PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY

A MAXIMAL $\mathbb{L}_p$-INEQUALITY FOR STATIONARY SEQUENCES AND ITS APPLICATIONS

MAGDA PELIGRAD, SERGEY UTEV, AND WEI BIAO WU

(Communicated by Richard C. Bradley)

Abstract. The paper aims to establish a new sharp Burkholder-type maximal inequality in $\mathbb{L}_p$ for a class of stationary sequences that includes martingale sequences, mixingales and other dependent structures. The case when the variables are bounded is also addressed leading to an exponential inequality for maximum of partial sums. As an application we present an invariance principle for partial sums of certain maps of Bernoulli shifts processes.

1. Introduction

In this paper we obtain a new Burkholder-type inequality for the $\mathbb{L}_p$-norm of the maximum of partial sums of stationary sequences. The bound is expressed in terms of the conditional expectation of sums with respect to an increasing field of sigma algebras, a quantity that is tractable in many examples. The method of proof is based on martingale approximation and diadic induction. We also analyze the bounded case and obtain an extension of Azuma's exponential bound to stationary and dependent sequences. When applied to mixingales the inequalities presented in this paper improve on several known results. A Markov chain example is constructed to comment on the sharpness of our $\mathbb{L}_p$ inequality. The applications contain an invariance principle for Bernoulli shifts.

Throughout the paper we consider strictly stationary sequences. Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, and $\mathbb{T} : \Omega \mapsto \Omega$ be a bijective bi-measurable transformation preserving the probability. Let $\mathcal{F}_0$ be a σ-algebra of $\mathcal{A}$ satisfying $\mathcal{F}_0 \subseteq \mathbb{T}^{-1}(\mathcal{F}_0)$, and define the nondecreasing filtration $(\mathcal{F}_i)_{i\in\mathbb{Z}}$ by $\mathcal{F}_i = \mathbb{T}^{-i}(\mathcal{F}_0)$. Let X_0 be a $\mathcal{F}_0$-measurable, centered real random variable. Define the strictly stationary sequence $(X_i)_{i\in\mathbb{Z}}$ by $X_i = X_0 \circ \mathbb{T}^i$; let $S_n = \sum_{k=1}^n X_k$ and $S_n^* = \max_{i\le n} |S_i|$. For $p \ge 2$, let $\|\cdot\|_p$ be the norm in $\mathbb{L}_p$, and $C_p \ge 2$ be the minimal constant such that

$$(1) \qquad \mathbb{E}\left[\max_{1\le k\le n} |Z_1 + \ldots + Z_k|^p\right] \le C_p n^{p/2} \|Z_1\|_p^p$$

holds for all n and stationary martingale differences $\{Z_k\}$. It is known that $C_p \le p^p$ by the Burkholder inequality (1973, 1988). For our particular setting (1) and

Received by the editors April 18, 2005.
2000 *Mathematics Subject Classification.* Primary 60F05, 60F17.
Key words and phrases. Martingale, maximal inequality, Markov chains, renewal sequences, Bernoulli shifts, invariance principle, stationary process.
The first author was supported by NSA grant.
The third author was supported by NSF grant DMS-0448704.

$p \geq 4$, a better constant $C_p \leq 4(2p)^{p/2}$ follows from Theorem 2.5 in Rio (2000) combined with the Doob maximal inequality. Maximal inequalities of type (1) play an important role in the theory of martingales.

The paper is structured as follows. Section 2 presents a generalization of (1) and some of its consequences and applications. In particular, an invariance principle is given for certain maps of Bernoulli shifts processes. Proofs are given in Section 3.

2. Main results

Recall $S_n^* = \max_{i \leq n} |S_i|$. We shall start by generalizing inequality (1) to stationary processes. To this end, we introduce the following two quantities

$$\delta_{n,p} = \sum_{j=1}^{n} j^{-3/2} \|\mathbb{E}(S_j|\mathcal{F}_0)\|_p \text{ and } \delta_{\infty,p} = \sum_{j=1}^{\infty} j^{-3/2} \|\mathbb{E}(S_j|\mathcal{F}_0)\|_p.$$

Theorem 1. *(A Maximal $\mathbb{L}_p$ inequality) Assume that $\mathbb{E}(|X_1|^p) < \infty$, $p \geq 2$. Then*

$$\|S_n^*\|_p \leq C_p^{1/p} n^{1/2} (\|X_1\|_p + 80\delta_{n,p}). \tag{2}$$

In the special case in which (X_i) are martingale differences with respect to the filter $(\mathcal{F}_i)$, then $\delta_{n,p} = 0$ and the inequality (5) reduces to the the classical one (1). It is immediate that $\|S_n^*\|_p = O(\sqrt{n})$ if $\delta_{\infty,p} < \infty$. In order to comment on the minimality of the condition $\delta_{\infty,p} < \infty$, we construct a Markov chain example, showing that, if this condition is barely altered the result fails to hold.

Proposition 1. *For any $p > 2$ and any non-negative sequence $a_n \to 0$ as $n \to \infty$ there exists a stationary ergodic discrete Markov chain $(Y_k)_{k \geq 0}$ and a functional g such that $X_i = g(Y_i)$; $i \geq 0$, $\mathbb{E}(X_1) = 0$, $\mathbb{E}(|X_1|^p) < \infty$ and*

$$\sum_{n=1}^{\infty} a_n n^{-3/2} \|\mathbb{E}(S_n|Y_0)\|_p < \infty \text{ but } \limsup_{n \to \infty} n^{-1/2} \|S_n^*\|_p = \infty. \tag{3}$$

Motivated by Maxwell and Woodroofe (2000), who proved that condition $\delta_{\infty,2} < \infty$ is enough for the central limit theorem, Peligrad and Utev (2005) established the $\mathbb{L}_2$-maximal inequality and used it to prove the invariance principle.

