Alea , – ()

Convergence rates in weighted L_1 spaces of kernel density estimators for linear processes

Anton Schick and Wolfgang Wefelmeyer

Anton Schick, Department of Mathematical Sciences, Binghamton University, Binghamton, NY 13902-6000, USA
E-mail address: anton@math.binghamton.edu, http://math.binghamton.edu/anton/index.html

Wolfgang Wefelmeyer, Mathematisches Institut, Universität zu Köln, Weyertal 86–90, 50931 Köln, Germany
E-mail address: wefelm@math.uni-koeln.de, http://www.mi.uni-koeln.de/∼wefelm/

Abstract. Pointwise and uniform convergence rates for kernel estimators of the stationary density of a linear process have been obtained by several authors. Here we obtain rates in weighted L_1 spaces. In particular, if infinitely many coefficients of the process are non-zero and the innovation density has bounded variation, then nearly parametric rates are achievable by proper choice of kernel and bandwidth.

1. Introduction

Consider a linear process $X_t = \sum_{s=0}^{\infty} a_s \varepsilon_{t-s}$ with independent and identically distributed (i.i.d.) innovations ε_t that have finite mean and density f. We assume that $a_0 = 1$ and that the coefficients are summable, $\sum_{s=0}^{\infty} |a_s| < \infty$. Then X_t has a stationary density g. It can be estimated by the kernel estimator

$$\hat{g}(x) = \frac{1}{n} \sum_{j=1}^{n} k_b(x - X_j), \quad x \in \mathbb{R}.$$

Here $k_b(v) = k(v/b)/b$, where k is a kernel and b is a bandwidth such that $b \to 0$ and $nb \to \infty$. Pointwise and uniform convergence rates have been studied by several authors, see for example Hall and Hart (1990), Tran (1992), Hallin and Tran (1996), Lu (2001), Wu and Mielniczuk (2002), Bryk and Mielniczuk (2005), and Schick and Wefelmeyer (2006).

The natural distance for densities is given by the L_1-norm. Convergence of $\hat{g}$ to g in this norm has been neglected for time series. Here we study rates of convergence in weighted L_1-norms under mild assumptions on f. More specifically, we consider the weight function $V(x) = (1+|x|)^{\gamma}$ for some non-negative γ and the corresponding weighted L_1-norm $\|h\|_V = \int |h(x)| V(x)\, dx$. We refer to this norm as the V-norm. The choice $\gamma = 0$ gives the usual L_1-norm. The weighted version is needed if we estimate expectations $E[v(X)]$ by $\int v(x)\hat{g}(x)\, dx$ for functions v bounded by V, for

[1]Anton Schick was supported by NSF Grant DMS 0405791.
Received by the editors , accepted .
2000 *Mathematics Subject Classification.* 62G07, 62G20, 62M05.
Key words and phrases. L_1-Lipschitz, smoothness of convolutions, variance bound.

example moments and absolute moments. Convergence of density estimators in the V-norm was studied in Müller, Schick and Wefelmeyer (2005) and Schick and Wefelmeyer (2007a) and (2008). The results of the present paper play a key role in Schick and Wefelmeyer (2008).

Let m denote a positive integer. Decompose X_t as $Y_t + Z_t$ with

$$Y_t = \sum_{s=0}^{m-1} a_s \varepsilon_{t-s} \quad \text{and} \quad Z_t = \sum_{s=m}^{\infty} a_s \varepsilon_{t-s}$$

and write f_m for the density of Y_t. Express $\hat{g} - g$ as the sum $S + T + B$ of three terms, where

$$S(x) = \frac{1}{n} \sum_{j=1}^{n} \big(k_b(x - X_j) - k_b * f_m(x - Z_j) \big), \tag{1.1}$$

$$T(x) = \frac{1}{n} \sum_{j=1}^{n} k_b * f_m(x - Z_j) - k_b * g(x), \tag{1.2}$$

$$B(x) = k_b * g(x) - g(x), \quad x \in \mathbb{R}. \tag{1.3}$$

For $m = 1$ this approach was used by Wu and Mielniczuk (2002), and for arbitrary m by Schick and Wefelmeyer (2006). We study the V-norms of the terms in (1.1)–(1.3) individually. Let

$$N = \sum_{s \geq 1} \mathbf{1}[a_s \neq 0] \tag{1.4}$$

denote the number of nonzero coefficients among a_s, $s \geq 1$. If $N = 0$ we have i.i.d. observations $X_t = \varepsilon_t$. If N is finite, the observations are m-dependent for some m. In those cases, we can choose $T = 0$ by taking m large enough. Thus the term T has to be dealt with only if $N = \infty$.

Under mild conditions on f, g and k we obtain the rates

$$\|S\|_V = O_P(n^{-1/2} b^{-1/2}), \quad \|T\|_V = O_P(n^{-1/2}), \quad \|B\|_V = O(b^r)$$

for a positive integer r. This yields the familiar rate

$$\|\hat{g} - g\|_V = O_P(n^{-1/2} b^{-1/2}) + O(b^r) \tag{1.5}$$

under such conditions.

For the special case $V = 1$ we are dealing with the usual L_1-norm and have the following results. We take a bounded kernel of order $\varrho \geq 2$. We distinguish the cases when N is finite and when N is infinite.

(i) If N is finite and f has bounded variation and a finite moment of order greater than one, then (1.5) holds for $V = 1$ with $r = \min\{\varrho, N + 1\}$.

