

NONPARAMETRIC DENSITY ESTIMATORS BASED ON NONSTATIONARY ABSOLUTELY REGULAR RANDOM SEQUENCES

MICHEL HAREL and MADAN L. PURI¹

*I.U.F.M. du Limousin
 U.R.A. 745 C.N.R.S., Toulouse, France
 and Indiana University, Dept. of Mathematics
 USA*

(Received May, 1995; Revised November, 1995)

ABSTRACT

In this paper, the central limit theorems for the density estimator and for the integrated square error are proved for the case when the underlying sequence of random variables is nonstationary. Applications to Markov processes and ARMA processes are provided.

Key words: Density Estimators, Nonstationary Absolutely Regular Random Sequences, Strong Mixing, φ -Mixing, Markov Processes, ARMA Processes.

AMS (MOS) subject classifications: 62G05, 60F05.

1. Introduction

Let $\{\mathbf{X}_i = (X_i^{(1)}, \dots, X_i^{(p)}), i \geq 1\}$ be a sequence of random variables with continuous d.f.'s (distribution functions) $F_i(\mathbf{x}), i \geq 1, \mathbf{x} \in \mathbb{R}^p$.

Assume that the processes satisfies the absolute regularity condition

$$\max_{j \geq 1} E \left\{ \sup_{A \in \sigma(\mathbf{X}_i, i \geq j+m)} |P(A | \sigma(\mathbf{X}_i, 1 \leq i \leq j)) - P(A)| \right\} = \beta(m) \downarrow 0 \text{ as } m \rightarrow \infty. \quad (1.1)$$

Here $\sigma(\mathbf{X}_i, 1 \leq i \leq j)$ and $\sigma(\mathbf{X}_i, i \geq j+m)$ are the σ -fields generated by $(\mathbf{X}_1, \dots, \mathbf{X}_j)$ and $(\mathbf{X}_{j+m}, \mathbf{X}_{j+m+1}, \dots)$ respectively. Also recall that the sequence $\{\mathbf{X}_i\}$ satisfies the strong mixing condition if

$$\max_{j \geq 1} [\sup \{ |P(A \cap B) - P(A)P(B)| ; A \in \sigma(\mathbf{X}_i, 1 \leq i \leq j), B \in \sigma(\mathbf{X}_i, i \geq j+m) \}] = \alpha(m) \downarrow 0 \text{ as } m \rightarrow \infty$$

and it satisfies the φ -mixing condition if

$$\max_{j \geq 1} [\sup \{ |P(A | B) - P(A)| ; B \in \sigma(\mathbf{X}_i, 1 \leq i \leq j), A \in \sigma(\mathbf{X}_i, i \geq j+m) \}] = \varphi(m) \downarrow 0 \text{ as } m \rightarrow \infty.$$

Since $\alpha(m) \leq \beta(m) \leq \varphi(m)$, it follows that if $\{\mathbf{X}_i\}$ is absolutely regular, then it is also strong mixing and if $\{\mathbf{X}_i\}$ is φ -mixing, it is also absolutely regular.

Suppose that the distribution function F_n has a density f_n , and F_n converges to a distribu-

¹Research supported by the Office of Naval Research Contract N00014-85-K-0648.

tion function F which admits a density f , and let f_n^* be an estimator of f based on $\mathbf{X}_1, \dots, \mathbf{X}_n$ defined below in (2.2).

In this paper, we establish the central limit theorems for the estimator f_n^* and for the integrated mean square error (I.M.S.E.) I_n defined by

$$I_n = \int_{\mathbb{R}^p} \{f_n^*(\mathbf{x}) - f(\mathbf{x})\}^2 d\mathbf{x}. \tag{1.2}$$

An additional asymptotic property of the I.M.S.E. is also studied in (2.5).

Several authors have proved central limit theorems for f_n^* and I_n when $\{\mathbf{X}_n, n \geq 1\}$ is a sequence of independent and identically distributed random variables (see, e.g., Csörgo and Révész [2], Rosenblatt [11] and Hall [8]). Later Takahata and Yoshihara [15] proved the central limit theorem for I_n when $\{\mathbf{X}_n, n \geq 1\}$ is an absolutely regular strictly stationary sequence. See also Tran [16, 17] and Roussas and Tran [13], and for a general theory, we refer to an excellent monograph of Devroyes and Györfi [5]. We may also mention the results of Roussas [12] for stationary Markov processes which are Doeblin recurrent and also the results of Doukhan and Ghindès [6] on the estimation in a Markov chain.

In this paper using some of the ideas of Takahata and Yoshihara [15], we prove the central limit theorem for f_n^* and I_n when the sequence is not stationary.

In section 5, we give applications of our results to Markov processes and ARMA processes for which the initial measure is not necessary to be the invariant measure. Under suitable conditions, any initial measure converges to the invariant measure. We estimate the density of this invariant measure by the estimator f_n^* defined below in (2.2).

2. Asymptotically Unbiased Estimation of the Invariant Density

Let $K(\mathbf{x})$ be a bounded, non-negative function on \mathbb{R}^p such that

$$\int K(\mathbf{x})d\mathbf{x} = 1, \quad \int x^{(i)}K(\mathbf{x})d\mathbf{x} = 0 \text{ and } \int x^{(i)}x^{(j)}K(\mathbf{x})d\mathbf{x} = 2\tau\delta_{ij}, \quad 1 \leq i, j \leq p, \tag{2.1}$$

and $\lim_{|\mathbf{x}| \rightarrow \infty} |K(\mathbf{x})| = 0.$

Here $\mathbf{x} = (x^{(1)}, \dots, x^{(p)})$, $d\mathbf{x} = dx^{(1)} \dots dx^{(p)}$, $|\mathbf{x}| = \sup_{1 \leq j \leq p} |x^{(j)}|$, τ is a constant which does not depend on i and j , and $\delta_{ij} = 1$ if $i = j$, and $= 0$, otherwise.

We define the estimator f_n^* of f by

$$f_n^*(\mathbf{x}) = (nh^p)^{-1} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h}\right) \tag{2.2}$$

where $h = h(n)$ is a sequence of positive constants such that $n^{\gamma/2}h^p \rightarrow \infty$, $0 < \gamma < 1$ and $h^p(\log n)^2 \rightarrow 0$ as $n \rightarrow \infty$ and $h \rightarrow 0$.

Let $F_{i,j}$ be the d.f. of $(\mathbf{X}_i, \mathbf{X}_j)$.

Theorem 2.1: *Suppose the sequence $\{X_i\}$ is absolutely regular with rates satisfying*

$$\beta(m) = \mathbf{O}(\rho^m) \text{ for some } 0 < \rho < 1. \tag{2.3}$$

Furthermore, assume that for any $j - i > 1$, there exists a continuous d.f. F_{j-i}^ on \mathbb{R}^{2p} with marginals F such that*

$$\|F_{i,j} - F_{j-i}^*\| = O(\rho_0^i), \quad 1 \leq i < j \leq n, \quad n \geq 1 \text{ for some } 0 < \rho_0 < 1 \tag{2.4}$$

where $\| \cdot \|$ denotes the norm of total variation. Then we have

$$\int E(f_n^*(\mathbf{x}) - f(\mathbf{x}))^2 d\mathbf{x} \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{2.5}$$

Proof: We have

$$\begin{aligned} & \int E(f_n^*(\mathbf{x}) - f(\mathbf{x}))^2 d\mathbf{x} \\ & \leq \int \left(\int \left(\frac{1}{nh^p} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right) - f(\mathbf{x}) \right)^2 \mathbf{Q}_n(d\mathbf{y}_1, \dots, d\mathbf{y}_n) \right) d\mathbf{x} \\ & \quad - \int \left(\int \left(\frac{1}{nh^p} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{z}_i}{h}\right) - f(\mathbf{x}) \right)^2 \mathbf{Q}_n^*(d\mathbf{z}_1, \dots, d\mathbf{z}_n) \right) d\mathbf{x} \\ & + 2 \int \left(\int \left(\frac{1}{nh^p} \left(\sum_{i=1}^n \left(K\left(\frac{\mathbf{x} - \mathbf{z}_i}{h}\right) - \int K\left(\frac{\mathbf{x}\mathbf{v}_i}{h}\right) f(\mathbf{v}_i) d\mathbf{v}_i \right) \right) \right)^2 \mathbf{Q}_n^*(d\mathbf{z}_1, \dots, d\mathbf{z}_n) \right) d\mathbf{x} \\ & \quad + 2 \int \left(\frac{1}{nh^p} \left(\sum_{i=1}^n \int \left(K\left(\frac{\mathbf{x} - \mathbf{z}_i}{h}\right) \right) f(\mathbf{z}_i) d\mathbf{z}_i - f(\mathbf{x}) \right) \right)^2 d\mathbf{x} \\ & = I_1 + I_2 + I_3 \end{aligned} \tag{2.6}$$

where \mathbf{Q}_n is the d.f. of $(\mathbf{X}_1, \dots, \mathbf{X}_n)$ and \mathbf{Q}_n^* is the d.f. of $(\mathbf{X}_1^*, \dots, \mathbf{X}_n^*)$ where $(\mathbf{X}_i^*, i \geq 1)$ is a strictly stationary sequence of random variables which is absolutely regular with a rate satisfying (2.3), for which the d.f. function of \mathbf{X}_1^* is F and the d.f. function of $(\mathbf{X}_1^*, \mathbf{X}_i^*)$ is F_{i-1}^* .

