An analogue of Khintchine's theorem for self-conformal sets

Simon Baker

Mathematics institute,
University of Warwick,
Coventry, CV4 7AL, UK.
Email: simonbaker412@gmail.com

May 24, 2018

Abstract

Khintchine's theorem is a classical result from metric number theory which relates the Lebesgue measure of certain limsup sets with the convergence/divergence of naturally occurring volume sums. In this paper we ask whether an analogous result holds for iterated function systems (IFS's). We say that an IFS is approximation regular if we observe Khintchine type behaviour, i.e., if the size of certain limsup sets defined using the IFS is determined by the convergence/divergence of naturally occurring sums. We prove that an IFS is approximation regular if it consists of conformal mappings and satisfies the open set condition. The divergence condition we introduce incorporates the inhomogeneity present within the IFS. We demonstrate via an example that such an approach is essential. We also formulate an analogue of the Duffin-Schaeffer conjecture and show that it holds for a set of full Hausdorff dimension.

Combining our results with the mass transference principle of Beresnevich and Velani [4], we prove a general result that implies the existence of exceptional points within the attractor of our IFS. These points are exceptional in the sense that they are "very well approximated". As a corollary of this result, we obtain a general solution to a problem of Mahler, and prove that there are badly approximable numbers that are very well approximated by quadratic irrationals.

The ideas put forward in this paper are introduced in the general setting of iterated function systems that may contain overlaps. We believe that by viewing iterated function systems from the perspective of metric number theory, one can gain a greater insight into the extent to which they overlap. The results of this paper should be interpreted as a first step in this investigation.[1]

Mathematics Subject Classification 2010: 11K60, 28A80, 37C45.

Key words and phrases: Self-conformal sets, overlapping fractal attractors, Khintchine's theorem.

1 Introduction

Let $V \subset \mathbb{R}^d$ be a closed set. A map $\phi : V \to V$ is called a contraction if there exists $r \in (0,1)$ such that $|\phi(x) - \phi(y)| \leq r|x - y|$ for all $x, y \in V$. An *iterated function system* (IFS) on V is a finite set of contractions $\Phi = \{\phi_i\}_{i \in \mathcal{D}}$. A well known result due to Hutchinson [12] states that for any iterated function system Φ, there exists a unique non-empty compact set $X \subset \mathbb{R}^d$ such that

$$X = \bigcup_{i \in \mathcal{D}} \phi_i(X). \tag{1.1}$$

The set X is called the *attractor* associated to Φ. Often one is interested in understanding the geometric properties of X. When the images of X under the elements of Φ are well separated, then the geometry of X is well understood. Moreover, under this assumption one can often calculate the dimension of X for the various notions of dimension. However, when the images of X overlap significantly, the geometry of X is much more complicated. One cannot draw as many conclusions as in the non-overlapping case.

[1]This research was supported by EPSRC grant EP/M001903/1.

This being said, one of the guiding principles within fractal geometry is that if the IFS generating X has no obvious mechanism preventing X from satisfying a certain property, then that property should be satisfied. Properties we might be interested in include whether the dimension of X satisfies a certain formula, whether the dimension of certain measures supported on X satisfy a certain formula, whether certain measures supported on X are absolutely continuous with respect to the Lebesgue measure, etc. Consequently, it is believed that despite the presence of significant overlaps within an IFS there should be much that we can say. For more on IFS's with overlaps we refer the reader to [8, 10, 18] and the references therein. In this paper we view IFS's from the perspective of metric number theory. We now take the opportunity to recall the relevant background from this area.

Given $\Psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$, we associate the set

$$J(\Psi) := \Big\{x \in \mathbb{R} : |x - p/q| \leq \Psi(q) \text{ for i.m. } (p,q) \in \mathbb{Z} \times \mathbb{N}\Big\}.$$

Here and throughout i.m. will be used as a shorthand for infinitely many. The following well known theorem is due to Khintchine [15].

Theorem 1.1. *(Khintchine's theorem) If $\Psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ is a decreasing function and*

$$\sum_{q=1}^{\infty} q\Psi(q) = \infty,$$

then Lebesgue almost every x is an element of $J(\Psi)$.

This result is complemented by the following straightforward consequence of the Borel-Cantelli lemma.

Theorem 1.2. *If $\sum_{q=1}^{\infty} q\Psi(q) < \infty$, then $J(\Psi)$ has zero Lebesgue measure.*

Note that by an example of Duffin and Schaeffer [6] the monotonicity assumption appearing in Theorem 1.1 cannot be removed entirely. In response to this example they proposed the following refinement of Theorem 1.1, now known as the Duffin-Schaeffer conjecture.

Conjecture 1.3. (Duffin-Schaeffer conjecture) If $\Psi : \mathbb{N} \to \mathbb{R}_{\geq 0}$ is an arbitrary function satisfying

$$\sum_{q=1}^{\infty} \#\{1 \leq p \leq q : \gcd(p,q) = 1\}\Psi(q) = \infty,$$

then Lebesgue almost every x is an element of $J(\Psi)$.

Our IFS analogue of the set $J(\Psi)$ is defined as follows. Given an IFS Φ let $\mathcal{D}^n := \{(i_1, \ldots, i_n) : i_j \in \mathcal{D} \text{ for } 1 \leq j \leq n\}$ and $\mathcal{D}^* := \cup_{n=1}^{\infty} \mathcal{D}^n$. Fix $z \in X$ and let $\Psi : \mathcal{D}^* \to \mathbb{R}_{\geq 0}$. The analogue of $J(\Psi)$ is

$$W(\Psi, z) := \Big\{x \in X : x \in B(\phi_I(z), \Psi(I)) \text{ for i.m. } I \in \mathcal{D}^*\Big\}.$$

Here and throughout we will use $I = (i_1, \ldots, i_n)$ to denote an element of $\mathcal{D}^*$, and ϕ_I will denote the concatenation $\phi_{i_1} \circ \cdots \circ \phi_{i_n}$. We will denote the length of a finite word I by $|I|$ and let X_I denote $\phi_I(X)$. Throughout this paper we will refer to X_I as a *cylinder set* and say that it has *rank* $|I|$. We call any function $\Psi : \mathcal{D}^* \to \mathbb{R}_{\geq 0}$ an *approximating function.*

Note that the set $W(\Psi, z)$ has some key differences when compared with the set $J(\Psi)$. First of all, we have introduced the additional variable z. Secondly, the function Ψ may depend on the specific digits of I not just the length of I. As we will see in Example 2.1, allowing Ψ to depend on the specific digits of I and not just the length of I is essential.

Theorem 1.1 quantifies the good distributional properties the rational numbers have within $\mathbb{R}$. Similarly, if an analogue of Khintchine's theorem were to hold for $W(\Psi, z)$, this would reflect the good distributional properties the images of z have within X. Therefore, if an analogue of Khintchine's theorem were to hold for every $z \in X$, this would tell us that from the perspective of metric number theory our IFS does not contain significant overlaps. One could then ask whether the presence of a Khintchine type theorem would imply any other nice properties for our IFS. These ideas are the motivation behind this paper.

With the set $W(\Psi, z)$ defined as above for a general IFS, it isn't obvious what the appropriate analogue of Khintchine's theorem should be. In this paper we narrow our scope to determining an analogue of Khintchine's theorem for IFS's consisting of conformal mappings. Restricting to IFS's consisting of conformal mappings we introduce the notion of an approximation regular IFS. Put simply, an IFS will be approximation regular if we observe Khintchine type behaviour, i.e., if the size of $W(\Psi, z)$ is determined by the convergence/divergence of naturally occurring sums. We prove that an IFS consisting of conformal mappings is approximation regular if it satisfies the open set condition. These results are of independent interest but should also be interpreted as a first step in our investigation into studying overlapping attractors from the perspective of metric number theory. In the final section of this paper we prove a complementary result which states that whenever our IFS contains an exact overlap then it cannot be approximation regular. This is relevant as the standard mechanism by which one can construct an IFS that fails to satisfy a certain property, that we would otherwise expect to be true, is to construct the IFS in such a way that it contains an exact overlap.

Notation. Throughout this paper we make use of the standard big $\mathcal{O}$ notation. Given two positive real valued functions f, g defined on some set S, we write $f \asymp g$ if there exists a positive constant C such that $C^{-1}f(x) \leq g(x) \leq Cf(x)$ for all $x \in S$.

1.1 Self-similar sets and self-conformal sets

In this section we describe the attractors we will be focusing on. Suppose $\phi : \mathbb{R}^d \to \mathbb{R}^d$ is a contraction and that it also satisfies the condition $|\phi(x) - \phi(y)| = r|x - y|$ for all $x, y \in \mathbb{R}^d$, for some $r \in (0, 1)$. If ϕ satisfies this condition we call ϕ a *similarity*. When our IFS Φ consists solely of similarities, we say that X is a *self-similar set*. The unique s for which

$$\sum_{i \in \mathcal{D}} r_i^s = 1 \tag{1.2}$$

is called the *similarity dimension*. We will denote the similarity dimension by $\dim_S(X)$. This choice of notation is a little misleading as the similarity dimension is a function of Φ not X. There maybe several IFS's with different similarity dimensions but each with X as their attractor. However, it should always be clear from our context which IFS we are referring to.

Self-conformal sets are a natural generalisation of self-similar sets. Let $V \subset \mathbb{R}^d$ be an open set, a C^1 map $\phi : V \to \mathbb{R}^d$ is a *conformal mapping* if it preserves angles. Equivalently ϕ is a conformal mapping if the differential ϕ' satisfies $|\phi'(x)y| = |\phi'(x)||y|$ for all $x \in V$ and $y \in \mathbb{R}^d$. We call Φ a *conformal iterated function system* on a compact set $Y \subset \mathbb{R}^d$, if each ϕ_i can be extended to an injective conformal contraction on some open connected neighbourhood V that contains Y and $0 < \inf_{x \in V} |\phi_i'(x)| \leq \sup_{x \in V} |\phi_i'(x)| < 1$. Throughout this paper we will also assume that the differentials are Hölder continuous, i.e., there exists $\alpha > 0$ and $c > 0$ such that

$$||\phi_i'(x)| - |\phi_i'(y)|| \leq c|x - y|^\alpha$$

for all $x, y \in V$. When Φ is a conformal iterated function system we call the attractor X a *self-conformal set*. The generalisation of the similarity dimension within the setting of self-conformal sets is the unique zero of Bowen's equation, that is the unique $s \in \mathbb{R}$ that satisfies $P(s) = 0$, where

$$P(s) := \lim_{n \to \infty} \frac{1}{n} \log \inf_{x \in X} \sum_{I \in \mathcal{D}^n} |\phi_I'(x)|^s = \lim_{n \to \infty} \frac{1}{n} \log \sup_{x \in X} \sum_{I \in \mathcal{D}^n} |\phi_I'(x)|^s. \tag{1.3}$$

When each element of Φ is a similarity then the equation $P(s) = 0$ reduces to (1.2). For notational convenience we will denote the unique s satisfying $P(s) = 0$ by $\dim_S(X)$ and also call it the similarity dimension.

When the images of X under the elements of Φ are well separated, the Hausdorff dimension of X is often equal to $\dim_S(X)$. Indeed when there exists an open set $O \subset \mathbb{R}^d$ such that $\phi_i(O) \subset O$ for all $i \in \mathcal{D}$, and $\phi_i(O) \cap \phi_j(O) = \emptyset$ for all $i \neq j$, then $\dim_H(X) = \dim_S(X)$ and X has positive and finite $\dim_S(X)$-dimensional Hausdorff measure. If there exists an open set O satisfying the above criteria then the IFS Φ is said to satisfy the *open set condition*. These results are well known and date back to the work of Ruelle [19]. For a proof see [7]. Note that without any separation assumptions we still have the upper bound $\dim_H(X) \leq \dim_S(X)$.

1.2 Statement of results

We now introduce the notion of an approximation regular pair and an approximation regular IFS. When defining an approximation regular pair we have to be careful whether the attractor X has zero or positive $\dim_H(X)$-dimensional Hausdorff measure. Consequently the following definition is split into two parts.

- Let Φ be a conformal iterated function system and suppose $\mathcal{H}^{\dim_H(X)}(X) > 0$. Given $z \in X$, we call (Φ, z) an *approximation regular pair* if whenever $\theta : \mathbb{N} \to \mathbb{R}_{\geq 0}$ is a decreasing function such that

$$\sum_{n=1}^{\infty} \sum_{I \in \mathcal{D}^n} (Diam(X_I)^{\dim_S(X)/\dim_H(X)} \theta(n))^{\dim_H(X)} = \infty, \tag{1.4}$$

  then $\mathcal{H}^{\dim_H(X)}$ almost every $x \in X$ is an element of $W(Diam(X_I)^{\dim_S(X)/\dim_H(X)} \theta(|I|), z)$.

