UNIQUENESS OF INVARIANT MEASURES FOR PLACE-DEPENDENT RANDOM ITERATIONS OF FUNCTIONS

ÖRJAN STENFLO

Abstract. We give a survey of some results within the convergence theory for iterated random functions with an emphasis on the question of uniqueness of invariant probability measures for place-dependent random iterations with finitely many maps. Some problems for future research are pointed out.

1. Introduction

Consider a finite set of continuous maps $\{w_i\}_{i=1}^N$ on some locally compact separable metric space (X, d) into itself. Associated to each map we are given continuous probability weights $p_i : X \to (0, 1)$, $p_i(x) > 0$, $i \in S := \{1, \ldots, N\}$ and

$$\sum_{i=1}^{N} p_i(x) = 1, \text{ for each } x \in X. \tag{1}$$

We call the set $\{(X, d);\ w_i(x),\ p_i(x), i \in S\}$ an IFS with place-dependent probabilities.

Specify a point $x \in X$. We are going to consider Markov chains $\{Z_n(x)\}$ heuristically constructed in the following way: Put $Z_0(x) := x$, and let $Z_n(x) := w_i(Z_{n-1}(x))$ with probability $p_i(Z_{n-1}(x))$, for each $n \geq 1$.

Let $C(X)$ denote the set of real-valued bounded continuous functions on X. Define the transfer operator $T : C(X) \to C(X)$ by

$$Tf(x) = \sum_{i=1}^{N} p_i(x) f(w_i(x)).$$

This operator characterizes the Markov chain. The fact that T maps $C(X)$ into itself is known as the Feller property. Markov chains with the Feller property are sometimes denoted Feller chains. We will mainly be interested in the problem of uniqueness/non-uniqueness of invariant probability measures. A probability measure π is invariant if

$$\int_X Tf d\pi = \int_X f d\pi, \tag{2}$$

Key words and phrases. Iterated Function Systems (IFS), Chains with complete connections, Ruelle-Perron-Frobenius operator, Markov chains, invariant measures, stationary measures, g-measures.

MSC2000: 28A80, 37H99, 60J05.

Supported by Hellmuth Hertz' Foundation and a HK RGC grant.

for all $f \in C(X)$. If we let $M(X)$ denote the set of Borel probability measures on X and define $T^\star : M(X) \to M(X)$ by requiring that $\int_X f dT^\star \nu = \int_X Tf d\nu$, for any $\nu \in M(X)$, and $f \in C(X)$, then (2) simply reads that $\pi \in M(X)$ is invariant iff $T^\star \pi = \pi$. ($T^\star$ is well defined by the Riesz representation theorem). Since T is assumed to have the Feller property it follows that $T^\star$ is continuous if we endow $M(X)$ with the topology of weak convergence (i.e. $\pi_n \xrightarrow{w} \pi \Leftrightarrow \int f d\pi_n \to \int f d\pi$, for all $f \in C(X)$).

It is known, see e.g. [51], that a Markov chain, $\{Z_n(x)\}$, with the Feller property always possesses at least one invariant probability measure under the mild condition that for any $\epsilon > 0$, and $x \in X$, there exists a compact set C such that $\liminf_{n\to\infty}(1/n)\sum_{j=0}^{n-1} P(Z_j(x) \in C) > 1 - \epsilon$. (Note that this condition holds trivially e.g. when (X, d) is compact.)

An invariant probability measure for the transfer operator is a stationary probability measure for the associated Markov chain. That is, a Markov chain "starting" according to a stationary probability measure will form a stationary stochastic process (with discrete time).

The first papers on random iterations were under the name "chains with complete connections". (Typically, the "index"-sequence is a stochastic sequence with "infinite connections").

Papers by Onicescu and Mihoc, e.g. [54], was motivated by applications to Urn models. In 1937 Doeblin and Fortet [20] published a paper which has had a great impact on future works in this subject. In 1950 Ionescu Tulcea and Marinescu [33] extended the work in [20]. We refer to [35] for a discussion of this and for further extensions. An important contribution was also given in Harris [29], whose ideas we are going to explore in Theorem 2 below. Place-dependent iterations has from the 50's also been studied under the name "learning models", see e.g. [18], [45], [38], [37], and [52].

The reader is referred to Kaijser [42] for an extensive survey of the literature up to 1980.

In the middle of the 80's there was a renewed attention in these kinds of models after Hutchinson [32], and Barnsley et al., [5] and [6] had demonstrated its importance within the theory of fractals. The concept of iterated function systems, introduced in [5], is nowadays the most widely used terminology. We refer to [28], [36], [27], [35] and [43] for results relating the convergence theory for IFS with results within the theory of chains with complete connections.

Recently it has also been realized that there is a strong link to the thermodynamic formalism of statistical mechanics. We are going to describe this connection briefly below. This important branch of symbolic dynamics started to develop in the 70's by works of Sinai [60], Ruelle [57],[58], Bowen [13] and others.

The present paper is organized as follows:

In Section 2 we review some results within the theory of place-independent iterations. Any Markov chain can be represented as an iterated function system with place-independent probabilities with (typically) an uncountable number of discontinuous maps (parameterized by the unit interval), see e.g. [47] or [2]. The results discussed in Section

2 can be considered as preliminaries for the next section where we are going to prove a convergence theorem for iterated function systems with place-dependent probabilities by making a place-independent representation and use techniques from the theory of place-independent iterations to obtain our result.

In Section 3 we consider place-dependent random iterations with "stable" maps.

We start in Section 3.1 by discussing results in the case when (X, d) is a symbolic space with finitely many symbols and the maps, $w_j,\ j \in S$, are a simple form of contractions. We present some smoothness conditions on the probability weights ensuring uniqueness of invariant measures and also, on the contrary, Bramson and Kalikow's example of a contractive IFS with place-dependent continuous (strictly positive) probabilities with more than one invariant probability measure. We also briefly describe the case when the probabilistic assumption (1) is relaxed. Such cases have been well-studied in statistical mechanics. In these cases we loose our probabilistic interpretation, but we can sometimes normalize the transfer operator and continue our analysis as in the probabilistic case.

In Section 3.2, we show how the results on symbolic spaces may be lifted to other compact spaces in case the maps in the IFS satisfy certain (deterministic) stability properties.

In Section 4 we discuss briefly some generalizations to stochastically stable situations, where the lifting method does not work out.

Finally in Section 5, we point out some problems for future research.

2. Iterated Function Systems with probabilities

Let (X, d) be a complete separable metric space, and let S be a measurable space. Consider a measurable function $w : X \times S \to X$. For each fixed $s \in S$, we write $w_s(x) := w(x, s)$. We call the set $\{(X, d);\ w_s, s \in S\}$ an iterated function system (IFS). (This generalizes the usual definition, as introduced in [5] (c.f. Section 1), where S typically is a finite set and the functions $w_s = w(\cdot, s) : X \to X$ typically have (Lipschitz) continuity properties.)

Let $\{I_n\}_{n=1}^{\infty}$ be a stochastic sequence with state space S. Specify a starting point $x \in X$. The stochastic sequence $\{I_n\}$ then controls the stochastic dynamical system $\{Z_n(x)\}_{n=0}^{\infty}$, where

$$Z_n(x) := w_{I_n} \circ w_{I_{n-1}} \circ \cdots \circ w_{I_1}(x),\ n \geq 1, \quad Z_0(x) = x. \tag{3}$$

We refer to [7], [24], [1], [12], and [59] for an overview of results in cases when $\{I_n\}$ has some dependence structure.

The particular case when $\{I_n\}$ is a sequence of independent and identically distributed (i.i.d.) random variables allows a richer analysis. See [47], [62] and [19] for surveys of this literature. We will assume that $\{I_n\}$ is i.i.d. in this section and concentrate on a result that will be useful in later sections.

