T E C H N I C A L R E P O R T

0626

UNIFORM IN BANDWITH EXACT RATES FOR A CLASS OF KERNEL ESTIMATORS

VARRON, D. and I. VAN KEILEGOM

I A P S T A T I S T I C S N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Uniform in bandwidth exact rates for a class of kernel estimators

Davit Varron[1] and Ingrid Van Keilegom[2]

Institute of Statistics, Catholic University of Louvain[3].

1

Abstract

Given an i.i.d sample (Y_i, Z_i), taking values in $\mathbb{R}^{d'} \times \mathbb{R}^d$, consider the quantities

$$W_n(g,h,z) := f_Z(z)^{-1/2} \sum_{i=1}^{n} \Big[\Big(< c_g(z), g(Y_i) > + d_g(z) \Big) K\Big(\frac{Z_i - z}{h}\Big) - \mathbb{E}\Big(\Big(< c_g(z), g(Y_i) > + d_g(z) \Big) K\Big(\frac{Z_i - z}{h}\Big) \Big) \Big], \tag{0.1}$$

where, z belongs to a compact set $H \subset \mathbb{R}^d$, f_Z is the density of Z_1, K is a kernel, $h > 0$ a bandwidth, g a Borel function on $\mathbb{R}^{d'}$ and $c_g(\cdot), d_g(\cdot)$ are continuous functions on $\mathbb{R}^d$. Given two bandwidth sequences $h_n < \mathfrak{h}_n$ fulfilling mild conditions, we prove that, for an explicit constant $\mathfrak{C}(\mathcal{G}, K)$, we have almost surely:

$$\lim_{n\to\infty} \sup_{z\in H,\, g\in\mathcal{G},\, h_n \le h \le \mathfrak{h}_n} \frac{|\, W_n(g,h,z) \,|}{\sqrt{2nh^d \log(h^{-d})}} = \mathfrak{C}(\mathcal{G}, K), \tag{0.2}$$

under mild conditions on the density f_Z, the class $\mathcal{G}$, the kernel K and the functions $c_g(\cdot), d_g(\cdot)$. We apply this result to the context of empirical likelihood, where regression parameters are estimated with a smoothed version of empirical likelihood, involving a kernel K and a bandwidth h. Namely, we prove that smoothed empirical likelihood can be used to build confidence intervals for conditional probabilities $\mathbb{P}\Big(Y \in C \mid Z = z\Big)$, that hold uniformly in $z \in H$, $C \in \mathcal{C}$, $h \in [h_n, \mathfrak{h}_n]$. Here $\mathcal{C}$ is a Vapnik-Chervonenkis class of sets.

Keywords: Local empirical processes, empirical likelihood, kernel smoothing, uniform in bandwidth consistency.

1 Introduction and statement of the main results

Consider an i.i.d sample $(Y_i, Z_i)_{i=1,\ldots,n}$ taking values in $\mathbb{R}^{d'} \times \mathbb{R}^d$, with the same distribution as a vector (Y, Z), and write $< \cdot, \cdot >$ for the usual inner product. In this paper, we investigate the limit behaviour of quantities of the following form (assuming that this expression is meaningfull):

$$W_n(g,h,z) := f_Z(z)^{-1/2} \sum_{i=1}^{n} \Big[\Big(< c_g(z), g(Y_i) > + d_g(z) \Big) K\Big(\frac{Z_i - z}{h}\Big) - \mathbb{E}\Big(\Big(< c_g(z), g(Y_i) > + d_g(z) \Big) K\Big(\frac{Z_i - z}{h}\Big) \Big) \Big]. \tag{1.1}$$

[1] *varron@stat.ucl.ac.be*, [2] *vankeilegom@stat.ucl.ac.be*.[3] Voie du Roman Pays 20, 1348 Louvain-la-Neuve, Belgium.

Here, K denotes a kernel, $h > 0$ is a smoothing parameter, g is a Borel function from $\mathbb{R}^{d'}$ to $\mathbb{R}^k$ and f_Z is (a version) of the density of Z. Given a class of functions $\mathcal{G}$ satisfying some Vapnik-Chervonenkis type conditions (see conditions (HG1) below), and given a compact set H, Einmahl and Mason [5] showed that somewhat recent tools in empirical processes theory could be used efficiently to provide exact rates of convergence of

$$\sup\Big\{ \mid W_n(g, h_n, z) \mid,\ g \in \mathcal{G},\ z \in H\Big\},$$

along a bandwidth sequence h_n fulfilling some mild conditions (see condition (HV) in the sequel). The exact content of their result is written in Theorem 1 below. The contribution of the present paper is twofold. As a first contribution, we provide an extension of the result of Einmahl and Mason, by enriching Theorem 1 with a uniformity in the bandwidth h, when h is allowed to vary into an interval $[h_n, \mathfrak{h}_n]$, with h_n and $\mathfrak{h}_n$ fulfilling conditions of Theorem 1. This extension is stated in Section 1.2 (Theorem 2), and is proved in Section 2. As a second contribution (Theorem 3), we apply our Theorem 2 to establish confidence intervals for quantities of the form

$$\mathbb{P}\Big(Y \in C \mid Z = z\Big),\ C \in \mathcal{C},\ z \in H,$$

by empirical likelihood techniques. Indeed, we prove that these confidence intervals can be built to hold uniformly in $z \in H$, $C \in \mathcal{C}$ and $h \in [h_n, \mathfrak{h}_n]$, under conditions that are very similar to those of Theorem 2. This result is stated in Section 1.3 and is proved in Section 3.

1.1 A result of Einmahl and Mason

As our first result is an extension of Theorem 1 in [5] we have to first introduce the notations and assumptions they made in their article. Consider a compact set $H \subset \mathbb{R}^d$ with nonempty interior. We shall make the following assumption on the law of (Y, Z).

(Hf) (Y, Z) has a density $f_{Y,Z}$ that is continuous in x on $\mathbb{R}^{d'} \times O'$, where $O' \subset \mathbb{R}^d$ is open and where $H \subset O'$.
Moreover f_Z is continuous and bounded away from zero and infinity on O'.

From now on, O will denote an open set fulfilling $H \subsetneq O \subsetneq O'$. Now consider a class $\mathcal{G}$ of functions from $\mathbb{R}^{d'}$ to $\mathbb{R}^k$. For $l = 1, \ldots, k$, write $\mathcal{G}_l := \Pi_l(\mathcal{G})$, where $\Pi_l(x_1, \ldots, x_l, \ldots, x_k) := x_l$ for $(x_1, \ldots, x_k) \in \mathbb{R}^k$.

(HG) Each class $\mathcal{G}_l$ is a pointwise separable VC subgraph class and has a finite valued measurable envelope function G_l satisfying, for some $p \in (2, \infty]$:
$\alpha := \max_{l=1,\ldots,k} \sup_{z \in O} \mid\mid G_l(\cdot) \mid\mid_{\mathcal{L}_{Y|Z=z},p} < \infty$,

where $\mid\mid G_l(\cdot) \mid\mid_{\mathcal{L}_{Y|Z=z},p}$ is le L^p-norm of G_l under the distribution of $Y\Big|Z = z$. For a definition of a pointwise separable VC subgraph class we refer to [11] (p. 110 and 141). Now, for any $g \in \mathcal{G}$, consider a pair of functions $(c_g(\cdot), d_g(\cdot))$, where c_g maps $\mathbb{R}^d$ to $\mathbb{R}^k$ and d_g maps $\mathbb{R}^d$ to $\mathbb{R}$, and assume that the classes of functions $\mathcal{D}_1 := \{c_g,\ g \in \mathcal{G}\}$ and $\mathcal{D}_2 := \{d_g,\ g \in \mathcal{G}\}$ are uniformly bounded and uniformly equicontinuous on O. We call this assumption $(H\mathcal{C})$. We now formulate our assumptions on the Kernel K, with the following definition.

$$\mathcal{K} := \Big\{K\big(\lambda \cdot -z\big),\ \lambda > 0,\ z \in \mathbb{R}^d\Big\}. \tag{1.2}$$

(HK1) K has bounded variation and the class $\mathcal{K}$ is VC subgraph

(HK2) $K(s) = 0$ when $s \notin [-1/2, 1/2]^d$,

(HK3) $\int_{\mathbb{R}^d} K(s)ds = 1$.

Note that (HK1) is fulfilled for a quite large class of kernels (see, e.g., Mason [8], Example F.1). In [5], Einmahl and Mason have studied the almost sure asymptotic behaviour of

$$\sup \Big\{ \mid W_n(g, h_n, z) \mid, \ g \in \mathcal{G}, \ z \in H \Big\}$$

(recall (0.1)), along a bandwidth sequence $(h_n)_{n \geq 1}$ that satisfies the following conditions (here we write $\log_2 n := \log\log(n \vee 3)$). :

$$(HV) \quad h_n \downarrow 0, \ n{h_n}^d \uparrow \infty, \ \log(1/h_n)/\log_2 n \to \infty, \ h_n^d \big(n/\log(1/h_n)\big)^{1-2/p} \to \infty,$$

where p is as in condition (HG). We also set

$$\Delta^2(g, z) := \mathbb{E}\Big(\big(< c_g(z), g(Y) > + d_g(z)\big)^2 \Big| Z = z\Big), \ z \in \mathbb{R}^d, \ g \in \mathcal{G}, \tag{1.3}$$

$$\Delta^2(g) := \sup_{z \in H} \Delta^2(g, z), \ g \in \mathcal{G} \tag{1.4}$$

$$\Delta^2(\mathcal{G}) := \sup_{g \in \mathcal{G}} \Delta^2(g). \tag{1.5}$$

Given a measurable space $(\chi, \mathcal{T})$, a measure Q and a Borel function $\psi : \ \chi \mapsto \mathbb{R}$, we write

$$|| \ \psi \ ||_{Q,p}^p = \int_\chi \mid \psi^p \mid dQ. \tag{1.6}$$

Under the above mentioned assumptions, Einmahl and Mason have proved the following theorem, λ denoting the Lebesgue measure.

