Available online at www.sciencedirect.com
SCIENCE DIRECT
stochastic processes and their applications
ELSEVIER Stochastic Processes and their Applications 111 (2004) 157–173
www.elsevier.com/locate/spa

On the convergence of moments in the almost sure central limit theorem for martingales with statistical applications

B. Bercu

Laboratoire de Statistique et Probabilités, UMR C5583, Université Paul Sabatier, 118 Route de Narbonne, 31062 Toulouse Cedex, France

Received 17 July 2001; received in revised form 10 June 2002; accepted 27 October 2002

Abstract

We establish new almost sure asymptotic properties for martingale transforms. It enables us to deduce the convergence of moments in the almost sure central limit theorem for martingales. Several statistical applications on the asymptotic behavior of stochastic regression models are also provided.

Keywords: Martingale transforms; Moments; Stochastic regression

1. Introduction

Let (ξ_n) be a sequence of independent and identically distributed random variables with $\mathbb{E}[\xi_n] = 0$, $\mathbb{E}[\xi_n^2] = \sigma^2$ and define the empirical measures

$$G_n = \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} \delta_{S_k/\sqrt{k}} \quad \text{with} \quad S_n = \sum_{k=1}^{n} \xi_k.$$

The celebrated almost sure central limit theorem (ASCLT) states that, with probability one, $G_n \Rightarrow G$ where G stands for the standard $\mathcal{N}(0, \sigma^2)$ distribution. It was simultaneously established by Brosamler (1988) and Schatte (1988, 1991) and in the present form by Lacey and Phillip (1990). In other words, for any bounded continuous

E-mail address: bercu@math.ups-tlse.fr (B. Bercu).

0304-4149/$ - see front matter
doi:10.1016/j.spa.2002.10.001

real function h

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} \frac{1}{k} h\left(\frac{S_k}{\sqrt{k}}\right) = \int_{\mathbb{R}} h(x)\,\mathrm{d}G(x) \quad \text{a.s.} \tag{1.1}$$

We also refer the reader to Berkes and Csáki (2001) for a remarkable universal ASCLT covering a large class of limit theorems for partial sums, extremes, empirical distribution functions and local times associated with (ξ_n).

One might wonder if convergence (1.1) holds for unbounded functions h. Schatte (1991) has shown that (1.1) is true with $h(x)=\exp(ax^2)$ and $a<1/4$. More recently, Ibragimov and Lifshits (1998, 1999) proved that (1.1) holds as soon as the integral on the right of (1.1) is finite together with a mild regularity assumption on h. On the other hand, we also know from the important contribution of Chaabane (1996, 2001) and Chaabane and Maaouia (2000) and Lifshits (2001, 2002) that the ASCLT holds in the martingale framework. More precisely, let (ε_n) be a martingale difference sequence adapted to an appropriate filtration $\mathbb{F}=(\mathscr{F}_n)$ and let (ϕ_n) be a sequence of random variables adapted to $\mathbb{F}$. We define the real martingale transform (M_n) by

$$M_n = \sum_{k=1}^{n} \phi_{k-1}\varepsilon_k.$$

We also define the explosion coefficient associated with (ϕ_n) by

$$f_n = \frac{\phi_n^2}{s_n} \quad \text{with} \quad s_n = \sum_{k=0}^{n} \phi_k^2.$$

In all the sequel, we assume that (s_n) increases a.s. to infinity. A simplified version of the ASCLT for martingales (Chaabane, 1996) is as follows.

Theorem 1. *Assume that (ε_n) is a martingale difference sequence such that $\mathbb{E}[\varepsilon_{n+1}^2|\mathscr{F}_n]=\sigma^2$ a.s. and satisfying for some $a>2$*

$$\sup_{n\geqslant 0} \mathbb{E}[|\varepsilon_{n+1}|^a|\mathscr{F}_n] < \infty \quad \textit{a.s.} \tag{1.2}$$

In addition, assume that for some $b>1$

$$\sum_{n=1}^{\infty} f_n^b < \infty \quad \textit{a.s.} \tag{1.3}$$

Then, (M_n) satisfies an ASCLT so that for any bounded continuous real function h

$$\lim_{n\to\infty} \frac{1}{\log s_n} \sum_{k=1}^{n} f_k h\left(\frac{M_k}{\sqrt{s_{k-1}}}\right) = \int_{\mathbb{R}} h(x)\,\mathrm{d}G(x) \quad \textit{a.s.} \tag{1.4}$$

Similarly to (1.1), a natural question is whether or not convergence (1.4) holds for unbounded functions h. It was already established by formula (2.4) of Wei (1987) that under the moment condition (1.2), if the explosion coefficient f_n tends to zero a.s., then

$$\lim_{n\to\infty} \frac{1}{\log s_n} \sum_{k=1}^{n} f_k \left(\frac{M_k^2}{s_{k-1}}\right) = \sigma^2 \quad \text{a.s.} \tag{1.5}$$

Consequently, (1.4) is true with $h(x)=x^2$. The purpose of this paper is to show that under a suitable additional moment assumption on (ε_n), convergence (1.4) holds for any functions h such that $|h(x)| \leqslant x^{2p}$ with $p \geqslant 1$. Several statistical applications on the asymptotic behavior of stochastic regression models are also provided. Finally, a recent application on the adaptive control of parametric nonlinear autoregressive models can be found in Bercu and Portier (2002).

The paper is organized as follows. In Section 2, we establish new almost sure asymptotic properties for powers of martingale transforms. Moreover, if the explosion coefficient f_n tends to zero a.s., we prove the convergence of moments in the ASCLT for martingales. We also propose similar results when the explosion coefficient f_n converges a.s. to a positive random variable. Statistical applications are developed in Section 4 and all technical proofs are collected in Sections 3 and 5.

