STUDIA MATHEMATICA 212 (1) (2012)

An invariance principle for the law of the iterated logarithm for some Markov chains

by

W. Bołt, A. A. Majewski and T. Szarek (Gdańsk)

Abstract. Strassen's invariance principle for additive functionals of Markov chains with spectral gap in the Wasserstein metric is proved.

1. Introduction. Suppose that (E, ρ) is a Polish space. We denote by $\mathcal{B}(E)$ the family of all Borel sets in E, by $\mathcal{M}_1$ the space of all probability Borel measures on E, and we let $\pi : E \times \mathcal{B}(E) \to [0, 1]$ be a transition probability on E. The Markov operator P is defined by $Pf(x) = \int_E f(y)\,\pi(x, dy)$ for every bounded Borel measurable function f on E. The same formula defines Pf for any Borel measurable function $f \geq 0$ which need not be finite. Denote by $B_b(E)$ the set of all bounded Borel measurable functions equipped with the supremum norm and let $C_b(E)$ be its subset consisting of all bounded continuous functions. The connection of the Markov operator to the topology is usually given by the Feller property $P(C_b(E)) \subset C_b(E)$.

Suppose that $(X_n)_{n\geq 0}$ is an E-valued Markov chain, given over some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with transition probability π and initial distribution μ_0. Denote by $\mathbb{E}$ the expectation corresponding to $\mathbb{P}$. We shall denote by μP the associated transfer operator describing the evolution of the law of X_n. To be precise, μP is defined by the formula $\int_E f(x)\,\mu P(dx) = \int_E Pf(x)\,\mu(dx)$ for any $f \in B_b(E)$ and $\mu \in \mathcal{M}_1$. To simplify the notation we shall write $\langle f, \mu\rangle$ instead of $\int_E f(y)\,\mu(dy)$.

Given a Lipschitz function $\psi : E \to \mathbb{R}$ we define

$$S_n(\psi) := \psi(X_0) + \psi(X_1) + \cdots + \psi(X_n) \quad \text{for } n \geq 0.$$

Our aim is to find conditions under which $S_n(\psi)$ satisfies the law of the iterated logarithm (LIL). This natural question is raised when central limit theorems (CLT) are verified. Since 1986 when Kipnis and Varadhan [12]

2010 *Mathematics Subject Classification*: Primary 60J25, 60H15; Secondary 76N10.
Key words and phrases: ergodicity of Markov families, invariant measures, law of iterated logarithm.

DOI: 10.4064/sm212-1-3

proved the central limit theorem for additive functionals of stationary reversible ergodic Markov chains, there has been a huge amount of reviving attempts to do this in various settings and under different conditions (see [16, 17]). A common factor of those results was that they were established with respect to the stationary probability law of the chain. In [2, Theorem IV.8.1] Gordin and Lifshits answered the question about the validity of the CLT with respect to the law of the Markov chain starting at some point x. Namely, they proved that for coboundaries the CLT holds for almost every x with respect to the invariant initial distribution (see also [6, 7]). On the other hand, Guivarc'h and Hardy [9] proved the CLT for a class of Markov chains associated with the transfer operator having spectral gap. Recently Komorowski and Walczuk [13] studied Markov processes with the transfer operator having spectral gap in the Wasserstein metric and proved the CLT in the non-stationary case. Other interesting results under similar assumptions were obtained by S. Kuksin and A. Shirikyan (see [14, 21]).

The LIL we study in this note was also considered in many papers. There are several results governing, for instance, Harris recurrent chains [3, 4, 18]. Similarly CLT results are formulated mostly for stationary ergodic chains (see for instance [1, 5, 19, 26, 27]). In the case when one is able to find a solution to the Poisson equation $h = f + Ph$, the problem may be reduced to the martingale case [8] (see also [18]). But the LIL for martingales was carefully examined in many papers (see [11, 10, 23, 24]) and a lot of satisfactory results were obtained.

Our note is aimed at proving the LIL for Markov chains that have the spectral gap property in the Wasserstein metric. It is worth mentioning here that many Markov chains enjoy this property, e.g. Markov chains associated with iterated function systems or with stochastic differential equations disturbed by Poisson noise (see [15]).

Our result is based upon the LIL for martingales due to Heyde and Scott (see Theorem 1 in [11]).

2. Assumptions and auxiliary results. For every measure $\nu \in \mathcal{M}_1$ the law of the Markov chain $(X_n)_{n\geq 0}$ with transition probability π and initial distribution ν is the probability measure $\mathbb{P}_\nu$ on $(E^{\mathbb{N}}, \mathcal{B}(E)^{\otimes\mathbb{N}})$ such that

$$\mathbb{P}_\nu[X_{n+1} \in A \mid X_n = x] = \pi(x, A) \quad \text{and} \quad \mathbb{P}_\nu[X_0 \in A] = \nu(A),$$

where $x \in E$ and $A \in \mathcal{B}(E)$. The expectation with respect to $\mathbb{P}_\nu$ is denoted by $\mathbb{E}_\nu$. For $\nu = \delta_x$, the Dirac measure at $x \in E$, we write just $\mathbb{P}_x$ and $\mathbb{E}_x$.

We will make the following assumption:

(H0) the Markov operator has the Feller property, i.e. $P(C_b(E)) \subset C_b(E)$.