Theorem 1 (Einmahl, Mason, 2000) *Under assumptions (HG), $(H\mathcal{C})$, (Hf), $(HK1)$–$(HK3)$ and (HV), we have almost surely*

$$\lim_{n \to \infty} \sup_{z \in H, \ g \in \mathcal{G}} \frac{\mid W_n(g, h_n, z) \mid}{\sqrt{2nh_n^d \log(h_n^{-d})}} = \Delta(\mathcal{G}) \ || \ K \ ||_{\lambda,2} \ . \tag{1.7}$$

We point out that (1.7) is slightly stronger than Theorem 1 of Einmahl and Mason [5], as $f_Z(z)^{-1/2}$ appears in our definition of $W_n(g, h, z)$ which is not the case in their paper. However, (1.7) is a consequence of their Theorem 1, as $f_Z^{-1/2}$ is uniformly continuous on H, by (Hf).

1.2 An extension of Theorem 1

Our first result states that Theorem 1 can be enriched by an additional uniformity in $h_n \leq h \leq \mathfrak{h}_n$ in the supremum appearing in (1.7), provided that $(h_n)_{n \geq 1}$ and $(\mathfrak{h}_n)_{n \geq 1}$ do fulfill assumption (HV). We also refer to Einmahl and Mason [6], where the authors provided some consistency results for kernel type function estimators that hold uniformly in the bandwidth (see also [12] for an improvement in the case of kernel density estimation).

Theorem 2 *Assume that (HG), (Hf), $(H\mathcal{C})$ and $(HK1) - (HK3)$ are satisfied. Let $(h_n)_{n \geq 1}$ and $(\mathfrak{h}_n)_{n \geq 1}$ be two sequences of constants fulfilling (HV) as well as $2h_n < \mathfrak{h}_n$. Then we have almost surely*

$$\lim_{n \to \infty} \sup_{z \in H, \ g \in \mathcal{G}, \ h_n \leq h \leq \mathfrak{h}_n} \frac{\mid W_n(g, h, z) \mid}{\sqrt{2nh^d \log(h^{-d})}} = \Delta(\mathcal{G}) \ || \ K \ ||_{\lambda,2} \ . \tag{1.8}$$

The proof of Theorem 2 is provided in Section 2.

Remark 1.1 *Einmahl and Mason [6] have proved a result strong enough to derive that, under weaker conditions than those of Theorem 2, we have almost surely*

$$\limsup_{n\to\infty} \sup_{\substack{z\in H,\, g\in\mathcal{G},\\ h\in[\frac{c\log n}{n},1]}} \frac{f_Z(z)^{1/2} W_n(g,h,z)}{\sqrt{nh^d\log(1/h)+\log\log n}} < \infty. \tag{1.9}$$

However, the finite constant appearing on the right hand side of (1.9) is not explicit in their result. The main advantage of Theorem 2 is that the right hand side of (1.9) is explicit, by paying the price of making stronger assumptions.

Remark 1.2 *As Theorem 2 is an extension of Theorem 1 of Einmahl and Mason, all the corollaries of Theorem 1 (see [6]) can be enriched with a uniformity in the bandwidth.*

1.3 Confidence bands by empirical likelihood

Empirical likelihood methods in statistical inference hav been introduced by Owen (see [10]). This nonparametric technique has suscitated much interest for several practical reasons, the most important one being that it directly provides confidence intervals without requiring further approximation methods, such as the estimation of dispersion parameters. Moreover, empirical likelihood is a very versatile tool which can be adapted in many different fields, for instance in estimation of densities or conditional expectations by kernel smoothing methods. The idea can be summarised as follows : consider an independent, identically distributed sample $(Y_i, Z_i)_{1\le i\le n}$ taking values in $\mathbb{R}^{d'}\times\mathbb{R}^d$. Given $h>0$, $z\in H$, a function g from $\mathbb{R}^{d'}$ to $\mathbb{R}^k$ and a (kernel) real function K, define the following centring parameter, which plays the role of a deterministic approximation of $\mathbb{E}\big(g(Y)\mid Z=z\big)$:

$$m(g,h,z) := \frac{\mathbb{E}\Big(g(Y)K\Big(\frac{Z-z}{h}\Big)\Big)}{\mathbb{E}\Big(K\Big(\frac{Z-z}{h}\Big)\Big)}. \tag{1.10}$$

This quantity is the root of the following equation in θ:

$$\mathbb{E}\Big(K\Big(\frac{Z-z}{h}\Big)\Big(g(Y)-\theta\Big)\Big) = 0, \tag{1.11}$$

which naturally leads to the following formula for a confidence interval (around $m(g,h,z)$) by empirical likelihood methods (for more details see, e.g., Owen [10], chapter 5) :

$$I_n(g,h,z,c) := \{\theta\in\mathbb{R},\ \mathcal{R}_n(\theta,g,h,z)\ge c\}, \tag{1.12}$$

where $c\in(0,1)$ is a given critical value that has to be chosen in practice, and where

$$\mathcal{R}_n(\theta,g,h,z) := \max\Big\{\prod_{i=1}^n np_i, \sum_{i=1}^n p_i K\Big(\frac{Z_i-z}{h}\Big)\big(g(Y_i)-\theta\big)=0,\ p_i\ge 0,\ \sum_{i=1}^n p_i = 1\Big\}. \tag{1.13}$$

It is known (see, e.g., [10], chapter 5) that, for fixed $z\in\mathbb{R}^d$ and fixed g, we can expect

$$m(g,h,z)\in I_n(g,h,z,c) \tag{1.14}$$

to hold with probability equal to $\mathbb{P}(\chi^2\le -2\log c)$, ultimately as $n\to\infty,\ h\to 0,\ nh^d\to\infty$ (see e.g., Owen, chapter 5). A natural arising question is:

- Can we expect (1.14) to hold uniformly in z, g and h?
- In that case, how much uniformity can we get?

Uniformity in g and z would allow to construct asymptotic *confidence bands* (instead of simple confidence intervals), while a uniformity in h would allow more flexibility in the practical choice of that smoothing parameter. Our Theorem 2 provides a tool strong enough to give some positive answers to these questions. Let us consider the case where $\mathcal{G} = \{1_C,\ C \in \mathcal{C}\}$ for a class of sets $\mathcal{C}$. We will make an abuse of notation, by identifying $\mathcal{C}$ and $\mathcal{G}$, and hence, we shall write $m(C, h, z)$ for $m\big(1_C, h, z\big)$ and so on. Write the conditional variances of $1_C(Y)$ given $Z = z$ as follows :

$$\sigma^2(C, z) := \mathbb{P}\big(Y \in C \mid Z = z\big) - \mathbb{P}^2\big(Y \in C \mid Z = z\big),\ C \in \mathcal{C},\ z \in H. \tag{1.15}$$

Our second result shows that we can construct, by empirical likelihood methods (recall (1.12)), confidence bands around the centring parameters $m(C, h, z)$ with lengths tending to zero at rate $\sqrt{2\sigma^2(C, z)\log(h^{-d})/nh^d}$ when $n \to \infty$ and $h_n \le h \le \mathfrak{h}_n$. We make the following assumptions on h_n, $\mathfrak{h}_n$ and $\mathcal{C}$

(HG') $\mathcal{C}$ is a VC class satisfying $\inf_{z \in H} \inf_{C \in \mathcal{C}} \sigma^2(C, z) =: \beta > 0$.

(HV') $h_n \downarrow 0,\ nh_n^d \uparrow \infty,\ \log(1/h_n)/\log_2 n \to \infty,\ nh_n^d/\log(1/\mathfrak{h}_n) \to \infty$.

Note that (HV') is equivalent to (HV) in the specific case where $p = \infty$.

Theorem 3 *Under assumptions (Hf), (HK1)-(HK3), (HG') and (HV'), as well as* $2h_n < \mathfrak{h}_n$*, we have almost surely:*

$$\lim_{n\to\infty} \sup_{\substack{z \in H,\ C \in \mathcal{C},\\ h_n \le h \le \mathfrak{h}_n}} \frac{-\log \mathcal{R}_n\Big(m(C, h, z), C, h, z\Big)}{\log(h^{-d})} = 1. \tag{1.16}$$

The proof of Theorem 3 is provided in Section 3.

Remark 1.3 *Theorem 3 shows that, ultimately as* $n \to \infty$*, we have*

$$\forall C \in \mathcal{C},\ \forall z \in H,\ \forall h \in [h_n, \mathfrak{h}_n] :\ m(C, h, z) \in I_n(C, h, z, h^d). \tag{1.17}$$

Hence, taking $c = h^d$ *when constructing confidence regions as in (1.12) ensures that each* $m(C, h, z)$ *belongs to its associated confidence interval* $I_n(C, h, z, c)$*. Moreover, this choice of* $c = h^d$ *is, in some sense, the best we can afford without loosing uniformity in* C, h *and* z *since, for fixed* $\epsilon > 0$*, taking* $c = h^{d-\epsilon}$ *would invalidate (1.17).*

Remark 1.4 *To calibrate the confidence region in order to obtain a coverage probability of, say,* $1 - \alpha$*, the weak convergence of the empirical likelihood process is required. This could be obtained if Theorem 2 was also enriched by a Bickel-Rosenblatt type limit law (refer to [1] for such a result in univariate kernel density estimation). However obtaining such a limit law is a real challenge in itself, and is beyond the scope of this paper. We leave that problem as an open question.*

2 Proof of Theorem 2

For ease of notations, we just prove Theorem 2 when $k = 1$. A close look at the proof shows that there is no loss of generality assuming $k = 1$.