2. Main results

We first propose new almost sure asymptotic properties for powers of martingales transforms generalizing a well-known strong law due to Neveu (1975) and Lai and Wei (1982). One can observe that (M_n) is a martingale transform whereas it is not necessarily a martingale except when (ε_n) and (ϕ_n) are both square integrable. In fact, all locally square integrable real martingales can be seen as particular martingale transforms. For any integer $p \geqslant 1$, set

$$v_n(p) = \frac{s_n^p - s_{n-1}^p}{s_n^p}.$$

Theorem 2. *Assume that (ε_n) is a martingale difference sequence such that for some integer $p \geqslant 1$*

$$\sup_{n \geqslant 0} \mathbb{E}[\varepsilon_{n+1}^{2p} | \mathscr{F}_n] < \infty \quad a.s. \tag{2.1}$$

Then, for any $\gamma > 0$

$$M_n^{2p} = \mathrm{o}\,(s_{n-1}^p (\log s_{n-1})^{1+\gamma}) \quad a.s. \tag{2.2}$$

$$\sum_{n=1}^{\infty} \frac{v_n(p)}{(\log s_n)^{1+\gamma}} \left(\frac{M_n^2}{s_{n-1}} \right)^p < \infty \quad a.s. \tag{2.3}$$

Moreover, assume that for some $a > 2p$

$$\sup_{n \geqslant 0} \mathbb{E}[|\varepsilon_{n+1}|^a | \mathscr{F}_n] < \infty \quad a.s. \tag{2.4}$$

Then

$$M_n^{2p} = \mathrm{O}\,(s_{n-1}^p \log s_{n-1}) \quad a.s. \tag{2.5}$$

$$\sum_{k=1}^{n} v_k(p) \left(\frac{M_k^2}{s_{k-1}} \right)^p = \mathrm{O}\,(\log s_n) \quad a.s. \tag{2.6}$$

Remark. On the one hand, in the particular case $p = 1$, the almost sure properties given above are exactly the scalar version of Theorem 1.3.24 of Duflo (1997), which is at the core of many proofs concerning the study of the asymptotic behavior of linear regression models. On the other hand, formula (2.30) of Wei (1987) implies that under (2.1)

$$M_n^{2p} = \mathrm{o}\,(s_{n-1}^p(\log s_{n-1})^\delta) \quad \text{a.s.} \tag{2.7}$$

with $\delta > 1$ which is exactly formula (2.2). However, we assume here that p is a positive integer while $p \geqslant 1$ can be a real number in Wei's result. Nevertheless, the proof of (2.7) is totally different from that of (2.2) and (2.5) as it mainly relies on the Burkholder–Davis–Gundy inequality. Finally, the most important results of Theorem 2 lie in (2.3) and (2.6) as we shall see below.

Theorem 3. *Assume that* (ε_n) *is a martingale difference sequence such that* $\mathbb{E}[\varepsilon_{n+1}^2|\mathscr{F}_n] = \sigma^2$ *a.s. and satisfying, for some integer* $p \geqslant 1$, *the moment condition* (2.4). *In addition, assume that the explosion coefficient* f_n *tends to zero a.s. Then*

$$\lim_{n\to\infty} \frac{1}{\log s_n} \sum_{k=1}^n f_k \left(\frac{M_k^2}{s_{k-1}}\right)^p = \frac{\sigma^{2p}(2p)!}{2^p p!} \quad \text{a.s.} \tag{2.8}$$

A straightforward application of Theorem 3 is as follows.

Corollary 4. *Assume that* (ε_n) *is a martingale difference sequence such that* $\mathbb{E}[\varepsilon_{n+1}^2|\mathscr{F}_n] = \sigma^2$ *a.s. and satisfying, for some integer* $p \geqslant 1$, *the moment condition* (2.4). *In addition, assume that* (1.3) *holds. Then, for any continuous real function* h *such that* $|h(x)| \leqslant x^{2p}$, *we have*

$$\lim_{n\to\infty} \frac{1}{\log s_n} \sum_{k=1}^n f_k h\left(\frac{M_k}{\sqrt{s_{k-1}}}\right) = \int_{\mathbb{R}} h(x)\,\mathrm{d}G(x) \quad \text{a.s.} \tag{2.9}$$

Theorem 3 establishes the convergence of moments in the ASCLT for martingales. It requires that the explosion coefficient f_n tends to zero a.s. One might wonder whether or not a similar convergence holds when f_n converges a.s. to a positive random variable f and it is the purpose of the following result to show that this is the case. First of all, for any integer $p \geqslant 1$, set

$$\sigma_n(p) = \mathbb{E}[\varepsilon_{n+1}^p|\mathscr{F}_n].$$

Theorem 5. *Assume that* (ε_n) *is a martingale difference sequence satisfying, for some integer* $p \geqslant 1$, *the moment condition* (2.4). *In addition, assume that for any* $2 \leqslant q \leqslant 2p$

$$\lim_{n\to\infty} \sigma_n(q) = \sigma(q) \quad \text{a.s.} \tag{2.10}$$

where $\sigma(q)=0$ if q is odd. If the explosion coefficient f_n converges a.s. to a random variable f with $0<f<1$, then

$$\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\left(\frac{M_k^2}{s_{k-1}}\right)^p = l(p,f) \quad a.s. \tag{2.11}$$

where $l(0,f)=1$ and, for $p\geqslant 1$, $l(p,f)$ is given by the recurrence equation

$$l(p,f)=\frac{1}{1-(1-f)^p}\sum_{k=1}^{p}C_{2p}^{2k}f^k(1-f)^{p-k}\sigma(2k)l(p-k,f).$$

Corollary 6. *For any integer $p\geqslant 1$, $l(p,f)$ does not depend upon the random variable f if and only if, for all $1\leqslant k\leqslant p$, the moments $\sigma(2k)$ coincide with those of an $\mathcal{N}(0,\sigma^2)$ random variable where $\sigma(2)=\sigma^2$. In this particular case, the limits in (2.8) and (2.11) are identical*

$$l(p,f)=l(p)=\frac{\sigma^{2p}(2p)!}{2^p p!}.$$

3. Proofs

3.1. Proof of Theorem 2

We shall prove Theorem 2 by induction on the power $p\geqslant 1$. First of all, Theorem 2 is already established for $p=1$ in Duflo (1997), Lai and Wei (1982) and Neveu (1975). Next, let $p\geqslant 2$ and assume that Theorem 2 holds for any power q with $1\leqslant q\leqslant p-1$. For $n\geqslant 0$, as $M_{n+1}=M_n+\phi_n\varepsilon_{n+1}$,

$$M_{n+1}^{2p}=\sum_{k=0}^{2p}C_{2p}^k\phi_n^k\varepsilon_{n+1}^k M_n^{2p-k}. \tag{3.1}$$

