RANDOM LOGISTIC MAPS AND LYAPUNOV EXPONENTS

DAVID STEINSALTZ

ABSTRACT. We prove that under certain basic regularity conditions, a random iteration of logistic maps converges to a random point attractor when the Lyapunov exponent is negative, and does not converge to a point when the Lyapunov exponent is positive.

1. INTRODUCTION

One of the fundamental questions about a random dynamical system in general, and an iterated function system in particular, is whether its path is absorbed into a single random attracting point. Almost equivalent is the question of when the iterates flatten out to approach a constant function. This is clearly the case when the individual functions are all contractions (discussed by J. Hutchinson [Hut81]), and these results may be extended by similar methods to "average contractive" systems — where the iterated maps do not shrink the distance between two points at every step, but do so everywhere, in expectation — as realized by M. Barnsley and J. Elton [BE88]. We have developed a somewhat new approach in [Ste99], which is viable for systems whose contraction is spatially inhomogeneous as well. (For an extensive review of other work on iterated function systems, see the survey paper of P. Diaconis and D. Freedman [DF99].) There may be regions of the space which are never contracted by the maps, and yet the iterates will converge if the orbit of a point wanders sufficiently around the space to pick up an average contraction. The earlier paper used a variant of Lyapunov drift functions to guarantee proper mixing. This technique has the advantage of being fairly straightforward, when it works, but it demands the hit-and-miss invention of a test function.

This average contraction is witnessed by a negative Lyapunov exponent. In this paper we apply very different methods to substantially resolve one class of examples, the iteration of random logistic maps. "Resolve" must here be understood in a conditional sense, to be sure, since we in fact only reduce it to the nontrivial problem of computing or estimating the Lyapunov exponent. Our methods are also incapable of dealing with distributions on the coefficients that are insufficiently spread out — those concentrated on two points, for instance — and a few other unpalatable restrictions have needed to be swallowed as well.

The largest Lyapunov exponent of a system often gives information about the overall expansion of the system. Negative Lyapunov exponents are associated with the long-term contraction of the space under the random transformation, and hence with the convergence to a random point attractor. This is unequivocal for random affine maps (cf. [AC92]). On the other hand, the information embedded in the Lyapunov exponents is purely local, so that arguments based on them may founder

Date: September 17, 2001.
1991 *Mathematics Subject Classification.* Primary: 39B12, 60G99.

on more global structures. For instance, negative Lyapunov exponents make it possible, but never certain, that a set will shrink to a point under the action of a Brownian flow (cf. [BH86] and [SS02]). The interpretation of Lyapunov exponents becomes particularly vexed when the transformations are not injective. Our goal in this paper is to show that in a paradigm noninjective case — iterated logistic maps of the unit interval — the Lyapunov exponent does arbitrate the existence of a random point attractor. While some computations are specific to this case, the methods are general enough that they could be applied to other discrete-time random iterations.

The discrete logistic family of maps on the unit interval, given by $x \mapsto ux(1-x)$, have long been studied as a simple but illustrative case of nonlinear iteration. (Many applications may be found in the book [Cvi84], and references therein.) As with most such smooth families of interval maps, this logistic family exhibits a wide range of behaviors, in this case as the parameter u rises from 0 to 4. (We will not consider here $u > 4$, when the map leaves the unit interval.) For $u \leq 1$ the iterates simply collapse to 0. Above 1, the fixed point at 0 becomes unstable, and a new fixed point arises which attracts the entire open interval $(0, 1)$. This behavior persists up through $u = 3$, when the period-doubling described by Feigenbaum [Fei84] begins: the fixed point splits into an attractive orbit of period 2, then period 4, and so on, until at last, above the critical parameter $3.57\ldots$ we arive at the realm of "chaotic" behavior, where there are aperiodic orbits. This is lucidly described in [May76], and at greater length in the book by R. Devaney [Dev89].

The behavior of long-term iterates is famously sensitive to the choice of u. There is a stable periodic orbit, but the period is often extremely long. It has been shown (see section V.6 of [dMvS93]) by Jakobson that when the Lyapunov exponent — defined as the single value taken on by

$$\lambda(u, x) := \lim_{n\to\infty} \frac{1}{n} \log\left|Df_u^n(x)\right|$$

for almost every x — is positive, the occupation measure of a generic orbit is absolutely continuous with respect to Lebesgue measure. On the other hand, the set of parameters in any neighborhood of the endpoint 4 for which the Lyapunov exponent is negative has positive Lebesgue measure.

What happens when we mix various parameter values together into the iteration? At first blush one might expect unbridled confusion, far more intractable than the iteration with a fixed parameter value. And yet, it is often the case with such problems that the individual peculiarities of different parameter values will cancel each other out, settling into characteristic behavior over a wide range of settings. We might hope that this would be the case when we iterate with independent randomly chosen parameter values, where the random choice is, in some sense, sufficiently spread out.

When iterating with changing parameter values, we find ourselves with a new ambiguity, which needs to be addressed at the outset. Suppose we have a sequence $u_1, u_2, \ldots$, and we define

$$f_i(x) = u_i x(1-x). \tag{1}$$

There are two ways that we may compose these functions:

$$F_n(x) := f_1\left(f_2(\cdots f_n(x)\cdots)\right) \tag{2}$$

$$\widetilde{F}_n(x) := f_n\left(f_{n-1}(\cdots f_1(x)\cdots)\right) \tag{3}$$

For many choices of the u_i, the "backward iterate" $F_n(x)$ converges as $n \to \infty$, to a constant independent of x. The "forward iterate" $\widetilde{F}_n(x)$, on the other hand, cannot converge, even when it is becoming flat, except in trivial cases.

In this paper we will be supposing the u_i to be i.i.d. choices from a distribution ν on $(0,4)$. The forward iterate is then a Markov chain for any fixed x. The backward iterates, though, despite having the same marginal distribution as the forward, exhibit a more complicated joint structure. Under some circumstances, this process has the property that we have elsewhere called "attractive", by which we mean that $\lim_{n\to\infty} F_n(x)$ exists and is independent of x almost surely. The function F_n then converges to a constant function. The distribution of this random constant is the unique stationary distribution of the Markov chain $\widetilde{F}_n$. Further discussion of these iterated function systems may be found in [Ste99], and in [BE88], where an application of the attractivity property to image-encoding is presented. Attractivity is another name for the existence of a one-point random attractor, in the language of random dynamical systems [Arn98].

Until very recently, this particular problem had received little attention. R. Bhattacharya and B. Rao [BR93] studied the interesting special case when the parameter u is chosen with equal probability from just two possible values. G. Letac and J.-F. Chamayou [CL91] have considered another special case, where $u_i/4$ has a β distribution with parameters $(a+\frac{1}{2}, a-\frac{1}{2})$, for $a \geq \frac{1}{2}$. They showed that $\beta_{a,a}$ is the stationary distribution for this system, but speculated that it is not attractive; that is, the forward iterates converge in distribution to $\beta_{a,a}$, but the backward iterates do not converge pointwise. In our recent paper [Ste99], we showed that the system is attractive for $a \geq 2$, but left the question open for smaller values of a.

While completing the present paper we have received preprints of two new works on related questions. K. B. Athreya and J. Dai have presented in [AD] have presented in a general form some basic results about the invariant measures of random iterations of logistic maps. The other preprint [Klü00], by M. Klünger, examines random logistic maps in the context of random-dynamical-system formalism. Some results of that work overlap with section 4 of this paper, where the attractivity of systems with negative Lyapunov exponents is considered. In one respect, Klünger's work is more general than ours, since it allows the sequence u_i to be an ergodic stationary sequence, not necessarily i.i.d.; the functions he considers are also slightly broader than the logistic family. His ν is also more general than ours, freed from the irreducibility condition that we need to impose on the Markov chain $\widetilde{F}_n$. On the other hand, his results for attractivity are only valid when u is concentrated on $[0,3]$. It is hardly surprising that it should be easier to prove the existence of random attractors in this case, when each f_i has a deterministic attractor. We discuss in section 2.2 why most of the heavy lifting of the present paper — in particular, the only significant use of the irreducibility and independence conditions — arises precisely from the need to incorporate parameter values over 3. (Klünger's paper also includes a different kind of result when the parameters are all in the range between 3 and $(\sqrt{5}+1)$, where the logistic maps have attractive orbits of period

2; and he proves attractivity when u is confined to a narrow interval straddling 3.) The assumption of independence, as opposed to stationarity which is assumed by Klünger, is also required to keep the action within the domain of Markov-chain theory.

One feature which is central to the current paper, but absent when $\nu\big((3,4]\big)=0$, is the Lyapunov exponent. When u_i is constrained to be less than 3, the Lyapunov exponent is always negative. The main result that we show here (Theorems 1 and 2) is that, under fairly general conditions, an iterated logistic function system is attractive precisely when its Lyapunov exponent is negative — except that the case in which the Lyapunov exponent is 0 remains undetermined. The precise results are

Theorem 1. *Suppose ν is logarithmically continuous and the Lyapunov exponent of the corresponding iterated function system is positive. Suppose, too, that the Markov chain $\widetilde{F}_n(x)$ is ψ-irreducible and aperiodic. Then $\lim_{n\to\infty} F_n(x)$ exists almost surely only if x is 0 or 1. In particular, the system is not attractive.*

Theorem 2. *Let ν define a random logistic system F_n with the following properties:*

- *The iterates of ν are dense.*
- *The Lyapunov exponent of the system is negative.*
- $\nu\big((0,3]\big)>0$.
- *For some $\alpha'\in(0,1)$,*

$$\eta_{\alpha'} := \int (4u-u^2)^{-\alpha'}\nu(du) < \infty. \tag{4}$$

Then the system is attractive.

The Lyapunov exponent, and the terms "logarithmically continuous", "dense iterates", and "ψ-irreducible", are defined in section 2. Throughout this paper, the Lyapunov exponent will be, as given by (5) and (6), spatially averaged with respect to the stationary distribution. This conforms to most standard usage, but the term has also been applied in the context of iterated function systems (e.g., [Elt90]) to a spatial supremum: $\lim_{n\to\infty} n^{-1}\log\sup_{x\neq y}\rho(\widetilde{F}_n(x),\widetilde{F}_n(y))/\rho(x,y)$. Except in the trivial case, where $\int\log u\,\nu(du)<0$ and the iterates converge almost surely to the constant 0, the Lipschitz constant of the iterates will always go to ∞, so this supremum Lyapunov exponent is not very useful in the present setting. The convergence to a flat function can only be expected to occur uniformly on compact subets of $(0,1)$.

A consequence is the following almost-complete resolution of the question posed by Letac and Chamayou:

Corollary 3. *When $\nu(\cdot/4)$ is the β distribution with parameters $a+\frac12, a-\frac12$ for some $a>\frac12$, the iterated logistic function system is attractive for $a>1$, and is not attractive for $\frac12<a<1$.*

Proof. Since the distribution is absolutely continuous with respect to Lebesgue measure, it is logarithmically continuous (a condition for Theorem 1, defined in section 2) and since ν is absolutely continuous with respect to Lebesgue measure, its iterates are *a fortiori* dense. The condition (4) is clearly satisfied, and $\int\log u\,\nu(du)=\psi(a+\frac12)-\psi(2a)>0$ as well, where ψ is the digamma function.

Attractivity is thus determined solely by the Lyapunov exponent. Since we know the stationary distribution, we may compute this directly:

$$\int_0^1 \int_0^1 \log(4y|1-2x|) d\beta_{a+\frac{1}{2},a-\frac{1}{2}}(y) d\beta_{a,a}(x).$$

This formula was evaluated, to a limited extent, in [Ste99]. D. Piau has pointed out, in a private communication, that the complicated expression given there can be simplified to

$$\lambda = \frac{1}{2}\big(\psi(1) - \psi(a)\big),$$

The Lyapunov exponent is thus positive precisely when $a > 1$, and negative when $a < 1$. □

2. Notation and preliminary facts

In what follows, $u_1, u_2, \ldots$ will be an i.i.d. sequence taking values in the open interval $(0, 4)$, with distribution ν. We will always use $\mathcal{F}_n$ to denote the σ-algebra generated by $\{u_1, \ldots, u_n\}$. The sequence defines an iterated function system associated to ν, comprising the sequences of random functions f_i, F_n, and $\widetilde{F}_n$ defined by (1)–(3).