Let μ denote the common distribution of the I_n's. We call the set $\{(X, d);\ w_s, s \in S, \mu\}$ an IFS with probabilities. The associated stochastic sequence

$\{Z_n(x)\}$ forms a Markov chain with transfer operator

$$Tf(x) = \int_S f(w_s(x))d\mu(s),\ f \in C(X).$$

For $x \in X$, define the reversed iterates

$$\hat{Z}_n(x) := w_{I_1} \circ w_{I_2} \circ \cdots \circ w_{I_n}(x),\ n \geq 1, \quad \hat{Z}_0(x) = x. \tag{4}$$

Since $\{I_n\}_{n=1}^{\infty}$ is i.i.d. it follows that $Z_n(x)$ and $\hat{Z}_n(x)$ defined in (3) and (4) respectively are identically distributed random variables for each fixed n and x. Thus in order to prove distributional limit results for the Markov chain $\{Z_n(x)\}$ as n tends to infinity we may instead study the pointwise more well behaved (but non-Markovian) sequence $\{\hat{Z}_n(x)\}$.

We say that a probability measure, π, is attractive if

$$P(Z_n(x) \in \cdot) \xrightarrow{w} \pi(\cdot)$$

for any $x \in X$, i.e. $T^n f(x) \to \int_X f d\pi$ for any $f \in C(X)$ and any $x \in X$.

Proposition 1. *An attractive probability measure for a Feller chain is uniquely invariant.*

Proof. Since $Tf \in C(X)$, for any $f \in C(X)$, the invariance of the attractive probability measure, π, follows immediately by taking limits in the equality $T^n(Tf(x)) = T^{n+1}f(x)$. Suppose ν is an arbitrary invariant probability measure. Then for any $f \in C(X)$,

$$\int_X f d\nu = \int_X T^n f d\nu \to \int_X \Big(\int_X f d\pi\Big) d\nu = \int_X f d\pi.$$

Therefore $\nu = \pi$. □

Corollary 1. *Suppose $\{(X,d);\ w_s, s \in S, \mu\}$ is an IFS with probabilities generating a Markov chain (3) with the Feller property. Suppose the limit*

$$\hat{Z} := \lim_{n\to\infty} \hat{Z}_n(x) \tag{5}$$

exists and does not depend on $x \in X$ a.s., then π defined by $\pi(\cdot) = P(\hat{Z} \in \cdot)$ is attractive and thus the unique invariant probability measure for $\{(X,d);\ w_s, s \in S, \mu\}$.

Proof. This can be seen from Proposition 1 by using that almost sure convergence implies convergence in distribution for $\hat{Z}_n(x)$ and by observing that $Z_n(x)$ and $\hat{Z}_n(x)$ are identically distributed for each fixed n and $x \in X$. □

Remark 1. *A slightly less general version of Corollary 1 was formulated as a principle in* [50].

In the case when the state space is compact, we obtain the following criteria for uniqueness of invariant probability measures.

Proposition 2. *Let (K, d) be a compact metric space and suppose $\{(K, d);\ w_s, s \in S, \mu\}$ is an IFS with probabilities generating a Markov chain (3) with the Feller property. Suppose*

$$diam(Z_n(K)) \xrightarrow{P} 0, \tag{6}$$

(where $\xrightarrow{P}$ denotes convergence in probability, and $diam(Z_n(K)) := \sup_{x,y\in K} d(Z_n(x), Z_n(y))$ is the diameter of the set $Z_n(K)$). Then there exists a unique invariant probability measure, π, for $\{(K, d);\ w_s, s \in S, \mu\}$, and π is uniformly attractive i.e.

$$\sup_{x\in K} |T^n f(x) - \int f d\pi| \to 0, \text{as } n \to \infty,$$

for any $f \in C(K)$.

Remark 2. *The criteria, (6), for uniqueness of invariant probability measures was introduced by Öberg in* [53].

Proof. Let $\{x_n\}$ be a sequence in K. It is sufficient to prove that the limit $\hat{Z} := \lim_{n\to\infty} \hat{Z}_n(x_n)$ exists and does not depend on $\{x_n\}$ *a.s.*

We are going to show that $\{\hat{Z}_n(x_n)\}$ is almost surely a Cauchy sequence. Since $\{I_n\}$ is i.i.d. it follows that condition (6) implies that $diam(\hat{Z}_n(K)) \xrightarrow{P} 0$. Since $\hat{Z}_{n+1}(K) \subset \hat{Z}_n(K)$, for any n, it follows that in fact $diam(\hat{Z}_n(K)) \xrightarrow{a.s.} 0$. For any positive integers n, and m with $n < m$ we have

$$d(\hat{Z}_n(x_n), \hat{Z}_m(x_m)) \le d(\hat{Z}_n(x_n), \hat{Z}_n(w_{I_{n+1}} \circ \cdots \circ w_{I_m}(x_m)) \le diam(\hat{Z}_n(K)). \tag{7}$$

Thus $\{\hat{Z}_n(x_n)\}$ is almost surely a Cauchy sequence which converges since K is complete. Since (7) holds uniformly in $\{x_n\}$ it follows that the *a.s.* limit $\hat{Z}$ is independent of $\{x_n\}$. Since almost sure convergence implies convergence in distribution, it follows that $T^n f(x_n) \to \int_K f d\pi$, for any $f \in C(K)$, where $\pi(\cdot) := P(\hat{Z} \in \cdot)$. This completes the proof of Proposition 2. □

We are going to use Proposition 2 in the section below.

Remark 3. *Note that in this section we do not require the family of maps $\{w_s\}$ to be finite or countable and that we do not require any of the maps in $\{w_s\}$ to be continuous. Thus in particular, we do not assume any global Lipschitz condition for any of the maps in $\{w_s\}$. Related results for locally contractive IFS can be found in* [40], [42], *and* [61].

3. Iterated Function Systems with place-dependent probabilities (deterministically stable cases)

Let $\{(X, d);\ w_i(x),\ p_i(x), i \in \{1, 2, ..., N\}\}$ be an IFS with place-dependent probabilities. We will suppose that the p_i's are strictly positive and uniformly continuous. For a uniformly continuous function $g : X \to (0, \infty)$, define the modulus of uniform continuity

$$\Delta_g(t) = \sup\{g(x) - g(y) : d(x, y) < t\}.$$

We are here going to present some uniform smoothness conditions on the p_i's and stability conditions on the family of maps $\{w_i\}$ that guarantee a unique invariant probability measure and see how a "phase transition" to non-uniqueness of invariant probability measures can occur if the smoothness conditions on the p_i's are relaxed for a fixed family of contractions $\{w_i\}$.

We start by discussing the important particular case when the state space is a symbolic space.

3.1. **The case when X is a symbolic space.** Let $\Sigma_N := \{1, 2, ..., N\}^{\mathbb{N}}$ and introduce a topology on Σ_N induced by the metric

$$\rho(\mathbf{i}, \mathbf{j}) := \begin{cases} 2^{-n}, & \text{if } \mathbf{i} \text{ and } \mathbf{j} \text{ differ for the first time in the} \\ & n^{\text{th}} \text{ digit} \\ 0, & \text{if } \mathbf{i} = \mathbf{j} \end{cases}$$

The space (Σ_N, ρ) is a compact metric space.

For $j \in \{1, 2, ..., N\}$ and $\mathbf{i} = i_1 i_2 ... \in \Sigma_N$, let $j\mathbf{i} = j i_1 i_2 ....$ Consider a continuous function $g : \Sigma_N \to (0, \infty)$, and suppose that g is normalized in the sense that

$$\sum_{j=1}^{N} g(j\mathbf{i}) = 1, \text{ for any } \mathbf{i} \in \Sigma_N. \tag{8}$$

(Such a function g is called a (strictly positive, continuous) g-function, see [46].)

