Lyapunov functions for Random Walks and Strings in Random Environment*

Francis Comets Mikhail Menshikov Serguei Popov

May 20, 1998

Abstract

We study two typical examples of countable Markov chains in random environment using the Lyapunov functions method: random walk and random string in random environment. In each case we construct an explicit Lyapunov function. Investigating the behavior of this function we get the classification for recurrence, transience, ergodicity. We obtain new results for random strings in random environment, though we simply review well known results for random walks using our approach.

AMS 1991 subject classification: 60J10, 60J80

Key words and phrases: random walk, random string, Lyapunov function, random medium, disordered systems, product of random matrices, Lyapunov exponents

Short Title: Random Walks and Strings in Random Environment

*F.C is from UPRES-A CNRS 7055 “Statistique et Modèles Aléatoires”; M.M and S.P acknowledge partial support under RFFI grant No. 98–01–00653; S.P. is also thankful to FAPESP for partial support

Introduction

The theory of countable Markov chains is currently developing in several directions. One of these is the qualitative analysis of Markov chains in a countable state space with a complicated structure. In relation to the forthcoming examples we mention two models: i) Multi-dimensional Markov chains with partial linear nonhomogeneities are considered in [1, 7], ii) Markov chains with the state space being equal to all finite sequences of letters from some alphabet (the so-called *strings*) are studied in [8, 9, 15]. For these problems a martingale method (method of Lyapunov functions) was developed.

Another branch of current interest is the theory of countable Markov chains in random environment. In this paper we study two examples using Lyapunov functions. Let us briefly describe what we mean by the "random environment". A given countable, time-homogeneous Markov chain $\mathcal{L} = \{\eta_t; t \geq 0\}$ can be defined by its state space $X = \{x_i\}$ and by the collection of transition probabilities $P_{ij} = \mathsf{P}\{\eta_{t+1} = x_j \mid \eta_t = x_i\}$. Assume that on some probability space $(\Omega, \mathcal{A}, \mathbf{P})$ we are given a collection of random variables $P_{ij}(\omega)$, $i, j \in \mathbf{N}$, $\omega \in \Omega$, such that for any fixed ω (which we call the *realization* of the random environment) the numbers $P_{ij}(\omega)$ are transition probabilities. Together with the state space X, these transition probabilities define a Markov chain $\mathcal{L}(\omega)$. One can view the Markov chain as the motion of a particle in a space-inhomogeneous medium $(P_{i,\cdot}(\omega))_i$. Of particular interest is when the field $(P_{i,\cdot}(\omega))_i$, viewed as a random field, is stationary in space; then the medium possesses the space homogeneity property at a statistical level. Some natural questions are: what is the probability (with respect to probability measure $\mathbf{P}$) for the Markov chain $\mathcal{L}(\omega)$ of being recurrent (resp., transient, ergodic)?

In the present paper we consider random strings in a random environment, which includes random walks in a random environment as a particular case. The main idea of our approach is the following: for a given $\omega \in \Omega$ we construct a Lyapunov function $f(x) = f(x; \omega)$ for the Markov chain $\mathcal{L}(\omega)$. It turns out that this function is a *spatially homogeneous* random field. Hence we can investigate its asymptotic behavior and, using criteria for countable Markov chains, we obtain the qualitative classification of the Markov chain $\mathcal{L}(\omega)$ for typical ω's .

The method can be applied to other models, like the branching random walk in random environment [4] where the Lyapunov functions are much easier analyzed than the process itself. Note that constructing a Lyapunov function is not an easy matter; this is the main restriction to the method.

In Section 2 we briefly study the one-dimensional random walk in random environment. Classification reduces to the investigation of the products of independent random variables. The results are well-known, but the interest there is rather in the extreme simplicity of the arguments which makes the approach more transparent. The model of the random strings in random environment, studied in Section 3, leads to the study of the products of i.i.d. random matrices. The results there are new, and they are formulated in terms of Lyapunov exponents.

The proofs in Sections 2 and 3 have many common points, so the rather elementary Section 2 is useful for a better understanding of the techniques and the ideas in Section 3.

1 Criteria for countable Markov chains

In this section we recall three martingale criteria for countable irreducible Markov chains. We state them in a rather simplified form which is sufficient for our purpose. More general criteria can be found in [7], [16] and [1] .

Let η_t be a discrete time-homogeneous Markov chain with a countable state space X, and a point $0 \in X$.

Proposition 1.1 *Suppose that there exists a function $f : X \mapsto \mathbf{R}$ which is bounded and non-constant, such that for all $x \neq 0$*

$$\mathsf{E}(f(\eta_{t+1}) - f(\eta_t) \mid \eta_t = x) = 0. \tag{1}$$

Then the Markov chain η_t is transient (i.e. non-recurrent).

This is a consequence of Theorem 2.2.2 of [7]. The next propositions are just simplified forms of Theorems 2.2.1 and 2.2.3 of [7] respectively.

Proposition 1.2 *If there exists a function $f : X \mapsto \mathbf{R}$ such that $f(x) \to +\infty$ as $x \to \infty$ and for all $x \neq 0$*

$$\mathsf{E}(f(\eta_{t+1}) - f(\eta_t) \mid \eta_t = x) \leq 0. \tag{2}$$

then the Markov chain η_t is recurrent.

