PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 22, Fasc. 1 (2002), pp. 101–113

ON HARTMAN'S LAW OF ITERATED LOGARITHM FOR EXPLOSIVE GAUSSIAN AUTOREGRESSIVE PROCESSES

BY

BERNARD BERCU (ORSAY) AND ABDERRAHMEN TOUATI (PARIS)

Abstract. A law of iterated logarithm is established for the maximum likelihood estimator of the unknown parameter of the explosive Gaussian autoregressive process. Outside the Gaussian case, we show that the law of iterated logarithm does not hold, except for a suitable averaging on the maximum likelihood estimator.

2000 AMS Subject Classification: Primary 60G15; Secondary 60F15, 62A10.

Key words and phrases: Autoregressive Gaussian process, law of iterated logarithm, maximum likelihood estimation.

1. INTRODUCTION

Since the pioneer works of Kolmogorov, Marcinkiewicz and Zygmund, and Hartman and Winter [10], [15], [16], [18], a wide literature concerning the Law of Iterated Logarithm (LIL) for sequences of independent random variables has been available. Many probabilists have attempted to find minimal conditions under which the LIL holds. A typical result is the general LIL established by Wittmann [23]. Let (ξ_n) be a sequence of independent random variables with $E[\xi_n] = 0$ and $E[\xi_n^2] = a_n$. Define

$$S_n = \sum_{k=1}^{n} \xi_k, \qquad A_n = \sum_{k=1}^{n} a_k.$$

THEOREM 1.1. *Assume that*

$$\lim_{n\to+\infty} A_n = \infty \quad \textit{and} \quad \limsup_{n\to+\infty} \frac{A_{n+1}}{A_n} < \infty,$$

and, for $2 < \alpha \leqslant 3$,

$$\sum_{n=1}^{\infty} \frac{E[|\xi_n|^\alpha]}{(A_n \log_2 A_n)^{\alpha/2}} < \infty. \tag{1.1}$$

Then we have the LIL

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2A_n \log_2 A_n}} = 1 \quad a.s. \tag{1.2}$$

On the one hand, if (ξ_n) are identically distributed, Theorem 1.1 contains the classical result of Hartman and Winter [10]. On the other hand, if $\sup E[|\xi_n|^\alpha] < \infty$ with $\alpha > 2$, then condition (1.1) immediately implies that $a_n = o(A_n \log_2 A_n)$. In fact, with additional suitable assumptions on (ξ_n), Tomkins has shown [19] that it is possible to establish the LIL when $a_n = o(A_n)$. One might wonder if such a result is true with a less restrictive assumption on (a_n). This question has been positively answered by Hartman [9] for Gaussian random variables.

THEOREM 1.2. *Assume that (ξ_n) are independent with $N(0, a_n)$ distribution. Moreover, assume that*

$$\lim_{n\to+\infty} A_n = \infty \quad \text{and} \quad \limsup_{n\to+\infty} \frac{a_n}{A_n} < 1. \tag{1.3}$$

Then we have the LIL

$$\limsup_{n\to\infty} \frac{|S_n|}{\sqrt{2A_n \log_2 A_n}} = 1 \quad a.s. \tag{1.4}$$

One can observe that if $A_n \sim ca_n$ with $c > 1$, the series given by (1.1) diverges and the strongest available result is Hartman's LIL.

The purpose of this paper is to prove LIL for explosive martingales, i.e. for martingales with increasing processes growing exponentially fast to infinity. Instead of proposing a rather technical general theory, we have deliberately chosen to enlighten our approach by focusing our attention on the very instructive explosive Gaussian autoregressive process.

The paper is organized as follows. Section 2 is devoted to the LIL for the maximum likelihood estimator of the unknown parameter of the explosive Gaussian autoregressive process. In Section 3, we prove that we can do without the normality assumption via a suitable averaging on the maximum likelihood estimator. However, without averaging, we show in Section 4 that for some simple explosive martingales, it is impossible to get rid of the normality assumption on (ξ_n).

2. MAIN RESULTS

Consider the autoregressive process of order $p \geqslant 1$ given, for all $n \geqslant 0$, by

$$X_n = \sum_{k=1}^{p} a_k X_{n-k} + \varepsilon_n, \tag{2.1}$$

where the initial state $(X_0, \ldots, X_{-p+1})$ is a square-integrable random vector and (ε_n) are i.i.d. distributed as $\mathcal{N}(0, \sigma^2)$ with $\sigma > 0$. Moreover, assume that (ε_n) are independent of $(X_0, \ldots, X_{-p+1})$. Let A be the companion matrix associated with (2.1)

$$A = \begin{bmatrix} a_1 & a_2 & \ldots & a_{p-1} & a_p \\ 1 & 0 & \ldots & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 & 0 \end{bmatrix}.$$

