Limit theorems for kernel density estimators under dependent samples

Yuexu Zhao, Zhengyan Lin
Department of Mathematics, Zhejiang University, Hangzhou 310027*, China*

Abstract. In this paper, we construct a moment inequality for mixing dependent random variables, it is of independent interest. As applications, the consistency of the kernel density estimation is investigated. Several limit theorems are established: First, the central limit theorems for the kernel density estimator $f_{n,K}(x)$ and its distribution function are constructed. Also, the convergence rates of $\|f_{n,K}(x) - Ef_{n,K}(x)\|_p$ in sup-norm loss and integral L^p-norm loss are proved. Moreover, the a.s. convergence rates of the supremum of $|f_{n,K}(x) - Ef_{n,K}(x)|$ over a compact set and the whole real line are obtained. It is showed, under suitable conditions on the mixing rates, the kernel function and the bandwidths, that the optimal rates for i.i.d. random variables are also optimal for dependent ones.

Keywords: Kernel density estimator; consistency; convergence rate; mixing rate
Mathematics Subject Classification (2000): Primary 62G07, 60G10; Secondary 60F15.

1. Introduction

Let $X, X_1, X_2, ...$ be independent and identically distributed (i.i.d.) random variables with common density f, and K be a bounded integrable kernel (a measurable function on $\mathbb{R}$), the classical kernel density estimators (KDEs) of f based on the observations $X_1, ..., X_n$ are defined as

$$f_{n,K}(x) = \frac{1}{nh_n}\sum_{i=1}^{n} K\Big(\frac{X_i - x}{h_n}\Big), \qquad x \in \mathbb{R}, \tag{1.1}$$

where the bandwidths $\{h_n, n \geq 1\}$ satisfy some regularity conditions.

Since the famous work done by Rosenblatt [25] and Parzen [18], the limit behavior for the KDEs has become an active subject. For the case of i.i.d. data, see, for example, Bickel and Rosenblatt [1], Silverman [34] and Stute [35, 36]. Using empirical process approach, Einmahl and Mason [4, 5] studied the uniform consistency and uniform consistency in bandwidth, respectively. Giné and Guillou [8, 9] investigated the exact rates of almost sure (a.s.) convergence of the supremum over adaptive intervals and all of $\mathbb{R}^d$, and Giné, Koltchinskii and Zinn [10] obtained weighted uniform consistency of KDEs, and so forth. As to weakly dependent observations, Földes [6], Rüschendorf [27], Sarda and Vieu [28], Peligrad [20] and Liebscher [13] studied the strong convergence of density estimators for $\phi-$mixing samples. Rosenblatt [25], Nze and Rios [17], Liebscher [15] investigated a.s. convergence of kernel estimators for $\alpha-$mixing random variables. For other results, one can refer to Neumann [16], Woodroofe [37, 38], Wu et al. [39], Yakowitz [40], and the reference therein. However, most of the work mentioned as above on a.s. convergence rates in sup-norm loss under dependent data are not optimal. Yu [41] obtained the best possible minimax rates for stationary sequences satisfying certain $\beta-$mixing conditions at the cost of sufficient smoothness for density functions. The purpose of the present article is to investigate the consistency of the KDEs, and tries to get the optimal convergence rates for certain dependent observations. More precisely, we require the random variables to be $\rho-$mixing, which is defined as follows:

Supported by the National Natural Science Foundation of China (Grant Nos. 11171303, 61273093) and the Specialized Research Fund for the Doctor Program of Higher Education (Grant No. 20090101110020).

Definition. *Suppose that* $X_1, X_2, ...$ *is a sequence of random variables on a probability space* $(\Omega, \mathscr{F}, P)$. *Set* $\mathscr{F}_n^- = \sigma(X_i, 1 \le i \le n)$, $\mathscr{F}_n^+ = \sigma(X_i, i \ge n)$, *define*

$$\rho(n) = \sup_{k\ge 1} \sup_{X \in L^2(\mathscr{F}_k^-)} \sup_{Y \in L^2(\mathscr{F}_{k+n}^+)} \frac{|EXY - EXEY|}{\sqrt{E(X-EX)^2 E(Y-EY)^2}}. \tag{1.2}$$

The sequence $X_1, X_2, ...$ *is said to be* $\rho-$*mixing if* $\rho(n) \to 0$ *as* $n \to \infty$.

This definition was introduced by Kolmogorov and Rozanov [12]. As is known, the asymptotic behavior of $\rho-$mixing sequences have received much well-deserved attention, and a variety of elegant results have been obtained. See, for example, Lin and Lu [15], Peligrad [19-21], Peligrad and Shao [23], Peligrad and Utev [24], Shao [29-33], and so forth.

Let $X_1, X_2, ...$ be a sequence of stationary $\rho-$mixing random variables with density f. replace the independent observations by the $\rho-$mixing ones in (1.1), one gets the corresponding density estimator of f for the dependent random variables.

In this article, we devote ourselves to doing three things. The first one is to study convergence in distribution of the estimator $f_{n,K}(x)$ both as an estimation for the true density function $f(x)$ and as an estimation $F_{n,K}(x) = \int_{-\infty}^{x} f_{n,K}(t)dt$ for the true distribution function $F(x)$ of X. The second is to investigate the convergence rates for the difference of $f_{n,K}(x)$ and its mean in sup-norm loss and integral L^p-norm loss. Our third goal is to discuss the strong uniform convergence rates of $|f_{n,K}(x) - Ef_{n,K}(x)|$ over a compact set of $\mathbb{R}$ and the whole real line $\mathbb{R}$, respectively. Of course, a natural question is posed as follows: Whether the optimal convergence rates could be achieved? The answer is affirmative for i.i.d. observations. As is known a variety of sharp results have been established, see, for example, Einmahl and Mason [4, 5], Giné and Guillou [8, 9], Giné, Koltchinskii and Zinn [10]. However, that in general is not the case for dependent samples. To obtain the best possible convergence rates, some different methods from those for i.i.d. case should be developed. The present paper tries to do this. Our technical proofs consist in applications of the blocking techniques, the martingale methods and some inequalities. It is showed that the optimal convergence rates for i.i.d. random variables are also optimal for dependent ones.

The remainder of the paper is structured as follows. Section 2 introduces some notation and assumptions. Section 3 formulates several results on the weak convergence. Section 4 constructs the rates of $\|f_{n,K}(x) - Ef_{n,K}(x)\|_p$ in the sup-norm loss and integral L^p-norm loss, while Section 5 derives the rates of strong uniform consistency for KDEs. Some useful results are stated in the Appendix.

2. Notation and assumptions

In this section, we present some basic notation and assumptions which will be used in the sequel. Let $X, X_1, X_2, ...$ be a sequence of non-degenerated and stationary $\rho-$mixing random variables. Denote $K_i(x) = K\big((X_i - x)/h_n\big)$ for fixed $n \in \mathbb{N}$, where K is a measurable function satisfying some regularity conditions. $f(x)$ is the unknown density function of X with respect to Lebesgue measure. For Borel measurable function g and Borel measure μ, let $L^p := L^p(\mu)$ be the usual Lebesgue spaces of real-valued functions normed by $\|\cdot\|_p$. As usual, write $\|g\|_p = \big(\int_{\mathbb{R}} |g(x)|^p d\mu(x)\big)^{1/p}$ for $1 \le p < \infty$. Define for any nonnegative integer s the spaces $\mathcal{C}^s(\mathbb{R})$ of all bounded continuous real-valued functions that are $s-$times continuously differentiable on $\mathbb{R}$. $\mathrm{I}(\cdot)$ is the indicator function. $[z]$ denotes the integer part of z, $\log x = \log(x \vee e)$. $a_n = O(b_n)$ means $\limsup_{n\to\infty} a_n/b_n < \infty$, $a_n = o(b_n)$ stands for $\limsup_{n\to\infty} a_n/b_n = 0$, and $a_n \asymp b_n$ means $0 < \liminf_{n\to\infty} a_n/b_n \le \limsup_{n\to\infty} a_n/b_n < \infty$. The letter C with subscripts denotes some finite and positive universal constants, it may take different values in each appearance.

Some assumptions are formulated below:
(B1) $h_n \searrow 0$ and $nh_n \to \infty$ as $n \to \infty$.
(B2) $h_n \asymp n^{-\delta} l(n)$ for $0 < \delta \leq 1$, $l(n)$ is a slowly varying function.
(C1) the density function $f(x)$ of X is uniformly bounded on $\mathbb{R}$.
(C2) the density function $f(x)$ of X is uniformly continuous and uniformly bounded on $\mathbb{R}$.
(K1) K is a real-valued measurable function satisfying $\sup_{x\in\mathbb{R}} |K(x)| < \infty$ and $\int_{-\infty}^{\infty} |K(x)|dx < \infty$.
(K2) K is a real-valued measurable function with compact support on $\mathbb{R}$, and satisfies Lipschitz condition.

Remark 1. Condition (B2) is a little more stronger than (B1). in other words, (B2) does not allow the bandwidths h_n to go to zero very slowly as $n \to \infty$. For example, the form of the bandwidths such as $h_n = 1/(\log n)^p$, for all $p > 0$, is excluded. But we would like to point out that most of the bandwidths including the optimal ones are contained in (B2).

3. Central limit theorems for KDEs and their distribution functions

Consider the KDE $f_{n,K}(x)$ defined in (1.1). The aim of this section is to investigate the CLT for $f_{n,K}(x)$ and $F_{n,K}(x)$. The classical theory of this subject was developed mostly in the 1950s, and it is an important theory in probability and statistics. Our first result reads as follows:

Theorem 3.1. *Suppose that conditions (B1), (C2) and (K1) hold. Further assume that $f(x) > 0$ and $\sum_{i=0}^{\infty} \rho(2^i) < \infty$. Then we have*

$$\sqrt{nh_n}(f_{n,K}(x) - Ef_{n,K}(x)) \xrightarrow{d} N(0, \|K\|_2^2 f(x)), \tag{3.1}$$

where " $\xrightarrow{d}$ " stands for convergence in distribution.
Proof. For any fixed $x \in \mathbb{R}$, we use the following decomposition:

$$f_{n,K}(x) - Ef_{n,K}(x) = \big[f_{n,K}(x) - \Gamma_{n,K}(x)\big] + \big[\Gamma_{n,K}(x) - Ef_{n,K}(x)\big], \tag{3.2}$$

where

$$\Gamma_{n,K}(x) := \frac{1}{nh_n}\sum_{i=1}^{n} E\big[K((X_i - x)/h_n)|\mathscr{F}_{i-1}\big], \quad \mathscr{F}_i := \sigma(X_j, j \leq i), \ \mathscr{F}_0 = \{\emptyset, \Omega\}.$$

Thus, (3.1) will be derived if one can show that

$$\Gamma_{n,K}(x) - Ef_{n,K}(x) = o_P\left(\frac{1}{\sqrt{nh_n}}\right) \tag{3.3}$$

and

$$\sqrt{nh_n}\big(f_{n,K}(x) - \Gamma_{n,K}(x)\big) \xrightarrow{d} N(0, \|K\|_2^2 f(x)). \tag{3.4}$$

We first prove (3.3). However, some preliminary work is needed. Denote $\mathbb{N}^+ = \{1, 2, ...\}$, and let $\mathbb{I}_k$ be the integer interval $[2^k, 2^{k+1})$. Clearly, for each $n \in \mathbb{N}^+$, there exists integer $k_n \geq 0$ such that $2^{k_n} \leq n < 2^{k_n+1}$. Moreover, for $0 < \beta < \alpha < 1$, let $p_k = [2^{\alpha k}], q_k = [2^{\beta k}], r_k = [2^k/(p_k + q_k)]$, then the integer set $\mathbb{I}_k$ can be blocked as follows:

$$\begin{aligned}
&\mathbb{I}_k(m) = [2^k + (m-1)(p_k + q_k), 2^k + (m-1)q_k + mp_k) \cap \mathbb{N}^+,\\
&\mathbb{J}_k(m) = [2^k + (m-1)q_k + mp_k, 2^k + m(p_k + q_k)) \cap \mathbb{N}^+,\\
&1 \leq m \leq r_k, \quad \mathbb{J}_k(r_k + 1) = [2^k + r_k(p_k + q_k), 2^{k+1}) \cap \mathbb{N}^+.
\end{aligned}$$

It is easy to see that $r_k \sim 2^{(1-\alpha)k}$. According to the symbols as above, there also exists some integer $m_n \geq 0$ such that $n \in \mathbb{I}_{k_n}(m_n) \cup \mathbb{J}_{k_n}(m_n)$. For simplicity, we introduce some extra notation as follows: Denote

$$W_k^m(x) := \sum_{i\in\mathbb{I}_k(m)} \Big[E(K_i(x)|\mathscr{F}_{i-1}) - EK_i(x)\Big], \quad V_k^m(x) := \sum_{i\in\mathbb{J}_k(m)} \Big[E(K_i(x)|\mathscr{F}_{i-1}) - EK_i(x)\Big].$$

Then it follows that

$$nh_n\big[\Gamma_{n,K}(x) - Ef_{n,K}(x)\big] = \left[\sum_{k=0}^{k_n-1}\sum_{m=1}^{r_k} W_k^m(x) + \sum_{m=1}^{m_n-1} W_{k_n}^m(x)\right]$$
$$+ \left[\sum_{k=0}^{k_n-1}\sum_{m=1}^{r_k+1} V_k^m(x) + \sum_{m=1}^{m_n-1} V_{k_n}^m(x)\right] + \sum_{i=N_n}^{n} \Big[E(K_i(x)|\mathscr{F}_{i-1}) - EK_i(x)\Big],$$

where $N_n = 2^{k_n} + (m_n - 1)(p_{k_n} + q_{k_n})$.

