The Annals of Probability
1995, Vol. 23, No. 3, 1204–1218

ON THE ALMOST SURE CONVERGENCE OF SERIES OF STATIONARY AND RELATED NONSTATIONARY VARIABLES[1]

BY CHRISTIAN HOUDRÉ

Georgia Institute of Technology

Dedicated to the memory of S. Cambanis

Let $\{X_n\}$ be, for example, a weakly stationary sequence or a lacunary system with finite pth moment, $1 \le p \le 2$, and let $\{a_n\}$ be a sequence of scalars. We obtain here conditions which ensure the almost sure convergence of the series $\sum a_n X_n$. When $\{X_n\}$ is an orthonormal sequence, the classical Rademacher–Menchov theorem is recovered. This is then applied to study the strong consistency of least squares estimates in multiple regression models.

1. Introduction. We study in the present note the problem of the almost sure convergence of the partial sums $S_N = \sum_{-N}^{N} a_n X_n$, where the X_n have some dependence structure. Although this problem is very classical, much remains to be known when the sequence $\{X_n\}$ is not assumed independent. This is certainly due to the fact that dependence structures assume many forms and are thus more intractable. Let us mention some known results, related to ours, and upon which we either improve or that we just complement. The first and foremost result we have in mind is the so-called Rademacher–Menchov theorem which asserts (see, e.g., [1] and [16]) that when $\{X_n\}$ is an orthonormal sequence, $\sum |a_n|^2 \log^2(1+|n|) < \infty$ guarantees the a.s. convergence of the corresponding weighted series $\sum a_n X_n$. Of course, for some particular orthonormal sequences, for example, martingale differences, the otherwise optimal factor $\log^2(1+|n|)$ is unnecessary (throughout, log is the base 2 logarithm). Other related results we have in mind deal with S_p (lacunary) systems (see, e.g., [6, 1, 16]), with quasistationary sequences [8] or large classes of weakly dependent variables [20, 3]. Two main features of the works just mentioned are the existence of moments of order p, $2 \le p$, and/or estimates of the covariance of the random sequence. In addition to providing a single framework for the above examples, these features will be reconsidered here and, in particular, we mainly assume $1 \le p \le 2$.

Received July 1993; revised October 1994.

[1]Research supported in part by a NSF Mathematical Sciences Postdoctoral Fellowship while at Department of Statistics, Stanford University, by NSF–NATO Postdoctoral Fellowships while at CEREMADE, Université Paris Dauphine, and at CERMA, Ecole Nationale des Ponts et Chaussées, and by AFOSR contract 91-0030.

AMS 1991 *subject classifications*. Primary 60F15, 60G10; secondary 60G12, 60E07.

Key words and phrases. Stationary sequences, series, almost sure convergence, Rademacher–Menchov theorem, SαS harmonizable sequences, strong consistency, least squares, multiple regression, S_p-systems, stationary Markov chains, weak dependence, mixing.

The form of dependency studied here includes, among others, weakly stationary sequences, harmonizable SαS sequences, $1 < \alpha \leq 2$, lacunary S_p-systems, $1 \leq p \leq 2$, as well as large classes of weakly dependent or quasistationary random variables. These classes of random sequences have in common a Fourier representation and an associated "spectrum" upon which our results rely. Furthermore, only the finiteness of pth moments, $1 \leq p \leq 2$, will be needed. Even for the particular cases just mentioned, our results are often optimal. For example, [8] provides optimal conditions for the a.s. convergence of series $\sum a_n X_n$, where $\{X_n\}$ is a quasistationary sequence. The conditions given there are expressed in terms of the covariance of the process; ours are expressed in terms of its spectrum (in the weakly stationary case). As is well known, and except in some exceptional cases, for example, $L^2((-\pi, \pi])$, functions in $L^r((-\pi, \pi])$, $r \neq 2$, cannot be characterized via (the decay of) their Fourier coefficients. Hence, when reduced to weakly stationary processes, our results and those of [8] complement one another.

Let us now briefly describe the methods and content of the paper. In the next section, we review some background material and provide examples. The third section is where the main a.s. convergence results are presented. At the heart of the methods developed there is a Rademacher–Menchov-type estimate as well as a dilation argument. The relevance of these results to the ergodic theory problem of pointwise convergence of $\sum_{n=1}^{N} a_n T^n$, where T is a contraction on a Hilbert space, is also briefly indicated. More generally, we wish to show that as far as a.s. convergence or summability is concerned, classical results valid for orthonormal sequences remain true (when properly modified) for random variables with more complicated dependence structure or without finite second moment. As a sample application of our results, we also study the strong consistency of least squares estimates in multiple regression models complementing the results of [2] and [4].

2. Preliminaries. Let us recall some background elements and refer the reader to [10] and [11] for more details and precise references. The class of random sequences under consideration are the so-called (p, q)-bounded sequences defined as follows:

Let (Ω, B, P) be a probability space and let $L^p(P)$, $1 \leq p \leq 2$, be the corresponding Lebesgue spaces (with E denoting expectation). A zero mean sequence $X = \{X_n\}_{n \in \mathbf{Z}} \subset L^p(\Omega, B, P)$ is (p, q)-bounded, $1 \leq p < +\infty$, $1 \leq q \leq +\infty$, if there exists a constant $C > 0$ (throughout, C is an absolute constant whose value might change from an expression to another) such that

$$\left(E\left|\sum_{n=-N}^{N} \alpha_n X_n\right|^p\right)^{1/p} \leq C\left(\int_{-\pi}^{\pi}\left|\sum_{n=-N}^{N} \alpha_n e^{in\theta}\right|^q d\theta\right)^{1/q} \tag{2.1}$$

for every finite sequence $\alpha_{-N}, \dots, \alpha_N \in \mathbf{C}$. From (2.1), a (p, q)-bounded sequence has a spectral representation, $X_n = \int_{-\pi}^{\pi} e^{in\theta}\, dZ_X(\theta)$, $n \in \mathbf{Z}$, where Z_X: $\mathscr{B}(-\pi, \pi] \longrightarrow L^p(P)$, the "random spectrum," is σ-additive ($\mathscr{B}(-\pi, \pi]$ is the Borel σ-algebra of $(-\pi, \pi]$).