Proposition 1.3 (Foster) *Assume that there exists a function $f : X \mapsto \mathbf{R}$, $f(x) \geq 0$, and $\delta > 0$ such that for all $x \neq 0$*

$$\mathsf{E}(f(\eta_{t+1}) - f(\eta_t) \mid \eta_t = x) \leq -\delta; \tag{3}$$

and $\mathsf{E}(f(\eta_{t+1}) \mid \eta_t = 0) < \infty$. Then the Markov chain η_t is ergodic.

2 Random walk on Z_+ in a random environment

Let us recall the model which was first considered by M.V. Kozlov [13] and F. Solomon [20], and investigated by many authors; see [18] for a review.

Let $\{\xi_i\}_{i=1,2,\ldots}$ be a sequence of i.i.d. random variables on Ω with values in $[0,1]$. Suppose also that $\mathbf{P}\{\xi_1 = 0\} = \mathbf{P}\{\xi_1 = 1\} = 0$. Then, for fixed environment, i.e. for a given realization of the sequence of independent and identically distributed (i.i.d.) random variables $\{\xi_i\}$, consider a Markov chain η_t on $\mathbf{Z}_+$ defined as follows: $\eta_0 = 0$,

$$\begin{aligned} p_n &= \mathsf{P}\{\eta_{t+1} = n-1 \mid \eta_t = n\} = \xi_n, \\ q_n &= \mathsf{P}\{\eta_{t+1} = n+1 \mid \eta_t = n\} = 1 - \xi_n, \ n = 1, 2 \ldots, \end{aligned}$$

and $\mathsf{P}\{\eta_{t+1} = 1 \mid \eta_t = 0\} = 1$. we will use $\mathbf{P}, \mathbf{E}$ to denote probability and expectation for the random environment ω, keeping the notations P, E for the Markov chain η_t itself.

Denote $\zeta_n = \log(p_n/q_n)$. The next theorem is due to Solomon [20].

Theorem 2.1 *Assume* $\mathbf{E}|\zeta_1| < \infty$*.*

- *If* $\mathbf{E}\zeta_1 < 0$*, then the random walk is transient for almost all* ω *(for almost all environments).*
- *If* $\mathbf{E}\zeta_1 \geq 0$*, then the random walk is recurrent for almost all* ω*.*
- *Moreover, if* $\mathbf{E}\zeta_1 > 0$*, then the random walk is ergodic for almost all* ω*.*

Sketch of proof. First, we prove the recurrence and the transience. Let us try to construct a function $f(x)$ satisfying (1) for fixed ω. Denote $\Delta_i = f(i) - f(i-1)$ and let $\Delta_0 = 1$, so $f(n) = \sum_{i=1}^{n} \Delta_i$, $f(0) = 0$. We have

$$\mathsf{E}(f(\eta_{t+1}) - f(\eta_t) \mid \eta_t = x) = -p_x\Delta_x + q_x\Delta_{x+1}, \tag{4}$$

so

$$\Delta_{x+1} = \frac{p_x}{q_x}\Delta_x = \exp\Big\{\sum_{i=1}^{x} \zeta_i\Big\}. \tag{5}$$

It is clear that the function $f(x)$ is positive and that if $\mathbf{E}\zeta_1 \geq 0$, then $f(x) \to +\infty$ for almost all ω , and if $\mathbf{E}\zeta_1 < 0$, then $f(x)$ is bounded. Applying Propositions 1.1 and 1.2 we complete the proof of the first two statements of Theorem 2.1.

To prove ergodicity, we construct the Lyapunov function $f(s)$ in the reverse way. Let $f(x) = \sum_{i=1}^{x} \Delta_i$, where

$$\Delta_i = \frac{1}{p_i} + \frac{q_i}{p_i p_{i+1}} + \frac{q_i q_{i+1}}{p_i p_{i+1} p_{i+2}} + \frac{q_i q_{i+1} q_{i+2}}{p_i p_{i+1} p_{i+2} p_{i+3}} + \cdots \tag{6}$$

Then one can easily check that if $\mathbf{E}\zeta_1 > 0$,

i) the series in the right-hand side of (6) converge, so the function $f(x)$ is well defined;

ii) $f(x)$ satisfies (3) with $\delta = 1$.

Applying now Proposition 1.3, we complete the proof of the last claim of Theorem 2.1. □

In the case $\mathbf{E}\zeta_1 = 0$ the random walk is null recurrent . As a simple illustration of Lyapunov functions techniques we prove a result (not optimal, compare to [19]) about the speed of escape of the walk.

Theorem 2.2 *Let $\mathbf{E}\zeta_1 = 0$, and $0 < \mathbf{E}\zeta_1^2 < \infty$. Then for any integer $k \geq 1$ and for any $\varepsilon > 0$ we have*

$$\eta_t / (\log t \log_2 t \ldots \log_k^{1+\varepsilon} t)^2 \longrightarrow 0 \tag{7}$$

almost surely as $t \to \infty$, with $\log_1 t := \log t$, $\log_{m+1} t = \log(\log_m t)$, $m \geq 1$. Also, for any $\varepsilon > 0$ and for any $p > 0$ we have

$$\frac{\eta_t}{\log^{2+\varepsilon} t} \longrightarrow 0 \tag{8}$$

in L^p, as $t \to \infty$.