Since our method of proof for obtaining the maximal inequalities is based on the martingale approximations, the dominant constant in all the inequalities are going to be the best known constants for the martingales. For the bounded case the optimal constants in the $\mathbb{L}_p$-martingale inequalities imply the following extension of Azuma's exponential inequality.

Proposition 2. *For $t \geq 0$, we have*

$$\mathbb{P}(S_n^* \geq t) \leq 4\sqrt{e} \exp[-t^2/2n(\|X_1\|_\infty + 80\delta_{n,\infty})^2]. \tag{4}$$

In the stationary case this bound improves upon several corresponding inequalities established by Deddens, Peligrad and Yang (1987), Rio (2000) and also Dedecker and Prieur (2005).

We refer to mixingales as to conditions imposed to the conditional expectation of individual variables X_j. Using the triangular inequality $\|\mathbb{E}(S_k|\mathcal{F}_0)\|_p \leq \sum_{i=1}^{k} \|\mathbb{E}(X_i|\mathcal{F}_0)\|_p$ and the elementary bound $\sum_{k=j}^{\infty} k^{-3/2} \leq 3j^{-1/2}$, valid for $j = 1, 2, \ldots$, we have the following inequality for mixingales:

Corollary 1. *Assume that* $\mathbb{E}(|X_1|^p) < \infty$, $p \geq 2$. *Then*

$$\|S_n^*\|_p \leq C_p^{1/p} n^{1/2} \left[\|X_1\|_p + 240 \sum_{k=1}^{n} k^{-1/2} \|\mathbb{E}(X_k|\mathcal{F}_0)\|_p \right]. \tag{5}$$

The characteristics $\|\mathbb{E}(X_k|\mathcal{F}_0)\|_p$ can be estimated by using the covariance inequalities derived in Dedecker and Doukhan (2003), via the representation motivated by Dedecker and Rio (2000), $\|\mathbb{E}(X_k|\mathcal{F}_0)\|_p^p = \text{cov}(Y, X_k)$, where

$$Y = |\mathbb{E}(X_k|\mathcal{F}_0)|^{p-1} [I_{(\mathbb{E}(X_k|\mathcal{F}_0)>0)} - I_{(\mathbb{E}(X_k|\mathcal{F}_0)\leq 0)}] \,.$$

Using Corollary 1 and the above observation, various maximal $\mathbb{L}_p$ inequalities are easily derived for mixingale type examples, including mixing structures. In particular, our Corollary 1 allows to improve Corollary 4 in Dedecker and Prior (2005) by relaxing their condition $\sum_{k=1}^{\infty} \phi(k) < \infty$ to $\sum_{k=1}^{\infty} \phi(k)/\sqrt{k} < \infty$. We direct the reader to the book by Bradley (2002) and the paper by Dedecker and Prieur (2005) for the corresponding definitions, a large number of examples and extensive literature on the topic.

For so-called ρ-mixing sequences, defined by the coefficient

$$\rho(n) = \sup\{\text{cov}(X, Y)/(\|X\|_2 \|Y\|_2) \ : \ X \in \mathbb{L}_2(\mathcal{F}_{-\infty}^0), \ Y \in \mathbb{L}_2(\mathcal{F}_n^{\infty})\} \,,$$

we derive

Lemma 1. *Let* $\mathbb{E}(|X_1|^p) < \infty$, $p \geq 2$. *If* $\sum_{k=1}^{\infty} \rho^{2/p}(2^k) < \infty$, *then* $\delta_{\infty,p} < \infty$.

The condition $\sum_{k=1}^{\infty} \rho^{2/p}(2^k) < \infty$, or equivalently $\sum_{n=1}^{\infty} \rho^{2/p}(n)/n < \infty$, requires that $\rho(n) \to 0$ at a logarithmic rate. However, for $p > 2$ the Burkholder-type maximal inequality derived in this way is different from the Rosenthal-type inequalities established in Peligrad (1985) and Shao (1995).

If the process (X_i) is stationary causal with the form $X_i = g(\ldots, \varepsilon_{i-1}, \varepsilon_i)$, where ε_i are i.i.d. random variables and g is a measurable function for which $\mathbb{E}(X_i) = 0$ and $X_i \in \mathbb{L}^p$, then there exists a simple bound for $\|\mathbb{E}(X_k|\mathcal{F}_0)\|_p$, where $\mathcal{F}_i = (\ldots, \varepsilon_{i-1}, \varepsilon_i)$. Let (ε_i') be an i.i.d. copy of (ε_i). For $k \geq 1$, by Jensen's inequality,

$$\|\mathbb{E}(X_k|\mathcal{F}_0)\|_p \leq \|g(\mathcal{F}_k) - g(\ldots, \varepsilon_{-1}', \varepsilon_0', \varepsilon_1, \ldots, \varepsilon_{k-1}, \varepsilon_k)\|_p. \tag{6}$$

The preceding bound is tractable since it is directly related to the data generating mechanism of the process X_i. Wu and Shao (2004) showed that, for a variety of nonlinear time series, the bound in (6) is $O(r^k)$ for some $r \in (0, 1)$. See Wu and Woodroofe (2000) and Dedecker and Prieur (2005) for more discussions.