- Let Φ be a conformal iterated function system and suppose $\mathcal{H}^{\dim_H(X)}(X) = 0$. Given $z \in X$, we call (Φ, z) an *approximation regular pair* if whenever $\theta : \mathbb{N} \to \mathbb{R}_{\geq 0}$ is a decreasing function such that (1.4) holds, then $\dim_H(W(Diam(X_I)^{\dim_S(X)/\dim_H(X)} \theta(|I|), z)) = \dim_H(X)$.

We call Φ an *approximation regular iterated function system* if (Φ, z) is an approximation regular pair for every $z \in X$.

We emphasise that in the definition of an approximation regular pair the attractor may contain considerable overlaps. So we could have $\dim_H(X) < \dim_S(X)$. It is also worth noting that originally our sets $W(\Psi, z)$ were defined for any function $\Psi : \mathcal{D}^* \to \mathbb{R}_{\geq 0}$. In the definition of an approximation regular pair we have restricted to functions of the form $\Psi(I) = Diam(X_I)^{\dim_S(X)/\dim_H(X)} \theta(|I|)$. It is a consequence of this restriction that if θ is such that (1.4) holds, then θ cannot decay to zero exponentially fast. We have done this because the function θ depends only on the length of I, and if θ were to contribute some exponential decay the resulting function would not necessarily take into account the inhomogeneity present within the IFS. Restricting to approximating functions that reflect the inhomogeneity of the IFS is essential. The importance of picking appropriate approximating functions is demonstrated in Section 2.

The approximation regularity of a class of overlapping attractors was studied by the author in [1] and [2]. These attractors are intimately related to the well known Bernoulli convolutions. In [1] the author used properties of the Bernoulli convolution to prove approximation regularity results.

Note that when Φ satisfies the open set condition then (1.4) reduces to

$$\sum_{n=1}^{\infty} \sum_{I \in \mathcal{D}^n} (Diam(X_I) \theta(n))^{\dim_H(X)} = \infty. \tag{1.5}$$

Moreover $\mathcal{H}^{\dim_H(X)}(X)$ will always be positive and finite, so we take the former definition of an approximation regular pair. Under this assumption the relevant limsup sets that appear in the definition of approximation regularity are of the form $W(Diam(X_I)\theta(|I|), z)$. For notational convenience we let $W(\theta, z)$ denote $W(Diam(X_I)\theta(|I|), z)$ throughout.

Our main result is the following.

Theorem 1.4. *Let Φ be an iterated function system with attractor X.*

1. *If $\Psi : \mathcal{D}^* \to \mathbb{R}_{\geq 0}$ is such that*

$$\sum_{n=1}^{\infty} \sum_{I \in \mathcal{D}^n} \Psi(I)^{\dim_H(X)} < \infty, \tag{1.6}$$

   then $\mathcal{H}^{\dim_H(X)}(W(\Psi, z)) = 0$ for all $z \in X$.

2. *If Φ is a conformal iterated function system and satisfies the open set condition, then Φ is an approximation regular IFS.*

In our introduction we mentioned an example of Duffin and Schaeffer [6] which demonstrates that one cannot remove all monotonicity assumptions from the statement of Khintchine's theorem. It is natural to ask whether we can weaken the monotonicity assumptions on θ appearing in the definition of an approximation regular pair. With this question in mind we make the following conjecture.

Conjecture 1.5. Theorem 1.4 holds with no underlying monotonicity assumptions on the class of θ.

In [17] Levesley, Salp, and Velani studied the approximation properties of balls centred at the left endpoint of the basic intervals generating the middle third cantor set. This fits into our setup with $\Phi = \{x/3, x/3+2/3\}$ and $z=0$. It is a consequence of their work that divergence in (1.5) is sufficient to prove that $\mathcal{H}^{\dim_H(X)}$ almost every $x\in X$ is contained in $W(\theta,0)$ with no monotonicity assumptions on the function θ. A key step in [17] was replacing the set $\{\phi_I(0)\}_{I\in\mathcal{D}^n}$ with its subset C_n which consists of those elements of $\{\phi_I(0)\}_{I\in\mathcal{D}^n}$ with coprime numerator and denominator. Importantly C_n has cardinality of the order $c\cdot 2^n$ and coprimeness guarantees good separation properties between the set C_n and C_m for $n\neq m$. These separation properties were important in their proof. In our setup it is not obvious what an appropriate analogue of C_n should be and if it even exists. Consequently we cannot prove Conjecture 1.5. We can however prove the following result which doesn't require θ to be decreasing.

Theorem 1.6. *Let $z\in X$ and Φ be a conformal iterated function system satisfying the open set condition. Suppose $\theta:\mathbb{N}\to\mathbb{R}_{\geq 0}$ is a function satisfying* (1.5) *and*

$$\sum_{n=1}^{Q}\theta(n)^{\dim_H(X)}\log\frac{1}{\theta(n)} = \mathcal{O}\Big(\Big(\sum_{n=1}^{Q}\theta(n)^{\dim_H(X)}\Big)^2\Big). \tag{1.7}$$

Then $\mathcal{H}^{\dim_H(X)}$ almost every $x\in X$ is an element of $W(\theta,z)$.

As an application of Theorem 1.6 we have the following corollary.

Corollary 1.7. *Let $z\in X$ and Φ be a conformal iterated function system satisfying the open set condition. Suppose $\theta:\mathbb{N}\to\mathbb{R}_{\geq 0}$ satisfies $\theta(n)\asymp n^{-1/\dim_H(X)}$. Then $\mathcal{H}^{\dim_H(X)}$ almost every $x\in X$ is an element of $W(\theta,z)$.*

Proof. We omit the details why for this choice of θ we have divergence in (1.5). Instead we give a brief argument explaining why (1.7) holds. For any θ satisfying our assumptions we have

$$\sum_{n=1}^{Q}\theta(n)^{\dim_H(X)}\log\frac{1}{\theta(n)} \asymp \sum_{n=1}^{Q}\frac{\log n}{n}\asymp (\log Q)^2.$$

Similarly, we have

$$\Big(\sum_{n=1}^{Q}\theta(n)^{\dim_H(X)}\Big)^2 \asymp \Big(\sum_{n=1}^{Q}\frac{1}{n}\Big)^2\asymp (\log Q)^2.$$

The final step in both of these equations can be seen to hold by approximating the summation with an integral. Combining these two equations we have (1.7). By Theorem 1.6 we may conclude our result. □

In the statement of Theorem 1.6 we introduced a new condition to replace the decreasing condition appearing in the definition of an approximation regular pair. Similarly, one can introduce a condition on the element z which allows one to prove that $\mathcal{H}^{\dim_H(X)}$ almost every $x\in X$ is an element of $W(\theta,z)$ with no monotonicity assumptions on θ. We postpone the statement of this condition until Section 4. We will show in Section 4 that the set of points with this property has full Hausdorff dimension within X. As a consequence of these results we have the following theorem.

Theorem 1.8. *If Φ is a conformal iterated function system satisfying the open set condition, then there exists $Y\subset X$ satisfying $\dim_H(Y)=\dim_H(X)$, such that if $z\in Y$ and $\theta:\mathbb{N}\to\mathbb{R}_{\geq 0}$ satisfies* (1.5)*, then $\mathcal{H}^{\dim_H(X)}$ almost every $x\in X$ is an element of $W(\theta,z)$.*

Remark 1.9. The well informed reader might rightly ask whether the general framework introduced by Beresnevich, Dickinson, and Velani [3] can be applied to give a proof of Theorem 1.4. This general approach relies on the introduction of an appropriate weight function which satisfies certain properties. For self-similar sets with a uniform contraction ratio such a weight function can be shown to exist. However, for more general self-similar sets and self-conformal sets it is not clear whether such a function exists. Thus we do not apply their techniques. In any case, our proof of Theorem 1.4 is the starting point for the proofs of Theorem 1.6 and Theorem 1.8. Both of these theorems do not follow from the work done in [3].

Remark 1.10. Under the assumptions of statement 2 from Theorem 1.4, it can be shown that $Diam(X_I)^{\dim_H(X)} \asymp \mu(X_I)$ for all $I \in \mathcal{D}^*$, where μ is a finite measure supported on X. This implies that

$$\sum_{I\in\mathcal{D}^n} Diam(X_I)^{\dim_H(X)} \asymp 1$$

for all $n \in \mathbb{N}$. Thus condition (1.5) is equivalent to

$$\sum_{n=1}^{\infty} \theta(n)^{\dim_H(X)} = \infty.$$

Remark 1.11. When X is a self-similar set where each similarity has the same contraction ratio r, then (1.5) can be rewritten as

$$\sum_{n=1}^{\infty} \#\mathcal{D}^n (r^n\theta(n))^{\dim_H(X)} = \infty. \tag{1.8}$$

The rephrased divergence condition stated in (1.8) is the same condition as that which appears in [9] and [17].

The rest of this paper is arranged as follows. In Section 2 we include some examples which demonstrate that restricting to approximating functions of the form appearing in the definition of an approximation regular pair is essential if one wants to expect the divergence of naturally occurring sums to provide any information about the size of $W(\Psi, z)$. In Section 3 we prove Theorem 1.4. We then prove Theorem 1.6 and Theorem 1.8 in Section 4. In Section 5 we combine Theorem 1.4 with the mass transference principle of Beresnevich and Velani [4]. The mass transference principle will allow us to determine the Hausdorff dimension of the set $W(\Psi, z)$ for a certain class of Ψ when we have convergence in (1.5). We apply this result to obtain a general criteria that allows one to deduce that X contains exceptional elements, see Proposition 5.5. Exceptional here means well approximated in a way that maybe defined independently from X. As an application of this result, we obtain a general solution to a problem of Mahler, and prove that there are badly approximable numbers that are very well approximated by quadratic irrationals. In Section 6 we prove that if a conformal IFS contains an exact overlap then there are no approximation regular pairs. We also discuss the overlapping case and suggest some future directions.

2 Examples

In this section we include two examples which demonstrate that in the definition of an approximation regular IFS it is necessary and perhaps natural to restrict to approximating functions of the form $\Psi(I) = Diam(X_I)^{\dim_S(X)/\dim_H(X)}\theta(|I|)$. The first example demonstrates that any inhomogeneity, i.e. different rates of contraction, that maybe present within our IFS should be taken into consideration.

Example 2.1. Let $\Phi = \{\phi_1, \phi_2\}$ where $\phi_1(x) = \frac{3x}{4}$ and $\phi_2(x) = \frac{x}{4} + \frac{3}{4}$. This iterated function system satisfies the open set condition and the corresponding attractor is the unit interval $[0, 1]$. For simplicity we take $z = 0$. The following argument can easily be adapted to an arbitrary $z \in [0, 1]$.

Note that the image $\phi_I(0)$ is always the left endpoint of the interval $\phi_I([0, 1])$. Also note that the set of intervals $\{\phi_I([0, 1])\}_{I\in\{1,2\}^n}$ always cover $[0, 1]$, and if two of these intervals have a non-empty intersection then they intersect at a mutual endpoint.

Consider the limsup set

$$W := \Big\{x \in [0, 1] : x \in B(\phi_I(0), 2^{-|I|}) \text{ for i.m. } I \in \{1, 2\}^*\Big\}.$$

Note that in the definition of W the radii of the defining balls only depends upon the length of the word I. There are 2^n elements of $\{1, 2\}^n$, therefore $\sum_{n=1}^{\infty}\sum_{I\in\{1,2\}^n} 2^{-n} = \infty$. Consequently, if an analogue of Khintchine's theorem held for IFS's where we didn't need to take into account the inhomogeneity of Φ, we would expect W to have full Lebesgue measure. We now show that this is not the case.

Given $x \in [0, 1]$, we say that $(i_n)_{n=1}^{\infty} \in \{1, 2\}^{\infty}$ is a coding for x if

$$x = \bigcap_{n=1}^{\infty} X_{i_1,\ldots,i_n}.$$

Every $x \in [0,1]$ has a coding. This coding is unique for every $x \in X$ apart from a countable set of x with precisely two codings. We let $\pi : \{1,2\}^{\mathbb{N}} \to [0,1]$ be the function which maps (i_n) to $\cap_{n=1}^{\infty} X_{i_1,\ldots,i_n}$. Let $\mathcal{P}$ be the Bernoulli measure on $\{1,2\}^{\mathbb{N}}$ which gives the digit 1 mass 3/4 and the digit 2 mass 1/4. The push forward of the measure $\mathcal{P}$ under the map π is precisely the Lebesgue measure restricted to $[0,1]$. Applying the strong law of large numbers, we may conclude that for Lebesgue almost every $x \in [0,1]$ its coding (i_n) satisfies

$$\lim_{m\to\infty} \frac{\#\{1 \le n \le m : i_n = 1\}}{m} \to \frac{3}{4} \text{ and } \lim_{m\to\infty} \frac{\#\{1 \le n \le m : i_n = 2\}}{m} \to \frac{1}{4}.$$

By the above, for any $\epsilon > 0$, we may pick a large $N \in \mathbb{N}$ such that the set

$$A := \Big\{x \in [0,1] : \frac{\#\{1 \le n \le m : i_n = 1\}}{m} \ge \frac{5}{8} \text{ and } \frac{\#\{1 \le n \le m : i_n = 2\}}{m} \le \frac{3}{8} \text{ for all } m \ge N\Big\}$$

has Lebesgue measure at least $1 - \epsilon$.

