Alea 6, 63–114 (2009)

Revisiting Révész's stochastic approximation method for the estimation of a regression function

Abdelkader Mokkadem, Mariane Pelletier and Yousri Slaoui

Université de Versailles-Saint-Quentin, Département de Mathématiques, 45, Avenue des Etats-Unis, 78035 Versailles Cedex, France
E-mail address: mokkadem@math.uvsq.fr, pelletier@math.uvsq.fr, slaoui@math.uvsq.fr

Abstract. In a pioneer work, Révész (1973) introduces the stochastic approximation method to build up a recursive kernel estimator of the regression function $x \mapsto \mathbb{E}(Y|X = x)$. However, according to Révész (1977), his estimator has two main drawbacks: on the one hand, its convergence rate is smaller than that of the nonrecursive Nadaraya-Watson's kernel regression estimator, and, on the other hand, the required assumptions on the density of the random variable X are stronger than those usually needed in the framework of regression estimation. We first come back on the study of the convergence rate of Révész's estimator. An approach in the proofs completely different from that used in Révész (1977) allows us to show that Révész's recursive estimator may reach the same optimal convergence rate as Nadaraya-Watson's estimator, but the required assumptions on the density of X remain stronger than the usual ones, and this is inherent to the definition of Révész's estimator. To overcome this drawback, we introduce the averaging principle of stochastic approximation algorithms to construct the averaged Révész's regression estimator, and give its asymptotic behaviour. Our assumptions on the density of X are then usual in the framework of regression estimation. We prove that the averaged Révész's regression estimator may reach the same optimal convergence rate as Nadaraya-Watson's estimator. Moreover, we show that, according to the estimation by confidence intervals point of view, it is better to use the averaged Révész's estimator rather than Nadaraya-Watson's estimator.

1. Introduction

The use of stochastic approximation algorithms in the framework of regression estimation was introduced by Kiefer and Wolfowitz (1952). It allows the construction of online estimators. The great advantage of recursive estimators on nonrecursive

Received by the editors December 6, 2007; accepted January 30, 2009.
2000 *Mathematics Subject Classification.* 62G08; 62L20.
Key words and phrases. Regression estimation, stochastic approximation algorithm, averaging principle.

ones is that their update, from a sample of size n to one of size $n+1$, requires considerably less computations. This property is particularly important in the framework of regression estimation, since the number of points at which the function is estimated is usually very large. The famous Kiefer and Wolfowitz algorithm allows the approximation of the point at which a regression function reaches its maximum. This pioneer work was widely discussed and extended in many directions (see, among many others, Blum, 1954; Fabian, 1967; Kushner and Clark, 1978; Hall and Heyde, 1980, Ruppert, 1982, Chen, 1988, Spall, 1988, Polyak and Tsybakov, 1990; Dippon and Renz, 1997; Spall, 1997; Chen et al., 1999; Dippon, 2003 and Mokkadem and Pelletier, 2007). The question of applying Robbins-Monro's procedure to construct a stochastic approximation algorithm, which allows the estimation of a regression function at a given point (instead of approximating its mode) was introduced by Révész (1973).

Let us recall that Robbins-Monro's procedure consists in building up stochastic approximation algorithms, which allow the search of the zero z^* of an unknown function $h\ :\mathbb{R}\to\mathbb{R}$. These algorithms are constructed in the following way : (i) $Z_0\in\mathbb{R}$ is arbitrarily chosen; (ii) the sequence (Z_n) is recursively defined by setting

$$Z_n = Z_{n-1} + \gamma_n \mathcal{W}_n, \tag{1.1}$$

where $\mathcal{W}_n$ is an observation of the function h at the point Z_{n-1}, and where the stepsize (γ_n) is a sequence of positive real numbers that goes to zero.

Let $(X,Y),(X_1,Y_1),\ldots,(X_n,Y_n)$ be independent, identically distributed pairs of random variables, and let f denote the probability density of X. In order to construct a stochastic algorithm for the estimation of the regression function $r: x\mapsto \mathbb{E}(Y|X=x)$ at a point x such that $f(x)\neq 0$, Révész (1973) defines an algorithm, which approximates the zero of the function $h\ : y\mapsto f(x)r(x)-f(x)y$. Following Robbins-Monro's procedure, this algorithm is defined by setting $r_0(x)\in\mathbb{R}$ and, for $n\geq 1$,

$$r_n(x) = r_{n-1}(x) + \frac{1}{n}\mathcal{W}_n(x), \tag{1.2}$$

where $\mathcal{W}_n(x)$ is an "observation" of the function h at the point $r_{n-1}(x)$. To define $\mathcal{W}_n(x)$, Révész (1973) introduces a kernel K (that is, a function satisfying $\int_{\mathbb{R}} K(x)dx = 1$) and a bandwidth (h_n) (that is, a sequence of positive real numbers that goes to zero), and sets

$$\mathcal{W}_n(x) = h_n^{-1} Y_n K(h_n^{-1}[x-X_n]) - h_n^{-1} K(h_n^{-1}[x-X_n]) r_{n-1}(x). \tag{1.3}$$

Révész (1977) chooses the bandwidth (h_n) equal to (n^{-a}) with $a\in]1/2,1[$, and establishes a central limit theorem for $r_n(x)-r(x)$ under the assumption $f(x)>(1-a)/2$, as well as an upper bound of the uniform strong convergence rate of r_n on any bounded interval I on which $\inf_{x\in I} f(x) > (1-a)/2$. The two drawbacks of his approach are the following. First, the condition $a>1/2$ on the bandwidth leads to a pointwise weak convergence rate of r_n slower than $n^{1/4}$, whereas the optimal pointwise weak convergence rate of the kernel estimator of a regression function introduced by Nadaraya (1964) and Watson (1964) is $n^{2/5}$ (and obtained by choosing $a=1/5$). Then, the condition $f(x)>(1-a)/2$ (or $\inf_{x\in I} f(x) > (1-a)/2$) is stronger than the condition $f(x)>0$ (or $\inf_{x\in I} f(x)>0$) usually required to establish the convergence rate of regression's estimators.

Our first aim in this paper is to come back on the study of the asymptotic behaviour of Révész's estimator. The technic we use is totally different from that employed by Révész (1977). Noting that the estimator r_n can be rewritten as

$$\begin{aligned} r_n(x) &= \left(1 - \frac{1}{n}h_n^{-1}K\left(\frac{x-X_n}{h_n}\right)\right) r_{n-1}(x) + \frac{1}{n}h_n^{-1}Y_nK\left(\frac{x-X_n}{h_n}\right) \\ &= \left(1 - \frac{1}{n}f(x)\right) r_{n-1}(x) + \frac{1}{n}\left[f(x) - h_n^{-1}K\left(\frac{x-X_n}{h_n}\right)\right] r_{n-1}(x) \\ &\quad + \frac{1}{n}h_n^{-1}Y_nK\left(\frac{x-X_n}{h_n}\right), \end{aligned}$$

we approximate the sequence (r_n) by the unobservable sequence (ρ_n) recursively defined by

$$\begin{aligned} \rho_n(x) &= \left(1 - \frac{1}{n}f(x)\right) \rho_{n-1}(x) + \frac{1}{n}\left[f(x) - h_n^{-1}K\left(\frac{x-X_n}{h_n}\right)\right] r(x) \\ &\quad + \frac{h_n^{-1}}{n}Y_nK\left(\frac{x-X_n}{h_n}\right). \end{aligned} \tag{1.4}$$

The asymptotic properties (pointwise weak and strong convergence rate, upper bound of the uniform strong convergence rate) of the approximating algorithm (1.4) are established by applying different results on the sums of independent variables and on the martingales. To show that the asymptotic properties of the approximating algorithm (1.4) are also satisfied by Révész's estimator, we use a technic of successive upper bounds. It turns out that our technique of demonstration allows the choice of the bandwidth $(h_n) = (n^{-1/5})$, which makes Révész's estimator converge with the optimal pointwise weak convergence rate $n^{2/5}$. However, to establish the asymptotic convergence rate of Révész's estimator, we need the same kind of conditions on the marginal density of X as Révész (1977) does. To understand why this second drawback is inherent in the definition of Révész's estimator, let us come back on the algorithm (1.1). The convergence rate of stochastic approximation algorithms constructed following Robbins-Monro's scheme, and used for the search of the zero z^* of an unknown function h, was widely studied. It is now well known (see, among many others, Nevelsón and Has'minskii, 1976; Kushner and Clark, 1978; Ljung et al., 1992 and Duflo, 1996) that the convergence rate of algorithms defined as (1.1) is obtained under the condition that the limit of the sequence $(n\gamma_n)$ as n goes to infinity is larger than a quantity, which involves the differential of h at the point z^*. Now, let us recall that the Révész's estimator (1.2) is an algorithm approximating the zero $y^* = r(x)$ of the function $y \mapsto f(x)r(x) - f(x)y$ (whose differential at the point y^* equals $-f(x)$), and let us enlighten that the stepsize used to define his algorithm is $(\gamma_n) = (n^{-1})$ (so that $\lim_{n\to\infty} n\gamma_n = 1$); the condition on the limit of $(n\gamma_n)$, which is usual in the framework of stochastic approximation algorithms, comes down, in the case of Révész's estimator, to a condition on the probability density f.

Our second aim in this paper is to introduce the averaging principle of stochastic approximation algorithms in the framework of regression estimation. As a matter of fact, in the framework of stochastic approximation, this principle is now well known to allow to get rid of tedious conditions on the stepsize. It was independently introduced by Ruppert (1988) and Polyak (1990), and then widely discussed and extended (see, among many others, Yin, 1991; Delyon and Juditsky, 1992; Polyak and Juditsky, 1992; Kushner and Yang, 1993; Le Breton, 1993; Le Breton and Novikov, 1995; Dippon and Renz, 1996, 1997 and Pelletier, 2000). This procedure consists in: (i) running the approximation algorithm by using slower stepsizes; (ii) computing an average of the approximations obtained in (i). We thus need to generalize the definition of Révész's estimator before defining the averaged Révész's estimator.

Let (γ_n) be a sequence of positive numbers going to zero. The generalized Révész's estimator is defined by setting $r_0(x) \in \mathbb{R}$, and, for $n \geq 1$,

$$r_n(x) = r_{n-1}(x) + \gamma_n \mathcal{W}_n(x), \tag{1.5}$$

where $\mathcal{W}_n(x)$ is defined in (1.3). (Révész's estimator clearly corresponds to the choice of the stepsize $(\gamma_n) = (n^{-1})$). The estimator (1.5) with (γ_n) not necessary equal to (n^{-1}) was introduced in Györfi et al. (2002), where the strong universal convergence rate of $r_n(x)$ is also proved. Now, let the stepsize in (1.5) satisfy $\lim_{n\to\infty} n\gamma_n = \infty$, and let (q_n) be a positive sequence such that $\sum q_n = \infty$. The averaged Révész's estimator is defined by setting

$$\overline{r}_n(x) = \frac{1}{\sum_{k=1}^n q_k} \sum_{k=1}^n q_k r_k(x) \tag{1.6}$$

(where the $r_k(x)$ are given by the algorithm (1.5)). Let us enlighten that the estimator $\overline{r}_n$ is still recursive.

We establish the asymptotic behaviour (pointwise weak and strong convergence rate, upper bound of the uniform strong convergence rate) of $\overline{r}_n$. The condition we require on the density f to prove the pointwise (respectively uniform) convergence rate of the averaged Révész's estimator is the usual condition $f(x) > 0$ (respectively $\inf_{x\in I} f(x) > 0$). Concerning the bandwidth, the choice $(h_n) = (n^{-1/5})$, which leads to the optimal convergence rate $n^{2/5}$, is allowed. Finally, we show that to construct confidence intervals by slightly undersmoothing, it is preferable to use the averaged Révész's estimator $\overline{r}_n$ (with an adequate choice of weights (q_n)) rather than Nadaraya-Watson's estimator, since the asymptotic variance of this lattest estimator is larger than that of $\overline{r}_n$.

Our paper is organized as follows. Our assumptions and main results are stated in Section 2, simulations study is performed in Section 3, the outlines of the proofs given in Section 4, whereas Section 5 is devoted to the proof of several lemmas.

2. Assumptions and main results

Let us first define the class of positive sequences that will be used in the statement of our assumptions.

Definition 2.1. Let $\gamma \in \mathbb{R}$ and $(v_n)_{n\geq 1}$ be a nonrandom positive sequence. We say that $(v_n) \in \mathcal{GS}(\gamma)$ if

$$\lim_{n\to\infty} n \left[1 - \frac{v_{n-1}}{v_n}\right] = \gamma. \tag{2.1}$$

Condition (2.1) was introduced by Galambos and Seneta (1973) to define regularly varying sequences (see alsoBojanic and Seneta, 1973); it was used in Mokkadem and Pelletier (2007) in the context of stochastic approximation algorithms. Typical sequences in $\mathcal{GS}(\gamma)$ are, for $b \in \mathbb{R}$, $n^\gamma (\log n)^b$, $n^\gamma (\log\log n)^b$, and so on.

Let $g(s,t)$ denote the density of the couple (X,Y) (in particular $f(x) = \int_{\mathbb{R}} g(x,t)\, dt$), and set $a(x) = r(x) f(x)$. The assumptions to which we shall refer for our pointwise results are the following.

(A1) $K : \mathbb{R} \to \mathbb{R}$ is a nonnegative, continuous, bounded function satisfying $\int_{\mathbb{R}} K(z)\,dz = 1$, $\int_{\mathbb{R}} zK(z)\,dz = 0$ and $\int_{\mathbb{R}} z^2K(z)\,dz < \infty$.

(A2) *i)* $(\gamma_n) \in \mathcal{GS}(-\alpha)$ with $\alpha \in \left]\frac{3}{4}, 1\right]$; moreover the limit of $(n\gamma_n)^{-1}$ as n goes to infinity exists.
ii) $(h_n) \in \mathcal{GS}(-a)$ with $a \in \left]\frac{1-\alpha}{4}, \frac{\alpha}{3}\right[$.

(A3) *i)* $g(s,t)$ is two times continuously differentiable with respect to s.
ii) For $q \in \{0,1,2\}$, $s \mapsto \int_{\mathbb{R}} t^q g(s,t)\,dt$ is a bounded function continuous at $s = x$.
For $q \in [2,3]$, $s \mapsto \int_{\mathbb{R}} |t|^q g(s,t)\,dt$ is a bounded function.
iii) For $q \in \{0,1\}$, $\int_{\mathbb{R}} |t|^q \left|\frac{\partial g}{\partial x}(x,t)\right| dt < \infty$, and $s \mapsto \int_{\mathbb{R}} t^q \frac{\partial^2 g}{\partial s^2}(s,t)\,dt$ is a bounded function continuous at $s = x$.

Remark 2.2. (A3) *ii)* says in particular that f is continuous at x.

For our uniform results, we shall also need the following additional assumption.

(A4) *i)*K is Lipschitz-continuous.
ii) There exists $t^* > 0$ such that $\mathbb{E}(\exp(t^*|Y|)) < \infty$.
iii) $a \in]1-\alpha, \alpha - 2/3[$.
iv) For $q \in \{0,1\}$, $x \mapsto \int_{\mathbb{R}} |t|^q \left|\frac{\partial g}{\partial x}(x,t)\right| dt$ is bounded on the set $\{x, f(x) > 0\}$.

Throughout this paper we shall use the following notation :

$$\xi = \lim_{n\to\infty} (n\gamma_n)^{-1}, \tag{2.2}$$

and, for $f(x) \neq 0$,

$$m^{(2)}(x) = \frac{1}{2f(x)}\left[\int_{\mathbb{R}} t\frac{\partial^2 g}{\partial x^2}(x,t)\,dt - r(x)\int_{\mathbb{R}} \frac{\partial^2 g}{\partial x^2}(x,t)\,dt\right]\int_{\mathbb{R}} z^2K(z)\,dz.$$

The asymptotic properties of the generalized Révész's estimator defined in (1.5) are stated in Section 2.1, those of the averaged estimator defined in (1.6) in Section 2.2.

2.1. *Asymptotic behaviour of the generalized Révész's estimator.* For stepsizes satisfying $\lim_{n\to\infty} n\gamma_n = \infty$, the strong universal consistency of the generalized Révész's estimator was established by Walk (2001) and Györfi et al. (2002). The aim of this section is to state the convergence rate of the estimator defined by (1.5). Theorems 2.3, 2.4, and 2.5 below give its weak pointwise convergence rate, its strong pointwise convergence rate, and an upper bound of its strong uniform convergence rate, respectively. Let us enlighten that the particular choice of stepsize $(\gamma_n) = (n^{-1})$ gives the asymptotic behaviour of Révész's estimator defined in (1.2).

Theorem 2.3 (Weak pointwise convergence rate of r_n)**.**
Let Assumptions (A1) − (A3) hold for $x \in \mathbb{R}$ such that $f(x) \neq 0$.

(1) *If there exists $c \geq 0$ such that $\gamma_n^{-1}h_n^5 \to c$, and if $\lim_{n\to\infty}(n\gamma_n) > (1 - a)/(2f(x))$, then*

$$\sqrt{\gamma_n^{-1}h_n}\,(r_n(x) - r(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\sqrt{c}f(x)\,m^{(2)}(x)}{f(x) - 2a\xi}, \frac{Var[Y|X=x]\,f(x)}{(2f(x) - (1-a)\xi)}\int_{\mathbb{R}} K^2(z)\,dz\right).$$

(2) *If* $\gamma_n^{-1}h_n^5 \to \infty$, *and if* $\lim_{n\to\infty}(n\gamma_n) > 2a/f(x)$, *then*

$$\frac{1}{h_n^2}(r_n(x) - r(x)) \xrightarrow{\mathbb{P}} \frac{f(x)\, m^{(2)}(x)}{(f(x) - 2a\xi)},$$

where $\xrightarrow{\mathcal{D}}$ *denotes the convergence in distribution,* $\mathcal{N}$ *the Gaussian-distribution and* $\xrightarrow{\mathbb{P}}$ *the convergence in probability.*

The combination of Parts 1 and 2 of Theorem 2.3 ensures that the optimal weak pointwise convergence rate of r_n equals $n^{2/5}$, and is obtained by choosing $a = 1/5$ and (γ_n) such that $\lim_{n\to\infty}(n\gamma_n) \in]2/(5f(x)), \infty[$.

Theorem 2.4 (Strong pointwise convergence rate of r_n)**.**
Let Assumptions $(A1) - (A3)$ *hold for* $x \in \mathbb{R}$ *such that* $f(x) \neq 0$.

(1) *If there exists* $c \geq 0$ *such that* $\gamma_n^{-1}h_n^5/\ln\left(\sum_{k=1}^n \gamma_k\right) \to c$, *and if*

$$\lim_{n\to\infty}(n\gamma_n) > (1-a)/(2f(x)),$$

then, with probability one, the sequence

$$\left(\sqrt{\frac{\gamma_n^{-1}h_n}{2\ln\left(\sum_{k=1}^n \gamma_k\right)}}(r_n(x) - r(x))\right)$$

is relatively compact and its limit set is the interval

$$\left[\sqrt{\frac{c}{2}}\frac{f(x)\, m^{(2)}(x)}{f(x) - 2a\xi} - \sqrt{\frac{Var[Y|X=x]\, f(x)\int_{\mathbb{R}} K^2(z)\, dz}{(2f(x) - (1-a)\xi)}},\right.$$
$$\left.\sqrt{\frac{c}{2}}\frac{f(x)\, m^{(2)}(x)}{f(x) - 2a\xi} + \sqrt{\frac{Var[Y|X=x]\, f(x)\int_{\mathbb{R}} K^2(z)\, dz}{(2f(x) - (1-a)\xi)}}\right].$$

(2) *If* $\gamma_n^{-1}h_n^5/\ln\left(\sum_{k=1}^n \gamma_k\right) \to \infty$, *and if* $\lim_{n\to\infty}(n\gamma_n) > 2a/f(x)$, *then, with probability one,*

$$\lim_{n\to\infty}\frac{1}{h_n^2}(r_n(x) - r(x)) = \frac{f(x)\, m^{(2)}(x)}{f(x) - 2a\xi}.$$

Theorem 2.5 (Strong uniform convergence rate of r_n)**.**
Let I *be a bounded open interval of* $\mathbb{R}$ *on which* $\varphi = \inf_{x\in I} f(x) > 0$, *and let Assumptions* $(A1) - (A4)$ *hold for all* $x \in I$.

(1) *If the sequence* $\left(\gamma_n^{-1}h_n^5/[\ln n]^2\right)$ *is bounded and if* $\lim_{n\to\infty}(n\gamma_n) > 1 - a/(2\varphi)$, *then*

$$\sup_{x\in I}|r_n(x) - r(x)| = O\left(\sqrt{\gamma_n h_n^{-1}}\ln n\right) \quad a.s.$$

(2) *If* $\lim_{n\to\infty}\left(\gamma_n^{-1}h_n^5/[\ln n]^2\right) = \infty$ *and if* $\lim_{n\to\infty}(n\gamma_n) > 2a/\varphi$, *then*

$$\sup_{x\in I}|r_n(x) - r(x)| = O\left(h_n^2\right) \quad a.s.$$

Parts 1 of Theorems 2.3 and 2.5 were obtained by Révész (1977) for the choices $(\gamma_n) = (n^{-1})$ and $(h_n) = (n^{-a})$ with $a \in]1/2, 1[$. Let us underline that, for this choice of stepsize, the conditions $\lim_{n\to\infty}(n\gamma_n) > (1-a)/(2f(x))$ and $\lim_{n\to\infty}(n\gamma_n) > (1-a)/(2\inf_{x\in I} f(x))$ come down to the following conditions on the unknown density f: $f(x) > (1-a)/2$ and $\inf_{x\in I} f(x) > (1-a)/2$. Let us also mention that our assumption (A2) implies $a \in]0, 1/3[$, so that our results on the generalized Révész's estimator do not include the results given in Révész (1977). However, our assumptions include the choice $(\gamma_n) = (n^{-1})$ and $a = 1/5$, which leads to the optimal weak convergence rate $n^{2/5}$, whereas the condition on the bandwidth required by Révész leads to a convergence rate of r_n slower than $n^{1/4}$.

Although the optimal convergence rate we obtain for the generalized Révész's estimator r_n has the same order as that of Nadaraya-Watson's estimator, this estimator has a main drawback: to make r_n converge with its optimal rate, one must set $a = 1/5$ and choose (γ_n) such that $\lim_{n\to\infty} n\gamma_n = \gamma^* \in]0, \infty[$ with $\gamma^* > 2/[5f(x)]$ whereas the density f is unknown. This tedious condition disappears as soon as the stepsize is chosen such that $\lim_{n\to\infty} n\gamma_n = \infty$ (for instance when $(\gamma_n) = ((\ln n)^b n^{-1})$ with $b > 0$), but the optimal convergence rate $n^{2/5}$ is then not reached any more.

2.2. *Asymptotic behaviour of the averaged Révész's estimator.* To state the asymptotic properties of the averaged Révész's estimator defined in (1.6), we need the following additional assumptions.

(A5) $\lim_{n\to\infty} n\gamma_n \left(\ln\left(\sum_{k=1}^n \gamma_k\right)\right)^{-1} = \infty$ and $a \in \left]1-\alpha, (4\alpha-3)/2\right[$.
(A6) $(q_n) \in \mathcal{GS}(-q)$ with $q < \min\{1-2a, (1+a)/2\}$.

Theorems 2.6, 2.7 and 2.8 below give the weak pointwise convergence rate, the strong pointwise convergence rate, and an upper bound of the strong uniform convergence rate of the averaged Révész's estimator.

Theorem 2.6 (Weak pointwise convergence rate of $\bar{r}_n$).
Let Assumptions (A1) − (A3), (A5) and (A6) hold for $x \in \mathbb{R}$ such that $f(x) \neq 0$.

(1) *If there exists $c \geq 0$ such that $nh_n^5 \to c$, then*

$$\sqrt{nh_n}\left(\bar{r}_n(x) - r(x)\right) \stackrel{\mathcal{D}}{\rightarrow} \mathcal{N}\left(c^{\frac{1}{2}}\frac{1-q}{1-q-2a}m^{(2)}(x), \frac{(1-q)^2}{1+a-2q}\frac{Var[Y|X=x]}{f(x)}\int_{\mathbb{R}} K^2(z)\,dz\right).$$

(2) *If $nh_n^5 \to \infty$, then*

$$\frac{1}{h_n^2}\left(\bar{r}_n(x) - r(x)\right) \stackrel{\mathbb{P}}{\rightarrow} \frac{1-q}{1-q-2a}m^{(2)}(x).$$

The combination of Parts 1 and 2 of Theorem 2.6 ensures that the optimal weak pointwise convergence rate of $\bar{r}_n$ is obtained by choosing $a = 1/5$, and equals $n^{2/5}$.

Theorem 2.7 (Strong pointwise convergence rate of $\bar{r}_n$).
Let Assumptions (A1) − (A3), (A5) and (A6) hold for $x \in \mathbb{R}$ such that $f(x) \neq 0$.

