge \varepsilon_n\right) < Cn^{(B+1.5)(\frac{1}{\zeta}+\theta)-0.5(B-d)+0.75}(log\, n)^{0.25+0.5(B-d)} h_n^{-1.75d-0.5d(d+B)}.$$

Since the expression on the right-hand side of the inequality converges to zero by assumption, then

$$P\left(\max_{1\le k\le l_n} \left|s^\tau(\mathbf{x}^k) - E(s^\tau(\mathbf{x}^k))\right| \ge \varepsilon_n\right) = o_p(1).$$

which completes the proof for the case where $p_1 = p_2 = p_3 = 0$. We now turn to the cases where p_1, p_2, p_3 may differ from zero. Consider first the case where only one of the exponents, say $p_1 = 1$ and $p_2 = p_3 = 0$. Provided that the bounds (5), (6), (7) continue to hold in the case where $s(\mathbf{x}) = \frac{1}{nh_n^d}\sum_{t=1}^{n} K\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_n}\right)\left(\frac{X_{ti}-x_i}{h_n}\right) w(\mathbf{X}_t-\mathbf{x};\mathbf{x}) g(\varepsilon_t)$ the proof follows in an analogous manner. Verification of the bounds, however, follows directly from the fact that $\mathcal{K}$ has compact support and is uniformly bounded. As such, whenever $\left|\frac{X_{ti}-x_i}{h_n}\right| > C$, $\mathcal{K}\left(\frac{X_{ti}-x_i}{h_n}\right) = 0$. All other cases, that is when $p_1 = p_2 = p_3 = 1$ and $i = j = l$, $i \ne j \ne l$, $i \ne j = l$, $i = j \ne l$ or $i = l \ne j$ are treated similarly. □

Lemma 2. *Assume that the kernel K_1 used to define $\hat{m}$ satisfies assumption A1 and assumptions A2 and A3 are holding. Assume also that the bandwidth h_{1n} used to define $\hat{m}$ satisfies equations (3) and (4). Then, if $E(|\varepsilon_t|^\zeta) < \infty$, $E(h^{1/2}(\mathbf{X}_t)^\zeta) < \infty$ for some $\zeta > 2$ and condition c) in Lemma 1 is holding*

$$\sup_{\mathbf{x}\in\mathcal{G}} |\hat{m}(\mathbf{x}) - m(\mathbf{x})| = O_p\left(L_{1n}\right), \tag{9}$$

where $L_{1n} = \left(\frac{log\, n}{nh_{1n}^d}\right)^{1/2} + h_{1n}^s$.

Proof. Note that $\hat{m}(\mathbf{x}) - m(\mathbf{x}) = e^T S_n^{-1}(\mathbf{x})\mathbf{c}_n(x)$ where $e^T = \begin{pmatrix} 1 & 0 & \cdots & 0 \end{pmatrix}$ is a $1 \times (d+1)$ vector,

$$S_n(x) = \begin{pmatrix} s_0(\mathbf{x}) & s_1(\mathbf{x}) & s_2(\mathbf{x}) & \cdots & s_d(\mathbf{x}) \\ s_1(\mathbf{x}) & s_{(1,1)}(\mathbf{x}) & s_{(1,2)}(\mathbf{x}) & \cdots & s_{(1,d)}(\mathbf{x}) \\ s_2(\mathbf{x}) & s_{(2,1)}(\mathbf{x}) & s_{(2,2)}(\mathbf{x}) & \cdots & s_{(2,d)}(\mathbf{x}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ s_d(\mathbf{x}) & s_{(d,1)}(\mathbf{x}) & s_{(d,2)}(\mathbf{x}) & \cdots & s_{(d,d)}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} s_0(\mathbf{x}) & \mathbf{S}_1^T(\mathbf{x}) \\ \mathbf{S}_1(\mathbf{x}) & \mathbf{S}_2(\mathbf{x}) \end{pmatrix} \text{ and } \mathbf{c}_n(\mathbf{x}) = \begin{pmatrix} c_0(\mathbf{x}) \\ c_1(\mathbf{x}) \\ \vdots \\ c_d(\mathbf{x}) \end{pmatrix} \tag{10}$$

with $s_0(\mathbf{x}) = \frac{1}{nh_n^d}\sum_{t=1}^{n} K\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{1n}}\right)$, $s_j(\mathbf{x}) = \frac{1}{nh_n^d}\sum_{t=1}^{n} K\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{1n}}\right)(X_{tj}-x_j)$, $c_0(\mathbf{x}) = \frac{1}{nh_{1n}^d}\sum_{t=1}^{n} K\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{1n}}\right) Y_t^*$, $c_j(\mathbf{x}) = \frac{1}{nh_{1n}^d}\sum_{t=1}^{n} K\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{1n}}\right)(X_{tj}-x_j) Y_t^*$ and $s_{(i,j)}(\mathbf{x}) = \frac{1}{nh_n^d}\sum_{t=1}^{n} K\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{1n}}\right)(X_{ti}-x_i)(X_{tj}-x_j)$ for $i, j = 1, \cdots, d$

and $Y_t^* = Y_t - m(\mathbf{x}) - m^{(1)}(\mathbf{x})(\mathbf{X}_t - \mathbf{x})$. Let $G_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & h_{1n}^{-s} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & h_{1n}^{-s} \end{pmatrix}$ and

$$\Sigma(\mathbf{x}) = \begin{pmatrix} f_{\mathbf{X}}(\mathbf{x}) & 0 & \cdots & 0 \\ \frac{\mu_{\mathcal{K},s}}{(s-1)!} D_1^{(s-1)} f_{\mathbf{X}}(\mathbf{x}) & \frac{\mu_{\mathcal{K},s}}{(s-2)!} D_1^{(s-2)} f_{\mathbf{X}}(\mathbf{x}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\mu_{\mathcal{K},s}}{(s-1)!} D_d^{(s-1)} f_{\mathbf{X}}(\mathbf{x}) & 0 & \cdots & \frac{\mu_{\mathcal{K},s}}{(s-2)!} D_d^{(s-2)} f_{\mathbf{X}}(\mathbf{x}) \end{pmatrix}$$

where $D_j^{(s)} f_{\mathbf{X}}(\mathbf{x}) = D_{i_1,\cdots,i_s} f_{\mathbf{X}}(\mathbf{x})$ where $i_1 = \cdots = i_s = j$. By partitioned inversion we obtain

$$\Sigma^{-1}(\mathbf{x}) = \begin{pmatrix} f_{\mathbf{X}}^{-1}(\mathbf{x}) & 0 & \cdots & 0 \\ -\frac{1}{(s-1) f_{\mathbf{X}}(\mathbf{x})} \frac{D_1^{(s-1)} f_{\mathbf{X}}(\mathbf{x})}{D_1^{(s-2)} f_{\mathbf{X}}(\mathbf{x})} & \frac{(s-2)!}{\mu_{\mathcal{K},s}} \frac{1}{D_1^{(s-2)} f_{\mathbf{X}}(\mathbf{x})} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -\frac{1}{(s-1) f_{\mathbf{X}}(\mathbf{x})} \frac{D_d^{(s-1)} f_{\mathbf{X}}(\mathbf{x})}{D_d^{(s-2)} f_{\mathbf{X}}(\mathbf{x})} & 0 & \cdots & \frac{(s-2)!}{\mu_{\mathcal{K},s}} \frac{1}{D_d^{(s-2)} f_{\mathbf{X}}(\mathbf{x})} \end{pmatrix}.$$

Then, $\hat{m}(\mathbf{x}) - m(\mathbf{x}) = e^T \left((G_n S_n(\mathbf{x}))^{-1} - \Sigma^{-1}(\mathbf{x}) \right) G_n c_n(\mathbf{x}) + e^T \Sigma^{-1}(\mathbf{x}) G_n c_n(\mathbf{x}) = \mathcal{I}_{1n}(\mathbf{x}) + \mathcal{I}_{2n}(\mathbf{x})$ By the Cauchy-Schwartz and Triangle inequalities we have

$$|\mathcal{I}_{1n}(\mathbf{x})| \leq \left(e^T \left((G_n S_n(\mathbf{x}))^{-1} - \Sigma^{-1}(\mathbf{x}) \right)^2 e \right)^{1/2} \left(|c_0(\mathbf{x})| + \left| \frac{1}{h_{1n}^s} \sum_{j=1}^{d} c_j(\mathbf{x}) \right| \right), \tag{11}$$

where we note that $(G_n S_n(\mathbf{x}))^{-1} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix}$ with

$$\mathbf{A}_{11} = \left(s_0(\mathbf{x}) - \mathbf{S}_1^T(\mathbf{x}) \left(\frac{1}{h_{1n}^s} \mathbf{S}_2(\mathbf{x}) \right)^{-1} \frac{1}{h_{1n}^s} \mathbf{S}_1(\mathbf{x}) \right)^{-1} \text{ and}$$
$$\mathbf{A}_{12} = -\frac{1}{s_0(\mathbf{x})} \mathbf{S}_1^T(\mathbf{x}) \left(\frac{1}{h_{1n}^s} \mathbf{S}_2(\mathbf{x}) - \frac{1}{h_{1n}^s s_0(\mathbf{x})} \mathbf{S}_1(\mathbf{x}) \mathbf{S}_1(\mathbf{x})^T \right)^{-1}. \tag{12}$$

By Lemma 1 we have $\sup_{\mathbf{x} \in \mathcal{G}} |s_0(\mathbf{x}) - E(s_0(\mathbf{x}))| = O_p\left(\left(\frac{\log n}{n h_{1n}^d} \right)^{1/2} \right)$, $\frac{1}{h_{1n}} \sup_{\mathbf{x} \in \mathcal{G}} |s_j(\mathbf{x}) - E(s_j(\mathbf{x}))| = O_p\left(\left(\frac{\log n}{n h_{1n}^d} \right)^{1/2} \right)$ and $\frac{1}{h_{1n}^2} \sup_{\mathbf{x} \in \mathcal{G}} |s_{(i,j)}(\mathbf{x}) - E(s_{(i,j)}(\mathbf{x}))| = O_p\left(\left(\frac{\log n}{n h_{1n}^d} \right)^{1/2} \right)$. Now, given A2 3) and A1 $\sup_{\mathbf{x} \in \mathcal{G}} |E(s_0(\mathbf{x})) - f_{\mathbf{X}}(\mathbf{x})| = O(h_{1n}^s)$, $\sup_{\mathbf{x} \in \mathcal{G}} \left| \frac{E(s_j(\mathbf{x}))}{h_{1n}^s} - \frac{\mu_{\mathcal{K},s}}{(s-1)!} D_j^{(s-1)} f_{\mathbf{X}}(\mathbf{x}) \right| = o(1)$, $\sup_{\mathbf{x} \in \mathcal{G}} \left| \frac{E(s_{(i,i)}(\mathbf{x}))}{h_{1n}^s} - \frac{\mu_{\mathcal{K},s}}{(s-2)!} D_j^{(s-2)} f_{\mathbf{X}}(\mathbf{x}) \right| = o(1)$ and lastly $\sup_{\mathbf{x} \in \mathcal{G}} \left| \frac{E(s_{(i,j)}(\mathbf{x}))}{h_{1n}^s} \right| = o(1)$ for $i \neq j$. Consequently, given that $h_{1n} \propto n^{-\frac{1}{2s+d}}$ we have

$$\sup_{\mathbf{x} \in \mathcal{G}} |s_0(\mathbf{x}) - f_{\mathbf{X}}(\mathbf{x})| = O_p(L_{1n}) \tag{13}$$

$$\sup_{\mathbf{x}\in\mathcal{G}}\left|\frac{s_j(\mathbf{x})}{h_{1n}^s}-\frac{\mu_{\mathcal{K},s}}{(s-1)!}D_j^{(s-1)}f_{\mathbf{X}}(\mathbf{x})\right|=o_p(1) \text{ which implies } \sup_{\mathbf{x}\in\mathcal{G}}|s_j(\mathbf{x})|=O_p(h_{1n}^s) \tag{14}$$

