ROOT-n CONSISTENCY IN WEIGHTED L_1-SPACES FOR DENSITY ESTIMATORS OF INVERTIBLE LINEAR PROCESSES

ANTON SCHICK AND WOLFGANG WEFELMEYER

Abstract. The stationary density of an invertible linear processes can be estimated at the parametric rate by a convolution of residual-based kernel estimators. We have shown elsewhere that the convergence is uniform and that a functional central limit theorem holds in the space of continuous functions vanishing at infinity. Here we show that analogous results hold in weighted L_1-spaces. We do not require smoothness of the innovation density.

AMS 2000 subject classification. Primary: 62G07, 62G20, 62M05, 62M10.

Key words and Phrases. Kernel estimator, plug-in estimator, tightness criteria, functional limit theorem, infinite-order moving average process, infinite-order autoregressive process.

1. Introduction.

Kernel estimators for the stationary density of linear processes are well-studied; see Chanda (1983), Hall and Hart (1990), Tran (1992), Hallin and Tran (1996), Coulon-Prieur and Doukhan (2000), Honda (2000), Lu (2001), Hallin, Lu and Tran (2001), Wu and Mielniczuk (2002), Bryk and Mielniczuk (2005) and Schick and Wefelmeyer (2006).

Kernel estimators are nonparametric estimators that do not use the structure of the underlying model. Sometimes the structure of the model can be exploited to construct estimators that converge at faster and even parametric rates. This was observed by Frees (1994) when estimating densities of certain functions $q(X_1, \dots, X_m)$ on the basis of independent observations $X_1, \dots, X_n$. Saavedra and Cao (2000) consider the special case $q(X_1, X_2) = X_1 + aX_2$. Schick and Wefelmeyer (2004b, 2007a) prove functional convergence for $q(X_1, \dots, X_m) = u_1(X_1) + \dots + u_m(X_m)$ and $q(X_1, X_2) = X_1 + X_2$, viewing their estimators as elements of L_1 or of the space C_0 of continuous functions on $\mathbb{R}$ vanishing at infinity. Giné and Mason (2007a, 2007b) obtain functional results and laws of the iterated logarithm in L_p, $1 \le p \le \infty$, and locally uniformly in the bandwidth, for general $q(X_1, \dots, X_m)$. Du and Schick (2007) obtain functional results in C_0 and L_p for estimators of derivatives of convolutions.

Special cases of the semiparametric time series model considered here have also been studied. Saavedra and Cao (1999) consider pointwise convergence of plug-in estimators for the stationary density of moving average processes of order one. Schick and Wefelmeyer

Anton Schick was supported in part by NSF Grant DMS 0405791.

(2004a) obtain asymptotic normality and efficiency, and Schick and Wefelmeyer (2004c) generalize this result to higher order moving average processes and to functional convergence in L_1 and C_0.

For general invertible linear processes, Schick and Wefelmeyer (2007b) construct $n^{1/2}$-consistent estimators and prove functional convergence in C_0. Here we obtain an analogous result in weighted L_1-spaces. We denote by L_V the space of functions with finite V-norm $\|a\|_V = \int V(x)|a(x)|\,dx$. Our main result is formulated for $V(x) = V_r(x) = (1 + |x|)^r$ for some non-negative r. Functional results for density estimators in L_V are useful if we want to estimate expectations under the stationary law of functions dominated by V by plugging in our density estimator, like moments and absolute moments. The choice $V = 1$ corresponds to the natural distance between densities.

As in Schick and Wefelmeyer (2007b), we consider a stationary linear process with infinite-order moving average representation

$$X_t = \varepsilon_t + \sum_{s=1}^{\infty} \varphi_s \varepsilon_{t-s}, \quad t \in \mathbb{Z}, \tag{1.1}$$

with summable coefficients, $\sum_{s=1}^{\infty} |\varphi_s| < \infty$, and i.i.d. innovations ε_t, $t \in \mathbb{Z}$, that have mean zero and finite variance. Suppose the innovations have a density f. Then X_0 has a density, say h. The usual estimator of this density from observations $X_1, \dots, X_n$ of the linear process is a kernel density estimator $\tilde{h}(x) = (1/n)\sum_{j=1}^{n} k_{b_n}(x - X_j)$, where b_n is a sequence of bandwidths and $k_b = k(x/b)/b$ for some kernel k and some $b > 0$. In order to construct a $n^{1/2}$-consistent estimator of h, we follow Schick and Wefelmeyer (2007b) and set

$$Y_t = X_t - \varepsilon_t = \sum_{s=1}^{\infty} \varphi_s \varepsilon_{t-s}, \quad t \in \mathbb{Z}.$$

We must assume that the representation $X_0 = \varepsilon_0 + Y_0$ is nondegenerate:

(C) *At least one of the moving average coefficients φ_s is nonzero.*

Then Y_0 has a density, say g. Since Y_0 is independent of ε_0, we can express the density h of X_0 as the convolution $h = f * g$ of f and g. We obtain an estimator of h as $\hat{h} = \hat{f} * \hat{g}$, where $\hat{f}$ and $\hat{g}$ are estimators of f and g. We base $\hat{f}$ and $\hat{g}$ on estimators of the innovations. For this we require invertibility of the process.

(I) *The function $\phi(z) = 1 + \sum_{s=1}^{\infty} \varphi_s z^s$ is bounded and bounded away from zero on the complex unit disk $D = \{z \in \mathbb{C} : |z| \le 1\}$.*

Then $\rho(z) = 1/\phi(z) = 1 - \sum_{s=1}^{\infty} \varrho_s z^s$ is also bounded and bounded away from zero on D. Hence the innovations have the infinite-order autoregressive representation

$$\varepsilon_t = X_t - \sum_{s=1}^{\infty} \varrho_s X_{t-s}, \quad t \in \mathbb{Z}. \tag{1.2}$$

Let p_n be positive integers with $p_n/n \to 0$. For $j = p_n+1, \dots, n$ we mimic the innovation ε_j by the residual

$$\hat{\varepsilon}_j = X_j - \sum_{i=1}^{p_n} \hat{\varrho}_i X_{j-i},$$

where $\hat{\varrho}_i$ is an estimator of ϱ_i for $i = 1, \dots, p_n$. We then estimate the innovation density by a kernel estimator based on the residuals,

$$\hat{f}(x) = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} k_{b_n}(x - \hat{\varepsilon}_j), \quad x \in \mathbb{R},$$

and we estimate the density g by a kernel estimator based on the differences $\hat{Y}_j = X_j - \hat{\varepsilon}_j$,

$$\hat{g}(x) = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} k_{b_n}(x - \hat{Y}_j), \quad x \in \mathbb{R}.$$

Our estimator is then $\hat{h} = \hat{f} * \hat{g}$. This estimator can be written as a V-statistic, see Schick and Wefelmeyer (2007b), and is therefore easy to calculate.

In addition to (C) and (I) we use the following assumptions.

(Q) *The autoregression coefficients fulfill* $\sum_{s>p_n} |\varrho_s| = O(n^{-1/2-\zeta})$ *for some* $\zeta > 0$.

If the autoregression coefficients are known to decay exponentially, condition (C) holds if $p_n/\log n$ tends to infinity. If the coefficients are known to decay polynomially, $|\varrho_s| = O(\varrho^{-\beta-1})$, then (Q) holds with $\zeta = \gamma\beta - 1/2$ if p_n is proportional to n^γ and $\gamma > 1/(2\beta)$. It would be interesting to find a data-driven selection of p_n.

(R) *The estimators* $\hat{\varrho}_i$ *of the autoregression coefficients* ϱ_i *fulfill*

$$\sum_{i=1}^{p_n} (\hat{\varrho}_i - \varrho_i)^2 = O_p(q_n n^{-1})$$

for some q_n *with* $1 \le q_n \le p_n$.

(S+) *The moving average coefficients satisfy* $\sum_{s=1}^{\infty} s^\beta |\varphi_s| < \infty$ *for some* $\beta > 1$.

If f also has a finite fourth moment and $np_n \sum_{s>p_n} \varrho_s^2 \to 0$ holds, then condition (R) with $q_n = p_n$ is met by the least squares estimators $\hat{\varrho}_1, \dots, \hat{\varrho}_{p_n}$ which minimize $\sum_{j=p_n+1}^{n} (X_j - \sum_{i=1}^{p_n} \varrho_i X_{j-i})^2$. Condition (R) can even be met with $q_n = 1$ in smooth parametric models for the autoregression coefficients. See Schick and Wefelmeyer (2007b) for details.

We say that a function a has *finite* V*-variation* if there are measures μ_1 and μ_2 of equal mass with $\int V d(\mu_1 + \mu_2)$ finite such that $a(x) = \mu_1((-\infty, x]) - \mu_2((-\infty, x])$ for Lebesgue almost all x. In this case, we call $\int V\, d(\mu_1 + \mu_2)$ the V*-variation of* a. Our assumptions on the innovation density are quite weak. Aside from moment conditions, we require only that f has finite V_{r+1}-variation. In particular, f need not be continuous.

(F) *The density f has mean zero, a finite moment of order $\xi > 2r+3$ and finite V_{r+1}-variation.*

We formulate our assumptions on the kernel and the bandwidth in terms of the moment order ξ and a positive integer m. This integer m plays the role of a (known) lower bound on the number N of non-zero coefficients among φ_s, $s \ge 1$,

$$N = \sum_{s=1}^{\infty} \mathbf{1}[\varphi_s \neq 0].$$

Note that (C) is equivalent to $N \ge 1$. Thus we can always take $m = 1$. But this choice may lead to an undersmoothed estimator. A possible solution is to select a data-driven lower bound m by testing whether the first few coefficients are non-zero.

(K) *The kernel k is twice continuously differentiable with bounded derivatives and k, k' and k'' have finite V_{2r+2}-norms. Furthermore, k has finite V_{r+m+1}-norm and satisfies $\int t^i k(t)\,dt = 0$ for $i = 1, \dots, m$.*

(B) *The bandwidth b_n satisfies*

$$nb_n^{2m+2} \to 0, \quad n^{-1/4-\zeta} b_n^{-1} \to 0, \quad n^{-1/4} b_n^{-1/2} \to 0,$$

and the sequences b_n, q_n and p_n satisfy

$$n^{-3/4} p_n q_n b_n^{-2} \to 0, \quad n^{-1/2} p_n q_n b_n^{-1} \to 0 \quad \text{and} \quad p_n q_n n^{-1+2/\xi} = O(1).$$

Under (B) we also have $n^{-1/4} q_n^{1/2} \to 0$ and $b_n^m q_n^{1/2} \to 0$, conditions that appear in some of our results. If $b_n \sim (n \log n)^{-1/(2m+2)}$, then (B) is implied by $\log n \; p_n q_n n^{-\beta} \to 0$ with β the smaller of $m/(2m+2)$ and $1 - 2/\xi$.

To describe our results, we define the processes $\mathbb{F}_n$ and $\mathbb{G}_n$ by

$$\mathbb{F}_n(x) = \frac{1}{n-p_n} \sum_{j=p_n+1}^{n} \big(f(x-Y_j) - E[f(x-Y_j)]\big), \quad x \in \mathbb{R},$$

$$\mathbb{G}_n(x) = \frac{1}{n-p_n} \sum_{j=p_n+1}^{n} \big(g(x-\varepsilon_j) - E[g(x-\varepsilon_j)]\big), \quad x \in \mathbb{R},$$

and functions $\nu_1, \nu_2, \dots$ by

$$\nu_i(x) = E[X_0 f(x - Y_i)], \qquad x \in \mathbb{R}.$$

We shall see that these functions are differentiable under our assumptions on f.

Theorem 1. *Let r be non-negative, m a positive integer and $N \geq m$. Suppose (I), (Q), (R), (S+), (F), (K) and (B) hold. Then*

$$\Big\|\hat{h} - h - \mathbb{F}_n - \mathbb{G}_n + \sum_{i=1}^{p_n} (\hat{\varrho}_i - \varrho_i)\nu_i'\Big\|_{V_r} = o_p(n^{-1/2}).$$

If we use the least squares estimators we have a more explicit result. With $\mathbf{X}_{j-1} = (X_{j-1}, \dots, X_{j-p_n})^\top$ and $M_n = E[\mathbf{X}_0 \mathbf{X}_0^\top]$, let

$$\tilde{\Delta} = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} M_n^{-1} \mathbf{X}_{j-1} \varepsilon_j. \tag{1.3}$$

Theorem 2. *In addition to the assumptions of Theorem 1, suppose that f has a finite fourth moment and $np_n \sum_{s>p_n} \varrho_s^2 \to 0$ holds, and that we use the least squares estimators $\hat{\varrho}_i$. Then*

$$\Big\|\hat{h} - h - \mathbb{F}_n - \mathbb{G}_n + \sum_{i=1}^{p_n} \tilde{\Delta}_i \nu_i'\Big\|_{V_r} = o_p(n^{-1/2}),$$

and $n^{1/2}(\hat{h} - h)$ is tight in L_{V_r} and converges weakly in L_{V_r} to a centered Gaussian process.

Of special interest is the case when we have a parametric model for the autocorrelation coefficients: There are functions $r_1, r_2, \dots$ from an open subset Θ of $\mathbb{R}^d$ into $\mathbb{R}$ such that $\varrho_i = r_i(\vartheta)$ for all i and some unknown ϑ in Θ. Then we can take $\hat{\varrho}_i = r_i(\hat{\vartheta})$ for all i and some estimator $\hat{\vartheta}$ of ϑ. Now let us impose the following conditions.

