A Law of the Iterated Logarithm for Stochastic Approximation Procedures in d–dimensional Euclidean Space

VALERY KOVAL*
Department of Mathematics
Zhytomyr Institute
of Engineering and Technology
Cherniakhovsky str. 103
10005 Zhytomyr, Ukraine

RAINER SCHWABE
Department of Medical Biometry
University of Tübingen
Westbahnhofstr. 55
D-72 070 Tübingen, Germany

Running head: LIL for d–dimensional stochastic approximation

Address for correspondence:

Rainer Schwabe
Department of Medical Biometry
University of Tübingen
Westbahnhofstr. 55
72 070 Tübingen, Germany
e–mail: `Rainer.Schwabe@uni-tuebingen.de`
Tel: ++49 / (0)7071 / 29 – 78 254
FAX: ++49 / (0)7071 / 29 – 50 75

* Work partly supported by a research grant of the *German Academic Exchange Service* (DAAD).

Abstract

In this paper we investigate the rate of convergence for general d–dimensional stochastic approximation procedures and present an explicit expression for the asymptotic bounds in the law of the iterated logarithm.

Key words: stochastic approximation, law of the iterated logarithm, rate of convergence

1 Introduction

Let $\mathbf{R}^d$ be the d–dimensional Euclidean space with corresponding norm $||x|| = \sqrt{x^\top x}$, where "$^\top$" denotes transposition. We consider the nonlinear stochastic difference equation

$$X_n = X_{n-1} - a_n \left(F(X_{n-1}) + \Psi_n \right) + b_n V_n, \quad n \geq 1, \ X_0 \in \mathbf{R}^d, \tag{1}$$

where $(a_n, \ n \geq 1)$ is a sequence of real numbers satisfying

$$a_n > 0, \ \lim_{n\to\infty} a_n = 0 \ \text{ and } \ \sum_{n=1}^{\infty} a_n = \infty. \tag{2}$$

Moreover, $(b_n, \ n \geq 1)$ is a sequence of real numbers, $(V_n, \ n \geq 1)$ and $(\Psi_n, \ n \geq 1)$ are sequences of d–dimensional random vectors such that $||\Psi_n|| \to 0$ almost surely (a. s.) as $n \to \infty$. Finally, we assume that the equation $F(x) = 0$ has a solution θ in $\mathbf{R}^d$.

The recursive scheme (1) includes, in particular, the well-known Robbins–Monro and Kiefer–Wolfowitz procedures (see e. g. Nevelson and Khasminskii, 1976; Kushner and Yin, 1997).

In this paper we investigate the rate of convergence of X_n towards θ as $n \to \infty$. For $d > 1$ and some special choices of the sequences $(a_n, \ n \geq 1)$ and $(b_n, \ n \geq 1)$ bounds were obtained for the rate of convergence by Korostelev (1983, 1984), Walk (1988) and Pelletier (1998). Koval (1998) proved that under rather general assumptions on (1) there exists a constant $L \in (0, \infty)$ such that

$$\limsup_{n\to\infty} \left(\frac{a_n}{b_n^2 \ell_n} \right)^{1/2} ||X_n - \theta|| = L \quad \text{a. s.}, \tag{3}$$

where $\ell_n = \log \left(\sum_{j=1}^n a_j \right)$. Here $\log t = \ln \max\{t, e\}$ for $t \geq 0$ where ln denotes the natural logarithm. Lower and upper bounds for L were given by Koval (1998), too. In the present paper we establish the exact value of L. Note that in the case $d = 1$ this problem was solved by Koval and Schwabe (1998).

Under very general assumptions on (1) the investigation of the relation (3) reduces to the analogous relation for the solution $(Y_n, \ n \geq 1)$ of the corresponding linear stochastic difference equation

$$Y_n = (I - a_n A) Y_{n-1} + b_n V_n, \quad n \geq 1, \ Y_0 = 0 \tag{4}$$

(see Koval, 1998, Lemma 2), where I is the identity matrix and A is a $d \times d$ matrix such that the real parts $re\ \lambda_i(A)$ of its eigenvalues $\lambda_1(A), ..., \lambda_d(A)$ satisfy the condition

$$\lambda^* = \min_{1 \le i \le d} \mathrm{re}\, \lambda_i(A) > 0. \tag{5}$$

The remainder of the paper is organized as follows: A law of the iterated logarithm (LIL) is given for the linear recursive equation (4) in Section 2 (Theorem 1). The proof of Theorem 1 is based on a LIL for the matrix-normalized sums of independent random vectors which is proved in Section 3. Theorem 1 is then proved in Section 4. Finally, the general nonlinear recursive scheme (1) is investigated in Section 5.

In the sequel we will use the following notations: For any matrix H the supremum norm $||\cdot||_\infty$ of H is defined by $||H||_\infty = \sqrt{\lambda_{\max}(H^\top H)}$, where $\lambda_{\max}(\cdot)$ denotes the largest eigenvalue of the corresponding symmetric matrix. In particular, if H is a positive semi–definite symmetric matrix, then this norm reduces to $||H||_\infty = \lambda_{\max}(H)$. We write $H > 0$ ($H \ge 0$) if the symmetric matrix H is positive definite (positive semi–definite). Moreover, we write $H_2 \ge H_1$ if $H_2 - H_1 \ge 0$. By c we denote a positive constant which may vary from line to line.