Application. Let $\{\varepsilon_k; k \in \mathbb{Z}\}$ be an i.i.d. sequence of Bernoulli variables, that is $\mathbb{P}(\varepsilon_1 = 0) = 1/2 = \mathbb{P}(\varepsilon_1 = 1)$ and let $S_n = \sum_{k=1}^{n} X_k$, where

$$X_n = g(Y_n) - \int_0^1 g(x)dx, \ Y_n = \sum_{k=0}^{\infty} 2^{-k-1} \varepsilon_{n-k}$$

and $g \in \mathbb{L}_2(0, 1)$, $(0, 1)$ being equipped with the Lebesgue measure. The transform Y_j is usually referred to as the Bernoulli shift of the i.i.d. sequence $\{\varepsilon_k; k \in \mathbb{Z}\}$ and it satisfies the recursion $Y_n = (Y_{n-1} + \varepsilon_n)/2$. To apply our maximal inequalities, it is necessary to bound $\delta_{n,p}$. In the following lemma, we relate the condition $\delta_{\infty,p} < \infty$ to the regularity properties of g.

Lemma 2. *Let $p \geq 2$ and $t > p/2 - 1$. Assume that $g \in \mathbb{L}_p(0,1)$ is a measurable function and define the Bernoulli shift process as above. Then, there exists a positive constant $c = c_{p,t}$ such that*

$$\left[\|X_1\|_p + \sum_{n=1}^{\infty} \frac{\|\mathbb{E}(X_n|Y_0)\|_p}{n^{1/2}}\right]^p \leq c \int_0^1 \int_0^1 \frac{|g(x)-g(y)|^p}{|x-y|} \left(\log \frac{1}{|x-y|}\right)^t dxdy \, . \tag{7}$$

As concrete examples of maps, we consider the example treated in Maxwell and Woodroofe (2000),

$$g(x) = x^{-\alpha} \sin(x^{-1}) \ , \ 0 < x < 1 \quad \text{and} \quad S_n(\alpha) = \sum_{k=1}^{n} X_k \ \ , 0 \leq \alpha < 1/2. \tag{8}$$

Maxwell and Woodroofe (2000) proved that for each $0 \leq \alpha < 1/2$, the normalized sums $S_n(\alpha)/\sqrt{n}$ weakly converges to the normal distribution. We strengthen the result by showing that $S_n(\alpha)/\sqrt{n}$, $\alpha \in [0, b]$ converges as a process in $\mathbb{C}[0, b]$ for each $b < 1/2$.

Proposition 3. *Let $S_n(\alpha)$ be defined by* (8) *and $b \in (0, 1/2)$. Then there exist real positives $p > 2$ and $\gamma > 1$, and a positive constant $c = c_{b,p,\gamma}$ such that*

$$n^{-p/2}\mathbb{E}\left[\max_{1\leq i\leq n} |S_i(\alpha) - S_i(\beta)|^p\right] \leq c|\alpha - \beta|^\gamma \tag{9}$$

for all $\alpha, \beta \in [0, b]$. Consequently, $S_n(\alpha)/\sqrt{n}$ weakly converges to a non-homogeneous Brownian motion B on $\mathbb{C}[0, b]$.

3. Proofs

Proof of Theorem 1. Let n, r be integers such that $n \geq 1$, $2^{r-1} \leq n < 2^r$; let $K = 5/\sqrt{2}$. As a matter of fact, we shall prove a slightly stronger inequality

$$\|S_n^*\|_q \leq C_p^{1/q} n^{1/2}[\|X_1 - \mathbb{E}(X_1|\mathcal{F}_0)\|_p + K\Delta_{r,p}],$$
$$\text{where } \Delta_{r,p} = \Delta_{r,p}(X, \mathcal{F}) = \sum_{j=0}^{r-1} 2^{-j/2}\|\mathbb{E}(S_{2^j}|\mathcal{F}_0)\|_p. \tag{10}$$

First, we notice that $\Delta_{r,p} \leq 9(\sqrt{2}+1)\delta_{n,p}$, which follows from the proof of Lemma 3.3 in Peligrad and Utev (2005) applied to the sub-additive sequence $V_k = \|\mathbb{E}(S_k|\mathcal{F}_0)\|_p$ for $k \leq n$ and $V_k = 0$ for $k > n$. Then, Proposition 1 follows from (10) applied with the inequality

$$\|X_1 - \mathbb{E}(X_1|\mathcal{F}_0)\|_p + K\Delta_{r,p} \leq \|X_1\|_p + [K9(\sqrt{2}+1) + 1]\delta_{n,p},$$

which explains the constant $K9(\sqrt{2}+1) + 1 \leq 80$.

We prove (10) by induction on n. For $n = 1$

$$\|X_1\|_p \leq \|X_1 - \mathbb{E}(X_1|\mathcal{F}_0)\|_p + \|\mathbb{E}(X_1|\mathcal{F}_0)\|_p = \|X_1 - \mathbb{E}(X_1|\mathcal{F}_0)\|_p + \Delta_{1,p}.$$

Assume that the inequality holds up to $n - 1$ and we shall prove it for n. By the triangle inequality

$$S_n^* \leq \max_{1\leq k\leq n} \left|\sum_{i=1}^{k}[X_i - \mathbb{E}(X_i|\mathcal{F}_{i-1})]\right| + \max_{1\leq k\leq n} \left|\sum_{i=1}^{k} \mathbb{E}(X_i|\mathcal{F}_{i-1})\right|. \tag{11}$$

By the Burkholder maximal inequality (1),

$$\left\| \max_{1\le k\le n} \left| \sum_{i=1}^{k} (X_i - \mathbb{E}(X_i|\mathcal{F}_{i-1})) \right| \right\|_p \le C_p^{1/p} \sqrt{n} \|X_1 - \mathbb{E}(X_1|\mathcal{F}_0)\|_p. \tag{12}$$

