Asymptotic normality for $L_1$-norm kernel estimator of conditional median under association dependence

Zhengyan Lin, Degui Li*

Department of Mathematics, Zhejiang University, Hangzhou 310027, China

Received 3 September 2005
Available online 20 September 2006

Abstract

Let $\{X_i, Y_i\}_{i=1}^{\infty}$ be a set of observations from a stationary jointly associated process and $\theta(x)$ be the conditional median, that is, $\theta(x) = \inf\{y : P(Y \leq y | X = x) \geq \frac{1}{2}\}$. We consider the problem of estimating $\theta(x)$ based on the $L_1$-norm kernel and establish asymptotic normality of the resulting estimator $\theta_n(x)$.

© 2006 Elsevier Inc. All rights reserved.

AMS 1991 subject classification: primary 62G07; secondary 60F05

Keywords: Associated processes; Asymptotic normality; Conditional median; $L_1$-norm kernel

1. Introduction

Since introduction by Koenker and Bassett [12], the conditional quantile, also called quantile regression or regression quantile, has been used widely over the last three decades in various disciplines, such as finance, economics, medicine and biology. We assume that $\{X_i, Y_i\}_{i=1}^{\infty}$ is a stationary sequence with $X_i$ being $\mathbf{R}^d$-valued and $Y_i$ being real-valued. Denote $F(y|x_0)$ the conditional distribution of $Y$ given $X = x_0$. The conditional quantile function of $F(y|x_0)$ at $x_0$ is defined as, for any $0 < \tau < 1$,

$$F^{-1}(\tau|x_0) = \inf\{y : P(Y \leq y|X = x_0) \geq \tau\}. \quad (1.1)$$

For more details about recent developments of the conditional quantile, we refer to [5,7,13,21] and the references therein.

* Corresponding author.
E-mail address: ldgofzju@tom.com (D. Li).
It is easy to see that the conditional median $\theta(x_0)$, corresponding to $F^{-1}(\frac{1}{2}|x_0)$, is a special conditional quantile. Next, we shall construct the $L_1$-norm kernel estimator $\theta_n(x_0)$ of $\theta(x_0)$. Let $J_n = J_n(x_0)$ be the set $\{i: ||X_i - x_0|| \leq h_n, 1 \leq i \leq n\}$, where $h_n \to 0$, and $N_n = N_n(x_0)$ be the number of points in $J_n$. Here and in the sequel, $\| \cdot \|$ denotes $L_2$-distance. $\theta_n(x_0)$ is defined as follows:

$$\theta_n(x_0) = \theta_{h_n}(x_0) = F_n^{-1}(\frac{1}{2}|x_0),$$

where $F_n(y|x_0) = \frac{1}{N_n} \sum_{i \in J_n} I(Y_i \leq y)$. It is easy to check that $\theta_n(x_0)$ is just the solution of the following problem:

$$\theta_n(x_0) = \arg \min_{z \in \mathbb{R}} \sum_{i=1}^{n} K_{n_i}(x_0)|Y_i - z|,$$

where $K_{n_i}(x_0) = K(\frac{X_i - x_0}{h_n}), K(x) = I(\|x\| \leq 1)$ and $I(\cdot)$ denotes the indicator function. From (1.3), we know that $\theta(x_0)$ is estimated by the method of least absolute deviations. Hence, $\theta_n(x_0)$ is the $L_1$-norm kernel estimator with kernel function $K(x) = I(\|x\| \leq 1)$. As is well known, estimators based on the method of least squares such as the Nadaraya–Watson estimator [15,20] do not perform well when the error distribution is heavy-tailed and are sensitive to outliers. Therefore, it is important to consider the problem of $L_1$-norm kernel estimator $\theta_n(x_0)$ of the conditional median $\theta(x_0)$, since estimators based on the $L_1$-norm are robust to heavy-tailed errors and outliers.

There is extensive literature concerning asymptotic properties of the $L_1$-norm kernel estimator $\theta_n(x_0)$. When the underlying process $\{X_i, Y_i\}_{i=1}^{\infty}$ is independent and identically distributed (i.i.d.), Hong [9,10] gave the asymptotic normality and a law of the iterated logarithm for $\theta_n(x_0)$. When the process $\{X_i, Y_i\}_{i=1}^{\infty}$ is stationary and $\pi$-mixing, Truong and Stone [19] derived the rates of convergence; Zhou and Liang [22] obtained the asymptotic normality of $\theta_n(x_0)$. Honda [8] considered the nonparametric estimation of the conditional median for long-range linear processes by the method of least absolute deviations. In this paper, we shall investigate the case of positive and negative association. The advantage here is that the results are established under summability condition on the covariance functions of the underlying process rather than on the mixing coefficient, which is difficult to verify.

**Definition.** For a finite index set $I$, the real-valued random variables $\{Z_i, i \in I\}$ are said to be (positively) associated if

$$\text{Cov}(G_1(Z_i, i \in I), G_2(Z_j, j \in I)) \geq 0$$

for any real-valued coordinate-wise nondecreasing functions $G_1$ and $G_2$ defined on $\mathbb{R}^I$, provided $EG_j^2(Z_i, i \in I) < \infty$, $j = 1, 2$. They are said to be negatively associated if for any two disjoint nonempty subsets $A$ and $B$ of $I$ and any coordinate-wise nondecreasing functions $G_1 : \mathbb{R}^A \to \mathbb{R}$ and $G_2 : \mathbb{R}^B \to \mathbb{R}$,

$$\text{Cov}(G_1(Z_i, i \in A), G_2(Z_j, j \in B)) \leq 0,$$

provided $EG_1^2(Z_i, i \in A) < \infty$ and $EG_2^2(Z_j, j \in B) < \infty$. If $I$ is not finite, $\{Z_i, i \in I\}$ are said to be associated or negatively associated if every finite subcollection is associated or negatively associated, respectively. In this paper, when no distinction is necessary, associated and negatively associated random variables will be referred to collectively as associated random variables.
The definition of association was introduced by Esary et al. [6] and the definition of negative association was due to Joag-Dev and Proschan [11]. Positive association occurs often in certain reliability theory problems, as well as in some important models employed in statistical mechanics. Furthermore, Pitt [17] showed that Gaussian processes are positively associated, if and only if their covariance functions are positive. We shall stress that the classes of $\alpha$-mixing processes and associated processes are distinct but may overlap through the following example. In the linear time series framework,

$$X_k = \sum_{j=0}^{\infty} a_j v_{k-j},$$

(1.6)

where $\{v_i\}_{i=-\infty}^{\infty}$ is a sequence of i.i.d. random variables with mean zero and variance $\sigma^2$, the sequence $\{X_k\}_{k=1}^{\infty}$ is positively associated if $a_j \geq 0$. On the other hand, Pham and Tran [16] showed that $\{X_k\}_{k=1}^{\infty}$ is $\alpha$-mixing under suitable conditions on $a_j$. In particular, Andrew [1] showed that when $\{v_i\}_{i=-\infty}^{\infty}$ is a sequence of i.i.d. Bernoulli random variables and $a_j = \gamma^j$, $0 < \gamma \leq \frac{1}{2}$, the sequence $\{X_k\}_{k=1}^{\infty}$ is not $\alpha$-mixing, whereas it is still positively associated. Negative association also appears in some reliability theory problems but to a less degree. Joag-Dev and Proschan [11] gave some examples which are negatively associated.

