The Annals of Statistics
1997, Vol. 25, No. 3, 982–1000

PARAMETRIC RATES OF NONPARAMETRIC ESTIMATORS AND PREDICTORS FOR CONTINUOUS TIME PROCESSES

BY DENIS BOSQ

University of Paris 6

We show that local irregularity of observed sample paths provides additional information which allows nonparametric estimators and predictors for continuous time processes to reach parametric rates in mean square as well as in a.s. uniform convergence. For example, we prove that under suitable conditions the kernel density estimator f_T associated with the observed sample path $(X_t,\ 0 \le t \le T)$ satisfies

$$\sup_{x\in\mathbb{R}} |f_T(x) - f(x)| = o\left(\ln_k T\left(\frac{\ln T}{T}\right)^{1/2}\right) \quad \text{a.s., } k \ge 1$$

where f denotes the unknown marginal density of the stationary process (X_t) and where $\ln_k$ denotes the kth iterated logarithm.

The proof uses a special Borel–Cantelli lemma for continuous time processes together with a sharp large deviation inequality. Furthermore the parametric rate obtained in (1) is preserved by using a suitable sampling scheme.

1. Introduction. Let $(X_t,\ t \in S)$ be a family of $\mathbb{R}^d$-valued random variables with a common unknown density f.

If (X_t) is a discrete time process ($S = \mathbb{Z}$), the problem of estimating f given a sample $X_1, \dots, X_n$ has been considered by many authors (see, e.g., [4], [7], [9], [18], [21], [22], [23], [24], [25]). They have proved that, under some mixing conditions, it is possible to construct a kernel estimator, say f_n, which achieves the same rates as in the i.i.d. case.

Thus, under classical regularity conditions concerning f, the mean square error (MSE) of f_n is $O(n^{-4/(4+d)})$ and its supnorm rate is $o(\ln_k n(\ln n/n)^{2/(4+d)})$ for all integers k, where $\ln_k$ denotes the kth iterated logarithm, defined recursively by $\ln_k x = \max(1, \ln_{k-1} x)$, $k \ge 2$, $x > 0$.

Now if (X_t) is a continuous time process ($S = \mathbb{R}$) observed over the time interval $[0, T]$, the MSE of the corresponding kernel estimator turns out to be $O(T^{-4/(4+d)})$ and its supnorm rate $o(\ln_k T(\ln T/T)^{2/(4+d)})$, $k \ge 1$. For details we refer to [7] (see also [1] and [12]).

However these rates can be sharpened provided that $(X_t,\ t \in \mathbb{R})$ should satisfy a local irregularity condition which will be specified and dis-

Received October 1995; revised August 1996.
AMS 1991 *subject classification*. 62G05.
Key words and phrases. Nonparametric estimation and prediction, parametric rate, regression, density, large deviation inequality.

cussed below. In that case the MSE is $O(1/T)$ and the supnorm rate is $o(\ln_k T(\ln T/T)^{1/2})$, $k \geq 1$. In the following, these rates will be called parametric rates since they correspond to the rate of the MLE in classical parametric continuous time models such as diffusion processes (see, e.g., [16]). It should be noticed that diffusion processes satisfy the local irregularity condition considered in this paper (see Section 2).

In the following, we will show that parametric rates are achieved by density estimators, regression estimators and nonparametric predictors.

The results are easily obtained for the MSE. On the contrary in the context of an a.s. uniform convergence, proofs use a special Borel–Cantelli lemma for continuous time processes together with a sharp exponential type inequality.

Finally we prove that a suitable sampling allows preserving parametric rates. This paradoxical result is obviously theoretical and must be clarified in a numerical context. Some comments about sampling appear at the end of Section 7.

The first parametric rate was pointed out by Castellana and Leadbetter [10] in 1986. They noticed that the variance of a density estimator is $O(1/T)$ for a real "locally irregular" process observed over the time interval $[0, T]$.

For other results of this kind, we refer to [5, 6], [14], [15] and [11]. Kutoyants [14] has recently shown that, if (X_t) is a diffusion process, the rate $O(1/T)$ is minimax and asymptotically efficient estimators can be constructed. Leblanc [15] has obtained parametric rates in L^p norm for the wavelets density estimator. Cheze [11] has obtained parametric rates for the kernel regression estimator. As far as we know, there are no results of this kind about almost sure uniform convergence until now.

The rest of the present paper is organized as follows. Section 2 contains some assumptions and notation. The main mathematical tools are introduced in Section 3. Sections 4, 5 and 6 are devoted, respectively, to density estimation, regression estimation and prediction. Section 7 deals with sampling and proofs are given in Section 8.

2. Notation and assumptions. Let $Z_t = (X_t, Y_t)$, $t \in \mathbb{R}$ be a $\mathbb{R}^d \times \mathbb{R}^{d'}$-valued measurable strictly stationary process defined on a probability space $(\Omega, \mathscr{A}, P)$. Let M be a locally bounded Borelian real function, defined on $\mathbb{R}^{d'}$, and such that $(\omega, t) \mapsto M(Y_t(\omega))$ belong to $L^2(P \otimes \lambda_T)$ for every positive T, where λ_T denotes Lebesgue measure over $[0, T]$.

Let us assume that (Z_0, Z_u) has a density $f_{(Z_0, Z_u)}$ for each strictly positive u and consider the following functional parameters:

$$f(x) = \int_{\mathbb{R}^{d'}} f_{Z_0}(x, y)\, dy, \qquad x \in \mathbb{R}^d \tag{2.1}$$

and

$$\varphi(x) = \int_{\mathbb{R}^{d}} M(y) f_{Z_0}(x, y)\, dy, \qquad x \in \mathbb{R}^d. \tag{2.2}$$

We may use f and φ for defining a version of the regression function $E(M(Y_0)\,|\,X_0 = \cdot)$ by setting

$$r(x) = \begin{cases} \varphi(x)/f(x), & \text{if } f(x) > 0, \\ EM(Y_0), & \text{if } f(x) = 0. \end{cases} \tag{2.3}$$

We wish to estimate f and r given the data $(Z_t,\ 0 \le t \le T)$. For this purpose we shall use a kernel $K = K_0^{\otimes d}$, where K_0 is a one-dimensional bounded symmetric density with continuous derivative. Furthermore, we assume that K_0 is strictly positive over $(-a, a)$, where $a \neq 0$, and vanishes elsewhere.

