

SINGULAR VECTORS ON FRACTALS AND PROJECTIONS OF SELF-SIMILAR MEASURES

OSAMA KHALIL

ABSTRACT. Singular vectors are those for which the quality of rational approximations provided by Dirichlet’s Theorem can be improved by arbitrarily small multiplicative constants. We provide an upper bound on the Hausdorff dimension of singular vectors lying on self-similar fractals in \mathbb{R}^d satisfying the open set condition. The bound is in terms of quantities which are closely tied to Frostman exponents of projections of the Hausdorff measure supported on the fractal. Our bound is optimal in the sense that it agrees with the exact dimension of singular vectors obtained by Cheung and Chevallier when the fractal is trivial (i.e. has non-empty interior). As a corollary, we show that if the fractal is the product of 2 copies of Cantor’s middle thirds set or the attractor of a planar homogeneous irrational IFS, then the upper bound is $2/3$ the dimension of the fractal. This addresses the upper bound part of a question raised by Bugeaud, Cheung and Chevallier. We apply our method in the setting of translation flows on flat surfaces to show that the dimension of non-uniquely ergodic directions belonging to a fractal is at most $1/2$ the dimension of the fractal.

1. INTRODUCTION

1.1. **Statement of results.** The goal of this article is to study the dimension of singular vectors on fractals. The notion of singular vectors is motivated by Dirichlet’s theorem. It states that for every vector $\mathbf{x} \in \mathbb{R}^d$ and for every $N \in \mathbb{N}$, the inequalities

$$\|\mathbf{q} \cdot \mathbf{x} + p\| \leq \varepsilon/N, \quad \|\mathbf{q}\| \leq N^{1/d}, \quad (1.1)$$

admit a non-zero solution $(p, \mathbf{q}) \in \mathbb{Z} \times \mathbb{Z}^d$ when $\varepsilon = 1$. A vector $\mathbf{x} \in \mathbb{R}^d$ is then said to be **singular** if for every $0 < \varepsilon < 1$, there exists $N_0 \in \mathbb{N}$ so that the inequalities in (1.1) admit a non-zero integral solution for all $N \geq N_0$. We will denote by $\text{Sing}(d)$ the set of singular vectors in \mathbb{R}^d . It is well-known [DS70a, KW08] that the Lebesgue measure of $\text{Sing}(d)$ is 0.

In a remarkable article, the Hausdorff dimension of $\text{Sing}(2)$ was determined by Cheung [Che11]. This result was extended by Cheung and Chevallier in [CC16] to higher dimensions. They showed that for $d \geq 2$,

$$\dim_H(\text{Sing}(d)) = \frac{d^2}{d+1}, \quad (1.2)$$

where \dim_H denotes Hausdorff dimension. More recently, a sharp upper bound on the dimension of singular $(m \times n)$ -matrices was found in [KKLM17] and the exact lower bound was later determined in [DFSU19]. On the other hand, the study of Diophantine properties of fractals has attracted a lot of interest in recent years, beginning with the work of Kleinbock, Lindenstrauss, and Weiss in [KLW04]. This study is motivated by Sprindžuk’s conjecture, resolved in [KM98], and concerns finding optimal conditions on measures and subsets of \mathbb{R}^d under which they inherit the Diophantine properties of the ambient space. Subsequently, Bugeaud, Cheung, and Chevallier raised the natural question of determining the Hausdorff dimension of singular vectors on fractals in \mathbb{R}^d .

Question 1.1 (Problem 6, [BCC18]). *What is the dimension of the set of vectors in $\text{Sing}(2)$ whose coordinates belong to Cantor’s middle thirds set?*

2010 *Mathematics Subject Classification.* 11J13, 11J83, 28A80, 37A17.

Key words and phrases. singular vectors, fractals, Hausdorff dimension, translation surfaces.

In this article, we give an upper bound on the dimension of singular vectors belonging to a large class of fractals which arise as the limit sets of iterated function systems (IFS) of contractive similarities on \mathbb{R}^d . This class includes such familiar examples as products of the same Cantor set, Koch snowflakes, Sierpiński gaskets, etc. Our result is new even in the setting of Question 1.1. We refer the reader to Section 2 for detailed definitions. For prior work on Diophantine properties of fractals, see Section 1.4.

To state the result, we need some notation. For $1 \leq k \leq d$, we denote by $\mathcal{A}(d, k)$ the collection of all affine subspaces of \mathbb{R}^d of dimension k . For $\mathcal{L} \in \mathcal{A}(d, k)$ and $\varepsilon > 0$, we use $\mathcal{L}^{(\varepsilon)}$ to denote the open ε -neighborhood of \mathcal{L} in the Euclidean metric. Given a compactly supported Borel measure μ on \mathbb{R}^d and $1 \leq \ell \leq d$, we define $\alpha_\ell(\mu)$ by

$$\alpha_\ell(\mu) = \liminf_{\varepsilon \rightarrow 0} \frac{\log \sup_{\mathcal{L} \in \mathcal{A}(d, d-\ell)} \mu(\mathcal{L}^{(\varepsilon)})}{\log \varepsilon}. \quad (1.3)$$

The quantities $\alpha_\ell(\mu)$ quantify the concentration of the support of μ near proper subspaces: the smaller $\alpha_\ell(\mu)$ is, the more concentrated its support is near affine subspaces of dimension $d - \ell$. On the other hand, proper rational affine subspaces of \mathbb{R}^d are contained in $\text{Sing}(d)$. Our first main result shows that one can provide an upper bound on the dimension of singular vectors on fractals in terms of these quantities. It is a special case of Theorem 1.12 below.

Theorem A. *Suppose \mathcal{F} is an IFS consisting of contractive similarities on \mathbb{R}^d and satisfying the open set condition. Let \mathcal{K} be the limit set of \mathcal{F} and μ be the restriction of the s -dimensional Hausdorff measure to \mathcal{K} , where $s = \dim_H(\mathcal{K})$. Then,*

$$\dim_H(\text{Sing}(d) \cap \mathcal{K}) \leq s - \min_{1 \leq \ell \leq d} \frac{(d - \ell + 1)\alpha_\ell(\mu)}{d + 1}.$$

Remark 1.2. It is known that in the setting of Theorem A, μ is finite and non-zero, and $\alpha_d(\mu) = s$ (Proposition 2.2). When \mathcal{F} is irreducible, i.e. no proper affine subspace is invariant by all the maps in \mathcal{F} , then $\alpha_\ell(\mu) > 0$ for each ℓ (Corollary A.10). We also show in Proposition A.7 that for self-similar measures, the lim inf in the definition of $\alpha_\ell(\mu)$ is in fact a limit.

Remark 1.3. If μ is the Lebesgue measure on \mathbb{R}^d , then $\alpha_\ell(\mu) = \ell$. Moreover, one can realize the unit cube as the attractor of an IFS as in Theorem A. Since singular vectors are invariant by integer translations, Theorem A shows that $\text{Sing}(d)$ has dimension at most $d^2/(d + 1)$, which agrees with the exact dimension obtained by Cheung and Chevallier.

In the setting of Question 1.1, Theorem A yields the following.

Corollary 1.4. *Suppose $\mathcal{K} = \mathcal{C} \times \mathcal{C}$, where $\mathcal{C} \subset [0, 1]$ is Cantor's middle thirds set. Then,*

$$\dim_H(\text{Sing}(2) \cap \mathcal{K}) \leq \frac{2 \dim_H(\mathcal{K})}{3} = \frac{4 \log 2}{3 \log 3}.$$

The quantities $\alpha_\ell(\mu)$ are closely tied to Frostman exponents of projections of μ , cf. [Shm16, Section 1.3] and Section 1.3 below. In a breakthrough article, these exponents were determined by Shmerkin for planar homogeneous irrational fractals in [Shm16, Theorem 8.2]. This result, combined with Theorem 1.9 below, allows us to evaluate the formula in Theorem A explicitly, yielding Corollary 1.5. Recall that a planar IFS $\mathcal{F} = \{h_i : i \in \Lambda\}$ is **homogeneous** if each h_i is of the form $\rho R_\alpha + b_i$ where $0 < \rho < 1$ and R_α is a rotation by angle α , each of which is independent of i . We say \mathcal{F} is **irrational** if $\alpha \notin \mathbb{Q}\pi$.

Corollary 1.5. *Suppose \mathcal{F} is a homogeneous irrational IFS on \mathbb{R}^2 satisfying the open set condition and let \mathcal{K} be its limit set. Then,*

$$\dim_H(\text{Sing}(2) \cap \mathcal{K}) \leq \frac{2 \dim_H(\mathcal{K})}{3}.$$

For more general self-similar fractals, when the co-dimension of \mathcal{K} is < 1 , it is possible to get an explicit, yet crude, estimate on $\alpha_\ell(\mu)$, yielding the following corollary.

Corollary 1.6. *Suppose \mathcal{F} , \mathcal{K} , and μ are as in Theorem A. Assume further that $s = \dim_H(\mathcal{K}) > d - 1$. Then, $\alpha_\ell(\mu) \geq s - d + \ell$ and, hence,*

$$\dim_H(\text{Sing}(d) \cap \mathcal{K}) \leq s - \frac{d(s - d + 1)}{d + 1}.$$

Remark 1.7. Corollary 1.6 along with (1.2) show that $\dim_H(\text{Sing}(d) \cap \mathcal{K}) < \dim_H(\text{Sing}(d))$, under the hypotheses of Theorem A, whenever $\dim_H(\mathcal{K}) < d$.

We emphasize that we do not expect the estimated upper bound in Corollary 1.6 to agree with the exact lower bound in general. However, we conjecture that the upper bound in Theorem A (and, in particular, Corollary 1.4) is sharp when the fractal contains a dense set of rational vectors, cf. Question 1.11.

It is worth noting that computing Frostman exponents of projections of self-affine (and in particular self-similar) measures is a rather delicate problem in general. For instance, when μ is the natural measure supported on the product of 2 Cantor sets with multiplicatively independent dissection ratios, this problem constitutes the content of Furstenberg's well-known intersection conjecture, recently resolved by Shmerkin in [Shm16], and independently by Wu in [Wu16].

1.2. Divergent orbits of the Teichmüller flow. Under Dani's correspondence, it is known that $\text{Sing}(d)$ corresponds to certain divergent orbits on the space of unimodular lattices in \mathbb{R}^{d+1} , see Theorem 1.12 below for details. In Section 7, we adapt our techniques to the closely related problem of divergent orbits of the Teichmüller geodesic flow.

In what follows, we fix a stratum \mathcal{H} of abelian differentials over a compact oriented surface (see Section 7.1 for definitions). Then, \mathcal{H} admits a natural action by $\text{SL}(2, \mathbb{R})$ and, in particular, by the following one parameter subgroups:

$$a_t = \begin{pmatrix} e^t & 0 \\ 0 & e^{-t} \end{pmatrix}, \quad u(s) = \begin{pmatrix} 1 & s \\ 0 & 1 \end{pmatrix}, \quad r_\theta = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}. \quad (1.4)$$

The a_t action induces the Teichmüller geodesic flow on \mathcal{H} . For $\omega \in \mathcal{H}$, we say the orbit $(a_t \omega)_{t \geq 0}$ is *divergent on average*, if for every compact set $Q \subset \mathcal{H}$, one has

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \chi_Q(a_t \omega) dt = 0, \quad (1.5)$$

where χ_Q denotes the indicator function of Q .

Theorem B. *Suppose \mathcal{F} is an IFS consisting of similarities on \mathbb{R} satisfying the open set condition and let \mathcal{K} be its limit set. Then, for every $\omega \in \mathcal{H}$, the Hausdorff dimension of the set of $s \in \mathcal{K}$ such that the orbit $(a_t u(s) \omega)_{t \geq 0}$ diverges on average in \mathcal{H} is at most $\frac{\dim_H(\mathcal{K})}{2}$.*

Masur showed in [Mas92] that the set of directions θ around any point $\omega \in \mathcal{H}$ for which the orbit $(a_t r_\theta \omega)_{t \geq 0}$ is divergent have dimension at most $1/2$. This was recently extended in [AAE⁺17] to show that this upper bound in fact holds for divergent on average directions. Theorem B generalizes both these results. More recently, the lower bound of $1/2$ on the dimension of divergent on average directions was established in [AM18].

The motivation for studying divergent Teichmüller geodesics comes from the study of the ergodic properties of billiard flows and interval exchange transformations (IETs). Masur's criterion states that if the vertical straight line flow on $r_\theta \omega$ is non-uniquely ergodic (NUE), then $(a_t r_\theta \omega)_{t \geq 0}$ diverges in \mathcal{H} [Mas92]. In this vein, Theorem B has the following corollary.

Corollary 1.8. *Let \mathcal{K} be as in Theorem B. Then, for every $\omega \in \mathcal{H}$, the set of directions $\theta \in \arctan(\mathcal{K})$ such that the vertical flow on $r_\theta \omega$ is not uniquely ergodic has dimension at most $\frac{\dim_H(\mathcal{K})}{2}$.*

The size of the set of NUE directions has been extensively studied. A celebrated theorem of Kerckhoff, Masur, and Smilie shows that the Lebesgue measure of the set of NUE directions is 0 [KMS86]. This result was generalized by Veech in [Vee99] to a broader class of measures which includes Lebesgue and the natural measure on a Cantor set. Additionally, it is known that first return maps of straight line flows give rise to IETs. This observation can be packaged in the form of a locally defined map from a stratum of abelian differentials to the space of parameters of IETs with a given permutation (cf. [Mas82] for details). Minsky and Weiss characterized the straight lines in the space of IETs that arise as the image of orbits of the horocycle flow $U = \{u(s) : s \in \mathbb{R}\}$ on strata [MW14] under such maps and showed that non-uniquely ergodic IETs belonging to these special lines have 0 mass with respect to a broad class of measures

1.3. Relation to Frostman exponents of projections. In the Appendix, we study the relationship between the quantities $\alpha_\ell(\mu)$ and the Frostman exponents of the projections of μ and prove Theorem 1.9 below. This result allows us to apply a result of Shmerkin to deduce Corollary 1.5 from Theorem A. To motivate the theorem, we first give an equivalent definition of the quantities $\alpha_\ell(\mu)$. Let $\text{Gr}(d, \ell)$ be the Grassmanian of vector subspaces of \mathbb{R}^d of dimension ℓ . We identify subspaces $\pi \in \text{Gr}(d, \ell)$ with the associated canonical orthogonal projection from \mathbb{R}^d . For $\pi \in \text{Gr}(d, \ell)$, $x \in \pi$, and $\varepsilon > 0$, let $B(x, \varepsilon)$ be the ball of radius ε around x . If X is a measure space with measure μ and $f : X \rightarrow Y$ is a measurable map, we denote by $f_*\mu$ the push-forward measure. Given a compactly supported Borel measure μ on \mathbb{R}^d and $1 \leq \ell \leq d$, $\alpha_\ell(\mu)$ can be alternatively defined by

$$\alpha_\ell(\mu) = \liminf_{\varepsilon \rightarrow 0} \frac{\log \sup_{\pi \in \text{Gr}(d, \ell)} \sup_{x \in \pi} \pi_*\mu(B(x, \varepsilon))}{\log \varepsilon}. \quad (1.6)$$

Given a finite measure ν on \mathbb{R}^ℓ , the **Frostman exponent** of ν , denoted $\dim_\infty(\nu)$, is defined by

$$\dim_\infty(\nu) = \liminf_{\varepsilon \rightarrow 0} \frac{\log \sup_{x \in \mathbb{R}^\ell} \nu(B(x, \varepsilon))}{\log \varepsilon}. \quad (1.7)$$

By definition, we have $\alpha_\ell(\mu) \leq \inf_{\pi_* \in \text{Gr}(d, \ell)} \dim_\infty(\pi_*\mu)$. For self-similar measures, we show the following.

Theorem 1.9. *Suppose \mathcal{F} is an irrational homogeneous IFS on \mathbb{R}^2 satisfying the open set condition with limit set \mathcal{K} . Let μ be the restriction of the s -dimensional Hausdorff measure to \mathcal{K} , where $s = \dim_H(\mathcal{K})$. Then, for Lebesgue almost every $\pi_\theta \in \text{Gr}(2, 1)$,*

$$\alpha_1(\mu) = \inf_{\pi_\beta \in \text{Gr}(2, 1)} \dim_\infty((\pi_\beta)_*\mu) = \dim_\infty((\pi_\theta)_*\mu).$$

It is reasonable to expect the first equality in Theorem 1.9 to hold for self-similar measures in greater generality. We hope to address this question in future work.

The proof proceeds by realizing the quantities in question as limits of a certain sub-additive cocycle over an irrational rotation and utilizing an extension of the classical sub-additive ergodic theorem due to Furman [Fur97]. This technique has been used for similar problems in [PS09, NPS12, GSSY16].

Remark 1.10. (1) Under the hypotheses of Theorem 1.9, it is shown in [Shm16, Theorem 8.2] that $\dim_\infty((\pi_\theta)_*\mu) = \min\{\dim_H(\mathcal{K}), 1\}$ for every $\theta \in [0, 2\pi)$. The point of Theorem 1.9 is establishing equality between $\alpha_1(\mu)$ and these Frostman exponents. This statement is perhaps not surprising to experts, though we could not locate a reference in the literature.

(2) Suppose $\mathcal{K} = \mathcal{C} \times \mathcal{C}$ where \mathcal{C} is Cantor's middle thirds set realized as the attractor of the natural IFS, denoted by \mathcal{F} . A remarkable result of Shmerkin in [Shm16, Theorem 6.2, Corollary 6.4] shows that $\dim_\infty((\pi_\theta)_*\mu) = 1$ whenever $\theta \notin \mathbb{Q}\pi$. On the other hand, the projections of μ on the coordinate axes have Frostman exponent equal to $\dim_H(\mathcal{C}) < 1$. This shows that the minimality assumption (irrationality of R_α) cannot be dropped for the second equality in the conclusion of Theorem 1.9.

Theorem 1.9 suggests an affirmative answer to the following question.

Question 1.11. *Suppose $\mathcal{F}, \mathcal{K}, s$, and μ are as in Theorem A. Assume that the group generated by the rotation parts of the maps in the IFS is dense in $\mathrm{SO}(d, \mathbb{R})$. Is it true that $\alpha_\ell(\mu) = \dim_\infty(\pi_*\mu) = \min(s, \ell)$ for every $\pi \in \mathrm{Gr}(d, \ell)$? If we further assume that the maps in \mathcal{F} are all defined by rational parameters, is it true that $\dim_H(\mathrm{Sing}(d) \cap \mathcal{K}) = sd/(d+1)$?*

1.4. Prior work. Kleinbock and Weiss showed in [KW05b] that irreducible self-similar measures satisfying the open set condition (OSC) give 0 mass to $\mathrm{Sing}(d)$. Indeed, they establish this result for the much wider class of *friendly measures* introduced in [KLW04]. When $0 < \varepsilon < 1$ is fixed, the set of vectors \mathbf{x} which admit non-trivial solutions to the inequalities in (1.1) for all large N are referred to as *Dirichlet ε -improvable* and are denoted DI_ε . In [SW16], generalizing Benoist and Quint's fundamental measure rigidity results for random walks on homogeneous spaces, Simmons and Weiss showed that irreducible self-similar measures with the OSC give 0 mass to DI_ε for every $0 < \varepsilon < 1$. Special cases of this result were obtained previously by Einsiedler, Fishman, and Shapira for measures admitting invariance by expanding maps [EFS11]. This latter result relied on entropy methods and measure rigidity results for higher rank diagonalizable actions.

The above results indicate scarcity of singular vectors (and indeed of DI_ε vectors) in the support of self-similar measures. On the other hand, Kleinbock and Weiss showed that badly approximable vectors have full dimension in the support of absolutely friendly measures (which include the measures in Theorem A) [KW05a], cf. [Fis09]. An observation of Davenport and Schmidt [DS70b, Theorem 2] shows that badly approximable vectors belong to $\cup_{0 < \varepsilon < 1} \mathrm{DI}_\varepsilon$. It follows that the set $\cup_{0 < \varepsilon < 1} \mathrm{DI}_\varepsilon$ has full dimension in the support of these measures. Finally, the reader may wish to consult [DFSU18] for more recent developments in the study of *extremality* of fractal measures.

On the fractal geometric side, dimensions of projections of self-similar sets and measures have been extensively studied, see [Shm15] for a survey. It is shown in [Shm16, Section 6] that the Frostman exponent of the projection in an irrational direction of the restriction of the Hausdorff measure to a Sierpiński carpet of dimension > 1 or to the 1-dimensional Sierpiński gasket in the plane is equal to 1. These results build on prior work of Hochman in [Hoc14] and the observation that the projections of such sets are themselves self-similar. Less is known about projections of these sets in rational directions. However, several regularity results on the dimension of slices of the aforementioned sets with lines of rational slopes have been established in [BFS12, BR14] and references therein. When \mathcal{K} is the limit set of an IFS \mathcal{F} satisfying a stronger separation condition than OSC and such that the rotation parts of the maps in \mathcal{F} generate a dense subgroup of $\mathrm{SO}(d, \mathbb{R})$, it is shown in [HS12] that the image of \mathcal{K} under a projection onto an ℓ -dimensional subspace has Hausdorff dimension $\min(\ell, \dim_H(\mathcal{K}))$. Moreover, under the same hypotheses, it is shown in *loc.cit.* that projections of the associated self-similar measures onto subspaces of dimension ℓ are exact-dimensional, with dimension equal to $\min(\ell, \dim_H(\mathcal{K}))$. The reader is referred to [PS09, NPS12, HS12] and references therein for more results in that direction.

1.5. Overview of the proof and reduction to dynamics. We will deduce Theorem A from a stronger dynamical statement, Theorem 1.12 below. Let $G = \mathrm{SL}(d+1, \mathbb{R})$, $\Gamma = \mathrm{SL}(d+1, \mathbb{Z})$, and $X = G/\Gamma$. For $t > 0$ and $\mathbf{x} \in \mathbb{R}^d$, define the following elements of G .

$$a_t = \begin{pmatrix} e^{dt} & \mathbf{0} \\ \mathbf{0} & e^{-t}\mathbf{I}_d \end{pmatrix}, \quad u(\mathbf{x}) = \begin{pmatrix} 1 & \mathbf{x} \\ \mathbf{0} & \mathbf{I}_d \end{pmatrix} \quad (1.8)$$

where \mathbf{I}_d denotes the $d \times d$ identity matrix. It was shown by Dani in [Dan85] that \mathbf{x} is singular if and only if the orbit $(a_t u(\mathbf{x})\Gamma)_{t \geq 0}$ diverges in X . Theorem A follows from the following result.

Theorem 1.12. *Suppose \mathcal{F} is an irreducible IFS on \mathbb{R}^d satisfying the open set condition and let \mathcal{K} be its limit set. Let $s = \dim_H(\mathcal{K})$ and μ be the restriction of the s -dimensional Hausdorff measure*

to \mathcal{K} . Then, for every $x_0 \in G/\Gamma$, the Hausdorff dimension of the set of vectors $\mathbf{x} \in \mathcal{K}$ such that the forward orbit $(a_t u(\mathbf{x})x_0)_{t \geq 0}$ diverges on average in G/Γ is at most

$$s - \min_{1 \leq \ell \leq d} \frac{(d - \ell + 1)\alpha_\ell(\mu)}{d + 1}.$$

For convenience of the reader, we outline the proof of Theorem 1.12. The proof has two main steps: a linear argument and a probabilistic scheme. The linear argument is concerned with estimating the average rate of expansion of vectors in the exterior powers of the standard representation of $\mathrm{SL}(d+1, \mathbb{R})$ with respect to any measure ν satisfying $\alpha_\ell(\nu) > 0$ for every ℓ . Roughly, we show that $\|a_t u(\mathbf{x})v\|^{-1} \in L^{\alpha_\ell(\nu) - \varepsilon}(\nu)$, for every $\varepsilon > 0$, where v is any non-zero vector in $\bigwedge^\ell \mathbb{R}^{d+1}$, equipped with the standard representation of G . This is the content of Section 5.

The proof of this step is based on a simple but crucial observation regarding *transversality* of the expanding coordinates of a_t . This is carried out in Propositions 4.1 and 4.2. Roughly speaking, we show that if the projections of $u(\mathbf{x})v$ onto each expanding coordinate of a_t are simultaneously small, this implies that \mathbf{x} belongs to a neighborhood of an affine subspace of low dimension.

The other ingredient is to use the method of integral inequalities, first introduced in [EMM98], to translate the results on linear expansion to recurrence results on the space G/Γ . This is carried out in Section 6. We construct a Margulis function (see Def. 3.1) $f : G/\Gamma \rightarrow \mathbb{R}_+$ which measures the depth of orbits into the cusp. We show that the average value of $f(a_t u(\mathbf{x})x_0)$, for a fixed $t > 0$, and with respect to any non-planar measure $d\nu(\mathbf{x})$, is a fraction of $f(x_0)$, whenever $x_0 \in G/\Gamma$ is sufficiently deep in the cusp. This concludes the linear argument.

