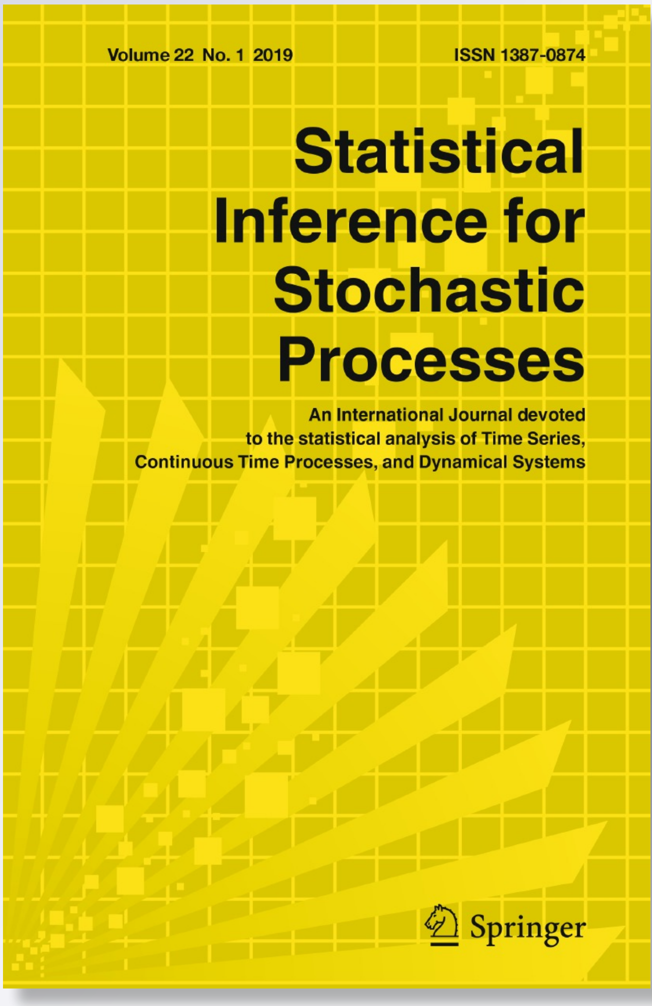
Nonparametric recursive estimation of the derivative of the regression function with application to sea shores water quality

Bernard Bercu, Sami Capderou & Gilles Durrieu

Statistical Inference for Stochastic Processes
An International Journal devoted to Time Series Analysis and the Statistics of Continuous Time Processes and Dynamical Systems

ISSN 1387-0874
Volume 22
Number 1

Stat Inference Stoch Process (2019) 22:17-40
DOI 10.1007/s11203-017-9169-1



Springer

Stat Inference Stoch Process (2019) 22:17–40
https://doi.org/10.1007/s11203-017-9169-1

Nonparametric recursive estimation of the derivative of the regression function with application to sea shores water quality

Bernard Bercu[1] · **Sami Capderou**[1] · **Gilles Durrieu**[2,3]

Received: 27 January 2017 / Accepted: 29 November 2017 / Published online: 5 December 2017

Abstract This paper is devoted to the nonparametric estimation of the derivative of the regression function in a nonparametric regression model. We implement a very efficient and easy to handle statistical procedure based on the derivative of the recursive Nadaraya–Watson estimator. We establish the almost sure convergence as well as the asymptotic normality for our estimates. We also illustrate our nonparametric estimation procedure on simulated data and real life data associated with sea shores water quality and valvometry.

Keywords Application and case studies · Mathematical statistics · Nonparametric methods · Smoothing and nonparametric regression

1 Introduction

Environmental and water protection should be tackled as a top priority of our society. It is forecasted that in 2035, nearly 60% of the world's population will live within 65 miles of the sea front (Haslett 2001). Water quality monitoring is therefore fundamental especially on the coastline. On the one hand, marine pollution comes mostly from land based sources. On the other hand, this pollution can lead to the collapse of coastal ecosystems and cause public

✉ Bernard Bercu
bernard.bercu@u-bordeaux.fr

Sami Capderou
sami.capderou@u-bordeaux.fr

Gilles Durrieu
gilles.durrieu@univ-nc.nc

1 Institut de Mathématiques de Bordeaux, UMR CNRS 5251, Université de Bordeaux, 351 Cours de la Libération, 33405 Talence, France

2 Laboratoire de Mathématiques de Bretagne Atlantique, UMR CNRS 6205, Université de Bretagne Sud, Campus de Tohannic, 56017 Vannes, France

3 Institut de Sciences Exactes et Appliquées, Université de la Nouvelle-Calédonie, 98851 Nouméa, Nouvelle-Calédonie, France

health issues. In this context, there is a critical need to develop a real-time reliable field assay to monitor the water quality within a decision making process. Among them, bioindicators are more and more commonly used. Endemic species are the most suitable bioindicators for the assessment of the quality of the coastal environment. For exemple, oysters, a well-known filter-feeding mollusc, feature a relevant sentinel organism to evaluate water quality. These animals being sedentary, they can witness the water quality evolution in a specific location.

The interest in investigating the bivalves activities by recording the valve movements has been explored for water quality surveillance. This area of interest is known as valvometry. The basic idea of valvometry is to use the bivalves ability to close its shell when exposed to a contaminant as an alarm signal (e.g. Doherty et al. (1987); Nagai et al. (2006); Sow et al. (2011)). Thus, recording the shell gaping activity of oysters is an effective method to study their behavior when facing water pollution (e.g. Riisgard et al. (2006); Garcia-March et al. (2008)). Nowadays, valvometric techniques produce high-frequency data, enabling online and in situ studies of the behavior of bivalve molluscs. They allow autonomous long-term recordings of valve movements without interfering their normal behavior. The goal of this paper is to propose a nonparametric statistical procedure based on the estimation of the derivative of the regression function in order to evaluate the velocity of the valve opening/closing activity.

A wide range of literature is available on nonparametric estimation of a regression function. We refer the reader to Nadaraya (1989), Devroye and Lugosi (2001), Györfi et al. (2002), Tsybakov (2009) for some excellent books on density and regression function estimation. Here, we shall focus our attention on the Nadaraya–Watson estimator of the regression function (Nadaraya 1964; Watson 1964). The almost sure convergence of this estimator was established by Noda (1976), while its asymptotic normality was proven by Schuster (1972). Later, Choi et al. (2000) proposed three data-sharpening versions of the Nadaraya–Watson estimator in order to reduce the asymptotic variance in the central limit theorem.

In this paper, we investigate an alternative approach, based on three recursive versions of the Nadaraya–Watson estimator (see Ahmad and Lin 1976; Amiri 2012; Bercu and Fraysse 2012; Devroye and Wagner 1980; Duflo 1997; Huang et al. 2014; Johnston 1982; Wand and Jones 1995). As it is well-known, recursive estimation procedures are specially useful when the observations are gathered sequentially. The three recursive versions of the Nadaraya–Watson estimator allow us to update the estimates as new observations are collected during the monitoring process, avoiding the need to recalculate a new estimate from the whole data. To the best of our knowledge, Ngerng (2011) is the only reference available on the derivative of the recursive Nadaraya–Watson estimator. However, the derivative is obtained by differentiation of the kernels, which is absolutely not necessary. In addition, some arguments are missing in the proof of the almost sure convergence and the explicit evaluation of the variance for the asymptotic normality is not provided. Consequently, our first goal is to carefully investigate the asymptotic behavior of the derivative of those three estimators. Our second goal is to illustrate our nonparametric estimation procedure on high-frequency valvometry data, in order to detect irregularities or abnormal behaviors of bivalves.

Another strategy for the recursive estimation of the regression function and of its derivatives is the local polynomial fitting approach, which was successfully implemented by Masry and Fan (1997), Vilar-Fernández and Vilar-Fernández (1998), Vilar-Fernández and Vilar-Fernández (2000). The advantages of local polynomial fitting are twofold. The estimates are easily computable and they have nice asymptotic properties. The main drawback of this approach is that it may be quantitatively affected by sparse regions of the random design (Seifert and Gasser 1996), as it requires compactly supported kernels.

The paper is organized as follows. Section 2 deals with our nonparametric estimation procedure of the derivative of the regression function. We establish in Sect. 3 the pointwise almost sure convergence as well as the asymptotic normality of our estimates and we compare their asymptotic variances. Section 4 deals with simulated data while Sect. 5 is devoted to a real data application on the survey of aquatic system using high-frequency valvometry. All the proofs of the nonparametric theoretical results are postponed to Appendices A and B.

2 Nonparametric estimation of the derivative

The relationship between the distance of two valves (Y_n) and the time of the measurement (X_n) can be seen as a nonparametric regression model given, for all $n \geq 1$, by

$$Y_n = f(X_n) + \varepsilon_n \tag{1}$$

where (ε_n) are unknown random errors. In all the sequel, we assume that (X_n) is a sequence of independent and identically distributed random variables with positive probability density function g. Our purpose is to estimate the derivative of the unknown regression function f which is directly associated with the velocity of the valve opening/closing activities of the oysters. For example, in an inhospitable environment, oysters behavior will be altered. Consequently, detecting changes of the closing and opening speed can provide insights about the health of oysters and so can be used as bioindicators of the water quality.

We recall that the Nadaraya–Watson estimator of the link function f is defined as

$$\widehat{f}_n^{NW}(x) = \frac{\sum_{k=1}^{n} Y_k K\left(\frac{x - X_k}{h_n}\right)}{\sum_{k=1}^{n} K\left(\frac{x - X_k}{h_n}\right)}, \tag{2}$$

where the kernel K is a chosen probability density function and the bandwidth (h_n) is a sequence of positive real numbers decreasing to zero. In our situation, we focus our attention on the recursive version of the Nadaraya–Watson estimator (Duflo 1997) of f given, for any $x \in \mathbb{R}$, by

$$\widehat{f}_n(x) = \frac{\sum_{k=1}^{n} \frac{Y_k}{h_k} K\left(\frac{x - X_k}{h_k}\right)}{\sum_{k=1}^{n} \frac{1}{h_k} K\left(\frac{x - X_k}{h_k}\right)}. \tag{3}$$

The denominator should, of course, be taken positive. It coincides with the recursive version of the Parzen–Rosenblatt estimator (Parzen 1962; Rosenblatt 1956) of the probability density function g. For any $x \in \mathbb{R}$, denote

$$\widehat{h}_n(x) = \frac{1}{n} \sum_{k=1}^{n} \frac{Y_k}{h_k} K\left(\frac{x - X_k}{h_k}\right), \widehat{g}_n(x) = \frac{1}{n} \sum_{k=1}^{n} \frac{1}{h_k} K\left(\frac{x - X_k}{h_k}\right) \tag{4}$$

which can be recursively calculated as

$$\widehat{h}_n(x) = \frac{n-1}{n} \hat{h}_{n-1}(x) + \frac{Y_n}{nh_n} K\left(\frac{x - X_n}{h_n}\right) \tag{5}$$

and

$$\widehat{g}_n(x) = \frac{n-1}{n}\,\widehat{g}_{n-1}(x) + \frac{1}{nh_n} K\left(\frac{x - X_n}{h_n}\right). \tag{6}$$

This modification allows dynamic updating of the estimates.