To estimate the impact of the second term in (11) we start by writing $n = 2m$, or $n = 2m+1$ according to a value odd or even of n. Notice that

$$\begin{aligned} \left\| \max_{1\le k\le n} \left| \sum_{j=1}^{k} \mathbb{E}(X_i|\mathcal{F}_{i-1}) \right| \right\|_p &\le \left\| \max_{1\le k\le m} \left| \sum_{i=1}^{2k} \mathbb{E}(X_i|\mathcal{F}_{i-1}) \right| \right\|_p \\ &\quad + \left\| \max_{0\le k\le m} |\mathbb{E}(X_{2k+1}|\mathcal{F}_{2k})| \right\|_p . \end{aligned} \tag{13}$$

The second term in the right hand side of (13) is estimated in a trivial way:

$$\left\| \max_{0\le k\le m} |\mathbb{E}(X_{2k+1}|\mathcal{F}_{2k})| \right\|_p \le \sqrt[p]{m+1} \|\mathbb{E}(X_1|\mathcal{F}_0)\|_p. \tag{14}$$

For the first term in the right hand side of (13) we apply the induction hypotheses to the stationary sequence $Y_0 = \mathbb{E}(X_{-1}|\mathcal{F}_{-2}) + \mathbb{E}(X_0|\mathcal{F}_{-1})$, the sigma algebra $\mathcal{G}_0 = \mathcal{F}_{-1}$ and the operator $\mathbb{T}^2$. Notice that the new filtration becomes $\{\mathcal{G}_i : i \in \mathbb{Z}\}$ where $\mathcal{G}_i = \mathcal{F}_{2i-1}$. By the induction hypotheses we obtain

$$\left\| \max_{1\le k\le m} \left| \sum_{i=1}^{2k} \mathbb{E}(X_i|\mathcal{F}_{i-1}) \right| \right\|_p \le C_p^{1/p} m^{p/2} [\mathbb{E}|Y_1 - \mathbb{E}(Y_1|\mathcal{G}_0)|^p \ + \ K\Delta_{r-1,p}(Y,\mathcal{G})] \ .$$

Clearly, $\|Y_1 - \mathbb{E}(Y_1|\mathcal{G}_0)\|_p \le 2\|\mathbb{E}(X_1|\mathcal{F}_{-1})\|_p \le 2\|\mathbb{E}(X_1|\mathcal{F}_0)\|_p$ and also

$$\begin{aligned} \Delta_{r-1,p}(Y,\mathcal{G}) &= \sum_{j=0}^{r-2} 2^{-j/2} \|\mathbb{E}(\sum_{k=1}^{2^{j+1}} Y_j|\mathcal{G}_0)\|_p \\ &= \sum_{j=0}^{r-2} 2^{-j/2} \|\mathbb{E}(S_{2^{j+1}}|\mathcal{F}_{-1})\|_p \\ &= \sqrt{2}[\Delta_{r,p}(X,\mathcal{F}) - \|\mathbb{E}(X_1|\mathcal{F}_0)\|_p]. \end{aligned}$$

And so

$$\begin{aligned} \left\| \max_{1\le k\le m} \left| \sum_{i=1}^{2k} \mathbb{E}(X_i|\mathcal{F}_{i-1}) \right| \right\|_p &\le C_p^{\frac{1}{p}} \{4\sqrt{m}\|\mathbb{E}(X_1|\mathcal{F}_0)\|_p \\ &\quad + \sqrt{2m} K[\Delta_{r,p}(X,\mathcal{F}) - \|\mathbb{E}(X_1|\mathcal{F}_0)\|_p]\}. \end{aligned}$$

This last relation, combined with (11), (12), (13), (14) and also with the fact that for all $m \ge 1$, we have $\sqrt[p]{m+1} \le \sqrt{m} C_p^{1/p}$, (since $C_p \ge 2$ for all $p \ge 2$) leads to

$$\begin{aligned} \|S_n^*\|_p &\le C_p^{1/p} \big\{ \sqrt{n} \|X_1 - \mathbb{E}(X_1|\mathcal{F}_0)\|_p + 5\sqrt{m}\|\mathbb{E}(X_1|\mathcal{F}_0)\|_p \\ &\quad + \sqrt{2m} K[\Delta_{r,p} - \|\mathbb{E}(X_1|\mathcal{F}_0\|_p)\big\} \\ &\le C_p^{1/p} \sqrt{n} [\|X_1 - \mathbb{E}(X_1|\mathcal{F}_0)\|_p + K\Delta_{r,p}] \end{aligned}$$

and so we take $K = 5/\sqrt{2}$ to guarantee the inequality. This completes the proof of the theorem. □

Remark 1. The proof of Theorem 1 is easily adapted to derive the following version of inequality (10), which incorporates the Burkholder quadratic variation term

$$\mathbb{E}\left[\max_{1\le i\le n}|S_i|^p\right]\le p^p\left[\left\|\left(\sum_{i=1}^n(X_i-\mathbb{E}(X_i|\mathcal{F}_{i-1}))^2\right)^{1/2}\right\|_p+\sqrt{n}(\frac{5}{\sqrt{2}}\Delta_{r,p})\right]^p,$$

where $p\ge 2$. The key change to make in the proof is to replace the right hand side of inequality (12) by the quadratic variation term and then carry on this term to the end of the proof.