We can write

$$\begin{aligned} I_1 & \leq 2 \left| \int \int \left((nh^p)^{-1} \sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right) \right) (dF_i(\mathbf{y}_i)) f(\mathbf{x}) d\mathbf{x} \right| \\ & + \left| (nh^p)^{-2} \sum_{1 \leq i \neq j \leq n} \int \int \int K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right) K\left(\frac{\mathbf{x} - \mathbf{y}_j}{h}\right) (dF_{i,j}(\mathbf{y}_i, \mathbf{y}_j) - dF_{|j-i|}^*(\mathbf{y}_i, \mathbf{y}_j)) d\mathbf{x} \right| \\ & + \left| \int \int (nh^p)^{-2} \left(\sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right) \right)^2 (dF_i(\mathbf{y}_i) - dF(\mathbf{y}_i)) d\mathbf{x} \right| \\ & \leq 2 \left| n^{-1} \sum_{i=1}^n \int \left(\int K(\mathbf{u}_i) f(\mathbf{y}_i + h\mathbf{u}_i) d\mathbf{u}_i \right) (dF_i(\mathbf{y}_i) - dF(\mathbf{y}_i)) \right| \\ & + (nh^p)^{-2} \sum_{1 \leq i \neq j \leq n} \left| \int \int \int K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right)^2 K\left(\frac{\mathbf{x} - \mathbf{y}_j}{h}\right) (dF_{i,j}(\mathbf{y}_i, \mathbf{y}_j) - dF_{|j-i|}^*(\mathbf{y}_i, \mathbf{y}_j)) d\mathbf{x} \right| \\ & \quad + (nh^p)^{-2} \sum_{i=1}^n \int \int K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right)^2 (dF_i(\mathbf{y}_i) - dF(\mathbf{y}_i)) d\mathbf{x}. \end{aligned} \tag{2.7}$$

From the decompositions

$$\sum_{i=1}^m = \sum_{i=1}^m + \sum_{i=m+1}^n$$

and

$$1 \leq i \neq j \leq n = \sum_{\substack{m < |j-i| < n \\ 1 \leq i \neq j \leq n}} + \sum_{\substack{1 < |j-i| \leq m \\ 1 \leq i \neq j \leq n}}$$

where $m = [n^{1-\gamma/2}]$ ($[a]$ is the integer part of a), we have by using (2.4)

$$I_1 \leq O(n^{-\gamma/2}) + O(n^{1+\gamma/2}h^p)^{-1} \tag{2.8}$$

and I_1 converges to zero as $n \rightarrow \infty$ by using (2.4).

Next

$$I_2 \leq \left| \int n^{-1}h^{-2p} \int K^2\left(\frac{\mathbf{x}-\mathbf{z}}{h}\right)f(\mathbf{z})d\mathbf{z} - n^{-1}(h^{-p} \int K\left(\frac{\mathbf{x}-\mathbf{z}}{h}\right)f(\mathbf{z})d\mathbf{z})^2 d\mathbf{x} \right| + 2(nh^p)^{-2} \int \left(\sum_{1 \leq i < j \leq n} \text{cov} \left(K\left(\frac{\mathbf{x}-\mathbf{X}_i^*}{h}\right), K\left(\frac{\mathbf{x}-\mathbf{X}_j^*}{h}\right) \right) \right) d\mathbf{x} = I_2^1 + I_2^2. \tag{2.9}$$

From Lemma 6.8², we deduce

$$h^{-p} \int K^2\left(\frac{\mathbf{x}-\mathbf{z}}{h}\right)f(\mathbf{z})d\mathbf{z} \rightarrow f(\mathbf{x}) \int K^2(\mathbf{z})d\mathbf{z} \text{ as } n \rightarrow \infty$$

and

$$h^{-p} \int K\left(\frac{\mathbf{x}-\mathbf{z}}{h}\right)f(\mathbf{z})d\mathbf{z} \rightarrow f(\mathbf{x}) \text{ as } n \rightarrow \infty.$$

As $(nh^p) \rightarrow \infty$ when $n \rightarrow \infty$ we deduce that I_2^1 converges to zero as $n \rightarrow \infty$.

From the condition of absolute regularity and Lemma 6.3, we can write

$$\text{Cov} \left(K\left(\frac{\mathbf{x}-\mathbf{X}_i^*}{h}\right), K\left(\frac{\mathbf{x}-\mathbf{X}_j^*}{h}\right) \right) \leq 2C\beta^{\delta/(2+\delta)}(j-i)[E(|K\left(\frac{\mathbf{x}-\mathbf{X}_1^*}{h}\right)|^{2+\delta})]^{2/(2+\delta)}$$

where C is some constant > 0 .

Thus

$$\begin{aligned} & (nh^p)^{2/(2+\delta)-1} \sum_{1 \leq i < j \leq n} \text{cov} \left(K\left(\frac{\mathbf{x}-\mathbf{X}_i^*}{h}\right), K\left(\frac{\mathbf{x}-\mathbf{X}_j^*}{h}\right) \right) \\ & \leq (n^{-1}h^{-2p/(2+\delta)}) \sum_{i=1}^n \sum_{j=1}^{n-i} (n-j)2C\beta^{\delta/(2+\delta)}(j-i)[E(|K\left(\frac{\mathbf{x}-\mathbf{X}_1^*}{h}\right)|^{2+\delta})]^{2/(2+\delta)} \\ & \leq 2C \left(\sum_{i=1}^{\infty} \beta^{\delta/(2+\delta)}(i) \right) [h^{-p}E(|K\left(\frac{\mathbf{x}-\mathbf{X}_1^*}{h}\right)|^{2+\delta})]^{2/(2+\delta)} = I_2^*. \end{aligned}$$

From Lemma 6.8, the above expression converges (as $n \rightarrow \infty$) as

$$2C \left(\sum_{i=1}^{+\infty} \beta^{\delta/(2+\delta)}(i) \right) \left(f(\mathbf{x}) \int K(\mathbf{z})d\mathbf{z} \right)^{2/(2+\delta)}$$

Thus

$$I_2^2 \leq 4n^{-1}(h^p)^{\delta/(2+\delta)}I_2^*. \tag{2.10}$$

It follows that I_2^2 converges to zero and from (2.8)-(2.10) that I_2 converges to zero.

Finally, again from Lemma 6.8, we deduce that $I_3 \rightarrow 0$ as $n \rightarrow \infty$.

Thus $I_1 + I_2 + I_3 \rightarrow 0$ as $n \rightarrow \infty$ and Theorem 2.1 is proved.

²The proofs of Lemmas 6.1-6.9 are discussed in the Appendix.

Theorem 2.2: *Suppose the sequence $\{\mathbf{X}_i\}$ satisfies the conditions of Theorem 2.1. Then, at every continuity point \mathbf{x} of f , we have*

$$E(f_n^*(\mathbf{x})) \rightarrow f(\mathbf{x}) \text{ as } n \rightarrow \infty \tag{2.11}$$

$$E(f_n^*(\mathbf{x}) - f(\mathbf{x}))^2 \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{2.12}$$

Proof: Since the proof of (2.12) is similar to that of (2.5), we only prove (2.11).

We have

$$\begin{aligned} E(f_n^*(\mathbf{x}) - f(\mathbf{x})) &= (nh^p)^{-1} \sum_{i=1}^n \int K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right) f_i(\mathbf{y}_i) d\mathbf{y}_i - f(\mathbf{x}) \\ &= (nh^p)^{-1} \sum_{i=1}^n \int K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right) (f_i(\mathbf{y}_i) d\mathbf{y}_i - f(\mathbf{y}_i) d\mathbf{y}_i) + (nh^p)^{-1} \sum_{i=1}^n \int K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right) f(\mathbf{y}_i) d\mathbf{y}_i - f(\mathbf{x}). \end{aligned}$$

The first term converges to zero from condition (2.4) and the second term converges to zero from Lemma 6.8.

3. Asymptotic Normality of the Estimator $f_n^*(\mathbf{x})$

Denote

$$\sigma_1^2(\mathbf{x}) = f(\mathbf{x}) \int K^2(\mathbf{z}) d\mathbf{z}. \tag{3.1}$$

Theorem 3.1: *Suppose the sequence $\{\mathbf{X}_i\}$ satisfies the conditions of Theorem 2.1, then at every continuity point \mathbf{x} of f , $(nh^p)^2 [f_n^*(\mathbf{x}) - E(f_n^*(\mathbf{x}))]$ converges in law to the normal distribution with mean 0 and variance $\sigma_1^2(\mathbf{x})$.*

Proof: First we prove that

$$E[(nh^p)^2 [f_n^*(\mathbf{x}) - E(f_n^*(\mathbf{x}))]^2] \text{ converges to } \sigma_1^2(\mathbf{x}). \tag{3.2}$$

We have

$$\begin{aligned} &E[(nh^p)^2 [f_n^*(\mathbf{x}) - E(f_n^*(\mathbf{x}))]^2] \\ &= nh^p (nh^p)^{-2} \int \left[\sum_{i=1}^n K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right) - \left(\int K\left(\frac{\mathbf{x} - \mathbf{z}_i}{h}\right) f_i(\mathbf{z}_i) d\mathbf{z}_i \right) \right]^2 \mathbf{Q}_n(d\mathbf{y}_1, \dots, d\mathbf{y}_n) \\ &= (nh^p)^{-1} \int \sum_{i=1}^n \left[K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right) - \int K\left(\frac{\mathbf{x} - \mathbf{z}_i}{h}\right) f_i(\mathbf{z}_i) d\mathbf{z}_i \right]^2 f_i(\mathbf{y}_i) d\mathbf{y}_i \\ &\quad + (nh^p)^{-1} \int \sum_{1 \leq k \neq j \leq n} \left(K\left(\frac{\mathbf{x} - \mathbf{y}_i}{h}\right) - \int K\left(\frac{\mathbf{x} - \mathbf{z}_i}{h}\right) f_i(\mathbf{z}_i) d\mathbf{z}_i \right) \\ &\quad \times \left(K\left(\frac{\mathbf{x} - \mathbf{y}_j}{h}\right) - \int K\left(\frac{\mathbf{x} - \mathbf{z}_j}{h}\right) f_j(\mathbf{z}_j) d\mathbf{z}_j \right) F_{i,j}(d\mathbf{y}_i, d\mathbf{y}_j) = J_1 + J_2. \end{aligned}$$

It follows from (2.4) that

$$J_1 \rightarrow f(\mathbf{x}) \int K^2(\mathbf{z}) d\mathbf{z} \text{ as } n \rightarrow \infty. \tag{3.3}$$

Now, define $A_{i,j}$ by

$$A_{i,j} = \int \left[K\left(\frac{\mathbf{x}-\mathbf{y}_i}{h}\right) - \int K\left(\frac{\mathbf{x}-\mathbf{z}_i}{h}\right) f_i(\mathbf{z}_i) d\mathbf{z}_i \right] \left[K\left(\frac{\mathbf{x}-\mathbf{y}_j}{h}\right) - \int K\left(\frac{\mathbf{x}-\mathbf{z}_j}{h}\right) f_j(\mathbf{z}_j) d\mathbf{z}_j \right] \\ \times F_{i,j}(d\mathbf{y}_i, d\mathbf{y}_j).$$

Letting $k = k_n = (\log n)^2$, where $[a]$ denotes the integer part of a , we have

$$J_2 \leq (nh^p)^{-1} \left(\sum_{\substack{1 \leq i \neq j \leq n \\ i-j \leq k}} |A_{i,j}| + \sum_{\substack{1 \leq i \neq j \leq n \\ j-i > k}} |A_{i,j}| \right) \\ \leq h^{-pk} M^2 h^{2p} + h^{-p} \sum_{i=k+1}^n [\beta(i+1)]^{\delta/(2+\delta)} (M_{2+\delta})^{2/(2+\delta)}$$

where $M = \sup_{x \in \mathbb{R}} |K(x)|$ and $M_{2+\delta} = \sup_{n \geq 1} \max_{1 \leq i \leq n} E |K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h}\right)|^{2+\delta}$. As $h^p(\log n)^2 \rightarrow 0$ as $n \rightarrow \infty$, we deduce $h^{pk} \rightarrow 0$ as $n \rightarrow \infty$. From (2.3), we deduce also that $h^{-p}[\beta(k+1)]^{\delta/(2+\delta)} \rightarrow 0$ as $n \rightarrow \infty$. Consequently,

$$J_2 \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{3.4}$$

From (3.3) and (3.4) we have (3.2).

