# Supplementary material to “Simultaneous inference for Berkson errors-in-variables regression under fixed design”

Katharina Proksch* Nicolai Bissantz† Hajo Holzmann‡

This paper contains supplementary material for the main document *Simultaneous inference for Berkson errors-in-variables regression under fixed design*, which was omitted therein for the sake of brevity. For better reference, we briefly recall the framework and give the relevant assumptions. All numbers in references to equations, lemmas or assumptions coincide with those given in the main document, except inequalities (S1), (S2) and (S3).

The Berkson errors-in-variables model with fixed design that we shall consider is given by

$$Y_j = g(w_j + \Delta_j) + \varepsilon_j, \tag{3}$$

where $w_j = j/(n a_n)$, $j = -n, \dots, n$, are the design points on a regular grid, $a_n$ is a design parameter that satisfies $a_n \to 0$, $na_n \to \infty$, and $\Delta_j$ and $\varepsilon_j$ are unobserved centered, independent and identically distributed errors for which $\mathrm{Var}[\varepsilon_1] = \sigma^2 > 0$ and $\mathbb{E}|\varepsilon_1|^M < \infty$ for some $M > 2$. The density $f_\Delta$ of the errors $\Delta_j$ is assumed to be known. Identification of $g$ on a given interval requires an infinitely supported design density if the error density is of infinite support. This corresponds to our assumption that asymptotically, the fixed design exhausts the whole real line. If we define $\gamma$ as the convolution of $g$ and $f_\Delta(-\cdot)$, that is, $\gamma(w) = \int_{\mathbb{R}} g(z) f_\Delta(z-w)\, dz$, then $\mathbb{E}[Y_j] = \gamma(w_j)$, and the calibrated regression model (Carroll et al., 2006) associated with (3) is given by

$$Y_j = \gamma(w_j) + \eta_j, \qquad \eta_j = g(w_j + \Delta_j) - \gamma(w_j) + \varepsilon_j. \tag{4}$$

Here the errors $\eta_j$ are independent and centered as well but no longer identically distributed since their variances $\nu^2(w_j) = \mathbb{E}[\eta_j^2]$ depend on the design points. To be precise, we have that

$$\nu^2(w_j) = \int \big(g(w_j + \delta) - \gamma(w_j)\big)^2 f_\Delta(\delta)\, d\delta + \sigma^2 \geq \sigma^2 > 0. \tag{5}$$

# I Proof of Lemma 5 in the main document

**Assumption 1.** The functions $g$ and $f_\Delta$ satisfy

*Department of Applied Mathematics, University of Twente, The Netherlands; ✉ k.proksch@utwente.nl

†Fakultät für Mathematik, Ruhr-Universität Bochum, Germany

‡Fachbereich Mathematik und Informatik, Philipps-Universität Marburg, Germany

(i) $g \in \mathscr{W}^m(\mathbb{R}) \cap L^r(\mathbb{R}) \quad$ for all $\quad r \leq M$ and for some $m > 5/2$.

(ii) $f_\Delta$ is a bounded, continuous, square-integrable density.

(iii) The function $\gamma$ decays sufficiently fast in the following sense:

$$\int_{|z|>1/a_n} \langle z \rangle^s |\gamma(z)|^2 \, dz < \infty,$$

for some $s > 1/2$, where $s$ may depend on $n$.

**Assumption 2.** Let $\Phi_k \in C^2(\mathbb{R})$ be symmetric, $\Phi_k(t) \equiv 1$ for all $t \in [-D,D]$, $0 < D < 1$, $|\Phi_k(t)| \leq 1$ and $\Phi_k(t) = 0$, $|t| > 1$.

**Lemma 5.** Let Assumptions 1 and 2 be satisfied. Further assume that $h/a_n \to 0$ as $n \to \infty$.

(i) Then for bias, we have that

$$\sup_{x\in[0,1]} \left|\mathbb{E}\left[\hat{g}_n(x;h)\right] - g(x)\right| = O\left(h^{m-\frac{1}{2}} + \frac{1}{na_n h^{\beta+\frac{3}{2}}}\right) + \begin{cases} o\left(a_n^{s+1/2} h^{1-\beta}\right), & \text{Ass. 3, (S),} \\ o\left(a_n^{s+1/2} h^{-\beta}\right), & \text{Ass. 5, (W).} \end{cases}$$

(ii) a) For the variance if Assumption 5, (W) holds and $na_n h^{1+\beta} \to \infty$, then we have that

$$\frac{\sigma^2}{2C\pi}(1+O(a_n)) \leq na_n h^{1+2\beta} \operatorname{Var}[\hat{g}_n(x;h)] \leq \frac{2^\beta \sup_{x\in\mathbb{R}} \nu^2(x)}{c\pi}.$$

(ii) b) If actually Assumption 3, (S) holds and $na_n h^{1+\beta} \to \infty$, then

$$\frac{\nu^2(x)}{C\pi}(1+O(a_n)) \leq na_n h^{1+2\beta} \operatorname{Var}[\hat{g}_n(x;h)] \leq \frac{\nu^2(x)}{c\pi}(1+O(h/a_n)),$$

Here $c, C$ and $\beta$ are the constants from Assumption 5 respectively 3.

**Proof of Lemma 5 in the main document**

(i) We have that

$$\begin{aligned}
\mathbb{E}\left[\hat{g}_n(x;h)\right] &= \frac{1}{na_n h} \sum_{j=-n}^{n} \gamma(w_j) K\left(\frac{w_j - x}{h}; h\right) \\
&= \frac{1}{h} \sum_{j=-n}^{n} \int_{w_j}^{w_j + \frac{1}{na_n}} \gamma(w_j) K\left(\frac{w_j - x}{h}; h\right) dz \\
&= \int_{-\frac{1}{a_n}}^{\frac{1}{a_n} + \frac{1}{na_n}} \gamma(z) K\left(\frac{z - x}{h}; h\right) dz + R_{n,1}(x) + R_{n,2}(x),
\end{aligned}$$

where

$$R_{n,1}(x) = \frac{1}{h}\sum_{j=-n}^{n}\int_{w_j}^{w_j+\frac{1}{na_n}} \frac{\mathrm{d}}{\mathrm{d}u}\left(\gamma(u)K\left(\frac{u-x}{h};h\right)\right)\Big|_{u=z}(w_j-z)\,dz$$

and

$$R_{n,2}(x) = \frac{1}{2h}\sum_{j=-n}^{n}\int_{w_j}^{w_j+\frac{1}{na_n}} \frac{\mathrm{d}^2}{\mathrm{d}u^2}\left(\gamma(u)K\left(\frac{u-x}{h};h\right)\right)\Big|_{u=\tilde{w}_j(z)}(w_j-z)^2\,dz.$$

Then,

$$\begin{aligned}
R_{n,1}(x) &= \frac{1}{h}\sum_{j=-n}^{n}\int_{w_j}^{w_j+\frac{1}{na_n}} \gamma'(z)K\left(\frac{z-x}{h};h\right)(w_j-z)\,dz \\
&\quad -\frac{1}{h^2}\sum_{j=-n}^{n}\int_{w_j}^{w_j+\frac{1}{na_n}} \gamma(z)K'\left(\frac{z-x}{h};h\right)(w_j-z)\,dz \\
&=: R_{n,1,1}(x)+R_{n,1,2}(x).
\end{aligned}$$

