Recursive kernel density estimation for time series

AMIR ABOUBACAR[1] and MOHAMED EL MACHKOURI[2*]
[1] Université de Lille
UMR 9221-LEM-Lille Economie Management
F-59000 Lille, France
e-mail: aboubacar.amiri@univ-lille3.fr
[2]Laboratoire de Mathématiques Raphaël Salem
UMR CNRS 6085, Université de Rouen Normandie, France.
e-mail: mohamed.elmachkouri@univ-rouen.fr

March 2, 2020

Abstract

We consider the recursive estimation of the probability density function of continuous random variables from a strongly mixing random sample. We revisit here earlier research on this subject by considering a more general class of recursive estimators, including the usual ones. We derive the quadratic mean error of the considered class of estimators. Moreover, we establish a central limit theorem by using Lindeberg's method resulting in a simplification of the existing assumptions on the sequence of smooth parameters and the mixing coefficient. This is the main contribution of this paper. Finally, the feasibility of the proposed estimator is illustrated throughout an empirical study.

Keywords: asymptotic normality; density function; quadratic mean error; recursive kernel estimators; strong mixing.

1 Introduction

A crucial question in statistics is how to estimate the probability density functions of continuous random variables from a random sample. This problem has been extensively analyzed in the nonparametric statistics literature for the cases of both independent and dependent data. Starting from the empirical distribution function, [23] and [26] introduced the well-known kernel density estimator, which is still very popular in the statistical community. However, although classical nonparametric methods are powerful and fairly well established in the statistical field, the *data streams problem* presents new challenges that are not easily solved with traditional methods. The term *data streams* refers to data sets that continuously and rapidly grow over time. At present, modern computing tools and acquisition techniques are allowing practitioners to collect large volumes of data over time. In fact, in many applications, the learning databases used are so large that it may be impossible for researchers to store them. This leads to a number of computational challenges. The topic of data streams is very recent. However, some tools that can help to solve various statistical problems raised by such data date back to the last century. Among these methods, sequential methods for successive experiments have been used for nonparametric density estimation. Sequential estimators have a decisive computational advantage because they can be updated easily as new data items arrive over time. This advantage arises from the fact that a standard estimator must always be completely recomputed when a new observation is collected, which is clearly a drawback in a sequential context, particularly when instantaneous estimations are required.

The first studies that focused on a recursive version of the Parzen-Rosenblatt estimator were presented by [10], [35] and [36]. Since these pioneering papers, recursive kernel density estimators have also been studied by [17], [20], [22], [27], [33], [34] and many others. There are already many published papers on the asymptotic properties of recursive kernel density and regression estimators for independent and identically distributed (i.i.d.) data. However, because most data streams show serial dependence, the assumption of i.i.d. repetitions is often violated in real-life problems. There are also some published papers on the asymptotic properties of recursive kernel density and regression estimators for dependent (weakly and strongly mixing) data: the density estimation problem has been investigated by [14], [15], [20], [28], [29], [30] and others, and the regression estimation problem has been studied by [2], [9], [13], [27], [31], [32] and others.

In this paper, we focus on a family of recursive estimators, introduced by [17], which includes those introduced by [1], [33] and [35]. In [1], the consistency and asymptotic normality of recursive kernel estimators for strongly mixing random variables were demonstrated. We refer to [1] and [9] for a more complete bibliography on the recursive estimation of the density and regression functions for stationary sequences. One can add, among others, a series of recent contributions extending previous works in several directions. For example, [21] extended the result reported in [1] to the larger class of η-dependent random variables introduced by [11], and [8] proposed recursive estimators for the invariant density and drift terms of some ergodic Hamiltonian systems. Density estimation in a data stream context has previously been considered by

[6], who proposed an algorithm based on kernel merging functions, while [7] presented an estimation method based on sequences of self-organizing maps. Recently, [3] also studied density estimation for directional data streams.

Our aim is to provide a new sufficient condition on the strong mixing coefficients for the consistency and asymptotic normality of recursive kernel density estimators. In particular, we improve upon the results of [1] in several ways. We have decided to revisit this earlier research for three major reasons. First, we consider a more general class of recursive estimators introduced by [17], including the family of estimators introduced by [1]. Second, the proof of the central limit theorem given in [1] is based on Bernstein's block decomposition. We propose to simplify this proof by using Lindeberg's method (see [19]). In particular, some key assumptions regarding the sequence of smooth parameters will be relaxed. Finally, we complete the series of previous results by deriving a mean square error convergence of the recursive estimator.

This paper is organized as follows. In Section 2.1, we present the model and the studied estimator. Section 2.2 describes the assumptions and theoretical results. In Section 3, we illustrate the feasibility and efficiency of the recursive estimator through a brief simulation study. Proofs of some technical lemmas and main results are given in Sections 4 and 5, respectively.

2 Asymptotic results

2.1 Presentation of the model

For d a positive integer, let $(X_i)_{i\in\mathbb{Z}}$ be a sequence of $\mathbb{R}^d$-valued random vectors, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. We assume that the random variables $(X_i)_{i\in\mathbb{Z}}$ are identically distributed with common law μ absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$. We denote by f the unknown probability density function of μ. Given two σ-algebras $\mathcal{U}$ and $\mathcal{V}$ of $\mathcal{F}$, we recall the α-mixing coefficient introduced by [25] defined by

$$\alpha(\mathcal{U}, \mathcal{V}) = \sup\{|\mathbb{P}(A \cap B) - \mathbb{P}(A)\mathbb{P}(B)|\,,\ A \in \mathcal{U},\ B \in \mathcal{V}\}.$$

For any positive integer n, we consider the mixing coefficient $\alpha(n)$ defined by

$$\alpha(n) = \sup_{k\in\mathbb{Z}} \alpha\left(\mathcal{F}_{-\infty}^{k-n}, \sigma(X_k)\right) \quad \text{where} \quad \mathcal{F}_{-\infty}^{k-n} = \sigma(X_i\,;\, i \leqslant k-n).$$

Here, we suppose that $\lim_{n\to\infty} \alpha(n) = 0$. In this case, we say that $(X_i)_{i\in\mathbb{Z}}$ is strongly mixing.

Let n be a positive integer and $x \in \mathbb{R}^d$. Consider a density function $K : \mathbb{R}^d \longrightarrow \mathbb{R}_+$ (kernel) and a sequence of positive real numbers $(h_n)_{n\geqslant 0}$ going to 0 as n goes to infinity (bandwidth parameter). In order to estimate f at x from the sample $(X_1, ..., X_n)$, Rosenblatt [26] and Parzen [23] introduced the well known kernel density estimator $f_n^{PR}(x)$ defined by

$$f_n^{PR}(x) = \frac{1}{nh_n^d}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h_n}\right). \tag{2.1}$$

In [1], for any $\ell \in [0, 1]$, Amiri introduced the recursive kernel density estimator $f_n^{\ell}(x)$ defined by

$$f_n^{\ell}(x) = \frac{1}{\sum_{k=1}^{n} h_k^{d(1-\ell)}}\sum_{i=1}^{n} h_i^{-d\ell} K\left(\frac{x - X_i}{h_i}\right) \tag{2.2}$$

which satisfies the autoregressive relation

$$f_{n+1}^{\ell}(x) = \frac{\sum_{i=1}^{n} h_i^{d(1-\ell)}}{\sum_{i=1}^{n+1} h_i^{d(1-\ell)}} f_n^{\ell}(x) + \frac{h_{n+1}^{d(1-\ell)}}{\sum_{i=1}^{n+1} h_i^{d(1-\ell)}} \tilde{f}_{n+1}(x) \text{ where } \tilde{f}_{n+1}(x) := \frac{1}{h_{n+1}^d} K\left(\frac{x-X_{n+1}}{h_{n+1}}\right).$$

The class (2.2) contains the recursive estimators introduced by Wolverton and Wagner [35] ($\ell = 1$) and Deheuvels [10] ($\ell = 0$) and a renormalized version of the one introduced by Wegman and Davies [33] ($\ell = 1/2$). The tuning parameter ℓ plays a role in regulating the quality improvement of the estimator with respect to the variance and the estimation errors. In fact, [1] prove that the MSE decreases with respect to ℓ. Nevertheless, as noted in the reference above, the choice of ℓ does not appear to have a major influence on the quality of the estimation. Therefore, in practice, one can simply use $\ell = 0$ or $\ell = 1$.

Let $(w_k)_{k\geqslant 1}$ be a nonincreasing sequence of positive real numbers satisfying $\sum_{k\geqslant 1} w_k = \infty$. In this work, we consider a general class of recursive estimators introduced by Hall and Patil [17] and defined by

$$f_n^{HP}(x) = \frac{1}{\sum_{k=1}^{n} w_k}\sum_{i=1}^{n} \frac{w_i}{h_i^d} K\left(\frac{x - X_i}{h_i}\right) \tag{2.3}$$

which satisfies

$$f_{n+1}^{HP}(x) = (1 - \gamma_{n+1}) f_n^{HP}(x) + \gamma_{n+1}\tilde{f}_{n+1}(x) \text{ where } \gamma_n := \frac{w_n}{\sum_{i=1}^{n} w_i}.$$

In particular, the class (2.2) is contained in the class (2.3) since $f_n^{\mathrm{HP}} = f_n^{\ell}$ if $w_k = h_k^{d(1-\ell)}$ and $\ell \in [0, 1]$.

2.2 Assumptions and theoretical results

Let $K : \mathbb{R}^d \longrightarrow \mathbb{R}_+$ such that $\int_{\mathbb{R}^d} K(t)dt = 1$ and $\sup_{x\in\mathbb{R}^d} K(x) < \infty$ (i.e. K is a bounded density function). Assume also that $\lim_{\|x\|\to\infty} \|x\|^d K(x) = 0$, $\int_{\mathbb{R}^d} \|u\|^2 K(u)du < \infty$ where $\|.\|$ is the usual norm on $\mathbb{R}^d$ and $\int_{\mathbb{R}^d} u_i K(u)du = 0$ for any $1 \leqslant i \leqslant d$. Let $(h_n)_{n\geqslant 1}$ be a nonincreasing sequence of positive real numbers going to 0 such that $nh_n^d \longrightarrow \infty$ as n goes to infinity. Let also $(w_n)_{n\geqslant 1}$ be a nonincreasing sequence of positive real numbers such that $(w_n h_n^{-d})_{n\geqslant 1}$ is nondecreasing and $\sum_{n\geqslant 1} w_n = \infty$. For any integer $n \geqslant 1$ and any (s,t) in $\mathbb{Z}^2$, we denote

$$A_{n,s,t} := \frac{1}{nh_n^s w_n^t} \sum_{i=1}^{n} h_i^s w_i^t.$$

We consider the following assumptions:

(H1) There exists $(\beta_{0,1}, \beta_{-d,2}) \in \mathbb{R}_+^2$ such that $\lim_{n\to\infty} A_{n,0,1} = \beta_{0,1}$ and $\lim_{n\to\infty} A_{n,-d,2} = \beta_{-d,2}$.

