

# FOURIER DECAY FROM $L^2$ -FLATTENING

SIMON BAKER, OSAMA KHALIL, AND TUOMAS SAHLSTEN

**ABSTRACT.** We develop a unified approach for establishing rates of decay for the Fourier transform of a wide class of dynamically defined measures. Among the key features of the method is the systematic use of the  $L^2$ -flattening theorem obtained in [49], coupled with non-concentration estimates for the derivatives of the underlying dynamical system. This method yields polylogarithmic Fourier decay for Diophantine self-similar measures, and polynomial decay for Patterson-Sullivan measures of convex cocompact hyperbolic manifolds, Gibbs measures associated to non-integrable  $C^2$  conformal systems, as well as stationary measures for carpet-like non-conformal iterated function systems. Applications include essential spectral gaps on convex cocompact hyperbolic manifolds, fractal uncertainty principles, and equidistribution properties of typical vectors in fractal sets.

## 1. INTRODUCTION

**1.1. Background and summary of main results.** The Fourier transform of a Borel probability measure  $\mu$  on  $\mathbb{R}^d$  is defined as follows:

$$\widehat{\mu}(\xi) := \int_{\mathbb{R}^d} e^{2\pi i \langle \xi, x \rangle} d\mu(x), \quad \xi \in \mathbb{R}^d. \quad (1.1)$$

Rates of decay of  $|\widehat{\mu}(\xi)|$  as  $\|\xi\| \rightarrow \infty$  for  $\mu$  arising from dynamical systems have been extensively studied in recent years (see survey [76] for history and recent developments). Beyond its intrinsic interest, this question has found many applications in other areas of mathematics; e.g. essential spectral gaps on hyperbolic manifolds [28, 14, 53], the uniqueness problem [56], quantum chaos and fractal uncertainty principles [26], normality to arbitrary integer bases [23], and geometric measure theory [79, 59] to name a few, see Section 1.7 for further discussion. Moreover, this problem has motivated the development of many new methods. These methods draw on a wide variety of tools ranging from spectral gaps of the underlying dynamics [5, 12], to renewal theory [52], sum-product phenomena [14, 53, 51], large deviation estimates for Fourier transforms [61, 2, 11], and many more; cf. [4, 3, 42, 77, 56, 55, 91, 18] for a non-exhaustive list.

In this article, we systematically develop a unified approach to obtain rates of Fourier decay for a wide class of dynamically defined measures. Our strategy can be summarized as follows:

### Step 1. Averaging.

Use the dynamics (or the multiscale/convolution structure of  $\mu$ ) to express the Fourier transform of  $\mu$  at frequency  $\xi$  as an *average* over the Fourier transforms of measures  $\mu_x$ ,  $x \in X$ , evaluated at a range of frequencies determined by  $\xi$  and the underlying dynamics. The family  $\{\mu_x\}$  is typically given by scaled images of  $\mu$  under the dynamics or rescaled restrictions of  $\mu$  to pieces of its support.

### Step 2. Flattening.

Find a mechanism to show that for each  $x \in X$ , the Fourier transform of  $\mu_x$  has the desired rate of decay for a large set of frequencies. Here it is crucial that the size of the set of exceptional frequencies can be bounded independently of  $x$ .

### Step 3. Separation.

Show (through a non-linearity or Diophantine assumption on the dynamics) that the frequencies appearing in the average in **Step 1** are sufficiently well-distributed that they have a suitably small intersection with the exceptional set of frequencies arising in **Step 2**.

A key to this strategy is provided by recent developments towards **Step 2**, where verifiable criteria have been shown to imply polynomial Fourier decay<sup>1</sup> outside a sparse set of frequencies; cf. Theorem 1.3 below. In particular, such criteria are known to hold for a very wide class of dynamically defined measures, thus distilling the difficulty in implementing the above strategy to achieving the separation in **Step 3**. In this article, we develop techniques for overcoming this difficulty, thus yielding the following broad range of results.

- (1) Polylogarithmic Fourier decay for **self-similar measures** satisfying Diophantine conditions either on their contraction ratios or their rotation parts: Theorems 1.6 and 1.9.
- (2) Polynomial Fourier decay for **Patterson-Sullivan measures** of convex cocompact hyperbolic manifolds: Theorem 1.11.
- (3) Polynomial Fourier decay for **Gibbs measures for conformal IFSs** satisfying a spectral gap assumption: Theorem 1.14.
- (4) Polynomial Fourier decay for the stationary measures of a class of **carpet like non-conformal IFSs** exhibiting non-linearity in each principal direction: Theorem 1.21.

We refer the reader to Remarks 1.10 and 1.12 for a discussion of the rates of decay in the case of self-similar and Patterson-Sullivan measures. Several special cases of our results were known before in low dimensions by a variety of methods. These are recalled below in detail. We only remark that our argument provides a new streamlined proof in all these cases in addition to extending them to arbitrary dimensions, which requires handling the significant added difficulty posed by the presence of subspaces of intermediate dimension.

The above results also have the following applications:

- (1) Normality and **equidistribution of vectors in fractal sets**: let  $\mu$  be one of the measures in Items 1-4 above viewed as a measure on  $\mathbb{R}^d$ . Let  $A$  be any expanding integer matrix on  $\mathbb{R}^d$ . Then,  $\mu$ -almost every  $x$  is  $A$ -normal, i.e., the orbit  $(A^n x)_{n \geq 1}$  is equidistributed mod 1 for  $\mu$ -almost every  $x$ : Corollary 1.23.
- (2) **Essential spectral gaps** and fractal uncertainty principles: let  $\Gamma$  be a Zariski-dense convex cocompact group of isometries of real hyperbolic space  $\mathbb{H}^{d+1}$  with limit set  $\Lambda_\Gamma$ , critical exponent  $\delta_\Gamma \leq d/2$ , and quotient manifold  $M = \mathbb{H}^{d+1}/\Gamma$ . Then,  $\Lambda_\Gamma$  satisfies a generalized Fractal Uncertainty Principle with exponent  $\beta = d/2 - \delta_\Gamma + \varepsilon$ ,  $\varepsilon > 0$  and, hence, the resolvent of the Laplace-Beltrami operator  $\Delta_M$  of  $M$  admits an essential spectral gap of size  $\beta$ . The improvement  $\varepsilon$  depends only on the non-concentration parameters of the PS measure of  $\Gamma$ , but not on its doubling constant: Corollary 1.24.

Details and context of the above applications are discussed in detail in Section 1.7 below.

**1.2. Discussion of the method.** As noted above, a very general result towards **Step 2** as obtained in [49, Corollary 11.5] using methods from additive combinatorics. Before recalling this result, it will be convenient to make the following definition.

**Definition 1.1.** We say that a Borel measure  $\mu$  on  $\mathbb{R}^d$  is *uniformly affinely non-concentrated* if for every  $\varepsilon > 0$ , there exists  $\delta(\varepsilon) > 0$  so that  $\delta(\varepsilon) \rightarrow 0$  as  $\varepsilon \rightarrow 0$  and for all  $x \in \text{supp}(\mu)$ ,  $0 < r \leq 1$ , and every affine hyperplane  $W < \mathbb{R}^d$ , we have

$$\mu(W^{(\varepsilon r)} \cap B(x, r)) \leq \delta(\varepsilon)\mu(B(x, r)), \quad (1.2)$$

where  $W^{(r)}$  and  $B(x, r)$  denote the  $r$ -neighborhood of  $W$  and the  $r$ -ball around  $x$  respectively. We refer to the function  $\delta(\varepsilon)$  as *the non-concentration parameters* of  $\mu$ .

<sup>1</sup>We say that a measure  $\mu$  has *polynomial Fourier decay* if  $|\hat{\mu}(\xi)| = \mathcal{O}(\|\xi\|^{-\kappa})$  for some  $\kappa > 0$  as  $\|\xi\| \rightarrow \infty$ . Similarly, we say that  $\mu$  has *polylogarithmic Fourier decay* if  $|\hat{\mu}(\xi)| = \mathcal{O}((\log \|\xi\|)^{-\kappa})$  for some  $\kappa > 0$  as  $\|\xi\| \rightarrow \infty$ .

**Remark 1.2.** It can be shown by an inductive argument that a measure  $\mu$  is uniformly affinely non-concentrated if and only if (1.2) holds with  $\delta(\varepsilon) = C\varepsilon^\alpha$  for some constants  $C \geq 1$  and  $\alpha > 0$ . Hence, we often refer to such  $\mu$  as being  $(C, \alpha)$ -affinely non-concentrated when we wish to emphasize the dependence on these parameters.

The following result shows that affine non-concentration is sufficient to guarantee polynomial Fourier decay outside a sparse set of frequencies.

**Theorem 1.3** ([49, Corollary 11.5]). *Let  $\mu$  be a compactly supported Borel probability measure on  $\mathbb{R}^d$  which is uniformly affinely non-concentrated. Then, for every  $\epsilon > 0$ , there exists effectively computable  $\tau = \tau(\epsilon) > 0$ , depending only on  $\epsilon$  and the non-concentration parameters on  $\mu$ , such that the set*

$$\{\|\xi\| \leq T : |\widehat{\mu}(\xi)| > T^{-\tau}\}$$

*can be covered by  $\mathcal{O}_\epsilon(T^\epsilon)$  balls of radius 1. Moreover, the implicit constants in our  $\mathcal{O}_\epsilon(T^\epsilon)$  bound only depend upon the non-concentration parameters of  $\mu$  as well as the diameter of its support.*

Special cases of Theorem 1.3 for measures on the real line were known before by different methods by works of Kaufman [48] and Tsujii [88] for self-similar measures, and by Rossi-Shmerkin [74] more generally.

**Remark 1.4.** It is shown in [49, Corollary 11.5] that the conclusion of Theorem 1.3 holds under a much weaker non-concentration hypothesis which allows concentration to happen at some scales and on small measure sets. This property is satisfied for instance by Patterson-Sullivan measures for *cusped* real hyperbolic manifolds. Moreover, in forthcoming work of Banaji and Yu [13], the authors are able to obtain the same conclusion as Theorem 1.3 for certain higher dimensional self-similar measures which are not known to satisfy the non-concentration hypotheses of the results in [49]. In particular, our strategy can be adapted to obtain rates of Fourier decay for even larger classes of dynamically defined measures that fail to satisfy the uniform affine non-concentration hypothesis.

As we shall see below, the non-concentration hypothesis in Theorem 1.3 is known to hold for a large class of dynamically defined measures. In particular, as noted before, the main difficulty in implementing the above strategy lies in achieving **Step 3** concerning separation of the images of the frequency under the dynamics. We note however that such separation is the heart of the difficulty in obtaining Fourier decay. Indeed, all the other steps of the above strategy, including the conclusion of Theorem 1.3, apply equally well to the Cantor-Lebesgue measure on Cantor's middle thirds set. However, it is well-known that the Fourier transform of this measure does not tend to 0 at infinity [59]. Nonetheless, we believe that the directness of the strategy will enable it to be applied in much broader contexts than those covered in this article.

Let us now describe in detail the main results of this article. To this end, and motivated by Theorem 1.3, it will be convenient to make the following definition.

**Definition 1.5.** We say that the Fourier transform of a Borel probability measure  $\mu$  on  $\mathbb{R}^d$  is *polynomially decaying on average* if for every  $\epsilon > 0$ , there exists  $\tau = \tau(\epsilon) > 0$ , such that the set

$$\{\|\xi\| \leq T : |\widehat{\mu}(\xi)| > T^{-\tau}\}$$

can be covered by  $\mathcal{O}_\epsilon(T^\epsilon)$  balls of radius 1.

**1.3. Self-similar systems.** Consider a self-similar iterated function system (IFS) of the form  $\Phi = \{x \mapsto f_a(x) = r_a O_a x + t_a : a \in \mathcal{A}\}$ , where  $\mathcal{A}$  is a finite set,  $|r_a| \in (0, 1)$ ,  $O_a \in O(d)$  and  $t_a \in \mathbb{R}^d$ . Given a probability vector  $\mathbf{p} = (p_a)_{a \in \mathcal{A}}$  then it is well-known that there exists a unique Borel probability measure satisfying

$$\mu = \sum_{a \in \mathcal{A}} p_a f_a \mu.$$

Here,  $f_a\mu$  denotes the pushforward of  $\mu$  under  $f_a$ . We call this measure the self-similar measure corresponding to  $\Phi$  and  $\mathbf{p}$ . Throughout this paper we will always assume that our IFS is non-trivial in the sense that the similarities do not all share a common fixed point. This ensures that the associated self-similar measures are non-atomic.

The Fourier decay properties of self-similar measures are well studied in the literature. Recently Rapaport [71] completed a classification of those self-similar measures satisfying

$$\limsup_{\|\xi\| \rightarrow \infty} |\widehat{\mu}(\xi)| > 0.$$

This built upon earlier work of Brémont [18] and Varjú-Yu [91].

Far less is known about explicit rates of decay when  $\widehat{\mu}(\xi) \rightarrow 0$  as  $\|\xi\| \rightarrow \infty$ . In dimension  $d = 1$ , a result of Solomyak [82] states that typical self-similar measures in  $\mathbb{R}$  have polynomial Fourier decay. That being said, polynomial Fourier decay has only been verified under certain restrictive algebraic assumptions on the contraction ratios. See the articles of Dai, Feng, and Wang [21], and Streck [85].

Li and the third author in [56] have shown that, under the assumption that the log-contraction ratios  $\log|r_a|/\log|r_b|$ ,  $a \neq b$ , contain a Diophantine number: for some  $c > 0$  and  $l > 2$  we have for all rational numbers  $p/q$  that:

$$\left| \frac{\log|r_a|}{\log|r_b|} - \frac{p}{q} \right| \geq \frac{c}{q^l},$$

then  $|\widehat{\mu}(\xi)| \rightarrow 0$  polylogarithmically. This happens for example if the IFS contains contractions by  $1/2$  and  $1/3$ . Li and the third author's proof used ideas from renewal theory. Varjú-Yu [91] considered the complementary case where the contraction ratios are all integer powers of a common number. In this setting, they proved polylogarithmic Fourier decay for self-similar measures satisfying certain number theoretic assumptions.

In higher dimensions, the work of Lindenstrauss and Varjú [58] introduced a representational theoretic method to bound Fourier transforms of self-similar measures with dense rotation parts  $O_a$ . In dimension  $d \geq 3$ , this method implies sub-polynomial Fourier decay for all self-similar measures, where the rotational components  $O_a$  of the contractions give a dense subgroup; cf. the survey [76]. Moreover, under a spectral gap condition for the transfer operator  $f \mapsto \sum_{a \in \mathcal{A}} p_a f \circ O_a^{-1}$  on  $L^2(\text{SO}(d))$ ,  $d \geq 3$ , it is possible to show that the associated self-similar measure has polynomial Fourier decay.

Our first result for self-similar measures is a higher dimensional generalisation of the aforementioned result of Li and Sahlsten [56].

**Theorem 1.6.** *Let  $\Phi = \{x \mapsto f_a(x) = r_a O_a x + t_a\}_{a \in \mathcal{A}}$  be a self-similar IFS such that two of the contractions have log-contraction ratio  $\frac{\log|r_a|}{\log|r_b|}$  which is Diophantine. Let  $\mu$  be a self-similar measure for  $\Phi$  with polynomially decaying on average Fourier transform. Then,  $\mu$  has polylogarithmic Fourier decay.*

We emphasise that in Theorem 1.6 the IFS does not have to satisfy any separation assumptions. Combining Theorem 1.6 with Theorem 1.3, and forthcoming work of Banaji and Yu [13] yields the following corollary. Recall that an IFS  $\Phi$  is said to be *irreducible* if the maps in  $\Phi$  do not preserve a common proper affine subspace.

**Corollary 1.7.** *Let  $\Phi = \{x \mapsto f_a(x) = r_a O_a x + t_a\}_{a \in \mathcal{A}}$  be a self-similar IFS acting on  $\mathbb{R}^d$  such that two of the contractions have log-contraction ratio  $\frac{\log|r_a|}{\log|r_b|}$  which is Diophantine. Let  $\mu$  be a self-similar measure for  $\Phi$ . Assume that at least one of the following holds:*

- (1)  $\mu$  is uniformly affinely non-concentrated, or
- (2)  $\Phi$  is irreducible and the orthogonal matrices  $\{O_a : a \in \mathcal{A}\}$  commute.

*Then,  $\mu$  has polylogarithmic Fourier decay.*

**Remark 1.8.** To demonstrate the scope of Corollary 1.7, we recall here several results where its hypotheses have been shown to hold.

- (1) The non-concentration hypothesis in Corollary 1.7(1) is known to hold for a large class of self-similar measures. In [13], Banaji and Yu prove that if a self-similar measure is doubling (which is satisfied when  $\Phi$  satisfies the strong separation condition) and not supported on an affine subspace (which holds when  $\Phi$  is irreducible), then  $\mu$  is uniformly affinely non-concentrated. In [50, Theorem 2.3], it is shown that if  $\Phi$  is irreducible and satisfies the open set condition and  $s$  is the Hausdorff dimension of its attractor, then the restriction of the  $s$ -dimensional Hausdorff measure to the attractor of  $\Phi$  is uniformly affinely non-concentrated.
- (2) Banaji and Yu [13] have also shown that the polynomial decay on average assumption appearing in Theorem 1.6 is satisfied under the second hypothesis of Corollary 1.7. In particular, they show that a self-similar measure  $\mu$  enjoys polynomial decay on average if  $\mu$  is not supported on an affine subspace, and if the underlying IFS satisfies a non-expanding property. The latter property is satisfied whenever the orthogonal parts of the similarities commute. Note that such measures are not known to satisfy affine non-concentration in the form needed to apply the flattening results in [49]. For further details we refer the reader to [13].

It is also possible to use the Diophantine properties of the orthogonal matrices in our IFS to prove Fourier decay results. Before stating our result in this direction we need to introduce some notation. Given a matrix  $O \in SO(2)$ , then  $O$  acts on  $\mathbb{R}^2$  by rotating anticlockwise by  $2\pi\theta$  for some  $\theta \in [0, 1]$ . We define  $\theta_O$  to be the unique  $\theta$  defined this way. Moreover, given a self-similar IFS  $\{x \mapsto f_a(x) = r_a O_a x + t_a\}_{a \in \mathcal{A}}$  acting on  $\mathbb{R}^2$  such that  $O_a \in SO(2)$  for all  $a \in \mathcal{A}$ , we define  $\theta_a = \theta_{O_a}$ . We say that a pair of real numbers  $(\theta_1, \theta_2)$  is Diophantine if there exists  $C, l > 0$  such that for any non-zero  $(p, q, r) \in \mathbb{Z}$  we have

$$|p\theta_1 + q\theta_2 + r| \geq \frac{C}{\max\{|p|, |q|\}^l}.$$

With this terminology we can state our theorem.

**Theorem 1.9.** *Let  $\Phi = \{x \mapsto f_a(x) = r_a O_a x + t_a\}_{a \in \mathcal{A}}$  be a self-similar IFS acting on  $\mathbb{R}^2$  such that  $O_a \in SO(2)$  for all  $a \in \mathcal{A}$ , and there exists two contractions such that  $(\theta_a, \theta_b)$  is Diophantine. Then every self-similar measure for this IFS has polylogarithmic Fourier decay.*

**Remark 1.10.** Although we do not pursue this here, we remark that the exponent appearing in our polylogarithmic rate of Fourier decay provided by our proof of Theorems 1.6 and 1.9 can be made explicit. Explicit rates obtained in this manner depend only on the Diophantine exponent  $l$  as well as on the function  $\tau = \tau(\epsilon)$  provided by the average polynomial decay hypothesis; cf. Def. 1.5. In particular, when the self-similar measure is uniformly affinely non-concentrated, the proof of Theorem 1.3 provides explicit bounds on the average decay rate  $\tau(\epsilon)$ , which in turn depend only on the non-concentration parameters of  $\mu$ . Additionally, in the case of one-dimensional self-similar measures, arguments of the first author together with Banaji in [11] can be used to provide explicit bounds on  $\tau(\epsilon)$  in cases where Theorem 1.3 may not apply (i.e. when (uniform) affine non-concentration may fail to hold). We also refer the reader to the forthcoming work of Banaji and Yu [13] where rates of average decay are studied for certain higher dimensional self-similar measures where affine non-concentration is not known to hold.

**1.4. Patterson-Sullivan measures.** Our next result concerns Patterson-Sullivan (PS for short) measures for convex cocompact hyperbolic manifolds. In this setting, the driving mechanism behind Fourier decay comes from non-linearity of the dynamics. Namely, the fact that the strong stable and unstable foliations of the geodesic flow are not jointly integrable.

To formulate our result, let  $\Gamma$  be a discrete, Zariski-dense, convex cocompact, group of isometries of real hyperbolic space  $\mathbb{H}^{d+1}$ ,  $d \geq 1$ . Let  $\Lambda_\Gamma$  be the limit set of  $\Gamma$  on  $\partial\mathbb{H}^{d+1}$  and  $\mu$  be the Patterson-Sullivan probability measure on  $\Lambda_\Gamma$ ; cf. Section 2.2 for detailed definitions. The following is our main result in this setting.

**Theorem 1.11.** *There exists  $\kappa > 0$  such that the following holds for all  $\varphi \in C^2(\partial\mathbb{H}^{d+1})$ ,  $\psi \in C^1(\partial\mathbb{H}^{d+1})$  satisfying*

$$\|\varphi\|_{C^2} + \|\psi\|_{C^1} \leq A, \quad \inf_{x \in \Lambda_\Gamma} \|\nabla_x \varphi\| > a,$$

*for some constants  $a > 0$  and  $A \geq 1$ . There exists a constant  $C = C(A, a, \mu) \geq 1$ , so that for all  $\lambda \neq 0$ , we have*

$$\left| \int_{\Lambda_\Gamma} e^{2\pi i \lambda \varphi(x)} \psi(x) d\mu(x) \right| \leq C |\lambda|^{-\kappa}.$$

Theorem 1.11 generalises prior work of Bourgain and Dyatlov in the case of hyperbolic surfaces [14] and of Li, Naud, and Pan in the case of Schottky hyperbolic 3-manifolds [53]. These results are based on Bourgain's sum-product theory, while the proof of Theorem 1.11 is based on Theorem 1.3, which was obtained using purely additive methods. Another feature of our proof of Theorem 1.11 compared to [14, 53] is that it does not require any symbolic coding of the action of  $\Gamma$  on  $\Lambda_\Gamma$ .

**Remark 1.12.** (1) Our proof shows that the rate  $\kappa$  provided by Theorem 1.11 depends only on non-concentration parameters of  $\mu$  in the sense of Definition 1.1. In particular, the rate of decay does not change upon replacing  $\Gamma$  by a finite index subgroup since the measure  $\mu$  remains the same in this case [73].

(2) PS measures for convex cocompact manifolds are known to be  $(C, \alpha)$ -uniformly affinely non-concentrated [22]. In this language, the exponent of polynomial Fourier decay obtained in [14, 53] is shown to depend only on the exponent  $\alpha$ , while the exponent provided by Theorem 1.11 depends on both  $C$  and  $\alpha$ . Roughly speaking, this difference comes from the fact that the rates provided by Bourgain's sum-product theorem depend only on  $\alpha$ , while those provided by the flattening theorem, Theorem 1.3, depend on both  $C$  and  $\alpha$ . On the other hand, the flexible nature of Theorem 1.3 enables our unified approach to the variety of different settings considered in this article. It is thus of interest to strengthen Theorem 1.3 to remove such dependence on the constant  $C$ ; cf. Remark 1.25 for further consequences.

To keep the presentation clear, we restricted our setup to the case of convex cocompact groups. Using the recurrence results obtained in [49], the proof of Theorem 1.11 can be adapted to handle the general case of geometrically finite manifolds.

Theorem 1.11 has immediate applications in the setting of Fractal Uncertainty Principles. We will explain this in detail in Section 1.7.2.

### 1.5. Gibbs measures for self-conformal iterated function systems and their subshifts.

Next, let us consider the problem of studying Fourier transforms of iterated function systems  $\Phi = \{f_a\}_{a \in \mathcal{A}}$  for general  $C^2$  contractions  $f_a : [0, 1]^d \rightarrow [0, 1]^d$ ,  $a \in \mathcal{A}$ , where  $d \geq 1$ . In this context, we say that an iterated function system is *conformal* if for every  $a \in \mathcal{A}$  and  $x \in [0, 1]^d$  the Jacobian of  $f_a$  evaluated at  $x$  satisfies

$$D_x f_a = \lambda_a(x) O_a(x)$$

for some  $\lambda_a(x) \in (-1, 1) \setminus \{0\}$  and  $O_a(x) \in O(d)$ , i.e. an iterated function system is conformal if, for every  $a \in \mathcal{A}$  and  $x \in [0, 1]^d$ , the Jacobian is a similarity map. Given  $x \in [0, 1]^d$  and  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}^n$ , we also define

$$\lambda_{\mathbf{a}}(x) = \prod_{i=1}^n \lambda_{a_i}(f_{a_{i+1}, \dots, a_n}(x)).$$

It follows from the chain rule that

$$|\lambda_{\mathbf{a}}(x)| = \|D_x f_{\mathbf{a}}\|$$

for any  $\mathbf{a} \in \cup_{n=1}^{\infty} \mathcal{A}^n$  and  $x \in [0, 1]^d$ .

In this context, we will prove Fourier decay results for Gibbs measures associated to subshifts of finite type and  $C^1$  potentials. This framework incorporates self-similar measures. However, instead of using the number-theoretic properties of an IFS, this section will focus on how the non-linearity within the IFS can be used to prove polynomial Fourier decay. Such an assumption rules out self-similar measures.

**1.5.1. Background.** The idea that the non-linearity present within an IFS can be used as a tool for proving Fourier decay is well-established within the literature. In the case  $d = 1$ , the first result in this direction was due to Kaufman [47]. For each  $N \in \mathbb{N}$ , he considered the set  $B_N$  consisting of those badly approximable numbers whose digits in the continued fraction expansion are bounded above by  $N$ . Kaufman proved that  $B_N$  supports a measure that has polynomial Fourier decay for  $N \geq 3$ . This result was later extended by Queffelec and Ramaré [70] to the case  $N = 2$ , and later to the case of more general invariant measures for the Gauss map by Jordan and the third author [43].

In [14], Bourgain and Dyatlov introduced a new method to study Fourier transforms of measures using ideas from Additive Combinatorics. They applied this method to prove polynomial Fourier decay for Patterson-Sullivan measures for limit sets of Fuchsian groups. This method was subsequently refined by the third author and Stevens [77], and applied to a more general class of iterated functions systems.

In recent years, more substantial progress has been made in this one-dimensional setting. In particular, the first and third authors in [12], and Algom, Rodriguez-Hertz, and Wang [5], established polynomial Fourier decay for a large class of self-conformal measures without any separation assumptions on the underlying IFS. These results were recently built upon by the first author and Banaji [11], and by Algom et al. [2], who independently proved that if an IFS consists of analytic contractions one of which is not an affine map, then every self-conformal measure has polynomial Fourier decay.

In an as yet unpublished work, Avila, Lyubich and Zhang [9] proved polylogarithmic Fourier decay for certain measures coming from  $C^{1+\alpha}$  iterated function systems. We also refer the reader to the papers [3, 4] for further results in this direction. In higher dimensions, we know far less. A result of Leclerc [51] gives sufficient conditions for the Julia set of a hyperbolic rational map to support a measure with polynomial Fourier decay.

**1.5.2. Statement of results.** As mentioned above, our results hold under a spectral gap hypothesis on the underlying conformal IFS. To properly formulate this result, we need to introduce some preliminaries.

Let  $\Phi = \{f_a\}_{a \in \mathcal{A}}$  be an iterated function system. As the maps are contracting, there exists a unique non-empty compact set satisfying the relation

$$X_{\Phi} = \bigcup_{a \in \mathcal{A}} f_a(X_{\Phi}).$$

For each  $a \in \mathcal{A}$  we let  $X_a = f_a(X)$ . We say that  $\Phi$  satisfies the strong separation condition if

$$X_a \cap X_b = \emptyset$$

for distinct  $a, b \in \mathcal{A}$ .  $X_{\Phi}$  can be viewed as the image of the full shift  $\mathcal{A}^{\mathbb{N}}$  under the map  $\pi : \mathcal{A}^{\mathbb{N}} \rightarrow X_{\Phi}$  given by

$$\pi(\mathbf{a}) = \lim_{n \rightarrow \infty} (f_{a_1} \circ \cdots \circ f_{a_n})(x_0),$$

where  $x_0 \in X_\Phi$  is an arbitrary point. By considering a small neighbourhood of each  $X_a$ , it can be shown that under the strong separation condition for each  $a \in \mathcal{A}$  there exists an open set  $U_a \subset \mathbb{R}^d$  satisfying the following properties:

- $U_a \cap U_b = \emptyset$  for  $a \neq b$
- If we let  $U := \cup_{a \in \mathcal{A}} U_a$  then  $f_a(U) \subset U_a$  for each  $a \in \mathcal{A}$ .

In what follows, when an IFS satisfies the strong separation condition we will fix a choice of  $U_a$  for each  $a \in \mathcal{A}$  and a corresponding  $U$ .

Given a  $\#\mathcal{A} \times \#\mathcal{A}$  matrix  $A$  taking values in  $\{0, 1\}$  we can define the associated *subshift of finite type* to mean the subset  $\Sigma_A \subset \mathcal{A}^\mathbb{N}$  given by

$$\Sigma_A = \{(a_i) \in \mathcal{A}^\mathbb{N} : A(a_i, a_{i+1}) = 1\}.$$

Thus each such matrix  $A$  defines a compact subset  $X_A \subset X_\Phi$  via the projection  $X_A = \pi(\Sigma_A)$ . We will always assume that  $\Sigma_A$  is topologically mixing, this is equivalent to  $A^n > 0$  for all  $n$  sufficiently large. We write  $\mathcal{W}_A$  as the set of all admissible words, that is, words that exist as a prefix of any word in  $\Sigma_A$ . Given  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{W}_A$  we define

$$X_{\mathbf{a}}^A = \pi(\{(b_i) \in \Sigma_A : b_1, \dots, b_n = \mathbf{a}\}).$$

When the choice of matrix  $A$  is implicit, we will by an abuse of notation write  $X_{\mathbf{a}}$  instead of  $X_{\mathbf{a}}^A$ . Given  $\mathbf{a}, \mathbf{b} \in \mathcal{W}_A$  we write  $\mathbf{a} \rightsquigarrow \mathbf{b}$  if  $\mathbf{ab} \in \mathcal{W}_A$ . Let us assume that  $\Phi$  satisfies the strong separation condition, given a point  $x \in U$  then  $x \in U_b$  for some unique  $b \in \mathcal{A}$ , we write  $\mathbf{a} \rightsquigarrow x$  for  $\mathbf{a} \in \mathcal{W}_A$  if  $\mathbf{a} \rightsquigarrow b$ .

Given a subshift of finite type  $\Sigma_A$  and a  $C^1$  function  $\psi : U \rightarrow \mathbb{R}$ , if  $x \in U$ ,  $a \in \mathcal{A}$  and  $a \rightsquigarrow x$ , then we define  $w_a(x) = e^{\psi(f_a(x))}$ . More generally, given  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{W}_A$  such that  $\mathbf{a} \rightsquigarrow x$ , we define  $w_{\mathbf{a}}(x) = e^{\sum_{j=0}^{n-1} \psi(f_{a_{j+1} \dots a_n}(x))}$ . We define the *pressure* of  $\psi$  to be

$$P(\psi) = \lim_{n \rightarrow \infty} \frac{1}{n} \log \left( \sum_{\mathbf{a} \in \mathcal{A}^n : \mathbf{a} \in \mathcal{W}_A} \sup_{x \in \Sigma_A : \mathbf{a} \rightsquigarrow x} w_{\mathbf{a}}(x) \right).$$

For a proof that the pressure of a  $C^1$  function always exists see [92]. For each  $\psi$ , there exists a unique probability measure  $\mu_\psi$  supported on  $X_A$  called the Gibbs measure of  $\psi$ . The defining property of this measure is that it satisfies the Gibbs condition: There exists  $C > 0$  such that for all  $\mathbf{a} \in \mathcal{W}_A$  and  $x \in X_A$  satisfying  $\mathbf{a} \rightsquigarrow x$  we have

$$C^{-1} \leq \frac{\mu_\psi(X_{\mathbf{a}})}{w_{\mathbf{a}}(x)e^{nP(\psi)}} \leq C. \quad (1.3)$$

Gibbs measures are well studied within the ergodic theory community. See [17] and [65] for a more thorough introduction to these measures.

Given a conformal IFS  $\Phi$  satisfying the strong separation condition, a subshift of finite type  $\Sigma_A$  and a  $C^1$  potential  $\psi : U \rightarrow \mathbb{R}$  satisfying  $P(\psi) = 0$ , we define the twisted transfer operators  $\mathcal{L}_{ib}$ ,  $b \in \mathbb{R}$ , by the formula

$$\mathcal{L}_{ib}h(x) := \sum_{\substack{a \in \mathcal{A} \\ a \rightsquigarrow x}} w_a(x) |\lambda_a(x)|^{ib} h(f_a(x)), \quad h \in C^{1,b}(U), x \in U,$$

in the Banach space  $C^{1,b}(U)$  of differentiable functions equipped with the norm

$$\|h\|_b := \|h\|_\infty + \frac{\sup_{x \in U} \|D_x h\|}{|b|}.$$

**Definition 1.13.** Let  $\Phi$ ,  $\Sigma_A$ , and  $\psi$  be given. We say that  $(\Phi, \Sigma_A, \psi)$  has a spectral gap if there exists  $0 < \rho, v < 1$  such that for all large enough  $|b|$ ,  $n \in \mathbb{N}$  and  $h \in C^{1,b}(U)$ , we have

$$\|\mathcal{L}_{ib}^n h\|_b \ll \rho^n |b|^v \|h\|_b.$$

Our main result in this section is the following general theorem.

**Theorem 1.14.** *Let  $\Phi$  be a conformal IFS satisfying the strong separation condition, let  $\Sigma_A$  be a subshift of finite type, and let  $\psi$  be a  $C^1$  potential satisfying  $P(\psi) = 0$ . Assume that  $(\Phi, \Sigma_A, \psi)$  has a spectral gap and suppose that  $\mu_\psi$  is uniformly affinely non-concentrated. Then  $\mu_\psi$  has polynomial Fourier decay.*

To the best of the authors' knowledge, Theorem 1.14 is the first polynomial Fourier decay result for self-conformal iterated function systems and Gibbs measures in arbitrary dimensions.

It is a well established fact that a suitable uniform non-integrability (UNI) condition can imply the existence of a spectral gap for a twisted transfer operator. See for instance the works of Naud [62] and Stoyanov [84, Theorem 1.1] who adapted Dolgopyat's method [25] to prove spectral gap results. We also refer the reader to the works of Avila, Gouëzel and Yoccoz [8], Sarkar-Winter [78] and Li-Pan [54] for similar spectral gap results. The uniform non-integrability condition we will work with is the following definition. It is inspired by the UNI condition introduced by Li and Pan [54, Lemma 4.5] in the context of Patterson-Sullivan measures.

**Definition 1.15** (UNI). Let  $\Phi$  be a conformal IFS satisfying the strong separation condition and  $\Sigma_A$  be a subshift of finite type. We say that the pair  $(\Phi, \Sigma_A)$  satisfies the *Uniform Non-Integrability* (UNI) condition if there exists  $\epsilon_0 > 0$  such that the following holds for infinitely many  $n \in \mathbb{N}$ : There exists  $x \in X_A$  such that for any unit vector  $e \in \mathbb{R}^d$  there exists  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  such that:

- We have

$$|\partial_e (\log |\lambda_{\mathbf{a}_1}(x)| - \log |\lambda_{\mathbf{a}_2}(x)|)| \geq \epsilon_0.$$

- $\mathbf{a}_1 \rightsquigarrow x$  and  $\mathbf{a}_2 \rightsquigarrow x$ .

Roughly speaking, the UNI condition holds if for infinitely many  $n$  one can find an  $x \in \Sigma_A$  at which we can observe non-linearity in all directions. We include some further discussion on this definition and some examples of iterated function systems satisfying it at the end of this subsection. The significance of our UNI condition is demonstrated by the following proposition.

**Proposition 1.16** (Spectral gap). *Let  $\Phi$  be a conformal IFS satisfying the strong separation condition and  $\Sigma_A$  be a subshift of finite type. Suppose that  $(\Phi, \Sigma_A)$  satisfies the UNI condition, then for any  $C^1$  potential  $\psi$  satisfying  $P(\psi) = 0$ , there exists  $0 < \rho < 1$  such that for all large enough  $|b|$ ,  $n \in \mathbb{N}$  and  $h \in C^{1,b}(U)$ , we have*

$$\|\mathcal{L}_{ib}^n h\|_b \ll \rho^n |b|^{1/2} \|h\|_b.$$

*In particular, if  $(\Phi, \Sigma_A)$  satisfies the UNI condition, then for any  $C^1$  potential  $\psi$  satisfying  $P(\psi) = 0$  the triple  $(\Phi, \Sigma_A, \psi)$  has a spectral gap.*

The proof of Proposition 1.16 is provided in Appendix C. Our proof of Proposition 1.16 relies heavily on arguments due to Li and Pan [54], and to Naud [62].

Given a Gibbs measure  $\mu_\psi$  associated to some  $C^1$  potential  $\psi$ , it can be shown that  $\mu_\psi$  is the Gibbs measure for another  $C^1$  function  $\psi'$  that satisfies  $P(\psi') = 0$ . Combining this observation with Theorem 1.14 and Proposition 1.16 yields the following statement.

**Theorem 1.17.** *Let  $\Phi$  be a conformal IFS satisfying the strong separation condition and  $\Sigma_A$  be a subshift of finite type. Suppose that  $(\Phi, \Sigma_A)$  satisfies the UNI condition. Then any Gibbs measure  $\mu_\psi$  that is uniformly affinely non-concentrated has polynomial Fourier decay.*

In [54] Li and Pan introduced the following seemingly stronger uniform non-integrability condition: there exists  $r > 0$  and  $\epsilon_0 > 0$  such that for any large  $n \in \mathbb{N}$ , any  $x \in X_A$  and unit vector  $e \in \mathbb{R}^d$ , there exist  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  such that:

- For all  $y \in B(x, r)$  we have

$$|\partial_e (\log |\lambda_{\mathbf{a}_1}(y)| - \log |\lambda_{\mathbf{a}_2}(y)|)| \geq \epsilon_0.$$

- For all  $y \in B(x, r)$  we have  $\mathbf{a}_1 \rightsquigarrow y$  and  $\mathbf{a}_2 \rightsquigarrow y$ .

This condition and our UNI condition are in fact equivalent. We include a proof of this fact in Appendix D. This proof follows an argument of Avila, Gouetzel and Yoccoz [8]. It will be this latter condition that we will use in Appendix C during our proof of Proposition 1.16.

**Remark 1.18.** (a) Our proof of Theorem 1.14 relies on showing that the derivatives arising from our IFS satisfy a non-concentration property. We prove that this property holds using our spectral gap assumption. A similar non-concentration statement is needed in the proof of Theorem 1.11. However, in our proof of Theorem 1.11 we do not use a spectral gap for the transfer operator and instead use explicit formulas for the derivatives directly. It would be interesting if the non-concentration of the derivatives in the setting of Theorem 1.14 could be also deduced directly from a suitable UNI condition without using a spectral gap assumption. We leave this as an interesting problem to pursue.

- (b) In dimension  $d = 1$ , Theorem 1.14 was proved in the work of the third author and Stevens [77], but as in Bourgain and Dyatlov [14], it relied on Bourgain's sum product theorem. Here we do not need it, so Theorem 1.14 also gives a new proof in the case  $d = 1$ . For dimension  $d = 2$ , Theorem 1.14 extends the recent work of Leclerc [51], regarding Fourier dimension of Julia sets, which used a spectral gap result for twisted transfer operators provided by Oh-Winter [63]. In dimension  $d = 2$ , we also highlight an upcoming work of Algom, Rodriguez Hertz, and Wang which provides a full characterization of polynomial Fourier decay in a similar spirit as the works of [12, 3, 4, 11, 2]. It would be interesting to obtain a similar classification in higher dimensions.
- (c) Theorem 1.17 is a statement that guarantees polynomial Fourier decay for Gibbs measures assuming the subshift satisfies a non-linearity assumption that is formulated in terms of the Jacobian matrices  $\{D_x f_a = \lambda_a(x) O_a(x)\}_{a \in \mathcal{A}}$ . In particular, we use properties of the contraction ratios  $\{\lambda_a(\cdot)\}_{a \in \mathcal{A}}$  to prove our result. In the spirit of Theorem 1.9, it would be interesting if an analogous result to Theorem 1.14 could be proved using properties of the orthogonal matrices  $\{O_a(\cdot)\}_{a \in \mathcal{A}}$  instead.

1.5.3. *Examples.* In this subsection, we discuss examples where the hypotheses of Theorems 1.14 and 1.17 are satisfied.

Whether a Gibbs measure satisfies the uniformly affinely non-concentrated property is a well investigated problem. In the context of the full shift, the simplest example of a Gibbs measure satisfying this property is when  $\psi$  is the geometric potential,  $\Phi$  satisfies the strong separation condition and  $\dim_H(X_\Phi) > d - 1$ . Under these assumptions it is well known that  $\mu_\psi(B(x, r)) \asymp r^{\dim_H(X)}$  for any  $x \in X_\Phi$ . It is straightforward to apply these properties to prove that  $\mu_\psi$  satisfies the uniformly affinely non-concentrated property. In the case of arbitrary Gibbs measures we refer the reader to [22, 90] and the references therein. In particular, Theorem 1.5 from [90] (see also Theorem 1.10 from [22]) shows that under natural assumptions on the conformal IFS and the Gibbs measure  $\mu_\psi$ , if the attractor of the IFS is not contained in a proper real analytic submanifold of  $\mathbb{R}^d$  then the Gibbs measure  $\mu_\psi$  will be uniformly affinely non-concentrated.