2 A LIL for the linear equation (4)

In this section we consider the linear recursive equation (4). We assume that the sequences $(a_n,\ n \ge 1)$ and $(b_n,\ n \ge 1)$ satisfy the following conditions:

$$b_n^2/a_n \ \ge\ b_{n+1}^2/a_{n+1}, \qquad n \ge 1 \tag{6}$$

$$\lim_{n\to\infty} \frac{1}{a_n}\left(1 - \frac{b_n^2/a_n}{b_{n-1}^2/a_{n-1}}\right) \ =\ \tau \ <\ 2\lambda^* \tag{7}$$

$$\limsup_{n\to\infty} \frac{|b_n - b_{n+1}|}{a_{n+1}|b_n|} \ <\ \infty. \tag{8}$$

Theorem 1. *Assume that conditions (2) and (5) to (8) are satisfied. If there exists a sequence of independent, identically distributed Gaussian random vectors* $(\Gamma_n,\ n \ge 1)$ *with zero expectation,* $\mathbf{E}(\Gamma_n) = 0$*, and covariance* *matrix* $\mathbf{E}(\Gamma_n \Gamma_n^\top) = D > 0,\ n \geq 1,$ *such that*

$$\left\| \sum_{i=1}^{n} V_i - \sum_{i=1}^{n} \Gamma_i \right\| = o\left((\ell_n / a_n)^{1/2} \right) \quad a.\,s.\ (n \to \infty), \tag{9}$$

then

$$\limsup_{n\to\infty} \left(\frac{a_n}{b_n^2 \ell_n} \right)^{1/2} \|Y_n\| = \left(2 \left\| \int_0^\infty e^{\left(\frac{\tau}{2} I - A\right)t} D e^{\left(\frac{\tau}{2} I - A^\top\right)t} dt \right\|_\infty \right)^{1/2} \quad a.\,s. \tag{10}$$

Consider the assumptions of Theorem 1. The conditions (6) to (8) are discussed in Koval and Schwabe (1998). The regularity condition $D > 0$ is commonly assumed (see Korostelev, 1984, p.161; Pelletier, 1998). Condition (9) means that the sequence $(V_n,\ n \geq 1)$ satisfies an almost sure invariance principle with an approximation error of order $o\left((\ell_n/a_n)^{1/2}\right)$. This allows for a reduction of the problem to prove (10) for the Gaussian case only (cf. Lemma 4). On the other hand a. s. invariance principles are known to hold for broad classes of random sequences of various types, such as martingale difference sequences, mixing sequences etc. (see Philipp, 1986, for a survey on this topic). Therefore the assumption (9) makes it possible to consider various classes of sequences $(V_n,\ n \geq 1)$ simultaneously. It is worth noticing that the proofs of a couple of strong limit theorems make use of assumptions of the type (9) (see, for example, Walk, 1988; Lacey and Philipp, 1990; Horváth, 1993).

In view of condition (9) stronger assumptions have to be imposed on the accuracy Δ_n in the approximation of the partial sum $\sum_{i=1}^n V_i$ by its Gaussian counterpart $\sum_{i=1}^n \Gamma_i$, if the steplengths $(a_n,\ n \geq 1)$ decrease at a slower rate, as $n \to \infty$, e. g.

a) $\Delta_n = o\left((n \log\log n)^{1/2}\right)$, if $a_n = 1/n$,
b) $\Delta_n = o\left((n^\alpha \log n)^{1/2}\right)$, if $a_n = 1/n^a$, $0 < a < 1$,
c) $\Delta_n = o\left((\log n)^{(1+\alpha)/2}\right)$, if $a_n = 1/(\log n)^\alpha$, $\alpha > 1$.

In the above cases a), b), and c), besides others, results were given e. g. by Einmahl (1987, 1989), Berger (1990), Berkes and Philipp (1979), Kuelbs and Philipp (1980), Morrow and Philipp (1982), and Eberlein (1986) for the approximation of $\sum_{i=1}^n V_i$ by $\sum_{i=1}^n \Gamma_i$ in $\mathbf{R}^d$ for various classes of sequences $(V_n,\ n \geq 1)$.

3 An auxiliary result

Let $(Z_n,\ n \geq 1)$ be a sequence of independent Gaussian random vectors in $\mathbf{R}^d$ with zero expectation, $\mathbf{E}(Z_n) = 0,\ n \geq 1$, and let $(A_n,\ n \geq 1)$ be an arbitrary sequence of $p \times d$ matrices ($A_n \neq 0, n \geq 1$). Denote by $S_n = \sum_{i=1}^n Z_i$ the partial sums and by $B_n = \mathbf{E}(S_n S_n^\top) = \sum_{i=1}^n \mathbf{E}(Z_i Z_i^\top)$ the corresponding covariance matrices, $n \geq 1$. Further, denote by $(e_n,\ n \geq 1)$ a sequence of normalized eigenvectors of the matrix $A_n B_n A_n^\top$ associated with the largest eigenvalue $q_n = \|A_n B_n A_n^\top\|_\infty$, i. e. $A_n B_n A_n^\top e_n = q_n e_n$ and $\|e_n\| = 1$, $n \geq 1$.

For the partial sums S_n of independent random vectors we establish next the LIL under matrix normalization which is of particular interest on its own (for more detailed information see Buldygin and Solntsev, 1997).

Proposition 1. *Let $B_{n_0} > 0$ for some $n_0 \geq 1$ and*

$$\|B_n\|_\infty \to \infty \qquad (n \to \infty). \tag{11}$$

Assume that there exists a sequence $(f_n,\ n \geq 1)$ of positive numbers such that

$$\limsup_{n\to\infty} \frac{f_{n+1}}{f_n} < \infty, \tag{12}$$

$$\log\log f_n \sim \log\log \|B_n\|_\infty \quad (n \to \infty), \tag{13}$$

and

$$\frac{e_m^\top A_m B_m A_n^\top e_n}{(q_m q_n)^{1/2}} \leq c \frac{f_m}{f_n} \tag{14}$$

for all $n > m \geq n_0$. Then

$$\limsup_{n\to\infty} \frac{\|A_n S_n\|}{(2 q_n \log\log \|B_n\|_\infty)^{1/2}} = 1 \quad a.\,s. \tag{15}$$

For the proof of Proposition 1 we need three additional lemmas. Moreover, it is helpful to note that the matrices B_n are regular for $n \geq n_0$ in view of the inequality $B_n \geq B_{n_0} > 0$. Hence, the inverse matrices B_n^{-1} exist for all $n \geq n_0$. Denote the positive (semi-)definite symmetric square root of B_n by $B_n^{1/2}$, i. e. $B_n^{1/2} B_n^{1/2} = B_n$, and its inverse by $B_n^{-1/2} = (B_n^{1/2})^{-1}$, $n \geq n_0$. Moreover, we write $\det(\cdot)$ for the determinant of a matrix.