2.1 Truncation

We start our proof of Theorem 2 as Einmahl and Mason did in their proof of Theorem 1. As the support of K is bounded and as $\mathfrak{h}_n \to 0$ we have almost surely, for all large n and for all $z \in H,\ g \in \mathcal{G},\ h_n \le h \le \mathfrak{h}_n$,

$$\begin{aligned} W_n(g,h,z) =& f_Z(z)^{-1/2}\Big[\sum_{i=1}^{n}\Big(c_g(z)g(\widetilde{Y}_i)+d_g(z)\Big)K\Big(\frac{Z_i-z}{h}\Big) \\ &-\mathbb{E}\Big(\Big(c_g(z)g(\widetilde{Y}_i)+d_g(z)\Big)K\Big(\frac{Z_i-z}{h}\Big)\Big)\Big], \end{aligned} \tag{2.1}$$

where $\widetilde{Y}_i := Y_i 1_{O'}(Z_i)$. Hence, we can suppose that $Y_i = Y_i 1_{O'}(Z_i)$ without changing the limiting behaviour of the processes we are studying here. Now consider a sequence of constants $(\gamma_n)_{n\ge1}$ fulfilling

$$\liminf_{n\to\infty} \frac{\gamma_n}{(n/\log(1/h_n))^{1/p}} > 0, \tag{2.2}$$

and consider the truncated expressions, with G denoting a measurable envelope function of $\mathcal{G}$ fulfilling (HG),

$$\begin{aligned} W_n^{\gamma_n}(g,h,z) :=& f_Z(z)^{-1/2}\sum_{i=1}^{n}\Big(c_g(z)g(Y_i)1_{\{G(Y_i)\le\gamma_n\}}+d_g(z)\Big)K\Big(\frac{Z_i-z}{h}\Big) \\ &-\mathbb{E}\Big(\Big(c_g(z)g(Y_i)1_{\{G(Y_i)\le\gamma_n\}}+d_g(z)\Big)K\Big(\frac{Z_i-z}{h}\Big)\Big). \end{aligned} \tag{2.3}$$

The following lemma allows us to study these truncated versions of the $W_n(g,h,z)$.

Lemma 2.1 *Under the assumptions of Theorem 2 and under (2.2) we have almost surely:*

$$\lim_{n\to\infty} \sup_{g\in\mathcal{G},\, z\in H,\, h_n\le h\le\mathfrak{h}_n} \frac{\Big|W_n^{\gamma_n}(g,h,z)-W_n(g,h,z)\Big|}{\sqrt{2nh^d\log(h^{-d})}} = 0. \tag{2.4}$$

Proof: A careful reading of the proof of Lemma 1 in Einmahl and Mason [5] shows that their assertions (2.8) and (2.9) remain true after adding a uniformity in $g \in \mathcal{G}$ and $h_n \le h \le \mathfrak{h}_n$, which readily implies Lemma 2.1. Note also that Lemma 2.1 is obvious when (HG) is fulfilled with $p = \infty$. □

The two next subsections are devoted to proving respectively the outer and inner bounds of Theorem 2.

2.2 Outer bounds

Fix $\epsilon > 0$. Our goal in this subsection is to show that, almost surely

$$\limsup_{n\to\infty} \sup_{g\in\mathcal{G},\, z\in H,\, h_n\le h\le\mathfrak{h}_n} \frac{|\ W_n(g,h,z)\ |}{\sqrt{2nh^d\log(h^{-d})}} \le \Delta(\mathcal{G})\ ||\ K\ ||_{\lambda,2}\ (1+4\epsilon). \tag{2.5}$$

To this aim, we shall first discretise each of the sets H, $[h_n, \mathfrak{h}_n]$ and $\mathcal{G}$ into properly chosen finite grids, then we shall control the oscillations between elements of the grids by a combination of a concentration inequality which is due to Talagrand (see also [9, 2, 7] for strongly improved versions) and of an upper bound for the first moment of these oscillations which is due to Einmahl and Mason [5].

2.2.1 Step 1: discretisations

Consider two parameters $\delta_1 \in (0,1)$, $\delta_2 \in (0,1)$ and $\rho \in (1,2)$ that will be chosen small enough in the sequel, and define the following subsequence

$$n_k := \Big[\exp\big(k/\log k\big)\Big],\ k \geq 5, \qquad N_k := \{n_{k-1}, n_{k-1}+1, \dots, n_k - 1\}. \tag{2.6}$$

Note that $n_k/n_{k-1} \to 1$ and

$$\log\log n_k = \log k(1+o(1)),\ k \to \infty. \tag{2.7}$$

We then construct the following finite grid for each $k \geq 1$

$$h_{n_k,R_k} := \mathfrak{h}_{n_{k-1}},\ \ h_{n_k,l} := \rho^l h_{n_k},\ l = 0, \dots, R_k - 1, \tag{2.8}$$

where $R_k := [\log(\mathfrak{h}_{n_{k-1}}/h_{n_k})/\log(\rho)] + 1$, and $[u]$ denotes the only integer q fulfilling $q \leq u < q+1$. Denote by $| z |_d := \max_{i=1,\dots,d} | z_i |$ the usual maximum norm on $\mathbb{R}^d$. Now, for fixed k and $0 \leq l \leq R_k$, we construct a finite grid $\mathcal{M}_{k,l} \subset H$ such that, given $z \in H$, there exists $\mathbf{z} \in \mathcal{M}_{k,l}$ fulfilling $| z - \mathbf{z} |_d < \delta_1 h_{n_k,l}$. Note that one can construct this grid so as $\sharp\mathcal{M}_{k,l} \leq C(\delta_1 h_{n_k,l})^{-d}$, where C is a constant that depends only on the volume of H. Now set $\gamma_n := \delta_2\big(n_k/\log(1/h^d_{n_k})\big)^{1/p}$, for each $k \geq 5$, $n \in N_k$. By Lemma 2.1, showing (2.5) is equivalent to showing that

$$\limsup_{n\to\infty} \sup_{z\in H,\, g\in\mathcal{G},\, h_n\leq h\leq \mathfrak{h}_n} \frac{| W_n^{\gamma_n}(g,h,z) |}{\sqrt{2nh^d\log(h^{-d})}} \leq \Delta(\mathcal{G}) \,||\, K \,||_{\lambda,2}\, (1+4\epsilon) \tag{2.9}$$

almost surely, for a proper choice of $\delta_2 > 0$.

2.2.2 Step 2: a discrete version of (2.5)

Given a real function ψ defined on a set S, we write:

$$||\, \psi \,||_S := \sup_{s\in S} | \psi(s) |\,. \tag{2.10}$$

Recall that, since f_Z is bounded away from 0 on H, we can define

$$\gamma := \inf_{z\in H} f_Z(z) > 0. \tag{2.11}$$

Also write, for convenience of notations

$$||\, c \,||_{\mathcal{G}\times H} := \sup_{g\in\mathcal{G},\, z\in H} | c_g(z) |, \quad ||\, d \,||_{\mathcal{G}\times H} := \sup_{g\in\mathcal{G},\, z\in H} | d_g(z) |\,. \tag{2.12}$$

Our first lemma is a discrete version of (2.9).

Lemma 2.2 *For any choice of*

$$0 < \delta_2 < \epsilon\Delta(\mathcal{G}) \,||\, K \,||_{\lambda,2} /(6\gamma^{1/2} \,||\, c \,||_{H\times\mathcal{G}} ||\, K \,||_{\mathbb{R}^d}), \tag{2.13}$$

and for any finite collection $\{g_1, \dots, g_q\} \subset \mathcal{G}$, *we have*

$$\limsup_{k\to\infty} \max_{\substack{n\in N_k,\, 1\leq\ell\leq q,\\ 0\leq l\leq R_k,\, \mathbf{z}\in\mathcal{M}_{k,l}}} \frac{| W_n^{\gamma_n}(g_\ell, h_{n_k,l}, \mathbf{z}) |}{\sqrt{2n_k h^d_{n_k,l}\log(1/h^d_{n_k,l})}} \leq \Delta(\mathcal{G}) \,||\, K \,||_{\lambda,2}\, (1+\epsilon). \tag{2.14}$$

Proof: We can assume here that $q=1$ with no loss of generality. We rename in this proof g_1 to g. We define, for $z \in H$, $h>0$ and $\overline{g} \in \mathcal{G}$,

$$\psi_{n_k,h,z,\overline{g}}: \ (y,x) \mapsto f_Z(z)^{-1/2}\Big[c_{\overline{g}}(z)\overline{g}(y)1_{\{G(y)\leq \gamma_{n_k}\}} + d_{\overline{g}}(z)\Big]K\Big(\frac{x-z}{h}\Big), \tag{2.15}$$

and set, for fixed $k \geq 1$ and $0 \leq l \leq R_k$,

$$\mathcal{G}_{k,l} := \Big\{\psi_{n_k,h_{n_k,l},\mathbf{z},g}, \ \mathbf{z} \in \mathcal{M}_{k,l}\Big\}. \tag{2.16}$$

