PROBABILITY
AND
MATHEMATICAL STATISTICS
Vol. 29, Fasc. 2 (2009), pp. 321–336

ASYMPTOTIC BEHAVIOR OF ULTIMATELY CONTRACTIVE ITERATED RANDOM LIPSCHITZ FUNCTIONS

BY

GEROLD ALSMEYER AND GERD HÖLKER (MÜNSTER)

Abstract. Let $(F_n)_{n \geqslant 0}$ be a random sequence of i.i.d. global Lipschitz functions on a complete separable metric space $(\mathbb{X}, d)$ with Lipschitz constants $L_1, L_2, \ldots$ For $n \geqslant 0$, denote by $M_n^x = F_n \circ \ldots \circ F_1(x)$ and $\hat{M}_n^x = F_1 \circ \ldots \circ F_n(x)$ the associated sequences of forward and backward iterations, respectively. If $\mathbb{E} \log^+ L_1 < 0$ (mean contraction) and $\mathbb{E} \log^+ d\big(F_1(x_0), x_0\big)$ is finite for some $x_0 \in \mathbb{X}$, then it is known (see [9]) that, for each $x \in \mathbb{X}$, the Markov chain M_n^x converges weakly to its unique stationary distribution π, while $\hat{M}_n^x$ is a.s. convergent to a random variable $\hat{M}_\infty$ which does not depend on x and has distribution π. In [2], renewal theoretic methods have been successfully employed to provide convergence rate results for $\hat{M}_n^x$, which then also lead to corresponding assertions for M_n^x via $M_n^x \overset{d}{=} \hat{M}_n^x$ for all n and x, where $\overset{d}{=}$ means equality in law. Here our purpose is to demonstrate how these methods are extended to the more general situation where only ultimate contraction, i.e. an a.s. negative Lyapunov exponent $\lim_{n\to\infty} n^{-1} \log l(F_n \circ \ldots \circ F_1)$ is assumed (here $l(F)$ denotes the Lipschitz constant of F). This not only leads to an extension of the results from [2] but in fact also to improvements of the obtained convergence rate.

2000 AMS Mathematics Subject Classification: Primary: 60J05; Secondary: 60K05, 60G17.

Key words and phrases: Random Lipschitz function, ultimately contractive, forward iterations, backward iterations, stationary distribution, Prokhorov metric, level γ ladder epochs, Lyapunov exponent.

1. INTRODUCTION

Iterations of random functions appear in a wide range of applied stochastic models ranging from queuing theory and financial times series to perfect simulation, the generation of fractal images and data compression; see Diaconis and Freedman [8] for an excellent survey including an extensive list of relevant literature. Not surprisingly, the question of stability of such iterated function systems (IFS) under suitable global or local contraction conditions is often of central interest and has therefore been studied extensively in the literature, e.g. in [3], [5],

[6], [8], [9], [11]–[14]. Here we will focus on a particularly nice subclass of IFS, namely iterations of i.i.d. (global) Lipschitz maps.

A very effective method of studying the asymptotic behavior of an ultimately contractive IFS of i.i.d. Lipschitz functions (to be defined below) is based on the identification of a strictly contractive, and thus well-behaved subsystem along a renewal sequence of stopping times followed by a subsequent analysis of the excursions of the system between these stopping times. The simple idea behind this approach is called *regeneration* in applied probability and calls for renewal theory as a natural ingredient. In [2] and also in [1], this idea has been successfully pursued and led to results on the rate of convergence of mean contractive systems to its stationary limit. Regeneration is also used in two related articles by Babillot et al. [4] and Silvestrov and Stenflo [13], but in a different vein. Roughly speaking, the purpose of the present work is to show how the regenerative arguments given in [2] may be refined in order for getting stronger and in fact more natural versions of the results from there.

We continue with a short review of the model assumptions and the notation from [2]. Let $(\mathbb{X}, d)$ be a complete separable metric space with Borel-σ-field $\mathfrak{B}(\mathbb{X})$ and $(M_n)_{n\geqslant 0}$ a temporally homogeneous $\mathbb{X}$-valued Markov chain of the form $M_n = F(\theta_n, M_{n-1})$ for $n \geqslant 1$, where

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

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

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

Its transition kernel P is given by $P(x, B) = \Lambda\big(F(\cdot, x) \in B\big)$ for $x \in \mathbb{X}$ and $B \in \mathfrak{B}(\mathbb{X})$, and we let P^n denote the n-step transition kernel. 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 by $\mathbb{E}_x$, 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.

For a Lipschitz continuous mapping $f : \mathbb{X} \to \mathbb{X}$, we put

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

and note that

$$L_n \stackrel{\text{def}}{=} l\big(F(\theta_n, \cdot)\big), \quad n \geqslant 1,$$

forms a sequence of i.i.d. random variables. Write $F_n(x)$ for $F(\theta_n, x)$ and put

$$F_{k:n} \stackrel{\text{def}}{=} F_k \circ \ldots \circ F_n, \qquad F_{n:k} \stackrel{\text{def}}{=} F_n \circ \ldots \circ F_k \qquad \text{for } 1 \leqslant k \leqslant n,$$

and

$$F_{0:1}(x) = F_{1:0}(x) \overset{\text{def}}{=} x.$$

Then M_n equals the n-th *forward iteration*, as $M_n = F_n(M_{n-1}) = F_{n:1}(M_0)$, and is closely related to the corresponding n-th *backward iteration*

$$\hat{M}_n \overset{\text{def}}{=} F_{1:n}(M_0),$$

owing to the obvious fact that

$$\mathbb{P}_x(M_n \in \cdot) = \mathbb{P}_x(\hat{M}_n \in \cdot)$$

for each $x \in \mathbb{X}$. Putting $M_n^x \overset{\text{def}}{=} F_{n:1}(x)$ and $\hat{M}_n^x \overset{\text{def}}{=} F_{1:n}(x)$ for $x \in \mathbb{X}$, we further have

$$\mathbb{P}\big((M_n^x, \hat{M}_n^x)_{n\geqslant 0} \in \cdot\big) = \mathbb{P}_x\big((M_n, \hat{M}_n)_{n\geqslant 0} \in \cdot\big).$$

The reason for introducing these additional sequences is that we will frequently do comparisons of $\hat{M}_n^x$ and $\hat{M}_n^y$, or M_n^x and M_n^y, for different x, y.