Turning our attention to our UNI hypothesis, the condition in Definition 1.15 is inspired by the work of Li and Pan [54], where a similar condition was introduced in the context of PS measures on limit sets of cusped, geometrically finite, groups of Möbius transformations. In light of the close relationship between such limit sets and attractors of conformal IFSs, it seems possible to extend the techniques of the proof of Theorem 1.14 to recover the polynomial Fourier decay result of Theorem 1.11 as well as its generalisations to other Gibbs measures. To this end, the results of [54] can be adapted to construct a coding of limit sets of convex cocompact groups using suitable Markov systems satisfying our UNI hypotheses. As noted before, this method yields worse dependence of the rates of decay on geometric non-concentration properties of the measure due to our use of spectral gap results in the proof.

Nonetheless, motivated by this connection, we provide explicit further examples of attractors of conformal IFSs in  $\mathbb{R}^d$  satisfying the hypotheses of Theorem 1.17. The examples we give are based on certain involution Möbius transformations on  $\mathbb{R}^d$  (see (1.4) below). We note that, in dimension  $d \geq 3$ , Liouville's Theorem (see [41, Chapter 5]) states that *any* conformal mapping is either of this form or is a similarity map, so the class of maps in this example is not very restrictive. Our examples will be formulated in the special case when  $\Sigma_A$  is the full shift  $\mathcal{A}^{\mathbb{N}}$ , but they can easily be adapted to incorporate more general subshifts of finite type.

**Example 1.19.** Let  $\Phi = \{f_a : [0, 1]^d \rightarrow [0, 1]^d\}_{a \in \mathcal{A}}$  be a conformal IFS. Assume that  $\Phi$  contains  $d + 1$  contractions of the form

$$f_{a_i}(x) = t_{a_i} + \frac{\lambda_{a_i} O_{a_i}(x - u_{a_i})}{|x - u_{a_i}|^2} \quad (1.4)$$

where  $t_{a_i} \in \mathbb{R}^d$ ,  $u_{a_i} \in \mathbb{R}^d \setminus [0, 1]^d$ ,  $\lambda_{a_i} \in (0, 1)$ , and  $O_{a_i} \in O(d)$ . Let us also suppose that the set of vectors  $\{u_{a_i}\}_{i=1}^{d+1}$  do not belong to an affine subspace of  $\mathbb{R}^d$ . Maps satisfying (1.4) have the following useful property: Let  $x \in [0, 1]^d$  and  $e \in \mathbb{R}^d$  be a unit vector. Then for each  $a_i$  we have

$$\partial_e \log |\lambda_{a_i}(x)| = -\frac{2\langle x - u_{a_i}, e \rangle}{|x - u_{a_i}|^2}. \quad (1.5)$$

This fact was proved in [54, Lemma 6.75]. It can be shown to hold by a direct calculation. Let us also assume now that  $\Phi$  contains a similarity given by:

$$f_{a_s}(x) = \lambda_{a_s} O_{a_s} x + t_{a_s}$$

for some  $t_{a_s} \in \mathbb{R}^d$ ,  $\lambda_{a_s} \in (0, 1)$ , and  $O_{a_s} \in O(d)$ .

We now bring our attention to proving that the UNI condition is satisfied by any conformal IFS satisfying the above conditions. Let  $n \in \mathbb{N}$  be arbitrary and  $x \in X_\Phi$ . Then for any unit vector  $e \in \mathbb{R}^d$  and  $a_i$  we have

$$\begin{aligned} \partial_e \left( \log |\lambda_{a_s^{n-1} a_i}(x)| - \log |\lambda_{a_s^n}(x)| \right) &= \partial_e \left( \log \lambda_{a_s}^{n-1} |\lambda_{a_i}(x)| - \log \lambda_{a_s}^n \right) \\ &= \partial_e \log |\lambda_{a_i}(x)| \\ &= -\frac{2\langle x - u_{a_i}, e \rangle}{|x - u_{a_i}|^2}, \end{aligned}$$

where the final line follows by (1.5). In particular, the expression obtained for  $\partial_e (\log |\lambda_{a_s^{n-1} a_i}(x)| - \log |\lambda_{a_s^n}(x)|)$  does not depend upon  $n$ . Thus, the UNI condition is satisfied if for any unit vector  $e \in \mathbb{R}^d$  we can find  $a_i$  such that

$$\langle x - u_{a_i}, e \rangle \neq 0.$$

Such an  $a_i$  has to exist because of our assumption that the set of vectors  $\{u_{a_i}\}_{i=1}^{d+1}$  do not belong to an affine subspace of  $\mathbb{R}^d$ .

We emphasise that it is possible to choose the contractions in this example in such a way that  $\dim_H(X_\Phi) > d - 1$  and the strong separation condition is satisfied by  $\Phi$ . Thus by the discussion in above, the Gibbs measure corresponding to the geometric potential would in this case provide an explicit example where Theorem 1.17 guarantees polynomial Fourier decay.

We end this discussion with several remarks and questions. Firstly, it is easy to check that the set of  $C^2$  conformal IFSs satisfying our UNI condition is open, while the above example shows that this set is non-empty. Moreover, close examination of Example 1.19 suggests that failure of the UNI condition places algebraic relations among the maps in the IFS. It is thus natural to expect that our UNI condition holds generically.

This discussion, along with Liouville's Theorem, raises the following question: Suppose  $\Phi$  is a conformal IFS acting on  $\mathbb{R}^d$  for  $d \geq 3$ . If  $\Phi$  contains a map of the form described by (1.4), then

must  $\Phi$  satisfy the UNI condition? If this question could be answered in the affirmative, then Theorem 1.17 would go a long way towards classifying polynomial Fourier decay for Gibbs measures in dimension  $d \geq 3$ .

Finally, in view of the restriction placed by Liouville's Theorem on conformal IFSs in dimensions  $d \geq 3$ , it is natural to consider generalisations of Theorem 1.17 by relaxing the conformality hypothesis to non-algebraic systems provided by *Quasiconformal-* or *Quasiregular Mappings*, with the price of lack of regularity of the distortion function  $x \mapsto \log |f'_a(x)|$ . The study of such systems has received a lot of attention in recent years; cf. [40, 60, 64]. In particular, it is of interest to extend the proofs of Theorems 1.14 and 1.17 to these closely related systems thus yielding a wider source of examples of polynomial Fourier decay; see Section 1.8.2 for further discussion of this direction.

**1.6. Non-conformal systems.** Our techniques also allow us to prove polynomial Fourier decay for certain stationary measures arising from non-conformal iterated function systems. It is well-known that such measures are more difficult to analyze than self-similar and self-conformal measures. Indeed the question of whether such a measure is uniformly affinely non-concentrated is far more delicate than in the self-similar and self-conformal setting. As such our knowledge of the Fourier decay properties of stationary measures for non-conformal iterated function systems is far less extensive.

A result of Solomyak [83] demonstrates that amongst parameterised families of self-affine measures one should expect a typical member to have polynomial Fourier decay. A result of Li and the third author [55] established polynomial Fourier decay for self-affine measures when the underlying IFS satisfies suitable algebraic assumptions. To the best of the authors' knowledge, our main result (Theorem 1.21) is the first general result in this setting that establishes polynomial Fourier decay using the non-linearity within the IFS.

The family of non-conformal IFSs we consider is defined as follows. Suppose that, for each  $1 \leq i \leq d$ , we are given a  $C^2$  IFS  $\{f_a^{(i)} : a \in \mathcal{A}_i\}$  acting on  $[0, 1]$ . Given  $\bar{a} = (a_1, \dots, a_d) \in \mathcal{A}_1 \times \dots \times \mathcal{A}_d$  we define the map

$$F_{\bar{a}}(x_1, \dots, x_d) = (f_{a_1}^{(1)}(x_1), \dots, f_{a_d}^{(d)}(x_d)).$$

Using this notation, given any  $\mathcal{A} \subset \mathcal{A}_1 \times \dots \times \mathcal{A}_d$  we can define a new IFS  $\{F_{\bar{a}}\}_{\bar{a} \in \mathcal{A}}$  acting on  $[0, 1]^d$ . We will call such an IFS a *restricted product IFS*.

Given a restricted product IFS  $\{F_{\bar{a}}\}_{\bar{a} \in \mathcal{A}}$  and a probability vector  $\mathbf{p} = (p_{\bar{a}})_{\bar{a} \in \mathcal{A}}$ , we let  $\mu_{\mathbf{p}}$  denote the unique Borel probability measure satisfying

$$\mu_{\mathbf{p}} = \sum_{\bar{a} \in \mathcal{A}} p_{\bar{a}} F_{\bar{a}} \mu_{\mathbf{p}}.$$

We call  $\mu_{\mathbf{p}}$  a stationary measure.

In this setting, it is reasonable to expect that if the IFS is suitably non-linear then every stationary measure will have polynomial Fourier decay. In this section our notion of non-linearity comes from the following definition.

**Definition 1.20.** Let  $\Phi = \{f_a\}$  be a  $C^2$  IFS acting on  $[0, 1]$  with attractor  $X_{\Phi}$ . We say that  $\Phi$  satisfies the Uniform Non-integrability (UNI) condition if there exists  $\epsilon_0 > 0$  such that for all  $n \in \mathbb{N}$  there exists  $\mathbf{a}, \mathbf{b} \in \mathcal{A}^n$  satisfying

$$\epsilon_0 < |(\log |f'_{\mathbf{a}}| - \log |f'_{\mathbf{b}}|)'(x)|^2$$

for all  $x \in X_{\Phi}$ .

<sup>2</sup>This definition often also includes the upper bound  $|(\log |f'_{\mathbf{a}}| - \log |f'_{\mathbf{b}}|)'(x)| < \epsilon_1$  for all  $x \in X_{\Phi}$  for some  $\epsilon_1 > 0$ . However the existence of such an  $\epsilon_1 > 0$  follows automatically from the regularity of the underlying IFS. Hence we omit it.

The Fourier decay properties of measures arising from iterated function systems satisfying this UNI condition have been well studied (see [5, 12, 77]). In the context of one dimensional  $C^2$  iterated function systems, it is known (see for example [5]) that if an IFS  $\Phi$  cannot be  $C^2$  conjugated to a linear IFS, i.e. there exists no  $C^2$  diffeomorphism  $h : \mathbb{R} \rightarrow \mathbb{R}$  such that  $\{g_a := h \circ f_a \circ h^{-1}\}_{a \in \mathcal{A}}$  satisfies  $g_a''(x) = 0$  for all  $x \in X_\Phi$  and  $a \in \mathcal{A}$ , then  $\Phi$  satisfies the UNI condition.

Our main result in this section is the following statement.

**Theorem 1.21.** *Let  $\{F_{\bar{a}}\}_{\bar{a} \in \mathcal{A}}$  be a restricted product IFS. Assume that the following properties are satisfied:*

- (1) *For each  $1 \leq i \leq d$ , the one-dimensional IFS  $\{f_a^{(i)}\}_{a \in \mathcal{A}_i}$  satisfies the strong separation condition.*
- (2) *For each  $\bar{a} \in \mathcal{A}$  and  $1 \leq i \leq d$ , there exists  $\bar{a}' \in \mathcal{A}$  such that  $a_j = a'_j$  for all  $j \neq i$  and  $a_i \neq a'_i$ .*
- (3) *For each  $1 \leq i \leq d$ , there exists  $\bar{a} \in \mathcal{A}$  such that the IFS*

$$\left\{ f_{a'_i}^{(i)} : (a'_1, \dots, a'_d) \in \mathcal{A}, a'_j = a_j \forall j \neq i \right\}$$

*satisfies the UNI condition.*

*Then, every stationary measure has polynomial Fourier decay.*

To help illuminate Theorem 1.21 we include an example of an IFS satisfying its assumptions: Let

$$\Phi := \left\{ f_{(i,j)}(x, y) = \left( \frac{1}{x+i}, \frac{1}{y+j} \right) : (i, j) \in \{(1, 2), (1, 3), (2, 1), (2, 3), (3, 1), (3, 2)\} \right\}.$$

Theorem 1.21 implies that all of the stationary measures for this IFS have polynomial Fourier decay.

**Remark 1.22.** The significance of defining our restricted product IFSs using products of one-dimensional IFSs, as opposed to products of  $d$ -dimensional IFSs, is that this property will allow us to disintegrate a stationary measure into an integral of random one-dimensional measures. These random measures were studied in [12]. In particular, our proof of Theorem 1.21 will rely upon a result from [12] which establishes a spectral gap for typical random compositions of transfer operators (see Proposition 6.5). Hence, a higher dimensional generalization of the results in [12] would lead to a version of Theorem 1.21 for products of higher dimensional IFSs.

**1.7. Applications.** We now discuss some of the consequences of the results presented here to equidistribution of vectors in fractal sets and quantum chaos.

**1.7.1. Equidistribution of vectors on fractals.** The study of Diophantine properties of typical points in the supports of dynamically defined measures has witnessed considerable activity in recent years; cf. [23, 81, 24, 39, 69] for instance and references therein.

In dimension one, an underlying principle behind many of the results in the subject can be summarized as follows: the continued fraction expansion as well as the digit expansions in different bases of a typical point should be independent of one another. Higher dimensional results also echo analogous principles.

This principle is exemplified by the following results. Host's theorem [39], and its generalization in [37], assert that if multiplicatively independent integers  $p, q \in \mathbb{N}$  and a  $\times p$ -invariant and ergodic measure  $\mu$  on  $[0, 1)$  of positive entropy are given, then the  $\times q$ -orbit of  $\mu$ -almost every point is equidistributed with respect to Lebesgue measure; see also [36] for a Fourier analytic proof of this result.

A second example is given by a result of Simmons and Weiss [81] asserting that if  $\mu$  is a non-trivial self-similar measure on  $[0, 1)$ , then the orbit of  $\mu$ -almost every point under the Gauss map is equidistributed towards the unique absolutely continuous Gauss measure.

Much less is known in higher dimensions. In [81], the authors obtain higher dimensional analogues of their aforementioned results for the Gauss map. In [24] Dayan, Ganguly and Weiss show almost sure  $A$ -normality for certain self-affine measures with linear parts given by negative powers of  $A$  and satisfying additional Diophantine conditions on their translation parts. In both cases, the results are based on measure rigidity results for certain random walks on homogeneous spaces.

In this vein, Fourier decay has provided an important avenue for obtaining new results for measures that are not amenable to analysis by other methods. Most notably, in [23], Davenport, Erdős, and LeVeque (DEL) showed that any<sup>3</sup> probability measure  $\mu$  on  $\mathbb{R}^d$  with polylogarithmic Fourier decay has that  $\mu$ -almost every  $x$  is  $A$ -normal with respect to any expanding integer matrix  $A$ , i.e., the orbit  $(A^n x)_{n \in \mathbb{N}}$  is equidistributed in the torus  $\mathbb{R}^d/\mathbb{Z}^d$  when reduced mod 1; cf. [69] for a recent extension of this result. Here, a matrix is expanding if its determinant has modulus larger than one. In light of this criterion, our results provide numerous new examples where the above principle holds.

**Corollary 1.23.** *Let  $\mu$  be any of the measures in the Theorems 1.6, 1.9, 1.11, 1.14 or 1.21. Then, we have that  $\mu$ -almost every  $x$  is  $A$ -normal with respect to any expanding integer valued matrix  $A$ .*

We also note that Corollary 1.23 opens the possibility of generalizing the results of [69] to higher dimensions.

**1.7.2. Fractal Uncertainty Principles, spectral gaps, and scattering resonances.** Our next application concerns essential spectral gaps for Laplacians on convex cocompact hyperbolic manifolds. To properly place our results in context, we introduce the relevant notions and previous progress on this problem.

Let  $\Gamma$  be a convex cocompact subgroup of  $\text{Isom}^+(\mathbb{H}^{d+1})$ ,  $d \geq 1$ , and denote by  $M$  the quotient manifold  $\mathbb{H}^{d+1}/\Gamma$ . As before, we let  $\delta_\Gamma$  denote the Hausdorff dimension of the limit set  $\Lambda_\Gamma$  of  $\Gamma$ . We let  $\Delta_M$  denote the Laplace-Beltrami operator on  $M$  and for  $\lambda \in \mathbb{C}$  with sufficiently large imaginary part, we define the resolvent of  $\Delta_M$  by  $R(\lambda) := (-\Delta_M - \frac{d^2}{4} - \lambda^2)^{-1}$ .

A topic of very active interest concerns meromorphic continuation of the analytic family of operators  $\lambda \mapsto R(\lambda)$  as well as the existence and location of its poles. Results of this type have important applications to rates of decay of solutions to wave equations in quantum chaos as well as to rates of decay of correlation in the field of hyperbolic dynamics. We refer the reader to [29] for an introduction to the modern aspects of the topic.

In this regard, Patterson-Sullivan theory based on studying fractal geometric properties of PS measures supported by  $\Lambda_\Gamma$  provides a powerful method for establishing such results. It follows from this theory that  $R(\lambda)$  is well-defined and analytic on the domain  $\text{Im}(\lambda) > -\beta_{PS}$ , for  $\beta_{PS} = \max\{0, d/2 - \delta_\Gamma\}$ . Moreover,  $\lambda = -\beta_{PS}$  is the only pole for  $R(\lambda)$  on the line  $\lambda = -\beta_{PS}$ . The constant  $\beta_{PS}$  is referred to as the *Patterson-Sullivan (PS) gap*. We say that  $R(\lambda)$  has an *essential spectral gap* of size  $\beta \geq \beta_{PS}$  if  $R(\lambda)$  admits a meromorphic continuation with at most finitely many poles to the domain  $\text{Im}(\lambda) \geq -\beta$ .

The topic of producing essential spectral gaps and studying their dependence on geometric invariants of the underlying manifold has received considerable interest in recent years. When  $\delta_\Gamma \leq d/2$ , Naud [62] in the case  $d = 1$  and Stoyanov [84] in higher dimensions proved the existence of an essential spectral gap using a refinement of the Dolgopyat method [25]. In a major breakthrough, Bourgain and Dyatlov [15] produced an improvement over the PS gap in the regime  $\delta_\Gamma > d/2$  and  $d = 1$  using tools from harmonic analysis by proving a Fractal Uncertainty Principle for Ahlfors-David regular sets (see the definition (1.6) below). Extending the results of [15] to higher dimensions remains a major open problem, due to counterexamples arising from affinely concentrated sets [26]. See the work by Cohen [20] for a recent breakthrough on this question and e.g. related works [10, 7, 19, 34].

<sup>3</sup>The reference [23] considered the case of dimension 1, see Appendix A for a proof in dimension  $d \geq 2$ .

In [28], Dyatlov and Zahl introduced the notion of *Fractal Uncertainty Principle* (FUP) that provides a way to produce essential spectral gaps with explicit dependence on the geometry of  $M$ . We recall this notion here. Two sets  $X, Y \subset \mathbb{R}^d$  are said to satisfy the *generalised FUP* with exponent  $\beta > 0$  if for any open set  $U \subset \mathbb{R}^d \times \mathbb{R}^d$ , compact subset  $V \subset U$ , smooth functions  $\Phi \in C^3(U, \mathbb{R})$  and  $G \in C^1(U, \mathbb{C})$  satisfying  $\text{supp}(G) \subset V$  and  $\|\Phi\|_{C^3} + \|G\|_{C^1} \leq C_{\Phi, G}$ ,  $\inf |\partial_{xy}^2 \Phi| \geq C_{\Phi, G}^{-1}$ , for some constant  $C_{\Phi, G} > 0$ , then for any  $0 < \rho < 1$  and  $h \ll 1$ , we have

$$\|\mathbf{1}_{X(h\rho)} \mathcal{B}(h) \mathbf{1}_{Y(h\rho)}\|_{L^2(\mathbb{R}^d) \rightarrow L^2(\mathbb{R}^d)} \leq Ch^\beta, \quad (1.6)$$

where  $\mathcal{B}(h)$  is the Fourier integral operator defined by

$$\mathcal{B}(h)u(x) = \frac{1}{(2\pi h)^{d/2}} \int_{\mathbb{R}^d} \exp\left(\frac{i\Phi(x, y)}{h}\right) G(x, y) u(y) dy,$$

for all  $u \in L^2(\mathbb{R}^d)$  and  $x \in \mathbb{R}^d$ . In [28], it is shown that if the generalised FUP holds for  $X = Y = \Lambda_\Gamma$  with some exponent  $\beta > 0$ , then there is an essential spectral gap of size  $\beta > 0$ . In particular, it follows from their results that if  $\delta_\Gamma \leq d/2$  and  $\mu_\Gamma$  has polynomial Fourier decay with rate  $\kappa > 0$ , then the generalised FUP holds with exponent  $\beta = \beta_{PS} + \varepsilon$ , for a constant  $\varepsilon > 0$  depending explicitly on  $\kappa$ .

When  $d = 1$ , using Dolgopyat's method [25], Dyatlov and Jin [27] proved that  $\Lambda_\Gamma$  satisfies a generalized FUP with exponent  $\beta = \beta_{PS} + \varepsilon$  for a constant  $\varepsilon > 0$  depending on the Hausdorff dimension of the limit set  $\delta_\Gamma$  and the *Ahlfors-David regularity constant*  $C_\Gamma$  of the PS measure  $\mu_\Gamma$ . Here,  $C_\Gamma \geq 1$  is a constant satisfying

$$C_\Gamma^{-1} r^{\delta_\Gamma} \leq \mu_\Gamma(B(x, r)) \leq C_\Gamma r^{\delta_\Gamma}, \quad r > 0, x \in \text{supp}(\mu). \quad (1.7)$$

Dependence on such constants is a byproduct of the use of Dolgopyat's method.

To remove the  $C_\Gamma$  dependency, Bourgain and Dyatlov [14] proved that, in the case  $d = 1$ , the PS measures have polynomial Fourier decay with a rate depending only on the Hausdorff dimension  $\delta_\Gamma$ . When  $\delta_\Gamma \leq 1/2$ , this in particular produced an essential spectral gap of size independent of the constant  $C_\Gamma$ . Their proof used Bourgain's sum-product theorem as well as the non-linearity of the action of  $\Gamma$  on its limit set  $\Lambda_\Gamma$  by Möbius transformations. See also [53] for an extension of this result to the case of Schottky Kleinian groups in the case  $d = 2$ .

In higher dimensions, Backus, Leng and Tao [10] generalised the work of Dyatlov and Jin [27] using Dolgopyat's method [25]. They obtained an improvement over the PS gap that depends on the non-concentration properties of  $\mu_\Gamma$  as well as the *upper doubling constant*  $C_D$  of the PS measure. Here,  $C_D \geq 1$  is a constant such that

$$\mu_\Gamma(B(x, 2r)) \leq C_D \mu_\Gamma(B(x, r)), \quad r > 0, x \in \text{supp}(\mu).$$

Note that we can always choose  $C_D \leq 2^{\delta_\Gamma} C_\Gamma^2$ , where  $C_\Gamma$  is as in (1.7). See also [19] for related results by different methods. However, the question of producing an improvement that is independent of such constant à la Bourgain and Dyatlov remains open in higher dimensions.

Now, the advantage of Theorem 1.11 is that it relies on the  $L^2$ -flattening theorem and non-concentration estimates for PS measures while avoiding Dolgopyat's method. In particular, we are able to produce an improvement over the PS gap that is independent of the doubling constant  $C_D$ .

**Corollary 1.24.** *Let  $\Gamma$  be a discrete, Zariski-dense, convex cocompact, group of isometries of real hyperbolic space  $\mathbb{H}^{d+1}$ ,  $d \geq 1$ . Assume  $\delta_\Gamma \leq d/2$ . Then there exists  $\varepsilon > 0$  such that the limit set  $\Lambda_\Gamma$  satisfies the generalised Fractal Uncertainty Principle with exponent  $\beta = \frac{d}{2} - \delta_\Gamma + \varepsilon$  and that there exists an essential spectral gap of size  $\frac{d}{2} - \delta_\Gamma + \varepsilon$ . Here  $\varepsilon$  depends on the constants  $C$  and  $\alpha$  such that the Patterson-Sullivan measure  $\mu_\Gamma$  is  $(C, \alpha)$ -uniformly affinely non-concentrated, but not on the doubling constant  $C_D$ .*

**Remark 1.25.** Corollary 1.24 is a partial generalization of the results of Bourgain and Dyatlov to higher dimensions. A full generalization requires the removal of the dependence on the constant  $C$  coming from uniform non-concentration. This can be achieved by removing the dependence of the function  $\tau(\epsilon)$  in the flattening Theorem 1.3 on  $C$ .

**1.8. Generalizations and future directions.** We end the introduction with some discussions of possible generalizations and future directions this work could take.

**1.8.1. Lower regularity.** Firstly given the interest in quantum chaos problems in variable curvature [29], it would be interesting to obtain generalisations of our Fourier decay and FUP results for Patterson-Sullivan measures in this context, where one has to contend with the additional difficulty arising from lower regularity of the dynamics. In a similar vein, our proof of Theorem 1.14 is restricted to  $C^2$  IFSs, so it would also be interesting to have polynomial decay for  $C^{1+\alpha}$  self-conformal IFSs, possibly using methods of Tsujii-Zhang [89]. The problem of studying lower regularity also relates deeply to the UNI condition we use and going beyond the conformal category, which we discuss below.

**1.8.2. UNI condition and Julia sets in quasiregular dynamics.** Given that Liouville's theorem on the rigidity of conformal maps in dimensions  $d \geq 3$ , in developing analogues of complex dynamics in higher dimensions, the study of analogues of this theory for wider classes of maps known as quasiconformal- and quasiregular maps has received considerable interest; see e.g. [41, 60, 64, 45, 46] and references therein. In dimension  $d = 2$ , by a result of Oh and Winter [63], if  $f : \mathbb{S}^2 \rightarrow \mathbb{S}^2$  is a hyperbolic rational map, then  $\tau = \log |f'|$  satisfies a *Non-Local Integrability* property on the Julia set  $J_f$  (which is related to UNI, see [62, Proposition 5.5]) if and only if  $f$  is not conjugated to the power map  $z \mapsto z^d$  for any integer  $d \neq 0$ . This was adapted by Leclerc [51] to prove polynomial Fourier decay for equilibrium states on Julia sets for such rational hyperbolic maps not conjugated to  $z \mapsto z^d$ .

In dimensions  $d \geq 3$ , analogues of hyperbolic rational maps are given by *Uniformly Quasiregular (UQR) Maps*  $f : M \rightarrow M$  on a Riemannian manifold  $M$  [64, 40]. These maps are known to have a fractal Julia set only when  $M$  is a rational homology sphere [45]. In this situation in  $d = 3$ , the analogues for the power maps  $z \mapsto z^d$  come from certain Lattès map constructions [60]. In particular, it is natural to expect that maps not conjugated to such Lattès maps provide a potential source of Julia sets with equilibrium states having polynomial Fourier decay. However, the problem with UQR maps is that they no longer satisfy the smoothness properties, which would allow us to define e.g. the UNI condition using  $\partial_e \log |f'|$  as  $\log |f'|$  is not generally even Hölder. Similar complications also appear in the case of Thurston maps of  $\mathbb{S}^2$ , where Li and Zheng [57] generalised the work of Oh and Winter with a notion of  $\alpha$ -strong non-integrability condition allowing them to prove a spectral gap theorem. It would be an interesting problem to pursue these notions in dimensions  $d \geq 3$  for Julia sets of UQR maps in order to prove and classify Fourier decay for the equilibrium states for these systems.

**1.8.3. Non-conformal measures.** In a different direction, it would be interesting to extend our methods to obtain rates of decay for self-affine measures recovering and generalizing [55]. Similarly, it would be interesting to see if the  $L^2$  flattening approach could improve the results on polynomial decay for Furstenberg measures on projective spaces arising as stationary measures for random matrix products in the cases that have not yet been studied.

**1.8.4. Self-similar measures.** Given that we are able to get *polylogarithmic* Fourier decay for a wide class of self-similar measures, this raises the prospect for pushing the method to find more examples of *polynomial* decay for self-similar measures beyond the ones studied in the cases of Bernoulli convolutions, see [77] for further discussion of this topic.

1.8.5. *Random measures.* Finally, going beyond the dynamically defined category, there has recently been much activity in proving the Fourier decay properties of randomly defined measures such as spatially independent martingales [80], Liouville Quantum Gravity [31], random multiplicative cascades [33] and random Cantor measures in the study of Fractal Uncertainty Principles [35]. It would be interesting to explore the potential applicability of the methods of this article for studying the Fourier transforms of such randomly defined measures.

1.9. **Organisation of the article.** For convenience of the reader, the proofs of the results for different classes of measures appearing in the introduction are presented in such a way that they can be read independently of one another.

After collecting general preliminaries and notations in Section 2, we prove Theorems 1.6 and 1.9 in Section 3. Section 4 is dedicated to the case of Patterson-Sullivan measures where we prove Theorem 1.11. In Section 5, the general non-linear self-conformal case is established and, finally, in Section 6, we address the non-conformal case.

The reader will find that the overall structure of the proof in the case of self-similar, self-conformal, and non-conformal systems is very similar due to the presence of a natural underlying symbolic coding of the dynamics. In the case of PS measures, we do not rely on any coding and instead prove Theorem 1.11 directly using dynamics of the geodesic flow as a substitute for the shift. However, the strategy remains the same in all cases.

The article has four appendices establishing for completeness some results not readily available in the literature. In Appendix A, we prove a higher dimensional generalization of the Davenport-Erdős-LeVeque criterion for equidistribution which yields Corollary 1.23 when combined with our results on Fourier decay. Appendix B is dedicated to the proof of an auxiliary estimate on multinomial distributions required for the proof of the self-similar case. In Appendix C, we provide a proof of Proposition 1.16 on the spectral gap of twisted transfer operators. Finally, in Appendix D we prove Proposition D.1 which establishes equivalence of the UNI criterion we use to the one introduced by Li and Pan [54] in their work on exponential mixing.

**Acknowledgements.** S.B. is supported by an EPSRC New Investigator Award (EP/W003880/1). O.K. is partially supported under NSF grant DMS-2247713 and DMS-2337911. T.S. is supported by the Academy of Finland via the project *Quantum chaos of large and many body systems*, grant Nos. 347365, 353738. The proof in Appendix A was done by the third author and Jonathan Fraser in an unpublished preprint, and we thank Jonathan for allowing us to include it in the appendix. We also thank Jialun Li for comments on an earlier draft.

## 2. PRELIMINARIES

In this section, we collect some notation and technical results that we use throughout the article.

2.1. **Preliminaries for iterated function systems and thermodynamic formalism.** Given a finite set  $\mathcal{A}$  we let  $\mathcal{A}^* = \cup_{n=1}^{\infty} \mathcal{A}^n$  denote the corresponding set of finite words. Given  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}^*$  we let  $|\mathbf{a}|$  denote its length and let  $\mathbf{a}^- = (a_1, \dots, a_{n-1})$  denote  $\mathbf{a}$  with the last digit removed. Given  $\mathbf{a}, \mathbf{b} \in \mathcal{A}^*$  we let  $\mathbf{a} \wedge \mathbf{b}$  denote the maximal common prefix of  $\mathbf{a}$  and  $\mathbf{b}$ . We define  $\mathbf{a} \wedge \mathbf{b}$  analogously for  $\mathbf{a}, \mathbf{b} \in \mathcal{A}^{\mathbb{N}}$ .

Suppose now that we are given an IFS  $\{f_a\}_{a \in \mathcal{A}}$  then given  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}^*$  we let

$$f_{\mathbf{a}} = f_{a_1} \circ \dots \circ f_{a_n}.$$

Suppose in addition that our IFS consisted of similarities so  $f_a(x) = r_a O_a x + t_a$  for some  $|r_a| \in (0, 1)$ ,  $O_a \in O(d)$  and  $t_a \in \mathbb{R}^d$ , then given  $\mathbf{a} = (a_1, \dots, a_n) \in \mathcal{A}^*$  we let

$$r_{\mathbf{a}} = \prod_{i=1}^n r_{a_i} \quad \text{and} \quad O_{\mathbf{a}} = O_{a_1} \cdots O_{a_n}.$$

Similarly, suppose we are given a probability vector  $\mathbf{p} = (p_a)_{a \in \mathcal{A}}$  then we for any  $\mathbf{a} \in \mathcal{A}^*$  we let

$$p_{\mathbf{a}} = \prod_{i=1}^n p_{a_i}.$$

Suppose now that  $\{f_a\}_{a \in \mathcal{A}}$  is a  $C^2$  conformal IFS acting on  $[0, 1]^d$ . Then, there exists  $C > 0$  such that

$$\|D_x f_a - D_y f_a\| \leq C \|x - y\|$$

for all  $x, y \in [0, 1]^d$  and  $a \in \mathcal{A}$ . Moreover, we have the following stronger statement that follows from [6, Lemma 7.3].

**Lemma 2.1.** *Let  $\{f_a\}_{a \in \mathcal{A}}$  be a  $C^2$  self-conformal IFS acting on  $[0, 1]^d$ . Then there exists  $C' > 0$  such that*

$$\|D_x f_{\mathbf{a}} - D_y f_{\mathbf{a}}\| \leq C \sup_{z \in [0, 1]^d} \|D_z f_{\mathbf{a}}\| \|x - y\|$$

for all  $x, y \in [0, 1]^d$  and  $\mathbf{a} \in \mathcal{A}^*$ .

Lemma 2.1 will play a crucial role in this paper when we want to linearize a function  $f_{\mathbf{a}}$ . We also have the following well known result [67, Lemma 2.2]

**Lemma 2.2.** *Let  $\{f_a\}_{a \in \mathcal{A}}$  be a  $C^2$  self-conformal IFS acting on  $[0, 1]^d$ . Then there exists  $C' > 0$  such that for any  $\mathbf{a} \in \mathcal{A}^*$  we have*

$$C^{-1} \text{Diam}(f_{\mathbf{a}}([0, 1]^d)) \leq \|D_x f_{\mathbf{a}}\| \leq C \text{Diam}(f_{\mathbf{a}}([0, 1]^d))$$

for all  $x \in [0, 1]^d$ .

We also recall some useful properties of Gibbs measures. Suppose  $\psi : U \rightarrow \mathbb{R}$  is a  $C^1$  potential satisfying  $P(\psi) = 0$ . Under this assumption it can be shown that a Gibbs measure  $\mu_\psi$  satisfies the following invariance property:

$$\int g d\mu_\psi = \int \sum_{a \rightsquigarrow x} w_a(x) g(f_a(x)) d\mu_\psi \quad (2.1)$$

for all  $g : X_A \rightarrow \mathbb{C}$  continuous. For a proof of (2.1) in the symbolic setting we refer the reader to [65]. Using the Gibbs property (1.3), it can be shown that Gibbs measures satisfy the following quasi-Bernoulli property. Let  $\mathbf{a}, \mathbf{b} \in \mathcal{W}_A$  be such that  $\mathbf{ab} \in \mathcal{W}_A$ . Then, there exists  $C > 1$  such that

$$C^{-1} \mu_\psi(X_{\mathbf{a}}) \mu_\psi(X_{\mathbf{b}}) \leq \mu_\psi(X_{\mathbf{ab}}) \leq C \mu_\psi(X_{\mathbf{a}}) \mu_\psi(X_{\mathbf{b}}). \quad (2.2)$$

**2.2. Preliminaries on Patterson-Sullivan measures.** We collect here some preliminary facts needed for the proof of Theorem 1.11.

**2.2.1. Convex cocompact manifolds.** The standard reference for the material in this section is [16]. Let  $G$  denote the group of orientation preserving isometries of real hyperbolic space, denoted  $\mathbb{H}^{d+1}$ , of dimension  $d \geq 1$ . In particular,  $G \cong \text{SO}(d+1, 1)^0$ .

Fix a basepoint  $o \in \mathbb{H}^{d+1}$ . Then,  $G$  acts transitively on  $\mathbb{H}^{d+1}$  and the stabilizer  $K$  of  $o$  is a maximal compact subgroup of  $G$ . We shall identify  $\mathbb{H}^{d+1}$  with  $K \backslash G$ . Denote by  $A = \{g_t : t \in \mathbb{R}\}$  a one parameter subgroup of  $G$  inducing the geodesic flow on the unit tangent bundle of  $\mathbb{H}^{d+1}$ . Let  $M < K$  denote the centralizer of  $A$  inside  $K$ .

Let  $\Gamma < G$  be an infinite discrete subgroup of  $G$ . The limit set of  $\Gamma$ , denoted  $\Lambda_\Gamma$ , is the set of limit points of the orbit  $\Gamma \cdot o$  on  $\partial \mathbb{H}^{d+1}$ . Note that the discreteness of  $\Gamma$  implies that all such limit points belong to the boundary. Moreover, this definition is independent of the choice of  $o$  in view of the negative curvature of  $\mathbb{H}^{d+1}$ . We often use  $\Lambda$  to denote  $\Lambda_\Gamma$  when  $\Gamma$  is understood from context. We say  $\Gamma$  is *non-elementary* if  $\Lambda_\Gamma$  is infinite.

The *non-wandering set* for the geodesic flow, denoted by  $\Omega \subseteq G/\Gamma$ , is the closure of the set of periodic  $A$ -orbits. We say  $\Gamma$  is *convex cocompact* if  $\Omega$  is compact, cf. [16]. Denote by  $N^+$  (resp.  $N^-$ ) the expanding (resp. contracting) horospherical subgroup of  $G$  associated to  $g_t$ ,  $t \geq 0$ .

Given  $g \in G$ , we denote by  $g^+$  the coset of  $P^-g$  in the quotient  $P^- \backslash G$ , where  $P^- = N^-AM$  is the stable parabolic group associated to  $\{g_t : t \geq 0\}$ . Similarly,  $g^-$  denotes the coset  $P^+g$  in  $P^+ \backslash G$ . Since  $M$  is contained in  $P^\pm$ , such a definition makes sense for vectors in the unit tangent bundle  $M \backslash G$ . Geometrically, for  $v \in M \backslash G$ ,  $v^+$  (resp.  $v^-$ ) is the forward (resp. backward) endpoint of the geodesic determined by  $v$  on the boundary of  $\mathbb{H}^{d+1}$ . Given  $x \in G/\Gamma$ , we say  $x^\pm$  belongs to  $\Lambda$  if the same holds for any representative of  $x$  in  $G$ ; this notion being well-defined since  $\Lambda$  is  $\Gamma$  invariant.

**2.2.2. Patterson-Sullivan measures.** The *critical exponent*, denoted  $\delta_\Gamma$ , is defined to be the infimum over all real number  $s \geq 0$  such that the Poincaré series

$$P_\Gamma(s, o) := \sum_{\gamma \in \Gamma} e^{-sd(o, \gamma \cdot o)} \quad (2.3)$$

converges. This exponent coincides with the Hausdorff dimension of the limit set as well as the topological entropy of the geodesic flow on the quotient orbifold  $\mathbb{H}^{d+1}/\Gamma$ . We shall simply write  $\delta$  for  $\delta_\Gamma$  when  $\Gamma$  is understood from context.

The *Busemann function* is defined as follows: given  $x, y \in \mathbb{H}^{d+1}$  and  $\xi \in \partial\mathbb{H}^{d+1}$ , let  $\gamma : [0, \infty) \rightarrow \mathbb{H}^{d+1}$  denote a geodesic ray terminating at  $\xi$  and define

$$\beta_\xi(x, y) = \lim_{t \rightarrow \infty} \text{dist}(x, \gamma(t)) - \text{dist}(y, \gamma(t)).$$

A  $\Gamma$ -invariant conformal density of dimension  $s$  is a collection of Radon measures  $\{\nu_x\}$  on the boundary indexed by  $x \in \mathbb{H}^{d+1}$  which satisfy the following equivariance property:

$$\gamma_* \nu_x = \nu_{\gamma x}, \quad \text{and} \quad \frac{d\nu_y}{d\nu_x}(\xi) = e^{-s\beta_\xi(x, y)}, \quad \forall x, y \in \mathbb{H}^{d+1}, \xi \in \partial\mathbb{H}^{d+1}, \gamma \in \Gamma.$$

Patterson [66] and Sullivan [86] showed the existence of a unique (up to scaling)  $\Gamma$ -invariant conformal density of dimension  $\delta_\Gamma$ , denoted  $\{\mu_x^{\text{PS}} : x \in \mathbb{H}^{d+1}\}$ . These measures are known as the *Patterson-Sullivan measures* (PS measures for short). We refer the reader to [72] and [68] and references therein for details of the construction in much greater generality.

**2.2.3. Stable and unstable foliations and leafwise measures.** Recall that we fixed a basepoint  $o \in \mathbb{H}^{d+1}$ . In what follows, we use the following notation for pullbacks of the Patterson-Sullivan measures to orbits of  $N^+$  under the visual map: for  $x \in G/\Gamma$ ,

$$d\mu_x^u(n) = e^{\delta_\Gamma \beta_{(nx)^+}(o, nx)} d\mu_o^{\text{PS}}((nx)^+). \quad (2.4)$$

These measures have simpler transformation formulas under the action of the geodesic flow and  $N^+$  which makes them relatively easier to analyze than the Patterson-Sullivan measures directly. In particular, they satisfy the following equivariance property under the geodesic flow:

$$\mu_{g_t x}^u = e^{\delta t} \text{Ad}(g_t)_* \mu_x^u. \quad (2.5)$$

Moreover, it follows readily from the definitions that for all  $n \in N^+$ ,

$$(n)_* \mu_{nx}^u = \mu_x^u, \quad (2.6)$$

where  $(n)_* \mu_{nx}^u$  is the pushforward of  $\mu_{nx}^u$  under the map  $u \mapsto un$  from  $N^+$  to itself. Finally, since  $M$  normalizes  $N^+$ , these conditionals are  $\text{Ad}(M)$ -invariant in the sense that for all  $m \in M$ ,

$$\mu_{mx}^u = \text{Ad}(m)_* \mu_x^u. \quad (2.7)$$

**2.2.4. Norms, metrics, and Lie algebras.** We denote by  $\mathfrak{n}^+$  and  $\mathfrak{n}^-$  the Lie algebras of  $N^+$  and  $N^-$  respectively. We fix an isomorphism of  $\mathfrak{n}^+$  and  $\mathfrak{n}^-$  using a Cartan involution sending  $g_t$  to  $g_{-t}$ . Moreover, we fix an isomorphism of  $\mathfrak{n}^+$  (and hence of  $\mathfrak{n}^-$ ) with  $\mathbb{R}^d$ . Finally, we fix a Euclidean inner product on  $\mathbb{R}^d \cong \mathfrak{n}^+ \cong \mathfrak{n}^-$  denoted with  $\langle \cdot, \cdot \rangle$  which is invariant by the Adjoint action of the group  $M \cong \mathrm{SO}_d(\mathbb{R})$  and induces the respective metrics on the groups  $N^+$  and  $N^-$ . Given  $r > 0$ , we let  $N_r^+$  (resp.  $P_r^-$ ) the neighborhood of identity of radius  $r$  inside  $N^+$  (resp.  $P^- := MAN^-$ ).

