INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2024, VOL. 70, NO. 4, PP. 1113–1123
Manuscript received October 15, 2024; revised October, 2024. doi: 10.24425/ijet.2024.152514

A computer scientist's perspective on approximation of IFS invariant sets and measures with the random iteration algorithm

Tomasz Martyn

***Abstract*—We study invariant sets and measures generated by iterated function systems defined on countable discrete spaces that are uniform grids of a finite dimension. The discrete spaces of this type can be considered as models of spaces in which actual numerical computation takes place. In this context, we investigate the possibility of the application of the random iteration algorithm to approximate these discrete IFS invariant sets and measures. The problems concerning a discretization of hyperbolic IFSs are considered as special cases of this more general setting.**

***Keywords*—IFS, Discrete Space; Markov Chain; Approximation; Invariant Set; Invariant Measure**

I. The problem

Let $\{\mathbb{R}^n; w_1, \ldots, w_m\}$, $m \in \mathbb{N}$, be a *hyperbolic iterated function system* (IFS) on a metric space $(\mathbb{R}^n, d)$, where d is a metric induced by a norm on $\mathbb{R}^n$, and the mappings w_i are contractions on $(\mathbb{R}^n, d)$. One can show that $\mathcal{H}(\mathbb{R}^n)$, the family of all compact and nonempty subsets of $\mathbb{R}^n$, when endowed with the Hausdorff metric h (induced by d), forms a complete metric space $(\mathcal{H}(\mathbb{R}^n), h)$. Moreover, if a contraction w on $(\mathbb{R}^n, d)$ is regarded as a set mapping $w : \mathcal{H}(\mathbb{R}^n) \to \mathcal{H}(\mathbb{R}^n)$, then w is a contraction operator on $(\mathcal{H}(\mathbb{R}^n), h)$ with the contractivity factor not greater than that of w acting on $(\mathbb{R}^n, d)$. This observation forms a basis for constructing the *Hutchison operator* $W : \mathcal{H}(\mathbb{R}^n) \to \mathcal{H}(\mathbb{R}^n)$ defined by $W(E) := \bigcup_{i=1}^{m} w_i(E)$. The Hutchinson operator a contraction on $(\mathcal{H}(\mathbb{R}^n), h)$ with the contractivity factor not exceeding the maximum of the contractivity factors of w_i. Finally, putting the contractivity of W and the completeness of $(\mathcal{H}(\mathbb{R}^n), h)$ together, by Banach's fixed-point theorem we get that W possesses exactly one fixed point $A_\infty = W(A_\infty)$, and moreover

$$\lim_{k\to\infty} W^{\circ k}(B) = A_\infty \tag{I.1}$$

regardless of $B \in \mathcal{H}(\mathbb{R}^n)$. The fixed point A_∞ is called the *attractor* of the IFS, and being an element of $\mathcal{H}(\mathbb{R}^n)$ it is a nonempty and compact subset of $(\mathbb{R}^n, d)$.

An analogous argument, also based on Banach's fixed-theorem, is utilized to prove the existence and the uniqueness of the *invariant measure* of the hyperbolic *IFS with probabilities*, $\{K; w_1, \ldots, w_m; p_1, \ldots, p_m\}$, where $\sum_{i=1}^{n} p_i = 1$, $p_i > 0$, and K is a compact subset of $(\mathbb{R}^n, d)$. This time the role of the complete space is taken on by the compact (and thus complete) metric space $(\mathcal{P}(K), d_{w^*})$, where $\mathcal{P}(K)$ is the family of Borel probability measures on K, and d_{w^*} is the Monge-Katorovich metric. The role of the contraction operator acting on $(\mathcal{P}(K), d_{w^*})$ is played by the Markov operator $T^* : \mathcal{P}(K) \to \mathcal{P}(K)$ defined by $T^*(\mu) := \sum_{i=1}^{N} p_i \mu \circ w_i^{-1}$. In this setting, by the Banach fixed point theorem

$$\lim_{k\to\infty} (T^*)^{\circ k}(\mu) = \pi_\infty \tag{I.2}$$

no matter what $\mu \in \mathcal{P}(K)$, that is, the IFS invariant measure $\pi_\infty \in \mathcal{P}(K)$ is the unique attractive fixed point of the Markov operator. Moreover, one can show that π_∞ is supported by the IFS attractor.

In actual implementations, however, the space upon which IFS mappings operate is always merely a discrete model of $\mathbb{R}^n$. From the point of view of computation the most basic model of the real numbers is a countable discrete space founded on floating-point arithmetic. Spaces of this kind are the natural setting within which implementation of the most popular algorithm for approximating IFS attractors and measures, the *random iteration algorithm* [1], generates its sequences of points.

The problem is that contraction mappings on $\mathbb{R}^n$ can lose their contraction properties when forced to act on a countable approximation of the space, even in such simple cases as contractive similarity transformations [2]. Moreover, this decline of contractivity caused by discretization is independent of precision, in the sense that no matter what a discretization granularity, the probability that a discrete version of an affine contraction is not a contraction remains unchanged. Therefore, we cannot get rid of this phenomenon by increasing the density of discretization. Since for an IFS to be hyperbolic all its component mappings have to be contractions, the probability that a discrete version of a hyperbolic affine IFS retains hyperbolicity decreases exponentially with the number of IFS mappings. As a consequence, actual implementations of an affine IFS are hardly ever hyperbolic even if the original IFS is so. Therefore, at least from the perspective of the standard hyperbolic IFS theoretical foundations, we cannot count on the discrete counterparts of the Hutchinson and Markov operators to be contractions in the discrete space, and thus, in such

T. Martyn is with Faculty of Electronics and Information Theory, Warsaw University of Technology, Poland (e-mail: tomasz.martyn@pw.edu.pl).

a situation, Banach's fixed-point theorem is not applicable, at least directly. In light of these facts, we can go even further and question the relevancy of outcomes supplied by the implementations of the algorithms built around the operators, asking how the outcomes are related, if at all, to the true attractor sets and measures of the original IFSs which are hyperbolic in $(\mathbb{R}^n, d)$?

The above problems direct us to the central subject of this paper, which concerns *discrete IFSs* that are not merely derivatives of hyperbolic IFSs but they are self-standing mathematical objects defined on a grid space from the very beginning. The main question around which this paper is organized is about the place of such self-standing discrete IFSs within the framework of the theory of iterated function systems. We treat the issues concerning discretization of hyperbolic IFSs as special instances of this general problem.

II. Minimal absorbing sets

In this section we develop a theory of minimal absorbing sets for maps acting in a discrete space. In some respects a minimal absorbing set of a map in the discrete space can be viewed as the counterpart of fixed points of contractions in $\mathbb{R}^n$. However we will define and analyse minimal absorbing sets in isolation from contractive maps in $\mathbb{R}^n$, as freestanding structures that can arise from the dynamics of a single map in a discrete space. Then we will show that discretization of a contraction in $\mathbb{R}^n$ leads to a map in a discrete space that can be regarded as a special case of the more general construction we will have considered earlier. We take advantage of the concept of minimal absorbing set throughout this paper, and such a generalization allows us to study, in Sec. III, discrete iteration function systems on their own, with the ones that arise from discretization of hyperbolic IFSs as a special case.

Hereafter, we will assume that the metric d on $\mathbb{R}^n$ is induced by a norm, that is, $d(x, y) := \|x - y\|$.

The following definition introduces the fundamental structure of this paper—a discrete counterpart of $\mathbb{R}^n$:

Definition II.1. *Let $\mathcal{G}^n(\delta)$ be the regular tiling of $\mathbb{R}^n$ by disjoint, half-open n-dimensional cubes of side $\delta > 0$, which are defined by*

$$C_\delta(m_1, \dots, m_n) := \Big[(m_1 - \tfrac{1}{2})\delta, (m_1 + \tfrac{1}{2})\delta\Big) \times \dots \times \Big[(m_n - \tfrac{1}{2})\delta, (m_n + \tfrac{1}{2})\delta\Big), \tag{II.1}$$

for $m_1, \dots, m_n \in \mathbb{Z}$, and thus

$$\mathcal{G}^n(\delta) := \big\{C_\delta(m_1, \dots, m_n) : m_1, \dots, m_n \in \mathbb{Z}\big\}.$$

We will call $\mathcal{G}^n(\delta)$ the δ-grid *in $\mathbb{R}^n$. We define the δ-*discretization *of $\mathbb{R}^n$, and denote it by $\mathcal{D}^n(\delta)$, as the set of the centers of the cubes in $\mathcal{G}^n(\delta)$, that is,*

$$\mathcal{D}^n(\delta) := \big\{\delta[m_1, \dots, m_n] : m_1, \dots, m_n \in \mathbb{Z}\big\}.$$

Moreover, for the sake of brevity, we will write θ to denote $\operatorname{diam}_d(C_\delta)/2$, half of the diameter of a δ-cube with respect to the metric d.

The discretization space $\mathcal{D}^n(\delta)$ is a formal model of the actual space upon which the concrete implementations of algorithms really act, and from which they yield their results. For example, if we regard squares $C_\delta(m_1, m_2)$ as pixels of size δ, then $\mathcal{D}^2(\delta)$ can be identified with an image space. However, when it comes to floating-point arithmetic, the issue is more complicated. Due to a fixed number of significant digits[1] in the floating point number representation, the representable numbers are not evenly spaced and the distance between consecutive numbers grows with scale[2]. In the context of the δ-grid model, this means that to represent the natural setting for floating point arithmetic we would have to use a grid with δ a monotonically increasing function of real numbers. Nevertheless, to keep things as simple as possible, in such a case we can assume constant δ set to a certain tiny number that roughly reflects the error carried by the floating point approximation in a bounded interval $[-a, a]$. For example, we can follow the approach used in numerical analysis and accept δ to be equal to the machine accuracy, which is the smallest (in magnitude) floating point number which, when added to the floating point 1.0, produces a floating point result different from 1.0 (see e.g. [3]).

