Statistical Inference for Stochastic Processes **3:** 31–40, 2000.

Marcinkiewicz–Zygmund Strong Laws for Infinite Variance Time Series

SANA LOUHICHI[1] and PHILIPPE SOULIER[2]
[1]*Université de Paris-Sud, Laboratoire de modélisation stochastique, 91405 Orsay Cedex, France*
[2]*Université d'Evry Val d'Essonne, Département de Mathématiques, 91025 Evry Cedex, France*

Abstract. We study rate of convergence in the strong law of large numbers for finite and infinite variance time series in both contexts of weak and strong dependence.

Key words: Marcinkiewicz–Zygmund strong laws, associated sequences, linear sequences, long-range dependence.

1. Introduction

For a centered stationary time series $(X_k)_{k\in\mathbb{Z}}$ with finite moment of order $q \in [1, 2]$, a rate of convergence in the strong law of large numbers (SLLN) is given by any integer $p < q$ such that $n^{-1/p}S_n$ converges almost surely to 0, with $S_n = X_1 + \cdots + X_n$. Such a result is usually called a Marcinkiewicz–Zygmund strong law. It is well known that if $(X_n)_{n\in\mathbb{Z}}$ is an i.i.d. sequence of random variables belonging to the domain of attraction of stable law of index α, $1 < \alpha \leqslant 2$, then $n^{-1/\alpha}S_n$ converges weakly to a stable law, while $n^{-1/p}S_n$ converges almost surely to zero for all $p < \alpha$. Related results for weakly dependent variables are known, see for instance Rao (1995) and the references therein, but this problem has not been studied in the context of long-range dependence.

The classical definition of long-range dependence of a time series is in terms of the sum of the autocovariances. If the sum diverges then the series is said long-range dependent. For infinite variance time series, this definition is inappropriate. The problem of defining long-range dependence for infinite variance time series is made even more ambiguous because of the fact that there is not a unique structure that can describe such time series. Contrarily to the finite variance case, linear and harmonizable processes form two disjoint classes of processes and their union does not cover all infinite variance processes (see definitions in Section 2). As suggested by Hall (1997), long-range dependence should be considered in view of a specific convergence problem, and a time series should be declared long-range dependent if the convergence rate in the problem of interest is strictly slowler than in the case of independent data.

In view of this, we propose the following definition of long-range dependence based on the rate of convergence in the SLLN. Let $(X_k)_{k\in\mathbb{Z}}$ be a stationary sequence of centered random variables with finite moment of order $q \in [1, 2]$. The sequence $(X_k)_{k\in\mathbb{Z}}$ is

- weakly dependent if $n^{-1/p} S_n$ converges almost surely to 0 for all $0 < p \leqslant q$ and $p < 2$;
- strongly or long-range dependent otherwise.

The interest of this definition is that it does not depend on a second order strucutre for the process $(X_k)_{k\in\mathbb{Z}}$ and that it is coherent with known weak convergent results as shown in Section 3.

The rest of the paper is organized as follows. Our results for almost sure convergence are stated in Section 2. We give results for three classes of processes, namely linear with symetric α-stable innovation, harmonizable and associated processes. Our results are new in the case of linear and associated processes, while in the case of harmonizable processes, they are a consequence of earlier results of Houdre (1995). The optimality of the rates of convergence obtained is then evaluated by comparison with known weak convergence results (Section 3). Proofs are postponed in Section 4.

