JOURNAL OF MULTIVARIATE ANALYSIS 44, 47–68 (1993)

Asymptotic Normality for Deconvolution Estimators of Multivariate Densities of Stationary Processes*

ELIAS MASRY

University of California at San Diego

We consider the estimation of the multivariate probability density functions of stationary random processes from noisy observations. The asymptotic normality of kernel-type deconvolution estimators is established for various classes of mixing processes. Classes of noise characteristic functions both with algebraic and with exponential decay are studied. © 1993 Academic Press, Inc.

I. INTRODUCTION

Let $\{X_i\}_{i=-\infty}^{\infty}$ be a stationary process and for each integer $p \geqslant 1$ let $f(\mathbf{x}; p) = f(x_1, ..., x_p; p)$ be the joint probability density function of the random variables $X_1, ..., X_p$ which is assumed to exist. Consider the deconvolution problem

$$Y_i = X_i + \varepsilon_i, \tag{1.1}$$

where the (noise) process $\{\varepsilon_i\}_{i=-\infty}^{\infty}$ consists of i.i.d. random variables, independent of the process $\{X_i\}_{i=-\infty}^{\infty}$, with known density $\bar{h}(x)$. Let $g(\mathbf{x}; p)$ be the joint probability density function of the random variables $Y_1, ..., Y_p$, which is given by the multidimensional convolution

$$g(\mathbf{x}; p) = \int_{R^p} f(\mathbf{x} - \mathbf{u}; p)\, h(\mathbf{u})\, d\mathbf{u}, \tag{1.2a}$$

where

$$h(\mathbf{u}) = \prod_{j=1}^{p} \bar{h}(u_j). \tag{1.2b}$$

Received April 2, 1991; revised April 13, 1992.

AMS (1991) subject classifications: 62G07, 62H10, 60F05.

Key words and phrases: Deconvolution of multivariate densities, mixing processes, asymptotic normality.

* This work was supported by the Office of Naval Research under Grant N00014-90-J-1175.

0047-259X/93 $5.00

The deconvolution problem is to estimate consistently the multivariate density $f(\mathbf{x}; p)$ from the noisy observations $\{Y_i\}_{i=1}^n$. The deconvolution problem arises in biological studies (Medgyessy [15], Mendelsohn and Rice [16]), communication theory (Wise *et al.* [23], Snyder *et al.* [20]), and applied physics and analytical chemistry (Jones and Misell [9], Harder and Galan [8]).

The case of i.i.d. observations (the X_i's are i.i.d. random variables with probability density function $f(x)$) has received considerable attention in recent years with contributions by Carroll and Hall [2], Liu and Taylor [11], Stefanski and Caroll [21], Zhang [24], and Fan [5]. The principal focus of these works is to provide bounds on the rate of quadratic-mean convergence for kernel-type estimates of $f(x)$. Recently Fan [6] established the asymptotic normality of kernel-based estimates of $f(x)$.

The general case of stationary processes $\{X_i\}_{i=-\infty}^{\infty}$ was considered recently by Masry [13], where bounds, as well as asymptotic expressions, for the mean-square error of kernel-type estimators of the multivariate probability densities $f(\mathbf{x}; p)$, $p \geqslant 1$, are established for several classes of mixing processes $\{X_i\}_{i=-\infty}^{\infty}$. In Masry [14], sharp rates of almost sure convergence for estimators of $f(\mathbf{x}; p)$ are established.

The purpose of this paper is to establish the asymptotic normality of kernel-type multivariate density estimators $\hat{f}_n(\mathbf{x}; p)$ based on the noisy observations $\{Y_i\}_{i=1}^n$. Asymptotic distributions are obviously useful for constructing confidence intervals. The principal results of the paper are given in Theorem 2.1 and Theorem 3.1 which are established for ρ-mixing and strongly mixing processes under mild conditions on the mixing coefficients. Section II considers the case where the tail of the noise characteristic function decays algebraically (e.g., gamma and Laplacian noise densities). In this case the asymptotic rate and constant of $\text{var}[\hat{f}_n(\mathbf{x}; p)]$ are known (Masry [13]), and therefore a central limit theorem having a classical form can be established (Theorem 2.1). In Section III we consider the case where the tail of the noise characteristic function decays exponentially fast (e.g., Gaussian and Cauchy noise densities); here no asymptotic expression for $\text{var}[\hat{f}_n(\mathbf{x}; p)]$ is known even in the i.i.d. case and only a weak version of the central limit theorem can be established (Theorem 3.1).

We first introduce some notation. Denote the characteristic functions of $f(\mathbf{x}; p)$, $g(\mathbf{x}; p)$, and $\bar{h}(x)$ by $\phi_f(\mathbf{t})$, $\phi_g(\mathbf{t})$, and $\bar{\phi}_h(t)$, respectively. Then

$$\phi_g(\mathbf{t}) = \phi_f(\mathbf{t})\, \phi_h(\mathbf{t}); \qquad \phi_h(\mathbf{t}) = \prod_{j=1}^{p} \bar{\phi}_h(t_j). \tag{1.3}$$

For simplicity we select product-type kernels as follows: Let $\bar{K}(x)$ be a real-valued, even, bounded density function on the real line satisfying

$\bar{K}(x) = O(|x|^{-1-\delta})$ for some $\delta > 0$ and denote its Fourier transform by $\bar{\phi}_K(t)$. For every $b > 0$ define

$$\bar{W}_b(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} [\bar{\phi}_K(t)/\bar{\phi}_h(t/b)]\, dt. \tag{1.4}$$

(Assumptions will be made on $\bar{\phi}_K(t)$ and $\bar{\phi}_h(t)$ which ensure that $\bar{\phi}_K(t)/\bar{\phi}_h(t/b) \in L_1 \cap L_\infty$.) Set

$$K(\mathbf{x}) = \prod_{j=1}^{p} \bar{K}(x_j); \qquad W_b(\mathbf{x}) = \prod_{j=1}^{p} \bar{W}_b(x_j) \tag{1.5}$$

so that

$$\phi_K(\mathbf{t}) = \prod_{j=1}^{p} \bar{\phi}_K(t_j). \tag{1.6}$$

The choice of product-type kernels in (1.5) is not essential, as shown in Masry [13], and is made for the sake of simplicity.

Let $\{b_n\}_{n=1}^{\infty}$ be a sequence of positive numbers such that $b_n \to 0$ as $n \to \infty$. Given the observations $\{Y_i\}_{i=1}^{n}$ we estimate $f(\mathbf{x}; p)$ by

$$\hat{f}_n(\mathbf{x}; p) = \frac{1}{(n-p+1)\, b_n^p} \sum_{j=0}^{n-p} W_{b_n}[(\mathbf{x} - \mathbf{Y}_j)/b_n], \tag{1.7}$$

where

$$\mathbf{Y}_j = (Y_{j+1}, ..., Y_{j+p})$$

and it is assumed of course that $n > p$. An alternative expression for $\hat{f}_n(\mathbf{x}; p)$ is

$$\hat{f}_n(\mathbf{x}; p) = (1/2\pi)^p \int_{R^p} e^{-i\mathbf{t}\cdot\mathbf{x}} \hat{\phi}_{g,n}(\mathbf{t}) \frac{\phi_K(b_n \mathbf{t})}{\phi_h(\mathbf{t})}\, d\mathbf{t},$$

where $\hat{\phi}_{g,n}(\mathbf{t})$ is the standard estimate of the characteristic function $\phi_g(\mathbf{t})$,

$$\hat{\phi}_{g,n}(\mathbf{t}) = \frac{1}{n-p+1} \sum_{j=0}^{n-p} e^{i\mathbf{t}\cdot\mathbf{Y}_j}$$

and $\mathbf{u} \cdot \mathbf{v} = \sum_{j=1}^{p} u_j v_j$.