(R1) *The estimator $\hat{\vartheta}$ of ϑ is $n^{1/2}$-consistent: $\hat{\vartheta} - \vartheta = O_p(n^{-1/2})$.*

(R2) *The functions $r_1, r_2, \dots$ are differentiable at ϑ with gradients $\dot{r}_1(\vartheta), \dot{r}_2(\vartheta), \dots$, and*

$$\sum_{i=1}^{\infty} \big(r_i(\vartheta + s) - r_i(\vartheta) - \dot{r}_i(\vartheta)^\top s\big)^2 = o(|s|^2) \quad \text{and} \quad \sum_{i=1}^{\infty} |\dot{r}_i(\vartheta)|^2 < \infty.$$

These conditions imply (R) with $q_n = 1$. As in Schick and Wefelmeyer (2007b) we obtain the expansion

$$\Big\|\hat{h} - h - \mathbb{F}_n - \mathbb{G}_n + (\hat{\vartheta} - \vartheta)^\top \sum_{i=1}^{\infty} \dot{r}_i(\vartheta)\nu_i'\Big\|_{V_r} = o_p(n^{-1/2}) \tag{1.4}$$

and hence tightness of $n^{1/2}(\hat{h} - h)$. Weak convergence of $n^{1/2}(\hat{h} - h)$ holds if $\hat{\vartheta}$ is *asymptotically linear* with *influence function* J, say,

$$\hat{\vartheta} - \vartheta = \frac{1}{n} \sum_{j=1}^{n} J(X_{j-1}, \varepsilon_j) + o_p(n^{-1/2}),$$

where $E(J(X_0,\varepsilon_1)|X_0)=0$ and $E[J(X_0,\varepsilon_1)J^\top(X_0,\varepsilon_1)]$ is positive definite and finite. A simple example is the AR(1) process $X_t=\vartheta X_{t-1}+\varepsilon_t$ with $|\vartheta|<1$ and $\vartheta\neq 0$. Then $r_1(\vartheta)=\vartheta$ and $r_s(\vartheta)=0$ for $s>1$, and $\sum_{i=1}^\infty \dot r_i(\vartheta)\nu_i'$ simplifies to ν_1', where

$$\nu_1(x)=E[X_0 f(x-\vartheta X_0)]=\int y f(x-\vartheta y)h(y)\,dy.$$

Other examples include MA(q) and ARMA(p,q).

The paper is organized as follows. In Section 2 we give some inequalities for V-norms. In Section 3 we study the space L_V. We characterize the compact subsets and consider continuity and Taylor expansions for shifts of functions. Section 5 presents conditions for tightness of sequences of L_V-valued random variables. In Sections 6 and 7 these results are applied to sequences of the form $n^{1/2}\mathbb{G}_n$ and $n^{1/2}\mathbb{F}_n$, respectively. Section 8 gives bounds on certain linear operators on L_V. In Section 9 we study how well the residuals approximate the true innovations, and obtain stochastic expansions in L_V for residual-based averages $(1/(n-p_n))\sum_{j=p_n+1}^n a_n(x-\hat\varepsilon_j)$ and $(1/(n-p_n))\sum_{j=p_n+1}^n a_n(x-\hat Y_j)$. The results of Sections 6–9 are used in Sections 10 and 11 to obtain convergence rates of $\hat f$ and $\hat g$ in L_V, stochastic expansions in L_V for linear functionals of the form $a*\hat f$ and $a*\hat g$, and tightness of $n^{1/2}\sum_{i=1}^{p_n}(\hat\varrho_i-\varrho_i)\nu_i'$ in L_V. Section 12 contains the proofs of Theorems 1 and 2 and of Lemma 13. Section 13 gives a variance bound used in Section 9.

2. The V-norm.

Throughout this paper, V is a continuous function on $\mathbb{R}$ satisfying $V(0)=1$ and

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

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

It follows from (2.2) that $V(x)\ge V(0)=1$ for all x in $\mathbb{R}$, that V is symmetric in the sense that $V(x)=V(-x)$ for all x in $\mathbb{R}$, and that $V(x)\ge V(y)$ if $|x|\ge|y|$. These properties and (2.1) yield

$$|V(x+s)-V(x)|\le V(x)(V(s)-1),\quad x,s\in\mathbb{R}. \tag{2.3}$$

Possible choices for V are $V(x)=\exp(|x|)$ and $V=V_r$ with $r\ge 0$.

For $\alpha>1$, we set $W_\alpha=V_\alpha V^2$ so that

$$W_\alpha(x)=(1+|x|)^\alpha V^2(x),\quad x\in\mathbb{R}.$$

The function W_α has the same properties as V.

We now present some inequalities for V-norms. It follows from the Cauchy–Schwarz inequality that

$$\Big(\int V(x)|a(x)|\,dx\Big)^2\le K_\alpha\int W_\alpha(x)a^2(x)\,dx$$

with $K_\alpha = \int (1+|x|)^{-\alpha}\,dx$. In other words,

$$\|a\|_V^2 \le K_\alpha \|a^2\|_{W_\alpha}. \tag{2.4}$$

In view of (2.1) the V-norm satisfies

$$\|a * b\|_V \le \|a\|_V \|b\|_V, \quad a, b \in L_V. \tag{2.5}$$

Moreover, since $(a * b)^2 \le \|b\|_1 (a^2 * |b|)$ by the Cauchy–Schwarz inequality, we obtain the inequality

$$\|(a * b)^2\|_V \le \|a^2\|_V \|b\|_V^2, \quad a^2 \in L_V, b \in L_V. \tag{2.6}$$

3. The space L_V.

In this section we study properties of the (Banach) space L_V of all (equivalence classes of) measurable functions a with finite V-norm. We begin by recalling the characterization of compact subsets given in Lemma 4 of Schick and Wefelmeyer (2007a). Introduce the shift $S_t a = a(\cdot - t)$.

Lemma 1. *A closed subset A of L_V is compact if and only if*

$$\sup_{a \in A} \|a\|_V < \infty, \tag{3.1}$$

$$\lim_{t \to 0} \sup_{a \in A} \|S_t a - a\|_V = 0, \tag{3.2}$$

$$\lim_{K \uparrow \infty} \sup_{a \in A} \int_{|x| > K} V(x)|a(x)|\,dx = 0. \tag{3.3}$$

From this one immediately obtains the following result. See Lemma 5 in Schick and Wefelmeyer (2007a).

Lemma 2. *Let k be a kernel with finite V-norm. Then $\sup_{a \in A} \|a * k_{b_n} - a\|_V \to 0$ for every compact subset A of L_V.*

Let us now give some simple sufficient conditions for compactness.

Lemma 3. *A closed subset A of L_V is compact if*

$$\sup_{a \in A} \int (1+|x|)^\beta V(x)|a(x)|\,dx < \infty \quad \textit{for some } \beta > 0 \tag{3.4}$$

and

$$\sup_{a \in A} \int_{|x| \le K} |a(x-t) - a(x)|\,dx \to 0 \quad \textit{as } t \to 0 \tag{3.5}$$

for all finite K.

Proof. We show that the present conditions imply (3.1) to (3.3). Condition (3.4) implies (3.1). It also implies (3.3) since

$$\int_{|x|\geq K} V(x)|a(x)|\,dx \leq (1+K)^{-\beta}\int_{|x|\geq K}(1+|x|)^{\beta}V(x)|a(x)|\,dx$$

for all positive K. With $g_t(x) = a(x-t) - a(x)$ we have

$$\|g_t\|_V \leq V(K)\int_{|x|\leq K}|g_t(x)|\,dx + \int_{|x|>K}V(x)|g_t(x)|\,dx, \quad K>0.$$

Since $V(x) \leq V(t)V(x-t)$ for all x and t, we also have

$$\int_{|x|>K}V(x)|g_t(x)|\,dx \leq (1+V(t))\int_{|x|>K-|t|}V(x)|a(x)|\,dx.$$

This shows that, in the presence of (3.3), condition (3.2) is equivalent to (3.5). □

In the remainder of this section we collect several convergence results for the space L_V. Let a be in L_V. Then (2.1) yields

$$\|S_t a\|_V = \int V(x+t)|a(x)|\,dx \leq V(t)\|a\|_V, \quad t\in\mathbb{R}, \tag{3.6}$$

and this and (2.2) imply

$$\sup_{|w|\leq 1}\|S_{wt}a\|_V \leq V(t)\|a\|_V, \quad t\in\mathbb{R}. \tag{3.7}$$

A measurable function a is called *V-Lipschitz (with constant L)* if

$$\|S_t a - a\|_V \leq L|t|V(t), \quad t\in\mathbb{R}. \tag{3.8}$$

We have the following connections between this concept and finite V-variation; see Schick and Wefelmeyer (2006, 2007a) for some of the details.

Lemma 4. *If a has finite V-variation M, then a is V-Lipschitz with constant M, and Va is bounded by M.*

Lemma 5. *If a is absolutely continuous and its a.e. derivative a' has finite V-norm, then a has finite V-variation $M = \|a'\|_V$.*

Lemma 6. *If a is V-Lipschitz with constant L and b has finite V-norm, then $a*b$ is V-Lipschitz with constant $L\|b\|_V$.*

Lemma 7. *If a is V_r-Lipschitz with constant L and has finite V_s-norm, where $0\leq s\leq r$, then a is V_s-Lipschitz with constant $LV_r(1)+2\|a\|_{V_s}$.*

A measurable function a is called *V-Lipschitz of order m (with constant L)* if there are measurable functions $a^{(1)}, \dots, a^{(m-1)}$ such that

$$\Big\| S_t a - a - \sum_{i=1}^{m-1} \frac{(-t)^i}{i!} a^{(i)} \Big\|_V \le L|t|^m V(t), \quad t \in \mathbb{R}. \tag{3.9}$$

If the functions $a^{(1)}, \dots, a^{(m-1)}$ have finite V-norms, then we say a is *strongly V-Lipschitz of order m*. Note that a is V-Lipschitz of order one if and only if a is V-Lipschitz.

We say that a is *absolutely continuous of order m* if a is $(m-1)$ times differentiable and its $(m-1)$-derivative is absolutely continuous. A sufficient condition for a to be V-Lipschitz of order m is that a is *V-regular of order m*: the function a is absolutely continuous of order $m-1$ and its $(m-1)$-derivative $a^{(m-1)}$ is V-Lipschitz. In this case the functions $a^{(1)}, \dots, a^{(m-1)}$ appearing in (3.9) are the derivatives of a, and the L in (3.9) is the Lipschitz constant of $a^{(m-1)}$. If, in addition, the derivatives $a^{(1)}, \dots, a^{(m-1)}$ have finite V-norms, then we say a is *strongly V-regular of order m*. In this case a is strongly V-Lipschitz of order m.

The following lemmas summarize results from Schick and Wefelmeyer (2006). The first provides a sufficient condition for a convolution to be V-Lipschitz of order two.

Lemma 8. *Let a have finite V-variation M, and let b have finite V-norm. Then $a * b$ is absolutely continuous, an a.e. derivative is given by*

$$(a*b)'(x) = \int b(x-y)(\mu_1(dy) - \mu_2(dy)), \quad x \in \mathbb{R},$$

and $\|(a*b)'\|_V \le M\|b\|_V$. Moreover, if Vb is bounded by B, then $V(a*b)'$ is bounded by BM, and if b is V-Lipschitz with constant L, then $(a*b)'$ is V-Lipschitz with constant ML and $a*b$ is strongly V-Lipschitz of order two with constant ML.

Lemma 9. *Let $a_1, \dots, a_m$ and b belong to L_V. If $a_1, \dots, a_m$ have finite V-variation, then the function $a = a_1 * \dots * a_m * b$ is absolutely continuous of order m with all m derivatives in L_V and is hence strongly V-regular of order m. If also b is V-Lipschitz, then the function a is strongly V-regular of order $m+1$.*

We say a kernel k has *V-order m* if

$$\int t^i k(t)\, dt = 0, \quad i = 1, \dots, m-1, \tag{3.10}$$

and

$$\int |t|^m V(t) |k(t)|\, dt < \infty. \tag{3.11}$$

The next lemma is a special case of Lemma 4.1 in Schick and Wefelmeyer (2006).

Lemma 10. *Let a be V-Lipschitz of order m and let k be an integrable function that satisfies the integrability conditions (3.10) and (3.11). Then*

$$\Big\| a * k_{b_n} - a \int k(t)\,dt \Big\|_V \le L b_n^m \int |t|^m V(b_n t)|k(t)|\,dt.$$

In particular, for a kernel k of V-order m, we have $\|a * k_{b_n} - a\|_V = O(b_n^m)$.

Lemma 11. *Let a be VV_1-Lipschitz with constant L and have finite VV_1-norm. Then the map b defined by $b(x) = xa(x)$ belongs to L_V and is V-Lipschitz with constant $2L+5\|a\|_{VV_1}$.*

Proof. We have $\|S_t b - b\|_V \le |t|\|S_t a\|_V + \|S_t a - a\|_{VV_1}$. By Lemma 4.3 of Schick and Wefelmeyer (2006) we have $\|S_t a - a\|_{VV_1} \le |t|V(t)(2L + 4\|a\|_{VV_1})$ and thus $\|S_t b - b\|_V \le |t|V(t)\|a\|_V + |t|V(t)(2L + 4\|a\|_{VV_1})$. The desired result is now immediate. □

4. A central limit theorem in L_V.

We recall the central limit theorem for L_1-spaces; see Ledoux and Talagrand (1991, Theorem 10.10) or van der Vaart and Wellner (1996, page 92).

Theorem 3. *Let μ be a σ-finite measure on the Borel-σ-field on $\mathbb{R}$. Let $Z_1, Z_2, \dots$ be independent and identically distributed zero-mean random elements in $L_1(\mu)$. Then the sequence $n^{-1/2}\sum_{i=1}^n Z_i$ converges in distribution (in $L_1(\mu)$) to a centered Gaussian process if and only if*

$$\lim_{t\to\infty} t^2 P\Big(\int |Z_1(x)|\,\mu(dx) > t\Big) = 0 \quad \textit{and} \quad \int E[Z_1^2(x)]^{1/2}\,\mu(dx) < \infty.$$

We now formulate a special case more suitable to our needs in the space L_V.