In the sequel, for convenience, we will often treat $\mathcal{D}^n(\delta)$ as a subset of $\mathbb{R}^n$, that is, without defining any explicit mapping of points from $\mathcal{D}^n(\delta)$ to $\mathbb{R}^n$; in other words, we assume that any point of $\mathcal{D}^n(\delta)$ is by definition a point of $\mathbb{R}^n$. The converse is in general not true and is the subject of the next definition which establishes a relationship between points and mappings on $\mathbb{R}^n$ and the ones of $\mathcal{D}^n(\delta)$.

Definition II.2. *We define the δ-*roundoff *of a point $x \in \mathbb{R}^n$ as the result of the operation $\tilde{\cdot} : \mathbb{R}^n \to \mathcal{D}^n(\delta)$ such that $\tilde{x} = \delta m$, where $m \in \mathbb{Z}^n$ such that $x \in C_\delta(m)$. Using the operator, the δ-*roundoff *of a mapping $w : \mathbb{R}^n \to \mathbb{R}^n$ is defined to be the mapping $\tilde{w} : \mathcal{D}^n(\delta) \to \mathcal{D}^n(\delta)$ such that $\tilde{w}(\tilde{x}) = \widetilde{w(\tilde{x})}$.*

In words, given a point in $\mathbb{R}^n$, the operator $\tilde{\cdot}$ finds a δ-cube within which the point resides and returns the cube's center. Obviously, it is only a conceptual model of actual rounding that goes in real computation environment, in which there are no such things as real numbers to be rounded. The actual process operates all the time only on some representations of ideal reals and the representations are accurate at most for a finite subset of rational numbers. Moreover, as to rounding a mapping w, our model is very optimistic because it assumes that the result of a computed and rounded value of w is within 0.5δ of the exact result (w.r.t. the maximum metric d_∞), regardless of the class of the mapping itself. This is the best possible accuracy of computation in a given δ-grid setup. However, in floating point computation (IEEE 754 standard) such precision[3] is guaranteed only for elementary arithmetic operations (addition, subtraction, multiplication, division, and square root). Since mappings are typically composed of more

[1] Typically the 23-bit and 52-bit mantissa for single- and double-precision, which translates to about 7 and 16 decimal digits.

[2] More strictly, the numbers are evenly spaced in the intervals $[2^j, 2^{j+1}]$, possibly excluding one or both endpoints, and at the endpoints the interval between adjacent numbers doubles [4].

[3] Here, we identify the value of δ with the ULP (acronym for *unit in the last place*), which is a measure of accuracy of floating point arithmetic, defined usually (but not always) as the gap between the two floating-point numbers nearest the number to be rounded [4], [5].

than one elementary operation, the error accumulates. Nevertheless, if a mapping consists of a finite number of elementary operations (such as affine mappings, for instance), then the error of rounding is bounded above by a constant, so our model of rounding remains valid up to a multiple of δ by a constant.

It is also worth noting that the operator $\tilde{\cdot}$ is a surjection and in fact any mapping on $\mathcal{D}^n(\delta)$ represents uncountable many mappings on $\mathbb{R}^n$. For now on, if, in a given context, we do not take advantage of some specific properties of the mapping $w : \mathbb{R}^n \to \mathbb{R}^n$ when discussing its δ-roundoff, for brevity we will often refer to $\tilde{w}$ as a mapping on $\mathcal{D}^n(\delta)$.

The next definition introduces the concept of an absorbing set—the set that has the ability to attract the orbits $\{\tilde{w}^{\circ i}(\tilde{x})\}_i$ and then trap them forever.

Definition II.3. *Let $\tilde{w} : \mathcal{D}^n(\delta) \to \mathcal{D}^n(\delta)$, and let C be a nonempty subset of $\mathcal{D}^n(\delta)$ such that $\tilde{w}(C) \subset C$. We will say that a set $\Lambda \subset C$ is an* absorbing set for $\tilde{w}$ in C *if*

$$\forall \tilde{x} \in C, \exists N \in \mathbb{N}, \forall i \geq N, \tilde{w}^{\circ i}(\tilde{x}) \in \Lambda.$$

In the sequel, if Λ is an absorbing set for $\tilde{w}$ in $\mathcal{D}^n(\delta)$, we will just write that Λ is an absorbing set for $\tilde{w}$, that is, without pointing out $\mathcal{D}^n(\delta)$ as a superset of Λ. Similarly, if it is obvious from the context which mapping $\tilde{w}$ is in question, for brevity we will often omit explicit indication of the mapping and write that Λ is an absorbing set in a set.

It is easy to prove[4] the properties of absorbing sets listed in the following theorem:

Theorem II.1. *For $\tilde{w} : \mathcal{D}^n(\delta) \to \mathcal{D}^n(\delta)$ and any nonempty set $C \supset \tilde{w}(C)$ the following statements hold:* (a) *Every absorbing set is nonempty (by definition).* (b) *Since C is nonempty, then by definition C is an absorbing set in C, so there is always at least one absorbing set in C.* (c) *If $\{\Lambda_i\}_{1 \leq i \leq K}$ is a finite family of K absorbing sets in C, then also $\bigcap_{i=1}^{K} \Lambda_i$ is an absorbing set in C.* (d) *If Λ is an absorbing set in C, then also $\tilde{w}(\Lambda)$ is an absorbing set in C.* (e) *For any absorbing set Λ in C, there exists $B \subset \Lambda$ such that $\tilde{w}(B) \subset B$ and B is an absorbing set in C.* (f) *If Λ is an absorbing set in C, and C is an absorbing set in D, then Λ is also an absorbing set in D.*

In search of the minimal absorbing set, in the theorem below we look at the result of the intersection of all absorbing sets in a given set.

Theorem II.2. *Let $\{\Lambda_\alpha\}_{\alpha \in \mathcal{I}}$ be the (possibly uncountable) family of all absorbing sets for $\tilde{w}$ in $C \subset \mathcal{D}^n(\delta)$. Then $\mathcal{M} := \bigcap_{\alpha \in \mathcal{I}} \Lambda_\alpha$ is the set of all periodic points of $\tilde{w}$ in C, and thus $\mathcal{M} = \tilde{w}(\mathcal{M})$.*

Remark II.1. *Note that although $\mathcal{M}$ arises from the intersection of absorbing sets, it does not follow that $\mathcal{M}$ is an absorbing set itself. In general, when $C \supset \tilde{w}(C)$ is unbounded (and hence the family of the absorbing sets in C may be uncountable), there is no guarantee that the intersection is nonempty, and even if it is not, $\mathcal{M}$ does not have to attract points lying outside. As an example of the latter case consider a δ-roundoff of the mapping $w(x) = \lambda x$, $\lambda > 1$, with $C = \mathcal{D}^n(\delta)$. There are infinitely many (unbounded) absorbing sets, each includes $\mathbf{0}$, and $\mathcal{M} = \{\mathbf{0}\}$, because $\mathbf{0}$ is the only periodic point of $\tilde{w}$. However, $\mathcal{M}$ does not have the attractive property of an absorbing set required by Def. II.3.*

[4] All proofs of the theorems and lemmas presented in the paper can be found in [12]

Now observe that for any nonempty set $B \subset \mathcal{D}^n(\delta)$ such that $B \supset \tilde{w}(B)$, an orbit $\{\tilde{w}^{\circ i}(\tilde{x})\}$ of $\tilde{x} \in B$ does not have to visit all points in B. This observation suggests that in general B can be divided into disjoint subsets that are domains for some orbits in B and which are omitted by the remaining orbits in B—the ones that stay in the boundaries of the other subsets of the division. To this end, we define a relation $\overset{orb}{\sim}$ on B, which given a couple of points in B checks if the orbits of the couple coincide at a certain point:

$$\tilde{x} \overset{orb}{\sim} \tilde{y} := \{(\tilde{x}, \tilde{y}) \in B^2 : \exists j, k \in \mathbb{N} \ s.t. \ \tilde{w}^{\circ j}(\tilde{x}) = \tilde{w}^{\circ k}(\tilde{y}) \}.$$

It is easy to see that the relation is an equivalence relation and thus uniquely decomposes B into the union of disjoint nonempty subsets, $B = \bigcup_i B_i$ with B_i being the equivalence classes of the relation. By definition, each B_i is composed of orbits in B that meet at a certain point and then, naturally, coincide from that point on. Furthermore, since $\mathcal{D}^n(\delta)$ is countable and B_i are disjoint, the decomposition consists of a countable number of sets B_i. Moreover, if B is additionally an absorbing set in a set C, then the decomposition of B extends to C, resulting in a countable partition of C into the sets of the form $\{\tilde{x} \in C : \exists i \in \mathbb{N} \ s.t. \ \tilde{w}^{\circ i}(\tilde{x}) \in B_i\}$.