Lemma 1. *For all $n \geq m \geq n_0$ the following inequality holds*

$$\|B_m^{-1/2} B_n^{1/2}\|_\infty^2 \leq \frac{\det(B_n)}{\det(B_m)}.$$

Proof. From $B_m^{-1/2} B_n B_m^{-1/2} = B_m^{-1/2}(B_n - B_m)B_m^{-1/2} + I$ and $B_m \leq B_n$ it follows that all eigenvalues $\lambda_1, ..., \lambda_d$ of the matrix $B_m^{-1/2} B_n B_m^{-1/2}$ exceed one, $\lambda_j \geq 1,\ j = 1, ..., d$. Hence, $\|B_m^{-1/2} B_n^{1/2}\|_\infty^2 = \lambda_{\max}(B_m^{-1/2} B_n B_m^{-1/2}) \leq \prod_{j=1}^d \lambda_j = \det(B_m^{-1/2} B_n B_m^{-1/2}) = \det(B_n)/\det(B_m)$. $\square$

In the following denote by Z a standard Gaussian random vector in $\mathbf{R}^d$, $\mathbf{E}(Z) = 0$, $\mathbf{E}(ZZ^\top) = I$.

Lemma 2. *For all $t > (2d)^{1/2}$ the following inequality holds*

$$\mathbf{P}(\|Z\| > t) \leq \exp\left\{-\tfrac{1}{2}\left(t - (2d)^{1/2}\right)^2\right\}.$$

Lemma 2 is a straightforward consequence of Lemma 3.1 in Ledoux and Talagrand (1991).

Lemma 3. *Let $(\xi(n),\ n \geq 1)$ be a real–valued Gaussian sequence with zero expectation, $\mathbf{E}(\xi(n)) = 0$, non–vanishing variances $\sigma^2(n) = \mathbf{E}(\xi^2(n)) > 0$ and correlations $r(m,n) = \mathbf{E}(\xi(m)\xi(n)) / (\sigma(m)\sigma(n))$. Assume that there exists a sequence $(f(n),\ n \geq 1)$ of positive numbers sucht that*

$$f(n) \to \infty \quad (n \to \infty), \tag{16}$$

$$\limsup_{n\to\infty} f(n+1)/f(n) < \infty, \tag{17}$$

and for all $n > m \geq 1$

$$|r(m,n)| \leq cf(m)/f(n). \tag{18}$$

Then

$$\limsup_{n\to\infty} \frac{|\xi(n)|}{(2\sigma^2(n)\log\log f(n))^{1/2}} \geq 1 \quad a.\ s.$$

Proof. Define a subsequence $(n_i,\ i \geq 1)$ of indices recursively by

$$n_1 = 1, \quad n_{i+1} = \min\{j:\ j > n_i,\ f(n_i)/f(j) < q\}, \quad i \geq 1 \tag{19}$$

for some fixed $q \in (0,1)$. Note that this subsequence is infinite by (16). To establish the present lemma it is sufficient to prove the assertion for the subsequence $(n_i,\ i \geq 1)$, i. e.

$$\limsup_{i\to\infty} \frac{|\xi(n_i)|}{(2\sigma^2(n_i)\log\log f(n_i))^{1/2}} \geq 1 \quad \text{a.s.} \tag{20}$$

Let $\eta(i) = \xi(n_i)/\sigma(n_i)$ and note that $(\eta(i),\ i \geq 1)$ is a normalized Gaussian sequence, $\mathbf{E}(\eta(i)) = 0$, $\mathbf{E}(\eta^2(i)) = 1$, with correlation $r_\eta(i,j) = \mathbf{E}(\eta(i)\eta(j)) = r(n_i, n_j)$. By (18) and (19) we have $|r_\eta(i,j)| < cq^{j-i}$ for all $j > i \geq 1$. This implies that $\lim_{k\to\infty}\sup_{|j-i|\geq k}|r_\eta(i,j)| = 0$ and, hence, by Lai (1973, Theorem 2)

$$\limsup_{i\to\infty} \frac{|\eta(i)|}{(2\log i)^{1/2}} = 1 \quad \text{a.s.} \tag{21}$$

It follows from (19) and (17), respectively, that $f(n_i - 1) \leq q^{-1} f(n_{i-1})$, $i \geq 2$, and $f(n) \leq c_1 f(n-1)$, $n > n_0$, for some $c_1 > 1$. This implies $f(n_i) \leq c_2^i f(n_0)$, $i \geq 1$, where $c_2 = c_1 q^{-1} > 1$. Consequently,

$$\limsup_{i\to\infty} \frac{|\eta(i)|}{(2\log\log f(n_i))^{1/2}} \geq 1 \quad \text{a.s.,}$$

in view of (21) which proves the assertion (20) for the subsequence $(n_i,\ i \geq 1)$.

□

Proof of Proposition 1. Without loss of generality we may asssume $n_0 = 1$. Set $t_n = (2\log\log \||B_n\|_\infty)^{1/2}$, $n \geq 1$. In a first step we prove

$$\limsup_{n\to\infty} \frac{\|A_n S_n\|}{q_n^{1/2} t_n} \leq 1 \quad \text{a.s.} \tag{22}$$

By the inequality $\|A_n S_n\| \leq \|A_n B_n^{1/2}\|_\infty \cdot \|B_n^{-1/2} S_n\|$ and the identity $\|A_n B_n^{1/2}\|_\infty = \|A_n B_n A_n^\top\|_\infty^{1/2} = q_n^{1/2}$ it will be sufficient to show

$$\limsup_{n\to\infty} t_n^{-1} \|B_n^{-1/2} S_n\| \leq 1 \quad \text{a.s.} \tag{23}$$

in order to establish (22).

The sequence $(\det(B_n),\ n \geq 1)$ of determinants is monotonically increasing with $\det(B_n) \to \infty$ (as $n \to \infty$) by (11). Therefore, for every $r > 1$, there exists a strictly increasing sequence of integers $(n_j,\ j \geq 1)$ satisfying

$$r\det(B_{n_j}) \leq \det(B_{n_{j+1}}) \leq r^3 \det(B_{n_j+1}) \tag{24}$$

by Lemma 3.3 in Wittmann (1985). Hence, in order to establish (23), it will be sufficient to show that for every $\varepsilon > 0$ there exists a number $r > 1$ such that

$$\sum_{j=1}^{\infty} \mathbf{P}\left(\max_{n_j < n \leq n_{j+1}} t_n^{-1} \|B_n^{-1/2} S_n\| > 1 + 2\varepsilon \right) < \infty \tag{25}$$

in view of the Borel–Cantelli Lemma.