There is a lot of literature on the nonparametric estimation for associated processes. Various statistical estimation problems under association have been studied in depth by many authors, say Cai [4]. From the definition above, we can give the definition of association for multivariate random processes $\{X_i, Y_i\}_{i=1}^{\infty}$ similarly. In fact, this is natural in the context of regression estimation. Masry [14] studied the local polynomial fitting for the multivariate (jointly) associated processes $\{X_i, Y_i\}_{i=1}^{\infty}$.

A technical problem arises when we establish asymptotic normality of $\theta_n(x_0)$. In fact, functions of (jointly) associated processes are not associated in general. In so doing, there are two basic issues to be dealt with. One is that of passing from characteristic functions to covariances, and the other is to study the behavior of covariances of functions of (jointly) associated processes. Both points can be resolved with the help of Lemma 1, which is widely used in the context of function estimation for associated processes. However, since the kernel function $K(x) = I(\|x\| \leq 1)$ is not continuous in $\mathbf{R}$, Lemma 1 cannot be applied directly. We shall first make some interesting transformations, which can be found in Section 3.

The organization of the paper is as follows. Some assumptions and the main results of our paper are provided in Section 2. The lemmas which are useful in our proofs are stated and proved in Section 3. The proofs of the main results are given in Section 4. Throughout the paper, convergence in distribution and convergence in probability are denoted by $\xrightarrow{d}$ and $\xrightarrow{p}$, respectively. $C$ denotes a positive constant which may change from one place to another.

2. Assumptions and main results

Before stating the main results, we first give the following assumptions.

A1. (i) In the neighborhood $U_{x_0} := U_{x_0}(\tilde{\xi}) = \{z : \|z - x_0\| \leq \xi\}$, $\xi > 0$ and $x_0 \in \mathbf{R}^d$, the distribution of $X$ is absolutely continuous and its density function $f(\cdot)$ is bounded away from zero, i.e.,

$$M_1^{-1} \leq f(x) \leq M_1 \quad \text{for } x \in U_{x_0},$$

(2.1)
where $M_1$ may depend on $x_0$. Furthermore, $f(x)$ has bounded partial derivatives in $U_{x_0}$.

(ii) $0 < \frac{\partial F(y|x_0)}{\partial x} \leq M_2$ and $\| \frac{\partial F(y|x)}{\partial x_j} \|_{\infty} \leq M_3$, $j = 1, \ldots, d$, where $M_2$ and $M_3$ are positive constants. Here and in the sequel, $\| \cdot \|_{\infty}$ stands for the sup norm.

A2. If $f_{1,j}$ is the joint density function of $(X_1, X_{j+1})$, $j \geq 1$, we have

$$\sup_{j,u,v} |f_{1,j}(u,v) - f(u)f(v)| \leq M_4,$$  \hspace{1cm} (2.2)

where $M_4$ is a positive constant.

A3. (i) Let $h_n > 0$ be bandwidths such that

$$h_n \rightarrow 0 \quad \text{and} \quad nh_n^d \rightarrow \infty \quad \text{as} \quad n \rightarrow \infty.$$  \hspace{1cm} (2.3)

(ii) The random variables $\{X_i, Y_i\}_{i=1}^{\infty}$ form strictly stationary (jointly) associated processes such that

$$\sum_{p=1}^{d} \sum_{q=1}^{d} \sum_{i=1}^{\infty} i^{\delta_1} |\text{Cov}(X_{1p}, X_{iq})| < \infty,$$  \hspace{1cm} (2.4)

$$\sum_{p=1}^{d} \sum_{i=1}^{\infty} i^{\delta_2} |\text{Cov}(Y_1, X_{ip})| < \infty,$$  \hspace{1cm} (2.5)

$$\sum_{i=1}^{\infty} i^{\delta_3} |\text{Cov}(Y_1, Y_i)| < \infty.$$  \hspace{1cm} (2.6)

where $X_i = (X_{i1}, \ldots, X_{id})$, $\delta_1 > 1 + 4/d$, $\delta_2 > 1 + 3/d$, $\delta_3 > 1 + 2/d$.

(iii) Let $s_n$ be a sequence of positive integers satisfying $s_n \rightarrow \infty$ and $s_n = o((nh_n^d)^{-1})$. And the covariance sequences of the stationary processes $\{X_i, Y_i\}_{i=1}^{\infty}$ satisfy

$$\frac{1}{h_n^{d+4}} \sum_{p=1}^{d} \sum_{q=1}^{d} \sum_{i=s_n}^{\infty} |\text{Cov}(X_{1p}, X_{iq})| \rightarrow 0,$$  \hspace{1cm} (2.7)

$$\frac{1}{h_n^{d+3}} \sum_{p=1}^{d} \sum_{i=s_n}^{\infty} |\text{Cov}(Y_1, X_{ip})| \rightarrow 0,$$  \hspace{1cm} (2.8)

$$\frac{1}{h_n^{d+2}} \sum_{i=s_n}^{\infty} |\text{Cov}(Y_1, Y_i)| \rightarrow 0.$$  \hspace{1cm} (2.9)

Remark 1 (Discussion of conditions).

(D1) Condition A1 is the same as Condition 1.1 in Zhou and Liang [22]. Condition A2 is assumed in many studies of nonparametric estimation for associated processes such as [14,18]. If $\sup_{j,u,v} f_{1,j}(u,v) < \infty$ and $\sup_{u} f(u) < \infty$, it is easy to check that A2 is satisfied.
(D2) Conditions A3(ii) and (iii) seem intricate, but it is natural in the context of nonparametric estimation for associated processes. For instance, both Masry [14] and Roussas [18] had similar conditions. Next, we shall give an example when A3 holds.