(1) *If there exists $c_1 \geq 0$ such that $nh_n^5/\ln\ln n \to c_1$, then, with probability one, the sequence*

$$\left(\sqrt{\frac{nh_n}{2\ln\ln n}}\left(\bar{r}_n(x) - r(x)\right)\right)$$

is relatively compact and its limit set is the interval

$$\left[c_1^{\frac{1}{2}}\frac{1-q}{1-q-2a}m^{(2)}(x)-\sqrt{\frac{(1-q)^2}{1+a-2q}\frac{Var[Y/X=x]}{f(x)}\int_{\mathbb{R}}K^2(z)\,dz},\right.$$

$$\left.c_1^{\frac{1}{2}}\frac{1-q}{1-q-2a}m^{(2)}(x)+\sqrt{\frac{(1-q)^2}{1+a-2q}\frac{Var[Y/X=x]}{f(x)}\int_{\mathbb{R}}K^2(z)\,dz}\right].$$

(2) *If* $nh_n^5/\ln\ln n \to \infty$*, then*

$$\lim_{n\to\infty}\frac{1}{h_n^2}\left(\bar{r}_n(x)-r(x)\right)=\frac{1-q}{1-q-2a}m^{(2)}(x) \quad a.s.$$

Theorem 2.8 (Strong uniform convergence rate of $\bar{r}_n$)**.**
Let I *be a bounded open interval of* $\mathbb{R}$ *on which* $\inf_{x\in I} f(x) > 0$*, and let Assumptions* $(A1)-(A6)$ *hold for all* $x \in I$*.*

(1) *If the sequence* $\left(nh_n^5/[\ln n]^2\right)$ *is bounded, and if* $\alpha > (3a+3)/4$*, then*

$$\sup_{x\in I}\left|\bar{r}_n(x)-r(x)\right| = O\left(\sqrt{n^{-1}h_n^{-1}}\ln n\right) \quad a.s.$$

(2) *If* $\lim_{n\to\infty}\left(nh_n^5/[\ln n]^2\right) = \infty$*, and if, in the case* $a \in [\alpha/5, 1/5]$*,* $\alpha > (4a+1)/2$*, then*

$$\sup_{x\in I}\left|\bar{r}_n(x)-r(x)\right| = O\left(h_n^2\right) \quad a.s.$$

Whatever the choices of the stepsize (γ_n) and of the weight (q_n) are, the convergence rate of the averaged Révész's estimator has the same order as that of the generalized Révész's estimator defined with a stepsize (γ_n) satisfying $\lim_{n\to\infty}(n\gamma_n)^{-1} \neq 0$ (and, in particular, of Révész's estimator). The main advantage of the averaged Révész's estimator on its nonaveraged version is that its convergence rate is obtained without tedious conditions on the marginal density f.
Now, to compare the averaged Révész's estimator with the nonrecursive Nadaraya-Watson's estimator defined as

$$\tilde{r}_n(x) = \frac{\sum_{k=1}^n Y_k K\left(h_n^{-1}[x-X_k]\right)}{\sum_{k=1}^n K\left(h_n^{-1}[x-X_k]\right)},$$

let us consider the estimation by confidence intervals point of view. In the context of density estimation, Hall (1992) shows that, to construct confidence intervals, slightly undersmoothing is more efficient than bias estimation; in the framework of regression estimation, the method of undersmoothing to construct confidence regions is used in particular by Neumann and Polzehl (1998) and Claeskens and van Keilegom (2003). To undersmooth, we choose (h_n) such that $\lim_{n\to\infty} nh_n^5 = 0$ (and thus $a \geq 1/5$). Moreover, to construct a confidence interval for $r(x)$, it is advised to choose the weight (q_n), which minimizes the asymptotic variance of $\bar{r}_n$. For a given a, the function $q \mapsto (1-q)^2/(1+a-2q)$ reaching its minimum at the point $q = a$, we can state the following corollary.

Corollary 2.9.
Let Assumptions $(A1)-(A3)$*,* $(A5)$ *and* $(A6)$ *hold for* $x \in \mathbb{R}$ *such that* $f(x) \neq 0$*,*

and with $a \geq 1/5$. To minimize the asymptotic variance of $\bar{r}_n$, q must be chosen equal to a. Moreover, if $\lim_{n\to\infty} nh_n^5 = 0$, we then have

$$\sqrt{nh_n}\left(\bar{r}_n(x) - r(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, (1-a)\frac{Var[Y|X=x]}{f(x)}\int_{\mathbb{R}} K^2(z)\,dz\right).$$

Let us recall that, when the bandwidth (h_n) is chosen such that $\lim_{n\to\infty} nh_n^5 = 0$, Nadaraya-Watson's estimator satisfies the central limit theorem

$$\sqrt{nh_n}\left(\tilde{r}_n(x) - r(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{Var[Y|X=x]}{f(x)}\int_{\mathbb{R}} K^2(z)\,dz\right). \tag{2.3}$$

It turns out that the averaged Révész's estimator defined with a weight (q_n) belonging to $\mathcal{GS}(-a)$ has a smaller asymptotic variance than Nadaraya-Watson's estimator. According to the estimation by confidence intervals point of view, it is thus better to use the averaged Révész's estimator rather than Nadaraya-Watson's one. This superiority of the recursive averaged Révész's estimator on the classical nonrecursive Nadaraya-Watson's estimator must be related to that of recursive density estimators on the classical nonrecursive Rosenblatt's estimator, and can be explained more easily in the framework of density estimation: roughly speaking, Rosenblatt's estimator can be seen as the average of n independent random variables, which all share the same variance v_n, whereas recursive estimators appear as the average of n independent random variables whose variances v_k^*, $1 \leq k \leq n$, satisfy $v_k^* < v_n$ for all $k < n$ and $v_n^* = v_n$.

3. Simulations

The object of this section is to provide a simulations study comparing Nadaraya-Watson's estimator and the averaged Révész's estimator. We consider the regression model

$$Y = r(X) + d\varepsilon,$$

where $d > 0$ and ε is $\mathcal{N}(0,1)$-distributed. Whatever the estimator is, we choose the kernel $K(x) = (2\pi)^{-1/2}\exp\left(-x^2/2\right)$, and the bandwidth equal to $(h_n) = n^{-1/5}(\ln n)^{-1}$ (which corresponds to a slight undersmoothing). The confidence intervals of $r(x)$ we consider are the following.

- When Nadaraya-Watson's estimator $\tilde{r}_n$ is used, we set

$$\tilde{I}_n = \left[\tilde{r}_n(x) - 1.96\sqrt{\frac{\sum_{i=1}^n \left(Y_i - \tilde{r}_n(X_i)\right)^2}{n^2 h_n \tilde{f}_n(x)}\int_{\mathbb{R}} K^2(z)\,dz},\right.$$
$$\left.\tilde{r}_n(x) + 1.96\sqrt{\frac{\sum_{i=1}^n \left(Y_i - \tilde{r}_n(X_i)\right)^2}{n^2 h_n \tilde{f}_n(x)}\int_{\mathbb{R}} K^2(z)\,dz}\right],$$

  where $\tilde{f}_n(x) = (nh_n)^{-1}\sum_{k=1}^n K\left(h_n^{-1}[x - X_k]\right)$ is Rosenblatt's density estimator. In view of (2.3), the asymptotic confidence level of $\tilde{I}_n$ is 95%.
- To define the averaged Révész's estimator $\bar{r}_n$, we choose the weights (q_n) equal to (h_n). This choice guarantees that (h_n) and (q_n) both belong to $\mathcal{GS}(-a)$ (with $a = 1/5$ here), so that, in view of Corollary 2.9, the asymptotic variance of $\bar{r}_n$ is minimal. We also let $(\gamma_n) = \left(n^{-0.9}\right)$ (this choice

being allowed by our assumptions). Moreover, we estimate the density f by the recursive estimator $\hat{f}_n(x)$ defined as

$$\hat{f}_n(x) = (1-\beta_n)\hat{f}_{n-1}(x) + \beta_n h_n^{-1} K\left(\frac{x - X_n}{h_n}\right),$$

where $(\beta_n) = \left(\frac{4}{5}n^{-1}\right)$; this choice of the stepsize (β_n) is known to minimize the variance of $\hat{f}_n$ (see Mokkadem et al., 2009). Finally, replacing $\tilde{r}_n$ and $\tilde{f}_n$ by the recursive estimators $\bar{r}_n$ and $\hat{f}_n$ in the definition of $\tilde{I}_n$, we get the recursive confidence interval

$$\begin{aligned} \bar{I}_n \quad = \quad & \left[\bar{r}_n(x) - 1.96\sqrt{\frac{\sum_{i=1}^n \left(Y_i - \bar{r}_n(X_i)\right)^2}{n^2 h_n \hat{f}_n(x)} \int_{\mathbb{R}} K^2(z)\,dz},\right. \\ & \left. \bar{r}_n(x) + 1.96\sqrt{\frac{\sum_{i=1}^n \left(Y_i - \bar{r}_n(X_i)\right)^2}{n^2 h_n \hat{f}_n(x)} \int_{\mathbb{R}} K^2(z)\,dz}\right]. \end{aligned}$$

The widths of the intervals $\tilde{I}_n$ and $\bar{I}_n$ are of the same order, but the asymptotic level of $\bar{I}_n$ is larger than that of $\tilde{I}_n$. More precisely, let Φ denote the distribution function of the standard normal; the application of Corollary 2.9 ensures that the asymptotic level of $\bar{I}_n$ is $2\Phi\left(1.96/\sqrt{4/5}\right) - 1 = 97.14\%$.

We consider three sample sizes $n = 50$, $n = 100$ and $n = 200$, three regression functions $r(x) = \cos(x)$, $r(x) = 0.3\exp(-4(x+1)^2) + 0.7\exp(-16(x-1)^2)$, and $r(x) = 1 + 0.4x$, three points $x = -0.5$, $x = 0$ and $x = 0.5$, two values of d, $d = 1$ and $d = 2$, and three densities of X, standard normal, normal mixture and student with 6 degrees of freedom. In each case the number of simulations is $N = 5000$. In each table, the first line corresponds to the use of Nadaraya-Watson's estimator $\tilde{r}_n$ and gives the empirical levels $\#\{r(x) \in \tilde{I}_n\}/N$; the second line corresponds to the use of the averaged Révész's estimator $\bar{r}_n$ and gives the empirical levels $\#\{r(x) \in \bar{I}_n\}/N$.

The simulations results confirm the theoretical ones: the coverage error of the intervals built up with the averaged Révész's estimator is smaller than the coverage error of the intervals built up with Nadaraya-Watson's estimator.

Model $r(x) = \cos(x)$.

Distribution of X: $\mathcal{N}(0,1)$

	$x=-0.5$			$x=0$			$x=0.5$	
$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$
				$d=1$				
96.5%	96.76%	96.5%	96.44%	96.62%	96.84%	96.7%	97.04%	96.92%
99.82%	99.9%	99.92%	99.8%	99.68%	99.76%	99.94%	99.86%	99.88%
				$d=2$				
95.42%	95.32%	95.7%	94.94%	95.44%	95.08%	95.4%	95.44%	96.2%
99.82%	99.86%	99.76%	99.66%	99.6%	99.44%	99.82%	99.9%	99.98%

Model $r(x) = 0.3\exp(-4(x+1)^2) + 0.7\exp(-16(x-1)^2)$.

Distribution of X: $\mathcal{N}(0,1)$

	$x=-0.5$			$x=0$			$x=0.5$	
$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$
				$d=1$				
95.04%	94.74%	95.08%	95.06%	95.28%	95.4%	95.44%	95.44%	95.84%
99.8%	99.62%	99.46%	99.24%	99.34%	99.06%	99.34%	99.34%	99.12%
				$d=2$				
95.26%	95.14%	95.34%	94.74%	94.88%	95.06%	94.48%	95.56%	95.62%
99.86%	99.76%	99.72%	99.64%	99.52%	99.38%	99.62%	99.74%	99.6%

Model $r(x) = 1 + 0.4x$.

Distribution of X: $\mathcal{N}(0,1)$

$x=-0.5$			$x=0$			$x=0.5$		
$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$
				$d=1$				
96.32%	95.94%	96.1%	96.24%	96.2%	96%	96.1%	96.24%	96.62%
99.84%	99.9%	99.6%	99.92%	99.82%	99.72%	99.86%	99.8%	99.76%
				$d=2$				
95.46%	94.76%	95.16%	95.56%	95.38%	95.54%	94.98%	94.96%	95.62%
99.82%	99.88%	99.62%	99.88%	99.78%	99.68%	99.88%	99.82%	99.68%

Model $r(x) = \cos(x)$.

Distribution of X: $1/2\mathcal{N}(-1/2,1) + 1/2\mathcal{N}(1/2,1)$

$x=-0.5$			$x=0$			$x=0.5$		
$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$
				$d=1$				
96.96%	97.06%	97.12%	97.26%	96.8%	97.1%	97.46%	96.94%	96.94%
99.96%	99.92%	99.88%	99.86%	99.8%	99.66%	99.96%	99.96%	99.8%
				$d=2$				
95.6%	95.32%	95.56%	95.08%	95.36%	95.64%	96.38%	95.7%	95.34%
99.82%	99.92%	99.74%	99.94%	99.78%	99.64%	99.96%	99.9%	99.64%

Model $r(x) = 0.3\exp(-4(x+1)^2) + 0.7\exp(-16(x-1)^2)$.

Distribution of X: $1/2\mathcal{N}(-1/2,1) + 1/2\mathcal{N}(1/2,1)$

$x=-0.5$			$x=0$			$x=0.5$		
$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$
				$d=1$				
94.9%	95.38%	95.3%	95.56%	94.56%	94.86%	95.24%	95.24%	95.48%
99.74%	99.62%	99.58%	99.44%	99.22%	99.1%	99.34%	99.28%	99.06%
				$d=2$				
94.54%	95.34%	94.92%	95.2%	94.4%	94.82%	95.24%	95.06%	95.14%
99.82%	99.78%	99.74%	99.84%	99.74%	99.6%	99.8%	99.78%	99.58%

Model $r(x) = 1 + 0.4x$.

Distribution of X: $1/2\mathcal{N}(-1/2,1) + 1/2\mathcal{N}(1/2,1)$

$x=-0.5$			$x=0$			$x=0.5$		
$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$
				$d=1$				
96.32%	96.66%	96.84%	96.46%	96.74%	96.64%	96.6%	96.72%	97.2%
99.92%	99.88%	99.8%	99.94%	99.98%	99.84%	99.88%	99.9%	99.86%
				$d=2$				
95.18%	95.46%	96.1%	95.08%	95.52%	95.6%	95.58%	95.44%	95.74%
99.94%	99.86%	99.78%	99.88%	99.96%	99.7%	99.9%	99.86%	99.8%

Model $r(x) = \cos(x)$.

Distribution of X: $\mathcal{T}(6)$

$x=-0.5$			$x=0$			$x=0.5$		
$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$
				$d=1$				
96.98%	97.54%	97.64%	97.02%	97.28%	97.52%	97.6%	97.1%	96.98%
99.9%	99.84%	99.62%	99.74%	99.86%	99.88%	99.98%	99.9%	99.86%
				$d=2$				
95.6%	95.96%	95.94%	95.4%	95.84%	96.06%	96.26%	95.62%	95.24%
99.88%	99.78%	99.82%	99.74%	99.72%	99.8%	99.98%	99.82%	99.68%

Model $r(x) = 0.3\exp(-4(x+1)^2) + 0.7\exp(-16(x-1)^2)$.

Distribution of X: $\mathcal{T}(6)$

$x=-0.5$			$x=0$			$x=0.5$		
$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$
				$d=1$				
95.3%	94.88%	95.08%	95.5%	95.06%	95.02%	95.28%	95.48%	95.56%
99.8%	99.68%	99.46%	99.16%	99.26%	99.18%	99.4%	99.24%	99.18%
				$d=2$				
94.88%	94.5%	94.8%	95.28%	94.8%	94.64%	95.06%	95.3%	95.3%
99.84%	99.82%	99.58%	99.64%	99.66%	99.58%	99.8%	99.7%	99.7%

Model $r(x) = 1 + 0.4x$.

Distribution of X: $\mathcal{T}(6)$								
$x=-0.5$			$x=0$			$x=0.5$		
$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$	$n=50$	$n=100$	$n=200$
				$d=1$				
96.62%	97.04%	97%	97.2%	97.08%	97.02%	96.36%	97.14%	97.22%
99.84%	99.9%	99.92%	99.94%	99.88%	99.82%	99.86%	99.84%	99.86%
				$d=2$				
95.04%	95.62%	95.54%	95.96%	95.58%	95.88%	94.94%	96.14%	95.86%
99.82%	99.9%	99.82%	99.84%	99.78%	99.66%	99.86%	99.84%	99.76%

4. Outlines of the proofs

From now on, we set $n_0 \geq 3$ such that $\forall k \geq n_0$, $\gamma_k \leq (2\|f\|_\infty)^{-1}$ and $\gamma_k h_k^{-1}\|K\|_\infty \leq 1$. Moreover, we introduce the following notations:

$$\begin{aligned} s_n &= \sum_{k=n_0}^{n} \gamma_k, \\ Z_n(x) &= h_n^{-1} K\left(\frac{x-X_n}{h_n}\right), \\ W_n(x) &= h_n^{-1} Y_n K\left(\frac{x-X_n}{h_n}\right) \end{aligned}$$

and, for $s > 0$,

$$\begin{aligned} \Pi_n(s) &= \prod_{j=n_0}^{n} (1 - s\gamma_j), \\ U_{k,n}(s) &= \Pi_n(s)\,\Pi_k^{-1}(s). \end{aligned}$$

Finally, we define the sequences (m_n) and $(\tilde{m}_n)$ by setting

$$(m_n) = \begin{cases} \left(\sqrt{\gamma_n h_n^{-1}}\right) & \text{if } \lim_{n\to\infty}\left(\gamma_n h_n^{-5}\right) = \infty, \\ \left(h_n^2\right) & \text{otherwise.} \end{cases} \tag{4.1}$$

$$(\tilde{m}_n) = \begin{cases} \left(\sqrt{\gamma_n h_n^{-1}}\ln n\right) & \text{if } \lim_{n\to\infty}\left(\gamma_n h_n^{-5}\ln n\right) = \infty, \\ \left(h_n^2\right) & \text{otherwise.} \end{cases} \tag{4.2}$$

Note that the sequences (m_n) and $(\tilde{m}_n)$ belong to $\mathcal{GS}(-m^*)$ with

$$m^* = \min\left\{\frac{\alpha - a}{2}, 2a\right\}. \tag{4.3}$$

Before giving the outlines of the proofs, we state the following technical lemma, which is proved in Mokkadem et al. (2009), and which will be used throughout the demonstrations.

Lemma 4.1.
Let $(v_n) \in \mathcal{GS}(v^*)$, $(\gamma_n) \in \mathcal{GS}(-\alpha)$ with $\alpha > 0$, and set $m > 0$. If $ms - v^*\xi > 0$ (where ξ is defined in (2.2)), then

$$\lim_{n\to\infty} v_n \Pi_n^m(s) \sum_{k=n_0}^{n} \Pi_k^{-m}(s)\frac{\gamma_k}{v_k} = \frac{1}{ms - v^*\xi}.$$

Moreover, for all positive sequence (α_n) *such that* $\lim_{n\to\infty}\alpha_n = 0$*, and all* C*,*

$$\lim_{n\to\infty} v_n \Pi_n^m(s)\left[\sum_{k=n_0}^{n} \Pi_k^{-m}(s)\frac{\gamma_k}{v_k}\alpha_k + C\right] = 0.$$

As explained in the introduction, we note that the stochastic approximation algorithm (1.5) can be rewritten as:

$$\begin{aligned} r_n(x) &= (1-\gamma_n Z_n(x))\, r_{n-1}(x) + \gamma_n W_n(x) && (4.4)\\ &= (1-\gamma_n f(x))\, r_{n-1}(x) + \gamma_n (f(x) - Z_n(x))\, r_{n-1}(x) + \gamma_n W_n(x). && (4.5)\end{aligned}$$

To establish the asymptotic behaviour of (r_n) and $(\bar{r}_n)$, we introduce the auxiliary stochastic approximation algorithm defined by setting $\rho_n(x) = r(x)$ for all $n \leq n_0 - 2$, $\rho_{n_0-1}(x) = r_{n_0-1}(x)$, and, for $n \geq n_0$,

$$\rho_n(x) = (1-\gamma_n f(x))\, \rho_{n-1}(x) + \gamma_n (f(x) - Z_n(x))\, r(x) + \gamma_n W_n(x). \quad (4.6)$$

We first give the asymptotic behaviour of (ρ_n) and of $(\bar{\rho}_n)$ in Section 4.1 and 4.2 respectively (and refer to Section 5 for the proof of the different lemmas). Then, we show in Section 4.3 how the asymptotic behaviour of (r_n) and $(\bar{r}_n)$ can be deduced from that of (ρ_n) and $(\bar{\rho}_n)$ respectively.

4.1. *Asymptotic behaviour of* ρ_n. The aim of this section is to give the outlines of the proof of the three following lemmas.

Lemma 4.2 (Weak pointwise convergence rate of ρ_n)**.**
Let Assumptions $(A1)-(A3)$ *hold for* $x \in \mathbb{R}$ *such that* $f(x) \neq 0$*.*

(1) *If there exists* $c \geq 0$ *such that* $\gamma_n^{-1}h_n^5 \to c$*, and if* $\lim_{n\to\infty}(n\gamma_n) > (1-a)/(2f(x))$*, then*

$$\sqrt{\gamma_n^{-1}h_n}\,(\rho_n(x) - r(x)) \xrightarrow{\mathcal{D}} \mathcal{N}\left(\frac{\sqrt{c}f(x)\,m^{(2)}(x)}{f(x)-2a\xi}, \frac{Var[Y|X=x]\,f(x)}{(2f(x)-(\alpha-a)\,\xi)}\int_{\mathbb{R}} K^2(z)\,dz\right). \quad (4.7)$$

(2) *If* $\gamma_n^{-1}h_n^5 \to \infty$*, and if* $\lim_{n\to\infty}(n\gamma_n) > 2a/f(x)$*, then*

$$\frac{1}{h_n^2}(\rho_n(x) - r(x)) \xrightarrow{\mathbb{P}} \frac{f(x)\,m^{(2)}(x)}{(f(x)-2a\xi)}. \quad (4.8)$$

Lemma 4.3 (Strong pointwise convergence rate of ρ_n)**.**
Let Assumptions $(A1)-(A3)$ *hold for* $x \in \mathbb{R}$ *such that* $f(x) \neq 0$*.*

(1) *If there exists* $c \geq 0$ *such that* $\gamma_n^{-1}h_n^5/\ln(s_n) \to c$*, and if* $\lim_{n\to\infty}(n\gamma_n) > (1-a)/(2f(x))$*, then, with probability one, the sequence*

$$\left(\sqrt{\frac{\gamma_n^{-1}h_n}{2\ln(s_n)}}\,(\rho_n(x) - r(x))\right)$$

is relatively compact and its limit set is the interval

$$\left[\sqrt{\frac{c}{2}}\frac{f(x)\,m^{(2)}(x)}{f(x)-2a\xi}-\sqrt{\frac{Var\left[Y|X=x\right]f(x)\int_{\mathbb{R}}K^{2}(z)\,dz}{\left(2f(x)-(\alpha-a)\,\xi\right)}},\right.$$
$$\left.\sqrt{\frac{c}{2}}\frac{f(x)\,m^{(2)}(x)}{f(x)-2a\xi}+\sqrt{\frac{Var\left[Y|X=x\right]f(x)\int_{\mathbb{R}}K^{2}(z)\,dz}{\left(2f(x)-(\alpha-a)\,\xi\right)}}\right].$$

(2) *If* $\gamma_n^{-1}h_n^5/\ln(s_n)\to\infty$, *and if* $\lim_{n\to\infty}(n\gamma_n)>2a/f(x)$ *then, with probability one,*

$$\lim_{n\to\infty}\frac{1}{h_n^2}\left(\rho_n(x)-r(x)\right)=\frac{f(x)\,m^{(2)}(x)}{f(x)-2a\xi}.$$

Lemma 4.4 (Strong uniform convergence rate of ρ_n)**.**
Let I *be a bounded open interval on which* $\varphi=\inf_{x\in I}f(x)>0$, *and let Assumptions* $(A1)-(A4)$ *hold for all* $x\in I$. *If* $\lim_{n\to\infty}n\gamma_n>\min\{(1-a)/(2\varphi),2a/\varphi\}$, *then*

$$\sup_{x\in I}\left|\rho_n(x)-r(x)\right|=O\left(\max\left\{\sqrt{\gamma_n h_n^{-1}}\ln n,h_n^2\right\}\right)\quad a.s.$$

To prove Lemmas 4.2 and 4.3, we first remark that, in view of (4.6), we have, for $n\geq n_0$,

$$\begin{aligned}\rho_n(x)-r(x) &= \left(1-\gamma_n f(x)\right)\left(\rho_{n-1}(x)-r(x)\right)+\gamma_n\left(W_n(x)-r(x)\,Z_n(x)\right)\\ &= \Pi_n\left(f(x)\right)\sum_{k=n_0}^{n}\Pi_k^{-1}\left(f(x)\right)\gamma_k\left(W_k(x)-r(x)\,Z_k(x)\right)\\ &\quad+\Pi_n\left(f(x)\right)\left(\rho_{n_0-1}(x)-r(x)\right)\\ &= \tilde{T}_n(x)+\tilde{R}_n(x),\end{aligned}\tag{4.9}$$

with, since $\rho_{n_0-1}=r_{n_0-1}$,

$$\begin{aligned}\tilde{T}_n(x) &= \sum_{k=n_0}^{n}U_{k,n}\left(f(x)\right)\gamma_k\left(W_k(x)-r(x)\,Z_k(x)\right),\\ \tilde{R}_n(x) &= \Pi_n\left(f(x)\right)\left(r_{n_0-1}(x)-r(x)\right).\end{aligned}$$

Noting that $\left|r_{n_0-1}(x)-r(x)\right|=O(1)$ a.s. and applying Lemma 4.1, we get

$$\begin{aligned}\left|\tilde{R}_n(x)\right| &= O\left(\Pi_n\left(f(x)\right)\right)\quad a.s.\\ &= o\left(m_n\right)\quad a.s.\end{aligned}$$

Lemmas 4.2 and 4.3 are thus straightforward consequences of the following lemmas, which are proved in Sections 5.1 and 5.2, respectively.