The probabilistic scheme takes as input the average height contraction established in the previous step and converts it into an upper dimension estimate on singular vectors in the support of the measure, Theorem 3.4. This step is the technical heart of this article and is carried out in Section 3. Before explaining the strategy, it is worth noting at this stage that we have not yet used the assumption that the measures under consideration are self-similar. However, this assumption is indispensable for the probabilistic scheme. For example, using Frostman's Lemma, and the fact that $\dim_H(\mathrm{Sing}(d)) > d - 1$, one can find a measure ν , whose support is contained $\mathrm{Sing}(d)$, and satisfying $\alpha_d(\nu) > d - 1$. This in particular implies that $\alpha_\ell(\nu) > 0$ for all ℓ (cf. proof of Lemma 2.6). The results of the previous step (the linear argument) apply to such a measure, however clearly singular vectors have full dimension in the support of ν .

The main step in proving Theorem 3.4 is Proposition 3.5. The key property of self-similar measures we use is a *renormalization* mechanism to convert global information (average contraction over the entire support) obtained via the linear argument at a small fixed time scale t to local information (average contraction over small pieces of the support) at large time scales. We make use of the existence of a faithful representation $\mathrm{Sim}(\mathbb{R}^d) \hookrightarrow N_G(U)$, where $\mathrm{Sim}(\mathbb{R}^d)$ is the group of similarities of \mathbb{R}^d , $U = \{u(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^d\}$, and $N_G(U)$ is the normalizer of U in G . Under this inclusion, the scaling subgroup of $\mathrm{Sim}(\mathbb{R}^d)$ corresponds to $\{a_t : t \in \mathbb{R}\}$.

Analogous steps have been studied before in [KKLM17] in the case of Lebesgue measure on \mathbb{R}^d and in [Kha18] for measures supported on curves. The strategy in those cases is to show that the probability that an orbit segment $(a_t u(\mathbf{x})x_0)_{0 \leq t \leq N}$ spends a large proportion of its time in the cusp decays exponentially with a precise rate. One then uses continuity of the flow to show that if \mathbf{x} is one such point, then a whole neighborhood of \mathbf{x} of radius $e^{-(d+1)N}$ has roughly the same behavior. This converts measure estimates into a count on covers and an estimate on the box dimension.

Unfortunately, this strategy seems to fail when the contraction ratios of the maps in the IFS are not all the same and this introduces considerable difficulties in our case. This is due to the fact that the natural neighborhoods used in constructing the covers are pieces of the fractal whose diameters have distinct exponential decay rates to 0. We introduce a method which does not rely on counting covers and rather works by estimating the Hausdorff dimension directly. The key starting inequality

of our estimates roughly takes the form:

$$\sum_{\mathcal{K}_\omega \cap B(N) \neq \emptyset} \text{diam}(\mathcal{K}_\omega)^{s-\gamma} \ll \int_{B(N-1)} \tau(\mathbf{x})^{-\gamma} f(a_{\tau(\mathbf{x})} u(\mathbf{x}) x_0) d\mu(\mathbf{x}), \quad (1.9)$$

where the sum is over self-similar pieces of the limit set \mathcal{K} which meet the set $B(N)$ of vectors \mathbf{x} whose orbit segment of length N spends a large proportion of its time in the cusp. Here, we use the fact that μ is a Hausdorff measure and, in particular, that $\mu(\mathcal{K}_\omega) \asymp \text{diam}(\mathcal{K}_\omega)^s$. This allows us to interpret this sum as an integral of a cocycle $\tau(\mathbf{x})^{-\gamma}$ measuring the diameter of the self-similar pieces. Moreover, and crucially to our method, we interpret the factor $\tau(\mathbf{x})^\gamma$ as being a non-constant contraction ratio of the averaging operator given by integrating against μ as in the right side of (1.9). This is to be contrasted with the more standard integral inequalities of Margulis functions of the form $\int f d\nu \leq af + b$, with a uniform contraction factor $0 < a < 1$. We also estimate the indicator of $B(N)$ by a product of the height function f and the indicator of $B(N-1)$. Accordingly, our linear argument is modified to take into account the additional cocycle factor inside the integral. An inductive procedure is then carried out to bound the $(s-\gamma)$ -Hausdorff measure of the set $\liminf_N B(N)$.

We conclude this introduction with a natural problem arising from our investigations.

Question 1.13. *What is the Hausdorff dimension of singular vectors belonging to the limit set of a Zariski-dense, convex cocompact, discrete group of Möbius transformations of \mathbb{R}^d ?*

We refer the reader to [BGSV18] where a related notion of *intrinsic* singular vectors on the limit set of geometrically finite manifolds is introduced and studied. We remark that the Diophantine problems studied in [BGSV18] correspond with the recurrence behavior of the geodesic flow on hyperbolic manifolds, while Question 1.13 is related to the recurrence of diagonal flows on $\text{SL}(d+1, \mathbb{R})/\text{SL}(d+1, \mathbb{Z})$ via Dani's correspondence in the spirit of Theorem 1.12.

Acknowledgements. I would like to thank the anonymous referees for a careful reading and for numerous valuable comments that improved the presentation and corrected several inaccuracies in the initial version of the article. I would like to thank Jon Chaika, Yitwah Cheung and Barak Weiss for their interest in the project and for comments on an earlier version of the manuscript. I would also like to thank Pablo Shmerkin for several valuable suggestions and for providing the proof of Lemma A.1.

2. PRELIMINARIES

We recall some properties of self-similar measures to be used in later sections.

2.1. Hausdorff dimension. We recall the definition of the Hausdorff dimension. Let A be a subset of a metric space X . For any $\kappa, s > 0$, we define

$$H_\kappa^s(A) = \inf \left\{ \sum_{I \in \mathcal{U}} \text{diam}(I)^s : \mathcal{U} \text{ is a cover of } A \text{ by balls of diameter } < \kappa \right\}. \quad (2.1)$$

The s -dimensional Hausdorff measure of A is defined to be

$$H^s(A) = \lim_{\kappa \rightarrow 0^+} H_\kappa^s(A) = \sup_{\kappa > 0} H_\kappa^s(A).$$

The Hausdorff dimension of A is defined by

$$\dim_H(A) = \inf \{ \gamma \geq 0 : H^\gamma(A) = 0 \} = \sup \{ \gamma \geq 0 : H^\gamma(A) = \infty \}.$$

2.2. IFS Notation. Fix a finite set Λ . An *iterated function system* (IFS for short) is a finite collection $\mathcal{F} = \{h_i : i \in \Lambda\}$ of contractive similarities of \mathbb{R}^d , i.e., for each $i \in \Lambda$, h_i has the form

$$h_i = \rho_i O_i + b_i,$$

where $0 < \rho_i < 1$, $O_i \in \text{SO}(d, \mathbb{R})$, and $b_i \in \mathbb{R}^d$. The **similarity dimension** of \mathcal{F} is defined to be the unique solution of the equation $\sum_{i \in \Lambda} \rho_i^s = 1$. It is shown in [Hut81] that there exists a unique compact set $\mathcal{K} \subset \mathbb{R}^d$ which is invariant by \mathcal{F} in the following sense.

$$\mathcal{K} = \bigcup_{i \in \Lambda} h_i(\mathcal{K}) \quad (2.2)$$

We refer to the set \mathcal{K} as the **limit set** of \mathcal{F} . Hutchinson introduced a notion of separation, the open set condition, that will allow us to treat the union in (2.2) as if it were disjoint, cf. Proposition 2.3 below. Following [Hut81], we say \mathcal{F} satisfies the **open set condition** (OSC for short) if there exists a non-empty open set $U \subset \mathbb{R}^d$ such that the following holds:

$$\begin{cases} h_i(U) \subseteq U, & \text{for every } i \in \Lambda, \\ h_i(U) \cap h_j(U) = \emptyset, & \text{for every } i \neq j \in \Lambda. \end{cases} \quad (2.3)$$

Given $\omega = (\omega_i) \in \Lambda^k$, we let

$$h_\omega = h_{\omega_1} \circ \cdots \circ h_{\omega_k}, \quad \mathcal{K}_\omega = h_\omega(\mathcal{K}), \quad \mathcal{P}_n = \{\mathcal{K}_\omega : \omega \in \Lambda^n\}. \quad (2.4)$$

The maps h_ω take the form $\rho_\omega O_\omega + b_\omega$, where

$$\rho_\omega = \prod_{i=0}^k \rho_{\omega_i}, \quad O_\omega = O_{\omega_1} \cdots O_{\omega_k}, \quad b_\omega = h_\omega(0) \in \mathbb{R}^d. \quad (2.5)$$

It will be convenient for us to consider the numbers ρ_ω for finite prefixes of infinite words ω . We do this by means of a multiplicative cocycle. To this end, let $\sigma : \Lambda^\mathbb{N} \rightarrow \Lambda^\mathbb{N}$ be the shift map, i.e., $\sigma((\omega_k)_{k \in \mathbb{N}}) = (\omega_{k+1})_{k \in \mathbb{N}}$. Consider the function $\rho : \Lambda^\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}_+$ defined as follows for $\omega = (\omega_k)_{k \in \mathbb{N}}$

$$\rho(\omega, n) = \prod_{k=1}^n \rho_{\omega_k}. \quad (2.6)$$

Then, ρ satisfies the cocycle relation: $\rho(\omega, m+n) = \rho(\omega, m)\rho(\sigma^m(\omega), n)$.

We wish to regard $\rho(\cdot, n)$ as a function on \mathcal{K} . Since members of \mathcal{P}_n may not be disjoint, there is ambiguity on the overlaps. For this purpose, we introduce a modified partition $\hat{\mathcal{P}}_n$ consisting of disjoint sets as follows. Using any fixed order on the elements of Λ , we can endow Λ^n with a lexicographic order and define for each $\omega \in \Lambda^n$:

$$\hat{\mathcal{K}}_\omega = \mathcal{K}_\omega \setminus \left(\bigcup_{\alpha > \omega} \mathcal{K}_\alpha \cap \mathcal{K}_\omega \right), \quad \hat{\mathcal{P}}_n = \{\hat{\mathcal{K}}_\omega : \omega \in \Lambda^n\}. \quad (2.7)$$

Let $\mathbf{x} \in \mathcal{K}$ and let $\omega \in \Lambda^n$ be the unique word such that $\mathbf{x} \in \hat{\mathcal{K}}_\omega$. Then, we define

$$\rho(\mathbf{x}, k) := \rho(\omega, k). \quad (2.8)$$

The following lemma shows that $\hat{\mathcal{P}}_n$ form a refining sequence of partitions. This fact is used in the inductive procedure of constructing covers for singular vectors by elements of $\hat{\mathcal{P}}_n$.

Lemma 2.1. *For every $m < n \in \mathbb{N}$ and every $\beta \in \Lambda^m$, we have $\hat{\mathcal{K}}_\beta = \bigcup \hat{\mathcal{K}}_\alpha$, where the union is taken over words $\alpha \in \Lambda^n$ such that $\beta \in \Lambda^m$ is the prefix of α of length m .*

Proof. Suppose $\alpha \in \Lambda^n$ is such that β is a prefix of α . Suppose $\omega \in \Lambda^m$ is such that $\omega > \beta$ and suppose $\mathbf{x} \in \mathcal{K}_\beta \cap \mathcal{K}_\omega$. By (2.2), we have $\mathcal{K}_\omega = \bigcup_{\tau \in \Lambda^{n-m}} \mathcal{K}_{\omega\tau}$, where $\mathcal{K}_{\omega\tau}$ denotes $h_\omega(\mathcal{K}_\tau)$. Let $\tau \in \Lambda^{n-m}$ be such that $\mathbf{x} \in \mathcal{K}_{\omega\tau}$. Then, since $\omega > \beta$ and β is a prefix of α , it follows that $\omega\tau > \alpha$ by definition of the lexicographic order. This implies that $\mathbf{x} \notin \hat{\mathcal{K}}_\alpha$. This shows that $\hat{\mathcal{K}}_\alpha \subseteq \hat{\mathcal{K}}_\beta$.

To show the reverse containment, fix $\mathbf{x} \in \hat{\mathcal{K}}_\beta$. Let $\alpha \in \Lambda^n$ be the maximal word in Λ^n in the lexicographic order having β as a prefix and satisfying 1) $\mathbf{x} \in \mathcal{K}_\alpha$, and 2) $\alpha \geq \alpha'$ for every $\alpha' \in \Lambda^n$ such that $\mathbf{x} \in \mathcal{K}_{\alpha'}$. Suppose for contradiction that $\mathbf{x} \notin \hat{\mathcal{K}}_\alpha$. Then, there exists some $\omega \in \Lambda^n$ such that $\omega > \alpha$ and $\mathbf{x} \in \mathcal{K}_\alpha \cap \mathcal{K}_\omega$. Let $\omega_0 \in \Lambda^m$ be the prefix of ω . By maximality of α , $\omega_0 \neq \beta$. However, by definition of the lexicographic order, since $\omega > \alpha$, it must be that $\omega_0 > \beta$. In particular, since $\mathbf{x} \in \mathcal{K}_\omega \subset \mathcal{K}_{\omega_0}$, it follows that $\mathbf{x} \notin \hat{\mathcal{K}}_\beta$, contrary to our assumption and concluding the proof. \square

2.3. Self-similar measures. Fix a probability vector λ with full support on Λ . That is $\lambda = (\lambda_i)_{i \in \Lambda}$, $\lambda_i > 0$ for every i , and $\sum_i \lambda_i = 1$. The *Markov-Feller Operator* P_λ is defined by

$$P_\lambda(\nu) = \sum_{i \in \Lambda} \lambda_i (h_i)_* \nu,$$

for all Borel measures ν on \mathbb{R}^d . It is shown in [Hut81] that there exists a unique probability measure μ supported on \mathcal{K} and satisfying

$$\mu = P_\lambda(\mu). \quad (2.9)$$

We refer to measures satisfying (2.9) as **self-similar measures** for (\mathcal{F}, λ) . Simple induction applied to (2.9) shows that

$$\mu = P_\lambda^k(\mu) = \sum_{\omega \in \Lambda^k} \lambda_\omega (h_\omega)_* \mu, \quad (2.10)$$

where $\lambda_\omega = \prod_{i=1}^k \lambda_i$. Under the open set condition, Hutchinson showed that the self-similar measure for the natural probability vector $(\rho_i^s)_{i \in \Lambda}$ is in fact the s -dimensional Hausdorff measure H^s .

Proposition 2.2 (Theorem 5.3(1), [Hut81]). *Suppose \mathcal{F} is an IFS satisfying the open set condition with similarity dimension s and let \mathcal{K} denote its limit set. Then, $0 < H^s(\mathcal{K}) < \infty$. In particular, $\dim_H(\mathcal{K}) = s$. Let μ denote the normalized restriction of H^s to \mathcal{K} . Then, μ is the self-similar measure for the probability vector $(\rho_i^s)_{i \in \Lambda}$. Moreover, there exist constants $a, b > 0$, such that for every $x \in \mathcal{K}$ and every $r \in (0, 1)$,*

$$ar^s \leq \mu(B(x, r)) \leq br^s.$$

2.4. Consequences of null overlaps. In general, the overlap between members of \mathcal{P}_n causes serious problems in the analysis. However, the following result, obtained in [Hut81], shows that the OSC insures that these overlaps are negligible from the point of view of self-similar measures.

Proposition 2.3 (Proposition 5.1(4), Theorem 5.3(1) [Hut81]). *Suppose \mathcal{F} satisfies the open set condition and let s be its similarity dimension. Let $\lambda = (\rho_i^s)_{i \in \Lambda}$ and μ be the self-similar probability measure for (\mathcal{F}, λ) . Then, for every $k \in \mathbb{N}$ and all $\alpha \neq \omega \in \Lambda^k$, $\mu(\mathcal{K}_\omega \cap \mathcal{K}_\alpha) = 0$.*

We now state two consequences of Proposition 2.3 which we use in our proof. For a Borel set A and a Borel measure μ , we denote by $\mu|_A$ the restriction of μ to A . That is for every Borel set B , $\mu|_A(B) = \mu(B \cap A)$. The following lemma will be useful in estimating Hausdorff dimension.

Lemma 2.4. *Suppose \mathcal{F} satisfies the open set condition. Then, for every $n \in \mathbb{N}$ and $\omega \in \Lambda^n$, \mathcal{K}_ω and $\hat{\mathcal{K}}_\omega$ have the same diameter.*

Proof. It suffices to show that $\hat{\mathcal{K}}_\omega$ is dense in \mathcal{K}_ω . For $\alpha \in \Lambda^\mathbb{N}$ and $k \in \mathbb{N}$, denote by $\alpha|_k$ its length k prefix. Let $\pi : \Lambda^\mathbb{N} \rightarrow \mathcal{K}$ denote the coding map defined by

$$\pi(\alpha) = \lim_{k \rightarrow \infty} h_{\alpha|_k}(0).$$

The space $\Lambda^\mathbb{N}$ is endowed with the product topology induced from the discrete topology on Λ . In this topology, the map π is onto and continuous [Hut81, Theorem 3.1.3(vii)]. Moreover, if $\lambda = (\rho_i^s)$ and μ is the unique self-similar measure satisfying (2.9), then

$$\pi_*(\lambda^\mathbb{N}) = \mu.$$

Let $A = \pi^{-1}(\mathcal{K}_\omega)$ and $\hat{A} = \pi^{-1}(\hat{\mathcal{K}}_\omega)$. It suffices to show that \hat{A} is dense in A by continuity of π . Note that A is the cylinder set consisting of all sequences whose prefix of length n is ω . Moreover, by Proposition 2.3, $\lambda^\mathbb{N}(\hat{A}) = \lambda^\mathbb{N}(A) > 0$. Since $\lambda^\mathbb{N}|_A$ has full support in A , \hat{A} is dense in A . \square

We record another useful consequence of the null overlaps.

Lemma 2.5. *Suppose \mathcal{F} satisfies the open set condition, $\lambda = (\rho_i^s)$ and $\mu = P_\lambda(\mu)$. Then, for every $k \in \mathbb{N}$ and all $\omega \in \Lambda^k$,*

$$\mu|_{\mathcal{K}_\omega} = \lambda_\omega(h_\omega)_*\mu.$$

Proof. This follows from Proposition 2.3 and equation (2.10). \square

2.5. Elementary facts on projections. We say \mathcal{F} is **irreducible** if no finite collection of proper affine subspaces of \mathbb{R}^d is invariant by each $h_i \in \mathcal{F}$. When s is the similarity dimension of \mathcal{F} , Proposition 2.2 implies that

$$\alpha_d(\mu) = s, \tag{2.11}$$

where μ is the restriction of the s -Hausdorff measure to \mathcal{K} . Moreover, it is easy to see that

$$i \leq j \implies \alpha_i(\mu) \leq \alpha_j(\mu). \tag{2.12}$$

The following lemma provides a simple lower estimate for $\alpha_\ell(\mu)$.

Lemma 2.6. *Suppose \mathcal{F} is an irreducible IFS on \mathbb{R}^d satisfying the open set condition and let \mathcal{K} be its limit set. Suppose $s = \dim_H(\mathcal{K}) > d - 1$. Then, for each $1 \leq \ell \leq d$, $\alpha_\ell(\mu) \geq s - d + \ell$, where μ is the restriction of H^s to \mathcal{K} .*

Proof. This result is well-known, we provide a proof for completeness. Given an affine subspace \mathcal{L} of dimension $d - \ell$ and $\varepsilon > 0$, the boundedness of \mathcal{K} implies that we can cover the set $\mathcal{K} \cap \mathcal{L}^{(\varepsilon)}$ with $O(\varepsilon^{-(d-\ell)})$ balls of radius ε . By Proposition 2.2, each such ball has measure at most $O(\varepsilon^s)$. In particular, $\mu(\mathcal{L}^{(\varepsilon)}) \ll \varepsilon^{s-d+\ell}$. As \mathcal{L} and ε were arbitrary, this completes the proof. \square

3. THE CONTRACTION HYPOTHESIS AND DIVERGENT TRAJECTORIES

In this section, we prove an abstract recurrence result for orbits of the diagonal flow in (1.8) starting from fractals in actions of $\mathrm{SL}(n, \mathbb{R})$ on metric spaces. Theorem 3.4 is the main result of this section establishing a bound on the dimension of divergent orbits. In later sections, we verify the hypotheses of this theorem in the settings of the results stated in the introduction.

3.1. The Contraction Hypothesis for Actions of $\mathrm{SL}(n, \mathbb{R})$. We fix the following more convenient parametrization of the diagonal subgroup a_t of $\mathrm{SL}(d+1, \mathbb{R})$, defined in (1.8), which we denote by g_t for $t > 0$,

$$g_t = \begin{pmatrix} t^{-d/(d+1)} & \mathbf{0} \\ \mathbf{0} & t^{1/(d+1)} \mathbf{I}_d \end{pmatrix}, \tag{3.1}$$

where I_d denotes the $d \times d$ identity matrix. Note that in this parametrisation, one has for every $t, s > 0$,

$$g_t \circ g_s = g_{ts}.$$

Recall the definition of self-similar measures in (2.9). The following is the key recurrence property for the action and the measure which underlies the results stated in the introduction.

Definition 3.1 (The Contraction Hypothesis). Suppose X is a metric space equipped with an action of $G = \mathrm{SL}(d+1, \mathbb{R})$ and let μ be a self-similar probability measure for an IFS $\mathcal{F} = \{h_i = \rho_i O_i + b_i : i \in \Lambda\}$ on \mathbb{R}^d . Given a collection of functions $\mathcal{F} = \{f_k : X \rightarrow (0, \infty] : k \in \mathbb{N}\}$ and real numbers $0 \leq \gamma_0 < \beta$, we say that μ satisfies the $(\mathcal{F}, \mathcal{F}, \beta, \gamma_0)$ -**contraction hypothesis** on X if the following properties hold:

- (1) The set $Z = \{f_k = \infty\}$ is independent of k and is G -invariant.
- (2) For every $k \in \mathbb{N}$, f_k is $\mathrm{SO}(d+1, \mathbb{R})$ -invariant and uniformly log Lipschitz with respect to the G action. That is for every bounded neighborhood \mathcal{O} of identity in G , there exists a constant $C_{\mathcal{O}} \geq 1$ such that for every $g \in \mathcal{O}$, $x \in X$ and $k \in \mathbb{N}$,

$$C_{\mathcal{O}}^{-1} f_k(x) \leq f_k(gx) \leq C_{\mathcal{O}} f_k(x). \quad (3.2)$$

- (3) There exists a constant $c \geq 1$ such that the following holds: for every $k \in \mathbb{N}$ and $\gamma_0 \leq \gamma \leq \beta$, there exists $T > 0$ such that for all $y \in X$ with $f_k(y) > T$,

$$\int_{\mathcal{K}} \rho(\mathbf{x}, k)^{-\gamma} f_k(g_{\rho(\mathbf{x}, k)} u(\mathbf{x})y) d\mu(\mathbf{x}) \leq c f_k(y) \left(\int_{\mathcal{K}} \rho(\mathbf{x}, k) d\mu(\mathbf{x}) \right)^{\beta-\gamma}, \quad (3.3)$$

where $\rho(\mathbf{x}, k)$ is the cocycle defined in (2.8) and \mathcal{K} is the limit set of the IFS \mathcal{F} .

The functions f_k will be referred to as **height functions**.

The notion of height functions was introduced in homogeneous dynamics in [EMM98] and was used in [KKLM17] to find a sharp upper bound on the dimension of singular systems of linear forms. The ‘‘Contraction Hypothesis’’ terminology is due to [BQ11].

Remark 3.2. (1) Inequality (3.3) should be thought of as a Margulis inequality with a *non-uniform* contraction ratio. To best illustrate this analogy, consider the case where $\rho(\cdot, 1) \equiv \rho$ for some fixed constant $\rho \in (0, 1)$. Setting $\gamma = \beta$ in (3.3) yields

$$\int_{\mathcal{K}} f_k(g_{\rho^k} u(\mathbf{x})y) d\mu(\mathbf{x}) \leq c \rho^{k\beta} f_k(y) + \rho^{-k\beta} T, \quad (3.4)$$

for all $k \in \mathbb{N}$ and all $y \in X$, where T is the constant in Def. 3.1(3). In particular, if k is large enough so that $\rho^{k\beta} c < 1$, the inequality in (3.4) recovers the classical form of Margulis inequalities of the form (cf. [EMM98, EM04, BQ11, EM01, EMM15, KKLM17, Kha18])

$$\int_{\mathcal{K}} f(g_t u(\mathbf{x})y) d\mu \leq a f(y) + b,$$

where $f = f_k$, $t = \rho^k$, $a = c \rho^{k\beta} < 1$, and $b = T \rho^{-k\beta}$.