Typical examples of (p,q)-bounded processes are second order weakly stationary processes ($p=2$, $q=+\infty$) and SαS harmonizable processes ($1 \le p < \alpha$, $q=+\infty$). Under more stringent conditions on the spectrum or the control measure, a weakly stationary sequence or a SαS harmonizable sequence is (p,q)-bounded $q < +\infty$. In fact, a weakly stationary sequence (resp. a SαS harmonizable sequence) is $(2,q)$-bounded, $2 \le q < +\infty$ [resp. (p,q)-bounded $1 \le p < \alpha$, $\alpha \le q < +\infty$] if and only if its spectrum (resp. its control measure) is in $L^{q/(q-2)}((-\pi,\pi])$ [resp. $\in L^{q/(q-\alpha)}((-\pi,\pi])$]. Further examples of $(2,2)$-bounded sequences are provided by large classes of time varying linear or ARMA processes [12]. For $q=2$, (2.1) corresponds to the defining property of lacunary S_p (also known as Hilbertian or $S_{p,2}$) systems. In particular, a quasistationary sequence $\{X_n\}$, that is, $\sup_{k\in\mathbf{Z}} |EX_{n+k}\overline{X}_k| \le \Phi(n)$ for some nonnegative Φ, such that $\sum \Phi(n) < +\infty$ is $(2,2)$-bounded (see [8]). It is also clear that by the Hausdorff–Young inequality, an $S_{p,r}$ system, $p \ge 1$, $r \ge 2$, is (p,r')-bounded, $1/r + 1/r' = 1$. Less typical examples of $(2,p)$-bounded sequences are given by large classes of stationary Markov chains (essentially by the proofs in [5] pages 225–227), or by large classes of sequences of bounded random variables. Let us now present some new classes of examples stemming from estimates obtained in [20] and in [3]. Let $\{X_n\}_{n\in\mathbf{Z}} \subset L^2(\Omega,B,P)$. Let also $\mathscr{F}_S = \sigma\{X_n,\ n\in S\}$ be the σ-field generated by the variables X_n, $n\in S \subset \mathbf{Z}$, and for S, $T \subset \mathbf{Z}$, let finally

$$r(\mathscr{F}_S,\mathscr{F}_T) = \sup\{(|EXY - EXEY|)/(E|X|^2)^{1/2}E(|Y|^2)^{1/2},\quad X \in L^2(\mathscr{F}_S),\ Y\in L^2(\mathscr{F}_T)\}.$$

The *maximal correlation coefficient* $\rho(n)$ is then defined via $\rho(n) = \sup r(\mathscr{F}_S,\mathscr{F}_T)$, where the supremum is taken over all finite subsets S of $[-k,k]$ and subsets T of $(-\infty,-n-k]\cup[n+k,+\infty)$. Similarly, the *strict maximal correlation coefficient* $\tilde\rho(n)$ is defined via $\tilde\rho(n) = \sup r(\mathscr{F}_S,\mathscr{F}_T)$, where the supremum is taken over all finite subsets S and T of $\mathbf{Z}$ which are at a distance at least n from one another. It is shown in [20] that the condition $\sum \rho(2^n) < \infty$ ensures that $E|\sum_{-N}^N X_n|^2 \le CN$. A simple modification of the proof given there actually gives $E|\sum_{-N}^N \alpha_n X_n|^2 \le C\sum_{-N}^N |\alpha_n|^2$, that is, a weakly dependent sequence such that $\sum\rho(2^n) < +\infty$ is $(2,2)$-bounded. More recently, it was proved in [3] that whenever $\tilde\rho(n) < 1$ for some n, and X_n is centered with finite pth moment, $p \ge 1$, then $E|\sum_k^N X_n|^p \le CE(\sum_k^N |X_n|^2)^{p/2}$, for all $N \ge 1$ and all $0 \le k \le N$. Again, a simple modification of the proof given in [3] as well as Hölder's inequality show that $(E|\sum_{-N}^N \alpha_n X_n|^p)^{1/p} \le C(\sum_{-N}^N |\alpha_n|^2)^{1/2}$. In other words, a centered sequence $\{X_n\} \subset L^2(P)\cap L^p(P)$, $1 \le p \le 2$, such that $E|X_n|^2 \le C$, for all n, and such that $\tilde\rho(n) < 1$, for some n, is $(p,2)$-bounded.

To finish this section, we present a rather useful decomposition of (p,q)-bounded processes (see [11] for more precise references, and note the range of p and q).

LEMMA 2.1. *A sequence $\{X_n\}_{n\in\mathbf{Z}}$ is (p,q)-bounded, $1 \le p \le 2 \le q \le +\infty$, if and only if $X_n = \Lambda Q Y_n$, $n\in\mathbf{Z}$, where Λ is a random variable in $L^{2p/(2-p)}(P)$,*

Q is the orthogonal projection from $L^2(\tilde{P}) \supset L^2(P)$ onto $L^2(P)$ and $\{Y_n\}_{n\in\mathbf{Z}} \subset L^2(\tilde{P})$ is a weakly stationary $(2,q)$-bounded, $2 \le q \le +\infty$, sequence.

Lemma 2.1 is a purely geometric result, which a priori has little to do with a.s. convergence. This lemma is, however, a crucial ingredient in extending the classical results.

3. Development. We now have the necessary preliminaries to prove the following theorem.

THEOREM 3.1. *Let $\{X_n\}_{n\in\mathbf{Z}}$ be (p,q)-bounded, $1 \le p \le 2 \le q \le +\infty$, and let $\sum |a_n|^2 |n|^{(q-2)/q} \log^2(1+|n|) < +\infty$. Then $S_N = \sum_{-N}^{N} a_n X_n$ converges a.s.*

PROOF. First, by Lemma 2.1, it is clear that we just need to prove the theorem for the nonstationary sequence $\{Z_n\}_{n\in\mathbf{Z}} = \{QY_n\}_{n\in\mathbf{Z}}$. To do so, we will adapt the proof of the classical Rademacher–Menchov theorem and first show that

$$E\delta_N^2 = E \max_{1\le j\le N} \left| \sum_{n=-j}^{j} a_n Z_n \right|^2 \le C N^{(q-2)/q} \log^2(1+N) \sum_{n=-N}^{N} |a_n|^2. \tag{3.1}$$