First note that, for each $k \geq 5$, $0 \leq l \leq R_k$ and $\mathbf{z} \in \mathcal{M}_{k,l}$, we have, for any $g \in \mathcal{G}$,

$$\begin{aligned} \| \ \psi_{n_k,h_{n_k,l},\mathbf{z},g} \ \|_{\mathbb{R}^{d'}\times\mathbb{R}^d} \leq & \big(\ \| \ c \ \|_{H\times\mathcal{G}} \ \gamma_{n_k} + \| \ d \ \|_{H\times\mathcal{G}} \ \big)\gamma^{-1/2} \ \| \ K \ \|_{\mathbb{R}^d} \\ \leq & 2 \ \| \ c \ \|_{H\times\mathcal{G}} \ \gamma^{-1/2} \ \| \ K \ \|_{\mathbb{R}^d} \ \delta_2(n_k h_{n_k}^d / \log(1/h_{n_k}^d))^{1/2} && (2.17) \\ \leq & \frac{\epsilon}{3} \ \| \ K \ \|_{\lambda,2}^2 \ \Delta(\mathcal{G})(n_k h_{n_k}^d/\log(1/h_{n_k}^d))^{1/2}, && (2.18) \end{aligned}$$

where (2.17) holds for all large k, uniformly in $0 \leq l \leq R_k$ and $\mathbf{z} \in \mathcal{M}_{k,l}$, according to assumption (HV), and where (2.18) holds by (2.13). Moreover we have (recall $(HK2)$)

$$\begin{aligned} & \mathrm{Var}\Big(\psi_{n_k,h_{n_k,l},\mathbf{z},g}(Y,Z)\Big) \\ \leq & \ \mathbb{E}\Big(\psi^2_{n_k,h_{n_k,l},\mathbf{z},g}(Y,Z)\Big) \\ \leq & \ \mathbb{E}\bigg(f_Z(\mathbf{z})^{-1}\Big(c_g(\mathbf{z})g(Y)+d_g(\mathbf{z})\Big)^2 K\Big(\frac{Z-\mathbf{z}}{h_{n_k,l}}\Big)^2\bigg) \\ & + f_Z(\mathbf{z})^{-1} \ \| \ d \ \|^2_{H\times\mathcal{G}} \| \ K \ \|^2_{\mathbb{R}^d} \ \mathbb{P}\Big(\big\{G(Y)\geq \gamma_{n_k}\big\}\cap\big\{ \ | \ Z-\mathbf{z} \ |_d \leq h_{n_k,l}/2\big\}\Big) && (2.19) \\ =: & A_1 + A_2. \end{aligned}$$

The first term on the right hand side of (2.19) is equal to

$$A_1 = \int_{|z-\mathbf{z}|_d \leq h_{n_k,l}/2} \mathbb{E}\bigg(\Big(c_g(\mathbf{z})g(Y)+d_g(\mathbf{z})\Big)^2\Big|Z=z\bigg)\frac{f_Z(z)}{f_Z(\mathbf{z})}K^2\Big(\frac{z-\mathbf{z}}{h_{n_k,l}}\Big)dz.$$

It follows, by making use of assumption $(H\mathcal{C})$, that there exists a function $r(\cdot)$ fulfilling $r(u) \to 0$ as $u \to 0$ and such that

$$\begin{aligned} A_1 \leq & \int_{|z-\mathbf{z}|_d \leq h_{n_k,l}/2} \Delta^2(g,z)\frac{f_Z(z)}{f_Z(\mathbf{z})}K^2\Big(\frac{z-\mathbf{z}}{h_{n_k,l}}\Big)dz + r(h_{n_k,l}) && (2.20) \\ \leq & \Delta^2(\mathcal{G})h^d_{n_k,l} \int_{[-1/2,1/2]^d} K^2(u)\frac{f_Z\big(\mathbf{z}+h_{n_k,l}u\big)}{f_Z(\mathbf{z})}du\big(1+r(h_{n_k,l})\big) && (2.21) \\ \leq & \Delta^2(\mathcal{G}) \ \| \ K \ \|^2_{\lambda,2} \ h^d_{n_k,l}(1+\varepsilon_{k,l}), && (2.22) \end{aligned}$$

where

$$\varepsilon_{k,l} := \sup_{z\in H, \ |u|_d \leq 1/2}\left|\frac{f_Z\big(z+h_{n_k,l}u\big)}{f_Z(z)}\big(1+r(h_{n_k,l})\big)-1\right|. \tag{2.23}$$

By assumption (Hf) and since $h_{n_k,l} \leq \mathfrak{h}_{n_{k-1}} \to 0$ we readily infer that

$$\lim_{k\to\infty} \max_{0\leq l\leq R_k} \varepsilon_{k,l} = 0.$$

Moreover we have, uniformly in $0 \leq l \leq R_k$ and $\mathbf{z} \in \mathcal{M}_{k,l}$ (recall (HG) and (Hf))

$$\begin{aligned}
&\mathbb{P}\Big(\{G(Y) \geq \gamma_{n_k}\} \cap \{\mid Z - \mathbf{z} \mid_d \leq h_{n_k,l}/2\}\Big) \\
\leq &\gamma_{n_k}^{-2} \int\limits_{|z-\mathbf{z}|_d \leq h_{n_k,l}/2} \mathbb{E}\Big(G^2(Y) \mid Z = z\Big) f_Z(z) dz \\
\leq &\gamma_{n_k}^{-2} h_{n_k,l}^d \ \alpha \int\limits_{[-1/2,1/2]^d} f_Z\big(\mathbf{z} + h_{n_k,l} u\big) du. \\
\leq &\gamma_{n_k}^{-2} h_{n_k,l}^d \ \alpha \mid\mid f_Z \mid\mid_O .
\end{aligned}$$

As $\gamma_{n_k} \to \infty$ we conclude that, for all large enough k and for each $0 \leq l \leq R_k$, $\mathbf{z} \in \mathcal{M}_{k,l}$,

$$\mathrm{Var}\Big(\psi_{n_k, h_{n_k,l}, \mathbf{z}, g}(Y, Z)\Big) \leq \Delta(\mathcal{G}) \mid\mid K \mid\mid_{\lambda,2} (1+\epsilon) h_{n_k,l}^d. \tag{2.24}$$

Given a real function $g : \mathbb{R} \times \mathbb{R}^d \mapsto \mathbb{R}$, we shall write

$$T_n(g) := \sum_{i=1}^n \Big\{ g(Y_i, Z_i) - \mathbb{E}\Big(g(Y_i, Z_i)\Big) \Big\}. \tag{2.25}$$

Combining (2.18) and (2.24) making use of the maximal version of Bernstein's inequality (see, e.g. Einmahl and Mason [4], Lemma 2.2) repeatedly for each $0 \leq l \leq R_k$, $\mathbf{z} \in \mathcal{M}_{k,l}$, we have, for all large k (recall that $\sharp \mathcal{M}_{k,l} \leq C \delta_1^{-d} h_{n_k,l}^{-d}$),

$$\begin{aligned}
&\mathbb{P}\Bigg(\max_{\substack{n \in N_k,\ 0 \leq l \leq R_k, \\ \mathbf{z} \in \mathcal{M}_{k,l}}} \frac{\mid W_n^{\gamma_n}(g, h_{n_k,l}, \mathbf{z}) \mid}{\sqrt{2 n_k h_{n_k,l}^d \log(1/h_{n_k,l}^d)}} > \Delta(\mathcal{G}) \mid\mid K \mid\mid_{\lambda,2} (1+\epsilon) \Bigg) \\
\leq &\sum_{l=0}^{R_k} \sharp \mathcal{M}_{k,l} \max_{\mathbf{z} \in \mathcal{M}_{k,l}} \mathbb{P}\Bigg(\max_{n \in N_k} \Big|\Big| T_n \Big|\Big|_{\mathcal{G}_{k,l}} \geq \Delta(\mathcal{G}) \mid\mid K \mid\mid_{\lambda,2} (1+\epsilon) \sqrt{2 n_k h_{n_k,l}^d \log(1/h_{n_k,l}^d)} \Bigg) \\
\leq &\sum_{l=0}^{R_k} \frac{C}{\delta_1^d h_{n_k,l}^d} 2 \exp\Big(- \big(1 + \epsilon^2/(1+\epsilon)\big) \log(1/h_{n_k,l}^d) \Big) \\
\leq &\frac{2C}{\delta_1^d} \sum_{l=0}^{R_k} h_{n_k,l}^{d\epsilon^2/2} \\
= &\frac{2C}{\delta_1^d} \sum_{l=0}^{R_k} \rho^{l d\epsilon^2/2} h_{n_k}^{d\epsilon^2/2} = \frac{2C}{\delta_1^d} h_{n_k}^{d\epsilon^2/2} \frac{\rho^{(R_k+1)d\epsilon^2/2} - 1}{\rho^{d\epsilon^2/2} - 1} \leq \frac{2C \rho^{d\epsilon^2/2}}{\delta_1^d (\rho^{d\epsilon^2/2} - 1)} \mathfrak{h}_{n_{k-1}}^{d\epsilon^2/2},
\end{aligned} \tag{2.26}$$

where the last inequality is a consequence of $R_k := [\log(\mathfrak{h}_{n_{k-1}}/h_{n_k})/\log(\rho)] + 1$. As $\log(1/\mathfrak{h}_{n_{k-1}})/\log\log n_{k-1} \to \infty$ (assumption (HV)), and by (2.7), the right hand side of expression (2.26) is summable in k. The proof of Lemma 2.2 now readily follows by making use of the Borel-Cantelli lemma. □

2.2.3 Step 3: end of the proof of Theorem 2

Our next lemma allows us to extend the uniformity in Lemma 2.2 to the whole sets $\mathcal{G}$, $[h_{n_k}, \mathfrak{h}_{n_{k-1}}]$ and H, provided that $\delta_1 > 0, \delta_2 > 0, \rho > 1$ and $\{g_1, \ldots, g_q\}$ have been properly chosen. Before stating our lemma, we need to recall three facts. We shall be able to properly discretise the class $\mathcal{G}$ by making use of the following result, which is a straightforward adaptation of Lemma 6 of Einmahl and Mason [5].