For fixed x, the sequence $\widetilde{F}_n(x)$ is a Markov chain. If this chain has a unique invariant measure, we will denote it by π, and call π *attractive* if every initial configuration converges in distribution to π. The system will be called *attractive* if $F_n(x)$ converges almost surely to a limit point $F_\infty(x)$, the limit being independent of x. If the system is attractive then the distribution of $F_\infty(x)$ is the unique invariant measure π for the Markov chain, and π is attractive. We define

$$\lambda_\nu(x) := \mathrm{E} \log |f'(x)| = \log |1-2x| + \int_0^4 \log u \, \nu(du), \tag{5}$$

and the Lyapunov exponent of the iterated function system is

$$\lambda_\nu := \int_0^1 \lambda_\nu(x) \pi(dx). \tag{6}$$

A Markov chain on a state-space $\mathcal{X}$ is called ψ-irreducible if there is a nonzero "irreducibility" measure ϕ defined on $\mathcal{X}$, such that if $A \subset \mathcal{X}$ is any set with $\phi(A) > 0$ and $x \in \mathcal{X}$, then there is an n such that

$$\mathrm{P}\big\{X_n \in A \,\big|\, X_0 = x\big\} > 0. \tag{7}$$

From this the "maximal irreducibility measure" ψ may be defined. (Precise definitions may be found in [MT93].) The Markov chain is aperiodic if the set of n satisfying (7) has greatest common divisor 1, for all x and A.

2.1. **Special notation for Theorem 1.** For x in the interval $[0, 1]$, we define the measure $\pi_x(A) = \mathrm{P}\big\{f(x) \in A\big\}$. A measure ν on $[0, 4]$ will be said to be *logarithmically continuous* if the function

$$\Lambda_\nu(z) := \int_0^4 \log |1 - zu| \nu(du) \tag{8}$$

is finite and continuous for $z \in \big[0, \frac{1}{2}\big]$. Note that finiteness and continuity are automatic on $[0, \frac{1}{4})$.

For a given $x \in (0,1)$, define the sets

$$A_x := \Big\{ y \in (0,1) \,:\, \liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \big|\widetilde{F}_i(x) - \widetilde{F}_i(y)\big| > 0 \text{ a. s. } \Big\}, \tag{9}$$

$$B_x := \Big\{ y \in (0,1) \,:\, \liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\big|\widetilde{F}_i(x) - \widetilde{F}_i(y)\big| > 0 \Big\}. \tag{10}$$

$$C_x^n(\epsilon) := \Big\{ y \in (0,1) \,:\, \mathrm{P}\big\{ \big|F_n(x) - F_n(y)\big| > \epsilon \big\} \geq \epsilon \Big\}, \tag{11}$$

$$C_x(\epsilon,\delta) := \Big\{ y \in (0,1) \,:\, \liminf_{n\to\infty} \frac{1}{n} \# \Big\{ i \leq n \,:\, y \in C_x^i(\epsilon) \Big\} \geq \delta \Big\}, \text{ and} \tag{12}$$

Lemma 4. *For every x,*

$$A_x \subset B_x \subset \bigcup_{\epsilon>0} C_x(\epsilon,\epsilon). \tag{13}$$

If the sequence $F_n(x)$ converges almost surely, then $\pi_x(A_x) = \pi_x(B_x) = 0$.

Proof. If $y \in A_x$ then

$$\mathrm{E}\big[\liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} |\widetilde{F}_i(x) - \widetilde{F}_i(y)|\big] > 0.$$

Fatou's Lemma then implies that $y \in B_x$; so $A_x \subset B_x$.

Suppose y is in B_x, and let

$$3\epsilon = \liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\big|F_i(x) - F_i(y)\big| > 0.$$

Since $X_i := \big|F_i(x) - F_i(y)\big| \leq 1$,

$$\mathrm{E}\big|F_i(x) - F_i(y)\big| \leq 2\epsilon + \mathbf{1}\big\{\mathrm{P}\{X_i \geq \epsilon\} > \epsilon\big\},$$

so that

$$3\epsilon = \liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathrm{E}\big|F_i(x) - F_i(y)\big| \leq 2\epsilon + \liminf_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}\big\{\mathrm{P}\{X_i > \epsilon\} > \epsilon\big\},$$

and y is in $C_x(\epsilon,\epsilon)$.

If $F_n(x)$ converges almost surely, the differences $\big|F_n(x) - F_{n+1}(x)\big|$ must go to zero in probability, and $\mathrm{P}\big\{f_{n+1}(x) \in C_x^n(\epsilon)\big\}$ goes to 0 for any positive ϵ, as n goes to ∞. Define the function $\phi_n(y) := n^{-1} \sum_{i=1}^{n} \mathbf{1}\big\{y \in C_x^i(\epsilon)\big\}$. Then

$$\int \phi_n(y) \pi_x(dy) = \frac{1}{n} \sum_{i=1}^{n} \pi_x\left(C_x^n(\epsilon)\right) \xrightarrow{n\to\infty} 0. \tag{14}$$

But we also know that if $y \in C_x(\epsilon,\epsilon)$, then $\liminf_{n\to\infty} \phi_n(y) \geq \epsilon$. Together with (14) this shows that $\pi_x(C_x(\epsilon,\epsilon)) = 0$. Since this is true for every positive ϵ, it follows that $\pi_x(A_x) = \pi_x(B_x) = 0$. $\square$

2.2. **Special notation for Theorem 2.** Theorem 2 relies fundamentally on the theory of general-state-space Markov chains, as expounded most thoroughly by S. Meyn and R. Tweedie in [MT93]. We have already introduced ψ-irreducibility. Another Markov-chain concept which will surface occasionally in this discussion is that of "petite" sets. A set $C \subset \mathfrak{X}$ is *petite* if a nontrivial Borel measure μ on $\mathfrak{X}$ may be found, together with a sequence $a_1, a_2, \dots$, where $\sum a_i = 1$, such that for any Borel set A,

$$\inf_{x \in C} \sum_{i=1}^{\infty} a_i \, \mathrm{P}_x^i(A) \geq \mu(A). \tag{15}$$

A Markov chain is *weakly Feller* if, for every open set U, the function $x \mapsto P_x(U)$ is lower semicontinuous. Our chain $\widetilde{F}_n(x)$ is weakly Feller. Proposition 6.2.8 of [MT93] implies that if $\widetilde{F}_n(x)$ is ψ-irreducible, and the support of ψ has nonempty interior, then all compact sets are petite.

We say the measure ν has *dense iterates* if there is an interval $\mathcal{I} \subset (0,1)$ such that for all $x \in (0,1)$,

$$\left\{ f_{u_n} \circ f_{u_{n-1}} \circ \cdots \circ f_{u_1}(x) \, : \, u_1, u_2, \dots, u_n \in \mathrm{supp}\ \nu \right\} \tag{16}$$

is dense in $\mathcal{I}$. It is shown in [BR93] that this can be the case if ν is supported on just two points, but it can also fail. One serviceable criterion is the following:

Proposition 5. *If the support of ν is dense in some interval and $\nu\big((0,3]\big) > 0$, then ν has dense iterates.*

Proof. Let u_0 be a point in $(\mathrm{supp}\ \nu) \cap (0,3]$, let $[a,b]$ be an interval where ν is dense, and let $\mathcal{I}' = \big[f_a(1-1/u_0), f_b(1-1/u_0)\big]$. Pick any x in $(0,1)$, $y \in \mathcal{I}'$, and $\epsilon > 0$. The function f_{u_0} has an attractive fixed point at $1 - 1/u_0$, so there is some n such that $\big|f_{u_0}^{n-1}(x) - (1-1/u_0)\big| < \epsilon/8$. Since the support of ν is dense in $[a,b]$, and f_u is continuous in the parameter u, there is some $u_n \in \mathrm{supp}\ \nu$ such that $|f_{u_n}(1-1/u_0) - y| \leq \epsilon/2$. Taking $u_i = u_0$ for $i \leq n-1$,

$$\begin{aligned} \big|f_{u_n} \circ \cdots \circ f_{u_1}(x) - y\big| &\leq \big|f_{u_n}(1-1/u_0) - y\big| + \big|f_{u_n} \circ \cdots \circ f_{u_1}(x) - f_{u_n}(1-1/u_0)\big| \\ &\leq \frac{\epsilon}{2} + 4\frac{\epsilon}{8}. \end{aligned}$$

□

The significance of this property derives from this further fact:

Proposition 6. *If ν has dense iterates and $\nu\big((0,3]\big) > 0$, the Markov chain $\widetilde{F}_n$ is ψ-irreducible and aperiodic, with the support of ψ having nonempty interior. Consequently, all compact sets are petite.*

Proof. Let $\mathcal{I}$ be an interval where the iterates are dense, and we take the irreducibility measure ϕ to be Lebesgue measure on $\mathcal{I}$. Then we need to show that for any $x \in (0,1)$, $y \in \mathcal{I}$, and $\epsilon > 0$, the set of n such that $\mathrm{P}\big\{\big|\widetilde{F}_n(x) - y\big| \leq \epsilon\big\} > 0$ is nonempty, and has greatest common divisor 1. For $u \in (\mathrm{supp}\ \nu) \cap (0,3]$ the function f_u has an attractive fixed point at $1 - 1/u$, so for all n' sufficiently large $\mathrm{P}\big\{\big|\widetilde{F}_{n'}(x) - (1-1/u)\big| \leq \epsilon/8\big\} > 0$. Since the iterates are dense, we may find n'' such that

$$\mathrm{P}\big\{\big|\widetilde{F}_{n''}(1 - \frac{1}{u}) - y\big| \leq \frac{\epsilon}{2}\big\} > 0.$$

In both cases we are using the fact that the functions f_u are continuous in the parameter u. Putting these together, along with the trivial bound $|f_u(a)-f_u(b)| \le 4|a-b|$, we get

$$\mathrm{P}\big\{\big|\widetilde{F}_{n'+n''}(x)-y\big| \le \epsilon\big|\big\} > 0.$$

Since n' could be any number sufficiently large, the periodicity is 1. □

These conditions guarantee that an iterated logistic function system converges to a stationary measure.

Lemma 7. *If $\int \log u\,\nu(du) > 0$ and $\int \log(4-u)\nu(du) < \infty$, and if the Markov chain $\widetilde{F}_n(x)$ is ψ-irreducible and aperiodic with the support of ψ having nonempty interior, then the Markov chain has a unique stationary probability π, and the chain converges in probability to π.*

Proof. This proof merely generalizes the one given for Letac and Chamayou's example in [Ste99]. We consider the Markov chain $X_n = \log \widetilde{F}_n(x)$. Theorem 9.2.2 of Meyn and Tweedie [MT93] tells us that the chain is Harris recurrent, implying existence of a unique stationary distribution, if there is a compact subset $A \subset (0,1)$ to which the chain returns infinitely often, with probability 1. By their Theorem 10.0.1 the chain is positive Harris recurrent (that is, the stationary distribution is finite) if in addition $\sup_{x\in A} \mathrm{E}_x\,\tau_A < \infty$, where $\tau_A = \min\{n \ge 1 : \widetilde{F}_n(x) \in A\}$, and A has positive irreducibility measure. Let $A = [x_0, 1-x_0]$, where x_0 is chosen small enough that $\delta := \mathrm{E}\big[\log u(1-x_0)\big] > 0$. For $x \in (0, x_0)$ then $\mathrm{E}\big[X_{n+1}-X_n \,\big|\, X_n = \log x\big] \ge \delta$, and $Y_n := X_{n\wedge\tau_A} - \delta(n \wedge \tau_a)$ is a bounded submartingale. By the optional stopping theorem (Theorem II-2-13 of [Nev75]), this means that for any positive n,

$$\log x \le \mathrm{E}\big[Y_{\tau_A\wedge n} \,\big|\, Y_0 = \log x\big] \le -\delta\,\mathrm{E}\big[\tau_A \wedge n \,\big|\, Y_0 = \log x\big].$$

By the monotone convergence theorem $\mathrm{E}\big[\tau_A \,\big|\, Y_0 = \log x\big] \le -(\log x)/\delta$. If $x > 1-x_0$, then $\mathrm{E}\big[\tau_A \,\big|\, Y_0 = \log x\big] \le -\log(1-x)/\delta$. Consequently

$$\begin{aligned}
\mathrm{E}_x\,\tau_A &= 1 + \mathrm{E}_x(\tau_A - 1)\mathbf{1}\{f_1(x) \notin A\} \\
&\le 1 + \delta^{-1}\Big(-\mathrm{E}\log u_1 - \log(x_0 - x_0^2) - \mathrm{E}\log\big(1-\frac{u_1}{4}\big)\Big),
\end{aligned}$$

which is finite. The convergence in distribution follows then from Theorem 13.0.1 of Meyn and Tweedie. □

Note that we have excluded distributions which put a positive probability on 0, by restricting the domain of the functions to the open interval $(0,1)$. This makes no significant difference, but it is a technically convenient definition, since it allows the Markov chain to be irreducible; otherwise, the point 0 is an absorbing set off on its own. Of course, if $\int \log u\,\nu(du) \le 0$, the iterates converge almost surely to 0, so there is a unique stationary distribution concentrated at $\{0\}$. Athreya and Dai show in [AD] that a stationary probability always exists when $\int \log u\,\nu(du) > 0$ and $\int \log(4-u)\nu(du) < \infty$. But uniqueness, and convergence in distribution, still require Harris recurrence.