Proceeding as in [1] or in [16] and without loss of generality let $N = 2^r$. Then, from the dyadic decomposition of N as well as the Cauchy–Schwarz inequality, it follows that

$$\delta_N^2 \le (r+1) \sum_{k=0}^{r} \sum_{m=-2^k+1}^{2^k-1} \left| \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} a_n Z_n \right|^2.$$

Now by (p,q)-boundedness (where again and throughout the paper, the sup norm corresponds to the case $q = \infty$)

$$E\left| \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} a_n Z_n \right|^2 \le C\left(\int_{-\pi}^{\pi} \left| \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} a_n e^{in\theta} \right|^q d\theta \right)^{2/q}. \tag{3.2}$$

Now, by the Hausdorff–Young inequality and since $2 \le q \le +\infty$,

$$\begin{aligned} \left(\int_{-\pi}^{\pi} \left| \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} a_n e^{in\theta} \right|^q d\theta \right)^{2/q} &\le C\left(\sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} |a_n|^{q'} \right)^{2/q'} \qquad \left(\frac{1}{q} + \frac{1}{q'} = 1 \right) \\ &\le C(2^{r-k})^{(2/q')-1} \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} |a_n|^2. \end{aligned} \tag{3.3}$$

Combining (3.2) and (3.3) we get

$$\begin{aligned} E\delta_N^2 &\le C(r+1) \sum_{k=0}^{r} \sum_{m=-2^k+1}^{2^k-1} (2^{r-k})^{(q-2)/q} \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} |a_n|^2 \\ &\le C(r+1)(2^r)^{(q-2)/q} \sum_{k=0}^{r} \sum_{m=-2^k+1}^{2^k-1} \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} |a_n|^2 \end{aligned}$$

$$\leq C(r+1)(2^r)^{(q-2)/q}\sum_{k=0}^{r}\sum_{n=-2^r}^{2^r}|a_n|^2$$

$$= C(r+1)^2(2^r)^{(q-2)/q}\sum_{n=-2^r}^{2^r}|a_n|^2,$$

and this proves (3.1) since $N = 2^r$. With this estimate, we can now show, in a very classical manner, that it is enough to prove the result along dyadic sequences. Let $S_n = \sum_{-n}^{n} a_k Z_k$. Then by (3.1) we have

$$\begin{aligned}
&\sum_{n=1}^{\infty} E \max_{2^n\leq m<2^{n+1}} |S_{2^n} - S_m|^2\\
&\quad\leq C\sum_{n=1}^{\infty} n^2(2^n)^{(q-2)/q}\Bigg(\sum_{k=2^n+1}^{2^{n+1}}|a_k|^2 + \sum_{k=-2^{n+1}}^{-2^n-1}|a_k|^2\Bigg)\\
&\quad\leq C\sum_{n=1}^{\infty}\Bigg(\sum_{k=2^n+1}^{2^{n+1}} k^{(q-2)/q}\log^2(1+k)|a_k|^2\\
&\qquad\qquad + \sum_{k=-2^{n+1}}^{-2^n-1}|k|^{(q-2)/q}\log^2(1+|k|)|a_k|^2\Bigg)\\
&\quad\leq C\sum_{n=-\infty}^{\infty}|n|^{(q-2)/q}\log^2(1+|n|)|a_n|^2 < +\infty;
\end{aligned}$$

hence $\lim_n \max_{2^n\leq m<2^{n+1}} |S_{2^n} - S_m| = 0$ a.s. P.

To finish the proof we thus just need to show that $S_{2^n} = S_1 + \sum_{k=0}^{n-1}(S_{2^{k+1}} - S_{2^k})$ converges with probability 1. However, since

$$\Bigg(\sum_{n=1}^{\infty} E\Big|S_{2^{n+1}} - S_{2^n}\Big|\Bigg)^2 \leq C\sum_{n=1}^{\infty} n^2 E\Bigg(\Bigg|\sum_{k=2^n+1}^{2^{n+1}} a_k Z_k\Bigg|^2 + \Bigg|\sum_{k=-2^{n+1}}^{-2^n-1} a_k Z_k\Bigg|^2\Bigg)\sum_{n=1}^{\infty}\frac{1}{n^2},$$

it is in turn enough to show that

$$\sum_{n=1}^{\infty} n^2 E\Bigg(\Bigg|\sum_{k=2^n+1}^{2^{n+1}} a_k Z_k\Bigg|^2 + \Bigg|\sum_{k=-2^{n+1}}^{-2^n-1} a_k Z_k\Bigg|^2\Bigg) < +\infty. \tag{3.4}$$

However, again using (p,q)-boundedness and proceeding as in the proof of (3.1), we see that the series (3.4) is less than or equal to

$$\begin{aligned}
&C\sum_{n=1}^{\infty} n^2(2^n)^{(q-2)/q}\Bigg(\sum_{k=2^n+1}^{2^{n+1}}|a_k|^2 + \sum_{k=-2^{n+1}}^{-2^n-1}|a_k|^2\Bigg)\\
&\qquad\leq C\sum_{n=-\infty}^{\infty}|n|^{(q-2)/q}\log^2(1+|n|)|a_n|^2 < +\infty. \qquad \square
\end{aligned}$$

REMARK 3.2. (i) Of course, the condition $\sum |a_n|^2 |n|^{(q-2)/q} \log^2(1+|n|) < +\infty$ does ensure the L^p-convergence, $1 \le p \le 2$, of $\sum a_n X_n$ to $\int_{-\pi}^{\pi} a(\theta)\, dZ_X(\theta)$, where $a(\theta) = \sum_{-\infty}^{\infty} a_n e^{in\theta}$. This is so because

$$\begin{aligned} \left(E\left|\sum_{n=-\infty}^{\infty} a_n X_n\right|^p\right)^{1/p} &\le C\left(\int_{-\pi}^{\pi}\left|\sum_{n=-\infty}^{\infty} a_n e^{in\theta}\right|^q d\theta\right)^{1/q} \\ &\le C\left(\sum_{n=-\infty}^{\infty} |a_n|^{q'}\right)^{1/q'} \qquad \left(\frac{1}{q}+\frac{1}{q'}=1\right) \\ &\le C\sum_{n\in\mathbf{Z}} |a_n|^2 |n|^{(q-2)/q} \log^2(1+|n|). \end{aligned}$$

(ii) Whenever $0 \le p < 1$, $2 \le q \le +\infty$ and $X_n = \int_{-\pi}^{\pi} e^{in\theta}\, dZ_X(\theta)$, where Z_X: $L^q((-\pi,\pi]) \to L^p(\Omega, B, P)$ is a continuous linear operator, a version of Lemma 2.1 continues to hold (see [12]) and so does a version of Theorem 3.1, namely, for such $\{X_n\}$ the condition $\sum |a_n|^2 |n|^{(q-2)/q} \log^2(1+|n|) < +\infty$ implies the a.s. convergence of $\sum a_n X_n$.

