ANGEWANDTE MATHEMATIK

UND

INFORMATIK

On the Harris Recurrence of Iterated Random Lipschitz Functions and Related Convergence Rate Results

Gerold Alsmeyer

FB 10, Institut für Mathematische Statistik
Einsteinstraße 62, D-48149 Münster, Germany
e-mail: gerolda@math.uni-muenster.de

Bericht Nr. 12/00-S

UNIVERSITÄT MÜNSTER

On the Harris Recurrence of Iterated Random Lipschitz Functions and Related Convergence Rate Results

Gerold Alsmeyer

Institut für Mathematische Statistik
Fachbereich Mathematik
Westfälische Wilhelms-Universität Münster
Einsteinstraße 62
D-48149 Münster, Germany

A result by Elton [6] states that iterated function systems

$$M_n = F_n(M_{n-1}), \quad n \geq 1,$$

of i.i.d. random Lipschitz maps $F_1, F_2, ...$ on a locally compact, complete separable metric space $(\mathbb{X}, d)$ converge weakly to its unique stationary distribution π if the pertinent Liapunov exponent is a.s. negative and $\mathbb{E} \log^+ d(F_1(x_0), x_0) < \infty$ for some $x_0 \in \mathbb{X}$. Diaconis and Freedman [5] showed the convergence rate be geometric in the Prokhorov metric if $\mathbb{E}L_1^p < \infty$ and $\mathbb{E}d(F_1(x_0), x_0)^p < \infty$ for some $p > 0$, where L_1 denotes the Lipschitz constant of F_1. The same and also polynomial rates have been recently obtained in [1] by different methods. In this article, necessary and sufficient conditions are given for the positive Harris recurrence of $(M_n)_{n \geq 0}$ on some absorbing subset $\mathbb{H}$ of $\mathbb{X}$. If $\mathbb{H} = \mathbb{X}$ and the support of π has nonempty interior, we further show that the same respective moment conditions ensuring the weak convergence rate results mentioned above now lead to polynomial, respectively geometric rate results for the convergence to π in total variation $\| \cdot \|$ or f-norm $\| \cdot \|_f$, $f(x) = 1 + d(x, x_0)^\eta$ for some $\eta \in (0, p]$. The results are applied to various examples that have been discussed in the literature, including the Beta walk, multivariate ARMA models and matrix recursions.

AMS 1991 subject classifications. 60J05, 60J15, 60G17.

Keywords and phrases. Iterated function system, Lipschitz map, Lipschitz constant, Liapunov exponent, Harris recurrence, total variation, f-ergodicity, geometric ergodicity, strictly contractive, drift condition, level γ ladder epoch.

1. Introduction

Consider a sequence of the form

$$M_n \;=\; F(\theta_n, M_{n-1}), \quad n \geq 1, \tag{1.1}$$

which satisfies the following assumptions:

(1) $M_0, \theta_1, \theta_2, ...$ are independent random elements on a common probability space $(\Omega, \mathfrak{A}, \mathbb{P})$;

(2) $\theta_1, \theta_2, ...$ are identically distributed with common distribution Λ and take values in a measurable space $(\Theta, \mathcal{A})$;

(3) $M_0, M_1, ...$ take values in a locally compact, complete separable metric space $(\mathbb{X}, d)$ with Borel-σ-field $\mathfrak{B}(\mathbb{X})$;

(4) $F : (\Theta \times \mathbb{X}, \mathcal{A} \otimes \mathfrak{B}(\mathbb{X})) \to (\mathbb{X}, \mathfrak{B}(\mathbb{X}))$ is jointly measurable and Lipschitz continuous in the second argument.

$(M_n)_{n\geq 0}$ clearly defines a temporally homogeneous Markov chain called an *iterated function system (IFS) of i.i.d. Lipschitz maps* hereafter. Its (n-step) transition kernel is denoted P (P^n). For $x \in \mathbb{X}$, let $\mathbb{P}_x$ be the probability measure on the underlying measurable space under which $M_0 = x$ a.s. The associated expectation is denoted $\mathbb{E}_x$, as usual. For an arbitrary distribution ν on $\mathbb{X}$, we put $\mathbb{P}_\nu(\cdot) \stackrel{\text{def}}{=} \int \mathbb{P}_x(\cdot)\,\nu(dx)$ with associated expectation $\mathbb{E}_\nu$. We use $\mathbb{P}$ and $\mathbb{E}$ for probabilities and expectations, respectively, that do not depend on the initial distribution.

Let us write F_n for $F(\theta_n, \cdot)$. Given a Lipschitz map $f : \mathbb{X} \to \mathbb{X}$, define its Lipschitz constant as

$$l(f) \;\stackrel{\text{def}}{=}\; \sup_{x\neq y} \frac{d(f(x), f(y))}{d(x,y)}.$$

Put further

$$L_n \;\stackrel{\text{def}}{=}\; l(F_n) \tag{1.2}$$

for $n \geq 1$ and note that $(L_n)_{n\geq 1}$ forms a sequence of i.i.d. random variables which is independent of M_0. Its distribution does therefore not depend on the distribution of M_0, that is, it is the same under every $\mathbb{P}_\nu$.

Elton [6] showed that, under every $\mathbb{P}_x$, $(M_n)_{n\geq 0}$ converges weakly to a stationary distribution π provided that its *Liapunov exponent* l^* is a.s. negative, i.e.

$$l^* \;\stackrel{\text{def}}{=}\; \lim_{n\to\infty} n^{-1} \log l(F_n \circ ... \circ F_1) \;<\; 0, \tag{1.3}$$

and furthermore

$$\mathbb{E}\log^+ L_1 < \infty \quad \text{and} \quad \mathbb{E}\log^+ d(F_1(x_0), x_0) < \infty \tag{1.4}$$

for some $x_0 \in \mathbb{X}$. He further showed that π is unique and $(M_n)_{n\geq 0}$ ergodic under $\mathbb{P}_\pi$. By Birkhoff's ergodic theorem, the latter implies for each $B \in \mathfrak{B}(\mathbb{X})$

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} \mathbf{1}_B(M_k) \;=\; \pi(B) \tag{1.5}$$

$\mathbb{P}_\pi$-a.s. and thus also $\mathbb{P}_x$-a.s. for π-almost all $x \in \mathbb{X}$. Hence, if $\pi(B) > 0$, then

$$\mathbb{P}_x(M_n \in B \text{ i.o.}) \;=\; 1 \tag{1.6}$$

for π-almost all $x \in \mathbb{X}$ and we would like to conclude that every π-positive set B is recurrent. Unfortunately, the π-null set of $x \in \mathbb{X}$ for which (1.6) fails to hold generally depends on the set B. On the other hand, if it does not, we infer the π-irreducibility of the chain $(M_n)_{n\geq 0}$ on some $\mathbb{H}$ with $\pi(\mathbb{H}) = 1$ and then, because of (1.6) for each π-positive B, further its Harris recurrence on $\mathbb{H}$. Provided additionally aperiodicity, this in turn implies that $\mathbb{P}_x(M_n \in \cdot)$ converges to π in total variation for every $x \in \mathbb{H}$ which, of course, is a much stronger conclusion than Elton's result. With regard to a further analysis of IFS, for instance the rate of convergence towards stationarity (in total variation), it also gives access to the highly developed theory for irreducible and Harris recurrent Markov chains on general state spaces.

Given an IFS of i.i.d. Lipschitz maps satisfying (1.3) and (1.4), two questions will be considered in this article and discussed in various examples. In Section 2, we state two equivalent conditions for the positive Harris recurrence on some absorbing subset $\mathbb{H}$ of $\mathbb{X}$ (Theorem 2.1) and also provide a sufficient condition for $\mathbb{H} = \mathbb{X}$ (Theorem 2.2). These conditions are quite often easy to check in applications when the stationary distribution is known to some extent. The proofs of Theorem 2.1 and 2.2 can be found in Section 4. Section 3 deals with the convergence towards stationarity for Harris recurrent IFS. Under additional moment conditions on L_1 and $d(F_1(x_0), x_0)$, we will show f-regularity and f-ergodicity for suitable functions f (Theorem 3.1) and provide polynomial as well as geometric rates of convergence towards stationarity (Theorem 3.2). While Theorem 3.1 is a rather straightforward consequence of results in [9], the proof of Theorem 3.2, given in Section 5, will take some effort and will be based upon regenerative arguments developed in [1] in combination with the use of Liapunov drift functions. It also requires the additional assumptions $\mathbb{H} = \mathbb{X}$ and *supp* $\pi \neq \emptyset$, where *supp* π denotes the support of π. Again, the result will be illustrated in a number of examples.

2. Necessary and Sufficient Criteria for Harris Recurrence

The following theorem and first main result of this article confirms that, given (1.3) and (1.4), the conclusions mentioned after (1.6) are indeed true under an additional absolute continuity condition on the transition kernel P. A set $B \in \mathfrak{B}(\mathbb{X})$ is called *π-positive*, if $\pi(B) > 0$, *π-full*, if $\pi(B) = 1$, and *(P-)absorbing*, if $P(x, B) = 1$ for all $x \in B$. Given two non-zero σ-finite measures ν, λ on $\mathbb{X}$, we say that *ν possesses a λ-continuous component* if $\nu(dx) \geq g(x)\lambda(dx)$ for some measurable function $g : \mathbb{X} \to [0, \infty)$ with $\int g\, d\lambda > 0$. For the definitions of irreducibility, Harris recurrence and related notions for Markov chains on general state spaces not explicitly repeated here, we always refer to the excellent and by now standard monograph by Meyn and Tweedie [9].

THEOREM 2.1. *Suppose $(M_n)_{n\ge 0}$ is an IFS of i.i.d. Lipschitz maps satisfying (1.3) and (1.4). Let π denote its stationary distribution. Consider the following assertions:*

(a) There exists a π-positive set $\mathfrak{X}$ and a non-zero σ-finite measure λ on $(\mathbb{X},\mathfrak{B}(\mathbb{X}))$ such that each $P(x,\cdot)$, $x\in\mathfrak{X}$, possesses a λ-continuous component.

(b) There exists a π-positive set $\mathfrak{X}$ and a non-zero σ-finite measure λ on $(\mathbb{X},\mathfrak{B}(\mathbb{X}))$ such that each $\sum_{n\ge 1}2^{-n}P^n(x,\cdot)$, $x\in\mathfrak{X}$, possesses a λ-continuous component.

(c) There exists a π-full, absorbing set $\mathbb{H}\in\mathfrak{B}(\mathbb{X})$, such that $(M_n)_{n\ge 0}$ is an aperiodic, positive Harris chain on $\mathbb{H}$.

Then (a)$\Rightarrow$(b)$\Leftrightarrow$(c) holds true.

Provided that $(M_n)_{n\ge 0}$ is a Harris chain on a set $\mathbb{H}$ as described in Theorem 2.1(c), this set is called a *Harris set (for $(M_n)_{n\ge 0}$)*. It is well-known that in this case there always exists a maximal absorbing set with this property, called *maximal Harris set*. Our second theorem contains some information on when this latter set is the whole space $\mathbb{X}$. Let $\mathrm{int}(B)$ denote the interior of a set $B\in\mathfrak{B}(\mathbb{X})$.

THEOREM 2.2. *Given the situation of Theorem 2.1, suppose $(M_n)_{n\ge 0}$ is Harris recurrent with maximal Harris set $\mathbb{H}$. Then the following assertions hold:*

(a) Either $\pi(\mathrm{int}(\mathbb{H}))=0$, or $\mathbb{H}=\mathbb{X}$.

(b) If (a) or (b) in Theorem 2.1 holds for some $\mathfrak{X}$ with $\pi(\mathrm{int}(\mathfrak{X}))>0$ and if $\mathrm{int}(supp\,\pi)\ne\emptyset$, then $\mathbb{H}=\mathbb{X}$.

In order to put our results in the right place within the extensive and well-established theory of Markov chains on general state spaces, the following comments and corollaries might be helpful.

REMARK A. Suppose we are given any Markov chain $(M_n)_{n\ge 0}$ on $(\mathbb{X},\mathfrak{B}(\mathbb{X}))$ with transition kernel P and stationary distribution π under which it forms an ergodic sequence. If $(M_n)_{n\ge 0}$ is π-irreducible on some P-absorbing set $\mathbb{X}_1$ then, by the recurrence/transience dichotomy [9, Theorem 8.3.4], $(M_n)_{n\ge 0}$ is also recurrent on $\mathbb{X}_1$, i.e. $U(x,B)\stackrel{\text{def}}{=}\sum_{n\ge 0}P^n(x,B)=\infty$ for all π-positive $B\in\mathfrak{B}(\mathbb{X}_1)$. Indeed, if it were not, there would be a uniformly transient, π-positive B, which thus satisfied $\sup_{x\in B}U(x,B)<\infty$. On the other hand, the ergodic theorem ensures $\lim_{n\to\infty}n^{-1}\sum_{k=1}^n P^k(x,B)=\pi(B)>0$ and therefore $U(x,B)=\infty$ for π-almost all $x\in B$. Hence B must be recurrent. We thus see that the π-irreducibility of the chain already entails its recurrence, which in turn [9, Theorem 9.0.1] implies its positive Harris recurrence on an absorbing and π-full set $\mathbb{H}$. Consequently, Theorem 2.1 is really about necessary and sufficient conditions for the π-irreducibility of $(M_n)_{n\ge 0}$.