Lemma 4.5. *The two parts of Lemma 4.2 hold when* $\rho_n(x)-r(x)$ *is replaced by* $\tilde{T}_n(x)$.

Lemma 4.6. *The two parts of Lemma 4.3 hold when* $\rho_n(x)-r(x)$ *is replaced by* $\tilde{T}_n(x)$.

In the same way, we remark that

$$\begin{aligned}\sup_{x\in I}\left|\tilde{R}_n(x)\right| &= O\left(\sup_{x\in I}\Pi_n(f(x))\right) \quad a.s.\\ &= O(\Pi_n(\varphi)) \quad a.s.\\ &= o(m_n) \quad a.s.,\end{aligned}$$

so that Lemma 4.4 is a straightforward consequence of the following lemma, which is proved in Section 5.3.

Lemma 4.7. *Lemma 4.4 holds when $\rho_n - r$ is replaced by $\tilde{T}_n$.*

4.2. *Asymptotic behaviour of $\bar{\rho}_n$.* The purpose of this section is to give the outlines of the proof of the three following lemmas.

Lemma 4.8 (Weak pointwise convergence rate of $\bar{\rho}_n$)**.**
Let Assumptions $(A1)-(A3)$, $(A5)$ and $(A6)$ hold for $x\in\mathbb{R}$ such that $f(x)\neq 0$.

(1) *If there exists $c\geq 0$ such that $nh_n^5\to c$, then*

$$\begin{aligned}&\sqrt{nh_n}\left(\bar{\rho}_n(x)-r(x)\right)\\ &\quad\xrightarrow{\mathcal{D}}\mathcal{N}\left(c^{\frac{1}{2}}\frac{1-q}{1-q-2a}m^{(2)}(x),\frac{(1-q)^2}{1+a-2q}\frac{Var[Y|X=x]}{f(x)}\int_{\mathbb{R}}K^2(z)\,dz\right).\end{aligned}$$

(2) *If $nh_n^5\to\infty$, then*

$$h_n^{-2}\left(\bar{\rho}_n(x)-r(x)\right)\xrightarrow{\mathbb{P}}\frac{1-q}{1-q-2a}m^{(2)}(x).$$

Lemma 4.9 (Strong pointwise convergence rate of $\bar{\rho}_n$)**.**
Let Assumptions $(A1)-(A3)$, $(A5)$ and $(A6)$ hold for $x\in\mathbb{R}$ such that $f(x)\neq 0$.

(1) *If there exists $c_1\geq 0$ such that $nh_n^5/\ln\ln n\to c_1$, then, with probability one, the sequence*

$$\left(\sqrt{\frac{nh_n}{2\ln\ln n}}\left(\bar{\rho}_n(x)-r(x)\right)\right)$$

is relatively compact and its limit set is the interval

$$\begin{aligned}&\left[\frac{1-q}{1-q-2a}\sqrt{\frac{c_1}{2}}m^{(2)}(x)-\sqrt{\frac{(1-q)^2}{1+a-2q}\frac{Var[Y/X=x]}{f(x)}\int_{\mathbb{R}}K^2(z)\,dz},\right.\\ &\quad\left.\frac{1-q}{1-q-2a}\sqrt{\frac{c_1}{2}}m^{(2)}(x)+\sqrt{\frac{(1-q)^2}{1+a-2q}\frac{Var[Y/X=x]}{f(x)}\int_{\mathbb{R}}K^2(z)\,dz}\right].\end{aligned}$$

(2) *If $nh_n^5/\ln\ln n\to\infty$, then*

$$\lim_{n\to\infty}h_n^{-2}\left(\bar{\rho}_n(x)-r(x)\right)=\frac{1-q}{1-q-2a}m^{(2)}(x)\quad a.s.$$

Lemma 4.10 (Strong uniform convergence rate of $\bar{\rho}_n$)**.**
Let I be a bounded open interval on which $\varphi = \inf_{x\in I} f(x) > 0$ and let Assumptions $(A1)-(A6)$ hold for all $x \in I$. We have

$$\sup_{x\in I} |\bar{\rho}_n(x) - r(x)| = O\left(\max\left\{\sqrt{n^{-1}h_n^{-1}}\ln n, h_n^2\right\}\right) \quad a.s.$$

To prove Lemmas 4.8-4.10, we note that (4.6) gives, for $n \geq n_0$,

$$\rho_n(x) - \rho_{n-1}(x) = -\gamma_n f(x)\left[\rho_{n-1}(x) - r(x)\right] + \gamma_n\left[W_n(x) - r(x) Z_n(x)\right],$$

and thus

$$\rho_{n-1}(x) - r(x) = \frac{1}{f(x)}\left[W_n(x) - r(x) Z_n(x)\right] - \frac{1}{\gamma_n f(x)}\left[\rho_n(x) - \rho_{n-1}(x)\right].$$

It follows that

$$\begin{aligned} \bar{\rho}_n(x) - r(x) &= \frac{1}{\sum_{k=1}^n q_k}\sum_{k=1}^n q_k\left[\rho_k(x) - r(x)\right] \\ &= \frac{1}{f(x)} T_n(x) - \frac{1}{f(x)} R_n^{(0)}(x), \end{aligned} \tag{4.10}$$

with

$$\begin{aligned} T_n(x) &= \frac{1}{\sum_{k=1}^n q_k}\sum_{k=n_0-1}^n q_k\left[W_{k+1}(x) - r(x) Z_{k+1}(x)\right], \\ R_n^{(0)}(x) &= \frac{1}{\sum_{k=1}^n q_k}\sum_{k=n_0-1}^n \frac{q_k}{\gamma_{k+1}}\left[\rho_{k+1}(x) - \rho_k(x)\right]. \end{aligned}$$

Let us note that $R_n^{(0)}$ can be rewritten as

$$\begin{aligned} & R_n^{(0)}(x) \\ &= \frac{1}{\sum_{k=1}^n q_k}\sum_{k=n_0-1}^n \frac{q_k}{\gamma_{k+1}}\left[\left(\rho_{k+1}(x) - r(x)\right) - \left(\rho_k(x) - r(x)\right)\right] \\ &= \frac{1}{\sum_{k=1}^n q_k}\sum_{k=n_0}^n \left(\frac{q_{k-1}}{\gamma_k} - \frac{q_k}{\gamma_{k+1}}\right)\left(\rho_k(x) - r(x)\right) \\ &\quad + \frac{1}{\sum_{k=1}^n q_k}\frac{q_n}{\gamma_{n+1}}\left(\rho_{n+1}(x) - r(x)\right) - \frac{1}{\sum_{k=1}^n q_k}\frac{q_{n_0-1}}{\gamma_{n_0}}\left(\rho_{n_0-1}(x) - r(x)\right) \\ &= \frac{1}{\sum_{k=1}^n q_k}\sum_{k=n_0}^n \frac{q_{k-1}}{\gamma_k}\left[1 - \frac{q_{k-1}^{-1}\gamma_k}{q_k^{-1}\gamma_{k+1}}\right]\left(\rho_k(x) - r(x)\right) \\ &\quad + \frac{1}{\sum_{k=1}^n q_k}\frac{q_n}{\gamma_{n+1}}\left(\rho_{n+1}(x) - r(x)\right) - \frac{1}{\sum_{k=1}^n q_k}\frac{q_{n_0-1}}{\gamma_{n_0}}\left(\rho_{n_0-1}(x) - r(x)\right). \end{aligned}$$

Since $\left(q_{k-1}^{-1}\gamma_k\right) \in \mathcal{GS}(q-\alpha)$, we have

$$\left[1 - \frac{q_{k-1}^{-1}\gamma_k}{q_k^{-1}\gamma_{k+1}}\right] = 1 - \left(1 - \frac{(q-\alpha)}{k} + o\left(\frac{1}{k}\right)\right) = O\left(k^{-1}\right)$$

and thus

$$\left|R_n^{(0)}(x)\right| = O\left(\frac{1}{\sum_{k=1}^n q_k}\left[\sum_{k=n_0}^n k^{-1} q_{k-1}\gamma_k^{-1}\left|\rho_k(x)-r(x)\right| + \frac{q_n}{\gamma_{n+1}}\left|\rho_{n+1}(x)-r(x)\right| + \frac{q_{n_0-1}}{\gamma_{n_0}}\left|\rho_{n_0-1}(x)-r(x)\right|\right]\right). \tag{4.11}$$

The application of Lemma 4.3 ensures that

$$\left|R_n^{(0)}(x)\right| = O\left(\frac{1}{\sum_{k=1}^n q_k}\left[\sum_{k=2}^n \left(k^{-1} q_k\gamma_k^{-1}\right)\left(\left(\gamma_k h_k^{-1}\ln(s_k)\right)^{\frac{1}{2}} + h_k^2\right) + \frac{q_n}{\gamma_{n+1}}\left(\left(\gamma_n h_n^{-1}\ln(s_n)\right)^{\frac{1}{2}} + h_n^2\right) + 1\right]\right) \quad a.s.$$

Now, let us recall that, if $(u_n) \in \mathcal{GS}(-u^*)$ with $u^* < 1$, then we have, for any fixed $k_0 \geq 1$,

$$\lim_{n\to\infty}\frac{nu_n}{\sum_{k=k_0}^n u_k} = 1-u^* \tag{4.12}$$

and, if $u^* \geq 1$, then for all $\epsilon > 0$, $u_n = O\left(n^{-1+\epsilon}\right)$ and thus

$$\sum_{k=1}^n u_k = O\left(n^\epsilon\right). \tag{4.13}$$

Now, set $\epsilon \in \left]0, \min\left\{1-q-2a, (1+a)/2-q\right\}\right[$ (the existence of such an ϵ being ensured by $(A6)$); in view of $(A5)$, we get

$$\begin{aligned}\left|R_n^{(0)}(x)\right| &= O\left(\frac{1}{nq_n}\left(n^\epsilon + q_n\gamma_n^{-\frac{1}{2}}h_n^{-\frac{1}{2}}\left(\ln(s_n)\right)^{\frac{1}{2}} + q_n\gamma_n^{-1}h_n^2\right) + \frac{1}{n\gamma_n}\left(\left(\gamma_n h_n^{-1}\ln(s_n)\right)^{\frac{1}{2}} + h_n^2\right)\right) \quad a.s.\\ &= O\left(\frac{n^\epsilon}{nq_n} + \frac{\sqrt{n^{-1}h_n^{-1}}}{\sqrt{n\gamma_n\left(\ln(s_n)\right)^{-1}}} + \frac{h_n^2}{n\gamma_n}\right) \quad a.s.\\ &= o\left(\sqrt{n^{-1}h_n^{-1}} + h_n^2\right) \quad a.s.\end{aligned}$$

In view of (4.10), Lemmas 4.8 and 4.9 are thus straightforward consequences of the two following lemmas, which are proved in Sections 5.4 and 5.5 respectively.

Lemma 4.11 (Weak pointwise convergence rate of T_n).
The two parts of Lemma 4.8 hold when $\bar{\rho}_n(x) - r(x)$ is replaced by $[f(x)]^{-1}T_n(x)$
.

Lemma 4.12 (Strong pointwise convergence rate of T_n).
The two parts of Lemma 4.9 hold when $\bar{\rho}_n(x) - r(x)$ is replaced by $[f(x)]^{-1}T_n(x)$
.

Now, in view of (4.11), the application of Lemma 4.4 ensures that

$$\begin{aligned}\sup_{x\in I}\left|R_n^{(0)}(x)\right| &= O\left(\frac{1}{\sum_{k=1}^n q_k}\left[\sum_{k=n_0}^n k^{-1}q_k\gamma_k^{-1}\sup_{x\in I}\left|\rho_k(x)-r(x)\right|\right.\right.\\&\left.\left.+\frac{q_n}{\gamma_{n+1}}\sup_{x\in I}\left|\rho_{n+1}(x)-r(x)\right|\right]\right)\quad a.s.\\&= O\left(\frac{1}{\sum_{k=1}^n q_k}\left[\sum_{k=n_0}^n\left(k^{-1}q_k\gamma_k^{-1}\right)\left(\left(\gamma_k h_k^{-1}\right)^{\frac{1}{2}}\ln k+h_k^2\right)\right.\right.\\&\left.\left.+\frac{q_n}{\gamma_{n+1}}\left(\left(\gamma_n h_n^{-1}\right)^{\frac{1}{2}}\ln n+h_n^2\right)\right]\right)\quad a.s.\end{aligned}$$

Setting $\epsilon\in\left]0,\min\left\{1-q-2a,(1+a)/2-q\right\}\right[$ again, we get, in view of $(A5)$,

$$\begin{aligned}\sup_{x\in I}\left|R_n^{(0)}(x)\right| &= O\left(\frac{1}{nq_n}\left(n^\epsilon+q_n\gamma_n^{-\frac{1}{2}}h_n^{-\frac{1}{2}}\ln n+q_n\gamma_n^{-1}h_n^2\right)\right.\\&\left.+\frac{1}{n\gamma_n}\left(\left(\gamma_n h_n^{-1}\right)^{\frac{1}{2}}\ln n+h_n^2\right)\right)\quad a.s.\\&= O\left(\frac{n^\epsilon}{nq_n}+\frac{\sqrt{n^{-1}h_n^{-1}}\,\ln n}{\sqrt{n\gamma_n}}+\frac{h_n^2}{n\gamma_n}\right)\quad a.s.\\&= o\left(\sqrt{n^{-1}h_n^{-1}}\,\ln n+h_n^2\right)\quad a.s.\end{aligned}$$

In view of (4.10), Lemma 4.10 is thus a straightforward consequence of the following lemma, which is proved in Section 5.6.

Lemma 4.13 (Strong uniform convergence rate of T_n).
Lemma 4.10 holds when $\bar{\rho}_n-r$ is replaced by T_n .

4.3. *How to deduce the asymptotic behaviour of r_n and $\bar{r}_n$ from that of ρ_n and $\bar{\rho}_n$.* Set

$$\Delta_n(x) = r_n(x)-\rho_n(x)$$

and

$$\begin{aligned}\bar{\Delta}_n(x) &= \frac{1}{\sum_{k=1}^n q_k}\sum_{k=1}^n q_k\Delta_k(x)\\&= \bar{r}_n(x)-\bar{\rho}_n(x).\end{aligned}$$

To deduce the asymptotic behaviour of r_n (respectively $\bar{r}_n$) from that of ρ_n (respectively $\bar{\rho}_n$), we prove that Δ_n (respectively $\bar{\Delta}_n$) is negligible in front of ρ_n (respectively $\bar{\rho}_n$). Note that, in view of (4.5) and (4.6), and since $\rho_{n_0-1}(x)=r_{n_0-1}(x)$, we have, for $n\geq n_0$,

$$\begin{aligned}\Delta_n(x) &= (1-\gamma_n f(x))\,\Delta_{n-1}(x)+\gamma_n\left(f(x)-Z_n(x)\right)\left(r_{n-1}(x)-r(x)\right)\\&= \sum_{k=n_0}^n U_{k,n}(f(x))\left(f(x)-Z_k(x)\right)\left(r_{k-1}(x)-r(x)\right).\end{aligned}\tag{4.14}$$

The difficulty which appears here is that Δ_n is expressed in function of the terms r_k-r, so that an upper bound of r_n-r is necessary for the obtention of an upper bound of Δ_n. Now, the key to overcome this difficulty is the following property

$(\mathcal{P})$: if $(r_n - r)$ is known to be bounded almost surely by a sequence (w_n), then it can be shown that (Δ_n) is bounded almost surely by a sequence (w'_n) such that $\lim_{n\to\infty} w'_n w_n^{-1} = 0$, which may allow to upper bound $r_n - r$ by a sequence smaller than (w_n). To deduce the asymptotic behaviour of r_n (respectively $\bar{r}_n$) from that of ρ_n (respectively $\bar{\rho}_n$), we thus proceed as follows. We first establish a rudimentary upper bound of $(r_n - r)$. Then, applying Property $(\mathcal{P})$ several times, we successively improve our upper bound of $(r_n - r)$, and this until we obtain an upper bound, which allows to prove that Δ_n (respectively $\bar{\Delta}_n$) is negligible in front of ρ_n (respectively $\bar{\rho}_n$).
We first establish the pointwise results on r_n and $\bar{r}_n$ (that is, Theorems 2.3, 2.4, 2.6, and 2.7) in Section 4.3.1, and then the uniform ones (that is, Theorems 2.5 and 2.8) in Section 4.3.2.

4.3.1. *Proof of Theorems 2.3, 2.4, 2.6 and 2.7.* The proof of Theorems 2.3, 2.4, 2.6 and 2.7 relies on the repeted application of the following lemma, which is proved in Section 5.7.

Lemma 4.14. *Let Assumptions $(A1) - (A3)$ hold, and assume that there exists $(w_n) \in \mathcal{GS}(w^*)$ such that $|r_n(x) - r(x)| = O(w_n)$ a.s.*

(1) *If the sequence $(n\gamma_n)$ is bounded, if $\lim_{n\to\infty} n\gamma_n > \min\{(1-a)/(2f(x)), 2a/f(x)\}$, and if $w^* \geq 0$, then, for all $\delta > 0$,*
$$|\Delta_n(x)| = O\left(m_n w_n (\ln n)^{\frac{(1+\delta)}{2}}\right) + o(m_n) \quad a.s.$$
(2) *If $\lim_{n\to\infty}(n\gamma_n) = \infty$, then, for all $\delta > 0$,*
$$|\Delta_n(x)| = O\left(m_n w_n \left(n^{1+\delta}\gamma_n\right)^{\frac{(1+\delta)}{2}}\right) \quad a.s.$$

We first establish a preliminary upper bound for $r_n(x) - r(x)$. Then, we successively prove Theorems 2.3 and 2.4 in the case $(n\gamma_n)$ is bounded, Theorems 2.3 and 2.4 in the case $\lim_{n\to\infty}(n\gamma_n) = \infty$, and finally Theorems 2.6 and 2.7.

Preliminary upper bound of $r_n(x) - r(x)$.
Since $0 \leq 1 - \gamma_n Z_n(x) \leq 1$ for all $n \geq n_0$, it follows from (4.4) that, for $n \geq n_0$,
$$\begin{aligned} |r_n(x)| &\leq |r_{n-1}(x)| + \gamma_n |Y_n| h_n^{-1} \|K\|_\infty \\ &\leq |r_{n_0-1}(x)| + \left(\sup_{k\leq n}|Y_k|\right)\|K\|_\infty \sum_{k=1}^n \gamma_k h_k^{-1}. \end{aligned} \tag{4.15}$$
Since
$$\mathbb{P}\left(\sup_{k\leq n}|Y_k| > n^2\right) \leq n\mathbb{P}\left(|Y| > n^2\right) \leq n^{-3}\mathbb{E}\left(|Y|^2\right),$$
we have $\sup_{k\leq n}|Y_k| \leq n^2$ a.s. Moreover, since $\left(\gamma_n h_n^{-1}\right) \in \mathcal{GS}(-\alpha + a)$ with $1 - \alpha + a > 0$, we note that $\sum_{k=1}^n \gamma_k h_k^{-1} = O\left(n\gamma_n h_n^{-1}\right)$. We thus deduce that
$$|r_n(x) - r(x)| = O\left(n^3\gamma_n h_n^{-1}\right) \quad a.s. \tag{4.16}$$

Proof of Theorems 2.3 and 2.4 in the case the sequence $(n\gamma_n)$ is bounded. In this case, $\alpha = 1$, and Lemmas 4.3 and 4.14 imply that:

- $|\rho_n(x) - r(x)| = O(m_n \ln n) \quad a.s.$ (4.17)
- If $\exists (w_n) \in \mathcal{GS}(w^*), w^* \geq 0,$ such that $|r_n(x) - r(x)| = O(w_n)$ $a.s.$, then $|\Delta_n(x)| = O(m_n w_n \ln n) + o(m_n) \quad a.s.$ (4.18)

Set $p_0 = \max\{p \text{ such that } -m^*p + 2 + a \geq 0\}$, set $j \in \{0, 1, \ldots, p_0 - 1\}$ and assume that

$$|r_n(x) - r(x)| \quad = \quad O\left(m_n^j \left(n^3\gamma_n h_n^{-1}\right)(\ln n)^j\right) \quad a.s. \tag{4.19}$$

Since the sequence $(w_n) = \left(m_n^j \left(n^3\gamma_n h_n^{-1}\right)(\ln n)^j\right)$ belongs to $\mathcal{GS}(-m^*j + 2 + a)$ with $-m^*j + 2 + a > 0$, the application of (4.18) implies that

$$|\Delta_n(x)| = O\left(m_n^{j+1}\left(n^3\gamma_n h_n^{-1}\right)(\ln n)^{j+1}\right) + o(m_n) \quad a.s.$$

Since $\left(m_n^{j+1}\left(n^3\gamma_n h_n^{-1}\right)(\ln n)^{j+1}\right) \in \mathcal{GS}(-m^*(j+1) + 2 + a)$ with $-m^*(j+1) + 2 + a \geq 0$, whereas $(m_n) \in \mathcal{GS}(-m^*)$ with $-m^* < 0$, it follows that

$$|\Delta_n(x)| = O\left(m_n^{j+1}\left(n^3\gamma_n h_n^{-1}\right)(\ln n)^{j+1}\right) \quad a.s.,$$

and the application of (4.17) leads to

$$\begin{aligned} |r_n(x) - r(x)| &\leq |\rho_n(x) - r(x)| + |\Delta_n(x)| \\ &= O\left(m_n^{j+1}\left(n^3\gamma_n h_n^{-1}\right)(\ln n)^{j+1}\right) \quad a.s. \end{aligned}$$

Since (4.16) ensures that (4.19) is satisfied for $j = 0$, we have proved by induction that

$$|r_n(x) - r(x)| = O\left(m_n^{p_0}\left(n^3\gamma_n h_n^{-1}\right)(\ln n)^{p_0}\right) \quad a.s.$$

Applying (4.18) with $(w_n) = \left(m_n^{p_0}\left(n^3\gamma_n h_n^{-1}\right)(\ln n)^{p_0}\right)$ and then (4.17), we obtain

$$|r_n(x) - r(x)| = O\left(m_n^{p_0+1}\left(n^3\gamma_n h_n^{-1}\right)(\ln n)^{p_0+1}\right) + O(m_n \ln n) \quad a.s.$$

Since the sequences $\left(m_n^{p_0+1}\left(n^3\gamma_n h_n^{-1}\right)(\ln n)^{p_0+1}\right)$ and $(m_n \ln n)$ are in $\mathcal{GS}(-m^*(p_0+1)+2+a)$ with $-m^*(p_0+1)+2+a < 0$ and $\mathcal{GS}(-m^*)$ with $-m^* < 0$ respectively, it follows that

$$|r_n(x) - r(x)| = O\left((\ln n)^{-2}\right) \quad a.s.$$

Applying once more (4.18) with $(w_n) = \left((\ln n)^{-2}\right) \in \mathcal{GS}(0)$, we get

$$|\Delta_n(x)| = O\left(m_n(\ln n)^{-1}\right) + o(m_n) = o(m_n) \quad a.s.$$

Theorem 2.3 (respectively Theorem 2.4) in the case $(n\gamma_n)$ is bounded then follows from the application of Lemma 4.2 (respectively Lemma 4.3).