In the above statements, "almost surely" means "for $\mathbf{P}$-almost every environment it holds P-a.s." and "convergence in L^p" stands for "convergence in $L^p(\mathsf{P})$ for $\mathbf{P}$-a.e. ω". P. Deheuvels and P. Revész proved (7) in [6], Theorem 4, with a completely different approach, but ours is much shorter. After finishing this paper we learnt that Y. Hu and Z. Shi [10] proved the exact upper limit result, $\limsup_t \eta_t / (\log^2 t \ \log_3 t) = 8/(\pi^2\sigma^2)$ a.s.

Proof. We use again the Lyapunov function

$$f(x) = \sum_{i=1}^{x} \exp\Big\{\sum_{j=1}^{i-1} \zeta_j\Big\}.$$

As shown in the proof of Theorem 2.1, the process $f(\eta_t)$ is a martingale except at 0 (this means that (1) holds for all $x \neq 0$). Thus we have that for some constant C' (depending on the environment) $\mathsf{E}(f(\eta_t) - f(\eta_{t-1})) \leq C'$ for all $t \geq 1$ (one can take $C' = p_1(\omega)/q_1(\omega)$). Therefore, for all $t \geq 1$ we have

$$\mathsf{E} f(\eta_t) \leq C' t \tag{9}$$

At this point we need a lower bound for the function $f(x)$. We use the so-called "inverse iterated logarithm law", recalling a result from E. Csáki:

Lemma 2.1 *([5], Th. 3.1). Let $\{\zeta_i\}_{i=1,2,\ldots}$ be a sequence of i.i.d. random variables, and let $\mathbf{E}\zeta_1 = 0$, $0 < \mathbf{E}\zeta_1^2 < \infty$. Let $a : [y_0, \infty) \to (0, \infty)$ be decreasing and such that $a(y)y^{1/2}$ increases, and $\int_1^\infty \frac{a(y)}{y} dy < \infty$. Then*

$$\lim_n \frac{\max_{i\le n} \sum_{j=1}^i \zeta_j}{a(n)\sqrt{n}} = +\infty \quad a.s. \tag{10}$$

It follows that for any such a there exists a positive constant $C = C(a,\omega)$ such that

$$f(x) \ge g(x) := C \exp\{x^{\frac{1}{2}} a(x)\} \tag{11}$$

for all x. We take for the moment $a(x) = (x+K)^{1/2-\varepsilon_1} x^{-1/2}$ and, for $p > 0$ we denote also $g_p(x) = C\exp\{(x+K)^{\frac{1}{p}(\frac{1}{2}-\varepsilon_1)}\}$, where K is chosen in such a way that the function $g_p(x)$ be convex on $[0,\infty)$. Using (9), (11), and Jensen inequality, we have

$$C't \ge \mathsf{E} f(\eta_t) \ge \mathsf{E} g(\eta_t) = \mathsf{E} g_p(\eta_t^p) \ge g_p(\mathsf{E}\eta_t^p),$$

so

$$(\mathsf{E}\eta_t^p)^{1/p} \le (g_p^{-1}(C't))^{1/p} \le C_3 \log^{2+\varepsilon_2} t,$$

for some C_3 and $\varepsilon_2 = (4\varepsilon_1)/(1-2\varepsilon_1)$, thus proving the convergence in L^p.

Let us prove a.s. convergence. Using Chebyshev inequality and (9), we have

$$\mathsf{P}\{f(\eta_t) > t^3\} \le \frac{\mathsf{E} f(\eta_t)}{t^3} \le C' \frac{1}{t^2}.$$

Taking now $a(y) = (\log y \log_2 y \dots \log_{k-1}^{1+\varepsilon} y)^{-1}$ in (11) we get

$$\begin{aligned} \mathsf{P}\{f(\eta_t) > t^3\} &\ge \mathsf{P}\{g(\eta_t) > t^3\} = \mathsf{P}\{\eta_t > g^{-1}(t^3)\} \\ &\ge \mathsf{P}\{\eta_t > C_4 (\log t \log_2 t \dots \log_k^{1+\varepsilon} t)^2\} \end{aligned}$$

for some $C_4 > 0$. Using Borel-Cantelli lemma, we prove the a.s. convergence, thus completing the proof of Theorem 2.2. □

3 Strings in a random environment

3.1 Dynamics of the string

Consider a finite alphabet $\mathcal{S} = \{1,\dots,k\}$. A string is just a finite sequence of symbols from $\mathcal{S}$. We write $|s|$ for the length of the string s, i.e. if $s = s_1 \dots s_n$, $s_i \in \mathcal{S}$, then $|s| = n$.

Consider a time-homogeneous Markov chain with the state space equal to the set of all finite strings. We describe the transition matrix of this Markov chain as follows: let $s = s_1 \dots s_n$, and $s_n = i \in \mathcal{S}$. Then

- we erase the rightmost symbol of s with probability $r_i^{(n)}$;
- we substitute the rightmost symbol i by j with probability $q_{ij}^{(n)}$;
- and we add the symbol j to the right end of the string with probability $p_{ij}^{(n)}$.