For the more general situation of a stationary sequence $(F_n)_{n\geqslant 1}$, Elton [9] showed that the law of M_n converges weakly to a unique stationary distribution π whenever

$$\mathbb{E}\log^+ L_1 < \infty \quad \text{and} \quad \mathbb{E}\log^+ d\big(F_1(x_0), x_0\big) < \infty \tag{1.1}$$

for some (and then all) $x_0 \in \mathbb{X}$ and the *Lyapunov exponent* $\log l^*$ is a.s. negative, where

$$\log l^* \overset{\text{def}}{=} \lim_{n\to\infty} n^{-1} \log l(F_{n:1}) \text{ a.s.} \tag{1.2}$$

exists by Kingman's subadditive ergodic theorem. By Kolmogorov's zero-one law, l^* is further a.s. constant in the present situation of i.i.d. $F_1, F_2, \ldots$ We call $(M_n)_{n\geqslant 0}$ *ultimately contractive* if $\log l^* < 0$ a.s., and *mean contractive* if the stronger condition $\mathbb{E}\log l(F_1) < 0$ holds true. Then the basic question posed at the outset of this work as opposed to [2] can be stated as follows: To what extent do the results obtained in [2] for mean contractive IFS of i.i.d. Lipschitz functions generalize to the more natural class of ultimately contractive IFS? We will show that the results of [2] not only persist to hold under ultimate contraction but may even be improved in that certain bounds on the rate of contraction appear to be sharper. The latter is accomplished by a refined renewal theoretic analysis an outline of which is next.

The basic idea for studying the asymptotic properties of an ultimately contractive, and thus weakly convergent IFS $(M_n)_{n\geqslant 0}$ is to go via the backward iterations $\hat{M}_n^x = F_{1:n}(x)$ for which convergence holds even true almost surely with a limit $\hat{M}_\infty$ not depending on x and (as it must) having distribution π. The obvious inequality

$$d(\hat{M}_{n+m}^x, \hat{M}_n^x) \leqslant l(F_{1:n})\, d\big(F_{n+1:n+m}(x), x\big) \text{ a.s.,} \tag{1.3}$$

valid for all $n, m \geqslant 0$ and $x \in \mathbb{X}$, forms a key tool in the necessary analysis which embarks on this very inequality together with the following observation: Assuming ultimate contraction ($\log l^* < 0$), fixing any $\gamma \in (l^*, 1)$ and using (1.2), we see that the ladder epochs $\sigma_0 \stackrel{\text{def}}{=} 0$,

$$\text{(1.4)} \quad \sigma_n \stackrel{\text{def}}{=} \inf \left\{ j > \sigma_{n-1} : \frac{1}{j - \sigma_{n-1}} \log l(F_{\sigma_{n-1}+1:j}) \leqslant \log \gamma \right\}, \quad n \geqslant 1,$$

are all a.s. finite and constituting an ordinary discrete renewal process. As a consequence, the subsequence $(M_{\sigma_n})_{n \geqslant 0}$ again forms an IFS of i.i.d. Lipschitz maps. Its associated backward iterations $\hat{M}^x_{\sigma_n} = F_{1:\sigma_n}(x)$ are *strictly contractive* because, by construction,

$$l(F_{1:\sigma_1}) \leqslant \gamma^{\sigma_1} \leqslant \gamma < 1.$$

Inequality (1.3) hence takes the very strong form

$$\text{(1.5)} \qquad d(\hat{M}^x_{\sigma_{n+m}}, \hat{M}^x_{\sigma_n}) \leqslant \gamma^n d\big(F_{\sigma_{n+1}+1:\sigma_{n+m}}(x), x\big)$$

for all $n, m \geqslant 0$ and $x \in \mathbb{X}$ and suggests the following procedure to prove convergence results for $(M_n)_{n \geqslant 0}$ and its associated backward iterations:

STEP 1. Given a set of conditions, find out what kind of results hold true for the strictly contractive sequences $(M_{\sigma_n})_{n \geqslant 0}$ or $(\hat{M}_{\sigma_n})_{n \geqslant 0}$ for any $\gamma \in (0, 1)$.

STEP 2. Analyze the excursions of $(M_n)_{n \geqslant 0}$ or $(\hat{M}_n)_{n \geqslant 0}$ between two successive ladder epochs and adjust the results with respect to $(M_n)_{n \geqslant 0}$ or $(\hat{M}_n)_{n \geqslant 0}$, respectively, if necessary.

In [2] the log-Lipschitz constant $l(F_{1:n})$ is first estimated from above by

$$\Gamma_n \stackrel{\text{def}}{=} \sum_{k=1}^{n} \log L_k,$$

which forms the n-th partial sum of an ordinary random walk having negative drift under the stronger mean contraction condition $\mathbb{E} \log l(F_1) < 0$. Hence the level $\log \gamma$ ladder epochs $\sigma_0(\gamma) \stackrel{\text{def}}{=} 0$,

$$\text{(1.6)} \qquad \sigma_n(\gamma) \stackrel{\text{def}}{=} \inf \left\{ k > \sigma_{n-1}(\gamma) : \frac{\Gamma_k - \Gamma_{\sigma_{n-1}(\gamma)}}{k - \sigma_{n-1}(\gamma)} \leqslant \log \gamma \right\}, \quad n \geqslant 1,$$

are all a.s. finite and constituting an ordinary discrete renewal process. These $\sigma_n(\gamma)$ then take the role of the σ_n defined in (1.4). However, besides requiring the stronger mean contraction condition, this latter approach also gives away too much information on the log-Lipschitz constants $l(F_{1:n})$ by estimating them at the outset through the sum of the log-Lipschitz constants $l(F_k)$ of the single factors. As a consequence, the results in [2], though looking quite similar to the ones given here, are actually weaker in almost all parts. In particular, a certain lower bound γ^* defined there through a certain optimal choice of γ in (1.6) and popping up in various places will here be replaced with the quite natural and really optimal constant l^*; see Theorem 2.1, parts (a)–(c) of Theorem 2.2 and part (a) of Theorem 2.3.

As in [2], we will work with two sets of conditions, namely that, for some $p > 0$ and some $x_0 \in \mathbb{X}$, either

$$\mathbb{E}\log^{p+1}(1 + L_1) < \infty \quad \text{and} \quad \mathbb{E}\log^{p+1}\Big(1 + d\big(F_1(x_0), x_0\big)\Big) < \infty \tag{1.7}$$

or

$$\mathbb{E}L_1^p < \infty \quad \text{and} \quad \mathbb{E}d\big(F_1(x_0), x_0\big)^p < \infty \tag{1.8}$$

holds true. Two major conclusions will concern the distance of $P^n(x, \cdot)$ for $x \in \mathbb{X}$ and π in the Prokhorov metric associated with d. Following [8], the latter is also denoted by d and defined, for two probability measures λ_1, λ_2 on $\mathbb{X}$, as the infimum over all $\delta \geqslant 0$ 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 $B^\delta \stackrel{\text{def}}{=} \{x \in \mathbb{X} : d(x, y) < \delta \text{ for some } y \in B\}$.

The further organization of the paper is as follows. The main results are presented in the next section. Section 3 collects some necessary prerequisites for their proofs, which in turn will be provided in Section 4. Plainly, details of the necessary arguments will often be omitted or stated in abridged form whenever these can essentially be copied from [2]. On the other hand, the reader will hopefully also notice some improvements in the presentation of technical details.