Define $p_j(\mathbf{i}) = g(j\mathbf{i})$, and $w_j(\mathbf{i}) = j\mathbf{i}$. Then $\{(\Sigma_N, \rho);\ w_j(\mathbf{i}),\ p_j(\mathbf{i}), j \in \{1, 2, ..., N\}\}$ is an IFS with place-dependent probabilities. Note that the maps w_j are contractions, and that this system can be represented by the function g. Invariant probability measures for IFSs associated to a g-function are called g-measures. In Section 3.2 below we are going to see how results for this particular IFS can be lifted to prove results for IFSs on other state spaces by establishing a semi-conjugacy.

If we define $\Phi(\mathbf{i}) = \log g(\mathbf{i})$ and let θ denote the shift map (i.e. $\theta(i_1 i_2 ...) = i_2 i_3 ....$), we see that the transfer operator can be written as

$$Tf(\mathbf{i}) = \sum_{j=1}^{N} p_j(\mathbf{i}) f(w_j(\mathbf{i})) = \sum_{j=1}^{N} e^{\Phi(j\mathbf{i})} f(j\mathbf{i}) = \sum_{y \in \theta^{-1}\mathbf{i}} e^{\Phi(y)} f(y). \tag{9}$$

The right hand version of (9) is how this operator is most commonly expressed. The transfer operator, sometimes also called the Ruelle-Perron-Frobenius operator, occurs naturally in statistical physics but then the normalization condition (8) is not so natural.

We call the system, $\{(\Sigma_N, \rho);\ w_j(\mathbf{i}),\ p_j(\mathbf{i}), j \in \{1, 2, ..., N\}\}$, a weighted IFS in cases when g does not necessarily satisfy condition (8).

The following theorem is a consequence of a version of the "Ruelle-Perron-Frobenius theorem" proved by Walters [66].

Theorem 1. *(Walters (1975)) Let* $g : \Sigma_N \to (0,\infty)$ *be continuous, and let* $\{(\Sigma_N,\rho);\ w_j(\mathbf{i}),\ p_j(\mathbf{i}), j \in \{1,2,...,N\}\}$ *be the associated weighted IFS. Suppose*

$$\sum_{k=0}^{\infty} \Delta_g(2^{-k}) < \infty. \tag{10}$$

Then there exists a constant $\lambda > 0$ *(the spectral radius of* T*), a unique* λ*-invariant probability measure i.e. distribution,* π*, satisfying*

$$\int_{\Sigma_N} Tf d\pi = \lambda \int_{\Sigma_N} f d\pi,$$

for all $f \in C(\Sigma_N)$*, and a unique function* $h \in C(\Sigma_N)$ *with* $h > 0$ *such that*

$$Th = \lambda h, \text{ and } \int_{\Sigma_N} h d\pi = 1.$$

The probability measure π *is uniformly attractive in the sense that*

$$\sup_{x\in\Sigma_N} |\lambda^{-n}T^n f(x) - h\int f d\pi| \to 0, \text{as } n \to \infty,$$

for any $f \in C(\Sigma_N)$*.*

Proof. We refer to [66] for a rigorous proof. We will here only give some idea of its structure.

The existence of λ and π, follows immediately by applying the Schauder-Tychonoff fix-point theorem (see [22], p.456) to the map $\nu \to (\int_{\Sigma_N}\sum_{i=1}^N p_i(x)d\nu(x))^{-1}T^\star\nu$, $\nu \in M(\Sigma_N)$.

The existence of h is more intricate. This is proved by finding a carefully chosen convex compact subset of the non-negative functions in $C(\Sigma_N)$ that a normalized version of the transfer operator leaves invariant. The Schauder-Tychonoff fix-point theorem then gives the existence of h. Given $h > 0$ it is possible to define a strictly positive and continuous g-function and proceed as in the probabilistic case. Indeed, it can be shown that the function $\hat{g}$ defined by

$$\hat{g}(\mathbf{i}) = \frac{g(\mathbf{i})h(\mathbf{i})}{\lambda h(\theta\mathbf{i})} \tag{11}$$

is a g-function satisfying the conditions of Theorem 2 below. □

Remark 4. *Note that* $\lambda = 1$*, and* $h \equiv 1$*, in the probabilistic case when* g *is normalized. (Invariant measures is a short notation for* 1*-invariant measures.)*

Remark 5. *Theorem 1 is of importance in the thermodynamic formalism in statistical mechanics. The functions* g *(or* Φ*) are sometimes called "potentials". Condition (10) means that* g *is Dini-continuous. This condition (posed on* ϕ*) is usually referred to as "summable variation" in the thermodynamic formalism literature. Observe that* g *is Dini-continuous iff* Φ *is Dini-continuous since* g *is assumed to be strictly positive and* *continuous and thus bounded away from zero. Observe also that Hölder-continuous functions are Dini-continuous. If g is assumed to be Hölder-continuous then the convergence rate is exponential, see* [13]. *The letter g refers to "Gibbs" since the probability measure $\tilde{\pi}$, defined by $\tilde{\pi}(B) := \int_B h d\pi$, for Borel sets B of Σ_N, can be shown to have the Gibbs property under these conditions, i.e. there exists a constant $C \geq 1$ such that*

$$C^{-1}\hat{g}(i_1...i_{n-1}i_nx)\cdots\hat{g}(i_{n-1}i_nx)\hat{g}(i_nx)$$
$$\leq \tilde{\pi}([i_1...i_n]) \leq C\hat{g}(i_1...i_{n-1}i_nx)\cdots\hat{g}(i_{n-1}i_nx)\hat{g}(i_nx),$$

for any $x \in \Sigma_N$, and cylinder set $[i_1...i_n] := \{i_1...i_ny;\ y \in \Sigma_N\}$, where $\hat{g}$ denotes the normalized g-function defined in (11). (The measure $\tilde{\pi}$ is the unique invariant probability measure for the IFS with place-dependent probabilities associated with $\hat{g}$). The Gibbs property is of importance in e.g. the multifractal analysis of measures.

See e.g. [13], [58], [55], [8], [26] *and* [3] *for more on this and for an overview of further results in this field.*

Note that the above result can be stated without introducing the concept of IFSs. We have deliberately chosen to state it in this form since it gives a convenient notation in later sections when the state space under consideration is no longer assumed to be Σ_N. The reader is encouraged to compare our formulations of the theorems with the original works to get familiar with the notation.

Theorem 1 can be strengthen in the probabilistic case when (8) holds:

Theorem 2. *Let $g : \Sigma_N \to (0,1)$ be a continuous strictly positive g-function, and let $\{(\Sigma_N, \rho);\ w_j(\mathbf{i}),\ p_j(\mathbf{i}),\ j \in \{1,2,...,N\}\}$ be the associated IFS with place-dependent probabilities. Suppose*

$$\sum_{m=l}^{\infty}\prod_{k=l}^{m}\Big(1-\frac{N(N-1)}{2}\Delta_g(2^{-k})\Big)=\infty,\quad \text{for some integer } l\geq 1. \tag{12}$$

Then there exists a uniformly attractive (and thus necessarily unique) invariant probability measure.

Remark 6. *Observe that (12) holds if the Dini-condition (10) holds. Condition (12), (see also the slightly weaker condition (15) below), was introduced by Harris* [29]. *A condition for uniqueness of invariant probability measures closely related to (12) and (15) can be found in Berbee* [9].

Proof. The IFS $\{\Sigma_N; f_s,\ s \in (0,1)\}$ with

$$f_s(x) = \ w_i(x),\ \text{ if } \textstyle\sum_{j=1}^{i-1} p_j(x) \leq s < \sum_{j=1}^{i} p_j(x) \tag{13}$$

together with the Lebesgue measure restricted to $(0,1)$ is an IFS with probabilities that generates "the same" Markov chain i.e. a Markov chain with the same transfer operator as the given place-dependent system. It is more well behaved in the sense that it has place-independent probabilities but the loss is that it generally has a denumerable set of discontinuous functions. Let $\{I_n\}$ be a sequence of independent random variables uniformly distributed in $(0,1)$. Define $Z_n(x)$ and $\hat{Z}_n(x)$ as in Section 2. By Proposition 2 it suffices to show that $diam(Z_n(\Sigma_N)) \xrightarrow{P} 0$.