**2.2.5. Local stable holonomy.** In this section, we recall the definition of (stable) holonomy maps which are essential for our arguments. We give a simplified discussion of this topic which is sufficient in our homogeneous setting. Let  $x = u^-y$  for some  $y \in \Omega$  and  $u^- \in N_2^-$ . Since the product map  $N^- \times A \times M \times N^+ \rightarrow G$  is a diffeomorphism near the identity, we can choose the norm on the Lie algebra so that the following holds. We can find maps  $p^- : N_1^+ \rightarrow P^- = N^-AM$  and  $u^+ : N_2^+ \rightarrow N^+$  so that

$$nu^- = p^-(n)u^+(n), \quad \forall n \in N_2^+. \quad (2.8)$$

Then, it follows by (2.4) that for all  $n \in N_2^+$ , we have

$$d\mu_y^u(u^+(n)) = e^{\delta\beta_{(nx)} + (u^+(n)y, nx)} d\mu_x^u(n).$$

Moreover, by further scaling the metrics if necessary, we can ensure that these maps are diffeomorphisms onto their images. In particular, writing  $\Phi(nx) = u^+(n)y$ , we obtain the following change of variables formula: for all  $f \in C(N_2^+)$ ,

$$\int f(n) d\mu_x^u(n) = \int f((u^+)^{-1}(n)) e^{-\delta\beta_{\Phi^{-1}(ny)} + (ny, \Phi^{-1}(ny))} d\mu_y^u(n). \quad (2.9)$$

**Remark 2.3.** To avoid cluttering the notation with auxiliary constants, we shall assume that the  $N^-$  component of  $p^-(n)$  belongs to  $N_2^-$  for all  $n \in N_2^+$  whenever  $u^-$  belongs to  $N_1^-$ .

**2.3. Notational convention.** Throughout this article, given two quantities  $A$  and  $B$  we will write  $A \ll B$  if there exists a constant  $C > 0$  such that  $A \leq CB$ . When we want to emphasize that  $C$  depends on another quantity, say  $x$ , we write  $A \ll_x B$ . We also write  $A = \mathcal{O}(B)$  to mean  $A \ll B$  and  $A = \mathcal{O}_x(B)$  to mean  $A \ll_x B$ . When  $A \ll B$  and  $B \ll A$  then we write  $A \asymp B$ .

### 3. SELF-SIMILAR MEASURES: PROOF OF THEOREMS 1.6 AND 1.9

**3.1. Sketch of the proof.** Before giving the proof of Theorems 1.6 and 1.9, we provide a sketch of the argument. The ideas underpinning this argument are also used in our later proofs. Suppose that our IFS is acting on  $\mathbb{R}$  and consists of two maps  $\{f_a(x) = r_ax + t_a, f_b(x) = r_bx + t_b\}$ . Moreover suppose that  $C, l > 0$  are such that

$$\left| \frac{\log |r_a|}{\log |r_b|} - \frac{p}{q} \right| \geq \frac{C}{q^l}$$

for all  $(p, q) \in \mathbb{Z} \times \mathbb{N}$ . Let  $\xi \in \mathbb{R}$  be given. We define the following cut-off set

$$\mathcal{A}_\xi := \left\{ \mathbf{a} \in \mathcal{A}^* : \prod_{j=1}^{|\mathbf{a}|} |r_{a_j}| < \frac{(\log \|\xi\|)^{3l}}{\|\xi\|} \text{ and } \prod_{j=1}^{|\mathbf{a}|-1} |r_{a_j}| \geq \frac{(\log \|\xi\|)^{3l}}{\|\xi\|} \right\}.$$

Appealing to the self-similarity of  $\mu$  it can be shown that

$$|\widehat{\mu}(\xi)| \leq \sum_{\mathbf{a} \in \mathcal{A}_\xi} p_{\mathbf{a}} |\widehat{\mu}(r_{\mathbf{a}}\xi)| \quad (3.1)$$

The significance of (3.1) is that it bounds  $|\widehat{\mu}(\xi)|$  from above by an average of  $|\widehat{\mu}(\cdot)|$  evaluated at different frequencies.

At this point, we can employ our Diophantine assumption to show that  $|r_{\mathbf{a}}\xi - r_{\mathbf{b}}\xi| \leq 1$  if and only if  $r_{\mathbf{a}} = r_{\mathbf{b}}$ . This implies that the frequencies we are averaging over in (3.1) form a large well

separated set. It is a consequence of this well-separated property and our assumption that  $\mu$  is  $L^2$ -flat that, for most  $\mathbf{a} \in \mathcal{A}_\xi$ , the term  $|\widehat{\mu}(r_{\mathbf{a}}\xi)|$  decays polylogarithmically in  $|\xi|$ . The contribution to  $\sum_{\mathbf{a} \in \mathcal{A}_\xi} p_{\mathbf{a}} |\widehat{\mu}(r_{\mathbf{a}}\xi)|$  coming from those remaining terms can be bounded by a probabilistic argument.

What makes this argument more challenging in the full generality of Theorem 1.6 is that to meaningfully apply the Diophantine assumption to bound the distance between  $r_{\mathbf{a}}\|\xi\|$  and  $r_{\mathbf{b}}\|\xi\|$  for some  $\xi \in \mathbb{R}^d$ , we require the products determining  $r_{\mathbf{a}}$  and  $r_{\mathbf{b}}$  to have the same number of contractions coming from digits belonging to  $\mathcal{A} \setminus \{a, b\}$ . This issue is overcome by a suitable conditioning argument.

**3.2. Proof of Theorem 1.6.** We fix  $\xi$  such that  $\|\xi\| \gg 1$ . Let  $a_1, a_2 \in \mathcal{A}$  be such that  $\frac{\log |r_{a_1}|}{\log |r_{a_2}|}$  is Diophantine. Let  $l \geq 2$  and  $C > 0$  be the corresponding parameters such that

$$\left| \frac{\log |r_{a_1}|}{\log |r_{a_2}|} - \frac{p}{q} \right| \geq \frac{C}{q^l}$$

for all  $(p, q) \in \mathbb{Z} \times \mathbb{N}$ . We define

$$\tilde{\xi} = \frac{(\log \|\xi\|)^{3l}}{\|\xi\|}.$$

Given a word  $\mathbf{a} \in \mathcal{A}^*$  and  $a \in \mathcal{A}$  we let

$$|\mathbf{a}|_a = \#\{1 \leq j \leq |\mathbf{a}| : a_j = a\}.$$

At this point we fix  $\delta, \epsilon > 0$  to both be sufficiently small that

$$\epsilon - 1/2(\#\mathcal{A} - 1) + (1/2 + \delta)(\#\mathcal{A} - 2) < 0. \quad (3.2)$$

which allow us to define

$$n(\xi) := \left\lfloor \left(1 - \frac{1}{(-\log \tilde{\xi})^{1/2-\delta}}\right) \frac{-\log \tilde{\xi}}{-\sum_{a \in \mathcal{A}} p_a \log |r_a|} \right\rfloor \quad \text{and} \quad E_a(\xi) := \frac{(-\log \tilde{\xi})^{1/2+\delta}}{-\#\mathcal{A} \log |r_a|}, \quad a \in \mathcal{A}.$$

We say that a word  $\mathbf{a}$  is *good* if  $\mathbf{a} \in \mathcal{A}^{n(\xi)}$  and for all  $a \in \mathcal{A}$  we have the following strong upper and lower bounds for  $|\mathbf{a}|_a$ :

$$p_a n(\xi) - E_a(\xi) \leq |\mathbf{a}|_a \leq p_a n(\xi) + E_a(\xi).$$

Let  $G_\xi$  be the set of good words. The following lemma records some useful properties of good words.

**Lemma 3.1.** *Let  $\mathbf{a} \in G_\xi$ . Then*

$$\tilde{\xi} \leq |r_{\mathbf{a}}| \ll \tilde{\xi} e^{2(\log \|\xi\|)^{1/2+\delta}}$$

*and there exists  $C_1, C_2 > 0$  such that*

$$\sum_{\mathbf{a} \in \mathcal{A}^{n(\xi)} : \mathbf{a} \notin G_\xi} p_{\mathbf{a}} \leq C_1 e^{-C_2(\log \|\xi\|)^{2\delta}}$$

*Proof.* The proof is a direct calculation using the definitions. Let  $\mathbf{a} \in G_\xi$ . By the definition of  $G_\xi$  we have

$$\begin{aligned} \log |r_{\mathbf{a}}| &= \sum_{a \in \mathcal{A}} \log |r_a|^{|\mathbf{a}|_a} \geq \sum_{a \in \mathcal{A}} (p_a n(\xi) + E_a(\xi)) \log |r_a| \\ &\geq \sum_{a \in \mathcal{A}} \log |r_a| \left( p_a \left(1 - \frac{1}{(-\log \tilde{\xi})^{1/2-\delta}}\right) \frac{-\log \tilde{\xi}}{-\sum_{a \in \mathcal{A}} p_a \log |r_a|} + E_a(\xi) \right) \\ &= \log \tilde{\xi} \left(1 - \frac{1}{(-\log \tilde{\xi})^{1/2-\delta}}\right) - (-\log \tilde{\xi})^{1/2+\delta} \\ &= \log \tilde{\xi}. \end{aligned}$$

This establishes the first property. For the second property, we similarly have

$$\begin{aligned}
|r_{\mathbf{a}}| &= \prod_{a \in \mathcal{A}} |r_a|^{|a|_a} \leq \prod_{a \in \mathcal{A}} |r_a|^{p_a n(\xi) - E_a(\xi)} \\
&\ll \exp \left( \sum_{a \in \mathcal{A}} \log |r_a| \left( p_a \left( 1 - \frac{1}{(-\log \tilde{\xi})^{1/2-\delta}} \right) - \frac{-\log \tilde{\xi}}{\sum_{a \in \mathcal{A}} p_a \log |r_a|} - E_a(\xi) \right) \right) \\
&= \exp \left( \log \tilde{\xi} \left( 1 - \frac{1}{(-\log \tilde{\xi})^{1/2-\delta}} \right) + (-\log \tilde{\xi})^{1/2+\delta} \right) \\
&= \exp \left( \log \tilde{\xi} + 2(-\log \tilde{\xi})^{1/2+\delta} \right).
\end{aligned}$$

Recalling the definition of  $\tilde{\xi}$ , we have that  $-\log \tilde{\xi} \leq \log \|\xi\|$ . Combining this with the above implies

$$|r_{\mathbf{a}}| \ll \tilde{\xi} e^{2(\log \|\xi\|)^{1/2+\delta}}.$$

The final statement in this lemma follows from an application of Hoeffding's inequality [38].  $\square$

We consider the following cut off set

$$\mathcal{A}_\xi := \left\{ \mathbf{a} \in \mathcal{A}^* : |r_{\mathbf{a}}| < \tilde{\xi} \text{ and } |r_{\mathbf{a}-}| \geq \tilde{\xi} \right\}, \quad (3.3)$$

and its subset

$$\tilde{\mathcal{A}}_\xi := \left\{ \mathbf{a} \in \mathcal{A}_\xi : a_1^{n(\xi)} \in G_\xi \right\}. \quad (3.4)$$

**Lemma 3.2.** *The following properties hold:*

- If  $\mathbf{a} \in \tilde{\mathcal{A}}_\xi$  then for all  $a \in \mathcal{A}$  we have

$$|a|_a \geq p_a n(\xi) - E_a(\xi)$$

and there exists  $C > 0$  such that for all  $a \in \mathcal{A}$  we have

$$|a|_a \leq p_a n(\xi) + E_a(\xi) + C(\log \|\xi\|)^{1/2+\delta}.$$

- There exists  $C_1, C_2 > 0$  such that

$$\sum_{\mathbf{a} \in \mathcal{A}_\xi : \mathbf{a} \notin \tilde{\mathcal{A}}_\xi} p_{\mathbf{a}} \leq C_1 e^{-C_2(\log \|\xi\|)^{2\delta}}.$$

*Proof.* Let  $\mathbf{a} \in \tilde{\mathcal{A}}_\xi$ . Then, the lower bound for  $|a|_a$  follows from the corresponding lower bound in the definition of a good word. The upper bound for  $|a|_a$  follows from the corresponding upper bound in the definition of a good word together with the fact that if  $\mathbf{a} \in \tilde{\mathcal{A}}_\xi$  then

$$|\mathbf{a}| - |\mathbf{a}|_1^{n(\xi)} \ll (\log \|\xi\|)^{1/2+\delta}. \quad (3.5)$$

(3.5) follows since if  $\mathbf{b} \in G_\xi$ , then for  $C$  sufficiently large we have

$$|r_{\mathbf{b}}| \left( \max_{a \in \mathcal{A}} \{|r_a|\} \right)^{C(\log \|\xi\|)^{1/2+\delta}} \ll \tilde{\xi} e^{2(\log \|\xi\|)^{1/2+\delta}} \left( \max_{a \in \mathcal{A}} \{|r_a|\} \right)^{C(\log \|\xi\|)^{1/2+\delta}} < \tilde{\xi},$$

where in the first inequality we have used the second property in Lemma 3.1.

To prove the final part of this lemma we begin by remarking that by Lemma 3.1 we know that if  $\mathbf{a} \in G_\xi$  we have  $r_{\mathbf{a}} \geq \tilde{\xi}$ . Thus all of the descendants of  $\mathbf{a}$  belonging to  $\mathcal{A}_\xi$  are elements of  $\tilde{\mathcal{A}}_\xi$ . Therefore

$$\sum_{\mathbf{a} \in \mathcal{A}_\xi : \mathbf{a} \notin \tilde{\mathcal{A}}_\xi} p_{\mathbf{a}} = \sum_{\mathbf{a} \in \mathcal{A}^{n(\xi)} : \mathbf{a} \notin G_\xi} p_{\mathbf{a}}.$$

Our upper bound for  $\sum_{\mathbf{a} \in \mathcal{A}_\xi : \mathbf{a} \notin \tilde{\mathcal{A}}_\xi} p_{\mathbf{a}}$  therefore follows from the corresponding bound in Lemma 3.1.  $\square$

We are now at a point where we want to use our Diophantine assumption. To use this assumption effectively we need to condition on the number of occurrences of a digit  $a \in \mathcal{A} \setminus \{a_1, a_2\}$ . With this goal in mind, we let

$$K = \left\{ (|\mathbf{a}|_a)_{a \in \mathcal{A} \setminus \{a_1, a_2\}} : \mathbf{a} \in \tilde{\mathcal{A}}_\xi \right\} \subset \mathbb{N}^{\#\mathcal{A}-2}. \quad (3.6)$$

It follows from Lemma 3.2 and the definition of  $\tilde{\xi}$  that if  $\mathbf{a} \in \tilde{\mathcal{A}}_\xi$ , then  $|\mathbf{a}|_a$  belongs to an interval of length at most a constant multiple of  $(\log \|\xi\|)^{1/2+\delta}$  for any  $a \in \mathcal{A}$ . Therefore we have the important bound

$$\#K \ll (\log \|\xi\|)^{(\#\mathcal{A}-2)(1/2+\delta)}. \quad (3.7)$$

The following lemma shows that if we condition on an element of  $K$  then the contraction ratios will be well separated. This is the only place where we use our Diophantine assumption.

**Lemma 3.3.** *Let  $\mathbf{a}, \mathbf{a}' \in \tilde{\mathcal{A}}_\xi$  be such that  $|\mathbf{a}|_a = |\mathbf{a}'|_a$  for all  $a \in \mathcal{A} \setminus \{a_1, a_2\}$ , and suppose that either  $|\mathbf{a}|_{a_1} \neq |\mathbf{a}'|_{a_1}$  or  $|\mathbf{a}|_{a_2} \neq |\mathbf{a}'|_{a_2}$ , then*

$$||r_{\mathbf{a}}||\xi|| - |r_{\mathbf{a}'}||\xi|| > 1$$

for  $\|\xi\|$  sufficiently large.

*Proof.* For convenience, we let

$$\varrho = (|\mathbf{a}'|_{a_1} - |\mathbf{a}|_{a_1}) \log |r_1| + (|\mathbf{a}'|_{a_2} - |\mathbf{a}|_{a_2}) \log |r_2|.$$

Then, we have the following

$$||r_{\mathbf{a}}||\xi|| - |r_{\mathbf{a}'}||\xi|| = |r_{\mathbf{a}}||\xi|| \left| 1 - \frac{|r_{\mathbf{a}'}|}{|r_{\mathbf{a}}|} \right| = |r_{\mathbf{a}}||\xi|| |1 - e^\varrho|. \quad (3.8)$$

It also follows from the definition of  $\mathcal{A}_\xi$  and  $\tilde{\xi}$  that

$$|r_{\mathbf{a}}||\xi|| \gg (\log \|\xi\|)^{3l}. \quad (3.9)$$

Using our Diophantine assumption we have

$$\begin{aligned} |1 - e^\varrho| &\geq |\varrho| = (|\mathbf{a}'|_{a_1} - |\mathbf{a}|_{a_1}) \log |r_2| \left| \frac{\log |r_1|}{\log |r_2|} - \frac{(|\mathbf{a}'|_{a_2} - |\mathbf{a}|_{a_2})}{(|\mathbf{a}'|_{a_1} - |\mathbf{a}|_{a_1})} \right| \\ &\gg \frac{1}{(|\mathbf{a}'|_{a_1} - |\mathbf{a}|_{a_1})^l} \\ &\gg \frac{1}{(\log \|\xi\|)^l}. \end{aligned} \quad (3.10)$$

In the final line we used that if  $\mathbf{a} \in \mathcal{A}_\xi$ , then we must have  $|\mathbf{a}| \ll \log \|\xi\|$ . Substituting (3.9) and (3.10) into (3.8), we obtain

$$||r_{\mathbf{a}}||\xi|| - |r_{\mathbf{a}'}||\xi|| \gg \frac{(\log \|\xi\|)^{3l}}{(\log \|\xi\|)^l}.$$

Our result follows.  $\square$

Before proceeding with the proof of our theorem we need the following upper bound for probabilities coming from a multinomial distribution.

**Lemma 3.4.** *Let  $(p_a)_{a \in \mathcal{A}}$  be a probability vector and  $n \in \mathbb{N}$ , then*

$$\frac{n!}{\prod_{a \in \mathcal{A}} k_a!} \prod_{a \in \mathcal{A}} p_a^{k_a} \ll n^{-1/2(\#\mathcal{A}-1)}$$

for any  $(k_a)_{a \in \mathcal{A}} \in \mathbb{N}^{\#\mathcal{A}}$  satisfying  $\sum_{a \in \mathcal{A}} k_a = n$ . Here the underlying constants only depend upon the probability vector.

The proof of Lemma 3.4 is straightforward but somewhat lengthy. Thus for the purpose of our exposition, we have deferred this argument to Appendix B. Equipped with Lemma 3.4 we can now complete our proof of Theorem 1.6.

*Proof of Theorem 1.6.* Using the self-similarity of the measure and Lemma 3.2 we have

$$\begin{aligned} |\widehat{\mu}(\xi)| &= \left| \int e^{2\pi i \langle \xi, x \rangle} d\mu \right| = \left| \sum_{\mathbf{a} \in \mathcal{A}_\xi} p_{\mathbf{a}} \int e^{2\pi i \langle \xi, r_{\mathbf{a}} O_{\mathbf{a}} x + t_{\mathbf{a}} \rangle} d\mu \right| \\ &\leq \sum_{\mathbf{a} \in \tilde{\mathcal{A}}_\xi} p_{\mathbf{a}} |\widehat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| + C_1 e^{-C_2 (\log \|\xi\|)^{2\delta}}. \end{aligned}$$

Thus to prove our theorem it suffices to show that

$$\sum_{\mathbf{a} \in \tilde{\mathcal{A}}_\xi} p_{\mathbf{a}} |\widehat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| \ll (\log \|\xi\|)^{-\nu}, \quad (3.11)$$

for some  $\nu > 0$ . At this point, we apply our polynomially decaying on average assumption for  $T = (\log \|\xi\|)^{3l}$ . In particular, for our fixed value of  $\epsilon$  satisfying (3.2), there exists  $\tau > 0$  such that the set

$$\left\{ \zeta \in \mathbb{R}^d : \|\zeta\| \leq (\log \|\xi\|)^{3l} \text{ and } |\widehat{\mu}(\zeta)| \geq (\log \|\xi\|)^{-\tau} \right\}$$

can be covered by  $\mathcal{O}_\epsilon((\log \|\xi\|)^\epsilon)$  balls of length 1. Taking the modulus of those frequencies  $\zeta$  satisfying  $|\widehat{\mu}(\zeta)| \geq (\log \|\xi\|)^{-\tau}$  given us the following more useful bound. Let

$$\text{Bad} := \left\{ n \in \mathbb{N} : \exists \zeta \in \mathbb{R}^d \text{ s.t. } \|\zeta\| \leq (\log \|\xi\|)^{3l}, |\widehat{\mu}(\zeta)| \geq (\log \|\xi\|)^{-\tau} \text{ and } \|\zeta\| \in [n, n+1] \right\}. \quad (3.12)$$

Then,  $\#\text{Bad} = \mathcal{O}_\epsilon((\log \|\xi\|)^\epsilon)$ .

To proceed, we use a conditioning argument on elements of the set  $K$  introduced in (3.6). To this end, given  $\bar{k} \in K$  and  $a \in \mathcal{A} \setminus \{a_1, a_2\}$ , we let  $k_a$  denote the component of  $\bar{k}$  corresponding to the letter  $a$ . For simplicity, let us use Good to denote the set  $(\cup_{n \in \text{Bad}} [n, n+1])^c$ . Hence, using the

definition of Bad, we obtain

$$\begin{aligned}
\sum_{\mathbf{a} \in \tilde{\mathcal{A}}_\xi} p_{\mathbf{a}} |\widehat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| &= \sum_{\bar{k} \in K} \sum_{\substack{\mathbf{a} \in \tilde{\mathcal{A}}_\xi \\ |\mathbf{a}|_a = k_a \forall a \in \mathcal{A} \setminus \{a_1, a_2\}}} p_{\mathbf{a}} |\widehat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| \\
&= \sum_{n \in \text{Bad}} \sum_{\bar{k} \in K} \sum_{\substack{\mathbf{a} \in \tilde{\mathcal{A}}_\xi \\ |\mathbf{a}|_a = k_a \forall a \in \mathcal{A} \setminus \{a_1, a_2\}}} p_{\mathbf{a}} \chi_{[n, n+1]}(|r_{\mathbf{a}}| \|\xi\|) |\widehat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| \\
&\quad + \sum_{\bar{k} \in K} \sum_{\substack{\mathbf{a} \in \tilde{\mathcal{A}}_\xi \\ |\mathbf{a}|_a = k_a \forall a \in \mathcal{A} \setminus \{a_1, a_2\}}} p_{\mathbf{a}} \chi_{\text{Good}}(|r_{\mathbf{a}}| \|\xi\|) |\widehat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| \\
&\leq \sum_{n \in \text{Bad}} \sum_{\bar{k} \in K} \sum_{\substack{\mathbf{a} \in \tilde{\mathcal{A}}_\xi \\ |\mathbf{a}|_a = k_a \forall a \in \mathcal{A} \setminus \{a_1, a_2\}}} p_{\mathbf{a}} \chi_{[n, n+1]}(|r_{\mathbf{a}}| \|\xi\|) |\widehat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| \\
&\quad + \sum_{\bar{k} \in K} \sum_{\substack{\mathbf{a} \in \tilde{\mathcal{A}}_\xi \\ |\mathbf{a}|_a = k_a \forall a \in \mathcal{A} \setminus \{a_1, a_2\}}} p_{\mathbf{a}} (\log \|\xi\|)^{-\tau} \\
&\leq \sum_{n \in \text{Bad}} \sum_{\bar{k} \in K} \sum_{\substack{\mathbf{a} \in \tilde{\mathcal{A}}_\xi \\ |\mathbf{a}|_a = k_a \forall a \in \mathcal{A} \setminus \{a_1, a_2\}}} p_{\mathbf{a}} \chi_{[n, n+1]}(|r_{\mathbf{a}}| \|\xi\|) |\widehat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| + (\log \|\xi\|)^{-\tau}.
\end{aligned}$$

Thus, our proof is complete if we can show that

$$\sum_{n \in \text{Bad}} \sum_{\bar{k} \in K} \sum_{\substack{\mathbf{a} \in \tilde{\mathcal{A}}_\xi \\ |\mathbf{a}|_a = k_a \forall a \in \mathcal{A} \setminus \{a_1, a_2\}}} p_{\mathbf{a}} \chi_{[n, n+1]}(|r_{\mathbf{a}}| \|\xi\|) |\widehat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| \ll (\log(\|\xi\|))^{-\nu}, \quad (3.13)$$

for some  $\nu > 0$ .

The crucial step towards verifying (3.13) is to note that by Lemma 3.3, for any  $n \in \mathbb{N}$  and  $\bar{k} \in K$ , if  $\mathbf{a}, \mathbf{b} \in \tilde{\mathcal{A}}_\xi$  satisfy  $|\mathbf{a}|_a = |\mathbf{b}|_a$  for all  $a \in \mathcal{A} \setminus \{a_1, a_2\}$  and  $|r_{\mathbf{a}}| \|\xi\|, |r_{\mathbf{b}}| \|\xi\| \in [n, n+1]$ , then  $|\mathbf{a}|_{a_1} = |\mathbf{b}|_{a_1}$  and  $|\mathbf{a}|_{a_2} = |\mathbf{b}|_{a_2}$ . Assuming such words exist, we denote the unique values for  $|\mathbf{a}|_{a_1}$  and  $|\mathbf{a}|_{a_2}$  by  $j_1(\bar{k}, n)$  and  $j_2(\bar{k}, n)$  respectively. Using this observation, we have

$$\begin{aligned}
&\sum_{n \in \text{Bad}} \sum_{\bar{k} \in K} \sum_{\substack{\mathbf{a} \in \tilde{\mathcal{A}}_\xi \\ |\mathbf{a}|_a = k_a \forall a \in \mathcal{A} \setminus \{a_1, a_2\}}} p_{\mathbf{a}} \chi_{[n, n+1]}(|r_{\mathbf{a}}| \|\xi\|) |\widehat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| \\
&\leq \sum_{n \in \text{Bad}} \sum_{\bar{k} \in K} \sum_{\substack{\mathbf{a} \in \tilde{\mathcal{A}}_\xi \\ |\mathbf{a}|_a = k_a \forall a \in \mathcal{A} \setminus \{a_1, a_2\} \\ |\mathbf{a}|_{a_1} = j_1(\bar{k}, n), |\mathbf{a}|_{a_2} = j_2(\bar{k}, n)}} p_{\mathbf{a}} \\
&= \sum_{n \in \text{Bad}} \sum_{\bar{k} \in K} \frac{\left( \sum_{a \in \mathcal{A} \setminus \{a_1, a_2\}} k_a + j_1(\bar{k}, n) + j_2(\bar{k}, n) \right)!}{\prod_{a \in \mathcal{A} \setminus \{a_1, a_2\}} k_a! j_1(\bar{k}, n)! j_2(\bar{k}, n)!} p_{a_1}^{j_1(\bar{k}, n)} p_{a_2}^{j_2(\bar{k}, n)} \prod_{a \in \mathcal{A} \setminus \{a_1, a_2\}} p_a^{k_a}.
\end{aligned}$$

Since  $\sum_{a \in \mathcal{A} \setminus \{a_1, a_2\}} k_a + j_1(\bar{k}, n) + j_2(\bar{k}, n)$  is the length of a word in  $\tilde{\mathcal{A}}_\xi$ , it follows that

$$\sum_{a \in \mathcal{A} \setminus \{a_1, a_2\}} k_a + j_1(\bar{k}, n) + j_2(\bar{k}, n) \asymp \log \|\xi\|.$$

Using this fact together with (3.7), Lemma 3.4, and the fact  $\#\text{Bad} = \mathcal{O}_\epsilon((\log \|\xi\|)^\epsilon)$  we have

$$\begin{aligned}
& \sum_{n \in \text{Bad}} \sum_{\bar{k} \in K} \sum_{\substack{\mathbf{a} \in \tilde{\mathcal{A}}_\xi \\ |\mathbf{a}|_a = k_a \forall a \in \mathcal{A} \setminus \{a_1, a_2\}}} p_{\mathbf{a}} \chi_{[n, n+1]}(r_a \|\xi\|) |\hat{\mu}(r_{\mathbf{a}} O_{\mathbf{a}}^T \xi)| \\
& \ll \sum_{n \in \text{Bad}} \sum_{\bar{k} \in K} (\log \|\xi\|)^{-1/2(\#\mathcal{A}-1)} \\
& \ll \sum_{n \in \text{Bad}} (\log \|\xi\|)^{-1/2(\#\mathcal{A}-1) + (1/2+\delta)(\#\mathcal{A}-2)} \\
& \ll (\log \|\xi\|)^{\epsilon - 1/2(\#\mathcal{A}-1) + (1/2+\delta)(\#\mathcal{A}-2)}
\end{aligned}$$

Recall now that by (3.2) we fixed  $\epsilon, \delta > 0$  to be sufficiently small so that the exponent appearing in the last line in the above is negative. Thus (3.13) holds and our proof is complete.  $\square$

**3.3. Proof of Theorem 1.9.** The proof of Theorem 1.9 is almost identical to the proof of Theorem 1.6 so we only indicate the small changes required.

Let  $a_1, a_2 \in \mathcal{A}$  be such that  $(\theta_{a_1}, \theta_{a_2})$  is Diophantine. We then fix  $\xi \in \mathbb{R}^d$ . We define  $\tilde{\xi}, G_\xi, \mathcal{A}_\xi$  and  $\tilde{\mathcal{A}}_\xi$  as above. The proof of Theorem 1.9 proceeds in an identical way to the proof of Theorem 1.6 until we get to Lemma 3.3. We have the following analogue of this lemma.

**Lemma 3.5.** *Let  $\{f_a(x) = r_a O_a x + t_a\}_{a \in \mathcal{A}}$  be an IFS acting on  $\mathbb{R}^2$  satisfying the assumptions of Theorem 1.9. Let  $\mathbf{a}, \mathbf{a}' \in \tilde{\mathcal{A}}_\xi$  be such that  $|\mathbf{a}|_a = |\mathbf{a}'|_a$  for all  $a \in \mathcal{A} \setminus \{a_1, a_2\}$  and suppose that either  $|\mathbf{a}|_{a_1} \neq |\mathbf{a}'|_{a_1}$  or  $|\mathbf{a}|_{a_2} \neq |\mathbf{a}'|_{a_2}$  then*

$$\|r_{\mathbf{a}} O_{\mathbf{a}}^T \xi - r_{\mathbf{a}'} O_{\mathbf{a}'}^T \xi\| > 1$$

for  $\|\xi\|$  sufficiently large.

*Proof.* Let  $\mathbf{a}, \mathbf{a}'$  satisfy the assumptions of the lemma. Iterating our self-similar IFS if necessary, we can assume without loss of generality that  $r_{a_1} > 0$  and  $r_{a_2} > 0$ . Using the fact that elements of  $SO(2)$  commute and are distance preserving we have

$$\left\| O_{\mathbf{a}}^T \frac{\xi}{\|\xi\|} - O_{\mathbf{a}'}^T \frac{\xi}{\|\xi\|} \right\| = \left\| (O_{a_1}^T)^{|\mathbf{a}|_{a_1}} (O_{a_2}^T)^{|\mathbf{a}|_{a_2}} e - (O_{a_1}^T)^{|\mathbf{a}'|_{a_1}} (O_{a_2}^T)^{|\mathbf{a}'|_{a_2}} e \right\|$$

where  $e = (1, 0)$ . Focusing on this latter expression, it follows from our Diophantine assumption and the fact  $|\mathbf{a}|, |\mathbf{a}'| \ll \log \|\xi\|$  that

$$\begin{aligned}
\left\| (O_{a_1}^T)^{|\mathbf{a}|_{a_1}} (O_{a_2}^T)^{|\mathbf{a}|_{a_2}} e - (O_{a_1}^T)^{|\mathbf{a}'|_{a_1}} (O_{a_2}^T)^{|\mathbf{a}'|_{a_2}} e \right\| & \gg d((|\mathbf{a}|_{a_1} - |\mathbf{a}'|_{a_1})\theta_1 + (|\mathbf{a}|_{a_2} - |\mathbf{a}'|_{a_2})\theta_2, \mathbb{Z}) \\
& \gg \frac{1}{(\log \|\xi\|)^l}.
\end{aligned}$$

Combining the bound above with the bounds  $|r_{\mathbf{a}}| \|\xi\| \gg (\log \|\xi\|)^{3l}$  and  $|r_{\mathbf{a}'}| \|\xi\| \gg (\log \|\xi\|)^{3l}$  we have

$$\begin{aligned}
\|r_{\mathbf{a}} O_{\mathbf{a}}^T \xi - r_{\mathbf{a}'} O_{\mathbf{a}'}^T \xi\| & = \left\| r_{\mathbf{a}} \|\xi\| O_{\mathbf{a}}^T \frac{\xi}{\|\xi\|} - r_{\mathbf{a}'} \|\xi\| O_{\mathbf{a}'}^T \frac{\xi}{\|\xi\|} \right\| \\
& \geq \|\xi\| \prod_{a \in \mathcal{A} \setminus \{a_1, a_2\}} |r_a|^{|\mathbf{a}|_a} \left\| r_{a_1}^{|\mathbf{a}|_{a_1}} r_{a_2}^{|\mathbf{a}|_{a_2}} O_{\mathbf{a}}^T \frac{\xi}{\|\xi\|} - r_{a_1}^{|\mathbf{a}'|_{a_1}} r_{a_2}^{|\mathbf{a}'|_{a_2}} O_{\mathbf{a}'}^T \frac{\xi}{\|\xi\|} \right\| \\
& \geq \min\{|r_{\mathbf{a}}| \|\xi\|, |r_{\mathbf{a}'}| \|\xi\|\} \left\| \left( O_{\mathbf{a}}^T \frac{\xi}{\|\xi\|} - O_{\mathbf{a}'}^T \frac{\xi}{\|\xi\|} \right) \right\| \\
& \gg \frac{(\log \|\xi\|)^{3l}}{(\log \|\xi\|)^l}.
\end{aligned}$$

In the penultimate line we have used the fact that if  $u, v \in \mathbb{R}^2 \setminus \{0\}$  have unit length and  $r_1, r_2 \geq 0$  satisfy  $r_1 \leq r_2$  then  $\|r_1 u - r_1 v\| \leq \|r_1 u - r_2 v\|$ . The final bound obtained above certainly exceeds 1 for  $\|\xi\|$  sufficiently large. Thus our result follows.  $\square$

To apply Lemma 3.5 in an analogous way to how we applied Lemma 3.3 in the proof of Theorem 1.6 we need the Fourier transform of our self-similar measure to satisfy a decay outside of sparse frequencies property as in the statement of Theorem 1.6. This decay is provided by Theorem 1.3 provided the self-similar measure is affinely non-concentrated in the sense of Definition 1.1. The latter property is guaranteed by a result of Banaji and Yu [13] which implies that if the orthogonal parts of our similarities commute and  $\mu$  is not supported on an affine subspace then this property is satisfied. Thus we can complete our proof of Theorem 1.9 by following the same strategy as used in our proof of Theorem 1.6.

**3.4. Remarks on Theorems 1.6 and 1.9.** Curiously our method of proof of reducing to random frequencies and using Fourier decay outside of a sparse set does *not* seem to give the Rajchman property if none of the log-contraction ratios  $\frac{\log |r_a|}{\log |r_b|}$  is a Diophantine number. That being said, if there exists  $a, b, a', b' \in \mathcal{A}$  such that both  $\frac{\log |r_a|}{\log |r_b|}$  and  $\frac{\log |r'_a|}{\log |r'_b|}$  are Liouville numbers, then under suitable conditions it is possible to establish a polylogarithmic decay rate. In particular, if those scales at which  $\frac{\log |r_a|}{\log |r_b|}$  observes very good rational approximations are sufficiently far apart from those scales at which  $\frac{\log |r'_a|}{\log |r'_b|}$  observes very good rational approximations, then one can adapt the proof of Theorem 1.6 to prove polylogarithmic Fourier decay.

#### 4. PATTERSON-SULLIVAN MEASURES: PROOF OF THEOREM 1.11

The goal of this section is to provide the proof of Theorem 1.11. Throughout this section, we fix a discrete, Zariski-dense, convex cocompact group  $\Gamma$  of isometries of  $\mathbb{H}^{d+1}$ ,  $d \geq 1$ .

**4.1. Reduction to linear phases.** Since PS measures are absolutely continuous with respect to the conditional measures  $\mu_x^u$  with smooth Radon-Nikodym derivatives, cf. (2.4), it suffices to prove the conclusion of Theorem 1.11 for  $\mu_x^u$  for some  $x \in G/\Gamma$  in place of the PS measure  $\mu$ . In fact, we prove a stronger statement which establishes uniform bounds for the measures  $\mu_x^u$  as  $x$  varies in the non-wandering set  $\Omega$ .

We begin with the following elementary lemma which reduces the proof to the study of linear phase functions.

**Lemma 4.1.** *To prove Theorem 1.11, it suffices to show that there exists  $\kappa > 0$  so that for all  $0 \neq \xi \in \mathbb{R}^d$ ,  $x \in \Omega$ , and  $\psi \in C_c^1(N_1^+)$ , we have*

$$\int_{N_1^+} e^{i\langle \xi, n \rangle} \psi(n) d\mu_x^u(n) \ll_{\Gamma} \|\psi\|_{C^1} \|\xi\|^{-\kappa}, \quad (4.1)$$

where, by abuse of notation, if  $n = \exp(v)$  for some  $v \in \mathfrak{n}^+ \cong \mathbb{R}^d$ , we let  $\langle \xi, n \rangle := \langle \xi, v \rangle$ .

**Remark 4.2.** In the remainder of this section, we fix  $\xi \in \mathbb{R}^d$  and  $\psi \in C_c^1(N_1^+)$ . Our goal is to prove the estimate (4.1).

*Proof of Lemma 4.1.* The proof is based on the uniformity of the estimate (4.1) as the basepoint  $x$  varies in  $\Omega$  which roughly translates to uniform Fourier decay (with linear phases) over pieces of the measure of size  $|\lambda|^{-1/2-\varepsilon}$ . We include a sketch of the argument for completeness.

Recall the notation of Theorem 1.11. Let  $\{\rho_j\}$  be a partition of unity of  $\text{supp}(\mu_x^u|_{N^+})$  with bounded multiplicity and such that each  $\rho_j$  is supported in a ball  $B_j$  of radius

$$r = |\lambda|^{-(1+\kappa)/(2+\kappa)}$$

around a  $n_j \in N_1^+$ . Here,  $\kappa$  is the exponent in (4.1). In view of [49, Prop. 9.9], we can choose such partition of unity so that each  $\rho_j$  has first derivatives with norm  $\mathcal{O}(r^{-1})$  and

$$\sum_j \mu_x^u(B_j) \ll \mu_x^u(N_1^+). \quad (4.2)$$

Then, Taylor expanding  $\varphi$  to the second order around each  $n_j$ , we obtain

$$\begin{aligned} \int_{N_1^+} e^{i\lambda\varphi(n)} \psi(n) d\mu_x^u &\leq \sum_j \left| \int_{B_j} \exp(i\langle \lambda \nabla \varphi(n_j), n n_j^{-1} \rangle) (\psi \rho_j)(n) d\mu_x^u \right| \\ &\quad + \mathcal{O}(\|\psi\|_{C^0} \|\varphi\|_{C^2} |\lambda| r^2 \mu_x^u(N_1^+)). \end{aligned}$$

Next, we use a change of variables sending  $B_j$  to  $N_1^+$  and apply (2.5) and (2.6). More precisely, let  $t = -\log r$ ,  $x_j = g_t n_j x$ ,  $\xi_j = r \lambda \nabla \varphi(n_j)$ , and  $\psi_j(n) := (\psi \rho_j)(\text{Ad}(g_{-t})(n) n_j)$ . Then, the  $j^{\text{th}}$  term in the above sum can be rewritten as follows:

$$\int_{B_j} \exp(i\langle \lambda \nabla \varphi(n_j), n n_j^{-1} \rangle) (\psi \rho_j)(n) d\mu_x^u = r^\delta \int_{N_1^+} \exp(i\langle \xi_j, n \rangle) \psi_j(n) d\mu_{x_j}^u.$$

Note that, since the geodesic flow scales the first derivatives of  $\rho_j$  by a factor of  $e^{-t}$ , each  $\psi_j$  has  $C^1$ -norm  $\mathcal{O}(\|\psi\|_{C^1})$ . Hence, since each  $x_j$  belongs to  $\Omega$ , (4.1) implies that

$$\int_{N_1^+} e^{i\lambda\varphi(n)} \psi(n) d\mu_x^u \ll_A |r\lambda|^{-\kappa} a^{-\kappa} \sum_j r^\delta + |\lambda| r^2.$$

Finally, in view of Sullivan's shadow lemma (cf. [49, Theorem 3.4]), we have that  $\mu_x^u(B_j) \asymp r^\delta$ . This concludes the proof in light of (4.2). The decay exponent obtained in this manner is  $\kappa/(2+\kappa)$ , with  $\kappa$  as in (4.1).  $\square$

**4.2. Polynomial non-concentration estimates.** We recall here well-known non-concentration estimates for PS measures. The first estimate is a direct consequence of Sullivan's shadow lemma.

**Proposition 4.3** (Sullivan's Shadow Lemma, cf. [86]). *For all  $y \in \Omega$  and all  $r > 0$ , we have*

$$\mu_y^u(N_r^+) \asymp_\Gamma r^\delta,$$

where  $\delta$  is the critical exponent of  $\Gamma$ .

We also recall the following quantitative decay property of the measure of hyperplane neighborhoods with respect to PS measures from [22]. Recall that  $N^+$  is an abelian group which we identify with its Lie algebra  $\mathfrak{n}^+ \cong \mathbb{R}^d$  via the exponential map. In light of this identification, the following result shows that PS measures (or, more precisely, their shadows  $\mu_x^u$ ) are non-concentrated in the sense of Definition 1.1. This will allow us to apply Theorem 1.3 in the proof of Theorem 1.11 as well as provide quantitative estimates on separation of frequencies arising over the course of implementing the strategy discussed in the introduction.