In the special case where g is known, a simplified version of the Nadaraya–Watson estimator of f, introduced by Johnston (1982), is given by

$$\widetilde{f}_n(x) = \frac{\widehat{h}_n(x)}{g(x)}. \tag{7}$$

In the same vein, an alternative estimator of f when g is known, was proposed by Wand and Jones (1995). It is defined, for any $x \in \mathbb{R}$, by

$$\check{f}_n(x) = \frac{1}{n}\sum_{k=1}^{n} \frac{Y_k}{g(X_k)h_k} K\left(\frac{x - X_k}{h_k}\right). \tag{8}$$

The derivatives of $\widehat{f}_n(x)$, $\widetilde{f}_n(x)$, and $\check{f}_n(x)$ are given, for any $x \in \mathbb{R}$ such that $g(x) > 0$, by

$$\widehat{f}_n'(x) = \frac{\widehat{h}_n'(x)}{\widehat{g}_n(x)} - \frac{\widehat{h}_n(x)\widehat{g}_n'(x)}{\widehat{g}_n^2(x)}, \tag{9}$$

$$\widetilde{f}_n'(x) = \frac{\widehat{h}_n'(x)}{g(x)} - \frac{\widehat{h}_n(x)g'(x)}{g^2(x)}, \tag{10}$$

$$\check{f}_n'(x) = \frac{1}{n}\sum_{k=1}^{n} \frac{Y_k}{g(X_k)h_k^2} K'\left(\frac{x - X_k}{h_k}\right). \tag{11}$$

3 Theoretical results

In order to investigate the asymptotic behavior of these derivative estimates, it is necessary to introduce several classical assumptions. First of all, denote by $\mathcal{F}_n$ the σ-algebra of the events occurring up to time n, $\mathcal{F}_n = \sigma(X_1, \varepsilon_1, \ldots, X_n, \varepsilon_n)$.

$(\mathcal{A}_1)$ The kernel K is a positive symmetric bounded function, differentiable with bounded derivative, satisfying

$$\int_{\mathbb{R}} K(x)dx = 1, \quad \int_{\mathbb{R}} K'(x)dx = 0, \quad \int_{\mathbb{R}} xK'(x)dx = -1, \int_{\mathbb{R}} x^2K'(x)dx = 0,$$
$$\int_{\mathbb{R}} x^4K(x)dx < \infty, \quad \int_{\mathbb{R}} x^4|K'(x)|dx < \infty.$$

$(\mathcal{A}_2)$ The regression function f and the density function g are bounded continuous, twice differentiable with bounded derivatives.

$(\mathcal{A}_3)$ The driven noise (ε_n) is a martingale difference sequence satisfying, for all $n \geq 1$, $\mathbb{E}[\varepsilon_n|\mathcal{F}_{n-1}] = 0$ and $\mathbb{E}[\varepsilon_n^2|\mathcal{F}_{n-1}] = \sigma^2$ a.s. where $\sigma^2 > 0$. Moreover, for all $n \geq 1$, X_n and ε_n are conditionally independent given $\mathcal{F}_{n-1}$.

On the one hand, it is not necessary to assume that the kernel K is compactly supported. On the other hand, we are not in the restrictive situation where the noise (ε_n) is a sequence of independent random variables. Our martingale assumption allows many general dependence structure of the random noise (ε_n). It is also important to notice that we are able to avoid the strict stationarity assumption usually made on the distribution of (X_n, Y_n). Finally, the bandwidth (h_n) is a sequence of positive real numbers, decreasing to zero, such that nh_n tends to infinity. For the sake of simplicity, we shall make use of $h_n = 1/n^\alpha$ with $0 < \alpha < 1$. Our first result on the almost sure convergence of our estimates is as follows.

Theorem 3.1 *Assume that* $(\mathcal{A}_1)$, $(\mathcal{A}_2)$ *and* $(\mathcal{A}_3)$ *hold. Then, if* $0 < \alpha < 1/3$, *we have for any* $x \in \mathbb{R}$ *such that* $g(x) > 0$,

$$\lim_{n\to\infty} \widehat{f}_n'(x) = f'(x) \quad a.s. \tag{12}$$

$$\lim_{n\to\infty} \widetilde{f}_n'(x) = f'(x) \quad a.s. \tag{13}$$

$$\lim_{n\to\infty} \check{f}_n'(x) = f'(x) \quad a.s. \tag{14}$$

Proof The proof is given in Appendix A. □

Remark 3.1 Under additional assumptions, it should be possible to establish almost sure rates of convergence of the uniform deviation of $\widehat{f}_n'(x) - f'(x)$ for x lying in a given compact set. In the same vein, uniform strong law of the logarithm were previously established for the non recursive version of the Nadaraya–Watson estimator (Blondin 2007; Deheuvels and Mason 2004; Mason 2004).

Our second result is devoted to the asymptotic normality of our estimates. Denote

$$\xi^2 = \int_{\mathbb{R}} \left(K'(x)\right)^2 dx. \tag{15}$$

Theorem 3.2 *Assume that* $(\mathcal{A}_1)$, $(\mathcal{A}_2)$ *and* $(\mathcal{A}_3)$ *hold and that the noise* (ε_n) *has a finite conditional moment of order* > 2. *Then, as soon as* $1/5 < \alpha < 1/3$, *we have for any* $x \in \mathbb{R}$ *such that* $g(x) > 0$, *the pointwise asymptotic normality*

$$\sqrt{nh_n^3}\left(\widehat{f}_n'(x) - f'(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\xi^2}{(1+3\alpha)g(x)}\sigma^2\right), \tag{16}$$

$$\sqrt{nh_n^3}\left(\widetilde{f}_n'(x) - f'(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\xi^2}{(1+3\alpha)g(x)}\left(f^2(x)+\sigma^2\right)\right), \tag{17}$$

$$\sqrt{nh_n^3}\left(\check{f}_n'(x) - f'(x)\right) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\xi^2}{(1+3\alpha)g(x)}\left(f^2(x)+\sigma^2\right)\right). \tag{18}$$

Proof The proof is given in Appendix B. □

Remark 3.2 One can realize that the derivate of the Nadaraya–Watson estimator $\widehat{f}_n'(x)$ is more efficient that $\widetilde{f}_n'(x)$ and $\check{f}_n'(x)$ as its asymptotic variance is the smallest one. The more $f(x)$ is far away from 0, the more one should make use of $\widehat{f}_n'(x)$. Moreover, the smallest values of ξ^2 are given by kernels with non-compact support. More precisely, for the Gaussian kernel

$$K(x) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right), \xi^2 = \frac{1}{4\sqrt{\pi}}$$

and for the Dirichlet kernel

$$K(x) = \frac{1}{\pi}\left(\frac{\sin x}{x}\right)^2, \xi^2 = \frac{4}{15\pi}.$$

Remark 3.3 Denote by $\widehat{f}_n^{LP}(x)$ the recursive local polynomial estimator of $f(x)$. It follows from Corollary 1 in Vilar-Fernández and Vilar-Fernández (1998) with $h_n = 1/n^{\alpha}$ where $1/5 < \alpha < 1/3$, that

$$\sqrt{nh_n^3}\big((\widehat{f}_n^{LP})'(x) - f'(x)\big) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{(1-2\alpha)^2\zeta^2}{(1-\alpha)g(x)}\,\sigma^2\right)$$

where

$$\zeta^2 = \left(\int_{\mathbb{R}} x^2 K(x)dx\right)^{-2} \int_{\mathbb{R}} x^2\big(K(x)\big)^2 dx.$$

The main drawback of this approach is that it may be quantitatively affected by sparse regions of the random design (X_n), as it requires K to be compactly supported (Seifert and Gasser 1996). For example, if K stands for the Epanechnikov kernel, one can easily check that $\zeta^2 = 15/7 \simeq 2.1429$. However, we already saw that for the standard Gaussian kernel, $\xi^2 = 1/(4\sqrt{\pi}) \simeq 0.1410$. Consequently, our approach offers much more flexibility in the choice of the kernel K, and the asymptotic variance for $\widehat{f}_n'(x)$ is in general smaller than that of the derivative of $\widehat{f}_n^{LP}(x)$.

4 Simulated data

This section is devoted to numerical experiments in order to evaluate the performances of our derivative estimates. The data are generated from the nonparametric regression model

$$Y_n = f(X_n) + \varepsilon_n, \tag{19}$$

where the regression function f and its derivative f' are defined, for all x in [0, 1], by

$$f(x) = (x+2)\sin(4\pi x^2) + 2\sin(8\pi x)$$

and

$$f'(x) = \sin(4\pi x^2) + 8\pi x(x+2)\cos(4\pi x^2) + 8\pi\cos(8\pi x).$$

The source of variation (ε_n) is a sequence of independent and identically distributed random variables with $\mathcal{N}(0,1)$ distribution. The random observation (X_n) is a sequence of independent random variables sharing the same distribution which is a mixture of three uniform distribution

$$g(x) = \frac{(1-p)}{2} g_1(x) + p g_2(x) + \frac{(1-p)}{2} g_3(x) \tag{20}$$

where $1/2 < p \leq 1$ and

$$g_1(x) = \frac{1}{1-p}\mathrm{I}_{x\in[0,1-p]}, \quad g_2(x) = \frac{1}{2p-1}\mathrm{I}_{x\in[1-p,p]}, \quad g_3(x) = \frac{1}{1-p}\mathrm{I}_{\{x\in[p,1]}.$$

One can observe that for $p = 1$, g coincides with the uniform distribution on [0, 1]. This distribution is introduced in order to illustrate the good performances of our statistical procedure in front of sparse regions of the random design (X_n). We implement our statistical procedure with a large sample size $n = 10\,000$ since we have large datasets in the application described in Sect. 5.