We summarize the above observations in the form of the following definition concerning the minimum of absorbing sets, their components and basins of attraction:

Definition II.4. *If the unique set $\mathcal{M} \subset C$ defined in Theorem II.2 meets the conditions of an absorbing set in C we will refer to it as the* minimal absorbing set for $\tilde{w}$ in C *and denote it by $\mathcal{M}[\tilde{w}, C]$. If $\mathcal{M}$ is the minimal absorbing set for $\tilde{w}$ in the whole space $\mathcal{D}^n(\delta)$, we will say that $\mathcal{M}$ is just the* minimal absorbing set for $\tilde{w}$ *and denote it by $\mathcal{M}[\tilde{w}]$. The disjoint, nonempty sets of the unique, countable decomposition of $\mathcal{M}[\tilde{w}, C]$ with respect to the relation $\overset{orb}{\sim}$ will be called the set's* components *and denoted by $\mathcal{M}_i[\tilde{w}, C]$, $i = 1, \ldots$. Finally, the* basin of attraction *of a component $\mathcal{M}_i[\tilde{w}, C]$ is defined by*

$$\mathcal{B}[\mathcal{M}_i[\tilde{w}, C]] := \{\tilde{x} \in C : \exists i \in \mathbb{N} \ s.t. \ \tilde{w}^{\circ i}(\tilde{x}) \in \mathcal{M}_i[\tilde{w}, C]\}.$$

Clearly, periodic points $\tilde{x}$ and $\tilde{y}$ are in the coincidence orbit relation, $\tilde{x} \overset{orb}{\sim} \tilde{y}$, if and only if they share the same periodic orbit. Since, by Theorem II.2, the minimum set $\mathcal{M}[\tilde{w}, C]$ consists of periodic points, we immediately get the following:

Corollary II.3. *Each component $\mathcal{M}_i[\tilde{w}, C]$ of the minimal absorbing set is equal to one of the periodic orbits for $\tilde{w}$ in C. Hence, the cardinality of every component is finite and*

$$\tilde{w}(\mathcal{M}_i[\tilde{w}, C]) = \mathcal{M}_i[\tilde{w}, C].$$

As shown in Remark II.1, an intersection of absorbing sets does not always yield an absorbing set. The following theorem gives a sufficient condition for the existence of the minimal absorbing set:

Theorem II.4. *If the set $C \subset \mathcal{D}^n(\delta)$ is bounded and $C \supset \tilde{w}(C)$, then the minimal absorbing set $\mathcal{M}[\tilde{w}, C]$ exists.*

The next theorem shows that the inheritance of absorbing sets by extensions of the original domain, asserted in Theorem II.1 (f), also holds for minimum of absorbing sets with preserving the property of minimality.

Theorem II.5. *If $\mathcal{M}$ is the minimal absorbing set for $\tilde{w}$ in C, and C is an absorbing set for $\tilde{w}$ in D, then $\mathcal{M}$ is also the minimal absorbing set for $\tilde{w}$ in D, that is,*

$$\mathcal{M}[\tilde{w}, C] = \mathcal{M}[\tilde{w}, D].$$

Now we will take a closer look at the existence of minimal absorbing sets for δ-roundoffs of contractions. We begin with the following single-map-orbit shadowing lemma, which is a special case of Lemma III.14 in Sec. III:

Lemma II.6. *Let $\tilde{w} : \mathcal{D}^n(\delta) \to \mathcal{D}^n(\delta)$ be the δ-roundoff of a contraction $w : \mathbb{R}^n \to \mathbb{R}^n$ with respect to the metric d. Let $\{\tilde{w}^{\circ i}(\tilde{x})\}_{i=1}^{\infty}$ and $\{w^{\circ i}(\tilde{x})\}_{i=1}^{\infty}$ be orbits of an arbitrary point $\tilde{x} \in \mathcal{D}^n(\delta)$. Then*

$$d(\tilde{w}^{\circ i}(\tilde{x}), w^{\circ i}(\tilde{x})) \le \theta(1-\lambda)^{-1}, \ \forall i \in \mathbb{N}, \qquad \text{(II.2)}$$

where $\lambda \in [0, 1)$ is the contractivity factor of w.

The upper bound of the form $\theta(1-\lambda)^{-1}$ is symptomatic of δ-discretization, and we will encounter it many times in this paper. On the basis of Eq. (II.2) we see that the orbit $\{\tilde{w}^{\circ i}(\tilde{x}))\}$ is shadowed by the orbit $\{w^{\circ i}(\tilde{x})\}$, so we expect that, in the limit, the fixed point $x_f = w(x_f)$ will be imitated by a bounded orbit in $\mathcal{D}^n(\delta)$.

Theorem II.7. *Let $\tilde{w} : \mathcal{D}^n(\delta) \to \mathcal{D}^n(\delta)$ be the δ-roundoff of a contraction $w : \mathbb{R}^n \to \mathbb{R}^n$ with respect to the metric d. Let $x_f \in \mathbb{R}^n$ be the fixed point of w, and $\lambda \in [0, 1)$ be the map's contractivity factor. Let $\Lambda(x_f, r) := \overline{B}(x_f, r) \cap \mathcal{D}^n(\delta)$, where $\overline{B}(x_f, r)$ is the closed ball in $(\mathbb{R}^n, d)$, centred at x_f and with radius r. Then $\Lambda(x_f, r_0)$ with $r_0 = \theta(1-\lambda)^{-1}$, that is,*

$$\Lambda(x_f, r_0) = \{\tilde{y} \in \mathcal{D}^n(\delta) : d(\tilde{y}, x_f) \le \theta(1-\lambda)^{-1}\}, \qquad \text{(II.3)}$$

is an absorbing set for $\tilde{w}$. Moreover, $\tilde{w}$ maps $\Lambda(x_f, r_0)$ into itself,

$$\tilde{w}(\Lambda(x_f, r_0)) \subset \Lambda(x_f, r_0). \qquad \text{(II.4)}$$

Corollary II.8. *If $\tilde{w} : \mathcal{D}^n(\delta) \to \mathcal{D}^n(\delta)$ is the roundoff of a contraction $w : \mathbb{R}^n \to \mathbb{R}^n$, then the minimal absorbing set $\mathcal{M}[\tilde{w}]$ exists. Morover, $\mathcal{M}[\tilde{w}]$ is finite and its cardinality is bounded from above by the cardinality of $\Lambda(x_f, r_0)$.*

III. Discrete Iterated Function System

In this section we will study invariant sets and measures that arise from iterated function systems acting on a discrete space $\mathcal{D}^n(\delta)$. We will define a discrete IFS in a place-dependent probability version, that is, the probabilities assigned to IFS mappings are real-valued functions rather than constant numbers.

Definition III.1. *Let S be a subset of $\mathcal{D}^n(\delta)$ endowed with a metric d. Let $\{\tilde{w}_i\}_{i=1}^{N}$ be a finite set of maps $\tilde{w}_i : S \to S$. In addition, associate with each map $\tilde{w}_i$ a function $p_i : S \to (0, 1]$ of place-dependent, positive probability weights such that $\sum_{i=1}^{N} p_i(\tilde{x}) = 1$ for every $\tilde{x} \in S$. We will refer to the set $\{S; \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ as a* discrete iterated function system with place-dependent probabilities *(DIFS for short).*

Hereafter, unless otherwise stated, when speaking about a DIFS we will mean a discrete iteration function system with place-dependent probabilities, whereas an IFS will always appear in the constant probabilities and, moreover, hyperbolic version in this paper. Also note that in the definition above we do not impose any metric properties on the DIFS mappings, in particular we do not require them to be contractions. For now, they are just mappings on a subset S of a discrete space $\mathcal{D}^n(\delta)$, and the only metric properties they own are those that follow from the properties of S itself, namely, the mappings are continuous (with respect to any metric, since S is discrete) and if S was additionally bounded, they would be Lipschitz.

The random iteration algorithm (RIA) in the version for a discrete IFS with place-dependent probabilities works as follows: Choose a point $\tilde{x}_0 \in S$ and generate a random orbit $\{\tilde{x}_k = \tilde{X}_k(\tilde{x}_0)\}_{k=0}^{\infty}$ by recursively setting $\tilde{X}_k(\tilde{x}_0) := \tilde{w}_{I_k}(\tilde{X}_{k-1}(\tilde{x}_0))$, where $\tilde{X}_0(\tilde{x}_0) = \tilde{x}_0$ and each mapping index I_k is drawn from the probability distribution on $\{1, \ldots, N\}$, specified by the values of the N probability weight functions at the "place" $\tilde{x}_{k-1} = \tilde{X}_{k-1}(\tilde{x}_0)$, that is, the distribution is $[p_1(\tilde{x}_{k-1}), \ldots, p_N(\tilde{x}_{k-1})]$. In a special case when the probability functions p_i are constant over S, and hence the random variables I_k are i.i.d. with a distribution independent of place, the DIFS with place-dependent probabilities reduces to a discrete IFS with constant probabilities.