In all the following, the smoothing parameter is $T^{-1/4}$; this choice is optimal for our purpose. Thus we set

$$K_T(u) = T^{d/4}K(uT^{1/4}), \qquad u \in \mathbb{R}^d. \tag{2.4}$$

and our Parzen–Rosenblatt type estimators (cf. [17] and [20]) are defined by

$$f_T(x) = \frac{1}{T}\int_0^T K_T(x - X_T)\, dt, \qquad x \in \mathbb{R}^d \tag{2.5}$$

and

$$r_T(x) = \begin{cases} \varphi_T(x)/f_T(x), & \text{if } f_T(x) > 0, \\ \dfrac{1}{T}\displaystyle\int_0^T M(Y_t)\, dt, & \text{if } f_T(x) = 0 \end{cases} \tag{2.6}$$

with

$$\varphi_T(x) = \frac{1}{T}\int_0^T M(Y_t) K_T(x - X_t)\, dt, \qquad x \in \mathbb{R}^d, \tag{2.7}$$

where the integrals in (2.5) and (2.7) are taken in the usual sense.

Regarding the functional parameters, we suppose that f_Z, f and φ are twice continuously differentiable and bounded as well as all their second partial derivatives.

In order to express the irregularity of sample paths we shall use the following functions:

$$g_u(x, x') = f_{(X_0, X_u)}(x, x') - f(x)f(x'), \qquad x, x' \in \mathbb{R}^d; \quad u > 0, \tag{2.8}$$

$$\begin{aligned} G_u(x, x') = \int_{\mathbb{R}^{2d'}} & M(y)M(y') \\ & \times \left[f_{(Z_0, Z_u)}(x, y; x', y') - f_{Z_0}(x, y) f_{Z_u}(x', y') \right] dy\, dy', \\ & \qquad x, x' \in \mathbb{R}^d; \quad u > 0. \end{aligned} \tag{2.9}$$

Now we will employ the following assumptions.

A. The expression $h(x, x') = \int_{(0, +\infty)} |g_u(x, x')|\, du$ does exist for all $x, x' \in \mathbb{R}^d$, is bounded and is continuous over the diagonal of $\mathbb{R}^d$.

B. $H(x, x') = \int_{(0, +\infty)} |G_u(x, x')|\, du$ does exist for all x, $x' \in \mathbb{R}^d$, is bounded and is continuous over the diagonal of $\mathbb{R}^d$.

C. The function g_u is continuous over the diagonal for all $u \neq 0$ and $\|g_u\|_\infty$ is integrable over $(0, +\infty)$.

D. The function f is ultimately decreasing with respect to $\|x\|$ and $\sup_{0 \le t \le 1} \|X_t\|^b \in L^1(P)$ for some $b > 0$.

E. (X_t) is geometrically strongly mixing (GSM); that is,

$$\alpha(v) = \sup_{\substack{A \in \sigma(X_t,\, t \le 0),\\ B \in \sigma(X_t,\, t \ge v)}} |P(A \cap B) - P(A)P(B)| \le \gamma \rho^v, \qquad v > 0,$$

where $\gamma > 0$, $0 < \rho < 1$ are constant.

F. (Z_t) is GSM.

G. (Z_t) is geometrically φ-reverse mixing; that is,

$$\sup |P(B \mid A) - P(B)| \le a \rho^v, \qquad v > 0,$$

where the supremum is taken over $A \in \sigma(\xi_t,\ t \ge v)$ with $P(A) > 0$ and $B \in \sigma(\xi_t,\ t \le 0)$ and where $a > 0$, $\rho \in (0, 1)$ are constant.

H. The term g_u is continuous over the diagonal of $\mathbb{R}^d$ and $\|g_u\|_\infty \le \Pi(u)$ where $(1 + u)\Pi(u)$ is integrable over $(0, +\infty)$ and $u\Pi(u)$ is bounded and ultimately decreasing.

I.

$$\sup_{(y, z) \in \mathbb{R}^{2d}} \left| \int_0^\infty g_u(y, z)\, du - \sum_{k=1}^{\infty} \delta_n g_{k\delta_n}(y, z) \right|_{\delta_n \to 0} \to 0.$$

Comments on assumptions. Assumption A contains an asymptotic independence condition (take u large) and a local irregularity condition (take u small). This local condition means that the information respectively provided by $(X_t, X_{t+\delta})$ and X_t differ significantly even if δ is small. It also means that sample paths are not smooth (see Section 4). Finally, local irregularity of the observed sample paths provides more information than discrete data. This partly explains parametric rates which appear below.

Assumption B is similar to assumption A and assumption C is slightly stronger than A.

Assumption D is a technical condition related to the extreme values of (X_t).

Assumptions E and F are classical mixing conditions which are satisfied by stationary diffusion processes (see [26]). Assumption G is stronger and is satisfied by some Markov processes (see [13]).

Assumption H is slightly stronger than C and I is a mild regularity condition. Assumptions H and I are valid if, for instance, (X_t) is an Ornstein–Uhlenbeck process.

3. Preliminary results. In the present section we state two lemmas which will be used below. Note that these propositions are interesting by themselves.

The first result is a Borel–Cantelli type lemma for continuous time processes.

LEMMA 3.1. *Let* $(U_t,\ t \geq 0)$ *be a real continuous time process such that we have the following*.

(a) *For all* $\eta > 0$, *there exists a real decreasing function* φ_n *integrable over* $[0, +\infty)$ *and satisfying*

$$P(|U_t| \geq \eta) \leq \varphi_\eta(t), \qquad t \geq 0; \tag{3.1}$$

(b) *The sample paths of* (U_t) *are uniformly continuous with probability* 1; *then*

$$\lim_{T \to +\infty} U_T = 0 \qquad a.s. \tag{3.2}$$

The following lemma provides a large deviation inequality for bounded dependent random variables.

LEMMA 3.2. *Let* $(\xi_i,\ i \in \mathbb{Z})$ *be a real-valued zero-mean strictly stationary sequence such that* $\sup_{1 \leq i \leq n} \|\xi_i\|_\infty \leq b$. *Then for all integer* q *such that* $1 \leq q \leq (n/2)$ *and all* $\varepsilon > 0$,

$$\begin{aligned} &P(|\xi_1 + \cdots + \xi_n| > n\varepsilon) \\ &\quad \leq 4\exp\left(-\frac{\varepsilon^2}{8v^2(q)}q\right) + 22\left(1 + \frac{4b}{\varepsilon}\right)^{1/2} q\alpha\left(\left[\frac{n}{2q}\right]\right), \end{aligned} \tag{3.3}$$

where

$$v^2(q) = \frac{2}{p^2}E\left(\xi_1 + \cdots + \xi_{[p]} + (p - [p])\xi_{[p+1]}\right)^2 + \frac{b\varepsilon}{2}, \tag{3.4}$$

$p = n/2q$ *and*

$$\alpha\left(\left[\frac{n}{2q}\right]\right) = \sup_{\substack{A \in \sigma(\xi_i,\ i \leq 0), \\ B \in \sigma\left(\xi_i,\ i \geq \left[\frac{n}{2q}\right]\right)}} |P(A \cap B) - P(A)P(B)|.$$

This lemma is an improvement of an inequality in [3]. Some ideas in the proof are taken from [19] and [24]. It should be noticed that (3.3) is an explicit inequality which is valid even if the sequence (ξ_i) is not strongly mixing.