COROLLARY 2.3. *Given the situation of Theorem 2.1, its assertions (b),(c) hold if, and only if, $(M_n)_{n\ge 0}$ is π-irreducible.*

Remark B. Condition 2.1(a) may be reformulated as $P(x,dy) \ge \mathbf{1}_{\mathfrak{X}}(x) g(x,y)\,\lambda(dy)$ for some π-positive set $\mathfrak{X}$, a non-zero σ-finite measure λ on $(\mathbb{X}, \mathfrak{B}(\mathbb{X}))$ and a product measurable function $g : \mathbb{X}^2 \to [0,\infty)$ satisfying $\int_{\mathbb{X}} g(x,y)\,\lambda(dy) > 0$ for all $x \in \mathfrak{X}$. The product measurability of g follows from the fact that $\mathfrak{B}(\mathbb{X})$ is countably generated. An equivalent statement of condition 2.1(b) is that, for each $x \in \mathfrak{X}$, there exists $n(x) \ge 1$ such that $P^{n(x)}(x,\cdot)$ is nonsingular with respect to λ. Since at least one of the sets $\mathfrak{X}_n \stackrel{\text{def}}{=} \{x \in \mathfrak{X} : n(x) = n\}$, $n \ge 1$, must be π-positive, 2.1(b) is also equivalent to condition 2.1(a) with P, $\mathfrak{X}$ replaced with P^n, $\mathfrak{X}_n$ for some $n \ge 1$, i.e.

$$P^n(x,dy) \;\ge\; \mathbf{1}_{\mathfrak{X}_n}(x) g_n(x,y)\,\lambda(dy) \tag{2.1}$$

for some product measurable function $g_n : \mathbb{X}^2 \to [0,\infty)$ satisfying $\int_{\mathbb{X}} g_n(x,y)\,\lambda(dy) > 0$ for all $x \in \mathfrak{X}_n$. We will show in the proof of Theorem 2.1 that λ can be chosen in such a way that $\lambda \le \pi$. Notice that $\mathfrak{X}_n$ forms a small set, if $\inf_{x \in \mathfrak{X}_n} g_n(x,y) \ge \delta$ for some $\delta > 0$, but that neither this nor the existence of any other small set is generally implied by (2.1). On the other hand, each Harris chain with stationary distribution π possesses π-positive small sets. We hence note as another nontrivial consequence of Theorem 2.1:

Corollary 2.4. *Given the situation of Theorem 2.1, its assertions (b),(c) hold if, and only if,* $(M_n)_{n\ge0}$ *possesses a* π*-positive small set.*

Remark C. As we are dealing with a special class of ergodic Markov chains, namely IFS of i.i.d. Lipschitz maps under an average contraction condition, the question seems natural, how this enters into our results. Obviously, $(M_n)_{n\ge0}$ is weak Feller, i.e. $Pf(x) \stackrel{\text{def}}{=} \int f(y)\,P(x,dy) = \mathbb{E}_x f(M_1)$ is a bounded continuous function whenever f has this property. However, $(M_n)_{n\ge0}$ even has a stronger property not shared by all weak Feller chains: Let τ be any a.s. finite stopping time and put $P^{(\tau)}f(x) \stackrel{\text{def}}{=} \mathbb{E}_x f(M_\tau)$. Then we see upon noting $P^{(\tau)}f(x) = \mathbb{E} f \circ F_{1:\tau}(x)$ that $P^{(\tau)}f$ is a also a bounded continuous function if f has this property. Of course, the nontrivial part of this implication is the inheritance of continuity.

Remark D. Given the Harris recurrence of $(M_n)_{n\ge0}$ on some Harris set $\mathbb{H}$, the stationary distribution is clearly a maximal irreducibility measure. On the other hand, the proof of Theorem 2.1 will show that $(M_n)_{n\ge0}$ is also $\lambda(\cdot \cap \mathfrak{X}')$-irreducible for some λ-positive $\mathfrak{X}'$ and every λ such that 2.1(a) or 2.1(b) holds. Consequently, π dominates $\lambda(\cdot \cap \mathfrak{X}')$, which in turn shows that *supp* π has inner points whenever this holds true for *supp* $\lambda(\cdot \cap \mathfrak{X}')$. The latter may sometimes be easier to check, for instance, if λ is Lebesgue measure on Euclidean space.

Remark E. The verification of 2.1(a) or 2.1(b) requires some knowledge of the stationary distribution π as we must be able to check π-positivity of the set $\mathfrak{X}$. This can be a problem. On the other hand, in many examples like 2.6(a)–(d) below this becomes unnecessary because $\mathfrak{X} = \mathbb{X}$. If further $\mathbb{X}$ is Euclidean and λ in 2.1(a) or 2.1(b) is Lebesgue measure, then, by the previous remark, *supp* π has nonempty interior, whence Theorem 2.2(c) renders the positive Harris recurrence of $(M_n)_{n\ge0}$ on whole $\mathbb{X}$.

REMARK F. It should be clear that $\mathbb{H}$ can be small compared to $\mathbb{X}$ in a set theoretic or topological sense. Take the trivial deterministic example where $F_n(x) = (1+x^2)/2$ on $\mathbb{X} = [0,1]$ for all $n \geq 1$. Then $\pi = \delta_1$ and $P(x,\cdot) = \delta_{(1+x^2)/2}$. We infer Harris recurrence on $\mathbb{H} = \{1\}$ but on no larger subset of $[0,1]$.

REMARK G. Let us finally point out that, given the situation of Theorem 2.1, we must have that $\sum_{n\geq 1} 2^{-n} P^n(x,\cdot)$ and π are mutually singular for π-almost all $x \in \mathbb{X}$ if 2.1(b) and thus also 2.1(c),(b) fail to hold. Consequently, $P^n(x,\cdot)$ converges weakly to π in this situation by Elton's result, while total variation convergence fails to hold for π-almost all $x \in \mathbb{X}$. This may be rephrased as the following zero-one law in which $\stackrel{w}{\to}$ denotes weak convergence and $\|\cdot\|$ the total variation distance.

COROLLARY 2.5. *Given the situation of Theorem 2.1, $P^n(x,\cdot) \stackrel{w}{\to} \pi$ for all $x \in \mathbb{X}$ and*

$$\pi\Big(\{x \in \mathbb{X} : \lim_{n\to\infty} \|P^n(x,\cdot) - \pi\| = 0\}\Big) \;=\; 0 \text{ or } 1,$$

the probability being 1 *iff 2.1(c)-(c) hold true.*

2.6. EXAMPLES. In the following, F always denotes a generic copy of $F_1, F_2, ...$ and $\lambda\!\!\lambda$ Lebesgue measure on $\mathbb{R}$ (or some subset).

(a) This is the motivating example in [5], called *Beta walk*, see 2.1 and 6.3 there. Let $\mathbb{X} \stackrel{\text{def}}{=} [0,1]$, $\phi_u(x) \stackrel{\text{def}}{=} ux$, $\psi_u(x) \stackrel{\text{def}}{=} x + u(1-x)$ for $u \in [0,1]$ and

$$F(x) \;=\; Z\phi_U(x) + (1-Z)\psi_U(x) \tag{2.2}$$

for independent random variables U, Z with a uniform distribution on $[0,1]$ and a Bernoulli(1/2) distribution, respectively. It is not difficult to verify that $(M_n)_{n\geq 0}$ satisfies the assumptions of Theorem 2.1 and has stationary distribution $\pi =$Beta$(1/2,1/2)$ with Lebesgue density $f(x) = \frac{1}{\pi\sqrt{x(1-x)}}$ on $(0,1)$, also called arcsine distribution. (Plainly, π in the denominator of f means the constant 3.14...) Now observe that $P(x,\cdot)$ is a mixture of a uniform distribution on $[0,x]$ and a uniform distribution on $[x,1]$. So it possesses a $\lambda\!\!\lambda$-continuous component for each $x \in [0,1]$. Theorem 2.1 and 2.2 therefore imply the Harris recurrence of $(M_n)_{n\geq 0}$ on $\mathbb{H} = \mathbb{X} = [0,1]$. The conclusion remains true in the biased case where Z has a Bernoulli(p) distribution for some $p \neq 1/2$. The stationary distribution in this case is a Beta(p,q) distribution with Lebesgue density $\frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} x^{p-1}(1-x)^{q-1}$ on $(0,1)$, where $q \stackrel{\text{def}}{=} 1-p$ and Γ is Euler's gamma function.

(b) Another well-studied example is the *autoregressive process of order 1 (AR(1)-process)*

$$M_n \;=\; \rho M_{n-1} + \theta_n, \quad n \geq 1 \tag{2.3}$$

on state space $\mathbb{X} = \mathbb{R}$, where $\theta_1, \theta_2, ...$ are i.i.d. with distribution $\mathbb{G}$, say. Hence

$$F(x) \;=\; \rho x + \theta \tag{2.4}$$

with $\theta \sim \mathbb{G}$. Here "$\sim$" means equality in distribution. Given $|\rho| < 1$ and $\mathbb{E}\log^+|\theta_1| < \infty$, the conditions of Theorem 2.1 hold and the stationary distribution is easily identified as the distribution of $\sum_{n\geq 0}\rho^n\theta_n$. Now, if $\mathbb{G}$ possesses an absolutely continuous component, then so does $P(x,\cdot) = \mathbb{G}(\cdot - \rho x)$ for every $x \in \mathbb{R}$. We hence infer the positive Harris recurrence of $(M_n)_{n\geq 0}$ on $\mathbb{H} = \mathbb{X} = \mathbb{R}$.

(c) The so-called *threshold AR(1)-process* is like the previous example of some interest in time series analysis, see [12, Example 4.4]. Consider

$$M_n = \begin{cases} \rho_+ M_{n-1} + \theta_n, & \text{if } M_{n-1} > 0 \\ \rho_- M_{n-1} + \theta_n, & \text{if } M_{n-1} \leq 0 \end{cases}, \quad n \geq 1 \tag{2.6}$$

on $\mathbb{X} = \mathbb{R}$, where $\theta_1, \theta_2, ...$ are again i.i.d. with some distribution $\mathbb{G}$. Obviously,

$$F(x) = (\rho_+ x + \theta)\mathbf{1}_{(0,\infty)}(x) + (\rho_- x + \theta)\mathbf{1}_{(-\infty,0]}(x) \tag{2.7}$$

with $\theta \sim \mathbb{G}$ in this case. Similar to the previous example, the conditions of Theorem 2.1 hold if $|\rho_+| < 1, |\rho_-| < 1$ and $\mathbb{E}\log^+|\theta_1| < \infty$. The positive Harris recurrence of the chain on $\mathbb{H} = \mathbb{X} = \mathbb{R}$ follows if $\mathbb{G}$ possesses an absolutely continuous component. However, concerning our assumptions on ρ_+, ρ_-, a stronger result was obtained by Petrucelli and Woolford [11]. Given $\mathbb{E}\theta_1 = 0$ and the absolute continuity of $\mathbb{G}$ with everywhere positive Lebesgue density, they showed that $(M_n)_{n\geq 0}$ is actually a positive Harris chain iff $\rho_+ < 1, \rho_- < 1$ and $\rho_+\rho_- < 1$.

(d) Let us next take a look at *matrix recursions* which have been studied by many authors, see 2.2 in [5] and the references given there. The defining equation is

$$M_n = \boldsymbol{A}_n M_{n-1} + B_n, \quad n \geq 1 \tag{2.8}$$

on $\mathbb{X} = \mathbb{R}^m$ for some $m \geq 1$, where $(\boldsymbol{A}_1, B_1), (\boldsymbol{A}_2, B_2), ...$ are i.i.d.; $\boldsymbol{A}_n$ is a $m \times m$ matrix and B_n a $m \times 1$ vector. So the associated random Lipschitz map is

$$F(x) = \boldsymbol{A}x + B \tag{2.9}$$

with $(\boldsymbol{A}, B)$ being a generic copy of $(\boldsymbol{A}_1, B_1)$. Let $\|\cdot\|$ be any norm on $\mathbb{R}^m$, define $\|\boldsymbol{A}\| \stackrel{\text{def}}{=} \sup\{\|\boldsymbol{A}x\|; x \in \mathbb{R}^m, \|x\| \leq 1\}$ for $m \times m$ matrices $\boldsymbol{A}$ and suppose that

$$\mathbb{E}\log^+\|\boldsymbol{A}\| < \infty \quad \text{and} \quad \mathbb{E}\log^+\|B\| < \infty.$$

Suppose further an a.s. negative Liapunov exponent l^*, here given by

$$l^* = \inf\{n^{-1}\mathbb{E}\log\|\boldsymbol{A}_1 \cdot ... \cdot \boldsymbol{A}_n\|; n \geq 1\}.$$

Then the conditions of Theorem 2.1 are satisfied (with $x_0 = 0$) whence, by Elton's result (for this situation earlier obtained by Vervaat [13], Brandt [3], see also [2] for a converse), M_n possesses a unique stationary distribution π which is the distribution of any solution M_∞ of the stochastic fixed point equation $M_\infty \sim \boldsymbol{A}M_\infty + B$, where $(\boldsymbol{A}, B)$ and M_∞ are independent. As one can easily see, we may take

$$M_\infty \;=\; \sum_{n\ge 1}\left(\prod_{k=1}^{n-1}\boldsymbol{A}_k\right)B_n. \tag{2.10}$$

If we now additionally assume that $(\boldsymbol{A}, B)$ is nonsingular with respect to $\lambda\!\!\lambda^{m\times m}\otimes\lambda\!\!\lambda^m$, then all $P(x,\cdot)$, $x\in\mathbb{R}^m$, are evidently nonsingular with respect to $\lambda\!\!\lambda^m$, whence Theorems 2.1 and 2.2 show the positive Harris recurrence of $(M_n)_{n\ge 0}$ on whole $\mathbb{X}=\mathbb{R}^m$. The same conclusion holds true provided that $\boldsymbol{A}$, B are independent and B is nonsingular with respect to $\lambda\!\!\lambda^m$.