$$\sup_{\mathbf{x}\in\mathcal{G}}\left|\frac{s_{(i,i)}(\mathbf{x})}{h_{1n}^s}-\frac{\mu_{\mathcal{K},s}}{(s-2)!}D_j^{(s-2)}f_{\mathbf{X}}(\mathbf{x})\right|=o_p(1) \text{ and } \sup_{\mathbf{x}\in\mathcal{G}}\left|\frac{s_{(i,j)}(\mathbf{x})}{h_{1n}^s}\right|=o_p(1) \text{ which implies}$$

$$\sup_{\mathbf{x}\in\mathcal{G}}|\mathbf{S}_2(\mathbf{x})|=O_p(h_{1n}^s), \tag{15}$$

where the absolute value and order in the last equation are taken element-wise. Now, in inequality (11) we have that

$$e^T\left((G_nS_n(\mathbf{x}))^{-1}-\Sigma^{-1}(\mathbf{x})\right)^2e=\left(\mathbf{A}_{11}(\mathbf{x})-\frac{1}{f_{\mathbf{X}}(\mathbf{x})}\right)^2+\sum_{j=1}^d\mathbf{A}_{12j}^2(\mathbf{x})$$

where $\mathbf{A}_{12j}(\mathbf{x})$ is the j^{th} element of $\mathbf{A}_{12}(\mathbf{x})$. Given (12), (13), (14) and (15) we have $\mathbf{A}_{11}(\mathbf{x})-\frac{1}{f_{\mathbf{X}}(\mathbf{x})}=O_p(L_{1n})$, $\mathbf{A}_{12j}(\mathbf{x})=O_p(h_{1n}^s)$ which gives $e^T\left((G_nS_n(\mathbf{x}))^{-1}-\Sigma^{-1}(\mathbf{x})\right)^2e=O_p(L_{1n}^2)$ and

$$\mathcal{I}_{1n}(\mathbf{x})\leq O_p(L_{1n})\left(|c_0(\mathbf{x})|+\left|\frac{1}{h_{1n}^s}\sum_{j=1}^d c_j(\mathbf{x})\right|\right).$$

Now,

$$\begin{aligned}
Y_t^*&=m(\mathbf{X}_t)-m(\mathbf{x})-m^{(1)}(\mathbf{x})(\mathbf{X}_t-\mathbf{x})+h^{1/2}(\mathbf{X}_t)\varepsilon_t\\
&=\frac{1}{2}\sum_{i_1=1}^d\sum_{i_2=1}^d D_{i_1i_2}m(\mathbf{x}+\lambda(\mathbf{X}_t-\mathbf{x}))(\mathbf{X}_{ti_1}-x_{i_1})(\mathbf{X}_{ti_2}-x_{i_2})+h^{1/2}(\mathbf{X}_t)\varepsilon_t\\
&=P_t(\mathbf{x})+h^{1/2}(\mathbf{X}_t)\varepsilon_t
\end{aligned}$$

and $c_0(\mathbf{x})=\frac{1}{nh_{1n}^d}\sum_{t=1}^n K_1\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{1n}}\right)P_t(\mathbf{x})+\frac{1}{nh_{1n}^d}\sum_{t=1}^n K_1\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{1n}}\right)h^{1/2}(X_t)\varepsilon_t=c_{0,1}(\mathbf{x})+c_{0,2}(\mathbf{x})$. By Lemma 1 we have $\frac{1}{h_{1n}^2}\sup_{\mathbf{x}\in\mathcal{G}}|c_{0,1}(\mathbf{x})-E(c_{0,1}(\mathbf{x}))|=O_p\left(\left(\frac{log\ n}{nh_{1n}^d}\right)^{1/2}\right)$. Now, we observe that

$$\begin{aligned}
E(c_{0,1}(\mathbf{x}))=\frac{1}{h_{1n}}E\Bigg(K_1\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{1n}}\right)\Bigg(&\frac{1}{2!}\sum_{i_1=1}^d\sum_{i_2=1}^d D_{i_1i_2}m(\mathbf{x})(\mathbf{X}_{ti_1}-x_{i_1})(\mathbf{X}_{ti_2}-x_{i_2})\\
&+\cdots+\frac{1}{s!}\sum_{i_1=1}^d\cdots\sum_{i_s=1}^d D_{i_1\cdots i_s}m(\mathbf{x}+\lambda(\mathbf{X}_t-\mathbf{x}))(\mathbf{X}_{ti_1}-x_{i_1})\cdots(\mathbf{X}_{ti_s}-x_{i_s})\Bigg)\Bigg)
\end{aligned}$$

and by A2 3) and A1 $\sup_{\mathbf{x}\in\mathcal{G}}|E(c_{0,1}(\mathbf{x}))|=O(h_{1n}^s)$. Hence, $\sup_{\mathbf{x}\in\mathcal{G}}|c_{0,1}(\mathbf{x})|=O_p(h_{1n}^s)$ given that $h_{1n}\propto n^{-\frac{1}{2s+d}}$. Similarly, by Lemma 1 and given that $E(c_{0,2}(\mathbf{x}))=0$, $\sup_{\mathbf{x}\in\mathcal{G}}|c_{0,2}(\mathbf{x})|=O_p\left(\left(\frac{log\ n}{nh_{1n}^d}\right)^{1/2}\right)$. Thus, $\sup_{\mathbf{x}\in\mathcal{G}}|c_0(\mathbf{x})|=O_p\left(L_{1n}\right)$. Similar arguments give

$$\sup_{\mathbf{x}\in\mathcal{G}}|c_j(\mathbf{x})|=O_p\left(h_{1n}^s+h_{1n}\left(\frac{log\ n}{nh_{1n}^d}\right)^{1/2}\right).$$

Consequently, $\sup_{\mathbf{x}\in\mathcal{G}}\mathcal{I}_{1n}(\mathbf{x}) = O_p(L_{1n})$. Lastly, $\mathcal{I}_{2n}(\mathbf{x}) = \frac{1}{f_{\mathbf{X}}(\mathbf{x})}c_0(\mathbf{x})$ and from the fact that $f_{\mathbf{X}}(\mathbf{x})$ is uniformly bounded away from 0 and the order of $c_0(\mathbf{x})$ we have $\sup_{\mathbf{x}\in\mathcal{G}}\mathcal{I}_{2n}(\mathbf{x}) = O_p(L_{1n})$. Therefore, $\sup_{\mathbf{x}\in\mathcal{G}}|\hat{m}(\mathbf{x}) - m(\mathbf{x})| = O_p(L_{1n})$. □

Lemma 3. *Assume that the kernel K_2 used to define $\hat{h}$ satisfies assumption A1 and assumptions A2 and A3 are holding. Assume also that the bandwidth h_{2n} used to define $\hat{h}$ satisfies equations (3) and (4). Then, under the assumptions in Lemma 2, if $E(|\varepsilon_t^2 - 1|^\zeta) < \infty$ and $E(h(\mathbf{X}_t)^\zeta) < \infty$ for some $\zeta > 2$,*

$$\mathop{sup}_{\mathbf{x}\in\mathcal{G}}|\hat{h}(\mathbf{x}) - h(\mathbf{x})| = O_p\left(L_{1n} + L_{2n}\right), \tag{16}$$

where $L_{1n} = \left(\frac{log\, n}{nh_{1n}^d}\right)^{1/2} + h_{1n}^s$ and $L_{2n} = \left(\frac{log\, n}{nh_{2n}^d}\right)^{1/2} + h_{2n}^s$.

Proof. As in Lemma 2, we write $\hat{h}(\mathbf{x}) - h(\mathbf{x}) = e^T\left((G_nS_n(\mathbf{x}))^{-1} - \Sigma^{-1}(\mathbf{x})\right)G_nq_n(\mathbf{x}) + e^T\Sigma^{-1}(\mathbf{x})G_nq_n(\mathbf{x}) = \mathcal{I}_{1n}^h(\mathbf{x}) + \mathcal{I}_{2n}^h(\mathbf{x})$ where $q_n(\mathbf{x})^T = \begin{pmatrix} q_0(\mathbf{x}) & q_1(\mathbf{x}) & \cdots & q_d(\mathbf{x}) \end{pmatrix}$, $q_0(\mathbf{x}) = \frac{1}{nh_{2n}^d}\sum_{t=1}^n K_2\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{2n}}\right)U_t^{*2}$, $q_j(\mathbf{x}) = \frac{1}{nh_{2n}^d}\sum_{t=1}^n K_2\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{2n}}\right)(X_{tj} - x_j)\,U_t^{*2}$, $U_t^{*2} = \hat{U}_t^2 - h(\mathbf{x}) - h^{(1)}(\mathbf{x})(\mathbf{X}_t - \mathbf{x})$ and $\hat{U}_t = Y_t - \hat{m}(\mathbf{X}_t)$. Hence, we write

$$U_t^{*2} = h(\mathbf{X}_t)(\varepsilon_t^2 - 1) + P_t^h(\mathbf{x}) - 2(\hat{m}(\mathbf{X}_t) - m(\mathbf{X}_t))h^{1/2}(\mathbf{X}_t)\varepsilon_t + (\hat{m}(\mathbf{X}_t) - m(\mathbf{X}_t))^2$$

where $P_t^h(\mathbf{x}) = \frac{1}{2}\sum_{i_1=1}^d\sum_{i_2=1}^d D_{i_1i_2}m(\mathbf{x} + \lambda(\mathbf{X}_t - \mathbf{x}))(\mathbf{X}_{ti_1} - x_{i_1})(\mathbf{X}_{ti_2} - x_{i_2})$. From Lemma 2 we have $|\mathcal{I}_{1n}^h(\mathbf{x})| \le O_p(L_{2n})\left(|q_0(\mathbf{x})| + \left|\frac{1}{h_{2n}^s}\sum_{j=1}^d q_j(\mathbf{x})\right|\right)$. Now,

$$\begin{aligned}
q_0(\mathbf{x}) &= \frac{1}{nh_{2n}^d}\sum_{t=1}^n K_2\left(\frac{\mathbf{X}_t - \mathbf{x}}{h_{2n}}\right)h(\mathbf{X}_t)(\varepsilon_t^2 - 1) + \frac{1}{nh_{2n}^d}\sum_{t=1}^n K_2\left(\frac{\mathbf{X}_t - \mathbf{x}}{h_{2n}}\right)P_t^h(\mathbf{x}) \\
&+ \frac{1}{nh_{2n}^d}\sum_{t=1}^n K_2\left(\frac{\mathbf{X}_t - \mathbf{x}}{h_{2n}}\right)(\hat{m}(\mathbf{X}_t) - m(\mathbf{X}_t))^2 - \frac{2}{nh_{2n}^d}\sum_{t=1}^n K_2\left(\frac{\mathbf{X}_t - \mathbf{x}}{h_{1n}}\right)(\hat{m}(\mathbf{X}_t) - m(\mathbf{X}_t))h^{1/2}(\mathbf{X}_t)\varepsilon_t \\
&= B_{1n}(\mathbf{x}) + B_{2n}(\mathbf{x}) + B_{3n}(\mathbf{x}) - B_{4n}(\mathbf{x}).
\end{aligned}$$

Given the conditions on this lemma and by the same arguments used to obtain the order of $c_{0,2}(\mathbf{x})$ in Lemma 2 we have that $\sup_{\mathbf{x}\in\mathcal{G}}|B_{1n}(\mathbf{x})| = O_p\left(\left(\frac{log\, n}{nh_{2n}^d}\right)^{1/2}\right)$. Also, by the same arguments used to obtain the order of $c_{0,1}(\mathbf{x})$ in Lemma 2 and provided $h_{2n} \propto n^{-\frac{1}{2s+d}}$ we have that $\sup_{\mathbf{x}\in\mathcal{G}}|B_{2n}(\mathbf{x})| = O_p\left(h_{2n}^s\right)$.