Let, now, $\varepsilon > 0$ be fixed and set $r = ((1+2\varepsilon)/(1+\varepsilon))^{2/3} > 1$. Note that the inequality $B_m \leq B_n$, for $m \leq n$, implies that

$$\|B_n^{-1/2} x\| \leq \|B_m^{-1/2} x\| \tag{26}$$

for all $x \in \mathbf{R}^d$. Further, note that there exists an integer j_0 such that $t_{n_j} > (2d)^{1/2}$ for all $j \geq j_0$. By (26), the Lévy inequality, Lemma 1, the right hand side of (24), and Lemma 2 we have for $j \geq j_0$

$$\begin{aligned}
&\mathbf{P}\left(\max_{n_j < n \leq n_{j+1}} t_n^{-1} \|B_n^{-1/2} S_n\| > 1 + 2\varepsilon \right) \\
\leq\ & \mathbf{P}\left(\max_{n_j < n \leq n_{j+1}} \|B_{n_j+1}^{-1/2} S_n\| > (1+2\varepsilon) t_{n_j} \right) \\
\leq\ & 2\mathbf{P}\left(\|B_{n_j+1}^{-1/2} B_{n_{j+1}}^{1/2}\|_\infty \cdot \|B_{n_{j+1}}^{-1/2} S_{n_{j+1}}\| > (1+2\varepsilon) t_{n_j} \right) \\
\leq\ & 2\mathbf{P}\left(r^{3/2} \|Z\| > (1+2\varepsilon) t_{n_j} \right) \\
=\ & 2\mathbf{P}\left(\|Z\| > (1+\varepsilon) t_{n_j} \right) \\
\leq\ & 2\left(\log \|B_{n_j}\|_\infty \right)^{-(1+\varepsilon)^2 (1-\delta_j)^2},
\end{aligned} \tag{27}$$

where $\delta_j = (2d)^{1/2} \left((1+\varepsilon) t_{n_j} \right)^{-1} \to 0$ as $j \to \infty$.

By the left hand side of (24) we have $\|B_{n_j}\|_\infty \geq r^{(j-1)/d} \left(\det(B_{n_1}) \right)^{1/d}$, $j \geq 1$, which together with (27) implies the assertion (25).

Finally, in order to establish (15) we have to prove

$$\limsup_{n \to \infty} \frac{\|A_n S_n\|}{q_n^{1/2} t_n} \geq 1 \quad \text{a.s.} \tag{28}$$

in view of (22). For this set $\zeta_n = e_n^\top A_n S_n$, $n \geq 1$. Note that $(\zeta_n,\ n \geq 1)$ is a Gaussian sequence with zero expectation, $\mathbf{E}(\zeta_n) = 0$, variances $\mathbf{E}(\zeta_n^2) = q_n$ and corresponding coefficients of correlation

$$r_\zeta(m,n) = \frac{e_m^\top A_m B_m A_n^\top e_n}{(q_m q_n)^{1/2}}, \quad n > m \geq 1.$$

By (11), (12) and (14) it follows from Lemma 3 that

$$\limsup_{n\to\infty} \frac{\left|e_n^\top A_n S_n\right|}{(q_n 2\log\log f_n)^{1/2}} \geq 1 \quad \text{a.s.} \tag{29}$$

By the Cauchy-Schwarz inequality we have $\|A_n S_n\| \geq |e_n^\top A_n S_n|$ and the assertion (28) follows from (13) and (29) which completes the proof. □

4 Proof of Theorem 1

In view of conditions (6) and $\lim_{n\to\infty} a_n = 0$ we may assume without loss of generality that $a_n < \min\{1,\ 1/(2\|A\|_\infty)\}$ and $b_n \neq 0$ for all $n \geq 1$. We adopt the convention $\prod_{j=n+1}^{n}(\cdot) = 1$ or I, respectively, for matrices.

In order to simplify the presentation of the proof we start with two lemmas.

First, consider the equation (4) when the sequence $(V_n,\ n \geq 1)$ is replaced by $(\Gamma_n,\ n \geq 1)$ defined in Theorem 1, i. e.

$$Y_n^* = (I - a_n A)Y_{n-1}^* + b_n \Gamma_n, \quad n \geq 1,\ Y_0^* = 0. \tag{30}$$

Lemma 4. *If the conditions (2), (5) and (7) to (9) are satisfied, then*

$$\|Y_n - Y_n^*\| = o\left((b_n^2 \ell_n / a_n)^{1/2}\right) \quad \textit{a. s. } (n \to \infty). \tag{31}$$

Proof. Denote by S_n and S_n^* the partial sums of the V_i and Γ_i, respectively, $S_n = \sum_{i=1}^n V_i$ and $S_n^* = \sum_{i=1}^n \Gamma_i$. We use Abel's identity to obtain

$$\begin{aligned}
Y_n - Y_n^* &= \sum_{k=1}^{n} b_k \prod_{j=k+1}^{n} (I - a_j A)(V_k - \Gamma_k) \\
&= \sum_{k=1}^{n-1} \left[b_k \prod_{j=k+1}^{n} (I - a_j A) - b_{k+1} \prod_{j=k+2}^{n} (I - a_j A) \right] (S_k - S_k^*) + b_n (S_n - S_n^*) \\
&= \sum_{k=1}^{n-1} \prod_{j=k+2}^{n} (I - a_j A) b_k a_{k+1} \left(\frac{b_k - b_{k+1}}{a_{k+1} b_k} I - A \right) (\ell_k / a_k)^{1/2} \left[(S_k - S_k^*)(a_k/\ell_k)^{1/2} \right] \\
&\quad + b_n (S_n - S_n^*).
\end{aligned}$$