2. Marcinkiewicz–Zygmund Strong Laws

2.1. LINEAR PROCESSES

Let $(\xi_j)_{j\in\mathbb{Z}}$ be a sequence of random variables which are either i.i.d. symmetric α-stable ($S\alpha S$) variables with $1 < \alpha < 2$, or uncorrelated with finite variance. The latter case will be refered to as the finite variance case or the case $\alpha = 2$ for convenience. Recall that the characteristic function of an $S\alpha S$ random variable ξ is $\mathbb{E}(e^{it\xi}) = e^{-\sigma^\alpha |t|^\alpha}$ for some $\sigma > 0$, and for all $p < \alpha$, $\mathbb{E}(|\xi|^p) < \infty$. In the sequel, we assume that $\sigma = 1$. Let $(b_j)_{j\in\mathbb{Z}}$ be a sequence of real numbers such that $\sum_{j\in\mathbb{Z}} |b_j|^\alpha < \infty$. Then we can define a stationary process $(X_k)_{k\in\mathbb{Z}}$ by

$$X_k = \sum_{j\in\mathbb{Z}} b_{k-j}\xi_j.$$

THEOREM 1. *Assume that there exists a real* $s \in [1, \alpha[$ *such that* $\sum_{j\in\mathbb{Z}} |b_j|^s < \infty$. *Then for all* p *such that* $1/p > 1 - 1/s + 1/\alpha$, $n^{-1/p} S_n$ *converges almost surely to zero.*

2.2. HARMONIZABLE PROCESSES

For $S\alpha S$ processes with $\alpha < 2$, the class of harmonizable processes and the class of processes having a moving average representation studied in the previous section are disjoint. Let $(X_k)_{k\in\mathbb{Z}}$ be an harmonizable process defined as follows

$$X_k = \int_{-\pi}^{\pi} c(x)e^{\mathrm{i}kx} M(\mathrm{d}x), \tag{1}$$

where c is a function such that $\int_{-\pi}^{\pi} |c(x)|^\alpha \,\mathrm{d}x < \infty$. If $1 < \alpha < 2$, we assume that M is an α-stable independently scattered random measure with Lebesgue control measure, i.e. for all function g such that $\int_{-\pi}^{\pi} |g(x)|^\alpha \,\mathrm{d}x < \infty$,

$$\mathbb{E}\left(\exp\left(\mathrm{i}t\int_{-\pi}^{\pi} g(x)M(\mathrm{d}x)\right)\right) = \exp\left(-|t|^\alpha \int_{-\pi}^{\pi} |g(x)|^\alpha \,\mathrm{d}x\right),$$

and $M(A)$ and $M(B)$ are independent whenever A and B are disjoint Borel sets. The function $f(x) = |c(x)|^\alpha$ is usually called the spectrum of the sequence $(X_k)_{k\in\mathbb{Z}}$. In the finite variance case, we only assume that M is an orthogonally scattered random measure with Lebesgue control measure, i.e. for all function g such that $\int_{-\pi}^{\pi} |g(x)|^2 \,\mathrm{d}x < \infty$,

$$\mathbb{E}\left[\left(\int_{-\pi}^{\pi} g(x)M(\mathrm{d}x)\right)^2\right] = \int_{-\pi}^{\pi} |g(x)|^2 \,\mathrm{d}x,$$

and $M(A)$ and $M(B)$ are uncorrelated whenever A and B are disjoint Borel sets. Note that in that case the process X is weakly stationnary with spectral density $f(x) = |c(x)|^2$. The following theorem is a particular case of Corollary 3.7 of Houdre (1995).

THEOREM 2. *Assume that there exists a real* $r > 1$ *such that* $f \in L^r([-\pi,\pi], \mathrm{d}x)$. *Then for all* p *such that* $1/p > 1 - (1 - 1/r)/\alpha$, $n^{-1/p}S_n$ *converges almost surely to zero.*

2.3. ASSOCIATED SEQUENCES

A finite collection $X_1, \dots, X_n$ of random variables is associated if for any coordinatewise nondecreasing functions $f,\ g\colon \mathbb{R}^n \to \mathbb{R}$

$$\mathrm{cov}(f(X_1,\dots,X_n), g(X_1,\dots,X_n)) \geqslant 0,$$

whenever the above covariance exists. An infinite collection is associated if every finite sub-collection is associated (Esary et al., 1967). Such an infinite collection will be called an associated process. For associated sequence $(X_i)_{i\in\mathbb{Z}}$, define

$$H_{i,j}(x, y) := P(X_i \geqslant x, X_j \geqslant y) - P(X_i \geqslant x)P(X_j \geqslant y). \tag{2}$$