We now obtain a bound for the cardinality of the set

$$\Sigma_m := \Big\{(i_n)_{n=1}^m \in \{1,2\}^m : \frac{\#\{1 \le n \le m : i_n = 1\}}{m} \ge \frac{5}{8}\Big\}.$$

This bound will rely on a result from probability theory. Suppose we have a sequence of independent random variables $X_1, X_2 \ldots, X_m$. Let

$$\overline{X} = \frac{1}{m}\sum_{n=1}^{m} X_n \text{ and } \mu = \frac{1}{m}\sum_{n=1}^{m} E(X_n).$$

The following bound is known as Hoeffding's inequality [11].

Lemma 2.2. *Suppose* $0 \le X_n \le 1$ *for all* $1 \le n \le m$*, then for* $0 < t < 1 - \mu$

$$Prob(\overline{X} - \mu \ge t) \le e^{-2mt^2}.$$

With Lemma 2.2 in mind we let $\mathcal{P}'$ be the unbiased probability measure that gives digit 1 mass 1/2 and digit 2 mass 1/2. Then

$$\#\Sigma_m = 2^m \cdot \mathcal{P}'\Big((i_n)_{n=1}^m : \frac{\#\{1 \le n \le m : i_n = 1\}}{m} \ge \frac{5}{8}\Big).$$

Applying Hoeffding's inequality we obtain

$$\#\Sigma_m = 2^m \cdot \mathcal{P}'\Big((i_n)_{n=1}^m : \frac{\#\{1 \le n \le m : i_n = 1\}}{m} - \frac{1}{2} \ge \frac{1}{8}\Big) \le \Big(\frac{2}{e^{2/64}}\Big)^m. \tag{2.1}$$

Equation (2.1) is the desired upper bound on the cardinality of Σ_m.

Returning to the interval $[0,1]$, we remark that if x is contained in $B(\phi_I(0), 2^{-|I|})$ for some $I \in \mathcal{D}^*$, then x must be contained in either $B(a_{i_1,\ldots,i_{|I|}}, 2^{-|I|})$ or $B(b_{i_1,\ldots,i_{|I|}}, 2^{-|I|})$, where (i_n) is the coding for x and $\phi_{i_1,\ldots,i_{|I|}}([0,1]) = [a_{i_1,\ldots,i_{|I|}}, b_{i_1,\ldots,i_{|I|}}]$. This is a consequence of how the intervals $\{\phi_I([0,1])\}_{I\in\{1,2\}^m}$ sit alongside one another in $[0,1]$, and because x is always contained in the interval $\phi_{i_1,\ldots,i_n}([0,1])$.

We now use the preceding observation to obtain estimates on $\mathcal{L}(A \cap W)$. Here and throughout $\mathcal{L}$ denotes the Lebesgue measure. For any $N \in \mathbb{N}$ we have

$$\begin{aligned}
\mathcal{L}(A \cap W) &\le \mathcal{L}\Big(\bigcup_{m=N}^{\infty} \bigcup_{I\in\Sigma_m} (B(a_{i_1,\ldots,i_m}, 2^{-m}) \cup B(b_{i_1,\ldots,i_m}, 2^{-m}))\Big)\\
&\le \sum_{m=N}^{\infty} \sum_{I\in\Sigma_m} \mathcal{L}(B(a_{i_1,\ldots,i_m}, 2^{-m})) + \mathcal{L}(B(b_{i_1,\ldots,i_m}, 2^{-m}))\\
&= \sum_{m=N}^{\infty} \#\Sigma_m \cdot 4 \cdot 2^{-m}\\
&\le \sum_{m=N}^{\infty} 4 \cdot \Big(\frac{1}{e^{2/64}}\Big)^m
\end{aligned}$$

$$< \infty.$$

In the penultimate inequality we used (2.1). Since we have convergence above, we can choose $N \in \mathbb{N}$ such that $\sum_{m=N}^{\infty} 4 \cdot e^{-2m/64}$ is arbitrarily small. Therefore $\mathcal{L}(A \cap W) = 0$. Since $\mathcal{L}(A) > 1 - \epsilon$ we have $\mathcal{L}(W) < \epsilon$. Since ϵ is arbitrary we may conclude that $\mathcal{L}(W) = 0$.

Our second example demonstrates that for a reasonable analogue of Khintchine's theorem to hold for IFS's, it is necessary that an approximation function gives weight to all words and is not concentrated on a subset of $\mathcal{D}^*$.

Example 2.3. Let $\Phi = \{\phi_1, \phi_2\}$ where $\phi_1(x) = \frac{x}{2}$ and $\phi_2(x) = \frac{x}{2} + \frac{1}{2}$. As in Example 2.1 our attractor is the interval $[0,1]$ and Φ satisfies the open set condition. We fix $z = 0$ and a word $J = (j_1, \ldots, j_m) \in \{1,2\}^m$. Let

$$\Psi(I) = \begin{cases} 2^{-|I|} & \text{if } I \text{ doesn't begin with } J \\ 2^{-|I|}|I|^{-2} & \text{if } I \text{ begins with } J. \end{cases}$$

Then Ψ satisfies

$$\sum_{\substack{I \in \mathcal{D}^* \\ I \text{ begins with } J}} \Psi(I) < \infty \tag{2.2}$$

and

$$\sum_{n=1}^{\infty} \sum_{I \in \mathcal{D}^n} \Psi(I) = \infty. \tag{2.3}$$

If an analogue of Khintchine's theorem held for IFS's where Ψ did not have to distribute weight evenly amongst the elements of $\mathcal{D}^*$, then (2.3) would suggest $W(\Psi, 0)$ has full Lebesgue measure. However, using the Borel Cantelli lemma we can show that (2.2) implies

$$\mathcal{L}\Big(\Big[\sum_{i=1}^{m} \frac{j_i}{2^i}, \sum_{i=1}^{m} \frac{j_i}{2^i} + \frac{1}{2^i}\Big] \cap W(\Psi, 0)\Big) = 0.$$

Thus $W(\Psi, 0)$ does not have full measure despite (2.3) being satisfied.

Bearing Example 2.1 and Example 2.3 in mind, we believe that the class of approximating functions we restrict to in the definition of an approximation regular pair is not so restrictive, and is in fact a natural class of approximating functions to study.

3 Proof of Theorem 1.4

3.1 Preliminaries

We start this section by recalling the definition of Hausdorff measure and Hausdorff dimension. Let $E \subset \mathbb{R}^d$, $s \geq 0$, and $\rho > 0$. We let

$$\mathcal{H}^s_\rho(E) := \inf \Big\{ \sum_{n=1}^{\infty} Diam(U_n)^s : \{U_n\}_{n=1}^{\infty} \text{ is a } \rho\text{-cover for } E \Big\}.$$

In the above the infimum is taken over all ρ-covers of E. The limit $\lim_{\rho \to 0} \mathcal{H}^s_\rho(E) := \mathcal{H}^s(E)$ exists and we call this limit the *s-dimensional Hausdorff measure* of E. It is a straightforward exercise to show that for any $E \subset \mathbb{R}^d$ the following equality holds

$$\inf\{s : \mathcal{H}^s(E) = 0\} = \sup\{s : \mathcal{H}^s(E) = \infty\}.$$

We call this coinciding value the *Hausdorff dimension* of E and denote it by $\dim_H(E)$.

Suppose E has Hausdorff dimension s. We say that E is *Ahlfors regular* if

$$\mathcal{H}^s(E \cap B(x,r)) \asymp r^s, \tag{3.1}$$

for all $x \in E$ and $0 < r < Diam(E)$. Importantly, under the open set condition of statement 2 from Theorem 1.4, the attractor X will always be Ahlfors regular.

In our proofs we will require the notion of a coding. This is the natural generalisation of what appeared in Example 2.1. Given an IFS Φ and $x \in X$, then there exists a sequence $(i_n) \in \mathcal{D}^{\mathbb{N}}$ such that

$$x = \bigcap_{n=1}^{\infty} X_{i_1,\ldots,i_n}.$$

We call such a sequence a *coding of* x. The coding of x is not necessarily unique. An x may well have a continuum of codings. As a final observation, we remark that if (i_n) is a coding for x, then $(j_1, \ldots, j_m, i_1, i_2, \ldots)$ is a coding for $\phi_J(x)$ where $J = (j_1, \ldots, j_m)$.

3.2 Statement 1

The proof of statement 1 from Theorem 1.4 is standard but we include it for completeness.

Proof of statement 1. Let Φ be an IFS with attractor X and let z be an arbitrary element of X. Let $\Psi : \mathcal{D}^* \to \mathbb{R}_{\geq 0}$ be an approximating function satisfying

$$\sum_{n=1}^{\infty} \sum_{I \in \mathcal{D}^n} \Psi(I)^{\dim_H(X)} < \infty. \tag{3.2}$$

Fix $\rho > 0$. Let $M \in \mathbb{N}$ be sufficiently large that $2\Psi(I) < \rho$ for all $I \in \mathcal{D}^n$ with $n \geq M$. Such an M exists because of (3.2). Therefore

$$\{B(\phi_I(z), \Psi(I))\}_{\substack{I \in \mathcal{D}^n \\ n \geq M}}$$

is a ρ cover of $W(\Psi, z)$. For any $\epsilon > 0$, one can assume that M is also sufficiently large that

$$\sum_{n=M}^{\infty} \sum_{I \in \mathcal{D}^n} \Psi(I)^{\dim_H(X)} < \epsilon.$$

This is a consequence of (3.2). Therefore

$$\begin{aligned} \mathcal{H}_{\rho}^{\dim_H(X)}(W(\Psi, z)) &\leq \sum_{n=M}^{\infty} \sum_{I \in \mathcal{D}^n} (2\Psi(I))^{\dim_H(X)} \\ &\leq 2^{\dim_H(X)} \epsilon. \end{aligned}$$

Since ϵ is arbitrary we must have $\mathcal{H}_{\rho}^{\dim_H(X)}(W(\Psi, z)) = 0$. Since ρ was arbitrary we have $\mathcal{H}^{\dim_H(X)}(W(\Psi, z)) = 0$.

$\square$

3.3 Statement 2

The proof of statement 2 from Theorem 1.4 follows a similar framework to the proof of Theorem 2 from [17]. In particular we make use of the following two lemmas.

Lemma 3.1. *Let X be a compact set in $\mathbb{R}^d$ and let μ be a finite doubling measure on X such that any open set is μ measurable. Let E be a Borel subset of X. Assume that there are constants $r_0, c > 0$ such that for any ball B with radius less than r_0 and centre in X we have*

$$\mu(E \cap B) > c\mu(B).$$

Then $\mu(X \setminus E) = 0$.

For a proof of Lemma 3.1 see [3, §8]. Note that a measure μ supported on a compact set X is *doubling* if there exists a constant $C > 1$ such that for any $x \in X$ and $r > 0$ we have

$$\mu(B(x, 2r)) \leq C\mu(B(x, r)).$$

Clearly if X is Ahlfors regular then the restriction of $\mathcal{H}^{\dim_H(X)}$ to X is a doubling measure.

Lemma 3.2. *Let X be a compact set in $\mathbb{R}^d$ and let μ be a finite measure on X. Also, let E_n be a sequence of μ-measurable sets such that $\sum_{n=1}^{\infty}\mu(E_n)=\infty$. Then*

$$\mu(\limsup_{n\to\infty} E_n) \geq \limsup_{Q\to\infty} \frac{(\sum_{n=1}^{Q}\mu(E_n))^2}{\sum_{n,m=1}^{Q}\mu(E_n\cap E_m)}.$$

For a proof of Lemma 3.2 see [22, Lemma 5].

In the proofs of each of our theorems we will need the following properties of self-conformal sets satisfying the open set condition. Let $\mu := \mathcal{H}^{\dim_H(X)}|_X$ be the $\dim_H(X)$-dimensional Hausdorff measure restricted to X, then:

- For any $n\in\mathbb{N}$ and $I,J\in D^n$ such that $I\neq J$ we have
$$\mu(X_I\cap X_J)=0. \tag{3.3}$$
- For any $I,J\in\mathcal{D}^*$
$$\mu(X_{IJ})\asymp\mu(X_I)\mu(X_J). \tag{3.4}$$
- For any $I,J\in\mathcal{D}^*$
$$Diam(X_{IJ})\asymp Diam(X_I)Diam(X_J). \tag{3.5}$$
- For any $I\in\mathcal{D}^*$
$$\mu(X_I)\asymp Diam(X_I)^{\dim_H(X)}. \tag{3.6}$$
- There exists $\gamma\in(0,1)$ such that for any $I\in\mathcal{D}^*$
$$\mu(X_I)=\mathcal{O}(\gamma^{|I|}). \tag{3.7}$$
- Let $x\in X$ and $(i_n)\in\mathcal{D}^{\mathbb{N}}$ be a coding of x. For any $0<r<Diam(X)$ there exists $N\in\mathbb{N}$ such that
$$X_{i_1,\ldots,i_N}\subset B(x,r) \text{ and } Diam(X_{i_1,\ldots,i_N})\asymp r. \tag{3.8}$$

In the above we have denoted the concatenation of two words I and J by IJ. Property (3.3) follows from [13, Theorem 3.7]. For a proof of the remaining properties and for a proof that X is Ahlfors regular see [7] and [19]. Properties (3.4), (3.6), and (3.7) are essentially a consequence of the fact that μ is equivalent to a suitably defined Gibbs measure for a particular Hölder continuous potential. Property (3.5) is a consequence of the differentials being Hölder continuous. The proof of (3.8) is standard.