Now let c_0 be a sufficient large number such that

$$\beta(m) = o(n^{-2})$$

where $m = m_n = [c_0 \log n]$. Further, let $\ell = \ell_n = [n^{1-\gamma}]$ and $q = q_n = [n/\ell + m]$. Define a sequence $\{(a'_i, a''_i), i = 1, \dots, q\}$ of pairs of integers inductively as follows:

$$a''_0 = 0, a'_i = a''_{i-1} + m, a''_i = a'_i + \ell - 1, \quad (i = 1, 2, \dots, q).$$

Define A_j by

$$A_j = K\left(\frac{x - X_j}{h}\right) - E\left(\left(K\left(\frac{x - X_j}{h}\right)\right)\right).$$

Using Lemma 6.1 (in Appendix), and Lemma 4 of Takahata and Yoshihara [15], we have

$$E\left(\exp\left\{it \frac{1}{n^{1/2} h^{p/2}} \sum_{j=1}^n A_j\right\}\right) \sim \prod_{\alpha=1}^q E\left(\exp\left\{it \frac{1}{n^{1/2} h^{p/2}} \sum_{j=a'_\alpha}^{a''_\alpha} A_j\right\}\right) + Cq\beta(m) \\ = \left\{1 - \frac{t^2}{2nh^p} E\left(\sum_{j=1}^{\ell c} A_j\right)^2 + O\left(\frac{|t^3|}{n^{3/2} h^{3p/2}} E\left|\sum_{j=1}^{\ell} A_j\right|^3\right)\right\}^q + o(n^{-2+\gamma}) \\ = \exp\left\{-\frac{t^2}{2} \frac{q}{nh^p} E\left(\sum_{j=1}^{\ell} A_j\right)^2 + O\left(q\left(\frac{\ell}{n}\right)^{3/2} h^{-3p/2} |t|^3\right)\right\} + o(n^{-2+\gamma}).$$

Thus by (3.2)

$$\lim_{n \rightarrow \infty} \frac{q}{nh^p} E\left(\sum_{j=1}^{\ell} A_j\right)^j = \sigma_1^2(x).$$

The results follows.

4. Asymptotic Normality of the Integrated Square Error I_n

We assume that

- (i) $\int |x^{(i)}x^{(j)}x^{(k)}| K(x)dx \leq M < \infty$ for each i, j and k ($1 \leq i, j, k \leq p$),
- (ii) the density functions $f_j^*(\mathbf{x}, \mathbf{y})$ of $F_j^*(\mathbf{x}, \mathbf{y})$ exist for each j ,
- (iii) second partial derivatives of $f(\mathbf{x})$ and $f_j^*(\mathbf{x}, \mathbf{y})$ exist, are uniformly bounded and satisfy the Lipschitz condition of order one. Furthermore, all the second order partial derivatives of $f(\mathbf{x})$ and of $f_j^*(\mathbf{x}, \mathbf{y})$ belong to the ball in $L^1(\mathbb{R}^p)$, and in $L^1(\mathbb{R}^p \times \mathbb{R}^p)$ respectively.

Denote

$$\begin{aligned} \sigma_2^2 &= \int \{[\Delta f(\mathbf{x})]^2\} f(\mathbf{x}) d\mathbf{x} - \left[\int \{\Delta f(\mathbf{x})\} f(\mathbf{x}) d\mathbf{x} \right]^2 \\ &+ 2 \sum_{i=1}^{\infty} \left(\int \int \{\Delta f(\mathbf{x}) \Delta f(\mathbf{y})\} dF_i^*(d\mathbf{x}, d\mathbf{y}) - \left[\int \{\Delta f(\mathbf{x})\} f(\mathbf{x}) d\mathbf{x} \right]^2 \right), \end{aligned} \tag{4.1}$$

where $\Delta = \sum_{i=1}^p \frac{\partial^2}{\partial x^{(i)2}}$ is Laplacian,

$$\tau^2 \sigma_3^2 = \left\{ \int f^2(\mathbf{x}) d\mathbf{x} \right\} \int \left\{ \int K(\mathbf{u}) K(\mathbf{u} + \mathbf{v}) d\mathbf{u} \right\}^2 d\mathbf{v}, \tag{4.2}$$

and let

$$d(n) = \begin{cases} n^{\frac{1}{2}} h^{-2} & \text{if } nh^{p+4} \rightarrow \infty \\ nh^{p/2} & \text{if } nh^{p+4} \rightarrow 0 \\ n^{(p+8)/2(p+4)} & \text{if } nh^{p+4} \rightarrow \lambda \ (0 < \lambda < \infty). \end{cases} \tag{4.3}$$

(Note that τ in (4.2) is the same as in (2.1)).

Then our main result is the following:

Theorem 4.1: *Suppose that the conditions of Theorem 2.1 hold. Then $\sigma_2 > 0$ and $\sigma_3 > 0$ exist, and*

$$d(n)\{I_n - E(I_n)\} \Rightarrow \begin{cases} 2\tau\sigma_2 Z & \text{if } nh^{p+4} \rightarrow \infty \\ \frac{1}{2}\sigma_3 Z & \text{if } nh^{p+4} \rightarrow 0 \\ (r\tau^2\sigma_2^2\lambda^{4/(p+4)} + 2\sigma_3^2\lambda^{-p/(p+4)})^{\frac{1}{2}} & \text{if } nh^{p+4} \rightarrow \lambda \ (0 < \lambda < \infty). \end{cases} \tag{4.4}$$

in distribution where Z has the standard normal distribution.

Proof: For brevity, we use the following notations

$$\begin{aligned} H_{i,j}(\mathbf{x}, \mathbf{y}) &= \int \left\{ K\left(\frac{\mathbf{u}-\mathbf{x}}{h}\right) - E\left(K\left(\frac{\mathbf{u}-\mathbf{X}_i}{h}\right)\right) \right\} \left\{ K\left(\frac{\mathbf{u}-\mathbf{y}}{h}\right) - E\left(K\left(\frac{\mathbf{u}-\mathbf{X}_j}{h}\right)\right) \right\} du, \quad 1 \leq i, j \leq n \\ \tilde{H}_{i,j}(\mathbf{X}_i, \mathbf{X}_j) &= H_{i,j}(\mathbf{X}_i, \mathbf{X}_j) - E(H_{i,j}(\mathbf{X}_i, \mathbf{X}_j)) \end{aligned} \tag{4.5}$$

$$K_j = \int \left\{ K\left(\frac{\mathbf{x}-\mathbf{X}_j}{h}\right) - E\left(K\left(\frac{\mathbf{x}-\mathbf{X}_j}{h}\right)\right) \right\} \{E(f_n^*(\mathbf{x})) - f(\mathbf{x})\} d\mathbf{x}, \quad 1 \leq j \leq n. \tag{4.6}$$

First, we decompose $I_n - E(I_n)$ as follows:

$$I_n - E(I_n) = 2(n^2 h^{2p})^{-1} \sum_{1 \leq i < j \leq n} \tilde{H}_{i,j}(\mathbf{X}_i, \mathbf{X}_j)$$

$$\begin{aligned}
 &+ 2(nh^p)^{-1} \sum_{j=1}^n K_j + (n^2h^{2p})^{-1} \sum_{j=1}^n \left[\int \left\{ K\left(\frac{\mathbf{x}-\mathbf{X}_j}{h}\right) - E\left(K\left(\frac{\mathbf{x}-\mathbf{X}_j}{h}\right)\right) \right\}^2 d\mathbf{x} \right. \\
 &\quad \left. - \int E\left\{ K\left(\frac{\mathbf{x}-\mathbf{X}_j}{h}\right) - E\left(K\left(\frac{\mathbf{x}-\mathbf{X}_j}{h}\right)\right) \right\}^2 d\mathbf{x} \right] \\
 &= I_1 + I_2 + I_3. \tag{4.7}
 \end{aligned}$$

The main part of the proof of the theorem is broken into proofs of the following four propositions. The first proposition uses Dvoretzsky’s theorem [7] and Proposition 3.1 of Takahata and Yoshihara [15].

Let c_1 be a sufficiently large number such that

$$\beta(m) = O(n^{-8})$$

where $m = m_n = [c_1 \log n]$.

Further, let $r = r_n = [n^{\frac{1}{4}}]$ and $k = k_n = [n/(r + m)]$. Define a sequence $\{(a_i, b_i), i = 1, \dots, k\}$ of pairs of integers inductively as follows:

$$b_0 = 0, \quad a_i = b_{i-1} + m, \quad b_i = a_i + r - 1 \quad (i = 1, 2, \dots, k).$$

Let $\mathcal{F}_\alpha = \sigma(X_i, 1 \leq i \leq a_\alpha - m)$, $(\alpha = 1, 2, \dots, k)$.