On the basis of above, RIA can be considered in terms of a collection of random variables $\{\tilde{X}_k : k \ge 0\}$ that forms a stochastic process on a countable set S, where $\tilde{X}_0$ is distributed with a point mass distribution concentrated at $\tilde{x}_0$ (i.e., the Dirac measure $\delta_{\tilde{x}_0}$). By the recursive definition of the random variables $\tilde{X}_k$ we get that, for any $k \ge 0$, $\tilde{X}_k$ depends only on $\tilde{X}_{k-1}$, that is, for any $\tilde{x}, \tilde{y} \in C$ and $k \ge 0$,

$$\begin{aligned}\Pr(\tilde{X}_k = \tilde{y} \mid \tilde{X}_{k-1} = \tilde{x}, \ldots, \tilde{X}_1, \tilde{X}_0) \\ = \Pr(\tilde{X}_k = \tilde{y} \mid \tilde{X}_{k-1} = \tilde{x}),\end{aligned}$$

and further that

$$\Pr(\tilde{X}_k = \tilde{y} \mid \tilde{X}_{k-1} = \tilde{x}) = \sum_{i=1}^{N} p_i(\tilde{x})\, \mathbf{1}_{\{\tilde{y}\}}(\tilde{w}_i(\tilde{x})), \qquad \text{(III.1)}$$

where $\mathbf{1}_A(.)$ denotes the indicator function of a set A. Therefore, the stochastic process is a (time-homogeneous) Markov chain with the transition matrix P whose entries $(P)_{\tilde{x}\tilde{y}} = \Pr(\tilde{X}_k = \tilde{y} \mid \tilde{X}_{k-1} = \tilde{x})$ are specified by Eq. (III.1).

In the following definition and theorem we gather some basic facts and notions pertaining to Markov chains that we will use further on (for more rigorous treatment see e.g. [10], [11]).

Definition III.2. *Let $\{\tilde{X}_k\}$ be a Markov chain on a countable set S, called the* state space *of the chain, and a transition matrix P.*
(a) *A state $\tilde{y}$ is said to be* accessible *from a state $\tilde{x}$ if there is a non-zero probability of reaching $\tilde{y}$ in a finite number of steps when starting from $\tilde{x}$:*

$$\exists k \in \mathbb{N},\ \Pr(\tilde{X}_k = \tilde{y} \mid \tilde{X}_0 = \tilde{x}) = (P^k)_{\tilde{x}\tilde{y}} > 0.$$

If $\tilde{x}$ and $\tilde{y}$ are accessible from each other, then they are said to communicate.
(b) *A Markov chain is called* irreducible *if any couple of its states communicate.*
(c) *A maximum subset C of S such that any two states in C communicate forms a* communication class. *A subset C is a* closed class *if the only states which are accessible from the states in C are those in C (i.e., if the chain starts in C, then it will stay in C with probability 1). Therefore, the Markov chain on any of its closed communication classes can be viewed as irreducible.*
(d) *A state $\tilde{x}$ is* recurrent *if, with probability 1, the chain starting in $\tilde{x}$ returns to $\tilde{x}$ in a finite number of steps. A recurrent state is* positive recurrent *if the expected number of steps to return is finite. A state that is not recurrent is called* transient.

Theorem III.1. *The state space S of a Markov chain uniquely decomposes into a countable union of disjoint subsets:*

$$S = \mathcal{T} \cup \mathcal{A}_1 \cup \mathcal{A}_2 \ldots$$

where $\mathcal{T}$ is the set of all transient states, and $\mathcal{A}_k$ are closed communication classes of the recurrent states. Moreover, if S is finite, then the union

$$\mathcal{A} = \bigcup_{k \le |S|} \mathcal{A}_k \tag{III.2}$$

is nonempty and consists of positive recurrent states. In addition, $\mathcal{A}$ is reached by the Markov chain in a finite number of steps with probability one and independently of the initial state.

Since all states in each set $\mathcal{A}_k$ are either recurrent or positive recurrent, the classes are called *recurrent communication classes* and *positive recurrent communication classes*, respectively. Similarly, a chain on a recurrent communication class and, respectively, positive recurrent communication class is called an *irreducible recurrent chain* and an *irreducible positive recurrent chain*, respectively.

The next theorem provides the view of closed and communication classes subjected to the action of the Hutchinson operator $\widetilde{W}(.) := \bigcup_{i=1}^N \tilde{w}_i(.)$ associated with a DIFS:

Theorem III.2. *Let $\{S; \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ be a DIFS, and let $\{\tilde{X}_k\}$ be the associated Markov chain.*
(a) *C is a closed class of $\{\tilde{X}_k\}$ if and only if $\bigcup_{i=1}^N \tilde{w}_i(C) \subset C$.*
(b) *If C is a communication class of $\{\tilde{X}_k\}$, then $\bigcup_{i=1}^N \tilde{w}_i(C) \supset C$.*

Theorem III.1 plays a similar role for DIFS as Banach's fixed point theorem for hyperbolic IFSs. In its first part it states that the presence of a recurrent state in a Markov chain implies that the chain possesses at least one set $\mathcal{A}$, which in light of Theorem III.2 is an invariant set of the DIFS Hutchinson operator. The second part provides a sufficient condition for the existence of a recurrent state, and thus also the set $\mathcal{A}$, and moreover it asserts that the set attracts orbits of the chain, in the sense that, with probability one, an orbit eventually visits and stays in $\mathcal{A}$, regardless of the orbit's intial point. Clearly, the properties of $\mathcal{A}$: the (forward) invariance under the DIFS Hutchison operator and the almost sure orbit attraction (if present) very much resemble the behavior of a hyperbolic IFS attractor. The key difference is that, in general, $\mathcal{A}$ decomposes into a countable collection of disjoint subsets $\mathcal{A}_k$, each of which is a fixed point of the DIFS Hutchinson operator and may act as an attractor within its own basin of attraction. This is a different situation from the one of a hyperbolic IFS in that, by Banach's fixed point theorem, the latter always possesses a single attractor being a unique fixed point of the associated Hutchison operator.

Remark III.1. *We have already had a glimpse of recurrent communication classes of Markov chains in Sec. II. Indeed, the components of a minimum absorbing set can be treated as recurrent communication classes of a trivial Markov chain that is generated by a single map (and hence the probability assigned to the map is 1) in a discrete space or its subset. Thereby, we have an example of a DIFS that consists of a single map and can possesses multiple attractors, even though—as we have seen in Sec. II—the map is a δ-roundoff of a contraction and, thus, the DIFS itself is a discretized version of a hyperbolic IFS.*

At this point, a natural question arises about upper bounds for the number of such invariant sets that a DIFS may possess.

Theorem III.3. *Let $\{S; \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ be a DIFS. Let $I \subset \{1, \ldots, N\}$ be the subset of the indices of the DIFS maps for which minimal absorbing sets exist. Suppose that I is nonempty and define $\mathcal{M}_{\mathcal{F}} := \{\mathcal{M}[\tilde{w}_i, S]\}_{i \in I}$. Let $\mathcal{A}_{\mathcal{F}} := \{\mathcal{A}_k\}$ be the family of the recurrent communication classes of the associated Markov chain. Then the following statements hold:*
(a) *The number of the sets in $\mathcal{A}_{\mathcal{F}}$ is not greater than the number of components of any set in $\mathcal{M}_{\mathcal{F}}$, that is,*

$$|\mathcal{A}_{\mathcal{F}}| \le \min_{i \in I} \mathcal{M}_{\#}[\tilde{w}_i, S],$$

where the subscript '$\#$' stands for the number of components of a minimal absorbing set.
(b) *If for some $i, j \in I$,*

$$\mathcal{M}[\tilde{w}_i, S] \subset \mathcal{B}[\mathcal{M}_k[\tilde{w}_j, S]]$$

for a certain $k \in \{1, \ldots, \mathcal{M}_{\#}[\tilde{w}_j, S]\}$, then

$$|\mathcal{A}_{\mathcal{F}}| \le 1.$$

A. On the existence of multiple attractors of a discretized hyperbolic IFS in practice

In [12] we showed that discretization of a hyperbolic IFS may lead to losing the contractive property characterizing the original and we also gave numerical evidence that from the statistical point of view, at least in the case of affine contractions on $(\mathbb{R}^2, d_E)$, such a situation can be considered typical and is not cured by increasing the precision of computation (expressed by the reciprocal of δ). As a consequence, even though we know that because of the fixed-point theorem the original hyperbolic IFS has an attractor and that this attractor is unique, we cannot use the theorem to show that the same facts hold for a discretized version of the IFS. Moreover, Theorem III.1 postulates, quite the opposite to Banach's theorem, that a discretized version of the hyperbolic IFS may possess more than one invariant set that attracts orbits generated by RIA, and Theorem III.3 (a) provides an upper bound for the number of the invariant sets in terms of the number of components of the minimal absorbing sets, which—as we know from Corollary II.8—δ-roundoffs of contractions always possess. On the other hand, the second part of Theorem III.3 gives a sufficient condition for a DIFS attractor (if it exists, and we show in the sequel that it is true) to be unique: *At least one of the minimal absorbing sets should be totally included in a basin of attraction of one of the remaining minimal absorbing sets.* Therefore conversely, a necessary condition for the DIFS to have more than one attractor is: *Each minimal absorbing set has to be intersected with at least two basins of attraction of every other minimal absorbing set.* Let us have a look at the possibility of the occurrence of such an event in practice.

Theorem II.7 gives us an upper bound on the diameter of a minimal absorbing set for a contraction and if we treat the value of $\mathrm{diam}_d(C_\delta)$ as a unit of length, it is clearly seen that the upper bound for the size of the minimal absorbing set depends only on the contractivity factor of the original contraction and remains constant when the precision of computation is changed. In this setting, altering the precision of computation affects only the unit of length in the way that the larger precision (or equivalently, the smaller δ), the smaller the unit of length. In effect, increasing the precision makes the extents of the minimal absorbing sets get smaller relative to the distances between the sets. At the same time, the basins of attractions spread radially from the minimal absorbing sets and their homogeneous regions get larger and larger as the distance from the minimal absorbing set increases (cf. the figures in the previous sections). Consequently, the higher precision of computation, the higher probability that the minimal absorbing set sits entirely in one of the basins of attraction of another minimal absorbing set. For resolutions of computation used in practice such as those offered by floating point arithmetic the minimal absorbing sets are usually so tiny relative to homogeneous regions of the basins of attraction that the configuration in which each of the sets is intersected by more than one basin of attraction of every other set is very special, not to say highly improbable. Therefore, even though a DIFS version of a hyperbolic IFS is usually not hyperbolic regardless of precision of computation in use but if the precision is reasonably high, then the DIFS not only has an attractor (we show it later on) but also this attractor is typically unique. Therefore the uniqueness of the DIFS attractor is not a result yielded by Banach's fixed point theorem (as it is sometimes suggested more or less explicitly in literature), but merely a typical resultant of the geometry of minimal absorbing sets and their basins of attraction when "embedded" in a suitable resolution of computation.