3.3.1 Proof of statement 2

We are now in a position to prove statement 2 from Theorem 1.4. We start by fixing $z\in X$ and let $(z_i)\in\mathcal{D}^{\mathbb{N}}$ be a coding of z. Let $\theta:\mathbb{N}\to\mathbb{R}$ be a decreasing function satisfying (1.5). We now fix a ball $B(y,r)$ where $y\in X$ and $0<r<Diam(X)$. We will show that

$$\mu(W(\theta,z)\cap B(y,r))>c\mu(B(y,r)), \tag{3.9}$$

for some constant $c>0$ that does not depend on our choice of ball. Applying Lemma 3.1 will then allow us to conclude our result.

The set $W(\theta,z)$ is defined to be a limsup set of a sequence of balls. It will be computationally easier to replace these balls with cylinder sets. For each $I\in\mathcal{D}^*$ we let $X_{I,\theta}$ be a cylinder set that satisfies the following two properties:

$$X_{I,\theta}\subseteq B(\phi_I(z),Diam(X_I)\theta(|I|)) \tag{3.10}$$
$$\mu(X_{I,\theta})\asymp(Diam(X_I)\theta(|I|))^{\dim_H(X)}. \tag{3.11}$$

Such a cylinder set exists by properties (3.6) and (3.8). Without loss of generality we may assume that

$$X_{I,\theta}=X_{I(z_1,\ldots,z_{n(I,\theta)})}$$

for some $n(I,\theta)\in\mathbb{N}$. We emphasise here that for any $I\in\mathcal{D}^n$ we have that I is a prefix of $I(z_1,\ldots,z_{n(I,\theta)})$, $|I(z_1,\ldots,z_{n(I,\theta)})| = n+n(I,\theta)$, and $X_{I,\theta}\subset X_I$. Importantly, by (3.3) we have $\mu(X_{I,\theta}\cap X_{J,\theta})=0$ for $I,J\in\mathcal{D}^n$ such that $I\neq J$. We now replace $W(\theta,z)$ with the following limsup set that is defined using cylinder sets instead of balls. Let

$$E(\theta,z) := \Big\{x\in X : x\in X_{I,\theta} \text{ for } i.m.\, I\in\mathcal{D}^*\Big\}$$

By (3.10) we have $E(\theta,z)\subseteq W(\theta,z)$. So to prove (3.9) it suffices to prove

$$\mu(E(\theta,z)\cap B(y,r)) > c\mu(B(y,r)). \tag{3.12}$$

To our ball $B(y,r)$ we associate the cylinder $X_{(y_1,\ldots,y_{n(r)})}$, where $(y_i)\in\mathcal{D}^{\mathbb{N}}$ is a coding for y and $X_{(y_1,\ldots,y_{n(r)})}$ satisfies (3.8). To each $n\geq n(r)$ we associate the sets

$$\mathcal{E}_n := \Big\{X_{I,\theta} : I\in\mathcal{D}^n \text{ and } (i_1,\ldots,i_{n(r)}) = (y_1,\ldots,y_{n(r)})\Big\}$$

and

$$E_n := \bigcup_{X_{I,\theta}\in\mathcal{E}_n} X_{I,\theta}.$$

Then

$$\limsup E_n \subset E(\theta,z)\cap X_{(y_1,\ldots,y_{n(r)})} \subset E(\theta,z)\cap B(y,r).$$

Therefore to prove (3.12) it suffices to show that

$$\mu(\limsup E_n) > c\mu(B(y,r)). \tag{3.13}$$

It is a consequence of (3.6), (3.8), and the fact that μ is Ahlfors regular that

$$\mu(X_{(y_1,\ldots,y_{n(r)})}) \asymp \mu(B(y,r)).$$

Therefore to prove (3.13) it suffices to prove

$$\mu(\limsup E_n) > c\mu(X_{(y_1,\ldots,y_{n(r)})}). \tag{3.14}$$

We now focus our attention on proving (3.14). To prove (3.14) we will apply Lemma 3.2 to the sequence of sets E_n. To successfully apply this lemma there are two steps, we must first demonstrate that $\sum\mu(E_n)=\infty$, we then obtain estimates for the measure of $E_n\cap E_m$ for $n\neq m$. We split these steps into the following two propositions.

Proposition 3.3. $\sum_{n=n(r)}^{\infty}\mu(E_n)=\infty$.

Proof. Recalling (1.5), our θ must satisfy

$$\sum_{n=n(r)}^{\infty}\sum_{I\in\mathcal{D}^n}(Diam(X_I)\theta(n))^{\dim_H(X)} = \infty. \tag{3.15}$$

Each $X_{I,\theta}$ satisfies $\mu(X_{I,\theta})\asymp (Diam(X_I)\theta(|I|))^{\dim_H(X)}$ by (3.11). Therefore

$$\sum_{n=n(r)}^{M}\sum_{I\in\mathcal{D}^n}(Diam(X_I)\theta(n))^{\dim_H(X)} \asymp \sum_{n=n(r)}^{M}\sum_{I\in\mathcal{D}^n}\mu(X_{I,\theta}) \tag{3.16}$$

for any $M\geq n(r)$. It is a consequence of (3.4) that for any $J\in\mathcal{D}^{n(r)}$ we have

$$\sum_{n=n(r)}^{M}\sum_{\substack{I\in\mathcal{D}^n\\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}}\mu(X_{I,\theta}) \asymp \sum_{n=n(r)}^{M}\sum_{\substack{I\in\mathcal{D}^n\\ (i_1,\ldots,i_{n(r)})=J}}\mu(X_{I,\theta}). \tag{3.17}$$

In (3.17) the implied constants may depend on $n(r)$. By (3.17) we see that

$$\sum_{n=n(r)}^{\infty}\sum_{\substack{I\in\mathcal{D}^n\\(i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}}\mu(X_{I,\theta})=\infty \text{ if and only if } \sum_{n=n(r)}^{\infty}\sum_{\substack{I\in\mathcal{D}^n\\(i_1,\ldots,i_{n(r)})=J}}\mu(X_{I,\theta})=\infty. \tag{3.18}$$

By (3.15) and (3.16) we must have

$$\sum_{n=n(r)}^{\infty}\sum_{\substack{I\in\mathcal{D}^n\\(i_1,\ldots,i_{n(r)})=J}}\mu(X_{I,\theta})=\infty$$

for at least one $J\in\mathcal{D}^{n(r)}$. Therefore by (3.18) we may conclude that

$$\sum_{n=n(r)}^{\infty}\sum_{\substack{I\in\mathcal{D}^n\\(i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}}\mu(X_{I,\theta})=\infty.$$

Since distinct elements of $\mathcal{E}_n$ intersect in a set of measure zero we have

$$\sum_{n=n(r)}^{\infty}\sum_{\substack{I\in\mathcal{D}^n\\(i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}}\mu(X_{I,\theta})=\sum_{n=n(r)}^{\infty}\mu(E_n).$$

Therefore $\sum_{n=n(r)}^{\infty}\mu(E_n)=\infty$ as required. □

Proposition 3.4. *Let $\theta:\mathbb{N}\to\mathbb{R}_{\geq 0}$ be decreasing. Then*

$$\sum_{n,m=n(r)}^{Q}\mu(E_n\cap E_m)=\mathcal{O}\Big(\mu(X_{(y_1,\ldots,y_{n(r)})})\Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}+\Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}\Big)^2\Big)\Big).$$

Proof. Let $I\in\mathcal{D}^n$ and $X_{I,\theta}\in\mathcal{E}_n$. When calculating $\mu(X_{I,\theta}\cap E_m)$ there are two cases that naturally arise, when $m>n+n(I,\theta)$ and when $n<m\leq n+n(I,\theta)$. Let us start with the case where $n<m\leq n+n(I,\theta)$. Since the rank of the cylinder $X_{I,\theta}$ is at least m, there exists at most one $X_{J,\theta}\in\mathcal{E}_m$ such that $X_{I,\theta}\cap X_{J,\theta}\neq\emptyset$. Moreover this J must be of the form $J=(i_1,\ldots,i_n,z_1,\ldots,z_{m-n})$. These observations give rise to the following useful bound on $\mu(X_{I,\theta}\cap E_m)$ for $n<m\leq n+n(I,\theta)$:

$$\begin{aligned}
\mu(X_{I,\theta}\cap E_m)&=\mu(X_{I,\theta}\cap X_{J,\theta})\\
&\leq\mu(X_{J,\theta})\\
&\asymp(Diam(X_J)\theta(m))^{\dim_H(X)} && \text{(By (3.11))}\\
&\asymp(Diam(X_I)Diam(X_{(z_1,\ldots,z_{m-n})})\theta(m))^{\dim_H(X)} && \text{(By (3.5))}\\
&\asymp\mu(X_I)\mu(X_{(z_1,\ldots,z_{m-n})})\theta(m)^{\dim_H(X)} && \text{(By (3.6))}\\
&\leq\mu(X_I)\mu(X_{(z_1,\ldots,z_{m-n})})\theta(n)^{\dim_H(X)} && \text{(Because } \theta \text{ is decreasing)}\\
&=\mathcal{O}(\mu(X_I)\theta(n)^{\dim_H(X)}\gamma^{m-n}) && \text{(By (3.7)).}
\end{aligned}$$

We have shown that if $n<m\leq n+n(I,\theta)$ then

$$\mu(X_{I,\theta}\cap E_m)=\mathcal{O}(\mu(X_I)\theta(n)^{\dim_H(X)}\gamma^{m-n}). \tag{3.19}$$

Now let us consider the case where $m>n+n(I,\theta)$. In this case

$$X_{I,\theta}\cap E_m=\Big\{X_{J,\theta}:J\in\mathcal{D}^m \text{ and } (j_1,\ldots,j_{n+n(I,\theta)})=(i_1,\ldots,i_n,z_1,\ldots,z_{n(I,\theta)})\Big\}.$$

Thus we obtain the following bounds

$$\mu(X_{I,\theta}\cap E_m)=\sum_{\substack{J\in\mathcal{D}^m\\(j_1,\ldots,j_{n+n(I,\theta)})=I(z_1,\ldots,z_{n(I,\theta)})}}\mu(X_{J,\theta})$$

$$
\begin{aligned}
&= \sum_{K\in\mathcal{D}^{m-n-n(\theta,I)}} \mu(X_{I(z_1,\ldots,z_{n(I,\theta)})K,\theta}) \\
&\asymp \sum_{K\in\mathcal{D}^{m-n-n(\theta,I)}} (Diam(X_{I(z_1,\ldots,z_{n(I,\theta)})K})\theta(m))^{\dim_H(X)} && \text{(By (3.11))} \\
&\asymp (Diam(X_{I,\theta}))\theta(m))^{\dim_H(X)} \sum_{K\in\mathcal{D}^{m-n-n(\theta,I)}} Diam(X_K)^{\dim_H(X)} && \text{(By (3.5))} \\
&\asymp \mu(X_{I,\theta})\theta(m)^{\dim_H(X)} \sum_{K\in\mathcal{D}^{m-n-n(\theta,I)}} \mu(X_K) && \text{(By (3.6))} \\
&\asymp (Diam(X_I)\theta(n)\theta(m))^{\dim_H(X)} && \text{(By (3.11))} \\
&\asymp \mu(X_I)\theta(n)^{\dim_H(X)}\theta(m)^{\dim_H(X)} && \text{(By (3.6)).}
\end{aligned}
$$

We have shown that if $m > n + n(I,\theta)$ then

$$
\mu(X_{I,\theta} \cap E_m) \asymp \mu(X_I)\theta(n)^{\dim_H(X)}\theta(m)^{\dim_H(X)}. \tag{3.20}
$$

Combining (3.19) and (3.20) we obtain the bound

$$
\mu(X_{I,\theta} \cap E_m) \leq \mathcal{O}\Big(\mu(X_I)\theta(n)^{\dim_H(X)}\gamma^{m-n} + \mu(X_I)\theta(n)^{\dim_H(X)}\theta(m)^{\dim_H(X)}\Big). \tag{3.21}
$$

Importantly this bound holds for all $m > n$.