Fact 2.1 (Einmahl, Mason, 2000) *Given $\varepsilon > 0$, there exists $h_{0,\varepsilon} > 0$ and a finite subclass $\{g_1, \dots, g_q\} \subset \mathcal{G}$ (that may depend on ε) fulfilling*

$$\sup_{\substack{0<h<h_{0,\varepsilon},\\ z\in H,\ g\in\mathcal{G}}} \min_{\ell=1,\dots,q} h^{-d} f_Z(z)^{-1}\mathbb{E}\Big[\Big(\big(c_g(z)g(Y)+d_g(z)\big)-\big(c_{g_\ell}(z)g_\ell(Y)+d_{g_\ell}(z)\big)\Big)^2 K^2\Big(\frac{Z-z}{h}\Big)\Big] \leq \varepsilon/2.$$

Now define the following distances on $\mathcal{G}$:

$$\begin{aligned} d^2(g_1,g_2) :=& \sup_{\substack{0<h<h_{0,\varepsilon},\\ z\in H}} h^{-d} f_Z(z)^{-1}\mathbb{E}\Big[\Big(\big(c_{g_1}(z)g_1(Y)+d_{g_1}(z)\big)-\big(c_{g_2}(z)g_2(Y)+d_{g_2}(z)\big)\Big)^2 \\ &\times K^2\Big(\frac{Z-z}{h}\Big)\Big], \\ \widetilde{d}(g_1,g_2) :=& \max\big\{d(g_1,g_2),\ \|\, c_{g_1}-c_{g_2}\,\|_H,\ \|\, d_{g_1}-d_{g_2}\,\|_H\ \big\}. \end{aligned} \tag{2.27}$$

We write $|\, K\,|_v$ for the total variation of K and we set, for $\psi : \mathbb{R}^d \mapsto \mathbb{R}$,

$$\omega_\psi(\delta) := \sup_{z_1,z_2\in H,\ |z_1-z_2|_d\leq\delta} \Big|\frac{\psi(z_2)}{f_Z(z_2)} - \frac{\psi(z_1)}{f_Z(z_1)}\Big|,\ \delta>0, \tag{2.28}$$

$$\beta_1 := \sup_{z\in O}\mathbb{E}\Big(\big(G^2(Y)+1\big)\Big|Z=z\Big) < \infty, \tag{2.29}$$

$$B := 4\beta_1\,\|\, f_Z\,\|_O \|\, f_Z^{-1}\,\|_O\ \Big(\ \|\, K\,\|^2_{\mathbb{R}^d} + \big(\sup_{g\in\mathcal{G}}\|\, c_g\,\|_O \vee \sup_{g\in\mathcal{G}}\|\, d_g\,\|_O\ \big)\ |\, K\,|_v^2\ \Big). \tag{2.30}$$

The following fact is a straightforward adaptation of Lemma 4 and Lemma 6 in [5].

Fact 2.2 (Einmahl, Mason, 2000) *Fix $\varepsilon > 0$. For any $\delta \in (0, 1/2)$ and $0 < h < h_{0,\varepsilon}$ fulfilling*

$$z + (2h)u \in O, \tag{2.31}$$

and for all large k we have, for each $z \in H$ and for each $u \in \mathbb{R}^d$ with $|\, u\,|_d \leq 1$, for each $\rho \in (1, 2]$, $z_1, z_2 \in H$ with $|\, z_1 - z_2\,|_d \leq (\delta h)$, and for each $g_1, g_2 \in \mathcal{G}$ fulfilling $\widetilde{d}^2(g_1, g_2) \leq \varepsilon$,

$$\begin{aligned} &\mathbb{E}\Big(\Big(\psi_{n_k,z_1,\rho h,g_1}(Y,Z) - \psi_{n_k,z_2,h,g_2}(Y,Z)\Big)^2\Big) \\ \leq& B\Big(\omega^2_{c_{g_2}}(\delta h) \vee \omega^2_{d_{g_2}}(\delta h) + \rho - 1 + \delta + \varepsilon\Big)h^d. \end{aligned} \tag{2.32}$$

Remarks: Assumption (2.31) is just technical, in order to have the continuity arguments of Einmahl and Mason valid. The presence of the term $\rho-1$ on the right hand side of (2.32) is due to the fact that we take care of the differences $h/h_{n_k,l} - 1$, which are implicitely handled in Lemma 6 of Einmahl and Mason [5].
The third fact is also largely inspired by the ideas of Einmahl and Mason [5]. We remind that the uniform entropy number of a class of functions $\mathcal{F}$ with measurable envelope F is defined as

$$\mathcal{N}(\epsilon, \mathcal{F}) := \sup_{Q \text{ proba}} \min\big\{p \geq 1,\ \exists (g_1, \dots, g_p) \in \mathcal{F}^p,\ \sup_{g\in\mathcal{F}} \min_{i=1,\dots,p} \|\, g - g_i\,\|_{Q,2} \leq \epsilon \,\|\, F\,\|_{Q,2}\ \big\},$$

where the supremum is taken over all probability measures Q. The following fact is proved in [12] (Proposition 2.1).

Fact 2.3 (Varron, 2006) *Let $\mathcal{F}$ be a class of functions on $\mathbb{R}^d$ with measurable envelope function F satisfying, for some constants $\tau > 0$ and $h \in (0,1)$,*

$$\sup_{g\in\mathcal{F}} \operatorname{Var}\big(g(Z_1)\big) \leq \tau^2 h^d.$$

Assume that there exists $\delta_0, C, v, \beta_0 > 0$ and $p > 2$ fulfilling, for all $0 < \epsilon < 1$,

$$\mathcal{N}(\epsilon, \mathcal{F}) \leq C\epsilon^{-v}, \tag{2.33}$$

$$\mathbb{E}\Big(F_Z(Y)^2\Big) \leq \beta_0^2, \tag{2.34}$$

$$\sup_{g\in\mathcal{F},\ z\in\mathbb{R}^d} \mid g(z) \mid \leq \delta_0 (nh^d/\log(h^{-d}))^{1/p}. \tag{2.35}$$

Then there exists a universal constant $A > 0$ and a parameter $D(v) > 0$ depending only on v such that, for fixed $\rho_0 > 0$, if $h > 0$ satisfies,

$$K_1 := \max\Big\{1, \big(4\delta_0\sqrt{v+1}/\tau\big)^{\frac{1}{1/2-1/p}}, \big(\rho_0\delta_0/\tau^2\big)^{\frac{1}{1/2-1/p}}\Big\} \leq \frac{nh^d}{\log(h^{-d})}, \tag{2.36}$$

$$K_2 := \min\big\{1/(\tau^2\beta_0), \tau^2\big\} \geq h^d, \tag{2.37}$$

then we have

$$\mathbb{P}\Big(\max_{1\leq m\leq n} \mid\mid T_m \mid\mid_{\mathcal{F}} \geq (\tau+\rho_0) D (nh^d \log(h^{-d}))^{1/2}\Big) \leq 4\exp\Big(-A(\frac{\rho_0}{\tau})^2 \log(h^{-d})\Big).$$

We can now state our second lemma, which will conclude the proof of the outer bounds of Theorem 2. Recall that $\epsilon > 0$ was fixed at the very beginning of our proof (see Section 2.2).

Lemma 2.3 *There exists a finite class $g_1, \ldots, g_q \in \mathcal{G}$ as well as two constants $\rho_\epsilon > 1$ and $\delta_{1,\epsilon} > 0$ small enough such that, for each $1 < \rho \leq \rho_\epsilon$ and each $0 < \delta_1 \leq \delta_{1,\epsilon}$, we have almost surely :*

$$\limsup_{k\to\infty} \max_{\substack{n\in N_k,\\ 0\leq l\leq R_k-1}} \sup_{g\in\mathcal{G}} \inf_{1\leq \ell\leq q} \sup_{\substack{z_1,z_2\in H,\ |z_1-z_2|<\delta,\\ h_{n_k,l}\leq h\leq \rho h_{n_k,l}}} \frac{\Big|W_n^{\gamma_n}(g, z_1, h_{n_k,l}) - W_n^{\gamma_n}(g_\ell, z_2, h)\Big|}{\sqrt{2n_k h^d_{n_k,l} \log(1/h^d_{n_k,l})}}$$
$$\leq \Delta(\mathcal{G}) \mid\mid K \mid\mid_{\lambda,2} \epsilon. \tag{2.38}$$

Proof :
Consider the class

$$\mathcal{G}' := \Big\{(y,z) \mapsto u_1\big(c_{g_1}(z_1) g_1(y) 1_{\{G(y)\leq t\}} + d_{g_1}(z_1)\big) K\Big(\frac{z-z_1}{h}\Big)$$
$$- u_2\big(c_{g_2}(z_2) g_2(y) 1_{\{G(y)\leq t\}} + d_{g_2}(z_2)\big) K\Big(\frac{z-z_2}{\mathfrak{h}}\Big),\ z_1, z_2 \in \mathbb{R}^d,\ g_1, g_2 \in \mathcal{G},$$
$$t \geq 0,\ (h, \mathfrak{h}) \in (0,1)^2,\ u_1, u_2 \in \Big[\inf_H f_Z^{-1/2}, \sup_H f_Z^{-1/2}\Big]\Big\}.$$

Recall that $\gamma = \inf_H f$ and note that $\mathcal{G}'$ admits the following function as an envelope function:

$$G'(y,z) := 2\gamma^{-1/2}\big(\mid\mid c \mid\mid_{H\times\mathcal{G}} G(y) + \mid\mid d \mid\mid_{H\times\mathcal{G}} \big) \mid\mid K \mid\mid_{\mathbb{R}^d} . \tag{2.39}$$

Set $\beta_4^2 := \mathbb{E}\big(G'^2(Y,Z)\big) < \infty$ (the finiteness of β_4 follows from (Hf) and (HG)). By an argument very similar to that used by Einmahl and Mason (see their Lemma 5 in [5]) we readily infer that there exist $C > 0$ and $v > 0$ fulfilling

$$\mathcal{N}(\epsilon, \mathcal{G}') \leq C\epsilon^{-v},\ \epsilon \in (0,1]. \tag{2.40}$$