We shall denote by $\mathcal{M}_{1,1}$ the space of all probability measures with finite first moment, i.e. $\nu \in \mathcal{M}_{1,1}$ iff $\nu \in \mathcal{M}_1$ and $\int_E \rho(x_0, x)\, \nu(dx) < \infty$ for some (thus all) $x_0 \in E$. For abbreviation we shall write $\rho_{x_0}(x) = \rho(x_0, x)$. Observe that every Lipschitz function (even unbounded) is integrable with respect to each ν in $\mathcal{M}_{1,1}$. We assume that:

(H1) for any $\nu \in \mathcal{M}_{1,1}$ we have $\nu P \in \mathcal{M}_{1,1}$.

It may be proved that $\mathcal{M}_{1,1}$ is a complete metric space when equipped with the Wasserstein metric

$$d(\nu_1, \nu_2) = \sup\{|\langle f, \nu_1\rangle - \langle f, \nu_2\rangle| : f : E \to \mathbb{R}, \operatorname{Lip} f \le 1\} \quad \text{for } \nu_1, \nu_2 \in \mathcal{M}_{1,1}.$$

Here $\operatorname{Lip} f$ denotes the Lipschitz constant of f. Convergence in the Wasserstein metric is equivalent to weak convergence plus convergence of the first moment (see e.g. [25, Theorem 6.9]). The main assumption made in our note is that the Markov operator P is contractive with respect to the Wasserstein metric, i.e.

(H2) there exist $\gamma \in (0, 1)$ and $c > 0$ such that

$$d(\nu_1 P^n, \nu_2 P^n) \le c\gamma^n d(\nu_1, \nu_2) \quad \text{for } n \ge 1,\ \nu_1, \nu_2 \in \mathcal{M}_{1,1}. \tag{2.1}$$

Assumption (H2) is called the *spectral gap property*. Let $\mu \in \mathcal{M}_{1,1}$. From now on we assume that the initial distribution of $(X_n)_{n\ge0}$ is μ. Moreover,

(H3) there exist $x_0 \in E$ and $\delta > 0$ such that

$$\sup_{n\ge0} \mathbb{E}_\mu \rho_{x_0}^{2+\delta}(X_n) = \sup_{n\ge0} \int \rho_{x_0}^{2+\delta}\, d(\mu P^n) < \infty. \tag{2.2}$$

It is easy to prove that under assumptions (H0)–(H2) there exists a unique invariant (ergodic) measure $\mu_* \in \mathcal{M}_1$. In particular, $\mu_* \in \mathcal{M}_{1,1}$. The proof in [13] was given for Markov processes with continuous time but it still remains valid in the discrete case. In the stationary case, (H3) means that $\rho_{x_0}^{2+\delta}$ is in $L_1(\mu_*)$.

Let $n_0 \ge 2$ be such that

$$\gamma_0 = c^2\gamma^{n_0} < 1.$$

We start this part of the paper with a rather technical lemma.

Lemma 2.1. *Let $g_{n,k} : E^{2(k+n)} \to \mathbb{R}$, for arbitrary $k, n \ge 1$, be Lipschitz continuous functions in each variable with the same Lipschitz constant L. Then there exists a constant $\tilde{L}$, depending only on L, such that the function*

$$\begin{aligned} H_{n,k}(x) = &\int_E \pi_1(x, dy_1) \int_E \pi_2(y_1, dy_2) \cdots \int_E \pi_{2(k+n)-1}(y_{2(k+n)-2}, dy_{2(k+n)-1}) \\ &\times \int_E \pi_{2(k+n)}(y_{2(k+n)-1}, dy_{2(k+n)})\, g_{n,k}(y_1, \ldots, y_{2(k+n)}), \end{aligned} \tag{2.3}$$

where $\pi_l(y_{l-1}, dy_l) = \delta_{y_{l-1}} P^{k_l}(dy_l)$, $k_l \geq 1$ *and additionally* $k_l \geq n_0 - 1$ *for all even* l, *is Lipschitzean with Lipschitz constant* $\tilde{L}$.

Proof. Define the functions $g_j : E^j \to \mathbb{R}$ by the formula

$$g_j(y_0, y_1, \ldots, y_{j-1}) = \int_E \pi_j(y_{j-1}, dy_j) \int_E \pi_{j+1}(y_j, dy_{j+1}) \times \cdots$$
$$\times \int_E \pi_{2(k+n)}(y_{2(k+n)-1}, dy_{2(k+n)})\, g_{n,k}(y_1, \ldots, y_{2(k+n)}) \quad \text{for } j = 1, \ldots, 2(k+n).$$

Let $\mathcal{L}_{j,l}$ for $j = 1, \ldots, 2(k+n)$ and $l = 0, \ldots, j-1$ denote the Lipschitz constant of g_j with respect to y_l. Then the Lipschitz constant of $H_{n,k}$ is equal to $\mathcal{L}_{1,0}$. It is obvious that $\mathcal{L}_{j,l} \leq L$ for $0 \leq l < j-1$, $j > 1$. To evaluate $\mathcal{L}_{j,j-1}$ fix $y_0, y_1, \ldots, y_{j-2}$ and $\tilde{y}_{j-1}, \hat{y}_{j-1}$. Then we have

