ON THE HAUSDORFF MEASURE OF SHRINKING TARGET SETS ON SELF-CONFORMAL SETS

DEMI ALLEN AND BALÁZS BÁRÁNY

ABSTRACT. In this article, we study the Hausdorff measure of shrinking target sets on self-conformal sets. The Hausdorff dimension of the sets we are interested in here was established by Hill and Velani in 1995. However, until recently, little more was known about the Hausdorff measure of these particular sets. In this paper we provide a complete characterisation of the Hausdorff measure of these sets, obtaining a dichotomy for the Hausdorff measure which is determined by the convergence or divergence of a sum depending on the radii of our "shrinking targets". Our main result complements earlier work of Levesley, Salp, and Velani (2007), and recent work of Baker (2019).

1. INTRODUCTION

The central theme of Diophantine approximation is understanding how well real numbers can be approximated by rationals and various generalisations thereof. Given an approximating function $\psi : \mathbb{N} \to \mathbb{R}^+ := [0, \infty)$, the classical set of ψ-*well-approximable points* is defined as the points $x \in [0, 1]$ such that $|x - p/q| < \psi(q)$ for infinitely many $(p, q) \in \mathbb{Z} \times \mathbb{N}$. In *metric* Diophantine approximation, there is a particular emphasis on understanding the "size" of such sets, in particular, the Lebesgue and Hausdorff measures. The most classical results in this direction can be traced back to Khintchine [17, 18] and Jarník [16] from the 1920s and 1930s who proved beautiful zero-full dichotomies for, respectively, the Lebesgue and Hausdorff measures of the ψ-well-approximable points in $\mathbb{R}^t$. Throughout we will denote the Hausdorff s-measure of F by $\mathcal{H}^s(F)$, and we will denote the Hausdorff dimension of F by $\dim_{\mathrm{H}} F$. For the definitions and properties of Hausdorff measure and dimension we refer the reader to [11].

In 1995, Hill and Velani [13] introduced a natural analogue of the classical ψ-well-approximable points from Diophantine approximation to the theory of dynamical systems, these are the so-called *shrinking target sets*. The basic idea is that we have a point, say $y \in X$ where X is a metric space, and a decreasing sequence $(r_n)_{n \in \mathbb{N}}$ of positive real numbers. The balls $B(y, r_n)$ are our *targets*. Given a map $T : X \to X$, we are interested in the set of points $x \in X$ for which $T^n x$ hits the target $B(y, r_n)$ infinitely often. Since their initial introduction, many more authors have contributed to the study of *shrinking targets*, *moving targets*, and *recurrence sets*. To name but a few, see [3, 13, 14, 19, 20, 31] and references within.

In this paper, we are particularly interested in the shrinking target problem on self-conformal sets. Fix $t \in \mathbb{N}$ and let $D \subset \mathbb{R}^t$ be a simply connected compact set such that $D = \overline{D^o}$. Here we denote the interior of a set $F \subset \mathbb{R}^t$ by F^o and we denote its closure by $\overline{F}$. We say that a map $f \colon D \mapsto D$ is contracting if there exists a constant $0 < c < 1$ such that

$$\|f(x) - f(y)\| \leqslant c\|x - y\|$$

for every $x, y \in D$. Throughout, $\|\cdot\| \colon \mathbb{R}^t \to \mathbb{R}^+$ will be any fixed norm on $\mathbb{R}^t$, and $B(x, r)$ denotes an open ball in $\mathbb{R}^t$, with respect to the fixed norm $\|\cdot\|$, centered at x with radius r.

Let Λ be a finite set of symbols and let $\Phi = \{f_i \colon D \mapsto D\}_{i \in \Lambda}$ be a finite set of contracting mappings. We call Φ an *iterated function system* (IFS). Hutchinson [15] showed that there exists a unique non-empty compact set X such that

$$X = \bigcup_{i \in \Lambda} f_i(X).$$

We call X the *attractor* of Φ. Throughout the paper we will assume that the maps $\{f_i\}_{i \in \Lambda}$ are $C^{1+\varepsilon}$-conformal regular strict contractions, i.e. the maps are $C^{1+\varepsilon}$-conformal and there exist constants $0 < a_{\min} \leqslant a_{\max} < 1$ such that

$$a_{\min} \leqslant \min_{x \in D} \|f_i'(x)\| \leqslant \max_{x \in D} \|f_i'(x)\| \leqslant a_{\max} \tag{1.1}$$

for every $i \in \Lambda$, where $\|f_i'(x)\|$ denotes the operator norm of the linear map $y \to f_i'(x)y$. We say that a mapping $f \colon D \mapsto D$ is $C^{1+\varepsilon}$-conformal if for every $x \in D$ and $y \in \mathbb{R}^t$,

$$\|f'(x)y\| = \|f'(x)\|\|y\| \text{ and } \|f'(x) - f'(y)\| \leqslant C\|x - y\|^\varepsilon \text{ for all } x, y \in D.$$

Note that the second condition above just says that the mapping $x \to f'(x)$ is Hölder continuous with exponent $\varepsilon > 0$. If the iterated function system Φ consists of $C^{1+\varepsilon}$-conformal mappings, we refer to Φ as a *conformal iterated function system* and we call the attractor of the conformal IFS Φ a *self-conformal set*.

In the special case when the IFS consists of similarities, i.e. there exist constants $0 < a_i < 1$ for each $i \in \Lambda$ such that for all $x, y \in D$ we have

$$\|f_i(x) - f_i(y)\| = a_i\|x - y\|,$$

we say that Φ is a *self-similar iterated function system* and that the attractor is a *self-similar set*. Clearly, all self-similar sets are self-conformal.

Let us denote by Λ^* the set of all finite sequences formed of the symbols in Λ, i.e. $\Lambda^* = \bigcup_{n=0}^\infty \Lambda^n$. We will denote the length of $\mathbf{i} \in \Lambda^*$ by $|\mathbf{i}|$. We will also use the following shorthand notation:

$$f_\mathbf{i} = f_{i_1} \circ f_{i_2} \circ \cdots \circ f_{i_{|\mathbf{i}|}}, \quad \text{and} \quad X_\mathbf{i} = f_\mathbf{i}(X).$$

We will adopt the convention that $f_\varnothing(x) = x$. We observe that by combining (1.1) with the chain rule, we also have for any $\mathbf{i} \in \Lambda^*$ that

$$a_{\min}^{|\mathbf{i}|} \leqslant \min_{x \in D} \|f_\mathbf{i}'(x)\| \leqslant \max_{x \in D} \|f_\mathbf{i}'(x)\| \leqslant a_{\max}^{|\mathbf{i}|}. \tag{1.2}$$

The dimension theory of shrinking target sets is well understood in the case of conformal iterated function systems which satisfy the open set condition. We say that Φ satisfies the *open set condition* (OSC) if there exists an open set $U \subset D$ such that

$$f_i(U) \subset U \text{ for every } i \in \Lambda, \text{ and } f_i(U) \cap f_j(U) = \varnothing \text{ for all } i \neq j \in \Lambda. \tag{1.3}$$

Given an approximating function $\Psi \colon \Lambda^* \mapsto \mathbb{R}^+$, for each $x \in X$ we will be interested in the set

$$W(x, \Psi) := \{y \in X : \|y - f_{\mathbf{i}}(x)\| < \Psi(\mathbf{i}) \text{ for infinitely many } \mathbf{i} \in \Lambda^*\}. \qquad (1.4)$$

We note that although the set $W(x, \Psi)$ is defined using strict inequalities, the geometric measure properties (in particular, Hausdorff measure and dimension) of the set would be unchanged if we instead defined the set using non-strict inequalities.

The Hausdorff dimension of $W(x, \Psi)$ was established by Hill and Velani in [13] for $\Psi : \Lambda^* \to \mathbb{R}^+$ of the form

$$\Psi(\mathbf{i}) = \operatorname{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|) \qquad (1.5)$$

where $\psi : \mathbb{N} \to \mathbb{R}^+$ is any monotonic decreasing approximating function (see Theorem 2.1 for a precise statement). Note that under the OSC, the maps f_i can be interpreted as the local inverses of some expanding map T on D, and so $W(x, \Psi)$ with Ψ of the form (1.5) is a proper generalisation of the shrinking target sets of expanding dynamics to iterated function systems.

In recent years, several other authors have also studied the problem of shrinking targets on fractals. For example, Chernov and Kleinbock [9] studied the measure of shrinking target sets with respect to ergodic measures, Chang, Wu and Wu [8] very recently studied the problem of recurrence sets on linear iterated iterated function systems consisting of maps with equal contraction ratios, Koivusalo and Ramírez [25] considered shrinking targets on self-affine sets, the second author and Rams computed the Hausdorff dimension for certain shrinking targets on Bedford-McMullen carpets [4], and Seuret and Wang considered some related problems in the setting of conformal iterated function systems [32]. However, the value of the $\dim_{\mathrm{H}}(W(x, \Psi))$-dimensional Hausdorff measure remained unknown except in some very special cases. The main result of the present paper is the following theorem which provides a complete characterisation of the Hausdorff s-measure of the set $W(x, \Psi)$ when Ψ takes the form given in (1.5).

Theorem 1.1 (Main result). *Let Φ be a conformal iterated function system which satisfies the open set condition and has attractor X. Let $x \in X$ and let $\psi \colon \mathbb{N} \mapsto \mathbb{R}^+$ be a monotonic decreasing function. Let $\Psi(\mathbf{i}) = \operatorname{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|)$ and let $W(x, \Psi)$ be as in (1.4). Then, for any open ball $B \subset \mathbb{R}^t$,*

$$\mathcal{H}^s(W(x, \Psi) \cap B) = \begin{cases} 0 & \text{if } \sum_{\mathbf{i} \in \Lambda^*} \Psi(\mathbf{i})^s < \infty, \\[2mm] \mathcal{H}^s(X \cap B) & \text{if } \sum_{\mathbf{i} \in \Lambda^*} \Psi(\mathbf{i})^s = \infty. \end{cases}$$

One of the previous examples where such complete characterisation was known is the work of Levesley, Salp and Velani [26], which was the original motivation of our paper. In [26], Levesley, Salp and Velani considered points in general *missing digit sets* which are ψ-well-approximable by rationals with appropriate denominators. A classical example of a missing digit set is the middle-third Cantor set and the work of Levesley, Salp, and Velani is just one of many works which aims to address the following question of Mahler posed in 1984 in his famous note "Some Suggestions for Further Research" [29]:

> *How close can irrational elements of Cantor's set be approximated by rational numbers*
> *(i) in Cantor's set, and*
> *(ii) by rational numbers not in Cantor's set?*

Suppose $b \geqslant 3$ is an integer and $J(b)$ is any proper subset of the set $S(b) := \{0, 1, \ldots, b-1\}$ with $\#J(b) \geqslant 2$. Denote by $K_{J(b)}$ the missing digit set consisting of the numbers $x \in [0, 1]$ for which there exists a base-b expansion consisting only of digits from $J(b)$. For example, the middle-third Cantor set is an example of a missing digit set corresponding to taking $b = 3$ and $J(b) = \{0, 2\}$. It can easily be verified that the Hausdorff dimension of $K_{J(b)}$ is $\gamma^* := \dim_{\mathrm{H}} K_{J(b)} = \frac{\log \#J(b)}{\log b}$.

Now, let $\mathcal{A}(b) = \{b^n : n = 0, 1, 2, \ldots\}$ and let $\phi : \mathbb{N} \to \mathbb{R}^+$ be an approximating function. Then let

$$W_{\mathcal{A}(b)}(\phi) := \left\{ x \in [0, 1] : \left| x - \frac{p}{q} \right| < \phi(q) \text{ for infinitely many } (p, q) \in \mathbb{Z} \times \mathcal{A}(b) \right\}.$$

It is easy to see that $W_{\mathcal{A}(b)}(\phi) \cap K_{J(b)}$ corresponds to the shrinking target set $W(0, \Psi) \cup W(1, \Psi)$ with $\psi(n) = b^n \phi(b^n)$. In [26], Levesley, Salp and Velani state the following result.[1]

Theorem 1.2 (Levesley – Salp – Velani, [26])**.** *Let $s \geqslant 0$. Then,*

$$\mathcal{H}^s(W_{\mathcal{A}(b)}(\phi) \cap K_{J(b)}) = \begin{cases} 0 & \text{if} \quad \sum_{n=1}^{\infty} \phi(b^n)^s \times (b^n)^{\gamma^*} < \infty, \\[2mm] \mathcal{H}^s(K_{J(b)}) & \text{if} \quad \sum_{n=1}^{\infty} \phi(b^n)^s \times (b^n)^{\gamma^*} = \infty. \end{cases}$$

The Mass Transference Principle due to Beresnevich and Velani [6] plays a crucial role in [26] in the proof of Theorem 1.2 to show the divergence part. However, interestingly, the Mass Transference Principle cannot be applied in the generality of our setting. More precisely, even in the self-similar case when the contraction ratios are inhomogeneous (that is there exist maps f_i, f_j with strictly different contraction ratios in absolute value) the conditions of the Mass Transference Principle are violated, see Section 3. That said, although it is impossible for us to apply the Mass Transference Principle in our general setting, it is worth remarking that the overarching strategy we use for proving Theorem 1.1 is still inspired by some of the underlying ideas in the proof of the Mass Transference Principle in [6].

More recently, complementary to the Hausdorff dimension result of Hill and Velani [13], Baker has made the first steps towards understanding the Hausdorff measure of the set $W(x, \Psi)$ on self-conformal sets.

Theorem 1.3 (Baker, [1])**.** *Let Φ be a conformal iterated function system which satisfies the open set condition and has attractor X. Given $x \in X$ and $\Psi : \Lambda^* \to \mathbb{R}^+$, let $W(x, \Psi)$ be defined as in* (1.4)*.*

(i) Let $s > 0$ and suppose that

$$\sum_{\mathbf{i} \in \Lambda^*} \Psi(\mathbf{i})^s < \infty.$$

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

(ii) Let $\psi : \mathbb{N} \to \mathbb{R}^+$ be a monotonic decreasing approximating function and suppose that $\Psi(\mathbf{i}) = \mathrm{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|)$. Then, if

$$\sum_{\mathbf{i} \in \Lambda^*} \Psi(\mathbf{i})^{\dim_{\mathrm{H}}(X)} = \infty$$

we have

$$\mathcal{H}^s(W(x, \Psi^{\frac{\dim_{\mathrm{H}} X}{s}})) = \mathcal{H}^s(X)$$

---

[1]In [26, Theorem 4] the result is stated for general gauge functions f whereas here for simplicity we opt to state the result only in terms of Hausdorff s-measure.

for all $x \in X$ and $s \leqslant \dim_{\mathrm{H}} X$. Here, $\Psi^t(\mathbf{i}) = \Psi(\mathbf{i})^t = (\mathrm{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|))^t$.

We note that Theorem 1.3 (i) holds for any iterated function systems, not just those which are conformal and satisfy the open set condition. More recently, Baker has extended Theorem 1.3 to also deal with overlapping iterated function systems in [2]. To prove Theorem 1.3 (ii) Baker establishes a Jarník-type zero-full dichotomy for the Hausdorff $\dim_{\mathrm{H}} X$-measure and, in the case $s < \dim_{\mathrm{H}} X$, Baker applies the Mass Transference Principle. It is important to note here the main difference between our main result (Theorem 1.1) and Baker's result. In particular, in order to study the Hausdorff-s measure of the shrinking target set in the divergence case, Baker needs to modify the radii of the targets, that is, he considers $\Psi(\mathbf{i})^t = (\mathrm{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|))^t$, while we can preserve them in the original form. Hence, our result does not follow directly from Baker's. In particular, we will show that in the case of self-similar iterated function systems with non-uniform contraction ratios that for any $\Psi(\mathbf{i}) = \mathrm{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|)$ with $\sum_{\mathbf{i}\in\Lambda^*} \Psi(\mathbf{i})^s = \infty$ there is no Ψ' of the form $\Psi'(\mathbf{i}) = \mathrm{diam}(X_{\mathbf{i}})\psi'(|\mathbf{i}|)$ such that $\sum_{\mathbf{i}\in\Lambda^*} \Psi'(\mathbf{i})^{\dim_{\mathrm{H}} X} = \infty$ and $W(x, (\Psi')^{\frac{\dim_{\mathrm{H}} X}{s}}) \subseteq W(x, \Psi)$, see Section 3.

1.1. **Structure of the paper.** In the next section, we introduce some further preliminaries and in Section 3 we provide a more detailed discussion of why we are unable to use the Mass Transference Principle directly in the present setting, and why our result (Theorem 1.1) does not follow from the result due to Baker (Theorem 1.3).

We present the proof of Theorem 1.1 in Sections 4–7. The key ingredient of the proof of Theorem 1.1 is Proposition 4.1. In Section 4, we provide a proof of Theorem 1.1 assuming Proposition 4.1. In Sections 5–7 we establish Proposition 4.1.

As hinted at previously, the ideas underlying our argument share some similarities with those present in the proof of the Mass Transference Principle [6]. Namely, Proposition 4.1 relies on the careful construction of a suitable Cantor-type set and a measure supported on this set satisfying certain conditions which enable us to use a version of the mass distribution principle. The existence of such a Cantor set together with an appropriate measure supported on this set is guaranteed by Proposition 5.2. In fact, establishing Proposition 5.2 is arguably the most substantial part of the proof of Theorem 1.1.

In Section 5, we describe the set $W(x, \Psi)$ in the language of symbolic dynamics and provide a proof of Proposition 4.1 subject to Proposition 5.2. In Section 6 we describe the required mass distribution, proving several technical lemmas along the way, before completing the proof of Proposition 5.2 in Section 7 by showing that the mass distribution we have constructed satisfies the requirements of Proposition 5.2. This completes the proof of Theorem 1.1.

Finally, in Section 8 we give an application of our main theorem to approximating badly approximable numbers by quadratic irrationals.

2. Preliminaries and Notations

2.1. **Conformal iterated function systems.** Throughout the paper, we will always assume that the iterated function system Φ consists of $C^{1+\varepsilon}$-conformal mappings, which map the simply connected compact set $D \subset \mathbb{R}^t$ with $D = \overline{D^o}$ into itself. If $f \colon D \mapsto D$ is a conformal map then by the simply connectedness of D, for every $x, y \in D$ there exist $\xi, \xi' \in D$ such that

$$\|f'(\xi)\|\|x - y\| \leqslant \|f(x) - f(y)\| \leqslant \|f'(\xi')\|\|x - y\|. \tag{2.1}$$

Notice that (2.1) is just a statement of the mean value theorem.

Since we insist that it consists of $C^{1+\varepsilon}$-conformal mappings, we have that the IFS Φ satisfies the *bounded distortion property*. That is, there exists a constant $C \geqslant 1$ such that for every $\mathbf{i} \in \Lambda^*$ and for all $x, y \in X$ we have

$$C^{-1} \leqslant \frac{\|f_{\mathbf{i}}'(x)\|}{\|f_{\mathbf{i}}'(y)\|} \leqslant C. \tag{2.2}$$

For a proof of this fact we refer the reader to work of Simon, Solomyak and Urbański [33, Lemma 5.8].

Combining (2.1) with the bounded distortion property we see that, for every $Y \subseteq D$, $\xi \in X$ and $\mathbf{i} \in \Lambda^*$,

$$C^{-1}\mathrm{diam}(f_{\mathbf{i}}(Y)) \leqslant \|f_{\mathbf{i}}'(\xi)\|\mathrm{diam}(Y) \leqslant C\mathrm{diam}(f_{\mathbf{i}}(Y)), \tag{2.3}$$

where C is the constant appearing in (2.2).