It follows from the association property that $H_{i,j}(x, y) \geqslant 0$ for all $x,\ y$ in $\mathbb{R}$. For $v > 0$, define $g_v(u) := (u \wedge v) \vee (-v)$ and

$$G_r(v) := \mathrm{cov}(g_v(X_1), g_v(X_r)) = \int_{|x|\leqslant v}\int_{|y|\leqslant v} H_{1,r}(x, y)\,\mathrm{d}x\,\mathrm{d}y. \tag{3}$$

The latter equality is a consequence of Remark 4 after Theorem 2.3 in Yu (1993).

THEOREM 3. *Let* $(X_n)_{n\in\mathbb{Z}}$ *be a centered stationary associated process. Assume that for some* $p \in [1, 2[$,

$$\sum_{r=2}^{\infty} \int_{r^{1/p}}^{+\infty} v^{p-3} G_r(v)\,\mathrm{d}v < \infty. \tag{4}$$

Then the following statements are equivalent:

(i) $\mathbb{E}|X_1|^p < \infty$;
(ii) *for all* $\epsilon > 0$, $\sum_{n=1}^{\infty} n^{-1} P(\max_{1\leqslant j\leqslant n} |S_j| > \epsilon n^{1/p}) < \infty$.

Under (4) and either (i) or (ii), it thus holds that $\lim_{n\to\infty} n^{-1/p} S_n = 0$, *almost surely.*

Note that this theorem is an extension of the result of Baum and Katz (1965) written for i.i.d. sequences (which are also associated by (P_4) of Esary et al. (1967)). We now consider some particular cases of Theorem 3. If the associated sequence $(X_k)_{k\in\mathbb{Z}}$ has a finite moment of order 2, letting v tend to infinity in the RHS of (3) implies that $G_r(v) \leqslant \mathrm{cov}(X_1, X_r)$, for all $v > 0$, and Theorem 3 yields the rate of convergence in the strong law of Birkel (1989):

COROLLARY 1. *Let p be a fixed real number in* $[1, 2[$. *Let* $(X_n)_{n\in\mathbb{N}}$ *be a stationary associated process such that* $\mathbb{E}[X_1^2] < \infty$ *and* $\mathbb{E}[X_1] = 0$. *If*

$$\sum_{r=2}^{\infty} r^{1-2/p} \mathrm{cov}(X_1, X_r) < \infty, \tag{5}$$

then $\lim_{n\to\infty} n^{-1/p} S_n = 0$ *almost surely.*

Consider now a strongly uniformly integrable i.i.d. sequence $(X_n)_{n\in\mathbb{Z}}$ (see Billingsley, 1968, page 32), i.e. assume that there exists a real $\alpha \in]1, 2[$ and a constant $K > 0$ such that

$$P(|X_1| \geqslant x) \leqslant K x^{-\alpha}. \tag{6}$$

We can then define, following in Dabrowski and Jakubowski (1994), Remark 2.9 and Proposition 2.10, for all $A > 0$ the following dependence coefficient

$$I_\alpha^A(X_i, X_j) = \sup_{a\geqslant A} a^{\alpha-2} \int_{-a}^{a} \int_{-a}^{a} H_{i,j}(x, y)\,\mathrm{d}x\,\mathrm{d}y.$$

Since for all $v > 1$, $G_r(v) \leqslant v^{2-\alpha} I_\alpha^1(X_1, X_r)$, Theorem 3 yields:

COROLLARY 2. *Let* $(X_n)_{n\in\mathbb{Z}}$ *be a centered sequence of stationary associated random variables fulfilling Condition (6) for some fixed* $\alpha \in]1,2[$. *If, for some* $p \in [1,\alpha[$,

$$\sum_{r=2}^{\infty} r^{1-\alpha/p} I_\alpha^1(X_1, X_r) < \infty, \tag{7}$$

then $\lim_{n\to\infty} n^{-1/p} S_n = 0$, *almost surely.*

3. Weak Convergence

In order to assess the optimality of the previous results, we compare them with related results on weak convergence. If for a centered sequence (X_n) we know that $n^{-1/q}\sum_{k=1}^n X_k$ converges weakly, we will say that our almost sure convergence results are optimal if we prove that for all $p < q$, $n^{-1/p}\sum_{k=1}^n X_k$ converges almost surely to zero.

