

# ON THE CONVERGENCE RATE OF THE CHAOS GAME

BALÁZS BÁRÁNY, NATALIA JURGA, AND ISTVÁN KOLOSSVÁRY

ABSTRACT. This paper studies how long it takes the orbit of the chaos game to reach a certain density inside the attractor of a strictly contracting iterated function system of which we only assume that its lower dimension is positive. We show that the rate of growth of this cover time is determined by the Minkowski dimension of the push-forward of the shift invariant measure with exponential decay of correlations driving the chaos game. Moreover, we bound the expected value of the cover time from above and below with multiplicative logarithmic correction terms. As an application, for Bedford-McMullen carpets we completely characterise the family of probability vectors which minimise the Minkowski dimension of Bernoulli measures. Interestingly, these vectors have not appeared in any other aspect of Bedford-McMullen carpets before.

## 1. INTRODUCTION

The *chaos game* is a simple random iterative procedure introduced by Barnsley [4] to generate the attractor of an *iterated function system* (IFS) which is a tuple  $\mathcal{F} = \{f_1, f_2, \dots, f_N\}$  of contracting transformations on  $\mathbb{R}^d$ . Given a fixed non-degenerate probability vector  $\mathbf{p} = (p_1, \dots, p_N)$ , the game starts from an initial point  $x_0$ , chooses a random function  $f_{i_1}$  from  $\mathcal{F}$  according to  $\mathbf{p}$  and returns the point  $x_1 = f_{i_1}(x_0)$ . It continues iteratively generating a sequence of points  $x_2, x_3, \dots$ , where  $x_n = f_{i_n}(x_{n-1})$  and the indices  $i_n$  are independently and identically distributed (i.i.d.) according to  $\mathbf{p}$ . The *attractor*  $\Lambda \subseteq \mathbb{R}^d$ , which is the unique non-empty, compact set satisfying  $\Lambda = \bigcup_{i=1}^N f_i(\Lambda)$  is then almost surely (regardless of the starting point) obtained as the  $\omega$ -limit set of the orbit  $(x_n)_{n=0}^\infty$ , i.e.  $\Lambda = \bigcap_{m=1}^\infty \overline{\{x_n : n \geq m\}}$ . The almost sure property is with respect to the Bernoulli measure  $\mu_{\mathbf{p}} = \mathbf{p}^{\mathbb{N}}$  defined on infinite sequences  $\mathbf{i}$  from the symbolic space  $\Sigma = \{1, 2, \dots, N\}^{\mathbb{N}}$ . In particular, if  $x_0 \in \Lambda$  (for example,  $x_0$  is a fixed point of one of the  $f_i$ ) then  $(x_n)_{n=0}^\infty$  is almost surely a dense subset of  $\Lambda$ . For the remainder we will always assume that  $x_0 \in \Lambda$ .

The chaos game can be naturally generalised to allow the indices  $i_n$  to be chosen in a non-i.i.d. way. The left shift  $\sigma : \Sigma \rightarrow \Sigma$  is defined  $\sigma(i_1 i_2 i_3 \dots) = i_2 i_3 \dots$  and a measure  $\mu$  on the symbolic space  $\Sigma$  is  $\sigma$ -invariant if  $\mu(\sigma^{-1}(\cdot)) = \mu(\cdot)$ . Then we also consider the chaos game with respect to a  $\sigma$ -invariant measure  $\mu$ , where for every  $n \in \mathbb{N}$  and  $i_1, \dots, i_n \in \{1, \dots, N\}^n$ , the first  $n$  indices in the chaos game are chosen to be  $i_1, \dots, i_n$  with probability  $\mu([i_1 \dots i_n])$ , where  $[i_1 \dots i_n]$  denotes the cylinder set  $\{\mathbf{j} \in \Sigma : j_l = i_l \text{ for } l = 1, \dots, n\}$ .

---

2010 *Mathematics Subject Classification*. Primary 28A80 Secondary 37A50 37C35.

*Key words and phrases*. chaos game, Minkowski dimension of measures, cover time, iterated function system, Bedford–McMullen carpet.

Balázs Bárány acknowledges support from grants OTKA K123782 and OTKA FK134251. 759/1). Natalia Jurga was supported by an EPSRC Standard Grant (EP/R015104/1). István Kolossváry was supported by a *Leverhulme Trust Research Project Grant* (RPG-2019-034).

The natural question we address in this paper is: given  $\mu$ , how long does it take the chaos game to reach a certain ‘resolution’ of the fractal? Moreover, for which  $\mu$  will it take the least number of steps to reach that resolution? We formulate these questions rigorously in a moment. We are only aware of [21] and [27] that study this direction. The former is an informal investigation, while the latter work of Morris and the second named author answer these questions for self-similar sets satisfying the open set condition in the i.i.d case. Our results greatly extend this scope and give additional context explaining and illustrating the results. A number of papers show convergence of the chaos game for more general IFSs, see [5–7, 33] for example, while others show ergodic theorems for time averages along the orbit [10, 16, 44]. We do not pursue these directions further here.

The question addressed in this article is a type of “covering problem”, which are problems with a rich history in the probability literature, originating with the classical coupon collector problem. Covering problems are concerned with the time taken for a stochastic process to “exhaust” its state space, in a precise sense which depends on the model being studied. Classical examples include the first time that a random walk on a finite path-connected graph visits all of the vertices [2] and the first time that an irreducible finite-state Markov chain visits all of its states [34, 36]. The geometric covering problems studied in [1, 38] are closely related to our problem, except that while our sequence of points  $x_1, x_2, x_3 \dots$  is determined in a dynamical way through the chaos game, the points belonging to the sequences in [1, 38] are chosen at random in an i.i.d. way. Another closely related field is the study of recurrence in chaotic dynamical systems through hitting time statistics [22, 35]. This is concerned with the study of the hitting time of an orbit under a chaotic dynamical system, to a ball which is shrinking down to a point in the state space.

**Setup.** Throughout, we assume that the IFS  $\mathcal{F}$  is strictly contracting in the sense that there exists  $0 < a < 1$  such that

$$\|f_i(x) - f_i(y)\| \leq a\|x - y\| \text{ for every } i = 1, \dots, N. \quad (1.1)$$

A priori we do not assume any separation condition on the first level cylinder sets  $f_i(\Lambda)$  nor any further smoothness conditions on the maps  $f_i$ .

The *orbit* of the starting point  $x_0 \in \Lambda$  according to  $\mathbf{i} \in \Sigma$  for the first  $n$  steps is

$$\mathcal{O}_n(\mathbf{i}, x_0) = \{x_0, f_{i_1}(x_0), f_{i_2} \circ f_{i_1}(x_0), \dots, f_{i_n} \circ \dots \circ f_{i_2} \circ f_{i_1}(x_0)\}.$$

For compositions of maps we use the standard notation  $f_{\mathbf{i}|n} := f_{i_1} \circ f_{i_2} \circ \dots \circ f_{i_n}$ , where  $\mathbf{i}|n = i_1 i_2 \dots i_n$ . Compositions are taken in reverse order in  $\mathcal{O}_n(\mathbf{i}, x_0)$ , for this we introduce the notation  $\overleftarrow{\mathbf{i}}|n := i_n i_{n-1} \dots i_1$  and so  $f_{\overleftarrow{\mathbf{i}}|n} = f_{i_n} \circ f_{i_{n-1}} \circ \dots \circ f_{i_1}$ . We measure the ‘resolution’ of  $\mathcal{O}_n(\mathbf{i}, x_0)$  as a subset of  $\Lambda$  according to the Hausdorff distance  $d_H$  between them. In particular, for any  $r > 0$

$$d_H(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda) < r \iff (\forall x \in \Lambda) (\exists y \in \mathcal{O}_n(\mathbf{i}, x_0)) \text{ such that } x \in B(y, r).$$

In this case, we also say that  $\mathcal{O}_n(\mathbf{i}, x_0)$  is *r-dense* in  $\Lambda$ . To measure the first instance when  $\mathcal{O}_n(\mathbf{i}, x_0)$  becomes *r-dense* in  $\Lambda$ , we introduce the waiting time

$$T_r(\mathbf{i}, x_0) := \inf\{n \geq 0 : d_H(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda) < r\}.$$

Let  $\mu$  be a left-shift invariant measure. We say that  $\mu$  has *one-sided exponential decay* if there exists  $\varepsilon > 0$  and  $\kappa > 0$  such that for any finite words  $\bar{i}, \bar{j} \in \Sigma^*$

$$\mu([\bar{i}] \cap \sigma^{-n}[\bar{j}]) \leq (1 + \kappa 2^{-\varepsilon(n-|\bar{i}|)})\mu([\bar{i}])\mu([\bar{j}]). \quad (1.2)$$

Such a measure  $\mu$  is necessarily strongly mixing. To see this, note that (1.2) immediately implies that for all finite words  $\bar{i}, \bar{j} \in \Sigma^*$ ,  $\mu$  satisfies the one-sided strong mixing property  $\limsup_{n \rightarrow \infty} \mu([\bar{i}] \cap \sigma^{-n}[\bar{j}]) \leq \mu([\bar{i}])\mu([\bar{j}])$ . The other side can then be obtained by applying the one-sided strong mixing property to  $\mu((\Sigma \setminus [\bar{i}]) \cap \sigma^{-n}[\bar{j}])$ . It is easy to see that Bernoulli measures satisfy the one-sided exponential decay property (1.2), but it is also satisfied by many other measures, such as Gibbs measures for Hölder continuous potentials on  $\Sigma$  and certain Käenmäki measures, both of which are important in the analysis of fractals and motivate the study of the chaos game beyond the i.i.d. case. A more detailed discussion of the applicability of our results to these classes of measures can be found in Section 2.1.

We define the *reversed measure of a  $\sigma$ -invariant measure  $\mu$* , denoted by  $\overleftarrow{\mu}$ , as

$$\overleftarrow{\mu}([\bar{i}]) = \mu([\overleftarrow{\bar{i}}]) \text{ for every } \bar{i} \in \Sigma^*.$$

Since  $\mu$  is  $\sigma$ -invariant it follows that  $\overleftarrow{\mu}$  is a well defined  $\sigma$ -invariant measure on  $\Sigma$ . We let  $\overleftarrow{\nu} := \pi_* \overleftarrow{\mu}$  where  $\pi_* \overleftarrow{\mu}$  is the pushforward measure of  $\overleftarrow{\mu}$  through the natural projection  $\pi : \Sigma \rightarrow \Lambda$  defined by

$$\pi(\mathbf{i}) := \lim_{n \rightarrow \infty} f_{\mathbf{i}|n}(0). \quad (1.3)$$

**Main contribution.** Informally, our main result states that if  $\Lambda$  has positive lower dimension  $\dim_{\text{L}} \Lambda > 0$ , see (2.1) for a definition, and  $\mu$  is an invariant measure with one-sided exponential decay then  $T_r(\mathbf{i}, x_0)$  asymptotically scales as  $r^{-\alpha(1+o(1))}$  as  $r \rightarrow 0$ , where the exponent  $\alpha$  is given by the Minkowski dimension of  $\overleftarrow{\nu}$  as defined by Falconer, Fraser and Käenmäki in [14], see Section 1.1 below. We obtain more precise bounds for the expected value of  $T_r(\mathbf{i}, x_0)$  with respect to  $\mu$  by showing that

$$(r(\log(1/r))^{2/\dim_{\text{L}} \Lambda})^{-\alpha+o(1)} \leq \mathbb{E}_{\mu} T_r(x_0) \leq (\log(1/r))^2 r^{-\alpha-o(1)}.$$

Moreover, we give a complete characterisation of the probability vectors  $\mathbf{p}$  which minimise the value of  $\alpha$  for the pushforward measure  $\nu_{\mathbf{p}}$  in the case of Bedford–McMullen carpets. See Section 2 for the precise formulation of these statements.

**1.1. Minkowski dimension of measures.** Let  $\nu$  be a fully supported finite Borel measure on a compact metric space  $X$ . Then the *upper Minkowski dimension* of  $\nu$  is defined as

$$\overline{\dim}_{\text{M}}(\nu) := \inf\{s \geq 0 : \text{there exists a constant } c > 0 \text{ such that} \\ \nu(B(x, r)) \geq cr^s \text{ for all } x \in X \text{ and } 0 < r < 1\}$$

and its *lower Minkowski dimension* is

$$\underline{\dim}_{\text{M}}(\nu) := \inf\{s \geq 0 : \text{there exist a constant } c > 0 \text{ and a sequence } (r_n)_{n \in \mathbb{N}} \\ \text{of positive real numbers such that } \lim_{n \rightarrow \infty} r_n = 0 \text{ and} \\ \nu(B(x, r_n)) \geq cr_n^s \text{ for all } x \in X \text{ and } n \in \mathbb{N}\}$$

(here  $B(x, r)$  denotes an open ball). If the two values coincide, then the common value, called the *Minkowski dimension* of  $\nu$ , is denoted by  $\dim_{\text{M}} \nu$ . It is also referred to as the box dimension of  $\nu$ , see [18, Chapter 4.2].

The upper and lower Minkowski dimension of a set  $X$  is denoted by  $\overline{\dim}_M X$  and  $\underline{\dim}_M X$ , respectively. The main result of [14] is to give an alternative characterization of these quantities by showing that

$$\overline{\dim}_M X = \min \{ \overline{\dim}_M(\nu) : \nu \text{ is a fully supported finite Borel measure on } X \}, \quad (1.4)$$

and an analogous claim holds for  $\underline{\dim}_M X$ . In general, for the attractor of an IFS, the measure that achieves the minimum is not invariant. However, an interpretation of the result in [27] is that for a self-similar set  $\Lambda$  satisfying the open set condition,  $\dim_M \Lambda$  is achieved by a self-similar measure  $\nu_{\mathbf{p}}$  if and only if  $\mathbf{p}$  is the so-called ‘natural measure’ associated to the IFS generating  $\Lambda$ . On the other hand, in the case of a family of self-affine sets called *Bedford–McMullen carpets* there is indeed a positive gap  $\dim_M \Lambda < \min_{\mathbf{p}} \dim_M \nu_{\mathbf{p}}$  unless the Hausdorff and Minkowski dimensions of  $\Lambda$  coincide, see Remark 2.6. A key contribution of the current paper is to identify the family of vectors  $\mathbf{p}$  which minimise  $\dim_M \nu_{\mathbf{p}}$ . Hence, these are the vectors for which the chaos game reaches a certain ‘resolution’ in the least number of steps.

Here we give a simple, yet useful equivalent characterization of  $\dim_M \nu$ . Let

$$\underline{\alpha}(\nu) := \liminf_{r \rightarrow 0} \max_{y \in \Lambda} \frac{\log \nu(B(y, r))}{\log r} \quad \text{and} \quad \overline{\alpha}(\nu) := \limsup_{r \rightarrow 0} \max_{y \in \Lambda} \frac{\log \nu(B(y, r))}{\log r}.$$

We note that  $\min_{y \in \Lambda} \nu(B(y, r))$  exists for each  $r$  by compactness of  $\Lambda$  and lower semi-continuity of  $y \mapsto \nu(B(y, r))$ , thus the above two definitions make sense.

**Lemma 1.1.** *Let  $\nu$  be a compactly supported finite Borel measure. Then*

$$\underline{\alpha}(\nu) = \underline{\dim}_M(\nu) \quad \text{and} \quad \overline{\alpha}(\nu) = \overline{\dim}_M(\nu).$$

*Proof.* The proof follows from the definition of the upper and lower Minkowski dimensions. Let  $X$  be the support of  $\nu$ . Let  $s > \overline{\dim}_M(\nu)$  be arbitrary. Then for every  $x \in X$ ,  $\frac{\log \nu(B(x, r))}{\log r} \leq s + \frac{\log c}{\log r}$  and so,  $\overline{\alpha}(\nu) \leq s$ . Since  $s$  was arbitrary, we have  $\overline{\alpha}(\nu) \leq \overline{\dim}_M(\nu)$ .

Now, let  $s < \overline{\dim}_M(\nu)$ . Then by definition for every  $c > 0$  there exists  $x_c \in X$  and  $0 < r_c < 1$  such that  $\nu(B(x_c, r_c)) < c(r_c)^s$ . Hence,

$$\max_{y \in X} \frac{\log \nu(B(y, r_c))}{\log r_c} \geq \frac{\log \nu(B(x_c, r_c))}{\log r_c} > s + \frac{\log c}{\log r_c}.$$

Letting  $c \rightarrow 0$ , necessarily we have that  $r_c \rightarrow 0$ , and since  $\frac{\log c}{\log r_c} \geq 0$ , we get  $\overline{\alpha}(\nu) \geq s$ . Since  $s$  was again arbitrary, we have  $\overline{\alpha}(\nu) \geq \overline{\dim}_M(\nu)$ .

The proof of the other equality is similar. Let  $s > \underline{\dim}_M(\nu)$ . Then there exist  $c > 0$  and a sequence  $r_n$  with  $r_n \rightarrow 0$  such that

$$\liminf_{r \rightarrow 0} \max_{x \in X} \frac{\log \nu(B(x, r))}{\log r} \leq \liminf_{n \rightarrow \infty} \max_{x \in X} \frac{\log \nu(B(x, r_n))}{\log r_n} \leq \liminf_{n \rightarrow \infty} \left( s + \frac{\log c}{\log r_n} \right) = s.$$

Now, let  $s < \underline{\dim}_M(\nu)$ . Then for every  $c > 0$  and every sequence  $r_n$  with  $r_n \rightarrow 0$  there exists  $x \in X$  such that  $\nu(B(x, r_n)) < c(r_n)^s$ . Hence, choosing  $r_n$  to be the sequence for which  $\underline{\alpha}(\nu)$  is achieved, we get that

$$\underline{\alpha}(\nu) = \lim_{n \rightarrow \infty} \max_{y \in X} \frac{\log \nu(B(y, r_n))}{\log r_n} \geq \lim_{n \rightarrow \infty} s + \frac{\log c}{\log r_n} = s.$$

□

The relationship of the Minkowski dimension of  $\nu$  to other well-studied notions becomes more apparent from this formulation. Recall that the lower and upper local dimensions of a measure  $\nu$  at a point  $x$  are

$$\underline{\dim}_{\text{loc}}(\nu, x) = \liminf_{r \rightarrow 0} \frac{\log \nu(B(x, r))}{\log r} \quad \text{and} \quad \overline{\dim}_{\text{loc}}(\nu, x) = \limsup_{r \rightarrow 0} \frac{\log \nu(B(x, r))}{\log r}.$$

The local dimension looks at the measure of a ball around a fixed point  $x$ , while the Minkowski dimension always takes the point  $y_r$  at scale  $r$  for which  $\nu(B(y_r, r))$  is minimal, the ‘least accessible part’ of the attractor. For self-similar sets taking a point  $x^*$  which maximises  $\underline{\dim}_{\text{loc}}(\nu_{\mathbf{p}}, x)$  automatically provides a sequence  $y_r \equiv x^*$  that gives  $\dim_{\text{M}} \nu_{\mathbf{p}}$ . However, for the Bedford–McMullen carpets we will show that  $\dim_{\text{M}} \nu_{\mathbf{p}}$  can be strictly larger than  $\max_x \underline{\dim}_{\text{loc}}(\nu_{\mathbf{p}}, x)$ , see Remark 5.2. Another related concept is the quantization problem for probability measures, see [31] for some background.

## 2. MAIN RESULTS

We begin with the asymptotic pointwise almost sure behaviour of  $T_r(\mathbf{i}, x_0)$ . Let  $N_r(X)$  denote the smallest number of open sets of diameter  $r$  required to cover the set  $X$ . The *lower dimension* of  $X$  is

$$\dim_{\text{L}} X = \sup \left\{ \alpha : (\exists C > 0) \text{ such that } (\forall 0 < r < R \leq |X| \text{ and } x \in X) \right. \\ \left. N_r(B(x, R) \cap X) \geq C(R/r)^\alpha \right\}. \quad (2.1)$$

Recall,  $\overleftarrow{\nu} = \pi_* \overleftarrow{\mu}$  is the pushforward measure of  $\overleftarrow{\mu}$  through the natural projection  $\pi$  defined in (1.3). If  $\mu$  is a Bernoulli measure  $\mu = \mu_{\mathbf{p}}$ , then since the  $\mu_{\mathbf{p}}$  measure of a cylinder set is independent of the order of digits, we clearly have  $\overleftarrow{\mu}_{\mathbf{p}} = \mu_{\mathbf{p}}$ . Thus, in case of the measure  $\nu_{\mathbf{p}} = \pi_* \mu_{\mathbf{p}}$ , which equivalently is the unique measure that satisfies

$$\nu_{\mathbf{p}}(\cdot) = \sum_{i=1}^N p_i \nu_{\mathbf{p}}(f_i^{-1}(\cdot)), \quad (2.2)$$

one can replace  $\overleftarrow{\nu}$  by  $\nu_{\mathbf{p}}$ .