$B_{4n}(\mathbf{x}) \le 2\sup_{\mathbf{x}\in\mathcal{G}}|\hat{m}(\mathbf{x}) - m(\mathbf{x})|\frac{1}{nh_{2n}^d}\sum_{t=1}^n\left|K_2\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{2n}}\right)\right|h^{1/2}(\mathbf{X}_t)|\varepsilon_t| = 2\sup_{\mathbf{x}\in\mathcal{G}}|\hat{m}(\mathbf{x}) - m(\mathbf{x})|\mathcal{M}_n(\mathbf{x})$. By Lemma

1 we have $\sup\limits_{\mathbf{x}\in\mathcal{G}}|\mathcal{M}_n(\mathbf{x})-E(\mathcal{M}_n(\mathbf{x}))|=O_p\left(\left(\frac{\log n}{nh_{2n}^d}\right)^{1/2}\right)$. Furthermore,

$$E(\mathcal{M}_n(\mathbf{x}))=E(|\varepsilon_t|)\int\prod_{j=1}^d|\mathcal{K}(u_j)|h^{1/2}(\mathbf{x}+h_{2n}\mathbf{u})f_{\mathbf{X}}(\mathbf{x}+h_{2n}\mathbf{u})d\mathbf{u}\to E(|\varepsilon_t|)h^{1/2}(\mathbf{x})f_{\mathbf{X}}(\mathbf{x})\int\prod_{j=1}^d|\mathcal{K}(u_j)|d\mathbf{u}<C$$

given A1, A2 1), A3 2) and the fact that $E(\varepsilon_t^2)=1$. Hence, $\sup\limits_{\mathbf{x}\in\mathcal{G}}|\mathcal{M}_n(\mathbf{x})|=O_p(1)$ and from Lemma 2 we conclude that $\sup\limits_{\mathbf{x}\in\mathcal{G}}|B_{4n}(\mathbf{x})|=O_p(L_{1n})$. Through similar arguments we show that $\sup\limits_{\mathbf{x}\in\mathcal{G}}|B_{3n}(\mathbf{x})|=O_p(L_{1n}^2)$ since $\sup\limits_{\mathbf{x}\in\mathcal{G}}\frac{1}{h_{2n}^d}\sum\limits_{t=1}^n\left|K_2\left(\frac{\mathbf{X}_t-\mathbf{x}}{h_{2n}}\right)\right|=O_p(1)$. Hence, $\sup\limits_{\mathbf{x}\in\mathcal{G}}|q_0(\mathbf{x})|=O_p(L_{2n}+L_{1n})$. Lastly, for the terms $q_j(\mathbf{x})$, we have in an analogous manner that

$$\sup_{\mathbf{x}\in\mathcal{G}}|q_j(\mathbf{x})|=O_p\left(h_{2n}^s+h_{2n}\left(\frac{\log n}{nh_{2n}^d}\right)^{1/2}\right)+h_{2n}O_p(L_{1n})+h_{2n}O_p(L_{1n}^2).$$

Consequently, $\sup\limits_{\mathbf{x}\in\mathcal{G}}|\mathcal{I}_{1n}^h(\mathbf{x})|=O_p\left(L_{2n}\right)$. Lastly, $\mathcal{I}_{2n}^h(\mathbf{x})=\frac{1}{f_{\mathbf{X}}(\mathbf{x})}q_0(\mathbf{x})$ and from the fact that $f_{\mathbf{X}}(\mathbf{x})$ is uniformly bounded away from 0 and the order of $q_0(\mathbf{x})$ we have $\sup\limits_{\mathbf{x}\in\mathcal{G}}\mathcal{I}_{2n}^h(\mathbf{x})=O_p(L_{1n}+L_{2n})$. Therefore, $\sup\limits_{\mathbf{x}\in\mathcal{G}}\left|\hat{h}(\mathbf{x})-h(\mathbf{x})\right|=O_p(L_{1n}+L_{2n})$. □

Corollary 1. *Under the assumptions of Lemma 3,*

$$\mathop{sup}_{\mathbf{x}\in\mathcal{G}}|\hat{h}^{1/2}(\mathbf{x})-h^{1/2}(\mathbf{x})|=O_p\left(L_{1n}+L_{2n}\right)\quad \text{and}\quad \mathop{sup}_{\mathbf{x}\in\mathcal{G}}|\chi_{\{\hat{h}(\mathbf{x})>0\}}-1|=O_p\left(L_{1n}+L_{2n}\right),$$

where $L_{1n}=\left(\frac{\log n}{nh_{1n}^d}\right)^{1/2}+h_{1n}^2$ *and* $L_{2n}=\left(\frac{\log n}{nh_{2n}^d}\right)^{1/2}+h_{2n}^2$.

Lemma 4. *Under assumptions A1-A6 and conditions FR1 and FR2, if* $\alpha\geq 1$ *we have*

$$N^{1/2}\left(\frac{\tilde{q}\left(a_n\right)-q_n\left(a_n\right)}{q\left(a_n\right)}\right)=O_p(1),\ \text{where } a_n=1-\tfrac{N}{n}.$$

Proof. We write

$$\sqrt{N}\left(\frac{\tilde{q}\left(a_n\right)-q_n\left(a_n\right)}{q\left(a_n\right)}\right)=\sqrt{N}\left(\frac{\tilde{q}\left(a_n\right)-q\left(a_n\right)}{q\left(a_n\right)}\right)-\sqrt{N}\left(\frac{q_n\left(a_n\right)-q\left(a_n\right)}{q\left(a_n\right)}\right)=T_{1n}-T_{2n}.$$

We first show that T_{2n} converges in distribution, which implies $T_{2n}=O_p(1)$. Note that

$$P(T_{2n}\leq z)=P\left(\frac{nk_0}{\sqrt{N}}(F(y_n)-a_n)\leq-\frac{nk_0}{\sqrt{N}}(F_n(y_n)-F(y_n))\right)$$

with $y_n=q(a_n)\left(1+\frac{z}{\sqrt{N}}\right)$. By the mean value theorem, $F(y_n)=a_n+f(q^*(a_n))\frac{q(a_n)}{\sqrt{N}}z$ where $q^*(a_n)=q(a_n)\left(1+\lambda\frac{z}{\sqrt{N}}\right)$ for some $\lambda\in(0,1)$. Thus,

$$\frac{nk_0}{\sqrt{N}}(F(y_n)-a_n)=\frac{nk_0}{N}f(q^*(a_n))q(a_n)z=k_0\frac{n(1-F(q^*(a_n)))}{N}\frac{q(a_n)f(q^*(a_n))}{1-F(q^*(a_n))}z.$$

Since $q^*(a_n) = q(a_n)(1+o(1))$ we have that $\underset{n\to\infty}{lim} \frac{n(1-F(q^*(a_n)))}{N} = 1$. In addition, given FR1 and by Proposition 1.15 in Resnick (1987) we have $\underset{n\to\infty}{lim} \frac{q(a_n)f(q^*(a_n))}{1-F(q^*(a_n))} = -\frac{1}{k_0}$, hence $\underset{n\to\infty}{lim} \frac{nk_0}{\sqrt{N}}(F(y_n)-a_n) = -z$. We now show that $\frac{n}{\sqrt{N}}(F_n(y_n)-F(y_n)) \stackrel{d}{\to} N(0,1)$. First, we observe that $\frac{n}{\sqrt{N}} - \frac{\sqrt{n}}{\sqrt{1-F(y_n)}} = o(1)$, hence we show that

$$\frac{\sqrt{n}}{\sqrt{1-F(y_n)}}(F_n(y_n)-F(y_n)) = \sum_{t=1}^{n} Z_{tn} \stackrel{d}{\to} N(0,1) \tag{17}$$

where $Z_{tn} = \frac{1}{\sqrt{n(1-F(y_n))}}\left(\chi_{\{\varepsilon_t \le y_n\}} - E\left(\chi_{\{\varepsilon_t \le y_n\}}\right)\right)$. It is readily verified that $E(Z_{tn}) = 0$ and $V(Z_{tn}) = n^{-1}F(y_n)$. Hence, given that $\sum_{t=1}^{n} E(|Z_{tn}|^3) \le 2(n(1-F(y_n)))^{-1/2} = o(1)$ we have by Liapounov's CLT that $\frac{n}{\sqrt{N}}(F_n(y_n)-F(y_n)) \stackrel{d}{\to} N(0,1)$. Hence, $T_{2n} \stackrel{d}{\to} N(0,k_0^2)$. We now show that $T_{1n} = O_p(1)$ by establishing that T_{1n} converges in distribution. As above,

$$P(T_{1n} \le z) = P\left(\frac{nk_0}{\sqrt{N}}(F(y_n)-a_n) \le -\frac{nk_0}{\sqrt{N}}(\tilde{F}(y_n)-F(y_n))\right) \tag{18}$$

and we establish that $\frac{n}{\sqrt{N}}(\tilde{F}(y_n)-F(y_n)) \stackrel{d}{\to} N(0,1)$. We start by noting that for some $\lambda_t \in (0,1)$

$$\begin{aligned}\tilde{F}(y_n) &= \int_{-\infty}^{y_n} \frac{1}{nh_{3n}}\sum_{t=1}^{n} K_3\left(\frac{y-\varepsilon_t}{h_{3n}}\right)dy - \int_{-\infty}^{y_n} \frac{1}{nh_{3n}^2}\sum_{t=1}^{n} K_3^{(1)}\left(\frac{y-\varepsilon_t}{h_{3n}}\right)dy\,(\hat{\varepsilon}_t-\varepsilon_t) \\ &+ \frac{1}{2}\int_{-\infty}^{y_n} \frac{1}{nh_{3n}^3}\sum_{t=1}^{n} K_3^{(2)}\left(\frac{y-\varepsilon_t^*}{h_{3n}}\right)dy\,(\hat{\varepsilon}_t-\varepsilon_t)^2 = Q_{1n} - Q_{2n} + Q_{3n}.\end{aligned}$$

where $\varepsilon_t^* = \lambda_t\varepsilon_t + (1-\lambda_t)\hat{\varepsilon}_t$. Therefore, $\frac{n}{\sqrt{N}}(\tilde{F}(y_n)-F(y_n)) = \frac{n}{\sqrt{N}}((Q_{1n}-F(y_n)) - Q_{2n} + Q_{3n})$. From equation (17) in the proof of Theorem 1 we have

$$\begin{aligned}Q_{3n} &\le \frac{1}{2nh_{3n}^2}\sum_{t=1}^{n}\left|K_3^{(1)}\left(\frac{y_n-\varepsilon_t^*}{h_{3n}}\right)\right|(O_p(L_{1n}) + O_p(L_{1n}+L_{2n})|\varepsilon_t|)^2 \\ &\le \frac{1}{2nh_{3n}^2}\sum_{t=1}^{n}\left|K_3^{(1)}\left(\frac{y_n-\varepsilon_t^*}{h_{3n}}\right)\right|\left(O_p(L_{1n}^2) + O_p(L_{1n}^2+L_{2n}^2)\varepsilon_t^2\right) \\ &= O_p\left(\frac{L_{1n}^2}{h_{3n}}\right)\frac{1}{nh_{3n}}\sum_{t=1}^{n}\left|K_3^{(1)}\left(\frac{y_n-\varepsilon_t^*}{h_{3n}}\right)\right| + O_p\left(\frac{L_{1n}^2+L_{2n}^2}{h_{3n}}\right)\frac{1}{nh_{3n}}\sum_{t=1}^{n}\left|K_3^{(1)}\left(\frac{y_n-\varepsilon_t^*}{h_{3n}}\right)\right|\varepsilon_t^2 \\ &= O_p\left(\frac{L_{1n}^2}{h_{3n}}\right)Q_{31n} + O_p\left(\frac{L_{1n}^2+L_{2n}^2}{h_{3n}}\right)Q_{32n}\end{aligned} \tag{19}$$