The bias of the estimator $\hat{f}_n(\mathbf{x}; p)$ is given by

PROPOSITION 1.1. (a) *For almost all* $\mathbf{x} \in R^p$ *we have*

$$E[\hat{f}_n(\mathbf{x}; p)] \to f(\mathbf{x}; p) \qquad \text{as} \quad n \to \infty.$$

(b) *If $f(\mathbf{x}; p)$ is twice differentiable and its second partial derivatives are bounded and continuous on R^p and the kernel $\bar{K}(x)$ satisfies $\int_{-\infty}^{\infty} u^2\bar{K}(u)\,du < \infty$, then*

$$(1/b_n^2)\,\text{bias}[\hat{f}_n(\mathbf{x}; p)] \to \frac{1}{2}\int_{R^p} \mathbf{u}G''(\mathbf{x}; p)\,\mathbf{u}^{\mathrm{T}}K(\mathbf{u})\,d\mathbf{u}$$

as $n \to \infty$, where the $p \times p$ matrix G'' is given by

$$G''(\mathbf{x}; p) = \left[\left(\frac{\partial^2 f(\mathbf{x}; p)}{\partial x_i\,\partial x_j}\right)_{ij}\right]$$

and $\mathbf{u}^{\mathrm{T}}$ is the transpose of the row vector $\mathbf{u}$.

II. Asymptotic Normality for the Class of $\bar{\phi}_h(t)$ with Algebraic Decay

Let $\mathscr{F}_i^k$ be the σ-algebra of events generated by the random variables $\{X_j, \varepsilon_j, i \leqslant j \leqslant k\}$ and let $L_2(\mathscr{F}_a^b)$ denote the collection of all second-order random variables which are $\mathscr{F}_a^b$-measurable. The stationary processes $\{X_i, \varepsilon_i\}$ are called strongly mixing (Rosenblatt, [18]) if

$$\sup_{\substack{A \in \mathscr{F}_{-\infty}^0,\\ B \in \mathscr{F}_k^{\infty}}} |P[AB] - P[A]\,P[B]| = \alpha(k) \to 0 \qquad \text{as} \quad k \to \infty, \tag{2.1a}$$

and are called ρ-mixing (Kolmogorov and Rozanov, [10]) if

$$\sup_{\substack{U \in L_2(\mathscr{F}_{-\infty}^0),\\ V \in L_2(\mathscr{F}_k^{\infty})}} \frac{|\text{cov}\{U, V\}|}{\text{var}^{1/2}[U]\,\text{var}^{1/2}[V]} = \rho(k) \to 0 \qquad \text{as} \quad k \to \infty, \tag{2.1b}$$

where $\alpha(k)$ is the strong mixing coefficient and $\rho(k)$ is the maximal correlation coefficient. It is well known that $\alpha(k) \leqslant \frac{1}{4}\rho(k)$.

An asymptotic expression for the variance of $\hat{f}_n(\mathbf{x}; \rho)$ was established in Masry [13] for the case where the noise characteristic function $\bar{\phi}_h(t)$ decays algebraically as $|t| \to \infty$. This includes, in particular, gamma and Laplacian noise probability density functions $\bar{h}(x)$. We make the following assumption on the characteristic functions $\bar{\phi}_h(t)$ and $\bar{\phi}_K(t)$.

Assumptions 2.1. $\bar{\phi}_h(t)$ and $\bar{\phi}_K(t)$ are twice continuously differentiable with bounded derivations such that

(i) $|\bar{\phi}_h(t)| > 0$ for all $t \in R$.

(ii) $t^{\beta}\bar{\phi}_h(t) \to B_1$ as $t \to \infty$ for some $\beta \geqslant 1$ and $|B_1| > 0$.

(iii) $\int_{-\infty}^{\infty} |t|^{\beta-2} |\bar{\phi}_K(t)| \, dt < \infty$ for $\beta > 1$; $\int_{-\infty}^{\infty} |t|^{2\beta} |\bar{\phi}_K(t)|^2 \, dt < \infty$.

(iv) $\int_{-\infty}^{\infty} |t|^{\beta-1} |\bar{\phi}'_K(t)| \, dt < \infty$; $\int_{-\infty}^{\infty} |t|^{\beta} |\bar{\phi}''_K(t)| \, dt < \infty$.

The parameter $\beta \geqslant 1$ is called the order of the noise density $\bar{h}(x)$. We make the following assumption on the processes $\{X_i, \varepsilon_i\}_{i=-\infty}^{\infty}$.

Assumption 2.2. (i) The probability density function $g(\mathbf{y}; q)$ of the vector $(Y_1, ..., Y_q)$ exists and is bounded by a constant for all $1 \leqslant q \leqslant 2p$.

(ii) The $2p$-dimensional probability density function $g(\mathbf{x}, \mathbf{y}; 2p, j)$ of the vectors $\mathbf{Y}_0$ and $\mathbf{Y}_j$, $j \geqslant p$, exists and

$$\sup_{\mathbf{x}, \mathbf{y}, j \geqslant p} |g(\mathbf{x}, \mathbf{y}; 2p, j) - g(\mathbf{x}; p) \, g(\mathbf{y}; p)| \leqslant M < \infty. \tag{2.2}$$

(iii) Either the processes $\{X_i, \varepsilon_i\}_{i=-\infty}^{\infty}$ are ρ-mixing with

$$\sum_{j=1}^{\infty} \rho(j) < \infty \tag{2.3}$$

or they are strongly mixing with

$$\sum_{j=1}^{\infty} j^a [\alpha(j)]^{1-2/v} < \infty \text{ for some } v > 2 \text{ and } a > 1 - 2/v. \tag{2.4}$$

Put

$$\mu_n = \frac{1}{b_n^p} E[W_{b_n}[(\mathbf{x} - \mathbf{Y}_j)/b_n]] \tag{2.5}$$

$$\tilde{Z}_{n,j} = \frac{1}{b_n^p} W_{b_n}[(\mathbf{x} - \mathbf{Y}_j)/b_n] - \mu_n. \tag{2.6}$$

We then have (Masry, [13, Theorem 3 and its proof]) the following.

LEMMA 2.1. *Under Assumption* 2.1 *and* 2.2 *we have*

(a) $\lim_{n \to \infty} b_n^{(2\beta+1)p} \operatorname{var}[\tilde{Z}_{n,i}] = \sigma^2(\mathbf{x})$

(b) $b_n^{(2\beta+1)p} \sum_{l=1}^{n} |\operatorname{cov}\{\tilde{Z}_{n,0}, \tilde{Z}_{n,l}\}| = o(1)$

(c) $\lim_{n \to \infty} n b_n^{(2\beta+1)p} \operatorname{var}[\hat{f}_n(\mathbf{x}; p)] = \sigma^2(\mathbf{x})$

at points of continuity of g $(\mathbf{x}; p)$, *where*

$$\sigma^2(\mathbf{x}) = \left\{ \frac{1}{2\pi |B_1|^2} \int_{-\infty}^{\infty} |t|^{2\beta} |\bar{\phi}_K(t)|^2 \, dt \right\}^p g(\mathbf{x}; p). \tag{2.7}$$

We remark that Part (c) follows from Parts (a), and (b) and stationarity. It is clear from Lemma 2.1 that $\operatorname{var}[\hat{f}_n(\mathbf{x}; p)] \to 0$ as $n \to \infty$ provided that

$nb_n^{(2\beta+1)p} \to \infty$ as $n \to \infty$. We also note that the summability condition (2.4) on the strong mixing coefficients can be slightly weakened as follows: with $c_n \to \infty$ such that $c_n b_n^p \to 0$ as $n \to \infty$, (2.4) can be replaced by

$$\frac{1}{b_n^{(1-2/v)p}} \sum_{j=c_n}^{n} [\alpha(j)]^{1-2/v} \to 0 \qquad \text{as} \quad n \to \infty \tag{2.4$'$}$$

(see the proof of Theorem 3 in [13]). We also note that (2.4) is equivalent to $\alpha(j) = O(1/j^{2+\delta})$ for some $\delta > 0$. Conditions (2.4) or (2.4)′ are standard conditions on the strong mixing coefficient in the context of quadratic-mean analysis of density estimation for dependent data: Robinson [17] assumes that $\alpha(j)$ satisfies $\sum_{j=N}^{\infty} \alpha(j) = o(1/N)$, which is equivalent to $\alpha(j) = o(1/j^2)$, and Roussas and Tran [19] assume precisely condition (2.4)′.