$$\begin{aligned}
&g_j(y_0, y_1, \ldots, y_{j-2}, \hat{y}_{j-1}) - g_j(y_0, y_1, \ldots, y_{j-2}, \tilde{y}_{j-1}) \\
&\quad = \int_E \pi_j(\hat{y}_{j-1}, dy_j)\, g_{j+1}(y_0, y_1, \ldots, \hat{y}_{j-1}, y_j) \\
&\qquad - \int_E \pi_j(\tilde{y}_{j-1}, dy_j)\, g_{j+1}(y_0, y_1, \ldots, \tilde{y}_{j-1}, y_j) \\
&\quad = \int_E \pi_j(\hat{y}_{j-1}, dy_j)\, (g_{j+1}(y_0, y_1, \ldots, \hat{y}_{j-1}, y_j) - g_{j+1}(y_0, y_1, \ldots, \tilde{y}_{j-1}, y_j)) \\
&\qquad + \int_E \pi_j(\hat{y}_{j-1}, dy_j)\, g_{j+1}(y_0, y_1, \ldots, \tilde{y}_{j-1}, y_j) \\
&\qquad - \int_E \pi_j(\tilde{y}_{j-1}, dy_j)\, g_{j+1}(y_0, y_1, \ldots, \tilde{y}_{j-1}, y_j),
\end{aligned}$$

and consequently

$$\begin{aligned}
&|g_j(y_0, y_1, \ldots, \hat{y}_{j-1}) - g_j(y_0, y_1, \ldots, \tilde{y}_{j-1})| \\
&\quad \leq \mathcal{L}_{j+1,j-1}\rho(\hat{y}_{j-1}, \tilde{y}_{j-1}) \int_E \pi_j(\hat{y}_{j-1}, dy_j) + |\langle P^{k_j} \tilde{g}_{j+1}, \delta_{\hat{y}_j}\rangle - \langle P^{k_j} \tilde{g}_{j+1}, \delta_{\tilde{y}_j}\rangle| \\
&\quad \leq L\rho(\hat{y}_{j-1}, \tilde{y}_{j-1}) + c_j \mathcal{L}_{j+1,j}\rho(\hat{y}_{j-1}, \tilde{y}_{j-1}),
\end{aligned}$$

where $c_j = c\gamma$ if j odd, $c_j = c\gamma^{n_0-1}$ if j even, and

$$\tilde{g}_{j+1}(\cdot) = g_{j+1}(y_0, y_1, \ldots, y_{j-2}, \tilde{y}_{j-1}, \cdot).$$

Hence

$$\mathcal{L}_{j,j-1} \leq L + c_j \mathcal{L}_{j+1,j} \quad \text{for } j = 1, \ldots, 2(k+n) - 1.$$

Since $\mathcal{L}_{2(k+n),2(k+n)-1} \leq L$, an easy computation shows that

$$\mathcal{L}_{1,0} \leq \frac{L(c\gamma + 1)}{1 - \gamma_0}.$$

This completes the proof. ∎

3. The law of the iterated logarithm

3.1. A martingale result. We start by recalling a classical result due to C. C. Heyde and D. J. Scott [11]. Let $\{M_n, \mathcal{F}_n : n \geq 0\}$ be a martingale on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$ where $\mathcal{F}_0 = \{\Omega, \emptyset\}$ and $\mathcal{F}_n$ is the σ-field generated by $M_1, \ldots, M_n$ for $n > 0$. Let $Z_0 = M_0 = 0$ $\mathbb{P}$-a.s. and $Z_n = M_n - M_{n-1}$ for $n \geq 1$. Further, let $s_n^2 = \mathbb{E}M_n^2 < \infty$.

We consider the metric space $(C, \tilde{\rho})$ of all real-valued continuous functions on $[0,1]$ with

$$\tilde{\rho}(x,y) = \sup_{0 \leq t \leq 1} |x(t) - y(t)| \quad \text{for } x, y \in C.$$

Let K be the set of absolutely continuous functions $x \in C$ such that $x(0) = 0$ and $\int_0^1 (x'(t))^2\,dt \leq 1$.

Define a real function g on $[0, \infty)$ by $g(s) = \sup\{n : s_n^2 \leq s\}$. We define a sequence of real random functions η_n on $[0,1]$, for $n > g(e)$, by

$$\eta_n(t) = \frac{M_k + (s_n^2 t - s_k^2)(s_{k+1}^2 - s_k^2)^{-1} Z_{k+1}}{\sqrt{2 s_n^2 \log\log s_n^2}}$$

if $s_k^2 \leq s_n^2 t \leq s_{k+1}^2$, $k = 1, \ldots, n-1$, and

$$\eta_n(t) = 0 \quad \text{for } n \leq g(e).$$

PROPOSITION 3.1 (Theorem 1 in [11]). *Under the above notations for the square integrable martingale (M_n), if $s_n^2 \to \infty$ and*

$$\sum_{n=1}^{\infty} s_n^{-4} \mathbb{E}[Z_n^4 \mathbf{1}_{\{|Z_n| < \gamma s_n\}}] < \infty \quad \textit{for some } \gamma > 0, \tag{3.1}$$

$$\sum_{n=1}^{\infty} s_n^{-1} \mathbb{E}[|Z_n| \mathbf{1}_{\{|Z_n| \geq \epsilon s_n\}}] < \infty \quad \textit{for all } \epsilon > 0, \tag{3.2}$$

$$s_n^{-2} \sum_{k=1}^{n} Z_k^2 \to 1 \quad \mathbb{P}\textit{-a.s. as } n \to \infty, \tag{3.3}$$

then $(\eta_n)_{n \geq 1}$ is relatively compact in C and the set of its limit points coincides with K.

3.2. Application to Markov chains. Let (E, ρ) be a Polish space, (X_n) a Markov chain with state space E, transition operator P satisfying conditions (H0)–(H2), and initial probability μ satisfying (H3). Let $\psi : E \to \mathbb{R}$ be a Lipschitz function with Lipschitz constant $L > 0$ such that $\langle \psi, \mu_* \rangle = 0$ (otherwise we could consider $\tilde{\psi} = \psi - \langle \psi, \mu_* \rangle$).