Lemma 12. *Let $U_1, U_2, \dots$ be independent and identically distributed random variables, a be a measurable function and*

$$\mathbb{H}(x) = \frac{1}{n}\sum_{j=1}^n \big(a(x - U_j) - E[a(x - U_j)]\big), \quad x \in \mathbb{R}.$$

If $\|a^2\|_{W_\alpha}$ and $E[W_\alpha(U_1)]$ are finite for some $\alpha > 1$, then $\sqrt{n}\mathbb{H}$ converges in distribution in the space L_V to a centered Gaussian process whose covariance structure matches that of $a(\cdot - U_1)$.

Proof. We apply the previous theorem with $\mu(dx) = V(x)\,dx$ and $Z_i(x) = a(x - U_i) - E[a(x - U_i)]$. Using (2.4) we find that

$$E\Big[\Big(\int |Z_1(x)|V(x)\,dx\Big)^2\Big] = E[\|Z_1\|_V^2] \le K_\alpha \int W_\alpha(x) E[Z_1^2(x)]\,dx$$

and

$$\Big(\int E[Z_1^2(x)]^{1/2} V(x)\,dx\Big)^2 \le K_\alpha \int W_\alpha(x) E[Z_1^2(x)]\,dx.$$

Since $t^2 P(X > t) \leq E[X^2 \mathbf{1}[X > t]]$ for $t > 0$, we need only show that $\int W_\alpha(x) E[Z_1^2(x)]\, dx$ is finite. Since $E[Z_1^2(x)] \leq E[a^2(x - U_1)]$, the integral in question can be bounded by $E[W_\alpha(U_1)] \|a^2\|_{W_\alpha}$ and is thus finite by our assumptions. □

5. Tightness in L_V.

The compactness conditions of Section 3 are now applied to obtain tightness of sequences of random variables in L_V. These will be used in Sections 6 and 7 to obtain tightness for averages of dependent L_V-valued random elements. Lemma 1 immediately implies the following characterization.

Proposition 1. *A sequence $\mathbb{A}_n$ of L_V-valued random variables is tight if and only if the following three conditions hold.*

(T1) *For every $\eta > 0$ there is a finite M such that for all (large) n,*

$$P(\|\mathbb{A}_n\|_V > M) < \eta.$$

(T2) *For every $\eta > 0$ there is a $\delta > 0$ such that for all (large) n,*

$$P\Big(\sup_{|t|<\delta} \|S_t \mathbb{A}_n - \mathbb{A}_n\|_V > \eta\Big) < \eta.$$

(T3) *For every $\eta > 0$ there is a finite K such that for all (large) n,*

$$P\Big(\int_{|x|>K} V(x) |\mathbb{A}_n(x)|\, dx > \eta\Big) < \eta.$$

From Lemma 3 we can derive the following sufficient conditions for tightness.

Proposition 2. *A sequence $\mathbb{A}_n$ of L_V-valued random variables is tight if the following two conditions are met.*

(T1′) *For some $\beta > 0$ and every $\eta > 0$ there is a finite M such that for all (large) n,*

$$P(\|V_\beta \mathbb{A}_n\|_V > M) < \eta.$$

(T2′) *For every $\eta > 0$ and finite K there is a $\delta > 0$ such that for all (large) n,*

$$P\Big(\sup_{|t|<\delta} \int_{|x|\leq K} |\mathbb{A}_n(x - t) - \mathbb{A}_n(x)|\, dx > \eta\Big) < \eta.$$

Let us now derive simple sufficient conditions for (T1′) and (T2′). The inequality (2.4) and the Markov inequality show that a sufficient condition for (T1′) is given by

$$\sup_n \int W_\alpha(x) E[\mathbb{A}_n^2(x)]\, dx < \infty \tag{5.1}$$

for some $\alpha > 1$. Since the process $\mathbb{X}_n$ defined by

$$\mathbb{X}_n(t) = \int_{|x|\leq K} |\mathbb{A}_n(x - t) - \mathbb{A}_n(x)|\, dx, \quad t \in [-1, 1],$$

has continuous sample paths, we can obtain sufficient conditions for (T2′) from sufficient conditions for tightness of the sequence $\mathbb{X}_n$. Since $\mathbb{X}_n(0) = 0$, one such condition is that $E[|\mathbb{X}_n(t) - \mathbb{X}_n(s)|^2] \le A|t-s|^\beta$ for some finite A, some $\beta > 1$ and all t and s in $[-1, 1]$; see Theorem 12.3 in Billingsley (1968). Another condition is that $E[(\mathbb{X}_n(t) - \mathbb{X}_n(t_1))^2(\mathbb{X}_n(t_2) - \mathbb{X}_n(t))^2] \le A|t_2 - t_1|^\beta$ for some finite A, some $\beta > 1$ and all $-1 \le t_1 \le t \le t_2 \le 1$; see Problem 7 on page 102 in Billingsley (1968). Since $(\mathbb{X}_n(t) - \mathbb{X}_n(s))^2 \le 2K \int (\mathbb{A}_n(x-t) - \mathbb{A}_n(x-s))^2\,dx = 2K \int (\mathbb{A}_n(x - |t-s|) - \mathbb{A}_n(x))^2\,dx$, we see that (T2′) follows if for some finite A and some $\beta > 1$ we have

$$\int E[(\mathbb{A}_n(x-t) - \mathbb{A}_n(x))^2]\,dx \le At^\beta, \quad 0 \le t \le 1, \tag{5.2}$$

or

$$\iint E[(\mathbb{A}_n(x-s) - \mathbb{A}_n(x))^2(\mathbb{A}_n(y-t) - \mathbb{A}_n(y))^2]\,dxdy \le At^\beta, \quad 0 \le s \le t \le 1. \tag{5.3}$$

Let us summarize this in the following theorem.

Theorem 4. *A sequence $\mathbb{A}_n$ of L_V-valued random variables is tight if (5.1) holds for some $\alpha > 1$ and if either (5.2) or (5.3) holds for some finite A and some $\beta > 1$.*

In the next two sections we use this theorem to establish tightness of some important sequences of random variables.

6. A class of tight sequences.

Let $Z_1, Z_2, \dots$ be i.i.d. random variables with distribution function D. Let $U_1, U_2, \dots$ be stationary random variables with $U_1, \dots, U_j$ independent of $Z_j, Z_{j+1}, \dots$ for all $j \ge 1$, and let $a_1, a_2, \dots$ be measurable functions. Set

$$\mathbb{A}_n(x) = n^{-1/2} \sum_{j=1}^n \Big(a_n(x - U_j - Z_j) - \int a_n(x - U_j - z)\,dD(z) \Big), \quad x \in \mathbb{R}.$$

Let $W = W_\alpha$ for some $\alpha > 1$. Then, with $U = U_1$ and $Z = Z_1$, we calculate

$$\begin{aligned} \int W(x) E[\mathbb{A}_n^2(x)]\,dx &\le \int W(x) E[\operatorname{Var}(a_n(x - U - Z)|U)]\,dx \\ &\le E[W(U)] \int W(x) \operatorname{Var}(a_n(x - Z))\,dx \end{aligned}$$

and

$$\begin{aligned} \int E[(\mathbb{A}_n(x-t) - \mathbb{A}_n(x))^2]\,dx &\le \int E[(a_n(x - t - U - Z) - a_n(x - U - Z))^2]\,dx \\ &= \int (a_n(x-t) - a_n(x))^2\,dx. \end{aligned}$$

Thus Theorem 4 yields the following result.

Proposition 3. *The sequence $\mathbb{A}_n$ is tight if the sequence $\int W_\alpha(x)\operatorname{Var}(a_n(x-Z))\,dx$ is bounded and $E[W_\alpha(U)]$ is finite for some $\alpha > 1$, and if*

$$\sup_n \int (a_n(x-t) - a_n(x))^2\,dx \le Bt^\kappa, \quad 0 < t < 1, \tag{6.1}$$

for some finite B and some $\kappa > 1$.

Let us now establish the bound (5.3). For this we first state the following bound. Its proof is deferred to Section 12.

Lemma 13. *Assume that (6.1) holds with $\kappa > 0$. Then for any random variables $U_{kij} = U_{kji}$ that are independent of (Z_i, Z_j, Z_k), the left-hand side of (5.3) is bounded by*

$$C = 5(4B)^2 t^{2\kappa} + 8(4B)^2 t^{3\kappa/2} \frac{1}{n^2} \sum_{1\le i<j<k\le n} \sqrt{E[\min\{t^\kappa, |U_k - U_{kij}|^\kappa\}]}.$$

Let us mention some consequences of this inequality.

Proposition 4. *Suppose the sequences $U_1, U_2, \ldots$ and $Z_1, Z_2, \ldots$ are independent. Then the sequence $\mathbb{A}_n$ is tight if $E[W_\alpha(U)]$ is finite and the sequence $\int W_\alpha(x)\operatorname{Var}(a_n(x-Z))\,dx$ is bounded for some $\alpha > 1$, and if (6.1) holds with $\kappa > 1/2$.*

Proof. If $\kappa > 1$, the result follows from Proposition 3. For $1/2 < \kappa \le 1$, use $U_{kij} = U_k$ in Lemma 13. Then $C = 5(4B)^2 t^{2\kappa}$, and (5.3) holds as $\kappa > 1/2$. □

Proposition 5. *Suppose that the random variables U_{kij} of Lemma 13 fulfill*

$$|U_k - U_{kij}| \le c_{k-i}|Z_i| + c_{k-j}|Z_j|, \quad 1 \le i < j \le k, \tag{6.2}$$

with positive numbers $c_1, c_2, \ldots$ satisfying

$$\sum_{j=1}^{\infty} \sqrt{c_j} < \infty. \tag{6.3}$$

Let Z have a finite mean. Then the sequence $\mathbb{A}_n$ is tight if $E[W_\alpha(U)]$ is finite and the sequence $\int W_\alpha(x)\operatorname{Var}(a_n(x-Z))\,dx$ is bounded for some $\alpha > 1$, and if (6.1) holds with $\kappa = 1$.

Proof. Since $\kappa = 1$, condition (5.3) will follow if we show that

$$\frac{1}{n^2} \sum_{1\le i<j<k\le n} \sqrt{E[c_{k-i}|Z_i| + c_{k-j}|Z_j|]} = O(1).$$

Since Z has a finite mean and $\sqrt{a+b} \le \sqrt{a} + \sqrt{b}$ for positive a and b, this follows as

$$\begin{aligned}
&\frac{1}{n^2} \sum_{1 \le i<j<k \le n} \sqrt{c_{k-i} + c_{k-j}} \\
&\quad \le \frac{1}{n^2} \sum_{1 \le i<k \le n} (k-i)\sqrt{c_{k-i}} + \frac{1}{n^2} \sum_{1 \le j<k \le n} (j-1)\sqrt{c_{k-j}} \\
&\quad = \frac{1}{n^2} \sum_{1 \le i<k \le n} (k-1)\sqrt{c_{k-i}} \le \sum_{j=1}^{\infty} \sqrt{c_j}.
\end{aligned}$$

□

Remark 1. A sufficient condition for (6.3) is that $\sum_{j=1}^{\infty} j^{\beta}|c_j|$ is finite for some $\beta > 1$.

Example 1. If we take $a_n(x) = \mathbf{1}[x \ge 0]$, then the process $\mathbb{A}_n$ becomes the empirical process $\mathbb{D}_n$ defined by

$$\mathbb{D}_n(x) = n^{-1/2} \sum_{j=1}^{n} \big(\mathbf{1}[Z_j - U_j \le x] - D(x - U_j)\big), \quad x \in \mathbb{R}.$$

In this case, $\int W(x) \operatorname{Var}(a_n(x - Z_1))\, dx = \int W(x) D(x)(1 - D(x))\, dx = \|D(1-D)\|_W$ and $a_n(x-t) - a_n(x) = -\mathbf{1}[0 \le x < t]$, so that inequality (6.1) holds with $B = 1$ and $\kappa = 1$. Now assume that $U_k = \sum_{j=1}^{\infty} d_j Z_{k-j}$ with coefficients d_j such that $\sum_{j=1}^{\infty} \sqrt{|d_j|}$ is finite. Upon taking $U_{kij} = U_k - d_{k-i}Z_i - d_{k-j}Z_j$, we see that (6.2) and (6.3) hold with $c_j = |d_j|$. Consequently, the empirical process $\mathbb{D}_n$ is tight in L_V if $D(1-D)$ has finite W-norm and $E[W(U)]$ is finite. For $V = V_m$ with m a non-negative integer, these conditions are equivalent to D having a finite moment of order $2m + 1 + \alpha$.

7. A second class of tight sequences.

Consider a linear process

$$U_t = \sum_{s=1}^{\infty} d_s Z_{t-s}, \quad t \in \mathbb{Z},$$

with independent and identically distributed innovations Z_t, $t \in \mathbb{Z}$, with finite mean and coefficients $d_1, d_2, \dots$ such that

$$\sum_{j=1}^{\infty} j|d_j| < \infty. \tag{7.1}$$

For bounded functions a and a' set

$$\mathbb{A}_n(x) = n^{-1/2} \sum_{j=1}^{n} \Big(a(x - U_j) - E[a(x - U_j)]\Big), \quad x \in \mathbb{R},$$

$$\mathbb{A}'_n(x) = n^{-1/2} \sum_{j=1}^{n} \Big(a'(x - U_j) - E[a'(x - U_j)]\Big), \quad x \in \mathbb{R}.$$

Now assume that a is absolutely continuous with a.e. derivative a'. Then we have

$$\mathbb{A}_n(x-t) - \mathbb{A}_n(x) = -t \int_0^1 \mathbb{A}_n'(x-st)\,ds, \quad t, x \in \mathbb{R}.$$

Hence Theorem 4 shows that the sequence $\mathbb{A}_n$ is tight in L_{V_r} if

$$\sup_n \int (1+|x|)^\nu E[\mathbb{A}_n^2(x)]\,dx < \infty \tag{7.2}$$

for some $\nu > 2r+1$, and if

$$\sup_n \int E[(\mathbb{A}_n'(x))^2]\,dx < \infty. \tag{7.3}$$

Sufficient conditions for (7.2) and (7.3) are given in Schick and Wefelmeyer (2007b) and are recalled in Section 13. More precisely, Lemma 23 applied with $h = a$ and $q = r+1$ and $q > p > r$, Lemma 24 applied with $h = a'$, and Lemmas 4 and 5 yield the following result.

