Electron. Commun. Probab. **23** (2018), no. 37, 1–12.
https://doi.org/10.1214/18-ECP140
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

Nonconventional random matrix products*

Yuri Kifer[†] Sasha Sodin[‡]

Abstract

Let $\xi_1, \xi_2, ...$ be independent identically distributed random variables and $F : \mathbb{R}^\ell \to SL_d(\mathbb{R})$ be a Borel measurable matrix-valued function. Set $X_n = F(\xi_{q_1(n)}, \xi_{q_2(n)}, ..., \xi_{q_\ell(n)})$ where $0 \le q_1 < q_2 < ... < q_\ell$ are increasing functions taking on integer values on integers. We study the asymptotic behavior as $N \to \infty$ of the singular values of the random matrix product $\Pi_N = X_N \cdots X_2 X_1$ and show, in particular, that (under certain conditions) $\frac{1}{N} \log \|\Pi_N\|$ converges with probability one as $N \to \infty$. We also obtain similar results for such products when ξ_i form a Markov chain. The essential difference from the usual setting appears since the sequence $(X_n,\, n \ge 1)$ is long-range dependent and nonstationary.

Keywords: random matrix products; large deviations; avalanche principle; nonconventional limit theorems.
AMS MSC 2010: 60B20; 60F15; 60F10; 82B44.
Submitted to ECP on March 25, 2018, final version accepted on May 29, 2018.
Supersedes arXiv:1803.09221.

1 Introduction

Products $\Pi_N = X_N \cdots X_2 X_1$ of random matrices $X_1, X_2, ...$ are extensively studied for more than half a century now. In the pioneering work [7], it was shown that when $X_1, X_2, ...$ form a stationary sequence with $E \ln^+ \|X_1\| < \infty$ then the limit $\gamma_1 = \lim_{N\to\infty} \frac{1}{N} \ln \|\Pi_N\|$ exists with probability one. Later, the more general Kingman's subadditive ergodic theorem became available and it yielded the above result as a corollary. Applying it to actions on the exterior products, the result was extended to all the singular values of Π_N, thus leading to the Oseledets multiplicative ergodic theorem.

In this paper we study similar questions for products of certain nonstationary sequences of random matrices. Namely, we start with a sequence of i.i.d. random variables $\xi_1, \xi_2, ...$ and a Borel measurable matrix valued function $F : \mathbb{R}^\ell \to SL_d(\mathbb{R})$ along with integer valued functions $0 \le q_1 < q_2 < ... < q_\ell$, and form the random matrices $X_n = F(\xi_{q_1(n)}, \xi_{q_2(n)}, ..., \xi_{q_\ell(n)})$. In particular, we allow arithmetic progressions $q_i(n) = in$, $i = 1, ..., \ell$. The sequence $X_1, X_2, ...$ is long range dependent and is not stationary, and so the study of the asymptotic behavior as $N \to \infty$ of the product $\Pi_N = X_N \cdots X_2 X_1$ is not described by the standard results mentioned above. Still, we

*S.Sodin was supported in part by the European Research Council starting grant SPECTRUM (639305) and by a Royal Society Wolfson Research Merit Award.

[†]Institute of Mathematics, Hebrew University of Jerusalem. E-mail: kifer@math.huji.ac.il

[‡]School of Mathematical Sciences, Queen Mary University of London & School of Mathematical Sciences, Tel Aviv University. E-mail: a.sodin@qmul.ac.uk

show that $\lim_{N\to\infty}\frac{1}{N}\ln\|\Pi_N\|$ exists with probability one and applying this to exterior products we will obtain corresponding results for all the singular values of Π_N. Similar results are obtained also for such products when ξ_i form a Markov chain satisfying certain conditions of the type of uniform geometric ergodicity.

The motivation for this paper is twofold. On one hand, it comes from the vast body of research on products of random matrices mentioned above (see [4] and [3]). In particular, our results provide a non-trivial family of random discrete Schrödinger equations $\psi_{n+1} = (\lambda - V_n)\psi_n - \psi_{n-1}$ which are not metrically transitive and yet the asymptotics of solutions can be described. Here, as usual, $\Delta\psi(n) = -(\psi(n+1)+\psi(n-1))$ is viewed as a discrete counterpart of the Laplacian. In our case, $V_n = \varphi(\xi_{q_1(n)}, \xi_{q_2(n)}, ..., \xi_{q_\ell(n)})$ and $\xi_1, \xi_2, ...$ are, say, i.i.d. random variables.

On the other hand, our motivation stems from the series of papers, originating in Furstenberg's proof of the Szemerédi theorem, on nonconventional ergodic and limit theorems which dealt with the sums of the form $\sum_{n=1}^{N}\varphi(\xi_{q_1(n)}, \xi_{q_2(n)}, ..., \xi_{q_\ell(n)})$ (see, for instance, [10] and references therein). Our results can be viewed as a counterpart of the nonconventional strong law of large numbers in the multiplicative setting.

2 Preliminaries and main results

2.1 I.i.d. case

Let $\xi_1, \xi_2, ...$ be i.i.d. random variables, and let $F : \mathbb{R}^\ell \to SL_d(\mathbb{R})$ be a Borel measurable matrix valued function where $\ell > 1$ (since for $\ell = 1$ the results of this paper are well known). Our setup also includes an ℓ-tuple of strictly increasing nonnegative functions $q_1 < q_2 < ... < q_\ell$ taking on integer values on integers with $q_1(1) \geq 1$. Set $X_n = F(\xi_{q_1(n)}, \xi_{q_2(n)}, ..., \xi_{q_\ell(n)})$ and observe that each X_n, $n \geq 1$ has the same distribution, since each ℓ-tuple $\xi_{q_1(n)}, \xi_{q_2(n)}, ..., \xi_{q_\ell(n)}$ has the same distribution as $\xi_1, \xi_2, ..., \xi_\ell$. Denote by μ the distribution of X_1 and by G_μ the support of μ. We will need the following

Assumption 2.1.

(i) *G_μ is strongly irreducible, i.e. there does not exist a finite union of proper subspaces of $\mathbb{R}^d$ that is preserved as a set by all matrices from G_μ (see [4]).*

(ii) *For some $\alpha > 0$,*

$$E\|X_1\|^\alpha < \infty. \tag{2.1}$$

(iii) *for any $\sigma > 0$ there exists $n_0(\sigma)$ such that for all $n \geq n_0(\sigma)$,*

$$q_{i+1}(n) \geq q_i(n + [\sigma \ln n]), \quad i = 1, ..., \ell - 1. \tag{2.2}$$

Clearly, (2.2) is satisfied, for instance, in the arithmetic progression case $q_i(n) = in$, $i = 1, ..., \ell$.