For every n and any Lipschitz function ψ, Minkowski's inequality in $L_{2+\delta}(\mu P^n)$ and (H3) yield

$$(3.4)\quad (\mathbb{E}_\mu[|\psi(X_n)|^{2+\delta}])^{1/(2+\delta)} = \Big(\int_E |\psi(x)|^{2+\delta}\, d(\mu P^n)\Big)^{1/(2+\delta)} \le |\psi(x_0)| + L\Big(\int_E \rho_{x_0}^{2+\delta}\, d(\mu P^n)\Big)^{1/(2+\delta)} < \infty$$

and consequently $\sup_{n\ge 0} \mathbb{E}_\mu[|\psi(X_n)|^{2+\delta}] < \infty$.

We have

$$\sum_{i=0}^{\infty} |P^i\psi(x)| = \sum_{i=0}^{\infty} |\langle \psi, \delta_x P^i\rangle - \langle \psi, \mu_* P^i\rangle| \le cd(\delta_x, \mu_*)\sum_{i=0}^{\infty}\gamma^i < \infty,$$

by (H2). Thus we may define the function

$$\chi(x) := \sum_{i=0}^{\infty} P^i\psi(x) \quad \text{for } x \in E.$$

We easily check that χ is a Lipschitz function (see Lemma 4.5 of [13]) and satisfies the Poisson equation $\chi - P\chi = \psi$.

It is well known (see e.g. [8]) that

$$M_n = \chi(X_n) - \chi(X_0) + \sum_{i=0}^{n-1} \psi(X_i) \quad \text{for } n \ge 0$$

is a martingale on the space $(E^{\mathbb{N}}, \mathcal{B}(E)^{\otimes\mathbb{N}}, \mathbb{P}_\mu)$ with respect to the natural filtration; its square integrable martingale differences are of the form

$$Z_n = \chi(X_n) - \chi(X_{n-1}) + \psi(X_{n-1}) \quad \text{for } n \ge 1.$$

Observe that $\mathbb{E}_{\mu_*} Z_1^2 < \infty$. Indeed, we easily check that $x \mapsto \mathbb{E}_x(Z_1^2 \wedge k)$ is a bounded continuous function for any $k \ge 1$. Further, as $\mathbb{E}_{\mu P^n}(Z_1^2 \wedge k) = \int_E \mathbb{E}_x(Z_1^2 \wedge k)\, \mu P^n(dx) \to \mathbb{E}_{\mu_*}(Z_1^2 \wedge k)$ for any $k \ge 1$ as $n \to \infty$ and $\sup_{n\ge 0} \mathbb{E}_{\mu P^n}(Z_1^2) < \infty$, we conclude that $\mathbb{E}_{\mu_*}(Z_1^2) < \infty$.

Set

$$\sigma^2 := \mathbb{E}_{\mu_*} Z_1^2.$$

We have

$$(3.5)\qquad \lim_{n\to\infty} \mathbb{E}_\mu Z_n^2 = \lim_{n\to\infty} \mathbb{E}_{\mu P^n} Z_1^2 = \sigma^2.$$

REMARK. The variance σ^2 defined above and the variance which appears as the variance of the limiting normal distribution in the CLT (see eg. [6] and [8]) are the same. One can prove that the function χ which solves the Poisson equation for ψ is in $L_2(\mu_*)$, just as we proved $E_{\mu_*}(Z_1^2) < \infty$. By [8] the martingale differences are in fact $Z_n = \chi(X_n) - P\chi(X_{n-1})$, and it can be easily computed that

$$\begin{aligned}(3.6)\qquad \sigma^2 = E_{\mu_*}(Z_1^2) &= E_{\mu_*}[\chi(X_1) - P\chi(X_0)]^2 \\ &= \int\limits_E \chi^2\, d\mu_* - \int\limits_E (P\chi)^2\, d\mu_*;\end{aligned}$$

the last term is precisely the variance of the limiting normal distribution given in [8]. The fact that χ is in $L_2(\mu_*)$ also yields for μ_*-almost every x the CLT in $(\Omega, \mathcal{F}, \mathbb{P}_x)$, by [2, Theorem IV.8.1].

In fact, since χ and ψ are Lipschitzean, we have $\sup_{n\ge1} \mathbb{E}_\mu |Z_n|^{2+\delta} < \infty$, by Minkowski's inequality and (3.4). Further, observe that

$$\sup_{n\ge1} \mathbb{E}_\mu (Z_n^2 \mathbf{1}_{\{|Z_n|^2 \ge k\}}) \le k^{-\delta/2} \sup_{n\ge1} \mathbb{E}_\mu |Z_n|^{2+\delta} \to 0$$

as $k \to \infty$. Therefore, condition (3.5) follows from the fact that $\mathbb{E}_{\mu P^n}(Z_1^2 \wedge k) \to \mathbb{E}_{\mu_*}(Z_1^2 \wedge k)$ as $n \to \infty$ for any $k \ge 1$. Finally, we obtain, by orthogonality of the martingale differences,

$$\begin{aligned}\lim_{n\to\infty} \frac{s_n^2}{n} &= \lim_{n\to\infty} \frac{\mathbb{E}_\mu M_n^2}{n} \\ &= \lim_{n\to\infty} \frac{\sum_{i=1}^n \mathbb{E}_\mu Z_i^2}{n} = \sigma^2.\end{aligned}$$

LEMMA 3.2. *The square integrable martingale differences* $(Z_n)_{n\ge1}$ *satisfy*

$$(3.7)\qquad \frac{1}{n} \sum_{l=1}^n Z_l^2 \to \sigma^2 \quad \mathbb{P}_\mu\text{-}a.s.\ as\ n \to \infty,$$

and consequently condition (3.3) *holds if* $\sigma^2 > 0$.