Proposition 6. *Let r be a non-negative number. Assume (7.1) holds and $E[|Z_0|^\nu]$ is finite for some $\nu > 2r+1$. Then the sequence $\mathbb{A}_n$ is tight in L_{V_r} if a and its a.e. derivative a' have finite V_{r+1}-norms and a' is bounded and 1-Lipschitz.*

8. Some operators on L_V.

We now return to the linear process introduced in (1.1). Throughout this section we assume that

$$E[|X_0|V(Y_i)] < \infty, \quad i = 1, 2, \dots$$

A sufficient condition for this is that $E[V^2(Y_1)]$ and $E[X_0^2]$ are finite. For the special case $V = V_r$ it suffices that ε_0 has a finite moment of order $r+1$. This follows from the following lemma.

Lemma 14. *Let m be a positive integer. For $i = 0, \dots, m$, let $\xi_i = \sum_{s\in\mathbb{Z}} c_{is}\varepsilon_s$ with $C = \sum_{s\in\mathbb{Z}}\sum_{i=0}^m |c_{is}|$ finite. Let $r, r_1, \dots, r_m$ be non-negative numbers and $t = \max(1, r + r_1 + \dots + r_m)$. Then*

$$E\Big[V_r(\xi_0)\prod_{i=1}^m |\xi_i|^{r_i}\Big] \le 2^{t-1}(1 + C^t E[|\varepsilon_0|^t]).$$

Proof. For $0 \le p \le 1 \le q$ one has

$$V_p(x) \le V_q(x) \le 2^{q-1}(1+|x|^q), \quad x \in \mathbb{R}.$$

Let $Z = \sum_{s\in\mathbb{Z}}\sum_{i=1}^m |c_{is}||\varepsilon_s|$. Then $|\xi_i| \le Z$ for all $i = 0, \dots, m$ and

$$V_r(\xi_0)\prod_{i=1}^m |\xi_i|^{r_i} \le V_t(Z) \le 2^{t-1}(1+Z^t).$$

The desired result follows now from the Minkowski inequality. □

Now we investigate some linear operators on L_V. For $i = 1, 2, \ldots$, we let T_i denote the operator which maps a in L_V to $T_i a$ in L_V defined by

$$T_i a(x) = E[X_0 a(x - Y_i)], \quad x \in \mathbb{R}.$$

These operators are bounded since

$$\|T_i a\|_V \le E[|X_0| V(Y_i)] \|a\|_V. \tag{8.1}$$

If N is finite, there is an integer τ such that $\varphi_s = 0$ for $s > \tau$. In this case $T_i a = 0$ for all $i > \tau$, as X_0 and Y_i are independent for such i and $E[X_0] = 0$.

With $\varphi_0 = 1$, we can express $T_i a$ as

$$T_i a(x) = \sum_{s=0}^{\infty} \varphi_s E[\varepsilon_{-s} a(x - Y_i)] = \sum_{s=0}^{\infty} \varphi_s D_{s+i} a(x)$$

where $D_j a(x) = E[\varepsilon_0 a(x - Y_j]$. Let m be a positive integer less than or equal to N. Denote by $\tau_1, \ldots, \tau_m$ the indices of the first m non-zero coefficients and set $\phi_i = \varphi_{\tau_i}$. Then we can write $Y_t = \phi_1 \varepsilon_{t-\tau_1} + \cdots + \phi_m \varepsilon_{t-\tau_m} + U_t$ with $U_t = \sum_{s>\tau_m} \varphi_s \varepsilon_{t-s}$, and obtain the representation

$$D_j a(x) = \sum_{i=1}^{m} \mathbf{1}[j = \tau_i] \bar{a}_i(x - U_j) + \mathbf{1}[j > \tau_m] E[\varepsilon_0 \bar{a}_0(x - U_j)]$$

with

$$\bar{a}_0(x) = E[a(x - \phi_1 \varepsilon_1 - \cdots - \phi_m \varepsilon_m)]$$

and

$$\bar{a}_i(x) = E[\varepsilon_i a(x - \phi_1 \varepsilon_1 - \cdots - \phi_m \varepsilon_m)], \quad i = 1, \ldots, m.$$

We can express $\bar{a}_i$ as the convolution $a * \psi_i$, where $\psi_i = \psi_{i1} * \cdots * \psi_{im}$ and

$$\psi_{ij}(x) = \frac{1}{|\phi_j|} f\Big(\frac{x}{\phi_j}\Big)\Big(\mathbf{1}[j \ne i] + \mathbf{1}[j = i] \frac{x}{\phi_j}\Big), \quad x \in \mathbb{R},$$

for $i = 0, \ldots, m$ and $j = 1, \ldots, m$.

Assume now that f has finite V_{r+1}-norm and finite V_{r+1}-variation and that a has finite V_r-norm. Then ψ_{ij} has finite V_{r+1}-norm and finite V_{r+1}-variation for $i \ne j$, while ψ_{ii} has finite V_r-norm and is V_r-Lipschitz; see Lemma 11. Thus, by Lemma 9, the functions $\psi_0, \ldots, \psi_m$ are strongly V_r-regular of order m with a common constant Λ; ψ_0 is even strongly V_{r+1}-regular of order m. This implies that the functions $\bar{a}_0, \ldots, \bar{a}_m$ are also strongly V_r-regular of order m with common constant $\Lambda \|a\|_{V_r}$. From this we derive that $D_j a$ is V_r-regular of order m with constant $E[(1 + |\varepsilon_0|) V_r(U_j)] \Lambda \|a\|_{V_r}$. For $i = 1, \ldots, m-1$, the i-th (almost everywhere) derivative $(D_j a)^{(i)}$ of $D_j a$ is given by

$$(D_j a)^{(i)}(x) = \sum_{l=1}^{m} \mathbf{1}[j = \tau_l] \bar{a}_l^{(i)}(x - U_j) + \mathbf{1}[j > \tau_m] E[\varepsilon_0 \bar{a}_0^{(i)}(x - U_j)]$$

and satisfies

$$\|(D_j a)^{(i)}\|_{V_r} \le E[(1+|\varepsilon_0|)V_r(U_j)] \max_{l=0,\dots,m} \|\bar{a}_l^{(i)}\|_{V_r}.$$

An alternative bound is available for large j. Indeed, for $j > \tau_m$, we have $(D_j a)^{(i)} = E[\varepsilon_0 \bar{a}_0^{(i)}(x-U_j)] = E[\varepsilon_0(\bar{a}_0^{(i)}(x-U_j) - \bar{a}_0^{(i)}(x-U_j+\varphi_j\varepsilon_0))]$ and obtain, since $\bar{a}_0^{(i)}$ is V_r-Lipschitz with constant L_i, that

$$\|(D_j a)^{(i)}\|_{V_r} \le |\varphi_j| E[\varepsilon_0^2] E[V_r(U_j - \varphi_j\varepsilon_0)] L_i.$$

If a has finite V_r-variation, then the functions $\bar{a}_0, \dots, \bar{a}_m$ are strongly V_r-regular of order $m+1$ and so are the functions $D_j a$, and the above holds also for $i = m$. It follows from Lemma 14 that $\sup_{j\ge 0} E[(1+|\varepsilon_0|)V_r(U_j)] < \infty$ and $\sup_{j \ge 0} E[V_r(U_j - \varphi_j\varepsilon_0)] < \infty$ as f has finite V_{r+1}-norm. Thus we have proved the following results.

Lemma 15. *Let $r \ge 0$ and let m be a positive integer. Suppose f has finite V_{r+1}-norm and finite V_{r+1}-variation. Let $N \ge m$. Then there is a constant C such that, for each a of finite V_r-norm, the functions $T_1 a, T_2 a, \dots$ are V_r-regular of order m with a common constant $C\|a\|_{V_r}$, and*

$$\sup_{0\le j\le m-1} \sum_{i=1}^{\infty} \|(T_i a)^{(j)}\|_{V_r} \le C\|a\|_{V_r}. \tag{8.2}$$

There is also a constant K such that, for all a with finite V_r-norm and finite V_r-variation M, the functions $T_1 a, T_2 a, \dots$ are V_r-regular of order $m+1$ with common constant KM and

$$\sum_{i=1}^{\infty} \|(T_i a)^{(m)}\|_{V_r} \le KM.$$

Lemma 16. *Let $r \ge 0$. Suppose a and f have finite V_{r+1}-variation, a has a finite V_{r+1}-norm and f a finite V_{r+2}-norm. Let $N \ge 4$. Then the functions $T_1 a, T_2 a, \dots$ are absolutely continuous of order three and*

$$\sup_{0\le j\le 3} \sum_{i=1}^{\infty} \|(T_i a)^{(j)}\|_{V_{r+1}} < \infty.$$

Proof. The above considerations with $m = 4$ show that $\bar{a}_0, \dots, \bar{a}_4$ have finite V_{r+1}-norms and are strongly V_{r+1}-regular of order four. The desired result is now immediate. □

Corollary 1. *Let $r \ge 0$ and let f have finite V_{r+1}-norm and finite V_{r+1}-variation. Let the kernel k have finite V_r-norm and be continuously differentiable with k' having finite V_{r+1}-norm. Then*

$$\sum_{i=1}^{\infty} \|T_i k'_{b_n}\|_{V_r} = O(1). \tag{8.3}$$

Proof. Note that $T_i k'_{b_n} = (T_i k_{b_n})'$. Thus the desired result follows from Lemma 15 with $m = 2$ if $N \geq 2$. If $N = 1$, the left-hand side of (8.3) simplifies to $\|k'_{b_n} * \psi_{11}\|_{V_r}$ and is bounded by $L \int |t| V(b_n t) |k'(t)|\, dt$ with L the V_r-Lipschitz constant of ψ_{11}; here we used Lemma 10 with $m = 1$ and $\int k'(t)\, dt = 0$. □

9. Behavior of residual-based processes.

Let $X_1, \dots, X_n$ be observations from the linear process (1.1). Throughout this section let a_n be twice continuously differentiable functions such that a'_n and a''_n have finite V-norms. Then we have the following inequalities.

$$\|S_t a_n - a_n\|_V \leq \|a'_n\|_V |t| V(t), \quad t \in \mathbb{R}, \tag{9.1}$$

$$\|S_t a'_n - a'_n\|_V \leq \|a''_n\|_V |t| V(t), \quad t \in \mathbb{R}, \tag{9.2}$$

$$\|S_t a_n - a_n + t a'_n\|_V \leq \|a''_n\|_V t^2 V(t), \quad t \in \mathbb{R}. \tag{9.3}$$

Set $\hat{\Delta} = (\hat{\varrho}_1 - \varrho_1, \dots, \hat{\varrho}_{p_n} - \varrho_{p_n})^\top$. Recall that $\mathbf{X}_{j-1} = (X_{j-1}, \dots, X_{j-p_n})^\top$. We first study the processes $\mathbb{A}_{n1}$ and $\mathbb{B}_{n1}$ defined by

$$\mathbb{A}_{n1}(x) = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} \big(a_n(x - \hat{\varepsilon}_j) - a_n(x - \varepsilon_j)\big),$$

$$\mathbb{B}_{n1}(x) = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} \mathbf{X}_{j-1} a'_n(x - \varepsilon_j), \quad x \in \mathbb{R}.$$

For this we introduce the following condition.

