A Chaos Game Algorithm for Generalized Iterated Function Systems

D. La Torre* and F. Mendivil†

Abstract

In this paper we provide an extension of the classical Chaos game for IFSP. The paper is divided into two parts: in the first one, we discuss how to determine the integral with respect to a measure which is a combination of a self-similar measure from an IFSP along with a density given by an IFSM. In the second part, we prove a version of the Ergodic Theorem for the integration of a continuous multifunction with respect to the invariant measure of an IFSP. These results are in line with some recent extensions of IFS theory to multifunctions.

1 Introduction

Since the introduction of the classical definition of Iterated Function Systems (IFSs) a large amount of work has been devoted to providing an efficient algorithm to determine its attractor. With this regard the Chaos Game algorithm probably represents the most important result: it allows one to compute integrals with respect to the invariant measure of an IFS with probabilities (or IFSP) by considering the average value along any particular trajectory of an easily defined random dynamical system. From a mathematical point of view, it can be considered an extension of the classical Ergodic Theorem. In this paper we aim to extend the Chaos game to the case of multifunctions.

The use of multifunctions in fractal analysis and fractal image coding is quite recent (see [17], [18]). According to the classical IFS theory, an image can be modeled either as an L^p function or as a probability measure. The motivations behind these two different constructions is quite intuitive. When a function-based representation is used, at each pixel in the domain the function assigns the color corresponding to that pixel. When this can not be done with precision because pixel colors are subject to noise and distortions, it is more convenient to use a measure-based approach which assigns the averaged color of a given block of pixels. In [13], [17], and [18], we illustrated the advantages of using the

*University of Milan, Department of Economics, Management and Quantitative Methods, Milan, Italy.

†Department of Mathematics and Statistics, Acadia University, Wolfville, Canada.

notion of multifunction in image analysis; through this formalism it is possible to describe in a unique framework possible uncertainty or noise on the color of a pixel. The easiest way to do this consists of assigning to any pixel an interval of possible values; from a statistical perspective this corresponds to providing an interval of possible outcomes for the random variable that describes the color of that pixel; the lower and upper values describe the range of possible color variations. This interval-valued formulation leads quite naturally to using multifunctions. Similar considerations justify the introduction of multimeasures. Bearing this in mind, it is quite natural to try to extend the Chaos Game algorithm to the case of multifunctions and multimeasures. In this paper we conduct the analysis for the multifunction case while we leave the multimeasure extension for a future work.

The paper is organized as follows. Section 2 is dedicated to refreshing the basic formalism of IFS and its variants. In Section 2 we also prove two new results involving IFS on functions and multifunctions where we assume only average conctractivity. In Section 3 we provide an extension of the chaos game for fractal measures with fractal densities, while Section 4 is devoted to proving the convergence of the Chaos game algorithm for multifunctions.

2 Iterated Function Systems and their generalizations

2.1 Geometric IFS

An (geometric) IFS is a finite collection of contractive maps $\{w_i : 1 \leq i \leq N\}$ which are defined on a complete metric space $(\mathbb{X}, d)$. This collection of maps w_i leads to the introduction of a mapping $\hat{W}(B) := \cup_i w_i(B)$ which is defined on the set of all non-empty compact subsets of $H(\mathbb{X})$. If one places on $H(\mathbb{X})$ the Hausdorff distance d_H defined as

$$d_H(A, B) = \max\{\sup_{u \in A} \inf_{v \in B} d(u, v), \sup_{u \in B} \inf_{v \in A} d(u, v)\}, \tag{1}$$

then $\hat{W}$ is contractive in the Hausdorff metric with fixed point A, which is called the attractor of the IFS. Thus, the existence and uniqueness of the attractor is a simple consequence of the Contraction Mapping theorem and it is the unique non-empty compact set $A \subset \mathbb{X}$ which satisfies the generalized self-similarity condition

$$A = \bigcup_i w_i(A). \tag{2}$$

This means that

$$\hat{W}^n(B) = \bigcup \{w_{\sigma_1} \circ w_{\sigma_2} \circ \cdots \circ w_{\sigma_n}(B) : \text{all sequences } \sigma_i \in \{1, 2, \dots, N\}\}$$

will converge to the attractor A for any initial compact set B. Since A is compact, this also means that the set

$$\text{closure}\left(B \cup \bigcup_{n\geq 1} \hat{W}^n(B)\right) \subset \mathbb{X} \tag{3}$$

is also compact (we use this fact later in the proof of Theorem 4.4). For more details about IFS we refer the reader to [3, 9, 15].

We pause here to comment on the use of the Contraction Mapping theorem as the fixed point theorem. It is certainly possible to use other fixed point theorems to guarantee the existence of a fixed point. However, the Contraction Mapping theorem has the benefit of ensuring geometric convergence of the iterates, which is very useful in practical situations. In addition, this fast convergence simplifies considerations involving Cesàro means, for example in considerations involving the Ergodic theorem.

2.2 IFS with probabilities

An Iterated Function System with Probabilities (IFSP) consists of the above collection of IFS maps w_i together with an associated set of probabilities p_i. An IFSP induces a Markov operator on the set of all Borel probability measures $B(\mathbb{X})$ on $\mathbb{X}$ whose action is given by

$$\mathcal{M}\nu(B) = \sum_i p_i \nu(w_i^{-1}(B)), \tag{4}$$

where ν is a Borel probability measure on $\mathbb{X}$ and B is an arbitrary Borel subset of $\mathbb{X}$. Let $M(X)$ denote the set of all Borel probability measures on X and define the Monge-Kantorovich metric by

$$d_{MK}(\mu, \nu) = \sup_{f \in Lip_1} \left[\int f d\mu - \int f d\nu \right]. \tag{5}$$

It is possible to prove that $(M(\mathbb{X}), d_{MK})$ is a complete metric space.

Definition 2.1. We say that the IFSP $\{w_i, p_i\}$ is *contractive on the average* if $\sum_i p_i s_i < 1$, where s_i is the contractivity factor of the map w_i.

This average contractivity condition on the IFSP is sufficient to ensure that the Markov operator $\mathcal{M}$ is contractive. The resulting unique fixed point of $\mathcal{M}$ is called the attractor of the IFSP (for more on IFSP see [3, 9, 15]).

The Ergodic Theorem for IFSP, proved by Elton (the theorem in [7] or theorem 3 in [8]), is the result which underlies the Chaos game. To describe the Chaos game, let $x_0 \in \mathbb{X}$ be arbitrary. For each n, choose $\sigma_n \in \{1, 2, \ldots, N\}$ according to the probabilities p_i. Having chosen σ_n, we define $x_{n+1} = w_{\sigma_n}(x_n)$. Then for any bounded continuous function f on $\mathbb{X}$ we have

$$\frac{1}{n}\sum_{i\leq n} f(x_i) \xrightarrow{n} \int_{\mathbb{X}} f(x)\ d\mu(x), \text{ for almost all sequences } \sigma_n,$$

where μ is the invariant probability measure associated with the IFSP. One major difference between the Ergodic Theorem for IFSP and the classical Ergodic Theorem is the fact that the Chaos game converges to the integral for any initial $x_0 \in \mathbb{X}$. In the classical Ergodic Theorem one only has the conclusion for μ almost all $x_0 \in \mathbb{X}$. To be fair, however, the Chaos game converges only for almost all sequences σ_n.