Moreover, for any $0\leqslant k\leqslant 2p$, set

$$\varphi_n(k)=s_n^{-p}\phi_n^k M_n^{2p-k}.$$

Via a standard truncation argument (Wei, 1987), we can assume without loss of generality that each $\varphi_n(k)$ is a bounded random variable. In addition, we can also assume $\sup_n\sigma_n(2p)\leqslant C$ a.s. for some constant $C\geqslant 1$. Denote

$$V_n=\frac{M_n^{2p}}{s_{n-1}^p} \quad\text{and}\quad W_n=\frac{\sigma_n(2p)\phi_n^{2p}}{s_n^p}.$$

It immediately follows from (3.1) that for any $n\geqslant 1$

$$\mathbb{E}[V_{n+1}|\mathscr{F}_n]\leqslant V_n-A_n+B_n+W_n \tag{3.2}$$

with

$$A_n=v_n(p)s_{n-1}^{-p}M_n^{2p} \quad\text{and}\quad B_n=\sum_{k=2}^{2p-1}C_{2p}^k|\sigma_n(k)\varphi_n(k)|.$$

Let (a_n) be the positive decreasing sequence given by $a_n^{-1} = (\log s_n)^{1+\gamma}$ with $\gamma > 0$. On the one hand, since $\phi_n^2 \leqslant s_n$, we have

$$\sum_{n=1}^{\infty} a_n W_n < \infty \quad \text{a.s.} \tag{3.3}$$

On the other hand, in order to use the Robbins–Siegmund theorem (see e.g. Duflo, 1997, p. 18), we claim that

$$\sum_{n=1}^{\infty} a_n B_n < \infty \quad \text{a.s.} \tag{3.4}$$

In fact, one can easily see from the Hölder inequality that each $|\sigma_n(k)| \leqslant C$ a.s. Moreover, for any integer $q \geqslant 1$, we have $s_n^q - s_{n-1}^q \geqslant \phi_n^{2r} s_n^{q-r}$ so that $v_n(q) \geqslant \phi_n^{2r} s_n^{-r}$ with $1 \leqslant r \leqslant q$. Consequently, we obtain from the induction assumption that for any $1 \leqslant q \leqslant p-1$ and $1 \leqslant r \leqslant q$

$$\sum_{n=1}^{\infty} \frac{a_n \phi_n^{2r} M_n^{2q}}{s_n^{q+r}} < \infty \quad \text{a.s.} \tag{3.5}$$

We shall apply (3.5) in the three following cases for proving (3.4).

Case 1: *Let* $2 \leqslant k \leqslant 2(p-1)$ *with* k *even.* We can find $1 \leqslant q \leqslant p-1$ such that $k = 2(p-q)$. Then, as $\phi_n^k \leqslant \phi_n^2 s_n^{p-q-1}$, we obtain from (3.5) with $r = 1$ that a.s.

$$\sum_{n=1}^{\infty} \frac{a_n |\phi_n^k M_n^{2p-k}|}{s_n^p} \leqslant \sum_{n=1}^{\infty} \frac{a_n \phi_n^2 M_n^{2q}}{s_n^{q+1}} < \infty \tag{3.6}$$

Case 2: *Let* $3 \leqslant k \leqslant 2p-3$ *with* k *odd.* First, assume that $k \leqslant p-1$. We can choose $2 \leqslant q \leqslant p-1$ such that $k = 2(p-q)+1$. Then, it follows from the Cauchy–Schwarz inequality and (3.5) with $2r = k+1$ that a.s.

$$\left(\sum_{n=1}^{\infty} \frac{a_n |\phi_n^k M_n^{2p-k}|}{s_n^p} \right)^2 \leqslant \sum_{n=1}^{\infty} \frac{a_n \phi_n^{2r} M_n^{2q}}{s_n^{q+r}} \sum_{n=1}^{\infty} \frac{a_n \phi_n^{2(r-1)} M_n^{2(q-1)}}{s_n^{q+r-2}} < \infty \tag{3.7}$$

Next, assume that $p \leqslant k \leqslant 2p-3$. We obviously have

$$\sum_{n=1}^{\infty} \frac{a_n |\phi_n^k M_n^{2p-k}|}{s_n^p} \leqslant \sum_{n=1}^{\infty} \frac{a_n |\phi_n^{2p-k} M_n^{2p-k}|}{s_n^{2p-k}} \quad \text{a.s.}$$

Hence, as before, we find that (3.7) also holds with $r = q$.

Case 3: *Let* $k = 2p-1$. We have a.s.

$$\left(\sum_{n=1}^{\infty} \frac{a_n |\phi_n^{2p-1} M_n|}{s_n^p} \right)^2 \leqslant \left(\sum_{n=1}^{\infty} \frac{a_n |\phi_n^3 M_n|}{s_n^2} \right)^2 \leqslant \sum_{n=1}^{\infty} \frac{a_n \phi_n^2 M_n^2}{s_n^2} \sum_{n=1}^{\infty} \frac{a_n \phi_n^4}{s_n^2} < \infty \tag{3.8}$$

Therefore, (3.4) follows from the conjunction of (3.6), (3.7) and (3.8). Finally, as $a_n = \mathrm{o}(1)$ a.s., we obtain from (3.2), (3.3), (3.4) together with the Robbins–Siegmund theorem that $a_{n-1}V_n = \mathrm{o}(1)$ a.s. and

$$\sum_{n=1}^{\infty} a_n A_n < \infty \quad \text{a.s.} \tag{3.9}$$

which completes the proof of (2.2) and (2.3). It now remains to prove (2.5) and (2.6). One can deduce from (3.1) that for any $n \geqslant 1$

$$V_{n+1} + \mathscr{A}_n = V_1 + \mathscr{B}_{n+1} + \mathscr{W}_{n+1} \tag{3.10}$$

where

$$\mathscr{A}_n = \sum_{k=1}^{n} v_k(p) s_{k-1}^{-p} M_k^{2p}, \qquad \mathscr{W}_{n+1} = \sum_{k=1}^{n} s_k^{-p} \phi_k^{2p} \varepsilon_{k+1}^{2p},$$

$$\mathscr{B}_{n+1} = \sum_{l=1}^{2p-1} C_{2p}^{l} B_{n+1}(l) \quad \text{with} \quad B_{n+1}(l) = \sum_{k=1}^{n} \varphi_k(l) \varepsilon_{k+1}^{l}.$$

On the one hand, (2.4) together with Chow's lemma (see e.g. Duflo, 1997, p. 22) imply that a.s.