3.1. LINEAR SEQUENCES

Let $(\xi_j)_{j\in\mathbb{Z}}$ be an i.i.d. sequence of $S\alpha S$ $(1 < \alpha < 2)$ random variables. Define $c_j = |j|^{-\gamma}$ $(j \neq 0)$ and $c_0 \in \mathbb{R}$, and $X_k = \sum_{j\in\mathbb{Z}} c_{k-j}\xi_j$. The sequence X_n is well defined if $\gamma > 1/\alpha$. If moreover $\gamma < 1$, Theorem 5.1 in Kasahara and Maejima (1988) yields $n^{\gamma-1-1/\alpha}\sum_{k=1}^n X_k$ converges weakly to a $S\alpha S$ distribution. Since in that case $\sum_{j\in\mathbb{Z}} |c_j|^s < \infty$ for any $s > 1/\gamma$, Theorem 1 shows that $n^{-1/p}\sum_{k=1}^n X_k$ converges almost surely to zero for any p such that $1/p > 1-\gamma+1/\alpha$. Hence Theorem 1 is optimal.

3.2. HARMONIZABLE PROCESSES

Let the function c of Section 2.2 be defined as $c(x) = \sum_{j\in\mathbb{Z}} c_j e^{ijx}$ where c_j is the sequence defined above and let (Z_k) be the harmonizable process of Section 2.2. Then Theorem 6.1 of Cambanis and Maejima (1989) yields that $n^{\gamma+1/\alpha-2}\sum_{k=1}^n Z_k$ converges weakly to a $S\alpha S$ distribution. The function c is regularly varying at zero and $|c(x)| \approx c_\gamma |x|^{\gamma-1}$ in a neighborhood of zero (cf. Zygmund) and thus $f = |c|^\alpha \in L^r$ if $r < 1/\alpha(1-\gamma)$. Theorem 2 yields that $n^{-1/p}\sum_{k=1}^n Z_k$ converges almost surely to 0 if $1/p > 2-\gamma-1/\alpha$. Hence Theorem 2 is also optimal.

3.3. ASSOCIATED SEQUENCES

In the finite variance case, our condition for almost sure convergence is (5): $\sum_{k\geqslant 2} k^{1-2/p}\mathrm{cov}(X_1, X_k) < \infty$. If $X_n = \sum_{j\geqslant 0} c_j\xi_{n-j}$ is a linear process with $c_j = j^{-\gamma}$, $j \geqslant 1$ and $\gamma > 1/2$, then Condition (5) holds if $1/p > 3/2-\gamma$ whereas weak convergence holds for $n^{-3/2+\gamma}\sum_{k=1}^n X_k$ (Davydov, 1970). If $\sigma^2 =$

$\mathbb{E}(X_1^2) + 2\sum_{r=2}^{\infty} \operatorname{cov}(X_1, X_r) < \infty$, then (5) holds for any $p < 2$ and weak convergence to a Gaussian distribution with variance σ^2 holds (Newman, 1980). Thus in the finite variance case, our results are optimal. Note that these results are also a consequence of Theorem 1.