2. RESULTS

Our first result provides additional information in Elton's ergodic theorem for the case of an ultimately contractive IFS of i.i.d. Lipschitz functions, which is considered here. It shows that, with high probability, the rate of decay of $d(\hat{M}_\infty, \hat{M}_n)$ towards 0 is of geometric order l, where l can be chosen arbitrarily close (from above) to l^*.

THEOREM 2.1. *Let $(M_n)_{n\geqslant 0}$ be an ultimately contractive IFS of i.i.d. Lipschitz maps satisfying* (1.1) *and with Lyapunov exponent* $\log l^*$. *Then*

$$\lim_{n\to\infty} \mathbb{P}_x\big(d(\hat{M}_\infty, \hat{M}_n) > l^n\big) = 0$$

holds true for all $x \in \mathbb{X}$ and $l \in (l^*, 1)$.

The next two theorems contain the announced refinements of similar results in [2], see Theorems 2.2 and 2.3 there.

THEOREM 2.2. *Assuming the situation of Theorem* 2.1 *and additionally condition* (1.7) *for some* $p > 0$, *the following assertions hold true*:

(a) *For each* $l \in (l^*, 1)$, $x \in \mathbb{X}$ *and some* $c_l \in (0, \infty)$,

$$\sum_{n \geqslant 1} n^{p-1} \mathbb{P}_x\big(d(\hat{M}_\infty, \hat{M}_n) > l^n\big) \leqslant c_l \Big(1 + \log^p\big(1 + d(x, x_0)\big)\Big)$$

and

$$\lim_{n \to \infty} n^p \, \mathbb{P}_x\big(d(\hat{M}_\infty, \hat{M}_n) > l^n\big) \;=\; 0.$$

(b) *For each* $l \in (l^*, 1)$ *and* $x \in \mathbb{X}$,

$$\limsup_{n \to \infty} n^{(p-1)/p} \left(\frac{1}{n} \log d(\hat{M}_\infty, \hat{M}_n) - \log l\right) \leqslant 0 \quad \mathbb{P}_x\text{-a.s.}$$

If $0 < p \leqslant 1$, *then this remains true for* $l = l^*$.

(c) *If* $p = 1$, *then* $\lim_{n \to \infty} l^{-n} d(\hat{M}_\infty, \hat{M}_n) = 0$ $\mathbb{P}_x$*-a.s. for all* $x \in \mathbb{X}$ *and all* $l \in (l^*, 1)$.

(d) $d\big(P^n(x, \cdot), \pi\big) \leqslant A_x (n+1)^{-p}$ *for all* $n \geqslant 0$, $x \in \mathbb{X}$ *and a constant* $A_x = \max\{A, 2d(x, x_0)\}$, *where* $A > 0$ *does not depend on* x *or* n.

(e) $\int_{\mathbb{X}} \log^p\big(1 + d(x, x_0)\big)\, \pi(dx) < \infty$.

THEOREM 2.3. *Assuming the situation of Theorem* 2.1 *and additionally condition* (1.8) *for some* $p > 0$, *the following assertions hold true*:

(a) *For each* $l \in (l^*, 1)$, $x \in \mathbb{X}$ *and some* $\alpha_l \in (0, 1)$,

$$\lim_{n \to \infty} \alpha_l^{-n} \mathbb{P}_x\big(d(\hat{M}_\infty, \hat{M}_n) > l^n\big) \;=\; 0.$$

(b) *There exists* $\eta > 0$ *such that for each* $q \in (0, \eta)$ *and some* $\alpha_q \in (0, 1)$,

$$\lim_{n \to \infty} \sup_{x \in \mathbb{X}} \alpha_q^{-n} \big(1 + d(x, x_0)\big)^{-q} \mathbb{E}_x d(\hat{M}_\infty, \hat{M}_n)^q \;=\; 0.$$

If $q = \eta$, *then the same holds true with* $\alpha_q = 1$.

(c) $d\big(P^n(x, \cdot), \pi\big) \leqslant A_x r^n$ *for all* $n \geqslant 0$, *some* $r \in (0, 1)$ *and a constant* $A_x = \max\{A, d(x, x_0)\}$, *where* r *and* A *do not depend on* x *or* n.

(d) $\int_{\mathbb{X}} d(x, x_0)^\eta \, \pi(dx) < \infty$ *for some* $\eta > 0$.

REMARK 2.1. (a) The constants c_l, α_l and α_q in the previous theorems generally depend further on $p > 0$ appearing in the respective moment assumption.

(b) Parts (d), (e) of Theorem 2.2 and parts (c), (d) of Theorem 2.3 have been taken from [2] without modification. The geometric convergence stated in Theorem 2.3 (c) was first given by Diaconis and Freedman [8].

(c) Since (M_n^x, M_n^y) and $(\hat{M}_n^x, \hat{M}_n^y)$ are identically distributed and

$$d(\hat{M}_n^x, \hat{M}_n^y) \ \leqslant \ d(\hat{M}_\infty^{x_0}, \hat{M}_n^x) \ + \ d(\hat{M}_\infty^{x_0}, \hat{M}_n^y)$$

for all $x, y \in \mathbb{X}$ and $n \geqslant 0$, the assertions on $d(\hat{M}_\infty, \hat{M}_n)$ in the previous two theorems are easily translated into similar results on $d(M_n^x, M_n^y)$ for the forward iterations started at different values x and y. This has been done in Corollaries 2.4 (a)–(c) and 2.5 (a) in [2], and we just note that these results could now be restated for each $l \in (l^*, 1)$.

(d) As further explained in [2] (cf. the end of Section 2), our results do also provide information on the distance of $M_n^x = F_{n:1}(x)$ for any $x \in \mathbb{X}$ to a stationary counterpart $M_n^\pi \stackrel{\text{def}}{=} F_{n:1}(M_0^\pi)$, where M_0^π has distribution π. For instance, one can easily infer that

$$\sum_{n \geqslant 1} n^{p-1}\, \mathbb{P}\Big(d(M_n^x, M_n^\pi) > l^n\Big) \ \leqslant \ c_l \Big(1 + \log^p \big(1 + d(x, x_0)\big)\Big)$$

and

$$\lim_{n \to \infty} n^p\, \mathbb{P}\Big(d(M_n^x, M_n^\pi) > l^n\Big) \ = \ 0$$

for each $l \in (l^*, 1)$, $x \in \mathbb{X}$ and some $c_l \in (0, \infty)$.

3. PREREQUISITES

This section will provide a number of auxiliary lemmata necessary to prove our stated results. We keep the notation from Section 2 and make the standing assumption that $(M_n)_{n \geqslant 0}$ is an ultimately contractive IFS of i.i.d. Lipschitz functions satisfying condition (1.1).

3.1. Strong contraction via subsampling. We fix any $\gamma \in (l^*, 1)$ and consider first the sequence $(\sigma_n)_{n \geqslant 0}$ of ladder epochs defined in (1.4).