For a closed set $A \subset \Sigma_N$, and for $k = 1, 2, ..., N-1$, define $A_k := \{s \in (0,1) : \inf_{x\in A} \sum_{j=1}^{k} p_j(x) \leq s \leq \sup_{x\in A} \sum_{j=1}^{k} p_j(x)\}$. Note that,

$$\sup_{x,y\in A} \rho(f_s(x), f_s(y)) = \; diam(A)/2 \quad \text{if } s \in \Sigma_N \setminus \cup_{k=1}^{N-1} A_k$$

Let μ_{Leb} denote the Lebesgue measure. Clearly $\mu_{Leb}(A_k) \leq k\Delta_g(diam(A)/2)$, and thus $\mu_{Leb}(\cup_{k=1}^{N-1} A_k) \leq \frac{N(N-1)}{2}\Delta_g(diam(A)/2)$. (Trivially, we also have $\mu_{Leb}(\cup_{k=1}^{N-1} A_k) \leq 1 - \epsilon_0$, where $\epsilon_0 := \min_{x\in\Sigma_N} p_1(x)$.)

It follows that

$$P(diam(w_{I_1}(A)) = \frac{1}{2}diam(A)) \geq 1 - \min\Big(\frac{N(N-1)}{2}\Delta_g(diam(A)/2), 1-\epsilon_0\Big).$$

Thus there exists a homogeneous Markov chain $\{Y_n\}$ with

$$P(Y_{n+1} = 2^{-(j+1)} \mid Y_n = 2^{-j}) = 1 - \min\Big(\frac{N(N-1)}{2}\Delta_g(2^{-(j+1)}), 1-\epsilon_0\Big),$$

and

$$P(Y_{n+1} = 2^{-1} \mid Y_n = 2^{-j}) = \min\Big(\frac{N(N-1)}{2}\Delta_g(2^{-(j+1)}), 1-\epsilon_0\Big), \; j \geq 1,$$

such that

$$diam(Z_n(\Sigma_N)) \leq Y_n, \; n \geq 0.$$

It follows that for any $\epsilon > 0$,

$$P(diam(Z_n(\Sigma_N)) \geq \epsilon) \leq P(Y_n \geq \epsilon), \tag{14}$$

and since, by assumption Y_n is a null-recurrent Markov chain, see e.g [56], p.80, ex.18, 18, it follows that $P(Y_n \geq \epsilon) \to 0$ and therefore by (14) also $P(diam(Z_n(\Sigma_N)) \geq \epsilon) \to 0$ i.e. $diam(Z_n(\Sigma_N)) \xrightarrow{P} 0$.

□

Remark 7. *The method of finding an IFS with place-independent probabilities generating the same Markov chain as an IFS with place-dependent probabilities in order to prove ergodic theorems was introduced in* [63]*. Note that there is in general not a unique way of doing this. This technique can be thought of as a variant of the coupling method. Coupling is the method of comparing random variables by defining them on the same probability space. The art of coupling is to do this in the "best possible way" for the purpose needed. The coupling method, as a tool for proving convergence theorems for random iterations, is discussed in some detail in* [43]*.*

By making a more "efficient" IFS representation than (13), e.g. by in the $(k+1)$ iteration step using the "optimal" IFS representation depending on $Z_k(\Sigma_N)$, it is possible *to prove, see* [65]*, that condition (12) can be relaxed to*

$$\sum_{m=l}^{\infty}\prod_{k=l}^{m}(1-(N-1)\Delta_g(2^{-k}))=\infty, \quad \textit{for some integer } l\geq 1. \tag{15}$$

Question: Can Theorem 1 be proved under (an analogue of) the Harris condition (15) ?

To merely assume that a strictly positive g-function is continuous is not sufficient for a unique invariant probability measure. The following theorem is a reformulation of a result proved by Bramson and Kalikow in [14]:

Theorem 3. *(Bramson and Kalikow (1993)) Let* $0<\epsilon<1/4$ *be a fixed constant. Define* $q_k=\frac{1}{2}\left(\frac{2}{3}\right)^k$*,* $k\geq 1$*. Then there exists a sequence* $\{m_k\}_{k=1}^{\infty}$ *of odd positive integers such that* $\{(\Sigma_2,\rho),w_i,p_i,i\in\{1,2\}\}$ *with*

$$w_1(\mathbf{i})=1\mathbf{i} \quad \text{and} \quad w_2(\mathbf{i})=2\mathbf{i},$$

$$p_1(\mathbf{i})=\sum_{k=1}^{\infty}q_k f(\mathbf{i},m_k), \tag{16}$$

and $p_2(\mathbf{i})=1-p_1(\mathbf{i})$*, where*

$$f(\mathbf{i},k)=\begin{cases}1-\epsilon & \text{if } \frac{\sum_{n=1}^{k}i_n}{k}<\frac{3}{2}\\ \epsilon & \text{otherwise}\end{cases}$$

is an IFS with place-dependent probabilities that has more than one invariant probability measure.

Remark 8. *Note the particular form of* p_1*. Kalikow* [44] *proved that in fact any continuous g-function admits a representation of the form*

$$g(\mathbf{i})=\sum_{k=1}^{\infty}q_k f(\mathbf{i},k),$$

where $\mathbf{i}=i_1i_2\ldots$*, and for fixed* k*,* $f(\mathbf{i},k)$ *is a function of* $(i_1,\ldots,i_k)$*, where* $0\leq f(\mathbf{i},k)\leq 1$*,* $0\leq q_k\leq 1$ *for any* k*, and* $\sum_{k=1}^{\infty}q_k=1$*.*

Such a representation is called a random Markov chain representation since the projection on the first coordinate of the g-generated Markov chain on Σ_N *forms a random Markov chain on* $\{1,\ldots,N\}$ *i.e., heuristically, a "n"-step Markov chain where* n *is random with* $P(n=k)=q_{k+1}$*.*

In [44]*, Kalikow also gave arguments implying that if* $\sum_{k=1}^{\infty}kq_k<\infty$ *then a strictly positive g-function has a unique g-measure. It is straightforward to check that g in fact is Dini-continuous under that condition.*

Remark 9. *In dynamical systems terminology, Bramson and Kalikow's theorem can be stated: There exists a sequence* $\{m_k\}_{k=1}^{\infty}$ *of odd positive integers such that the continuous function* $g:\Sigma_2\to(0,1)$*, defined by* $g(1\mathbf{i}):=p_1(\mathbf{i})$*, and* $g(2\mathbf{i})=p_2(\mathbf{i})$*, for any* $\mathbf{i}\in\Sigma_2$*, where* p_1 *is defined as in (16), and* $p_2=1-p_1$*, is a continuous strictly positive g-function with more than one g-measure.*

Proof. (Reformulation of the proof in [14]) We shall first define the sequence $\{m_k\}$. We do this inductively; Given that $m_1, m_2, .., m_{j-1}$ are already defined we define

$$\tilde{p}_1^{(j)}(\mathbf{i})=\sum_{k=1}^{j-1}q_kf(\mathbf{i},m_k)+\epsilon q_j+(1-\epsilon)\sum_{k=j+1}^{\infty}q_k,$$

and let $\{Z_n^{(j)}(\mathbf{i})\}$ denote a Markov chain starting at $\mathbf{i}\in\Sigma_2$, with transfer operator

$$T_{(j)}f(\mathbf{i})=\tilde{p}_1^{(j)}(\mathbf{i})f(1\mathbf{i})+(1-\tilde{p}_1^{(j)}(\mathbf{i}))f(2\mathbf{i}).$$

(Note that $\{Z_n^{(j)}(\mathbf{i})\}$ has a unique stationary probability measure since Kalikow's condition is fulfilled.)