In all the sequel, we assume that we are in an explosive situation, i.e. all the eigenvalues of A lie outside the unit circle. Relation (2.1) can be rewritten as

$$\Phi_n = A\Phi_{n-1} + e_n, \qquad \Phi_n = A^n \Phi_0 + \sum_{k=1}^{n} A^{n-k} e_k, \tag{2.2}$$

where $\Phi_n^t = (X_n, \ldots, X_{n-p+1})$ and $e_n^t = (\varepsilon_n, 0, \ldots, 0)$. Hence, it follows from (2.2) that $Y_n = A^{-n}\Phi_n$ converges a.s. and in mean square to

$$Y = \Phi_0 + \sum_{k=1}^{\infty} A^{-k} e_k. \tag{2.3}$$

The maximum likelihood estimator of the unknown parameter $\theta^t = (a_1, \ldots, a_p)$, which coincides with the least-squares estimator, is given by

$$\hat{\theta}_n = Q_{n-1}^{-1} \sum_{k=1}^{n} \Phi_{k-1} X_k, \qquad Q_n = \sum_{k=0}^{n} \Phi_k \Phi_k^t + I_p, \tag{2.4}$$

where the identity matrix I_p is added in order to avoid useless invertibility assumptions. First of all, it was shown in [8] and [14] that

$$\lim_{n\to\infty} A^{-n} Q_{n-1} (A^{-n})^t = L \text{ a.s.}, \tag{2.5}$$

where L is the a.s. invertible matrix

$$L = \sum_{k=1}^{\infty} A^{-k} Y Y^t (A^{-k})^t. \tag{2.6}$$

Moreover, the asymptotic behavior in law of $\hat{\theta}_n$, properly normalized, is known from the earlier papers of Anderson [1], Rao [17], White [22] and the recent extensions obtained by Touati [20], [21]. To be more precise, we have

$$A^{-n} Q_{n-1} (\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} H, \qquad (A^n)^t (\hat{\theta}_n - \theta) \xrightarrow{\mathcal{L}} L^{-1} H$$

with

$$H = \sum_{k=1}^{\infty} A^{-k} Y \zeta_k,$$

where (ζ_n) are i.i.d. with $\mathcal{N}(0, \sigma^2)$ distribution, and (ζ_n) are independent of Φ_0 and (ε_n). In the particular case $p = 1$ and $X_0 = 0$, the limiting distribution $L^{-1} H$ is Cauchy. It is an important question to know, in addition, whether or not an LIL holds and it is the purpose of this paper to show that it is the case. Surprisingly, in contrast with the estimation theory for supercritical Galton–Watson processes illustrated by the important contribution of Heyde [5], [11], [12], [13], this question has not been tackled in the literature for explosive autoregressive processes.

THEOREM 2.1. *For any left eigenvector f of A associated with the eigenvalue λ, we have the LIL*

$$\limsup_{n\to\infty} \frac{|f^* Q_{n-1} (\hat{\theta}_n - \theta)|}{|\lambda|^n \sqrt{2 \log n}} = \frac{\sigma |\langle f, Y\rangle|}{\sqrt{|\lambda|^2 - 1}} \quad a.s., \tag{2.7}$$

$$\limsup_{n\to\infty} \frac{|f^* L (A^n)^t (\hat{\theta}_n - \theta)|}{\sqrt{2 \log n}} = \frac{\sigma |\langle f, Y\rangle|}{\sqrt{|\lambda|^2 - 1}} \quad a.s., \tag{2.8}$$

where Y and L are given by (2.3) *and* (2.6), *respectively.*

COROLLARY 2.1. *In the scalar case $p = 1$, the maximum likelihood estimator $(\hat{\theta}_n)$ follows the LIL*

$$\limsup_{n\to\infty} \frac{|\theta|^n |\hat{\theta}_n - \theta|}{\sqrt{2 \log n}} = \frac{\sigma \sqrt{\theta^2 - 1}}{|Y|} \quad a.s.$$

Remark. It was shown in [14] that, for all $v \in \mathbf{R}^p$ with $v \neq 0$, $\langle v, Y\rangle$ has a continuous distribution. Hence, in the scalar case, $Y \neq 0$ a.s. Moreover, one can also estimate θ by the Yule–Walker estimator

$$\tilde{\theta}_n = Q_n^{-1} \sum_{k=1}^{n} X_k X_{k-1}. \tag{2.9}$$

It is easy to see that $\tilde{\theta}_n$ converges a.s. to θ^{-1} and we infer from Corollary 2.1 that

$$\limsup_{n\to\infty} \frac{|\theta|^n |\tilde{\theta}_n - \theta^{-1}|}{\sqrt{2 \log n}} = \frac{\sigma \sqrt{\theta^2 - 1}}{\theta^2 |Y|}.$$