2.3 Simple consequences of average contractivity

We now make some observations, which will be very useful later in this paper, about IFSP which are average contractive.

Let $\Lambda = \{0, 1, \ldots, N\}^\infty$ be the code space associated with a given average contractive IFSP $\{w_i, p_i\}$ on $\mathbb{X}$ and denote by $\mathbb{P}$ the product probability on Λ induced by $\{p_i\}$ on each factor. Now,

$$d(w_{\sigma_1} \circ w_{\sigma_2} \circ \cdots \circ w_{\sigma_n}(x),\ w_{\sigma_1} \circ w_{\sigma_2} \circ \cdots \circ w_{\sigma_n}(y) \leq s_{\sigma_1} s_{\sigma_2} \cdots s_{\sigma_n} d(x, y)$$

and for any $\tau > \sum_i p_i s_i$ we have

$$\mathbb{P}(s_{\sigma_1} s_{\sigma_2} \cdots s_{\sigma_n} > \tau^n) = \mathbb{P}(\frac{1}{n} \sum_{i=1}^{n} \ln(s_{\sigma_i}) > \ln(\tau)).$$

Let ξ the the discrete random variable given by $Pr(\xi = \ln(s_i)) = p_i$. Then since $\{w_i, p_i\}$ is contractive on average, we have $\mathbb{E}(\xi) = \sum_i p_i \ln(s_i) < 0$. Thus, by Cramèr's Theorem (see [6, page 5]) applied to the random variable ξ, for any $\kappa < 1$ there is a $0 < \gamma < 1$ so that for all $n \in \mathbb{N}$

$$\mathbb{P}\{d(w_{\sigma_1} \circ w_{\sigma_2} \circ \cdots \circ w_{\sigma_n}(x),\ w_{\sigma_1} \circ w_{\sigma_2} \circ \cdots \circ w_{\sigma_n}(y)) \leq \kappa^n d(x, y)\} \geq 1 - \gamma^n. \quad (6)$$

This means that there is a set $\Theta \subseteq \Lambda$ with $\mathbb{P}(\Theta) = 1$ such that the address map $a_w : \Theta \to \mathbb{X}$ is well-defined. Usually $a_w(\Lambda) = A \subset \mathbb{X}$ is the attractor of the geometric IFS $\{w_i\}$. In our case, $a_w(\Theta)$ is a subset of $\mathbb{X}$ of full μ-measure. However, the "full" geometric attractor might not be well-defined.

2.4 IFS on functions

To define an IFS on maps (IFSM) and set-valued maps (multifunctions), we start with a geometric IFS $\{w_i\}$ and Lipschitz maps $\phi_i : \mathbb{R} \to \mathbb{R}$, and then we define the action of the corresponding IFSM operator on functions $f : \mathbb{X} \to \mathbb{R}$ by (see [12])

$$T(f)(x) = \sum_i \phi_i(f(w_i^{-1}(x))), \quad (7)$$

with the understanding that, in the sum, we skip those i for which $x \notin w_i(\mathbb{X})$. This defines an operator $T : L^p(\mathbb{X}) \to L^p(\mathbb{X})$ and, under appropriate conditions on w_i and ϕ_i, T will be contractive and thus have a unique fixed point, the attractor of this IFSM.

We will need contractivity in $L^\infty(\mathbb{X},\mu)$, so we give a result which ensures this contractivity. As we will only be interested in the case where the geometric IFS $\{w_i\}$ is non-overlapping, we restrict to this case.

Proposition 2.2. *Suppose that $\{w_i, p_i\}$ for $i = 1, 2, \ldots, N$ is a non-overlapping IFSP on $\mathbb{X}$ which is contractive. Suppose further that $\{\phi_i, p_i\}$ is an IFSP on $\mathbb{R}$ which is contractive on average. Let μ be the invariant measure of $\{w_i, p_i\}$ on $\mathbb{X}$. Define the IFSM operator T as in (7).*

Then T has a unique fixed point ψ in $L^\infty(\mathbb{X},\mu)$. Furthermore, for any $g \in L^\infty(\mathbb{X},\mathbb{R})$ we have $T^n(g) \to \psi$ in $L^\infty(\mathbb{X},\mu)$.

Proof. By the comments above in section 2.3, since both $\{w_i, p_i\}$ and $\{\phi_i, p_i\}$ are average contractive, there is a set $\Theta \subset \Lambda$ with $\mathbb{P}(\Theta) = 1$ and such that the two address maps $a_w : \Lambda \to \mathbb{X}$ and $a_\phi : \Theta \to \mathbb{R}$ are well-defined. We define the function $\psi : \mathbb{X} \to \mathbb{R}$ by

$$\psi(x) = \begin{cases} 0, & \text{if } x \notin a_w(\Theta) \\ a_\phi(\sigma), & \text{if } x = a_w(\sigma). \end{cases}$$

For $x = a_w(\sigma) = a_w(\sigma_1, \sigma_2, \sigma_3, \ldots)$, we have $\psi(x) = a_\phi(\sigma) = a_\phi(\sigma_1, \sigma_2, \sigma_3, \ldots)$. In addition, we have $w_i^{-1}(x) = a_w(\sigma_2, \sigma_3, \ldots)$ and $\phi_i(\psi(w_i^{-1}(x))) = a_\phi(\sigma_2, \sigma_3, \ldots)$. This means that

$$\psi(x) = a_\phi(\sigma_1, \sigma_2, \sigma_3, \ldots) = \phi_i(\psi(w_i^{-1}(x)))$$

and so $\psi(x) = T(\psi)(x)$ for μ-almost all $x \in \mathbb{X}$.

Next we show that $T^n(g) \to \psi$ in $L^\infty(\mathbb{X},\mu)$ for any $g \in L^\infty(\mathbb{X},\mu)$. Since we assumed that $w_i(A) \cap w_j(A) = \emptyset$ for $i \neq j$ (where $A = a_w(\Lambda)$ is the geometric attractor), we know that $w_i^{-1}(x)$ only makes sense if $x \in w_i(\mathbb{X})$. Furthermore, since $\mu(A) = 1$, we can restrict attention to $x \in w_i(A)$. Now,

$$T^n(g)(x) = \sum_{\sigma_1,\sigma_2,\ldots,\sigma_n \in \{1,2,\ldots,N\}} \phi_{\sigma_n} \circ \cdots \phi_{\sigma_1} \circ g(w_{\sigma_1}^{-1} \circ \cdots \circ w_{\sigma_n}^{-1}(x)), \tag{8}$$

where the sum is over all possible choices $\sigma_1, \sigma_2, \ldots, \sigma_n \in \{1, 2, \ldots, N\}$. Because the IFS maps w_i are disjoint, for any fixed $x \in A$ there is only one term in the sum from (8) which is relevant. In particular, this is the choice of σ_i so that $x \in w_{\sigma_1} \circ \cdots \circ w_{\sigma_n}(A)$. We have

$$\phi_{\sigma_n} \circ \cdots \circ \phi_{\sigma_1} \circ g(w_{\sigma_1}^{-1} \circ \cdots \circ w_{\sigma_n}^{-1}(x)) = (\phi_{\sigma_n} \circ \cdots \circ \phi_{\sigma_1}) \circ g \circ (w_{\sigma_n} \circ \cdots \circ w_{\sigma_1})^{-1}(x).$$