**Theorem 2.1.** *Let  $\mathcal{F}$  be an arbitrary IFS which satisfies (1.1) and whose attractor  $\Lambda$  satisfies  $\dim_{\text{L}} \Lambda > 0$ . Then for any measure  $\mu$  on  $\Sigma$  with one-sided exponential decay of correlations (1.2) and for  $\mu$ -a.e.  $\mathbf{i}$  and every  $x_0 \in \Lambda$*

$$\liminf_{r \rightarrow 0} \frac{\log T_r(\mathbf{i}, x_0)}{-\log r} = \underline{\dim}_{\text{M}}(\overleftarrow{\nu}) \quad \text{and} \quad \limsup_{r \rightarrow 0} \frac{\log T_r(\mathbf{i}, x_0)}{-\log r} = \overline{\dim}_{\text{M}}(\overleftarrow{\nu}).$$

*In particular, if  $\mu = \mu_{\mathbf{p}}$  is Bernoulli, then one can replace  $\overleftarrow{\nu}$  by  $\nu_{\mathbf{p}}$ .*

Theorem 2.1 can easily be flipped over to obtain the decay rate of  $d_{\text{H}}(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda)$ .

**Corollary 2.2.** *Let  $\mathcal{F}$  be an arbitrary IFS which satisfies (1.1) and whose attractor  $\Lambda$  satisfies  $\dim_{\text{L}} \Lambda > 0$ . Then for any measure  $\mu$  with one-sided exponential decay of correlations (1.2) and for  $\mu$ -a.e.  $\mathbf{i}$  and every  $x_0 \in \Lambda$*

$$\liminf_{n \rightarrow \infty} \frac{\log d_{\text{H}}(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda)}{-\log n} = \frac{1}{\underline{\dim}_{\text{M}}(\overleftarrow{\nu})} \quad \text{and} \quad \limsup_{n \rightarrow \infty} \frac{\log d_{\text{H}}(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda)}{-\log n} = \frac{1}{\overline{\dim}_{\text{M}}(\overleftarrow{\nu})}.$$

*Again, if  $\mu = \mu_{\mathbf{p}}$  is Bernoulli, then one can replace  $\overleftarrow{\nu}$  by  $\nu_{\mathbf{p}}$ .*

*Proof.* From the definition of  $T_r(\mathbf{i}, x_0)$  it follows that  $r \leq d_H(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda)$  for every  $n < T_r(\mathbf{i}, x_0)$ , in particular, we have  $n \leq T_{d_H(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda)/2}(\mathbf{i}, x_0)$ . This and Theorem 2.1 imply that

$$\begin{aligned} \frac{1}{\dim_M(\overleftarrow{\nu})} &= \liminf_{r \rightarrow 0} \frac{-\log r}{\log T_r(\mathbf{i}, x_0)} \geq \liminf_{r \rightarrow 0} \frac{-\log d_H(\mathcal{O}_{T_r(\mathbf{i}, x_0)-1}(\mathbf{i}, x_0), \Lambda)}{\log T_r(\mathbf{i}, x_0)} \\ &\geq \liminf_{n \rightarrow \infty} \frac{-\log d_H(\mathcal{O}_{n-1}(\mathbf{i}, x_0), \Lambda)}{\log n} = \liminf_{n \rightarrow \infty} \frac{-\log d_H(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda)}{\log n} \\ &\geq \liminf_{n \rightarrow \infty} \frac{-\log d_H(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda)}{\log T_{d_H(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda)/2}(\mathbf{i}, x_0)} \geq \liminf_{r \rightarrow 0} \frac{-\log r}{\log T_{r/2}(\mathbf{i}, x_0)} = \frac{1}{\dim_M(\overleftarrow{\nu})}. \end{aligned}$$

The proof of the second part is similar.  $\square$

**Remark 2.3.** The condition of one-sided exponential decay seems necessary. The other condition  $\dim_L \Lambda > 0$  seems purely technical and we believe the result should still hold without it. In particular, we will show in Section 3.2 some examples when this condition can be omitted. It is assumed only to ensure that there are sufficiently many small balls, where  $\dim_M \pi_* \overleftarrow{\mu}$  is attained and so the initial point  $x_0$  cannot cause a strict drop in the approximation by beginning the chaos game at the least accessible part of  $\Lambda$ . Note, however, that  $\dim_L \Lambda > 0$  is not a very restrictive condition. For example, all self-affine sets are uniformly perfect [45] which is equivalent to having positive lower dimension [29, Lemma 2.1]. As an application, we look at Bedford–McMullen carpets, see Section 2.2.

Next we consider the expected value of the cover time  $\mathbb{E}_\mu T_r(x_0)$ , which denotes the expectation of  $T_r(\mathbf{i}, x_0)$  with respect to the measure  $\mu$ . We find that, roughly speaking, this can be bounded in terms of  $(1/r)^\alpha$  (the reciprocal of the measure of the ball of minimum measure at scale  $r$ ), up to some logarithmic correction factors. Define

$$\underline{\alpha}(r) := \max_{x \in \Lambda} \frac{\log \overleftarrow{\nu}(B(x, r))}{\log r} - \underline{\alpha} \quad \text{and} \quad \bar{\alpha}(r) := \max_{x \in \Lambda} \frac{\log \overleftarrow{\nu}(B(x, r))}{\log r} - \bar{\alpha}. \quad (2.3)$$

**Theorem 2.4.** *Let  $\mathcal{F}$  be an arbitrary IFS which satisfies (1.1). Fix a  $\sigma$ -invariant measure  $\mu$  on  $\Sigma$  and let  $\underline{\alpha} = \dim_M(\nu)$  and  $\bar{\alpha} = \overline{\dim}_M(\nu)$ .*

- (a) *If  $\mu$  has one-sided exponential decay, there exists a constant<sup>1</sup>  $C_1$  such that for all  $x_0 \in \Lambda$  and  $r > 0$  such that  $|\bar{\alpha}(r/4)| < \bar{\alpha}/2$  and  $(r/4)^{\bar{\alpha}/2} < 1/2\kappa$ ,*

$$\mathbb{E}_\mu T_r(x_0) \leq C_1 (\log(4/r))^2 (r/4)^{-\bar{\alpha} - \bar{\alpha}(r/4)},$$

*where  $\kappa$  is the constant defined in (1.2).*

- (b) *If  $\dim_L \Lambda > 0$  then for all  $r > 0$  sufficiently small<sup>2</sup> and all  $x_0 \in \Lambda$ ,*

$$\mathbb{E}_\mu T_r(x_0) \geq R_r^{-\alpha + \underline{\alpha}(R_r)}$$

*where  $C_2(r(\log(1/r))^{2/\dim_L \Lambda}) \leq R_r \leq C_3(r(\log(1/r))^{2/\dim_L \Lambda})$  for some uniform constants  $C_2, C_3$ . The dependence of  $R_r$  on  $r$  and other parameters will be made explicit in the proof of Lemma 4.3.*

<sup>1</sup>The constant  $C_1$  will be made explicit in the proof of Lemma 4.2.

<sup>2</sup>The assumptions on the size of  $r$  will be made explicit in Lemma 4.3.

Since  $r^{-\underline{\alpha}+\underline{\varrho}(r)}$  and  $r^{-\overline{\alpha}-\overline{\varrho}(r)}$  are lower and upper bounds respectively on the measure of the ball of minimum measure at scale  $r$ , the main terms in Theorem 2.4(a) and (b) are essentially analogous to the main theorem in [27]. Moreover, Theorem 2.4(a) is essentially analogous to the upper bound on the cover time for the chaos game induced by a Bernoulli measure on self-similar sets in [27] (except we have gained an extra logarithmic factor  $\log(1/r)$  due to the fact that the measure is not necessarily Bernoulli but only has exponential decay). However, the lower bound in Theorem 2.4(b) is essentially just the trivial bound which estimates the expected cover time from below by the expected time to hit the ball of minimum measure. In order to improve this (to be analogous to the lower bound in [27] for instance), one would need some extra information on where the balls of minimum measure were located, in particular to guarantee that they are sufficiently far from each other, in a dynamical sense.

**2.1. Examples with optimal rate.** From an applied point of view, the chaos game provides an efficient algorithm to produce images well-approximating the attractor of an IFS. The choice of the measure  $\mu$  driving the chaos game influences the “quality” of the image we obtain. For any  $r > 0$ , consider a maximal  $r$ -packing of the attractor  $\Lambda$ , i.e. a collection of sets of diameter  $r$  with disjoint interiors that cover  $\Lambda$ . The orbit  $\mathcal{O}_n(\mathbf{i}, x_0)$  becomes  $r$ -dense in  $\Lambda$  once it has visited all elements of the  $r$ -packing. Therefore, the practicality of the algorithm depends on how easy is it to define the  $r$ -packing and to keep track which elements of the  $r$ -packing the orbit has visited.

**2.1.1. Self-similar sets with SSP.** If the IFS satisfies the *strong separation property*, i.e.  $f_i(\Lambda) \cap f_j(\Lambda) = \emptyset$  for every  $i \neq j$ , then the natural projection  $\pi : \Sigma \rightarrow \Lambda$  defined in (1.3) is a bijection between the symbolic space and the attractor. Hence, there is a chance to define the packing in terms of finite length strings. When the  $f_i$  are all similarity mappings with contraction ratios  $0 < \lambda_i < 1$ , then the symbolic  $r$ -packing consists of those cylinder sets  $[i_1, \dots, i_m]$  for which  $\lambda_{i_1} \cdot \dots \cdot \lambda_{i_m} \leq r < \lambda_{i_1} \cdot \dots \cdot \lambda_{i_{m-1}}$ . If  $j$  is the next chosen index, then the chaos game transitions from the current state  $[i_1, \dots, i_m]$  to the unique cylinder of the packing containing the cylinder  $[j, i_1, \dots, i_m]$ . As mentioned before,  $\dim_{\text{M}} \Lambda$  is achieved by a self-similar measure  $\nu_{\mathbf{p}}$  if and only if  $p_i = \lambda_i^s$ , where  $s$ , often called the *similarity dimension*, is the unique solution to the equation  $\sum_i \lambda_i^s = 1$ . By Theorem 2.1 and (1.4), the shortest running time to approximate  $\Lambda$  with resolution  $r$  is to choose  $\mathbf{p}$  this way. Note that this also follows from [27].

**2.1.2. Self-similar sets with overlaps.** The separation condition is important, but can be circumvented in some cases in order to achieve the optimal possible rate. Suppose that the similarity dimension  $s$  of the self-similar IFS  $\{f_i(x) = \lambda_i O_i x + t_i\}_{i=1}^N$  on  $\mathbb{R}^d$  is smaller than the dimension of the state space. Moreover, suppose that  $\dim_{\text{M}} \Lambda = s$ , which happens generically, see [23, 24]. Then for the self-similar measure  $\nu_{\mathbf{p}}$  with  $p_i = \lambda_i^s$  we have

$$\frac{\log \nu_{\mathbf{p}}(B(x, r))}{\log r} \leq \frac{\log \nu_{\mathbf{p}}(\pi([i_1, \dots, i_m]))}{\log r} \leq s + \frac{s \min_i \log \lambda_i}{\log r},$$

and so  $\overline{\dim}_{\text{M}} \nu_{\mathbf{p}} \leq s$ . On the other hand, by our assumption  $\underline{\dim}_{\text{M}} \nu_{\mathbf{p}} \geq \underline{\dim}_{\text{M}} \Lambda = s$ .

**2.1.3. Bernoulli convolution.** We saw that the most efficient convergence rate can be found for a wide class of self-similar systems if the similarity dimension is smaller than the dimension



of the state space. However, this is not the case for Bernoulli convolutions. For  $\lambda \in (1/2, 1)$ , consider the overlapping IFS

$$\mathcal{F} = \{\lambda x - 1, \lambda x + 1\}, \quad \text{with attractor } \Lambda = \left[ \frac{-1}{1-\lambda}, \frac{1}{1-\lambda} \right]. \quad (2.4)$$

The self-similar measure  $\nu_{\mathbf{p}}$  associated to  $\mathcal{F}$  is the well-known biased *Bernoulli convolution*. There is extensive literature on it, especially for the non-biased  $\mathbf{p} = (1/2, 1/2)$  case, see the survey [39] or the recent influential papers [41, 43] and references therein. Clearly,  $\dim_{\text{M}} \Lambda = s = 1$ , however it is easy to see that

$$\dim_{\text{M}} \nu_{\mathbf{p}} \geq \lim_{r \rightarrow 0} \max_{x \in \{-(1-\lambda)^{-1}, (1-\lambda)^{-1}\}} \frac{\log \nu_{\mathbf{p}}(B(x, r))}{\log r} = \frac{\min\{\log p_1, \log p_2\}}{\log \lambda} \geq \frac{\log 2}{\log(1/\lambda)} > 1$$

implying that there is no Bernoulli measure which achieves  $\dim_{\text{M}} \Lambda$ .

There are two further natural classes of measures satisfying one-sided exponential decay (1.2) which motivate the study of the chaos game beyond the i.i.d. and self-similar case.

**2.1.4. Self-conformal sets.** For conformal IFS  $\{f_1, \dots, f_N\}$  on  $\mathbb{R}^d$ , where the contractions are  $C^{1+\epsilon}$ , the natural measure  $\overleftarrow{\nu}$  on the attractor  $\Lambda$  can be expressed as the push-forward of the Gibbs measure  $\overleftarrow{\mu}$  for the Hölder continuous potential  $\mathbf{i} \mapsto s \log |f'_{i_1}(\pi(\sigma \mathbf{i}))|$ , i.e.  $\overleftarrow{\nu} = \pi_* \overleftarrow{\mu}$ , where  $s$  is the conformality dimension, see [13, p. 89]. Such Gibbs measures with Hölder continuous potentials satisfy (1.2) (see [9, p. 15]), thus our results are applicable to the reversed Gibbs measure  $\mu$  (which clearly also satisfies (1.2)). In particular, there exists a constant  $C > 0$  such that for every  $x \in \Lambda$  and every finite word  $\bar{i}$

$$C^{-1} \leq \frac{\overleftarrow{\mu}(\bar{i})}{|f'_{\bar{i}}(x)|^s} \leq C. \quad (2.5)$$

Let us further assume that  $\dim_{\text{M}} \Lambda = s \leq d$ , which holds for instance under the strong separation property (see [13, p. 89]) or for typical systems (see [42, Theorem 6.1]). Similarly to the self-similar case, it is easy to show in this case that  $\dim_{\text{M}} \overleftarrow{\nu} = s$ . Hence, the chaos game driven by  $\mu$  will distribute mass most uniformly over  $\Lambda$ .

**2.1.5. Self-affine sets with small dimension.** Let  $\{S_i(\cdot) = A_i(\cdot) + t_i\}_{i=1}^N$  be a self-affine IFS, where  $A_i \in \mathcal{GL}(d, \mathbb{R})$  and  $t_i \in \mathbb{R}^d$ . The attractor of this IFS is called a *self-affine set*  $E$ . The *affinity dimension*  $s_0$  is the expected Hausdorff and Minkowski dimension of  $E$  [11]. Generically, the measure of maximal Hausdorff dimension on  $E$  is the pushforward  $\overleftarrow{\nu} = \pi_* \overleftarrow{\mu}$  of a measure  $\overleftarrow{\mu}$  on  $\Sigma$  which is the equilibrium state for a particular sub-additive potential (which depends on  $s_0$ ) [28], where by “generically” we mean in the sense that if a set of linear parts  $A_i$  are fixed (where each  $\|A_i\| < 1/2$ ) then the conclusion holds for Lebesgue typical choices of translations  $t_i$ , see [25]. We shall refer to this measure  $\overleftarrow{\mu}$  as a *Käenmäki measure*. Piraino proved that if the linear parts  $\{A_i\}_{i=1}^N$  generate a strongly irreducible semigroup which contains a matrix with a simple leading eigenvalue, then (1.2) holds for the Käenmäki measure  $\overleftarrow{\mu}$ , see [40, Theorem 3.3]. Therefore, our results are applicable to the reversed Käenmäki measure  $\mu$ , which is a natural candidate that we expect to optimise the cover time of the chaos game on  $E$ . If we additionally assume that  $s_0 \leq 1$  then  $\mu$  does indeed optimise the cover time of the chaos game, i.e.  $\dim_{\text{M}} \overleftarrow{\nu} = \dim_{\text{M}} E = s_0$ . To see this, note that under the assumptions of  $\|A_i\| < 1/2$ ,  $\dim_{\text{M}} E = s_0$  for Lebesgue almost every translations, see [12], therefore  $\dim_{\text{M}} \overleftarrow{\nu} \geq s_0$ . On the other hand, the strong irreducibility implies that  $\overleftarrow{\mu}$  satisfies a type of “Gibbs property” for the weighted norm potential [30, Remark 4.2], and using this and



the definition of the Minkowski dimension, it is not difficult to show that  $\dim_{\mathbb{M}} \overleftarrow{\mathcal{V}} \leq s_0$ . As a result,  $\dim_{\mathbb{M}} \overleftarrow{\mathcal{V}} = \dim_{\mathbb{M}} E = s_0$ , i.e.  $\mu$  optimises the cover time. Since  $\dim_{\mathbb{M}} E = s_0$  under the assumptions of strong irreducibility and strong open set condition for planar systems, see [3], we can repeat the argument above for that situation as well.

Leaving the self-similar setting, except for special cases, the construction of the  $r$ -packing and keeping track of the orbit is difficult. One such special setting is the family of planar self-affine carpets. We continue by introducing the class of Bedford–McMullen carpets, give a complete characterisation of the vectors  $\mathbf{p}$  which minimise  $\dim_{\mathbb{M}} \nu_{\mathbf{p}}$  in Theorem 2.5, and later in Section 5.4 present how to keep track of the orbit on the appropriately defined symbolic  $r$ -packing together with some simulations.

**2.2. Application to Bedford–McMullen carpets.** Bedford–McMullen carpets are self-affine sets on the plane introduced independently by Bedford [8] and McMullen [37]. There is an abundant amount of literature on them, we refer to the recent survey [19] and references therein, thanks to the simplicity of their construction and at the same time exhibiting many interesting features. Here we show another such interesting feature regarding the optimisation problem

$$\min_{\mathbf{p}} \dim_{\mathbb{M}} \nu_{\mathbf{p}}. \quad (2.6)$$

As discussed in Remark 2.3, our results are applicable. In fact, an explicit formula is known for the lower dimension of Bedford–McMullen carpets [17]. Thus, any vector  $\mathbf{p}^*$  minimising  $\dim_{\mathbb{M}} \nu_{\mathbf{p}}$  has the interpretation that the chaos game run with  $\mathbf{p}^*$  has the fastest running time among Bernoulli measures to reach a certain resolution.

Split  $R = [0, 1]^2$  into  $m$  columns of equal width and  $n$  rows of equal height for some integers  $n > m \geq 2$  and consider orientation preserving maps on  $R$  of the form

$$f_{(i,j)}(\underline{x}) := \begin{pmatrix} 1/m & 0 \\ 0 & 1/n \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} i/m \\ j/n \end{pmatrix}$$

for the index set  $(i, j) \in \mathcal{A} \subseteq \{0, \dots, m-1\} \times \{0, \dots, n-1\}$ . The attractor  $\Lambda$  of the IFS  $\mathcal{F} = \{f_{(i,j)}\}_{(i,j) \in \mathcal{A}}$  is called a *Bedford–McMullen carpet*, see Figure 1 for three examples. For our purposes it only matters how many maps there are in each column. Therefore, the input parameters of a carpet for us are the following.