In the infinite variance case, the optimality of our result is more difficult to assess. Theorem 2.8 in Dabrowski and Jakubowski (1994) states that if the stationary associated centered sequence $(X_n)_{n\in\mathbb{N}^*}$ belongs to the domain of strict normal attraction of a jointly α-stable stationary sequence (cf. definition in Dabrowski and Jakubowski, 1994, p. 4–5), $(Y_n)_{n\in\mathbb{N}^*}$, then Condition (6) holds and $n^{-1/\alpha}S_n$ converges weakly to a nondegenerate strictly α stable distribution as soon as $\sum_{r=2}^{\infty} I_\alpha^1(X_1, X_r) < \infty$. In that case, Corollary 2 yields that $n^{-1/p}S_n$ converges almost surely to zero for all $p < \alpha$. If the series $\sum_{r=2}^{\infty} I_\alpha^1(X_1, X_r)$ is divergent, we do not know when weak convergence of S_n, properly renormalized, holds. If we compare with results for linear processes with regularly varying coefficients, then Condition (4) or (7) do not yield the exact rate of convergence in the strong law of large numbers. In view of the finite variance case and of the case when the series $\sum_{r=2}^{\infty} I_\alpha^1(X_1, X_r)$ is convergent, we nevertheless conjecture that Condition (4) cannot be improved and give the exact rate of convergence for associated sequences.

4. Proofs

Proof of Theorem 1. The proof of this theorem is based on a maximal inequality which is an extension of the Rademacher–Menčov inequality.

PROPOSITION 1. *Let $(a_i)_{i\in\mathbb{Z}}$ be a sequence of real numbers. If the assumptions of Theorem 1 hold, then for any real $p \in]s, \alpha[$,*

$$\mathbb{E}\left(\max_{1\leqslant k\leqslant n}\left|\sum_{i=1}^{k} a_i X_i\right|^p\right) \leqslant C_{p,\alpha}\left(\sum_{j\in\mathbb{Z}} |b_j|^s\right)^{p/s} \log_2(2n) n^{p(1-1/s)} \sum_{i=1}^{n} |a_i|^p. \quad (8)$$

Proof. Recall that $\sum_{j\in\mathbb{Z}} |b_j|^s < \infty$ and let $1/t = 1 - 1/s$, with the convention that $t = \infty$ and $1/t = 0$ if $s = 1$. In the case $\alpha < 2$, the random variables ξ_j are assumed to be i.i.d. $S\alpha S$, thus there exists a constant $C_{p,\alpha}$ (cf. for instance Samorodnitsky and Taqqu, 1994, Property 1.2.17) such that

$$\mathbb{E}\left|\sum_{k=1}^{n} a_k X_k\right|^p = C_{p,\alpha}\left(\sum_{j\in\mathbb{Z}}\left|\sum_{k=1}^{n} a_k b_{k-j}\right|^\alpha\right)^{p/\alpha} \leqslant C_{p,\alpha}\sum_{j\in\mathbb{Z}}\left|\sum_{k=1}^{n} a_k b_{k-j}\right|^p.$$

In the finite variance case, since $p/2 < 1$, applying Jensen inequality and the uncorrelatedness of the ξ_j's, we get

$$\mathbb{E}\left|\sum_{k=1}^{n} a_k X_k\right|^p \leqslant \left(\sum_{j\in\mathbb{Z}}\left|\sum_{k=1}^{n} a_k b_{k-j}\right|^2\right)^{p/2} \leqslant \sum_{j\in\mathbb{Z}}\left|\sum_{k=1}^{n} a_k b_{k-j}\right|^2 .$$

Applying first Hölder and then Jensen inequalities yields

$$\begin{aligned}\left|\sum_{k=1}^{n} a_k b_{k-j}\right|^p &\leqslant n^{p/t}\left(\sum_{k=1}^{n}|a_k|^s|b_{k-j}|^s\right)^{p/s}\\ &\leqslant n^{p/t}\sum_{k=1}^{n}|b_{k-j}|^s|a_k|^p\left(\sum_{j\in\mathbb{Z}}|b_j|^s\right)^{p/s-1}.\end{aligned}$$

Thus, setting $C_{p,2} = 1$, we obtain

$$\mathbb{E}\left|\sum_{i=1}^{n} a_i X_i\right|^p \leqslant C_{p,\alpha}\left(\sum_{j\in\mathbb{Z}}|b_j|^s\right)^{p/s} n^{p(1-1/s)}\sum_{i=1}^{n}|a_i|^p. \tag{9}$$

The maximal inequality (8) is a consequence of (9), see for instance Móricz et al. (1982).

