Journal of Sciences, Islamic Republic of Iran 17(1): 75-81 (2006) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

Wavelet Based Estimation of the Derivatives of a Density for a Discrete-Time Stochastic Process: L_p-Losses

H. Doosti,* H.A. Niroumand, and M. Afshari

Department of Statistics, School of Mathematical Sciences, Ferdowsi University, P.O. Box 1159-91775, Mashhad, Islamic Republic of Iran

Abstract

We propose a method of estimation of the derivatives of probability density based on wavelets methods for a sequence of random variables with a common one-dimensional probability density function and obtain an upper bound on L_p-losses for such estimators. We suppose that the process is strongly mixing and we show that the rate of convergence essentially depends on the behavior of a special quadratic characteristic.

Keywords: Mixing sequences; Multiresolution analysis; Besov space; wavelets; nonparametric estimation of derivative of density

1. Introduction

Methods of estimation of density and regression function are quite common in statistical applications. Recently, there has been a lot of interest in nonparametric estimation of such functions based on wavelets. The reader may be referred to [17] and [32] for a detailed coverage of wavelet theory in statistics and to [25-26] for a recent comprehensive review and application of these and other methods of nonparametric functional estimation.

Antoniadis *et al.* [2] and Masry [22] among others discussed the estimation of regression and density function using the wavelets. Prakasa Rao [24] considered the use of wavelets for estimating the derivatives of a density and investigated further their use for estimating the integrated squared density [24]. Walter and Ghorai [33] discussed the advantages and disadvantages of wavelet based methods of nonparametric estimation from *i.i.d.* sequences of random variables. Prakasa Rao [27] echoed the same advantages and disadvantages for the case of associated sequences. Furthermore, he pointed out that these methods allow one to obtain precise limits on the asymptotic mean squared error for the estimator of density and its derivatives as well as some other functional of the density [24,25]. Recently, Doosti *et al.* [14] have shown that the results in [27] can be extended to the case of negatively associated sequences.

Chaubey *et al.* [5,6] extended results of [24] for a sequence of associated and negatively associated random variables. Doosti and Nezakati [15] and Chaubey and Doosti [7] obtained upper bounds for the L_p-losses for a sequence of m-dependent random variables in density function estimation and estimation of derivatives of density functions, respectively. We show that the L_p-error of a linear wavelet density estimator for some stochastic process (the estimator is constructed from dependent data) attains the same rate as when the observations are independent. The organization of the paper is as follows. In Section 2, we discuss the preliminaries of the wavelet based

* *E-mail: doosti@math.um.ac.ir*

estimation of the derivatives of the density along with the necessary underlying setup considered in [24]. Then in Section 3, we extend results of [21].

In Section 4, we study some conditions for some processes which are not necessarily Markov.

2. Preliminaries

Let $\{X_n, n \geq 1\}$ be a sequence of random variables on the probability space $(\Omega, \aleph, P)$. We suppose that X_i has a bounded and compactly supported marginal density $f(.)$, with respect to Lebesgue measure, which does not depend on *i*. We estimate this density from *n* observations $X_i, i = 1, \ldots, n$. For any function $f \in \mathbf{L}_2(\mathbf{R})$, we can write a formal expansion (see [8]):

$$f = \sum_{k \in Z} \alpha_{j_0,k} \phi_{j_0,k} + \sum_{j \geq j_0} \sum_{k \in Z} \delta_{j,k} \psi_{j,k} = P_{j_0} f + \sum_{j \geq j_0} D_j f$$

where the functions

$$\phi_{j_0,k}(x) = 2^{j_0/2} \phi(2^{j_0} x - k)$$

and

$$\psi_{j,k}(x) = 2^{j/2} \psi(2^j x - k)$$

Constitute an (inhomogeneous) orthonormal basis of $\mathbf{L}_2(\mathbf{R})$. Here $\phi(x)$ and $\psi(x)$ are the scale function and the orthogonal wavelet, respectively. Wavelet coefficients are given by the integrals

$$\alpha_{j_0,k} = \int f(x) \phi_{j_0,k}(x) dx, \quad \delta_{j,k} = \int f(x) \psi_{j,k} dx$$

We suppose that both ϕ and $\psi \in \mathbf{C}^{r+1}$, $r \in \mathbf{N}$, have compact supports included in $[-\delta, \delta]$. Note that, by corollary 5.5.2 in [9], ψ is orthogonal to polynomials of degree $\leq r$, *i.e.*

$$\int \psi(x) x^l dx = 0, \forall l = 0.1, \ldots, r$$

We suppose that *f* belongs to the Besov class (see [23], §VI.10), for some $0 \leq s \leq r+1$, $p \geq 1$ and $q \geq 1$, where

$$F_{s,p,q} = \{ f \in B^s_{p,q}, \| f \|_{B^s_{p,q}} \leq M \}$$

$$\| f \|_{B^s_{p,q}} = \| P_{j_0} f \|_p + (\sum_{j \geq j_0} (\| D_j f \|_p 2^{js})^q)^{1/q}$$