Considering all non-empty columns, assume that the number of maps in a column take  $M_0$  different values. In ascending order, let  $N_1 < N_2 < \dots < N_{M_0}$  denote these different values. Moreover, let  $R_i$  denote the number of columns with  $N_i$  number of maps. If  $M_0 = 1$ , then we say that the carpet has *uniform vertical fibres*. The total number of non-empty columns is  $M = R_1 + \dots + R_{M_0} \leq m$  and the total number of maps is  $N = R_1 N_1 + \dots + R_{M_0} N_{M_0} \leq nm$ . For a distinguished index  $1 \leq K \leq M_0$ , let  $|\mathcal{R}_K| := R_1 + \dots + R_K$ , i.e. the number of columns with at most  $N_K$  rectangles, and  $\|\mathcal{R}_K^C\| := R_{K+1} N_{K+1} + \dots + R_{M_0} N_{M_0}$ , i.e. the total number of rectangles in columns with strictly more than  $N_K$  rectangles.

Recall, every non-degenerate probability vector  $\mathbf{p}$  defines the measure  $\nu_{\mathbf{p}}$  on  $\Lambda$  via (2.2). Let  $\mathcal{Q} := \{\mathbf{q} = (q_1, \dots, q_{M_0}) : R_1 q_1 + \dots + R_{M_0} q_{M_0} = 1 \text{ and } q_k > 0 \text{ for all } 1 \leq k \leq M_0\}$  and define the function

$$\alpha(\mathbf{q}) := \max_{1 \leq k, \ell \leq M_0} \left\{ \frac{\log(q_k/N_k)}{-\log n} + \left(1 - \frac{\log m}{\log n}\right) \frac{\log q_\ell}{-\log m} \right\}. \quad (2.7)$$

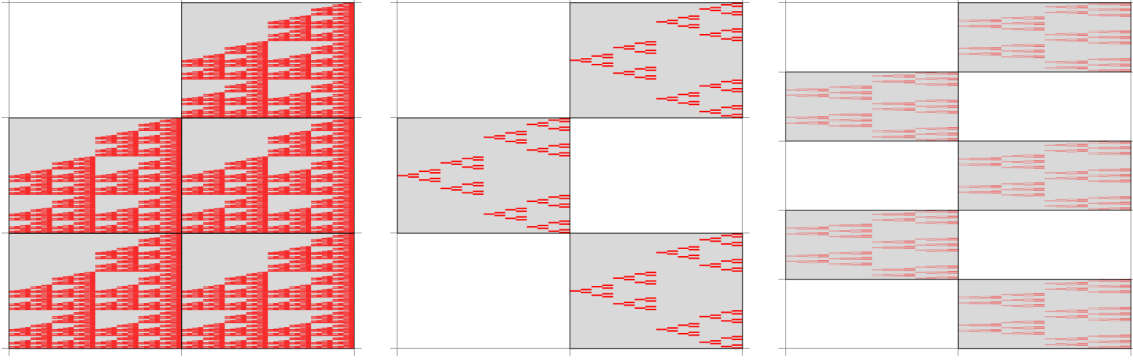


FIGURE 1. Three Bedford–McMullen carpets (in red) and the images of  $[0, 1]^2$  under the maps in each IFS (shaded rectangles). In each case,  $\dim_{\mathbb{M}} \nu_{\mathbf{p}}$  is minimised by a different vector  $\mathbf{p}$ , see Table 1.

**Theorem 2.5.** *Let  $\Lambda$  be a Bedford–McMullen carpet with non-uniform vertical fibres. Then*

$$\min_{\mathbf{p}} \dim_{\mathbb{M}} \nu_{\mathbf{p}} = \min_{\mathbf{q} \in \mathcal{Q}} \alpha(\mathbf{q}).$$

Moreover,  $\min_{\mathbf{q} \in \mathcal{Q}} \alpha(\mathbf{q})$  is attained at a vector either of the form  $\mathbf{q}_K = (q_{K,1}, \dots, q_{K,M_0})$  for a distinguished  $1 \leq K \leq M_0$  defined by

$$q_{K,k} = \begin{cases} N_K \cdot (N_K |\mathcal{R}_K| + \|\mathcal{R}_K^C\|)^{-1}, & \text{for all } k \leq K, \\ N_k \cdot (N_K |\mathcal{R}_K| + \|\mathcal{R}_K^C\|)^{-1}, & \text{for all } k > K; \end{cases} \quad (2.8)$$

or of the form  $\mathbf{Q}_K = (Q_{K,1}, \dots, Q_{K,M_0})$  for a  $1 \leq K \leq M_0 - 1$  defined by

$$Q_{K,k} = \left(1 - \frac{\log m}{\log n}\right) \frac{1}{|\mathcal{R}_K|} \text{ for all } k \leq K \quad \text{and} \quad Q_{K,k} = \frac{\log m}{\log n} \frac{N_k}{\|\mathcal{R}_K^C\|} \text{ for all } k > K. \quad (2.9)$$

Furthermore, if  $\alpha(\mathbf{q}^*) = \min_{\mathbf{q} \in \mathcal{Q}} \alpha(\mathbf{q})$ , then  $\mathbf{q}^*$  defines a solution  $\mathbf{p}^*$  to (2.6) by defining  $p_i^* := q_k^*/N_k$  if  $i$  belongs to a column with  $N_k$  rectangles (i.e. mass in each column is distributed uniformly amongst the rectangles within it).

**Remark 2.6.**

- (1) We give a procedure to determine which vector minimises  $\alpha(\mathbf{q})$  in Proposition 5.4. For all three examples in Figure 1 a different vector is the minimiser, see Table 1.
- (2) A Bedford–McMullen carpet  $\Lambda$  has uniform vertical fibres if and only if its Hausdorff and Minkowski dimensions are equal. It is easy to see that in this case the vector minimising  $\dim_{\mathbb{M}} \nu_{\mathbf{p}}$  is the uniform vector  $\mathbf{p}^* = (1/N, \dots, 1/N)$ . Moreover,  $\dim_{\mathbb{M}} \nu_{\mathbf{p}^*} = \dim_{\mathbb{M}} \Lambda$ . If  $\Lambda$  has non-uniform vertical fibres, then there is a positive gap  $\min_{\mathbf{p}} \dim_{\mathbb{M}} \nu_{\mathbf{p}} > \dim_{\mathbb{M}} \Lambda$ . See Claim 5.1 for details.
- (3) The vector minimizing  $\dim_{\mathbb{M}} \nu_{\mathbf{p}}$  is not necessarily unique, see Section 5.5.

**Structure of paper.** In Section 3 we prove Theorem 2.1, while Section 4 contains the proof of Theorem 2.4. The results for Bedford–McMullen carpets are proved in Section 5. It also contains concrete examples and results of simulations. Section 6 contains a short list of questions that arise naturally from our investigations.

## 3. PROOF OF THEOREM 2.1

We begin by recalling and introducing necessary notation. Elements of the symbolic space  $\Sigma = \{1, \dots, N\}^{\mathbb{N}}$  are denoted  $\mathbf{i}, \mathbf{j}$ . The set of all finite words is  $\Sigma^* = \bigcup_{n=0}^{\infty} \{1, \dots, N\}^n$ , where for  $n=0$  we get the empty word and let  $\bar{\Sigma} := \Sigma \cup \Sigma^*$ . Elements of  $\Sigma^*$  are denoted by  $\bar{i}, \bar{j}$  or as a truncation  $\mathbf{i}|n = i_1 \dots i_n$  of an infinite word. For  $\bar{i} = i_1 \dots i_n$  we let  $|\bar{i}|$  denote the length  $n$  of the word  $\bar{i}$ . The left shift operator on  $\bar{\Sigma}$  is  $\sigma$ , i.e.  $\sigma(i_1 i_2 i_3 \dots) = i_2 i_3 \dots$ . The cylinder set defined by the finite word  $\bar{i}$  of length  $n$  is  $[\bar{i}] = \{\mathbf{i} \in \Sigma : \mathbf{i}|n = \bar{i}\}$ . The natural projection  $\pi : \Sigma \rightarrow \Lambda$  is well defined by the limit

$$\pi(\mathbf{i}) := \lim_{n \rightarrow \infty} f_{\mathbf{i}|n}(x_0),$$

where the limit is independent of the starting point  $x_0$ . The map  $\pi$  is continuous, surjective, but may fail to be injective.

Let  $\bar{i}_-$  denote the finite word which is obtained from  $\bar{i}$  by dropping the last symbol of  $\bar{i}$ . Let

$$\mathcal{P}_r := \{\bar{i} \in \Sigma^* : |f_{\bar{i}}(\Lambda)| \leq r < |f_{\bar{i}_-}(\Lambda)|\}.$$

For much of the time it is enough to work along subsequences of  $\{2^{-n}\}_{n \in \mathbb{N}}$ . Slightly abusing notation, we write  $\mathcal{P}_n := \mathcal{P}_{2^{-n}}$  and  $T_n(\mathbf{i}, x_0) := T_{2^{-n}}(\mathbf{i}, x_0)$ . This should make no confusion, since from the context it should be clear whether the subscript will tend to  $\infty$  or 0. Similarly, let  $\mathcal{Q}_r(A)$  and  $\mathcal{Q}_n(A)$  denote a maximal packing of  $A$  by balls centered in  $A$  with radius  $r$  and  $2^{-n}$  respectively. The definition of the Minkowski dimension implies that there exists  $C_0 > 0$  such that

$$\#\mathcal{Q}_r(\Lambda) \leq C_0 r^{-2D} \quad \text{for every } n, \quad (3.1)$$

where  $D = \overline{\dim}_M \Lambda$ . In particular  $\#\mathcal{Q}_n(\Lambda) \leq C_0 2^{2Dn}$  for every  $n$ .

It readily follows from our assumption (1.1) that for all  $r > 0$ ,

$$L(r) := \frac{\log r}{\log a} - \frac{\log(|\Lambda|/a)}{\log a} \geq \max\{|\bar{i}| : \bar{i} \in \mathcal{P}_r\} \quad (3.2)$$

In particular for every  $n \in \mathbb{N}$

$$L(n) := n \frac{\log 2}{-\log a} - \frac{\log(|\Lambda|/a)}{\log a} \geq \max\{|\bar{i}| : \bar{i} \in \mathcal{P}_n\}. \quad (3.3)$$

For  $\mathbf{i} \in \Sigma$ , let  $\mathcal{P}_r(\mathbf{i})$  be the unique element  $\bar{i} \in \mathcal{P}_r$  such that  $\mathbf{i} \in [\bar{i}]$ . Let us define the symbolic ball as

$$\tilde{B}(x, r) := \{\bar{i} \in \mathcal{P}_r : B(x, r) \cap \pi[\bar{i}] \neq \emptyset\}.$$

Then

$$B(x, r) \cap \Lambda \subseteq \pi \tilde{B}(x, r) \subseteq B(x, 2r) \cap \Lambda.$$

Recall the definition of lower dimension from (2.1). If  $d := \dim_L(\Lambda) > 0$  then there exists a constant  $c_0 > 0$  such that for every  $0 < r < R < |\Lambda|$  and  $x \in \Lambda$

$$N_r(\Lambda \cap B(x, R)) \geq c_0 \left(\frac{R}{r}\right)^{d/2}. \quad (3.4)$$

Theorem 2.1 follows from the following two propositions.

**Proposition 3.1.** *Suppose that  $\dim_{\mathbb{L}} \Lambda > 0$  and let  $\mu$  be an arbitrary  $\sigma$ -invariant measure. Let  $n_k \rightarrow \infty$  be a sequence such that the limit*

$$\lim_{k \rightarrow \infty} \max_{y \in \Lambda} \frac{\log \overleftarrow{\nu} (B(y, ((L(n_k) + 2)/c_0)^{2/d} 2^{-n_k+2}))}{-n_k \log 2}$$

*exists. Let  $\alpha$  denote this limit. Then for  $\mu$ -a.e.  $\mathbf{i}$  and every  $x_0 \in \Lambda$*

$$\liminf_{k \rightarrow \infty} \frac{\log T_{n_k}(\mathbf{i}, x_0)}{n_k \log 2} \geq \alpha.$$

**Proposition 3.2.** *Let  $\mu$  be a  $\sigma$ -invariant measure with exponential decay of correlation. Let  $n_k \rightarrow \infty$  be a sequence such that the limit*

$$\lim_{k \rightarrow \infty} \max_{y \in \Lambda} \frac{\log \overleftarrow{\nu} (B(y, 2^{-n_k}))}{-n_k \log 2}$$

*exists. Let  $\alpha$  denote this limit. Then for  $\mu$ -a.e.  $\mathbf{i}$  and every  $x_0 \in \Lambda$*

$$\limsup_{k \rightarrow \infty} \frac{\log T_{n_k}(\mathbf{i}, x_0)}{n_k \log 2} \leq \alpha.$$

*Proof of Theorem 2.1.* Let  $n_k \rightarrow \infty$  be a sequence for which

$$\underline{\dim}_{\mathbb{M}}(\overleftarrow{\nu}) = \lim_{k \rightarrow \infty} \max_{y \in \Lambda} \frac{\log \overleftarrow{\nu} (B(y, 2^{-n_k}))}{-n_k \log 2}.$$

Then by Proposition 3.2 and Lemma 1.1, for  $\mu$ -almost every  $\mathbf{i}$  and every  $x_0 \in \Lambda$

$$\liminf_{n \rightarrow \infty} \frac{\log T_n(\mathbf{i}, x_0)}{n \log 2} \leq \liminf_{k \rightarrow \infty} \frac{\log T_{n_k}(\mathbf{i}, x_0)}{n_k \log 2} \leq \underline{\dim}_{\mathbb{M}}(\overleftarrow{\nu}).$$

For an arbitrary  $r > 0$  there exists a unique integer  $n$  such that  $2^{-n} \leq r < 2^{-(n-1)}$ . Hence,

$$\frac{n-1}{n} \frac{\log T_{n-1}(\mathbf{i}, x_0)}{(n-1) \log 2} \leq \frac{\log T_r(\mathbf{i}, x_0)}{-\log r}$$

and so  $\liminf_{n \rightarrow \infty} \frac{1}{n \log 2} \log T_n(\mathbf{i}, x_0) = \liminf_{r \rightarrow 0} \frac{1}{-\log r} \log T_r(\mathbf{i}, x_0)$ . This gives one direction. On the other hand

$$\liminf_{r \rightarrow 0} \max_{y \in \Lambda} \frac{\log \overleftarrow{\nu} (B(y, r))}{\log r} = \liminf_{n \rightarrow \infty} \max_{y \in \Lambda} \frac{\log \overleftarrow{\nu} (B(y, ((L(n) + 2)/c_0)^{2/d} 2^{-n+2}))}{-n \log 2}.$$

By Lemma 1.1, this common value equals  $\underline{\dim}_{\mathbb{M}}(\overleftarrow{\nu})$ . Hence, Proposition 3.1 implies that for  $\mu$ -a.e.  $\mathbf{i}$  and every  $x_0 \in \Lambda$

$$\liminf_{n \rightarrow \infty} \frac{\log T_n(\mathbf{i}, x_0)}{n \log 2} \geq \underline{\dim}_{\mathbb{M}}(\overleftarrow{\nu}),$$

proving the first assertion. The proof of the other assertion is analogous.  $\square$

The next two subsections contain the proofs of the two remaining propositions. Both proofs are Borel-Cantelli arguments.

**3.1. Lower bound, proof of Proposition 3.1.** In this proof we will estimate the cover time  $T_n$  from below in terms of the time to hit the “least accessible” part of  $\Lambda$  (i.e. the ball of minimum measure at scale  $2^{-n}$ ). Moreover, we will use the positive lower dimension assumption to deduce that there are sufficiently many small balls of measure comparable to the ball of minimum measure, so that the choice of initial point for the chaos game cannot cause the cover time to drop by beginning in the least accessible part of  $\Lambda$ .

Recall the notations  $L(n)$  from (3.3) and  $c_0$  from (3.4). The first time an orbit hits a ball around a point  $y$  with radius  $r$  is denoted by

$$\mathcal{T}_r(\mathbf{i}, y, x_0) := \min\{n \geq 1 : f_{\mathbf{i}|n}^{\leftarrow}(x_0) \in B(y, r)\}. \quad (3.5)$$

Observe that  $\mathcal{T}_r(\mathbf{i}, y, x_0) \leq T_r(\mathbf{i}, x_0)$  for every  $y \in \Lambda$ . Let  $n_k \rightarrow \infty$  be a sequence such that

$$\alpha := \lim_{k \rightarrow \infty} \max_{y \in \Lambda} \frac{\log \overleftarrow{\mathcal{V}}(B(y, ((L(n_k) + 2)/c_0)^{2/d} 2^{-n_k+2}))}{-n_k \log 2},$$

moreover, let  $y'_k \in \Lambda$  be a point for which the maximum is attained for  $n_k$ . Therefore, for every integer  $K \geq 1$ , we can find  $N(K)$  such that

$$\overleftarrow{\mathcal{V}}(B(y'_k, ((L(n_k) + 2)/c_0)^{2/d} 2^{-n_k+2})) \leq 2^{-n_k(\alpha - \frac{1}{2K})} \quad \text{for every } k \geq N(K). \quad (3.6)$$

Choosing  $r = 2^{-n_k+2}$  and  $R = ((L(n_k) + 2)/c_0)^{2/d} \cdot r$  and letting  $D$  denote the dimension of the state space  $\mathbb{R}^D$  containing  $\Lambda$ , we see that

$$L(n_k) + 2 \leq N_r(\Lambda \cap B(y'_k, R)) \leq \#\mathcal{Q}_{n_k-1}(\Lambda \cap B(y'_k, R)) \leq \left(\frac{2R}{r}\right)^D = 2^D ((L(n_k) + 2)/c_0)^{2D/d}, \quad (3.7)$$

where the upper bound is trivial and the lower bound follows from (3.4). Hence, for each  $\mathbf{i} \in \Sigma$  there exists  $y_k(\mathbf{i}) \in \Lambda$  such that  $B(y_k(\mathbf{i}), 2^{-n_k+1}) \cap \mathcal{O}_{\lceil L(n_k) \rceil}(\mathbf{i}, x_0) = \emptyset$  and

$$B(y_k(\mathbf{i}), 2^{-n_k+1}) \subset B(y'_k, ((L(n_k) + 2)/c_0)^{2/d} 2^{-n_k+2}).$$

Let  $\mathbf{j}$  be an arbitrary coding of  $x_0$ . Then,

$$\begin{aligned} & \mu \left( \left\{ \mathbf{i} : T_{2^{-n_k}}(\mathbf{i}, x_0) < 2^{n_k(\alpha - 1/K)} \right\} \right) \\ & \leq \mu \left( \left\{ \mathbf{i} : \mathcal{T}_{2^{-n_k}}(\mathbf{i}, y_k(\mathbf{i}), x_0) < 2^{n_k(\alpha - 1/K)} \right\} \right) \\ & \leq \mu \left( \left\{ \mathbf{i} : \text{there exists } \lceil L(n_k) \rceil + 1 \leq \ell < 2^{n_k(\alpha - 1/K)} \text{ s.t. } f_{\mathbf{i}|\ell}^{\leftarrow}(x_0) \in B(y_k(\mathbf{i}), 2^{-n_k}) \right\} \right) \\ & \leq \mu \left( \left\{ \mathbf{i} : \text{there exists } \lceil L(n_k) \rceil + 1 \leq \ell < 2^{n_k(\alpha - 1/K)} \text{ s.t. } \mathcal{P}_{n_k}(\mathbf{i}|\ell; \mathbf{j}) \in \tilde{B}(y_k(\mathbf{i}), 2^{-n_k}) \right\} \right) \end{aligned} \quad (3.8)$$

where the first inequality follows because for all  $\mathbf{i} \in \Sigma$ ,  $T_{2^{-n_k}}(\mathbf{i}, x_0) \geq \mathcal{T}_{2^{-n_k}}(\mathbf{i}, y_k(\mathbf{i}), x_0)$ , the second inequality is because for all  $\mathbf{i} \in \Sigma$ ,  $y_k(\mathbf{i})$  is chosen such that  $B(y_k(\mathbf{i}), 2^{-n_k+1}) \cap \mathcal{O}_{\lceil L(n_k) \rceil}(\mathbf{i}, x_0) = \emptyset$  and the final inequality is because  $B(y_k(\mathbf{i}), 2^{-n_k}) \subseteq \tilde{B}(y_k(\mathbf{i}), 2^{-n_k})$ .