We have,

$$\tilde{p}_1^{(j)}(\mathbf{i})-(\sum_{k=1}^{j-1}q_kf(\mathbf{i},m_k)+\frac{1}{2}\sum_{k=j}^{\infty}q_k)=\Big(\frac{1}{2}-\epsilon\Big)\Big(\sum_{k=j+1}^{\infty}q_k-q_j\Big)=q_j\Big(\frac{1}{2}-\epsilon\Big)>\frac{q_j}{4}.$$

Thus, by the ergodic theorem and comparison with the "symmetric process" associated with $\sum_{k=1}^{j-1}q_kf(\mathbf{i},m_k)+\frac{1}{2}\sum_{k=j}^{\infty}q_k$, it follows that the process $\{Z_n^{(j)}(\mathbf{i})\}$ will have

$$\lim_{n\to\infty}P\Big(Z_n^{(j)}(\mathbf{i})\in\Big\{\mathbf{i}\in\Sigma_2:\frac{\sum_{k=1}^n i_k}{n}\le\frac{3}{2}-\frac{q_j}{4}\Big\}\Big)=1.$$

Choose an odd positive integer m_j with $m_j>8m_{j-1}/q_j$, such that

$$P(Z_{m_j}^{(j)}(\mathbf{i})\in A_j)\le 3^{-(j+1)},\quad\text{uniformly in }\mathbf{i}\in\Sigma_2,$$

where

$$A_j:=\Big\{\mathbf{i}\in\Sigma_2:\frac{\sum_{n=1}^{m_j}i_n}{m_j}\ge\frac{3}{2}-\frac{q_j}{8}\Big\}.$$

Let $\mathbf{1}\in\Sigma_2$ denote the infinite sequence 111..., and let $W_{\mathbf{1}}$ denote the class of probability measures μ on Σ_2 that are weakly $\mathbf{1}$-concentrated in the sense that,

$$\mu\in W_{\mathbf{1}}\Leftrightarrow\mu(A_j)\le 3^{-j},\text{for all } j\ge 1.$$

Let $\{Z_n(\mathbf{1})\}$ be the Markov chain starting in $\mathbf{1}\in\Sigma_2$ generated by the given IFS with place-dependent probabilities determined by (16), and define

$$\mu_n^{\mathbf{1}}(\cdot):=P(Z_n(\mathbf{1})\in\cdot),\ n\ge 0.$$

The idea is to show that $\mu_n^{\mathbf{1}}\in W_{\mathbf{1}}$ for any n from which it follows that there exists an invariant probability measure, $\pi_1\in W_{\mathbf{1}}$.

This gives non-uniqueness in invariant probability measures by symmetry reasons since, we can analogously introduce a class of weakly $\mathbf{2}$-concentrated probability measures and argue in the same way by starting a Markov chain in the sequence $\mathbf{2} = 222...$, to obtain a weakly $\mathbf{2}$-concentrated invariant probability measure, π_2.

The proof that $\mu_n^{\mathbf{1}} \in W_{\mathbf{1}}$ for any n is by induction. Suppose that $\mu_n \in W_{\mathbf{1}}$, for $n < n_0$ i.e. $\mu_n(A_j) \leq 3^{-j}$ for all $n < n_0$, and $j \geq 1$. Fix an arbitrary integer $k \geq 1$. By conditioning on the values of $Z_{n_0-m_k}(\mathbf{1})$ (understanding $Z_n(\mathbf{1}) = \mathbf{1}$, for $n \leq 0$), and using the induction hypothesis, we obtain

$$\begin{aligned}
\mu_{n_0}(A_k) &= P(Z_{n_0}(\mathbf{1}) \in A_k) \\
&\leq P(Z_{n_0}(\mathbf{1}) \in A_k \mid Z_{n_0-m_k}(\mathbf{1}) \in (\Sigma_2 \backslash (\cup_{j=k+1}^{\infty} A_j))) \\
&\quad + P(Z_{n_0-m_k}(\mathbf{1}) \in \cup_{j=k+1}^{\infty} A_j) \\
&\leq P(Z_{n_0}(\mathbf{1}) \in A_k \mid Z_{n_0-m_k}(\mathbf{1}) \in (\Sigma_2 \backslash (\cup_{j=k+1}^{\infty} A_j))) + \sum_{j=k+1}^{\infty} 3^{-j} \\
&\leq P(Z_{m_k}(\mathbf{i}) \in A_k \mid \mathbf{i} \in (\Sigma_2 \backslash (\cup_{j=k+1}^{\infty} A_j))) + 2 \cdot 3^{-(k+1)}, \qquad (17)
\end{aligned}$$

where we in the last step used the Markov property. Since for $\mathbf{i} \in (\Sigma_2 \backslash (\cup_{j=k+1}^{\infty} A_j))$ we have that $p_1(\mathbf{i}) \geq \tilde{p}_1^{(k)}(\mathbf{i})$ or more generally

$p_1(a_1 a_2 ... a_n \mathbf{i}) \geq \tilde{p}_1^{(k)}(a_1 a_2 ... a_n \mathbf{i})$, for any $n \leq m_k$, and $a_j \in \{1, 2\}$, $1 \leq j \leq n$, it follows that

$$P(Z_{m_k}(\mathbf{i}) \in A_k \mid \mathbf{i} \in (\Sigma_2 \backslash (\cup_{j=k+1}^{\infty} A_j))) \leq P(Z_{m_k}^{(k)}(\mathbf{i}) \in A_k) \leq 3^{-(k+1)},$$

and we thus obtain from (17) that $\mu_{n_0}(A_k) \leq 3^{-k}$. Since k was arbitrary, the proof of Theorem 3 now follows by using the induction principle. □

3.2. **Lifting the symbolic space results to other compact spaces.** In this section we are going to consider cases when the limit in (5) exist also in a deterministic sense, i.e. the limit

$$\hat{Z}(\mathbf{i}) = \lim_{n \to \infty} w_{i_1} \circ w_{i_2} \circ \cdots \circ w_{i_n}(x), \qquad (18)$$

exists and is independent of $x \in X$, for any sequence $\mathbf{i} = i_1 i_2 ... \in \Sigma_N$, and the map $\hat{Z} : \Sigma_N \to X$ is continuous, i.e. the limit in (18) is uniform in $x \in X$.

As an example, see e.g. [30] or [4], this is the case if $\{(K, d);\ w_i(x), i \in S = \{1, 2, ..., N\}\}$ is a weakly contractive IFS, i.e. $d(w_i(x), w_i(y)) < d(x, y)$ for all $x, y \in K$ and $i \in S$, and (K, d) is a compact metric space.

We shall assume in what follows, that $(X, d) := (K, d)$ is a compact metric space where $K = \hat{Z}(\Sigma_N)$. Since $\hat{Z}(\Sigma_N)$ is compact when $\hat{Z}$ is continuous, this gives no further restrictions.

We will demonstrate how it in this case is possible to "lift" the results from Section 3.1 on symbolic spaces to other compact spaces by establishing a (semi)-conjugacy. This technique was first used in Fan and Lau [26] and was explored in further detail in [49] and [64].

Define

$$Diam_n(K) = \sup_{\mathbf{i}\in\Sigma_N} diam(w_{i_1} \circ w_{i_2} \circ \cdots \circ w_{i_n}(K)),$$

and let

$$\Delta_p(t) := \max_{i\in S} \Delta_{p_i}(t).$$

(Note that since we have assumed that $\hat{Z}$ is continuous, it follows that $Diam_n(K) \to 0$ by Dini's theorem.)