Now,

$$na_nh|R_{n,1,1}(x)| \le \int_{-\frac{1}{a_n}}^{\frac{1}{a_n}}\left|\gamma'(z)K\left(\frac{z-x}{h};h\right)\right|dz \le \|\gamma'\|_2\left\|K\left(\frac{\cdot-x}{h};h\right)\right\|_2 = O\left(\frac{1}{h^{-\frac{1}{2}+\beta}}\right).$$

Analogously,

$$|R_{n,1,2}(x)| = O\left(\frac{1}{na_nh^{\frac{3}{2}+\beta}}\right).$$

Furthermore

$$\begin{aligned}
|R_{n,2}(x)| &\le \left[\frac{1}{n^2a_n^2h}\sum_{j=-n}^{n}\int_{w_j}^{w_j+\frac{1}{na_n}}\left|\gamma''(\tilde{w}_j(z))K\left(\frac{\tilde{w}_j(z)-x}{h};h\right)\right|dz\right. \\
&\quad \left.+\frac{2}{n^2a_n^2h^2}\sum_{j=-n}^{n}\int_{w_j}^{w_j+\frac{1}{na_n}}\left|\gamma'(\tilde{w}_j(z))K'\left(\frac{\tilde{w}_j(z)-x}{h};h\right)\right|dz\right] \\
&\quad +\frac{1}{n^2a_n^2h^3}\sum_{j=-n}^{n}\int_{w_j}^{w_j+\frac{1}{na_n}}\left|\gamma(\tilde{w}_j(z))K''\left(\frac{\tilde{w}_j(z)-x}{h};h\right)\right|dz \\
&=: [R_{n,2,1}(x)]+R_{n,2,2}(x).
\end{aligned}$$

Then, $R_{n,2,1}(x) = O\left(\frac{1}{n^2a_n^3h^{2+\beta}}+\frac{1}{n^2a_n^3h^{2+\beta}}\right) = o\left(\frac{1}{na_nh^{\frac{3}{2}+\beta}}\right)$, since $h/a_n \to 0$. By Assumption 1 (iii) $\mathscr{F}\gamma \in \mathscr{W}^s, s > \frac{1}{2}$, therefore $\gamma \in L^1(\mathbb{R})$ and hence

$$\begin{aligned}
R_{n,2,2} &\le \frac{\mathscr{C}}{h^{3+\beta}(na_n)^2}\left(\|\gamma\|_1+\frac{1}{na_n^2}\right) = O\left(\frac{1}{n^2a_n^2h^{3+\beta}}+\frac{1}{n^3a_n^4h^{3+\beta}}\right) \\
&= O\left(\frac{1}{na_nh^{\frac{3}{2}+\beta}}\left(\frac{1}{na_nh^{\frac{3}{2}}}+\frac{1}{n^2a_n^3h^{\frac{3}{2}+\beta}}\right)\right) = o\left(\frac{1}{na_nh^{\frac{3}{2}+\beta}}\right),
\end{aligned}$$

where we used again that $h/a_n \to 0$. Hence, in total we find

$$\mathbb{E}[\hat{g}_n(x;h)] = \frac{1}{h}\int_{-\frac{1}{a_n}}^{\frac{1}{a_n}} \gamma(z) K\Big(\frac{z-x}{h};h\Big)\,dz + O\Big(\frac{1}{na_n h^{\frac{3}{2}+\beta}}\Big).$$

Next, we enlarge the domain of integration and estimate the remainder as follows. By the Cauchy-Schwarz inequality we obtain

$$\int_{|z|>\frac{1}{a_n}} \gamma(z) K\left(\frac{z-x}{h};h\right)dz \le \left(\int_{|z|>\frac{1}{a_n}} |\gamma(z)|^2\,dz\right)^{\frac{1}{2}} \times \left(\int_{|z|>\frac{1}{a_n}} \left|K\left(\frac{z-x}{h};h\right)\right|^2 dz\right)^{\frac{1}{2}}.$$

By Assumption 1 (iii)

$$\int_{|z|>\frac{1}{a_n}} |\gamma(z)|^2\,dz \le \int_{|z|>\frac{1}{a_n}} \frac{(1+z^2)^s}{(1+\frac{1}{a_n^2})^s}|\gamma(z)|^2\,dz \le \mathscr{C} a_n^{2s}.$$

By Lemma 4

$$\int_{\{|z|>\frac{1}{a_n}\}} \left(K\left(\frac{z-x}{h};h\right)\right)^2 dz \le \mathscr{C} a_n \begin{Bmatrix} h^{-2\beta}, & (W) \\ h^{-2\beta+2}, & (S) \end{Bmatrix} = O(a_n \|K(\cdot;h)\|_2^2).$$

Hence,

$$\mathbb{E}[\hat{g}_n(x;h)] = \frac{1}{h}\int \gamma(z) K\Big(\frac{z-x}{h};h\Big)\,dz + o\Big(\frac{1}{na_n h^{\frac{3}{2}+\beta}}\Big) + \begin{cases} O\left(a_n^{s+1/2} h^{1-\beta}\right), & \text{Ass. 3, (S)}, \\ O\left(a_n^{s+1/2} h^{-\beta}\right), & \text{Ass. 5, (W)}. \end{cases}$$

Furthermore, by Plancherel's equality and the convolution theorem,

$$\begin{aligned}
&\frac{1}{h}\int \gamma(z) K\Big(\frac{z-x}{h};h\Big)\,dz = \frac{1}{h}\int \gamma(z) \widetilde{K}\Big(\frac{z-x}{h};h\Big)\,dz \\
&= \frac{1}{2\pi h}\int \overline{\Phi_\gamma(z)}\Phi_{\widetilde{K}(\frac{\cdot-x}{h};h)}(z)\,dz = \frac{1}{2\pi}\int \exp(ixz)\overline{\Phi_\gamma(z)}\Phi_{\widetilde{K}(\cdot;h)}(zh)\,dz \\
&= \frac{1}{2\pi}\int \exp(ixz)\overline{\Phi_{f_\Delta}(z)\Phi_g(z)}\frac{\Phi_k(hz)}{\Phi_{f_\Delta}(-z)}\,dz = \frac{1}{2\pi}\int \exp(ixz)\Phi_g(-z)\Phi_k(hz)\,dz.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\frac{1}{h}\int \gamma(z) K\Big(\frac{z-x}{h};h\Big)\,dz \\
&= \frac{1}{2\pi}\int \exp(ixz)\Phi_g(-z)\,dz + \frac{1}{2\pi}\int \exp(ixz)\Phi_g(-z)\left(\Phi_k(hz)-1\right)dz \\
&= g(x) + \frac{1}{2\pi}\int \exp(ixz)\Phi_g(z)\left(\Phi_k(hz)-1\right)dz = g(x) + R(x;h),
\end{aligned}$$

where

$$R(x;h) = \frac{1}{2\pi}\int \exp(ixz)\Phi_g(z)\left(\Phi_k(hz)-1\right)dz.$$

Finally,

$$|R_n(x)| \leq \mathscr{C}\int_{|z|>D/h}|\Phi_g(z)|\,dz \leq \left(\int_{|z|>D/h}\left(\frac{1}{1+z^2}\right)^m dz\right)^{\frac{1}{2}} \times\left(\int_{|z|>D/h}|\langle z\rangle^{2m}|\Phi_g(z)|^2\,dz\right)^{\frac{1}{2}},$$

which yields the estimate $R(x;h) = O(h^{m-\frac{1}{2}})$, uniformly with respect to $x$
(ii) In the situation of both, (ii)a) and (ii)b), we have

$$\begin{aligned}\operatorname{Var}\left[\hat{g}_n(x;h)\right] &= \frac{1}{n^2a_n^2h^2}\sum_{j=-n}^{n}\nu^2(w_j)\left|K\Big(\frac{w_j-x}{h};h\Big)\right|^2 \\ &= \frac{1}{na_nh^2}\int_{-\frac{1}{a_n}}^{\frac{1}{a_n}+\frac{1}{na_n}}\nu^2(z)\left(K\Big(\frac{z-x}{h};h\Big)\right)^2 dz + R_n(x),\end{aligned}$$

where

$$R_n(x) = \frac{1}{na_nh^2}\sum_{j=-n}^{n}\int_{w_j}^{w_j+\frac{1}{na_n}}\left[\nu^2(w_j)\left(K\Big(\frac{w_j-x}{h};h\Big)\right)^2 - \nu^2(z)\left(K\Big(\frac{z-x}{h};h\Big)\right)^2\right]dz.$$

