ISSN 0032-9460, Problems of Information Transmission, 2010, Vol. 46, No. 1, pp. 22–37. 

METHODS OF SIGNAL PROCESSING

# Wide-Sense Nonparametric Semirecursive Identification of Strong Mixing Processes[1]

**A. V. Kitaeva**[a] **and G. M. Koshkin**[b]

[a] *Tomsk Polytechnical University*
olz@mail.tomsknet.ru, kit1157@yandex.ru
[b] *Tomsk State University*
*Department of Informatization Problems, Tomsk Scientific Center, Siberian Branch of the RAS, Tomsk*
kgm@mail.tsu.ru

Received January 21, 2009; in final form, November 13, 2009

**Abstract**—We find principal parts of asymptotic mean-square errors of semirecursive nonparametric estimators of functionals of a multidimensional density function under the assumption that observations satisfy a strong mixing condition. Results are illustrated by an example of a nonlinear autoregression process.

**DOI:** 10.1134/S0032946010010047

## 1. INTRODUCTION

Let $(X_k)_{k\ge 0}$ be a sequence of random variables defined on a probability space $(\Omega, \mathrm{F}, \mathbf{P})$, and let a $\sigma$-algebra $\mathrm{F}_{i,j} = \sigma\{X_k,\ i \le k \le j\}$ be generated by $(X_i, \ldots, X_j)$.

**Definition 1.** A strictly stationary sequence $(X_k)_{k\ge 0}$ satisfies the *strong mixing condition* (notation: $(X_k)_{k\ge 0} \in \mathcal{S}(\alpha)$) if

$$\alpha(\tau) = \sup_{\substack{A\in \mathrm{F}_{0,\ell}\\ B\in \mathrm{F}_{\ell+\tau,\infty}}} |\mathbf{P}(AB) - \mathbf{P}(A)\,\mathbf{P}(B)| \downarrow 0, \quad \tau \to \infty, \quad \tau > 0. \tag{1}$$

The parameter $\alpha(\tau)$ is referred to as the *strong mixing coefficient*.

Time series satisfying condition (1) are often used in modeling economic, financial, physical, and technical processes [1–3].

The wide-sense identification problem for stochastic systems [4,5] is often reduced to estimating, based on observed sequences of output and input variables, functions of the form

$$H(A) = H\big(\{a_i(x)\}, \{a_i^{(1j)}(x)\},\ i = \overline{1,s},\ j = \overline{1,m}\big) = H\big(a(x), a^{(1j)}(x)\big), \tag{2}$$

where $x \in \mathbb{R}^m$, $H(\cdot)\colon \mathbb{R}^{(m+1)s} \to \mathbb{R}^1$ is a given function, $a^{(1j)}(x) = \big(a_1^{(1j)}(x), \ldots, a_s^{(1j)}(x)\big)$, and $a(x) \equiv a^{(0j)}(x) = (a_1(x), \ldots, a_s(x))$. Functionals $a_i(x)$ and their derivatives are defined as follows:

$$a_i(x) = \int g_i(y) f(x,y)\,dy, \quad a_i^{(1j)}(x) = \frac{\partial a_i(x)}{\partial x_j}, \quad i = \overline{1,s}, \quad j = \overline{1,m}, \tag{3}$$

where $g_1, \ldots, g_s$ are known functions, $f(x,y)$ is an unknown density function of the observed random vector $Z = (X,Y) \in \mathbb{R}^{m+1}$, $X = (X_1, \ldots, X_m)$ are input variables, and $Y$ is an output variable. Integration in (3) is over the whole number axis, and in what follows we assume that $\int \equiv \int\limits_{\mathbb{R}^1}$.

[1] Supported in part by the Russian Foundation for Basic Research, project no. 09-08-00595a.

Since for $g_s(y) \equiv 1$ we have $a_s(x) = \int f(x,y)\,dy = p(x)$, where $p(x)$ is the density function of $X$, in the form (2) we can represent any function of the conditional functionals

$$b_i(x) = a_i(x)/p(x) = \int g_i(y) f(y\,|\,x)\,dy \tag{4}$$

and their derivatives $b_i^{(1j)}(x) = \dfrac{\partial b_i(x)}{\partial x_j}$, $i = \overline{1, s-1}$.

In the case $H(a_1, a_2) = a_1/a_2$, $g_1(y) = y^m$, $m \geq 1$, and $g_2(y) = 1$ we obtain first conditional moments $\mu_m(x) = \int y^m f(y\,|\,x)\,dy$, in particular, for $m = 1$, the regression function

$$r(x) = \mathbf{E}(Y\,|\,X = x) = \mathbf{E}(Y\,|\,x) = \int y f(y\,|\,x)\,dy = \frac{\int y f(x,y)\,dy}{p(x)}. \tag{5}$$

Expression (2) covers also functions of conditional central moments, for instance, the conditional variance (volatility function)

$$D(x) = \mathbf{D}(Y\,|\,x) = \mu_2(x) - r^2(x),$$
$$H(a_1, a_2, a_3) = a_1/a_3 - (a_2/a_3)^2, \qquad g_1(y) = y^2, \qquad g_2(y) = y, \qquad g_3(y) = 1.$$

In stochastic systems, the influence of input variables on the output can be studied with the help of response functions. In particular, for the regression model (5), the response function with respect to the $j$th input [6] is defined as follows:

$$\begin{gathered} T_j(x) = \frac{\partial r(x)}{\partial x_j}, \\ H(a_1, a_2, a_1^{(1j)}, a_2^{(1j)}) = \frac{a_1^{(1j)}}{a_2} - \frac{a_1 a_2^{(1j)}}{a_2^2} = b_1^{(1j)}, \qquad g_1(y) = y, \qquad g_2(y) = 1. \end{gathered} \tag{6}$$

Other examples of using expressions of the form (2) can be found in [7, 8].

As recursive nonparametric estimators for the functionals $a^{(0j)}(x) \equiv a(x)$ and their derivatives $a^{(1j)}(x)$ at $x$ we take the statistics

$$\begin{aligned} a_n^{(rj)}(x) &= \frac{1}{n}\sum_{i=1}^{n} \frac{g(Y_i)}{h_i^{m+r}} \boldsymbol{K}^{(rj)}\Big(\frac{x - X_i}{h_i}\Big) \\ &= a_{n-1}^{(rj)}(x) - \frac{1}{n}\left[a_{n-1}^{(rj)}(x) - \frac{g(Y_n)}{h_n^{m+r}} \boldsymbol{K}^{(rj)}\Big(\frac{x - X_n}{h_n}\Big)\right], \end{aligned} \tag{7}$$

where $r = 0, 1$, $Z_\ell = (X_\ell, Y_\ell)$, $\ell = \overline{1, n}$, is an $(m+1)$-dimensional sample characterized by a density $f(x,y)$, $(Z_j)_{j\geq 1} \in \mathcal{S}(\alpha)$, $\boldsymbol{K}^{(0j)}(u) \equiv \boldsymbol{K}(u) = \prod_{i=1}^{m} K(u_i)$ is an $m$-dimensional product-form kernel, $\boldsymbol{K}^{(1j)}(u) = \dfrac{\partial \boldsymbol{K}(u)}{\partial u_j} = K(u_1)\ldots K(u_{j-1}) K^{(1)}(u_j) K(u_{j+1})\ldots K(u_m)$, $K^{(1)}(u_j) = \dfrac{dK(u_j)}{du_j}$, $(h_n) \downarrow 0$ is a number sequence, $g(y) = (g_1(y), \ldots, g_s(y))$, $a_n^{(rj)}(x) = (a_{1n}^{(rj)}(x), \ldots, a_{sn}^{(rj)}(x))$, and $a_n^{(0j)}(x) \equiv a_n(x)$.

Recursive procedures have a number of advantages: as a rule, they are easily computer implemented, they are memory-saving, provide a finished result at every step of the algorithm, newly obtained measurements do not lead to cumbersome re-computations; thus, real-time data processing is possible.

Recursive kernel estimators were first proposed and studied in [9, 10] for a one-dimensional density ($m = 1$, $s = 1$, $g(y) = 1$, and $H(a_1) = a_1$).

Semi-recursive kernel substitution estimators for the conditional functionals $b(x) = (b_1(x), \dots, b_{s-1}(x))$ at $x$ are of the form

$$b_n(x) = \frac{\sum\limits_{i=1}^{n} \frac{g(Y_i)}{h_i^m} \boldsymbol{K}\Big(\frac{x - X_i}{h_i}\Big)}{\sum\limits_{i=1}^{n} \frac{1}{h_i^m} \boldsymbol{K}\Big(\frac{x - X_i}{h_i}\Big)} = \frac{a_n(x)}{p_n(x)}, \quad g(y) = (g_1(y), \dots, g_{s-1}(y)). \tag{8}$$

For $g_1(y) = y$, $g_2(y) = 1$ $(s = 2)$, and $m = 1$ we obtain a semi-recursive analog of the Nadaraya–Watson kernel estimate [11, 12] for the one-dimensional regression function (3), which in the case of independent observations was considered in [13–15]. Various types of convergence for such estimators were studied in [13–19]. Estimators of this type are said to be semi-recursive [19], since only the numerator and denominator are computed recursively.

A semi-recursive (the functionals $a_i(x)$ and their derivatives are computed recursively) substitution estimator for (1) is of the form

$$H(A_n) = H\big(\{a_n^{(rj)}(x)\},\ j = \overline{1, m},\ r = 0, 1\big). \tag{9}$$

In the present paper, problems of instability of estimators (8) and, possibly, (9) (depending on the form of a function $H$) will be solved with the help of piecewise smooth approximations $\widetilde{H}(A_n, \delta_n)$ of estimators $H(A_n)$ [20]:

$$\widetilde{H}(A_n, \delta_n) = \frac{H(A_n)}{(1 + \delta_n |H(A_n)|^\tau)^\rho},$$

where $\tau > 0$, $\rho > 0$, $\rho\tau \geq 1$, and $(\delta_n) \downarrow 0$ as $n \to \infty$.