Using Taylor's Theorem we can write for some $\lambda \in (0,1)$ and $\varepsilon_t^{**} = \lambda_t\varepsilon_t + (1-\lambda_t)\varepsilon_t^*$ that

$$\begin{aligned}Q_{32n} &\le \frac{1}{nh_{3n}}\sum_{t=1}^{n}\left|K_3^{(1)}\left(\frac{y_n-\varepsilon_t}{h_{3n}}\right)\right|\varepsilon_t^2 + \frac{1}{nh_{3n}^2}\sum_{t=1}^{n}\left|K_3^{(2)}\left(\frac{y_n-\varepsilon_t^{**}}{h_{3n}}\right)\right||\varepsilon_t^*-\varepsilon_t|\varepsilon_t^2 \\ &= Q_{321n} + Q_{322n}.\end{aligned}$$

Note that

$$E(Q_{321n}) = y_n(1-F(y_n)) \int |K_3^{(1)}(\psi)| \frac{(y_n - h_{3n}\psi) f(y_n - h_{3n}\psi)}{1-F(y_n - h_{3n}\psi)} \frac{1-F(y_n - h_{3n}\psi)}{1-F(y_n)} \frac{(y_n - h_{3n}\psi)}{y_n} d\,\psi$$

and since by part a) of Proposition 1.15 in Resnick (1987)

$$\frac{(y_n - h_{3n}\psi) f(y_n - h_{3n}\psi)}{1-F(y_n - h_{3n}\psi)} \to -1/k_0,\ \frac{1-F(y_n - h_{3n}\psi)}{1-F(y_n)} \to 1$$

and $\frac{(y_n - h_{3n}\psi)}{y_n} \to 1$ we have that $E(Q_{311n}) = y_n(1-F(y_n)) \int |K_3^{(1)}(\psi)| d\psi (1+o(1))$. By part c) of the same Proposition 1.15, $1-F(x) = C\,exp\,(-g(x))$ where $g(x) = \int_{z_0}^{x} t^{-1}\alpha(t)dt$ and $\alpha(t) \to \alpha$ as $t \to \infty$ for some z_0 and all $x > z_0$. Hence, $g(x) \approx \alpha log \frac{x}{z_0}$ and $1-F(x) \approx Cexp\left(-\alpha log \frac{x}{z_0}\right)$. Consequently, $\lim_{x\to\infty} x(1-F(x)) = C \lim_{x\to\infty} \frac{x}{exp(g(x))} = C \lim_{x\to\infty} \frac{x}{(x/z_0)^{\alpha}\alpha z_0^2} = C$ for $\alpha \geq 1$. Thus, we conclude that $y_n(1-F(y_n)) = O(1)$ and $Q_{321n} = O_p(1)$. For Q_{322n}, note that $|\varepsilon_t^* - \varepsilon_t| \leq \lambda_t \left(O_p(L_{1n}) + O_p(L_{1n} + L_{2n})|\varepsilon_t|\right)$ and since $\lambda_t \in (0,1)$ we have

$$Q_{322n} \leq \frac{O_p(L_{1n})}{h_{3n}^2} \frac{1}{n} \sum_{t=1}^{n} \left| K_3^{(2)}\left(\frac{y_n - \varepsilon_t^{**}}{h_{3n}}\right)\right| \varepsilon_t^2 + \frac{O_p(L_{1n}+L_{2n})}{h_{3n}^2} \frac{1}{n} \sum_{t=1}^{n} \left| K_3^{(2)}\left(\frac{y_n - \varepsilon_t^{**}}{h_{3n}}\right)\right| |\varepsilon_t|^3.$$

Given that $|K_3^{(2)}(x)| < C$, $E(|\varepsilon_t|^3) < \infty$ and the fact that $O_p((L_{1n}+L_{2n})h_{3n}^{-2}) = o_p(1)$ given the orders of h_{1n}, h_{2n} and h_{3n} we conclude that $Q_{32n} = O_p(1)$. Following similar arguments we have that $Q_{31n} = O_p(1)$ and consequently $Q_{3n} = O_p\left(\frac{L_{1n}^2}{h_{3n}}\right) + O_p\left(\frac{(L_{1n}+L_{2n})^2}{h_{3n}}\right)$. Thus,

$$\frac{n^{1/2}}{(1-F(y_n))^{1/2}} Q_{3n} = \frac{n}{\sqrt{N} h_{3n}} C \left(\frac{log\,n}{nh_{in}^d} + h_{in}^{2s}\right) = o_p(1) \tag{20}$$

for $i = 1, 2$ given the orders of h_{1n}, h_{2n}, h_{3n} and N. We now turn to the study of Q_{2n}. Using equation (16) in the proof of Theorem 1 we write

$$\begin{aligned} Q_{2n} = \frac{1}{nh_{3n}^2} \sum_{t=1}^{n} \int_{-\infty}^{y_n} K_3^{(1)}\left(\frac{y-\varepsilon_t}{h_{3n}}\right) dy \Big(&(m(\mathbf{X}_t) - \hat{m}(\mathbf{X}_t))(\hat{h}^{-1/2}(\mathbf{X}_t) - h^{-1/2}(\mathbf{X}_t))\chi_{\{\hat{h}(\mathbf{X}_t)>0\}} \\ &+ h^{-1/2}(\mathbf{X}_t)(m(\mathbf{X}_t) - \hat{m}(\mathbf{X}_t))\left(\chi_{\{\hat{h}(\mathbf{X}_t)>0\}} - 1\right) + h^{-1/2}(\mathbf{X}_t)(m(\mathbf{X}_t) - \hat{m}(\mathbf{X}_t)) \end{aligned}$$

$$\begin{aligned} &+ \left(\frac{h^{1/2}(\mathbf{X}_t)}{\hat{h}^{1/2}(\mathbf{X}_t)} - 1\right)\chi_{\{\hat{h}(\mathbf{X}_t)>0\}}\varepsilon_t + \left(\chi_{\{\hat{h}(\mathbf{X}_t)>0\}} - 1\right)\varepsilon_t \Big) \\ &= \sum_{j=1}^{5} Q_{2jn}. \end{aligned} \tag{21}$$

We investigate the order of each Q_{2jn} separately. First, we note that from Lemmas 3, 4 and Corollary 1

$$Q_{21n} \leq \frac{1}{h_{3n}} O_p(L_{1n}) O_p(L_{1n}+L_{2n}) \frac{1}{nh_{3n}} \sum_{t=1}^{n} \left| \int_{-\infty}^{y_n} K_3^{(1)}\left(\frac{y-\varepsilon_t}{h_{3n}}\right) dy \right| \tag{22}$$

with $\frac{1}{h_{3n}} O_p(L_{1n}) O_p(L_{1n}+L_{2n}) \leq \frac{1}{h_{3n}} O_p(L_{1n}^2+L_{2n}^2) = o_p\left(\frac{(1-F(y_n))^{1/2}}{n^{1/2}}\right)$. Furthermore,

$$E\left(\frac{1}{nh_{3n}} \sum_{t=1}^{n} \left| \int_{-\infty}^{y_n} K_3^{(1)}\left(\frac{y-\varepsilon_t}{h_{3n}}\right) dy \right|\right) = \frac{1}{h_{3n}} E\left(\left| K_3^{(1)}\left(\frac{y-\varepsilon_t}{h_{3n}}\right)\right|\right) = \int |K_3(\psi)| f(y_n - h_{3n}\psi) d\psi$$
$$= o(1)$$

since $f(y_n + h_{3n}\psi) \to 0$ as $n \to \infty$ and $|K_3(\psi)| < C$. Hence, $\frac{1}{nh_{3n}} \sum_{t=1}^{n} \left| \int_{-\infty}^{y_n} K_3^{(1)}\left(\frac{y-\varepsilon_t}{h_{3n}}\right) dy \right| = o_p(1)$ and

$$\frac{n^{1/2}}{(1-F(y_n))^{1/2}} Q_{21n} = o_p(1). \tag{23}$$

Since,

$$Q_{22n} \leq \frac{1}{nh_{3n}^2} \sum_{t=1}^{n} \left| \int_{-\infty}^{y_n} K_3^{(1)}\left(\frac{y-\varepsilon_t}{h_{3n}}\right) dy \right| \sup_{\mathbf{x}\in\mathcal{G}} h^{-1/2}(\mathbf{x}) \sup_{\mathbf{x}\in\mathcal{G}} |m(\mathbf{x}) - \hat{m}(\mathbf{x})| \sup_{\mathbf{x}\in\mathcal{G}} \left| \chi_{\{\hat{h}(\mathbf{x})>0\}} - 1 \right|,$$

we immediately conclude by Lemmas 3, 4, Corollary 1, the fact that $h(\mathbf{x})$ is bounded away from zero and $\frac{1}{nh_{3n}} \sum_{t=1}^{n} \left| \int_{-\infty}^{y_n} K_3^{(1)}\left(\frac{y-\varepsilon_t}{h_{3n}}\right) dy \right| = o_p(1)$ that

$$\frac{n^{1/2}}{(1-F(y_n))^{1/2}} Q_{22n} = o_p(1). \tag{24}$$

From Theorem 2 in Martins-Filho and Yao (2009) we write

$$\hat{m}(\mathbf{X}_t) - m(\mathbf{X}_t) = \frac{1}{nh_{1n}^d f_{\mathbf{X}}(\mathbf{X}_t)} \sum_{i=1}^{n} K_1\left(\frac{\mathbf{X}_i - \mathbf{X}_t}{h_{1n}}\right) Y_i^*(1+o_p(1)) \tag{25}$$

uniformly on $\mathcal{G}$, where Y_i^* is as defined in Lemma 2. Therefore,

$$\begin{aligned} Q_{23n} &= -\frac{1}{nh_{3n}} \sum_{t=1}^{n} \frac{1}{h_{3n}} \int_{-\infty}^{y_n} K_3^{(1)}\left(\frac{y-\varepsilon_t}{h_{3n}}\right) dy\, h^{-1/2}(\mathbf{X}_t) \frac{1}{nh_{1n}^d f_{\mathbf{X}}(\mathbf{X}_t)} \sum_{i=1}^{n} K_1\left(\frac{\mathbf{X}_i - \mathbf{X}_t}{h_{1n}}\right) Y_i^*(1+o_p(1)) \\ &= -(1+o_p(1)) \frac{1}{n^2} \sum_{t=1}^{n} \sum_{i=1}^{n} \frac{1}{h_{1n}^d h_{3n} f_{\mathbf{X}}(\mathbf{X}_t)} K_3\left(\frac{y_n-\varepsilon_t}{h_{3n}}\right) K_1\left(\frac{\mathbf{X}_i - \mathbf{X}_t}{h_{1n}}\right) \left(\frac{h^{1/2}(\mathbf{X}_i)}{h^{1/2}(\mathbf{X}_t)} \varepsilon_i \right. \\ &\quad \left. + \frac{1}{2} h^{-1/2}(\mathbf{X}_t)(\mathbf{X}_i - \mathbf{X}_t)^T m^{(2)}(\mathbf{Z}_{ti})(\mathbf{X}_i - \mathbf{X}_t)\right) \end{aligned}$$

where $\mathbf{Z}_{ti} = \lambda \mathbf{X}_i + (1-\lambda)\mathbf{X}_t$. Now, let

$$\begin{aligned} \psi_n(\mathbf{w}_t, \mathbf{w}_i) &= \frac{1}{h_{1n}^d h_{3n} f_{\mathbf{X}}(\mathbf{X}_t)} K_3\left(\frac{y_n-\varepsilon_t}{h_{3n}}\right) K_1\left(\frac{\mathbf{X}_i - \mathbf{X}_t}{h_{1n}}\right) \\ &\quad \times \left(\frac{h^{1/2}(\mathbf{X}_i)}{h^{1/2}(\mathbf{X}_t)} \varepsilon_i + \frac{1}{2} h^{-1/2}(\mathbf{X}_t)(\mathbf{X}_i - \mathbf{X}_t)^T m^{(2)}(\mathbf{Z}_{ti})(\mathbf{X}_i - \mathbf{X}_t)\right) \end{aligned}$$

for $\mathbf{w}_t = (\mathbf{X}_t \ \varepsilon_t)$ and define $\phi_n(\mathbf{w}_t, \mathbf{w}_i) = \psi_n(\mathbf{w}_t, \mathbf{w}_i) + \psi_n(\mathbf{w}_i, \mathbf{w}_t)$. Then, we write $Q_{23n} = -(1 + o_p(1))\frac{1}{2}\frac{1}{n^2}\sum_{t=1}^{n}\sum_{i=1}^{n}\phi_n(\mathbf{w}_t, \mathbf{w}_i)$ and

$$\begin{aligned}\frac{1}{2n^2}\sum_{t=1}^{n}\sum_{i=1}^{n}\phi_n(\mathbf{w}_t, \mathbf{w}_i) &= \frac{1}{2n^2}\sum_{t=1}^{n}\phi_n(\mathbf{w}_t, \mathbf{w}_t) + \frac{1}{2n^2}\sum_{t=1}^{n}\sum_{\substack{i=1\\ i\neq t}}^{n}\phi_n(\mathbf{w}_t, \mathbf{w}_i)\\ &= \frac{1}{2n^2}\sum_{t=1}^{n}\phi_n(\mathbf{w}_t, \mathbf{w}_t) + \frac{1}{n^2}\sum_{t=1}^{n}\sum_{\substack{i=1\\ t<i}}^{n}\phi_n(\mathbf{w}_t, \mathbf{w}_i).\end{aligned}$$