Let $\delta_n = (a_n/\ell_n)^{1/2}\|S_n - S_n^*\|$, $n \geq 1$, and note that

$$\delta_n \to 0 \quad \text{a.s. } (n \to \infty) \tag{32}$$

in view of (9). Next we make use of condition (8) and obtain

$$\begin{aligned} \|Y_n - Y_n^*\| \leq & \sum_{k=1}^{n-1} \left\| \prod_{j=k+2}^{n} (I - a_j A) \right\|_\infty |b_k| a_{k+1} \\ & \cdot \left(\left| \frac{b_k - b_{k+1}}{a_{k+1} b_k} \right| + \|A\|_\infty \right) (\ell_k/a_k)^{1/2} \delta_k \\ & + \left(b_n^2 \ell_n / a_n \right)^{1/2} \delta_n. \end{aligned} \tag{33}$$

It follows from (7) (see Koval and Schwabe, 1998) that for some ν, $0 < \nu < 2$, and for all sufficiently large n

$$\frac{b_n^2}{a_n} \geq \frac{b_{n-1}^2}{a_{n-1}} (1 - \lambda^* a_n)^\nu. \tag{34}$$

Now we set

$$\sigma = (\lambda^* - \delta) - (\lambda^* + \varepsilon)\nu/2, \tag{35}$$

where $\delta \in (0, \lambda^*)$ and $\varepsilon > 0$ are chosen sufficiently small such that $\sigma > 0$. By Walk and Zsidó (1989, Lemma 3) it follows from (2) and (5) that

$$\left\| \prod_{j=k}^{n} (I - a_j A) \right\|_\infty \leq c \prod_{j=k}^{n} (1 - a_j)^{\lambda^* - \delta} \tag{36}$$

for all $n \geq k \geq 1$. It follows from (35) and (2) that

$$\begin{aligned} \prod_{j=k}^{n} (1 - a_j)^{\lambda^* - \delta} & \leq \prod_{j=k}^{n} (1 - a_j)^\sigma \exp \left\{ -\frac{\nu}{2} (\lambda^* + \varepsilon) \sum_{j=k}^{n} a_j \right\} \\ & \leq c \prod_{j=k}^{n} (1 - a_j)^\sigma \prod_{j=k}^{n} (1 - \lambda^* a_j)^{\nu/2}, \quad 1 \leq k \leq n. \end{aligned} \tag{37}$$

By (36) and (37) it follows from (33) that

$$\begin{aligned} & \|Y_n - Y_n^*\| \\ & \leq c\, \ell_n^{1/2} \sum_{k=1}^{n-1} \left(|b_k| a_k^{-1/2} \prod_{j=k+2}^{n} (1 - \lambda^* a_j)^{\nu/2} \right) a_{k+1} \prod_{j=k+2}^{n} (1 - a_j)^\sigma \delta_k \\ & \quad + \left(b_n^2 \ell_n / a_n \right)^{1/2} \delta_n. \end{aligned}$$

Next, we make use of the relation $|b_k|a_k^{-1/2}\prod_{j=k+2}^n(1-\lambda^* a_j)^{\nu/2} \le c|b_n|a_n^{-1/2}$, $1 \le k \le n-1$ which follows form (34) and obtain

$$\|Y_n - Y_n^*\| \le c\left(b_n^2\ell_n/a_n\right)^{1/2}(t_n+\delta_n), \tag{38}$$

where $t_n = \sum_{k=1}^{n-1} a_{k+1}\prod_{j=k+2}^n(1-a_j)^\sigma\delta_k$. Set $c_{nk} = a_{k+1}\prod_{j=k+2}^n(1-a_j)^\sigma$, $1 \le k \le n-1$. Then

$$c_{nk} \to 0 \quad (n\to\infty), \quad \text{for every } k \ge 1, \tag{39}$$

and

$$\begin{aligned}\sum_{k=1}^{n-1} c_{nk} &= \sum_{k=1}^{n-1}\frac{a_{k+1}}{1-(1-a_{k+1})^\sigma}\left(\prod_{j=k+2}^n(1-a_j)^\sigma - \prod_{j=k+1}^n(1-a_j)^\sigma\right)\\ &\to 1/\sigma \quad (n\to\infty)\end{aligned} \tag{40}$$

by the Toeplitz Lemma. Thus, again by the Toeplitz Lemma, it follows from (32), (39) and (40) that $t_n = \sum_{k=1}^{n-1} c_{nk}\delta_k \to 0$ a. s. as $n\to\infty$. This, togehter with (32) and (38), implies relation (31) which completes the proof. $\square$

Lemma 5. *If the conditions (2) and (5) to (7) are satisfied, then*

$$\limsup_{n\to\infty}\left(\frac{a_n}{b_n^2\ell_n}\right)^{1/2}\|Y_n^*\| = \sqrt{2}M \quad a.\ s. \tag{41}$$

for the recursive equation (30), where

$$M = \left\|\int_0^\infty e^{(\frac{\tau}{2}I-A)t}De^{(\frac{\tau}{2}I-A^\top)t}dt\right\|_\infty^{1/2}.$$

Proof. Set $A_n = \prod_{j=1}^n(I-a_jA)$, $n\ge 1$. With this notation we may derive from (30) that $Y_n^* = A_n\sum_{k=1}^n b_kA_k^{-1}\Gamma_k$. Now we set $S_n = \sum_{k=1}^n b_kA_k^{-1}\Gamma_k$ and denote by $B_n = \mathbf{E}(S_nS_n^\top) = \sum_{k=1}^n b_k^2A_k^{-1}D(A_k^{-1})^\top$ and $Q_n = \mathbf{E}(Y_n^*(Y_n^*)^\top)$ the covariance matrices of S_n and Y_n^*, respectively. Note that $Y_n^* = A_nS_n$ and $Q_n = A_nB_nA_n^\top$. Moreover, let $q_n = \|Q_n\|_\infty$.