LEMMA 3.1. *Under the stated assumptions*, σ_1 *has finite mean. Furthermore*, *for any* $p > 0$, $\mathbb{E} \log^{p+1}(1 + L_1) < \infty$ *implies* $\mathbb{E}\sigma_1^{p+1} < \infty$, *while* $\mathbb{E}L_1^p < \infty$ *implies* $\mathbb{E}s^{\sigma_1} < \infty$ *for some* $s > 1$.

Proof. As $\mathbb{E} \log^+ L_1 < \infty$, $n^{-1}\mathbb{E} \log l(F_{1:n}) \to \log l^*$ by Proposition 2 in [9]. Hence there exists $m \in \mathbb{N}$ such that

$$m^{-1}\mathbb{E} \log l(F_{1:m}) \ < \ \log \gamma. \tag{3.1}$$

By subadditivity,

$$\frac{1}{nm} \log l(F_{1:nm}) \ \leqslant \ \frac{1}{n} \sum_{j=1}^{n} \frac{1}{m} \log l(F_{(j-1)m+1:jm})$$

and $\sigma_1 \leqslant m\sigma_1^*$, where

$$\sigma_1^* \stackrel{\text{def}}{=} \inf\left\{n \geqslant 1 : \frac{1}{n}\sum_{j=1}^{n}\frac{1}{m}\log l(F_{(j-1)m+1:jm}) < \log\gamma\right\}.$$

Observe that σ_1^* is the first descending ladder epoch of the random walk

$$\left(\sum_{j=1}^{n}\frac{1}{m}\log l(F_{(j-1)m+1:jm}) - n\log\gamma\right)_{n\geqslant 0}$$

which, by (3.1), has negative drift. Hence $\mathbb{E}\sigma_1 \leqslant m\mathbb{E}\sigma_1^* < \infty$. To prove the remaining moment assertions, recall that

$$\log l(F_{1:m}) \;\leqslant\; \sum_{j=1}^{m}\log L_j$$

which further gives

$$\begin{aligned}&\big(m^{-1}\log l(F_{1:m}) - \log\gamma\big)^+\\ &\qquad\leqslant m^{-1}\sum_{j=1}^{m}\log^+ L_j + |\log\gamma| \leqslant m^{-1}\sum_{j=1}^{m}\log(1+L_j) + |\log\gamma|.\end{aligned}$$

This shows that $\mathbb{E}\log^{p+1}(1+L_1) < \infty$ implies

$$\mathbb{E}\Big(\big(m^{-1}\log l(F_{1:m}) - \log\gamma\big)^+\Big)^{p+1} < \infty,$$

and then $\mathbb{E}\sigma_1^{p+1} \leqslant m^{p+1}\mathbb{E}(\sigma_1^*)^{p+1} < \infty$ by Theorem III.3.1 in [10]. Furthermore,

$$\begin{aligned}\mathbb{E}\exp\left(mp\left(\frac{1}{m}\log l(F_{1:m}) - \log\gamma\right)\right) &\leqslant \gamma^{-mp}\,\mathbb{E}l(F_{1:m})^p\\ &\leqslant \gamma^{-mp}\mathbb{E}\big(\prod_{j=1}^{m}L_j\big)^p = \gamma^{-mp}\left(\mathbb{E}L_1^p\right)^m,\end{aligned}$$

where the last equality holds because the L_j are i.i.d. Now, if $\mathbb{E}L_1^p$ is finite, then the same holds true for $\mathbb{E}\exp\big(mp\left(m^{-1}\log l(F_{1:m}) - \log\gamma\right)\big)$, and we infer that $\mathbb{E}s^{\sigma_1} < \infty$ for some $s > 1$ by an appeal to Theorem III.3.2 in [10]. ■

The following notation is taken from [2], however, with the σ_n as in the previous lemma. Keeping $\gamma \in (l^*, 1)$ fixed, put

$$\begin{aligned}\tau(n) &\stackrel{\text{def}}{=} \inf\{j \geqslant 0 : \sigma_j \geqslant n\},\\ C_{n+1} &\stackrel{\text{def}}{=} \max\big\{d\big(F_{\sigma_n+1:\sigma_{n+1}}(x_0), x_0\big);\\ &\qquad\qquad d\big(F_{\sigma_n+1:\sigma_{n+1}}(x_0), F_{\sigma_n+1:k}(x_0)\big), \sigma_n < k < \sigma_{n+1}\big\},\\ D_n &\stackrel{\text{def}}{=} \sum_{j\geqslant 0}\gamma^j d\big(F_{\sigma_{n+j}+1:\sigma_{n+j+1}}(x_0), x_0\big)\end{aligned}$$

for $n \geqslant 0$, where x_0 is given by (1.1). We continue by listing a number of facts that have been pointed out or proved in [2] and carry over to the present situation without further ado:

(P1) As a consequence of (1.1), $\mathbb{E}\log(1+C_1) < \infty$ and the D_n are a.s. finite.

(P2) $(C_n)_{n\geqslant 1}$ and $(F_{\sigma_n+1:\sigma_{n+1}})_{n\geqslant 0}$ are both sequences of i.i.d. random variables.

(P3) $C_{\tau(n)}$ converges in distribution to a limiting variable C_∞ with distribution function

$$\mathbb{P}(C_\infty \leqslant t) \;=\; (\mathbb{E}\sigma_1)^{-1}\mathbb{E}\sigma_1 \mathbf{1}_{\{C_1\leqslant t\}}, \quad t \geqslant 0.$$

Furthermore, $\mathbb{P}(C_{\tau(n)} \in \cdot) \leqslant \mathbb{E}\sigma_1\,\mathbb{P}(C_\infty \in \cdot) = \mathbb{E}\sigma_1 \mathbf{1}_{\{C_1\in\,\cdot\}}$ for all $n \geqslant 0$.

(P4) As a consequence of (P2), $(D_n)_{n\geqslant 0}$ is stationary, and it is autoregressive of order one, viz. $D_n = d\big(F_{\sigma_n+1:\sigma_{n+1}}(x_0), x_0\big) + \gamma D_{n+1}$.

(P5) For each $n \geqslant 0$, $D_{\tau(n)}$ is independent of $\tau(n)$ and $(F_j, L_j)_{1\leqslant j\leqslant \sigma_{\tau(n)}}$ with the same distribution as D_0, for $(F_{\sigma_{\tau(n)}+k})_{k\geqslant 1} \stackrel{d}{=} (F_k)_{k\geqslant 1}$ for each n.

(P6) The C_n and D_n are linked by the inequality

$$D_n \;\leqslant\; \sum_{j\geqslant 1} \gamma^{j-1} C_{n+j} \quad \text{a.s.}$$

The next lemma provides the crucial estimates for *strongly contractive* IFS of i.i.d. Lipschitz maps and will be subsequently utilized for the system obtained by subsampling our given IFS along $(\sigma_n)_{n\geqslant 0}$.