Proof of Proposition 2. We can follow the proof of the $\mathbb{L}_p$ inequality from Theorem 1 for the variables in $\mathbb{L}_\infty$. Here we only point out the differences. Instead of Burkholder-type bound (1) for $\mathbb{L}_p$ moments of martingales with differences in $\mathbb{L}_p$ we shall use the following near optimal bound for martingales with differences in $\mathbb{L}_\infty$

$$\begin{aligned}\mathbb{E}\left[\max_{1\le k\le n}|Z_1+\ldots+Z_k|^p\right] &\le p^p(p-1)^{-p}\mathbb{E}[|Z_1+\ldots+Z_n|^p]\\ &\le p^p(p-1)^{-p}(\mathbb{E}|\eta|^p)n^{p/2}\|Z_1\|_\infty^p \qquad (15)\end{aligned}$$

valid for all $p\ge 2$, integers n and stationary martingale differences $\{Z_k\}$, where η is a standard normal variable. The first part is the Doob maximal inequality. The second part follows from the Eaton-Hoeffding argument (Eaton, 1974, pp. 612-613) applied together with the Bentkus martingale generalization (2004, Lemma 4.3) of the Hunt comparison (1955).

From (15), by using the approach of Proposition 1, since $p^p(p-1)^{-p}\le 4$, for $p\ge 2$, we derive the inequality $\|S_n^*\|_p\le 4^{1/p}\|\eta\|_p\sqrt{n}(\|X_1\|_\infty+80\delta_{n,\infty})$. To obtain the exponential inequality it remains to combine this bound with the last part of Theorem 2.4 in Rio (2000). □

Proof of Proposition 1. Let $(Y_k)_{k\ge 0}$ be a discrete Markov chain with the state space $\mathbb{Z}^+=\{0,1,\ldots\}$ and transition matrix $P=(p_{ij})$ given by $p_{k(k-1)}=1$ for $k\ge 1$ and $p_j=p_{0(j-1)}=\mathbb{P}(\tau=j)$, $j=1,2,\ldots$ (i.e. whenever the chain hits 0 it regenerates with the probability p_j). Notice that if $p_{n_j}>0$ along a subsequence $n_j\to\infty$; then the chain is irreducible. Moreover the stationary distribution exists if and only if $\mathbb{E}(\tau)<\infty$ and it is given by

$$\pi_j=\pi_0\sum_{i=j+1}^{\infty}p_i\ ,\ j=1,2\ldots,\ \text{ where } \pi_0=1/\mathbb{E}(\tau).$$

With this in mind, we shall construct first a stationary Markov chain $(Y_k)_{k\ge 0}$ by specifying the sequence $(p_j)_{j\ge 1}$. We consider a non-negative sequence $a_n\to 0$ and define $p_i=cj/u_j^{1+p/2}$, when $i=u_j$ for some $j\ge 1$, and $p_i=0$, elsewhere, where the sequence $\{u_k;k=1,2,\ldots\}$ of positive integers satisfies

$$u_1=1,\ u_2=2,\ u_k^4+1<u_{k+1}\ \text{ for }\ k\ge 3\ \text{ and }\ a_t\le k^{-2}\ \text{ for }\ t\ge u_k. \qquad (16)$$

Notice that $\mathbb{E}(\tau^2)<\infty$ but $\mathbb{E}(\tau^p)=\infty$. We start the chain with the stationary distribution $(\pi_j)_{j\ge 0}$. Further, we take $g(x)=I_{(x=0)}-\pi_0$, where $\pi_0=\mathbb{P}(Y_0=0)$.

The stationary sequence is defined by $X_j = I_{(Y_j=0)} - \pi_0$ so that $S_n = \sum_{j=1}^n X_j = \sum_{j=1}^n I_{(Y_j=0)} - n\pi_0$. Let us denote by

$$\nu = \min\{m \geq 1 \;:\; Y_m = 0\},\ A_n = \mathbb{E}_0(S_n),\ x \wedge y = \min(x, y),$$

where, $\mathbb{P}_0$ and $\mathbb{E}_0$ denote the probability and the expectation operator when the Markov chain is started at 0 i.e. $\mathbb{P}_0(Y_0 = 0) = 1$. We shall check first the convergence part in Proposition 1. Similarly to Proposition 3.1 in Peligrad and Utev (2005) we notice that

$$\|\mathbb{E}(S_n|Y_0)\|_p \leq \|\nu \wedge n\|_p \;+\; \max_{1\leq i\leq n} |A_i| = I_n + J_n \text{ (say), where } \|x\|_p^p = \sum_{k=0}^{\infty} |x_k|^p \pi_k.$$

Clearly, up to a positive constant, the probabilities p_i defined above are smaller than those considered in Peligrad and Utev (2005), where the case $p = 2$ is dealt with. Thus, immediately it follows that $\sum_{n=1}^\infty a_n J_n/n^{3/2} < \infty$. Furthermore, the convergence $\sum_{n=1}^\infty a_n I_n/n^{3/2} < \infty$ is proved using similar computations as in Peligrad and Utev (2005). This completes the convergence part of the example.

It remains to prove that

$$\limsup_{n\to\infty} n^{-p/2} \|S_n^*\|_p^p = \infty. \tag{17}$$

With this aim we define

$$T_0 = 0, T_k = \min\{t > T_{k-1} : Y_t = 0\},\ \tau_k = T_k - T_{k-1}\ ,k = 1, 2, \ldots.$$

Then, $\{\tau_j\}$ are independent variables equally distributed as τ [see, for example, Breiman (1968, p.146)]. Let $\xi_j = 1 - \pi_0\tau_j$ and we introduce the stopping time $\nu_n = \min\{j \geq 1 : T_j \geq n\}$. Clearly, $S_{T_k} = \sum_{j=1}^k \xi_j = k - \pi_0 T_k$, $\nu_n \leq n$, and since $\mathbb{E}(\tau) = 1/\pi_0$, it follows that $\mathbb{E}_0(\xi_1) = 0$.