(e) Let us finally look at an example, in fact a one-dimensional special case of the previous one ($\boldsymbol{A} = (a)$ and $B_n = \theta_n$) and again taken from [5], with a negative answer as to Harris recurrence. Put $f_0(x) \overset{\text{def}}{=} ax-1$, $f_1(x) \overset{\text{def}}{=} ax+1$ for $x\in\mathbb{R}$ and some $a\in(0,1)$ and consider

$$F(x) \;=\; f_\theta(x) \tag{2.11}$$

where θ is 0 or 1 with probability 1/2 each. The associated IFS $(M_n)_{n\ge 0}$ with state space $\mathbb{X}=\mathbb{R}$ thus satisfies the recursive equation

$$M_n \;=\; aM_{n-1}+\theta_n, \quad n\ge 1 \tag{2.12}$$

where $\theta_1,\theta_2,\ldots$ are independent symmetric variables on $\{-1,1\}$. Its unique stationary distribution π is the distribution of the infinite series $\sum_{n\ge 1}a^{n-1}\theta_n$. It is known that π is continuous for every $a\in(0,1)$, singular for $a\in(0,1/2)\cup N$, $N\subset(1/2,1)$ a nonempty $\lambda\!\!\lambda$-null set, and absolutely continuous, otherwise. If $a=1/2$, π is the uniform distribution on $[-2,2]$. The question which values of $a\in(1/2,1)$ give a singular π remains open, see 2.5 in [5] for further information and references.

We claim that $(M_n)_{n\ge 0}$ is never Harris recurrent. If it were, by Theorem 2.1, we could find a π-positive set $\mathfrak{X}_0$, necessarily uncountable because π is continuous, such that the

$$P(x,\cdot) \;=\; \frac{1}{2}\delta_{ax+1}+\frac{1}{2}\delta_{ax-1}, \quad x\in\mathfrak{X}_0$$

were dominated by some σ-finite measure λ. By a well-known result of Halmos and Savage [7], we could then find a countable subset $\mathfrak{X}_1$ of $\mathfrak{X}_0$ such that $(P(x,\cdot))_{x\in\mathfrak{X}_0}$ and $(P(x,\cdot))_{x\in\mathfrak{X}_1}$ were equivalent, that is $P(x,N)=0$ for all $x\in\mathfrak{X}_0$ iff $P(x,N)=0$ for all $x\in\mathfrak{X}_1$. On the other hand, given any countable $\mathfrak{X}_1=\{x_n; n\ge 1\}$, the set of x such that $P(x,\cdot)$ is nonsingular with respect to some $P(x_n,\cdot)$ is easily identified as $\mathfrak{X}_1\cup\{x\in\mathbb{X}: x=x_n\pm\frac{2}{a}$ for some $n\}$ which is again countable. Consequently, the uncountable $\mathfrak{X}_0$ contains elements x such that $P(x,\cdot)$ is orthogonal to each $P(x_n,\cdot)$, a contradiction to the equivalence of $(P(x,\cdot))_{x\in\mathfrak{X}_0}$ and $(P(x,\cdot))_{x\in\mathfrak{X}_1}$. We conclude with the help of Corollary 2.5 that M_n converges to π in distribution under every $\mathbb{P}_x$ while convergence in total variation fails to hold for π-almost all x.

3. The Rate of Convergence Towards Stationarity

As already mentioned above, Theorem 2.1(c) implies, by invoking the ergodic theorem for aperiodic, positive Harris chains, see [9, Theorem 13.0.1], that

$$\lim_{n\to\infty} \|\mathbb{P}_x(M_n \in \cdot) - \pi\| \;=\; 0 \tag{3.1}$$

for all $x \in \mathbb{H}$ where $\|\cdot\|$ denotes the total variation distance. A weaker metric considered in [5] and [1] is the Prokhorov metric associated with d. Following [5], the latter is also denoted d and defined, for two probability measures λ_1, λ_2 on $\mathbb{X}$, as the infimum over all $\delta \geq 0$ such that

$$\lambda_1(B) < \lambda_2(B^\delta) + \delta \quad \text{and} \quad \lambda_2(B) < \lambda_1(B^\delta) + \delta$$

for all $B \in \mathfrak{B}(\mathbb{X})$, where $B^\delta \stackrel{\text{def}}{=} \{x \in \mathbb{X} : d(x,y) < \delta \text{ for some } y \in B\}$. It has been shown in [1] that, for each $p > 0$,

$$E \log^+ L_1 < 0 \tag{3.2}$$

together with

$$\mathbb{E}\log^{p+1}(1 + L_1) < \infty \quad \text{and} \quad \mathbb{E}\log^{p+1} d(F_1(x_0), x_0) < \infty \tag{3.3}$$

implies

$$\int_{\mathbb{X}} \log^p(1 + d(x, x_0))\; \pi(dx) \;<\; \infty \tag{3.4}$$

and

$$d(P^n(x,\cdot), \pi) \;\leq\; A_x(n+1)^{-p} \tag{3.5}$$

for all $x \in \mathbb{X}$, $n \geq 0$ and a constant A_x of the form $\max\{A, 2d(x, x_0)\}$ for some $A \in (0,\infty)$ and $x_0 \in \mathbb{X}$. If (3.2) and

$$\mathbb{E}L_1^p < \infty \quad \text{and} \quad \mathbb{E}d(F_1(x_0), x_0)^p < \infty \tag{3.6}$$

for some $p > 0$ are satisfied, then

$$\int_{\mathbb{X}} d(x, x_0)^\eta\; \pi(dx) \;<\; \infty \tag{3.7}$$

for some $0 < \eta \leq p$ and

$$d(P^n(x,\cdot), \pi) \;\leq\; A_x r^n \tag{3.8}$$

hold true for all $x \in \mathbb{X}$, $n \geq 0$, some $r \in (0,1)$ not depending on x and a constant A_x of the same form as in (3.5). This result is due to Diaconis and Freedman [5] and reproved in [1] by different methods. If $(M_n)_{n\geq 0}$ is Harris recurrent, it is natural to ask in view of (3.5) and (3.8), whether or not similar conclusions hold when replacing the Prokhorov distance with the total variation distance. The positive answer is provided in Theorem 3.2 for the case $\mathbb{H} = \mathbb{X}$ and under the additional assumption that the support of the stationary distribution π has nonempty interior.

Weaker conclusions, stated as Theorem 3.1, can be drawn much more easily from (3.4) and (3.7) concerning the f-regularity of $(M_n)_{n\ge 0}$. Following [9], a set $C \in \mathfrak{B}(\mathbb{X})$ is called *f-regular* for a function $f : \mathbb{X} \to [1,\infty)$ if for each π-positive $B \in \mathfrak{B}(\mathbb{X})$

$$\sup_{x\in C} \mathbb{E}_x \left(\sum_{n=0}^{\varrho(B)-1} f(M_n) \right) \;<\; \infty,$$

where $\varrho(B) \stackrel{\text{def}}{=} \inf\{n \ge 1 : M_n \in B\}$. $(M_n)_{n\ge 0}$ is called f-regular on a P-absorbing set $\mathbb{H}$ if it is π-irreducible and $\mathbb{H}$ admits a countable cover of f-regular sets. Defining the f-norm $\|\nu\|_f$ for a signed measure ν as

$$\|\nu\|_f \;\stackrel{\text{def}}{=}\; \sup_{|g|\le f} |\nu(g)|, \quad \nu(g) \;\stackrel{\text{def}}{=}\; \int g \, d\nu,$$

$(M_n)_{n\ge 0}$ is called *f-ergodic* on $\mathbb{H}$ if it is positive Harris on $\mathbb{H}$ with invariant distribution π satisfying $\pi(f) < \infty$ and if

$$\lim_{n\to\infty} \|P^n(x,\cdot) - \pi\|_f \;=\; 0$$

for all $x \in \mathbb{H}$. Now put

$$f(x) \;\stackrel{\text{def}}{=}\; 1 + \log^p(1 + d(x,x_0)) \tag{3.9}$$

if (3.3) holds for $p > 0$, and

$$f(x) \;\stackrel{\text{def}}{=}\; 1 + d(x,x_0)^\eta \tag{3.10}$$

if (3.6) holds for $p > 0$ and with η defined by (3.7). These will be standing definitions in the sequel. Observe that (3.4) and (3.7) may then be restated as $\pi(f) < \infty$. By using Meyn and Tweedie's main result on f-regularity, see [9, Theorem 14.3.3], the following result is now immediate and hence stated without proof.

THEOREM 3.1. *Let $(M_n)_{n\ge 0}$ be an IFS of i.i.d. Lipschitz maps satisfying (1.3) and (1.4). Suppose further that $(M_n)_{n\ge 0}$ is an aperiodic positive Harris chain on a π-full, absorbing set $\mathbb{H}$ and that either (3.3) or (3.6) holds for some $p > 0$. Then $\mathbb{H}$ may be chosen such that $(M_n)_{n\ge 0}$ is f-regular and f-ergodic on $\mathbb{H}$ with f according to (3.9), respectively (3.10).*

It is to be understood that the Harris set $\mathbb{H}$ on which $(M_n)_{n\ge 0}$ is f-regular need not be the maximal Harris set.

Let us now turn to the rate of convergence towards stationarity for $(M_n)_{n\ge 0}$. As already mentioned above, our result will need $\mathbb{H} = \mathbb{X}$ and the interior of *supp* π be nonempty. In order to get some feeling for the problem, the reader should notice that the contractive behavior of an IFS $(M_n)_{n\ge 0}$, even in the strictly contractive case $L_1 \le \gamma < 1$ a.s., is a topological property which does not automatically translate into rapid convergence in the very strong total variation norm. When thinking in terms of coupling rates, the latter property means that two appropriately constructed versions $(M_n^x)_{n\ge 0}$ and $(M_n^y)_{n\ge 0}$ of $(M_n)_{n\ge 0}$ with different starting points $x, y \in \mathbb{H}$ may be *exactly coupled* in very short time, i.e. $M_n^x = M_n^y$ for all $n \ge T$

with a coupling epoch T satisfying some high order moment condition. Strict contraction, on the other hand, only ensures that the distance $d(M_n^x, M_n^y)$ can be made decreasing at a geometric rate (by just choosing $M_n^x = F_{n:1}(x)$ and $M_n^y = F_{n:1}(y)$), but this may come along with one process rapidly entering a small or petite set (the sets for regeneration and thus canonical candidates for coupling attempts, see [9] for definitions) while the other process takes a much longer time to get there so that the overall time it takes to glue both processes together may be large (although they have been close for a long time already). For this unfortunate case to happen we must envisage Harris recurrent IFS for which all small and petite sets look topologically "bad". As it turns out, this means that they all have empty interior. Indeed, the proof of Theorem 3.2 in Section 5 will require the existence of at least one $x \in \mathbb{X}$ and $\varepsilon > 0$ such that the ε-ball $\mathbb{B}_\varepsilon(x)$ is π-positive and (1-)regular, which by [9, Theorem 14.2.4] implies that $\mathbb{B}_\varepsilon(x)$ is petite. Unfortunately, this condition seems difficult to verify in general. We will therefore resort to a further result of Meyn and Tweedie [8, Theorem 3.4] which yields that all compact subsets of $\mathbb{X}$, and thus all ε-balls ($\mathbb{X}$ locally compact) as well, are petite provided $\mathbb{H} = \mathbb{X}$ and $\mathrm{int}(\mathit{supp}\,\pi) \neq \emptyset$. Note, however, that our results still apply if $\mathbb{H}$ is a closed subset of $\mathbb{X}$ and $\mathit{supp}\,\pi$ has nonempty interior in the relative topology on $\mathbb{H}$. We then consider $(M_n)_{n\ge 0}$ as an IFS on the reduced state space $\mathbb{H}$ which inherits all topological properties of $\mathbb{X}$.

THEOREM 3.2. *Let $(M_n)_{n\ge0}$ be an IFS of i.i.d. Lipschitz maps with a.s. negative Liapunov exponent l^* and stationary distribution π. Suppose further that $(M_n)_{n\ge0}$ is a positive Harris chain on all of $\mathbb{X}$ and that $\mathrm{int}(\mathit{supp}\,\pi) \neq \emptyset$. Then the following assertions hold:*

(a) If $(M_n)_{n\ge0}$ satisfies (3.3) for some $p > 0$, then

$$\sum_{n\ge1} n^{p-1} \|\mathbb{P}_x(M_n \in \cdot) - \pi\| \;<\; \infty \tag{3.11}$$

as well as

$$\lim_{n\to\infty} n^p \|\mathbb{P}_x(M_n \in \cdot) - \pi\| \;=\; 0. \tag{3.12}$$

for all $x \in \mathbb{X}$.

(b) If $(M_n)_{n\ge0}$ satisfies (3.6) for some $p > 0$, then

$$\sum_{n\ge0} r^{-n} \|\mathbb{P}_x(M_n \in \cdot) - \pi\|_f \;<\; \infty \tag{3.13}$$

for all $x \in \mathbb{X}$ and some $r \in (0,1)$ not depending on $x \in \mathbb{X}$, where f is defined as in (3.10).