Recall that the singular values $s_1(g) \geq s_2(g) \geq ... \geq s_d(g) \geq 0$ of a $d \times d$ matrix g are the square roots of the eigenvalues $s_i^2(g)$ of g^*g. The first singular value $s_1(g)$ is the Euclidean operator norm of g,

$$s_1(g) = \max_{x\in\mathbb{R}^d\setminus\{0\}} \frac{\|gx\|}{\|x\|} = \|g\|.$$

If $X \in SL_d(\mathbb{R})$ then $1 = s_1(X)s_2(X)\cdots s_d(X) \leq s_1^{d-1}(X)s_d(X)$, and so $\|X^{-1}\| = s_d^{-1}(X) \leq s_1^{d-1}(X) = \|X\|^{d-1}$. Hence, (2.1) implies also that

$$E\|X_1^{-1}\|^{\alpha'} < \infty \quad \text{with} \quad \alpha' = \frac{\alpha}{d-1},\ d > 1.$$

Since $F \equiv 1$ if $d = 1$ and the problems discussed here become trivial then, we assume without loss of generality that $d > 1$.

Let $Y_1, Y_2, ...$ be an i.i.d. sequence of random matrices having the distribution μ, and so satisfying (i) and (ii) of Assumption 2.1 with Y_1 in place of X_1. Hence (cf. [4, 3]), the limits

$$\gamma_i = \lim_{N\to\infty} \frac{1}{N} \ln s_i(Y_N \cdots Y_2 Y_1),\ i = 1, ..., d, \tag{2.3}$$

exist with probability one; in particular, $\gamma_1 = \lim_{N\to\infty} \frac{1}{N} \ln \|Y_N \cdots Y_2 Y_1\|$. The following theorem asserts that the similar result holds true for $\Pi_N = X_N \cdots X_2 X_1$.

Theorem 2.2. *Suppose that Assumption 2.1 holds true. Then with probability one*

$$\lim_{N\to\infty} \frac{1}{N} \ln s_i(\Pi_N) = \gamma_i,\ i = 1, ..., d, \tag{2.4}$$

where $\gamma_1, ..., \gamma_d$ are the same as in (2.3). In particular, $\lim\limits_{N\to\infty} \frac{1}{N} \ln \|\Pi_N\| = \gamma_1$.

Observe that it suffices to prove Theorem 2.2 only for the largest singular value, i.e. for $i = 1$. Indeed, note that (i) and (ii) of Assumption 2.1 remain valid for the exterior powers $\wedge^i \Pi_N$, $i = 1, ..., d$ of Π_N (defined by $\wedge^i \Pi_N (x_1 \wedge ... \wedge x_i) = \Pi_N x_1 \wedge ... \wedge \Pi_N x_i$) if these were true for Π_N itself (see [4]). Hence, proving Theorem 2.2 for each $s_i(\wedge^i \Pi_N)$ we will obtain that

$$\lim_{N\to\infty} \frac{1}{N} \ln s_1(\wedge^i \Pi_N) = \lim_{N\to\infty} \frac{1}{N} \ln \prod_{j=1}^{i} s_j(\Pi_N) = \sum_{j=1}^{i} \gamma_j \tag{2.5}$$

which yields (2.4).

The proof of Theorem 2.2, presented in Sections 3 and 4, is based on two main ingredients. The first one is a large deviations bound for products of random matrices which was first proved by Le Page under the additional contraction assumption. We rely on a version of this result from Theorem 14.19 in [3] which does not require the contraction condition. In fact, the upper bound of large deviations from Theorem 6.2 on p.131 of [4] suffices for our purposes, as well. The second ingredient playing a decisive role in our proof of the lower bound below is the avalanche principle proved originally for two dimensional matrices in [8] and extended (in a strengthened form) to the multidimensional case in [6]. It is not difficult to see that the convergence in Theorem 2.2 holds true also in mean which does not require large deviations estimates but only a subadditivity argument together with the avalanche principle.

2.2 Markov case

Next, we discuss the case when $\xi_0, \xi_1, \xi_2, ...$ form a Markov chain on a Polish space $\mathcal{E}$ (to conform with the standard notation, we start the indices from 0), $F : \mathcal{E}^\ell \to SL_d(\mathbb{R})$ is a Borel measurable matrix function and $X_n = F(\xi_{q_1(n)}, \xi_{q_2(n)}, ..., \xi_{q_\ell(n)})$ with $q_i(n)$, $i = 1, ..., \ell$ satisfying Assumption 2.1(iii). Let $P(n, x, \cdot)$, $x \in \mathcal{E}$ be the n-step transition probability of the Markov chain above, $P(x, \cdot) = P(1, x, \cdot)$ and assume that there exists a probability measure ν on $\mathcal{E}$ such that for some $R, \rho > 0$, all $n \geq 1$ and any bounded Borel function f on $\mathcal{E}$,

$$\sup_{x\in\mathcal{E}} \left| \int P(n, x, dy) f(y) - \int f d\nu \right| \leq R e^{-\rho n} \sup_{x\in\mathcal{E}} |f(x)|. \tag{2.6}$$

This assumption will be satisfied for an aperiodic Markov chain if, for instance, a version of the Doeblin condition holds true (see, for instance, [5], Section 21.23). It follows that ν is the unique invariant measure of this Markov chain, i.e. the only measure ν

satisfying $\int d\nu(x)P(x,\Gamma) = \nu(\Gamma)$ for any Borel set $\Gamma \subset \mathcal{E}$, and so ν is ergodic. Taking ν as the initial distribution of the Markov chain, i.e. as the distribution of ξ_0, makes it a stationary ergodic process. Still, the condition (2.6) will enable us to obtain stronger results for the Markov chain starting at any initial point $x \in \mathcal{E}$.

Let $\{\xi_n^{(i)},\, n \geq 0\}$, $i = 1, ..., \ell$ be ℓ independent copies of the Markov chain $\{\xi_n,\, n \geq 0\}$ which produces an ℓ-component Markov chain $\Xi_n = (\xi^{(1)}_{q_1(n)}, \xi^{(2)}_{q_2(n)}, ..., \xi^{(\ell)}_{q_\ell(n)})$, $n \geq 0$ with the transition probabilities $P_\Xi(\bar{x}, \Gamma_1 \times \Gamma_2 \times \cdots \times \Gamma_\ell) = \prod_{i=1}^{\ell} P(x_i, \Gamma_i)$ where $\bar{x} = (x_1, ..., x_\ell)$. Set $Y_n = F(\xi^{(1)}_{q_1(n)}, \xi^{(2)}_{q_2(n)}, ..., \xi^{(\ell)}_{q_\ell(n)})$, $n \geq 0$ and assume that for some $\alpha > 0$,

$$\sup_{\bar{x} \in \mathcal{E}^\ell} E_{\bar{x}} \|Y_1\|^\alpha < \infty \tag{2.7}$$

where $E_{\bar{x}}$, $\bar{x} = (x_1, ..., x_\ell)$ is the expectation with respect to the probability $P_{\bar{x}}$ of the Markov chain Ξ_n, $n \geq 0$ starting at $\bar{x}$.