Dependence of observations makes analysis of properties of estimators much more complicated: for example, the principal part mean-square error (MSE) of the Nadaraya–Watson estimator for strongly mixed sequences was found only in 1999 [21]; in the same paper, convergence of this estimator with probability 1 was also proved. Estimators (9) in the case of independent observations were considered by the authors in [8].

In the present paper we find principal parts of the MSE and find the mean-square rate (improved by a choice of a kernel) of convergence of estimators (9) and their piecewise smooth approximation to $H(A)$. The obtained results are illustrated by a wide-sense identification problem for a nonlinear autoregression process.

## 2. ASYMPTOTIC PROPERTIES OF ESTIMATORS OF FUNCTIONALS AND THEIR DERIVATIVES

Let us introduce necessary definitions and notation.

**Definition 2.** A function $H(\cdot)\colon \mathbb{R}^s \to \mathbb{R}^1$ *belongs to the class* $\mathcal{N}_\nu(x)$ $(H(\cdot) \in \mathcal{N}_\nu(x))$ if it and all of its partial derivatives (up to the $\nu$th inclusive) are continuous at $x$. We say that $H(\cdot) \in \mathcal{N}_\nu(\mathbb{R})$ if these properties of $H(\cdot)$ are satisfied for sll $x \in \mathbb{R}^s$.

Definition 2 is related to smoothness conditions for the estimated function $H$, and the next two definitions are related to the estimation procedure.

Introduce the following notation: $\sup\limits_x = \sup\limits_{x \in \mathbb{R}^m}$, $T_j = \int u^j K(u)\, du$, $j = 1, 2, \dots$.

**Definition 3.** A function $K(\cdot)$ *belongs to the class* $\mathcal{A}^{(r)}$ *of normalized kernels*, $r = 0, 1$, if $\int |K^{(r)}(u)|\, du < \infty$ and $\int K(u)\, du = 1$. A function $K(\cdot)$ *belongs to the class* $\mathcal{A}_\nu^{(r)}$ if $K(\cdot) \in \mathcal{A}^{(r)}$ and $K(\cdot)$ satisfies the conditions $\int |u^\nu K(u)|\, du < \infty$, $T_j = 0$, $j = 1, \dots, \nu - 1$, $T_\nu \neq 0$, and $K(u) = K(-u)$.

The parameter $\nu$ in Definition 3 determines the mean-square convergence rate for estimators (9). For simplicity, we use the notation $\mathcal{A}^{(0)} = \mathcal{A}$ and $\mathcal{A}_\nu^{(0)} = \mathcal{A}_\nu$.

**Definition 4.** A sequence $(h_n)$ *belongs to the class* $\mathcal{H}(\alpha)$ if

$$\frac{1}{n}\sum_{i=1}^{n} h_i^\alpha = S_\alpha h_n^\alpha + o(h_n^\alpha), \tag{10}$$

where $\alpha$ is a real number and $S_\alpha$ is a constant independent of $n$.

Condition (10) is related to the recursive structure of estimators and is satisfied, for instance, fir $h_i = O(i^{-\gamma})$, $0 < \gamma < 1$ (this is precisely the form of the optimal diffusion parameters (12)); the constant $S_\alpha$ can be found by the Euler–Maclaurin formula [22, Section 12.6.466]. In particular, for any $p \neq -1$ we obtain

$$\sum_{j=1}^{n} j^p = \frac{n^{p+1}}{p+1} + o(n^{p+1});$$

for the case of a positive integer $p$, see this equality in [23, equation 0.121].

In Lemma 1, let $g(\cdot)\colon \mathbb{R}^1 \to \mathbb{R}^1$, i.e., let $a(x)$ be a scalar function. Introduce the following notation: $a^{s+}(x) = \int |g^s(y)| f(x,y)\,dy$, $\mathbf{b}(a_n^{(1j)}(x)) = \mathbf{E}\,a_n^{(1j)}(x) - a^{(1j)}(x)$ is the bias of the estimator $a_n^{(1j)}(x)$, and $\omega_\nu^{(rj)}(x) = \dfrac{T_\nu}{\nu!}\sum\limits_{\ell=1}^{m}\dfrac{\partial^\nu a^{(rj)}(x)}{\partial x_\ell^\nu}$, $r = 0, 1$, $j = \overline{1,m}$.

**Lemma 1** (bias convergence rate). *Assume that* $(h_n)$ *is a sequence of the class* $\mathcal{H}(\nu)$ *and that* $\sup\limits_x a^{1+}(x) < \infty$. *Assume that for* $r = 1$ *we have* $\sup\limits_x |a^{(1j)}(x)| < \infty$ *and* $\lim\limits_{|u|\to\infty} K(u) = 0$, *and for* $r = 0$ (*or* $r = 1$),

$$a^{(rj)}(\cdot) \in \mathcal{N}_\nu(\mathbb{R}), \quad \sup_x \left|\frac{\partial^\nu a^{(rj)}(x)}{\partial x_\ell \partial x_t \ldots \partial x_q}\right| < \infty, \quad \ell, t, \ldots, q = \overline{1,m}, \quad K(\cdot) \in \mathcal{A}_\nu^{(r)}.$$

*Then*, *as* $n \to \infty$, *for* $r = 0$ (*or*, *respectively*, $r = 1$) *we have*

$$\left|\mathbf{b}(a_n^{(rj)}(x)) - S_\nu \omega_\nu^{(rj)}(x) h_n^\nu\right| = o(h_n^\nu).$$

In the proof of Lemma 1, dependence of sample observations $Z_1, \ldots, Z_n$ plays no role (see [8, Lemmas 1–3]).

Let us find principal parts of covariances of the estimators $a_{tn}^{(rj)}(x)$ and $a_{pn}^{(qk)}(x)$, which we need in the computation of the MSE of the estimator $H(A_n)$ (see Theorems 2 and 3). We turn back to a vector function $g(y)$. Introduce the notation

$$\omega_{i\nu}^{(rj)}(x) = \frac{T_\nu}{\nu!}\sum_{\ell=1}^{m}\frac{\partial^\nu a_i^{(rj)}(x)}{\partial x_\ell^\nu}, \quad r, q = 0, 1, \quad t, p, i = \overline{1,s},$$

$$L^{(r,q)} = \int K^{(r)}(u) K^{(q)}(u)\,du, \qquad a_{t,p}(x) = \int g_t(y) g_p(y) f(x,y)\,dy,$$

$$a_{t,p}^{1+}(x) = \int |g_t(y) g_p(y)| f(x,y)\,dy, \qquad \mathcal{B}_{t,p}^{(r,q)} = L^{(r,q)} \left(L^{(0,0)}\right)^{m-1} a_{t,p}(x),$$

$$a_{i(i+\tau),tp}^{+}(x,y) = \int\limits_{\mathbb{R}^2} |g_t(v) g_p(q)| f_{i(i+\tau)}(x,v,y,q)\,dv\,dq,$$

where $f_{i(i+\tau)}$, $\tau \geq 1$, is a $2(m+1)$-dimensional density function of the sample variables $(Z_i, Z_{i+\tau})$. Note that $\int\limits_{\mathbb{R}^m} \boldsymbol{K}^{(1j)}(u)\,du = \int\limits_{\mathbb{R}^m} \boldsymbol{K}^{(1k)}(u)\,du \triangleq \int\limits_{\mathbb{R}^m} \boldsymbol{K}^{(1)}(u)\,du$ for any $j, k = \overline{1,m}$ by the product form of the kernel.

**Lemma 2** (covariance of the estimators $a_{tn}^{(rj)}(x)$ and $a_{pn}^{(qk)}(x)$). *Assume that an index $\theta$ takes values $t$ and $p$, index $\gamma$ takes values $r$ and $q$, and the following conditions are fulfilled*:

(1) $(Z_j)_{j\geq 1} \in \mathcal{S}(\alpha)$, *and* $\int\limits_0^\infty [\alpha(\tau)]^\lambda\, d\tau < \infty$ *for some* $\lambda \in (0, 1/2)$;

(2) $a_{t,p}(\cdot) \in \mathcal{N}_0(\mathbb{R})$, $a_\theta(\cdot) \in \mathcal{N}_0(\mathbb{R})$, *and* $a_\theta^{\frac{2}{1-2\lambda}+}(\cdot) \in \mathcal{N}_0(x)$;

(3) $\sup\limits_x a_\theta^{1+}(x) < \infty$, $\sup\limits_x a_\theta^{\frac{2}{1-2\lambda}+}(x) < \infty$, $\sup\limits_x a_{t,p}^{1+}(x) < \infty$, *and* $\sup\limits_{x,y} a_{i(i+\tau),tp}^{+}(x,y) < C$;

(4) $K(\cdot) \in \mathcal{A}^{(\gamma)}$, $\sup\limits_{u\in\mathbb{R}^1} |K^{(\gamma)}(u)| < \infty$, *and* $\sup\limits_{u\in\mathbb{R}^1} |K(u)| < \infty$ *for* $m > 1$ *and* $rq = 1$;

(5) $(h_n) \in \mathcal{H}(-m-r-q)$ $\;1/(nh_n^{m+r+q}) \downarrow 0$.