Recalling the notations of Fact 2.3, we set $\varepsilon = D(v)^{-1}(1+\sqrt{2/A})^{-1}\epsilon\Delta(\mathcal{G}) \mid\mid K \mid\mid_{\lambda,2}$. By Fact 2.1 and by $(H\mathcal{C})$, for any $\varepsilon > 0$, we can choose a finite subclass $\{g_1, \dots, g_q\} \subset \mathcal{G}$ such that $\mathcal{G}$ is included in the finite reunion of the corresponding balls with $\widetilde{d}$-radius smaller than $\varepsilon/2$. For fixed $k \geq 5,\ 0 \leq l \leq R_k - 1,\ 1 \leq \ell \leq q$ and $\delta > 0$, define the following class of functions:

$$\begin{aligned}\mathcal{G}_{k,l,q,\delta} :=& \Big\{\psi_{n_k,z_1,h,g} - \psi_{n_k,z_2,h_{n_k,l},g_\ell},\ z_1, z_2 \in H,\ \mid z_1 - z_2 \mid \leq \delta,\\ &\widetilde{d}(g, g_\ell) \leq \varepsilon/2,\ h_{n_k,l} \leq h \leq \rho h_{n_k,l}\Big\}.\end{aligned}$$

Obviously we always have $\mathcal{G}_{k,l,\ell,\delta} \subset \mathcal{G}'$. By inclusion, all the classes $\mathcal{G}_{k,l,\ell,\delta}$ inherit properties (2.39) and (2.40). Moreover, proving Lemma 2.3 is equivalent to showing that, almost surely

$$\limsup_{k\to\infty} \max_{\substack{0\leq l\leq R_k-1,\\ 1\leq \ell\leq q}} \frac{\max_{n\in N_k} \mid\mid T_n \mid\mid_{\mathcal{G}_{k,l,\ell,\delta}}}{\sqrt{2n_k h_{n_k,l}^d \log(1/h_{n_k,l}^d)}} \leq \Delta(\mathcal{G}) \mid\mid K \mid\mid_{\lambda,2} \epsilon. \tag{2.41}$$

As $h_{n_k,l} \leq \mathfrak{h}_{n_{k-1}} \to 0$ and by Fact 2.2, we can choose $\delta_{1,\epsilon} > 0$ such that, for each $0 < \delta_1 < \delta_{1,\epsilon}$, for all large k and for all $0 \leq l \leq R_k - 1$,

$$\sup_{\psi\in\mathcal{G}_{k,l,\ell,\delta_{1,\epsilon}}} h_{n_k,l}^{-d}\mathbb{E}\big(\psi^2(Y,Z)\big) \leq \varepsilon^2. \tag{2.42}$$

Recalling that $h_{n_k} \leq h_{n_k,l} \leq \mathfrak{h}_{n_{k-1}}$ and assumption (HV), we can choose k large enough so that each class $\mathcal{G}_{k,l,\ell,\delta_1}$ fulfills conditions (2.36) and (2.37) with $\beta_0 := \beta_4,\ h := h_{n_k,l}$, $n := n_k,\ \tau := \varepsilon,\ \rho := \sqrt{2/A}\tau$ and C, v appearing in (2.40). Hence, we have, uniformly in $0 \leq l \leq R_k - 1$ and $1 \leq \ell \leq q$,

$$\begin{aligned}&\mathbb{P}\Big(\max_{n\in N_k} \mid\mid T_n \mid\mid_{\mathcal{G}_{k,l,\delta_{1,\varepsilon}}} > \Delta(\mathcal{G}) \mid\mid K \mid\mid_{\lambda,2} \epsilon\sqrt{2n_k h_{n_k,l}^d \log(1/h_{n_k,l}^d)}\Big)\\ \leq&\mathbb{P}\Big(\max_{n\in N_k} \mid\mid T_n \mid\mid_{\mathcal{G}_{k,l,\delta_1}} \geq D(v)(\tau+\rho)\sqrt{2n h_{n_k,l}^d \log(1/h_{n_k,l}^d)}\Big)\\ \leq&4\exp\Big(-2\log(1/h_{n_k,l}^d)\Big).\end{aligned}$$

Now, by Bonferroni's inequality we have, for all large k,

$$\begin{aligned}&\mathbb{P}\Big(\bigcup_{l=0}^{R_k-1}\bigcup_{j=1}^{J_l} \max_{n\in N_k} \mid\mid T_n \mid\mid_{\mathcal{G}_{k,l,\delta_{1,\varepsilon}}} > \Delta(\mathcal{G}) \mid\mid K \mid\mid_{\lambda,2} \epsilon\sqrt{2n_k h_{n_k,l}^d \log(1/h_{n_k,l}^d)}\Big)\\ \leq& \sum_{l=0}^{R_k-1} 4\sharp\mathcal{M}_{k,l} h_{n_k,l}^{2d} \leq \frac{4C}{\delta_1}\sum_{l=0}^{R_k-1} h_{n_k,l}^d \leq \frac{4C}{1-\rho^d}\rho^{dR_k} \leq \frac{4C\rho^d}{1-\rho^d}\mathfrak{h}_{n_{k-1}}.\end{aligned}$$

As $\log(1/h_{n_k})/\log\log(n_k) \to \infty$ by (HV) and (2.7), the proof of Lemma 2.3 is concluded by a straightforward use of the Borel-Cantelli lemma. □

Combining Lemmas 2.2 and 2.3 we get, for any choice of $\delta_1,\ \delta_2 > 0$ and $\rho > 1$ small enough,

$$\limsup_{k\to\infty} \max_{\substack{n\in N_k,\\ 0\le l\le R_k-1}} \sup_{\substack{z\in H,\\ h_{n_k,l}\le h\le \rho h_{n_k,l}}} \frac{\left|W_n^{\gamma_n}(g,h,z)\right|}{\sqrt{2n_k h_{n_k,l}^d \log(1/h_{n_k,l}^d)}} \le \Delta(\mathcal{G})(1+3\epsilon) \text{ a.s.} \tag{2.43}$$

Now, as $n_k/n_{k-1}\to 1$,assertion (2.5) is implied by the following assertion

$$\limsup_{k\to\infty} \sup_{\substack{1<\rho'\le\rho,\\ h\in(h_{n_k},\mathfrak{h}_{n_{k-1}})}} \left|\frac{\rho h^d \log(1/\rho h^d)}{h^d\log(h^{-d})} - 1\right| \le \epsilon/(1+5\epsilon), \tag{2.44}$$

which, by routine computations, turns out to be true if we choose $\rho > 1$ small enough. This concludes the proof of the outer bounds of Theorem 2.□

2.3 Inner bounds

Proving the inner bounds of Theorem 2 is a simple consequence of Theorem 1, since, almost surely,

$$\liminf_{n\to\infty} \sup_{\substack{z\in H,\, C\in\mathcal{C},\\ h_n\le h\le\mathfrak{h}_n}} \frac{\mid W_n(g,h,z)\mid}{\sqrt{2nh^d\log(h^{-d})}} \ge \liminf_{n\to\infty} \sup_{\substack{z\in H,\\ C\in\mathcal{C}}} \frac{\mid W_n(g,h_n,z)\mid}{\sqrt{2nh_n^d\log(h_n^{-d})}} = \Delta(\mathcal{G}) \mid\mid K \mid\mid_{\lambda,2}, \tag{2.45}$$

where (2.45) is a consequence of Theorem 1.

3 Proof of Theorem 3

Our proof of Theorem 3 is inspired by Owen (see [10], chapter 5) and borrows some ideas of Härdle *et al.* (see their Lemma 1 in [3]). Set, for $n\ge 1$, $C\in\mathcal{C}$, $h>0$ and $z\in H$,

$$X_n(C,h,z) := \sum_{i=1}^n K\Big(\frac{Z_i-z}{h}\Big)\Big(1_C(Y_i) - m(C,h,z)\Big), \tag{3.1}$$

$$S_n(C,h,z) := f_Z(z)^{-1}\sum_{i=1}^n \Big[K\Big(\frac{Z_i-z}{h}\Big)\Big(1_C(Y_i)-m(C,h,z)\Big)\Big]^2, \tag{3.2}$$

$$w_{i,n}(C,h,z) := K\Big(\frac{Z_i-z}{h}\Big)\Big(1_C(Y_i)-m(C,h,z)\Big). \tag{3.3}$$

The proof of Theorem 3 consists in showing that the quantities

$$-2\log\Big(\mathcal{R}_n\big(m(C,h,z),C,h,z\big)\Big),\ C\in\mathcal{C},\ z\in H,\ h\in[h_n,\mathfrak{h}_n]$$

are asymptotically equivalent to

$$U_n(C,h,z) := \frac{X_n(C,h,z)^2}{f_Z(z)S_n(C,h,z)},\ C\in\mathcal{C},\ z\in H,\ h\in[h_n,\mathfrak{h}_n], \tag{3.4}$$

and in establishing the almost sure limit behaviour of the quantities $U_n(C,h,z)$. Recall that $\sigma^2(C,z) := \mathrm{Var}\big(1_C(Y)\mid Z=z\big)$ and write

$$r(C,z) := \mathbb{E}\big(1_C(Y)\mid Z=z\big). \tag{3.5}$$

By (Hf) together with Scheffé's lemma, both $\sigma^2(C,\cdot)$ and $r(C,\cdot)$ are equicontinuous uniformly in $C \in \mathcal{C}$, namely

$$\lim_{\delta\to 0}\sup_{C\in\mathcal{C}}\sup_{\substack{z_1,z_2\in H\\ |z_1-z_2|\le\delta}} | r(C,z_1) - r(C,z_2) | = 0, \tag{3.6}$$

$$\lim_{\delta\to 0}\sup_{C\in\mathcal{C}}\sup_{\substack{z_1,z_2\in H\\ |z_1-z_2|\le\delta}} | \sigma^2(C,z_1) - \sigma^2(C,z_2) | = 0. \tag{3.7}$$

3.1 Step 1: an application of Theorem 2

Recall that $\sigma^2(C,z) := \mathrm{Var}\big(1_C(Y_i) \mid Z=z\big)$ and that $r(C,z) := \mathbb{P}\big(Y \in C \mid Z = z\big)$. In this first step we prove that, given $\epsilon > 0$, we have $(2\log(h^{-d}))^{-1}U_n(C,h,z) \le (1+\epsilon)$ uniformly in C, h, z, ultimately as $n \to \infty$.