- (2) The collection of height functions we use in our applications will be equivalent in the following sense: for all $k, \ell \in \mathbb{N}$, there exists a constant $C = C(k, \ell) \geq 1$, such that

$$C^{-1} f_k \leq f_\ell \leq C f_k.$$

Allowing this flexibility in Def. 3.1 is necessary however for verifying the contraction inequality (3.3) in the settings of flows on homogeneous spaces and strata of abelian differentials.

We refer the reader to the proofs of Theorem 6.3 and Corollary 7.3 where this dependence of the height functions on k is exploited. For the purposes of the discussion in this section, this flexibility does not play a role in the proofs and the reader may wish to regard the collection \mathcal{F} as consisting of a single height function.

(3) In our applications, the constant T , in Def. 3.1(3), can be chosen to be independent of γ .

We note that allowing height functions to assume the value ∞ has proven useful in several important applications [BQ11, EMM15].

Definition 3.3. In the presence of a collection $\mathcal{F} = \{f_k\}$ of height functions on a metric space X with a G -action, we say an orbit $(a_t x)_{t \geq 0}$ is **\mathcal{F} -divergent on average**, if for every $k \in \mathbb{N}$ and every $M > 0$,

$$\frac{1}{T} \int_0^T \chi_{M,k}(a_t x) dt \rightarrow 0,$$

where $\chi_{M,k}$ is the indicator function of $\{y \in X : f_k(y) \leq M\}$. We often drop \mathcal{F} from the notation: \mathcal{F} -divergent on average when \mathcal{F} is understood from context.

The following is the main result of this section.

Theorem 3.4. *Let X be a metric space equipped with an action by $G = \mathrm{SL}(d+1, \mathbb{R})$. Suppose \mathcal{F} is an IFS on \mathbb{R}^d satisfying the open set condition and denote by \mathcal{K} its limit set. Let $s = \dim_H(\mathcal{K})$ and let μ be the restriction of the s -dimensional Hausdorff measure to \mathcal{K} . Assume that μ satisfies the $(\mathcal{F}, \mathcal{F}, \beta, \gamma_0)$ -contraction hypothesis on X for a collection $\mathcal{F} = \{f_k : k \in \mathbb{N}\}$ of height functions and real numbers $0 \leq \gamma_0 < \beta \leq s$. Then, for all $x_0 \in X \setminus \{f_1 = \infty\}$,*

$$\dim_H(\mathbf{x} \in \mathcal{K} : (a_t u(\mathbf{x})x_0)_{t \geq 0} \text{ is } \mathcal{F}\text{-divergent on average}) \leq s - \beta.$$

The main applications of our results are to the G action on the space of unimodular lattices in \mathbb{R}^d and to $\mathrm{SL}(2, \mathbb{R})$ actions on moduli spaces of abelian differentials. The remainder of this section is dedicated to the proof of Theorem 3.4.

3.2. Notation. Throughout the remainder of this section, we let X, G, \mathcal{F} and μ be as in Theorem 3.4. We fix a finite set Λ so that $\mathcal{F} = \{h_i : i \in \Lambda\}$ and denote by \mathcal{K} the limit set of \mathcal{F} . We use the notation of Definition 3.1 pertaining to the height functions f_k . In particular, for $s = \dim \mathcal{K}$, we have

$$\gamma_0 < \beta \leq s.$$

Moreover, we retain the iterated function systems notation of Section 2. In particular, given a word $\omega \in \Lambda^n$, we use the notation:

$$\mathcal{K}_\omega := h_\omega(\mathcal{K}),$$

where h_ω is given by (2.4). For an integrable function φ on \mathbb{R}^d , we use $\int \varphi d\mu$ and $\int_{\mathcal{K}} \varphi d\mu$ interchangeably.

3.3. Integral inequalities and covering estimates. The goal of this section is to use the contraction hypothesis to control the Hausdorff dimension of divergent orbits.

Using the log Lipschitz property (2) of Definition 3.1, one can easily verify that the orbit $(a_t u(\mathbf{x})x_0)_{t \geq 0}$ diverges on average if and only if

$$\frac{1}{N} \sum_{l=1}^N \chi_{M,k}(g_{\rho(\mathbf{x}, lk)} u(\mathbf{x})x_0) \rightarrow 0, \quad \text{as } N \rightarrow \infty, \quad (3.5)$$

for all $k \in \mathbb{N}$ and $M > 0$, where $\chi_{M,k}$ is the indicator of $\{f_k \leq M\}$. This observation is very useful in handling the case where the contraction ratios of maps in \mathcal{F} are not all the same. This case poses

significant difficulties in the proof. We are thus naturally led to studying the following sets: for $0 < \delta < 1$, $M > 0$, $x \in X$, and $N, k \in \mathbb{N}$, define $Z_x(M, N, k, \delta)$ by

$$Z_x(M, N, k, \delta) := \{\mathbf{x} \in \mathcal{K} : \#\{1 \leq l \leq N : f_k(g_{\rho(\mathbf{x}, lk)}u(\mathbf{x})x) > M\} > \delta N\}. \quad (3.6)$$

The following proposition is one of the main technical results of this article.

Proposition 3.5. *There exists a constant $c_0 \geq 1$, depending only on the support of μ and on the constant c in Def. 3.1(3), such that the following holds. For every $k \in \mathbb{N}$, $\gamma \in (\gamma_0, \beta)$ and $x \in X \setminus \{f_1 = \infty\}$, there exists $M_0 = M_0(k, x, \gamma) > 0$, so that for all $M > M_0$ the following holds. For all $0 < \delta < 1$, $N \geq 1$, one has that*

$$\sum \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq c_0^N \left(\int_{\mathcal{K}} \rho(\mathbf{x}, k)^{-\gamma} d\mu \right)^{(1-\delta)N} \left(\int_{\mathcal{K}} \rho(\mathbf{x}, k) d\mu \right)^{\delta(\beta-\gamma)N},$$

where the sum is taken over words $\omega \in \Lambda^N$ satisfying $\hat{\mathcal{K}}_\omega \cap Z_x(M, N, k, \delta) \neq \emptyset$. Moreover, M_0 can be chosen uniformly as x varies in a fixed sub-level set $\{f_k \leq L\}$ for any $L > 0$.

Notational Convention. For the remainder of this section, we use f to denote f_1 to simplify notation.

We need technical preparation before the proof which occupies the next 3 subsections. Define the following constants:

$$K = \text{diam}(\mathcal{K}), \quad R = \sup_{\mathbf{x} \in \mathcal{K}} \|\mathbf{x}\|. \quad (3.7)$$

Using (2) of Definition 3.1, we can find $A \geq 1$ such that

$$A^{-1}f_k(y) \leq f_k(u(\mathbf{x})y) \leq Af_k(y), \quad (3.8)$$

for all $k \in \mathbb{N}$, $y \in X$ and all \mathbf{x} in a ball around $\mathbf{0} \in \mathbb{R}^d$ of radius $2R$. We also fix a constant $B \geq 1$ so that

$$B^{-1}f_k(y) \leq f_k(g_{\rho(\mathbf{x}, 1)}y) \leq Bf_k(y), \quad (3.9)$$

for all $k \in \mathbb{N}$, $\mathbf{x} \in \mathcal{K}$ and $y \in X$. For $x \in X$, $M > 0$ and natural numbers $m, n \in \mathbb{N}$, we define the following sets:

$$B_x(M, m; n) = \{\mathbf{x} \in \mathcal{K} : f_1(g_{\rho(\mathbf{x}, m+l)}u(\mathbf{x})x) \geq M, \text{ for } 1 \leq l \leq n\}.$$

Recall the definition of the sets $\hat{\mathcal{K}}_\omega$ and the partitions $\hat{\mathcal{P}}_n$ in (2.7). We frequently use the fact that $\text{diam}(\mathcal{K}_\omega) = \text{diam}(\hat{\mathcal{K}}_\omega) = K\rho_\omega$ for every $\omega \in \cup_k \Lambda^k$ and that

$$\mu(\mathcal{K}_\omega) = \mu(\hat{\mathcal{K}}_\omega) = \rho_\omega^s. \quad (3.10)$$

These facts follow from Lemmas 2.4 and 2.5.

3.4. Averages of multiplicative cocycles. We record the following cocycle relation.

Lemma 3.6. *Let $\omega \in \Lambda^n$. Then, for every $m \in \mathbb{N}$ and $\mathbf{x} \in \hat{\mathcal{K}}_\omega$,*

$$\rho(\mathbf{x}, n+m) = \rho(h_\omega^{-1}(\mathbf{x}), m)\rho(\mathbf{x}, n).$$

Proof. Let $u \in \Lambda^m$ be such that $\mathbf{x} \in \hat{\mathcal{K}}_{\omega u}$, where ωu is the concatenated word. In particular, by definition

$$\mathcal{K}_{\omega u} = h_\omega(h_u(\mathcal{K})).$$

Then, we have $\rho(\mathbf{x}, m+n) = \rho_\omega \rho_u$, where ρ_ω, ρ_u are the contraction ratios of h_ω and h_u respectively. Moreover, Lemma 2.1 implies that $h_\omega^{-1}(\mathbf{x}) \in \hat{\mathcal{K}}_u$. Hence, $\rho(h_\omega^{-1}(\mathbf{x}), m) = \rho_u$. Finally, since $\hat{\mathcal{K}}_u \subset \mathcal{K}$, we see that $\mathbf{x} \in \hat{\mathcal{K}}_\omega$ and, in particular, $\rho(\mathbf{x}, n) = \rho_\omega$. \square

The next lemma is a special case of a general principle: averages of "locally constant" submultiplicative cocycles form a submultiplicative sequence.

Lemma 3.7. *For all $\gamma \in \mathbb{R}$ and all $n \in \mathbb{N}$,*

$$\int \rho(\mathbf{x}, n)^\gamma d\mu(\mathbf{x}) = \left[\int \rho(\mathbf{x}, 1)^\gamma d\mu(\mathbf{x}) \right]^n$$

Proof. Let $a_n = \int \rho(\mathbf{x}, n)^\gamma d\mu$. Then, for all $m, n \in \mathbb{N}$, by Lemma 3.6, we obtain

$$\begin{aligned} a_{m+n} &= \int \rho(\mathbf{x}, m+n)^\gamma d\mu = \sum_{\omega \in \Lambda^m} \int_{\hat{\mathcal{K}}_\omega} \rho(\mathbf{x}, m)^\gamma \rho(h_\omega^{-1}(\mathbf{x}), n)^\gamma d\mu \\ &= \sum_{\omega \in \Lambda^m} \rho_\omega^\gamma \int_{\hat{\mathcal{K}}_\omega} \rho(h_\omega^{-1}(\mathbf{x}), n)^\gamma d\mu. \end{aligned}$$

By Lemma 2.5, applied with $\lambda_\omega = \rho(\omega, m)^s$, it follows that

$$(h_\omega^{-1})_* \left(\mu|_{\hat{\mathcal{K}}_\omega} \right) = \rho_\omega^s \mu = \mu(\hat{\mathcal{K}}_\omega).$$

Hence, we get that

$$a_{m+n} = \int \rho(\mathbf{x}, n)^\gamma d\mu \sum_{\omega \in \Lambda^m} \rho_\omega^\gamma \mu(\hat{\mathcal{K}}_\omega) = a_n a_m.$$

The lemma follows by induction. \square

The following lemma allows us to control complete sums over covers.

Lemma 3.8. *Let $k, m \in \mathbb{N}$. For every $\alpha \in \Lambda^k$ and $\gamma \in \mathbb{R}$,*

$$\sum_{\omega \in \Lambda^m} \text{diam} \left(\hat{\mathcal{K}}_{\alpha\omega} \right)^{s-\gamma} \leq \text{diam} \left(\hat{\mathcal{K}}_\alpha \right)^{s-\gamma} \int \rho(\mathbf{x}, m)^{-\gamma} d\mu(\mathbf{x}),$$

where, for $\omega \in \Lambda^m$, $\alpha\omega$ denotes the concatenation of α and ω . In particular, $\mathcal{K}_{\alpha\omega} = h_\alpha(\mathcal{K}_\omega)$.

Proof. By Lemma 2.4 and the cocycle property of ρ , we have

$$\text{diam} \left(\hat{\mathcal{K}}_{\alpha\omega} \right) = \text{diam} \left(\mathcal{K}_{\alpha\omega} \right) = K \rho(\alpha\omega, k+m) = K \rho(\alpha, k) \rho(\omega, m).$$

Moreover, we have that $\rho(\mathbf{x}, m+k)$ is constant almost everywhere on $\hat{\mathcal{K}}_{\alpha\omega}$ and equal to $\rho(\alpha\omega, m+k)$ for every $\omega \in \Lambda^m$. Thus, using the fact that $\mu(\mathcal{K}_{\alpha\omega}) = \rho(\alpha\omega, m+k)^s$, we obtain

$$\sum_{\omega \in \Lambda^m} \text{diam} \left(\hat{\mathcal{K}}_{\alpha\omega} \right)^{s-\gamma} = K^{s-\gamma} \sum_{\omega \in \Lambda^m} \rho_{\alpha\omega}^{s-\gamma} = K^{s-\gamma} \sum_{\omega \in \Lambda^m} \rho_{\alpha\omega}^{-\gamma} \mu(\mathcal{K}_{\alpha\omega}) = K^{s-\gamma} \int_{\hat{\mathcal{K}}_\alpha} \rho(\mathbf{x}, m+k)^{-\gamma} d\mu.$$

By Lemma 2.8, for almost every $\mathbf{x} \in \mathcal{K}_\alpha$

$$\rho(\mathbf{x}, m+k) = \rho(\mathbf{x}, k) \rho(h_\alpha^{-1}(\mathbf{x}), m) = \rho(\alpha, k) \rho(h_\alpha^{-1}(\mathbf{x}), m).$$

It follows that

$$\sum_{\omega \in \Lambda^m} \text{diam} \left(\hat{\mathcal{K}}_{\alpha\omega} \right)^{s-\gamma} = K^{s-\gamma} \rho(\alpha, k)^{-\gamma} \int_{\hat{\mathcal{K}}_\alpha} \rho(h_\alpha^{-1}(\mathbf{x}), m)^{-\gamma} d\mu$$

By Lemma 2.5, for every integrable function φ , we have $\int_{\mathcal{K}_\alpha} \varphi(h_\alpha^{-1}(\mathbf{x})) d\mu = \mu(\mathcal{K}_\alpha) \int \varphi(\mathbf{x}) d\mu$. This implies that

$$\begin{aligned} \sum_{\omega \in \Lambda^m} \text{diam} \left(\hat{\mathcal{K}}_{\alpha\omega} \right)^{s-\gamma} &= K^{s-\gamma} \rho(\alpha, k)^{-\gamma} \mu(\hat{\mathcal{K}}_\alpha) \int \rho(\mathbf{x}, m)^{-\gamma} d\mu \\ &= K^{s-\gamma} \rho(\alpha, k)^{s-\gamma} \int \rho(\mathbf{x}, m)^{-\gamma} d\mu = \text{diam} \left(\hat{\mathcal{K}}_\alpha \right)^{s-\gamma} \int \rho(\mathbf{x}, m)^{-\gamma} d\mu. \end{aligned}$$

□

3.5. Consequences of the log-Lipschitz property. The next 3 lemmas provide us with simple consequences of the log-Lipschitz property of the function f in Definition 3.1.

Lemma 3.9. *Suppose $\hat{\mathcal{K}}_\omega \in \hat{\mathcal{P}}_{m+n}$ is such that $\hat{\mathcal{K}}_\omega \cap B_x(M, m; n) \neq \emptyset$. Then,*

$$\hat{\mathcal{K}}_\omega \subseteq B_x(M/A, m; n).$$

Proof. Suppose $\mathbf{x}_0 \in \hat{\mathcal{K}}_\omega \cap B_x(M, m; n)$. Let $N = m + n$ and let $m < l \leq N$ be such that $f(g_{\rho(\mathbf{x}_0, l)} u(\mathbf{x}_0) x)$ is greater than M . Let $\mathbf{x} \in \hat{\mathcal{K}}_\omega$ be any other vector. Note that $\rho(\cdot, l)$ is constant on elements of $\hat{\mathcal{P}}_N$. This implies

$$g_{\rho(\mathbf{x}, l)} u(\mathbf{x}) = u \left(\rho(\mathbf{x}_0, l)^{-1}(\mathbf{x} - \mathbf{x}_0) \right) g_{\rho(\mathbf{x}_0, l)} u(\mathbf{x}_0).$$

Let $\mathbf{y}, \mathbf{y}_0 \in \hat{\mathcal{K}}$ be such that $\mathbf{x} = h_\omega(\mathbf{y})$ and $\mathbf{x}_0 = h_\omega(\mathbf{y}_0)$. The invariance of the Euclidean norm by $\text{SO}(d, \mathbb{R})$ implies

$$\|\mathbf{x} - \mathbf{x}_0\| = \|h_\omega(\mathbf{y}) - h_\omega(\mathbf{y}_0)\| = \rho(\omega, N) \|\mathbf{y} - \mathbf{y}_0\| \leq 2\rho(\omega, N)R.$$

Thus, since $\rho(\mathbf{x}_0, l) \geq \rho(\mathbf{x}_0, N) = \rho(\omega, N)$, the choice of the constant A in (3.8) implies

$$f(g_{\rho(\mathbf{x}, l)} u(\mathbf{x}) x) \geq M/A.$$

This being true for all $\mathbf{x} \in \hat{\mathcal{K}}_\omega$ concludes the proof. □

Lemma 3.10. *Suppose $f(g_{\rho(\mathbf{x}_0, \ell)} u(\mathbf{x}_0) y) > M$ for some $y \in X$, $\mathbf{x}_0 \in \hat{\mathcal{K}}_\omega$ and some $\omega \in \Lambda^\ell$. Then,*

$$f(g_{\rho(\omega, \ell)} u(b_\omega) y) > M/A,$$

where $b_\omega = h_\omega(0)$.

Proof. The proof is completely analogous to that of Lemma 3.9. □

Lemma 3.11. *Let $\omega \in \Lambda^\ell$. Then, for all $y \in X$,*

$$f(g_{\rho(\omega, \ell)} u(b_\omega) y) \leq \frac{A}{\mu(\hat{\mathcal{K}}_\omega)} \int_{\hat{\mathcal{K}}_\omega} f(g_{\rho(\mathbf{x}, \ell)} u(\mathbf{x}) y) d\mu,$$

where A is given by (3.8) and $b_\omega = h_\omega(0)$.

Proof. The proof follows from the fact that $\rho(\mathbf{x}, \ell) = \rho(\omega, \ell)$ everywhere on $\hat{\mathcal{K}}_\omega$ and the following estimate:

$$\rho(\omega, \ell)^{-1} \|b_\omega - \mathbf{x}\| \leq R,$$

for all $\mathbf{x} \in \hat{\mathcal{K}}_\omega$, where R is given by (3.7). \square

3.6. Consequences of the contraction property. For every $\ell \in \mathbb{N}$ and $\omega \in \Lambda^\ell$, we define the following elements of $\text{SO}(d+1, \mathbb{R})$:

$$k_\omega = \begin{pmatrix} 1 & \mathbf{0} \\ \mathbf{0} & O_\omega \end{pmatrix}, \quad (3.11)$$

where O_ω is the rotation part of the similarity h_ω and is given by (2.5). Note that $O_\omega \in \text{SO}(d, \mathbb{R})$ and that each k_ω commutes with g_t . The following lemma is the first main step in the proof of Proposition 3.5.

Lemma 3.12. *Let $A > 0$ be the constant in (3.8). Let $\gamma \in [\gamma_0, \beta]$ and let $T > 0$ be the constant provided by (3) of Definition 3.1 with $k = 1$. Suppose that $f(g_{\rho(\mathbf{x}_0, \ell)} u(\mathbf{x}_0) y) > AT$ for some $y \in X$, $\ell \in \mathbb{N}$, $\omega \in \Lambda^\ell$, and $\mathbf{x}_0 \in \hat{\mathcal{K}}_\omega$. Then,*

$$\begin{aligned} \int_{\hat{\mathcal{K}}_\omega} \rho(\mathbf{x}, \ell + 1)^{-\gamma} f(g_{\rho(\mathbf{x}, \ell + 1)} u(\mathbf{x}) y) d\mu(\mathbf{x}) \\ \leq cA \int_{\hat{\mathcal{K}}_\omega} \rho(\mathbf{x}, \ell)^{-\gamma} f(g_{\rho(\mathbf{x}, \ell)} u(\mathbf{x}) y) d\mu(\mathbf{x}) \left(\int \rho(\mathbf{x}, 1) d\mu(\mathbf{x}) \right)^{\beta - \gamma}, \end{aligned}$$

where c is as in (3) of Definition 3.1.

Proof. By Lemma 3.6, it follows that

$$\begin{aligned} \int_{\hat{\mathcal{K}}_\omega} \rho(\mathbf{x}, \ell + 1)^{-\gamma} f(g_{\rho(\mathbf{x}, \ell + 1)} u(\mathbf{x}) y) d\mu &= \int_{\hat{\mathcal{K}}_\omega} \rho(\mathbf{x}, \ell)^{-\gamma} \rho(h_\omega^{-1}(\mathbf{x}), 1)^{-\gamma} f(g_{\rho(\mathbf{x}, \ell + 1)} u(\mathbf{x}) y) d\mu \\ &= \rho(\omega, \ell)^{-\gamma} \int_{\hat{\mathcal{K}}_\omega} \rho(h_\omega^{-1}(\mathbf{x}), 1)^{-\gamma} f(g_{\rho(\mathbf{x}, \ell + 1)} u(\mathbf{x}) y) d\mu, \end{aligned} \quad (3.12)$$

where on the second line we used the fact that $\rho(\mathbf{x}, \ell) = \rho(\omega, \ell)$ everywhere on $\hat{\mathcal{K}}_\omega$.

Note that since $O_\omega \in \text{SO}(d, \mathbb{R})$, we have $O_\omega^{-1} = O_\omega^t$, where O_ω^t denotes the transpose of O_ω . Moreover, $h_\omega^{-1}(\mathbf{x}) = \rho(\omega, \ell)^{-1} O_\omega^{-1}(\mathbf{x} - b_\omega)$, where $b_\omega = h_\omega(0)$. Thus, the following identity holds.

$$g_{\rho(\mathbf{x}, \ell + 1)} k_\omega u(\mathbf{x}) = g_{\rho(h_\omega^{-1}(\mathbf{x}), 1)} u(h_\omega^{-1}(\mathbf{x})) g_{\rho(\omega, \ell)} k_\omega u(b_\omega).$$

Observe that k_ω commutes with g_t and recall that the function f is $\text{SO}(d+1, \mathbb{R})$ -invariant. This implies

$$f(g_{\rho(\mathbf{x}, \ell + 1)} u(\mathbf{x}) y) = f(g_{\rho(h_\omega^{-1}(\mathbf{x}), 1)} u(h_\omega^{-1}(\mathbf{x})) g_{\rho(\omega, \ell)} k_\omega u(b_\omega) y). \quad (3.13)$$

By Lemma 2.5, for every $\varphi \in L^1(\mu)$,

$$\frac{1}{\mu(\hat{\mathcal{K}}_\omega)} \int_{\hat{\mathcal{K}}_\omega} \varphi(h_\omega^{-1}(\mathbf{x})) d\mu(\mathbf{x}) = \int \varphi(\mathbf{x}) d\mu(\mathbf{x}).$$

Combining this fact with (3.13), we obtain, for $z = g_{\rho(\omega, \ell)} k_\omega u(b_\omega) y$,

$$\int_{\hat{\mathcal{K}}_\omega} \rho(h_\omega^{-1}(\mathbf{x}), 1)^{-\gamma} f(g_{\rho(\mathbf{x}, \ell + 1)} u(\mathbf{x}) y) d\mu = \mu(\hat{\mathcal{K}}_\omega) \int \rho(\mathbf{x}, 1)^{-\gamma} f(g_{\rho(\mathbf{x}, 1)} u(\mathbf{x}) z) d\mu. \quad (3.14)$$

Since k_ω commutes with g_t and f is invariant by k_ω , it follows that $f(z) = f(g_{\rho(\omega,\ell)}u(b_\omega)y)$. In particular, by Lemma 3.10, we have $f(z) > T$. Thus, the contraction property of f in (3) of Definition 3.1 implies

$$\int \rho(\mathbf{x}, 1)^{-\gamma} f(g_{\rho(\mathbf{x},1)}u(\mathbf{x})z) d\mu \leq c f(z) \left(\int \rho(\mathbf{x}, 1) d\mu \right)^{\beta-\gamma}.$$

Finally, we apply Lemma 3.11 to get

$$f(z) \leq \frac{A}{\mu(\hat{\mathcal{K}}_\omega)} \int_{\hat{\mathcal{K}}_\omega} f(g_{\rho(\mathbf{x},\ell)}u(\mathbf{x})y) d\mu. \quad (3.15)$$

Combining (3.12), (3.14), and (3.15), along with the fact that $\rho(\mathbf{x}, \ell) = \rho(\omega, \ell)$ for all $\mathbf{x} \in \hat{\mathcal{K}}_\omega$ yields the desired estimate and concludes the proof. \square

The following lemma uses Lemma 3.12 as a base step in an inductive procedure to establish an exponentially decaying estimate for similar averages over points with long cusp excursions.