2.3. A few words about the conditions and the strategy.

When ν is concentrated at a single point, the relationship between Lyapunov exponent and long-term behavior of the iterates is far more complicated than our simple-minded theorems would admit. (For more details, see section V.4 of [dMvS93].) The case of measures supported on two points was itself already worth a paper by R. Bhattacharya and

B. Rao [BR93]. Fortunately, as is often the case, adding more randomness smooths out and simplifies the problem. The conditions "logarithmically continuous" and ψ-irreducible guarantee the necessary quantum of randomness for Theorems 1 and 2 respectively. They are clearly stronger than necessary, but they seem appropriate to the methods that we are applying. Logarithmically continuous rules out atoms between 2 and 4, and goes a bit further in requiring smoothness in the distribution.

For Theorem 2 we need to assume that ν places nonzero mass on the subinterval $(0, 3]$. This may seem unduly restrictive; but in fact, some such condition is required. These are the values of u for which the deterministic iteration has an attractive fixed point. If this interval has nonzero mass, then there is a positive probability of randomly picking a long run of functions with nearly the same fixed point. This tells us that eventually there will be some kind of contraction, if we wait long enough. This clearly need not be the case if ν is supported away from this region. For instance, suppose ν were uniform on the interval $[3.05, 3.051]$. All u in this interval give rise to maps with stable points of period 2. In the long run the random iterates become flat, reflecting a negative Lyapunov exponent, but do not converge to a constant function. Rather, the iterates converge to a (slightly) random step function with two steps. This is merely to say that much of the intricate range of behavior available to iterated logistic maps is maintained in the random case, even when we move beyond the trivial case of δ measures. What is perhaps surprising is that even a small overlap with the stable-fixed-point region $(0, 3]$, and sufficient randomness to make the Markov chain ψ-irreducible (with ψ adequately spread out), suffice to drive these systems into the very simple behavior of uniform convergence to a random fixed point. We get ψ-irreducibility from Lemma 5 and Lemma 6, under the assumption that ν is dense on an interval.

Theorem 1 relies on the tastelessly high-level condition of ψ-irreducibility itself, to avoid assuming that $\nu\big((0, 3]\big) > 0$. There must be a more aesthetic way around this problem, but I have not yet found it. There seemed little disadvantage, on the other hand, in using the more easily checked conditions which imply ψ-irreducibility in Theorem 2, since $\nu\big((0, 3]\big) > 0$ is required there for other reasons.

These conditions are not imposed in the paper of Klünger; that work contents itself as well with conditions for $\int \log u\nu(du)$ and $\int \log(4-u)\nu(du)$ instead of our stronger versions, which involve $(4u-u^2)^{-\alpha'}$ for some positive α'. It is worth taking a moment to reflect on where these assumptions enter the proof of Theorem 2.

Why is the result not simply trivial? After all,

$$\big|F_{n+1}(x) - F_n(x)\big| = \big|F_n\big(f_{n+1}(x)\big) - F_n(x)\big|,$$

so that attractivity depends fundamentally on the range of F_n (restricted to a compact interval) contracting sufficiently quickly to a point. This will follow if the derivative at every point converges exponentially to 0. The Markov chain $\widetilde{F}_n(x)$ is supposed to converge in distribution to π. It follows by the chain rule that the derivative at a fixed point should satisfy

$$\begin{aligned} n^{-1}\log\big|\widetilde{F}_n'(x)\big| &= n^{-1}\sum_{i=1}^{n}\log\big|f_i'\big(\widetilde{F}_{i-1}(x)\big)\big| \\ &= n^{-1}\sum_{i=1}^{n}\log u_i + n^{-1}\sum_{i=0}^{n-1}\log\big|2\widetilde{F}_i(x) - 1\big| \xrightarrow{n\to\infty} \lambda_\nu, \end{aligned}$$

as long as the Markov chain is ergodic. (To be sure, $\log|1-2x|$ is not a bounded function, but this is only a symptom of a larger problem.) Pointwise, the derivative of the n-th iterate should be growing exponentially when the Lyapunov exponent is positive, and shrinking exponentially when the Lyapunov exponent is negative. In the positive case the usual arguments which settle the question for affine maps, as in [AC92], must be augmented to allow for the noninjectivity: Even when the derivative is blowing up locally at every point, the function could in principle just happen to fold over to stay within an ever-shrinking span. On the other hand, this folding should, if anything, only make the negative case easier.

What we need, though, is *uniform* exponential shrinking of the derivatives. Pointwise exponential shrinking is useless without information about the size of the exceptional sets where the derivative gets very large. We cannot infer anything if, say, $|\widetilde{F}_n'(x)|^{1/n}$ converges always to a number $r<1$, but there is a set of x with measure about $e^{-n/2}$ where the derivative is as large as e^n. The logarithms of the derivatives are being added along a random Markov path, and each point corresponds to a separate path. To clarify this, it will help to place the problem in a more general context. Consider a general $\mathbb{R}^d$-valued iterated function system, with $f_i \in \mathcal{C} := \mathcal{C}^1(\mathcal{X},\mathcal{X})$, where $\mathcal{X}\subset\mathbb{R}^d$. Define a Markov chain with state space $\mathcal{C}\times\mathcal{X}$, defined by $Y_n := \big(f_n, \widetilde{F}_{n-1}(x)\big)$, where x is a fixed starting point. Then $\log D_x\widetilde{F}_n \le \sum_{i=1}^n g(Y_i)$, where $g(f,x) := D_x f$, and $D_x f$ is the local Lipschitz constant of f at x. As we explained in [Ste99], the iterated function system is attractive if for fixed paths $\gamma:(0,1)\to(0,1)$,

$$\sum_{n=1}^{\infty}\int D_{\gamma(t)}F_n\,dt$$

is finite almost surely. Ignoring for a moment the switch from F_n to $\widetilde{F}_n$, which does raise nontrivial problems, we expect that the integral will fall off exponentially with n, so satisfying the condition for attractivity, if

$$\limsup_{n\to\infty}\sup_s\left\{\frac{1}{n}\log\mathrm{Leb}\big\{x\in\gamma\ :\ \log D_x\widetilde{F}_n \ge sn\big\}+s\right\}<0. \tag{17}$$

Let $I(s)$ be the large-deviation rate function for the partial sums of g along Y_n:

$$\lim_{n\to\infty}\frac{1}{n}\log\mathrm{P}\big\{\sum_{i=1}^n g(Y_i)\ge sn\big\}=\lim_{n\to\infty}\frac{1}{n}\log\mathrm{P}\big\{D_x\widetilde{F}_n\ge e^{sn}\big\}=I(s).$$

In the worst case, these exceptionally bad (from the point of view of attractivity) points would show up in every realization in proportional strength; that is, the Lebesgue measure of the set of points where the derivative is at least e^{sn} is always about the same as the probability for any individual point. The condition for attractivity (17) then becomes

$$\sup_s I(s)+s<0.$$

By the general theory of large deviations for Markov chains, (see [Din93] for I. Dinwoodie's generalization of a theorem of S. Varadhan [Var84]) this is equivalent to the condition that

$$\mathrm{E}\big[D_xF_n\big]=\mathrm{E}_x\exp\big\{\sum_{i=1}^n g(Y_i)\big\}\big]\le\phi(x)r^n \tag{18}$$

for some $r < 1$ and $\phi : \mathcal{X} \to [1, \infty)$ which is bounded on compact subsets of $\mathcal{X}$.

This has been a tenuous chain of speculation, but at the end of it we arrive on solid ground: The condition (18) is the one that we called "locally contractive" in the paper [Ste99], and we showed there that, under mild conditions (which would always be satisfied when $\mathcal{X}$ is bounded), it implies that the system is attractive. It has the advantage of being easily checked in many cases, by means of a drift criterion; we repeat this criterion, in an improved form, at the end of this section. We used this criterion to show that the Chamayou-Letac logistic system is attractive for $a \geq 2$.

The reason for rederiving local contractivity here is to show why, for all its benefits, it imposes too strong a condition to be appropriate for random logistic maps. Exceptional behavior of sample paths of these maps will tend not to be isolated. Each iteration involves at most one folding; otherwise, nothing but monotonic mapping. We would expect the points with exceptionally large derivatives to arise *en masse* in some realizations, and in others not at all. That is, the exponentially small probability that the derivative at x is very large should be a result of exceptional realizations of the system, not of x being an exceptional point in an otherwise typical realization. Local contractivity ignores the coherence of these unimodal maps.

Our approach will be to ignore the derivatives at individual points, and instead to follow the development of the endpoints of the image of an interval $[x_0, 1 - x_0]$, where x_0 is any number between 0 and $\frac{1}{2}$. We divide up the behavior of the forward iteration $\widetilde{F}_n\big([x_0, 1 - x_0]\big)$ into three stages: Stage I, when x_0 is very small; Stage II, when x_0 has reached an intermediate value, but the image is still not very small; and stage III, when the image of $[x_0, 1 - x_0]$ has been squeezed into an interval of size no more than a given ϵ_0. The idea is that the time of commencing stage III has geometric tails, and once stage III has been reached, there is a nonzero chance that the diameter of the image will fall exponentially without ever returning above ϵ_0. An interval of size below ϵ_0 is close enough to being a point that its size will tend to shrink exponentially, on average, just like the derivative at a point, with rate close to the Lyapunov exponent. At any step, the image of the interval is expanded by a factor whose logarithm is no more than

$$\log\big|1 - 2\widetilde{F}_n(x) + \epsilon_0\big| - \log\big(1 - \widetilde{F}_n(x)\big),$$

where x stays constant during stage III. Note that this averages out to the Lyapunov exponent (except for the disturbing term ϵ_0), since

$$\begin{aligned}\int \log(1-x)\pi(dx) &= \int\int \log\big[ux(1-x)\big]\,\pi(dx)\nu(du) - \int \log x\,\pi(dx) - \int \log u\nu(du)\\ &= -\int \log u\,\nu(du),\end{aligned}$$

by the invariance of π.

It is not enough to check that the process eventually enters stage III and remains there. At the end, we will need to convert the result about $\widetilde{F}_n$ to one about F_n; for this purpose we need reasonable tail bounds for the time when this last entry into stage III occurs.

It is only here, in stage III, that we need the stronger conditions. We are summing a function along a path of the Markov chain, and trying to estimate the probability that it runs off to infinity without ever dropping below a certain value. We know that the long-term average should be close to the integral with respect to the stationary distribution — the Lyapunov exponent — but we need sufficient mixing conditions to tell us that the short-term averages of $\log|1-2X_n+\epsilon_0|-\log(1-X_n)$ will reach the stationary value quickly enough. This is technically arduous because the function $\log|1-2x|-\log(1-x)$ is unbounded, and because the chain is not uniformly ergodic. We note here that this problem simply does not arise when ν is restricted to $(0,3)$. For any choice of x and $u\in(0,3)$,

$$\frac{\bigl|1-2x\bigr|}{1-x}\frac{\bigl|1-2ux(1-x)\bigr|}{1-ux(1-x)}<1,$$

so any two steps in a row automatically give the desired contraction, regardless of any mixing properties.