Next, let  $\mathcal{Y}_k$  denote the set of centres of the balls in  $\mathcal{Q}_{n_k-1}(\Lambda \cap B(y'_k, R))$ . The subset of  $\Sigma$  that appears in (3.8) is easily seen to be contained inside

$$\bigcup_{y \in \mathcal{Y}_k} \bigcup_{\ell = \lceil L(n_k) \rceil + 1}^{\lfloor 2^{n_k(\alpha-1/K)} \rfloor} \bigcup_{\bar{i} \in \tilde{B}(y, 2^{-n_k})} \left\{ \mathbf{i} : \mathcal{P}_{n_k}(\mathbf{i} \overleftarrow{[\ell] \mathbf{j}}) = \bar{i} \right\}.$$

Thus we can bound (3.8) by sum of measures of sets appearing in the above union to obtain that for  $k \geq N(K)$

$$\begin{aligned} \mu \left( \left\{ \mathbf{i} : T_{2^{-n_k}}(\mathbf{i}, x_0) < 2^{n_k(\alpha-1/K)} \right\} \right) &\leq \sum_{y \in \mathcal{Y}_k} \sum_{\ell = \lceil L(n_k) \rceil + 1}^{\lfloor 2^{n_k(\alpha-1/K)} \rfloor} \sum_{\bar{i} \in \tilde{B}(y, 2^{-n_k})} \mu \left( \left\{ \mathbf{i} : \mathcal{P}_{n_k}(\mathbf{i} \overleftarrow{[\ell] \mathbf{j}}) = \bar{i} \right\} \right) \\ &\leq 2^{n_k(\alpha-1/K)} \sum_{y \in \mathcal{Y}_k} \sum_{\bar{i} \in \tilde{B}(y, 2^{-n_k})} \mu(\overleftarrow{[\bar{i}]}) \\ &\leq 2^{n_k(\alpha-1/K)} \sum_{y \in \mathcal{Y}_k} \overleftarrow{\nu}(B(y, 2^{-n_k+1})) \\ &\leq 2^D ((L(n_k) + 2)/c_0)^{2D/d} 2^{n_k(\alpha-1/K)} 2^{-n_k(\alpha-1/(2K))} \\ &= 2^D ((L(n_k) + 2)/c_0)^{2D/d} 2^{-n_k/(2K)}. \end{aligned}$$

The second inequality follows because  $\mathcal{P}_{n_k}(\mathbf{i} \overleftarrow{[\ell] \mathbf{j}})$  depends only on  $\mathbf{i}$  for  $\ell \geq \lceil L(n_k) \rceil$  by (3.3). The third inequality follows because  $\mu(\overleftarrow{[\bar{i}]}) = \overleftarrow{\mu}(\bar{i})$  and  $\pi \tilde{B}(y, 2^{-n_k}) \subseteq B(y, 2^{-n_k+1})$ . The fourth inequality follows by (3.6), (3.7), and the fact that for any  $y \in \mathcal{Y}_k$ ,  $B(y, 2^{-n_k+1}) \subset B(y'_k, ((L(n_k) + 2)/c_0)^{2/d} 2^{-n_k+2})$ . Thus, the Borel-Cantelli lemma implies that

$$\mu \left( \left\{ \mathbf{i} : T_{2^{-n_k}}(\mathbf{i}, x_0) < 2^{n_k(\alpha-1/K)} \text{ for infinitely many } k \text{'s} \right\} \right) = 0.$$

Since  $K \geq 1$  was arbitrary, we get

$$\mu \left( \left\{ \mathbf{i} : \liminf_{k \rightarrow \infty} \frac{\log T_{2^{-n_k}}(\mathbf{i}, x_0)}{n_k \log 2} \geq \alpha \right\} \right) = 1.$$

**3.2. Notes on positive lower dimension.** Let us observe that the condition  $\dim_L \Lambda > 0$  is purely technical and used only to show the independence of the initial point  $x_0 \in \Lambda$  by providing sufficiently large collection of balls with approximately the smallest possible measure. However, this condition can be circumvented by some other conditions. For example, the lower dimension of the attractor of the system  $\{f_1(x) = x^{1/x}/3, f_2(x) = (x+2)/3\}$  is 0, see [18, Section 6.3] but the claim of Proposition 3.1, and in particular Theorem 2.2, holds for this system as well with measures with one-sided exponential decay.

**Proposition 3.3.** *Let  $\mathcal{F}$  be an IFS satisfying (1.1) and the strong separation property. Moreover, let us assume that there exist  $\bar{i} \neq \bar{j} \in \Sigma^*$  and a constant  $0 < b < 1$  such that*

$$\|f_{\bar{i}}(x) - f_{\bar{i}}(y)\|, \|f_{\bar{j}}(x) - f_{\bar{j}}(y)\| > b\|x - y\| \text{ for all } x, y \in \Lambda.$$

*Let  $\mu$  be a  $\sigma$ -invariant measure such that there exists a constant  $C > 0$  such that for every  $\bar{i}, \bar{j} \in \Sigma^*$*

$$\mu(\overleftarrow{[\bar{i} \bar{j}]} \Sigma^*) \leq C \mu(\overleftarrow{[\bar{i}]}) \mu(\overleftarrow{[\bar{j}]} \Sigma^*). \quad (3.9)$$

Let  $n_k \rightarrow \infty$  be a sequence such that the limit

$$\lim_{k \rightarrow \infty} \max_{y \in \Lambda} \frac{\log \overleftarrow{\nu} \left( B(y, b^{-1} L(n_k)^{-\log b 2^{-n_k}}) \right)}{-n_k \log 2}$$

exists. Let  $\alpha$  denote this limit. Then for  $\mu$ -a.e.  $\mathbf{i}$  and every  $x_0 \in \Lambda$

$$\liminf_{k \rightarrow \infty} \frac{\log T_{n_k}(\mathbf{i}, x_0)}{n_k \log 2} \geq \alpha.$$

*Proof.* Let  $y'_k \in \Lambda$  be where  $\max_{y \in \Lambda} \frac{\log \overleftarrow{\nu} \left( B(y, L(n_k)^{-\log b 2^{-n_k}}) \right)}{-n_k \log 2}$  is attained. By the strong separation, there exists  $\delta > 0$  such that  $d(f_{\bar{i}}(\Lambda), f_{\bar{j}}(\Lambda)) > \delta$ . Let  $\bar{h}_1, \dots, \bar{h}_{2^{m_k}} \in \{\bar{i}, \bar{j}\}^{m_k}$  be words such that  $\log L(n_k) \leq m_k < \log L(n_k) + 1$ . By the strong separation

$$B(f_{\bar{h}_i}(y'_k), 2^{-n_k}) \cap B(f_{\bar{h}_j}(y'_k), 2^{-n_k}) = \emptyset$$

for every  $i \neq j$  and  $k$  sufficiently large. Indeed, if the intersection above was non-empty then

$$2^{-n_k+1} \geq \|f_{\bar{h}_i}(y'_k) - f_{\bar{h}_j}(y'_k)\| \geq b^{m_k} \delta \geq b L(n_k)^{-\log b} \delta,$$

which is absurd as  $k \rightarrow \infty$ . Hence, by the definition of  $m_k$  for every  $\mathbf{i} \in \Sigma$  there exists  $y_k(\mathbf{i}) \in \{f_{\bar{h}_i}(y'_k)\}_{i=1}^{2^{m_k}}$  such that  $B(y_k(\mathbf{i}), 2^{-n_k}) \cap \mathcal{O}_{\lceil L(n_k) \rceil}(\mathbf{i}, x_0) = \emptyset$ .

It is easy to see that  $\overleftarrow{\mu}$  satisfies (3.9) too, and by using the strong separation, we get that for every  $\overleftarrow{\nu}$ -measurable set  $E$  and finite word  $\bar{h} \in \Sigma_*$

$$\overleftarrow{\nu}(f_{\bar{h}}(E)) = \overleftarrow{\mu}(\sigma^{-|\bar{h}|} \pi^{-1}(E) \cap [\bar{h}]) \leq C \overleftarrow{\mu}(\sigma^{-|\bar{h}|} \pi^{-1}(E)) \overleftarrow{\mu}([\bar{h}]) \leq C \overleftarrow{\mu}(\pi^{-1}(E)) = C \overleftarrow{\nu}(E). \quad (3.10)$$

Moreover, by our assumptions on  $f_{\bar{i}}, f_{\bar{j}}$

$$B(f_{\bar{h}_i}(y'_k), 2^{-n_k}) \subseteq f_{\bar{h}_i} \left( B(y'_k, b^{-m} 2^{-n_k}) \right) \subseteq f_{\bar{h}_i} \left( B(y'_k, b^{-1} L(n_k)^{-\log b 2^{-n_k}}) \right),$$

for all  $i = 1, \dots, 2^{m_k}$ . Hence, by (3.10)

$$\overleftarrow{\nu} \left( f_{\bar{h}_i} \left( B(y'_k, b^{-1} L(n_k)^{-\log b 2^{-n_k}}) \right) \right) \leq C \overleftarrow{\nu} \left( B(y'_k, b^{-1} L(n_k)^{-\log b 2^{-n_k}}) \right)$$

and so, for any  $K \geq 1$  and every sufficiently large  $k$

$$\frac{\log \overleftarrow{\nu} \left( B(f_{\bar{h}_i}(y'_k), 2^{-n_k}) \right)}{-n_k \log 2} \geq \frac{\log \overleftarrow{\nu} \left( B(y'_k, b^{-1} L(n_k)^{-\log b 2^{-n_k}}) \right)}{-n_k \log 2} + \frac{\log C}{-n_k \log 2} \geq \alpha - \frac{1}{2K}.$$

Now, the proof now can be finished as in Proposition 3.1 and left for the reader.  $\square$

Observe that the strong separation played a crucial role in the Proof of Proposition 3.3. Since in the overlapping case it might happen that (3.10) fails and the  $\overleftarrow{\nu}$  measure of  $f_{\bar{h}}(B(y, r))$  is much larger than the measure of the set  $B(y, r)$ .

**3.3. Upper bound, proof of Proposition 3.2.** The upper bound is also a Borel-Cantelli argument, and for this we will need to estimate the probability that the cover time is asymptotically larger than what is claimed. In order to estimate this probability, we will use the one-sided exponential decay of correlations assumption to allow us to consider subsequent visits independently of one another.



Let  $n_k \rightarrow \infty$  be a sequence such that

$$\alpha := \lim_{k \rightarrow \infty} \max_{y \in \Lambda} \frac{\log \overleftarrow{\mathcal{V}}(B(y, 2^{-n_k}))}{-n_k \log 2}.$$

Since

$$\lim_{k \rightarrow \infty} \max_{y \in \Lambda} \frac{\log \overleftarrow{\mathcal{V}}(B(y, 2^{-(n_k+3)}))}{-n_k \log 2} = \alpha$$

we can, for every  $K \geq 1$ , define  $N(K)$  such that

$$\min_{y \in \Lambda} \overleftarrow{\mathcal{V}}(B(y, 2^{-(n_k+3)})) \geq 2^{-n_k(\alpha+1/(2K))} \quad (3.11)$$

for every  $k \geq N(K)$ .

Let us consider the packing  $\mathcal{Q}_{n_k+2}(\Lambda)$  of  $\Lambda$  with balls of radius  $2^{-n_k-2}$ . Let  $\mathcal{Y}_k$  denote the centres of balls in the packing  $\mathcal{Q}_{n_k+2}$ . Define

$$t_{n_k}(\mathbf{i}) := \min \{m \geq 1 : (\forall B \in \mathcal{Q}_{n_k+2}) (\exists \ell \leq m) \text{ such that } f_{\mathbf{i}|\ell}^{\leftarrow}(x_0) \in B\}.$$

First notice that for any  $\mathbf{i} \in \Sigma$ ,  $T_{2^{-n_k}}(\mathbf{i}, x_0) \leq t_{n_k}(\mathbf{i})$ . Indeed, since  $\{2B\}_{B \in \mathcal{Q}_{n_k+2}}$  is a cover of  $\Lambda$ , for every  $x \in \Lambda$  there exists  $B \in \mathcal{Q}_{n_k+2}$  such that  $x \in 2B$  and, by definition of  $t_{n_k}(\mathbf{i})$ , there exists  $1 \leq \ell \leq t_{n_k}(\mathbf{i})$  such that  $f_{\mathbf{i}|\ell}^{\leftarrow}(x_0) \in B$ . In particular  $\|x - f_{\mathbf{i}|\ell}^{\leftarrow}(x_0)\| \leq 2^{-n_k-1} + 2^{-n_k-2} < 2^{-n_k}$ .

For short, let

$$m_k := \lceil L(n_k + 3) + \frac{(\alpha + 1)}{\varepsilon} n_k \rceil = \lceil (n_k + 3) \frac{\log 2}{-\log a} - \frac{\log(|\Lambda|/a)}{\log a} + \frac{(\alpha + 1)}{\varepsilon} n_k \rceil,$$

where  $\varepsilon > 0$  is defined in (1.2). Let  $\mathbf{j}$  be an arbitrary coding of  $x_0$ .

Then

$$\begin{aligned} & \mu \left( \left\{ \mathbf{i} : T_{2^{-n_k}}(\mathbf{i}, x_0) \geq \lceil 2^{(\alpha+1/K)n_k} \rceil m_k \right\} \right) \\ & \leq \mu \left( \left\{ \mathbf{i} : t_{n_k}(\mathbf{i}, x_0) \geq \lceil 2^{(\alpha+1/K)n_k} \rceil m_k \right\} \right) \\ & \leq \mu \left( \left\{ \mathbf{i} : \exists y \in \mathcal{Y}_k \text{ s.t. } f_{\mathbf{i}|\ell m_k}^{\leftarrow}(x_0) \notin B(y, 2^{-n_k-2}) \text{ for every } 1 \leq \ell \leq \lceil 2^{(\alpha+1/K)n_k} \rceil \right\} \right) \\ & \leq \mu \left( \left\{ \mathbf{i} : \exists y \in \mathcal{Y}_k \text{ s.t. } \mathcal{P}_{2^{-n_k-3}}(\overleftarrow{\mathbf{i}|\ell m_k \mathbf{j}}) \notin \tilde{B}(y, 2^{-n_k-3}) \text{ for every } 1 \leq \ell \leq \lceil 2^{(\alpha+1/K)n_k} \rceil \right\} \right), \end{aligned} \quad (3.12)$$

where the first inequality is because  $T_{2^{-n}}(\mathbf{i}, x_0) \leq t_n(\mathbf{i})$  and the final inequality follows because  $\pi \tilde{B}(y, 2^{-n_k-3}) \subseteq B(y, 2^{-n_k-2})$ . Note that since  $m_k \geq L(n_k + 3)$ , it follows that for any  $\mathbf{i} \in \Sigma$ ,  $\mathcal{P}_{2^{-n_k-3}}(\overleftarrow{\mathbf{i}|\ell m_k \mathbf{j}})$  depends only on  $\mathbf{i}$ .

Next, observe that by (1.2), (3.3) and the fact that  $m_k - L(n_k + 3) \geq \frac{\alpha+1}{\varepsilon} n_k$  then for any  $M \in \mathbb{N}$  and  $\bar{i}_1, \dots, \bar{i}_M \in \mathcal{P}_{n_k+3}$ ,

$$\begin{aligned} & \mu \left( \sigma^{-(m_k - |\bar{i}_1|)}(\overleftarrow{[\bar{i}_1]}) \cap \sigma^{-(2m_k - |\bar{i}_2|)}(\overleftarrow{[\bar{i}_2]}) \cap \dots \cap \sigma^{-(Mm_k - |\bar{i}_M|)}(\overleftarrow{[\bar{i}_M]}) \right) \\ & \leq (1 + 2^{-(\alpha+1)n_k} \kappa)^M \mu(\overleftarrow{[\bar{i}_1]}) \cdots \mu(\overleftarrow{[\bar{i}_M]}). \end{aligned}$$

Applying this to (3.12) we obtain that

$$\begin{aligned}
& \mu \left( \left\{ \mathbf{i} : T_{2^{-n_k}}(\mathbf{i}, x_0) \geq \lceil 2^{(\alpha+1/K)n_k} \rceil m_k \right\} \right) \\
& \leq (1 + 2^{-(\alpha+1)n_k/K})^{\lceil 2^{(\alpha+1/K)n_k} \rceil} \sum_{y \in \mathcal{Y}_k} \left( 1 - \sum_{\bar{z} \in \tilde{B}(y, 2^{-n_k-3})} \mu(\lceil \bar{z} \rceil) \right)^{\lceil 2^{(\alpha+1/K)n_k} \rceil} \\
& \leq (1 + 2^{-(\alpha+1)n_k/K})^{\lceil 2^{(\alpha+1/K)n_k} \rceil} \sum_{y \in \mathcal{Y}_k} (1 - \lceil \mathcal{V}(B(y, 2^{-n_k-3})) \rceil)^{\lceil 2^{(\alpha+1/K)n_k} \rceil}, \tag{3.13}
\end{aligned}$$

where the final inequality follows because  $B(y, 2^{-n_k-3}) \subseteq \pi \tilde{B}(y, 2^{-n_k-3})$ .

Next, applying (3.11) to (3.13) we obtain that for  $k \geq N(K)$ ,

$$\begin{aligned}
& \mu \left( \left\{ \mathbf{i} : T_{2^{-n_k}}(\mathbf{i}, x_0) \geq \lceil 2^{(\alpha+1/K)n_k} \rceil m_k \right\} \right) \\
& \leq \# \mathcal{Q}_{n_k+2} (1 + 2^{-(\alpha+1)n_k/K})^{\lceil 2^{(\alpha+1/K)n_k} \rceil} (1 - 2^{-n_k(\alpha+1/(2K))})^{\lceil 2^{(\alpha+1/K)n_k} \rceil} \\
& \leq C_0 2^{2D(n_k+2)} \exp \left( (2^{-(\alpha+1)n_k/K} - 2^{-n_k(\alpha+1/(2K))}) \lceil 2^{(\alpha+1/K)n_k} \rceil \right) \\
& \leq C_0 2^{2D(n_k+2)} \exp \left( -2^{n_k/(2K)} + 2^{-(1-1/K)n_k/K} + 2^{-(\alpha+1)n_k/K} - 2^{-n_k(\alpha+1/(2K))} \right),
\end{aligned}$$

where in the second inequality we use (3.1) and the fact that for all  $x \in \mathbb{R}$ ,  $1 + x \leq \exp(x)$ .

By the Borel-Cantelli lemma

$$\mu \left( \left\{ \mathbf{i} : T_{2^{-n_k}}(\mathbf{i}, x_0) \geq \lceil 2^{(\alpha+1/K)n_k} \rceil m_k \text{ for inf. many } k \text{'s} \right\} \right) = 0.$$

Since  $K$  was arbitrary and  $\lim_{k \rightarrow \infty} \frac{\log \lceil 2^{(\alpha+1/K)n_k} \rceil m_k}{n_k \log 2} = \alpha + 1/K$ , we get

$$\mu \left( \left\{ \mathbf{i} : \limsup_{k \rightarrow \infty} \frac{\log T_{2^{-n_k}}(\mathbf{i}, x_0)}{n_k \log 2} \leq \alpha \right\} \right) = 1.$$

#### 4. EXPECTED COVER TIME, PROOF OF THEOREM 2.4

Fix  $x_0 \in \Lambda$  and let  $\mathcal{I}$  be a finite set of pairwise disjoint Euclidean balls in  $\mathbb{R}^d$ . For  $\mathbf{i} \in \Sigma$  we define  $\mathcal{T}_{\mathcal{I}}(\mathbf{i}, x_0)$  to be the first time that the orbit  $\mathcal{O}_n(\mathbf{i}, x_0)$  has visited each of the balls in  $\mathcal{I}$ :

$$\mathcal{T}_{\mathcal{I}}(\mathbf{i}, x_0) := \inf \{ n \geq 0 : \forall I \in \mathcal{I}, \exists y \in \mathcal{O}_n(\mathbf{i}, x_0) \text{ such that } y \in I \}.$$

Similarly for  $\mathbf{i} \in \Sigma$  and  $I \in \mathcal{I}$  we define  $\mathcal{T}_I(\mathbf{i}, x_0)$  to be the first time that the orbit  $\mathcal{O}_n(\mathbf{i}, x_0)$  visits the ball  $I$ :

$$\mathcal{T}_I(\mathbf{i}, x_0) := \inf \{ n \geq 0 : \exists y \in \mathcal{O}_n(\mathbf{i}, x_0) \text{ such that } y \in I \}.$$

The expected values of  $\mathcal{T}_{\mathcal{I}}(\mathbf{i}, x_0)$  and  $\mathcal{T}_I(\mathbf{i}, x_0)$  with respect to a measure  $\mu$  on  $\Sigma$  are then denoted by  $\mathbb{E}_{\mu} \mathcal{T}_{\mathcal{I}}(x_0)$  and  $\mathbb{E}_{\mu} \mathcal{T}_I(x_0)$  respectively.