The *pressure function* $P\colon \mathbb{R} \mapsto \mathbb{R}$ corresponding to the IFS Φ is defined as follows,

$$P(s) = \lim_{n \to \infty} \frac{1}{n} \log \sum_{\mathbf{i} \in \Lambda^n} \|f_{\mathbf{i}}'(x)\|^s. \tag{2.4}$$

Note that by the bounded distortion property (2.2), the function P is independent of the choice of $x \in X$. It is easy to see that P is strictly monotonically decreasing, convex and continuous.

Peres, Rams, Simon and Solomyak [30, Theorem 1.1] showed that Φ satisfies the open set condition if and only if

$$0 < \mathcal{H}^d(X) < \infty, \tag{2.5}$$

where d is the unique solution of the equation $P(d) = 0$. In this case $d = \dim_{\mathrm{H}} X$. This result was later generalized by Käenmäki and Rossi [21, Proposition 3.5].

Let $\psi\colon \mathbb{N} \mapsto \mathbb{R}^+$ be a monotonically decreasing function and let us define the *shrinking rate* of the function ψ as

$$\alpha = \alpha(\psi) = \liminf_{n \to \infty} -\frac{\log \psi(n)}{n}.$$

Note that α can be $+\infty$. The following theorem can be deduced from [13, Theorem 7].

Theorem 2.1 (Hill – Velani, [13])**.** *Let Φ be a conformal iterated function system which satisfies the open set condition and has attractor X. Let $x \in X$ and let $\psi\colon \mathbb{N} \mapsto \mathbb{R}^+$ be a monotonic decreasing function. Denote by $\alpha := \alpha(\psi)$ the shrinking rate of ψ. Then, for the approximating function $\Psi(\mathbf{i}) = \mathrm{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|)$ and the set $W(x, \Psi)$, we have*

$$\dim_{\mathrm{H}}(W(x, \Psi)) = \begin{cases} 0 & \text{if } \alpha = \infty, \\ s & \text{if } 0 \leqslant \alpha < \infty, \end{cases}$$

where s is the unique root of the equation $P(s) = s\alpha$.

Note that the equation $P(s) = s\alpha$ has always a unique solution $s \in [0, d]$ for every $\alpha \geqslant 0$, since the map $s \mapsto P(s) - s\alpha$ is strictly monotonically decreasing, $P(0) = \log(\#\Lambda)$ and $P(d) - d\alpha = -d\alpha \leqslant 0$.

2.2. **Symbolic approach.**
Throughout, let $\Sigma = \Lambda^{\mathbb{N}}$ and let $\sigma\colon \Sigma \mapsto \Sigma$ denote the usual left-shift operator on Σ; namely, for $\mathbf{i} = (i_1, i_2, i_3, i_4, \dots) \in \Sigma$,

$$\sigma\mathbf{i} = \sigma(i_1, i_2, i_3, i_4, \dots) = (i_2, i_3, i_4, \dots).$$

For $\mathbf{i} = (i_1, \dots, i_k) \in \Lambda^*$, the cylinder set $[\mathbf{i}]$ is defined as

$$[\mathbf{i}] = [i_1, \dots, i_k] := \{\mathbf{j} \in \Sigma : i_1 = j_1, \dots, i_k = j_k\}.$$

By convention, $[\varnothing] = \Sigma$. For a sequence $\mathbf{i} \in \Sigma$ and $n, m \in \mathbb{N}$ with $n \leqslant m$, let

$$\mathbf{i}|_n^m := (i_n, \ldots, i_m).$$

For $n > m$ we define $\mathbf{i}|_n^m = \varnothing$. Next, for $\mathbf{i}, \mathbf{j} \in \Sigma$ let $|\mathbf{i} \wedge \mathbf{j}| = \min\{k \geqslant 1 : i_k \neq j_k\} - 1$ and let $\mathbf{i} \wedge \mathbf{j} = i_1, \ldots, i_{|\mathbf{i} \wedge \mathbf{j}|}$ be the common part of $\mathbf{i}$ and $\mathbf{j}$. If $|\mathbf{i} \wedge \mathbf{j}| = 0$ then we define $\mathbf{i} \wedge \mathbf{j}$ as the empty word.

For any $\alpha \in [0, \infty)$, there exists a unique ergodic σ-invariant probability measure, $\mathbb{P}$, and a constant $C \geqslant 1$ such that for every $\mathbf{i} \in \Lambda^*$ and $x \in X$

$$C^{-1} \leqslant \frac{\mathbb{P}([\mathbf{i}])}{e^{-\alpha s|\mathbf{i}|} \|f_{\mathbf{i}}'(x)\|^s} \leqslant C, \tag{2.6}$$

where s is the unique solution of the equation $P(s) = s\alpha$, see for example [7, Theorem 1.2].

The elements of Σ and X can be associated in a natural way. More precisely, for every $\mathbf{i} \in \Sigma$ let

$$\pi(\mathbf{i}) = \lim_{n \to \infty} f_{i_1} \circ \cdots \circ f_{i_n}(0).$$

We call the function $\pi \colon \Sigma \mapsto X$ the *natural projection*. It is easy to see that $\pi(\mathbf{i}) = f_{i_1}(\pi(\sigma\mathbf{i}))$. In particular, for any $n \in \mathbb{N}$, we have

$$\pi(\mathbf{i}) = f_{\mathbf{i}|_1^n}(\pi(\sigma^n \mathbf{i})). \tag{2.7}$$

For a probability measure μ on Σ, we will denote by $\pi_* \mu$ the *pushforward* of the measure μ; that is, for a set $A \subseteq X$ we have $\pi_* \mu(A) = \mu(\pi^{-1}(A))$.

3. A Limitation of the Mass Transference Principle

Before we prove the main result of this paper, we show two phenomena. First, we show that the Mass Transference Principle, originally introduced by Beresnevich and Velani in [6], is not applicable if the underlying iterated function system consists of similarities with different contraction ratios, by showing that the assumptions of the Mass Transference Principle are violated in this case. Secondly, we show that Theorem 1.1 cannot be deduced from Baker's result Theorem 1.3 [1].

More precisely, in this section, we consider iterated function systems of the form

$$\Phi = \{f_i \colon x \in \mathbb{R}^t \mapsto a_i O_i x + b_i\}_{i \in \Lambda}, \tag{3.1}$$

where $a_i \in (0, 1)$, $b_i \in \mathbb{R}^t$, and O_i is a rotation. We will be particularly interested in the case when there exist $i, j \in \Lambda$ such that $a_i \neq a_j$. We note that for iterated function systems of the form in (3.1), the pressure equation from (2.4) simplifies to

$$P(s) = \log\left(\sum_{i \in \Lambda} a_i^s\right). \tag{3.2}$$

In this case, we also note that the value of s appearing in the Hausdorff dimension result due to Hill and Velani (Theorem 2.1) is the unique solution of $\sum_{i \in \Lambda} a_i^s e^{-s\alpha} = 1$.

Before proceeding, we state a corresponding version of the Mass Transference Principle which is the most relevant to our current setting. The statement we give below, Theorem 3.1, can be deduced from [6, Theorem 3], which was the result used by Levesley, Salp, and Velani in [26] to study Diophantine approximation on the middle-third Cantor set. Indeed, since its initial discovery, the Mass Transference Principle has become a widely used tool with profound consequences Diophantine approximation.

Theorem 3.1 (Beresnevich – Velani, [6])**.** *Let Φ be a conformal iterated function system of the form* (3.1) *which satisfies the open set condition and has attractor X. Let us write $d = \dim_H X$ and let $s \in (0, d]$ be arbitrary. Let $\psi \colon \mathbb{N} \mapsto \mathbb{R}^+$ be an arbitrary bounded map. For $\Psi(\mathbf{i}) = \mathrm{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|)$ and $x \in X$, denote by $W(x, \Psi)$ the set defined in* (1.4)*. If, for any ball $B \subset \mathbb{R}^t$,*

$$\mathcal{H}^d(B \cap W(x, \Psi^{s/d})) = \mathcal{H}^d(B \cap X)$$

then, for any ball $B \subset \mathbb{R}^t$,

$$\mathcal{H}^s(B \cap W(x, \Psi)) = \mathcal{H}^s(B \cap X).$$

It follows from the result of Hill and Velani (Theorem 2.1) that $\dim_H W(x, \Psi)$ is the unique solution of the equation $P(s) = \alpha(\psi)s$, where $\alpha(\psi)$ is the shrinking rate of ψ. Hence, the case when the value of the Hausdorff measure is in question is exactly at this choice of s. We show that if $0 < \alpha(\psi) < \infty$ and there exist $i, j \in \Lambda$ such that $a_i \neq a_j$ then the d-dimensional Hausdorff measure of the limsup set $W(x, \Psi^{s/d})$ would be zero at this critical choice of s. In such cases the Theorem 3.1 is not useful.

For simplicity, we will work with iterated function systems $\Phi = \{f_i\}_{i \in \Lambda}$ satisfying the *strong separation condition* (SSC). That is, we assume that

$$f_i(X) \cap f_j(X) = \varnothing \text{ for every } i \neq j \in \Lambda,$$

where X is the attractor of Φ. Clearly any IFS satisfying the SSC also satisfies the OSC.

Proposition 3.2. *Let Φ be a conformal iterated function system of the form* (3.1) *which satisfies the strong separation condition and has attractor X. Suppose that there exist $i, j \in \Lambda$ such that $a_i \neq a_j$. Let $\psi \colon \mathbb{N} \mapsto \mathbb{R}^+$ be a monotonic decreasing function with shrinking rate $\alpha(\psi) \in (0, \infty)$. Let s be the solution to $\sum_{i \in \Lambda} e^{-\alpha s} a_i^s = 1$ and write $d = \dim_H X$. In this case it is well-known that d is the unique solution of $\sum_{i \in \Lambda} a_i^d = 1$.*

Let $x \in X$ and let $\Psi(\mathbf{i}) = \mathrm{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|)$. Then,

$$\mathcal{H}^d(W(x, \Psi^{s/d})) = 0.$$

Proof. Without loss of generality, we may assume that $\mathrm{diam}(X) = 1$. Throughout the proof, let $\mathbf{x}$ be the unique coding of x, i.e. $\pi(\mathbf{x}) = x$. The uniqueness of this encoding is guaranteed by the strong separation condition.

Let us define λ to be the natural σ-invariant ergodic probability measure on Σ, whose projection is equivalent to $\mathcal{H}^d|_X$. That is, $\pi_*\lambda = \lambda \circ \pi^{-1} = \frac{\mathcal{H}^d|_X}{\mathcal{H}^d(X)}$. Moreover, for each $\mathbf{i} = (i_1, \ldots, i_n) \in \Lambda^*$,

$$\lambda([i_1, \ldots, i_n]) = (a_{\mathbf{i}})^d.$$

Here we use the notation $a_{\mathbf{i}} = a_{i_1} a_{i_2} \ldots a_{i_{|\mathbf{i}|}}$.

Let $\chi := -\sum_{i \in \Lambda} a_i^d \log a_i = -\int \log a_{i_1} d\lambda(\mathbf{i})$ and let

$$F = \left\{ \mathbf{i} \in \Sigma : \lim_{n \to \infty} \frac{-1}{n} \log a_{\mathbf{i}|_1^n} = \lim_{n \to \infty} \frac{-1}{n} \sum_{k=1}^{n} \log a_{i_k} = \chi \right\}.$$

By Birkhoff's Ergodic Theorem (see, for example, [10, Theorem 2.30] or [34, Theorem 1.14]), $\lambda(F) = 1$.

Now, let

$$\widehat{W}(x, \Psi^{s/d}) := \left\{ \mathbf{j} \in \Sigma : \|\pi(\mathbf{j}) - \pi(\mathbf{i}\mathbf{x})\| < (a_{\mathbf{i}})^{s/d}\psi(|\mathbf{i}|)^{s/d} \text{ for infinitely many } \mathbf{i} \in \Lambda^* \right\}.$$

Since Φ satisfies the strong separation condition and hence points in the symbolic space Σ uniquely encode points in X, we have that

$$\pi\widehat{W}(x, \Psi^{s/d}) = W(x, \Psi^{s/d}). \tag{3.3}$$

Thus, since $\lambda(F) = 1$, it is sufficient for us to show that $\lambda(\widehat{W}(x, \Psi^{s/d}) \cap F) = 0$. We make the following claim.

Claim 1. *If* $\mathbf{j} \in \widehat{W}(x, \Psi^{s/d}) \cap F$ *then* $|\mathbf{j} \wedge \mathbf{ix}| \geqslant |\mathbf{i}|$ *for infinitely many* $\mathbf{i} \in \Lambda^*$ *such that* $\|\pi(\mathbf{j}) - \pi(\mathbf{ix})\| < (a_\mathbf{i})^{s/d}\psi(|\mathbf{i}|)^{s/d}$. *Recall that* $\mathbf{j} \wedge \mathbf{ix}$ *is the common part of* $\mathbf{j}$ *with the concatenation* $\mathbf{ix}$.

Proof of Claim 1. Suppose to the contrary that there exists $N = N(\mathbf{j}) \geqslant 1$ such that if $|\mathbf{i}| > N$ and $\|\pi(\mathbf{j}) - \pi(\mathbf{ix})\| \leqslant (a_\mathbf{i})^{s/d}\psi(|\mathbf{i}|)^{s/d}$ then $|\mathbf{j} \wedge \mathbf{ix}| < |\mathbf{i}|$.

Let

$$\delta = \min_{i \neq j} d(f_i(X), f_j(X)) > 0$$

where, for subsets $A, B \subset \mathbb{R}^t$, $d(A, B) = \min\{\|a - b\| : a \in A \text{ and } b \in B\}$. Then from the definitions of the common part $\mathbf{i} \wedge \mathbf{j}$ and $\delta > 0$ and (2.7) we have

$$\begin{aligned}
\delta a_{\mathbf{j} \wedge \mathbf{ix}} &\leqslant a_{\mathbf{j} \wedge \mathbf{ix}}\|\pi(\sigma^{|\mathbf{j} \wedge \mathbf{ix}|}\mathbf{j}) - \pi(\sigma^{|\mathbf{j} \wedge \mathbf{ix}|}\mathbf{ix})\| \\
&= \|f_{\mathbf{j} \wedge \mathbf{ix}}(\pi(\sigma^{|\mathbf{j} \wedge \mathbf{ix}|}\mathbf{j})) - f_{\mathbf{j} \wedge \mathbf{ix}}(\pi(\sigma^{|\mathbf{j} \wedge \mathbf{ix}|}\mathbf{ix}))\| \\
&= \|\pi(\mathbf{j}) - \pi(\mathbf{ix})\| \\
&\leqslant (a_\mathbf{i})^{s/d}\psi(|\mathbf{i}|)^{s/d}.
\end{aligned}$$

The last inequality above holds by assumption.

Since $s < d$ and we are assuming that $|\mathbf{j} \wedge \mathbf{ix}| < |\mathbf{i}|$, it follows from the previous inequality that

$$\delta \leqslant (a_{\mathbf{j} \wedge \mathbf{ix}})^{s/d-1}(a_{\sigma^{|\mathbf{j} \wedge \mathbf{ix}|}\mathbf{i}})^{s/d}\psi(|\mathbf{i}|)^{s/d} \leqslant (a_{\mathbf{j} \wedge \mathbf{ix}})^{s/d-1}\psi(|\mathbf{i}|)^{s/d} \leqslant (a_{\mathbf{j}|_1^{|\mathbf{i}|}})^{s/d-1}\psi(|\mathbf{i}|)^{s/d}.$$

Thus, since $\mathbf{j} \in F$, we have

$$0 \geqslant \liminf_{n \to \infty} \frac{-1}{n} \log\left((a_{\mathbf{j}|_1^n})^{s/d-1}\psi(n)^{s/d}\right) = \left(\frac{s}{d} - 1\right)\chi + \frac{s}{d}\alpha. \tag{3.4}$$

Let

$$D_{KL}(\lambda\|\mathbb{P}) := \sum_{i \in \Lambda} \lambda(i) \log \frac{\lambda(i)}{\mathbb{P}(i)}$$

denote the *Kullback-Leibler divergence* (or *relative entropy*) of the measure λ with respect to $\mathbb{P}$, where $\mathbb{P}$ is the measure defined in (2.6). See [28, Section 2.6] for a definition of Kullback-Leibler divergence. It is a property of the Kullback-Leibler divergence that $0 \leqslant D_{KL}(\lambda\|\mathbb{P})$ and $0 = D_{KL}(\lambda\|\mathbb{P})$ if and only if $\lambda = \mathbb{P}$.

Since $\alpha > 0$, we have that $s < d$. By the assumption that there exist contraction ratios $a_i \neq a_j$, it follows that there exists $i \in \Lambda$ such that $e^{\alpha s}a_i^s \neq a_i^d$. Indeed, otherwise we would have that $a_i^{d-s} = e^{\alpha s} = a_j^{d-s}$ for every $i, j \in \Lambda$, which is impossible. Thus, $\lambda \neq \mathbb{P}$ and it follows from the definitions of the measures $\mathbb{P}$ and λ that

$$0 < D_{KL}(\lambda\|\mathbb{P}) = (s - d)\chi + s\alpha, \tag{3.5}$$

but this contradicts (3.4). □

Claim 1 combined with (2.7) implies that for every $\mathbf{j} \in \widehat{W}(x, \Psi^{s/d}) \cap F$ there are infinitely many $\mathbf{i} \in \Lambda^*$ such that $\mathbf{j}|_1^{|\mathbf{i}|} = \mathbf{i}$ and

$$(a_{\mathbf{i}})^{s/d} \psi(|\mathbf{i}|)^{s/d} > \|\pi(\mathbf{j}) - \pi(\mathbf{i}x)\| = \|f_{\mathbf{i}}(\pi(\sigma^{|\mathbf{i}|}\mathbf{j})) - f_{\mathbf{i}}(\pi(\mathbf{x}))\| = a_{\mathbf{i}}\|\pi(\sigma^{|\mathbf{i}|}\mathbf{j}) - \pi(\mathbf{x})\|.$$

Hence,

$$\widehat{W}(x, \Psi^{s/d}) \cap F \subseteq \{\mathbf{i} \in \Sigma : \|\pi(\sigma^n\mathbf{i}) - x\| < (a_{\mathbf{i}|_1^n})^{s/d-1}\psi(n)^{s/d} \text{ for infinitely many } n \in \mathbb{N}\} \cap F. \tag{3.6}$$

By the definition of F, it follows from Egorov's Theorem (see, for example, [23, Theorem 12.1]) that, for every $\varepsilon > 0$ there exists a set $E \subset F$ such that $\lambda(E) > 1 - \varepsilon$ and the sequence of functions $\mathbf{i} \mapsto \frac{-1}{n} \log a_{\mathbf{i}|_1^n}$ converges uniformly to χ on E.