Proof. First observe that it is enough to show that for any $i \in \{1, \ldots, n_0\}$ we have

$$\frac{1}{n} \sum_{l=1}^n Z_{i+ln_0}^2 \to \sigma^2 \quad \mathbb{P}_\mu\text{-a.s. as } n \to \infty.$$

If we show that both the functions

$$x \mapsto \mathbb{E}_x\biggl(\biggl|\liminf_{n\to\infty}\biggl(\frac{1}{n}\sum_{l=1}^n Z_{i+ln_0}^2\biggr) - \sigma^2\biggr| \wedge 1\biggr)$$

and

$$x \mapsto \mathbb{E}_x\biggl(\biggl|\limsup_{n\to\infty}\biggl(\frac{1}{n}\sum_{l=1}^n Z_{i+ln_0}^2\biggr) - \sigma^2\biggr| \wedge 1\biggr)$$

are continuous, we shall be done. Indeed, then we have

$$\begin{aligned}
&\mathbb{E}_\mu\left(\left|\liminf_{n\to\infty}\left(\frac{1}{n}\sum_{l=1}^{n} Z_{i+ln_0}^2\right)-\sigma^2\right|\wedge 1\right)\\
&\qquad\qquad= \int_E \mathbb{E}_x\left(\left|\liminf_{n\to\infty}\left(\frac{1}{n}\sum_{l=1}^{n} Z_{i+ln_0}^2\right)-\sigma^2\right|\wedge 1\right)\mu(dx)\\
&\qquad\qquad= \int_E \mathbb{E}_x\left(\left|\liminf_{n\to\infty}\left(\frac{1}{n}\sum_{l=1}^{n} Z_{i+ln_0}^2\right)-\sigma^2\right|\wedge 1\right)\mu P^{i+mn_0}(dx)\\
&\qquad\qquad\to \mathbb{E}_{\mu_*}\left(\left|\liminf_{n\to\infty}\left(\frac{1}{n}\sum_{l=1}^{n} Z_{i+ln_0}^2\right)-\sigma^2\right|\wedge 1\right)
\end{aligned}$$

as $m\to\infty$, since μP^{i+mn_0} converges weakly to μ_* as $m\to\infty$. On the other hand, from the Birkhoff individual ergodic theorem we have

$$\mathbb{E}_{\mu_*}\left(\left|\liminf_{n\to\infty}\left(\frac{1}{n}\sum_{l=1}^{n} Z_{i+ln_0}^2\right)-\sigma^2\right|\wedge 1\right)=0$$

and consequently

$$\mathbb{E}_{\mu}\left(\left|\liminf_{n\to\infty}\left(\frac{1}{n}\sum_{l=1}^{n} Z_{i+ln_0}^2\right)-\sigma^2\right|\wedge 1\right)=0,$$

which, in turn, gives

$$\liminf_{n\to\infty}\left(\frac{1}{n}\sum_{l=1}^{n} Z_{i+ln_0}^2\right)=\sigma^2 \qquad \mathbb{P}_\mu\text{-a.s.}$$

Analogously we may show that

$$\limsup_{n\to\infty}\left(\frac{1}{n}\sum_{l=1}^{n} Z_{i+ln_0}^2\right)=\sigma^2 \qquad \mathbb{P}_\mu\text{-a.s.}$$

The remainder of the proof is devoted to showing the continuity of the relevant functions. Again, we restrict to the first function, since the proof for the second one goes in almost the same manner.

Observe that

$$\begin{aligned}
&\mathbb{E}_x\left(\left|\liminf_{n\to\infty}\left(\frac{1}{n}\sum_{l=1}^{n} Z_{i+ln_0}^2)-\sigma^2\right|\wedge 1\right)\\
&\ = \lim_{n\to\infty}\lim_{k\to\infty}\mathbb{E}_x\left(\left|\min\left\{\frac{1}{n}\sum_{l=1}^{n} Z_{i+ln_0}^2-\sigma^2,\dots,\frac{1}{n+k}\sum_{l=1}^{n+k} Z_{i+ln_0}^2-\sigma^2\right\}\right|\wedge 1\right)\\
&\ = \lim_{n\to\infty}\lim_{k\to\infty} H_{n,k}(x),
\end{aligned}$$

where

$$H_{n,k}(x) = \mathbb{E}_x\Big(\Big|\min\Big\{\frac{1}{n}\Big(\sum_{l=1}^{n} Z^2_{i+ln_0} \wedge n(1+\sigma^2)\Big) - \sigma^2, \dots, \frac{1}{n+k}\Big(\sum_{l=1}^{n+k} Z^2_{i+ln_0} \wedge (n+k)(1+\sigma^2) - \sigma^2\Big)\Big\}\Big| \wedge 1\Big)$$

$$= \mathbb{E}_x\Big(\Big|\min\Big\{\frac{1}{n}\Big(\sum_{l=1}^{n}(\chi(X_{i+ln_0}) - \chi(X_{i-1+ln_0}) + \psi(X_{i-1+ln_0}))^2 \wedge n(1+\sigma^2)\Big) - \sigma^2, \dots, \frac{1}{n+k}\Big(\sum_{l=1}^{n+k}(\chi(X_{i+ln_0}) - \chi(X_{i-1+ln_0}) + \psi(X_{i-1+ln_0}))^2 \wedge (n+k)(1+\sigma^2) - \sigma^2\Big)\Big\}\Big| \wedge 1\Big).$$