LEMMA 3.2. *Let $(M_n)_{n\geqslant 0}$ be an IFS of i.i.d. Lipschitz maps satisfying* (1.1) *and the strong contraction condition*

$$L_1 \;\leqslant\; \gamma \quad \textit{a.s.}$$

for some $\gamma \in (0,1)$. Then the D_n are a.s. finite and

$$\sup_{m\geqslant 1} d\big(F_{n+1:n+m}(x_0), x_0\big) \;\leqslant\; D_n \quad \textit{a.s.} \tag{3.2}$$

for all $n \geqslant 0$. Furthermore,

$$d(\hat{M}_\infty^{x_0}, \hat{M}_n^{x_0}) \;\leqslant\; L_{1:n} D_n \quad \textit{a.s.}, \tag{3.3}$$

$$d(\hat{M}_\infty^{x_0}, \hat{M}_n^{x}) \;\leqslant\; L_{1:n}\big(D_n + d(x, x_0)\big) \quad \textit{a.s.} \tag{3.4}$$

for all $n \geqslant 0$ and $x \in \mathbb{X}$.

Proof. The proof is essentially the same as the one for Lemma 3.1 in [2], and therefore omitted. ∎

LEMMA 3.3. *Given the previously introduced notation*,

$$d(\hat{M}_\infty^{x_0}, \hat{M}_n^x) \;\leqslant\; \gamma^{\sigma_{\tau(n)-1}} C_{\tau(n)} + \gamma^{\sigma_{\tau(n)}} D_{\tau(n)} + l(F_{1:n}) d(x, x_0) \quad \textit{a.s.}$$

holds true for each $n \geqslant 0$.

Proof. Putting $F_n' \stackrel{\text{def}}{=} F_{\sigma_{n-1}+1:\sigma_n}$ and $L_n' \stackrel{\text{def}}{=} l(F_n')$, we have by (1.4)

$$L_n' \;\leqslant\; \gamma^{\sigma_n - \sigma_{n-1}} \quad \text{a.s.}$$

The (F_n', L_n'), $n \geqslant 1$, are further i.i.d., so that $M_n' \stackrel{\text{def}}{=} F_{n:1}'(M_0)$, $n \geqslant 0$, is a strongly contractive IFS with backward iterations $\hat{M}_n'$ satisfying

$$\hat{M}_n' \;=\; F_{1:n}'(M_0) \;=\; \hat{M}_{\sigma_n}, \quad n \geqslant 0.$$

Notice, however, that M_n' generally differs from M_{σ_n}. An application of (3.3) to $(\hat{M}_n')_{n \geqslant 0}$ leads to

$$d(\hat{M}_\infty^{x_0}, \hat{M}_{\sigma_{\tau(n)}}^{x_0}) \;\leqslant\; L_{1:\tau(n)}' D_{\tau(n)} \;\leqslant\; \gamma^{\sigma_{\tau(n)}} D_{\tau(n)} \quad \text{a.s.}$$

Moreover,

$$\begin{aligned} d(\hat{M}_{\sigma_{\tau(n)}}^{x_0}, \hat{M}_n^{x_0}) &= d\Big(F_{1:\tau(n)-1}'\big(F_{\tau(n)}'(x_0)\big), F_{1:\tau(n)-1}'\big(F_{\sigma_{\tau(n)-1}+1:n}(x_0)\big)\Big) \\ &\leqslant L_{1:\tau(n)-1}' d\big(F_{\tau(n)}'(x_0), F_{\sigma_{\tau(n)-1}+1:n}(x_0)\big) \\ &\leqslant \gamma^{\sigma_{\tau(n)-1}} C_{\tau(n)} \quad \text{a.s.} \end{aligned}$$

and $d(\hat{M}_n^{x_0}, \hat{M}_n^x) \leqslant l(F_{1:n}) d(x, x_0)$ a.s. By combining these estimates with

$$d(\hat{M}_\infty^{x_0}, \hat{M}_n^x) \;\leqslant\; d(\hat{M}_\infty^{x_0}, \hat{M}_{\sigma_{\tau(n)}}^{x_0}) + d(\hat{M}_{\sigma_{\tau(n)}}^{x_0}, \hat{M}_n^{x_0}) + d(\hat{M}_n^{x_0}, \hat{M}_n^x)$$

we arrive at the assertion of the lemma. ∎

3.2. Moment and tail probability results. In order to prove our theorems, the following moment and tail probability results are needed, which have also been stated in [2] (see Section 3 therein). Proofs are therefore omitted except for Lemma 3.8 which requires an extra argument. Regarding Lemmata 3.4 and 3.6, let us further note that the stated inequalities for $\sup_{n \geqslant 0} \mathbb{E} \log^p(1 + C_{\tau(n)})$ and $\sup_{n \geqslant 0} \mathbb{E} C_{\tau(n)}^\eta$ in terms of the respective moments of C_∞ are always true under our standing assumptions on the given IFS and in fact direct consequences of the final inequality in (P3). Hence, it is the finiteness of $\mathbb{E} \log^p(1 + C_\infty)$ under (1.7) and of $\mathbb{E} C_\infty^\eta$ for some $\eta > 0$ under (1.8), which needs really to be verified in these lemmata.

LEMMA 3.4. *Given $p > 0$, suppose that* (1.7) *holds true. Then*

$$\mathbb{E}\log^{p+1}(1+C_1) \;<\; \infty, \qquad \mathbb{E}\log^{p}(1+D_0) \;<\; \infty,$$

and the family $\{\log^p(1+C_{\tau(n)}) : n \geqslant 0\}$ is uniformly integrable with

$$\sup_{n\geqslant 0}\mathbb{E}\log^p(1+C_{\tau(n)}) \;\leqslant\; \mathbb{E}\sigma_1\,\mathbb{E}\log^p(1+C_\infty) \;=\; \mathbb{E}\sigma_1\log^p(1+C_1) \;<\; \infty.$$

LEMMA 3.5. *Given $p > 0$, suppose that $\mathbb{E}\log^{p+1}(1+L_1) < \infty$. Then*

$$\sum_{n\geqslant 1} n^{p-1}\,\mathbb{P}\big(l(F_{1:n}) > l^n\big) \;<\; \infty$$

and

$$\lim_{n\to\infty} n^p\,\mathbb{P}\big(l(F_{1:n}) > l^n\big) \;=\; 0$$

hold true for any $l \in (l^*, 1)$. Furthermore,

$$\sum_{n\geqslant 1} n^{p-1}\,\mathbb{P}\big(\sigma_{\tau(n)-1} \leqslant (1-\rho)n\big) \;<\; \infty$$

and

$$\lim_{n\to\infty} n^p\,\mathbb{P}\big(\sigma_{\tau(n)-1} \leqslant (1-\rho)n\big) \;=\; 0$$

hold true for all $\rho > 0$.