In order to establish relation (41) we will make use of Proposition 1. For this purpose we first prove the following asymptotic relations

$$\log\log\|B_n\|_\infty \sim \ell_n \quad (n\to\infty) \tag{42}$$

$$q_n \sim M^2b_n^2/a_n \quad (n\to\infty). \tag{43}$$

Cosider first (42). Let $D^{1/2}$ denote the symmetric square root of D. With this notation we may obtain a lower bound

$$\begin{aligned}
\|B_n\|_\infty &\geq d^{-1}\,\mathrm{tr}\,B_n \\
&= d^{-1}\sum_{k=1}^n b_k^2\,\mathrm{tr}\,\left[A_k^{-1}D^{1/2}\left(A_k^{-1}D^{1/2}\right)^\top\right] \\
&\geq d^{-1}\sum_{k=1}^n b_k^2\|A_k^{-1}D^{1/2}\|_\infty^2 \\
&\geq \frac{\|D\|_\infty}{d}\sum_{k=1}^n b_k^2/\|A_k\|_\infty^2.
\end{aligned}$$

Hence, by (36) and (6) we get

$$\begin{aligned}
\|B_n\|_\infty &\geq c\sum_{k=1}^n b_k^2\prod_{j=1}^k(1-a_j)^{2(\delta-\lambda^*)} \\
&= c\sum_{k=1}^n \frac{b_k^2}{a_k}\frac{a_k}{1-(1-a_k)^{2(\lambda^*-\delta)}}\left[\prod_{j=1}^k(1-a_j)^{2(\delta-\lambda^*)}-\prod_{j=1}^{k-1}(1-a_j)^{2(\delta-\lambda^*)}\right] \\
&\geq c\frac{b_n^2}{a_n}\prod_{j=1}^n(1-a_j)^{2(\delta-\lambda^*)}.
\end{aligned}$$

It follows from (34) that $\frac{b_n^2}{a_n}\prod_{j=1}^n(1-\lambda^*a_j)^{-\nu}\geq cb_1^2/a_1$, $n\geq 1$. In view of (37) this finally implies

$$\|B_n\|_\infty \geq c\exp\left(2\sigma\sum_{j=1}^n a_j\right),\quad n\geq 1. \tag{44}$$

Now we calculate an upper bound for $\|B_n\|_\infty$. It follows from (36) and (37) that

$$\begin{aligned}
q_n &\leq \|D\|_\infty\sum_{k=1}^n b_k^2\left\|\prod_{j=k+1}^n(I-a_jA)\right\|_\infty^2 \\
&\leq c\|D\|_\infty\sum_{k=1}^n\frac{b_k^2}{a_k}\prod_{j=k+1}^n(1-\lambda^*a_j)^\nu a_k\prod_{j=k+1}^n(1-a_j)^{2\sigma}.
\end{aligned}$$

We make use of the inequality

$$\frac{b_k^2}{a_k}\prod_{j=k+1}^n(1-\lambda^*a_j)^\nu\leq c\frac{b_n^2}{a_n},\quad 1\leq k\leq n, \tag{45}$$

which follows from (34) and obtain

$$q_n \leq c\left(b_n^2/a_n\right)\sum_{k=1}^{n} a_k \prod_{j=k+1}^{n}(1-a_j)^{2\sigma} \leq c\left(b_n^2/a_n\right), \quad n \geq 1 \tag{46}$$

(see also (40). By (46) and (6) we have

$$\begin{aligned}
\|B_n\|_\infty &\leq \|Q_n\|_\infty \cdot \|A_n^{-1}\|_\infty^2 \\
&\leq c\left(b_n^2/a_n\right)\left\|\prod_{j=1}^{n}(I-a_jA)^{-1}\right\|_\infty \\
&\leq c\left(b_1^2/a_1\right)\prod_{j=1}^{n}\left\|(1-a_jA)^{-1}\right\|_\infty \\
&\leq c\prod_{j=1}^{n}\frac{1}{1-a_j\|A\|_\infty} \\
&\leq c\exp\left\{(\|A\|_\infty+\widetilde{\varepsilon})\sum_{j=1}^{n}a_j\right\} \qquad (47)
\end{aligned}$$

for some $\widetilde{\varepsilon} > 0$ and all sufficiently large n. The asymptotic relation (42) then follows from (44) and (47).

Next, we consider the second asymptotic relation (43). Set $G_n = (a_n/b_n^2)\, Q_n$, $n \geq 1$. Obviously the sequence $(G_n,\ n \geq 1)$ satisfies the recursive relation

$$G_n = \alpha_n(I-a_nA)G_{n-1}(I-a_nA^\top)+a_nD, \quad n \geq 1,$$

where $\alpha_n = (a_n/b_n^2)\left(b_{n-1}^2/a_{n-1}\right)$, $n \geq 2$, $\alpha_1 = a_1/b_1^2$, $G_0 = 0$. We can rewrite this recursion as

$$G_n = \left[I-a_n\left(A-\frac{\tau}{2}I\right)\right]G_{n-1}\left[I-a_n\left(A^\top-\frac{\tau}{2}I\right)\right]+a_nD+a_nU_n, \tag{48}$$

where

$$\begin{aligned}
U_n &= \left(\frac{\alpha_{n-1}}{a_n}-\tau\right)G_{n-1}-\frac{\tau^2}{4}a_nG_{n-1}+\frac{\tau}{2}a_n\left(AG_{n-1}+G_{n-1}A^\top\right) \\
&\quad+(\alpha_n-1)\left[a_nAG_{n-1}A^\top-AG_{n-1}-G_{n-1}A^\top\right].
\end{aligned}$$

It follows from (7) that $\alpha_n \to 1$ and $(\alpha_n-1)/a_n \to \tau$ as n tends to infinity. Observe also that $\sup_{n\geq 1}\|G_n\|_\infty < \infty$ in view of (46). This and $\lim_{n\to\infty} a_n = 0$ imply

$$\|U_n\|_\infty \to 0 \quad (n \to \infty). \tag{49}$$

Set $\overline{A} = A - \frac{\tau}{2}I$ and note that by (7) the eigenvalues $\lambda_1(\overline{A}), ..., \lambda_d(\overline{A})$ of the matrix $\overline{A}$ satisfy the condition

$$\overline{\lambda} = \min_{1\leq i\leq d} \operatorname{re}\lambda_i(\overline{A}) = \lambda^* - \frac{\tau}{2} > 0 \tag{50}$$

Consider the slightly simpler recursive equation

$$G_n^* = (I - a_n\overline{A})G_{n-1}^*(I - a_n\overline{A}^\top) + a_n D, \quad n \geq 1, \tag{51}$$

where $G_0^* = 0$. From (48) and (51) we can derive that

$$\|G_n - G_n^*\|_\infty \leq \sum_{k=1}^n a_k \left\| \prod_{j=k+1}^n (I - a_j\overline{A}) \right\|_\infty^2 \cdot \|U_k\|_\infty.$$