We now state some particular cases and consequences of Theorem 3.1.

COROLLARY 3.3. (i) *Let $\{X_n\}_{n\in\mathbf{Z}}$ be a weakly stationary sequence and let $\sum |a_n|^2 |n| \log^2(1+|n|) < +\infty$. Then $\sum a_n X_n$ converges with probability* 1. *Furthermore, if $\{X_n\}_{n\in\mathbf{Z}}$ has its spectrum in $L^r((-\pi,\pi])$, $1 < r \le +\infty$, and if $\sum |a_n|^2 |n|^{1/r} \log^2(1+|n|) < +\infty$, then $\sum a_n X_n$ converges almost surely.*

(ii) *Let $\{X_n\}_{n\in\mathbf{Z}}$ be an S_p system, $1 \le p \le 2$, and let $\sum |a_n|^2 \log^2(1+|n|) < +\infty$. Then $\sum a_n X_n$ converges almost surely.*

PROOF. A weakly stationary sequence is $(2,\infty)$-bounded. It is $(2,q)$-bounded, $2 \le q < +\infty$, if and only if its spectrum is in $L^{q/(q-2)}((-\pi,\pi])$. □

REMARK 3.4. (i) It is shown in [8] that for a weakly stationary sequence $\{X_n\}$ with covariance $R(n) = \int_{-\pi}^{\pi} e^{in\theta}\, d\mu(\theta)$, $n \in \mathbf{Z}$, the optimal condition $\sum |a_n|^2 \sum_{k=-n}^{n} |R(k)| \log^2(1+|n|) < +\infty$ guarantees the a.s. convergence of the series $\sum a_n X_n$. Of course, when $\mu = \delta_{\theta_k}$ this condition and the one of the corollary are identical. However, in general they are independent of one another since they respectively involve the spectrum μ and its Fourier coefficients. Since

$$\sum |a_n|^2 \sum_{k=-n}^{n} |R(k)| \log^2(1+|n|) \le C \sum |a_n|^2 |n| \log^2(1+|n|),$$

the conditions of Corollary 3.3 are stronger than the conditions in [8]. However, as is well known (see [17] Chapter 1, Section 4), the Fourier coefficient functions in $L^1((-\pi,\pi])$ can decrease to zero arbitrarily slowly. Hence, if $d\mu = f\, d\theta$, where $f \in L^1((-\pi,\pi])$ is such that its Fourier coefficients $|R(n)|$

decrease to zero sufficiently slowly, we also have

$$\sum |a_n|^2 |n||R(n)| \log^2(1+|n|) \le C \sum |a_n|^2 \sum_{k=-n}^{n} |R(k)| \log^2(1+|n|).$$

Similarly, no necessary and sufficient conditions on the Fourier coefficients can characterize $L^r((-\pi,\pi])$, $r \neq 2$, so Corollary 3.3 and the results of [8] complement one another.

(ii) Corollary 3.3(ii) has to be contrasted with the case $p > 2$, where a.s. convergence only requires $\sum |a_n|^2 < +\infty$ (see [6, 16]).

(iii) The conditions of Theorem 3.1 or the ones of Corollary 3.3 are sharp. If $b_n = o(|n|^{1/2r} \log |n|)$, there exists a stationary sequence $\{X_n\}$ with L^r spectrum, $1 \le r \le +\infty$, such that $\sum a_n X_n$ diverges a.s. and such that $\sum |a_n|^2 |b_n|^2 < +\infty$. A way of constructing such counterexamples is to start with a Menchov–Tandori system of divergence ([1], Chapter 2, Section 4) and then to essentially proceed as in [7] by building a moving average sequence based on this system of divergence.

(iv) Let $\{X_n\} \subset L^2(\Omega, B, P)$ be a sequence such that $\sup_{n\in\mathbf{Z}} E|X_n|^2 < +\infty$. Then it is clear from the proof of Theorem 3.1 that the condition $\sum |a_n|^2 |n| \log^2(1+|n|) < +\infty$ implies the a.s. convergence of $\sum a_n X_n$. This last condition is independent of the results of [22] (see also [21]) requiring $\sum_{n,m=-\infty}^{+\infty} |a_n||a_m||EX_n\overline{X}_m| \log(1+|m|)\log(1+|m|) < +\infty$ for the a.s. convergence of $\sum a_n X_n$.

COROLLARY 3.5. *Let T be a contraction on $L^2(P)$, that is, T: $L^2(P) \longrightarrow L^2(P)$ is linear, bounded, with $\|T\| \le 1$, and let $\sum_{n=1}^{\infty} |a_n|^2 n \log^2(1+n) < +\infty$. For any $f \in L^2(P)$, $\sum_{n=1}^{\infty} a_n T^n f(x)$ converges for a.a. x. The result remains true if T: $L^2(P) \to L^2(P)$ is an invertible bounded linear operator such that* $\sup_{n\in\mathbf{Z}} \|T^n\| < +\infty$.