Then

$$\begin{aligned}R_n(x) &= \frac{1}{na_nh^2}\sum_{j=-n}^{n}\int_{w_j}^{w_j+\frac{1}{na_n}}\nu^2(w_j)\left[\left(K\Big(\frac{w_j-x}{h};h\Big)\right)^2 - \left(K\Big(\frac{z-x}{h};h\Big)\right)^2\right]dz \\ &+ \frac{1}{na_nh^2}\sum_{j=-n}^{n}\int_{w_j}^{w_j+\frac{1}{na_n}}\left(K\Big(\frac{z-x}{h};h\Big)\right)^2\left[\nu^2(w_j)-\nu^2(z)\right]dz \\ &=: R_{n,1}(x) + R_{n,2}(x).\end{aligned}$$

By uniform Lipschitz continuity of $\nu^2$ (see Lemma 3 (ii)), it is immediate that

$$\begin{aligned}|R_{n,2}(x)| &\leq \frac{\mathscr{C}}{n^2a_n^2h^2}\int\left(K\Big(\frac{z-x}{h};h\Big)\right)^2 \leq \frac{\mathscr{C}}{n^2a_n^2h}\|K(\cdot,h)\|^2 \\ &= O\left(\frac{1}{n^2a_n^2h^{1+2\beta}}\right).\end{aligned}$$

Next, we consider the term $R_{n,1}$ for which we will use a Taylor expansion of $K^2(\cdot;h)$. To this end, notice first from (7) that for any $l$

$$K^{(l)}(w;h) = \frac{(-1)^l}{2\pi}\int\frac{e^{-itw}\Phi_k(t)\cdot t^l}{\Phi_{f_\Delta(-t/h)}}\,dt,$$

where the functions $F_l : t \mapsto \Phi_k(t) \cdot t^l$ is uniformly bounded by 1 and twice continuously differentiable by Assumption 2 for any $l \in \mathbb{N}$. It follows that $K^2(\cdot;h)$ is smooth with integrable derivatives of all orders $l \in \mathbb{N}$, since

$$(K^2)^{(l)}(w;h) = (K \cdot K)^{(l)}(w;h) = \sum_{k=0}^{l} \binom{l}{k} K^{(l-k)}(w;h) K^{(k)}(w;h),$$

by the general Leibniz rule. This yields

$$\int \left| (K^2)^{(l)}(w;h) \right| dw \leq \sum_{k=0}^{l} \binom{l}{k} \left( \int |K^{(l-k)}(w;h)|^2 \, dw \right)^{\frac{1}{2}} \times \left( \int |K^{(k)}(w;h)|^2 \, dw \right)^{\frac{1}{2}} \leq \frac{\mathscr{C}}{h^{2\beta}}, \tag{S1}$$

by Lemma 4 and the previous discussion. Let $M \in \mathbb{N}$ be such that $M \geq \frac{2}{\beta} - 1$. It follows that

$$|R_{n,1}(x)| \leq \frac{\mathscr{C}}{na_n h^2} \left( \sum_{j=-n}^{n} \int_{w_j}^{w_j + \frac{1}{na_n}} \left[ \sum_{l=1}^{M} \left| (K^2)^{(l)} \left( \frac{z-x}{h}; h \right) \right| \left( \frac{1}{na_n h} \right)^l \right] dz + \left( \frac{1}{na_n h} \right)^{M+1} \cdot \frac{1}{a_n h^{2\beta}} \right).$$

By (S1), we deduce

$$\begin{aligned} |R_{n,1}(x)| &\leq \frac{\mathscr{C}}{na_n h^{1+2\beta}} \left( \frac{1}{na_n h} + \frac{1}{ha_n} \left( \frac{1}{na_n h} \right)^{M+1} \right) \\ &\leq \frac{\mathscr{C}}{na_n h^{1+2\beta}} \left( \frac{1}{na_n h} + \left( \frac{1}{na_n h^{1+\frac{2}{M+1}}} \right)^{M+1} \right). \end{aligned}$$

Since $M \geq \frac{2}{\beta} - 1$, we finally obtain

$$|R_{n,1}(x)| = o\left( \frac{1}{na_n h^{1+2\beta}} \right).$$

An application of Plancherel's theorem and Assumption 5 give

$$\frac{1}{\pi C h^{2\beta}} \leq \|K(\cdot;h)\|_2^2 = \frac{1}{2\pi} \left\| \frac{\Phi_k}{\Phi_{f_\Delta}(\cdot/h)} \right\|_2^2 \leq \frac{1}{\pi c} \left( 1 + \frac{1}{h^2} \right)^{\beta}. \tag{S2}$$

Now, if (S) holds, an application of Lemma 4 yields

$$\sup_{x \in [0,1]} \left| \int_{\frac{1}{a_n}}^{\infty} \left( K\left( \frac{z-x}{h}; h \right) \right)^2 dz \right| = O\left( \frac{a_n h^2}{h^{2\beta}} \right) = O\left( a_n h^2 \|K(\cdot;h)\|_2^2 \right).$$

Thus,

$$\int_{-\frac{1}{a_n}}^{\frac{1}{a_n}+\frac{1}{na_n}} \nu^2(z)\left(K\Big(\frac{z-x}{h};h\Big)\right)^2 dz = h\nu^2(x)\|K(\cdot;h)\|_2^2(1+O(ha_n))+R_{n,3}(x),$$

where

$$\begin{aligned}
|R_{n,3}(x)| &:= \left|\int_{-\frac{1}{a_n}}^{\frac{1}{a_n}+\frac{1}{na_n}} \big(\nu^2(z)-\nu^2(x)\big)\left(K\Big(\frac{z-x}{h};h\Big)\right)^2 dz\right| \\
&\leq \mathscr{C}h\int_{-\frac{2}{a_nh}}^{\frac{2}{a_nh}} \big|\nu^2(zh-x)-\nu^2(x)\big|\big(K(z;h)\big)^2 dz \\
&\leq \mathscr{C}h\int_{-\frac{2}{a_nh}}^{\frac{2}{a_nh}} \big|zh\big|\big(K(z;h)\big)^2 dz.
\end{aligned}$$