In order to establish a more general result on the LIL associated with $(\hat{\theta}_n)$, we need some algebraic preliminaries on the Jordan canonical form associated with the companion matrix A. One can find an invertible matrix P such that PAP^{-1} is a Jordan matrix J, i.e. a direct sum of q Jordan blocs $J_1, J_2, \ldots, J_q$, where, for $1 \leqslant i \leqslant q$, J_i is either a scalar $\lambda_i \in C$ or a square matrix of order $d_i > 1$ with diagonal terms each equal to $\lambda_i \in C$, $d_i - 1$ terms each equal to one in the subdiagonal and all other entries equal to zero. We have $d_1 + d_2 + \ldots + d_q = p$ and the orders d_i may not be distinct and the values λ_i may not be distinct either. Moreover, for $1 \leqslant i \leqslant q$, let f_i be the left eigenvector associated with the Jordan block J_i and set $f = (f_1^t, f_2^t, \ldots, f_q^t)$. Furthermore, for any vector v of C^p, set $v = (v_1^t, v_2^t, \ldots, v_q^t)$, where, for $1 \leqslant i \leqslant q$, v_i is a vector of C^{d_i}. Finally, for any vectors u and v of C^p, define

$$\Delta(u, v) = \sum_{i=1}^{q} \|u_i\|_1 \|v_i\|_1 \quad \text{and} \quad \Delta(u) = \sup\{\Delta(u, v) \text{ with } \|v\|_2 = 1\}.$$

THEOREM 2.2. *Let $\varrho(A)$ be the spectral radius of A and denote by ν the index associated with $\varrho(A)$. Then, for any vector v of C^p, we have the LIL*

(2.10)
$$\limsup_{n\to\infty} \frac{|v^* P Q_{n-1}(\hat{\theta}_n - \theta)|}{n^{(\nu-1)} \varrho(A)^n \sqrt{2\log n}} \leqslant \frac{\sigma \|f\|_\infty}{\sqrt{\varrho(A)^2 - 1}} \left(\frac{\varrho(A)^{-(\nu-1)}}{(\nu-1)!}\right) \Delta(PY, v) \text{ a.s.},$$

which implies that

(2.11)
$$\limsup_{n\to\infty} \frac{\|Q_{n-1}(\hat{\theta}_n - \theta)\|^2}{2n^{2(\nu-1)} \varrho(A)^{2n} \log n} \leqslant \frac{\sigma^2 \|f\|_\infty^2}{\varrho(A)^2 - 1} \left(\frac{\varrho(A)^{-(\nu-1)}}{(\nu-1)!}\right)^2 \frac{\Delta^2(PY)}{\lambda_{\min}(P^* P)} \text{ a.s.}$$

In addition, for any vector v of C^p, we also have

$$\limsup_{n\to\infty} \frac{|v^* L(A^n)^t(\hat{\theta}_n - \theta)|}{\sqrt{2\log n}} = \sigma\Big(\sum_{k=1}^{\infty} (v^* A^{-k} Y)^2\Big)^{1/2} \text{ a.s.} \tag{2.12}$$

More particularly,

$$\limsup_{n\to\infty} \frac{\|(A^n)^t(\hat{\theta}_n - \theta)\|^2}{2\log n} \leqslant \sigma^2 \sum_{k=1}^{\infty} \|L^{-1} A^{-k} Y\|^2 \text{ a.s.} \tag{2.13}$$

Proof. First of all, it follows from (2.4) that $Q_{n-1}(\hat{\theta}_n - \theta) = M_n$ with

$$M_n = M_0 + \sum_{k=1}^{n} \Phi_{k-1} \varepsilon_k \quad \text{and} \quad M_0 = -\theta.$$

Moreover, relation (2.2) can be rewritten as

$$\Phi_n = A^n Y - R_n, \tag{2.14}$$

where

$$Y = \Phi_0 + \sum_{k=1}^{\infty} A^{-k} \varepsilon_k \quad \text{and} \quad R_n = \sum_{k=1}^{\infty} A^{-k} \varepsilon_{n+k}.$$

Consequently, we infer from (2.14) that $M_n = M_0 + U_n Y - V_n$ with

$$U_n = \sum_{k=1}^{n} A^{k-1} \varepsilon_k \quad \text{and} \quad V_n = \sum_{k=1}^{n} R_{k-1} \varepsilon_k. \tag{2.15}$$

On the one hand, we have

$$A V_n = \sum_{k=1}^{n} e_k \varepsilon_k + \sum_{k=1}^{n} R_k \varepsilon_k. \tag{2.16}$$