Thus for $x = a_w(\sigma) = a_w(\sigma_1, \sigma_2, \ldots, \sigma_n, \sigma_{n+1}, \ldots)$

$$T^n(g)(x) = (\phi_{\sigma_n} \circ \cdots \circ \phi_{\sigma_1}) \circ g(a_w(\sigma_{n+1}, \sigma_{n+2}, \ldots)) \tag{9}$$

and

$$\psi(x) = T^n(\psi)(x) = (\phi_{\sigma_n} \circ \cdots \circ \phi_{\sigma_1}) \circ \psi(a_w(\sigma_{n+1}, \sigma_{n+2}, \ldots)). \tag{10}$$

However, then by (6) applied to $\{\phi_i, p_i\}$ we have the desired result for $\sigma \in \Theta$. $\square$

We note that because we assume that the IFS $\{w_i\}$ is disjoint, the proof of Proposition 2.2 applies to the much more general situation of an IFS on functions $f : \mathbb{X} \to \mathbb{Y}$ where both $\mathbb{X}$ and $\mathbb{Y}$ are complete metric spaces, $\{w_i, p_i\}$ is a contractive IFS on $\mathbb{X}$ and $\{\phi_i, p_i\}$ is an average contractive IFS on $\mathbb{Y}$. In this case, choosing $f_0 : \mathbb{X} \to \mathbb{Y}$ to be any function, we can define an IFS operator on such functions by

$$T(f)(x) = \begin{cases} \phi_i(f(w_i^{-1}(x))), & \text{if } x \in w_i(A) \\ f_0(x), & \text{if } x \notin w_i(A) \text{ for any } i. \end{cases}$$

Then the proof of Proposition 2.2 shows that there is a unique fixed function $\psi : \mathbb{X} \to \mathbb{Y}$ and for any $g : \mathbb{X} \to \mathbb{Y}$ we have $T^n(g) \to \psi$ "uniformly μ-almost everywhere". We will use this observation in the next section.

2.5 IFS on set-valued functions

An IFS on multifunctions is defined in a similar way as IFS on functions (see [17]). Define $\mathcal{K}_c$ by

$$\mathcal{K}_c = \{C \subset \mathbb{R}^d : \emptyset \neq C \text{ is compact and convex}\}. \tag{11}$$

For us, a *multifunction* (or set-valued function) will always be a function $F : \mathbb{X} \to \mathcal{K}_c$, so that $F(x)$ is a non-empty compact and convex subset of $\mathbb{R}^d$ for each $x \in \mathbb{X}$. There are several different complete metrics one can place on $\mathcal{K}_c$, including L^p type metrics and the Hausdorff metric. We will use the Hausdorff metric.

Given the geometric IFS $\{w_i\}$ and, in addition, the linear maps $\alpha_i : \mathbb{R}^d \to \mathbb{R}^d$ and a fixed multifunction $F_0 : \mathbb{X} \to \mathcal{K}_c$, we define the IFS operator

$$T(F)(x) = \sum_i \alpha_i(F(w_i^{-1}(x))) + F_0(x), \tag{12}$$

where again we skip those i for which $x \notin w_i(\mathbb{X})$. As before, under suitable conditions on w_i and α_i, we obtain a contraction and a unique fixed point for T. In this case, the fixed point is a self-similar multifunction.

We will again only use the situation where the IFS $\{w_i\}$ is non-overlapping. Proposition 2.2 can easily be extended to the more general situation of IFS on multifunctions.

Proposition 2.3. *Suppose that $\{w_i, p_i\}$ for $i = 1, 2, \ldots, N$ is a non-overlapping IFSP on $\mathbb{X}$ which is contractive. Suppose further that $\alpha_i : \mathbb{R}^d \to \mathbb{R}^d$ are linear maps with $\sum_i p_i \|a_i\| < 1$. Let μ be the invariant measure of $\{w_i, p_i\}$ on $\mathbb{X}$ and let $F_0 : \mathbb{X} \to \mathcal{K}_c$ be a fixed multifunction. Let T be the IFS operator defined in (12).*

Then T has a unique fixed point Ψ in $L^\infty(\mathbb{X}, \mathcal{K}_c, \mu)$. Furthermore, for any $g \in L^\infty(\mathbb{X}, \mathcal{K}_c, \mu)$ we have $T^n(g) \to \Psi$ in $L^\infty(\mathbb{X}, \mathcal{K}_c, \mu)$.

Proof. First we see that the summation in the definition of T is unnecessary since each point x is in only one $w_i(A)$. Thus, we can rewrite the definition of T as

$$T(F)(x) = \begin{cases} \alpha_i(F(w_i^{-1}(x)) + F_0(x), & \text{if } x \in w_i(A) \\ F_0(x), & \text{if } x \notin w_i(A) \text{ for any } i. \end{cases} \tag{13}$$

Now we use the observation immediately after the proof of Proposition 2.2 to see that the conclusion holds. □

Example 2.4. Take $\mathbb{X} = [0,1]$, $w_1(x) = x/2$, $w_2(x) = x/2 + 1/2$ so that the attractor of $\{w_1, w_2\}$ is $[0,1]$. For simplicity we will choose $p_1 = p_2 = 1/2$. Let $\theta_1 = 0$ and $\theta_2 = 1/100$ so that θ_2/π is irrational. We use these two angles to define the two linear transformations $\alpha_1, \alpha_2 : \mathbb{R}^2 \to \mathbb{R}^2$ by setting α_i to be a rotation around the origin by an angle of θ_i followed by a uniform scaling with scaling factor $1/2$. Finally, let the multifunction $F_0 : [0,1] \rightrightarrows \mathbb{R}^2$ be defined by $F_0(x) = [-1,1]^2$ for all $x \in [0,1]$.

This data defines an IFS operator T on multifunctions as in (12). The fixed point multifunction Ψ of this operator is rather difficult to describe precisely. For any x, the compact and convex set $\Psi(x)$ is an infinite sum of reduced and rotated copies of $[-1,1]^2$. The precise rotation of each term in the sum depends in an intricate way on the binary expansion of $x \in [0,1]$. Since the multiples of $1/100$ are dense in $[0, 2\pi)$, the set $\Psi(x)$ has no "corners".

To conclude this brief review of IFS formulations and extensions, it is worth recalling that in previous work Chaos games have been constructed for rendering the attractor of an IFSM (see [12]) and for performing both a wavelet analysis of an arbitrary L^2 function and rendering an approximation to an arbitrary L^2 function via wavelet synthesis (see [19]).