PROOF. The result is true for unitary operators (equivalently for stationary processes). Then a contraction can be dilated to a unitary operator, while T such that $\sup_{n\in\mathbf{Z}} \|T^n\| < +\infty$ is similar to a unitary operator. In both cases, we then proceed as in the proof of Corollary 4.9 in [13]. □

COROLLARY 3.6. *Let $\{X_n\}_{n\in\mathbf{Z}}$ be harmonizable* SαS, $1 < \alpha < 2$, *with control measure μ. Let $\sum |a_n|^2 |n| \log^2(1+|n|) < +\infty$. Then $\sum a_n X_n$ converges* a.s. *Furthermore, if $d\mu = f\,d\theta$ with $f \in L^r((-\pi,\pi])$, $1 < r \le 2/(2-\alpha)$, then $\sum |a_n|^2 |n|^{(\alpha r - 2r + 2)/\alpha r} \log^2(1+|n|) < \infty$ ensures the result.*

PROOF. A harmonizable SαS sequence is (p,∞)-bounded, $1 \le p < \alpha$. When $d\mu = f\,d\theta$ with f satisfying the given hypothesis, then $\{X_n\}$ is (p,q)-bounded, $1 \le p < \alpha$, $2 \le q = \alpha r/(r-1) \le +\infty$. □

Of course, from the above results and via Kronecker's lemma the a.s. convergence to zero of various averages follows. In particular, we get the following corollary.

COROLLARY 3.7. *Let $X = \{X_n\}_{n\in\mathbf{Z}}$ be (p,q)-bounded, $1 \le p \le 2 \le q < +\infty$. Then for any $\varepsilon > 0$,*

$$\lim_{N\to\infty} \frac{1}{N^{(1-1/q)}\log^{3/2+\varepsilon} N} \sum_{n=1}^{N} X_n = 0, \quad \text{a.s.}$$

One of our main purposes in writing this note was to illustrate how to transfer results valid for orthonormal series to the (p,q)-bounded ones. In fact, a great deal of the results of [1] and [16], Chapters 8–10, can be similarly generalized. Such is the case for more recent results on, for example, Cesáro a.s. convergence obtained in [18]. To further illustrate our point, let us sketch the proofs of the following versions of three classical results.

THEOREM 3.8. *Let $\{X_n\}_{n\in\mathbf{Z}}$ be $(p,2)$-bounded, $1 \le p \le 2$, and let $\sum |a_n|^2(\log\log(2+|n|))^2 < +\infty$. Then the arithmetic means $\sigma_N X = (S_0 + \dots + S_N)/(N+1)$ converge a.s.*

PROOF. By Lemma 2.1 it is enough to prove the result for the $(2,2)$-bounded sequence $\{Z_n = QY_n\}$, and we first show that

$$\sum_{n=1}^{\infty} E \max_{2^n \le m < 2^{n+1}} |\sigma_{2^n} Z - \sigma_m Z|^2 < +\infty. \tag{3.5}$$

For any integer $k \ge 2$, we have

$$\sigma_k Z - \sigma_{k-1} Z = \frac{1}{k(k+1)} \sum_{r=-k}^{k} r a_r Z_r.$$

Hence,

$$\begin{aligned} E \max_{2^n \le m < 2^{n+1}} |\sigma_{2^n} Z - \sigma_m Z|^2 &\le E\left(\sum_{k=2^n+1}^{2^{n+1}} |\sigma_k Z - \sigma_{k-1} Z| \right)^2 \\ &\le \sum_{k=2^n+1}^{2^{n+1}} kE|\sigma_k Z - \sigma_{k-1} Z|^2 \sum_{k=2^n+1}^{2^{n+1}} \frac{1}{k} \\ &\le C \sum_{k=2^n+1}^{2^{n+1}} \frac{k}{k^2(k+1)^2} E\left| \sum_{r=-k}^{k} r a_r Z_r \right|^2 \\ &\le C \sum_{k=2^n+1}^{2^{n+1}} \frac{1}{k^3} \sum_{r=-k}^{k} |r a_r|^2, \end{aligned} \tag{3.6}$$

where the last inequality follows from $(2,2)$-boundedness. From (3.6), we now get

$$\begin{aligned} \sum_{n=1}^{\infty} E \max_{2^n \le m < 2^{n+1}} |\sigma_{2^n} Z - \sigma_m Z|^2 &\le C \sum_{n=1}^{\infty} \sum_{k=2^n+1}^{2^{n+1}} \frac{1}{k^3} \sum_{r=-k}^{k} |r a_r|^2 \\ &\le C \sum_{n=1}^{\infty} \frac{1}{n^3} \sum_{r=-n}^{n} |r a_r|^2 \end{aligned}$$

$$\le C \sum_{r=-\infty}^{\infty} |ra_r|^2 \sum_{n=|r|}^{\infty} \frac{1}{n^3}$$

$$\le C \sum_{r=-\infty}^{\infty} |a_r|^2 < +\infty.$$

Hence (3.5) is proved and we just need to show that $\sigma_{2^n} Z$ converges a.s. To do so, we note that with probability 1, $\lim_{n\to+\infty}(S_{2^n}Z - \sigma_{2^n}Z) = 0$. Indeed, proceeding as above we have

$$\sum_{n=1}^{\infty} E|S_{2^n}Z - \sigma_{2^n}Z|^2 = \sum_{n=1}^{\infty} \frac{1}{(1+2^n)^2} E\left|\sum_{k=-2^n}^{2^n} |k| a_k Z_k\right|^2$$

$$\le C \sum_{n=1}^{\infty} \frac{1}{2^{2n}} \sum_{k=-2^n}^{2^n} |ka_k|^2$$

$$\le C \sum_{n=-\infty}^{\infty} |a_n|^2 < +\infty.$$

To finish the proof, it is thus enough to show that, under the stated hypotheses, $S_{2^n}Z$ converges a.s. First, it is easily seen that the random sequence $V_n = A_n^{-1} \sum_{|k|=2^n+1}^{2^{n+1}} a_k Z_k$, where $A_n^2 = \sum_{|k|=2^n+1}^{2^{n+1}} |a_k|^2$, is $(2,2)$-bounded since it verifies (2.1) with $p = q = 2$. Hence, by Theorem 3.1, $S_{2^N}Z = \sum_0^{N-1} A_n V_n$ converges a.s. if $\sum_n A_n^2 \log^2 n < +\infty$, that is, if $\sum_n |a_n|^2 (\log\log(2+|n|))^2 < +\infty$. □

REMARK 3.9. (i) Again [see Remark 3.4(iii)] the factor $\log\log$ is optimal.