For any vector v of C^p, if $w_k = v^* R_k \varepsilon_k$, it is not hard to see that (w_k) is a sequence of square-integrable random variables such that

$$E[w_k] = 0, \quad E[w_k \bar{w}_l] = 0 \text{ if } k \neq l$$

and

$$E[|w_k|^2] \leqslant \sigma^4 \|v\|^2 \varrho^2 (1-\varrho^2)^{-1} \quad \text{with } 0 < \varrho(A^{-1}) < \varrho < 1,$$

where $\varrho(A^{-1})$ denotes the spectral radius of A^{-1}. Then, it follows from the Rademacher–Menchoff theorem (see e.g. Theorem 2.3.2 of [18]) together with the usual strong law of large numbers that, for any vector v of C^p,

$$\lim_{n\to\infty} n^{-1} v^* A V_n = \sigma^2 \langle v, \delta_1 \rangle \text{ a.s.,} \tag{2.17}$$

where $\delta_1^t = (1, 0, \ldots, 0)$. Hence, we obtain $\|V_n\| = O(n)$ a.s. On the other hand, for the sequence (U_n), assume first of all that f is a left eigenvector of A associated with the eigenvalue λ. For any vector u of C^p, if $\xi_n = \lambda^{n-1} \langle f, u \rangle \varepsilon_n$, then (ξ_n) is a sequence of independent random variables distributed as $N(0, a_n)$ with $a_n = \sigma^2 |\lambda|^{2(n-1)} |\langle f, u\rangle|^2$. Consequently, we find by the use of Theorem 1.2 that

$$\limsup_{n\to\infty} \frac{|f^* U_n u|}{|\lambda|^n \sqrt{2 \log n}} = \frac{\sigma |\langle f, u\rangle|}{\sqrt{|\lambda|^2 - 1}} \text{ a.s.} \tag{2.18}$$

Hence (2.7) and (2.8) follow from the conjunction of (2.5), (2.17) and (2.18). More generally, for any vector v of C^p, we infer from (2.15) that

$$v^* P U_n P^{-1} = \sum_{i=1}^{q} \sum_{k=1}^{n} v_i^* J_i^{k-1} \varepsilon_k. \tag{2.19}$$

We shall focus our attention on a particular Jordan block J of dimension d, associated with the eigenvalue $\lambda \in C$ with $|\lambda| > 1$, where the index i is omitted in order to avoid heaviness in the notation. There exists a basis $(f_1, f_2, \ldots, f_d)$ of C^d such that $f_1^* J = \lambda f_1^*$ and, for $2 \leqslant j \leqslant d$, $f_j^* J = \lambda f_j^* + f_{j-1}^*$. Hence $f_1^* J^k = \lambda^k f_1^*$ and, for $2 \leqslant j \leqslant d$,

$$f_j^* (J - \lambda I_d)^k = f_{j-k}^* \text{ if } 1 \leqslant k \leqslant j-1 \quad \text{and} \quad f_j^* (J - \lambda I_d)^k = 0 \text{ if } j \leqslant k \leqslant d.$$

Therefore, one can easily check that for all $1 \leqslant j \leqslant d$

$$\sum_{k=1}^{n} f_j^* J^{k-1} \varepsilon_k = \sum_{k=1}^{n} \sum_{l=1}^{j-1} \varphi_k(l) f_l^* \varepsilon_k,$$

where $\varphi_k(1) = \lambda^{k-1}$ and, for $k \geqslant d$, $\varphi_k(l) = C_{k-1}^{j-l} \lambda^{k-j+l-1}$. Furthermore, for any vector u of C^d and for all $1 \leqslant j \leqslant d$, set

$$\xi_n = \sum_{l=1}^{j-1} \varphi_n(l) \langle f_l, u \rangle \varepsilon_n.$$

As before, (ξ_n) is a sequence of independent random variables distributed as $N(0, a_n)$ with

$$a_n = \sigma^2 \Big| \sum_{l=1}^{j-1} \varphi_n(l) \langle f_l, u \rangle \Big|^2.$$

We can show after some straightforward calculations that

$$A_n \sim \frac{\sigma^2 |\langle f_l, u \rangle|^2}{((j-1)!)^2} \left(\frac{|\lambda|^2}{|\lambda|^2 - 1} \right) n^{2(j-1)} |\lambda|^{2(n-j)}.$$

Consequently, it immediately follows from Theorem 1.2 that for any vector u of C^d and for all $1 \leqslant j \leqslant d$

$$\limsup_{n \to \infty} \frac{\big| \sum_{k=1}^{n} f_j^* J^{k-1} \varepsilon_k u \big|}{n^{(j-1)} |\lambda|^{(n-j)} \sqrt{2 \log n}} = \frac{\sigma |\langle f_1, u \rangle|}{(j-1)!} \left(\frac{|\lambda|^2}{|\lambda|^2 - 1} \right)^{1/2} \quad \text{a.s.} \tag{2.20}$$