(H2) There exists $(\theta, \kappa_\theta) \in]0,1] \times \mathbb{R}_+$ such that $h_n^{d(1-\theta)} \sum_{i=1}^n w_i^2 \leqslant \kappa_\theta n w_n^2$ and $\sum_{k>0} k^{1/\theta} \alpha(k) < \infty$.

(H3) (i) The law of X_0 is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d$ and its density f is Lipschitz and twice differentiable with bounded second derivatives.

(ii) For any $i \in \mathbb{Z}\backslash\{0\}$, the law of (X_0, X_i) is absolutely continuous with respect to the Lebesgue measure on $\mathbb{R}^d \times \mathbb{R}^d$ and there exists $c > 0$ such that $\sup_{i\in\mathbb{Z}\backslash\{0\}} |f_{0,i}(x,y) - f(x)f(y)| \leqslant c$ for any $(x,y) \in \mathbb{R}^d \times \mathbb{R}^d$ where $f_{0,i}$ is the joint density function of (X_0, X_i).

Assumptions (H1) and (H2) are classical in the context of recursive kernel density estimators (see [2], [20], [21], [33] and many others). In fact, a link between the bandwidth parameters h_n, the weights w_n and the strong mixing coefficients $\alpha(n)$ is necessary for the consistency and the asymptotic normality of the considered estimator. In particular, these conditions are satisfied if $h_n^d = n^{-\gamma}$ for $0 < \gamma < 1$ and $w_n = n^{-\nu}$ for $0 < \nu \leqslant \gamma \wedge (1/2 + \gamma(1-\theta)/2)$ and $0 < \theta < 1$ (see Proposition 2 below). Assumption (H3) is also standard in kernel density estimation (see [2], [4], [5], [21] and references therein). However, we lay emphasis on that our assumption (H3)(ii) is less restrictive than the condition $\sup_{i\neq j} \|f_{i,j}\|_\infty < \infty$ and $\|f\|_\infty < \infty$ assumed for example in [4] (Proposition 4.2 on page 161).

The main contribution of this work is to lay emphasis on that the consistency and the asymptotic normality of the recursive estimator (2.3) hold under mild conditions on the bandwidth parameter h_n and the strong mixing coefficient $\alpha(n)$ (see assumptions (H1) and (H2)). In the sequel, for any x in $\mathbb{R}^d$, we denote

$$\sigma_x^2 := \frac{\beta_{-d,2} f(x)}{\beta_{0,1}^2} \int_{\mathbb{R}^d} K^2(t)dt. \tag{2.4}$$

Our first result is the following.

Proposition 1. *If* (H1), (H2) *and* (H3) *hold then for any* $x \in \mathbb{R}^d$,

$$\lim_{n\to\infty} |nh_n^d \mathbb{V}[f_n^{HP}(x)] - \sigma_x^2| = 0 \tag{2.5}$$

where σ_x^2 *is defined by* (2.4).

Proposition 1 is an extension of Theorem 2.2.6 in [1] where the particular case $w_n = h_n^{d(1-\ell)}$ with $\ell \in [0,1]$ was considered. More precisely, with our notations, it is proved in [1] that if $nh_n^{d+2} \longrightarrow \infty$ as n goes to infinity, $\alpha(n) = O(n^{-\rho})$ with $\rho > 0$ and there exists $\beta_{s,0} \in \mathbb{R}_+$ such that $\lim_{n\to\infty} A_{n,s,0} = \beta_{s,0}$ for any $s \in]-\infty, d+2]$ then (2.5) holds when

$$(\ell,\rho) \in [((1/2 - 1/d)^+, 1] \times]2, +\infty[\text{ and } d \geqslant 1 \tag{2.6}$$

or

$$(\ell,\rho) \in [0, 1/2 - 1/d[\times]1 + d/2, +\infty[\text{ and } d \geqslant 3 \tag{2.7}$$

where $x^+ = \max\{x, 0\}$ for any real x. One can notice that (2.6) implies (H2) with $\theta = 1$. So, (2.5) holds when $\sum_{n>0} n\alpha(n) < \infty$ which is weaker than $\alpha(n) = O(n^{-\rho})$ with $\rho > 2$. Similarily, (2.7) implies (H2) with $\theta = 2(1 + d\ell)/d$ and consequently (2.5) holds when $\sum_{n\geqslant 1} n^{\frac{d}{2(d\ell+1)}} \alpha(n) < \infty$ which is weaker than $\alpha(n) = O(n^{-\rho})$ with $\rho > 1 + d/2$.

Proposition 2. *Assume that* (H1) *and* (H3) *hold. If* (H2) *holds with* $0 < \theta < 1$ *and there exist* $0 < \gamma < 1$ *such that* $h_n^d = n^{-\gamma}$ *and* $0 < \nu \leqslant \nu^*$ *where* $\nu^* = \gamma \wedge (1/2 + \gamma(1-\theta)/2)$ *such that* $w_n = n^{-\nu}$ *then* $\arg\min_{0<\nu\leqslant\nu^*} \sigma_x^2 = \nu^*$ *and*

$$\min_{0<\nu\leqslant\nu^*} \sigma_x^2 = \psi(\nu^*) f(x) \int_{\mathbb{R}^d} K^2(t)dt \quad \textit{where} \quad \psi(\nu^*) = \frac{(1-\nu^*)^2}{1 - 2\nu^* + \gamma}$$

and σ_x^2 *is defined by* (2.4).

In Proposition 2, we note that if $\gamma \leqslant 1/(1+\theta)$ then $\nu^* = \gamma$ and $\theta/(1+\theta) \leqslant \psi(\nu^*) = 1 - \gamma < 1$ whereas $\nu^* = 1/2 + \gamma(1-\theta)/2$ and $\theta/4 < \psi(\nu^*) = (1 - \gamma(1-\theta))^2/(4\theta\gamma) \leqslant \theta/(1+\theta) \leqslant 1/2$ if $\gamma \geqslant 1/(1+\theta)$. So, we understand that the recursive estimator f_n^{HP} allows us to halve at least the asymptotic variance of the non recursive estimator f_n^{PR} when $w_n = h_n^d = n^{-\gamma}$ with $1/(1+\theta) \leqslant \gamma < 1$.

We obtain also the convergence to zero of the mean square error of the recursive estimator. For any sequences $(p_n)_{n\geqslant 1}$ and $(q_n)_{n\geqslant 1}$ of positive numbers, the notation $p_n \lesssim q_n$ means that there exists $c > 0$ (not depending on n) such that $p_n \leqslant cq_n$.

Proposition 3. *If* (H1), (H2) *and* (H3) *hold then for any* $x \in \mathbb{R}^d$,

$$|\mathbb{E}[f_n^{HP}(x)] - f(x)| \lesssim \frac{\sum_{i=1}^n w_i h_i^2}{\sum_{k=1}^n w_k} \quad \textit{and} \quad \mathbb{V}[f_n^{HP}(x)] \lesssim \frac{1}{nh_n^d}.$$

So, if $A_{n,2,1} \lesssim 1$ *then* $\left|\mathbb{E}[f_n^{HP}(x)] - f(x)\right| \lesssim h_n^2$ *and* $\mathbb{E}[(f_n^{HP}(x) - f(x))^2] \lesssim n^{-\frac{4}{d+4}}$ *for* $h_n = n^{-\frac{1}{d+4}}$.

The asymptotic normality of the estimator (2.3) is given by the following result which is the main contribution of this paper.

Theorem 1. *Assume that* (H1) *and* (H3) *hold. If* (H2) *holds with* $\theta \in]0,1]$ *such that* $nh_n^{d(1+\theta)} \longrightarrow \infty$ *then for any* $x \in \mathbb{R}^d$ *such that* $f(x) > 0$,

$$\sqrt{nh_n^d}(f_n^{HP}(x) - \mathbb{E}[f_n^{HP}(x)]) \xrightarrow[n\to\infty]{Law} \mathcal{N}(0, \sigma_x^2)$$

where σ_x^2 *is defined by* (2.4).

Theorem 1 is also an extension of Theorem 2.2.7 in [1] where the author obtained the asymptotic normality of the recursive estimator (2.2) when $\alpha(n) = O(\rho^n)$ with $0 < \rho < 1$. Analogous results of Theorem 1 have been established for the (non recursive) Parzen-Rosenblatt kernel density estimator for strong mixing random fields (see [12]) and weak dependent time series (see [4]) using the Lindeberg's method which appears to be a flexible approach for proving central limit theorems under mild assumptions both for recursive and non recursive estimators. The following result is a direct consequence of Proposition 3, Theorem 1 and Slutsky's lemma.

Theorem 2. *Assume that* (H1) *and* (H3) *hold. If* (H2) *holds with* $\theta \in]0, 4/d[$ *such that* $nh_n^{d(1+\theta)} \longrightarrow \infty$, $nh_n^{d+4} \longrightarrow 0$ *and* $A_{n,2,1} \lesssim 1$ *then for any* $x \in \mathbb{R}^d$ *such that* $f(x) > 0$,

$$\sqrt{nh_n^d}(f_n^{HP}(x) - f(x)) \xrightarrow[n\to\infty]{Law} \mathcal{N}(0, \sigma_x^2)$$

where σ_x^2 *is defined by* (2.4).