Example 1. Suppose that $h_n \to 0$, $\frac{nh_n^d}{(\log n)^2} \to \infty$ and $a(i) \leq C \rho^i$ for some $0 < \rho < 1$ and each $i \geq 1$, where

$$a(i) = \max \left\{ |\text{Cov}(Y_1, Y_i)|, \max_p |\text{Cov}(Y_1, X_{1p})|, \max_{p,q} |\text{Cov}(X_{1p}, X_{iq})| \right\}. $$

A3(i) and (ii) are easy to verify. Now, let $s_n \sim m \log n$, where $m \geq 1 + 4/d$ and $x_n \sim y_n$ stands for $\lim \frac{x_n}{y_n} = 1$. Noting that $m \geq \frac{1}{\log \rho} (1 + 4/d)$, we have

$$\frac{1}{h_n^{d+4}} \sum_{p=1}^d \sum_{q=1}^d \sum_{i=s_n}^\infty |\text{Cov}(X_{1p}, X_{iq})| \leq C \frac{\rho^m \log n}{h_n^{d+4}} \leq C \frac{\rho^m \log n}{h_n^{d+4}} = O(n^m \log \rho h_n^{-d-4}) = o(1).$$

Hence, (2.7) is satisfied. (2.8) and (2.9) can be proved by analogous argument.

Theorem 1. Suppose that Conditions A1–A3 hold and $nh_n^{d+2} \to 0$. Then, we have

$$(nh_n^d)^{\frac{1}{2}} (\theta_n(x_0) - \theta(x_0)) \overset{d}{\to} N(0, \sigma^2),$$

(2.10)

where $\sigma^2 = \frac{4L(1)f^2(\theta(x_0)|x_0)f(x_0)}{L(\zeta)}$, $L(\zeta)$ denotes the volume of $d$-dimensional sphere with radius $\zeta > 0$ and $f(y|x)$ is the conditional density function of $Y$ given $X = x$.

Remark 2. Theorem 1 still holds for the case of negative association. Since the proof is analogous, we omit the details.

Remark 3. Supposed that the process $\{X_i, Y_i\}_{i=1}^\infty$ is a sequence of i.i.d. random vectors, it is easy to check that (2.10) holds under Conditions A1, A3(i) and $nh_n^{d+2} \to 0$ since Conditions A2, A3(ii) and (iii) are obviously satisfied. When the process $\{X_i, Y_i\}_{i=1}^\infty$ is a stationary $\alpha$-mixing sequence, $nh_n^{d+2} \to 0$ and

$$\frac{nh_n^d}{(\log n)^3} \to \infty,$$

(2.11)

Zhou and Liang [22] proved that (2.10) holds under appropriate conditions. We consider the case of association dependence in this paper. Furthermore, Example 1 showed that (2.11) can be weakened to $\frac{nh_n^d}{(\log n)^2} \to \infty$ when $a(i) = O(\rho^i)$, $0 < \rho < 1$.

Remark 4. If the assumption $nh_n^{d+2} \to 0$ is replaced by $nh_n^{d+4} \to \Theta$ for some positive constant $\Theta$, we can establish asymptotic normality of $\theta_n(x_0)$ under the following condition which is a little stronger than A3.
A3'. (i) Let $h_n > 0$ be bandwidths such that $h_n \to 0$.

(ii) The random variables $\{X_i, Y_i\}_{i=1}^\infty$ form strictly stationary (jointly) associated processes such that (2.4)–(2.6) hold with $\delta_1 > 1 + 6/d$, $\delta_2 > 1 + 5/d$, $\delta_3 > 1 + 4/d$.

(iii) Let $s_n$ be a sequence of positive integers satisfying $s_n \to \infty$ and $s_n = o((nh_n^d)^{1/2})$. Furthermore, the covariance sequences of the stationary processes $\{X_i, Y_i\}_{i=1}^\infty$ satisfy

$$\frac{1}{h_n^{d+6+\tau}} \sum_{p=1}^d \sum_{q=1}^d \sum_{i=s_n}^\infty |\text{Cov}(X_{1p}, X_{iq})| \to 0,$$

(2.12)

$$\frac{1}{h_n^{d+5+\tau}} \sum_{p=1}^d \sum_{i=s_n}^\infty |\text{Cov}(Y_1, X_{ip})| \to 0,$$

(2.13)

$$\frac{1}{h_n^{d+4+\tau}} \sum_{i=s_n}^\infty |\text{Cov}(Y_1, Y_i)| \to 0$$

(2.14)

for some $\tau > 0$.

If A3 and $nh_n^{d+2} \to 0$ are replaced by A3' and $nh_n^{d+4} \to \Theta$, respectively, we have the following result. Suppose that for $(x, y) \in B(x_0)$, $f(x)$ and $F(y|x)$ have bounded second partial derivatives, where $B(x) = U_x \times U_{\theta(x)}$ and $U_{\theta(x)} := U_{\theta(x)}(\xi) = \{y : |y - \theta(x)| \leq \xi\}$. Then

$$(nh_n^d)^{1/2} (\theta_n(x_0) - \theta(x_0)) \to \mathcal{N}(-\mu, \sigma^2),$$

(2.15)

where $\mu = \Gamma(\Theta, x_0, d)\kappa$, $\kappa = \sum_{j=1}^d \left( \frac{\partial F(\theta(x_0)|x)}{\partial x_j} \right) \frac{2}{f(x)} \frac{\partial f(x)}{\partial x_j} \frac{\partial f(\theta(x_0)|x)}{\partial x_j} \big|_{x=x_0}$, $x = (x_1, \ldots, x_d)$,

$\Gamma(\Theta, x_0, d)$ is a constant depending on $(\Theta, x_0, d)$ and $\sigma^2$ is defined as above. Since the proof is analogous to that of Theorem 1 with some modifications, details are omitted.

Next, we shall give an example such that A3' holds.

**Example 2.** Suppose that $h_n \to 0$, $nh_n^{d+2\beta} \to \infty$ for some $0 < \beta < 1$ and $a(i) \leq Ci^{-v}$ for $v > 1 + (d + 6 + \tau)/(\beta - \alpha)$, $0 < \alpha < \beta$ and $i \geq 1$, where $a(i)$ is defined in Example 1. A3'(i) and (ii) are easy to verify. Now, let $s_n \sim h_n^{-\beta + \alpha}$. Noting that $v > 1 + (d + 6 + \tau)/(\beta - \alpha)$, we have

$$\frac{1}{h_n^{d+6+\tau}} \sum_{p=1}^d \sum_{q=1}^d \sum_{i=s_n}^\infty |\text{Cov}(X_{1p}, X_{iq})| \leq \frac{C}{h_n^{d+6+\tau}} \sum_{i=s_n}^\infty i^{-v} \leq C h_n^{(\beta - \alpha)(1-v) - d - 6 - \tau} = o(1).$$

Hence, (2.12) is satisfied. (2.13) and (2.14) can be proved by analogous argument.