Thus, in order to verify (3.3), it suffices to show that the sums on the big blocks satisfy

$$E\left[\sum_{k=0}^{k_n-1}\sum_{m=1}^{r_k} W_k^m(x) + \sum_{m=1}^{m_n-1} W_{k_n}^m(x)\right]^2 = o\big(nh_n\big). \tag{3.5}$$

Note that the left-hand side of (3.5) is controlled by

$$2E\left[\sum_{k=0}^{k_n-1}\sum_{m=1}^{r_k} W_k^m(x)\right]^2 + 2E\left[\sum_{m=1}^{m_n-1} W_{k_n}^m(x)\right]^2 =: \Sigma_1 + \Sigma_2. \tag{3.6}$$

So we only need to show that

$$\Sigma_1 = o\big(nh_n\big), \qquad \Sigma_2 = o\big(nh_n\big). \tag{3.7}$$

Using the towering property and Jensen's inequality for the conditional expectations together with Lemma A.4 with $p = 2$, we have

$$\begin{aligned}
\Sigma_1 &= 2E\left[\sum_{k=0}^{k_n-1} E\left(\sum_{m=1}^{r_k} W_k^m(x)\Big|\mathscr{F}_{2^k+r_k-1}\right)\right]^2 \\
&\leq C\log^2(2k_n)\sum_{k=0}^{k_n-1}\rho^2(q_k)E\left[\sum_{m=1}^{r_k} W_k^m(x)\right]^2 \\
&= C\log^2(2k_n)\sum_{k=0}^{k_n-1}\rho^2(q_k)E\left[\sum_{m=1}^{r_k} E\Big(W_k^m(x)\big|\mathscr{F}_{t_k(m)}\Big)\right]^2 \\
&\leq C\log^2(2k_n)\sum_{k=0}^{k_n-1}\rho^4(q_k)\log^2(2r_k)\sum_{m=1}^{r_k} E\Big(W_k^m(x)\Big)^2 \\
&\leq C\log^2(2k_n)\sum_{k=0}^{k_n-1} 2^k\rho^4(q_k)\log^2(2r_k)\log^2(2p_k)\|K_1(x)\|_2^2,
\end{aligned}$$

where $t_k(m) = 2^k + (-q_{k-1})\mathrm{I}(m=1) + (m-2)(p_k+q_k)\mathrm{I}(m \neq 1) + p_k, 1 \leq m \leq r_k$.

Recalling the condition imposed on the mixing rates, without loss of generality (w.l.o.g.), suppose that $\rho(n) \leq 1/\log n$, and observe that

$$EK_1^2(x) = h_n\int_{-\infty}^{\infty} K^2(u)f(x+h_nu)du \leq h_n\|f\|_\infty\|K\|_\infty\|K\|_1. \tag{3.8}$$

Then, applying Lemma A.2, we can get

$$\Sigma_1 \le C\log^2(2k_n)EK_1^2(x)\sum_{k=0}^{k_n-1}2^k k^{-2} \le Cnh_n(\log\log n)^2(\log n)^{-1} = o(nh_n). \tag{3.9}$$

Similarly, we have

$$\Sigma_2 \le Cnh_n(\log\log n)^2(\log n)^{-2} = o(nh_n). \tag{3.10}$$

Combining (3.9) and (3.10) yields (3.3).

As to (3.4), note that

$$\sqrt{nh_n}\big(f_{n,K}(x)-\Gamma_{n,K}(x)\big) = \frac{1}{\sqrt{nh_n}}\sum_{i=1}^n\Big[K_i(x)-E\big(K_i(x)|\mathscr{F}_{i-1}\big)\Big]. \tag{3.11}$$

We next estimate the conditional variance

$$\begin{aligned}&\frac{1}{nh_n}\sum_{i=1}^n E\Big\{\big[K_i(x)-E\big(K_i(x)|\mathscr{F}_{i-1}\big)\big]^2|\mathscr{F}_{i-1}\Big\}\\ &= \frac{1}{nh_n}\sum_{i=1}^n E\big[K_i^2(x)|\mathscr{F}_{i-1}\big] - \frac{1}{nh_n}\sum_{i=1}^n\big[E\big(K_i(x)|\mathscr{F}_{i-1}\big)\big]^2 =: \Xi_1-\Xi_2.\end{aligned}$$

For Ξ_1, observe that

$$P\big(|\Xi_1-\|K\|_2^2f(x)|>\epsilon\big) = P\big(|\Xi_1-h_n^{-1}EK_1^2(x)+h_n^{-1}EK_1^2(x)-\|K\|_2^2f(x)|>\epsilon\big). \tag{3.12}$$

Clearly, on account of condition (C2) and Bonchner's lemma, we have

$$h_n^{-1}EK_1^2(x) = \int_{-\infty}^{\infty}K^2(u)f(x+h_nu)du \to \|K\|_2^2f(x). \tag{3.13}$$

Therefore, by Lemma A.2 and Jensen's inequality together with condition (B1), the right-hand side of (3.12), for large n, is controlled by

$$P\big(|\Xi_1-h_n^{-1}EK_1^2(x)|>\epsilon/2\big) \le C\epsilon^{-2}n^{-1}h_n^{-2}EK_1^4(x) \le C\epsilon^{-2}(nh_n)^{-1}\|f\|_\infty \to 0. \tag{3.14}$$

For Ξ_2, we have with probability one,

$$h_n^{-1}E\big(K_i(x)|\mathscr{F}_{i-1}\big) = \int_{-\infty}^{\infty}K(u)f(x+h_nu|\mathscr{F}_{i-1})du \le \|K\|_\infty\|f\|_\infty.$$

Then, it follows that

$$E\Xi_2 \le h_n\|K\|_\infty\|f\|_\infty \to 0. \tag{3.15}$$

Combining (3.12)−(3.15) yields

$$\frac{1}{nh_n}\sum_{i=1}^n E\Big\{\big[K_i(x)-E\big(K_i(x)|\mathscr{F}_{i-1}\big)\big]^2|\mathscr{F}_{i-1}\Big\} \xrightarrow{P} \|K\|_2^2f(x). \tag{3.16}$$

Moreover, applying the C_r-inequality, we have

$$E\big(K_i(x)-E\big(K_i(x)|\mathscr{F}_{i-1}\big)\big)^2 \le 4EK_i^2(x) \le 4h_n\|f\|_\infty. \tag{3.17}$$

Thus on account of (B1), the Lindeberg condition

$$\begin{aligned}&\frac{1}{nh_n}\sum_{i=1}^n E\Big[K_i(x)-E\big(K_i(x)|\mathscr{F}_{i-1}\big)\Big]^2\mathrm{I}(|K_i(x)-E(K_i(x)|\mathscr{F}_{i-1})|>\epsilon\sqrt{nh_n})\\ &\le 4h_n^{-1}EK_1^2(x)\mathrm{I}(|K_1(x)|>\epsilon\sqrt{nh_n}/2) \le 4\|f\|_\infty\int_{|K(u)|>\epsilon\sqrt{nh_n}/2}K(u)du = o(1)\end{aligned} \tag{3.18}$$

holds for any $\epsilon > 0$.

Finally, according to (3.16) and (3.18), then using the martingale central limit theorem together with Slutsky's theorem gives (3.1).

Remark 2. Let us consider the deviation of the kernel density estimator with respect to the true density function. Note that

$$f_{n,K}(x) - f(x) = \big[f_{n,K}(x) - Ef_{n,K}\big] + \big[Ef_{n,K}(x) - f(x)\big]. \tag{3.19}$$

The first term on the right-hand side of (3.19) is the probabilistic term, while the second term is the bias. If (C2) and the conditions imposed on the kernel K in Theorem 3.1 are replaced by

(C3) the density function $f(x)$ is uniformly bounded, $f(x) \in \mathcal{C}(\mathbb{R})$ and $\sup_x |f^{'}(x)| < \infty$,
(K3) K satisfies $\sup_x |K(x)| < \infty$, $\int_{-\infty}^{\infty} |xK(x)|dx < \infty$,
then applying Taylor's expansion, we have for some $0 < \upsilon < 1$,

$$\begin{aligned} \big|Ef_{n,K}(x) - f(x)\big| &= \bigg|\int_{-\infty}^{\infty} K(y)\big[f(x - h_n y) - f(x)\big]dy\bigg| \\ &= h_n\bigg|\int_{-\infty}^{\infty} yK(y)f^{'}(x - \upsilon h_n y)dy\bigg| = O(h_n). \end{aligned}$$

Therefore, (3.1) holds whenever $nh_n^3 \to 0$ as $n \to \infty$.

In fact, the bias can always be balanced with the probabilistic term by calibrating the normalizing sequence $\{h_n, n \geq 1\}$, provided enough regularity for K and f are assumed.

Another interesting problem is the weak convergence for the distribution function of the KDE. More precisely, denote $F_{n,K}(x) = \int_{-\infty}^{x} f_{n,K}(t)dt$, we construct the CLT for the difference between $F_{n,K}(x)$ and its mean.

Theorem 3.2. *Suppose that condition (B1) holds, and that $\sum_{i=0}^{\infty} \rho(2^i) < \infty$. Further, assume that the density function $f(x)$ is continuous and positive on $\mathbb{R}$, and that the kernel K is symmetric and $\int_{-\infty}^{\infty} K(x)dx = 1$. Then we have*

$$\sqrt{n}(F_{n,K}(x) - EF_{n,K}(x)) \xrightarrow{d} N(0, F(x)(1 - F(x))), \tag{3.20}$$

where $F(x)$ is the true distribution function of X.
Proof. Observe that

$$F_{n,K}(x) - EF_{n,K}(x) = [F_{n,K}(x) - \Lambda_{n,K}(x)] + [\Lambda_{n,K}(x) - EF_{n,K}(x)], \tag{3.21}$$

where

$$\Lambda_{n,K}(x) := \frac{1}{nh_n}\sum_{i=1}^{n} E\bigg[\int_{-\infty}^{x} K_i(t)dt\bigg|\mathscr{F}_{i-1}\bigg].$$

Along the similar proof lines as those of (3.3), one can get

$$\Lambda_{n,K}(x) - EF_{n,K}(x) = o_P\bigg(\frac{1}{\sqrt{n}}\bigg). \tag{3.22}$$

We next show that

$$F_{n,K}(x) - \Lambda_{n,K}(x) \xrightarrow{d} N(0, F(x)(1 - F(x))). \tag{3.23}$$

Note that $F_{n,K}(x) - \Lambda_{n,K}(x)$ is a martingale with respect to an increasing σ-algebra $\mathscr{F}_n = \sigma(X_1, ..., X_n)$. So in order to verify (3.23), we only need to check the conditions on the CLT for martingales. For simplicity, set

$$E_{n,2}^K(x) := h_n^{-2} E\Big(\int_{-\infty}^x K_1(t)dt\Big)^2. \tag{3.24}$$

We claim that the limit of $E_{n,2}^K(x)$ exists for any fixed $x \in \mathbb{R}$ as $n \to \infty$. The proof is as follows: Denote

$$G_K(x) := \int_{-\infty}^x K(u)du.$$

Recalling that $\int_{-\infty}^{\infty} K(x)dx = 1$. Obviously, $G_K(x)$ is the distribution function of a finite measure. Then by the symmetry of kernel K, we have

$$h_n^{-1}\int_{-\infty}^x K_1(t)dt = h_n^{-1}\int_{-\infty}^x K\Big(\frac{X_1-t}{h_n}\Big)dt = \int_{-\infty}^{\frac{x-X_1}{h_n}} K(u)du = G_K\Big(\frac{x-X_1}{h_n}\Big). \tag{3.25}$$