By (24) we have $|z\cdot K\big(z;h\big)| \leq \mathscr{C}/h^{\beta}$ and

$$|R_{n,3}(x)| \leq \mathscr{C}h^{2-\beta}\int_{-\frac{2}{a_nh}}^{\frac{2}{a_nh}} \big|K(z;h)\big|\,dz = O\left(\frac{h^2\ln(n)}{h^{2\beta}}\right),$$

since, by (24) and (25) in the proof of Lemma 4 and (S),

$$\begin{aligned}
&\int_{-\frac{2}{a_nh}}^{\frac{2}{a_nh}} \big|K\big(z;h\big)\big|\,dz \leq \frac{\mathscr{C}}{h^{\beta}} + \int_{1\leq|z|\leq\frac{2}{a_nh}} \big|K\big(z;h\big)\big|\,dz \\
&\leq \frac{\mathscr{C}}{h^{\beta}}\left(1+\int_{1\leq|z|\leq\frac{2}{a_nh}} \frac{1}{|z|}\,dz\right) \leq \frac{\mathscr{C}}{h^{\beta}}\left(1+\ln(2/a_nh)\right) = O(\ln(n)/h^{\beta}). \qquad \text{(S3)}
\end{aligned}$$

Assertion (ii)b) now follows.
(ii)a) Under (W)

$$\int_{-\frac{1}{a_n}}^{\frac{1}{a_n}+\frac{1}{na_n}} \nu^2(z)\left(K\Big(\frac{z-x}{h};h\Big)\right)^2 dz \leq \sup_{y\in\mathbb{R}} \nu^2(y)h\int |K(z;h)|^2\,dz,$$

and the second inequality of (ii)a) follows by (S2). Furthermore,

$$\begin{aligned}
\int_{-\frac{1}{a_n}}^{\frac{1}{a_n}+\frac{1}{na_n}} \nu^2(z)\left(K\Big(\frac{z-x}{h};h\Big)\right)^2 dz &\geq \sigma^2\int_{-\frac{1}{a_n}}^{\frac{1}{a_n}+\frac{1}{na_n}} \left(K\Big(\frac{z-x}{h};h\Big)\right)^2 dz \\
&\geq \frac{h\sigma^2}{2}\|K(\cdot;h)\|_2^2,
\end{aligned}$$

for sufficiently large $n$ by Lemma 4. Now, the first inequality of (ii)a) follows by (S2), which concludes the proof of this lemma. □

# II More general design

For ease of notation, we considered an equally spaced design in the main document. However, this can be relaxed to more general designs. In this section, we restate the main results (Theorem 1, Theorem 2, Lemma 5) and adjust their proofs to the case where the design is generated by a known positive design density $f_{D,n}$ on $[0,\infty)$ as follows

$$\frac{j}{n+1} = \int_0^{w_j} f_{D,n}(z)\,dz, \quad j = 1,\dots,n,$$

and $w_j = -w_{-j}$. Note that, given the latter definition, we have $f_{D,n}(z) = (n+1)/na_n I_{[0,1/a_n]}(z)$. Furthermore, we require the following regularity assumptions.

**Assumption 3.**

1. The density $f_{D,n}$ is continuously differentiable, $f_{D,n} \in C^1(\operatorname{supp}(f_{D,n}))$.
2. There exist constants $c_D$ and $C_D$ such that $c_D a_n \le f_{D,n} \le C_D a_n$.
3. The derivative $f'_{D,n}$ is uniformly bounded, $|f'_{D,n}| \le a_n C_{D'}$.

Regarding our estimator, we need to make the following adjustment to accommodate the more general design

$$\hat{g}_n(x;h) = \frac{1}{nh}\sum_{j=-n}^{n} \frac{Y_j}{f_{D,n}(w_j)} K\left(\frac{w_j - x}{h}; h\right).$$

This yields the following adjusted Lemma 5 and adjusted proof.

**Lemma 6.** Let Assumptions 1 and 2 be satisfied. Further assume that $h/a_n \to 0$ as $n \to \infty$.

(i) Then for bias, we have that

$$\sup_{x\in[0,1]} \left|\mathbb{E}\left[\hat{g}_n(x;h)\right] - g(x)\right| = O\left(h^{m-\frac{1}{2}} + \frac{1}{na_n h^{\beta+\frac{3}{2}}}\right) + \begin{cases} o\left(a_n^{s+1/2} h^{1-\beta}\right), & \text{Ass. 3, (S)},\\ o\left(a_n^{s+1/2} h^{-\beta}\right), & \text{Ass. 5, (W)}.\end{cases}$$

(ii) a) For the variance if Assumption 5, (W) holds and $na_n h^{1+\beta} \to \infty$, then we have that

$$\frac{1}{C_D}\cdot\frac{\sigma^2}{2C\pi}(1+O(a_n)) \le na_n h^{1+2\beta}\operatorname{Var}[\hat{g}_n(x;h)] \le \frac{2^\beta \sup_{x\in\mathbb{R}} \nu^2(x)}{c\pi}\cdot\frac{1}{c_D}.$$

(ii) b) If actually Assumption 3, (S) holds and $na_n h^{1+\beta} \to \infty$, then

$$\frac{1}{f_{D,n}(x)}\cdot\frac{\nu^2(x)}{C\pi}(1+O(a_n)) \le na_n h^{1+2\beta}\operatorname{Var}[\hat{g}_n(x;h)] \le \frac{\nu^2(x)}{c\pi}(1+O(h/a_n))\cdot\frac{1}{f_{D,n}(x)},$$

Here $c, C$ and $\beta$ are the constants from Assumption 5 respectively 3.

*Proof of Lemma 6.* (i) We have that

$$\mathbb{E}\left[\hat{g}_n(x;h)\right] = \frac{1}{nh}\sum_{j=-n}^{n}\frac{\gamma(w_j)}{f_{D,n}(w_j)}K\left(\frac{w_j-x}{h};h\right)$$
$$= \frac{1}{nh}\sum_{j=-n}^{n-1}\int_{w_j}^{w_{j+1}}\frac{\gamma(w_j)}{f_{D,n}(w_j)(w_{j+1}-w_j)}K\left(\frac{w_j-x}{h};h\right)\,dz.$$

Next, observe

$$f_{D,n}(w_j)(w_{j+1}-w_j) = F_{D,n}(w_{j+1})-F_{D,n}(w_j)-\frac{1}{2}f'_{D,n}(w_j^*)(w_{j+1}-w_j)^2,$$
$$= \frac{1}{n+1}+\frac{1}{2}f'_{D,n}(w_j^*)(w_{j+1}-w_j)^2,$$

where $F_{D,n}$ is the primitive of $f_{D,n}$. This yields

$$\left|f_{D,n}(w_j)(w_{j+1}-w_j)-\frac{1}{n}\right| \le \frac{1}{n^2}+\frac{1}{2}C_{D'}\frac{1}{n^2a_n^2c_D}$$

and thus

$$\frac{1}{f_{D,n}(w_j)(w_{j+1}-w_j)} = n\left(1+O\left(\tfrac{1}{na_n^2}\right)\right). \tag{S4}$$

Replacement of $1/f_{D,n}(w_j)(w_{j+1}-w_j)$ by the latter estimate now yields

$$\mathbb{E}\left[\hat{g}_n(x;h)\right] = \frac{1}{h}\int_{-\frac{1}{a_n}}^{\frac{1}{a_n}}\gamma(z)K\left(\frac{z-x}{h};h\right)\,dz\left(1+O\left(\tfrac{1}{na_n^2}\right)\right)+R_{n,1}(x)+R_{n,2}(x),$$

where

$$R_{n,1}(x) = \frac{1}{h}\sum_{j=-n}^{n-1}\int_{w_j}^{w_{j+1}}\frac{\mathrm{d}}{\mathrm{d}u}\left(\gamma(u)K\left(\frac{u-x}{h};h\right)\right)\Big|_{u=z}(w_j-z)\,dz$$

and

$$R_{n,2}(x) = \frac{1}{2h}\sum_{j=-n}^{n-1}\int_{w_j}^{w_{j+1}}\frac{\mathrm{d}^2}{\mathrm{d}u^2}\left(\gamma(u)K\left(\frac{u-x}{h};h\right)\right)\Big|_{u=\tilde{w}_j(z)}(w_j-z)^2\,dz.$$