As usual, the conditions $nh_n^{d+4} \longrightarrow 0$ and $A_{n,2,1} \lesssim 1$ are only assumed in order to control the bias of the estimator. In particular, one can notice that this control is achieved when $w_n = h_n^d = n^{-\gamma}$ with $d/(d+4) < \gamma < d/(d+2)$. Finally, we recall that Masry [20] and Khardani and Slaoui [18] obtained also the asymptotic normality of the recursive estimator introduced by Wolverton and Wagner [35] and for a large class of recursive estimators defined by a stochastic approximation algorithm respectively for strong mixing time series. Their results are obtained under some entrelaced conditions in a complicated way between the bandwidth parameters and the strong mixing coefficients which are difficult to compare with our assumptions (see Theorem 8 p. 265 in [20] and Theorem 2.6 p. 36 in [18]). Such kind of conditions are inherent to the Bernstein's method and could be relaxed by using Lindeberg's method.

3 A toy example

In this section, we provide a simulation study to illustrate the asymptotic results of the recursive kernel density estimator obtained in Section 2. We generated $M = 100$ times random samples of size $n = 300$ from: (i) a normal mixture distribution made up of observations from $\mathcal{N}(\mu = 1, \sigma^2 = 0.25)$ and $\mathcal{N}(\mu = -1, \sigma^2 = 0.25)$, each with probability 0.5. (ii) an AR(1) model given by $X_t = \frac{1}{\sqrt{2}} X_{t-1} + \varepsilon_t$, where ε_t are iid standard normal variables. For both cases, we calculate from the simulated data set the kernel density estimator (2.3) at a fixed point $x \in \mathbb{R}$, using the Gaussian kernel, a bandwidth parameter $h_n = n^{-1/5}$ and weights functions $w_n = h_n$. These parameters are arbitrarily fixed to illustrate the feasibility of the estimator trough a simple example. We do not investigate in this work any procedure for a data-driven choice of parameters. Such a study is an important task and will be done in a forthcoming paper. Figure 1 shows kernel estimates of the density for these data with recursive (dashed curve) and non recursive (dotted curve) procedure, along with the true underlying density (the solid curve) for $x = -1$. The 300 data points are marked by vertical lines on the horizontal axis. Inspection of Figure 1 shows that the recursive estimator fits well to the target density function. Also, estimator (2.3) is competitive with respect to its natural competitor introduced by [26] and [23].

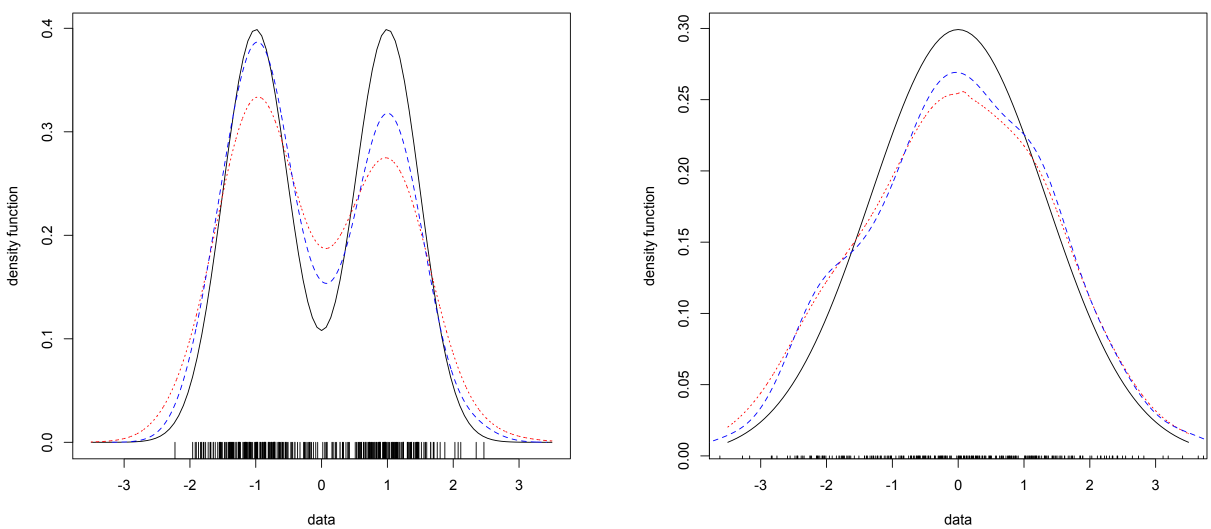

Figure 1: Kernel density estimates using recursive (dashed curve) and non recursive (dotted curve) methods along with the true underlying density (solid curve) of a normal mixture distribution (left) and an AR(1) model (right).

In order to illustrate the result obtained in Proposition 3, we calculate $(f_n^{HP}(x) - f(x))^2$ for a given sample size n and x randomly selected in [-100, 100]. Then, we obtain the value of the mean square error $MSE := \mathbb{E}[(f_n^{HP}(x) - f(x))^2]$ by taking the arithmetic average of the above value over 100 replications of a Monte-Carlo procedure. Results are summarized in Table 1 for several values of sample size. These results confirm empirically the tightness of the mean square error of the recursive estimator f_n^{HP} as n goes to infinity.

n	50	100	150	200	250	300
normal mixture	0.0042	0.0039	0.0034	0.0026	0.0026	0.0020
AR (1)	0.0056	0.0051	0.0040	0.0034	0.0033	0.0029

Table 1: MSE for kernel density estimates using recursive method (results are multiplied by 100).

Turning to the asymptotic normality result given in Theorem 1, our purpose is to compare the distribution of the random variable $S_n(x) := \sqrt{nh_n^d/\sigma_x^2}(f_n^{HP}(x) - \mathbb{E}[f_n^{HP}(x)])$ with a standard normal law. In our simulation study, let $x = -1$ and $n = 300$. The value of σ_{-1}^2 is totally computable, whereas the value of $\mathbb{E}[f_{300}^{HP}(-1)]$ is approximated by the average over a sample of 100 realizations of the random variable $f_{300}^{HP}(-1)$. Figure 2 displays the histogram of 100 copies of the random variable $S_{300}(-1)$ along with the standard normal density. One can observe that the obtained result seems to fit well to the target distribution, that is the standard normal distribution.

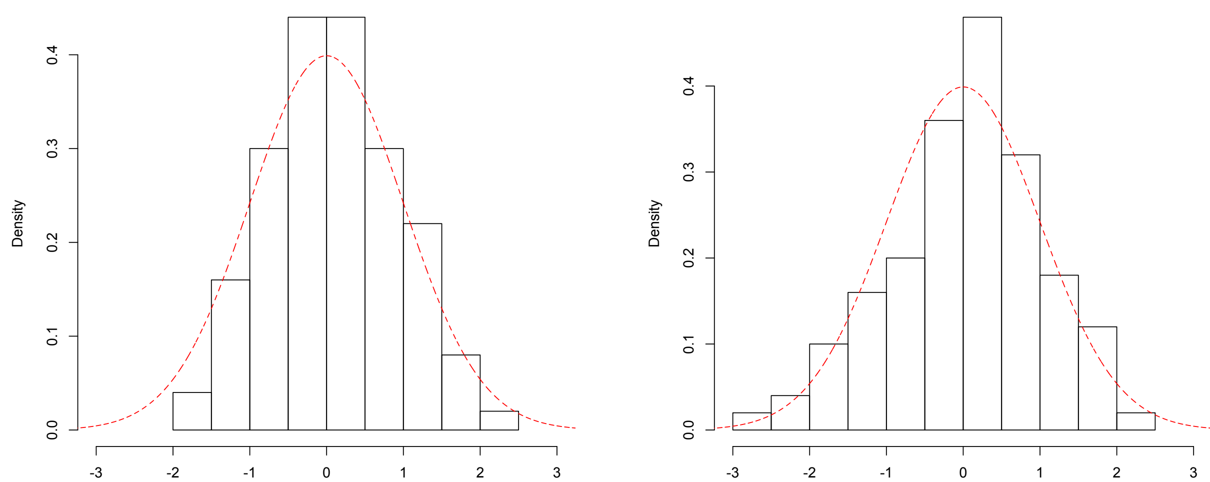

Figure 2: Histograms of $S_{300}(-1)$ along with the standard normal density when the density is estimated using data from a normal mixture (left) and AR (1) model (right).

4 Preliminary lemmas

The following technical lemmas will be usefull in the proof of our main results in section 5.

Lemma 1. *If there exists* $\theta > 0$ *such that* $\sum_{k>0} k^{1/\theta}\alpha(k) < \infty$ *then there exists a sequence* $(m_n)_{n\geqslant 1}$ *of positive integers going to infinity as n goes to infinity such that*

$$\lim_{n\to\infty} m_n h_n^{d\theta} = \lim_{n\to\infty} \frac{1}{h_n^d} \sum_{j\geqslant m_n} \alpha(j) = 0.$$

One can notice that $m_n = o(n)$ since $nh_n^d \longrightarrow \infty$ as $n \longrightarrow \infty$.

Proof of Lemma 1. For any positive integer n, we define

$$m_n := \max\left\{ v_n\,, \left\lfloor h_n^{-d\theta} \left(\sum_{j\geqslant v_n} j^{1/\theta}\alpha(j) \right)^{\theta/2} \right\rfloor + 1 \right\}$$

where $v_n := \lfloor h_n^{\frac{-d\theta}{2}} \rfloor$ and $\lfloor s \rfloor$ stands for the largest integer less than s for any s in $\mathbb{R}$. Since $\sum_{k>0} k^{1/\theta}\alpha(k) < \infty$, we have $m_n h_n^{d\theta} \le \max\left\{ h_n^{\frac{d\theta}{2}}, \left(\sum_{j\geqslant v_n} j^{1/\theta}\alpha(j)\right)^{\theta/2} + h_n^{d\theta} \right\} \xrightarrow[n\to\infty]{} 0$. Noting that $v_n \leqslant m_n$, we have $m_n h_n^{d\theta} \geqslant \left(\sum_{j\geqslant m_n} j^{1/\theta}\alpha(j)\right)^{\theta/2}$. Consequently, we derive

$$\frac{1}{h_n^d} \sum_{j\geqslant m_n} \alpha(j) \leqslant \frac{1}{(m_n h_n^{d\theta})^{1/\theta}} \sum_{j\geqslant m_n} j^{1/\theta}\alpha(j) \leqslant \left(\sum_{j\geqslant m_n} j^{1/\theta}\alpha(j) \right)^{1/2} \xrightarrow[n\to\infty]{} 0.$$

The proof of Lemma 1 is complete. □

Lemma 2. *Let* $(a_n)_{n\geqslant 1}$ *be a sequence of real numbers going to* $a \in \mathbb{R}$ *as n goes to infinity. If* $\lim_{n\to\infty} A_{n,-d,2} = \beta_{-d,2} \in \mathbb{R}$ *then*

$$\lim_{n\to\infty} \frac{h_n^d}{n w_n^2} \sum_{i=1}^{n} \frac{w_i^2 a_i}{h_i^d} = a\beta_{-d,2}.$$

Proof of Lemma 2. It follows from Toepliz's lemma (see for example Masry [20]). It is left to the reader. □

For any $1 \leqslant i \leqslant n$ and any x in $\mathbb{R}^d$, we denote

$$\Delta_i(x) = \frac{n h_n^{d/2} w_i Z_i(x)}{h_i^d \sum_{k=1}^n w_k} \tag{4.1}$$

where $Z_i(x) = K_i(x, X_i) - \mathbb{E}[K_i(x, X_i)]$.