$$\mathscr{W}_{n+1} = \mathrm{O}(F_n(p)) \quad \text{with} \quad F_n(p) = \sum_{k=1}^{n} f_k^{p}.$$

In addition, by the elementary inequality $x \leqslant -\log(1-x)$ with $0 < x < 1$, we deduce that $f_n^p \leqslant f_n \leqslant -\log(1-f_n)$ so that $f_n^p \leqslant \log s_n - \log s_{n-1}$. Consequently, $F_n(p) \leqslant \log s_n$ which ensures that

$$\mathscr{W}_{n+1} = \mathrm{O}(\log s_n) \quad \text{a.s.} \tag{3.11}$$

On the other hand, we claim that

$$|\mathscr{B}_{n+1}| = \mathrm{O}(\log s_n) \quad \text{a.s.} \tag{3.12}$$

In order to prove relation (3.12), we have to show that for any $1 \leqslant l \leqslant 2p-1$, $|B_{n+1}(l)| = \mathrm{O}(\log s_n)$ a.s. First of all, split $B_{n+1}(l)$ into two terms, $B_{n+1}(l) = C_{n+1}(l) + D_n(l)$ with

$$C_{n+1}(l) = \sum_{k=1}^{n} \varphi_k(l) e_{k+1}(l) \quad \text{and} \quad D_n(l) = \sum_{k=1}^{n} \varphi_k(l) \sigma_k(l)$$

where $e_{n+1}(l) = \varepsilon_{n+1}^{l} - \sigma_n(l)$. For any $1 \leqslant l \leqslant p$, one can easily deduce from (2.2) and (2.3) together with Kronecker's lemma that, for all $\gamma > 0$,

$$\tau_n(l) = \sum_{k=1}^{n} |\varphi_k(l)|^2 = \mathrm{o}((\log s_n)^{d+\gamma}) \quad \text{a.s.}$$

with $d = (2p-1)/p$. Consequently, by virtue of the standard strong law of large numbers for martingales, we obtain that $|C_{n+1}(l)|^2 = \mathrm{O}(\tau_n(l) \log \tau_n(l))$ so that $|C_{n+1}(l)|^2 = \mathrm{o}((\log s_n)^{d+\gamma} \log_2 s_n)$ a.s. Since $d < 2$, it implies that, for any $1 \leqslant l \leqslant p$, $C_{n+1}(l) =$

$o(\log s_n)$ a.s. Moreover, for any $p+1 \leqslant l \leqslant 2p-1$, we find via Chow's lemma that either $(C_{n+1}(l))$ converges a.s. or $C_{n+1}(l) = o(v_n(l))$ a.s. where

$$v_n(l) = \sum_{k=1}^{n} |\varphi_k(l)|^{\delta} = \sum_{k=1}^{n} f_k^p \left(\frac{M_k^2}{s_k}\right)^{p(\delta-1)}$$

with $\delta = 2p/l$. One can observe that we always have $1 < \delta < 2$. Therefore, it follows from the Hölder inequality together with the induction assumption that $v_n(l) = O(\log s_n)$ a.s. which leads to $C_{n+1}(l) = o(\log s_n)$ a.s. Consequently, we infer that for any $1 \leqslant l \leqslant 2p-1$

$$C_{n+1}(l) = o(\log s_n) \quad \text{a.s.} \tag{3.13}$$

In order to prove (3.12), it remains to show that for any $1 \leqslant l \leqslant 2p-1$

$$|D_n(l)| = O(\log s_n) \quad \text{a.s.} \tag{3.14}$$

We shall proceed as in the proof of (3.4). Similarly to (3.5), we deduce from the induction assumption that for any $1 \leqslant q \leqslant p-1$ and $1 \leqslant r \leqslant q$

$$\sum_{k=1}^{n} \frac{\phi_k^{2r} M_k^{2q}}{s_k^{q+r}} = O(\log s_n) \quad \text{a.s.} \tag{3.15}$$

We shall apply (3.15) in the three following cases for proving (3.14).

Case 1: *Let* $2 \leqslant l \leqslant 2(p-1)$ *with* l *even.* We can find $1 \leqslant q \leqslant p-1$ such that $l = 2(p-q)$. Then, as $\phi_k^l \leqslant \phi_k^2 s_k^{p-q-1}$, we obtain from (3.15) with $r = 1$ that a.s.

$$|D_n(l)| = O\left(\sum_{k=1}^{n} \frac{\phi_k^2 M_k^{2q}}{s_k^{q+1}}\right) = O(\log s_n). \tag{3.16}$$

Case 2: *Let* $3 \leqslant l \leqslant 2p-3$ *with* l *odd.* One can note that $D_n(1)=0$ since $\sigma_n(1)=0$ a.s. Hence, we can take $l \geqslant 3$. First, assume that $l \leqslant p-1$. We can choose $2 \leqslant q \leqslant p-1$ such that $l = 2(p-q)+1$. Then, it follows from the Cauchy–Schwarz inequality and (3.15) with $2r = l+1$ that a.s.

$$|D_n(l)| = O\left(\sum_{k=1}^{n} \frac{\phi_k^{2r} M_k^{2q}}{s_k^{q+r}} \sum_{k=1}^{n} \frac{\phi_k^{2(r-1)} M_k^{2(q-1)}}{s_k^{q+r-2}}\right)^{1/2} = O(\log s_n). \tag{3.17}$$

Next, assume that $p \leqslant l \leqslant 2p-3$. We obviously have

$$|D_n(l)| = O\left(\sum_{k=1}^{n} \frac{|\phi_k^{2p-l} M_k^{2p-l}|}{s_k^{2p-l}}\right) \quad \text{a.s.}$$

Hence, we obtain that (3.17) also holds with $r = q$.