For $z\in[w_j,w_{j+1}]$ we obtain the following estimates by Assumption 3:

$$|w_j-z| \le |w_j-w_{j+1}| \le 1/(na_nc_D)$$

. Therefore, the rest of the proof of claim (i) follows along the lines of the proof of Lemma 5 (i).
(ii) In the situation of both, (ii)a) and (ii)b), we have

$$\mathrm{Var}\left[\hat{g}_n(x;h)\right] = \frac{1}{n^2h^2}\sum_{j=-n}^{n}\frac{\nu^2(w_j)}{f_{D,n}(w_j)^2}\left|K\left(\frac{w_j-x}{h};h\right)\right|^2$$
$$= \frac{1}{nh^2}\int_{-\frac{1}{a_n}}^{\frac{1}{a_n}}\frac{\nu^2(z)}{f_{D,n}(z)}\left(K\left(\frac{z-x}{h};h\right)\right)^2dz\cdot\left(1+O\left(\frac{1}{na_n^2}\right)\right)+R_n(x),$$

where

$$R_n(x) = \frac{1}{nh^2} \sum_{j=-n}^{n-1} \int_{w_j}^{w_{j+1}} \left[ \frac{\nu^2(w_j)}{n f_{D,n}(w_j)^2 (w_{j+1}-w_j)} \left( K\Big(\frac{w_j - x}{h}; h\Big) \right)^2 - \frac{\nu^2(z)}{f_{D,n}(z)} \left( K\Big(\frac{z-x}{h}; h\Big) \right)^2 \right] dz.$$

Then

$$\begin{aligned} R_n(x) &= \frac{1}{nh^2} \sum_{j=-n}^{n-1} \int_{w_j}^{w_{j+1}} \frac{\nu^2(w_j)}{n f_{D,n}(w_j)^2 (w_{j+1}-w_j)} \left[ \left( K\Big(\frac{w_j - x}{h}; h\Big) \right)^2 - \left( K\Big(\frac{z-x}{h}; h\Big) \right)^2 \right] dz \\ &\quad + \frac{1}{nh^2} \sum_{j=-n}^{n} \int_{w_j}^{w_{j+1}} \left( K\Big(\frac{z-x}{h}; h\Big) \right)^2 \left[ \frac{\nu^2(w_j)}{n f_{D,n}(w_j)^2 (w_{j+1}-w_j)} - \frac{\nu^2(z)}{f_{D,n}(z)} \right] dz. \end{aligned}$$

Using (S4), we further obtain

$$\begin{aligned} R_n(x) &= \frac{1}{nh^2} \sum_{j=-n}^{n-1} \int_{w_j}^{w_{j+1}} \frac{\nu^2(w_j)}{f_{D,n}(w_j)} \left[ \left( K\Big(\frac{w_j - x}{h}; h\Big) \right)^2 - \left( K\Big(\frac{z-x}{h}; h\Big) \right)^2 \right] dz \cdot \left(1 + O\Big(\tfrac{1}{na_n}\Big)\right) \\ &\quad + \frac{1}{nh^2} \sum_{j=-n}^{n-1} \int_{w_j}^{w_{j+1}} \left( K\Big(\frac{z-x}{h}; h\Big) \right)^2 \left[ \frac{\nu^2(w_j)}{f_{D,n}(w_j)} - \frac{\nu^2(z)}{f_{D,n}(z)} \right] dz \\ &\quad + \frac{\mathscr{C}}{n^2 a_n h^2} \sum_{j=-n}^{n-1} \int_{w_j}^{w_{j+1}} \left( K\Big(\frac{z-x}{h}; h\Big) \right)^2 \frac{\nu^2(w_j)}{f_{D,n}(w_j)} \, dz =: R_{n,1}(x) + R_{n,2}(x) + R_{n,3}(x). \end{aligned}$$

It holds that

$$|R_{n,3}(x)| \le \frac{\mathscr{C}}{n^2 a_n^2 h} \|K(\cdot, h)\|_2^2 = O\left( \frac{1}{n^2 a_n h^{2\beta+1}} \right).$$

Furthermore,

$$\begin{aligned} R_{n,2}(x) &= \frac{1}{nh^2} \sum_{j=-n}^{n-1} \int_{w_j}^{w_{j+1}} \left( K\Big(\frac{z-x}{h}; h\Big) \right)^2 \nu^2(z) \frac{f_{D,n}(z) - f_{D,n}(w_j)}{f_{D,n}(z) f_{D,n}(w_j)} \, dz \\ &\quad + \frac{1}{nh^2} \sum_{j=-n}^{n-1} \int_{w_j}^{w_{j+1}} \left( K\Big(\frac{z-x}{h}; h\Big) \right)^2 \left[ \frac{\nu^2(w_j) - \nu^2(z)}{f_{D,n}(w_j)} \right] dz \end{aligned}$$

By uniform Lipschitz continuity of $\nu^2$ (see Lemma 3 (ii)) and $f_{D,n}$ (by Assumption 3), it is immediate that

$$\begin{aligned} |R_{n,2}(x)| &\le \frac{\mathscr{C}}{n^2 a_n^2 h^2} \int \left( K\Big(\frac{z-x}{h}; h\Big) \right)^2 \le \frac{\mathscr{C}}{n^2 a_n^2 h} \|K(\cdot, h)\|^2 \\ &= O\left( \frac{1}{n^2 a_n^2 h^{1+2\beta}} \right). \end{aligned}$$

Using again that $|w_j - z| \le |w_j - w_{j+1}| \le 1/(n a_n c_D)$, the rest of the proof of claim (ii) follows along the lines of the proof of Lemma 5 (ii). □

# III Role of the hyperparameters

In this section, we discuss the setting presented in Example 1 of the main document in more detail, in order to shed some light on the role of the parameters $a_n$ and $h$, as well as on the assumptions made for our theoretical considerations. In particular, we show that, in a typical setting, the conditions listed in Assumption 4 are satisfied if the bandwidth $h$ is chosen to be of rate $h \sim h^*/\log(n)$, where $h^* \sim \left(\frac{\ln(n)}{na_n}\right)^{\frac{1}{2(\beta+m)}}$ balances bias and standard deviation of $\hat{g}_n(\cdot;h)$. As a specific example, we consider the case of a function $g \in \mathscr{W}^m(\mathbb{R}), m > 5/2$, of bounded support, $[-1,1]$, say, $f_\Delta$ as in Definition (8) in the main document with $a = 1$, i.e.,

$$f_\Delta(x) = \tfrac{1}{2}e^{-|x|} \quad \text{with} \quad \Phi_{f_\Delta}(t) = \langle t \rangle^{-2},$$

and $\mathbb{E}[\varepsilon_1^4] < \infty$, i.e., $M = 4$. Here, the parameter $\beta$, which gives the degree of ill-posedness of the problem and which is defined in Assumption 3 in the main document is given by $\beta = 2$. The design parameter $a_n$ ensures that asymptotically, observations on the whole real line are available. This is necessary since the function $\gamma$ will typically be of unbounded support, even if the function $g$ itself is of bounded support as it is the case in this example. To give some more intuition, we now provide some computations for our specific example, for which it will turn out that Assumption 1 (iii) is met for any $s = O(1/(\ln\ln(n)a_n))$. We find

$$\begin{aligned}
\int_{|z|>1/a_n} \langle z \rangle^{2s} (\gamma(z))^2 \, dz &= \int_{|z|>1/a_n} \langle z \rangle^{2s} \left( \int g(t) f_\Delta(t-z) \, dt \right)^2 dz \\
&= \frac{1}{4} \int_{|z|>1/a_n} \langle z \rangle^{2s} \left( \int g(t) \exp(-|t-z|) \, dt \right)^2 dz \\
&= \frac{1}{4} \int_{|z|>1/a_n} \langle z \rangle^{2s} \left( \int_{|t|\le 1} g(t) \exp(-|t-z|) \, dt \right)^2 dz.
\end{aligned}$$