Set

$$g_{n,k}(y_1, \dots, y_{2(n+k)}) = \Big|\min\Big\{\frac{1}{n}\Big(\sum_{l=1}^{n}(\chi(y_{2l}) - \chi(y_{2l-1}) + \psi(y_{2l-1}))^2 \wedge n(1+\sigma^2)\Big) - \sigma^2, \dots, \frac{1}{n+k}\Big(\sum_{l=1}^{n+k}(\chi(y_{2l}) - \chi(y_{2l-1}) + \psi(y_{2l-1}))^2 \wedge (n+k)(1+\sigma^2) - \sigma^2\Big)\Big\}\Big| \wedge 1$$

so that

$$H_{n,k}(x) = \mathbb{E}_x(g_{n,k}(X_{i+n_0-1}, X_{i+n_0}, X_{i+2n_0-1}, X_{i+2n_0}, \dots, X_{i+2(n+k)n_0})).$$

Observe that $H_{n,k}$ is given by formula (2.3). If we show that there exists L such that $g_{n,k}$ is Lipschitz continuous in each variable with Lipschitz constant L (independent of n, k), then all $H_{n,k}$ are Lipschitzean with the same Lipschitz constant $\tilde{L}$, by Lemma 1. Consequently, $\lim_{n\to\infty} \lim_{k\to\infty} H_{n,k}$ is Lipschitzean and in particular continuous. Since the minimum of any finite family of functions which are Lipschitz continuous in each variable with Lipschitz constant L is Lipschitz continuous in each variable with the same Lipschitz constant, to finish the proof it is enough to observe that the function

$$(y_1, \dots, y_{2p}) \mapsto \frac{1}{p}\Big(\sum_{l=1}^{p}(\chi(y_{2l}) - \chi(y_{2l-1}) + \psi(y_{2l-1}))^2 \wedge p(1+\sigma^2)\Big) - \sigma^2$$

is Lipschitz continuous in each variable with Lipschitz constant L for some fixed $L > 0$. On the other hand, each term in the above sum is Lipschitz continuous in each variable with Lipschitz constant

$$(1/p)(\operatorname{Lip}\chi + \operatorname{Lip}\psi)2p(1+\sigma^2) = 2(\operatorname{Lip}\chi + \operatorname{Lip}\psi)(1+\sigma^2).$$

Observe that each variable appears in one term in the above sum. Hence $L \le 2(\operatorname{Lip}\chi + \operatorname{Lip}\psi)(1+\sigma^2)$, which finishes the proof. ∎

Note that following the proof of the previous lemma we could show that the Markov chain considered satisfies the strong law of large numbers (SLLN). This, however, follows directly from Theorem 2.1 in [20].

LEMMA 3.3. *Let $\sigma^2 > 0$. Under assumptions* (H0)–(H3) *the square integrable martingale differences $(Z_n)_{n\ge1}$ satisfy conditions* (3.1), (3.2).

Proof. Since $\sup_{n\ge1}\mathbb{E}_\mu|Z_n|^{2+\delta} < \infty$, where δ is the constant given in (H3), we have

$$\begin{aligned}\sum_{n=1}^{\infty} s_n^{-4}\mathbb{E}_\mu[Z_n^4\mathbf{1}_{\{|Z_n|<\gamma s_n\}}] &\le \sum_{n=1}^{\infty} s_n^{-4}\gamma^{2-\delta}s_n^{2-\delta}\mathbb{E}_\mu|Z_n|^{2+\delta}\\ &\le \gamma^{2-\delta}\sup_{n\ge1}\mathbb{E}_\mu|Z_n|^{2+\delta}\sum_{n=1}^{\infty}s_n^{-2-\delta}.\end{aligned}$$

On the other hand, the condition $s_n^2/n \to \sigma^2$ as $n\to\infty$ gives $\sum_{n=1}^{\infty} s_n^{-2-\delta} < \infty$, which completes the proof of condition (3.1).

To show condition (3.2) observe that

$$\begin{aligned}\sum_{n=1}^{\infty} s_n^{-1}\mathbb{E}_\mu[|Z_n|\mathbf{1}_{\{|Z_n|\ge\epsilon s_n\}}] &\le \sum_{n=1}^{\infty} s_n^{-1}\mathbb{E}_\mu[|Z_n|^{2+\delta}/(\epsilon s_n)^{1+\delta}]\\ &\le \epsilon^{-1-\delta}\sup_{n\ge1}\mathbb{E}_\mu|Z_n|^{2+\delta}\sum_{n=1}^{\infty}s_n^{-2-\delta} < \infty.\end{aligned}$$ ∎

3.3. The law of the iterated logarithm for Markov chains

THEOREM 3.4. *Let (E,ρ) be a Polish space, (X_n) a Markov chain with state space E, transition operator P satisfying conditions* (H0)–(H2), *and initial probability μ satisfying* (H3). *If ψ is a Lipschitz function with $\langle\psi,\mu_*\rangle = 0$ and $\sigma^2 > 0$, then $\mathbb{P}_\mu$-a.s. the sequence*

$$\theta_n(t) = \frac{\sum_{i=1}^{k}\psi(X_i) + (nt-k)\psi(X_{k+1})}{\sigma\sqrt{2n\log\log n}}$$

for $k \le nt \le k+1$, $k = 1,\dots,n-1$ and $t>0$, $n>e$, and $\theta_n(t) = 0$ otherwise, is relatively compact in C and the set of its limit points coincides with K.