LEMMA 3.6. *Given $p > 0$, suppose that* (1.8) *holds true. Then there exists $\eta > 0$ such that*

$$\mathbb{E}C_1^{2\eta} \;<\; \infty, \qquad \mathbb{E}D_0^{2\eta} \;<\; \infty,$$

and the family $\{C_{\tau(n)}^\eta : n \geqslant 0\}$ is uniformly integrable with

$$\sup_{n\geqslant 0}\mathbb{E}C_{\tau(n)}^\eta \;\leqslant\; \mathbb{E}\sigma_1\,\mathbb{E}C_\infty^\eta \;=\; \mathbb{E}\sigma_1 C_1^\eta \;<\; \infty.$$

LEMMA 3.7. *Given $p > 0$, suppose that $\mathbb{E}L_1^p < \infty$. Then*

$$\lim_{n\to\infty} \alpha^{-n}\,\mathbb{P}\big(\sigma_{\tau(n)-1} \leqslant (1-\rho)n\big) \;=\; 0$$

for all $\rho > 0$ and some $\alpha = \alpha_\rho \in (0,1)$.

LEMMA 3.8. *Given $p > 0$, suppose that $\mathbb{E}L_1^p < \infty$. Then there exist $q \in (0,p]$ and $k \geqslant 1$ such that $\mathbb{E}l(F_{1:k})^q < 1$ and $\{l(F_{1:n})^q : n \geqslant 1\}$ is uniformly integrable. Furthermore,*

$$\mathbb{P}\big(l(F_{1:n}) > \varepsilon l^n\big) \;\leqslant\; K_\varepsilon l^n$$

for all $n \geqslant 1$, $\varepsilon > 0$ and suitable $l \in (0,1)$, $K_\varepsilon > 0$.

Proof. Since $n^{-1}\,\mathbb{E}\log l(F_{1:n}) \to \log l^* < 0$ as $n \to \infty$, we can fix $k \geqslant 1$ such that $\mathbb{E}\log l(F_{1:k}) < 0$. Moreover, $\mathbb{E}l(F_{1:k})^p \leqslant \mathbb{E}L_{1:k}^p = (\mathbb{E}L_1^p)^k$, by subadditivity of $\log l(F_{1:n})$. Consequently, the function $[0,p] \ni q \mapsto \mathbb{E}l(F_{1:k})^q$ is everywhere finite and convex with value 1 and negative right-hand derivative at 0. This allows us to pick a q with $m_q \stackrel{\text{def}}{=} \mathbb{E}l(F_{1:k})^q < 1$. Put $K = 1 \vee \max_{1\leqslant j<k} \mathbb{E}l(F_{1:j})^q$ and note that $\mathbb{E}l(F_{1:n})^q \leqslant K m_q^j$ if $n = jk + r$ with $r \in \{0,\ldots,k-1\}$. Hence $l(F_{1:n})^q \to 0$ a.s. and in $\mathfrak{L}_1$, which particularly ensures uniform integrability. Also, by an appeal to Markov's inequality, we infer for all $l \in (0,1)$ sufficiently large and all $\varepsilon > 0$ that

$$\mathbb{P}\big(l(F_{1:n}) > \varepsilon l^n\big) \;\leqslant\; \frac{K}{m_q \varepsilon^q}\left(\frac{m_q^{1/k}}{l^q}\right)^n \;\leqslant\; K_\varepsilon l^n$$

for all $n \geqslant 0$, where $K_\varepsilon \stackrel{\text{def}}{=} K/(m_q\varepsilon^q)$. ∎

4. PROOFS OF THE MAIN RESULTS

Proof of Theorem 2.1. Pick any $l \in (l^*,1)$, and then $\gamma \in (l^*, l)$ for which we define the σ_n, C_n and D_n as in Section 3. It follows that $l = \gamma^{1-\varepsilon}$ for some $\varepsilon > 0$, and we infer with the help of Lemma 3.3 that

$$\begin{aligned}
\mathbb{P}_x\big(d(\hat M_\infty, \hat M_n) > l^n\big) \;&=\; \mathbb{P}\big(d(\hat M_\infty^{x_0}, \hat M_n^x) > \gamma^{n(1-\varepsilon)}\big)\\
&\leqslant\; \mathbb{P}\left(\gamma^{\sigma_{\tau(n)-1}-n(1-\varepsilon)}(C_{\tau(n)} + D_{\tau(n)}) > \frac{1}{2}\right)\\
&\quad + \mathbb{P}\left(\left(\frac{l(F_{1:n})^{1/n}}{l}\right)^n d(x,x_0) > \frac{1}{2}\right).
\end{aligned}$$

But the latter two probabilities converge to 0 as $n \to \infty$, for $l(F_{1:n})^{1/n} \to l^* < l$ a.s., $(C_{\tau(n)} + D_{\tau(n)})_{n\geqslant 0}$ forms a tight sequence by (P3)–(P5), and

$$\sigma_{\tau(n)-1} - n(1-\varepsilon) \;=\; \tau(n)\left(\frac{\sigma_{\tau(n)-1}}{\tau(n)} - (1-\varepsilon)\frac{n}{\tau(n)}\right) \;\to\; \infty \quad \text{a.s.}$$

For the last convergence we have used that $\tau(n) \to \infty$ a.s., $n^{-1}\tau(n) \to (\mathbb{E}\sigma_1)^{-1}$ a.s. by the elementary renewal theorem, and $n^{-1}\sigma_n \to \mathbb{E}\sigma_1$ a.s. by the strong law of large numbers. ∎

Proof of Theorem 2.2. (a) Pick l, γ, ε as before and embark on the inequality

$$\begin{aligned}
(4.1)\quad \mathbb{P}\big(d(\hat M_\infty^{x_0}, \hat M_n^x) > l^n\big) \;&\leqslant\; \mathbb{P}\big(\sigma_{\tau(n)-1} \leqslant (1-\rho)n\big)\\
&\quad +\; \mathbb{P}(\gamma^{(1-\rho)n} C_{\tau(n)} > l^n/3) + \mathbb{P}(\gamma^{(1-\rho)n} D_{\tau(n)} > l^n/3)\\
&\quad +\; \mathbb{P}\big(l(F_{1:n}) > \gamma^n\big) \;+\; \mathbf{1}\big(\gamma^n d(x,x_0) > l^n/3\big),
\end{aligned}$$

which is again easily obtained with Lemma 3.3. As $\gamma < l$, the last indicator becomes 0 for $n > n_0 \stackrel{\text{def}}{=} \log 3d(x,x_0)/\log(l/\gamma)$, and this gives

$$\sum_{n\geqslant 1} n^{p-1}\mathbf{1}\big(\gamma^n d(x,x_0) > l^n/3\big) \leqslant n_0^p \;\leqslant\; c\,\log^p\big(1+d(x,x_0)\big)$$

for some $c > 0$. Moreover, fixing any $\rho \in (1,\varepsilon)$, we see that

$$\sum_{n\geqslant 1} n^{p-1}\mathbb{P}\big(\sigma_{\tau(n)-1} \leqslant (1-\rho)n\big) \quad \text{and} \quad \sum_{n\geqslant 1} n^{p-1}\mathbb{P}\big(l(F_{1:n}) > \gamma^n\big)$$