4. Density estimation. This section deals with estimation of the finite dimensional distributions of $(X_t,\ t \in \mathbb{R})$. In fact we will only consider estimation of the marginal density since estimation of the density $f_{(X_{t_1}, \ldots, X_{t_l})}$ of $(X_{t_1}, \ldots, X_{t_l})$ reduces to estimation of the dl-dimensional marginal density of the process $(X_{t_1+s}, \ldots, X_{t_l+s})$, $s \in \mathbb{R}$.

The first result concerning f_T is an extension of a Castellana–Leadbetter ([10]) result.

PROPOSITION 4.1. *If* A *holds, then*

$$E(f_T(x) - f(x))^2 = O\left(\frac{1}{T}\right), \qquad x \in \mathbb{R}^d. \tag{4.1}$$

It can be shown (cf. [7]) that A is necessary for the full rate $1/T$. In particular if (X_t) is a Gaussian process we have the following alternative.

PROPOSITION 4.2. *Let (X_t) be a real stationary Gaussian process, continuous in mean square and such that*

$$|\mathrm{Cov}(X_0, X_u)| < VX_0, \qquad u > 0, \tag{4.2}$$

and

$$\int_{u_0}^{+\infty} |\mathrm{Cov}(X_0, X_u)|\, du < \infty, \qquad u_0 > 0. \tag{4.3}$$

(i) *If there exists $u_1 > 0$ such that*

$$\int_0^{u_1} \left[E(X_u - X_0)^2\right]^{-1/2} du < \infty \tag{4.4}$$

then

$$TE(f_T(x) - f(x))^2 \to l < \infty, \qquad x \in \mathbb{R}.$$

(ii) *If there exists $u_1 > 0$ such that*

$$\int_0^{u_1} \left[E(X_u - X_0)^2\right]^{-1/2} du = +\infty \tag{4.5}$$

then

$$TE(f_T(x) - f(x))^2 \to +\infty.$$

An easy consequence of Proposition 4.2 is the following: if (X_t) has differentiable sample paths, then (4.5) holds. In fact, under some regularity conditions, it can be shown that $E(f_T(x) - f(x))^2 \simeq \ln T/T$ (see [2]). On the contrary (4.4) is satisfied by regular diffusion processes and in particular by the Ornstein–Uhlenbeck process.

An extreme example should be a Gaussian process with an autocorrelation $\rho(u)$ such that

$$1 - \rho(u) \simeq (\ln u)^{\beta - 1}, \qquad u \to 0(+), 0 < \beta < 1.$$

Then sample paths are not continuous; however (4.4) holds provided (4.2) and (4.3) are satisfied.

Concerning the supnorm we have the following.

PROPOSITION 4.3. *If $d = 1$ and if* C, D, E *hold then*

$$\sup_{x\in\mathbb{R}} |f_T(x) - f(x)| = o\left(\ln_k T\left(\frac{\ln T}{T}\right)^{1/2}\right) \quad a.s. \tag{4.6}$$

for all positive integers k.

A similar result may be obtained for $d > 1$ but with stronger assumptions (see [7]).

5. Regression estimation. The asymptotic quadratic error of r_T is given by the following proposition.

PROPOSITION 5.1. *Suppose that* A *and* B *hold and that $f(x) > 0$. Then we have the following*:

(i) *if M is bounded we have*

$$E(r_T(x) - r(x))^2 = O\left(\frac{1}{T}\right); \tag{5.1}$$

(ii) *if* $\exp[\sup_{0\le t\le 1}|M(Y_t)|^s] \in L^1(P)$ *for some $s > 0$, we have*

$$E(r_T(x) - r(x))^2 = O\left(\frac{(\ln T)^{2/s}}{T}\right). \tag{5.2}$$

We now turn to uniform convergence. Note that uniform convergence of r_T may be obtained over compact sets but, in general, not over the whole space since the behavior of r for $\|x\|$ large may be very irregular. Now if $d = 1$ and if Δ is a compact interval such that $\inf_{x\in\Delta} f(x) > 0$, we obtain the following rate.

PROPOSITION 5.2. *If* A *and* B *hold*, *(Z_t) is GSM and M is bounded*, *then*

$$\sup_{x\in\Delta} |r_T(x) - r(x)| = o\left(\ln_k T\left(\frac{\ln T}{T}\right)^{1/2}\right) \quad a.s. \tag{5.3}$$

for all positive integers k.

6. Prediction. Let $(\xi_t,\ t\in\mathbb{R})$ be a strictly stationary measurable process. Given the data $(\xi_t,\ 0\le t\le T)$, we would like to predict the nonobserved square integrable real random variable $\zeta_{T+l} = M(\xi_{T+l})$ where the horizon l satisfies $0 < l < T$ and where M is measurable and bounded on compact sets.

In order to simplify the exposition, we suppose that (ξ_t) is a real Markov process with sample paths which are continuous on the left. We now define the kernel predictor by

$$\hat{\zeta}_{T+l} = \varphi_{T-l}(\xi_T)/f_T(\xi_T), \tag{6.1}$$

where

$$\varphi_{T-l}(\xi_T) = \int_0^{T-l} M(\xi_{t-l}) K_T(\xi_T - \xi_t)\, dt, \tag{6.2}$$

and

$$f_T(\xi_T) = \int_0^T K_T(\xi_T - \xi_t)\,dt. \tag{6.3}$$

Note that the special form of K together with left continuity of paths entail that $f_T(\xi_T)$ is strictly positive.