We shall now put together the results obtained for the various components in (2.19). Let $\varrho(A)$ be the spectral radius of A and denote by ν the index associated with $\varrho(A)$, that is the order of the largest block corresponding to $\varrho(A)$. Then, we deduce from (2.19) together with (2.20) that, for any vectors u and v of C^p,

$$\limsup_{n \to \infty} \frac{|v^* P U_n P^{-1} u|}{n^{(\nu-1)} \varrho(A)^n \sqrt{2 \log n}} \leqslant \frac{\sigma \|f\|_\infty}{\sqrt{\varrho(A)^2 - 1}} \left(\frac{\varrho(A)^{-(\nu-1)}}{(\nu-1)!} \right) \Delta(u, v) \quad \text{a.s.} \tag{2.21}$$

Therefrom, (2.17) and (2.21) imply (2.10) and (2.11). Furthermore, for any vectors u and v of C^p, set

$$W_n = v^* A^{-n} U_n u = \sum_{k=1}^{n} v^* A^{-(n-k+1)} u \varepsilon_k.$$

The random variable W_n is distributed as $\mathcal{N}(0, \sigma_n^2)$ with

$$\sigma_n^2 = \sigma^2 \sum_{k=1}^{n} (v^* A^{-k} u)^2.$$

Hence, it is well known that

$$\limsup_{n\to\infty} \frac{|W_n|}{\sqrt{2\log n}} = \sigma\Big(\sum_{k=1}^{\infty} (v^* A^{-k} u)^2\Big)^{1/2} \text{ a.s.} \tag{2.22}$$

Finally, we deduce (2.12) and (2.13) from (2.17) and (2.22), completing the proof of Theorem 2.2. ∎

3. EXTENSION

One might wonder if an LIL similar to that of Theorem 2.2 holds without the normality assumption on (ε_n). As we shall see in Section 4, we cannot do without this normality assumption. However, we shall now prove that, via a suitable averaging on $(\hat{\theta}_n)$, it is possible to obtain an LIL with the only hypothesis that (ε_n) is a martingale difference sequence with constant conditional variance σ^2 and finite conditional moment of order greater than 2. The main tool for proving LIL in the martingale framework is given by the following lemma (see [7] and [18]).

Let $F = (\mathscr{F}_n)$ be a sequence of nondecreasing σ-algebras. Moreover, let (ξ_n) be a martingale difference sequence adapted to F such that, for all $n \geqslant 0$, $E[\xi_{n+1}^2 \mid \mathscr{F}_n] = \sigma^2$ with $\sigma > 0$ and $\sup E[|\xi_{n+1}|^\alpha \mid \mathscr{F}_n] < \infty$ a.s. for some $\alpha > 2$. For a p-dimensional sequence of random vectors (φ_n) adapted to F, define

$$M_n = \sum_{k=1}^{n} \varphi_{k-1} \xi_k, \quad P_n = \sum_{k=0}^{n} \varphi_k \varphi_k^t + I_p,$$

where the identity matrix I_p is added in order to avoid useless invertibility assumptions.

LEMMA 3.1. *Let (c_n) be a deterministic real sequence increasing to infinity. Assume that $c_n^{-1} P_{n-1}$ converges a.s. to a finite random matrix L and that*

$$\sum_{n=1}^{\infty} \left(\frac{\|\varphi_n\|}{\sqrt{c_n}}\right)^{\beta} < \infty \text{ a.s.} \tag{3.1}$$

with $2 < \beta \leqslant \alpha$. Then, for any vector $v \in \mathbf{R}^p$ such that $v^t L v > 0$, we have the LIL

$$\limsup_{n\to\infty} \left(\frac{1}{2c_n \log_2 c_n}\right)^{1/2} |v^t M_n| = \sigma (v^t L v)^{1/2} \text{ a.s.} \tag{3.2}$$

Consider the autoregressive process of order $p \geqslant 1$ given, for all $n \geqslant 0$, by

$$X_n = \sum_{k=1}^{p} a_k X_{n-k} + \varepsilon_n.$$

Denote by $F = (\mathscr{F}_n)$ the natural filtration $\mathscr{F}_n = \sigma(\Phi_0, \varepsilon_1, \ldots, \varepsilon_n)$. Assume that the initial state Φ_0 is a square-integrable random vector independent of (ε_n).