Lemma 3. *If* (H1) *and* (H3) *hold then for all* $x \in \mathbb{R}^d$*, there exists* $\kappa = \kappa(x) > 0$ *such that for all* $1 \leqslant i \neq j \leqslant n$*,*

$$\mathbb{E}[|\Delta_i(x)\Delta_j(x)|] \leqslant \frac{\kappa w_i w_j h_n^d}{w_n^2}.$$

Proof of Lemma 3. Let $1 \leqslant i \neq j \leqslant n$ and $x \in \mathbb{R}^d$ be fixed. Then, we have

$$\mathbb{E}[|Z_i(x)Z_j(x)|] \leqslant \mathbb{E}[K_i(x, X_i)K_j(x, X_j)] + 3\mathbb{E}[K_i(x, X_i)]\mathbb{E}[K_j(x, X_j)].$$

Moreover,

$$\mathbb{E}[K_i(x, X_i)] = h_i^d \int_{\mathbb{R}^d} K(v)f(x - vh_i)dv$$

and using assumption (H3)(ii), we have

$$\begin{aligned}
\mathbb{E}[K_i(x, X_i)K_j(x, X_j)] &= \iint_{\mathbb{R}^d \times \mathbb{R}^d} K_i(x, u)K_j(x, v)f_{0,j-i}(u, v)dudv \\
&\lesssim \iint_{\mathbb{R}^d \times \mathbb{R}^d} K_i(x, u)K_j(x, v)dudv + \iint_{\mathbb{R}^d \times \mathbb{R}^d} K_i(x, u)K_j(x, v)f(u)f(v)dudv \\
&\leqslant h_i^d h_j^d \left(\left(\int_{\mathbb{R}^d} K(v)dv \right)^2 + \int_{\mathbb{R}^d} K(v)f(x - vh_i)dv \times \int_{\mathbb{R}^d} K(v)f(x - vh_j)dv \right).
\end{aligned}$$

Since f is continuous, using Theorem 1A in [23], we have

$$\lim_{n \to \infty} \int_{\mathbb{R}^d} K(v)f(x - vh_n)dv = f(x) \tag{4.2}$$

and consequently, we get

$$\mathbb{E}[K_i(x, X_i)K_j(x, X_j)] \lesssim h_i^d h_j^d \quad \text{and} \quad \mathbb{E}[K_i(x, X_i)] \lesssim h_i^d. \tag{4.3}$$

Consequently, we obtain

$$\mathbb{E}[|Z_i(x)Z_j(x)|] \lesssim h_i^d h_j^d. \tag{4.4}$$

Now, keeping in mind (4.1), we have

$$\Delta_i(x)\Delta_j(x) = \frac{A_{n,0,1}^{-2} h_n^d w_i w_j Z_i(x)Z_j(x)}{h_i^d h_j^d w_n^2}. \tag{4.5}$$

Combining (4.4) and (4.5) and using (H1)(i), we derive

$$\mathbb{E}[|\Delta_i(x)\Delta_j(x)|] \lesssim \frac{w_i w_j h_n^d}{w_n^2}.$$

The proof of Lemma 3 is complete. $\square$

For any real random variable X and any $p > 0$, we denote $\|X\|_p := (\mathbb{E}[|X|^p])^{1/p}$. The following inequality is a classical result for strongly mixing random variables (see for example [16]).

Lemma 4. *Let $\mathcal{U}$ et $\mathcal{V}$ be two σ-algebras and let X be a real random variable measurable with respect to $\mathcal{U}$. If $1 \leqslant p \leqslant r \leqslant \infty$ then $\|\mathbb{E}[X / \mathcal{V}] - \mathbb{E}[X]\|_p \leqslant 2\left(2^{1/p} + 1\right)\|X\|_r\, \alpha(\mathcal{U}, \mathcal{V})^{\frac{1}{p} - \frac{1}{r}}$.*

Proof of Lemma 4. See [16]. $\square$

5 Proofs of the main results

Let n be a positive integer and let $x \in \mathbb{R}^d$ be fixed. Recall that

$$f_n^{HP}(x) = \frac{1}{\sum_{k=1}^n w_k} \sum_{i=1}^n \frac{w_i K_i(x, X_i)}{h_i^d} \quad \text{where} \quad K_i(x,v) = K\left(\frac{x-v}{h_i}\right) \text{ for any } v \in \mathbb{R}^d.$$

Proof of Proposition 1. In the sequel, we consider the notations $\Delta_i(x)$ and $Z_i(x)$ defined by (4.1).

$$nh_n^d \mathbb{V}\left[f_n^{HP}(x)\right] = \mathbb{E}\left[\sum_{i=1}^n \frac{\Delta_i(x)}{\sqrt{n}}\right]^2 = \frac{1}{n}\sum_{i=1}^n \mathbb{E}[\Delta_i^2(x)] + \frac{2}{n}\sum_{i=1}^n \sum_{j=i+1}^n \mathbb{E}[\Delta_i(x)\Delta_j(x)]. \tag{5.1}$$

Moreover,

$$\begin{aligned}\left|\frac{1}{n}\sum_{i=1}^n \left(\mathbb{E}[\Delta_i^2(x)] - \sigma_x^2\right)\right| &\leqslant \left|\frac{A_{n,0,1}^{-2} h_n^d}{n w_n^2} \sum_{i=1}^n \frac{w_i^2}{h_i^d}\left(h_i^{-d}\mathbb{E}[Z_i^2(x)] - f(x)\int_{\mathbb{R}^d} K^2(v)dv\right)\right| \\ &\quad + \left|\frac{A_{n,-d,2}}{A_{n,0,1}^2} - \frac{\beta_{-d,2}}{\beta_{0,1}^2}\right| f(x)\int_{\mathbb{R}^d} K^2(v)dv.\end{aligned}$$

Using (H1), we have

$$\lim_{n\to\infty}\left|\frac{A_{n,-d,2}}{A_{n,0,1}^2} - \frac{\beta_{-d,2}}{\beta_{0,1}^2}\right| = 0.$$

Since f is Lipschitz, using (4.2), we have

$$\begin{aligned}\left|h_i^{-d}\mathbb{E}[Z_i^2(x)] - f(x)\int_{\mathbb{R}^d} K^2(v)dv\right| &= \left|\int_{\mathbb{R}^d} K^2(v)\left(f(x - vh_i) - f(x)\right)dv - h_i^{2d}\left(\int_{\mathbb{R}^d} K(v) f(x - vh_i)dv\right)^2\right| \\ &\lesssim h_i \int_{\mathbb{R}^d} \|v\| K^2(v)dv + h_i^{2d} \lesssim h_i.\end{aligned}$$

Consequently,

$$\limsup_{n\to\infty}\left|\frac{1}{n}\sum_{i=1}^n \left(\mathbb{E}[\Delta_i^2(x)] - \sigma_x^2\right)\right| \lesssim \limsup_{n\to\infty} \frac{A_{n,0,1}^{-2} h_n^d}{n w_n^2}\sum_{i=1}^n \frac{w_i^2}{h_i^d} h_i = 0 \quad \text{(by Lemma 2)}. \tag{5.2}$$

Now, it suffices to prove

$$\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n \sum_{j=i+1}^n \mathbb{E}[\Delta_i(x)\Delta_j(x)] = 0.$$

Let m_n be defined by Lemma 1 and recall that m_n/n goes to zero as n goes to infinity. Then,

$$\frac{1}{n}\sum_{i=1}^n \sum_{j=i+1}^n \mathbb{E}[\Delta_i(x)\Delta_j(x)] = \frac{1}{n}\sum_{i=1}^n \sum_{j=i+1}^{(i+m_n)\wedge n} \mathbb{E}[\Delta_i(x)\Delta_j(x)] + \frac{1}{n}\sum_{i=1}^n \sum_{j=(i+m_n)\wedge n+1}^n \mathbb{E}[\Delta_i(x)\Delta_j(x)].$$

Consequently, using Lemma 3 and keeping in mind that $(w_k)_{k\geqslant 1}$ is nonincreasing, we derive

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{(i+m_n)\wedge n}\left|\mathbb{E}[\Delta_i(x)\Delta_j(x)]\right| \lesssim \frac{h_n^d}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{(i+m_n)\wedge n}\frac{w_i^2}{w_n^2} \leqslant m_n h_n^{d\theta}\times\frac{h_n^{d(1-\theta)}}{nw_n^2}\sum_{i=1}^{n}w_i^2.$$

Using (H2), we obtain

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{(i+m_n)\wedge n}\left|\mathbb{E}[\Delta_i(x)\Delta_j(x)]\right| \lesssim m_n h_n^{d\theta}. \tag{5.3}$$

Now,

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=(i+m_n)\wedge n+1}^{n}\left|\mathbb{E}[\Delta_i(x)\Delta_j(x)]\right| \lesssim \frac{h_n^d}{nw_n^2}\sum_{i=1}^{n}\sum_{j=(i+m_n)\wedge n+1}^{n}\frac{w_i^2\left|\mathbb{E}[Z_i(x)Z_j(x)]\right|}{h_i^d h_j^d}.$$