*Then, as* $n \to \infty$, *we have*

$$\left|\operatorname{cov}\left(a_{tn}^{(rj)}(x), a_{pn}^{(qk)}(x)\right) - \frac{S_{-m-r-q}}{nh_n^{m+r+q}}\mathcal{B}_{t,p}^{(r,q)}(x)\right| = o\left(\frac{1}{nh_n^{m+r+q}}\right), \tag{11}$$

*in particular, for* $t = p$,

$$\mathbf{D}\, a_{tn}^{(rj)}(x) \sim \frac{S_{-m-2r}}{nh_n^{m+2r}}\mathcal{B}_{t,t}^{(r,r)}(x). \tag{12}$$

Lemma 2 shows that principal parts of variances of the estimators $a_{tn}^{(rj)}(x)$ for independent (see[8, corollary]) and strongly mixed (equation (12)) estimators coincide. Thus, since for the MSE $\mathbf{u}^2(a_{tn}^{(rj)}(x))$ we have $\mathbf{u}^2(a_{tn}^{(rj)}(x)) = \mathbf{D}\, a_{tn}^{(rj)}(x) + \mathbf{b}^2(a_{tn}^{(rj)}(x))$, we obtain the following result.

**Theorem 1** (MSE of optimal estimators of functionals). *If the conditions of Lemma* 1, *conditions of Lemma* 2 *for* $q = r$ *and* $\theta = p = t$, *and an extra condition* $\omega_{t\nu}^{(rj)}(x) \neq 0$ *are satisfied, then, as* $n \to \infty$, *principal parts of the mean-square optimal parameters* $h_{tn}^{(rj)\mathrm{o}}$ *and of the MSE* $\mathbf{u}^2(a_{tn}^{(rj)\mathrm{o}}(x))$ *are given by*

$$h_{tn}^{(rj)\mathrm{o}} = \operatorname*{arg\,min}_{h_{tn}^{(rj)}>0} \mathbf{u}^2(a_{tn}^{(rj)}(x)) \sim \left[\frac{(m+2r)(m+\nu+2r)\mathcal{B}_{t,t}^{(r,r)}}{4n\nu(m+2\nu+2r)\left[\omega_{t\nu}^{(rj)}(x)\right]^2}\right]^{\frac{1}{m+2(\nu+r)}}, \tag{13}$$

$$\begin{aligned}
&\mathbf{u}^2\Big(a_{tn}^{(rj)}(x)\big|_{h_{tn}^{(rj)}=h_{tn}^{(rj)\mathrm{o}}}\Big) = \mathbf{u}^2(a_{tn}^{(rj)\mathrm{o}}(x))\\
&\sim (m+1+2r)\left[\frac{m+2(\nu+r)}{m+\nu+2r}\right]^{\frac{2(m+\nu+2r)}{m+2(\nu+r)}}\left[\frac{\mathcal{B}_{t,t}^{(r,r)}}{4n\nu}\right]^{\frac{2\nu}{m+2(\nu+r)}}\left[\frac{\left[\omega_{t\nu}^{(rj)}(x)\right]^2}{m+2r}\right]^{\frac{m+2r}{m+2(\nu+r)}}\\
&= O\big(n^{-\frac{2\nu}{m+2(\nu+r)}}\big).
\end{aligned} \tag{14}$$

The proof of Theorem 1 for strongly mixed sequences completely follows the proof of Theorem 1 for independent sequences from [8]. According to Theorem 1, the convergence rate of optimal nonparametric estimators $a_{in}^{(rj)\mathrm{o}}(x)$ for strongly mixed observations, which equals $n^{-\frac{2\nu}{m+2(\nu+r)}}$, for large $\nu$ approaches the usual convergence rate of parametric estimators, equal to $n^{-1}$.

Introduce the following notation: $f_{1(i+1)(i+j+1)(i+j+k+1)}$ is the density function of the sample variables $(Z_1, Z_{i+1}, Z_{i+j+1}, Z_{i+j+k+1})$;

$$\begin{aligned}
&a_{1(i+1)(i+j+1)(i+j+k+1),t}^{+}(x,y,x',y')\\
&\quad= \int\limits_{\mathbb{R}^4} |g_t(v)g_t(u)g_t(v')g_t(u')| f_{1(i+1)(i+j+1)(i+j+k+1)}(x,v,y,u,x',v',y',u')\, dv\, du\, dv'\, du',
\end{aligned}$$

$$1 \leq i,j,k < n, \quad i+j+k \leq n-1;$$

$$a_{1(1+j)(1+j+k),t}^{(2+\delta)+}(x,y,x') = \int\limits_{\mathbb{R}^3} |g_t(v)g_t(u)g_t(v')|^{2+\delta} f_{1(1+j)(1+j+k)}(x,v,y,u,x',v')\, dv\, du\, dv',$$

$$a^{(2+\delta)+}_{1(i+1),t}(x,x') = \int\limits_{\mathbb{R}^2} |g_t(v)g_t(u)|^{2+\delta} f_{1(1+i)}(x,v,u,x')\,dv\,du,$$

$$M_4(a^{(rj)}_{tn}) = \mathbf{E}\big[a^{(rj)}_{tn}(x) - a^{(rj)}_t(x)\big]^4, \qquad S^{(rj)}_{tn} = a^{(rj)}_{tn}(x) - \mathbf{E}\,a^{(rj)}_{tn}(x).$$

Let us find the order of convergence of $M_4(a^{(rj)}_{tn})$ to zero (see Theorem 3, condition (3)). To this end, we first formulate the following lemma.

**Lemma 3** (convergence order of the fourth central moments of the estimators $a^{(rj)}_{tn}(x)$). *Let for $r = 0$ (or $r = 1$) the following conditions be fulfilled*:

(1) $(Z_j)_{j\ge1} \in \mathcal{S}(\alpha)$, *and* $\int\limits_0^\infty \tau^2[\alpha(\tau)]^{\frac{\delta}{2+\delta}}\,d\tau < \infty$ *for some* $0 < \delta < 2$;
(2) $\sup\limits_{u\in\mathbb{R}^1} |K^{(r)}(u)| < \infty$, $\int |K^{(r)}(u)|\,du < \infty$;
(3) $(h_n)$ *is monotone nonincreasing and* $1/(nh_n^{m+2r}) \downarrow 0$;
(4) $\sup\limits_x a^{\beta+}_t(x) < \infty$, $\beta = 0, 4$;
(5) $\sup\limits_x a^{+}_{1(i+1)(i+j+1)(i+j+k+1),t}(x,x,x,x) < \infty$, $\sup\limits_x a^{(2+\delta)+}_{1(1+j)(1+j+k),t}(x,x,x) < \infty$, *and* $\sup\limits_x a^{(2+\delta)+}_{1(i+1),t}(x,x) < \infty$.

*Then, as $n \to \infty$, for $r = 0$ (or, respectively, $r = 1$) we have*

$$\mathbf{E}(S^{(rj)}_{tn})^4 = O\Big(\frac{1}{n^2 h_n^{2(m+2r)}}\Big). \tag{15}$$

**Lemma 4** (convergence order of the fourth moments of the deviations $M_4(a^{(rj)}_{tn})$). *If the conditions of Lemmas* 1 *and* 3 *are fulfilled for $r = 0$ (or $r = 1$), then, as $n \to \infty$, we have*

$$M_4(a^{(rj)}_{tn}) = O\Big(\frac{1}{n^2 h_n^{2(m+2r)}} + h_n^{4\nu}\Big). \tag{16}$$

Proofs of Lemmas 2 and 3 are given in the Appendix, and Lemma 4 follows from Lemma 3 (see the proof of [8, Lemma 6]).

## 3. MSE OF SUBSTITUTION ESTIMATORS AND OF THEIR PIECEWISE SMOOTH APPROXIMATIONS

To find the MSE, we use results of [20], where a function $H(t) = H(t_1,\ldots,t_s)$ is considered, $t = t(x) = (t_1(x),\ldots,t_s(x))$ being a bounded vector function, $x \in \mathbb{R}^\alpha$. Under certain conditions, there are found asymptotics of moments of deviations of $H(t_n)$ from $H(t)$, where $t_n = t_n(x) = (t_{1n},\ldots,t_{sn}) = (t_{1n}(x),\ldots,t_{sn}(x))$ is a vector statistic constructed from a sample $X_1,\ldots,X_n$, $X_i \in \mathbb{R}^\alpha$, of not necessarily independent but identically distributed random variables.

Denote $H_j(t) = \partial H(z)/\partial z_j\big|_{z=t}$, $j = \overline{1,s}$; let $\|t\|$ be the Euclidean norm of a vector $t$.

Following [20], we formulate a theorem which allows us to find the MSE of estimator (9).

**Theorem 2** ([20, Theorem 1] with $k = 2$ and $m = 4$). *Assume that*

(1) $H(z) \in \mathcal{N}_2(t)$;
(2) *For any values of the variables $X_1,\ldots,X_n$, the sequence $\{|H(t_n)|\}$ is majorized by a number sequence $(C_0 d_n^\gamma)$, $(d_n) \uparrow \infty$, $C_0$ being a constant,* $0 \le \gamma \le 1/4$;
(3) $\mathbf{E}\,\|t_n - t\|^4 = O(d_n^{-2})$.

*Then, as* $n \to \infty$, *we have*

$$\mathbf{u}^2(H(t_n)) = \sum_{j,p=1}^{s} H_j H_p[\operatorname{cov}(t_{jn}, t_{pn}) + \mathbf{b}(t_{jn})\mathbf{b}(t_{pn})] + O(d_n^{-3/2}).$$

Introduce the following notation ($r = 0, 1$, $j = \overline{1, m}$, $i = \overline{1, s}$):

$$H_{ijr} = \partial H(A)/\partial a_i^{(rj)}, \qquad Q = \begin{cases} \{0\} & \text{if } \forall j \;\; r = 0, \\ \{1\} & \text{if } \forall j \;\; r = 1, \\ \{0, 1\} & \text{if } \exists j \;\; r = 0 \wedge r = 1, \end{cases}$$

$$\max(r) = \max_{r \in Q}(r), \qquad d_n = \left[\frac{1}{nh_n^{m+2\max(r)}} + h_n^{2\nu}\right]^{-1}.$$

Theorem 2, taking into account Lemmas 1, 2, and 4, implies the following result.