**Theorem 4.4** ([49, Theorem 12.1]). *There exist constants  $C \geq 1$  and  $\alpha > 0$  such that for all  $\varepsilon, r > 0$ ,  $x \in \Omega$ , and all affine hyperplanes  $L < N^+$ , we have that*

$$\mu_x^u(N_r^+ \cap L^{(\varepsilon r)}) \leq C \varepsilon^\alpha \mu_x^u(N_r^+),$$

where  $L^{(\varepsilon r)}$  denotes the  $\varepsilon r$ -neighborhood of  $L$  in  $N^+$ .

**Remark 4.5.** The reference [49, Theorem 12.1] shows that, in the more general setting of geometrically finite groups, one has  $\mu_x^u(N_r^+ \cap L^{(\varepsilon r)}) \leq t(\varepsilon) V(x) \mu_x^u(N_r^+)$ , for a function  $t(\varepsilon) \rightarrow 0$  as  $\varepsilon \rightarrow 0$  and for a function  $V(x)$  that is uniformly bounded above and below on compact sets. The proof is much simpler in the case  $\Gamma$  is convex cocompact and in fact yields the apriori stronger bound  $t(\varepsilon) \leq C \varepsilon^\alpha$ . In fact, when  $\Gamma$  is convex cocompact, Theorem 4.4 can be deduced directly from the

fact that PS measures give 0 mass to proper subvarieties of the boundary ([30, Corollary 9.4]) using the argument in [50, Section 8].

**4.3. Proof of Theorem 1.11.** The remainder of the section is dedicated to the proof of the estimate (4.1). We fix  $\xi \in \mathbb{R}^d$ ,  $x \in \Omega$ , and  $\psi \in C_c^1(N_1^+)$ . Our argument is dynamical in nature using the self-similar structure of the measures  $\mu_x^u$  under the geodesic flow, cf. (2.5) and (2.6), to implement the strategy described in the introduction.

**Partitions of unity and flow prisms.** As a first step, we find convenient partitions of the space by flow boxes. Namely, we refer to sets of the form  $P_r^- N_s^+ \cdot x$  for  $r, s > 0$  and  $x \in G/\Gamma$  as *flow boxes*. We say that a collection of sets  $\{S_i\}$  has multiplicity bounded by a constant  $C \geq 1$  if for all  $x$ :

$$\sum_i \mathbb{1}_{S_i}(x) \leq C \mathbb{1}_{\cup_i S_i}(x).$$

Let  $\iota$  denote the smaller of  $1/2$  and the injectivity radius of  $G/\Gamma$  and set

$$\iota_\xi := \iota / \|\xi\|^{1/3}. \quad (4.3)$$

The following lemma provides us with an efficient cover of  $\Omega$  by “thin flow boxes” in the unstable direction.

**Lemma 4.6.** *The collection  $\{P_\iota^- N_{\iota_\xi}^+ \cdot x : x \in \Omega\}$  admits a finite subcover  $\mathcal{B}_\xi$  such that*

$$\#\mathcal{B}_\xi \ll_\Gamma \|\xi\|^{\delta/3},$$

where  $\delta$  is the critical exponent of  $\Gamma$ . Moreover,  $\mathcal{B}_\xi$  has uniformly bounded multiplicity on  $\Omega$ ; i.e. for all  $x \in \Omega$ ,  $\sum_{B \in \mathcal{B}_\xi} \mathbb{1}_B(x) \ll_\Gamma 1$ .

*Proof.* Let  $\mathcal{Q}$  denote a cover of  $G/\Gamma$  by flow boxes of the form  $N_\iota^+ P_\iota^- \cdot x$ , where  $\iota$  is a fixed lower bound on the injectivity radius of  $G/\Gamma$  as above. With the help of the Vitali covering lemma, such cover can be chosen to have multiplicity  $C_G \geq 1$  depending only on the dimension of  $G$ . We will build our collection of boxes  $\mathcal{B}_\xi$  by refining this cover as follows.

Let  $\mathcal{Q}^0$  denote the collection of boxes  $Q \in \mathcal{Q}$  such that  $Q \cap \Omega \neq \emptyset$ . By convex cocompactness, we have that  $\#\mathcal{Q}^0 \asymp_\Gamma 1$ . For each  $Q \in \mathcal{Q}^0$ , we fix some  $x_Q \in Q \cap \Omega$ . Then, we can find a finite set of points  $\{u_i : i \in I_Q\} \subset N_{2\iota}^+$  such that the points  $x_i := u_i x_Q$  belong to  $\Omega$ . Moreover, these points be chosen so that the balls  $N_{\iota_\xi}^+ \cdot x_i$  provide a cover of  $\Omega \cap N_\iota^+ \cdot x_Q$  with multiplicity bounded by a constant  $C_\Gamma \geq 1$  depending only  $\Gamma$  (i.e.  $C_\Gamma$  is independent of  $\xi$ ). This is again possible thanks to the Vitali covering lemma and the fact that the measure  $\mu_{x_Q}^u \big|_{N_{3\iota}^+}$  is doubling by Proposition 4.3.

With this notation, we define  $\mathcal{B}_\xi$  as follows:

$$\mathcal{B}_\xi := \left\{ P_\iota^- N_{\iota_\xi}^+ \cdot u_i x_Q : i \in I_Q, Q \in \mathcal{Q}^0 \right\}.$$

This bounded multiplicity in particular implies that

$$\iota_\xi^\delta \times \#I_Q \asymp \sum_{i \in I_Q} \mu_{x_i}^u(N_{\iota_\xi}^+) \asymp \mu_{x_Q}^u(N_\iota^+) \asymp 1.$$

This estimate implies the desired the bound on the cardinality of  $\mathcal{B}_\xi$  in light of (4.3). To bound the multiplicity of  $\mathcal{B}_\xi$ , let  $x \in \Omega$  be arbitrary, and note that

$$\sum_{B \in \mathcal{B}_\xi} \mathbb{1}_B(x) = \sum_{Q \in \mathcal{Q}^0} \sum_{i \in I_Q} \mathbb{1}_{P_\iota^- N_{\iota_\xi}^+ \cdot u_i x_Q}(x) \leq C_\Gamma \sum_{Q \in \mathcal{Q}^0} \mathbb{1}_{\cup_{i \in I_Q} P_\iota^- N_{\iota_\xi}^+ \cdot u_i x_Q}(x) \leq C_\Gamma \#\mathcal{Q}^0 \ll_\Gamma 1.$$

This concludes the proof.  $\square$

Let  $\mathcal{B}_\xi$  be the finite cover provided by Lemma 4.6 and let  $\mathcal{P}_\xi$  denote a partition of unity subordinate to it. For each  $\rho \in \mathcal{P}_\xi$ , we denote by  $B_\rho$  the element of  $\mathcal{B}_\xi$  containing the support of  $\rho$ . In particular, such partition of unity can be chosen so that for all  $\rho \in \mathcal{P}_\xi$ , we have

$$\|\rho\|_{C^1} \ll \|\xi\|^{1/3}. \quad (4.4)$$

Moreover, by Lemma 4.6, we have

$$\#\mathcal{P}_\xi \leq \#\mathcal{B}_\xi \ll_\Gamma \|\xi\|^{\delta/3}. \quad (4.5)$$

**Transversals.** We fix a system of transversals  $\{T_\rho\}$  to the strong unstable foliation inside the boxes  $B_\rho$ . Since  $B_\rho$  meets  $\Omega$  for all  $\rho \in \mathcal{P}_\xi$ , we fix some  $y_\rho$  in the intersection  $B_\rho \cap \Omega$ . In this notation, we write

$$B_\rho = P_\iota^- N_{\iota\xi}^+ \cdot y_\rho, \quad T_\rho = P_\iota^- \cdot y_\rho. \quad (4.6)$$

We also let  $M_\rho, A_\rho$ , and  $N_\rho^-$  be neighborhoods of identity in  $M, A$  and  $N^-$  respectively so that  $P_\iota^- = M_\rho A_\rho N_\rho^-$ .

**Saturation.** Fix  $t > 0$  to be chosen so that  $e^t$  is a small positive power of  $\|\xi\|$ ; cf. (4.38). Using our partition of unity, we can write

$$\int_{N_1^+} e^{i\langle \xi, n \rangle} \psi(n) d\mu_x^u(n) = \sum_{\rho \in \mathcal{P}_\xi} \int_{N_1^+} e^{i\langle \xi, n \rangle} \psi(n) \rho(g_t n x) d\mu_x^u(n). \quad (4.7)$$

Here, we are using the fact that, since  $x \in \Omega$ , then the restriction of the support of  $\mu_x^u$  to  $N_1^+$  consists of points  $n \in N_1^+$  with  $nx \in \Omega$  (or equivalently, that  $g_t n x \in \Omega$ ) and that  $\sum_\rho \rho(y) = 1$  for all  $y \in \Omega$ .

Our first step is to partition the integrals on the right side of (4.7) over  $N_1^+$  into pieces according to the flow box they land in under flowing by  $g_t$ . To simplify notation, we write

$$x_t := g_t x. \quad (4.8)$$

We denote by  $N_1^+(t)$  a neighborhood of  $N_1^+$  defined by the property that the intersection

$$B_\rho \cap (\text{Ad}(g_t)(N_1^+(t)) \cdot x_t)$$

consists entirely of full local strong unstable leaves in  $B_\rho$ . We note that since  $\text{Ad}(g_t)$  expands  $N^+$  and  $B_\rho$  has radius  $< 1$ ,  $N_1^+(t)$  is contained inside  $N_2^+$ . Since  $\psi$  is compactly supported inside  $N_1^+$ , we have

$$\chi_{N_1^+}(n) \psi(n) = \chi_{N_1^+(t)}(n) \psi(n), \quad \forall n \in N^+. \quad (4.9)$$

For simplicity, we set

$$\xi_t := e^{-t} \xi, \quad \psi_t(n) := \psi(\text{Ad}(g_t)^{-1} n), \quad \mathcal{A}_t := \text{Ad}(g_t)(N_1^+(t)).$$

For  $\rho \in \mathcal{P}_\xi$ , we let  $\mathcal{W}_{\rho,t}$  denote the collection of connected components of the set

$$\{n \in \mathcal{A}_t : nx_t \in B_\rho\}.$$

In view of (4.9), changing variables using (2.5) yields

$$\sum_{\rho \in \mathcal{P}_\xi} \int_{N_1^+} e^{i\langle \xi, n \rangle} \psi(n) \rho(g_t n x) d\mu_x^u = e^{-\delta t} \sum_{\rho \in \mathcal{P}_\xi, W \in \mathcal{W}_{\rho,t}} \int_{n \in W} e^{i\langle \xi_t, n \rangle} \psi_t(n) \rho(nx_t) d\mu_{x_t}^u. \quad (4.10)$$

**Centering the integrals.** It will be convenient to center all the integrals in (4.10) so that their basepoints belong to the transversals  $T_\rho$  of the respective flow box  $B_\rho$ ; cf. (4.6).

Let  $I_{\rho,t}$  denote an index set for  $\mathcal{W}_{\rho,t}$ . For  $W \in \mathcal{W}_{\rho,t}$  with index  $\ell \in I_{\rho,t}$ , let  $n_{\rho,\ell} \in W$ ,  $m_{\rho,\ell} \in M_\rho$ ,  $n_{\rho,\ell}^- \in N_\rho^-$ , and  $t_{\rho,\ell}$  with  $|t_{\rho,\ell}| \ll \iota$  be such that

$$x_{\rho,\ell} := n_{\rho,\ell} \cdot x_t = n_{\rho,\ell}^- m_{\rho,\ell} g_{t_{\rho,\ell}} \cdot y_\rho \in T_\rho. \quad (4.11)$$

Note that since  $x$  and  $y_\rho$  both belong to  $\Omega$ , we have that

$$x_{\rho,\ell} \in \Omega, \quad n_{\rho,\ell}^- y_\rho \in \Omega. \quad (4.12)$$

For each such  $\ell$  and  $W$ , let us denote  $W_\ell = W n_{\rho,\ell}^{-1}$  and set

$$\tilde{\chi}_{\rho,\ell}(t, n) := \exp(i \langle \xi_t, n n_{\rho,\ell} \rangle). \quad (4.13)$$

Changing variables using (2.5) and (2.6), we can rewrite the right side of (4.10) as follows:

$$\begin{aligned} e^{-\delta t} \sum_{\rho \in \mathcal{P}_\xi, W \in \mathcal{W}_{\rho,t}} \int_{n \in W} e^{i \langle \xi_t, n \rangle} \psi_t(n) \rho(n x_t) d\mu_{x_t}^u(n) \\ = e^{-\delta t} \sum_{\rho \in \mathcal{P}_\xi} \sum_{\ell \in I_{\rho,t}} \int_{n \in W_\ell} \tilde{\chi}_{\rho,\ell}(t, n) \psi_t(n n_{\rho,\ell}) \rho(n x_{\rho,\ell}) d\mu_{x_{\rho,\ell}}^u(n). \end{aligned} \quad (4.14)$$

**Mass estimates.** We record here certain counting estimates which will allow us to sum error terms in later estimates over  $\mathcal{P}_\xi$ . Note that by definition of  $N_1^+(j)$ , we have  $\bigcup_{\rho \in \mathcal{P}_\xi, W \in \mathcal{W}_{\rho,t}} W \subseteq \mathcal{A}_t$ . Thus, it follows that

$$\sum_{\rho \in \mathcal{P}_\xi, \ell \in I_{\rho,t}} \mu_{x_{\rho,\ell}}^u(W_\ell) \ll \mu_{x_t}^u(\mathcal{A}_t) = e^{\delta t} \mu_x^u(N_1^+(t)) \ll e^{\delta t} \mu_x^u(N_2^+) \ll e^{\delta t} \mu_x^u(N_1^+), \quad (4.15)$$

where the last inequality follows since  $N_1^+(j) \subseteq N_2^+$  using the doubling property of PS measures [49, Proposition 3.1]. We also used the fact that the partition of unity  $\mathcal{P}_\xi$  has uniformly bounded multiplicity.

**Weak-stable holonomy.** Fix some  $\rho \in \mathcal{P}_\xi$ . Recall the points  $y_\rho \in T_\rho$  and  $n_{\rho,\ell}^- \in N_\rho^-$  satisfying (4.11). Let

$$p_{\rho,\ell}^- := n_{\rho,\ell}^- m_{\rho,\ell} g_{t_{\rho,\ell}}. \quad (4.16)$$

The product map  $N^- \times A \times M \times N^+ \rightarrow G$  is a diffeomorphism on a ball of radius 1 around the identity; cf. Section 2.2.5. Hence, given  $\ell \in I_{\rho,t}$ , we can define maps  $\phi_\ell$  and  $\tilde{p}_\ell^-$  from  $W_\ell$  to  $N^+$  and  $P^-$  respectively by the following formula

$$n p_{\rho,\ell}^- = \tilde{p}_\ell^-(n) \phi_\ell(n). \quad (4.17)$$

We suppress the dependence on  $\rho$  and  $t$  to ease notation. Then,  $\phi_\ell$  induces a map between the strong unstable manifolds of  $x_{\rho,\ell}$  and  $y_\rho$ , also denoted  $\phi_\ell$ , and defined by

$$\phi_\ell(n x_{\rho,\ell}) = \phi_\ell(n) y_\rho.$$

In particular, this induced map coincides with the local weak stable holonomy map inside  $B_\rho$ .

Note that we can find a neighborhood  $W_\rho \subset N^+$  of identity of radius  $\asymp \iota_\xi$  such that

$$\phi_\ell(W_\ell) \subseteq W_\rho, \quad (4.18)$$

for all  $\ell \in I_{\rho,t}$ . Moreover, by shrinking the radius  $\iota_\xi$  of the flow boxes by an absolute amount (depending only on the metric on  $G$ ) if necessary, we may assume that all the maps  $\phi_\ell$  are invertible

on  $W_\rho$ . Hence, we can define the following:

$$\begin{aligned} p_\ell^-(n) &:= \tilde{p}_\ell^-(\phi_\ell^{-1}(n)) \in P^-, & \tilde{\psi}_{\rho,\ell}(t, n) &:= J\phi_\ell(n) \times \psi_t(\phi_\ell^{-1}(n)n_{\rho,\ell}), \\ \chi_{\rho,\ell}(t, n) &:= \tilde{\chi}_{\rho,\ell}(t, \phi_\ell^{-1}(n)), & \rho_\ell(n) &:= \rho(p_\ell^-(n)ny_\rho), \end{aligned} \quad (4.19)$$

where  $J\phi_\ell$  denotes the Jacobian of the change of variable  $\phi_\ell$ ; cf. (2.9).

Changing variables in the right side of (4.14), we obtain

$$\sum_{\ell \in I_{\rho,t}} \int_{n \in W_\ell} \tilde{\chi}_{\rho,\ell}(t, n) \tilde{\psi}_{\rho,\ell}(t, n) \rho(nx_{\rho,\ell}) d\mu_{x_{\rho,\ell}}^u = \sum_{\ell \in I_{\rho,t}} \int_{W_\rho} \chi_{\rho,\ell}(t, n) \tilde{\psi}_{\rho,\ell}(t, n) \rho_\ell(n) d\mu_{y_\rho}^u. \quad (4.20)$$

**Phase formula.** The following lemma provides a formula for the stable holonomy maps  $\phi_\ell$  defined above (4.18) which are responsible for the oscillation of  $\chi_{\rho,\ell}$  along  $N^+$ . The elementary proof of this lemma is given in Section 4.4.

**Lemma 4.7.** *Let  $p_{\rho,\ell}^-$  be as in (4.16) and let  $w_{\rho,\ell} \in \mathfrak{n}^-$  be such that  $n_{\rho,\ell}^- = \exp(w_{\rho,\ell})$ . Define vectors  $z_{\rho,\ell} \in \mathfrak{n}^-$  by*

$$z_{\rho,\ell} := -e^{t_{\rho,\ell}} m_{\rho,\ell}^{-1} \cdot w_{\rho,\ell}. \quad (4.21)$$

Then, for every  $n = \exp(v) \in N_{1/2}^+$ , we have

$$\log \phi_\ell^{-1}(n) = e^{t_{\rho,\ell} - \tilde{\tau}_\ell(v)} m_{\rho,\ell} \cdot \left( v + \frac{\|v\|^2}{2} z_{\rho,\ell} \right),$$

where  $\log \phi_\ell^{-1}(n)$  is viewed as an element of  $\mathfrak{n}^+$  and  $\tilde{\tau}_\ell : N_{1/2}^+ \rightarrow \mathbb{R}_+$  is given by

$$\tilde{\tau}_\ell(v) = \log \left( 1 + \langle v, z_{\rho,\ell} \rangle + \frac{\|v\|^2 \|z_{\rho,\ell}\|^2}{4} \right).$$

It will be convenient for our estimates to simplify the expression for  $\tilde{\tau}_\ell$  by removing the quadratic term. This is the reason for our choice of flow boxes of width  $\asymp \|\xi\|^{-1/3}$  along the strong unstable manifold. In what follows, to simplify notation, we set

$$\tau_\ell(v) = -t_{\rho,\ell} + \log(1 + \langle v, z_{\rho,\ell} \rangle), \quad \Gamma_\ell(v) := e^{-\tau_\ell(v)} m_{\rho,\ell} \cdot \left( v + \frac{\|v\|^2}{2} z_{\rho,\ell} \right). \quad (4.22)$$

Recall the centering points  $n_{\rho,\ell}$  defined in (4.11). The following corollary provides a first step towards linearizing the phase in the oscillatory functions  $\chi_{\rho,\ell}$  by replacing  $\tilde{\tau}_\ell$  in Lemma 4.7 with  $\tau_\ell$  in (4.22).

**Corollary 4.8.** *With the same notation as in Lemma 4.7, we have for all  $n = \exp(v) \in W_\rho$  that*

$$\chi_{\rho,\ell}(t, n) = \mathbf{a}_{\rho,\ell}(t, n) + \mathcal{O}(e^{-t}),$$

where

$$\mathbf{a}_{\rho,\ell}(t, n) := \exp(i\langle \xi_t, n_{\rho,\ell} \rangle) \times \exp(i\langle \xi_t, \Gamma_\ell(v) \rangle). \quad (4.23)$$

*Proof.* Recall from (4.19) and (4.13) that  $\chi_{\rho,\ell}(t, n) = \exp(i\langle \xi_t, \phi_\ell^{-1}(n)n_{\rho,\ell} \rangle)$ . We also recall that  $\xi_t = e^{-t}\xi$ . Then, for all  $n = \exp(v) \in W_\rho$ , we have that

$$|\chi_{\rho,\ell}(t, n) - \mathbf{a}_{\rho,\ell}(t, n)| \ll \|v\|^2 \|\xi_t\| \|\Gamma_\ell(v)\|$$

Since  $W_\rho$  has radius  $\asymp \iota_\xi \asymp \|\xi\|^{-1/3}$  (cf. (4.3)), we have that both  $v$  and  $\Gamma_\ell(v)$  have norm  $\ll \|\xi\|^{-1/3}$ . In particular, the upper bound above is  $\mathcal{O}(e^{-t})$  as desired.  $\square$

Let us summarize our progress so far. To simplify notation, set

$$\psi_{\rho,\ell}(t, n) := \tilde{\psi}_{\rho,\ell}(t, n) \times \rho_\ell(n). \quad (4.24)$$

Then, in light of (4.7), (4.10), (4.14), (4.20), and Corollary 4.8, we find that

$$\int_{N_1^+} e^{i\langle \xi, n \rangle} \psi(n) d\mu_x^u = e^{-\delta t} \sum_{\rho \in \mathcal{P}_\xi} \int_{W_\rho} \sum_{\ell \in I_{\rho,t}} \mathbf{a}_{\rho,\ell}(t, n) \psi_{\rho,\ell}(t, n) d\mu_{y_\rho}^u + \mathcal{O}(e^{-t}). \quad (4.25)$$

**Cauchy-Schwarz.** We are left with estimating integrals of the form:

$$\int_{W_\rho} \Psi_\rho(t, n) d\mu_{y_\rho}^u, \quad \Psi_\rho(t, n) := \sum_{\ell \in I_{\rho,t}} \mathbf{a}_{\rho,\ell}(t, n) \psi_{\rho,\ell}(t, n). \quad (4.26)$$

By Cauchy-Schwarz, we get

$$\left| \int_{W_\rho} \Psi_\rho(t, n) d\mu_{y_\rho}^u \right|^2 \leq \mu_{y_\rho}^u(W_\rho) \int_{W_\rho} |\Psi_\rho(t, n)|^2 d\mu_{y_\rho}^u \quad (4.27)$$

We begin by noting the following apriori bounds on  $\Psi_\rho$ :

$$\|\psi_{\rho,\ell}\|_{L^\infty(W_\rho)} \ll 1, \quad \|\Psi_\rho\|_{L^\infty(W_\rho)} \ll \#I_{\rho,t}. \quad (4.28)$$

**Partitioning the support.** Using [49, Proposition 9.9], we can find a cover  $\{A_j\}$  of  $W_\rho$  with balls of radius

$$r = \|\xi\|^{-1/2} \quad (4.29)$$

centered around  $u_j \in W_\rho \cap \text{supp}(\mu_{y_\rho}^u)$  and satisfying  $\sum_j \mu_{y_\rho}^u(A_j) \ll \mu_{y_\rho}^u(W_\rho)$ . By the triangle inequality<sup>4</sup> we have

$$\int_{W_\rho} |\Psi_\rho(t, n)|^2 d\mu_{y_\rho}^u \leq \sum_j \int_{A_j} |\Psi_\rho(t, n)|^2 d\mu_{y_\rho}^u. \quad (4.30)$$

For  $k, \ell \in I_{\rho,t}$ , we let

$$\psi_{k,\ell}(t, n) := \psi_{\rho,k}(t, n) \overline{\psi_{\rho,\ell}(t, n)}, \quad \mathbf{a}_{k,\ell}(t, n) := \mathbf{a}_{\rho,k}(t, n) \overline{\mathbf{a}_{\rho,\ell}(t, n)}.$$

Expanding the square, we get

$$\sum_j \int_{A_j} |\Psi_\rho(t, n)|^2 d\mu_{y_\rho}^u \leq \sum_j \sum_{k, \ell \in I_{\rho,t}} \left| \int_{A_j} \mathbf{a}_{k,\ell}(t, n) \psi_{k,\ell}(t, n) d\mu_{y_\rho}^u \right|.$$

Using (2.5) and (2.6), we change variables in the integrals using the maps taking each  $A_j$  onto  $N_1^+$ . More precisely, recall that  $A_j$  is a ball of radius  $r$  around  $u_j$  such that  $u_j y_\rho \in \Omega$ . Letting

$$\begin{aligned} \tau &= -\log r, & y_\rho^j &= g_\tau u_j y_\rho, & \mathbf{a}_{k,\ell}^j(t, n) &= \mathbf{a}_{k,\ell}(t, \text{Ad}(g_{-\tau})(n) u_j), \\ \psi_{k,\ell}^j(t, n) &= \psi_{k,\ell}(t, \text{Ad}(g_{-\tau})(n) u_j), \end{aligned} \quad (4.31)$$

we can rewrite the above sum as

$$\sum_j \sum_{k, \ell \in I_{\rho,t}} \left| \int_{A_j} \mathbf{a}_{k,\ell}(t, n) \psi_{k,\ell}(t, n) d\mu_{y_\rho}^u \right| \leq r^\delta \sum_j \sum_{k, \ell \in I_{\rho,t}} \left| \int_{N_1^+} \mathbf{a}_{k,\ell}^j(t, n) \psi_{k,\ell}^j(t, n) d\mu_{y_\rho^j}^u \right|. \quad (4.32)$$

One advantage of flowing forward by  $g_\tau$  is that it provides smoothing of the amplitude functions  $\psi_{k,\ell}$ . In particular, it follows by (4.4) that

$$\left\| \psi_{k,\ell}^j \right\|_{C^1} \ll \|\psi\|_{C^1} \times r \times \|\xi\|^{1/3} \ll \|\psi\|_{C^1} \|\xi\|^{-1/6}.$$

<sup>4</sup>Cauchy-Schwarz allows us to have a non-negative integrand which in turn enables this step.

Applied to the right side of (4.32), we obtain

$$\int_{W_\rho} |\Psi_\rho(t, n)|^2 d\mu_{y_\rho}^u = r^\delta \sum_j \sum_{k, \ell \in I_{\rho, t}} \left| \int_{N_1^+} \mathbf{a}_{k, \ell}^j(t, n) d\mu_{y_\rho}^u \right| + \mathcal{O}(\|\psi\|_{C^1} \|\xi\|^{-1/6} \# I_{\rho, t}^2 \mu_{y_\rho}^u(W_\rho)). \quad (4.33)$$

**Linearizing the phase.** We now turn to estimating the sum of oscillatory integrals in (4.33). Recall that  $u_j$  denotes the center of the ball  $A_j$  for each  $j$  and let  $v_j \in \mathfrak{n}^+$  be such that

$$u_j = \exp(v_j).$$

Then, given  $n = \exp(v) \in A_j$ , and recalling the maps  $\Gamma_\ell$  in (4.22), we get

$$\Gamma_\ell(v) = \Gamma_\ell(v_j) + D(\Gamma_\ell(v_j))(v - v_j) + \mathcal{O}(r^2),$$

where  $D(\Gamma_\ell)$  denotes the derivative of  $\Gamma_\ell$ .

The following elementary lemma uses the explicit expression for  $\Gamma_\ell$  in (4.22) to simplify the form of  $D\Gamma_\ell(v_j)$ .

**Lemma 4.9.** *For all  $\ell$  and  $j$ , we have*

$$D\Gamma_\ell(v_j) = e^{-\tau_\ell(v_j)} m_{\rho, \ell} + \mathcal{O}(\|\xi\|^{-2/3}).$$

Let

$$\mathbf{b}_{k, \ell}^j := r \xi_t \cdot \left( e^{-\tau_k(v_j)} m_{\rho, k} - e^{-\tau_\ell(v_j)} m_{\rho, \ell} \right). \quad (4.34)$$

Recall that  $\xi_t = e^{-t} \xi$  so that  $\|\xi_t\| r^2 = e^{-t}$ . Hence, by absorbing the constant terms into the absolute value, we obtain from (4.33) and Lemma 4.9 that

$$\begin{aligned} & \int_{W_\rho} |\Psi_\rho(t, n)|^2 d\mu_{y_\rho}^u \\ &= r^\delta \sum_j \sum_{k, \ell \in I_{\rho, t}} \left| \int_{N_1^+} \exp(i \langle \mathbf{b}_{k, \ell}^j, v \rangle) d\mu_{y_\rho}^u \right| + \mathcal{O}((e^{-t} + e^{-t} \|\xi\|^{-1/6} + \|\psi\|_{C^1} \|\xi\|^{-1/6}) \# I_{\rho, t}^2 \mu_{y_\rho}^u(W_\rho)). \end{aligned} \quad (4.35)$$

*Proof of Lemma 4.9.* Recall the definition of the vectors  $z_{\rho, \ell} \in \mathfrak{n}^-$ . To simplify notation, set

$$\lambda_\ell(v_j) = \frac{1}{1 + \langle v_j, z_{\rho, \ell} \rangle}.$$

In particular,  $e^{-\tau_\ell(v_j)} = e^{t_{\rho, \ell}} \lambda_\ell(v_j)$ . Then, using the formula for  $\Gamma_\ell$  in (4.22), we obtain

$$D\Gamma_\ell(v_j) = e^{t_{\rho, \ell}} \lambda_\ell(v_j) m_{\rho, \ell} \left[ -\lambda_\ell(v_j) \left( v_j \cdot z_{\rho, \ell}^t + \frac{\|v_j\|^2}{2} z_{\rho, \ell} \cdot z_{\rho, \ell}^t \right) + \text{Id} + z_{\rho, \ell} \cdot v_j^t \right].$$

Here, we are viewing  $v_j$  and  $z_{\rho, \ell}$  as  $(d \times 1)$ -column vectors and use  $v_j^t$  and  $z_{\rho, \ell}^t$  to denote the transpose of  $v_j$  and  $z_{\rho, \ell}$  respectively. Now, observe that

$$\lambda_\ell(v_j) v_j = v_j - \frac{\langle v_j, z_{\rho, \ell} \rangle}{1 + \langle v_j, z_{\rho, \ell} \rangle} v_j = v_j + \mathcal{O}(\|v_j\|^2).$$

The lemma now follows upon recalling that  $\exp(v_j)$  belongs to  $W_\rho$  so that  $\|v_j\| \ll \|\xi\|^{-1/3}$  in view of our choice of flow boxes; cf. (4.3) and the discussion around it.  $\square$

**Separation of frequencies.** To apply Theorem 1.3, it will be important to understand the distribution of the frequencies  $\mathbf{b}_{k,\ell}^j$ . To this end, we have the following lemma which allows us to avoid studying the separation of the holonomy matrices  $m_{\rho,\ell}$ .

**Lemma 4.10.** *For all  $j, k, \ell$ , we have*

$$\|\mathbf{b}_{k,\ell}^j\| \gg \|r\xi_t\| |e^{-\tau_\ell(v_j)} - e^{-\tau_k(v_j)}|,$$

where  $\tau_\ell(v_j)$  and  $\tau_k(v_j)$  are defined in (4.22).

*Proof.* In what follows, to simplify notation, we let

$$m_k := m_{\rho,k}, \quad c_k := e^{-\tau_k(v_j)}, \quad Q_k := c_k m_k,$$

with the similar notation for the index  $\ell$  in place of  $k$  defined analogously. The lemma is evident when  $c_k = c_\ell$ . Hence, we may assume without loss of generality that  $c_k > c_\ell$ , and recall that these functions are non-negative by definition; cf. (4.23).

Recall the elementary estimate  $\|g \cdot v\| \geq \|v\| / \|g^{-1}\|$  for any invertible linear map  $g$  and any vector  $v \in \mathbb{R}^d$ . This estimate implies the following lower bound:

$$\|\mathbf{b}_{k,\ell}^j\| \geq \frac{r \|\xi_t\|}{\|(Q_k - Q_\ell)^{-1}\|} = \frac{r \|\xi_t\| c_k}{\|(\text{Id} - \frac{c_\ell}{c_k} m_\ell m_k^{-1})^{-1}\|}.$$

That  $\text{Id} - \frac{c_\ell}{c_k} m_\ell m_k^{-1}$  (and hence  $Q_k - Q_\ell$ ) is invertible follows at once from the following estimate on the norm of its inverse. Recall that the rotation matrices  $m_k$  and  $m_\ell$  have spectral radius 1. In particular, since  $c_\ell < c_k$ , we may use the power series expansion of  $\text{Id} - Q$ , for matrices  $Q$  with spectral radius  $< 1$ , to get that

$$\|(\text{Id} - \frac{c_\ell}{c_k} m_\ell m_k^{-1})^{-1}\| \ll \sum_{n \geq 0} \left(\frac{c_\ell}{c_k}\right)^n = \frac{c_k}{c_k - c_\ell}.$$

The lemma follows by combining the above two estimates.  $\square$

To proceed, we recall that  $t_{\rho,\ell}$ ,  $m_{\rho,\ell}$ , and  $n_{\rho,\ell}^- = \exp(w_{\rho,\ell})$  parametrize respectively the geodesic flow,  $M$ , and strong stable coordinates of the transverse intersections of the expanded horospherical disk  $g_t N_1^+ x$  with a fixed transversal  $T_\rho$  of the flow box  $B_\rho$ . We also recall that  $z_{\rho,\ell} = e^{t_{\rho,\ell}} m_{\rho,\ell}^{-1} w_{\rho,\ell}$  and  $\tau_\ell(v_j) = e^{t_{\rho,\ell}} / (1 + \langle v_j, z_{\rho,\ell} \rangle)$ .

Lemma 4.10 motivates the definition of the following subset of  $I_{\rho,t}^2$  parametrizing pairs  $(k, \ell)$  for which the frequencies  $\mathbf{b}_{k,\ell}^j$  are too small. Namely, we set

$$\text{Small} := \left\{ (k, \ell, j) : \|\mathbf{b}_{k,\ell}^j\| < 1 \right\}. \quad (4.36)$$

Roughly speaking, elements of  $C_{k,\ell}$  correspond to points  $v_j$  lying in a small neighborhood of a hyperplane orthogonal to  $m_{\rho,\ell}^{-1} w_{\rho,\ell} - m_{\rho,k}^{-1} w_{\rho,k}$ . Theorem 4.4 will then provide us with a counting estimate on  $C_{k,\ell}$ . This is done in the following lemma.

**Lemma 4.11.** *Let  $\alpha > 0$  be the exponent provided by Theorem 4.4. Then, for every fixed  $k, \ell \in I_{\rho,t}$ , we have*

$$\sum_{j: (k,\ell,j) \in \text{Small}} r^\delta \ll \left( \frac{\|\xi\|^{-1/6} e^t}{\|m_{\rho,\ell}^{-1} w_{\rho,\ell} - m_{\rho,k}^{-1} w_{\rho,k}\|} \right)^\alpha \mu_{y_\rho}^u(W_\rho).$$

*Proof.* Let  $j$  be such that  $(k, \ell, j) \in \text{Small}$  and recall that  $\xi_t = e^{-t} \xi$  and  $r = \|\xi\|^{-1/2}$ . To simplify notation, we also let

$$u_{k,\ell} := m_{\rho,\ell}^{-1} w_{\rho,\ell} - m_{\rho,k}^{-1} w_{\rho,k}.$$

Then, Lemma 4.10 and a direct calculation show that

$$|e^{t\rho,\ell} - e^{t\rho,k} + e^{t\rho,\ell+t\rho,k} \langle v_j, u_{k,\ell} \rangle| \ll \|\xi\|^{-1/2} e^t.$$

Let  $\epsilon_1 = \|\xi\|^{-1/2} e^t / \|u_{k,\ell}\|$ . It follows that  $v_j$  belongs to a neighborhood of radius  $\mathcal{O}(\epsilon_1)$  around an affine hyperplane  $L$  parallel to the kernel of the linear functional  $v \mapsto \langle v, u_{k,\ell} \rangle$ .

Recall that  $A_j$  denotes the ball of radius  $r$  around  $\exp(v_j) \in W_\rho$  and that  $W_\rho$  has radius  $\asymp \|\xi\|^{-1/3}$ . It follows we can find a radius  $\epsilon_2 \asymp \|\xi\|^{-1/3}$  such that

$$\bigcup_{j:(k,\ell,j) \in \text{Small}} A_j \subseteq L^{(\epsilon_1+r)} \cap N_{\epsilon_2}^+,$$

where  $L^{(\epsilon_1+r)}$  denotes the  $(\epsilon_1+r)$ -neighborhood of  $L$ . Furthermore, by the bounded multiplicity of the cover  $\{A_j\}$  of  $W_\rho$  and the fact that each  $A_j$  has measure  $\asymp r^\delta$  (cf. Proposition 4.3), we get that

$$\sum_{j:(k,\ell,j) \in \text{Small}} r^\delta \ll \mu_{y_\rho}^u \left( \bigcup_{j:(k,\ell,j) \in \text{Small}} A_j \right).$$

Hence, Theorem 4.4 implies that the above sum is  $\mathcal{O}(\mu_{y_\rho}^u(W_\rho)(\epsilon_1+r)^\alpha/\epsilon_2^\alpha)$ , which concludes the proof since  $r \ll \epsilon_1$ .  $\square$

To apply Lemma 4.11, we need the following counting estimate on close by vectors of the form  $m_{\rho,\ell}^{-1} w_{\rho,\ell}$ . It is a consequence of Theorem 4.4.

**Lemma 4.12.** *For every  $k \in I_{\rho,t}$  and  $\eta > 0$ , we have*

$$\#\left\{\ell \in I_{\rho,t} : \left\| m_{\rho,\ell}^{-1} w_{\rho,\ell} - m_{\rho,k}^{-1} w_{\rho,k} \right\| < \|\xi\|^{-\eta} \right\} \ll (e^{-t} + \|\xi\|^{-\eta})^\alpha e^{\delta t}, \quad (4.37)$$

where  $\alpha > 0$  is the exponent provided by Theorem 4.4.

*Proof.* Let  $B_k$  denote the set on the left side of (4.37) and let  $\ell$  be some element of  $B_k$ . Then, by  $M$ -invariance of the norm, we have<sup>5</sup>

$$|\|w_{\rho,\ell}\| - \|w_{\rho,k}\|| \ll \|\xi\|^{-\eta}.$$

In particular, the vectors  $w_{\rho,\ell}$  with  $\ell \in B_k$  all belong to a neighborhood of width  $\ll \|\xi\|^{-\eta}$  of the sphere  $S$  of radius  $\|w_{\rho,k}\|_\infty$  around the origin in the norm metric.

The next ingredient is to note that the points  $w_{\rho,\ell}$  are separated by an amount  $\gg e^{-t}$ . This follows by a similar argument to the proof of [49, Proposition 9.13]<sup>6</sup>. In particular, there is  $\epsilon_1 \asymp (\|\xi\|^{-\eta} + e^{-t})$  and  $\epsilon_2 \asymp e^{-t}$  such that

$$\bigsqcup_{\ell \in B_k} N_{\epsilon_2}^- \cdot \exp(w_{\rho,\ell}) y_\rho \subseteq N_{\epsilon_1}^- \cdot S,$$

where  $N_{\epsilon_1}^- \cdot S$  is the  $\epsilon_1$ -neighborhood of  $S$ .

To conclude the proof, let  $\mu_{y_\rho}^s$  denote the shadow of the PS measure on  $N^- \cdot y_\rho$  defined analogously to the measures  $\mu_{y_\rho}^u$  in (2.4). The above discussion implies that

$$\#B_k \ll \frac{\mu_{y_\rho}^s(N_{\epsilon_1}^- \cdot S)}{\min \mu_{y_\rho}^s(N_{\epsilon_2}^- \cdot \exp(w_{\rho,\ell}) y_\rho)}.$$

By (4.12), the points  $\exp(w_{\rho,\ell}) \cdot y_\rho$  all belong to  $\Omega$ . In particular, by Proposition 4.3, we have

$$\mu_{y_\rho}^s(N_{e^{-t}}^- \cdot \exp(w_{\rho,\ell}) y_\rho) \asymp e^{-\delta t}.$$

<sup>5</sup>This estimate is again done to bypass studying the separation of the rotation matrices  $w_{\rho,\ell}$ .

<sup>6</sup>This proof is based on injectivity radius considerations along with the fact that  $g_t$  expands the stable manifold by  $e^t$  in backward time.

On the other hand, by [22, Lemma 3.8], we have that  $N_{\epsilon_1}^- \cdot S$  has measure  $\mathcal{O}(\epsilon_1^\alpha)^7$ . The lemma now follows.  $\square$

**Reduction to  $L^2$ -flattening.** To simplify our error terms, we make the following choices:

$$\eta = 1/12, \quad e^t = \|\xi\|^{1/24}. \quad (4.38)$$

In view of Lemmas 4.11 and 4.12, we introduce the following notation:

$$\text{Close}_\eta := \left\{ (k, \ell) \in I_{\rho, t}^2 : \left\| m_{\rho, \ell}^{-1} w_{\rho, \ell} - m_{\rho, k}^{-1} w_{\rho, k} \right\| < \|\xi\|^{-\eta} \right\}. \quad (4.39)$$

We also define the following set of indices parametrizing measures  $\mu_{y_\rho}^u$  for which many of the frequencies  $\mathbf{b}_{k, \ell}^j$  are close together. Let  $\mathcal{J}$  denote the index set for the indices  $j$  of the measures  $\mu_{y_\rho}^u$  and set

$$\text{Bad}_\eta := \left\{ j \in \mathcal{J} : \# \{ (k, \ell) \in I_{\rho, t}^2 : (k, \ell, j) \in \text{Small} \} > \|\xi\|^{-\alpha/48} \times (e^{\delta t} + \#I_{\rho, t}) \#I_{\rho, t} \right\}. \quad (4.40)$$

The following corollary allows us to estimate the part of the sum corresponding to  $\text{Bad}_\eta$ .