In the following proposition, we show that  $\mathbb{E}_{\mu} \mathcal{T}_{\mathcal{I}}(x_0)$  can be bounded above and below in terms of uniform upper and lower bounds on  $\mathbb{E}_{\mu} \mathcal{T}_{\mathcal{I}}(x)$  over  $x \in \Lambda$ . This is an adaptation of a method of Matthews [36] for bounding the expected cover time of a Markov chain, meaning the expected time for the Markov chain to visit all of its states, in terms of the expected hitting times of individual states, see also [34].

**Proposition 4.1.** *Fix  $x_0 \in \Lambda$  and a finite set  $\mathcal{I}$  of pairwise disjoint Euclidean balls in  $\mathbb{R}^d$ .*

- (a) *Suppose there exists  $C > 1$  such that for all  $\bar{i}, \bar{j} \in \Sigma^*$ ,  $\mu([\bar{i}\bar{j}]) \leq C\mu([\bar{i}])\mu([\bar{j}])$ .<sup>3</sup> Additionally assume that  $\sup_{I \in \mathcal{I}} \sup_{x \in \Lambda} \mathbb{E}_\mu \mathcal{T}_I(x) \leq T$ . Then*

$$\mathbb{E}_\mu \mathcal{T}_{\mathcal{I}}(x_0) \leq CT \left( 1 + \frac{1}{2} + \cdots + \frac{1}{\#\mathcal{I}} \right).$$

- (b) *Suppose there exists a constant  $c > 0$  such that  $\mu([\bar{i}\bar{j}]) \geq c\mu([\bar{i}])\mu([\bar{j}])$  for all  $\bar{i}, \bar{j} \in \Sigma^*$ . Additionally assume that  $\inf_{I \in \mathcal{I}} \inf_{x \in \Lambda \setminus I} \mathbb{E}_\mu \mathcal{T}_I(x) \geq t$ . Then*

$$\mathbb{E}_\mu \mathcal{T}_{\mathcal{I}}(x_0) \geq ct \left( 1 + \frac{1}{2} + \cdots + \frac{1}{\#\mathcal{I}} \right).$$

*Proof.* Write  $N = \#\mathcal{I}$  and write  $\mathcal{I} = \{I_1, \dots, I_N\}$ . Let  $S_N$  denote the set of permutations  $\sigma$  of  $\{1, \dots, N\}$ . Let  $m$  be the uniform measure on  $S_N$ , so that for all  $\sigma \in S_N$ ,  $m(\sigma) = \frac{1}{N!}$ .

Given  $\sigma \in S_N$ , for  $\mathbf{i} \in \Sigma$  we let  $\mathcal{T}_{\sigma(1)}^{\sigma(k)}(\mathbf{i}, x_0)$  denote the first time that the orbit  $\mathcal{O}_n(\mathbf{i}, x_0)$  has visited each of the balls  $I_{\sigma(1)}, \dots, I_{\sigma(k)}$ :

$$\mathcal{T}_{\sigma(1)}^{\sigma(k)}(\mathbf{i}, x_0) := \min\{n \geq 0 : \forall 1 \leq j \leq k, \exists y \in \mathcal{O}_n(\mathbf{i}, x_0) \text{ such that } y \in I_{\sigma(j)}\}.$$

The expected value of  $\mathcal{T}_{\sigma(1)}^{\sigma(k)}(\mathbf{i}, x_0)$  with respect to  $\mu$  is denoted by  $\mathbb{E}_\mu \mathcal{T}_{\sigma(1)}^{\sigma(k)}(x_0)$ . For brevity, throughout this proof we will use the notation  $\mathcal{T}_{\sigma(k)}$  to mean  $\mathcal{T}_{I_{\sigma(k)}}$ .

Given  $\sigma \in S_N$  and  $2 \leq k \leq N$  we define

$$A_{\sigma,k} := \{\mathbf{i} \in \Sigma : \mathcal{T}_{\sigma(1)}^{\sigma(k-1)}(\mathbf{i}, x_0) < \mathcal{T}_{\sigma(k)}(\mathbf{i}, x_0)\}.$$

We also consider the following decomposition of  $A_{\sigma,k}$  into cylinder sets. In particular there exists a set of words  $C_{\sigma,k} \subset \Sigma^*$  such that  $A_{\sigma,k} = \bigcup_{\bar{i} \in C_{\sigma,k}} [\bar{i}]$  and for each  $\bar{i} \in C_{\sigma,k}$  and each  $\mathbf{i} \in [\bar{i}]$ ,  $\mathcal{T}_{\sigma(1)}^{\sigma(k-1)}(\mathbf{i}, x_0) = |\bar{i}|$ .

$\mathbb{E}_\mu \mathcal{T}_{\mathcal{I}}(x_0) = \mathbb{E}_\mu \mathcal{T}_{\sigma(1)}^{\sigma(N)}(x_0)$  for any  $\sigma \in S_N$ . Therefore,

$$\begin{aligned} \mathbb{E}_\mu \mathcal{T}_{\mathcal{I}}(x_0) &= \int \mathbb{E}_\mu \mathcal{T}_{\sigma(1)}^{\sigma(N)}(x_0) dm \\ &= \int \mathbb{E}_\mu \mathcal{T}_{\sigma(1)}(x_0) dm + \sum_{k=2}^N \int \mathbb{E}_\mu (\mathcal{T}_{\sigma(1)}^{\sigma(k)}(x_0) - \mathcal{T}_{\sigma(1)}^{\sigma(k-1)}(x_0)) dm \\ &= \int \mathbb{E}_\mu \mathcal{T}_{\sigma(1)}(x_0) dm + \sum_{k=2}^N \int \mathbb{E}_\mu (\mathcal{T}_{\sigma(1)}^{\sigma(k)}(\mathbf{i}, x_0) - \mathcal{T}_{\sigma(1)}^{\sigma(k-1)}(\mathbf{i}, x_0) : \mathbf{i} \in A_{\sigma,k}) dm \\ &= \int \mathbb{E}_\mu \mathcal{T}_{\sigma(1)}(x_0) dm + \sum_{k=2}^N \int \sum_{\bar{i} \in C_{\sigma,k}} \mathbb{E}_\mu (\mathcal{T}_{\sigma(1)}^{\sigma(k)}(\mathbf{i}, x_0) - \mathcal{T}_{\sigma(1)}^{\sigma(k-1)}(\mathbf{i}, x_0) : \mathbf{i} \in [\bar{i}]) dm \end{aligned}$$

where the penultimate equality holds because if  $\mathbf{i} \notin A_{\sigma,k}$  then  $\mathcal{T}_{\sigma(1)}^{\sigma(k)}(\mathbf{i}, x_0) = \mathcal{T}_{\sigma(1)}^{\sigma(k-1)}(\mathbf{i}, x_0)$ .

<sup>3</sup>Note that this is clearly satisfied by any invariant measure satisfying the one-sided exponential decay property (1.2).

For each  $\bar{i} \in C_{\sigma,k}$  there exists  $D_{\bar{i}} \subset \Sigma^*$  such that  $[\bar{i}] = \bigcup_{\bar{j} \in D_{\bar{i}}} [\bar{i}\bar{j}]$  and for all  $\mathbf{i} \in [\bar{i}]$ ,  $\mathcal{T}_{\sigma(k)}(\mathbf{i}, x_0) = |\bar{i}| + |\bar{j}|$ . In particular,  $\mathbb{E}_\mu \mathcal{T}_{\sigma(k)}(f_{\bar{i}}^{\leftarrow}(x_0)) = \sum_{\bar{j} \in D_{\bar{i}}} \mu([\bar{i}\bar{j}])|\bar{j}|$ .

To prove (a), notice that since  $\mu([\bar{i}\bar{j}]) \leq C\mu([\bar{i}])\mu([\bar{j}])$  for all  $\bar{i}, \bar{j} \in \Sigma^*$ ,

$$\begin{aligned} \sum_{\bar{i} \in C_{\sigma,k}} \mathbb{E}_\mu(\mathcal{T}_{\sigma(1)}^{\sigma(k)}(\mathbf{i}, x_0) - \mathcal{T}_{\sigma(1)}^{\sigma(k-1)}(\mathbf{i}, x_0) : \mathbf{i} \in [\bar{i}]) &= \sum_{\bar{i} \in C_{\sigma,k}} \sum_{\bar{j} \in D_{\bar{i}}} \mu([\bar{i}\bar{j}])|\bar{j}| \\ &\leq C \sum_{\bar{i} \in C_{\sigma,k}} \sum_{\bar{j} \in D_{\bar{i}}} \mu([\bar{i}])\mu([\bar{j}])|\bar{j}| = C \sum_{\bar{i} \in C_{\sigma,k}} \mu([\bar{i}])\mathbb{E}_\mu \mathcal{T}_{\sigma(k)}(f_{\bar{i}}^{\leftarrow}(x_0)) \leq CT\mu(A_{\sigma,k}), \end{aligned}$$

where in the final inequality we used that  $\sup_{I \in \mathcal{I}} \sup_{x \in \Lambda} \mathbb{E}_\mu \mathcal{T}_I(x) \leq T$ . Therefore

$$\mathbb{E}_\mu \mathcal{T}_{\mathcal{I}}(x_0) \leq CT \left( 1 + \sum_{k=2}^N \int \mu(A_{\sigma,k}) dm \right).$$

To prove (b) notice that since  $\mu([\bar{i}\bar{j}]) \geq c\mu([\bar{i}])\mu([\bar{j}])$  for all  $\bar{i}, \bar{j} \in \Sigma^*$ , we similarly obtain

$$\sum_{\bar{i} \in C_{\sigma,k}} \mathbb{E}_\mu(\mathcal{T}_{\sigma(1)}^{\sigma(k)}(\mathbf{i}, x_0) - \mathcal{T}_{\sigma(1)}^{\sigma(k-1)}(\mathbf{i}, x_0) : \mathbf{i} \in [\bar{i}]) \geq c \sum_{\bar{i} \in C_{\sigma,k}} \mu([\bar{i}])\mathbb{E}_\mu \mathcal{T}_{\sigma(k)}(f_{\bar{i}}^{\leftarrow}(x_0)) \geq ct\mu(A_{\sigma,k}),$$

where in the final inequality we used that  $\inf_{I \in \mathcal{I}} \inf_{x \in \Lambda \setminus I} \mathbb{E}_\mu \mathcal{T}_I(x) \geq t$ . Therefore

$$\mathbb{E}_\mu \mathcal{T}_{\mathcal{I}}(x_0) \geq ct \left( 1 + \sum_{k=2}^N \int \mu(A_{\sigma,k}) dm \right).$$

Therefore to prove the proposition it suffices to show that  $\int \mu(A_{\sigma,k}) dm = \frac{1}{k}$ . Fix  $2 \leq k \leq N$ . For each  $\sigma \in S_N$  consider the unordered set  $\{\sigma(1), \dots, \sigma(k)\}$ . Note that there are  $\frac{N(N-1)\dots(N-(k-1))}{k!}$  possible values that this set can take. For each possible value  $\{i_1, \dots, i_k\} \subset \{1, \dots, N\}$  that this set can take, let  $S_N(\{i_1, \dots, i_k\})$  denote the set of all  $\sigma$  for which  $\{\sigma(1), \dots, \sigma(k)\} = \{i_1, \dots, i_k\}$ , thinking of these as unordered sets.

Next, we can further separate each  $S_N(\{i_1, \dots, i_k\})$  into  $k$  subsets  $S_N^{i_j}(\{i_1, \dots, i_k\})$ , ( $1 \leq j \leq k$ ), which determines the set of all  $\sigma \in S_N(\{i_1, \dots, i_k\})$  for which  $\sigma(k) = i_j$ . Note that each  $S_N^{i_j}(\{i_1, \dots, i_k\})$  contains  $(N-k)!(k-1)!$  permutations, corresponding to  $(N-k)!$  ways to order the last  $N-k$  terms and  $(k-1)!$  ways to arrange the first  $k-1$  terms. Over each  $\sigma \in S_N^{i_j}(\{i_1, \dots, i_k\})$ , the set  $A_{\sigma,k}$  is constant. If for each  $1 \leq j \leq k$  we choose a representative  $\sigma_j \in S_N^{i_j}(\{i_1, \dots, i_k\})$  then since the balls in  $\mathcal{I}$  are pairwise disjoint, it follows that  $\{A_{\sigma_j,k}\}_{j=1}^k$  are pairwise disjoint and  $\bigcup_{j=1}^k A_{\sigma_j,k} = \Sigma$ .

Hence for any choice of  $\{i_1, \dots, i_k\} \subset \{1, \dots, N\}$ ,

$$\int_{S_N(\{i_1, \dots, i_k\})} \mu(A_{\sigma,k}) dm = \sum_{n=1}^k \frac{(k-1)!(N-k)!}{N!} \mu(A_{\sigma_n,k}), \quad (4.1)$$

where the factor  $\frac{1}{N!}$  comes from the fact that  $m$  is uniformly distributed. Now, since  $\bigcup_{j=1}^k A_{\sigma_j,k} = \Sigma$  and  $\{A_{\sigma_j,k}\}_{j=1}^k$  are pairwise disjoint we have

$$\mu(A_{\sigma_k,k}) = 1 - \sum_{j=1}^{k-1} \mu(A_{\sigma_j,k})$$

and substituting this into (4.1) we obtain

$$\begin{aligned} \int_{S_N(\{i_1, \dots, i_k\})} \mu(A_{\sigma, k}) dm &= \sum_{j=1}^{k-1} \frac{(k-1)!(N-k)!}{N!} \mu(A_{\sigma_j, k}) + \frac{(k-1)!(N-k)!}{N!} \left(1 - \sum_{j=1}^{k-1} \mu(A_{\sigma_j, k})\right) \\ &= \frac{(k-1)!(N-k)!}{N!}. \end{aligned}$$

Therefore,

$$\int_{S_N} \mu(A_{\sigma, k}) dm = \frac{N(N-1) \cdots (N-(k-1))}{k!} \cdot \frac{(k-1)!(N-k)!}{N!} = \frac{1}{k}.$$

□

We now establish the upper bound for  $\mathbb{E}_\mu T_r(x_0)$  from Theorem 2.4(a). By Proposition 4.1 it is sufficient to estimate from above the expected hitting time to the ball  $I$  of minimum measure at scale  $r$ . This can be estimated by bounding the probability of slow hitting times to  $I$ , and for this estimate the one-sided exponential decay of correlations assumption will be required to allow us to consider different segments of the orbit under the chaos game independently of each other.

**Lemma 4.2** (Proof of Theorem 2.4(a)). *Suppose  $\mu$  has one-sided exponential decay. Let  $\bar{\alpha} = \overline{\dim}_M(\overleftarrow{\mathcal{V}})$ . There exists a constant  $C_1$  (which will be made explicit) such that for all  $x_0 \in \Lambda$  and  $r > 0$  such that  $|\bar{o}(r/4)| < \bar{\alpha}/2$  and  $(r/4)^{\bar{\alpha}/2} < 1/2\kappa$ ,*

$$\mathbb{E}_\mu T_r(x_0) \leq C_1 (\log(4/r))^2 (r/4)^{-\bar{\alpha} - \bar{o}(r/4)},$$

where  $\kappa$  is the constant defined in (1.2).

*Proof.* Recall from (2.3) that  $\bar{o}(r)$  was defined as

$$\bar{o}(r) := \max_{x \in \Lambda} \frac{\log \overleftarrow{\mathcal{V}}(B(x, r))}{\log r} - \bar{\alpha}.$$

Fix  $\epsilon > 0$ . For each  $r > 0$  for short, let  $\mathcal{I}_r := \mathcal{Q}_{r/4}(\Lambda)$  be a maximal centred packing of  $\Lambda$  by disjoint balls of radius  $\frac{r}{4}$ . By (3.1)  $\log \#\mathcal{I}_r \leq \log C_0 + 2D \log(4/r)$  for all  $r > 0$ . Also, it'll be useful to keep in mind that  $\log(4/r) > \frac{2 \log 2\kappa}{\bar{\alpha}}$  by our assumptions on  $r$ . Since the collection of balls of radius  $\frac{r}{2}$  given by  $\{2I\}_{I \in \mathcal{I}_r}$  forms a cover of  $\Lambda$ , it is easy to see that  $\mathbb{E}_\mu T_r \leq \mathbb{E}_\mu \mathcal{T}_{\mathcal{I}_r}$ . We will use Proposition 4.1 to bound  $\mathbb{E}_\mu \mathcal{T}_{\mathcal{I}_r}$  from above. Therefore, we begin by showing that there exists a constant  $C'_1$  such that for all  $r$  such that  $|\bar{o}(r/4)| < \bar{\alpha}/2$  and  $(r/4)^{\bar{\alpha}/2} < 1/2\kappa$ ,

$$\sup_{I \in \mathcal{I}_r} \sup_{x \in \Lambda} \mathbb{E}_\mu \mathcal{T}_I(x) \leq \frac{C'_1 \log(4/r)}{(r/4)^{\bar{\alpha} + \bar{o}(r/4)}}.$$

By definition of  $\bar{o}(r)$ ,  $\overleftarrow{\mathcal{V}}(I) \geq (r/4)^{\bar{\alpha} + \bar{o}(r/4)}$  for all  $I \in \mathcal{I}_r$ . Recall from (3.2) that

$$L(r) := \frac{\log r}{\log a} - \frac{\log(|\Lambda|/a)}{\log a} \geq \max\{|\bar{i}| : \bar{i} \in \mathcal{P}_r\}$$

and define

$$\ell(r) := L(r) + \frac{2\bar{\alpha}}{\epsilon \log 2} \log(1/r).$$

Now, fix any  $x \in \Lambda$  and let  $\mathbf{j}$  be an arbitrary coding of  $x$ . Then by one-sided exponential decay (1.2)

$$\begin{aligned}
\mu(\{\mathbf{i} : \mathcal{T}_I(x) > n\}) &\leq \mu\left(\left\{\mathbf{i} : \mathcal{P}_r(\overleftarrow{\mathbf{i}}|_{k\ell(r/4)}\mathbf{j}) \notin I \text{ for every } 1 \leq k \leq \lceil n/\ell(r/4) \rceil\right\}\right) \\
&\leq (1 + 2^{-\varepsilon(\ell(r/4)-L(r/4))\kappa})^{\lceil n/\ell(r/4) \rceil} (1 - \overleftarrow{\nu}(I))^{\lceil n/\ell(r/4) \rceil} \\
&= (1 + (r/4)^{2\bar{\alpha}\kappa})^{\lceil n/\ell(r/4) \rceil} (1 - \overleftarrow{\nu}(I))^{\lceil n/\ell(r/4) \rceil} \\
&= (1 + (r/4)^{2\bar{\alpha}\kappa})^{\lceil n/\ell(r/4) \rceil} (1 - (r/4)^{\bar{\alpha}+\bar{o}(r/4)})^{\lceil n/\ell(r/4) \rceil} \\
&= (1 - (r/4)^{\bar{\alpha}+\bar{o}(r/4)} + (r/4)^{2\bar{\alpha}\kappa} - (r/4)^{3\bar{\alpha}+o(r/4)}\kappa)^{\lceil n/\ell(r/4) \rceil} \\
&\leq (1 - (r/4)^{\bar{\alpha}+\bar{o}(r/4)}/2)^{n/\ell(r/4)},
\end{aligned}$$

where in the final inequality we have used that

$$(r/4)^{2\bar{\alpha}\kappa} - (r/4)^{3\bar{\alpha}+o(r/4)}\kappa < (r/4)^{2\bar{\alpha}\kappa} < (r/4)^{\bar{\alpha}+\bar{o}(r/4)}/2,$$

since  $\bar{o}(r/4) < \bar{\alpha}/2$  and  $(r/4)^{\bar{\alpha}/2} < 1/2\kappa$ . Hence,

$$\begin{aligned}
\mathbb{E}_\mu \mathcal{T}_I(x) &= \sum_{n=0}^{\infty} \mu(\mathcal{T}_I(x) > n) \leq \sum_{n=0}^{\infty} (1 - (r/4)^{\bar{\alpha}+\bar{o}(r/4)}/2)^{n/\ell(r/4)} \\
&= \frac{1}{1 - (1 - (r/4)^{\bar{\alpha}+\bar{o}(r/4)}/2)^{1/\ell(r/4)}} \leq \frac{2\ell(r/4)}{(r/4)^{\bar{\alpha}+\bar{o}(r/4)}},
\end{aligned}$$

where in the last inequality we used the Bernoulli inequality (Mitrinovic inequality),  $(1+y)^\beta \leq 1 + \beta y$  for  $0 < \beta < 1$  and  $y \geq -1$ , which is applicable since  $\bar{o}(r/4) > -\bar{\alpha}$ . Therefore we have proved (4), where  $C'_1 = \frac{4\bar{\alpha}}{\varepsilon \log 2} - \frac{2}{\log a} - \bar{\alpha} \frac{\log |\Lambda|/a}{\log a \log 2\kappa}$ , using that  $\log(4/r) > \frac{2 \log 2\kappa}{\bar{\alpha}}$  and the definition of  $\ell(r/4)$ .