If $z > 1/a_n$ in the outer integral, we have $t < z$ in the inner integral, implying $\exp(-|t-z|) = \exp(-z+t)$ and thus

$$\int_{|z|>1/a_n} \exp(-2z) \langle z \rangle^{2s} \left( \int_{|t|\le 1} g(t) \exp(t) \, dt \right)^2 dz = \mathscr{C} \int_{|z|>1/a_n} \exp(-2z) \langle z \rangle^{2s} \, dz.$$

Integration by parts yields

$$\begin{aligned}
\int_{|z|>1/a_n} \exp(-2z) \langle z \rangle^{2s} \, dz &= \frac{1}{2} \left( 1 + \frac{1}{a_n^2} \right)^s \exp\left( -\frac{2}{a_n} \right) + \frac{s}{2} \int_{|z|>1/a_n} \exp(-2z) z (1+z^2)^{s-1} \, dz \\
&\le \frac{1}{2} \left( 1 + \frac{1}{a_n^2} \right)^s \exp\left( -\frac{2}{a_n} \right) + \frac{s a_n}{2} \int_{|z|>1/a_n} \exp(-2z)(1+z^2)^s \, dz.
\end{aligned}$$

The latter estimate implies

$$\int_{|z|>1/a_n} \exp(-2z) \langle z \rangle^{2s} \, dz \le \frac{1}{2} \frac{(1+\frac{1}{a_n^2})^s}{1 - s a_n/2} \exp\left( -\tfrac{2}{a_n} \right).$$

If $s = O(1/(\ln\ln(n)a_n))$, we find that

$$\int_{|z|>1/a_n} \exp(-2z)\langle z\rangle^{2s}\,dz \le \mathscr{C}\left(1+\frac{1}{a_n^2}\right)^s \exp\left(-\frac{2}{a_n}\right). \tag{1}$$

for sufficiently large $n$. Taking the logarithm on the right hand side of (1) gives

$$\frac{1}{a_n}\left[a_n\ln(\mathscr{C}) + sa_n\ln(1+1/a_n^2) - 2\right],$$

which tends to $-\infty$ if $s = O(1/(\ln\ln(n)a_n))$, implying that $\left(1+\frac{1}{a_n^2}\right)^s \exp\left(-\frac{2}{a_n}\right) \to 0$. Analogously,

$$\int_{z<-1/a_n} \exp(2z)\langle z\rangle^{2s}\left(\int_{|t|\le 1} g(t)\exp(-t)\,dt\right)^2 dz \to 0,$$

if $s = O(1/(\ln\ln(n)a_n))$. If we choose an undersmoothing bandwidth of order $h \sim h^*/\ln(n)$ and $a_n \sim s_n \sim 1/\ln(n)$, Assumption 4 (ii) becomes

$$\ln(n)^{1-2m-2\beta} + o\left(\ln(n)^{1-2m-2\beta}\right) = o(1),$$

which is satisfied since $m > 5/2$ and $\beta > 0$.

# IV Extensions: Details

Our theoretical developments for the procedure in Section 5 in the main document actually involve a sample splitting. To this end, let $(d_n)_{n\in\mathbb{N}}$ be a sequence of natural numbers, $d_n \to \infty$ and $d_n = o(n)$, $1/d_n = o(1/\ln(n)^2)$ and let $\mathscr{J}_n := \{-n,\dots,n\}\setminus\{-n+k\cdot d_n \,|\, 1 \le k \le 2n/d_n\}$, i.e., we remove each $d_n$-th data point from our original sample. Now, set $\mathscr{Y}_1 := \{Y_j \,|\, j \in \mathscr{J}_n\}$ as well as $\mathscr{Y}_2 := \{Y_j \,|\, j \in \{-n,\dots,n\}\setminus\mathscr{J}_n\}$. This way, the asymptotic properties of the estimator based on the main part of the sample, $\mathscr{Y}_1$, remain the same. Let further, for $j \in \mathscr{J}_n$, $\triangle_j$ denote the difference of $\omega_j$ and its left neighbor, that is, $\triangle_j = 1/(na_n)$ if $j-1 \in \mathscr{J}_n$ and $\triangle_j = 2/(na_n)$ else. Define the estimator $\tilde{g}_n$, based on $\mathscr{Y}_1$ by

$$\tilde{g}_n(x;h) = \frac{1}{h}\sum_{j\in\mathscr{J}_n} Y_j \triangle_j K\left(\frac{w_j - x}{h}; h\right).$$

We now formulate an analog of Theorem 1 under Assumption 5.

**Theorem 1.** Let Assumptions 2, 4 (i) and 5 be satisfied. Let further $\tilde{\nu}_n$ be a nonparametric estimator of the standard deviation in model (4) based on $\mathscr{Y}_2$ such that for some sequence of positive numbers $b_n \to 0$ for which $a_n/b_n = o(1/\ln(n))$ we have that

$$\mathbb{E}\left[\sup_{|j|\le b_n n} |\tilde{\nu}_n(\omega_j) - \nu(\omega_j)|\right] = o\big(1/(\ln(n)\ln\ln(n))\big) \quad \text{and} \quad \tilde{\nu}_n > \tilde{\sigma} > 0 \tag{S7}$$

for some constant $\tilde{\sigma} > 0$.

1. There exists a sequence of independent standard normally distributed random variables $(Z_n)_{n\in\mathbb{Z}}$, independent of $\tilde{\nu}_n$, such that for

$$\widetilde{\mathbb{D}}_n(x) = \frac{\sqrt{na_nh^{1+2\beta}}}{\tilde{\nu}(x)}\left(\tilde{g}_n(x;h) - \mathbb{E}[\tilde{g}_n(x;h)]\right),$$
$$\widetilde{\mathbb{G}}_n(x) = \frac{\sqrt{na_nh^{1+2\beta}}}{h\,\tilde{\nu}_n(x)}\sum_{j\in\mathscr{J}_n,|j|\le nb_n}\tilde{\nu}_n(\omega_j)\,\triangle_j\,Z_jK\left(\frac{w_j-x}{h};h\right), \tag{S8}$$

we have that

$$\forall\,\alpha\in(0,1): \lim_{n\to\infty}\mathbb{P}\left(\|\widetilde{\mathbb{D}}_n\|\le q_{\|\widetilde{\mathbb{G}}_n\|}(\alpha)\right) = \alpha,$$

where $q_{\|\widetilde{\mathbb{G}}_n\|}(\alpha)$ denotes the $\alpha$-quantile of $\|\widetilde{\mathbb{G}}_n\|$.

2. If, in addition, $\sqrt{na_nh^{2m+2\beta}} + \sqrt{na_n^{2s+1}} + 1/\sqrt{na_nh^2} = o(1/\sqrt{\ln(n)})$. and if Assumption 1 is satisfied, $\mathbb{E}[\tilde{g}_n(x;h)]$ in (11) can be replaced by $g(x)$.