3.3. EXAMPLES. (a) [5, Section 6.3] The *Beta-Walk* is a generalization of Example 2.6(a) and obtained by replacing the uniform variable U in (2.2) by a Beta(α,α) variable V, $\alpha \in [0,\infty]$. Here Beta$(0,0) \stackrel{\text{def}}{=} \frac{1}{2}(\delta_0 + \delta_1)$ and Beta$(\infty,\infty) \stackrel{\text{def}}{=} \delta_{1/2}$. Example 2.6(a) is the case $\alpha = 1$. As one can easily see with Theorem 2.1 and 2.2, $(M_n)_{n\ge0}$ is a positive Harris chain on $\mathbb{X} = [0,1]$ for $\alpha \in (0,\infty]$, but is not for $\alpha = 0$. Diaconis and Freedman [5, Theorem 6.1] show that π equals Beta$(\frac{\alpha}{\alpha+1}, \frac{\alpha}{\alpha+1})$ for $\alpha \in \{0,1,\infty\}$, but differs from it otherwise, although sharing the first three moments. Except for the case $\alpha = 0$, where $\pi = \frac{1}{2}(\delta_0 + \delta_1)$, π is further absolutely continuous with therefore nonempty $\mathrm{int}(\mathit{supp}\,\pi)$. Since $\mathbb{X}$ is compact, condition (3.6) with $x_0 = 0$ holds for every $p > 0$, whence Theorem 3.2 implies geometric ergodicity of the chain for every $\alpha \in (0, \infty)$. If $\alpha = 0$, the same conclusion yields by observing that, starting from any $x \in [0, 1]$, it takes a geometric time to enter the absorbing closed Harris set $\mathit{supp}\,\pi = \{0, 1\}$ and that Theorem 3.2 gives geometric ergodicity on that set.

(b) We next take a look at *multivariate ARMA* models as discussed in [12, Section 4.3.1] and [2, Section 4], where our presentation follows the latter source. Example 2.6(b) forms a very simple, univariate special case of such models. Given real matrices $\boldsymbol{F}_i$, $1 \le i \le k$, and $\boldsymbol{G}_j$, $0 \le j \le l$, of dimension $d \times d$ and $d \times m$, respectively, a $\mathbb{R}^d$-valued random process $(Y_n)_{n>-k}$ is a *nonanticipative solution* of an ARMA(k, l) equation, if

$$Y_n \;=\; \sum_{i=1}^{k} \boldsymbol{F}_i Y_{n-i} \;+\; \sum_{j=0}^{l} \boldsymbol{G}_j \theta_{n-j}, \quad n \ge 1, \tag{3.14}$$

where θ_n, $n > -l$, are i.i.d. $\mathbb{R}^m$-valued and independent of $(Y_{-k+1}, ..., Y_0)$.

With $\boldsymbol{I}_d$ denoting the d-dimensional identity matrix, put

$$\boldsymbol{F}(z) \;\stackrel{\text{def}}{=}\; \boldsymbol{I}_d \;-\; \sum_{i=1}^{k} \boldsymbol{F}_i z^i \quad \text{and} \quad \boldsymbol{G}(z) \;\stackrel{\text{def}}{=}\; \sum_{j=1}^{l} \boldsymbol{G}_j z^j, \quad z \in \mathbb{C}.$$

Suppose that $\boldsymbol{F}(z)^{-1}\boldsymbol{G}(z)$ is irreducible in the sense that every matrix function $\boldsymbol{D}(z)$ which is a common left divisor of $\boldsymbol{F}(z)$ and $\boldsymbol{G}(z)$, has a constant determinant. It is shown in [2, Theorem 4.1] that, under the assumptions

(b.1) θ_1 is not carried by a fixed hyperplane,

(b.2) $\mathbb{E}\log^+\|\theta_1\| < \infty$ ($\|\cdot\|$ any norm on $\mathbb{R}^m$),

and

(b.3) $\boldsymbol{F}(z)^{-1}\boldsymbol{G}(z)$ is irreducible,

(3.14) possesses a stationary nonanticipative solution iff

(b.4) all zeros of the polynomial $\det\boldsymbol{F}(z)$ lie outside the closed unit disk of $\mathbb{C}$.

Let ${}^t\boldsymbol{M}$ denote the transpose of a matrix $\boldsymbol{M}$. In order to discuss (3.14) within our framework of IFS of i.i.d. Lipschitz functions, we first need a so-called state space representation of the model, given by

$$M_{n+1} \;=\; \boldsymbol{A}M_n \;+\; \boldsymbol{B}\theta_{n+1}, \quad n \ge 0 \tag{3.15}$$

and

$$Y_{n+1} \;=\; \boldsymbol{C}M_n \;+\; \boldsymbol{D}\theta_{n+1}, \quad n \ge 0, \tag{3.16}$$

where $M_n \stackrel{\text{def}}{=} {}^t({}^tY_n, ..., {}^tY_{n-k+1}, {}^t\theta_n, ..., {}^t\theta_{n-l+1})$,

$$\Psi \stackrel{\text{def}}{=} (\boldsymbol{F}_1, ..., \boldsymbol{F}_{k-1}), \qquad\qquad \Gamma \stackrel{\text{def}}{=} (\boldsymbol{G}_1, ..., \boldsymbol{G}_{l-1}),$$

$$\boldsymbol{A} \stackrel{\text{def}}{=} \begin{pmatrix} \Psi & \boldsymbol{F}_k & \Gamma & \boldsymbol{G}_l \\ \boldsymbol{I}_{(k-1)d} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \boldsymbol{I}_{m(l-1)} & 0 \end{pmatrix}, \qquad \boldsymbol{B} \stackrel{\text{def}}{=} \begin{pmatrix} \boldsymbol{G}_0 \\ 0 \\ \boldsymbol{I}_m \\ 0 \end{pmatrix},$$

$$\boldsymbol{C} \stackrel{\text{def}}{=} (\Psi, \boldsymbol{F}_k, \Gamma, \boldsymbol{G}_l), \qquad\qquad \boldsymbol{D} \stackrel{\text{def}}{=} \boldsymbol{G}_0,$$

see [2]. The dimensions of $\boldsymbol{A}$ and $\boldsymbol{B}$ are $(kd+lm)\times(kd+lm)$ and $(kd+lm)\times m$, respectively.

Now, $(M_n)_{n\geq 0}$ clearly defines an IFS of i.i.d. Lipschitz maps. Under conditions (b.1–4), which are always assumed hereafter, its Liapunov exponent is a.s. negative. Indeed, all eigenvalues of $\boldsymbol{A}$ have modulus less than 1, as following from [2, Theorem 4.1] in combination with [2, eq. (15)]

$$\boldsymbol{F}(z)^{-1}\boldsymbol{G}(z) \;=\; \boldsymbol{C}(z^{-1}\boldsymbol{I}_{kd+lm} - \boldsymbol{A})^{-1}\boldsymbol{B} \;+\; \boldsymbol{D}.$$

The latter implies that z is an eigenvalue of $\boldsymbol{A}$ iff $\det\boldsymbol{F}(1/z) = 0$.

Following [2], let $\mathbb{H} = w + \mathbb{V}$ be the minimal affine subspace of $\mathbb{R}^{kd+lm}$ such that the stationary distribution π of $(M_n)_{n\geq 0}$ is concentrated on $\mathbb{H}$. Since π is unique, $\mathbb{H}$ is also the affine subspace of minimal dimension which is invariant for $(M_n)_{n\geq 0}$, i.e. $\boldsymbol{A}(w+\mathbb{V}) + \boldsymbol{B}\theta_n \subset w + \mathbb{V}$ a.s. It follows that $\boldsymbol{A}w - w \in \mathbb{V}$, $\boldsymbol{A}\mathbb{V} \subset \mathbb{V}$ and $\mathrm{Im}(\boldsymbol{B}) \subset \mathbb{V}$. Hence, for some invertible matrix $\boldsymbol{S}$, $\boldsymbol{S}M_n$ takes the form $\boldsymbol{S}M_n = {}^t({}^t\tilde{M}_n, 0)$ for each $n \geq 0$ and satisfies

$$\boldsymbol{S}M_{n+1} \;=\; \boldsymbol{SAS}^{-1}\boldsymbol{S}M_n \;+\; \boldsymbol{SB}\theta_{n+1}, \quad n \geq 0, \tag{3.17}$$

with

$$\boldsymbol{SAS}^{-1} \;=\; \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} \\ 0 & \boldsymbol{A}_{22} \end{pmatrix}, \quad \boldsymbol{SB} \;=\; \begin{pmatrix} \boldsymbol{B}_1 \\ 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{S}(\boldsymbol{A}w - w) \;=\; \begin{pmatrix} x_1 \\ 0 \end{pmatrix},$$

where $\boldsymbol{A}_{11}$, $\boldsymbol{B}_1$ and x_1 have dimensions $s \times s$, $s \times d$ and $s \times 1$, respectively ($s \stackrel{\text{def}}{=} \dim(\mathbb{V})$). Now, $(\tilde{M}_n)_{n\geq 0}$ satisfies

$$\tilde{M}_{n+1} \;=\; \boldsymbol{A}_{11}\tilde{M}_n \;+\; \boldsymbol{B}_1\theta_{n+1} \;+\; x_1, \quad n \geq 0, \tag{3.18}$$

and forms again an IFS of i.i.d. Lipschitz maps with a.s. negative Liapunov exponent. Its state space is $\mathbb{R}^s$ and its unique stationary distribution $\tilde{\pi} \sim \sum_{n\geq 1} \boldsymbol{A}_{11}^{n-1}(\boldsymbol{B}_1\theta_n + x_1)$ is not concentrated on any proper affine subspace of $\mathbb{R}^s$, by minimality of $\mathbb{H}$. Consequently, the linear hull of $\mathrm{Im}(\boldsymbol{B}_1)$, $\mathrm{Im}(\boldsymbol{A}_{11}\boldsymbol{B}_1)$,..., $\mathrm{Im}(\boldsymbol{A}_{11}^{s-1}\boldsymbol{B}_1)$ must be $\mathbb{R}^s$ and, by (b.1) and the independence of the θ_n, the distribution of $\sum_{n=1}^{s} \boldsymbol{A}_{11}^{n-1}\boldsymbol{B}_1\theta_n$ cannot be carried by a fixed hyperplane of $\mathbb{R}^s$. Replacing (b.1) with the stronger condition

(b.1') the distribution of θ_1 is nonsingular with respect to $\lambda\!\!\!\lambda^d$, we infer that the distribution of $\sum_{n=1}^{s} \boldsymbol{A}_{11}^{n-1}\boldsymbol{B}_1\theta_n$ is nonsingular with respect to $\lambda\!\!\lambda^s$. Since further

$$\tilde{M}_{n+s} \;=\; \boldsymbol{A}_{11}^{s}\tilde{M}_n \;+\; \sum_{j=1}^{s}\boldsymbol{A}_{11}^{s-j}\boldsymbol{B}_1\theta_{n+j} \;+\; \sum_{j=0}^{s-1}\boldsymbol{A}_{11}^{j}x_1, \quad n\ge 0,$$

validity of condition (c) of Theorem 2.1 with $\mathfrak{X}=\mathbb{R}^s$ and $\lambda=\lambda\!\!\lambda^s$ is now easily verified. This in turn shows the positive Harris recurrence of $(\tilde{M}_n)_{n\ge 0}$ on $\mathbb{X}=\mathbb{R}^s$, by invoking Theorems 2.1 and 2.2(b). Note that *supp* $\tilde{\pi}$ has nonempty interior because $\tilde{\pi}$ dominates a $\lambda\!\!\lambda^s$-continuous measure, see Remarks D and E. We further infer that part (a), respectively part (b) of Theorem 3.2 applies to $(\tilde{M}_n)_{n\ge 0}$, provided that

(b.5) $\quad \mathbb{E}\log^{p+1}(1+\|\boldsymbol{B}_1\theta_1\|)<\infty,$

respectively

(b.6) $\quad \mathbb{E}\|\boldsymbol{B}_1\theta_1\|^p<\infty$

are satisfied in addition to (b.1') and (b.2–4). The function f arising in (3.13) here takes the form $f(x)=1+\|x\|^p$ for an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^s$. Corresponding conclusions for the ARMA process $(Y_n)_{n\ge 0}$ then follow under the same respective conditions because

$$\begin{aligned}\|\mathbb{P}(\tilde{M}_n\in\cdot|\tilde{M}_0=x)-\tilde{\pi}\| \;&=\; \|\mathbb{P}(\boldsymbol{S}M_n\in\cdot|\boldsymbol{S}M_0={}^t(x,0))-\tilde{\pi}\otimes\delta_0^{(kd+lm)-s}\|\\ &=\; \|\mathbb{P}(M_n\in\cdot|M_0=\boldsymbol{S}^{-1}x)-\pi\|\\ &\ge\; \|\mathbb{P}(Y_n\in\cdot|M_0=\boldsymbol{S}^{-1}x)-\xi\|\end{aligned}$$

and

$$\|\mathbb{P}(\tilde{M}_n\in\cdot|\tilde{M}_0=x)-\tilde{\pi}\|_f \;\ge\; \|\mathbb{P}(Y_n\in\cdot|M_0=\boldsymbol{S}^{-1}x)-\xi\|_g$$

where $g(y)\stackrel{\text{def}}{=}c(1+\|y\|^p)$ for an arbitrary norm $\|\cdot\|$ on $\mathbb{R}^d$ and a sufficiently small constant $c>0$ depending on the function f.