This implies the following

$$
\begin{aligned}
\sum_{n,m=n(r)}^{Q} \mu(E_n \cap E_m) &= 2\sum_{n=n(r)}^{Q} \mu(E_n) + 2\sum_{n=n(r)}^{Q-1} \sum_{m=n+1}^{Q} \mu(E_n \cap E_m) \\
&= 2\sum_{n=n(r)}^{Q} \mu(E_n) + 2\sum_{n=n(r)}^{Q-1} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n+1}^{Q} \mu(X_{I,\theta} \cap E_m) \\
&= 2\sum_{n=n(r)}^{Q} \mu(E_n) + \mathcal{O}\Big(\sum_{n=n(r)}^{Q-1} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n+1}^{Q} \mu(X_I)\theta(n)^{\dim_H(X)}\gamma^{m-n}\Big) \\
&+ \mathcal{O}\Big(\sum_{n=n(r)}^{Q-1} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n+1}^{Q} \mu(X_I)\theta(n)^{\dim_H(X)}\theta(m)^{\dim_H(X)}\Big).
\end{aligned} \tag{3.22}
$$

We now consider each of the three terms appearing in (3.22) individually. For the first term we have

$$
\begin{aligned}
\sum_{n=n(r)}^{Q} \mu(E_n) &= \sum_{n=n(r)}^{Q} \sum_{I\in\mathcal{D}^{n-n(r)}} \mu(X_{(y_1,\ldots,y_{n(r)})I,\theta}) \\
&\asymp \sum_{n=n(r)}^{Q} \sum_{I\in\mathcal{D}^{n-n(r)}} (Diam(X_{(y_1,\ldots,y_{n(r)})I})\theta(n))^{\dim_H(X)} && \text{(By (3.11))} \\
&\asymp \sum_{n=n(r)}^{Q} \sum_{I\in\mathcal{D}^{n-n(r)}} (Diam(X_{(y_1,\ldots,y_{n(r)})})Diam(X_I)\theta(n))^{\dim_H(X)} && \text{(By (3.5))} \\
&\asymp \mu(X_{(y_1,\ldots,y_{n(r)})}) \sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)} \sum_{I\in\mathcal{D}^{n-n(r)}} \mu(X_I) && \text{(By (3.6))} \\
&\asymp \mu(X_{(y_1,\ldots,y_{n(r)})}) \sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)}.
\end{aligned}
$$

Thus we have shown that

$$\sum_{n=n(r)}^{Q}\mu(E_n)\asymp\mu(X_{(y_1,\ldots,y_{n(r)})})\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}. \tag{3.23}$$

We now focus on the second term in (3.22):

$$\begin{aligned}
&\sum_{n=n(r)}^{Q-1}\sum_{\substack{I\in\mathcal{D}^n\\(i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}}\sum_{m=n+1}^{Q}\mu(X_I)\theta(n)^{\dim_H(X)}\gamma^{m-n}\\
&\asymp\mu(X_{(y_1,\ldots,y_{n(r)})})\sum_{n=n(r)}^{Q-1}\sum_{J\in\mathcal{D}^{n-n(r)}}\sum_{m=n+1}^{Q}\mu(X_J)\theta(n)^{\dim_H(X)}\gamma^{m-n} && \text{(By (3.4))}\\
&=\mu(X_{(y_1,\ldots,y_{n(r)})})\sum_{n=n(r)}^{Q-1}\theta(n)^{\dim_H(X)}\sum_{J\in\mathcal{D}^{n-n(r)}}\mu(X_J)\sum_{m=n+1}^{Q}\gamma^{m-n}\\
&=\mathcal{O}\Big(\mu(X_{(y_1,\ldots,y_{n(r)})})\sum_{n=n(r)}^{Q-1}\theta(n)^{\dim_H(X)}\Big). && (3.24)
\end{aligned}$$

In the last line above we used the fact that $\gamma\in(0,1)$ to conclude that $\sum_{m=n+1}^{Q}\gamma^{m-n}$ can be bounded above by a constant independent of n and Q. We now turn our attention to the third term in (3.22):

$$\begin{aligned}
&\sum_{n=n(r)}^{Q-1}\sum_{\substack{I\in\mathcal{D}^n\\(i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}}\sum_{m=n+1}^{Q}\mu(X_I)\theta(n)^{\dim_H(X)}\theta(m)^{\dim_H(X)}\\
&\asymp\mu(X_{(y_1,\ldots,y_{n(r)})})\sum_{n=n(r)}^{Q-1}\sum_{J\in\mathcal{D}^{n-n(r)}}\sum_{m=n+1}^{Q}\mu(X_J)\theta(n)^{\dim_H(X)}\theta(m)^{\dim_H(X)} && \text{(By (3.4))}\\
&=\mu(X_{(y_1,\ldots,y_{n(r)})})\sum_{n=n(r)}^{Q-1}\theta(n)^{\dim_H(X)}\sum_{J\in\mathcal{D}^{n-n(r)}}\mu(X_J)\sum_{m=n+1}^{Q}\theta(m)^{\dim_H(X)}\\
&=\mu(X_{(y_1,\ldots,y_{n(r)})})\sum_{n=n(r)}^{Q-1}\theta(n)^{\dim_H(X)}\sum_{m=n+1}^{Q}\theta(m)^{\dim_H(X)}\\
&=\mathcal{O}\Big(\mu(X_{(y_1,\ldots,y_{n(r)})})\Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}\Big)^2\Big). && (3.25)
\end{aligned}$$

Substituting (3.23), (3.24), and (3.25) into (3.22) we obtain

$$\sum_{n,m=n(r)}^{Q}\mu(E_n\cap E_m)=\mathcal{O}\Big(\mu(X_{(y_1,\ldots,y_{n(r)})})\Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}+\Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}\Big)^2\Big)\Big)$$

as required.

□

Proof of statement 2 from Theorem 1.4. By Proposition 3.4 and (3.23) there exists a $C>0$ such that

$$\frac{(\sum_{n=n(r)}^{Q}\mu(E_n))^2}{\sum_{n,m=n(r)}^{Q}\mu(E_n\cap E_m)}$$

can be bounded below by

$$\frac{\Big(\mu(X_{(y_1,\ldots,y_{n(r)})})\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}\Big)^2}{C\mu(X_{(y_1,\ldots,y_{n(r)})})\Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}+\Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}\Big)^2\Big)}. \tag{3.26}$$

It is a consequence of our approximating function θ satisfying the divergence condition (1.5), that for Q sufficiently large

$$\sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)} > 1,$$

and therefore

$$\sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)} < \Big(\sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)}\Big)^2.$$

Taking limits in (3.26) we obtain

$$\limsup_{Q\to\infty} \frac{(\sum_{n=n(r)}^{Q} \mu(E_n))^2}{\sum_{n,m=n(r)}^{Q} \mu(E_n \cap E_m)} \geq \frac{\mu(X_{(y_1,\ldots,y_{n(r)})})}{2C}. \tag{3.27}$$

By Proposition 3.3 we may apply Lemma 3.2. Applying Lemma 3.2 and (3.27) we obtain

$$\mu(\limsup E_n) \geq \frac{\mu(X_{(y_1,\ldots,y_{n(r)})})}{2C}.$$

Thus we may conclude that (3.14) holds and we have completed our proof. □

4 Proofs of Theorem 1.6 and Theorem 1.8

The critical part of the proof of Theorem 1.4 was obtaining estimates for $\mu(X_{I,\theta} \cap E_m)$ when $n < m \leq n + n(I,\theta)$. Indeed this was the only point in our proof where the decreasing assumption on θ was used. To weaken the monotonicity assumptions used in the proof of Theorem 1.4, we need new methods to obtain bounds on $\mu(X_{I,\theta} \cap E_m)$ for $n < m \leq n + n(I,\theta)$. The hypothesis appearing in Theorem 1.6 and the condition appearing in Theorem 1.8 provide different methods for bounding $\mu(X_{I,\theta} \cap E_m)$ for $n < m \leq n + n(I,\theta)$.

4.1 Proof of Theorem 1.6

Let $z \in X$, $y \in X$ and $r > 0$ all be as in the proof of Theorem 1.4. As in the proof of Theorem 1.4, to prove Theorem 1.6, it suffices to show that (3.14) holds. The first step in proving Theorem 1.6 is to prove the following more general analogue of Proposition 3.4.

Proposition 4.1. *For any $\theta : \mathbb{N} \to (0, 1/2)$ we have*

$$\sum_{n,m=n(r)}^{Q} \mu(E_n \cap E_m) = \mathcal{O}\Big(\mu(X_{(y_1,\ldots,y_{n(r)})})\Big(\sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)} + \sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)} \log\Big(\frac{1}{\theta(n)}\Big) + \Big(\sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)}\Big)^2\Big)\Big). \tag{4.1}$$

Proof. The quantity $n(I,\theta)$ can be taken to be the smallest $N \in \mathbb{N}$ for which

$$Diam(X_{I(z_1,\ldots,z_N)}) < Diam(X_I)\theta(n).$$

It is a consequence of (3.5) that for each $N \in \mathbb{N}$ we have

$$Diam(X_{I(z_1,\ldots,z_N)}) = \mathcal{O}(Diam(X_I)\kappa^N).$$

Where $\kappa \in (0,1)$ is some constant independent of I and z. Combining these two statements we can prove the following bound

$$n(I,\theta) = \mathcal{O}\Big(\log\Big(\frac{1}{\theta(n)}\Big)\Big). \tag{4.2}$$

In the derivation of (4.2) we use the fact that θ only takes values in the interval $(0, 1/2)$. This assumption means we don't need to worry about constants or $\log(\theta(n)^{-1})$ being negative. Applying (4.2) we have

$$\sum_{m=n+1}^{n+n(I,\theta)} \mu(X_{I,\theta} \cap E_m) \leq n(I,\theta)\mu(X_{I,\theta}) = \mathcal{O}\Big(\mu(X_{I,\theta}) \log\Big(\frac{1}{\theta(n)}\Big)\Big). \tag{4.3}$$

By similar arguments to those given in Section 3, we can combine properties (3.4)–(3.8) with (4.3) to prove

$$\sum_{m=n+1}^{n+n(I,\theta)} \mu(X_{I,\theta} \cap E_m) = \mathcal{O}\Big(\mu(X_I)\theta(n)^{\dim_H(X)} \log\Big(\frac{1}{\theta(n)}\Big)\Big). \tag{4.4}$$

We now obtain an analogue of (3.22):

$$\begin{aligned}
\sum_{n,m=n(r)}^{Q} \mu(E_n \cap E_m) &= 2\sum_{n=n(r)}^{Q} \mu(E_n) + 2\sum_{n=n(r)}^{Q-1}\sum_{m=n+1}^{Q} \mu(E_n \cap E_m) \\
&\asymp \sum_{n=n(r)}^{Q} \mu(E_n) + \sum_{n=n(r)}^{Q-1} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n+1}^{Q} \mu(X_{I,\theta} \cap E_m) \\
&\asymp \sum_{n=n(r)}^{Q} \mu(E_n) + \sum_{n=n(r)}^{Q-1} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n+1}^{n+n(I,\theta)} \mu(X_{I,\theta} \cap E_m) + \\
&+ \sum_{n=n(r)}^{Q-1} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n+n(I,\theta)+1}^{Q} \mu(X_{I,\theta} \cap E_m) \\
&= \sum_{n=n(r)}^{Q} \mu(E_n) + \mathcal{O}\Big(\sum_{n=n(r)}^{Q-1} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \mu(X_I)\theta(n)^{\dim_H(X)} \log\Big(\frac{1}{\theta(n)}\Big)\Big) \\
&+ \mathcal{O}\Big(\sum_{n=n(r)}^{Q-1} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n+n(I,\theta)+1}^{Q} \mu(X_I)\theta(n)^{\dim_H(X)}\theta(m)^{\dim_H(X)}\Big) \qquad (4.5) \\
&\leq \sum_{n=n(r)}^{Q} \mu(E_n) + \mathcal{O}\Big(\sum_{n=n(r)}^{Q-1} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \mu(X_I)\theta(n)^{\dim_H(X)} \log\Big(\frac{1}{\theta(n)}\Big)\Big) \\
&+ \mathcal{O}\Big(\sum_{n=n(r)}^{Q} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n}^{Q} \mu(X_I)\theta(n)^{\dim_H(X)}\theta(m)^{\dim_H(X)}\Big). \qquad (4.6)
\end{aligned}$$

In (4.5) we used the bounds given by (4.4) and (3.20). We now focus on the three terms in (4.6) individually. By identical arguments to those given in Proposition 3.4, we have the following bounds for the first and third term:

$$\sum_{n=n(r)}^{Q} \mu(E_n) \asymp \mu(X_{(y_1,\ldots,y_{n(r)})}) \sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)} \tag{4.7}$$

$$\sum_{n=n(r)}^{Q} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n}^{Q} \mu(X_I)\theta(n)^{\dim_H(X)}\theta(m)^{\dim_H(X)} \asymp \mu(X_{(y_1,\ldots,y_{n(r)})})\Big(\sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)}\Big)^2. \tag{4.8}$$

It remains to bound the second term:

$$
\begin{aligned}
&\sum_{n=n(r)}^{Q-1} \sum_{\substack{I\in\mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \mu(X_I)\theta(n)^{\dim_H(X)}\log\Big(\frac{1}{\theta(n)}\Big) \\
&\asymp \mu(X_{(y_1,\ldots,y_{n(r)})}) \sum_{n=n(r)}^{Q-1} \sum_{J\in\mathcal{D}^{n-n(r)}} \mu(X_J)\theta(n)^{\dim_H(X)}\log\Big(\frac{1}{\theta(n)}\Big) \qquad \text{(By (3.4))} \\
&\asymp \mu(X_{(y_1,\ldots,y_{n(r)})}) \sum_{n=n(r)}^{Q-1} \theta(n)^{\dim_H(X)}\log\Big(\frac{1}{\theta(n)}\Big). \qquad (4.9)
\end{aligned}
$$

Substituting (4.7), (4.8), and (4.9) into (4.6) we obtain (4.1). □

Equipped with Proposition 4.1 we are now in a position to prove Theorem 1.6.