We now study the asymptotic behavior of $\hat{\zeta}_{T+l}$ as T tends to infinity, l remaining fixed. As usual $\hat{\zeta}_{T+l}$ is considered to be an approximation of

$$r(\xi_T) = E(\zeta_{T+l} \mid \xi_s, s \le T) = E(\zeta_{T+l} \mid \xi_T). \tag{6.4}$$

Let us first indicate an almost sure convergence rate

PROPOSITION 6.1. *If the conditions in Proposition* 5.2 *hold with* $Z_t = (M(\xi_{t+l}), \xi_t)$ *then*

$$\left|\hat{\zeta}_{T+l} - r(\xi_T)\right| \mathbb{1}_{\xi_T \in \Delta} = o\left(\ln_k T\left(\frac{\ln T}{T}\right)^{1/2}\right) \quad a.s., k \ge 1. \tag{6.5}$$

In order to to study convergence in mean square we need for technical reasons the stronger mixing assumption G.

On the other hand, we slightly modify our predictor by setting

$$\zeta^*_{T+l} = r_{T'}(\zeta_T), \tag{6.6}$$

where $T' = T - l - (\ln T)\ln_2 T$. Then we have the following result.

PROPOSITION 6.2. *If* A, B *and* G *hold*, *then*

$$E\left[(\zeta^*_{T+l} - r(\xi_T))^2 \mathbb{1}_{\zeta_T \in \Delta}\right] = O\left(\frac{1}{T}\right), \tag{6.7}$$

where Δ *is any compact interval such that* $\inf_{x \in \Delta} f(x) > 0$.

7. Sampling. In continuous time, data are often collected by using a sampling scheme. We consider a process $(Z_t,\ t \in \mathbb{R})$ with "irregular" paths observed at sampling instants.

In order to model the fact that observations are frequent during a long time, we assume that these sampling instants are $\delta_n, 2\delta_n, \dots, n\delta_n$ where $\delta_n \to 0$ and $T_n = n\delta_n \to \infty$ as $n \to \infty$. (T_n) is a given sequence of positive real numbers and one must select (δ_n) so that the rate of kernel estimators should be "parametric."

Here the kernel estimators are defined by

$$f_n^*(x) = \frac{1}{n}\sum_{j=1}^{n} K_{T_n}(x - X_{j\delta_n}), \qquad x \in \mathbb{R}^d, \tag{7.1}$$

and

$$r_n^*(x) = \begin{cases} \varphi_n^*(x)/f_n^*(x), & \text{if } f_n^*(x) > 0, \\ \dfrac{1}{n}\sum_{j=1}^{n} M(Y_{j\delta_n}), & \text{if } f_n^*(x) = 0, \end{cases} \tag{7.2}$$

where

$$\varphi_n^*(x) = \frac{1}{n}\sum_{j=1}^{n} M(Y_{j\delta_n}), \qquad x \in \mathbb{R}^d. \tag{7.3}$$

We then have the following propositions.

PROPOSITION 7.1. (i) *If* A, G *and* H *hold then*

$$E(f_n^*(x) - f(x))^2 = O\left(\frac{1}{T_n}\right) \tag{7.4}$$

provided $f(x) > 0$ *and* $\delta_n = T_n^{-d/4}$.

(ii) *If in addition* B *holds and* M *is bounded then*

$$E(r_n^*(x) - r(x))^2 = O\left(\frac{1}{T_n}\right). \tag{7.5}$$

PROPOSITION 7.2. (i) *Under conditions of Proposition* 4.3 *and if* $\delta_n = T_n^{-1/4}$, *then*

$$\sup_{x\in\mathbb{R}^d} |f_n^*(x) - f(x)| = o\left(\ln_k T_n \left(\frac{\ln_k T_n}{T_n}\right)^{1/2}\right) \quad a.s., k \geq 1. \tag{7.6}$$

(ii) *Under conditions of Proposition* 5.2 *and if* $\delta_n = T_n^{-1/4}$, *then*

$$\sup_{x\in\Delta} |r_n^*(x) - f(x)| = o\left(\ln_k T_n \left(\frac{\ln_k T_n}{T_n}\right)^{1/2}\right) \quad a.s., k \geq 1. \tag{7.7}$$

Note that the choice of δ_n in Proposition (7.1) implies $T_n = n^{4/(d+4)}$; consequently, although the obtained rates are parametric with respect to T_n they are, in fact, nonparametric with respect to n. However, it is interesting to note that the n discrete observations are sampled over an interval of length $T_n \ll n$. On the other hand, it is easy to see that the sequence (δ_n) is optimal in the following sense: if (δ_n') is such that the MSE of (f_n) is $O(1/T_n)$, then $\delta_n' = O(\delta_n)$. Thus $n = T_n^{(d+4)/4}$ is asymptotically the smallest sample size which ensures the rate $1/T_n$ over $[0, T_n]$. Moreover, this rate cannot be improved.

8. Proofs.

PROOF OF LEMMA 3.1. First let (T_n) be a sequence of real numbers which satisfies $T_{n+1} - T_n \geq a > 0$ where a is some constant.

Since φ_η is decreasing we have

$$\int_{T_N}^{+\infty} \varphi_\eta(t)\,dt \ge \sum_{n\ge N} (T_{n+1} - T_n)\varphi_\eta(T_{n+1}) \ge a \sum_{n\ge N} \varphi_\eta(T_{n+1}),$$

thus $\sum_n \varphi_\eta(T_n) < +\infty$ and the classical Borel–Cantelli lemma yields $P(\lim_n \sup\{|U_{T_n}| > \eta\}) = 0$, $\eta > 0$ which in turn implies $U_{T_n} \to 0$ a.s.

Let now (T_n) be any sequence of real numbers satisfying $T_n \uparrow +\infty$.

To each positive integer k we may associate a subsequence $(T_p^{(k)})$ of (T_n) defined as follows:

$$T_1^{(k)} = T_{n_1} \quad \text{where } n_1 = 1,$$

$$T_2^{(k)} = T_{n_2} \quad \text{where } T_{n_2} - T_{n_1} \ge \frac{1}{k},\ T_{n_2} - T_{n_2-1} < \frac{1}{k},$$

$$\vdots$$

$$T_p^{(k)} = T_{n_p} \quad \text{where } T_{n_p} - T_{n_{p-1}} \ge \frac{1}{k},\ T_{n_p} - T_{n_p-1} < \frac{1}{k},$$

$$\vdots$$

The first part of the current proof shows that $U_{T_p^{(k)}} \to_{p\to\infty} 0$ a.s. for each k.

Now let us set

$$\Omega_0 = \left\{\omega : t \mapsto U_t(\omega) \text{ is uniformly continuous, } U_{T_p^{(k)}}(\omega) \to 0,\ k \ge 1\right\};$$

clearly $P(\Omega_0) = 1$.