Proof of Theorems 2.3 and 2.4 in the case $\lim_{n\to\infty}(n\gamma_n)=\infty$.
In this case, Lemmas 4.3 and 4.14 imply that, for all $\delta>0$,

- $|\rho_n(x)-r(x)|=O(\tilde{m}_n)\quad a.s.$ (4.20)
- If there exists $(w_n)\in\mathcal{GS}(w^*)$ such that $|r_n(x)-r(x)|=O(w_n)\,a.s.$, then $|\Delta_n(x)|=O\left(m_n\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}w_n\right)\quad a.s.$ (4.21)

Now, set $\delta>0$ such that $c(\delta)=-m^*+(1+\delta)(1+\delta-\alpha)/2<0$ (the existence of such a δ being ensured by $(A2)$). In view of (4.16), the application of (4.21) with $(w_n)=\left(n^3\gamma_nh_n^{-1}\right)$ ensures that

$$|\Delta_n(x)| = O\left(m_n\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}n^3\gamma_nh_n^{-1}\right)\quad a.s.$$

and, in view of (4.20), it follows that

$$|r_n(x)-r(x)| = O(\tilde{m}_n)+O\left(m_n\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}n^3\gamma_nh_n^{-1}\right)\quad a.s.$$

Set $p\geq1$, and assume that

$$|r_n(x)-r(x)| = O(\tilde{m}_n)+O\left(m_n^p\left(n^{1+\delta}\gamma_n\right)^{p\left(\frac{1+\delta}{2}\right)}n^3\gamma_nh_n^{-1}\right)\quad a.s.$$

The application of (4.21) with $(w_n)=(\tilde{m}_n)$ and with

$$(w_n)=\left(m_n^p\left(n^{1+\delta}\gamma_n\right)^{p\left(\frac{1+\delta}{2}\right)}n^3\gamma_nh_n^{-1}\right)$$

ensures that

$$|\Delta_n(x)|=O\left(m_n\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}\tilde{m}_n\right)+O\left(m_n^{p+1}\left(n^{1+\delta}\gamma_n\right)^{(p+1)\frac{1+\delta}{2}}n^3\gamma_nh_n^{-1}\right)\quad a.s.$$

The sequence $\left(m_n\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}\right)$ being in $\mathcal{GS}(c(\delta))$ with $c(\delta)<0$, it follows that

$$|\Delta_n(x)| = o(\tilde{m}_n)+O\left(m_n^{p+1}\left(n^{1+\delta}\gamma_n\right)^{(p+1)\frac{1+\delta}{2}}n^3\gamma_nh_n^{-1}\right)\quad a.s.$$

and, in view of (4.20), we obtain

$$|r_n(x)-r(x)| = O(\tilde{m}_n)+O\left(m_n^{p+1}\left(n^{1+\delta}\gamma_n\right)^{(p+1)\frac{1+\delta}{2}}n^3\gamma_nh_n^{-1}\right)\quad a.s.$$

We have thus proved by induction that, for all $p\geq1$,

$$|r_n(x)-r(x)| = O(\tilde{m}_n)+O\left(m_n^p\left(n^{1+\delta}\gamma_n\right)^{p\left(\frac{1+\delta}{2}\right)}n^3\gamma_nh_n^{-1}\right)\quad a.s.$$

By setting p large enough, we deduce that

$$|r_n(x)-r(x)| = O(\tilde{m}_n)\quad a.s.$$

Applying once more (4.21) with $(w_n)=(\tilde{m}_n)$, we get

$$\begin{aligned}|\Delta_n(x)| &= O\left(m_n\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}\tilde{m}_n\right)\quad a.s. \\ &= o(m_n)\quad a.s.\end{aligned}\tag{4.22}$$

Theorem 2.3 (respectively Theorem 2.4) in the case $\lim_{n\to\infty}(n\gamma_n)=\infty$ then follows from the application of Lemma 4.2 (respectively Lemma 4.3).

Proof of Theorems 2.6 and 2.7.
In view of (4.22), and applying (4.12) and (4.13), we get, for all $\delta>0$,

$$\begin{aligned}\left|\bar{\Delta}_n(x)\right| &= O\left(\frac{1}{\sum_{k=1}^n q_k}\sum_{k=1}^n q_k\tilde{m}_k^2\left(k^{1+\delta}\gamma_k\right)^{\frac{1+\delta}{2}}\right) \quad a.s.\\ &= O\left(\frac{1}{nq_n}\left[n^\delta+nq_n\tilde{m}_n^2\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}\right]\right) \quad a.s.\\ &= O\left(n^{\delta-1}q_n^{-1}+\tilde{m}_n^2\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}\right) \quad a.s. \qquad (4.23)\end{aligned}$$

• Let us first consider the case when the sequence $\left(nh_n^5\right)$ is bounded. In this case, we have $a\geq 1/5$, so that $a\geq\alpha/5$ and $m^*=(\alpha-a)/2$. Noting that $(A2)$ implies $a<3\alpha-2$, and applying $(A6)$, we can set $\delta>0$ such that

$$\delta-\frac{(1+a)}{2}+q<0 \text{ and } \frac{(1-a)}{2}-2m^*+\frac{(1+\delta)}{2}(1+\delta-\alpha)<0.$$

In view of (4.23), we then obtain

$$\sqrt{nh_n}\left|\bar{\Delta}_n(x)\right|=o(1) \quad a.s.$$

The first part of Theorem 2.6 (respectively of Theorem 2.7) then follows from the application of the first part of Lemma 4.8 (respectively of Lemma 4.9).
• Let us now consider the case when $\lim_{n\to\infty}\left(nh_n^5\right)=\infty$. Noting that $(A2)$ then ensures that $6a<3\alpha-1$, and applying $(A6)$, we can set $\delta>0$ such that

$$2a+\delta-1+q<0 \text{ and } 2a-2m^*+\frac{(1+\delta)}{2}(1+\delta-\alpha)<0.$$

It then follows from (4.23) that

$$h_n^{-2}\left|\bar{\Delta}_n(x)\right|=o(1) \quad a.s.$$

The second part of Theorem 2.6 (respectively of Theorem 2.7) then follows from the application of the second part of Lemma 4.8 (respectively of Lemma 4.9).

4.3.2. *Proof of Theorems 2.5 and 2.8.* Set

$$B_n=n\gamma_nh_n^{-1}\ln n. \qquad (4.24)$$

The proof of Theorems 2.5 and 2.8 relies on the repeted application of the following lemma, which is proved in Section 5.8.

Lemma 4.15. *Let I be a bounded open interval on which $\varphi=\inf_{x\in I}f(x)>0$, let Assumptions $(A1)-(A4)$ hold for all $x\in I$, and assume that there exists $(w_n)\in\mathcal{GS}(w^*)$ such that $\sup_{x\in I}|r_n(x)-r(x)|=O(w_n)$ a.s. Moreover,*
• *in the case when $(n\gamma_n)$ is bounded, assume that $\lim_{n\to\infty}n\gamma_n>m^*/\varphi$ and that $w^*\geq 0$;*
• *in the case when $\lim_{n\to\infty}n\gamma_n=\infty$, assume that $\left(w_n^{-1}B_n\sqrt{\gamma_nh_n^{-1}\ln n}\right)$ is a bounded sequence. Then, we have*

$$\sup_{x\in I}|\Delta_n(x)|=O\left(m_nw_n\sqrt{\ln n}\right) \quad a.s.$$

We first establish a preliminary upper bound of $\sup_{x\in I}\left(|r_n(x)-r(x)|\right)$ (which is better than the pointwise upper bound (4.16) since the random variables Y_k are assumed to have a finite exponential moment in Theorems 2.5 and 2.8). Then, we successively prove Theorem 2.5 in the case when $(n\gamma_n)$ is bounded, Theorem 2.5 in the case when $\lim_{n\to\infty}(n\gamma_n)=\infty$, and finally Theorem 2.8.

Preliminary upper bound.
Proceeding as for the proof of (4.16), we note that, for all $n\geq n_0$,

$$\sup_{x\in I}|r_n(x)|\leq\sup_{x\in I}|r_{n_0-1}(x)|+\left(\sup_{k\leq n}|Y_k|\right)\|K\|_\infty\sum_{k=1}^n\gamma_kh_k^{-1},$$

with, this time, in view of $(A4)$,

$$\mathbb{P}\left[\sup_{k\leq n}|Y_k|>\frac{3}{t^*}\ln n\right]\leq n\mathbb{P}\left[\exp\left(t^*|Y|\right)>n^3\right]\leq n^{-2}\mathbb{E}\left(\exp\left(t^*|Y|\right)\right).$$

We deduce that

$$\sup_{x\in I}|r_n(x)-r(x)|=O(B_n)\quad a.s.$$

Proof of Theorem 2.5 in the case $(n\gamma_n)$ is bounded.
In this case, we have $\alpha=1$, $(B_n)\in\mathcal{GS}(a)$ (with $a>0$); the application of Lemma 4.15 with $(w_n)=(B_n)$ ensures that

$$\sup_{x\in I}|\Delta_n(x)|=O\left(m_nB_n\sqrt{\ln n}\right)\quad a.s.$$

Applying Lemma 4.4, we get

$$\begin{aligned}\sup_{x\in I}|r_n(x)-r(x)| &\leq \sup_{x\in I}|\rho_n(x)-r(x)|+\sup_{x\in I}|\Delta_n(x)|\\ &= O\left(m_nB_n\sqrt{\ln n}\right)\quad a.s.\end{aligned}$$

Since $\left(m_nB_n\sqrt{\ln n}\right)\in\mathcal{GS}(-m^*+a)$ (with $-m^*+a<0$), it follows that

$$\sup_{x\in I}|r_n(x)-r(x)| = O\left([\ln n]^{-1}\right)\quad a.s.$$

Applying once more Lemma 4.15 with $(w_n)=\left([\ln n]^{-1}\right)$, we get

$$\begin{aligned}\sup_{x\in I}|\Delta_n(x)| &= O\left(m_n(\ln n)^{-\frac{1}{2}}\right)\quad a.s.\\ &= o(\tilde{m}_n)\quad a.s.\end{aligned}$$

Theorem 2.5 in the case when $(n\gamma_n)$ is bounded then follows from the application of Lemma 4.4.

Proof of Theorem 2.5 in the case $\lim_{n\to\infty}(n\gamma_n)=\infty$.
The sequence $\left(\sqrt{\gamma_nh_n^{-1}\ln n}\right)$ being clearly bounded, we can apply Lemma 4.15 with $(w_n)=(B_n)$; we then obtain

$$\sup_{x\in I}|\Delta_n(x)|=O\left(m_nB_n\sqrt{\ln n}\right)\quad a.s.$$

The application of Lemma 4.4 then ensures that

$$\begin{aligned}\sup_{x\in I}\left|r_n(x)-r(x)\right| &= O\left(\tilde{m}_n\right)+O\left(m_nB_n\sqrt{\ln n}\right) \quad a.s.\\ &= O\left(m_nB_n\sqrt{\ln n}\right) \quad a.s.\end{aligned}$$

Since $\left(m_nB_n\sqrt{\ln n}\right)^{-1}B_n\sqrt{\gamma_nh_n^{-1}\ln n}=m_n^{-1}\sqrt{\gamma_nh_n^{-1}}=O(1)$, we can apply once more Lemma 4.15 with $(w_n)=\left(m_nB_n\sqrt{\ln n}\right)$; we get

$$\sup_{x\in I}\left|\Delta_n(x)\right| = O\left(m_n^2B_n\ln n\right) \quad a.s. \tag{4.25}$$

Noting that $(m_nB_n\ln n)\in\mathcal{GS}\left(-m^*+1-\alpha+a\right)$ with, in view of $(A4)\,iii)$, $-m^*+1-\alpha+a<0$, it follows that

$$\sup_{x\in I}\left|\Delta_n(x)\right| = o\left(m_n\right) \quad a.s.$$

Theorem 2.5 in the case when $\lim_{n\to\infty}(n\gamma_n)=\infty$ then follows from the application of Lemma 4.4.

Proof of Theorem 2.8.
• In the case when the sequence $\left(nh_n^5/\ln n\right)$ is bounded, we have, in view of (4.25),

$$\sqrt{nh_n}\left(\ln n\right)^{-1}\sup_{x\in I}\left|\Delta_n(x)\right|=O\left(\sqrt{nh_n}m_n^2B_n\right) \quad a.s.$$

Now, in this case, we have $a\geq 1/5\geq\alpha/5$ and thus $m^*=(\alpha-a)/2$. It follows that $\left(\sqrt{nh_n}m_n^2B_n\right)\in\mathcal{GS}\left(3(1+a)/2-2\alpha\right)$ with $3(1+a)/2-2\alpha<0$, and thus

$$\sup_{x\in I}\left|\Delta_n(x)\right|=O\left(\sqrt{n^{-1}h_n^{-1}}\,\ln n\right) \quad a.s.$$

The first part of Theorem 2.8 then follows from the application of Lemma 4.10.
• In the case when $\lim_{n\to\infty}\left(nh_n^5/\ln n\right)=\infty$, (4.25) ensures that

$$h_n^{-2}\sup_{x\in I}\left|\Delta_n(x)\right|=O\left(h_n^{-2}m_n^2B_n\ln n\right) \quad a.s.,$$

with $\left(h_n^{-2}m_n^2B_n\ln n\right)\in\mathcal{GS}\left(3a-2m^*+1-\alpha\right)$. Noting that the assumptions of Theorem 2.8 ensure that $3a-2m^*+1-\alpha<0$, we deduce that

$$\sup_{x\in I}\left|\Delta_n(x)\right|=O\left(h_n^2\right) \quad a.s.$$

The second part of Theorem 2.8 then follows from the application of Lemma 4.10.

5. Proof of Lemmas

5.1. *Proof of Lemma 4.5.* We establish that, under the condition $\lim_{n\to\infty}(n\gamma_n)>(\alpha-a)/(2f(x))$,

- if $a\geq\alpha/5$, then $\sqrt{\gamma_n^{-1}h_n}\left(\tilde{T}_n(x)-\mathbb{E}\left(\tilde{T}_n(x)\right)\right)$
$$\stackrel{\mathcal{D}}{\to}\mathcal{N}\left(0,\frac{Var\left[Y|X=x\right]f(x)}{\left(2f(x)-(\alpha-a)\xi\right)}\int_{\mathbb{R}}K^2(z)\,dz\right), \tag{5.1}$$
- if $a>\alpha/5$, then $\sqrt{\gamma_n^{-1}h_n}\mathbb{E}\left(\tilde{T}_n(x)\right)\to 0, \quad (5.2)$

and prove that, under the condition $\lim_{n\to\infty}(n\gamma_n) > 2a/f(x)$,

- if $a \leq \alpha/5$, then $h_n^{-2}\mathbb{E}\left(\tilde{T}_n(x)\right) \to \dfrac{f(x)\, m^{(2)}(x)}{f(x) - 2a\xi}$, (5.3)
- if $a < \alpha/5$, then $h_n^{-2}\left(\tilde{T}_n(x) - \mathbb{E}\left(\tilde{T}_n(x)\right)\right) \xrightarrow{\mathbb{P}} 0.$ (5.4)

As a matter of fact, the combination of (5.1) and (5.2) (respectively, of (5.1) and (5.3)) gives Part 1 of Lemma 4.5 in the case $a > \alpha/5$ (respectively, $a = \alpha/5$), that of (5.3) and (5.4) (respectively, of (5.1) and (5.3)) gives Part 2 of Lemma 4.5 in the case $a < \alpha/5$ (respectively, $a = \alpha/5$). We prove (5.1), (5.4), (5.3), and (5.2) successively.

Proof of (5.1)**.** Set

$$\tilde{\eta}_k(x) = \Pi_k^{-1}(f(x))\,\gamma_k\left[(W_k(x) - r(x)\, Z_k(x))\right], \tag{5.5}$$

so that $\tilde{T}_n(x) - \mathbb{E}\left(\tilde{T}_n(x)\right) = \Pi_n(f(x))\sum_{k=n_0}^{n}\left[\tilde{\eta}_k(x) - \mathbb{E}(\tilde{\eta}_k(x))\right]$. We have

$$\begin{aligned} Var(\tilde{\eta}_k(x)) &= \Pi_k^{-2}(f(x))\,\gamma_k^2\left[Var(W_k(x)) + r^2(x)\, Var(Z_k(x))\right.\\ &\quad \left.-2r(x)\, Cov(W_k(x), Z_k(x))\right]. \end{aligned}$$

In view of $(A3)$, classical computations give

$$Var(W_k(x)) = \frac{1}{h_k}\left[\mathbb{E}\left[Y^2|X=x\right] f(x)\int_{\mathbb{R}} K^2(z)\,dz + o(1)\right], \tag{5.6}$$

$$Var(Z_k(x)) = \frac{1}{h_k}\left[f(x)\int_{\mathbb{R}} K^2(z)\,dz + o(1)\right], \tag{5.7}$$

$$Cov(W_k(x), Z_k(x)) = \frac{1}{h_k}\left[r(x)\, f(x)\int_{\mathbb{R}} K^2(z)\,dz + o(1)\right]. \tag{5.8}$$

It follows that

$$Var(\tilde{\eta}_k(x)) = \frac{\Pi_k^{-2}(f(x))\,\gamma_k^2}{h_k}\left[Var[Y|X=x]\, f(x)\int_{\mathbb{R}} K^2(z)\,dz + o(1)\right], \tag{5.9}$$

and, since $\lim_{n\to\infty}(n\gamma_n) > (\alpha - a)/(2f(x))$, Lemma 4.1 ensures that

$$\begin{aligned} v_n^2 &= \sum_{k=n_0}^{n} Var(\tilde{\eta}_k(x))\\ &= \sum_{k=n_0}^{n}\frac{\Pi_k^{-2}(f(x))\,\gamma_k^2}{h_k}\left[Var[Y|X=x]\, f(x)\int_{\mathbb{R}} K^2(z)\,dz + o(1)\right]\\ &= \frac{\gamma_n\Pi_n^{-2}(f(x))}{h_n[2f(x) - (\alpha - a)\xi]}\left[Var[Y|X=x] f(x)\int_{\mathbb{R}} K^2(z)dz + o(1)\right]. \end{aligned} \tag{5.10}$$

For all $p \in]0,1]$ and in view of $(A3)$, we have

$$\begin{aligned}
&\mathbb{E}\left(\left|Y_k - r(x)\right|^{2+p} K^{2+p}\left(\frac{x - X_k}{h_k}\right)\right)\\
&= h_k \int_{\mathbb{R}^2} \left|y - r(x)\right|^{2+p} K^{2+p}(s)\, g(x - h_k s, y)\, dy ds\\
&\leq 2^{1+p} h_k \int_{\mathbb{R}} \left\{\int_{\mathbb{R}} |y|^{2+p} g(x - h_k s, y)\, dy + \left|r(x)\right|^{2+p} \int_{\mathbb{R}} g(x - h_k s, y)\, dy\right\}\\
&\quad K^{2+p}(s)\, ds\\
&= O(h_k). \qquad (5.11)
\end{aligned}$$

Now, set $p \in]0,1]$ such that $\lim_{n\to\infty}(n\gamma_n) > (1+p)(\alpha - a)/((2+p) f(x))$. Applying Lemma 4.1, we get

$$\begin{aligned}
&\sum_{k=n_0}^{n} \mathbb{E}\left[\left|\tilde{\eta}_k(x)\right|^{2+p}\right]\\
&= O\left(\sum_{k=n_0}^{n} \frac{\Pi_k^{-2-p}(f(x))\,\gamma_k^{2+p}}{h_k^{2+p}} \mathbb{E}\left(\left|Y_k - r(x)\right|^{2+p} K^{2+p}\left(\frac{x - X_k}{h_k}\right)\right)\right)\\
&= O\left(\sum_{k=n_0}^{n} \frac{\Pi_k^{-2-p}(f(x))\,\gamma_k^{2+p}}{h_k^{1+p}}\right)\\
&= O\left(\frac{1}{\Pi_n^{2+p}(f(x))} \frac{\gamma_n^{1+p}}{h_n^{1+p}}\right). \qquad (5.12)
\end{aligned}$$

Using (5.10), we deduce that

$$\frac{1}{v_n^{2+p}} \sum_{k=n_0}^{n} \mathbb{E}\left[\left|\tilde{\eta}_k(x)\right|^{2+p}\right] = O\left(\left(\frac{\gamma_n}{h_n}\right)^{\frac{p}{2}}\right) = o(1),$$

and (5.1) follows by application of Lyapounov Theorem.

Proof of (5.4)***.*** In view of (5.9), and since $a < \alpha/5$ and $\lim_{n\to\infty}(n\gamma_n) > 2a/f(x)$, the application of Lemma 4.1 ensures that

$$\begin{aligned}
&Var\left(\tilde{T}_n(x)\right)\\
&= \Pi_n^2(f(x)) \sum_{k=n_0}^{n} \frac{\Pi_k^{-2}(f(x))\,\gamma_k^2}{h_k}\left[Var[Y|X=x] f(x) \int_{\mathbb{R}} K^2(z)\, dz + o(1)\right]\\
&= \Pi_n^2(f(x)) \sum_{k=n_0}^{n} \Pi_k^{-2}(f(x))\,\gamma_k o\left(h_k^4\right) = o\left(h_n^4\right),
\end{aligned}$$

which gives (5.4).

Proof of (5.3)***.*** We have

$$\begin{aligned}
&\mathbb{E}\left(\tilde{T}_n(x)\right)\\
&= \Pi_n(f(x)) \sum_{k=n_0}^{n} \Pi_k^{-1}(f(x))\gamma_k\left[(\mathbb{E}(W_k(x)) - a(x) - r(x)(\mathbb{E}(Z_k(x)) - f(x))\right].
\end{aligned}$$

In view of $(A3)$ we obtain

$$\begin{aligned}\mathbb{E}\left(W_k\left(x\right)\right)-a\left(x\right) &= \frac{1}{2}h_k^2\int_{\mathbb{R}}y\frac{\partial^2 g}{\partial x^2}\left(x,y\right)dy\left[1+o\left(1\right)\right]\int_{\mathbb{R}}z^2K\left(z\right)dz, \quad (5.13)\\ \mathbb{E}\left(Z_k\left(x\right)\right)-f\left(x\right) &= \frac{1}{2}h_k^2\int_{\mathbb{R}}\frac{\partial^2 g}{\partial x^2}\left(x,y\right)dy\left[1+o\left(1\right)\right]\int_{\mathbb{R}}z^2K\left(z\right)dz. \quad (5.14)\end{aligned}$$

Since $\lim_{n\to\infty}\left(n\gamma_n\right)>2a/f\left(x\right)$, it follows from the application of Lemma 4.1 that

$$\begin{aligned}\mathbb{E}\left(\tilde{T}_n\left(x\right)\right) &= \Pi_n\left(f\left(x\right)\right)\sum_{k=n_0}^{n}\Pi_k^{-1}\left(f\left(x\right)\right)\gamma_k h_k^2\int_{\mathbb{R}}z^2K\left(z\right)dz\\ &\quad\left[\frac{1}{2}\left(\int_{\mathbb{R}}y\frac{\partial^2 g}{\partial x^2}\left(x,y\right)dy-r\left(x\right)\int_{\mathbb{R}}\frac{\partial^2 g}{\partial x^2}\left(x,y\right)dy\right)+o\left(1\right)\right]\\ &= \Pi_n\left(f\left(x\right)\right)\sum_{k=n_0}^{n}\Pi_k^{-1}\left(f\left(x\right)\right)\gamma_k h_k^2\left[m^{(2)}\left(x\right)f\left(x\right)+o\left(1\right)\right]\\ &= \frac{1}{f\left(x\right)-2a\xi}h_n^2\left[m^{(2)}\left(x\right)f\left(x\right)+o\left(1\right)\right],\end{aligned}$$

which gives (5.3).

Proof of (5.2)**.** Since $a>\alpha/5$ and $\lim_{n\to\infty}\left(n\gamma_n\right)>\left(1-a\right)/\left(2f\left(x\right)\right)$, we have

$$\begin{aligned}\mathbb{E}\left(\tilde{T}_n\left(x\right)\right) &= \Pi_n\left(f\left(x\right)\right)\sum_{k=n_0}^{n}\Pi_k^{-1}\left(f\left(x\right)\right)\gamma_k o\left(\sqrt{\gamma_k h_k^{-1}}\right)\\ &= o\left(\sqrt{\gamma_n h_n^{-1}}\right),\end{aligned}$$

which gives (5.2).

5.2. *Proof of Lemma 4.6.* Set

$$S_n\left(x\right) = \sum_{k=n_0}^{n}\left[\tilde{\eta}_k\left(x\right)-\mathbb{E}\left(\tilde{\eta}_k\left(x\right)\right)\right],$$

where $\tilde{\eta}_k$ is defined in (5.5).