To conclude the proof of Theorem 1, note that the maximal inequality (8) implies the almost sure convergence of the series $\sum_{k\geqslant 1} a_k X_k$ as soon as $\sum_{n\geqslant 1}|a_n|^{p'} n^{p'/t}\log_2^{p'}(2n) < \infty$ for some $p' < \alpha$ (cf. for instance Loève, 1978, Section 36.1). Finally, applying Kroneker's Lemma concludes the proof of Theorem 1. □

Proof of Theorem 3. We only prove (ii) → (i). For a proof of (i) → (ii), we refer to Proposition 1 in Louhichi (1998). Obviously (ii) yields

$$\sum_{n=1}^{\infty} n^{-1}\mathbb{P}\left(\max_{1\leqslant j\leqslant n}|X_j| > \epsilon n^{1/p}\right) < \infty, \quad \text{for all } \epsilon > 0,$$

which is equivalent to

$$\sum_{k=1}^{\infty}\mathbb{P}\left(\max_{1\leqslant j\leqslant 2^k}|X_j| > \epsilon 2^{k/p}\right) < \infty \quad \text{for all } \epsilon > 0. \tag{10}$$

We must prove that $\sum_{n=1}^{\infty}\mathbb{P}(|X_1| > n^{1/p}) < \infty$, or, equivalently,

$$\sum_{k=0}^{\infty} 2^k\mathbb{P}(|X_1| > 2^{k/p}) < \infty.$$

The following lemmas are very useful in the sequel.

LEMMA 1. *Let $(X_n)_{n\in\mathbb{Z}}$ be a sequence of identically distributed random variables. Then for any $x > 0$ and $m \in \mathbb{Z}$ it holds:*

$$\mathbb{P}(X_1 > x) = \frac{1}{m}\mathbb{P}\left(\max_{1\leqslant i\leqslant m} X_i > x\right) + \frac{1}{m}\sum_{j=2}^{m}\mathbb{P}\left(X_j > x,\ \max_{1\leqslant i<j} X_i > x\right). \tag{11}$$

The proof of this lemma is very classical and is omitted, cf. for example Equation (2.4) in Peligrad and Gut (1999). The next lemma gives a bound for the last term in Equation (11) under association.

LEMMA 2. *Let $(X_n)_{n\in\mathbb{Z}}$ be a sequence of stationary and associated sequence. For $x > 0$ recall that we defined $g_x(z) = (x\wedge z)\vee(-x)$. Then for any $a \in]0,1[$, it holds that*

$$\begin{aligned}\mathbb{P}\left(X_j > x,\ \max_{1\leqslant i<j} X_i > x\right) \ &\leqslant\ \frac{2}{(1-a)x^2}\sum_{r=2}^{j}\mathrm{cov}(g_x(X_1), g_x(X_r)) + \\ &\quad + \mathbb{P}(X_1 > ax)\mathbb{P}\left(\max_{1\leqslant i<j} X_i > x/2\right).\end{aligned}$$

Proof of Lemma 2. We generalize a result of Vronski (1999). Since the function $(x_1,\dots,x_j) \to (x_j, \max_{1\leqslant i<j} X_i)$ is coordinatewise nondecreasing, the vector $(X_j, \max_{1\leqslant i<j} X_i)$ is associated. Hence for all s,t,

$$\mathbb{P}\left(X_j > t,\ \max_{1\leqslant i<j} X_i > s\right) - \mathbb{P}(X_j > t)\mathbb{P}\left(\max_{1\leqslant i<j} X_i > s\right) \geqslant 0. \tag{12}$$