Since  $\mu$  satisfies the one-sided exponential decay property (1.2), there exists  $C > 1$  such that for all  $\bar{i}, \bar{j} \in \Sigma^*$ ,  $\mu(\bar{i}\bar{j}) \leq C\mu(\bar{i})\mu(\bar{j})$ . Hence by Proposition 4.1(a),

$$\begin{aligned}
\mathbb{E}_\mu T_r(x_0) &\leq \mathbb{E}_\mu \mathcal{T}_{\mathcal{I}_r}(x_0) \leq CC'_1 \log(1/r)(r/4)^{-\bar{\alpha}-\bar{o}(r/4)} \left(1 + \dots + \frac{1}{\#\mathcal{I}_r}\right) \\
&\leq CC'_1 \log(4/r)(r/4)^{-\bar{\alpha}-\bar{o}(r/4)} (1 + \log \#\mathcal{I}_r) \\
&\leq CC'_1 \log(4/r)(r/4)^{-\bar{\alpha}-\bar{o}(r/4)} (1 + \log C_0 + 2D \log(4/r)),
\end{aligned}$$

which completes the proof of the lemma by setting  $C_1 = CC'_1(2D + \bar{\alpha} \frac{1+\log C_0}{2 \log 2\kappa})$ , where again we have used that  $\log(4/r) > \frac{2 \log 2\kappa}{\bar{\alpha}}$ .  $\square$

Next we establish the lower bound on  $\mathbb{E}_\mu T_r(x_0)$  that appears in Theorem 2.4(b). Recall  $L(r)$  in (3.2) and  $d = \dim_{\mathbb{L}} \Lambda > 0$  in (3.4). Again by Proposition 4.1 it is sufficient to estimate from below the expected hitting time to the ball of minimum measure at scale  $r$ . We will use the assumption of positive lower dimension to ensure there are sufficiently many balls of measure comparable to that of the ball of minimum measure at scale  $r$ , so that starting the chaos game with an initial point  $x_0$  lying in the least accessible part of  $\Lambda$  will not cause a significant drop in the expectation  $\mathbb{E}_\mu T_r(x_0)$ .

**Lemma 4.3** (Proof of Theorem 2.4(b)). *Assume  $\dim_{\mathbb{L}} \Lambda > 0$ . Denote  $R_r := 2r \left( \frac{L(r)+2}{c_0} \right)^{2/d}$ . Then for all  $r > 0$  such that  $\min_{x \in \Lambda} \check{\nu}(B(x, R_r)) < \frac{1}{4}$  and all  $x_0 \in \Lambda$*

$$\mathbb{E}_{\mu} T_r(x_0) \geq \frac{1}{4} R_r^{-\alpha + \varrho(R_r)}.$$

*Proof.* Recall that  $\varrho(r)$  was defined as  $\varrho(r) := \max_{x \in \Lambda} \frac{\log \check{\nu}(B(x, r))}{\log r} - \underline{\alpha}$ . Let  $y' \in \Lambda$  be such that  $\check{\nu}(B(y', R_r))$  is minimal and note that

$$\check{\nu}(B(y', R_r)) = R_r^{-\alpha - \varrho(R_r)}. \quad (4.2)$$

Similarly to (3.7),

$$N_{2r}(\Lambda \cap B(y', R_r)) = N_{2r} \left( \Lambda \cap B \left( y', 2r \left( \frac{L(r)+2}{c_0} \right)^{2/d} \right) \right) \geq L(r) + 2.$$

Hence,  $\#\mathcal{Q}_r(\Lambda \cap B(y', R_r)) \geq L(r) + 2$ . So for every  $\mathbf{i} \in \Sigma$  there exists  $y(\mathbf{i}) \in \Lambda$  such that

$$B(y(\mathbf{i}), r) \cap \mathcal{O}_{\lceil L(n_k) \rceil}(\mathbf{i}, x_0) = \emptyset \text{ and } B(y(\mathbf{i}), r) \subset B(y', R_r).$$

Let  $\mathcal{Y}_r$  be the set of centres of the balls in  $\mathcal{Q}_r(\Lambda \cap B(y', R_r))$ . Recall the definition (3.5) of  $\mathcal{T}_r(\mathbf{i}, y, x_0)$ . So

$$\begin{aligned} \mathbb{E}_{\mu}(T_r(x_0)) &= \sum_{n=0}^{\infty} \mu(\mathbf{i} \in \Sigma : T_r(\mathbf{i}, x_0) > n) \geq \sum_{n=0}^{\infty} \mu(\mathbf{i} \in \Sigma : \mathcal{T}_r(\mathbf{i}, y(\mathbf{i}), x_0) > n) \\ &= \sum_{n=0}^{\infty} (1 - \mu(\mathbf{i} \in \Sigma : \mathcal{T}_r(\mathbf{i}, y(\mathbf{i}), x_0) \leq n)) = \sum_{n=0}^{\infty} \left( 1 - \sum_{y \in \mathcal{Y}_r} \mu(\mathbf{i} \in [\mathbf{i}] : \mathcal{T}_r(\mathbf{i}, y, x_0) \leq n) \right) \\ &\geq \sum_{n=0}^{\infty} \max\{0, 1 - n \sum_{y \in \mathcal{Y}_r} \check{\nu}(B(y, r))\} \geq \sum_{n=0}^{\infty} \max\{0, 1 - n \check{\nu}(B(y', R_r))\} \\ &= \sum_{n=0}^{\lceil \check{\nu}(B(y', R_r))^{-1} \rceil} \max\{0, 1 - n \check{\nu}(B(y', R_r))\} \\ &\geq \check{\nu}(B(y', R_r))^{-1} \left( \frac{1}{2} - \check{\nu}(B(y', R_r)) \right), \end{aligned}$$

which completes the proof by (4.2) and since  $\min_{x \in \Lambda} \check{\nu}(B(x, R_r)) < \frac{1}{4}$ .  $\square$

## 5. BEDFORD–MCMULLEN CARPETS

In this section, we give an explicit procedure to determine a vector solving the optimisation problem (2.6) in Proposition 5.4, prove the claims of Theorem 2.5 and provide some additional insight through examples. Recall all the notation introduced in Section 2.2, in particular, the function  $\alpha(\mathbf{q})$  and vectors  $\mathbf{q}_K, \mathbf{Q}_K$  from (2.7), (2.8), and (2.9), respectively.

To each vector  $\mathbf{q}_K$  and  $\mathbf{Q}_K$  we associate the vector  $\mathbf{p}_K = (p_{K,1}, \dots, p_{K,N})$  and  $\mathbf{P}_K = (P_{K,1}, \dots, P_{K,N})$  by distributing mass within columns uniformly, i.e. we set  $p_{K,i} = q_{K,k}/N_k$  if  $i$  belongs to a column with  $N_k$  rectangles and similarly  $P_{K,i} = Q_{K,k}/N_k$ . According to Theorem 2.5, one of these vectors solves the optimisation problem (2.6).



Observe that  $\mathbf{p}_1$  gives the *uniform measure*  $\mathbf{p}_1 = (1/N, \dots, 1/N)$  and  $\mathbf{p}_{M_0}$  gives the *coordinate uniform measure*  $\mathbf{p}_{M_0} = (1/(MN_{\psi(1)}), \dots, 1/(MN_{\psi(N)}))$ , where the function  $\psi : \{1, 2, \dots, N\} \rightarrow \{1, 2, \dots, M_0\}$  is defined

$$\psi(i) := k, \quad \text{if } i \text{ belongs to a column with } N_k \text{ rectangles.}$$

Moreover, the interpretation of the vectors  $\mathbf{P}_K$  is that all columns with at most  $N_K$  maps are given mass  $1 - \frac{\log m}{\log n}$  and this mass is further distributed within these columns in a coordinate uniform way. While the remaining  $\frac{\log m}{\log n}$  weight is given to columns with more than  $N_K$  rectangles and this mass is distributed uniformly between all rectangles in these columns.

The uniform measure is the ‘natural measure’ in the uniform vertical fibre case (recall, when all non-empty columns have the same number of rectangles). In the non-uniform vertical fibre case the coordinate uniform measure has the property that its lower and Assouad dimensions simultaneously realise the lower and Assouad dimension of the attractor provided the ‘*very strong separation condition*’ holds [20], see [18, Section 8.6] for additional information and definitions. In case of the Minkowski dimension there does not exist a self-affine measure  $\nu_{\mathbf{p}}$  for which  $\dim_{\mathbb{M}} \nu_{\mathbf{p}} = \dim_{\mathbb{M}} \Lambda$  (unless  $\Lambda$  has uniform vertical fibres).

**Claim 5.1.** *For a vector  $\mathbf{p} = (p_1, \dots, p_N)$ , let  $\mathbf{q}_{\mathbf{p}} = (q_{\mathbf{p},1}, \dots, q_{\mathbf{p},M})$  denote the vector defined as*

$$q_{\mathbf{p},j} := \text{sum of probabilities } p_i \text{ in the } \hat{j}\text{-th column.}$$

*Then, for any self-affine measure  $\nu_{\mathbf{p}}$  on a Bedford–McMullen carpet*

$$\dim_{\mathbb{M}} \nu_{\mathbf{p}} = \max_{\substack{1 \leq i \leq N \\ 1 \leq j \leq M}} \left\{ \frac{\log p_i}{-\log n} + \left(1 - \frac{\log m}{\log n}\right) \frac{\log q_{\mathbf{p},j}}{-\log m} \right\}. \quad (5.1)$$

*As a result,  $\min_{\mathbf{p}} \dim_{\mathbb{M}} \nu_{\mathbf{p}} = \dim_{\mathbb{M}} \Lambda$  if and only if  $\Lambda$  has uniform vertical fibres.*

Formula (5.1) is stated in [18, Theorem 8.6.2] in the case of the very strong separation condition. For convenience of the reader, we provide the short argument for any Bedford–McMullen carpet  $\Lambda$ .

*Proof.* For an index  $i \in \{1, \dots, N\}$  let  $\phi(i) \in \{1, \dots, M\}$  denote the index of the column to which the rectangle  $f_i([0, 1]^2)$  belongs to. Furthermore, for  $\mathbf{i} \in \Sigma = \{1, \dots, N\}^{\mathbb{N}}$  define  $\Phi(\mathbf{i}) = \phi(i_1)\phi(i_2)\dots$ . Then  $\Sigma$  endowed with the metric  $d(\mathbf{i}, \mathbf{j}) := m^{-|\Phi(\mathbf{i}) \wedge \Phi(\mathbf{j})|} + n^{-|\mathbf{i} \wedge \mathbf{j}|}$  is a complete metric space, where  $\mathbf{i} \wedge \mathbf{j}$  denotes the longest common prefix of  $\mathbf{i}$  and  $\mathbf{j}$ . A level  $K$  ball according to this metric is

$$B_K(\mathbf{i}) = \left\{ \mathbf{j} \in \Sigma : |\mathbf{i} \wedge \mathbf{j}| \geq L(K) \text{ and } |\Phi(\mathbf{i}) \wedge \Phi(\mathbf{j})| \geq K \right\}, \quad (5.2)$$

where  $L(K)$  is the unique integer such that  $m^{-K} \leq n^{-L(K)} < m^{-(K-1)}$ . We call  $B_K(\mathbf{i})$  a *symbolic approximate square* at level  $K$ . Let  $\mathcal{B}_K$  denote the set of level  $K$  approximate squares. Each  $B_K(\mathbf{i})$  can be identified with the sequence  $(i_1, \dots, i_{L(K)}, \phi(i_{L(K)+1}), \dots, \phi(i_K))$ .

As before, let  $\mu_{\mathbf{p}} = \mathbf{p}^{\mathbb{N}}$ . The  $\mu_{\mathbf{p}}$  measure of an approximate square  $B_K(\mathbf{i})$  is equal to

$$\mu_{\mathbf{p}}(B_K(\mathbf{i})) = \prod_{\ell=1}^{L(K)} p_{i_{\ell}} \cdot \prod_{\ell=L(K)+1}^K q_{\mathbf{p}, \phi(i_{\ell})}.$$

Let  $p^* = \min_i p_i$  and  $q^* = \min_j q_{\mathbf{p},j}$ . Then

$$\min \{ \mu_{\mathbf{p}}(B_K(\mathbf{i})) : B_K(\mathbf{i}) \in \mathcal{B}_K \} = (p^*)^{L(K)} \cdot (q^*)^{K-L(K)} = (m^{-K})^{\frac{\log p^*}{-\log n} + \left(1 - \frac{\log m}{\log n}\right) \frac{\log q^*}{-\log m}}.$$

This immediately implies that  $\dim_{\mathbb{M}} \mu_{\mathbf{p}}$  equals the formula in (5.1).

The next step is to show that  $\dim_{\mathbb{M}} \nu_{\mathbf{p}} = \dim_{\mathbb{M}} \mu_{\mathbf{p}}$ . The way that  $d(\mathbf{i}, \mathbf{j})$  is defined implies that up some uniform multiplicative constant  $d(\mathbf{i}, \mathbf{j}) \approx \|\pi(\mathbf{i}) - \pi(\mathbf{j})\|$ , where  $\pi : \Sigma \rightarrow \Lambda$  is the natural projection defined in (1.3). Hence, for any approximate square  $\text{diam}(\pi(B_K(\mathbf{i}))) \approx m^{-K} \approx n^{-L(K)}$ . Since  $\pi$  can only increase the measure of a ball, we get  $\dim_{\mathbb{M}} \nu_{\mathbf{p}} \leq \dim_{\mathbb{M}} \mu_{\mathbf{p}}$ .

To see the other inequality, consider the approximate square  $\pi(B_K(\mathbf{i})) \subset \Lambda$  whose  $\mu_{\mathbf{p}}$  measure is minimal. By assumption, on the first level at least two columns are non-empty and at least one of them has at least two maps. Thus, there is an  $x \in \pi(B_K(\mathbf{i}))$  and a constant  $c$  independent of  $K$  such that  $B(x, cn^{-(L(K)+2)}) \cap \Lambda \subset \pi(B_K(\mathbf{i}))$ . As a result,

$$\nu_{\mathbf{p}}(B(x, cn^{-(L(K)+2)})) \leq \mu_{\mathbf{p}}(B_K(\mathbf{i})) \approx (cn^{-(L(K)+2)})^{\dim_{\mathbb{M}} \mu_{\mathbf{p}}}.$$

This implies that  $\dim_{\mathbb{M}} \nu_{\mathbf{p}} \geq \dim_{\mathbb{M}} \mu_{\mathbf{p}}$ . Thus,  $\dim_{\mathbb{M}} \nu_{\mathbf{p}} = \dim_{\mathbb{M}} \mu_{\mathbf{p}}$ .

Finally,  $\dim_{\mathbb{M}} \Lambda = \frac{\log N}{\log n} + \left(1 - \frac{\log m}{\log n}\right) \frac{\log M}{\log m}$ . Hence,  $\min_{\mathbf{p}} \dim_{\mathbb{M}} \nu_{\mathbf{p}} = \dim_{\mathbb{M}} \Lambda$  if and only if both  $\mathbf{p}$  and  $\mathbf{q}_{\mathbf{p}}$  are the uniform vectors on  $\{1, \dots, N\}$  and  $\{1, \dots, M\}$ , respectively. This can happen only if each column has the same number of rectangles, i.e.  $\Lambda$  has uniform vertical fibres.  $\square$

**Remark 5.2.** The local dimension spectrum of Bedford–McMullen carpets was studied in [32] and [26]. In particular, they showed that the upper end of the spectrum equals

$$\max_{x \in \Lambda} \dim_{\text{loc}}(\nu_{\mathbf{p}}, x) = \max_{1 \leq i \leq N} \left\{ \frac{\log p_i}{-\log n} + \left(1 - \frac{\log m}{\log n}\right) \frac{\log q_{\mathbf{p}, \psi(i)}}{-\log m} \right\}.$$

This differs from (5.1) only in that the coordinate of  $\mathbf{q}_{\mathbf{p}}$  can not be chosen independently, it has to be the coordinate corresponding to the column of  $i$ . Hence,  $\max_{x \in \Lambda} \dim_{\text{loc}}(\nu_{\mathbf{p}}, x) \leq \dim_{\mathbb{M}} \nu_{\mathbf{p}}$  and there is a strict inequality if  $p_i$  and  $q_{\mathbf{p},j}$  are minimised in different columns. This is the case for the vectors  $\mathbf{P}_K$ . Such a phenomena does not hold for self-similar sets.

The next claim shows that we can reduce the minimisation problem  $\min_{\mathbf{p}} \dim_{\mathbb{M}} \nu_{\mathbf{p}}$  to the lower dimensional problem  $\min_{\mathbf{q} \in \mathcal{Q}} \alpha(\mathbf{q})$ , where  $\alpha(\mathbf{q})$  was defined in (2.7) and  $\mathcal{Q} = \{\mathbf{q} = (q_1, \dots, q_{M_0}) : R_1 q_1 + \dots + R_{M_0} q_{M_0} = 1 \text{ and } q_k > 0 \text{ for all } 1 \leq k \leq M_0\}$ .

**Claim 5.3.** *If  $\mathbf{q} = (q_1, \dots, q_{M_0})$  is such that  $\alpha(\mathbf{q}) = \min_{\mathbf{q}' \in \mathcal{Q}} \alpha(\mathbf{q}')$ , then  $\mathbf{q}$  has the following two properties:*

- (1)  $\min_{\mathbf{p}} \dim_{\mathbb{M}} \nu_{\mathbf{p}} = \alpha(\mathbf{q})$ , where  $\mathbf{q}$  gives a solution  $\mathbf{p}^* = (p_1^*, \dots, p_N^*)$  to (2.6) by setting  $p_i^* := q_k/N_k$  if  $i$  belongs to a column with  $N_k$  rectangles (i.e. mass is distributed uniformly within columns);
- (2) There exists a unique  $1 \leq K \leq M_0 - 1$  for which

$$q_1 = q_2 = \dots = q_K \leq \min_{K+1 \leq k \leq M_0} q_k \quad \text{and} \quad \frac{q_{K+1}}{N_{K+1}} = \frac{q_{K+2}}{N_{K+2}} = \dots = \frac{q_{M_0}}{N_{M_0}} \leq \min_{1 \leq k \leq K} \frac{q_k}{N_k}. \quad (5.3)$$

*Proof.* First observe that the maximization over the indices  $k$  and  $\ell$  in the definition (2.7) of  $\alpha(\mathbf{q})$  is independent. Moreover,  $-\log q_k/N_k$  and  $-\log q_\ell$  are maximal if and only if  $q_k/N_k$  and  $q_\ell$  are minimal, respectively.