Lemma 3.1 *Under the assumptions of Theorem 3, we have almost surely :*

$$\lim_{n\to\infty}\sup_{\substack{z\in H,\, C\in\mathcal{C},\\ h_n\le h\le\mathfrak{h}_n}} \Big| \frac{S_n(C,h,z)}{nh^d\sigma^2(C,z) \mid\mid K \mid\mid_{\lambda,2}^2} - 1 \Big| = 0, \tag{3.8}$$

$$\lim_{n\to\infty}\sup_{\substack{z\in H,\, C\in\mathcal{C},\\ h_n\le h\le\mathfrak{h}_n}} \Big| \frac{X_n(C,h,z)}{\sqrt{2f_Z(z)\sigma^2(C,z) \mid\mid K \mid\mid_{\lambda,2}^2\, nh^d\log(h^{-d})}} \Big| = 1. \tag{3.9}$$

As a consequence we have

$$\lim_{n\to\infty}\sup_{\substack{z\in H,\, C\in\mathcal{C},\\ h_n\le h\le\mathfrak{h}_n}} \Big| \frac{U_n(C,h,z)}{2\log(h^{-d})} \Big| = 1 \ \ a.s. \tag{3.10}$$

Proof:
Note that (3.10) is a consequence of (3.8) and (3.9). Set $L(\cdot) = K^2(\cdot) \mid\mid K \mid\mid_{\lambda,2}^{-2}$. To apply Theorem 2 we write $\big(1_C(Y) - r(C,z)\big)^2 = 1_C(Y)(1-2r(C,z)) + r^2(C,z)$. Notice that, under (HG') and (HV'), the class $\mathcal{C}$ and the sequence $(h_n)_{n\ge 1}$ satisfy the conditions of Theorem 2 with $p=\infty$. By Scheffé's lemma together with assumption (Hf) and (HG'), the two following collections of functions are uniformly equicontinuous on H :

$$\mathcal{D}_1 := \big\{f_Z^{-1/2}(\cdot)(1-2r(C,\cdot)),\ C\in\mathcal{C}\big\},\ \ \mathcal{D}_2 := \big\{f_Z^{-1/2}(\cdot)r^2(C,\cdot),\ C\in\mathcal{C}\big\}. \tag{3.11}$$

We can hence apply Theorem 2 to the class $\mathcal{C}$, with $\mathcal{D}_1, \mathcal{D}_2$ defined as above, and with the kernel L to obtain, with probability one,

$$\lim_{n\to\infty}\sup_{z\in H,\, C\in\mathcal{C},\, h_n\le h\le\mathfrak{h}_n} \frac{| \widetilde{W}_n(C,h,z) |}{\sqrt{2nh^d\log(h^{-d})}} < \infty, \tag{3.12}$$

with

$$\widetilde{W}_n(C,h,z) := f_Z(z)^{-1}\sum_{i=1}^n \Big\{ \big(1_C(Y_i) - r(C,z)\big)^2 L\Big(\frac{Z_i-z}{h}\Big) - \mathbb{E}\Big[\big(1_C(Y_i) - r(C,z)\big)^2 L\Big(\frac{Z_i-z}{h}\Big)\Big]\Big\}. \tag{3.13}$$

Now write

$$\mathbb{E}\bigg(\big(1_C(Y_i) - r(C,z)\big)^2 L\Big(\frac{Z_i-z}{h}\Big)\bigg) =: \widetilde{r}(C,h,z). \tag{3.14}$$

By assumptions (HG'), (HV') and (Hf) together with Scheffé's lemma, we can infer that

$$\lim_{n\to\infty} \sup_{\substack{z\in H,\, C\in\mathcal{C},\\ h\in[h_n,\mathfrak{h}_n]}} \Big| \frac{\widetilde{r}(C,h,z)}{h^d f_Z(z)\sigma^2(C,z)} - 1 \Big| = 0 \tag{3.15}$$

$$\lim_{n\to\infty} \sup_{\substack{z\in H,\, C\in\mathcal{C},\\ h\in[h_n,\mathfrak{h}_n]}} \mid m(C,h,z) - r(C,z) \mid = 0, \tag{3.16}$$

$$\lim_{n\to\infty} \sup_{h\in[h_n,\mathfrak{h}_n]} \frac{\log(h^{-d})}{nh^d} = 0. \tag{3.17}$$

Writing

$$\Big(S_n(C,h,z) - nf_Z(z)^{-1} \mid\mid K \mid\mid_{\lambda,2}^2 \widetilde{r}(C,h,z)\Big) - \widetilde{W}_n(C,h,z)$$
$$= f_Z(z)^{-1}\Big[\big(m^2(C,h,z) - r^2(C,z)\big) - 2\big(m(C,h,z) - r(C,z)\big)\Big] \mid\mid K \mid\mid_{\lambda,2}^2 \sum_{i=1}^{n} L\Big(\frac{Z_i - z}{h}\Big),$$

we conclude by Theorem 2 and (3.16) that

$$\lim_{n\to\infty} \sup_{\substack{z\in H,\, C\in\mathcal{C},\\ h_n\le h\le\mathfrak{h}_n}} \frac{\Big|\Big(S_n(C,h,z) - nf_Z(z)^{-1} \mid\mid K \mid\mid_{\lambda,2}^2 \widetilde{r}(C,h,z)\Big) - \widetilde{W}_n(C,h,z)\Big|}{\sqrt{2nh^d\log(h^{-d})}} = 0$$

with probability one, from where we obtain with (3.14) and (3.15) that

$$\lim_{n\to\infty} \sup_{\substack{z\in H,\, C\in\mathcal{C},\\ h_n\le h\le\mathfrak{h}_n}} \frac{\Big|S_n(C,h,z) - nh^d\sigma^2(C,z)\Big|}{\sqrt{2nh^d\log(h^{-d})}} < \infty. \tag{3.18}$$

The proof of (3.8) is now concluded, by (3.17), (3.18) and (HG'). Assertion (3.9) can be proved in a very similar way, taking care that the class $\mathcal{D} := \big\{f_Z(\cdot)^{-1/2}\sigma(C,\cdot)^{-1}\big\}$ is uniformly equicontinuous and bounded away from zero and infinity on H. We omit details. □

3.2 Step 2: convex hull condition

The second step of our proof of Theorem 3 is usually called the "convex hull condition".

Lemma 3.2 *With probability one, we have, for all large n and for all $C \in \mathcal{C}$, $z \in H$, $h_n \le h \le \mathfrak{h}_n$,*

$$\sharp\Big\{i:\ K\Big(\frac{Z_i - z}{h}\Big)\big(1_C(Y_i) - m(C,h,z)\big) > 0\Big\} \in \{1,2,\ldots,n-1\}. \tag{3.19}$$

Proof: It is sufficient to prove that

$$\liminf_{n\to\infty} \inf_{\substack{z\in H,\, C\in\mathcal{C},\\ h\in[h_n,\mathfrak{h}_n]}} \mathbb{P}\Big(\pm\big(1_C(Y) - m(C,h,z)\big)K\Big(\frac{Z - z}{h}\Big) > 0\Big) > 0, \tag{3.20}$$

and that the following class is Glivenko-Cantelli:

$$\mathcal{A} := \Big\{\Big\{(y,\widetilde{z}) \in \mathbb{R}^{d'} \times \mathbb{R}^d,\ \big(1_C(y) - m(C,h,z)\big)K\Big(\frac{\widetilde{z} - z}{h}\Big) > 0\Big\},\ C \in \mathcal{C},\ h > 0,\ z \in H\Big\}. \tag{3.21}$$

First note that $\mathcal{A}\subset\mathcal{B}$, where

$$\mathcal{B}:=\left\{\left\{(y,\widetilde{z})\in\mathbb{R}^{d'}\times\mathbb{R}^d,\ \big(1_C(y)-a\big)K\Big(\frac{\widetilde{z}-z}{h}\Big)>0\right\},\ C\in\mathcal{C},\ h>0,\ z\in\mathbb{R}^d,\ a\in\mathbb{R}\right\}.$$

By (HK1) and by Lemma 2.6.18 in [11], the two following classes of sets are VC:

$$\mathcal{B}_{\pm}:=\left\{\left\{\widetilde{z}\in\mathbb{R}^d,\ \pm K\big(h^{-d}(\widetilde{z}-z)\big)>0\right\},\ z\in\mathbb{R}^d,\ h>0\right\}.$$

Moreover, as $\mathcal{C}$ is a VC class of sets, we straightforwardly deduce that the following class is also VC:

$$\mathcal{M}_{\mathcal{G}}:=\left\{\{z\in\chi,\ 1_C(z)>a\},\ C\in\mathcal{C},\ a\in\mathbb{R}\right\}.$$

By a combination of points (i) and (ii) of Lemma 2.6.17 in [11], we conclude that $\mathcal{B}$ is VC, which entails that $\mathcal{A}$ is Glivenko-Cantelli. We now have to prove (3.20). Define the following family of random variables

$$\mathcal{H}_{\mathfrak{h}}:=\left\{\big(1_C(Y)-m(C,h,z)\big)K\Big(\frac{Z-z}{h}\Big),\ z\in H,\ C\in\mathcal{C},\ 0<h\leq\mathfrak{h}\right\}.$$