Of course we assume that for all i and for all $n = 1, 2, \dots$

$$r_i^{(n)} + \sum_j q_{ij}^{(n)} + \sum_j p_{ij}^{(n)} = 1. \tag{12}$$

These parameters do not define the evolution when the string is empty (its length equals 0), but we simply assume that the jumps from the empty string are somehow defined and can only occur to strings of length 1. Clearly, these "boundary conditions" do not affect the asymptotic behaviour of the string. So we see that the process is completely defined by the collection of numbers $\{r_i^{(n)}, q_{ij}^{(n)}, p_{ij}^{(n)}\}$, $n = 1, 2, \dots$, $i = 1, \dots, k$, and $j = 1, \dots, k$.

For the case when the quantities $r_i^{(n)}$, $q_{ij}^{(n)}$ and $p_{ij}^{(n)}$ do not depend on n, this model was investigated by A.S. Gajrat, V.A. Malyshev, M.V. Menshikov and K.D. Pelih in [8]. In fact, they investigated even more general model, when the maximal value of jump may be greater than one. All these are models for LIFO (last in, first out) queuing systems [8], and they are random walks on trees. In the case $k = 1$ we recover the random walk from Section 2. Note that the length $|\eta_t|$ of the string is *not* a Markov chain itself. We consider here only the *one-sided* evolution of the string; Two-sided *homogeneous* strings were studied by A.S. Gajrat, V.A. Malyshev and A.A. Zamyatin in [9].

Let's now describe the random environment. First we define $\xi_n = (r_i^{(n)}, q_{ij}^{(n)}, p_{ij}^{(n)})_{i,j}$, and we will say that a vector ξ is non-negative ($\xi \geq 0$) if all its components are non-negative. We will assume that the vectors $\xi_n = \xi_n(\omega), n \geq 0$, are i.i.d. random vectors on $(\Omega, \mathcal{A}, \mathbf{P})$, with non-negative values satisfying (12). This defines the

Markov chain describing the evolution of the string , for each fixed environment $\omega = \{\xi_1, \xi_2, \ldots\}$.

3.2 Lyapunov exponents and products of random matrices

In this section we review some properties of products of random matrices that we need below.

Let $\langle \cdot, \cdot \rangle$ be the standard scalar product in $\mathbf{R}^k$. For $x \in \mathbf{R}^k$ define its norm by $\|x\| = \langle x, x \rangle^{1/2}$. The transpose of matrix A is denoted by A^*. Euclidean (operator) norm of a $k \times k$ real matrix A can be defined by any of the following, equivalent formulas:

$$\begin{aligned} \|A\| &:= \sup_{\|x\|=1} \|Ax\| \\ &= [\text{largest eigenvalue of } A^*A]^{1/2}. \end{aligned}$$

Consider a sequence of i.i.d. random matrices A_n. We assume that the matrices A_n satisfy the following condition:

Condition A. $\mathbf{E} \log^+ \|A\| < \infty$, where $\log^+ x = \max\{\log x, 0\}$.

Let $(A_1(\omega), A_2(\omega), \ldots)$ be a realization of a sequence of i.i.d. random matrices. Let $a_1(n) \geq a_2(n) \geq \cdots \geq a_k(n) \geq 0$ be the square roots of the (random) eigenvalues of $(A_n \ldots A_1)^*(A_n \ldots A_1)$. Then the following limit exists for almost all ω (and it is the same for almost all ω):

$$\lim_{n\to\infty} \frac{1}{n} \log a_j(n) = \gamma_j(A) \tag{13}$$

for $j = 1, \ldots, k$ (see Proposition 5.6 of [3], A does not need to be invertible). The numbers $\gamma_j, -\infty \leq \gamma_k \leq \cdots \leq \gamma_1 < \infty$ are called the *Lyapunov exponents* of the sequence of random matrices $\{A_n\}$. In particular,

$$\gamma_1(A) = \lim_{n\to\infty} \frac{1}{n} \log a_1(n) = \lim_{n\to\infty} \frac{1}{n} \log \|A_n \ldots A_1\| \qquad \text{a.s.} \tag{14}$$

is the *top Lyapunov exponent*.

The following simple lemma establishes relations between Lyapunov exponents of A and A^{-1}.

Lemma 3.1 *Assume that A is a.s. invertible, and that both A and A^{-1} satisfy Condition A. Then for $j = 1, \dots, k$*

$$\gamma_j(A^{-1}) = -\gamma_{k-j+1}(A). \tag{15}$$

Proof. Let $a_1(n, A^{-1}), \dots, a_k(n, A^{-1})$ be the square roots of eigenvalues of

$$(A_n^{-1} \dots A_1^{-1})^*(A_n^{-1} \dots A_1^{-1}) = [(A_1 \dots A_n)(A_1 \dots A_n)^*]^{-1}.$$

Since UV has the same eigenvalues as VU, $(a_1(n, A^{-1}))^{-1}, \dots, (a_k(n, A^{-1}))^{-1}$ are the square roots of eigenvalues of $(A_1 \dots A_n)^*(A_1 \dots A_n)$. But $A_1 \dots A_n$ has the same distribution law as $A_n \dots A_1$. So for all $j = 1, \dots, k$ we have $(a_j(n, A^{-1}))^{-1} = a_{k-j+1}(n, A)$ in law, and consequently, $\gamma_j(A^{-1}) = -\gamma_{k-j+1}(A)$. $\square$