Set $H_n = Y_n \cdots Y_2 Y_1$. It follows from [2] (see also Section 5) that the limits

$$\lim_{N \to \infty} \frac{1}{N} \ln s_i(H_N) = \gamma_i, \;\; i = 1, ..., d \tag{2.8}$$

exist $P_{\bar{x}}$-almost surely (a.s.) for each $\bar{x} \in \mathcal{E}^\ell$ where, again, $s_i(g)$ is the i-th singular value of a matrix g. Viewing (2.8) as a definition of γ_i's we assume also that, for some $1 < k \leq d$,

$$\gamma_1 > \gamma_2 > \cdots > \gamma_k\ ; \tag{2.9}$$

sufficient conditions for this can be found in [1] and [12]. In addition, following [2] we assume quasi-irreducibility which means that the subspaces

$$V(\bar{x}) = \{u \in \mathbb{R}^d :\ \lim_{N \to \infty} \frac{1}{N} \ln \|H_N u\| \leq \gamma_2 \quad P_{\bar{x}} - \text{a.s.}\}$$

are trivial for almost all $\bar{x} = (x_1, ..., x_\ell)$ with respect to the product measure $\bar{\nu} = \nu \times \cdots \times \nu$. Denote by P_x the path space probability of the Markov chain ξ_n, $n \geq 0$ provided that $\xi_0 = x$.

Theorem 2.3. *Assume the above conditions (2.6), (2.7), (2.8), (2.9) and quasi-irreducibility. The singular values $s_i(\Pi_N)$, $i = 1, ..., k-1$ of $\Pi_N = X_N \cdots X_2 X_1$ satisfy*

$$\lim_{N \to \infty} \frac{1}{N} \ln s_i(\Pi_N) = \gamma_i \quad P_x - \text{a.s.} \tag{2.10}$$

for each $x \in \mathcal{E}$.

The proof of this result will be given in Section 5 relying on the large deviations theorem for products of Markov dependent random matrices from [2] and an additional argument enabling us to compare large deviations estimates for the products H_m and for $\Pi_{n+m}\Pi_n^{-1}$ in spite of the fact that the latter is not a product of Markov dependent random matrices.

3 Upper bound

There are two cases in the proof of Theorem 2.2: $\gamma_1 = 0$ and $\gamma_1 > 0$. The first case requires only the upper bound since $\ln \|A\| \geq 0$ for any $A \in SL_d(\mathbb{R})$. The second case will require both a lower and an upper bound so we will start with the latter which will serve in both cases. In fact, by Furstenberg's theorem (see Theorem 6.3 on p.66 in [4]) under the strong irreducibility condition, $\gamma_1 = 0$ if and only if G_μ is contained in a compact subgroup; then each X_n belongs to this subgroup too and Theorem 2.2 follows in this case directly.

It follows from the large deviations theorem for products of i.i.d. random matrices (see [4, p.131, Theorem 6.2] and [3, Theorem 14.19]) that for any $\varepsilon > 0$ there exists $\kappa(\varepsilon) > 0$ and $n_1(\varepsilon) \geq 1$ such that

$$P\{\frac{1}{n}\ln\|Y_n\cdots Y_2Y_1\| > \gamma_1+\varepsilon\} \leq e^{-\kappa(\varepsilon)n} \tag{3.1}$$

for all $n \geq n_1(\varepsilon)$. Without loss of generality we assume that $\kappa(\epsilon) < 1$. Fix $\varepsilon > 0$ and set $r(n) = r_\varepsilon(n) = [\frac{2}{\kappa(\varepsilon)}\ln n]$. Observe that if $r(n) \geq 1$ and $q_i(n+r(n)) \leq q_{i+1}(n)$ for $i = 1, ..., \ell-1$ then $X_n, X_{n+1}, ..., X_{n+r(n)-1}$ is an i.i.d. tuple having the same distribution as $(Y_1, Y_2, ..., Y_{r(n)})$. Set $n_2(\varepsilon) = \min\{m \geq n_0(\frac{2}{\kappa(\varepsilon)}) : r(m) \geq n_1(\varepsilon)\}$ where n_0 comes from Assumption 2.1(iii). Then for all $n \geq n_2(\varepsilon)$,

$$P\{\frac{1}{r(n)}\ln\|X_{n+r(n)-1}\cdots X_{n+1}X_n\| > \gamma_1+\varepsilon\} \leq e^{-\kappa(\varepsilon)r(n)}. \tag{3.2}$$

This together with (2.2) and the Borel-Cantelli lemma yields existence of a finite with probability one random variable $M_1(\varepsilon) = M_1(\varepsilon,\omega)$ such that for any $n \geq M_1(\varepsilon)$,

$$\ln\|X_{n+r(n)-1}\cdots X_{n+1}X_n\| \leq r(n)(\gamma_1+\varepsilon). \tag{3.3}$$

Set $m_1 = n_2(\varepsilon)$ and recursively $m_{i+1} = m_i + r(m_i)$. Then $m_1 < m_2 < m_3 < ...$ and $m_n \to \infty$ as $n \to \infty$. Hence, $X_{m_i}, X_{m_i+1}, ..., X_{m_{i+1}-1}$ is a tuple of i.i.d. random matrices for each $i \geq 1$. In particular, when $m_i \geq M_1(\varepsilon)$ we have by (3.3) that

$$\ln\|X_{m_{i+1}-1}\cdots X_{m_i+1}X_{m_i}\| \leq (m_{i+1}-m_i)(\gamma+\varepsilon). \tag{3.4}$$