Lemma 3.13. *Let $A > 0$ and $B > 0$ be the constants in (3.8) and (3.9) respectively. Let $\gamma \in [\gamma_0, \beta]$ and let $T > 0$ be the constant provided by (3) of Definition 3.1 with $k = 1$. For all $M > ABT$, $m, n \in \mathbb{N}$, $\alpha \in \Lambda^m$ and $x \in X$,*

$$\begin{aligned} \int_{B_x(M, m; n-1) \cap \hat{\mathcal{K}}_\alpha} \rho(\mathbf{x}, m+n)^{-\gamma} f(g_{\rho(\mathbf{x}, m+n)}u(\mathbf{x})x) d\mu &\leq \\ &\leq \theta^n \int_{B_x(M, m; n-1) \cap \hat{\mathcal{K}}_\alpha} \rho(\mathbf{x}, m)^{-\gamma} f(g_{\rho(\mathbf{x}, m)}u(\mathbf{x})x) d\mu, \end{aligned}$$

where θ is given by:

$$\theta = cA \left(\int \rho(\mathbf{x}, 1) d\mu \right)^{(\beta-\gamma)}, \quad (3.16)$$

and c is as in (3) of Definition 3.1.

Proof. If $n = 1$, then $B_x(M, m; n-1) = \emptyset$ and the statement follows trivially. Thus, we may assume that $n > 1$ and that $B_x(M, m, n-1) \neq \emptyset$. Let $H = B_x(M, m; n-1) \cap \hat{\mathcal{K}}_\alpha$. Lemma 2.1 implies that $H \subseteq \bigcup \hat{\mathcal{K}}_\omega$, where the union is taken over words $\omega \in \Lambda^{m+n-1}$ so that $\hat{\mathcal{K}}_\omega \cap H \neq \emptyset$. In particular, we get

$$\begin{aligned} \int_H \rho(\mathbf{x}, m+n)^{-\gamma} f(g_{\rho(\mathbf{x}, m+n)}u(\mathbf{x})x) d\mu &\leq \\ &\leq \sum_{\substack{\omega \in \Lambda^{m+n-1} \\ \hat{\mathcal{K}}_\omega \cap H \neq \emptyset}} \int_{\hat{\mathcal{K}}_\omega} \rho(\mathbf{x}, m+n)^{-\gamma} f(g_{\rho(\mathbf{x}, m+n)}u(\mathbf{x})x) d\mu \leq \\ &\leq \theta \sum_{\substack{\omega \in \Lambda^{m+n-1} \\ \hat{\mathcal{K}}_\omega \cap H \neq \emptyset}} \int_{\hat{\mathcal{K}}_\omega} \rho(\mathbf{x}, m+n-1)^{-\gamma} f(g_{\rho(\mathbf{x}, m+n-1)}u(\mathbf{x})x) d\mu \quad \text{by Lemma 3.12,} \end{aligned} \quad (3.17)$$

where θ is given by (3.16). Next, we note that the following inclusion holds by Lemma 2.1.

$$\bigcup_{\substack{\omega \in \Lambda^{m+n-1} \\ \hat{\mathcal{K}}_\omega \cap H \neq \emptyset}} \hat{\mathcal{K}}_\omega \subseteq \bigcup_{\substack{\omega \in \Lambda^{m+n-2} \\ \hat{\mathcal{K}}_\omega \cap H \neq \emptyset}} \hat{\mathcal{K}}_\omega. \quad (3.18)$$

Moreover, Proposition 2.3 shows that the indicator functions of the above unions are equal almost everywhere to sums of the indicator functions of the members of the union. Hence, combining (3.17), and (3.18) yields the following estimate

$$\begin{aligned} \int_H \rho(\mathbf{x}, m+n)^{-\gamma} f(g_{\rho(\mathbf{x}, m+n)} u(\mathbf{x}) x) d\mu \\ \leq \theta \sum_{\substack{\omega \in \Lambda^{m+n-2} \\ \hat{\mathcal{K}}_\omega \cap H \neq \emptyset}} \int_{\hat{\mathcal{K}}_\omega} \rho(\mathbf{x}, m+n-1)^{-\gamma} f(g_{\rho(\mathbf{x}, m+n-1)} u(\mathbf{x}) x) d\mu. \end{aligned}$$

By an iterated application of (3.17) and (3.18), we obtain

$$\begin{aligned} \int_H \rho(\mathbf{x}, m+n)^{-\gamma} f(g_{\rho(\mathbf{x}, m+n)} u(\mathbf{x}) x) d\mu \\ \leq \theta^{n-1} \sum_{\substack{\omega \in \Lambda^m \\ \hat{\mathcal{K}}_\omega \cap H \neq \emptyset}} \int_{\hat{\mathcal{K}}_\omega} \rho(\mathbf{x}, m+1)^{-\gamma} f(g_{\rho(\mathbf{x}, m+1)} u(\mathbf{x}) x) d\mu. \end{aligned}$$

To apply Lemma 3.12 at this stage, we need to ensure that for each $\omega \in \Lambda^m$ such that $\hat{\mathcal{K}}_\omega \cap H \neq \emptyset$, we have that $f(g_{\rho(\mathbf{x}_0, m)} u(\mathbf{x}_0) x) > AT$ for some $\mathbf{x}_0 \in \hat{\mathcal{K}}_\omega$. Recall that we are assuming that $n > 1$ and $H \neq \emptyset$. Let $\mathbf{x}_0 \in \hat{\mathcal{K}}_\omega \cap H$ for some $\omega \in \Lambda^m$. Then, we have that

$$f(g_{\rho(\mathbf{x}_0, m+1)} u(\mathbf{x}_0) x) \geq M > ABT.$$

Moreover, by Lemma 3.6, we can write $\rho(\mathbf{x}_0, m+1) = \rho(\mathbf{x}_0, m)\rho(\mathbf{y}, 1)$, for some $\mathbf{y} \in \mathcal{K}$. Thus, by the choice of B in (3.9), this implies that $f(g_{\rho(\mathbf{x}_0, m)} u(\mathbf{x}_0) x) > AT$. Hence, we obtain

$$\int_H \rho(\mathbf{x}, m+n)^{-\gamma} f(g_{\rho(\mathbf{x}, m+n)} u(\mathbf{x}) x) d\mu \leq \theta^n \sum_{\substack{\omega \in \Lambda^m \\ \hat{\mathcal{K}}_\omega \cap H \neq \emptyset}} \int_{\hat{\mathcal{K}}_\omega} \rho(\mathbf{x}, m)^{-\gamma} f(g_{\rho(\mathbf{x}, m)} u(\mathbf{x}) x) d\mu.$$

Finally, by Lemma 3.9, it follows that

$$\bigcup_{\substack{\omega \in \Lambda^m \\ \hat{\mathcal{K}}_\omega \cap H \neq \emptyset}} \hat{\mathcal{K}}_\omega \subseteq B_x(M/A, m; n-1) \cap \hat{\mathcal{K}}_\alpha.$$

□

The next ingredient is to provide an upper estimate of the sum appearing in Proposition 3.5 using integral estimates of the height function f . The next lemma is a first step in that direction.

Lemma 3.14. *For all $x \in X$, $\gamma \in \mathbb{R}$, $m, n \in \mathbb{N}$, and any $\alpha \in \Lambda^m$,*

$$\begin{aligned} \sum_{\substack{\omega \in \Lambda^{m+n} \\ \hat{\mathcal{K}}_\omega \cap B_x(M, m; n) \cap \hat{\mathcal{K}}_\alpha \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq \\ \frac{K^{s-\gamma}}{M/A} \int_{B_x(M/A, m; n-1) \cap \hat{\mathcal{K}}_\alpha} \rho(\mathbf{x}, m+n)^{-\gamma} f(g_{\rho(\mathbf{x}, m+n)} u(\mathbf{x}) x) d\mu. \end{aligned}$$

where K is given by (3.7) and A is as in (3.8).

Proof. For any finite word $\omega \in \Lambda^N$, we have

$$\text{diam}(\hat{\mathcal{K}}_\omega) = K\rho(\omega, N).$$

Moreover, we have that $\mu(\hat{\mathcal{K}}_\omega) = \mu(\mathcal{K}_\omega) = \rho(\omega, N)^s$. Indeed, this follows from the self-similarity of μ in (2.9) and Proposition 2.3 showing that the distinct sets \mathcal{K}_ω have null overlap. To simplify notation, let $H = B_x(M, m; n) \cap \hat{\mathcal{K}}_\omega$, $N = m + n$ and define

$$B = \bigcup_{\omega \in \Lambda^N: \hat{\mathcal{K}}_\omega \cap H \neq \emptyset} \hat{\mathcal{K}}_\omega.$$

The other ingredient is to note that $\rho(\mathbf{x}, N) = \rho(\omega, N)$ for every $\mathbf{x} \in \hat{\mathcal{K}}_\omega$. Hence, we get

$$\sum_{\omega \in \Lambda^N: \hat{\mathcal{K}}_\omega \cap H \neq \emptyset} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq K^{s-\gamma} \int \rho(\mathbf{x}, m+n)^{-\gamma} \chi_B(\mathbf{x}) d\mu(\mathbf{x}).$$

In view of Lemma 3.9, we have that

$$B \subseteq B_x(M/A, m; n).$$

Finally, we observe that the following inequality

$$\chi_{B_x(M/A, m; n)}(\mathbf{x}) \leq \frac{f(g_{\rho(\mathbf{x}, m+n)} u(\mathbf{x})x)}{M/A}$$

holds for all $\mathbf{x} \in \mathbb{R}^d$ by definition, where for a set G , χ_G denotes its indicator function. \square

3.7. Proof of Proposition 3.5. Fix some $k \in \mathbb{N}$ and define

$$\mathcal{F}^k = \{h_\alpha : \alpha \in \Lambda^k\}, \quad \mathcal{F}^k := \{f_{\ell k} : \ell \in \mathbb{N}\} \subseteq \mathcal{F}.$$

Our hypothesis implies that μ satisfies the $(\mathcal{F}^k, \mathcal{F}^k, \beta, \gamma_0)$ -contraction hypothesis (note that μ is self-similar with respect to \mathcal{F}^k). In particular, Lemmas 3.6 - 3.14 hold with \mathcal{F}^k and \mathcal{F}^k in place of \mathcal{F} and \mathcal{F} respectively. In our proof below, the only dependence of the constants on k is in the constant B defined in (3.9) (since the cocycle $\rho(\cdot, 1)$ for \mathcal{F}^k is given by $\rho(\cdot, k)$ for \mathcal{F}), and T given by (3) of Definition 3.1. This dependence will appear only in the choice of the constant M_0 in (3.19) below. Since our conclusion states that M_0 depends on k and by replacing \mathcal{F} and \mathcal{F} by \mathcal{F}^k and \mathcal{F}^k respectively, we may hence assume that $k = 1$. For simplicity, we use the following notation:

$$Z_x(M, N, \delta) := Z_x(M, N, 1, \delta), \quad f := f_1.$$

Let $x \in X \setminus \{f = \infty\}$ and $\gamma \in (\gamma_0, \beta)$ be given. Let $T > 0$ be the constant provided by (3) of Definition 3.1 with $k = 1$. We define M_0 as follows

$$M_0 = \max\{ABT, f(x)\}, \tag{3.19}$$

where A and B are the constants in (3.8) and (3.9) respectively. Suppose $M > M_0$, $N \in \mathbb{N}$ and $0 < \delta < 1$ are given. To simplify notation, let

$$\xi = \int \rho(\mathbf{x}, 1)^{-\gamma} d\mu, \quad \zeta = \left(\int \rho(\mathbf{x}, 1) d\mu \right)^{\beta-\gamma}, \quad \theta = cA\zeta, \tag{3.20}$$

where $c \geq 1$ is the constant in (3) of Definition 3.1.

Consider a subset $Q \subseteq \{1, \dots, N\}$ containing at least δN elements. Define the following set of trajectories whose behavior is determined by Q :

$$Z(N, Q) = \{\mathbf{x} \in Z_x(M, N, \delta) : f(g_{\rho(\mathbf{x}, l)} u(\mathbf{x})x) \geq M \text{ iff } l \in Q\}.$$

We decompose the set Q and its complement into maximal "connected" intervals as follows

$$Q = \bigsqcup_{i=1}^q B_i, \quad \{1, \dots, N\} \setminus Q = \bigsqcup_{j=1}^p G_j,$$

for some integers $p, q \geq 0$. Note that $|p - q| \leq 1$. We claim that

$$\sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq (cA)^{3N} (\zeta^\delta \xi^{1-\delta})^N. \quad (3.21)$$

Since the set $Z_x(M, N, \delta)$ is a union of at most 2^N subsets of the form $Z(N, Q)$, the claim of the proposition follows by taking $c_0 = 2(cA)^3$.

Order the intervals B_i and G_j in the way they appear in the sequence $1 \leq \dots \leq N$. Write I_r for the r^{th} interval in this sequence for $1 \leq r \leq p + q$. For a subset $J \subseteq \{1, \dots, N\}$, we use $|J|$ to denote its cardinality. For purposes of induction, we write Λ^0 for a set with one element and $\hat{\mathcal{K}}_\alpha := \mathcal{K}$ whenever $\alpha \in \Lambda^0$.

Case 1: $I_{p+q} \subseteq Q$ so that $I_{p+q} = B_q$. Let $\alpha \in \Lambda^{N-|B_q|}$ be such that $\hat{\mathcal{K}}_\alpha \cap Z(N, Q) \neq \emptyset$. Then, we note that

$$Z(N, Q) \cap \hat{\mathcal{K}}_\alpha \subseteq B_x(M, N - |B_q|; |B_q|) \cap \hat{\mathcal{K}}_\alpha,$$

where the sets $B_x(\cdot, \cdot, \cdot)$ were defined in (3.3). Hence, we may apply Lemma 3.14 with $m = N - |B_q|$ and $n = |B_q|$ to get

$$\begin{aligned} \sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \cap \hat{\mathcal{K}}_\alpha \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} &\leq \sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap B_x(M, m; n) \cap \hat{\mathcal{K}}_\alpha \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \\ &\leq \frac{K^{s-\gamma}}{M/A} \int_{B_x(M/A, m; n-1) \cap \hat{\mathcal{K}}_\alpha} \rho(\mathbf{x}, m+n)^{-\gamma} f(g_{\rho(\mathbf{x}, m+n)} u(\mathbf{x})x) d\mu, \end{aligned}$$

where $K = \text{diam}(\mathcal{K})$. We can then apply Lemma 3.13 to get

$$\sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \cap \hat{\mathcal{K}}_\alpha \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq \frac{\theta^n K^{s-\gamma}}{M/A} \int_{B_x(M, m; n-1) \cap \hat{\mathcal{K}}_\alpha} \rho(\mathbf{x}, m)^{-\gamma} f(g_{\rho(\mathbf{x}, m)} u(\mathbf{x})x) d\mu,$$

where θ is defined in (3.20).

Recall that $\alpha \in \Lambda^m$ was chosen so that $\hat{\mathcal{K}}_\alpha \cap Z(N, Q) \neq \emptyset$. Moreover, the choice of M_0 implies that $1 \in G_1$. In particular, since $I_{p+q} = B_q$ is a maximal sub-interval of Q , we see that $f(g_{\rho(\mathbf{x}_0, m)} u(\mathbf{x}_0)x) < M$ for some $\mathbf{x}_0 \in \hat{\mathcal{K}}_\alpha$. The choice of the constant A in (3.8) then implies that $f(g_{\rho(\mathbf{x}, m)} u(\mathbf{x})x) \leq MA$ for all $\mathbf{x} \in \hat{\mathcal{K}}_\alpha$.

Moreover, $\rho(\mathbf{x}, m)$ is constant everywhere on $\hat{\mathcal{K}}_\alpha$. It follows that

$$\sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \cap \hat{\mathcal{K}}_\alpha \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq \frac{\theta^n K^{s-\gamma}}{M/A} M A \rho(\alpha, m)^{-\gamma} \mu(\hat{\mathcal{K}}_\alpha) = \theta^n A^2 \text{diam}(\hat{\mathcal{K}}_\alpha)^{s-\gamma}.$$

In the last equality, we used the fact that $\text{diam}(\hat{\mathcal{K}}_\alpha) = K\rho(\alpha, m)$ and $\mu(\hat{\mathcal{K}}_\alpha) = \mu(\mathcal{K}_\alpha) = \rho(\alpha, m)^s$. This follows from Lemmas 2.4 and 2.5 respectively. We thus arrive at the following estimate

$$\begin{aligned} \sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} &\leq \sum_{\substack{\alpha \in \Lambda^{N-|B_q|} \\ \hat{\mathcal{K}}_\alpha \cap Z(N, Q) \neq \emptyset}} \sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \cap \hat{\mathcal{K}}_\alpha \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \\ &\leq \theta^{|B_q|} A^2 \sum_{\substack{\alpha \in \Lambda^{N-|B_q|} \\ \hat{\mathcal{K}}_\alpha \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\alpha)^{s-\gamma}. \end{aligned}$$

In view of the inclusion $Z(N, Q) \subseteq Z(N - |B_q|, Q \setminus B_q)$, it follows that

$$\sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq \theta^{|B_q|} A^2 \sum_{\substack{\alpha \in \Lambda^{N-|B_q|} \\ \hat{\mathcal{K}}_\alpha \cap Z(N-|B_q|, Q \setminus B_q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\alpha)^{s-\gamma}. \quad (3.22)$$

Case 2: $I_{p+q} \subseteq \{1, \dots, N\} \setminus Q$ so that $I_{p+q} = G_p$. In this case, we apply Lemma 3.8 to obtain

$$\begin{aligned} \sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} &\leq \sum_{\substack{\alpha \in \Lambda^{N-|G_p|} \\ \hat{\mathcal{K}}_\alpha \cap Z(N, Q) \neq \emptyset}} \sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \cap \hat{\mathcal{K}}_\alpha \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \\ &\leq \sum_{\substack{\alpha \in \Lambda^{N-|G_p|} \\ \hat{\mathcal{K}}_\alpha \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\alpha)^{s-\gamma} \int \rho(\mathbf{x}, |G_p|)^{-\gamma} d\mu(\mathbf{x}). \end{aligned}$$

Then, using Lemma 3.7, it follows that

$$\sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq \xi^{|G_p|} \sum_{\substack{\alpha \in \Lambda^{N-|G_p|} \\ \hat{\mathcal{K}}_\alpha \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\alpha)^{s-\gamma}.$$

Finally, using that $Z(N, Q) \subseteq Z(N - |G_p|, Q)$, we obtain

$$\sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq \xi^{|G_p|} \sum_{\substack{\alpha \in \Lambda^{N-|G_p|} \\ \hat{\mathcal{K}}_\alpha \cap Z(N-|G_p|, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\alpha)^{s-\gamma}. \quad (3.23)$$

Equipped with the estimates in (3.22) and (3.23), we iteratively bound the sum in (3.21) by similar sums over covers of sets of the form $Z(L, V)$ with $L < N$ and $V \subseteq Q$ yielding the following upper bound by downward induction on N :

$$\sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq \theta^{|Q|} A^{2q} \xi^{(N-|Q|)}.$$

Recall that $\theta = cA\zeta$ and $q \leq N$. Moreover, since γ is positive and strictly less than β , we have that $\zeta < 1$, $\xi > 1$. Hence, since $|Q| \geq \delta N$, it follows that

$$\sum_{\substack{\omega \in \Lambda^N \\ \hat{\mathcal{K}}_\omega \cap Z(N, Q) \neq \emptyset}} \text{diam}(\hat{\mathcal{K}}_\omega)^{s-\gamma} \leq (cA)^{3N} (\zeta^\delta \xi^{1-\delta})^N. \quad (3.24)$$

This implies the main claim (3.21).

3.8. Proof of Theorem 3.4. Having established Proposition 3.5, the proof of Theorem 3.4 follows from the definition of Hausdorff dimension. Recall the definition of the Hausdorff (outer) measures given in Section 2.1.

Let $x \in X$ and let $Z_x \subseteq \mathcal{K}$ denote the set of vectors \mathbf{x} for which the trajectory $a_t u(\mathbf{x})x$ diverges on average. As we noted in (3.5), in view of the log Lipschitz property (2) of Definition 3.1, the set Z_x is related to the sets $Z_x(M, N, k, \delta)$, defined in (3.6), via the following inclusion:

$$Z_x \subseteq \liminf_{N \rightarrow \infty} Z_x(M, N, k, \delta) = \bigcup_{N_0 \geq 1} \bigcap_{N \geq N_0} Z_x(M, N, k, \delta), \quad (3.25)$$

for all $M, \delta > 0$ and $k \in \mathbb{N}$. We wish to apply Proposition 3.5. Fix some $\gamma \in (\gamma_0, \beta)$ and let $c_0 \geq 1$ be as in the conclusion of the proposition. Recall the definition of $\xi > 1$ and $\zeta < 1$ in (3.20). Let $\delta \in (0, 1)$ be sufficiently close to 1 such that

$$\zeta^\delta \xi^{1-\delta} < 1.$$

Note that $\delta \rightarrow 1$ as $\gamma \rightarrow \beta$. Choose $k \in \mathbb{N}$ large enough so that

$$(\zeta^\delta \xi^{1-\delta})^k = \left(\int \rho(\mathbf{x}, k) d\mu \right)^{(\beta-\gamma)\delta} \left(\int \rho(\mathbf{x}, k)^{-\gamma} d\mu \right)^{1-\delta} < 1/c_0, \quad (3.26)$$

where we used Lemma 3.7 in the first equality. Let $M_0 = M_0(k, x, \gamma) > 0$ be as in Proposition 3.5 and suppose $M > M_0$. For each $N \in \mathbb{N}$, define

$$\kappa(N) = \max \left\{ \text{diam}(\mathcal{K}_\omega) : \omega \in \Lambda^{kN} \right\}.$$

Suppose $N_0 \in \mathbb{N}$ is given. Then, for every $J \geq N_0$, Proposition 3.5 and the first equality in (3.26) show that

$$H_{\kappa(N_0)}^{s-\gamma} \left(\bigcap_{N \geq N_0} Z_x(M, N, k, \delta) \right) \leq H_{\kappa(N_0)}^{s-\gamma} (Z_x(M, J, k, \delta)) \leq \left(c_0 (\zeta^\delta \xi^{1-\delta})^k \right)^J.$$

By taking J to infinity, it follows that

$$H^{s-\gamma} \left(\bigcap_{N \geq N_0} Z_x(M, N, k, \delta) \right) = 0.$$

Recall that $\dim_H(\cup_n A_n) = \sup_n \dim_H(A_n)$ for any countable collection of Borel sets A_n . Since Z_x is contained in the union of countably many sets of the form $\bigcap_{N \geq N_0} Z_x(M, N, k, \delta)$ by (3.25), it follows that $\dim_H(Z_x) \leq s - \gamma$. This being true for $\gamma_0 \leq \gamma < \beta$, this implies that $\dim_H(Z_x) \leq s - \beta$ as desired.

4. TRANSVERSALITY OF EXPANDING COORDINATES

In this section, we establish the first step towards verifying the contraction hypothesis on the space of lattices $\mathrm{SL}(d+1, \mathbb{R})/\mathrm{SL}(d+1, \mathbb{Z})$. The goal is to prove a key observation which allows us to obtain optimal average contraction rates with respect to any measure μ for which $\alpha_\ell(\mu) > 0$ for every ℓ , where $\alpha_\ell(\mu)$ is defined in (1.3). The main results are Proposition 4.1 and 4.2.

4.1. The exterior power representation. We begin by giving a description of the coordinates of the fundamental representation of $G = \mathrm{SL}(d+1, \mathbb{R})$ on the following vector space

$$V = \bigoplus_{\ell=1}^d \bigwedge^{\ell} \mathbb{R}^{d+1}.$$

An element $g \in G$ acts on V via the linear map $\bigoplus_{\ell=1}^d \bigwedge^{\ell} g$. Consider the basis $\{\mathbf{e}_0, \dots, \mathbf{e}_d\}$ of \mathbb{R}^{d+1} , where \mathbf{e}_i denotes the i^{th} standard basis element. For each index set $I = \{i_1 < \dots < i_\ell\} \subset \{0, \dots, d\}$, we let

$$\mathbf{e}_I := \mathbf{e}_{i_1} \wedge \dots \wedge \mathbf{e}_{i_\ell}. \quad (4.1)$$

The collection of monomials \mathbf{e}_I gives a basis of $V_\ell = \bigwedge^{\ell} \mathbb{R}^{d+1}$ for each $1 \leq \ell \leq d+1$. We denote by $\langle \cdot, \cdot \rangle$ the standard Euclidean inner product on V , making the monomials \mathbf{e}_I an orthonormal basis of V . Note that this basis consists of joint eigenvectors of the linear maps $\bigoplus_{\ell=1}^d \bigwedge^{\ell} g$, where $g \in G$ is any diagonal matrix.