We will also need the following theorem [17]

Theorem 3.1 (Oseledec's Multiplicative Ergodic Theorem) *Let A_n, $n = 1, 2, \dots$ be a stationary ergodic sequence of $k \times k$ real matrices on the probability space $(\Omega, \mathcal{B}, m)$ and suppose that $\mathsf{E} \log^+ \|A_1\| < \infty$. Let $\gamma_1 \geq \gamma_2 \geq \cdots \geq \gamma_k$ be the Lyapunov exponents of the sequence A_n. Consider the strictly increasing nonrandom sequence of integers $1 = i_1 < i_2 < \cdots < i_p < i_{p+1} = k + 1$ such that $\gamma_{i_{q-1}} > \gamma_{i_q}$, if $q = 2, 3, \dots, p$, and $\gamma_i = \gamma_j$, if $i_q \leq i, j < i_{q+1}$ (the i_q's mark the points of decrease of the γ_i). Then for almost all $\omega \in \Omega$:*

- *for every $v \in \mathbf{R}^k$*

  $$\lim_{n\to\infty} \frac{1}{n} \log \|A_n \dots A_1 v\|$$

  exists or is $-\infty$;

- *for $q \leq p$*

  $$V(q, \omega) = \left\{ v \in \mathbf{R}^k : \lim_{n\to\infty} \frac{1}{n} \log \|A_n \dots A_1 v\| \leq \gamma_{i_q} \right\}$$

  is a random linear subspace of $\mathbf{R}^k$ with dimension $k - i_q + 1$ ($V(1, \omega) = \mathbf{R}^k$);

- *for $q \geq 1$, $v \in V(q, \omega) \ominus V(q + 1, \omega)$ implies that*

  $$\lim_{n\to\infty} \frac{1}{n} \log \|A_n \dots A_1 v\| = \gamma_{i_q}.$$

3.3 Main results

Introduce two sequences of random $k \times k$ matrices $\{B_n\}$ and $\{D_n\}$, $B_n = (p_{ij}^{(n)})_{i,j=1,\dots,k}$ and $D_n = (d_{ij}^{(n)})_{i,j=1,\dots,k}$, where $d_{ij}^{(n)} = -q_{ij}^{(n)}$ for $i \neq j$ and

$$d_{ii}^{(n)} = r_i^{(n)} + \sum_{j:j\neq i} q_{ij}^{(n)},$$

with $p_{ij}^{(n)}$, $q_{ij}^{(n)}$ and $r_i^{(n)}$ defined in Section 3.1. Important ingredients are the sequence of i.i.d. random matrices A_n,

$$A_n = D_n^{-1} B_n$$

and its Lyapunov exponents $\gamma_1, \dots, \gamma_k$. Besides Condition A we will consider

Condition D. $\mathbf{E} \log(1/r_i) < \infty, i = 1, \dots, k$

The interest of this condition appears in the

Proposition 3.1 *If Condition D holds then the matrix D is a.s. invertible, the Condition A holds and* $\mathbf{E} \log^+ \|D^{-1}\| < \infty$

Proof. With probability one $r_i > 0, i = 1, \dots, k$, and then any non-zero vector v has a non-zero image Dv: if $|v_i| = \max(|v_j|; j = 1, \dots, k)$ we may assume $v_i > 0$ without loss of generality, and we have $(Dv)_i = r_i v_i + \sum_{j\neq i} q_{ij}(v_i - v_j) \geq r_i v_i > 0$. Moreover

$$\|Dv\| \geq (Dv)_i \geq r_i v_i \geq k^{-1/2} r_i \|v\|$$

which implies $\|D^{-1}\| \leq k^{1/2} \max_i(1/r_i)$.

Therefore $\mathbf{E} \log^+ \|D^{-1}\| \leq \sum_i \mathbf{E} \log(1/r_i) < \infty$. The last claim follows easily from this, from $\|A\| = \|D^{-1}B\| \leq \|D^{-1}\| \|B\|$ and boundedness of B. □

Now we are ready to formulate the main results. In the next one we use two different sets of assumptions, each one being of interest for applications.

Theorem 3.2 *Assume that Condition A holds for the matrix A_n, and that*

- *either B_n is a.s. invertible and Condition A holds for A_n^{-1}*
- *or* $\mathbf{E} \log(1/p_{ii}) < \infty, i = 1, \dots, k$

If $\gamma_1 > 0$, then the Markov chain, describing the evolution of the string, is transient (for almost all ω).

Theorem 3.3 *Assume Condition D for A_n. If $\gamma_1 < 0$, then the process is a.s. ergodic.*

Theorem 3.4 *Let $\gamma_1 = 0$. In addition to Condition D assume that A_1 is a.s. invertible, and that no finite union of proper subspaces of $\mathbf{R}^k$ is a.s stable by A_1. Then the process is a.s. recurrent.*

3.4 Proof of Theorem 3.2

In this section we shall make use of Proposition 1.1 again.