Moreover, assume that (ε_n) is a martingale difference sequence adapted to F such that, for all $n \geqslant 0$, $E[\varepsilon_{n+1}^2 | \mathcal{F}_n] = \sigma^2 > 0$ and, for some $\alpha > 2$,

$$\sup_n E[|\varepsilon_{n+1}|^\alpha | \mathcal{F}_n] < \infty \text{ a.s.} \tag{3.3}$$

THEOREM 3.1. *For any vector $v \in \mathbf{R}^p$, we have the LIL*

$$\limsup_{n\to\infty} \left(\frac{1}{2n\log_2 n}\right)^{1/2} |\sum_{k=1}^{n} v^t A^{-k} Q_{k-1}(\hat{\theta}_k - \theta)| = \sigma |v^t (A - I_p)^{-1} Y| \text{ a.s.,} \tag{3.4}$$

$$\limsup_{n\to\infty} \left(\frac{1}{2n\log_2 n}\right)^{1/2} |\sum_{k=1}^{n} v^t (A^{-k})^t (\hat{\theta}_k - \theta)| = \sigma |v^t (AL - L)^{-1} Y| \text{ a.s.} \tag{3.5}$$

with Y and L given by (2.3) *and* (2.6), *respectively.*

Remark. It is possible to establish a similar result to Theorem 3.1 for multitype supercritical branching processes.

COROLLARY 3.1. *In the scalar case $p = 1$, the least-squares estimator $(\hat{\theta}_n)$ follows the LIL*

$$\limsup_{n\to\infty} \left(\frac{1}{2n\log_2 n}\right)^{1/2} |\sum_{k=1}^{n} \theta^k (\hat{\theta}_k - \theta)| = \sigma \left|\frac{\theta+1}{Y}\right| \text{ a.s.}$$

Remark. As in the previous section, since $\liminf E[\varepsilon_{n+1}^2 | \mathcal{F}_n] = \sigma^2$ with $\sigma > 0$, the random variable $\langle v, Y \rangle$ has a continuous distribution for all $v \in \mathbf{R}^p$ with $v \neq 0$. Hence, in the scalar case, $Y \neq 0$ a.s. Moreover, we immediately infer from Corollary 3.1 that the Yule–Walker estimator $\tilde{\theta}_n$ given by (2.9) satisfies

$$\limsup_{n\to\infty} \left(\frac{1}{2n\log_2 n}\right)^{1/2} |\sum_{k=1}^{n} \theta^k (\tilde{\theta}_k - \theta^{-1})| = \sigma \left|\frac{\theta+1}{\theta^2 Y}\right| \text{ a.s.}$$

Proof. As in the previous section, we have $\hat{\theta}_n - \theta = Q_{n-1}^{-1} M_n$ with

$$M_n = M_0 + \sum_{k=1}^{n} \Phi_{k-1} \varepsilon_k \quad \text{and} \quad M_0 = -\theta.$$

If $Z_n = A^{-n} M_n$, we clearly have

$$Z_n = A^{-1} Z_{n-1} + \varphi_{n-1} \varepsilon_n, \tag{3.6}$$

where $\varphi_n = A^{-1} Y_n$ and $Y_n = A^{-n} \Phi_n$. Set

$$T_n = \sum_{k=1}^{n} \varphi_{k-1} \varepsilon_k, \quad P_n = \sum_{k=0}^{n} \varphi_k \varphi_k^t + I_p.$$

We have already seen that Y_n converges a.s. to Y given by (2.3). Hence, it follows from Kronecker's lemma that

$$\lim_{n\to\infty} n^{-1} P_n = A^{-1} Y Y^t (A^{-1})^t \text{ a.s.}$$

Moreover, as (φ_n) is almost surely bounded, we immediately have for all $\beta > 2$

$$\sum_{n=1}^{\infty}\left(\frac{\|\varphi_n\|}{\sqrt{n}}\right)^{\beta} < \infty \text{ a.s.}$$

Consequently, we find via Lemma 3.1 that for any $v \in \boldsymbol{R}^p$

$$\limsup_{n\to\infty}\left(\frac{1}{2n\log_2 n}\right)^{1/2} |v^t T_n| = \sigma |v^t A^{-1} Y| \text{ a.s.} \tag{3.7}$$

Furthermore, (3.6) can be rewritten as $Z_{n+1} - Z_n = -BZ_n + \varphi_n \varepsilon_{n+1}$ with $B = I_p - A^{-1}$. As the matrix B is invertible, we obtain

$$\sum_{k=1}^{n} Z_k = B^{-1}(Z_1 - Z_{n+1}) + B^{-1}(T_{n+1} - T_1). \tag{3.8}$$

In addition, we can easily deduce from (3.6) together with (3.3) that

$$\|Z_n\| = O(\max_{1\leqslant k\leqslant n} |\varepsilon_k|) = o(n^{1/\beta}) \text{ a.s.} \tag{3.9}$$

with $2 < \beta < \alpha$. Hence, we infer from (3.7)–(3.9) that for any $v \in \boldsymbol{R}^p$

$$\limsup_{n\to\infty}\left(\frac{1}{2n\log_2 n}\right)^{1/2} \left|v^t \sum_{k=1}^{n} Z_k\right| = \sigma |v^t (A - I_p)^{-1} Y| \text{ a.s.} \tag{3.10}$$