We first illustrate the pointwise almost sure convergence of the three estimators $\widehat{f}_n'(x)$, $\widetilde{f}_n'(x)$ and $\check{f}_n'(x)$ when the random observations (X_n) are uniformly distributed over [0, 1],

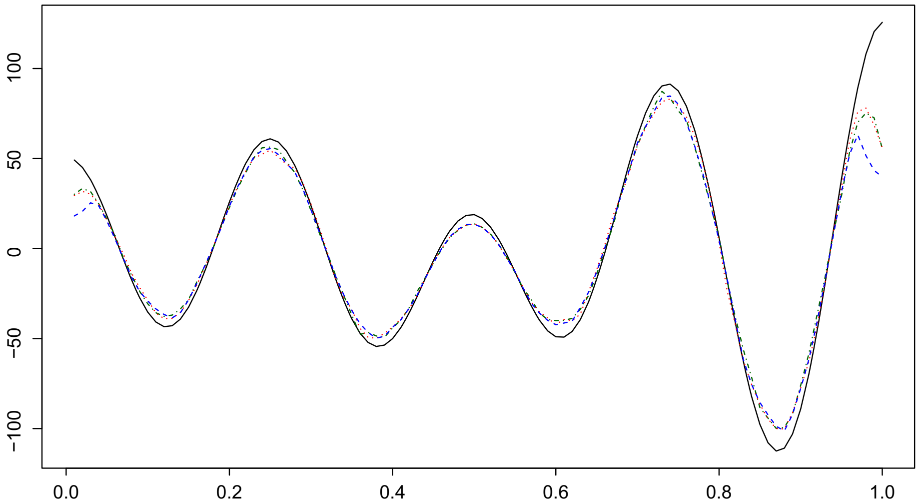


Fig. 1 Illustration of the almost sure convergence of $\widehat{f}_n'(x)$ (dotted line), $\widetilde{f}_n'(x)$ (dashed line), and $\breve{f}_n'(x)$ (dash-dotted line), to $f'(x)$ (solid line)

which means that $p = 1$. The choice of the kernel K is not crucial for the pointwise almost sure convergence and we have chosen to make use of the Epanechnikov kernel. In order to select the parameter α of the bandwidth, we use the standard cross validation method. Figure 1 shows that the three estimators $\widehat{f}_n'(x)$, $\widetilde{f}_n'(x)$ and $\breve{f}_n'(x)$ approximate very well the true derivative $f'(x)$ after selecting $\alpha = 0.3$ by cross validation.

In order to illustrate the pointwise asymptotic normality of our estimates, we implement a simulation study based on $N = 2\,000$ realizations. We numerically check the asymptotic normality at points $x = 0.5$ and $x = 0.8$ for our three estimators. One can see in Fig. 2 that the distributions of our three estimators are normally distributed and centered around 0. We observe the effect of $f^2(x)$ on the asymptotic variance of $\widetilde{f}_n'(x)$ and $\breve{f}_n'(x)$. Indeed, for $x = 0.5$, we have $f^2(x) = 0$, while for $x = 0.8$, we have $f^2(x) = 21.65$ which explains the differences between the asymptotic variances. It can be also shown that the mean squared error (MSE) of $\widehat{f}_n'(x)$ is much more smaller than the MSE of the non-recursive version of the Nadaraya–Watson estimator. In term of asymptotic variance, it is clear that $\widehat{f}_n'(x)$ performs better than $\widetilde{f}_n'(x)$ and $\breve{f}_n'(x)$. This is the reason why we have chosen to make use of $\widehat{f}_n'(x)$ to estimate the derivative $f'(x)$ for our real life data experiments.

We next focus our attention on simulations with sparse regions of the random design (X_n) by choosing $p = 0.8$ in (20). One can observe in Fig. 3 the sparse regions $[0, 0.2]$ and $[0.8, 1]$.

It was pointed out by Seifert and Gasser (1996) that local polynomial estimators with compactly supported kernels have shaky behaviour in sparse data areas. Figure 4 represents the recursive Nadaraya–Watson estimator $\widehat{f}_n'(x)$ and the recursive local polynomial estimator $\widehat{f}_n'^{LP}(x)$ for $x \in [0, 1]$, both with the same Epanechnikov kernel. One can see that $\widehat{f}_n'(x)$ is less affected by sparse regions than $\widehat{f}_n'^{LP}(x)$.

Finally, in order to underline the differences between $\widehat{f}_n'(x)$ and $\widehat{f}_n'^{LP}(x)$, we compute the average MSE on three points : $x = 0.12$, $x = 0.51$ and $x = 0.88$ for $N = 500$ realizations. The first and third points are located in sparse data areas, while the second one is in the dense center area. One can observe in Table 1 that while both estimators share the same MSE in the

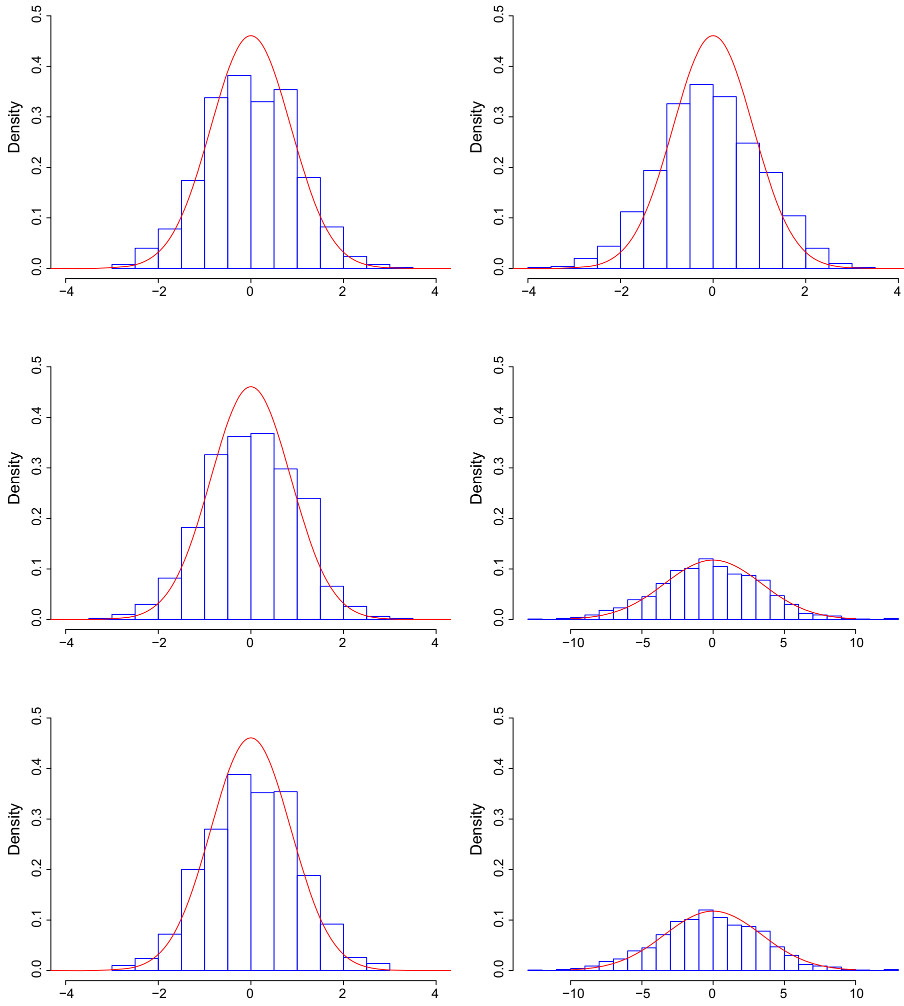


Fig. 2 Asymptotic normality of $\widehat{f}_n'(x)$ (first row), $\widetilde{f}_n'(x)$ (second row) and $\check{f}_n'(x)$ (third row) at point $x = 0.5$ (left column) and point $x = 0.8$ (right column). The density curves represent the asymptotic normal distributions given in Theorem 3.2

central region, the recursive Nadaraya–Watson estimator has smaller MSE in the sparse data areas for all kernels. A more elaborate comparison is desirable, and it should be performed someday.

5 High-frequency valvometry data

The motivation of this paper is to monitor sea shores water quality. For that purpose, we study bivalves activities by recording the valve movements. We use a high frequency, noninvasive valvometry electronic system developed by the UMR CNRS 5805 EPOC laboratory in Arcachon (France). The electronic principle of valvometry is described by Tran et al.

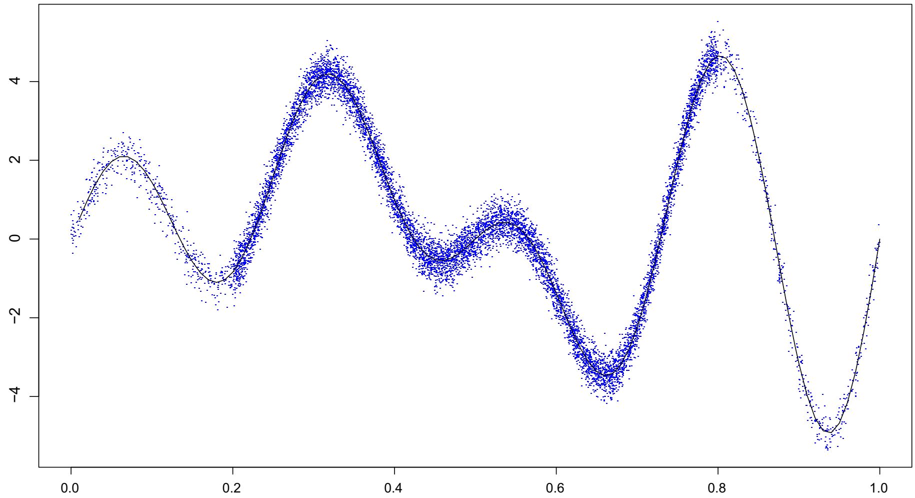


Fig. 3 Illustration of the (X_n, Y_n), $n = 10\,000$, following the model (19) with sparse region. The true curve of $f(x)$ is in plain line

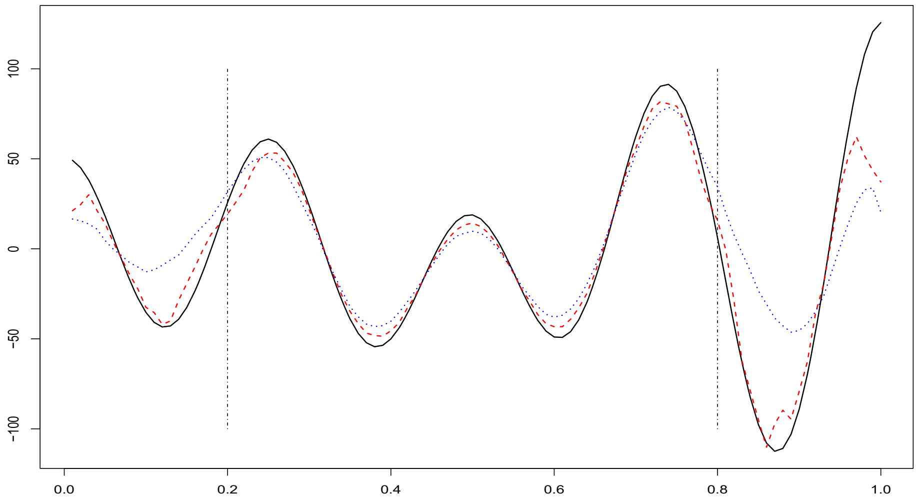


Fig. 4 Illustration of $\widehat{f}_n'(x)$ (dashed line) and $\widehat{f}_n'^{LP}(x)$ (dotted line) with sparse region in [0, 0.2] and [0.8, 1]. The true curve of $f'(x)$ is in solid line