We conclude with an improved version — necessary and sufficient, whereas the previous version was merely sufficient — of our earlier criterion for local contractivity:

Proposition 8. *An iterated function system is locally contractive if and only if there exists a drift function $\phi_*:\mathcal{X}\to[1,\infty)$ which is bounded on bounded subsets of $\mathcal{X}$, and some $r_*<1$ such that for all $x\in\mathcal{X}$,*

$$\mathrm{E}\bigl[\phi_*\bigl(f(x)\bigr)D_xf\bigr]\le r_*\phi_*(x). \tag{19}$$

Proof. Assume first that the system is locally contractive. Then there is a function $\phi:\mathcal{X}\to[1,\infty)$ and $r<1$ satisfying $\mathrm{E}\bigl[D_xF_n\bigr]\le r^n\phi(x)$. Let $G_n(x):=\mathrm{E}\bigl[D_xF_n\bigr]$, let $r_*:=(1+r)/2$, and define

$$\phi_*(x):=\sum_{n=0}^{\infty}r_*^{-n}G_n(x).$$

We have

$$\mathrm{E}\bigl[G_n\bigl(f(x)\bigr)D_xf\bigr]=\mathrm{E}\Bigl[\mathrm{E}\bigl[D_{f_{n+1}(x)}F_n\cdot D_xf_{n+1}\bigm|\mathcal{F}_n\bigr]\Bigr]\le G_{n+1}(x).$$

Thus

$$\mathrm{E}\bigl[\phi_*\bigl(f(x)\bigr)D_xf\bigr]\le\sum_{n=0}^{\infty}r_*^{-n}G_{n+1}(x)\le r_*\phi_*(x).$$

Now assume that (19) holds. Redefine

$$G_n(x):=\mathrm{E}\bigl[\phi_*\bigl(\widetilde{F}_n(x)\bigr)D_x\widetilde{F}_n\bigr].$$

Then, applying (19),

$$\begin{aligned}G_n(x)&\le\mathrm{E}\Bigl[\phi_*\Bigl(f_n\bigl(\widetilde{F}_{n-1}(x)\bigr)\Bigr)D_{\widetilde{F}_{n-1}(x)}f_n\cdot D_x\widetilde{F}_{n-1}\Bigr]\\&\le\mathrm{E}\bigl[r_*\phi_*\bigl(\widetilde{F}_{n-1}(x)\bigr)D_x\widetilde{F}_{n-1}\bigr]\\&=r_*G_{n-1}(x),\end{aligned}$$

So finally, since $\phi_*\ge1$,

$$\mathrm{E}\bigl[D_xF_n\bigr]=\mathrm{E}\bigl[D_x\widetilde{F}_n\bigr]\le G_n(x)\le r_*^nG_0(x)=r_*^n\phi_*(x).$$

3. Proof of Theorem 1

If the system were attractive, then the image of an interval would contract, until eventually it started to behave like a single point. But single points expand locally, in the long run, since the Lyapunov exponent is positive. It is this intuition that underlies the proof.

Let x be any point in $(0,1)$, and $y \in (0,1)$ a point distinct from x and from $1-x$. Define

$$Y_n := \bigl|\widetilde{F}_n(x) - \widetilde{F}_n(y)\bigr|, \quad \text{and} \quad X_n := \log Y_n.$$

Since y is neither x nor $1-x$, and $\mathrm{E}\log u_1$ is finite, $\mathrm{E}\,X_1$ is also finite. For each n,

$$\begin{aligned}
X_{n+1} &= \log\bigl|f_{n+1}\bigl(\widetilde{F}_n(x)\bigr) - f_{n+1}\bigl(\widetilde{F}_n(y)\bigr)\bigr| \\
&= \log\Bigl[u_{n+1}\bigl|\widetilde{F}_n(x) - \widetilde{F}_n(y)\bigr| \cdot \bigl|1 - \widetilde{F}_n(x) - \widetilde{F}_n(y)\bigr|\Bigr], \text{ and}
\end{aligned}$$

$$\begin{aligned}
\mathrm{E}\bigl[X_{n+1} - X_n \bigm| \mathcal{F}_{n-1}\bigr] &= \mathrm{E}\Bigl[\log u_{n+1}\bigl|1 - 2f_n\bigl(\widetilde{F}_{n-1}(x)\bigr)\bigr| \Bigm| \mathcal{F}_{n-1}\Bigr] \\
&\quad + \mathrm{E}\left[\log \frac{\bigl|1 - f_n\bigl(\widetilde{F}_{n-1}(x)\bigr) - f_n\bigl(\widetilde{F}_{n-1}(y)\bigr)\bigr|}{\bigl|1 - 2f_n\bigl(\widetilde{F}_{n-1}(x)\bigr)\bigr|} \Bigm| \mathcal{F}_{n-1}\right] \\
&= \mathrm{E}\bigl[\lambda_\nu\bigl(\widetilde{F}_n(x)\bigr) \bigm| \mathcal{F}_{n-1}\bigr] + \Lambda_\nu\Bigl(\widetilde{F}_{n-1}(x) - \widetilde{F}_{n-1}(x)^2 + \widetilde{F}_{n-1}(y) - \widetilde{F}_{n-1}(y)^2\Bigr) \\
&\quad - \Lambda_\nu\Bigl(2\bigl(\widetilde{F}_{n-1}(x) - \widetilde{F}_{n-1}(x)^2\bigr)\Bigr) \\
&\geq \mathrm{E}\bigl[\lambda_\nu\bigl(\widetilde{F}_n(x)\bigr) \bigm| \mathcal{F}_{n-1}\bigr] - \rho\bigl(Y_{n-1}\bigr),
\end{aligned}$$

where ρ is the modulus of continuity for Λ_ν. Note that Λ_ν is continuous on a compact interval, so $\rho(0) = 0$; in addition, ρ is continuous, nondecreasing, sublinear, and bounded by 1 (cf. page 101 of [Tim66]). A function with these properties has a concave majorant ρ_*, defined as the infimum of all concave functions which are $\geq \rho$, which is concave, continuous, and such that $\rho_*(0) = 0$. Thus

$$\begin{aligned}
-\mathrm{E}\,X_1 \geq \mathrm{E}\bigl[X_n - X_1\bigr] &= \sum_{i=2}^{n} \mathrm{E}\bigl[X_i - X_{i-1}\bigr] \\
&\geq \sum_{i=2}^{n} \mathrm{E}\,\lambda_\nu\bigl(\widetilde{F}_{i-1}(x)\bigr) - \sum_{i=1}^{n} \mathrm{E}\,\rho\bigl(Y_{i-2}\bigr).
\end{aligned}$$

By assumption, $\widetilde{F}_n(x)$ converges in distribution to μ, which implies that

$$\mathrm{E}\,\lambda_\nu\bigl(\widetilde{F}_i(x)\bigr) \xrightarrow{n\to\infty} \int_0^1 \lambda_\nu(z)\pi(dz) = \lambda_\nu.$$

Since $\mathrm{E}\,X_1$ is finite, this means that

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\,\rho_*\bigl(Y_i\bigr) \geq \liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\,\rho\bigl(Y_i\bigr) \geq \lambda_\nu.$$

By an application of Jensen's inequality,

$$\liminf_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} \mathrm{E}\,Y_i \geq \rho_*^{-1}(\lambda_\nu) > 0,$$

where ρ_* is the concave majorant of ρ on $[0,1]$. Thus y is in B_x. But this is true for every y which is neither x nor $1-x$, so B_x contains all but these two points. If

$F_n(x)$ converges then $\pi_x\{x, 1-x\}$ must be 1, by Lemma 4. This is impossible if the chain is ψ-irreducible.

4. Proof of Theorem 2

Suppose first that $\int \log u\, \nu(du) \le 0$. Since ν is not a delta distribution at 1,

$$\liminf_{n\to\infty} \log \mathrm{Lip}\ F_n \le \liminf_{n\to\infty} \log\ u_1 + \cdots + \log\ u_n = -\infty.$$

The sequence $\sup_x F_n(x)$ is nonincreasing in n, and is bounded above by $\frac{1}{2}\mathrm{Lip}\ F_n$, so $\lim_{n\to\infty}\sup_x F_n(x) = 0$. This means that the system is attractive, with the trivial limit 0.

Suppose now that $\int \log u\, \nu(du) > 0$. For any $\alpha \in (0, \alpha')$, an application of the elementary inequality $e^x - 1 - x \le \frac{x^2}{2}(e^x + e^{-x})$ with $x = -\alpha \log u$ shows that

$$\int u^{-\alpha}\nu(du) \le 1 + \alpha \int \log u\, \nu(du) + \frac{\alpha^2}{2}\int \left(\log^2 u\right)\left(u^\alpha + u^{-\alpha}\right)\nu(du).$$

Thus, for α sufficiently small but positive,

$$\int u^{-\alpha}\nu(du) > 1, \tag{20}$$

and of course (4) still holds with α in place of α'. (This simple computation was suggested by a similar one in [Wu00].)

It will be convenient to be able to treat the logistic functions as monotone, by folding all the points back onto the interval $(0, \frac{1}{2}]$. To this end we use the tent map

$$\phi(x) = \min\{x, 1-x\}.$$

Let $X_0 \in (0, \frac{1}{2})$ be chosen, and let $Z_0 = \frac{1}{2}$. We define recursively

$$\begin{aligned} X_{n+1} &= \min_{x\in[X_n, Z_n]} \phi\big(f_{n+1}(x)\big) \\ Z_{n+1} &= \max_{x\in[X_n, Z_n]} \phi\big(f_{n+1}(x)\big). \end{aligned}$$

Thus $\phi\big(\widetilde{F}_n([X_0, 1-X_0])\big) = [X_n, Z_n]$ for $n \ge 1$. We want to define X_n^* and Z_n^* with simpler dynamics such that $[X_n, Z_n] \subset [X_n^*, Z_n^*] \subset (0, \frac{1}{2}]$. There will also be $\epsilon_n^* \ge \sqrt{Z_n^*/X_n^*} - 1$. These definitions imply that

$$\sup_{X_0 \le x \le y \le 1-X_0} \left|\widetilde{F}_n(x) - \widetilde{F}_n(y)\right| \le 2\epsilon_n^* \sqrt{X_n^*}\left(\sqrt{X_n^*} + \sqrt{Z_n^*}\right) \le 2\epsilon_n^*. \tag{21}$$

We claim that there are positive constants η' and c, independent of n and X_0 (but depending on the chain and the choice of ϵ_0 and x_0), such that for every positive integer p there is a positive constant B_p, with

$$\mathrm{P}\big\{\log(\epsilon_n^*/\epsilon_0) \ge -\eta' n\big\} \le B_p X_0^{-c} n^{-p}, \tag{22}$$

We show first that the theorem follows from this claim.