We may also say *f* belongs to the Besov class if and only if

$\| \alpha_{j_0 .} \|_{l_p} < \infty$, **and**

$$(\sum_{j \geq j_0} (\| \delta_{j .} \|_{l_p} 2^{j(s+1/2-1/p)})^q)^{1/q} < \infty \quad (2.1)$$

where We consider Besov spaces essentially because of their executional expressive power [see [31] and the discussion in [13]. We construct the density estimator see [27],

$\hat{f} = \sum_{k \in K_{j_0}} \hat{\alpha}_{j_0,k} \phi_{j_0,k}$, **with**

$$\hat{\alpha}_{j_0,k} = \frac{1}{n} \sum_{i=1}^{n} \phi_{j_0,k}(X_i), \quad (2.2)$$

where K_{j_0} is the set of k such that $\sup p(f) \cap \sup p(\phi_{j_0,k}) \neq \phi$. The fact that ϕ has a compact support implies that K_{j_0} is finite and $card(K_{j_0}) = O(2^{j_0})$. Wavelet density estimators aroused much interest in the recent literature, see [12] and [16]. In the case of independent samples, the properties of the linear estimator (2.2) have been studied for a variety of error measures and density classes; see [19,21,30]. In the setup considered by Prakasa Rao [24], we assume that ϕ is a scaling function generating an *r*-regular multiresolution analysis and $f^{(d)} \in \mathbf{L}_2(\mathbf{R})$. Furthermore, we assume that there exists $C_m \geq 0$ and $\beta_m \geq 0$ such that

$$| f^{(m)}(x) | \leq C_m (1 + |x|)^{-\beta_m}, \quad 0 \leq m \leq d. \quad (2.3)$$

Prakasa Rao [24] showed that the projection of $f^{(d)}$ on V_{j_0} is

$$f^{(d)}_{n,d}(x) = \sum_k a_{j_0,k} \phi_{j_0,k}(x).$$

where

$$a_{j_0,k} = (-1)^d \int \phi^{(d)}_{j_0,k}(x) f_X(x) dx.$$

So its estimator is

$$\hat{f}^{(d)}_{n,d}(x) = \sum_k \hat{a}_{j_0,k} \phi_{j_0,k}(x), \quad (2.4)$$

where

$$\hat{a}_{j_0,k} = \frac{(-1)^d}{n} \sum_{i=1}^{n} \phi^{(d)}_{j_0,k}(X_i),$$

For the estimator in Equation (2.4), the sum is considered for $k \in K_{j_0}$.

If we want the estimator of derivative of density function in Equation (2.4) for a stochastic process to attain the same result as for the associated, negatively associated and *m*-dependent cases, we have to impose certain *weak dependence* conditions on the considered process $\{X_n, n \ge 1\}$ defined on the $(\Omega, \aleph, P)$. Let $\mathbf{N}_k^m$ denote the σ-algebra generated by the events

$$\{X_k \in A_k, ..., X_m \in A_m\}.$$

We consider the following classical mixing conditions:

1. *strong mixing* (s.m.), also called α-mixing,

$$\sup_m \sup_{A \in \mathbf{N}_1^m, B \in \mathbf{N}_{m+s}^\infty} |P(AB) - P(A)P(B)| = \alpha(s) \to 0 \quad \text{as } s \to \infty,$$

2. *complete regularity* (c.r.), also called β-mixing,

$$\sup_m E\{Var_{B \in \mathbf{N}_{m+s}^\infty} |P(B \mid \mathbf{N}_1^m) - P(B)|\} = \beta(s) \to 0 \quad \text{as } s \to \infty,$$

3. *uniformly strong mixing* (u.s.m.), also called ϕ-mixing

$$\sup_m \sup_{A \in \mathbf{N}_1^m, P(A)>0, B \in \mathbf{N}_{m+s}^\infty} \frac{|P(AB) - P(A)P(B)|}{P(A)} = \phi(s) \to 0 \quad \text{as } s \to \infty,$$

4. *ρ-mixing*

$$\sup_m \sup_{X \in L^2(\mathbf{N}_1^m), Y \in L^2(\mathbf{N}_{m+s}^\infty)} |corr(X, Y)| = \rho \to 0 \quad s \to \infty.$$

Following [10] we denote by $\text{var}_{A \in \mathbf{F}}\, \mu(A)$ the total variation of the restriction of the measure μ defined on some σ-algebra $\mathbf{N}$ to the σ-algebra $\mathbf{F}$. We call the corresponding values $\alpha(s)$, $\beta(s)$ and $\phi(s)$ the s.m., c.r. and u.s.m, coefficients, respectively.