B. Invariant measure and stationary distribution

Now we consider the behavior of an orbit generated by RIA within a recurrent communication class. Suppose that a DIFS possesses at least one recurrent communication class $\mathcal{A}_k$ and assume that the associated Markov chain arrived in the set at a certain step while wandering through S or it just started in one of the points of the set. Because $\mathcal{A}_k$ is a closed communication class the orbit will never leave the set and the Markov chain on $\mathcal{A}_k$ is irreducible (Def. III.2 (c)). In the context of the problem of approximating the geometry of $\mathcal{A}_k$ it is crucial whether the orbit visits all points of the set during the evolution of the chain. It is not hard to show (see e.g. [13], p. 391) that if the chain is irreducible recurrent than for any state $\tilde{x} \in \mathcal{A}_k$,

$$\Pr(\exists i \in \mathbb{N} : \tilde{X}_i = \tilde{x}) = 1 \qquad \text{(III.3)}$$

regardless of the initial state $\tilde{X}_0 = \tilde{x}_0 \in \mathcal{A}_k$. In other words, starting from any arbitrary state in $\mathcal{A}_k$, the orbit is certain to pass through every other state in $\mathcal{A}_k$.

Moreover, the orbits of the Markov chain on a recurrent communication class possess some *average* statistical properties reflected in an *invariant measure* supported by the set. Writing P for the transition matrix of the chain restricted to a recurrent communication class $\mathcal{A}_k$ (an irreducible recurrent chain), an invariant measure π can be represented by a row vector $[\pi] = [(\pi)_{\tilde{x}}]_{\tilde{x} \in \mathcal{A}_k}$ of the values $(\pi)_{\tilde{x}} = \pi\{\tilde{x}\}$, which satisfies the equation $[\pi] = [\pi]P$. It can be shown (see [10], [13], also [14] for the treatment from the standpoint of Perron-Frobenius theory) that such a vector for a recurrent communication class always exists, its entries are positive and finite, and equal to the *expected* number of visits the chain makes to states $\tilde{x}$ while returning to a state $\tilde{y} \in \mathcal{A}_k$. Because of the relationship between the values of π and a given baseline state $\tilde{y}$, there can be, in general, more than one invariant measure for a recurrent communication class, however the measures are equal up to multiplication by a constant. Another thing is that the value of the invariant measures on a infinite recurrent communication set can be infinite. But if an invariant measure π is finite, then after normalization becomes a probability measure referred to as a *stationary distribution*. It is known from from the Markov chain theory that the normalization of an invariant measure on a recurrent communication class is possible if and only if the class is *positive* recurrent. (As asserted in the second part of Theorem III.1, this positive recurrence condition is always satisfied for finite closed communication classes.) The values of the stationary distribution entries $(\pi)_{\tilde{x}}$ are reciprocals of the expected number of steps to return, they are finite and characterize each positive recurrent state $\tilde{x}$ (Def. III.2 (d)). From this follows that a stationary distribution of a Markov chain restricted to a positive recurrent communication class must be unique. Clearly, if for a certain $i \geq 0$, the distribution of $\tilde{X}_i$ of a chain $\{\tilde{X}_k : k \geq 0\}$ is a stationary distribution π, then the distributions of the random variables that follow in

$\{\tilde{X}_k \ : k \geq i\}$ do not change, they are identical and equal to π—the process is said to be in a steady (or equilibrium) state.

C. To be attractive (or at least ergodic)

It is known (see [8], Theorem 2.1; also [6], [7]) that if an IFS is hyperbolic on $(\mathbb{R}^n, d)$ or even only contractive on average (with some additional constraints on place dependent p_i's such as Dini's continuity) the stationary distribution is unique and attractive in the sense that regardless of the initial distribution of X_0, the distributions μ_k of the random variables in the associated chain $\{X_k : k \geq 0\}$ converges weakly to the stationary distribution π_∞, that is, $\mathbf{E}[f(X_k)] \to \mathbf{E}[f(X)] = \int_{\mathbb{R}^n} f(x) d\pi_\infty$ for any bounded continuous function $f : \mathbb{R}^n \to \mathbb{R}$, where X is the limiting (in distribution) random variable distributed as π_∞. As a consequence, no matter what an initial distribution, the Markov process generated by RIA heads, with probability 1, for the steady state expressed by the attractive distribution π_∞, that is the IFS invariant measure. This fact forms the basis for Elton's ergodic theorem (see [9], Corollary 1), which states that for any $x \in \mathbb{R}^n$ and f defined above, the space average $\mathbf{E}[f(X)]$ is equal with probability 1 to the time average of $\{f(X_k) : X_0 = x, k \geq 0\}$, that is,

$$\lim_{m \to \infty} \frac{1}{m} \sum_{k=0}^{m-1} f(X_k) = \int_{\mathbb{R}^n} f(x) d\pi_\infty \ \ a.s. \qquad \text{(III.4)}$$

In other words, empirical probability measures $\mu_m = \frac{1}{m} \sum_{k=0}^{m-1} \delta_{x_k}$ of almost every orbit of the chain $\{X_k\}$ converge weakly to π_∞.

In effect, given a bounded continuous function $f : \mathbb{R}^n \to \mathbb{R}$, one can determine the right-hand side integral in the equation above by plugging an orbit generated by RIA into the left-hand side of the equation, and we are guaranteed that the result will be correct with probability one. In particular, given a Borel subset A, one can approximate the value of $\pi_\infty(A)$ by $\mu_m(A)$ for a certain large m (under an additional technical requirement of the zero π_∞-measure of A's boundary), which technically is equivalent to counting and then normalizing the number of points that pop into A in a finite-time realization (m steps) of the algorithm. Such an approach leads to a popular method for visualizing the IFS invariant measure on a discrete grid of pixels. Below we investigate the possible effects of the application of RIA in its variant for visualizing invariant measures when the algorithm is applied in the context of a DIFS without any metric properties imposed on the mappings.

As pointed out in Sec. III-B, a stationary distribution of an *irreducible* Markov chain is strictly positive and exists if and only if the chain is positive recurrent. This stationary distribution extends trivially to a stationary probability measure of any *reducible* Markov chain whose state space includes the recurrent positive state space of this smaller irreducible chain as one of its closed communication classes (the measure of the complement of the closed communication class is zero). A positive recurrent communication class exists if the chain generated by a DIFS possesses a positive recurrent state. A positive recurrent state exists, first of all, if the state space of the associated chain is, or can be appropriately restricted to a finite set (cf. the second part of Theorem III.1—we will show that this holds for DIFSs arising from hyperbolic IFSs) or if some other, more general criteria for positive recurrence hold (e.g., the existence of a closed class satisfying Doeblin's criterion [11] or Foster's criterion [10]). Now, suppose that the associated chain has more than one class $\mathcal{A}_k$, $k \in I$, each being positive recurrent, so the chain has at least $|I| \in \mathbb{N}$ stationary distributions. But if π_k's are stationary distributions of the chain, then any convex combination of these distributions is also a stationary distribution, because

$$\Big(\sum_{k \in I} \alpha_k [\pi_k]\Big) P = \sum_{k \in I} \alpha_k ([\pi_k] P) = \sum_{k \in I} \alpha_k [\pi_k], \qquad \text{(III.5)}$$

where P is the transition matrix of the chain, and $\sum_{k \in I} \alpha_k = 1$, $\alpha_k \geq 0$.

Let us summarize our considerations on the DIFS stationary measures with the following:

Corollary III.4. *Let $\{S; \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ be a DIFS, and let $\mathcal{A}_\mathcal{F}$ be the family of the recurrent communication classes of the associated Markov chain. Denote by $\mathcal{A}_\mathcal{F}^+$ the positive recurrent subfamily of $\mathcal{A}_\mathcal{F}$. Then:*
(a) *The Markov chain has a stationary probability measure if $|\mathcal{A}_\mathcal{F}^+| \geq 1$.*
(b) *if $|\mathcal{A}_\mathcal{F}^+| = 1$, then the stationary probability measure is unique and supported by the single positive recurrent communication class in $\mathcal{A}_\mathcal{F}^+$.*
(c) *If $|\mathcal{A}_\mathcal{F}^+| > 1$, then the chain has uncountable number of stationary probability measures of the form (III.5), that is, the stationary distributions are convex combinations of the stationary probability measures on the sets in $\mathcal{A}_\mathcal{F}^+$ and supported by their unions.*

Obviously, by the definition of an attractive distribution, a stationary distribution can be globally attractive only if the stationary distribution is unique. Therefore, a necessary condition for a stationary probability measure π supported by the sets in $\mathcal{A}_\mathcal{F}^+$ to be the globally attractive distribution is $|\mathcal{A}_\mathcal{F}^+| = 1$, that is, there has to be exactly one positive recurrent communication class, say $\mathcal{A}_1^+$, in the state space of the chain. However, in order to provide weak convergence of probability measures to π independently of the initial distribution, in addition there cannot be other closed recurrent and closed transient classes in the state space but $\mathcal{A}_1^+$ and, moreover, the probability for the chain to escape from the transient states $\mathcal{T}$ has to be one. As long as the latter "escape-from-transient-class" condition is always met when the space upon which a DIFS acts is, or can be rightly restricted to a finite set, we usually cannot ensure that there will be only one recurrent communication class in the state space even in the case of DIFSs arising from hyperbolic IFSs. Therefore, even if a DIFS is a discretized version of hyperbolic IFS, it can have uncountable many stationary distributions (and thus no globally attractive distribution), and hence DIFSs are not in general contractive even on average (because in the latter case the stationary distribution is unique).