In particular, there exists a natural number $N = N(E)$ such that for every $\mathbf{i} \in E$ and every $n \geqslant N$

$$a_{\mathbf{i}|_1^n}^{s/d-1}\psi(n)^{s/d} < e^{-\frac{n}{2d}D_{KL}(\lambda\|\mathbb{P})}.$$

To see this, recall the right-hand equalities of (3.4) and (3.5). Thus,

$$\widehat{W}(x, \Psi^{s/d}) \cap E \subseteq W'' := \{\mathbf{j} \in \Sigma : \|\pi(\sigma^n\mathbf{j}) - x\| < e^{-\frac{n}{2d}D_{KL}(\lambda\|\mathbb{P})} \text{ for infinitely many } n \in \mathbb{N}\}$$

$$= \{\mathbf{j} \in \Sigma : \|\pi(\mathbf{j}) - f_{\mathbf{i}}(x)\| < a_{\mathbf{i}}e^{-\frac{|\mathbf{i}|}{2d}D_{KL}(\lambda\|\mathbb{P})} \text{ for infinitely many } \mathbf{i} \in \Lambda^*\}.$$

Since

$$\sum_{\mathbf{i}\in\Lambda^*} (a_{\mathbf{i}}e^{-\frac{|\mathbf{i}|}{2d}D_{KL}(\lambda\|\mathbb{P})})^d = \sum_{n=1}^{\infty} e^{-nD_{KL}(\lambda\|\mathbb{P})/2} \left(\sum_{i\in\Lambda} a_i^d\right)^n = \sum_{n=1}^{\infty} e^{-nD_{KL}(\lambda\|\mathbb{P})/2} < \infty,$$

we have that $\lambda(W'') = \mathcal{H}^d(\pi(W'')) = 0$ by Theorem 1.3. Hence, by (3.3) and the fact that $\lambda(F) = 1$ we have

$$\begin{aligned}
\mathcal{H}^d(W(x, \Psi^{s/d})) &= \mathcal{H}^d(X) \cdot \lambda(\widehat{W}(x, \Psi^{s/d})) \\
&= \mathcal{H}^d(X) \cdot \lambda(\widehat{W}(x, \Psi^{s/d}) \cap F) \\
&\leqslant \mathcal{H}^d(X) \cdot \left(\lambda(W'') + \lambda(F\backslash E)\right) \\
&< \mathcal{H}^d(X) \cdot \varepsilon.
\end{aligned}$$

Finally, since $\varepsilon > 0$ was arbitrary and $\mathcal{H}^d(X)$ is finite by (2.5), the statement follows. $\square$

Although Baker's result Theorem 1.3(ii) relies on the Mass Transference Principle, *a priori* it might happen that for a given function ψ and $s > 0$ with $\sum_{\mathbf{i}\in\Lambda^*} \Psi(\mathbf{i})^s = \infty$ one could construct another function $\overline{\psi}$ for which $\sum_{\mathbf{i}\in\Lambda^*} \overline{\Psi}(\mathbf{i})^d = \infty$ implies $\mathcal{H}^s(W(x, \Psi)) = \infty$ for any $x \in X$. One way to show this for example would be to show the containment $W(x, (\overline{\Psi})^{d/s}) \subseteq W(x, \Psi)$ and apply Theorem 1.3(ii). However, the next proposition shows that such a containment is not possible in some cases.

Proposition 3.3. *Let Φ be a conformal iterated function system of the form* (3.1) *which satisfies the strong separation condition and has attractor X. Write $d = \dim_{\mathrm{H}} X$. Suppose that there exist $i, j \in \Lambda$ such that $a_i \neq a_j$. Let $\psi \colon \mathbb{N} \mapsto \mathbb{R}^+$ be a monotonic decreasing function such that the shrinking rate is achieved as a limit, i.e.*

$$0 < \alpha(\psi) = \lim_{n\to\infty} \frac{-\log\psi(n)}{n} < d.$$

Moreover, suppose that $\sum_{\mathbf{i}\in\Lambda^*} \Psi(\mathbf{i})^s = \infty$, where s is the unique solution of the equation $\sum_{i\in\Lambda} e^{-\alpha s}a_i^s = 1$ and $\Psi(\mathbf{i}) = \mathrm{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|)$. Then there exists $x \in X$ such that there is

no function $\overline{\Psi}\colon \Lambda^* \mapsto \mathbb{R}^+$ *of the form* $\overline{\Psi}(\mathbf{i}) = \operatorname{diam}(X_{\mathbf{i}})\overline{\psi}(|\mathbf{i}|)$, *where* $\overline{\psi}\colon \mathbb{N} \mapsto \mathbb{R}^+$ *is such that* $\alpha(\overline{\psi}) = \liminf_{n\to\infty} \frac{-\log\overline{\psi}(n)}{n} \geqslant 0$ *and* $\sum_{\mathbf{i}\in\Lambda^*} \overline{\Psi}(\mathbf{i})^d = \infty$ *and* $W(x, (\overline{\Psi})^{d/s}) \subseteq W(x, \Psi)$.

Observe that such a function ψ exists, for example take $\psi(n) = e^{-\alpha n}$.

Proof. Without loss of generality, we may assume that $\operatorname{diam}(X) = 1$, $a_1 = \max_{i\in\Lambda} a_i$. Now, let $\mathbf{x} := (2, 2, 2, \dots)$ and $x := \pi(\mathbf{x})$. In order to reach a contradiction, we suppose that there does exist $\overline{\Psi}\colon \Lambda^* \mapsto \mathbb{R}^+$ of the form $\overline{\Psi}(\mathbf{i}) = \operatorname{diam}(X_{\mathbf{i}})\overline{\psi}(|\mathbf{i}|)$, where $\overline{\psi}\colon \mathbb{N} \mapsto \mathbb{R}_+$ is such that $\alpha(\overline{\psi}) = \liminf_{n\to\infty} \frac{-\log\overline{\psi}(n)}{n} \geqslant 0$ and $\sum_{\mathbf{i}\in\Lambda^*} \overline{\Psi}(\mathbf{i})^d = \infty$ and $W(x, (\overline{\Psi})^{d/s}) \subseteq W(x, \Psi)$.

Since $\infty = \sum_{\mathbf{i}\in\Lambda^*} \overline{\Psi}(\mathbf{i})^d = \sum_{n=0}^{\infty} \overline{\psi}(n)^d$ and $\alpha(\overline{\psi}) \leqslant 0$ we get $\alpha(\overline{\psi}) = 0$. Let n_k be a sequence along which $\lim_{k\to\infty} \frac{-\log\overline{\psi}(n_k)}{n_k} = 0$. Let

$$\ell_k = \left\lceil \left(\frac{d}{s} - 1\right)\left(\left(n_k - \sum_{i=1}^{k-1} \ell_i\right)\frac{\log a_1}{\log a_2} + \sum_{i=1}^{k-1} \ell_i\right) + \frac{d\log\overline{\psi}(n_k)}{s\log a_2} \right\rceil. \tag{3.7}$$

By taking a sufficiently fast growing subsequence of n_k and replacing n_k with it, we can assume without loss of generality that $n_k + \ell_k < n_{k+1}$ and $\frac{\sum_{i=1}^{k-1}\ell_i}{n_k} \to 0$ as $k \to \infty$.

Now, let us define $\mathbf{i} \in \Sigma$ such that

$$i_j = \begin{cases} 2 & \text{if } n_k \leqslant j \leqslant n_k + \ell_k \\ 1 & \text{if } n_k + \ell_k < j < n_{k+1} \end{cases} \quad \text{for some } k \in \mathbb{N}.$$

It is easy to see that $\pi(\mathbf{i}) \in W(\pi(\mathbf{x}), (\overline{\Psi})^{d/s})$. Indeed, by (3.7)

$$\|\pi(\sigma^{n_k}\mathbf{i}) - \pi(\mathbf{x})\| \leqslant a_2^{\ell_k} \leqslant \left(a_{\mathbf{i}|_{n_k}}\right)^{d/s-1} \overline{\psi}(n_k)^{d/s},$$

and so $\|\pi(\mathbf{i}) - f_{\mathbf{i}|_{n_k}}(\pi(\mathbf{x}))\| \leqslant \left(\overline{\Psi}(\mathbf{i})\right)^{d/s}$. We will show that $\pi(\mathbf{i}) \notin W(\pi(\mathbf{x}), \Psi)$.

By the disjointness of the cylinders, $\pi(\mathbf{i}) \in W(\pi(\mathbf{x}), \Psi)$ if and only if

$$\|\pi(\sigma^{n_k}\mathbf{i}) - \pi(\mathbf{x})\| \leqslant \psi(n_k) \text{ for infinitely many } k \in \mathbb{N}.$$

However, by the strong separation condition $\|\pi(\sigma^{n_k}\mathbf{i}) - \pi(\mathbf{x})\| \geqslant \delta a_2^{\ell_k-1}$, where $\delta = \min_{i\neq j} d(f_i(X), f_j(X))$. So $\psi(n_k) \geqslant \delta a_2^{\ell_k-1}$ for infinitely many $k \in \mathbb{N}$. Taking logarithms, dividing by n_k, and letting $k \to \infty$, we get

$$-\alpha \geqslant \left(\frac{d}{s} - 1\right)\log a_1.$$

However, by the convexity of the pressure (3.2) we have

$$P(s) \geqslant P(d) + (s - d)P'(d),$$

where $P'(d) = \sum_{i\in\Lambda} a_i^d \log a_i$. Since there exists $a_i \neq a_1$ we have that $P'(d) < \log a_1$. This implies

$$(d - s)(-\log a_1) \geqslant \alpha s = P(s) \geqslant (s - d)P'(d) > (d - s)(-\log a_1),$$

which is a contradiction. Hence, $\pi(\mathbf{i}) \notin W(\pi(\mathbf{x}), \Psi)$. $\qquad\square$

Note that the Hausdorff measure of the sets $W(x, \Psi)$ in both Proposition 3.2 and Proposition 3.3 can be calculated by using Theorem 1.1.

4. Proof of the Main Result (Theorem 1.1)

We first note that the convergence part of Theorem 1.1 is contained in greater generality in Theorem 1.3 (i). When $\sum_{\mathbf{i}\in\Lambda^*}\Psi(\mathbf{i})^s < \infty$, the proof that $\mathcal{H}^s(W(x,\Psi)) = 0$ follows from a standard covering argument combined with the definition of Hausdorff measure. For further details see the argument given in [1, §3.2]. Thus, it remains to prove the divergence part of Theorem 1.1.

An observation that is central to proving the convergence part of the result, and which we will also make use of for the divergence case, is that the set $W(x,\Psi)$ is the $\limsup$ set of the family of balls $\{B(f_{\mathbf{i}}(x),\Psi(\mathbf{i}))\}_{\mathbf{i}\in\Lambda^*}$. That is,

$$W(x,\Psi) = \bigcap_{n=0}^{\infty} \bigcup_{\substack{\mathbf{i}\in\Lambda^* \\ |\mathbf{i}|\geqslant n}} B(f_{\mathbf{i}}(x),\Psi(\mathbf{i})).$$

In proving Theorem 1.1, we first show that by using Theorems 2.1 and 1.3 the problem can be reduced to the case when $P(s) = s\alpha$ where $\alpha := \alpha(\psi)$ is the shrinking rate of ψ and P is the pressure function defined in (2.4). To tackle the proof in the remaining case, we use the following proposition.

Proposition 4.1. *Let Φ be a conformal iterated function system which satisfies the open set condition and has attractor X. Let $x \in X$ and let $\psi\colon \mathbb{N} \mapsto \mathbb{R}^+$ be a monotonic decreasing function with shrinking rate $0 < \alpha(\psi) < \infty$. Let $\Psi(\mathbf{i}) = \operatorname{diam}(X_{\mathbf{i}})\psi(|\mathbf{i}|)$ and let s be the unique solution of the equation $P(s) = s\alpha(\psi)$. For the set $W(x,\Psi)$, if*

$$\sum_{n=1}^{\infty} \psi(n)^s \sum_{\mathbf{i}\in\Lambda^n} \|f'_{\mathbf{i}}(x)\|^s = \infty,$$

then

$$\mathcal{H}^s(W(x,\Psi)) = \infty.$$

In the remainder of this section, we will give the proof of Theorem 1.1 assuming Proposition 4.1. The rest of the paper will then be devoted to establishing Proposition 4.1 and a number of other required technical lemmas. The key to establishing Proposition 4.1 is proving the existence of a suitable mass distribution as outlined in Proposition 5.2.

Proof of Theorem 1.1 (Divergence). Recall that we are given

$$\sum_{\mathbf{i}\in\Lambda^*} \Psi(\mathbf{i})^s = \infty. \tag{4.1}$$

It follows from (2.3) together with the bounded distortion property (2.2) that, for every $x \in X$, we have

$$\sum_{n=1}^{\infty} \psi(n)^s \sum_{\mathbf{i}\in\Lambda^n} \|f'_{\mathbf{i}}(x)\|^s = \infty. \tag{4.2}$$

Next, let α denote the shrinking rate of ψ. First of all, let us consider the case when $\alpha = \infty$. It follows from (1.2) that

$$\infty = \sum_{n=1}^{\infty} \psi(n)^s \sum_{\mathbf{i}\in\Lambda^n} \|f'_{\mathbf{i}}(x)\|^s \leqslant \sum_{n=1}^{\infty} \psi(n)^s (\sharp\Lambda a_{\max}^s)^n,$$

which is possible if and only if $s = 0$. Otherwise, it follows from the definition of α that the terms in the sum on the far right-hand side become too small, thus forcing the sum to converge. However, it can be seen that $W(x,\Psi)$ has continuum many elements. Indeed,

it is a countable intersection of open and dense sets and, hence, it is a dense G_δ set by Baire's category theorem, see [12, Theorem 6.54]. Thus, $\mathcal{H}^0(W(x, \Psi) \cap B) = \infty$ for every open ball B in X and so we may assume that $\alpha < \infty$.

Next, observe that, by the root test, it follows from (4.2) that

$$1 \leqslant \limsup_{n \to \infty} \sqrt[n]{\psi(n)^s \sum_{\mathbf{i} \in \Lambda^n} \|f_\mathbf{i}'(x)\|^s} = e^{-\alpha s} e^{P(s)}. \tag{4.3}$$

If $\alpha = 0$ then $P(s) \geqslant 0$. Furthermore, in this case, $\dim_{\mathrm{H}}(W(x, \Psi)) = d := \dim_{\mathrm{H}} X$ by Theorem 2.1 and (2.5), which yields that $\dim_{\mathrm{H}} X = d$ where d is the unique solution of the equation $P(d) = 0$. Since $P(s)$ is strictly monotonically decreasing, if $P(s) > 0$ then $s < d$ and, hence, $\mathcal{H}^s(W(x, \Psi)) = \infty$. If $P(s) = 0$ then $s = d = \dim_{\mathrm{H}} X$ and the statement follows from Theorem 1.3(ii) since $W(x, \Psi)$ has full $\mathcal{H}^d|_X$-measure. Thus, we may assume that $0 < \alpha < \infty$.

Note that it follows from (4.3) that $P(s) \geqslant s\alpha$. Since $s \mapsto P(s) - s\alpha$ is strictly monotonically decreasing, if $P(s) > s\alpha$ then by Theorem 2.1, $\dim_{\mathrm{H}}(W(x, \Psi)) > s$ and thus the statement follows again. So, for the remainder of the proof, suppose that $P(s) = s\alpha$.

Now, let B be an open ball such that $X \cap B \neq \varnothing$. Since the maps of Φ are uniformly contracting (this is essentially what (1.1) tells us), there exists $\mathbf{i} \in \Lambda^*$ so that $f_\mathbf{i}(X) \subseteq B$. Let $\widetilde{\Psi}(\mathbf{j}) = \mathrm{diam}(X_\mathbf{j})\widetilde{\psi}(|\mathbf{j}|)$, where $\widetilde{\psi}(n) = C^{-1}\psi(n + |\mathbf{i}|)$ and $C > 1$ is the constant in (2.2). From (4.2) and the bounds on f' given in (1.2), we have

$$\sum_{n=0}^{\infty} \widetilde{\psi}(n)^s \sum_{\mathbf{j} \in \Lambda^n} \|f_\mathbf{j}'(x)\|^s = C^{-s} \sum_{n=0}^{\infty} \psi(n + |\mathbf{i}|)^s \sum_{\mathbf{j} \in \Lambda^n} \|f_\mathbf{j}'(x)\|^s$$

$$\geqslant C^{-s}(\sharp\Lambda)^{-|\mathbf{i}|} a_{\max}^{-s|\mathbf{i}|} \sum_{n=0}^{\infty} \psi(n + |\mathbf{i}|)^s \sum_{\mathbf{j} \in \Lambda^{n+|\mathbf{i}|}} \|f_\mathbf{j}'(x)\|^s = \infty.$$

Thus, by Proposition 4.1, we have $\mathcal{H}^s(W(x, \widetilde{\Psi})) = \infty$.

Next, suppose that $\|y - f_\mathbf{j}(x)\| < \widetilde{\Psi}(\mathbf{j})$. Employing the mean value theorem (2.1) and (2.3), we see that

$$\begin{aligned}
\|f_\mathbf{i}(y) - f_{\mathbf{ij}}(x)\| &\leqslant \|f_\mathbf{i}'(\xi)\|\|y - f_\mathbf{j}(x)\| \\
&< \|f_\mathbf{i}'(\xi)\|\widetilde{\Psi}(|\mathbf{j}|) \\
&= C^{-1}\|f_\mathbf{i}'(\xi)\|\mathrm{diam}(X_\mathbf{j})\psi(|\mathbf{ij}|) \\
&\leqslant \mathrm{diam}(f_\mathbf{i}(X_\mathbf{j}))\psi(|\mathbf{ij}|) \\
&= \mathrm{diam}(X_{\mathbf{ij}})\psi(|\mathbf{ij}|).
\end{aligned}$$

Hence, if $y \in W(x, \widetilde{\Psi})$ then $f_\mathbf{i}(y) \in W(x, \Psi)$ and thus

$$\mathcal{H}^s(W(x, \Psi) \cap B) \geqslant \mathcal{H}^s(f_\mathbf{i}(W(x, \widetilde{\Psi})) \cap B) = \mathcal{H}^s(f_\mathbf{i}(W(x, \widetilde{\Psi}))).$$

The last inequality above follows since $\mathbf{i} \in \Lambda^*$ was chosen so that $f_\mathbf{i}(X) \subset B$ and so, consequently, we also have $f_\mathbf{i}(W(x, \widetilde{\Psi}) \subset B$.