Applying (3) and (12), we now get

$$\begin{aligned}&\mathrm{cov}\left(g_x(X_j), g_x\left(\max_{1\leqslant i<j} X_i\right)\right)\\ &\quad= \int_{-x}^{x}\int_{-x}^{x}\left(\mathbb{P}\left(X_j > t,\ \max_{1\leqslant i<j} X_i > s\right) - \mathbb{P}(X_j > t)\mathbb{P}\left(\max_{1\leqslant i<j} X_i > s\right)\right)\mathrm{d}t\,\mathrm{d}s\\ &\quad\geqslant \int_{ax}^{x}\int_{x/2}^{x}\left(\mathbb{P}\left(X_j > t,\ \max_{1\leqslant i<j} X_i > s\right) - \mathbb{P}(X_j > t)\mathbb{P}\left(\max_{1\leqslant i<j} X_i > s\right)\right)\mathrm{d}t\,\mathrm{d}s\\ &\quad\geqslant \frac{x^2(1-a)}{2}\left[\mathbb{P}\left(X_j > x,\ \max_{1\leqslant i<j} X_i > x\right) - \mathbb{P}(X_j > ax)\mathbb{P}\left(\max_{1\leqslant i<j} X_i > x/2\right)\right].\end{aligned} \tag{13}$$

Since the function $h(x_1,\dots,x_{j-1}) := g_x(x_1)+\cdots+g_x(x_{j-1})-g_x(\max_{1\leqslant i<j-1} x_i)$ is coordinatewise nondecreasing, association yields

$$\mathrm{cov}(g_x(X_j), h(X_1,\dots,X_{j-1})) = \sum_{r=1}^{j-1}\mathrm{cov}(g_x(X_j), g_x(X_r)). \tag{14}$$

By stationarity, the right hand side of (14) also writes

$$\operatorname{cov}\left(g_x(X_j),\, g_x\left(\max_{1\leqslant i<j} X_i\right)\right) \leqslant \sum_{r=2}^{j} \operatorname{cov}(g_x(X_1),\, g_x(X_r)). \tag{15}$$

Finally, we combine inequalities (13) and (15) to get the desired result. We now proceed with the proof of Theorem 3. Let $p_k := 2^k \mathbb{P}(X_1 > 2^{k/p})$ and let c_p denotes a positive constant that depends only on p and may be different from line to line. Lemmas 1 and 2 applied with $x = 2^{k/p}$, $m = 2^k$ and $a = 2^{-1/p}$ yield:

$$\begin{aligned} p_k &\leqslant 2\mathbb{P}\left(\max_{1\leqslant i\leqslant 2^k} X_i > 2^{k/p-1}\right) p_{k-1} + \mathbb{P}\left(\max_{1\leqslant i\leqslant 2^k} X_i > 2^{k/p}\right) + \\ &\quad + c_p 2^{k(1-2/p)} \sum_{r=2}^{2^k} \operatorname{cov}(g_{2^{k/p}}(X_1),\, g_{2^{k/p}}(X_r)) \\ &=: 2\mathbb{P}\left(\max_{1\leqslant i\leqslant 2^k} X_i > 2^{k/p-1}\right) p_{k-1} + a_k + c_p b_k. \end{aligned} \tag{16}$$

It follows from the summability condition (10) applied with $\epsilon = 1/2$ that there exists an integer k_0 such that for all $k \geqslant k_0$,

$$\mathbb{P}\left(\max_{1\leqslant i\leqslant 2^k} X_i > 2^{k/p-1}\right) \leqslant 1/4.$$

The last inequality and (16) yield for all $k \geqslant k_0$, $p_k \leqslant p_{k-1}/2 + a_k + c_p b_k$, which implies, for all $K > k_0$,

$$\frac{1}{2}\sum_{k=k_0}^{K-1} p_k \leqslant \frac{p_{k_0-1}}{2} + \sum_{k=k_0}^{K} a_k + c_p \sum_{k=k_0}^{K} b_k. \tag{17}$$