As corollaries of Theorem 1 and the stronger form of Theorem 2 (as given in Remark 7), we obtain

Corollary 2. *Let* $\{(K, d);\ w_i(x),\ p_i(x), i \in \{1, 2, ..., N\}\}$ *be a weighted IFS, i.e.* $p_i : K \to (0, \infty)$ *are continuous, and* $\sum_{j=1}^N p_j(x)$ *is not necessarily assumed to be* 1*. Suppose*

$$\sum_{n=0}^{\infty} \Delta_p(Diam_n(K)) < \infty.$$

Then there exists a constant $\lambda > 0$ *(the spectral radius of* T*), a unique* λ*-invariant probability measure i.e. distribution,* π*, satisfying*

$$\int_K Tf d\pi = \lambda \int_K f d\pi,$$

for all $f \in C(K)$*, and a unique function* $h \in C(K)$ *with* $h > 0$ *such that*

$$Th = \lambda h, \text{ and } \int_K h d\pi = 1.$$

The probability measure π *is uniformly attractive in the sense that*

$$\sup_{x\in K} |\lambda^{-n} T^n f(x) - h \int f d\pi| \to 0, \text{as } n \to \infty,$$

for any $f \in C(K)$*.*

Corollary 3. *Suppose* $\{(K, d);\ w_i(x),\ p_i(x), i \in \{1, 2, ..., N\}\}$ *is an IFS with place-dependent strictly positive continuous probabilities and suppose*

$$\limsup_{n\to\infty} n\Delta_p(Diam_n(K)) < (N-1)^{-1}.$$

Then there exists a uniformly attractive (and thus necessarily unique) invariant probability measure.

Proof. Define $g : \Sigma_N \to (0, 1)$ by

$$g(\mathbf{i}) := p_{i_1}(\hat{Z}(\theta(\mathbf{i})))$$

Then $\Delta_g(2^{-n}) \leq \Delta_p(Diam_{n-2}(K))$, $n \geq 2$. We can now apply Theorem 1, and the stronger form of Theorem 2 as given in Remark 7 respectively, to obtain a unique (λ)-invariant probability measure, $\tilde{\pi}$ (and eigen-function $\tilde{h} \in C(\Sigma_N)$) for the IFS with probabilities on the symbolic space discussed in Section 3.1. The probability measure

$\pi(\cdot) = \tilde{\pi}(\mathbf{i} : \hat{Z}(\mathbf{i}) \in \cdot)$ is uniquely (λ)-invariant, and $h(\hat{Z}(\mathbf{i})) := \tilde{h}(\mathbf{i})$ has the desired properties. (We refer to [26] for further details.) This proves Corollary 2 and Corollary 3. □

Remark 10. *Similar results and extensions of Corollary 2 have been proved in* [26] *and* [49].

Remark 11. *If the IFS with place-dependent probabilities satisfy certain monotonicity conditions that makes the generated Markov chain stochastically monotone, then it is possible to relax the regularity conditions on the p_i's and still prove that there is a unique invariant probability measure. See e.g.* [21], [41], [11] [31], [17] *and* [10].

As a consequence of Bramson and Kalikow's result (Theorem 3 above), we obtain the following theorem;

Theorem 4. *(Stenflo (2001)) Let w_1 and w_2 be two maps from $[0,1]$ into itself defined by*

$$w_1(x) = \sigma x \quad \text{and} \quad w_2(x) = \alpha + (1-\alpha)x,$$

where both $0 < \sigma < \alpha < 1$ are constant parameter values. Then there exists a continuous function $p_1 : [0,1] \to (0,1)$ such that the IFS $\{[0,1];\ w_i(x), i \in \{1,2\}\}$ with probabilities $p_1(x)$ and $p_2(x) := 1 - p_1(x)$ generates a Markov chain with more than one stationary probability measure.

Proof. (Sketch) For a sequence $\mathbf{i} = i_1 i_2 ... \in \Sigma_2$, define

$$\hat{Z}(\mathbf{i}) = \lim_{n\to\infty} w_{i_1} \circ w_{i_2} \circ \cdots \circ w_{i_n}(0).$$

The map $\hat{Z} : \Sigma_2 \to [0,1]$ is continuous and $1-1$ and the image of Σ_2 is a Cantor set, C. Define, for $x \in \mathrm{C}$, $p_1(x) := \hat{p_1}(\hat{Z}^{-1}(x))$, where $\hat{p_1}(\mathbf{i})$ is defined as in (16) and extend p_1 for points $x \in [0,1]\backslash\mathrm{C}$ by linear interpolation. Then p_1 will have the desired properties.

We refer to Stenflo [64], for further details. □

Remark 12. *Theorem 4 constitutes a counterexample to the conjecture that an IFS on the unit interval with two contractive maps and place-dependent strictly positive continuous probabilities necessarily has a unique invariant probability measure. See* [42], [43] *and* [64] *for accounts on the history of that conjecture.*

3.3. **E-chains.** Suppose that (X,d) is a locally compact separable metric space. Let $C_c(X)$ denote the set of continuous functions with compact support. We say (following the notion of [51]) that a Markov chain is an e-chain if for any $f \in C_c(X)$, $\{T^n f\}$ is equi-continuous on compact sets. It follows from the Arzela-Ascoli theorem, see e.g. [25], or [51], that Feller chains with an attractive invariant measure are in fact e-chains.

Conversely, we have

Theorem 5. *Let (K, d) is a compact metric space. Suppose $\{(K, d);\ w_i(x),\ p_i(x), i \in S = \{1, 2, ..., N\}\}$ is an IFS with place-dependent strictly positive probabilities generating an e-chain and the map $\hat{Z} : \Sigma_N \to K$ of (18) exists and is continuous and onto. Then there exists a uniformly attractive (and thus necessarily unique) invariant probability measure.*

Proof. We will make a slight generalization of a proof by Keane [46]; Equip the set of continuous functions on K, $C(K)$, with the supremum norm, $\| \cdot \|$. Let $f \in C(K)$. Note that $\|T^n f\| \leq \|f\|$ for any $n \in \mathbb{N}$. Thus $\{T^n f\}$ is a bounded equi-continuous sequence in $C(K)$ and we obtain from the Arzela-Ascoli theorem, that there exists a function $f_\star \in C(K)$ and an increasing sequence $\{n_i\}$ of positive integers, such that $\|T^{n_i} f - f_\star\| \to 0$ as $i \to \infty$.

Clearly

$$\min_{x \in K} f(x) \leq \min_{x \in K} Tf(x) \leq ... \leq \min_{x \in K} f_\star(x).$$

Note that $\min_{x \in K} f_\star(x) = \min_{x \in K} Tf_\star(x)$. Assume $\min_{x \in K} f_\star(x) = Tf_\star(y_1) = \sum_{i \in S} p_i(y_1) f_\star(w_i(y_1))$ for some $y_1 \in K$. Then it follows that $f_\star(w_i(y_1)) = \min_{x \in K} f_\star(x)$ for all $i \in S$ and similarly for any finite sequence $\{i_k\}_{k=1}^m$, $m \geq 1$, of integers in S, $f_\star(w_{i_1} \circ \cdots \circ w_{i_m}(y_m)) = \min_{x \in K} f_\star(x)$, for some $y_m \in K$. Since $\hat{Z}$ is continuous, it follows that $f_\star$ is constant. Thus it follows that in fact $\|T^n f - f_\star\| \to 0$ as $n \to \infty$, and thus, by the Riesz representation theorem, there exists a probability measure, π, such that $\|T^n f - \int f d\pi\| \to 0$ as $n \to \infty$ and therefore we see from Proposition 1 that π must be uniquely invariant. □

Remark 13. *It is surprisingly difficult to construct Feller chains that are not e-chains, see* [51]*. Note however that the system in Bramson and Kalikow's theorem is an IFS with strictly positive place-dependent probabilities that generates a Feller chain that is not an e-chain. See* [67] *for further examples.*

4. Iterated Function Systems with place-dependent probabilities (stochastically stable cases)

In Section 3.2 above we treated the case when the limit in (5) exists in a deterministic sense. This is the case when the maps $\{w_i\}$ are (weakly) contractive maps and (X, d) is compact. In this section we are going to discuss cases when the limit in (5) does not necessarily exist a priori, and cases when the state space (X, d) is no longer assumed to be compact.