Case 3: *Let* $l = 2p-1$. We have a.s.

$$|D_n(l)| = O\left(\sum_{k=1}^{n} \frac{|\phi_k^3 M_k|}{s_k^2}\right) = O\left(\sum_{k=1}^{n} \frac{\phi_k^2 M_k^2}{s_k^2} \sum_{k=1}^{n} \frac{\phi_k^4}{s_k^2}\right)^{1/2} = O(\log s_n). \tag{3.18}$$

Consequently, (3.14) follows from the conjunction of (3.16), (3.17) and (3.18) and it immediately implies (3.12). Finally, we find from (3.10), (3.11) and (3.12) that

$V_{n+1} = \mathrm{O}(\log s_n)$ a.s. and $\mathscr{A}_n = \mathrm{O}(\log s_n)$ a.s. which completes the proof of Theorem 2. □

3.2. Proof of Theorem 3

We shall use the same notations as those in the proof of Theorem 2. First of all, it immediately follows from (1.5) that Theorem 3 is true for $p=1$. Next, let $p \geqslant 2$ and assume that Theorem 3 holds for any power q with $1 \leqslant q \leqslant p-1$. On the one hand, via (2.7), we clearly have $V_{n+1} = \mathrm{o}(\log s_n)$ a.s. On the other hand, as the explosion coefficient f_n tends to zero a.s., we obtain that $F_n(p) = \mathrm{o}(\log s_n)$ a.s. which ensures that $\mathscr{W}_{n+1} = \mathrm{o}(\log s_n)$ a.s. Furthermore, we already saw from (3.13) that for any $1 \leqslant l \leqslant 2p-1$, $C_{n+1}(l) = \mathrm{o}(\log s_n)$ a.s. In addition, it is not hard to see from the induction assumption that for any $3 \leqslant l \leqslant 2p-1$, $D_n(l) = \mathrm{o}(\log s_n)$ a.s. We are now in a position to prove (2.8). One can observe that

$$D_n(2) = \sigma^2 \sum_{k=1}^{n} f_k \left(\frac{M_k^2}{s_k}\right)^{p-1}.$$

As f_n tends to zero, s_n is a.s. equivalent to s_{n-1}. Thus, we deduce from the induction assumption that a.s.

$$\lim_{n\to\infty} \frac{1}{\log s_n} D_n(2) = \sigma^2 l(p-1) \quad \text{with} \quad l(p) = \frac{\sigma^{2p}(2p)!}{2^p p!}. \tag{3.19}$$

Hence, we obtain from (3.19) that

$$\lim_{n\to\infty} \frac{1}{\log s_n} \mathscr{B}_{n+1} = C_{2p}^2 \sigma^2 l(p-1) = pl(p) \quad \text{a.s.} \tag{3.20}$$

Consequently, we find from (3.10) that

$$\lim_{n\to\infty} \frac{1}{\log s_n} \mathscr{A}_n = pl(p) \quad \text{a.s.} \tag{3.21}$$

We recall that

$$\mathscr{A}_n = \sum_{k=1}^{n} v_k(p) \left(\frac{M_k^2}{s_{k-1}}\right)^p \tag{3.22}$$

with

$$v_n(p) = \frac{s_n^p - s_{n-1}^p}{s_n^p} = f_n \sum_{q=0}^{p-1} \left(\frac{s_{n-1}}{s_n}\right)^{p-1-q}.$$

Hereafter, as $v_n(p)$ is a.s. equivalent to pf_n, convergence (2.8) follows from (3.21) and (3.22). □

3.3. Proof of Theorem 5

We also prove Theorem 5 by induction on the power $p \geqslant 1$. As before, Theorem 5 is already established for $p=1$ by formula (2.3) of Wei (1987). Next, let $p \geqslant 2$

and assume that Theorem 5 holds for any power q with $1 \leqslant q \leqslant p-1$. Recall that we have the decomposition

$$V_{n+1} + \mathscr{A}_n = V_1 + \mathscr{B}_{n+1} + \mathscr{W}_{n+1}. \tag{3.23}$$

On the one hand, as the explosion coefficient f_n converges a.s. to f, s_{n-1}/s_n tends a.s. to $1-f$ and $\log s_n$ is a.s. equivalent to $-n\log(1-f)$. Consequently, we deduce from (2.7) that $V_{n+1} = \mathrm{o}(n)$ a.s. On the other hand, it follows from Chow's lemma that

$$\lim_{n\to\infty} \frac{1}{n} \mathscr{W}_{n+1} = f^p \sigma(2p) \quad \text{a.s.} \tag{3.24}$$

Moreover, we already saw that for any $1 \leqslant l \leqslant 2p-1$, $C_{n+1}(l) = \mathrm{o}(n)$ a.s. which implies that a.s.

$$\mathscr{B}_{n+1} = \sum_{l=2}^{2p-1} C_{2p}^{l} D_n(l) + \mathrm{o}(n) \quad \text{with} \quad D_n(l) = \sum_{k=1}^{n} \varphi_k(l)\sigma_k(l).$$

We shall now study the asymptotic behavior of $D_n(l)$ in the two following cases.

Case 1: *Let* $2 \leqslant l \leqslant 2(p-1)$ *with* l *even.* We can find $1 \leqslant q \leqslant p-1$ such that $l = 2(p-q)$. Notice that $D_n(l)$ can be rewritten as

$$D_n(l) = \sum_{k=1}^{n} \sigma_k(2(p-q)) f_k^{p-q} \left(\frac{s_{k-1}}{s_k}\right)^q \left(\frac{M_k^2}{s_{k-1}}\right)^q,$$

it follows from (2.10) together with the induction assumption that

$$\lim_{n\to\infty} \frac{1}{n} D_n(l) = \sigma(2(p-q)) f^{p-q} (1-f)^q l(q,f) \quad \text{a.s.} \tag{3.25}$$

Case 2: *Let* $3 \leqslant l \leqslant 2p-1$ *with* l *odd.* As $\sigma_n(l)$ tends a.s. to zero,

$$|D_n(l)| = \mathrm{O}(1) + \mathrm{o}\left(\sum_{k=1}^{n} |\varphi_k(l)|\right) \quad \text{a.s.}$$

In addition, via the same arguments as in the proof of (3.14), we find that

$$\sum_{k=1}^{n} |\varphi_k(l)| = \mathrm{O}(n) \quad \text{a.s.}$$

which immediately ensures that $D_n(l) = \mathrm{o}(n)$ a.s. Therefore, we deduce from the conjunction of (3.23), (3.24) and (3.25) that

$$\lim_{n\to\infty} \frac{1}{n} \mathscr{A}_n = \sum_{k=1}^{p} C_{2p}^{2k} \sigma(2k) f^k (1-f)^{p-k} l(p-k,f) \quad \text{a.s.} \tag{3.26}$$

Finally, as $v_n(p)$ converges a.s. to $1-(1-f)^p$, we infer (2.11) from (3.22) and (3.26), which completes the proof of Theorem 5. □