We now proceed to establish the asymptotic normality of $\hat{f}_n(\mathbf{x}; p)$ under Assumptions 2.1 and 2.2. The general procedure is as follows. Put

$$Z_{n,j} = b_n^{(2\beta+1)p/2} \left\{ \frac{1}{b_n^p} W_{b_n}[(\mathbf{x} - \mathbf{Y}_j)/b_n] - \mu_n \right\} = b_n^{(2\beta+1)p/2} \tilde{Z}_{n,j} \tag{2.8a}$$

and

$$S_n = \sum_{j=0}^{n-1} Z_{n,j}. \tag{2.8b}$$

Then

$$\begin{aligned} &\{nb_n^{(2\beta+1)p}\}^{1/2} \{\hat{f}_n(\mathbf{x}; p) - E[\hat{f}_n(\mathbf{x}; p)]\} \\ &\quad = \left(\frac{n}{n-p+1}\right)^{1/2} \left(\frac{1}{\sqrt{n-p+1}} S_{n-p+1}\right). \end{aligned} \tag{2.9}$$

It suffices to show that

$$\frac{1}{\sqrt{n}} S_n \xrightarrow{L} N(0, \sigma^2(\mathbf{x})) \qquad \text{as} \quad n \to \infty.$$

Partition the set $\{1, 2, ..., n\}$ into $2k+1$ subsets with large blocks of size $r = r_n$ and small blocks of size $s = s_n$, where

$$k = k_n = \left\lfloor \frac{n}{r_n + s_n} \right\rfloor \tag{2.10}$$

and $\lfloor x \rfloor$ denotes the integer part of x. Define the random variables

$$\eta_j = \sum_{i=j(r+s)+1}^{j(r+s)+r} Z_{n,i}, \qquad 0 \leqslant j \leqslant k-1, \tag{2.11}$$

$$\xi_j = \sum_{i=j(r+s)+r+1}^{(j+1)(r+s)} Z_{n,i}, \qquad 0 \leqslant j \leqslant k-1, \tag{2.12}$$

and

$$\zeta_k = \sum_{i=k(r+s)+1}^{n} Z_{n,i}. \tag{2.13}$$

Write

$$\begin{aligned} S_n &= \sum_{j=0}^{k-1} \eta_j + \sum_{j=0}^{k-1} \xi_j + \zeta_k \\ &\equiv S'_n + S''_n + S'''_n. \end{aligned} \tag{2.14}$$

The small block-big block procedure is to show that, as $n \to \infty$,

$$\frac{1}{n} E[S''_n]^2 \to 0, \qquad \frac{1}{n} E[S'''_n]^2 \to 0 \tag{2.15a}$$

$$\left| E[e^{itS'_n}] - \prod_{j=0}^{k-1} E[e^{it\eta_j}] \right| \to 0 \tag{2.15b}$$

$$\frac{1}{n} \sum_{j=0}^{k-1} E[\eta_j^2] \to \sigma^2(\mathbf{x}) \tag{2.15c}$$

$$\frac{1}{n} \sum_{j=0}^{k-1} E[\eta_j^2 I\{|\eta_j| > \varepsilon\sigma(x)\sqrt{n}\}] \to 0 \tag{2.15d}$$

for every $\varepsilon > 0$. Equation (2.15a) implies that S''_n and S'''_n are asymptotically negligible; Eq. (2.15b) shows that the summands $\{\eta_j\}$ in S'_n are asymptotically independent; and Eqs. (2.15c)–(2.15d) are the standard Lindeberg–Feller conditions for asymptotic normality of S'_n under independence. We note that $Z_{n,i}$ is a function of the random vector $\mathbf{Y}_i = (Y_{i+1}, ..., Y_{i+p})$ and thus $Z_{n,i}$ and $Z_{n,i+1}$ share all their random variables except their end points. This overlap makes the analysis more complex for $p > 1$.

The principal result of this section is the following.

THEOREM 2.1. *Let Assumptions 2.1 and 2.2 hold and assume that $b_n \to 0$ such that $nb_n^{(2\beta+1)p} \to \infty$ as $n \to \infty$. Let $\{s_n\}$ be a sequence of positive integers, $s_n \to \infty$, such that $s_n = o((nb_n^p)^{1/2})$ as $n \to \infty$.*

(a) *For ρ-mixing processes, let $\rho(k)$ satisfy*

$$(n/h_n^p)^{1/2}\,\rho(s_n) \to 0 \qquad as \quad n \to \infty. \tag{2.16}$$

(b) *For strongly mixing processes, let $\alpha(k)$ satisfy*

$$(n/h_n^p)^{1/2}\,\alpha(s_n) \to 0 \qquad as \quad n \to \infty. \tag{2.17}$$

Then

$$(nh_n^{(2\beta+1)p})^{1/2}\,\{\hat{f}_n(\mathbf{x}; p) - E[\hat{f}_n(\mathbf{x}; p)]\} \xrightarrow{L} N(0, \sigma^2(\mathbf{x}))$$

as $n \to \infty$ at continuity points of $g(\mathbf{x}; p)$ with $g(\mathbf{x}; p) > 0$, where the asymptotic constant $\sigma^2(\mathbf{x})$ is given in (2.7).

Corollary 2.1. *If, in addition, $f(\mathbf{x}; p)$ is twice continuously differentiable, as in Part (b) of Proposition 1.1, and the bandwidth parameter h_n is such that*

$$nh_n^{(2\beta+1)p+4} \to 0 \qquad as \quad n \to \infty,$$

then

$$(nh_n^{(2\beta+1)p})^{1/2}\,\{\hat{f}_n(\mathbf{x}; p) - f(\mathbf{x}; p)]\} \xrightarrow{L} N(0, \sigma^2(\mathbf{x}))$$

as $n \to \infty$ at all points $\mathbf{x}$ such that $g(\mathbf{x}; p) > 0$.

Remark 2.1. For the case $p = 1$ with i.i.d. observations, such a result was obtained by Fan [6]. Theorem 2.1 can be viewed as a generalization to dependent observations and to multivariate densities of arbitrary order p.

Remark 2.2. Since convergence in distribution does not imply convergence in quadratic-mean, the scale parameter $\sigma^2(\mathbf{x})$ in the asymptotic distribution does not usually represent the asymptotic variance of the standardized estimate. In our case, however, quadratic-mean convergence holds and, by Lemma 2.1(c), $\sigma^2(\mathbf{x})$ is indeed the asymptotic variance of the standardized estimate.

Remark 2.3. Condition (2.4) on the strong mixing coefficient is equivalent to $\alpha(j) = O(j^{-(2+\delta)})$ for some $\delta > 0$. Under $nh_n^{(2\beta+1)p} \to \infty$, condition (2.17) on the strong mixing coefficient is satisfied when $\alpha(j) = O([(\log j)\, j^{1+1/\beta}]^{-1})$. It follows that for $\beta \geqslant 1$, Theorem 2.1 for strongly mixing processes holds whenever the strong mixing coefficient satisfies $\alpha(j) = O(j^{-(2+\delta)})$ for some $\delta > 0$. Note that this is compatible with the conditions imposed in the literature on the strong mixing coefficient in the context of standard multivariate density/regression estimation. See Robinson [17] and Roussas and Tran [19] who assume $\alpha(j)=o(1/j^2)$ and $\alpha(j)=O(1/j^{2+\delta})$, respectively. We also note that the form of condition (2.17) on the strong mixing coefficient is commonly found in the literature as in Castellana and Leadbetter [3]. Regarding the ρ-mixing case, we are imposing conditions (2.3) and (2.16) on the mixing coefficient. (2.3) is satisfied when $\rho(j)=O(1/j^{1+\delta})$ for some $\delta>0$ and condition (2.16) is satisfied when $\rho(j)=O([\log j)\, j^{1+1/\beta}]^{-1})$. It follows that in the ρ-mixing case, Theorem 2.1 holds whenever $\rho(j)=O([\log j)\, j^{1+1/\beta}]^{-1})$. In the ρ-mixing case, it appears possible to considerably weaken the conditions on the ρ-mixing coefficient by appealing to Theorem 2 of Bradley [1], where asymptotic normality of *univariate* density estimates is considered under the condition $\sum_{j=1}^{\infty}\rho(2^j)<\infty$. We did not do so since our method of proof (the large-block small-block argument) applies to both strongly mixing and ρ-mixing processes whereas Bradley's distinct approach is shown to work only for ρ-mixing processes.