Now,

$$\begin{aligned}\frac{1}{n^2}\sum_{t=1}^{n}\sum_{\substack{i=1\\ t<i}}^{n}\phi_n(\mathbf{w}_t, \mathbf{w}_i) &= \frac{1}{n^2}\sum_{t=1}^{n}\sum_{\substack{i=1\\ t<i}}^{n}\Phi_n(\mathbf{w}_t, \mathbf{w}_i) + \frac{1}{n^2}\sum_{t=1}^{n}\sum_{\substack{i=1\\ t<i}}^{n}(\theta_{ni} + \theta_{nt} - \theta_n)\\ &= Q_{231n} + Q_{232n} \qquad (26)\end{aligned}$$

where $\theta_{ni} = \int \phi_n(\mathbf{w}_t, \mathbf{w}_i)dP(\mathbf{w}_t)$, $\theta_n = \int \phi_n(\mathbf{w}_t, \mathbf{w}_i)dP(\mathbf{w}_t)dP(\mathbf{w}_i)$, $\Phi_n(\mathbf{w}_t, \mathbf{w}_i) = \phi_n(\mathbf{w}_t, \mathbf{w}_i) - \theta_{ni} - \theta_{nt} + \theta_n$ and $P(\mathbf{w}_t)$ be the probability measure associated with the vector $\mathbf{w}_t$. Then, $\Phi_n(\mathbf{w}_t, \mathbf{w}_i)$ is a symmetric function, and for fixed $\mathbf{w}_i$, $E(\Phi_n(\mathbf{w}_t, \mathbf{w}_i)) = 0$. By part (ii) of Lemma A.2 in Gao (2007), for $\delta > 0$, $E\left(\frac{1}{n^2}\sum_{t=1}^{n}\sum_{\substack{i=1\\ t<i}}^{n}\Phi_n(\mathbf{w}_t, \mathbf{w}_i)\right)^2 \leq \frac{C}{n^2}M^{1/(1+\delta)}$ where

$$M = \max_{1<i\leq n}\max\left\{E\left(|\Phi_n(\mathbf{w}_1, \mathbf{w}_i)|^{2(1+\delta)}\right), \int\int |\Phi_n(\mathbf{w}_1, \mathbf{w}_i)|^{2(1+\delta)}dP(\mathbf{w}_1)dP(\mathbf{w}_i)\right\}.$$

By the c_r and Cauchy-Schwartz inequalities we have that

$$E\left(|\Phi_n(\mathbf{w}_1, \mathbf{w}_i)|^{2(1+\delta)}\right) \leq C\left(E\left(|\phi_n(\mathbf{w}_1, \mathbf{w}_i)|^{2(1+\delta)}\right) + \int\int |\phi_n(\mathbf{w}_1, \mathbf{w}_i)|^{2(1+\delta)}dP(\mathbf{w}_1)dP(\mathbf{w}_i)\right)$$

and

$$\begin{aligned}\int\int |\Phi_n(\mathbf{w}_1, \mathbf{w}_i)|^{2(1+\delta)}dP(\mathbf{w}_1)dP(\mathbf{w}_i) \leq C\Bigg(&\int\int |\phi_n(\mathbf{w}_1, \mathbf{w}_i)|^{2(1+\delta)}dP(\mathbf{w}_1)dP(\mathbf{w}_i)\\ &+ E|\phi_n(\mathbf{w}_1, \mathbf{w}_i)|^{2(1+\delta)}\Bigg).\end{aligned}$$

Hence, we investigate the order of $E|\phi_n(\mathbf{w}_i,\mathbf{w}_i)|^{2(1+\delta)}$ for $t>i$. Note that,

$$\begin{aligned}
E|\phi_n(\mathbf{w}_i,\mathbf{w}_t)|^{2(1+\delta)} &\le CE|\psi_n(\mathbf{w}_i,\mathbf{w}_t)|^{2(1+\delta)}\\
&\le Ch_{3n}^{-(1+2\delta)}h_{1n}^{-d(1+\delta)}\frac{1}{h_{3n}h_{1n}^d}E\left(\frac{1}{f_{\mathbf{X}}(\mathbf{X}_i)^{2(1+\delta)}}K_3^{2(1+\delta)}\left(\frac{y_n-\varepsilon_i}{h_{3n}}\right)\right.\\
&\times K_1^{2(1+\delta)}\left(\frac{\mathbf{X}_t-\mathbf{X}_i}{h_{1n}}\right)\left(\frac{h^{1/2}(\mathbf{X}_t)}{h^{1/2}(\mathbf{X}_i)}\varepsilon_t\right)^{2(1+\delta)}\\
&+\left.\left(\frac{1}{2}h^{-1/2}(\mathbf{X}_i)(\mathbf{X}_t-\mathbf{X}_i)^Tm^{(2)}(\mathbf{Z}_{it})(\mathbf{X}_t-\mathbf{X}_i)\right)^{2(1+\delta)}\right)\\
&= C\,h_{3n}^{-(1+2\delta)}h_{1n}^{-d(1+\delta)}(Q_{2311}+Q_{2312}).
\end{aligned}$$

Since $t>i$ and ε_t is independent of $\mathbf{X}_t$ we have

$$\begin{aligned}
\frac{y_n}{1-F(y_n)}Q_{2311} &< CE(\varepsilon_t^{2(1+\delta)})\frac{y_n}{1-F(y_n)}\int\int\int K_3^{2(1+\delta)}(\psi_1)f(y_n-h_{3n}\psi_1)\left(\frac{h^{1/2}(\mathbf{X}_t)}{h^{1/2}(\mathbf{X}_t-h_{1n}\psi_2)}\right)^{2(1+\delta)}\\
&\times K_1^{2(1+\delta)}(\psi_2)f_{\mathbf{X}_i\mathbf{X}_t|\varepsilon_i=y_n-h_{3n}\psi_1}(\mathbf{X}_t-h_{1n}\psi_2,\mathbf{X}_t)d\psi_1d\psi_2d\mathbf{X}_t\\
&< CE(\varepsilon_t^{2(1+\delta)})\int\int\int K_3^{2(1+\delta)}(\psi_1)f(y_n-h_{3n}\psi_1)\frac{y_n-h_{3n}\psi_1}{1-F(y_n-h_{3n}\psi_1)}\\
&\times\frac{1-F(y_n-h_{3n}\psi_1)}{1-F(y_n)}K_1^{2(1+\delta)}(\psi_2)\left(\frac{h^{1/2}(\mathbf{X}_t)}{h^{1/2}(\mathbf{X}_t-h_{1n}\psi_2)}\right)^{2(1+\delta)}\frac{y_n}{y_n-h_{3n}\psi_1}\\
&\times\sup_{\varepsilon}f_{\mathbf{X}_i\mathbf{X}_t|\varepsilon_i=\varepsilon}(\mathbf{X}_t-h_{1n}\psi_2,\mathbf{X}_t)d\psi_1d\psi_2d\mathbf{X}_t.
\end{aligned}$$

Consequently, $Q_{2311}=O\left(\frac{1-F(y_n)}{y_n}\right)$ by A5 1) and 3). Now,

$$\begin{aligned}
\frac{y_n}{1-F(y_n)}Q_{2312} &< C\,\frac{y_n}{1-F(y_n)}\int K_3^{2(1+\delta)}(\psi_1)f(y_n-h_{3n}\psi_1)h_{1n}^{4(1+\delta)}\\
&\times\int\int\sup_{\varepsilon}K_1^{2(1+\delta)}(\psi_2)(\psi_2^T\psi_2)^{2(1+\delta)}f_{\mathbf{X}_i\mathbf{X}_t|\varepsilon_i=\varepsilon}(\mathbf{X}_t-h_{1n}\psi_2,\mathbf{X}_t)d\psi_1d\psi_2d\mathbf{X}_t\\
&= O(h_{1n}^{4(1+\delta)})
\end{aligned}$$

since $h^{1/2}(\mathbf{x}), f_{\mathbf{X}}(\mathbf{x})>0$, and every element of $m^{(2)}$ is uniformly bounded. Hence, $Q_{2312}=O\left(\frac{1-F(y_n)}{y_n}h_{1n}^{4(1+\delta)}\right)$ and $E(|\phi_n(\mathbf{w}_i,\mathbf{w}_t)|^{2(1+\delta)})=O\left(h_{3n}^{-(1+2\delta)}h_{1n}^{-d(1+2\delta)}\frac{1-F(y_n)}{y_n}\right)$. Now, for $t>i$

$$\int\int|\phi_n(\mathbf{w}_i,\mathbf{w}_t)|^{2(1+\delta)}dP(\mathbf{w}_i)dP(\mathbf{w}_t)\le\int\int|\psi_n(\mathbf{w}_i,\mathbf{w}_t)|^{2(1+\delta)}dP(\mathbf{w}_i)dP(\mathbf{w}_t)$$

$$\leq C\, h_{3n}^{-(1+2\delta)} h_{1n}^{-d(1+\delta)} \left(\frac{1}{h_{3n} h_{1n}^{d}} \int\int \frac{1}{f_{\mathbf{X}}(\mathbf{X}_i)^{2(1+\delta)}} \right.$$
$$\times K_3^{2(1+\delta)} \left(\frac{y_n - \varepsilon_i}{h_{3n}} \right) K_1^{2(1+\delta)} \left(\frac{\mathbf{X}_t - \mathbf{X}_i}{h_{1n}} \right) \left(\left(\frac{h^{1/2}(\mathbf{X}_t)}{h^{1/2}(\mathbf{X}_i)} \varepsilon_t \right)^{2(1+\delta)} \right.$$
$$\left. + \left(\frac{1}{2} h^{-1/2}(\mathbf{X}_i)(\mathbf{X}_t - \mathbf{X}_i)^T m^{(2)}(\mathbf{Z}_{it})(\mathbf{X}_t - \mathbf{X}_i) \right)^{2(1+\delta)} \right)$$
$$\left. \times\, dP(\mathbf{w}_i) dP(\mathbf{w}_t) \right) = C\, h_{3n}^{-(1+2\delta)} h_{1n}^{-d(1+2\delta)} (Q_{2313} + Q_{2314})$$