By (21), the claim tells us that for all $X_0 \in (0, \frac{1}{2})$ and positive integers n,

$$\mathrm{P}\Big\{\sup_{X_0 \le x \le y \le 1-X_0} \left|\widetilde{F}_n(x) - \widetilde{F}_n(y)\right| \ge 2\epsilon_0 e^{-\eta' n} \;\Big|\; \mathcal{F}_0\Big\} \le B_p X_0^{-c} n^{-p}.$$

Since F_n and $\widetilde{F}_n$ have the same distribution, the same inequality holds when $\widetilde{F}_n$ is replaced by F_n. The fact that the constants do not depend on X_0 or n, furthermore, allows us to take $X_0 = n^{-r}$, where $r > 1/\alpha$ — recall that α was specified to be a positive constant such that (20) and (4) hold — obtaining for $n \geq 2$,

$$\mathrm{P}\big\{ \sup_{n^{-r} \leq x \leq y \leq 1-n^{-r}} \big|F_n(x) - F_n(y)\big| \geq 2\epsilon_0 e^{-\eta' n} \big\} \leq B_p n^{-p+rc}. \tag{23}$$

Since p is arbitrary, we may take it to be larger than $rc+1$. Also,

$$\begin{aligned}\mathrm{P}\big\{f_n(x) \notin \big(n^{-r}, 1-n^{-r}\big)\big\} &= \mathrm{P}\big\{u_n \leq n^{-r}(x-x^2)^{-1} \quad \text{or} \quad u_n \geq \big(1-n^{-r}\big)(x-x^2)^{-1}\big\} \\ &\leq \eta_\alpha \left(\frac{4}{x-x^2}\right)^\alpha n^{-r\alpha}\end{aligned}$$

By the Borel-Cantelli Lemma, it follows that for all ϵ_0 sufficiently small and $x \in (0,1)$, since $F_{n+1}(x) = F_n(f_{n+1}(x))$,

$$\begin{aligned}&\mathrm{P}\big\{\exists \text{ infinitely many } n \text{ s.t. } \big|F_{n+1}(x) - F_n(x)\big| \geq \epsilon_0 e^{-\eta' n/4}\big\} \\ &\quad \leq \mathrm{P}\big\{\exists\, \infty \text{ many } n \text{ s.t. } \sup_{n^{-r} \leq y \leq 1-n^{-r}} \big|F_n(x) - F_n(y)\big| \geq \epsilon_0 \text{ or } f_{n+1}(x) \notin \big(n^{-r}, 1-n^{-r}\big)\big\},\end{aligned}$$

which is 0. Thus $\big(F_n(x)\big)$ is almost surely a Cauchy sequence, and $|F_n(x) - F_n(y)| \to 0$ for all $x, y \in (0,1)$. Thus $\lim_{n\to\infty} F_n(x)$ exists for every $x \in (0,1)$, and is independent of the choice of x.

We now need to prove the claim. Since $\int \log u\, \nu(du) > 0$ and $\int \log(4-u)\nu(du) < \infty$, by the Monotone Convergence Theorem we may find an $x_0 \in (0, \frac{1}{2})$ and $\eta > 0$ such that

$$\int \log\big[ux_0(1-x_0) \,\wedge\, (1-u/4)\big]\nu(du) \geq \log x_0 + \eta. \tag{24}$$

Also, since the Lyapunov exponent is negative,

$$\int \log u\, \nu(du) + \int \log|1-2x|\pi(dx) < 0.$$

Since π is the stationary distribution for the Markov chain, it must be that

$$\int\int \log ux(1-x)\nu(du)\pi(dx) = \int \log x\, \pi(dx),$$

so that

$$\int \log|1-2x|\pi(dx) - \int \log(1-x)\pi(dx) = \int \log u\, \nu(du) + \int \log|1-2x|\pi(dx) < 0.$$

This means that for all $\epsilon_0 > 0$ sufficiently small, by the Monotone Convergence Theorem,

$$-\eta' := \frac{1}{4}\log(1+\epsilon_0) \,+\, \int \log\left(\frac{|1-2x| + \epsilon_0}{(1-x)(1-2\epsilon_0)}\right)\pi(dx) \tag{25}$$

is negative. We fix such an $\epsilon_0 \in (0, \frac{1}{64})$.

We define a sequence of stopping times $0 = \tau_0 < \sigma_1 < \tau_1 < \sigma_2 < \cdots$, by the following rules:

$$\begin{aligned}\sigma_{i+1} &:= \min\big\{n > \tau_i \,:\, \epsilon_n^* \leq r_0\epsilon_0 \text{ and } X_n^* \geq x_0\big\}, \text{ and} \\ \tau_i &:= \min\big\{n > \sigma_i \,:\, \epsilon_n^* \geq \epsilon_0\big\}.\end{aligned}$$

Here r_0 is a parameter between 0 and 1, which will be specified presently. We split up the definition of X_n^* into three "stages": If $\tau_i \leq n < \sigma_{i+1}$ for some i and $X_n^* < x_0$, we say that the process is in Stage I at time n; it is in Stage II if $\tau_i \leq n < \sigma_{i+1}$ and $X_n^* \geq x_0$. When $\sigma_i \leq n < \tau_i$ the process is in Stage III. Intuitively, the process is in Stage I as long as the left endpoint of the interval is still stuck in the corner, close to 0. Once the interval has achieved sufficient separation from 0, Stage II commences, whereby we regard the width of the interval, waiting for it to shrink below a small fraction of ϵ_0. That achieved, Stage III continues, unless the interval swells up larger than a width of ϵ_0 again, at which time the process would revert to Stage I or II (depending on the location of the left endpoint).

In Stage I, we define $X^*_{n+1} = \left[u_{n+1}X^*_n(1-X^*_n)\right] \wedge (1-u_{n+1}/4)$ and $Z^*_{n+1}=\frac{1}{2}$. In Stage II it is

$$
\begin{aligned}
X^*_{n+1} &= \min\left\{\phi\left(u_{n+1}X^*_n(1-X^*_n)\right),\ \phi\left(u_{n+1}Z^*_n(1-Z^*_n)\right)\right\}, \text{ and}\\
Z^*_{n+1} &= \max\left\{\phi\left(u_{n+1}X^*_n(1-X^*_n)\right),\ \phi\left(u_{n+1}Z^*_n(1-Z^*_n)\right)\right\}.
\end{aligned}
$$

In both of these stages $\epsilon^*_n := \sqrt{X^*_n/Y^*_n}$. Finally, in Stage III, we use an auxiliary process $Y^{(i)}_n$ to generate the others. the idea is that $Y^{(i)}_n$ is a point in the middle of the short interval, acting as a surrogate for the whole, while ϵ^* is the extension on either side around $Y^{(i)}_n$:

$$
\begin{aligned}
Y^{(i)}_{\sigma_i} &= \sqrt{X^*_{\sigma_i}Z^*_{\sigma_i}}\,,\\
\epsilon^*_{\sigma_i} &= \sqrt{Z^*_{\sigma_i}/X^*_{\sigma_i}}-1,\\
\text{and for all } n\ge\sigma_i,\quad Y^{(i)}_{n+1} &= u_{n+1}Y^{(i)}_n(1-Y^{(i)}_n).
\end{aligned}
$$

Then for $\sigma_i < n \le \tau_i$ we define $Y^*_n = \phi(Y^{(i)}_n)$ and

$$
\begin{aligned}
\epsilon^*_{n+1} &= \epsilon^*_n\cdot\begin{cases}\frac{1+\epsilon^*_n-2Y^*_n}{(1-Y^*_n)(1-\epsilon^*_n)} & \text{if } Y^{(i)}_{n+1}\le\frac{1}{2}\,;\\ \frac{u_{n+1}Y^*_n\left(1+\epsilon^*_n-2Y^*_n\right)}{1-u_{n+1}Z^*_n(1-Z^*_n)} & \text{if } Y^{(i)}_{n+1}>\frac{1}{2}\,;\end{cases}\\
X^*_{n+1} &= \frac{Y^*_{n+1}}{1+\epsilon^*_{n+1}};\text{ and}\\
Z^*_{n+1} &= \min\left\{\frac{1}{2}, Y^*_{n+1}(1+\epsilon^*_{n+1})\right\}.
\end{aligned}
$$

We show in Lemma 11 that these definitions do indeed imply that

$$
\phi\Big(f_{n+1}([X^*_n,Z^*_n])\Big)\subset[X^*_{n+1},Z^*_{n+1}]. \tag{26}
$$

The differences $\sigma_i-\tau_{i-1}$ have geometric tails, while the probability of $\tau_i-\sigma_i$ being finite but larger than some n falls off faster than any power of n. These and other useful properties are given in Lemma 9. That lemma provides us with a value of r_0 which guarantees that $\mathrm{P}\left\{\tau_i=\infty \,\middle|\, \mathcal{F}_{\sigma_i}\right\}$ is bounded away from 0.

Let $I:=\min\{i : \tau_i=\infty\}$. By the tail bound (34), I is almost surely finite, with

$$
\mathrm{P}\left\{I\ge i\,\middle|\,\mathcal{F}_0\right\}\le(1-c_4)^{i-1}\text{ almost surely.}
$$

By (31) and (32), for all positive integers i, p, and n, on the event $\{\tau_{i-1}<\infty\}$,

$$
\begin{aligned}
\mathrm{P}\left\{\infty>\tau_i-\tau_{i-1}\ge n\,\middle|\,\mathcal{F}_{\tau_{i-1}}\right\} &\le \mathrm{P}\left\{\sigma_i-\tau_{i-1}\ge\frac{n}{2}\,\middle|\,\mathcal{F}_{\tau_{i-1}}\right\}+\mathrm{P}\left\{\infty>\tau_i-\sigma_i\ge n\,\middle|\,\mathcal{F}_{\tau_{i-1}}\right\}\\
&\le c_1e^{-c_2n/2}(X^*_{\tau_{i-1}})^{-c_5}+2^pb_pn^{-p},
\end{aligned}
$$

where the b's and c's are positive constants. By (33), we may find positive constants b'_p such that

$$\mathrm{P}\big\{\infty > \tau_1 \geq n \,\big|\, \mathcal{F}_0\big\} \leq b'_p X_0^{-c_5} n^{-p} \text{ for all } n \text{ and } p. \tag{27}$$

For any $i \geq 2$, on the event $\{\tau_{i-1} < \infty\}$,

$$\begin{aligned} \mathrm{P}\big\{\infty > \tau_i - \tau_{i-1} \geq n \,\big|\, \mathcal{F}_0\big\} &\leq \mathrm{E}\Big[\mathrm{E}\big[c_1 e^{-c_2 n/2}(X^*_{\tau_{i-1}})^{-c_5} + 2^p b_p n^{-p} \,\big|\, \mathcal{F}_{\sigma_{i-1}}\big] \,\Big|\, \mathcal{F}_0\Big] \\ &\leq c_1 c_3 e^{-c_2 n/2} + 2^p b_p n^{-p}, \end{aligned} \tag{28}$$

so we may choose the constants b'_p to satisfy

$$\mathrm{P}\big\{\infty > \tau_i - \tau_{i-1} \geq n \,\big|\, \mathcal{F}_0\big\} \leq b'_p n^{-p} \quad \text{for all } i \geq 2. \tag{29}$$

Then

$$\begin{aligned} \mathrm{P}\big\{\sigma_I \geq n \,\big|\, \mathcal{F}_0\big\} &\leq \sum_{i=1}^{n} \mathrm{P}\big\{I \geq i \,\big|\, \mathcal{F}_0\big\}\, \mathrm{P}\big\{\sigma_i \geq n \,\big|\, I \geq i\,,\, \mathcal{F}_0\big\} \\ &\leq \sum_{i=1}^{n} (1-c_4)^{i-1} \sum_{k=1}^{i} \mathrm{P}\big\{\infty > \tau_k - \tau_{k-1} \geq \frac{n}{i} \,\big|\, \mathcal{F}_0\big\} \\ &\leq \sum_{i=2}^{n} (1-c_4)^{i-1} b'_p i^p (i + X_0^{-c_5}) n^{-p} \\ &\leq B'_p X_0^{-c_5} n^{-p}, \end{aligned}$$

where B'_p and c_5 are positive constants depending on the choice of ϵ_0 and x_0, and on the Markov chain, but not on the starting point X_0.

Observe now that

$$\begin{aligned} &\mathrm{P}\big\{\log(\epsilon^*_n/\epsilon_0) \geq -\eta' n\big\} \\ &\qquad \leq \mathrm{P}\big\{\sigma_I \geq \frac{n}{2}\big\} + \sum_{i=1}^{n/2} \mathrm{P}\big\{\sigma_i \leq \frac{n}{2}, \quad \tau_i = \infty \text{ and } \log\big(\epsilon^*_n/\epsilon_0\big) \geq -\eta' n\big\} \\ &\qquad \leq 2^p B'_p X_0^{-c_5} n^{-p} + 2^p a_{p+1} n^{-p}, \end{aligned} \tag{30}$$

by (35). □

5. Technical lemmata

Lemma 9. *With the notation of the proof of Theorem 2, for ϵ_0 and x_0 sufficiently small there are positive constants c_1, c_2, c_3, c_4, c_5 and $b_1, b_2, b_3, \ldots$ such that for all natural numbers i, n, and p,*

$$\mathrm{P}\big\{\sigma_{i+1} - \tau_i \geq n \,\big|\, \mathcal{F}_{\tau_i}\big\} \leq c_1 e^{-c_2 n}\big(X^*_{\tau_i}\big)^{-c_5} \text{ on } \{\tau_i < \infty\}, \tag{31}$$

$$\mathrm{P}\big\{\infty > \tau_i - \sigma_i \geq n \,\big|\, \mathcal{F}_{\sigma_i}\big\} \leq b_p n^{-p} \text{ on } \{\tau_{i-1} < \infty\}, \tag{32}$$

$$\mathrm{E}\big[\big(X^*_{\tau_i}\big)^{-c_5} \mathbf{1}_{\{\tau_i < \infty\}} \,\big|\, \mathcal{F}_{\sigma_i}\big] \leq c_3 \text{ on } \{\sigma_i < \infty\}. \tag{33}$$

almost surely.