**Theorem 3** (MSE of the estimator $H(A_n)$). *Assume that for* $t, p = \overline{1, s}$, $j = \overline{1, m}$, *and* $r \in Q$ *we have*

(1) $(Z_i)_{i \geq 1} \in \mathcal{S}(\alpha)$, *and* $\int\limits_0^\infty \tau^2[\alpha(\tau)]^{\frac{\delta}{2+\delta}}\, d\tau < \infty$ *for some* $0 < \delta < 2$;

(2) $a_{t,p}(\cdot) \in \mathcal{N}_0(\mathbb{R})$, $a_t^{(2+\delta)+}(\cdot) \in \mathcal{N}_0(x)$, $\sup\limits_x a_{t,p}^{+}(x) < \infty$, *and* $\sup\limits_x a_t^{\beta+}(x) < \infty$, $\beta = 0, 4$;

(3) $K(\cdot) \in \mathcal{A}_\nu^{(r)}$; $\sup\limits_{u \in \mathbb{R}^1} |K^{(r)}(u)| < \infty$; *if* $1 \in Q$, *then* $\lim\limits_{|u| \to \infty} K(u) = 0$; *for* $m > 1$, $\sup\limits_{u \in \mathbb{R}^1} |K(u)| < \infty$;

(4) $a_t^{(rj)}(\cdot) \in \mathcal{N}_\nu(\mathbb{R})$, $\sup\limits_x \bigl|a_t^{(rj)}(x)\bigr| < \infty$, *and* $\sup\limits_x \Bigl|\dfrac{\partial^\nu a_t^{(rj)}(x)}{\partial x_\ell \ldots \partial x_q}\Bigr| < \infty$, $\ell, \ldots, q = \overline{1, m}$;

(5) *A monotone nonincreasing sequence* $(h_n)$ *belongs to the class* $\mathcal{H}(\nu)$, $(h_n) \in \mathcal{H}(-m - 2k)$ *for integers* $0 \leq k \leq \max(r)$, *and if* $Q = \{1\}$, *then* $k = 1$;

(6) $\sup\limits_x a_{1(i+1)(i+j+1)(i+j+k+1),t}(x, x, x, x) < \infty$, $\sup\limits_x a^{(2+\delta)+}_{1(i+1)(i+j+1),t}(x, x, x) < \infty$, $\sup\limits_x a^{(2+\delta)+}_{1(i+1),t}(x, x) < \infty$, *and* $\sup\limits_{x,y} a^{+}_{1(i+1),tp}(x, y) < \infty$, *for any* $i, j, k \geq 1$;

(7) $H(\cdot) \in \mathcal{N}_2(A)$;

(8) *For any values of the sample* $Z_1, \ldots, Z_n$, *the sequence* $\{|H(A_n)|\}$ *is majorized by a number sequence* $(C_0 d_n^\gamma)$, $(d_n) \uparrow \infty$, $C_0$ *being a constant*, $0 \leq \gamma \leq 1/4$.

*Then*

$$\mathbf{u}^2(H(A_n)) = \sum_{t,p,r,q,j,k} H_{tjr} H_{pkq}\left[S_{-(m+r+q)} \frac{\mathcal{B}_{t,p}^{(r,q)}(x)}{nh_n^{m+r+q}} + S_\nu^2 \omega_{t\nu}^{(rj)}(x)\omega_{p\nu}^{(qk)}(x) h_n^{2\nu}\right] + O(d_n^{-3/2}). \quad (17)$$

Note that validity of the condition

$$\int\limits_0^\infty \tau^2[\alpha(\tau)]^{\frac{\delta}{2+\delta}}\, d\tau < \infty, \quad 0 < \delta < 2,$$

implies the validity of the condition of Lemma 2

$$\int\limits_0^\infty [\alpha(\tau)]^\lambda\, d\tau < \infty, \quad 0 < \lambda < 1/2,$$

and the condition $\sup\limits_x a_t^{\beta+}(x) < \infty$, $\beta = 0, 4$, implies $\sup\limits_x a_t^{(2+\delta)+}(x) < \infty$. Conditions (1)–(6) of Theorem 3 ensure the validity of condition (3) of Theorem 2.

**Theorem 4** (MSE of the estimator $\widetilde{H}(A_n,\delta_n)$). *Assume that the conditions of Theorem 3 are fulfilled, with condition* (8) *replaced by the condition*

(8*) $H(A)\neq 0$ *or* $\tau=4,6,\dots$.

*Then, as* $n\to\infty$, *we have*

$$\mathbf{u}^2(\widetilde{H}(A_n,\delta_n))\sim\mathbf{u}^2(H(A_n)).$$

Optimal sequences of the diffusion parameter and the corresponding MSE of the estimators $H(A_n)$ and their piecewise smooth approximations $\widetilde{H}(A_n,\delta_n)$ for strongly mixed observations can be found in the same way as in Theorem 1 and are not presented here because of awkwardness of expressions.

Theorem 4 follows from [20, Corollary 4] with $k=2$ and $m=4$.

## 4. IDENTIFICATION OF A NONLINEAR AUTOREGRESSION

Let a sequence $(X_t)_{t=m,m+1,\dots}$ be generated by a nonlinear autoregression of order $m$

$$X_t=\Psi(X_{t-1},\dots,X_{t-m})+\Phi(X_{t-1},\dots,X_{t-m})\xi_t, \tag{18}$$

where $\xi_t$ is a sequence of zero-mean i.i.d. random variables with unit variance, and $\Psi$ and $\Phi>0$ are unknown functions defined on $\mathbb{R}^m$.

Denote $U_{t-1}=(X_{t-1},\dots,X_{t-m})$. Note that for $x\in\mathbb{R}^m$ we have the conditional expectation $\mathbf{E}(X_t\,|\,U_{t-1}=x)=\Psi(x)$ and conditional variance $\mathbf{D}(X_t\,|\,U_{t-1}=x)=\Phi^2(x)$. Thus, the function $\Psi(x)$ can be interpreted as a prognostic value of the process (18), and $\Phi(x)$ reflects the prognosis risk.

We presume that assumptions 3.1 and 3.2 from [24, p. 263] are fulfilled (here we have a sequence of indices $(i_1,i_2,\dots,i_{q+1})=(0,1,\dots,m)$, $g_1(\cdot)\triangleq\Psi(\cdot)$, $g_2(\cdot)\triangleq\Phi(\cdot)$, $Y\triangleq X$, and $e\triangleq\xi$). In this case, according to [24, Lemma 3.1], $(X_t)$ is a strictly stationary process satisfying the strong mixing condition with a strong mixing coefficient

$$\alpha(\tau)\sim e^{-\delta\tau},\quad \delta>0. \tag{19}$$

Sufficient conditions of geometric ergodicity of a nonlinear autoregression process, implying strong mixing with geometrically decaying coefficients, were considered, e.g., in [25–27].

The strong mixing coefficients (19), as well as those geometrically decaying, satisfy the conditions of Theorems 3 and 4.

Let $f(x,y)$ be a stationary density function of the vector $(U_{t-1},X_t)$. In the framework of model (4), we have $\Psi(x)=H(a(x))=a_1(x)/a_2(x)$, $a(x)=(a_1(x),a_2(x))$, where $a_1(x)=\int yf(x,y)\,dy$ and $a_2(x)=p(x)$.

We estimate $\Psi(x)$ by the statistic

$$\Psi_n(x)=\sum_{t=m+1}^{n+m}\frac{X_t}{h_t^m}\boldsymbol{K}\Big(\frac{x-U_{t-1}}{h_t}\Big)\Big/\sum_{t=m+1}^{n+m}\frac{1}{h_t^m}\boldsymbol{K}\Big(\frac{x-U_{t-1}}{h_t}\Big)=\frac{a_{1n}(x)}{p_n(x)}. \tag{20}$$

A piecewise smooth approximation for (20) is of the form

$$\widetilde{\Psi}_n(x)=\frac{\Psi_n(x)}{(1+\delta_n|\Psi_n(x)|^\tau)^\rho}, \tag{21}$$

where by Theorems 2 and 3 we have $\delta_n=O\Big(\dfrac{1}{nh_n^m}+h_n^{2\nu}\Big)$, $(\delta_n)\downarrow 0$ as $n\to\infty$.

We estimate the conditional variance for model (18) by a statistic similar to (20):

$$D_n(x) = \sum_{t=m+1}^{n+m} \frac{X_t^2}{h_t^m} \boldsymbol{K}\Big(\frac{x-U_{t-1}}{h_t}\Big) \Big/ \sum_{t=m+1}^{n+m} \frac{1}{h_t^m} \boldsymbol{K}\Big(\frac{x-U_{t-1}}{h_t}\Big) - \Psi_n^2(x), \tag{22}$$

and estimate the response functions, which show the degree of relationship between variations of the input, $x_{t-j}$, and output, $x_t$, variables of model (18), $j = 1, \dots, m$, by the statistics

$$T_{(t-j)n}(x) = \left[ \frac{\sum\limits_{t=m+1}^{n+m} \frac{X_t}{h_t^{m+1}} \boldsymbol{K}^{(1j)}\Big(\frac{x-U_{t-1}}{h_t}\Big)}{\sum\limits_{t=m+1}^{n+m} \frac{1}{h_t^m} \boldsymbol{K}\Big(\frac{x-U_{t-1}}{h_t}\Big)} - \frac{\sum\limits_{t=m+1}^{n+m} \frac{X_t}{h_t^m} \boldsymbol{K}\Big(\frac{x-U_{t-1}}{h_t}\Big) \sum\limits_{t=m+1}^{n+m} \frac{1}{h_t^{m+1}} \boldsymbol{K}^{(1j)}\Big(\frac{x-U_{t-1}}{h_t}\Big)}{\Big[\sum\limits_{t=m+1}^{n+m} \frac{1}{h_t^m} \boldsymbol{K}\Big(\frac{x-U_{t-1}}{h_t}\Big)\Big]^2} \right]. \tag{23}$$

Estimator (22) is a semi-recursive analog of kernel estimators (3.6) and (6) of the volatility function considered, respectively, in [3] and [28].