Proof of Theorem 1.6. We assume that $\theta(n) \to 0$ as $n \to \infty$. Consequently, dividing by a positive constant if necessary, we may assume that θ satisfies the hypothesis of Proposition 4.1. The case where θ does not converge to 0 is easily dealt with. Under this assumption it can be shown that $W(\theta, z)$ is always a set of full μ measure for any $z \in X$. We omit the details of this fact.

Proposition 4.1 and (4.7) implies that there exists a $C > 0$ such that

$$\frac{(\sum_{n=n(r)}^{Q}\mu(E_n))^2}{\sum_{n,m=n(r)}^{Q}\mu(E_n\cap E_m)}$$

can be bounded below by

$$\frac{\mu(X_{(y_1,\ldots,y_{n(r)})})\Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}\Big)^2}{C\Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}+\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}\log\Big(\frac{1}{\theta(n)}\Big)+\Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}\Big)^2\Big)}. \qquad (4.10)$$

For Q sufficiently large

$$\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)} > 1$$

and therefore

$$\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)} < \Big(\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}\Big)^2.$$

We take the limsup in (4.10). It is a consequence of the above remark and our additional assumption (1.7), that there exists a $C' > 0$ such that

$$\limsup_{Q\to\infty}\frac{(\sum_{n=n(r)}^{Q}\mu(E_n))^2}{\sum_{n,m=n(r)}^{Q}\mu(E_n\cap E_m)} \geq \frac{\mu(X_{(y_1,\ldots,y_{n(r)})})}{C'}.$$

Importantly C' is independent of y and r. Proposition 3.3 still holds for this choice of θ, so we may apply Lemma 3.2. Applying Lemma 3.2 we may conclude that

$$\mu(\limsup E_n) \geq \frac{\mu(X_{(y_1,\ldots,y_{n(r)})})}{C'}.$$

Thus (3.14) holds for any θ satisfying (1.7) and we have completed our proof. □

4.2 Proof of Theorem 1.8

Assume that $y \in X$ and $r > 0$ are all as in the proof of Theorem 1.4. We will show that for any z satisfying an additional condition, whenever θ satisfies (1.5) then (3.14) will hold.

As in the proof of Theorem 1.6 we need to control $\mu(X_{I,\theta} \cap E_m)$ for $n < m \leq n + n(I,\theta)$. We start with a simple observation. Let $I \in \mathcal{D}^n$ and $J \in \mathcal{D}^m$ for $n < m \leq n + n(I,\theta)$. Consider the words $I_\theta := (i_1, \ldots, i_n, z_1, \ldots, z_{n(I,\theta)})$ and $J_\theta := (j_1, \ldots, j_m, z_1, \ldots, z_{n(J,\theta)})$. If $X_{I,\theta} \cap X_{J,\theta} \neq \emptyset$ then either I_θ is a prefix of J_θ, or J_θ is a prefix of I_θ, i.e.,

$$I_\theta = (j_1, \ldots, j_m, z_1, \ldots, z_{n(I,\theta)-(m-n)}) \tag{4.11}$$

or

$$J_\theta = (i_1, \ldots, i_n, z_1, \ldots, z_{n(J,\theta)+(m-n)}). \tag{4.12}$$

If (4.11) holds then

$$(z_1, \ldots, z_{n(I,\theta)-(m-n)}) = (z_{m-n+1}, \ldots, z_{n(I,\theta)}). \tag{4.13}$$

Alternatively, if (4.12) holds then

$$(z_1, \ldots, z_{n(J,\theta)}) = (z_{m-n+1}, \ldots, z_{n(J,\theta)+(m-n)}). \tag{4.14}$$

Consequently, we see that the nonempty intersection of $X_{I,\theta}$ and $X_{J,\theta}$ implies some nontrivial relations for the coding (z_i). This leads to the following definition.

We say that z has a *leading block coding* if there exists a sequence $(z_i) \in \mathcal{D}^{\mathbb{N}}$, such that (z_i) is a coding of z and there exists $l \in \mathbb{N}$ such that the initial word $(z_1, \ldots, z_l)$ appears only finitely many times in (z_i). Our first step in the proof of Theorem 1.8 is the following proposition.

Proposition 4.2. *Suppose z has a leading block coding and $\theta(n) \to 0$ as $n \to \infty$. Then*

$$\sum_{n,m=n(r)}^{Q} \mu(E_n \cap E_m) = \mathcal{O}\Big(\mu(X_{(y_1,\ldots,y_{n(r)})})\Big(\sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)} + \Big(\sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)}\Big)^2\Big)\Big).$$

Proof. Replicating the arguments used in the proof of Proposition 3.4, it suffices to show that

$$\sum_{n=n(r)}^{Q-1} \sum_{\substack{I \in \mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n+1}^{n+n(I,\theta)} \mu(X_{I,\theta} \cap E_m) = \mathcal{O}\Big(\mu(X_{(y_1,\ldots,y_{n(r)})}) \sum_{n=n(r)}^{Q} \theta(n)^{\dim_H(X)}\Big). \tag{4.15}$$

Assume that $(z_1, \ldots, z_l)$ appears only finitely many times in the coding (z_i). Since $\theta(n) \to 0$ we must have $n(I,\theta) \to \infty$ as $|I| \to \infty$. It follows that for n sufficiently large, if $X_{I,\theta} \cap X_{J,\theta} \neq \emptyset$ and (4.14) occurs, then

$$(z_1, \ldots, z_l) = (z_{m-n+1}, \ldots, z_{m-n+l}).$$

This can only occur finitely many times by definition. Suppose $X_{I,\theta} \cap X_{J,\theta} \neq \emptyset$ and (4.13) occurs, then if $n < m \leq n + n(I,\theta) - l$ we have

$$(z_1, \ldots, z_l) = (z_{m-n+1}, \ldots, z_{m-n+l}).$$

Which can only occur finitely many times by definition. As there at most $l-1$ values of m satisfying $n + n(I,\theta) - l < m \leq n + n(I,\theta)$, it follows that

$$\sup_{n \in \mathbb{N}} \sup_{I \in \mathcal{D}^n} \#\Big\{m \in \mathbb{N} : n < m \leq n + n(I,\phi) \text{ and } X_{I,\theta} \cap E_m \neq \emptyset\Big\} < \infty.$$

Therefore,

$$\sum_{n=n(r)}^{Q-1} \sum_{\substack{I \in \mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \sum_{m=n+1}^{n+n(I,\theta)} \mu(X_{I,\theta} \cap E_m) = \mathcal{O}\Big(\sum_{n=n(r)}^{Q-1} \sum_{\substack{I \in \mathcal{D}^n \\ (i_1,\ldots,i_{n(r)})=(y_1,\ldots,y_{n(r)})}} \mu(X_{I,\theta})\Big)$$

$$= \mathcal{O}\Big(\mu(X_{(y_1,\ldots,y_{n(r)})})\sum_{n=n(r)}^{Q-1}\sum_{J\in\mathcal{D}^{n-n(r)}}\mu(X_J)\theta(n)^{\dim_H(X)}\Big)$$

$$= \mathcal{O}\Big(\mu(X_{(y_1,\ldots,y_{n(r)})})\sum_{n=n(r)}^{Q}\theta(n)^{\dim_H(X)}\Big).$$

The proof of the second equality above follows from the same arguments used in the proof of Theorem 1.6. Thus we have proved (4.15) and our proof is complete.

□

By an analogous argument to that given in the proof of Theorem 1.4, we may combine Proposition 4.2 with Lemma 3.2 and Proposition 3.3, to conclude that if z has a leading block coding and θ is any function satisfying $\theta(n) \to 0$ as $n \to \infty$ and (1.5), then (3.14) holds and consequently μ almost every x is an element of $W(\theta, z)$. As we previously remarked upon, the case where θ does not converge to 0 is easily dealt with. Under this assumption we always have that $W(\theta, z)$ is a set of full μ measure for any $z \in X$. Summarising the above, we have the following result.

Proposition 4.3. *Suppose z has a leading block coding and $\theta : \mathbb{N} \to \mathbb{R}_{\geq 0}$ satisfies* (1.5)*, then μ almost every x is an element of $W(\theta, z)$.*

To prove Theorem 1.8 we need to prove that the set of z with a leading block coding has full Hausdorff dimension.

Proposition 4.4. $\dim_H(\{z \in X : z$ *has a leading block coding*$\}) = \dim_H(X)$.

Proof. Let $Y := \{z \in X : z \text{ has a leading block coding}\}$. Let $N \in \mathbb{N}$ and let us fix a digit $i \in \mathcal{D}$. Consider the word $i^N = (i, \ldots, i) \in \mathcal{D}^N$ and the IFS $\{\phi_I\}_{I\in\mathcal{D}^N\setminus\{i^N\}}$ with corresponding attractor $X^{i,N}$. For any $\epsilon > 0$ we can pick N sufficiently large that

$$\dim_H(X) - \epsilon < \dim_H(X^{i,N}) \leq \dim_H(X). \tag{4.16}$$

Now consider the set $\phi_{i^{2N}}(X^{i,N})$, where $i^{2N} \in \mathcal{D}^{2N}$ is the word consisting of $2N$ consecutive i's. Every element of $\phi_{i^{2N}}(X^{i,N})$ has a coding that begins with i^{2N} and for which i^{2N} occurs only finitely many times. This is because every sequence in $(\mathcal{D}^N \setminus \{i^N\})^{\mathbb{N}}$ cannot contain the word i^{2N}. Therefore $\phi_{i^{2N}}(X^{i,N}) \subseteq Y$. Since $\phi_{i^{2N}}$ is a bi-Lipschitz map it follows from (4.16) that

$$\dim_H(X) - \epsilon < \dim_H(\phi_{i^{2N}}(X^{i,N})) \leq \dim_H(X).$$

Consequently,

$$\dim_H(X) - \epsilon < \dim_H(Y) \leq \dim_H(X).$$

Since ϵ is arbitrary we must have $\dim_H(Y) = \dim_H(X)$ as required. □

Combining Proposition 4.3 and Proposition 4.4 we may conclude Theorem 1.8.

5 The mass transference principle and some applications

Suppose $\Psi : \mathcal{D}^* \to \mathbb{R}$ is such that $\sum_{n=1}^{\infty}\sum_{I\in\mathcal{D}^n}\Psi(I)^{\dim_H(X)} < \infty$, then by Theorem 1.4 we know that $\mathcal{H}^{\dim_H(X)}(W(\Psi, z)) = 0$ for any $z \in X$. Under this assumption it is natural to ask what is the Hausdorff dimension of $W(\Psi, z)$. In this section we obtain a partial solution to this question by employing the mass transference principle introduced by Beresnevich and Velani [4]. We end this section with some applications of this result.

5.1 The mass transference principle

The first result of this section is the following theorem.

Theorem 5.1. *Let Φ be an iterated function system with attractor X.*

1. *Let $\Psi : \mathcal{D}^* \to \mathbb{R}_{\geq 0}$ and $s > 0$. Suppose*

$$\sum_{n=1}^{\infty} \sum_{I \in \mathcal{D}^n} \Psi(I)^s < \infty,$$

then $\mathcal{H}^s(W(\Psi, z)) = 0$ for all $z \in X$.

2. *If Φ is a conformal iterated function system satisfying the open set condition and $\theta : \mathbb{N} \to \mathbb{R}_{\geq 0}$ is a decreasing function that satisfies* (1.5)*, then*

$$\mathcal{H}^{\dim_H(X)/t}(W((Diam(X_I)\theta(|I|))^t, z)) = \mathcal{H}^{\dim_H(X)/t}(X)$$

for all $z \in X$ and $t \geq 1$.

Taking $t = 1$ in the above we observe that statement 2 from Theorem 1.4 is a consequence of Theorem 5.1. It is in fact the case that the opposite implication is true. Theorem 1.4 implies Theorem 5.1. This is because of the mass transference principle. We now briefly detail this technique.

Let $X \subset \mathbb{R}^d$ and assume that X is Ahlfors regular. Given $s > 0$ and a ball $B = B(x, r)$, we let $B^s = B(x, r^{s/\dim_H(X)})$. The following theorem is a simplified version of Theorem 3 proved in [4]. It will allow us to prove that Theorem 1.4 implies Theorem 5.1.