By Walk and Zsidó (1989, Lemma 3) it follows from (2) and (50) that $\left\|\prod_{j=k}^n (I - a_j\overline{A})\right\|_\infty^2 \leq c\prod_{j=k}^n (1-a_j)^{2(\overline{\lambda}-\delta)}$, for all $n \geq k \geq 1$, where $0 < \delta < \overline{\lambda}$. Hence, $\|G_n - G_n^*\|_\infty \leq c\sum_{k=1}^n a_k \prod_{j=k+1}^n (1-a_j)^{2(\overline{\lambda}-\delta)}\|U_k\|_\infty$. This, together with (49) and the Toeplitz Lemma implies

$$\|G_n - G_n^*\|_\infty \to 0 \quad (n \to \infty). \tag{52}$$

Moreover, we have $G_n^* = \sum_{k=1}^n a_k \prod_{j=k+1}^n (I - a_j\overline{A}) D \prod_{j=k+1}^n (I - a_j\overline{A}^\top)$, $n \geq 1$, from the definition (51). By Zhu (1996, Lemmas 9 and 10) it follows from (2) and (50) that $\|G_n^*\|_\infty \to M^2$ and, hence, $\|G_n\|_\infty \to M^2$ in view of (52), as n tends to infinity, which completes the proof of the second asymptotic relation (43).

In view of the relations (42) and (43) it is sufficient to check the conditions of Prosition 1 for the equation (30) in order to establish the limiting behaviour (41). Obviously, $B_1 = b_1^2 A_1^{-1} D (A_1^{-1})^\top > 0$ and the condition (11) is fulfilled by (42) and (2). Set $f_n = q_n^{1/2}/\prod_{j=1}^n (1-a_j)^\gamma$, $n \geq 1$, where $\gamma = \lambda^* - \delta$ (see (35)). Then $\lim_{n\to\infty} f_{n+1}/f_n = 1$ by (2), (7) and (43) which establishes condition (12).

By (43) we have $f_n \geq c\left(b_n/a_n^{1/2}\right)\prod_{j=1}^n (1-a_j)^{-\gamma}$. Analogously to (44) we obtain $f_n \geq c\exp\left(\sigma\sum_{j=1}^n a_j\right)$, $n \geq 1$. Conversely, by (43) and (6) we have $f_n \leq c\prod_{j=1}^n (1-a_j)^{-\gamma} \leq c\exp\left\{\gamma(1+\tilde{\varepsilon})\sum_{j=1}^n a_j\right\}$ for some $\tilde{\varepsilon} > 0$ and all sufficiently large n. Consequnetly, these two inequalities imply $\log\log f_n \sim \ell_n$ as n tends to infinity which, together with (42) establishes condition (13).

Finally, we get condition (14) by invoking inequality (36):

$$
\begin{aligned}
\frac{e_m^\top A_m B_m A_n^\top e_n}{(q_m q_n)^{1/2}} &\leq \frac{q_m \|A_n A_m^{-1}\|_\infty}{(q_m q_n)^{1/2}} \\
&\leq c\left(\frac{q_m}{q_n}\right)^{1/2} \prod_{j=m+1}^{n} (1-a_j)^\gamma \\
&= c\frac{f_m}{f_n},
\end{aligned}
$$

for all $n > m \geq 1$. □

Proof of Theorem 1. The result is immediate from Lemmas 4 and 5. □

5 A LIL for the nonlinear equation (1)

Theorem 2. *Assume that*

$$X_n \to \theta \quad a.\,s. \quad (n \to \infty), \tag{53}$$

and either

$$\|\Psi_n\| = O(b_n^2/a_n^{1/2}) \quad a.\,s. \quad (n \to \infty) \tag{54}$$

or

$$\|\Psi_n\| = o\left(\|X_{n-1} - \theta\|\right) \quad a.\,s. \quad (n \to \infty) \tag{55}$$

are satisfied. Assume further that

$$F(x) = A(x-\theta) + o\left(\|x-\theta\|\right) \qquad as\ x \to \theta \tag{56}$$

and that the conditions (2) and (5) to (9) are fulfilled. Then

$$\limsup_{n\to\infty} \left(\frac{a_n}{b_n^2 \ell_n}\right)^{1/2} \|X_n - \theta\| = \left(2\left\|\int_0^\infty e^{\left(\frac{\tau}{2}I - A\right)t} D e^{\left(\frac{\tau}{2}I - A^\top\right)t} dt\right\|_\infty\right)^{1/2} \quad a.\,s. \tag{57}$$

Proof. By Lemma 2 in Koval (1998) it follows from (2), (7), (10) and (53) to (56) that $||X_n - Y_n|| = o\left((b_n^2 \ell_n / a_n)^{1/2}\right)$ a.s., for $n \to \infty$, which, together with (10) implies the result (57). □

Remark. Note that the Lemma 2 in Koval (1998) is formulated for $\Psi \equiv 0$. However, its proof goes through if either condition (54) or (55) is satisifed. In fact, in the case of (54) we need additionally (45) which is ensured by condition (7).

We will give a short discussion of the assumptions imposed in Theorem 2. The consistency condition (53) is commonly assumed for investigations of the rate of convergence in the situation of a nonlinear response function F (see e. g. Mark, 1982, pp.36–37; Korostelev, 1984, p.161; Pelletier, 1998). The problem (53) has been investigated in great detail by Nevel'son and Khasminskii (1976), Korostelev (1984) and Kushner and Yin (1997), besides others. Concerning the condition (56) it has been used by Korostelev (1984, pp.160–161). Pelletier (1998) imposes a more stringent assumption. An analogous condition to (54) has been used e. g. by Walk (1988) and Zhu (1996), where the latter needs it in the central limit theorem (see also Nevel'son and Khasminskii, 1976, Chapter 4, Section 2). Conditions similar to (55) appear e. g. in Korostelev (1984) and Pelletier (1998).

Finally, we derive to corollaries of Thorem 2 (compare Korostelev, 1984, pp.160–162, Walk, 1988, Section 4, and Pelletier, 1998).