Nonetheless, given a DIFS such that the associated chain leaves the transient class $\mathcal{T}$ with probability one, that is, $\Pr(\exists i \in \mathbb{N} : \tilde{X}_i \notin \mathcal{T}) = 1$ independently of $\tilde{X}_0 \in \mathcal{T}$, we immediately get that the recurrent family $\mathcal{A}_\mathcal{F}$ is nonempty (since $\mathcal{T}$ is not a closed class, so there must be at least one recurrent state) and thus the orbit generated by RIA is guaranteed (almost surely) to reach one of the sets $\mathcal{A}_k$ in $\mathcal{A}_\mathcal{F}$. Since the sets are closed classes, the orbit will never leave such a set after it gets to the set. Moreover, due to Eq. (III.3) it will visit all points $\tilde{x}$ in $\mathcal{A}_k$. Yet the orbit is not guaranteed to stabilize according to a stationary distribution even if the set is positive recurrent and, hence, the irreducible Markov chain on the set possesses a stationary distribution. The reason is that in order for a stationary distribution to be attractive for distributions on a positive recurrent communication class, the class should additionally contain an *aperiodic* state, a state for which the greatest common divisor of possible numbers of steps to return (periods) is 1 (which would imply that all states of the class were aperiodic too). It is also worth noting that in contrast to positive recurrence, aperiodicity is not guaranteed by finiteness of a state space. Therefore, even if a domain on which a DIFS acts is finite, there is no guarantee that the DIFS has a stationary probability measure that is attractive at the very least locally for distributions on a positive recurrent communication class. Anyway, any positive recurrent Markov chain is ergodic (see remark below), which means that despite the fact the associated chain within a positive recurrent communication class $\mathcal{A}_k^+$ does not necessarily reach a stationary distribution, the orbit generated by RIA can still be used to determine the value of the counterpart of the integral on the right-hand side of Eq. (III.4). In particular, if π is a stationary distribution on $\mathcal{A}_k^+$, then for any state $\tilde{y} \in \mathcal{A}_k^+$,

$$\lim_{m\to\infty} \frac{1}{m} \sum_{i=0}^{m-1} \mathbf{1}_{\{\tilde{y}\}}(\tilde{X}_i) = \pi\{\tilde{y}\} \;\; a.s.$$

provided that $\Pr(\tilde{X}_0 \in \mathcal{A}_k^+) = 1$. Hence, for any $\tilde{y} \in \mathcal{A}_k^+$,

$$\Pr\left(\lim_{m\to\infty} \frac{1}{m} \sum_{i=0}^{m-1} \mathbf{1}_{\{\tilde{y}\}}(\tilde{X}_i) = \pi\{\tilde{y}\} | \tilde{X}_j \in \mathcal{A}_k^+ \right) = 1,$$

provided that there is a certain $j \in \mathbb{N}$ such that $\Pr(\tilde{X}_j \in \mathcal{A}_k^+) > 0$. As a consequence, if the orbit generated by RIA enters $\mathcal{A}_k^+$, one can use the orbit to render an image of the stationary distribution supported by the set. On the other hand, if the orbit does not enter a positive recurrent communication class, that is, it arrives either in $\mathcal{A}_k$ that is not positive recurrent or it stays in a transient class, then for any $\tilde{x} \in S$,

$$\lim_{m\to\infty} \frac{1}{m} \sum_{k=0}^{m-1} \mathbf{1}_{\{\tilde{x}\}}(X_k) = 0 \;\; a.s.$$

no matter what the initial distribution (see [11], pp. 72–74). In such a case, an attempt to visualize a measure with RIA would yield a gradually vanishing image of a measure on points in $S \subset \mathcal{D}^n(\delta)$.

Remark III.2 (On egodicity of Markov chains)**.** *It is common in the literature on countable Markov chains that an ergodic chain is defined as an irreducible chain that is both positive recurrent and aperiodic, which is a sufficient condition for the chain to have an attractive (limiting) and, thus, unique stationary distribution. Such a chain is also ergodic in the sense of ergodic theory by some standard arguments. However, this usage of the term 'ergodic' in this context is somewhat misleading because it suggests that aperiodicity is necessary for a Markov chain to be ergodic in the sense of ergodic theory, whereas—as we pointed out earlier—every irreducible positive recurrent chain is ergodic (see e.g. [11], also a remark on Markov ergodicity by Elton in [9]). An instructive instance of an irreducible positive recurrent chain that is not aperiodic, but still ergodic is a 2-state chain with the transition matrix $\left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. The chain is positive recurrent and periodic with period 2, it has a unique stationary distribution $[\frac{1}{2}, \frac{1}{2}]$, but the distribution is not attractive. Nevertheless, it is clearly seen that the chain is ergodic.*

Corollary III.5. *Let $\{S; \tilde{w}_1, \dots, \tilde{w}_N; p_1, \dots, p_N\}$ be a DIFS such that the associated Markov chain $\{\tilde{X}_i : i \geq 0\}$ satisfies $\Pr(\exists i \in \mathbb{N} : \tilde{X}_i \notin \mathcal{T}) = 1$ for any $\tilde{X}_0 = \tilde{x} \in S$. Let $\{\tilde{x}_i\}_{i=0}^\infty$ be an orbit generated by RIA. Then, for a certain $m \in \mathbb{N}$, with probability one,*

$$\{\tilde{x}_i\}_{i=m}^\infty = \mathcal{A}_k,$$

where $\mathcal{A}_k \in \mathcal{A}_\mathcal{F}$, that is, the set is one of the recurrent communication classes of $\{\tilde{X}_i\}$.

In addition, if the set is positive recurrent, $\mathcal{A}_k \in \mathcal{A}_\mathcal{F}^+$, and $\Pr(\tilde{X}_j \in \mathcal{A}_k) = 1$ for a certain $j \in \mathbb{N}$, then, with probability one, for any $\tilde{y} \in S$, the ratio of the points in $\{\tilde{x}_i\}_{i=0}^M$ that coincide with $\tilde{y}$ to their total number M converges to the value of $\pi_k\{\tilde{y}\}$ as $M \to \infty$, where π_k is the stationary probability measure of the chain on $\mathcal{A}_k$; that is

$$\Pr\left(\lim_{M\to\infty} \frac{1}{M} \sum_{i=0}^{M-1} \delta_{\tilde{X}_i} = \pi_k | \tilde{X}_j \in \mathcal{A}_k \right) = 1$$

and

$$\Pr\left(\lim_{M\to\infty} \frac{1}{M} \sum_{i=0}^{M-1} \mathbf{1}_{\{\tilde{y}\}}(\tilde{X}_i) = \pi_k\{\tilde{y}\} | \tilde{X}_j \in \mathcal{A}_k \right) = 1,$$

if only $\Pr(\tilde{X}_j \in \mathcal{A}_k) > 0$ for a certain $j \in \mathbb{N}$.

In particular, both statements of the corollary hold if S is a finite set.

D. DIFSs for hyperbolic IFSs

In this section we examine the case of DIFSs consisting of maps being δ-roundoffs of contractions, in particular, DIFSs that are discretized versions of hyperbolic IFSs with constant probabilities. We start with the following theorem that establishes important facts concerning a DIFS composed of δ-roundoffs of contractions.

Theorem III.6. *Let $\{\mathcal{D}^n(\delta); \tilde{w}_1, \dots, \tilde{w}_N; p_1, \dots, p_N\}$ be a DIFS such that the mappings $\tilde{w}_i$ are δ-roundoffs of contractions w_i on $(\mathbb{R}^n, d)$. Let o be any point of $\mathbb{R}^n$. Denote $r_{max} := \max_i d(x_f^{(i)}, o)$, $\lambda_{max} := \max_i \lambda_i$ and $\alpha := \frac{1+\lambda_{max}}{1-\lambda_{max}}$, where $x_f^{(i)}$ and λ_i are the fixed points and, respectively, the contractivity factors of the mappings w_i. Then, for any $\varepsilon > 0$ and $S = B(o, r + \varepsilon) \cap \mathcal{D}^n(\delta)$, where $B(o, r)$ is an open ball in $(\mathbb{R}^n, d)$, centred at o and with radius $r = \alpha r_{max} + \theta(1 - \lambda_{max})^{-1}$, the DIFS transformations map*

S into itself, $\tilde{w}_i(S) \subset S$ for every $i \in \{1, \ldots, N\}$. Moreover, the DIFS possesses nonempty $\mathcal{A}$, the set of all recurrent states of the associated Markov chain, and $\mathcal{A} \subset S$.