Put

$$T_\alpha = T_{n\alpha} = \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha - m} H_{i,j}(\mathbf{X}_i, \mathbf{X}_j), \quad \alpha = 1, \dots, k \tag{4.7'}$$

and

$$U_n = \sum_{\alpha=1}^k (T_{n\alpha} - E(T_{n\alpha})), \quad S_n = \sum_{1 \leq i < j \leq n} \tilde{H}_{i,j}(\mathbf{X}_i, \mathbf{X}_j).$$

Proposition 4.1: *If the conditions of Theorem 4.1 are satisfied, then $(nh^{3p/2})^{-1} S_n$ converges in law to the normal distribution with mean 0 and variance σ_3^2 defined in (4.2).*

Proof: Let $s_n = \text{Var} U_n$. If we prove

$$s_n^{-1} \sum_{\alpha=1}^k E\{T_\alpha | \mathcal{F}_\alpha\} \rightarrow 0 \text{ in probability as } n \rightarrow \infty, \tag{4.8}$$

$$(s_n^2)^{-1} \sum_{\alpha=1}^k [E\{T_\alpha^2 | \mathcal{F}_\alpha\} - (E\{T_\alpha | \mathcal{F}_\alpha\})^2] \rightarrow 1 \text{ in probability as } n \rightarrow \infty, \tag{4.9}$$

and

$$(s_n^4)^{-1} \sum_{\alpha=1}^k E(T_\alpha)^4 \rightarrow 0 \text{ as } n \rightarrow \infty, \tag{4.10}$$

then, it will follow from Dvoretzsky’s theorem [7] that $s_n^{-1} U_n$ converges in law to a $\mathcal{N}(0, 1)$ random variable.

The proofs of (4.8) and (4.9) are given in Lemma 6.4 and that of (4.10) in Lemma 6.7 in the Appendix.

Proposition 4.1 will now follow if we show that

$$2s_n^2(n^2h^{3p})^{-1} = \sigma_3^2(1 + o(1)) \tag{4.11}$$

and

$$s_n^{-2} E(S_n - U_n)^2 \rightarrow 0 \text{ in probability as } n \rightarrow \infty. \tag{4.12}$$

(4.11) and (4.12) are proved in Lemmas 6.4 and 6.6, respectively (see Appendix).

Proposition 4.2: *If conditions of Theorem 4.1 are satisfied and if $nh^{p+4} \rightarrow \infty$ as $n \rightarrow \infty$, then $n^2 h^{-2} I_2$ converges in law to the normal distribution with mean 0 and variance $2\tau^2 \sigma_2^2$ defined in (4.1).*

Proof: We first prove that

$$E(nh^{-4} I_2^2) \text{ converges to } 2\tau^2 \sigma_2^2 \text{ as } n \rightarrow \infty. \quad (4.13)$$

We have

$$E(nh^{-4} I_2^2) = 4(nh^{2p+4})^{-1} E\left(\sum_{j=1}^n K_j\right)^2. \quad (4.14)$$

First, we prove that

$$\lim_{n \rightarrow \infty} n^{-1} \left| \sum_{1 \leq i < j \leq n} ((h^{2p+4})^{-1} E(K_i K_j) - C_{j-i}) \right| = 0 \quad (4.15)$$

where

$$C_\ell = \tau^2 \int \int \{\Delta f(\mathbf{x}) \Delta f(\mathbf{y})\} F_\ell^*(d\mathbf{x}, d\mathbf{y}) - \left[\int \{\Delta f(\mathbf{x})\} f(\mathbf{x}) d\mathbf{x} \right]^2, \quad \ell > 1.$$

We can write

$$E(K_i K_j) = \int \int \left[\int \left\{ K\left(\frac{\mathbf{x}-\mathbf{y}_i}{h}\right) - E\left(K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h}\right)\right) \right\} \{E(f_n^*(\mathbf{x})) - f(\mathbf{x})\} d\mathbf{x} \right] \\ \left[\int \left\{ K\left(\frac{\mathbf{z}-\mathbf{y}_j}{h}\right) - E\left(K\left(\frac{\mathbf{z}-\mathbf{X}_j}{h}\right)\right) \right\} \{E(f_n^*(\mathbf{z})) - f(\mathbf{z})\} d\mathbf{z} \right] F_{i,j}(d\mathbf{y}_i, d\mathbf{y}_j).$$

Since

$$E(f_n^*(\mathbf{x}) - f(\mathbf{x})) = (nh^p)^{-1} \sum_{i=1}^n \int K\left(\frac{\mathbf{x}-\mathbf{u}}{h}\right) (dF_i(\mathbf{u}) - dF(\mathbf{u})) \\ + (nh^p)^{-1} \sum_{i=1}^n \int K\left(\frac{\mathbf{x}-\mathbf{u}}{h}\right) dF(\mathbf{u}) - f(\mathbf{x})$$

$= O(n^{-\gamma}) + h^2 \tau \Delta f(\mathbf{x}) + O(h^3)$ (from conditions (2.4) and (i)-(iii)), we obtains

$$E(K_i K_j) = \int \int \left[\int \left\{ K\left(\frac{\mathbf{x}-\mathbf{y}_i}{h}\right) - E\left(K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h}\right)\right) \right\} \{h^2 \tau \Delta f(\mathbf{x}) + O(h^3) + O(n^{-\gamma})\} d\mathbf{x} \right] \\ \left[\int \left\{ K\left(\frac{\mathbf{z}-\mathbf{y}_j}{h}\right) - E\left(K\left(\frac{\mathbf{z}-\mathbf{X}_j}{h}\right)\right) \right\} \{h^2 \tau \Delta f(\mathbf{z}) + O(h^3) + O(n^{-\gamma})\} d\mathbf{z} \right] dF_{i,j}(\mathbf{y}_i, \mathbf{y}_j).$$

Put

$$C_\ell^* = \int \int \left[\int \left\{ K\left(\frac{\mathbf{x}-\mathbf{y}}{h}\right) - E\left(K\left(\frac{\mathbf{x}-\mathbf{X}_1^*}{h}\right)\right) \right\} \{h^2 \tau \Delta f(\mathbf{x})\} d\mathbf{x} \right] \\ \left[\int \left\{ K\left(\frac{\mathbf{z}-\mathbf{u}}{h}\right) - E\left(K\left(\frac{\mathbf{z}-\mathbf{X}_1^*}{h}\right)\right) \right\} \{h^2 \tau \Delta f(\mathbf{z})\} d\mathbf{z} \right] dF_\ell^*(d\mathbf{y}, d\mathbf{u})$$

which implies that

$$n^{-1} \left| (h^{2p+4})^{-1} \sum_{1 \leq i < j \leq n} E(K_i K_j) - \sum_{i=1}^n (n-i) C_i^* \right| \\ \leq n^{-1} \left| (h^{2p+4})^{-1} \sum_{1 \leq j-i \leq m} E(K_i K_j) - \sum_{i=1}^m (n-i) C_i^* \right| \\ + n^{-1} \left| (h^{2p+4})^{-1} \sum_{m < j-i \leq n} E(K_i K_j) - \sum_{i=m+1}^n (n-i) C_i^* \right|$$

$$\begin{aligned} &\leq n^{-1} |(h^{2p+4})^{-1} \sum_{1 \leq j-i \leq m} \int \int \left[\int \left\{ K\left(\frac{\mathbf{x}-\mathbf{y}_i}{h}\right) - E\left(K\left(\frac{\mathbf{x}-\mathbf{X}_i}{h}\right)\right) \right\} h^{2\tau} \Delta f(\mathbf{x}) d\mathbf{x} \right] \\ &\quad \times \left[\int \left\{ K\left(\frac{\mathbf{z}-\mathbf{y}_j}{h}\right) - E\left(K\left(\frac{\mathbf{z}-\mathbf{X}_j}{h}\right)\right) \right\} h^{2\tau} \Delta f(\mathbf{z}) d\mathbf{z} \right] F_{i,j}(d\mathbf{y}_i, d\mathbf{y}_j) - \sum_{i=1}^m (n-i) C_i^* | \\ &\quad + m [\mathbf{O}(h) + \mathbf{O}((n\gamma h^2)^{-1}) + \mathbf{O}(h^2) + \mathbf{O}((n\gamma h)^{-1}) + \mathbf{O}(n^{-2\gamma})] \\ &\quad + n^{-1} |(h^{2p+4})^{-1} \sum_{m < j-i \leq n} E(K_i K_j) - \sum_{i=m+1}^n (n-i) C_i^* | \\ &= I_n + J_n + L_n, \text{ say.} \end{aligned}$$

From condition (2.4), we deduce

$$I_n = m \mathbf{O}(n^{-1}).$$

Also

$$J_n = m [\mathbf{O}(h) + \mathbf{O}((n\gamma h^2)^{-1})].$$

From condition (2.3), we have

$$L_n = \mathbf{O}(\rho^m).$$

For $\epsilon > 0$, let m be fixed such that

$$L_n < \epsilon/3. \tag{4.16}$$

We can find n_0 sufficiently large such that for any $n \geq n_0$

$$J_n < \epsilon/3 \text{ and } I_n < \epsilon/3. \tag{4.17}$$

From (4.16) and (4.17) we deduce that

$$\lim_{n \rightarrow \infty} \left| \sum_{1 \leq i < j \leq n} E(K_i K_j) - C_{j-i}^* \right| = 0. \tag{4.18}$$

From Remark 2 in Takahata and Yoshihara [15], we deduce

$$\lim_{n \rightarrow \infty} \left| \sum_{1 \leq i < j \leq n} (C_{j-i}^* - C_{j-i}) \right| = 0 \tag{4.19}$$

(4.18) and (4.19) entail (4.15).

Following the arguments similar to deriving (4.18) and (4.19), we obtain

$$\lim_{n \rightarrow \infty} \left| (nh^{2p+4})^{-1} \sum_{i=1}^n E(K_i)^2 - C_0 \right| = 0 \tag{4.20}$$

where

$$C_0 = \int \{\Delta f(\mathbf{x})\}^2 f(\mathbf{x}) d\mathbf{x} - \left[\int \{\Delta f(\mathbf{x})\} f(\mathbf{x}) d\mathbf{x} \right]^2.$$

Finally we have

$$\begin{aligned} E(nh^{-4} I_2^2) - 2\tau^2 \sigma_2^2 &\leq 2n^{-1} \left| \sum_{1 \leq i \neq j \leq n} (h^{2p+4})^{-1} E(K_i K_j) - \sum_{i=1}^n (n-i) C_i \right| \\ &\quad + \sum_{i=n+1}^{+\infty} |C_i| + \sum_{i=1}^n \sum_{k=i}^{+\infty} |C_k| \\ &= A_n + B_n + C_n, \text{ say.} \end{aligned}$$

From (4.15) and (4.20), $A_n \rightarrow 0$ as $n \rightarrow \infty$ and from Lemma 6.3 in the Appendix and condition (2.3) we easily deduce that $B_n \rightarrow 0$ and $C_n \rightarrow 0$ as $n \rightarrow \infty$.