• Let us first consider the case $a\geq\alpha/5$ (in which case $\lim_{n\to\infty}n\gamma_n>(\alpha-a)/(2f(x))$). We set $H_n^2\left(f\left(x\right)\right)=\Pi_n^2\left(f\left(x\right)\right)\gamma_n^{-1}h_n$, and note that, since $\left(\gamma_n^{-1}h_n\right)\in\mathcal{GS}\left(\alpha-a\right)$, we have

$$\begin{aligned}&\ln\left(H_n^{-2}\left(f\left(x\right)\right)\right)\\ &= -2\ln\left(\Pi_n(f(x))\right)+\ln\left(\prod_{k=n_0}^{n}\frac{\gamma_{k-1}^{-1}h_{k-1}}{\gamma_k^{-1}h_k}\right)+\ln\left(\gamma_{n_0-1}h_{n_0-1}^{-1}\right)\\ &= -2\sum_{k=n_0}^{n}\ln\left(1-f\left(x\right)\gamma_k\right)+\sum_{k=n_0}^{n}\ln\left(1-\frac{\alpha-a}{k}+o\left(\frac{1}{k}\right)\right)+\ln\left(\gamma_{n_0-1}h_{n_0-1}^{-1}\right)\\ &= \sum_{k=n_0}^{n}\left(2f\left(x\right)\gamma_k+o\left(\gamma_k\right)\right)-\sum_{k=n_0}^{n}\left(\left(\alpha-a\right)\xi\gamma_k+o\left(\gamma_k\right)\right)+\ln\left(\gamma_{n_0-1}h_{n_0-1}^{-1}\right)\\ &= \left(2f\left(x\right)-\xi\left(\alpha-a\right)\right)s_n+o\left(s_n\right). \quad (5.15)\end{aligned}$$

Since $2f(x) - \xi(\alpha - a) > 0$, it follows in particular that $\lim_{n\to\infty} H_n^{-2}(f(x)) = \infty$. Moreover, we clearly have $\lim_{n\to\infty} H_n^2(f(x))/H_{n-1}^2(f(x)) = 1$, and by (5.10),

$$\begin{aligned}
&\lim_{n\to\infty} H_n^2(f(x)) \sum_{k=n_0}^{n} Var[\tilde{\eta}_k(x)] \\
&= [2f(x) - (\alpha - a)\xi]^{-1} Var[Y|X=x] f(x) \int_{\mathbb{R}} K^2(z)\,dz
\end{aligned}$$

and, in view of (5.11),

$$\mathbb{E}\left[|\tilde{\eta}_n(x)|^3\right] = O\left(\Pi_n^{-3}(f(x))\gamma_n^3 h_n^{-2}\right).$$

Now, since $\left(\gamma_n^{-1}h_n\right) \in \mathcal{GS}(\alpha - a)$, applying Lemma 4.1 and (5.15), we get

$$\begin{aligned}
&\frac{1}{n\sqrt{n}} \sum_{k=n_0}^{n} \mathbb{E}\left(|H_n(f(x))\tilde{\eta}_k(x)|^3\right) \\
&= O\left(\frac{H_n^3(f(x))}{n\sqrt{n}}\left(\sum_{k=n_0}^{n} \frac{\Pi_k^{-3}(f(x))\gamma_k^3}{h_k^2}\right)\right) \\
&= O\left(\frac{\Pi_n^3(f(x))\gamma_n^{-\frac{3}{2}}h_n^{\frac{3}{2}}}{n\sqrt{n}}\left(\sum_{k=n_0}^{n} \Pi_k^{-3}(f(x))\gamma_k o\left(\left(\gamma_k h_k^{-1}\right)^{\frac{3}{2}}\right)\right)\right) \\
&= o\left(\frac{1}{n\sqrt{n}}\right) \\
&= o\left(\left[\ln\left(H_n^{-2}(f(x))\right)\right]^{-1}\right).
\end{aligned}$$

The application of Theorem 1 in Mokkadem and Pelletier (2008) then ensures that, with probability one, the sequence

$$\left(\frac{H_n(f(x))S_n(x)}{\sqrt{2\ln\ln\left(H_n^{-2}(f(x))\right)}}\right) = \left(\frac{\sqrt{\gamma_n^{-1}h_n}\left(\tilde{T}_n(x) - \mathbb{E}\left(\tilde{T}_n(x)\right)\right)}{\sqrt{2\ln\ln\left(H_n^{-2}(f(x))\right)}}\right)$$

is relatively compact and its limit set is the interval

$$\left[-\sqrt{\frac{Var[Y|X=x]f(x)}{2f(x)-(\alpha-a)\xi}\int_{\mathbb{R}} K^2(z)dz}, \sqrt{\frac{Var[Y|X=x]f(x)}{2f(x)-(\alpha-a)\xi}\int_{\mathbb{R}} K^2(z)dz}\right]. \quad (5.16)$$

In view of (5.15), we have $\lim_{n\to\infty} \ln\ln\left(H_n^{-2}(f(x))\right)/\ln s_n = 1$. It follows that, with probability one, the sequence $\left(\sqrt{\gamma_n^{-1}h_n}\left(\tilde{T}_n(x) - \mathbb{E}\left(\tilde{T}_n(x)\right)\right)/\sqrt{2\ln s_n}\right)$ is relatively compact, and its limit set is the interval given in (5.16). The application of (5.2) (respectively of (5.3)) concludes the proof of Lemma 4.6 in the case $a > \alpha/5$ (respectively $a = \alpha/5$).

• Let us now consider the case $a < \alpha/5$ (in which case $\lim_{n\to\infty}(n\gamma_n) > 2a/f(x)$). Set $H_n^{-2}(f(x)) = \Pi_n^{-2}(f(x))h_n^4\left(\ln\ln\left(\Pi_n^{-2}(f(x))h_n^4\right)\right)^{-1}$, and note that, since

$\left(h_n^{-4}\right) \in \mathcal{GS}\left(4a\right)$, we have

$$\begin{aligned}
&\ln\left(\Pi_n^{-2}\left(f\left(x\right)\right)h_n^4\right)\\
&= -2\ln\left(\Pi_n\left(f\left(x\right)\right)\right)+\ln\left(\prod_{k=n_0}^{n}\frac{h_{k-1}^{-4}}{h_k^{-4}}\right)+\ln\left(h_{n_0-1}^4\right)\\
&= -2\sum_{k=n_0}^{n}\ln\left(1-\gamma_k f\left(x\right)\right)+\sum_{k=n_0}^{n}\ln\left(1-\frac{4a}{k}+o\left(\frac{1}{k}\right)\right)+\ln\left(h_{n_0-1}^4\right)\\
&= \sum_{k=n_0}^{n}\left(2\gamma_k f\left(x\right)+o\left(\gamma_k\right)\right)-\sum_{k=n_0}^{n}\left(4a\xi\gamma_k+o\left(\gamma_k\right)\right)+\ln\left(h_{n_0-1}^4\right)\\
&= \left(2f\left(x\right)-4a\xi\right)s_n+o\left(s_n\right).
\end{aligned} \tag{5.17}$$

Since $2f(x)-4a\xi>0$, it follows in particular that $\lim_{n\to\infty}\Pi_n^{-2}(f(x))h_n^4=\infty$, and thus $\lim_{n\to\infty}H_n^{-2}(f(x))=\infty$. Moreover, we clearly have $\lim_{n\to\infty}H_n^2(f(x))/H_{n-1}^2(f(x))=1$. Now, set $\epsilon\in]0,\alpha-5a[$ such that $\lim_{n\to\infty}n\gamma_n>2a/f(x)+\epsilon/2$; in view of (5.10), and applying Lemma 4.1, we get

$$\begin{aligned}
&H_n^2\left(f\left(x\right)\right)\sum_{k=n_0}^{n}Var\left[\tilde{\eta}_k\left(x\right)\right]\\
&= O\left(\Pi_n^2\left(f\left(x\right)\right)h_n^{-4}\ln\ln\left(\Pi_n^{-2}\left(f\left(x\right)\right)h_n^4\right)\sum_{k=n_0}^{n}\frac{\Pi_k^{-2}\left(f\left(x\right)\right)\gamma_k^2}{h_k}\right)\\
&= O\left(\Pi_n^2\left(f\left(x\right)\right)h_n^{-4}\ln\ln\left(\Pi_n^{-2}\left(f\left(x\right)\right)h_n^4\right)\sum_{k=n_0}^{n}\Pi_k^{-2}\left(f\left(x\right)\right)\gamma_k o\left(h_k^4k^{-\epsilon}\right)\right)\\
&= o\left(1\right).
\end{aligned}$$

Moreover, in view of (5.11) we have

$$\mathbb{E}\left[\left|\tilde{\eta}_n\left(x\right)\right|^3\right]=O\left(\Pi_n^{-3}\left(f\left(x\right)\right)\gamma_n^3h_n^{-2}\right),$$

and thus in view of (5.17), we get

$$\begin{aligned}
&\frac{1}{n\sqrt{n}}\sum_{k=n_0}^{n}\mathbb{E}\left(\left|H_n\left(f\left(x\right)\right)\tilde{\eta}_k\left(x\right)\right|^3\right)\\
&= O\left(\frac{\Pi_n^3\left(f\left(x\right)\right)h_n^{-6}}{n\sqrt{n}}\left(\ln\ln\left(\Pi_n^{-2}\left(f\left(x\right)\right)h_n^4\right)\right)^{\frac{3}{2}}\left(\sum_{k=n_0}^{n}\frac{\Pi_k^{-3}\left(f\left(x\right)\right)\gamma_k^3}{h_k^2}\right)\right)\\
&= O\left(\frac{\Pi_n^3\left(f\left(x\right)\right)h_n^{-6}}{n\sqrt{n}}\left(\ln\ln\left(\Pi_n^{-2}\left(f\left(x\right)\right)h_n^4\right)\right)^{\frac{3}{2}}\left(\sum_{k=n_0}^{n}\Pi_k^{-3}\left(f\left(x\right)\right)\gamma_k o\left(h_k^6\right)\right)\right)\\
&= o\left(\frac{\left(\ln\ln\left(\Pi_n^{-2}\left(f\left(x\right)\right)h_n^4\right)\right)^{\frac{3}{2}}}{n\sqrt{n}}\right)\\
&= o\left(\left[\ln\left(H_n^{-2}\left(f\left(x\right)\right)\right)\right]^{-1}\right).
\end{aligned}$$

The application of Theorem 1 in Mokkadem and Pelletier (2008) then ensures that, with probability one,

$$\lim_{n\to\infty}\frac{H_n\left(f\left(x\right)\right)S_n\left(x\right)}{\sqrt{2\ln\ln\left(H_n^{-2}\left(f\left(x\right)\right)\right)}} = \lim_{n\to\infty} h_n^{-2}\frac{\sqrt{\ln\ln\left(\Pi_n^{-2}\left(f\left(x\right)\right)h_n^4\right)}}{\sqrt{2\ln\ln\left(H_n^{-2}\left(f\left(x\right)\right)\right)}}\left(\tilde{T}_n\left(x\right)-\mathbb{E}\left(\tilde{T}_n\left(x\right)\right)\right)=0.$$

Noting that (5.17) ensures that

$$\lim_{n\to\infty}\ln\ln\left(H_n^{-2}\left(f\left(x\right)\right)\right)/\ln\ln\left(\Pi_n^{-2}\left(f\left(x\right)\right)h_n^4\right)=1,$$

we get

$$\lim_{n\to\infty}h_n^{-2}\left[\tilde{T}_n\left(x\right)-\mathbb{E}\left(\tilde{T}_n\left(x\right)\right)\right]=0 \quad a.s.$$

and Lemma 4.6 in the case $a<\alpha/5$ follows from (5.3).

5.3. *Proof of Lemma 4.7.* Let us write

$$\tilde{T}_n\left(x\right)=\tilde{T}_n^{(1)}\left(x\right)-r\left(x\right)\tilde{T}_n^{(2)}\left(x\right),$$

with

$$\begin{aligned}\tilde{T}_n^{(1)}\left(x\right) &= \sum_{k=n_0}^{n}U_{k,n}\left(f\left(x\right)\right)\gamma_k h_k^{-1}Y_kK\left(\frac{x-X_k}{h_k}\right),\\ \tilde{T}_n^{(2)}\left(x\right) &= \sum_{k=n_0}^{n}U_{k,n}\left(f\left(x\right)\right)\gamma_k h_k^{-1}K\left(\frac{x-X_k}{h_k}\right).\end{aligned}$$

Lemma 4.7 is proved by showing that, for $i\in\{1,2\}$, under the condition $\lim_{n\to\infty} n\gamma_n>(\alpha-a)/(2\varphi)$,

- if $a\geq\alpha/5$, then $\sup_{x\in I}\left|\tilde{T}_n^{(i)}(x)-\mathbb{E}\left(\tilde{T}_n^{(i)}(x)\right)\right|=O\left(\sqrt{\gamma_nh_n^{-1}}\,\ln n\right)$ $a.s.$, (5.18)
- if $a>\alpha/5$, then $\sup_{x\in I}\left|\mathbb{E}\left(\tilde{T}_n\left(x\right)\right)\right|=o\left(\sqrt{\gamma_nh_n^{-1}}\,\ln n\right)$ (5.19)

and by proving that, under the condition $\lim_{n\to\infty}\left(n\gamma_n\right)>2a/\varphi$,

- if $a<\alpha/5$, then $\sup_{x\in I}\left|\tilde{T}_n^{(i)}(x)-\mathbb{E}\left(\tilde{T}_n^{(i)}(x)\right)\right|=o\left(h_n^2\right)$ $a.s.$, (5.20)
- if $a\leq\alpha/5$, then $\sup_{x\in I}\mathbb{E}\left(\tilde{T}_n\left(x\right)\right)=O\left(h_n^2\right)$. (5.21)

As a matter of fact, Lemma 4.7 follows from the combination of (5.18) and (5.19) in the case $a>\alpha/5$, from that of (5.18) and (5.21) in the case $a=\alpha/5$, and from that of (5.20) and (5.21) in the case $a<\alpha/5$.

The proof of (5.19) and (5.21) is similar to that of (5.2) and (5.3), and is omitted. Moreover the proof of (5.18) and (5.20) for $i=2$ is similar to that for $i=1$, and is omitted too. To prove simultaneously (5.18) and (5.20) for $i=1$, we introduce the sequence (v_n) defined as

$$(v_n) = \begin{cases} \left(\sqrt{\gamma_n^{-1} h_n}\right) & \text{if } a \geq \alpha/5, \\ \left(h_n^{-2} (\ln n)^2\right) & \text{if } a < \alpha/5. \end{cases} \tag{5.22}$$

As a matter of fact, (5.18) and (5.20) are proved for $i=1$ by establishing that

$$\sup_{x \in I} \left| \tilde{T}_n^{(1)}(x) - \mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right) \right| = O\left(v_n^{-1} \ln n\right) \quad a.s. \tag{5.23}$$

To this end we first state the following lemma.

Lemma 5.1. *There exists $s>0$ such that, for all $C>0$,*

$$\sup_{x \in I} \mathbb{P}\left[\frac{v_n}{\ln n} \left| \tilde{T}_n^{(1)}(x) - \mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right) \right| \geq C \right] = O\left(n^{-\frac{C}{s}}\right).$$

We first show how (5.23) can be deduced from Lemma 5.1. Set $(M_n) \in \mathcal{GS}(\tilde{m})$ with $\tilde{m} > 0$, and note that, for all $C>0$, we have

$$\begin{aligned} &\mathbb{P}\left[\frac{v_n}{\ln n} \sup_{x \in I} \left| \tilde{T}_n^{(1)}(x) - \mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right) \right| \geq C \right] \\ &\leq \mathbb{P}\left[\frac{v_n}{\ln n} \sup_{x \in I} \left| \tilde{T}_n^{(1)}(x) - \mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right) \right| \geq C \text{ and } \sup_{k \leq n} |Y_k| \leq M_n \right] \\ &\quad + \mathbb{P}\left[\sup_{k \leq n} |Y_k| \geq M_n \right]. \end{aligned} \tag{5.24}$$

Let (d_n) be a positive sequence satisfying $d_n < 1$ for all n and $\lim_{n\to\infty} \gamma_n^{-1} d_n = 0$. Let $I_i^{(n)}$ be $N(n)$ intervals of length d_n such that $\cup_{i=1}^{N(n)} I_i^{(n)} = I$, and for all $i \in \{1, \dots, N(n)\}$, set $x_i^{(n)} \in I_i^{(n)}$. We have

$$\begin{aligned} &\mathbb{P}\left[\frac{v_n}{\ln n} \sup_{x \in I} \left| \tilde{T}_n^{(1)}(x) - \mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right) \right| \geq C \text{ and } \sup_{k \leq n} |Y_k| \leq M_n \right] \\ &\leq \sum_{i=1}^{N(n)} \mathbb{P}\left[\frac{v_n}{\ln n} \sup_{x \in I_i^{(n)}} \left| \tilde{T}_n^{(1)}(x) - \mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right) \right| \geq C \text{ and } \sup_{k \leq n} |Y_k| \leq M_n \right]. \end{aligned}$$

Let us prove that there exists c^* such that, for all $x, y \in I$ such that $|x-y| \leq d_n$, and on $\left\{\sup_{k \leq n} |Y_k| \leq M_n\right\}$,

$$\left| \tilde{T}_n^{(1)}(x) - \tilde{T}_n^{(1)}(y) \right| \leq c^* M_n h_n^{-1} \gamma_n^{-1} d_n. \tag{5.25}$$

To this end, we write

$$\left| \tilde{T}_n^{(1)}(x) - \tilde{T}_n^{(1)}(y) \right| \leq A_{n,1}(x,y) + A_{n,2}(x,y),$$

with

$$\begin{aligned} A_{n,1}(x,y) &= \left| \sum_{k=n_0}^{n} U_{k,n}(f(x)) \gamma_k h_k^{-1} Y_k \left[K\left(\frac{x - X_k}{h_k}\right) - K\left(\frac{y - X_k}{h_k}\right) \right] \right|, \\ A_{n,2}(x,y) &= \left| \sum_{k=n_0}^{n} U_{k,n}(f(y)) \gamma_k h_k^{-1} Y_k K\left(\frac{y - X_k}{h_k}\right) \left[\frac{U_{k,n}(f(x))}{U_{k,n}(f(y))} - 1 \right] \right|. \end{aligned}$$

• Since K is Lipschitz-continuous, and by application of Lemma 4.1, there exist $k^*, c_1^* > 0$ such that, for all $x, y \in \mathbb{R}$ satisfying $|x-y| \leq d_n$ and on $\{\sup_{k\leq n}|Y_k| \leq M_n\}$, we have:

$$\begin{aligned} A_{n,1}(x,y) &\leq k^* M_n \sum_{k=n_0}^{n} U_{k,n}(\varphi)\gamma_k h_k^{-2} d_n \\ &\leq c_1^* M_n h_n^{-2} d_n. \end{aligned} \quad (5.26)$$

• Now, let c_i^* be positive constants; for $j \geq n_0$ and for $p \in \{1,2\}$ we have

$$\begin{aligned} \left(\frac{1-\gamma_j f(x)}{1-\gamma_j f(y)}\right)^p &= \left(1 + \frac{\gamma_j (f(y) - f(x))}{1-\gamma_j f(y)}\right)^p \\ &\leq \left(1 + \frac{c_2^* \gamma_j d_n}{1-\gamma_j \|f\|_\infty}\right)^p \\ &\leq (1 + 2c_2^* \gamma_j d_n)^p \\ &\leq 1 + c_3^* \gamma_j d_n. \end{aligned} \quad (5.27)$$

We deduce that, for k and n such that $n_0 \leq k \leq n$,

$$\begin{aligned} \left[\frac{U_{k,n}^p(f(x))}{U_{k,n}^p(f(y))} - 1\right] &= \left(\prod_{j=k+1}^{n} \frac{1-\gamma_j f(x)}{1-\gamma_j f(y)}\right)^p - 1 \\ &\leq \left(\prod_{j=k+1}^{n} \exp(c_3^* \gamma_j d_n)\right) - 1 \\ &\leq \exp\left[c_3^* d_n \gamma_n^{-1} \sum_{j=k+1}^{n} \gamma_j \gamma_n\right] - 1 \\ &\leq \exp\left[c_4^* d_n \gamma_n^{-1} \sum_{n\geq 1} \gamma_n^2\right] - 1 \\ &\leq \exp\left[c_5^* d_n \gamma_n^{-1}\right] - 1 \\ &\leq c_5^* d_n \gamma_n^{-1} \exp\left[c_5^* d_n \gamma_n^{-1}\right] \\ &\leq c_6^* d_n \gamma_n^{-1}. \end{aligned} \quad (5.28)$$

The application of (5.28) with $p = 1$ and of Lemma 4.1 ensures that, for all $x, y \in I$ satisfying $|x-y| \leq d_n$, and on $\left\{\sup_{k\leq n}|Y_k| \leq M_n\right\}$, we have:

$$\begin{aligned} A_{n,2}(x,y) &\leq \|K\|_\infty M_n \sum_{k=n_0}^{n} U_{k,n}(\varphi)\gamma_k h_k^{-1}\left(c_6^* d_n \gamma_n^{-1}\right) \\ &\leq c_7^* M_n h_n^{-1} \gamma_n^{-1} d_n. \end{aligned} \quad (5.29)$$

The upper bound (5.25) follows from the combination of (5.26) and (5.29).
Now, it follows from (5.25) that, for all $x \in I_i^{(n)}$ and on $\left\{\sup_{k\leq n}|Y_k| \leq M_n\right\}$, we

have

$$\begin{aligned}
&\left|\tilde{T}_n^{(1)}(x)-\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)\right| \\
&\leq \left|\tilde{T}_n^{(1)}(x)-\tilde{T}_n^{(1)}\left(x_i^{(n)}\right)\right|+\left|\tilde{T}_n^{(1)}\left(x_i^{(n)}\right)-\mathbb{E}\left(\tilde{T}_n^{(1)}\left(x_i^{(n)}\right)\right)\right| \\
&\quad+\left|\mathbb{E}\left(\tilde{T}_n^{(1)}\left(x_i^{(n)}\right)\right)-\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)\right| \\
&\leq 2c^*M_nh_n^{-1}\gamma_n^{-1}d_n+\left|\tilde{T}_n^{(1)}\left(x_i^{(n)}\right)-\mathbb{E}\left(\tilde{T}_n^{(1)}\left(x_i^{(n)}\right)\right)\right|.
\end{aligned}$$

In view of (5.24), we obtain, for all $C>0$,

$$\begin{aligned}
&\mathbb{P}\left[\frac{v_n}{\ln n}\sup_{x\in I}\left|\tilde{T}_n^{(1)}(x)-\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)\right|\geq C\right] \\
&\leq \sum_{i=1}^{N(n)}\mathbb{P}\left[\frac{v_n}{\ln n}\left|T_n^{(1)}\left(x_i^{(n)}\right)-\mathbb{E}\left(T_n^{(1)}\left(x_i^{(n)}\right)\right)\right|+2c^*M_n\frac{v_n}{\ln n}h_n^{-1}\gamma_n^{-1}d_n\geq C\right] \\
&\quad+n\mathbb{P}\left[|Y|\geq M_n\right].
\end{aligned}$$

Now, note that, in view of (5.22), $(v_n)\in\mathcal{GS}(m^*)$ where m^* is defined in (4.3). Set $(d_n)\in\mathcal{GS}(-(\tilde{m}+m^*+a+\alpha))$ such that, for all n, $2c^*M_nv_n(\ln n)^{-1}h_n^{-1}\gamma_n^{-1}d_n\leq C/2$; in view of Lemma 5.1 and Assumption $(A4)\,ii)$, there exists $s>0$ such that

$$\begin{aligned}
&\mathbb{P}\left[\frac{v_n}{\ln n}\sup_{x\in I}\left|\tilde{T}_n^{(1)}(x)-\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)\right|\geq C\right] \\
&\leq N(n)\sup_{x\in I}\mathbb{P}\left[\frac{v_n}{\ln n}\left|\tilde{T}_n^{(1)}(x)-\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)\right|\geq\frac{C}{2}\right] \\
&\quad+n\exp(-t^*M_n)\,\mathbb{E}\left(\exp(t^*|Y|)\right) \\
&= O\left(d_n^{-1}n^{-\frac{C}{2s}}+n\exp(-t^*M_n)\right).
\end{aligned}$$

Since $(M_n)\in\mathcal{GS}(\tilde{m})$ with $\tilde{m}>0$, and since $\left(d_n^{-1}\right)\in\mathcal{GS}(\tilde{m}+m^*+a+\alpha)$, we can choose C large enough so that

$$\sum_{n\geq 0}\mathbb{P}\left[\frac{v_n}{\ln n}\sup_{x\in I}\left|\tilde{T}_n^{(1)}(x)-\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)\right|\geq C\right]<\infty,$$

which gives (5.23).
It remains to prove Lemma 5.1. For all $x\in I$ and all $s>0$, we have

$$\begin{aligned}
&\mathbb{P}\left[\frac{v_n}{\ln n}\left(\tilde{T}_n^{(1)}(x)-\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)\right)\geq C\right] \\
&= \mathbb{P}\left[\exp\left[s^{-1}v_n\left(\tilde{T}_n^{(1)}(x)-\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)\right)\right]\geq n^{\frac{C}{s}}\right] \\
&\leq n^{-\frac{C}{s}}\mathbb{E}\left(\exp\left[s^{-1}v_n\left(\tilde{T}_n^{(1)}(x)-\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)\right)\right]\right) \\
&\leq n^{-\frac{C}{s}}\prod_{k=n_0}^{n}\mathbb{E}\left(\exp\left(s^{-1}V_{k,n}(x)\right)\right), \qquad (5.30)
\end{aligned}$$

with

$$V_{k,n}(x) = v_nU_{k,n}(f(x))\gamma_kh_k^{-1}\left[Y_kK\left(\frac{x-X_k}{h_k}\right)-\mathbb{E}\left(Y_kK\left(\frac{x-X_k}{h_k}\right)\right)\right].$$