For r_0 sufficiently small (but still positive), there is a positive constant c_4 such that for all i, on the event $\{\sigma_i < \infty\}$,

$$\mathrm{P}\big\{\tau_i = \infty \,\big|\, \mathcal{F}_{\sigma_i}\big\} \geq c_4 \ a.s. \tag{34}$$

There are also positive constants $a_1, a_2, \dots$, such that for every i and n,

$$\mathrm{P}\bigl\{\tau_i - \sigma_i \geq n \text{ and } \log \epsilon^*_{\sigma_i+n} - \log \epsilon_0 \geq -2\frac{\eta'}{n} \bigm| \mathcal{F}_{\sigma_i}\bigr\} \leq a_p n^{-p} \text{ on } \{\sigma_i < \infty\}. \tag{35}$$

Here η' is the positive constant given in (25). The a's, b's, and c's are "constant" in that they depend only on the distribution ν, and on the choice of r_0, ϵ_0 and x_0, not on the starting point X_0.

Proof of (31). Let $\sigma = \sigma_{i+1}$ and $\tau = \tau_i$ for some given i. Let n_0 and q be given as in Lemma 10. (The condition $\nu\bigl((1,4)\bigr) > 0$ is guaranteed by our assumption $\mathrm{E}\log u > 0$.) On the event $\{\tau < \infty\}$ we define a new sequence of stopping times ρ_j which interpolate between τ and σ:

$$\begin{aligned}
\rho_0 &:= \tau - n_0,\\
\rho_1 &:= \min\bigl\{k \geq \tau \,:\, X^*_k \geq x_0\bigr\}, \text{ and}\\
\rho_j &:= \min\bigl\{k \geq \rho_{j-1} + n_0 \,:\, X^*_k \geq x_0\bigr\} \text{ for } j \geq 2.
\end{aligned}$$

We define $J = \min\{j \,:\, \sigma \leq \rho_j + n_0\}$. Suppose that we can find a $\gamma \in (0, 2\alpha/3]$ such that $\bigl(X^*_{(\rho_j+n_0+k)\wedge\rho_{j+1}}\bigr)^{-\gamma} e^{\gamma^2 k}$ is a supermartingale, and in addition

$$\Bigl(8^\gamma e^{\gamma^2}\eta_\gamma\Bigr) \leq (1-q)^{-1/2n_0}, \text{ and} \qquad x_0^{-\gamma} \leq (1-q)^{-1}. \tag{36}$$

By the optional stopping theorem for positive supermartingales (Theorem II-2-13 of [Nev75]), it follows that for each $j \geq 0$, on the event $\{J > j\}$,

$$\mathrm{E}\bigl[e^{\gamma^2(\rho_{j+1}-\rho_j-n_0)}\bigl(X^*_{\rho_{j+1}}\bigr)^{-\gamma} \bigm| \mathcal{F}_{\rho_j+n_0}\bigr] \leq \bigl(X^*_{\rho_j+n_0}\bigr)^{-\gamma}.$$

For any k, on the event $\rho_{j+1} - \rho_j > k$, using η_γ defined in (4),

$$\begin{aligned}
&\mathrm{E}\bigl[\bigl(X^*_{\rho_j+k+1}\bigr)^{-\gamma} \bigm| \mathcal{F}_{\rho_j+k}\bigr]\\
&\qquad \leq \mathrm{E}\bigl[u^{-\gamma}_{\rho_j+k+1}\bigl(X^*_{\rho_j+k}\bigr)^{-\gamma}\bigl(1 - X^*_{\rho_j+k}\bigr)^{-\gamma} \vee \bigl(1 - \frac{u_{\rho_j+k+1}}{4}\bigr)^{-\gamma} \bigm| \mathcal{F}_{\rho_j+k}\bigr]\\
&\qquad \leq 8^\gamma \eta_\gamma \bigl(X^*_{\rho_j+k}\bigr)^{-\gamma},
\end{aligned}$$

since $X^*_{\rho_j+k} \leq \frac{1}{2}$. (Note: We have used the stage I definition of $X^*_{\rho_j+k+1}$, but it serves as well as a lower bound in stage II.) This yields

$$\begin{aligned}
\mathrm{E}\bigl[e^{\gamma^2(\rho_{j+1}-\rho_j)}\bigl(X^*_{\rho_{j+1}}\bigr)^{-\gamma}\mathbf{1}_{\{\sigma\geq\rho_{j+1}\}} \bigm| \mathcal{F}_{\rho_j}\bigr] &\leq e^{\gamma^2 n_0}\,\mathrm{E}\bigl[\bigl(X^*_{\rho_j+n_0}\bigr)^{-\gamma} \bigm| \mathcal{F}_{\rho_j}\bigr]\mathbf{1}_{\{\sigma\geq\rho_j\}}\\
&\leq (1-q)^{-1/2}\bigl(X^*_{\rho_j}\bigr)^{-\gamma}\mathbf{1}_{\{\sigma\geq\rho_j\}}.
\end{aligned}$$

Iterating this conditioning, we see that for all positive j,

$$\mathrm{E}\bigl[e^{\gamma^2(\rho_j-\tau)}\mathbf{1}_{\{\sigma\geq\rho_j\}} \bigm| \mathcal{F}_\tau\bigr] \leq (1-q)^{-j/2}\bigl(X^*_\tau\bigr)^{-\gamma}.$$

By Lemma 10, we have always $\mathrm{P}\bigl\{\sigma \leq \rho_j + n_0 \bigm| \mathcal{F}_{\rho_j}\bigr\} \geq q$. (The definition of σ excludes the possibility that $\rho_j + n_0 < \sigma < \rho_{j+1}$, since for all k in that range $X^*_k < x_0$.) Thus for all positive j,

$$\mathrm{P}\bigl\{\sigma \geq \rho_j \bigm| \mathcal{F}_\tau\bigr\} \leq (1-q)^{j-1}.$$

We now see that J is almost surely finite, and

$$\begin{aligned}
\mathrm{E}\big[\exp\{\frac{\gamma^2}{2}(\sigma-\tau)\}\,\big|\,\mathcal{F}_\tau\big] &\le e^{\gamma^2 n_0/2}\,\mathrm{E}\left[\sum_{j=1}^{\infty}\mathbf{1}_{\{J=j\}}\exp\{\frac{\gamma^2}{2}(\rho_j-\tau)\}\,\big|\,\mathcal{F}_\tau\right]\\
&\le e^{\gamma^2 n_0/2}\sum_{j=1}^{\infty}\mathrm{P}\{J\ge j\,|\,\mathcal{F}_\tau\}^{1/2}\,\mathrm{E}\big[\exp\{\gamma^2(\rho_j-\tau)\}\,\big|\,\mathcal{F}_\tau\big]^{1/2}\\
&\le e^{\gamma^2 n_0/2}\big(X^*_\tau\big)^{-\gamma/2}(1-q)^{-1}\sum_{j=1}^{\infty}(1-q)^{j/4}.
\end{aligned}$$

This proves the claim (31) with $c_1=e^{\gamma^2 n_0/2}(1-q)^{-1}(1-\sqrt[4]{1-q})^{-1}$, $c_2=\gamma^2/2$, and $c_5=\gamma/2$, once we have shown that the necessary constant γ exists.

For all $1\le k\le \rho_j-\rho_{j-1}-n_0$, by (24), and the fact that $X^*_{\tau+k-1}<x_0$,

$$0<\eta\le \mathrm{E}\big[\log X^*_{\tau+k}-\log X^*_{\tau+k-1}\,\big|\,\mathcal{F}_{\tau+k-1}\big],\text{ and}$$

$$\begin{aligned}
&\mathrm{E}\Big[\exp\{\alpha\big|\log X^*_{\tau+k}-\log X^*_{\tau+k-1}\big|\}\,\big|\,\mathcal{F}_{\tau+k-1}\Big]\\
&\quad\le \mathrm{E}\Big[\big(X^*_{\tau+k}\big)^{\alpha}\,\big|\,\mathcal{F}_{\tau+k-1}\Big]\big(X^*_{\tau+k-1}\big)^{-\alpha}+\mathrm{E}\Big[\big(X^*_{\tau+k}\big)^{-\alpha}\,\big|\,\mathcal{F}_{\tau+k-1}\Big]\big(X^*_{\tau+k-1}\big)^{\alpha}\\
&\quad\le \mathrm{E}\big[u^{\alpha}_{\tau+k}\big]\;+\;\mathrm{E}\Big[\min\Big\{u_{\tau+k}(1-X^*_{\tau+k-1}),\,(1-u_{\tau+k}/4)\Big\}^{-\alpha}\Big]\\
&\quad\le 4^{\alpha}+(1-x_0)^{-\alpha}\,\mathrm{E}\big[u^{-\alpha}\big]+\mathrm{E}\big[(1-u/4)^{-\alpha}\big].
\end{aligned}$$

Under these conditions Lemma 2.6 of [SS02] provides a positive γ_0, depending only on the distribution ν, such that $\big(X^*_{(\rho_j+n_0+k)\wedge\rho_{j-1}}\big)^{-\gamma}e^{\gamma^2 k}$ is a supermartingale for all $\gamma\in[0,\gamma_0]$. The conditions (36) are satisfied for all γ sufficiently small. $\square$

Proof of (34), (32), *and* (35). We need to show first that $Y_n:=Y^{(i)}_{\sigma_i+n}$, stopped when $n=\tau_i-\sigma_i$, is a V-uniformly ergodic Markov chain, with $V(x)=(x-x^2)^{-\alpha}$. By Lemma 15.2.8 and Theorem 16.1.2 of [MT93], we need to show that the sets $\{V(x)\le a\}$ are petite sets and

$$PV\le\lambda V+L, \tag{37}$$

where P is the Markov operator and λ and L are positive constants, with $\lambda<1$. The first we have already shown in Proposition 6, since the sets $\{x\,:\,V(x)\le a\}$ are compact subsets of $(0,1)$. The second condition is

$$\int\Big(u(x-x^2)\big(1-u(x-x^2)\big)\Big)^{-\alpha}\nu(du)\le\lambda(x-x^2)^{-\alpha}+L,$$

which becomes

$$\int(u-yu^2)^{-\alpha}\nu(du)\le\lambda+Ly^{\alpha}$$

when we set $y=x-x^2$. Using the convexity of the function $y\mapsto(1-yu)^{-\alpha}$, and the fact that $y\le y^{\alpha}$, we see that

$$\begin{aligned}
\int(u-yu^2)^{-\alpha}\nu(du)&\le\int u^{-\alpha}\big(1-4y+4y(1-u/4)^{-\alpha}\big)\nu(du)\\
&\le(1-4y)\int u^{-\alpha}\nu(du)+4^{1+\alpha}\eta_\alpha y^{\alpha},
\end{aligned}$$

so that (37) holds with $\lambda=\int u^{-\alpha}\nu(du)$ and $L=4^{1+\alpha}\eta_\alpha$.