**Remark 5.** From Examples 1 and 2, we can find that there is a tradeoff between the conditions on $h_n$ and the summability conditions on the covariance functions.
3. Some useful lemmas

In this section, we present some lemmas which are useful in the proof of Theorem 1. First, we give Lemma 1 in Bulinski [3], which is a generalization of Lemma 3.1 in Birkel [2].

Lemma 1. Let \{X_j, j \in I\} be a finite collection of associated random variables. Let \(I_1\) and \(I_2\) be two disjoint subsets of \(I\) and let \(H_i\) be functions on \(R^{I_i}\), \(i = 1, 2\), with bounded first order partial left and right derivatives and that for any cube in \(R^{I_i}\) and for \(j \in I_1\) and \(k \in I_2\), there are only finite number of points at which the left and right partial derivatives are not equal. Then

\[
|\text{Cov}(H_1(X_j, j \in I_1), H_2(X_k, k \in I_2))| \\
\leq \sum_{j \in I_1} \sum_{k \in I_2} \left\| \frac{\partial H_1}{\partial x_j} \right\|_\infty \left\| \frac{\partial H_2}{\partial x_k} \right\|_\infty |\text{Cov}(X_j, X_k)|,
\]

(3.1)

where \(\left\| \frac{\partial H_i}{\partial x_j} \right\|_\infty = \max\{\left\| \frac{\partial^+ H_i}{\partial x_j} \right\|_\infty, \left\| \frac{\partial^- H_i}{\partial x_j} \right\|_\infty\}\).

The following lemma, which is of independent interest, is critical in our proof.

Lemma 2. If Conditions A1(i), A2, A3(i) and (2.4) are satisfied and \(nh_n^{d+2} \to 0\), we have

\[
(nh_n^d)^{-1} N_n - L(1)f(x_0) \xrightarrow{p} 0,
\]

(3.2)

where \(N_n\) is defined as before.

Proof. Noting that \(nh_n^d \to \infty\), there exists a sequence \(\phi(n)\) such that \(\phi(n) \to \infty\) and \(\frac{nh_n^d}{\phi(n)} \to \infty\). Therefore, in order to prove (3.2), it suffices to show that

\[
P\{|(nh_n^d)^{-1} N_n - L(1)f(x_0)| > \eta_n L(1)f(x_0)\} \to 0,
\]

(3.3)

where \(\eta_n = (nh_n^d)^{-\frac{1}{2}} \phi^\frac{1}{2}(n)\). It is easy to verify that \(h_n = o(\eta_n)\) by noting \(nh_n^{d+2} \to 0\) and \(\phi(n) \to \infty\). Note that

\[
P\{|(nh_n^d)^{-1} N_n - L(1)f(x_0)| > \eta_n L(1)f(x_0)\}
\]

\[
= P\{(nh_n^d)^{-1} N_n > L(1)f(x_0) + \eta_n L(1)f(x_0)\}
\]

\[
+ P\{(nh_n^d)^{-1} N_n < L(1)f(x_0) - \eta_n L(1)f(x_0)\}.
\]

Our aim is to show that both of the above two terms tend to zero. We only prove the conclusion for the first term since the proof for the second term is similar.

Recall that \(N_n = \sum_{i=1}^{\infty} K_{ni}(x_0)\), where \(K_{ni}(x_0) = I(\|X_i - x_0\| \leq h_n)\). The first term equals

\[
P\left\{\frac{1}{n} \sum_{i=1}^{n} (K_{ni}(x_0) - EK_{ni}(x_0)) > h_n^d L(1)f(x_0)(1 + \eta_n) - EK_{n1}(x_0)\right\}.
\]

(3.4)
By Taylor’s expansion, it follows that

$$f(x) = f(x_0) + \sum_{i=1}^{d} \frac{\partial f(x)}{\partial x_i} \bigg|_{x=x_0}(x_i - x_{i0}),$$

(3.5)

where \(x = (x_1, x_2, \ldots, x_d)\), \(x_0 = (x_{10}, x_{20}, \ldots, x_{d0})\) and \(\phi\) is some real-valued vector between \(x\) and \(x_0\), i.e., \(\phi = x_0 + \delta(x - x_0)\) for some \(0 < \delta < 1\). By (3.5), Condition A1(i) and \(h_n = o(\eta_n)\), we have for \(n\) large enough,

$$EK_{n1}(x_0) = f(x_0)L(h_n) + \sum_{i=1}^{d} \frac{\partial f(x)}{\partial x_i} \bigg|_{x=x_0} \int_{\|x-x_0\| \leq h_n} (x_i - x_{i0}) \, dx_1 \cdots dx_d$$

$$= f(x_0)L(h_n) + O(h_nL(h_n))$$

$$\leq L(1)f(x_0)h_n^d + \frac{1}{2} \eta_n L(1)f(x_0)h_n^d$$

$$= L(1)f(x_0)h_n^d(1 + \eta_n) - \frac{1}{2} \eta_n L(1)f(x_0)h_n^d.$$

Hence, in order to prove that (3.4) tends to zero, it suffices to prove that

$$P \left\{ \frac{1}{n} \sum_{i=1}^{n} (K_{ni}(x_0) - EK_{ni}(x_0)) > \frac{1}{2} \eta_n L(1)f(x_0)h_n^d \right\} \to 0.$$  (3.6)

Define \(Z_{ni} = \frac{1}{n}(K_{ni}(x_0) - EK_{ni}(x_0))\) and \(g_1(z)\) as follows:

$$g_1(z) = \begin{cases} 
1, & z < h_n - \frac{h_n^2}{2}, \\
\frac{1}{2} \sin \left( \frac{\pi}{h_n^2} z + \frac{3}{2} \pi - \frac{\pi}{h_n^2} \right) + \frac{1}{2}, & h_n - \frac{h_n^2}{2} \leq z < h_n, \\
0, & z \geq h_n.
\end{cases}$$

Let \(Z_{ni}^\prime = \frac{1}{n}[g_1(\|X_i - x_0\|) - E g_1(\|X_i - x_0\|)]\) and \(Z_{ni}'' = Z_{ni} - Z_{ni}^\prime\). Next, we shall give a bound for \(E(\sum_{i=1}^{n} Z_{ni})^2\). It is easy to check that

$$E \left( \sum_{i=1}^{n} Z_{ni} \right)^2 \leq 2E \left( \sum_{i=1}^{n} Z_{ni}' \right)^2 + 2E \left( \sum_{i=1}^{n} Z_{ni}'' \right)^2.$$  (3.7)