are finite and the appearing probabilities of the order $o(n^{-p})$ for $n \to \infty$ by Lemma 3.5. Next, by using (P3) and recalling $l = \gamma^{1-\varepsilon}$, we obtain

$$\begin{aligned}
&\sum_{n\geqslant 1} n^{p-1}\mathbb{P}(\gamma^{(1-\rho)n}C_{\tau(n)} > l^n/3)\\
&\qquad\leqslant\; \sum_{n\geqslant 1} n^{p-1}\mathbb{P}(3C_{\tau(n)} > \gamma^{n(\rho-\varepsilon)})\\
&\qquad\leqslant\; \sum_{n\geqslant 1} n^{p-1}\mathbb{P}\big(\log(1+3C_{\tau(n)}) > n(\varepsilon-\rho)|\log\gamma|\big)\\
&\qquad\leqslant\; \mathbb{E}\sigma_1 \sum_{n\geqslant 1} n^{p-1}\mathbb{P}\big(\log(1+3C_\infty) > n(\varepsilon-\rho)|\log\gamma|\big)\\
&\qquad\leqslant\; K\,\mathbb{E}\log^p(1+C_\infty)
\end{aligned}$$

for some $K > 0$, and the last moment is finite by Lemma 3.4. But the last fact further implies

$$\mathbb{P}(\gamma^{(1-\rho)n}C_{\tau(n)} > l^n/3) \leqslant \mathbb{E}\sigma_1\,\mathbb{P}\big(\log(1+3C_\infty) > n(\varepsilon-\rho)|\log\gamma|\big) = o(n^{-p})$$

as $n \to \infty$. Finally, as $D_{\tau(n)} \stackrel{d}{=} D_0$ for all $n \geqslant 0$ by (P5), we have

$$\begin{aligned}
&\sum_{n\geqslant 1} n^{p-1}\mathbb{P}(\gamma^{(1-\rho)n}D_{\tau(n)} > l^n/3)\\
&\qquad\leqslant\; \sum_{n\geqslant 1} n^{p-1}\mathbb{P}\big(\log(1+3D_0) > n(\varepsilon-\rho)|\log\gamma|\big)\\
&\qquad\leqslant\; K\,\mathbb{E}\log^p(1+D_0)
\end{aligned}$$

for some $K > 0$, the last moment being finite by another appeal to Lemma 3.4. As a consequence, we get

$$\mathbb{P}(\gamma^{(1-\rho)n}D_{\tau(n)} > l^n/3) \leqslant \mathbb{P}\big(\log(1+3D_0) > n(\varepsilon-\rho)|\log\gamma|\big) = o(n^{-p})$$

as $n \to \infty$. A combination of these facts with (4.1) proves (a).

(b) By another appeal to our key inequality, stated as Lemma 3.3, we infer that

$$\begin{aligned}(4.2)\quad & n^{(p-1)/p}\left(\frac{1}{n}\log d(\hat M_\infty^{x_0},\hat M_n^x)-\log l\right)\\ &\leqslant\ n^{(p-1)/p}\log\gamma\left(\frac{\sigma_{\tau(n)-1}}{n}-\frac{\log l}{\log\gamma}\right)\\ &\quad+\ n^{-1/p}\log\big(C_{\tau(n)}+D_{\tau(n)}+\gamma^{-\sigma_{\tau(n)-1}}l(F_{1:n})d(x,x_0)\big)\quad \text{a.s.}\end{aligned}$$

for each $n\geqslant 0$ (with l,γ as before). Therefore, we must prove that the limsup of the right-hand side of this inequality is less than or equal to 0 as $n\to\infty$. By the arguments given at the end of the proof of Theorem 2.1, we obtain

$$\lim_{n\to\infty}\frac{\sigma_{\tau(n)-1}}{n}-\frac{\log l}{\log\gamma}\ =\ 1-\frac{\log l}{\log\gamma}\ >\ 0\quad\text{a.s.},$$

and thus (as $\log\gamma<0$)

$$\limsup_{n\to\infty} n^{(p-1)/p}\log\gamma\left(\frac{\sigma_{\tau(n)-1}}{n}-\frac{\log l}{\log\gamma}\right)\ \leqslant\ 0\quad\text{a.s.}$$

If $p\in(0,1)$, this remains even true with l^* instead of l (in this case pick any $\gamma\in(l^*,1)$) because $n^{-(p-1)/p}\to 0$. As for the second term on the right-hand side in (4.2), it suffices to note that it converges a.s. to 0, for this holds true for $\gamma^{-\sigma_{\tau(n)-1}}l(F_{1:n})d(x,x_0)$ and $\{\log^p(C_{\tau(n)}+D_{\tau(n)}):n\geqslant 0\}$ is uniformly integrable by Lemma 3.4 and (P5).

(c) As (b) for $p=1$ and any fixed $x\in\mathbb{X}$ may be restated as

$$d(\hat M_\infty^{x_0},\hat M_n^x)\ =\ (l^*R_n)^n,\quad n\geqslant 0,$$

for suitable random variables $R_n\geqslant 0$ satisfying $\limsup_{n\to\infty}R_n\leqslant 1$ a.s., we have for any $l\in(l^*,1)$

$$\lim_{n\to\infty} l^{-n}d(\hat M_\infty^{x_0},\hat M_n^x)\leqslant\lim_{n\to\infty}\left(\frac{l^*R_n}{l}\right)^n\ =\ 0\quad\text{a.s.},$$

as claimed.

(d) Fixing any l,γ as before and $\rho\in(0,1)$, we can choose a sufficiently large constant $A>1$ such that, by Lemma 3.5,

$$\begin{aligned}&\sup_{n\geqslant 0}(n+1)^p\,\mathbb{P}\big(l(F_{1:n})>(n+1)^{-p}/2\big)\leqslant A/3,\\ &\sup_{n\geqslant 0}(n+1)^p\,\mathbb{P}\big(\sigma_{\tau(n)-1}\leqslant(1-\rho)n\big)\leqslant A/3,\end{aligned}$$

and, by uniform integrability of $\{\log^p(C_{\tau(n)} + D_{\tau(n)}) : n \geqslant 0\}$,

$$\sup_{n \geqslant 0} (n+1)^p \, \mathbb{P}\big(\gamma^{(1-\rho)n}(C_{\tau(n)} + D_{\tau(n)}) > A(n+1)^{-p}/2\big) \leqslant A/3$$

holds true. Put $A_x = \max\{A, 2d(x, x_0)\}$. Then, by a similar estimation as in (4.1),