Table 1 MSE associated to the Recursive Nadaraya–Watson (RNW) and recursive local polynomial (RLP) estimators for design points $x = 0.12$ (sparse region), $x = 0.51$ (dense region) and $x = 0.88$ (sparse region) with Epanechnikov, Quartic and Gaussian kernels

Design points	Epanechnikov kernel		Quartic kernel		Gaussian kernel	
	RNW	RLP	RNW	RLP	RNW	RLP
x = 0.12	0.1056	0.4378	0.1801	0.4951	0.3891	0.5865
x = 0.51	0.3945	0.4429	0.6695	0.5215	0.7119	0.7031
x = 0.88	0.0429	0.2942	0.0675	0.3228	0.1429	0.3871

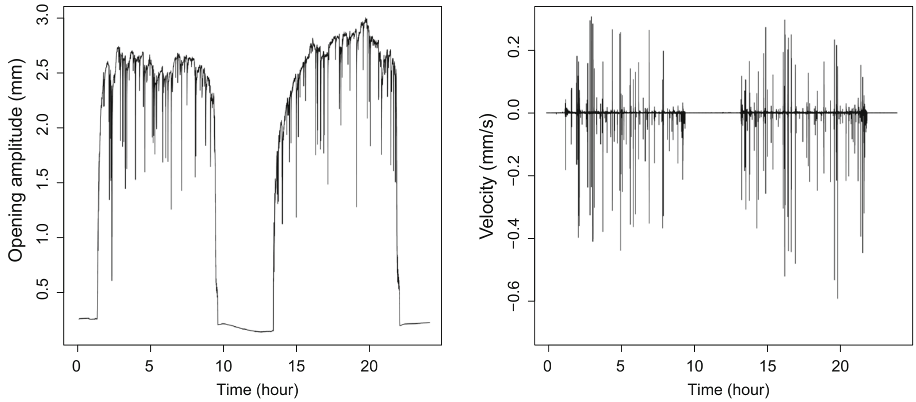


Fig. 5 A typical example of valvometric data for one oyster the June 2, 2011. In the left hand side, relationship between the opening amplitude (in millimeters) and the time of the experiment (over 24 h period). In the right hand side, the closing and opening velocity (millimeters per second) according to time (over the same period)

(2003), Chambon et al. (2007) and on the website http://molluscan-eye.epoc.u-bordeaux1.fr. This electronic system works autonomously without human intervention for a long period of time (at least one full year). Each animal is equipped with two light coils (sensors), of approximately 53 mg each (unembedded), fixed on the edge of each valve. One of the coils emits a high-frequency, sinusoidal signal which is received by the other coil. The strength of the electric field produced between the two coils being proportional to the inverse of distance between the point of measurement and the center of the transmitting coil, the distance between coils can be measured and the accuracy of the measurements is a few μm.

For each sixteen animals, one measurement is received every 0.1 s (10 Hz). So, the activity of oyster is measured every 1.6 s and each day, we obtain 864,000 triplets of data: the time of the measurement, the distance between the two valves and the animal number. A first electronic card in a waterproof case next to the animals manages the electrodes and a second electronic card handles the data acquisition. The valvometry system uses a GSM/GPRS modem and Linux operating system for the data storage, the internet access, and the data transmission. After each 24h period or any other programmed period of time, the data are transmitted to a workstation server and then inserted in a SQL database which is accessible with the software R (Development Core Team 2015) or a text terminal.

Several valvometric systems have been installed around the world: southern lagoon of New Caledonia, Spain, Ny Alesund Svalbard at 1300 km from the north pole, the north east of Murmansk in Russia on the Barents sea and at several sites in France with various species but we concentrate here on the Locmariaquer site situated in south Brittany based on sixteen oysters placed in a single bag. Locmariaquer (GPS coordinates 47°34 N, 2°56 W) is an important oyster farming area located near the narrow tidal pass which connects the gulf of Morbihan to the ocean, on the right side of the Auray river's mouth. Thus, oysters are close to the seasonal high traffic of the navigation channel and are potentially exposed to pollution as chemical residues of intensive agricultural practices.

As argued in Ahmed et al. (2015), Durrieu et al. (2015) and Durrieu et al. (2016), pollution can affect the activity of oysters and in particular the shells opening and closing velocities and so the movement speeds can be considered as an indicator of the animal stress activity since its movements are associated to aquatic system perturbations. In Ahmed et al. (2015),

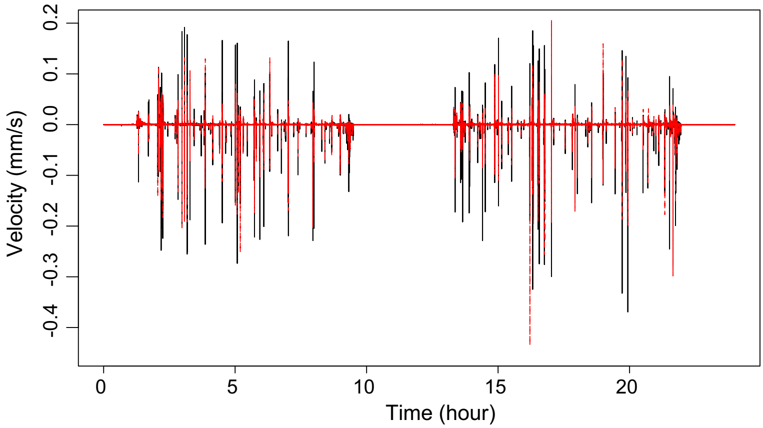


Fig. 6 The dashed line displays for June 2, 2011, the estimated $f'(x)$ using estimator $\widehat{f'_n}(x)$ versus the time x and the solid lines represent the observed speeds of valve openings and closings. The closing and opening velocity are measured in millimeters per second

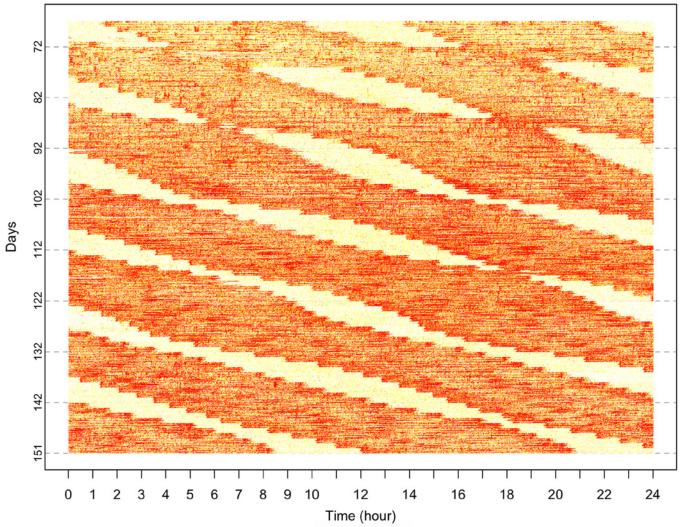


Fig. 7 Velocities estimation using $\widehat{f'_n}(x)$ from the 63th to the 151th days of 2011, considering the 16 oysters in Locmariaquer. The x-axis represents the time in a 24 h time period and the y-axis represents the number of days since January 1, 2011

the authors propose an interesting deterministic alternative method for the estimation of movement velocity based on differentiator estimators.

An example of valves activity and opening/closing velocity recordings June 2, 2011 is given in Fig. 5. Moreover, as the recursive estimator (9) defined in random design regression have smaller asymptotic variance than for fixed design regression (Bercu et al. 2017), we select the estimate $\widehat{f'_n}$ of f'.

Figure 6 displays for the same day the plot of the estimate $\widehat{f'_n}$ of f' of the valve closing and opening velocity for one oyster at the Locmariaquer site. The bandwidth parameter was selected by a standard the cross validation method.

We propose in Fig. 7 a visualization of the opening and closing velocity estimations of the 16 oysters for the period between the 63rd (March, 4th) and the 151st (May, 31th) of 2011. For each day, we compute for each oyster the estimator $\widehat{f}_n'(x)$ of $f'(x)$ for times x between 0 and 24h. Each velocity is represented by a color code : yellow for the smallest velocities, red for the highest and orange for the intermediate. Therefore, we can process the day velocity vector into a color line made of red, orange and yellow. We do it for each animal and superimpose multi-coloured dotted line. Hence, one day corresponds to 16 multi-coloured dotted lines. We repeat the process the day after and put the new 16 lines under the ones already obtained. This graphical representation reveals different clusters of global behaviors of the animals. First, one can notice yellow diagonal zones corresponding to the closed states of the animals. These states are highly correlated to the tidal amplitude, the animals being closed at low tide. Until the 100th day the animals have a normal behaviour but then we can observe a predominance of red in activity periods. It can be explained by a sudden change in temperature in the environment associated to the modifications of the specific activity of two enzymatic biomarkers meaning a possible pollution as described in Durrieu et al. (2016). We have performed many other analyses of these data using extreme value theory and other nonparametric statistical methods, all of which point the same conclusion (Coudret et al. 2015; Durrieu et al. 2015, 2016). Altogether, we anticipate that this approach could have a significant contribution providing *in situ* instant diagnosis of the bivalves behavior and thus appears to be an effective, early warning tool in ecological risk assessment.

Acknowledgements The authors would like to thank the Editor, the Associate Editor and the two anonymous Referees for their helpful comments and well-pointed remarks which helped to improve the paper substantially.

Appendix A: Proofs of the almost sure convergence results

The proofs of the almost sure convergence results rely on the following lemma. We also refer the reader to Silverman (1986) for the estimation of the derivative of the Parzen–Rosenblatt estimator.