Remark 2.4. Note that the optimal choice of the bandwidth h_n, which minimizes the asymptotic mean-square error, is of the order $h_n \sim n^{-1/[(2\beta+1)p+4]}$. In Corollary 2.1 we have selected a slightly faster rate of decay for h_n which makes the variance of the estimator to be the dominant term. As a consequence, under the conditions of Corollary 2.1, it is easy to obtain confidence intervals for the estimator $\hat{f}_n(\mathbf{x};p)$: We only need a consistent estimate for $\sigma^2(\mathbf{x})$; by (2.7) we have $\sigma^2(\mathbf{x})=Dg(\mathbf{x};p)$, where the constant D is known, since $\bar{\phi}_K(t)$ and B_1 are assumed known. A consistent estimate for $g(\mathbf{x};p)$ can easily be obtained from the data $\{Y_i\}_{i=0}^{n}$ using a standard kernel approach:

$$\hat{g}_n(\mathbf{x};p)=\frac{1}{(n-p+1)\,h_n^p}\sum_{j=0}^{n-p} K[(\mathbf{x}-\mathbf{Y}_j)/h_n].$$

For ρ-mixing and strongly mixing processes satisfying the summability conditions (2.3) and (2.4), respectively, we have, at continuity points of $g(\mathbf{x};p)$,

$$nh_n^p \operatorname{var}[\hat{g}_n(\mathbf{x};p)] \to g(\mathbf{x};p)\int_{R^p} K^2(\mathbf{u})\,d\mathbf{u}.$$

Note that the convergence rate here is faster than that of Lemma 2.1 for $\hat{f}_n(\mathbf{x};p)$ (for sharp rates of almost sure convergence for standard multivariate density estimators of stationary processes—see, for example, [12]).

Proof of Theorem 2.1. We focus the proof on strongly mixing processes (which is more involved) and we remark on the differences for ρ-mixing processes. We first choose the block size. The condition $s_n=o((nh_n^p)^{1/2})$ and

(2.17) imply that there exist integers $q_n \to \infty$ such that for strongly mixing processes, we have

$$q_n s_n = o((nb_n^p)^{1/2}) \tag{2.18a}$$

and

$$q_n (n/b_n^p)^{1/2} \alpha(s_n) \to 0 \quad \text{as} \quad n \to \infty, \tag{2.18b}$$

and for ρ-mixing processes,

$$q_n s_n = o((nb_n^p)^{1/2}); \qquad q_n (n/b_n^p)^{1/2} \rho(s_n) \to 0 \quad \text{as} \quad n \to \infty. \tag{2.18c}$$

Now define the large-block size r_n by

$$r_n = \lfloor (nb_n^p)^{1/2}/q_n \rfloor. \tag{2.19}$$

Then simple algebra shows that the following properties hold as $n \to \infty$:

$$s_n/r_n \to 0 \tag{2.20a}$$

$$r_n/n \to 0 \tag{2.20b}$$

$$r_n/[(nb_n^p)^{1/2}] \to 0 \tag{2.20c}$$

$$\frac{n}{r_n} \alpha(s_n) \to 0. \tag{2.20d}$$

We now proceed to establish (2.15a)–(2.15d). Since $E[Z_{n,i}] = 0$ we have $E[\eta_i] = E[\xi_i] = E[\zeta_k] = 0$. Consider first

$$E[S_n'']^2 = \operatorname{var}\left[\sum_{j=0}^{k-1} \xi_j\right] = \sum_{j=0}^{k-1} \operatorname{var}[\xi_j] + \sum_{\substack{i=0 \\ i \neq j}}^{k-1} \sum_{j=0}^{k-1} \operatorname{cov}\{\xi_i, \xi_j\} \equiv F_1 + F_2. \tag{2.21}$$

With $m_j = j(r+s) + r$ we have by (2.12)

$$\begin{aligned} \operatorname{var}[\xi_j] &= \sum_{i_1=1}^{s} \sum_{i_2=1}^{s} \operatorname{cov}\{Z_{n,m_j+i_1}, Z_{n,m_j+i_2}\} \\ &= s \operatorname{var}[Z_{n,0}] + 2s \sum_{i=1}^{s-1} \left(1 - \frac{i}{s}\right) \operatorname{cov}\{Z_{n,0}, Z_{n,i}\} \end{aligned} \tag{2.22}$$

independent of j by stationarity. By Lemma 2.1(b) the sum above is $o(1)$ and by Lemma 2.1(a) we have

$$\operatorname{var}[\xi_j] = \sigma^2(\mathbf{x})\, s\{1 + o(1)\}. \tag{2.23}$$

It then follows by (2.21) that

$$F_1 \leqslant Cks, \tag{2.24}$$

where C is a generic constant. Next consider the term F_2 in (2.21):

$$F_2 = \sum_{i=0}^{k-1} \sum_{\substack{j=0 \\ i \neq j}}^{k-1} \sum_{l_1=1}^{s} \sum_{l_2=1}^{s} \operatorname{cov}\{Z_{n,m_i+l_1}, Z_{n,m_j+l_2}\}, \tag{2.25}$$

but since $i \neq j$, $|m_i - m_j + l_1 - l_2| \geqslant r$ so that

$$|F_2| \leqslant 2 \sum_{l_1=1}^{n-r} \sum_{l_2=l_1+r}^{n} |\operatorname{cov}\{Z_{n,l_1}, Z_{n,l_2}\}|.$$

Since $r = r_n \to \infty$ we can assume that $r_n > p$ so that the random vectors $\mathbf{Y}_{l_1}$ and $\mathbf{Y}_{l_2}$ (appearing in Z_{n,l_1} and Z_{n,l_2} respectively) do not have common components. By stationarity and Lemma 2.1(b)

$$\begin{aligned} |F_2| &\leqslant 2 \sum_{l_1=1}^{n-r} \sum_{l_2=l_1+r}^{n} |\operatorname{cov}\{Z_{n,0}, Z_{n,l_2-l_1}\}| \\ &\leqslant 2n \sum_{j=r}^{n-1} |\operatorname{cov}\{Z_{n,0}, Z_{n,j}\}| = o(n). \end{aligned} \tag{2.26}$$

Hence by (2.21), (2.24), and (2.26) we have

$$\frac{1}{n} E[S_n'']^2 \leqslant \frac{C}{n} \{ks + o(n)\} \to 0 \qquad \text{as} \quad n \to \infty,$$

where the last step follows by (2.10) and (2.20a), establishing the first part of (2.15a). Using a similar argument, we find

$$\begin{aligned} \frac{1}{n} E[S_n''']^2 &= \frac{1}{n} \sum_{j=1}^{n-k(r+s)} \operatorname{var}[Z_{n,m_k+j}] \\ &\quad + \frac{1}{n} \sum_{\substack{i=1 \\ i \neq j}}^{n-k(r+s)} \sum_{j=1}^{n-k(r+s)} \operatorname{cov}\{Z_{n,m_k+i}, Z_{n,m_k+j}\} \end{aligned} \tag{2.27}$$

$$\leqslant \frac{C}{n} [n - k(r+s)] \leqslant C \frac{r_n + s_n}{n} \to 0 \qquad \text{as} \quad n \to \infty \tag{2.28}$$

by (2.20a) and (2.20b). This proves the second part of (2.15a).

In order to establish (2.15b) we make use of the following lemma due to Volkonskii and Rozanov [22].