(ii) If N is infinite, the series $\sum_{s=1}^{\infty} s|a_s|$ converges, f has a finite moment of order greater than one, and the function $x \mapsto (1 + |x|)f(x)$ has bounded variation, then (1.5) holds for $V = 1$ with $r = \varrho$.

The paper is organized as follows. In Section 2 we state the results. Sections 3, 4 and 5 treat the terms S, B and T, respectively. Relations between various smoothness conditions for V-norms are studied in Section 6. An auxiliary result used in Section 5 is proved in Section 7. This result is of independent interest. Together with the generalization in Corollary 7.1 it is used in Schick and Wefelmeyer (2007b) and (2008).

2. Results

Let V be a measurable function satisfying

$$V(x+y) \le V(x)V(y), \quad x, y \in \mathbb{R}, \tag{2.1}$$

and

$$1 \le V(zx) \le V(x), \quad x \in \mathbb{R}, |z| \le 1. \tag{2.2}$$

Then the V-norm of a measurable function h is defined by

$$\|h\|_V = \int V(x)|h(x)|\,dx.$$

Let $v(x) = 1 + |x|$ and $W_\alpha = V^2 v^\alpha$. We are mainly interested in the case when V is a non-negative power of v, say $V = v^\gamma$. In this case $W_\alpha = v^{2\gamma+\alpha}$. The reason for restricting attention to this case is that we can rely on the moment inequality (3.2) below and can then give conditions in terms of finiteness of moments of f.

We now state some inequalities on the V-norm. An application of the Cauchy–Schwarz inequality yields

$$\|h\|_V^2 \le K_\alpha \|h^2\|_{W_\alpha} \tag{2.3}$$

for all $\alpha > 1$, with $K_\alpha = \int v^{-\alpha}(x)\,dx$. It follows from (2.1) that

$$\|h(\cdot - t)\|_V \le V(t)\|h\|_V, \quad t \in \mathbb{R}; \tag{2.4}$$

see Schick and Wefelmeyer (2007a). From this inequality we derive that

$$\int V(x)\Big|\int h(x-y)\mu(dy)\Big|\,dx \le \|h\|_V \int V\,d\mu \tag{2.5}$$

for every measure μ such that $\int V\,d\mu < \infty$, and every h with finite V-norm. In particular, the V-norm of a convolution $h_1 * h_2$ of two functions with finite V-norms satisfies the inequality

$$\|h_1 * h_2\|_V \le \|h_1\|_V \|h_2\|_V. \tag{2.6}$$

Since $(h_1 * h_2)^2 \le \|h_2\|_1 (h_1^2 * |h_2|)$ in view of the Cauchy–Schwarz inequality, we obtain from the last inequality that

$$\|(h_1 * h_2)^2\|_V \le \|h_1^2\|_V \|h_2\|_V \|h_2\|_1 \tag{2.7}$$

if h_1^2 and h_2 have finite V-norms.

To state our results we introduce the following definitions. These concepts and their relations are studied in Section 6.

Definition 2.1. A function h is *V-Lipschitz* (*with constant L*) if

$$\int V(x)|h(x-t) - h(x)|\,dx \le L|t|V(t), \quad t \in \mathbb{R}.$$

A function h is *V-Lipschitz of order r (with constant L)* for some positive integer r if there are functions $h^{(1)}, \dots, h^{(r-1)}$ such that

$$\int V(x)\Big|h(x+t) - h(x) - \sum_{i=1}^{r-1} \frac{t^i}{i!} h^{(i)}(x)\Big|\,dx \le L|t|^r V(t), \quad t \in \mathbb{R}.$$

If the functions $h^{(1)}, \dots, h^{(r-1)}$ also have finite V-norms, then h is *strongly V-Lipschitz of order r*.

Definition 2.2. A function h has *finite V-variation* if there are finite measures μ_1 and μ_2 satisfying $\mu_1(\mathbb{R}) = \mu_2(\mathbb{R})$ and $\int V\,d(\mu_1+\mu_2) < \infty$ such that $h(x) = \mu_1((-\infty,x]) - \mu_2((-\infty,x])$ for Lebesgue-almost-all x. In this case we call $\mu = \mu_1+\mu_2$ a *measure of V-variation* of h.

We need the following strengthened concept of a kernel of order ϱ. For $V = 1$ this definition reduces to the usual definition of a kernel of order ϱ if we also assume that $\int x^\varrho k(x)\,dx \neq 0$.

Definition 2.3. A kernel k is of *V-order ϱ* if ϱ is an integer greater than one,

$$\int x^i k(x)\,dx = 0, \quad i = 1, \dots, \varrho - 1,$$

and $\int (1+|x|)^\varrho V(x)|k(x)|\,dx$ is finite.

We have the following results. The first result treats the case of independent observations and is essentially contained in Müller, Schick and Wefelmeyer (2005) and in Schick and Wefelmeyer (2007a).

Theorem 2.1. *If $N = 0$, f is V-Lipschitz, f and k^2 have finite W_α-norms for some $\alpha > 1$, and $\int |t|V(t)|k(t)|\,dt$ is finite, then*

$$\|\hat{g} - g\|_V = O_P(n^{-1/2}b^{-1/2}) + O(b).$$

Proof: Since $N = 0$, we have $g = f$. If we take $m = 1$, we obtain $T = 0$ and $\hat{g} - g = S + B$. Hence Theorem 2.1 follows from Propositions 3.1 and 4.1.