(ii) Let $\{R_Y(n)\}$ be the covariance sequence of the dominating stationary sequence $\{Y_n\}$ as given in Lemma 2.1. Then it is clear from the proof given above, as well as from Remark 3.4(i), that $\sum |a_n|^2 \sum_{k=-n}^{n} |R_Y(k)| (\log\log(2+n))^2 < +\infty$ implies the a.s. convergence of $\sigma_N Y$. This in turn gives the a.s. convergence of $\sigma_N QY$, which ensures the a.s. convergence of $\sigma_N X$. A similar comment also applies to the results given below.

THEOREM 3.10. *Let $\{X_n\}_{n\in\mathbf{Z}}$ be (p,q)-bounded, $1 \le p \le 2 \le q \le +\infty$, and let*

$$\sum_{n=0}^{\infty} \left(\sum_{|k|=2^{2^n}+1}^{2^{2^{n+1}}} |a_k|^2 |k|^{(q-2)/q} \log^2 |k| \right)^{1/2} < +\infty.$$

Then $\sum a_n X_n$ converges unconditionally a.s.

PROOF. The proof of this result is a modification of the proof of the corresponding result for orthonormal variables. It can be obtained by replacing in the proof of the classical result (as given in [16], Chapter 8) the maximal inequality there (which is a subset of our case $q = 2$) by the maximal inequality (3.1). □

COROLLARY 3.11. *Let $\{X_n\}$ be (p,q)-bounded, $1 \le p \le 2 \le q \le +\infty$, and let $\{b_n\}$ be a nondecreasing sequence such that $\sum 1/((1+|n|)|b_n|) < +\infty$. Let $\sum |a_n|^2|n|^{(2-q)/q}\log^2(1+|n|)\log\log(2+|n|)|b_{\log\log n}| < +\infty$. Then $\sum a_n X_n$ converges unconditionally a.s.*

PROOF. The statement follows from $\sum 1/((1+|n|)|b_n|) < +\infty$ and the fact that $\{b_n\}$ is nondecreasing. □

For orthogonal sequences, the following result is due to Garsia (see [9]). For $q = 2$, the extensions presented below are rather immediate: The sequence $\{Y_n\}$ of Lemma 2.1 can be defined on a probability space $\Omega \oplus \Omega'$ and furthermore can be chosen orthogonal with constant second moment. Hence, since Λ in the lemma preserves a.s. convergence, the orthogonal result implies the desired conclusion. In particular, this result complements the one obtained in [19].

THEOREM 3.12. *Let $\{X_n\}$ be (p,q)-bounded, $1 \le p \le 2 \le q \le +\infty$, and let also $\sum |a_n|^2|n|^{(q-2)/q} < +\infty$. There exists a permutation π of the integers such that $\sum a_{\pi(n)} X_{\pi(n)}$ converges a.s.*

PROOF. First, use Lemma 2.1 to reduce the problem to studying the sequence $\{Z_n = QY_n\}$. Then a modification of the classical proof allows us to incorporate the factor $|n|^{(q-2)/q}$ in the maximal inequality there. The result follows from this, still proceeding as in [9]. □

To date, in our approach to the a.s. convergence problem, we have only been concerned with the case $p \le 2 \le q$. For other ranges of p and q and except [see Remark 3.2(ii)] for $0 \le p < 1$, no version of Lemma 2.1 holds. Nevertheless, it is possible to state partial results similar to Theorem 3.1 by examining the proof of the Rademacher–Menchov theorem.

THEOREM 3.13. *Let $\{X_n\}_{n\in\mathbf{Z}}$ be (p,q)-bounded, $1 \le p < +\infty$, $1 \le q \le +\infty$, and let $\{b_n\}$ be a nondecreasing sequence such that $\sum 1/(|nb_n|) < +\infty$.*

(i) *Let $p,q \ge 2$, $(p,q) \ne (2,2)$, and let $\sum |a_n|^p|n|^{p(1-1/p-1/q)}|b_{\log(1+|n|)}|^{p/p'} \times \log^{p/p'}(1+|n|) < +\infty$. Then $\sum a_n X_n$ converges a.s.*

(ii) *Let $p > 2$, $q \le 2$, and let $\sum |a_n|^p|n|^{(p/2)-1}|b_{\log(1+|n|)}|^{p/p'}\log^{p/p'}(1+|n|) < +\infty$. Then $\sum a_n X_n$ converges a.s.*

(iii) *Let $p,q \le 2$ and let $\sum |a_n|^p\log^p(1+|n|) < +\infty$. Then $\sum a_n X_n$ converges a.s.*

PROOF. Proceeding as in the proof of Theorem 3.1 (with its notation), we have, when $p \ge 1$,

$$\delta_N^p = (r+1)^{(p/p')}\sum_{k=0}^{r}\sum_{m=-2^k+1}^{2^k-1}\left|\sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} a_n Z_n\right|^p, \tag{3.7}$$

where $1/p + 1/p' = 1$. Then by (p,q)-boundedness,

$$E\delta_N^p \le C(r+1)^{p/p'} \sum_{k=0}^{r} \sum_{m=-2^k+1}^{2^k-1} \left(\int_{-\pi}^{\pi} \left| \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} a_n e^{in\theta} \right|^q d\theta \right)^{p/q}. \tag{3.8}$$

However, now, for $q \ge 2$ and by the Hausdorff–Young inequality, the r.h.s. of (3.8) is itself less than or equal to

$$C(r+1)^{p/p'} \sum_{k=0}^{r} \sum_{m=-2^k+1}^{2^k-1} \left(\sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} |a_n|^{q'} \right)^{p/q'}, \tag{3.9}$$

where again $1/q + 1/q' = 1$. Under the conditions of (i), $p/q' > 1$, hence we get, from (3.9),

$$\begin{aligned} E\delta_N^p &\le C(r+1)^{p/p'} \sum_{k=0}^{r} \sum_{m=-2^k+1}^{2^k-1} (2^{r-k})^{(p/q')-1} \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} |a_n|^p \\ &\le C(r+1)^{(p/p')} \sum_{k=0}^{r} (2^{r-k})^{(p/q')-1} \sum_{n=-2^r}^{2^r} |a_n|^p \\ &\le C \log^{p/p'} N N^{p(1-(1/p)-(1/q))} \sum_{n=-N}^{N} |a_n|^p. \end{aligned} \tag{3.10}$$