Proof. First observe that since $s_n^2/n \to \sigma^2 > 0$ as $n\to\infty$, we have

$$\frac{\sqrt{2s_n^2\log\log s_n^2}}{\sigma\sqrt{2n\log\log n}} \to 1 \quad \text{as } n\to\infty.$$

Consequently, from Lemmas 3.2 and 3.3 it follows that the sequence

$$\eta_n(t) = \frac{M_k + (s_n^2 t - s_k^2)(s_{k+1}^2 - s_k^2)^{-1} Z_{k+1}}{\sigma\sqrt{2n \log\log n}}$$

for $s_k^2 \le s_n^2 t \le s_{k+1}^2$, $k = 1, \ldots, n-1$ and $t > 0, n > e$, and $\eta_n(t) = 0$ otherwise, is relatively compact in C and the set of its limit points coincides with K (see Heyde and Scott [11]). Let $t \in (0,1]$ and $n \ge 1$. Observe that if $k \le nt \le k+1$, then

$$\frac{k\sigma^2}{s_k^2} s_k^2 \le \frac{n\sigma^2}{s_n^2} t s_n^2 \le \frac{(k+1)\sigma^2}{s_{k+1}^2} s_{k+1}^2.$$

Set

$$\hat{\eta}_n(t) = \frac{M_k + (nt - k)Z_{k+1}}{\sigma\sqrt{2n \log\log n}},$$

where $k \ge 1$ is such that $k \le nt \le k+1$. Since $n\sigma^2/s_n^2 \to 1$ as $n \to \infty$, for any $\varepsilon > 0$ we have

$$(1-\varepsilon)s_k^2 \le (1+\varepsilon)s_n^2 t \le (1+\varepsilon)^2(1-\varepsilon)^{-1} s_{k+1}^2$$

for all n large enough. Hence there is $t_* \in [t(1-\varepsilon)(1+\varepsilon)^{-1}, t(1+\varepsilon)(1-\varepsilon)^{-1}]$ such that $s_k^2 \le s_n^2 t_* \le s_{k+1}^2$. On the other hand, the diameter of the interval $[s_k^2/s_n^2, s_{k+1}^2/s_n^2]$ for a fixed $k = 1, \ldots, n-1$ converges to 0 as $n \to \infty$. Consequently, for any $t > 0$ and $n > e$ there exists $t_n > 0$ such that $\hat{\eta}_n(t) = \eta_n(t_n)$ and $t_n \to t$ as $n \to \infty$. Since the sequence $(\eta_n(t))_{n>e}$ is relatively compact in C and the set of its limit points coincides with K, the sequence $(\hat{\eta}_n(t))_{n>e}$ is also relatively compact and has the same set of limit points.

Fix $\varepsilon > 0$. Define

$$\begin{aligned} A_n = &\left\{\omega \in \Omega : \frac{|M_n - \sum_{i=1}^{n-1} \psi(X_i)|}{\sqrt{n}} \ge \varepsilon/2\right\} \\ &\cup \left\{\omega \in \Omega : \frac{|Z_{n+1} - \psi(X_n)|}{\sqrt{n}} \ge \varepsilon/2\right\} \quad \text{for } n \ge 1. \end{aligned}$$

Now we are going to show that $\sum_{n=1}^{\infty} \mathbb{P}_\mu(A_n) < \infty$. Indeed, keeping in mind that χ is Lipschitzean, by the Chebyshev inequality we obtain

$$\begin{aligned} \mathbb{P}_\mu&\left(\left\{\omega \in \Omega : \frac{|M_n - \sum_{i=1}^{n-1} \psi(X_i)|}{\sqrt{n}} \ge \varepsilon/2\right\}\right) \\ &= \mathbb{P}_\mu\left(\left\{\omega \in \Omega : \frac{|\chi(X_n) - \chi(X_0)|}{\sqrt{n}} \ge \varepsilon/2\right\}\right) \\ &\le c_0 \frac{\mathbb{E}(\rho_{x_0}(X_n))^{2+\delta} + \mathbb{E}(\rho_{x_0}(X_0))^{2+\delta}}{n^{1+\delta/2}} \le \frac{\tilde{c}}{n^{1+\delta/2}}, \end{aligned}$$

by (H3), for some constant $\tilde{c} > 0$ independent of n.

Analogously, we may check that there exists a positive constant $\tilde{C}$ (independent of n) such that

$$\mathbb{P}_\mu\biggl(\biggl\{\omega\in\Omega: \frac{|Z_{n+1}-\psi(X_n)|}{\sqrt{n}}\ge\varepsilon/2\biggr\}\biggr) = \mathbb{P}_\mu\biggl(\biggl\{\omega\in\Omega: \frac{|\chi(X_{n+1})-\chi(X_n)|}{\sqrt{n}}\ge\varepsilon/2\biggr\}\biggr)\le\frac{\tilde{C}}{n^{1+\delta/2}},$$

by (H3) and the Lipschitz property of χ. Thus the series $\sum_{n=1}^{\infty}\mathbb{P}_\mu(A_n)$ is convergent.