Theorem 2.2. *Let N be positive and finite and let $V = v^\gamma$ for some $\gamma \geq 0$. If f has finite V-variation, f and k^2 have finite moments of order $\beta > 2\gamma+1$, and k is of V-order ϱ, then*

$$\|\hat{g} - g\|_V = O_P(n^{-1/2}b^{-1/2}) + O(b^r),$$

where r is the minimum of $N+1$ and ϱ.

Proof: Since N is finite, we can pick m so large that $T = 0$. It follows from Lemma 6.5 that g is strongly V-Lipschitz of order $N+1$. We have $v(ax) \leq v(a)v(x)$ and therefore $E[W_\alpha(a\varepsilon_0)] \leq W_\alpha(a)E[W_\alpha(\varepsilon_0)]$. Hence we derive Theorem 2.2 from Propositions 3.2 and 4.1.

Theorem 2.3. *Let N be infinite and let $\sum_{j=1}^\infty j|a_j|$ be finite. Set $V = v^\gamma$ for some $\gamma \geq 0$. Suppose that for some non-negative p and q with $p + q > 2\gamma + 1$ and some positive integer m, the density f has a finite moment of order $p + \max(q,1)$ and finite v^γ-variation, f_m is v^q-Lipschitz, and $v^p f_m$ is bounded. Suppose k is a bounded kernel of V-order r and $\int v^\beta(x)|k(x)|\,dx$ is finite for some $\beta > 2\gamma+1$. Then*

$$\|\hat{g} - g\|_V = O_P(n^{-1/2}b^{-1/2}) + O(b^r).$$

Proof: It follows from Proposition 3.3 that $\|S\|_V = O_P(n^{-1/2}b^{-1/2})$. The density g is V-Lipschitz of order r by Lemma 6.5. Hence we obtain $\|B\|_V = O(b^r)$ by Proposition 4.1. Finally, Proposition 5.1 yields $\|T\|_V = O_P(n^{-1/2})$.

By Lemma 6.1, the assumptions on f_m in Theorem 2.3 are met if they are met by f. By Lemma 6.2, the assumptions on f_m are met if f_m has finite v^s-variation, with $s \geq \max(q,p)$. Thus, the assumptions on f and f_m in Theorem 2.3 are met if f has a finite moment of order greater than $2\gamma+1$ and finite $v^{\gamma+1}$-variation. We formulate the corresponding result in the following corollary.

Corollary 2.1. *Let N be infinite and let $\sum_{j=1}^{\infty} j|a_j|$ be finite. Set $V = v^{\gamma}$ for some $\gamma \ge 0$. Let the density f have a finite moment of order $\beta > 2\gamma + 1$ and finite $v^{\gamma+1}$-variation. Suppose k is a bounded kernel of V-order r and $\int v^{\beta}(x)|k(x)|\,dx$ is finite for some $\beta > 2\gamma + 1$. Then*

$$\|\hat{g} - g\|_V = O_P(n^{-1/2}b^{-1/2}) + O(b^r).$$

Suppose we know that N is infinite. Under the assumptions of Theorem 2.3, we can control the rate $O(b^r)$ of the bias by choosing a kernel of high order r. A choice of bandwidth $b \sim n^{-1/(2r+1)}$ yields the rate

$$\|\hat{g} - g\|_V = O_P(n^{-r/(2r+1)}).$$

Thus we can achieve a rate close to the parametric rate $n^{-1/2}$. For invertible processes, even the parametric rate $n^{-1/2}$ can be achieved using the above results and constructing estimators that exploit the linear structure of the process; see Schick and Wefelmeyer (2007b) for the supremum norm and Schick and Wefelmeyer (2008) for the V-norm.

Note that if vf has bounded variation, then vf is bounded, f has bounded variation, and a simple argument shows that f is v-Lipschitz. Thus we derive from Theorem 2.3 the following result for the case $V = 1$.

Corollary 2.2. *Let N be infinite and let $\sum_{j=1}^{\infty} j|a_j|$ be finite. Suppose f has a finite moment of order greater than one, vf has bounded variation, and the kernel k is bounded and of order r. Then*

$$\|\hat{g} - g\|_1 = O_P(n^{-1/2}b^{-1/2}) + O(b^r).$$

3. Behavior of S

Let us first deal with the term S defined in (1.1). Since the i-th and j-th summands of S are uncorrelated if $|i - j| \ge m$, we obtain that $nE[S^2(x)] \le 2mE[k_b^2(x - X_1)]$. Using this and the inequalities (2.3) and (2.5), the latter with W_α in place of V, we find

$$\begin{aligned} nE[\|S\|_V^2] &\le K_\alpha \|nE[S^2]\|_{W_\alpha} \\ &\le 2mK_\alpha \int W_\alpha(x)E[k_b^2(x - X_1)]\,dx \\ &\le 2mK_\alpha E[W_\alpha(X_1)] \int W_\alpha(x)k_b^2(x)\,dx \\ &\le \frac{2m}{b} K_\alpha E[W_\alpha(X_1)] \int W_\alpha(bx)k^2(x)\,dx \end{aligned}$$

for all $\alpha > 1$. Thus we have the following result.

Proposition 3.1. *Suppose $E[W_\alpha(X_1)]$ and $\|k^2\|_{W_\alpha}$ are finite for some $\alpha > 1$. Then $\|S\|_V = O_P(n^{-1/2}b^{-1/2})$.*

Now consider a finite N. Since W_α inherits (2.1) and (2.2) from V, we obtain with the aid of these inequalities the bound $E[W_\alpha(X_1)] \le (E[W_\alpha(a\varepsilon_0)])^N$ with $a = \sup\{|a_s| : s \ge 0\}$. Hence we have the following consequence of Proposition 3.1.