*Proof of Theorem 1.* We require that

$$\|\widetilde{\mathbb{D}}_n - \widetilde{\mathbb{G}}_n\| = o_{\mathbb{P}}(1/\sqrt{\ln(n)}), \tag{Step 1}$$

as well as

$$\mathbb{E}\left[\|\widetilde{\mathbb{G}}_n\|\right] = O_{\mathbb{P}}(\sqrt{\ln(n)}). \tag{Step 2}$$

**Step 1 a: Gaussian Approximation**

Lemmas 7 and 8 are in preparation for the Gaussian approximation where the target process $\widetilde{\mathbb{G}}_n$ is first approximated by processes $\widetilde{\mathbb{G}}_{n,0}^{b_n}$ and $\widetilde{\mathbb{G}}_{n,0}$.

**Lemma 7.** We shall show that

$$\|\nu\,\widetilde{\mathbb{G}}_{n,0}^{b_n} - \tilde{\nu}_n\,\widetilde{\mathbb{G}}_n\| = o_{\mathbb{P}}(1/\sqrt{\ln(n)}), \tag{2}$$

where

$$\widetilde{\mathbb{G}}_{n,0}^{b_n} := \frac{\sqrt{na_nh^{1+2\beta}}}{h\nu(x)}\sum_{j\in\mathscr{J}_n,|j|\le nb_n}\nu(\omega_j)\,\triangle_j\,Z_jK\left(\frac{w_j-x}{h};h\right).$$

*Proof.* Let $\mathscr{Z} := \{Z_{-n},\ldots,Z_n\}$ be iid standard Gaussian random variables as in the proof of Theorem 1, and $c_n := 1/\ln(n)\ln\ln(n)$.

$$\mathbb{P}_{\mathscr{Z},\mathscr{Y}_2}\left(\|\widetilde{\mathbb{G}}_{n,0}^{b_n} - \widetilde{\mathbb{G}}_n\| > \delta/\sqrt{\ln(n)}\right) \le \mathbb{P}_{\mathscr{Z},\mathscr{Y}_2}\left(\sup_{|j|\le b_nn}|\nu(w_j)-\tilde{\nu}_n(w_j)| > c_n\right)$$
$$+\mathbb{P}_{\mathscr{Z},\mathscr{Y}_2}\left(\|\widetilde{\mathbb{G}}_{n,0}^{b_n} - \widetilde{\mathbb{G}}_n\| > \delta/\sqrt{\ln(n)};\ \sup_{|j|\le b_nn}|\nu(w_j)-\tilde{\nu}_n(w_j)|\le c_n\right) = P_{n,1}+P_{n,2}.$$

By assumption, $P_{n,1} = o(1)$.

$$
\begin{aligned}
P_{n,2} &\le \mathbb{P}_{\mathscr{X},\mathscr{Y}_2}\left( \|\widetilde{\mathbb{G}}_{n,0}^{b_n} - \widetilde{\mathbb{G}}_n\| I_{\{\sup_{|j|\le b_n n} |\nu(w_j)-\tilde{\nu}_n(w_j)|\le c_n\}} > \delta/\sqrt{\ln(n)} \right) \\
&= \mathbb{E}_{\mathscr{Y}_2}\left[ \mathbb{P}_{\mathscr{X},\mathscr{Y}_2}\left( \|\widetilde{\mathbb{G}}_{n,0}^{b_n} - \widetilde{\mathbb{G}}_n\| I_{\{\sup_{|j|\le b_n n} |\nu(w_j)-\tilde{\nu}_n(w_j)|\le c_n\}} > \delta/\sqrt{\ln(n)} \,\middle|\, \mathscr{Y}_2 \right) \right] \\
&\le \mathbb{E}_{\mathscr{Y}_2}\left[ \mathbb{E}_{\mathscr{X},\mathscr{Y}_2}\left( \|\widetilde{\mathbb{G}}_{n,0}^{b_n} - \widetilde{\mathbb{G}}_n\| I_{\{\sup_{|j|\le b_n n} |\nu(w_j)-\tilde{\nu}_n(w_j)|\le c_n\}} \,\middle|\, \mathscr{Y}_2 \right) \right] \cdot \frac{\sqrt{\ln(n)}}{\delta},
\end{aligned}
$$

by the conditional Markov inequality. Set

$$
R_n(x) := \big(\widetilde{\mathbb{G}}_{n,0}^{b_n}(x) - \widetilde{\mathbb{G}}_n(x)\big) I_{\{\sup_{|j|\le b_n n} |\nu(w_j)-\tilde{\nu}_n(w_j)|\le c_n\}}
$$

and

$$
\widetilde{R}_n(x) := \frac{2h^{\beta} c_n}{\sqrt{n a_n h}} \sum_{|j|\le b_n n} Z_j K\left( \frac{w_j - x}{h}; h \right).
$$

Since $d_{R_n}(s,t) \le d_{\widetilde{R}_n}(s,t)$ for all $s,t \in [0,1]$ and for all samples $\mathscr{Y}_2$ and $\mathbb{E}\big[\|\widetilde{R}_n\| \,|\, \mathscr{Y}_2\big] = \mathbb{E}\|\widetilde{R}_n\| = O(c_n\sqrt{\ln(n)}) = o(1/\sqrt{\ln(n)})$, it follows by Lemma 7 that $\mathbb{E}_{\mathscr{X},\mathscr{Y}_2}[\|R_n\| \,|\, \mathscr{Y}_2] \le \mathbb{E}[\|\widetilde{R}_n\|]$ for all samples. Therefore, $P_{n,2} = o(1/\sqrt{\ln(n)})$ and the claim follows. □

**Lemma 8.** We have that

$$
\|\nu\widetilde{\mathbb{G}}_{n,0}^{b_n} - \nu\widetilde{\mathbb{G}}_{n,0}\| = o_{\mathbb{P}}(1/\sqrt{\ln(n)}).
$$

where

$$
\widetilde{\mathbb{G}}_{n,0}(x) := \frac{h^{\beta}}{\nu(x)\sqrt{n a_n h}} \sum_{\substack{j\in\mathscr{J}_n \\ nb_n<|j|\le n}} \nu(w_j)\Delta_j Z_j K\left( \frac{w_j - x}{h}; h \right). \tag{S9}
$$

*Proof.* Let

$$
R_n(x) := \nu\widetilde{\mathbb{G}}_{n,0}^{b_n} - \nu\widetilde{\mathbb{G}}_{n,0} = \frac{\sqrt{n a_n h^{1+2\beta}}}{h} \sum_{\substack{j\in\mathscr{J}_n \\ nb_n<|j|\le n}} \nu(w_j)\Delta_j Z_j K\left( \frac{w_j - x}{h}; h \right).
$$

Then, since $\Delta_j^2 \le 4/(n^2 a_n^2)$ and $\mathscr{J}_n \subset \{-n, \ldots, n\}$,

$$
\operatorname{Var}[R_n(x)] \le \frac{4h^{2\beta}}{n a_n h} \sum_{nb_n<|j|\le n} \nu^2(w_j) K^2\left( \frac{w_j - x}{h}; h \right).
$$

From the proof of Lemma 5 we deduce

$$
\operatorname{Var}[R_n(x)] \le \mathscr{C}\left( \frac{h a_n}{b_n} + \frac{1}{n a_n h} \right) = o\left( \frac{1}{\ln(n)^2} \right).
$$

An application of the following Lemma 9 concludes this proof.