3 Chaos game for fractal measures with fractal densities

As announced in the introduction, the Chaos game was originally conceived to produce an image of the attractor of the IFS (see [3]), as the support of μ is the attractor so plotting each x_n in turn will produce a sequence of approximations to the attractor which converge (in the Hausdorff distance). The image would seem to appear as if from a cloud of smoke, since the first few iterations produce scattered points with no particular structure. The aim of this section is to extend the Chaos game to measures ν of the form

$$\nu(B) = \int_B \psi(x)\ d\mu(x),$$

where μ is the attractor of the IFS with probabilities $\{w_i, p_i\}$ on $\mathbb{X}$ and ψ is the attractor of the IFSM $\{w_i, \phi_i\}$. We assume that $w_i(\mathbb{X}) \cap w_j(\mathbb{X}) = \emptyset$ for $i \neq j$ and that each $\phi_i : \mathbb{R} \to \mathbb{R}$ is contractive. This implies that there is some $M > 0$

with $|\psi(x)| \leq M$ for all $x \in \mathbb{X}$. Throughout the paper we make the standing assumption that $\mathbb{X}$ is a locally compact, separable and complete metric space.

Theorem 3.1. *Suppose that $\{w_i, p_i\}$ is a contractive IFSP on $\mathbb{X}$ and $\{\phi_i, p_i\}$ is an average contractive IFSP on $\mathbb{R}$. Let $x_0 \in \mathbb{X}$, $y_0 \in \mathbb{R}$ and suppose that $f : \mathbb{X} \to \mathbb{R}$ is bounded and continuous. Let $\sigma_n \in \{1, 2, \ldots, N\}$ be a sequence of independent and identically distributed random variables which are chosen according to the probabilities $\{p_i\}$ and define $x_{n+1} = w_{\sigma_n}(x_n)$ and $y_{n+1} = \phi_{\sigma_n}(y_n)$.*

Then for all choices of $x_0 \in \mathbb{X}$ and all $y_0 \in \mathbb{R}$ and $\mathbb{P}$-a.e. σ we have that

$$\lim_n 1/n \sum_{i \leq n} y_i f(x_i) = \int_{\mathbb{X}} f(x) d\nu(x) = \int_{\mathbb{X}} f(x)\psi(x) d\mu(x).$$

Proof. As in (6), with except for a set of exponentially vanishing probability we have

$$|\phi_{\sigma_n} \circ \phi_{\sigma_{n-1}} \circ \cdots \circ \phi_{\sigma_1}(y_0) - \phi_{\sigma_n} \circ \phi_{\sigma_{n-1}} \circ \cdots \circ \phi_{\sigma_1}(z_0)| \leq \kappa^n \; |y_0 - z_0| \to 0$$

for any $y_0, z_0 \in \mathbb{R}$. Thus without loss of generality we assume that $y_0 = \psi(x_0)$, and notice that

$$y_{n+1} = \phi_{\sigma_n}(y_n) = \phi_{\sigma_n}(\psi(x_n)) = \psi(x_{n+1}).$$

The set $\Omega = \mathbb{X} \times \mathbb{R}$ equipped with the metric

$$d(\ (x_1, y_1),\ (x_2, y_2)\ ) = d(x_1, x_2) + |y_1 - y_2|$$

is complete and locally compact. If we define the functions $F_i : \Omega \to \Omega$ as $F_i(x, y) = (w_i(x), \phi_i(y))$, then the IFS $\{F_i, p_i\}$ is average contractive on Ω. Take $\xi : \Omega \to \mathbb{R}$ defined as

$$\xi(x, y) = \begin{cases} yf(x), & \text{if } |y| \leq 2M \\ -2Mf(x), & \text{if } y < -2M \\ 2Mf(x), & \text{if } y > 2M. \end{cases} \tag{14}$$

Let $A \subset \mathbb{X}$ be the attractor of $\{w_i\}$ and η be the invariant measure of $\{F_i, p_i\}$ on Ω. Then, as before, there is a set $\Theta \subseteq \Lambda$ with $\mathbb{P}(\Theta) = 1$ where the address map $a_\phi : \Theta \to \mathbb{X}$ is well-defined. Define

$$\Delta = \{(x, \psi(x)) : x \in a_w(\Theta)\}. \tag{15}$$

It is easy to see that $\eta(\Delta) = 1$, η projects from Δ onto the measure μ on $a_w(\Theta) \subseteq A \subset \mathbb{X}$, and $\mu(A \setminus a_w(\Theta)) = 0$. Notice that $\xi(x, y) = yf(x)$ for any $(x, y) \in \Delta$.

Let $z_0 = (x_0, y_0) \in \Omega$ and consider the random sequence $z_{n+1} = F_{\sigma_n}(z_n) = (x_{n+1}, y_{n+1})$. Using Elton's Theorem, we obtain

$$1/n \sum_{i \leq n} y_i f(x_i) \quad = \quad 1/n \sum_{i \leq n} \xi(x_i, y_i) = 1/n \sum_{i \leq n} \xi(z_i)$$

$$\to \quad \int_{\Omega} \xi(z)\ d\theta(z) = \int_{\mathbb{X}} f(x)\psi(x)\ d\mu(x)$$

which proves the theorem. □

Remark 3.2. If one has a (finite) measure ν of the form

$$\nu(B) = \int_B \theta(x)\ d\mu(x),$$

where θ is any bounded function in $L^1(\mathbb{X}, \mu)$, it is simple to see that

$$\frac{1}{n}\sum_{i \le n} f(x_i)\theta(x_i) \to \int_{\mathbb{X}} f(x)\theta(x)\ d\mu(x) = \int_{\mathbb{X}} f(x)\ d\nu(x)$$

for all bounded continuous f.

The advantage of using a fractal density is that the IFSM maps ϕ_i automatically give the value of the density at any point.

It is not possible, in general, to allow the IFS maps w_i to overlap.

Example 3.3. Let $\mathbb{X} = [0,1]$, $w_0(x) = 2x/3$ and $w_1(x) = 2x/3 + 1/3$ with $p_0 = p_1 = 1/2$, and $\phi_0(t) = 1$ and $\phi_1(t) = 1$. Choose the function $f(x) = 1$. Then clearly ϕ_0, ϕ_1 are contractive and $\psi(x) = \chi_{[0,1/3]} + 2\chi_{[1/3,2/3]} + \chi_{[2/3,1]}$ is the fixed point of the IFSM operator. The invariant measure μ of the IFSP is symmetric, so $\mu([0,1/2]) = \mu([1/2,1]) = 1/2$. Using the invariance of μ, we compute that

$$\begin{aligned}
\int_0^1 f(x)\psi(x)\ d\mu(x) &= \mu([0,1/3]) + 2\mu([1/3,2/3]) + \mu([2/3,1]) \\
&= p_0(\mu([0,1/2]) + 2\mu([1/2,1])) + p_1(2\mu([0,1/2]) + \mu([1/2,1])) \\
&= 3/2.
\end{aligned}$$

However, $y_n = 1$ whatever the choice of σ_n and thus $y_n f(x_n) = 1$ for all n and so the average along the trajectory is 1 as well.

The problem with having an overlap is that if x_n is in the region of overlap (so that there are at least two different addresses for x_n), then it is generally not true that $y_n = \psi(x_n)$. We see this in the Example 3.3 where $\psi(x) = 2$ for any x in the overlap but $y_n = 1$ always.