Note that G_K is bounded from above by one almost surely, it then turns out that

$$E_{n,2}^K(x) = E\Big[G_K\Big(\frac{x-X_1}{h_n}\Big)\Big]^2 \le 1. \tag{3.26}$$

Thus we have for any fixed $x \in \mathbb{R}$,

$$\begin{aligned} E\Big[G_K\Big(\frac{x-X_1}{h_n}\Big)\Big]^2 &= \int_{-\infty}^{\infty}\Big[G_K\Big(\frac{x-u}{h_n}\Big)\Big]^2 f(u)du \\ &= \Big(\int_{-\infty}^x + \int_x^{\infty}\Big)\Big\{\Big[G_K\Big(\frac{x-u}{h_n}\Big)\Big]^2 f(u)du\Big\} \to F(x). \end{aligned} \tag{3.27}$$

The conditional variance

$$\begin{aligned} &\frac{1}{nh_n^2}\sum_{i=1}^n E\Big\{\Big[\int_{-\infty}^x K_i(t)dt - E\Big(\int_{-\infty}^x K_i(t)dt\Big|\mathscr{F}_{i-1}\Big)\Big]^2\Big|\mathscr{F}_{i-1}\Big\} \\ &= \frac{1}{nh_n^2}\sum_{i=1}^n E\Big[\Big(\int_{-\infty}^x K_i(t)dt\Big)^2\Big|\mathscr{F}_{i-1}\Big] - \frac{1}{nh_n^2}\sum_{i=1}^n\Big[E\Big(\int_{-\infty}^x K_i(t)dt\Big|\mathscr{F}_{i-1}\Big)\Big]^2 =: \Xi^{'} - \Xi^{''}. \end{aligned}$$

For $\Xi^{'}$, by Lemma A.2 and Jensen's inequality, it follows for any $x \in \mathbb{R}$,

$$\begin{aligned} &P\big(|\Xi^{'} - F(x)| > \epsilon\big) \\ &= P\big(|\Xi^{'} - E_{n,2}^K(x) + E_{n,2}^K(x) - F(x)| > \epsilon\big) \\ &\le P\big(|\Xi^{'} - E_{n,2}^K(x)| > \epsilon/2) \le C\epsilon^{-2}n^{-2}EG_K^4((x-X_1)/h_n\big) \\ &\le C\epsilon^{-2}n^{-1} \to 0, \quad n \to \infty. \end{aligned} \tag{3.28}$$

As for $\Xi^{''}$. First, similarly to that of (3.27), we have

$$E_{n,1}^K := E\Big[G_K\Big(\frac{x-X_1}{h_n}\Big)\Big] \to F(x), \quad n \to \infty. \tag{3.29}$$

Moreover, we have for large n,

$$\begin{aligned} &P\big(|\Xi^{''} - F^2(x)| > \epsilon\big) \\ &= P\big(|\Xi^{''} - (E_{n,1}^K(x))^2 + (E_{n,1}^K(x))^2 - F^2(x)| > \epsilon\big) \\ &\le P\big(|\Xi^{''} - (E_{n,2}^K(x))^2| > \epsilon/2\big). \end{aligned} \tag{3.30}$$

Further, note that

$$\Xi'' - (E_{n,2}^K(x))^2 = \frac{1}{n}\sum_{i=1}^{n}\left[G_K\left(\frac{x-X_1}{h_n}\right) - E_{n,1}^K(x)\right]\left[G_K\left(\frac{x-X_1}{h_n}\right) + E_{n,1}^K(x)\right]. \tag{3.31}$$

By the a.s. boundness of G_K, (3.29) and Lemma A.2, the right-hand side of (3.30) is less than or equal to

$$P\left(\frac{1}{n}\sum_{i=1}^{n}\left[G_K\left(\frac{x-X_1}{h_n}\right) - E_{n,1}^K(x)\right] > \epsilon/4\right) \to 0, \quad n \to \infty. \tag{3.32}$$

Then it turns out that

$$\sum_{i=1}^{n} E\Big\{\big[K_i(x) - E\big(K_i(x)|\mathscr{F}_{i-1}\big)\big]^2|\mathscr{F}_{i-1}\Big\} \xrightarrow{P} F(x)(1-F(x)). \tag{3.33}$$

Similarly to that of (3.18), one can show that the Lindeberg condition holds. Finally, by the CLT for martingales and Slutsky's theorem, we obtains (3.20).

Before stating the next result, we introduce the following condition:

(B3) $h_n \searrow 0$, $nh_n \to \infty$ and $\sqrt{n}\omega(n)h_n \to 0$, where $\omega(n)$ is a nonnegative real function satisfying $\omega(n) \nearrow \infty$ as $n \to \infty$.

Theorem 3.3. *Suppose that condition (B3) holds, and that $\sum_{i=0}^{\infty}\rho(2^i) < \infty$. Further, assume that the density function $f(x)$ is positive and Lipschitz continuous on $\mathbb{R}$, and that the kernel K is a symmetric function with bounded support, $\int_{-\infty}^{\infty} K(x)dx = 1$ and $\int_{-\infty}^{\infty}|K(x)|dx < \infty$. Then we have*

$$\sqrt{n}(F_{n,K}(x) - F(x)) \xrightarrow{d} N(0, F(x)(1-F(x))). \tag{3.34}$$

Proof. Observe that

$$F_{n,K}(x) - F(x) = [F_{n,K}(x) - \Lambda_{n,K}(x)] + [\Lambda_{n,K}(x) - EF_{n,K}(x)] + [EF_{n,K}(x) - F(x)], \tag{3.35}$$

where $\Lambda_{n,K}(x)$ is defined in the proof of Theorem 3.2. In fact, according to Theorem 3.2, one only needs to show that

$$EF_{n,K}(x) - F(x) = o\left(\frac{1}{\sqrt{n}}\right). \tag{3.36}$$

Observe that $f(x)$ is integrable on $\mathbb{R}$. Then for any given $\epsilon > 0$, it can be decomposed as follows:

$$f(x) = f_1(x) + f_2(x), \tag{3.37}$$

where $f_1(x)$ is continuous on a compact interval $[c_1, c_2]$ (say), and $f_2(x)$ satisfies

$$\int_{-\infty}^{\infty}|f_2(x)|dx < \epsilon. \tag{3.38}$$

Denote $\tau_n = 1/(\sqrt{n}\omega(n))$. Recalling that $\int_{-\infty}^{\infty}K(x)dx = 1$, we then have for any fixed $x \in \mathbb{R}$,

$$\begin{aligned}|EF_{n,K}(x) - F(x)| &= \left|\int_{-\infty}^{x}\left[\int_{-\infty}^{\infty}K(u)\big[f(t+uh_n) - f(t)\big]du\right]dt\right| \\ &\leq \int_{-\infty}^{x}\int_{|u|>\tau_n h_n^{-1}}|K(u)||f_2(t+uh_n) - f_2(t)|dudt \\ &\quad + \int_{-\infty}^{x}\int_{|u|\leq\tau_n h_n^{-1}}|K(u)||f_2(t+uh_n) - f_2(t)|dudt =: \Theta_1 + \Theta_2.\end{aligned}$$

Recall that the kernel K is supported on the bounded interval. Clearly, Θ_1 is controlled by

$$\begin{aligned}
&\int_{|u|>\tau_n h_n^{-1}} |K(u)| \Big[\int_{-\infty}^{\infty} |f_2(t+uh_n) - f_2(t)| dt \Big] du \\
&\quad \le 2 \int_{|u|>\tau_n h_n^{-1}} |K(u)| \Big[\int_{-\infty}^{\infty} |f_2(t)| dt \Big] du \to 0, \quad n \to \infty. \qquad (3.39)
\end{aligned}$$

For Θ_2, note that f_1 is Lipschitz continuous on $[c_1, c_2]$. Subsequently,

$$\begin{aligned}
\Theta_2 &\le \int_{|u|\le\tau_n h_n^{-1}} |K(u)| \Big[\int_{c_1-\tau_n}^{(c_2+\tau_n)\wedge x} \sup_{|s|\le\tau_n} |f_1(t+s) - f_1(t)| dt \Big] du \\
&\le C(c_2 - c_1)\tau_n + o\Big(\frac{1}{\sqrt{n}}\Big) = o\Big(\frac{1}{\sqrt{n}}\Big). \qquad (3.40)
\end{aligned}$$

Therefore, we complete the proof of (3.36).

4. Convergence rates of $\|f_{n,K}(x) - Ef_{n,K}(x)\|_p$ in sup-norm and integral L^p-norm

One may be interested in the consistency of $\|f_{n,K}(x) - Ef_{n,K}(x)\|_p$, which are investigated in this section. Among which, the uniform convergence rate with respect to L^p-norm distance is established in Theorem 4.1, while the convergence rate for integral L^p-norm is given in Theorem 4.2.

Theorem 4.1. *Let $p \ge 2$. Suppose that $\sum_{i=0}^{\infty} \rho^{2/p}(2^i) < \infty$, and that the conditions (B1), (C1), (K1) are satisfied. Then we have*

$$\sup_{x\in\mathbb{R}} \big\| f_{n,K}(x) - Ef_{n,K}(x) \big\|_p = O\Big[\Big(\frac{1}{nh_n} \Big)^{1/2} \Big]. \qquad (4.1)$$

Proof. We will use the symbols such as $p_k, q_k, r_k, \mathbb{I}_k(m), \mathbb{J}_k(m)$ etc. appeared in the proof of Theorem 3.1. However, the values of α and β are different from those as in the proof of Theorem 3.1. Here we select some $(p-2)/(p-1) < \alpha < 1$ and $0 < \beta < \min(\alpha, 1 + p\alpha - p)$. In fact, β allows taking the value $1 + p\alpha - p$. For simplicity, define

$$Y_k^m(x) = \sum_{i\in\mathbb{I}_k(m)} K_i(x),\ Z_k^m(x) = \sum_{i\in\mathbb{J}_k(m)} K_i(x),\ 1 \le m \le r_k;\ Z_k^{r_k+1}(x) = \sum_{i\in\mathbb{J}_k(r_k+1)} K_i(x),$$

$\xi_k^m(x) = Y_k^m(x) - E[Y_k^m(x)|\mathscr{F}_k(m-1)]$, where

$$\mathscr{F}_k(m) = \sigma\big(X_r, r \le 2^k - q_{k-1}\mathrm{I}(m=1) + [(m-1)(p_k+q_k) - q_k]\mathrm{I}(m \ne 1)\big).$$

Then we have for any fixed $x \in \mathbb{R}$,

$$\begin{aligned}
f_{n,K}(x) - Ef_{n,K}(x) = \frac{1}{nh_n} \Bigg\{ & \Bigg[\sum_{k=0}^{k_n-1} \sum_{m=1}^{r_k} \xi_k^m(x) + \sum_{m=1}^{m_n-1} \xi_{k_n}^m(x) \Bigg] \\
& + \Bigg[\sum_{k=0}^{k_n-1} \sum_{m=1}^{r_k+1} (Z_k^m(x) - EZ_k^m(x)) + \sum_{m=1}^{m_n-1} (Z_{k_n}^m(x) - EZ_{k_n}^m(x)) \Bigg] \\
& + \Bigg[\sum_{k=0}^{k_n-1} \sum_{m=1}^{r_k} U_k^m(x) + \sum_{m=1}^{m_n-1} U_{k_n}^m(x) \Bigg] + \sum_{i=N_n}^{n} (K_i(x) - EK_i(x)) \Bigg\} \\
=: \frac{1}{nh_n} (\mathrm{I}_1(x) + \mathrm{I}_2(x) + & \mathrm{I}_3(x) + \mathrm{I}_4(x)),
\end{aligned}$$

where $U_k^m(x) = E(Y_k^m(x)|\mathscr{F}_k(m-1)) - EY_k^m(x)$.