Now if $\omega \in \Omega_0$ and $\eta > 0$, there exists $k = k(\eta, \omega)$ such that $|t - s| \le 1/k$ implies $|U_t(\omega) - U_s(\omega)| < \eta/2$. Consider the sequence $T_p^{(k)}$: for each p and each n such that $n_p \le n < n_{p/1}$, we have $|T_n - T_{n_p}| < 1/k$, and consequently $|U_{T_n}(\omega)| < \eta$ for n large enough. This is valid for each $\eta > 0$ and each $\omega \in \Omega_0$; thus $U_{T_n} \to 0$ a.s. □

Proof of Lemma 3.2. We shall use the following lemma essentially due to Bradley [8].

Lemma 8.1. *Let (X, Y) be an $\mathbb{R}^d \times R$-valued random vector such that $Y \in L^p(P)$ for some $p \in [1, +\infty]$. Let c be a real number such that $\|Y + c\|_p > 0$ and $\xi \in (0, \|Y + c\|_p]$. Then there exists a random variable Y^* such that*:

$$\text{(i)}\ \ P_{Y^*} = P_Y \text{ and } Y^* \text{ is independent of } X,$$

$$\text{(ii)}\ \ P(|Y^* - Y| > \xi) \le 11\left(\xi^{-1}\|Y + c\|_p\right)^{p/(2p+1)} \times \left[\alpha(\sigma(X), \sigma(Y))\right]^{2p/(2p+1)}. \tag{8.1}$$

In the original statement of this lemma, 11 is replaced by 18 and $c = 0$ but the proof is not different.

Now in order to prove Lemma 3.2, we consider the auxiliary continuous time process $\eta_t = \xi_{[t+1]}$, $t \in \mathbb{R}$. We clearly have $\sum_{i=1}^n \xi_i = \int_0^n \eta_u \, du$.

Let us now define "blocks" as follows:

$$V_1 = \int_0^p \eta_u \, du, \qquad V_1' = \int_p^{2p} \eta_u \, du,$$

$$\vdots$$

$$V_q = \int_{2(q-1)p}^{(2q-1)p} \eta_u \, du, \qquad V_q' = \int_{(2q-1)p}^{2qp} \eta_u \, du,$$

where $p = n/2q$.

Using Lemma 8.1 recursively we may define independent random variables $W_1, \dots, W_q$ such that

$$P\big(|W_j - V_j| > \zeta\big) \le 11 \left(\frac{\|V_j + c\|_\infty}{\zeta} \right)^{1/2} \alpha([p]), \tag{8.2}$$

where $c = \delta bp$, $\zeta = \min(n\varepsilon/4q, (\delta - 1)bp)$ for some $p > 1$ which will be specified below.

Note that, for each j,

$$\|V_j + c\|_\infty \ge c - \|V_j\|_\infty \ge (\delta - 1) bp > 0$$

so that $0 < \zeta \le \|V_j + c\|_\infty$ as required in Lemma 8.1.

Now according to the choice of c and ζ, (8.2) may be written as

$$P\big(|W_j - V_j| > \zeta\big) \le 11 \left(\frac{(\delta + 1) bp}{\min(n\varepsilon/4q, (\delta - 1) bp)} \right)^{1/2} \alpha([p]).$$

If $\delta = 1 + \varepsilon/2b$ then

$$P\big(|W_j - V_j| > \zeta\big) \le 11 \left(1 + \frac{4b}{\varepsilon} \right)^{1/2} \alpha([p]). \tag{8.3}$$

On the other hand we may apply Bernstein's inequality ([7]) to $\sum_1^q W_j$. We obtain

$$P\left(\left| \sum_1^q W_j \right| > \frac{n\varepsilon}{4} \right) \le 2 \exp\left(- \frac{n^2 \varepsilon^2 / 16}{4 \sum_1^q EW_j^2 + 2bpn\varepsilon/4} \right). \tag{8.4}$$

Now since $P_{W_j} = P_{V_j}$ we have

$$EW_j^2 = EV_j^2 = E\left(\int_{jp}^{(j+1)p} \eta_u \, du \right)^2. \tag{8.5}$$

By (8.4) and (8.5) it follows that

$$P\left(\left| \sum_1^q W_j \right| > \frac{n\varepsilon}{4} \right) \le 2 \exp\left(- \frac{\varepsilon^2 q}{8 v^2(q)} \right), \tag{8.6}$$

where $v^2(q)$ is given by (3.4).

On the other hand, elementary computations show that

$$(8.7)\quad P(|\xi_1 + \cdots + \xi_n| > n\varepsilon) \le P\left(\left|\sum_1^q V_j\right| > \frac{n\varepsilon}{2}\right) + P\left(\left|\sum_1^q V_j'\right| > \frac{n\varepsilon}{2}\right)$$

and that

$$(8.8)\quad P\left(\left|\sum_1^q V_j\right| > \frac{n\varepsilon}{2}\right) \le P\left(\left|\sum_1^q W_j\right| > \frac{n\varepsilon}{4}\right) + \sum_1^q P(|V_j - W_j| > \zeta).$$

Finally collecting the bounds in (8.7), (8.3), (8.6), (8.8) and the analogous bound for the V_j''s, we obtain (3.3). The proof of Lemma 3.2 is therefore complete. □

Proofs of Propositions 4.1 and 4.2 follow from standard argument and are therefore omitted.

PROOF OF PROPOSITION 4.3. (i) Let us consider the process

$$U_T = \frac{1}{\ln_k T}\left(\frac{T}{\ln T}\right)^{1/2}(f_T(x) - Ef_T(x)),$$

where k and x are fixed.

We first show that

$$(8.9)\qquad P(|U_T| \ge \eta) \le c_\eta T^{-c_\eta'(\ln_k T)^2}, \qquad \eta > 0,$$

where c_η and c_η' are strictly positive and do not depend on x. For this purpose, we shall use Lemma 3.2. Let us set

$$\xi_{jn}' = \frac{1}{\delta}\int_{(j-1)\delta}^{j\delta} K_T(x - X_t)\,dt; \qquad j = 1, \dots, n,$$

where $\delta = n^{-1}T$ and $n = [T]$. Thus we have

$$f_T(x) - Ef_T(x) = \frac{1}{n}\sum_{j=1}^n (\xi_{jn}' - E\xi_{jn}') = \frac{1}{n}\sum_{j=1}^n \xi_{jn}.$$

We now evaluate $v^2(q)$. First

$$\begin{aligned} p\delta V&\left(\frac{1}{p\delta}\int_0^{p\delta} K_T(x - X_t)\,dt\right) \\ &= 2p\delta\int_{(0,\,p\delta)}\left(1 - \frac{u}{T}\right)du\int_{\mathbb{R}^{2d}} K_T(x-y)K_T(x-z)g_u(y,z)\,dy\,dz \\ &\le 2\int_0^\infty \|g_u\|_\infty\,du \end{aligned}$$

by B.