4 Chaos game for multifunctions

In this section we aim at extending the Chaos game to set-valued functions and IFS on set-valued functions. More precisely, if $F : \mathbb{X} \rightrightarrows \mathbb{R}^d$ is a multifunction which takes values in $\mathcal{K}_c$, the space of all non-empty compact and convex subsets of $\mathbb{R}^d$, we aim at computing

$$\int_{\mathbb{X}} F(x)\ d\mu(x)$$

via a Chaos Game algorithm. The above integral has to be understood in the Aumann sense (see [2]), that is

$$\int_{\mathbb{X}} F(x) d\mu(x) = \left\{ \int_{\mathbb{X}} f(x) d\mu(x) : f \in L^1(\mathbb{X}, \mu), f(x) \in F(x) \text{ for } \mu\text{-a.e. } x \in \mathbb{X} \right\}.$$

Continuity of a multifunction F is defined using the Hausdorff distance on the range. If there is a compact set $K \subset \mathbb{R}^d$ with $F(x) \subseteq K$ for all x then F is *bounded*. The *support function* of a compact and convex set K is the function $\mathrm{supp}(\cdot, K) : \mathbb{R}^d \to \mathbb{R}$ defined by

$$\mathrm{supp}(q, K) = \sup_{k \in K} q \cdot k.$$

Because $\mathrm{supp}(\lambda q, K) = \lambda\ \mathrm{supp}(q, K)$ for $\lambda \geq 0$, the support function is clearly defined by its values on $S^{d-1} = \{y \in R^d : \|y\| = 1\}$. A compact and convex K may be recovered from its support function as

$$K = \bigcap_{q \in S^{d-1}} \{z \in \mathbb{R}^d : z \cdot q \leq \mathrm{supp}(q, K)\}.$$

The *norm* of a set $K \in \mathcal{K}_c$ is defined as $\|K\| = \sup\{\|x\| : x \in K\}$. For more properties of multifunction we refer to [2].

Before proving our main results, we slightly modify the usual ergodic theorem for the case of continuous multifunctions.

Proposition 4.1. *Let $\{w_i, p_i\}$ be an average contractive IFSP on $\mathbb{X}$. Choose $x_0 \in \mathbb{X}$ and define $x_{n+1} = w_{\sigma_n}(x_n)$, with $\sigma_n \in \{1, 2, \ldots, N\}$ being independent and identically distributed random variables chosen according to $\{p_i\}$. Finally, let $F : \mathbb{X} \to \mathcal{K}_c$ be any bounded continuous multifunction and $q \in S^{d-1}$ be fixed.*

Then for any choice of x_0 and $\mathbb{P}$-a.e. σ, we have that

$$\lim_n \mathrm{supp}\left(q, (1/n) \sum_{i \leq n} F(x_i)\right) = \mathrm{supp}\left(q, \int_{\mathbb{X}} F(x)\ d\mu(x)\right).$$

Proof. Classical properties of the support function supp imply that

$$\mathrm{supp}\left(q, 1/n \sum_{i \leq n} F(x_i)\right) = 1/n \sum_{i \leq n} \mathrm{supp}(q, F(x_i))$$

and

$$\mathrm{supp}\left(q, \int_{\mathbb{X}} F(x)\ d\mu(x)\right) = \int_{\mathbb{X}} \mathrm{supp}(q, F(x))\ d\mu(x).$$

On the other hand, the continuity and the boundedness of F imply that $x \mapsto \mathrm{supp}(p, F(x))$ is continuous and bounded for all p. We can now apply Elton's Ergodic theorem for IFSP [7], which leads to

$$\begin{aligned}
\lim_n \mathrm{supp}\left(q, 1/n \sum_{i\le n} F(x_i)\right) &= \lim_n 1/n \sum_{i\le n} \mathrm{supp}(q, F(x_i)) \\
&= \int_{\mathbb{X}} \mathrm{supp}(q, F(x))\ d\mu(x) \\
&= \mathrm{supp}\left(q, \int_{\mathbb{X}} F(x)\ d\mu(x)\right)
\end{aligned}$$

for $\mathbb{P}$-a.e. σ, as desired. ☐

Now we would like to use the above result to conclude that the sequence of compact and convex sets $1/n \sum_{i\le n} F(x_i)$ converges to the compact and convex set $\int_{\mathbb{X}} F(x)\ d\mu(x)$. The issue is that for each $q \in S^{d-1}$, we have convergence only for $\sigma \in C_q$ where C_q is a subset of the code space Λ of full $\mathbb{P}$-measure. But then it is conceivable that $\mathbb{P}(\bigcap_q C_q) < 1$. The next result shows that this cannot happen.

Theorem 4.2. *Let $\{w_i, p_i\}$ be an average contractive IFSP on $\mathbb{X}$. Choose $x_0 \in \mathbb{X}$ and define $x_{n+1} = w_{\sigma_n}(x_n)$, with $\sigma_n \in \{1, 2, \ldots, N\}$ being independent and identically distributed random variables chosen according to $\{p_i\}$. Finally, let $F : \mathbb{X} \to \mathcal{K}_c$ be any bounded continuous multifunction and $q \in S^{d-1}$ be fixed.*

Then for all choices of $x_0 \in \mathbb{X}$ and $\mathbb{P}$-a.e. σ, the limit

$$\lim_n \mathrm{supp}\left(q, (1/n) \sum_{i\le n} F(x_i)\right) = \mathrm{supp}\left(q, \int_{\mathbb{X}} F(x)\ d\mu(x)\right)$$

exists and is uniform over all $q \in S^{d-1}$

Proof. Let us consider a countable dense subset $\{q_m\}$ of S^{d-1} . Then for each q_m, we have a subset $C_m \subseteq \Lambda$ such that for all $\sigma \in C_m$

$$\lim_n \mathrm{supp}\left(q_m, (1/n) \sum_{i\le n} F(x_i)\right) = \mathrm{supp}\left(q_m, \int_{\mathbb{X}} F(x)\ d\mu(x)\right) \tag{16}$$

and with $\mathbb{P}(C_m) = 1$. Let $C = \bigcap_m C_m$. It is trivial to prove that $\mathbb{P}(C) = 1$ with (16) holding for all q_m and all $\sigma \in C$.