Finally, it follows from (1.2) and (2.3) combined with the definition of Hausdorff s-measure that

$$\mathcal{H}^s(f_\mathbf{i}(W(x, \widetilde{\Psi}))) \geqslant a_{\min}^{|\mathbf{i}|s} \mathcal{H}^s(W(x, \widetilde{\Psi})).$$

This completes the proof. $\qquad\qquad\square$

5. Proof of Proposition 4.1: A Symbolic Approach

For the rest of the paper, we fix an $x \in X$ and a symbolic representation $\mathbf{x} \in \Sigma$ for which $\pi(\mathbf{x}) = x$. Next, let us define $\rho \colon \mathbb{N} \mapsto \mathbb{N}$ as follows; let $\rho(n)$ be the unique natural number such that

$$\operatorname{diam}(X_{\mathbf{x}|_1^{\rho(n)}}) \leqslant C^{-2}\operatorname{diam}(X)\psi(n) < \operatorname{diam}(X_{\mathbf{x}|_1^{\rho(n)-1}}). \tag{5.1}$$

Note that ρ is monotonically increasing. By combining (5.1) with (2.3) and the bounded distortion property (2.2), it can be seen that if α is the shrinking rate of ψ defined earlier, then

$$\alpha = \liminf_{n \to \infty} \frac{-\log \psi(n)}{n} = \liminf_{n \to \infty} \frac{-1}{n} \log \operatorname{diam}(X_{\mathbf{x}|_1^{\rho(n)}}) = \liminf_{n \to \infty} \frac{-1}{n} \log \|f'_{\mathbf{x}|_1^{\rho(n)}}(x)\|. \tag{5.2}$$

For a monotonic increasing function $\rho \colon \mathbb{N} \mapsto \mathbb{N}$ let

$$\widehat{W}(\mathbf{x}, \rho) = \{\mathbf{i} \in \Sigma : \sigma^n \mathbf{i} \in [\mathbf{x}|_1^{\rho(n)}] \text{ for infinitely many } n \in \mathbb{N}\}. \tag{5.3}$$

Lemma 5.1. *Let $\mathbf{x} \in \Sigma$, let $\psi \colon \mathbb{N} \mapsto \mathbb{R}^+$ be a monotonic decreasing function, and let ρ be as defined in (5.1). Then,*

$$\pi\widehat{W}(\mathbf{x}, \rho) \subseteq W(\pi(\mathbf{x}), \Psi).$$

Proof. If $\mathbf{i} \in \widehat{W}(\mathbf{x}, \rho)$, then $\sigma^n \mathbf{i} \in [\mathbf{x}|_1^{\rho(n)}]$ for infinitely many $n \in \mathbb{N}$. For each such $n \in \mathbb{N}$, we have $\pi(\sigma^n \mathbf{i}) \in \pi([\mathbf{x}|_1^{\rho(n)}]) = X_{\mathbf{x}|_1^{\rho(n)}}$. Hence, by the definition of ρ, for infinitely many $n \in \mathbb{N}$ we have

$$\|\pi(\sigma^n \mathbf{i}) - \pi(\mathbf{x})\| \leqslant \operatorname{diam}(X_{\mathbf{x}|_1^{\rho(n)}}) \leqslant C^{-2}\psi(n)\operatorname{diam}(X).$$

Then, using (2.1), (2.3), (2.7) and the fact that $\pi(\mathbf{x}) = x$, for infinitely many $n \in \mathbb{N}$ we have

$$\begin{aligned}
\|\pi(\mathbf{i}) - f_{\mathbf{i}|_1^n}(x)\| &\leqslant C\|f'_{\mathbf{i}|_1^n}(\xi)\|\|\pi(\sigma^n \mathbf{i}) - x\| \\
&\leqslant C^{-1}\|f'_{\mathbf{i}|_1^n}(\xi)\|\operatorname{diam}(X)\psi(n) \\
&\leqslant \operatorname{diam}(X_{\mathbf{i}|_1^n})\psi(n) \\
&= \Psi(\mathbf{i}|_1^n).
\end{aligned}$$

Thus, $\pi(\mathbf{i}) \in W(x, \Psi)$. $\qquad\square$

Our goal now is to prove the following proposition, which implies Proposition 4.1.

For simplicity, throughout the rest of the paper, we use the Vinogradov notation and write $A \ll B$ to denote that $A \leqslant dB$ for some constant $d > 0$. When we refer to explicit constants C, these may not always be the same constant but will typically be related to the bounded distortion property (2.2) or the constant arising in (2.6).

Proposition 5.2. *Let $\mathbf{x} \in \Sigma$ and let $\rho \colon \mathbb{N} \mapsto \mathbb{N}$ be a monotonic increasing function with*

$$\liminf_{n \to \infty} \frac{-1}{n} \log \|f'_{\mathbf{x}|_1^{\rho(n)}}(\pi(\mathbf{x}))\| =: \alpha \in (0, \infty).$$

Let s be the unique solution of the equation $P(s) = s\alpha$. Suppose that

$$\sum_{n=0}^{\infty} e^{\alpha s n} \|f'_{\mathbf{x}|_1^{\rho(n)}}(\pi(\mathbf{x}))\|^s = \infty.$$

Then, there exists a probability measure η such that $\eta(\widehat{W}(\mathbf{x}, \rho)) = 1$ and, for every $\delta > 0$, there exists a $K \geqslant 1$ such that for every $\mathbf{i} \in \Lambda^*$ with $|\mathbf{i}| \geqslant K$,

$$\eta([\mathbf{i}]) \ll \delta \cdot (\operatorname{diam}(X_{\mathbf{i}}))^s,$$

where the implicit constant is independent of $\mathbf{i}$ and δ.

For $r > 0$, denote by Θ_r the sequences $\mathbf{i} \in \Lambda^*$ for which the cylinders $f_{\mathbf{i}}(X)$ have diameter approximately equal to r. More precisely,

$$\Theta_r = \{\mathbf{i} \in \Lambda^* : \operatorname{diam}(X_{\mathbf{i}}) \leqslant r < \operatorname{diam}(X_{\mathbf{i}|_1^{|\mathbf{i}|-1}})\}.$$

Note that the collection of cylinders $\{[\mathbf{i}] : \mathbf{i} \in \Theta_r\}$ partitions Σ.

Proof of Proposition 4.1. First let us show that the assumption on divergence implies the divergence of the series in Proposition 5.2. Take $\mathbb{P}$ to be the measure described in (2.6) and let $\rho : \mathbb{N} \to \mathbb{N}$ be as defined in (5.1). Then, from the definitions of $\mathbb{P}$ and ρ it follows by (2.2) and (2.3) that

$$\infty = \sum_{n=0}^{\infty} \psi(n)^s \sum_{\mathbf{i} \in \Lambda^n} \|f_{\mathbf{i}}'(x)\|^s \leqslant C^{2s+1} \operatorname{diam}(X)^{-s} \sum_{n=0}^{\infty} \operatorname{diam}(X_{\mathbf{x}|_1^{\rho(n)-1}})^s \sum_{\mathbf{i} \in \Lambda^n} e^{\alpha s n} \mathbb{P}([\mathbf{i}])$$

$$\leqslant C^{3s+1} \sum_{n=0}^{\infty} \|f_{\mathbf{x}|_1^{\rho(n)-1}}'(f_{x_{\rho(n)}}(x))\|^s e^{\alpha s n}$$

$$\leqslant C^{3s+1} \sum_{n=0}^{\infty} a_{\min}^{-s} \|f_{x_{\rho(n)}}'(x)\|^s \|f_{\mathbf{x}|_1^{\rho(n)-1}}'(f_{x_{\rho(n)}}(x))\|^s e^{\alpha s n}$$

$$= C^{3s+1} a_{\min}^{-s} \sum_{n=0}^{\infty} \|f_{\mathbf{x}|_1^{\rho(n)}}'(x)\|^s e^{\alpha s n}.$$

To obtain the penultimate line of the above we employ the bounded distortion property (2.2) and to obtain the final inequality we use the chain rule.

Applying Proposition 5.2, let η be the probability measure described, let $\delta > 0$ be arbitrary, and let K be the corresponding index given in Proposition 5.2. Choose $R > 0$ sufficiently small such that $\min\{|\mathbf{i}| : \mathbf{i} \in \Theta_r\} \geqslant K$ for every $0 < r < R$.

By [22, Corollary 5.8 and Theorem 3.9], there exists a constant $C \geqslant 0$ such that for any bounded Borel subset I of $\mathbb{R}^t$

$$\sharp\{\mathbf{i} \in \Theta_{\operatorname{diam}(I)} : f_{\mathbf{i}}(X) \cap I \neq \varnothing\} \leqslant C.$$

Let I be a bounded Borel subset of $\mathbb{R}^t$ such that $\operatorname{diam}(I) < R$. Now, by Proposition 5.2, we have

$$\pi_*\eta(I) \leqslant \pi_*\eta\left(\bigcup_{\substack{\mathbf{i} \in \Theta_{\operatorname{diam}(I)} \\ f_{\mathbf{i}}(X) \cap I \neq \varnothing}} f_{\mathbf{i}}(X)\right) \leqslant \sum_{\substack{\mathbf{i} \in \Theta_{\operatorname{diam}(I)} \\ f_{\mathbf{i}}(X) \cap I \neq \varnothing}} \eta([\mathbf{i}]) \ll \sum_{\substack{\mathbf{i} \in \Theta_{\operatorname{diam}(I)} \\ f_{\mathbf{i}}(X) \cap I \neq \varnothing}} \delta \operatorname{diam}(X_{\mathbf{i}})^s \leqslant C\delta \operatorname{diam}(I)^s.$$

Let $\{I_i\}_i$ be such that $\pi\widehat{W}(\mathbf{x}, \rho) \subseteq \bigcup_i I_i$ and $\operatorname{diam}(I_i) < R$, i.e. let $\{I_i\}_i$ be an R-cover for $\pi\widehat{W}(\mathbf{x}, \rho)$. Also recall that, by Proposition 5.2, we have $\eta(\widehat{W}(\mathbf{x}, \rho)) = 1$. Hence, using the above inequality, we have

$$\sum_i \operatorname{diam}(I_i)^s \gg \sum_i \frac{1}{\delta} \pi_*\eta(I_i) \geqslant \frac{1}{\delta} \pi_*\eta\left(\bigcup_i I_i\right) \geqslant \frac{1}{\delta} \pi_*\eta(\pi\widehat{W}(\mathbf{x}, \rho)) = \frac{1}{\delta}.$$

Therefore, by the definition of Hausdorff s-measure, $\mathcal{H}^s(\pi\widehat{W}(\mathbf{x},\rho)) \gg \frac{1}{\delta}$. Since $\delta > 0$ was arbitrary, this implies that $\mathcal{H}^s(\pi\widehat{W}(\mathbf{x},\rho)) = \infty$. By Lemma 5.1, $\mathcal{H}^s(\pi\widehat{W}(\mathbf{x},\rho)) \leqslant \mathcal{H}^s(W(x,\Psi))$, and thus the proof is complete. $\qquad\square$

Before we turn to the proof of Proposition 5.2, we prove a technical lemma. We say that $\mathbf{i} = (i_1, i_2, \ldots) \in \Sigma$ is m-*periodic* if $i_k = i_{k+m}$ for every $k \geqslant 1$. We say that $\mathbf{i} \in \Sigma$ is m-*periodic on* (ℓ, n), where $n - \ell \geqslant m$, if $i_k = i_{k+m}$ for $\ell \leqslant k \leqslant n - m$.

Lemma 5.3. *Let* $\mathbf{x} \in \Sigma$ *and let* $n \in \mathbb{N}$. *Suppose that*

$$m(\mathbf{x}, n) := \min\{k : 1 \leqslant k \text{ and } \mathbf{x}|_1^{n-k} = \mathbf{x}|_{k+1}^n\} < n/2.$$

Then, $\mathbf{x}$ *is* $m(\mathbf{x}, n)$-*periodic on* $(1, n)$. *Moreover, for each* $1 \leqslant k \leqslant n - m(\mathbf{x}, n)$, *we have* $\mathbf{x}|_1^{n-k} = \mathbf{x}|_{k+1}^n$ *if and only if there exists* $p \in \mathbb{N}$ *such that* $k = p \cdot m(\mathbf{x}, n)$.

Proof. For convenience, let us write $m = m(\mathbf{x}, n)$. By the definition of m, we have $x_\ell = x_{m+\ell}$ for every $\ell = 1, \ldots, n - m$ and, thus, the proof of the first part of the lemma is complete.

Next, let $q := \left\lfloor \frac{n}{m} \right\rfloor$. Note that $q \geqslant 2$ since $m < \frac{n}{2}$. By using the m-periodicity of $\mathbf{x}$, we have that $x_\ell = x_{qm+\ell}$ for every $\ell = 1, \ldots, n - qm$. In other words, $\mathbf{x}|_1^{n-qm} = \mathbf{x}|_{qm+1}^n$. Thus, again using the m-periodicity of $\mathbf{x}$, there are words $\boldsymbol{\tau} \in \Lambda^{n-qm}$ and $\boldsymbol{\omega} \in \Lambda^{(q+1)m-n}$ such that $\mathbf{x} = \boldsymbol{\tau}\boldsymbol{\omega}\boldsymbol{\tau}\ldots\boldsymbol{\omega}\boldsymbol{\tau}$. Hence, for every $p = 1, \ldots, q$, we have $\mathbf{x}|_1^{n-pm} = \mathbf{x}|_{pm+1}^n$. In particular, this yields that $\mathbf{x}|_1^{n-k} = \mathbf{x}|_{k+1}^n$ if $k = pm$ for some $p = 1, \ldots, q$.

For the other direction, we argue by contradiction. Let us suppose that there exists some k such that $m \nmid k$ and $\mathbf{x}|_1^{n-k} = \mathbf{x}|_{k+1}^n$. By the definition of m, it follows that $m < k$. In order to obtain a contradiction, it is enough to show that

$$\mathbf{x}|_1^{n-k+m} = \mathbf{x}|_{k+1-m}^n.$$

Then, by induction, one can find $\ell \in \mathbb{N}$ such that for $k' = k - \ell m < m$, we have $\mathbf{x}|_{k'+1}^n = \mathbf{x}|_1^{n-k'}$, which is a contradiction.

Since $\mathbf{x}|_{n-m+1}^n = \boldsymbol{\omega}\boldsymbol{\tau}$ and $\mathbf{x}|_1^{n-k} = \mathbf{x}|_{k+1}^n$, by using the m-periodicity of $\mathbf{x}$ we have $\boldsymbol{\omega}\boldsymbol{\tau} = \mathbf{x}|_{n-m+1}^n = \mathbf{x}|_{n-k-m+1}^{n-k} = \mathbf{x}|_{n-k+1}^{n-k+m}$. Similarly, $\boldsymbol{\tau}\boldsymbol{\omega} = \mathbf{x}|_1^m = \mathbf{x}|_{k+1}^{k+m} = \mathbf{x}|_{k+1-m}^k$. Hence, using the m-periodicity of $\mathbf{x}$ and the fact that $\mathbf{x}|_1^m = \boldsymbol{\tau}\boldsymbol{\omega}$, we have

$$\mathbf{x}|_1^{n-k+m} = \mathbf{x}|_1^{n-k}\boldsymbol{\omega}\boldsymbol{\tau} = \boldsymbol{\tau}\boldsymbol{\omega}\mathbf{x}|_{m+1}^{n-k}\boldsymbol{\omega}\boldsymbol{\tau} = \boldsymbol{\tau}\boldsymbol{\omega}\mathbf{x}|_1^{n-k-m}\boldsymbol{\omega}\boldsymbol{\tau}$$

$$= \boldsymbol{\tau}\boldsymbol{\omega}\mathbf{x}|_{k+1}^{n-m}\boldsymbol{\omega}\boldsymbol{\tau} = \boldsymbol{\tau}\boldsymbol{\omega}\mathbf{x}|_{k+1}^n = \mathbf{x}|_{k-m+1}^n,$$

as required. $\qquad\square$

6. Construction of the Mass Distribution

Let $\Phi = \{f_i\}_{i\in\Lambda}$ be a conformal iterated function system satisfying the open set condition. Throughout the next three sections, we fix an $\mathbf{x} \in \Sigma$ and a function $\rho \colon \mathbb{N} \mapsto \mathbb{N}$ such that $n \mapsto \rho(n)$ is monotonically increasing and for which

$$\liminf_{n\to\infty} \frac{-1}{n} \log \|f'_{\mathbf{x}|_1^{\rho(n)}}(\xi)\| =: \alpha \in (0, \infty),$$

where $\xi \in X$. Note that by the bounded distortion property (2.2), we may take ξ to be any element of X. We extend the function $\rho \colon \mathbb{N} \mapsto \mathbb{N}$ to a map $\rho \colon \mathbb{R}^+ \mapsto \mathbb{N}$ in a natural way; that is, $\rho(x) := \rho(\lfloor x \rfloor)$.

Let s be the unique solution of the equation $P(s) = s\alpha$ and suppose that

$$\sum_{n=0}^{\infty} e^{\alpha s n} \|f'_{\mathbf{x}|_1^{\rho(n)}}(\xi)\|^s = \infty.$$

For a strictly monotonic increasing sequence $\underline{A} = (A_k)$ of natural numbers, let

$$C_{\underline{A}} = \{\mathbf{j} \in \Sigma : \sigma^\ell \mathbf{j} \notin [\mathbf{x}|_1^{\rho(\ell)}] \text{ for every } A_k \neq \ell \geqslant 0 \text{ but } \sigma^{A_k}\mathbf{j} \in [\mathbf{x}|_1^{\rho(A_k)}] \text{ for every } k \geqslant 1\}.$$

Observe that for every strictly monotonic increasing sequence $\underline{A}$, $C_{\underline{A}}$ is compact and for $\underline{A} \neq \underline{A}'$, $C_{\underline{A}} \cap C_{\underline{A}'} = \varnothing$.

In order to achieve the correct dimension (as given by Theorem 2.1), we restrict ourselves to the sequences $\underline{A}$, which are rapidly growing. By taking sequences $\underline{A}$ which are rapidly growing, we ensure that $C_{\underline{A}} \subset \widehat{W}(\mathbf{x}, \rho)$ but at the same time benefits from as much freedom as possible between consecutive "hits" of the shrinking target set. The next lemma will be used to show that there exists an uncountable set of such sequences. To save on notation, let us write

$$\varepsilon(n) := e^{\alpha s n} \|f'_{\mathbf{x}|_1^{\rho(n)}}(\xi)\|^s.$$

Lemma 6.1. *There exist sequences $\{n_k\}$ and $\{m_k\}$ such that*

(i) $n_1 > \max\left\{4, \frac{4}{\alpha^2}\right\}$,

(ii) $n_1 > \left(\dfrac{-4\log a_{\min} + \alpha}{\alpha}\right)^2$,

(iii) $n_1 > \dfrac{-8\log a_{\min}}{\alpha}\left(\dfrac{-2\log a_{\min}}{\alpha} + 2\right)$,

(iv) $C\left(a_{\max}^s e^{-\alpha s}\right)^{\min\{1, \frac{\alpha}{-2\log a_{\min}}\}}\sqrt{1 + \frac{\alpha}{-2\log a_{\min}}}\sqrt{n_1} < 1$,

and for every $n \geqslant n_1 - \sqrt{n_1}$, we have $\rho(n) \geqslant \frac{\alpha}{-2\log a_{\min}}n$. Moreover, for every $k \geqslant 1$,

(1) $n_k + \rho(n_k) < m_k$,

(2) $\max\{m_{k-1} + \rho(m_{k-1}) + 2, (2m_{k-1} + \rho(m_{k-1}))^2\} < n_k$,

(3) $\displaystyle\lim_{\ell \to \infty} \frac{C^{(1+s)\ell}a_{\min}^{-2s\ell}e^{s\alpha(\sum_{j=1}^{\ell-1}(m_j + \rho(m_j)) + 2\ell)}}{\prod_{j=1}^{\ell}\sum_{k=n_j}^{m_j}\varepsilon(k)} = 0$,

(4) $\displaystyle\lim_{\ell \to \infty} \frac{C^{(1+s)\ell}a_{\min}^{-2s\ell}e^{-s\alpha(n_\ell - \sum_{j=1}^{\ell-1}(m_j + \rho(m_j)) - 2\ell)}}{\prod_{j=1}^{\ell-1}\sum_{k=n_j}^{m_j}\varepsilon(k)} = 0$.

Throughout this lemma, C is the constant arising from (2.2).

Essentially, the terms (3) and (4) of Lemma 6.1 will play an important role in the proof of Proposition 5.2. The construction of such sequences is possible basically because in (4) n_ℓ appears only in the numerator in the term $e^{-s\alpha n_\ell}$ and in (3) m_ℓ appears only in the denominator in $\sum_{k=n_\ell}^{m_\ell}\varepsilon_k$. This allows us to choose a sufficiently rapidly growing sequence such that these terms converge to zero.