(10) implies that $\sum_{k=1}^{\infty} a_k < \infty$ and we now prove that $\sum_{k=1}^{\infty} b_k < \infty$. Fubini's theorem and the integral representation (3) yield

$$\begin{aligned} \sum_{k=1}^{\infty} b_k &= \sum_{k=1}^{\infty} \left(2^{k(1-2/p)} \sum_{r=2}^{2^k} \operatorname{cov}(g_{2^{k/p}}(X_1),\, g_{2^{k/p}}(X_r))\right) \\ &\leqslant c_p \sum_{r=2}^{\infty} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} (r \vee |x|^p \vee |y|^p)^{1-2/p} H_{1,r}(x,\, y)\, \mathrm{d}x\, \mathrm{d}y \\ &= c_p \sum_{r=2}^{\infty} \int_{r^{1/p}}^{\infty} v^{p-3} G_r(v)\, \mathrm{d}v < \infty \end{aligned}$$

under Condition (4) (for the proof of the last equality we refer to Lemma 4 in Louhichi (1998)). Altogether, letting K tend to infinity in (17), we obtain that

$$\sum_{k=0}^{\infty} 2^k \mathbb{P}(X_1 > 2^{k/p}) = \sum_{k=0}^{\infty} p_k < \infty.$$

Since association of the sequence $(X_i)_i$ implies association for the sequence $(-X_i)_i$, we can obtain in the same way that

$$\sum_{k=0}^{\infty} 2^k \mathbb{P}(-X_1 > 2^{k/p}) < \infty.$$

This concludes the proof of Theorem 3.

References

Baum, L. E. and Katz, M.: Convergence rates in the law of large numbers, *Trans. Amer. Math. Soc.* **120** (1965), 108–123.

Billingsley, P.: *Convergence of Probability Measures*, Wiley, New York, 1968.

Birkel, T.: A note on the strong law of large numbers for positively dependent random variables, *Statist. Probab. Lett.* **7** (1989), 17–20.

Cambanis, S. and Maejima, M.: Two classes of self-similar stable processes with stationary increments, *Stoch. Proc. Appl.* **32** (1989), 305–329.

Dabrowski, A. R. and Jakubowski, A.: Stable limits for associated random variables, *Ann. Probab.* **22**(1) (1994), 1–16.

Davydov, Yu. A.: The invariance principle for stationary processes, *Theor. Probab. Appl.* **15** (1970), 487–498.

Esary, J., Proschan, F. and Walkup, D.: Association of random variables with applications, *Ann. Math. Statist.* **38** (1967), 1466–1476.

Hall, P.: On defining and measuring long-range dependence, *Fields Institute Communications* **11** (1997), 153–160.

Houdre, C.: On the almost sure convergence of series of stationary and related non stationary variables, *Ann. Probab.* **23**(3) (1995), 1204–1218.

Kasahara, Y. and Maejima, M.: Weighted sums of i.i.d. random variables attracted to integrals of stable processes, *Probab. Theory Relat. Fields* **78** (1988), 75–96.

Loève, M.: *Probability Theory*, Springer, Berlin, 1978.

Louhichi, S.: Convergence rates in the strong law for associated random variables, *Preprint*, 1998.

Móricz, F., Serfling, R. and Stout, W.: Moments and probability bounds with quasi superadditive structurefor the maximum cumulative partial sum, *Ann. Probab.* **10** (1982), 1032–1040.

Newman, C. M.: Normal fluctuations and the FKG inequalities, *Comm. Math. Phys.* **74** (1980), 119–128.

Peligrad, M. and Gut, A.: Almost sure results for a class of dependent random variables, *J. Theor. Probab.* **12** (1999), 87–104.

Rio, E.: A maximal inequality and dependent Marcinkiewicz–Zygmund strong laws, *Ann. Probab.* **23**(2) (1995), 918–937.

Samorodnitsky, G. and Taqqu, M. S.: *Stable non-Gaussian Processes: Stochastic Models with Infinite Variance*, Chapmann and Hall, New York, 1994.

Vronski, M. A.: The rate of convergence in the SLLN for associated processes and fields, *to appear in Theor. Probab. Appl.*

Yu, H.: A Glivenko-Cantelli lemma and weak convergence for empirical processes of associated sequences, *Probab. Theory Relat. Fields* **95**(3) (1993), 357–370.