LEMMA. *Let $V_1, ..., V_L$ be random variables measurable with respect to the σ-algebras $\mathscr{F}_{i_1}^{j_1}, ..., \mathscr{F}_{i_L}^{j_L}$, respectively, with $1 \leqslant i_1 < j_1 < i_2 < \cdots < j_L \leqslant n$, $i_{l+1} - j_l \geqslant w \geqslant 1$ and $|V_j| \leqslant 1$ for $j = 1, ..., L$.*
Then

$$\left| E\left[\prod_{j=1}^{L} V_j\right] - \prod_{j=1}^{L} E[V_j] \right| \leqslant 16(L-1)\,\alpha(w).$$

We note that η_l is a function of the random variables $\{Y_{l(r+s)+2}, ..., Y_{p+l(r+s)+r}\}$ or η_l is $\mathscr{F}_{i_l}^{j_l}$-measurable with $i_l = l(r+s)+2$, $j_l = p + l(r+s)+r$. Also $i_{l+1} - j_l = s + 2 - p$. Hence with $V_j = e^{it\eta_j}$ we have

$$\left| E\left[\prod_{j=0}^{k-1} e^{it\eta_j}\right] - \prod_{j=0}^{k-1} E[e^{it\eta_j}] \right| \leqslant 16 k_n \alpha(s_n + 2 - p) \tag{2.29}$$

and (2.15b) follows since

$$k_n \alpha(s_n + 2 - p) \sim \frac{n}{r_n + s_n}\,\alpha(s_n + 2 - p) \to 0$$

as $n \to \infty$ by (2.20a) and (2.20d).

Next we establish (2.15c). By stationarity and (2.23), with r replacing s, we have

$$\frac{1}{n}\sum_{j=0}^{k-1} E[\eta_j^2] = \frac{k_n}{n} E[\eta_0^2] = \frac{k_n r_n}{n} \operatorname{var}[Z_{n,0}](1 + o(1)) \to \sigma^2(\mathbf{x}) \tag{2.30}$$

as $n \to \infty$ by Lemma 2.1(a), (2.10), and (2.20a).

It remains to establish (2.15d). With η_j given by (2.11) and $Z_{n,i}$ by (2.8) we have

$$|Z_{n,i}| \leqslant \frac{b_n^{(2\beta+1)p/2}}{b_n^p} \left\{ \sup_{\mathbf{u}} |W_{b_n}(\mathbf{u})| + |\mu_n| \right\}.$$

It is easy to check that $|W_{b_n}(\mathbf{u})| \leqslant C/b_n^{\beta p}$ (as in Lemma 3.2 of [13]) and since $\mu_n \to f(\mathbf{x}; p)$ by Proposition 1.1 we have $|Z_{n,i}| \leqslant C/b_n^{p/2}$ uniformly in i. Hence by (2.20c)

$$\max_{0 \leqslant j \leqslant k-1} |\eta_j|/\sqrt{n} \leqslant \frac{C r_n / b_n^{p/2}}{\sqrt{n}} \to 0, \qquad \text{as} \quad n \to \infty. \tag{2.31}$$

Thus the set $\{|\eta_j| > \sigma\varepsilon\sqrt{n}\}$ becomes empty for large n. (2.15d) now follows by

$$\frac{1}{n}\sum_{j=0}^{k-1} E[\eta_j^2 I\{|\eta_j| > \sigma\varepsilon\sqrt{n}\}] \leqslant \frac{C^2}{n} k_n (r_n/b_n^{p/2})^2\, P[|\eta_0| > \sigma\varepsilon\sqrt{n}]. \tag{2.32}$$

This completes the proof of the theorem. ∎

III. Asymptotic Normality for the Class of $\bar{\phi}_h(t)$ with Exponential Decay

In this section we establish the asymptotic normality of $\hat{f}_n(\mathbf{x}; p)$ when the noise characteristic function $\bar{\phi}_h(t)$ decays exponentially fast. Specifically we assume that

Assumption 3.1. (i) $|\bar{\phi}_h(t)| > 0$ for all t.

(ii) $B_1 |t|^{\beta_0} e^{-a|t|^\beta} \leqslant |\bar{\phi}_h(t)| \leqslant B_2 |t|^{\beta_0} e^{-a|t|^\beta}$ as $|t| \to \infty$ for some $a > 0$, $\beta > 0$, β_0 real, and positive constants B_i.

(iii) $\bar{\phi}_K(t)$ has a finite support $(-d, d)$.

(iv) $|\bar{\phi}_K(t)| \leqslant B_3(d-t)^m$ for $t \in (d-\delta, d)$ for some positive constants m and δ.

(v) $\bar{\phi}_K(t) \geqslant B_4(d-t)^m$ for $t \in (d-\delta, d)$, where B_4 is a positive constant.

(vi) Either $\bar{I}_h(t) = o(\bar{R}_h(t))$ or $\bar{R}_h(t) = o(\bar{I}_h(t))$ as $t \to \infty$, where $\bar{R}_h(t)$ and $\bar{I}_h(t)$ are the real and imaginary parts of $\bar{\phi}_h(t)$.

The parameter $\beta > 0$ is called the order of the noise density $\bar{h}(x)$ ($\beta_0 = 0$ with $\beta = 2$ gives the Gaussian case; $\beta_0 = 0$ with $\beta = 1$ gives the Cauchy density case). Since $t^j \bar{\phi}_h(t) \in L_1$ it follows that $\bar{h}(x)$ has bounded derivatives of all orders and thus $g(\mathbf{x}; p)$ of (1.2a) is certainly bounded and continuous,

$$g(\mathbf{x}; p) \leqslant M < \infty.$$

Conditions (v) and (vi) are used to establish lower bounds on the variance $\sigma_0^2(n)$ of (3.1). Condition (vi) is fairly weak: it says that at the tail, the characteristic function $\bar{\phi}_h(t)$ is either purely real or purely imaginary. In particular, this is satisfied for symmetric densities $\bar{h}(x)$. Put

$$\sigma_0^2(n) = \operatorname{var}[\tilde{Z}_{n,0}] \tag{3.1}$$

and

$$\sigma^2(n) = \operatorname{var}[\hat{f}_n(\mathbf{x}; p)]. \tag{3.2}$$

The asymptotic rates and constants of $\sigma_0^2(n)$ and $\sigma^2(n)$ are not available in this case; upper bounds were given in Masry [13] along with quadratic-mean consistency of $\hat{f}_n(\mathbf{x}; p)$ for a suitable bandwidth selection. As a consequence, no central limit theorem having a classical form, as in Theorem 2.1, can be established. Instead we show that

$$\frac{\hat{f}_n(\mathbf{x}; p) - E[\hat{f}_n(\mathbf{x}; p)]}{\sigma(n)} \xrightarrow{L} N(0, 1)$$

as $n \to \infty$.

We make use of the following two lemmas whose proofs are given in the Appendix. Lemma 3.1 provides fairly tight upper and lower bounds on $\sigma_0^2(n)$. Lemma 3.2 gives an asymptotic relationship between $\sigma^2(n)$ and $\sigma_0^2(n)$ which is similar in character to that of Lemma 2.1.

LEMMA 3.1. *Under Assumption* 3.1 *we have*

$$\sigma_0^2(n) \geqslant B_5 b_n^{2p[(m+1)\beta+\beta_0-1/2]} e^{2ap(d/b_n)^\beta} \tag{3.3}$$

and

$$\sigma_0^2(n) \leqslant B_6 b_n^{2p[(m+1)\beta+\beta_0-1]} [\ln(1/b_n)]^{2mp} e^{2ap(d/b_n)^\beta} \tag{3.4}$$

for some positive constants B_5 *and* B_6.

Assumption 3.2. The processes $\{X_i, \varepsilon_i\}$ are either ρ-mixing with $\sum_{j=1}^{\infty} \rho(j) < \infty$ or strongly mixing with

$$\sum_{j=1}^{\infty} j^{\lambda} [\alpha(j)]^{1-2/\nu} < \infty \qquad \text{for some} \quad \nu > 2 \text{ and } \lambda > 0.$$

Note that the above condition is equivalent to $\alpha(j) = O(1/j^{1+\delta})$ for some $\delta > 0$ and is weaker than the corresponding condition (2.4) in Assumption 2.2 for the algebraic case.

LEMMA 3.2. *Under Assumptions* 3.1 *and* 3.2 *we have*

(a) $\sum_{j=1}^{n} |\text{cov}\{\tilde{Z}_{n,0}, \tilde{Z}_{n,j}\}| = o(\sigma_0^2(n))$ *and*

(b) $\sigma^2(n) = (1/n)\, \sigma_0^2(n)(1 + o(1))$.