Proof. First of all note that, by the definition of α, there exists an $N \in \mathbb{N}$ such that for all natural numbers $n \geqslant N$,

$$\frac{\alpha}{2} \leqslant \frac{-\log\|f'_{\mathbf{x}|_1^{\rho(n)}}(\xi)\|}{n}$$

for any $\xi \in X$. Combining this with the bounds in (1.2) we see that

$$\frac{\alpha}{2} \leqslant \frac{-\log a_{\min}^{\rho(n)}}{n} = \frac{-\rho(n) \log a_{\min}}{n}.$$

Thus, for all $n \geqslant N$,

$$\rho(n) \geqslant \frac{\alpha}{-2 \log a_{\min}} n.$$

We now construct sequences $\{n_k\}$ and $\{m_k\}$ inductively. Let us fix an arbitrary sequence converging to 0, say $p_n = 2^{-n}$. We begin by choosing n_1 sufficiently large so that (i)-(iv) hold and $n_1 - \sqrt{n_1} \geqslant N$. Then choose m_1 such that $n_1 + \rho(n_1) < m_1$. We then proceed by induction. Suppose that n_k and m_k satisfying (1)–(4) have already been defined for $k = 1, \ldots, \ell - 1$. Next, find n_ℓ such that (2) holds and

$$\frac{C^{(1+s)\ell} a_{\min}^{-2s\ell} e^{-s\alpha(n_\ell - \sum_{j=1}^{\ell-1}(m_j + \rho(m_j)) - 2\ell)}}{\prod_{j=1}^{\ell-1} \sum_{k=n_j}^{m_j} \varepsilon(k)} < p_\ell.$$

This is possible since $s, \alpha > 0$. We then find m_ℓ so that (1) holds and

$$\frac{C^{(1+s)\ell} a_{\min}^{-2s\ell} e^{s\alpha(\sum_{j=1}^{\ell-1}(m_j + \rho(m_j)) + 2\ell)}}{\prod_{j=1}^{\ell} \sum_{k=n_j}^{m_j} \varepsilon(k)} < p_\ell.$$

This is possible by the divergence of $\sum_{k=1}^{\infty} \varepsilon(k)$. By construction, the sequences $\{n_k\}$ and $\{m_k\}$ satisfy all of the required conditions, thus completing the proof of the lemma. $\square$

Let Ξ be the set of sequences such that $A_k \in [n_k, m_k]$ for every $k \geqslant 1$. In the rest of the paper, we construct the mass distribution η as follows. We define a family of probability measures $\{\mu_{\underline{A}}\}_{\underline{A} \in \Xi}$, where $\mu_{\underline{A}}$ is supported on $C_{\underline{A}}$, and an appropriate probability measure ν on Ξ. We will then set $\eta = \int \mu_{\underline{A}} d\nu(\underline{A})$. Clearly, $\eta(\widehat{W}(\mathbf{x}, \rho)) = 1$ since, for every $\underline{A} \in \Xi$, $C_{\underline{A}} \subset \widehat{W}(\mathbf{x}, \rho)$.

Let us define $R \colon \mathbb{R}^+ \mapsto \mathbb{N}$ as follows

$$R(x) := \max\{m \in \mathbb{N} : x \geqslant m + \rho(m)\}.$$

We adopt the convention that $R(n) = 1$ if $n < 1 + \rho(1)$. Since $\rho(n) \geqslant \frac{\alpha}{-2 \log a_{\min}} n$ for every $n \geqslant n_1 - \sqrt{n_1}$, it follows from the definition of $R(n)$ that

$$n \geqslant R(n) + \rho(R(n)) \geqslant \left(1 + \frac{\alpha}{-2 \log a_{\min}}\right) R(n)$$

whenever $R(n) \geqslant n_1 - \sqrt{n_1}$. In particular, for every $n \geqslant n_1 - \sqrt{n_1} + \rho(n_1 - \sqrt{n_1})$ we have

$$R(n) \leqslant \frac{-2 \log a_{\min}}{\alpha - 2 \log a_{\min}} n. \tag{6.1}$$

Let $p + \rho(p) + 2 < q$ be integers such that $p \geqslant 0$. Let

$$\Omega_{p,q} := \Big\{\mathbf{i} \in \Lambda^{q-p-2} : \mathbf{i}|_1^{\ell+\rho(\ell)-p-1} \neq \mathbf{x}|_{p-\ell+2}^{\rho(\ell)} \text{ for every } \ell = R(p + \sqrt{p} + 1) + 1, \ldots, p$$

$$\mathbf{i}|_{\ell+1}^{\ell+\rho(\ell+p+1)} \neq \mathbf{x}|_1^{\rho(\ell+p+1)} \text{ for } \ell = 0, \ldots, R(q-1) - p - 1, \text{ and} \quad (6.2)$$

$$\mathbf{i}|_{\ell+1}^{q-p-2} \neq \mathbf{x}|_1^{q-p-\ell-2} \text{ for } \ell = R(q-1) - p, \ldots, q - \sqrt{q}\Big\}.$$

By the choice of p and q we see that the subwords in the definition of the set $\Omega_{p,q}$ are well defined, since $\ell + \rho(\ell) - p - 1 \leqslant \rho(p) - 1 \leqslant q - p - 2$ and $\ell + \rho(\ell + p + 1) \leqslant$

$R(q-1) - p - 1 + \rho(R(q-1)) \leqslant q - p - 2$ and $R(q-1) - p \geqslant 0$ by the definition of R. We will later show that the set $\Omega_{p,q}$ is non-empty by giving a lower bound on its measure.

Let us recall that

$$m(\mathbf{x}, n) = \min\{k : 1 \leqslant k \text{ and } \mathbf{x}|_1^{n-k} = \mathbf{x}|_{k+1}^n\}.$$

Observe that if $\mathbf{x}|_1^{n-m(\mathbf{x},n)} = \mathbf{x}|_{m(\mathbf{x},n)+1}^n$ then $\mathbf{x}|_1^{n-m(\mathbf{x},n)-1} = \mathbf{x}|_{m(\mathbf{x},n)+1}^{n-1}$, and hence,

$$m(\mathbf{x}, n-1) \leqslant m(\mathbf{x}, n) \text{ for every } n \geqslant 3. \tag{6.3}$$

Next, we define two functions $\boldsymbol{\omega}, \boldsymbol{\tau} \colon \Xi \mapsto \Sigma$. For $\underline{A} \in \Xi$, the kth coordinates of $\boldsymbol{\omega}$ and $\boldsymbol{\tau}$ will depend solely upon on the kth coordinate A_k of $\underline{A}$, we will use the notation

$$\boldsymbol{\omega}(\underline{A}) = (\omega_1(A_1), \omega_2(A_2), \ldots, \omega_k(A_k), \ldots),$$
$$\text{and} \quad \boldsymbol{\tau}(\underline{A}) = (\tau_1(A_1), \tau_2(A_2), \ldots, \tau_k(A_k), \ldots).$$

For a sequence $\underline{A} \in \Xi$, we define the kth coordinate of $\boldsymbol{\omega}$ as to be arbitrary but *not* equal to $x_{m(\mathbf{x},\rho(A_k - \sqrt{A_k}))}$. Similarly, for $\underline{A} \in \Xi$, define $\tau_k(A_k)$ as to be arbitrary but *not* equal to $x_{\rho(A_k) - \left\lfloor \frac{\rho(A_k)}{m(\mathbf{x},\rho(A_k))} \right\rfloor \cdot m(\mathbf{x},\rho(A_k)) + 1}$.

Lemma 6.2. *Let $y \in \Lambda$ be such that $y \neq x_1$ and let $\mathbf{y} = (y, y, \ldots)$. For an $\underline{A} \in \Xi$, let*

$$\Gamma_{\underline{A}} := [\mathbf{y}|_1^{A_1-1}] \times \prod_{k=1}^{\infty} \left(\{\omega_k(A_k)\mathbf{x}|_1^{\rho(A_k)}\tau_k(A_k)\} \times \Omega_{A_k + \rho(A_k), A_{k+1}} \right). \tag{6.4}$$

Then $\Gamma_{\underline{A}} \subseteq C_{\underline{A}}$ for every $\underline{A} \in \Xi$.

Note that $\Gamma_{\underline{A}} \subseteq \widehat{W}(\mathbf{x}, \rho)$ for every $\underline{A} \in \Xi$ clearly by definition. However, later in the proof we need the more technical containment $\Gamma_{\underline{A}} \subseteq C_{\underline{A}}$, since then for any $\mathbf{i} \in \bigcup_{\underline{A} \in \Xi} \Gamma_{\underline{A}}$ there exists a unique $\underline{A} \in \Xi$ such that $\mathbf{i} \in \Gamma_{\underline{A}}$, and this is heavily used during the proof of Proposition 5.2.

Proof of Lemma 6.2. Observe that if $\mathbf{i} = (i_1, i_2, \ldots) \in \Gamma_{\underline{A}}$, then $\mathbf{i}|_1^{A_1-1} = \mathbf{y}|_1^{A_1-1}$ and, for every $k \geqslant 1$, we have:

- $i_{A_k} = \omega_k(A_k)$,
- $i_{A_k + \rho(A_k) + 1} = \tau_k(A_k)$,
- $\mathbf{i}|_{A_k+1}^{A_k+\rho(A_k)} = \mathbf{x}|_1^{\rho(A_k)}$, and
- $\mathbf{i}|_{A_k+\rho(A_k)+2}^{A_{k+1}-1} \in \Omega_{A_k+\rho(A_k),A_{k+1}}$.

In particular, $\sigma^{A_k}\mathbf{i} \in [\mathbf{x}|_1^{\rho(A_k)}]$ for every $k \geqslant 1$.

Now let $\ell \notin \underline{A}$. To show that $\mathbf{i} \in C_{\underline{A}}$, we need to show that $\sigma^\ell \mathbf{i} \notin [\mathbf{x}|_1^{\rho(\ell)}]$. We argue by contradiction, so suppose to the contrary that $\sigma^\ell \mathbf{i} \in [\mathbf{x}|_1^{\rho(\ell)}]$. It is easy to see that $\ell < A_1$ is not possible since $y \neq x_1$. Hence, we may assume that $A_k < \ell < A_{k+1}$ for some $k \geqslant 1$.

There are five possible cases to consider:

(1) $A_k < \ell \leqslant R\left(A_k + \rho(A_k) + \left\lfloor \sqrt{A_k + \rho(A_k)} \right\rfloor + 1\right)$

(2) $R\left(A_k + \rho(A_k) + \left\lfloor \sqrt{A_k + \rho(A_k)} \right\rfloor + 1\right) + 1 \leqslant \ell \leqslant A_k + \rho(A_k)$

(3) $A_k + \rho(A_k) + 1 \leqslant \ell \leqslant R(A_{k+1} - 1)$,

(4) $R(A_{k+1} - 1) + 1 \leqslant \ell \leqslant A_{k+1} - 1 - \left\lfloor \sqrt{A_{k+1}} \right\rfloor$,

(5) $A_{k+1} - \left\lfloor \sqrt{A_{k+1}} \right\rfloor \leqslant \ell < A_{k+1}$

First, consider the case (1). Recall that $\sigma^{A_k}\mathbf{i} \in [\mathbf{x}|_1^{\rho(A_k)}]$ and we are assuming, in order to eventually reach a contradiction, that $\sigma^{\ell}\mathbf{i} \in [\mathbf{x}|_1^{\rho(\ell)}]$. Consequently,

$$\mathbf{i}|_{A_k+1}^{A_k+\rho(A_k)} = \mathbf{x}|_1^{\rho(A_k)} \qquad \text{and} \qquad \mathbf{i}|_{\ell+1}^{\ell+\rho(\ell)} = \mathbf{x}|_1^{\rho(\ell)}.$$

So, it follows that

$$\mathbf{x}|_{\ell-A_k+1}^{\rho(A_k)} = \mathbf{i}|_{A_k+1+(\ell-A_k)}^{A_k+\rho(A_k)} = \mathbf{i}|_{\ell+1}^{\rho(A_k)+A_k} = \mathbf{x}|_1^{\rho(A_k)+A_k-(\ell+1)+1} = \mathbf{x}|_1^{\rho(A_k)-\ell+A_k}.$$

Thus, we have

$$m(\mathbf{x}, \rho(A_k)) \leqslant \ell - A_k$$
$$\leqslant R\left(A_k + \rho(A_k) + 1 + \left\lfloor \sqrt{A_k + \rho(A_k)} \right\rfloor \right) - A_k$$
$$\leqslant \rho(A_k) + 1 + \sqrt{A_k + \rho(A_k)} - \rho\left(R\left(A_k + \rho(A_k) + 1 + \left\lfloor \sqrt{A_k + \rho(A_k)} \right\rfloor \right)\right)$$
$$\leqslant \rho(A_k) + 1 + \sqrt{A_k + \rho(A_k)} - \rho\left(R\left(A_k + \rho(A_k)\right)\right)$$
$$= \rho(A_k) + 1 + \sqrt{A_k + \rho(A_k)} - \rho\left(A_k\right)$$
$$= \sqrt{A_k + \rho(A_k)} + 1.$$

By Lemma 6.1(iii) and the fact that $\rho(A_k) \geqslant \frac{\alpha A_k}{-2\log a_{\min}}$, we have

$$A_k < \frac{\alpha}{-2\log a_{\min}} A_k \left(\frac{\alpha}{-8\log a_{\min}} A_k - 2\right) < \rho(A_k)\left(\frac{\rho(A_k)}{4} - 2\right) + 1.$$

Hence, $m(\mathbf{x}, \rho(A_k)) \leqslant \ell - A_k \leqslant \sqrt{A_k + \rho(A_k)} + 1 < \rho(A_k)/2 \leqslant \rho(A_k) - m(\mathbf{x}, \rho(A_k))$. So, it follows from Lemma 5.3 that there exists $p \in \mathbb{N}$ such that $\ell - A_k = p \cdot m(\mathbf{x}, \rho(A_k))$ and also that $\mathbf{x}$ is $m(\mathbf{x}, \rho(A_k))$-periodic on $(1, \rho(A_k))$. However, since $\mathbf{i}|_{\ell+1}^{\ell+\rho(\ell)} = \mathbf{x}|_1^{\rho(\ell)}$, we have

$$\tau_k(A_k) = i_{A_k+\rho(A_k)+1} = i_{A_k-\ell+\ell+\rho(A_k)+1} = i_{\rho(A_k)-p\cdot m(\mathbf{x}, \rho(A_k))+\ell+1}$$
$$= x_{\rho(A_k)-p\cdot m(\mathbf{x}, \rho(A_k))+1} = x_{\rho(A_k)-\left\lfloor \frac{\rho(A_k)}{m(\mathbf{x}, \rho(A_k))} \right\rfloor \cdot m(\mathbf{x}, \rho(A_k))+1} \neq \tau_k(A_k),$$

which contradicts the definition of $\tau_k(A_k)$.

Suppose that (2) holds. Since $\sigma^{\ell}\mathbf{i} \in [\mathbf{x}|_1^{\rho(\ell)}]$, we have $\mathbf{i}|_{A_k+\rho(A_k)+2}^{\ell+\rho(\ell)} = \mathbf{x}|_{A_k+\rho(A_k)-\ell+2}^{\rho(\ell)}$. However, $\mathbf{i}|_{A_k+\rho(A_k)+2}^{A_{k+1}-1} \in \Omega_{A_k+\rho(A_k), A_{k+1}}$ and so, by definition, we have $\mathbf{i}|_{A_k+\rho(A_k)+2}^{\ell+\rho(\ell)} \neq \mathbf{x}|_{A_k+\rho(A_k)-\ell+2}^{\rho(\ell)}$, which is a contradiction.

Similarly, if (3) holds then again by definition $\mathbf{i}|_{\ell+1}^{\ell+\rho(\ell)} \neq \mathbf{x}|_1^{\rho(\ell)}$, and hence $\sigma^{\ell}\mathbf{i} \notin [\mathbf{x}|_1^{\rho(\ell)}]$, a contradiction.

If (4) holds then, by $\sigma^{\ell}\mathbf{i} \in [\mathbf{x}|_1^{\rho(\ell)}]$, we have $\mathbf{i}|_{\ell+1}^{A_{k+1}-1} = \mathbf{x}|_1^{A_{k+1}-1-\ell}$. However, it follows from the fact that $\mathbf{i}|_{A_k+\rho(A_k)+2}^{A_{k+1}-1} \in \Omega_{A_k+\rho(A_k), A_{k+1}}$ that $\mathbf{i}|_{\ell+1}^{A_{k+1}-(A_k+\rho(A_k))-2} \neq \mathbf{x}|_1^{A_{k+1}-(A_k+\rho(A_k))-\ell-2}$. In particular, this implies that $\mathbf{i}|_{\ell+1}^{A_{k+1}-1} \neq \mathbf{x}_1^{A_{k+1}-1-\ell}$, which is again a contradiction.

Finally, let us consider the case (5), which proof is similar to the case (1). Since $\sigma^{\ell}\mathbf{i} \in [\mathbf{x}|_1^{\rho(\ell)}]$ and $\sigma^{A_k+}\mathbf{i} \in [\mathbf{x}|_1^{\rho(A_{k+1})}]$, we have

$$\mathbf{x}|_{A_{k+1}+1-\ell}^{\rho(\ell)} = \mathbf{i}|_{A_{k+1}+1}^{\rho(\ell)+\ell} = \mathbf{x}|_1^{\rho(\ell)+\ell-A_{k+1}}.$$

Thus, $m(\mathbf{x}, \rho(\ell)) \leqslant A_{k+1} - \ell \leqslant \sqrt{A_{k+1}}$. Observe that, by (6.3), we have $m(\mathbf{x}, \rho(\ell)) \geqslant m(\mathbf{x}, \rho(A_{k+1}-\lfloor\sqrt{A_{k+1}}\rfloor))$ and also note that $\mathbf{x}|_{A_{k+1}+1-\ell}^{\rho(A_{k+1}-\lfloor\sqrt{A_{k+1}}\rfloor)} = \mathbf{x}|_1^{\rho(A_{k+1}-\lfloor\sqrt{A_{k+1}}\rfloor)+\ell-A_{k+1}}.$

By Lemma 6.1, since $\underline{A} \in \Xi$, we have

$$\rho(A_{k+1} - \lfloor\sqrt{A_{k+1}}\rfloor) \geqslant \frac{\alpha}{-2\log a_{\min}}\left(A_{k+1} - \sqrt{A_{k+1}}\right).$$

By Lemma 6.1(ii), we have $\sqrt{A_{k+1}} < \frac{\alpha}{-4\log a_{\min}}\left(A_{k+1} - \sqrt{A_{k+1}}\right)$ and, hence,

$$m(\mathbf{x}, \rho(A_{k+1} - \lfloor\sqrt{A_{k+1}}\rfloor)) \leqslant A_{k+1} - \ell < \frac{\rho(A_{k+1} - \lfloor\sqrt{A_{k+1}}\rfloor)}{2}$$
$$\leqslant \rho(A_{k+1} - \sqrt{A_{k+1}}) - m(\mathbf{x}, \rho(A_{k+1} - \lfloor\sqrt{A_{k+1}}\rfloor)).$$

Again, it follows from Lemma 5.3 that $\mathbf{x}$ is $m(\mathbf{x}, \rho(A_{k+1} - \lfloor\sqrt{A_{k+1}}\rfloor))$-periodic on $(1, \rho(A_{k+1} - \lfloor\sqrt{A_{k+1}}\rfloor))$. Furthermore, we also have that there exists $q \in \mathbb{N}$ such that $q \cdot m(\mathbf{x}, \rho(A_{k+1} - \lfloor\sqrt{A_{k+1}}\rfloor)) = A_{k+1} - \ell$. Hence, by the definition of $\omega_{k+1}(A_{k+1})$ and using that $\mathbf{i}|_{\ell+1}^{\rho(\ell)+\ell} = \mathbf{x}|_1^{\rho(\ell)}$, we have

$$\omega_{k+1}(A_{k+1}) = i_{A_{k+1}} = i_{A_{k+1}-\ell+\ell} = i_{q \cdot m(\mathbf{x},\rho(A_{k+1}-\lfloor\sqrt{A_{k+1}}\rfloor))+\ell}$$
$$= x_{q \cdot m(\mathbf{x},\rho(A_{k+1}-\lfloor\sqrt{A_{k+1}}\rfloor))} = x_{m(\mathbf{x},\rho(A_{k+1}-\lfloor\sqrt{A_{k+1}}\rfloor))} \neq \omega_{k+1}(A_{k+1}),$$

which is a contradiction. $\qquad\square$

For $\mathbf{i} \in \Lambda^{q-p-2}$, let

$$\mathbb{P}_{p,q}(\mathbf{i}) = \mathbb{P}([\mathbf{i}]),$$

where $\mathbb{P}$ is the measure defined by (2.6). Thus, $\mathbb{P}_{p,q}$ is the probability measure supported on Λ^{q-p-2} corresponding to the equilibrium state $\mathbb{P}$.