Theorem 1 from [Ste01] implies that if $g:(0,1)\to\mathbb{R}$ with $|g|\le c\log V$ for some constant c, then for every positive integer p there is a constant c_p^*, depending on the distribution ν and the function g, such that the normalized partial sums along the Markov chain Y_{σ_i+n},

$$S_k(g):=\sum_{n=0}^{k-1}\Bigl(g(Y_{\sigma_i+n})\;-\;\pi(g)\Bigr),$$

satisfy for every positive integer k,

$$\mathrm{E}\bigl[\bigl|S_k(g)\bigr|^p\;\big|\;Y_{\sigma_i}=y\bigr]\le c_p^*k^{p/2}V(y). \tag{38}$$

We apply this to the function

$$g(y)=\log\frac{|1-2y|+\epsilon_0}{(1-y)(1-2\epsilon_0)}. \tag{39}$$

The defining characteristic of ϵ_0 is (25), which tells us that $\log(1+\epsilon_0)+\pi(g)=:-\eta'<0$.

Observe now that for $\sigma_i\le n<\tau_i$, since $\epsilon_n^*<\epsilon_0$, and using the fact that $Y_n^{(i)}=1-Y_n^*$ when $Y_n^{(i)}>\frac12$,

$$\begin{aligned}\log\epsilon_{n+1}^*-\log\epsilon_n^*&\le g\bigl(Y_n^*\bigr)+\mathbf{1}\{Y_{n+1}^{(i)}>\frac12\}\log\frac{u_{n+1}^{(i)}Y_n^{(i)}(1-Y_n^{(i)})}{1-u_{n+1}Z_n^*(1-Z_n^*)}\\&=g(Y_n^{(i)})+\mathbf{1}\{Y_n^{(i)}>\frac12\}\log\frac{1-Y_n^{(i)}}{Y_n^{(i)}}\\&\qquad+\mathbf{1}\{Y_{n+1}^{(i)}>\frac12\}\log\frac{Y_{n+1}^{(i)}}{1-u_{n+1}Z_n^*(1-Z_n^*)}\end{aligned}$$

Consequently, for $0<k\le\tau_i-\sigma_i$,

$$\begin{aligned}&\log\epsilon_{\sigma_i+k}^*-\log\epsilon_{\sigma_i}^*\\&\quad\le\sum_{n=\sigma_i}^{\sigma_i+k-1}\left[g(Y_n^{(i)})+\mathbf{1}\{Y_n^{(i)}>\frac12\}\log\frac{1-Y_n^{(i)}}{Y_n^{(i)}}+\mathbf{1}\{Y_{n+1}^{(i)}>\frac12\}\log\frac{Y_{n+1}^{(i)}}{1-u_{n+1}Z_n^*(1-Z_n^*)}\right]\\&\quad=S_{\sigma_i+k}(g)-S_{\sigma_i}(g)+k\pi(g)+\sum_{n=\sigma_i+1}^{\sigma_i+k-1}\mathbf{1}\{Y_n^{(i)}>\frac12\}\log\frac{1-Y_n^{(i)}}{1-u_nZ_{n-1}^*(1-Z_{n-1}^*)}\\&\qquad+\mathbf{1}\{Y_{\sigma_i}^{(i)}>\frac12\}\log\frac{1-Y_{\sigma_i}^{(i)}}{Y_{\sigma_i}^{(i)}}\;+\;\mathbf{1}\{Y_{\sigma_i+k}^{(i)}>\frac12\}\log\frac{Y_{\sigma_i+k}^{(i)}}{1-u_{\sigma_i+k}Z_{\sigma_i+k-1}^*(1-Z_{\sigma_i+k-1}^*)}.\end{aligned}$$

When $Y_n^{(i)}>\frac12$ and $\sigma_i<n<\tau_i$, by (26) we know that $1-u_nZ_{n-1}^*(1-Z_{n-1}^*)\ge X_n^*$, so

$$\frac{1-Y_n^{(i)}}{1-u_nZ_{n-1}^*(1-Z_{n-1}^*)}\le\frac{Y_n^*}{X_n^*}\le1+\epsilon_0.$$

Thus

$$\log\epsilon_{\sigma_i+k}^*-\log\epsilon_{\sigma_i}^*\le S_{\sigma_i+k}(g)-S_{\sigma_i}(g)+4\eta'k-\log\bigl(1-\frac{u_{k+\sigma_i}}{4}\bigr).$$

The event $\{\tau_i = \sigma_i + k\}$ occurs only when $\log \epsilon^*_{\sigma_i+k} - \log \epsilon^*_{\sigma_i} \geq -\log r_0$, so either

$$\begin{aligned} S_{\sigma_i+k}(g) - S_{\sigma_i}(g) &\geq \frac{1}{2}\big(-\log r_0 - 4\eta' k\big) \text{ or} \\ u_{\sigma_i+k} &\geq 4 - 4\sqrt{r_0}e^{2\eta' k}. \end{aligned}$$

By (38), and $V\big(Y_{\sigma_i}\big) \leq V(x_0)$, the former event has probability bounded by

$$2^{2p+2}c^*_{2p+2}V(x_0)\big(-\log r_0 - 4\eta' k\big)^{-2p-2}k^{p+1},$$

while the latter has probability no more than $\eta_\alpha 4^\alpha r_0^{\alpha/2}e^{2\eta' k\alpha}$. (These probabilities are almost sure, conditioned on $\mathcal{F}_{\sigma_i}$). Thus, on $\{\sigma_i < \infty\}$,

$$\begin{aligned} \mathrm{P}\big\{\infty > \tau_i - \sigma_i \geq n \,\big|\, \mathcal{F}_{\sigma_i}\big\} &= \sum_{k=n}^{\infty} \mathrm{P}\big\{\tau_i - \sigma_i = k \,\big|\, \mathcal{F}_{\sigma_i}\big\} \\ &\leq c^*_{2p+2}\sum_{k=n}^{\infty}\big(-\log r_0 - 4\eta' k\big)^{-2p-2}k^{p+1} \;+\; 4^\alpha r_0^{\alpha/2}\eta_\alpha \sum_{k=n}^{\infty} e^{4\eta' k\alpha/2} \end{aligned}$$

which is bounded by a constant times n^{-p} for each $p \geq 2$. Furthermore, the right side converges to 0 as r_0 goes to 0, so we may choose r_0 to make it smaller than 1. When $n = 1$, this is a uniform upper bound on the conditional probability that τ_i is finite, so we have taken care of (34) as well as (32). The same bounds show as well that

$$\begin{aligned} &\mathrm{P}\big\{\tau_i - \sigma_i \geq n \text{ and } \log \epsilon^*_{\sigma_i+n} - \log \epsilon_0 \geq -2\eta' n \,\big|\, \mathcal{F}_{\sigma_i}\big\} \\ &\qquad \leq \mathrm{P}\big\{S_n(g) \geq -\eta' n \,\big|\, \mathcal{F}_{\sigma_i}\big\} + \mathrm{P}\Big\{\frac{u_n}{4} \geq 1 - e^{-\eta' n} \,\Big|\, \mathcal{F}_{\sigma_i}\Big\}, \end{aligned}$$

so we can apply the same argument to prove (35). □

Proof of (33). For any $\sigma_i \leq k \leq \tau_i$, we have, from the definition of ϵ^*_k,

$$\begin{aligned} X^*_k &= \big(Y^*_k\big)\big(1 + \epsilon^*_k\big)^{-1} \\ &\geq \Big(Y^{(i)}_k\big(1 - Y^{(i)}_k\big)\Big)\Big(1 + \epsilon^*_{k-1}\cdot 4(1+\epsilon_0)\big[(1 - Y^*_{k-1}) \wedge \big(1 - \frac{u_k}{4}\big)\big]^{-1}\Big)^{-1} \\ &\geq Y^{(i)}_k\big(1 - Y^{(i)}_k\big)\Big(1 + 16(\epsilon_0 + \epsilon_0^2)\big(4 - u_k\big)^{-1}\Big)^{-1}. \end{aligned}$$

Because of the V-uniform ergodicity of the Markov chain $Y^{(i)}_k$, for any positive $\alpha' \leq \alpha$ and any $j \geq 0$,

$$\begin{aligned} \mathrm{E}\Big[\Big(Y^{(i)}_{j+\sigma_i}\big(1 - Y^{(i)}_{j+\sigma_i}\big)\Big)^{-\alpha'}\,\Big|\,\mathcal{F}_{\sigma_i}\Big] &\leq \mathrm{E}\big[V\big(Y^{(i)}_{j+\sigma_i}\big)\,\big|\,\mathcal{F}_{\sigma_i}\big]^{\alpha'/\alpha} \\ &\leq \big[cV\big(Y^{(i)}_{\sigma_i}\big)\big]^{\alpha'/\alpha} + \pi(V)^{\alpha'/\alpha} \\ &\leq \big[cV(x_0)\big]^{\alpha'/\alpha} + \pi(V)^{\alpha'/\alpha} =: C(\alpha'), \end{aligned} \tag{40}$$

almost surely, where c is a deterministic quantity depending on α' and the Markov chain, but not on j or i. If $c_5 = \alpha/3$, we may apply Hölder's inequality to get

$$\begin{aligned}
\mathrm{E}\big[X_{\tau_i}^{-c_5}\mathbf{1}\{\tau_i<\infty\}\,\big|\,\mathcal{F}_{\sigma_i}\big] &= \mathrm{E}\left[\sum_{k=\sigma_i+1}^{\infty}\big(X_k^*\big)^{-c_5}\mathbf{1}\{\tau_i=k\}\,\Big|\,\mathcal{F}_{\sigma_i}\right]\\
&\le \sum_{j=1}^{\infty}\mathrm{E}\Big[\big(Y_{\sigma_i+j}^{(i)}(1-Y_{\sigma_i+j}^{(i)})\big)^{-\alpha}\,\big|\,\mathcal{F}_{\sigma_i}\Big]^{1/3}\\
&\qquad\qquad \times\ \mathrm{E}\Big[\Big(1+16(\epsilon_0+\epsilon_0^2)\big(4-u_{\sigma_i+j}\big)^{-1}\Big)^{\alpha}\,\Big|\,\mathcal{F}_{\sigma_i}\Big]^{1/3}\\
&\qquad\qquad\qquad \times\ \mathrm{P}\big\{\infty>\tau_i-\sigma_i\ge j\,\big|\,\mathcal{F}_{\sigma_i}\big\}^{1/3}\\
&\le \sum_{j=1}^{\infty}C(\alpha)^{1/3}\Big(1+\eta_\alpha\Big)^{1/3}\ \big(b_p j^{-p}\big)^{1/3}
\end{aligned}$$

for any positive p. If $p>3$, this sum is finite, giving the desired bound. □

Lemma 10. *Let ν be a probability on $(0,4)$ such that $\nu\big((0,3]\big)>0$ and $\nu\big((1,4)\big)>0$, and u_i an i.i.d. sequence with distribution ν. Then for any positive x_0 and ϵ_0 sufficiently small there is a positive integer n_0, real $q>0$, and $y_0\in[x_0+\epsilon_0, 1-x_0-\epsilon_0]$, such that*

$$\mathrm{P}\big\{\forall x\in[x_0,1-x_0],\ \big|\widetilde{F}_{n_0}(x)-y_0\big|\le\epsilon_0\big\}\ge q \tag{41}$$

Proof. Suppose that $\nu\big((1,2]\big)>0$. Then we may choose x_0 to be small enough that ν puts positive mass on the interval $J:=[(1-x_0)^{-1},2]$. If $u_i\in J$ for all $i\le m$, the iterates up to m are monotonic on $(0,\frac12]$. This means that $\min_{x_0\le x\le 1-x_0}\widetilde{F}_m(x)=\widetilde{F}_m(x_0)\ge x_0$ and $\max_{x_0\le x\le 1-x_0}\widetilde{F}_m(x)=\widetilde{F}_m(1/2)$. If $\widetilde{F}_m(1/2)-\widetilde{F}_m(x_0)>\epsilon_0$, then

$$\frac{\widetilde{F}_{m+1}(\frac12)}{\widetilde{F}_{m+1}(x_0)}=\frac{1-\widetilde{F}_m(\frac12)}{1-\widetilde{F}_m(x_0)}\cdot\frac{\widetilde{F}_m(\frac12)}{\widetilde{F}_m(x_0)}\le(1-\epsilon_0)\cdot\frac{\widetilde{F}_m(\frac12)}{\widetilde{F}_m(x_0)}.$$

Consequently, letting $n_0=\lceil\log(\epsilon_0/2x_0)/\log(1-\epsilon_0)\rceil$, the image $\widetilde{F}_{n_0}\big([x_0,1-x_0]\big)$ falls into some subinterval of $[x_0,1-x_0]$ of length ϵ_0 with probability at least $\nu(J)^{n_0}$. Thus for some y_0 it must have positive probability of lying in $[y_0-\epsilon_0,y_0+\epsilon_0]$.