Assume that (X, d) is a locally compact separable metric space where sets of finite diameter are relatively compact. We are going to consider systems that are contractive on the average. (Convergence theorems for place-dependent random iterations with non-expansive maps on general state spaces can be found in [48]).

The following theorem is a consequence of a theorem proved by Barnsley et al. [6].

Theorem 6. *(Barnsley et al. (1988)) Let $\{(X, d);\ w_i(x),\ p_i(x),\ i \in S = \{1, 2, ..., N\}\}$ be an IFS with place-dependent probabilities with all $w_i, i \in S$ being*

Lipschitz-continuous and where all p_i's are Dini-continuous and bounded away from 0. *Suppose*

$$\sup_{x\neq y}\sum_{i=1}^{N} p_i(x)\log\frac{d(w_i(x),w_i(y))}{d(x,y)}<0. \tag{19}$$

Then the generated Markov chain has an attractive (and thus necessarily unique) stationary probability measure.

Remark 14. *A local version of the log-average contraction condition, (19), was used in Kaijser* [42] *and a proof of Theorem 6 in the case when (X,d) is a compact metric space also follows from* [42].

By Jensen's inequality, condition (19) is more general than the average contraction condition

$$\sup_{x\neq y}\sum_{i=1}^{N} p_i(x)\frac{d(w_i(x),w_i(y))}{d(x,y)}<1, \tag{20}$$

introduced by Isaac [38]. *Isaac proved the special case of Theorem 6, when (X,d) is assumed to be compact, condition (20) holds, and the p_i's are assumed to be Lipschitz-continuous.*

Question: Can Theorem 6 be proved if the p_i's satisfy Harris condition (15) ?

The following result was proved in [23].

Theorem 7. *(Elton (1987)) Assume the conditions of Theorem 6. Let $\{Z_n(x)\}$ denote the associated Feller chain and let π denote its unique stationary probability measure. Then*

$$\lim_{n\to\infty}\frac{\sum_{k=0}^{n-1} f(Z_k(x))}{n}=\int_X f\,d\pi,\ \text{ for all } f\in C(X) \quad a.s., \tag{21}$$

for any $x\in X$.

Remark 15. *In the case when (X,d) is compact, Breiman* [15] *proved that*

$$\lim_{n\to\infty}\frac{\sum_{k=0}^{n-1} f(Z_k(x))}{n}=\int_X f\,d\pi \quad a.s., \tag{22}$$

for any fixed $x\in X$ and $f\in C(X)$, in fact holds for any Feller chain with a unique stationary probability measure π.

Question: Does (21), or the (in general) slightly weaker assertion (22), hold for an e-chain with a unique invariant measure ?

5. The future

We have briefly mentioned the coupling method as an important tool to prove limit theorems for IFSs with place-dependent probabilities. This method, in this context, is described in [34], and [43]. We expect that this method, as a tool here, will be explored in further detail. There are of course a variety of further questions related to the ergodic behavior of stochastic sequences arising from place-dependent iterations that we have not treated here. It is typically possible to give exponential rates of convergence and central limit theorems in cases when the probabilities are Hölder continuous. An interesting work investigating mixing properties and convergence rates as a function of the smoothness of the probabilities (on symbolic spaces) is given in [16]. Coupling techniques is a basic tool there. An analogue of their work for not necessarily contractive systems on other state spaces would be interesting.

We have seen, that when the limit in (5) exists in a deterministic sense, then we can apply the machinery on symbolic space by lifting methods.

We have also seen that uniqueness of invariant probabilities can be proved also in cases when this limit does not necessarily exist in a deterministic sense.

A thorough investigation whether a probabilistic analogue of (5) still holds in these known cases seems to be lacking in the literature.

It would be interesting to find a correct generalization of the "lifting" method in cases when the limit in (5) only exist in a probabilistic sense. A result in this direction is given in [39].

Acknowledgments: This paper was written during a postdoctoral visit at the Department of Mathematics at the Chinese University of Hong Kong. I am grateful to Ka-Sing Lau for making this research possible, for valuable discussions, and for creating a stimulating research environment to work in. Thanks also to Michael Barnsley, John Elton and De-Jun Feng for helpful comments and discussions.

References

[1] L. Arnold. *Random dynamical systems.* Springer-Verlag, Berlin, 1998.

[2] K. B. Athreya and Ö. Stenflo. Perfect sampling for Doeblin chains. *Preprint*, 2001.

[3] V. Baladi. *Positive transfer operators and decay of correlations.* World Scientific Publishing Co. Inc., River Edge, NJ, 2000.

[4] M. F. Barnsley, A. Deliu, and R. Xie. Stationary stochastic processes and fractal data compression. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, 7(3):551–567, 1997.

[5] M. F. Barnsley and S. Demko. Iterated function systems and the global construction of fractals. *Proc. Roy. Soc. London Ser. A*, 399(1817):243–275, 1985.

[6] M. F. Barnsley, S. G. Demko, J. H. Elton, and J. S. Geronimo. Invariant measures for Markov processes arising from iterated function systems with place-dependent probabilities. *Ann. Inst. H. Poincaré Probab. Statist.*, 24:367–394, 1988 (Erratum: 25:589–590, 1989).

[7] M. F. Barnsley, J. H. Elton, and D. P. Hardin. Recurrent iterated function systems. *Constr. Approx.*, 5(1):3–31, 1989. Fractal approximation.