**Corollary 4.13.** *We have the following counting estimate on  $\text{Bad}_\eta$ :*

$$\sum_{j \in \text{Bad}_\eta} r^\delta \ll \|\xi\|^{-\alpha/48} \mu_{y_\rho}^u(W_\rho).$$

*Proof.* The corollary will follow from an application of Markov's inequality to the estimates in Lemmas 4.11 and 4.12 as follows. First, we note that

$$\begin{aligned} & \sum_{j \in \mathcal{J}} \sum_{(k, \ell) \in I_{\rho, t}^2} r^\delta \mathbb{1}_{\text{Small}}(k, \ell, j) \\ &= \underbrace{\sum_{(k, \ell) \in \text{Close}_\eta} \sum_{j \in \mathcal{J}} r^\delta \mathbb{1}_{\text{Small}}(k, \ell, j)}_{\text{(I)}} + \underbrace{\sum_{(k, \ell) \notin \text{Close}_\eta} \sum_{j \in \mathcal{J}} r^\delta \mathbb{1}_{\text{Small}}(k, \ell, j)}_{\text{(II)}}. \end{aligned}$$

Then, by Lemma 4.12 and our choices in (4.38), the first sum is estimated as follows:

$$\text{(I)} \ll (e^{-t} + \|\xi\|^{-\eta})^\alpha \times e^{\delta t} \#I_{\rho, t} \mu_{y_\rho}^u(W_\rho) \ll \|\xi\|^{-\alpha/24} \times e^{\delta t} \#I_{\rho, t} \mu_{y_\rho}^u(W_\rho).$$

For the second sum, we use Lemma 4.11 and the definition of  $\text{Close}_\eta$  to get

$$\text{(II)} \ll \|\xi\|^{(\eta-1/6)\alpha} e^{\alpha t} \times \#I_{\rho, t}^2 \mu_{y_\rho}^u(W_\rho) \ll \|\xi\|^{-\alpha/24} \times \#I_{\rho, t}^2 \mu_{y_\rho}^u(W_\rho).$$

Hence, the corollary follows by Markov's inequality.  $\square$

To simplify notation, we set

$$E_1 := \max \left\{ \|\xi\|^{-1/24}, \|\xi\|^{-\alpha/48} \right\}. \quad (4.41)$$

For  $w \in \mathbb{R}^d$ , we let

$$\nu_j := \mu_{y_\rho}^u \Big|_{N_1^+}, \quad \widehat{\nu}_j(w) := \int_{N_1^+} e^{i\langle w, n \rangle} d\nu_j(n). \quad (4.42)$$

Then, by (4.33) and Corollary 4.13, we obtain

$$\int_{W_\rho} |\Psi_\rho(t, n)|^2 d\mu_{y_\rho}^u = r^\delta \sum_{j \notin \text{Bad}_\eta} \sum_{k, \ell \in I_{\rho, t}} |\widehat{\nu}_j(\mathbf{b}_{k, \ell}^j)| + O \left( (\|\psi\|_{C^1} + 1) \times E_1 \times \#I_{\rho, t}^2 \mu_{y_\rho}^u(W_\rho) \right). \quad (4.43)$$

<sup>7</sup>Note that, similarly to the case of affine subspaces in Theorem 4.4, this estimate can be deduced from the fact that PS measures give 0 mass to proper subvarieties of the boundary using the argument in [50, Section 8].

For each  $j$ , the sum on the right side of the above estimate can be viewed as an average, when properly normalized, over Fourier coefficients of the measure  $\nu_j$ . Moreover, Corollary 4.13 guarantees that the frequencies  $\mathbf{b}_{k,\ell}^j$  are sampled from a well-separated set. Hence, this average can be estimated using the  $L^2$ -Flattening Theorem, Theorem 1.3.

**The role of  $L^2$ -flattening.** Let  $\eta_2 > 0$  be a small parameter to be chosen using Lemma 4.14 below. Note that the total mass of  $\nu_j$ , denoted  $|\nu_j|$ , is  $\mu_{y_\rho}^u(N_1^+)$ . For each  $k \in I_{\rho,t}$ , define the following set, which roughly speaking, consists of frequencies where  $\widehat{\nu}_j$  is large:

$$B(j, k, \eta_2) := \left\{ \ell \in I_{\rho,t} : |\widehat{\nu}_j(\mathbf{b}_{k,\ell}^j)| > \|\xi\|^{-\eta_2} |\nu_j| \right\}. \quad (4.44)$$

Then, splitting the sum over frequencies according to the size of the Fourier transform  $\widehat{\nu}_j$  and reversing our change variables to go back to integrating over  $A_j$ , we obtain

$$r^\delta \sum_{j \notin \text{Bad}_\eta} \sum_{k, \ell \in I_{\rho,t}} |\widehat{\nu}_j(\mathbf{b}_{k,\ell}^j)| \ll \left( \max_{j \notin \text{Bad}_\eta, k \in I_{\rho,t}} \#B(j, k, \eta_2) + \|\xi\|^{-\eta_2} \#I_{\rho,t} \right) \#I_{\rho,t} \mu_{y_\rho}^u(W_\rho), \quad (4.45)$$

The following key counting estimate for  $B(j, k, \eta_2)$  is a consequence of the  $L^2$ -flattening theorem, Theorem 1.3.

**Lemma 4.14.** *For every  $\varepsilon > 0$ , there is  $\eta_2 > 0$  such that for all  $j \notin \text{Bad}_\eta$  and  $k \in I_{\rho,t}$ , we have*

$$\#B(j, k, \eta_2) \ll_\varepsilon \|\xi\|^{\varepsilon - \alpha/96} \times \sqrt{(e^{\delta t} + \#I_{\rho,t}) \#I_{\rho,t}},$$

where  $\alpha > 0$  is the exponent provided by Theorem 4.4. Here,  $\eta_2$  is the constant provided by Theorem 1.3 (denoted by  $\delta$  in the notation of the theorem).

*Proof.* Recall the definition of the frequencies  $\mathbf{b}_{k,\ell}^j$  in (4.34). The rough idea behind the proof is that, since  $j \notin \text{Bad}_\eta$ , the frequencies  $\mathbf{b}_{k,\ell}^j$  are well-separated. This allows us to apply Theorem 1.3 to the set of frequencies where the Fourier transform is large to conclude that the sets  $B(j, k, \eta_2)$  are relatively small in size.

More precisely, Theorem 4.4 and Theorem 1.3 imply that there exists  $\eta_2 > 0$ , depending on  $\varepsilon$  (but not on the index  $j$ ), such that the set

$$Q := \left\{ \mathbf{b}_{k,\ell}^j : \ell \in B(j, k, \eta_2) \right\}$$

can be covered by  $\mathcal{O}_\varepsilon(\|\xi\|^\varepsilon)$  balls  $B_i$  of radius  $1/2$ .

Let  $\tilde{B}_i$  denote the set of indices  $\ell \in B(j, k, \eta_2)$  such that  $\mathbf{b}_{k,\ell}^j \in B_i$ . In particular, we have

$$\#B(j, k, \eta_2) \leq \sum_i \#\tilde{B}_i. \quad (4.46)$$

Moreover, we note that for  $\ell_1, \ell_2 \in \tilde{B}_i$ , we have that  $\mathbf{b}_{k,\ell_1}^j - \mathbf{b}_{k,\ell_2}^j = \mathbf{b}_{\ell_2,\ell_1}^j$ . Since  $B_i$  has radius  $1/2$ , we get that  $\|\mathbf{b}_{\ell_2,\ell_1}^j\| < 1$ . Hence, recalling the definition of the sets  $\text{Small}$  in (4.36), and letting  $\text{Small}_j$  denote the set of pairs  $(p, q) \in I_{\rho,t}^2$  with  $(p, q, j) \in \text{Small}$ , we obtain

$$\#\tilde{B}_i^2 \leq \#\text{Small}_j.$$

On the other hand, since  $j \notin \text{Bad}_\eta$ , then by definition, we have that

$$\#\text{Small}_j \leq \|\xi\|^{-\alpha/48} \times (e^{\delta t} + \#I_{\rho,t}) \#I_{\rho,t}.$$

Since the sum in (4.46) has at most  $\mathcal{O}_\varepsilon(\|\xi\|^\varepsilon)$  terms, this estimate completes the proof.  $\square$

**Combining estimates and concluding the proof.** Recall that  $\alpha$  is the exponent provided by Theorem 4.4. Let  $\eta_2 > 0$  be the exponent provided by Lemma 4.14 when applied with  $\varepsilon = \alpha/200$  and let  $\kappa$  be defined as follows:

$$\kappa = \min \{1/24, \alpha/200, \eta_2\}. \quad (4.47)$$

Then, by combining (4.25), (4.27), (4.43), (4.45), and Lemma 4.14, we obtain the following bound:

$$\begin{aligned} \int_{N_1^+} e^{i\langle \xi, n \rangle} \psi(n) d\mu_x^u(n) &\ll_{\Gamma} (\|\psi\|_{C^1} + 1) \times \|\xi\|^{-\kappa} \\ &\times \left( e^{-\delta t} \sum_{\rho} \mu_{y_{\rho}}^u(W_{\rho}) \# I_{\rho, t} + e^{-3\delta t/4} \sum_{\rho} \mu_{y_{\rho}}^u(W_{\rho}) (\# I_{\rho, t})^{3/4} \right). \end{aligned}$$

The first sum on the right side is  $\mathcal{O}_{\Gamma}(1)$  in light of (4.15) and the fact that  $\mu_{y_{\rho}}^u(W_{\rho}) \asymp m_{x_{\rho, \ell}}^u(W_{\ell})$  for all  $\ell \in I_{\rho, t}$ . That the second sum is also  $\mathcal{O}_{\Gamma}(1)$  is proved in the following lemma. This concludes the proof of Theorem 1.11 apart from Lemma 4.7 which is proved in the next section.

**Lemma 4.15.** *For every  $p \in (1, \infty)$ , we have that*

$$e^{-\delta t/p} \sum_{\rho \in \mathcal{P}_{\xi}} \mu_{y_{\rho}}^u(W_{\rho}) (\# I_{\rho, t})^{1/p} \ll_{\Gamma} 1.$$

*Proof.* Indeed, letting  $q$  be such that  $1/p + 1/q = 1$ , we obtain by Hölder's inequality that

$$e^{-\delta t/p} \sum_{\rho \in \mathcal{P}_{\xi}} \mu_{y_{\rho}}^u(W_{\rho}) (\# I_{\rho, t})^{1/p} \leq e^{-\delta t/p} \left( \sum_{\rho \in \mathcal{P}_{\xi}} \mu_{y_{\rho}}^u(W_{\rho}) \right)^{1/q} \times \left( \sum_{\rho \in \mathcal{P}_{\xi}} \mu_{y_{\rho}}^u(W_{\rho}) \# I_{\rho, t} \right)^{1/p}.$$

Since  $\mu_{y_{\rho}}^u(W_{\rho}) \asymp \mu_{x_{\rho, \ell}}^u(W_{\ell})$  for all  $\ell \in I_{\rho, t}$ , it follows by (4.15) that

$$e^{-\delta t} \sum_{\rho \in \mathcal{P}_{\xi}} \mu_{y_{\rho}}^u(W_{\rho}) \# I_{\rho, t} \ll_{\Gamma} 1.$$

Moreover, Proposition 4.3 implies that  $\mu_{y_{\rho}}^u(W_{\rho}) \asymp \|\xi\|^{-\delta/3}$ . Hence, the lemma follows in light of (4.5).  $\square$

**4.4. Explicit formula for stable holonomy maps and Proof of Lemma 4.7.** In this section, we give explicit formulas for the commutation relations between stable and unstable subgroups which we need for the proof of Lemma 4.7.

Consider the following quadratic form on  $\mathbb{R}^{d+2}$ : for  $x = (x_i) \in \mathbb{R}^{d+2}$ ,

$$Q(x) = 2x_0x_{d+1} - |x_1|^2 - \dots - |x_d|^2.$$

Let  $\mathrm{SO}_{\mathbb{R}}(Q) \cong \mathrm{SO}(d+1, 1)$  be the orthogonal group of  $Q$ ; i.e. the subgroup of  $\mathrm{SL}_{d+2}(\mathbb{R})$  preserving  $Q$ . Then, we have a surjective homomorphism  $\mathrm{SO}_{\mathbb{R}}(Q) \rightarrow G = \mathrm{Isom}^+(\mathbb{H}^{d+1})$  with finite kernel. The geodesic flow is induced by the diagonal group

$$A = \{g_t = \mathrm{diag}(e^t, I_d, e^{-t}) : t \in \mathbb{R}\},$$

where  $I_d$  denotes the identity matrix in dimension  $d$ .

For  $x \in \mathbb{R}^d$ , viewed as a row vector, we write  $x^t$  for its transpose. We let  $\|x\|^2 := x \cdot x$ , and  $x \cdot x$  denotes the sum of coordinate-wise products in the standard basis on  $\mathbb{R}^d$ . Hence,  $N^+$  can be parametrized as follows:

$$N^+ = \left\{ n^+(x) := \begin{pmatrix} 1 & x & \frac{\|x\|^2}{2} \\ \mathbf{0} & I_d & x^t \\ 0 & \mathbf{0} & 1 \end{pmatrix} : x \in \mathbb{R}^d \right\}. \quad (4.48)$$

The group  $N^-$  is parametrized by the transpose of the elements of  $N^+$ . Recall that  $M = \text{SO}_d(\mathbb{R})$  denotes the centralizer of  $A$  inside the standard maximal compact subgroup  $K \cong \text{SO}_{d+1}(\mathbb{R})$  of  $G$ . In particular,  $M$  is given by

$$M = \{m(O) := \text{diag}(1, O, 1) : O \in \text{SO}_d(\mathbb{R})\}.$$

Finally, we recall that the product map  $N^- \times A \times M \times N^+ \rightarrow G$  is a diffeomorphism near to the identity.

We are now ready for the proof. Recall from (4.17) that  $\phi_\ell(n)$  is defined to be the element of  $N^+$  satisfying  $np_{\rho,\ell}^- \in N^-AM\phi_\ell(n)$ . In particular,  $\phi_\ell^{-1}(n)$  is the unique element of  $N^+$  satisfying

$$n(p_{\rho,\ell}^-)^{-1} \in N^-AM\phi_\ell^{-1}(n).$$

Hence, in view of the explicit parametrization above, in order to compute  $\phi_\ell^{-1}(n)$ , it suffices to compute the top row of the matrix  $n(p_{\rho,\ell}^-)^{-1}$  and to note that

$$g_s n^+(x) = \begin{pmatrix} e^s & e^s x & \frac{e^s \|x\|^2}{2} \\ \mathbf{0} & \text{Id} & x^t \\ 0 & \mathbf{0} & e^{-s} \end{pmatrix},$$

for all  $s \in \mathbb{R}$  and  $x \in \mathbb{R}^d$ . This allows us to extract the  $N^+$  component from the top row of  $n(p_{\rho,\ell}^-)^{-1}$  by scaling it suitably so that the top left entry is 1. In particular, the claimed formula follows by a direct calculation upon recalling from (4.16) that  $p_{\rho,\ell}^- = \exp(w_{\rho,\ell})m_{\rho,\ell}gt_{\rho,\ell}$ .

## 5. SELF-CONFORMAL MEASURES: PROOF OF THEOREM 1.14

Throughout this section we fix  $\Phi$ ,  $\Sigma_A$ , and  $\psi$  satisfying the assumptions of Theorem 1.14. To simplify our notation we will denote the Gibbs measure  $\mu_\psi$  by  $\mu$ .

**5.1. The Gibbs property and bounding the Fourier transform by an average.** Let  $\xi \in \mathbb{R}^d \setminus \{0\}$ . The first step in our proof is to bound  $|\widehat{\mu}(\xi)|$  from above by an average of the Fourier transform over multiple frequencies using (2.1).

Let us define the following partition of  $\Sigma_A$  depending upon  $\xi$ :

$$\mathcal{A}_\xi := \left\{ \mathbf{b} \in \mathcal{W}_A : \sup_{x \in [0,1]^d} \|Df_{\mathbf{b}}(x)\| \leq |\xi|^{-2/3} \text{ and } \sup_{x \in [0,1]^d} \|Df_{\mathbf{b}^-}(x)\| > |\xi|^{-2/3} \right\}.$$

For a word  $\mathbf{b} \in \mathcal{W}_A$  we let  $\mu_{\mathbf{b}}$  denote the normalised restriction of  $\mu_{\mathbf{b}}$  to the cylinder  $X_{\mathbf{b}}$ , i.e.

$$\mu_{\mathbf{b}} := \frac{\mu|_{X_{\mathbf{b}}}}{\mu(X_{\mathbf{b}})}.$$

For each  $\mathbf{b} \in \mathcal{W}_A$  we also fix a point  $x_{\mathbf{b}} \in X_{\mathbf{b}}$  arbitrarily. Given  $\xi \in \mathbb{R}^d \setminus \{0\}$  and a word  $\mathbf{b} \in \mathcal{W}_A$  we let

$$\mu_{\mathbf{b},\xi} := S_{\|\xi\|^{2/3}}^* \mu_{\mathbf{b}},$$

where  $S_{\|\xi\|^{2/3}} : \mathbb{R}^d \rightarrow \mathbb{R}^d$  is given by  $S_{\|\xi\|^{2/3}}(x) := \|\xi\|^{2/3} \cdot x$ . We then have the following key decomposition of the Fourier transform using the Fourier transforms of  $\mu_{\mathbf{b},\xi}$  at random frequencies sampled using the derivative matrices:

**Proposition 5.1.** *If  $\xi \in \mathbb{R}^d \setminus \{0\}$  and  $n(\xi) = \lfloor c \log \|\xi\| \rfloor$  for some constant  $c > 0$ , we have*

$$|\widehat{\mu}(\xi)| \leq \sum_{\mathbf{b} \in \mathcal{A}_\xi} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \sim \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \left| \widehat{\mu_{\mathbf{b},\xi}} \left( (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right) \right| + \mathcal{O}(\|\xi\|^{-1/3}).$$

*Proof.* Applying (2.1) and then splitting our integral over cylinder sets  $X_{\mathbf{b}}$  we have

$$\begin{aligned} \int e^{2\pi i \langle \xi, x \rangle} d\mu(x) &= \int \sum_{\substack{\mathbf{a} \in \mathcal{A}^n(\xi) \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) e^{2\pi i \langle \xi, f_{\mathbf{a}}(x) \rangle} d\mu(x) \\ &= \sum_{\mathbf{b} \in \mathcal{A}_{\xi}} \sum_{\substack{\mathbf{a} \in \mathcal{A}^n(\xi) \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} \int_{X_{\mathbf{b}}} w_{\mathbf{a}}(x) e^{2\pi i \langle \xi, f_{\mathbf{a}}(x) \rangle} d\mu(x) \\ &= \sum_{\mathbf{b} \in \mathcal{A}_{\xi}} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^n(\xi) \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} \int w_{\mathbf{a}}(x) e^{2\pi i \langle \xi, f_{\mathbf{a}}(x) \rangle} d\mu_{\mathbf{b}}(x) \end{aligned}$$

In the penultimate line we used that for a point  $x \in X_{\mathbf{b}}$  we have  $\mathbf{a} \rightsquigarrow x$  if and only if  $\mathbf{a} \rightsquigarrow \mathbf{b}$ . Using our assumption  $P(\psi) = 0$ , we see that the Gibbs property (1.3) gives a constant  $C > 0$  such that  $|w_{\mathbf{a}}(x)| \leq C$  for all  $x \in X$ . Using this inequality together with the fact our potential  $\psi$  is  $C^1$  we can show that for any  $x, y \in X_A$  we have

$$|w_{\mathbf{a}}(x) - w_{\mathbf{a}}(y)| = |w_{\mathbf{a}}(x)| |1 - e^{S_n \psi(f_{\mathbf{a}}(y)) - S_n \psi(f_{\mathbf{a}}(x))}| \ll |x - y|. \quad (5.1)$$

Thus for any  $\mathbf{b} \in \mathcal{A}_{\xi}$  and  $x \in X_{\mathbf{b}}$  we have

$$|w_{\mathbf{a}}(x) - w_{\mathbf{a}}(x_{\mathbf{b}})| \ll \|\xi\|^{-2/3}$$

since  $|x - x_{\mathbf{b}}| \ll \|\xi\|^{-2/3}$  by Lemma 2.2. Thus

$$\begin{aligned} &\sum_{\mathbf{b} \in \mathcal{A}_{\xi}} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^n(\xi) \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} \int w_{\mathbf{a}}(x) e^{2\pi i \langle \xi, f_{\mathbf{a}}(x) \rangle} d\mu_{\mathbf{b}}(x) \\ &= \sum_{\mathbf{b} \in \mathcal{A}_{\xi}} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^n(\xi) \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \int e^{2\pi i \langle \xi, f_{\mathbf{a}}(x) \rangle} d\mu_{\mathbf{b}}(x) + \mathcal{O}(\|\xi\|^{-2/3}). \end{aligned}$$

Now, towards the claim, we need to linearize  $f_{\mathbf{a}}(x)$  for  $x \in X_{\mathbf{b}}$  using the value of the derivative at  $x_{\mathbf{b}} \in X_{\mathbf{b}}$ . For this purpose, given  $t \in [0, 1]$  let  $x_t = x_{\mathbf{b}} + t(x - x_{\mathbf{b}})$ . Using this notation we have

$$f_{\mathbf{a}}(x) = f_{\mathbf{a}}(x_{\mathbf{b}}) + \left( \int_0^1 D_{x_t} f_{\mathbf{a}} dt \right) (x - x_{\mathbf{b}}). \quad (5.2)$$

By Lemma 2.1 we have

$$\|D_{x_t} f_{\mathbf{a}} - D_{x_{\mathbf{b}}} f_{\mathbf{a}}\| \ll \sup_{x \in [0, 1]^d} \|D f_{\mathbf{a}}(x)\| \cdot \|x_t - x_{\mathbf{b}}\| = \mathcal{O}(\|\xi\|^{-2/3})$$

because  $\|x_t - x_{\mathbf{b}}\| \ll \|\xi\|^{-2/3}$  for all  $t \in [0, 1]$ . This in turn implies that for  $x \in X_{\mathbf{b}}$  we have

$$\left\| \left( \int_0^1 D_{x_t} f_{\mathbf{a}} - D_{x_{\mathbf{b}}} f_{\mathbf{a}} dt \right) (x - x_{\mathbf{b}}) \right\| \ll \|\xi\|^{-2/3} \|x - x_{\mathbf{b}}\| \ll \|\xi\|^{-4/3}.$$

Thus by (5.2) we have

$$f_{\mathbf{a}}(x) = f_{\mathbf{a}}(x_{\mathbf{b}}) + D_{x_{\mathbf{b}}} f_{\mathbf{a}}(x - x_{\mathbf{b}}) + \mathcal{O}(\|\xi\|^{-4/3}).$$

By the 1-Lipschitz property of  $x \mapsto e^{ix}$  and Cauchy-Schwartz inequality, we have for the matrix  $A_{\mathbf{b}} = D_{x_{\mathbf{b}}} f_{\mathbf{a}}$  and a vector  $v_{\mathbf{b}} = f_{\mathbf{a}}(x_{\mathbf{b}}) - D_{x_{\mathbf{b}}} f_{\mathbf{a}} x_{\mathbf{b}}$  that

$$\left| \int e^{2\pi i \langle \xi, f_{\mathbf{a}}(x) \rangle} - e^{2\pi i \langle \xi, A_{\mathbf{b}} x + v_{\mathbf{b}} \rangle} d\mu_{\mathbf{b}}(x) \right| \ll \|\xi\| \sup_{x \in X_{\mathbf{b}}} \|f_{\mathbf{a}}(x) - A_{\mathbf{b}} x - v_{\mathbf{b}}\| \ll \|\xi\|^{-1/3}.$$

Thus we have proved:

$$\int e^{2\pi i \langle \xi, x \rangle} d\mu(x) = \sum_{\mathbf{b} \in \mathcal{A}_\xi} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \int e^{2\pi i \langle \xi, f_{\mathbf{a}}(x_{\mathbf{b}}) + D_{x_{\mathbf{b}}} f_{\mathbf{a}}(x - x_{\mathbf{b}}) \rangle} d\mu_{\mathbf{b}}(x) + \mathcal{O}(\|\xi\|^{-1/3}).$$

Taking absolute values and applying the triangle inequality this implies:

$$\int e^{2\pi i \langle \xi, x \rangle} d\mu(x) \leq \sum_{\mathbf{b} \in \mathcal{A}_\xi} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \left| \int e^{2\pi i \langle \xi, D_{x_{\mathbf{b}}} f_{\mathbf{a}} x \rangle} d\mu_{\mathbf{b}}(x) \right| + \mathcal{O}(\|\xi\|^{-1/3}).$$

Recalling the definition of  $\mu_{\mathbf{b}, \xi} = S_{\|\xi\|^{2/3}}^* \mu_{\mathbf{b}}$ , where  $S_{\|\xi\|^{2/3}}(x) = \|\xi\|^{2/3} \cdot x$ , and taking the transpose of the matrices appearing in our inner product yields:

$$\begin{aligned} \int e^{2\pi i \langle \xi, x \rangle} d\mu(x) &\leq \sum_{\mathbf{b} \in \mathcal{A}_\xi} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \left| \int e^{2\pi i \langle (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}}, x \rangle} d\mu_{\mathbf{b}, \xi}(x) \right| + \mathcal{O}(\|\xi\|^{-1/3}) \\ &= \sum_{\mathbf{b} \in \mathcal{A}_\xi} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \left| \widehat{\mu_{\mathbf{b}, \xi}} \left( (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right) \right| + \mathcal{O}(\|\xi\|^{-1/3}). \end{aligned}$$

This completes the proof.  $\square$

**5.2. Affine non-concentration of the averaged measures.** To successfully apply Theorem 1.3 to the measures  $\mu_{\mathbf{b}, \xi}$  appearing in Proposition 5.1, we need to show that they are uniformly affinely non-concentrated and that the non-concentration parameters can be taken to be independent of  $\mathbf{b}$  and  $\xi$ . This property is guaranteed by the following proposition.

**Proposition 5.2.** *Assume that  $\mu$  is uniformly affinely non-concentrated. Then there exists  $\delta' : (0, \infty) \rightarrow (0, \infty)$  satisfying the following:*

- (1)  $\delta'(\epsilon) \rightarrow 0$  as  $\epsilon \rightarrow 0$ .
- (2) *For every  $\xi \neq 0$ ,  $\mathbf{b} \in \mathcal{A}_\xi$ ,  $x \in \text{supp}(\mu_{\mathbf{b}, \xi})$ ,  $0 < r \leq 1$ ,  $\epsilon > 0$  and affine hyperplane  $W < \mathbb{R}^d$  we have*

$$\mu_{\mathbf{b}, \xi}(W^{(\epsilon r)} \cap B(x, r)) \leq \delta'(\epsilon) \mu(B(x, r))$$

*Proof.* Fix  $\xi \neq 0$  and  $\mathbf{b} \in \mathcal{A}_\xi$ . Let  $r \in (0, 1]$  and  $\epsilon > 0$ . Then for any  $x \in \text{supp}(\mu_{\mathbf{b}, \xi})$  and affine hyperplane  $W < \mathbb{R}^d$  there exists  $V < \mathbb{R}^d$  and  $y \in X_{\mathbf{b}}$  such that

$$\begin{aligned} \mu_{\mathbf{b}, \xi}(W^{(\epsilon r)} \cap B(x, r)) &= \mu_{\mathbf{b}}(V^{(\epsilon \cdot \|\xi\|^{-2/3} r)} \cap B(y, r \cdot \|\xi\|^{-2/3})) \\ &= \frac{\mu|_{X_{\mathbf{b}}}(V^{(\epsilon \cdot \|\xi\|^{-2/3} r)} \cap B(y, r \cdot \|\xi\|^{-2/3}))}{\mu(X_{\mathbf{b}})} \\ &\leq \frac{\mu(V^{(\epsilon \cdot \|\xi\|^{-2/3} r)} \cap B(y, r \cdot \|\xi\|^{-2/3}))}{\mu(X_{\mathbf{b}})} \\ &\leq \frac{\delta(\epsilon) \mu(B(y, r \cdot \|\xi\|^{-2/3}))}{\mu(X_{\mathbf{b}})}. \end{aligned}$$

Where in the final line we let  $\delta : (0, \infty) \rightarrow (0, \infty)$  be as in the definition of uniformly affinely non-concentrated. Therefore to complete our proof it remains to show that

$$\frac{\mu(B(y, r \cdot \|\xi\|^{-2/3}))}{\mu(X_{\mathbf{b}})} \ll \mu_{\mathbf{b}, \xi}(B(x, r)),$$

which by the definition of  $\mu_{\mathbf{b}, \xi}$  is equivalent to

$$\mu(B(y, r \cdot \|\xi\|^{-2/3})) \ll \mu(X_{\mathbf{b}} \cap B(y, r \cdot \|\xi\|^{-2/3})). \quad (5.3)$$

Recall that  $\text{Diam}(X_{\mathbf{b}}) \asymp \|\xi\|^{-2/3}$  since  $\mathbf{b} \in \mathcal{A}_\xi$ . It follows from this observation and the Strong Separation Condition that there exists  $k \in \mathbb{N}$  depending only upon the underlying IFS such that one of two cases must occur:

- (1)  $X_{\mathbf{ba}} \subset B(y, r\|\xi\|^{-2/3}) \cap X_A \subset X_{b_1, \dots, b_{|\mathbf{b}|-k}}$  for some  $\mathbf{a} \in \mathcal{A}^k$ .
- (2) There exists  $\mathbf{a} \in \mathcal{A}^*$  and  $\mathbf{a}^* \in \mathcal{A}^k$  such that  $X_{\mathbf{baa}^*} \subset B(y, r\|\xi\|^{-2/3}) \cap X_A \subset X_{\mathbf{ba}}$ .

In the first case, it follows from the quasi-Bernoulli property of  $\mu$  (2.2) that

$$\mu(X_{\mathbf{b}} \cap B(y, r \cdot \|\xi\|^{-2/3})) \geq \mu(X_{\mathbf{ba}}) \gg \mu(X_{b_1, \dots, b_{|\mathbf{b}|-k}}) \geq \mu(B(y, r\|\xi\|^{-2/3})).$$

Again appealing to the quasi-Bernoulli property of  $\mu$ , in the second case we have

$$\mu(X_{\mathbf{b}} \cap B(y, r \cdot \|\xi\|^{-2/3})) \geq \mu(X_{\mathbf{baa}^*}) \gg \mu(X_{\mathbf{ba}}) \geq \mu(B(y, r\|\xi\|^{-2/3})).$$

Thus (5.3) holds and our proof is complete.  $\square$

**5.3. Spectral gap and non-concentration of frequencies.** We now fix the parameter  $c$  appearing in Proposition 5.1. We assume that  $c$  is sufficiently small that

$$\left\| (Df_{\mathbf{a}}(x_{\mathbf{b}}))^T \frac{\xi}{\|\xi\|^{2/3}} \right\| \geq \|\xi\|^{1/6} \quad (5.4)$$

for any  $\mathbf{a} \in \mathcal{A}^{n(\xi)}$  and  $\mathbf{b} \in \mathcal{W}_A$  with  $\mathbf{a} \rightsquigarrow \mathbf{b}$ . The final part we need for our proof of Theorem 1.14 is the following proposition. It is in the proof of this proposition that we will use our spectral gap assumption.

**Proposition 5.3.** *There exists  $\gamma > 0$  such that for any  $\mathbf{b} \in \mathcal{A}_\xi$  and integer  $n \in \mathbb{N}$  satisfying  $\|\xi\|^{1/6} \leq n \leq \|\xi\|^{1/3}$  we have*

$$\sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \chi_{[n, n+1]} \left( \left\| (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right\| \right) = \mathcal{O}(\|\xi\|^{-\gamma}).$$

*Proof.* Let  $n \in \mathbb{N}$  satisfy  $\|\xi\|^{1/6} \leq n \leq \|\xi\|^{1/3}$ . Observe that

$$\begin{aligned} \left\| (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right\| \in [n, n+1] &\Leftrightarrow \left| \prod_{i=1}^{n(\xi)} \lambda_{a_i}(f_{a_{i+1} \dots a_n}(x_{\mathbf{b}})) \right| \cdot \left\| \frac{\xi}{\|\xi\|^{2/3}} \right\| \in [n, n+1] \\ &\Leftrightarrow \log \left| \prod_{i=1}^{n(\xi)} \lambda_{a_i}(f_{a_{i+1} \dots a_n}(x_{\mathbf{b}})) \right| \in I_{n, \xi} \end{aligned}$$

for

$$I_{n, \xi} = [\log(\|\xi\|^{-1/3}n), \log(\|\xi\|^{-1/3}(n+1))].$$

Thus we have

$$\sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \chi_{[n, n+1]} \left( \left\| (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right\| \right) = \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \chi_{I_{n, \xi}} \left( \log \left| \prod_{i=1}^{n(\xi)} \lambda_{a_i}(f_{a_{i+1} \dots a_n}(x_{\mathbf{b}})) \right| \right) \quad (5.5)$$

By the mean value theorem

$$|I_{n, \xi}| \asymp \frac{1}{n} \leq |\xi|^{-1/6}.$$

We replace  $I_{n, \xi}$  with an interval  $\tilde{I}_{n, \xi}$  that contains  $I_{n, \xi}$  and is of length  $|\xi|^{-\delta}$  where

$$\delta < \min \{1/6, -c \log \rho\}.$$

Here  $\rho \in (0, 1)$  is as in the definition of spectral gap and  $c$  is as in (5.4). We let  $h_{n,\xi}$  be a smooth mollifier approximating  $\chi_{\tilde{I}_{n,\xi}}$  such that  $h_{n,\xi} \geq \chi_{\tilde{I}_{n,\xi}}$ ,  $\|h_{n,\xi}\|_1 \ll |\xi|^{-\delta}$  and  $\|h''_{n,\xi}\|_1 \ll |\xi|^\delta$  as in [12, 77]. Using Fourier inversion we have the following

$$\begin{aligned}
& \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \chi_{I_{n,\xi}} \left( \log \left| \prod_{i=1}^{n(\xi)} \lambda_{a_i}(f_{a_{i+1} \dots a_n}(x_{\mathbf{b}})) \right| \right) \\
& \leq \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) h_{n,\xi} \left( \log \left| \prod_{i=1}^{n(\xi)} \lambda_{a_i}(f_{a_{i+1} \dots a_n}(x_{\mathbf{b}})) \right| \right) \\
& = \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \int_{-\infty}^{\infty} \widehat{h_{n,\xi}}(\eta) e^{-2\pi i \eta \log |\prod_{i=1}^{n(\xi)} \lambda_{a_i}(f_{a_{i+1} \dots a_n}(x_{\mathbf{b}}))|} d\eta \\
& = \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \int_{-\infty}^{\infty} \widehat{h_{n,\xi}}(\eta) \prod_{i=1}^{n(\xi)} |\lambda_{a_i}(f_{a_{i+1} \dots a_n}(x_{\mathbf{b}}))|^{2\pi i \eta} d\eta \\
& = \int_{-\infty}^{\infty} \widehat{h_{n,\xi}}(\eta) \mathcal{L}_{2\pi i \eta}^{n(\xi)}(1)(x_{\mathbf{b}}) d\eta.
\end{aligned}$$

Let  $\Theta > 0$  be some parameter such that our spectral gap bound can be applied for  $|\eta| \geq \Theta$ . Splitting the integral above we have:

$$\begin{aligned}
\int_{-\infty}^{\infty} \widehat{h_{n,\xi}}(\eta) \mathcal{L}_{2\pi i \eta}^{n(\xi)}(1)(x_{\mathbf{b}}) d\eta &= \int_{|\eta| < \Theta} \widehat{h_{n,\xi}}(\eta) \mathcal{L}_{2\pi i \eta}^{n(\xi)}(1)(x_{\mathbf{b}}) d\eta + \int_{|\eta| \geq \Theta} \widehat{h_{n,\xi}}(\eta) \mathcal{L}_{2\pi i \eta}^{n(\xi)}(1)(x_{\mathbf{b}}) d\eta \\
&\ll |\widehat{h_{n,\xi}}(0)| + \int_{|\eta| \geq \Theta} \widehat{h_{n,\xi}}(\eta) \mathcal{L}_{2\pi i \eta}^{n(\xi)}(1)(x_{\mathbf{b}}) d\eta \\
&\ll |\xi|^{-\delta} + \int_{|\eta| \geq \Theta} \widehat{h_{n,\xi}}(\eta) \mathcal{L}_{2\pi i \eta}^{n(\xi)}(1)(x_{\mathbf{b}}) d\eta
\end{aligned}$$

Applying our spectral gap bound we have

$$\|\mathcal{L}_{2\pi i \eta}^{n(\xi)} 1\|_{\infty} \ll \rho^{n(\xi)} |\eta|^v$$

for some  $v \in (0, 1)$ . Since  $n(\xi) = \lfloor c \log \|\xi\| \rfloor$ , this gives us:

$$\begin{aligned}
\left| \int_{|\eta| \geq \Theta} \widehat{h_{n,\xi}}(\eta) \mathcal{L}_{2\pi i \eta}^{n(\xi)}(1)(x_{\mathbf{b}}) d\eta \right| &\leq \int_{|\eta| \geq \Theta} |\widehat{h_{n,\xi}}(\eta)| \rho^{n(\xi)} |\eta|^v d\eta \\
&\ll \|\xi\|^{c \log \rho} \int_{|\eta| \geq \Theta} |\widehat{h_{n,\xi}}(\eta)| |\eta|^v d\eta.
\end{aligned}$$

Using integration by parts it can be shown that

$$|\widehat{h_{n,\xi}}(\eta)| \ll \frac{1}{1 + |\eta|^2} (\|h_{n,\xi}\|_1 + \|h''_{n,\xi}\|_1).$$

Substituting this bound into the above yields:

$$\begin{aligned} \|\xi\|^{c \log \rho} \int_{|\eta| \geq \Theta} |\widehat{h_{n,\xi}}(\eta)| |\eta|^v d\eta &\leq \|\xi\|^{c \log \rho} \int_{|\eta| \geq \Theta} \frac{|\eta|^v}{1 + |\eta|^2} (\|h_{n,\xi}\|_1 + \|h''_{n,\xi}\|_1) d\eta \\ &\ll \|\xi\|^{c \log \rho} (\|h_{n,\xi}\|_1 + \|h''_{n,\xi}\|_1) \\ &\ll \|\xi\|^{\delta + c \log \rho}. \end{aligned}$$

In the penultimate line we used that  $v \in (0, 1)$  thus  $\int_{|\eta| \geq \Theta} \frac{|\eta|^v}{1 + |\eta|^2} d\eta \ll 1$ . In the final line we used that  $\|h_{n,\xi}\|_1 \leq \|\xi\|^{-\delta}$  and  $\|h''_{n,\xi}\|_1 \leq \|\xi\|^\delta$ . By the definition of  $\delta$  we know that  $\delta + c \log \rho < 0$ . This completes the proof of this proposition.  $\square$

**5.4. Proof of Theorem 1.14.** Equipped with the propositions above we can now prove Theorem 1.14. By Proposition 5.1, we have

$$|\widehat{\mu}(\xi)| \leq \sum_{\mathbf{b} \in \mathcal{A}_\xi} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \left| \widehat{\mu_{\mathbf{b},\xi}} \left( (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right) \right| + \mathcal{O}(\|\xi\|^{-1/3}).$$

Thus to prove our result it suffices to show that

$$\sum_{\mathbf{b} \in \mathcal{A}_\xi} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \left| \widehat{\mu_{\mathbf{b},\xi}} \left( (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right) \right| = \mathcal{O}(\|\xi\|^{-\eta}), \quad (5.6)$$

for some  $\eta > 0$ .

Let  $\gamma$  be as in Proposition 5.3. Applying Theorem 1.3, we can assert that there exists  $\tau > 0$  such that for each  $\mathbf{b} \in \mathcal{A}_\xi$ , if we let

$$\text{Bad}_{\xi, \mathbf{b}} := \left\{ n \in \mathbb{Z} : |n| \leq |\xi|^{1/3}, \exists \eta \in \mathbb{R}^d \text{ such that } \|\eta\| \in [n, n+1] \text{ and } |\widehat{\mu_{\mathbf{b},\xi}}(\eta)| \geq \|\xi\|^{-\tau} \right\}$$

then

$$\#\text{Bad}_{\xi, \mathbf{b}} = \mathcal{O}_\gamma(\|\xi\|^{\gamma/2}). \quad (5.7)$$

We emphasise that the underlying constants in (5.7) do not depend upon  $\mathbf{b}$  or  $\xi$ . This follows from Proposition 5.2 which asserts that the non-concentration parameters can be taken to be independent of  $\mathbf{b}$  and  $\xi$ . By our choice of  $c$  we know that

$$\left\| (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right\| \geq \|\xi\|^{1/6}$$

for all  $\mathbf{b} \in \mathcal{A}_\xi$  and  $\mathbf{a} \in \mathcal{A}^{n(\xi)}$  satisfying  $\mathbf{a} \rightsquigarrow \mathbf{b}$ . Therefore we can apply Proposition 5.3 to the frequencies appearing in (5.6). In particular, using this proposition and (5.7) we have

$$\begin{aligned}
& \sum_{\mathbf{b} \in A_\xi} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \left| \widehat{\mu_{\mathbf{b}, \xi}} \left( (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right) \right| \\
& \leq \sum_{\mathbf{b} \in A_\xi} \mu(X_{\mathbf{b}}) \sum_{n \in \text{Bad}_{\xi, \mathbf{b}}} \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \chi_{[n, n+1]} \left( \left\| (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right\| \right) \\
& \quad + \sum_{\mathbf{b} \in A_\xi} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \left| \widehat{\mu_{\mathbf{b}, \xi}} \left( (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right) \right| \chi_{(\cup_{n \in \text{Bad}_{\xi, \mathbf{b}}} [n, n+1])^c} \left( \left\| (D_{x_{\mathbf{b}}} f_{\mathbf{a}})^T \frac{\xi}{\|\xi\|^{2/3}} \right\| \right) \\
& \ll \sum_{\mathbf{b} \in A_\xi} \mu(X_{\mathbf{b}}) \sum_{n \in \text{Bad}_{\xi, \mathbf{b}}} \|\xi\|^{-\gamma} + \sum_{\mathbf{b} \in A_\xi} \mu(X_{\mathbf{b}}) \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \|\xi\|^{-\tau} \\
& \ll \left( \|\xi\|^{-\gamma/2} + \|\xi\|^{-\tau} \right).
\end{aligned}$$

In the final line we used the trivial estimates:

$$\sum_{\mathbf{b} \in A_\xi} \mu(X_{\mathbf{b}}) = 1 \quad \text{and} \quad \sum_{\substack{\mathbf{a} \in \mathcal{A}^{n(\xi)} \\ \mathbf{a} \rightsquigarrow \mathbf{b}}} w_{\mathbf{a}}(x_{\mathbf{b}}) \ll 1.$$

Therefore, (5.6) holds and our proof is complete.