We note that by Lemmas 3.1 and 3.2 and Proposition (1.1), $\hat{f}_n(\mathbf{x}; p)$ is consistent in quadratic mean whenever the bandwidth parameter satisfies $b_n = d(2ap/[\theta \ln n])^{1/\beta}$ for some $0 < \theta < 1$. Then $\text{var}[\hat{f}_n(\mathbf{x}; p)] = O(1/n^{1-\theta})$ so that the bias term becomes dominant. Under the assumptions of Proposition (1.1)(b), the mean-square convergence rate is of the order $1/(\ln n)^{2/\beta}$ for *all* $p \geqslant 1$. Fan [5] has shown that this rate is optimal for $p = 1$ under the assumption of i.i.d. observations.

The main result of this section is given by the following theorem.

THEOREM 3.1. *Let Assumptions* 3.1 *and* 3.2 *hold. Assume that* $b_n \to 0$ *such that* $nb_n^{p\gamma} \to \infty$ *as* $n \to \infty$ *for some* $\gamma > 1$. *Put* $s_n = \lfloor (nb_n^{p\gamma})^{1/2} \rfloor$.

(a) *For* ρ-*mixing processes, let* $\rho(k)$ *satisfy*

$$(n/b_n^{p\gamma})^{1/2} \rho(s_n) \to 0 \qquad \text{as} \quad n \to \infty. \tag{3.5}$$

(b) *For strongly mixing processes, let $\alpha(k)$ satisfy*

$$(n/b_n^{p\gamma})^{1/2}\,\alpha(s_n) \to 0 \qquad as \quad n \to \infty. \tag{3.6}$$

Then

$$\sqrt{n}\,\frac{\{\hat{f}_n(\mathbf{x};p) - E[\hat{f}_n(\mathbf{x};p)]\}}{\sigma_0(n)} \xrightarrow{L} N(0,1) \tag{3.7}$$

as $n \to \infty$.

Remark 3.1. If $b_n = d(2ap/[\theta \ln n])^{1/\beta}$ for some $0 < \theta < 1$ (which ensures quadratic-mean convergence), then conditions (3.5) and (3.6) on the mixing coefficients are automatically satisfied whenever $\alpha(j)$ or $\rho(j)$ are $O(1/j^{1+\delta})$ for some $\delta > 0$.

Remark 3.2. Note that the precise rate and asymptotic variance are not specified in Theorem 3.1. However, fairly tight upper and lower bounds on $\sigma_0^2(n)$ are provided in Lemma 3.1 We also remark that Theorem 3.1 generalizes the corresponding result obtained by Fan [6] for $p = 1$ and i.i.d. observations. It should be noted that the convergence in distribution in (3.7) can be quite slow in view of the potentially slow rate of convergence of $\operatorname{var}[\hat{f}_n(\mathbf{x};p)]$. Moreover, as discussed earlier, the bias term is frequently dominant. It follows that in the exponential decay case, asymptotic normality for the estimator $\hat{f}_n(\mathbf{x};p)$ does not appear to be as practically useful as in the algebraic case.

Proof. We normalize $\tilde{Z}_{n,j}$ of (2.6) by

$$Z_{n,j} = \tilde{Z}_{n,j}/\sigma_0(n); \qquad S_n = \sum_{j=0}^{n-1} Z_{n,j} \tag{3.8}$$

so that

$$\operatorname{var}[Z_{n,j}] = 1 \tag{3.9}$$

and by Lemma 3.2,

$$\sum_{j=1}^{n} |\operatorname{cov}\{Z_{n,0}, Z_{n,j}| = o(1); \qquad \operatorname{var}[S_n] = n(1 + o(1)). \tag{3.10}$$

Then

$$\sqrt{n}\,\frac{\{\hat{f}_n(\mathbf{x};p) - E\hat{f}_n(\mathbf{x};p)]\}}{\sigma_0(n)} = \left(\frac{n}{n-p+1}\right)^{1/2}\left(\frac{1}{\sqrt{n-p+1}}\,S_{n-p+1}\right).$$

We need to show that

$$\frac{1}{\sqrt{n}} S_n \xrightarrow{L} N(0, 1). \tag{3.11}$$

Define the large block size $r = r_n$ by $r_n = \lfloor (nb_n^{p\gamma_1})^{1/2} \rfloor$, where $1 < \gamma_1 < \gamma$. Then it is easy to verify the following properties as $n \to \infty$:

$$s_n/r_n \to 0, \qquad r_n/n \to 0 \tag{3.12}$$

$$\frac{r_n}{(nb_n^p)^{1/2}} [\ln(1/b_n)]^{mp} \to 0 \tag{3.13}$$

and

$$\frac{n}{r_n} \alpha(s_n) \to 0. \tag{3.14}$$

Let η_j, ξ_j, ζ_k and S'_n, S''_n, S'''_n be as in Section II but with $Z_{n,j}$ given by (3.8). We need to show that as $n \to \infty$

$$\frac{1}{n} E[S''_n]^2 \to 0, \qquad \frac{1}{n} E[S'''_n]^2 \to 0 \tag{3.15a}$$

$$\left| E[e^{itS'_n}] - \prod_{j=0}^{k-1} E[e^{it\eta_j}] \right| \to 0 \tag{3.15b}$$

$$\frac{1}{n} \sum_{j=0}^{k-1} E[\eta_j^2] \to 1 \tag{3.15c}$$

$$\frac{1}{n} \sum_{j=0}^{k-1} E[\eta_j^2 I\{|\eta_j| > \varepsilon \sqrt{n}\}] \to 0 \tag{3.15d}$$

for every $\varepsilon > 0$.

As in (2.22), using (3.9) and (3.10), we find

$$\operatorname{var}[\xi_j] = s(1 + o(1)) \tag{3.16}$$

so that by (2.21) $F_1 = ks(1 + o(1))$ and as in (2.26) we have by (3.10) that $F_2 = o(n)$. Thus

$$\frac{1}{n} E[S''_n]^2 \to 0 \qquad \text{as} \quad n \to \infty. \tag{3.17}$$

Similarly by (2.27), using (3.9) and (3.10), we have

$$\frac{1}{n} E[S'''_n]^2 \leqslant \frac{1}{n} [n - k(r + s)] (1 + o(1)) \to 0 \qquad \text{as} \quad n \to \infty. \tag{3.18}$$

Next as in (2.29)

$$\left| E[e^{itS_n'}] - \prod_{j=0}^{k-1} E[e^{it\eta_j}] \right| \leqslant 16k_n\alpha(s_n+2-p) \sim \frac{n}{r_n}\alpha(s_n+2-p) \to 0 \qquad (3.19)$$

by (3.14). Now by (3.10) and (3.16) (with s replaced by r) we have $E[\eta_0^2]=r_n(1+o(1))$ and by stationarity and (3.12), $(1/n)\sum_{j=0}^{k-1} E[\eta_j^2]=(k_n r_n/n)(1+o(1))\to 1$. Finally we verify the Lindeberg-Feller condition (3.15d). We have, since $\mu_n \to f(\mathbf{x}; p)$,

$$|Z_{n,i}| < \text{Const.}(\|\bar{W}_{b_n}\|_\infty/b_n)^p/\sigma_0(n)$$

uniformly in i so that

$$|\eta_0| \leqslant \text{Const. } r_n(\|\bar{W}_{b_n}\|_\infty/b_n)^p/\sigma_0(n).$$

Using the upper bound on $\|\bar{W}_{b_n}\|_\infty$ in Lemma A.1 and the lower bound on $\sigma_0(n)$ given in Lemma 3.1, we find

$$|\eta_0|/\sqrt{n} \leqslant \text{Const.} \frac{r_n}{(nb_n^p)^{1/2}}[\ln(1/b_n)]^{mp} \to 0$$

by (3.13). Thus $P[|\eta_0|>\varepsilon\sqrt{n}]=0$ for sufficiently large n. (3.15d) now follows by

$$\frac{1}{n}\sum_{j=0}^{k-1} E[\eta_j^2 I\{|\eta_j|>\varepsilon\sqrt{n}\}]$$

$$\leqslant \text{Const.} \frac{k_n}{n}\left(\frac{r_n^2}{b_n^p}\right)[\ln(1/b_n)]^{2mp} P[|\eta_0|/\sqrt{n}>\varepsilon]. \quad \blacksquare$$

APPENDIX

We make use of the following bounds on the kernel $\bar{W}_{b_n}(x)$ of (1.4) (Fan and Masry, [7]).