By the submultiplicative property of the Euclidean operator matrix norm,

$$\begin{aligned}\ln\|X_N\cdots X_2X_1\| \leq &\sum_{N\geq j\geq m_{k(N)}}\ln\|X_j\| \\ &+ \sum_{i:\, M_1(\varepsilon)\leq m_i<k(N)}\ln\|X_{m_{i+1}-1}\cdots X_{m_i+1}X_{m_i}\| \\ &+ \sum_{1\leq j\leq \max(n_2(\varepsilon),M_1(\varepsilon)+r(M_1(\varepsilon))}\ln\|X_j\|\end{aligned} \tag{3.5}$$

where $k(N) = k_\varepsilon(N) = \max\{i : m_i \leq N\}$. Since the last sum is a fixed random variable (depending on ε) which is finite with probability one then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{1\leq j\leq \max(n_2(\varepsilon),M_1(\varepsilon)+r(M_1(\varepsilon))}\ln\|X_j\| = 0 \quad \text{almost surely.} \tag{3.6}$$

Next, we observe that by the Chebyshev inequality

$$P\{\ln\|X_n\| \geq \frac{2}{\alpha}\ln n\} = P\{\|X_1\|^\alpha \geq n^2\} \leq Dn^{-2} \tag{3.7}$$

where $D = E\|X_1\|^\alpha < \infty$ by (2.1). Hence, by the Borel-Cantelli lemma there exists a finite with probability one random variable $M_2 = M_2(\omega)$ such that for all $n \geq M_2$,

$$\ln\|X_n\| < \frac{2}{\alpha}\ln n. \tag{3.8}$$

Observe that

$$N - m_{k(N)} < r(m_{k(N)}) \leq r(N) = [\frac{2}{\kappa(\varepsilon)}\ln N], \tag{3.9}$$

and so, in particular, $m_{k(N)} \to \infty$ as $N \to \infty$. Thus, it suffices to estimate the first expression in the right hand side of (3.5) on the events $\Gamma_N = \{\omega : M_2(\omega) \le m_{k(N)}\}$. By (3.8) and (3.9) on the event Γ_N,

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{N\ge j\ge m_{k(N)}} \ln \|X_j\| \le \limsup_{N\to\infty} \frac{2}{N\alpha} r(N) \ln n = 0. \tag{3.10}$$

Finally, collecting (3.4)–(3.6) and (3.8)–(3.10) we see that with probability one

$$\limsup_{N\to\infty} \frac{1}{N} \ln \|X_N \cdots X_2 X_1\| \le \gamma_1 + \varepsilon.$$

Since $\varepsilon > 0$ is arbitrary we obtain the required upper bound

$$\limsup_{N\to\infty} \frac{1}{N} \ln \|\Pi_N\| = \limsup_{N\to\infty} \frac{1}{N} \ln \|X_N \cdots X_2 X_1\| \le \gamma_1 \tag{3.11}$$

with probability one. If $\gamma_1 = 0$ this already implies (2.4) while in the case $\gamma_1 > 0$ we shall also need the corresponding lower bound.

4 Lower bound

First, observe that without loss of generality we can assume here that $\gamma_1 > \gamma_2$ where the γ_i's were defined in (2.3). Indeed, either $\gamma_1 = \gamma_2 = ... = \gamma_d$ and then $\gamma_i = 0$ for all i's since all the matrices here have determinant equal one, or $\gamma_1 = ... = \gamma_k > \gamma_{k+1} \ge ... \ge \gamma_d$ for some $1 \le k < d$. Then we can prove the result for the first singular value of the k-th exterior power $\wedge^k \Pi_N$ of Π_N obtaining that with probability one,

$$\lim_{N\to\infty} \frac{1}{N} \ln s_1(\wedge^k \Pi_N) = \lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{k} \ln s_i(\Pi_N) = k\gamma_1.$$

Since $s_1(\Pi_N) \ge s_2(\Pi_N) \ge ... \ge s_d(\Pi_N)$ and $\lim_{N\to\infty} \frac{1}{N} s_1(\Pi_N) \le \gamma_1$ with probability one by the upper bound we obtain that, in fact, the last inequality is the equality. Thus, we obtain Theorem 2.2 for $s_1(\Pi_N)$ which is sufficient for its full statement as explained in Section 2.

Hence, we can and will assume here that $\gamma_1 > \gamma_2$, $\gamma_1 > 0$ and start with another bound of large deviations for products of i.i.d. random matrices (see [3]) which in the same notation as in Section 3 says that for any $\varepsilon > 0$ there exists $\kappa(\varepsilon) > 0$ and $n_1(\varepsilon) \ge 1$ such that

$$P\{\frac{1}{n} \ln \|Y_n \cdots Y_2 Y_1\| < \gamma_1 - \varepsilon\} \le e^{-\kappa(\varepsilon)n} \tag{4.1}$$

for all $n \ge n_1(\varepsilon)$. Let $r(n)$ and $n_2(\varepsilon)$ be the same as in Section 3. Then, for all $n \ge n_2(\varepsilon)$ we obtain

$$P\{\frac{1}{r(n)} \ln \|X_{n+r(n)-1} \cdots X_{n+1} X_n\| < \gamma_1 - \varepsilon\} \le e^{-\kappa(\varepsilon)r(n)}. \tag{4.2}$$

Since there exists no inequality similar to (3.5) to employ for a proof of the lower bound we will need a more advanced argument in order to make use of the splitting of the product $X_N \cdots X_2 X_1$ into appropriate products of i.i.d. matrices. Namely, we will rely on the avalanche principle which appears for products of multidimensional matrices in [6]. Following [6] for each $g \in GL_d(\mathbb{R})$ we set

$$gr(g) = \frac{s_1(g)}{s_2(g)}$$

which is called the gap of $g \in GL_d(\mathbb{R})$. Now we have (see [6, §2.4]),

Theorem 4.1. *(Avalanche Principle). There exist universal constants* $c, C > 0$ *such that whenever* $a \geq cb > c$ *and* $g_j \in GL_d(\mathbb{R})$, $j = 1, ..., l$ *satisfy*

(i) $gr(g_j) \geq a$, $j = 1, ..., l$ *and*
(ii) $\ln \|g_{j+1}g_j\| - \ln \|g_{j+1}\| - \ln \|g_j\| \geq -\frac{1}{2} \ln b$

then

$$\ln \|g_l \cdots g_2 g_1\| + \sum_{j=2}^{l-1} \ln \|g_j\| \geq \sum_{j=1}^{l-1} \ln \|g_{j+1}g_j\| - Cl\frac{b}{a},\ j = 1, ..., l-1. \tag{4.3}$$

Observe that from (ii) and (4.3) we obtain

$$\ln \|g_l \cdots g_2 g_1\| \geq \sum_{j=1}^{l} \ln \|g_j\| - \frac{1}{2} l \ln b - Cl\frac{b}{a}. \tag{4.4}$$

Let us take

$$g_j = X_{m_{j+1}-1} \cdots X_{m_{j+1}} X_{m_j} \tag{4.5}$$

where $m_1 < m_2 < ... < m_{k(N)}$ are as in Section 3. This together with (4.4) will yield the required lower bound of the form

$$\liminf_{N\to\infty} \frac{1}{N} \ln \|X_N \cdots X_2 X_1\| \geq \gamma_1 - \varepsilon \tag{4.6}$$

provided that we can obtain appropriate bounds on parameters a and b in the avalanche principle above.