3.4. Proof of Corollary 6

First of all, Corollary 6 is clearly true for $p=1$ as $l(1,f)=\sigma(2)=\sigma^2$. On the one hand, assume that for any $2 \leqslant k \leqslant p$, $\sigma(2k)=l(k)$. We have by induction

$$l(p,f)=\frac{1}{1-(1-f)^p}\left(\sum_{k=1}^{p} C_{2p}^{2k} f^k(1-f)^{p-k}\,\frac{\sigma^{2k}(2k)!}{2^k k!}\,l(p-k,f)\right),$$

$$=\frac{1}{1-(1-f)^p}\left(\frac{\sigma^{2p}(2p)!}{2^p}\sum_{k=1}^{p}\frac{f^k(1-f)^{p-k}}{k!(p-k)!}\right),$$

$$=\frac{1}{1-(1-f)^p}\left(l(p)\sum_{k=1}^{p} C_p^k f^k(1-f)^{p-k}\right)=l(p).$$

On the other hand, assume that $l(p,f)$ does not depend upon the random variable f. For any $p \geqslant 1$, as $l(p,f)$ is a continuous function of f, we necessarily have $l(p,f)=l(p,1)=\sigma(2p)$. Moreover, after some tedious calculation, we can prove the expansion

$$l(p,f)=l(p)+\sum_{k=2}^{p} P_{p,k}(f)\delta_k \tag{3.27}$$

with $\delta_k=\sigma(2k)-l(k)$, where the rational functions $P_{p,k}(f)$ may be explicitly calculated. For example,

$$P_{p,p}(f)=\frac{f^p}{1-(1-f)^p},$$

$$P_{p,p-1}(f)=\frac{C_{2p}^2 l(1)}{1-(1-f)^p}[f(1-f)^{p-1}P_{p-1,p-1}(f)+f^{p-1}(1-f)].$$

It remains to show that for any $2 \leqslant k \leqslant p$, $\delta_k=0$. We shall only carry out the proof that $\delta_p=0$ inasmuch as the rest of the proof follows essentially the same arguments than those for δ_p. First, for $p=2$, one can easily see that (3.27) reduces to $f(1-f)\delta_2=0$ which immediately implies that $\delta_2=0$. Next, for $p \geqslant 3$, we deduce from (3.27) that

$$(1-f^p-(1-f)^p)\delta_p=\sum_{k=2}^{p-1}(1-(1-f)^p)P_{p,k}(f)\delta_k.$$

Consequently, dividing this identity by $f(1-f)$, we find that

$$Q_p(f)\delta_p=\sum_{k=2}^{p-1} R_{p,k}(f)\delta_k \tag{3.28}$$

where

$$Q_p(f)=\frac{1-f^p-(1-f)^p}{f(1-f)}=\sum_{k=0}^{p-2} C_p^{k+1} f^k(1-f)^{p-2-k}$$

and the rational functions $R_{p,k}(f)=(f(1-f))^{-1}(1-(1-f)^p)P_{p,k}(f)$ may be explicitly calculated. For example,

$$R_{p,p-1}(f)=C_{2p}^2 l(1)[(1-f)^{p-2}P_{p-1,p-1}(f)+f^{p-2}].$$

The key point here is that $Q_p(0)=p$ whereas $R_{p,k}(0)=0$. Hence, the constant term in relation (3.28) is $p\delta_p$ which ensures that $\delta_p=0$. □

4. Statistical applications

Consider the stochastic regression model given, for all $n \geqslant 1$, by

$$X_n=\theta\phi_{n-1}+\varepsilon_n \tag{4.1}$$

where X_n, ϕ_n, and ε_n are the observation, the regression variable and the driven noise of the system, respectively. Assume that (ε_n) is a martingale difference sequence such that $\mathbb{E}[\varepsilon_{n+1}^2|\mathscr{F}_n]=\sigma^2$ a.s. In order to estimate the unknown real parameter θ, we use the least-squares estimator

$$\hat{\theta}_n=s_{n-1}^{-1}\sum_{k=1}^{n}\phi_{k-1}X_k \quad \text{where} \quad s_n=\sum_{k=0}^{n}\phi_k^2.$$

It immediately follows from (4.1) that $s_{n-1}(\hat{\theta}_n-\theta)=M_n$ with

$$M_n=\sum_{k=1}^{n}\phi_{k-1}\varepsilon_k.$$

Hence, (4.1) can be rewritten as

$$X_n-\hat{\theta}_{n-1}\phi_{n-1}=\pi_{n-1}+\varepsilon_n \tag{4.2}$$

with $\pi_n=-s_{n-1}^{-1}\phi_n M_n$. If (ε_n) satisfies the moment condition (1.2) and

$$\Delta_n=\frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^2,$$

then Δ_n converges a.s. to σ^2. First of all, we assume that (s_n) increases a.s. to infinity and that the explosion coefficient f_n tends to zero a.s. which leads to $\log s_n=o(n)$ a.s. Convergence (1.5) clearly implies that

$$\lim_{n\to\infty}\frac{1}{\log s_n}\sum_{k=1}^{n}\pi_k^2=\sigma^2 \quad \text{a.s.} \tag{4.3}$$

Consequently, we deduce from (4.2) and (4.3) that if

$$\Gamma_n=\frac{1}{n}\sum_{k=1}^{n}(X_k-\hat{\theta}_{k-1}\phi_{k-1})^2,$$

then Γ_n is a strongly consistent estimator of σ^2 with

$$\lim_{n\to\infty}\frac{n}{\log s_n}(\Gamma_n-\Delta_n)=\sigma^2 \quad \text{a.s.} \tag{4.4}$$

The purpose of this section is to propose strongly consistent estimators of higher-order moments of (ε_n). Recall that for any $q \geqslant 0$, $\sigma_n(q)=\mathbb{E}[\varepsilon_{n+1}^q|\mathscr{F}_n]$ with $\sigma_n(0)=1$, $\sigma_n(1)=0$ and $\sigma_n(2)=\sigma^2$. Moreover, define

$$\Gamma_n(q) = \frac{1}{n}\sum_{k=1}^{n}(X_k - \hat{\theta}_{k-1}\phi_{k-1})^q \quad \text{and} \quad \Delta_n(q) = \frac{1}{n}\sum_{k=1}^{n}\varepsilon_k^q.$$