For k and n such that $n_0 \leq k \leq n$, set

$$\alpha_{k,n} = v_n U_{k,n}(\varphi)\gamma_k h_k^{-1}.$$

We have, for all $x \in I$,

$$\begin{aligned}
&\mathbb{E}\left(\exp\left[s^{-1}V_{k,n}(x)\right]\right)\\
&\leq 1+\frac{1}{2}\mathbb{E}\left[s^{-2}V_{k,n}^2(x)\right]+\mathbb{E}\left(s^{-3}\left|V_{k,n}^3(x)\right|\right)\exp\left[\left|V_{k,n}(x)\right|\right]\\
&\leq 1+\frac{1}{2}s^{-2}\alpha_{k,n}^2 Var\left[Y_k K\left(\frac{x-X_k}{h_k}\right)\right]\\
&\quad +s^{-3}\alpha_{k,n}^3\|K\|_\infty^3\mathbb{E}\left[\left(|Y_k|^3+\left(\mathbb{E}\left(|Y_k|\right)\right)^3\right)\exp\left(s^{-1}\alpha_{k,n}\|K\|_\infty\left(|Y_k|+\mathbb{E}\left(|Y_k|\right)\right)\right)\right].
\end{aligned}$$

Now, note that $\alpha_{k,n}$ can be rewritten as:

$$\alpha_{k,n} = \frac{v_n\Pi_n(\varphi)}{v_k\Pi_k(\varphi)}v_k\gamma_k h_k^{-1}.$$

Since $(v_n) \in \mathcal{GS}(m^*)$ with $\varphi - m^*\xi > 0$ (where ξ is defined in (2.2)), we have

$$\begin{aligned}
\frac{\Pi_n(\varphi)}{\Pi_{n-1}(\varphi)}\frac{v_n}{v_{n-1}} &= (1-\gamma_n\varphi)\left(1+m^*\frac{1}{n}+o\left(\frac{1}{n}\right)\right)\\
&= (1-\gamma_n\varphi)\left(1+m^*\xi\gamma_n+o(\gamma_n)\right)\\
&= 1-(\varphi-m^*\xi)\gamma_n+o(\gamma_n)\\
&\leq 1 \qquad \text{for } n \text{ large enough.}
\end{aligned}$$

Writing

$$\frac{v_n\Pi_n(\varphi)}{v_k\Pi_k(\varphi)} = \prod_{i=k}^{n-1}\frac{v_{i+1}\Pi_{i+1}(\varphi)}{v_i\Pi_i(\varphi)},$$

we obtain

$$\sup_{n_0\leq k\leq n}\frac{v_n\Pi_n(\varphi)}{v_k\Pi_k(\varphi)} < \infty$$

and, since $\lim_{n\to\infty} v_k\gamma_k h_k = 0$, we deduce that $\sup_{n_0\leq k\leq n}\alpha_{k,n} < \infty$. Thus, in view of Assumption $(A4)\,ii)$, there exist $s > 0$ and $c^* > 0$ such that, for all k and n such that $n_0 \leq k \leq n$,

$$\mathbb{E}\left[\left(|Y_k|^3+\left(\mathbb{E}\left(|Y_k|\right)\right)^3\right)\exp\left(s^{-1}\alpha_{k,n}\|K\|_\infty\left(|Y_k|+\mathbb{E}\left(|Y_k|\right)\right)\right)\right] \leq c^*.$$

From classical computations, we have $\sup_{x\in I} Var\left[Y_k K\left((x-X_k)h_k^{-1}\right)\right] = O(h_k)$. We then deduce that there exist $C_1^*, C_2^* > 0$ such that, for all $x \in I$, for all k and n such that $n_0 \leq k \leq n$,

$$\begin{aligned}
\mathbb{E}\left(\exp\left[s^{-1}V_{k,n}(x)\right]\right) &\leq 1+C_1^*v_n^2U_{n,k}^2(\varphi)\gamma_k^2h_k^{-1}+C_2^*v_n^3U_{k,n}^3(\varphi)\gamma_k^3h_k^{-3}\\
&\leq \exp\left[C_1^*v_n^2U_{k,n}^2(\varphi)\gamma_k^2h_k^{-1}+C_2^*v_n^3U_{k,n}^3(\varphi)\gamma_k^3h_k^{-3}\right].
\end{aligned}$$

Applying Lemma 4.1, we deduce from (5.30) that, for all $C > 0$,

$$\begin{aligned}
&\sup_{x\in I}\mathbb{P}\left[\frac{v_n}{\ln n}\left(\tilde{T}_n^{(1)}(x)-\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)\right)\geq C\right]\\
&\leq\quad n^{-\frac{C}{s}}\exp\left[C_1^*v_n^2\sum_{k=n_0}^{n}U_{k,n}^2(\varphi)\gamma_kO\left(v_k^{-2}\right)+C_2^*v_n^3\sum_{k=n_0}^{n}U_{k,n}^3(\varphi)\gamma_kO\left(v_k^{-3}\right)\right]\\
&=\quad O\left(n^{-\frac{C}{s}}\right).
\end{aligned}$$

We establish exactly in the same way that, for all $C > 0$,

$$\sup_{x\in I}\mathbb{P}\left[\frac{v_n}{\ln n}\left(\mathbb{E}\left(\tilde{T}_n^{(1)}(x)\right)-\tilde{T}_n^{(1)}(x)\right)\geq C\right]\quad=\quad O\left(n^{-\frac{C}{s}}\right),$$

which concludes the proof of Lemma 5.1.

5.4. *Proof of Lemma 4.11.* Set

$$\eta_k(x)=\left(W_{k+1}(x)-r(x)Z_{k+1}(x)\right).\tag{5.31}$$

In order to prove Lemma 4.11, we first establish a central limit theorem for

$$T_n(x)-\mathbb{E}\left(T_n(x)\right)=\frac{1}{\sum_{k=1}^{n}q_k}\sum_{k=n_0-1}^{n}q_k\left[\eta_k(x)-\mathbb{E}\left(\eta_k(x)\right)\right].$$

In view of (5.6)-(5.8) and since $h_k/h_{k+1}=1+o(1)$, we have

$$\begin{aligned}
&Var\left(\eta_k(x)\right)\\
&=\quad Var\left(W_{k+1}(x)\right)+r^2(x)Var\left(Z_{k+1}(x)\right)-2r(x)Cov\left(W_{k+1}(x),Z_{k+1}(x)\right)\\
&=\quad\frac{1}{h_k}\left[Var[Y|X=x]f(x)\int_{\mathbb{R}}K^2(z)dz+o(1)\right].
\end{aligned}$$

Noting that $\left(q_n^2h_n^{-1}\right)\in\mathcal{GS}\left(-2q+a\right)$ with $q<(1+a)/2$, and using (4.12), we get

$$\begin{aligned}
v_n^2\quad&=\quad\sum_{k=n_0-1}^{n}q_k^2Var\left(\eta_k(x)\right)\\
&=\quad\sum_{k=n_0-1}^{n}\frac{q_k^2}{h_k}\left[Var\left[Y|X=x\right]f(x)\int_{\mathbb{R}}K^2(z)\,dz+o(1)\right]\\
&=\quad\frac{nq_n^2h_n^{-1}}{1-2q+a}\left[Var\left[Y|X=x\right]f(x)\int_{\mathbb{R}}K^2(z)\,dz+o(1)\right].
\end{aligned}\tag{5.32}$$

Now, set $p\in\,]0,1]$ such that $q<\left(1+a\left(1+p\right)\right)/\left(2+p\right)$; it follows from (5.11) that

$$\begin{aligned}
&\sum_{k=n_0-1}^{n}q_k^{2+p}\mathbb{E}\left[\left|\eta_k(x)\right|^{2+p}\right]\\
&=\quad O\left(\sum_{k=n_0-1}^{n}\frac{q_k^{2+p}}{h_k^{2+p}}\mathbb{E}\left(\left|Y_k-r(x)\right|^{2+p}K^{2+p}\left(\frac{x-X_k}{h_k}\right)\right)\right)\\
&=\quad O\left(\sum_{k=n_0-1}^{n}\frac{q_k^{2+p}}{h_k^{1+p}}\right).
\end{aligned}$$

In view of (5.32) and using (4.12), we get

$$\frac{1}{v_n^{2+p}}\sum_{k=n_0-1}^{n} q_k^{2+p}\mathbb{E}\left[\left|\eta_k\left(x\right)\right|^{2+p}\right] = O\left(\frac{nq_n^{2+p}h_n^{-(1+p)}}{\left(nq_n^2h_n^{-1}\right)^{1+\frac{p}{2}}}\right) = O\left(\frac{1}{n^{\frac{p}{2}}h_n^{\frac{p}{2}}}\right) = o\left(1\right).$$

The application of Lyapounov Theorem gives

$$\frac{\sum_{k=1}^{n} q_k}{\sqrt{nq_n^2h_n^{-1}}}\left(T_n\left(x\right)-\mathbb{E}\left[T_n\left(x\right)\right]\right) \stackrel{\mathcal{D}}{\rightarrow} \mathcal{N}\left(0, \frac{1}{1+a-2q}Var\left[Y|X=x\right]f\left(x\right)\int_{\mathbb{R}}K^2\left(z\right)dz\right)$$

and applying (4.12), we obtain

$$\sqrt{nh_n}\left(T_n(x)-\mathbb{E}\left[T_n(x)\right]\right) \stackrel{\mathcal{D}}{\rightarrow} \mathcal{N}\left(0, \frac{(1-q)^2}{1+a-2q}Var[Y|X=x]f(x)\int_{\mathbb{R}}K^2(z)dz\right). \tag{5.33}$$

Now, note that

$$\mathbb{E}\left(T_n(x)\right) = \frac{1}{\sum_{k=1}^{n} q_k}\sum_{k=n_0-1}^{n} q_k\left[\left(\mathbb{E}\left(W_{k+1}(x)\right)-a(x)\right)-r(x)\left(\mathbb{E}\left(Z_{k+1}(x)\right)-f(x)\right)\right].$$

Since $h_{n+1}/h_n = 1+o\left(1\right)$, it follows from (5.13) and (5.14) that

$$\begin{aligned}\mathbb{E}\left(T_n\left(x\right)\right) &= \frac{1}{\sum_{k=1}^{n} q_k}\sum_{k=n_0-1}^{n} q_kh_{k+1}^2\int_{\mathbb{R}}z^2K\left(z\right)dz \\ &\quad \left[\frac{1}{2}\left(\int_{\mathbb{R}}y\frac{\partial^2 g}{\partial x^2}\left(x,y\right)dy - r\left(x\right)\int_{\mathbb{R}}\frac{\partial^2 g}{\partial x^2}\left(x,y\right)dy\right)+o\left(1\right)\right] \\ &= \frac{1}{\sum_{k=1}^{n} q_k}\sum_{k=n_0-1}^{n} q_kh_k^2\left[m^{(2)}\left(x\right)f\left(x\right)+o\left(1\right)\right].\end{aligned}$$

Applying (4.12), we obtain

$$\lim_{n\to\infty}\frac{1}{h_n^2}\mathbb{E}\left(T_n\left(x\right)\right) = \frac{1-q}{1-2a-q}m^{(2)}\left(x\right)f\left(x\right) \tag{5.34}$$

and Lemma 4.11 follows from the combination of (5.33) and (5.34).

5.5. *Proof of Lemma 4.12.* Set

$$S_n\left(x\right) = \sum_{k=n_0-1}^{n} q_k\left[\eta_k\left(x\right)-\mathbb{E}\left(\eta_k\left(x\right)\right)\right]$$

where η_k is defined in (5.31), and $H_n^{-2} = nh_n^{-1}q_n^2$. Let us first note that, since $\left(nh_n^{-1}q_n^2\right) \in \mathcal{GS}\left(1+a-2q\right)$ with $1+a-2q > 0$, we have $\lim_{n\to\infty} H_n^{-2} = \infty$.

Moreover, we have $\lim_{n\to\infty} H_n^2/H_{n-1}^2 = 1$ and, by (5.32),

$$\lim_{n\to\infty} H_n^2 \sum_{k=n_0-1}^{n} q_k^2 Var\left[\eta_k(x)\right] = \left[1+a-2q\right]^{-1} Var\left[Y|X=x\right] f(x) \int_{\mathbb{R}} K^2(z)\,dz$$

and, by (5.11), $\mathbb{E}\left[\left|q_n\eta_n(x)\right|^3\right] = O\left(q_n^3 h_n^{-2}\right)$. Since, for all $\epsilon > 0$,

$$\begin{aligned}
\frac{1}{n\sqrt{n}} \sum_{k=n_0-1}^{n} \mathbb{E}\left(\left|H_n q_k \eta_k(x)\right|^3\right) &= O\left(\frac{H_n^3}{n\sqrt{n}} \sum_{k=n_0-1}^{n} \frac{q_k^3}{h_k^2}\right) \\
&= O\left(n^{-3} h_n^{\frac{3}{2}} q_n^{-3}\left(n^{\epsilon} + n q_n^3 h_n^{-2}\right)\right),
\end{aligned}$$

we have

$$\frac{1}{n\sqrt{n}} \sum_{k=n_0-1}^{n} \mathbb{E}\left(\left|H_n q_k \eta_k(x)\right|^3\right) = o\left(\left[\ln\left(H_n^{-2}\right)\right]^{-1}\right).$$

The application of Theorem 1 of Mokkadem and Pelletier (2008) ensures that, with probability one, the sequence

$$\left(\frac{H_n S_n(x)}{\sqrt{2\ln\ln\left(H_n^{-2}\right)}}\right) = \left(\frac{\sum_{k=1}^{n} q_k}{n q_n} \frac{\sqrt{n h_n}\left(T_n(x) - \mathbb{E}\left(T_n(x)\right)\right)}{\sqrt{2\ln\ln\left(H_n^{-2}\right)}}\right)$$

is relatively compact and its limit set is the interval

$$\left[-\sqrt{\frac{1}{1+a-2q}\sigma(x)}, \sqrt{\frac{1}{1+a-2q}\sigma(x)}\right],$$

with $\sigma(x) = Var\left[Y|X=x\right] f(x) \int_{\mathbb{R}} K^2(z)\,dz$. Since $\lim_{n\to\infty} \ln\ln\left(H_n^{-2}\right)/\ln\ln n = 1$, and using (4.12), it follows that, with probability one, the sequence $(\sqrt{nh_n}(T_n(x) - \mathbb{E}(T_n(x)))/\sqrt{2\ln\ln n})$ is relatively compact, and its limit set is the interval

$$\left[-\sqrt{\frac{(1-q)^2}{1+a-2q}\sigma(x)}, \sqrt{\frac{(1-q)^2}{1+a-2q}\sigma(x)}\right].$$

The application of (5.34) concludes the proof of Lemma 4.12.

5.6. *Proof of Lemma 4.13.* Let us write $T_n(x)$ as

$$T_n(x) = T_{n,1}(x) - r(x)\,T_{n,2}(x)$$

with

$$\begin{aligned}
T_{n,1}(x) &= \frac{1}{\sum_{k=1}^{n} q_k} \sum_{k=n_0-1}^{n} \frac{q_k}{h_{k+1}} Y_{k+1} K\left(\frac{x - X_{k+1}}{h_{k+1}}\right) \\
T_{n,2}(x) &= \frac{1}{\sum_{k=1}^{n} q_k} \sum_{k=n_0-1}^{n} \frac{q_k}{h_{k+1}} K\left(\frac{x - X_{k+1}}{h_{k+1}}\right).
\end{aligned}$$

Lemma 4.13 is proved by showing that, for $i \in \{1,2\}$,

$$\sup_{x\in I}\left|T_{n,i}(x) - \mathbb{E}\left(T_{n,i}(x)\right)\right| = O\left(\sqrt{n^{-1}h_n^{-1}}\ln n\right) \quad a.s. \tag{5.35}$$

and that

$$\sup_{x\in I}\left|\mathbb{E}\left(T_{n,1}\left(x\right)\right)-r\left(x\right)\mathbb{E}\left(T_{n,2}\left(x\right)\right)\right|=O\left(h_n^2\right). \tag{5.36}$$

The proof of (5.36) relies on classical computations and is omitted. Moreover the proof of (5.35) for $i=2$ is similar to that for $i=1$, and is omitted too. We now prove (5.35) for $i=1$. To this end, we first state the following lemma.

Lemma 5.2. *There exists $s>0$ such that, for all $C>0$,*

$$\sup_{x\in I}\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\left|T_{n,1}\left(x\right)-\mathbb{E}\left(T_{n,1}\left(x\right)\right)\right|\geq C\right]=O\left(n^{-\frac{C}{s}}\right).$$

We first show how (5.35) for $i=1$ can be deduced from Lemma 5.2, and then prove Lemma 5.2. Set $(M_n)\in\mathcal{GS}\left(\tilde{m}\right)$ with $\tilde{m}>0$, and note that, for all $C>0$, we have

$$\begin{aligned}
&\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\sup_{x\in I}\left|T_{n,1}\left(x\right)-\mathbb{E}\left(T_{n,1}\left(x\right)\right)\right|\geq C\right]\\
&\leq \mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\sup_{x\in I}\left|T_{n,1}\left(x\right)-\mathbb{E}\left(T_{n,1}\left(x\right)\right)\right|\geq C \text{ and } \sup_{k\leq n}\left|Y_{k+1}\right|\leq M_n\right]\\
&\quad+\mathbb{P}\left[\sup_{k\leq n}\left|Y_{k+1}\right|\geq M_n\right].
\end{aligned}$$

Let $I_i^{(n)}$ be $N\left(n\right)$ intervals of length d_n such that $\cup_{i=1}^{N(n)}I_i^{(n)}=I$, and for all $i\in\{1,\ldots,N\left(n\right)\}$, set $x_i^{(n)}\in I_i^{(n)}$. We have

$$\begin{aligned}
&\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\sup_{x\in I}\left|T_{n,1}\left(x\right)-\mathbb{E}\left(T_{n,1}\left(x\right)\right)\right|\geq C\right]\\
&\leq \sum_{i=1}^{N(n)}\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\sup_{x\in I_i^{(n)}}\left|T_{n,1}\left(x\right)-\mathbb{E}\left(T_{n,1}\left(x\right)\right)\right|\geq C \text{ and } \sup_{k\leq n}\left|Y_{k+1}\right|\leq M_n\right]\\
&\quad+\mathbb{P}\left[\sup_{k\leq n}\left|Y_{k+1}\right|\geq M_n\right].
\end{aligned}\tag{5.37}$$

Since K is Lipschitz-continuous, there exist $k^*,c^*>0$, such that, for all $x,y\in\mathbb{R}$ satisfying $|x-y|\leq d_n$ and on $\left\{\sup_{k\leq n}\left|Y_{k+1}\right|\leq M_n\right\}$, we have:

$$\begin{aligned}
&\left|T_{n,1}\left(x\right)-T_{n,1}\left(y\right)\right|\\
&=\left|\frac{1}{\sum_{k=1}^{n}q_k}\sum_{k=n_0-1}^{n}q_kh_{k+1}^{-1}Y_{k+1}\left[K\left(\frac{x-X_{k+1}}{h_{k+1}}\right)-K\left(\frac{y-X_{k+1}}{h_{k+1}}\right)\right]\right|\\
&\leq k^*M_nd_n\frac{1}{\sum_{k=1}^{n}q_k}\sum_{k=1}^{n}q_kh_{k+1}^{-2}\\
&\leq c^*M_nh_n^{-2}d_n.
\end{aligned}$$

It follows that, for all $x \in I_i^{(n)}$, on $\left\{\sup_{k\leq n}|Y_{k+1}|\leq M_n\right\}$, we have

$$\begin{aligned}
&|T_{n,1}(x)-\mathbb{E}(T_{n,1}(x))|\\
&\leq \left|T_{n,1}(x)-T_{n,1}\left(x_i^{(n)}\right)\right|+\left|T_{n,1}\left(x_i^{(n)}\right)-\mathbb{E}\left(T_{n,1}\left(x_i^{(n)}\right)\right)\right|\\
&\quad+\left|\mathbb{E}\left(T_{n,1}\left(x_i^{(n)}\right)\right)-\mathbb{E}(T_{n,1}(x))\right|\\
&\leq 2c^*M_nh_n^{-2}d_n+\left|T_{n,1}\left(x_i^{(n)}\right)-\mathbb{E}\left(T_{n,1}\left(x_i^{(n)}\right)\right)\right|.
\end{aligned}$$

In view of (5.37), we obtain, for all $C>0$,

$$\begin{aligned}
&\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\sup_{x\in I}|T_{n,1}(x)-\mathbb{E}(T_{n,1}(x))|\geq C\right]\\
&\leq \sum_{i=1}^{N(n)}\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\left|T_{n,1}\left(x_i^{(n)}\right)-\mathbb{E}\left(T_{n,1}\left(x_i^{(n)}\right)\right)\right|+2c^*M_n\frac{\sqrt{nh_n}}{\ln n}h_n^{-2}d_n\geq C\right]\\
&\quad+n\mathbb{P}[|Y|\geq M_n].
\end{aligned}$$

Now, set $(d_n)\in\mathcal{GS}\left(-\frac{1}{2}-\frac{3}{2}a-\tilde{m}\right)$ such that, for all n, $2c^*M_n\sqrt{nh_n}(\ln n)^{-1}h_n^{-2}d_n \leq C/2$; in view of Lemma 5.2 and Assumption $(A4)\,ii)$, there exists $s>0$ such that

$$\begin{aligned}
&\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\sup_{x\in I}|T_{n,1}(x)-\mathbb{E}(T_{n,1}(x))|\geq C\right]\\
&\leq N(n)\sup_{x\in I}\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}|T_{n,1}(x)-\mathbb{E}(T_{n,1}(x))|\geq\frac{C}{2}\right]\\
&\quad+n\exp(-t^*M_n)\,\mathbb{E}(\exp(t^*|Y|))\\
&= O\left(d_n^{-1}n^{-\frac{C}{2s}}+n\exp(-t^*M_n)\right).
\end{aligned}$$

Since $\left(d_n^{-1}\right)\in\mathcal{GS}\left(\frac{1}{2}+\frac{3}{2}a+\tilde{m}\right)$ and since $(M_n)\in\mathcal{GS}(\tilde{m})$ with $\tilde{m}>0$, we can choose C large enough so that

$$\sum_{n\geq 0}\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\sup_{x\in I}|T_{n,1}(x)-\mathbb{E}(T_{n,1}(x))|\geq C\right]<\infty,$$

which gives (5.35) for $i=1$.