By Rio's covariance inequality ([24], Theorem 1.1), for $j > i$, we have

$$\left|\mathbb{E}[Z_i(x)Z_j(x)]\right| \leqslant 2\int_0^{2\alpha(j-i)} \mathrm{Q}_{Z_i(x)}(u)\,\mathrm{Q}_{Z_j(x)}(u)\,du$$

where $\mathrm{Q}_{Z_i(x)}(u) = \inf\{t > 0 \,/\, \mathbb{P}(|Z_i(x)| \geqslant t) \leqslant u\}$, $u \in [0,1]$. Since K is bounded, we have $\sup_{0\leqslant u\leqslant 1}\sup_{x\in\mathbb{R}^d}\mathrm{Q}_{Z_i(x)}(u) \lesssim 1$ and consequently

$$\left|\mathbb{E}[Z_i(x)Z_j(x)]\right| \lesssim \alpha(j-i). \tag{5.4}$$

Keeping in mind that $(h_n)_{n\geqslant 1}$ is nonincreasing and using (H1), we derive

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=(i+m_n)\wedge n+1}^{n}\left|\mathbb{E}[\Delta_i(x)\Delta_j(x)]\right| \lesssim \frac{1}{h_n^d}\sum_{j\geqslant m_n}\alpha(j). \tag{5.5}$$

Combining (5.3) and (5.5) and using Lemma 1, we derive

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{j=i+1}^{n}\left|\mathbb{E}[\Delta_i(x)\Delta_j(x)]\right| \lesssim m_n h_n^{d\theta} + \frac{1}{h_n^d}\sum_{j\geqslant m_n}\alpha(j) \xrightarrow[n\to\infty]{} 0. \tag{5.6}$$

Combining (5.1), (5.2) and (5.6), the proof of Proposition 1 is complete. □

Proof of Proposition 2. Let $(\gamma,\theta) \in]0,1[^2$ such that $h_n^d = n^{-\gamma}$ and $\sum_{n\geqslant 1} n^{1/\theta}\alpha(n) < \infty$. Let also $0 < \nu \leqslant \nu^* := \gamma \wedge (1/2 + \gamma(1-\theta)/2) < 1$ such that $w_n = n^{-\nu}$. One can notice that

$$A_{n,0,1} = \frac{1}{nw_n}\sum_{i=1}^{n} w_i = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{(i/n)^{\nu}} \xrightarrow[n\to\infty]{} \frac{1}{1-\nu} =: \beta_{0,1}$$

and

$$A_{n,-d,2} = \frac{h_n^d}{nw_n^2}\sum_{i=1}^{n}\frac{w_i^2}{h_i^d} = \frac{1}{n}\sum_{i=1}^{n}\frac{1}{(i/n)^{2\nu-\gamma}} \xrightarrow[n\to\infty]{} \frac{1}{1-2\nu+\gamma} =: \beta_{-d,2}.$$

Moreover,

$$B_n := \frac{h_n^{d(1-\theta)}}{n w_n^2} \sum_{i=1}^{n} w_i^2 = \frac{n^{-\gamma(1-\theta)}}{n} \sum_{i=1}^{n} \frac{1}{(i/n)^{2\nu}}.$$

If $\nu < 1/2$ then $B_n \lesssim 1$. If $1/2 < \nu \leqslant 1/2 + \gamma(1-\theta)/2$ then $B_n \lesssim n^{-1+2\nu-\gamma(1-\theta)} \lesssim 1$. Finally, if $\nu = 1/2$ then $B_n \lesssim n^{-\gamma(1-\theta)} \log n \lesssim 1$ since $\theta < 1$. Consequently, assumptions (H1) and (H2) are satisfied when $\theta < 1$. Let ψ be the function defined for any $t \in \mathbb{R}$ by $\psi(t) = (1-t)^2/(1-2t+\gamma)$. Since $\beta_{-d,2}/\beta_{0,1}^2 = \psi(\nu)$ and ψ is decreasing from $\psi(0) = 1/(1+\gamma)$ to $\psi(\nu^*)$ on the interval $[0, \nu^*]$, we get $\arg\min_{0<\nu\leqslant\nu^*} \sigma_x^2 = \nu^*$ and

$$\min_{0<\nu\leqslant\nu^*} \sigma_x^2 = \psi(\nu^*) f(x) \int_{\mathbb{R}^d} K^2(t)\, dt.$$

The proof of Proposition 2 is complete. □

Proof of Proposition 3. Let $x \in \mathbb{R}^d$ and let n be a positive integer. We have

$$\mathbb{E}\left[\left(f_n^{HP}(x) - f(x)\right)^2\right] = \mathbb{V}\left(f_n^{HP}(x)\right) + \left(\mathbb{E}[f_n^{HP}(x)] - f(x)\right)^2.$$

By Proposition 1, we have $\mathbb{V}\left[f_n^{HP}(x)\right] \lesssim 1/(nh_n^d)$. Moreover, we have also

$$\left|\mathbb{E}[f_n^{HP}(x)] - f(x)\right| = \left|\frac{1}{\sum_{k=1}^{n} w_k} \sum_{i=1}^{n} w_i \int_{\mathbb{R}^d} K(v)\left(f(x - vh_i) - f(x)\right) dv\right|.$$

Using Taylor's formula and (H3)(i), we derive

$$\left|\mathbb{E}[f_n^{HP}(x)] - f(x)\right| \lesssim \frac{\sum_{i=1}^{n} w_i h_i^2}{\sum_{k=1}^{n} w_k}.$$

Using (H1) and $A_{n,2,1} \lesssim 1$, we obtain $\left|\mathbb{E}[f_n^{HP}(x)] - f(x)\right| \lesssim \frac{h_n^2 A_{n,2,1}}{A_{n,0,1}} \lesssim h_n^2$. Finally, for $h_n = n^{\frac{-1}{d+4}}$, we get $\mathbb{E}[\left(f_n^{HP}(x) - f(x)\right)^2] \lesssim n^{\frac{-4}{d+4}}$. The proof of Proposition 3 is complete. □

Proof of Theorem 1. Let n be a positive integer and $x \in \mathbb{R}^d$ such that $f(x) > 0$. Then

$$\sqrt{nh_n^d}\left(f_n^{HP}(x) - \mathbb{E}[f_n^{HP}(x)]\right) = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \Delta_i$$

where $\Delta_i = \Delta_i(x)$ is defined by (4.1). Let $\varphi : \mathbb{R} \longrightarrow \mathbb{R}$ be a function with compact support and two continuous derivatives. Let $(Y_i)_{i\geqslant 1}$ be independent Gaussian random variables with mean zero and variance $\mathbb{E}[Y_i^2] = \mathbb{E}[\Delta_i^2]$ which are assumed to be independent of the sequence $(X_i)_{i\geqslant 1}$. Keeping in mind (5.2), it suffices to show that the term

$$I_n(\varphi) = \mathbb{E}\left[\varphi\left(\sum_{i=1}^{n} \frac{\Delta_i}{\sqrt{n}}\right)\right] - \mathbb{E}\left[\varphi\left(\sum_{i=1}^{n} \frac{Y_i}{\sqrt{n}}\right)\right]$$

goes to zero as n goes to infinity. Let $1 \leqslant k \leqslant n$ be a positive integer and define

$$U_k = \Delta_1 + \ldots + \Delta_k + Y_{k+1} + \ldots + Y_n.$$

Applying Lindeberg's method, we write

$$I_n(\varphi) = \sum_{k=1}^{n} \mathbb{E}\left[\varphi\left(\frac{U_k}{\sqrt{n}}\right)\right] - \mathbb{E}\left[\varphi\left(\frac{U_{k-1}}{\sqrt{n}}\right)\right] = \sum_{k=1}^{n} a_k(\varphi) \tag{5.7}$$

where

$$a_k(\varphi) = \mathbb{E}\left[\varphi\left(W_k + \frac{\Delta_k}{\sqrt{n}}\right) - \varphi(W_k)\right] - \mathbb{E}\left[\varphi\left(W_k + \frac{Y_k}{\sqrt{n}}\right) - \varphi(W_k)\right]$$

and

$$W_k = \frac{1}{\sqrt{n}}\left(\Delta_1 + \ldots + \Delta_{k-1} + Y_{k+1} + \ldots + Y_n\right).$$

Using Taylor's formula, we have

$$a_k(\varphi) = \mathbb{E}[\varphi'(W_k) \times \frac{\Delta_k}{\sqrt{n}}] + \mathbb{E}[\frac{\varphi''(\theta_k)\Delta_k^2}{2n}] - \mathbb{E}[\varphi'(W_k) \times \frac{Y_k}{\sqrt{n}}] - \mathbb{E}[\frac{\varphi''(\tau_k)Y_k^2}{2n}]$$

with

$$|\theta_k - W_k| \leqslant \frac{|\Delta_k|}{\sqrt{n}} \quad \text{and} \quad |\tau_k - W_k| \leqslant \frac{|Y_k|}{\sqrt{n}}.$$

Moreover,

$$\begin{aligned} a_k(\varphi) =& \mathbb{E}\left[\varphi'(W_k) \times \frac{\Delta_k - Y_k}{\sqrt{n}}\right] + \mathbb{E}\left[\frac{\left(\varphi''(\theta_k) - \varphi''(W_k)\right)\Delta_k^2}{2n}\right] \\ &- \mathbb{E}\left[\frac{\left(\varphi''(\tau_k) - \varphi''(W_k)\right)Y_k^2}{2n}\right] + \mathbb{E}\left[\varphi''(W_k) \times \frac{\Delta_k^2 - Y_k^2}{2n}\right]. \end{aligned}$$

Let $\varepsilon > 0$ be fixed and let $\delta > 0$ such that $|\varphi''(x) - \varphi''(y)| \leqslant \varepsilon$ for any (x, y) satisfying $|x - y| < \delta$. In particular, if $|\Delta_k| < \delta\sqrt{n}$ and $|Y_k| < \delta\sqrt{n}$ then $|\theta_k - W_k| < \delta$ and $|\tau_k - W_k| < \delta$ and consequently $|\varphi''(\theta_k) - \varphi''(W_k)| \leqslant \varepsilon$ and $|\varphi''(\tau_k) - \varphi''(W_k)| \leqslant \varepsilon$. So,

$$a_k(\varphi) = n^{-1/2}\mathbb{E}[(\Delta_k - Y_k)\varphi'(W_k)] + (2n)^{-1}\mathbb{E}[(\Delta_k^2 - Y_k^2)\varphi''(W_k)] + \mathbb{E}[R_k]$$

where

$$|R_k| \leqslant \frac{\|\varphi''\|_\infty}{n}\left(\Delta_k^2 \mathbb{1}_{|\Delta_k| > \delta\sqrt{n}} + Y_k^2 \mathbb{1}_{|Y_k| > \delta\sqrt{n}}\right) + \frac{\varepsilon}{2n}\left(\Delta_k^2 + Y_k^2\right).$$