[8] T. Bedford, M. Keane, and C. Series, editors. *Ergodic theory, symbolic dynamics, and hyperbolic spaces*. The Clarendon Press Oxford University Press, New York, 1991. Papers from the Workshop on Hyperbolic Geometry and Ergodic Theory held in Trieste, April 17–28, 1989.
[9] H. Berbee. Chains with infinite connections: uniqueness and Markov representation. *Probab. Theory Related Fields*, 76(2):243–253, 1987.
[10] R. Bhattacharya and M. Majumdar. On a theorem of Dubins and Freedman. *J. Theoret. Probab.*, 12(4):1067–1087, 1999.
[11] R. N. Bhattacharya and O. Lee. Asymptotics of a class of Markov processes which are not in general irreducible. *Ann. Probab.*, 16(3):1333–1347, 1988 (Correction 25:1541–1543,1997).
[12] A. A. Borovkov. *Ergodicity and stability of stochastic processes*. John Wiley & Sons Ltd., Chichester, 1998. Translated from the 1994 Russian original by V. Yurinsky [V. V. Yurinskiĭ].
[13] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Springer-Verlag, Berlin, 1975. Lecture Notes in Mathematics, Vol. 470.
[14] M. Bramson and S. Kalikow. Nonuniqueness in g-functions. *Israel J. Math.*, 84:153–160, 1993.
[15] L. Breiman. The strong law of large numbers for a class of Markov chains. *Ann. Math. Statist.*, 31:801–803, 1960.
[16] X. Bressaud, R. Fernández, and A. Galves. Decay of correlations for non-Hölderian dynamics. A coupling approach. *Electron. J. Probab.*, 4:no. 3, 19 pp. (electronic), 1999.
[17] R. M. Burton and G. Keller. Stationary measures for randomly chosen maps. *J. Theoret. Probab.*, 6:1–16, 1993.
[18] R. R. Bush and F. Mosteller. A stochastic model with applications to learning. *Ann. Math. Statistics*, 24:559–585, 1953.
[19] P. Diaconis and D. Freedman. Iterated random functions. *SIAM Rev.*, 41(1):45–76 (electronic), 1999.
[20] W. Doeblin and R. Fortet. Sur des chaines a liaisons completès. *Bull. Soc. Math. France*, 65:132–148, 1937.
[21] L. E. Dubins and D. A. Freedman. Invariant probabilities for certain Markov processes. *Ann. Math. Statist.*, 37:837–848, 1966.
[22] N. Dunford and J. T. Schwartz. *Linear Operators. I. General Theory.* Interscience Publishers, Inc., New York, 1958.
[23] J. H. Elton. An ergodic theorem for iterated maps. *Ergodic Theory Dynam. Systems*, 7(4):481–488, 1987.
[24] J. H. Elton. A multiplicative ergodic theorem for Lipschitz maps. *Stochastic Process. Appl.*, 34(1):39–47, 1990.
[25] J. H. Elton and Z. Yan. Approximation of measures by Markov processes and homogeneous affine iterated function systems. *Constr. Approx.*, 5(1):69–87, 1989. Fractal approximation.
[26] A. H. Fan and K.-S. Lau. Iterated function system and Ruelle operator. *J. Math. Anal. Appl.*, 231(2):319–344, 1999.
[27] Ş. Grigorescu. Limit theorems for Markov chains arising from iterated function systems. *Rev. Roumaine Math. Pures Appl.*, 37(10):887–899, 1992.
[28] Ş. Grigorescu and G. Popescu. Random systems with complete connections as a framework for fractals. *Stud. Cerc. Mat.*, 41(6):481–489, 1989.
[29] T. E. Harris. On chains of infinite order. *Pacific J. Math.*, 5:707–724, 1955.
[30] M. Hata. On the structure of self-similar sets. *Japan J. Appl. Math.*, 2(2):381–414, 1985.
[31] P. Hulse. Uniqueness and ergodic properties of attractive g-measures. *Ergodic Theory Dynam. Systems*, 11(1):65–77, 1991.
[32] J. E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, 30(5):713–747, 1981.
[33] C. T. Ionescu Tulcea and G. Marinescu. Théorie ergodique pour des classes d'opérations non complètement continues. *Ann. of Math. (2)*, 52:140–147, 1950.

[34] M. Iosifescu. A coupling method in the theory of dependence with complete connections according to Doeblin. *Rev. Roumaine Math. Pures Appl.*, 37(1):59–66, 1992.

[35] M. Iosifescu. A basic tool in mathematical chaos theory: Doeblin and Fortet's ergodic theorem and Ionescu Tulcea and Marinescu's generalization. In *Doeblin and modern probability (Blaubeuren, 1991)*, pages 111–124. Amer. Math. Soc., Providence, RI, 1993.

[36] M. Iosifescu and Ş. Grigorescu. *Dependence with complete connections and its applications.* Cambridge University Press, Cambridge, 1990.

[37] M. Iosifescu and R. Theodorescu. *Random processes and learning.* Springer-Verlag, New York, 1969. Die Grundlehren der mathematischen Wissenschaften, Band 150.

[38] R. Isaac. Markov processes and unique stationary probability measures. *Pacific J. Math.*, 12:273–286, 1962.

[39] A. Johansson and A. Öberg. *preprint*, 2001.

[40] T. Kaijser. A limit theorem for Markov chains in compact metric spaces with applications to products of random matrices. *Duke Math. J.*, 45(2):311–349, 1978.

[41] T. Kaijser. Another central limit theorem for random systems with complete connections. *Rev. Roumaine Math. Pures Appl.*, 24:383–412, 1979.

[42] T. Kaijser. On a new contraction condition for random systems with complete connections. *Rev. Roumaine Math. Pures Appl.*, 26:1075–1117, 1981.

[43] T. Kaijser. On a theorem of Karlin. *Acta Appl. Math.*, 34:51–69, 1994.

[44] S. Kalikow. Random Markov processes and uniform martingales. *Israel J. Math.*, 71(1):33–54, 1990.

[45] S. Karlin. Some random walks arising in learning models. I. *Pacific J. Math.*, 3:725–756, 1953.

[46] M. Keane. Strongly mixing g-measures. *Invent. Math.*, 16:309–324, 1972.

[47] Y. Kifer. *Ergodic theory of random transformations.* Birkhäuser Boston Inc., Boston, MA, 1986.

[48] A. Lasota and J. A. Yorke. Lower bound technique for Markov operators and iterated function systems. *Random Comput. Dynam.*, 2(1):41–77, 1994.

[49] K.-S. Lau and Y.-L. Ye. Ruelle operator with weakly contractive IFS. *Studia Math.*, to appear.

[50] G. Letac. A contraction principle for certain Markov chains and its applications. In *Random matrices and their applications (Brunswick, Maine, 1984)*, pages 263–273. Amer. Math. Soc., Providence, RI, 1986.

[51] S. P. Meyn and R. L. Tweedie. *Markov chains and stochastic stability.* Springer-Verlag London Ltd., London, 1993.

[52] M. F. Norman. *Markov processes and learning models.* Academic Press, New York, 1972. Mathematics in Science and Engineering, Vol. 84.

[53] A. Öberg. Approximation of invariant measures for random iterations. *U.U.D.M. Rep. 6, Dept. of Mathematics, Uppsala University*, 1997.

[54] O. Onicescu and G. Mihoc. Sur les chaînes de variables statistiques. *Bull. Sci. Math. de France*, 59:174–192, 1935.

[55] W. Parry and M. Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque*, (187-188):268, 1990.

[56] N. U. Prabhu. *Stochastic processes. Basic theory and its applications.* The Macmillan Co., New York, 1965.

[57] D. Ruelle. Statistical mechanics of a one-dimensional lattice gas. *Comm. Math. Phys.*, 9:267–278, 1968.

[58] D. Ruelle. *Thermodynamic formalism.* Addison-Wesley Publishing Co., Reading, Mass., 1978. The mathematical structures of classical equilibrium statistical mechanics, With a foreword by Giovanni Gallavotti and Gian-Carlo Rota.

[59] D. S. Silvestrov and Ö. Stenflo. Ergodic theorems for iterated function systems controlled by regenerative sequences. *J. Theoret. Probab.*, 11(3):589–608, 1998.

[60] Ja. G. Sinaĭ. Gibbs measures in ergodic theory. *Uspehi Mat. Nauk*, 27(4(166)):21–64, 1972.

[61] D. Steinsaltz. Locally contractive iterated function systems. *Ann. Probab.*, 27(4):1952–1979, 1999.
[62] Ö. Stenflo. *Ergodic theorems for iterated function systems controlled by stochastic sequences.* Doctoral thesis No 14, Dept. of Mathematics, Umeå University, 1998.
[63] Ö. Stenflo. Ergodic theorems for Markov chains represented by iterated function systems. *Bull. Polish. Acad. Sci. Math.*, 49(1):27–43, 2001.
[64] Ö. Stenflo. A note on a theorem of Karlin. *Statist. Probab. Lett.*, 54(2):183–187, 2001.
[65] Ö. Stenflo. Uniqueness in g-measures. *Preprint*, 2001.
[66] P. Walters. Ruelle's operator theorem and g-measures. *Trans. Amer. Math. Soc.*, 214:375–387, 1975.
[67] R. Zaharopol. Iterated function systems generated by strict contractions and place-dependent probabilities. *Bull. Polish Acad. Sci. Math.*, 48(4):429–438, 2000.

Department of Mathematics, Umeå University, SE-90187 Umeå, Sweden

Department of Mathematics, Lady Shaw Building, The Chinese University of Hong Kong, Shatin, N.T., Hong Kong
E-mail address: stenflo@math.umu.se