Then using (3.4) it is easy to see that

$$v^2(q) \le \frac{4}{p\delta}\int_0^\infty \|g_u\|_\infty\,du + \|K\|_\infty \varepsilon T^{1/4}.$$

We now choose $q = [(n\varepsilon T^{1/4})/2] + 1$, hence $p = 1/(\varepsilon T^{1/4} + \theta)$ where $0 < \theta \le 2$ and $v^2(q) \le a\varepsilon T^{1/4}$ where a is constant.

Substituting in (3.3), we obtain

$$
\begin{aligned}
P\left(\left|\sum_{j=1}^{n} \xi_{jn}\right| > n\varepsilon\right) &\le 4\exp\left(-c\varepsilon q T^{-1/4}\right) \\
&\quad + 22\left(1 + \frac{8T^{1/4}\|K\|_\infty}{\varepsilon}\right)^{1/2} q\alpha([p]) \qquad (8.10)\\
&\le u_T + v_T,
\end{aligned}
$$

where c is a strictly positive constant.

We now choose $\varepsilon = (\ln T/T)^{1/2}\ln_k T\eta$ so that

$$u_T \le 4T^{-c_1(\ln_k T)^2}, \tag{8.11}$$

where c_1 is constant.

On the other hand

$$v_T \le 22\left[1 + \left(\frac{8T^{1/4}\|K\|_\infty}{\varepsilon}\right)^{1/2}\right] q\gamma\rho^p,$$

which is clearly $O(T^{-c_1(\ln_k T)^2})$; hence (8.9).

(ii) We now prove that (U_T) satisfies condition (b) in Lemma 3.1. Let us consider

$$V_T = \frac{1}{\ln_k T}\left(\frac{T}{\ln T}\right)^{1/2}\frac{1}{T}I_T,$$

where

$$I_T = \int_0^T K\left(T^{1/4}(x - X_t)\right)dt$$

then $U_T = V_T - EV_T$.

We are going to prove that

$$\sup_{x\in\mathbb{R}^d,\ \omega\in\Omega} \left|V_T(x,\omega) - V_S(x,\omega)\right| \le \Lambda|T - S|, \tag{8.12}$$

$T > 1$, $S > 1$, where Λ is constant.

Note that (8.12) implies a similar result for (U_T) since

$$|EV_T - EV_S| \le E|V_T - V_S| \le \sup_{x,\omega}|V_T - V_S|.$$

Now we set

$$\ln V_T = A_T + B_T,$$

where

$$A_T = -\ln_{k+1} T + \frac{1}{2}\ln\left(\frac{T}{\ln T}\right) - \ln T$$

and

$$B_T = \ln I_T,$$

where I_T is supposed to be strictly positive.

The derivative of A_T is clearly $O(T^{-1})$.

Concerning B_T we first have

$$I'_T = \int_0^T \frac{\partial K}{\partial T}\big(T^{1/4}(x - X_t)\big)\,dt + K\big(T^{1/4}(x - X_T)\big).$$

On the other hand

$$\frac{\partial K_0}{\partial T}\big(T^{1/4}(x_j - X_{t,j})\big) = \frac{1}{4}T^{-3/4}(x_j - X_{t,j})K'_0\big(T^{1/4}(x_j - X_{t,j})\big)$$

for each component $x_j - X_{t,j}$ of the vector $x - X_t$. Then since $K'_0(u) = 0$ if $|u| \geq a$ we infer that

$$\left|\frac{\partial K_0}{\partial T}\big(T^{1/4}(x_j - X_{t,j})\big)\right| \leq \frac{a\|K'_0\|_\infty}{4}\frac{1}{T},$$

therefore

$$\left|\frac{\partial K}{\partial T}\big(T^{1/4}(x - X_t)\big)\right| = O\left(\frac{1}{T}\right)$$

and finally $|I'_T| = O(1)$. Hence,

$$|(\ln V_T)'| \leq c_1 T^{-1} + c_2 I_T^{-1},$$

where c_1 and c_2 are positive constants.

Using the relation $V'_T = V_T(\ln V_T)'$, it is then easy to see that

$$|V'_T| \leq cT^{-1/2}, \qquad T > 1, \tag{8.13}$$

where c is constant. By continuity the bound in (8.13) remains valid if $I_T = 0$.

Finally V'_T is bounded, hence (8.12).

(iii) We are now in a position to apply Lemma 3.1: using (8.9) and (8.12), we obtain

$$\frac{1}{\ln_k T}\left(\frac{T}{\ln T}\right)^{1/2}(f_T(x) - Ef_T(x)) \to 0 \quad \text{a.s.}$$

and since the bias is $O(T^{-1/2})$, we find

$$\frac{1}{\ln_k T}\left(\frac{T}{\ln T}\right)^{1/2}|f_T(x) - f(x)| \to 0 \quad \text{a.s.}$$

(iv) Using a covering of $\mathscr{B}_T = \{x: \|x\| \leq T^\gamma\}$, with $\gamma > 0$, and the fact that K is Lipschitz, it is easy to show that the convergence is uniform over $\mathscr{B}_T$. Finally condition C allows establishing uniform convergence over $\mathbb{R}^d$. For details we refer to [7].

Proof of Proposition 5.1. (i) Since $f(x) > 0$, the following decomposition, where x is omitted, is valid for T large enough:

$$r_T - \frac{E\varphi_T}{Ef_T} = r_T\frac{Ef_T - f_T}{Ef_T} + \frac{\varphi_T - E\varphi_T}{Ef_T}, \tag{8.14}$$

hence

$$E\left(r_T - \frac{E\varphi_T}{Ef_T}\right)^2 \le \frac{2\left(\|m\|_\infty^2 + 1\right)}{\|Ef_T\|^2}\left[Vf_T + V\varphi_T\right].$$

By Proposition 4.1 $Vf_T = O(T^{-1})$. Similarly it can be checked that $V\varphi_T = O(T^{-1})$.