Proof of Theorem 3.2. We start with the case when B (and then A) is a.s. invertible and when $\mathbf{E}\log^+\|A^{-1}\| < \infty$. All that we need is to construct a function $f(s)$ as in Proposition 1.1. We show that there exists a sequence of column vectors $v_n = (v_n^i)_{i=1,\dots,k}$ such that the function $f(s)$ can be defined as

$$f(s) = \sum_{j=1}^{|s|} v_j^{s_j}. \tag{16}$$

Indeed, for $f(s)$ defined by (16) we have

$$\begin{aligned} &\mathsf{E}(f(\eta_{t+1}) - f(\eta_t) \mid \eta_t = s, |s| = n, s_n = i) \\ &\quad = \quad -r_i^{(n)} v_n^i + \sum_{j=1}^{k} q_{ij}^{(n)}(-v_n^i + v_n^j) + \sum_{j=1}^{k} p_{ij}^{(n)} v_{n+1}^j, \end{aligned}$$

so

$$\mathsf{E}(f(\eta_{t+1}) - f(\eta_t) \mid \eta_t = s, |s| = n, s_n = \cdot) = -D_n v_n + B_n v_{n+1}. \tag{17}$$

Thus, since $f(s)$ must satisfy (1), we have the following relation between v_n and v_{n+1} (compare this with (4) and (5); A_n^{-1} is an analog of p_n/q_n):

$$v_{n+1} = A_n^{-1} v_n, \tag{18}$$

where the matrices A_n, B_n, and D_n were introduced in Section 3.3. From (18) it follows that

$$v_{n+1} = A_n^{-1} \dots A_1^{-1} v_0$$

for some vector v_0. According to Theorem 3.1, for almost every environment ω there exists a random vector $v_0 = v_0(\omega) \neq 0$ such that

$$\lim_{n\to\infty} \frac{1}{n} \log \|A_n^{-1} \dots A_1^{-1} v_0\| = \gamma_k(A^{-1}) = -\gamma_1. \tag{19}$$

From the last relation it follows that there exists two constants $C_{1,2} = C_{1,2}(\omega) > 0$ such that $\|v_n\| \leq C_1 \exp\{-C_2 n\}$ for all $n \geq 0$. Taking (16) into account, we get

$$|f(s)| \leq \frac{C_1 \exp\{-C_2\}}{1 - \exp\{-C_2\}}$$

So this function satisfies all the conditions of Proposition 1.1 and Theorem 3.2 is proved under the first set of assumptions.

Under the second assumption we compare the random string of interest, say $(\eta_t)_{t=0,1,\dots}$, with a family of new random strings, $(\eta_t^\varepsilon)_{t=0,1,\dots}$ indexed by $\varepsilon \in (0,1)$, with transition probabilities

$$r_i^{(n,\varepsilon)} = r_i^n, \quad q_{ij}^{(n,\varepsilon)} = q_{ij}^{(n)}, \qquad p_{ij}^{(n,\varepsilon)} = p_{ij}^{(n)} - \delta_{ij}\varepsilon \min_i p_{ii}^{(n)},$$

where δ_{ij} is the Kronecker delta (so $p_{ij}^{(n,\varepsilon)} \in (0, p_{ij}^{(n)}])$, and the new string stays at the same place with probability $\sum_j (p_{ij}^{(n)} - p_{ij}^{(n,\varepsilon)}) = \varepsilon \min_i p_{ii}^{(n)}$.

Clearly the new string $\eta_\cdot^\varepsilon$ is "more recurrent" than $\eta_\cdot$ for $\varepsilon > 0$, and we simply need to prove that $\eta_\cdot^\varepsilon$ is transient for some ε. Let $\varphi(\lambda) = |\det(B - \lambda I)|$ be the absolute value of the characteristic polynomial of B. Recalling that $\|(B - \lambda I)^{-1}\|^2$ is the largest eigenvalue of $(B^* - \lambda I)^{-1}(B - \lambda I)^{-1}$ we see that

$$\begin{aligned} \varphi(\lambda) &= [\det(B - \lambda I)(B^* - \lambda I)]^{1/2} \\ &\leq \|(B - \lambda I)^{-1}\|^{-1} \|(B - \lambda I)\|^{k-1} \end{aligned}$$

On the other hand, since φ is (the absolute value of) a polynomial with coefficients smooth in B

$$\int_0^1 \log^+(1/\varphi(\lambda)) d\lambda \leq C(\|B\|) < \infty$$

Corresponding to the new string we define $B^\varepsilon = B - \varepsilon \min_i(p_{ii})I$. From the last two estimates we get

$$\begin{aligned}
\mathbf{E}\int_0^1 \log^+ \|(B^\varepsilon)^{-1}\| d\varepsilon &\leq C' + \mathbf{E}\int_0^1 \log^+ \left(1/\varphi(\varepsilon \min_i p_{ii})\right) d\varepsilon \\
&\leq C' + \mathbf{E}(\min_i p_{ii})^{-1} \int_0^1 \log^+ (1/\varphi(\lambda))\, d\lambda \\
&\leq C' + C(1)\sum_i \mathbf{E}(p_{ii}^{-1}) < \infty
\end{aligned}$$

Hence for Lebesgue-a.e. ε, the matrix A_n^ε is $\tilde{Q}$-a.s. invertible, and $\mathbf{E}\log^+ \|(B_n^\varepsilon)^{-1}\|$, $\mathbf{E}\log^+ \|(A_n^\varepsilon)^{-1}\|$ are finite.