(F0) *The density f has a finite moment of order $\xi > 2$ and $p_n q_n n^{-1+2/\xi}$ is bounded.*

From Schick and Wefelmeyer (2007b) we recall some properties of the average

$$\overline{\mathbf{X}} = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} \mathbf{X}_{j-1}$$

and of the residuals $\hat{\varepsilon}_j$ and the closely related quantities

$$\hat{\varepsilon}^*_j = \varepsilon_j - \hat{\Delta}^\top \mathbf{X}_{j-1} = \hat{\varepsilon}_j - \sum_{i=p_n+1}^{\infty} \varrho_i X_{j-i}, \quad j = p_n + 1, \dots, n. \tag{9.4}$$

Lemma 17. *Suppose (I), (Q) and (R) hold. Then*

$$\textstyle\sum_{j=p_n+1}^{n} (\hat{\varepsilon}_j - \varepsilon_j)^2 = O_p(p_n q_n), \tag{9.5}$$

$$\textstyle\sum_{j=p_n+1}^{n} (\hat{\varepsilon}_j - \hat{\varepsilon}^*_j)^2 = O_p(n^{-2\zeta}), \tag{9.6}$$

$$\hat{\Delta}^\top \overline{\mathbf{X}} = O_p(n^{-1} p_n^{1/2} q_n^{1/2}). \tag{9.7}$$

If also (F0) holds, then

$$\max_{p_n+1\le j\le n} |\hat\varepsilon_j - \varepsilon_j| = o_p(1), \tag{9.8}$$

Lemma 18. *Suppose (I), (Q), (R) and (F0) hold and f has finite V^2-norm. Then*

$$\|\mathbb{A}_{n1} - \hat\Delta^\top \mathbb{B}_{n1}\|_V = O_p(n^{-1/2-\zeta}\|a_n'\|_V) + O_p(n^{-1}p_n q_n \|a_n''\|_V).$$

Proof. By continuity of V we have

$$\max_{p_n+1\le j\le n} V(\eta_j) = 1 + o_p(1) \quad \text{if} \quad \max_{p_n+1\le j\le n} |\eta_j| = o_p(1). \tag{9.9}$$

The properties of f imply that $E[V^2(\varepsilon_1)]$ is finite and $E[V(\varepsilon_1)|\mathbf{X}_0|^2] = \|f\|_V E[\|\mathbf{X}_0\|^2] = O(p_n)$. Thus (2.1), (9.8) and (9.9) imply

$$U_{n1} = \frac{1}{n-p_n}\sum_{j=p_n+1}^n V^2(\hat\varepsilon_j) \le \frac{1}{n-p_n}\sum_{j=p_n+1}^n V^2(\varepsilon_j)V^2(\hat\varepsilon_j - \varepsilon_j) = O_p(1),$$

$$U_{n2} = \frac{1}{n-p_n}\sum_{j=p_n+1}^n V(\varepsilon_j)|\mathbf{X}_{j-1}|^2 = O_p(p_n).$$

Let $\mathbb{A}_{n1}^*$ be defined as $\mathbb{A}_{n1}$, but with $\hat\varepsilon_j$ replaced by $\hat\varepsilon_j^*$ given in (9.4). Then, by (3.6) and (9.1),

$$\begin{aligned}
\|\mathbb{A}_{n1} - \mathbb{A}_{n1}^*\|_V &\le \frac{1}{n-p_n}\sum_{j=p_n+1}^n \|a_n(\cdot - \hat\varepsilon_j) - a_n(\cdot - \hat\varepsilon_j^*)\|_V \\
&\le \|a_n'\|_V \frac{1}{n-p_n}\sum_{j=p_n+1}^n V(\hat\varepsilon_j)|\hat\varepsilon_j - \hat\varepsilon_j^*| \max_{p_n+1\le j\le n} V(\hat\varepsilon_j - \hat\varepsilon_j^*) \\
&\le \|a_n'\|_V \Big(U_{n1}\frac{1}{n-p_n}\sum_{j=p_n+1}^n (\hat\varepsilon_j - \hat\varepsilon_j^*)^2\Big)^{1/2} \max_{p_n+1\le j\le n} V(\hat\varepsilon_j - \hat\varepsilon_j^*).
\end{aligned}$$

Since (9.6) implies $\max_{p_n+1\le j\le n} |\hat\varepsilon_j^* - \hat\varepsilon_j| = o_p(1)$, relations (9.6) and (9.9) give

$$\|\mathbb{A}_{n1} - \mathbb{A}_{n1}^*\|_V = O_p(n^{-1/2-\zeta}\|a_n'\|_V).$$

From (3.6) and (9.3) we obtain the bound

$$\|\mathbb{A}_{n1}^* - \hat\Delta^\top \mathbb{B}_{n1}\|_V \le \frac{1}{n-p_n}\sum_{j=p_n+1}^n V(\varepsilon_j)\|a_n''\|_V(\hat\varepsilon_j^* - \varepsilon_j)^2 V(\hat\varepsilon_j^* - \varepsilon_j).$$

Note that assumption (R) implies that $|\hat\Delta|^2 = O_p(n^{-1}q_n)$. Thus (9.6), (9.8), (9.9) and the identity $\hat\varepsilon_j^* - \varepsilon_j = -\hat\Delta^\top \mathbf{X}_{j-1}$ give

$$\|\mathbb{A}_{n1}^* - \hat\Delta^\top \mathbb{B}_{n1}\|_V = O_p(\|a_n''\|_V |\hat\Delta|^2 U_{n2}) = O_p(n^{-1}p_n q_n \|a_n''\|_V).$$

The desired result is now immediate. □

Lemma 19. *Suppose (I), (Q) and (R) hold. Let f and $(a_n')^2$ have finite W_α-norms for some $\alpha > 1$. Then*

$$\|\hat{\Delta}^\top \mathbb{B}_{n1}\|_V = O_p\big(n^{-1} p_n^{1/2} q_n^{1/2} \|(a_n')^2\|_{W_\alpha}^{1/2}\big).$$

Proof. Let $W = W_\alpha$. Set $\bar{\mathbb{B}}_{n1}(x) = \overline{\mathbf{X}}(a_n' * f)(x)$ for $x \in \mathbb{R}$. In view of (9.7), (2.5) and (2.4) we have

$$\|\hat{\Delta}^\top \bar{\mathbb{B}}_{n1}\|_V = O_p(n^{-1} p_n^{1/2} q_n^{1/2} \|a_n' * f\|_V) = O_p\big(n^{-1} p_n^{1/2} q_n^{1/2} \|(a_n')^2\|_W^{1/2}\big).$$

Since $\mathbb{B}_{n1}(x) - \bar{\mathbb{B}}_{n1}(x)$ is a martingale, we have

$$(n - p_n) E[|\mathbb{B}_{n1}(x) - \bar{\mathbb{B}}_{n1}(x)|^2] \le E[|\mathbf{X}_0|^2] \int (a_n'(x - z))^2 f(z)\, dz$$

and thus, as $W(x + y) \le W(x) W(y)$ and $E[|\mathbf{X}_0|^2] \le p_n E[X_0^2]$,

$$\int W(x) E[|\mathbb{B}_{n1}(x) - \bar{\mathbb{B}}_{n1}(x)|^2]\, dx \le \frac{p_n}{n - p_n} E[X_0^2] \|(a_n')^2\|_W \|f\|_W.$$

This and (2.4) show that

$$\|\hat{\Delta}^\top (\mathbb{B}_{n1} - \bar{\mathbb{B}}_{n1})\|_V = O_p(n^{-1} p_n^{1/2} q_n^{1/2} \|(a_n')^2\|_W^{1/2}).$$

This completes the proof. □

The previous two lemmas will be applied to $\hat{f}$. For $\hat{g}$ we need analogous results with $Y_j = X_j - \varepsilon_j$ and $\hat{Y}_j = X_j - \hat{\varepsilon}_j$ in place of ε_j and $\hat{\varepsilon}_j$. The corresponding processes are now

$$\mathbb{A}_{n2}(x) = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} \big(a_n(x - \hat{Y}_j) - a_n(x - Y_j)\big),$$

$$\mathbb{B}_{n2}(x) = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} \mathbf{X}_{j-1} a_n'(x - Y_j), \quad x \in \mathbb{R}.$$

Lemma 20. *Suppose (I), (Q), (R) and (F0) hold. Let $E[V(Y_1)|\mathbf{X}_0|^2] = O(p_n)$ and let $E[V^2(Y_1)]$ be finite. Then*

$$\|\mathbb{A}_{n2} + \hat{\Delta}^\top \mathbb{B}_{n2}\|_V = O_p\big(n^{-1/2-\zeta} \|a_n'\|_V\big) + O_p\big(n^{-1} p_n q_n \|a_n''\|_V\big).$$

Proof. Stationarity, finiteness of $E[V^2(Y_1)]$, and (2.1) and (9.9) give

$$\frac{1}{n - p_n} \sum_{j=p_n+1}^{n} V^2(\hat{Y}_j) = O_p(1),$$

while stationarity and $E[V(Y_1)|\mathbf{X}_0|^2] = O(p_n)$ give

$$\frac{1}{n - p_n} \sum_{j=p_n+1}^{n} V(Y_j)|\mathbf{X}_{j-1}|^2 = O_p(p_n).$$

The desired result now follows as in the proof of Lemma 18. □

For $r \geq 0$, $E[V_r^2(Y_1)]$ is finite and $E[V_r(Y_1)|\mathbf{X}_0|^2] = O(p_n)$ if ε_0 has a finite moment of order $\xi \geq \max(2r, r+2)$. This follows from Lemma 14.

Lemma 21. *Suppose (C), (I), (Q) and (R) hold. Let $\sum_{s>0} s|\varphi_s|$ be finite and r be a non-negative number. Suppose that f has a finite moment of order $\xi > 2r+3$, and that a_n' and a_n'' have finite V_{r+1}-norms. Then*

$$\|\hat{\Delta}^\top(\mathbb{B}_{n2} - E[\mathbb{B}_{n2}])\|_{V_r} = O_p\Big(n^{-1}p_n^{1/2}q_n^{1/2}\Big(\|a_n'\|_{V_{r+1}} + \|a_n''\|_{V_{r+1}}\Big)\Big).$$

Proof. Lemmas 4 and 5 imply that a_n' is V_{r+1}-Lipschitz with constant $\|a_n''\|_{V_{r+1}}$ and that $V_{r+1}a_n'$ is bounded by $\|a_n''\|_{V_{r+1}}$. We may assume that $\xi < 2r+4$. Then $\alpha = \xi - 2r - 2$ satisfies $1 < \alpha < 2$. By (2.4),

$$\|\hat{\Delta}^\top(\mathbb{B}_{n2} - E[\mathbb{B}_{n2}])\|_{V_r}^2 \leq K_\alpha |\hat{\Delta}|^2 \sum_{i=1}^{p_n} \int V_{2r+\alpha}(x) B_i^2(x)\, dx$$

with $B_i(x)$ the i-th coordinate of $\mathbb{B}_{n2}(x) - E[\mathbb{B}_{n2}(x)]$. It now follows from Lemma 25 applied with $h = a_n'$, $q = r+1$ and $p = r + \alpha - 1 < r+1$ that

$$\sum_{i=1}^{p_n} \int V_{2r+\alpha}(x)(n-p_n)E[B_i^2(x)]\, dx \leq \frac{Cp_n}{n}\|V_p a_n'\|_\infty (\|a_n'\|_{V_{r+1}} + \|a_n''\|_{V_{r+1}})$$

for some $C > 0$. The above inequalities and the rate $|\hat{\Delta}|^2 = O_p(n^{-1}q_n)$ yield the desired result. □

10. Estimating the innovation density.

In this section we study rates of convergence in L_V of the residual-based kernel estimator $\hat{f}$ and of functionals $a * \hat{f}$. We impose the following conditions on the innovation density f and the kernel k.

(F1) *The density f has finite W_α-norm for some $\alpha > 1$.*

(K1) *The kernel k is twice continuously differentiable with bounded derivatives. Moreover, k, k' and k'' have finite W_2-norms.*

Under (K1) the i-th derivative $k_{b_n}^{(i)}$ of k_{b_n} satisfies

$$\|k_{b_n}^{(i)}\|_{W_2} = O(b_n^{-i}) \quad \text{and} \quad \|(k_{b_n}^{(i)})^2\|_{W_2} = O(b_n^{-1-2i}), \quad i = 0,1,2. \tag{10.1}$$

These rates stay valid if we replace W_2 by V, VV_1, VV_2 or W_α with $\alpha < 2$. Since k' and k'' have finite VV_2-norms, $\int k'(t)\, dt = \int k''(t)\, dt = 0$, and $\int |t|V(t)V_1(t)|k_{b_n}^{(i)}(t)|\, dt = O(b_n^{1-i})$ for $i = 1,2$, Lemma 10 yields the following rates.

Lemma 22. *Let (K1) hold and let $U = VV_\beta$ with $0 \leq \beta \leq 1$. Suppose a has finite U-norm and is U-Lipschitz. Then $\|a * k_{b_n}'\|_U = O(1)$ and $\|a * k_{b_n}''\|_U = O(b_n^{-1})$.*

Theorem 5. *Suppose that (I), (Q), (R), (F0), (F1) and (K1) hold. Then*

$$\|\hat{f} - f * k_{b_n}\|_V = O_p(n^{-1/2-\zeta} b_n^{-1}) + O_p(n^{-1} p_n q_n b_n^{-2}) + O_p(n^{-1/2} b_n^{-1/2}).$$

Proof. We may assume that $\alpha \le 2$. Set $W = W_\alpha$. Let $\bar{f} = f * k_{b_n}$ and let $\tilde{f}$ denote the kernel estimator based on the actual innovations $\varepsilon_{p_n+1}, \dots, \varepsilon_n$,

$$\tilde{f}(x) = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} k_{b_n}(x - \varepsilon_j), \quad x \in \mathbb{R}.$$

Then $E[\tilde{f}(x)] = \bar{f}(x)$ and $n \operatorname{Var} \tilde{f}(x) \le k_{b_n}^2 * f(x)$, and in view of inequalities (2.4) and (2.5) (the latter applied with W in place of V) we have $nE[\|\tilde{f} - \bar{f}\|_V^2] \le K_\alpha \|k_{b_n}^2\|_W \|f\|_W = O(b_n^{-1})$. This shows that $\|\tilde{f} - \bar{f}\|_V = O_p(n^{-1/2} b_n^{-1/2})$. Thus we are left to show that

$$\|\hat{f} - \tilde{f}\|_V = O_p(n^{-1/2-\zeta} b_n^{-1}) + O_p(n^{-1} p_n q_n b_n^{-2}).$$

But this follows from Lemmas 18 and 19, applied with $a_n = k_{b_n}$, and from the rates $\|k'_{b_n}\|_V = O(b_n^{-1})$, $\|k''_{b_n}\|_V = O(b_n^{-2})$ and $\|(k'_{b_n})^2\|_{W_\alpha} = O(b_n^{-3})$ shown above. □

Theorem 6. *Suppose that (I), (Q), (R), (F0), (F1) and (K1) hold. Let $n^{-1/2} p_n q_n b_n^{-1} \to 0$. Suppose a is V-Lipschitz and $\|a^2\|_{W_\alpha}$ is finite. Then*

$$\|a * (\hat{f} - f * k_{b_n}) - \mathbb{A}_n\|_V = o_p(n^{-1/2})$$

and $n^{1/2}\mathbb{A}_n$ is tight in L_V, where

$$\mathbb{A}_n(x) = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} \big(a(x - \varepsilon_j) - E[a(x - \varepsilon_j)]\big), \quad x \in \mathbb{R}.$$