## 6. NON-CONFORMAL SYSTEMS: PROOF OF THEOREM 1.21

**6.1. Outline of the argument.** We begin this section by outlining our proof of Theorem 1.21. Given a frequency  $\xi \in \mathbb{R}^d$ , we will begin by picking a dominant direction  $1 \leq i^* \leq d$  such that  $\max_{1 \leq i \leq d} |\xi_i| = |\xi_{i^*}|$ . Without loss of generality we can take  $i^* = 1$ . We will then disintegrate our stationary measure in this direction to express it as an integral of random measures. Each of these random measures will be of the form  $\mu_{\mathbf{b}} \times \delta_{x_{\mathbf{b}}}$ , where  $\mu_{\mathbf{b}}$  is a probability measure on  $[0, 1]$  and  $\delta_{x_{\mathbf{b}}}$  is simply a Dirac mass supported on a point  $x_{\mathbf{b}} \in [0, 1]^{d-1}$ .

The fact that our random measures are of this form and  $\max_{1 \leq i \leq d} |\xi_i| = |\xi_1|$  will mean that to prove Theorem 1.21, it is sufficient to show that a typical  $\mu_{\mathbf{b}}$  will have polynomial Fourier decay. This will be done by using the strategy employed throughout this paper.

To this end, we will bound  $\widehat{\mu_{\mathbf{b}}}(\xi_1)$  from above by an average of Fourier transforms over a set of measures and a set of frequencies. We will show that the measures appearing in this average are all uniformly affinely non-concentrated and the underlying non-concentration parameters can be taken to be independent of the measure (Proposition 6.4). We will also show that the set of frequencies appearing in this average typically satisfy a non-concentration property (Proposition 6.8). We will then apply Theorem 1.3 to prove the desired Fourier decay result.

The major difference in our proof when compared to Theorem 1.14 is that the measures  $\mu_{\mathbf{b}}$  are no longer self-conformal measures and only satisfy a weaker form of dynamical self-conformality (see Lemma 6.1). It is a consequence of this weaker form of self-conformality that we do not consider a single transfer operator in our analysis, but instead must consider a family of transfer operators and how they behave under random compositions. Here we will rely heavily on work on the first and third authors [12], which establishes that a spectral gap typically exists for these random compositions under the third assumption of Theorem 1.21.

The aforementioned technique of disintegrating a stationary measure in terms of simpler, albeit random, measures has its origins in a paper of Galicer et al [32]. It has subsequently been applied to problems relating to absolute continuity of self-similar measures [44, 75], normal numbers in fractal sets [1], and Fourier transforms of fractal measures [11].

**6.2. Beginning our proof and disintegrating  $\mu$ .** Let us begin our proof of Theorem 1.21 by fixing a restricted product IFS  $\{F_{\bar{a}}\}_{\bar{a} \in \mathcal{A}}$  satisfying the assumptions of this theorem and a probability vector  $\mathbf{p} = (p_{\bar{a}})_{\bar{a} \in \mathcal{A}}$ . To ease our notation we will denote the corresponding stationary measure by  $\mu$  instead of  $\mu_{\mathbf{p}}$ . Let us also fix  $\xi \in \mathbb{R}^d$ . Without loss of generality we will assume that  $\max_{1 \leq i \leq d} |\xi_i| = |\xi_1|$ . Thus to prove our result it suffices to show that

$$|\widehat{\mu}(\xi)| \ll |\xi_1|^{-\eta} \quad (6.1)$$

for some  $\eta > 0$ . Our first step towards proving (6.1) is to disintegrate our measure  $\mu$ . This is done in Proposition 6.2 after introducing the necessary notation.

Let  $\pi : \mathcal{A} \rightarrow \mathcal{A}_2 \times \cdots \times \mathcal{A}_d$  be the map given by  $\pi(a_1, \dots, a_d) = (a_2, \dots, a_d)$ . We define  $\mathcal{B} = \pi(\mathcal{A})$  and define a new probability vector  $\mathbf{q} = (q_{\bar{b}})_{\bar{b} \in \mathcal{B}}$  by the formula

$$q_{\bar{b}} = \sum_{\bar{a} : \pi(\bar{a}) = \bar{b}} p_{\bar{a}}.$$

We let  $Q$  be the probability measure on  $\mathcal{B}^{\mathbb{N}}$  corresponding to  $\mathbf{q}$ . Given  $\bar{b} = (b_1, \dots, b_{d-1}) \in \mathcal{B}$  we define  $\tilde{F}_{\bar{b}} : [0, 1]^{d-1} \rightarrow [0, 1]^{d-1}$  to be the contraction given by

$$\tilde{F}_{\bar{b}}(x_1, \dots, x_{d-1}) = (f_{b_1}^{(2)}(x_1), \dots, f_{b_{d-1}}^{(d)}(x_{d-1})).$$

Given  $\bar{b} \in \mathcal{B}$ , we let  $\mathcal{A}_{\bar{b}} = \{a \in \mathcal{A}_1 : (a, \bar{b}) \in \mathcal{A}\}$ . For  $\mathbf{b} = (\bar{b}_i) \in \mathcal{B}^{\mathbb{N}}$ , we also define  $\Sigma_{\mathbf{b}} = \prod_{i=1}^{\infty} \mathcal{A}_{\bar{b}_i}$  and let  $x_{\mathbf{b}} \in [0, 1]^{d-1}$  be the point given by the formula

$$x_{\mathbf{b}} := \lim_{n \rightarrow \infty} (\tilde{F}_{\bar{b}_1} \circ \cdots \circ \tilde{F}_{\bar{b}_n})(0)$$

We define a Bernoulli measure on  $\Sigma_{\mathbf{b}}$  according to the rule

$$m_{\mathbf{b}}([a_1, \dots, a_n]) = \prod_{i=1}^n \frac{p_{a_i \bar{b}_i}}{q_{\bar{b}_i}}$$

and  $\pi_{\mathbf{b}} : \Sigma_{\mathbf{b}} \rightarrow [0, 1]$  according to the rule

$$\pi_{\mathbf{b}}(\mathbf{a}) = \lim_{n \rightarrow \infty} (f_{a_1}^{(1)} \circ \cdots \circ f_{a_n}^{(1)})(0).$$

We let

$$X_{\mathbf{b}} := \pi_{\mathbf{b}}(\Sigma_{\mathbf{b}}) \text{ and } \mu_{\mathbf{b}} := \pi_{\mathbf{b}} m_{\mathbf{b}}.$$

We remark that it follows from the second assumption in Theorem 1.21 that  $\#\mathcal{A}_{\bar{b}} \geq 2$  for all  $\bar{b} \in \mathcal{B}$ . Consequently, there exists  $\gamma \in (0, 1)$  such that

$$\max_{\bar{b} \in \mathcal{B}} \max_{a \in \mathcal{A}_{\bar{b}}} \frac{p_{a\bar{b}}}{q_{\bar{b}}} \leq \gamma. \quad (6.2)$$

We let  $\sigma : \mathcal{B}^{\mathbb{N}} \rightarrow \mathcal{B}^{\mathbb{N}}$  be the left shift map given by  $\sigma((\bar{b}_i)_{i=1}^{\infty}) = (\bar{b}_{i+1})_{i=1}^{\infty}$ . The following lemma shows that the measures  $\mu_{\mathbf{b}}$  exhibit a form of dynamical self-conformality.

**Lemma 6.1.** *Let  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$ . Then for all  $n \in \mathbb{N}$  we have*

$$\mu_{\mathbf{b}} = \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_n}} m_{\mathbf{b}}([\mathbf{a}]) f_{\mathbf{a}} \mu_{\sigma^n(\mathbf{b})}.$$

*Proof.* Let  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$  and  $n \in \mathbb{N}$ . By considering  $\mu_{\mathbf{b}}$  and  $\mu_{\sigma^n(\mathbf{b})}$  as weak star limits of weighted Dirac masses we have

$$\begin{aligned}
\mu_{\mathbf{b}} &= \lim_{M \rightarrow \infty} \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_M}} m_{\mathbf{b}}([\mathbf{a}]) \delta_{f_{\mathbf{a}}(0)} \\
&= \lim_{M \rightarrow \infty} \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{a}' \in \mathcal{A}_{\bar{b}_{n+1}} \times \dots \times \mathcal{A}_{\bar{b}_M}} m_{\sigma^n(\mathbf{b})}([\mathbf{a}']) \delta_{f_{\mathbf{a}\mathbf{a}'}(0)} \\
&= \lim_{M \rightarrow \infty} \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n}} m_{\mathbf{b}}([\mathbf{a}]) f_{\mathbf{a}} \left( \sum_{\mathbf{a}' \in \mathcal{A}_{\bar{b}_{n+1}} \times \dots \times \mathcal{A}_{\bar{b}_M}} m_{\sigma^n(\mathbf{b})}([\mathbf{a}']) \delta_{f_{\mathbf{a}'}(0)} \right) \\
&= \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n}} m_{\mathbf{b}}([\mathbf{a}]) f_{\mathbf{a}} \mu_{\sigma^n(\mathbf{b})}.
\end{aligned}$$

□

The following proposition is our aforementioned disintegration result.

**Proposition 6.2.** *We have*

$$\mu = \int_{\mathcal{B}^{\mathbb{N}}} \mu_{\mathbf{b}} \times \delta_{x_{\mathbf{b}}} dQ.$$

*Proof.* Consider the Borel probability measure  $\tilde{\mu}$  given by

$$\tilde{\mu} = \int_{\mathcal{B}^{\mathbb{N}}} \mu_{\mathbf{b}} \times \delta_{x_{\mathbf{b}}} dQ.$$

We will show that  $\tilde{\mu}$  satisfies

$$\tilde{\mu} = \sum_{\bar{a} \in \mathcal{A}} p_{\bar{a}} F_{\bar{a}} \tilde{\mu}. \quad (6.3)$$

Since the stationary measure  $\mu$  is the unique Borel probability measure satisfying this equation, our result will follow.

Using Lemma 6.1, we have

$$\begin{aligned}
\tilde{\mu} &= \int_{\mathcal{B}^{\mathbb{N}}} \mu_{\mathbf{b}} \times \delta_{x_{\mathbf{b}}} dQ = \int_{\mathcal{B}^{\mathbb{N}}} \sum_{a \in \mathcal{A}_{\bar{b}_1}} m_{\mathbf{b}}([a]) f_a \mu_{\sigma(\mathbf{b})} \times \delta_{x_{\mathbf{b}}} dQ \\
&= \int_{\mathcal{B}^{\mathbb{N}}} \sum_{a \in \mathcal{A}_{\bar{b}_1}} m_{\mathbf{b}}([a]) f_a \mu_{\sigma(\mathbf{b})} \times \tilde{F}_{\bar{b}_1} \delta_{x_{\sigma(\mathbf{b})}} dQ \\
&= \int_{\mathcal{B}^{\mathbb{N}}} \sum_{a \in \mathcal{A}_{\bar{b}_1}} m_{\mathbf{b}}([a]) F_{a\bar{b}_1} \left( \mu_{\sigma(\mathbf{b})} \times \delta_{x_{\sigma(\mathbf{b})}} \right) dQ \\
&= \sum_{\bar{b} \in \mathcal{B}} \sum_{a \in \mathcal{A}_{\bar{b}}} \frac{p_{a\bar{b}}}{q_{\bar{b}}} \int_{\mathbf{b}: \bar{b}_1 = \bar{b}} F_{a\bar{b}} \left( \mu_{\sigma(\mathbf{b})} \times \delta_{x_{\sigma(\mathbf{b})}} \right) dQ.
\end{aligned}$$

In the final line we used that

$$m_{\mathbf{b}}([a]) = \frac{p_{a\bar{b}}}{q_{\bar{b}}}$$

if  $\bar{b}_1 = \bar{b}$ . Now using the fact that  $Q$  is a  $\sigma$ -invariant product measure and  $F_{a\bar{b}}(\mu_{\sigma(\mathbf{b})} \times \delta_{x_{\sigma(\mathbf{b})}})$  does not depend on the first digit of  $\mathbf{b}$ , we have

$$\begin{aligned} \sum_{\bar{b} \in \mathcal{B}} \sum_{a \in \mathcal{A}_{\bar{b}}} \frac{p_{a\bar{b}}}{q_{\bar{b}}} \int_{\mathbf{b}: \bar{b}_1 = \bar{b}} F_{a\bar{b}}(\mu_{\sigma(\mathbf{b})} \times \delta_{x_{\sigma(\mathbf{b})}}) dQ &= \sum_{\bar{b} \in \mathcal{B}} \sum_{a \in \mathcal{A}_{\bar{b}}} \frac{p_{a\bar{b}}}{q_{\bar{b}}} \times q_{\bar{b}} \int_{\mathcal{B}^{\mathbb{N}}} F_{a\bar{b}}(\mu_{\mathbf{b}} \times \delta_{x_{\mathbf{b}}}) dQ \\ &= \sum_{\bar{b} \in \mathcal{B}} \sum_{a \in \mathcal{A}_{\bar{b}}} p_{a\bar{b}} \int_{\mathcal{B}^{\mathbb{N}}} F_{a\bar{b}}(\mu_{\mathbf{b}} \times \delta_{x_{\mathbf{b}}}) dQ \\ &= \sum_{\bar{a} \in \mathcal{A}} p_{\bar{a}} \int_{\mathcal{B}^{\mathbb{N}}} F_{\bar{a}}(\mu_{\mathbf{b}} \times \delta_{x_{\mathbf{b}}}) dQ \\ &= \sum_{\bar{a} \in \mathcal{A}} p_{\bar{a}} F_{\bar{a}} \tilde{\mu}. \end{aligned}$$

Thus,  $\tilde{\mu}$  satisfies (6.3) and our proof is complete.  $\square$

We finish this subsection by recording one consequence of Proposition 6.2 with regards to the Fourier transform of  $\mu$ :

$$\begin{aligned} |\widehat{\mu}(\xi)| &= \left| \int e^{2\pi i \langle \xi, x \rangle} d\mu \right| = \left| \int_{\mathcal{B}^{\mathbb{N}}} \int e^{2\pi i \langle \xi, x \rangle} d(\mu_{\mathbf{b}} \times \delta_{x_{\mathbf{b}}}) dQ(\mathbf{b}) \right| \\ &= \left| \int_{\mathcal{B}^{\mathbb{N}}} e^{2\pi i \langle (\xi_i)_{i=2}^d, x_{\mathbf{b}} \rangle} \int e^{2\pi i \xi_1 x} d\mu_{\mathbf{b}} dQ(\mathbf{b}) \right| \\ &\leq \int_{\mathcal{B}^{\mathbb{N}}} |\widehat{\mu_{\mathbf{b}}}(\xi_1)| dQ(\mathbf{b}). \end{aligned} \quad (6.4)$$

The significance of (6.4) is that it tells us that to prove (6.18) it is sufficient to show that outside of a set of small  $Q$  measure, each  $\mu_{\mathbf{b}}$  has polynomial Fourier decay.

**6.3. Bounding the Fourier transform by an average.** Let us now focus on  $\widehat{\mu_{\mathbf{b}}}(\xi_1)$ . We show in Proposition 6.3 that  $\widehat{\mu_{\mathbf{b}}}(\xi_1)$  is bounded above by an expression involving an average of Fourier transforms. Importantly Theorem 1.3 can be meaningfully applied to this average.

To this end, we let

$$n(\xi_1) = \lfloor c \log |\xi_1| \rfloor \quad (6.5)$$

for some  $c > 0$  chosen to be sufficiently small that

$$\min_{\mathbf{a} \in \mathcal{A}_1^{n(\xi_1)}} \min_{x \in [0,1]} |f_{\mathbf{a}}^{(1)}(x)| \geq \xi_1^{-1/3}. \quad (6.6)$$

To each  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$  we associate the set

$$\mathcal{A}_{\xi_1, \mathbf{b}} := \left\{ \mathbf{c} \in \bigcup_{j=1}^{\infty} \prod_{l=1}^j \mathcal{A}_{\bar{b}_{n(\xi_1)+l}} : \text{Diam}(f_{\mathbf{c}}([0,1])) < \xi_1^{-2/3}, \text{Diam}(f_{\mathbf{c}^-}([0,1])) \geq \xi_1^{-2/3} \right\}.$$

For each  $\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}$  we pick  $x_{\mathbf{c}} \in \pi_{\sigma^{n(\xi_1)}(\mathbf{b})}([c])$ . Given  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$  and  $\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}$  we let  $\mu_{\mathbf{c}, \sigma^{n(\xi_1)}(\mathbf{b}), \xi_1}$  be the pushforward of

$$\frac{\mu_{\sigma^{n(\xi_1)}(\mathbf{b})}|_{\pi_{\sigma^{n(\xi_1)}(\mathbf{b})}([c])}}{\mu_{\sigma^{n(\xi_1)}(\mathbf{b})}(\pi_{\sigma^{n(\xi_1)}(\mathbf{b})}([c]))}$$

under  $x \rightarrow \xi_1^{2/3} x$ . The following statement bounds  $|\widehat{\mu_{\mathbf{b}}}(\xi_1)|$  by an average of  $\widehat{\mu_{\mathbf{c}, \sigma^{n(\xi_1)}(\mathbf{b}), \xi_1}}$  taken over all  $\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}$  and a range of frequencies plus a polynomially small error.

**Proposition 6.3.** *For any  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$  we have*

$$|\widehat{\mu}_{\mathbf{b}}(\xi_1)| \leq \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n(\xi_1)}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} m_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}]) |\mu_{\mathbf{c}, \sigma^{n(\xi_1)}(\mathbf{b}), \xi_1}(\xi_1^{1/3} (f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}}))| + \mathcal{O}(|\xi_1|^{-1/3})$$

*Proof.* Fix  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$ . Using Lemma 6.1 we have

$$|\widehat{\mu}_{\mathbf{b}}(\xi_1)| = \left| \int e^{2\pi i \xi_1 x} d\mu_{\mathbf{b}} \right| = \left| \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n(\xi_1)}} m_{\mathbf{b}}([\mathbf{a}]) \int e^{2\pi i \xi_1 f_{\mathbf{a}}^{(1)}(x)} d\mu_{\sigma^{n(\xi_1)}(\mathbf{b})} \right|$$

We now split the latter integrals over cylinder sets determined by element of  $\mathcal{A}_{\xi_1, \mathbf{b}}$ :

$$\begin{aligned} & \left| \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n(\xi_1)}} m_{\mathbf{b}}([\mathbf{a}]) \int e^{2\pi i \xi_1 f_{\mathbf{a}}^{(1)}(x)} d\mu_{\sigma^{n(\xi_1)}(\mathbf{b})} \right| \\ &= \left| \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n(\xi_1)}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} \int_{\pi_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}])} e^{2\pi i \xi_1 f_{\mathbf{a}}^{(1)}(x)} d\mu_{\sigma^{n(\xi_1)}(\mathbf{b})} \right| \end{aligned}$$

For each  $\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n(\xi_1)}$  we can linearize  $f_{\mathbf{a}}^{(1)}(x)$  around  $x_{\mathbf{c}}$  and use Lemma 2.1 to derive the bound

$$f_{\mathbf{a}}^{(1)}(x) = f_{\mathbf{a}}^{(1)}(x_{\mathbf{c}}) + (f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})(x - x_{\mathbf{c}}) + \mathcal{O}(|\xi_1|^{-4/3}) \quad (6.7)$$

for all  $x \in \pi_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}])$ . Here we have used the definition of  $\mathcal{A}_{\xi_1, \mathbf{b}}$ . Applying (6.7) in the above we have

$$\begin{aligned} & \left| \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n(\xi_1)}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} \int_{\pi_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}])} e^{2\pi i \xi_1 f_{\mathbf{a}}^{(1)}(x)} d\mu_{\sigma^{n(\xi_1)}(\mathbf{b})} \right| \\ &\leq \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n(\xi_1)}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} \left| \int_{\pi_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}])} e^{2\pi i \xi_1 (f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})x} d\mu_{\sigma^{n(\xi_1)}(\mathbf{b})} \right| + \mathcal{O}(|\xi_1|^{-1/3}) \\ &= \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n(\xi_1)}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} m_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}]) \left| \int e^{2\pi i \xi_1^{1/3} (f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})x} d\mu_{\mathbf{c}, \sigma^{n(\xi_1)}(\mathbf{b}), \xi_1} \right| + \mathcal{O}(|\xi_1|^{-1/3}) \\ &= \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \dots \times \mathcal{A}_{\bar{b}_n(\xi_1)}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} m_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}]) \left| \mu_{\mathbf{c}, \sigma^{n(\xi_1)}(\mathbf{b}), \xi_1}(\xi_1^{1/3} (f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})) \right| + \mathcal{O}(|\xi_1|^{-1/3}). \end{aligned}$$

This completes our proof.  $\square$

**6.4. Affine non-concentration of the averaged measures.** In light of the above proposition, to successfully apply Theorem 1.3 to prove Theorem 1.21, we need to show that each  $\mu_{\mathbf{c}, \sigma^{n(\xi_1)}(\mathbf{b}), \xi_1}$  is uniformly affinely non-concentrated and the non-concentration parameters do not depend upon our choice of measure. This fact is established in the following proposition.

**Proposition 6.4.** *There exists  $C, \alpha > 0$  such that for any  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$ ,  $\xi_1 \in \mathbb{R}$ ,  $\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}$ ,  $x \in \text{supp}(\mu_{\mathbf{c}, \sigma^{n(\xi_1)}(\mathbf{b}), \xi_1})$ ,  $y \in \mathbb{R}$ ,  $r \in (0, 1]$  and  $\epsilon > 0$  we have*

$$\mu_{\mathbf{c}, \sigma^{n(\xi_1)}(\mathbf{b}), \xi_1}(B(x, r) \cap B(y, \epsilon r)) \leq C\epsilon^{\alpha} \mu_{\mathbf{c}, \sigma^{n(\xi_1)}(\mathbf{b}), \xi_1}(B(x, r)).$$

*Proof.* Let us fix  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$ ,  $\xi_1 \in \mathbb{R}$  and  $\mathbf{c} \in A_{\xi_1, \mathbf{b}}$ . Let  $x \in \text{supp}(\mu_{\mathbf{c}, \sigma^n(\xi_1)(\mathbf{b}), \xi_1})$ ,  $r \in (0, 1]$  and  $\epsilon > 0$ . Then, by definition, for any  $y \in \mathbb{R}$  we have

$$\begin{aligned}
& \mu_{\mathbf{c}, \sigma^n(\xi_1)(\mathbf{b}), \xi_1}(B(x, r) \cap B(y, \epsilon r)) \\
&= \frac{\mu_{\sigma^n(\xi_1)(\mathbf{b})|_{\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}] )}(B(x \cdot \xi_1^{-2/3}, r \cdot \xi_1^{-2/3}) \cap B(y \cdot \xi_1^{-2/3}, \epsilon r \cdot \xi_1^{-2/3}))}{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}] ))} \\
&\leq \frac{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(x \cdot \xi_1^{-2/3}, r \cdot \xi_1^{-2/3}) \cap B(y \cdot \xi_1^{-2/3}, \epsilon r \cdot \xi_1^{-2/3}))}{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}] ))} \\
&= \frac{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(x', r \cdot \xi_1^{-2/3}) \cap B(y', \epsilon r \cdot \xi_1^{-2/3}))}{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}] ))}. \tag{6.8}
\end{aligned}$$

Where in the last line we have adopted the notation:  $x' = x \cdot \xi_1^{-2/3} \in \pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}])$  and  $y' = y \cdot \xi_1^{-2/3}$ . We begin our proof by analysing the sets appearing in the numerator in (6.8). Without loss of generality we can assume that  $y' \in X_{\sigma^n(\xi_1)(\mathbf{b})}$ .

It is useful at this point to recall that the IFS  $\{f_a^{(1)}\}_{a \in \mathcal{A}}$  satisfies the strong separation condition. Therefore, if  $X_1 \subset \mathbb{R}$  denotes the attractor for this IFS, we have  $f_a^{(1)}(X_1) \cap f_{a'}^{(1)}(X_1) = \emptyset$  for distinct  $a, a' \in \mathcal{A}_1$ . Therefore

$$\min_{a \neq a'} d(f_a^{(1)}(X_1), f_{a'}^{(1)}(X_1)) > 0.$$

It is a consequence of Lemma 2.2 and this property that for  $\mathbf{a}, \mathbf{a}' \in \mathcal{A}_1^*$  we have

$$d(f_{\mathbf{a}}^{(1)}(X_1), f_{\mathbf{a}'}^{(1)}(X_1)) \asymp \text{Diam}(f_{|\mathbf{a} \wedge \mathbf{a}'|}([0, 1])). \tag{6.9}$$

Since  $\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{d}]) \subset f_{\mathbf{d}}^{(1)}(X_1)$  for all  $\mathbf{d} \in \cup_{j=1}^{\infty} \prod_{i=1}^j \mathcal{A}_{\bar{b}_{n(\xi_1)+i}}$ , (6.9) implies

$$d(\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{d}]), \pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{d}'])) \asymp \text{Diam}(f_{|\mathbf{d} \wedge \mathbf{d}'|}([0, 1])) \tag{6.10}$$

for distinct  $\mathbf{d}, \mathbf{d}' \in \cup_{j=1}^{\infty} \prod_{i=1}^j \mathcal{A}_{\bar{b}_{n(\xi_1)+i}}$ .

We let  $\mathbf{d} = (d_i) \in \prod_{i=1}^{\infty} \mathcal{A}_{\bar{b}_{n(\xi_1)+i}}$  be the unique sequence satisfying  $\pi_{\sigma^n(\xi_1)(\mathbf{b})}(\mathbf{d}) = x'$ . Since  $x' \in \pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}])$  the sequence  $\mathbf{d}$  must begin with  $\mathbf{c}$ . Let  $m \in \mathbb{N}$  be minimal such that

$$\pi_{\sigma^n(\xi_1)(\mathbf{b})}([d_1, \dots, d_m]) \subset B(x', r \cdot \xi_1^{-2/3}).$$

It follows from (6.10) that a bounded number of  $(d'_1, \dots, d'_m) \in \prod_{i=1}^m \mathcal{A}_{\bar{b}_{n(\xi_1)+i}}$  satisfy

$$\pi_{\sigma^n(\xi_1)(\mathbf{b})}([d'_1, \dots, d'_m]) \cap B(x', r \cdot \xi_1^{-2/3}) \neq \emptyset. \tag{6.11}$$

It also follows from (6.10) that there exists  $l \in \mathbb{N}$  depending only on our IFS, such that if  $\mathbf{d}'$  satisfies (6.11) then  $d_i = d'_i$  for all  $1 \leq i \leq m - l$ . Therefore for such a  $(d'_1, \dots, d'_m)$  we have

$$m_{\sigma^n(\xi_1)(\mathbf{b})}([d_1, \dots, d_m]) \asymp m_{\sigma^n(\xi_1)(\mathbf{b})}([d'_1, \dots, d'_m]).$$

Combining these observations we have

$$\mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(x', r \cdot \xi_1^{-2/3})) \asymp m_{\sigma^n(\xi_1)(\mathbf{b})}([d_1, \dots, d_m]). \tag{6.12}$$

Let us now consider  $B(y', \epsilon r \cdot \xi_1^{-2/3})$ . Let  $\mathbf{e} = (e_i) \in \Sigma_{\sigma^n(\xi_1)(\mathbf{b})}$  be the unique sequence satisfying  $\pi_{\sigma^n(\xi_1)(\mathbf{b})}(\mathbf{e}) = y'$ . Let  $m' \geq m$  be minimal such that

$$\pi_{\sigma^n(\xi_1)(\mathbf{b})}([e_1, \dots, e_{m'}]) \subset B(y', \epsilon r \cdot \xi_1^{-2/3}).$$

By an analogous argument to that given above, we must have

$$\mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(y', \epsilon r \xi_1^{-2/3})) \asymp m_{\sigma^n(\xi_1)(\mathbf{b})}([e_1, \dots, e_{m'}]). \quad (6.13)$$

Using the fact that our IFS is uniformly contracting, it can be shown that  $m' - m \geq \lfloor d \log \epsilon^{-1} \rfloor$  for some  $d > 0$  depending only on our IFS. Using this inequality and recalling the definition of  $\gamma$  given by (6.2), we have

$$\begin{aligned} \mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(y', \epsilon r \xi_1^{-2/3})) &\ll m_{\sigma^n(\xi_1)(\mathbf{b})}([e_1, \dots, e_{m'}]) \\ &\ll m_{\sigma^n(\xi_1)(\mathbf{b})}([e_1, \dots, e_m]) \epsilon^{-d \log \gamma} \\ &\ll m_{\sigma^n(\xi_1)(\mathbf{b})}([d_1, \dots, d_m]) \epsilon^{-d \log \gamma}. \end{aligned} \quad (6.14)$$

In the last line we used that  $(e_1, \dots, e_m)$  must satisfy (6.11) therefore  $e_i = d_i$  for  $1 \leq i \leq m - l$  and so  $m_{\sigma^n(\xi_1)(\mathbf{b})}([e_1, \dots, e_m]) \asymp m_{\sigma^n(\xi_1)(\mathbf{b})}([d_1, \dots, d_m])$ . Summarising, it follows from (6.8), (6.12), (6.13) and (6.14) that

$$\begin{aligned} \mu_{\mathbf{c}, \sigma^n(\xi_1)(\mathbf{b}), \xi}(B(x, r) \cap B(y, \epsilon r)) &\leq \frac{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(x', r \cdot \xi_1^{-2/3}) \cap B(y', \epsilon r \cdot \xi_1^{-2/3}))}{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}]))} \\ &\ll \frac{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(y', \epsilon r \cdot \xi_1^{-2/3}))}{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}]))} \\ &\ll \epsilon^{-d \log \gamma} \frac{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(x', r \cdot \xi_1^{-2/3}))}{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}]))}. \end{aligned}$$

Taking  $\alpha = -d \log \gamma$ , we see that to complete our proof it remains to show that

$$\frac{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(x', r \cdot \xi_1^{-2/3}))}{\mu_{\sigma^n(\xi_1)(\mathbf{b})}(\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}]))} \ll \mu_{\mathbf{c}, \sigma^n(\xi_1)(\mathbf{b}), \xi}(B(x, r)).$$

This is equivalent to

$$\mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(x', r \cdot \xi_1^{-2/3})) \ll \mu_{\sigma^n(\xi_1)(\mathbf{b})}|_{\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}])}(B(x', r \cdot \xi_1^{-2/3})). \quad (6.15)$$

It is useful to recall at this point that by the definition of  $A_{\xi_1, \mathbf{b}}$  we know that

$$\text{Diam}(\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}])) \asymp \xi_1^{-2/3}. \quad (6.16)$$

Therefore, it follows from (6.10) that there exists  $r_0 > 0$  depending only on our IFS such that if  $r < r_0$  then

$$\mu_{\sigma^n(\xi_1)(\mathbf{b})}|_{\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}])}(B(x', r \cdot \xi_1^{-2/3})) = \mu_{\sigma^n(\xi_1)(\mathbf{b})}(B(x', r \cdot \xi_1^{-2/3})).$$

Therefore (6.15) holds for  $r \leq r_0$ . For  $r \geq r_0$ , one can appeal to (6.16) and the fact  $\mathbf{d}$  begins with  $\mathbf{c}$  to show that

$$m_{\sigma^n(\xi_1)(\mathbf{b})}([d_1, \dots, d_m]) \asymp m_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}]).$$

It is also straightforward to show that

$$\mu_{\sigma^n(\xi_1)(\mathbf{b})}|_{\pi_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}])}(B(x', r \cdot \xi_1^{-2/3})) \asymp m_{\sigma^n(\xi_1)(\mathbf{b})}([\mathbf{c}]).$$

Thus (6.15) follows for  $r \geq r_0$  by these observations and (6.12). This completes our proof.  $\square$

**6.5. Spectral gap and non-concentration of frequencies.** The final step in the proof of Theorem 1.21 is to show that the set of frequencies

$$\left\{ \xi_1^{1/3} (f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}}) : \mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_n(\xi_1)} \right\}$$

appearing in Proposition 6.3 satisfy a non-concentration property. This is established in Proposition 6.8 below. The key ingredient is a spectral gap for certain random transfer operators (Proposition 6.5).

More precisely, to each  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$  and  $b \in \mathbb{R}$ , we associate a transfer operator that acts on  $C^1(U)$ :

$$\mathcal{L}_{ib}^{(\mathbf{b})} h(x) = \sum_{a \in \mathcal{A}_{\bar{b}_1}} m_{\mathbf{b}}([a]) |(f_a^{(1)})'(x)|^{ib} h(f_a(x)).$$

Iterating this formula we have

$$(\mathcal{L}_{ib}^{(\sigma^{n-1}(\mathbf{b}))} \circ \cdots \circ \mathcal{L}_{ib}^{(\mathbf{b})}) h(x) = \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_n}} m_{\mathbf{b}}([\mathbf{a}]) |(f_{\mathbf{a}}^{(1)})'(x)|^{ib} h(f_{\mathbf{a}}^{(1)}(x)) \quad (6.17)$$

for any  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$ ,  $h \in C^1(U)$ ,  $n \in \mathbb{N}$  and  $x \in U$ .

The following statement asserts that a random operator  $\mathcal{L}_{ib}^{(\sigma^{n-1}(\mathbf{b}))} \circ \cdots \circ \mathcal{L}_{ib}^{(\mathbf{b})}$  has a spectral gap for all  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$  outside of a small measure set. It is in the proof of this proposition where we use the third assumption in Theorem 1.21.

**Proposition 6.5.** *There exists  $\rho_1 \in (0, 1)$  and  $\Theta > 0$  such that the following statements are true:*

- *For all  $n \in \mathbb{N}$  there exists  $\Omega_n \in \mathcal{B}^{\mathbb{N}}$  such that*

$$\|\mathcal{L}_{ib}^{(\sigma^{n-1}(\mathbf{b}))} \circ \cdots \circ \mathcal{L}_{ib}^{(\mathbf{b})}\|_{\infty} \ll \rho_1^n |b|^{1/2} \|h\|_b$$

*for all  $\bar{\mathbf{b}} \in \Omega_n$ ,  $h \in C^1(U)$  and  $|b| > \Theta$ .*

- *There exists  $\delta > 0$  such that  $\Omega_n$  satisfies  $Q(\Omega_n^c) \ll e^{-\delta n}$  for all  $n \in \mathbb{N}$ .*

*Sketch of proof of Proposition 6.5.* We do not include a full proof of Proposition 6.5 because it is essentially contained in [12]. See Theorem 3.2 and (3.9) from this paper. We give a rough outline of the argument for convenience of the reader. The third assumption in Theorem 1.21 implies that the IFS  $\{f_a^{(1)} : a \in \mathcal{A}_{\bar{b}^*}\}$  must satisfy the UNI property for some  $\bar{b}^* \in B$ . A large deviation argument tells us that outside of an exponentially small measure set, for a typical  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$ , a significant proportion of the operators appearing in the composition  $\mathcal{L}_{ib}^{(\sigma^{n-1}(\mathbf{b}))} \circ \cdots \circ \mathcal{L}_{ib}^{(\mathbf{b})}$  will equal  $\mathcal{L}_{ib}^{(\bar{b}^*)}$ . The existence of the desired spectral gap for the single operator  $\mathcal{L}_{ib}^{(\bar{b}^*)}$  goes back to Naud [62] and Stoyanov [84]. By adapting Naud's argument, the first author and third author showed in [12] that if  $\mathcal{L}_{ib}^{(\bar{b}^*)}$  appeared sufficiently many times within  $\mathcal{L}_{ib}^{(\sigma^{n-1}(\mathbf{b}))} \circ \cdots \circ \mathcal{L}_{ib}^{(\mathbf{b})}$  then we have the desired spectral gap. Combining this with our large deviation bound implies Proposition 6.5.  $\square$

**Remark 6.6.** We remark that in [12] the separation assumption for our IFS is that  $f_a^{(1)}([0, 1]) \cap f_{a'}^{(1)}([0, 1]) = \emptyset$  for distinct  $a, a' \in \mathcal{A}_1$ . This is a stronger assumption than the strong separation condition. Nevertheless it is possible to adapt the arguments in [12] to weaken this assumption to the strong separation condition.

**Remark 6.7.** We note that Proposition 6.5 gives a spectral gap in terms of the  $\|\cdot\|_{\infty}$  norm instead of the usual  $\|\cdot\|_b$  norm. Despite the spectral gap being phrased in terms of this norm, it causes no issues in our proof of Theorem 1.21.

Equipped with Proposition 6.5, we can now prove our non-concentration statement for frequencies.

**Proposition 6.8.** *There exists  $\kappa > 0$  such that for any  $\mathbf{b} \in \Omega_{n(\xi_1)}$ ,  $\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}$  and  $n \in \mathbb{N}$  satisfying  $|\xi_1|^{1/6} \leq n \leq |\xi_1|^{1/3}$  we have*

$$\sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) \chi_{[n, n+1]}(\xi_1^{1/3} |(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})|) = \mathcal{O}(|\xi_1|^{-\kappa}).$$

*Proof.* The proof of Proposition 6.8 is analogous to Proposition 5.3. Thus, we only explain how the random transfer operators appear in our analysis. This is the only point where the proof differs.

As in the proof of Proposition 5.3, we can replace  $\chi_{[n, n+1]}$  with a smooth function  $h_{n, \xi_1}$  that satisfies

$$\chi_{[n, n+1]}(\xi_1^{1/3} |(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})|) \leq h_{n, \xi_1}(\log |(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})|)$$

and for which we have useful bounds on the  $L^1$  norm of its second derivative. Applying the above and the Fourier inversion formula, we have the following:

$$\begin{aligned} & \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) \chi_{[n, n+1]}(\xi_1^{1/3} |(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})|) \\ & \leq \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) h_{n, \xi}(\log |(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})|) \\ & \leq \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) \int \widehat{h_{n, \xi}}(\eta) e^{-2\pi i \eta \log |(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})|} d\eta \\ & = \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) \int \widehat{h_{n, \xi}}(\eta) |(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})|^{-2\pi i \eta} d\eta \\ & = \int \widehat{h_{n, \xi}}(\eta) \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) |(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})|^{-2\pi i \eta} d\eta \\ & = \int \widehat{h_{n, \xi}}(\eta) (\mathcal{L}_{-2\pi i \eta}^{(\sigma^{n(\xi_1)} - 1)(\mathbf{b})} \circ \cdots \circ \mathcal{L}_{-2\pi i \eta}^{(\mathbf{b})})(1)(x_{\mathbf{c}}) d\eta, \end{aligned}$$

where, in the last line, we used (6.17). The rest of our proof is identical to the proof of Proposition 5.3. We split out integral into two parts, one part that can be controlled using the bound  $|\widehat{h_{n, \xi}}(\eta)| \leq |\widehat{h_{n, \xi}}(0)|$ , and one part on which we use our assumption  $\mathbf{b} \in \Omega_{n(\xi_1)}$  and the spectral gap guaranteed by Proposition 6.5.  $\square$

**6.6. Conclusion of the proof of Theorem 1.21.** Recall that, by our assumption, we have  $\max_{1 \leq i \leq d} |\xi_i| = |\xi_1|$  and, hence, it suffices to show that (6.1) holds. By (6.4) and Proposition 6.3, we have

$$\begin{aligned} |\widehat{\mu}(\xi)| & \leq \mathcal{O}(|\xi_1|^{-1/3}) \\ & + \int \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} m_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}]) |\mu_{\mathbf{c}, \widehat{\sigma^{n(\xi_1)}(\mathbf{b})}, \xi_1}(\xi_1^{1/3} (f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}}))| dQ(\mathbf{b}). \end{aligned}$$

So to complete our proof it is sufficient to show that

$$\int \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} m_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}]) |\mu_{\mathbf{c}, \widehat{\sigma^{n(\xi_1)}(\mathbf{b})}, \xi_1}(\xi_1^{1/3} (f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}}))| dQ(\mathbf{b}) \ll |\xi_1|^{-\eta} \quad (6.18)$$

for some  $\eta > 0$ . We begin by splitting our outer integral. Let  $\Omega_{n(\xi)}$  be as in Proposition 6.5. It follows from this proposition that

$$\begin{aligned}
& \int \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} m_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}]) |\mu_{\mathbf{c}, \widehat{\sigma^{n(\xi_1)}(\mathbf{b})}, \xi_1}(\xi_1^{1/3}(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}}))| dQ(\mathbf{b}) \\
& \ll |\xi_1|^{-\eta_1} \\
& + \int_{\Omega_{n(\xi_1)}} \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} m_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}]) |\mu_{\mathbf{c}, \widehat{\sigma^{n(\xi_1)}(\mathbf{b})}, \xi_1}(\xi_1^{1/3}(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}}))| dQ(\mathbf{b}) \\
& = |\xi_1|^{-\eta_1} \\
& + \int_{\Omega_{n(\xi_1)}} \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} m_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}]) \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) |\mu_{\mathbf{c}, \widehat{\sigma^{n(\xi_1)}(\mathbf{b})}, \xi_1}(\xi_1^{1/3}(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}}))| dQ(\mathbf{b}),
\end{aligned} \tag{6.19}$$

for some  $\eta_1 > 0$ . For each  $\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}$ ,  $\mathbf{b} \in \mathcal{B}^{\mathbb{N}}$  and  $\tau > 0$ , we let

$$\text{Bad}_{\mathbf{c}, \mathbf{b}, \tau} := \left\{ n \in \mathbb{Z} : |n| \leq |\xi_1|^{1/3} \text{ and } \exists \zeta \in [n, n+1] \text{ satisfying } |\mu_{\mathbf{c}, \widehat{\sigma^{n(\xi_1)}(\mathbf{b})}, \xi_1}(\zeta)| \geq |\xi_1|^{-\tau} \right\}.$$

By Proposition 6.4, we know that each  $\mu_{\mathbf{c}, \sigma^{n(\xi_1)}(\mathbf{b}), \xi_1}$  is uniformly affinely non-concentrated and that the non-concentration parameters can be chosen independently of the measure. As such, we can apply Theorem 1.3 to assert that there exists  $\tau > 0$  such that

$$\#\text{Bad}_{\mathbf{c}, \mathbf{b}, \tau} = \mathcal{O}_{\kappa}(|\xi_1|^{\kappa/2}), \tag{6.20}$$

where  $\kappa$  is as in Proposition 6.8. Crucially the underlying constants in (6.20) do not depend upon  $\mathbf{c}$ ,  $\mathbf{b}$  or  $\xi$ . Using Proposition 6.8 and (6.6), we see that for any  $\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}$  and  $\mathbf{b} \in \Omega_{n(\xi_1)}$  we have

$$\begin{aligned}
& \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) |\mu_{\mathbf{c}, \widehat{\sigma^{n(\xi_1)}(\mathbf{b})}, \xi_1}(\xi_1^{1/3}(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}}))| \\
& \leq \sum_{n \in \text{Bad}_{\mathbf{c}, \mathbf{b}, \tau}} \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) \chi_{[n, n+1]}(\xi_1^{1/3}(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})) \\
& + \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) \chi_{(\cup_{n \in \text{Bad}_{\mathbf{c}, \mathbf{b}, \tau}} [n, n+1])^c}(\xi_1^{1/3}(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}})) |\mu_{\mathbf{c}, \widehat{\sigma^{n(\xi_1)}(\mathbf{b})}, \xi_1}(\xi_1^{1/3}(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}}))| \\
& \leq \sum_{n \in \text{Bad}_{\mathbf{c}, \mathbf{b}, \tau}} |\xi_1|^{-\kappa} + \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) |\xi_1|^{-\tau} \\
& \ll |\xi_1|^{-\min\{\kappa/2, \tau\}}.
\end{aligned} \tag{6.21}$$

In the last line, we used (6.20). Substituting (6.21) into (6.19), we have

$$\begin{aligned}
& \int_{\Omega_{n(\xi_1)}} \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} m_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}]) \sum_{\mathbf{a} \in \mathcal{A}_{\bar{b}_1} \times \cdots \times \mathcal{A}_{\bar{b}_{n(\xi_1)}}} m_{\mathbf{b}}([\mathbf{a}]) |\mu_{\mathbf{c}, \widehat{\sigma^{n(\xi_1)}(\mathbf{b})}, \xi_1}(\xi_1^{1/3}(f_{\mathbf{a}}^{(1)})'(x_{\mathbf{c}}))| dQ + |\xi_1|^{-\eta_1} \\
& \ll \int_{\Omega_{n(\xi_1)}} |\xi_1|^{-\min\{\kappa/2, \tau\}} \sum_{\mathbf{c} \in \mathcal{A}_{\xi_1, \mathbf{b}}} m_{\sigma^{n(\xi_1)}(\mathbf{b})}([\mathbf{c}]) dQ + |\xi_1|^{-\eta_1} \\
& \leq |\xi_1|^{-\min\{\kappa/2, \tau, \eta_1\}}.
\end{aligned}$$

Thus, (6.18) holds and our proof is complete.