Since $\mathbf{X}_t$ and ε_t are independent $dP(\mathbf{w}_t) = f_{\mathbf{X}}(\mathbf{X}_t) f(\varepsilon_t) d\mathbf{X}_t d\varepsilon_t$ and

$$Q_{2313} = \frac{1}{h_{3n}} \int K_3^{2(1+\delta)} \left(\frac{y_n - \varepsilon_i}{h_{3n}} \right) f(\varepsilon_i) d\varepsilon_i E(\varepsilon_t^{2(1+\delta)}) \frac{1}{h_{1n}^{d}} \int\int \left(\frac{h^{1/2}(\mathbf{X}_t)}{h^{1/2}(\mathbf{X}_i)} \right)^{2(1+\delta)}$$
$$\times \frac{1}{f_{\mathbf{X}}(\mathbf{X}_i)^{2(1+\delta)}} K_1^{2(1+\delta)} \left(\frac{\mathbf{X}_t - \mathbf{X}_i}{h_{1n}} \right) f_{\mathbf{X}}(\mathbf{X}_i) f_{\mathbf{X}}(\mathbf{X}_t) d\mathbf{X}_t d\mathbf{X}_i.$$

Note that $E(\varepsilon_t^{2(1+\delta)}) < C$ and $\int K_3^{2(1+\delta)} \left(\frac{y_n - \varepsilon_i}{h_{3n}} \right) f(\varepsilon_i) d\varepsilon_i = O\left(\frac{1-F(y_n)}{y_n} \right)$. The remaining integral can be written as

$$\int\int \left(\frac{h^{1/2}(\mathbf{X}_t)}{h^{1/2}(\mathbf{X}_t - h_{1n}\psi)} \right)^{2(1+\delta)} \frac{1}{f_{\mathbf{X}}(\mathbf{X}_t - h_{1n}\psi)^{2(1+\delta)}} K_1^{2(1+\delta)} (\psi)\, f_{\mathbf{X}}(\mathbf{X}_t - h_{1n}\psi) f_{\mathbf{X}}(\mathbf{X}_t) d\psi d\mathbf{X}_t$$
$$\to \int K_1^{2(1+\delta)}(\psi) d\psi \int f_{\mathbf{X}}^{-2\delta}(\mathbf{X}_t) d\mathbf{X}_t < C$$

since $\sup_{\mathbf{x} \in \mathcal{G}} f_{\mathbf{X}}(\mathbf{x})$ and $h(\mathbf{x}) > C > 0$. Consequently, $Q_{2313} = O\left(\frac{1-F(y_n)}{y_n} \right)$. Given that all elements of $m^{(2)}(\mathbf{x})$ are uniformly bounded above, similar arguments give $Q_{2314} = O\left(\frac{1-F(y_n)}{y_n} h_{1n}^{4(1+\delta)} \right)$. Combining the orders of Q_{2313}, Q_{2314} we have $\int\int |\phi_n(\mathbf{w}_i, \mathbf{w}_t)|^{2+\delta} dP(\mathbf{w}_i) dP(\mathbf{w}_t) = O\left(\frac{1-F(y_n)}{y_n} h_{3n}^{-(1+2\delta)} h_{1n}^{-d(1+2\delta)} \right)$. Hence, we can write $E(Q_{231n}^2) = O\left(n^{-2} \left(\frac{1-F(y_n)}{y_n} \right)^{1/(1+\delta)} h_{3n}^{-(1+2\delta)/(1+\delta)} h_{1n}^{-d(1+2\delta)/(1+\delta)} \right)$ and consequently $Q_{231n} = O_p\left(n^{-1} \left(\frac{1-F(y_n)}{y_n} \right)^{1/2(1+\delta)} h_{3n}^{-(1+2\delta)/2(1+\delta)} h_{1n}^{-d(1+2\delta)/2(1+\delta)} \right)$. Thus,

$$\frac{n^{1/2}}{(1-F(y_n))^{1/2}} Q_{231n} = O_p\left(\left(n h_{3n}^{1+2\delta} h_{1n}^{d(1+2\delta)} (n(1-F(y_n)))^{\delta} y_n \right)^{-1/2(1+\delta)} \right).$$

Since $y_n \to \infty$ and $n(1-F(y_n)) \to \infty$ we have that $\frac{n^{1/2}}{(1-F(y_n))^{1/2}} Q_{231n} = o_p(1)$ provided $n h_{3n}^{1+2\delta} h_{1n}^{d(1+2\delta)} \to \infty$, which is verified given that $h_{1n} \propto n^{-\frac{1}{2s+d}}$ and $h_{3n} \propto n^{-\frac{s}{2(2s+d)}+\delta}$. We now consider $Q_{232n} = \frac{1}{n^2} \sum_{\substack{t=1 \\ t<i}}^{n} \sum_{i=1}^{n} (\theta_{ni} +$

$\theta_{nt}-\theta_n$). Note that

$$Q_{232n} = \frac{1}{2n^2}\sum_{t=1}^{n}\sum_{\substack{i=1\\ t\neq i}}^{n}(\theta_{ni}+\theta_{nt}-\theta_n) = \frac{n-1}{n^2}\sum_{t=1}^{n}\theta_{nt} - \frac{1}{2n^2}\frac{n^2-n}{2}\theta_n$$
$$= \frac{n-1}{n^2}\sum_{t=1}^{n}\int \phi_n(\mathbf{w}_t,\mathbf{w}_i)dP(\mathbf{w}_i) + O(1)\theta_n.$$

Let $\theta_n = \mathcal{I}_{1n}+\mathcal{I}_{2n}$ where

$$\mathcal{I}_{1n} = \frac{2}{h_{3n}h_{1n}^d}\int\int \frac{1}{f_{\mathbf{X}}(\mathbf{X}_i)}K_3\left(\frac{y_n-\varepsilon_i}{h_{3n}}\right)K_1\left(\frac{\mathbf{X}_t-\mathbf{X}_i}{h_{1n}}\right)\frac{h^{1/2}(\mathbf{X}_t)}{h^{1/2}(\mathbf{X}_i)}\varepsilon_t dP(\mathbf{w}_i)dP(\mathbf{w}_t)$$
$$\mathcal{I}_{2n} = \frac{1}{h_{3n}h_{1n}^d}\int\int \frac{1}{h^{1/2}(\mathbf{X}_i)f_{\mathbf{X}}(\mathbf{X}_i)}K_3\left(\frac{y_n-\varepsilon_i}{h_{3n}}\right)K_1\left(\frac{\mathbf{X}_t-\mathbf{X}_i}{h_{1n}}\right)(\mathbf{X}_t-\mathbf{X}_i)^T m^{(2)}(\mathbf{Z}_{it})(\mathbf{X}_t-\mathbf{X}_i)$$
$$\times\, dP(\mathbf{w}_i)dP(\mathbf{w}_t).$$

We note that $E(\mathcal{I}_{1n})=0$ since $E(\varepsilon_t)=0$ and

$$E(\mathcal{I}_{2n}) = \frac{2}{h_{3n}}E\left(K_3\left(\frac{y_n-\varepsilon_i}{h_{3n}}\right)\right)\frac{1}{h_{1n}^d}\int\int K_1\left(\frac{\mathbf{X}_t-\mathbf{X}_i}{h_{1n}}\right)\frac{1}{h^{1/2}(\mathbf{X}_i)}(\mathbf{X}_t-\mathbf{X}_i)^T m^{(2)}(\mathbf{Z}_{it})(\mathbf{X}_t-\mathbf{X}_i)$$
$$\times\, f_{\mathbf{X}}(\mathbf{X}_t)d\mathbf{X}_t d\mathbf{X}_i = \mathcal{I}_{21n}\mathcal{I}_{22n}.$$

Using Proposition 1.15 in Resnick (1987) we have $\frac{y_n}{1-F(y_n)}\mathcal{I}_{21n}\to -2k_0$. Furthermore, given that $h(\mathbf{x}) > C > 0$ and m and $f_{\mathbf{X}}$ have s bounded partial derivatives $\mathcal{I}_{22n} = O(h_{1n}^s)$. Hence, $\theta_n = O\left(h_{1n}^s\frac{1-F(y_n)}{y_n}\right)$. Now,

$$\frac{n-1}{n}\sum_{t=1}^{n}\int \phi_n(\mathbf{w}_t,\mathbf{w}_i)dP(\mathbf{w}_i) = (1+O(n^{-1}))\frac{1}{n}\sum_{t=1}^{n}\left(\int \psi_n(\mathbf{w}_t,\mathbf{w}_i)dP(\mathbf{w}_i)\right.$$
$$\left. + \int \psi_n(\mathbf{w}_t,\mathbf{w}_i)dP(\mathbf{w}_i)\right)$$
$$= (1+O(n^{-1}))\frac{1}{n}\sum_{t=1}^{n}\left(\mathcal{I}_{11t}+\mathcal{I}_{12t}+\mathcal{I}_{13t}\right)$$

where,

$$\mathcal{I}_{11t} = \frac{1}{h_{3n}}E\left(K_3\left(\frac{y_n-\varepsilon_i}{h_{3n}}\right)\right)\frac{1}{h_{1n}^d}\int K_1\left(\frac{\mathbf{X}_t-\mathbf{X}_i}{h_{1n}}\right)\frac{1}{h^{1/2}(\mathbf{X}_i)}d\mathbf{X}_i h^{1/2}(\mathbf{X}_t)\varepsilon_t$$
$$\mathcal{I}_{12t} = \frac{1}{h_{3n}}E\left(K_3\left(\frac{y_n-\varepsilon_i}{h_{3n}}\right)\right)\frac{1}{h_{1n}^d}\int K_1\left(\frac{\mathbf{X}_t-\mathbf{X}_i}{h_{1n}}\right)\frac{1}{2\,h^{1/2}(\mathbf{X}_i)}$$
$$\times\,(\mathbf{X}_t-\mathbf{X}_i)^T m^{(2)}(\mathbf{Z}_{it})(\mathbf{X}_t-\mathbf{X}_i)d\mathbf{X}_i$$
$$\mathcal{I}_{13t} = \frac{1}{h_{3n}}E\left(K_3\left(\frac{y_n-\varepsilon_t}{h_{3n}}\right)\right)\frac{1}{f_{\mathbf{X}}(\mathbf{X}_t)}\frac{1}{h_{1n}^d}\int K_1\left(\frac{\mathbf{X}_i-\mathbf{X}_t}{h_{1n}}\right)\frac{1}{2\,h^{1/2}(\mathbf{X}_t)}$$
$$\times\,(\mathbf{X}_i-\mathbf{X}_t)^T m^{(2)}(\mathbf{Z}_{ti})(\mathbf{X}_i-\mathbf{X}_t)f_{\mathbf{X}}(\mathbf{X}_i)d\mathbf{X}_i.$$