Proof. We may assume that $\alpha \le 2$. Let $\bar{f}$ and $\tilde{f}$ be as in the previous proof. It is easy to check that $a * (\tilde{f} - \bar{f}) = \mathbb{A}_n * k_{b_n}$. It follows from Lemma 12 and the finiteness of $\|a^2\|_{W_\alpha}$ and $\|f\|_{W_\alpha}$ that $n^{1/2}\mathbb{A}_n$ is tight in L_V. Then Lemma 2 gives $\|n^{1/2}(\mathbb{A}_n * k_{b_n} - \mathbb{A}_n)\|_V = o_p(1)$. In other words,

$$\|a * (\tilde{f} - \bar{f}) - \mathbb{A}_n\|_V = o_p(n^{-1/2}).$$

One verifies that

$$a * (\hat{f} - \tilde{f}) = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} \big(a_n(x - \hat{\varepsilon}_j) - a_n(x - \varepsilon_j)\big)$$

with $a_n = a * k_{b_n}$. Since $a'_n = a * k'_{b_n}$ and $a''_n = a * k''_{b_n}$, Lemma 22 yields $\|a'_n\|_V = O(1)$ and $\|a''_n\|_V = O(b_n^{-1})$. Using (2.6) and (10.1) we find $\|(a'_n)^2\|_{W_\alpha} \le \|a^2\|_{W_\alpha} \|k'_{b_n}\|_{W_\alpha}^2 = O(b_n^{-2})$. Thus Lemmas 18 and 19 yield

$$\|a * (\hat{f} - \tilde{f})\|_V = O_p(n^{-1/2-\zeta}) + O_p(n^{-1} p_n q_n b_n^{-1}) = o_p(n^{-1/2}).$$

The desired result follows from the above. □

In the above proof we have seen that $\|n^{1/2}(\mathbb{A}_n * k_{b_n} - \mathbb{A}_n)\|_V = o_p(1)$. Thus we obtain the following result.

Corollary 2. *Under the assumptions of the previous theorem we have*

$$\Big\| k_{b_n} * a * (\hat{f} - f * k_{b_n}) - \mathbb{A}_n \|_V = o_p(n^{-1/2}).$$

11. Estimating the density g.

Now we study convergence rates in L_V of the kernel estimator $\hat{g}$ based on $\hat{Y}_j = X_j - \hat{\varepsilon}_j$, $j = p_n + 1, \dots, n$, and of functionals of the form $a * \hat{g}$. Here we restrict ourselves to the case $V = V_r$ for some non-negative r. Then $W_\alpha = V_{2r+\alpha}$.

Theorem 7. *Suppose that (C), (I), (Q) and (R) hold, that $\sum_{s>0} s|\varphi_s|$ is finite, that (K1) holds with $W_2 = V_{2r+2}$ and that (F0) holds with $\xi > 2r+3$. Let f have finite V_{r+1}-variation. Then, with $m_n = n^{-1} p_n q_n b_n^{-2}$,*

$$\|\hat{g} - g * k_{b_n}\|_{V_r} = O_p(n^{-1/2-\zeta} b_n^{-1}) + O_p(m_n) + O_p(n^{-1/2} b_n^{-1/2}) + O_p(n^{-1/2} q_n^{1/2}).$$

Proof. Let $\tilde{g}$ denote the kernel density estimator based on $Y_{p_n+1}, \dots, Y_n$,

$$\tilde{g}(x) = \frac{1}{n - p_n} \sum_{j=p_n+1}^{n} k_{b_n}(x - Y_j), \quad x \in \mathbb{R}.$$

We may assume that $\alpha \le 2$. Then $\|f\|_{W_\alpha}$, $\|k^2\|_{W_\alpha}$, $\|k\|_{V_{r+1}}$ and $\sum_{s>0} s|\varphi_s|$ are finite and f has finite V_{r+1}-variation. Thus Proposition 3.3 and Corollary 5.1 in Schick and Wefelmeyer (2006) yield the rate

$$\|\tilde{g} - g * k_{b_n}\|_{V_r} = O_p(n^{-1/2} b_n^{-1/2}).$$

Let $\bar{\Gamma}_n(x) = E[\mathbf{X}_0 k'_{b_n}(x - Y_1)]$. It suffices to show

$$\|\hat{g} - \tilde{g} + \hat{\Delta}^\top \bar{\Gamma}_n\|_{V_r} = O_p(n^{-1/2-\zeta} b_n^{-1}) + O_p(n^{-1} p_n q_n b_n^{-2}), \tag{11.1}$$

$$\|\hat{\Delta}^\top \bar{\Gamma}_n\|_{V_r} = O_p(n^{-1/2} q_n^{1/2}). \tag{11.2}$$

Statement (11.1) follows from Lemmas 20 and 21 applied with $a_n = k_{b_n}$ and the rates $\|k'_{b_n}\|_{V_{r+1}} = O(b_n^{-1})$ and $\|k''_{b_n}\|_{V_{r+1}} = O(b_n^{-2})$; see (10.1) for the latter.

The i-th component of $\bar{\Gamma}_n$ is $T_i k'_{b_n}$. Thus the Cauchy–Schwarz inequality yields the bound

$$\|\hat{\Delta}^\top \bar{\Gamma}_n\|_{V_r}^2 \le |\hat{\Delta}|^2 \sum_{i=1}^{\infty} \|T_i k'_{b_n}\|_{V_r}^2.$$

Thus statement (11.2) follows from Corollary 1 and (R). □

By assumption (C) we have $N \ge 1$. Thus the following result always applies with $m = 1$.

Theorem 8. *Suppose that (C), (I) (Q), (R) and (S+) hold, that (K1) holds with* $W_2 = V_{2r+2}$ *and that (F0) holds with* $\xi > 2r+3$. *Let* f *have finite* V_{r+1}*-variation, and let* a *have finite* V_{r+1}*-variation and finite* $V_{2r+\alpha}$*-norm for some* $\alpha > 1$. *Assume the kernel* k *is of* V_r*-order* m *for some positive integer* m. *Let* $b_n^m q_n^{1/2} \to 0$ *and* $n^{-1/2} p_n q_n b_n^{-1} \to 0$. *Then, if* $N \geq m$,

$$\Big\| a * (\hat{g} - g * k_{b_n}) - \mathbb{K}_n + \sum_{i=1}^{p_n} (\hat{\varrho}_i - \varrho_i)(T_i a)' \Big\|_{V_r} = o_p(n^{-1/2})$$

and $n^{1/2}\mathbb{K}_n$ *is tight in* L_{V_r}, *where*

$$\mathbb{K}_n(x) = \frac{1}{n-p_n} \sum_{j=p_n+1}^{n} \big(a(x - Y_j) - E[a(x - Y_j)] \big), \quad x \in \mathbb{R}.$$

Proof. We may assume that $\alpha \leq 2$. Set $W = W_\alpha$. Let $\tau = \inf\{s \geq 1 : \varphi_s \neq 0\}$. Write $Y_t = Z_t + U_t$ with $Z_t = \varphi_\tau \varepsilon_{t-\tau}$ and $U_t = \sum_{s=1}^\infty d_s Z_{t-s}$, where $d_s = \varphi_{\tau+s}/\varphi_\tau$. We can express $\mathbb{K}_n = \mathbb{K}_{n1} + \mathbb{K}_{n2}$ with

$$\mathbb{K}_{n1}(x) = \frac{1}{n-p_n} \sum_{j=p_n+1}^{n} \big(a(x - Z_j - U_j) - \bar{a}(x - U_j) \big),$$
$$\mathbb{K}_{n2}(x) = \frac{1}{n-p_n} \sum_{j=p_n+1}^{n} \big(\bar{a}(x - U_j) - E[\bar{a}(x - U_j)] \big),$$

and $\bar{a}(x) = E[a(x - Z_\tau)]$. Then $\bar{a} = a * \psi$, where ψ is the density of $Z_\tau = \varphi_\tau \varepsilon_0$. Let us now show that $n^{1/2}\mathbb{K}_{n1}$ and $n^{1/2}\mathbb{K}_{n2}$ (and hence $n^{1/2}\mathbb{K}_n$) are tight in L_{V_r}. We use Proposition 5 to establish tightness of $n^{1/2}\mathbb{K}_{n1}$. The assumptions of this proposition hold with $U_{kij} = U_k - d_{k-i}Z_i - d_{k-j}Z_j$ and $c_j = |d_j|$. Indeed, (6.3) holds in view of (S+) and Remark 1, the moment assumptions on ε_0 and Lemma 14 yield that $E[W(U)]$ and $E[W(Z)]$ are finite, and so is

$$\int W(x) \operatorname{Var}(a(x - Z))\, dx \leq \|a^2\|_W E[W(Z)],$$

and (6.1) holds with $\kappa = 1$ as a is bounded and 1-Lipschitz. Thus Proposition 5 yields tightness of $n^{1/2}\mathbb{K}_{n1}$.

We use Proposition 6 to establish tightness of $n^{1/2}\mathbb{K}_{n2}$. Since a and f, and hence ψ, have finite V_{r+1}-norms and have finite V_{r+1}-variations, $\bar{a}$ has finite V_{r+1}-norm and is absolutely continuous, $\bar{a}'$ is V_{r+1}-Lipschitz and has finite V_{r+1}-norm, and $V_{r+1}\bar{a}'$ is bounded; see Lemmas 4 and 8. This, (S+) and the moment assumptions on f give the required assumptions for this proposition, and thus tightness of $n^{1/2}\mathbb{K}_{n2}$.

Let $\tilde{g}$ be as in the previous proof and set $\bar{g} = E[\tilde{g}] = g * k_{b_n}$. It is easy to check that $a * (\tilde{g} - \bar{g}) = \mathbb{K}_n * k_{b_n}$. Since $n^{1/2}\mathbb{K}_n$ is tight in L_{V_r}, we obtain from Lemma 2 that

$\|n^{1/2}(\mathbb{K}_n * k_{b_n} - \mathbb{K}_n)\|_{V_r} = o_p(1)$. In other words,

$$\|a * (\tilde{g} - \bar{g}) - \mathbb{K}_n\|_{V_r} = o_p(n^{-1/2}).$$

Next, one verifies that $a * (\hat{g} - \tilde{g})$ equals $\mathbb{A}_{n2}$ with $a_n = a * k_{b_n}$. We obtain from Lemma 22 that $\|a_n'\|_{V_{r+1}} = O(1)$ and $\|a_n''\|_{V_{r+1}} = O(b_n^{-1})$. Thus Lemmas 20 and 21 yield

$$\Big\|a * (\hat{g} - \tilde{g}) + \sum_{i=1}^{p_n} (\hat{\varrho}_i - \varrho_i) T_i a_n'\Big\|_{V_r} = O_p(n^{-1/2-\zeta}) + O_p(n^{-1} p_n q_n b_n^{-1}) = o_p(n^{-1/2}).$$

We have $T_i a_n' = T_i(a * k_{b_n}') = (T_i a) * k_{b_n}' = (T_i a)' * k_{b_n}$. Let us now show that

$$\sum_{i=1}^{\infty} \|(T_i a_n)' - (T_i a)'\|_{V_r} = \sum_{i=1}^{\infty} \|(T_i a)' * k_{b_n} - (T_i a)'\|_{V_r} = O_p(b_n^m). \tag{11.3}$$

By Lemma 15 the functions $(T_1 a)'$, $(T_2 a)', \dots$ are V_r-regular of order m with constants $L_1, L_2, \dots$ (bounded by some L), so that Lemma 10 yields

$$\|(T_i a)' * k_{b_n} - (T_i a)'\|_{V_r} \le L_i V_r(b_n) b_n^m \int V_r(t) |t^m k(t)|\, dt.$$

This is the desired result if N is finite as in this case $T_i a = 0$ for all but finitely many i. If N is infinite, then we obtain again from Lemma 15 that the Lipschitz constants are summable as we can take $L_i = \|(T_i a)^{(m+1)}\|_{V_r}$.

Note that (R) implies $\max_{1 \le i \le p_n} |\hat{\varrho}_i - \varrho_i| = O_p(q_n^{1/2} n^{-1/2})$. This, (11.3) and $b_n^m q_n^{1/2} \to 0$ yield

$$\Big\|\sum_{i=1}^{p_n} (\hat{\varrho}_i - \varrho_i)\big(T_i a_n' - (T_i a)'\big)\Big\|_{V_r} = O(b_n^m q_n^{1/2} n^{-1/2}) = o_p(n^{-1/2}).$$

Combining the above yields the desired result. □

In the above proof we have seen that $\|n^{1/2}(\mathbb{K}_n * k_{b_n} - \mathbb{K}_n)\|_{V_r} = o_p(1)$ and that $\| \sum_{i=1}^{p_n} (\hat{\varrho}_i - \varrho_i)((T_i a)' * k_{b_n} - (T_i a)')\|_{V_r} = o_p(n^{-1/2})$. Thus we obtain the following result.