With the maximal inequality (3.10), we first reduce the problem to studying a.s. convergence along dyadic sequences by proving, as in Theorem 3.1, that

$$\begin{aligned} &\sum_{n=1}^{\infty} E \max_{2^n \le m < 2^{n+1}} |S_{2^n} - S_m|^p \\ &\quad \le C \sum_{n=-\infty}^{+\infty} |a_n|^p |n|^{p(1-(1/p)-(1/q))} \log^{p/p'}(1+|n|) \\ &\quad \le C \sum_{n=-\infty}^{+\infty} |a_n|^p |n|^{p(1-(1/p)-(1/q))} |b_{\log(1+|n|)}|^{p/p'} \log^{p/p'}(1+|n|) \\ &\quad < +\infty. \end{aligned}$$

To conclude the proof of (i), we then see that

$$\begin{aligned} &\left(\sum_{n=1}^{\infty} E|S_{2^{n+1}} - S_{2^n}| \right)^p \\ &\quad \le C \sum_{n=1}^{\infty} n^{p/p'} |b_n|^{p/p'} E \left| \sum_{|k|=2^n+1}^{2^{n+1}} a_k X_k \right|^p \left(\sum \frac{1}{|nb_n|} \right)^{p/p'} \\ &\quad \le C \sum_{n=-\infty}^{\infty} |a_n|^p |n|^{p(1-(1/p)-(1/q))} |b_{\log(1+|n|)}|^{p/p'} \log^{p/p'}(1+|n|) < +\infty. \end{aligned}$$

When $p > 2$, $q \le 2$, we get, from (3.8),

$$
\begin{aligned}
E\delta_N^p &\le C(r+1)^{p/p'} \sum_{k=0}^{r} \sum_{m=-2^k+1}^{2^k+1} \left(\sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} |a_n|^2 \right)^{p/2} \\
&\le C(r+1)^{p/p'} \sum_{k=0}^{r} \sum_{m=-2^k+1}^{2^k-1} (2^{r-k})^{(p/2)-1} \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} |a_n|^p \\
&\le C(r+1)^{p/p'} \sum_{k=0}^{r} (2^{r-k})^{(p/2)-1} \sum_{n=-2^{-r}}^{2^r} |a_n|^p \\
&\le C(r+1)^{p/p'} (2^r)^{(p/2)-1} \sum_{n=-N}^{N} |a_n|^p.
\end{aligned}
\tag{3.11}
$$

We thus have in the case (ii) the following maximal inequality:

$$
E\delta_N^p \le C \log^{p/p'} N N^{(p/2)-1} \sum_{n=-N}^{N} |a_n|^p. \tag{3.12}
$$

It now follows from (3.12) that

$$
\begin{aligned}
&\sum_{n=1}^{\infty} E \max_{2^n \le m < 2^{n+1}} |S_{2^n} - S_m|^p \\
&\qquad \le C \sum_{n=-\infty}^{\infty} |a_n|^p |n|^{(p/2)-1} |b_{\log(1+|n|)}|^{p/p'} \log^{p/p'}(1+|n|) < +\infty,
\end{aligned}
$$

and to finish the proof of (ii) we again notice that, as in the proof of (i),

$$
\left(\sum_{n=1}^{\infty} E|S_{2^{n+1}} - S_{2^n}| \right)^p < +\infty.
$$

For (iii) and since p and $q \le 2$, we get, from (3.8),

$$
\begin{aligned}
E\delta_N^p &\le C(r+1)^{p/p'} \sum_{k=0}^{r} \sum_{m=-2^k+1}^{2^k-1} \left(\sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} |a_n|^2 \right)^{p/2} \\
&\le C(r+1)^{p/p'} \sum_{k=0}^{r} \sum_{m=-2^k+1}^{2^k-1} \sum_{n=m2^{r-k}+1}^{(m+1)2^{r-k}} |a_n|^p \\
&\le C(r+1)^{p/p'} \sum_{k=0}^{r} \sum_{n=-2^r}^{2^r} |a_n|^p \\
&\le C \log^p N \sum_{n=-N}^{N} |a_n|^p.
\end{aligned}
\tag{3.13}
$$

Using (3.13) and proceeding as in the proof of (i) or (ii), we get (iii). □

As already mentioned, the conditions of Theorem 3.1 are sharp. One might nevertheless wonder if a.s. convergence is ever possible when these conditions are violated, for example, when $a_n = 1/n, n \neq 0, a_0 = 0$. Rather optimal results on that question are presented in [14]. It is shown there that a.s. convergence can happen and, in particular, the results of [14] do recover the necessary and sufficient condition for the a.s. convergence of $\sum_{n\neq 0} X_n/n$ which is obtained, for $\{X_n\}$ weakly stationary, in [15]. This will not be further discussed here and instead, to finish the paper, we apply the results presented above to study some strong consistency questions arising from [2] and [4].

Let us recall some elements of the framework we have in mind. Let

$$y_n = \beta_1 x_{n1} + \cdots + \beta_m x_{nm} + \varepsilon_n, \qquad n = 1, 2, \ldots, \tag{3.14}$$

where the $\beta_1, \ldots, \beta_m$ are unknown scalar parameters, where the x_{nj} are known scalars and where the ε_n are zero mean unobservable random disturbances. Then for any $n \geq m$, the least squares estimate $\mathbf{b}_n = (b_{n1}, \ldots, b_{nm})'$ of $\beta = (\beta_1, \ldots, \beta_m)'$ based on the design matrix $\mathbf{X}_n = (x_{ij})_{1\leq i\leq n,\ 1\leq j\leq m}$ and the response vector $\mathbf{y}_n = (y_1, \ldots, y_n)'$ is given by

$$\mathbf{b}_n = (\mathbf{X}_n'\mathbf{X}_n)^{-1}\mathbf{X}_n'\mathbf{y}_n, \tag{3.15}$$

provided the matrix $\mathbf{A}_n = (\mathbf{X}_n'\mathbf{X}_n)^{-1}$ is nonsingular. In this framework, the following result complements the ones in [2] and [4].