Proposition 3.2. *Let N be finite and suppose that $E[W_\alpha(a\varepsilon_0)]$ and $\|k^2\|_{W_\alpha}$ are finite for some $\alpha > 1$. Then $\|S\|_V = O_P(n^{-1/2}b^{-1/2})$.*

Now consider the case that V is a non-negative power of v. Then W_α is also a power of v. We have

$$(1+|x|)^r \le 2^{r-1}(1+|x|^r), \quad x \in \mathbb{R}, r \ge 1. \tag{3.1}$$

This and the Minkowski inequality give

$$E[v^r(X_1)] \le 2^{r-1}\Big(1 + \Big(\sum_{s=0}^{\infty} |a_s|\Big)^r E[|\varepsilon_0|^r]\Big), \quad r \ge 1. \tag{3.2}$$

Thus we have the following result.

Proposition 3.3. *Suppose* $E[|\varepsilon_0|^\beta]$ *and* $\int v^\beta(x)k^2(x)\,dx$ *are finite for some* $\beta > 2\gamma + 1$ *with* $\gamma \ge 0$. *Then* $\|S\|_V = O_P(n^{-1/2}b^{-1/2})$ *for* $V = v^\gamma$.

4. Behavior of the bias

Next we deal with the bias term B defined in (1.3). For this we shall use the following lemma.

Lemma 4.1. *Suppose* $h, h_1, \dots, h_{r-1}$, w *and* U *are measurable functions such that*

$$\int V(x)\Big|h(x+t) - h(x) - \sum_{i=1}^{r-1} \frac{t^i}{i!} h_i(x)\Big|\,dx \le U(t), \quad t \in \mathbb{R},$$

and $c_i = \int t^i w(t)\,dt$, $i = 0, \dots, r-1$, *and* $A = \int U(-t)|w(t)|\,dt$ *are finite. Then*

$$\Big\|h * w - \sum_{i=0}^{r-1} \frac{(-1)^i c_i}{i!} h_i\Big\|_V \le A. \tag{4.1}$$

Proof: Let Δ denote the left-hand side of (4.1). Then

$$\begin{aligned}\Delta &= \int V(x)\Big|\int \Big(h(x-t) - h(x) - \sum_{i=1}^{r-1} \frac{(-t)^i}{i!} h_i(x)\Big) w(t)\,dt\Big|\,dx \\ &\le \iint V(x)\Big|h(x-t) - h(x) - \sum_{i=1}^{r-1} \frac{(-t)^i}{i!} h_i(x)\Big|\,dx\,|w(t)|\,dt\end{aligned}$$

and hence $\Delta \le A$.

Proposition 4.1. *Suppose* g *is* V*-Lipschitz of order* r *and the kernel* k *is of* V*-order* r. *Then* $\|B\|_V = O(b^r)$.

Proof: This follows from Lemma 4.1 applied with $w = k_b$ and $U(t) = L|t|^r V(t)$. Note that $c_0 = 1$ and $c_i = 0$ for $i = 1, \dots, r-1$ and

$$\int U(-t)|k_b(t)|\,dt = L\int |bt|^r V(bt)|k(t)|\,dt \le Lb^r \int |t|^r V(t)|k(t)|\,dt$$

for $b \le 1$.

Sufficient conditions for g to be V-Lipschitz of order $r \ge 2$ are given in Section 6.

5. Behavior of T

Finally we consider the term T introduced in (1.2). As shown in the Introduction, we need to treat only the case when N is infinite.

Proposition 5.1. *Let $V = v^\gamma$ for some $\gamma \geq 0$. Let p and q be non-negative numbers with $p + q > 2\gamma + 1$. Set $\beta = p + \max(1, q)$ and $r = \max(p, q)$. Suppose N is infinite, $\sum_{j=1}^\infty j|a_j|$ is finite, f has finite v^β-norm, $v^p f_m$ is bounded, f_m is v^q-Lipschitz, and $v^r k$ is integrable. Then $\|T\|_V = O_P(n^{-1/2})$.*

Proof: In view of (2.3) it suffices to show that $\|E[nT^2]\|_{v^{p+q}}$ is bounded. For this we apply Lemma 7.1 below with $h = f_m * k_b$, $c_j = a_{j+i}$ and $U_j = \varepsilon_{j-i}$, with $i = \inf\{j \geq m : a_j \neq 0\}$. Since f has finite v^β-norm, U_0 has finite moment of order β. Moreover, f_m has finite v^q-norm and $\|f_m * k_b\|_{v^q} \leq \|f_m\|_{v^q}\|k_b\|_{v^q}$. Since f_m is v^q-Lipschitz with constant L and $v^p f_m$ is bounded by C, say, we obtain from Remark 6.1 below that $f_m * k_b$ is v^q-Lipschitz with constant $L\|k_b\|_{v^q}$ and $v^p f_m * k_b$ is bounded by $C\|k_b\|_{v^p}$. Note that $\|k_b\|_{v^s} \leq \|k_b\|_{v^r} \leq v^r(b)\|k\|_{v^r}$ for $s \leq r$. Our assumptions on the coefficients a_j guarantee that D of Lemma 7.1 is finite. From this lemma we obtain that $\|E[nT^2]\|_{v^{p+q}} = O(1)$, which is the desired result.