This proves (4.13).

Now using Lemma 6.9 (in Appendix) and Lemma 4 of Takahata and Yoshihara [15]

$$E\left(\sum_{j=1}^n K_j\right)^3 \leq C\left(\sum_{j=1}^n E(K_j^6)^{1/3}\right)^{3/2} \leq Cn^{3/2}h^{3(p+2)}$$

where C is some constant > 0 .

Hence, using Lemma 6.1 (in Appendix) we have

$$\begin{aligned} E\left(\exp\left\{it\frac{1}{n^{1/2}h^{p+2}}\sum_{j=1}^n K_j\right\}\right) &\sim \prod_{\alpha=1}^k E\left(\exp\left\{it\frac{1}{n^{1/2}h^{p+2}}\sum_{j=\alpha_\alpha}^{b_\alpha} K_j\right\}\right) + Ck\beta(m) \\ &= \left\{1 - \frac{t^2}{2nh^{2(p+2)}}E\left(\sum_{j=1}^r K_j\right)^2 + O\left(\frac{|t^3|}{n^{3/2}h^{3(p+2)/2}}E\left|\sum_{j=1}^r K_j\right|^3\right)\right\}^k + o(n^{-1}) \\ &= \exp\left\{-\frac{t^2}{2}\frac{k}{nh^{2(p+2)}}E\left(\sum_{j=1}^r K_j\right)^2 + O(k\left(\frac{r}{n}\right)^{3/2}|t|^3)\right\} + o(n^{-1}). \end{aligned}$$

Thus by (4.13)

$$\lim_{n \rightarrow \infty} \frac{k}{hh^{2(p+2)}}E\left(\sum_{j=1}^r K_j\right)^2 = 2\tau^2\sigma_2^2.$$

The result follows.

Proposition 4.3: *If $nh^{p+4} \rightarrow \lambda$, ($0 < \lambda < \infty$) as $n \rightarrow \infty$, then $n^{(p+8)/2(p+4)}I_2$ and $n^{(p+8)/2(p+4)}n^{-2}h^{-2p}S_n$ are asymptotically uncorrelated as $n \rightarrow \infty$.*

Proof: By Lemma 6.1, Schwarz's inequality and (6.3), we have

$$\begin{aligned} |E(nh^p I_2 S_n)| &\leq \sum_{i=1}^n \sum_{1 \leq i < k \leq n} |E(K_i \tilde{H}_{j,k}(\mathbf{X}_j, \mathbf{X}_k))| \\ &\leq \left\{ \sum_{\max(|i-j|, |j-k|, |k-i|) \leq m} + \sum_{\max(|i-j|, |j-k|, |k-i|) > m} \right\} |E(K_i \tilde{H}_{j,k}(\mathbf{X}_j, \mathbf{X}_k))| \\ &\leq C[mn \sup_{1 \leq i \leq n} \|K_i\|_2 \max_{j,k} \|\tilde{H}_{j,k}(\mathbf{X}_j, \mathbf{X}_k)\|_2 + n^3 \beta(m)] \\ &\leq C[nm^2 h^p + o(n^{-5})] \end{aligned} \tag{4.21}$$

since $\sup_{1 \leq i \leq n} \|K_i\|_2 \leq Ch^{p+2}$ where $\|K_i\|_2 = (E(K_i^2))^{1/2}$ and C is a constant > 0 .

From (4.21), we deduce

$$\begin{aligned} E(n^{(p+8)/2(p+4)} I_2 n^{(p+8)/2(p+4)} n^{-2} h^{-2p} S_n) \\ \leq Cn^{1+\frac{4}{p+4}} n^{-2} h^{-2p} n^{-1} h^{-p} n h^p + 2h^p m^2 \\ = \frac{Ch^2 m^2}{(nh^{p+4})^{p/p+4}}. \end{aligned} \tag{4.22}$$

From $nh^{p+4} \rightarrow \lambda$ as $n \rightarrow \infty$ and $m = O(\log n)$, we deduce that

$$\frac{h^2 m^2}{(nh^p + 4)^{p/(p+4)}} \rightarrow 0 \text{ as } n \rightarrow \infty \tag{4.23}$$

which proves Proposition 4.3.

Proposition 4.4: *If the conditions of Theorem 4.1 are satisfied, then*

$$\text{Var}(I_3) = O(n^{-3}h^{-2p}).$$

Proof: Let

$$M_i = \int \left\{ K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h}\right) - E\left(K\left(\frac{\mathbf{x} - \mathbf{X}_i}{h}\right)\right) \right\}^2 d\mathbf{x}.$$

From Lemma 2 in Hall [8], it follows that

$$\sup_{1 \leq i \leq n} E((M_i)^j) = O(h^{jp}).$$

By Lemma 4 of Takahata and Yoshihara [15], we have

$$\begin{aligned} n^4 h^{4p} \text{Var}(I_3) &= E \left[\sum_{j=1}^n \{M_j - E(M_j)\} \right]^2 \leq Cn \sup_{1 \leq j \leq n} \|M_j - E(M_j)\|_3^2 \\ &\leq Cn \sup_{1 \leq j \leq n} \|M_j\|_3^2 \leq Cnh^{2p} \text{ where } C \text{ is a constant } > 0 \end{aligned}$$

and Proposition 4.4 is proved.

5. Applications

5.1

Consider a sequence $\{\mathbf{X}_i, i \geq 1\}$ of \mathbb{R}^p -valued random variables which is a Markov process with transition probability $P(\mathbf{x}; A)$ where $A \in \mathfrak{B}$, \mathfrak{B} is the Borel σ -field of \mathbb{R}^p and $\mathbf{x} \in \mathbb{R}^p$.

Recall that the Markov process is *geometrically ergodic* if it is ergodic and if there exists $0 < \rho < 1$ such that

$$\|P^n(\mathbf{x}; \cdot) - \mu(\cdot)\| = O(\rho^n) \text{ for all a.s. } \mathbf{x} \in \mathbb{R}^p \tag{5.1}$$

where μ is the invariant measure and P^n the n -step transition probability.

We say that the process $\{\mathbf{X}_i\}_{i \geq 1}$ has ν for initial probability measure if the law of probability of \mathbf{X}_1 is defined by ν and for any $i > 1$, the law of probability P_i of \mathbf{X}_i is defined by νP^{i-1} .

For any probability measure ν and any transition probability \mathbf{Q} we denote by $\mathbf{Q} \otimes \nu$ the probability measure defined on \mathbb{R}^{2p} by

$$\mathbf{Q} \otimes \nu(A \times B) = \int_B \mathbf{Q}(\mathbf{x}; A) \nu(d\mathbf{x}) \text{ for any } A \times B \in \mathfrak{B} \times \mathfrak{B}.$$

The Markov process is called *strongly aperiodic* if for any $\mathbf{x} \in \mathbb{R}^p$, the transition probability $P(\mathbf{x}; \cdot)$ is equivalent to the Lebesgue measure.

The Markov process is called *Harris recurrent* if there exists a σ -finite measure ν on \mathbb{R}^p with $\nu(\mathbb{R}^p) > 0$ such that $\nu(A) > 0$ implies $(P(\mathbf{x}; \mathbf{X}_i \in A \text{ i.o.}) = 1$ for all $\mathbf{x} \in \mathbb{R}^p$.

Theorem 5.1: *Let $\{\mathbf{X}_i\}_{i \geq 1}$ be a Markov process which is strongly aperiodic, Harris*

recurrent, and geometrically ergodic. Suppose that

- (j) the invariant measure μ has a density f which admits bounded second partial derivatives which are integrable, and furthermore

$$\int |x^{(j)}| f(\mathbf{x}) d\mathbf{x} < \infty,$$

$$\int |x^{(j)}| \frac{\partial}{\partial x^{(j)}} f(\mathbf{x}) d\mathbf{x} < \infty, 1 \leq j \leq p.$$

- (jj) the transition probability $P(\cdot; \cdot)$ has a transition density $p(\mathbf{x}; \mathbf{y})$ which admits bounded third partial derivatives. Moreover, the first and second derivatives are bounded and integrable with respect to \mathbf{y} for each \mathbf{x} ; they also satisfy

$$\int |y^{(j)}| p(\mathbf{x}; \mathbf{y}) d\mathbf{y} < \infty$$

$$\int |y^{(j)}| \frac{\partial}{\partial y^{(j)}} p(\mathbf{x}; \mathbf{y}) d\mathbf{y} < \infty$$

$$\sup_{n \in \mathbb{N}^*} \int |y^{(j)}| p^{(n)}(\mathbf{x}; \mathbf{y}) d\mathbf{y} \leq A |x^{(j)}|, \quad 1 \leq j \leq p, \mathbf{x} \in \mathbb{R}^p$$

where A is some constant > 0 and $p^{(n)}$ is the transition density of P^n . Then, for any initial measure ν , the conclusions of Theorems 2.1, 2.2, 3.1 and 4.1 hold for the nonparametric estimator f_n^* defined in (2.2).

Proof: From Theorems 2.1, 2.2 and 3.1, we have only to prove (2.3) and (2.4).

First we prove (2.3). From Davydov [4] and the condition of strong aperiodicity, we have

$$\begin{aligned} \beta(m) &= \sup_n \int P_n(d\mathbf{x}) \|P^m(\mathbf{x}; \cdot) - P_{m+n}(\cdot)\| \\ &\leq \sup_n \int P_n(d\mathbf{x}) \|P^m(\mathbf{x}; \cdot) - \mu(\cdot)\| + \|P_{m+n}(\cdot) - \mu(\cdot)\|. \end{aligned}$$

As the process is geometrically ergodic, we can find $0 < \rho < 1$ such that

$$\|P^m(\mathbf{x}; \cdot) - \mu(\cdot)\| = O(\rho^m) \text{ for all a.s. } \mathbf{x} \in \mathbb{R}^p. \tag{5.2}$$

From Theorem 2.1 of Nummelin and Touminen [10], we deduce

$$\beta(m) = O(\rho^m) \tag{5.3}$$

which is the same as (2.3).