Corollary 3. *Under the assumptions of the previous theorem we have*

$$\Big\|k_{b_n} * a * (\hat{g} - g * k_{b_n}) - \mathbb{K}_n + \sum_{i=1}^{p_n} (\hat{\varrho}_i - \varrho_i)(T_i a)'\Big\|_{V_r} = o_p(n^{-1/2}).$$

To simplify notation let $\gamma_i = (T_i a)'$. Let us now take a closer look at the term

$$J = \sum_{i=1}^{p_n} (\hat{\varrho}_i - \varrho_i) \gamma_i.$$

We shall first look at the case when we are dealing with a parametric model for the autoregressive coefficients, say $\varrho_i = r_i(\vartheta)$ for a differentiable function r_i defined on an open subset Θ of $\mathbb{R}^d$ for $i = 1, 2, \dots$, where ϑ is an unknown parameter. Then it is natural to take $\hat{\varrho}_i = r_i(\hat{\vartheta})$ with $\hat{\vartheta}$ a $n^{1/2}$-consistent estimator of ϑ. Let (R2) hold. Then under the assumptions of the previous theorem $\sum_{i=1}^\infty \|\gamma_i\|_{V_r}$ is finite and so is $\sum_{i=1}^\infty \|\gamma_i\|_{V_r}^2$. Thus we obtain that

$$\|J - D^\top(\hat\vartheta - \vartheta)\|_{V_r} = o_p(n^{-1/2})$$

with $D = \sum_{i=1}^\infty \dot r_i(\vartheta)\gamma_i$. It is now easy to see that $n^{1/2}D^\top(\hat\vartheta-\vartheta)$ is tight.

Let us now look at the nonparametric situation. Let $M_n = E[\mathbf{X}_0\mathbf{X}_0^\top]$. Then M_n is invertible, and the operator norm of its inverse M_n^{-1} is bounded by some C. We consider the case when $\hat\Delta - \tilde\Delta = o_p(n^{-1/2})$ with $\tilde\Delta$ as given in (1.3).

Theorem 9. *Suppose that (C), (I) and (R) hold, that a and f have finite V_{r+1}-variation, a has a finite V_{r+1}-norm and f a finite V_{r+2}-norm. Let $\hat\Delta = \tilde\Delta + o_p(n^{-1/2})$ with $\tilde\Delta$ as in (1.3). Then*

$$\Big\| n^{1/2}\sum_{i=1}^{p_n}(\hat\Delta_i - \tilde\Delta_i)\gamma_i\Big\|_{V_r} = o_p(1), \tag{11.4}$$

and $n^{1/2}\sum_{i=1}^{p_n}\tilde\Delta_i\gamma_i$ is tight in L_{V_r}.

Proof. Note that $\gamma_i = 0$ if for some τ we have $\varphi_s = 0$ for all $s > \tau$. In this case N is finite and the conclusion is obvious. Now assume that $N \ge 4$. Then, in view of Lemma 15, we have

$$\Big\|\sum_{i=1}^{p_n}(\hat\Delta_i - \tilde\Delta_i)\gamma_i\Big\|_{V_r} \le \sum_{i=1}^\infty \|\gamma_i\|_{V_r} \max_{1\le i\le p_n}|\hat\Delta_i - \tilde\Delta_i| = O_p(|\hat\Delta - \tilde\Delta|) = o_p(n^{-1/2}).$$

This proves (11.4).

In view of Lemma 16, we have

$$B_j = \sum_{i=1}^\infty \|\gamma_i^{(j)}\|_{V_{r+1}} < \infty, \quad j = 0,1,2.$$

Thus the functions γ_i and γ_i' are of finite V_{r+1}-variation bounded by B_1 and B_2, respectively, see Lemma 5. Hence, by Lemma 4, $V_{r+1}\gamma_i$ is bounded by B_1 and $V_{r+1}\gamma_i'$ is bounded by B_2, and

$$\sum_{i=1}^\infty \|\gamma_i^2\|_{W_2} \le B_1 \sum_{i=1}^\infty \|\gamma_i\|_{V_{r+1}} = B_0B_1. \tag{11.5}$$

Moreover, we have

$$\begin{aligned}\int(\gamma_i(x-t) - \gamma_i(x))^2\,dx &\le t^2\int\Big(\int_0^1 \gamma_i'(x-st)\,ds\Big)^2\,dx\\ &\le t^2\int_0^1\int(\gamma_i'(x-st))^2\,dx\,ds \le t^2\|\gamma_i'\|_\infty\|\gamma_i'\|_1\end{aligned}$$

and thus

$$\sum_{i=1}^{\infty} \int (\gamma_i(x-t) - \gamma_i(x))^2 \, dx \le t^2 B_1 B_2. \tag{11.6}$$

We can write $n^{1/2} \sum_{i=1}^{p_n} \tilde{\Delta}_i \gamma_i = n^{1/2} \tilde{\Delta}^\top \tilde{\gamma}$ where $\tilde{\gamma} = (\gamma_1, \dots, \gamma_{p_n})^\top$. The matrix $E[n\tilde{\Delta}\tilde{\Delta}^\top]$ is given by $\sigma^2 M_n^{-1}$ with σ^2 the variance of ε_0. Since the spectral norm of M_n^{-1} is bounded by some C for all n, we obtain that

$$E[n(\tilde{\Delta}^\top \tilde{\gamma}(x))^2] = E[\tilde{\gamma}^\top(x) n \tilde{\Delta}\tilde{\Delta}^\top \tilde{\gamma}(x)] = \sigma^2 \tilde{\gamma}^\top(x) M_n^{-1} \tilde{\gamma}(x) \le \sigma^2 C |\tilde{\gamma}(x)|^2.$$

Using this, (11.5) and (11.6), we derive

$$\int W_2(x) E[(n^{1/2} \tilde{\Delta}^\top \tilde{\gamma}(x))^2] \, dx \le \sigma^2 C B_0 B_1,$$
$$\int E[(n^{1/2} \tilde{\Delta}^\top (\tilde{\gamma}(x-t) - \tilde{\gamma}(x))^2] \, dx \le \sigma^2 C B_1 B_2 t^2.$$

Thus tightness of $n^{1/2} \tilde{\Delta}^\top \gamma$ follows from Theorem 4. □

12. Some proofs.

This section contains the proofs of Theorems 1 and 2 and of Lemma 13. Under (C) and (F), the density g inherits the properties of f, see Lemmas 4, 5, 8 and 14. In particular, g is bounded and $\|g\|_{V_\xi}$ is finite.

Proof of Theorem 1. Set $\bar{f} = f * k_{b_n}$, $\bar{g} = g * k_{b_n}$ and $\bar{h} = \bar{f} * \bar{g}$. Write

$$\hat{h} - h = \bar{h} - h + \bar{g} * (\hat{f} - \bar{f}) + \bar{f} * (\hat{g} - \bar{g}) + (\hat{f} - \bar{f}) * (\hat{g} - \bar{g}). \tag{12.1}$$

We study the four right-hand terms, beginning with the last. Condition (B) and Theorems 5 and 7 imply $\|\hat{f} - \bar{f}\|_{V_r} = o_p(n^{-1/4})$ and $\|\hat{g} - \bar{g}\|_{V_r} = o_p(n^{-1/4})$. Inequality (2.5) then gives

$$\|(\hat{f} - \bar{f}) * (\hat{g} - \bar{g})\|_{V_r} \le \|\hat{f} - \bar{f}\|_{V_r} \|\hat{g} - \bar{g}\|_{V_r} = o_p(n^{-1/2}). \tag{12.2}$$

An application of Corollary 3 with $a = f$ gives

$$\Big\| \bar{f} * (\hat{g} - \bar{g}) - \mathbb{F}_n + \sum_{i=1}^{p_n} (\hat{\varrho}_i - \varrho_i) \nu_i' \Big\|_{V_r} = o_p(n^{-1/2}). \tag{12.3}$$

An application of Corollary 2 with $a = g$ yields

$$\|\bar{g} * (\hat{f} - \bar{f}) - \mathbb{G}_n\|_{V_r} = o_p(n^{-1/2}). \tag{12.4}$$

Since h is V_r-smooth of order $m+1$ by Lemma 9 and k has V_r-order $m+1$, we obtain from Lemma 10 and $nb_n^{2m+2} \to 0$ that

$$\|h * k_b - h\|_{V_r} = O(b_n^{m+1}) = o(n^{-1/2}).$$

This completes the proof.

Proof of Theorem 2. The least squares estimators $\hat{\varrho}_i$ fulfill (R) and $\hat{\Delta} - \tilde{\Delta} = o_p(n^{-1/2})$; see Lemma 1 in Schick and Wefelmeyer (2007b). Theorem 9, applied with $a = f$, now implies that $n^{1/2} \sum_{i=1}^{p_n} \tilde{\Delta}_i \nu_i'$ is tight in L_{V_r}, and that $\|n^{1/2} \sum_{i=1}^{p_n} (\hat{\Delta}_i - \tilde{\Delta}_i)\nu_i'\|_{V_r} = o_p(1)$. The latter and Theorem 1 now give the stochastic expansion of Theorem 2. The sequences $n^{1/2}\mathbb{F}_n$ and $n^{1/2}\mathbb{G}_n$ are tight in L_{V_r} by Theorems 8 and 6, applied with $a = f$ and $a = g$, respectively. This establishes the desired tightness.

Proof of Lemma 13. It follows from (6.1) that

$$\int \big(a_n(x-s-t) - a_n(x-s) - a_n(x-t) + a_n(x)\big)^2\,dx \le 4B\min\{|s|^\kappa, |t|^\kappa\}$$

for all s and t in $[-1,1]$. Let $d_s(x) = a_n(x-s) - a_n(x)$ and $\bar{d}_s(x) = E[d_s(x - Z_1)]$, and set $\xi_j(x,s) = d_s(x - Z_j) - \bar{d}_s(x)$. Then we can express the i-th summand in $\mathbb{A}_n(x-s) - \mathbb{A}_n(x)$ as $\xi_i(x - U_i, s)$. The left-hand side of (5.3) can be expressed as $(1/n^2)\sum T_{ijkl}$, where

$$T_{ijkl} = \iint E[\xi_i(x-U_i,s)\xi_j(x-U_j,s)\xi_k(y-U_k,t)\xi_l(y-U_l,t)]\,dx\,dy$$

and the summation is over all four indices, each ranging from 1 to n. Since multiplication is commutative, the term T_{ijkl} does not change its value if we switch i and j or k and l. It is easy to see that $T_{ijkl} = 0$ if one index is larger than the other three indices. Since

$$\int \xi_k(y-U_k,t)^2\,dy = \int \xi_k^2(y,t)\,dy = \int (d_t(y-Z_k) - \bar{d}_t(y))^2\,dy$$

is independent of $(\xi_i(y-U_i,t), \xi_j(y-U_j,t))$ for i and j less than k, we have

$$T_{ijkk} = \int E[\xi_i(x-U_i,s)\xi_j(x-U_j,s)]\,dx \int E[\xi_k^2(y,t)]\,dy = 0, \quad i<j<k.$$

By the same argument, $T_{kkij} = 0$ for the same indices. Thus we have

$$\sum T_{ijkl} = \sum_i T_{iiii} + \sum_{i<j}(4T_{ijij} + 2T_{ijjj} + 2T_{jjij} + T_{iijj} + T_{jjii}) + \sum_{i<j<k} 4(T_{ikjk} + T_{jkik}).$$

We have

$$\int (d_t(y-Z_k) - \bar{d}_t(y))^2\,dy \le \int (2d_t^2(y-Z_k) + 2E[d_t^2(y-Z_k)])\,dy = 4\int d_t^2(x)\,dx.$$

With (6.1) we therefore get

$$\int \xi_k(y-U_k,t)^2\,dy \le 4Bt^\kappa. \tag{12.5}$$

From this we immediately obtain that each term whose four indices take on at most two distinct values, is bounded by $(4B)^2 s^\kappa t^\kappa$. This is clear for T_{iiii}, T_{iijj} and T_{jjii}, but requires an application of the Cauchy–Schwarz inequality for the other terms; for example, $T_{ijij}^2 \le$

$T_{iijj}T_{jjii}$ and $T_{ijjj}^2 \le T_{iijj}T_{jjjj}$. Now let us look at the term T_{ikjk} with $i < j < k$. Since U_{kij} is independent of (Z_i, Z_j, Z_k), we have

$$T_{ikjk} = \iint E[\xi_i(x - U_i, s)\xi_j(y - U_j, s)\Delta_{kij}(x, y)]\, dx\, dy$$

with

$$\Delta_{kij}(x, y) = \xi_k(x - U_k, s)\xi_k(y - U_k, t) - \xi_k(x - U_{kij}, s)\xi_k(y - U_{kij}, t),$$

and thus we get from the Cauchy–Schwarz inequality and (12.5) that

$$\begin{aligned} T_{ikjk}^2 &\le 2(4B)^2(st)^\kappa \iint E[(\xi_k(x - U_k, s) - \xi_k(x - U_{kij}, s))^2(\xi_j(y - U_j, t))^2]\, dx\, dy \\ &\quad + 2(4B)^2(st)^\kappa \iint E[(\xi_k(y - U_k, t) - \xi_k(y - U_{kij}, t))^2(\xi_i(x - U_i, s))^2]\, dx\, dy \\ &\le 2(4B)^3 s^\kappa t^{2\kappa} M_{kij}(s) + 2(4B)^3 s^{2\kappa} t^\kappa M_{kij}(t), \end{aligned}$$

where

$$\begin{aligned} M_{kij}(u) &= \int E[(\xi_k(x - U_k, u) - \xi_k(x - U_{kij}, u))^2]\, dx \\ &\le \int E[(d_u(x - U_k - Z_k) - d_u(x - U_{kij} - Z_k))^2]\, dx \\ &\le E\Big[\int (d_u(x - U_k + U_{kij}) - d_u(x))^2\, dx\Big] \\ &\le 4BE[\min\{|U_k - U_{kij}|^\kappa, u^\kappa\}], \quad u > 0. \end{aligned}$$

This establishes the bound $|T_{ikjk}| \le 2(4B)^2 t^{3\kappa/2}\sqrt{E[\min\{|U_k - U_{kij}|^\kappa, t^\kappa\}]}$, since $s < t$. This is also a bound for $|T_{kikj}|$. This completes the proof of the desired bound.