Moreover, we will show that under certain conditions of weak dependence (more precisely, under strong mixing conditions) the rate of convergence of wavelet estimators is the same (up to a constant) as for the independent case. As we will see, for the estimators to attain the "independent" rates of convergence, we should require the stochastic process to satisfy some local regularity conditions.

3. Main Results

First, we present a bound for the moment of order *p* of the sum of *N* random variables which depends on the second moment and mixing coefficients. This bound constitutes the basis of the main results of this paper-Theorems 1, 2 and 3. This is a Rosenthal-type inequality (see [28,29] for other inequalities of this kind). We suppose that (ξ_i) is a strong mixing sequence of real random variables on the probability space $(\Omega, \aleph, P)$.

Lemma 3.1. *[21] Let $2 \le p < \infty$ and $\xi_1, ..., \xi_n$ be a sequence of real-valued random variable such that $\mathbf{E}(\xi_i) = 0$, $\|\xi_i\|_\infty < S$, and $\mathbf{E}(\xi_i^2) \le \sigma^2$. Then there exists C such that:*

$$\mathbf{E}(|\sum_{i=1}^n \xi_i|^p) \le C\{(\frac{n\sigma^2}{l})^{p/2}\sigma_l^p + \frac{n\sigma^2}{l}\sigma_l^2(lS)^{p-2} + S^p n^p \alpha(l)\},$$

where $l \in \mathbf{N}$, $2 \le l \le n/2$, $\sigma_l^2 = \max\{\max_{1 \le u \le n} \sigma_u^2(l)$, $\max_{1 \le u \le n} \sigma_u^2(l-1)\}$ and $\sigma_u^2(l) = \mathbf{E}(\sum_{i=u}^{u+l-1} \xi_i)^2$.

In what follows, $\alpha(l)$ is the strong mixing coefficient defined in the introduction. We denote

$$\sigma_l^2 = \max_{1 \le u \le n-l+1} \max(\sigma_u^2(l), \sigma_u^2(l-1)),$$
$$\sigma_u^2(l) = \max_{k \in K_{j_0}} \mathbf{E}(\sum_{i=u}^{u+l-1}(\phi_{j_0,k}(X_i) - \mathbf{E}\phi_{j_0,k}(X_i)))^2. \tag{3.1}$$

In Theorems 3.1 and 3.2, the results of [21] are obtained by letting $d = 0$.

Theorem 3.1. *Let $f_{n,d}^{(d)}(x) \in F_{s,p,q}$ with $s \ge 1/p$, $p \ge 1$, and $q \ge 1$. Suppose that there exist constants $\alpha > 1$ and c_α such that for any l, $\alpha(l) \le c_\alpha \alpha^{-1}$. Furthermore, suppose that there is a function g with $g(l) \ge G$ (G is a positive constant), such that for any $l = O(\ln(n))$, $\sigma_l^2 \le \lg(l)$. Then for $p \ge \max(2, p)$, there exists a constant C such that*

$$\mathbf{E}\left\| f_{n,d}^{(d)}(x) - \hat{f}_{n,d}^{(d)}(x) \right\|_p^2 \le C[\frac{n}{g(\ln(n))}]^{-\frac{2(s'-d)}{1+2s'}},$$

where $s' = s + 1/p' - 1/p$ and $2^{j_0} = [\frac{n}{g(\ln(n))}]^{\frac{1}{1+2s}}$.

Theorem 3.2. *Let $f_{n,d}^{(d)}(x) \in F_{s,p,q}$ with $s \ge 1/p$, $p \ge 1$, and $q \ge 1$. Suppose that $\alpha(l) \le c_\alpha l^{-\alpha}$, for any $l \in \mathbf{N}$, $2 \le l \le n/2$. Let us set $\mu = p(s+1)/[\alpha(1+2s)]$ $\alpha \ge p(1+s)/s$ and suppose that there is a function g with $g(l) \ge G$ (G is a positive constant), such that for* *any* $l = O(\ln(n))$, $\sigma_l^2 \le \lg(l)$. *Then for* $p \ge \max(2, p)$, *there exists a constant C such that*

$$\mathbf{E}\left\| f_{n,d}^{(d)}(x) - \hat{f}_{n,d}^{(d)}(x) \right\|_p^2 \le C[\frac{n}{g(n^{\mu})}]^{-\frac{2(s'-d)}{1+2s'}},$$

where $s' = s + 1/p' - 1/p$.