Thus in order to prove (4.1), it is enough to show that

$$\sup_{x\in\mathbb{R}} \|\mathrm{I}_2(x) + \mathrm{I}_3(x) + \mathrm{I}_4)(x)\|_p = o\Big[(nh_n)^{1/2}\Big] \tag{4.2}$$

and

$$\sup_{x\in\mathbb{R}} \|\mathrm{I}_1(x)\|_p = O\Big[(nh_n)^{1/2}\Big]. \tag{4.3}$$

The proof of (4.2) will be divided into two steps:

Step 1. We first show for any $x \in \mathbb{R}$,

$$\left\| \sum_{m=1}^{r_k+1} (Z_k^m(x) - EZ_k^m(x)) \right\|_p = o\Big[(2^k h_{2^k})^{1/2}\Big]. \tag{4.4}$$

Using Lemma A.4, and note that $\sum_{i=0}^{\infty} \rho^{2/p}(2^i) < \infty$, it follows that

$$E|Z_k^m(x) - EZ_k^m(x)|^p \le Cq_k^{p/2} \|K_1(x)\|_2^p + Cq_k \|K_1(x)\|_p^p. \tag{4.5}$$

Recalling conditions (C1) and (K1), then by a simple calculation, we have

$$\|K_1(x)\|_2^p \le (\|f\|_\infty \|K\|_\infty \|K\|_1 h_{2^k})^{p/2} \tag{4.6}$$

and

$$\|K_1(x)\|_p^p \le \|f\|_\infty \|K\|_\infty^{p-1} \|K\|_1 h_{2^k}. \tag{4.7}$$

Applying (4.5)−(4.7) and Minkowski's inequality yields

$$\begin{aligned}
&\left\| \sum_{k=0}^{k_n-1} \sum_{m=1}^{r_k+1} (Z_k^m(x) - EZ_k^m(x)) \right\|_p \\
&\le C \sum_{k=0}^{k_n-1} \Big(r_k h_{2^k}^{1/2} q_k^{1/2} + r_k h_{2^k}^{1/p} q_k^{1/p} \Big) \\
&= O\Big[(nh_n)^{1/2} \Big(n^{(1+\beta-2\alpha)/2} + o(n^{1-\alpha-(1-\beta)/p}) \Big)\Big] = o\Big[(nh_n)^{1/2}\Big],
\end{aligned}$$

where the first equality is obtained by $1/h_n = o(n)$, and the second one is due to $\beta \le 1 + p\alpha - p$. Note that $h_n \searrow 0$ and $nh_n \to \infty$ implies that $h_n/h_{n+1} \le 2$ for $n \ge 1$. Similarly, we have

$$\sup_{x\in\mathbb{R}} \left\| \sum_{m=1}^{m_n-1} (Z_{k_n}^m(x) - EZ_{k_n}^m(x)) \right\|_p = o\Big[(nh_n)^{1/2}\Big]. \tag{4.8}$$

Step 2. We next prove for any $x \in \mathbb{R}$,

$$\left\| \sum_{m=1}^{r_k} U_k^m(x) \right\|_p = o\Big[(2^k h_{2^k})^{1/2}\Big]. \tag{4.9}$$

Recalling that $\sum_{i=0}^{\infty} \rho^{2/p}(2^i) < \infty$, hence w.l.o.g., we suppose that $\rho(n) \le (\log n)^{-p/2}$. By Lemma A.4, we have

$$\begin{aligned}
E|U_k^m(x)|^p &\le L(\log 2p_k)^p \Bigg[\sum_{i=1}^{p_k} \rho^{2/(p-1)}(q(i/2)) \Bigg] \|K_1(x)\|_p^p \\
&\quad + L(\log 2p_k)^p \Bigg[\sum_{i=1}^{p_k} \rho^2(q(i/2)) \Bigg]^{p/2} \|K_1(x)\|_2^p \\
&\le L\Big[k^p 2^{-\alpha k/(p-1)} h_{2^k} + k^p 2^{\alpha(1-p)p/2} h_{2^k}^{p/2}\Big],
\end{aligned}$$

where $q(x)$ is the linear interpolating function of q_k. Subsequently,

$$\|U_k^m(x)\|_p \le L\Big[k2^{-\alpha k/p(p-1)}h_{2^k}^{1/p} + k2^{\alpha(1-p)/2}h_{2^k}^{1/2}\Big]. \tag{4.10}$$

Then taking (4.10) back into (4.9), a standard computation leads to

$$\begin{aligned}
&\sup_{x\in\mathbb{R}}\Big\|\sum_{k=0}^{k_n-1}\sum_{m=1}^{r_k}U_k^m(x)\Big\|_p \\
&\le C\sum_{k=0}^{k_n-1}\Big[r_k k2^{-\alpha k/p(p-1)}h_{2^k}^{1/p} + r_k k2^{\alpha(1-p)/2}h_{2^k}^{1/2}\Big] \\
&= C\sum_{k=0}^{k_n-1}\Big[k2^{(1-\alpha-\alpha/p(p-1))k}h_{2^k}^{1/p} + k2^{(1-\alpha+\alpha(1-p)/2)k}h_{2^k}^{1/2}\Big] \\
&= O\Big[(nh_n)^{1/2}(\log n)\Big(n^{1/2-\alpha-\alpha/p(p-1)}h_n^{1/p-1/2} + n^{(1-\alpha-p\alpha)/2}\Big)\Big] \\
&= O\Big[(nh_n)^{1/2}(\log n)\Big(n^{1-\alpha-1/p} + n^{(1-\alpha-p\alpha)/2}\Big)\Big] = o\Big[(nh_n)^{1/2}\Big],
\end{aligned}$$

where the third equality is obtained by $nh_n \to \infty$, and the last equality is due to $\alpha > (p-1)/p$.

Similarly, we have

$$\sup_{x\in\mathbb{R}}\Big\|\sum_{m=1}^{m_n-1}U_{k_n}^m(x)\Big\|_p = o\Big[(nh_n)^{1/2}\Big] \tag{4.11}$$

and

$$\sup_{x\in\mathbb{R}}\Big\|\sum_{i=N_n}^{n}(K_i(x) - EK_i(x))\Big\|_p = o\Big[(nh_n)^{1/2}\Big]. \tag{4.12}$$

According to the two steps as above, we complete the proof of (4.2).

In order to prove (4.3), we first show for any integer $s \ge 1$,

$$\sup_{x\in\mathbb{R}}\|\mathrm{I}_1(x)\|_{2^s} = O\Big[(nh_n)^{1/2}\Big]. \tag{4.13}$$

Note that

$$\Big\|\sum_{k=0}^{k_n-1}\sum_{m=1}^{r_k}\xi_k^m(x)\Big\|_{2^s} \le \sum_{k=0}^{k_n-1}\Big\|\sum_{m=1}^{r_k}\xi_k^m(x)\Big\|_{2^s}. \tag{4.14}$$

Hence, we only need to show that

$$\Big\|\sum_{m=1}^{r_k}\xi_k^m(x)\Big\|_{2^s} = O\Big[(2^k h_{2^k})^{1/2}\Big]. \tag{4.15}$$

(4.15) will be derived by induction on s: If $s=1$, using the orthogonal property of the martingale sequences and Lemma A.2, we have

$$E\Big[\sum_{m=1}^{r_k}\xi_k^m(x)\Big]^2 = \sum_{m=1}^{r_k}E(\xi_k^m(x))^2 \le 4\sum_{m=1}^{r_k}E(Y_k^m(x))^2 = O\big(2^k h_{2^k}\big). \tag{4.16}$$

Suppose that (4.15) holds true for any integer less that s, we next show that it remains valid for s itself. Applying the the Marcinkiewicz−Zygmund−Burkholder inequality yields

$$\begin{aligned}
E\Big[\sum_{m=1}^{r_k}\xi_k^m(x)\Big]^{2^s} &\le CE\Big[\sum_{m=1}^{r_k}(\xi_k^m(x))^2\Big]^{2^{s-1}}\\
&= CE\Big\{\sum_{m=1}^{r_k}\Big[(\xi_k^m(x))^2 - E\big((\xi_k^m(x))^2|\mathscr{F}_k(m-1)\big) + E\big((\xi_k^m(x))^2|\mathscr{F}_k(m-1)\big)\Big]\Big\}^{2^{s-1}}\\
&\le CE\Big\{\sum_{m=1}^{r_k}\Big[(\xi_k^m(x))^2 - E\big((\xi_k^m(x))^2|\mathscr{F}_k(m-1)\big)\Big\}^{2^{s-1}}\\
&\quad + CE\Big\{\sum_{m=1}^{r_k}E\big((\xi_k^m(x))^2|\mathscr{F}_k(m-1)\big)\Big]\Big\}^{2^{s-1}} =: \mathrm{II}_1 + \mathrm{II}_2,
\end{aligned}$$

where the definition of $\mathscr{F}_k(m-1)$ can be referred to the beginning of the proof.

Note that $\xi_k^m(x))^2 - E\big((\xi_k^m(x))^2|\mathscr{F}_k(m-1), m = 1, 2, ...,$ are martingale differences. By the induction hypothesis, II_1 is of order $O\big((2^k h_{2^k})^{2^{s-2}}\big)$.

Using Lemma A.4 and Jensen's inequality, we have

$$\begin{aligned}
\mathrm{II}_2 &\le CK(\log 2r_k)^{2^{s-1}}\Big[\sum_{i=1}^{r_k}\rho^{2^{-s}}(q(i/2))\Big]\|(\xi_k^m(x))^2\|_{2^{s-1}}^{2^{s-1}}\\
&\quad + CK(\log 2r_k)^{2^{s-1}}\Big[\sum_{i=1}^{r_k}\rho^2(q(i/2))\Big]^{2^{s-2}}\|(\xi_k^m(x))^2\|_2^{2^{s-1}}\\
&\le CK(\log 2r_k)^{2^{s-1}}\Big[\sum_{i=1}^{r_k}\rho^{2^{-s}}(q(i/2))\Big]\|Y_k^m(x)\|_{2^s}^{2^s}\\
&\quad + CK(\log 2r_k)^{2^{s-1}}\Big[\sum_{i=1}^{r_k}\rho^2(q(i/2))\Big]^{2^{s-2}}\|Y_k^m(x)\|_4^{2^s} =: \mathrm{II}_{21} + \mathrm{II}_{22}.
\end{aligned}$$

Note that $\rho(n) \le (\log n)^{-2^{s-1}}$. By Lemma A.2, a standard calculation yields

$$\mathrm{II}_{21} = O\Big[k^{2^{s-1}}2^{(1-\alpha)k/2}\Big(2^{\alpha 2^{s-1}k}h_{2^k}^{2^{s-1}} + 2^{\alpha k}h_{2^k}\Big)\Big] = o\big(2^{2^{s-1}k}h_{2^k}^{2^{s-1}}\big). \tag{4.17}$$

Similarly,

$$\begin{aligned}
\mathrm{II}_{22} &= O\Big[k^{2^{s-1}}2^{(1-\alpha)(1-2^s)2^{s-2}k}\Big(2^{\alpha 2^{s-1}k}h_{2^k}^{2^{s-1}} + 2^{\alpha 2^{s-2}k}h_{2^k}^{2^{s-2}}\Big)\Big]\\
&= o\Big[\big(2^{2^{s-1}k}h_{2^k}^{2^{s-1}}\big)\times\big(2^{(\alpha-1)k} + 2^{\alpha-\alpha 2^{s-2}}\big)\Big] = o\big(2^{2^{s-1}k}h_{2^k}^{2^{s-1}}\big).
\end{aligned} \tag{4.18}$$

Combining (4.17) and (4.18) gives (4.15).

We next show that (4.15) holds true for any $p > 2$. In fact, there exists an integer $s \ge 1$ such that $p \in (2^s, 2^{s+1})$. Then it follows from Lyapunov's inequality,

$$\Big\|\sum_{m=1}^{r_k}\xi_k^m(x)\Big\|_p^p \le \Big\|\sum_{m=1}^{r_k}\xi_k^m(x)\Big\|_{2^s}^{2^{s+1}-p}\Big\|\sum_{m=1}^{r_k}\xi_k^m(x)\Big\|_{2^{s+1}}^{2p-2^{s+1}} = O\Big[\big(2^k h_{2^k}\big)^{p/2}\Big]. \tag{4.19}$$

Finally, we have

$$\sup_{x\in\mathbb{R}}\Big\|\sum_{k=0}^{k_n-1}\sum_{m=1}^{r_k}\xi_k^m(x)\Big\|_p = O\Big[(nh_n)^{1/2}\Big]. \tag{4.20}$$

Similarly, we have

$$\sup_{x\in\mathbb{R}}\Big\|\sum_{m=1}^{m_n-1}\xi_{k_n}^m(x)\Big\|_p=O\Big[(nh_n)^{1/2}\Big]. \tag{4.21}$$

According to the proof as above, we claim that (4.1) holds true.