By the Cauchy-Schwarz inequality we have $\mathbb{P}(X>0)\geq\mathbb{E}(X^2)^{-1}\mathbb{E}(X1_{X>0})^2$. Hence it is sufficient to prove that, for $\mathfrak{h}$ small enough we have

$$\inf_{X\in\mathcal{H}_{\mathfrak{h}}}\mathbb{E}\Big(X1_{X>0}\Big)=\frac{1}{2}\inf_{X\in\mathcal{H}_{\mathfrak{h}}}\mathbb{E}\Big(\mid X\mid\Big)>0,\tag{3.22}$$

$$\sup_{X\in\mathcal{H}_{\mathfrak{h}}}\mathbb{E}\Big(X^2\Big)<\infty.\tag{3.23}$$

Note that the equality appearing in (3.22) is a consequence of $\mathbb{E}(X)=0$ for each $X\in\mathcal{H}_{\mathfrak{h}}$. By (HG'), (Hf) and (3.6), routine analysis shows that, for $\mathfrak{h}$ small enough, both (3.23) and the following assertion are true:

$$\inf_{X\in\mathcal{H}_h}\mathbb{E}\big(X^2\big)>\frac{1}{2}\inf_{z\in H,\ C\in\mathcal{C}}\sigma^2(C,z)f_Z(z)\mid\mid K\mid\mid_{\lambda,2}=:\alpha_0>0.\tag{3.24}$$

Now, as $\mathcal{H}_{\mathfrak{h}}$ is uniformly bounded by some constant $M>0$ we get that $\alpha_0\leq M\mathbb{E}\big(\mid X\mid\big)$ for all $X\in\mathcal{H}_{\mathfrak{h}}$, and hence (3.22) is proved. This concludes the proof of Lemma 3.2. □

3.3 Step 3: end of the proof of Theorem 3

Lemma 3.2 ensures us (see, e.g., [10], p. 219) that almost surely, for all large n and for each $z\in H,\ C\in\mathcal{C},\ h_n\leq h\leq\mathfrak{h}_n$, the maximum value in $\mathcal{R}_n\big(m(C,h,z),C,h,z\big)$ is obtained by choosing the following weights (recall (1.13))

$$p_i(C,h,z):=\frac{1}{n}\frac{1}{1+\lambda_n(C,h,z)w_{i,n}(C,h,z)},\tag{3.25}$$

where $\lambda_n(C,h,z)$ is the unique solution of

$$\sum_{i=1}^n\frac{w_{i,n}(C,h,z)}{1+\lambda_n(C,h,z)w_{i,n}(C,h,z)}=0.\tag{3.26}$$

Our next lemma gives an asymptotic control of

$$\sup_{\substack{C\in\mathcal{C},z\in H,\\ h_n\leq h\leq\mathfrak{h}_n}}\mid\lambda_n(C,h,z)\mid.$$

It is largely inspired by Lemma 1 in [3].

Lemma 3.3 *Under the assumptions of Theorem 2 we have almost surely*

$$\sup_{\substack{C\in\mathcal{C},\, z\in H,\\ h_n\le h\le \mathfrak{h}_n}} \mid \lambda_n(C,h,z) \mid = O\Big(\sqrt{\frac{\log(h^{-d})}{nh^d}}\Big). \tag{3.27}$$

Proof : Following the proof of Owen (see [10], p. 220), Lemma 3.3 will be proved if we check the following three conditions:

$$\max_{1\le i\le n} \sup_{\substack{z\in H,\, C\in\mathcal{C},\\ h_n\le h\le \mathfrak{h}_n}} \sqrt{\frac{\log(h^{-d})}{nh^d}} \mid w_{i,n}(C,h,z) \mid = o_{a.s.}(1), \tag{3.28}$$

$$\sup_{\substack{z\in H,\, C\in\mathcal{C},\\ h_n\le h\le \mathfrak{h}_n}} \frac{\mid X_n(C,h,z) \mid}{\sqrt{nh^d \log(h^{-d})}} = O_{a.s.}(1), \tag{3.29}$$

$$\liminf_{n\to\infty} \inf_{\substack{z\in H,\, C\in\mathcal{C},\\ h_n\le h\le \mathfrak{h}_n}} \frac{S_n(C,h,z)}{nh^d} > 0 \ a.s. \tag{3.30}$$

As each $w_{i,n}(C,h,z)$ is almost surely bounded by $2 \mid\mid K \mid\mid_{\mathbb{R}^d}$, and by (3.17), condition (3.28) is readily satisfied. Now note that condition (3.29) is a straightforward consequence of Theorem 2, and that (3.30) is a consequence of Lemma 3.1 and (HG'). The remainder of the proof of Lemma 3.3 is done by following Owen (see [10], p. 220).□

Now set

$$V_{i,n}(C,h,z) := \lambda_n(C,h,z) w_{i,n}(C,h,z).$$

By Lemma 3.3, we have

$$\lim_{n\to\infty} \max_{1\le i\le n} \sup_{z\in H,\, C\in\mathcal{C}} \mid V_{i,n}(C,z) \mid = 0_{a.s.}, \tag{3.31}$$

which entails, almost surely, for all large n and for each $z \in H,\ C \in \mathcal{C}$,

$$\begin{aligned}
0 =& \sum_{i=1}^{n} \frac{w_{i,n}(C,h,z)}{1+V_{i,n}(C,h,z)}\\
=& \sum_{i=1}^{n} w_{i,n}(C,h,z)\Big(1 - V_{i,n}(C,h,z) + V_{i,n}^2(C,h,z)/(1+V_{i,n}(C,h,z))\Big)\\
=& X_n(C,h,z) - f_Z(z) S_n(C,h,z)\lambda_n(C,h,z) + \sum_{i=1}^{n} \frac{w_{i,n}(C,h,z)V_{i,n}^2(C,h,z)}{1+V_{i,n}(C,h,z)}\\
=& X_n(C,h,z) - f_Z(z) S_n(C,h,z)\lambda_n(C,h,z) + \sum_{i=1}^{n} \frac{w_{i,n}^3(C,h,z)}{1+V_{i,n}(C,h,z)}\lambda_n^2(C,h,z).
\end{aligned} \tag{3.32}$$

From (3.28), (3.29) and (3.30), we conclude that there exists a random sequence ϵ_n such that, almost surely, we have $\epsilon_n \to 0$ and

$$\begin{aligned}
\sum_{i=1}^{n} \frac{w_{i,n}^3(C,h,z)}{1+V_{i,n}(C,h,z)}\lambda_n^2(C,h,z) \le& X_n(C,h,z) \max_i w_{i,n}^2(C,h,z)\\
&\times \Big(\min_{1\le i\le n} \mid 1+V_{i,n}(C,h,z) \mid\Big)^{-1} \lambda_n^2(C,h,z)\\
\le& \epsilon_n \sqrt{nh^d \log(h^{-d})},
\end{aligned} \tag{3.33}$$

uniformly in $C \in \mathcal{C}, z \in H,\ h \in [h_n, \mathfrak{h}_n]$. Hence, dividing the right hand side of (3.32) by $S_n(C, h, z)$, recalling (3.30) and (3.17), we obtain with probability one

$$\lambda_n(C, h, z) = \frac{X_n(C, h, z)}{f_Z(z) S_n(C, h, z)} + \beta_n(C, h, z), \tag{3.34}$$

with $\beta_n(C, h, z) \leq M\epsilon_n \sqrt{\log(h^{-d})/nh^d}$ uniformly in $C \in \mathcal{C},\ z \in H$ and $h \in [h_n, \mathfrak{h}_n]$, for some almost surely finite random variable M. We can now conclude that (recall (3.4))

$$\lim_{n\to\infty} \sup_{\substack{z\in H,\ C\in\mathcal{C},\\ h\in[h_n,\mathfrak{h}_n]}} \left| \frac{-2\log\Big(\mathcal{R}_n\big(g, z, m(C, h, z)\big)\Big)}{U_n(C, h, z)} - 1 \right| = 0, \tag{3.35}$$

by reasoning as in Owen [10], p.221. The proof of Theorem 3 is then concluded by (3.10). □

References

[1] P.J. Bickel and M Rosenblatt. On some global measures of the deviations of density function estimates. *Ann. Statist.*, 1:1071–1095, 1973.

[2] O. Bousquet. A Bennett concentration inequality and its application to suprema of empirical processes. *C.R. Acad. Sci. Paris, Ser. I*, 334:495–500, 2002.

[3] S.X. Chen, W. Härdle, and M. Li. An empirical likelihood goodness-of-fit test for time series. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 65(3):663–678, 2003.

[4] U. Einmahl and D.M. Mason. Some universal results on the behavior of the increments of partial sums. *Ann. Prob.*, 24:1388–1407, 1996.

[5] U. Einmahl and D.M. Mason. An empirical process approach to the uniform consistency of kernel type estimators. *J. Theoret. Probab.*, 13:1–13, 2000.

[6] U. Einmahl and D.M. Mason. Uniform in bandwidth consistency of variable bandwidth kernel estimators. *Preprint*, 2003.

[7] T. Klein. Une inégalité de concentration à gauche pour les processus empiriques. *C.R. Acad. Sci. Paris, Ser. I*, 2002.

[8] D.M. Mason. A uniform functional law of the iterated logarithm for the local empirical process. *Ann. Probab.*, 32(2):1391–1418, 2004.

[9] P. Massart. Strong approximations for multivariate empirical and related processes. *Ann. Probab.*, 17:266–291, 1989.

[10] A.B. Owen. *Empirical Likelihood.* Chapman and Hall/CRC, 2001.

[11] A.W. Van der Vaart and J.A. Wellner. *Weak convergence and empirical processes.* Springer, 1996.

[12] D. Varron. A limited in bandwidth uniformity for the functional limit law for the increments of the empirical process. *Discussion Paper DP0602, Université Cathiliquede Louvain*, 2006.