Now, (4.2) together with the definition of $r(n) = r_\varepsilon(n)$ and the Borel-Cantelli lemma yield that there exists a finite with probability one random variable $M_1(\varepsilon)$ such that for any $n \geq M_1(\varepsilon)$,

$$\ln \|X_{n+r(n)-1} \cdots X_{n+1} X_n\| \geq r(n)(\gamma_1 - \varepsilon). \tag{4.7}$$

In particular, for each $i < k(N)$ such that $m_i \geq M_1(\varepsilon)$ we have

$$\ln \|X_{m_{i+1}-1} \cdots X_{m_i+1} X_{m_i}\| \geq (m_{i+1} - m_i)(\gamma_1 - \varepsilon). \tag{4.8}$$

Next, set $j_N = \min\{j : m_j \geq \sqrt{N}\}$. By the submultiplicative property of the Euclidean matrix norm,

$$\begin{aligned} \ln \|X_N \cdots X_2 X_1\| &\geq \ln \|X_{m_{k(N)}-1} \cdots X_{m_{j_N}+1} X_{m_{j_N}}\| \\ &- \ln \|(X_N \cdots X_{m_{k(N)}+1} X_{m_{k(N)}})^{-1}\| - \ln \|(X_{m_{j_N}-1} \cdots X_2 X_1)^{-1}\|. \end{aligned} \tag{4.9}$$

As explained in Section 2 the condition (2.1) implies also that $D' = E\|X_1^{-1}\|^{\alpha'} < \infty$ where $\alpha' = \frac{\alpha}{d-1}$. Thus, in the same way as in (3.7) we have

$$P\{\ln \|X_n^{-1}\| \geq \frac{2}{\alpha'} \ln n\} \leq D' n^{-2}, \tag{4.10}$$

and so by the Borel-Cantelli lemma there exists a finite with probability one random variable $M_2' = M_2'(\omega)$ such that for all $n \geq M_2'$,

$$\ln \|X_n^{-1}\| < \frac{2}{\alpha'} \ln n. \tag{4.11}$$

Thus, similarly to (3.10) we obtain that on the event $\Gamma_N' = \{\omega : M_2'(\omega) \leq m_{k(N)}\}$,

$$\limsup_{N\to\infty} \frac{1}{N} \ln \|(X_N \cdots X_{m_{k(N)+1}} X_{m_{k(N)}})^{-1}\| \leq \limsup_{N\to\infty} \frac{1}{N} \sum_{N\geq j\geq m_{k(N)}} \ln \|X_j^{-1}\| = 0. \tag{4.12}$$

On the other hand,

$$\begin{aligned}&\limsup_{N\to\infty}\frac{1}{N}\ln\|(X_{m_{j_N-1}}\cdots X_2X_1)^{-1}\|\\ &\le \limsup_{N\to\infty}\frac{1}{N}\sum_{1\le j<M_2'}\|X_j^{-1}\|+\limsup_{N\to\infty}\frac{1}{N}\sum_{M_2'\le j\le m_{j_N}}\|X_j^{-1}\|=0.\end{aligned}\tag{4.13}$$

Indeed, the first limit in the right hand side of (4.13) is zero since the sum there is a fixed random variable which is finite with probability one. The second limit there is zero in view of (4.12) and the estimate $m_{j_N}\le\sqrt{N}+r([\sqrt{N}])$.

Applying the avalanche principle we will show that in the above case with probability one

$$\begin{aligned}&\liminf_{N\to\infty}\tfrac{1}{N}\ln\|X_{m_{k(N)}-1}\cdots X_{m_{j_N}+1}X_{m_{j_N}}\|\\ &=\liminf_{N\to\infty}\tfrac{1}{N}\ln\|g_{k(N)-1}\cdots g_{j_N+1}g_{j_N}\|\ge\gamma_1-7\varepsilon.\end{aligned}\tag{4.14}$$

First, we estimate the avalanche principle parameters $a=a(\varepsilon,N)$ and $b=b(\varepsilon,N)$ which will depend on ε and N. Set $g(n)=X_{n+r(n)-1}\cdots X_{n+1}X_n$ so that $g_j=g(m_j)$, and let $s_1(g(n))\ge s_2(g(n))\ge...\ge s_d(g(n))>0$ be the singular values of $g(n)$. The second exterior power $\wedge^2 g(n)$ of $g(n)$ acting on the second exterior power $\wedge^2\mathbb{R}^d$ of $\mathbb{R}^d$ has the biggest singular value equal to $s_1(g(n))s_2(g(n))$. Hence

$$gr(g(n))=\frac{s_1(g(n))}{s_2(g(n))}=\frac{s_1^2(g(n))}{s_1(g(n))s_2(g(n))}=\frac{\|g(n)\|^2}{\|\wedge^2 g(n)\|}\tag{4.15}$$

where $\|\cdot\|$ is the Euclidean operator norm.

Now, set $H_n=Y_n\cdots Y_2Y_1$ with $Y_1,Y_2,...$ introduced in Section 2. Under our conditions with probability one

$$\lim_{n\to\infty}\frac{1}{n}\ln\|H_n\|=\gamma_1\ \text{ and }\ \lim_{n\to\infty}\frac{1}{n}\ln\|\wedge^2 H_n\|=\gamma_1+\gamma_2\tag{4.16}$$

and, recall that $\gamma_2<\gamma_1$. Applying the large deviations bounds to $\|H_n\|$ and to $\|\wedge^2 H_n\|$ we obtain that for any $\varepsilon>0$ there exists $\kappa(\varepsilon)>0$ (which could be different from before but we denote it by the same letter) and $n_3(\varepsilon)\ge 1$ such that

$$P\{\frac{1}{n}\ln\|\wedge^2 H_n\|>\gamma_1+\gamma_2+\varepsilon\}\le e^{-\kappa(\varepsilon)n}\tag{4.17}$$

for all $n\ge n_3(\varepsilon)$. Hence, if $r(n)\ge n_3(\varepsilon)$ and $n\ge n_0(\frac{2}{\kappa(\varepsilon)})$ then

$$P\{\frac{1}{r(n)}\ln\|\wedge^2 g(n)\|>\gamma_1+\gamma_2+\varepsilon\}\le e^{-\kappa(\varepsilon)r(n)}.\tag{4.18}$$