From the order of $\mathcal{I}_{21n}$ and the fact that $\begin{pmatrix} \mathbf{X}_t^T & \varepsilon_t \end{pmatrix}$ is strictly stationary, $\mathbf{X}_t$ independent of ε_t, we have for $t \neq i$ that $\frac{1}{n}\sum_{t=1}^{n}\mathcal{I}_{11t} = O_p\left(n^{-1/2}\frac{1-F(y_n)}{y_n}\right)$. Now,

$$\frac{y_n}{1-F(y_n)}\frac{1}{n}\sum_{t=1}^{n}\mathcal{I}_{12t} = \frac{y_n}{1-F(y_n)}E\left(K_3\left(\frac{y_n-\varepsilon_i}{h_{3n}}\right)\right)$$
$$\times\frac{1}{n}\sum_{t=1}^{n}\int\frac{h_{1n}^2}{2h^{1/2}(\mathbf{X}_t-h_{1n}\phi_t)}K_1(\phi_t)\phi_t^T m^{(2)}(\mathbf{X}_t+\lambda h_{1n}\phi_t)\phi_t d\phi_t$$
$$= O(1)\frac{1}{n}\sum_{t=1}^{n}\gamma_{nt}$$

and $E\left(\left(\frac{1}{n}\sum_{t=1}^{n}\gamma_{nt}\right)^2\right) = \frac{1}{n^2}\sum_{t=1}^{n}E(\gamma_{nt}^2)+\frac{1}{n^2}\sum_{t=1}^{n}\sum_{\substack{i=1\\t\neq i}}^{n}E(\gamma_{nt}\gamma_{ni})$. From previous arguments we immediately conclude that $E(\gamma_{nt}^2)=O(h_{1n}^{2s})$. For the second term, note that $\left|\frac{1}{n^2}\sum_{t=1}^{n}\sum_{\substack{i=1\\t\neq i}}^{n}E(\gamma_{nt}\gamma_{ni})\right| \leq \frac{1}{n}\sum_{t=1}^{n}\sup_t\frac{1}{n}\sum_{\substack{i=1\\t\neq i}}^{n}|E(\gamma_{nt}\gamma_{ni}))|$. Letting $J_n=\sum_{i=1}^{n}|E(\gamma_{nt}\gamma_{n,t+i})|$, we observe that for $\delta_1>2$, by Theorem 3(1) in Doukhan (1994) we have

$$|E(\gamma_{nt}\gamma_{n,t+i})| \leq 8(E(|\gamma_{nt}|^{\delta_1})^{2/\delta_1})\alpha(i)^{1-2/\delta_1} \tag{27}$$

where $E(|\gamma_{nt}|^{\delta_1})=O(h_{1n}^{s\delta_1})$. Hence, by assumption A2 1) $J_n \leq O(h_{1n}^{2s})\sum_{i=1}^{n}\alpha(i)^{1-2/\delta_1}=O(h_{1n}^{2s})$. Similarly, $\sum_{i=1}^{n}|E(\gamma_{nt}\gamma_{n,t-i})| = O(h_{1n}^{2s})$ and consequently $\left|\frac{1}{n^2}\sum_{t=1}^{n}\sum_{\substack{i=1\\t\neq i}}^{n}E(\gamma_{nt}\gamma_{ni})\right| \leq O(n^{-1}h_{1n}^{2s})$. As a result, we write $E\left(\left(\frac{1}{n}\sum_{t=1}^{n}\gamma_{nt}\right)^2\right)=O(n^{-1}h_{1n}^{2s})$ and $\frac{y_n}{1-F(y_n)}\frac{1}{n}\sum_{t=1}^{n}\mathcal{I}_{12t}=O_p\left(n^{-1/2}h_{1n}^{s}\right)$. Now, we let $\frac{1}{n}\sum_{t=1}^{n}\mathcal{I}_{13t}=O(1)\frac{1}{n}\sum_{t=1}^{n}\eta_{nt}$ where $\eta_{nt}=\frac{1}{f_{\mathbf{X}}(\mathbf{X}_t)}\frac{1}{2h^{1/2}(\mathbf{X}_t)}\frac{1}{h_{1n}^d}\int K_1\left(\frac{\mathbf{X}_i-\mathbf{X}_t}{h_{1n}}\right)(\mathbf{X}_i-\mathbf{X}_t)^T m^{(2)}(\mathbf{Z}_{it})(\mathbf{X}_i-\mathbf{X}_t)f_{\mathbf{X}}(\mathbf{X}_i)d\mathbf{X}_i$. It is immediate from arguments used for $\mathcal{I}_{12t}$ that $\frac{1}{n^2}\sum_{t=1}^{n}E(\eta_{nt}^2)=O\left(n^{-1}h_{3n}h_{1n}^{2s}\frac{1-F(y_n)}{y_n}\right)$. Furthermore, for $\delta_1>2$ we have

$$\sum_{i=1}^{n}E(|\eta_{nt}\eta_{n,t+1}|) \leq 8\sum_{i=1}^{n}(E(|\eta_{nt}|^{\delta_1}))^{2/\delta_1}\alpha(i)^{1-2/\delta_1} = O\left(h_{3n}^{2(1-\delta_1)/\delta_1}h_{1n}^{2s}\left(\frac{1-F(y_n)}{y_n}\right)^{2/\delta_1}\right)$$
$$= o\left(h_{3n}^{-2}h_{1n}^{2s}\left(1-F(y_n)\right)^{2/\delta_1}\right)$$

as $\delta_1>2$ and $y_n\to\infty$. Thus, provided that $\frac{h_{1n}^{2s}}{h_{3n}^2(1-F(y_n)}\to 0$, which is verified when $h_{1n}\propto n^{-1/(2s+d)}$, $h_{3n}\propto n^{-s/2(2s+d)+\delta}$, we have $\sum_{i=1}^{n}E(|\eta_{nt}\eta_{n,t+i}|)=o(1-F(y_n))$. Thus, $E\left(\left(\frac{1}{n}\sum_{t=1}^{n}\eta_{nt}\right)\right)^2=O\left(n^{-1}h_{3n}^{-1}h_{1n}^{-2s}\frac{1-F(y_n)}{y_n}\right)+$

$o(n^{-1}(1-F(y_n)))$. Hence, $\frac{1}{n}\sum_{t=1}^{n}\mathcal{I}_{13t} = o_p\left(\frac{(1-F(y_n))^{1/2}}{n^{1/2}}\right)$ and

$$\frac{n-1}{n^2}\sum_{t=1}^{n}\int \phi_n(\mathbf{w}_i,\mathbf{w}_t)dP(\mathbf{w}_t) = O_p\left(n^{-1/2}\frac{1-F(y_n)}{y_n}\right) + O_p\left(n^{-1/2}h_{1n}^s\frac{1-F(y_n)}{y_n}\right)$$
$$+ o_p\left(n^{-1/2}(1-F(y_n))^{1/2}\right) = o_p\left(n^{-1/2}(1-F(y_n))^{1/2}\right).$$

Consequently, $Q_{232n} = O\left(h_{1n}^s\frac{1-F(y_n)}{y_n}\right) + o_p\left(n^{-1/2}(1-F(y_n))^{1/2}\right)$ and

$$Q_{23n} = O_p\left(h_{1n}^s\frac{1-F(y_n)}{y_n}\right) + o_p\left(n^{-1/2}(1-F(y_n))\right).$$

Now, whenever $i=t$ we have $Q_{23n} = -\frac{1}{nh_{1n}^d}K_1(\mathbf{0})\frac{1}{n}\sum_{t=1}^{n}\nu_{nt}$ where $\nu_{nt} = \frac{1}{n}\sum_{t=1}^{n}K_3\left(\frac{y_n-\varepsilon_t}{h_{3n}}\right)\frac{\varepsilon_t}{f_{\mathbf{X}}(\mathbf{X}_t)}$. $E(\nu_{nt}^2) = O(y_n(1-F(y_n))h_{3n}^{-1}) = O(h_{3n}^{-1})$ since $y_n(1-F(y_n)) = O(1)$ if $\alpha > 1$. Also, for $\delta_1 > 2$ we have $\sum_{i=1}^{n}|E(\nu_{nt}\nu_{n,t+i})| \leq 8(E(|\nu_{nt}|_1^{\delta}))^{2/\delta_1}\alpha(i)^{1-2/\delta_1}$ where $E(|\nu_{nt}|_1^{\delta}) = O(h_{3n}^{-\delta_1})$ if $E(|\varepsilon|^{\delta_1}) < \infty$. Hence, $\sum_{i=1}^{n}|E(\nu_{nt}\nu_{n,t+i})| = O(h_{3n}^{-2})$ and consequently $E\left(\left(\frac{1}{n}\sum_{t=1}^{n}\nu_{nt}\right)^2\right) = O((nh_{3n}^2)^{-1})$. Thus, we have that $Q_{23n} = O_p\left(n^{-3/2}h_{1n}^{-d}h_{3n}^{-1}\right)$ and $Q_{23n} = o_p\left(n^{-1/2}(1-F(y_n))^{1/2}\right)$ since $n(1-F(y_n)) \to \infty$ and $n^{-1/2}h_{1n}^{-d}h_{3n}^{-1} = o(1)$. Overall, we have $Q_{23n} = o_p\left(n^{-1/2}(1-F(y_n))^{1/2}\right) + O_p\left(h_{1n}^s(1-F(y_n))\right)$ and consequently

$$\frac{n^{1/2}}{(1-F(y_n))^{1/2}}Q_{23n} = o_p(1). \tag{28}$$

We now consider Q_{24n} which can be written as

$$Q_{24n} = \frac{1}{nh_{3n}}\sum_{t=1}^{n}K_3\left(\frac{y_n-\varepsilon_t}{h_{3n}}\right)\left(\frac{h^{1/2}(\mathbf{X}_t)}{\hat{h}^{1/2}(\mathbf{X}_t)}-1\right)\left(\chi_{\{\hat{h}(\mathbf{X}_t)>0\}}-1\right)\varepsilon_t$$
$$+\frac{1}{nh_{3n}}\sum_{t=1}^{n}K_3\left(\frac{y_n-\varepsilon_t}{h_{3n}}\right)\left(\frac{h^{1/2}(\mathbf{X}_t)}{\hat{h}^{1/2}(\mathbf{X}_t)}-1\right)\varepsilon_t = Q_{241n}+Q_{242n}.$$

By Lemmas 3, 4 and Corollary 1 we have $\sup_{\mathbf{x}\in\mathcal{G}}\left|\chi_{\{\hat{h}(\mathbf{X}_t)>0\}}-1\right| = O_p(L_{1n}+L_{2n})$ and $\sup_{\mathbf{x}\in\mathcal{G}}\left|\frac{h^{1/2}(\mathbf{X}_t)}{\hat{h}^{1/2}(\mathbf{X}_t)}-1\right| = O_p(L_{1n}+L_{2n})$. Thus, $Q_{41n} \leq O_p(L_{1n}^2+L_{2n}^2)\frac{1}{nh_{3n}}\sum_{t=1}^{n}\left|K_3\left(\frac{y_n-\varepsilon_t}{h_{3n}}\right)\right||\varepsilon_t|$. Since $1-F(y_n) = o(1)$, by Proposition 1.15 in Resnick (1987) we have $\frac{1}{nh_{3n}}\sum_{t=1}^{n}\left|K_3\left(\frac{y_n-\varepsilon_t}{h_{3n}}\right)\right||\varepsilon_t| = o_p(1)$. Hence, $Q_{241n} = o_p(L_{1n}^2+L_{2n}^2)$ and $\frac{n^{1/2}}{(1-F(y_n))^{1/2}}Q_{241n} = o_p(1)$. Now, we can write Q_{242n} as

$$Q_{242n} = -\frac{1}{2nh_{3n}}\sum_{t=1}^{n}K_3\left(\frac{y_n-\varepsilon_t}{h_{3n}}\right)\frac{1}{h(\mathbf{X}_t)}(\hat{h}(\mathbf{X}_t)-h(\mathbf{X}_t)) + o_p\left(\frac{(1-F(y_n))^{1/2}}{n^{1/2}}\right),$$

hence it suffices to study the order of the first term. Given that $\hat{h}(\mathbf{X}_t)$ is a local linear estimator, and under assumptions A1-A3 we can write

$$\hat{h}(\mathbf{X}_t) - h(\mathbf{X}_t) = (1+o_p(1))\frac{1}{nh_{2n}^d f_{\mathbf{X}}(\mathbf{X}_t)}\sum_{i=1}^n K_2\left(\frac{\mathbf{X}_i - \mathbf{X}_t}{h_{2n}}\right)\big(h(\mathbf{X}_i)(\varepsilon_i^2 - 1)$$
$$+ \frac{1}{2}(\mathbf{X}_i - \mathbf{X}_t)^T h^{(2)}(\mathbf{Z}_{it})(\mathbf{X}_i - \mathbf{X}_t) - \frac{2}{nh_{1n}^d f_{\mathbf{X}}(\mathbf{X}_i)} h^{1/2}(\mathbf{X}_i)\varepsilon_i \sum_{j=1}^n K_1\left(\frac{\mathbf{X}_j - \mathbf{X}_i}{h_{1n}}\right)$$
$$\times (h^{1/2}(\mathbf{X}_j)\varepsilon_j + \frac{1}{2}(\mathbf{X}_j - \mathbf{X}_i)^T m^{(2)}(\mathbf{Z}_{ji})(\mathbf{X}_j - \mathbf{X}_i))(1+o_p(1)) + O_p(L_{2n}^2)\Big). \quad (29)$$

Following arguments that are similar to those used in the study of Q_{23n}, we show that

$$\frac{n^{1/2}}{(1-F(y_n))^{1/2}} Q_{24n} = o_p(1). \quad (30)$$

Due to the similarity of the arguments, we omit the details. However, it is worth pointing out that in the case of Q_{24n}, three dimensional U-Statistics appear due to the structure of (29). As such, we must use part (i) of Lemma A.2 in Gao (2007). This in turn requires assumptions A5 2) and 4).