Finally, recalling that $Z_n = A^{-n}(M_0 + Q_{n-1}(\hat{\theta}_n - \theta))$, (3.4) and (3.5) follow from (3.10) together with (2.5), completing the proof of Theorem 3.1. ■

4. CONCLUSION

In this paper, we have studied the LIL for (S_n) of the typical form

$$S_n = \sum_{k=1}^{n} a^k \xi_k,$$

where $|a| > 1$ and (ξ_n) are independent with $\mathcal{N}(0, \sigma^2)$ distribution. A natural question is, without averaging, is it possible to do without the normality assumption on (ξ_n)? We shall show in this section that, except for bounded random variables, it is impossible to get rid of this normality assumption. Assume that (ξ_n) are i.i.d. with mean zero and variance σ^2 with $\sigma > 0$. It is easy to see that

$$a^{-n} S_n \xrightarrow{\mathcal{L}} H \quad \text{with } H = \sum_{k=0}^{\infty} a^{-k} \xi_{k+1}. \tag{4.1}$$

In order to prove an LIL for (S_n), a classical approach is to introduce a time change as follows. Let Λ be the function defined, for all $t \in \boldsymbol{R}$ with $t \geqslant 1$, by

$\Lambda(t) = a^{2t}$. We obviously have $\Lambda^{-1}(t) = (2\log|a|)^{-1}\log(t)$. Consequently, if

$$\nu_t = \left[\frac{\log t}{2\log|a|}\right],$$

then it follows from (4.1) that

$$|a|^{-\nu_n}|S_{\nu_n}| \xrightarrow{\mathscr{L}} |H|. \tag{4.2}$$

Moreover, we have

$$E[S_n^2] = \frac{a^2\sigma^2}{a^2-1}(a^{2n}-1),$$

which implies that $E[S_{\nu_n}^2] \sim a^2\sigma^2(a^2-1)^{-1}a^{2\nu_n}$. However, we shall now show that the sequence $(n^{-1}a^{2\nu_n})$ is such that $(n^{-1}E[S_{\nu_n}^2])$ is bounded by 1 but it never converges. As a matter of fact, let Ψ be the function given, for all $t \in \boldsymbol{R}$, by $\Psi(t) = t-[t]-1/2$. One can easily check that

$$a^{2\nu_t} = \frac{t}{|a|}|a|^{-2\Psi(\Lambda^{-1}(t))}.$$

As Ψ is a periodic function with period one, the sequence $(t^{-1}a^{2\nu_t})$ never converges. Hence, an LIL similar to that of Lemma 3.1 never holds for (S_n). Furthermore, we have

$$|S_{\nu_{n+1}}-S_{\nu_n}| = \Big|\sum_{k=\nu_n+1}^{\nu_{n+1}} a^k\xi_k\Big| = O\big(\max_{\nu_n<k\leqslant\nu_{n+1}}|\xi_k|\big)|a|^{\nu_n}$$
$$= O\big(\max_{\nu_n<k\leqslant\nu_{n+1}}|\xi_k|\big)\sqrt{n} = o(\sqrt{n\log_2 n}) \text{ a.s.}$$

as soon as $|\xi_n| = o(\sqrt{\log n})$ a.s. Then we can deduce from Proposition 5 of [7] that

$$\limsup_{n\to\infty}\left(\frac{1}{2n\log_2 n}\right)^{1/2}|S_{\nu_n}| \leqslant 1 \text{ a.s.} \tag{4.3}$$

However, we have already seen that (4.3) is no longer true when replacing ν_n by n.

THEOREM 4.1. *Assume that (ξ_n) is a sequence of independent and identically distributed random variables with $|\xi_n| = O(\sqrt{\log n})$ a.s. Then we have*

$$\limsup_{n\to\infty}\frac{|S_n|}{|a|^n\sqrt{\log n}} \leqslant \left(\frac{|a|}{|a|-1}\right)\limsup_{n\to\infty}\frac{|\xi_n|}{\sqrt{\log n}} \text{ a.s.} \tag{4.4}$$

Proof. If $Z_n = a^{-n}S_n/\sqrt{\log n}$, we clearly have

$$Z_{n+1} = a^{-1}Z_n\left(\frac{\log n}{\log(n+1)}\right)+\frac{\xi_{n+1}}{\sqrt{\log(n+1)}},$$

$$|Z_{n+1}| \leqslant |a|^{-1}|Z_n|+\frac{|\xi_{n+1}|}{\sqrt{\log(n+1)}}.$$

Hence, for all $n \geqslant 2$, we find that

$$|Z_n| \leqslant |a|^{-(n-1)}|Z_1| + \sum_{k=2}^{n} |a|^{-(n-k)} \frac{|\xi_k|}{\sqrt{\log k}},$$

which implies (4.4). ■

Remark. If

$$\limsup_{n\to\infty} \frac{|\xi_n|}{\sqrt{\log n}} = \infty,$$

one can observe that

$$\limsup_{n\to\infty} \frac{|S_n|}{|a|^n \sqrt{\log n}} = \infty \text{ a.s.}$$

REFERENCES

[1] T. W. Anderson, *On the asymptotic distributions of estimates of parameters of stochastic difference equations*, Ann. Math. Statist. 30 (1959), pp. 676–687.