Corollary 1. *Assume conditions (53), (56) and (5). Let*

$$a_n = 1/n, \qquad b_n = 1/n^{(1+\beta)/2}, \tag{58}$$

where $0 < \beta < 2\lambda^*$. Further assume that either

$$\|\Psi_n\| = O\left(1/n^{\beta/2}\right) \quad a.\,s. \text{ as } n \to \infty$$

or (55) is satisfied. If, additionally, (9) holds, i. e.

$$\left\|\sum_{i=1}^{n} V_i - \sum_{i=1}^{n} \Gamma_i\right\| = o\left((\log\log n)^{1/2}\right) \quad a.\,s.(n \to \infty),$$

then

$$\limsup_{n\to\infty} \left(\frac{n^\beta}{\log\log n}\right)^{1/2} ||X_n - \theta|| = \left(2\left\|\int_0^\infty e^{\left(\frac{\beta}{2}I-A\right)t} D e^{\left(\frac{\beta}{2}I-A^\top\right)t} dt\right\|_\infty\right)^{1/2} \quad a.\,s.$$

Corollary 2. *Assume conditions (53), (56) and (5). Let*

$$a_n = 1/n^a, \quad 0 < a < 1, \quad \text{and} \quad b_n = 1/n^b, \quad \text{where } b > a/2. \tag{59}$$

Further assume that either

$$\|\Psi_n\| = O\left(n^{(a/2)-b}\right) \quad a.\ s.\ as\ n \to \infty$$

or (55) is satisfied. If, additionally, (9) holds, i. e.

$$\left\|\sum_{i=1}^{n} V_i - \sum_{i=1}^{n} \Gamma_i\right\| = o\left((n^a \log n)^{1/2}\right) \quad a.\ s. (n \to \infty),$$

then

$$\limsup_{n\to\infty} \left(\frac{n^{2b-a}}{\log n}\right)^{1/2} \|X_n - \theta\| = \left(2(1-a)\left\|\int_0^\infty e^{-At} D e^{-A^\top t} dt\right\|_\infty\right)^{1/2} \quad a.\ s.$$

The proofs of the corollaries are immediate as each of (58) and (59) implies the conditions (2) and (6) to (8) with $\tau = \beta$ and $\tau = 0$, respectively.

References

1. Berger, E. (1990). An almost sure invariance principle for stationary ergodic sequences of Banach spaced valued random variables. *Probab. Thory Related Fields* **84**, 161–201.

2. Berkes, I. and Philipp, W. (1979). Approximation theorems for independent and weakly dependent random vectors. *Ann. Probab.* **7**, 29–54.

3. Buldygin, V. and Solntsev, S. (1997). *Asymptotic behaviour of linearly transformed sums of random variables.* Kluwer, Dordrecht.

4. Eberlein, E. (1986). On strong invariance principles under dependence assumptions. *Ann. Probab.* **14**, 260–270.

5. Einmahl, U. (1987a). Strong invariance principles for partial sums of independent random vectors. *Ann. Probab.* **15**, 1419–1440.

6. Einmahl, U. (1987b). A useful estimate in the multidimensional invariance principle. *Probab. Thory Related Fields* **76**, 81–101.

7. Einmahl, U. (1989). Extensions of results of Komlós, Major and Tusnády to the multivariate case. *J. Multivariate Anal.* **28**, 20–68.

8. Horváth, L. (1993). A note on the law of the iterated logarithm for Abel sums. *C. R. Math. Rep. Acad. Sci. Canada* **15**, 213–217.

9. Korostelev, A. P. (1983). On upper functions for stochastic approximation procedues. *Theory Probab. Appl.* **28**, 806–811.

10. Korostelev, A. P. (1984). *Stochastic recursive procedures. Local properties.* Nauka, Moscow (in Russian).

11. Koval, V. A. (1998). Rate of convergence of stochastic approximation procedures in a Banach space. *Cybernetics Syst. Anal.* **34**, 386–394.

12. Koval, V. and Schwabe, R. (1998). Exact bounds for the rate of convergence in general stochastic approximation procedures. *Stochastic Anal. Appl.* **16**, 501–515.

13. Kuelbs, J. and Philipp, W. (1980). Almost sure invariance principles for partial sums of mixing B–valued random variables. *Ann. Probab.* **8**, 1003–1036.

14. Kushner, H. J. and Yin, G. G. (1997). *Stochastic approximation algorithms and applications.* Springer, New York.

15. Lacey, M. T. and Philipp, W. (1990). A note on the almost sure limit theorem. *Statistics Probab. Letters* **9**, 201–205.

16. Lai, T. L. (1973). Gaussian processes, moving averages and quick detection problems. *Ann. Probab.* **1**, 825–837.

17. Ledoux, M. and Talagrand, M. (1991). *Probability in Banach spaces.* Springer, Berlin.

18. Mark, G. (1982). Loglog–Invarianzprinzipien für Prozesse der stochastischen Approximation. *Mitteil. Math. Sem. Giessen* **153**.

19. Morrow, G. and Philipp, W. (1982). An almost sure invariance principle for Hilbert space valued matingales. *Trans. Amer. Math. Soc.* **273**, 231–251.

20. Nevel'son, M. B. and Khasminskii, R. Z. (1976). *Stochastic approximation and recursive estimation.* Amer. Math. Soc., Providence, RI.

21. Pelletier, M. (1998). On the almost sure asymptotic behaviour of stochastic algorithms. *Stochastic Processes Appl.* **78**, 217–244.

22. Philipp, W. (1986). Invariance principles for independent and weakly dependent random variables. In: *Dependence in Probability and Statistics 11.* (E. Eberlein and M. S. Taqqu, eds.). Birkhäuser, Basel, 225–268.

23. Walk, H. (1988). Limit behaviour of stochastic approximation procedures. *Statist. Decis.* **6**, 109–128.

24. Walk, H. and Zsidó, L. (1989). Convergence of the Robbins-Monro method for linear problems in a Banach space. *J. Math. Anal. Appl.* **139**, 152–177.

25. Wittmann, R. (1985). A general law of iterated logarithm. *Z. Wahrsch. verw. Gebiete* **68**, 521–543.

26. Zhu, Y. M. (1996). Asymptotic normality for a vector stochastic difference equation with applications in stochastic approximation. *J. Multivar. Anal.* **57**, 101–118.