Suppose now that $\nu\big((2,3]\big)>0$. Let u be a point in $(2,3]$ which is in the support of ν. The function f_u has a stable fixed point $z_u=1-\frac1u\in(\frac12,\frac23)$. Consider x_0 and ϵ_0 such that $x_0+\epsilon_0\le z_u\le 1-x_0-\epsilon_0$. The fixed point is a universal attractor, so there is a number n_0 such that

$$\widetilde{F}_{n_0}\big([x_0,1-x_0]\big)\subset\big(z_u-\frac{\epsilon_0}{2},z_u+\frac{\epsilon_0}{2}\big).$$

Since $f_u(x)$ is uniformly continuous as a function of u, if J is a small enough interval around u, and $u_1,\dots,u_n\in J$, then

$$\widetilde{F}_n\big((x_0,1-x_0)\big)\subset\big(z_u-\epsilon_0,z_u+\epsilon_0\big).$$

If neither of these intervals has positive measure, it must be that $\nu\big((0,1]\big)$ and $\nu\big((3,4)\big)$ are both positive. We may find a $u\in(0,1]$ such that $\nu\big([u,1]\big)>0$. Let $J_1=[u,1]$ and $J_2=(3,4)$. We define $\{1,2\}$-valued $\mathcal{F}_{m-1}$-measurable random variables μ_m by the rule: $\mu_m=2$ if $\widetilde{F}_{m-1}(1/2)\le w_0:=1/2-1/\sqrt{8}$; otherwise $\mu_m=1$. We require that $x_0+\epsilon_0\le u/8$.

Suppose now that we have a sequence of u_i's with $u_i \in J_{\mu_i}$. Whenever $\mu_m = 2$,

$$\widetilde{F}_m(\frac{1}{2}) > 3(1-w_0)\widetilde{F}_{m-1}(\frac{1}{2}).$$

When $\mu_m = 1$ there is a drop, but $\widetilde{F}_m(\frac{1}{2})$ never goes below $u \cdot w_0(1-w_0) = u/8$. The wait between successive indices m with $\mu_m = 1$ is never more than $K := \lceil \log(8w_0/u)/\log 3(1-w_0) \rceil$. As in the first part, $\widetilde{F}_m(x) \le \frac{1}{2}$ for all m and x, so the functions act monotonically, and $\widetilde{F}_m(\frac{1}{2})/\widetilde{F}_m(x_0)$ is decreasing in m. If $\mu_m = 1$ and $\widetilde{F}_{m-1}(\frac{1}{2}) - \widetilde{F}_{m-1}(x_0) > \epsilon_0$, then

$$\frac{\widetilde{F}_m(\frac{1}{2})}{\widetilde{F}_m(x_0)} = \frac{1-\widetilde{F}_{m-1}(\frac{1}{2})}{1-\widetilde{F}_{m-1}(x_0)}\frac{\widetilde{F}_{m-1}(\frac{1}{2})}{\widetilde{F}_{m-1}(x_0)} < \left(1-\epsilon_0\right)\frac{\widetilde{F}_{m-1}(\frac{1}{2})}{\widetilde{F}_{m-1}(x_0)}.$$

If these conditions have been met k times before episode m, we have

$$(1-\epsilon_0)^k \frac{1}{2x_0} > \frac{\widetilde{F}_m(\frac{1}{2})}{\widetilde{F}_m(x_0)} > 1.$$

If we take $n = K \cdot \lceil \log(2x_0)/\log(1-\epsilon_0) \rceil + 1$, then

$$\begin{aligned} \mathrm{P}\big\{\exists n_0 \le n \text{ s.t. } \widetilde{F}_{n_0}([x_0, 1-x_0]) \subset [x_0, 1-x_0] \;\&\; \operatorname{diam} \widetilde{F}_{n_0}([x_0,1-x_0]) \le \epsilon_0\big\} & \\ \ge \mathrm{P}\big\{u_i \in J_{\mu_i} \text{ for all } i \le n\big\} & \\ \ge \min\big\{\nu(J_1), \nu(J_2)\big\}^n. & \qquad (42) \end{aligned}$$

□

Lemma 11. *The process (X_n^*, Z_n^*) satisfies*

$$\phi\Big(f_{n+1}\big([X_n^*, Z_n^*]\big)\Big) \subset [X_{n+1}^*, Z_{n+1}^*]. \qquad (43)$$

Proof. Stages I and II are obvious. Suppose now that the process is in Stage III, so $\sigma_i \le n < \tau_i$. To simplify the notation, we define $u = u_{n+1}$ and

$$\begin{aligned} x &= X_n^*, & x' &= X_{n+1}^*,\\ z &= Z_n^*, & z' &= Z_{n+1}^*,\\ \epsilon &= \epsilon_n^*, & \epsilon' &= \epsilon_{n+1}^*,\\ y &= Y_n^*, & y' &= Y_{n+1}^{(i)},\\ \mathcal{I} &= [x,z], & \mathcal{I}' &= \phi\big(f_{n+1}([x,z])\big). \end{aligned}$$

What we need to show is that, for all $w \in [x,z]$, letting $w' = uw(1-w)$,

$$\frac{\phi(y')}{1+\epsilon'} \le \phi(w') \le (1+\epsilon')\phi(y').$$

This is equivalent to showing that

$$\sup_{x \le w \le z} \max\left\{\frac{\phi(w')-\phi(y')}{\phi(y')}, \frac{\phi(y')-\phi(w')}{\phi(w')}\right\} \le \epsilon',$$

which will be a consequence of

$$\epsilon' \ge \tilde{\epsilon}(w) := \max\left\{\frac{w'-y'}{y'}, \frac{y'-w'}{w'}, \frac{|y'-w'|}{1-y'}, \frac{|y'-w'|}{1-w'}\right\}. \qquad (44)$$

Since $y' - w' = u(y-w)(1-y-w)$ and $y/(1+\epsilon) \le w \le y(1+\epsilon)$, and $\epsilon \le \epsilon_0$ (because the process is in stage III),

$$\frac{w'-y'}{y'} = \frac{(w-y)(1-y-w)}{y(1-y)} \le \frac{\epsilon\big(1-y(2+\epsilon)/(1+\epsilon)\big)}{1-y} \le \epsilon\frac{1-2y+\epsilon}{1-y} \le \epsilon', \text{ and}$$
$$\frac{y'-w'}{w'} = \frac{(y-w)(1-y-w)}{w(1-w)} \le \frac{\epsilon\big(1-y(2+\epsilon)/(1+\epsilon)\big)}{1-w} \le \epsilon\frac{1-2y+\epsilon}{(1-y)(1-\epsilon)} \le \epsilon'.$$

This takes care of the cases when y' or w' is the smallest of $\{y', w', 1-y', 1-w'\}$. Now consider the case when $y' \le \frac{1}{2} \le w'$. Since y and w are both in $(0, \frac{1}{2}]$, it must be that $y \le w \le (1+\epsilon)y$, which implies that $1-w' \ge 2y'-w' \ge uy(1-y-\epsilon)$. It follows that

$$\frac{w'-y'}{1-w'} \ge \frac{(w-y)(1-y-w)}{y(1-y-\epsilon)} \le \epsilon\frac{1-2y+\epsilon}{(1-y)(1-2\epsilon)} \le \epsilon'.$$

Finally, there is the case when $y' \ge \frac{1}{2}$. The denominator denominator of $\tilde{\epsilon}(w)$ is at least $1-uz(1-z)$, and

$$\tilde{\epsilon}(w) \le \frac{|y'-w'|}{1-uz(1-z)} \le \frac{u\epsilon y(1-y-y/(1+\epsilon))}{1-uz(1-z)} \le \epsilon'.$$

□

Acknowledgement – I would like to thank Persi Diaconis, who first suggested I look at Gerard Letac's conjectures about iterated logistic maps.

References

[AC92] Ludwig Arnold and Hans Crauel. Iterated function systems and multiplicative ergodic theory. In M. Pinksy and V. Wihstutz, editors, *Diffusion processes and related problems in analysis*, volume II, pages 283–305. Birkhäuser, 1992.

[AD] K. B. Athreya and Jack Dai. Random logistic maps – I. Preprint.

[Arn98] Ludwig Arnold. *Random Dynamical Systems*. Springer-Verlag, New York, Heidelberg, Berlin, 1998.

[BE88] Michael F. Barnsley and John H. Elton. A new class of Markov processes for image encoding. *The Annals of Applied Probability*, 20:14–32, 1988.

[BH86] Peter Baxendale and Theodore E. Harris. Isotropic stochastic flows. *The Annals of Probability*, 14(4):1155–1179, 1986.

[BR93] R. N. Bhattacharya and B. V. Rao. Random iterations of two quadratic maps. In et al. Stamatis Cambanis, editor, *Stochastic Processes: A Festscrift in honour of Gopinath Kallianpur*, pages 13–21. Springer-Verlag, 1993.

[CL91] Jean-François Chamayou and Gérard Letac. Explicit stationary distributions for compositions of random functions and products of random matrices. *Journal of Theoretical Probability*, 4(1):3–36, 1991.

[Cvi84] Predrag Cvitanovic. *Universality in Chaos*. Adam Hilger, Ltd., Bristol, 1984.

[Dev89] Robert L. Devaney. *An introduction to chaotic dynamical systems*. Addison-Wesley, 2nd edition, 1989.

[DF99] Persi Diaconis and David Freedman. Iterated function systems. *SIAM Review*, 41(1):45–76, 1999.

[Din93] I. H. Dinwoodie. Identifying a large deviation rate function. *The Annals of Probability*, pages 216–231, 1993.

[dMvS93] Wellington de Melo and Sebastian van Strien. *One-dimensional dynamics*. Number 3:25 in Ergebnisse der Mathematik und ihrer Grenzgebiete. Springer-Verlag, New York, Heidelberg, Berlin, 1993.

[Elt90] John H. Elton. A multiplicative ergodic theorem for Lipschitz maps. *Stochastic Processes Appl.*, 34(1):39–47, 1990.

[Fei84] Mitchell Feigenbaum. Universal behavior in nonlinear systems. In Predrag Cvitanovic, editor, *Universality in Chaos*, pages 48–84. Adam Hilger, Ltd., 1984. Originally published in *Los Alamos Science*, 1980.

[Hut81] J. E. Hutchinson. Fractals and self-similarity. *Indiana University Mathematics Journal*, 30:713–747, 1981.

[Klü00] Marc Klünger. Stochastic population models. Preprint, 2000.

[May76] Robert M. May. Simple mathematical models with very complicated dynamics. *Nature*, 261(5560):459–467, 1976.

[MT93] Sean Meyn and Richard Tweedie. *Markov chains and stochastic stability*. Springer Verlag, New York, Heidelberg, Berlin, 1993.

[Nev75] Jacques Neveu. *Discrete-Parameter Martingales*. North-Holland Publishing Company, Amsterdam, 1975. Trans. T. P. Speed.

[SS02] Michael Scheutzow and David Steinsaltz. Chasing balls through martingale fields. *Annals of Probability*, 2002. To appear.

[Ste99] David Steinsaltz. Locally contractive iterated function systems. *The Annals of Probability*, pages 1952–1979, 1999.

[Ste01] David Steinsaltz. Convergence of moments in a Markov-chain central limit theorem. *Indagationes Mathematicae*, 2001. To appear.

[Tim66] A. F. Timin. *Theory of approximations of a real variable*. Hindustan Publishing corporation, Delhi, 1966.

[Var84] Srinivasa Varadhan. *Large Deviations with Applications*. SIAM, Philadelphia, 1984.

[Wu00] Wei Bao Wu. Iterated function systems: Part i. Stationarity and path properties. Technical report, University of Michigan, 2000.

Department of Statistics, 367 Evans Hall, University of California, Berkeley, CA 94720, USA

E-mail address: `dstein@stat.berkeley.edu`