Also, by continuity of the top Lyapunov exponent in a suitable weak topology, $\gamma_1(A^\varepsilon) \to \gamma_1(A) > 0$ as $\varepsilon \to 0$ (the reader is referred to Corollary III.1.2 in [14]; it is not difficult to check that the proof therein extends to our case, using the following remark: the Furstenberg formula III.1.1 for γ_1 is valid for positive matrices with any invariant probability measure concentrated on $P_+^{d-1} = \{\dot{t} \in P^{d-1} : \exists t \in \mathbf{R}^d, \dot{t} = t/\|t\|, t_i > 0,\ i = 1, \ldots, k\}$).

Choosing ε small enough the first set of assumptions is satisfied and we derive from the previous proof that the string $\eta^\varepsilon_.$ is transient, and therefore $\eta_.$ has the same property. $\square$

3.5 Proof of Theorem 3.3

We shall use Proposition 1.3 to prove ergodicity. Also we need one supplementary fact.

Lemma 3.2 *Matrix D_1^{-1} is almost surely non-negative.*

Proof. We need to prove that any vector w with at least a negative component has image Dw with a negative component. Assume first

$$d_{ii} > -\sum_{j:j\neq i} d_{ij},\ i = 1, \ldots, k \tag{20}$$

In this case the statement is easily checked: $(Dw)_i < 0$ if $w_i = \min_j w_j < 0$. In the general case D is a limit of matrices satisfying (20); their inverse is non-negative, and so is D^{-1} by continuity. $\square$

Proof of Theorem 3.3. In contrast to the previous section, here we explicitly construct the function $f(s)$. Denote by I the identity matrix of order k, and by $\mathbf{1}$ the column vector with all coordinates being equal to 1. Then we again define $f(s)$ by the formula (16), where

$$\begin{aligned} v_n &= D_n^{-1}\mathbf{1} + \sum_{m:m\geq n} A_n \ldots A_m D_{m+1}^{-1}\mathbf{1} \\ &= (D_n^{-1} + A_n D_{n+1}^{-1} + A_n A_{n+1} D_{n+2}^{-1} + \cdots)\mathbf{1} \end{aligned}$$

(compare this formula with (6)). To prove that v_n is finite we first notice that the property

$$\mathbf{E}\log^+\|D^{-1}\| = \int_0^\infty \mathbf{P}\left(\log\|D^{-1}\| > u\right) du < \infty$$

implies that for every $C > 0$

$$\sum_{m>0} \mathbf{P}\left(C^{-1}\log(\|D_m^{-1}\| > m\right) < \infty$$

and from Borel-Cantelli lemma, that

$$\|D_m^{-1}\| < \exp mC$$

for m large enough. Also, since $\gamma_1 < 0$, one can easily get by using (14) and choosing $C < -\gamma_1$ that

$$\|v_n\| \leq (\|D_n^{-1}\| + \|A_n\|\,\|D_{n+1}^{-1}\| + \|A_n A_{n+1}\|\,\|D_{n+2}^{-1}\| + \cdots)\|\mathbf{1}\| < \infty \qquad \text{a.s.}$$

for all n. With other respects we have $v_n \geq 0, v_n \neq 0$ according to Lemma 3.2. According to (17) and keeping in mind that $A_n = D_n^{-1}B_n$ we have

$$\begin{aligned} &\mathsf{E}(f(\eta_{t+1}) - f(\eta_t) \mid \eta_t = s, |s| = n, s_n = \cdot) \\ &= (-I - D_n A_n D_{n+1}^{-1} - D_n A_n A_{n+1} D_{n+2}^{-1} - \cdots)\mathbf{1} \\ &\quad + (B_n D_{n+1}^{-1} + B_n A_{n+1} D_{n+2}^{-1} + B_n A_{n+1} A_{n+2} D_{n+3}^{-1} + \cdots)\mathbf{1} \\ &= -I\mathbf{1} - D_n(A_n D_{n+1}^{-1} - A_n A_{n+1} D_{n+2}^{-1} - \cdots)\mathbf{1} \\ &\quad + D_n(A_n D_{n+1}^{-1} + A_n A_{n+1} D_{n+2}^{-1} + \cdots)\mathbf{1} \\ &= -\mathbf{1}. \end{aligned}$$

Therefore the function $f(s)$ satisfies all the conditions of Proposition 1.3 and Theorem 3.3 is proved. □

3.6 Proof of Theorem 3.4

From a theorem in [2] the assumptions imply

$$\liminf_{n\to\infty} \|A_1 \dots A_n\| = 0 \quad a.s.$$

and therefore

$$B_1 A_2 \dots A_n \mathbf{1} \leq D_1 \mathbf{1} \quad \text{for some finite} \quad n \tag{21}$$

Define recursively the random times $\tau_0 = 1$, and

$$\tau_{k+1} = \min\{n > \tau_k : \ -D_{\tau_k}\mathbf{1} + B_{\tau_k} A_{\tau_k+1} \dots A_n \mathbf{1} \leq 0 \ \}$$

for $k = 0, 1, \dots$ From (21) the times τ_k are a.s. finite. We still define $f(s)$ by (16) with $v_{\tau_k} = \mathbf{1}$ and

$$v_n = A_n A_{n+1} \dots A_{\tau_{k+1}} \mathbf{1} \quad , \quad \tau_k < n \leq \tau_{k+1} \tag{22}$$