Lemma A.1 *Assume that* $(\mathcal{A}_1)$, $(\mathcal{A}_2)$ *and* $(\mathcal{A}_3)$ *hold. Then, the estimators* $\widehat{g}_n$ *and* $\widehat{h}_n$*, given by* (4)*, satisfy for any* $x \in \mathbb{R}$*,*

$$\lim_{n\to\infty} \widehat{g}_n(x) = g(x) \quad a.s. \tag{A.1}$$

$$\lim_{n\to\infty} \widehat{h}_n(x) = f(x)g(x) \quad a.s. \tag{A.2}$$

Moreover, as soon as $0 < \alpha < 1/3$*, we also have for any* $x \in \mathbb{R}$*,*

$$\lim_{n\to\infty} \widehat{g}_n'(x) = g'(x) \quad a.s. \tag{A.3}$$

$$\lim_{n\to\infty} \widehat{h}_n'(x) = \big(f(x)g(x)\big)' \quad a.s. \tag{A.4}$$

Proof We shall only prove the almost sure convergence (A.4) inasmuch as (A.1) and (A.2) are well-known and the proof of (A.3) is more easy to handle and follow the same lines as the proof of (A.4). We deduce from (1) and (4) that for any $x \in \mathbb{R}$,

$$\widehat{h}_n(x) = \frac{1}{n}\sum_{k=1}^{n} \frac{f(X_k)}{h_k} K\left(\frac{x - X_k}{h_k}\right) + \frac{1}{n}\sum_{k=1}^{n} \frac{\varepsilon_k}{h_k} K\left(\frac{x - X_k}{h_k}\right).$$

Hence, by derivation, we have the decomposition

$$n\widehat{h}_n'(x) = A_n(x) + B_n(x) \tag{A.5}$$

where

$$A_n(x) = \sum_{k=1}^{n} a_k(x) = \sum_{k=1}^{n} f(X_k)v_k(X_k, x),$$
$$B_n(x) = \sum_{k=1}^{n} b_k(x) = \sum_{k=1}^{n} \varepsilon_k v_k(X_k, x)$$

with

$$v_n(X_n, x) = \frac{1}{h_n^2} K'\left(\frac{x - X_n}{h_n}\right). \tag{A.6}$$

On the one hand, we have for any $x \in \mathbb{R}$,

$$\begin{aligned} \mathbb{E}[a_n(x)] &= \int_{\mathbb{R}} f(x_n)v_n(x_n, x)g(x_n)dx_n \\ &= \frac{1}{h_n}\int_{\mathbb{R}} f(x - h_n y)g(x - h_n y)K'(y)dy. \end{aligned} \tag{A.7}$$

The regression function f as well as the density function g are bounded continuous and twice differentiable with bounded derivatives. Consequently, it follows from Taylor's formula that it exist θ_f, θ_g in the interval $]0, 1[$ such that, for any $x \in \mathbb{R}$,

$$f(x - h_n y) = f(x) - h_n y f'(x) + \frac{h_n^2 y^2}{2} f''(x - h_n y\theta_f),$$

and

$$g(x - h_n y) = g(x) - h_n y g'(x) + \frac{h_n^2 y^2}{2} g''(x - h_n y\theta_g).$$

By a careful analysis of each term in the product $f(x - h_n y)g(x - h_n y)$, we deduce from (A.7) together with assumption $(\mathcal{A}_1)$ that

$$\begin{aligned} \mathbb{E}[a_n(x)] &= -\big(f(x)g(x)\big)'\int_{\mathbb{R}} yK'(y)dy + h_n f'(x)g'(x)\int_{\mathbb{R}} y^2K'(y)dy + R_n(x) \\ &= \big(f(x)g(x)\big)' + R_n(x) \end{aligned} \tag{A.8}$$

where the remainder $R_n(x)$ satisfies

$$\sup_{x\in\mathbb{R}} |R_n(x)| = O(h_n).$$

Consequently, (A.8) immediately leads to

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[A_n(x)] = \big(f(x)g(x)\big)', \tag{A.9}$$

which is the limit we are looking for. By the same token,

$$\begin{aligned} \mathbb{E}[a_n^2(x)] &= \int_{\mathbb{R}} f^2(x_n)v_n^2(x_n, x)g(x_n)dx_n \\ &= \frac{1}{h_n^3}\int_{\mathbb{R}} f^2(x - h_n y)g(x - h_n y)\big(K'(y)\big)^2 dy \end{aligned}$$

$$= \frac{1}{h_n^3}\xi^2 f^2(x)g(x) + \zeta_n(x), \tag{A.10}$$

where ξ^2 is defined in (15) and the remainder $\zeta_n(x)$ is such that

$$\sup_{x\in\mathbb{R}} |\zeta_n(x)| = O\Big(\frac{1}{h_n^2}\Big).$$

Therefore, we deduce from (A.8) and (A.10) that

$$\lim_{n\to\infty} \frac{1}{n^{1+3\alpha}} \mathrm{Var}(A_n(x)) = \frac{\xi^2 f^2(x)g(x)}{1+3\alpha}. \tag{A.11}$$

On the other hand, we assume that for any $n \geq 1$, X_n and ε_n are conditionally independent given $\mathcal{F}_{n-1}$ where $\mathcal{F}_n = \sigma(X_1, \varepsilon_1, \ldots, X_n, \varepsilon_n)$. Consequently, we have for any $x \in \mathbb{R}$,

$$\begin{aligned}\mathbb{E}[b_n(x)|\mathcal{F}_{n-1}] &= \mathbb{E}[\varepsilon_n v_n(X_n, x)|\mathcal{F}_{n-1}] = \mathbb{E}[\varepsilon_n|\mathcal{F}_{n-1}]\mathbb{E}[v_n(X_n, x)|\mathcal{F}_{n-1}]\\ &= \mathbb{E}[\varepsilon_n|\mathcal{F}_{n-1}]\mathbb{E}[v_n(X_n, x)] = 0.\end{aligned}$$

Moreover,

$$\begin{aligned}\mathbb{E}[b_n^2(x)|\mathcal{F}_{n-1}] &= \mathbb{E}[\varepsilon_n^2 v_n^2(X_n, x)|\mathcal{F}_{n-1}] = \mathbb{E}[\varepsilon_n^2|\mathcal{F}_{n-1}]\mathbb{E}[v_n^2(X_n, x)|\mathcal{F}_{n-1}]\\ &= \sigma^2\mathbb{E}[v_n^2(X_n, x)].\end{aligned}$$

Furthermore, we have

$$\begin{aligned}\mathbb{E}[v_n^2(X_n, x)] &= \int_{\mathbb{R}} v_n^2(x_n, x)g(x_n)dx_n = \frac{1}{h_n^3}\int_{\mathbb{R}} g(x-h_n y)\big(K'(y)\big)^2 dy\\ &= \frac{1}{h_n^3}\int_{\mathbb{R}}\Big(g(x) - h_n y g'(x) + \frac{h_n^2 y^2}{2} g''(x-h_n y\theta_g)\Big)\big(K'(y)\big)^2 dy\\ &= \frac{1}{h_n^3}\xi^2 g(x) + \Delta_n(x)\end{aligned} \tag{A.12}$$

where ξ^2 is defined in (15) and the remainder $\Delta_n(x)$ is such that

$$\sup_{x\in\mathbb{R}} |\Delta_n(x)| = O\Big(\frac{1}{h_n^2}\Big).$$

Consequently, denoting

$$W_n(x) = \sum_{k=1}^{n} v_k^2(X_k, x),$$

it follows from (A.12) that

$$\lim_{n\to\infty} \frac{1}{n^{1+3\alpha}} \mathbb{E}[W_n(x)] = \frac{\xi^2 g(x)}{1+3\alpha}. \tag{A.13}$$

We are now in the position to prove the almost sure convergence (A.4). The decomposition (A.5) can be rewritten as

$$n\widehat{h}_n'(x) = M_n^A(x) + \mathbb{E}[A_n(x)] + B_n(x), \tag{A.14}$$

where $M_n^A(x) = A_n(x) - \mathbb{E}[A_n(x)]$. One can observe that $(M_n^A(x))$ and $(B_n(x))$ are both square integrable martingale difference sequences with predictable quadratic variations respectively given by $\langle M^A(x)\rangle_n = \text{Var}(A_n(x))$ and $\langle B(x)\rangle_n = \sigma^2\mathbb{E}[W_n(x)]$. Consequently, (A.11) together with (A.13) immediately lead to

$$\lim_{n\to\infty} \frac{\langle M^A(x)\rangle_n}{n^{1+3\alpha}} = \frac{\xi^2 f^2(x)g(x)}{1+3\alpha} \text{and} \quad \lim_{n\to\infty} \frac{\langle B(x)\rangle_n}{n^{1+3\alpha}} = \frac{\sigma^2\xi^2 g(x)}{1+3\alpha}. \tag{A.15}$$

Hence, we obtain from the strong law of large numbers for martingales given e.g. by Theorem 1.3.15 of Duflo (1997) that, for any $\gamma > 0$, $(M_n^A(x))^2 = o(n^{1+3\alpha}(\log n)^{1+\gamma})$ a.s. and $(B_n(x))^2 = o(n^{1+3\alpha}(\log n)^{1+\gamma})$ a.s. Therefore, as $0 < \alpha < 1/3$, it ensures that, for any $x \in \mathbb{R}$

$$\lim_{n\to\infty} \frac{1}{n} M_n^A(x) = 0 \quad \text{a.s.} \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n} B_n(x) = 0 \quad \text{a.s.} \tag{A.16}$$

Finally, we deduce from decomposition (A.14) together with (A.9) and (A.16) that for any $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \widehat{h}_n'(x) = \big(f(x)g(x)\big)' \quad \text{a.s.}$$

Thus Lemma A.1 is proven. □

Proof of Theorem 3.1 We shall now proceed to the proof of the Theorem 3.1. It clearly follows from relation (9) and Lemma A.1 that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\begin{aligned}\lim_{n\to+\infty} \widehat{f}_n'(x) &= \lim_{n\to+\infty}\left(\frac{\widehat{h}_n'(x)}{\widehat{g}_n(x)} - \frac{\widehat{h}_n(x)\widehat{g}_n'(x)}{\widehat{g}_n^2(x)}\right) = \frac{\big(f(x)g(x)\big)'}{g(x)} - \frac{f(x)g(x)g'(x)}{g^2(x)} \quad \text{a.s.}\\ &= \frac{f'(x)g(x) + f(x)g'(x) - f(x)g'(x)}{g(x)} = f'(x) \quad \text{a.s.}\end{aligned}$$

By the same token, relation (10) and Lemma A.1 immediately lead to

$$\lim_{n\to+\infty} \widetilde{f}_n'(x) = \lim_{n\to+\infty}\left(\frac{\widehat{h}_n'(x)}{g(x)} - \frac{\widehat{h}_n(x)g'(x)}{g(x)^2}\right) = f'(x) \quad \text{a.s.}$$

It only remains to prove (14). We obtain from relation (11) that

$$n\breve{f}_n'(x) = C_n(x) + D_n(x) \tag{A.17}$$

where

$$\begin{aligned} C_n(x) &= \sum_{k=1}^n c_k(x) = \sum_{k=1}^n \frac{f(X_k)}{g(X_k)} v_k(X_k, x),\\ D_n(x) &= \sum_{k=1}^n d_k(x) = \sum_{k=1}^n \frac{\varepsilon_k}{g(X_k)} v_k(X_k, x).\end{aligned}$$