Notice that, for a positive integer K

$$\max_{0\leq j\leq Kn} |S_j| \geq \max_{T_j\leq Kn} |S_{T_j}| = \pi_0 \max_{T_j\leq Kn} |T_j - \mathbb{E}(T_j)| \geq \pi_0|T_n - \mathbb{E}(T_n)|I_{(T_n\leq Kn)}.$$

Thus, in order to show (17), it is enough to prove that

$$\limsup_{n\to\infty} \frac{1}{n^{p/2}} \mathbb{E}\left[|T_n - \mathbb{E}(T_n)|^p I_{(T_n\leq Kn)}\right] = \infty. \tag{18}$$

We shall use a truncation argument. Starting with the decomposition

$$\tau_i - \mathbb{E}(\tau) = \{\tau_i I_{(\tau_i\leq\sqrt{n})} - \mathbb{E}[\tau I_{(\tau\leq\sqrt{n})}]\} - \mathbb{E}[\tau I_{(\tau>\sqrt{n})}] + \tau_i I_{(\tau_i>\sqrt{n})}$$

and letting $X_{in} = n^{-1/2}\{\tau_i I_{(\tau_i\leq\sqrt{n})} - \mathbb{E}[\tau I_{(\tau\leq\sqrt{n})}]\}$, $Y_{in} = n^{-1/2}\tau_i I_{(\tau_i>\sqrt{n})} I_{(T_n\leq Kn)}$ and $W_n = X_{1n} + \ldots + X_{nn}$, we write

$$n^{-1/2}\|[T_n - \mathbb{E}(T_n)]I_{(T_n\leq Kn)}\|_p \geq \left\|\sum_{i=1}^n Y_{in}\right\|_p - n^{1/2}\mathbb{E}[\tau I_{(\tau>\sqrt{n})}] - \|W_n\|_p.$$

Notice that $n^{1/2}\mathbb{E}[\tau I_{(\tau>\sqrt{n})}] \leq \mathbb{E}(\tau^2) < \infty$, $|X_{in}| \leq 1$ and $\mathrm{Var}(W_n) = n\mathrm{Var}(X_{in}) \leq \mathbb{E}(\tau^2) < \infty$. So, by independence and Rosenthal inequality, $\|W_n\|_p = O(1)$. Thus, the divergence (18) holds if

$$\limsup_{n\to\infty} n^{-1/2} \left\|\sum_{i=1}^n \tau_i I_{(\tau_i>\sqrt{n})} I_{(T_n\leq Kn)}\right\|_p = \infty.$$

Since $(\tau_n)_{n\geq 1}$ are independent and identically distributed and $p>2$, we can write

$$\begin{aligned}\mathbb{E}\left|\sum_{i=1}^{n}\tau_i I_{(\tau_i>\sqrt{n})}I_{(T_n\leq Kn)}\right|^p &\geq \sum_{i=1}^{n}\mathbb{E}[\tau_i^p I_{(\tau_i>\sqrt{n})}I_{(T_n\leq Kn)}]\\ &\geq n\mathbb{E}[\tau_n^p I_{(\sqrt{n}<\tau_n<n)}I_{(T_n\leq Kn)}].\end{aligned}$$

Since $\{\tau_n\leq n\}\cap\{\tau_1+\ldots+\tau_{n-1}\leq (K-1)n\}\subset\{\tau_1+\ldots+\tau_n\leq Kn\}$, by using again the independence of $(\tau_n)_{n\geq 1}$ we obtain

$$\begin{aligned}\mathbb{E}\left[\tau_n^p I_{(\tau_n>\sqrt{n})}I_{(T_n\leq Kn)}\right] &\geq \mathbb{E}\left[\tau_n^p I_{(\tau_n>\sqrt{n})}I_{(\tau_n\leq n)}I_{(\tau_1+\ldots+\tau_{n-1}\leq (K-1)n)}\right]\\ &= \mathbb{E}[\tau_n^p I_{(\tau_n>\sqrt{n})}I_{(\tau_n\leq n)}]\mathbb{P}[\tau_1+\ldots+\tau_{n-1}\leq (K-1)n].\end{aligned}$$

By the law of large numbers we have $\mathbb{P}[\tau_1+\ldots+\tau_{n-1}\leq (K-1)n]\to 1$ as $n\to\infty$ when $K-1>\mathbb{E}(\tau)$. Overall, for all n sufficiently large

$$\begin{aligned}\mathbb{E}\left[\tau_n^p I_{(\tau_n>\sqrt{n})}I_{(T_n\leq Kn)}\right] &\geq \frac{1}{2}n\mathbb{E}[\tau_n^p I_{(\tau_n>\sqrt{n})}I_{(\tau_n\leq n)}]\\ &= \frac{1}{2}n\mathbb{E}[\tau_n^p I_{(\sqrt{n}<\tau_n\leq n)}].\end{aligned}$$

So, it remains to prove that $\limsup_{n\to\infty} n^{1-p/2}\mathbb{E}[\tau^p I(\sqrt{n}<\tau\leq n)]=\infty$. To this end, we observe that along the subsequence $(u_j)_{j\geq 1}$ defined by (16) and $j\to\infty$

$$\frac{u_j}{u_j^{p/2}}\mathbb{E}[\tau^p I(\sqrt{u_j}<\tau\leq u_j)]\geq u_j^{1-p/2}u_j^p\mathbb{P}(\tau=u_j)=u_j^{1-p/2}u_j^p cju_j^{-1-p/2}=cj\to\infty.$$