Remark III.3. *One can easily check that if $\{\mathbb{R}^n; w_1, \ldots, w_N; p_1, \ldots, p_N\}$ is a hyperbolic IFS, then for any $\varepsilon > 0$ and $S = B(o, r + \varepsilon)$, where $r = \lim_{\delta \to 0} \alpha r_{max} + \theta(1 - \lambda_{max})^{-1} = \alpha r_{max}$, the IFS transformations map S into itself, $w_i(S) \subset S$ for every $i \in \{1, \ldots, N\}$. Hence, the space upon which the IFS acts can be restricted to any of the compact sets (closed balls) $\overline{S}$, and the IFS attractor $A_\infty \subset \overline{S}$.*

Corollary III.7. *If $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ is a DIFS in which the mappings $\tilde{w}_i$ are δ-roundoffs of contractions, then $\mathcal{A}$ is nonempty and finite and hence consists of positive recurrent states; thus $\mathcal{A}_\mathcal{F} = \mathcal{A}^+_\mathcal{F} \neq \varnothing$. Therefore, the DIFS possesses stationary probability measures supported by unions of the finite positive recurrent communication classes $\mathcal{A}^+_k$ of which $\mathcal{A}$ is composed. Moreover, for every $\tilde{X}_0 = \tilde{x} \in \mathcal{D}^n(\delta)$, $\Pr(\exists i \in \mathbb{N} : \tilde{X}_i \notin \mathcal{T}) = 1$. Therefore, by Corollary III.5, each run of RIA results, with probability one, in rendering a stationary probability measure supported by one of the classes $\mathcal{A}^+_k$. The stationary probability measure is unique if for a certain $i \in \{1, \ldots, N\}$, the minimal absorbing set for $\tilde{w}_i$ consists of a single component, or there are $i, j \in \{1, \ldots, N\}$, such that $\mathcal{M}[\tilde{w}_i, S] \subset \mathcal{B}[\mathcal{M}_k[\tilde{w}_j, S]]$ for a certain component $\mathcal{M}_k[\tilde{w}_j]$ of the minimal absorbing set $\mathcal{M}_k[\tilde{w}_j]$ (Theorem III.3).*

The lemma below establishes a connection between a hyperbolic IFS and a DIFS composed of δ-roundoffs of the IFS mapping, in the form of an assertion on mutual shadowing of orbits of Markov chains generated by the DIFS and the corresponding IFS.

Lemma III.8. *Let $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ be a DIFS in which $\tilde{w}_i$'s are δ-roundoffs of contractions w_i on $(\mathbb{R}^n, d)$. Let $\{\mathbb{R}^n; w_1, \ldots, w_n; q_1, \ldots, q_N\}$ be an IFS with strictly positive probabilities $q_i \in (0, 1]$, $\sum_{i=1}^N q_i = 1$. Let $\tilde{x}_0$ be any point of $\mathcal{D}^n(\delta)$. Let $X = \{X_k : X_0 = \tilde{x}_0\}$ and $\tilde{X} = \{\tilde{X}_k : \tilde{X}_0 = \tilde{x}_0\}$ be the chains generated by the IFS and the DIFS, respectively. Then for any orbit $\{x_k = w_{i_k}(x_{k-1})\}$ of X (respectively, any orbit $\{\tilde{x}_k = \tilde{w}_{i_k}(\tilde{x}_{k-1})\}$ of $\tilde{X}$), there exists an orbit $\{\tilde{x}_k = \tilde{w}_{i_k}(\tilde{x}_{k-1})\}$ of $\tilde{X}$ (respectively, an orbit $\{x_k = w_{i_k}(x_{k-1})\}$ of X) such that, for any $k \in \mathbb{N}$,*

$$d(x_k, \tilde{x}_k) \leq \theta(1 - \lambda_{max})^{-1}, \tag{III.6}$$

where $\lambda_{max} = \max_i \lambda_i$, λ_i is the contractivity factor of w_i.

In turn, the next theorem answers the question about geometrical resemblance between DIFS invariant sets $\mathcal{A}^+_k \in \mathcal{A}_\mathcal{F}$ and the attractor A_∞ of a hyperbolic IFS, expressed in terms of the Hausdorff distance.

Theorem III.9. *Let $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ be a DIFS in which $\tilde{w}_i$'s are δ-roundoffs of contractions w_i on $(\mathbb{R}^n, d)$, and let $\{\mathbb{R}^n; w_1, \ldots, w_n; q_1, \ldots, q_N\}$ be a corresponding hyperbolic IFS with strictly positive probabilities q_i. Let $\mathcal{A}_\mathcal{F}$ be the family of all positive recurrent classes of the Markov chain associated with the DIFS, and let A_∞ be the attractor of the IFS. Then, independently of the values of probabilities $p_i(.)$ and q_i, for each $\mathcal{A}^+_k \in \mathcal{A}_\mathcal{F}$, the Hausdorff distance between $\mathcal{A}^+_k$ and A_∞ is bounded from above as*

$$h(\mathcal{A}^+_k, A_\infty) \leq \theta(1 - \lambda_{max})^{-1}. \tag{III.7}$$

Corollary III.10. *Let $\{\mathbb{R}^n; w_1, \ldots, w_n; p_1, \ldots, p_N\}$ be a hyperbolic IFS with the attractor A_∞. Let $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ be the corresponding DIFSs (with the same constant probabilities as in the IFS), parametrized by $\delta > 0$, in which $\tilde{w}_i$'s are δ-roundoffs of contractions w_i. Let $\mathcal{A}_\mathcal{F}(\delta) = \{\mathcal{A}^+_k(\delta)\}_k$ be the family of all positive recurrent classes of the Markov chain associated with a DIFS for fixed δ. Then*

$$\lim_{\delta \to 0} \mathcal{A}^+_k(\delta) = A_\infty \text{ (in the Hausdorff metric)},$$

where $\mathcal{A}^+_k(\delta)$ is any set from $\mathcal{A}_\mathcal{F}(\delta)$ for fixed δ.

We also have a theorem concerning a relationship between DIFS and IFS measures:

Theorem III.11. *Let $\{\mathbb{R}^n; w_1, \ldots, w_n; p_1, \ldots, p_N\}$ be a hyperbolic IFS, and let π_∞ be the IFS invariant measure. Let $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ be the corresponding DIFSs (with the same constant probabilities as in the IFS), parametrized by $\delta > 0$, in which $\tilde{w}_i$'s are δ-roundoffs of contractions w_i. Let $\mathcal{A}_\mathcal{F}(\delta)$ be the family of all positive recurrent classes of the Markov chain associated with a DIFS for fixed δ, and $\Pi(\delta) = \{\pi_k(\delta) : \operatorname{supp}(\pi_k(\delta)) \in \mathcal{A}_\mathcal{F}(\delta)\}$ be the family of the chain's stationary distributions supported by the sets in $\mathcal{A}_\mathcal{F}(\delta)$. Then for any continuous and bounded $f : \mathbb{R}^n \to \mathbb{R}$,*

$$\lim_{\delta \to 0} \sum_{\tilde{x} \in \mathcal{D}^n(\delta)} f(\tilde{x}) \pi_k(\delta)\{\tilde{x}\} = \int_{\mathbb{R}^n} f(x) d\pi_\infty, \tag{III.8}$$

where $\pi_k(\delta)$ is any stationary distribution from $\Pi(\delta)$ for fixed δ. In other words, $\pi_k(\delta)$'s converge weakly to π_∞ as $\delta \to 0$.

E. DIFSs for hyperbolic IFSs

In this section we examine the case of DIFSs consisting of maps being δ-roundoffs of contractions, in particular, DIFSs that are discretized versions of hyperbolic IFSs with constant probabilities. We start with the following theorem that establishes important facts concerning a DIFS composed of δ-roundoffs of contractions.

Theorem III.12. *Let $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ be a DIFS such that the mappings $\tilde{w}_i$ are δ-roundoffs of contractions w_i on $(\mathbb{R}^n, d)$. Let o be any point of $\mathbb{R}^n$. Denote $r_{max} := \max_i d(x_f^{(i)}, o)$, $\lambda_{max} := \max_i \lambda_i$ and $\alpha := \frac{1+\lambda_{max}}{1-\lambda_{max}}$, where $x_f^{(i)}$ and λ_i are the fixed points and, respectively, the contractivity factors of the mappings w_i. Then, for any $\varepsilon > 0$ and $S = B(o, r + \varepsilon) \cap \mathcal{D}^n(\delta)$, where $B(o, r)$ is an open ball in $(\mathbb{R}^n, d)$, centred at o and with radius $r = \alpha r_{max} + \theta(1 - \lambda_{max})^{-1}$, the DIFS transformations map S into itself, $\tilde{w}_i(S) \subset S$ for every $i \in \{1, \ldots, N\}$. Moreover, the DIFS possesses nonempty $\mathcal{A}$, the set of all recurrent states of the associated Markov chain, and $\mathcal{A} \subset S$.*

Remark III.4. *One can easily check that if* $\{\mathbb{R}^n; w_1, \ldots, w_N; p_1, \ldots, p_N\}$ *is a hyperbolic IFS, then for any* $\varepsilon > 0$ *and* $S = B(o, r + \varepsilon)$, *where* $r = \lim_{\delta\to 0} \alpha r_{max} + \theta(1 - \lambda_{max})^{-1} = \alpha r_{max}$, *the IFS transformations map* S *into itself,* $w_i(S) \subset S$ *for every* $i \in \{1, \ldots, N\}$. *Hence, the space upon which the IFS acts can be restricted to any of the compact sets (closed balls)* $\overline{S}$, *and the IFS attractor* $A_\infty \subset \overline{S}$.