As in the proof of Lemma A.1, we find that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\lim_{n\to\infty} \frac{1}{n}\mathbb{E}[C_n(x)] = f'(x) \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n^{1+3\alpha}} \text{Var}(C_n(x)) = \frac{\xi^2 f^2(x)}{(1+3\alpha)g(x)}. \tag{A.18}$$

Hereafter, we split $n\breve{f}_n'(x)$ into three terms

$$n\breve{f}_n'(x) = M_n^C(x) + \mathbb{E}[C_n(x)] + D_n(x) \tag{A.19}$$

where $M_n^C(x) = C_n(x) - \mathbb{E}[C_n(x)]$. One can observe that $(M_n^C(x))$ and $(D_n(x))$ are both square integrable martingale difference sequences with predictable quadratic variations satisfying, for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\lim_{n\to\infty} \frac{\langle M^C(x)\rangle_n}{n^{1+3\alpha}} = \frac{\xi^2 f^2(x)}{(1+3\alpha)g(x)} \quad \text{and} \quad \lim_{n\to\infty} \frac{\langle D(x)\rangle_n}{n^{1+3\alpha}} = \frac{\xi^2\sigma^2}{(1+3\alpha)g(x)}. \tag{A.20}$$

Therefore, we deduce from the strong law of large numbers for martingales that, as soon as $0 < \alpha < 1/3$,

$$\lim_{n\to\infty} \frac{1}{n} M_n^C(x) = 0 \quad \text{a.s.} \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n} D_n(x) = 0 \quad \text{a.s.} \tag{A.21}$$

Finally, it follows from (A.19) together with (A.18) and (A.21) that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\lim_{n\to\infty} \breve{f}_n'(x) = f'(x) \quad \text{a.s.}$$

which achieves the proof of Theorem 3.1. □

Appendix B: Proofs of the asymptotic normality results

In order to prove Theorem 3.2, we shall make use of the central limit theorem for martingales given e.g. by Theorem 2.1.9 of Duflo (1997). First of all, we focus our attention on convergence (18) since it is the easiest convergence to prove.

Proof of convergence (18). It follows from (A.19) that

$$\sqrt{nh_n^3}\big(\breve{f}_n'(x) - f'(x)\big) = \frac{\sqrt{nh_n^3}}{n}\Big(M_n^C(x) + \mathbb{E}[C_n(x)] + D_n(x) - nf'(x)\Big),$$

which implies the martingale decomposition

$$\sqrt{nh_n^3}\big(\breve{f}_n'(x) - f'(x)\big) = \frac{1}{\sqrt{n^{1+3\alpha}}}\big(\langle e, \mathcal{M}_n(x)\rangle + \breve{R}_n(x)\big) \tag{B.1}$$

where

$$e = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathcal{M}_n(x) = \begin{pmatrix} M_n^C(x) \\ D_n(x) \end{pmatrix},$$

and the remainder

$$\breve{R}_n(x) = \mathbb{E}[C_n(x)] - nf'(x) = \sum_{k=1}^{n}\big(\mathbb{E}[c_k(x)] - f'(x)\big). \tag{B.2}$$

It follows from Taylor's formula that it exists $\theta_f \in]0, 1[$ such that, for any $x \in \mathbb{R}$,

$$\begin{aligned}\mathbb{E}[c_n(x)] &= \int_{\mathbb{R}} f(x_n)v_n(x_n, x)dx_n = \frac{1}{h_n}\int_{\mathbb{R}} f(x - h_n y)K'(y)dy \\ &= f'(x) + \frac{h_n}{2}\int_{\mathbb{R}} f''(x - h_n y\theta_f)y^2 K'(y)dy,\end{aligned}$$

where v_n is defined in (A.6). Since f'' is bounded, we have

$$\sup_{x\in\mathbb{R}}\big|\mathbb{E}[c_n(x)] - f'(x)\big| \leq M_f \tau^2 h_n \tag{B.3}$$

where

$$M_f = \sup_{x\in\mathbb{R}} |f''(x)| \quad \text{and} \quad \tau^2 = \frac{1}{2}\int_{\mathbb{R}} y^2 \big|K'(y)\big| dy.$$

Hence, we deduce from (B.2) and (B.3) that

$$\sup_{x\in\mathbb{R}} \big|\breve{R}_n(x)\big| \le \tau^2 M_f \sum_{k=1}^{n} h_k.$$

However, it is easily seen that

$$\sum_{k=1}^{n} h_k \le \frac{1}{1-\alpha} n^{1-\alpha}.$$

Therefore, as soon as $\alpha > 1/5$, we obtain that

$$\sup_{x\in\mathbb{R}} \big|\breve{R}_n(x)\big| = o(\sqrt{n^{1+3\alpha}}). \tag{B.4}$$

Hereafter, the predictable quadratic variation (Duflo 1997) of the two-dimensional real martingale $(\mathcal{M}_n(x))$ is given, for all $n \ge 1$, by the diagonal matrix

$$\langle \mathcal{M}(x)\rangle_n = \begin{pmatrix} \langle M^C(x)\rangle_n & 0 \\ 0 & \langle D(x)\rangle_n \end{pmatrix}.$$

Then, it follows from (A.20) that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\lim_{n\to\infty} \frac{1}{n^{1+3\alpha}} \langle \mathcal{M}(x)\rangle_n = \frac{\xi^2}{(1+3\alpha)g(x)} \begin{pmatrix} f^2(x) & 0 \\ 0 & \sigma^2 \end{pmatrix}. \tag{B.5}$$

Furthermore, it is not hard to see that the martingale $(\mathcal{M}_n(x))$ satisfies the Lindeberg condition. As a matter of fact, we assume that the sequence (ε_n) has a finite conditional moment of order $p > 2$. Let $a > 0$ be such that $p = 2(1+a)$. If we denote $\Delta\mathcal{M}_n(x) = \mathcal{M}_n(x) - \mathcal{M}_{n-1}(x)$, we have for all $n \ge 1$,

$$\begin{aligned}\mathbb{E}\big[\|\Delta\mathcal{M}_n(x)\|^p|\mathcal{F}_{n-1}\big] &= \mathbb{E}\Big[\Big(\big(\Delta M_n^C(x)\big)^2 + \big(\Delta D_n(x)\big)^2\Big)^{1+a}|\mathcal{F}_{n-1}\Big] \\ &\le 2^a\, \mathbb{E}\Big[\big|\Delta M_n^C(x)\big|^p + \big|\Delta D_n(x)\big|^p\Big)|\mathcal{F}_{n-1}\Big]. \end{aligned} \tag{B.6}$$

On the one hand,

$$\begin{aligned}\mathbb{E}\big[\big|\Delta M_n^C(x)\big|^p|\mathcal{F}_{n-1}\big] &= \mathbb{E}\big[|c_n(x) - \mathbb{E}[c_n(x)]|^p|\mathcal{F}_{n-1}\big] \\ &\le 2^{p-1}\Big(\mathbb{E}\big[\big|c_n(x)\big|^p\big] + \big|\mathbb{E}[c_n(x)]\big|^p\Big). \end{aligned} \tag{B.7}$$

However, as f' is bounded, it follows from (B.3) that

$$\sup_{x\in\mathbb{R}} \big|\mathbb{E}[c_n(x)]\big| \le m_f + M_f\tau^2$$

where $m_f = \sup_{x\in\mathbb{R}} |f'(x)|$. Consequently, it exists a positive constant C_p such that

$$\sup_{x\in\mathbb{R}} \big|\mathbb{E}[c_n(x)]\big|^p \le C_p. \tag{B.8}$$

Moreover,

$$\mathbb{E}\big[\big|c_n(x)\big|^p\big] = \int_{\mathbb{R}} \frac{f(x_n)^p}{g(x_n)^{p-1}} \big|v_n(x_n, x)\big|^p dx_n$$
$$= \frac{1}{h_n^{2p-1}} \int_{\mathbb{R}} \frac{f(x - h_n y)^p}{g(x - h_n y)^{p-1}} \big|K'(y)\big|^p dy.$$

Hence, for any $x \in \mathbb{R}$ such that $g(x) > 0$, it exist a positive constant c_p such that

$$\mathbb{E}\big[\big|c_n(x)\big|^p\big] \le \frac{c_p}{h_n^{2p-1}}. \tag{B.9}$$

Therefore, we deduce from (B.7) together with (B.8) and (B.9) that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\mathbb{E}\big[\big|\Delta M_n^C(x)\big|^p|\mathcal{F}_{n-1}\big] \le 2^{p-1}\Big(\frac{c_p}{h_n^{2p-1}} + C_p\Big). \tag{B.10}$$

On the other hand, we also have from assumption $(\mathcal{A}_3)$ that

$$\mathbb{E}\big[\big|\Delta D_n(x)\big|^p|\mathcal{F}_{n-1}\big] = \mathbb{E}\big[|d_n(x)|^p|\mathcal{F}_{n-1}\big] = \mathbb{E}\big[|\varepsilon_n|^p|w_n(X_n, x)|^p|\mathcal{F}_{n-1}\big]$$
$$= \mathbb{E}\big[|\varepsilon_n|^p|\mathcal{F}_{n-1}\big]\mathbb{E}\big[|w_n(X_n, x)|^p\big]$$
$$\le E_p\mathbb{E}\big[|w_n(X_n, x)|^p\big] \tag{B.11}$$

where

$$w_n(X_n, x) = \frac{v_n(X_n, x)}{g(X_n)} = \frac{1}{h_n^2 g(X_n)} K'\left(\frac{x - X_n}{h_n}\right)$$

and E_p stands for the finite random variable

$$E_p = \sup_{n\ge1} \mathbb{E}\big[|\varepsilon_n|^p|\mathcal{F}_{n-1}\big].$$

Moreover, as in the proof of (B.9), we obtain that for any $x \in \mathbb{R}$ such that $g(x) > 0$, it exist a positive constant w_p such that

$$\mathbb{E}\big[\big|w_n(X_n, x)\big|^p\big] \le \frac{w_p}{h_n^{2p-1}}. \tag{B.12}$$

Hence, it follows from (B.11) and (B.12) that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\mathbb{E}\big[\big|\Delta D_n(x)\big|^p|\mathcal{F}_{n-1}\big] \le \frac{E_p w_p}{h_n^{2p-1}}. \tag{B.13}$$