[2] S. Asmussen and N. Keiding, *Martingale central limit theorems and asymptotic theory for multitype branching processes*, Adv. in Appl. Probab. 10 (1978), pp. 109–129.

[3] K. B. Athreya and P. E. Ney, *Branching Processes*, Springer, Berlin 1972.

[4] B. Bercu, *On large deviations in the Gaussian autoregressive process: stable, unstable and explosive cases*, Bernoulli 7 (2001), pp. 299–316.

[5] B. M. Brown and C. C. Heyde, *An invariance principle and some convergence rates for branching processes*, Z. Wahrscheinlichkeitstheorie verw. Gebiete 20 (1971), pp. 271–278.

[6] M. Duflo, *Random Iterative Models*, Springer, Berlin 1997.

[7] M. Duflo, R. Senoussi and A. Touati, *Sur la loi des grands nombres pour les martingales vectorielles et l'estimateur des moindres carrés d'un modèle de régression*, Ann. Inst. H. Poincaré 26 (1990), pp. 549–566.

[8] M. Duflo, R. Senoussi and A. Touati, *Propriétés asymptotiques presque sûres de l'estimateur des moindres carrés d'un modèle autorégressif vectoriel*, Ann. Inst. H. Poincaré 27 (1990), pp. 1–25.

[9] P. Hartman, *Normal distributions and the law of iterated logarithm*, Amer. J. Math. 63 (1941), pp. 584–588.

[10] P. Hartman and A. Winter, *On the law of iterated logarithm*, Amer. J. Math. 63 (1941), pp. 169–176.

[11] C. C. Heyde, *Some almost sure convergence theorems for branching processes*, Z. Wahrscheinlichkeitstheorie verw. Gebiete 20 (1971), pp. 189–192.

[12] C. C. Heyde and J. R. Leslie, *Improved classical limit analogues for Galton–Watson processes with or without immigration*, Bull. Austral. Math. Soc. 5 (1971), pp. 145–155.

[13] R. M. Huggins, *Laws of iterated logarithm for time changed Brownian motion with an application to branching processes*, Ann. Probab. 13 (1985), pp. 1148–1156.

[14] T. L. Lai and C. Z. Wei, *Asymptotic properties of general autoregressive models and strong consistency of least-squares estimates of their parameters*, J. Multivariate Anal. 13 (1983), pp. 1–23.

[15] J. Marcinkiewicz and A. Zygmund, *Remarque sur la loi du logarithme itéré*, Fund. Math. 29 (1937), pp. 215–222.
[16] V. V. Petrov, *Limit Theorems of Probability Theory*, Oxford University Press, Oxford 1995.
[17] M. M. Rao, *Consistency and limit distribution of estimators of parameters in explosive stochastic difference equations*, Ann. Math. Statist. 32 (1961), pp. 195–218.
[18] W. F. Stout, *Almost sure convergence*, Academic Press, New York 1974.
[19] R. J. Tomkins, *Some iterated logarithm results related to the central limit theorem*, Trans. Amer. Math. Soc. 156 (1971), pp. 185–192.
[20] A. Touati, *Two theorems on convergence in distribution for stochastic integrals and statistical applications*, Probab. Theory Appl. 38 (1993), pp. 95–117.
[21] A. Touati, *Vitesse de convergence en loi de l'estimateur des moindres carrés d'un modèle autoregressif (cas mixte)*, Ann. Inst. H. Poincaré 32 (1996), pp. 211–230.
[22] J. S. White, *The limit distribution of the serial correlation in the explosive case*, Ann. Math. Statist. 29 (1958), pp. 1188–1197.
[23] R. Wittmann, *A general law of iterated logarithm*, Z. Wahrscheinlichkeitstheorie verw. Gebiete 68 (1985), pp. 521–543.

Bernard Bercu
Laboratoire de Mathématiques
UMR 8628 CNRS
Bâtiment 425, Université Paris-Sud
91405 Orsay Cedex, France
E-mail: Bernard.Bercu@math.u-psud.fr

Abderrahmen Touati
Laboratoire de Mathématiques
UMR 7534 CNRS
Place du Maréchal Delattre de Tassigny
75775 Paris Cedex 16, France
E-mail: abdertouati@yahoo.com

Received on 12.10.2001