By Lemma 6.2, if f_m has finite v^q-variation, then $v^q f_m$ is bounded and f_m is v^q-Lipschitz. Thus, taking $\gamma < p < \gamma + 1 = q$, we arrive at the following result.

Corollary 5.1. *Let $V = v^\gamma$ for some $\gamma \geq 0$, and let $\beta > 2\gamma + 1$. Suppose N is infinite, $\sum_{j=1}^\infty j|a_j|$ is finite, f has finite v^β-norm, f_m has finite $v^{\gamma+1}$-variation, and $v^{\gamma+1}k$ is integrable. Then $\|T\|_V = O_P(n^{-1/2})$.*

6. Smoothness in the V-norm

Here we study finite V-variation and the V-Lipschitz property and their relations. Our first lemma shows that these properties are preserved under convolutions with a measure ν for which $\int V d\nu$ is finite.

Lemma 6.1. *Let ν be a measure with $\int V d\nu$ finite. Let h be a function for which*

$$h_*(x) = \int h(x - y)\nu(dy), \quad x \in \mathbb{R},$$

is well-defined. Then the following are true.

1. *If h has finite V-norm, then the V-norm of h_* is bounded by $\|h\|_V \int V\, d\nu$.*
2. *If h is V-Lipschitz with constant L, then h_* is V-Lipschitz with constant $L \int V\, d\nu$.*
3. *If h is strongly V-Lipschitz of order r with constant L, then h_* is strongly V-Lipschitz of order r with constant $L \int V\, d\nu$.*
4. *If h has finite V-variation with measure of variation μ, then h_* has finite V-variation with measure of variation $\mu * \nu$.*
5. *If Vh is bounded by C, then Vh_* is bounded by $C \int V\, d\nu$.*

Proof: Conclusion (1) is a consequence of (2.5). Conclusion (2) follows from the bound

$$\begin{aligned}\int V(x)|h_*(x - t) - h_*(x)|\, dx &\leq \iint V(x)|h(x - t - y) - h(x - y)|\, dx\, \nu(dy) \\ &\leq \int V(x)|h(x - t) - h(x)|\, dx \int V(y)\, \nu(dy).\end{aligned}$$

To verify (3), take $h_*^{(i)}(x) = \int h^{(i)}(x-y)\nu(dy)$. Then, by (2.5), the functions $h_*^{(1)}, \dots, h_*^{(r-1)}$ have finite V-norms, and

$$\int V(x)\Big|h_*(x+t) - h_*(x) - \sum_{i=1}^{r-1} \frac{t^i}{i!} h_*^{(i)}(x)\Big|\,dx$$
$$\leq \int V\,d\nu \int V(x)\Big|h(x+t) - h(x) - \sum_{i=1}^{r-1} \frac{t^i}{i!} h^{(i)}(x)\Big|\,dx,$$

and (3) follows. To verify (4) we may assume that $h(x) = \mu_1((-\infty,x]) - \mu_2((-\infty,x])$ for all $x \in \mathbb{R}$, where μ_1 and μ_2 are finite measures with $\mu_1(\mathbb{R}) = \mu_2(\mathbb{R})$ and $\int V\,d\mu_1$ and $\int V\,d\mu_2$ finite. We now derive $h_*(x) = (\mu_1 - \mu_2) * \nu((-\infty,x])$ and hence (4). Finally, (5) follows from the bound $|V(x)h_*(x)| \leq \int V(y)V(x-y)|h(x-y)|\nu(dy) \leq C\int V\,d\nu$.

Remark 6.1. Let h and u be measurable functions with $h*u$ well-defined and $\|u\|_V$ finite. Then the conclusions of Lemma 6.1 hold with h_* replaced by $h*u$ and $\nu(dx) = |u(x)|\,dx$ so that $\int V\,d\nu$ becomes $\|u\|_V$. To see this, write $u = u_+ - u_-$, where u_+ and u_- are the positive and negative part of u, and apply Lemma 6.1 with $\nu(dx) = u_+(x)\,dx$ and $\nu(dx) = u_-(x)\,dx$.

Remark 6.2. An integrable function of bounded variation has finite 1-variation. Hence densities of bounded variation are 1-Lipschitz. Moreover, an integrable absolutely continuous function h with $\|h'\|_V$ finite has finite V-variation (with μ_1 having density $h'_+ = \max(h',0)$ and μ_2 having density $h'_- = \max(-h',0)$).

The next lemma gives consequences of finite V-variation.

Lemma 6.2. *If h has finite V-variation, then Vh is bounded by $\int V\,d\mu$ and h is V-Lipschitz with constant $\int V\,d\mu$, where μ is a measure of V-variation of h. If h has finite vV-variation, then h has finite V-norm $\|h\|_V \leq \int vV\,d\mu$.*

Proof: We may assume that $h(x) = \mu_1((-\infty,x]) - \mu_2((-\infty,x])$ for all $x \in \mathbb{R}$, where μ_1 and μ_2 are finite measures with $\mu_1(\mathbb{R}) = \mu_2(\mathbb{R})$. Then we have $h(x) = \mu_2((x,\infty)) - \mu_1((x,\infty))$ for all x. By (2.2), we have $V(x) \leq V(y)$ for $|x| \leq |y|$. Then, with $\mu = \mu_1 + \mu_2$, we obtain for $x \geq 0$ the inequalities

$$V(x)|h(x)| \leq V(x)\int_{x\leq y} \mu(dy) \leq \int_{x \leq y} V(y)\mu(dy)$$

and

$$\int_0^\infty V(x)|h(x)|\,dx \leq \iint_{0\leq x\leq y} V(y)\mu(dy)\,dx \leq \int_0^\infty yV(y)\mu(dy).$$

For $x \leq 0$, we obtain the inequalities

$$V(x)|h(x)| \leq V(x)\int_{y\leq x} \mu(dy) \leq \int_{y \leq x} V(y)\mu(dy)$$

and

$$\int_{-\infty}^0 V(x)|h(x)|\,dx \leq \int_{-\infty}^0 |y|V(y)\mu(dy).$$

Thus Vh is bounded by $\int V\,d\mu$, and $\|h\|_V \leq \int vV\,d\mu$. The arguments in the proof of Lemma 8 of Schick and Wefelmeyer (2007a) show that a function with finite V-variation is V-Lipschitz with constant $\int V\,d\mu$.