By assumption F is a bounded multifunction and so there is a a compact and convex $K \subset \mathbb{R}^d$ with $F(x) \subset K$ for all x and also a constant $M > 0$ with $\mathrm{supp}(q, K) \le M$ for all $q \in S^{d-1}$. Because K is convex, we have for all n

$$1/n \sum_{i\le n} F(x_i) \subset K$$

and therefore

$$\left|\operatorname{supp}\left(q, 1/n\sum_{i\leq n} F(x_i)\right)\right| \leq M.$$

On the other hand, we have

$$\int_{\mathbb{X}} F(x)\ d\mu(x) \subset K$$

and

$$\left|\operatorname{supp}\left(q, \int_{\mathbb{X}} F(x)\ d\mu(x)\right)\right| \leq M.$$

This implies that for all q and all n, the convex and bounded functions (by M, see [21]) given by

$$q \mapsto \operatorname{supp}\left(q, 1/n\sum_{i\leq n} F(x_i)\right)$$

and

$$q \mapsto \operatorname{supp}\left(q, \int_{\mathbb{X}} F(x)\ d\mu(x)\right)$$

are Lipschitz with factor at most M. Let us take $\epsilon > 0$ and choose $\delta = \epsilon/(3M)$ and $q_{m_1}, q_{m_2}, \ldots, q_{m_l}$ such that they form a δ-cover of the compact set S^{d-1}. Let N large enough so that $n \geq N$ implies that

$$\left|\operatorname{supp}\left(q_{m_i}, 1/n\sum_{i\leq n} F(x_i)\right) - \operatorname{supp}\left(q_{m_i}, \int_{\mathbb{X}} F(x)\ d\mu(x)\right)\right| < \epsilon/3$$

for $i = 1, 2, \ldots, l$. Then for a;; $n \geq N$ and all $q \in S^{d-1}$ we have that

$$\left|\operatorname{supp}\left(q, 1/n\sum_{i\leq n} F(x_i)\right) - \operatorname{supp}\left(q, \int_{\mathbb{X}} F(x)\ d\mu(x)\right)\right| \leq$$

$$\left|\operatorname{supp}\left(q, 1/n\sum_{i\leq n} F(x_i)\right) - \operatorname{supp}\left(q_{m_i}, 1/n\sum_{i\leq n} F(x_i)\right)\right| +$$

$$\left|\operatorname{supp}\left(q_{m_i}, 1/n\sum_{i\leq n} F(x_i)\right) - \operatorname{supp}\left(q_{m_i}, \int_{\mathbb{X}} F(x)\ d\mu(x)\right)\right| +$$

$$\left|\operatorname{supp}\left(q_{m_i}, \int_{\mathbb{X}} F(x)\ d\mu(x)\right) - \operatorname{supp}\left(q, \int_{\mathbb{X}} F(x)\ d\mu(x)\right)\right| <$$

$$M\frac{\epsilon}{3M} + \frac{\epsilon}{3} + M\frac{\epsilon}{3M} = \epsilon$$

which proves the theorem. $\square$

The following corollary gives us the result we seek.

Corollary 4.3. *Let us suppose that the hypotheses of the previous theorem are satisfied. Then we have*

$$1/n \sum_{i \leq n} F(x_i) \to \int_{\mathbb{X}} F(x)\ d\mu(x)$$

where the convergence is in the Hausdorff metric.

Proof. The proof is trivial and easily follows once we observe that for two compact and convex sets $A, B \subset \mathbb{R}^d$ it happens

$$d_H(A, B) = \sup_{q \in S^{d-1}} |\mathrm{supp}(q, A) - \mathrm{supp}(q, B)|.$$

□

The main arguments used in Theorems 3.1 and 4.2 and Proposition 4.1 can be combined to produce a Chaos game for integrals of the form

$$\phi(B) = \int_B f(x)\Psi(x)\ d\mu(x) \tag{17}$$

where $f : \mathbb{X} \to \mathbb{R}$ is continuous, $\Psi : \mathbb{X} \rightrightarrows \mathbb{R}^d$ is the multifunction attractor of the IFS on multifunctions $\{w_i, \alpha_i, F_0\}$ from (12) and μ is the invariant probability measure for $\{w_i, p_i\}$. Since we will assume that $\{w_i\}$ are non-overlapping, we can reformulate this operator in a way similar to that of (13) by using

$$\Phi_i(F)(x) = \alpha_i F(w_i^{-1}(x)) + F_0(x) \quad \text{ for } i = 1, 2, \ldots, N.$$

We will assume that the fixed point multifunction Ψ is bounded. Notice that if each $\alpha_i : \mathbb{R}^d \to \mathbb{R}^d$ is a linear contraction and F_0 is bounded then this will be the case. With no loss of generality, we assume that the bounding set K is of the form $K = \{x \in \mathbb{R}^d : \|x\| \leq r\}$ for some $r > 0$.

Let us choose $x_0 \in \mathbb{X}$ and $S_0 \in \mathcal{K}_c$ and with σ_n, as before, being independent and identically distributed. Let $x_{n+1} = w_{\sigma_n}(x_n)$ and $S_{n+1} = \alpha_{\sigma_n}(S_n)$. We aim at proving that

$$1/n \sum_{i \leq n} f(x_i) S_i \to \int_{\mathbb{X}} f(x)\Psi(x)\ d\mu(x).$$

We now proceed by first proving the case where the function f is non-negative.

Theorem 4.4. *Let $\{w_i, p_i\}$ be a contractive IFSP on $\mathbb{X}$ and $\{w_i, \alpha_i, F_0\}$ be an IFS on multifunctions with $\sum_i p_i \|\alpha_i\| < 1$. Choose $x_0 \in \mathbb{X}$, $S_0 \in \mathcal{K}_c$, and define $x_{n+1} = w_{\sigma_n}(x_n)$ and $S_{n+1} = \Phi_{\sigma_n}(S_n)$ where $\sigma_n \in \{1, 2, \ldots, N\}$ is a sequence of independent and identically distributed random variables chosen according to the probabilities $\{p_i\}$. Finally, suppose that $f : \mathbb{X} \to [0, \infty)$ is a bounded and continuous function.*

Then for all x_0 and $\mathbb{P}$-a.e. σ we have that

$$\lim_n 1/n \sum_{i \le n} S_i f(x_i) = \int_{\mathbb{X}} f(x)\Psi(x) d\mu(x),$$

where convergence is in the Hausdorff metric.

Proof. To prove this theorem we are going to use the same arguments as in the proof of Theorem 3.1 . It is natural to try to use $\Omega = \mathbb{X} \times \mathcal{K}_c$, but this is not locally compact as required by Elton's theorem. We choose a suitable locally compact subset of $\mathbb{X} \times \mathcal{K}_c$ instead. As in the proof of Theorem 3.1, there is a subset $\Theta \subseteq \Lambda$ with $\mathbb{P}(\Theta) = 1$ and the address map $a_\Phi : \Theta \to \mathcal{K}_c$ exists. Notice that Θ can be chosen to be invariant under the left shift and all the right shifts, that is $\sigma \mapsto (i, \sigma)$, on Λ. Let Γ be the closure in $\mathcal{K}_c$ of the set

$$S_0 \cup \bigcup \{\Phi_{\sigma_1} \circ \Phi_{\sigma_2} \circ \cdots \circ \Phi_{\sigma_n}(S_0) : n \in \mathbb{N}, \sigma \in \Theta\}.$$

This implies that Γ is a compact subset of $\mathcal{K}_c$. The important feature of Γ is that it is closed under the action of each Φ_i. Let $\Omega = \mathbb{X} \times \Gamma$ be equipped with the metric

$$d(\,(x_1, D_1), (x_2, D_2)\,) = d(x_1, x_2) + d_H(D_1, D_2).$$

We note that Ω is locally compact, separable and complete. Define on Ω the family of contractions $\Upsilon_i : \Omega \to \Omega$ specified by $\Upsilon_i(x, D) = (w_i(x), \Phi_i(D))$. Then the IFS $\{\Upsilon, p_i\}$ is average contractive on Ω. Let A be the attractor of the IFS $\{w_i\}$ on $\mathbb{X}$. Then it is easy to see that the invariant measure η of $\{\Upsilon_i, p_i\}$ is supported on the set

$$\Delta = \{(x, \Psi(x)) : x \in a_w(\Theta)\}$$

onto the measure μ on $\mathbb{X}$. For each $q \in S^{d-1}$, define $\xi_q : \Omega \to \mathbb{R}$ by

$$\xi_q(x, D) = \begin{cases} f(x) \cdot \mathrm{supp}(q, D), & \text{if } D \subset 2K \\ 2f(x) \cdot \mathrm{supp}(q, K), & \text{otherwise.} \end{cases}$$