Therefore we established the desired result. □

Proof of Lemma 1. We first mention the following estimate of the absolute moments of order p of the partial sums contained in papers by Peligrad (1985) or Shao (1995). For a certain constant K depending only on p and $A=\sum_{j=0}^{\infty}\rho^{2/p}(2^j)$, we have

$$\|S_n\|_p\leq Kn^{1/2}\|X\|_p \tag{19}$$

By the triangular inequality and stationarity, we derive

$$\|\mathbb{E}(S_{2n}|\mathcal{F}_0)\|_p\leq\|\mathbb{E}(S_n|\mathcal{F}_0)\|_p+\|\mathbb{E}(S_n|\mathcal{F}_{-n})\|_p \tag{20}$$

We combine now Lemma 4.3 and Theorem 4.12 in Bradley (2002) to obtain

$$\|\mathbb{E}(S_n|\mathcal{F}_{-n})\|_p\leq 2^{1-2/p}\rho^{2/p}(n)\|S_n\|_p$$

and so, by (19), $\|\mathbb{E}(S_n|\mathcal{F}_{-n})\|_p\leq c\rho^{2/p}(n)n^{1/2}$ where $c=2^{1-2/p}K^{1/p}\|X\|_p$. Thus, by iterating (20), we have $\|\mathbb{E}(S_{2^{k+1}}|\mathcal{F}_0)\|_p\leq c\sum_{i=0}^{k}2^{i/2}\rho^{2/p}(2^i)$. Hence

$$\sum_{j\geq 0}2^{-j/2}\|\mathbb{E}(S_{2^j}|\mathcal{F}_0)\|_p\leq c\sum_{j\geq 0}2^{-j/2}\sum_{i=0}^{j-1}2^{i/2}\rho^{2/p}(2^i)\leq 4c\sum_{i\geq 0}\rho^{2/p}(2^i)\ .$$

Since $\|\mathbb{E}(S_n|\mathcal{F}_0)\|_p$ is sub-additive, (see Peligrad and Utev (2005), Lemma 2.7), it follows that $\sum_r 2^{-r/2}\|\mathbb{E}(S_{2^r}|\mathcal{F}_0)\|_p<\infty$ is equivalent to $\sum_n n^{-3/2}\|\mathbb{E}(S_n|\mathcal{F}_0)\|_p<\infty$ and the result follows. □

Proof of Lemma 2. Without loss of generality assume $\mathbb{E}[g(Y_0)]=0$. Let $D_k=\{j2^{-k}:j=0,\ldots,2^k-1\}$. Following Maxwell and Woodroofe (2000), we notice

$$\mathbb{E}|\mathbb{E}(g(Y_k)|Y_0)|^p=\int_0^1\left|2^{-k}\sum_{z\in D_k}\int_0^1[g(x2^{-k}+z)-g(y2^{-k}+z)]dy\right|^p dx.$$

By Jensen's inequality and change of measure

$$\begin{aligned}
\mathbb{E}|\mathbb{E}(g(Y_k)|Y_0)|^p &\leq 2^{-k}\sum_{z\in D_k}\int_0^1\int_0^1 |g(x2^{-k}+z)-g(y2^{-k}+z)|^p dydx \\
&= 2^k\sum_{z\in D_k}\int_0^{2^{-k}}\int_0^{2^{-k}} |g(x+z)-g(y+z)|^p dydx \\
&\leq 2^k\int_0^1\int_0^1 I_{(|x-y|\leq 2^{-k})}|g(x)-g(y)|^p dydx.
\end{aligned}$$

Let $t > p/2-1$ and $q = p/(p-1)$. Then

$$Q := \sum_{k=0}^{\infty}(1+k)^t\mathbb{E}|\mathbb{E}(g(Y_k)|\mathcal{F}_0)|^p \leq \int_0^1\int_0^1 J(|x-y|)|g(x)-g(y)|^p dydx,$$

where $J(z) = \sum_{k:2^{-k}\geq z} 2^k(1+k)^t \leq Cz^{-1}[\log(1/z)]^t$ for some constant $C > 0$. Since $t > p/2-1$, $\gamma := -q(t/p+1/2) < -1$. By Hölder's inequality,

$$\begin{aligned}
\|g(Y_1)\|_p + \sum_{n=1}^{\infty} n^{-1/2}\|\mathbb{E}(X_n|\mathcal{F}_0)\|_p &\leq \sqrt{2}\sum_{k=0}^{\infty}(k+1)^{-1/2}\|\mathbb{E}(X_k|\mathcal{F}_0)\|_p \\
&\leq \sqrt{2}Q^{1/p}\left[\sum_{k=0}^{\infty}(k+1)^{\gamma}\right]^{1/q} := k_{t,p}Q^{1/p}
\end{aligned}$$

for some $k_{t,p} < \infty$. So the lemma follows. □

Proof of Proposition 3. First we notice that the weak convergence involved in this proposition will immediately follow from the central limit theorem from Maxwell and Woodroofe (2000) along with (9) and Theorem 12.3 in Billingsley (1968).

To prove (9), we apply the maximal inequality of Corollary 1 and Lemma 2 to stationary sequence $\{g_\alpha(Y_k) - g_\beta(Y_k) - \mathbb{E}[g_\alpha(Y_0) - g_\beta(Y_0)]\ ;\ k \in \mathbb{Z}\}$. We choose $2 < p \leq 3$ such that $pb < 1$ and take $\gamma = p/2$, $\varepsilon = (p/2-1)/4 > 0$ so that $p - \gamma = p/2 = 1 + 4\varepsilon > 1 + 2\varepsilon$.