LEMMA A.1. (a) *Under Assumption* 3.1(i)–(iv) *we have as* $n\to\infty$

$$\|\bar{W}_{b_n}\|_\infty = O[b_n^{(m+1)\beta+\beta_0}\{\ln(1/b_n)\}^m e^{a(d/b_n)^\beta}].$$

(b) *Under Assumption* 3.1 *we have*

$$|\bar{W}_{b_n}(x)| \geqslant B_7\bar{H}(x)\, b_n^{(m+1)\beta+\beta_0} e^{a(d/b_n)^\beta}$$

for some $B_7 > 0$ *uniformly in* x *on a bounded interval, where*

$$\bar{H}(x) = \begin{cases} |\cos dx| & \text{if } \ \bar{I}_h(t) = o(\bar{R}_h(t)) \\ |\sin dx| & \text{if } \ \bar{R}_h(t) = o(\bar{I}_h(t)). \end{cases}$$

Proof of Lemma 3.1. Since μ_n is bounded, we have

$$\sigma_0^2(n) = E[\tilde{Z}_{n,0}^2] = \frac{1}{b_n^{2p}} E[W_{b_n}^2[(\mathbf{x} - \mathbf{Y}_0)/b_n]] + O(1)$$

$$= \frac{1}{b_n^p} \int_{R^p} W_{b_n}^2(\mathbf{u})\, g(\mathbf{x} - b_n \mathbf{u})\, d\mathbf{u} + O(1)$$

and by Lemma A.1 and the factorization (1.5)

$$\sigma_0^2(n) \geqslant \frac{B_7^{2p} b_n^{2p[(m+1)\beta + \beta_0]} e^{2ap(d/b_n)^\beta}}{b_n^p}$$

$$\times \int_{[-1,1]^p} g(\mathbf{x} - b_n \mathbf{u}; p) \prod_{j=1}^{p} \bar{H}^2(u_j)\, d\mathbf{u} + O(1)$$

and since g is continuous and $b_n \to 0$,

$$\sigma_0^2(n) \geqslant B_7^{2p} b_n^{2p[(m+1)\beta + \beta_0 - 1/2]} e^{2ap(d/b_n)^\beta} g(\mathbf{x}; p) \left[\int_{-1}^{1} \bar{H}^2(u)\, du \right]^p (1 + o(1))$$

$$\geqslant B_5 b_n^{2p[(m+1)\beta + \beta_0 - 1/2]} e^{2ap(d/b_n)^\beta}$$

for some constant $B_5 > 0$. The upper bound follows from

$$\sigma_0^2(n) \leqslant b_n^{-2p} E[W_{b_n}^2[(\mathbf{x} - \mathbf{Y}_0)/b_n]] \leqslant b_n^{-2p} \| \bar{W}_{b_n} \|_\infty^{2p}$$

and from the upper bound on $\| \bar{W}_{b_n} \|_\infty$ given in Lemma A.1. ∎

Proof of Lemma 3.2. By stationarity

$$\sigma^2(n) = \frac{1}{n-p+1} \sigma_0^2(n) + \frac{2}{n-p+1} \sum_{l=1}^{n-p} \left(1 - \frac{l}{n-p+1}\right) \tilde{I}_{n,l}, \qquad \text{(A.1)}$$

where

$$\tilde{I}_{n,l} = \operatorname{cov}\{\tilde{Z}_{n,0}, \tilde{Z}_{n,l}\}. \qquad \text{(A.2)}$$

Put

$$J = \sum_{l=1}^{n-p} |\tilde{I}_{n,l}| = \sum_{l=1}^{p-1} + \sum_{l=p}^{c_n} + \sum_{l=c_n+1}^{n-p} \equiv J_1 + J_2 + J_3, \qquad \text{(A.3)}$$

where

$$c_n = e^{ap(d/b_n)^\beta}. \tag{A.4}$$

Note that J_l is vacuous when $p=1$. Consider J_l: When $1 \leqslant l \leqslant p-1$ there is an overlap between the vectors $\mathbf{Y}_0$ and $\mathbf{Y}_l$. Put

$$\mathbf{X}' = (X_1, ..., X_p); \qquad \mathbf{X}'' = (X_{l+1}, ..., X_p); \qquad \mathbf{X}''' = (X_{p+1}, ..., X_{p+l})$$

and similarly for $\boldsymbol{\varepsilon}'$, $\boldsymbol{\varepsilon}''$, and $\boldsymbol{\varepsilon}'''$. Then

$$\mathbf{Y}' = \mathbf{X}' + \boldsymbol{\varepsilon}', \qquad \mathbf{Y}'' = \mathbf{X}'' + \boldsymbol{\varepsilon}'', \qquad \mathbf{Y}''' = \mathbf{X}''' + \boldsymbol{\varepsilon}'''.$$

Define

$$q_1(\mathbf{Y}') = \prod_{j=1}^{l} \bar{W}_{b_n}[(x_j - Y_j)/b_n] \tag{A.5a}$$

$$q_2(\mathbf{Y}'') = \prod_{j=l+1}^{p} \bar{W}_{b_n}[(x_j - Y_j)/b_n] \prod_{i=l+1}^{p} \bar{W}_{b_n}[(x_{i-l} - Y_i)/b_n] \tag{A.5b}$$

and

$$q_3(\mathbf{Y}''') = \prod_{j=p+1}^{p+l} \bar{W}_{b_n}[(x_{j-l} - Y_j)/b_n]. \tag{A.5c}$$

Then since $\mu_n \to f(\mathbf{x}; p)$ as $n \to \infty$,

$$\tilde{I}_{n,l} = \frac{1}{b_n^{2p}} E[q_2(\mathbf{Y}'')\, E[q_1(\mathbf{Y}')\, q_3(\mathbf{Y}''') \mid \mathbf{X}'', \boldsymbol{\varepsilon}'']] + O(1). \tag{A.6}$$

By (1.4) and Fubini's theorem the inner expectation is

$$\begin{aligned} J_5 = \frac{1}{(2\pi)^{2l}} \int_{R^{2l}} E\Bigg[\exp\Bigg\{ -\frac{i}{b_n} \Bigg\{ \sum_{j=1}^{l} t_j(x_j - X_j - \varepsilon_j) \\ + \sum_{j=p+1}^{p+l} t_j(x_{j-l} - X_j - \varepsilon_j)\Bigg\}\Bigg\} \Bigg| \mathbf{X}'', \boldsymbol{\varepsilon}''\Bigg] \\ \times \prod_{j=1}^{l} [\bar{\phi}_K(t_j)/\bar{\phi}_h(t_j/b_n)] \\ \times \prod_{j=p+1}^{p+l} [\bar{\phi}_K(t_j)/\bar{\phi}_h(t_j/b_n)]\, d\mathbf{t}'\, d\mathbf{t}'''. \end{aligned}$$

Let $\phi_{\mathbf{X}',\mathbf{X}'''|\mathbf{X}''}(\mathbf{t}', \mathbf{t}'''|\mathbf{X}'')$ be the conditional characteristic function of $(\mathbf{X}', \mathbf{X}'')$ given $\mathbf{X}''$. By the independence of $\{X_i\}$ and $\{\varepsilon_i\}$ we have

$$J_5 = \frac{1}{(2\pi)^{2l}} \int_{R^{2l}} \exp\left\{\frac{-i}{b_n}\left[\sum_{j=1}^{l} t_j x_j + \sum_{j=p+1}^{p+l} t_j x_{j-l}\right]\right\}$$
$$\times \phi_{\mathbf{X}',\mathbf{X}'''|\mathbf{X}''}(\mathbf{t}'/b_n, \mathbf{t}'''/b_n|\mathbf{X}'') \prod_{j=1}^{l} \bar\phi_K(t_j)$$
$$\times \prod_{j=p+1}^{p+l} \bar\phi_K(t_j)\, d\mathbf{t}'\, d\mathbf{t}'''$$

so that

$$|J_5| \leqslant (2\pi)^{-2l}\left[\int_{-d}^{d} |\bar\phi_K(t)|\, dt\right]^{2l} \leqslant (d/\pi)^{2l}.$$