Notice that ξ_q is continuous and bounded for each q. It is easy to check that for any $(x, D) \in \Delta$, $\xi_q(x, D) = f(x) \cdot \mathrm{supp}(q, D)$, since $D \subset K$. We thus obtain

$$\begin{aligned}
\mathrm{supp}\left(q, 1/n \sum_{i \le n} f(x_i) S_i\right) &= 1/n \sum_{i \le n} f(x_i)\mathrm{supp}(q, S_i) = 1/n \sum_{i \le n} f(x_i)\mathrm{supp}(q, \Psi(x_i)) \\
&= 1/n \sum_{i \le n} \xi_q(x_i, \Psi(x_i)) \\
&\to \int_\Omega \xi_q(\omega)\; d\theta(\omega) = \int_{\mathbb{X}} f(x) \cdot \mathrm{supp}(q, \Psi(x))\; d\mu(x) \\
&= \mathrm{supp}\left(q, \int_{\mathbb{X}} f(x) \cdot \Psi(x)\; d\mu(x)\right).
\end{aligned}$$

Now by using the same idea as in the proof of Theorem 4.2, we easily get that

$$1/n \sum_{i \le n} f(x_i) S_i \to \int_{\mathbb{X}} f(x) \cdot \Psi(x) \; d\mu(x)$$

and the convergence is understood in the Hausdorff sense. $\square$

The following result follows from the previous one once we separate the sum into the two parts corresponding to when $f(x_i) < 0$ and when $f(x_i) \ge 0$.

Theorem 4.5. *Let $\{w_i, p_i\}$ be a contractive IFSP on $\mathbb{X}$ and $\{w_i, \alpha_i, F_0\}$ be an IFS on multifunctions with $\sum_i p_i \|\alpha_i\| < 1$. Choose $x_0 \in \mathbb{X}$, $S_0 \in \mathcal{K}_c$, and define $x_{n+1} = w_{\sigma_n}(x_n)$ and $S_{n+1} = \Phi_{\sigma_n}(S_n)$ where $\sigma_n \in \{1, 2, \ldots, N\}$ is a sequence of independent and identically distributed random variables chosen according to the probabilities $\{p_i\}$. Finally, suppose that $f : \mathbb{X} \to \mathbb{R}$ is a bounded and continuous function.*

Then for all x_0 and $\mathbb{P}$-a.e. σ we have that

$$\lim_n 1/n \sum_{i \le n} S_i f(x_i) = \int_{\mathbb{X}} f^+(x) \Psi(x) d\mu(x) - \int_{\mathbb{X}} f^-(x) \Psi(x) d\mu(x),$$

where the above convergence is in the Hausdorff metric.

We now compare Theorem 4.2 and its corollary to the classical situation of continuous functions. Corresponding to an IFSP there is a Markov operator $\mathcal{M}$ mapping probability measures to probability measures; this Markov operator possesses an adjoint which acts on continuous functions via the duality induced by integration, that is

$$\int_{\mathbb{X}} f(x) \; d\mathcal{M}\nu(x) = \int_{\mathbb{X}} f(x) \; d\left(\sum_i p_i \nu \circ w_i^{-1} \right)(x) =$$

$$\int_{\mathbb{X}} \left(\sum_i p_i f(w_i(y)) \right) \; d\nu(y) = \int_{\mathbb{X}} \mathcal{U}(f)(y) \; d\nu(y).$$

Clearly $\mathcal{U}$ maps continuous functions to continuous functions. In the case of multifunction there is an analogous adjoint operator given by

$$\mathcal{U}(F)(x) = \sum_i p_i F(w_i(x)).$$

If we denote by μ the fixed point of the Markov operator $\mathcal{M}$, we have that

$$\int_{\mathbb{X}} F(x) \; d\mu(x) = \sum_i p_i \int_{\mathbb{X}} F(w_i(x)) \; d\mu(x) = \sum_\sigma p_\sigma \int_{\mathbb{X}} F(w_\sigma(x)) \; d\mu(x),$$

where the last sum is done over all $\sigma \in \{1, 2, \ldots, N\}^L$ for finite L. Notice that we are using the notation

$$p_\sigma = p_{\sigma_1} p_{\sigma_2} \cdots p_{\sigma_L}, \quad w_\sigma = w_{\sigma_1} \circ w_{\sigma_2} \circ \cdots \circ w_{\sigma_L}.$$

The last sum above can be also interpreted as the integral of F with respect to the measure

$$\mu_L = \sum_{\sigma} p_\sigma \delta_{w_\sigma(x)},$$

which is a finite convex combination of point masses. By construction μ_L converges weakly to μ and so the integrals will also converge.

In this way, Theorem 4.2 and its corollary state that instead of considering all possible finite products of the p_i and all possible finite compositions of the w_i it is sufficient to take only a single infinite trajectory as long as the individual maps are chosen correctly.

5 Some computational considerations

To determine an approximated value of the sum $1/n \sum_i F(x_i)$ can be quite complicated in practice. This is true even in the simplest situation in which we assume that $F(x)$ is a polyhedral set for each x; in fact an infinite convex combination of polyhedral sets is not polyhedral in general. One alternative way to approximate this sum consists of choosing a finite number of points $q_j \in S^{d-1}$, $j = 1, 2, \ldots, L$, which are nicely distributed and then compute $1/n \sum_i \operatorname{supp}(q_j, F(x_i))$ (this is a scalar for each q_j). Using these quantities we can then construct the approximation

$$1/n \sum_{i \le n} F(x_i) \subseteq \bigcap_{j=1}^{L} \{z \in \mathbb{R}^d : z \cdot q_j \le 1/n \sum_{i \le n} \operatorname{supp}(q_j, F(x_i))\}. \tag{18}$$

The larger set on the right side of the above inclusion is given by the intersection and it turns out that it is always polyhedral. Further, the higher the number of points q_j, the better the approximation. In fact, since the support function of a compact and convex set is a convex function of the direction, it is even possible to obtain bounds on the error in the Hausdorff distance. To see this, take the points q_j to form an ϵ-net on S^{d-1}. This means that all points on S^{d-1} are within a distance of at most ϵ from some q_j. Since F is bounded, there is a compact and convex set K with $F(x) \subseteq K$ for all x. As in the above results, let $M = \sup_{p \in S^{d-1}} \operatorname{supp}(p, K)$ which then implies that the function $p \mapsto \operatorname{supp}(p, K)$ defined on S^{d-1} has Lipschitz factor at most M. Furthermore, all the support functions for both the sums and for the limiting integral have Lipschitz factor at most M. Let now $p \in S^{d-1}$ be arbitrary and q_j be within a distance of ϵ from p. Then

$$\left| \operatorname{supp}(p, 1/n \sum_{i \le n} F(x_i)) - \operatorname{supp}(q_j, 1/n \sum_{i \le n} F(x_i)) \right| \le \epsilon M$$

and thus

$$d_H(1/n \sum_{i \le n} F(x_i), \ \bigcap_{j=1}^{L} \{z \in \mathbb{R}^d : z \cdot q_j \le 1/n \sum_{i \le n} \operatorname{supp}(q_j, F(x_i))\}) \le 2\epsilon M.$$

The approximation from (18) is within a Hausdorff distance of $2\epsilon M$ from the true set and by increasing the number of points we can make this error as small as we wish.