Remark 3. If conditions (C1) and (K1) in Theorem 4.1 are replaced by (C3) and (K3), respectively, we have

$$\sup_{x\in\mathbb{R}}\Big\|f_{n,K}(x)-f(x)\Big\|_p=O\Big[\Big(\frac{1}{nh_n}\Big)^{1/2}+h_n\Big]. \tag{4.22}$$

Theorem 4.2. *Under the conditions of Theorem 4.1, we have for $p\geq 2$,*

$$\Big[\int_{-\infty}^{\infty}\|f_{n,K}(x)-Ef_{n,K}(x)\|_p^p dx\Big]^{1/p}=O\Big[\Big(\frac{1}{nh_n}\Big)^{1/2}\Big]. \tag{4.23}$$

Proof. According to the decomposition as above, we only need to prove

$$\Big[\int_{-\infty}^{\infty}\big\|\mathrm{I}_1(x)+\mathrm{I}_2(x)+\mathrm{I}_3(x)+\mathrm{I}_4(x)\big\|_p^p dx\Big]^{1/p}=O\Big[\big(nh_n\big)^{1/2}\Big]. \tag{4.24}$$

Obviously, (4.24) will be derived if one can show that

$$\sum_{j=1}^{4}\Big[\int_{-\infty}^{\infty}\|\mathrm{I}_j(x)\|_p^p dx\Big]^{1/p}=O\Big[\big(nh_n\big)^{1/2}\Big]. \tag{4.25}$$

In fact, the proof of (4.25) is contained in that of Theorem 4.1. For example, noting that

$$\Big\|\sum_{k=0}^{k_n-1}\sum_{m=1}^{r_k}\xi_k^m(x)\Big\|_p^p\leq\Big[\sum_{k=0}^{k_n-1}\Big\|\sum_{m=1}^{r_k}\xi_k^m(x)\Big\|_p\Big]^p. \tag{4.26}$$

Then we have

$$\Big[\int_{-\infty}^{\infty}\Big\|\sum_{k=0}^{k_n-1}\sum_{m=1}^{r_k}\xi_k^m(x)\Big\|_p^p dx\Big]^{1/p}=O\Big[\big(nh_n\big)^{1/2}\Big], \tag{4.27}$$

and the order of $\big\|\sum_{m=1}^{r_k}\xi_k^m(x)\big\|_p$ has been obtained in Theorem 4.1. Similarly, one can derive the order for other terms, here they are omitted.

5. Rates of strong uniform consistency for KDEs

In this section, we consider the a.s. convergence rates for KDEs, which are an active subject in probability and statistics these years. Among these results, Theorem 5.1 establishes the uniform rate over a compact set, while Theorem 5.2 gives the same rate over the whole real line. It is showed that the uniform convergence rates for mixing dependent observations are as good as those for i.i.d. ones.

Theorem 5.1. *Let D be a compact subset of $\mathbb{R}$. Suppose that $\rho(1)\leq 1/4$ and $\sum_{i=0}^{\infty}\rho(2^i)<\infty$, and that the conditions (B2), (C1), (K2) are satisfied. Then we have*

$$\sup_{x\in D}\big|f_{n,K}(x)-Ef_{n,K}(x)\big|=O_{a.s.}\Big[\Big(\frac{|\log h_n|}{nh_n}\Big)^{1/2}\Big]. \tag{5.1}$$

Proof. We first introduce some notation: Let $\mathbb{H}_k$ denote the set of all integers in the interval

$[2^k, 2^{k+1}), k \geq 0$. For $0 < b < a < 1/2$, whose values will be specified later. Define $p_k = [2^{ak}], q_k = [2^{bk}], r_k = [2^k/(p_k + q_k)]$, and blocks

$$\mathbb{I}_k(j) = [2^k + (j-1)(p_k + q_k), 2^k + (j-1)q_k + jp_k) \cap \mathbb{N}^+,$$
$$\mathbb{J}_k(j) = [2^k + (j-1)q_k + jp_k, 2^k + j(q_k + p_k)) \cap \mathbb{N}^+, \quad 1 \leq j \leq r_k,$$
$$\mathbb{J}_k(r_k + 1) = [2^k + r_k(p_k + q_k), 2^{k+1}) \cap \mathbb{N}^+.$$

Note that (5.1) will be derived if one can show that

$$\max_{2^k \leq n < 2^{k+1}} \sup_{x \in D} \left| f_{n,K}(x) - Ef_{n,K}(x) \right| = O_{a.s.}\left[\Big(\frac{|\log h_{2^k}|}{2^k h_{2^k}} \Big)^{1/2} \right]. \tag{5.2}$$

In order to prove (5.2), observe that D is a compact set, so one can choose finite open balls with centers at $x_1, ..., x_{l_k}$, and radius $d_k = (h_{2^k}^3 |\log h_{2^k}|/2^k)^{1/2}$ to cover D. Obviously, the numbers of the balls are of order $O((2^k/(h_{2^k}^3|\log h_{2^k}|))^{1/2})$. Further denote the ith ball by $B_i = B(x_i, d_k), 1 \leq i \leq l_k$, it is easy to see that $D \subset \bigcup_{i=1}^{l_k} B_i$. For simplicity, write

$$S_n(x) := \frac{1}{\sqrt{h_n}} \sum_{i=1}^{n} \big(K_i(x) - EK_i(x) \big). \tag{5.3}$$

By the Lipschitz condition on kernel K, we have for any $x \in B_m$ and some $U > 0$,

$$\max_{2^k \leq n < 2^{k+1}} \left| S_n(x) - S_n(x_m) \right| \leq U d_k \max_{2^k \leq n < 2^{k+1}} n h_n^{-3/2} \leq U \sqrt{2^{k+1} |\log h_{2^{k+1}}|}. \tag{5.4}$$

Denote $\lambda_n = \sqrt{n|\log h_n|}$, we have for $M > U$,

$$\begin{aligned} &P\Big(\max_{2^k \leq n < 2^{k+1}} \sup_{x \in D} |S_n(x)| \geq 2M\lambda_{2^{k+1}} \Big) \\ &\leq \sum_{m=1}^{l_k} \max_{1 \leq m \leq l_k} P\Big(\max_{2^k \leq n < 2^{k+1}} |S_n(x_m)| \geq M\lambda_{2^{k+1}} \Big) \qquad (5.5) \\ &\quad + \sum_{m=1}^{l_k} P\Big(\max_{2^k \leq n < 2^{k+1}} \sup_{x \in B_m} |S_n(x) - S_n(x_m)| \geq M\lambda_{2^{k+1}} \Big). \qquad (5.6) \end{aligned}$$

Clearly, (5.6) vanishes on account of (5.4), so we only need to consider (5.5). Let us first introduce some extra notation, define

$$U_k(j, x_m) = \frac{1}{\sqrt{h_{2^k}}} \sum_{i \in \mathbb{I}_k(j)} \big[K_i(x_m) - EK_i(x_m) \big], \quad V_k(j, x_m) = \frac{1}{\sqrt{h_{2^k}}} \sum_{i \in \mathbb{J}_k(j)} \big[K_i(x_m) - EK_i(x_m) \big],$$

$$1 \leq j \leq r_k; \quad V_k(r_k + 1, x_m) = \frac{1}{\sqrt{h_{2^k}}} \sum_{i \in \mathbb{J}_k(r_k+1)} \big[K_i(x_m) - EK_i(x_m) \big],$$

$$\zeta_k(j, x_m) = U_k(j, x_m) - E\big[U_k(j, x_m) | \mathscr{F}_k(j-1) \big], \quad \mathscr{F}_k(j) = \sigma\big(X_r, r \leq 2^k + (j-1)(p_k + q_k) \big).$$

Then, it is easy to give the following decomposition:

$$\begin{aligned} S_n(x_m) &= \sum_{j=1}^{r_k} \zeta_k(j, x_m) + \sum_{j=1}^{r_k} E\big[U_k(j, x_m) | \mathscr{F}_k(j-1) \big] \\ &\quad + \sum_{j=1}^{r_k+1} E\big[V_k(j, x_m) | \mathscr{F}_k(j-1) \big] =: \Sigma_1 + \Sigma_2 + \Sigma_3. \end{aligned}$$

The main ideas of the proof are as follows: First, note that Σ_1 is a martingale, in order to obtain good estimation of the tail probability for Σ_1, the exponential inequality is necessary.

Clearly, the Freedman inequality is suitable for the present setting. However, some preliminary work is required. More precisely, one needs to derive the growth rates of the bound for the martingale differences, and the bound of the sum of the conditional variances. Second, we will show that Σ_2 and Σ_3 can be negligible, that is $\Sigma_2 + \Sigma_3$ is of order $o_{a.s.}(\lambda_{2^k})$.

The procedure as above follows from three facts below:

(F1) We first show that $\max_{1\le j\le r_k} |\zeta_k(j, x_m)| \le p_k$ *with probability one.* We use the following decomposition:

$$\zeta_k(j,x_m) = \frac{1}{\sqrt{h_{2^k}}}\sum_{i\in\mathbb{I}_k(j)}\big(A_i(x_m) - E[A_i(x_m)|\mathscr{F}_k^{i-1}]\big) + \frac{1}{\sqrt{h_{2^k}}}\sum_{i\in\mathbb{I}_k(j)}\big(E[A_i(x_m)|\mathscr{F}_k^{i-1}] - E[A_i(x_m)|\mathscr{F}_k(j-1)]\big) =: \Sigma_{11} + \Sigma_{12},$$

where

$$A_i(x_m) = K_i(x_m) - EK_i(x_m), \quad \mathscr{F}_k^i(j) = \sigma(X_r, r \le 2^k + (j-1)(p_k+q_k) + i - 1).$$

Thus by Markov's inequality, it follows for some $s > 0$,

$$\sum_{k=1}^{\infty} P\Big(\max_{1\le j\le r_k}\big|\zeta_k(j,x_m)\big| > p_k\Big) \le C\sum_{k=1}^{\infty} r_k p_k^{-s}\Big(E\big|\Sigma_{11}\big|^s + E\big|\Sigma_{12}\big|^s\Big). \tag{5.7}$$

It suffices to prove the term on the right-hand side of (5.7) is finite. In fact, one only needs to estimate $E\big|\Sigma_{11}\big|^s$ and $E\big|\Sigma_{12}\big|^s$. For the first term, denote $B_{i,1}(x_m) = A_i(x_m) - E[A_i(x_m)|\mathscr{F}_k^{i-1}]$. Furthermore define recursively $B_{i,l}(x_m) = B_{i,l-1}^2(x_m) - E[B_{i,l-1}^2(x_m)|\mathscr{F}_k^{i-1}], l \in \mathbb{N}$. Clearly, $B_{i,l}(x_m), i = 2^k + (j-1)(p_k+q_k)+1, ..., 2^k + (j-1)q_k + jp_k$, are martingale differences. We will show by the induction method that

$$E\big|\Sigma_{11}\big|^{2^l} = O\Big[p_k^{2^{l-1}}\Big], \quad l \in \mathbb{N}^+. \tag{5.8}$$

If $l = 1$, recalling the conditions on K and f, it turns out that

$$E\Sigma_{11}^2 \le \frac{4}{h_{2^k}}\sum_{i\in\mathbb{I}_k(j)} EK_i^2(x_m) = O(p_k). \tag{5.9}$$

Suppose that (5.8) holds true for $3 \le l \in \mathbb{N}^+$. Then using the Marcinkiewicz−Zygmund−Burkholder inequality, we have for $l \in \mathbb{N}^+$,

$$\begin{aligned} E\big|\Sigma_{11}\big|^{2^l} &\le c_l h_{2^k}^{-2^{l-1}} E\Big|\sum_{i\in\mathbb{I}_k(j)} B_{i,1}^2(x_m)\Big|^{2^{l-1}} \\ &\le c_l h_{2^k}^{-2^{l-1}} E\Big|\sum_{i\in\mathbb{I}_k(j)} B_{i,2}(x_m)\Big|^{2^{l-1}} + c_l h_{2^k}^{-2^{l-1}} E\Big|\sum_{i\in\mathbb{I}_k(j)} E[B_{i,1}^2(x_m)|\mathscr{F}_{i-1}]\Big|^{2^{l-1}}. \end{aligned} \tag{5.10}$$

According to the induction hypothesis, one can easily get the first term in (5.10) is of order $O(p_k^{2^{l-2}})$. As for the second term in (5.10), note that with probability one,

$$E[B_{i,1}^2(x_m)|\mathscr{F}_{i-1}] \le 4h_{2^k}\|K\|_2^2\|f\|_\infty, \tag{5.11}$$

which is of order $O(p_k^{2^{l-1}})$.