Since W_k is independent of Y_k and $\mathbb{E}[Y_k] = 0$, we have

$$n^{-1/2}\mathbb{E}[(\Delta_k - Y_k)\varphi'(W_k)] = n^{-1/2}\mathbb{E}[\Delta_k\varphi'(W_k)]$$

and

$$(2n)^{-1}\mathbb{E}[(\Delta_k^2 - Y_k^2)\varphi''(W_k)] = (2n)^{-1}\mathbb{E}[(\Delta_k^2 - \mathbb{E}[\Delta_k^2])\varphi''(W_k)].$$

So, we obtain

$$a_k(\varphi) = n^{-1/2}\mathbb{E}[\Delta_k\varphi'(W_k)] + (2n)^{-1}\mathbb{E}[(\Delta_k^2 - \mathbb{E}[\Delta_k^2])\varphi''(W_k)] + \mathbb{E}[R_k]. \tag{5.8}$$

Since K is bounded, using assumption (H1), we have

$$|\Delta_k| = \left|\frac{w_k n Z_k h_n^{d/2}}{h_k^d \sum_{i=1}^n w_i}\right| = \left|\frac{w_k Z_k h_n^{d/2} A_{n,0,1}^{-1}}{h_k^d w_n}\right| \lesssim \frac{h_n^{d/2} w_k}{h_k^d w_n}$$

where $Z_k := K_k(x, X_k) - \mathbb{E}[K_k(x - X_k)]$. Since $(w_n h_n^{-d})_{n\geqslant 1}$ is nondecreasing, we obtain $|\Delta_k| \lesssim h_n^{-d/2}$. Consequently,

$$\sum_{k=1}^n \mathbb{E}[|R_k|] \leqslant \frac{\|\varphi''\|_\infty}{\delta\sqrt{nh_n^d}} \times \frac{1}{n}\sum_{k=1}^n \mathbb{E}[\Delta_k^2] + \frac{\|\varphi''\|_\infty}{\delta\sqrt{n}} \times \frac{1}{n}\sum_{k=1}^n \mathbb{E}[|Y_k|^3] + \frac{\varepsilon}{n}\sum_{k=1}^n \mathbb{E}[\Delta_k^2].$$

Noting that $\mathbb{E}[|Y_k|^3] = \sqrt{8/\pi} \times \left(\mathbb{E}[Y_k^2]\right)^{3/2} \lesssim h_n^{-d/2}\mathbb{E}[\Delta_k^2]$ and using (5.2), we obtain

$$\limsup_{n\longrightarrow\infty} \sum_{k=1}^n \mathbb{E}[|R_k|] \lesssim \varepsilon\sigma_x^2. \tag{5.9}$$

Let $(m_n)_{n\geqslant 1}$ be the sequence defined in Lemma 1 and recall that $m_n/n \longrightarrow 0$ as $n \longrightarrow \infty$. Since Δ_k is independent of $(Y_j)_{j\geqslant k}$ and $\mathbb{E}[\Delta_k] = 0$, we have

$$\left|\frac{1}{\sqrt{n}}\sum_{k=1}^{m_n} \mathbb{E}[\Delta_k\varphi'(W_k)]\right| = \left|\frac{1}{\sqrt{n}}\sum_{k=1}^{m_n} \mathbb{E}\left[\Delta_k\left(\varphi'(W_k) - \varphi'\left(\sum_{j=k+1}^n \frac{Y_j}{\sqrt{n}}\right)\right)\right]\right| \lesssim \frac{1}{n}\sum_{k=1}^{m_n}\sum_{j=1}^{k-1} \mathbb{E}[|\Delta_k\Delta_j|].$$

Using Lemma 1 and Lemma 3 and assumption (H2), we derive

$$\frac{1}{n}\sum_{k=1}^{m_n}\sum_{j=1}^{k-1} \mathbb{E}[|\Delta_k\Delta_j|] \lesssim m_n h_n^{d\theta} \times \frac{h_n^{d(1-\theta)}}{n w_n^2}\sum_{j=1}^n w_j^2 \lesssim m_n h_n^{d\theta} \xrightarrow[n\to\infty]{} 0$$

and consequently

$$\lim_{n\longrightarrow\infty} \frac{1}{\sqrt{n}}\sum_{k=1}^{m_n} \mathbb{E}[\Delta_k\varphi'(W_k)] = 0. \tag{5.10}$$

For any function $\psi : \mathbb{R} \longrightarrow \mathbb{R}$, we adopt the notation

$$\psi_{s,t} = \psi\left(\sum_{j=1}^s \frac{\Delta_j}{\sqrt{n}} + \sum_{j=t}^n \frac{Y_j}{\sqrt{n}}\right)$$

for any $0 \leqslant s < t \leqslant n+1$ with the usual convention

$$\psi_{0,t} = \psi\left(\sum_{j=t}^{n} \frac{Y_j}{\sqrt{n}}\right) \quad \text{and} \quad \psi_{s,n+1} = \psi\left(\sum_{j=1}^{s} \frac{\Delta_j}{\sqrt{n}}\right).$$

Now,

$$\sum_{k=m_n+1}^{n} \mathbb{E}[\Delta_k \varphi'(W_k)] = \sum_{k=m_n+1}^{n} \mathbb{E}[\Delta_k \varphi'_{k-m_n,k+1}] + \sum_{k=m_n+1}^{n} \mathbb{E}[\Delta_k(\varphi'_{k-1,k+1} - \varphi'_{k-m_n,k+1})]. \tag{5.11}$$

Moreover,

$$\sum_{k=m_n+1}^{n} \mathbb{E}[\Delta_k \varphi'_{k-m_n,k+1}] = \sum_{k=m_n+1}^{n} \sum_{i=1}^{k-m_n} \mathrm{Cov}\left(\Delta_k, \varphi'_{i,k+1} - \varphi'_{i-1,k+1}\right).$$

Applying again Rio's covariance inequality, for $k > i$, we have

$$\left|\mathrm{Cov}\left(\Delta_k, \varphi'_{i,k+1} - \varphi'_{i-1,k+1})\right)\right| \leqslant 2\int_0^{2\alpha(k-i)} \mathrm{Q}_{|\Delta_k|}(u)\, \mathrm{Q}_{|\varphi'_{i,k+1} - \varphi'_{i-1,k+1}|}(u)\, du.$$

Since $|\Delta_k| \lesssim h_n^{-d/2}$ and $|\varphi'_{i,k+1} - \varphi'_{i-1,k+1}| \lesssim \left(nh_n^d\right)^{-1/2}$, we have

$$\sup_{u\in[0,1]} \mathrm{Q}_{|\Delta_k|}(u) \lesssim h_n^{-d/2} \quad \text{and} \quad \sup_{u\in[0,1]} \mathrm{Q}_{|\varphi'_{i,k+1} - \varphi'_{i-1,k+1}|}(u) \lesssim \left(nh_n^d\right)^{-1/2}$$

and consequently

$$\left|\mathrm{Cov}\left(\Delta_k, \varphi'_{i,k+1} - \varphi'_{i-1,k+1}\right)\right| \lesssim \frac{\alpha(k-i)}{\sqrt{n}h_n^d}.$$

So, we derive

$$n^{-1/2}\left|\sum_{k=m_n+1}^{n} \mathbb{E}\left[\Delta_k \varphi'_{k-m_n,k+1}\right]\right| \lesssim \frac{1}{\sqrt{n}} \sum_{k=m_n+1}^{n} \sum_{i=1}^{k-m_n} \frac{\alpha(k-i)}{\sqrt{n}h_n^d} \leqslant \frac{1}{h_n^d} \sum_{j\geqslant m_n} \alpha(j).$$

Using Lemma 1, we obtain

$$\lim_{n\to\infty} \frac{1}{\sqrt{n}} \sum_{k=m_n+1}^{n} \mathbb{E}[\Delta_k \varphi'_{k-m_n,k+1}] = 0. \tag{5.12}$$

Now,

$$\left|\frac{1}{\sqrt{n}} \sum_{k=m_n+1}^{n} \mathbb{E}\left[\Delta_k(\varphi'_{k-1,k+1} - \varphi'_{k-m_n,k+1})\right]\right| \lesssim \frac{1}{n} \sum_{k=m_n+1}^{n} \sum_{j=k-m_n+1}^{k-1} \mathbb{E}[|\Delta_k \Delta_j|].$$

Noting that

$$\frac{1}{n} \sum_{k=m_n+1}^{n} \sum_{j=k-m_n+1}^{k-1} \mathbb{E}[|\Delta_k \Delta_j|] = \frac{1}{n} \sum_{j=2}^{n-1} \sum_{k=(j+1)\vee(m_n+1)}^{(j+m_n-1)\wedge n} \mathbb{E}[|\Delta_k \Delta_j|]$$

and using Lemma 3, we obtain

$$\frac{1}{n}\sum_{k=m_n+1}^{n}\sum_{j=k-m_n+1}^{k-1}\mathbb{E}[|\Delta_k\Delta_j|] \lesssim m_n h_n^{d\theta}\times\frac{h_n^{d(1-\theta)}}{nw_n^2}\sum_{j=1}^{n}w_j^2.$$

Using assumption (H2) and Lemma 1, we derive

$$\lim_{n\to\infty} n^{-1/2}\left|\sum_{k=m_n+1}^{n}\mathbb{E}\left[\Delta_k\left(\varphi'_{k-1,k+1}-\varphi'_{k-m_n,k+1}\right)\right]\right| = 0. \tag{5.13}$$

Combining (5.10), (5.11), (5.12) and (5.13), we obtain

$$\lim_{n\to\infty}\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\mathbb{E}[\Delta_k\varphi'(W_k)] = 0. \tag{5.14}$$