Corollary III.13. *If* $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ *is a DIFS in which the mappings* $\tilde{w}_i$ *are* δ*-roundoffs of contractions, then* $\mathcal{A}$ *is nonempty and finite and hence consists of positive recurrent states; thus* $\mathcal{A}_\mathcal{F} = \mathcal{A}^+_\mathcal{F} \neq \varnothing$. *Therefore, the DIFS possesses stationary probability measures supported by unions of the finite positive recurrent communication classes* $\mathcal{A}^+_k$ *of which* $\mathcal{A}$ *is composed. Moreover, for every* $\tilde{X}_0 = \tilde{x} \in \mathcal{D}^n(\delta)$, $\Pr(\exists i \in \mathbb{N} : \tilde{X}_i \notin \mathcal{T}) = 1$. *Therefore, by Corollary III.5, each run of RIA results, with probability one, in rendering a stationary probability measure supported by one of the classes* $\mathcal{A}^+_k$. *The stationary probability measure is unique if for a certain* $i \in \{1, \ldots, N\}$, *the minimal absorbing set for* $\tilde{w}_i$ *consists of a single component, or there are* $i, j \in \{1, \ldots, N\}$, *such that* $\mathcal{M}[\tilde{w}_i, S] \subset \mathcal{B}[\mathcal{M}_k[\tilde{w}_j, S]]$ *for a certain component* $\mathcal{M}_k[\tilde{w}_j]$ *of the minimal absorbing set* $\mathcal{M}_k[\tilde{w}_j]$ *(Theorem III.3).*

The lemma below establishes a connection between a hyperbolic IFS and a DIFS composed of δ-roundoffs of the IFS mapping, in the form of an assertion on mutual shadowing of orbits of Markov chains generated by the DIFS and the corresponding IFS.

Lemma III.14. *Let* $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ *be a DIFS in which* $\tilde{w}_i$*'s are* δ*-roundoffs of contractions* w_i *on* $(\mathbb{R}^n, d)$. *Let* $\{\mathbb{R}^n; w_1, \ldots, w_n; q_1, \ldots, q_N\}$ *be an IFS with strictly positive probabilities* $q_i \in (0, 1]$, $\sum_{i=1}^N q_i = 1$. *Let* $\tilde{x}_0$ *be any point of* $\mathcal{D}^n(\delta)$. *Let* $X = \{X_k : X_0 = \tilde{x}_0\}$ *and* $\tilde{X} = \{\tilde{X}_k : \tilde{X}_0 = \tilde{x}_0\}$ *be the chains generated by the IFS and the DIFS, respectively. Then for any orbit* $\{x_k = w_{i_k}(x_{k-1})\}$ *of* X *(respectively, any orbit* $\{\tilde{x}_k = \tilde{w}_{i_k}(\tilde{x}_{k-1})\}$ *of* $\tilde{X}$*), there exists an orbit* $\{\tilde{x}_k = \tilde{w}_{i_k}(\tilde{x}_{k-1})\}$ *of* $\tilde{X}$ *(respectively, an orbit* $\{x_k = w_{i_k}(x_{k-1})\}$ *of* X*) such that, for any* $k \in \mathbb{N}$,

$$d(x_k, \tilde{x}_k) \leq \theta(1 - \lambda_{max})^{-1}, \qquad \text{(III.9)}$$

where $\lambda_{max} = \max_i \lambda_i$, λ_i *is the contractivity factor of* w_i.

In turn, the next theorem answers the question about geometrical resemblance between DIFS invariant sets $\mathcal{A}^+_k \in \mathcal{A}_\mathcal{F}$ and the attractor A_∞ of a hyperbolic IFS, expressed in terms of the Hausdorff distance.

Theorem III.15. *Let* $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ *be a DIFS in which* $\tilde{w}_i$*'s are* δ*-roundoffs of contractions* w_i *on* $(\mathbb{R}^n, d)$, *and let* $\{\mathbb{R}^n; w_1, \ldots, w_n; q_1, \ldots, q_N\}$ *be a corresponding hyperbolic IFS with strictly positive probabilities* q_i. *Let* $\mathcal{A}_\mathcal{F}$ *be the family of all positive recurrent classes of the Markov chain associated with the DIFS, and let* A_∞ *be the attractor of the IFS. Then, independently of the values of probabilities* $p_i(.)$ *and* q_i, *for each* $\mathcal{A}^+_k \in \mathcal{A}_\mathcal{F}$, *the Hausdorff distance between* $\mathcal{A}^+_k$ *and* A_∞ *is bounded from above as*

$$h(\mathcal{A}^+_k, A_\infty) \leq \theta(1 - \lambda_{max})^{-1}. \qquad \text{(III.10)}$$

Corollary III.16. *Let* $\{\mathbb{R}^n; w_1, \ldots, w_n; p_1, \ldots, p_N\}$ *be a hyperbolic IFS with the attractor* A_∞. *Let* $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ *be the corresponding DIFSs (with the same constant probabilities as in the IFS), parametrized by* $\delta > 0$, *in which* $\tilde{w}_i$*'s are* δ*-roundoffs of contractions* w_i. *Let* $\mathcal{A}_\mathcal{F}(\delta) = \{\mathcal{A}^+_k(\delta)\}_k$ *be the family of all positive recurrent classes of the Markov chain associated with a DIFS for fixed* δ. *Then*

$$\lim_{\delta\to 0} \mathcal{A}^+_k(\delta) = A_\infty \ \text{(in the Hausdorff metric)},$$

where $\mathcal{A}^+_k(\delta)$ *is any set from* $\mathcal{A}_\mathcal{F}(\delta)$ *for fixed* δ.

We also have a theorem concerning a relationship between DIFS and IFS measures:

Theorem III.17. *Let* $\{\mathbb{R}^n; w_1, \ldots, w_n; p_1, \ldots, p_N\}$ *be a hyperbolic IFS, and let* π_∞ *be the IFS invariant measure. Let* $\{\mathcal{D}^n(\delta); \tilde{w}_1, \ldots, \tilde{w}_N; p_1, \ldots, p_N\}$ *be the corresponding DIFSs (with the same constant probabilities as in the IFS), parametrized by* $\delta > 0$, *in which* $\tilde{w}_i$*'s are* δ*-roundoffs of contractions* w_i. *Let* $\mathcal{A}_\mathcal{F}(\delta)$ *be the family of all positive recurrent classes of the Markov chain associated with a DIFS for fixed* δ, *and* $\Pi(\delta) = \{\pi_k(\delta) : \operatorname{supp}(\pi_k(\delta)) \in \mathcal{A}_\mathcal{F}(\delta)\}$ *be the family of the chain's stationary distributions supported by the sets in* $\mathcal{A}_\mathcal{F}(\delta)$. *Then for any continuous and bounded* $f : \mathbb{R}^n \to \mathbb{R}$,

$$\lim_{\delta\to 0} \sum_{\tilde{x}\in\mathcal{D}^n(\delta)} f(\tilde{x})\pi_k(\delta)\{\tilde{x}\} = \int_{\mathbb{R}^n} f(x) d\pi_\infty, \qquad \text{(III.11)}$$

where $\pi_k(\delta)$ *is any stationary distribution from* $\Pi(\delta)$ *for fixed* δ. *In other words,* $\pi_k(\delta)$*'s converge weakly to* π_∞ *as* $\delta \to 0$.

References

[1] M.F. Barnsley, *Fractals everywhere, 2nd ed.*, Boston Academic Press, 1993.

[2] M. Peruggia, *Discrete Iterated Function Systems*. A. K. Peters Wellesley MA, 1993.

[3] W.H. Press, S.A. Teukolsky, W.T. Vetterling, and B.P. Flannery, *Numerical Recipes in C (2nd Ed.)*, Cambridge University Press, 1992.

[4] J. Harrison, "A machine-checked theory of floating point arithmetic". In: Bertot Y, Dowek G, Théry L, Hirschowitz A, Paulin C (eds). *Theorem Proving in Higher Order Logics. TPHOLs 1999. Lecture Notes in Computer Science*, vol. 1690. Springer, Berlin, Heidelberg, 1999, pp. 113–130. .

[5] J-M. Muller, "On the definition of ulp(x)", *Research Report*, INRIA, LIP, 2005.

[6] P. Diaconis and D. Freedman, "Iterated Random Functions", *SIAM Rev.*, vol. 41, pp. 45–76, 1999.

[7] Ö. Stenflo, "Uniqueness of invariant measures for place-dependent random iterations of functions", *IMA Vol. Math. Appl.*, vol. 132, pp. 13–32, 2002.

[8] 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.*, vol. 24, pp. 367–394, 1988.

[9] J.H. Elton, "An ergodic theorem for iterated maps", *Ergodic Theory and Dynam. Systems*, vol. 7, pp. 481–488, 1987.

[10] R.F. Serfozo, *Basics of Applied Stochastic Processes*, Springer-Verlag New York-Berlin, 2009.

[11] D.W. Stroock DW, *An Introduction to Markov Processes*, Springer-Verlag New York-Berlin, 2005.

[12] T. Martyn, "The discrete charm of iterated function systems". [Online]. Available: https://doi.org/10.48550/arXiv.2410.15139

[13] W. Feller, *An Introduction to Probability Theory and Its Applications (3th Ed.), Volume I*, John Wiley & Sons, 1968.

[14] C. Meyer, *Matrix analysis and applied linear algebra*, SIAM, 2000.