We next show (5.8) holds true for any integer s on $(2^l, 2^{l+1})$. By Lyapunov's inequality, it turns out for large k,

$$E\big|\Sigma_{11}\big|^{s} = E\big|\Sigma_{11}\big|^{(2s-2^{l+1})+(2^{l+1}-s)}$$
$$\le \big[E\Sigma_{11}^{2^l}\big]^{(2^{l+1}-s)/2^l}\big[E\Sigma_{11}^{2^{l+1}}\big]^{(s-2^l)/2^l} = O\big[p_k^{s/2}\big]. \tag{5.12}$$

As to Σ_{12}. Recalling the conditions on K and f, we have with probability one,

$$E[A_i(x_m)|\mathscr{F}_k^{i-1}] = h_{2^k}\int_{-\infty}^{\infty} K(u)f(x_m+uh_{2^k}|\mathscr{F}_k^{i-1})du = O(h_{2^k}). \tag{5.13}$$

Thus it follows for large k,

$$E\big|\Sigma_{12}\big|^{s} = O(p_k^s h_{2^k}^{s/2}). \tag{5.14}$$

On account of condition (B2), one can choose large s, then taking (5.12) and (5.14) back into (5.7), the desired result is obtained by the Borel-Cantelli lemma.

(F2) We also have

$$\sum_{j=1}^{r_k} E\big[\zeta_k^2(j,x_m)|\mathscr{F}_k(j-1)\big] = O_{a.s.}\big[2^k\big]. \tag{5.15}$$

Which is proved as follows: Observe that

$$E\big[\zeta_k^2(j,x_m)|\mathscr{F}_k(j-1)\big] = E\big[U_k^2(j,x_m)|\mathscr{F}_k(j-1)\big] - \big(E\big[U_k(j,x_m)|\mathscr{F}_k(j-1)\big]\big)^2 =: \Delta_1 - \Delta_2.$$

For Δ_2, by the definition of $\rho-$mixing (or Lemma A.1), we have

$$E\big(E\big[U_k(j,x_m)|\mathscr{F}_k(j-1)\big]\big)^2$$
$$= E\big\{U_k(j,x_m)E[U_k(j,x_m)|\mathscr{F}_k(j-1)]\big\}$$
$$\le 4\rho(q_k)E\big[U_k(j,x_m)\big]^2 = O\big[p_k\rho(q_k)\big].$$

Using $\sum_{i=0}^{\infty}\rho(2^i)<\infty$, it turns out that

$$\sum_{k=1}^{\infty}\sum_{j=1}^{r_k} E\big(E[U_k(j,x_m)|\mathscr{F}_k(j-1)]\big)^2/2^k < \infty, \tag{5.16}$$

which implies

$$\sum_{k=1}^{\infty}\sum_{j=r_{k-1}+1}^{r_k} E\big(E[U_k(j,x_m)|\mathscr{F}_k(j-1)]\big)^2/2^k < \infty. \tag{5.17}$$

Applying the Borel-Cantelli lemma and the Kronecher lemma gives

$$\sum_{j=1}^{r_n}\big(E[U_n(j,x_m)|\mathscr{F}_n(j-1)]\big)^2 = o_{a.s.}(2^n). \tag{5.18}$$

Thus, the estimation for Δ_2 is finished.

For Δ_1, let

$$U_k^{'}(j,x_m) = U_k(j,x_m)\mathrm{I}(|U_k(j,x_m)|\le 2^{k/2}),\quad U_k^{''}(j,x_m) = U_k(j,x_m)\mathrm{I}(|U_k(j,x_m)|>2^{k/2}).$$

Obviously, we have

$$E\big[U_k(j,x_m)|\mathscr{F}_k(j-1)\big] = E\big[U_k^{'}(j,x_m)|\mathscr{F}_k(j-1)\big] + E\big[U_k^{''}(j,x_m)|\mathscr{F}_k(j-1)\big]. \tag{5.19}$$

Note that

$$E\big(E[U_k^{''2}(j,x_m)|\mathscr{F}_k(j-1)]\big) = EU_k^2(j,x_m)\mathrm{I}(|U_k(j,x_m)|>2^{k/2})$$
$$\le 2^{-k\delta/2}E|U_k(j,x_m)|^{2+\delta} \le 2^{-k\delta/2}\big[2^{ak(2+\delta)/2} + 2^{ak}h_{2^k}^{-\delta/2}\big]. \tag{5.20}$$

Consequently,

$$\sum_{k=1}^{\infty}\sum_{j=1}^{r_k} 2^{-k}E\big(E[U_k^{''2}(j,x_m)|\mathscr{F}_k(j-1)]\big) < \infty. \tag{5.21}$$

Using the Kronecher lemma yields

$$\sum_{k=1}^{n}\sum_{j=1}^{r_k} E\big(E[U_k^{''2}(j,x_m)|\mathscr{F}_k(j-1)]\big) = o_{a.s.}\big(2^n\big). \tag{5.22}$$

As for the first term on the right-hand side of (5.19), it follows that

$$\begin{aligned}
&P\Big(\sum_{j=1}^{r_k}\big[(U_k^{'}(j,x_m)-EU_k^{'}(j,x_m))|\mathscr{F}_k(j-1)\big] \geq \epsilon 2^k\Big)\\
&\leq C2^{-2k}E\Big[\sum_{j=1}^{r_k}\big[(U_k^{'}(j,x_m)-EU_k^{'}(j,x_m))|\mathscr{F}_k(j-1)\big]\Big]^2\\
&\leq C2^{-2k}\sum_{j=1}^{r_k}\big(\rho(q_k)\big)^2E[U_k^{'}(j,x_m)]^2(\log r_k)^2\\
&\leq C2^{-2k}\sum_{j=1}^{r_k}k^{-2}2^{2ak}(\log r_k)^2 \leq C2^{(a-1)k}.
\end{aligned}$$

The second inequality as above follows from Lemma A.4 with $p = 2$.

Thus by the Borel-Cantelli lemma, we have

$$\sum_{j=1}^{r_k}\big[(U_k^{'}(j,x_m)-EU_k^{'}(j,x_m))|\mathscr{F}_k(j-1)\big] = O_{a.s.}(2^k). \tag{5.23}$$

According to (5.18), (5.22) and (5.23), the proof of (5.15) is complete. Furthermore, w.l.o.g., there exists some constant $V > 0$, such that

$$2^{-k}\sum_{j=1}^{r_k}E\big[\zeta_k^2(j,x_m)|\mathscr{F}_k(j-1)\big] \leq V \qquad a.s. \tag{5.24}$$

(F3) Since Σ_3 is a sum on the small blocks, we only need to consider Σ_2. We have

$$\sum_{j=1}^{r_k}E[U_k(j,x_m)|\mathscr{F}_k(j-1)] = o_{a.s.}\Big[\sqrt{2^k|\log h_{2^k}|}\Big]. \tag{5.25}$$

Note that $2^k|\log h_{2^k}| \geq 2^k$ for $k \geq k_0$, clearly, (5.25) follows from the stronger result below.

$$\sum_{j=1}^{r_k}E[U_k(j,x_m)|\mathscr{F}_k(j-1)] = o_{a.s.}\big(\sqrt{2^k}\big). \tag{5.26}$$

Furthermore, (5.26) can be derived from (5.17).

On account of the preliminary work as above, we next consider (5.5), observe that

$$P\Big(\max_{2^k\leq n<2^{k+1}}|S_n(x_m)| \geq M\lambda_{2^{k+1}}\Big) = P\Big(\max_{1\leq j<2^k}|S_{2^k+j}(x_m)| \geq M\lambda_{2^{k+1}}\Big). \tag{5.27}$$

Then similarly to the proof of Lemma 3.1 in Herrndorf [11], we have

$$\begin{aligned} & P\Big(\max_{1\le j<2^k}|S_{2^k+j}(x_m)|\ge M\lambda_{2^{k+1}}\Big) \\ & \le P\big(|S_{2^{k+1}}(x_m)|\ge M\lambda_{2^{k+1}}/3\big) \\ & \quad \times\Big[1-4\rho^2(1)-\max_{1\le j<2^k}P\big(|S_{2^{k+1}}(x_m)-S_{2^k+j}(x_m)|\ge M\lambda_{2^{k+1}}/3\big)\Big]^{-1}. \end{aligned}$$

By the Markov inequality, we have

$$\max_{1\le j<2^k}P\big(|S_{2^{k+1}}-S_{2^k+j}|\ge M\lambda_{2^{k+1}}/3\big)\le 1/2. \tag{5.28}$$

Recalling $\rho(1)\le 1/4$, which together with (5.28) yields

$$P\Big(\max_{1\le j<2^k}|S_{2^k+j}(x_m)|\ge M\lambda_{2^{k+1}}\Big)\le 4P\big(|S_{2^{k+1}}(x_m)|\ge M\lambda_{2^{k+1}}/3\big). \tag{5.29}$$

Finally, with the help of (F1)−(F3) together with (5.29), then by the Freedman inequality (see, e.g., Lemma A.3 in the Appendix), we have for $M\ge\sqrt{74V}$,

$$\begin{aligned} & \sum_{k=1}^{\infty}\sum_{m=1}^{l_k}P\Big(\max_{2^k\le n<2^{k+1}}|S_n(x_m)|\ge M\lambda_{2^{k+1}}\Big) \\ & \le C\sum_{k=1}^{\infty}l_kP\Big(\Big|\sum_{j=1}^{r_k}\zeta_k(j,x_m)\Big|\ge M\lambda_{2^{k+1}}/6\Big) \\ & \le C\sum_{k=1}^{\infty}l_k\exp\left\{\frac{-M^2 2^k|\log h_{2^k}|}{72Mp_k2^{k/2}|\log h_{2^k}|^{1/2}+36V2^k}\right\} \\ & \le C\sum_{k=1}^{\infty}l_kh_{2^k}^{M^2/37V}<\infty. \end{aligned} \tag{5.30}$$

Thus, applying the Borel-Cantelli lemma yields (5.1).

Remark 4. We compare (5.1) with those of Peligrad [22] and Shao [30] for mixing observations. Peligrad [22] obtained the following result: Let D be a compact support subset of $\mathbb{R}^d$ and $\{X_n, n\ge 1\}$ be a sequence of $\mathbb{R}^d$-valued $\phi-$mixing random variables with common unknown density function $f(x)=f(x_1,...,x_d)$. Suppose that (B1) holds, f is continuous in a ε-neighborhood of D, and kernel K satisfies:
1) K is a density function on $\mathbb{R}^d$,
2) for any $x\in\mathbb{R}^d$, $K(x)\le K_1<\infty$,
3) $\|x\|^{d+1}K(x)\to 0$ as $x\to\infty$,
4) $\int\|x\|K(x)dx<\infty$,
5) K is Lipschitzian continuous of order γ on $\mathbb{R}^d$.
Further, assume that

$$\sum_{n=1}^{\infty}\phi^{1/2}(2^n)<\infty, \tag{5.31}$$

then it follows that

$$\sup_{x\in D}|f_n(x)-Ef_n(x)|=O_{a.s.}\Big[\Big(\frac{\log^2 n}{nh_n^d}\Big)^{1/2}\Big]. \tag{5.32}$$

Especially, if $h_n = O((\log^2 n/n)^{1/(d+2)})$, it turns out that

$$\sup_{x\in D} |f_n(x) - f(x)| = O_{a.s.}\Big[\Big(\frac{\log^2 n}{n}\Big)^{1/(d+2)}\Big]. \tag{5.33}$$

If the mixing rates are strengthened to $\phi(n) = O(n^{-2-d})$, Shao [30] obtained

$$\sup_{x\in D} |f_n(x) - Ef_n(x)| = O_{a.s.}\Big[\Big(\frac{\log n}{nh_n^d}\Big)^{1/2}\Big]. \tag{5.34}$$

Especially, if $h_n = O((\log n/n)^{1/(d+2)})$, it turns out that

$$\sup_{x\in D} |f_n(x) - f(x)| = O_{a.s.}\Big[\Big(\frac{\log n}{n}\Big)^{1/(d+2)}\Big]. \tag{5.35}$$

Under conditions 3) and 4), by Bochner−Parzen theorem, we have $Ef_{n,K}(x) - f(x) = O(h_n)$. Note that $\phi-$mixing is contained in $\rho-$mixing, so Theorem 5.1 holds true for $\phi-$mixing data. Obviously, the rate in (5.1) is better than that in (5.32) under the condition (5.31); (5.34) achieves the best possible a.s. convergence rate, however, the mixing rate $\phi(n) = O(n^{-2-d})$ is more stronger than that in (5.31).