Finally, from the Borel–Cantelli lemma it follows that $\mathbb{P}_\mu$-a.s.

$$\limsup_{n\to\infty}\sup_{0\le t\le 1}\biggl|\frac{M_k+(nt-k)Z_{k+1}}{\sigma\sqrt{2n\log\log n}}-\frac{\sum_{i=1}^{k}\psi(X_i)+(nt-k)\psi(X_{k+1})}{\sigma\sqrt{2n\log\log n}}\biggr|<\varepsilon,$$

where $k\le nt\le k+1$. Since $\varepsilon>0$ was arbitrary, the proof is complete. ∎

Acknowledgements. The research of T.S. was partly supported by NSF (grant no. N N201 419139). The authors wish to express their gratitude to an anonymous referee for a thorough reading of the manuscript and valuable remarks. We have learned a lot from the report.

References

[1] M. A. Arcones, *The law of the iterated logarithm over a stationary Gaussian sequence of random vectors*, J. Theoret. Probab. 12 (1999), 615–641.

[2] A. N. Borodin and I. Ibragimov, *Limit theorems for functionals of random walks*, Trudy Mat. Inst. Steklova 195 (1994) (in Russian).

[3] X. Chen, *Limit theorems for functionals of ergodic Markov chains with general state space*, Mem. Amer. Math. Soc. 139 (1999), no. 664.

[4] X. Chen, *The law of the iterated logarithm for functionals of Harris recurrent Markov chains: self-normalization*, J. Theoret. Probab. 12 (1999), 421–445.

[5] C. Cuny, *Pointwise theorems with rate with applications to limit theorems for stationary processes*, Stoch. Dynam. 11 (2011), 135–155.

[6] Y. Derriennic and M. Lin, *Variance bounding Markov chains, L_2-uniform mean ergodicity and the CLT*, Stoch. Dynam. 11 (2011), 81–94.

[7] Y. Derriennic and M. Lin, *The central limit theorem for Markov chains started at a point*, Probab. Theory Related Fields 125 (2003), 73–76.

[8] M. I. Gordin and B. A. Lifšic, *The central limit theorem for stationary Markov processes*, Soviet Math. Dokl. 19 (1978), 392–394.

[9] Y. Guivarc'h et J. Hardy, *Théorèmes limites pour une classe de chaînes de Markov et application aux difféomorphismes d'Anosov*, Ann. Inst. H. Poincaré 24 (1988), 73–98.

[10] P. Hall and C. C. Heyde, *Martingale Limit Theory and its Application*, Academic Press, New York, 1980.

[11] C. C. Heyde and D. J. Scott, *Invariance principles for the law of the iterated logarithm for martingales and processes with stationary increments*, Ann. Probab. 1 (1973), 428–436.

[12] C. Kipnis and S. R. Varadhan, *Central limit theorem for additive functionals of reversible Markov processes and applications to simple exclusions*, Comm. Math. Phys. 104 (1986), 1–19.

[13] T. Komorowski and A. Walczuk, *Central limit theorem for Markov processes with spectral gap in the Wasserstein metric*, Stoch. Process. Appl. 122 (2012), 2155–2184.

[14] S. B. Kuksin, *Ergodic theorems for 2D statistical hydrodynamics*, Rev. Math. Phys. 14 (2002), 1–16.

[15] A. Lasota, *From fractals to stochastic differential equations*, in: Chaos—The Interplay Between Stochastic and Deterministic Behaviour (Karpacz, 1995), Lecture Notes in Phys. 457, Springer, 1995, 235–255.

[16] M. Maxwell, *Local and global central limit theorems for stationary ergodic sequences*, Ph.D. dissertation, Univ. Michigan, 1997.

[17] M. Maxwell and M. Woodroofe, *Central limit theorems for additive functionals of Markov chains*, Ann. Probab. 28 (2000), 713–724.

[18] S. P. Meyn and R. L. Tweedie, *Markov Chains and Stochastic Stability*, Springer, Berlin, 1993.

[19] E. Rio, *The functional law of the iterated logarithm for stationary strongly mixing sequences*, Ann. Probab. 23 (1995), 1188–1203.

[20] A. Shirikyan, *A version of the law of large numbers and applications*, in: Probabilistic Methods in Fluids, World Sci., 2003, 263–271.

[21] A. Shirikyan, *Law of large numbers and central limit theorem for randomly forced PDE's*, Probab. Theory Related Fields 134 (2006), 215–247.

[22] W. F. Stout, *A martingale analogue of Kolmogorov's law of the iterated logarithm*, Z. Wahrsch. Verw. Gebiete 15 (1970), 279–290.

[23] V. Strassen, *An invariance principle for the law of the iterated logarithm*, Z. Wahrsch. Verw. Gebiete 3 (1964), 211–226.

[24] V. Strassen, *Almost sure behavior of sums of independent random variables and martingales*, in: Proc. 5th Berkeley Sympos. Math. Statist. and Probab. 2 (1965), 315–344.

[25] C. Villani, *Optimal Transport, Old and New*, Springer, 2008.

[26] M. Woodroofe and O. Zhao, *Law of the iterated logarithm for stationary processes*, Ann. Probab. 36 (2008), 127–142.

[27] L. Wu, *Functional law of iterated logarithm for additive functionals of reversible Markov processes*, Acta Math. Appl. Sinica (English Ser.) 16 (2000), 149–161.

W. Bołt, A. A. Majewski, T. Szarek
Institute of Mathematics
University of Gdańsk
Wita Stwosza 57
80-952 Gdańsk, Poland
E-mail: ja@hope.art.pl
aamajewski@gmail.com
szarek@intertele.pl

Received March 10, 2012
Revised version October 13, 2012 (7451)