(c) Returning to Example 2.6(d) of *matrix recursions* $M_n=\boldsymbol{A}_nM_{n-1}+B_n$, $n\ge 1$, with a.s. negative Liapunov exponent and satisfying

(c.1) $\quad \mathbb{E}\log^+\|\boldsymbol{A}\|<\infty$ and $\mathbb{E}\log^+\|B\|<\infty$,

we recall that its positive Harris recurrence on whole $\mathbb{X}=\mathbb{R}^m$ follows if further

(c.2) $\quad (\boldsymbol{A}_1,B_1)$ is nonsingular with respect to $\lambda\!\!\lambda^{m\times m}\otimes\lambda\!\!\lambda^m$,

or

(c.2') $\quad \boldsymbol{A}_1$, B_1 are independent and B_1 is nonsingular with respect to $\lambda\!\!\lambda^m$

holds true. Given any $p>0$, it is then immediate to conclude the assertion of Theorem 3.2(a) and of 3.2(b) with $f(x)=\|x\|^p$, provided that additionally

(c.3) $\quad \mathbb{E}\log^{p+1}(1+\|\boldsymbol{A}_1\|)<\infty$ and $\mathbb{E}\log^{p+1}(1+\|B_1\|)<\infty$,

respectively

(c.4) $\quad \mathbb{E}\|\boldsymbol{A}_1\|^p<\infty$ and $\mathbb{E}\|B_1\|^p<\infty$.

(d) We finally consider *stochastic difference equations with additive noise*

$$M_{n+1} \;=\; T(M_n) \;+\; \theta_{n+1}, \quad n\ge 0,$$

for which ergodicity results were obtained by Chan and Tong [4]. Here $T:\mathbb{R}^m\to\mathbb{R}^m$ is a (nonrandom) Lipschitz function and the θ_n, $n\ge 1$, are i.i.d. random vectors in $\mathbb{R}^m$. Hence $(M_n)_{n\ge 0}$ is an IFS of i.i.d. Lipschitz maps F_n, $n\ge 1$, with generic element $F(x)=T(x)+\theta$. As one can easily see, it satisfies Elton's conditions (1.3), (1.4) and is thus ergodic provided that

(d.1) T is strictly contractive, i.e. $l(T)<1$;

(d.2) $\mathbb{E}\log^+\|\theta_1\|<\infty$.

If furthermore

(d.3) θ_1 is nonsingular with respect to m-dimensional Lebesgue measure $\lambda\!\!\lambda^m$,

we infer, for some $\lambda\!\!\lambda^m$-positive $\mathfrak{X}'$, the $\lambda\!\!\lambda^m(\cdot\cap\mathfrak{X}')$-irreducibility and the positive Harris recurrence of $(M_n)_{n\ge 0}$ on some absorbing set $\mathbb{H}$ (Remark D and Theorem 2.1 with $\mathfrak{X}=\mathbb{X}=\mathbb{R}^m$). Moreover, $\mathrm{int}(\mathit{supp}\,\pi)\ne\emptyset$ because π (as a maximal irreducibility measure) dominates $\lambda\!\!\lambda^m(\cdot\cap\mathfrak{X}')$. Hence $\mathbb{H}=\mathbb{R}^m$ by Theorem 2.2(b). Now Theorem 3.2(a), respectively 3.2(b) with $f(x)=\|x\|^p$ applies, provided additionally

(d.4) $\mathbb{E}\log^{p+1}(1+\|\theta_1\|)<\infty$,

respectively

(d.5) $\mathbb{E}\|\theta_1\|^p<\infty$.

Notice that the threshold $AR(1)$-model in 2.6(c) with $|\rho_+|<1$ and $|\rho_-|<1$ is a special case of the given model.

Geometric ergodicity of $(M_n)_{n\ge 0}$ is shown in [4] for the situation where (d.1) and (d.3) are replaced with

(d.1') $T(0)=0$ and T is exponentially stable in the large, i.e. $\|T^n(x)\|\le ke^{-cn}\|x\|$ for all $x\in\mathbb{R}^m$ and some constants $K,c\in(0,\infty)$,

(d.3') θ_1 is absolutely continuous with respect to $\lambda\!\!\lambda^m$ with an everywhere positive density,

and where (d.5) holds with $p=1$. Plainly, (d.1') is weaker than (d.1), while (d.3') is stronger than (d.3) and ensures the irreducibility of the model. The Harris recurrence and geometric ergodicity may then be inferred from (d.1') by showing the existence of a suitable Liapunov function and applying Theorem 15.0.1 in [9], see [4] for details and also Section 5.

4. Proofs of Theorems 2.1 and 2.2

Proof of Theorem 2.1. Since "(a)⇒(b)" is trivial we first show "(b)⇒(c)". So we suppose (b) which provides us with the existence of some $n_0\ge 1$, a non-zero σ-finite measure $\lambda^{(n_0)}$, a product measurable function $g_{n_0}:\mathbb{X}^2\to[0,\infty)$ and a substochastic kernel $Q^{(n_0)}$ such that

$$P^{n_0}(x,dy)\;=\;g_{n_0}(x,y)\lambda^{(n_0)}(dy)\;+\;Q^{(n_0)}(x,dy),\quad x\in\mathbb{X},\tag{4.1}$$

where $\int_{\mathbb{X}}g_{n_0}(x,y)\,\lambda^{(n_0)}(dy)>0$ for all $x\in\mathfrak{X}$. By stationarity,

$$\pi(dy)\;=\;\overline{g}_{n_0}(y)\lambda^{(n_0)}(dy)+\pi Q^{(n_0)}(dy)$$

where $\overline{g}_{n_0}(y) \stackrel{\text{def}}{=} \int g_{n_0}(x,y)\,\pi(dx)$. So $\pi \geq \overline{\lambda}^{(n_0)} \stackrel{\text{def}}{=} \overline{g}_{n_0}\lambda^{(n_0)}$ and the latter measure is nonzero because

$$\int \overline{g}_{n_0}(y)\ \lambda^{(n_0)}(dy)\ \geq\ \int_{\mathfrak{X}}\int g_{n_0}(x,y)\ \lambda^{(n_0)}(dy)\ \pi(dx)\ >\ 0.$$

Possibly after replacing $\mathfrak{X}$ with the π-positive set $\mathfrak{X}\cap\{\overline{g_{n_0}}>0\}$, we may thus assume hereafter that $\lambda^{(n_0)}$ is absolutely continuous w.r.t. π.

For an arbitrary stopping time $\tau = h((M_n)_{n\geq 0})$ for $(M_n)_{n\geq 0}$, let $\vartheta^k\tau \stackrel{\text{def}}{=} k+h((M_{k+n})_{n\geq 0})$ for $k\in\mathbb{N}_0$ and $P^{(\tau)}(x,\cdot) \stackrel{\text{def}}{=} \mathbb{P}_x(M_\tau\in\cdot)$ for $x\in\mathbb{X}$, hence $P^{(\tau)} = P^\tau$ if τ is a.s. constant. We claim that, given any $\mathbb{P}_\pi$-a.s. finite stopping time τ for $(M_n)_{n\geq 0}$, the family $(P^{(\vartheta^{n_0}\tau)}(x,\cdot))_{x\in\mathfrak{X}}$ is nonsingular with respect to $\lambda^{(n_0,\tau)}$ defined through

$$\lambda^{(n_0,\tau)}(B)\ \stackrel{\text{def}}{=}\ \int_{\mathbb{X}}\mathbb{P}(F_{\tau:1}(y)\in B)\ \lambda^{(n_0)}(dy),\quad B\in\mathfrak{B}(\mathbb{X}).$$

This follows because

$$\begin{aligned} P^{(\vartheta^{n_0}\tau)}(x,B)\ &=\ \int_{\mathbb{X}} P^{(\tau)}(y,B)g_{n_0}(x,y)\ \lambda^{(n_0)}(dy)\ +\ \int_{\mathbb{X}} P^{(\tau)}(y,B)\ Q^{(n_0)}(x,dy)\\ &=\ \int_{\mathbb{X}}\mathbb{P}(F_{\tau:1}(y)\in B)g_{n_0}(x,y)\ \lambda^{(n_0)}(dy)\ +\ \int_{\mathbb{X}} P^{(\tau)}(y,B)\ Q^{(n_0)}(x,dy)\end{aligned}$$

for all $B\in\mathfrak{B}(\mathbb{X})$. The measure defined by the first integral in the previous line (as a function of B) is clearly absolutely continuous with respect to $\lambda^{(n_0,\tau)}$ with density $g_{n_0,\tau}$, say. Let $Q^{(n_0,\tau)}$ be the measure defined by the second integral. It follows that

$$P^{(\vartheta^{n_0}\tau)}(x,dy)\ =\ g_{n_0,\tau}(x,y)\lambda^{(n_0,\tau)}(dy)\ +\ Q^{(n_0,\tau)}(x,dy),\quad x\in\mathbb{X}, \tag{4.2}$$

where $\int_{\mathbb{X}} g_{n_0,\tau}(x,y)\,\lambda^{(n_0,\tau)}(dy) > 0$ for all $x\in\mathfrak{X}$. We have thus particularly shown that, if (c) holds for some $n_0\geq 1$, then it also holds for all $n\geq n_0$ (with the same $\mathfrak{X}$ but in general different λ).

Next recall from (1.6) that, for each π-positive B,

$$\mathbb{X}_0(B)\ \stackrel{\text{def}}{=}\ \{x\in\mathbb{X}:\mathbb{P}_x(M_n\in B\text{ i.o.})=1\}$$

satisfies $\pi(\mathbb{X}_0(B))=1$ and thus also $P(x,\mathbb{X}_0(B))=1$ for π-almost all $x\in\mathbb{X}$. Recursively, define

$$\mathbb{X}_{n+1}(B)\ \stackrel{\text{def}}{=}\ \{x\in\mathbb{X}_n(B):P(x,\mathbb{X}_n(B))=1\}$$

for $n\geq 0$. Then $\pi(\mathbb{X}_n(B))=1$ for all $n\geq 0$ and $\mathbb{X}_n(B)\downarrow\mathbb{X}_\infty(B)\stackrel{\text{def}}{=}\cap_{k\geq 0}\mathbb{X}_k(B)$, as $n\to\infty$, giving $\pi(\mathbb{X}_\infty(B))=1$. $\mathbb{X}_\infty(B)$ is further absorbing because, by construction, $P(x,\mathbb{X}_n(B))=1$ for all $x\in\mathbb{X}_\infty(B)$ and $n\geq 0$, and thus $P(x,\mathbb{X}_\infty(B)) = \lim_{n\to\infty}P(x,\mathbb{X}_n(B)) = 1$ for all $x\in\mathbb{X}_\infty(B)$.

Put $\tau\stackrel{\text{def}}{=}\inf\{n\geq 1: M_k\in\mathfrak{X}\}$ and notice that τ is clearly $\mathbb{P}_\pi$-a.s. finite. Since $\pi(\mathbb{X}_\infty(\mathfrak{X})^c)$ $=0$, also $\lambda^{(n_0)}(\mathbb{X}_\infty(\mathfrak{X})^c)=0$. It is now obvious from the previous considerations that we can choose $\delta > 0$ sufficiently small such that

$$\int_{\mathbb{X}_\infty(\mathfrak{X})} \int_{\mathfrak{X}} \int_{\mathbb{X}_\infty(\mathfrak{X})} \mathbf{1}_{\{g_{n_0,\tau} \geq \delta\}}(x,y) \mathbf{1}_{\{g_{n_0} \geq \delta\}}(y,z) \ \lambda^{(n_0)}(dz) \ \lambda^{(n_0,\tau)}(dy) \ \pi(dx) \ > \ 0.$$

Hence we may invoke Lemma 4.3 by Niemi and Nummelin [10] to infer the existence of a π-positive set $\mathfrak{X}_1 \subset \mathbb{X}_\infty(\mathfrak{X})$ and a $\lambda^{(n_0)}$-positive set $\mathfrak{X}_2 \subset \mathbb{X}_\infty(\mathfrak{X})$ satisfying

$$\alpha \ \stackrel{\text{def}}{=} \ \inf_{x \in \mathfrak{X}_1, z \in \mathfrak{X}_2} \lambda^{(n_0,\tau)}(y \in \mathfrak{X} : g_{n_0,\tau}(x,y) \geq \delta, g_{n_0}(y,z) \geq \delta) \ > \ 0.$$

A combination of this result with (4.1) and (4.2) implies

$$\begin{aligned} P^{(\vartheta^{n_0}(\tau+n_0))}(x,B) \ &= \ \int_{\mathfrak{X}} P^{n_0}(y,B) \ P^{(\vartheta^{n_0}\tau)}(x,dy) \\ &\geq \ \int_{\mathfrak{X}} g_{n_0,\tau}(x,y) \int_{B \cap \mathfrak{X}_2} g_{n_0}(y,z) \ \lambda^{(n_0)}(dz) \ \lambda^{(n_0,\tau)}(dy) \\ &\geq \ \alpha\delta^2 \lambda^{(n_0)}(B \cap \mathfrak{X}_2) \end{aligned} \tag{4.3}$$

for all $x \in \mathfrak{X}_1$ and $B \in \mathfrak{B}(\mathbb{X})$. By defining $\mathbb{H} \stackrel{\text{def}}{=} \mathbb{X}_\infty(\mathfrak{X}_1)$, we obtain an absorbing set such that $\mathfrak{X}_1$ is a regeneration set for $(M_n)_{n\geq 0}$ restricted to $\mathbb{H}$, i.e., $\mathfrak{X}_1$ is recurrent and satisfies a minorization condition, namely (4.3). This proves the Harris recurrence of $(M_n)_{n\geq 0}$ on $\mathbb{H}$. Note also that this implies that the chain is $\lambda^{(n_0)}(\cdot \cap \mathfrak{X}_2)$-irreducible (see Remark D).