13. A bound.

Let U_t, $t \in \mathbb{Z}$, be independent and identically distributed random variables with finite mean. For summable coefficients $c_0, c_1, \dots$ and $d_0, d_1, \dots$, with $d_0 \ne 0$, let us consider the linear processes

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

and let us set

$$\|c\| = \sum_{j=0}^\infty |c_j|, \quad \|d\| = \sum_{j=0}^\infty |d_j|,$$

$$D_c = \sum_{j=0}^\infty (j+1)|c_j|, \quad D_d = \sum_{j=0}^\infty (j+1)|d_j|.$$

For a measurable function h and $x \in \mathbb{R}$ we define

$$K(x) = n^{-1/2} \sum_{j=1}^{n} \Big(h(x - T_j) - E[h(x - T_j)) \Big),$$

$$H(x) = n^{-1/2} \sum_{j=1}^{n} \Big(S_j h(x - T_j) - E[S_j h(x - T_j)] \Big).$$

Set

$$A(\alpha, \beta) = 2^{\beta-1}(1 + \alpha^{\beta} E[|U_0|^{\beta}]), \quad \alpha > 0,\, \beta \geq 1,$$

Schick and Wefelmeyer (2006) have shown the following two results.

Lemma 23. *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_0 has a finite moment of order $p + q_*$. Let D_d be finite. Then*

$$\int V_{p+q}(x) E[K^2(x)]\, dx \leq 8\Lambda \|V_p h\|_\infty D_d A^4,$$

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

Lemma 24. *Suppose h is bounded and 1-Lipschitz with constant L. Let D_d be finite. Then*

$$\int E[K^2(x)]\, dx \leq 4L\|h\|_\infty D_d E[|U_0|].$$

We shall now obtain similar results for the process H.

Lemma 25. *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_0 has a finite moment of order $\beta = p + 2 + q_*$. Let $D = D_c + D_d$ 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 $A = A(\alpha, p + 2 + q_*)$ with $\alpha = \max(1, 2\|c\| + 2\|d\|)$.

Proof. We can write $H(x) = n^{-1/2} \sum_{j=1}^{n} (Z_j(x) - E[Z_j(x)])$ where

$$Z_j(x) = S_j h(x - T_j), \quad x \in \mathbb{R}.$$

Now set

$$S_j^* = \sum_{s=0}^{j-1} c_s U_{j-s}, \quad \bar{S}_j = \sum_{s=j}^{\infty} c_s U_{j-s}, \quad T_j^* = \sum_{s=0}^{j-1} d_s U_{j-s}, \quad \bar{T}_j = \sum_{s=j}^{\infty} d_s U_{j-s}.$$

Then we have the bounds $|S_j^*| + |\bar{S}_j| \leq R_j$ and $|T_j^*| + |\bar{T}_j| \leq R_j$ with

$$R_j = \sum_{s=0}^{\infty} \alpha_s |U_{j-s}| \quad \text{and} \quad \alpha_j = \max(|c_j|, |d_j|).$$

Set $R = R_0$ and $U = U_0$. By Lemma 14, for every $t \in [0, \beta]$,

$$E[V_t(R_j)] \le E[V_t(R + R_j)] \le E[V_\beta(R + R_j)] \le A$$

and

$$E[V_t(\alpha U)] \le E[V_\beta(\alpha U)] \le A.$$

Next we can write

$$Z_j(x) = S_j^* h(x - T_j^* - \bar{T}_j) + \bar{S}_j h(x - T_j^* - \bar{T}_j)$$

and obtain, with $\mathcal{F}$ denoting the σ-field generated by $\{U_t : t \le 0\}$, that

$$\bar{Z}_j(x) = E(Z_j(x)|\mathcal{F}) = h_j^*(x - \bar{T}_j) + \bar{S}_j h_j(x - \bar{T}_j) \tag{13.1}$$

where h_j^* and h_j are the functions defined by

$$h_j^*(x) = E[S_j^* h(x - T_j^*)] \quad \text{and} \quad h_j = E[h(x - T_j^*)], \quad x \in \mathbb{R}.$$

These functions have finite V_q-norm and inherit the V_q-Lipschitz property of h. More precisely, with $B_j = E[V_q(T_j^*)]$ and $B_j^* = E[|S_j^*|V_q(T_j^*)]$, we obtain the bounds

$$\|h_j\|_{V_q} \le B_j \|h\|_{V_q} \quad \text{and} \quad \|h_j^*\|_{V_q} \le B_j^* \|h\|_{V_q},$$

and find that h_j is V_q-Lipschitz with constant LB_j and h_j^* is V_q-Lipschitz with constant LB_j^*. Since $|T_j^*| \le R_j$, $|S_j^*| \le R_j$ and $q \le \beta - 1$, we obtain the inequalities $B_j \le E[V_q(R_j)] \le A$ and $B_j^* \le E[V_{q+1}(R_j)] \le E[V_\beta(R_j)] \le A$. Thus, by Lemma 4.3 in Schick and Wefelmeyer (2006), we have

$$\|h_j^*(\cdot - t) - h_j^*\|_{V_q} \le B|t|V_{q_*-1}(t) \quad \text{and} \quad \|h_j(\cdot - t) - h_j\|_{V_q} \le B|t|V_{q_*-1}(t) \tag{13.2}$$

with $B = 2A\Lambda$. To simplify notation, we abbreviate S_0 by S, T_0 by T, and Z_0 by Z. Using stationarity and a conditioning argument, we obtain

$$E[H^2(x)] = \mathrm{Var}(Z(x)) + \frac{2}{n} \sum_{j=1}^{n-1} (n - j) \,\mathrm{Cov}(Z(x), \bar{Z}_j(x)).$$

Thus

$$\int V_{p+q}(x) E[H^2(x)]\, dx \le 2 \sum_{j=0}^{n-1} \Gamma_j$$

where

$$\Gamma_j = \int V_{p+q}(x) E\big[|Z(x)|\big|\bar{Z}_j(x) - E[\bar{Z}_j(x)]\big|\big]\, dx.$$

Since $V_p h$ is bounded, $V_p(x + y) \le V_p(x)V_p(y)$, $|S| \le R \le V_1(R)$ and $|T| \le R$, we derive the bound

$$V_p(x)|Z(x)| \le |S|V_p(T)V_p(x - T)|h(x - T)| \le V_{p+1}(R)\|V_p h\|_\infty, \quad x \in \mathbb{R}.$$

Thus we get

$$\Gamma_j \le \|V_p h\|_\infty \int V_q(x) E\big[V_{p+1}(R)|\bar Z_j(x) - E[\bar Z_j(x)]|\big]\,dx.$$

Using the expression (13.1) for $\bar Z_j(x)$, we obtain the bound

$$\Gamma_j \le \|V_p h\|_\infty (\Gamma_{j1} + \Gamma_{j2} + \Gamma_{j3} + A\Gamma_{j4} + A\Gamma_{j5} + A\|h_j\|_{V_q} E[|\bar S_j|]), \tag{13.3}$$

where

$$\begin{aligned}
\Gamma_{j1} &= \int V_q(x) E[V_{p+1}(R)|h_j^*(x-\bar T_j) - h_j^*(x)|]\,dx,\\
\Gamma_{j2} &= \int V_q(x) E[V_{p+1}(R)|\bar S_j||h_j(x-\bar T_j) - h_j(x)|]\,dx,\\
\Gamma_{j3} &= \int V_q(x) E[V_{p+1}(R)|\bar S_j||h_j(x)|]\,dx = \|h_j\|_{V_q} E[V_{p+1}(R)|\bar S_j|],\\
\Gamma_{j4} &= \int V_q(x) E[|h_j^*(x-\bar T_j) - h_j^*(x)|]\,dx,\\
\Gamma_{j5} &= \int V_q(x) E[|\bar S_j||h_j(x-\bar T_j) - h_j(x)|]\,dx.
\end{aligned}$$

Now use Fubini's theorem, the inequalities (13.2) and $V_{q_*-1}(\bar T_j)(1+|\bar S_j|) \le V_{q_*}(R_j)$ to obtain $\Gamma_{j1} + \Gamma_{j2} \le BE[V_{p+1}(R)V_{q_*}(R_j)|\bar T_j|] \le BE[V_{\beta-1}(R+R_j)|\bar T_j|]$ and $\Gamma_{j4} + \Gamma_{j5} \le BE[V_{q_*}(R_j)|\bar T_j|] \le BE[V_{\beta-1}(R+R_j)|\bar T_j|]$. We also have $E[|\bar S_j|] \le E[V_{p+1}(R_j)|\bar S_j|] \le E[V_{\beta-1}(R+R_j)|\bar S_j|]$ and $\|h_j\|_{V_q} \le A\|h\|_{V_q} \le B$. Plugging these inequalities into (13.3) yields

$$\Gamma_j \le \|V_p h\|_\infty (1+A) BE[V_{\beta-1}(R+R_j)(|\bar T_j| + |\bar S_j|)].$$

Using (2.1) with $V = V_{\beta-1}$, the independence of $U_{ij} = (\alpha_i + \alpha_{i+j})|U_{-i}|$ and $R + R_j - U_{ij}$ for $i \ge 0$, and the inequalities $\alpha_i + \alpha_{i+j} \le \alpha$ and $0 \le R + R_j - U_{ij} \le R + R_j$, we obtain

$$\begin{aligned}
E[V_{\beta-1}(R+R_j)|U_{-i}|] &\le E[|U_{-i}|V_{\beta-1}(U_{ij})]E[V_{\beta-1}(R+R_j-U_{ij})]\\
&\le E[V_\beta(\alpha U)]E[V_{\beta-1}(R+R_j)] \le A^2
\end{aligned}$$

and thus

$$\begin{aligned}
E[V_{\beta-1}(R+R_j)|\bar T_j| + |\bar S_j|] &\le \sum_{i=0}^\infty (|d_{i+j}| + |c_{i+j}|) E[V_{\beta-1}(R+R_j)|U_{-i}|]\\
&\le A^2 \sum_{i=0}^\infty (|d_{i+j}| + |c_{i+j}|).
\end{aligned}$$

Thus we have the bound

$$\Gamma_j \le \|V_p h\|_\infty (1+A) BA^2 \sum_{s=j}^\infty (|c_s| + |d_s|), \quad j \ge 0.$$

Since $B = 2\Lambda A$, $A \leq A^2$, and

$$\sum_{j=0}^{\infty}\sum_{s=j}^{\infty}(|c_s| + |d_s|) = \sum_{j=0}^{\infty}(j+1)(|c_j| + |d_j|) = D,$$

the desired result follows. □

References

[1] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, Chichester.

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

[3] Chanda, K. C. (1983). Density estimation for linear processes. *Ann. Inst. Statist. Math.* **35**, 439–446.

[4] Coulon-Prieur, C. and Doukhan, P. (2000). A triangular central limit theorem under a new weak dependence condition. *Statist. Probab. Lett.* **47**, 61–68.

[5] Du, J. and Schick, A. (2007). Root-n consistency and functional central limit theorems for estimators of derivatives of convolutions of densities. *Internat. J. Statist. Management Syst.* **2**, 67–87.

[6] Frees, E. W. (1994). Estimating densities of functions of observations. *J. Amer. Statist. Assoc.* **89**, 517–525.

[7] Giné, E. and Mason, D. (2007a). On local U-statistic processes and the estimation of densities of functions of several variables. To appear in: *Ann. Statist.* **35**.

[8] Giné, E. and Mason, D. (2007b). Laws of the iterated logarithm for the local U-statistic process. To appear in: *J. Theoret. Probab.*

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

[10] Hallin, M., Lu, Z. and Tran, L. T. (2001). Density estimation for spatial linear processes. *Bernoulli* **7**, 657–668.

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

[12] Honda, T. (2000). Nonparametric density estimation for a long-range dependent linear process. *Ann. Inst. Statist. Math.* **52**, 599–611.

[13] Ledoux, M. and Talagrand, M. (1991). *Probability in Banach Spaces. Isoperimetry and Processes.* Ergebnisse der Mathematik und ihrer Grenzgebiete (3) 23, Springer, Berlin.

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

[15] Saavedra, A. and Cao, R. (1999). Rate of convergence of a convolution-type estimator of the marginal density of an MA(1) process. *Stochastic Process. Appl.* **80**, 129–155.

[16] Saavedra, A. and Cao, R. (2000). On the estimation of the marginal density of a moving average process. *Canad. J. Statist.* **28**, 799–815.

[17] Schick, A. and Wefelmeyer, W. (2004a). Root n consistent and optimal density estimators for moving average processes. *Scand. J. Statist.* **31**, 63–78.

[18] Schick, A. and Wefelmeyer, W. (2004b). Root n consistent density estimators for sums of independent random variables. *J. Nonparametr. Statist.* **16**, 925–935.

[19] Schick, A. and Wefelmeyer, W. (2004c). Functional convergence and optimality of plug-in estimators for stationary densities of moving average processes. *Bernoulli* **10**, 889–917.

[20] Schick, A. and Wefelmeyer, W. (2006). Convergence rates in weighted L_1 spaces of kernel density estimators for linear processes. Technical Report, Department of Mathematical Sciences, Binghamton University.

[21] Schick, A. and Wefelmeyer, W. (2007a). Root-n consistent density estimators of convolutions in weighted L_1-norms. *J. Statist. Plann. Inference* **37**, 1765–1774.

[22] Schick, A. and Wefelmeyer, W. (2007b). Uniformly root-n consistent density estimators for weakly dependent invertible linear processes. *Ann. Statist.* **35**, 815–843.

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

[24] van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes. With Applications to Statistics.* Springer Series in Statistics. Springer, New York.

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

Anton Schick, Binghamton University, Department of Mathematical Sciences, Binghamton, NY 13902-6000, USA

Wolfgang Wefelmeyer, Mathematisches Institut, Universität zu Köln, Weyertal 86-90, 50931 Köln, Germany