Denote by $\widehat{\Omega}_{p,q}$ the subset of Σ such that $\mathbf{i} \in \widehat{\Omega}_{p,q}$ if and only if $\mathbf{i}|_{p+2}^{q-1} \in \Omega_{p,q}$. Since $\mathbb{P}$ is σ-invariant by definition, $\mathbb{P}(\widehat{\Omega}_{p,q}) = \mathbb{P}_{p,q}(\Omega_{p,q})$.

Lemma 6.3. *There exists a constant $C' > 0$ such that for every $p > n_1$ (with n_1 as defined in Lemma 6.1) and for every $q > \max\{p + \rho(p) + 2, (p+2)^2\}$,*

$$\mathbb{P}_{p,q}(\Omega_{p,q}) \geqslant 1 - C'\left(a_{\max}^s e^{-\alpha s}\right)^{\min\left\{\sqrt{p}, \sqrt{q}-p, \frac{\alpha p}{-2\log a_{\min}}\right\}}.$$

Proof. First, note that $\min\{\rho(\ell + p + 1), q - p - 2 - \ell\} = \rho(\ell + p + 1)$ if and only if $\ell \leqslant R(q-1) - p - 1$. Furthermore, note that $a_{\max}^s e^{-\alpha s} < 1$ since $\alpha > 0$ and $s > 0$.

Denote the complement of $\widehat{\Omega}_{p,q}$ by $\widehat{\Omega}_{p,q}^c$. Since $\widehat{\Omega}_{p,q} = \bigcup_{\mathbf{i}\in\Omega_{p,q}} \sigma^{-(p+1)}[\mathbf{i}]$, using the definition of $\Omega_{p,q}$ in (6.2) we have that

$$\widehat{\Omega}_{p,q}^c \subseteq \left(\bigcup_{\ell=R(p+\sqrt{p}+1)+1}^{p} \sigma^{-(p+1)}\left[\mathbf{x}|_{p-\ell+2}^{\rho(\ell)}\right]\right) \bigcup \left(\bigcup_{\ell=0}^{q-\sqrt{q}} \sigma^{-(\ell+p+1)}\left[\mathbf{x}|_1^{\min\{\rho(\ell+p+1),q-p-2-\ell\}}\right]\right).$$
(6.5)

Then by the σ-invariance of $\mathbb{P}$, for every $q > p + \rho(p) + 2$ we have

$$\mathbb{P}(\widehat{\Omega}_{p,q}^c) \leqslant \sum_{\ell=R(p+\sqrt{p}+1)+1}^{p} \mathbb{P}\left(\left[\mathbf{x}|_{p-\ell+2}^{\rho(\ell)}\right]\right) + \sum_{\ell=0}^{q-\sqrt{q}} \mathbb{P}\left(\left[\mathbf{x}|_1^{\min\{\rho(\ell+p+1),q-p-2-\ell\}}\right]\right).$$

Here, the assumption that $q > p + \rho(p) + 2$ ensures that $\left[\mathbf{x}\big|_1^{\min\{\rho(\ell+p+1),q-p-2-\ell\}}\right] \neq \varnothing$ for $\ell \geqslant 0$. It then follows from (2.6) and (1.2) that

$$
\mathbb{P}(\widehat{\Omega}_{p,q}^c) \leqslant C \sum_{\ell=R(p+\sqrt{p}+1)+1}^{p} \|f'_{\mathbf{x}|_{p-\ell+2}^{\rho(\ell)}}(\xi)\|^s e^{-(\rho(\ell)-p+\ell-1)\alpha s}
$$

$$
+ C \sum_{\ell=0}^{q-\sqrt{q}} \|f'_{\mathbf{x}|_1^{\min\{\rho(\ell+p+1),q-p-2-\ell\}}}(\xi)\|^s e^{-\min\{\rho(\ell+p+1),q-p-2-\ell\}\alpha s}
$$

$$
\leqslant C \sum_{\ell=R(p+\sqrt{p}+1)+1}^{p} \left(a_{\max}^s e^{-\alpha s}\right)^{\rho(\ell)+\ell-p-1} + C \sum_{\ell=0}^{q-\sqrt{q}} \left(a_{\max}^s e^{-\alpha s}\right)^{\min\{\rho(\ell+p+1),q-p-2-\ell\}}.
$$

Since $\ell \mapsto \rho(\ell) + \ell - p - 1$ is a strictly monotonic increasing sequence of integers and

$$
\rho(R(p+\sqrt{p}+1)+1) + R(p+\sqrt{p}+1) + 1 - p - 1 > p + \sqrt{p} + 1 - p - 1 = \sqrt{p},
$$

we have

$$
\sum_{\ell=R(p+\sqrt{p}+1)+1}^{p} \left(a_{\max}^s e^{-\alpha s}\right)^{\rho(\ell)+\ell-p-1} \leqslant \sum_{k=\sqrt{p}}^{\infty} a_{\max}^{ks} e^{-\alpha s k}
$$

$$
= (a_{\max}^s e^{-\alpha s})^{\sqrt{p}} \sum_{k=0}^{\infty} a_{\max}^{ks} e^{-\alpha s k}
$$

$$
= \frac{(a_{\max}^s e^{-\alpha s})^{\sqrt{p}}}{1 - a_{\max}^s e^{-\alpha s}}.
$$

Next, note that, by Lemma 6.1, since $p > n_1$ by assumption, we have

$$
\rho(\ell + p + 1) \geqslant \frac{\alpha(\ell+p+1)}{-2\log a_{\min}}
$$

for any $\ell \geqslant 0$. Hence, we have

$$
\mathbb{P}(\widehat{\Omega}_{p,q}^c) \leqslant \frac{C\,(a_{\max}^s e^{-\alpha s})^{\sqrt{p}}}{1 - a_{\max}^s e^{-\alpha s}} + C \sum_{\ell=R(q-1)-p}^{q-\lfloor\sqrt{q}\rfloor} \left(a_{\max}^s e^{-\alpha s}\right)^{q-p-2-\ell} + C \sum_{\ell=0}^{R(q-1)-p-1} \left(a_{\max}^s e^{-\alpha s}\right)^{\rho(\ell+p+1)}
$$

$$
\leqslant \frac{C\,(a_{\max}^s e^{-\alpha s})^{\sqrt{p}}}{1 - a_{\max}^s e^{-\alpha s}} + C \sum_{\ell=\sqrt{q}-p-2}^{q-R(q-1)-2} \left(a_{\max}^s e^{-\alpha s}\right)^{\ell} + C \sum_{\ell=0}^{R(q-1)-p-1} \left(a_{\max}^s e^{-\alpha s}\right)^{\frac{\alpha(\ell+p+1)}{-2\log a_{\min}}}
$$

$$
\leqslant \frac{C\,(a_{\max}^s e^{-\alpha s})^{\sqrt{p}}}{1 - a_{\max}^s e^{-\alpha s}} + C \sum_{\ell=\sqrt{q}-p-2}^{\infty} \left(a_{\max}^s e^{-\alpha s}\right)^{\ell} + C \sum_{\ell=0}^{\infty} \left(a_{\max}^s e^{-\alpha s}\right)^{\frac{\alpha(\ell+p+1)}{-2\log a_{\min}}}
$$

$$
\leqslant \frac{C\,(a_{\max}^s e^{-\alpha s})^{\sqrt{p}}}{1 - a_{\max}^s e^{-\alpha s}} + \frac{C\,(a_{\max}^s e^{-\alpha s})^{\sqrt{q}-p-2}}{1 - a_{\max}^s e^{-\alpha s}} + \frac{C\,(a_{\max}^s e^{-\alpha s})^{\frac{\alpha(p+1)}{-2\log a_{\min}}}}{1 - (a_{\max}^s e^{-\alpha s})^{\frac{\alpha}{-2\log a_{\min}}}}
$$

$$
\leqslant C' \left(a_{\max}^s e^{-\alpha s}\right)^{\min\left\{\sqrt{p},\sqrt{q}-p,\frac{\alpha p}{-2\log a_{\min}}\right\}}.
$$

Note that, since $\ell > 0$, we require the condition $q > (p+2)^2$ to ensure the middle term above is valid.

Finally, since $\mathbb{P}_{p,q}(\Omega_{p,q}) = \mathbb{P}(\widehat{\Omega}_{p,q}) = 1 - \mathbb{P}(\widehat{\Omega}_{p,q}^c)$, the statement follows. $\qquad\square$

Lemma 6.4. *Let $0 < p_n < 1$ be a sequence such that $\sum_{n=1}^{\infty} p_n < \infty$ and $\max_n\{p_n\} < 1$. Then $\prod_{n=1}^{\infty}(1 - p_n) > 0$.*

Proof. Using the Taylor expansion of $\log(1-x)$, we see that

$$\log(1-x) = -\sum_{k=1}^{\infty}\frac{x^k}{k} \geqslant -x - \sum_{k=2}^{\infty}\frac{x^k}{2} = -x - \frac{x^2}{2(1-x)} \geqslant -x\left(1 + \frac{1}{2(1-x)}\right)$$

for every $0 < x < 1$. Therefore,

$$\prod_{n=1}^{\infty}(1-p_n) = \prod_{n=1}^{\infty} e^{\log(1-p_n)} \geqslant \prod_{n=1}^{\infty} e^{-p_n\left(1+\frac{1}{2(1-\max_n p_n)}\right)} = e^{-\left(1+\frac{1}{2(1-\max_n p_n)}\right)\sum_{n=1}^{\infty} p_n} > 0. \quad \square$$

Lemma 6.5. *There exists a constant $Q > 0$ such that for every $\underline{A} \in \Xi$,*

$$\prod_{j=1}^{\infty}\mathbb{P}_{A_j+\rho(A_j),A_{j+1}}(\Omega_{A_j+\rho(A_j),A_{j+1}}) \geqslant Q.$$

Proof. By Lemma 6.1(1) and Lemma 6.1(2),

$$\sqrt{n_{j+1}} - m_j - \rho(m_j) \geqslant m_j \geqslant n_j + \rho(n_j) \geqslant \sqrt{n_j + \rho(n_j)} \geqslant \sqrt{\left(1 + \frac{\alpha}{-2\log a_{\min}}\right)n_j}.$$

Since $A_j \in [n_j, m_j]$,

$$\min\left\{\sqrt{A_j + \rho(A_j)}, \sqrt{A_{j+1}} - A_j - \rho(A_j), \frac{\alpha(A_j + \rho(A_j))}{-2\log a_{\min}}\right\}$$

$$\geqslant \min\left\{\sqrt{n_j + \rho(n_j)}, \sqrt{n_{j+1}} - m_j - \rho(m_j), \frac{\alpha(n_j + \rho(n_j))}{-2\log a_{\min}}\right\}$$

$$\geqslant \min\left\{\sqrt{\left(1 + \frac{\alpha}{-2\log a_{\min}}\right)n_j}, \frac{\alpha}{-2\log a_{\min}}\left(1 + \frac{\alpha}{-2\log a_{\min}}\right)n_j\right\}$$

$$\geqslant \min\left\{1, \frac{\alpha}{-2\log a_{\min}}\right\}\sqrt{\left(1 + \frac{\alpha}{-2\log a_{\min}}\right)n_j}.$$

Hence, by Lemma 6.3,

$$\mathbb{P}_{A_j+\rho(A_j),A_{j+1}}(\Omega_{A_j+\rho(A_j),A_{j+1}}) \geqslant 1 - C\left(a_{\max}^s e^{-\alpha s}\right)^{\min\left\{1,\frac{\alpha}{-2\log a_{\min}}\right\}\sqrt{\left(1+\frac{\alpha}{-2\log a_{\min}}\right)n_j}}. \quad (6.6)$$

By (6.6), we have

$$\prod_{j=1}^{\infty}\mathbb{P}_{A_j+\rho(A_j),A_{j+1}}(\Omega_{A_j+\rho(A_j),A_{j+1}}) \geqslant \prod_{j=1}^{\infty}\left(1 - C\left(a_{\max}^s e^{-\alpha s}\right)^{\sqrt{n_j}\min\left\{1,\frac{\alpha}{-2\log a_{\min}}\right\}\sqrt{1+\frac{\alpha}{-2\log a_{\min}}}}\right)$$

$$\geqslant \prod_{n=n_1}^{\infty}\left(1 - C\left(a_{\max}^s e^{-\alpha s}\right)^{\sqrt{n}\min\{1,\frac{\alpha}{-2\log a_{\min}}\}\sqrt{1+\frac{\alpha}{-2\log a_{\min}}}}\right),$$

which is a positive constant by Lemma 6.1(iv) and Lemma 6.4. $\quad \square$

Let

$$\mathbb{S}_{p,q} := \frac{\mathbb{P}_{p,q}|_{\Omega_{p,q}}}{\mathbb{P}_{p,q}(\Omega_{p,q})}.$$

For a sequence $\underline{A} \in \Xi$, let $\mathbb{Q}_k$ be the probability measure on $\Lambda^{\rho(A_k)+2}$ such that

$$\mathbb{Q}_k(\mathbf{i}) = \begin{cases} 1 & \text{if } \mathbf{i} = \omega_k(A_k)\mathbf{x}|_1^{\rho(A_k)}\tau_k(A_k), \\ 0 & \text{otherwise;} \end{cases}$$

and let δ be the probability measure on Λ^{A_1-1} such that

$$\delta(\mathbf{i}) = \begin{cases} 1 & \text{if } \mathbf{i} = \underline{\mathbf{y}}|_1^{A_1-1}, \\ 0 & \text{otherwise.} \end{cases}$$

For each $\underline{A} \in \Xi$, we define a probability measure $\mu_{\underline{A}}$ as follows:

$$\mu_{\underline{A}} = \delta \times \prod_{k=1}^{\infty} \mathbb{Q}_k \times \mathbb{S}_{A_k+\rho(A_k),A_{k+1}}. \tag{6.7}$$

It follows from Lemma 6.5 that $\mu_{\underline{A}}$ is a well-defined probability measure on Σ with respect to the σ-algebra generated by the cylinder sets, since Lemma 6.5 guarantees that the normalising factor in the definition of $\mu_{\underline{A}}$ is non-zero. Moreover, by construction, $\mu_{\underline{A}}(\Gamma_{\underline{A}}) = 1$, where $\Gamma_{\underline{A}}$ is the set defined in Lemma 6.2.

Recall that we defined $\varepsilon(n) := e^{\alpha s n}\|f'_{\mathbf{x}|_1^{\rho(n)}}(\xi)\|^s$ and assumed that $\sum_{k=1}^{\infty} \varepsilon(n)$ is a divergent series. Finally, we define the probability measure ν on Ξ as

$$\nu([A_1,\ldots,A_\ell]) = \prod_{j=1}^{\ell} \frac{\varepsilon(A_j)}{\sum_{k=n_j}^{m_j} \varepsilon(k)}. \tag{6.8}$$

By its construction, the measure $\eta := \int \mu_{\underline{A}} d\nu(\underline{A})$ is a well-defined probability measure on Σ with respect to the σ-algebra generated by the cylinder sets. We conclude this section by observing that $\eta(\widehat{W}(\mathbf{x},\rho)) = 1$. In fact, we actually have the stronger statement that $\eta\left(\bigcup_{\underline{A} \in \Xi} \Gamma_{\underline{A}}\right) = 1$ since

$$\int_\Xi \mu_{\underline{A}}(\bigcup_{\underline{A} \in \Xi} \Gamma_{\underline{A}}) d\nu(\underline{A}) = \int_\Xi \mu_{\underline{A}}(\Gamma_{\underline{A}}) d\nu(\underline{A}) = \int_\Xi 1 d\nu(\underline{A}) = 1.$$

The conclusion that $\eta(\widehat{W}(\mathbf{x},\rho)) = 1$ follows from the fact that $\eta\left(\bigcup_{\underline{A} \in \Xi} \Gamma_{\underline{A}}\right) = 1$ upon recalling that $\Gamma_{\underline{A}} \subseteq C_{\underline{A}} \subseteq \widehat{W}(\mathbf{x},\rho)$ for every $\underline{A} \in \Xi$.

7. Proof of Proposition 5.2

Before we turn to the proof, we give estimates for $\mu_{\underline{A}}([\mathbf{i}|_1^k])$.

Lemma 7.1. *For every $\underline{A} \in \Xi$ and every $\mathbf{i} \in \Gamma_{\underline{A}}$,*

$$\mu_{\underline{A}}([\mathbf{i}|_1^k]) \ll \begin{cases} \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \dfrac{C^{(1+s)\ell}a_{\min}^{-2\ell s}e^{-s\alpha(A_\ell-\sum_{j=1}^{\ell-1}\rho(A_j)-2\ell)}}{\|f'_{\mathbf{x}|_1^{k-A_\ell}}(x)\|^s \prod_{j=1}^{\ell-1}\|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s} & \text{if } A_\ell \leqslant k < A_\ell + \rho(A_\ell), \\[3ex] \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \dfrac{C^{(1+s)\ell}a_{\min}^{-2\ell s}e^{-s\alpha(k-\sum_{j=1}^{\ell}\rho(A_j)-2\ell)}}{\prod_{j=1}^{\ell}\|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s} & \text{if } A_\ell + \rho(A_\ell) \leqslant k < A_{\ell+1}, \end{cases}$$

where $C \geqslant 1$ is the constant taken here to be the maximum of the constants appearing in (2.2) and (2.6).

Note that it follows from Lemma 6.1(1) and 6.1(2) that $A_\ell - \sum_{j=1}^{\ell-1} \rho(A_j) - 2\ell > 0$ and so the exponent of e is negative.