Remark. In the case of independent random variables, $\sigma_l^2 = O(l)$. Moreover, in the dependent case a rough bound $\sigma_l^2 = O(l^2)$ can be easily obtained. If some additional conditions are imposed on the process (X_i), the bound $\sigma_l^2 = O(l)$ can be achieved (see next section). Let us consider the following condition: $\mathbf{C}_\sigma$: $\sigma_l^2 = O(l)$.

When the condition $\mathbf{C}_\sigma$ is satisfied, the same rate as for the independent case, $O(n^{-2s'/(2s'+l)})$, is attained. We study $\mathbf{C}_\sigma$ in the next section. Theorems 3.1 and 3.2 are corollaries of the following lemmas:

Lemma 3.2. *Let* $f_{n,d}^{(d)}(x) \in F_{s,p,q}$ *with* $s \ge 1/p$, $p \ge 1$, $q \ge 1$ *and* $p \ge \max(2, p)$. *Then there exists a constant C such that*

$$\mathbf{E}\left\| f_{n,d}^{(d)}(x) - \hat{f}_{n,d}^{(d)}(x) \right\|_p^2 \le C\{2^{-2j_0 s} + \frac{2^{j_0}}{n}\frac{\sigma_l^2}{l} + (\frac{2^{j_0}}{n})^{2-2/p} l^{2/P(p-3)} \sigma_l^{4/p} + 2^{2j_0}\alpha(l)^{2/p}\}, \tag{3.1}$$

where $l \in \mathbf{N}$, $2 \le l \le n/2$, *and* $s' = s + 1/p' - 1/p$.

Proof. First, we decompose $\mathbf{E}\left\| \hat{f}_{n,d}^{(d)}(x) - f_{n,d}^{(d)}(x) \right\|_{p'}^2$ into a bias term and a stochastic term

$$\begin{aligned}\mathbf{E}\left\| \hat{f}_{n,d}^{(d)}(x) - f_{n,d}^{(d)}(x) \right\|_{p'}^2 &\le 2(\left\| f_{n,d}^{(d)} - P_{j_0} f_{n,d}^{(d)} \right\|_{p'}^2 \\ &+ \mathbf{E}\left\| \hat{f}_{n,d}^{(d)} - P_{j_0} f_{n,d}^{(d)} \right\|_{p'}^2) \\ &= 2(T_1 + T_2)\end{aligned} \tag{3.2}$$

Now, we want to find upper bounds for T_1 and T_2.

$$\sqrt{T_1} = \left\| \sum_{j \ge j_0} D_j f_{n,d}^{(d)} \right\|_{p'} \le \sum_{j \ge j_0} (\left\| D_j f_{n,d}^{(d)} \right\|_{p'} 2^{js'}) 2^{-js'}$$

$$\le \{\sum_{j \ge j_0} (\left\| D_j f_{n,d}^{(d)} \right\|_{p'} 2^{js'})^q\}^{1/q} \{\sum_{j \ge j_0} 2^{-js'q'}\}^{1/q'}$$

By Holder's inequality, with $1/q + 1/q' = 1$, from the above equation, we have

$$T_1 \le C \left\| f_{n,d}^{(d)} \right\|_{B_{p',q}^{s'}} 2^{-s'j_0} \le C \left\| f_{n,d}^{(d)} \right\|_{B_{p,q}^{s}} 2^{-s'j_0}. \tag{3.3}$$

The last inequality holds, because of the continuous Sobolev injection, see [31] and the discussion in [12], which implies that for $B_{p,q}^s \subset B_{p',q}^{s'}$, one gets,

$$\left\| f_{n,d}^{(d)} \right\|_{B_{p',q}^{s'}} \le \left\| f_{n,d}^{(d)} \right\|_{B_{p,q}^{s}}.$$

Therefore, we get from Equation (3.3)

$$T_1 \le K\, 2^{-2s'j_0}. \tag{3.4}$$

Next, we have

$$\begin{aligned}T_2 &= \mathbf{E}\left\| \hat{f}_{n,d}^{(d)} - P_{j_0} f_{n,d}^{(d)} \right\|_{p'}^2 \\ &= \mathbf{E}\left\| \sum_{k \in K_{j_0}} (\hat{a}_{j_0,k} - a_{j_0,k}) \phi_{j_0,k}(x) \right\|_{p'}^2.\end{aligned}$$

This gives by using Lemma 1 in [21], p. 82 (using [23]),

$$T_2 \le C\, \mathbf{E}\{\left\| \hat{a}_{j_0,k} - a_{j_0,k} \right\|_{lp'}^2\} 2^{2j_0(1/2-1/p')}.$$