Since $(M_n)_{n\geq 0}$ possesses a stationary distribution, it is clearly positive Harris recurrent whence we are left with the proof of aperiodicity. However, if $(M_n)_{n\geq 0}$ were q-periodic with cyclic classes $\mathbb{X}_1, ..., \mathbb{X}_q$, say, then the q-skeleton $(M_{nq})_{n\geq 0}$ would have stationary distributions $\frac{\pi(\cdot \cap \mathbb{X}_k)}{\pi(\mathbb{X}_k)}$, $k = 1, ..., q$. On the other hand, the latter is also an IFS of i.i.d. Lipschitz maps satisfying (1.3), (1.4) and thus possesses only one stationary distribution. Consequently, $q = 1$ and $(M_n)_{n\geq 0}$ is aperiodic.

"(c)⇒(b)" If $(M_n)_{n\geq 0}$ is Harris recurrent then, for some π-positive $\mathfrak{X}$ (called small set), there exist $n_0 \geq 1$ and a nonzero measure λ with $\lambda(\mathfrak{X}^c) = 0$ such that $P^{n_0}(x, \cdot) \geq \lambda$ for all $x \in \mathfrak{X}$. Consequently, (b) is satisfied. ♢

Proof of Theorem 2.2. (a) We must show that if $\mathbb{H}$ has a π-positive interior, i.e. $\mathbb{B}_\varepsilon(x) = \{y \in \mathbb{X} : d(x,y) \leq \varepsilon\} \subset \mathbb{H}$ and $\pi(\mathbb{B}_\varepsilon(x)) > 0$ for some $x \in \mathbb{H}$ and $\varepsilon > 0$, then $\mathbb{H} = \mathbb{X}$. Recall that $\varrho(B)$ denotes the first return time to B for $(M_n)_{n\geq 0}$. It suffices to prove $\mathbb{P}_y(\varrho(\mathbb{B}_\varepsilon(x)) < \infty) = 1$ for all $y \in \mathbb{X}$. Pick $\eta \in (0, \varepsilon)$ such that $\pi(\mathbb{B}_\eta(x)) > 0$. As will be seen in the next section, $(M_n)_{n\geq 0}$ contains a subsequence $(M_{\sigma_n})_{n\geq 0}$ which is also an IFS of i.i.d. Lipschitz maps with the same stationary distribution and further strictly contractive, i.e. $L_{1:\sigma_1} \leq \gamma < 1$ a.s. The $\sigma_n = \sigma_n(\gamma)$ are defined in (5.3) and do not depend on M_0. Recalling (1.6), $\mathbb{P}_w(M_{\sigma_n} \in \mathbb{B}_\eta(x) \text{ i.o.}) = 1$ for π-almost all $w \in \mathbb{X}$. Fix any such w and put $M_n^y \stackrel{\text{def}}{=} F_{n:1}(y)$ for $n \geq 0$ and $y \in \mathbb{X}$. Then $(M_n^y)_{n\geq 0}$ is a copy of $(M_n)_{n\geq 0}$ under $\mathbb{P}$, and $d(M_{\sigma_n}^y, M_{\sigma_n}^w) \leq \gamma^n d(y,w)$ a.s. for all $n \geq 0$. Consequently, $\{M_{\sigma_n}^w \in \mathbb{B}_\eta(x)\} \subset \{M_{\sigma_n}^y \in \mathbb{B}_\varepsilon(x)\}$

for all n sufficiently large implying

$$\mathbb{P}(M_n^y \in \mathbb{B}_\varepsilon(x) \text{ i.o.}) \ \geq \ \mathbb{P}(M_{\sigma_n}^y \in \mathbb{B}_\varepsilon(x) \text{ i.o.}) \ \geq \ \mathbb{P}(M_{\sigma_n}^w \in \mathbb{B}_\eta(x) \text{ i.o.}) \ = \ 1$$

and thus the desired $\mathbb{P}_y(\varrho(\mathbb{B}_\varepsilon(x)) < \infty) = 1$ for all $y \in \mathbb{X}$.

(b) Note first that, by the previous argument, we generally have $\mathbb{P}_x(M_n \in B \text{ i.o.}) = 1$ for all $x \in \mathbb{X}$ and all π-positive open $B \subset \mathbb{X}$. Hence $\pi(\text{int}(\mathfrak{X})) > 0$ ensures $\mathbb{P}_x(\varrho(\mathfrak{X}) < \infty) = 1$ for all $x \in \mathbb{X}$, which together with 2.1(a) or (c) easily yields the λ-irreducibility of $(M_n)_{n\geq 0}$. Combining this with $\text{int}(\mathit{supp}\,\pi) \neq \emptyset$, we infer that every compact set is petite [8, Theorem 3.4] and that $(M_n)_{n\geq 0}$ forms a T-chain [8, Theorem 3.2]. Now the positive Harris recurrence on $\mathbb{X}$ follows if $(M_n)_{n\geq 0}$ is also nonevanescent [9, Theorem 9.2.2]. For that latter property, we must show that $\mathbb{P}_x(M_n \in C \text{ i.o.}) = 1$ for all $x \in \mathbb{X}$ and some compact set C. But $\text{int}(\mathit{supp}\,\pi) \neq \emptyset$ implies the existence of a $w \in \mathbb{X}$ with $\pi(\mathbb{B}_\varepsilon(w)) > 0$ for all $\varepsilon > 0$ whence $\mathbb{P}_x(M_n \in \mathbb{B}_\varepsilon(w) \text{ i.o.}) = 1$ for all $\varepsilon > 0$. We obtain the desired conclusion because the closure of some $\mathbb{B}_\varepsilon(w)$ is compact by the local compactness of $\mathbb{X}$. ♢

5. Proof of Theorem 3.2

Throughout this section, we are always given an IFS $(M_n)_{n\geq 0}$ of i.i.d. Lipschitz maps satisfying (1.3) and (1.4). Since the Liapunov exponent l^* is a.s. negative, the definition of l^* implies $\mathbb{E}\log^+ l(F_{m:1}) < 0$ for some $m \geq 1$. Let us assume $m = 1$, that is

$$\mathbb{E}\log^+ l(F_1) \ < \ 0, \tag{5.1}$$

because the modifications of the subsequent arguments in case $m \geq 2$ are totally straightforward but notationally rather tedious. Put $L_n \stackrel{\text{def}}{=} l(F_n)$ for $n \geq 1$ and note that these random variables are i.i.d. and also independent of M_0. Moreover,

$$l(F_{1:n}) \ \leq \ L_{1:n} \ \stackrel{\text{def}}{=} \ \prod_{k=1}^n L_k \quad \text{a.s.} \tag{5.2}$$

Given (5.1), the sequence $(\sum_{k=1}^n \log L_k)_{n\geq 0}$ forms an ordinary zero-delayed random walk with negative drift whence, for any $\gamma \in (0,1)$, the level $\log\gamma$ ladder epochs $\sigma_0(\gamma) \stackrel{\text{def}}{=} 0$ and

$$\begin{aligned} \sigma_n(\gamma) \ &\stackrel{\text{def}}{=} \ \inf\left\{k > \sigma_{n-1}(\gamma) : \sum_{j=\sigma_{n-1}(\gamma)+1}^{k} \log L_j \leq \log\gamma\right\} \\ &= \ \inf\{k > \sigma_{n-1}(\gamma) : L_{\sigma_{n-1}(\gamma)+1:k} \leq \gamma\}, \quad n \geq 1 \end{aligned} \tag{5.3}$$

are all a.s. finite. Now fix any $\gamma \in (0,1)$ and write σ_n for $\sigma_n(\gamma)$. For $(\sigma_n)_{n\geq 0}$ constitutes a renewal process, the random mappings

$$F_n' \stackrel{\text{def}}{=} F_{\sigma_n:\sigma_{n-1}+1}, \quad n \geq 1,$$

are also i.i.d. random Lipschitz functions. They are (a.s.) *strictly (γ-)contractive* because, by (5.2) and the definition of the σ_n,

$$d(F_1'(x), F_1'(y)) \;\le\; L_{1:\sigma_1} d(x,y) \;\le\; \gamma d(x,y) \quad \text{a.s.}$$

for all $x, y \in \mathbb{X}$. Furthermore,

$$\mathbb{E} \log^+ d(F_1'(x_0), x_0) \;<\; \infty$$

as has been shown in [1], see their proof of Lemma 3.2. Consequently, the subsequence $(M_{\sigma_n})_{n\ge 0} = (F_{1:n}'(M_0))_{n\ge 0}$ is an IFS of i.i.d. strictly contractive Lipschitz maps satisfying Elton's conditions. Let P' and π' denote its transition kernel and stationary distribution, respectively. We will show in the following lemma that actually $\pi' = \pi$ holds true.

LEMMA 5.1. *Let $(M_n)_{n\ge 0}$ be an IFS of i.i.d. Lipschitz maps satisfying (1.3) and (1.4). If $(M_n)_{n\ge 0}$ is an aperiodic positive Harris chain on some P-absorbing set $\mathbb{H}$ with stationary distribution π, then, given any $\gamma \in (0,1)$, the same holds true for the subsequence $(M_{\sigma_n})_{n\ge 0}$, i.e., it is also Harris recurrent on $\mathbb{H}$ and has the same stationary distribution.*

It is important for the further derivation that $(M_{\sigma_n})_{n\ge 0}$ does indeed share Harris set and stationary distribution with $(M_n)_{n\ge 0}$ because this guarantees that the additional assumptions $\mathbb{H} = \mathbb{X}$ as well as $\mathrm{int}(\mathit{supp}\,\pi) \ne \emptyset$ carry over from the latter to the former sequence.

PROOF. If $(M_n)_{n\ge 0}$ is aperiodic and positive Harris on $\mathbb{H}$, then $(P^m(x,\cdot))_{x\in\mathbb{H}}$ is nonsingular with respect to π for some $m \ge 1$. Note that $\sigma_m \ge m$ and ${P'}^m = P^{(\sigma_m)}$. As obtained in the proof of Theorem 2.1, the family $({P'}^m(x,\cdot))_{x\in\mathbb{H}}$ is also nonsingular with respect to a non-zero σ-finite measure λ' whence Theorem 2.1 implies that $(M_{\sigma_n})_{n\ge 0}$ is an aperiodic positive Harris chain on some P'-absorbing set $\mathbb{H}'$. Since $\mathbb{H}$ is also P'-absorbing, we can choose $\mathbb{H}' \subset \mathbb{H}$.

Now consider the backward process $\hat{M}_n \stackrel{\text{def}}{=} F_{1:n}(M_0)$, $n \ge 0$, associated with $(M_n)_{n\ge 0}$ which converges a.s. to a random variable $\hat{M}_\infty$ with distribution π, regardless of the initial distribution, see [6]. Consequently, the same holds true for the subsequence $\hat{M}_{\sigma_n}$, $n \ge 0$. Since the F_n are i.i.d. and independent of M_0, we further have for all $n \ge 0$

$$\hat{M}_{\sigma_n} \;=\; F_{1:\sigma_n}(M_0) \;\sim\; F_{\sigma_1:1} \circ \ldots \circ F_{\sigma_n:\sigma_{n-1}+1}(M_0) \;=\; F_{1:n}'(M_0),$$

where $(F_{1:n}'(M_0))_{n\ge 0}$ is the backward process of $(M_{\sigma_n})_{n\ge 0}$. Hence $\hat{M}_{\sigma_n} \to \hat{M}_\infty \sim \pi$ a.s. implies that $F_{1:n}'(M_0)$ converges weakly to π (in fact, it also converges a.s. to some π-distributed random variable) which is therefore also the stationary distribution of $(M_{\sigma_n})_{n\ge 0}$.