THEOREM 3.14. *Let $\{\varepsilon_n\}$ be $(p,2)$-bounded, $1 \leq p \leq 2$. Let $\mathbf{A}_n$ be nonsingular for all $n \geq m$ and such that $\lim_{n\to+\infty} a_{jj}^{(n)} = 0$, $j = 1, \ldots, m$. Then,*

$$b_j^{(n)} - \beta_j = o\left(\sqrt{f\big(a_{jj}^{(n)}\big)}\log n\right) \quad a.s., \tag{3.16}$$

for any positive function f on $(0, +\infty)$ such that $\int_0^C f^{-1}(t)\,dt < +\infty$, for some $C > 0$, and such that $t^{-2}f(t) \uparrow +\infty$ as $t \downarrow 0$.

PROOF. Once more, via Lemma 2.1, $\varepsilon_n = \Lambda Q\nu_n$ and we just need to prove the result for the sequence $\{Q\nu_n\}$. It is also clear from the proof of Lemma 2.1 (see [11]) that the sequence $\{\nu_n\}$ can be defined on a probability space $\Omega \oplus \Omega'$ and thus it is enough to prove the result for the disturbance sequence $\{\nu_n\}$. Furthermore, $\{\nu_n\}$ can be chosen to have zero mean if $\{\varepsilon_n\}$ has zero mean and is orthogonal with constant variance (because $q = 2$). So the problem has been reduced to studying strong consistency for a model where the disturbances $\{\nu_n\}$ form a zero mean orthogonal sequence with constant variance. For such a model and, say, for $j = 1$, we have (see [2])

$$b_1^{(n)} - \beta_1 = \frac{\sum_{k=1}^{n-m+1} c_k \nu_k}{\sum_{k=1}^{n-m+1} c_k^2}, \tag{3.17}$$

where $\sum_{k=1}^{n-m+1} c_k^2 = 1/a_{11}^{(n)}$. However, for such a sequence, it is known (see [1], Chapter 2, Section 3) that whenever f satisfies the requirements of the theorem,

$$\frac{1}{\sum_{k=1}^{n-m+1} c_k^2} \sum_{k=1}^{n-m+1} c_k \nu_k = o\left(\sqrt{f\left(\frac{1}{\sum_{k=1}^{n-m+1} c_k^2}\right)} \log(n-m+1)\right) \quad \text{a.s.} \tag{3.18}$$

Hence from (3.15), (3.16) and the form of $a_{11}^{(n)}$, the desired result follows. □

To finish, let us also mention that when ε_n is stationary with spectrum in L^r, $1 \le r \le +\infty$, or is (p,q)-bounded, $1 \le p \le 2 \le q \le +\infty$, and if $r = q/(q-2)$, then $b_j^{(n)} - \beta_j = o\left(\sqrt{f(a_{jj}^{(n)})} n^{1/2r} \log n\right)$. In particular, it is enough that $a_{jj}^{(n)} = o(n^{1/r}(\log n)^{-2-\delta})$, $\delta > 0$, to have strong consistency.

REFERENCES

[1] ALEXITS, G. (1961). *Convergence Problems of Orthogonal Series.* Pergamon, New York.

[2] ANDERSON, T. W. and TAYLOR, J. B. (1976). Strong consistency of least squares estimates in normal linear regression. *Ann. Statist.* **4** 788–790.

[3] BRYC, W. and SMOLENSKI, W. (1993). Moment conditions for almost sure convergence of weakly correlated random variables. *Proc. Amer. Math. Soc.* **119** 629–635.

[4] CHEN, G. C., LAI T. L. and WEI, C. Z. (1981). Convergence systems and strong consistency of least squares estimates in regression models. *J. Multivariate Anal.* **11** 319–333.

[5] DOOB, J. L. (1953). *Stochastic Processes.* Wiley, New York.

[6] GAPOSHKIN, V. F. (1966). Lacunary series and independent functions. *Russian Math. Surveys* **21** 1–82.

[7] GAPOSHKIN, V. F. (1973). On the strong law of large numbers for second order stationary processes and sequences. *Theory Probab. Appl.* **18** 372–375.

[8] GAPOSHKIN, V. F. (1975). Convergence of series connected with stationary sequences. *Math. USSR-Izv.* **9** 1297–1321.

[9] GARSIA, A. M. (1970). *Topics in Almost Everywhere Convergence.* Markham, Chicago, IL.

[10] HOUDRÉ, C. (1990). Harmonizability, V-boundedness and $(2,p)$-boundedness of stochastic processes. *Probab. Theory Related Fields* **84** 39–45.

[11] HOUDRÉ, C. (1990). Linear Fourier and stochastic analysis. *Probab. Theory Related Fields* **87** 167–188.

[12] HOUDRÉ, C. (1990). Factorization algorithms and non-stationary Wiener filtering. *Stochastics Stochastics Rep.* **33** 49–62.

[13] HOUDRÉ, C. (1992). On the Spectral SLLN and pointwise ergodic theorem in L^α. *Ann. Probab.* **20** 1731–1753.

[14] HOUDRÉ, C. and LACEY, M. T. (1995). Spectral criteria, SLLNS and a.s. convergence of series of stationary variables. Preprint.

[15] JAJTE, R. (1987). On the existence of the ergodic Hilbert transform. *Ann. Probab.* **15** 831–835.

[16] KASHIN, B. S. and SAAKYAN, A. A. (1989). *Orthogonal Series. Translations of Mathematical Monographs* **75**. Amer. Math. Soc., Providence, R.I.

[17] KATZNELSON, Y. (1968). *An Introduction to Harmonic Analysis*. Wiley, New York.

[18] MÓRICZ, F. (1983). On the Cesáro means of orthogonal sequences of random variables. *Ann. Probab.* **11** 827–832.

[19] ØRNO, P. (1976). A note on unconditionally converging series in L_p. *Proc. Amer. Math. Soc.* **59** 252–254.

[20] PELIGRAD, M. (1982). Invariance principles for mixing sequences of variables. *Ann. Probab.* **10** 968–981.

[21] SERFLING, R. J. (1970). Convergence properties of S_n under moment restrictions. *Ann. Math. Statist.* **41** 1235–1248.

[22] SZÉP, A. (1972). The non-orthogonal Menchoff–Rademacher theorem. *Acta Sci. Math. (Szeged)* **33** 231–235.

GEORGIA INSTITUTE OF TECHNOLOGY
SCHOOL OF MATHEMATICS
ATLANTA, GEORGIA 30332-0160
houdre@math.gatech.edu