We now give sufficient conditions for a function to be V-Lipschitz of order $r \geq 2$. For this we make the following definitions.

Definition 6.1. A function h is *absolutely continuous of order* r if h is $(r-1)$-times differentiable and if its $(r-1)$-th derivative $h^{(r-1)}$ is absolutely continuous with almost everywhere derivative $h^{(r)}$.

Definition 6.2. A function h is *V-regular of order r* if h is absolutely continuous of order $r-1$ and $h^{(r-1)}$ is V-Lipschitz. If also $h^{(1)}, \dots, h^{(r-1)}$ have finite V-norms, we call h *strongly V-regular of order r*.

Lemma 6.3. *If h is (strongly) V-regular of order $r \geq 2$, then h is (strongly) V-Lipschitz of the same order.*

Proof: Let

$$\Delta(x,t) = h(x+t) - h(x) - \sum_{i=1}^{r-1} \frac{t^i}{i!} h^{(i)}(x).$$

We have

$$\Delta(x,t) = t^{r-1} \int_0^1 \frac{(1-u)^{r-2}}{(r-2)!} \big(h^{(r-1)}(x+ut) - h^{(r-1)}(x)\big)\, du.$$

Since $h^{(r-1)}$ is V-Lipschitz and $V(ut) \leq V(t)$ for $|u| \leq 1$,

$$\int V(x)|\Delta(x,t)|\, dx \leq L|t|^r V(t) \int_0^1 \frac{(1-u)^{r-2}}{(r-2)!}\, du. \tag{6.1}$$

The desired results are now immediate.

Remark 6.3. In the case of strong V-regularity, the bound (6.1) can be replaced by

$$\int V(x)|\Delta(x,t)|\, dx \leq 2\Lambda \frac{|t|^r}{1+|t|} V(t) \int_0^1 \frac{(1-u)^{r-2}}{(r-2)!}\, du,$$

with $\Lambda = \max(L, 2\|h^{(r-1)}\|_V)$. This follows from the fact that the map $x \mapsto x/(1+x)$ is increasing on the interval $(0,\infty)$ and the following lemma. This alternative bound is better if $|t|$ is large.

Lemma 6.4. *If h has finite V-norm and is V-Lipschitz with constant L, then*

$$\|h(\cdot - t) - h\|_V \leq 2\Lambda \frac{|t|}{1+|t|} V(t), \quad t \in \mathbb{R},$$

where $\Lambda = \max(L, 2\|h\|_V)$.

Proof: The statement is clear if $|t| \leq 1$, and it follows from the bound

$$\|h(\cdot - t) - h\|_V \leq (V(t)+1)\|h\|_V \leq 2V(t)\|h\|_V$$

for $|t| > 1$.

Sufficient condition for V-regularity of g are given next.

Lemma 6.5. *Suppose f has finite V-norm and finite V-variation for $V = v^\gamma$ with $\gamma \geq 0$. Let $N \geq p$. Then g is strongly V-regular of order $p+1$ and hence strongly V-Lipschitz of that order.*

Proof: Let $\alpha_1, \dots, \alpha_p$ denote the first p nonzero numbers among a_s, $s \geq 1$. Let g_p denote the density of $\varepsilon_0 + \sum_{i=1}^p \alpha_i \varepsilon_i$. Then $g(x) = E[g_p(x - Z)]$, $x \in \mathbb{R}$, for some random variable Z with $E[V(Z)] < \infty$. Thus by (3) of Lemma 6.1 it suffices to show that g_p is strongly V-regular of order $p+1$. This is true for $p = 0$. In this case, g_0 equals f, and the latter is V-Lipschitz with constant $\int V\, d\mu$ by Lemma 6.2. The desired result now follows by induction using the following lemma. Keep in mind that the density of $a\varepsilon_0$ inherits the properties of the density f for non-zero a.

Lemma 6.6. *Suppose the functions h_1 and h_2 have finite V-norms, h_1 is V-Lipschitz with constant L and h_2 has finite V-variation. Then $h = h_1 * h_2$ is absolutely continuous and h' has finite V-norm $\|h'\|_V \leq \|h_1\|_V \int V\, d\mu$ and is V-Lipschitz with constant $L \int V\, d\mu$, where μ is a measure of V-variation. Hence h is strongly V-Lipschitz of order 2.*

Proof: We may assume that $h_2(x) = \mu_1((-\infty, x]) - \mu_2((-\infty, x])$ for all x, where μ_1 and μ_2 are measures such that $\int V\, d(\mu_1 + \mu_2)$ is finite. For $i = 1, 2$, set $q_i(x) = \int h_1(x-y)\, \mu_i(dy)$, $x \in \mathbb{R}$. By Lemma 6.1, q_i has finite V-norm and q_i is V-Lipschitz. Since $q_1 - q_2$ is an almost everywhere derivative of h, as shown in Lemma 1 of Schick and Wefelmeyer (2006), we obtain the desired result.