Theorem 5.2. *Let $X \in L^p$ for $p \geq 2$. Under the conditions of Theorem 5.1, we have*

$$\sup_{x\in\mathbb{R}} \big|f_{n,K}(x) - Ef_{n,K}(x)\big| = O_{a.s.}\Big[\Big(\frac{|\log h_n|}{nh_n}\Big)^{1/2}\Big]. \tag{5.36}$$

Proof. In order to prove (5.36), it suffices to show that

$$\sup_{x\leq n^{3/p}} \big|f_{n,K}(x) - Ef_{n,K}(x)\big| = O_{a.s.}\Big[\Big(\frac{|\log h_n|}{nh_n}\Big)^{1/2}\Big] \tag{5.37}$$

and

$$\sup_{x> n^{3/p}} \big|f_{n,K}(x) - Ef_{n,K}(x)\big| = O_{a.s.}\Big[\Big(\frac{|\log h_n|}{nh_n}\Big)^{1/2}\Big]. \tag{5.38}$$

Along the similar proof lines as those in Theorem 5.1, one can complete the proof of (5.37). Therefore, we only need to consider (5.38). Note that K is bounded with compact support and $X \in L^p$, using the Markov inequality, it follows that

$$\begin{aligned} E\Big[\sup_{x>n^{3/p}} \big|f_{n,K}(x) - Ef_{n,K}(x)\big|\Big] &\leq 2h_n^{-1}E\Big[\sup_{x>n^{3/p}} \big|K_i(x)\big|\Big] \\ &\leq Ch_n^{-1}P(|X_i| \geq n^{3/p}/2) \leq Cn^{-3}h_n^{-1}, \end{aligned} \tag{5.39}$$

which implies for some $M > 0$,

$$\sum_{n=1}^{\infty} P\Big(\sup_{x>n^{3/p}} \big|f_{n,K}(x) - Ef_{n,K}(x)\big| \geq M(|\log h_n|/nh_n)^{1/2}\Big) \leq \sum_{n=1}^{\infty} \frac{1}{n^2}. \tag{5.40}$$

Thus, (5.38) is obtained by the Borel-Cantelli lemma.

Appendix

We list the following basic lemmas, the first one comes from Bradley and Bryc [2], the second can be found in Shao [33], while the third one is due to Freedman [7].

Lemma A.1. *Let $p,\ q > 1$ with $1/p + 1/q = 1$. Suppose that $X \in L^p(\mathscr{F}_1^k)$ and $Y \in L^q(\mathscr{F}_{k+n}^{\infty})$ are two $\rho-$mixing random variables. Then we have*

$$|EXY - EXEY| \le 10\big(\rho(n)\big)^{\frac{2}{p}\wedge\frac{2}{q}}\|X\|_p\|Y\|_q.$$

Lemma A.2. *Let $\{X_n, n \ge 1\}$ be a sequence of $\rho-$mixing random variables with mean zero and $\|X_i\|_p < \infty$ for some $p \ge 2$. Then there exists a constant L depending only on p and $\rho(\cdot)$ such that for any $k \ge 0, n \ge 1$,*

$$\begin{aligned} E|S_k(n)|^p \le{}& Ln^{p/2}\exp\left[L\sum_{i=0}^{[\log n]}\rho(2^i)\right]\max_{k<i\le k+n}\big\|X_i\big\|_2^p \\ &+ Ln\exp\left[L\sum_{j=0}^{[\log n]}\rho^{2/p}(2^j)\right]\max_{k<i\le k+n}\big\|X_i\big\|_p^p. \end{aligned}$$

Lemma A.3. *Let $\{X_n, \mathscr{F}_n, n \ge 1\}$ be a martingale difference sequence with $S_n = \sum_{i=1}^n X_i$. Suppose that τ is a stopping time, and L a positive real number. Suppose $P(|X_i| \le L, i \le \tau) = 1$. Then for all positive real numbers a and b,*

$$P(S_n \ge a, T_n \le b \quad for\ some \quad n \le \tau) \le \exp\left[\frac{-a^2}{2(La+b)}\right],$$

where $T_n = \sum_{i=1}^n Var(X_i|\mathscr{F}_{i-1})$.

Before formulating the next Lemma, we give some extra notation: Let p_k and q_k be sequences of positive integers satisfying $q_k = o(p_k)$ and $q_k \nearrow \infty$ as $k \to \infty$. Then the successively blocks which only include integers are defined as follows:

$\mathbb{I}_k = [(k-1)(p_k+q_k)+1, (k-1)q_k + kp_k) \cap \mathbb{N}^+,$

$\mathbb{J}_k = [(k-1)q_k + kp_k, k(p_k+q_k)) \cap \mathbb{N}^+, \quad k = 1, 2, ...$

Furthermore, let $\{X_n, n \ge 1\}$ be a sequence of $\rho-$mixing random variables with mean zero, and $E|X_n|^p < \infty$ for some $p \ge 2$. Denote

$$\xi_k = \sum_{i\in\mathbb{I}_k} X_i, \qquad \eta_k = \sum_{i\in\mathbb{J}_k} X_i.$$

Lemma A.4. *Let $\{\xi_n, n \ge 1\}$ be as above. Suppose that $\sum_{i=0}^{\infty}\rho(2^i) < \infty$, then for $p \ge 2$, there exists a positive constant $L = L(p, \rho(\cdot))$ depending only on p and $\rho(\cdot)$ such that for every $k \ge 0, n \ge 1$,*

$$\begin{aligned} E|G_m(n)|^p \le{}& L(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\big\|\xi_k\big\|_2^2\right]^{p/2} \\ &+ L(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-1)}(q(k/2))\big\|\xi_k\big\|_p^p\right], \end{aligned} \tag{A.1}$$

where $G_m(n) = \sum_{k=m+1}^{m+n} E(\xi_k|\mathscr{F}_{k-1})$, $\mathscr{F}_k = \sigma(X_i, i \le (k-1)(p_k+q_k)+p_k)$, and $q(x)$ is the linear interpolating function of q_k.

Proof. We claim that p may not be integer. However, we only consider the integral case in the present article since the proof for the non-integral part is similar. We first show that (A.1)

holds true for all even numbers, then verify that it is the case for all odd numbers. The proof is decomposed into the following three steps by induction on p.
Step 1. If $p = 2$, (A.1) can be rewrote as follows:

$$EG_m^2(n) \le L(\log 2n)^2 \left[\sum_{k=m+1}^{m+n} \rho^2(q(k/2)) \big\| \xi_k \big\|_2^2 \right]. \tag{A.2}$$

A similar result as (A.2) can be found in Shao [31]. However, for the reader's convenience, we give its proof by induction on n below: If $n = 1$, by Lemma A.1, it follows that

$$\begin{aligned} EG_m^2(1) &= E(E(\xi_{m+1}|\mathscr{F}_m))^2 \\ &= E(\xi_{m+1} E(\xi_{m+1}|\mathscr{F}_m)) \\ &\le 10\rho(q_m)\|\xi_{m+1}\|_2 \|E(\xi_{m+1}|\mathscr{F}_m)\|_2. \end{aligned} \tag{A.3}$$

Then a simple calculation leads to

$$EG_m^2(1) \le 100\rho^2(q_m)\|\xi_{m+1}\|_2^2. \tag{A.4}$$

Suppose that (A.2) holds true for any integer less than n. We next show it remains valid for n itself. Let $n_1 = [n/2], n_2 = n - n_1$. Clearly,

$$\begin{aligned} EG_m^2(n) &= EG_m^2(n_1) + EG_{m+n_1}^2(n_2) + 2EG_m(n_1)G_{m+n_1}(n_2) \\ &\le EG_m^2(n_1) + EG_{m+n_1}^2(n_2) + 2\rho(q_{m+n_1})\|G_m(n_1)\|_2 \|G_{m+n_1}(n_2)\|_2. \end{aligned} \tag{A.5}$$

By the induction hypothesis and Lemma A.2 with $\sum_{i=0}^{\infty} \rho(2^i) < \infty$, we have

$$\begin{aligned} EG_m^2(n) \le{}& L(\log 2n_1)^2 \left[\sum_{k=m+1}^{m+n_1} \rho^2(q(k/2)) \big\|\xi_k\big\|_2^2 \right] \\ &+ L(\log 2n_2)^2 \left[\sum_{k=m+n_1+1}^{m+n} \rho^2(q(k/2)) \big\|\xi_k\big\|_2^2 \right] \\ &+ L\log(2n_1)\rho(q_{m+n_1}) \left[\sum_{k=m+n_1+1}^{m+n} \big\|\xi_k\big\|_2^2 \right]^{1/2} \left[\sum_{k=m+1}^{m+n_1} \rho^2(q(k/2)) \big\|\xi_k\big\|_2^2 \right]^{1/2} \\ \le{}& L(\log 2n)^2 \left[\sum_{k=m+1}^{m+n} \rho^2(q(k/2)) \big\|\xi_k\big\|_2^2 \right]. \end{aligned} \tag{A.6}$$

Step 2. Let $p = 2l$ for $l \ge 2$. Suppose that (A.1) holds true for all even numbers less than p. We will show that it is also valid for p. To this end, the following preliminary work is needed.

(i) We will derive an upper bound for the pth moment of $G_k(n)$. The basic inequalities below are useful: For any $x \ge 0$ and $p > 1$, we have

$$(1+x)^p \le 1 + x^p + 4^p(x + x^{p-1}). \tag{A.7}$$

Moreover, let $\alpha,\ \beta > 0$ and $\alpha + \beta = 1$. Applying Young's inequality, we have for any $x,\ y > 0$,

$$x^\alpha y^\beta \le \alpha x + \beta y \le x + y. \tag{A.8}$$

Thus by (A.7), it follows for $n \ge 2$,

$$\begin{aligned} EG_m^p(n) &= E(G_m(n_1) + G_{m+n_1}(n_2))^p \\ &\le E(G_m(n_1))^p + E(G_{m+n_1}(n_2))^p \\ &\quad + 4^p\big(EG_m(n_1)(G_{m+n_1}(n_2))^{p-1} + E(G_m(n_1))^{p-1}G_{m+n_1}(n_2)\big). \end{aligned} \tag{A.9}$$

Using Lemma A.1 and (A.8), we give the following stronger upper bound,

$$\begin{aligned}
&EG_m(n_1)(G_{m+n_1}(n_2))^{p-1} \\
&\quad \le E|G_m(n_1)|E|G_{m+n_1}(n_2)|^{p-1} + 10\rho^{2/p}(q_{m+n_1})\|G_m(n_1)\|_p\|G_{m+n_1}(n_2)\|_p^{p-1} \\
&\quad \le E|G_m(n_1)|E|G_{m+n_1}(n_2)|^{p-1} + 10\rho^{2/p}(q_{m+n_1})\Big[\|G_m(n_1)\|_p^p + \|G_{m+n_1}(n_2)\|_p^p\Big].
\end{aligned}$$

Similarly, we have

$$\begin{aligned}
&E(G_m(n_1))^{p-1}G_{m+n_1}(n_2) \\
&\quad \le E|G_m(n_1)|^{p-1}E|G_{m+n_1}(n_2)| + 10\rho^{2/p}(q_{m+n_1})\Big[\|G_m(n_1)\|_p^p + \|G_{m+n_1}(n_2)\|_p^p\Big].
\end{aligned}$$

Hence combining the estimations as above, we get

$$\begin{aligned}
EG_m^p(n) &\le \big(1 + 20\times 4^p\rho^{2/p}(q_{m+n_1})\big)\Big[EG_m^p(n_1) + EG_{m+n_1}^p(n_2)\Big] \\
&\quad + E|G_m(n_1)|^{p-1}E|G_{m+n_1}(n_2)| + E|G_m(n_1)|E|G_{m+n_1}(n_2)|^{p-1} \\
&\le \big(1 + 20\times 4^p\rho^{2/p}(q_{m+n_1})\big)\Big[E|G_m(n_1)|^p + E|G_{m+n_1}(n_2)|^p\Big] \\
&\quad + \parallel G_m(n_1) \parallel_{p-1}^{p-1}\parallel G_{m+n_1}(n_2) \parallel_2 + \parallel G_m(n_1) \parallel_2\parallel G_{m+n_1}(n_2) \parallel_{p-1}^{p-1}.
\end{aligned}$$

(ii) We next show for each $n \ge 1$,

$$\begin{aligned}
EG_m^p(n) &\le L(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{p/2} \\
&\quad + L(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-2)}(q(k/2))\|\xi_k\|_{p-1}^{p-1}\right]\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{1/2} \\
&\quad + L(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-1)}(q(k/2))\|\xi_k\|_p^p\right]. \qquad \text{(A.10)}
\end{aligned}$$

To prove (A.10), we apply the induction method on n: If $n = 1$, it follows from Lemma A.1,

$$EG_m^p(1) = E\big[G_m(1)(G_m(1))^{p-1}\big] \le 10\rho^{2/p}(q_m)\parallel \xi_{m+1} \parallel_p\parallel G_m(1) \parallel_p^{p-1}. \qquad \text{(A.11)}$$

Then a standard calculation leads to

$$EG_m^p(1) \le 10\rho^2(q_m)\parallel \xi_{m+1} \parallel_p^p. \qquad \text{(A.12)}$$

Observe that $\rho^2(q_m) \le \rho^{2/(p-1)}(q_m)$ for $l \ge 1$, it is easy to see that (A.10) holds true.