Let $v \in V_I \setminus \{0\}$ and write

$$v = \sum_{I \subset \{1, \dots, d+1\}} v_I \mathbf{e}_I. \quad (4.2)$$

where the sum is over index sets of cardinality l . Let $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_d) \in \mathbb{R}^d$. First, we note that $u(\mathbf{x})$ fixes \mathbf{e}_0 and maps \mathbf{e}_i to $\mathbf{e}_i + \mathbf{x}_i \mathbf{e}_0$ for $i = 1, \dots, d$. This implies the following.

$$u(\mathbf{x})\mathbf{e}_I = \begin{cases} \mathbf{e}_I & 0 \in I, \\ \mathbf{e}_I + \sum_{i \in I} \pm \mathbf{x}_i \mathbf{e}_{(I \cup \{0\}) \setminus \{i\}}, & \text{otherwise,} \end{cases}$$

where the sign depends on I . In particular, we get that

$$u(\mathbf{x})v = \sum_{\substack{I \subset \{0, \dots, d\} \\ 0 \notin I}} v_I \mathbf{e}_I + \sum_{\substack{I \subset \{0, \dots, d\} \\ 0 \in I}} \left(v_I + \sum_{i \notin I} \pm v_{(I \cup \{i\}) \setminus \{0\}} \mathbf{x}_i \right) \mathbf{e}_I. \quad (4.3)$$

4.2. Transversality of the expanding coordinates. Let $v \in V_I$ and write $v_I = \langle v, \mathbf{e}_I \rangle$ for each index set I . For each index set I containing 0, let $\mathcal{L}_I(v)$ be the affine subspace defined by

$$\mathcal{L}_I(v) = \left\{ \mathbf{x} \in \mathbb{R}^d : \langle u(\mathbf{x})v, \mathbf{e}_I \rangle = 0 \right\} = \left\{ \mathbf{x} \in \mathbb{R}^d : v_I + \sum_{i \notin I} \pm v_{(I \cup \{i\}) \setminus \{0\}} \mathbf{x}_i = 0 \right\}.$$

Note that it is possible that $\mathcal{L}_I(v) = \emptyset$ or $\mathcal{L}_I(v) = \mathbb{R}^d$. Denote by n_I the normal vector of $\mathcal{L}_I(v)$ given by

$$n_I = \sum_{i \notin I} (\langle u(\mathbf{e}_i)v, \mathbf{e}_I \rangle - v_I) \mathbf{e}_i = \sum_{i \notin I} \pm v_{(I \cup \{i\}) \setminus \{0\}} \mathbf{e}_i, \quad (4.4)$$

where e_i denotes the standard basis of \mathbb{R}^d and the choice of the signs is the same as in (4.3). Given an index set $J \subset \{1, \dots, d\}$ ($0 \notin J$) of size $\ell = |J|$, define $\mathcal{J}(J)$ by

$$\mathcal{J}(J) = \{I \subset \{0, \dots, d\} : 0 \in I, J = (I \cup \{i\}) \setminus \{0\}, \text{ for some } i \notin I\}. \quad (4.5)$$

We note if $v_J \neq 0$, then $I \in \mathcal{J}(J)$ if and only if $\pm v_J$ appears as a coordinate of n_I .

The following elementary proposition is a key observation for our proof.

Proposition 4.1 (Transversality). *Suppose $v \in V_\ell$ and let $I \subset \{1, \dots, d\}$ be an index set. Let $\ell = \#\mathcal{J}(I)$. Then,*

$$\left\| \bigwedge_{J \in \mathcal{J}(I)} n_J \right\| \geq |v_I|^\ell.$$

Proof. Note that $0 \notin I$ by definition and in particular $\mathcal{J}(I) \neq \emptyset$. Consider the map $\varphi : I \rightarrow \mathcal{J}(I)$ defined by

$$\varphi(j) = (I \cup \{0\}) \setminus \{j\},$$

for every $j \in I$. One easily verifies that φ is a bijection. In particular, $\ell = |I|$. We claim that for each $j \in I$, the following holds:

$$|\langle n_J, e_j \rangle| = \begin{cases} |v_I| & J = \varphi(j), \\ 0 & J \in \mathcal{J}(I), J \neq \varphi(j). \end{cases} \quad (4.6)$$

Indeed, when $J = \varphi(j)$, then $|\langle n_J, e_j \rangle| = |v_I|$ by definition of n_J in (4.4). Otherwise, if $J' \in \mathcal{J}(I)$ satisfies $J' \neq \varphi(j)$, then we observe that $j \in J'$. Indeed, $J' = (I \cup \{0\}) \setminus \{j'\}$ for some $j' \in I$, not equal to j . In this case, the definition of $n_{J'}$ in (4.4) shows that the coefficient of e_j is 0. This completes the proof of (4.6).

Equation (4.6) implies that the $(\ell \times \ell)$ -matrix $(\langle n_J, e_j \rangle)_{J \in \mathcal{J}(I), j \in I}$ is diagonal, up to a permutation of the rows, with $\pm v_I$ on the diagonal. It then follows from the definition of the inner product on $\bigwedge^\ell \mathbb{R}^{d+1}$ that

$$\left| \left\langle \bigwedge_{J \in \mathcal{J}(I)} n_J, e_I \right\rangle \right| = |\det((\langle n_J, e_j \rangle)_{J \in \mathcal{J}(I), j \in I})| = |v_I|^\ell.$$

This completes the proof. \square

For an affine subspace $\mathcal{L} \subset \mathbb{R}^d$ and $\delta > 0$, recall that we denote by $\mathcal{L}^{(\delta)}$ the open δ -neighborhood of \mathcal{L} . More precisely,

$$\mathcal{L}^{(\delta)} = \left\{ \mathbf{x} \in \mathbb{R}^d : d(\mathbf{x}, \mathcal{L}) < \delta \right\},$$

where $d(\cdot, \cdot)$ is the Euclidean distance.

Proposition 4.2 (From Transversality to Integrability). *Suppose $\mathcal{L}_1, \dots, \mathcal{L}_\ell$ are affine hyperplanes in \mathbb{R}^d with $1 \leq \ell \leq d-1$. Suppose that n_k is a unit normal vector of \mathcal{L}_k for each $1 \leq k \leq \ell$. Assume further that*

$$\|n_1 \wedge \dots \wedge n_\ell\| \geq \kappa,$$

for some $\kappa > 0$. Then, there exists $C \geq 1$, depending only on κ and d , so that for all $\varepsilon > 0$,

$$\bigcap_{k=1}^{\ell} \mathcal{L}_k^{(\varepsilon)} \subseteq \left(\bigcap_{k=1}^{\ell} \mathcal{L}_k \right)^{(C\varepsilon)}. \quad (4.7)$$

Proof. Let $P_\ell = \bigcap_{k=1}^\ell \mathcal{L}_k$. First, we claim that $P_\ell \neq \emptyset$. For each k , let $v_k \in \mathbb{R}$ be such that $\mathcal{L}_k = \{\mathbf{x} : \langle n_k, \mathbf{x} \rangle = v_k\}$. Let A' denote the $\ell \times d$ matrix whose k^{th} row is n_k and write $\mathbf{v} \in \mathbb{R}^\ell$ for the column vector whose k^{th} entry is v_k . Then, P_ℓ is the set of solutions of the system $A'\mathbf{x} = \mathbf{v}$. The assumption that $n_1 \wedge \cdots \wedge n_\ell \neq 0$ implies that A' has full rank. In particular, this implies that P_ℓ is non-empty.

By applying a translation, we may assume that $0 \in P_\ell$. Let $\pi : \mathbb{R}^d \rightarrow \mathbb{R}^d/P_\ell \cong \mathbb{R}^\ell$ denote the canonical projection parallel to P_ℓ . Let $L_k = \pi(\mathcal{L}_k)$ for each k . We note that it suffices to show

$$\bigcap_{k=1}^\ell (L_k)^{(\varepsilon)} \subseteq \left(\bigcap_{k=1}^\ell L_k \right)^{(C\varepsilon)},$$

for every $\varepsilon > 0$. Note that $\pi(n_k)$ is a unit vector orthogonal to L_k for each k . We continue to denote by n_k the image of n_k under π .

Suppose $v \in \bigcap_{k=1}^\ell (L_k)^{(\varepsilon)}$. Since each L_k passes through the origin and each n_k is a unit normal to L_k , the Euclidean distance of v to L_k is given by $d(v, L_k) = |\langle n_k, v \rangle|$. In particular, $|\langle n_k, v \rangle| < \varepsilon$. Note that $\bigcap_{k=1}^\ell L_k = \{0\}$. Thus, our task is to show that $\|v\| \leq C\varepsilon$ for an appropriate uniform constant $C > 0$.

Denote by A the square matrix whose rows are n_k . As n_k are unit vectors, there is a constant $K > 0$, depending only on d , such that $\|A\| \leq K$. Moreover, we have that $|\det A| = \|n_1 \wedge \cdots \wedge n_\ell\| \geq \kappa$ and in particular that A is invertible. Recall that if $g \in \text{SL}(\ell, \mathbb{R})$, then $\|g^{-1}\| \leq \|g\|^\ell$. Thus, the following norm estimates follow:

$$\|A^{-1}\| \leq |\det A|^{-1-1/\ell} \|A\|^\ell \leq \kappa^{-1-1/\ell} K^\ell.$$

Moreover, we have that

$$\|A^{-1}\|^{-1} \|v\| \leq \|Av\| \leq \max_k |\langle n_k, v \rangle| < \varepsilon.$$

Together those two inequalities imply that $\|v\| \leq C\varepsilon$, where $C = \kappa^{-1-1/\ell} K^\ell$. □

5. DECAY EXPONENTS, TRANSVERSALITY, AND EXPANSION

This section is dedicated to proving estimates on the average rate of expansion of vectors in linear representations with respect to general measures, Proposition 5.1. The main point of the result is the precise integrability exponent for the functions $\|g_{\tau(\mathbf{x})}u(\mathbf{x})v\|^{-1}$. A key ingredient in the proof is the transversality result obtained in the previous section. Recall the definition of $\alpha_\ell(\mu)$ in 1.3 and the parametrization of g_t in (3.1).

Proposition 5.1. *Let $V_\ell = \bigwedge^\ell \mathbb{R}^{d+1}$ for some $1 \leq \ell < d+1$. Suppose μ is a compactly supported Borel probability measure on \mathbb{R}^d and suppose $\alpha_\ell = \alpha_\ell(\mu) > 0$. Let $0 < \lambda \leq \alpha_\ell$ be given. Then, for all $0 < \delta < 1$, there exists a constant $C = C(\delta, \mu) \geq 1$ so that the following holds: for every $v \in V_\ell \setminus \{0\}$, $\gamma \in \mathbb{R}$, and every measurable function $\tau : \mathbb{R}^d \rightarrow (0, 1)$:*

$$\int \tau(\mathbf{x})^\gamma \|g_{\tau(\mathbf{x})}u(\mathbf{x})v\|^{-\delta\lambda} d\mu(\mathbf{x}) \leq C \|v\|^{-\delta\lambda} \left(\int \tau(\mathbf{x})^{p(\gamma+\kappa)} d\mu \right)^{1/p},$$

where

$$\kappa = \frac{\delta\lambda(d-\ell+1)}{d+1}, \quad p = \frac{1+\delta}{1-\delta}.$$

Remark 5.2. It is worth noting that the constant C in Proposition 5.1 has a delicate dependence on the measure μ . In particular, it depends on the rate of convergence of the \liminf in the definition

of $\alpha_\ell(\mu)$. Moreover, in our proof, $C \rightarrow \infty$ as $\delta\lambda \rightarrow \alpha_\ell$. In particular, it is not clear whether Proposition 5.1 holds with $\delta\lambda = \alpha_\ell$ and with a finite constant C .

Before the proof, we state two elementary lemmas which will be useful for us. The first lemma is immediate from the definition.

Lemma 5.3. *Suppose μ is a Borel measure on \mathbb{R}^d such that $\alpha_\ell(\mu)$ exists for some $1 \leq \ell \leq d$. Then, for every $\eta > 0$, there exists $0 < \varepsilon_0 < 1$ so that for all $0 < \varepsilon \leq \varepsilon_0$*

$$\sup \left\{ \mu \left(\mathcal{L}^{(\varepsilon)} \right) : \mathcal{L} \text{ is a proper affine subspace of dimension } d - \ell \right\} \leq \varepsilon^{\alpha_\ell - \eta}.$$

The next lemma is a simple application of Fubini's Theorem.

Lemma 5.4. *Let μ be a Borel measure and f a non-negative Borel function on a separable metric space X . Then,*

$$\int_X f \, d\mu = \int_0^\infty \mu(\{x \in X : f(x) \geq t\}) \, dt$$

Proof of Proposition 5.1. Let $0 \neq v \in V_\ell$ and without loss of generality assume $\|v\| = 1$. Denote by \mathcal{B} the following collection of vectors

$$\mathcal{B} = \{\mathbf{e}_I : 0 \in I \subset \{0, \dots, d\}\}. \quad (5.1)$$

We use V^+ to denote the linear span of \mathcal{B} and we let $V_\ell^+ = V^+ \cap V_\ell$. Then, V^+ is the expanding subspace corresponding to g_t . Denote by $\pi_+ : V_\ell \rightarrow V_\ell^+$ be the canonical projection. Denote by v^+ the image of v under π_+ and let $v^- = v - v^+$. Define K as follows:

$$K = \max \left\{ 1, \sup_{\mathbf{x} \in \text{supp}(\mu)} \|\mathbf{x}\| \right\}.$$

Then, $K < \infty$ since μ is compactly supported. Observe that for $t > 0$, we have that

$$\|g_t u(\mathbf{x})v\| \geq \|g_t \pi_+(u(\mathbf{x})v)\| = t^{-\frac{d-\ell+1}{d+1}} \|\pi_+(u(\mathbf{x})v)\| \quad (5.2)$$

Fix some $\delta \in (0, 1)$ and let $0 < \lambda \leq \alpha_\ell$ be given. We wish to apply Hölder's inequality. To this end, let

$$\beta = \delta\lambda, \quad q = \frac{1 + \delta}{2\delta}, \quad p = \frac{1 + \delta}{1 - \delta}.$$

Note that p and q are Hölder conjugates. Then, Hölder's inequality and (5.2) imply

$$\begin{aligned} \int \tau(\mathbf{x})^\gamma \|g_{\tau(\mathbf{x})} u(\mathbf{x})v\|^{-\beta} \, d\mu &\leq \int \tau(\mathbf{x})^{\gamma+\kappa} \|\pi_+(u(\mathbf{x})v)\|^{-\beta} \, d\mu \\ &\leq \left(\int \tau(\mathbf{x})^{p(\gamma+\kappa)} \, d\mu \right)^{1/p} \left(\int \|\pi_+(u(\mathbf{x})v)\|^{-q\beta} \, d\mu \right)^{1/q} \end{aligned} \quad (5.3)$$

Hence, it remains to show that the integral of $\|\pi_+(u(\mathbf{x})v)\|^{-q\beta}$ is uniformly bounded. We split the analysis into two cases based on the size of v^- . Recall the expression of $u(\mathbf{x})v$ in standard coordinates given in (4.3).

Case 1: $\|v^-\| \leq 1/3K$. Then, since $\|v\| = 1$, there is an index set I containing 0 so that $|v_I| \geq 2/3$. It follows that for each $\mathbf{x} \in \text{supp}(\mu)$

$$\|\pi_+(u(\mathbf{x})v)\| \geq \left| v_I + \sum_{i \notin I} \pm v_{(I \cup \{i\}) \setminus \{0\}} \mathbf{x}_i \right| \geq |v_I| - \|v^-\| \|x\| \geq \frac{2}{3} - \frac{1}{3} = \frac{1}{3}.$$

This implies that $\left(\int \|\pi_+(u(\mathbf{x})v)\|^{-q\beta} d\mu \right)^{1/q} \leq 3^{-\beta}$ and concludes the proof in this case.

Case 2: $\|v^-\| > 1/3K$. Then, there exists some index set J , not containing 0 so that $|v_J| > 1/3K$. Define $\mathcal{J}(J)$ by

$$\mathcal{J}(J) = \{I \subset \{0, \dots, d\} : 0 \in I, J = (I \cup \{i\}) \setminus \{0\}, \text{ for some } i \notin I\}.$$

Let $I \in \mathcal{J}(J)$ and define n_I by

$$n_I = \sum_{i \notin I} \pm v_{(I \cup \{i\}) \setminus \{0\}} e_i,$$

where the choice of signs is as in (4.3). Note that $\pm v_J$ appears as a coordinate of n_I . In particular, $\|n_I\| \geq |v_J| > 1/3K \neq 0$. Moreover, Proposition 4.1 then shows that $\left\| \bigwedge_{I \in \mathcal{J}(J)} n_I \right\| \geq (3K)^{-\ell}$.

Consider the hyperplane $\mathcal{L}_I = \{\mathbf{x} : \langle n_I, \mathbf{x} \rangle = -v_I\}$. Then, a simple calculation shows

$$d(\mathbf{x}, \mathcal{L}_I) = \frac{|\langle n_I, \mathbf{x} \rangle + v_I|}{\|n_I\|},$$

where $d(\cdot, \cdot)$ denotes Euclidean distance. For each $\varepsilon > 0$, we define the set $E(v, \varepsilon)$ as follows:

$$E(v, \varepsilon) = \{\mathbf{x} \in \text{supp}(\mu) : \|\pi_+(u(\mathbf{x})v)\| \leq \varepsilon\}.$$

Suppose that $\mathbf{x} \in E(v, \varepsilon)$. Then, for each index set I containing 0, the following holds.

$$|\langle n_I, \mathbf{x} \rangle + v_I| = \left| v_I + \sum_{i \notin I} \pm v_{(I \cup \{i\}) \setminus \{0\}} \mathbf{x}_i \right| \leq \varepsilon.$$

It follows that $\mathbf{x} \in \bigcap_{I \in \mathcal{J}(J)} \mathcal{L}_I^{(3K\varepsilon)}$. For simplicity, denote by $\mathcal{L}(J) := \bigcap_{I \in \mathcal{J}(J)} \mathcal{L}_I$. Hence, Proposition 4.2, applied with $\kappa = (3K)^{-\ell}$, implies

$$E(v, \varepsilon) \subseteq (\mathcal{L}(J))^{(C_1\varepsilon)}, \tag{5.4}$$

for some constant $C_1 \geq 1$ depending only on d and K . Applying Lemma 5.4, we obtain

$$\int \|\pi_+(u(\mathbf{x})v)\|^{-q\beta} d\mu = \int_0^\infty \mu \left(E \left(v, r^{-1/q\beta} \right) \right) dr. \tag{5.5}$$

The next ingredient is to apply Lemma 5.3. We observe that $q\beta = (1 + \delta)\lambda/2$ is strictly smaller than α_ℓ , since $0 < \delta < 1$ and $\beta = \delta\lambda$. Let $\eta = (\alpha_\ell - q\beta)/2$. Then, $0 < \eta < \alpha_\ell$. Applying Lemma 5.3 with this η , we get that there exists $\varepsilon_0 > 0$ so that for all $0 < \varepsilon < \varepsilon_0$,

$$\mu \left((\mathcal{L}(J))^{(\varepsilon)} \right) \leq \varepsilon^{\alpha_\ell - \eta}.$$

Here, we use the fact that $\mathcal{L}(J)$ is an affine subspace of dimension $d - \ell$ since it is the intersection of $\#(\mathcal{J}(J)) = \ell$ transverse affine hyperplanes by Proposition 4.1. Let $R > 1$ be sufficiently large

such that $C_1 R^{-1/q\beta} < \varepsilon_0$. Note that the choice of R here depends only on δ and μ . Since μ is a probability measure, it follows that

$$\int_0^R \mu \left(E \left(v, r^{-1/q\beta} \right) \right) dr \leq R. \quad (5.6)$$

Moreover, by (5.4), we obtain

$$\int_R^\infty \mu \left(E \left(v, r^{\frac{-1}{q\beta}} \right) \right) dr \leq \int_R^\infty \mu \left((\mathcal{L}(J)) \left(C_1 r^{\frac{-1}{q\beta}} \right) \right) dr \leq (C_1)^{\alpha_\ell - \eta} \int_R^\infty r^{-\frac{\alpha_\ell - \eta}{q\beta}} dr =: C_2. \quad (5.7)$$

Finally, note that our choice of η implies that $\frac{\alpha_\ell - \eta}{q\beta} = 1 + \frac{\eta}{q\beta} > 1$. In particular, we have that $C_2 < \infty$. Combining (5.3), (5.5), (5.6), and (5.7) concludes the proof. \square

6. HEIGHT FUNCTIONS AND INTEGRAL INEQUALITIES

In this section, we construct a proper function on the space of unimodular lattices. Using the results in the previous section, we verify that this function satisfies the properties listed in Definition 3.1. The key idea that allows converting integral estimates in linear representations into integral estimates over the space of lattices is the use of the so-called systems of integral inequalities which first appeared in [EMM98]. The main result of this section is Theorem 6.3.

6.1. Preliminary notation. Throughout this section, we set

$$G = \mathrm{SL}(d+1, \mathbb{R}), \quad \Gamma = \mathrm{SL}(d+1, \mathbb{Z}), \quad X = G/\Gamma.$$

In view of Proposition 5.1, the results of this section apply to general Borel measures μ and time parametrizations τ on \mathbb{R}^d . However, for our application, we restrict ourselves to the setting of Theorem A. We fix a finite set Λ and an irreducible IFS $\mathcal{F} = \{\rho_i O_i + b_i : i \in \Lambda\}$ on \mathbb{R}^d satisfying the open set condition with limit set \mathcal{K} . We denote by s the Hausdorff dimension of \mathcal{K} and μ the unique self-similar probability measure supported on \mathcal{K} for the canonical probability vector $(\rho_i^s)_i$. Recall that μ in this case coincides with the normalized restriction of the s -dimensional Hausdorff measure to \mathcal{K} .

We denote by V and V_ℓ the following vector spaces, endowed with the standard representations of G ,

$$V_\ell = \bigwedge^\ell \mathbb{R}^{d+1}, \quad V = \bigoplus_{\ell=1}^d V_\ell.$$

Motivated by Proposition 5.1, we define exponents of the form β_ℓ , for $1 \leq \ell \leq d$ as follows:

$$\varpi = \min_{1 \leq \ell \leq d} \alpha_\ell(\mu)(d - \ell + 1), \quad \beta_\ell := \frac{d - \ell + 1}{\varpi}, \quad (6.1)$$

where $\alpha_\ell(\mu)$ was defined in (1.3). Corollary A.10 shows that $\alpha_\ell(\mu) > 0$ since the IFS \mathcal{F} is irreducible.

The space $X = G/\Gamma$ is identified with the space of unimodular lattices in \mathbb{R}^{d+1} via the map $g\mathrm{SL}(d+1, \mathbb{Z}) \mapsto g\mathbb{Z}^{d+1}$. For $x \in X$, let $P(x)$ denote the set of all **primitive** subgroups of the lattice x . Recall that a subgroup L of a lattice x in \mathbb{R}^{d+1} is primitive if $L = \mathbb{Z}^{d+1} \cap \mathrm{span}_{\mathbb{R}}(L)$, where $\mathrm{span}_{\mathbb{R}}(L)$ is the \mathbb{R} -span of any \mathbb{Z} -basis of L . We say a monomial $v_1 \wedge \cdots \wedge v_\ell \in V_\ell$ is **x -integral** if the abelian subgroup of \mathbb{R}^{d+1} generated by $\{v_1, \dots, v_\ell\}$ is primitive, i.e. belongs to $P(x)$.

For every $0 < \ell < d + 1$, we define a function $\varphi_\ell : X \rightarrow [1, \infty)$ as follows: ,

$$\varphi_\ell(x) = \max \left\{ \|v\|^{-1} : v \in V_\ell \text{ is an } x\text{-integral monomial} \right\}. \quad (6.2)$$

For $\ell = 0, d + 1$, set $\varphi_\ell \equiv 1$. For a compact set $Q \subset G$, define

$$\|Q\| = \sup_{g \in Q} \max(\|g\|, \|g^{-1}\|)^{d+1}. \quad (6.3)$$

where $\|\cdot\|$ is the operator norm induced by the Euclidean norm on V . It follows from the definitions that for every $h \in Q$, $\varepsilon > 0$ and every $v \in \bigwedge^* \mathbb{R}^{d+1}$,

$$\|Q\|^{-1} \varphi_\ell(x) \leq \varphi_\ell(hx) \leq \|Q\| \varphi_\ell(x). \quad (6.4)$$

6.2. The Contraction Hypothesis on X . We recall the following Lemma from [EMM98] which underlies the main property of Margulis functions we prove later in the section.