Further, by using Jensen's inequality the above equation implies,

$$T_2 \le C\, 2^{2j_0(1/2-1/p')} \{\sum_{k \in K_{j_0}} \mathbf{E}\left| \hat{a}_{j_0,k} - a_{j_0,k} \right|^{p'}\}^{2/p'}. \tag{3.5}$$

To complete the proof, it is sufficient to estimate $\mathbf{E}\left| \hat{a}_{j_0,k} - a_{j_0,k} \right|^{p'}$. We know that

$$\hat{a}_{j_0,k} - a_{j_0,k} = \frac{1}{n}\sum_{i=1}^{n} \{[\phi_{j_0,k}^{(d)}(X_i) - a_{j_0,k}]\}.$$

Denote $\xi_i = [\phi_{j_0,k}^{(d)}(X_i) - a_{j_0,k}]$. Note that $\|\xi_i\|_\infty \le K.2^{j_0(1/2+d)}\|\phi\|_\infty$, $\mathbf{E}\xi_i = 0$, $\mathbf{E}\xi_i^2 \le \|f\|_\infty 2^{2j_0 d} \int_{-\infty}^{\infty} \phi^{2(d)}(v)dv$ and $|\hat{\alpha}_{j_0,k}^{(d)} - \alpha_{j_0,k}| = \frac{(-1)^d}{n} |\sum_{i=1}^{n} \xi_i|$. Hence applying the result in Equation (3.1) and using $card(K_{j_0}) = O(2^{j_0})$ we have

$$\{ \sum_{k \in K_{j_0}} \mathbf{E} | \hat{a}_{j_0,k} - a_{j_0,k} |^{p'} \}^{2/p'}$$

$$\leq \{C 2^{j_0} \frac{1}{n^{p'}} (n 2^{(j_0/2)(p'-2+2dp')} c_1 + n^{p'/2} 2^{j_0 dp'} c_2)\}^{2/p'}$$

$$\leq K_1 \{ \frac{2^{j_0(1+2d)}}{n^{2(1-1/p')}} + \frac{2^{2j_0 d + 2(j_0/p')}}{n} \}.$$

Now by substituting the above bound in (3.5), we get

$$T_2 \leq K_1 2^{2j_0(1/2-1/p')} \{ \frac{2^{j_0(1+2d)}}{n^{2(1-1/p')}} + \frac{2^{2j_0 d+2(j_0/p')}}{n} \}$$

$$= K_1 \{ \frac{2^{2j_0 - 2(j_0/p') + 2j_0 d}}{n^{2-2/p'}} + \frac{2^{j_0(1+2d)}}{n} \}$$

$$= K_1 \{ \frac{2^{j_0}}{n} (\frac{2^{j_0(1+2d)}}{n})^{1-2/p'} + \frac{2^{j_0(1+2d)}}{n} \}.$$

Since $n \geq 2^{j_0}$ and $1 - 2/p' \geq 0$ imply $(\frac{2^{j_0}}{n})^{1-2/p'} \leq 1$, we have the inequality

$$T_2 \leq \frac{K_2 2^{j_0(1+2d)}}{n}. \tag{3.6}$$

By using the bounds obtained in (3.5) and (3.6), and choosing j_0 such that $2^{j_0} = n^{\frac{1}{1+2s'}}$ in (3.3), the theorem is proved.

Proof of Theorems 3.1 and 3.2. To obtain the results it is sufficient to balance the terms in the upper bound (3.1) by choosing the parameters 2^{j_0} and *l*. We first choose 2^{j_0} to obtain equilibrium between the two first terms of the right-hand side of (3.1), and next, we choose *l* to attain the correct rate in the last terms.

4. Discussion of the Theorem's Assumptions

We study some additional conditions which the bound $\sigma_l^2 = O(l)$ can be achieved.

Comment. Since the inequality $\rho(t) \leq 2[\phi(t)]^{1/2}$ holds (see [16]) If the process is ϕ-mixing and $\Sigma_{t=1}^{\infty} [\phi(t)]^{1/2} < \Phi < \infty$, then the process is ρ-mixing and $\Sigma_{t=1}^{\infty} \phi(t) < R < \infty$.

For Gaussian processes ϕ-mixing is equivalent to *m*-dependence (see [18], Section 1), whereas ρ-mixing is equivalent to α-mixing (see [20], Section 2.1).

Theorem 4.1. Let $(X_n, n \geq 1)$ be a stochastic process on R and suppose the process is ρ-mixing and $\Sigma_{t=1}^{\infty} \phi(t) < \Phi < \infty$. Suppose X_n admits a bounded marginal density which is common for any *n*, then there exists a constant such that for any *n*, $\sigma_l^2 \leq Gl$.