$$\begin{aligned}
&\mathbb{P}\big(d(\hat{M}_\infty^{x_0}, \hat{M}_n^x) > A_x(n+1)^{-p}\big) \\
&\quad \leqslant \mathbb{P}\big(\sigma_{\tau(n)-1} \leqslant (1-\rho)n\big) + \mathbb{P}\big(l(F_{1:n})d(x, x_0) > A_x(n+1)^{-p}/2\big) \\
&\qquad + \mathbb{P}\big(\gamma^{(1-\rho)n}(C_{\tau(n)} + D_{\tau(n)}) > A_x(n+1)^{-p}/2\big) \\
&\quad \leqslant \mathbb{P}\big(\sigma_{\tau(n)-1} \leqslant (1-\rho)n\big) + \mathbb{P}\big(l(F_{1:n}) > (n+1)^{-p}/2\big) \\
&\qquad + \mathbb{P}\big(\gamma^{(1-\rho)n}(C_{\tau(n)} + D_{\tau(n)}) > A(n+1)^{-p}/2\big) \\
&\quad \leqslant A(n+1)^{-p} \leqslant A_x(n+1)^{-p}
\end{aligned}$$

for all $n \geqslant 0$, and this leads to the desired conclusion by invoking Lemma 5.8 in [8] (also stated as Lemma 3.6 in [2]).

(e) This can be copied verbatim from [2] and is therefore omitted. ■

P r o o f o f T h e o r e m 2.3. In view of the previously provided arguments it is now rather straightforward to adapt the proof of Theorem 2.3 in [2] to the present situation, and we therefore restrict ourselves to a proof of part (b).

(b) Choose $\eta \in (0, p \wedge 1)$ such that Lemma 3.6 is valid and fix an arbitrary $q \in (0, \eta]$. By Lemma 3.8, $m_q = \mathbb{E}l(F_{1:k})^q < 1$ for some $k \geqslant 1$. By another appeal to Lemma 3.3 and a simple estimation, we get

$$\big(1 + d(x, x_0)\big)^{-q} d(\hat{M}_\infty^{x_0}, \hat{M}_n^x)^q \leqslant \gamma^{q\sigma_{\tau(n)-1}}(C_{\tau(n)} + D_{\tau(n)})^q + l(F_{1:n})^q \quad \text{a.s.}$$

Observe that the right-hand side does not depend on $x \in \mathbb{X}$, converges to 0 in probability and is uniformly integrable by Lemmata 3.6 and 3.8. This proves the assertion for $q = \eta$. If $q < \eta$, use Hölder's inequality to obtain

$$\begin{aligned}
&\mathbb{E}\gamma^{q\sigma_{\tau(n)-1}}(C_{\tau(n)} + D_{\tau(n)})^q \\
&\qquad\qquad \leqslant \big(\mathbb{E}\gamma^{\eta q\sigma_{\tau(n)-1}/(\eta-q)}\big)^{(\eta-q)/\eta} \big(\mathbb{E}(C_{\tau(n)} + D_{\tau(n)})^\eta\big)^{q/\eta}
\end{aligned}$$

for each $n \geqslant 1$. As $\mathbb{E}l(F_{1:n})^q \leqslant K m_q^{n/k}$ (see the proof of Lemma 3.8), it remains to show that

$$\lim_{n\to\infty} \alpha_q^{-n} \big(\mathbb{E}\gamma^{\eta q\sigma_{\tau(n)-1}/(\eta-q)}\big)^{(\eta-q)/\eta} = 0$$

for some $\alpha_q \in [m_q, 1)$. To this end we further estimate

$$\begin{aligned}&\big(\mathbb{E}\gamma^{\eta q \sigma_{\tau(n)-1}/(\eta-q)}\big)^{(\eta-q)/\eta}\\ &\qquad\qquad \leqslant \big(\mathbb{E}\gamma^{\eta q \sigma_{\tau(n)-1}/(\eta-q)}\mathbf{1}_{\{\sigma_{\tau(n)-1}\leqslant(1-\rho)n\}}\big)^{(\eta-q)/\eta} + \gamma^{(1-\rho)n}\\ &\qquad\qquad \leqslant \mathbb{P}\big(\sigma_{\tau(n)-1} \leqslant (1-\rho)n\big)^{(\eta-q)/\eta} + \gamma^{(1-\rho)n},\end{aligned}$$

which holds for any $\rho \in (0,1)$ and $n \geqslant 0$. By invoking Lemma 3.7, we thus arrive at the desired conclusion. ∎

REFERENCES

[1] G. Alsmeyer, *On the Harris recurrence of iterated random Lipschitz functions and related convergence rate results*, J. Theoret. Probab. 16 (2003), pp. 217–247.

[2] G. Alsmeyer and C. D. Fuh, *Limit theorems for iterated random functions by regenerative methods*, Stochastic Process. Appl. 96 (2001), pp. 123–142. *Corrigendum*: 97 (2002), pp. 341–345.

[3] L. Arnold and H. Crauel, *Iterated function systems and multiplicative ergodic theory*, in: *Diffusion Processes and Related Problems in Analysis* 2, M. Pinsky and V. Wihstutz (Eds.), Birkhäuser, Boston 1992, pp. 283–305.

[4] M. Babillot, P. Bougerol and L. Elie, *The random difference equation $X_n = A_n X_{n-1} + B_n$ in the critical case*, Ann. Probab. 25 (1997), pp. 478–493.

[5] M. F. Barnsley and J. H. Elton, *A new class of Markov processes for image encoding*, Adv. in Appl. Probab. 20 (1988), pp. 14–32.

[6] M. Benda, *A central limit theorem for contractive dynamical systems*, J. Appl. Probab. 35 (1998), pp. 200–205.

[7] Y. S. Chow and T. L. Lai, *Some one-sided theorems on the tail distribution of sample sums with applications to the last time and largest excess of boundary crossings*, Trans. Amer. Math. Soc. 208 (1975), pp. 51–72.

[8] P. Diaconis and D. Freedman, *Iterated random functions*, SIAM Rev. 41 (1999), pp. 45–76.

[9] J. H. Elton, *A multiplicative ergodic theorem for Lipschitz maps*, Stochatic Process. Appl. 34 (1990), pp. 39–47.

[10] A. Gut, *Stopped Random Walks: Limit Theorems and Applications*, Springer, New York 1988.

[11] J. E. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. 30 (1981), pp. 713–747.

[12] G. Letac, *A contraction principle for certain Markov chains and its applications*, Contemp. Math. 50 (1986), pp. 263–273.

[13] D. Silvestrov and Ö. Stenflo, *Ergodic theorems for iterated function systems controlled by regenerative sequences*, J. Theoret. Probab. 11 (1998), pp. 589–608.

[14] D. Steinsaltz, *Locally contractive iterated function systems*, Ann. Probab. 27 (1999), pp. 1952–1979.

Institut für Mathematische Statistik, FB 10
Westfälische Wilhelms-Universität Münster
Einsteinstrasse 62, 48149 Münster, Germany
E-mail: gerolda@math.uni-muenster.de
E-mail: gerd.hoelker@web.de

Received on 20.6.2008;
revised version on 20.3.2009