This together with (4.2) and (4.15) yields that

$$P\{gr(g(n))<e^{(\gamma_1-\gamma_2-2\varepsilon)r(n)}\}\le 2e^{-\kappa(\varepsilon)r(n)}.\tag{4.19}$$

Taking into account that $r(n)=[\frac{2}{\kappa(\varepsilon)}\ln n]$ we conclude from (4.19) and the Borel-Cantelli lemma that there exists a finite with probability one random variable $M_3(\varepsilon)$ such that for any $n\ge M_3(\varepsilon)$,

$$gr(g(n))\ge e^{(\gamma_1-\gamma_2-2\varepsilon)r(n)}.\tag{4.20}$$

Next, we use that by our choice of $r(n)$ there exists $n_4(\varepsilon)\ge 1$ such that if $n\ge n_4(\varepsilon)$ then $X_n,X_{n+1},...,X_{n+r(n)+r(n+r(n+r(n)))-1}$ is an i.i.d. tuple having the same distribution as $Y_1, Y_2, ..., Y_{r(n)+r(n+r(n+r(n)))}$. Thus, similarly to the above, relying on the large deviations bound (4.1) together with the Borel-Cantelli lemma we conclude that there exists a finite with probability one random variable $M_4(\varepsilon)$ such that for any $n \geq M_4(\varepsilon)$,

$$\begin{aligned} &\|X_{n+r(n)+r(n+r(n+r(n)))-1} \cdots X_n X_{n+1}\| \\ &= \|g(n+r(n))g(n)\| \geq e^{(\gamma_1-\varepsilon)(r(n)+r(n+r(n)))}. \end{aligned} \tag{4.21}$$

Applying the large deviations estimate (3.3) to $\|g(n)\|$ and to $\|g(n+r(n))\|$ together with the Borel-Cantelli lemma we obtain that there exists a finite with probability one random variable $M_5(\varepsilon)$ such that for any $n \geq M_5(\varepsilon)$,

$$\|g(n)\| \leq e^{(\gamma_1+\varepsilon)r(n)} \text{ and } \|g(n+r(n))\| \leq e^{(\gamma_1+\varepsilon)r(n+r(n))}. \tag{4.22}$$

Let $n_5(\varepsilon)$ be such that $\frac{2}{\kappa(\varepsilon)}\ln(1+\frac{r_\varepsilon(n)}{n}) \leq 1$ for any $n \geq n_5(\varepsilon)$. Then, by (2.2), (4.21) and (4.22) for any $n \geq \max(n_5(\varepsilon), M_5(\varepsilon))$,

$$\frac{\|g(n+r(n))g(n)\|}{\|g(n)\|\|g(n+r(n))\|} \geq e^{-3\varepsilon(r(n)+r(n+r(n)))} \geq e^{-6\varepsilon(r(n)+1)}. \tag{4.23}$$

Observe that for $n \geq \sqrt{N}$ the numbers $k(N) = \max\{i: m_i < N\}$ and $j_N = \min\{j : m_j \geq \sqrt{N}\}$ satisfy

$$k(N) - j_N \leq \frac{N}{r([\sqrt{N}])} = \frac{N}{[\frac{1}{\kappa(\varepsilon)}\ln N]}. \tag{4.24}$$

When $n \geq \sqrt{N}$ then (4.20) and (4.23) hold true for

$$\omega \in \Omega_{\varepsilon,N} = \{\omega : \max(M_3(\varepsilon,\omega), M_4(\varepsilon,\omega), M_4(\varepsilon,\omega)(\omega), n_5(\varepsilon)) \leq \sqrt{N}\}.$$

Clearly, $\Omega_{\varepsilon,N} \uparrow \tilde{\Omega}$ with $P(\tilde{\Omega}) = 1$. Thus we can estimate the parameters of the avalanche principle for $\omega \in \Omega_{\varepsilon,N}$ and each fixed N large enough and then let $N \to \infty$.

It follows from (4.20), (4.23) and (4.24) that applying the avalanche principle to $g_{j_N}, g_{j_N+1}, ..., g_{k(N)-1}$ we can take in (4.4),

$$\begin{aligned} l = l(N) = k(N) - j_N,\ a = a(\varepsilon, N) = e^{(\gamma_1-\gamma_2-2\varepsilon)r([\sqrt{N}])} \\ \text{and } b = b(\varepsilon, N) = e^{6\varepsilon(r(N)+1)}. \end{aligned}$$

Choosing ε much smaller than $\frac{1}{8}(\gamma_1 - \gamma_2)$ we let $N \to \infty$ to obtain from (4.4), (4.8), (4.20), (4.23) and (4.24) together with the avalanche principle that (4.14) holds true. These together with (4.9), (4.12) and (4.13) yield that

$$\liminf_{N\to\infty} \frac{1}{N}\|X_N \cdots X_2 X_1\| \geq \gamma_1 - 7\varepsilon. \tag{4.25}$$

Now we let $\varepsilon \to 0$ and obtain

$$\liminf_{N\to\infty} \frac{1}{N} \ln \|X_N \cdots X_2 X_1\| \geq \gamma_1$$

which together with (3.11) yields (2.4) and completes the proof of Theorem 2.2. □

5 Products with Markov dependence

As in the case of Theorem 2.2 it suffices to prove Theorem 2.3 only for the biggest singular value $s_1(\Pi_N)$. It is easy to see that the condition of the form (2.6) remains true also for the product Markov chain Ξ_n, $n \geq 0$. Hence, it follows by the large deviations result of Theorem 4.3 in [2] applied to the products $H_N = Y_N \cdots Y_2 Y_1$ of

Markov dependent random matrices that for any $\varepsilon > 0$ there exists $\kappa(\varepsilon) > 0$ and $n(\varepsilon) \geq 1$ such that

$$P_{\bar{x}}\{|\frac{1}{n}\ln \|H_n\| - \gamma_1| > \varepsilon\} \leq e^{-\kappa(\varepsilon)n} \tag{5.1}$$

for any $n \geq n(\varepsilon)$ and $\bar{x} = (x_1, ..., x_\ell) \in E^\ell$ where, recall, $P_{\bar{x}}$ is the probability conditioned on $\Xi_0 = \bar{x}$. Note that (5.1) together with the Borel–Cantelli lemma yields (2.7).