So, in order to complete the proof, we have to prove that

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}[\left(\Delta_k^2-\mathbb{E}[\Delta_k^2]\right)\varphi''(W_k)] = 0. \tag{5.15}$$

Let $M = \lceil h_n^{-d\theta}\rceil$ where $\lceil x\rceil$ is the smallest integer larger than x for any $x\in\mathbb{R}$. Since $m_n h_n^{d\theta}\longrightarrow 0$ and $nh_n^{d\theta}\longrightarrow\infty$, without loss of generality, we assume that $m_n\leqslant M\leqslant n$. Then,

$$\frac{1}{n}\sum_{k=1}^{n}\mathbb{E}[\left(\Delta_k^2-\mathbb{E}[\Delta_k^2]\right)\varphi''(W_k)] = D_n+E_n+F_n$$

where

$$\begin{aligned}
D_n &= \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}\left[\left(\Delta_k^2-\mathbb{E}[\Delta_k^2|\mathcal{F}_{-\infty}^{k-M}]\right)\varphi''_{(k-M)^+,k+1}\right] = 0,\\
E_n &= \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}\left[\left(\Delta_k^2-\mathbb{E}[\Delta_k^2|\mathcal{F}_{-\infty}^{k-M}]\right)\left(\varphi''_{k-1,k+1}-\varphi''_{(k-M)^+,k+1}\right)\right],\\
F_n &= \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}\left[\left(\mathbb{E}[\Delta_k^2|\mathcal{F}_{-\infty}^{k-M}]-\mathbb{E}[\Delta_k^2]\right)\varphi''_{k-1,k+1}\right]
\end{aligned}$$

with $\mathcal{F}_{-\infty}^{k-M} = \sigma\left(X_j\,;\, j\leqslant k-M\right)$ and $(k-M)^+ = \max\{k-M,0\}$ for any $1\leqslant k\leqslant n$. Moreover,

$$|E_n| \leqslant \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}\left[\left(\Delta_k^2+\mathbb{E}[\Delta_k^2|\mathcal{F}_{-\infty}^{k-M}]\right)\left(2\wedge\left|\sum_{j=(k-M)^++1}^{k-1}\frac{\Delta_j}{\sqrt{n}}\right|\right)\right] \leqslant \frac{1}{n}\sum_{k=1}^{n}\left(E_{1,k}+E_{2,k}+E_{3,k}\right).$$

where

$$E_{1,k} = \mathbb{E}\left[\Delta_k^2\left|\sum_{j=(k-M)^++1}^{k-1}\frac{\Delta_j}{\sqrt{n}}\right|\right],\quad E_{2,k} = 2\|\mathbb{E}[\Delta_k^2|\mathcal{F}_{-\infty}^{k-M}]-\mathbb{E}[\Delta_k^2]\|_1 \ \text{ and }\ E_{3,k} = \mathbb{E}[\Delta_k^2]\left\|\sum_{j=(k-M)^++1}^{k-1}\frac{\Delta_j}{\sqrt{n}}\right\|_1.$$

Noting that $|F_n| \leqslant \frac{1}{2n}\sum_{k=1}^n E_{2,k}$, we obtain

$$\frac{1}{n}\left|\sum_{k=1}^{n} \mathbb{E}\left[(\Delta_k^2 - \mathbb{E}[\Delta_k^2])\varphi''(W_k)\right]\right| \lesssim |E_n|. \tag{5.16}$$

Using $|\Delta_k| \lesssim h_n^{-d/2}$ and Lemma 3, we have

$$\begin{aligned}\frac{1}{n}\sum_{k=1}^{n} E_{1,k} &\lesssim \frac{1}{n^{3/2}h_n^{d/2}}\sum_{k=1}^{n}\sum_{j=(k-M)^+ +1}^{k-1}\mathbb{E}[|\Delta_k\Delta_j|] \lesssim \frac{h_n^{d/2}}{n^{3/2}w_n^2}\sum_{k=1}^{n}\sum_{j=(k-M)^+ +1}^{k-1} w_j^2 \\ &= \frac{h_n^{d/2}}{n^{3/2}w_n^2}\left(\sum_{k=1}^{M}\sum_{j=1}^{k-1} w_j^2 + \sum_{k=M+1}^{n}\sum_{j=k-M+1}^{k-1} w_j^2\right) \leqslant \frac{2Mh_n^{d\theta}}{\sqrt{nh_n^d}} \times \frac{h_n^{d(1-\theta)}}{nw_n^2}\sum_{j=1}^{n} w_j^2.\end{aligned}$$

Using (H2), we obtain

$$\frac{1}{n}\sum_{k=1}^{n} E_{1,k} \lesssim \frac{Mh_n^{d\theta}}{\sqrt{nh_n^d}}. \tag{5.17}$$

In the other part, using Lemma 4 with $p = 1$ and $r = \infty$, we have

$$\frac{1}{n}\sum_{k=1}^{n} E_{2,k} \lesssim h_n^{-d}\alpha(M). \tag{5.18}$$

Now, we note that

$$\frac{1}{n}\sum_{k=1}^{n} E_{3,k} \leqslant \frac{1}{n^{3/2}}\sum_{k=1}^{n}\mathbb{E}[\Delta_k^2] \times \sqrt{\sum_{j=(k-M)^+ +1}^{k-1}\mathbb{E}[\Delta_j^2] + 2\sum_{j=(k-M)^+ +1}^{k-1}\sum_{i=(k-M)^+ +1}^{j-1}\mathbb{E}[\Delta_i\Delta_j]}. \tag{5.19}$$

Keeping in mind that $|\Delta_j| \lesssim h_n^{-d/2}$, we have

$$\sum_{j=(k-M)^+ +1}^{k-1}\mathbb{E}[\Delta_j^2] \lesssim Mh_n^{-d}. \tag{5.20}$$

Lemma 5.1. *For any positive integer M such that $m_n \leqslant M \leqslant n$, we have*

$$\sup_{1\leqslant k\leqslant n}\sum_{j=(k-M)^+ +1}^{k-1}\sum_{i=(k-M)^+ +1}^{j-1}|\mathbb{E}[\Delta_i\Delta_j]| \lesssim n\left(m_n h_n^{d\theta} + h_n^{-d}\sum_{j\geqslant m_n}\alpha(j)\right).$$

Combining (5.2), (5.19), (5.20) and Lemma 5.1, we obtain

$$\frac{1}{n}\sum_{k=1}^{n} E_{3,k} \lesssim \sqrt{\frac{M}{nh_n^d} + m_n h_n^{d\theta} + h_n^{-d}\sum_{j\geqslant m_n}\alpha(j)}. \tag{5.21}$$

From (5.16), (5.17), (5.18) and (5.21) and $M = \lceil h_n^{-d\theta} \rceil$, we derive

$$\frac{1}{n}\left|\sum_{k=1}^{n} \mathbb{E}\left[\left(\Delta_k^2 - \mathbb{E}[\Delta_k^2]\right)\varphi''(W_k)\right]\right| \lesssim \frac{1}{\sqrt{nh_n^d}} + M^{1/\theta}\alpha(M) + \sqrt{\frac{1}{nh_n^{d(1+\theta)}} + m_n h_n^{d\theta} + h_n^{-d}\sum_{j\geqslant m_n}\alpha(j)}.$$

Using Lemma 1 and keeping in mind that $nh_n^{d(1+\theta)} \longrightarrow \infty$ and $\sum_{k>0} k^{1/\theta}\alpha(k) < \infty$, we obtain (5.15). Finally, combining (5.7), (5.8), (5.9), (5.14) and (5.15), we derive

$$\limsup_{n\longrightarrow\infty} |I_n(\varphi)| \lesssim \varepsilon\sigma_x^2.$$

Since ε is arbitrarily small, we obtain $\lim_{n\to\infty} |I_n(\varphi)| = 0$. The proof of Theorem 1 is complete. □

Proof of Lemma 5.1. Let $1 \leqslant k \leqslant n$ and $m_n \leqslant M \leqslant n$ be fixed. In the sequel, we denote

$$S(k) := \sum_{j=(k-M)^+ +1}^{k-1} \sum_{i=(k-M)^+ +1}^{j-1} |\mathbb{E}[\Delta_i\Delta_j]|.$$

Keeping in mind Lemma 3 and (5.4), for any $1 \leqslant i < j \leqslant n$, we have

$$\mathbb{E}[|\Delta_i\Delta_j|] \lesssim \frac{w_i^2 h_n^d}{w_n^2} \quad \text{and} \quad |\mathbb{E}[\Delta_i\Delta_j]| \lesssim \frac{w_i^2 h_n^d \alpha(j-i)}{h_i^d h_j^d w_n^2}.$$

Assume that $k \leqslant m_n \leqslant M \leqslant n$. Using (H2), we have

$$S(k) \lesssim \sum_{j=1}^{k-1}\sum_{i=1}^{j-1} \frac{w_i^2 h_n^d}{w_n^2} \leqslant m_n \sum_{i=1}^{n} \frac{w_i^2 h_n^d}{w_n^2} = n m_n h_n^{d\theta} \times \frac{h_n^{d(1-\theta)}}{n w_n^2}\sum_{i=1}^{n} w_i^2 \lesssim n m_n h_n^{d\theta}.$$