In order for $\mathbb{H}' = \mathbb{H}$ it must still be shown that $\mathbb{P}_x(M_{\sigma_n} \in A \text{ i.o.}) = 1$ for all π-positive $A \in \mathfrak{B}(\mathbb{X})$ and $x \in \mathbb{H} \cap {\mathbb{H}'}^c$ which in turn holds if $\mathbb{P}_x(T < \infty) = 1$ for $T \stackrel{\text{def}}{=} \inf\{n \ge 0 : M_{\sigma_n} \in$

$\mathbb{H}'\}$ and all $x \in \mathbb{H} \cap \mathbb{H}'^c$ because then, by the strong Markov property,

$$\mathbb{P}_x(M_{\sigma_n} \in A \text{ i.o.}) \;=\; \int_{\mathbb{H}'} \mathbb{P}_y(M_{\sigma_n} \in A \text{ i.o.})\, \mathbb{P}_x(M_{\sigma_T} \in dy) \;=\; 1.$$

We will now prove the latter assertion for an arbitrary fixed $x \in \mathbb{H} \cap \mathbb{H}'^c$. To that end, we first show the existence of a P-absorbing set $\mathbb{X}_\infty \subset \mathbb{H}'$. Define $\mathbb{X}_1 \stackrel{\text{def}}{=} \{y \in \mathbb{H}' : P(y, \mathbb{H}') = 1\}$ and then recursively $\mathbb{X}_{n+1} \stackrel{\text{def}}{=} \{y \in \mathbb{X}_n : P(y, \mathbb{X}_n) = 1\}$ for $n \geq 1$. Using $\pi = \pi P$ and $\pi(\mathbb{H}') = 1$, we inductively obtain $\pi(\mathbb{X}_n) = 1$ for all $n \geq 1$. The $\mathbb{X}_n$ are obviously nonincreasing, and $\mathbb{X}_\infty \stackrel{\text{def}}{=} \cap_{n\geq 1}\mathbb{X}_n$ satisfies $P(y, \mathbb{X}_\infty) = \lim_{n\to\infty} P(y, \mathbb{X}_n) = 1$ for all $y \in \mathbb{X}_\infty$. Now observe that $\pi(\mathbb{X}_\infty) = 1$ implies $\mathbb{P}_x(\varrho(\mathbb{X}_\infty) = 1) = 1$ which in turn yields

$$\mathbb{P}_x(M_n \in \mathbb{X}_\infty \text{ for all } n \geq 1) \;=\; 1$$

because $\mathbb{X}_\infty$ is P-absorbing. In particular, we infer the a.s. finiteness of T under $\mathbb{P}_x$. $\diamondsuit$

Based upon the previous preliminary observations, the further derivation of Theorem 3.2 is divided into two parts: We first prove a stronger result when $(M_n)_{n\geq 0}$ is strictly contractive (Proposition 5.2). In the general situation, this stronger result still holds for the subsequence $(M_{\sigma_n})_{n\geq 0}$, and the second part will show how this combines with certain regenerative arguments similar to those in [1] in order to get Theorem 3.2.

1. The strictly contractive case. Suppose now $(M_n)_{n\geq 0}$ satisfies (1.3), (1.4) and is further strictly γ-contractive, i.e. $L_1 \leq \gamma < 1$ a.s. For our later convenience we assume $\gamma < 1/2$. Our goal is to show *V-geometric ergodicity* of $(M_n)_{n\geq 0}$ on an absorbing set $\mathbb{H}$ for a suitable π-integrable function $V : \mathbb{H} \to [1, \infty)$, that is

$$\sum_{n\geq 0} r^{-n} \|P^n(x, \cdot) - \pi\|_V \;<\; \infty \tag{5.4}$$

for all $x \in \mathbb{H}$ and some $r \in (0, 1)$ not depending on x. Provided the moment assumptions (3.3) (Case I) or (3.6) (Case II) for some $p > 0$, we define a new "flattened", but still complete metric $\hat{d}$ through

$$\hat{d}(x, y) \;\stackrel{\text{def}}{=}\; \begin{cases} \log^{p\wedge 1}\left(1 + \frac{d(x,y)}{1+d(x,y)}\right) & \text{in Case I} \\ d(x, y)^{\eta\wedge 1} & \text{in Case II,} \end{cases} \tag{5.5}$$

where $\eta \in (0, p]$ is as in (3.7). Let $(\hat{L}_n)_{n\geq 1}$ be the sequence of Lipschitz constants for $(F_n)_{n\geq 1}$ under $\hat{d}$. As one can easily check, $L_1 \leq \gamma < 1$ a.s. implies $\hat{L}_1 \leq \hat{\gamma} < 1$ a.s.; we may in fact choose $\hat{\gamma} = \max_{t\in[0,1]} \frac{\log^{p\wedge 1}(1+\beta x)}{\log^{p\wedge 1}(1+x)}$ for any $\beta \in (\gamma, 1)$ in Case I and $\hat{\gamma} = \gamma^{p\wedge 1}$ in Case II. Notice that $\hat{d}$ is bounded in Case I whence $\mathbb{E}\hat{d}(F_1(x_0), x_0)^q < \infty$ for all $q > 0$. In Case II, we have $\mathbb{E}\hat{d}(F_1(x_0), x_0)^{p\vee 1} < \infty$ by (3.6). Consequently, replacing d with $\hat{d}$ leaves us with an IFS $(M_n)_{n\geq 0}$ which still satisfies (1.3), (1.4), and furthermore condition (3.6) for some $p \geq 1$ in *both* cases. The function V is now defined as

$$V(x) \stackrel{\text{def}}{=} \begin{cases} 1+\hat{d}(x,x_0) & \text{in Case I} \\ 1+\hat{d}(x,x_0)^{\eta\vee 1} = 1+d(x,x_0)^\eta & \text{in Case II} \end{cases} \tag{5.8}$$

and indeed π-integrable by (3.4), respectively (3.7). Notice that V coincides with f in (3.10) in Case II.

PROPOSITION 5.2. *Let $(M_n)_{n\ge 0}$ be a strictly γ-contractive IFS satisfying (1.3), (1.4), $\gamma < 1/2$ and (3.3) or (3.6) for some $p > 0$. If $(M_n)_{n\ge 0}$ is further Harris recurrent on whole $\mathbb{X}$ and $\mathrm{int}(supp\,\pi) \neq \emptyset$ holds, then $(M_n)_{n\ge 0}$ is V-geometrically ergodic with V as in (5.8). More precisely,*

$$\sum_{n\ge 0} r^{-n}\|P^n(x,\cdot)-\pi\|_V \;\le\; AV(x) \tag{5.9}$$

for all $x \in \mathbb{X}$ and suitable constants $r \in (0,1)$ and $A > 0$.

In Case I, the V-norm just coincides with the total variation norm $\|\cdot\|$ because V is bounded. As already discussed in some detail, the proof of the proposition will require that the balls $\mathbb{B}_R = \mathbb{B}_R(x_0) \stackrel{\text{def}}{=} \{x \in \mathbb{X} : d(x,x_0) \le R\}$ are petite for all $R > 0$ with $\pi(\mathbb{B}_R) > 0$.

PROOF. The result follows directly from [9, Theorem 15.0.1] if we can verify the drift condition

$$\Delta V(x) \;\le\; -\beta V(x) \;+\; b\mathbf{1}_C(x), \quad x \in \mathbb{X}, \tag{5.10}$$

for a petite set C and constants $\beta, b \in (0,\infty)$, where

$$\Delta V(x) \stackrel{\text{def}}{=} PV(x) - V(x) \;=\; \int_{\mathbb{X}} (V(y)-V(x))\,P(x,dy).$$

We only consider Case II, the first one being similar and even a little easier. It follows upon using Minkowski's inequality, the γ-strict contractivity, $b \stackrel{\text{def}}{=} 2^\eta \mathbb{E}d(F_1(x_0),x_0)^\eta < \infty$ (by (3.6)) and $\beta \stackrel{\text{def}}{=} (1-(2\gamma)^\eta)/2 > 0$ (since $\gamma < 1/2$)

$$\begin{aligned} \Delta V(x) &= \mathbb{E}_x d(F_1(x),x_0)^\eta - d(x,x_0)^\eta \\ &\le \mathbb{E}_x[d(F_1(x),F_1(x_0)) + d(F_1(x_0),x_0)]^\eta - d(x,x_0)^\eta \\ &\le ((2\gamma)^\eta - 1)V(x) \;+\; 2^\eta\mathbb{E}d(F_1(x_0),x_0)^\eta \\ &\le -\beta V(x) \;+\; b\mathbf{1}_{\mathbb{B}_R}(x) \end{aligned} \tag{5.11}$$

for sufficiently large $R > 0$. So it only remains to verify the petiteness of $\mathbb{B}_R$ which is accomplished by the subsequent lemma. ♢

LEMMA 5.3. *Under the assumptions of Proposition 5.2, each $\mathbb{B}_R(x)$, $x \in \mathbb{X}$ and $R > 0$, is petite.*

PROOF. Since $(M_n)_{n\ge 0}$ is a weak Feller chain, its π-irreducibility and $\mathrm{int}(\mathit{supp}\,\pi)\ne\emptyset$ imply that all compact subsets of $\mathbb{X}$ are petite, see [8, Theorem 3.4]. $\mathrm{int}(\mathit{supp}\,\pi)\ne\emptyset$ further implies the existence of some $w\in\mathbb{X}$ such that $\pi(\mathbb{B}_R(w))>0$ for all $R>0$. Next, the local compactness of $\mathbb{X}$ guarantees the existence of a compact and thus petite neighborhood C of w. Consequently, every $\mathbb{B}_R(w)$ contained in C is also petite. Fix any $R_0>0$ having $\mathbb{B}_{2R_0}(w)$ petite and further an arbitrary $x\in\mathbb{X}$ and $R>0$. We must show petiteness of $\mathbb{B}_R(x)$. Since $\pi(\mathbb{B}_{R_0}(w))>0$, the Harris recurrence of $(M_n)_{n\ge 0}$ ensures the existence of some $N\ge 1$ such that $\gamma^N R<R_0$ and $P^N(x,\mathbb{B}_{R_0}(w))>0$. Now pick an arbitrary $y\in\mathbb{B}_R(x)$, $y\ne x$, and consider the two coupled chains $M_n^x\stackrel{\mathrm{def}}{=}F_{n:1}(x)$ and $M_n^y\stackrel{\mathrm{def}}{=}F_{n:1}(y)$ for $n\ge 0$. Since $d(M_n^x,M_n^y)\le\gamma^n d(x,y)\le\gamma^n R$ a.s. for all $n\ge 0$, the choice of N gives

$$d(M_N^y,w)\ \le\ d(M_N^x,M_N^y)+d(M_N^x,w)\ <\ R_0+d(M_N^x,w)\quad\text{a.s.}$$

and therefore $\{M_N^x\in\mathbb{B}_{R_0}(w)\}\subset\{M_N^y\in\mathbb{B}_{2R_0}(w)\}$ a.s. for all $y\in\mathbb{B}_R(x)$. We thus conclude

$$P^N(y,\mathbb{B}_{2R_0}(w))\ =\ \mathbb{P}(M_N^y\in\mathbb{B}_{2R_0}(w))\ \ge\ \mathbb{P}(M_N^x\in\mathbb{B}_{R_0}(w))\ =\ P^N(x,\mathbb{B}_{R_0}(w))\ >\ 0$$

for all $y\in\mathbb{B}_R(x)$, whence the petiteness of $\mathbb{B}_{2R_0}(w)$ implies the same for $\mathbb{B}_R(x)$. ♢

The next lemma will be needed for the proof of Theorem 3.2(b) and supplements the previous one by showing that the π-positive $\mathbb{B}_R(x)$ are also (1-)regular.

LEMMA 5.4. *Let $(M_n)_{n\ge 0}$ be a strictly contractive IFS of i.i.d. Lipschitz maps satisfying (1.3) and (1.4). Let $w\in\mathbb{X}$ and $R_0>0$ so large that $\pi(\mathbb{B}_{R_0}(w))>0$. Then*

$$\mathbb{E}_x\varrho(\mathbb{B}_R(w))\ \le\ K_R(1+\log(1+d(x,w)))\mathbb{E}_{\pi(\cdot|R_0)}\varrho(\mathbb{B}_{R_0}(w))\ <\ \infty \tag{5.12}$$

for all $x\in\mathbb{X}$, $R>R_0$ and some $K_R\in(0,\infty)$, where $\pi(\cdot|R)\stackrel{\mathrm{def}}{=}\frac{\pi(\cdot\cap\mathbb{B}_R)}{\pi(\mathbb{B}_R)}$. In particular,

$$\sup_{x\in\mathbb{B}_R(w)}\ \mathbb{E}_x\varrho(\mathbb{B}_R(w))\ <\ \infty. \tag{5.13}$$

Given the conditions of Proposition 5.2, $\mathbb{B}_R(w)$ is regular.

PROOF. Note first that $(M_{\varrho_n(\mathbb{B}_R(w))},\varrho_{n+1}(\mathbb{B}_R(w))-\varrho_n(\mathbb{B}_R(w)))_{n\ge 0}$ is stationary and ergodic under $\mathbb{P}_{\pi(\cdot|R)}$ and that $\mathbb{E}_{\pi(\cdot|R)}\varrho_1(\mathbb{B}_R(w))<\infty$ for all R. Moreover, by the ergodic theorem, $n^{-1}\varrho_n(\mathbb{B}_R(w))\to\mathbb{E}_{\pi(\cdot|R)}\varrho(\mathbb{B}_R(w))$ a.s. and in mean under $\mathbb{P}_{\pi(\cdot|R)}$. Hence there exists $z\in\mathbb{B}_{R_0}(w)$ such that

$$\sup_{n\ge 1} n^{-1}\mathbb{E}_z\varrho_n(\mathbb{B}_{R_0}(w))\ \le\ K\,\mathbb{E}_{\pi(\cdot|R_0)}\varrho(\mathbb{B}_{R_0}(w)) \tag{5.14}$$

for a suitable constant $K>0$. Suppose w.l.o.g. $z=w$.