Corollary 7. *Assume that* (ε_n) *satisfies, for some integer* $p \geqslant 1$, *the moment condition* (2.4). *If one can find some* $2 \leqslant q \leqslant 2p$ *such that* $\sigma_n(q) = \sigma(q)$ *a.s., then* $\Gamma_n(q)$ *is a strongly consistent estimator of* $\sigma(q)$ *with*

$$(\Gamma_n(q) - \Delta_n(q))^2 = \mathrm{O}\left(\frac{\log s_n}{n}\right) \quad \text{a.s.} \tag{4.5}$$

Remark. It follows from Chow's lemma that if, for some $p \geqslant 1$, (ε_n) satisfies (2.4) with $a > 2p$, then, for all $2 \leqslant q \leqslant 2p$, $\Delta_n(q)$ converges a.s. to $\sigma(q)$ with the rate of convergence

$$|\Delta_n(q) - \sigma(q)| = \mathrm{o}\left(\frac{n^c}{n}\right) \quad \text{a.s.}$$

where c is such that $2pa^{-1} < c < 1$. Consequently, as soon as $\log s_n = \mathrm{o}(n^c)$ a.s., we infer from (4.5) that

$$(\Gamma_n(q) - \sigma(q))^2 = \mathrm{o}\left(\frac{n^c}{n}\right) \quad \text{a.s.}$$

Another application of Theorem 3 concerns the convergence in average of the estimation error $(\hat{\theta}_n - \theta)^{2p}$.

Corollary 8. *Assume that* (ε_n) *satisfies, for some integer* $p \geqslant 1$, *the moment condition* (2.4). *Then*

$$\lim_{n\to\infty} \frac{1}{\log s_n}\sum_{k=1}^{n} f_k s_k^p (\hat{\theta}_k - \theta)^{2p} = \frac{\sigma^{2p}(2p)!}{2^p p!} \quad a.s. \tag{4.6}$$

In addition, assume that for some positive constant τ

$$\lim_{n\to\infty} \frac{1}{n} s_n = \tau \quad a.s. \tag{4.7}$$

Then

$$\lim_{n\to\infty} \frac{1}{\log n}\sum_{k=1}^{n} k^{p-1}(\hat{\theta}_k - \theta)^{2p} = \frac{\sigma^{2p}(2p)!}{\tau^p 2^p p!} \quad a.s. \tag{4.8}$$

Example 1. If we choose $\phi_n = X_n$, we can rewrite (4.1) as the linear autoregressive model

$$X_n = \theta X_{n-1} + \varepsilon_n. \tag{4.9}$$

On the one hand, in the stable case $|\theta| < 1$, $f_n \to 0$ a.s. and s_n/n converges a.s. to $\sigma^2/(1-\theta^2)$ so that $\log s_n$ is a.s. equivalent to $\log n$ (see e.g. Duflo, 1997 and Lai and Wei, 1983). Thus, (4.5) holds replacing $\log s_n$ by $\log n$. Moreover, it follows from (4.8) that

$$\lim_{n\to\infty} \frac{1}{\log n} \sum_{k=1}^{n} k^{p-1}(\hat{\theta}_k - \theta)^{2p} = \frac{(1-\theta^2)^p(2p)!}{2^p p!} \quad \text{a.s.}$$

On the other hand, in the unstable case $|\theta| = 1$, once again $f_n \to 0$ but s_n/n^2 diverges. However, by formula (3.5) of Wei, 1987, $\log s_n$ is a.s. equivalent to $2\log n$. Consequently, (4.5) is true replacing $\log s_n$ by $\log n$.

Example 2. If we choose $\phi_n = f(X_n)$ where f is a known real function, we can rewrite (4.1) as the parametric functional autoregressive model

$$X_n = \theta f(X_{n-1}) + \varepsilon_n.$$

Assume that for all x in $\mathbb{R}$,

$$c|x| + d \leqslant |f(x)| \leqslant a|x| + b$$

where $0 < a|\theta| < 1$, $b, c \geqslant 0$ and $d > 0$ if $c = 0$, $d \geqslant 0$ otherwise. One can easily check that $n = O(s_n)$, $f_n \to 0$ and $s_n = O(n)$ so that $\log s_n = O(\log n)$ a.s which implies that (4.5) holds replacing $\log s_n$ by $\log n$.

Hereafter, we assume that the explosion coefficient f_n converges a.s. to a random variable f with $0 < f < 1$. On the one hand, s_n grows exponentially fast to infinity and the expression of $l(p, f)$, given by Theorem 5, depends on all the moments $\sigma(2k)$ with $1 \leqslant k \leqslant p$. Consequently, it is possible but rather intricated to estimate the even moments of (ε_n) as in Corollary 7. On the other hand, even if s_n grows exponentially fast to infinity, the estimator $\hat{\theta}_n$ is self-normalized. Therefore, we can propose a direct application of Theorem 5 similar to Corollary 8.

Corollary 9. *Assume that (ε_n) satisfies, for some integer $p \geqslant 1$, the moment conditions (2.4) and (2.10). Then*

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} s_{k-1}^p(\hat{\theta}_k - \theta)^{2p} = l(p, f) \quad a.s. \tag{4.10}$$

In addition, assume that for some positive random variable τ

$$\lim_{n\to\infty} (1-f)^n s_n = \tau \quad a.s. \tag{4.11}$$

Then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \frac{(\hat{\theta}_k - \theta)^{2p}}{(1-f)^{kp}} = \frac{l(p, f)}{\tau^p(1-f)^p} \quad a.s. \tag{4.12}$$

Example 3. Consider once again the linear autoregressive model given by (4.9). In the explosive case $|\theta| > 1$, $\theta^{-n}X_n$ converges a.s. and in mean square to the nonzero random variable

$$Y = X_0 + \sum_{k=1}^{\infty} \theta^{-k} \varepsilon_k.$$

Hence, we directly obtain via Toeplitz's lemma that $f_n \to (\theta^2 - 1)/\theta^2$ a.s. and s_n/θ^{2n} converges a.s. to $\theta^2 Y^2/(\theta^2 - 1)$ (see e.g. Bercu, 2001; Duflo, 1997 and Lai and Wei, 1983). Consequently, it follows from (4.12) that

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} (\theta^k(\hat{\theta}_k - \theta))^{2p} = \frac{(\theta^2 - 1)^p l(p, f)}{Y^{2p}} \quad \text{a.s.}$$