To find the MSE of the estimator $\Psi_n(x)$, we use Theorem 3. In formula (20), let the kernel be $K(\cdot) \in \mathcal{A}_\nu$, $\sup\limits_{u\in\mathbb{R}^1} |K(u)| < \infty$, and $(h_n) \in \mathcal{H}(\nu) \cap \mathcal{H}(-m)$. Assume that the functions $a_i(x)$, $i = 1, 2$, and their derivatives of orders up to $\nu$ inclusive are continuous and bounded on $\mathbb{R}^m$, the functions $\int y^2 f(x, y)\, dy$ and $\int y^4 f(x, y)\, dy$ are bounded on $\mathbb{R}^m$, and, moreover, $\int y^2 f(x, y)\, dy$ and $\int |y|^{2+\delta} f(x, y)\, dy$ are continuous at $x$. Then conditions (1)–(5) of Theorem 3 are fulfilled; we also assume that condition (6) holds. If $p(x) > 0$, then condition (7) is also fulfilled.

If the random variable $X_t$ is bounded and a nonnegative kernel is chosen, then it is easy to show that $\Psi_n(x)$ is bounded for $\nu = 2$, which is equivalent to existence of a majorizing sequence with $\gamma = 0$, and hence condition (8) of Theorem 3 holds. As a result, we obtain

$$\mathbf{u}^2(\Psi_n(x)) = \sum_{i,p=1}^{2} H_i H_p \bigg( S_{-m} \frac{\mathcal{B}_{i,p}^{(0,0)}(x)}{n h_n^m} + S_2^2 \omega_{i2}^{(0)}(x) \omega_{p2}^{(0)}(x) h_n^4 \bigg) + O\bigg( \Big[ \frac{1}{n h_n^m} + h_n^4 \Big]^{3/2} \bigg)$$

as $n \to \infty$, where

$$H_1 = \frac{1}{p(x)}, \qquad H_2 = -\frac{\Psi(x)}{p(x)}, \qquad \mathcal{B}_{i,p}^{(0,0)}(x) = (L^{(0,0)})^m a_{i,p}(x), \qquad L^{(0,0)} = \int K^2(u)\, du,$$

$$a_{1,1}(x) = \int y^2 f(x, y)\, dy, \qquad a_{1,2}(x) = a_{2,1}(x) = \int y f(x, y)\, dy, \qquad a_{2,2}(x) = p(x),$$

$$\omega_{12}^{(0)}(x) = \frac{T_2}{2} \sum_{\ell=1}^{m} \frac{\partial^2 a_1(x)}{\partial x_\ell^2}, \qquad \omega_{22}^{(0)}(x) = \frac{T_2}{2} \sum_{\ell=1}^{m} \frac{\partial^2 p(x)}{\partial x_\ell^2}.$$

For $\nu > 2$, computing the MSE by methods presented in [11, 29] is impossible. The possibility of vanishing of the denominator of (20) makes it difficult to find a majorizing sequence $(d_n)$ in condition (8) of Theorem 3.

In this case, the problem can be solved by using the piecewise smooth approximation (21), for which, according to Theorem 4, condition (8*) is fulfilled if we take an even $\tau \ge 4$. Then, as $n \to \infty$, we have

$$\mathbf{u}^2(\widetilde{\Psi}_n(x)) = \sum_{i,p=1}^{2} H_i H_p \bigg( S_{-m} \frac{\mathcal{B}_{i,p}^{(0,0)}(x)}{n h_n^m} + S_\nu^2 \omega_{i\nu}^{(0)}(x) \omega_{p\nu}^{(0)}(x) h_n^{2\nu} \bigg) + O\bigg( \Big[ \frac{1}{n h_n^m} + h_n^{2\nu} \Big]^{3/2} \bigg),$$

where

$$\omega_{1\nu}^{(0)}(x) = \frac{T_\nu}{\nu!}\sum_{\ell=1}^{m}\frac{\partial^\nu a_1(x)}{\partial x_\ell^\nu}, \qquad \omega_{2\nu}^{(0)}(x) = \frac{T_\nu}{\nu!}\sum_{\ell=1}^{m}\frac{\partial^\nu p(x)}{\partial x_\ell^\nu}.$$

For the response function $T_{t-j}(x) = \dfrac{\partial \Psi(x)}{\partial x_{t-j}}$, difficulties related to finding a majorizing sequence force one to use a piecewise smooth approximation

$$\widetilde{T}_{(t-j)n}(x) = \frac{T_{(t-j)n}(x)}{(1+\delta_n|T_{(t-j)n}(x)|^\tau)^\rho},$$

where $T_{(t-j)n}(x)$ is given by formula (23), $\delta_n = O(h_n^{2\nu} + 1/(nh_n^{m+2}))$.

Here $K(u)$ must additionally satisfy the conditions $\sup\limits_{u\in\mathbb{R}^1}|K^{(1)}(u)| < \infty$ and $\lim\limits_{|u|\to\infty} K(u)$, and $(h_n) \in \mathcal{H}(\nu)\cap\mathcal{H}(-m-2)$. To use the result of Theorem 4 and find $\mathbf{u}^2(\widetilde{T}_{(t-j)n}(x))$, we also have to require that the functions $a_1(x)$ and $a_2(x)$ have continuous and bounded on $\mathbb{R}^m$ derivatives of order $\nu+1$.

*APPENDIX*

Let $\|X\|_p = (\mathbf{E}\,|X|^p)^{\frac{1}{p}}$. We need an auxiliary statement.

**Proposition** [30]. *If random variables $X$ and $Y$ are measurable with respect to the $\sigma$-algebras $\mathrm{F}_{0,t}$ and $\mathrm{F}_{t+\tau,\infty}$, $\tau > 0$, respectively, and satisfy the strong mixing condition* (1), $1 \le p,q,r \le \infty$, $p^{-1}+q^{-1}+r^{-1} = 1$, *then*

$$|\mathbf{E}(XY) - \mathbf{E}\,X\,\mathbf{E}\,Y| \le 12\alpha^{\frac{1}{r}}(\tau)\|X\|_p\|Y\|_q.$$

**Proof of Lemma 2.** Denote

$$\xi_{ti}^{(rj)}(x) = \frac{1}{h_i^{m+r}} g_t(Y_i)\boldsymbol{K}^{(rj)}\Big(\frac{x-X_i}{h_i}\Big).$$

We represent the covariance in the form

$$\begin{aligned}
\operatorname{cov}\big(a_{tn}^{(rj)}(x), a_{pn}^{(qk)}(x)\big) &= \frac{1}{n^2}\sum_{i=1}^{n}\sum_{\ell=1}^{n}\operatorname{cov}\big(\xi_{ti}^{(rj)}(x), \xi_{pl}^{(qk)}(x)\big)\\
&= \frac{1}{n}\sum_{i=1}^{n}\operatorname{cov}\big(\xi_{ti}^{(rj)}(x), \xi_{pi}^{(qk)}(x)\big) + \frac{2}{n^2}\sum_{\tau=1}^{n-1}\sum_{i=1}^{n-\tau}\operatorname{cov}\big(\xi_{ti}^{(rj)}(x), \xi_{p(\tau+i)}^{(qk)}(x)\big)\\
&= A_n(x) + R_n(x). \qquad (24)
\end{aligned}$$

By Lemma 4 from [8], for the term $A_n(x)$ as $n\to\infty$ we have

$$\left|A_n(x) - \frac{S_{-(m+r+q)}}{nh_n^{m+r+q}}\mathcal{B}_{t,p}^{(r,q)}(x)\right| = o\Big(\frac{1}{nh_n^{m+r+q}}\Big).$$

Denoting $U = \xi_{ti}^{(rj)}(x)$ and $V = \xi_{p(\tau+i)}^{(qk)}(x)$, let us estimate the term $R_n(x)$. Applying the auxiliary proposition with $r = \dfrac{2+\delta}{\delta}$ and $p = q = 2+\delta$, where $\delta > 0$ is arbitrary, we obtain

$$|\operatorname{cov}(U,V)| \le 12[\alpha(\tau)]^{\frac{\delta}{2+\delta}}\big[\mathbf{E}\,|U|^{2+\delta}\,\mathbf{E}\,|V|^{2+\delta}\big]^{\frac{1}{2+\delta}}. \qquad (25)$$