Now, we prove (2.4). We have from (5.2)

$$\begin{aligned} \|P^m \otimes P_n - P^m \otimes \mu\| &= 2 \sup_{A \times B \in \mathfrak{B} \times \mathfrak{B}} \left| \int_B P^m(\mathbf{x}; A) P_n(d\mathbf{x}) - \int_B P^m(\mathbf{x}; A) \mu(d\mathbf{x}) \right| \\ &\leq 2 \sup_{A \times B \in \mathfrak{B} \times \mathfrak{B}} P^m(\mathbf{x}; A) \left| \int_B P_n(d\mathbf{x}) - \mu(d\mathbf{x}) \right| \\ &\leq 2 \|P_n - \mu\| = O(\rho^n) \end{aligned}$$

that is

$$\|P^m \otimes P_n - P^m \otimes \mu\| = O(\rho^n). \quad (5.4)$$

Thus the conclusions of Theorems 2.1, 2.3 and 3.1 hold.

To prove Theorem 4.1, we have only to verify the conditions (i)-(iii) of Section 4, but they are easy consequences of conditions (i) and (ii) of Theorem 5.1.

Example 5.2: We consider an ARMA process

$$\mathbf{X}_i = a\mathbf{X}_{i-1} + b\epsilon_i + \epsilon_{i-1}, \quad i \in \mathbb{N} \quad (5.5)$$

where \mathbf{X}_0 admits a strictly positive density, $\{\epsilon_i, i \in \mathbb{N}\}$ is a sequence of independent and identically distribution (i.i.d.) \mathbb{R}^p -valued random variables with strictly positive density such that $E(\epsilon_i) = 0$, and a and b are real numbers such that $|a| < 1$.

If the density function g of ϵ_0 has three bounded first partial derivatives such that the first and second derivatives are integrable and satisfy

$$\int |y^{(j)}| g(\mathbf{y}) d\mathbf{y} < \infty \text{ and } \int |y^{(j)}| \frac{\partial}{\partial y^{(j)}} g(\mathbf{y}) d\mathbf{y} < \infty, \quad 1 \leq j \leq p$$

and if moreover, the density of the invariant measure satisfies condition (j) in Theorem 5.1, then the conditions of Theorem 5.1 are satisfied for the process defined in (5.5), because we have here a particular case of Markov process satisfying our conditions. The law of the process on which observations are taken is defined by the initial measure (i.e., the measure which defines the law of \mathbf{X}_0) and the transitional measures (defined from the formula (5.5)). From the fact that regardless of which is the initial measure, the density function of the measure of \mathbf{X}_n converges to the density function of the invariant measure, it is clear that if the process defined by (5.5) satisfies the above conditions of derivability, we can estimate the density f of the invariant measure by the estimator f_n^* defined in (2.2) for any initial measure of \mathbf{X}_0 which admits a strictly positive density. Moreover, we can also apply the central limit theorem to f_n^* and I_n to study the confidence regions based on these statistics. For example, if the initial measure is Gaussian, then \mathbf{X}_0 admits strictly positive density.

5.2 Applications to φ -mixing Markov processes

Theorem 5.2: Let $\{\mathbf{X}_i\}_{i \geq 1}$ be a Markov process which is aperiodic and Doebelin recurrent. Suppose that conditions (j) and (jj) of Theorem 5.1 are satisfied. Then, for any initial measure, the conclusions of Theorems 2.1, 2.2, 3.1 and 4.1 hold for the nonparametric estimator f_n^* .

Proof: From Theorem 4.1 of Davydov [4] the process $\{\mathbf{X}_i\}$ is geometrically φ -mixing which implies geometrically absolute regularity. The proof is now similar to Theorem 5.1.

Example 5.2: We consider the process

$$\mathbf{X}_i = f(\mathbf{X}_{i-1}) + \epsilon_{i-1}, \quad i \in \mathbb{N}^* \quad (5.6)$$

where \mathbf{X}_0 admits some strictly positive density and $\{\epsilon_i, i \in \mathbb{N}\}$ is a sequence of i.i.d. \mathbb{R}^p -valued random variables with strictly positive density such that $E(\epsilon_i) = 0$ and f is a bounded continuous function.

If the density function g of ϵ_0 and the function f admit the three bounded first partial derivatives and if the density of the invariant measure has bounded second derivatives which are integrable and the first derivatives are also integrable, then we are in the same situation as Theorem 5.2.

We can also under these conditions, estimate the density f by f_n^* for any initial measure which admits a strictly positive density.

6. Appendix

The Lemmas (6.1 to 6.3) are well known results and their proofs are not given.

Lemma 6.1: *Let $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ be random \mathbb{R}^p vectors satisfying an absolutely regular condition with mixing rate $\beta(m)$.*

Let $h(\mathbf{x}_1, \dots, \mathbf{x}_k)$ be a bounded Borel measurable function, i.e., $|h(\mathbf{x}_1, \dots, \mathbf{x}_k)| \leq C_1$, then

$$|E(h(\mathbf{Y}_{i_1}, \dots, \mathbf{Y}_{i_k})) - \int \dots \int h(\mathbf{x}_1, \dots, \mathbf{x}_k) dF^{(1)}(\mathbf{x}_1, \dots, \mathbf{x}_j) dF^{(2)}(\mathbf{x}_{j+1}, \dots, \mathbf{x}_n)| \leq 2C_1\beta(i_{j+1} - i_j)$$

where $F^{(1)}$ and $F^{(2)}$ are respectively d.f.'s of $(\mathbf{Y}_{i_1}, \dots, \mathbf{Y}_{i_j})$ and $(\mathbf{Y}_{i_{j+1}}, \dots, \mathbf{Y}_{i_k})$ for $i_1 < i_2 < \dots < i_k$.

This Lemma is an extension of Lemma 2.1 of Yoshihara [18] and is proved in Harel and Puri [9].

Lemma 6.2: (Takahata and Yoshihara [14]). *Let $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ be a random vector as in Lemma 6.1. Let $h(\mathbf{y}, \mathbf{z})$ be a Borel measurable function such that $|h(\mathbf{y}, \mathbf{z})| \leq C_1$ for all \mathbf{y} and \mathbf{z} . Let \mathbf{Z}_1 be a $\sigma(\mathbf{X}_i; 1 \leq i \leq k)$ -measurable random variable, and \mathbf{Z}_2 be a $\sigma(\mathbf{X}_i; i \geq k + m)$ -measurable random variable. Further let $H(\mathbf{y}) = E(h(\mathbf{y}, \mathbf{Z}_2))$ then*

$$E | E\{h(\mathbf{Z}_1, \mathbf{Z}_2) | \sigma(\mathbf{X}_i; 1 \leq i \leq k)\} - H(\mathbf{Z}_1) | \leq 2C_1\beta(m).$$

Lemma 6.3: (Davydov [3]). *Let $\mathbf{Y}_1, \dots, \mathbf{Y}_n$ be \mathbb{R}^p -valued random vectors satisfying a strong mixing condition with rate $\alpha(m)$. If $\|\mathbf{Y}_i\|_s$ exists for all i and $s > 2$ and $E(\mathbf{Y}_i) = 0, i \geq 1$, then*

$$E |\mathbf{Y}_i \mathbf{Y}_j| \leq C_2 \alpha^{1 - \frac{1}{s} - \frac{1}{t}}(j - i) \|\mathbf{Y}_i\|_s \|\mathbf{Y}_j\|_t, i \leq j, s > 2, t > 2 \tag{6.1}$$

where C_2 is a constant > 0 , and of course, if the sequence $\{\mathbf{Y}_i\}_{i \geq 1}$ is absolutely regular with rate $\beta(m)$, the inequality (6.1) holds when we replace $\alpha^{1 - \frac{1}{s} - \frac{1}{t}}(j - 1)$ by $\beta^{1 - \frac{1}{s} - \frac{1}{t}}(j - i)$.

In what follows, we always assume that the conditions of Theorem 4.1 are satisfied and C denotes a universal constant.

Let $\{\tilde{\mathbf{X}}_i\}_{i \geq 1}$ be independent random vectors each having the same d.f. as that of \mathbf{X}_i .

Put

$$G_{j,k}^{(i)}(\mathbf{x}, \mathbf{y}) = \int H_{i,j}(\mathbf{u}, \mathbf{x}) H_{i,k}(\mathbf{y}) dF_i(\mathbf{u})$$

where $H_{i,j}$ is defined in (4.5), and

$$\mathbf{Y}_{\alpha,i}(\mathbf{x}) = \sum_{j=1}^{a_\alpha - m} H_{i,j}(\mathbf{x}, \mathbf{X}_j), \alpha = 1, \dots, k.$$

Let \mathbf{Q}_α be the distribution of $(\mathbf{X}_{a_\alpha}, \dots, \mathbf{X}_{b_\alpha})$.

From Hall [8] and Lemma 6.1, the following are easily obtained:

$$EH_{i,j}(\tilde{\mathbf{X}}_i, \tilde{\mathbf{X}}_j) = 0, E(H_{i,j}(\tilde{\mathbf{X}}_i, \tilde{\mathbf{X}}_j) | \tilde{\mathbf{X}}_i) = 0 \text{ a.s.} \tag{6.2}$$

$$1 \leq \max_{i \neq j \leq n} E(H_{i,j}^{2k}(\tilde{\mathbf{X}}_i, \tilde{\mathbf{X}}_j)) = O(h^{2kp+1}) \tag{6.3}$$

$$1 \leq \max_{i,j \leq n} \sup_{x,y} H_{i,j}(\mathbf{x}, \mathbf{y}) = O(h^p) \tag{6.4}$$

$$1 \leq \max_{i,j \leq n} E(H_{i,j}^2(\mathbf{X}_i, \mathbf{X}_j)) \leq Ch^{2p}, \tag{6.5}$$

$$|E(H_{i,j}(\mathbf{X}_i, \mathbf{X}_j))| \leq Ch^p\beta(|i-j|) \text{ for all } i \text{ and } j. \tag{6.6}$$

$$E(G_{j,k}^{(i)}(\tilde{\mathbf{X}}_j, \tilde{\mathbf{X}}_k)) = 0, E([G_{j,k}^{(i)}(\tilde{\mathbf{X}}_j, \tilde{\mathbf{X}}_k)]^2) = O(h^{7p}) \tag{6.7}$$

$$|E(G_{j,k}^{(i)}(\mathbf{X}_j, \mathbf{X}_k))| \leq C\beta^{\frac{1}{2}}(|k-j|)h^{7p/2}. \tag{6.8}$$

The proofs of Lemmas 6.4-6.7 are in general similar to the proofs of Lemmas 5-8 in Takahata and Yoshihara [15]. For reasons of brevity and to avoid repetitious arguments, we give brief outlines of the proofs.