5. Proofs of statistical results

5.1. Proof of Corollary 7

We already saw via (4.4) that Corollary 7 holds for $q = 2$. Next, let $q \geqslant 3$ and assume that $\sigma_n(q) = \sigma(q)$ a.s. It follows from (4.2) that for any $n \geqslant 1$

$$n(\Gamma_n(q) - \Delta_n(q)) = P_{n-1}(q) + Q_n(q) \tag{5.1}$$

where

$$P_n(q) = \sum_{k=0}^{n} \pi_k^q,$$

$$Q_n(q) = \sum_{l=1}^{q-1} C_q^l R_n(l) \quad \text{with} \quad R_n(l) = \sum_{k=0}^{n-1} \pi_k^{q-l} \varepsilon_{k+1}^l.$$

As f_n tends to zero, s_n is a.s. equivalent to s_{n-1}. Thus, we deduce from (2.8) that for any $2 \leqslant r \leqslant p$

$$\lim_{n\to\infty} \frac{1}{\log s_n} \sum_{k=1}^{n} f_k \frac{\pi_k^{2r}}{f_k^r} = \frac{\sigma^{2r}(2r)!}{2^r r!} \quad \text{a.s.}$$

which implies that

$$\sum_{k=0}^{n} \pi_k^{2r} = \mathrm{o}(\log s_n) \quad \text{a.s.} \tag{5.2}$$

On the one hand, if q is even, we can find $2 \leqslant r \leqslant p$ such that $q = 2r$. Hence, we immediately obtain from (5.2) that

$$P_n(q) = \mathrm{o}(\log s_n) \quad \text{a.s.} \tag{5.3}$$

On the other hand, if q is odd, we also derive from (5.2) together with the Cauchy–Schwarz inequality that (5.3) still holds. Next, via the same way that (3.12) is established in the proof of Theorem 2, we find that

$$|Q_n(q)|^2 = \mathrm{O}(n \log s_n) \quad \text{a.s.} \tag{5.4}$$

Consequently, (4.5) follows from (5.1), (5.3) and (5.4), which completes the proof of Corollary 7. □

5.2. Proof of Corollary 8

As (4.6) is a straightforward application of (2.8), we only have to prove (4.8). For any sequence (a_n), it is not hard to see that

$$\sum_{k=1}^{n} a_k \phi_k^2 = a_n(s_n - n\tau) - a_1 s_0 + \tau \sum_{k=1}^{n} a_k + r_n \tag{5.5}$$

where

$$r_n = \sum_{k=1}^{n-1} (a_k - a_{k+1})(s_k - k\tau).$$

We shall now choose

$$a_n = \frac{1}{s_n} \frac{M_n^{2p}}{s_{n-1}^p} = \frac{s_{n-1}^p (\hat{\theta}_n - \theta)^{2p}}{s_n}.$$

On the one hand, (2.5) directly implies that $a_n(s_n - n\tau) = \mathrm{o}(\log s_n)$ a.s. On the other hand, proceeding exactly as in the proof of the second part of Theorem 2, we infer that $r_n = \mathrm{o}(\log s_n)$ a.s. Therefore, we deduce from (2.8) and (5.5) that a.s.

$$\lim_{n\to\infty} \frac{1}{\log s_n} \sum_{k=1}^{n} a_k = \frac{l(p)}{\tau} \quad \text{with} \quad l(p) = \frac{\sigma^{2p}(2p)!}{2^p p!}. \tag{5.6}$$

Finally, we find via Toeplitz's lemma together with (4.7) and (5.6) that (4.8) holds, which completes the proof of Corollary 8. □

Acknowledgements

The author is deeply grateful to J.C. Fort, F. Gamboa, M. Lifshits and A. Touati for fruitful discussions. He also thanks the anonymous referees for their very careful reading of the manuscript.

References

Bercu, B., 2001. On large deviations in the Gaussian autoregressive process: stable, unstable and explosive cases. Bernoulli 7, 299–316.

Bercu, B., Portier, B., 2002. Adaptive control of parametric nonlinear autoregressive models via a new martingale approach. IEEE Trans. Automat. Control 47, 1524–1528.

Berkes, I., Csáki, E., 2001. A universal result in almost sure central limit theory. Stochastic Process. Appl. 94, 105–134.

Brosamler, G.A., 1988. An almost sure everywhere central limit theorem. Math. Proc. Cambridge Philos. Soc. 104, 561–574.

Chaabane, F., 1996. Version forte du théorème de la limite centrale fonctionnel pour les martingales. C.R. Acad. Sci. Paris 323, 195–198.

Chaabane, F., 2001. Invariance principle with logarithm averaging for martingales. Studia Math. Sci. Hungar. 37, 21–52.

Chaabane, F., Maaouia, F., 2000. Théorèmes limites avec poids pour les martingales vectorielles. Esaim Prob. Stat. 4, 137–189.

Duflo, M., 1997. Random Iterative Models. Springer, Berlin.
Ibragimov, I.A., Lifshits, M.A., 1998. On the convergence of generalized moments in almost sure central limit theorem. Statist. Probab. Lett. 40, 343–351.
Ibragimov, I.A., Lifshits, M.A., 1999. On almost sure limit theorems. Theory Probab. Appl. 44, 254–272.
Lai, T.L., Wei, C.Z., 1982. Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems. Ann. Statist. 10, 154–166.
Lai, T.L., Wei, C.Z., 1983. Asymptotic properties of general autoregressive models and strong consistency of least squares estimates of their parameters. J. Multivariate Anal. 13, 1–23.
Lacey, M.T., Phillip, W., 1990. A note on the almost sure central limit theorem. Statist. Probab. Lett. 9, 201–205.
Lifshits, M., 2001. Lecture Notes on Almost Sure Limit Theorems, Vol. 54. Publications IRMA, Lille, pp. 1–25.
Lifshits, M., 2002. Almost sure limit theorem for martingales. In: Berkes, I., Csáki, E., Csörgő, M. (Eds.), Limit Theorems in Probability and Statistics II. J. Bolyai Mathematical Society, Budapest, pp. 367–390.
Neveu, J., 1975. Discrete Parameter Martingales. North-Holland, Amsterdam.
Schatte, P., 1988. On strong versions of the almost sure central limit theorem. Math. Nachr. 137, 249–256.
Schatte, P., 1991. On the central limit theorem with almost sure convergence. Probab. Math. Statist. 11, 237–246.
Wei, C.Z., 1987. Adaptive prediction by least squares predictors in stochastic regression models with applications to time series. Ann. Statist. 15, 1667–1682.