7. A bound

Consider a linear process

$$S_t = \sum_{s=0}^{\infty} c_s U_{t-s}, \quad t \in \mathbb{Z},$$

with independent and identically distributed innovations U_t, $t \in \mathbb{Z}$, with finite mean and summable coefficients $c_0, c_1, \dots$ with $c_0 \neq 0$. For a bounded measurable function h, set

$$H(x) = n^{-1/2} \sum_{j=1}^{n} \big(h(x - S_j) - E[h(x - S_j)] \big), \quad x \in \mathbb{R}.$$

In this section we derive bounds for $\int v^r(x) E[H^2(x)]\, dx$ with $r \geq 0$. For this we set

$$\|c\| = \sum_{j=0}^{\infty} |c_j| \quad \text{and} \quad D = \sum_{j=0}^{\infty} (j+1)|c_j|.$$

To simplify notation, we abbreviate U_0 by U. Also, let us set

$$A(\alpha, \beta) = 2^{\beta-1}(1 + \alpha^{\beta} E[|U|^{\beta}]), \quad \alpha \geq 0, \beta \geq 1.$$

Lemma 7.1. *Let p and q be non-negative and $q_* = \max(q, 1)$. Suppose h has finite v^q-norm and is v^q-Lipschitz with constant L, $v^p h$ is bounded, and U has a finite moment of order $p + q_*$. Let D be finite. Then*

$$\int v^{p+q}(x) E[H^2(x)]\, dx \leq 8\Lambda \|v^p h\|_\infty D A^4,$$

where $\Lambda = \max(L, 2\|h\|_{v^q})$ and $A = A\big(\max(1, 2\|c\|), p + q_*\big)$.

Proof: Let $\beta = p + q_*$. For $j = 0, 1, \dots$, let

$$Q_j = \sum_{s=0}^{\infty} |c_s U_{j-s}|, \quad T_j = \sum_{s=0}^{j-1} c_s U_{j-s}, \quad R_j = S_j - T_j = \sum_{s=j}^{\infty} c_s U_{j-s},$$

and $h_j(x) = E[h(x - T_j)]$, $x \in \mathbb{R}$. Note that $T_0 = 0$, $R_0 = S_0$ and $h_0 = h$. The absolute values of T_j and R_j are bounded by Q_j so that for non-negative t and all $j = 0, 1, \dots,$

$$E[v^t(T_j)] \le E[v^t(Q_j)] \quad \text{and} \quad E[v^t(R_j)] \le E[v^t(Q_j)]. \tag{7.1}$$

Let $Q = Q_0$. The argument leading to (3.2) yields that, for every $t \in [0, \beta]$ and every $j = 0, 1, \dots,$

$$E[v^t(Q)] \le E[v^\beta(Q)] \le A \quad \text{and} \quad E[v^t(Q + Q_j)] \le E[v^\beta(Q + Q_j)] \le A. \tag{7.2}$$

Using stationarity and a conditioning argument, we obtain

$$\begin{aligned} E[H^2(x)] &= \operatorname{Var}(h(x - S_0)) + \frac{2}{n}\sum_{j=1}^{n-1}(n-j)\operatorname{Cov}(h(x - S_0), h(x - S_j)) \\ &= \operatorname{Var}(h_0(x - R_0)) + \frac{2}{n}\sum_{j=1}^{n-1}(n-j)\operatorname{Cov}(h_0(x - R_0), h_j(x - R_j)). \end{aligned}$$

Thus

$$\int v^{p+q}(x)E[H^2(x)]\,dx \le 2\sum_{j=0}^{\infty}\Gamma_j$$

where

$$\Gamma_j = \int v^{p+q}(x)E[|h(x - R_0) - E[h(x - R_0)]||h_j(x - R_j) - h_j(x)|]\,dx.$$

Since $v^p h$ is bounded and $v(x + y) \le v(x)v(y)$, we derive the bound

$$v^p(x)|h(x - R_0)| \le v^p(x - R_0)v^p(R_0)|h(x - R_0)| \le \|v^p h\|_\infty v^p(Q)$$

which implies

$$v^p(x)|E[h(x - R_0)]| \le \|v^p h\|_\infty E[v^p(Q)] \le \|v^p h\|_\infty A.$$

Using these bounds and $v \ge 1$ and $A \ge 1$, we obtain for $j \ge 0$,

$$\Gamma_j \le 2A\|v^p h\|_\infty E\Big[v^p(Q)\int v^q(x)|h_j(x - R_j) - h_j(x)|\,dx\Big].$$

Note that $\|h_j\|_{v^q} \le \|h\|_{v^q}E[v^q(T_j)]$ and h_j is v^q-Lipschitz with constant $L_j = LE[v^q(T_j)]$. Thus, by Lemma 6.4 and the inequalities (7.1) and (7.2), we obtain the bound

$$\int v^q(x)|h_j(x - R_j) - h_j(x)|\,dx \le 2\Lambda A\big(v^{q_*-1}(Q_j)|R_j|\big).$$