Consequently, we deduce from (B.6) together with (B.10) and (B.13) that for any $x \in \mathbb{R}$ such that $g(x) > 0$, one can find a positive finite random variable M_p such that, for all $n \ge 1$,

$$\mathbb{E}\big[\|\Delta \mathcal{M}_n(x)\|^p|\mathcal{F}_{n-1}\big] \le \frac{M_p}{h_n^{2p-1}} \quad \text{a.s.} \tag{B.14}$$

We recall that $p = 2(1 + a)$. For any $\varepsilon > 0$, if $\mathcal{A}_k(x, \varepsilon, n) = \big\{\|\Delta \mathcal{M}_k(x)\| \ge \varepsilon\sqrt{n^{1+3\alpha}}\big\}$, we have from (B.14),

$$\frac{1}{n^{1+3\alpha}}\sum_{k=1}^{n}\mathbb{E}\big[\|\Delta\mathcal{M}_k(x)\|^2\mathrm{I}_{\mathcal{A}_k(x,\varepsilon,n)}|\mathcal{F}_{k-1}\big] \leq \frac{1}{\varepsilon^{p-2}n^b}\sum_{k=1}^{n}\mathbb{E}\big[\|\Delta\mathcal{M}_k(x)\|^p|\mathcal{F}_{k-1}\big]$$
$$\leq \frac{M_p}{\varepsilon^{p-2}n^b}\sum_{k=1}^{n}\frac{1}{h_k^{2p-1}} \quad \text{a.s.}$$
$$\leq \frac{M_p n^c}{\varepsilon^{p-2}} \quad \text{a.s.}$$

where $b = (a+1)(1+3\alpha)$ and $c = a(\alpha-1)$. Since $c < 0$, the Lindeberg condition is clearly satisfied. Finally, we can conclude from the central limit theorem for martingales (Duflo 1997) that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\frac{1}{\sqrt{n^{1+3\alpha}}}\mathcal{M}_n(x) \xrightarrow{\mathcal{D}} \mathcal{N}\big(0, \Gamma(x)\big), \tag{B.15}$$

where

$$\Gamma(x) = \frac{\xi^2}{(1+3\alpha)g(x)}\begin{pmatrix} f^2(x) & 0 \\ 0 & \sigma^2 \end{pmatrix}.$$

Hence, (B.1) together with (B.4) and (B.15) immediately leads to

$$\sqrt{nh_n^3}\big(\breve{f}_n'(x) - f'(x)\big) \xrightarrow{\mathcal{D}} \mathcal{N}\Big(0, \frac{\xi^2}{1+3\alpha}\frac{f^2(x)+\sigma^2}{g(x)}\Big).$$

□

Proof of convergence (17). It follows from (4) that

$$n\widehat{h}_n(x) = P_n(x) + Q_n(x) = M_n^P(x) + \mathbb{E}[P_n(x)] + Q_n(x) \tag{B.16}$$

where $M_n^P(x) = P_n(x) - \mathbb{E}[P_n(x)]$,

$$P_n(x) = \sum_{k=1}^{n} p_k(x) = \sum_{k=1}^{n} f(X_k)u_k(X_k, x),$$
$$Q_n(x) = \sum_{k=1}^{n} q_k(x) = \sum_{k=1}^{n} \varepsilon_k u_k(X_k, x)$$

with

$$u_n(X_n, x) = \frac{1}{h_n}K\left(\frac{x - X_n}{h_n}\right).$$

Hence, for any $x \in \mathbb{R}$ such that $g(x) > 0$, we obtain from (10), (A.14) and (B.16)

$$n(\widetilde{f}_n'(x) - f'(x)) = \frac{1}{g(x)}\big(M_n^A(x) + \mathbb{E}[A_n(x)] + B_n(x)\big)$$
$$- \frac{g'(x)}{g^2(x)}\big(M_n^P(x) + \mathbb{E}[P_n(x)] + Q_n(x)\big) - nf'(x)$$

which leads to the martingale decomposition

$$\sqrt{nh_n^3}\big(\widetilde{f}_n'(x) - f'(x)\big) = \frac{1}{\sqrt{n^{1+3\alpha}}}\big(\langle\widetilde{e}(x), \mathcal{M}_n(x)\rangle + \widetilde{R}_n(x)\big) \tag{B.17}$$

where

$$\widetilde{e}(x) = \frac{1}{g^2(x)} \begin{pmatrix} g(x) \\ g(x) \\ -g'(x) \\ -g'(x) \end{pmatrix}, \quad \mathcal{M}_n(x) = \begin{pmatrix} M_n^A(x) \\ B_n(x) \\ M_n^P(x) \\ Q_n(x) \end{pmatrix},$$

and the remainder

$$\widetilde{R}_n(x) = \frac{1}{g(x)} \mathbb{E}[A_n(x)] - \frac{g'(x)}{g^2(x)} \mathbb{E}[P_n(x)] - nf'(x).$$

We saw in (A.8) that $\mathbb{E}[a_n(x)] = \big(f(x)g(x)\big)' + R_n(x)$ where $\sup_{x\in\mathbb{R}}\big|R_n(x)\big| = O(h_n)$. By the same token, $\mathbb{E}[P_n(x)] = f(x)g(x) + \zeta_n(x)$ where $\sup_{x\in\mathbb{R}}\big|\zeta_n(x)\big| = O(h_n^2)$. Therefore, as soon as $\alpha > 1/5$, we obtain that

$$\sup_{x\in\mathbb{R}}\big|\widetilde{R}_n(x)\big| = O\left(\sum_{k=1}^{n} h_k\right) = o\left(\sqrt{n^{1+3\alpha}}\right). \tag{B.18}$$

Furthermore, as in the proof of (A.15) and (B.5), the predictable quadratic variation of the four-dimensional real martingale $(\mathcal{M}_n(x))$ satisfies, for any $x \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{1}{n^{1+3\alpha}} \langle \mathcal{M}(x)\rangle_n = \Gamma(x) \tag{B.19}$$

where $\Gamma(x)$ is the four-dimensional covariance matrix given by

$$\Gamma(x) = \frac{\xi^2 g(x)}{(1+3\alpha)} \begin{pmatrix} f^2(x) & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Moreover, via the same lines as in the proof of (B.14), we can also show that $(\mathcal{M}_n(x))$ satisfies the Lindeberg condition. Finally, we find from the central limit theorem for martingales (Duflo 1997) that for any $x \in \mathbb{R}$,

$$\frac{1}{\sqrt{n^{1+3\alpha}}} \mathcal{M}_n(x) \xrightarrow{\mathcal{D}} \mathcal{N}\big(0, \Gamma(x)\big),$$

which implies, from (B.17) and (B.18), that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\sqrt{nh_n^3}\big(\widetilde{f}_n'(x) - f'(x)\big) \xrightarrow{\mathcal{D}} \mathcal{N}\left(0, \frac{\xi^2}{1+3\alpha}\frac{f^2(x)+\sigma^2}{g(x)}\right). \qquad \square$$

Proof of convergence (16). First of all, for any $x \in \mathbb{R}$, denote $h(x) = f(x)g(x)$. It follows from (3) and (4) that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\begin{aligned} \widehat{f}_n(x) - f(x) &= \frac{\widehat{h}_n(x)}{\widehat{g}_n(x)} - \frac{h(x)}{g(x)} = \frac{1}{g(x)\widehat{g}_n(x)}\Big(\widehat{h}_n(x)g(x) - h(x)\widehat{g}_n(x)\Big) \\ &= \frac{1}{g(x)\widehat{g}_n(x)}\Big(g(x)\Big(\widehat{h}_n(x) - h(x)\Big) - h(x)\Big(\widehat{g}_n(x) - g(x)\Big)\Big) \\ &= \frac{1}{\widehat{g}_n(x)}\Big(\widehat{h}_n(x) - h(x)\Big) - \frac{h(x)}{g(x)\widehat{g}_n(x)}\Big(\widehat{g}_n(x) - g(x)\Big). \end{aligned} \tag{B.20}$$

By the same token, we obtain from (9) together with tedious but straightforward calculation that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\begin{aligned}\widehat{f}_n'(x) - f'(x) &= \Big(\frac{\widehat{h}_n'(x)}{\widehat{g}_n(x)} - \frac{\widehat{h}_n(x)\widehat{g}_n'(x)}{\widehat{g}_n^2(x)}\Big) - \Big(\frac{h'(x)}{g(x)} - \frac{h(x)g'(x)}{g^2(x)}\Big)\\ &= \frac{1}{\widehat{g}_n(x)}\Big(\widehat{h}_n'(x) - h'(x)\Big) - \frac{f'(x)}{\widehat{g}_n(x)}\Big(\widehat{g}_n(x) - g(x)\Big)\\ &\quad - \frac{\widehat{h}_n(x)}{\widehat{g}_n^2(x)}\Big(\widehat{g}_n'(x) - g'(x)\Big) - \frac{g'(x)}{\widehat{g}_n(x)}\Big(\widehat{f}_n(x) - f(x)\Big).\end{aligned} \tag{B.21}$$

Hence, we obtain from (B.20) and (B.21) that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\begin{aligned}\widehat{f}_n'(x) - f'(x) &= \frac{1}{\widehat{g}_n(x)}\Big(\widehat{h}_n'(x) - h'(x)\Big) + \frac{\widehat{\ell}_n(x)}{\widehat{g}_n^2(x)}\Big(\widehat{g}_n(x) - g(x)\Big)\\ &\quad - \frac{\widehat{h}_n(x)}{\widehat{g}_n^2(x)}\Big(\widehat{g}_n'(x) - g'(x)\Big) - \frac{g'(x)}{\widehat{g}_n^2(x)}\Big(\widehat{h}_n(x) - h(x)\Big)\end{aligned} \tag{B.22}$$

where $\widehat{\ell}_n(x) = f(x)g'(x) - f'(x)\widehat{g}_n(x)$. Therefore, we deduce from identity (B.22) the martingale decomposition

$$\sqrt{nh_n^3}\big(\widehat{f}_n'(x) - f'(x)\big) = \frac{1}{\sqrt{n^{1+3\alpha}}}\big(\langle \widehat{e}_n(x), \mathcal{M}_n(x)\rangle + \widehat{R}_n(x)\big) \tag{B.23}$$

with

$$\widehat{e}_n(x) = \frac{1}{\widehat{g}_n^2(x)}\begin{pmatrix}\widehat{g}_n(x)\\ \widehat{g}_n(x)\\ -\widehat{h}_n(x)\\ \widehat{\ell}_n(x)\\ -g'(x)\\ -g'(x)\end{pmatrix}, \quad \mathcal{M}_n(x) = \begin{pmatrix}M_n^A(x)\\ B_n(x)\\ M_n^V(x)\\ M_n^U(x)\\ M_n^P(x)\\ Q_n(x)\end{pmatrix},$$

where the martingale difference sequences $(M_n^A(x))$, $(B_n(x))$ and $(M_n^P(x))$, $(Q_n(x))$ were previously defined in (A.14) and (B.16), while the martingale difference sequences $(M_n^U(x))$ and $(M_n^V(x))$ are given by $M_n^U(x) = U_n(x) - \mathbb{E}[U_n(x)]$ and $M_n^V(x) = V_n(x) - \mathbb{E}[V_n(x)]$ with

$$U_n(x) = \sum_{k=1}^n u_k(X_k, x) = \sum_{k=1}^n \frac{1}{h_n}K\left(\frac{x - X_n}{h_n}\right),$$

$$V_n(x) = \sum_{k=1}^n v_k(X_k, x) = \sum_{k=1}^n \frac{1}{h_n^2}K'\left(\frac{x - X_n}{h_n}\right).$$