Lemma 6.1 (Lemma 5.6 in [EMM98]). *Let $x \in X$ and let $\Lambda_1, \Lambda_2 \in P(x)$. Then,*

$$\|\Lambda_1\| \|\Lambda_2\| \geq \|\Lambda_1 \cap \Lambda_2\| \|\Lambda_1 + \Lambda_2\|.$$

The following proposition establishes the fundamental property of Margulis functions obtained in [EMM98]. It is obtained via the method of integral inequalities first introduced in [EMM98]

Proposition 6.2. *For every $0 < \varrho < 1$, there exists a constant C , depending only on ϱ and μ , such that for every $k \in \mathbb{N}$, there exists $\omega_0 = \omega_0(k, \mu) \geq 1$, so that for all $x_0 \in X$ and all $\gamma \in \mathbb{R}$*

$$\begin{aligned} \int \rho(\mathbf{x}, k)^\gamma \varphi_\ell^{\varrho/\beta_\ell}(g_{\rho(\mathbf{x}, k)} u(\mathbf{x}) x_0) d\mu(\mathbf{x}) &\leq C \varphi_\ell^{\varrho/\beta_\ell}(x_0) \left(\int \rho(\mathbf{x}, k)^{p(\gamma+\kappa)} d\mu(\mathbf{x}) \right)^{1/p} \\ &\quad + \omega_0^{2\varrho/\beta_\ell} \max_{1 \leq j \leq \min\{\ell, d+1-\ell\}} \left(\sqrt{\varphi_{\ell+j}(x) \varphi_{\ell-j}(x)} \right)^{\varrho/\beta_\ell}, \end{aligned}$$

where

$$\kappa = \frac{\varrho\varpi}{d+1}, \quad p = \frac{1+\varrho}{1-\varrho}, \quad (6.5)$$

and ϖ was defined in (6.1).

Proof. Let $0 < \varrho < 1$ be given. Fix $k \in \mathbb{N}$ and let

$$Q = \{g_{\rho(\mathbf{x}, k)} u(\mathbf{x}) : \mathbf{x} \in \text{supp}(\mu)\}.$$

Let $\omega = \|Q\|$ as defined in (6.3) for Q as above. Following [EMM98], let Ψ_ℓ denote the finite subset of $P(x_0)$ of rank ℓ subgroups L of x_0 satisfying

$$\|L\|^{-1} \geq \omega^{-1} \varphi_\ell(x_0).$$

The finiteness of Ψ follows from the discreteness of the lattice x_0 . Suppose that Ψ_ℓ consists of a single element and denote it by Λ_ℓ . In this case, by (6.4), we see that for all $\mathbf{x} \in \text{supp}(\mu)$,

$$\varphi_\ell(g_{\rho(\mathbf{x}, k)} u(\mathbf{x}) x_0) = \|g_{\rho(\mathbf{x}, k)} u(\mathbf{x}) \Lambda_\ell\|^{-1}.$$

Observe further that the definition of the exponents β_ℓ implies $\lambda = \frac{1}{\beta_\ell} \leq \alpha_\ell(\mu)$, for each ℓ . In particular, upon applying Proposition 5.1 with $v = \Lambda_\ell$, $\lambda = 1/\beta_\ell$, $\tau(\cdot) = \rho(\cdot, k)$, and $\delta = \varrho$, we

obtain

$$\begin{aligned} \int \rho(\mathbf{x}, k)^\gamma \varphi_\ell^{\varrho/\beta_\ell}(g_{\rho(\mathbf{x}, k)}u(\mathbf{x})x_0) d\mu &= \int \rho(\mathbf{x}, k)^\gamma \|g_{\rho(\mathbf{x}, k)}u(\mathbf{x})\Lambda_\ell\|^{-\varrho/\beta_\ell} d\mu \\ &\leq C \varphi_\ell^{\varrho/\beta_\ell}(x_0) \left(\int \rho(\mathbf{x}, k)^{p(\gamma+\kappa)} d\mu(\mathbf{x}) \right)^{1/p}, \end{aligned} \quad (6.6)$$

where $C \geq 1$ is the constant in Proposition 5.1 and p and κ are as in (6.5).

Alternatively, suppose the cardinality of Ψ_ℓ is at least 2. Let $\Lambda_\ell \in \Psi_\ell$ be such that $\varphi_\ell(x_0) = \|\Lambda_\ell\|^{-1}$, and let $L \neq \Lambda_\ell$ be another element of Ψ_ℓ . Then, the group $\Lambda_\ell + L$ has rank $\ell + j$ for some $j > 0$. Moreover, by definition of Ψ_ℓ , $\|\Lambda_\ell\|^{-1} \leq \omega \|L\|^{-1}$. Hence, in view of Lemma 6.1, for all $\mathbf{x} \in \text{supp}(\mu)$,

$$\begin{aligned} \varphi_\ell(g_{\rho(\mathbf{x}, k)}u(\mathbf{x})x_0) &\leq \omega \varphi_\ell(x_0) = \frac{\omega}{\|\Lambda_\ell\|} \leq \frac{\omega^2}{\sqrt{\|\Lambda_\ell\| \|L\|}} \leq \frac{\omega^2}{\sqrt{\|\Lambda_\ell + L\| \|\Lambda_\ell \cap L\|}} \\ &\leq \omega^2 \max_{1 \leq j \leq \min\{\ell, d+1-\ell\}} \sqrt{\varphi_{\ell+j}(x_0) \varphi_{\ell-j}(x_0)}. \end{aligned}$$

Combining this estimate with (6.6), we get the desired conclusion with

$$\omega_0 = \omega \max \{ \rho(\mathbf{x}, k)^\gamma : \mathbf{x} \in \text{supp}(\mu) \}.$$

□

Given $\varepsilon > 0$ and $0 < \varrho < 1$, we define the function $f_{\varepsilon, \varrho} : X \rightarrow \mathbb{R}_+$ by

$$f_{\varepsilon, \varrho}(x) = 2 + \sum_{\ell=1}^d \varepsilon^\ell \varphi_\ell^{\varrho/\beta_\ell}(x). \quad (6.7)$$

Define β as follows

$$\beta = \min_{1 \leq \ell \leq d} \frac{\alpha_\ell(\mu)(d - \ell + 1)}{d + 1} = \frac{\varpi}{d + 1}. \quad (6.8)$$

We are now ready to verify the contraction hypothesis in our context. The idea of the deduction of the following result from Proposition 6.2 is due to Margulis and first appeared in [EMM98]. In that article, all the exponents of φ_ℓ were the same. Since our exponents $1/\beta_\ell$ are distinct, this introduces a complication which we address in the proof. The reader may wish to consult variants of this idea in [KKLM17, Proposition 4.1] and [BQ11, Claim 5.9].

Theorem 6.3. *For every $0 < \varrho < 1$, there exists a constant C_1 , depending only on ϱ and μ , such that for every $k \in \mathbb{N}$, there exists $b = b(k, \mu) \geq 1$ and $0 < \varepsilon_0 < 1$, depending only on k and μ , so that for all $x_0 \in X$ and all $\gamma \in \mathbb{R}$ satisfying $\varrho\beta - \frac{1-\varrho}{1+\varrho} \leq \gamma \leq \varrho\beta$,*

$$\int \rho(\mathbf{x}, k)^{-\gamma} f_{\varepsilon_0, \varrho}(g_{\rho(\mathbf{x}, k)}u(\mathbf{x})x_0) d\mu(\mathbf{x}) \leq C_1 \left(\int \rho(\mathbf{x}, k) d\mu(\mathbf{x}) \right)^{\varrho\beta-\gamma} f_{\varepsilon_0, \varrho}(x_0) + b.$$

Remark 6.4. The proof of Theorem 6.3 will show that the lower bound restriction on γ is only for aesthetic reasons.

We first prove the following elementary but crucial lemma.

Lemma 6.5. *For all natural numbers $0 < j, \ell < d + 1$, we have*

$$\ell - \frac{\beta_{\ell-j}(\ell - j) + \beta_{\ell+j}(\ell + j)}{2\beta_\ell} \geq \frac{1}{d}. \quad (6.9)$$

Proof. We note that the function $q(x) = x(d+1-x)$ satisfies the following concavity property:

$$\frac{q(\ell-j) + q(\ell+j)}{2} = q(\ell) - j^2.$$

It follows that

$$\ell - \frac{\beta_{\ell-j}(\ell-j) + \beta_{\ell+j}(\ell+j)}{2\beta_\ell} = \ell - \frac{q(\ell) - j^2}{d-\ell+1} = \frac{j^2}{d-\ell+1}.$$

Since $j^2 \geq 1$ and $d+1-\ell \leq d$, the lemma follows. \square

Proof of Theorem 6.3. Fix $k \in \mathbb{N}$ and $0 < \varrho < 1$. Let $C, \omega_0 \geq 1$ be the constants provided by Proposition 6.2. To simplify notation, let

$$a = C \left(\int \rho(\mathbf{x}, k)^{p(\kappa-\gamma)} d\mu(\mathbf{x}) \right)^{1/p},$$

where p and κ are as in (6.5). Note that a depends on k . Let $0 < \varepsilon_0 < 1$ be a constant to be determined. Suppose $x_0 \in X$. It follows from Proposition 6.2 that

$$\begin{aligned} \int \rho(\mathbf{x}, k)^{-\gamma} f_{\varepsilon_0, \varrho}(g_{\rho(\mathbf{x}, k)} u(\mathbf{x}) x_0) d\mu &= 2 + \sum_{\ell=1}^d \varepsilon_0^\ell \int \rho(\mathbf{x}, k)^{-\gamma} \varphi_\ell^{\varrho/\beta_\ell}(g_{\rho(\mathbf{x}, k)} u(\mathbf{x}) x_0) d\mu \leq \\ &\leq 2 + a \sum_{\ell=1}^d \varepsilon_0^\ell \varphi_\ell^{\varrho/\beta_\ell}(x_0) + \omega_0^{2\varpi} \sum_{\ell=1}^d \varepsilon_0^\ell \max_{1 \leq j \leq \min\{\ell, d+1-\ell\}} \left(\sqrt{\varphi_{\ell+j}(x_0) \varphi_{\ell-j}(x_0)} \right)^{\varrho/\beta_\ell} \\ &= a f_{\varepsilon_0, \varrho}(x_0) + 2(1-a) + \omega_0^{2\varpi} \sum_{\ell=1}^d \varepsilon_0^\ell \max_{1 \leq j \leq \min\{\ell, d+1-\ell\}} \left(\sqrt{\varphi_{\ell+j}(x_0) \varphi_{\ell-j}(x_0)} \right)^{\varrho/\beta_\ell}, \end{aligned} \quad (6.10)$$

where we used the fact that $\omega_0^{2\varrho/\beta_\ell} \leq \omega_0^{2\varpi}$. We observe that the exponents β_ℓ satisfy the following relation:

$$\beta_{\ell-j} + \beta_{\ell+j} = 2\beta_\ell,$$

for all $0 < j, \ell < d+1$. In particular, this implies

$$\left(\sqrt{\varphi_{\ell+j}(x_0) \varphi_{\ell-j}(x_0)} \right)^{\varrho/\beta_\ell} \leq \varepsilon_0^{\frac{-(\beta_{\ell-j}(\ell-j) + \beta_{\ell+j}(\ell+j))}{2\beta_\ell}} f_{\varepsilon_0, \varrho}(x_0). \quad (6.11)$$

Moreover, Lemma 6.5 shows that for all $0 < j, \ell < d+1$:

$$\ell - \frac{\beta_{\ell-j}(\ell-j) + \beta_{\ell+j}(\ell+j)}{2\beta_\ell} \geq \frac{1}{d}. \quad (6.12)$$

Applying the estimates (6.11) and (6.12) to the last sum in (6.10), using the fact that $\varepsilon_0 < 1$, yields

$$\int \rho(\mathbf{x}, k)^{-\gamma} f_{\varepsilon_0, \varrho}(g_{\rho(\mathbf{x}, k)} u(\mathbf{x}) x_0) d\mu \leq a f_{\varepsilon_0, \varrho}(x_0) + 2(1-a) + d\varepsilon_0^{1/d} \omega_0^{2\varpi} f_{\varepsilon_0, \varrho}(x_0).$$

Choosing $\varepsilon_0^{1/d} = \frac{a}{\omega_0^{2\varpi d}}$, $b = 2(1-a)$, and $C_1 = 2C$, we obtain

$$\int_{\mathcal{K}} \rho(\mathbf{x}, k)^{-\gamma} f_{\varepsilon_0, \varrho}(g_{\rho(\mathbf{x}, k)} u(\mathbf{x}) x_0) d\mu(\mathbf{x}) \leq C_1 \left(\int_{\mathcal{K}} \rho(\mathbf{x}, k)^{p(\kappa-\gamma)} d\mu(\mathbf{x}) \right)^{1/p} f_{\varepsilon_0, \varrho}(x_0) + b.$$

Note that b also depends on k . Since $\varrho\beta - \frac{1}{p} \leq \gamma \leq \varrho\beta = \kappa$, Jensen's inequality implies that

$$\left(\int_{\mathcal{K}} \rho(\mathbf{x}, k)^{p(\kappa-\gamma)} d\mu(\mathbf{x}) \right)^{1/p} \leq \left(\int_{\mathcal{K}} \rho(\mathbf{x}, k) d\mu(\mathbf{x}) \right)^{\varrho\beta-\gamma},$$

thus completing the proof. \square

To demonstrate the power of the method of integral inequalities, we state a consequence of the above analysis, which can be obtained with a little more work. We do not need this statement for our purposes, so we omit the proof and refer the interested reader to [BQ11, Claim 5.9] for the proof of a similar statement. Given $\varepsilon > 0$, consider the function $F_\varepsilon : X \rightarrow \mathbb{R}_+$ defined by

$$F_\varepsilon(x) = \max \varepsilon^{\varpi\ell} \|v\|^{\frac{-1}{\beta_\ell}}, \quad (6.13)$$

where the maximum is taken over all x -integral monomials $v \in V_\ell$ and all $0 < \ell < d + 1$.

Proposition 6.6. *For every compact set $Q \subset G$, there exist constants $C_1 \geq 1$ and $\varepsilon_0 > 0$, depending only on Q , such that for every $x \in X$, whenever $F_{\varepsilon_0}(x) > C_1$, the set $\Psi(x)$ of x -integral monomials v satisfying*

$$\varepsilon_0^{\varpi\ell} \|v\|^{\frac{-1}{\beta_\ell}} \geq F_{\varepsilon_0}(x) / \|Q\|^{2\varpi}, \quad (6.14)$$

contains at most one primitive vector up to a sign in each V_ℓ with $0 < \ell < d + 1$.

6.3. Proof of Theorem 1.12 and its corollaries. Let $0 < \varrho < 1$ be given. For every k , let $\varepsilon_0(k)$ be the constant provided by Theorem 6.3. Consider the collection of height functions:

$$\mathcal{F} = \{f_k = f_{\varepsilon_0(k), \varrho} : k \in \mathbb{N}\}.$$

Theorem 6.3 shows that the action of G on $X = G/\Gamma$ satisfies the $(\mathcal{F}, \mathcal{F}, \beta, \gamma_0)$ -contraction hypothesis with β and γ_0 given by:

$$\beta = \frac{\varrho\varpi}{d+1}, \quad \gamma_0 = \frac{\varrho\varpi}{d+1} - \frac{1-\varrho}{1+\varrho},$$

where ϖ was defined in (6.1). Indeed, the contraction property (3) of Definition 3.1 follows from Theorem 6.3, where T can be chosen as follows:

$$T = \frac{b}{C_1 \left(\int \rho(\mathbf{x}, k) d\mu(\mathbf{x}) \right)^{\varrho\beta-\gamma}}.$$

The remaining properties follow directly from the definition of $f_{\varepsilon_0, \varrho}$.

Hence, Theorem 3.4 applies and shows that the dimension of the set in the conclusion of Theorem 1.12 is at most $s - \varrho\varpi/(d+1)$. Since $0 < \varrho < 1$ was arbitrary, this completes the proof.

Theorem A follows from Theorem 1.12 by taking x_0 to be the identity coset and using Dani's correspondence. Corollary 1.4 follows from Theorem A along with Lemma A.1 showing that $\alpha_1(\mu) = \log 2 / \log 3$ in this case. Finally, in the setting of Corollary 1.5, it is shown in [Shm16, Theorem 8.2] that $\dim_\infty(\pi_\theta\mu) = \min\{\dim_H(\mathcal{K}), 1\}$ for every $\theta \in [0, 2\pi)$, where $\pi_\theta\mu$ is the projection of μ in direction θ and $\dim_\infty(\pi_\theta\mu)$ is the Frostman exponent of $\pi_\theta\mu$ defined in (1.7). Hence, applying Theorem 1.9, we get that $\alpha_1(\mu) = \min\{\dim_H(\mathcal{K}), 1\}$. Corollary 1.5 then follows from Theorem A.

7. FRACTALS AND THE TEICHMÜLLER FLOW

The goal of this section is to prove Theorem B. The height functions needed to apply Theorem 3.4 in this context were constructed by Eskin and Masur in [EM01]. In using this construction, we apply

Proposition 5.1 with $G = \mathrm{SL}(2, \mathbb{R})$. We remark that the proof of Proposition 5.1 simplifies in this case and does not require the results of Section 4.

7.1. Background and Definitions. For background on Teichmüller dynamics and translation flows, the reader is referred to [FM14] for an excellent survey. Suppose S is a compact oriented surface of genus $g \geq 1$. An abelian differential on S is an isotopy class of pairs (M, ω) , where M is a Riemann surface structure on S and ω is a holomorphic 1-form. Then, ω induces a (possibly singular) flat metric on S . A unit area abelian differential is one in which M has area 1 in the induced flat metric. A saddle connection of (M, ω) is a flat geodesic segment joining zeros of ω .

If $\Sigma \subset S$ denotes the set of zeros of ω , then $S \setminus \Sigma$ admits an atlas of charts to the complex plane so that all transition maps are given by translations. In these coordinates, ω is given by the pull-back of the canonical holomorphic 1-form dz on \mathbb{C} . Moreover, the “vertical” vector field (parallel to the imaginary axis) on \mathbb{C} induces a well-defined vector field on $S \setminus \Sigma$. The flow defined by this vector field is referred to as the vertical flow on (M, ω) . The induced area measure by ω is invariant for this flow. The flow is said to be uniquely ergodic if this is the only invariant measure.

Let $\alpha = (\alpha_1, \dots, \alpha_n)$ be an integral partition of $2g - 2$, i.e. $\alpha_i \in \mathbb{N}$ and $\sum \alpha_i = 2g - 2$. By a stratum of abelian differentials of order α , we mean the space of unit area abelian differentials on S whose zeroes have multiplicities $\alpha_1, \dots, \alpha_n$. Strata of abelian differentials are non-compact. This can be seen by taking a sequence of abelian differentials in which the systole on the associated Riemann surface tends to 0. If the integral partition defining the stratum contains two distinct elements, then a sequence of abelian differentials may “diverge” if the distance between two distinct zeros tends to 0.

There are local coordinates on a stratum into \mathbb{C}^N for appropriate N , called period coordinates (e.g., see [FM14, Section 2.3] for details), such that all changes of coordinates are given by affine maps. In period coordinates, $\mathrm{SL}(2, \mathbb{R})$ acts naturally on each copy of \mathbb{C} . The Teichmüller geodesic flow is the flow induced by the action of the diagonal group in $\mathrm{SL}(2, \mathbb{R})$. The behavior of the orbit of (M, ω) under $\mathrm{SL}(2, \mathbb{R})$ and its various subgroups determines many of the ergodic properties of the vertical flow. Most relevant to our application is Masur’s criterion asserting that the a_t -orbit of (M, ω) diverges if the vertical flow defined by ω is not uniquely ergodic.

In the sequel, we fix one such stratum and denote it by \mathcal{H} . For simplicity, we use ω to denote elements of \mathcal{H} .

7.2. The Contraction Hypothesis on strata of abelian differentials. Suppose ν is a Borel probability measure on $\mathrm{SL}(2, \mathbb{R})$. We say ν is (a, α) -linearly expanding on \mathbb{R}^2 for some constants $a, \alpha > 0$ if for all $v \in \mathbb{R}^2 \setminus \{0\}$ the following holds

$$\int \|gv\|^{-\alpha} d\nu(g) \leq a \|v\|^{-\alpha}.$$

For any $g \in \mathrm{SL}(2, \mathbb{R})$, denote by $\|g\|$ the operator norm of g in its action on \mathbb{R}^2 . Given a compact set $Q \subset \mathrm{SL}(2, \mathbb{R})$, define $\|Q\|$ as follows:

$$\|Q\| = \sup_{g \in Q} \max(\|g\|, \|g^{-1}\|). \quad (7.1)$$

Theorem 7.1 (Lemma 7.5, [EM01]). *There exist $p_0 = p_0(\mathcal{H}) \in \mathbb{N}$ such that the following holds. Suppose $\alpha > 0$ is given. Then, there exist functions $f_{i,\alpha} : \mathcal{H} \rightarrow [1, \infty)$ for each $1 \leq i \leq p_0$ such that $f_{1,\alpha}$ is a proper function and each $f_{i,\alpha}$ satisfies the log-Lipschitz property in Def. 3.1(2). Moreover, suppose ν is a compactly supported Borel probability measure on $\mathrm{SL}(2, \mathbb{R})$ which is (a, α) -linearly expanding on \mathbb{R}^2 . Then, there exist constants ω and b , depending only on a , α and $\|\mathrm{supp}(\nu)\|$*

(cf. (7.1)), such that for every $1 \leq i \leq p_0$ and $x_0 \in \mathcal{H}$:

$$\int f_{i,\alpha}(gx_0) d\nu(g) \leq a f_{i,\alpha}(x_0) + b + w \sum_{j>i} f_{j,\alpha}(x_0).$$

Proof. Theorem 7.1 was obtained in [EM01, Lemma 7.5] for the special measures $d\nu = \delta_{\alpha_t r_\theta} d\theta$ for any $t > 0$, where $d\theta$ is the normalized Lebesgue measure on $[0, 2\pi)$. The main part of the proof is the construction of the functions $f_{i,\alpha}$ in [EM01, p.464] (denoted α_i in *loc. cit.*) using the notion of admissible complexes. The definition of these functions depends on a parameter δ which we take to be $\alpha - 1$ in our notation. Inspection of the (short) proof of [EM01, Lemma 7.5] shows that the only input used to establish the desired contraction property is the (a, α) -linear expansion and the compactness of the support of the measure. The other parts of the argument are independent of the shape of the measure. \square

Remark 7.2. The function $f_{1,\alpha}$ in Theorem 7.1 is given by a power of the reciprocal of the length of the shortest saddle connection.

Corollary 7.3 (Lemma 2.10, [Ath06]). *Suppose p_0, ν, b and ω are as in Theorem 7.1. Then, there exist constants $\varepsilon_i > 0$, depending only on a and ω , such that the following holds. Let $\varepsilon = (\varepsilon_i) \in \mathbb{R}^{p_0}$ and let $f_\varepsilon = \sum_i \varepsilon_i f_{i,\alpha}$. Then, for all $x_0 \in \mathcal{H}$:*

$$\int f_\varepsilon(gx_0) d\nu(g) \leq 2a f_\varepsilon(x_0) + b p_0.$$

Proof. Let $\varepsilon_i = (1 + a/\omega)^{i-1}$. Then, one verifies that $\sum_{j<i} \varepsilon_j \leq \varepsilon_i a/\omega$ for each i . It follows that

$$\begin{aligned} \int f_\varepsilon(gx_0) d\nu(g) &\leq a f_\varepsilon(x_0) + p_0 b + \omega \sum_{i=1}^{p_0} \varepsilon_i \sum_{j>i} f_{j,\alpha}(x_0) \\ &= a f_\varepsilon(x_0) + p_0 b + \omega \sum_{i=1}^{p_0} f_{i,\alpha}(x_0) \sum_{j<i} \varepsilon_j \leq 2a f_\varepsilon(x_0) + p_0 b. \end{aligned}$$

\square

7.3. Proof of Theorem B and Corollary 1.8. We wish to apply Theorem 3.4. Let $\mathcal{F} = \{h_i : i \in \Lambda\}$ be an IFS as in the statement of the theorem and let \mathcal{K} be its limit set. Denote by $s = \dim_H(\mathcal{K})$ and μ the normalized restriction of H^s to \mathcal{K} . Then, $\alpha_1(\mu) = s$ by Proposition 2.2.