It remains to prove Lemma 5.2. For all $x\in I$ and all $s>0$, we have

$$\begin{aligned}
&\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}(T_{n,1}(x)-\mathbb{E}(T_{n,1}(x)))\geq C\right] &(5.38)\\
&= \mathbb{P}\left[\exp\left[s^{-1}\sqrt{nh_n}(T_{n,1}(x)-\mathbb{E}(T_{n,1}(x)))\right]\geq n^{\frac{C}{s}}\right]\\
&\leq n^{-\frac{C}{s}}\mathbb{E}\left(\exp\left[s^{-1}\sqrt{nh_n}(T_{n,1}(x)-\mathbb{E}(T_{n,1}(x)))\right]\right)\\
&\leq n^{-\frac{C}{s}}\prod_{k=n_0-1}^{n}\mathbb{E}\left(\exp\left(s^{-1}U_{k,n}(x)\right)\right), &(5.39)
\end{aligned}$$

with

$$\begin{aligned}
&U_{k,n}(x)\\
&= \frac{\sqrt{nh_n}}{\sum_{k=1}^n q_k}\frac{q_k}{h_{k+1}}\left[Y_{k+1}K\left(\frac{x-X_{k+1}}{h_{k+1}}\right)-\mathbb{E}\left(\left(Y_{k+1}K\left(\frac{x-X_{k+1}}{h_{k+1}}\right)\right)\right)\right].
\end{aligned}$$

For k and n such that $k \leq n$, set

$$\alpha_{k,n} = \frac{\sqrt{nh_n}}{\sum_{k=1}^n q_k} \frac{q_k}{h_{k+1}}.$$

We have, for all $x \in I$,

$$\begin{aligned}
&\mathbb{E}\left(\exp\left[s^{-1}U_{k,n}(x)\right]\right) \\
&\leq 1 + \frac{1}{2}\mathbb{E}\left[s^{-2}U_{k,n}^2(x)\right] + \mathbb{E}\left[s^{-3}\left|U_{k,n}^3(x)\right|\right]\exp\left[\left|U_{k,n}(x)\right|\right] \\
&\leq 1 + \frac{1}{2}s^{-2}\alpha_{k,n}^2 Var\left[Y_{k+1}K\left(\frac{x-X_{k+1}}{h_{k+1}}\right)\right] + \left\{s^{-3}\alpha_{k,n}^3\|K\|_\infty^3\right. \\
&\qquad \left.\mathbb{E}\left[\left(\left|Y_{k+1}\right|^3 + \left(\mathbb{E}(|Y_{k+1}|)\right)^3\right)\exp\left(s^{-1}\alpha_{k,n}\|K\|_\infty\left(|Y_{k+1}| + \mathbb{E}\left(|Y_{k+1}|\right)\right)\right)\right]\right\}.
\end{aligned}$$

Now, note that

$$\begin{aligned}
\alpha_{k,n} &= \left(\frac{nq_n}{\sum_{k=1}^n q_k}\right)\frac{q_k\sqrt{kh_k^{-1}}}{q_n\sqrt{nh_n^{-1}}}\sqrt{k^{-1}h_k h_{k+1}^{-2}} \\
&= \left(\frac{nq_n}{\sum_{k=1}^n q_k}\right)\left(\prod_{j=k}^{n-1}\frac{q_j\sqrt{jh_j^{-1}}}{q_{j+1}\sqrt{(j+1)\,h_{j+1}^{-1}}}\right)\sqrt{k^{-1}h_k h_{k+1}^{-2}}.
\end{aligned}$$

Since $\left(q_j\sqrt{jh_j^{-1}}\right) \in \mathcal{GS}\left(-q + (1+a)/2\right)$ with $-q + (1+a)/2 > 0$, we have

$$\begin{aligned}
\frac{q_j\sqrt{jh_j^{-1}}}{q_{j+1}\sqrt{(j+1)\,h_{j+1}^{-1}}} &= 1 - \left(-q + \frac{1+a}{2}\right)\frac{1}{j} + o\left(\frac{1}{j}\right) \\
&\leq 1 \quad \text{for } j \text{ large enough.}
\end{aligned}$$

It follows that $\sup_{k \leq n}\alpha_{k,n} < \infty$. Consequently, in view of Assumption $(A4)\,ii)$, there exist $s > 0$ and $c^* > 0$ such that, for all k and n such that $k \leq n$,

$$\mathbb{E}\left[\left(\left|Y_{k+1}\right|^3 + \left(\mathbb{E}\left(|Y_{k+1}|\right)\right)^3\right)\exp\left(s^{-1}\alpha_{k,n}\|K\|_\infty\left(|Y_{k+1}| + \mathbb{E}\left(|Y_{k+1}|\right)\right)\right)\right] \leq c^*.$$

Recall that $\sup_{x\in I} Var\left[Y_k K\left((x - X_k)\,h_k^{-1}\right)\right] = O\left(h_k\right)$. We then deduce that there exist positive constants C_i^*, such that, for all $x \in I$, and for all k and n such that $k \leq n$,

$$\begin{aligned}
\mathbb{E}\left(\exp\left[s^{-1}U_{k,n}(x)\right]\right) &\leq 1 + C_1^*\frac{nh_n}{\left(\sum_{k=1}^n q_k\right)^2}\frac{q_k^2}{h_k} + C_2^*\frac{\left(nh_n\right)^{\frac{3}{2}}}{\left(\sum_{k=1}^n q_k\right)^3}\frac{q_k^3}{h_k^3} \\
&\leq \exp\left[C_3^*\frac{h_n}{nq_n^2}\frac{q_k^2}{h_k} + C_4^*\frac{h_n^{\frac{3}{2}}}{n^{\frac{3}{2}}q_n^3}\frac{q_k^3}{h_k^3}\right].
\end{aligned}$$

Then, it follows from (5.38) that, for all $C > 0$,

$$\begin{aligned}
&\sup_{x\in I}\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\left(T_{n,1}(x) - \mathbb{E}\left(T_{n,1}(x)\right)\right) \geq C\right] \\
&\leq n^{-\frac{C}{s}}\exp\left[C_3^*\frac{h_n}{nq_n^2}\sum_{k=1}^n\frac{q_k^2}{h_k} + C_4^*\frac{h_n^{\frac{3}{2}}}{n^{\frac{3}{2}}q_n^3}\sum_{k=1}^n\frac{q_k^3}{h_k^3}\right] = O\left(n^{-\frac{C}{s}}\right).
\end{aligned}$$

We establish exactly in the same way that, for all $C > 0$,

$$\sup_{x\in I}\mathbb{P}\left[\frac{\sqrt{nh_n}}{\ln n}\left(\mathbb{E}\left(T_{n,1}\left(x\right)\right)-T_{n,1}\left(x\right)\right)\geq C\right] \quad = \quad O\left(n^{-\frac{C}{s}}\right),$$

which concludes the proof of Lemma 5.2.

5.7. *Proof of Lemma 4.14*. In view of (4.14), we have

$$\Delta_n\left(x\right) \quad = \quad \Delta_n^{(1)}\left(x\right)+\Delta_n^{(2)}\left(x\right),$$

with

$$\Delta_n^{(1)}(x) \quad = \quad \sum_{k=n_0}^{n}U_{k,n}(f(x)t)\gamma_k\left(\mathbb{E}\left[Z_k(x)\right]-Z_k(x)\right)\left(r_{k-1}(x)-r(x)\right), \tag{5.40}$$

$$\Delta_n^{(2)}(x) \quad = \quad \sum_{k=n_0}^{n}U_{k,n}(f(x))\gamma_k\left(f(x)-\mathbb{E}\left[Z_k(x)\right]\right)\left(r_{k-1}(x)-r(x)\right). \tag{5.41}$$

Let us first note that, in view of (5.14) and by application of Lemma 4.1, we have

$$\begin{aligned}
\left|\Delta_n^{(2)}\left(x\right)\right| \quad &= \quad O\left(\Pi_n\left(f\left(x\right)\right)\sum_{k=n_0}^{n}\Pi_k^{-1}\left(f\left(x\right)\right)\gamma_k h_k^2 w_k\right) \quad a.s.\\
&= \quad O\left(\Pi_n\left(f\left(x\right)\right)\sum_{k=n_0}^{n}\Pi_k^{-1}\left(f\left(x\right)\right)\gamma_k O\left(m_k\right)w_k\right) \quad a.s.\\
&= \quad O\left(m_n w_n\right) \quad a.s.
\end{aligned}$$

Let us now bound $\Delta_n^{(1)}\left(x\right)$. To this end, we set

$$\begin{aligned}
\varepsilon_k\left(x\right) \quad &= \quad \mathbb{E}\left(Z_k\left(x\right)\right)-Z_k\left(x\right),\\
G_k\left(x\right) \quad &= \quad r_k\left(x\right)-r\left(x\right),\\
S_n\left(x\right) \quad &= \quad \sum_{k=1}^{n}\Pi_k^{-1}\left(f\left(x\right)\right)\gamma_k\varepsilon_k\left(x\right)G_{k-1}\left(x\right)
\end{aligned}$$

and $\mathcal{F}_k = \sigma\left(\left(X_1,Y_1\right),\ldots,\left(X_k,Y_k\right)\right)$. In view of (5.7) and of Lemma 4.1, the increasing process of the martingale $\left(S_n\left(x\right)\right)$ satisfies

$$\begin{aligned}
< S >_n\left(x\right) \quad &= \quad \sum_{k=n_0}^{n}\mathbb{E}\left[\Pi_k^{-2}\left(f\left(x\right)\right)\gamma_k^2\varepsilon_k^2\left(x\right)G_{k-1}^2\left(x\right)|\mathcal{F}_{k-1}\right]\\
&= \quad \sum_{k=n_0}^{n}\Pi_k^{-2}\left(f\left(x\right)\right)\gamma_k^2 G_{k-1}^2\left(x\right)Var\left[Z_k\left(x\right)\right]\\
&= \quad O\left(\sum_{k=n_0}^{n}\Pi_k^{-2}\left(f\left(x\right)\right)\gamma_k^2 w_k^2\frac{1}{h_k}\right) \quad a.s.\\
&= \quad O\left(\sum_{k=n_0}^{n}\Pi_k^{-2}\left(f\left(x\right)\right)\gamma_k m_k^2 w_k^2\right) \quad a.s.\\
&= \quad O\left(\Pi_n^{-2}\left(f\left(x\right)\right)m_n^2 w_n^2\right) \quad a.s.
\end{aligned}$$

• Let us first consider the case the sequence $(n\gamma_n)$ is bounded. We then have $\left(\Pi_n^{-1}\left(f\left(x\right)\right) \in \mathcal{GS}\left(\xi^{-1}f\left(x\right)\right)\right)$, and thus $\ln\left(< S >_n\left(x\right)\right) = O\left(\ln n\right)$ a.s. Theorem 1.3.15 in Duflo (1997) then ensures that, for any $\delta > 0$,

$$\begin{aligned} \left|S_n\left(x\right)\right| &= o\left(< S >_n^{\frac{1}{2}}\left(x\right)\left(\ln < S >_n\left(x\right)\right)^{\frac{1+\delta}{2}}\right) + O\left(1\right) \quad a.s. \\ &= o\left(\Pi_n^{-1}\left(f\left(x\right)\right)m_nw_n\left(\ln n\right)^{\frac{1+\delta}{2}}\right) + O\left(1\right) \quad a.s. \end{aligned}$$

It follows that, for any $\delta > 0$,

$$\begin{aligned} \left|\Delta_n^{(1)}\left(x\right)\right| &= o\left(m_nw_n\left(\ln n\right)^{\frac{1+\delta}{2}}\right) + O\left(\Pi_n\left(f\left(x\right)\right)\right) \quad a.s. \\ &= o\left(m_nw_n\left(\ln n\right)^{\frac{1+\delta}{2}}\right) + o\left(m_n\right) \quad a.s., \end{aligned}$$

which concludes the proof of Lemma 4.14 in this case.
• Let us now consider the case $\lim_{n\to\infty}(n\gamma_n) = \infty$. In this case, for all $\delta > 0$, we have

$$\begin{aligned} \ln\left(\Pi_n^{-2}\left(f\left(x\right)\right)\right) &= \sum_{k=n_0}^{n}\ln\left(1-\gamma_kf\left(x\right)\right)^{-2} \\ &= \sum_{k=n_0}^{n}\left(2\gamma_kf\left(x\right) + o\left(\gamma_k\right)\right) \\ &= O\left(\sum_{k=1}^{n}\gamma_kk^{\delta}\right). \end{aligned}$$

Since $\left(\gamma_nn^{\delta}\right) \in \mathcal{GS}\left(-\left(\alpha-\delta\right)\right)$ with $(\alpha-\delta) < 1$, we have

$$\lim_{n\to\infty}\frac{n\left(\gamma_nn^{\delta}\right)}{\sum_{k=1}^{n}\gamma_kk^{\delta}} = 1-\left(\alpha-\delta\right).$$

It follows that $\ln\left(\Pi_n^{-2}\left(f\left(x\right)\right)\right) = O\left(n^{1+\delta}\gamma_n\right)$. The sequence (m_nw_n) being in $\mathcal{GS}\left(-m^*+w^*\right)$, we deduce that, for all $\delta > 0$, we have

$$\ln\left(< S >_n\left(x\right)\right) = O\left(n^{1+\delta}\gamma_n\right) \quad a.s.$$

Theorem 1.3.15 in Duflo (1997) then ensures that, for any $\delta > 0$,

$$\begin{aligned} \left|S_n\left(x\right)\right| &= o\left(< S >_n^{\frac{1}{2}}\left(x\right)\left(\ln < S >_n\left(x\right)\right)^{\frac{1+\delta}{2}}\right) + O\left(1\right) \quad a.s. \\ &= o\left(\Pi_n^{-1}\left(f\left(x\right)\right)m_nw_n\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}\right) + O\left(1\right) \quad a.s. \end{aligned}$$

It follows from the application of Lemma 4.1 that, for any $\delta > 0$,

$$\begin{aligned} \left|\Delta_n^{(1)}\left(x\right)\right| &= o\left(m_nw_n\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}\right) + O\left(\Pi_n\left(f\left(x\right)\right)\right) \quad a.s. \\ &= o\left(m_nw_n\left(n^{1+\delta}\gamma_n\right)^{\frac{1+\delta}{2}}\right) \quad a.s., \end{aligned}$$

which concludes the proof of Lemma 4.14.

5.8. *Proof of Lemma 4.15.* Let us first note that, in view of (5.41), and by application of Lemma 4.1, we have

$$\begin{aligned}\sup_{x\in I}\left|\Delta_n^{(2)}(x)\right| &= O\left(\sum_{k=n_0}^{n}\left(\sup_{x\in I}U_{k,n}(f(x))\right)\gamma_k h_k^2 w_k\right) \quad a.s.\\ &= O\left(\sum_{k=n_0}^{n}U_{k,n}(\varphi)\gamma_k m_k w_k\right) \quad a.s.\\ &= O(m_n w_n) \quad a.s.\end{aligned}$$

Now, set

$$A_n = \frac{3}{t^*}\ln n \tag{5.42}$$

(where t^* is defined in $(A4)\,ii)$) and write $\Delta_n^{(1)}$ (defined in (5.40)) as

$$\Delta_n^{(1)}(x) = \Pi_n(f(x))\,M_n^{(n)}(x) + \Pi_n(f(x))\,S_n(x),$$

with

$$S_n(x) = \sum_{k=n_0}^{n}\Pi_k^{-1}(f(x))\gamma_k\left(\mathbb{E}\left[Z_k(x)\right]-Z_k(x)\right)\left(r_{k-1}(x)-r(x)\right)\mathbb{1}_{\sup_{l\leq k-1}|Y_l|>A_n},$$

$$M_k^{(n)}(x) = \sum_{j=n_0}^{k}\Pi_j^{-1}(f(x))\gamma_j\left(\mathbb{E}[Z_j(x)]-Z_j(x)\right)\left(r_{j-1}(x)-r(x)\right)\mathbb{1}_{\sup_{l\leq j-1}\|Y_l|\leq A_n}.$$

Let us first prove a uniform strong upper bound for S_n. For any $c>0$, we have

$$\begin{aligned}\sum_{n\geq 0}\mathbb{P}\left[\sup_{x\in I}m_n^{-1}w_n^{-1}\left|S_n(x)\right|\geq c\right] &= O\left(\sum_{n\geq 0}\mathbb{P}\left(\sup_{l\leq n-1}|Y_l|>A_n\right)\right)\\ &= O\left(\sum_{n\geq 0}n\mathbb{P}\left(|Y|>A_n\right)\right)\\ &= O\left(\sum_{n\geq 0}n\exp\left(-t^*A_n\right)\right)\\ &< \infty.\end{aligned}$$

It follows that

$$\sup_{x\in I}\left|S_n(x)\right| = O(m_n w_n) \quad a.s.$$

To establish the strong uniform bound of $M_n^{(n)}$, we shall apply the following result given in Duflo (1997, page 209).

Lemma 5.3.
Let $\left(M_k^{(n)}\right)_k$ *be a martingale such that, for all* $k\leq n$, $\left|M_k^{(n)}-M_{k-1}^{(n)}\right|\leq c_n$, *and set*
$\Phi_c(\lambda) = c^{-2}\left(e^{\lambda c}-1-\lambda c\right)$. *For all* λ_n *such that* $\lambda_n c_n\leq 1$ *and all* $\alpha_n>0$, *we have*

$$\mathbb{P}\left(\lambda_n\left(M_n^{(n)}-M_0^{(n)}\right)\geq\Phi_{c_n}(\lambda_n)<M^{(n)}>_n+\alpha_n\lambda_n\right)\leq e^{-\alpha_n\lambda_n}.$$

In view of (4.15), there exists $C^*>0$ such that, on $\left\{\sup_{l\leq k}|Y_l|\leq A_n\right\}$,

$$|r_k(x)-r(x)| \leq C^*k\gamma_k h_k^{-1}A_n.$$

Consequently, there exists $C_1>0$ such that

$$\begin{aligned}
&\left|M_k^{(n)}(x)-M_{k-1}^{(n)}(x)\right| \\
&\leq \Pi_k^{-1}(f(x))\gamma_k\left|Z_k(x)-\mathbb{E}(Z_k(x))\right|\left|(r_{k-1}(x)-r(x))\mathbb{1}_{\sup_{l\leq k-1}|Y_l|\leq A_n}\right| \\
&\leq \Pi_k^{-1}(f(x))\gamma_k\left(2h_k^{-1}\|K\|_\infty\right)\left(C^*(k-1)\gamma_{k-1}h_{k-1}^{-1}A_n\right) \\
&\leq C_1\Pi_k^{-1}(f(x))k\gamma_k^2h_k^{-2}A_n.
\end{aligned}$$

• In the case $\lim_{n\to\infty}(n\gamma_n)=\infty$, since $\left(n\gamma_n^2h_n^{-2}\right)\in\mathcal{GS}(1-2\alpha+2a)$ there exists $(u_k)\to 0$ such that

$$\begin{aligned}
\frac{(k-1)\gamma_{k-1}^2h_{k-1}^{-2}}{k\gamma_k^2h_k^{-2}} &= 1-[1-2\alpha+2a]\frac{1}{k}+o\left(\frac{1}{k}\right) \\
&= 1+u_k\gamma_k.
\end{aligned}$$

It follows that there exists $k_0\geq n_0$ such that, for all $k\geq k_0$ and for all $x\in I$,

$$\begin{aligned}
\frac{\Pi_{k-1}^{-1}(f(x))(k-1)\gamma_{k-1}^2h_{k-1}^{-2}}{\Pi_k^{-1}(f(x))k\gamma_k^2h_k^{-2}} &= (1-\gamma_kf(x))(1+u_k\gamma_k) \\
&= 1-\gamma_kf(x)+u_k\gamma_k(1-\gamma_kf(x)) \\
&\leq 1-\gamma_k\varphi+u_k\gamma_k(1+\gamma_k\|f\|_\infty) \\
&\leq 1.
\end{aligned}$$

Consequently, there exists $C>0$ such that, for all $x\in I$ and all $k\leq n$,

$$\left|M_k^{(n)}(x)-M_{k-1}^{(n)}(x)\right| \leq C\Pi_n^{-1}(f(x))n\gamma_n^2h_n^{-2}A_n. \tag{5.43}$$

• In the case $\lim_{n\to\infty}(n\gamma_n)<\infty$ (in which case $\alpha=1$), we set $\epsilon\in]0,\min\{(1-3a)/2;\varphi\xi^{-1}-m^*\}[$ (where m^* is defined in (4.3)), and write

$$\left|M_k^{(n)}(x)-M_{k-1}^{(n)}(x)\right| \leq C_1\left[\Pi_k^{-1}(f(x))k^{-\epsilon}m_k\right]A_n\left[m_k^{-1}k^{1+\epsilon}\gamma_k^2h_k^{-2}\right].$$

Since $\left(m_n^{-1}n^{1+\epsilon}\gamma_n^2h_n^{-2}\right)\in\mathcal{GS}(m^*+1+\epsilon-2\alpha+2a)$ with

$$\begin{aligned}
m^*+1+\epsilon-2\alpha+2a &\leq \frac{1-a}{2}+\epsilon-1+2a \\
&\leq \epsilon-\frac{1}{2}(1-3a)<0,
\end{aligned}$$

the sequence $\left(m_n^{-1}n^{1+\epsilon}\gamma_n^2h_n^{-2}\right)$ is bounded. On the other hand, since $(n^{-\epsilon}m_n)\in\mathcal{GS}(-\epsilon-m^*)$, there exists $(u_k)\to 0$ and $k_0\geq n_0$ such that, for all $k\geq k_0$ and for all $x\in I$,

$$\begin{aligned}
\frac{\Pi_{k-1}^{-1}(f(x))(k-1)^{-\epsilon}m_{k-1}}{\Pi_k^{-1}(f(x))k^{-\epsilon}m_k} &= (1-\gamma_kf(x))\left(1+(m^*+\epsilon)\frac{1}{k}+o\left(\frac{1}{k}\right)\right) \\
&= (1-\gamma_kf(x))(1+(m^*+\epsilon)\xi\gamma_k+u_k\gamma_k) \\
&\leq (1-\gamma_k\varphi)(1+(m^*+\epsilon)\xi\gamma_k+|u_k|\gamma_k) \\
&\leq 1-\frac{(\varphi-(m^*+\epsilon)\xi)\gamma_k}{2} \\
&\leq 1.
\end{aligned}$$

It follows that there exists $C > 0$ such that, for all $x \in I$ and all $k \leq n$,

$$\left| M_k^{(n)}(x) - M_{k-1}^{(n)}(x) \right| \quad \leq \quad C\Pi_n^{-1}(f(x))\, n^{-\epsilon} m_n A_n. \tag{5.44}$$

From now on, we set

$$c_n(x) = \begin{cases} C\Pi_n^{-1}(f(x))\, n\gamma_n^2 h_n^{-2} A_n & \text{if } \lim_{n\to\infty}(n\gamma_n) = \infty, \\ C\Pi_n^{-1}(f(x))\, m_n n^{-\epsilon} A_n & \text{if } \lim_{n\to\infty}(n\gamma_n) < \infty, \end{cases}$$

so that in view of (5.43) and (5.44), for all $x \in I$ and all $k \leq n$, we have

$$\left| M_k^{(n)}(x) - M_{k-1}^{(n)}(x) \right| \quad \leq \quad c_n(x).$$

Now, let (u_n) be a positive sequence such that, for all n,

$$\begin{cases} u_n \leq C^{-1} n^{-1} \gamma_n^{-2} h_n^2 A_n^{-1} & \text{if } \lim_{n\to\infty}(n\gamma_n) = \infty, \\ u_n \leq C^{-1} m_n^{-1} n^{\epsilon} A_n^{-1} & \text{if } \lim_{n\to\infty}(n\gamma_n) < \infty \end{cases} \tag{5.45}$$

and set

$$\lambda_n(x) = u_n \Pi_n(f(x)).$$

Let us at first assume that the following lemma holds.

Lemma 5.4.
There exist $C_2 > 0$ and $\rho > 0$ such that for all $x, y \in I$ such that $|x - y| \leq C_2 n^{-\rho}$, we have

$$\begin{aligned} \left| \lambda_n(x) M_n^{(n)}(x) - \lambda_n(y) M_n^{(n)}(y) \right| \quad &\leq \quad 1, \\ \left| \Phi_{c_n(x)}(\lambda_n(x)) < M^{(n)} >_n (x) - \Phi_{c_n(y)}(\lambda_n(y)) < M^{(n)} >_n (y) \right| \quad &\leq \quad 1. \end{aligned}$$

Set

$$\begin{aligned} d_n &= C_2 n^{-\rho}, \\ V_n(x) &= \lambda_n(x) M_n^{(n)}(x) - \Phi_{c_n(x)}(\lambda_n(x)) < M^{(n)} >_n (x), \\ \alpha_n(x) &= \frac{(\rho + 2)\ln n}{\lambda_n(x)}. \end{aligned}$$

Let $I_i^{(n)}$ be $N(n)$ intervals of length d_n such that $\cup_{i=1}^{N(n)} = I$, and for all $i \in \{1, \ldots, N(n)\}$, set $x_i^{(n)} \in I_i^{(n)}$. Applying Lemma 5.4, we get, for n large enough,

$$\begin{aligned} \mathbb{P}\left[\sup_{x \in I} V_n(x) \geq 2(\rho + 2)\ln n\right] \quad &\leq \quad \sum_{i=1}^{N(n)} \mathbb{P}\left[\sup_{x \in I_i^{(n)}} V_n(x) \geq 2(\rho + 2)\ln n\right] \\ &\leq \quad \sum_{i=1}^{N(n)} \mathbb{P}\left[V_n\left(x_i^{(n)}\right) + 2 \geq 2(\rho + 2)\ln n\right] \\ &\leq \quad N(n) \sup_{x \in I} \mathbb{P}\left[V_n(x) \geq (\rho + 2)\ln n\right]. \end{aligned}$$

Now, the application of Lemma 5.3 ensures that, for all $x \in I$,

$$\begin{aligned} &\mathbb{P}\left[V_n(x) \geq (\rho + 2)\ln n\right] \\ &\leq \quad \mathbb{P}\left[\lambda_n(x) M_n^{(n)}(x) - \Phi_{c_n(x)}(\lambda_n(x)) < M^{(n)} >_n (x) \geq \alpha_n(x) \lambda_n(x)\right] \\ &\leq \quad \exp\left[-\alpha_n(x) \lambda_n(x)\right] \\ &\leq \quad n^{-(\rho+2)}. \end{aligned}$$

It follows that

$$\sum_{n\geq 1}\mathbb{P}\left[\sup_{x\in I}V_n(x)\geq 2(\rho+2)\ln n\right]=O\left(\sum_{n\geq 1}n^{-2}\right)<+\infty$$

and, applying Borel-Cantelli Lemma, we obtain

$$\sup_{x\in I}\lambda_n(x)\,M_n^{(n)}(x)\quad\leq\quad\sup_{x\in I}\Phi_{c_n(x)}(\lambda_n(x))<M^{(n)}>_n(x)+2(\rho+2)\ln n\quad a.s.$$

Since $\Phi_c(\lambda)\leq\lambda^2$ as soon as $\lambda c\leq 1$, and since $\lambda_n(x)=u_n\Pi_n(f(x))$, it follows that

$$u_n\sup_{x\in I}\Pi_n(f(x))M_n^{(n)}(x)\leq u_n^2\sup_{x\in I}\Pi_n^2(f(x))<M^{(n)}>_n(x)+2(\rho+2)\ln n\quad a.s.$$

Establishing the same upper bound for the martingale $\left(-M_k^{(n)}\right)$, we obtain

$$\sup_{x\in I}\Pi_n(f(x))\left|M_n^{(n)}(x)\right|\leq u_n\sup_{x\in I}\Pi_n^2(f(x))<M^{(n)}>_n(x)+2\frac{(\rho+2)\ln n}{u_n}\quad a.s.$$

Now, since $\sup_{x\in I}Var(Z_k(x))=O\left(h_k^{-1}\right)$, we have

$$\begin{aligned}&\sup_{x\in I}\Pi_n^2(f(x))<M^{(n)}>_n(x)\\&=O\left(\sum_{k=n_0}^{n}\sup_{x\in I}U_{k,n}^2(f(x))\,\gamma_k^2\sup_{x\in I}|r_{k-1}(x)-r(x)|^2\sup_{x\in I}(Var[Z_k(x)])\right)\\&=O\left(\sum_{k=n_0}^{n}U_{k,n}^2(\varphi)\,\gamma_k^2h_k^{-1}w_k^2\right)\quad a.s.\end{aligned}\tag{5.46}$$

• Let us first consider the case when the sequence $(n\gamma_n)$ is bounded. In this case, (5.46) and Lemma 4.1 imply that

$$\begin{aligned}\sup_{x\in I}\Pi_n^2(f(x))<M^{(n)}>_n(x)&=O\left(\sum_{k=n_0}^{n}U_{k,n}^2(\varphi)\,\gamma_km_k^2w_k^2\right)\quad a.s.\\&=O\left(m_n^2w_n^2\right)\quad a.s.\end{aligned}$$

In this case, we have thus proved that, for all positive sequence (u_n) satisfying (5.45), we have

$$\sup_{x\in I}\Pi_n(f(x))\left|M_n^{(n)}(x)\right|=O\left(u_nm_n^2w_n^2+\frac{\ln n}{u_n}\right)\quad a.s.\tag{5.47}$$

Now, since the sequence $\left(\left[m_n^{-1}w_n^{-1}\sqrt{\ln n}\right]m_nn^{-\epsilon}A_n\right)$ belongs to $\mathcal{GS}(-(w^*+\epsilon))$ with $w^*+\epsilon>0$, there exists $u_0>0$ such that, for all n,

$$u_0m_n^{-1}w_n^{-1}\sqrt{\ln n}\leq C^{-1}m_n^{-1}n^{\epsilon}A_n^{-1}$$

(where C is defined in (5.45)). Applying (5.47) with $(u_n)=\left(u_0m_n^{-1}w_n^{-1}\sqrt{\ln n}\right)$, we obtain

$$\sup_{x\in I}\Pi_n(f(x))\left|M_n^{(n)}(x)\right|\quad=\quad O\left(m_nw_n\sqrt{\ln n}\right)\quad a.s.,$$

which concludes the proof of Lemma 4.15 in this case.