Theorem 5.2. *Let X be as above and $(B_l)_{l=1}^{\infty}$ be a sequence of balls in X with radii tending to zero. Let $s > 0$ and suppose that for any ball B in X we have*

$$\mathcal{H}^{\dim_H(X)}(B \cap \limsup_{l\to\infty} B_l^s) = \mathcal{H}^{\dim_H(X)}(B).$$

Then, for any ball B in X

$$\mathcal{H}^s(B \cap \limsup_{l\to\infty} B_l) = \mathcal{H}^s(B).$$

We now prove Theorem 5.1 using Theorem 5.2.

Proof of Theorem 5.1. The proof of statement 1 is analogous to the proof of statement 1 from Theorem 1.4 so we omit it.

Suppose Φ is a conformal iterated function system satisfying the open set condition. Assume $\theta : \mathbb{N} \to \mathbb{R}_{\geq 0}$ is a decreasing function satisfying (1.5). Let $z \in X$, $t \geq 1$, and $(B_l)_{l=1}^{\infty}$ be an enumeration of the set of balls $\{B(\phi_I(z), (Diam(X_I)\theta(|I|))^t)\}_{I \in D^*}$. It follows from the definition that $(B_l^{\dim_H(X)/t})_{l=1}^{\infty}$ is an enumeration of the set of balls $\{B(\phi_I(z), Diam(X_I)\theta(|I|))\}_{I \in \mathcal{D}^*}$. In which case, by Theorem 1.4 we know that

$$\mathcal{H}^{\dim_H(X)}(B \cap \limsup_{l\to\infty} B_l^{\dim_H(X)/t}) = \mathcal{H}^{\dim_H(X)}(B)$$

for any ball B. Applying Theorem 5.2 we may conclude that

$$\mathcal{H}^{\dim_H /t}(B \cap \limsup_{l\to\infty} B_l) = \mathcal{H}^{\dim_H(X)/t}(B),$$

for any ball B, and

$$\mathcal{H}^{\dim_H(X)/t}(W((Diam(X_I)\theta(|I|))^t, z)) = \mathcal{H}^{\dim_H(X)/t}(X).$$

□

Statement 2 from Theorem 5.1 implies a lower bound for the Hausdorff dimension of certain limsup sets. It is a consequence of statement 1 that this bound is in fact optimal and implies the following result.

Corollary 5.3. *If Φ is a conformal iterated function system satisfying the open set condition and $\theta : \mathbb{N} \to \mathbb{R}_{\geq 0}$ is a decreasing function that satisfies* (1.5)*, then*

$$\dim_H(W((Diam(X_I)\theta(|I|))^t, z)) = \frac{\dim_H(X)}{t}$$

for all $z \in X$ and $t \geq 1$

We leave the proof of Corollary 5.3 to the interested reader.

Remark 5.4. One would like to improve Theorem 5.1 to a more general statement that covered arbitrary dimension functions. Such a result appeared in [17]. However, since the radii of our balls take the restrictive form $Diam(X_I)\theta(|I|)$, we cannot prove such a general statement. In the more general case, the analogue of the normalised ball B^s does not take a form we can work with.

5.2 Applications

We now include some applications of Theorem 5.1. Before giving these applications it is useful to build some general theory.

Let E be some subset of $\mathbb{R}^d$. We call a function $H : E \to \mathbb{R}_{>0}$ a *height function* on E. We say that an IFS Φ *respects* E *and* H if the following two conditions hold:

- $\phi(x) \in E$ for all $x \in E$ and $\phi \in \Phi$.
- There exists $C > 1$ such that $H(\phi(x)) < CH(x)$ for all $x \in E$ and $\phi \in \Phi$.

Proposition 5.5. *Let Φ be a conformal iterated function system satisfying the open set condition and H be a height function on a set E. Suppose $E \cap X \neq \emptyset$ and Φ respects E and H, then for any $l > 0$ we have*

$$\dim_H(X)\Big(\Big\{x \in X : |x - e| < H(e)^{-l} \textit{ for i.m. } e \in E\Big\}\Big) > 0.$$

Proof. Fix Φ, H, E and l. Let $a \in E \cap X$. We will prove that

$$W(Diam(X_I)^t, a) \subseteq \Big\{x \in X : |x - e| < H(e)^{-l} \text{ for i.m. } e \in E\Big\} \tag{5.1}$$

for some suitable choice of $t \geq 1$. Our result will then follow from Corollary 5.3. Since each element of Φ is a contraction there exists $\gamma \in (0, 1)$ such that

$$Diam(X_I) < \gamma^{|I|} Diam(X) \tag{5.2}$$

for all $I \in \mathcal{D}^*$. Since Φ respects E and H we have

$$H(\phi_I(a)) \leq C^{|I|} H(a) \tag{5.3}$$

for all $I \in \mathcal{D}^*$. Let us now choose t sufficiently large that

$$(\gamma^n Diam(X))^t < \Big(\frac{1}{C^n H(a)}\Big)^l, \tag{5.4}$$

for all n sufficiently large. As a consequence of (5.2), (5.3) and (5.4) the following inclusions hold for $|I|$ sufficiently large

$$B(\phi_I(a), Diam(X_I)^t) \subseteq B(\phi_I(a), (\gamma^{|I|} Diam(X))^t) \subseteq B(\phi_I(a), (C^{|I|} H(a))^{-l}) \subseteq B(\phi_I(a), H(\phi_I(a))^{-l}).$$

These inclusions along with the fact $\phi_I(a) \in E$ for all $I \in \mathcal{D}^*$ imply (5.1). $\square$

Proposition 5.5 allows us to deduce the existence of elements within our attractor X that are "very well approximable," where the definition of "very well approximable" may be defined in a way that is independent from our attractor X. In the next section we give two applications which exhibit this phenomenon.

Note that in the proof of Proposition 5.5 we didn't require the full strength of Corollary 5.3. One can prove this result by combining elementary arguments with the mass transference principle. Therefore, Proposition 5.5 and the applications below should be interpreted as consequence of the mass transference principle within the setting of IFS's, rather than as a consequence of approximation regularity.

5.2.1 A problem of Mahler

A classical theorem due to Dirichlet states that for any $\alpha = (\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$ and $Q \in \mathbb{N}$, there exists $(p_1, \ldots, p_d) \in \mathbb{Z}^d$ and $1 \leq q < Q^d$ such that

$$\max_{1 \leq i \leq d} |q\alpha_i - p_i| \leq \frac{1}{Q}.$$

This theorem implies that for any $(\alpha_1, \ldots, \alpha_d) \in \mathbb{R}^d$ with at least one α_i irrational, there exists infinitely many $(p_1/q, \ldots, p_d/q) \in \mathbb{Q}^d$ such that

$$\max_{1 \leq i \leq d} \Big|\alpha_i - \frac{p_i}{q}\Big| \leq \frac{1}{q^{1+1/d}}.$$

We call $\alpha \in \mathbb{R}^d$ *very well approximable* if there exists $\tau > 1 + 1/d$ for which there exists infinitely many $(p_1/q, \ldots, p_d/q) \in \mathbb{Q}^d$ satisfying

$$\max_{1 \leq i \leq d} \left|\alpha_i - \frac{p_i}{q}\right| \leq \frac{1}{q^\tau}. \tag{5.5}$$

We call α *Liouville* if it satisfies (5.5) for all $\tau > 1 + 1/d$. It is well known that the set of very well approximable numbers has Lebesgue measure zero, and the set of Liouville numbers has Hausdorff dimension zero. In [5] the following assertion was made and attributed to Mahler.

There exists very well approximable numbers, other than Liouville numbers, in the middle third Cantor set.

This assertion was proved to be correct in [17]. Related work appears in [9]. Applying Proposition 5.5 we now obtain a more general version of this result. We call Φ a *rational iteration function system* if each element of Φ is a similarity of the form $\phi_i(x) = \frac{p_i x}{q_i} + (\frac{a_{i,1}}{b_{i,1}}, \ldots, \frac{a_{i,d}}{b_{i,d}})$ where $p_i, q_i \in \mathbb{Z}$ and $a_{i,1}, b_{i,1}, \ldots, a_{i,d}, b_{i,d} \in \mathbb{Z}$.

Theorem 5.6. *Let Φ be a rational iterated function system satisfying the open set condition, then X contains very well approximable numbers that are not Liouville.*

Proof. By Proposition 5.5 and the fact that the Liouville numbers have Hausdorff dimension zero, it suffices to show that $X \cap \mathbb{Q}^d \neq \emptyset$ and Φ respects $\mathbb{Q}^d$ and H. Where $H : \mathbb{Q}^d \to \mathbb{N}$ is the height function defined to be $H((\frac{p_1}{q_1}, \ldots, \frac{p_d}{q_d})) = \mathrm{lcm}(q_1, \ldots, q_d)$, where in this expression each p_i/q_i is assumed to be in its reduced form and lcm denotes the lowest common multiple.

To see that $X \cap \mathbb{Q}^d \neq \emptyset$, we remark that for any $\phi_i \in \Phi$ the unique fixed point satisfying $\phi_i(x) = x$ is contained in X. Proving that this fixed point is contained in $\mathbb{Q}^d$ follows immediately from the definition of a rational iterated function system.

We now show that Φ respects $\mathbb{Q}^d$ and H. Clearly any element of Φ maps $\mathbb{Q}^d$ to $\mathbb{Q}^d$. It remains to show that we satisfy the required growth condition. Fix $\overline{x} = (\frac{x_1}{y_1}, \ldots, \frac{x_d}{y_d}) \in \mathbb{Q}^d$. Let $y = \mathrm{lcm}(y_1, \ldots, y_m)$, by an abuse of notation we write $\overline{x} = (\frac{x_1}{y}, \ldots, \frac{x_d}{y})$. We observe the following

$$\begin{aligned}\phi_i(\overline{x}) &= \Big(\frac{p_i x_1}{q_i y}, \ldots, \frac{p_i x_d}{q_i y}\Big) + \Big(\frac{a_{i,1}}{b_{i,1}}, \ldots, \frac{a_{i,d}}{b_{i,d}}\Big)\\ &= \Big(\frac{b_{i,1} p_i x_1 + a_{i,1} q_i y}{b_{i,1} q_i y}, \ldots, \frac{b_{i,d} p_i x_d + a_{i,d} q_i y}{b_{i,d} q_i y}\Big).\end{aligned} \tag{5.6}$$

Each term in (5.6) can be rewritten to have denominator $q_i y \cdot \prod_{j=1}^d b_{i,j}$. Therefore

$$H(\phi_i(\overline{x})) \leq q_i y \cdot \prod_{j=1}^d b_{i,j}.$$

Taking $C = \sup_{i \in \mathcal{D}} q_i \cdot \prod_{j=1}^d b_{i,j}$ we have

$$H(\phi_i(\overline{x})) \leq C H(\overline{x}),$$

for all $\overline{x} \in \mathbb{Q}^d$ and $\phi_i \in \Phi$. Thus Φ respects $\mathbb{Q}^d$ and H. $\square$

5.2.2 Approximating badly approximable numbers by quadratic irrationals

For our next application we prove that there exist badly approximable numbers that are "very well approximated" by quadratic irrationals. We start by recalling what it means to be badly approximable and give an overview of approximation by algebraic numbers.

We call $x \in (0, 1)$ *badly approximable* if there exists $\kappa(x) > 0$ such that

$$\left|x - \frac{p}{q}\right| > \frac{\kappa(x)}{q^2} \text{ for all } (p, q) \in \mathbb{Z} \times \mathbb{N}.$$

Every $x \in (0, 1) \setminus \mathbb{Q}$ has a unique continued fraction expansion, that is a unique sequence $(a_n(x)) \in \mathbb{N}^\mathbb{N}$ such that

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cdots}}} := \lim_{n \to \infty} \cfrac{1}{a_1 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}.$$

We refer to the sequence (a_n) as the partial quotients of x. Badly approximable numbers are characterised by their continued fraction expansion. It is well known that x is badly approximable if and only if its sequence of partial quotients is bounded [5].

We now detail some of the background behind approximation by algebraic numbers. What follows is taken from [5]. The height of an algebraic number α, denoted by $H(\alpha)$, is the maximum of the moduli of the coefficients of its minimal polynomial. For example, we have $H(\sqrt{2}) = 2$ since the minimal polynomial of $\sqrt{2}$ is $x^2 - 2$. The degree of an algebraic number α, denoted by $\deg(\alpha)$, is the degree of the minimal polynomial.

Given $x \in \mathbb{R}$ and $n \in \mathbb{N}$, we define $w_n(x)$ to be the supremum of the real numbers ω for which there exists i.m. real algebraic numbers α with $\deg(\alpha) \leq n$ satisfying

$$0 < |x - \alpha| < H(\alpha)^{-\omega-1}.$$

We then let

$$w(x) := \limsup_{n\to\infty} \frac{w_n(x)}{n}.$$

We will not use the quantity $w(x)$. We merely remark that it can be used to give a classification of the real numbers in terms of how transcendental they are. This classification is known as Koksma's classification [16]. The following result describes the generic behaviour of $w_n(x)$.