Fix $k \in \mathbb{N}$ and $0 < \delta < 1$. Let $\kappa = \delta s/2$ and $p = (1 + \delta)/(1 - \delta)$. Suppose γ is any number satisfying $\kappa - (1/p) \leq \gamma \leq \kappa$ and let $\tau(\mathbf{x}) = \rho(\mathbf{x}, k)$. Define a Borel measure ν on $\text{SL}(2, \mathbb{R})$ by

$$\int \varphi d\nu = \int \varphi(g_{\tau(\mathbf{x})} u(\mathbf{x})) \tau(\mathbf{x})^{-\gamma} d\mu(\mathbf{x}),$$

for every compactly supported continuous function φ on $\text{SL}(2, \mathbb{R})$.

Applying Proposition 5.1 with $G = \text{SL}(2, \mathbb{R})$, $V_1 = \mathbb{R}^2$, and $\lambda = \alpha_1(\mu) = s$ then shows that ν is $(a, \delta s)$ -linearly expanding on \mathbb{R}^2 , with

$$a = C \left(\int \tau(\mathbf{x})^{p(\kappa-\gamma)} d\mu \right)^{1/p},$$

for a constant $C \geq 1$ depending only on δ and μ . Hence, Theorem 7.1 and Corollary 7.3 apply and provide, for every $k \in \mathbb{N}$, constants $\varepsilon(k) = (\varepsilon_i) \in \mathbb{R}_{>0}^{p_0}$ and \bar{b} (depending on k) and a function $f_{\varepsilon(k)}$

such that for every $\omega \in \mathcal{H}$,

$$\int \tau(\mathbf{x})^{-\gamma} f_{\varepsilon(k)}(g_{\tau(\mathbf{x})}u(\mathbf{x})\omega) d\mu(\mathbf{x}) \leq 2af_{\varepsilon(k)}(\omega) + \bar{b} \leq 2Cf_{\varepsilon(k)}(\omega) \left(\int \tau(\mathbf{x}) d\mu \right)^{\kappa-\gamma} + \bar{b},$$

by Jensen's inequality. For $k \in \mathbb{N}$, define

$$f_k := f_{\varepsilon(k)}, \quad T = \frac{\bar{b}}{4C \left(\int \tau(\mathbf{x}) d\mu \right)^{\kappa-\gamma}}.$$

Fix $k \in \mathbb{N}$ and suppose that $f_k(\omega) > T$. Then, the above estimate becomes

$$\int \tau(\mathbf{x})^{-\gamma} f_k(g_{\tau(\mathbf{x})}u(\mathbf{x})\omega) d\mu(\mathbf{x}) \leq 2Cf_k(\omega) \left(\int \tau(\mathbf{x}) d\mu \right)^{\kappa-\gamma} + \bar{b} \leq 4Cf_k(\omega) \left(\int \tau(\mathbf{x}) d\mu \right)^{\kappa-\gamma}.$$

Let $c = 4C$, $\beta = \kappa$, $\gamma_0 = \kappa - (1/p)$, and $\mathcal{F} = \{f_k : k \in \mathbb{N}\}$. The above argument shows that μ satisfies the $(\mathcal{F}, \mathcal{F}, \beta, \gamma_0)$ -contraction hypothesis. In particular, this completes the verification of the hypotheses of Theorem 3.4 in this setting and shows that the dimension of divergent on average directions belonging to \mathcal{K} is at most $\delta s/2$. Since $\delta \in (0, 1)$ was arbitrary, Theorem B follows.

To prove Corollary 1.8, we first observe that the restriction of the map \arctan to (the compact set) \mathcal{K} is bi-Lipschitz onto its image. Moreover, we note that for $\theta \in (-\pi/2, \pi/2)$,

$$r_\theta = {}^t u(-\tan \theta) a_{\log \cos \theta} u(\tan \theta),$$

where for $g \in \mathrm{SL}(2, \mathbb{R})$, ${}^t g$ denotes its transpose. Since a_t contracts ${}^t u(-\tan \theta)$ and commutes with $a_{\log \cos \theta}$, it follows that the orbit $(a_t r_\theta \omega)_{t \geq 0}$ diverges on average in \mathcal{H} if and only if the orbit $(a_t u(\tan \theta) \omega)_{t \geq 0}$ does. Since bi-Lipschitz maps preserve Hausdorff dimension, the corollary follows from Theorem B.

APPENDIX A. FROSTMAN EXPONENTS OF PROJECTIONS OF SELF-SIMILAR MEASURES

The goal of this section is to complete the proofs of Corollary 1.4 by computing $\alpha_1(\mu)$ in this case in Lemma A.1. We also provide a proof of Theorem 1.9 relating $\alpha_1(\mu)$ to the Frostman exponents of projections of μ for planar homogeneous fractal measures, thus completing the proof of Corollary 1.5. Finally, we show that the limit in the definition of the exponents $\alpha_\ell(\mu)$ exists for every ℓ (Proposition A.7) in full generality.

Notation. If (X, μ) is a measure space and $f : X \rightarrow Y$ is a measurable map, we denote by $f\mu$ the push-forward measure.

A.1. Projections of products of Cantor sets. Consider the IFS \mathcal{F} on \mathbb{R}^2 given by maps of the form

$$h_v(\mathbf{x}) = \frac{\mathbf{x} + v}{3}, \quad v \in \Lambda := \{0, 2\}^2.$$

The limit set \mathcal{K} of \mathcal{F} coincides with a product of 2 copies of Cantor's middle thirds set. For convenience, let Λ^0 denote the set consisting of a single point and, for $\omega \in \Lambda^0$, denote by h_ω the identity mapping. For $\omega \in \Lambda^n$, h_ω takes the form $\mathbf{x} \mapsto 3^{-n} \cdot \mathbf{x} + b_\omega$, where $b_\omega \in \mathcal{C} \times \mathcal{C}$ is a rational vector satisfying $3^n b_\omega \in \mathbb{Z}^2$.

The following lemma computes the value of $\alpha_1(\mu)$ in this case and completes the proof of Corollary 1.4. We thank Pablo Shmerkin for providing its proof.

Lemma A.1. *Suppose $\mathcal{K} = \mathcal{C} \times \mathcal{C}$, where \mathcal{C} is Cantor's middle-thirds set and let $s = \dim_H(\mathcal{K})$. Let μ denote $H^s|_{\mathcal{K}}$. Then, $\alpha_1(\mu) = \log 2 / \log 3$.*

Proof. That $\alpha_1(\mu) \leq \log 2 / \log 3$ follows from projecting μ onto the coordinate axes. For the reverse inequality, it suffices to show that for each n and each affine line \mathcal{L}

$$\mu\left(\mathcal{L}^{(3^{-n})}\right) \ll 2^{-n}, \quad (\text{A.1})$$

where the implied constant is absolute. Let $I = [0, 1]$. We claim that for every line \mathcal{L} , the number of squares of the form $h_\omega(I^2)$ with $\omega \in \Lambda^n$ which meet \mathcal{L} is $\ll 2^n$, with an absolute implied constant. Assuming the claim, fix a line \mathcal{L} and $n \in \mathbb{N}$, and let $\mathcal{L}_1, \mathcal{L}_2$ denote the 2 lines parallel to \mathcal{L} and bounding the neighborhood $\mathcal{L}^{(3^{-n})}$. Then, we note that $\mathcal{L}^{(3^{-n})} \cap \mathcal{K}$ is contained in the union of the squares $h_\omega(I^2)$ which meet either of $\mathcal{L}, \mathcal{L}_1$, or \mathcal{L}_2 , for some $\omega \in \Lambda^n$. Moreover, μ assigns mass 4^{-n} to each such square. It follows that $\mu\left(\mathcal{L}^{(3^{-n})}\right) \ll 3 \times 2^{-n}$, and hence (A.1) follows.

For a line \mathcal{L} , denote by $b_\mathcal{L}$ the x -coordinate of the intersection of \mathcal{L} with the x -axis. For $i = 0, 1$, denote by \mathcal{L}_i the line of slope 1 such that $b_{\mathcal{L}_i} = (-1)^i/3$. Denote by \mathcal{L}'_i the line of slope -1 such that $b_{\mathcal{L}'_i} = (-1)^i/3 + 1$. We say a line \mathcal{L} is exceptional if

$$\mathcal{L} \in \left\{ h_\omega(\mathcal{L}_i) : \omega \in \bigcup_{n \geq 0} \Lambda^n, i = 1, 2 \right\} \cup \left\{ h_\omega(\mathcal{L}'_i) : \omega \in \bigcup_{n \geq 0} \Lambda^n, i = 1, 2 \right\}.$$

We note that each of the exceptional lines \mathcal{L}_i and \mathcal{L}'_i meet 3 of the squares $h_v(I^2)$ at their corners. Moreover, a line \mathcal{L}' meets three squares of the form $h_\omega(h_v(I^2)) \subset h_\omega(I^2)$, $\omega \in \cup_n \Lambda^n$ if and only if \mathcal{L}' is exceptional.

Suppose \mathcal{L} is an affine line which is not exceptional. Then, we observe that \mathcal{L} meets at most 2 squares of the form $h_v(I^2)$, where $v \in \Lambda$. It follows by induction that \mathcal{L} meets at most 2^n squares of the form $h_\omega(I^2)$, for $n \in \mathbb{N}$ and $\omega \in \Lambda^n$. Indeed, if $\omega' \in \Lambda^{n-1}$ is the prefix of ω , then $h_{\omega'}^{-1}(\mathcal{L})$ is a non-exceptional affine line and hence meets at most 2 squares of the form $h_v(I^2)$, $v \in \Lambda$.

Finally, assume \mathcal{L} is an exceptional line of slope = 1, the case of -1 slope being identical. Let $\omega \in \Lambda^k$ be such that $\mathcal{L} = h_\omega(\mathcal{L}_i)$ for $i \in \{0, 1\}$. Writing $b_\omega = (b_{\omega_1}, b_{\omega_2})$, the line $h_\omega(\mathcal{L}_i)$ has the form:

$$h_\omega(\mathcal{L}_i) = \left\{ \mathbf{x} = (x_1, x_2) \in \mathbb{R}^2 : x_2 = x_1 - b_{\omega_1} + b_{\omega_2} + (-1)^i \cdot 3^{-(k+1)} \right\}. \quad (\text{A.2})$$

Since $b_{\omega_1}, b_{\omega_2} \in 3^{-k}\mathbb{Z}$, it follows that $b_\mathcal{L} = a/3^{k+1}$, for $a \in \mathbb{Z}$ coprime to 3. Moreover, suppose $\alpha \in \Lambda^n$, for $n \neq k$. Since $h_\alpha(\mathcal{L}_i)$ is parallel to \mathcal{L} , $h_\alpha(\mathcal{L}_i) = \mathcal{L}$ if and only if $b_\mathcal{L} = b_{h_\alpha(\mathcal{L}_i)}$. A calculation similar to that yielding (A.2) shows that $b_{h_\alpha(\mathcal{L}_i)} = m \cdot 3^{-n-1}$, where $m \in \mathbb{Z}$ is coprime to 3. Thus, it follows that

$$\mathcal{L} \notin \left\{ h_\alpha(\mathcal{L}_i) : \alpha \in \bigcup_{n \geq 0, n \neq k} \Lambda^n, i = 1, 2 \right\}. \quad (\text{A.3})$$

Let B_n denote the set of squares of the form $h_\alpha(I^2)$ which meet \mathcal{L} with $\alpha \in \Lambda^n$ and denote by q_n the cardinality of B_n . We show that (A.3) implies that $q_n \leq (3/2) \cdot 2^n$, which will conclude the proof. Indeed, suppose $n \neq k$ and let $h_\alpha(I^2) \in B_n$. Then, since $h_\alpha^{-1}(\mathcal{L}) \neq \mathcal{L}_0, \mathcal{L}_1$, it follows that $h_\alpha^{-1}(\mathcal{L})$ meets at most 2 squares of the form $h_v(I^2)$. Thus, $q_{n+1} \leq 2q_n$. Alternatively, if $n = k$, then $h_\alpha^{-1}(\mathcal{L})$ meets at most 3 such squares and, hence, $q_{k+1} \leq 3q_k$. By induction, we conclude that $q_n \leq (3/2) \cdot 2^n$ as desired. \square

A.2. Frostman exponents of projections and Theorem 1.9. Let the notation be as in the statement of the theorem. A key ingredient in the argument is a generalization of the sub-additive ergodic theorem for uniquely ergodic systems due to Furman in [Fur97]. It gives information about

the behavior of every orbit, as opposed to the almost everywhere statement of the subadditive ergodic theorem. Recall that if (X, μ, T) is a measure preserving system, a sequence of functions (ϕ_n) is a **sub-additive cocycle** over T if for all m, n and almost every $x \in X$:

$$\phi_{m+n}(x) \leq \phi_m(x) + \phi_n(T^m x).$$

Theorem A.2 (Theorem 1, [Fur97]). *Suppose (X, T, ν) is a uniquely ergodic probability measure preserving system, where X is a compact metric space and T is continuous. Let $\phi_n : X \rightarrow \mathbb{R}$ be a continuous sub-additive cocycle over T . Then, the following holds*

$$\limsup_{n \rightarrow \infty} \frac{\phi_n(x)}{n} \leq \inf_{n \in \mathbb{N}} \left\{ \int \frac{\phi_n}{n} d\nu \right\}, \quad (\text{A.4})$$

uniformly over all $x \in X$.

Given an affine line \mathcal{L} and $\theta \in [0, 2\pi)$, we write $\mathcal{L} \perp \theta$ whenever \mathcal{L} is orthogonal to any line of slope $\tan \theta$. Recall that $\mathcal{A}(d, \ell)$ denotes the collection of all affine subspaces of \mathbb{R}^d of dimension ℓ . For every $\theta \in [0, 2\pi)$ and $n \in \mathbb{N}$, $\varepsilon > 0$, we define

$$t(\theta, \varepsilon) := \sup_{\mathcal{L} \in \mathcal{A}(2,1): \mathcal{L} \perp \theta} \mu(\mathcal{L}^{(\varepsilon)}), \quad \tau(\theta, n) := t(\theta, \rho^n).$$

The first step in applying Theorem A.2 is the following continuity result.

Proposition A.3. *For every $\varepsilon > 0$, the function $\log t(\cdot, \varepsilon)$ is continuous on $[0, 2\pi)$.*

The key ingredient is the compactness of the support of μ , which is used in the proof of the following lemma.

Lemma A.4. *Suppose \mathcal{L}_1 is an affine line and suppose $0 \leq \delta \leq \pi/2$ is given. Let $K = \text{diam}(\mathcal{K})$. Then, there exists an affine line \mathcal{L}_2 which meets \mathcal{L}_1 at an angle δ and such that for every $\varepsilon > 0$,*

$$\mathcal{L}_1^{(\varepsilon)} \cap \mathcal{K} \subset \mathcal{L}_2^{(\varepsilon + A\delta)},$$

where $A = 4(K + \varepsilon)$.

Proof. If $\delta = 0$, we can take $\mathcal{L}_2 = \mathcal{L}_1$. Hence, we will assume $\delta > 0$. Let B be a ball of radius $K + \varepsilon$ containing \mathcal{K} . If $\mathcal{L}_1 \cap B = \emptyset$, then $\mathcal{L}_1^{(\varepsilon)} \cap \mathcal{K} = \emptyset$ and the statement follows trivially. Otherwise, let $x_0 \in \mathcal{L}_1 \cap B$. Let \mathcal{L}_2 be the line passing through x_0 at an angle δ with \mathcal{L}_1 .

For $i = 1, 2$, let n_i be the unit normal vectors to \mathcal{L}_i so that the angle between n_1 and n_2 is δ . Then, for $x \in \mathbb{R}^2$, $d(x, \mathcal{L}_i) = |\langle x - x_0, n_i \rangle|$. Let $x \in \mathcal{L}_1^{(\varepsilon)} \cap \mathcal{K}$. We then have

$$d(x, \mathcal{L}_2) = |\langle x - x_0, n_2 \rangle| \leq |\langle x - x_0, n_1 \rangle| + |\langle x - x_0, n_2 - n_1 \rangle| \leq \varepsilon + \|x - x_0\| \|n_2 - n_1\|.$$

Since $x, x_0 \in B$, we have $\|x - x_0\| \leq 2(K + \varepsilon)$. Moreover, using the law of sines, one verifies that

$$\|n_2 - n_1\| = \frac{\sin \delta}{\sin(\frac{\pi - \delta}{2})} \leq \frac{\pi \delta}{\pi - \delta} \leq 2\delta,$$

where we used the fact that $0 < \delta \leq \pi/2$. □

Proof of Proposition A.3. We may assume without loss of generality that the similarity dimension of \mathcal{F} is positive. Indeed, otherwise, \mathcal{F} consists of a single map, \mathcal{K} is a single point, and all the quantities in question are 0. We then observe that the irrationality of the rotation angle α implies that \mathcal{F} is irreducible and in particular, that $\alpha_1(\mu) > 0$ by Corollary A.10. Fix some $0 < \beta < \alpha_1(\mu)$.

It follows from the definition of $\alpha_1(\mu)$ that there exists $r_0 > 0$ so that for all $0 < r < r_0$ and all affine lines \mathcal{L} ,

$$\mu\left(\mathcal{L}^{(r)}\right) \leq r^\beta. \quad (\text{A.5})$$

Moreover, by Proposition 2.2, there exists a constant $C \geq 1$ such that

$$\frac{1}{C}r^s \leq \mu(B(x, r)) \leq Cr^s, \quad (\text{A.6})$$

for every $x \in \mathcal{K}$ and every $r > 0$, where $s = \dim_H(\mathcal{K})$. Fix $\varepsilon > 0$ and assume $0 < \eta < 1$ is given. Denote by $K = \text{diam}(\mathcal{K})$ and let $A = 4(K + \varepsilon)$. Let $0 < \delta \leq \pi/2$ be sufficiently small so that

$$\delta^{\beta/2} \leq \min \left\{ \frac{\varepsilon^s}{2(A+1)^\beta C}, \eta \right\}.$$

Suppose $\varphi, \theta \in [0, 2\pi)$ are two angles at distance at most δ in \mathbb{S}^1 . Let \mathcal{L}_1 be a line satisfying $\mathcal{L}_1 \perp \theta$. By Lemma A.4, there exists a line $\mathcal{L}_2 \perp \varphi$ such that

$$\mathcal{L}_1^{(\varepsilon)} \cap \mathcal{K} \subseteq \mathcal{L}_2^{(\varepsilon+A\delta)}.$$

Moreover, we can find lines $\mathcal{L}_3, \mathcal{L}_4$ parallel to \mathcal{L}_2 and satisfying $\cup_{j=3,4} \mathcal{L}_j \subset \mathcal{L}_2^{(\varepsilon+A\delta)} \setminus \mathcal{L}_2^{(\varepsilon)}$ and $\mathcal{L}_2^{(\varepsilon+A\delta)} \setminus \mathcal{L}_2^{(\varepsilon)} \subseteq \cup_{j=3,4} \mathcal{L}_j^{((A+1)\delta)}$. This, along with (A.5), imply

$$\mu\left(\mathcal{L}_1^{(\varepsilon)}\right) \leq \mu\left(\mathcal{L}_2^{(\varepsilon)}\right) + \mu\left(\mathcal{L}_3^{((A+1)\delta)}\right) + \mu\left(\mathcal{L}_4^{((A+1)\delta)}\right) \leq t(\varphi, \varepsilon) + 2((A+1)\delta)^\beta.$$

The next ingredient is to observe that $t(\varphi, \varepsilon) \geq \varepsilon^s/C$. This follows by taking $\mathcal{L}_0 \perp \varphi$ to be a line passing through a point in \mathcal{K} . Then, $\mathcal{L}_0^{(\varepsilon)}$ contains a ball of radius ε centered in \mathcal{K} . The claim thus follows from the estimate in (A.6). Our choice of δ hence implies that

$$\mu\left(\mathcal{L}_1^{(\varepsilon)}\right) \leq (1 + \delta^{\beta/2})t(\varphi, \varepsilon).$$

Since \mathcal{L}_1 was arbitrary, we see that $\log t(\theta, \varepsilon) \leq \eta + \log t(\varphi, \varepsilon)$. Since δ is independent of θ and φ , one can run the above argument with the roles of φ and θ reversed, to get the reverse inequality and conclude the proof. \square

The next ingredient in the proof of Theorem 1.9 is establishing the following cocycle property of the functions τ .

Proposition A.5. *There exists a constant $D \geq 1$, such that for every $\theta \in [0, 2\pi)$ and $m, n \in \mathbb{N}$,*

$$\tau(\theta, m+n) \leq D\tau(\theta, m)\tau(R_{-\alpha}^m(\theta), n).$$

We will need the following doubling property of the functions $t(\cdot)$.

Lemma A.6. *For every $A \geq 1$, there exists $D \geq 1$ such that for every θ and $\varepsilon > 0$,*

$$t(\theta, A\varepsilon) \leq Dt(\theta, \varepsilon).$$

Proof. Let D denote the cardinality of a finite cover of a ball of radius 1 in \mathbb{R} by balls of radius $1/A$. By applying scaling and translation, it follows that for every $\varepsilon > 0$, any ball of radius $A\varepsilon$ in \mathbb{R} can be covered by at most D balls of radius ε . Hence, for every Borel measure ν on \mathbb{R} , one has

$$\nu(B(x, A\varepsilon)) \leq D \sup_{y \in \mathbb{R}} \nu(B(y, \varepsilon)),$$

for all $x \in \mathbb{R}$ and all $\varepsilon > 0$. □

Proof of Proposition A.5. Let $\mathcal{L} \perp \theta$ and $m, n \in \mathbb{N}$ be given and let $P = \Lambda^m$. For every $\xi \in P$ and $\varepsilon > 0$, observe that $h_\xi^{-1}(\mathcal{L}^{(\varepsilon)}) = (h_\xi^{-1}(\mathcal{L}))^{(\varepsilon/\rho^m)}$. Moreover, one has $h_\xi^{-1}(\mathcal{L}) \perp R_{-\alpha}^m(\theta)$. Then, by equation (2.10) and Lemma 2.5, it follows that

$$\begin{aligned} \mu\left(\mathcal{L}^{(\rho^{m+n})}\right) &= \sum_{\xi \in P} \mu\left(\mathcal{L}^{(\rho^{m+n})} \cap \mathcal{K}_\xi\right) = \sum_{\xi \in P} \rho^{ms} \mu\left(h_\xi^{-1}\left(\mathcal{L}^{(\rho^{m+n})}\right)\right) \\ &= \sum_{\xi \in P} \rho^{ms} \mu\left(h_\xi^{-1}(\mathcal{L})^{(\rho^n)}\right) \leq \tau(\theta - m\alpha, n) \sum_{\substack{\xi \in P \\ \mathcal{K}_\xi \cap \mathcal{L}^{(\rho^{m+n})} \neq \emptyset}} \rho^{ms}. \end{aligned}$$

The other ingredient is the observation that for every $\xi \in P$ satisfying $\mathcal{K}_\xi \cap \mathcal{L}^{(\rho^{m+n})} \neq \emptyset$, we have that $\mathcal{K}_\xi \subset \mathcal{L}^{((1+K)\rho^m)}$, where $K = \text{diam}(\mathcal{K})$. Indeed, this follows from the fact that the diameter of \mathcal{K}_ξ is $K\rho^m$. Moreover, we have $\mu(\mathcal{K}_\xi) = \rho^{ms}$ for every ξ . Hence, Proposition 2.3 on the null overlaps between the distinct \mathcal{K}_ξ implies that

$$\sum_{\substack{\xi \in P \\ \mathcal{K}_\xi \cap \mathcal{L}^{(\rho^{m+n})} \neq \emptyset}} \rho^{ms} \leq \mu\left(\bigcup_{\substack{\xi \in P \\ \mathcal{K}_\xi \cap \mathcal{L}^{(\rho^{m+n})} \neq \emptyset}} \mathcal{K}_\xi\right) \leq \mu\left(\mathcal{L}^{((1+K)\rho^m)}\right) \leq t(\theta, (1+K)\rho^m).$$

Finally, we apply the doubling property from Lemma A.6 with $A = K + 1$ to get that $t(\theta, (1+K)\rho^m) \leq D\tau(\theta, m)$, for a constant $D \geq 1$ depending only on K . □

We are now ready for the proof of Theorem 1.9.