To analyze the integral (7), we shall change the variables from x to $1/x$ and from y to $1/y$, and introduce the function $G(x) = (x^\beta - x^\alpha)\sin(x)$. Notice that $xy \geq |x-y|$ for $x, y \geq 1$. Then, the integral (7) is reduced to

$$J = \int\int_{x>1,y>1} |G(x)-G(y)|^p\frac{1}{xy|x-y|}\log\left(\frac{xy}{|x-y|}\right)dxdy. \tag{21}$$

Next, without loss of generality we assume that $\beta = \alpha + \delta$ with $\delta > 0$. Since $|x^\beta - x^\alpha| = x^\beta(1 - e^{-\delta\ln(x)}) \leq x^\beta\delta|\ln(x)|$, we then derive that there exists a positive constant c such that for all $x, y \geq 1$ and all α, β with $0 \leq \alpha, \beta \leq b$,

$$|G(x)-G(y)|^p \leq c\delta^\gamma\min(|x-y|^{p-\gamma}, 1)(x^{pb}[\ln(x)]^p + y^{pb}[\ln(y)]^p)$$

Since, for every $\varepsilon > 0$ there exists a positive non-decreasing function k_ε such that $\ln(x) \leq k_\epsilon x^\varepsilon$ for all $x \geq 1$, J in (21) is bounded by $2k_\varepsilon c\delta^\gamma I$, where

$$I = \int\int_{x>1,y>1}\min(|x-y|, 1)^{p-\gamma}\ \frac{x^u+y^u}{x^2y^2}\left(\frac{xy}{|x-y|}\right)^{1+\varepsilon}dxdy\ ,$$

$u = (1+pb)/2$, $0 < u < 1$ and $p - \gamma > 1 + \varepsilon$. The fact that $I < \infty$ can be shown in a similar way as it was indicated in Maxwell and Woodroofe (2000). □

Acknowledgment

The authors are grateful to the referee and the editor for helpful suggestions.

References

1. Bentkus, V. (2004). On Hoeffding's inequalities. *Ann. Probab.* **32** 1650–1673.
2. Billingsley, P. (1968). *Convergence of Probability Measures*, Wiley, New York.
3. Bradley, R. C. (2002). *Introduction to strong mixing conditions.* Volume 1, Technical Report, Department of Mathematics, Indiana University, Bloomington. *Custom Publishing of I.U.*, Bloomington.
4. Breiman, L. (1968). *Probability.* Addison–Wesley.
5. Burkholder, D. L. (1973). Distribution function inequalities for martingales. *Ann. Probab.* **1** 19–42.
6. Burkholder, D. L. (1988). Sharp inequalities for martingales and stochastic integrals. *Astrisque No. 157-158*, 75–94.
7. Dedecker, J. and Doukhan, P. (2003). A new covariance inequality and applications, *Stochastic Process. Appl.* **106** 63–80.
8. Dedecker, J., Prieur, C. (2005). New dependence coefficients. Examples and applications to statistics. *Probab. Theory Relat. Fields* **132** 203–236
9. Dedecker, J., Rio, E. (2000). On the functional central limit theorem for stationary processes. (English. English, French summary) *Ann. Inst. H. Poincaré Probab. Statist.* **36** 1–34.
10. Deddens, J., Peligrad, M. and Yang, T (1987) On strong consistency of kernel estimators under dependence assumptions. Proceedings of the Sixth Pannonian Symposium of Mathematical Statistics, (1987), D. Reidel Publishing Company, 33–43.
11. Eaton, M. L. (1974). A probability inequality for linear combinations of bounded random variables. *Ann. Statist.* **2** 609–614.
12. Hunt, G. A. (1955). An inequality in probability theory. *Proc. Amer. Math. Soc.* **6** 506–510.
13. Maxwell, M., Woodroofe, M. (2000). Central limit theorems for additive functionals of Markov chains. *Ann. Probab.* **28** 713–724.
14. Peligrad, M. (1987). The convergence of moments in the central limit theorem for ρ-mixing sequences of random variables. *Proc. Amer. Math. Soc.* **101** 142–148.
15. Peligrad, M. (1985). Convergence rates of the strong law for stationary mixing sequences. *Z. Wahrsch. Verw. Gebiete* **70** 307–314.
16. Peligrad, M. and Utev, S. (2005). A new maximal inequality and invariance principle for stationary sequences. *Ann. Probab.* **33** 798–815.
17. Rio, E. (2000). *Theorie asymptotique des processus aleatoires faiblement dependants.* Mathematiques et applications de la SMAI. 31 Springer.
18. Shao, Qi Man. (1995). Maximal inequalities for partial sums of ρ-mixing sequences. *Ann. Probab.* **23** 948–965.
19. Wu, W. B., Shao X. (2004). Limit theorems for iterated random functions. *J. Appl. Probab.* **41** 425–436.
20. Wu, W. B., Woodroofe, M. (2000): A central limit theorem for iterated random functions. *J. Appl. Probab.* **37** 748-755.

Department of Mathematical Sciences, University of Cincinnati, PO Box 210025, Cincinnati, OH 45221-0025, USA

School of Mathematical Sciences, University of Nottingham, Nottingham, NG7 2RD, ENGLAND
E-mail address: `sergey.utev@nottingham.ac.uk`

Department of Statistics, The University of Chicago, 5734 S. University Avenue, Chicago, IL 60637, USA
E-mail address: `wbwu@galton.uchicago.edu`