Since

$$\mathbf{E}\,|U|^{2+\delta} = \frac{1}{h_i^{(m+r)(2+\delta)}} \int\limits_{\mathbb{R}^{m+1}} \Big| g_t(z) \boldsymbol{K}^{(rj)}\Big(\frac{x-t}{h_i}\Big)\Big|^{2+\delta} f(t,z)\,dt\,dz,$$

we obtain, as in the case of Lemma 1 from [8],

$$\begin{aligned} \mathbf{E}\,|U|^{2+\delta} &= \frac{1}{h_i^{(m+r)(2+\delta)-m}} a_t^{(2+\delta)+}(x) \int\limits_{\mathbb{R}^m} |\boldsymbol{K}^{(r)}(z)|^{2+\delta}\,dz + o\bigg(\frac{1}{h_i^{(m+r)(2+\delta)-m}}\bigg), \\ \mathbf{E}\,|V|^{2+\delta} &= \frac{1}{h_{\tau+i}^{(m+q)(2+\delta)-m}} a_p^{(2+\delta)+}(x) \int\limits_{\mathbb{R}^m} |\boldsymbol{K}^{(q)}(z)|^{2+\delta}\,dz + o\bigg(\frac{1}{h_{\tau+i}^{(m+q)(2+\delta)-m}}\bigg). \end{aligned} \tag{26}$$

Choose $\delta = \dfrac{4\lambda}{1-2\lambda}$, $0 < \lambda < 1/2$. Then, taking into account that $\alpha(\tau) \downarrow 0$, we have

$$\sum_{\tau=1}^{\infty} [\alpha(\tau)]^{\frac{\delta}{2+\delta}} \le \int\limits_0^1 [\alpha(\tau)]^{\frac{\delta}{2+\delta}}\,d\tau + \int\limits_1^2 [\alpha(\tau)]^{\frac{\delta}{2+\delta}}\,d\tau + \ldots = \int\limits_0^{\infty} [\alpha(\tau)]^{2\lambda}\,d\tau < \infty. \tag{27}$$

Let us prove (11). Equation (24) implies

$$\begin{aligned} |R_n(x)| \le \frac{2}{n^2} \sum_{\tau=1}^{c(n)-1} \sum_{i=1}^{n-\tau} \Big|\mathrm{cov}\big(\xi_{ti}^{(rj)}(x), \xi_{p(\tau+i)}^{(qk)}(x)\big)\Big| \\ + \frac{2}{n^2} \sum_{\tau=c(n)}^{n} \sum_{i=1}^{n-\tau} \Big|\mathrm{cov}\big(\xi_{ti}^{(rj)}(x), \xi_{p(\tau+i)}^{(qk)}(x)\big)\Big| = J_1 + J_2. \end{aligned}$$

Let $c(n)$ be positive integers such that $c(n)h_n^m \to 0$ and $c(n)h_n^{2m\lambda} \to \infty$ as $n \to \infty$ (for instance, we may take $c(n) \sim h_n^{m(\varepsilon-1)}$, $0 < \varepsilon < 1 - 2\lambda$, $0 < \lambda < 1/2$). Then, as $n \to \infty$, we have

$$\begin{aligned} J_1 &\le \frac{2}{n^2} \sum_{\tau=1}^{c(n)-1} \sum_{i=1}^{n-\tau} \frac{1}{h_i^{m+r} h_{\tau+i}^{m+q}} \int\limits_{\mathbb{R}^{2(m+1)}} \Big| g_t(z) \boldsymbol{K}^{(rj)}\Big(\frac{x-u}{h_i}\Big) g_p(y) \boldsymbol{K}^{(qk)}\Big(\frac{x-v}{h_{\tau+i}}\Big)\Big| \\ &\quad \times \big| f_{i(i+\tau)}(u,z,v,y) - f(u,z) f(v,y)\big|\,du\,dz\,dv\,dy \\ &\le \frac{2}{n^2} \sum_{\tau=1}^{c(n)-1} \sum_{i=1}^{n-\tau} \frac{1}{h_i^{m+r} h_{\tau+i}^{m+q}} \Big[\sup_x a_{i(i+\tau),tp}^{+}(x,x) + \sup_x a_t^{1+}(x) \sup_x a_p^{1+}(x)\Big] \\ &\quad \times \int\limits_{\mathbb{R}^{2m}} \Big| \boldsymbol{K}^{(rj)}\Big(\frac{x-u}{h_i}\Big) \boldsymbol{K}^{(qk)}\Big(\frac{x-v}{h_{\tau+i}}\Big)\Big|\,du\,dv \\ &\le \frac{C}{n^2} \sum_{\tau=1}^{c(n)-1} \sum_{i=1}^{n-\tau} \frac{1}{h_i^r h_{\tau+i}^q} \int\limits_{\mathbb{R}^m} |\boldsymbol{K}^{(r)}(u)|du \int\limits_{\mathbb{R}^m} |\boldsymbol{K}^{(q)}(u)|du \\ &\le \frac{Cc(n)}{n} \frac{1}{h_n^{r+q}} = O\bigg(\frac{h_n^m c(n)}{n h_n^{m+r+q}}\bigg) = o\bigg(\frac{1}{n h_n^{m+r+q}}\bigg). \end{aligned}$$

Hereafter, by $C$ we denote any constant, which need not be the same even within a single reasoning.

Using relations (25)–(27) and expressing $\delta$ through $\lambda$, we obtain

$$\begin{aligned}
J_2 &= \frac{2}{n^2}\sum_{\tau=c(n)}^{n}\sum_{i=1}^{n-\tau}|\mathrm{cov}(\xi_{ti}^{(rj)}(x),\xi_{p(\tau+i)}^{(qk)}(x))| \\
&\le \frac{24}{n^2}[a_t^{(2+\delta)+}(x)a_p^{(2+\delta)+}(x)]^{\frac{1}{2+\delta}}\left[\int\limits_{\mathbb{R}^m}|\boldsymbol{K}^{(r)}(z)|^{2+\delta}\,dz\int\limits_{\mathbb{R}^m}|\boldsymbol{K}^{(q)}(z)|^{2+\delta}\,dz\right]^{\frac{1}{2+\delta}} \\
&\quad\times\sum_{\tau=c(n)}^{n}\sum_{i=1}^{n-\tau}\frac{[\alpha(\tau)]^{\frac{\delta}{2+\delta}}}{h_i^{[(m+r)(2+\delta)-m]/(2+\delta)}}\frac{1}{h_{\tau+i}^{[(m+q)(2+\delta)-m]/(2+\delta)}} \\
&\quad+o\biggl(\frac{1}{nh_n^{[(2m+r+q)(2+\delta)-2m]/(2+\delta)}}\biggr) \\
&\le \frac{24}{n^2}\Bigl[a_t^{\frac{2}{1-2\lambda}+}(x)a_p^{\frac{2}{1-2\lambda}+}(x)\Bigr]^{\frac{1-2\lambda}{2}}\left[\int\limits_{\mathbb{R}^m}|\boldsymbol{K}^{(r)}(u)|^{\frac{2}{1-2\lambda}}\,du\int\limits_{\mathbb{R}^m}|\boldsymbol{K}^{(q)}(u)|^{\frac{2}{1-2\lambda}}\,du\right]^{\frac{1-2\lambda}{2}} \\
&\quad\times\sum_{\tau=c(n)}^{n}[\alpha(\tau)]^{2\lambda}\sum_{i=\tau}^{n}\frac{1}{h_i^{m(1+2\lambda)+r+q}}+o\biggl(\frac{1}{nh_n^{m(1+2\lambda)+r+q}}\biggr) \\
&\le \frac{24\Bigl[a_t^{\frac{2}{1-2\lambda}+}(x)a_p^{\frac{2}{1-2\lambda}+}(x)\Bigr]^{\frac{1-2\lambda}{2}}}{nh_n^{m(1+2\lambda)+r+q}}\left[\int\limits_{\mathbb{R}^m}|\boldsymbol{K}^{(r)}(u)|^{\frac{2}{1-2\lambda}}\,du\int\limits_{\mathbb{R}^m}|\boldsymbol{K}^{(q)}(u)|^{\frac{2}{1-2\lambda}}\,du\right]^{\frac{1-2\lambda}{2}} \\
&\quad\times\sum_{\tau=c(n)}^{n}[\alpha(\tau)]^{2\lambda}+o\biggl(\frac{1}{nh_n^{m(1+2\lambda)+r+q}}\biggr).
\end{aligned}$$

Furthermore,

$$\begin{aligned}
J_2 &\le \frac{C}{c(n)nh_n^{m(1+2\lambda)+r+q}}\sum_{\tau=c(n)}^{\infty}c(n)[\alpha(\tau)]^{2\lambda} \\
&\le \frac{C}{c(n)nh_n^{m(1+2\lambda)+r+q}}\sum_{\tau=c(n)}^{\infty}\tau[\alpha(\tau)]^{2\lambda} \\
&\le \frac{C}{c(n)nh_n^{m(1+2\lambda)+r+q}}\sum_{\tau=1}^{\infty}\tau[\alpha(\tau)]^{2\lambda}.
\end{aligned}$$

Since $\alpha(\tau)$ is nonincreasing, i.e., $1\ge\alpha(1)\ge\alpha(2)\ge\ldots$, it follows from

$$\sum_{\tau=1}^{\infty}[\alpha(\tau)]^{\lambda}<\infty,\quad 0<\lambda<\frac{1}{2},$$

that

$$\sum_{\tau=1}^{\infty}\tau[\alpha(\tau)]^{2\lambda}=\sum_{\gamma=1}^{\infty}\sum_{t=\gamma}^{\infty}[\alpha(t)]^{2\lambda}\le\sum_{\gamma=1}^{\infty}[\alpha(\gamma)]^{\lambda}\sum_{t=\gamma}^{\infty}[\alpha(t)]^{\lambda}\le\left[\sum_{\tau=1}^{\infty}[\alpha(\tau)]^{\lambda}\right]^2<\infty.$$

Thus,

$$J_2=O\biggl(\frac{1}{c(n)nh_n^{m(1+2\lambda)+r+q}}\biggr)=O\biggl(\frac{1}{nh_n^{m+r+q}c(n)h_n^{2m\lambda}}\biggr)=o\biggl(\frac{1}{nh_n^{m+r+q}}\biggr).\quad\triangle$$