Now

$$r - \frac{E\varphi_T}{Ef_T} = r\frac{Ef_T - f}{Ef_T} + \frac{\varphi - E\varphi_T}{Ef_T},$$

thus

$$\left(r - \frac{E\varphi_T}{Ef_T}\right)^2 \le \frac{2\left(\|m\|_\infty^2 + 1\right)}{(Ef_T)^2}\left[(Ef_T - f)^2 + (\varphi - E\varphi_T)^2\right].$$

Again by Proposition 4.1 $(Ef_T - f)^2 = O(T^{-1})$ and similarly, $E(\varphi - E\varphi_T)^2 = O(T^{-1})$, hence (5.1).

(ii) In order to establish (5.2) we need the following lemma.

LEMMA 8.2. *Let $(\xi_t, t \in \mathbb{R}_+)$ be a stochastic process such that $\sup_{0\le t\le T}|\xi_t|$ is measurable for each positive T and which satisfies the condition*

$$c = \sup_{j\ge 0} E\left(\exp a|\eta_j|^s\right) < \infty \tag{8.15}$$

for some $a > 0$, some $s > 0$ and where $\eta_j = \sup_{j\le t\le j+1}|\xi_t|$.

Then

$$E\left(\sup_{0\le t\le T}|\xi_t|^p\right) \le \left(\frac{\ln c([T]+1)}{a}\right)^{p/s}, \qquad p > 0, T > 0. \tag{8.16}$$

PROOF. Consider the strictly concave function

$$l(x) = a^{-p/s}(\ln x)^{p/s}, \qquad x > 0$$

and set

$$\zeta = \sup_{0\le t\le [T]+1}|\xi_t|^p = \max_{0\le j\le [T]}\eta_j^p;$$

then Jensen's inequality entails

$$\begin{aligned} E\left(\sup_{0\le t\le T}|\xi_t|^p\right) &\le E\zeta \le E\left[l\left(l^{-1}(\zeta)\right)\right] \\ &\le lE\left[l^{-1}(\zeta)\right] \\ &\le lE\left[\max_{0\le j\le [T]}\exp a\eta_j^s\right] \\ &\le lE\left[\sum_{j=0}^{[T]}\exp a\eta_j^s\right] \\ &\le a^{-p/s}\left[\ln([T]+1)c\right]^{p/s}. \end{aligned}$$

□

We now apply Lemma 8.2 to $\xi_t = M(Y_t)$, $t \geq 0$. We have, for every $p > 0$,

$$E(r_T^{2p}) = E\left(\int_0^T p_{tT}\,\xi_t\,dt\right)^{2p},$$

where

$$p_{tT} = \begin{cases} K_T(x - X_t)\Big/\int_0^T K_T(x - X_t)\,dt, & \text{if } f_T(x) > 0,\\ \dfrac{1}{T}, & \text{if } f_T(x) = 0.\end{cases}$$

Thus

$$E(r_T^{2p}) \leq E\Big(\sup_{0 \leq t \leq T} |\xi_t|^{2p}\Big)$$

and by Lemma 8.2 it follows that

$$\big(E(r_T^{2p})\big)^{1/p} \leq \left\{[\ln c([T] + 1)]^{2p/s}\right\}^{1/p}. \tag{8.17}$$

On the other hand, setting $q = 1 + \varepsilon_T$ where $(1/p) + (1/q) = 1$ we obtain

$$\left(E|f_T - Ef_T|^{2q}\right)^{1/q} \leq \left((2\|K\|_\infty)^{2\varepsilon_T} T^{2\varepsilon_T/4} V f_T\right)^{1/q}$$

and since $Vf_T \leq kT^{-1}$ where k is constant, we get

$$\left(E|f_T - Ef_T|^{2q}\right)^{1/q} \leq \left(2\|K\|_\infty^{2\varepsilon_T} k\right)^{1/q} T^{(\varepsilon_T/2 - 1)/q}.$$

Now we choose $\varepsilon_T \to_{T\to\infty} 1$ and more precisely

$$\frac{1 - (\varepsilon_T/2)}{1 + \varepsilon_T} = 1 - \frac{1}{\ln T}.$$

Therefore since $T^{1/\ln T} = O(1)$, we have

$$\left(E|f_T - Ef_T|^{2q}\right)^{1/q} = O\left(\frac{1}{T}\right). \tag{8.18}$$

Now using Hölder's inequality together with (8.17) and (8.18), we obtain

$$E\left(r_T^2 (f_T - Ef_T)^2\right) = O\left(\frac{(\ln T)^{2/s}}{T}\right).$$

On the other hand, it is easy to prove using A that

$$E(\varphi_T - E\varphi_T)^2 = O(T^{-1}),$$

hence (5.2). □

The proof of Proposition 5.2 is similar to the proof of Proposition 4.3 and therefore omitted.

PROOF OF PROPOSITION 6.1. It suffices to notice that

$$|\varphi_{T-l}(\xi_T) - \varphi(\xi_T)|\mathbb{1}_{\xi_T \in \Delta} \leq \sup_{x \in \Delta} |\varphi_{T-l}(x) - \varphi(x)|$$

and that

$$|f_T(\xi_T) - f(\xi_T)|\mathbb{1}_{\xi_T \in \Delta} \le \sup_{x\in\Delta}|f_T(x) - f(x)|$$

and then apply Proposition 5.2. Details are omitted.

PROOF OF PROPOSITION 6.2.

$$\begin{aligned} D_T &= E(r_{T'}(\xi_T) - r(\xi_T))^2\mathbb{1}_{\xi_T\in\Delta} \\ &= \int_\Delta E(r_{T'}(x) - r(x))^2 f(x)\,dx, \end{aligned} \tag{8.19}$$

where f is the density of ξ_T.

By a Cai–Roussas lemma ([9]) for φ_{rev}-mixing processes, it follows that for P_{ξ_0}-almost every x in Δ,

$$\begin{aligned} &E(r_{T'}(x) - r(x))^2 \mid \zeta_T = x) \\ &\quad\le E(r_{T'}(x) - r(x))^2 + 8\sup_{x\in\Delta}|M(x)|\varphi_{\text{rev}}(T - T'). \end{aligned}$$

Using this bound in (8.19), we get

$$D + \le \sup_{x\in\Delta} E(r_{T'}(x) - r(x))^2 + 8\sup_{x\in\Delta}|M(x)|a\rho^{(T-T')}.$$

Then B' entails

$$\begin{aligned} D_T &= O\left(\frac{1}{T'}\right) + O(\rho^{(T-T')}) \\ &= O\left(\frac{1}{T}\right) \end{aligned}$$

which proves (6.7). □

PROOF OF PROPOSITION 7.1. (i) First it is easy to prove that

$$\lim_{n\to\infty}\|H_n - G_n\|_\infty = 0, \tag{8.20}$$

where $H_n(y, z) = \sum_{i=1}^{\infty}\delta_n g_{i\delta_n}(y, z)$ and

$$G_n(y, z) = \sum_{i=1}^{n-1}\left(1 - \frac{i}{n}\right)\delta_n g_{i\delta_n}(y, z).$$