Proof. First suppose that $A_\ell \leqslant k < A_\ell + \rho(A_\ell)$. Then $\mathbf{i}|_{A_\ell}^k = \omega_\ell(A_\ell)\mathbf{x}|_1^{k-A_\ell}$ (see proof of Lemma 6.2) and thus,

$$\mu_{\underline{A}}([\mathbf{i}|_1^k]) = \prod_{j=1}^{\ell-1} \mathbb{S}_{A_j+\rho(A_j),A_{j+1}}([\mathbf{i}|_{A_j+\rho(A_j)+2}^{A_{j+1}-1}]).$$

By Lemma 6.5 we have $\prod_{j=1}^{\ell-1} \mathbb{P}_{A_j+\rho(A_j),A_{j+1}}(\Omega_{A_j+\rho(A_j),A_{j+1}}) \geqslant Q > 0$, and hence

$$\mu_{\underline{A}}([\mathbf{i}|_1^k]) \leqslant Q^{-1} \prod_{j=1}^{\ell-1} \mathbb{P}_{A_j+\rho(A_j),A_{j+1}}([\mathbf{i}|_{A_j+\rho(A_j)+2}^{A_{j+1}-1}]).$$

By (2.6) and since $A_1 \leqslant m_1$ and (1.1),

$$\mu_{\underline{A}}([\mathbf{i}|_1^k]) \leqslant \frac{C^\ell \|f'_{\mathbf{i}|_1^{A_1-1}}(\xi_0)\|^s e^{-\alpha s(A_1-1)}}{Q(a_{\min}^s e^{-\alpha s})^{m_1}} \prod_{j=1}^{\ell-1} \|f'_{\mathbf{i}|_{A_j+\rho(A_j)+2}^{A_{j+1}-1}}(\xi_j)\|^s e^{-\alpha s(A_{j+1}-A_j-\rho(A_j)-2)},$$

where $\xi_0, \cdots, \xi_{\ell-1} \in X$ are arbitrary. So one can choose $\xi_j = f'_{\mathbf{i}|_{A_j-1}^k}(x)$, for $j = 0\ldots,\ell-1$ and thus, by the chain rule and (2.2), we have

$$\|f'_{\mathbf{i}|_1^{A_1-1}}(\xi_0)\| \prod_{j=1}^{\ell-1} \|f'_{\mathbf{i}|_{A_j+\rho(A_j)+2}^{A_{j+1}-2}}(\xi_j)\|$$

$$= \|f'_{\mathbf{i}|_1^k}(x)\| \|f'_{\omega_\ell(A_\ell)\mathbf{x}|_1^{k-A_\ell}}(x)\|^{-1} \prod_{j=1}^{\ell-1} \|f'_{\omega_j(A_j)\mathbf{x}|_1^{\rho(A_j)}\tau_j(A_j)}(f_{\mathbf{i}|_{A_j+\rho(A_j)+2}^k}(x))\|^{-1}$$

$$\leqslant \|f'_{\mathbf{i}|_1^k}(x)\| C^{\ell-1} a_{\min}^{-(2\ell-1)} \|f'_{\mathbf{x}|_1^{k-A_\ell}}(x)\|^{-1} \prod_{j=1}^{\ell-1} \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^{-1},$$

where in the last inequality we used (2.2) and (1.2). Thus,

$$\mu_{\underline{A}}([\mathbf{i}|_1^k]) \ll \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \cdot C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(A_\ell-\sum_{j=1}^{\ell-1}\rho(A_j)-2\ell)} \|f'_{\mathbf{x}|_1^{k-A_\ell}}(x)\|^{-s} \prod_{j=1}^{\ell-1} \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^{-s}.$$

Now, consider the case when $A_\ell + \rho(A_\ell) \leqslant k < A_{\ell+1}$. Then, similarly to the previous argument, we have

$$\mu_{\underline{A}}([\mathbf{i}|_1^k]) = \left(\prod_{j=1}^{\ell-1} \mathbb{S}_{A_j+\rho(A_j),A_{j+1}}([\mathbf{i}|_{A_j+\rho(A_j)+2}^{A_{j+1}-1}])\right) \cdot \mathbb{S}_{A_\ell+\rho(A_\ell),A_{\ell+1}}([\mathbf{i}|_{A_\ell+\rho(A_\ell)+2}^k])$$

$$\ll \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \cdot C^{(1+s)\ell} a_{\min}^{-2s\ell} e^{-s\alpha(k-\sum_{j=1}^\ell \rho(A_j)-2\ell)} \prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^{-s}.$$

$\square$

Lemma 7.2. *Let $\underline{A} \in \Xi$ and $\mathbf{i} \in \Gamma_{\underline{A}}$ be arbitrary. Then, for every $k \geqslant n_1$,*

$$
\mu_{\underline{A}}([\mathbf{i}|_1^k]) \ll \begin{cases} \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \dfrac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(A_\ell - \sum_{j=1}^{\ell-1}\rho(A_j) - 2\ell)}}{\prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s} & \text{if } A_\ell \leqslant k < n_{\ell+1} \\[4ex] \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \dfrac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(n_{\ell+1} - \sum_{j=1}^{\ell}\rho(A_j) - 2\ell)}}{\prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s} & \text{if } n_{\ell+1} \leqslant k < A_{\ell+1}, \end{cases}
$$

where $C \geqslant 1$ is as in Lemma 7.1.

Proof. Let $\ell \geqslant 1$ be such that $A_\ell \leqslant k < A_{\ell+1}$. If $A_\ell \leqslant k < n_{\ell+1}$ then either $k < A_\ell + \rho(A_\ell)$ or $k \geqslant A_\ell + \rho(A_\ell)$. If $A_\ell \leqslant k < A_\ell + \rho(A_\ell)$ then, by Lemma 7.1,

$$
\mu_{\underline{A}}([\mathbf{i}|_1^k]) \ll \|f'_{\mathbf{i}|_1^k}(\xi)\|^s C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(A_\ell - \sum_{j=1}^{\ell-1}\rho(A_j) - 2\ell)} \cdot \prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^{-s}.
$$

If $A_\ell + \rho(A_\ell) \leqslant k < n_{\ell+1}$ then, by Lemma 7.1 again,

$$
\mu_{\underline{A}}([\mathbf{i}|_1^k]) \ll \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(k - \sum_{j=1}^{\ell}\rho(A_j) - 2\ell)}}{\prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s}
$$

$$
\leqslant \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(A_\ell - \sum_{j=1}^{\ell-1}\rho(A_j) - 2\ell)}}{\prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s}.
$$

On the other hand, if $n_{\ell+1} \leqslant k < A_{\ell+1}$ then $k \geqslant n_{\ell+1} > m_\ell + \rho(m_\ell) \geqslant A_\ell + \rho(A_\ell)$. Thus, in this case, by Lemma 7.1,

$$
\mu_{\underline{A}}([\mathbf{i}|_1^k]) \ll \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(k - \sum_{j=1}^{\ell}\rho(A_j) - 2\ell)}}{\prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s}
$$

$$
\leqslant \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(n_{\ell+1} - \sum_{j=1}^{\ell}\rho(A_j) - 2\ell)}}{\prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s}.
$$

$\square$

Finally, we turn to the proof of our main proposition.

Proof of Proposition 5.2. Let $\delta > 0$ be arbitrary but fixed. One can find $L \geqslant 1$ such that all the terms in Lemma 6.1(3) and Lemma 6.1(4) are smaller than δ for every $\ell \geqslant L$. Let us choose $K := n_L + 1$.

Let $k \geqslant K$ and $\mathbf{j} \in \Lambda^k$ be arbitrary. It can be seen that if $[\mathbf{j}] \cap \bigcup_{\underline{A} \in \Xi} \Gamma_{\underline{A}} = \varnothing$ then $\eta([\mathbf{j}]) = 0$ and thus, the bound holds trivially. So, without loss of generality, we may assume that $[\mathbf{j}] \cap \bigcup_{\underline{A} \in \Xi} \Gamma_{\underline{A}} \neq \varnothing$ and pick $\mathbf{i} \in \bigcup_{\underline{A} \in \Xi} \Gamma_{\underline{A}}$ such that $\mathbf{i}|_1^k = \mathbf{j}$.

There are two possible cases to consider: either $m_\ell \leqslant k < n_{\ell+1}$ or $n_\ell \leqslant k < m_\ell$ for some $\ell \geqslant 1$.

$\underline{\text{Case 1: } m_\ell \leqslant k < n_{\ell+1}}$

If $m_\ell \leqslant k < n_{\ell+1}$ then by the pairwise disjointness of the sets $\{\Gamma_{\underline{A}}\}_{\underline{A} \in \Xi}$ there exists a unique sequence $A_1, \ldots, A_\ell$ such that $\sigma^{A_j}\mathbf{i} \in [\mathbf{x}|_1^{\rho(A_j)}]$ for $j = 1, \ldots, \ell$, and for every

sequence $\underline{A}^* \in \Xi$ such that $A_j^* \neq A_j$ for some $j = 1, \ldots, \ell$, we have $[\mathbf{i}|_1^k] \cap \Gamma_{\underline{A}^*} = \varnothing$. Moreover, $A_\ell \leqslant k < n_{\ell+1}$. So, by Lemma 7.2,

$$
\begin{aligned}
\eta([\mathbf{i}|_1^k]) &= \int_\Xi \mu_{\underline{A}'}([\mathbf{i}|_1^k]) d\nu(\underline{A}') = \int_{[A_1,\ldots,A_\ell]} \mu_{\underline{A}'}([\mathbf{i}|_1^k]) d\nu(\underline{A}') \\
&\ll \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(A_\ell - \sum_{j=1}^{\ell-1}\rho(A_j) - 2\ell)}}{\prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s} \nu([A_1,\ldots,A_\ell]) \\
&= \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(A_\ell - \sum_{j=1}^{\ell-1}\rho(A_j) - 2\ell)}}{\prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s} \frac{\prod_{j=1}^\ell \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s e^{\alpha s A_j}}{\prod_{j=1}^\ell \sum_{k=n_j}^{m_j} \varepsilon(k)} \\
&= \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{s\alpha(\sum_{j=1}^{\ell-1}(A_j + \rho(A_j)) + 2\ell)}}{\prod_{j=1}^\ell \sum_{k=n_j}^{m_j} \varepsilon(k)} \\
&\leqslant \delta \cdot \|f'_{\mathbf{i}|_1^k}(\xi)\|^s.
\end{aligned}
$$

The final inequality above follows from Lemma 6.1(3) and our choice of δ.

Case 2: $n_\ell \leqslant k < m_\ell$

If $n_\ell \leqslant k < m_\ell$ then again by the pairwise disjointness of the sets $\{\Gamma_{\underline{A}}\}_{\underline{A} \in \Xi}$ there exists a unique sequence $A_1, \ldots, A_{\ell-1}$ such that $\sigma^{A_j}\mathbf{i} \in [\mathbf{x}|_1^{\rho(A_j)}]$ for $j = 1, \ldots, \ell-1$, and for every sequence $\underline{A}^* \in \Xi$ such that $A_j^* \neq A_j$ for some $j = 1, \ldots, \ell-1$, we have $[\mathbf{i}|_1^k] \cap \Gamma_{\underline{A}^*} = \varnothing$. Moreover, there is at most one $n_\ell \leqslant B \leqslant R(k)$ such that $[\mathbf{i}|_1^k] \cap \Gamma_{A_1,\ldots,A_{\ell-1},B,\ldots} \neq \varnothing$ and in that case $[\mathbf{i}|_1^k] \cap \Gamma_{A_1,\ldots,A_{\ell-1},B',\ldots} = \varnothing$ for any $B' \neq B$. To see this, note that if $B > R(k)$ then, by definition, $\rho(B) + B > k$ and so it is impossible for $\mathbf{i}|_1^k$ to contain $\mathbf{x}|_1^{\rho(B)}$ completely. Motivated by this, we decompose $[n_\ell, m_\ell)$ into three parts with respect to k:

- $[n_\ell, R(k)]$,

- $[R(k) + 1, k)$, where $\mathbf{i}|_1^k$ may contain part, but not all, of $\mathbf{x}|_1^{\rho(B)}$, and

- $[k, m_\ell)$, in which case $\mathbf{i}|_1^k$ does not contain any of $\mathbf{x}|_1^{\rho(B)}$.

First, suppose that there exists a $B \in [n_\ell, R(k)]$ such that $[\mathbf{i}|_1^k] \cap \Gamma_{A_1,\ldots,A_{\ell-1},B,\ldots} \neq \varnothing$. Then, by Lemma 7.2,

$$
\begin{aligned}
\eta([\mathbf{i}|_1^k]) &= \int_\Xi \mu_{\underline{A}'}([\mathbf{i}|_1^k]) d\nu(\underline{A}') \\
&= \int_{[A_1,\ldots,A_{\ell-1},B]} \mu_{\underline{A}'}([\mathbf{i}|_1^k]) d\nu(\underline{A}') \\
&\ll \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(B - \sum_{j=1}^{\ell-1}\rho(A_j) - 2\ell)}}{\|f'_{\mathbf{x}|_1^{\rho(B)}}(x)\|^s \prod_{j=1}^{\ell-1} \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s} \frac{\varepsilon(B) \prod_{j=1}^{\ell-1} \varepsilon(A_j)}{\prod_{j=1}^\ell \sum_{k=n_j}^{m_j} \varepsilon(k)} \\
&= \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{s\alpha(\sum_{j=1}^{\ell-1}(A_j + \rho(A_j)) + 2\ell)}}{\prod_{j=1}^\ell \sum_{k=n_j}^{m_j} \varepsilon(k)} \\
&\leqslant \delta \cdot \|f'_{\mathbf{i}|_1^k}(\xi)\|^s,
\end{aligned}
$$

where, again, the last inequality follows from Lemma 6.1(3) and our choice of δ.

If for every $B \in [n_\ell, R(k)]$ we have $[\mathbf{i}|_1^k] \cap \Gamma_{A_1,\dots,A_{\ell-1},B,\dots} = \varnothing$, then

$$\eta([\mathbf{i}|_1^k]) = \int_\Xi \mu_{\underline{A}'}([\mathbf{i}|_1^k]) d\nu(\underline{A}') = \sum_{B=R(k)+1}^{m_\ell} \int_{[A_1,\dots,A_{\ell-1},B]} \mu_{\underline{A}}([\mathbf{i}|_1^k]) d\nu(\underline{A}).$$

First, we give an estimate for the part $B = k, \dots, m_\ell$. By Lemma 7.2 we have

$$\sum_{B=k}^{m_\ell} \int_{[A_1,\dots,A_{\ell-1},B]} \mu_{\underline{A}}([\mathbf{i}|_1^k]) d\nu(\underline{A})$$

$$\ll \sum_{B=k}^{m_\ell} \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(n_\ell - \sum_{j=1}^{\ell-1}\rho(A_j) - 2\ell)}}{\prod_{j=1}^{\ell-1} \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s} \frac{\varepsilon(B) \prod_{j=1}^{\ell-1} \varepsilon(A_j)}{\prod_{j=1}^{\ell} \sum_{k=n_j}^{m_j} \varepsilon(k)}$$

$$= \sum_{B=k}^{m_\ell} \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(n_\ell - \sum_{j=1}^{\ell-1}(A_j + \rho(A_j)) - 2\ell)} \varepsilon(B)}{\prod_{j=1}^{\ell} \sum_{k=n_j}^{m_j} \varepsilon(k)}$$

$$= \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(n_\ell - \sum_{j=1}^{\ell-1}(A_j + \rho(A_j)) - 2\ell)}}{\prod_{j=1}^{\ell-1} \sum_{k=n_j}^{m_j} \varepsilon(k)} \frac{\sum_{B=k}^{m_\ell} \varepsilon(B)}{\sum_{k=n_\ell}^{m_\ell} \varepsilon(k)}$$

$$\leqslant \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(n_\ell - \sum_{j=1}^{\ell-1}(A_j + \rho(A_j)) - 2\ell)}}{\prod_{j=1}^{\ell-1} \sum_{k=n_j}^{m_j} \varepsilon(k)}$$

$$\leqslant \delta \cdot \|f'_{\mathbf{i}|_1^k}(\xi)\|^s.$$

The last inequality above follows from Lemma 6.1(4) and our choice of δ.

Finally, to estimate the part corresponding to $B = R(k), \dots, k-1$, note that we have $B < k < B + \rho(B)$. Also note that it can be shown via the chain rule and the bounded distortion property (2.2) that $\|f'_{\mathbf{x}|_1^{k-B}}(x)\|^{-s} \|f'_{\mathbf{x}|_1^{\rho(B)}}(x)\|^s = \|f'_{\mathbf{x}|_{k-B+1}^{\rho(B)}}(x)\|^s$. Hence, using Lemma 7.1 directly, we see that

$$\sum_{B=R(k)+1}^{k-1} \int_{[A_1,\dots,A_{\ell-1},B]} \mu_{\underline{A}}([\mathbf{i}|_1^k]) d\nu(\underline{A})$$

$$\ll \sum_{B=R(k)+1}^{k-1} \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{-s\alpha(B - \sum_{j=1}^{\ell-1}\rho(A_j) - 2\ell)} \cdot \|f'_{\mathbf{x}|_1^{k-B}}(x)\|^{-s}}{\prod_{j=1}^{\ell-1} \|f'_{\mathbf{x}|_1^{\rho(A_j)}}(x)\|^s} \frac{\varepsilon(B) \prod_{j=1}^{\ell-1} \varepsilon(A_j)}{\prod_{j=1}^{\ell} \sum_{k=n_j}^{m_j} \varepsilon(k)}$$

$$= \sum_{B=R(k)+1}^{k-1} \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \frac{C^{(1+s)\ell} a_{\min}^{-2\ell s} e^{s\alpha(\sum_{j=1}^{\ell-1}(A_j + \rho(A_j)) + 2\ell)} \cdot \|f'_{\mathbf{x}|_{k-B+1}^{\rho(B)}}(x)\|^s}{\prod_{j=1}^{\ell} \sum_{k=n_j}^{m_j} \varepsilon(k)}.$$

Thus, it follows from Lemma 6.1(3) and our choice of δ, combined with (1.2), that

$$
\begin{aligned}
\sum_{B=R(k)+1}^{k-1} \int_{[A_1,\dots,A_{\ell-1},B]} \mu_{\underline{A}}([\mathbf{i}|_1^k]) d\nu(\underline{A}) &\ll \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \cdot \delta \sum_{B=R(k)+1}^{k-1} (a_{\max})^{s(\rho(B)+B-k)} \\
&\leqslant \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \delta \sum_{B=0}^{\infty} (a_{\max})^{s(\rho(R(k)+B+1)+R(k)+1+B-k)} \\
&\leqslant \|f'_{\mathbf{i}|_1^k}(\xi)\|^s \delta \sum_{B=0}^{\infty} (a_{\max})^{Bs} \\
&\ll \delta \|f'_{\mathbf{i}|_i^k}(\xi)\|^s.
\end{aligned}
$$

The proof of Proposition 5.2 is complete upon noting that $\|f'_{\mathbf{i}|_1^k}(\xi)\| \leqslant C \operatorname{diam}(X)^{-1}\operatorname{diam}(X_{\mathbf{i}|_1^k})$ by (2.3). $\qquad\square$

8. An example: Badly Approximable Numbers and Quadratic Irrationals

In this section we discuss an application of our main theorem (Theorem 1.1) to the problem of approximating badly approximable numbers by quadratic irrationals. Previously, Baker demonstrated the existence of badly approximable numbers which are "very well approximable" by quadratic irrationals [1]. Here, we extend this result by showing that Theorem 1.1 can be applied to obtain a Jarník-type statement for the set of badly approximable numbers which are "well-approximable" by a fixed quadratic irrational. We achieve this by utilising the correspondence between badly approximable numbers and partial quotients of continued fraction expansions, and by expressing the numbers with continued fraction expansions with partial quotients bounded by $M \in \mathbb{N}$ as the attractor of a conformal iterated function system.