Proof. We use the decomposition

$$\sigma_{k,u}^2 (l) \leq V_{u,k}(l) + C_{u,k}(l), \tag{4.1}$$

with

$$V_{u,k}(l) = \Sigma_{i=1}^{l+u-1} E_f (\varphi_{j0,k}^{(d)}(X_i) - E_f(\varphi_{j0,k}^{(d)}(X_i))^2, \tag{4.2}$$

$$C_{u,k}(l) = 2 \Sigma \Sigma_{u \leq m < t \leq l+u-1} \left| \text{cov}(\varphi_{j0,k}^{(d)}(X_m), \varphi_{j0,k}^{(d)}(X_t)) \right|. \tag{4.3}$$

The term in (4.2) can be estimated as follows:

$$V_{u,k}(l) \leq l \max_{u \leq i \leq (l+u-1)} E_f (\varphi_{j0,k}^{(d)}(X_i))^2 \leq l \| f \|_\infty \tag{4.4}$$

To bound the term $C_{u,k}(l)$ we apply a ρ-mixing covariance inequality (see [16], Section 1.2.2) that is,

$$| \text{cov}(\varphi_{j_0,k}^{(d)}(X_m), \varphi_{j_0,k}^{(d)}(X_t)) |$$

$$\leq 2\rho(t-m)(E_f(\varphi_{j_0,k}^{(d)}(X_m))^2)^{1/2} (E_f(\varphi_{j_0,k}^{(d)}(X_t)^2)^{1/2}$$

$$\leq 2 \| f \|_\infty \rho(t-m).$$

So we obtain

$$C_{u,k}(l) = 2\Sigma\Sigma_{u \leq m \leq t \leq l+u-1} | \text{cov}(\varphi_{j_0,k}^{(d)}(X_m), \varphi_{j_0,k}^{(d)}(X_t)) |$$

$$\leq 2 \| f \|_\infty 2\Sigma\Sigma_{u \leq m < t \leq l+u-1} \rho(t-m) \leq \rho \| f \|_\infty Rl.$$

Theorem 4.2. Let $(X_n, n \geq 1)$ be an R-valued stocastic process. Suppose that X_n admits a bounded marginal density which is common for all $1 \leq n \leq N$, if the distribution of (X_m, X_t) has a joint density $f_{m,t}(.,.)$ such that for all m and t, $m \neq t$ $(\int | f_{m,t}(x,y) |^\nu)^{1/\nu} = \| f_{m,t}(.,.) \|_\nu \leq F_\nu < \infty$ for some ν<2, (with the usual modification for $\nu = \infty$) then there exists a constant G such that for all $l \leq 2^{j_0(1-(2/\nu))}$, $\sigma_l^2 \leq Gl$.

Proof. Denote $l_{j_0} \leq 2^{j_0(1-(2/\nu))}$. We bound $\sigma_{k,u}^2$ as it has been done in (4.1). Next we use the partition $\Gamma = \{(t,m) \in N^2, u \leq m < t < (l+u-1)\} = \Gamma_1 \bigcup \Gamma_2$ where

$$\Gamma = \{(t,m) \in \Gamma, 0 < t - m \leq \min(l, l_{j_0})\}$$

and Γ_2 is the compliment of Γ_1 in Γ. Due to the inequality (4.4) it is enough to estimate the $C_{(u,k)}(l)$. To bound the covariance sum over the domain Γ_1 we use

$$|\operatorname{cov}(\varphi^{(d)}_{j_0,k}(X_m)\varphi^{(d)}_{j_0,k}(X_t)|$$
$$\le (\int|\varphi^{(d)}_{j_0,k}(x)\varphi^{(d)}_{j_0,k}(y)|f_{m,t}(x,y)dxdy \tag{4.5}$$
$$+(\int|\varphi^{(d)}_{j_0,k}(x)|f(x)dx))^2$$
$$\le F_\nu\left\|\varphi^{(d)}_{j_0,k}\right\|^{2/\nu'}_{\nu'}+(E_f\ |\varphi^{(d)}_{j_0,k}(X_1))^2 \tag{4.5}$$

with $1/\nu'+1/\nu=1$.