Assertion (1) simply follows from the fact that the value of  $-\log \min_i p_i$  can only increase in the formula (5.1) for  $\dim_{\mathbf{M}} \nu_{\mathbf{p}}$  if mass within each column is not distributed uniformly. This argument also implies that for each  $k = 1, \dots, M_0$ , the mass within the collection of  $R_k$  columns should also be uniformly distributed amongst the columns. Thus, it is enough to consider the  $M_0 - 1$  dimensional problem  $\min_{\mathbf{q} \in \mathcal{Q}} \alpha(\mathbf{q})$  to solve  $\min_{\mathbf{p}} \dim_{\mathbf{M}} \nu_{\mathbf{p}}$ .

Assume  $\mathbf{q} = (q_1, \dots, q_{M_0})$  is such that  $\alpha(\mathbf{q}) = \min_{\mathbf{q}' \in \mathcal{Q}} \alpha(\mathbf{q}')$ . To see (2), first notice that for each  $k = 1, \dots, M_0$  at least one of the following holds:

$$(a) \ q_k = \min_{1 \leq \ell \leq M_0} q_\ell, \quad (b) \ q_k/N_k = \min_{1 \leq \ell \leq M_0} q_\ell/N_\ell.$$

This is because if there existed an index  $1 \leq k_0 \leq M_0$  such that  $q_{k_0} > \min_{\ell} q_\ell$  and  $q_{k_0}/N_{k_0} > \min_{\ell} q_\ell/N_\ell$ , then mass could be transferred from that column into the columns attaining (either) minimum, therefore reducing the maximum in the definition of  $\alpha(\mathbf{q})$ .

Secondly, observe that if  $1 \leq K_1 \neq K_2 \leq M_0$  are two distinct indices such that  $q_{K_1} = \min_{1 \leq \ell \leq M_0} q_\ell$  and  $q_{K_2}/N_{K_2} = \min_{1 \leq \ell \leq M_0} q_\ell/N_\ell$ , then  $K_1 < K_2$ . Indeed by the choice of  $K_1$  and  $K_2$

$$q_{K_2}/N_{K_2} \leq q_{K_1}/N_{K_1} \leq q_{K_2}/N_{K_1} \implies N_{K_1} \leq N_{K_2}.$$

Since  $N_k$  are in ascending order and  $K_1 \neq K_2$ , it follows that  $K_1 < K_2$ . In particular (2) holds.

The uniqueness of  $K$  follows by the simple observation that since  $N_1 < \dots < N_{M_0}$  and  $q_{K+i}/N_{K+i} = q_{K+i+1}/N_{K+i+1}$  for  $i = 1, \dots, M_0 - K - 1$  we have  $q_{K+i+1} > q_{K+i}$  for  $i = 1, \dots, M_0 - K - 1$ .  $\square$

**5.1. Finding a vector minimising  $\dim_{\mathbf{M}} \nu_{\mathbf{p}}$ .** Claim 5.3 shows that to find a vector that minimises  $\dim_{\mathbf{M}} \nu_{\mathbf{p}}$ , it is enough to find a vector that minimises  $\alpha(\mathbf{q})$  and then distribute mass within columns evenly amongst rectangles. The next proposition shows how to find a vector which minimises  $\alpha(\mathbf{q})$ .

Recall, for an index  $1 \leq K \leq M_0$ , we denote  $|\mathcal{R}_K| = R_1 + \dots + R_K$  and  $\|\mathcal{R}_K^C\| = R_{K+1}N_{K+1} + \dots + R_{M_0}N_{M_0}$ . Also recall the definitions of  $\mathbf{q}_K$  and  $\mathbf{Q}_K$  from (2.8) and (2.9). For each  $1 \leq K \leq M_0 - 1$  let

$$A_K := \left( \frac{\log n}{\log m} - 1 \right) \frac{\|\mathcal{R}_K^C\|}{|\mathcal{R}_K|}. \quad (5.4)$$

**Proposition 5.4.** *Assume the parameters  $n, m, N_1, \dots, N_{M_0}, R_1, \dots, R_{M_0}$  define a Bedford–McMullen carpet  $\Lambda$  with non-uniform vertical fibres. Then*

$$\min_{\mathbf{p}} \dim_{\mathbf{M}} \nu_{\mathbf{p}} = \min \{ \alpha_1, \alpha_2, \dots, \alpha_{M_0-1} \},$$

where for  $1 \leq K \leq M_0 - 1$ ,

$$\alpha_K := \min \{ \alpha(\mathbf{q}_K), \alpha(\mathbf{Q}_K), \alpha(\mathbf{q}_{K+1}) \} = \begin{cases} \alpha(\mathbf{q}_K), & \text{if } A_K < N_K, \\ \alpha(\mathbf{Q}_K), & \text{if } N_K \leq A_K \leq N_{K+1}, \\ \alpha(\mathbf{q}_{K+1}), & \text{if } A_K > N_{K+1}. \end{cases}$$

We note that  $\alpha(\mathbf{q}_K) = \alpha(\mathbf{Q}_K)$  if  $A_K = N_K$ , similarly,  $\alpha(\mathbf{q}_{K+1}) = \alpha(\mathbf{Q}_K)$  if  $A_K = N_{K+1}$ . Before turning to the proofs, we give examples when  $M_0 = 2$  or  $3$  and when the optimiser solving  $\min_{\mathbf{p}} \dim_{\mathbf{M}} \nu_{\mathbf{p}}$  is not unique.

**5.2. Proof of Theorem 2.5 and Proposition 5.4.** Recall from Claim 5.3 that it is enough to consider a vector  $\mathbf{q} = (q_1, \dots, q_{M_0})$  which satisfies (5.3) for some  $1 \leq K \leq M_0 - 1$ . Recall that  $|\mathcal{R}_K| = R_1 + \dots + R_K$  and  $\|\mathcal{R}_K^C\| = R_{K+1}N_{K+1} + \dots + R_{M_0}N_{M_0}$ . Therefore, from (5.3) we can express  $q_{K+1}/N_{K+1}$  in terms of  $q_K$  using that

$$1 = \sum_{\ell=1}^{M_0} R_\ell q_\ell = |\mathcal{R}_K|q_K + \|\mathcal{R}_K^C\| \frac{q_{K+1}}{N_{K+1}} \iff \frac{q_{K+1}}{N_{K+1}} = \frac{1 - |\mathcal{R}_K|q_K}{\|\mathcal{R}_K^C\|}. \quad (5.5)$$

Combining (5.3) and (5.5) yields the inequality

$$\frac{q_K}{N_{K+1}} \stackrel{(5.3)}{\leq} \frac{q_{K+1}}{N_{K+1}} \stackrel{(5.5)}{=} \frac{1 - |\mathcal{R}_K|q_K}{\|\mathcal{R}_K^C\|} \stackrel{(5.3)}{\leq} \frac{q_K}{N_K}.$$

After rearranging and using that  $|\mathcal{R}_{K+1}| = |\mathcal{R}_K| + R_{K+1}$  and  $\|\mathcal{R}_{K+1}^C\| = \|\mathcal{R}_K^C\| - R_{K+1}N_{K+1}$ , we get the condition

$$\frac{N_K}{N_K|\mathcal{R}_K| + \|\mathcal{R}_K^C\|} \leq q_K \leq \frac{N_{K+1}}{N_{K+1}|\mathcal{R}_{K+1}| + \|\mathcal{R}_{K+1}^C\|}. \quad (5.6)$$

Also, substituting (5.5) back into  $\alpha(\mathbf{q})$ , recall (2.7), we obtain a one variable function in  $q_K$ :

$$f_K(q_K) := \frac{\log \frac{1 - |\mathcal{R}_K|q_K}{\|\mathcal{R}_K^C\|}}{-\log n} + \left(1 - \frac{\log m}{\log n}\right) \frac{\log q_K}{-\log m}.$$

Therefore, to obtain a solution to  $\min_{\mathbf{q}} \alpha(\mathbf{q})$  it is enough to minimise  $f_K(q_K)$  subject to condition (5.6).

To minimise  $f_K(q_K)$ , first observe that the equation  $\frac{d}{dq_K} f_K(q_K) = 0$  yields the unique solution  $q_K^* = \left(1 - \frac{\log m}{\log n}\right) / |\mathcal{R}_K|$ . This is indeed a minimum, since the second derivative

$$\frac{d^2}{d(q_K)^2} f_K(q_K) = \frac{|\mathcal{R}_K|^2}{(1 - |\mathcal{R}_K|q_K)^2 \log n} + \left(1 - \frac{\log m}{\log n}\right) \frac{1}{(q_K)^2 \log m} > 0$$

for any  $q_K$ , in particular, also for  $q_K = q_K^*$ . Hence,

$$f_K(q_K) \text{ strictly decreases for } q_K < q_K^* \text{ and } f_K(q_K) \text{ strictly increases for } q_K > q_K^*. \quad (5.7)$$

Recall  $A_K$  from (5.4). The condition  $N_K \leq A_K \leq N_{K+1}$  from Proposition 5.4 is equivalent to

$$\frac{N_K}{N_K|\mathcal{R}_K| + \|\mathcal{R}_K^C\|} \leq \left(1 - \frac{\log m}{\log n}\right) \frac{1}{|\mathcal{R}_K|} \leq \frac{N_{K+1}}{N_{K+1}|\mathcal{R}_{K+1}| + \|\mathcal{R}_{K+1}^C\|},$$

which can be seen by rearranging and using again that  $|\mathcal{R}_{K+1}| = |\mathcal{R}_K| + R_{K+1}$  and  $\|\mathcal{R}_{K+1}^C\| = \|\mathcal{R}_K^C\| - R_{K+1}N_{K+1}$ . In particular, this implies that the global minimum  $q_K^*$  for  $f_K$  satisfies the bounds in (5.6). Substituting  $q_K = \left(1 - \frac{\log m}{\log n}\right) / |\mathcal{R}_K|$  into (5.5) we recover the measure  $\mathbf{Q}_K$  defined in (2.8). Therefore, if  $N_K \leq A_K \leq N_{K+1}$  then

$$\min_{\substack{\mathbf{q}: \\ (5.3) \text{ holds for } K}} \alpha(\mathbf{q}) = \alpha(\mathbf{Q}_K).$$

Next, note that if  $A_K < N_K$  then similarly to the above, we can deduce that the global minima of  $f_K$  satisfies

$$q_K^* = \left(1 - \frac{\log m}{\log n}\right) \frac{1}{|\mathcal{R}_K|} < \frac{N_K}{N_K |\mathcal{R}_K| + \|\mathcal{R}_K^C\|}.$$

In particular, the global minima for  $f_K$  is not in the range determined by (5.6). Instead, (5.7) implies that the minimum of  $f_K$  subject to (5.6) is obtained at  $q_K = \frac{N_K}{N_K |\mathcal{R}_K| + \|\mathcal{R}_K^C\|}$ . Substituting  $q_K = \frac{N_K}{N_K |\mathcal{R}_K| + \|\mathcal{R}_K^C\|}$  into (5.5), we recover the measure  $\mathbf{q}_K$  defined in (2.8). Hence, if  $A_K < N_K$  then

$$\min_{\mathbf{q}: \text{(5.3) holds for } K} \alpha(\mathbf{q}) = \alpha(\mathbf{q}_K).$$

Finally, if  $A > N_{K+1}$  then the global minima of  $f_K$  satisfies

$$q_K^* = \left(1 - \frac{\log m}{\log n}\right) \frac{1}{|\mathcal{R}_K|} > \frac{N_{K+1}}{N_{K+1} |\mathcal{R}_{K+1}| + \|\mathcal{R}_{K+1}^C\|}.$$

This time (5.7) implies that  $q_K = \frac{N_{K+1}}{N_{K+1} |\mathcal{R}_{K+1}| + \|\mathcal{R}_{K+1}^C\|}$  is where the minimum of  $f_K$  subject to (5.6) is attained. Substituting this  $q_K$  into (5.5), we recover the measure  $\mathbf{q}_{K+1}$ . Hence, if  $A > N_{K+1}$  then

$$\min_{\mathbf{q}: \text{(5.3) holds for } K} \alpha(\mathbf{q}) = \alpha(\mathbf{q}_{K+1}).$$

To conclude, we deduce from Claim 5.3 that  $\min_{\mathbf{p}} \dim_{\mathbb{M}} \nu_{\mathbf{p}} = \min \{\alpha_1, \alpha_2, \dots, \alpha_{M_0-1}\}$ . In particular by Claim 5.3 item (2) we can deduce the form that the optimising vector  $\mathbf{p}^*$  for (2.6) takes. This completes the proof of Theorem 2.5 and Proposition 5.4.

**5.3. Special case with two different columns.** The input parameters of a Bedford–McMullen carpet with two different columns are:  $n > m, N_1 < N_2 \leq n$  and  $R_1 + R_2 \leq m$ . Assume the indices of the maps  $f_i$  defining the carpet are ordered such that the first  $R_1 N_1$  belong to columns with  $N_1$  rectangles. By Proposition 5.4,

$$\min_{\mathbf{p}} \dim_{\mathbb{M}} \nu_{\mathbf{p}} = \alpha_1 = \min\{\alpha(\mathbf{q}_1), \alpha(\mathbf{q}_2), \alpha(\mathbf{Q}_1)\},$$

therefore the optimising vector is either the uniform measure  $\mathbf{p}_1 = (1/N, \dots, 1/N)$ , the coordinate uniform measure  $\mathbf{p}_2 = (1/(MN_{\psi(1)}), \dots, 1/(MN_{\psi(N)}))$  or  $\mathbf{P}_1$  defined by

$$P_{1,i} = \left(1 - \frac{\log m}{\log n}\right) \frac{1}{R_1 N_1} \text{ for all } i \leq R_1 N_1 \quad \text{and} \quad P_{1,i} = \frac{\log m}{\log n} \frac{1}{R_2 N_2} \text{ for all } i > R_1 N_1.$$

In this case  $A_1$  is given by

$$A_1 = \left(\frac{\log n}{\log m} - 1\right) \frac{R_2 N_2}{R_1}.$$

Table 1 shows three examples where  $\mathbf{p}_1$ ,  $\mathbf{P}_1$  and  $\mathbf{p}_2$  are the optimising vectors, respectively. Note that in the last example,  $\mathbf{q}_2$  is the optimizer regardless of the choice of  $1 \leq N_1 < N_2 \leq n$ .

	$R_1$	$R_2$	$N_1$	$N_2$	$m$	$n$	$\alpha(\mathbf{q}_1)$	$\alpha(\mathbf{Q}_1)$	$\alpha(\mathbf{q}_2)$	$\dim_{\mathbb{M}} \Lambda$
$A_1 < N_1$	1	1	2	3	2	3	1.95286	–	2	1.83404
$N_1 \leq A_1 \leq N_2$	1	1	1	2	2	3	1.58496	1.58089	1.63093	1.36907
$N_2 < A_1$	1	1	2	3	2	5	1.75260	–	1.68261	1.56932

TABLE 1. Three examples in which each minimises  $\dim_{\mathbb{M}} \nu_{\mathbf{p}}$  for a different vector.

**5.4. Simulations.** We demonstrate how to keep track of the orbit and how the choice of the measure  $\mu$  driving the chaos game influences the “quality” of the image on one of the examples presented in Table 1.

For any  $r > 0$  consider a maximal  $r$ -packing of the attractor  $\Lambda$ , i.e. a collection of sets of diameter  $r$  with disjoint interiors that cover  $\Lambda$ . The orbit  $\mathcal{O}_n(\mathbf{i}, x_0)$  becomes  $r$ -dense in  $\Lambda$  once it has visited all elements of the  $r$ -packing. For Bedford–McMullen carpets we can keep track of the orbit using the collection of symbolic approximate squares of level  $K$  introduced in (5.2), where  $K$  is chosen so that  $m^{-K} \leq r < m^{-K+1}$ . Recall, each approximate square  $B_K(\mathbf{i})$  is identified with the sequence  $(i_1, \dots, i_{L(K)}; \phi(i_{L(K)+1}), \dots, \phi(i_K))$ , where  $\phi(i) \in \{1, \dots, M\}$  denotes the index of the column to which the rectangle  $f_i([0, 1]^2)$  belongs to. One step of the chaos game corresponds to the transition

$$(i_1, \dots, i_{L(K)}; \phi(i_{L(K)+1}), \dots, \phi(i_K)) \mapsto (j, i_1, \dots, i_{L(K)-1}; \phi(i_{L(K)}), \dots, \phi(i_{K-1}))$$

if  $j$  was the next chosen index.

Table 2 shows the cover times in the middle example of Table 1 for various vectors at two different scales  $K = 6$  and  $K = 9$ . The cover times are averaged out over 400 independent runs of the chaos game when  $K = 6$ , and averaged out over 100 independent runs for  $K = 9$ . The vectors  $\mathbf{p}_1$ ,  $\mathbf{P}_1$  and  $\mathbf{p}_2$  are the same as in Section 5.3, while  $\widehat{\mathbf{p}}$  corresponds to the McMullen measure which maximises the Hausdorff dimension of  $\nu_{\mathbf{p}}$  defined as

$$\widehat{p}_k = N_i^{\frac{\log m}{\log n} - 1} \cdot \left( \sum_{j=1}^M N_j^{\frac{\log m}{\log n}} \right)^{-1} \quad \text{if } \phi(k) = \hat{i}.$$

The pair  $(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}})$  corresponds to the uniform vectors on  $\{1, \dots, N\}$  and  $\{1, \dots, M\}$ , respectively. In this case, we modified the chaos game to “two dimensions”: in a transition step a new map *and* a new column are chosen *independently* of each other according to  $\widetilde{\mathbf{p}}$  and  $\widetilde{\mathbf{q}}$ . That is, if  $k \in \{1, \dots, N\}$  and  $\ell \in \{1, \dots, M\}$  are chosen uniformly and independently, then a transition step is

$$(i_1, \dots, i_{L(K)}; j_{L(K)+1}, \dots, j_K) \mapsto (k, i_1, \dots, i_{L(K)-1}; \ell, j_{L(K)+1}, \dots, j_{K-1}).$$

The choice  $(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}})$  is optimal in the sense that it minimises

$$\alpha(\mathbf{p}, \mathbf{q}) := \max_{\substack{1 \leq i \leq N \\ 1 \leq j \leq M}} \left\{ \frac{\log p_i}{-\log n} + \left( 1 - \frac{\log m}{\log n} \right) \frac{\log q_j}{-\log m} \right\}$$

with value  $\alpha(\widetilde{\mathbf{p}}, \widetilde{\mathbf{q}}) = \dim_{\mathbb{M}} \Lambda$ . Hence, is the most efficient possible. However, the drawback of this modified chaos game is that it does not have any clear geometric interpretation.

Indeed, Table 2 shows that the runtime is substantially faster with  $(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$  than with the other vectors. Table 1 suggests that the runtime with  $\mathbf{P}_1$  should be second fastest since it corresponds to the smallest exponent  $\alpha$ , however, this is not supported by the empirical data in Table 2. One explanation for this could be the fact that for small values of  $K$  the measure of  $B_K(\mathbf{i}^*)$ , which denotes the level  $K$  approximate square of minimum measure, does not yet reflect the asymptotic behaviour inferred from the exponent  $\alpha$ . For example, by direct computation we obtain that at level  $K = 6$

$$0.00137 = \mu_{\mathbf{p}_1}(B_6(\mathbf{i}^*)) < \mu_{\mathbf{P}_1}(B_6(\mathbf{i}^*)) < \mu_{\hat{\mathbf{p}}}(B_6(\mathbf{i}^*)) < \mu_{\mathbf{p}_2}(B_6(\mathbf{i}^*)) = 0.00195,$$

whereas for level  $K = 100$

$$8.55 \times 10^{-50} = \mu_{\mathbf{p}_2}(B_K(\mathbf{i}^*)) < \mu_{\mathbf{P}_1}(B_K(\mathbf{i}^*)) < \mu_{\hat{\mathbf{p}}}(B_K(\mathbf{i}^*)) < \mu_{\mathbf{p}_1}(B_K(\mathbf{i}^*)) = 2.61 \times 10^{-48}.$$

Another reason could be that the impact of the subexponential error terms on the runtime is amplified for small values of  $K$ . For larger  $K$ , we expect the runtime would follow the order of the exponents but, since the number of approximate squares grows exponentially in  $K$ , it is computationally impossible to do simulations for much larger  $K$ .