It is not hard to see that the remainder $\widehat{R}_n(x)$, which can be explicitely calculated, plays a negligible role since, as soon as $\alpha > 1/5$,

$$\sup_{x\in\mathbb{R}}\big|\widehat{R}_n(x)\big| = o\left(\sqrt{n^{1+3\alpha}}\right). \tag{B.24}$$

It remains to establish the asymptotic behavior of the six-dimensional real martingale $(\mathcal{M}_n(x))$. As in the proof of (B.5) and (B.19), we can show that for any $x \in \mathbb{R}$,

$$\lim_{n\to\infty}\frac{1}{n^{1+3\alpha}}\langle\mathcal{M}(x)\rangle_n = \Gamma(x) \tag{B.25}$$

where $\Gamma(x)$ is the six-dimensional covariance matrix given by

$$\Gamma(x) = \frac{\xi^2 g(x)}{(1+3\alpha)} \begin{pmatrix} f^2(x) & 0 & f(x) & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 & 0 & 0 \\ f(x) & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Moreover, via the same lines as in the proof of (B.14), $(\mathcal{M}_n(x))$ satisfies the Lindeberg condition. Hence, we obtain from the central limit theorem for martingales (Duflo 1997) that for any $x \in \mathbb{R}$,

$$\frac{1}{\sqrt{n^{1+3\alpha}}} \mathcal{M}_n(x) \xrightarrow{\mathcal{D}} \mathcal{N}\big(0, \Gamma(x)\big). \tag{B.26}$$

Furthermore, it follows from Lemma A.1 that for any $x \in \mathbb{R}$ such that $g(x) > 0$, $\widehat{e}_n(x)$ converges a.s. to $e(x)$ where

$$e(x) = \frac{1}{g^2(x)} \begin{pmatrix} g(x) \\ g(x) \\ -h(x) \\ \ell(x) \\ -g'(x) \\ -g'(x) \end{pmatrix} \tag{B.27}$$

with $\ell(x) = f(x)g'(x) - f'(x)g(x)$. Finally, we deduce from (B.23), (B.24), (B.26) and (B.27) together with Slutsky's lemma that for any $x \in \mathbb{R}$ such that $g(x) > 0$,

$$\sqrt{nh_n^3}\big(\widehat{f}_n'(x) - f'(x)\big) \xrightarrow{\mathcal{D}} \mathcal{N}\big(0, \sigma^2(x)\big) \tag{B.28}$$

where $\sigma^2(x) = \langle e(x), \Gamma(x)e(x)\rangle$. However, as $h(x) = f(x)g(x)$, it is not hard to see that

$$\begin{aligned} \sigma^2(x) &= \frac{\xi^2}{(1+3\alpha)g^3(x)} \begin{pmatrix} g(x) \\ g(x) \\ -f(x)g(x) \end{pmatrix}^T \begin{pmatrix} f^2(x) & 0 & f(x) \\ 0 & \sigma^2 & 0 \\ f(x) & 0 & 1 \end{pmatrix} \begin{pmatrix} g(x) \\ g(x) \\ -f(x)g(x) \end{pmatrix} \\ &= \frac{\xi^2}{(1+3\alpha)g(x)} \begin{pmatrix} 1 \\ 1 \\ -f(x) \end{pmatrix}^T \begin{pmatrix} f^2(x) & 0 & f(x) \\ 0 & \sigma^2 & 0 \\ f(x) & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -f(x) \end{pmatrix} \\ &= \frac{\xi^2 \sigma^2}{(1+3\alpha)g(x)}, \end{aligned}$$

which completes the proof of Theorem 3.2. □

References

Ahmad IA, Lin PE (1976) Nonparametric sequential estimation of a multiple regression function. Bull Math Stat 17(1):63–75

Ahmed H, Ushirobira R, Efimov D, Tran D, Massabuau JC (2015) Velocity estimation of valve movement in oysters for water quality surveillance. IFAC-PapersOnLine 48(11):333–338

Amiri A (2012) Recursive regression estimators with application to nonparametric prediction. J Nonparametr Stat 24(1):169–186

Bercu B, Capderou S, Durrieu G (2017) A nonparametric statistical procedure for the detection of marine pollution. Submitted for publication

Bercu B, Fraysse P (2012) A Robbins–Monro procedure for estimation in semiparametric regression models. Ann Stat 40(2):666–693

Blondin D (2007) Rates of strong uniform consistency for local least squares kernel regression estimators. Stat Probab Lett 77(14):1526–1534

Chambon C, Legeay A, Durrieu G, Gonzalez P, Ciret P, Massabuau JC (2007) Influence of the parasite worm Polydora sp on the behaviour of the oyster crassostrea gigas: a study of the respiratory impact and associated oxidative stress. Mar Biol 152(2):329–338

Choi E, Hall P, Rousson V (2000) Data sharpening methods for bias reduction in nonparametric regression. Ann Stat 28(2):1339–1355

Coudret R, Durrieu G, Saracco J (2015) Comparison of kernel density estimators with assumption on number of modes. Commun Stat Simul Comput 44(1):196–216

Deheuvels P, Mason DM (2004) General asymptotic confidence bands based on kernel-type function estimators. Stat Inference Stoch Process 7(3):225–277

Devroye L, Lugosi G (2001) Combinatorial methods in density estimation. Springer series in statistics. Springer, New York

Devroye L, Wagner TJ (1980) On the L_1 convergence of kernel estimators of regression functions with applications in discrimination. Z Wahrsch Verw Gebiete 51(1):15–25

Doherty FG, Cherry DS, Cairns J Jr (1987) Valve closure responses of the asiatic clam corbicula fluminea exposed to cadmium and zinc. Hydrobiologia 153(2):159–167

Duflo M (1997) Random iterative models, vol 34 of Applications of mathematics (New York), Springer, Berlin

Durrieu G, Grama I, Pham QK, Tricot JM (2015) Nonparametric adaptive estimation of conditional probabilities of rare events and extreme quantiles. Extremes 18(3):437–478

Durrieu G, Pham QK, Foltete AS, Maxime V, Grama I, Le Tilly V, Duval H, Tricot JM, Ben Naceur C, Sire O (2016) Dynamic extreme values modeling and monitoring by means of sea shores water quality biomarkers and valvometry. Environ Monit Assess 188(7):401–408

Garcia-March J, Sanchis Solsona M, Garcia-Carrascosa A (2008) Shell gaping behavior of pinna nobilis l., 1758: circadian and circalunar rhythms revealed by in situ monitoring. Mar Biol 153(4):689–698

Györfi L, Kohler M, Krzyżak A, Walk H (2002) A distribution-free theory of nonparametric regression. Springer series in statistics. Springer, New York

Haslett SK (2001) Coastal systems. Routledge, London

Huang Y, Chen X, Wu WB (2014) Recursive nonparametric estimation for time series. IEEE Trans Inf Theory 60(2):1301–1312

Johnston GJ (1982) Probabilities of maximal deviations for nonparametric regression function estimates. J Multivar Anal 12(3):402–414

Mason DM (2004) A uniform functional law of the logarithm for the local empirical process. Ann Probab 32(2):1391–1418

Masry E, Fan J (1997) Local polynomial estimation of regression functions for mixing processes. Scand J Stat 24(2):165–179

Nadaraya EA (1964) On estimating regression. Theory Probab Its Appl 9(1):141–142

Nadaraya EA (1989) Nonparametric estimation of probability densities and regression curves, vol 20 of Mathematics and its applications (Soviet Series). Kluwer Academic Publishers Group, Dordrecht

Nagai K, Honjo T, Go J, Yamashita H, Oh SJ (2006) Detecting the shellfish killer heterocapsa circularisquama (dinophyceae) by measuring bivalve valve activity with a hall element sensor. Aquaculture 255(1):395–401

Ngerng MH (2011) Recursive nonparametric estimation of local first derivative under dependence conditions. Commun Stat Theory Methods 40(7):1159–1168

Noda K (1976) Estimation of a regression function by the parzen kernel-type density estimators. Ann Inst Stat Math 28(1):221–234

Parzen E (1962) On estimation of a probability density function and mode. Ann Math Stat 33(3):1065–1076

Development Core Team R (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria

Riisgard HU, Lassen J, Kittner C (2006) Valve-gape response times in mussels (mytilus edulis): effects of laboratory preceding-feeding conditions and in situ tidally induced variation in phytoplankton biomass. J Shellfish Res 25(3):901–911

Rosenblatt M (1956) Remarks on some nonparametric estimates of a density function. Ann Math Stat 27(3):832–837

Schuster EF (1972) Joint asymptotic distribution of the estimated regression function at a finite number of distinct points. Ann Math Stat 43(1):84–88

Seifert B, Gasser T (1996) Finite-sample variance of local polynomials: analysis and solutions. J Am Stat Assoc 91(433):267–275

Silverman BW (1986) Density estimation for statistics and data analysis. In: Monographs on statistics and applied probability. Chapman & Hall, London

Sow M, Durrieu G, Briollais L (2011) Water quality assessment by means of HFNI valvometry and high-frequency data modeling. Environ Monit Assess 182(4):155–170

Tran D, Ciret P, Ciutat A, Durrieu G, Massabuau JC (2003) Estimation of potential and limits of bivalve closure response to detect contaminants: application to cadmium. Environ Toxicol Chem 22(4):914–920

Tsybakov AB (2009) Introduction to nonparametric estimation. Springer series in statistics. Springer, New York

Vilar-Fernández JA, Vilar-Fernández JM (1998) Recursive estimation of regression functions by local polynomial fitting. Ann Inst Stat Math 50(4):729–754

Vilar-Fernández JM, Vilar-Fernández JA (2000) Recursive local polynomial regression under dependence conditions. Test 9(1):209–232

Wand MP, Jones MC (1995) Kernel smoothing, In: Monographs on statistics and applied probability (vol 60). Chapman & Hall, Ltd., London

Watson GS (1964) Smooth regression analysis. Sankhyā Indian J Stat Ser A 26:359–372