Next, since

$$\left\|\varphi^{(d)}_{j_0,k}\right\|^{2/\nu'}_{\nu'}=2^{j_0(1-2/\nu')}(\int(|\varphi^{(d)}(y)|)^{\nu'}dy)^{2/\nu'}$$
$$\le(2\delta)^{2(1-1/\nu)}\left\|\varphi^{(d)}\right\|^2_\infty 2^{j_0(1-2/\nu)},$$

and $E_f\ |\varphi^{(d)}_{j_0,k}(X_1)|\le 2\delta\left\|\varphi^{(d)}\right\|_\infty\|f\|_\infty 2^{-j_0/2}$ we have by (4.5)

$$|\operatorname{cov}(\varphi^{(d)}_{j_0,k}(X_m),\varphi^{(d)}_{j_0,k}(X_t))|$$
$$\le F_\nu(2\delta)^{2(1-1/\nu)}\left\|\varphi^{(d)}\right\|^2_\infty 2^{-j_0(1-2/\nu)}$$
$$+4\left\|\varphi^{(d)}\right\|^2_\infty(\delta)^2\|f\|^2_\infty 2^{-j_0}.$$

Thus

$$\sum_{(\Gamma_1)}|\operatorname{cov}(\varphi^{(d)}_{j_0,k}(X_m),\varphi^{(d)}_{j_0,k}(X_t))|$$
$$\le l\min(l_{j_0},l)[F_\nu(2\delta)^{2(1-1/\nu)}\left\|\varphi^{(d)}\right\|^2_\infty 2^{-j_0(1-2/\nu)} \tag{4.6}$$
$$+4\left\|\varphi^{(d)}\right\|^2_\infty(\delta)^2\|f\|^2_\infty 2^{-j_0}]$$
$$\le[F_\nu(2\delta)^{2(1-1/\nu)}\left\|\varphi^{(d)}\right\|^2_\infty+4\left\|\varphi^{(d)}\right\|^2_\infty\delta^2\|f\|^2_\infty]l$$

For $l\le l_{j_0}$ the set Γ_2 is empty and (4.6) together with (4.4) complete the proof.

Theorem 4.3. Let $(X_n,n\ge1)$ be an R-valued stochastic process and (X_n) is α-mixing with $\alpha(l)\le c_\alpha l^{-\alpha}$, $\alpha>[\nu/\nu-2]+1$. Suppose that (X_n) admits a bounded marginal density which is common for all $1\le n\le N$ then there exists a positive G such that for all l, $\sigma^2_l\le Gl$.

Proof. We estimate the covariance sum over the set Γ_2. By using the Davylov inequality (see [16], Section 1.2.2)

$$|\operatorname{cov}(\varphi^{(d)}_{j_0,k}X(m),\varphi^{(d)}_{j_0,k}X(t))|$$
$$\le 8\alpha(t-m)2^{j_0}\left\|\varphi^{(d)}\right\|^2_\infty$$
$$\le 8C_\alpha\left\|\varphi^{(d)}\right\|^2_\infty 2^{j_0}(t-m)^{-\alpha}$$

We obtain

$$\sum_{\Gamma_2}|\operatorname{cov}(\varphi^{(d)}_{j_0,k}(X_m),\varphi^{(d)}_{j_0,k}(X_t))|$$
$$\le 8C_\alpha\left\|\varphi^{(d)}\right\|^2_\infty\sum_{\Gamma_2}2^{j0}(t-m)^{-(\alpha-\frac{\nu}{\nu-2})}(t-m)^{-(\nu/(\nu-2))}$$
$$\le 8C_\alpha\left\|\varphi^{(d)}\right\|^2_\infty\sum_u^{l+u-1}2^{j_0}\sum_{s>l_{j0}}s^{-(\alpha-\frac{\nu}{\nu-2})}s^{-(\nu/(\nu-2))} \tag{4.7}$$
$$\le 8C_\alpha\left\|\varphi^{(d)}\right\|^2_\infty\sum_u^{l+u-1}2^{j_0}l_{j_0}^{-(\nu/\nu-2)}\int_1^\infty x^{-(\alpha-\nu/(\nu-2))}$$
$$\le 8C_\alpha\left\|\varphi^{(d)}\right\|^2_\infty(\alpha-\frac{\nu}{\nu-2}-1)^{-1}l$$

because $\alpha>\frac{\nu}{\nu-2}$.

When substituting (4.6) and (4.7) in (4.3) we obtain the upper bound for $C_{u,k}(l)$.

References

1. Ango Nze P. and Doukhan P. functional estimation for time series: A general approach, prepublication de l'Universite Paris-Sudon, 43-93 (1993).
2. Antoniadis A., Gregoire G., and McKeague I. Wavelet methods for curve estimation. *Journal of the Americal Statistical Association*, **89**: 1340-1353 (1994).
3. Bradley R.C. Basic properties of strong mixing conditions, In: *Dependence in Probability and Statistics, A Survey of Resent Results*. (Oberwolfach,Birkhauser) (1985).
4. Barlow R.F. and Proschan F. *Statistical Theory of Reliability and Life Testing: Probability Models*. Holt Reinhart and Winston, New York (1975).
5. Chaubey Y.P. and Doosti H., and Prakasa Rao B.L.S. Wavelet based estimation of the derivatives of a density with associated variables. submitted (2005).
6. Chaubey Y.P., Doosti H., and Prakasa Rao B.L.S. Wavelet based estimation of the derivatives of a density for a negetively associated process. submitted (2005).
7. Chaubey Y.P. and Doosti H. Wavelet based estimation of the derivatives of a density for *m*-dependent random variables. *Journal of Iranian Statistical Society*, **4**(2) (2005).