Now, let $A_t = \left\{\omega : \left|\chi_{\hat{h}(\mathbf{X}_t)>0)} - 1\right| = 0\right\}$ and $B_t = \{\omega : h(\mathbf{X}_t) - \hat{h}(\mathbf{X}_t) < \delta\}$ for some $\delta > 0$. Clearly, $A_t^c \subseteq B_t^c$ so that $\chi_{A_t^c} \le \chi_{B_t^c} \le \frac{h(\mathbf{X}_t) - \hat{h}(\mathbf{X}_t)}{\delta}$. Thus,

$$\begin{aligned} Q_{25n} &= \frac{1}{nh_{3n}}\sum_{t=1}^n K_3\left(\frac{y_n - \varepsilon_t}{h_{3n}}\right)\left(\chi_{\hat{h}(\mathbf{X}_t)>0)} - 1\right)\varepsilon_t \chi_{A_t^c} \\ &\le \frac{1}{nh_{3n}}\sum_{t=1}^n \left|K_3\left(\frac{y_n - \varepsilon_t}{h_{3n}}\right)\right|\left|\chi_{\hat{h}(\mathbf{X}_t)>0)} - 1\right| |\varepsilon_t| \frac{h(\mathbf{X}_t) - \hat{h}(\mathbf{X}_t)}{\delta} \\ &\le O_p(L_{1n}^2 + L_{2n}^2)\frac{1}{nh_{3n}}\sum_{t=1}^n \left|K_3\left(\frac{y_n - \varepsilon_t}{h_{3n}}\right)\right| |\varepsilon_t| \end{aligned}$$

where the last inequality follows from Lemma 3 and Corollary 1. Given that $\frac{1}{nh_{3n}}\sum_{t=1}^n \left|K_3\left(\frac{y_n - \varepsilon_t}{h_{3n}}\right)\right| |\varepsilon_t| = O_p\left(\frac{1-F(y_n)}{y_n}\right)$ we have

$$\frac{n^{1/2}}{(1-F(y_n))^{1/2}} Q_{25n} = o_p(1). \quad (31)$$

Combining the orders obtained in equations (23), (24), (28), (30) and (31) we have

$$\frac{n^{1/2}}{(1-F(y_n))^{1/2}} Q_{2n} = o_p(1). \quad (32)$$

Lastly, we show that $\frac{n}{\sqrt{N}}(Q_{1n} - F(y_n)) \xrightarrow{d} N(0,1)$. First, we put $q_{1n} = \frac{1}{h_{3n}}\int_{-\infty}^{y_n} K_3\left(\frac{y - \varepsilon_t}{h_{3n}}\right) dy$ and write

$$\begin{aligned} \frac{n}{\sqrt{N}}(Q_{1n} - F(y_n)) &= \sum_{t=1}^n \frac{1}{\sqrt{n(1-F(y_n))}}(q_{1n} - E(q_{1n})) + \sum_{t=1}^n \frac{1}{\sqrt{n(1-F(y_n))}}(E(q_{1n}) - F(y_n)) \\ &= I_{1n} + I_{2n}. \end{aligned}$$

Clearly, $E\left(\frac{1}{\sqrt{n(1-F(y_n))}}\left(q_{1n}-E(q_{1n})\right)\right)=0$ and $V\left(\frac{1}{\sqrt{n(1-F(y_n))}}\left(q_{1n}-E(q_{1n})\right)\right)=\frac{s_n^2}{n(1-F(y_n))}$ where

$$s_n^2=\int \frac{1}{h_{3n}} b\left(\frac{y_n-u}{h_{3n}}\right) F(u)du-\left(\int \frac{1}{h_{3n}} K_2\left(\frac{y_n-u}{h_{3n}}\right) F(u)du\right)^2$$

and $b(x)=2K_3(x)\int_{-\infty}^{x} K_3(y)dy$. Define $s^2=F(y_n)(1-F(y_n))$ and write $\frac{s_n^2}{(1-F(y_n))}=\frac{s_n^2-s^2}{1-F(y_n)}+F(y_n)$. Since, $\frac{s_n^2-s^2}{1-F(y_n)}=o(h_{3n})$ and $F(y_n)\to 1$ as $n\to\infty$ we have $\frac{s_n^2}{1-F(y_n)}\to 1$. By Liapounov's CLT, $I_{1n}\overset{d}{\to} N(0,1)$ provided that $nE(|Z_{in}|^3)\to 0$ as $n\to\infty$, where

$$Z_{in}=\frac{1}{\sqrt{n(1-F(y_n))}}\left(\frac{1}{h_{3n}}\int_{-\infty}^{y_n} K_3\left(\frac{y-\varepsilon_t}{h_{3n}}\right)dy-E\left(\frac{1}{h_{3n}}\int_{-\infty}^{y_n} K_3\left(\frac{y-\varepsilon_t}{h_{3n}}\right)dy\right)\right).$$

The condition is verified by noting that

$$\left|\left(\frac{1}{h_{3n}}\int_{-\infty}^{y_n} K_3\left(\frac{y-\varepsilon_t}{h_{2n}}\right)dy-E\left(\frac{1}{h_{3n}}\int_{-\infty}^{y_n} K_3\left(\frac{y-\varepsilon_t}{h_{3n}}\right)dy\right)\right)\right|\le 2$$

since $\frac{1}{h_{3n}}\int_{-\infty}^{y_n} K_3\left(\frac{y-\varepsilon_t}{h_{2n}}\right)dy\le 1$. Consequently, $|Z_{in}|\le \frac{2}{\sqrt{n(1-F(y_n))}}$ and

$$n\,E(|Z_{in}|^3)\le \frac{2n}{\sqrt{n(1-F(y_n))}}\frac{s_n^2}{n(1-F(y_n))}\to 0 \text{ as } n\to\infty.$$

Integrating by parts we have

$$\begin{aligned}|E\left(q_{1n}\right)-F(y_n)|=&\left|\int(-h_{3n})\psi K_3(\psi)f(y_n)+\sum_{j=1}^{m_1-1}\frac{(-h_{3n}\psi)^{j+1}}{(j+1)!}\frac{d}{dy_n^j}f(y_n)\right.\\&\left.+\frac{(-h_{3n}\psi)^{m_1+1}}{(m_1+1)!}\frac{d}{dy_n^{m_1}}f(y_n^*)d\psi\right|,\end{aligned}$$

where $y_n^*=\lambda(y_n-h_{3n}\psi)+(1-\lambda)y_n$ for some $\lambda\in(0,1)$. Since K_3 is an m_1^{th}-order kernel and $\left|\frac{d}{du^{m_1}}f(u)\right|<C$, we have that $|E\left(q_{1n}\right)-F(y_n)|\le C\,\frac{h_{3n}^{m_1+1}}{(m_1+1)!}\int|\psi^{m_1+1}K_3(\psi)|d\psi=O(h_{3n}^{m_1+1})$. Hence, $I_{2n}=O\left(\frac{n}{\sqrt{N}}h_{3n}^{m_1+1}\right)=o(1)$, given the orders of h_{3n} and N, and

$$\frac{n}{\sqrt{N}}(Q_{1n}-F(y_n))\overset{d}{\to} N(0,1). \tag{33}$$

Equations (20), (32) and (33) show that $\frac{n}{\sqrt{N}}(\tilde{F}(y_n)-F(y_n))\overset{d}{\to} N(0,1)$, and by consequence $T_{1n}=O_p(1)$ which completes the proof. □

Lemma 5. *Let $a_n=1-\frac{N}{n}$ and for $i=1,\cdots,N$ define $Z_i=\varepsilon_i-q_n(a_n)$ whenever $\varepsilon_i>q_n(a_n)$ and for $i=1,\cdots,N_1$ define $Z_i'=\varepsilon_i-q(a_n)$ whenever $\varepsilon_i>q(a_n)$. If $\Delta_\sigma=\frac{1}{N}\sum_{i=1}^{N}\frac{\partial}{\partial\sigma}log\,g(Z_i;\sigma_N,k_0)\sigma_N-$*

$\frac{1}{N}\sum_{i=1}^{N_1}\frac{\partial}{\partial\sigma}log\, g(Z_i';\sigma_N,k_0)\sigma_N$ *and* $\Delta_k = \frac{1}{N}\sum_{i=1}^{N}\frac{\partial}{\partial k}log\, g(Z_i;\sigma_N,k_0) - \frac{1}{N}\sum_{i=1}^{N_1}\frac{\partial}{\partial k}log\, g(Z_i';\sigma_N,k_0)$, *then* $N^{1/2}\Delta_\sigma = b_1\sqrt{N}\frac{q_n(a_n)-q(a_n)}{q(a_n)} + o_p(1)$ *and* $N^{1/2}\Delta_k = b_2\sqrt{N}\frac{q_n(a_n)-q(a_n)}{q(a_n)} + o_p(1)$, *where* $b_1 = -\frac{\alpha(1+\alpha)}{2+\alpha}$, $b_2 = \left(-\frac{\alpha^2(1+\alpha)}{2+\alpha} + \frac{\alpha^3}{1+\alpha}\right)$.

Proof. The proof is identical to that of Lemma 3 in Martins-Filho et al. (2014) by substituting their $U_{(n-N)}$ with $\varepsilon_{(n-N)}$. □

Lemma 6. $E\left(log\left(1-\frac{k_0Z_i'}{\sigma_N}\right)\left(1-\frac{k_0Z_i'}{\sigma_N}\right)^{-1}\left(\frac{k_0Z_i'}{\sigma_N}\right)\right) = -\frac{1}{\alpha} + \frac{\alpha}{(1+\alpha)^2} + O(\phi(\varepsilon_{(n-N)}))$

Proof. The proof is identical to that of Lemma 4 in Martins-Filho et al. (2014) by substituting their $U_{(n-N)}$ with $\varepsilon_{(n-N)}$. □

References

Bosq, D., 1996. Nonparametric statistics for stochastic processes: estimation and prediction. No. 110 in Lecture Notes in Statistics. Springer Verlag, Berlin.

Doukhan, P., 1994. Mixing: properties and examples. Springer-Verlag, New York.

Gao, J., 2007. Nonlinear time series: nonparametric and parametric methods. Chapman and Hall, New York.

Martins-Filho, C., Yao, F., 2009. Nonparametric regression estimation with general parametric error covariance. Journal of Multivariate Analysis 100, 309–333.

Martins-Filho, C., Yao, F., Torero, M., 2013. Nonparametric estimation of conditional value-at-risk and expected shortfall based on extreme value theory. http://spot.colorado.edu/∼martinsc/Research.html., University of Colorado at Boulder.

Martins-Filho, C., Yao, F., Torero, M., 2014. High order conditional quantile estimation based on nonparametric models of regression. Econometric Reviews, forthcoming.

Resnick, S. I., 1987. Extreme values, regular variation and point processes. Springer Verlag, New York.

White, H., 2001. Asymptotic Theory for Econometricians, 2nd Edition. Academic Press, San Diego.