**Proof of Lemma 3.** We use methods of the proofs of Lemmas 4 and 1 from [31, Sections 4.20 and 4.22]. Denote

$$\eta_{ir} = \frac{1}{h_i^{m+r}}\left[g_t(Y_i)\boldsymbol{K}^{(rj)}\left(\frac{x-X_i}{h_i}\right) - \mathbf{E}\left[g_t(Y_i)\boldsymbol{K}^{(rj)}\left(\frac{x-X_i}{h_i}\right)\right]\right].$$

The sequence $(Z_j)_{j\ge 1}$ is stationary, and therefore

$$\mathbf{E}(S_{tn}^{(rj)})^4 = \frac{1}{n^4}\,\mathbf{E}\left(\sum_{i=1}\eta_{ir}\right)^4 \le \frac{4!}{n^4}\sum_{s,i,j,k}|\mathbf{E}\,\eta_{sr}\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r}|, \tag{28}$$

where the sum is over $s,i,j,k \ge 1$, $s+i+j+k \le n$. By the proposition with $r = \dfrac{2+\delta}{\delta}$ and $p = q = 2+\delta$, taking into account condition (2) of the lemma and the fact that $\mathbf{E}\,\eta_{sr} = 0$, upon changing variables in the integrals we obtain

$$\begin{aligned}
&\big|\mathbf{E}\{\eta_{sr}(\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\}\big| \\
&\qquad \le 12[\alpha(i)]^{\frac{\delta}{2+\delta}}\left[\mathbf{E}\,|\eta_{sr}|^{2+\delta}\,\mathbf{E}\big|\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r}\big|^{2+\delta}\right]^{\frac{1}{2+\delta}} \\
&\qquad \le C[\alpha(i)]^{\frac{\delta}{2+\delta}} h_n^{\frac{4m}{2+\delta}-4(m+r)}\left(\int\limits_{\mathbb{R}^m}|\boldsymbol{K}^{(r)}(u)|^{2+\delta}\,du\right)^{\frac{4}{2+\delta}} \\
&\qquad\qquad \times \sup_x\Big[a_t^{(2+\delta)+}(x)a_{s(s+j)(s+j+k),t}^{(2+\delta)+}(x,x,x)\Big]^{\frac{1}{2+\delta}} \le Ch_n^{\frac{4m}{2+\delta}-4(m+r)}[\alpha(i)]^{\frac{\delta}{2+\delta}}
\end{aligned}$$

and

$$\big|\mathbf{E}\{(\eta_{sr}\eta_{(i+s)r}\eta_{(i+j+s)r})\eta_{(i+j+k+s)r}\}\big| \le Ch_n^{\frac{4m}{2+\delta}-4(m+r)}[\alpha(k)]^{\frac{\delta}{2+\delta}}.$$

Similarly,

$$\begin{aligned}
\big|\mathbf{E}\{(\eta_{sr}\eta_{(i+s)r})(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\}\big| &= \big|\mathbf{E}\{(\eta_{sr}\eta_{(i+s)r})(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\} \\
&- \mathbf{E}(\eta_{sr}\eta_{(i+s)r})\,\mathbf{E}(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r}) + \mathbf{E}(\eta_{sr}\eta_{(s+i)r})\,\mathbf{E}(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\big| \\
&\le Ch_n^{\frac{4m}{2+\delta}-4(m+r)}[\alpha(j)]^{\frac{\delta}{2+\delta}} + \big|\mathbf{E}(\eta_{sr}\eta_{(s+i)r})\,\mathbf{E}(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\big|.
\end{aligned} \tag{29}$$

Also, we have

$$\big|\mathbf{E}(\eta_{sr}\eta_{(i+s)r})\big| \le Ch_n^{\frac{2m}{2+\delta}-2(m+r)}[\alpha(i)]^{\frac{\delta}{2+\delta}},$$

$$\big|\mathbf{E}(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\big| \le Ch_n^{\frac{2m}{2+\delta}-2(m+r)}[\alpha(k)]^{\frac{\delta}{2+\delta}}.$$

Substituting these two inequalities into (29), we obtain

$$\big|\mathbf{E}\{(\eta_{sr}\eta_{(i+s)r})(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\}\big| \le Ch_n^{\frac{4m}{2+\delta}-4(m+r)}\left([\alpha(i)]^{\frac{\delta}{2+\delta}}[\alpha(k)]^{\frac{\delta}{2+\delta}} + [\alpha(j)]^{\frac{\delta}{2+\delta}}\right).$$

Thus,

$$\begin{aligned}
&C^{-1}h_n^{\frac{-4m}{2+\delta}+4(m+r)}\sum_{s,i,j,k}\big|\mathbf{E}\{\eta_{sr}\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r}\}\big| \\
&\qquad \le 2n\sum_{i=1}^{n}\sum_{j,k=1}^{i}[\alpha(i)]^{\frac{\delta}{2+\delta}} + n\sum_{j=1}^{n}\sum_{i,k=1}^{\infty}[\alpha(i)]^{\frac{\delta}{2+\delta}}[\alpha(k)]^{\frac{\delta}{2+\delta}} \\
&\qquad \le n\sum_{i=1}^{n}i^2[\alpha(i)]^{\frac{\delta}{2+\delta}} + n^2\sum_{i=1}^{\infty}[\alpha(i)]^{\frac{\delta}{2+\delta}}\sum_{k=1}^{\infty}[\alpha(k)]^{\frac{\delta}{2+\delta}}
\end{aligned}$$

$$\le n \sum_{i=1}^{\infty} i^2 [\alpha(i)]^{\frac{\delta}{2+\delta}} + n^2 \left( \sum_{i=1}^{\infty} [\alpha(i)]^{\frac{\delta}{2+\delta}} \right)^2$$
$$\le n \int_1^{\infty} \tau^2 [\alpha(\tau)]^{\frac{\delta}{2+\delta}}\, d\tau + n^2 \left( \int_1^{\infty} [\alpha(\tau)]^{\frac{\delta}{2+\delta}}\, d\tau \right)^2, \tag{30}$$

where we have taken into account that $1 \ge \alpha(0) \ge \alpha(1) \ge \alpha(2) \ge \ldots$.

Equations (28) and (30) imply

$$\mathbf{E}(S_{tn}^{(rj)})^4 \le \frac{4!\, nCh_n^{\frac{4m}{2+\delta}}}{n^4 h_n^{4(m+r)}} \left[ n \left( \sum_{i=1}^{\infty} [\alpha(i)]^{\frac{\delta}{2+\delta}} \right)^2 + 3 \sum_{k=1}^{\infty} k^2 [\alpha(k)]^{\frac{\delta}{2+\delta}} \right]$$
$$\le \frac{Ch_n^{-\frac{2m\delta}{2+\delta}}}{n^2 h_n^{2(m+2r)}} \left( \sum_{i=1}^{\infty} [\alpha(i)]^{\frac{\delta}{2+\delta}} \right)^2 + \frac{Ch_n^{\frac{m(2-\delta)}{2+\delta}}}{n^2 h_n^{2(m+2r)} nh_n^m} \sum_{k=1}^{\infty} k^2 [\alpha(k)]^{\frac{\delta}{2+\delta}}. \tag{31}$$

Consider the first term on the right-hand side of inequality (31). Choose a sequence of positive integers $c(n)$ such that $c(n) = o(n)$ and $c(n) = O(h_n^{-m})$. Then, taking into account that $c^2(n) h_n^{\frac{2m\delta}{2+\delta}} \to \infty$ as $n \to \infty$, we obtain

$$\frac{h_n^{-\frac{2m\delta}{2+\delta}}}{n^2 h_n^{2(m+2r)}} \left( \sum_{i=c(n)}^{\infty} [\alpha(i)]^{\frac{\delta}{2+\delta}} \right)^2 = \frac{h_n^{-\frac{2m\delta}{2+\delta}}}{n^2 h_n^{2(m+2r)} c^2(n)} \left( \sum_{i=c(n)}^{\infty} c(n) [\alpha(i)]^{\frac{\delta}{2+\delta}} \right)^2$$
$$\le \frac{C}{n^2 h_n^{2(m+2r)} c^2(n) h_n^{\frac{2m\delta}{2+\delta}}} \left( \sum_{\tau=1}^{\infty} \tau [\alpha(\tau)]^{\frac{\delta}{2+\delta}} \right)^2 = o\!\left( \frac{1}{n^2 h_n^{2(m+2r)}} \right).$$

Now we show that

$$\frac{Ch_n^{-\frac{2m\delta}{2+\delta}}}{n^2 h_n^{2(m+2r)}} \left( \sum_{i=1}^{c(n)} [\alpha(i)]^{\frac{\delta}{2+\delta}} \right)^2 \le O\!\left( \frac{1}{n^2 h_n^{2(m+2r)}} \right).$$