## APPENDIX A. HIGHER DIMENSIONAL DAVENPORT-ERDŐS-LEVEQUE CRITERION

The purpose of this appendix is to give the proof of the following:

**Theorem A.1.** *Let  $\mu$  be a probability measure on  $\mathbb{R}^d$  with polylogarithmic Fourier decay. Then,  $\mu$  almost every  $x$  is  $A$ -normal for any expanding integer valued matrix  $A$ .*

The strategy of proof for Theorem A.1 is similar to that implemented by Queffélec and Ramaré [70] for the one dimensional version of the DEL criterion. The third author and Jonathan Fraser wrote a proof of Theorem A.1 in an unpublished note. We thank Jonathan for his permission to include the argument here.

To prove Theorem A.1, we need the following Lemma.

**Lemma A.2** (Lemma 7.3 [70]). *Suppose  $(r_N) \subset (0, \infty)$  is a sequence of reals such that  $\sum_{N=1}^{\infty} \frac{r_N}{N} < \infty$ . Then there is a subsequence  $N_j \rightarrow \infty$ ,  $j \in \mathbb{N}$ , such that  $\sum_{j=1}^{\infty} r_{N_j} < \infty$  and  $\lim_{j \rightarrow \infty} \frac{N_{j+1}}{N_j} = 1$ .*

*Proof of Theorem A.1.* Let  $\mu$  be a measure on the  $d$ -torus  $\mathbb{T}^d$  for which there exists some  $\alpha > 0$  and  $C > 0$  such that the Fourier transform satisfies  $|\widehat{\mu}(\xi)| \leq C|\log|\xi||^{-\alpha}$  for all  $\xi \in \mathbb{R}^d$  with  $|\xi| > 1$ . Also, let  $T(x) = Ax \bmod 1$ ,  $x \in \mathbb{T}^d$  for some  $A \in \mathbb{Z}^{d \times d}$ . Fix  $\mathbf{k} \in \mathbb{Z}^d \setminus \{0\}$  and for  $N \in \mathbb{N}$  write

$$S_N(x) := \frac{1}{N} \sum_{n=1}^N \exp(2\pi i \mathbf{k} \cdot T^n(x)),$$

where  $\mathbf{k}$  has been suppressed from the notation. Weyl's equidistribution criterion in  $\mathbb{T}^d$  (see e.g. [87, Proposition 1.1.2]) says that  $(x_n) \subset \mathbb{T}^d$  is equidistributed in  $\mathbb{T}^d$  if and only if for any  $\mathbf{k} \in \mathbb{Z}^d \setminus \{0\}$  we have  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \exp(2\pi i \mathbf{k} \cdot x_n) = 0$ . Thus it is enough to prove that  $S_N(x) \rightarrow 0$  as  $N \rightarrow \infty$  for  $\mu$  almost every  $x$ . Note that from this one can deduce that the result holds simultaneously for *all* expanding toral endomorphisms since such maps are represented by two dimensional matrices with integer coefficients and so there are only countably many possibilities.

Let  $r_N := \int |S_N(x)|^2 d\mu(x)$ . For each  $x \in \mathbb{T}^d$  as  $A$  has integer entries, we know that  $\exp(2\pi i \mathbf{k} \cdot T^n(x)) = \exp(2\pi i (A^*)^n(\mathbf{k}) \cdot x)$ . Then we may estimate

$$\begin{aligned} r_N &= \frac{1}{N^2} \sum_{m,n=1}^N \int \exp(-2\pi i \mathbf{k} \cdot (T^m(x) - T^n(x))) d\mu(x) \\ &= \frac{1}{N^2} \sum_{m,n=1}^N \int \exp(-2\pi i ((A^*)^m(\mathbf{k}) - (A^*)^n(\mathbf{k})) \cdot x) d\mu(x) \\ &\leq \frac{1}{N} + \frac{2}{N^2} \sum_{m=2}^N \sum_{n=1}^{m-1} |\widehat{\mu}((A^*)^m(\mathbf{k}) - (A^*)^n(\mathbf{k}))|. \end{aligned}$$

Let  $\sigma_d = \inf\{|Ax| : |x| = 1\}$  be the smallest singular value of  $A$ , which is the same as the smallest singular value of  $A^*$ . Thus we have for any  $n \in \mathbb{N}$  and  $y \in \mathbb{R}^d$  that  $|(A^*)^n y| \geq \sigma_d^n |y|$ . Since  $\sigma_d > 1$ , we have  $\sigma_d^n \rightarrow \infty$  as  $n \rightarrow \infty$ . Thus there exists  $n_0 \in \mathbb{N}$  such that whenever  $m, n \in \mathbb{N}$  with  $m > n \geq n_0$ , we have

$$\sigma_d^n \geq 1 + \frac{1}{\sigma_d - 1} \geq 1 + \frac{1}{\sigma_d^{m-n} - 1} = \frac{\sigma_d^{m-n}}{\sigma_d^{m-n} - 1}.$$

Thus for  $m > n \geq n_0$  we can bound  $(\sigma_d^{m-n} - 1)\sigma_d^n \geq \sigma_d^{m-n}$ , giving

$$|(A^*)^m(\mathbf{k}) - (A^*)^n(\mathbf{k})| \geq ||(A^*)^m(\mathbf{k})| - |(A^*)^n(\mathbf{k})|| \geq (\sigma_d^{m-n} - 1)\sigma_d^n |\mathbf{k}| \geq \sigma_d^{m-n}$$

again, valid for  $m > n \geq n_0$ . By the Fourier decay and  $|(A^*)^n y| \geq \sigma_d^n |y|$  we obtain

$$\begin{aligned}
\sum_{m=2}^N \sum_{n=1}^{m-1} |\widehat{\mu}((A^*)^m(\mathbf{k}) - (A^*)^n(\mathbf{k}))| &\leq N(n_0 - 1) + \sum_{m=2}^N \sum_{n=n_0}^{m-1} |\widehat{\mu}((A^*)^m(\mathbf{k}) - (A^*)^n(\mathbf{k}))| \\
&\lesssim N + \sum_{m=2}^N \sum_{n=n_0}^{m-1} (\log |(A^*)^m(\mathbf{k}) - (A^*)^n(\mathbf{k})|)^{-\alpha} \\
&\leq N + \sum_{m=2}^N \sum_{n=n_0}^{m-1} (\log \sigma_d^{m-n})^{-\alpha} \\
&= N + \log(\sigma_d)^{-\alpha} \sum_{m=2}^N \sum_{n=n_0}^{m-1} \frac{1}{(m-n)^\alpha} \\
&= \mathcal{O}(N + N^{2-\alpha}),
\end{aligned}$$

Since  $\alpha > 0$  we have that  $\sum_{N=1}^\infty \frac{r_N}{N} < \infty$ . Thus by Lemma A.2 there is a subsequence  $N_j \rightarrow \infty$ ,  $j \in \mathbb{N}$ , such that  $\sum_{j=1}^\infty r_{N_j} < \infty$  and  $\lim_{j \rightarrow \infty} \frac{N_{j+1}}{N_j} = 1$ . In particular, the former condition yields that

$$\int \sum_{j=1}^\infty |S_{N_j}(x)|^2 d\mu(x) < \infty$$

and so  $S_{N_j}(x) \rightarrow 0$  for  $\mu$  almost every  $x$  as  $j \rightarrow \infty$ . Now we just check that the latter condition actually yields  $S_N(x) \rightarrow 0$  for  $\mu$  almost every  $x$  as  $N \rightarrow \infty$ , which is what we need. Fix  $N \in \mathbb{N}$  and find  $j \in \mathbb{N}$  such that  $N_j \leq N \leq N_{j+1}$ . This yields that

$$|NS_N(x) - N_j S_{N_j}(x)| = \left| \sum_{n=N_j+1}^N \exp(-2\pi i \mathbf{k} \cdot T^n(x)) \right| \leq N - N_j \leq N_{j+1} - N_j.$$

Hence

$$|S_N(x)| \leq |S_{N_j}(x)| + \frac{N_{j+1} - N_j}{N_j},$$

which converges to 0 as  $j \rightarrow \infty$  at  $\mu$  almost every  $x$ .  $\square$

## APPENDIX B. PROOF OF LEMMA 3.4

In this section we will prove lemma 3.4. Let us recall its statement.

**Lemma B.1.** *Let  $(p_a)_{a \in \mathcal{A}}$  be a probability vector and  $n \in \mathbb{N}$ , then*

$$\frac{n!}{\prod_{a \in \mathcal{A}} k_a!} \prod_{a \in \mathcal{A}} p_a^{k_a} \ll n^{-1/2(\#\mathcal{A}-1)}$$

for any  $(k_a)_{a \in \mathcal{A}} \in \mathbb{N}^{\#\mathcal{A}}$  satisfying  $\sum_{a \in \mathcal{A}} k_a = n$ . Here the underlying constants only depend upon the probability vector.

*Proof.* Let  $(k_a^*)_{a \in \mathcal{A}}$  be such that

$$\max_{(k_a)_{a \in \mathcal{A}}: \sum_{a \in \mathcal{A}} k_a = n} \frac{n!}{\prod_{a \in \mathcal{A}} k_a!} \prod_{a \in \mathcal{A}} p_a^{k_a} = \frac{n!}{\prod_{a \in \mathcal{A}} k_a^*!} \prod_{a \in \mathcal{A}} p_a^{k_a^*}.$$

We claim that  $(k_a^*)_{a \in \mathcal{A}}$  must satisfy

$$k_a^* \in \{\lfloor p_a n \rfloor - 10\#\mathcal{A}, \dots, \lfloor p_a n \rfloor + 10\#\mathcal{A}\} \quad (\text{B.1})$$

for all  $a \in \mathcal{A}$ . Suppose for instance that  $k_{a'}^* < \lfloor p_{a'} n \rfloor - 10\#\mathcal{A}$  for some  $a' \in \mathcal{A}$ , then there must exists  $a'' \in \mathcal{A}$  such that  $k_{a''}^* \geq \lfloor p_{a''} n \rfloor + 2$ . We have

$$\begin{aligned} & \frac{n!}{(k_{a'}^* + 1)!(k_{a''}^* - 1)! \prod_{a \in \mathcal{A} \setminus \{a', a''\}} k_a^*!} p_{a'}^{k_{a'}^* + 1} p_{a''}^{k_{a''}^* - 1} \prod_{a \in \mathcal{A} \setminus \{a', a''\}} p_a^{k_a^*} \\ &= \frac{k_{a''}^* p_{a'}}{(k_{a'}^* + 1)p_{a''}} \frac{n!}{\prod_{a \in \mathcal{A}} k_a^*!} \prod_{a \in \mathcal{A}} p_a^{k_a^*} \end{aligned} \quad (\text{B.2})$$

However, we have the following lower bound for  $\frac{k_{a''}^* p_{a'}}{(k_{a'}^* + 1)p_{a''}}$ :

$$\frac{k_{a''}^* p_{a'}}{(k_{a'}^* + 1)p_{a''}} \geq \frac{(\lfloor p_{a''} n \rfloor + 2)p_{a'}}{(\lfloor p_{a'} n \rfloor - 10\#\mathcal{A})p_{a''}} \geq \frac{p_{a''} p_{a'} n + p_{a'}}{p_{a'} p_{a''} n - 10\#\mathcal{A} p_{a''}} > 1. \quad (\text{B.3})$$

Substituting (B.3) into (B.2) we have

$$\frac{n!}{(k_{a'}^* + 1)!(k_{a''}^* - 1)! \prod_{a \in \mathcal{A} \setminus \{a', a''\}} k_a^*!} p_{a'}^{k_{a'}^* + 1} p_{a''}^{k_{a''}^* - 1} \prod_{a \in \mathcal{A} \setminus \{a', a''\}} p_a^{k_a^*} > \frac{n!}{\prod_{a \in \mathcal{A}} k_a^*!} \prod_{a \in \mathcal{A}} p_a^{k_a^*}.$$

However this contradicts our assumption that  $(k_a^*)$  was maximising. The case where  $k_{a'}^* > \lfloor p_{a'} n \rfloor - 10\#\mathcal{A}$  follows similarly. Thus we have proved our claim.

We now claim that if  $(k_a)_{a \in \mathcal{A}}$  satisfies

$$k_a \in \{\lfloor p_a n \rfloor - 10\#\mathcal{A}, \dots, \lfloor p_a n \rfloor + 10\#\mathcal{A}\} \quad (\text{B.4})$$

for all  $a \in \mathcal{A}$  then

$$\frac{n!}{\prod_{a \in \mathcal{A}} k_a^*!} \prod_{a \in \mathcal{A}} p_a^{k_a^*} \ll \frac{n!}{\prod_{a \in \mathcal{A}} k_a!} \prod_{a \in \mathcal{A}} p_a^{k_a}. \quad (\text{B.5})$$

We remark that if  $(k_a)$  satisfies (B.4) then

$$\frac{n!}{\prod_{a \in \mathcal{A}} k_a^*!} \prod_{a \in \mathcal{A}} p_a^{k_a^*}$$

can be obtained from

$$\frac{n!}{\prod_{a \in \mathcal{A}} k_a!} \prod_{a \in \mathcal{A}} p_a^{k_a}$$

by multiplying the latter term a bounded number of times by terms each of which is bounded above by

$$\frac{\max_{a \in \mathcal{A}} \{\lfloor p_a n \rfloor + 10\#\mathcal{A}\} \max_{a \in \mathcal{A}} \{p_a\}}{\min_{a \in \mathcal{A}} \{\lfloor p_a n \rfloor - 10\#\mathcal{A}\} \min_{a \in \mathcal{A}} \{p_a\}}. \quad (\text{B.6})$$

Since (B.6) is uniformly bounded from above our claim follows.

It follows from (B.5) that to prove our lemma we just need to show that the desired bound holds for a specific  $(k_a)$  satisfying (B.4). This we do now. Pick  $a' \in \mathcal{A}$  arbitrarily, then by Stirling's formula we have

$$\begin{aligned} & \frac{n!}{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor)! \prod_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor!} p_{a'}^{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor)} \prod_{a \in \mathcal{A} \setminus \{a'\}} p_a^{\lfloor p_a n \rfloor} \\ & \ll \frac{\sqrt{n} n^n p_{a'}^{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor)}}{\sqrt{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor) (n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor)^{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor)}}} \prod_{a \in \mathcal{A} \setminus \{a'\}} \frac{p_a^{\lfloor p_a n \rfloor}}{\sqrt{\lfloor p_a n \rfloor \lfloor p_a n \rfloor^{\lfloor p_a n \rfloor}}}. \end{aligned}$$

The following bounds are straightforward to verify

$$\frac{(p_a n)^{\lfloor p_a n \rfloor}}{\lfloor p_a n \rfloor^{\lfloor p_a n \rfloor}} \ll 1 \quad \text{and} \quad \frac{(p_{a'} n)^{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor)}}{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor)^{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor)}} \ll 1. \quad (\text{B.7})$$

Substituting (B.7) into the above yields:

$$\begin{aligned} & \frac{n!}{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor)! \prod_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor!} p_{a'}^{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor)} \prod_{a \in \mathcal{A} \setminus \{a'\}} p_a^{\lfloor p_a n \rfloor} \\ & \ll \frac{\sqrt{n}}{\sqrt{(n - \sum_{a \in \mathcal{A} \setminus \{a'\}} \lfloor p_a n \rfloor) \prod_{a \in \mathcal{A} \setminus \{a'\}} \sqrt{p_a n}}} \\ & \ll n^{-1/2(\#\mathcal{A}-1)}. \end{aligned}$$

Thus our result follows.  $\square$

### APPENDIX C. PROOF OF PROPOSITION 1.16

The purpose of this section is to prove Proposition 1.16, which we recall here:

**Proposition C.1** (Spectral gap). *Let  $\Phi$  be a conformal IFS satisfying the strong separation condition and  $\Sigma_A$  be a subshift of finite type. Suppose that  $(\Phi, \Sigma_A)$  satisfies the UNI condition, then for any  $C^1$  potential  $\psi$  satisfying  $P(\psi) = 0$ , there exists  $0 < \rho < 1$  such that for all large enough  $|b|$ ,  $n \in \mathbb{N}$  and  $h \in C^{1,b}(U)$ , we have*

$$\|\mathcal{L}_{ib}^n h\|_b \ll \rho^n |b|^{1/2} \|h\|_b.$$

*In particular, if  $(\Phi, \Sigma_A)$  satisfies the UNI condition, then for any  $C^1$  potential  $\psi$  satisfying  $P(\psi) = 0$  the triple  $(\Phi, \Sigma_A, \psi)$  has a spectral gap.*

In our proof of this proposition we will make use of a uniform non-integrability condition introduced in [54]. It follows from Proposition D.1 that this condition is equivalent to our UNI condition. Thus in this section, by Proposition D.1 we may freely assume the existence of  $r > 0$  and  $\epsilon_0 > 0$  such that the following holds for any large  $n \in \mathbb{N}$ . For any  $x \in X_A$  and unit vector  $e \in \mathbb{R}^d$ , there exist  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  such that:

- For all  $y \in B(x, r)$  we have

$$|\partial_e (\log |\lambda_{\mathbf{a}_1}(y)| - \log |\lambda_{\mathbf{a}_2}(y)|)| \geq \epsilon_0.$$

- For all  $y \in B(x, r)$  we have  $\mathbf{a}_1 \rightsquigarrow y$  and  $\mathbf{a}_2 \rightsquigarrow y$ .

Without loss of generality we may assume that our (untwisted) transfer operator satisfies

$$\sum_{\substack{a \in \mathcal{A} \\ a \rightsquigarrow x}} w_a(x) = 1 \quad (\text{C.1})$$

for all  $x \in X_A$ . This is permissible as one can conjugate our transfer operator by a multiplication operator so that the resulting transfer operator satisfies (C.1). Moreover, it can be shown that if this new operator satisfies the conclusion of Proposition 1.16, then our original transfer operator also satisfies this conclusion. Thus there is no loss of generality in making this assumption. For further details of this reduction we refer the reader to [62, Section 5.1].

Recall that  $U = \bigcup_{a \in \mathcal{A}} U_a$  is a choice of open set satisfying  $X_\Phi \subset U$ , where the  $U_a$  are disjoint open sets with  $f_a(U) \subset U_a$  for all  $a \in \mathcal{A}$ .

In our proof of Proposition 1.16 we will make regular use of the following set of functions: For  $A > 0$  let

$$C_A := \{u \in C^1(U) : u > 0, \|D_x u\| \leq A u(x)\}.$$

We record here a useful property of the cone  $C_A$ . If  $u \in C_A$  then

$$e^{-A|x-y|} \leq \frac{u(x)}{u(y)} \leq e^{A|x-y|} \quad (\text{C.2})$$

for all  $x, y \in U$ .

We will also make regular use of the following lemma. It records various well known bounded distortion estimates whose proofs we omit.

**Lemma C.2.** *Let  $\Phi = \{f_a\}_{a \in \mathcal{A}}$  be a conformal IFS satisfying the strong separation condition and  $\psi : U \rightarrow \mathbb{R}$  be a  $C^1$  potential. Then there exists  $C_1, C_2, C_3 > 0$  such that the following statements are true:*

- For any  $x, y \in X_\Phi$  and  $\mathbf{a} \in \mathcal{A}^*$  we have

$$e^{-C_1|x-y|} \leq \frac{w_{\mathbf{a}}(x)}{w_{\mathbf{a}}(y)} \leq e^{C_1|x-y|}.$$

- For any  $x \in X_\Phi$  and  $\mathbf{a} \in \mathcal{A}^*$  we have

$$\|D_x w_{\mathbf{a}}\| \leq C_2 w_{\mathbf{a}}(x).$$

- For any  $x \in X_\Phi$  and  $\mathbf{a} \in \mathcal{A}^*$  we have

$$\|D_x \log |\lambda_{\mathbf{a}}|\| \leq C_3.$$

**C.1. Construction of Dolgopyat operators.** In what follows we will fix an IFS  $\Phi$  and a subshift of finite type  $\Sigma_A$  such that the assumptions of Proposition 1.16 hold. We will also fix a  $C^1$  potential  $\psi$  satisfying  $P(\psi) = 0$ . In this section we will denote the corresponding Gibbs measure by  $\mu$ .

In this section we will define a family of *Dolgopyat operators*  $(\mathcal{N}_b^J)_{J \in \mathcal{E}_b}$  that will allow us to prove Proposition 1.16 via properties of complex transfer operators. First of all, let us fix  $A > 1$  sufficiently large such that

$$(C_2 + 2|b| + A|b|\gamma^n) \leq A|b| \quad (\text{C.3})$$

and

$$(2C_2 + 2C_3|b| + 2A|b|\gamma^n) \leq A|b| \quad (\text{C.4})$$

for all  $|b| > 1$  and  $n$  sufficiently large. This is slightly abusing notation with the matrix  $A$  defining the subshift of finite type. Here  $\gamma \in (0, 1)$  is the uniform contraction factor in our IFS and  $C_2, C_3 > 0$  are as in Lemma C.2.

We now state the key proposition where the UNI condition manifests.

**Proposition C.3.** *There exist  $\eta_0 = \eta_0(\Phi, \Sigma_A) \in (0, 1)$ ,  $D = D(\Phi, \Sigma_A) > 0$ ,  $\epsilon_1 = \epsilon_1(\Phi, \Sigma_A)$  and  $\epsilon_2 = \epsilon_2(\Phi, \Sigma_A) > 0$  such that for  $N \in \mathbb{N}$  sufficiently large,  $|b| > 1$ ,  $H \in C_{A|b|}$ , and  $u \in C^1(U)$ , if  $u$  satisfies  $|u| \leq H$  and  $\|Du\| \leq A|b|H$ , then for any  $y \in X_A$  there exists  $x' \in X_A \cap B\left(y, \frac{\epsilon_1 D}{|b|}\right)$  such that*

$$B\left(x', \frac{\epsilon_2}{|b|}\right) \subset B\left(y, \frac{\epsilon_1 D}{|b|}\right)$$

and for which we have the following: there exists  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^N$  such that  $\mathbf{a}_1 \rightsquigarrow x$  and  $\mathbf{a}_2 \rightsquigarrow x$  for all  $x \in B\left(y, \frac{\epsilon_1 D}{|b|}\right)$ , and one of the following holds for all  $x \in B\left(x', \frac{\epsilon_2}{|b|}\right)$ :

$$\begin{aligned} & \left| w_{\mathbf{a}_1}(x) |\lambda_{\mathbf{a}_1}(x)|^{ib} u(f_{\mathbf{a}_1}(x)) + w_{\mathbf{a}_2}(x) |\lambda_{\mathbf{a}_2}(x)|^{ib} u(f_{\mathbf{a}_2}(x)) \right| \\ & \leq \eta_0 w_{\mathbf{a}_1}(x) H(f_{\mathbf{a}_1}(x)) + w_{\mathbf{a}_2}(x) H(f_{\mathbf{a}_2}(x)), \end{aligned} \quad (\text{C.5})$$

or

$$\begin{aligned} & \left| w_{\mathbf{a}_1}(x) |\lambda_{\mathbf{a}_1}(x)|^{ib} u(f_{\mathbf{a}_1}(x)) + w_{\mathbf{a}_2}(x) |\lambda_{\mathbf{a}_2}(x)|^{ib} u(f_{\mathbf{a}_2}(x)) \right| \\ & \leq w_{\mathbf{a}_1}(x) H(f_{\mathbf{a}_1}(x)) + \eta_0 w_{\mathbf{a}_2}(x) H(f_{\mathbf{a}_2}(x)) \end{aligned} \quad (\text{C.6})$$

The proof of Proposition C.3 will be given in Section C.3 below. Having found the parameters  $\eta_0, D, \varepsilon_1, \varepsilon_2$  in Proposition C.3, we can now define the family of Dolgopyat operators.

**Definition C.4** (Dolgopyat operators  $(\mathcal{N}_b^J)_{J \in \mathcal{E}_b}$ ). Assume  $|b| > 1$  is fixed and  $N \in \mathbb{N}$  is a large number so that Proposition 1.16 applies. We now define our Dolgopyat operators. We fix a maximal set of points  $\{y_l\}_{l \in S(b)} \subset X_A$  with some index set  $S(b)$  such that

$$B\left(y_l, \frac{\varepsilon_1 D}{|b|}\right) \cap B\left(y_{l'}, \frac{\varepsilon_1 D}{|b|}\right) = \emptyset$$

for distinct  $l, l' \in S(b)$ . By Proposition C.3, applied to  $y_l$  for each  $l \in S(b)$ , we can always find a point  $x'_l \in B\left(y_l, \frac{\varepsilon_1 D}{|b|}\right)$  satisfying

$$B\left(x'_l, \frac{\varepsilon_2}{|b|}\right) \subset B\left(y_l, \frac{\varepsilon_1 D}{|b|}\right)$$

and a word  $\mathbf{a}_l \in \mathcal{A}^N$  satisfying  $\mathbf{a}_l \rightsquigarrow x$  for all  $x \in B\left(y_l, \frac{\varepsilon_1 D}{|b|}\right)$  such that either (C.5) holds with  $\mathbf{a}_1 = \mathbf{a}_l$  or (C.6) holds with  $\mathbf{a}_2 = \mathbf{a}_l$ . Let  $\eta_0 \in (0, 1)$  be the value from Proposition C.3 and fix  $\eta \in [\eta_0, 1)$ , then we can choose a  $C^1$  function  $\chi : U \rightarrow [\eta, 1]$  such that  $\chi(x) = 1$  outside of  $\cup_{l \in S(b)} f_{\mathbf{a}_l}(B(x'_l, \frac{\varepsilon_2}{|b|}))$  and  $\chi(x) = \eta$  within  $\cup_{l \in S(b)} f_{\mathbf{a}_l}(B(x'_l, \frac{\varepsilon_2}{|b|}))$ . Moreover, we can choose  $\eta$  in such a way that  $\|D(\chi \circ f_{\mathbf{a}})\| \leq |b|$  for all  $\mathbf{a} \in \mathcal{A}^N$ . We can also freely assume  $\eta \geq 1/2$ . We emphasize that the function  $\chi$  only depends upon the pairs  $\{(x'_l, \mathbf{a}_l)\}_{l \in S(b)}$ . We let

$$\mathcal{E}_b := \left\{ (x'_l, \mathbf{a}_l)_{l \in S(b)} : \begin{array}{l} B\left(x'_l, \frac{\varepsilon_2}{|b|}\right) \subset B\left(y_l, \frac{\varepsilon_1 D}{|b|}\right) \text{ and} \\ \mathbf{a}_l \in \mathcal{A}^N \text{ satisfies } \mathbf{a}_l \rightsquigarrow x \forall x \in B\left(y_l, \frac{\varepsilon_1 D}{|b|}\right) \end{array} \right\}.$$

Given  $J \in \mathcal{E}_b$  we define the *Dolgopyat operator*  $\mathcal{N}_b^J$  on  $C^1(U)$  according to the rule

$$\mathcal{N}_b^J(H) := \mathcal{L}_0^N(\chi H).$$

Crucially, Dolgopyat operators satisfy the following key properties, which we will prove using Proposition C.3, the doubling property of the Gibbs measure  $\mu$ , and basic inequalities for complex transfer operators.

**Lemma C.5** (Key properties of Dolgopyat operators). *There exists  $N \in \mathbb{N}$ ,  $A > 1$  and  $\rho \in (0, 1)$  such that for  $|b|$  sufficiently large the Dolgopyat operators  $(\mathcal{N}_b^J)_{J \in \mathcal{E}_b}$  on  $C^1(U)$  satisfy the following properties:*

(1) *The cone  $C_{A|b|}$  is stable under  $\mathcal{N}_b^J$  for all  $J \in \mathcal{E}_b$ , that is, if  $H \in C_{A|b|}$ , then*

$$\|D_x \mathcal{N}_b^J(H)\| \leq A|b| \mathcal{N}_b^J(H)(x), \quad x \in U.$$

(2) *For all  $H \in C_{A|b|}$  and  $J \in \mathcal{E}_b$ ,*

$$\int |\mathcal{N}_b^J H|^2 d\mu \leq \rho \int |H|^2 d\mu.$$

(3) *Given  $H \in C_{A|b|}$  and  $u \in C^1(U)$  such that  $|u| \leq H$  and  $\|Du\| \leq A|b|H$ , there exists  $J \in \mathcal{E}_b$  such that*

$$|\mathcal{L}_{ib}^N u| \leq \mathcal{N}_b^J(H)$$

and

$$\|D(\mathcal{L}_{ib}^N u)\| \leq A|b| \mathcal{N}_b^J(H).$$

The proof of this lemma will be given in Section C.4 below.

**C.2. Reduction of spectral gap to Dolgopyat operators.** Let us now show how Proposition 1.16 follows using Lemma C.5. Our proof is based upon arguments of Naud [62]. For the purposes of brevity we will omit certain calculations from our proof.

*Proof of Proposition 1.16.* We start by proving that Lemma C.5 implies an  $L^2$  decay statement. Let  $u \in C^1(U)$  satisfy  $\|u\|_b \leq 1$ . Set  $H = \|u\|_b 1$ . Then  $H \in C_{A|b|}$  and  $|u| \leq H$ ,  $\|Du\| \leq |b|\|u\|_b \leq A|b|H$ . Thus by the definition of  $\mathcal{L}_{ib}$  and iterating Lemma C.5(3), we get for all  $n \geq 1$ ,  $|\mathcal{L}_{ib}^{nN} u| \leq \mathcal{N}_b^{J_n} \mathcal{N}_b^{J_{n-1}} \dots \mathcal{N}_b^{J_1} H$  and  $\|D\mathcal{L}_{ib}^{nN} u\| \leq A|b|\mathcal{N}_b^{J_n} \mathcal{N}_b^{J_{n-1}} \dots \mathcal{N}_b^{J_1} H$  for some  $J_1, \dots, J_n \in \mathcal{E}_b$ . Hence

$$\int_X |\mathcal{L}_{ib}^{nN} u|^2 d\mu \leq \int_X |\mathcal{N}_b^{J_n} \mathcal{N}_b^{J_{n-1}} \dots \mathcal{N}_b^{J_1} H|^2 d\mu \leq \rho^n \int_X |H|^2 d\mu \leq \rho^n. \quad (\text{C.7})$$

Now we will show how (C.7) leads to a spectral gap. We will use the following well known inequalities: Let  $u \in C^1(U)$ . There exists  $C_1, C_2 > 0$  and  $\rho_1 \in (0, 1)$  such that

- (i)  $\sup_{x \in U} \|D_x \mathcal{L}_{ib}^n u\| \leq C_1 |b| \|\mathcal{L}_0^n u\|_\infty + \rho_1^n \|\mathcal{L}_0^n (\|Du\|)\|_\infty$
- (ii)  $\|\mathcal{L}_{ib}^n u\|_\infty \leq \int |u| d\mu + C_2 \rho_1^n \sup_{x \in U} \|D_x u\|$

We remark that the key inequality for complex transfer operators (i) was proved in [65] in the symbolic setting and (ii) follows from the quasi-compactness of  $\mathcal{L}_{ib}$ .

Now we prove our spectral gap result. Fix  $h \in C^{1,b}(U)$  with  $\|h\|_b \leq 1$ . Using the above and the Cauchy-Schwartz inequality, we have

$$\begin{aligned} |\mathcal{L}_{ib}^{2nN} h(x)| &\leq |\mathcal{L}_{ib}^{nN} (\mathcal{L}_{ib}^{nN} h)(x)| \leq \mathcal{L}_0^{nN} (|\mathcal{L}_{ib}^{nN} h|)(x) \\ &\ll (\mathcal{L}_0^{nN} (|\mathcal{L}_{ib}^{nN} h|^2)(x))^{1/2} \\ &\leq \left( \int_X |\mathcal{L}_{ib}^{nN} h|^2 d\mu + C_2 \rho_1^{nN} \sup_{x \in U} \|D_x |\mathcal{L}_{ib}^{nN} h|^2\| \right)^{1/2} \\ &\ll \left( \int_X |\mathcal{L}_{ib}^{nN} h|^2 d\mu + C_2 \rho_1^{nN} |b| \right)^{1/2}. \end{aligned}$$

In the penultimate line we used the second of the inequalities listed above. In the final line we used the product rule, the first of the inequalities listed above, and the fact  $\|h\|_b \leq 1$  to assert that  $\sup_{x \in U} \|D_x |\mathcal{L}_{ib}^{nN} h|^2\| \ll |b|$ . Now using (C.7) we have

$$|\mathcal{L}_{ib}^{2nN} h(x)| \ll (\rho^n + C_2 \rho_1^{nN} |b|)^{1/2} \ll \max\{\rho^{1/4N}, \rho_1^{1/4}\}^{2nN} |b|^{1/2}.$$

This implies

$$\|\mathcal{L}_{ib}^{2nN}\|_\infty \ll \max\{\rho^{1/4N}, \rho_1^{1/4}\}^{2nN} |b|^{1/2} \|h\|_b. \quad (\text{C.8})$$

This establishes the sup norm part of our spectral gap bound. It remains to prove the derivative part. Using the first of the inequalities listed above we have:

$$\begin{aligned} \frac{1}{|b|} \sup_{x \in U} \|D_x \mathcal{L}_{ib}^{2nN} h\| &\leq C_1 \|\mathcal{L}_0^{nN} (\mathcal{L}_{ib}^{nN} h)\|_\infty + \frac{\rho_1^n}{|b|} \|\mathcal{L}_0^{nN} (\|D\mathcal{L}_{ib}^{nN} h\|)\|_\infty \\ &\ll C_1 \|\mathcal{L}_0^{nN} (\mathcal{L}_{ib}^{nN} h)\|_\infty + \rho_1^n \|\mathcal{L}_{ib}^{nN} h\|_b. \end{aligned}$$

The first term in the above can be bounded using similar ideas to those used above to establish (C.8). The second term can be bounded by evaluating  $D_x \mathcal{L}_{ib}^{nN} h$  and bounding accordingly.  $\square$

**C.3. Proof of Proposition C.3.** Let us now prove Proposition C.3. To this end we will need the following two lemmas. Firstly, we need the following higher dimensional analogue of Lemma 5.11 from [62]:

**Lemma C.6.** *Let  $Z \subset U$  be a set satisfying  $\text{Diam}(Z) \leq \frac{c}{|b|}$ . Let  $H \in C_{A|b|}$  and  $u \in C^1(U)$  satisfy  $|u| \leq H$  and  $\|Du\| \leq A|b|H$ . Then for  $c$  sufficiently small, we have either  $|u(x)| \leq \frac{3}{4}H(x)$  for all  $x \in Z$ , or  $|u(x)| \geq \frac{1}{4}H(x)$  for all  $x \in Z$ .*

*Proof.* The proof is same as in [62] and we include it for completeness. If  $x_0 \in Z$  has  $|u(x_0)| \leq \frac{1}{4}H(x_0)$ , then at any  $x \in Z$ , we have for all small enough  $c > 0$ :

$$|u(x)| \leq |u(x) - u(x_0)| + \frac{1}{4}H(x_0) \leq A|b|\text{diam}(Z) \sup_{x \in Z} H + \frac{1}{4}H(x_0) \leq (cA + \frac{1}{4})e^{Ac}H(x) \leq \frac{3}{4}H(x).$$

In the penultimate inequality we used (C.2).  $\square$

We also recall the following triangle lemma that follows from applying trigonometric identities:

**Lemma C.7** (Triangle lemma). *Let  $z_1, z_2 \geq 0$  be two complex numbers such that  $|z_1/z_2| \leq L$  and  $2\pi - \epsilon \geq |\arg(z_1) - \arg(z_2)| \geq \epsilon > 0$ . Then there exists  $0 < \delta(L, \epsilon) < 1$  such that*

$$|z_1 + z_2| \leq (1 - \delta)|z_1| + |z_2|.$$

*Proof of Proposition C.3.* Let  $N \in \mathbb{N}$  be large and  $|b| > 1$ . Let us also fix  $H \in C_{A|b|}$  and  $u \in C^1(U)$  satisfying the assumptions of Proposition C.3. Let  $y \in X_A$ . Let  $\epsilon_1 > \epsilon_2 > 0$  be parameters that only depend upon  $\Phi$  and  $\Sigma_A$ . Using properties of  $\Phi$  and the fact  $\Sigma_A$  is topologically mixing, by the strong separation of the IFS  $\Phi$ , there exists  $D > D' > 2$  depending only on  $\Phi$  and  $\Sigma_A$  such that there exists  $x' \in X_A$  satisfying

$$x' \in B\left(y, \frac{\epsilon_1 D}{|b|}\right) \setminus B\left(y, \frac{\epsilon_1 D'}{|b|}\right)$$

and

$$B\left(x', \frac{\epsilon_2}{|b|}\right) \subset B\left(y, \frac{\epsilon_1 D}{|b|}\right).$$

Let  $B_1 = B\left(y, \frac{\epsilon_2}{|b|}\right)$ ,  $B_2 = B\left(x', \frac{\epsilon_2}{|b|}\right)$ . Let  $\hat{B}$  be the smallest ball containing  $B_1 \cup B_2$ . We can assume that  $\epsilon_1$  and  $\epsilon_2$  are sufficiently small that  $\hat{B} \subset B(y, r)$ . Here  $r$  is as in the uniform non-integrability condition stated at the beginning of this section (recall that by Proposition D.1 this is equivalent to our UNI condition). We remark that for any  $x_1 \in B_1$  and  $x_2 \in B_2$  we have

$$\frac{\epsilon_1}{b} \leq \|x_1 - x_2\| \leq (D + 2)\frac{\epsilon_1}{b}. \quad (\text{C.9})$$

Let  $e_0 = \frac{(y-x')}{\|y-x'\|}$ . Applying the uniform non-integrability condition stated at the start of this section for this  $e_0$  and  $y$ , we see that there exists  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^N$  such that

$$|\partial_{e_0}(\log |\lambda_{\mathbf{a}_1}(x)| - \log |\lambda_{\mathbf{a}_2}(x)|)| \geq \epsilon_0 \quad (\text{C.10})$$

and

$$\mathbf{a}_1 \rightsquigarrow x \quad \text{and} \quad \mathbf{a}_2 \rightsquigarrow x$$

for all  $x \in B(y, r)$ .

Since our IFS is contracting, and we can choose  $\epsilon_1$  to be small, we can assume that for  $m \in \{1, 2\}$  the set  $f_{\mathbf{a}_m}(\hat{B})$  is sufficiently small so that we can apply Lemma C.6 and assert that for  $m \in \{1, 2\}$  either  $|u(f_{\mathbf{a}_m}(x))| \leq \frac{3}{4}H(f_{\mathbf{a}_m}(x))$  for all  $x \in \hat{B}$  or  $|u(f_{\mathbf{a}_m}(x))| \geq \frac{1}{4}H(f_{\mathbf{a}_m}(x))$  for all  $x \in \hat{B}$ . If  $|u(f_{\mathbf{a}_m}(x))| \leq \frac{3}{4}H(f_{\mathbf{a}_m}(x))$  for all  $x \in \hat{B}$  for some  $m \in \{1, 2\}$  then our result follows immediately. Thus we will assume in what follows that

$$|u(f_{\mathbf{a}_m}(x))| \geq \frac{1}{4}H(f_{\mathbf{a}_m}(x)) \quad (\text{C.11})$$

for all  $x \in \hat{B}$  for  $m \in \{1, 2\}$ .