**Lemma 9.** Let $(\mathbb{X}_n(t), t \in T)$ be an almost surely bounded Gaussian process on a compact index set $T$ with $\sigma_{T,n}^2 := \|\mathrm{Var}(\mathbb{X}_n)\| = o(1/\ln(n)^2)$ such that $N(T, \delta, d_{\mathbb{X}_n}) \leq (n/\delta)^a$ for all $\delta \leq \sigma_{T,n}$ and some $a \in (0, \infty)$. Then

$$\|\mathbb{X}_n\| = o_{\mathbb{P}}(1/\sqrt{\ln(n)}).$$

**Proof of Lemma 9.** Fix $\delta > 0$. An application of Theorem 4.1.2 in Adler and Taylor (2007) yields, for large enough $n \in \mathbb{N}$ and a universal constant $K$

$$\mathbb{P}\left(\|\mathbb{X}_n\|_T > \frac{\delta}{\sqrt{\ln(n)}}\right) \leq 2\left(\frac{Kn\delta}{2\sqrt{\ln(n)}\sigma_{T,n}^2}\right)^a \cdot \frac{2\sqrt{\ln(n)}\sigma_{T,n}}{\sqrt{2\pi}\delta} \cdot \exp\left(-\frac{\delta^2}{8\ln(n)\sigma_{T,n}^2}\right).$$

Now

$$\mathbb{P}\left(\|\mathbb{X}_n\|_T > \frac{\delta}{\sqrt{\ln(n)}}\right) \leq \mathscr{C} n^{a+1}\delta^{a-1}\sigma_{n,T}^{-2a+1}\exp\left(-\sigma_{T,n}^{-1}\right)\left(\exp\left(-\sigma_{T,n}^{-1}\right)\right)^{\frac{\delta^2}{o(1)}-1}.$$

Since $\sigma_{n,T}^{-2a+1}\exp\left(-\sigma_{T,n}^{-1}\right) \to 0$ as $\sigma_{n,T} \to 0$ and $\left(\exp\left(-\sigma_{T,n}^{-1}\right)\right)^{\frac{\delta^2}{o(1)}-1} = o(n^{-b})$ for any fixed $b \in (0, \infty)$, the claim of the lemma now follows. □

**Lemma 10.** We have that

$$\|\widetilde{\mathbb{G}}_n - \widetilde{\mathbb{D}}_n\| = o_{\mathbb{P}}(1/\sqrt{\ln(n)}).$$

*Proof of Lemma 10.* Since by assumption (S7), $\tilde{v}_n$ is bounded away from zero, it suffices to show that

$$\|\tilde{v}_n\widetilde{\mathbb{G}}_n - \tilde{v}_n\widetilde{\mathbb{D}}_n\| = o_{\mathbb{P}}(1/\sqrt{\ln(n)}).$$

We estimate

$$\begin{aligned}
\|\tilde{v}_n\widetilde{\mathbb{G}}_n - \tilde{v}_n\widetilde{\mathbb{D}}_n\| &\leq \|\tilde{v}_n\widetilde{\mathbb{G}}_n - v\widetilde{\mathbb{G}}_{n,0}^{b_n}\| + \|v\widetilde{\mathbb{G}}_{n,0}^{b_n} - v\widetilde{\mathbb{G}}_{n,0}\| + \|v\widetilde{\mathbb{G}}_{n,0} - \tilde{v}_n\widetilde{\mathbb{D}}_n\| \\
&= o_{\mathbb{P}}\left(1/\sqrt{\ln(n)}\right) + \|v\widetilde{\mathbb{G}}_{n,0} - \tilde{v}_n\widetilde{\mathbb{D}}_n\|
\end{aligned}$$

The claim now follows along the lines of the Gaussian approximation in the proof of Theorem 1. □

### Step 2: Expectation of the maximum

**Lemma 11.**

$$\mathbb{E}\|\widetilde{\mathbb{G}}_n\| = O(\sqrt{\ln(n)}).$$

*Proof.* Write $\mathbb{E}[\|\widetilde{\mathbb{G}}_n\|] = \mathbb{E}\big[\mathbb{E}[\|\widetilde{\mathbb{G}}_n\|] \,\big|\, \mathscr{Y}_2\big]$ and define

$$\mathbb{X}_n(t) := \frac{2M_n h^{\beta-1/2}}{\sqrt{na_n}\tilde{\sigma}} \sum_{j=-n}^{n} Z_j \nu(w_j) K\left(\frac{w_j - t}{h}; h\right),$$

where $M_n := \sqrt{\max_{|j|\le b_n n} |\tilde{\nu}^2(w_j) - \nu^2(w_j)| \frac{1}{\sigma^2} + 1}$. Conditionally on $\mathscr{Y}_2$, $(\widetilde{\mathbb{G}}_n(t), t \in [0,1])$ is a Gaussian process and we find for $s,t \in [0,1]$ and for all possible samples $\mathscr{Y}_2$ the following set of inequalities hold

$$\begin{aligned}
&\mathbb{E}\big[\big|\widetilde{\mathbb{G}}_n(s) - \widetilde{\mathbb{G}}_n(t)\big|^2 \,\big|\, \mathscr{Y}_2\big] \le \frac{na_n h^{2\beta-1}}{\tilde{\sigma}^2} \sum_{|j|\le nb_n} \tilde{\nu}^2(w_j)\Delta_j^2 \left|K\left(\frac{w_j - s}{h}; h\right) - K\left(\frac{w_j - t}{h}; h\right)\right|^2 \\
&\le \frac{4h^{2\beta-1}}{na_n\tilde{\sigma}^2} \sum_{|j|\le nb_n} \left[|\tilde{\nu}^2(w_j) - \nu^2(w_j)| \frac{\nu^2(w_j)}{\sigma^2} + \nu^2(w_j)\right] \left|K\left(\frac{w_j - s}{h}; h\right) - K\left(\frac{w_j - t}{h}; h\right)\right|^2 \\
&\le \mathbb{E}\big[\big|\mathbb{X}_n(s) - \mathbb{X}_n(t)\big|^2 \,\big|\, \mathscr{Y}_2\big].
\end{aligned}$$

An application of Lemma 7 yields, for all samples, $\mathbb{E}\big[\|\widetilde{\mathbb{G}}_n\| \,\big|\, \mathscr{Y}_2\big] \le \mathbb{E}\big[\|X_n\| \,\big|\, \mathscr{Y}_2\big]$. Therefore, $\mathbb{E}[\|\widetilde{\mathbb{G}}_n\|] \le \mathbb{E}\Big[\mathbb{E}[\|X_n\|] \,\big|\, \mathscr{Y}_2\big]\Big] \le \mathscr{C}\mathbb{E}\|\mathbb{G}_n\| \cdot \mathbb{E}M_n$.

An application of Jensen's inequality and (S7) yield $\mathbb{E}M_n \le 2$ for sufficiently large $n \in \mathbb{N}$. □

**Step 3: Anti-Concentration**

Following the arguments given in the proof of Theorem 1 concludes the proof of this theorem. □

# References


Adler, R. J. and Taylor, J. E. (2007). *Random Fields and Geometry*. Springer, New York.

Carroll, R. J., Ruppert, D., Stefanski, L. A., and Crainiceanu, C. M. (2006). *Measurement error in nonlinear models*, volume 105 of *Monographs on Statistics and Applied Probability*. Chapman & Hall/CRC, Boca Raton, FL, second edition. A modern perspective.