Hence we have $v_n = A_n v_{n+1}$ when $\tau_k < n < \tau_{k+1}$, $k = 0, 1, \dots$, and

$$-D_{\tau_k} v_{\tau_k} + B_{\tau_k} v_{\tau_k+1} = -D_{\tau_k}\mathbf{1} + B_{\tau_k} A_{\tau_k+1} \dots A_{\tau_{k+1}} \mathbf{1} \leq 0$$

by definition of τ_{k+1}. From (17) this means that $f(\eta_t)$ is a supermartingale except at $x = 0$, see (2). From Lemma 3.2 and from the definition (22) the vector v_n is non-negative, and therefore $f(\cdot) \geq 0$. Finally, since the τ_k's are a.s. finite and $v_{\tau_k} = \mathbf{1}$, $f(x) \to \infty$ as (the length of) $x \to \infty$. Proposition 1.2 applies, and the random string is recurrent. □

References

[1] S. Aspandiiarov, R. Iasnogorodski and M.V. Menshikov (1996) Passage-time moments for non-negative stochastic processes and an application to reflected random walk in a quadrant. *Ann. Prob.* **24** (2), 932–960.

[2] P. Bougerol (1986) Oscillation de produits de matrices aléatoires dont l'exposant de Lyapunov est nul. In: *Lyapunov exponents* (L. Arnold and V. Wihstutz, ed.) *Lect. Notes Math.* **1186**, 27–36, Springer-Verlag, Berlin.

[3] P. Bougerol and J. Lacroix (1985) *Products of Random Matrices with Application to Schrödinger Operators.* Birkhäuser Boston.

[4] F. Comets, M. Menshikov, S. Popov (1997) *One-dimensional Branching Random Walk in a Random Environment: a Classification.* Preprint

[5] E. Csáki (1978) On the lower limits of maxima and minima of Wiener process and partial sums. *Z. Wahrsch. verw. Geb.* **43**, 205–221.

[6] P. Deheuvels and P. Revész (1986) Simple random walk on the line in random environment. *Probab. Th. Relat. Fields* **72**, 215–230.

[7] G. Fayolle, V.A. Malyshev and M.V. Menshikov (1995) *Topics in the Constructive Theory of Countable Markov Chains.* Cambridge University Press.

[8] A.S. Gajrat, V.A. Malyshev, M.V. Menshikov and K.D. Pelih (1995) Classification of Markov chains describing the evolution of a string of characters. *Uspehi Mat. Nauk* **50** No. 2 (302), 5–24 (in Russian).

[9] A.S. Gajrat, V.A. Malyshev and A.A. Zamyatin (1995) Two-sided evolution of a random chain. *Markov Processes Relat. Fields* **1**, 281–316.

[10] Y. Hu, Z. Shi (1997) The limits of Sinai's simple random walk in random environment. Preprint Univ. Paris VI.

[11] H. Kesten, M.V. Kozlov and F. Spitzer (1975) A limit law for random walk in a random environment. *Compositio Mathematica* **30**, Fasc. 2, 145–168.

[12] E.S. Key (1984) Recurrence and transience criteria for random walk in a random environment. *Ann. Prob.* **12**, No. 2, 529–560.

[13] M.V. Kozlov (1973) Random walk in one dimensional random medium. *Theory Probab. Appl.* **18**, 387–388.

[14] F. Ledrappier (1984) Quelques propriétés des exposants caractéristiques. *Lect. Notes Math.* **1097**, Springer, Berlin.

[15] V.A. Malyshev (1997) Interacting strings. *Uspehi Mat. Nauk* **52** 2(314), 59–86 (in Russian).

[16] S.P. Meyn, R. Tweedie (1993) *Markov chains and stochastic stability*, Springer-Verlag, London.

[17] I.V. Oseledec (1968) A multiplicative ergodic theorem. Lyapunov characteristic numbers for dynamical systems. *Trans. Moscow Math. Soc.* **19**, 197–231.

[18] P. Revesz (1990) *Random walk in random and nonrandom environments.* World Scientific, Teaneck NJ.

[19] Ya.G. Sinaï (1982) The limiting behavior of one-dimensional random walk in random medium. *Theory Probab. Appl.* **27**, 256–268.

[20] F. Solomon (1972) Random walks in a random environment. Ph.D. Thesis, Cornell University. See also *Ann. Prob.* **3** (1975), 1–31.

Francis Comets

Université Paris 7 — Denis Diderot, Mathématiques, case 7012, 2 place Jussieu, 75251 Paris Cedex 05, France.

Mikhail Menshikov

Laboratory of Large Random Systems, Faculty of Mathematics and Mechanics, Moscow State University, 119899 Moscow, Russia
and
Department of Statistics, Institute for Mathematics and Statistics, University of São Paulo, Box Postal 66.281, São Paulo, SP, Brasil.

Serguei Popov

Institute for Problems of Information Transmission RAS, 101447, B. Karetnyj 19, Moscow, Russia.
and
Department of Statistics, Institute for Mathematics and Statistics, University of São Paulo, Box Postal 66.281, São Paulo, SP, Brasil.