Suppose that (A.10) is valid for any integer less than n, we next show that it is the case for

n itself. On account of the induction hypothesis, $EG_m(n)^p$ is less than or equal to

$$L(\log 2n)^p\left\{\left[\sum_{k=m+1}^{m+n_1}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{p/2}+\left[\sum_{k=m+n_1+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{p/2}\right\}$$

$$+L(\log 2n)^p\left\{\left[\sum_{k=m+1}^{m+n_1}\rho^{2/(p-2)}(q(k/2))\|\xi_k\|_{p-1}^{p-1}\right]\left[\sum_{k=m+1}^{m+n_1}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{1/2}\right.$$

$$\left.+\left[\sum_{k=m+n_1+1}^{m+n}\rho^{2/(p-2)}(q(k/2))\|\xi_k\|_{p-1}^{p-1}\right]\left[\sum_{k=m+n_1+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{1/2}\right\}$$

$$+\left[L\sum_{k=m+1}^{m+n_1}\rho^{2/(p-1)}(q(k/2))\|\xi_k\|_p^p+L\sum_{k=m+n_1+1}^{m+n}\rho^{2/(p-1)}(q(k/2))\|\xi_k\|_p^p\right]$$

$$+4^{3p}L\left\{\log^{1/2}(2n)\left(\sum_{k=k+1}^{k+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right)^{1/2}(\log 2n)^{p-1}\right.$$

$$\left.\times\left\{\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-2)}(q(k/2))\|\xi_k\|_{p-1}^{p-1}\right]+\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{\frac{p-1}{2}}\right\}\right\}$$

$$\leq L(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{p/2}$$

$$+L(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-1)}(q(k/2)))\|\xi_k\|_p^p\right]$$

$$+L(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-2)}(q(k/2))\|\xi_k\|_{p-1}^{p-1}\right]\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{1/2}.$$

Therefore the proof of (A.10) is complete.

(iii) We finally verify that (A.1) holds true for $p = 2l$. Observe that

$$(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-2)}(q(k/2))\|\xi_k\|_{p-1}^{p-1}\right]\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{1/2}$$

$$\leq(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-1)}(q(k/2))\|\xi_k\|_{p-1}^{p-1}\right]\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{1/2}. \quad (A.13)$$

Then by Lyapunov's inequality, it follows that

$$\sum_{k=m+1}^{m+n}\rho^{2/(p-1)}(q(k/2))\|\xi_k\|_{p-1}^{p-1}\leq\sum_{k=m+1}^{m+n}\rho^{2/(p-1)}(q(k/2))\|\xi_k\|_2^{2/(p-2)}\|\xi_k\|_p^{p(p-3)/(p-2)}. \quad (A.14)$$

Furthermore note that

$$\frac{2}{p-1}=\frac{2}{p-2}+\frac{2(p-3)}{(p-1)(p-2)}. \quad (A.15)$$

Therefore by the Hölder inequality, the right hand side of (A.14) is less than or equal to

$$\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{1/(p-2)}\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-1)}(q(k/2))\|\xi_k\|_p^p\right]^{(p-3)/(p-2)}. \quad (A.16)$$

On account of (A.14) and (A.16), the right hand side of (A.13) is controlled by

$$
\begin{aligned}
&(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{\frac{1}{2}+\frac{1}{p-2}}\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-1)}(q(k/2))\|\xi_k\|_p^p\right]^{\frac{p-3}{p-2}}\\
&=(\log 2n)^p\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{\frac{p}{2}\times\frac{1}{p-2}}\left[\sum_{k=m+1}^{m+n}\rho^{2/(p-1)}(q(k/2))\|\xi_k\|_p^p\right]^{\frac{p-3}{p-2}}\\
&\le(\log 2n)^p\left\{\left[\sum_{k=m+1}^{m+n}\rho^2(q(k/2))\|\xi_k\|_2^2\right]^{p/2}+\sum_{k=m+1}^{m+n}\rho^{2/(p-1)}(q(k/2))\|\xi_k\|_p^p\right\}. \quad (A.17)
\end{aligned}
$$

Step 3. Assume that p is an odd number. Clearly, there exists integer $l \geq 2$ such that $2l-2 < p < 2l$. Again using Lyapunov's inequality and (A.8), we have

$$
\begin{aligned}
E|G_m(n)|^p &= E|G_m(n)|^{2l-1} \qquad (A.18)\\
&\le \left(EG_m(n)^{2l-2}\right)^{1/2}\left(EG_m(n)^{2l}\right)^{1/2}\\
&\le EG_m(n)^{2l-2}+EG_m(n)^{2l}. \qquad (A.19)
\end{aligned}
$$

According to the procedures as above, the proof of (A.1) is complete.

References

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

[2] Bradley, R. C., and Bryc, W. (1985). Multilinear forms and measures of dependence between random variables. J. Multivariate Anal. 16, 335-367. measures of the deviations of density function estimates. Ann. Statist. 1, 1071-1095.

[3] Deheuvels, P. (2000). Uniform limit laws for kernel density estimators on possibly unbounded intervals. In Recent Advances in Reliability Theory: Methodology, Practice and Inference (N. Limnios and M. Nikulin, eds.) 477-492, Birkhäuser, Boston.

[4] Einmahl, U., and Mason, D. M. (2000). An empirical process approach to the uniform consistency of kernel-type function estimators. J. Theoret. Probab. 13, 1-37.

[5] Einmahl, U., and Mason, D. M. (2005). Uniform in bandwidth consistency of kernel-type function estimators. Ann. Statist. 33, 1380-1403.

[6] Földes, A. (1974). Density estimation for dependent sample. Studia Scientiarum Mathematicarum Hungarica 9, 443-452.

[7] Freedman, D. A. (1975). On tail probabilities for martingales. Ann. Probab. 3, 100-118.

[8] Giné, E., and Guillou, A. (2002). On consistency of kernel density estimators for randomly censored data: Rates holding uniformly over adaptive intervals. Ann. Inst. H. Poincaré Probab. Statist. 37, 503-522.

[9] Giné, E., and Guillou, A. (2002). Rates of strong consistency for multivariate kernel density estimators. Ann. Inst. H. Poincaré Probab. Statist, 38, 907-922.

[10] Giné, E., Koltchinskii, V., and Zinn, J. (2004). Weighted uniform consistency of kernel density estimators. Ann. Probab. 32, 2570-2605.

[11] Herrndorf, N. (1983). The Invariance Principle for p-mixing Sequences. Z. Wahr. verw. Gebiete 63, 97-108.

[12] Kolmogorov, A. N., and Rozanov, U. A. (1960). On the strong mixing conditions of a stationary Gaussian process. Probab Theory Appl. 2, 222-227.

[13] Liebscher, E. (1995). Strong convergence of sums of $\varphi-$mixing random variables. Math. Methods Statist. 4, 216-229.

[14] Liebscher, E. (1996). Strong convergence of sums of $\alpha-$mixing random variables with applications to density estimation. Stochastic Process. Appl. 65, 69-80.

[15] Lin, Z., and Lu, C. (1997). Limit Theory on Mixing Dependent Random Variables, Science Press and Kluwer Academic Publishers.

[16] Neumann, M. H. (1998). Strong approximation of density estimators from weakly dependent observations by density estimators from independent observations. Ann. Statist. 26, 2014-2048.

[17] Nze, P. A., and Rios, R. (1995). Density estimation in the L^{∞} norm for mixing processes. C. R. Acad. Sci. Paris 320, 1259-1262.
[18] Parzen, E. (1962). On the estimation of a probability density function and the mode. Ann. Math. Statist. 33, 1065-1076.
[19] Peligrad, M. (1982). Invariance principles for mixing sequences of random variables. Ann. Probab. 10, 968-981.
[20] Peligrad, M. (1986). Recent advances in the central limit theorem and its weak invariance principle for mixing sequences of random variables. Dependence in Probab. Statist. Eberlin, E. and Taqqu, M. S. (eds) Progress in Probab. Statist. Birckhauser, 11, 193-223.
[21] Peligrad, M. (1987). On the central limit theorem for $\rho-$mixing sequences of random variables. Ann. Probab. 15, 1387-1394.
[22] Peligrad, M. (1992). Properties of uniform consistency of the kernel estimators of density and of regression functions under dependent assumptions. Stoch. Stoch. Reports 40, 147-168.
[23] Peligrad, M., and Shao, Q. M. (1995). Estimation of the variance of partial sums for $\rho-$mixing random variables. J. Multivariate Anal. 52, 140-157.
[24] Peligrad, M., and Utev, S. (2005). A new maximal inequality and invariance principle for stationary sequences. Ann. Probab. 33, 798-815.
[25] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of a density function. Ann. Math. Statist. 27, 832-835.
[26] Roussas, G. G. (1988). nonparametric estimation in mixing sequences of random variables. J. Statist. Plann. Inference 18, 135-149.
[27] Rüschendorf, L. (1977). Consistency of estimators for multivariate density functions and for the mode. Sankhya 39, 243-250.
[28] Sarda, P., and Vieu, P. (1989). Empirical distribution function for mixing random variables. Statistics 20, 559-571.
[29] Shao, Q. M. (1989). On the invariance principle for $\rho-$mixing sequences of random variables. Chin. Ann. Math. 10B, 427-433.
[30] Shao, Q. M. (1990). Exponential inequalities and density estimation under dependent assumptions. Manuscript.
[31] Shao, Q. M. (1993). On the invariance principle for stationary $\rho-$mixing sequence with infinite variance. Chin. Ann. Math. 14B, 27-42.
[32] Shao, Q. M. (1993). Almost sure invariance principles for mixing sequences of random variables. Stochastic Process Appl. 48, 319-334.
[33] Shao, Q. M. (1995). Maximal inequality for partial sums of $\rho-$mixing sequences. Ann. Probab. 23, 948-965.
[34] Silverman, B. W. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. Ann. Statist. 6, 177-184.
[35] Stute, W. (1982). A law of the logarithm for kernel density estimators, Ann. Probab. 10, 414-422.
[36] Stute, W. (1984). The oscillation behavior of empirical processes: The multivariate case. Ann. Probab. 12, 361-379.
[37] Woodroofe, M. (1967). On the maximum deviation of the sample density. Ann. Math. Statist. 38, 475-481.
[38] Woodroofe, M. (1970). Discussion of "Density estimates and Markov sequences" by M. Rosenblatt. In nonparametric Techniques in Statistical Inference (M. Puri, ed.) 211-213, Cambridge Univ. Press.
[39] Wu, W. B., Huang, Y., and Huang, Y. (2010). Kernel estimation for time series: An asymptotic theory. Stochastic Process. Appl. 120, 2412-2431.
[40] S. Yakowitz, nonparametric density estimation, prediction and regression for Markov sequences, J. Amer. Statist. Assoc. 80 (1985) 215-221.
[41] B. Yu, Density estimation in the L^{∞} norm for dependent data with applications to the Gibbs sampler, Ann. Statist. 21 (1993) 711-735.