Example 5.1. We continue the discussion of Example 2.4 in the context of approximating the integral in Theorem 4.4. Since the multifunction F_0 take values in $\mathbb{R}^2$, we need to choose $q_j \in S^1$ which are "nicely distributed," by which we mean fairly uniformly distributed with small spacing. In this instance, we do this by choosing $q_j = (2\pi j)/\ell$ for $j = 0, 1, \ldots, \ell - 1$ and for some large enough $\ell \in \mathbb{N}$. We have $\|F_0(x)\| = \|[0,1]\| = \sqrt{2}$ for all x and thus because the contraction factors of both α_i are equal to $1/2$, $\|\Psi(x)\| \leq 2\sqrt{2}$ for any x. Thus we have M (from the above discussion) to be $M = 2$.

In particular, what this means is that we keep track of ℓ different values ρ_j, one for each direction q_j. We initialize $\rho_j^0 = \mathrm{supp}(q_j, S_0)$. At each step the "true" update to ρ_j^n should be

$$\rho_j^{n+1} = \rho_j^n + \mathrm{supp}(q_j, \alpha_{\sigma_n}(S_n) + F_0(x_n)) = \rho_j^n + \frac{1}{2}\mathrm{supp}(q_j, R_{\theta_{\sigma_n}} S_n) + \mathrm{supp}(q_j, [-1,1]^2),$$

where R_θ is a rotation of angle θ around the origin in $\mathbb{R}^2$. However, as we are not storing the sets S_n but only the values ρ_j^n, we can't compute this update exactly. In particular, we cannot compute $\mathrm{supp}(q_j, R_{\theta_{\sigma_n}} S_n) = \mathrm{supp}(R_{-\theta_{\sigma_n}} q_j, S_n)$ exactly if $\sigma_n = 2$ because $R_{-\theta_2} q_j$ will be in a direction "in between" the directions of two successive q_i's. Thus we have to approximate this value as well. The support function is convex and positively homogeneous as a function of the direction, so a linear interpolation is a reasonable first approximation. For a given j, there is some $\gamma(j) \in \{0, 1, \ldots, \ell\}$ with

$$\frac{2\pi\gamma(j)}{\ell} < \frac{2\pi j}{\ell} - \frac{1}{100} < \frac{2\pi(\gamma(j)+1)}{\ell}.$$

The approximated update to ρ_j^n is

$$\rho_j^{n+1} = \rho_j^n + \frac{1}{2}\rho_{\gamma(j)}^n + \mathrm{supp}(q_j, [-1,1]^2).$$

After some specified number of iterations L, the approximation to the integral is given by the set

$$\bigcap_{j=1}^{\ell} \{x \in \mathbb{R}^2 : x \cdot q_j \leq \rho_j^L / L\}.$$

Thus we see that the approximation is a polygon with at most ℓ sides.

Another alternative approach consists of approximating the functions $x \mapsto \mathrm{supp}(q_m, F(x))$ by polynomials and then using the recursive formula for the moments of the invariant measure for an IFS (see [4, 11]). This will give exact values for the polynomials and they will then approximate the set-valued integral. Unfortunately, extreme care must be taken with this approach since methods based on moments are often numerically unstable and give poor results in practice.

Acknowledgments

FM gratefully acknowledges support for this work from the Natural Sciences and Engineering Research Council of Canada (NSERC).

References

[1] Z. Artstein, Set-valued measures, *Trans AMS*, 165 (1927), 103-125.

[2] J.P. Aubin and H. Frankowska, *Set-Valued Analysis*, Birkhauser, Boston (1990).

[3] M. F. Barnsley, *Fractals Everywhere*, Academic Press, New York (1988).

[4] M.F. Barnsley and S. Demko, Iterated function systems and the global construction of fractals, *Proc. Roy. Soc. London* 399 (1985), 243-275.

[5] C. Choirat, C. Hess and R. Seri, A functional version of the Birkhoff ergodic theorem for a normal integrand: a variational approach, *Ann. Prob.* 31 (2003), no.1, 63-92.

[6] F. den Hollander, *Large Deviations*, AMS, Providence, RI (2000).

[7] J. Elton, An Ergodic Theorem for Iterated Maps, *Journal of Ergodic Theory and Dynamical Systems* 7 (1987), 481-488.

[8] J. Elton, A multiplicative ergodic theorem for Lipschitz maps, *Stochastic Process. Appl.* 34 (1990), no. 1, 39-47.

[9] K. Falconer, *Fractal Geometry - Mathematical Foundations and Applications*, John Wiley, Second Edition (2003).

[10] B. Forte and E.R. Vrscay, Solving the inverse problem for functions and image approximation using iterated function systems, *Dyn. Cont. Impul. Sys.* 1 (1995), 177-231.

[11] B. Forte and E.R. Vrscay, Solving the inverse problem for measures using iterated function systems: A new approach, *Adv. Appl. Prob.* 27 (1995), 800-820.

[12] B. Forte, F. Mendivil and E.R. Vrscay, "Chaos Games" for Iterated Function Systems with grey level maps, *SIAM J. Math. Anal.* 29 (1998), no.4, 878-890.

[13] H. Kunze, D. La Torre, F. Mendivil, and E.R. Vrscay, *Fractal-based methods in analysis*, Springer, Berlin (2012).

[14] F. Hiai, Radon-Nikodym Theorems for Set-Valued Measures, *Journal of Multivariate Analysis* 8 (1978), 96-118.

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

[16] D. La Torre and F. Mendivil, IFS Markov operators on multimeasures, submitted.

[17] D. La Torre and F. Mendivil, Iterated Function Systems on Multifunctions and Inverse Problems, *J. Math. Anal. App.*, 340 (2008), 1469-1479.

[18] D. La Torre and F. Mendivil, Minkowski-additive multimeasures, monotonicity and self-similarity, *Image Anal. Stereology*, 30 (2001), no. 3, 135-142.

[19] F. Mendivil and J. Silver, Chaos games for wavelet analysis and wavelet synthesis, *Comput. Math. Appl.* 46 (2003), no. 5-6, 863-873.

[20] F. Mendivil and E.R. Vrscay, Fractal vector measures and vector calculus on planar fractal domains, *Chaos, Solitons and Fractals* 14 (2002), 1239-1254.

[21] R. T. Rockafellar, R. J-B. Wets, *Variational Analysis*, Springer (1998).