First, we give a bound for \(E(\sum_{i=1}^{n} Z_{ni}')^2\). By stationarity, we have

$$E \left( \sum_{i=1}^{n} Z_{ni}' \right)^2 = \sum_{i=1}^{n} E(Z_{ni}')^2 + 2\sum_{i=2}^{n} (n - i + 1) \text{Cov}(Z_{ni}', Z_{ni}').$$  (3.8)

Noting that \(E(Z_{ni}')^2 \leq C \frac{h_n^d}{n^2}\), we have

$$\sum_{i=1}^{n} E(Z_{ni}')^2 \leq C \frac{h_n^d}{n}.$$  (3.9)
Moreover,
\[
\sum_{i=2}^{n} (n - i + 1) \text{Cov}(Z_{n1}', Z_{ni}') \\
= \sum_{i=2}^{n} (n - i + 1) \text{Cov}(Z_{n1}', Z_{ni}') + \sum_{i=c_n+1}^{n} (n - i + 1) \text{Cov}(Z_{n1}', Z_{ni}') \\
=: I_{n1} + I_{n2},
\]
where \( c_n = [h_n^{d+\gamma}], 0 < \gamma < d - \frac{d+4}{\delta_1}. \) By A2 and the definition of \( g_1(\cdot), \) we have
\[
I_{n1} = \frac{1}{n^2} \sum_{i=2}^{c_n} (n - i + 1) \int \cdots \int_{\|x\| \leq h_n, \|y\| \leq h_n} g_1(\|x\|) g_1(\|y\|) \\
\times (f_{1,i}(x, y) - f(x) f(y)) \, dx \, dy \\
\leq M_3 \frac{c_n}{n} \int \cdots \int_{\|x\| \leq h_n, \|y\| \leq h_n} g_1(\|x\|) g_1(\|y\|) \, dx \, dy \\
= O \left( \frac{c_n h_n^{2d}}{n} \right) = O \left( \frac{h_n^{d+\gamma}}{n} \right) = O \left( \frac{h_n^d}{n} \right).
\]
On the other hand, by Lemma 1 and (2.4),
\[
I_{n2} \leq \frac{1}{n^2} \sum_{p=1}^{d} \sum_{q=1}^{d} \sum_{i=c_n+1}^{n} h_n^{-4} \text{Cov}(X_{1p}, X_{iq}) \\
\leq \frac{h_n^d}{n} \sum_{p=1}^{d} \sum_{q=1}^{d} \sum_{i=c_n+1}^{n} i^{\delta_1} \text{Cov}(X_{1p}, X_{iq}) \\
\leq o \left( \frac{h_n^d}{n} \right).
\]
Hence,
\[
\sum_{i=2}^{n} (n - i + 1) \text{Cov}(Z_{n1}', Z_{ni}') \leq o \left( \frac{h_n^d}{n} \right). \tag{3.10}
\]
In view of (3.8)–(3.10), we have
\[
E \left( \sum_{i=1}^{n} Z_{ni}' \right)^2 \leq C \frac{h_n^d}{n}. \tag{3.11}
\]
By the similar method, we have
\[
E \left( \sum_{i=1}^{n} Z_{ni}'' \right)^2 = \sum_{i=1}^{n} E(Z_{ni}'')^2 + 2 \sum_{i=2}^{n} (n - i + 1) \text{Cov}(Z_{n1}'', Z_{ni}'') \\
\leq C \left( \frac{h_n^{d+1}}{n} + h_n^{2d+2} \right) \\
= o \left( \frac{h_n^d}{n} \right). \tag{3.12}
\]
By (3.7), (3.11) and (3.12), we have
\[
E \left( \sum_{i=1}^{n} Z_{ni} \right)^2 \leq C \frac{h_n^d}{n}.
\] (3.13)

Letting \( \kappa_n = \frac{1}{2} L(1) f(x_0) h_n^d \eta_n \), by the Markov inequality, we have
\[
P \left\{ \sum_{i=1}^{n} Z_{ni} > \kappa_n \right\} \leq E \left( \sum_{i=1}^{n} Z_{ni} \right)^2 \frac{\kappa_n^2}{\kappa_n^2} \leq \frac{C h_n^d}{n \kappa_n^2} \leq \frac{C}{\phi(n)} \to 0.
\] (3.14)

Therefore, (3.3) holds. \( \square \)

Before stating the following lemma, we give some notations. Let
\[
T_{ni} = \left( \frac{4}{f(x_0) L(h_n)} \right)^{\frac{1}{2}} K_{ni}(x_0) \left\{ I(Y_i > \theta(x_0) + \varepsilon) - [1 - F(\theta(x_0) + \varepsilon | X_i)] \right\},
\] (3.15)

where \( \varepsilon = \sigma t(nh_n^d)^{-\frac{1}{2}} \), \( t \in R \), \( i = 1, 2, \ldots, n \) and define \( g_2(z) \) as follows:
\[
g_2(z) = \begin{cases} 0, & z \leq \theta(x_0) + \varepsilon, \\ \frac{z - (\theta(x_0) + \varepsilon)}{h_n}, & \theta(x_0) + \varepsilon < z \leq \theta(x_0) + \varepsilon + h_n, \\ 1, & z > \theta(x_0) + \varepsilon + h_n. \end{cases}
\]

Lemma 3 shall give the estimation of the variance of \( S_n = \sum_{i=1}^{n} T_{ni} \). Lemma 1, which is crucial in the proof of Lemma 3, cannot be applied immediately since \( T_{ni} \) is discontinuous. Hence, some transformations on \( T_{ni} \) are needed in the proof of Lemma 3.