Theorem 5.7. *Lebesgue almost every $x \in \mathbb{R}$ satisfies $w_n(x) = n$ for each $n \in \mathbb{N}$.*

For a proof of Theorem 5.7 see [5]. The proof of this theorem is originally due to Sprindžuk [20, 21]. For our applications we will only require Theorem 5.7 when $n = 2$.

We call $x \in \mathbb{R}$ *quadratically very well approximable* if $w_2(x) > 2$. We call $x \in \mathbb{R}$ *quadratically Liouville* if $w_2(x) = \infty$. It is a consequence of a result due to Kasch and Volkmann [14] that the set of quadratically Liouville numbers has Hausdorff dimension zero. The following theorem demonstrates that there are badly approximable numbers that are quadratically very well approximable but not quadratically Liouville.

Theorem 5.8. *Let $D \subset \mathbb{N}$ be a finite set that contains at least two elements and $X_D := \{x \in (0,1) \setminus \mathbb{Q} : (a_n(x)) \in D^{\mathbb{N}}\}$. Then X_D contains numbers that are quadratically very well approximable but not quadratically Liouville.*

Proof. We start our proof by remarking that X_D can be identified with the unique attractor for the IFS $\Phi = \{\phi_i\}_{i\in D}$ where $\phi_i(x) = \frac{1}{x+i}$. Moreover, this Φ is a conformal iterated function system satisfying the open set condition. Therefore we can use Proposition 5.5. Let E be the set of quadratic irrationals. We know by the aforementioned result of Kasch and Volkmann that the set of quadratically Liouville numbers has Hausdorff dimension zero. Therefore to complete our proof it suffices to show that we satisfy the remaining hypothesis of Proposition 5.5. Namely we need to show that $E \cap X_D \neq \emptyset$ and Φ respects E and H. We have $E \cap X_D \neq \emptyset$ since every eventually periodic $(a_n) \in D^{\mathbb{N}}$ is the continued fraction expansion of a quadratic irrational. It remains to show that Φ respects E and H. Fix $\alpha \in E$ with minimal polynomial $ax^2 + bx + c$. Then $H(\alpha) = \max\{|a|, |b|, |c|\}$. A simple calculation shows that $\phi_i(\alpha)$ is a root of the polynomial

$$(ai^2 - bi + c)x^2 + (b - 2ia)x + a. \tag{5.7}$$

Therefore $\phi_i(\alpha)$ is either a quadratic irrational or a rational number. However, if $\phi_i(\alpha)$ is rational then it can be shown that α is also rational, a contradiction. Therefore each element of Φ maps E to E and it remains to show we satisfy the growth condition. Since $\phi_i(\alpha)$ is a quadratic irrational, (5.7) cannot be factorised into two linear factors. Consequently, (5.7) can be written in the form $nM(x)$, where $n \in \mathbb{Z}\setminus\{0\}$ and $M(x) \in \mathbb{Z}[x]$ is the minimal polynomial of $\phi_i(\alpha)$. It follows that

$$\begin{aligned} H(\phi_i(\alpha)) &\leq \max\{|ai^2 - bi + c|, |b - 2ia|, |a|\} \\ &\leq \max\{(|a| + |b| + |c|)i^2, (|b| + 2|a|)i, |a|\} \\ &\leq 3i^2 H(\alpha). \end{aligned}$$

Taking $C = \max_{i\in D}\{3i^2\}$ we have

$$H(\phi_i(\alpha)) \leq CH(\alpha)$$

for all $i \in D$ and $\alpha \in E$. Therefore Φ respects E and H and our proof is complete. □

6 The overlapping case and further directions

We conclude this paper by proving that whenever our IFS contains an exact overlap there are no approximation regular pairs. We also include some discussion of the overlapping case and suggest some future directions.

Understanding the structure of overlapping attractors is a classical problem. For Φ a conformal iterated function system we always have $\dim_H(X) \leq \dim_S(X)$. We also trivially have the upper bound $\dim_H(X) \leq d$ where d is the dimension of the ambient Euclidean space. These two bounds imply

$$\dim_H(X) \leq \min\{\dim_S(X), d\}. \tag{6.1}$$

Determining conditions under which we have equality in (6.1) is an active area of research, see [10, 18] and the references therein. There is a standard way of constructing examples for which we have strict inequality in (6.1). We construct an IFS in such a way that there is an exact overlap, i.e., there exists $I, J \in \mathcal{D}^n$ such that $\phi_I = \phi_J$. This means we can remove one of these maps from our IFS and still be left with the same attractor. This new IFS has a strictly smaller similarity dimension which can lead to a strict inequality in (6.1). It is conjectured that exact overlaps are the only mechanism by which we can have strict inequality in (6.1). Interestingly, as far as the author knows, the only known condition which can result in a pair (Φ, z) failing to be approximation regular is when there is an exact overlap. This result is proved in the following theorem.

Theorem 6.1. *Let Φ be a conformal iterated function system containing an exact overlap. Then there are no approximation regular pairs.*

Proof. Suppose that $I, J \in \mathcal{D}^k$ are such that $\phi_I = \phi_J$. Let us fix $z \in X$. It suffices to construct a function $\theta : \mathbb{N} \to \mathbb{R}_{\geq 0}$ such that we have divergence in (1.4), yet

$$\dim_H(W(Diam(X_I)^{\dim_{S,\Phi}(X)/\dim_H(X)}\theta(|I|), z)) < \dim_H(X). \tag{6.2}$$

Note that in (6.2) we have included the subscript Φ in the similarity dimension. This is to emphasise the dependence on Φ. This dependence will be important in what follows.

Let $\theta(|I|) = 1$ for all $I \in \mathcal{D}^*$, so our approximating function is simply $\Psi(I) = Diam(X_I)^{\dim_{S,\Phi}(X)/\dim_H(X)}$. The summation in (1.4) reduces to

$$\sum_{n=1}^{\infty} \sum_{I \in \mathcal{D}^n} Diam(X_I)^{\dim_{S,\Phi}(X)}.$$

It is a consequence of Bowen's equation and (1.3) that this series diverges. It remains to show that (6.2) holds.

The attractor X is also the attractor for the iterated function system $\Phi^k := \{\phi_I\}_{I \in \mathcal{D}^k \setminus \{J\}}$. This follows by iterating (1.1) and using the fact that we have an exact overlap. Φ^k is a conformal iterated function system, so we can consider it's similarity dimension $\dim_{S,\Phi^k}(X)$. As a consequence of the exact overlap we have

$$\dim_{S,\Phi^k}(X) < \dim_{S,\Phi}(X).$$

Now let us choose $\epsilon > 0$ such that

$$\dim_{S,\Phi^k}(X) < \frac{\dim_{S,\Phi}(X)}{\dim_H(X)} \cdot (\dim_H(X) - \epsilon) \tag{6.3}$$

Replicating arguments from the proof of statement 1 from Theorem 1.4, we can show that for any $\rho > 0$ there exists $M \in \mathbb{N}$ sufficiently large such that we have the following bound

$$\mathcal{H}_\rho^{\dim_H(X)-\epsilon}(W(Diam(X_I)^{\frac{\dim_{S,\Phi}(X)}{\dim_H(X)}}, z)) \leq \sum_{n=M}^{\infty} \sum_{\substack{I \in \mathcal{D}^n \\ J \text{ is not a subword of } I}} Diam(X_I)^{\frac{\dim_{S,\Phi}(X)}{\dim_H(X)} \cdot (\dim_H(X) - \epsilon)}. \tag{6.4}$$

It is a consequence of (6.3) and (1.3) that the second summation on the right hand side of (6.4) tends to zero exponentially fast. Therefore the right hand side converges and M can be chosen to make this summation arbitrarily small. It follows that $\mathcal{H}_\rho^{\dim_H(X)-\epsilon}(W(Diam(X_I)^{\frac{\dim_{S,\Phi}(X)}{\dim_H(X)}}, z)) = 0$, and since ρ is arbitrary we must have $\mathcal{H}^{\dim_H(X)-\epsilon}(W(Diam(X_I)^{\frac{\dim_{S,\Phi}(X)}{\dim_H(X)}}, z)) = 0$. Thus (6.2) holds and our proof is complete. □

Theorem 6.1 and the discussion at the start of this section give rise to several natural questions.

1. Is the only condition under which an IFS fails to be approximation regular when there is an exact overlap?

2. Can one relate the approximation regularity properties of a conformal IFS to other nice properties of the attractor? For example, can it be related to equality in (6.1)? Can it be related to the absolute continuity of certain natural measures supported on X?

3. Another natural direction to pursue is to consider more general attractors. In particular, one can ask what is the analogue of the above theory in the setting of self-affine sets, IFS's consisting of infinitely many contracting maps, and for randomly defined attractors. We currently have no results in this direction. We expect that for a self-affine set the analogue of the divergence condition (1.4) would have to take into account the rotations that might be present with the IFS. This was something we didn't have to consider for conformal iterated function systems.

We believe that approximation regularity can be used as an effective tool to measure how much an attractor overlaps. It is possible however that such an approach is to blunt and a more subtle approach is required. Expecting the divergence of certain sums to be the deciding criteria in determining whether a limsup set has full measure/Hausdorff dimension could be wishful thinking. Instead of looking solely at the divergence of certain sums, one should perhaps put a greater emphasis on determining those approximating functions Ψ for which $W(\Psi, z)$ is of full measure/Hausdorff dimension.

Acknowledgements. The author would like to thank the anonymous referee whose comments improved the exposition of this article.

References

[1] S. Baker, *Approximation properties of β-expansions,* Acta Arith. 168 (2015), 269–287

[2] S. Baker, *Approximation properties of β-expansion II,* Ergodic Theory and Dynamical Systems (to appear).

[3] V. Beresnevich, D. Dickinson, S. Velani, *Measure theoretic laws for lim sup sets,* Mem. Amer. Math. Soc. 179 (2006), no. 846, x+91 pp.

[4] V. Beresnevich, S. Velani, *A mass transference principle and the Duffin-Schaeffer conjecture for Hausdorff measures,* Ann. of Math. (2) 164 (2006), no. 3, 971-992.

[5] Y. Bugeaud, *Approximation by algebraic numbers,* Cambridge Tracts in Mathematics, 2004.

[6] R. J. Duffin, A. C. Schaeffer, *Khintchine's problem in metric Diophantine approximation,* Duke Math. J., 8 (1941). 243-255.

[7] K. Falconer, *Techniques in fractal geometry,* John Wiley, 1997.

[8] D. J. Feng, H. Hu, *Dimension theory of iterated function systems,* Comm. Pure Appl. Math. 62 (2009), no. 11, 1435-1500.

[9] L. Fishman, D. Simmons,*Intrinsic approximation for fractals defined by rational iterated function systems: Mahler's research suggestion* Proc. Lond. Math. Soc. (3) 109 (2014), no. 1, 189-212.

[10] M. Hochman, *On self-similar sets with overlaps and inverse theorems for entropy,* Ann. of Math. (2) 180 (2014), no. 2, 773-822.

[11] W. Hoeffding, *Probability inequalities for sums of bounded random variables,* J. Amer. Statist. Assoc. 58 1963 13-30.

[12] J. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. 30 (1981), no. 5, 713-747.

[13] A. Käenmäki, *On natural invariant measures on generalised iterated function systems*, Ann. Acad. Sci. Fenn. Math. 29 (2004), no. 2, 419–458.

[14] F. Kasch, B. Volkmann, Zur Mahlerschen Vermutung über S-Zahlen, Math. Ann., 136, (1958) 442-453.

[15] A. Khintchine, *Einige Sätze über Kettenbruche, mit Anwendungen auf die Theorie der Diophantischen Approximationen,* Math. Ann. 92 (1924), 115125.

[16] J. F. Koksma, *Über die Mahlersche Klasseneinteilung der transzendenten Zahlen und die Approximation komplexer Zahlen durch algebraische Zahlen,* Monatsh. Math. Phys. 48, (1939). 176-189.

[17] J. Levesley, C. Salp, S. Velani, *On a problem of K. Mahler: Diophantine approximation and Cantor sets,* Math. Ann. 338 (2007), no. 1, 97-118.

[18] Y. Peres, B. Solomyak, *Problems on self-similar sets and self-affine sets: an update,* Fractal geometry and stochastics, II (Greifswald/Koserow, 1998), volume 46 of Progr. Probab., pages 95-106. Birkhuser, Basel, 2000.

[19] D. Ruelle, *Repellers for real analytic maps,* Ergodic Theory Dynamical Systems 2 (1982), no. 1, 99-107.

[20] V. G. Sprindžuk, *A proof of Mahlers conjecture on the measure of the set of S-numbers,* Izv. Akad. Nauk SSSR Ser. Mat. 29 1965 379-436.

[21] V. G. Sprindžuk, *Mahlers Problem in Metric Number Theory,* Izdat. Nauka i Tehnika, Minsk, 1967, 181 pp.

[22] V. G. Sprindžuk, *Metric theory of Diophantine approximation* V. H. Winston & Sons, 1979.