Taking into account conditions (2) and (5) of the lemma, we find

$$\left|\mathbf{E}\{\eta_{sr}(\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\}\right| \le Ch_n^{-4r} \left( \int |\boldsymbol{K}^{(r)}(u)|\, du \right)^4$$
$$\times \sup_x a^+_{s(i+s)(i+j+s)(i+j+k+s),t}(x,x,x,x) \le Ch_n^{-4r} \overline{a}_{si(\cdot\cdot),t},$$
$$\left|\mathbf{E}\{(\eta_{sr}\eta_{(i+s)r}\eta_{(i+j+s)r})\eta_{(i+j+k+s)r}\}\right| \le Ch_n^{-4r} \overline{a}_{s(\cdot\cdot)k,t},$$
$$\left|\mathbf{E}\{(\eta_{sr}\eta_{(i+s)r})(\eta_{(i+j+s)r}\eta_{(i+j+k+s)r})\}\right| \le Ch_n^{-4r} \overline{a}_{s(\cdot)j(\cdot),t},$$

where

$$\overline{a}_{si(\cdot\cdot),t} = \max_{j,k} \sup_x a^+_{s(i+s)(i+j+s)(i+j+k+s),t}(x,x,x,x),$$
$$\overline{a}_{s(\cdot\cdot)k,t} = \max_{i,j} \sup_x a^+_{s(i+s)(i+j+s)(i+j+k+s),t}(x,x,x,x),$$
$$\overline{a}_{s(\cdot)j(\cdot),t} = \max_{i,k} \sup_x a^+_{s(i+s)(i+j+s)(i+j+k+1s),t}(x,x,x,x).$$

Similarly to (30), we obtain

$$\sum_{s,i,j,k} \big|\mathbf{E}\,\eta_{sr}\eta_{(i+s)r}\eta_{(i+j+s)r}\eta_{(i+j+k+s)r}\big| \le 3Ch_n^{-4r}\sum_{s=1}^{n}\sum_{i,j,k=1}^{s}\overline{a}_{si,t}$$
$$\le 3Ch_n^{-4r}\sum_{s=1}^{n} s^3\overline{a}_{s,t} \le 3Ch_n^{-4r} n\sum_{k=1}^{\infty}k^2\overline{a}_{k,t}$$
$$= 3Ch_n^{-4r} n\left(\sum_{k=1}^{c(n)-1}k^2\overline{a}_{k,t} + \sum_{k=c(n)}^{\infty}k^2\overline{a}_{k,t}\right), \qquad (32)$$

where $\overline{a}_{si,t} = \max(\overline{a}_{si(\cdot\cdot),t}, \overline{a}_{s(\cdot\cdot)i,t}, \overline{a}_{s(\cdot)i(\cdot),t})$ and $\overline{a}_{s,t} = \max\limits_{i=1,s}\overline{a}_{si,t}$.

The second term on the right-hand side of (32) is of the order of $o\left(\dfrac{1}{n^2h_n^{2(m+2r)}}\right)$, and the first term satisfies $\sum\limits_{k=1}^{c(n)-1} k^2\overline{a}_{k,t} \le Cnc^2(n)$. Since $\lim\limits_{n\to\infty} c(n)h_n = o(1)$, we have

$$\mathbf{E}(S_{nr}^{(rj)})^4 \le \frac{n^2Cc^2(n)h_n^4}{n^4h_n^{4(m+r)}} + o\left(\frac{1}{n^2h_n^{2(m+2r)}}\right) = O\left(\frac{1}{n^2h_n^{2(m+2r)}}\right). \quad \triangle$$

## REFERENCES


1. Härdle, W. and Tsybakov, A., Local Polynomial Estimators of the Volatility Function in Nonparametric Autoregression, *J. Econometrics*, 1997, vol. 81, no. 1, pp. 223–242.
2. Härdle, W., Tsybakov, A., and Yang, L., Nonparametric Vector Autoregression, *J. Statist. Plann. Inference*, 1998, vol. 68, no. 2, pp. 221–245.
3. Chen, G., Choi, Y.K., and Zhou, Y., Nonparametric Estimation of Structural Change Points in Volatility Models for Time Series, *J. Econometrics*, 2005, vol. 126, no. 1, pp. 79–114.
4. Raibman, N.S., *Chto takoe identifikatsiya?* (What Is Identification?), Moscow: Nauka, 1970.
5. Eykhoff, P., System Identification: Parameter and State Estimation, London: Wiley, 1974. Translated under the title *Osnovy identifikatsii sistem upravleniya. Otsenivanie parametrov i sostoyaniya*, Moscow: Mir, 1975.
6. Pashchenko, F.F., Response Function and Its Application in Optimal Model Choice, in *Sistemy upravleniya* (Control Systems), Moscow: Nauka, 1973, pp. 72–78.
7. Kitaeva, A.V., Koshkin, G.M., and Piven, I.G., Nonparametric Identification in Economic Systems, *Obozr. Prikl. Prom. Mat.*, 2008, vol. 15, no. 4, pp. 588–612.
8. Kitaeva, A.V. and Koshkin, G.M., Recurrent Nonparametric Estimation of Functions from Functionals of Multidimensional Density and Their Derivatives, *Autom. Remote Control*, 2009, no. 3, pp. 389–407 [*Avtomat. i Telemekh.* (Engl. Transl.), 2009, no. 3, pp. 48–67].
9. Wolverton, C.T. and Wagner, T.J., Asymptotically Optimal Discriminant Functions for Pattern Classification, *IEEE Trans. Inform. Theory*, 1969, vol. 15, no. 2, pp. 258–265.
10. Banon, G., Sur un estimateur non paramétrique de la densité de probabilité, *Rev. Statist. Appl.*, 1976, vol. 24, no. 4, pp. 61–73.
11. Nadaraya, E.A., On Estimation of Regression, *Teor. Veroyatnost. i Primenen.*, 1964, vol. 9, no. 1, pp. 157–159.
12. Watson, G.S., Smooth Regression Analysis, *Sankhyā, Ser. A*, 1964, vol. 26, pp. 359–372.
13. Ahmad, J.A. and Lin, P.E., Nonparametric Sequential Estimation of a Multiple Regression Function, *Bull. Math. Statist.*, 1976, vol. 17, no. 1–2, pp. 63–75.

14. Buldakov, V.M. and Koshkin, G.M., On Recursive Estimates of Probability Density and Regression Line, *Probl. Peredachi Inf.*, 1977, vol. 13, no. 1, pp. 58–66 [*Probl. Inf. Trans.* (Engl. Transl.), 1977, vol. 13, no. 1, pp. 41–48].

15. Devroye, L.P. and Wagner, T.J., On the $L_1$ Convergence of Kernel Estimator of Regression Functions with Applications in Discrimination, *Z. Wahrsch. Verw. Gebiete*, 1980, vol. 51, no. 1, pp. 15–25.

16. Krzyźak, A. and Pawlak, M., Almost Everywhere Convergence of a Recursive Regression Function Estimate and Classification, *IEEE Trans. Inform. Theory*, 1984, vol. 30, no. 1, pp. 91–93.

17. Greblicki, W. and Pawlak, M., Necessary and Sufficient Consistency Conditions for a Recursive Kernel Regression Estimate, *J. Multivariate Anal.*, 1987, vol. 23, no. 1, pp. 67–76.

18. Krzyźak, A., Global Convergence of the Recursive Kernel Estimates with Applications in Classification and Nonlinear System Estimation, *IEEE Trans. Inform. Theory*, 1992, vol. 38, no. 4, pp. 1323–1338.

19. Walk, H., Strong Universal Pointwise Consistency of Recursive Kernel Regression Estimates, *Ann. Inst. Statist. Math.*, 2001, vol. 53, no. 4, pp. 691–707.

20. Koshkin, G.M., Deviation Moments of the Substitution Estimator and of Its Piecewise-Smooth Approximations, *Sibirsk. Mat. Zh.*, 1999, vol. 40, no. 3, pp. 605–618 [*Siberian Math. J.* (Engl. Transl.), 1999, vol. 40, no. 3, pp. 515–527].

21. Bosq, D. and Cheze-Payaud, N., Optimal Asymptotic Quadratic Error of Nonparametric Regression Function Estimates for a Continuous-Time Process from Sampled-Data, *Statistics*, 1999, vol. 32, no. 3, pp. 229–247.

22. Fikhtengol'ts, G.M., *Kurs differentsial'nogo i integral'nogo ischisleniya* (A Course in Differential and Integral Calculus), Moscow: Nauka, 1966, vol. 2.

23. Gradshtein, I.S. and Ryzhik, I.M., *Tablitsy integralov, summ, ryadov i proizvedenii*, Moscow: Nauka, 1971, 5th ed. Translated under the title *Table of Integrals, Series, and Products*, New York: Academic, 1980.

24. Masry, E. and Tjøstheim, D., Nonparametric Estimation and Identification of Nonlinear ARCH Time Series, *Econometric Theory*, 1995, vol. 11, no. 2, pp. 258–289.

25. Bhattacharya, R.N. and Lee, C., Ergodicity of Nonlinear First Order Autoregressive Models, *J. Theoret. Probab.*, 1995, vol. 8, no. 1, pp. 207–219.

26. Lu, Z.D., On the Geometric Ergodicity of a Non-linear Autoregressive Model with an Autoregressive Conditional Heteroscedastic Term, *Statist. Sinica*, 1998, vol. 8, no. 4, pp. 1205–1217.

27. Chen, M. and Chen, G., Geometric Ergodicity of Nonlinear Autoregressive Models with Changing Conditional Variances, *Canad. J. Statist.*, 2000, vol. 28, no. 3, pp. 605–613.

28. Tjøstheim, D. and Auestad, B.H., Nonparametric Identification of Nonlinear Time Series: Projections, *J. Amer. Statist. Association*, 1994, vol. 89, no. 428, pp. 1398–1409.

29. Collomb, G., Estimation non paramétrique de la régression par la méthode du noyau, *Thèse, Univ. Paul Sabatier, Toulouse*, 1976.

30. Davydov, Yu.A., The Convergence of Distributions Which Are Generated by Stationary Random Processes, *Teor. Veroyatnost. i Primenen.*, 1968, vol. 13, no. 4, pp. 730–737.

31. Billingsley, P., *Convergence of Probability Measures*, New York: Wiley, 1968. Translated under the title *Skhodimost' veroyatnostnykh mer*, Moscow: Nauka, 1977.