Next, we observe that (2.6) implies ϕ-mixing of the Markov chain $\xi_0, \xi_1, \xi_2, ...$ with the ϕ-dependence coefficient satisfying $\phi(n) \leq 2Re^{-\rho n}$ (see [5]). This seems to be well known (see p.p.365-366 in [11] for the case of finite Markov chains and Theorem 21.1 in [5] for a general stationary Markov chain) but we claim this for each probability P_x, $x \in \mathcal{E}$, and so for readers' convenience we will elaborate this here. For any $0 \leq m \leq n$ let $\mathcal{F}_{mn}$ be the σ-algebra generated by $\xi_m, \xi_{m+1}, ..., \xi_n$. Then the ϕ_x-dependence coefficient for $x \in \mathcal{E}$ is defined by

$$\phi_x(n) = \sup_{m \geq 0} \big\{\big|\frac{P_x(\Gamma \cap \Delta)}{P_x(\Gamma)} - P_x(\Delta)\big| : \Gamma \in \mathcal{F}_{0m},\, \Delta \in \mathcal{F}_{m+n,\infty},\, P_x(\Gamma) > 0\big\} \tag{5.2}$$

where, recall, P_x is the probability corresponding to the initial condition $\xi_0 = x$. In order to show that

$$\phi_x(n) \leq 2Re^{-\rho n} \tag{5.3}$$

when (2.6) holds true observe that it suffices to consider Γ and Δ of the form

$$\Gamma = \bigcap_{i=1}^{k} \{\xi_{m_i} \in G_i\} \text{ and } \Delta = \bigcap_{j=k+1}^{l} \{\xi_{m_j} \in G_j\}$$

where $m_1 < m_2 < ... < m_k < m_k + n \leq m_{k+1} < m_{k+2} < ... < m_l$. Set

$$f(y) = \mathbb{I}_{G_{k+1}}(y) \int_{G_{k+2}} P(m_{k+2} - m_{k+1}, y, dy_1) \int_{G_{k+3}} P(m_{k+3} - m_{k+2}, y_1, dy_2) \\ \cdots \int_{G_{l-1}} P(m_l - m_{l-1}, y_{l-k-2}, G_l),$$

where $\mathbb{I}_G$ is the indicator of G, and observe that $P_\nu(\Delta) = \int_E f(x) d\nu(x)$. Then

$$P_x(\Gamma \cap \Delta) = \int_{G_1} P(m_1, x, dy_1) P(m_2 - m_1, y_1, dy_2) \int_{G_2} P(m_3 - m_2, y_2, dy_3) \int_{G_3} \\ \cdots P(m_k - m_{k-1}, y_{k-1}, dy_k) \int_{G_k} P(m_{k+1} - m_k, y_k, dy_{k+1}) \int_{G_{k+1}} f(y_{k+1}) \\ = P_x(\Gamma) P_x(\Delta) + Q$$

where by (2.6),

$$|Q| \leq P_x(\Gamma) \sup_y \big|\int_{\mathcal{E}} P(m_{k+1} - m_k, y, dz) f(z) - \int_{\mathcal{E}} P(m_{k+1}, x, dz) f(z)\big| \\ \leq P_x(\Gamma) \sup_y \big|\int_{\mathcal{E}} P(m_{k+1} - m_k, y, dz) f(z) - \int f(z) d\nu(z)\big| \\ + P_x(\Gamma) \big|\int_{\mathcal{E}} P(m_{k+1}, x, dz) f(z) - \int f(z) d\nu(z)\big| \leq 2Re^{-\rho n} P_x(\Gamma)$$

yielding (5.3).

Next, set $r(n) = r_\varepsilon(n) = [\frac{2}{\delta(\varepsilon)} \ln n]$, where $\delta(\varepsilon) = \min(\kappa(\varepsilon), \rho)$, and observe that for large n,

$$q_{i+1}(n) \geq q_i(n + 2r(n)) \geq q_i(n + r(n)) + r(n) \quad \text{for all } i = 1, ..., \ell - 1. \tag{5.4}$$

Consider vectors $\bar{x}^{(i)} = (x_1^{(i)}, x_2^{(i)}, ..., x_m^{(i)})$, $x_j^{(i)} \in \mathcal{E}$, $i = 1, ..., \ell$ and view products

$$Q_m(\bar{x}^{(1)}, \bar{x}^{(2)}, ..., \bar{x}^{(\ell)}) = \prod_{j=1}^{m} F(x_j^{(1)}, x_j^{(2)}, ..., x_j^{(\ell)})$$

as functions of vectors $\bar{x}^{(i)}$, $i=1,...,\ell$. Introduce another function

$$\varphi_m(\bar{x}^{(1)},...,\bar{x}^{(\ell)})=\mathbb{I}_{\{|\frac{1}{m}\ln\|Q_m(\bar{x}^{(1)},...,\bar{x}^{(\ell)})\|-\gamma_1|>\varepsilon\}}$$

which takes on the values 0 and 1 only. We are going to plug in place of $\bar{x}^{(i)}$ in φ_m with $m=r(n)$ the vectors $\bar{\xi}^{(i)}=(\xi_{q_i(n)},\xi_{q_i(n+1)},...,\xi_{q_i(n+r(n)-1)})$ observing that $\bar{\xi}^{(i)}$ is $\mathcal{F}_{q_i(n),q_i(n+r(n)-1)}$-measurable and that by (5.4) there is a gap of at least $r(n)$ between the intervals $[q_i(n),\,q_i(n+r(n)-1)]$ for different $i=1,...,\ell$.

To use the above observation we will need the following result which is a particular case of Corollary 1.3.11 in [10] (see also Corollary 3.3 in [9]).

Lemma 5.1. *Let Z_i be $\wp_i$-dimensional $\mathcal{E}^{\wp_i}$-valued random vectors with a distribution μ_i, $i=1,...,k$ defined on the same probability space $(\Omega,\mathcal{F},P)$ and such that Z_i is $\mathcal{F}_{m_in_i}$-measurable where $n_{i-1}<m_i\le n_i<m_{i+1}$, $i=1,...,k$, $n_0=0$, $m_{k+1}=\infty$. Then for any bounded Borel function $h=h(x_1,...,x_k)$, $x_i\in\mathcal{E}^{\wp_i}$,*

$$\begin{aligned}&|Eh(Z_1,Z_2,...,Z_k)-\textstyle\int h(x_1,x_2,...,x_k)d\mu_1(x_1)d\mu_2(x_2)...d\mu(x_k)|\\&\qquad\le 4\sup_{x_1,...,x_k}|h(x_1,...,x_k)|\textstyle\sum_{i=2}^k\phi(m_i-n_{i-1})\end{aligned}\tag{5.5}$$

with the ϕ-dependence coefficient defined in (5.2). In particular, if $Z_1^{(1)},Z_2^{(2)},...,Z_k^{(k)}$ are independent copies of $Z_1,Z_2,...,Z_k$, respectively, then taking $h=\mathbb{I}_\Gamma$ for a Borel set $\Gamma\subset\mathcal{E}^{\wp_1+\wp_2+\cdots+\wp_k}$ it follows that

$$|P\{(Z_1,Z_2,...,Z_k)\in\Gamma\}-P\{(Z_1^{(1)},Z_2^{(2)},...,Z_k^{(k)})\in\Gamma\}|\le 4\sum_{i=2}^k\phi(m_i-n_{i-1}).\tag{5.6}$$