If $m_n < k \leqslant M \leqslant n$ then

$$\begin{aligned}
S(k) = \sum_{j=1}^{k-1}\sum_{i=1}^{j-1} |\mathbb{E}[\Delta_i\Delta_j]| &= \sum_{j=1}^{m_n}\sum_{i=1}^{j-1} |\mathbb{E}[\Delta_i\Delta_j]| + \sum_{j=m_n+1}^{k-1}\sum_{i=1}^{j-m_n} |\mathbb{E}[\Delta_i\Delta_j]| + \sum_{j=m_n+1}^{k-1}\sum_{i=j-m_n+1}^{j-1} |\mathbb{E}[\Delta_i\Delta_j]| \\
&\lesssim \sum_{j=1}^{m_n}\sum_{i=1}^{j-1} \frac{w_i^2 h_n^d}{w_n^2} + \sum_{j=m_n+1}^{k-1}\sum_{i=1}^{j-m_n} \frac{w_i^2 h_n^d \alpha(j-i)}{h_i^d h_j^d w_n^2} + \sum_{j=m_n+1}^{k-1}\sum_{i=j-m_n+1}^{j-1} \frac{w_i^2 h_n^d}{w_n^2} \\
&\lesssim \frac{m_n h_n^d}{w_n^2}\sum_{i=1}^{n} w_i^2 + \frac{1}{w_n^2}\sum_{i=1}^{k-1-m_n} \frac{w_i^2}{h_i^d}\sum_{j=i+m_n}^{k-1}\alpha(j-i) + \sum_{i=2}^{k-2}\frac{w_i^2 h_n^d}{w_n^2}\sum_{j=(i+1)\vee(m_n+1)}^{(i+m_n-1)\wedge(k-1)} 1 \\
&\lesssim n m_n h_n^{d\theta} \times \frac{h_n^{d(1-\theta)}}{n w_n^2}\sum_{i=1}^{n} w_i^2 + \times\frac{h_n^d}{n w_n^2}\sum_{i=1}^{n}\frac{w_i^2}{h_i^d} \times \frac{n}{h_n^d}\sum_{j\geqslant m_n}\alpha(j) + n m_n h_n^{d\theta} \times \frac{h_n^{d(1-\theta)}}{n w_n^2}\sum_{i=1}^{n} w_i^2.
\end{aligned}$$

Using (H1) and (H2), we get

$$S(k) \lesssim n\left(m_n h_n^{d\theta} + h_n^{-d}\sum_{j\geqslant m_n}\alpha(j)\right). \tag{5.22}$$

If $m_n \leqslant M < k \leqslant n$ then

$$
\begin{aligned}
S(k) &= \sum_{j=k-M+1}^{k-M+m_n} \sum_{i=k-M+1}^{j-1} |\mathbb{E}[\Delta_i \Delta_j]| + \sum_{j=k-M+m_n+1}^{k-1} \sum_{i=k-M+1}^{j-m_n} |\mathbb{E}[\Delta_i \Delta_j]| + \sum_{j=k-M+m_n+1}^{k-1} \sum_{i=j-m_n+1}^{j-1} |\mathbb{E}[\Delta_i \Delta_j]| \\
&\lesssim \sum_{j=k-M+1}^{k-M+m_n} \sum_{i=k-M+1}^{j-1} \frac{w_i^2 h_n^d}{w_n^2} + \sum_{j=k-M+m_n+1}^{k-1} \sum_{i=k-M+1}^{j-m_n} \frac{w_i^2 h_n^d \alpha(j-i)}{h_i^d h_j^d w_n^2} + \sum_{j=k-M+m_n+1}^{k-1} \sum_{i=j-m_n+1}^{j-1} \frac{w_i^2 h_n^d}{w_n^2} \\
&\lesssim \frac{m_n h_n^d}{w_n^2} \sum_{i=1}^{n} w_i^2 + \sum_{i=k-M+1}^{k-1-m_n} \sum_{j=i+m_n}^{k-1} \frac{w_i^2 \alpha(j-i)}{h_i^d w_n^2} + \sum_{i=k-M+2}^{k-2} \frac{w_i^2 h_n^d}{w_n^2} \sum_{j=(i+1)\vee(k-M+m_n+1)}^{(k-1)\wedge(i+m_n-1)} 1 \\
&\lesssim n m_n h_n^{d\theta} \times \frac{h_n^{d(1-\theta)}}{n w_n^2} \sum_{i=1}^{n} w_i^2 + \frac{h_n^d}{n w_n^2} \sum_{i=1}^{n} \frac{w_i^2}{h_i^d} \times \frac{n}{h_n^d} \sum_{j \geqslant m_n} \alpha(j) + n m_n h_n^{d\theta} \times \frac{h_n^{d(1-\theta)}}{n w_n^2} \sum_{i=1}^{n} w_i^2 .
\end{aligned}
$$

Using (H1) and (H2), we get again the bound (5.22). The proof of Lemma 5.1 is complete. □

Proof of Theorem 2. The proof is a direct consequence of Proposition 3, Theorem 1 and Slutsky's lemma. It is left to the reader. □

Acknowledgements

We would like to thank two anonymous reviewers for their very careful reading of the manuscript and for their insightful comments and suggestions.

References

[1] A. Amiri. *Estimateurs fonctionnels récursifs et leurs applications à la prévision.* 2010. Thèse de Doctorat.

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

[3] A. Amiri, B. Thiam, and T. Verdebout. On the estimation of the density of a directional data stream. *Scand. J. Stat.*, 44(1):249–267, 2017.

[4] J.-M. Bardet, P. Doukhan, G. Lang, and N. Ragache. Dependent Lindeberg central limit theorem and some applications. *ESAIM Probab. Stat.*, 12:154–172, 2008.

[5] G. Biau and B. Cadre. Nonparametric spatial prediction. *Stat. Inference Stoch. Process.*, 7(3):327–349, 2004.

[6] Z. Cai, W. Qian, L. Wei, and A. Zhou. M-kernel merging: towards density estimation over data streams. *Proc. Database Syst. Adv. Appl. 18th Int. Conf.*, pages 285–292, 2003.

[7] Y. Cao, H. He, and H. Man. SOMKE: kernel density estimation over data streams by sequences of self-organizing maps. *IEEE Trans. on Neural Networks and Learning Syst.*, 23(8):1254–1268, 2012.

[8] P. Cattiaux, José R. León, and C. Prieur. Recursive estimation for stochastic damping hamiltonian systems. *J. Nonparametr. Stat.*, 27(3):401–424, 2015.

[9] X. Chen, Y. Huang, and W.B. Wu. Recursive nonparametric estimation for time series. *IEEE Trans. Inform. Theory*, 60:1301–1312, 2013.

[10] P. Deheuvels. Sur l'estimation séquentielle de la densité. *C. R. Acad. Sci. Paris Sér. A-B*, 276:A1119–A1121, 1973.

[11] P. Doukhan and S. Louhichi. A new weak dependence condition and applications to moment inequalities. *Stochastic Process. Appl.*, 84(2):313–342, 1999.

[12] M. El Machkouri. Asymptotic normality for the Parzen-Rosenblatt density estimator for strongly mixing random fields. *Statistical Inference for Stochastic Processes*, 14(1):73–84, 2011.

[13] L. Gyorfi, M. Kohler, and H. Walk. Weak and strong universal consistency of semirecursive kernel and partitioning regression estimates. *Statist. Decisions*, 16:1–18, 1998.

[14] L. Gyorfi and E. Masry. Strong consistency and rates for recursive probability density estimators of stationary processes. *J. Multivariate Anal.*, 22:79–93, 1987.

[15] L. Gyorfi and E. Masry. The l_1 and l_2 strong consistency of recursive kernel density estimation from dependent samples. *IEEE Trans. Inform. Theory.*, 36:531–539, 1990.

[16] P. Hall and C. C. Heyde. *Martingale limit theory and its application.* Academic Press, New York, 1980.

[17] P. Hall and P. Patil. On the efficiency of on-line density estimators. *IEEE Trans. Information Theory*, 40(5):1504–1512, 1994.

[18] S. Khardani and Y. Slaoui. Recursive kernel density estimation and optimal bandwidth selection under α: mixing data. *J. Stat. Theory Pract.*, 13(2):Art. 36, 21, 2019.

[19] J. W. Lindeberg. Eine neue Herleitung des Exponentialgezetzes in der Wahrscheinlichkeitsrechnung. *Mathematische Zeitschrift*, 15:211–225, 1922.

[20] E. Masry. Recursive probability density estimation for weakly dependent stationary processes. *IEEE Trans. Inform. Theory*, 32(2):254–267, 1986.

[21] K. Mezhoud, Z. Mohdeb, and S. Louhichi. Recursive kernel estimation of the density under η-weak dependence. *J. Korean Statist. Soc.*, 43(3):403–414, 2014.

[22] A. Mokkadem, M. Pelletier, and Y. Slaoui. The stochastic approximation method for the estimation of a multivariate probability density. *J. Statist. Plann. Inference*, 139(7):2459–2478, 2009.

[23] E. Parzen. On the estimation of a probability density and the mode. *Ann. Math. Statist.*, 33:1965–1976, 1962.

[24] E. Rio. Covariance inequalities for strongly mixing processes. *Ann. Inst. H. Poincaré Probab. Statist.*, 29(4):587–597, 1993.

[25] M. Rosenblatt. A central limit theorem and a strong mixing condition. *Proc. Nat. Acad. Sci. USA*, 42:43–47, 1956.

[26] M. Rosenblatt. Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.*, 27:832–837, 1956.

[27] G. Roussas and L.T. Tran. Asymptotic normality of the recursive kernel regression estimate under dependence conditions. *Ann. Stat*, 20:98–120, 1992.

[28] H. Takahata. Almost sure convergence of density estimators for weakly dependent stationary processes. *Bull. Tokyo. Gakugei Univ. Ser. IV*, 32:11–32, 1980.

[29] L.T. Tran. Recursive density estimation under dependence. *IEEE Trans. Inform. Theory.*, 35:1103–1108, 1989.

[30] L.T. Tran. Kernel density estimation on random fields. *J. Multivariate Anal.*, 34:37–53, 1990.

[31] H. Walk and S. Yakowitz. Iterative nonparametric estimation of a log-optimal portfolio selection function. *IEEE Trans. Inform. Theory*, 48:324–333, 2002.

[32] L. Wang and H. Liang. Strong uniform convergence of the recursive regression estimator under ϕ-mixing conditions. *Metrika*, 59:245?261, 2004.

[33] E. J. Wegman and H. I. Davies. Remarks on some recursive estimators of a probability density. *Ann. Statist.*, 7(2):316–327, 1979.

[34] W. Wertz. Recursive and sequential density estimation. In *Sequential methods in statistics*, volume 16 of *Banach Center Publ.*, pages 515–544. PWN, Warsaw, 1985.

[35] C. T. Wolverton and T. J. Wagner. Recursive estimates of probability densities. *IEEE Trans. Syst. Man. Cybern.*, 5:307–308, 1969.

[36] H. Yamato. Sequential estimation of a continuous probability density function and mode. *Bull. Math. Statist*, 14:1–12, 1971.