Now the variance of f_n^* has the classical decomposition

$$Vf_n^*(x) = \tilde{V}_n + C_n,$$

where $\tilde{V}_n$ is the sum of variances and where

$$C_n = \frac{2}{n\delta_n}\int K_{T_n}(x - y)K_{T_n}(x - z)G_n(y, z)\,dy\,dz.$$

For $\tilde{V}_n$ we have the well known

$$\tilde{V}_n = O\left(\frac{1}{nT_n^{-d/4}}\right).$$

Concerning C_n, note that

$$\left| \frac{n\delta_n}{2} C_n - \int K_{T_n}(x-y) K_{T_n}(x-z) H_n(y,z)\, dy\, dz \right| \le \|H_n - G_n\|_\infty$$

and

$$\left| \int K_{T_n}(x-y) K_{T_n}(x-z) [H_n(y,z) - G(y,z)]\, dy\, dz \right| \le \|H_n - G\|_\infty,$$

where $G(y,z) = \int_0^{+\infty} g_u(y,z)\, du$.

By (8.20) it follows that

$$\left| \frac{n\delta_n}{2} C_n - \int K_{T_n}(x-y) K_{T_n}(x-z) G(y,z)\, dy\, dz \right| \to 0.$$

Since G is continuous at (x,x) we find

$$n\delta_n C_n \to 2\int_0^{+\infty} g_u(x,x)\, du.$$

Finally, we have

$$E(f_n^*(x) - f(x))^2 = O\left(\frac{1}{nT_n^{-d/4}} + \frac{1}{n\delta_n} + \frac{1}{T_n} \right) \tag{8.21}$$

since the bias is as usual $O(T_n^{-1/2})$.

Now since $\delta_n = T_n^{-d/4}$ and $nT_n^{-d/4} = T_n$ the result follows.

(ii) The proof of (7.5) is similar and therefore omitted.

The proof of Proposition 7.2 uses Lemma 3.2 essentially as in Proposition 4.3 and 5.1. Details are omitted.

Acknowledgment. I would like to thank the referees for many valuable comments.

REFERENCES

[1] BANON, G. and NGUYEN, H. T. (1978). Sur l'estimation récurrente de la densité et de sa dérivée pour un processus de Markov. *C.R. Acad. Sci. Paris Sér. A* **286** 691–694.

[2] BLANKE, D. and BOSQ, D. (1997). Accurate rates of density estimators for continuous time processes. *Statist. Probab. Lett.* To appear.

[3] BOSQ, D. (1993). Bernstein-type large deviation inequalities for partial sums of strong mixing processes. *Statistics* **24** 59–70.

[4] BOSQ, D. (1995). Optimal asymptotic quadratic error of density estimators for strong mixing or chaotic data. *Statist. Probab. Lett.* **22** 339–347.

[5] BOSQ, D. (1993). Vitesses optimales et superoptimales des estimateurs fonctionnels pour un processus à temps continu. *C.R. Acad. Sci. Paris Sér. I Math.* **317** 1075–1078.

[6] BOSQ, D. (1995). Sur le comportement exotique de l'estimateur à noyau de la densité marginale d'un processus à temps continu. *C.R. Acad. Sci. Paris Sér. I Math.* **320** 369–372.

[7] BOSQ, D. (1996). *Nonparametric Statistics for Stochastic Processes. Lecture Notes in Statist.* **110** 1–169. Springer, Berlin.

[8] BRADLEY, R. (1983). Approximation theorems for strongly mixing random variables. *Michigan Math. J.* **30** 69–81.

[9] CAI, Z. and ROUSSAS, G. (1992). Uniform strong estimation under α-mixing, with rates. *Statist. Probab. Lett.* **15** 47–55.

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

[11] CHEZE-PAYAUD, N. (1994). Nonparametric regression and prediction for continuous time process. *Publ. Inst. Statistics Univ. Paris* **38** 37–58.

[12] DELECROIX, M. (1980). Sur l'estimation des densités d'un processus stationaire à temps continu. *Publ. Inst. Statist. Univ. Paris* **25** 17–39.

[13] DOUKHAN, P. (1994). *Mixing-Properties and Examples. Lecture Notes in Statist.* Springer, Berlin.

[14] KUTOYANTS, Y. (1996). Efficient density estimation for ergodic diffusion. Research memorandum 607, Inst. Statistical Mathematics, Tokyo.

[15] LEBLANC, F. (1995). Ph.D. dissertation, Univ. Paris 6.

[16] LIPTSER, R. S. and SHIRYAYEV, A. N. (1978). *Statistics of Random Processes* I, II. Springer, New York.

[17] PARZEN, E. (1962). On the estimation of probability density function and mode. *Ann. Math. Statist.* **33** 1065–1076.

[18] PRAKHASA-RAO, B. L. S. (1982). *Nonparametric Functional Estimation.* Wiley, New York.

[19] RHOMARI, N. (1994). Ph.D. dissertation, Univ. Paris 6.

[20] ROSENBLATT, M. (1956). Remarks on some nonparametric estimates of a density function. *Ann. Math. Statist.* **27** 832–837.

[21] ROUSSAS, G. (1988). Nonparametric estimation in mixing sequences of random variables. *J. Statist. Plann. Inference* **18** 135–149.

[22] ROUSSAS, G. (1990). Nonparametric regression estimation under mixing conditions. *Stochastic Process. Appl.* **36** 107–116.

[23] TRAN, L. T. (1990). Kernel density and regression estimation for dependent random variables and time series. Technical report, Univ. Indiana.

[24] TRAN, L. T. (1993). Nonparametric function estimation for time series by local average estimations. *Ann. Statist.* **21** 1040–1057.

[25] TRUONG, Y. K. and STONE, C. J. (1992). Nonparametric function estimation involving time series. *Ann. Statist.* **20** 77–98.

[26] VERETENNIKOV, A. YU. (1990). On hypoellipticity conditions and estimates of the mixing rate for stochastic differential equations. *Soviet Math. Dokl.* **40** 94–97.

LABORATOIRE DE STATISTIQUE THÉORIQUE ET APPLIQUÉE
UNIVERSITÉ PIERRE ET MARIE CURIE—UFR 920
4, PLACE JUSSIEU
75252 PARIS CEDEX 05
FRANCE