**Lemma 3.** Define \( \sigma_n^2 := ES_n^2 \), then, under the conditions of Theorem 1,
\[
\sigma_n^2 \sim n.
\] (3.16)

**Proof.** Let \( T_{ni}' = \left( \frac{4}{f(x_0) L(h_n)} \right)^{\frac{1}{2}} g_1(\|X_i - x_0\|) \{g_2(Y_i) - [1 - F(\theta(x_0) + \varepsilon | X_i)]\} \) and \( T_{ni}'' = T_{ni} - T_{ni}' \), where \( g_1(x) \) and \( g_2(x) \) are defined as before. It is easy to check that
\[
ES_n^2 = E \left( \sum_{i=1}^{n} T_{ni}' \right)^2 + E \left( \sum_{i=1}^{n} T_{ni}'' \right)^2 + 2E \left( \sum_{i=1}^{n} T_{ni}' \right) \left( \sum_{i=1}^{n} T_{ni}'' \right).
\] (3.17)

By stationarity, we have
\[
E \left( \sum_{i=1}^{n} T_{ni}' \right)^2 = n E(T_{n1}')^2 + 2 \sum_{i=2}^{n} (n - i + 1) \text{Cov}(T_{n1}', T_{ni}')
\]
\[
= n E(T_{n1}')^2 + 2 \sum_{i=2}^{d_n} (n - i + 1) \text{Cov}(T_{n1}', T_{ni}')
\]
\[+2 \sum_{i=d_n+1}^{n} (n-i+1) \text{Cov}(T'_{ni}, T'_{ni}) =: J_{n1} + J_{n2} + J_{n3},\]

where \(d_n\), specified later, is a sequence of integers. Following the proof of Zhou and Liang [22], we have

\[J_{n1} = n + o(n). \quad (3.18)\]

For \(2 \leq i \leq d_n\), we have

\[|\text{Cov}(T'_{n1}, T'_{ni})| \leq C h_n^{-d} \int \ldots \int_{\|x\| \leq h_n, \|y\| \leq h_n} g_1(\|x\|)g_1(\|y\|) \, dx \, dy \leq C h_n^{-d} h_n^{2d} = O(h_n^d).\]

Hence, we have

\[J_{n2} \leq O(nd_nh_n^d). \quad (3.19)\]

Denote \(\Gamma(x, y) = g_1(\|x - x_0\|)(g_2(y) - (1 - F(\theta(x_0) + \varepsilon|x)))\), where \(x = (x_1, x_2, \ldots, x_d)\). By an elementary calculation, we have

\[\left\| \frac{\partial \Gamma}{\partial y} \right\|_{\infty} \leq C \frac{1}{h_n}, \quad \left\| \frac{\partial \Gamma}{\partial x_i} \right\|_{\infty} \leq C \frac{1}{h_n^2}, \quad i = 1, 2, \ldots, d. \quad (3.20)\]

By Lemma 1 and (3.20), we have

\[J_{n3} \leq Cn \left\{ \sum_{p=1}^{d} \sum_{q=1}^{d} \sum_{i=d_n+1}^{n} h_n^{-(d+4)}|\text{Cov}(X_{1p}, X_{iq})| + \sum_{p=1}^{d} \sum_{i=d_n+1}^{n} h_n^{-(d+3)}|\text{Cov}(Y_1, X_{ip})| \right.\]
\[+ \left. \sum_{p=1}^{d} \sum_{i=d_n+1}^{n} h_n^{-(d+3)}|\text{Cov}(X_{1p}, Y_i)| + \sum_{i=d_n+1}^{n} h_n^{-(d+2)}|\text{Cov}(Y_1, Y_i)| \right\}.\]

Taking \(d_n = [h_n^{d+\gamma}]\), where \(0 < \gamma < \min\{d - \frac{d+4}{\delta_1}, d - \frac{d+3}{\delta_2}, d - \frac{d+2}{\delta_3}\}\), by (2.4)–(2.6),

\[J_{n3} \leq Cn \sum_{p=1}^{d} \sum_{q=1}^{d} \sum_{i=[h_n^{\delta_1}]}^{n} i^{\delta_1} |\text{Cov}(X_{1p}, X_{iq})| \]
\[+ Cn \sum_{p=1}^{d} \sum_{i=[h_n^{\delta_2}]}^{n} i^{\delta_2} |\text{Cov}(Y_1, X_{ip})| \]
\[+ Cn \sum_{p=1}^{d} \sum_{i=[h_n^{\delta_3}]}^{n} i^{\delta_2} |\text{Cov}(X_{1p}, Y_i)| + Cn \sum_{i=[h_n^{\delta_5}]}^{n} i^{\delta_3} |\text{Cov}(Y_1, Y_i)| \]

\[= o(n). \quad (3.21)\]
On the other hand, by the definition of $d_n$ and (3.19),

$$J_{n2} = O(nh_n^\frac{1}{2}) = o(n).$$

(3.22)

Therefore, by (3.18), (3.21) and (3.22), we have

$$E \left( \sum_{i=1}^{n} T'_{ni} \right)^2 = n + o(n).$$

(3.23)

Analogously,

$$E \left( \sum_{i=1}^{n} T''_{ni} \right)^2 = o(n).$$

(3.24)

$$E \left( \sum_{i=1}^{n} T'_{ni} \right) \left( \sum_{i=1}^{n} T''_{ni} \right) \leq E \left[ \left( \sum_{i=1}^{n} T'_{ni} \right)^2 \right]^{\frac{1}{2}} E \left[ \left( \sum_{i=1}^{n} T''_{ni} \right)^2 \right]^{\frac{1}{2}} = o(n).$$

(3.25)

In view of (3.23)–(3.25), (3.16) holds. □

4. Proof of Theorem 1

Recall that $\varepsilon = \sigma t (nh_n^d)^{-\frac{1}{2}}$. By simplifying, we have

$$P\{ (nh_n^d)^{\frac{1}{2}} (\theta_n(x_0) - \theta(x_0)) < t \sigma \}$$

$$= P \left\{ \frac{1}{2} < F_n(\theta(x_0) + \varepsilon|x_0) \right\}$$

$$= P \left\{ \frac{1}{N_n} \sum_{i \in J_n} I(Y_i > \theta(x_0) + \varepsilon) < \frac{1}{2} \right\}$$

$$= P \left\{ \frac{1}{N_n} \sum_{i \in J_n} (I(Y_i > \theta(x_0) + \varepsilon) - [1 - F(\theta(x_0) + \varepsilon|X_i)]) \right\}$$

$$< \frac{1}{N_n} \sum_{i \in J_n} F(\theta(x_0) + \varepsilon|X_i) - \frac{1}{2} \right\}. \quad (4.1)$$

Noting that $nh_n^{d+2} \to 0$, we have

$$\frac{1}{N_n} \sum_{i \in J_n} F(\theta(x_0) + \varepsilon|X_i) - \frac{1}{2}$$

$$= \frac{1}{N_n} \sum_{i \in J_n} \{ F(\theta(x_0) + \varepsilon|X_i) - F(\theta(x_0) + \varepsilon|x_0) \} + \left\{ F(\theta(x_0) + \varepsilon|x_0) - \frac{1}{2} \right\}$$

$$= o(h_n) + \varepsilon f(\theta(x_0)|x_0) + o((nh_n^d)^{-\frac{1}{2}}) = \varepsilon f(\theta(x_0)|x_0) + o((nh_n^d)^{-\frac{1}{2}}).$$
Hence, the last term of (4.1) equals
\[
P \left\{ \frac{1}{N_n} \sum_{i \in J_n} \left( I(Y_i > \theta(x_0) + \varepsilon) - [1 - F(\theta(x_0) + \varepsilon | X_i)] \right) < \varepsilon f(\theta(x_0)|x_0) + o((nh_n^d)^{-\frac{1}{2}}) \right\}.
\]