8. Daubechies I. Ten Lectures on Wavelets. CBMS-NSF regional conferences series in applied mathematics, SIAM, Philadelphia (1992).
9. Daubechies I. Orthogonal bases of compactly supported wavelets. *Communication in Pure and Applied Mathematics*, **41**: 909-996 (1988).
10. Davydov Y.A. Mixing conditions for Markov chains. *Theory Probab. Appl.*, **18**: 312-328 (1973).
11. Delyon B. Limit theorems for mixing processes. Publication inter no. 546 I.R.I.S.A. (1990).
12. Donoho D.L., Johnstone I.M., Kerkyacharian G., and Picard D. Density estimation by wavelet thresholding. *The Annals of Statistics*, **2**: 508-539 (1996).
13. Donoho D.L., Johnstone I.M., Kerkyacharian G., and Picard D. Wavelet shrinkage: asymptopia (with discussion). *Journal of Royal Statistical Society*, Ser. B, **57**(2): 301-370 (1995).
14. Doosti H., Fakoor V., and Chaubey Y.P. "Wavelet linear density estimation for negative associated sequences." Preprint, Concordia University (2005).
15. Doosti H. and Nezakati A. Wavelet lnear density estimation for *m*-dependent random variables. Submitted (2005).
16. Doukhan P. and Leon J.R. Une note sur la deviation quadratique destimateurs de densites par projections orthogonales. *C.R. Acad. Sci.Paris*, t310, serie 1, 425-430 (1990).
17. Hardle W., Kerkyacharian G., Picard D., and Tsybakov A. Wavelets, Approximations, and Statistical Applications, Lecture Notes in Statistics. **129**, Springer-Verlag, Berlin (1998).
18. Ibragimov I.A. and Linnik Y.V. Independent and Stationary Sequences of Random Variables (Walter Nordhoof Pub., Groningen) (1971).
19. Kerkyacharian G. and Picard D. Density estimation in Besov spaces. *Statist. Probab. Lett.*, **13**: 15-24 (1992).
20. Kolmogorov A.N. and Rozanov Y.A. On the strong mixing conditions for stationary Gaussian sequences. *Theory Probab. Appl.*, **5**: 204-207 (1960).
21. Leblanc F. Wavelet linear density estimator for a discrete-time stochastic process: L_p-losses. *Statistics and Probability Letters*, **27**: 71-84 (1996).
22. Masry E. Probability density estimation from dependent observations using wavelets orthonormal bases. *Statist. Probab.Lett.*, **21**: 181-194 (1994).
23. Meyer Y. *Ondelettes et Operateurs*. Hermann, Paris (1990).
24. Prakasa Rao B.L.S. Nonparametric estimation of the drivatives of a density by the method of wavelets. *Bull. Inform. Cyb.*, **28**: 91-100 (1996).
25. Prakasa Rao B.L.S. Estimation of the integrated squared density derivative by wavelets. *Ibid.*, **31**: 47-65 (1999a).
26. Prakasa Rao B.L.S. Nonparametric functional estimation: An overview. In: *Asymptotics, Nonparametrics and Time Series*. Ghosh S. (Ed.), Marcel Dekker Inc., New York, p. 461-509 (1999b).
27. Prakasa Rao B.L.S. Wavelet linear density estimation for associated sequences. *Journal of the Indian Statistical Association*, **41**: 369-379 (2003).
28. Romano J.P. and Wolf M. A more general central limit theorem for m-dependent random variables with unbounded m. *Statistics and Probability Letters*, **47**: 115-124 (2000).
29. Shao Q.M. A comparison theorem on moment inequalities between negatively associated and independent random variables. *J. Theoretical. Probab.*, **13**: 343-356 (2000).
30. Tribouley K. Density estimation by cross-validation with wavelet method. *Statistica Neerlandica*, **45**: 41-62 (1995).
31. Triebel H. *Theory of Function Spaces II*, volume 84 of Monographs in mathematics. Birkhauser Verlag, Basel (1992).
32. Vidakovic B. *Statistical Modeling by Wavelets*. Wiley, New York (1999).
33. Walter G. and Ghorai J. Advantages and disadvantages of density estimation with wavelets. In: *Proceedings of the 24th Symp. on the Interface*, Joseph Newton H. (Ed.), Interface FNA, VA **24**: 234-343 (1992).