Lemma 6.4: As $n \rightarrow \infty$

$$s_n^2 \sim \frac{1}{2}n^2h^{3p}\sigma_3^2 \tag{6.9}$$

where \sim means that the ration of the two sides $\rightarrow 1$ as $n \rightarrow \infty$.

Proof: We have

$$s_n^2 = E\left(\sum_{\alpha=1}^k T_\alpha\right)^2 - \left\{E\left(\sum_{\alpha=1}^k T_\alpha\right)\right\}^2$$

where T_α is defined in (4.7'). By (6.5) we note that

$$|E\left(\sum_{\alpha=1}^k T_\alpha\right)| \leq \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-m} |E(H_{i,j}(\mathbf{X}_i, \mathbf{X}_j))| \leq Cn^2\beta(m) = o(n^{-6}). \tag{6.10}$$

Now consider

$$E\left(\sum_{\alpha=1}^k T_\alpha\right)^2 = \sum_{\alpha=1}^k E(T_\alpha)^2 + 2 \sum_{a \geq \alpha \leq \alpha' \leq k} E(T_\alpha T_{\alpha'}) = I_{11} + I_{12}.$$

By Lemma 6.1,

$$E(T_\alpha T_{\alpha'}) \leq Cr^2n^2\beta(m)$$

which implies

$$I_{12} = o(n^{-1}). \tag{6.11}$$

By Lemma 6.1, we have

$$\begin{aligned} I_{11} &= E(T_\alpha^2) = E\left(\sum_{i=a_\alpha}^{b_\alpha} \mathbf{Y}_{\alpha,i}(\mathbf{X}_i)\right)^2 \sim E\left(\int\left(\sum_{i=a_\alpha}^{b_\alpha} \mathbf{Y}_{\alpha,i}(\mathbf{x}_i)\right) d\mathbf{Q}_\alpha(\mathbf{x}_{a_\alpha}, \dots, \mathbf{x}_{b_\alpha})\right)^2 + Cr^2n^2\beta(m) \\ &= \sum_{i=a_\alpha}^{b_\alpha} \int E(\mathbf{Y}_{\alpha,i}^2(\mathbf{x}_i))dF_i(\mathbf{x}_i) + 2 \sum_{a_\alpha \leq i < j \leq b_\alpha} \int E(\mathbf{Y}_{\alpha,i}(\mathbf{x}_i)(\mathbf{Y}_{\alpha,j}(\mathbf{x}_j)))d\mathbf{Q}_\alpha(\mathbf{x}_{a_\alpha}, \dots, \mathbf{x}_{b_\alpha}) + o(n^{-5}) \end{aligned}$$

$$= J_{1\alpha} + J_{2\alpha} + o(n^{-5}).$$

By (6.8) we have

$$\begin{aligned} J_{1,\alpha} &= \sum_{i=a_\alpha}^{b_\alpha} \left[\sum_{j=1}^{a_\alpha-m} \int E(H_{i,j}^2(\mathbf{x}_i, \mathbf{X}_j)) dF_i(\mathbf{x}_i) \right. \\ &\quad \left. + 2 \sum_{1 \leq j < j' \leq a_\alpha-m} \int E(H_{i,j}(\mathbf{x}_i, \mathbf{X}_j) H_{i,j'}(\mathbf{x}_i, \mathbf{X}_{j'})) dF_i(\mathbf{x}_i) \right] \\ &\leq \sum_{i=a_\alpha}^{b_\alpha} \left[\sum_{j=1}^{a_\alpha-m} E(H_{i,j}^2(\tilde{\mathbf{X}}_i, \tilde{\mathbf{X}}_j)) + 2 \sum_{1 \leq j < j' \leq a_\alpha-m} E(G_{j,j'}^{(i)}(\mathbf{X}_j, \mathbf{X}_{j'})) \right] \\ &= \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-m} E(H_{i,j}^2(\tilde{\mathbf{X}}_i, \tilde{\mathbf{X}}_j)) + O(h^{7p/2}). \end{aligned}$$

From Lemma 3 of Hall [8], we obtain

$$E(H_{i,j}^2(\tilde{\mathbf{X}}_i, \tilde{\mathbf{X}}_j)) \sim h^{3p} \left\{ \int f_{i,j}(\mathbf{x}) f_{j,j}(\mathbf{x}) d\mathbf{x} \right\} \left[\int \{K(\mathbf{u})K(\mathbf{u} + \mathbf{v})\}^2 d\mathbf{v} \right].$$

Thus

$$\begin{aligned} J_{1,\alpha} &\sim h^{3p} \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-m} \left\{ \int f_{i,j}(\mathbf{x}) f_{j,j}(\mathbf{x}) d\mathbf{x} \right\} \left[\int \{K(\mathbf{u})K(\mathbf{u} + \mathbf{v})\}^2 d\mathbf{v} \right] + O(h^{7p/2}) \\ &= h^{3p} \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-m} D_{i,j} + O(h^{7p/2}). \end{aligned}$$

On the other hand, by Lemma 6.1, (6.2), (6.5) and (6.7) we get

$$|J_{2\alpha}| \leq o(nr h^{3p}).$$

Thus

$$I_{11} = \sum_{\alpha=1}^k (J_{1\alpha} + J_{2\alpha}) = h^{3p} \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-m} D_{i,j} + O(h^{7p/2}) + o(nr h^{3p})$$

and from condition (2.4) and Lemma 6.8, we can obtain

$$I_{11} = \frac{n^2}{2} h^{3p} \sigma_3^2 (1 + o(1)). \tag{6.12}$$

Now (6.9) follows from (6.11) and (6.12) and the proof is complete.

Lemma 6.5:

$$s_n^{-1} \sum_{\alpha=1}^k E\{T_\alpha | \mathfrak{F}_\alpha\} \rightarrow 0 \text{ in probability as } n \rightarrow \infty. \tag{6.13}$$

$$s_n^{-2} \sum_{\alpha=1}^k E\{T_\alpha^2 | \mathfrak{F}_\alpha\} - (E\{T_\alpha | \mathfrak{F}_\alpha\})^2 \rightarrow 1 \text{ in probability as } n \rightarrow \infty. \tag{6.14}$$

Proof: By Lemma 6.2 and (6.2), we obtain

$$E \left| \sum_{\alpha=1}^k E\{T_\alpha | \mathcal{F}_\alpha\} \right| \leq \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=1}^{a_\alpha-m} \{ |E(H_{i,j}(\mathbf{X}_i, \mathbf{X}_j))| + C\beta(m) \} \\ \leq Cn^2\beta(m) = o(n^{-1}s_n).$$

(6.13) follows.

To prove (6.14), it suffices to show that

$$I_{21} = s_n^{-2} \sum_{\alpha=1}^k E | E\{T_\alpha^2 | \mathcal{F}_\alpha\} |^2 - (ET_\alpha)^2 \rightarrow 0 \text{ as } n \rightarrow \infty, \tag{6.15}$$

and

$$I_{12} = s_n^{-2} \sum_{\alpha=1}^k [E | E\{T_\alpha | \mathcal{F}_\alpha\} |^2 - (ET_\alpha)^2] \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{6.16}$$

(6.15) follows since by Lemmas 6.1 and 6.2, we obtain after some computations that

$$I_{21} \leq Cn^3r\beta(m) = o(n^{-1}s_n^2).$$

On the other hand, by Lemma 1 (after some computations) we get

$$|E[E\{H_{i,j}(\mathbf{X}_i, \mathbf{X}_j) | \mathcal{F}_\alpha\} E\{H_{\ell,p}(\mathbf{X}_\ell, \mathbf{X}_p) | \mathcal{F}_\alpha\}]| \leq C\beta(m)$$

which implies

$$\sum_{\alpha=1}^k E(E(T_\alpha | \mathcal{F}_\alpha))^2 \leq Cn^3r\beta(m) + o(n^{-1}s_n^2) \tag{6.17}$$

(6.16) follows from (6.10) and (6.17).

Lemma 6.6:

$$s_n^{-2} E(S_n - U_n)^2 \rightarrow 0 \text{ as } n \rightarrow \infty. \tag{6.18}$$

Proof: Since

$$S_n - U_n = \sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=a_\alpha-m+1}^{a_\alpha-1} \{H_{i,j}(\mathbf{X}_i, \mathbf{X}_j) - E(H_{i,j}(\mathbf{X}_i, \mathbf{X}_j))\}.$$

The proof follows by showing that

$$\sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=a_\alpha-m+1}^{a_\alpha-1} |E(H_{i,j}(\mathbf{X}_i, \mathbf{X}_j))| \leq cn^{3/4}(\log n)^2 h^{3p/2} \tag{6.19}$$

and

$$E \left(\sum_{\alpha=1}^k \sum_{i=a_\alpha}^{b_\alpha} \sum_{j=a_\alpha-m+1}^{a_\alpha-1} H_{i,j}(\mathbf{X}_i, \mathbf{X}_j) \right) \leq 2k^2c[mr + r^2m^2\beta m] + m^4 = o(s_n^2). \tag{6.20}$$

Lemma 6.7:

$$\sum_{\alpha=1}^k E(T_\alpha)^4 = o(s_n^r) \text{ as } n \rightarrow \infty. \tag{6.21}$$