By Lemma 2, we have
\[
\frac{1}{nhd_n^d} \sum_{i=1}^{n} J_n p \rightarrow L(1) f(x_0).
\]

Therefore, in order to show that (2.10) holds, it suffices to prove that
\[
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} T_{ni} \overset{d}{\rightarrow} N(0, 1),
\]
where $T_{ni}$ is defined as in Lemma 3. In order to apply Lemma 1, we shall make some transformations on $T_{ni}$. Define $T_{ni}'$ and $T_{ni}''$ as in Lemma 3. Hence, we have
\[
\sum_{i=1}^{n} T_{ni} = \sum_{i=1}^{n} T_{ni}' + \sum_{i=1}^{n} T_{ni}''.
\]

Recall that
\[
E \left( \sum_{i=1}^{n} T_{ni}'' \right)^2 = o(n).
\]

Hence, in order to prove (4.4), it is enough to show that
\[
\frac{1}{\sqrt{n}} \sum_{i=1}^{n} T_{ni}' \overset{d}{\rightarrow} N(0, 1).
\]

Next, we employ the big-block and small-block procedure to prove (4.7). Partition the set \{1, 2, \ldots, n\} into $2k_n + 1$ subsets with large blocks of size $u_n := \left\lfloor \left(\frac{nh_n^d}{q_n} \right)^{\frac{1}{2}} \right\rfloor$ and small blocks of size $s_n$, where $s_n$ is defined in Condition A3(iii), $q_n$ is a sequence of positive numbers satisfying $q_n s_n = o((nh_n^d)^{\frac{1}{2}})$ and $k_n = \left\lfloor \frac{n}{u_n + s_n} \right\rfloor$. Define the following random variables for $1 \leq j \leq k_n$:
\[
U_j = \sum_{i=(j-1)(u_n+s_n)+u_n}^{(j-1)(u_n+s_n)+u_n} T_{ni}', \quad V_j = \sum_{i=(j-1)(u_n+s_n)+u_n+1}^{(j-1)(u_n+s_n)+u_n+1} T_{ni}',
\]
\[
U_{k_n+1} = \sum_{i=k_n(u_n+s_n)+1}^{n} T_{ni}'.
\]

Write
\[
S_n = \sum_{i=1}^{n} T_{ni}' = \sum_{j=1}^{k_n} U_j + \sum_{j=1}^{k_n} V_j + U_{k_n+1} =: S_{n1} + S_{n2} + S_{n3}.
\]
By Condition A3(iii), simple algebra shows that as \( n \to \infty \),
\[
\frac{s_n}{u_n} \to 0, \quad \frac{u_n}{n} \to 0, \quad \frac{u_n}{(nh_n^d)^2} \to 0.
\]
\[(4.10)\]

Next, we shall calculate the bound for \( E(S_n^2) \). By the definition of \( S_n \),
\[
E(S_n^2) = E\left( \sum_{j=1}^{k_n} V_j \right)^2
\]
\[
= \sum_{j=1}^{k_n} EV_j^2 + 2 \sum_{j=2}^{k_n} (k_n - j + 1) \text{Cov}(V_1, V_j)
\]
\[
=: H_1 + H_2.
\]

Following the proof of Lemma 3, \( EV_j^2 \sim s_n \). Hence, we have
\[
H_1 \sim k_ns_n \sim \frac{nS_n}{u_n + s_n} \sim \frac{nS_n}{u_n} = o(n).
\]
\[(4.11)\]

On the other hand, let \( \rho_j = (j - 1)(u_n + s_n), 1 \leq j \leq k_n \),
\[
H_2 = 2 \sum_{j=2}^{k_n} (k_n - j + 1) \sum_{i_1=\rho_1+u_n+1}^{\rho_2} \sum_{i_2=\rho_j+u_n+1}^{\rho_{j+1}} |\text{Cov}(T_{ni_1}^\prime, T_{ni_2}^\prime)|
\]
\[
\leq 2 \sum_{j=\rho_n+1}^{\infty} k_nu_n |\text{Cov}(T_{n1}^\prime, T_{nj}^\prime)|
\]
\[
\leq Cn \sum_{j=\rho_n+1}^{\infty} |\text{Cov}(T_{n1}^\prime, T_{nj}^\prime)| = o(n).
\]
\[(4.12)\]

In view of (4.11) and (4.12), we have
\[
E(S_n^2) = o(n).
\]
\[(4.13)\]

By a similar argument,
\[
E(S_n^3) = o(n).
\]
\[(4.14)\]

Eqs. (4.13) and (4.14) imply that \( \frac{1}{\sqrt{n}}(S_{n2} + S_{n3}) \) are asymptotically negligible. Therefore, it is enough to establish asymptotic normality for \( \frac{1}{\sqrt{n}}S_{n1} \). Next, we shall show that the summands \( U_j \) in \( S_{n1} \) are asymptotically independent, i.e.,
\[
\left| E \exp \left\{ it \frac{1}{\sqrt{n}} S_{n1} \right\} - \prod_{j=1}^{k_n} E \exp \left\{ it \frac{1}{\sqrt{n}} U_j \right\} \right| \to 0.
\]
\[(4.15)\]

The proof of (4.15) is far more involved than that for \( \alpha \)-mixing processes. We shall prove it with the help of Lemma 1.
\[
\left| E \exp \left\{ it \frac{1}{\sqrt{n}} S_{n1} \right\} - \prod_{j=1}^{k_n} E \exp \left\{ it \frac{1}{\sqrt{n}} U_j \right\} \right|
\]
By induction, we obtain

\[
\begin{align*}
&\leq \left| E \exp \left\{ \frac{i}{\sqrt{n}} S_{n1} \right\} - \prod_{j=1}^{k_n} E \exp \left\{ \frac{i}{\sqrt{n}} U_j \right\} \right| \\
&\leq \sum_{l=1}^{k_n-1} \left| \text{Cov} \left( \exp \left\{ \frac{i}{\sqrt{n}} \sum_{j=1}^{k_n-l} U_j \right\}, \exp \left\{ \frac{i}{\sqrt{n}} U_{k_n-l+1} \right\} \right| \right|.
\end{align*}
\]