	$R_1$	$R_2$	$N_1$	$N_2$	$m$	$n$	$\mathbf{p}_1$	$\mathbf{P}_1$	$\mathbf{p}_2$	$\hat{\mathbf{p}}$	$(\tilde{\mathbf{p}}, \tilde{\mathbf{q}})$
$K = 6$	1	1	1	2	2	3	2787	2202	2442	2060	1288
$K = 9$	1	1	1	2	2	3	118057	86445	112666	78910	33855

TABLE 2. The runtime of the chaos game on the same example with different vectors until it visits all approximate squares at level  $K$ .

In Figure 2 we plotted the orbit starting from  $(1, 1) \in \Lambda$  using three different vectors; from left to right  $\mathbf{P}_1, \mathbf{p}_1$  and  $(0.6, 0.25, 0.15)$ . The orbits were terminated once the game with  $\mathbf{P}_1$  visited all approximate squares at level  $K = 7$ . It is difficult to see any difference between the first two figures with the naked eye. However, in the third case where the vector was chosen deliberately to be very different from the optimal, it is apparent that it hasn't visited many approximate squares in the same number of steps.

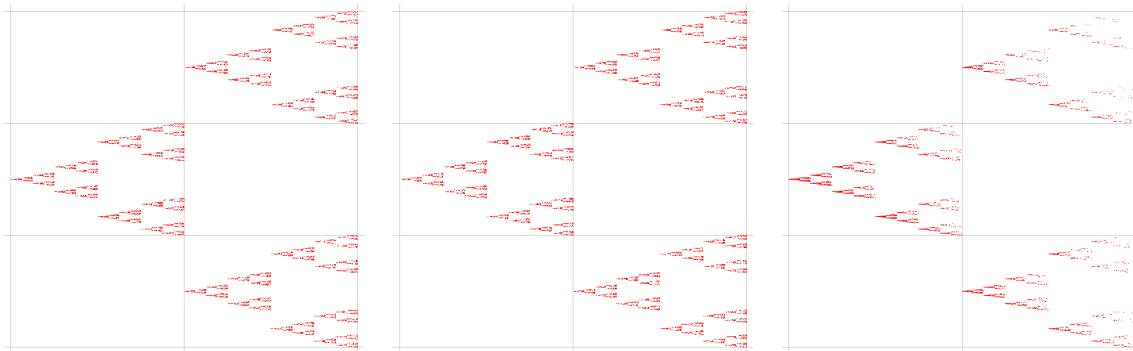


FIGURE 2. Plots of orbits using three different vectors (from left to right  $\mathbf{P}_1, \mathbf{p}_1$  and  $(0.6, 0.25, 0.15)$ ) terminated when the game with  $\mathbf{P}_1$  visited all level-7 approximate squares.



**5.5. Non-unique optimiser.** Let us again consider a Bedford–McMullen carpet with two different columns, where  $m = 2$ ,  $n = 4$ ,  $R_1 = R_2 = 1$  and  $N_1 = 2$ ,  $N_2 = 4$ . Assume the indices of the maps  $f_i$  defining the carpet are ordered such that the first 2 belong to the column that contains 2 rectangles. We know that  $\min_{\mathbf{p}} \dim_{\mathbb{M}} \nu_{\mathbf{p}} = \alpha_1 = \alpha(\mathbf{Q}_1)$  since  $A_1 = 4 = N_2$ . Moreover  $\mathbf{Q}_1 = \mathbf{q}_2 = (\frac{1}{2}, \frac{1}{2})$ , meaning that the co-ordinate uniform vector

$$\mathbf{P}_1 = \left( \frac{1}{4}, \frac{1}{4}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8} \right)$$

is an optimiser for (2.6).

Now, for  $0 \leq \varepsilon \leq \frac{1}{8}$  define the perturbed vector

$$\mathbf{P}_1^\varepsilon = \left( \frac{1}{4} - \varepsilon, \frac{1}{4} + \varepsilon, \frac{1}{8}, \frac{1}{8}, \frac{1}{8}, \frac{1}{8} \right).$$

From (5.1) one can check that  $\dim_{\mathbb{M}} \nu_{\mathbf{P}_1^\varepsilon}$  is actually independent of  $\varepsilon$ , thus each  $\mathbf{P}_1^\varepsilon$  is also an optimising vector for  $0 \leq \varepsilon \leq \frac{1}{8}$ . In particular, the optimising vector for (2.6) is not unique.

**5.6. Special case with three different columns.** The input parameters of a Bedford–McMullen carpet with three different columns are:  $n > m$ ,  $N_1 < N_2 < N_3 \leq n$  and  $R_1 + R_2 + R_3 \leq m$ . The parameter space, shown in Figure 3, consists of vectors  $\mathbf{q} = (q_1, q_2, q_3)$  such that  $R_1 q_1 + R_2 q_2 + R_3 q_3 = 1$ . It may be that  $N_2/N > 1/M$ , however, the line connecting  $\mathbf{q}_1$  to  $(1/R_1, 0)$  always intersects the line  $q_1 = q_2$  (defining  $\mathbf{q}_2$ ) and never the line connecting  $\mathbf{q}_3$  with  $(0, 1/R_2)$ .

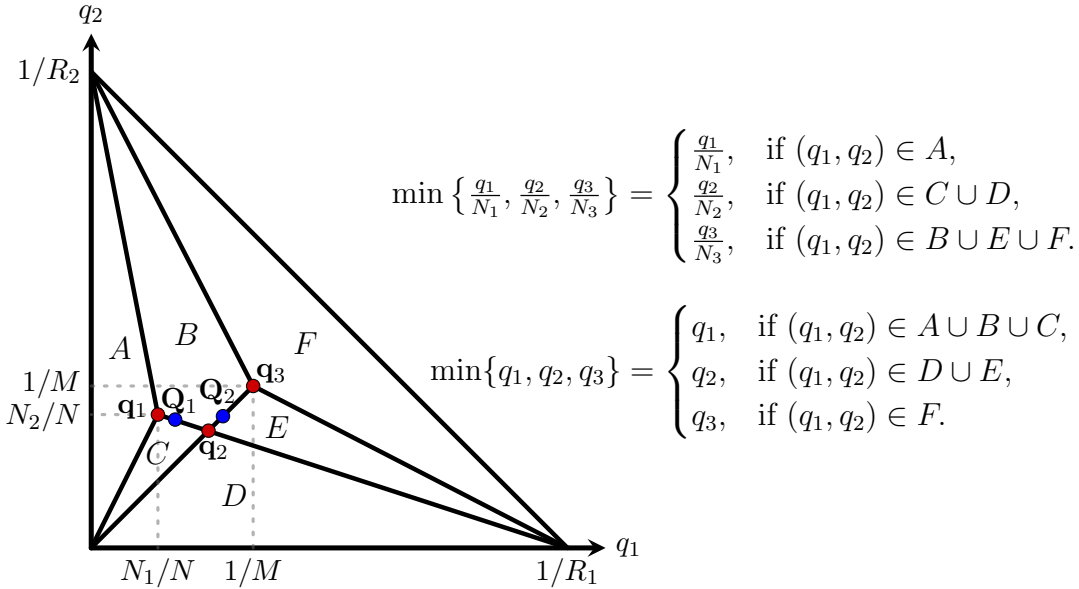


FIGURE 3. The parameter space of vectors  $\mathbf{q} = (q_1, q_2, (1 - R_1 q_1 - R_2 q_2)/R_3)$  shown with the six regions where  $\alpha(\mathbf{q})$  takes different values.

The function  $\alpha(\mathbf{q})$  takes different values according to which region the vector  $\mathbf{q}$  falls into. The candidates that minimise  $\alpha(\mathbf{q})$  are the vectors  $\mathbf{q}_1, \mathbf{q}_2, \mathbf{q}_3, \mathbf{Q}_1$  and  $\mathbf{Q}_2$ .  $\mathbf{Q}_1$  is a valid candidate if and only if it lies on the line between  $\mathbf{q}_1$  and  $\mathbf{q}_2$ . Likewise for  $\mathbf{Q}_2$  between  $\mathbf{q}_2$

and  $\mathbf{q}_3$ . We leave it to the reader to find five different examples, where each one minimizes  $\alpha(\mathbf{q})$  for a different candidate. We note that the McMullen measure which maximizes the Hausdorff dimension of  $\nu_{\mathbf{p}}$  lies somewhere in the interior of region  $B$ .

## 6. OPEN QUESTIONS

Here we suggest some possible directions for future investigation appearing naturally from the results of the presented paper.

**Question 6.1.** Can Theorem 2.1 be extended beyond the class of measures that satisfy the one-sided exponential decay of correlations property? For example, to measures which satisfy a weaker mixing property, or even to all ergodic measures?

**Question 6.2.** In Proposition 3.1, is it possible to omit the assumption that  $\dim_{\mathbb{L}} \Lambda > 0$  if the IFS is overlapping? In point of view of Proposition 3.3 is it possible to relax the conditions even further?

**Question 6.3.** Is it possible to relax the hyperbolicity condition in Theorem 2.1? That is, is it possible to relax the statement to allow for parabolic iterated function systems  $\mathcal{F}$ , where  $\mathcal{F}$  is no longer uniformly contracting, but it contains a map with a neutral fixed point?

**Question 6.4.** At least in special cases, as was shown in [27] for self-similar sets, is it possible to obtain tighter bounds on the expected value of the cover time in Theorem 2.4?

**Question 6.5.** Let us consider a simple iterated function system of similarities with the strong separation condition, and let us consider the natural self-similar measure  $\mu$  defined by the similarity dimension, and fix an  $r > 0$ . Is it possible to give tight bounds on  $n$  for which  $\mu(\{\mathbf{i} : d_{\mathbb{H}}(\mathcal{O}_n(\mathbf{i}, x_0), \Lambda) < r\}) \geq 0.95$ ? Is it possible in the case of the weak separation condition? Or even without any separation condition?

**Question 6.6.** In case of Bernoulli convolutions, recall (2.4), all self-similar measures have Minkowski dimension strictly larger than 1. Can it be achieved if the probability vector  $\mathbf{p}$  is allowed to depend on the current position of the orbit? More precisely, let  $p : \Lambda \rightarrow [0, 1]$ . Assuming some regularity on  $p$ , see [15], there exists a unique measure  $\nu$  satisfying

$$\nu(B) = \int_{(B+1)/\lambda} p(x) d\nu(x) + \int_{(B-1)/\lambda} (1 - p(x)) d\nu(x)$$

for any Borel set  $B$ . Can  $\dim_{\mathbb{M}} \nu = 1$  be achieved with an appropriate choice of  $p$ ?

**Question 6.7.** Recall that the Minkowski dimensions of Bernoulli measures supported on a Bedford-McMullen carpet  $\Lambda$  obtain a maximum at  $\dim_{\mathbb{M}} \Lambda$  if and only if  $\Lambda$  has uniform vertical fibres (Claim 5.1). Therefore it is interesting to ask whether, in the non-uniform vertical fibres case, there exists an invariant measure supported on  $\Lambda$  whose Minkowski dimension equals  $\dim_{\mathbb{M}} \Lambda$  (thus which optimises the cover time for the chaos game)? Note that although it is known that  $\dim_{\mathbb{M}} \Lambda$  is always achieved by the Minkowski dimension of some measure supported on  $\Lambda$ , this measure is not necessarily invariant [14].

## REFERENCES

- [1] D. J. Aldous, *Covering a compact space by fixed-radius or growing random balls*, arXiv e-prints (2021), arXiv:2101.12671, available at [2101.12671](https://arxiv.org/abs/2101.12671).

- [2] D. J. Aldous, *An introduction to covering problems for random walks on graphs*, J. Theoret. Probab. **2** (1989), no. 1, 87–89.
- [3] B. Bárány, M. Hochman, and A. Rapaport, *Hausdorff dimension of planar self-affine sets and measures*, Invent. Math. **216** (2019), no. 3, 601–659.
- [4] M. Barnsley, *Fractals everywhere*, Academic Press Professional, Inc., USA, 1988.
- [5] M. F. Barnsley, K. Leśniak, and M. Rypka, *Chaos game for IFSs on topological spaces*, J. Math. Anal. Appl. **435** (2016), no. 2, 1458–1466.
- [6] M. F. Barnsley and K. Leśniak, *The chaos game on a general iterated function system from a topological point of view*, Internat. J. Bifur. Chaos Appl. Sci. Engrg. **24** (2014), no. 11, 1450139.
- [7] M. F. Barnsley and A. Vince, *The chaos game on a general iterated function system*, Ergodic Theory Dynam. Systems **31** (2011), no. 4, 1073–1079.
- [8] T. Bedford, *Crinkly curves, markov partitions and box dimensions in self-similar sets*, Ph.D. Thesis, 1984.
- [9] R. Bowen, *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, revised, Lecture Notes in Mathematics, vol. 470, Springer-Verlag, Berlin, 2008. With a preface by David Ruelle, Edited by Jean-René Chazottes.
- [10] J. H. Elton, *An ergodic theorem for iterated maps*, Ergodic Theory Dynam. Systems **7** (1987), no. 4, 481–488.
- [11] K. J. Falconer, *The Hausdorff dimension of self-affine fractals*, Math. Proc. Cambridge Philos. Soc. **103** (1988), no. 2, 339–350.
- [12] ———, *The Hausdorff dimension of self-affine fractals*, Mathematical Proceedings of the Cambridge Philosophical Society **103** (1988), no. 2, 339–350.
- [13] ———, *Techniques in Fractal Geometry*, John Wiley, England, 1997.
- [14] K. J. Falconer, J. M. Fraser, and A. Käenmäki, *Minkowski dimension for measures*, arXiv e-prints (2020), arXiv:2001.07055, available at [2001.07055](https://arxiv.org/abs/2001.07055).
- [15] A. H. Fan and K.-S. Lau, *Iterated Function System and Ruelle Operator*, J. Math. Anal. Appl. **231** (1999), 319–344.
- [16] B. Forte and F. Mendivil, *A classical ergodic property for IFS: a simple proof*, Ergodic Theory Dynam. Systems **18** (1998), no. 3, 609–611.
- [17] J. M. Fraser, *Assouad type dimensions and homogeneity of fractals*, Trans. Amer. Math. Soc. **366** (2014), no. 12, 6687–6733.
- [18] J. M. Fraser, *Assouad dimension and fractal geometry*, Cambridge Tracts in Mathematics, Cambridge University Press, 2020.
- [19] J. M. Fraser, *Fractal geometry of Bedford-McMullen carpets*, Thermodynamic formalism, 2021, pp. 495–516, M. Pollicott and S. Vaienti editors.
- [20] J. M. Fraser and D. Howroyd, *Assouad type dimensions for self-affine sponges*, Ann. Acad. Sci. Fenn. Math. **42** (2017), 149–174.
- [21] J. Gutiérrez, A. Iglesias, and M. A. Rodríguez, *A multifractal analysis of IFSP invariant measures with application to fractal image generation*, Fractals **04** (1996), no. 01, 17–27.
- [22] N. T. A. Haydn, *Entry and return times distribution*, Dyn. Syst. **28** (2013), no. 3, 333–353.
- [23] M. Hochman, *On self-similar sets with overlaps and inverse theorems for entropy in  $\mathbb{R}^d$* , ArXiv e-prints **1503.09043** (March 2015), available at [1503.09043](https://arxiv.org/abs/1503.09043).
- [24] M. Hochman, *On self-similar sets with overlaps and inverse theorems for entropy*, Annals of Mathematics **180** (2014), no. 2, 773–822.
- [25] T. Jordan, M. Pollicott, and K. Simon, *Hausdorff dimension for randomly perturbed self affine attractors*, Comm. Math. Phys. **270** (2007), no. 2, 519–544.
- [26] T. Jordan and M. Rams, *Multifractal analysis for Bedford–McMullen carpets*, Math. Proc. Cambridge Philos. Soc. **150** (2011), no. 1, 147–156.
- [27] N. Jurga and I. D. Morris, *How long is the Chaos Game?*, to appear in Bull. Lond. Math. Soc. (2021), arXiv:2007.11517, available at [2007.11517](https://arxiv.org/abs/2007.11517).
- [28] A. Käenmäki, *Measures of full dimension on self-affine sets*, Acta Univ. Carolin. Math. Phys. **45** (2004), no. 2, 45–53. 32nd Winter School on Abstract Analysis.
- [29] A. Käenmäki, J. Lehrback, and M. Vuorinen, *Dimensions, Whitney covers, and tubular neighborhoods*, Indiana Univ. Math. J. **62** (2013), 1861–1889.
- [30] A. Käenmäki and I. D. Morris, *Structure of equilibrium states on self-affine sets and strict monotonicity of affinity dimension*, Proc. Lond. Math. Soc. (3) **116** (2018), no. 4, 929–956. MR3789836

- [31] M. Kesseböhmer and S. Zhu, *Some recent developments in quantization of fractal measures*, Fractal Geometry and Stochastics V, 2015, pp. 105–120.
- [32] J. F. King, *The singularity spectrum for general Sierpiński carpets*, Adv. Math. **116** (1995), no. 1, 1–11.
- [33] K. Leśniak, *Random iteration for infinite nonexpansive iterated function systems*, Chaos **25** (2015), no. 8, 083117.
- [34] D. A. Levin and Y. Peres, *Markov chains and mixing times*, American Mathematical Society, Providence, RI, 2009.
- [35] V. Lucarini, D. Faranda, A. C. G. M. M. de Freitas, J. M. M. de Freitas, M. Holland, T. Kuna, M. Nicol, M. Todd, and S. Vaienti, *Extremes and recurrence in dynamical systems*, Pure and Applied Mathematics (Hoboken), John Wiley & Sons, Inc., Hoboken, NJ, 2016.
- [36] P. Matthews, *Covering problems for brownian motion on spheres*, Ann. Probab. **16** (1988), no. 1, 189–199.
- [37] C. McMullen, *The Hausdorff dimension of general Sierpiński carpets*, Nagoya Math. J. **96** (1984), 1–9.
- [38] M. D. Penrose, *Random Euclidean coverage from within*, arXiv e-prints (2021), arXiv:2101.06306, available at [2101.06306](https://arxiv.org/abs/2101.06306).
- [39] Y. Peres, W. Schlag, and B. Solomyak, *Sixty Years of Bernoulli Convolutions*, Fractal Geometry and Stochastics II, 2000, pp. 39–65.
- [40] M. Piraino, *The weak bernoulli property for matrix gibbs states*, Ergodic Theory Dynam. Systems **40** (2020), no. 8, 2219–2238.
- [41] P. Shmerkin, *On Furstenberg’s intersection conjecture, self-similar measures, and the  $L_q$  norms of convolutions*, Annals of Mathematics **189** (2019), no. 2, 319–391.
- [42] K. Simon, B. Solomyak, and M. Urbański, *Hausdorff dimension of limit sets for parabolic IFS with overlaps*, Pacific J. Math. **201** (2001), no. 2, 441–478.
- [43] P. Varjú, *Absolute continuity of Bernoulli convolutions for algebraic parameters*, J. Amer. Math. Soc. **32** (2019), 351–397.
- [44] I. Werner, *Ergodic theorem for contractive Markov systems*, Nonlinearity **17** (2004), no. 6, 2303–2313.
- [45] F. Xie, Y. Yin, and Y. Sun, *Uniform perfectness of self-affine sets*, Proc. Amer. Math. Soc. **131** (2003), no. 10, 3053–3057.

BALÁZS BÁRÁNY; BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS, MTA-BME STOCHASTICS RESEARCH GROUP, P.O. BOX 91, 1521 BUDAPEST, HUNGARY

*Email address:* `balubs@math.bme.hu`

NATALIA JURGA; UNIVERSITY OF ST ANDREWS, SCHOOL OF MATHEMATICS AND STATISTICS, ST ANDREWS, KY16 9SS, SCOTLAND

*Email address:* `naj1@st-andrews.ac.uk`

ISTVÁN KOLOSSVÁRY; UNIVERSITY OF ST ANDREWS, SCHOOL OF MATHEMATICS AND STATISTICS, ST ANDREWS, KY16 9SS, SCOTLAND

*Email address:* `itk1@st-andrews.ac.uk`