Thus by (A.6)

$$|\tilde I_{n,l}| \leqslant \frac{(d/\pi)^{2l}}{b_n^{2p}} E|q_2(\mathbf{Y}'')| + O(1)$$
$$\leqslant \frac{\text{Const.}}{b_n^{2p}} \|\bar W_{b_n}\|_\infty^{2(p-l)} + O(1), \qquad 1 \leqslant l \leqslant p-1,$$

where the last step follows from (A.5b). Hence

$$J_1 \leqslant \text{Const.}\, b_n^{-2p} \sum_{l=1}^{p-1} \|\bar W_{b_n}\|_\infty^{2(p-l)} + O(1) = O(b_n^{-2p} \|\bar W_{b_n}\|_\infty^{2(p-1)}).$$

Using the upper bound on $\|\bar W_{b_n}\|_\infty$ of Lemma A.1 and the lower bound on $\sigma_0^2(n)$ of Lemma 3.1 gives

$$\frac{J_1}{\sigma_0^2(n)} \leqslant \text{Const.}\, b_n^{-[p+2(m+1)\beta+2\beta_0]} [\ln(1/b_n)]^{2m(p-1)} e^{-2a(d/b_n)^\beta} \to 0 \qquad \text{(A.7)}$$

as $n \to \infty$. For J_2, an argument similar to the above shows that $|\tilde I_{n,l}| = O(b_n^{-2p})$ uniformly in $l \geqslant p$ so that $J_2 = O(c_n b_n^{-2p})$. Hence by (A.4) and the lower bound on $\sigma_0^2(n)$ of Lemma 3.1,

$$\frac{|J_2|}{\sigma_0^2(n)} \leqslant \text{Const.}\, b_n^{-2p[(m+1)\beta+\beta_0+1/2]} e^{-ap(d/b_n)^\beta} \to 0 \qquad \text{(A.8)}$$

as $n \to \infty$. We now consider J_3. Here we have

$$|\tilde I_{n,l}| \leqslant \begin{cases} \rho(l-p+1)\, \sigma_0^2(n) & \rho\text{-mixing} \\ 8[\alpha(l-p+1)]^{1-2/v} \{E|\tilde Z_{n,0}|^v\}^{2/v} & \text{strongly mixing} \end{cases}$$

where the bound for strongly mixing processes follows by Davydov's lemma [4]. For ρ-mixing processes we then have

$$\frac{J_3}{\sigma_0^2(n)} \leqslant \sum_{l=c_n+1}^{\infty} \rho(l-p+1) \to 0 \qquad \text{as} \quad n \to \infty \tag{A.9a}$$

by the summability of $\{\rho(j)\}$. For strongly mixing processes we note that by (2.6)

$$E|\tilde{Z}_{n,0}|^{\nu} \leqslant b_n^{-p\nu} \| W_{b_n} \|_{\infty}^{\nu} + O(1)$$

and thus

$$|J_3| \leqslant \text{Const.}\, b_n^{-2p} \| \bar{W}_{b_n} \|_{\infty}^{2p} \sum_{l=c_n+1}^{\infty} [\alpha(l-p+1)]^{1-2/\nu}$$

$$\leqslant \frac{\text{Const.}}{b_n^{2p} c_n^{\lambda}} \| \bar{W}_{b_n} \|_{\infty}^{2p} \sum_{l=c_n}^{\infty} l^{\lambda} [\alpha(l-p+1)]^{1-2/\nu}.$$

Using the upper bound on $\| \bar{W}_{b_n} \|_{\infty}$ in Lemma A.1 and the lower bound on $\sigma_0^2(n)$ in Lemma 3.1 we find for the strongly mixing case

$$\frac{|J_3|}{\sigma_0^2(n)} \leqslant \frac{\text{Const.}}{c_n^{\lambda} b_n^{p}} [\ln(1/b_n)]^{2mp} \sum_{l=c_n}^{\infty} l^{\lambda} [\alpha(l-p+1)]^{1-2/\nu} \to 0 \qquad \text{as} \quad n \to \infty \tag{A.9b}$$

by the exponential growth of c_n (see (A.4)) and $\sum_{j=1}^{\infty} j^{\lambda} [\alpha(j)]^{1-2/\nu} < \infty$.

It follows by (A.3) and (A.7)–(A.9) that $J = o(\sigma_0^2(n))$ and part (a) of the Lemma follows. Part (b) then follows from (A.1). ∎

ACKNOWLEDGMENT

I thank the referee for helpful and constructive comments which led to an improved presentation of the paper.

REFERENCES

[1] Bradley, R. C. (1983). Asymptotic normality of some kernel-type estimators of probability density. *Statist. Probab. Lett.* **1** 295–300.

[2] Carroll, R. J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *J. Amer. Statist. Assoc.* **83** 1184–1186.

[3] Castellana, J. V., and Leadbetter, M. R. (1986). On smoothed probability density estimation for stationary processes. *Stochastic Process. Appl.* **21** 179–193.

[4] Davydov, Yu. A. (1968). Convergence of distributions generated by stationary stochastic processes. *Theory Probab. Appl.* **13** 691–696.

[5] Fan, J. (1991a). On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.* **19** 1257–1272.

[6] Fan, J. (1991b). Asymptotic normality for deconvolving kernel density estimators. *Sankhya Ser. A* **53** 97–110.

[7] Fan, J. and Masry, E. (1991). Multivariate regression estimation with errors-in-variables: Asymptotic normality for mixing processes. *J. Multivariate Anal.*, in press.

[8] den Harder, A., and de Galan, L. (1974). Evaluation of a method of real-time deconvolution. *Anal. Chem.* **46** 1464–1470.

[9] Jones, A. F., and Misell, D. L. (1967). A practical method for deconvolution of experimental curves. *Brit. J. Appl. Phys.* **18** 1479–1483.

[10] Kolmogorov, A. N., and Rozanov, Yu. A. (1960). On strong mixing conditions for stationary Gaussian processes. *Theory Probab. Appl.* **5** 204–207.

[11] Liu, M. C., and Taylor, R. (1989). A consistent nonparametric density estimator for the deconvolution problem. *Canad. J. Statist.* **17** 399–410.

[12] Masry, E. (1989). Nonparametric estimation of conditional probability densities and expectations of stationary processes: Strong consistency and rates. *Stochastic Process. Appl.* **32** 109–127.

[13] Masry, E. (1991a). Multivariate probability density deconvolution for stationary random processes. *IEEE Trans. Inform. Theory* **IT-37** 1105–1115.

[14] Masry, E. (1991b). Strong consistency and rates for deconvolution of multivariate densities of stationary processes. *Stochastic Process. Appl.*, in press.

[15] Medgyessy, P. (1977). *Decomposition of Superpositions of Density Functions and Discrete Distributions.* Wiley, New York.

[16] Mendelsohn, J., and Rice, J. (1982). Deconvolution of microfluorometric histograms with B-splines. *J. Amer. Statist. Assoc.* **77** 748–753.

[17] Robinson, P. M. (1983). Nonparametric estimators for time series. *J. Time Series* **4** 185–207.

[18] Rosenblatt, M. (1956). A central limit theorem and strong mixing conditions. *Proc. Nat. Acad. Sci. U.S.A.* **4** 43–47.

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

[20] Snyder, D. L., Miller, M. I., and Schultz, T. J. (1988). Constrained probability density estimation from noisy data. In *Proc.* 1988 *Conf. Inform. Sci. and Systems*, pp. 170–172.

[21] Stefanski, L., and Carroll, R. J. (1990). Deconvolving kernel density estimators. *Statistics* **21** 169–184.

[22] Volkonskii, V. A., and Rozanov, Yu. A. (1959). Some limit theorems for random functions. *Theory Probab. Appl.* **4** 178–197.

[23] Wise, G. L., Traganitis, A. P., and Thomas, J. B. (1977). The estimation of a probability density function from measurements corrupted by Poisson noise. *IEEE Trans. Inform. Theory* **IT-23** 764–766.

[24] Zhang, C. H. (1990). Fourier methods for estimating mixing densities and distributions. *Ann. Statist.* **18** 806–830.