Since $v^s(x)v^t(y) \le (v(x + y))^{s+t}$ for non-negative s, t, x, y, the above shows that

$$\Gamma_j \le 4\Lambda A^2\|v^p h\|_\infty E[v^{\beta-1}(Q + Q_j)|R_j|].$$

Using $v(x+y) \le v(x)v(y)$ and the independence of U_{-i} and $Q_{j,i} = Q + Q_j - |c_i U_{-i}| - |c_{j+i} U_{-i}|$ for $i \ge 0$, we obtain

$$\begin{aligned} E[v^{\beta-1}(Q+Q_j)|R_j|] &\le E\Big[\sum_{s=j}^{\infty} |c_s U_{j-s}| v^{\beta-1}(Q+Q_j)\Big] \\ &\le \sum_{s=j}^{\infty} |c_s| E[|U_{j-s}| v^{\beta-1}((|c_{s-j}| + |c_s|)U_{j-s})] E[v^{\beta-1}(Q_{j,s-j})] \\ &\le \sum_{s=j}^{\infty} |c_s| E[v^{\beta}(\alpha U)] E[v^{\beta}(Q+Q_j)], \end{aligned}$$

with $\alpha = \max(1, 2\|c\|)$. We have $E[v^{\beta}(\alpha U)] \le A$ by (3.1). Using inequality (7.2) again, we obtain the bound

$$\Gamma_j \le 4\Lambda \|v^p h\|_\infty A^4 \sum_{s=j}^{\infty} |c_s|, \quad j \ge 0.$$

Note also that

$$\sum_{j=0}^{\infty} \sum_{s=j}^{\infty} |c_s| = \sum_{s=0}^{\infty} (1+s)|c_s| = D.$$

The desired result is now immediate.

Using the inequality $\|h_j(\cdot - t) - h_j\|_{v^q} \le AL|t|v^q(t)$ instead of the inequality provided by Lemma 6.4, we can avoid the assumption that h has finite v^q-norm at the price of (possibly) increasing the moment condition from $p + q_*$ to $p + q + 1$. More precisely, we have the following result.

Lemma 7.2. *Let p and q be non-negative. Suppose h is v^q-Lipschitz with constant L, $v^p h$ is bounded, and U has a finite moment of order $\beta = p + q + 1$. Let D be finite. Then*

$$\int v^{p+q}(x) E[H^2(x)]\, dx \le 4LA^4 \|v^p h\|_\infty D,$$

where now $A = A(\max(1, 2\|c\|), p + q + 1)$.

Repeating the above proof with $v^p = v^q = 1$, we obtain the following result.

Lemma 7.3. *Suppose h is bounded and 1-Lipschitz with constant L. Let D be finite. Then $E[H^2(x)] \le 4\|h\|_\infty \Gamma(x)$ for all x, and $\|\Gamma\|_1 \le DE[|U|]$, where*

$$\Gamma(x) = \sum_{j=0}^{\infty} E[|h(x - R_j) - h(x)|], \quad x \in \mathbb{R}.$$

Consequently,

$$\int E[H^2(x)]\, dx \le 4L\|h\|_\infty DE[|U|].$$

If we take $h(x) = 1[0 \le x]$, then H becomes the empirical process

$$D_n(x) = n^{-1/2} \sum_{j=1}^{n} (\mathbf{1}[S_t \le x] - P(S_t \le x)), \quad x \in \mathbb{R}.$$

This choice of h is bounded by 1 and 1-Lipschitz with constant $L = 1$. Thus we have the following result.

Corollary 7.1. *Let D be finite. Then there exists an integrable function Ψ with $\|\Psi\|_1 \le 4DE[|U|]$ such that $E[D_n^2(x)] \le \Psi(x)$ for all n and x.*

Acknowledgment. We thank two referees whose suggestions led us to rewrite the paper completely.

References

[1] A. Bryk and J. Mielniczuk. Asymptotic properties of density estimates for linear processes: application of projection method. *J. Nonparametr. Stat.* **17**, 121–133 (2005).

[2] P. Hall and J. D. Hart. Convergence rates in density estimation for data from infinite-order moving average processes. *Probab. Theory Related Fields* **87**, 253–274 (1990).

[3] M. Hallin and L. T. Tran. Kernel density estimation for linear processes: Asymptotic normality and optimal bandwidth derivation. *Ann. Inst. Statist. Math.* **48**, 429–449 (1996).

[4] Z. Lu. Asymptotic normality of kernel density estimators under dependence. *Ann. Inst. Statist. Math.* **53**, 447–468 (2001).

[5] U. U. Müller, A. Schick and W. Wefelmeyer. Weighted residual-based density estimators for nonlinear autoregressive models. *Statist. Sinica* **15**, 177–195 (2005).

[6] A. Schick and W. Wefelmeyer. Pointwise convergence rates and central limit theorems for kernel density estimators in linear processes. *Statist. Probab. Lett.* **76**, 1756–1760 (2006).

[7] A. Schick and W. Wefelmeyer. Root-n consistent density estimators of convolutions in weighted L_1-norms. *J. Statist. Plann. Inference* **137**, 1765–1774 (2007a).

[8] A. Schick and W. Wefelmeyer. Uniformly root-n consistent density estimators for weakly dependent invertible linear processes. *Ann. Statist.* **35**, 815–84 (2007b).

[9] A. Schick and W. Wefelmeyer. Root-n consistency in weighted L_1-spaces for density estimators of invertible linear processes. To appear in: *Stat. Inference Stoch. Process.* (2008).

[10] L. T. Tran. Kernel density estimation for linear processes. *Stochastic Process. Appl.* **41**, 281–296 (1992).

[11] W. B. Wu and J. Mielniczuk. Kernel density estimation for linear processes. *Ann. Statist.* **30**, 1441–1459 (2002).