Next, pick an arbitrary $x \in \mathbb{X}$, $x \neq w$, and consider once more the two coupled chains M_n^x, M_n^w for $n \geq 0$. Denote by $\varrho_n^x(\mathbb{B}_R(w))$ and $\varrho_n^w(\mathbb{B}_R(w))$, $n \geq 1$, the associated return times to $\mathbb{B}_R(w)$. Since $d(M_n^x, M_n^w) \leq \gamma^n d(x,w)$ a.s. for all $n \geq 0$, we infer

$$d(M^x_{\varrho_n^w(\mathbb{B}_{R_0}(w))}, M^w_{\varrho_n^w(\mathbb{B}_{R_0}(w))}) \;\leq\; \gamma^n d(x,w) \quad \text{a.s.}$$

and therefore

$$d(M^x_{\varrho_n^w(\mathbb{B}_{R_0}(w))}, w) \;\leq\; R_0 + \gamma^n d(x,w) \quad \text{a.s.}$$

which in turn implies $\varrho^x(\mathbb{B}_R(w)) \leq \varrho^w_{N(x,R-R_0)}(\mathbb{B}_{R_0}(w))$ a.s. for all $R > R_0$ and

$$N(x, R-R_0) \;\stackrel{\text{def}}{=}\; \left[\frac{\log d(x,w) - \log(R-R_0)}{\log(1/\gamma)}\right]^+ + 1,$$

where $[t]^+ \stackrel{\text{def}}{=} \sup\{k \in \mathbb{Z} : k \leq t\} \vee 0$. (5.12) follows easily from the inequality

$$\begin{aligned} \mathbb{E}_x \varrho(\mathbb{B}_R(w)) \;&=\; \mathbb{E}\varrho^x(\mathbb{B}_R(w)) \;\leq\; \mathbb{E}\varrho_{N(x,R-R_0)}(\mathbb{B}_{R_0}(w)) \\ &\leq\; KN(x,R-R_0)\, \mathbb{E}_{\pi(\cdot|R_0)}\varrho(\mathbb{B}_{R_0}(w)), \end{aligned}$$

where (5.14) has been utilized. Given the conditions of Proposition 5.2, we have that $\mathbb{B}_R(w)$ is petite (Lemma 5.3) and *"self-regular"* (property (5.13)), whence its regularity follows from [9, Theorem 14.2.4]. ♢

2. The general case (Proof of Theorem 3.2). The assertions of Theorem 3.2 are now fairly easily established. We begin with part (a) of the theorem.

Proof of Theorem 3.2(a). For a fixed $\gamma \in (0,1)$, let $(M_{\sigma_n})_{n \geq 0}$ be the strictly contractive IFS with transition kernel P' defined before Lemma 5.1 with σ_n, $n \geq 0$, being the level $\log\gamma$ ladder epochs in (5.3). Proposition 5.2 yields $K \stackrel{\text{def}}{=} \sum_{n \geq 0} r^{-n} \|P'^n(x,\cdot) - \pi\| < \infty$ for all $x \in \mathbb{X}$ and some $r \in (0,1)$. Put $T(n) \stackrel{\text{def}}{=} \sup\{k \geq 0 : \sigma_k \leq n\}$ for $n \geq 0$ and $\Delta_{x,\pi} \stackrel{\text{def}}{=} \mathbb{P}_x - \mathbb{P}_\pi$ for $x \in \mathbb{X}$. Notice that the $T(n)$ are independent of M_0. We have for all $B \in \mathfrak{B}(\mathbb{X})$, $x \in \mathbb{X}$, $n \geq 0$ and $a > 0$

$$\begin{aligned} |P^n(x,B) - \pi(B)| \;&=\; |\mathbb{P}_x(M_n \in B) - \mathbb{P}_\pi(M_n \in B)| \\ &\leq\; \mathbb{P}(T(n) \leq an) \;+\; |\Delta_{x,\pi}(M_n \in B, T(n) > an)|. \end{aligned} \tag{5.15}$$

Note that, given arbitrary probability measures Q_1, Q_2 on a measurable space $(\Omega, \mathfrak{A})$ and a measurable partition $(B_k)_{1 \leq k \leq n}$ of Ω, the equality $\|Q_1 - Q_2\| = \sum_{k=1}^n \|Q_1(\cdot \cap B_k) - Q_2(\cdot \cap B_k)\|$ holds. Therefore, we further infer for all $B \in \mathfrak{B}(\mathbb{X})$, $x \in \mathbb{X}$, $n \geq 0$ and $a > 0$ that

$$\begin{aligned} &|\Delta_{x,\pi}(M_n \in B, T(n) > an)| \\ &\quad\leq\; \sum_{k > an} \sum_{j=1}^n |\Delta_{x,\pi}(M_n \in B, \sigma_k = j, \sigma_{k+1} > n)| \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k>an} \sum_{j=1}^{n} \left| \int \mathbb{P}_y(M_{n-j} \in B, \sigma_1 > n-j)\ \Delta_{x,\pi}(M_{\sigma_k} \in dy, \sigma_k = j) \right| \\
&\le \sum_{k>an} \sum_{j=1}^{n} \|\Delta_{x,\pi}(M_{\sigma_k} \in \cdot, \sigma_k = j)\| \\
&= n \sum_{k>an} \|\mathbb{P}_x(M_{\sigma_k} \in \cdot) - \mathbb{P}_\pi(M_{\sigma_k} \in \cdot)\| \\
&= n \sum_{k>an} \|P'^k(x,\cdot) - \pi\| \\
&\le n r^{an} K
\end{aligned}
$$

and thus

$$\|P^n(x,\cdot) - \pi\| \ \le\ \mathbb{P}(T(n) \le an) \ +\ n r^{an} k$$

for all $x \in \mathbb{X}$, $n \ge 0$ and $a > 0$. Assertions (3.11) and (3.12) are now immediate because, by (3.3), we can choose $a > 0$ such that

$$\sum_{n\ge 1} n^{p-1}\mathbb{P}(T(n) \le an) \ <\ \infty \quad \text{and} \quad \lim_{n\to\infty} n^p \mathbb{P}(T(n) \le an) \ =\ 0.$$

This has been shown in [1, Lemma 3.5]. ♢

REMARK. If (3.6) holds, Lemma 3.5 in [1] further provides

$$\lim_{n\to\infty} b^n \mathbb{P}(T(n) \le an) \ =\ 0$$

for some $a > 0$ and $b > 1$. Hence the previous proof under moment assumption (3.6) immediately yields

$$\sum_{n\ge 0} s^n \|P^n(x,\cdot) - \pi\| \ <\ \infty.$$

However, this is weaker than the result asserted in (3.13) of Theorem 3.2(b) which we are going to prove next.

PROOF OF THEOREM 3.2(b). Recall that $f(x) = V(x) = 1 + d(x,x_0)^p$ in the present situation. The counterpart of (5.15) for an estimation of $\|P^n(x,\cdot) - \pi\|_f$ obviously reads

$$|P^n g(x) - \pi(g)| \ \le\ \int_{\{T(n)\le an\}} f(M_n)\ d(\mathbb{P}_x + \mathbb{P}_\pi) \ +\ \left| \int_{\{T(n)>an\}} g(M_n)\ d\Delta_{x,\pi} \right|$$

for arbitrary functions g satisfying $|g| \le f$. Unfortunately, the first term on the right-hand side seems to be difficult for further estimation. Therefore, instead of pursuing this argument any further, another use of the drift condition (5.1), now verified for the m-skeleton $(M_{mn})_{n\ge 0}$ for suitable $m \ge 1$, appears to be more convenient. Before we can do so, we must however show that the $\mathbb{B}_R$, $R > 0$ are also petite for these skeletons. We begin by noting that all petite sets C for $(M_n)_{n\ge 0}$ are also small (i.e. $P^n(x,\cdot) \ge \alpha\nu_n$ for all $x \in C$ and some $n \ge 1$, $\alpha \in (0,1)$ and a probability distribution ν_n concentrated on C), because $(M_n)_{n\geq 0}$ is aperiodic, see [9, Theorem 5.5.7]. Aperiodicity further guarantees that small sets for $(M_n)_{n\geq 0}$ are also small (and thus petite) for each skeleton. Consequently, we are left with proof that the π-positive $\mathbb{B}_R$ are petite or, equivalently [9, Theorem 14.2.4], regular for $(M_n)_{n\geq 0}$. But the latter is easy with Lemma 5.4. Let $\varrho'(B)$ be the first return time to $B \in \mathfrak{B}(\mathbb{X})$ of $(M_{\sigma_n})_{n\geq 0}$. Clearly, $\varrho(B) \leq \varrho'(B)$ for each B so that the regularity of $\mathbb{B}_R$ for $(M_{\sigma_n})_{n\geq 0}$ (Lemma 5.4) immediately implies the same property for $(M_n)_{n\geq 0}$.

Now we can verify the drift condition (5.10) for some m-skeleton and with V as given. Since $\mathbb{E}\log^+ L_1 < 0$ and $\mathbb{E}L_1^p < \infty$, there exists $m \geq 1$ such that $\gamma_m \stackrel{\text{def}}{=} \mathbb{E}L_{1:m}^p < 1/2$. Notice also that $d(F_{1:j}(x_0), F_{1:j-1}(x_0)) \leq L_{1:j-1} d(F_j(x_0), x_0)$ together with the independence of $L_{1:j-1}$ and $F_j(x_0)$ for each $j \geq 1$ implies (setting $\gamma_0 \stackrel{\text{def}}{=} 1$)

$$\begin{aligned} b \stackrel{\text{def}}{=} (\mathbb{E}d(F_{1:m}(x_0), x_0)^p)^{1/p} &\leq \sum_{j=1}^m (\mathbb{E}d(F_{1:j}(x_0), F_{1:j-1}(x_0))^p)^{1/p} \\ &\leq \sum_{j=1}^m (\mathbb{E}[L_{1:j-1} d(F_j(x_0), x_0)]^p)^{1/p} \\ &= (\mathbb{E}d(F_1(x_0), x_0)^p)^{1/p} \sum_{j=0}^{m-1} \gamma_j^{1/p} < \infty. \end{aligned} \tag{5.16}$$

Hence a similar estimation as in (5.11) yields with $\beta \stackrel{\text{def}}{=} (1 - (2\gamma_m)^p)/2 > 0$

$$\begin{aligned} P^m f(x) - f(x) &\leq ((2\gamma_m)^p - 1) f(x) + 2^p \mathbb{E}d(F_{1:m}(x_0), x_0)^p \\ &\leq -\beta f(x) + (2b)^p \mathbf{1}_{\mathbb{B}_R}(x) \end{aligned}$$

for sufficiently large $R > 0$, so that, by a further appeal to [9, Theorem 15.0.1],

$$\sum_{n\geq 0} r^{-n} \|P^{mn}(x, \cdot) - \pi\|_f \leq A f(x)$$

for all $x \in \mathbb{X}$ and suitable constants $r \in (0, 1)$ and $A > 0$. The fact that condition (3.6), if valid for one $x_0 \in \mathbb{X}$, already holds for all $x \in \mathbb{X}$ in combination with (5.16) implies $P^j f(x) = \mathbb{E}d(F_{1:j}(x_0), x_0)^p < \infty$ for all $j \geq 1$, whence we finally conclude

$$\begin{aligned} \sum_{n\geq 0} r^{-n} \|P^n(x, \cdot) - \pi\|_f &\leq \sum_{j=0}^{m-1} r^{-j} \int \left(\sum_{n\geq 0} r^{-mn} \|P^{mn}(y, \cdot) - \pi\|_f \right) P^j(x, dy) \\ &\leq \sum_{j=0}^{m-1} r^{-j} \int A f(y)\, P^j(x, dy) \\ &\leq A \sum_{j=0}^{m-1} P^j f(x) < \infty \end{aligned}$$

for all $x \in \mathbb{X}$. The proof of (3.13) is herewith complete. $\diamondsuit$

Acknowledgement

A major part of this work was done during a visit of the author at the Academia Sinica (Taipei) in Spring 2000. Its kind hospitality and stimulating environment is gratefully acknowledged. I would also like to thank two anonymous referees for their numerous insightful and constructive comments which led to a substantial improvement of the presentation.

References

[1] ALSMEYER, G. and FUH, C.D. (2001). Limit theorems for iterated random functions by regenerative methods. *Stoch. Proc. Appl.* **96**, 123-142. Corrigendum in **97**, 341-345.

[2] BOUGEROL, P. and PICARD, N. (1992). Strict stationarity of generalized autoregressive processes. *Ann. Probab.* **20**, 1714-1730.

[3] BRANDT, A. (1986). The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Adv. Appl. Probab.* **18**, 211-220.

[4] CHAN, K.S. and TONG, H. (1985). On the use of the deterministic Lyapunov function for the ergodicity of stochastic difference equations. *Adv. Appl. Probab.* **17**, 666-678.

[5] DIACONIS, P. and FREEDMAN, D. (1999). Iterated random functions. *SIAM Review* **41**, 45-76.

[6] ELTON, J.H. (1990). A multiplicative ergodic theorem for Lipschitz maps. *Stoch. Proc. Appl.* **34**, 39-47.

[7] HALMOS, P. and SAVAGE, L.J. (1948). Application of the Radon-Nikodym theorem to the theory of sufficient statistics. *Ann. Math. Statist.* **20**, 225-241.

[8] MEYN, S.P. and TWEEDIE, R.L. (1992). Stability of Markovian processes, I. *Adv. Appl. Probab.* **24**, 542-574.

[9] MEYN, S.P. and TWEEDIE, R.L. (1993). *Markov Chains and Stochastic Stability.* Springer, London.

[10] NIEMI, S. and NUMMELIN, E. (1986). On non-singular renewal kernels with an application to a semigroup of transition kernels. *Stoch. Proc. Appl.* **22**, 177-202.

[11] PETRUCELLI, J.D. and WOOLFORD, S.W. (1984). A threshold AR(1) model. *J. Appl. Probab.* **21**, 270-286.

[12] TONG, H. (1990). *Non-linear Time Series. A Dynamical Systems Approach.* Oxford Univ. Press, Oxford.

[13] VERVAAT, W. (1979). On a stochastic difference equation and a representation of non-negative infinitely divisible random variables. *Adv. Appl. Probab.* **11**, 750-783.