Now, applying (5.6) to $Z_i=\bar{\xi}^{(i)}$, $i=1,...,\ell$ and

$$\Gamma=\big\{(\bar{x}^{(1)},...,\bar{x}^{(\ell)}):\,|\frac{1}{r(n)}\ln\|Q_{r(n)}(\bar{x}^{(1)},...,\bar{x}^{(\ell)})\|-\gamma_1|>\varepsilon\big\}$$

and taking into account (5.1), (5.3) and (5.4) we obtain that for each $x\in\mathcal{E}$,

$$P_x\big\{|\frac{1}{r(n)}\ln\|X_{n+r(n)-1}\cdots X_{n+1}X_n\|-\gamma_1|>\varepsilon\big\}\le e^{-\kappa(\varepsilon)r(n)}+8R\ell n^{-2}\tag{5.7}$$

whenever $r(n)\ge n(\varepsilon)$.

The remaining part of the proof of Theorem 2.3 proceeds in the same way as in the i.i.d. case of Theorem 2.2 except for the arguments leading to (3.10), (4.12) and (4.13). Namely, we cannot use the Chebyshev inequality in order to obtain (3.7) and (4.10) since, in general, in the present situation $E_\nu\|X_n\|^\alpha$ and $E_\nu\|X_n^{-1}\|^{\alpha'}$ may be not equal to $E_\nu\|X_1\|^\alpha$ and $E_\nu\|X_1^{-1}\|^{\alpha'}$, respectively, and the latter expectations may be not equal to $E_{\bar\nu}\|Y_1\|^\alpha$ and $E_{\bar\nu}\|Y_1^{-1}\|^{\alpha'}$ where E_ν and $E_{\bar\nu}$ are the expectations corresponding to the path space probabilities P_ν and $P_{\bar\nu}$ of the Markov chains ξ_n and Ξ_n having initial distributions ν and $\bar\nu=\nu\times\cdots\nu$, respectively. But applying Lemma 5.1 in the same way as in (5.7) we obtain by the Chebyshev inequality that

$$\begin{aligned}&P_x\{\ln\|X_n\|\ge\tfrac{2}{\alpha}\ln n\}\le P_{\bar x}\{\ln\|Y_n\|\ge\tfrac{2}{\alpha}\ln n\}+8R\ell n^{-2}\\&\le n^{-2}E_{\bar x}\|Y_n\|^\alpha+8R\ell n^{-2}\le n^{-2}(\textstyle\sup_{\bar z}E_{\bar z}\|Y_1\|^\alpha+8R\ell)\quad\forall\,x\in\mathcal{E}\end{aligned}$$

since

$$E_{\bar x}\|Y_n\|^\alpha=\int_{\mathcal{E}}P_\Xi(n,\bar x,d\bar y)\|F(\bar y)\|^\alpha\le\sup_{\bar z}\int P_\Xi(\bar z,d\bar v)\|F(\bar v)\|^\alpha=\sup_{\bar z}E_{\bar z}\|Y_1\|^\alpha<\infty.$$

Similarly, for any $x \in \mathcal{E}$,

$$P_x\{\ln \|X_n^{-1}\| \geq \frac{2}{\alpha'} \ln n\} \leq (D' + 8R\ell)n^{-2}$$

where $D' = E_{\bar{x}}\|Y_1^{-1}\|^{\alpha'} < \infty$ and $\alpha' = \frac{\alpha}{d-1}$. Now, the corresponding versions of (3.10), (4.12) and (4.13) follow in the same way as in Section 3 and 4 while the arguments related to the avalanche principle remain the same. □

Remark 5.2. Since Lemma 5.1 is quite general the Markov dependence in the sequence $\xi_n,\, n \geq 0$ is needed only to rely on large deviations result from [2], and so our method will go through whenever large deviations estimates (actually, only upper bounds in the form (5.1)) for products of stationary sufficiently fast ϕ-mixing sequences of random matrices become available.

References

[1] P. Bougerol, *Comparison des exposants de Lyapunov des processus markoviens multiplicatifs*, Ann. Inst. H. Poincaré (sec. B) 24 (1988), 439–489. MR-0978021

[2] P. Bougerol, *Théorémes limite pour les systémes linéares á coefficients markovien*, Probab. Theor. Rel. Fields 78 (1988), 193–221. MR-0945109

[3] Y. Benoist and J.-F. Quint, *Random Walks on Reductive Groups*, Springer, Heidelberg, 2016. MR-3560700

[4] P. Bougerol and J. Lacroix, *Products of Random Matrices and Applications to Schrödinger Operators*, Birkhäuser, Boston, 1985. MR-0886674

[5] R.C. Bradley, *Introduction to Strong Mixing Conditions,* Kendrick Press, Heber City, 2007. MR-2325296

[6] P. Duarte and S. Klein, *Lyapunov Exponents of Linear Cocycles, Continuity via Large Deviations*, Atlantis Press, 2016. MR-3468528

[7] H. Furstenberg and H. Kesten, *Products of random matrices*, Ann. Math. Stat. 31 (1960), 457–469. MR-0121828

[8] M. Goldstein and W. Schlag, *Hölder continuity of the integrated density of states for quasi-periodic Schrödinger equation and averages of shifts of subharmonic functions*, Ann. Math. 154 (2001), 155–203. MR-1847592

[9] Y. Hafouta, *Stein's method for nonconventional sums*, arXiv 1704.01094

[10] Y. Hafouta and Y. Kifer, *Nonconventional Limit Theorems and Random Dynamics*, World Scientific, Singapore, 2018.

[11] I.A. Ibragimov and Y.V. Linnik, *Independent and Stationary Sequences of Random Variables*, Wolters-Noordhoff, Groningen, 1971. MR-0322926

[12] A.D. Virtser, *On the simplicity of the spectrum of the Lyapunov characteristic indices of a product of random matrices*, Theor. Probab. Appl. 28 (1984), 122–135. MR-0691471