Proof: Since from (4.7'), $|T_\alpha^4| \leq Cn^4r^4$, it follows from Lemma 6.1 that

$$\begin{aligned}
 E(T_\alpha)^4 &\leq E \left\{ \int \left(\sum_{i=a_\alpha}^{b_\alpha} Y_{\alpha,i}(\mathbf{x}_i) \right)^4 d\mathbf{Q}_\alpha \right\} + Cn^4r^4\beta(m) \\
 &= \sum_{i=a_\alpha}^{b_\alpha} E(Y_{\alpha,i}^4(\mathbf{x}_i))d\mathbf{Q}_\alpha + \int_{a_\alpha \leq i, i' \leq b_\alpha} \int E(Y_{\alpha,i}^3(\mathbf{x}_i)Y_{\alpha,i'}(\mathbf{x}_{i'}))d\mathbf{Q}_\alpha \\
 &\quad + 2 \sum_{\substack{a_\alpha \leq i, i' \leq b_\alpha \\ i \neq i'}} \int E(Y_{\alpha,i}^2(\mathbf{x}_i)Y_{\alpha,i'}^2(\mathbf{x}_{i'}))d\mathbf{Q}_\alpha \\
 &\quad + \sum_{\substack{a_\alpha \leq i_1, i_2, i_3 \leq b_\alpha \\ i_s \neq i_\ell (s \neq \ell)}} \int E(Y_{\alpha,i_1}^2(\mathbf{x}_{i_1})Y_{\alpha,i_2}(\mathbf{x}_{i_2})Y_{\alpha,i_3}(\mathbf{x}_{i_3}))d\mathbf{Q}_\alpha \\
 &\quad + \sum_{\substack{a_\alpha \leq i_1, i_2, i_3, i_4 \leq b_\alpha \\ i_s \neq i_\ell (s \neq \ell)}} \int E\left(\prod_{d=1}^4 Y_{\alpha,i_d}(\mathbf{x}_{i_d})\right)d\mathbf{Q}_\alpha + o(n^{-1}) \\
 &= I_{\alpha,1} + I_{\alpha,2} + I_{\alpha,3} + I_{\alpha,4} + I_{\alpha,5} + o(n^{-1}).
 \end{aligned}$$

Using Lemmas 6.1 and 6.4, Hölder's inequality and Schwartz's inequality, we get after some computations

$$\sum_{\alpha=1}^k I_{\alpha,j} = o(s_n^4), \quad 1 \leq j \leq p$$

which implies (6.21).

Lemma 6.8: (Cacoullos [1]). *Suppose $M(\mathbf{y})$ is a Borel scalar function on \mathbb{R}^p such that*

$$\sup_{\mathbf{y} \in \mathbb{R}^p} |M(\mathbf{y})| < \infty \tag{6.22}$$

$$\int |M(\mathbf{y})| d\mathbf{y} < \infty \tag{6.23}$$

$$\lim_{|\mathbf{y}| \rightarrow \infty} |\mathbf{y}|^p M(\mathbf{y}) = 0. \tag{6.24}$$

Let $g(\mathbf{y})$ be another scalar function on \mathbb{R}^p such that

$$\int |g(\mathbf{y})| d\mathbf{y} < \infty. \tag{6.25}$$

and define

$$g_n(\mathbf{x}) = (h^{-p}) \int M\left(\frac{\mathbf{y}}{h}\right)g(\mathbf{x} - \mathbf{y})d\mathbf{y}.$$

Then at every point \mathbf{x} of continuity of g

$$\lim_{n \rightarrow \infty} g_n(\mathbf{x}) = g(\mathbf{x}) \int M(\mathbf{y}) d\mathbf{y}. \tag{6.26}$$

Proof: Choose $\delta > 0$ and split the region of integration in two regions $|\mathbf{y}| \leq \delta$ and $|\mathbf{y}| > \delta$. Then we have

$$\begin{aligned} |g_n(\mathbf{x}) - g(\mathbf{x}) \int M(\mathbf{y}) d\mathbf{y}| &= |h^{-p} \int [g(\mathbf{x} - \mathbf{y}) - g(\mathbf{x})] M\left(\frac{\mathbf{y}}{h}\right) d\mathbf{y}| \\ &\leq \max_{|\mathbf{y}| \leq \delta} |g(\mathbf{x} - \mathbf{y}) - g(\mathbf{x})| \int_{|\mathbf{z}| \leq \delta/h} |M(\mathbf{z})| d\mathbf{z} \\ &+ \int_{|\mathbf{y}| > \delta} \frac{|g(\mathbf{x} - \mathbf{y})|}{|\mathbf{y}|^p} \frac{|\mathbf{y}|^p}{h^p} |M\left(\frac{\mathbf{y}}{h}\right)| d\mathbf{y} + |g(\mathbf{x})| \int_{|\mathbf{y}| > \delta} h^{-p} |M\left(\frac{\mathbf{y}}{h}\right)| d\mathbf{y} \\ &\leq \max_{|\mathbf{y}| \leq \delta} |g(\mathbf{x} - \mathbf{y}) - g(\mathbf{x})| \int |M(\mathbf{z})| d\mathbf{z} \\ &+ \delta^{-p} \sup_{|\mathbf{z}| \geq \delta/h} |\mathbf{z}|^p |M(\mathbf{z})| \int |g(\mathbf{y})| d\mathbf{y} + |g(\mathbf{x})| \int_{|\mathbf{z} > \delta/h} |M(\mathbf{z})| d\mathbf{z} \\ &= I_1 + I_2 + I_3. \end{aligned}$$

From the continuity of g at x and (6.23), I_1 tends to 0 if we let first $n \rightarrow \infty$ and then $\delta \rightarrow 0$. From (6.24) and (6.25), I_2 tends to 0 and from (6.23), I_3 tends to 0 as $n \rightarrow \infty$. The proof follows.

Lemma 6.9: *We have*

$$E(K_j^6) = O(h^{6(p+2)}), \quad j \geq 1 \tag{6.27}$$

where K_j is defined in (4.6).

Proof: Define

$$B_j = \int K\left(\frac{x - \mathbf{X}_j}{h}\right) \{E(f_n^*(x)) - f(x)\} dx.$$

Then, we have for any $k \geq 1$

$$\begin{aligned} |E(B_j^k)| &\leq \left| \int_{\mathbb{R}^k} \left\{ \prod_{i=1}^k |E f_n^*(x^{(i)}) - f(x^{(i)})| \int_{\left[\prod_{i=1}^k K\left(\frac{x^{(i)} - z}{h}\right) \right]} f_j(z) dz \right\} dx^{(1)} \dots dx^{(k)} \right| \\ &= \left| \int_{\mathbb{R}^k} \left\{ (O(n^{-1}) + h^2 \tau \Delta f(x) + o(h^3))^k \int_{\left[\prod_{i=1}^k K\left(\frac{x^{(i)} - z}{h}\right) \right]} f_j(z) dz \right\} dx^{(1)} \dots dx^{(k)} \right| \\ &\leq C h^{2k} \left| \int_{\mathbb{R}^k} \left\{ \left[\prod_{i=1}^k K\left(\frac{x^{(i)} - z}{h}\right) \right] f_j(z) dz \right\} dx^{(1)} \dots dx^{(k)} \right| \end{aligned}$$

$$\leq Ch^{k(p+2)} \text{ where } C \text{ is some constant } > 0.$$

The desired results follow immediately on noting that

$$\begin{aligned} E(K_j^6) &= E(B_j^6) - 6E(B_j^5)E(B_j) + 15E(B_j^4)E^2(B_j) - 20E(B_j^3)E^3(B_j) \\ &\quad + 15E(B_j^2)E^4(B_j) - 5E^6(B_j). \end{aligned}$$

References

- [1] Cacoullos, T., Estimation of a multivariate density, *Ann. Inst. Statist. Math.* **18** (1966), 179-189.
- [2] Csörgö, M. and Révész, P., *Strong Approximations in Probability and Statistics*, Academic Press, New York 1981.
- [3] Davydov, Yu.A., Invariance principle for empirical stationary processes, *Theory Prob. Appl.* **14** (1970), 487-498.
- [4] Davydov, Yu.A., Mixing conditions for Markov chains, *Theor. Prob. Appl.* **18** (1973), 312-328.
- [5] Devroyes, L. and Györfi, L., *Nonparametric Density Estimation: The L_1 View*, Wiley, New York 1984.
- [6] Doukhan, P. and Ghindès, M., Estimation de la transition de probabilité d'une chaîne de Markov Doeblin-récurrente. Étude du cas du processus autorégressif général d'ordre 1, *Stoch. Processes Appl.* **15** (1983), 271-293.
- [7] Dvoretzky, A., Central limit theorems for dependent random variables, In: *Proc. Sixth Berkeley Symp. Math. Statist. Probab.* (ed. by L. LeCam, et al.), University of California Press, Los Angeles, CA **2** (1972), 513-555.
- [8] Hall, P., Central limit theorem for integrated square error of multivariate nonparametric density estimators, *J. Multivariate Anal.* **14** (1984), 1-16.
- [9] Harel, M. and Puri, M.L., Limiting behavior of U -statistics, V -statistics and one sample rank order statistics for nonstationary absolutely regular processes, *J. Multivariate Anal.* **30** (1989), 181-204.
- [10] Nummelin, E. and Tuominen, P., Geometric ergodicity of Harris recurrent Markov chains with applications to renewal theory, *Stoch. Processes Appl.* **12** (1982), 187-202.
- [11] Rosenblatt, M., A quadratic measure of derivation of two-dimensional density estimates and a test of independence, *Ann. Statist.* **3** (1975), 1-14.
- [12] Roussas, G., Nonparametric estimation of the transition distribution of a Markov process, *Ann. Math. Statist.* **40:4** (1969), 1386-1400.
- [13] Roussas, G. and Tran, L., Joint asymptotic normality of kernel estimates under dependence conditions, with application to hazard rate, *Nonparametric Statistics* **1** (1992), 335-355.
- [14] Takahata, H. and Yoshihara, K.I., Asymptotic normality of a recursive stochastic algorithm with observations satisfying some absolute regularity conditions, *Yokohama Math. Jour.* **33** (1985), 139-159.
- [15] Takahata, M. and Yoshihara, K.I., Central limit theorems for integrated square error of nonparametric density estimators based on absolutely regular random sequences, *Yokohama Math. Jour.* **35** (1987), 95-111.
- [16] Tran, L.T., Density estimation under dependence, *Statist. Probab. Lett.* **10** (1990), 193-201.
- [17] Tran, L.T., Kernel density estimation for linear processes, *Stoch. Processes Appl.* **41** (1992), 281-296.

- [18] Yoshihara, K.I., Limiting behavior of U -statistics for stationary absolutely regular processes, *Z. Wahrscheinlichkeitstr.* **35** (1976), 237-252.