Proof of Theorem 1.9. Let $\phi_n(\theta) := \log \tau(\theta, n) + \log D$, where D is the constant in the conclusion of Proposition A.5. One can then verify that for every $\theta \in [0, 2\pi)$:

$$\dim_\infty(\pi_\theta \mu) = \liminf_{n \rightarrow \infty} \frac{\log \tau(\theta, n)}{n \log \rho} = \liminf_{n \rightarrow \infty} \frac{\log \phi_n(\theta)}{n \log \rho}.$$

Let ν be the Lebesgue probability measure on \mathbb{S}^1 . Then, ϕ_n is a sub-additive cocycle over the transformation $T = R_{-\alpha}$ of $X = \mathbb{S}^1$. In particular, Kingman's sub-additive ergodic theorem implies that

$$\dim_\infty(\pi_\theta \mu) = \phi_* := \sup_{n \geq 1} \int \frac{\phi_n}{n \log \rho} d\nu = \lim_{n \rightarrow \infty} \int \frac{\phi_n}{n \log \rho} d\nu, \quad \text{for } \nu\text{-almost every } \theta.$$

Moreover, the cocycle (ϕ_n) is continuous by Proposition A.4. Hence, Theorem A.2 implies that

$$\dim_\infty(\pi_\theta \mu) \geq \phi_*, \quad \text{for every } \theta. \tag{A.7}$$

It remains to show that $\alpha_1(\mu) = \phi_*$. That $\alpha_1(\mu) \leq \phi_*$ follows by definition. For the reverse inequality, we use the uniformity in convergence provided by Theorem A.2. Fix $0 < \varepsilon < 1$. For every $n \in \mathbb{N}$, let θ_n be such that

$$\log \sup_{\theta} \tau(\theta, n) = \sup_{\theta} \log \tau(\theta, n) \leq (1 - \varepsilon) \log \tau(\theta_n, n). \tag{A.8}$$

Here we used the fact that $\log \tau(\cdot) \leq 0$ since μ is a probability measure. From uniform convergence of the \liminf in θ in (A.7), we can find $n_0 \in \mathbb{N}$ so that for all $n \geq n_0$ and for all θ ,

$$\frac{\log \tau(\theta, n)}{n \log \rho} \geq \phi_* - \varepsilon.$$

Combining this with (A.8) and the fact that $\log \rho < 0$, we get

$$\frac{\log \sup_{\theta \in \mathbb{S}^1} \tau(\theta, n)}{n \log \rho} \geq (1 - \varepsilon)(\phi_* - \varepsilon),$$

for all $n \geq n_0$. Since ε was arbitrary, it follows that $\alpha_1(\mu) \geq \phi_*$ as desired. \square

A.3. Existence of the limit in Definition 1.3. Throughout this section, we fix a finite set Λ and an IFS $\mathcal{F} = \{h_i : i \in \Lambda\}$ on \mathbb{R}^d satisfying the open set condition. We denote by \mathcal{K} the limit set of \mathcal{F} , s the similarity dimension of \mathcal{F} , and μ the restriction of H^s to \mathcal{K} .

Proposition A.7. *Suppose μ is as above. Then, for every $1 \leq \ell \leq d$, the limit in the definition of $\alpha_\ell(\mu)$ exists.*

For every $1 \leq \ell \leq d$, $\omega \in \Lambda^{\mathbb{N}}$ and $n \in \mathbb{N}$, $\varepsilon > 0$, we define

$$t_\ell(\varepsilon) = \sup_{\mathcal{L} \in \mathcal{A}(d, \ell)} \mu(\mathcal{L}^{(\varepsilon)}), \quad \tau_\ell(\omega, n) = t_\ell(\rho(\omega, n)),$$

where $\rho(\omega, n)$ is defined by (2.6). The functions t_ℓ satisfy the following doubling property.

Lemma A.8. *For every $A \geq 1$, there exists a constant $D \geq 1$, depending only on A and ℓ , such that for every $\varepsilon > 0$,*

$$t_\ell(A\varepsilon) \leq Dt_\ell(\varepsilon).$$

Proof. The proof is completely analogous to Lemma A.6. \square

The key step in proving Proposition A.7 is to show that τ_ℓ is submultiplicative. This was essentially shown in [KLW04]. We include a proof for completeness.

Lemma A.9. *There exists a constant $C_\ell \geq 1$ so that*

$$\tau_\ell(\omega, m+n) \leq C_\ell \tau_\ell(\omega, m) \tau_\ell(\sigma^m \omega, n)$$

for all $\omega \in \Lambda^{\mathbb{N}}$ and all $m, n \in \mathbb{N}$.

Before proving this lemma, we record the following useful corollary.

Corollary A.10 (Lemma 8.2, [KLW04]). *Suppose \mathcal{F} is an irreducible IFS on \mathbb{R}^d satisfying the open set condition. Let \mathcal{K} be its limit set and let $s = \dim_H(\mathcal{K})$. Then, for each $1 \leq \ell \leq d$, $\alpha_\ell(\mu) > 0$, where μ is the restriction of H^s to \mathcal{K} .*

Proof. Since Lemma 8.2 of [KLW04] is stated in a different form, we provide here a proof for completeness. It is shown over the course of the proof of [KLW04, Theorem 2.3] that the irreducibility of \mathcal{F} implies that $\mu(\mathcal{L}) = 0$ for all proper affine subspaces $\mathcal{L} \subset \mathbb{R}^d$. We claim that this implies that

$$\lim_{\varepsilon \rightarrow 0} t_\ell(\varepsilon) = 0.$$

Indeed, suppose not. Then, there exist $\delta > 0$ and sequences $\varepsilon_i \rightarrow 0$ and $\mathcal{L}_i \in \mathcal{A}(d, \ell)$ such that $\mathcal{L}_i \cap \mathcal{K} \neq \emptyset$ and

$$\mu(\mathcal{L}_i^{(\varepsilon_i)}) > \delta.$$

Since \mathcal{K} is compact, the subset of $\mathcal{A}(d, \ell)$ consisting of affine subspaces that meet \mathcal{K} is compact. In particular, by passing to a subsequence if necessary, we may assume that $\mathcal{L}_i \rightarrow \mathcal{L}$ for some $\mathcal{L} \in \mathcal{A}(d, \ell)$ such that $\mathcal{L} \cap \mathcal{K} \neq \emptyset$. Hence, it follows by the dominated convergence theorem that $\mu(\mathcal{L}) \geq \delta > 0$, which is a contradiction.

Let $C_\ell \geq 1$ be the constant in Lemma A.9. Let $\varepsilon_0 > 0$ be sufficiently small so that

$$t_\ell(\varepsilon) < 1/2C_\ell,$$

for all $0 < \varepsilon \leq \varepsilon_0$. Let ρ_{\max} and ρ_{\min} denote the largest and smallest contraction ratios of the maps in \mathcal{F} respectively. Then, $0 < \rho_{\min} \leq \rho_{\max} < 1$. In particular, we can find $k \in \mathbb{N}$ sufficiently large so that $\rho_{\max}^k < \varepsilon_0$. Fix one such k . Then, for all $\alpha \in \Lambda^{\mathbb{N}}$, we have

$$\rho(\alpha, k) \leq \rho_{\max}^k < \varepsilon_0. \quad (\text{A.9})$$

Fix some $\omega \in \Lambda^{\mathbb{N}}$. It follows from Lemma A.9 and (A.9) that for all $n \in \mathbb{N}$,

$$\tau_\ell(\omega, nk) \leq C_\ell^n \prod_{i=0}^{n-1} \tau_\ell(\sigma^i \omega, k) \leq C_\ell^n t_\ell(\varepsilon_0)^n < 2^{-n}.$$

Moreover, note that $t_\ell(\varepsilon) \leq 1$ for all $\varepsilon > 0$ since μ is a probability measure. In particular, $\log t_\ell(\varepsilon)/\log \varepsilon \geq 0$ for all $\varepsilon > 0$. Combined with that fact that $\rho(\omega, nk) \geq \rho_{\min}^{nk}$, this implies that

$$\frac{\log \tau_\ell(\omega, nk)}{\log \rho(\omega, nk)} \geq \frac{\log \tau_\ell(\omega, nk)}{nk \log \rho_{\min}} \geq \frac{-\log 2}{k \log \rho_{\min}} > 0,$$

for all $n \in \mathbb{N}$. This proves that $\liminf_{n \rightarrow \infty} \frac{\log \tau_\ell(\omega, nk)}{\log \rho(\omega, nk)} > 0$. To conclude the proof, suppose that $0 < \varepsilon < \rho(\omega, k)$ is given and let $n(\varepsilon) \in \mathbb{N}$ be such that

$$\rho(\omega, (n(\varepsilon) + 1)k) \leq \varepsilon < \rho(\omega, n(\varepsilon)k).$$

In particular, we have $t_\ell(\varepsilon) \leq \tau_\ell(\omega, (n(\varepsilon)k))$. Moreover, the choice of $n(\varepsilon)$ implies

$$\varepsilon \geq \rho(\omega, n(\varepsilon)k) \rho_{\min}^k.$$

Thus, we obtain the following estimate:

$$\frac{\log t_\ell(\varepsilon)}{\log \varepsilon} \geq \frac{\log \tau_\ell(\omega, n(\varepsilon)k)}{\log \varepsilon} \geq \frac{\log \tau_\ell(\omega, n(\varepsilon)k)}{\log \rho(\omega, n(\varepsilon)k) + k \log \rho_{\min}}.$$

Note that $n(\varepsilon) \rightarrow \infty$ as $\varepsilon \rightarrow 0$. In particular, $\log \rho(\omega, n(\varepsilon)k) \rightarrow -\infty$. This shows that

$$\alpha_\ell(\mu) = \liminf_{\varepsilon \rightarrow 0} \frac{\log t_\ell(\varepsilon)}{\log \varepsilon} \geq \liminf_{n \rightarrow \infty} \frac{\log \tau_\ell(\omega, nk)}{\log \rho(\omega, nk)} > 0,$$

as desired. □

Following [KLW04], we say a finite set $P \subset \cup_{k \geq 1} \Lambda^k$ is a **complete prefix set** if every $\omega \in \Lambda^{\mathbb{N}}$, there is a unique word $\alpha \in P$ which occurs as a prefix for ω . It is easy to check that the set

$$P(\varepsilon) := \left\{ \alpha \in \cup_{k \geq 1} \Lambda^k : \varepsilon \rho_{\min} \leq \rho_\alpha \leq \varepsilon \right\} \quad (\text{A.10})$$

forms a complete prefix set, where $\rho_{\min} = \min \{\rho_i : i \in \Lambda\}$, and ρ_α is as in (2.5). The following lemma allows us to handle the case where the contraction ratios are not all the same.

Lemma A.11. *Suppose P is a complete prefix set. Then, for every continuous function h on \mathbb{R}^d ,*

$$\int h d\mu = \sum_{\alpha \in P} \int_{\mathcal{K}_\alpha} h d\mu.$$

Proof. Proposition 2.3 implies that $\{\mathcal{K}_\alpha : \alpha \in P\}$ forms a measurable partition of the support of μ with null overlaps. \square

Proof of Lemma A.9. Suppose $\omega \in \Lambda^\mathbb{N}$ and $m, n \in \mathbb{N}$ are given. Let $P = P(\rho(\omega, m))$ be the set defined in (A.10) with $\varepsilon = \rho(\omega, m)$. Let \mathcal{L} be an affine subspace of dimension $d - \ell$. For simplicity, we write $\mathcal{L}^{(\omega, k)} := \mathcal{L}^{(\rho(\omega, k))}$ for every $k \in \mathbb{N}$. Note that for each $\xi \in P$, $\rho(\omega, m+n)/\rho_\xi \leq \rho_{\min}^{-1} \rho(\sigma^m(\omega), n)$. By Lemma A.11 and a similar argument to the proof of Proposition A.5, one obtains

$$\begin{aligned} \mu\left(\mathcal{L}^{(\omega, m+n)}\right) &= \sum_{\xi \in P} \mu\left(\mathcal{L}^{(\omega, m+n)} \cap \mathcal{K}_\xi\right) = \sum_{\xi \in P} \rho_\xi^s \left(h_\xi^{-1}\left(\mathcal{L}^{(\omega, m+n)}\right)\right) \\ &= \sum_{\xi \in P} \rho_\xi^s \mu\left(h_\xi^{-1}\left(\mathcal{L}\right)^{(\rho(\omega, m+n)/\rho_\xi)}\right) \leq t_\ell(\rho_{\min}^{-1} \rho(\sigma^m(\omega), n)) \sum_{\substack{\xi \in P \\ \mathcal{K}_\xi \cap \mathcal{L}^{(\omega, m+n)} \neq \emptyset}} \rho_\xi^s. \end{aligned} \quad (\text{A.11})$$

Proposition 2.3 implies that $\mu(\mathcal{K}_{\xi_1} \cap \mathcal{K}_{\xi_2}) = 0$ for every $\xi_1, \xi_2 \in P$ with $\xi_1 \neq \xi_2$. Moreover, by definition of P , for every $\xi \in P$ satisfying $\mathcal{K}_\xi \cap \mathcal{L}^{(\omega, m+n)} \neq \emptyset$, we have that $\mathcal{K}_\xi \subset \mathcal{L}^{((1+K)\rho(\omega, m))}$, where $K = \text{diam}(\mathcal{K})$. Hence, arguing as in the proof of Proposition A.5, we obtain

$$\sum_{\substack{\xi \in P \\ \mathcal{K}_\xi \cap \mathcal{L}^{(\omega, m+n)} \neq \emptyset}} \rho_\xi^s \leq t_\ell((1+K)\rho(\omega, m)). \quad (\text{A.12})$$

In view of the doubling property of t_ℓ provided by Lemma A.8, the conclusion follows by combining (A.11) and (A.12). \square

We now deduce Proposition A.7 from Lemma A.9.

Proof of Proposition A.7. Let C_ℓ be the constant given by Lemma A.9. Consider the function $\varphi_\ell : \Lambda^\mathbb{N} \times \mathbb{N} \rightarrow \mathbb{R}_+$ defined by $\varphi_\ell = \log C_\ell + \log \tau_\ell$. Lemma A.9 implies that φ_ℓ is a subadditive cocycle over $\sigma : \Lambda^\mathbb{N} \rightarrow \Lambda^\mathbb{N}$.

Consider the constant sequence $\omega = (i)_k \in \Lambda^\mathbb{N}$, for some fixed $i \in \Lambda$. Then, the sequence $a_n = \varphi_\ell(\omega, n)$ is a subadditive sequence. This follows from the fact that $\sigma\omega = \omega$. Thus, by Fekete's lemma, we obtain the following equalities $\lambda_\ell = \lim_{n \rightarrow \infty} \frac{a_n}{n} = \inf_{n \geq 1} \frac{a_n}{n}$.

We claim that $\alpha_\ell(\mu) = \lambda_\ell / \log \rho_i$ and in particular that the limit defining $\alpha_\ell(\mu)$ exists. To see this, suppose ε_k is a sequence tending to 0. For each k , let $n_k \in \mathbb{N}$ be such that $\rho_i^{n_k+1} < \varepsilon_k \leq \rho_i^{n_k}$. Since $0 < \rho_i < 1$, it follows that

$$\frac{\log \sup_{\mathcal{L}: \dim \mathcal{L} = d - \ell} \mu\left(\mathcal{L}^{(\rho_i^{n_k})}\right)}{n_k \log \rho_i} \leq \frac{\log \sup_{\mathcal{L}: \dim \mathcal{L} = d - \ell} \mu\left(\mathcal{L}^{(\varepsilon_k)}\right)}{\log \varepsilon_k} \frac{n_k + 1}{n_k}.$$

In particular, we see that

$$\liminf_{\varepsilon \rightarrow 0} \frac{\log \sup_{\mathcal{L}: \dim \mathcal{L} = d - \ell} \mu\left(\mathcal{L}^{(\varepsilon)}\right)}{\log \varepsilon} \geq \lambda_\ell / \log \rho_i.$$

The opposite inequality involving the lim sup follows analogously. □

REFERENCES

- [AAE⁺17] H. Al-Saqban, P. Apisa, A. Erchenko, O. Khalil, S. Mirzadeh, and C. Uyanik, *Exceptional directions for the Teichmüller geodesic flow and Hausdorff dimension*, Journal of the Eur. Math. Soc., to appear. (2017), 1–48.
- [AM18] Paul Apisa and Howard Masur, *Divergent on average directions of Teichmüller geodesic flow*, arXiv e-prints (2018), arXiv:1803.00093.
- [Ath06] Jayadev S. Athreya, *Quantitative recurrence and large deviations for Teichmüller geodesic flow*, Geom. Dedicata **119** (2006), 121–140. MR 2247652
- [BCC18] Yann Bugeaud, Yitwah Cheung, and Nicolas Chevallier, *Hausdorff dimension and uniform exponents in dimension two*, Mathematical Proceedings of the Cambridge Philosophical Society (2018), 1–36.
- [BFS12] Balázs Bárány, Andrew Ferguson, and Károly Simon, *Slicing the Sierpiński gasket*, Nonlinearity **25** (2012), no. 6, 1753–1770. MR 2929601
- [BGSV18] Victor Beresnevich, Anish Ghosh, David Simmons, and Sanju Velani, *Diophantine approximation in Kleinian groups: singular, extremal, and bad limit points*, J. Lond. Math. Soc. (2) **98** (2018), no. 2, 306–328. MR 3873110
- [BQ11] Yves Benoist and Jean-Francois Quint, *Random walks on finite volume homogeneous spaces*, Inventiones mathematicae **187** (2011), no. 1, 37–59.
- [BR14] Balázs Bárány and Michał Rams, *Dimension of slices of Sierpiński-like carpets*, J. Fractal Geom. **1** (2014), no. 3, 273–294. MR 3276836
- [CC16] Yitwah Cheung and Nicolas Chevallier, *Hausdorff dimension of singular vectors*, Duke Math. J. **165** (2016), no. 12, 2273–2329. MR 3544282
- [Che11] Yitwah Cheung, *Hausdorff dimension of the set of singular pairs*, Ann. of Math. (2) **173** (2011), no. 1, 127–167. MR 2753601
- [Dan85] S. G. Dani, *Divergent trajectories of flows on homogeneous spaces and Diophantine approximation*, J. Reine Angew. Math. **359** (1985), 55–89. MR 794799
- [DFSU18] Tushar Das, Lior Fishman, David Simmons, and Mariusz Urbański, *Extremality and dynamically defined measures, part I: Diophantine properties of quasi-decaying measures*, Selecta Math. (N.S.) **24** (2018), no. 3, 2165–2206. MR 3816502
- [DFSU19] Tushar Das, Lior Fishman, David Simmons, and Mariusz Urbański, *A variational principle in the parametric geometry of numbers*, arXiv e-prints (2019), arXiv:1901.06602.
- [DS70a] H. Davenport and W. M. Schmidt, *Dirichlet’s theorem on diophantine approximation. II*, Acta Arith. **16** (1969/1970), 413–424. MR 0279040
- [DS70b] H. Davenport and Wolfgang M. Schmidt, *Dirichlet’s theorem on diophantine approximation*, Symposia Mathematica, Vol. IV (INDAM, Rome, 1968/69), Academic Press, London, 1970, pp. 113–132. MR 0272722
- [EFS11] Manfred Einsiedler, Lior Fishman, and Uri Shapira, *Diophantine approximations on fractals*, Geom. Funct. Anal. **21** (2011), no. 1, 14–35. MR 2773102
- [EM01] Alex Eskin and Howard Masur, *Asymptotic formulas on flat surfaces*, Ergodic Theory Dynam. Systems **21** (2001), no. 2, 443–478. MR 1827113
- [EM04] Alex Eskin and Gregory Margulis, *Recurrence properties of random walks on finite volume homogeneous manifolds*, Random Walks and Geometry: Proceedings of a Workshop at the Erwin Schrödinger Institute, Vienna, June 18–July 13, 2001, Walter de Gruyter, 2004, p. 431.
- [EMM98] Alex Eskin, Gregory Margulis, and Shahar Mozes, *Upper bounds and asymptotics in a quantitative version of the oppenheim conjecture*, Annals of Mathematics **147** (1998), no. 1, 93–141.
- [EMM15] Alex Eskin, Maryam Mirzakhani, and Amir Mohammadi, *Isolation, equidistribution, and orbit closures for the $SL(2, \mathbb{R})$ action on moduli space*, Ann. of Math. (2) **182** (2015), no. 2, 673–721. MR 3418528
- [Fis09] Lior Fishman, *Schmidt’s game on fractals*, Israel J. Math. **171** (2009), 77–92. MR 2520102
- [FM14] Giovanni Forni and Carlos Matheus, *Introduction to Teichmüller theory and its applications to dynamics of interval exchange transformations, flows on surfaces and billiards*, J. Mod. Dyn. **8** (2014), no. 3–4, 271–436. MR 3345837
- [Fur97] Alex Furman, *On the multiplicative ergodic theorem for uniquely ergodic systems*, Ann. Inst. H. Poincaré Probab. Statist. **33** (1997), no. 6, 797–815. MR 1484541
- [GSSY16] Daniel Galicer, Santiago Saglietti, Pablo Shmerkin, and Alexia Yavicoli, *L^q dimensions and projections of random measures*, Nonlinearity **29** (2016), no. 9, 2609–2640. MR 3544801

- [Hoc14] Michael Hochman, *On self-similar sets with overlaps and inverse theorems for entropy*, Ann. of Math. (2) **180** (2014), no. 2, 773–822. MR 3224722
- [HS12] Michael Hochman and Pablo Shmerkin, *Local entropy averages and projections of fractal measures*, Ann. of Math. (2) **175** (2012), no. 3, 1001–1059. MR 2912701
- [Hut81] John E. Hutchinson, *Fractals and self-similarity*, Indiana Univ. Math. J. **30** (1981), no. 5, 713–747. MR 625600
- [Kha18] Osama Khalil, *Bounded and Divergent Trajectories And Expanding Curves on Homogeneous Spaces*, arXiv e-prints (2018), arXiv:1806.06832.
- [KKLM17] S. Kadyrov, D. Kleinbock, E. Lindenstrauss, and G. A. Margulis, *Singular systems of linear forms and non-escape of mass in the space of lattices*, J. Anal. Math. **133** (2017), 253–277. MR 3736492
- [KLW04] Dmitry Kleinbock, Elon Lindenstrauss, and Barak Weiss, *On fractal measures and Diophantine approximation*, Selecta Math. (N.S.) **10** (2004), no. 4, 479–523. MR 2134453
- [KM98] D. Y. Kleinbock and G. A. Margulis, *Flows on homogeneous spaces and diophantine approximation on manifolds*, The Annals of Mathematics **148** (1998), no. 1, 339.
- [KMS86] Steven Kerckhoff, Howard Masur, and John Smillie, *Ergodicity of billiard flows and quadratic differentials*, Ann. of Math. (2) **124** (1986), no. 2, 293–311. MR 855297
- [KW05a] Dmitry Kleinbock and Barak Weiss, *Badly approximable vectors on fractals*, Israel J. Math. **149** (2005), 137–170, Probability in mathematics. MR 2191212
- [KW05b] ———, *Friendly measures, homogeneous flows and singular vectors*, Algebraic and topological dynamics, Contemp. Math., vol. 385, Amer. Math. Soc., Providence, RI, 2005, pp. 281–292. MR 2180240
- [KW08] ———, *Dirichlet’s theorem on Diophantine approximation and homogeneous flows*, J. Mod. Dyn. **2** (2008), no. 1, 43–62. MR 2366229
- [Mas82] Howard Masur, *Interval exchange transformations and measured foliations*, Ann. of Math. (2) **115** (1982), no. 1, 169–200. MR 644018
- [Mas92] ———, *Hausdorff dimension of the set of nonergodic foliations of a quadratic differential*, Duke Math. J. **66** (1992), no. 3, 387–442. MR 1167101
- [MW14] Yair Minsky and Barak Weiss, *Cohomology classes represented by measured foliations, and Mahler’s question for interval exchanges*, Ann. Sci. Éc. Norm. Supér. (4) **47** (2014), no. 2, 245–284. MR 3215923
- [NPS12] Fedor Nazarov, Yuval Peres, and Pablo Shmerkin, *Convolutions of Cantor measures without resonance*, Israel J. Math. **187** (2012), 93–116. MR 2891700
- [PS09] Yuval Peres and Pablo Shmerkin, *Resonance between Cantor sets*, Ergodic Theory Dynam. Systems **29** (2009), no. 1, 201–221. MR 2470633
- [Shm15] Pablo Shmerkin, *Projections of self-similar and related fractals: a survey of recent developments*, Fractal geometry and stochastics V, Progr. Probab., vol. 70, Birkhäuser/Springer, Cham, 2015, pp. 53–74. MR 3558150
- [Shm16] Pablo Shmerkin, *On Furstenberg’s intersection conjecture, self-similar measures, and the L^q -norms of convolutions*, Annals of Mathematics (to appear). (2016), arXiv:1609.07802.
- [SW16] David Simmons and Barak Weiss, *Random walks on homogeneous spaces and diophantine approximation on fractals*, arXiv e-prints (2016), arXiv:1611.05899.
- [Vee99] William A. Veech, *Measures supported on the set of uniquely ergodic directions of an arbitrary holomorphic 1-form*, Ergodic Theory Dynam. Systems **19** (1999), no. 4, 1093–1109. MR 1709433
- [Wu16] Meng Wu, *A proof of Furstenberg’s conjecture on the intersections of $\times p$ and $\times q$ -invariant sets*, Annals of Mathematics (to appear) (2016), arXiv:1609.08053.

DEPARTMENT OF MATHEMATICS, OHIO STATE UNIVERSITY, COLUMBUS, OH
E-mail address: khalil.37@osu.edu