Recall that a number $x \in [0,1]$ is said to be *badly approximable* if there exists a constant $c = c(x) > 0$, dependent on x, such that for every $\frac{p}{q} \in \mathbb{Q}$ we have

$$
\left| x - \frac{p}{q} \right| > \frac{c}{q^2}.
$$

It is well known that a number is badly approximable if and only if its continued fraction expansion has bounded partial quotients (see, for example, [5, Theorem 1.15]).

For $x \in [0,1]$, let $[a_1, a_2, \dots]$ denote its continued fraction expansion. Note that this expansion will be finite if x is rational. Recall that the a_i's are called the *partial quotients* of x and are the numbers which arise when we write x in the form

$$
x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{\ddots}}}}
$$

with $a_i \in \mathbb{N}$ for each $i \in \mathbb{N}$. The partial quotients, a_i, can also be defined via the Gauss map. The *Gauss map* is the map $T : [0,1] \to [0,1]$ defined by

$$
T(x) = \begin{cases} 0 & \text{if } x = 0, \\ \frac{1}{x} - \left\lfloor \frac{1}{x} \right\rfloor & \text{for } x \in (0,1]. \end{cases}
$$

For each $n \geqslant 1$, $a_n = \left\lfloor \dfrac{1}{T^{n-1}(x)} \right\rfloor$.

For our present purposes, another useful way for us to view the continued fraction expansion of $x \in [0, 1]$ is the following. For every integer $a \geqslant 1$, let

$$f_a(y) = \frac{1}{a + y}.$$

We have

$$x = [a_1, a_2, \ldots] = \lim_{n \to \infty} f_{a_1} \circ \cdots \circ f_{a_n}(1).$$

If $x \in (0, 1]$ is badly approximable, then there exists $Q \geqslant 1$ such that $a_n \leqslant Q$ for every $n \geqslant 1$. Let us denote the set of badly approximable numbers in $(0, 1]$ by **Bad**, and denote by $\mathbf{Bad}_Q$ the numbers $x \in \mathbf{Bad}$ such that $a_n(x) \leqslant Q$ for every $n \geqslant 1$ (where $a_n(x)$ is the nth partial quotient of x). By definition, $\mathbf{Bad}_Q$ is the attractor of the IFS $\{f_a \circ f_b\}_{a,b=1}^Q$. Moreover, it is easy to see that this IFS is conformal and satisfies the open set condition.

Recall that $x \in [0, 1]$ is a *quadratic irrational* if it is irrational and is a root of a quadratic polynomial $ax^2 + bx + c = 0$, where $a \neq 0$ and a, b, c are integers. It is well known that $x \in [0, 1]$ is a quadratic irrational if and only if x has an *eventually periodic* continued fraction expansion; that is, there exist finite sequences $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_\ell)$ and $\boldsymbol{\tau} = (\tau_1, \ldots, \tau_m)$ such that

$$x = [\omega_1, \ldots, \omega_\ell, \overline{\tau_1, \ldots, \tau_m}],$$

where $\overline{\boldsymbol{\tau}}$ denotes the infinite periodic sequence formed by repeating $\boldsymbol{\tau}$. We will denote the set of quadratic irrationals in $[0, 1]$ by **QI**.

Given a monotonically decreasing approximating function $\psi \colon \mathbb{N} \mapsto \mathbb{R}^+$, let

$$W(\psi; \mathbf{Bad}; \mathbf{QI}) := \{x \in \mathbf{Bad} : \exists \alpha \in \mathbf{QI} : |T^n(x) - \alpha| < \psi(n) \text{ for infinitely many } n \in \mathbb{N}\}.$$

Thus, $W(\psi; \mathbf{Bad}; \mathbf{QI})$ is the set of badly approximable numbers which are "well-approximable" by a fixed quadratic irrational. We will investigate the Hausdorff measure of $W(\psi; \mathbf{Bad}; \mathbf{QI})$. Notice that it is sensible to fix the quadratic irrational in the definition of $W(\psi; \mathbf{Bad}; \mathbf{QI})$ corresponding to a given $x \in \mathbf{Bad}$, otherwise we would necessarily have $W(\psi; \mathbf{Bad}; \mathbf{QI}) = \mathbf{Bad}$ since quadratic irrationals are dense in the reals.

Let us now adapt some standard notation from the usual theory of continued fractions. For the proofs and more details, see [27, Section 2] or [24].

For a sequence of integers $\{a_n\}_{n=1}^{\infty}$, let

$$q_{n+1}(a_1, \ldots, a_{n+1}) := q_{n+1} = a_{n+1}q_n + q_{n-1}, \text{ and}$$
$$p_{n+1}(a_1, \ldots, a_{n+1}) := p_{n+1} = a_{n+1}p_n + p_{n-1}$$

for $n \geqslant 1$, where we define $p_{-1} = q_0 = 1$ and $p_0 = q_{-1} = 0$. Then,

$$f_{a_1} \circ \cdots \circ f_{a_n}(x) = \frac{p_{n-1}x + p_n}{q_{n-1}x + q_n}.$$

Moreover, for every $0 < k < n$ and $a_1, \ldots, a_n \in \mathbb{N}$ we have

$$1 \leqslant \frac{q_n(a_1, \ldots, a_n)}{q_k(a_1, \ldots, a_k)q_{n-k}(a_{k+1}, \ldots, a_n)} \leqslant 2, \tag{8.1}$$

and, for every $x \in [0, 1]$,

$$\frac{1}{4q_n^2} \leqslant |(f_{a_1} \circ \cdots \circ f_{a_n})'(x)| \leqslant \frac{1}{q_n^2}.$$

Thus, combining the above bounds with the bounded distortion property (2.2), there exists a constant $K \geqslant 1$, depending on Q but independent of the sequence $a_1, a_2, \ldots, a_n$, such

that

$$\frac{1}{K} \times \frac{1}{q_n(a_1, \ldots, a_n)^2} \leqslant \mathrm{diam}(f_{a_1} \circ \cdots \circ f_{a_n}(\mathbf{Bad}_Q)) \leqslant \frac{K}{q_n(a_1, \ldots, a_n)^2}. \qquad (8.2)$$

Now, we are ready to state a corollary of our main result.

Theorem 8.1. *Let $W(\psi; \mathbf{Bad}; \mathbf{QI})$ be the set defined above. We have*

$$\mathcal{H}^s(W(\psi; \mathbf{Bad}; \mathbf{QI})) = \begin{cases} 0 & \text{if, } \forall\, Q \geqslant 1,\ \sum_{n=1}^{\infty} \sum_{a_1, \ldots, a_n = 1}^{Q} \frac{\psi(n)^s}{q_n(a_1, \ldots, a_n)^{2s}} < \infty, \\[2em] \infty & \text{if } \exists\, Q \geqslant 1 \text{ such that } \sum_{n=1}^{\infty} \sum_{a_1, \ldots, a_n = 1}^{Q} \frac{\psi(n)^s}{q_n(a_1, \ldots, a_n)^{2s}} = \infty. \end{cases}$$

Proof. If $x \in W(\psi; \mathbf{Bad}; \mathbf{QI})$ then there exists $\alpha = \alpha(x) \in \mathbf{QI}$ such that $|T^n(x) - \alpha| < \psi(n)$ for infinitely many $n \in \mathbb{N}$. On the other hand, there exists $Q_1 \geqslant 1$ such that $x \in \mathbf{Bad}_{Q_1}$ and since α is quasi-periodic there exists $Q_2 \geqslant 1$ such that $\alpha \in \mathbf{Bad}_{Q_2}$. Hence, $x \in \mathbf{Bad}_Q$ and $\alpha \in \mathbf{QI} \cap \mathbf{Bad}_Q$, where $Q = \max\{Q_1, Q_2\}$. Thus, $x \in W_Q(\psi; \alpha)$ where

$$W_Q(\psi; \alpha) = \{x \in \mathbf{Bad}_Q : |T^n(x) - \alpha| < \psi(n) \text{ for infinitely many } n \in \mathbb{N}\}.$$

So, we have that

$$W(\psi; \mathbf{Bad}; \mathbf{QI}) = \bigcup_{Q=1}^{\infty} \bigcup_{\alpha \in \mathbf{QI} \cap \mathbf{Bad}_Q} W_Q(\psi; \alpha). \qquad (8.3)$$

It is easy to see that if $x \in W_Q(\psi; \alpha)$ then we must have $|T^{2n}(x) - \alpha| < \psi(2n)$ for infinitely many $n \in \mathbb{N}$ or $|T^{2n+1}(x) - \alpha| < \psi(2n + 1)$ for infinitely many $n \in \mathbb{N}$. Hence, we can decompose $W_Q(\psi; \alpha)$ into

$$W_Q(\psi; \alpha) = U_Q(\psi_0; \alpha) \cup T(U_Q(\psi_1; \alpha)),$$

where $\psi_0(n) = \psi(2n)$ and $\psi_1(n) = \psi(2n + 1)$ and

$$U_Q(\psi_i; \alpha) = \{x \in \mathbf{Bad}_Q : |T^{2n}(x) - \alpha| < \psi_i(n) \text{ for infinitely many } n \in \mathbb{N}\}.$$

Since the set $\mathbf{Bad}_Q$ is the attractor of the conformal IFS $\{f_a \circ f_b\}_{a,b=1}^{Q}$, it follows from Theorem 1.1 taken together with (8.2) that, for any $\alpha \in \mathbf{QI} \cap \mathbf{Bad}_Q$, we have

$$\mathcal{H}^s(U_Q(\psi_i, \alpha)) = \begin{cases} 0 & \text{if } \sum_{n=1}^{\infty} \sum_{a_1=1, \ldots, a_{2n}=1}^{Q} \frac{\psi_i(n)^s}{q_{2n}(a_1, \ldots, a_{2n})^{2s}} < \infty \\[1.5em] \infty & \text{if } \sum_{n=1}^{\infty} \sum_{a_1=1, \ldots, a_{2n}=1}^{Q} \frac{\psi_i(n)^s}{q_{2n}(a_1, \ldots, a_{2n})^{2s}} = \infty. \end{cases} \qquad (8.4)$$

On the other hand,

$$\sum_{n=1}^{\infty} \sum_{a_1, \ldots, a_n = 1}^{Q} \frac{\psi(n)^s}{q_n(a_1, \ldots, a_n)^{2s}} = \sum_{n=1}^{\infty} \sum_{a_1, \ldots, a_{2n} = 1}^{Q} \frac{\psi(2n)^s}{q_{2n}(a_1, \ldots, a_{2n})^{2s}} + \sum_{n=0}^{\infty} \sum_{a_1, \ldots, a_{2n+1} = 1}^{Q} \frac{\psi(2n + 1)^s}{q_{2n+1}(a_1, \ldots, a_{2n+1})^{2s}}.$$

By (8.1),

$$\sum_{n=0}^{\infty} \sum_{a_1, \ldots, a_{2n+1} = 1}^{Q} \frac{\psi(2n + 1)^s}{q_{2n+1}(a_1, \ldots, a_{2n+1})^{2s}} \leqslant 2 \sum_{n=0}^{\infty} \sum_{a_1, \ldots, a_{2n+1} = 1}^{Q} \frac{\psi(2n + 1)^s}{q_1(a_{2n+1})^{2s} q_{2n}(a_1, \ldots, a_{2n})^{2s}}$$

$$= 2 \left(\sum_{a=1}^{Q} q_1(a)^{-2s} \right) \sum_{n=0}^{\infty} \sum_{a_1, \ldots, a_{2n} = 1}^{Q} \frac{\psi(2n + 1)^s}{q_{2n}(a_1, \ldots, a_{2n})^{2s}},$$

and the other inequality

$$\sum_{n=0}^{\infty} \sum_{a_1,\ldots,a_{2n+1}=1}^{Q} \frac{\psi(2n+1)^s}{q_{2n+1}(a_1,\ldots,a_{2n+1})^{2s}} \geqslant \left(\sum_{a=1}^{Q} q_1(a)^{-2s}\right) \sum_{n=0}^{\infty} \sum_{a_1,\ldots,a_{2n}=1}^{Q} \frac{\psi(2n+1)^s}{q_{2n}(a_1,\ldots,a_{2n})^{2s}}$$

follows by similar argument. Thus, $\displaystyle\sum_{n=1}^{\infty} \sum_{a_1,\ldots,a_n=1}^{Q} \frac{\psi(n)^s}{q_n(a_1,\ldots,a_n)^{2s}}$ is finite if and only if

$$\sum_{n=1}^{\infty} \sum_{a_1=1,\ldots,a_{2n}=1}^{Q} \frac{\psi_i(n)^s}{q_{2n}(a_1,\ldots,a_{2n})^{2s}} < \infty$$

for $i = 0, 1$. Then the statement follows by (8.3) and (8.4). $\qquad\square$

References

[1] S. Baker. An analogue of Khintchine's theorem for self-conformal sets. *Math. Proc. Cambridge Philos. Soc.*, 167 (2019), no. 3, 567–597.

[2] S. Baker. Overlapping iterated function systems from the perspective of metric number theory Memoirs of the AMS (to appear), *arXiv preprint arXiv:1901.07875*.

[3] S. Baker and M. Farmer. Quantitative recurrence properties for self-conformal sets. *arXiv preprint arXiv:1909.08913*, 2019, to appear in Proc. Amer. Math. Soc.

[4] B. Bárány and M. Rams. Shrinking targets on Bedford-McMullen carpets. *Proc. Lond. Math. Soc. (3)*, 117(5):951–995, 2018.

[5] V. Beresnevich, F. Ramírez, and S. Velani. Metric Diophantine approximation: aspects of recent work. In *Dynamics and analytic number theory*, volume 437 of *London Math. Soc. Lecture Note Ser.*, pages 1–95. Cambridge Univ. Press, Cambridge, 2016.

[6] V. Beresnevich and S. Velani. A mass transference principle and the Duffin-Schaeffer conjecture for Hausdorff measures. *Ann. of Math. (2)*, 164(3):971–992, 2006.

[7] R. Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*, volume 470 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, revised edition, 2008. With a preface by David Ruelle, Edited by Jean-René Chazottes.

[8] Y. Chang, M. Wu, and W. Wu. Quantitative recurrence properties and homogeneous self-similar sets. *Proc. Amer. Math. Soc.*, 147(4):1453–1465, 2019.

[9] N. Chernov and D. Kleinbock. Dynamical Borel-Cantelli lemmas for Gibbs measures. *Israel J. Math.*, 122:1–27, 2001.

[10] M. Einsiedler and T. Ward. *Ergodic theory with a view towards number theory*, volume 259 of *Graduate Texts in Mathematics*. Springer-Verlag London, Ltd., London, 2011.

[11] K. Falconer. *Fractal geometry*. John Wiley & Sons, Inc., Hoboken, NJ, second edition, 2003. Mathematical foundations and applications.

[12] E. Hewitt and K. Stromberg. *Real and abstract analysis*. Springer-Verlag, New York-Heidelberg, 1975. A modern treatment of the theory of functions of a real variable, Third printing, Graduate Texts in Mathematics, No. 25.

[13] R. Hill and S. L. Velani. The ergodic theory of shrinking targets. *Invent. Math.*, 119(1):175–198, 1995.

[14] R. Hill and S. L. Velani. The shrinking target problem for matrix transformations of tori. *J. London Math. Soc. (2)*, 60(2):381–398, 1999.
[15] J. E. Hutchinson. Fractals and self-similarity. *Indiana Univ. Math. J.*, 30(5):713–747, 1981.
[16] V. Jarník. Uber die simultanen diophantischen approximationen. *Math. Z.*, 33(1):505–543, 1931.
[17] A. Khintchine, *Einige Sätze über Kettenbrüche, mit Anwendungen auf die Theorie der Diophantischen Approximationen*, Math. Ann. 92 (1924), 115–125.
[18] A. Khintchine, *Über die angenäherte Auflösung linearer Gleichungen in ganzen Zahlen*, Rec. Math. Soc. Moscou, 32 (1925), 203–218.
[19] E. Järvenpää, M. Järvenpää, H. Koivusalo, B. Li, V. Suomala, and Y. Xiao. Hitting probabilities of random covering sets in tori and metric spaces. *Electron. J. Probab.*, 22:Paper No. 1, 18, 2017.
[20] M. Järvenpää. Random covering sets, hitting probabilities and variants of the covering problem. In *Recent developments in fractals and related fields*, Trends Math., pages 175–187. Birkhäuser/Springer, Cham, 2017.
[21] A. Käenmäki and E. Rossi. Weak separation condition, Assouad dimension, and Furstenberg homogeneity. *Ann. Acad. Sci. Fenn. Math.*, 41(1):465–490, 2016.
[22] A. Käenmäki and M. Vilppolainen. Separation conditions on controlled Moran constructions. *Fund. Math.*, 200(1):69–100, 2008.
[23] A. B. Kharazishvili. *Strange functions in real analysis*, volume 272 of *Pure and Applied Mathematics (Boca Raton)*. Chapman & Hall/CRC, Boca Raton, FL, second edition, 2006.
[24] A. Y. Khintchine. *Continued fractions*. Translated by Peter Wynn. P. Noordhoff, Ltd., Groningen, 1963.
[25] H. Koivusalo and F. A. Ramírez. Recurrence to shrinking targets on typical self-affine fractals. *Proc. Edinb. Math. Soc. (2)*, 61(2):387–400, 2018.
[26] J. Levesley, C. Salp, and S. L. Velani. On a problem of K. Mahler: Diophantine approximation and Cantor sets. *Math. Ann.*, 338(1):97–118, 2007.
[27] B. Li, B.-W. Wang, J. Wu, and J. Xu. The shrinking target problem in the dynamical system of continued fractions. *Proc. Lond. Math. Soc. (3)*, 108(1):159–186, 2014.
[28] D. J. C. MacKay. *Information theory, inference and learning algorithms*. Cambridge University Press, New York, 2003.
[29] K. Mahler. Some suggestions for further research. *Bull. Austral. Math. Soc.*, 29(1):101–108, 1984.
[30] Y. Peres, M. Rams, K. Simon, and B. Solomyak. Equivalence of positive Hausdorff measure and the open set condition for self-conformal sets. *Proc. Amer. Math. Soc.*, 129(9):2689–2699, 2001.
[31] T. Persson and M. Rams. On shrinking targets for piecewise expanding interval maps. *Ergodic Theory Dynam. Systems*, 37(2):646–663, 2017.
[32] S. Seuret and B.-W. Wang. Quantitative recurrence properties in conformal iterated function systems. *Adv. Math.*, 280:472–505, 2015.
[33] K. Simon, B. Solomyak, and M. Urbański. Hausdorff dimension of limit sets for parabolic IFS with overlaps. *Pacific J. Math.*, 201(2):441–478, 2001.
[34] P. Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

(Demi Allen) SCHOOL OF MATHEMATICS, UNIVERSITY OF BRISTOL, FRY BUILDING, WOODLAND ROAD, BRISTOL, BS8 1UG, UK, AND THE HEILBRONN INSTITUTE FOR MATHEMATICAL RESEARCH, BRISTOL, UK

Email address: demi.allen@bristol.ac.uk

(Balázs Bárány) BUDAPEST UNIVERSITY OF TECHNOLOGY AND ECONOMICS, DEPARTMENT OF STOCHASTICS, MTA-BME STOCHASTICS RESEARCH GROUP, P.O.Box 91, 1521 BUDAPEST, HUNGARY

Email address: balubsheep@gmail.com