• Let us now consider the case $\lim_{n\to\infty}(n\gamma_n) = \infty$. In this case, (5.46) implies that

$$\sup_{x\in I} \Pi_n^2 (f(x)) < M >_n^{(n)} (x) \quad = \quad O\left(\gamma_n h_n^{-1} w_n^2\right) \quad a.s.$$

In this case, we have thus proved that, for all positive sequence (u_n) satisfying (5.45), we have

$$\sup_{x\in I} \Pi_n (f(x)) \left|M_n^{(n)}(x)\right| \quad = \quad O\left(u_n\gamma_n h_n^{-1} w_n^2 + \frac{\ln n}{u_n}\right) \quad a.s. \tag{5.48}$$

Now, in view of (4.24), of (5.42), and of the assumptions of Lemma 4.15, we have

$$\begin{aligned}\left[\sqrt{\gamma_n^{-1} h_n \ln n}\, w_n^{-1}\right]\left[n\gamma_n^2 h_n^{-2} A_n\right] \quad &= \quad O\left(\sqrt{\gamma_n h_n^{-1}\ln n}\, w_n^{-1} B_n\right) \\ &= \quad O(1).\end{aligned}$$

Thus, there exists $u_0 > 0$ such that, for all n,

$$u_0\sqrt{\gamma_n^{-1} h_n \ln n}\, w_n^{-1} \leq C^{-1} n^{-1}\gamma_n^{-2} h_n^2 A_n^{-1}.$$

Applying (5.48) with $(u_n) = \left(u_0\sqrt{\gamma_n^{-1} h_n \ln n}\, w_n^{-1}\right)$, we obtain

$$\begin{aligned}\sup_{x\in I} \Pi_n (f(x)) \left|M_n^{(n)}(x)\right| \quad &= \quad O\left(\sqrt{\gamma_n h_n^{-1}\ln n}\, w_n\right) \quad a.s. \\ &= \quad O\left(m_n w_n \sqrt{\ln n}\right) \quad a.s.,\end{aligned}$$

which concludes the proof of Lemma 4.15.

Proof of Lemma 5.4. Let $(\delta_n) \in \mathcal{GS}(-\delta^*)$, set $x, y \in I$ such that $|x-y| \leq \delta_n$, and let c_i^* denote generic constants. Let us first note that $|Z_k(x) - Z_k(y)| \leq c_1^*\delta_n h_k^{-2}$, that $|Z_k(x)| \leq c_2^* h_k^{-1}$, that

$$\begin{aligned}|Var[Z_k(x)] - Var[Z_k(y)]| \quad &= \quad |\mathbb{E}\{(Z_k(x) - Z_k(y) - [\mathbb{E}(Z_k(x)) - \mathbb{E}(Z_k(y))]) \\ &\qquad (Z_k(x) + Z_k(y) - [\mathbb{E}(Z_k(x)) + \mathbb{E}(Z_k(y))])\}| \\ &\leq \quad 8c_2^* h_k^{-1}\mathbb{E}[|Z_k(x) - Z_k(y)|] \\ &\leq \quad c_3^* h_k^{-3}\delta_n\end{aligned}$$

and that, in view of (4.15), on $\left\{\sup_{l\leq k}|Y_l| \leq A_n\right\}$,

$$|r_k(x) - r(x)| \leq c_4^* k\gamma_k h_k^{-1} A_n.$$

Now, in view of (4.4), we have

$$r_k(x) - r(x) = [1 - \gamma_k Z_k(x)][r_{k-1}(x) - r(x)] + \gamma_k [W_k(x) - r(x) Z_k(x)],$$

so that

$$\begin{aligned}&[r_k(x) - r(x)] - [r_k(y) - r(y)] \\ &\quad = \quad [1-\gamma_k Z_k(x)][r_{k-1}(x) - r(x)] - [1-\gamma_k Z_k(y)][r_{k-1}(y) - r(y)] \\ &\qquad\quad +\gamma_k([W_k(x) - W_k(y)] - [r(x) Z_k(x) - r(y) Z_k(y)]) \\ &\quad = \quad [1-\gamma_k Z_k(x)]([r_{k-1}(x) - r(x)] - [r_{k-1}(y) - r(y)]) \\ &\qquad\quad -\gamma_k [Z_k(x) - Z_k(y)][r_{k-1}(y) - r(y)] \\ &\qquad\quad +\gamma_k([W_k(x) - W_k(y)] - r(x)[Z_k(x) - Z_k(y)] - Z_k(y)[r(x) - r(y)]).\end{aligned}$$

For $k \geq n_0$, we have $|1-\gamma_k Z_k(x)| \leq 1$. It follows that, for $k \geq n_0$ and on $\left\{\sup_{l\leq k}|Y_l| \leq A_n\right\}$,

$$\begin{aligned}
&|[r_k(x)-r(x)]-[r_k(y)-r(y)]| \\
&\leq |[r_{k-1}(x)-r(x)]-[r_{k-1}(y)-r(y)]| + \gamma_k|Z_k(x)-Z_k(y)||r_{k-1}(y)-r(y)| \\
&\quad +\gamma_k(|Y_k|+|r(x)|)|Z_k(x)-Z_k(y)| + \gamma_k|Z_k(y)||r(x)-r(y)| \\
&\leq |[r_{k-1}(x)-r(x)]-[r_{k-1}(y)-r(y)]| + \left(c_5^* k\gamma_k^2 h_k^{-3}\delta_n A_n\right) + \left(c_6^*\gamma_k A_n h_k^{-2}\delta_n\right) \\
&\quad + \left(c_7^*\gamma_k h_k^{-1}\delta_n\right) \\
&\leq |[r_{k-1}(x)-r(x)]-[r_{k-1}(y)-r(y)]| + c_8^* k\gamma_k^2 h_k^{-3}\delta_n A_n \\
&\leq c_9^*\sum_{j=1}^{k} j\gamma_j^2 h_j^{-3}\delta_n A_n \\
&\leq c_{10}^* k^2\gamma_k^2 h_k^{-3}\delta_n A_n.
\end{aligned}$$

Moreover, we note that, on $\left\{\sup_{l\leq k}|Y_l| \leq A_n\right\}$,

$$\begin{aligned}
&\left|[r_k(x)-r(x)]^2-[r_k(y)-r(y)]^2\right| \\
&\leq |[r_k(x)-r(x)]-[r_k(y)-r(y)]||[r_k(x)-r(x)]+[r_k(y)-r(y)]| \\
&\leq c_{11}^* k^3\gamma_k^2 h_k^{-4}\delta_n A_n^2.
\end{aligned}$$

Using all the previous upper bounds, as well as (5.28) with $p=1$, we get

$$\begin{aligned}
&\left|\lambda_n(x) M_n^{(n)}(x) - \lambda_n(y) M_n^{(n)}(y)\right| \\
&= u_n\left|\sum_{k=n_0}^{n} U_{k,n}(f(x))\gamma_k(\mathbb{E}[Z_k(x)]-Z_k(x))[r_{k-1}(x)-r(x)]\right. \\
&\quad \left. - \sum_{k=n_0}^{n} U_{k,n}(f(y))\gamma_k(\mathbb{E}[Z_k(y)]-Z_k(y))[r_{k-1}(y)-r(y)]\right| \mathbb{1}_{\sup_{l\leq k-1}|Y_l|\leq A_n} \\
&\leq u_n\sum_{k=n_0}^{n}\left\{U_{k,n}(f(x))\gamma_k|\mathbb{E}[Z_k(x)]-Z_k(x)|\right. \\
&\quad \left. |(r_{k-1}(x)-r(x))-(r_{k-1}(y)-r(y))|\mathbb{1}_{\sup_{l\leq k-1}|Y_l|\leq A_n}\right\} \\
&\quad +u_n\sum_{k=n_0}^{n}\left\{U_{k,n}(f(x))\gamma_k|r_{k-1}(y)-r(y)|\right. \\
&\quad \left. (|Z_k(x)-Z_k(y)|+\mathbb{E}[|Z_k(x)-Z_k(y)|])\mathbb{1}_{\sup_{l\leq k-1}|Y_l|\leq A_n}\right\} \\
&\quad +u_n\sum_{k=n_0}^{n}\left\{U_{k,n}(f(y))\left|\frac{U_{k,n}(f(x))}{U_{k,n}(f(y))}-1\right|\gamma_k|r_{k-1}(y)-r(y)|\right. \\
&\quad \left. (|Z_k(x)|+\mathbb{E}[|Z_k(x)|])\mathbb{1}_{\sup_{l\leq k-1}|Y_l|\leq A_n}\right\}
\end{aligned}$$

$$
\begin{aligned}
&\leq c_{12}^{*} u_n \sum_{k=n_0}^{n} U_{k,n}(\varphi) \gamma_k \left(h_k^{-1}\right) \left(k^2 \gamma_k^2 h_k^{-3} \delta_n A_n\right) \\
&\quad + c_{13}^{*} u_n \sum_{k=n_0}^{n} U_{k,n}(\varphi) \gamma_k (k \gamma_k h_k^{-1} A_n) \left(\delta_n h_k^{-2}\right) \\
&\quad + c_{14}^{*} u_n \sum_{k=n_0}^{n} U_{k,n}(\varphi) \left(\delta_n \gamma_n^{-1}\right) \gamma_k (k \gamma_k h_k^{-1} A_n) \left(h_k^{-1}\right) \\
&\leq c_{15}^{*} u_n n^2 \gamma_n^2 h_n^{-4} \delta_n A_n + c_{16}^{*} u_n n \gamma_n h_n^{-3} \delta_n A_n + c_{17}^{*} u_n n h_n^{-2} \delta_n A_n.
\end{aligned}
$$

In view of (5.42) and (5.45), it follows that there exist $s_1^* > 0$ and $\tilde{S}_n^{(1)} \in \mathcal{G}S(s_1^*)$ such that

$$
\left| \lambda_n(x) M_n^{(n)}(x) - \lambda_n(y) M_n^{(n)}(y) \right| \leq \delta_n \tilde{S}_n^{(1)}. \tag{5.49}
$$

Now, we have

$$
\begin{aligned}
&\Phi_{c_n(x)}(\lambda_n(x)) < M^{(n)} >_n (x) - \Phi_{c_n(y)}(\lambda_n(y)) < M^{(n)} >_n (y) \\
&= \frac{\Phi_{c_n(x)}(\lambda_n(x))}{\lambda_n^2(x)} \lambda_n^2(x) < M^{(n)} >_n (x) - \frac{\Phi_{c_n(y)}(\lambda_n(y))}{\lambda_n^2(y)} \lambda_n^2(y) < M^{(n)} >_n (y) \\
&= \frac{\Phi_{c_n(x)}(\lambda_n(x))}{\lambda_n^2(x)} \left[\lambda_n^2(x) < M^{(n)} >_n (x) - \lambda_n^2(y) < M^{(n)} >_n (y) \right] \\
&\quad + \left[\frac{\Phi_{c_n(x)}(\lambda_n(x))}{\lambda_n^2(x)} - \frac{\Phi_{c_n(y)}(\lambda_n(y))}{\lambda_n^2(y)} \right] \lambda_n^2(y) < M^{(n)} >_n (y).
\end{aligned}
$$

Since $c_n(x) \lambda_n(x) = c_n(y) \lambda_n(y) = \tilde{B}_n$, we have

$$
\frac{\Phi_{c_n(x)}(\lambda_n(x))}{\lambda_n^2(x)} = \tilde{B}_n^{-2} \left(\exp\left(\tilde{B}_n\right) - 1 - \tilde{B}_n \right) = \frac{\Phi_{c_n(y)}(\lambda_n(y))}{\lambda_n^2(y)}.
$$

Using the fact that $\Phi_c(\lambda) \leq \lambda^2$ for $\lambda c \leq 1$, and applying (5.28) with $p = 2$, we deduce that

$$
\begin{aligned}
&\left| \Phi_{c_n(x)}(\lambda_n(x)) < M^{(n)} >_n (x) - \Phi_{c_n(y)}(\lambda_n(y)) < M^{(n)} >_n (y) \right| \\
&\leq \left| \lambda_n^2(x) < M^{(n)} >_n (x) - \lambda_n^2(y) < M^{(n)} >_n (y) \right| \\
&\leq u_n^2 \left| \sum_{k=n_0}^{n} U_{k,n}^2(f(x)) \gamma_k^2 Var\left[Z_k(x)\right] \left[r_{k-1}(x) - r(x)\right]^2 \right. \\
&\quad \left. - \sum_{k=n_0}^{n} U_{k,n}^2(f(y)) \gamma_k^2 Var\left[Z_k(y)\right] \left[r_{k-1}(y) - r(y)\right]^2 \right| \mathbb{1}_{\sup_{l \leq k-1} |Y_l| \leq A_n}
\end{aligned}
$$

$$
\begin{aligned}
\leq\;& u_n^2 \sum_{k=n_0}^{n} \left\{ U_{k,n}^2 (f(x)) \gamma_k^2 Var\left[Z_k(x)\right] \right. \\
& \left. \left| \left(r_{k-1}(x) - r(x)\right)^2 - \left(r_{k-1}(y) - r(y)\right)^2 \right| \mathbb{1}_{\sup_{l \leq k-1} |Y_l| \leq A_n} \right\} \\
& + u_n^2 \sum_{k=n_0}^{n} \left\{ U_{k,n}^2 (f(x)) \gamma_k^2 \left(r_{k-1}(y) - r(y)\right)^2 \right. \\
& \left. \left| Var\left[Z_k(x)\right] - Var\left[Z_k(y)\right] \right| \mathbb{1}_{\sup_{l \leq k-1} |Y_l| \leq A_n} \right\} \\
& + u_n^2 \sum_{k=n_0}^{n} \left\{ U_{k,n}^2 (f(y)) \left| \frac{U_{k,n}^2 (f(x))}{U_{k,n}^2 (f(y))} - 1 \right| \gamma_k^2 \left(r_{k-1}(y) - r(y)\right)^2 \right. \\
& \left. Var\left[Z_k(y)\right] \mathbb{1}_{\sup_{l \leq k-1} |Y_l| \leq A_n} \right\} \\
\leq\;& c_{18}^* u_n^2 \sum_{k=n_0}^{n} U_{k,n}^2 (\varphi) \gamma_k^2 \left(h_k^{-1}\right) \left(k^3 \gamma_k^2 h_k^{-4} \delta_n A_n^2\right) \\
& + c_{19}^* u_n^2 \sum_{k=n_0}^{n} U_{k,n}^2 (\varphi) \gamma_k^2 \left(k^2 \gamma_k^2 h_k^{-2} A_n^2\right) \left(\delta_n h_k^{-3}\right) \\
& + c_{20}^* u_n^2 \sum_{k=n_0}^{n} U_{k,n}^2 (\varphi) \left(\delta_n \gamma_n^{-1}\right) \gamma_k^2 \left(k^2 \gamma_k^2 h_k^{-2} A_n^2\right) h_k^{-1}.
\end{aligned}
$$

In view of (5.42) and (5.45), it follows that there exist $s_2^* > 0$ and $\tilde{S}_n^{(2)} \in \mathcal{G}S(s_2^*)$ such that

$$
\left| \Phi_{c_n(x)} \left(\lambda_n(x)\right) < M^{(n)} >_n (x) - \Phi_{c_n(y)} \left(\lambda_n(y)\right) < M^{(n)} >_n (y) \right| \leq \delta_n \tilde{S}_n^{(2)}. \tag{5.50}
$$

Lemma 5.4 follows from the combination of (5.49) and (5.50).

Acknowledgments We deeply thank two anonymous referees for their helpful comments.

References

J. R. Blum. Multidimensional stochastic approximation methods. *Ann. Math. Statist.* **25**, 737–744 (1954). MR0065092.

R. Bojanic and E. Seneta. A unified theory of regularly varying sequences. *Math. Z.* **134**, 91–106 (1973). MR0333082.

H. F. Chen. Lower rate of convergence for locating a maximum of a function. *Ann. Statist.* **16**, 1330–1334 (1988). MR0959206.

H. F. Chen, T. E. Duncan and B. Pasik-Duncan. A Kiefer-Wolfowitz algorithm with randomized differences. *IEEE Trans. Automat. Control.* **44**, 442–453 (1999). MR1680258.

G. Claeskens and I. van Keilegom. Bootstrap confidence bands for regression curves and their derivatives. *Ann. Statist.* **31**, 1852–1884 (2003). MR2036392.

B. Delyon and A. B. Juditsky. Stochastic optimization with averaging of trajectories. *Stochastics Stochastic Rep.* **39**, 107–118 (1992). MR1275360.

J. Dippon. Accelerated randomized stochastic optimization. *Ann. Statist.* **31**, 1260–1281 (2003). MR2001650.

J. Dippon and J. Renz. Weighted means of processes in stochastic approximation. *Math. Meth. Statist.* **5**, 32–60 (1996). MR1386824.

J. Dippon and J. Renz. Weighted means in stochastic approximation of minima. *SIAM J. Control Optim.* **35**, 1811–1827 (1997). MR1466929.

M. Duflo. Algorithmes stochastiques. In *Mathématiques et applications*. Springer (1996). MR1612815.

M. Duflo. Random iterative models. In *Applications of Mathematics*. Springer (1997). MR1485774.

V. Fabian. Stochastic approximation of minima with improved asymptotic speed. *Ann. Math. Statist.* **38**, 191–200 (1967). MR0207136.

J. Galambos and E. Seneta. Regularly varying sequences. *Proc. Amer. Math. Soc.* **41**, 110–116 (1973). MR0323963.

L. Györfi, M. Kohler, A. Krzyzak and H. Walk. *A distribution-free theory of nonparametric regression*. Springer-Verlag, New York (2002). MR1920390.

P. Hall. Effect of bias estimation on coverage accuracy of bootstrap confidence intervals for a probability density. *Ann. Statist.* **20**, 675–694 (1992). MR1165587.

P. Hall and C. C. Heyde. *Martingale limit theory and its application*. Academic Press Inc, New York-London (1980). MR0624435.

J. Kiefer and J. Wolfowitz. Stochastic estimation of the maximum of a regression function. *Ann. Math. Statist.* **23**, 462–466 (1952). MR0050243.

H. J. Kushner and D. S. Clark. *Stochastic approximation methods for constrained and unconstrained systems*. Springer, New York (1978). MR0499560.

H. J. Kushner and J. Yang. Stochastic approximation with averaging of the iterates: optimal asymptotic rate of convergence for general processes. *SIAM J. Control Optim.* **31**, 1045–1062 (1993). MR1227546.

A. Le Breton. About the averaging approach schemes for stochastic approximation. *Math. Meth. Statist.* **2**, 295–315 (1993). .

A. Le Breton and A. Novikov. Some results about averaging in stochastic approximation. *Metrika* **42**, 153–171 (1995). MR1360979.

L. Ljung, G. Pflug and H. Walk. *Stochastic approximation and optimization of random systems*. Birkhäuser (1992). MR1162311.

A. Mokkadem and M. Pelletier. A companion for the Kiefer-Wolfowitz-Blum stochastic approximation algorithm. *Ann. Statist.* **35**, 1749–1772 (2007). MR2351104.

A. Mokkadem and M. Pelletier. Compact law of the iterated logarithm for matrix-normalized sums of random vectors. *Theory Probab. Appl.* **52**, 636–650 (2008).

A. Mokkadem, M. Pelletier and Y. Slaoui. The stochastic approximation method for the estimation of a multivariate probability density. *J. Statist. Plann. Inference* **139**, 2459–2478 (2009).

E. A. Nadaraya. On estimating regression. *Theory Probab. Appl.* **10**, 186–190 (1964).

M. H. Neumann and J. Polzehl. Simultaneous bootstrap confidence bands in nonparametric regression. *J. Nonparametr. Statist.* **9**, 307–333 (1998). MR1646905.

M. B. Nevelsón and R. Z. Has'minskii. *Stochastic approximation and recursive estimation*. Amer. Math. Soc., Providence, RI (1976). MR0423714.

M. Pelletier. Asymptotic almost sure efficiency of averaged stochastic algorithms. *SIAM J. Control Optim.* **39**, 49–72 (2000). MR1780908.

B. T. Polyak. New method of stochastic approximation type. *Automat. Remote Control.* **51**, 937–946 (1990). MR1071220.

B. T. Polyak and A. B. Juditsky. Acceleration of stochastic approximation by averaging. *SIAM J. Control Optim.* **30**, 838–855 (1992). MR1167814.

B. T. Polyak and A. B. Tsybakov. Optimal orders of accuracy for search algorithms of stochastic optimization. *Problems Inform. Transmission* **26**, 126–133 (1990). MR1074128.

P. Révész. Robbins-Monro procedure in a Hilbert space and its application in the theory of learning processes I. *Studia Sci. Math. Hung.* **8**, 391–398 (1973). MR0373198.

P. Révész. How to apply the method of stochastic approximation in the non-parametric estimation of a regression function. *Math. Aperationsforsch. Statist. (Ser. Statistics)* **8**, 119–126 (1977). MR0501557.

D. Ruppert. Almost sure approximations for the Robbins-Monro and Kiefer-Wolfowitz processes with dependent noises. *Ann. Probab.* **10**, 178–187 (1982). MR0637384.

D. Ruppert. Efficient estimators from a slowly converging Robbins-Monro process. Technical Report 781, School of Oper. Res. and Ind. Engrg., Cornell University, Ithaca, NY (1988).

J. C. Spall. A stochastic approximation algorithm for large-dimensional systems in the Kiefer-Wolfowitz setting. In *Proc. Conference on Decision and Control*, pages 1544–1548. IEEE, New York (1988).

J. C. Spall. A one-measurement form of simultaneous perturbation stochastic approximation. *Automatica J. IFAC* **33**, 109–112 (1997). MR1437902.

H. Walk. Strong universal pointwise consistency of recursive regression estimates. *Ann. Inst. Statist. Math.* **53**, 691–707 (2001). MR1879605.

G. S. Watson. Smooth regression analysis. *Sakhya Ser. A* **26**, 359–372 (1964). MR0185765.

G. Yin. On extensions of Polyak's averaging approach to stochastic approximation. *Stochastics Stochastic Rep.* **33**, 254–164 (1991). MR1128497.