Write \( \Gamma_1(x, y) = \exp \left\{ \frac{i}{\sqrt{n}} \left( \frac{4}{f(x) f(y)} \right)^{1/2} \Gamma(x, y) \right\} \), where \( \Gamma(x, y) \) is defined as in the proof of Lemma 3 and \( x = (x_1, x_2, \ldots, x_d) \). It is easy to check that

\[
\left\| \frac{\partial \Gamma_1(x, y)}{\partial y} \right\| \leq C \frac{t}{(nh_n^{d+2})^{1/2}}, \quad \left\| \frac{\partial \Gamma_1(x, y)}{\partial x_i} \right\| \leq C \frac{t}{(nh_n^{d+4})^{1/2}}, \quad i=1, 2, \ldots, d. \quad (4.16)
\]

Denote the \( j \)th large block by \( \Delta_j = \{ i : (j-1)(u_n + s_n) + 1 \leq i \leq (j-1)(u_n + s_n) + u_n \} \), \( j = 1, 2, \ldots, k_n \). Then, with the help of Lemma 1 and (4.16), we have

\[
\begin{align*}
&\leq \frac{Ct^2}{nh_n^{d+2}} \sum_{l=1}^{k_n-1} \sum_{i \in \Delta_l} \sum_{j \in \Delta_{k_n-l+1}} |\text{Cov}(Y_i, Y_j)| \\
&+ \frac{Ct^2}{nh_n^{d+3}} \sum_{l=1}^{k_n-1} \sum_{p=1}^{d} \sum_{i \in \Delta_l} \sum_{j \in \Delta_{k_n-l+1}} |\text{Cov}(Y_i, X_{jp})| \\
&+ \frac{Ct^2}{nh_n^{d+3}} \sum_{l=1}^{k_n-1} \sum_{p=1}^{d} \sum_{i \in \Delta_l} \sum_{j \in \Delta_{k_n-l+1}} |\text{Cov}(X_{ip}, Y_j)| \\
&+ \frac{Ct^2}{nh_n^{d+4}} \sum_{l=1}^{k_n-1} \sum_{p=1}^{d} \sum_{q=1}^{d} \sum_{i \in \Delta_l} \sum_{j \in \Delta_{k_n-l+1}} |\text{Cov}(X_{ip}, X_{jq})| \\
&=: \Pi_1 + \Pi_2 + \Pi_3 + \Pi_4.
\end{align*}
\]
For $\Pi_1$, by stationarity, we have

$$\Pi_1 \leq \frac{C_2}{nh_n^{d+2}} \left\{ (k_n - 1) \sum_{i \in \Delta_1} \sum_{j \in \Delta_2} |\text{Cov}(Y_i, Y_j)| + (k_n - 2) \sum_{i \in \Delta_1} \sum_{j \in \Delta_3} |\text{Cov}(Y_i, Y_j)| \right. $$

$$+ \cdots + \sum_{i \in \Delta_1} \sum_{j \in \Delta_{kn}} |\text{Cov}(Y_i, Y_j)| \right\}$$

$$= \frac{C_2}{nh_n^{d+2}} \sum_{l=1}^{k_n-1} (k_n - l) \sum_{i \in \Delta_1} \sum_{j \in \Delta_{l+1}} |\text{Cov}(Y_i, Y_j)|. \quad (4.17)$$

Once again, by stationarity,

$$\sum_{i \in \Delta_1} \sum_{j \in \Delta_{l+1}} |\text{Cov}(Y_i, Y_j)|$$

$$= u_n \text{Cov}(Y_1, Y_{(j-1)(u_n+s_n)+1}) + (u_n - 1)\text{Cov}(Y_1, Y_{(j-1)(u_n+s_n)+2})$$

$$+ \text{Cov}(Y_1, Y_{(j-1)(u_n+s_n)}) + \cdots + \text{Cov}(Y_1, Y_{(j-1)(u_n+s_n)+u_n})$$

$$+ \text{Cov}(Y_1, Y_{(j-1)(u_n+s_n)-u_n+1}).$$

By means of the above expression, (4.17) becomes

$$\Pi_1 \leq \frac{C_2}{nh_n^{d+2}} \sum_{l=1}^{k_n-1} (k_n - l) \left\{ u_n \text{Cov}(Y_1, X_{i,p}) \right\} \to 0,$$

by Condition A3(iii) and $\frac{k_n u_n}{n} \to 1$. Similar arguments show that

$$\Pi_2 \leq \frac{C_2}{h_n^{d+3}} \sum_{p=1}^{d} \sum_{i=s_n}^{\infty} |\text{Cov}(Y_1, X_{i,p})| \to 0, \quad (4.18)$$

$$\Pi_3 \leq \frac{C_2}{h_n^{d+3}} \sum_{p=1}^{d} \sum_{i=s_n}^{\infty} |\text{Cov}(X_{1,p}, Y_i)| \to 0, \quad (4.19)$$

$$\Pi_4 \leq \frac{C_2}{h_n^{d+4}} \sum_{p=1}^{d} \sum_{q=1}^{d} \sum_{i=s_n}^{\infty} |\text{Cov}(X_{1,p}, X_{i,q})| \to 0. \quad (4.20)$$

Therefore, (4.15) holds. Now, we construct a new sequence $\{U'_j\}$ which is i.i.d. with $\{U_j\}$. It is easy to check that

$$\frac{1}{n} E \left( \sum_{j=1}^{k_n} U'_j \right)^2 \leq \frac{1}{n} \sum_{j=1}^{k_n} E U'_j^2 \sim \frac{k_n u_n}{n} \to 1. \quad (4.21)$$
Since $\max_{1 \leq j \leq kn} |U_j| \leq Ch_n^{-\frac{4}{7}} u_n$, which together with (4.10) implies that $\{|U_j| \geq \varepsilon \sqrt{n}\}$ is an empty set when $n$ is large enough, we have

$$\frac{1}{n} \sum_{j=1}^{kn} E\{U_j' U_j | |U_j'| \geq \varepsilon \sqrt{n}\} \to 0.$$  \hfill (4.22)

Therefore, by (4.13)–(4.15), (4.21) and (4.22), (4.7) holds. \hfill \square

Acknowledgments

The authors are grateful to the editor and the two anonymous referees for their valuable advice and encouragement. Thanks also go to Dr. Jia Chen for some constructive comments. This research was supported by National Natural Science Foundation of China (Grant no. 10571159) and Specialized Research Fund for the Doctor Program of Higher Education (Grant no. 2002335090).

References