For  $x \in \hat{B}$ , we set

$$z_1(x) = w_{\mathbf{a}_1}(x)|\lambda_{\mathbf{a}_1}(x)|^{ib}u(f_{\mathbf{a}_1}(x)) \quad \text{and} \quad z_{\mathbf{a}_2}(x) = w_{\mathbf{a}_2}(x)|\lambda_{\mathbf{a}_2}(x)|^{ib}u(f_{\mathbf{a}_2}(x)).$$

We claim that there exists  $M > 0$  such that either  $|z_1(x)|/|z_2(x)| \leq M$  for all  $x \in \hat{B}$  or  $|z_2(x)/z_1(x)| \leq M$  for all  $x \in \hat{B}$ . Using (C.11) and our assumptions on  $u$  we see that for all  $x \in \hat{B}$  we have

$$\frac{w_{\mathbf{a}_1}(x)H(f_{\mathbf{a}_1}(x))}{4w_{\mathbf{a}_2}(x)H(f_{\mathbf{a}_2}(x))} \leq \left| \frac{z_1(x)}{z_2(x)} \right| \leq \frac{4w_{\mathbf{a}_1}(x)H(f_{\mathbf{a}_1}(x))}{w_{\mathbf{a}_2}(x)H(f_{\mathbf{a}_2}(x))}.$$

If there exists  $x_0 \in \hat{B}$  such that

$$\frac{w_{\mathbf{a}_1}(x_0)H(f_{\mathbf{a}_1}(x_0))}{w_{\mathbf{a}_2}(x_0)H(f_{\mathbf{a}_2}(x_0))} \leq 1,$$

then using (C.2), and Lemma C.2, for all  $x \in \hat{B}$  we have

$$\frac{w_{\mathbf{a}_1}(x)H(f_{\mathbf{a}_1}(x))}{w_{\mathbf{a}_2}(x)H(f_{\mathbf{a}_2}(x))} \leq \frac{w_{\mathbf{a}_1}(x_0)e^{(2D+2)\epsilon_1(A+C_1)}H(f_{\mathbf{a}_1}(x_0))}{w_{\mathbf{a}_2}(x_0)e^{-(2D+2)\epsilon_1(A+C_1)}H(f_{\mathbf{a}_2}(x_0))} \leq e^{2(2D+2)\epsilon_1(A+C_1)}.$$

Here  $C_1 > 0$  is as in Lemma C.2. Thus in this case we can take  $M := 4e^{2(2D+2)\epsilon_1(A+C_1)}$ . Similarly it can be shown that a suitable  $M'$  exists such that if

$$\frac{w_{\mathbf{a}_1}(x)H(f_{\mathbf{a}_1}(x))}{w_{\mathbf{a}_2}(x)H(f_{\mathbf{a}_2}(x))} \geq 1$$

for all  $x \in \hat{B}$  then  $|z_2(x)/z_1(x)| \leq M'$  for all  $x \in \hat{B}$ . This completes the proof of our claim.

For each  $m \in \{1, 2\}$  define the function  $g_m : [0, 1] \rightarrow \mathbb{C}$  by  $g_m(t) = z_m(y + t(x' - y))$ . Then define  $L_m : [0, 1] \rightarrow \mathbb{C}$  by

$$L_m(t) = \int_0^t \frac{g'_m(s)}{g_m(s)} ds + l_{0,m}$$

where  $l_{0,m}$  satisfies  $e^{l_{0,m}} = z_m(y)$ .  $L_m$  is well defined as  $g_m(t) \neq 0$  for all  $t \in [0, 1]$  and  $e^{L_m(t)} = z_m(t)$  for all  $t \in [0, 1]$ . Set  $\Phi(t) = \text{Im}(L_1(t)) - \text{Im}(L_2(t))$ . Since the potential  $\psi$  defining the Gibbs measure is differentiable, we have:

$$\begin{aligned} \Phi'(t) &= \text{Im} \left( \frac{g'_1(t)}{g_1(t)} - \frac{g'_2(t)}{g_2(t)} \right) \\ &= \text{Im} \left( \frac{\partial_{e_0} w_{\mathbf{a}_1}(y + t(x' - y)) \|x' - y\| |\lambda_{\mathbf{a}_1}(y + t(x' - y))|^{ib} u(f_{\mathbf{a}_1}(y + t(x' - y)))}{w_{\mathbf{a}_1}(y + t(x' - y)) |\lambda_{\mathbf{a}_1}(y + t(x' - y))|^{ib} u(f_{\mathbf{a}_1}(y + t(x' - y)))} \right) \\ &\quad + \text{Im} \left( \frac{w_{\mathbf{a}_1}(y + t(x' - y)) ib \partial_{e_0} (\log |\lambda_{\mathbf{a}_1}(y + t(x' - y))|) |x' - y| u(f_{\mathbf{a}_1}(y + t(x' - y)))}{w_{\mathbf{a}_1}(y + t(x' - y)) u(f_{\mathbf{a}_1}(y + t(x' - y)))} \right) \\ &\quad + \text{Im} \left( \frac{w_{\mathbf{a}_1}(y + t(x' - y)) |\lambda_{\mathbf{a}_1}(y + t(x' - y))|^{ib} D_{f_{\mathbf{a}_1}(y+t(x'-y))} u D_{y+t(x'-y)} f_{\mathbf{a}_1}(x' - y)}{w_1(y + t(x' - y)) |\lambda_{\mathbf{a}_1}(y + t(x' - y))|^{ib} u(f_{\mathbf{a}_1}(y + t(x' - y)))} \right) \\ &\quad - \text{Im} \left( \frac{\partial_{e_0} w_{\mathbf{a}_2}(y + t(x' - y)) \|x' - y\| |\lambda_{\mathbf{a}_2}(y + t(x' - y))|^{ib} u(f_{\mathbf{a}_2}(y + t(x' - y)))}{w_{\mathbf{a}_2}(y + t(x' - y)) |\lambda_{\mathbf{a}_2}(y + t(x' - y))|^{ib} u(f_{\mathbf{a}_2}(y + t(x' - y)))} \right) \\ &\quad - \text{Im} \left( \frac{w_{\mathbf{a}_2}(y + t(x' - y)) ib \partial_{e_0} (\log |\lambda_{\mathbf{a}_2}(y + t(x' - y))|) |x' - y| u(f_{\mathbf{a}_2}(y + t(x' - y)))}{w_{\mathbf{a}_2}(y + t(x' - y)) u(f_{\mathbf{a}_2}(y + t(x' - y)))} \right) \\ &\quad - \text{Im} \left( \frac{w_{\mathbf{a}_2}(y + t(x' - y)) |\lambda_{\mathbf{a}_2}(y + t(x' - y))|^{ib} D_{f_{\mathbf{a}_2}(y+t(x'-y))} u D_{y+t(x'-y)} f_{\mathbf{a}_2}(x' - y)}{w_{\mathbf{a}_2}(y + t(x' - y)) |\lambda_{\mathbf{a}_2}(y + t(x' - y))|^{ib} u(f_{\mathbf{a}_2}(y + t(x' - y)))} \right). \end{aligned}$$

Note that the first and fourth term are equal to 0 as we are taking imaginary parts of real numbers. Thus after some cancellation we are left with the identity:

$$\begin{aligned}\Phi'(t) &= b\|x' - y\| \left( \partial_{e_0}(\log |\lambda_{\mathbf{a}_1}(y + t(x' - y))| - \log |\lambda_{\mathbf{a}_2}(y + t(x' - y))|) \right) \\ &\quad + Im \left( \frac{D_{f_{\mathbf{a}_1}(y+t(x'-y))} u D_{y+t(x'-y)} f_{\mathbf{a}_1}(x' - y)}{u(f_{\mathbf{a}_1}(y + t(x' - y)))} \right) \\ &\quad - Im \left( \frac{D_{f_{\mathbf{a}_2}(y+t(x'-y))} u D_{y+t(x'-y)} f_{\mathbf{a}_2}(x' - y)}{u(f_{\mathbf{a}_2}(y + t(x' - y)))} \right).\end{aligned}$$

By our assumptions on  $u$  and (C.11), we have

$$\left| \frac{D_{f_{\mathbf{a}_1}(y+t(x'-y))} u D_{y+t(x'-y)} f_{\mathbf{a}_1}(x' - y)}{u(f_{\mathbf{a}_1}(y + t(x' - y)))} \right| \leq \frac{4A|b|H(f_{\mathbf{a}_1}(y + t(x' - y)))\gamma^{n_0}|x' - y|}{H(f_{\mathbf{a}_1}(y + t(x' - y)))} = 4A|b|\gamma^N|x' - y|.$$

Similarly,

$$\left| \frac{D_{f_{\mathbf{a}_2}(y+t(x'-y))} u D_{y+t(x'-y)} f_{\mathbf{a}_2}(x' - y)}{u(f_{\mathbf{a}_2}(y + t(x' - y)))} \right| \leq 4A|b|\gamma^N|x' - y|.$$

Appealing to bounded distortions it can be shown that there exists  $m > 0$  depending only on our IFS such that

$$|\partial_{e_0}(\log |\lambda_{\mathbf{a}_1}(y + t(x' - y))| - \log |\lambda_{\mathbf{a}_2}(y + t(x' - y))|)| \leq m, \quad (\text{C.12})$$

for any  $t \in [0, 1]$  (see also (D.1)). Combining (C.10) and (C.12) with the above we have

$$(\epsilon_0 - 8A\gamma^N) |b|\|x' - y\| \leq |\Phi'(t)| \leq (m + 8A\gamma^N) |b|\|x' - y\|.$$

Thus for  $N$  sufficiently large we have

$$\frac{\epsilon_0|b|}{2}\|x' - y\| \leq |\Phi'(t)| \leq 2m|b|\|x' - y\|.$$

Therefore

$$\frac{\epsilon_0|b|}{2}\|x' - y\| \leq |\Phi(1) - \Phi(0)| \leq 2m|b|\|x' - y\|.$$

By (C.9) we have

$$\frac{\epsilon_1}{|b|} \leq \|x' - y\| \leq (D + 2) \frac{\epsilon_1}{|b|}.$$

This implies

$$\frac{\epsilon_0\epsilon_1}{2} \leq |\Phi(1) - \Phi(0)| \leq 2m(D + 2)\epsilon_1.$$

Now choose  $\epsilon_1$  sufficiently small that  $(2m(D + 2) + \epsilon_0/2)\epsilon_1 \leq \pi$  and let  $\epsilon' = \frac{\epsilon_0\epsilon_1}{8}$ . Suppose now that

$$\Phi(1), \Phi(0) \in \cup_{k \in \mathbb{Z}} [2\pi k - \epsilon', 2\pi k + \epsilon'].$$

Since  $|\Phi(1) - \Phi(0)| \leq 2m(D + 2)\epsilon_1$  we cannot have  $\Phi(1) \in [2\pi k_1 - \epsilon', 2\pi k_1 + \epsilon']$  and  $\Phi(0) \in [2\pi k_2 - \epsilon', 2\pi k_2 + \epsilon']$  for  $k_1 \neq k_2$ . As we would then have

$$2m(D + 2)\epsilon_1 \geq |\Phi(1) - \Phi(0)| \geq 2\pi - 2\epsilon' = 2\pi - \frac{\epsilon_0\epsilon_1}{4}$$

which is not possible given our choice of  $\epsilon_1$ . Therefore we must have

$$\frac{\epsilon_0\epsilon_1}{2} \leq |\Phi(1) - \Phi(0)| \leq 2\epsilon' = \frac{\epsilon_0\epsilon_1}{4}.$$

Which is also not possible. Recall now the definition of  $\Phi(t) = Im(L_1(t)) - Im(L_2(t))$ , where for  $m \in \{1, 2\}$  one defined  $L_m(t) = \int_0^t \frac{g'_m(s)}{g_m(s)} ds + l_{0,m}$  and  $g_m(t) = z_m(y + t(x' - y))$ . It follows from the above that we cannot have

$$|arg(z_1(y)) - arg(z_2(y))| < \epsilon'$$

and

$$|\arg(z_1(x')) - \arg(z_2(x'))| < \epsilon'$$

for both  $y$  and  $x'$  simultaneously.

Let us assume that

$$|\arg(z_1(x')) - \arg(z_2(x'))| \geq \epsilon'.$$

The other case when  $|\arg(z_1(y)) - \arg(z_2(y))| \geq \epsilon'$  follows similarly. Appealing to the bound

$$|\partial_e (\log |\lambda_{\mathbf{a}_1}(y)| - \log |\lambda_{\mathbf{a}_2}(y)|)| \leq m$$

which holds for all  $y \in U$  and unit vectors  $e \in \mathbb{R}^d$ , it can be shown using a similar argument to that given above describing how arguments change along paths that there exists  $\epsilon_2 > 0$  depending only on our IFS such that if  $x \in B\left(x', \frac{\epsilon_2}{|b|}\right) \subset B\left(y, \frac{\epsilon_1 D}{|b|}\right)$  then

$$|\arg(z_1(x)) - \arg(z_2(x))| \geq \epsilon'/2.$$

Applying Lemma C.7 and our assumption  $|u| \leq H$ , we can assert that there exists  $\eta_0 = \eta_0(M, \epsilon')$  such that one of the following holds for all  $x \in B\left(x', \frac{\epsilon_2}{|b|}\right)$ : either

$$\left| w_{\mathbf{a}_1}(x) |\lambda_{\mathbf{a}_1}(x)|^{ib} u(f_{\mathbf{a}_1}(x)) + w_{\mathbf{a}_2}(x) |\lambda_{\mathbf{a}_2}(x)|^{ib} u(f_{\mathbf{a}_2}(x)) \right| \leq \eta_0 w_{\mathbf{a}_1}(x) H(f_{\mathbf{a}_1}(x)) + w_{\mathbf{a}_2}(x) H(f_{\mathbf{a}_2}(x)),$$

or

$$\left| w_{\mathbf{a}_1}(x) |\lambda_{\mathbf{a}_1}(x)|^{ib} u(f_{\mathbf{a}_1}(x)) + w_{\mathbf{a}_2}(x) |\lambda_{\mathbf{a}_2}(x)|^{ib} u(f_{\mathbf{a}_2}(x)) \right| \leq w_{\mathbf{a}_1}(x) H(f_{\mathbf{a}_1}(x)) + \eta_0 w_{\mathbf{a}_2}(x) H(f_{\mathbf{a}_2}(x)),$$

depending on whether  $|z_1(x)|/|z_2(x)| \leq M$  for all  $x \in B\left(x', \frac{\epsilon_2}{|b|}\right)$  or  $|z_2(x)|/|z_1(x)| \leq M$  for all  $x \in B\left(x', \frac{\epsilon_2}{|b|}\right)$ . This completes our proof.  $\square$

**C.4. Verifying Lemma C.5.** Let us now verify all the parts of Lemma C.5. We first note that Proposition C.3 gives the following immediate corollary:

**Corollary C.8.** *Let  $N \in \mathbb{N}$  be sufficiently large and  $|b| > 1$ . If  $u \in C^1(U)$  satisfies  $|u| \leq H$  and  $\|Du\| \leq A|b|H$ , then there exists  $J \in \mathcal{E}_b$  such that*

$$|\mathcal{L}_{ib}^N u| \leq \mathcal{N}_b^J(H).$$

Notice the corollary above gives us the first part of item 3 from Lemma C.5. It remains to prove the remaining parts.

We now prove the first part of Lemma C.5.

*Proof of Lemma C.5(1).* Let  $J \in \mathcal{E}_b$  and  $H \in C_{A|b|}$ . Then we have

$$\mathcal{N}_b^J(H) = \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) \chi(f_{\mathbf{a}}(x)) H(f_{\mathbf{a}}(x)).$$

Applying the product rule for differentiation this yields

$$D_x \mathcal{N}_b^J(H) = \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} D_x w_{\mathbf{a}} \chi(f_{\mathbf{a}}(x)) H(f_{\mathbf{a}}(x)) + w_{\mathbf{a}}(x) D_x \chi(f_{\mathbf{a}}) H(f_{\mathbf{a}}(x)) + w_{\mathbf{a}}(x) \chi(f_{\mathbf{a}}(x)) D_x H(f_{\mathbf{a}}).$$

We will focus on each of the terms appearing in this summation individually. Applying Lemma C.2 to the first term we have

$$\left\| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} D_x w_{\mathbf{a}} \chi(f_{\mathbf{a}}(x)) H(f_{\mathbf{a}}(x)) \right\| \leq C_2 \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) \chi(f_{\mathbf{a}}(x)) H(f_{\mathbf{a}}(x)) \leq C_2 \mathcal{N}_b^J(H)(x).$$

Here  $C_2$  is as in the statement of Lemma C.2. For the second term, it follows by our assumptions on  $\chi$  that

$$\begin{aligned} \left\| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) D_x \chi(f_{\mathbf{a}}) H(f_{\mathbf{a}}(x)) \right\| &\leq |b| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) H(f_{\mathbf{a}}(x)) \\ &\leq 2|b| \mathcal{N}_b^J(H)(x). \end{aligned}$$

In the above we used that  $H = \frac{\chi H}{\chi} \leq \frac{\chi H}{\eta} \leq 2\chi H$  (recall  $\chi$  takes values in  $[\eta, 1]$  and  $\eta \geq 1/2$ ). For the third term we have

$$\begin{aligned} \left\| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) \chi(f_{\mathbf{a}}(x)) D_x H(f_{\mathbf{a}}) \right\| &= \left\| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) \chi(f_{\mathbf{a}}(x)) D_{f_{\mathbf{a}}(x)} H D_x f_{\mathbf{a}} \right\| \\ &\leq \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) \chi(f_{\mathbf{a}}(x)) A |b| H(f_{\mathbf{a}}(x)) \gamma^N \\ &\leq A |b| \gamma^N \mathcal{N}_b^J(H)(x). \end{aligned}$$

Combining the bounds collected above we have

$$\|D\mathcal{N}_b H\| \leq (C_2 + 2|b| + A|b|\gamma^N) \mathcal{N}_b^J(H).$$

Therefore (C.3) implies that for  $N$  sufficiently large

$$\|D\mathcal{N}_b^J H\| \leq A|b| \mathcal{N}_b^J(H).$$

Thus the cone is preserved and our proof is complete.  $\square$

We now set out to prove the second part of Lemma C.5.

*Proof of Lemma C.5(2).* Applying the Cauchy Schwartz inequality and the fact  $\chi \leq 1$ , for all  $x \in X_A$  we have

$$\begin{aligned} |\mathcal{N}_b^J H(x)|^2 &= \left| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) \chi(f_{\mathbf{a}}(x)) H(f_{\mathbf{a}}(x)) \right|^2 \\ &\leq \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) \chi(f_{\mathbf{a}}(x))^2 \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) H(f_{\mathbf{a}}(x))^2 \\ &\leq \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) \chi(f_{\mathbf{a}}(x)) \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) H(f_{\mathbf{a}}(x))^2. \end{aligned}$$

Using (C.1) and the fact  $\chi \leq 1$ , we see that the above implies the following bound for all  $x \in X_A$ :

$$|\mathcal{N}_b^J H(x)|^2 \leq (\mathcal{L}_0 H^2)(x).$$

If  $x \in B(x'_l, \frac{\epsilon_2}{3|b|}) \cap X_A$  then  $\chi(f_{\mathbf{a}_l}(x)) = \eta$ , by the definition of the Dolgopyat operators, thus for these  $x$  the above implies that

$$\begin{aligned} |\mathcal{N}_b^J H(x)|^2 &= \left( \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) - w_{\mathbf{a}_l}(x)(1 - \eta) \right) \left( \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) H(f_{\mathbf{a}}(x))^2 \right) \\ &\leq \kappa (\mathcal{L}_0^N H^2)(x) \end{aligned}$$

where  $\kappa \in (0, 1)$  is given by

$$\kappa = 1 - (1 - \eta) \min_{\mathbf{a} \in \mathcal{A}^N, x \in X_A} \{w_{\mathbf{a}}(x)\}.$$

Here we have again used our assumption (C.1).

Combining the above bounds, we have

$$\begin{aligned} \int |\mathcal{N}_b^J H|^2 d\mu &\leq \kappa \int_{\cup_{l \in S(b)} B(x'_l, \frac{\epsilon_2}{3|b|})} \mathcal{L}_0^N H^2 d\mu + \int_{(\cup_{l \in S(b)} B(x'_l, \frac{\epsilon_2}{3|b|}))^c} \mathcal{L}_0^N H^2 d\mu \\ &= \int \mathcal{L}_0^N H^2 d\mu - (1 - \kappa) \int_{\cup_{l \in S(b)} B(x'_l, \frac{\epsilon_2}{3|b|})} \mathcal{L}_0^N H^2 d\mu \\ &= \int H^2 d\mu - (1 - \kappa) \int_{\cup_{l \in S(b)} B(x'_l, \frac{\epsilon_2}{3|b|})} \mathcal{L}_0^N H^2 d\mu \end{aligned}$$

In the final line we used (2.1). Therefore to complete our proof it suffices to show that there exists  $\delta > 0$  such that

$$\int_{\cup_{l \in S(b)} B(x'_l, \frac{\epsilon_2}{3|b|})} \mathcal{L}_0^N H^2 d\mu \geq \delta \int H^2 d\mu.$$

We start by remarking that since we choose  $\{y_l\}_{l \in S(b)} \subset X_A$  to be a maximal set for which  $B(y_l, \frac{\epsilon_1 D}{|b|}) \cap B(y_{l'}, \frac{\epsilon_1 D}{|b|}) = \emptyset$  for distinct  $l, l' \in S(b)$ , we must have

$$X_A \subset \bigcup_{l \in S(b)} B\left(y_l, \frac{2\epsilon_1 D}{|b|}\right). \quad (\text{C.13})$$

Moreover, using the doubling property of Gibbs measures, there must exists  $C > 0$  such that

$$\mu\left(B\left(y_l, \frac{2\epsilon_1 D}{|b|}\right)\right) \leq C \mu\left(B\left(x'_l, \frac{\epsilon_2}{3|b|}\right)\right).$$

Combining this bound with (2.1), (C.2), (C.13) and the fact  $\mathcal{L}_0^N H^2 \in C_{A|b|}$  for  $N$  sufficiently large, we have

$$\begin{aligned} \int H^2 d\mu &= \int \mathcal{L}_0^N H^2 d\mu \leq \sum_{l \in S(b)} \int_{B(y_l, \frac{2\epsilon_1 D}{|b|})} \mathcal{L}_0^N H^2 d\mu \\ &\leq e^{4\epsilon_1 D A} \sum_{l \in S(b)} \mu\left(B\left(y_l, \frac{2\epsilon_1 D}{|b|}\right)\right) \min_{x \in B(y_l, \frac{2\epsilon_1 D}{|b|})} \mathcal{L}_0^N H^2(x) \\ &\leq e^{4\epsilon_1 D A} C \sum_{l \in S(b)} \mu\left(B\left(x'_l, \frac{\epsilon_2}{3|b|}\right)\right) \min_{x \in B(y_l, \frac{2\epsilon_1 D}{|b|})} \mathcal{L}_0^N H^2(x) \\ &\leq e^{4\epsilon_1 D A} C \int_{\cup_{l \in S(b)} B(x'_l, \frac{\epsilon_2}{3|b|})} \mathcal{L}_0^N H^2 d\mu \end{aligned}$$

In the last line we used that  $B(x'_l, \frac{\epsilon_2}{3|b|}) \subset B(y_l, \frac{2\epsilon_1 D}{|b|})$ . Taking  $\delta = (e^{4\epsilon_1 D A} C)^{-1}$  this completes our proof.  $\square$

We will now show that the second part of the third statement in Lemma C.5 is satisfied and in doing so complete our proof of this lemma.

*Proof of Lemma C.5(3).* We have

$$\|D_x \mathcal{L}_{ib}^N u\| = \left\| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} D_x w_{\mathbf{a}} |\lambda_{\mathbf{a}}(x)|^{ib} u(f_{\mathbf{a}}(x)) + w_{\mathbf{a}}(x) D_x |\lambda_{\mathbf{a}}|^{ib} u(f_{\mathbf{a}}(x)) + w_{\mathbf{a}}(x) |\lambda_{\mathbf{a}}(x)|^{ib} D_x u(f_{\mathbf{a}}) \right\|.$$

We will consider each of the three terms above individually. In what follows  $C_2$  and  $C_3$  are as in the statement of Lemma C.2. Applying Lemma C.2 to the first term we have

$$\left\| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} D_x w_{\mathbf{a}} |\lambda_{\mathbf{a}}(x)|^{ib} u(f_{\mathbf{a}}(x)) \right\| \leq C_2 \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) H(f_{\mathbf{a}}(x)).$$

Here we have used that  $|u| \leq H$ . Recalling that  $H \leq 2\chi H$ , the above implies

$$\left\| \sum_{\mathbf{a} \in \mathcal{A}^N} D_x w_{\mathbf{a}} |\lambda_{\mathbf{a}}(x)|^{ib} u(f_{\mathbf{a}}(x)) \right\| \leq 2C_2 \mathcal{N}_b^J H.$$

For the second term, using Lemma C.2 and  $H \leq 2\chi H$  we have

$$\left\| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) D_x |\lambda_{\mathbf{a}}|^{ib} u(f_{\mathbf{a}}(x)) \right\| \leq C_3 |b| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) |u(f_{\mathbf{a}}(x))| \leq 2C_3 |b| \mathcal{N}_b^J H.$$

For the third term we have

$$\begin{aligned} \left\| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) |\lambda_{\mathbf{a}}(x)|^{ib} D_x u(f_{\mathbf{a}}) \right\| &= \left\| \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) |\lambda_{\mathbf{a}}(x)|^{ib} D_{f_{\mathbf{a}}(x)} u D_x f_{\mathbf{a}} \right\| \\ &\leq \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) \|D_{f_{\mathbf{a}}(x)} u\| \gamma^n \\ &\leq \sum_{\substack{\mathbf{a} \in \mathcal{A}^N \\ \mathbf{a} \rightsquigarrow x}} w_{\mathbf{a}}(x) A |b| H(f_{\mathbf{a}}(x)) \gamma^n \\ &\leq 2A |b| \gamma^n \mathcal{N}_b^J H(x). \end{aligned}$$

In the final line we used  $H \leq 2\chi H$ . Summarising the above, we have

$$\|D_x \mathcal{L}_{ib}^N u\| \leq (2C_2 + 2C_3 |b| + 2A |b| \gamma^n) \mathcal{N}_b^J H.$$

Which by (C.4) implies

$$\|D_x \mathcal{L}_{ib}^N u\| \leq A |b| \mathcal{N}_b^J H(x).$$

This completes our proof.  $\square$

#### APPENDIX D. EQUIVALENCE OF THE UNI CONDITIONS

In this section we will prove that our UNI condition is equivalent to the UNI condition stated by Li and Pan [54]. To help with our exposition we recall its statement below. Our proof of this proposition is based upon an argument of Avila, Gouëzel and Yoccoz [8].

**Proposition D.1.** *Let  $\Phi = \{f_a : a \in \mathcal{A}\}$  be a conformal IFS (i.e.  $D_x f_a = \lambda_a(x) O_a(x)$ ,  $x \in \mathbb{R}^d$ , for some  $\lambda_a(x) \in (-1, 1)$  and  $O_a(x) \in \text{SO}(d)$ ) satisfying the strong separation condition and  $\Sigma_A$  be a subshift of finite type. Assume that there exists  $\epsilon_0 > 0$  such that the following holds for infinitely many  $n \in \mathbb{N}$ : There exists  $x \in X_A$  such that for any unit vector  $e \in \mathbb{R}^d$  there exists  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  such that:*

- We have

$$|\partial_e (\log |\lambda_{\mathbf{a}_1}(x)| - \log |\lambda_{\mathbf{a}_2}(x)|)| \geq \epsilon_0.$$

- $\mathbf{a}_1 \rightsquigarrow x$  and  $\mathbf{a}_2 \rightsquigarrow x$ .

*Then there exists  $r > 0$  and  $\epsilon_0 > 0$  such that for any large  $n \in \mathbb{N}$ , any  $x \in X_A$  and unit vector  $e \in \mathbb{R}^d$ , there exist  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  such that:*

- For all  $y \in B(x, r)$  we have

$$|\partial_e (\log |\lambda_{\mathbf{a}_1}(y)| - \log |\lambda_{\mathbf{a}_2}(y)|)| \geq \epsilon_0.$$

- For all  $y \in B(x, r)$  we have  $\mathbf{a}_1 \rightsquigarrow y$  and  $\mathbf{a}_2 \rightsquigarrow y$ .

*Proof.* Recall that  $U = \bigcup_{a \in \mathcal{A}} U_a$  is a choice of open set satisfying  $X_\Phi \subset U$ , where the  $U_a$  are disjoint open sets with  $f_a(U) \subset U_a$  for all  $a \in \mathcal{A}$ . We begin our proof by observing the following equality:

$$\log |\lambda_{\mathbf{a}}(x)| = \sum_{i=1}^n \log |\lambda_{a_i}(f_{a_{i+1} \dots a_n}(x))|$$

for any  $x \in U$  and  $\mathbf{a} \in \mathcal{W}_A \cap \mathcal{A}^n$  satisfying  $\mathbf{a} \rightsquigarrow x$ . Consequently, for any unit vector  $e \in \mathbb{R}^d$ , if we apply the chain rule we have

$$\partial_e \log |\lambda_{\mathbf{a}}(x)| = \left( \sum_{i=1}^n \frac{D_{f_{a_{i+1} \dots a_n}}(x) |\lambda_{a_i}| D_x f_{a_{i+1} \dots a_n}}{|\lambda_{a_i}(f_{a_{i+1} \dots a_n}(x))|} \right) \cdot e. \quad (\text{D.1})$$

Now using the fact that our IFS is uniformly contracting and therefore  $\|D_x f_{a_{i+1} \dots a_n}\| \ll \gamma^{n-i}$  for some  $\gamma \in (0, 1)$ , we can assert that for any  $x \in U$  and  $\mathbf{a} \in \mathcal{W}_A \cap \mathcal{A}^n$  satisfying  $\mathbf{a} \rightsquigarrow x$  we have the following for any  $1 \leq m \leq n$ :

$$\left| \partial_e \log |\lambda_{\mathbf{a}}(x)| - \left( \sum_{i=m}^n \frac{D_{f_{a_{i+1} \dots a_n}}(x) |\lambda_{a_i}| D_x f_{a_{i+1} \dots a_n}}{|\lambda_{a_i}(f_{a_{i+1} \dots a_n}(x))|} \right) \cdot e \right| \ll \gamma^{n-m}.$$

This is equivalent to

$$|\partial_e \log |\lambda_{\mathbf{a}}(x)| - \partial_e \log |\lambda_{a_{m+1} \dots a_n}(x)|| \ll \gamma^{n-m} \quad (\text{D.2})$$

for  $1 \leq m \leq n$ . Equation (D.2) has two important consequences. The first is that for  $n$  sufficiently large depending on  $\epsilon_0$ , if

$$|\partial_e (\log |\lambda_{\mathbf{a}_1}(x)| - \log |\lambda_{\mathbf{a}_2}(x)|)| \geq \epsilon_0 \quad (\text{D.3})$$

for some  $x \in U$  and  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  satisfying  $\mathbf{a}_1 \rightsquigarrow x$  and  $\mathbf{a}_2 \rightsquigarrow x$ , then

$$|\partial_e (\log |\lambda_{\mathbf{b}_1 \mathbf{a}_1}(x)| - \log |\lambda_{\mathbf{b}_2 \mathbf{a}_2}(x)|)| \geq \frac{\epsilon_0}{2}$$

for any  $\mathbf{b}_1, \mathbf{b}_2 \in \mathcal{W}_A$  satisfying  $\mathbf{b}_1 \rightsquigarrow \mathbf{a}_1$  and  $\mathbf{b}_2 \rightsquigarrow \mathbf{a}_2$ . Consequently, at the cost of swapping  $\epsilon_0$  with a potentially smaller constant, we can assume that the hypothesis of our lemma holds for all  $n$  sufficiently large. The second consequence of (D.2) is that there exists  $r_1 > 0$  depending only on our IFS and  $\epsilon_0$ , such that if (D.3) holds for some  $x \in X_A$ ,  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  satisfying  $\mathbf{a}_1 \rightsquigarrow x$  and  $\mathbf{a}_2 \rightsquigarrow x$  and unit vector  $e \in \mathbb{R}^d$ , then for all  $y \in B(x, r_1)$  we have

$$|\partial_e (\log |\lambda_{\mathbf{a}_1}(y)| - \log |\lambda_{\mathbf{a}_2}(y)|)| \geq \frac{\epsilon_0}{2}, \quad (\text{D.4})$$

$\mathbf{a}_1 \rightsquigarrow y$  and  $\mathbf{a}_2 \rightsquigarrow y$ .

Now let  $n \in \mathbb{N}$  be a large number and  $x \in X_A$  be such that for all unit vectors  $e \in \mathbb{R}^d$  there exists  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  such that (D.3) is satisfied,  $\mathbf{a}_1 \rightsquigarrow x$ , and  $\mathbf{a}_2 \rightsquigarrow x$ . Using that  $\Sigma_A$  is topologically mixing and  $r_1$  depends only on our IFS and  $\epsilon_0$ , we can assert that there exists  $m \in \mathbb{N}$  depending only on our IFS,  $\Sigma_A$  and  $\epsilon_0$  such that for any  $x' \in X_A$  there exists  $\mathbf{c} \in \mathcal{W}_A \cap \mathcal{A}^m$  so that  $\mathbf{c} \rightsquigarrow x'$  and

$$f_{\mathbf{c}}(x') \in B(x, r_1). \quad (\text{D.5})$$

For such an  $x'$  and  $\mathbf{c}$  we have the following for any unit vector  $e \in \mathbb{R}^d$  and  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  satisfying  $\mathbf{a}_1 \rightsquigarrow \mathbf{c}$  and  $\mathbf{a}_2 \rightsquigarrow \mathbf{c}$ :

$$\begin{aligned} & |\partial_e (\log |\lambda_{\mathbf{a}_1 \mathbf{c}}(x')| - \log |\lambda_{\mathbf{a}_2 \mathbf{c}}(x')|)| \\ &= |\partial_e (\log |\lambda_{\mathbf{a}_1}(f_{\mathbf{c}}(x'))| - \log |\lambda_{\mathbf{a}_2}(f_{\mathbf{c}}(x'))|)| \\ &= \left( \sum_{i=1}^n \frac{D_{f_{a_{i+1},1} \dots a_{n,1}}(x) |\lambda_{a_{i,1}}| D_x f_{a_{i+1},1} \dots a_{n,1}}{|\lambda_{a_{i,1}}(f_{a_{i+1},1} \dots a_{n,1}(x))|} - \sum_{i=1}^n \frac{D_{f_{a_{i+1},2} \dots a_{n,2}}(x) |\lambda_{a_{i,2}}| D_x f_{a_{i+1},2} \dots a_{n,2}}{|\lambda_{a_{i,2}}(f_{a_{i+1},2} \dots a_{n,2}(x))|} \right) D_x f_{\mathbf{c}} \cdot e. \end{aligned}$$

In the final line we used the chain rule and (D.1). If we let  $v_{x',e} = D_{x'} f_{\mathbf{c}} \cdot e$  then it follows from the above and (D.1) that

$$|\partial_e (\log |\lambda_{\mathbf{a}_1 \mathbf{c}}(x')| - \log |\lambda_{\mathbf{a}_2 \mathbf{c}'}(x')|)| = \left| \partial_{v_{x',e}} (\log |\lambda_{\mathbf{a}_1}(f_{\mathbf{c}}(x'))| - \log |\lambda_{\mathbf{a}_2}(f_{\mathbf{c}'}(x'))|) \right|. \quad (\text{D.6})$$

The vector  $v_{x',e}$  is not necessarily of unit length and instead satisfies the lower bound  $\|v_{x',e}\| \geq \kappa^m$  for some  $\kappa \in (0, 1)$  depending on the IFS.

Using our assumptions on  $x$ , we can assert that there exists  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A$  such that  $\mathbf{a}_1 \rightsquigarrow x$ ,  $\mathbf{a}_2 \rightsquigarrow x$  and

$$\left| \frac{\partial_{v_{x',e}}}{\|v_{x',e}\|} (\log |\lambda_{\mathbf{a}_1}(x)| - \log |\lambda_{\mathbf{a}_2}(x)|) \right| \geq \epsilon_0.$$

Now using the above together with our lower bound for the norm of  $v_{x',e}$ , (D.4), (D.5), and (D.6), yields that for this choice of  $\mathbf{a}_1, \mathbf{a}_2$  we have  $\mathbf{a}_1 \mathbf{c} \rightsquigarrow x'$ ,  $\mathbf{a}_2 \mathbf{c} \rightsquigarrow x'$  and

$$|\partial_e (\log |\lambda_{\mathbf{a}_1 \mathbf{c}}(x')| - \log |\lambda_{\mathbf{a}_2 \mathbf{c}'}(x')|)| \geq \frac{\epsilon_0 \kappa^m}{2}.$$

Summarising the above, we have shown that for any  $n \in \mathbb{N}$  sufficiently large, for any  $x \in X_A$  and unit vector  $e \in \mathbb{R}^d$  there exists  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  with  $\mathbf{a}_1 \rightsquigarrow x$ , and  $\mathbf{a}_2 \rightsquigarrow x$  such that

$$|\partial_e (\log |\lambda_{\mathbf{a}_1}(x)| - \log |\lambda_{\mathbf{a}_2}(x)|)| \geq \frac{\epsilon_0 \kappa^m}{2}.$$

Appealing again to (D.2) as in the derivation of (D.4), it can be shown that there exists  $r > 0$  depending only on our IFS and  $\epsilon_0$ , such that for any  $n \in \mathbb{N}$  sufficiently large, for any  $x \in X_A$  and unit vector  $e \in \mathbb{R}^d$ , there exists  $\mathbf{a}_1, \mathbf{a}_2 \in \mathcal{W}_A \cap \mathcal{A}^n$  such that for all  $y \in B(x, r)$  and  $\mathbf{a}_1 \rightsquigarrow y$ , and  $\mathbf{a}_2 \rightsquigarrow y$ , we have

$$|\partial_e (\log |\lambda_{\mathbf{a}_1}(y)| - \log |\lambda_{\mathbf{a}_2}(y)|)| \geq \frac{\epsilon_0 \kappa^m}{4}.$$

Recalling that  $m$  and  $\kappa$  only depend on our IFS,  $\Sigma_A$  and  $\epsilon_0$ , we see that we can take  $\frac{\epsilon_0 \kappa^m}{4}$  as our choice of  $\epsilon_0$ . Taking  $r$  as chosen above completes our proof.  $\square$

## REFERENCES

- [1] Amir Algom, Simon Baker, and Pablo Shmerkin. On normal numbers and self-similar measures. *Advances in Mathematics*, 399:108276, 2022.
- [2] Amir Algom, Yuanyang Chang, Meng Wu, and Yu-Liang Wu. Van der Corput and metric theorems for geometric progressions for self-similar measures. *arXiv e-prints*, page arXiv:2401.01120, January 2024.
- [3] Amir Algom, Federico Rodriguez Hertz, and Zhiren Wang. Logarithmic Fourier decay for self conformal measures. *J. Lond. Math. Soc. (2)*, 106(2):1628–1661, 2022.
- [4] Amir Algom, Federico Rodriguez Hertz, and Zhiren Wang. Pointwise normality and Fourier decay for self-conformal measures. *Adv. Math.*, 393:Paper No. 108096, 72, 2021.
- [5] Amir Algom, Federico Rodriguez Hertz, and Zhiren Wang. Polynomial Fourier decay and a cocycle version of Dolgopyat’s method for self conformal measures. *arXiv e-prints*, page arXiv:2306.01275, June 2023.
- [6] Jasmina Angelevska, Antti Käenmäki, and Sascha Troscheit. Self-conformal sets with positive hausdorff measure. *Bulletin of the London Mathematical Society*, 52(1):200–223, 2020.
- [7] Jayadev Athreya, Semyon Dyatlov, and Nicholas Miller. Semiclassical measures for complex hyperbolic quotients. *arXiv e-prints*, page arXiv:2402.06477, February 2024.
- [8] Artur Avila, Sébastien Gouëzel, and Jean-Christophe Yoccoz. Exponential mixing for the Teichmüller flow. *Publ. Math. Inst. Hautes Études Sci.*, 104:143–211, 2006.
- [9] Artur Avila, Mikhail Lyubich, and Zhiyuan Zhang. To appear.
- [10] Aidan Backus, James Leng, and Zhongkai Tao. The fractal uncertainty principle via Dolgopyat’s method in higher dimensions. *arXiv e-prints*, page arXiv:2302.11708, February 2023.
- [11] Simon Baker and Amlan Banaji. Polynomial Fourier decay for fractal measures and their pushforwards. *arXiv e-prints*, page arXiv:2401.01241, January 2024.
- [12] Simon Baker and Tuomas Sahlsten. Spectral gaps and Fourier dimension for self-conformal sets with overlaps. *arXiv e-prints*, page arXiv:2306.01389, June 2023.
- [13] Amlan Banaji and Han Yu. Nonlinear images of self-similar sets and measures (to appear).

- [14] Jean Bourgain and Semyon Dyatlov. Fourier dimension and spectral gaps for hyperbolic surfaces. *Geom. Funct. Anal.*, 27(4):744–771, 2017.
- [15] Jean Bourgain and Semyon Dyatlov. Spectral gaps without the pressure condition. *Ann. of Math. (2)*, 187(3):825–867, 2018.
- [16] B.H. Bowditch. Geometrical finiteness for hyperbolic groups. *Journal of Functional Analysis*, 113(2):245–317, 1993.
- [17] Rufus Bowen. *Equilibrium states and the ergodic theory of Anosov diffeomorphisms*. Edited by Jean-René Chazottes, volume 470 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, revised edition, 2008. With a preface by David Ruelle.
- [18] Julien Brémont. Self-similar measures and the Rajchman property. *Ann. H. Lebesgue*, 4:973–1004, 2021.
- [19] Laura Cladek and Terence Tao. Additive energy of regular measures in one and higher dimensions, and the fractional uncertainty principle. *Ars Inveniendi Analytica*, 6 2021.
- [20] Alex Cohen. Fractal uncertainty in higher dimensions. *arXiv e-prints*, page arXiv:2305.05022, May 2023.
- [21] Xin-Rong Dai, De-Jun Feng, and Yang Wang. Refinable functions with non-integer dilations. *J. Funct. Anal.*, 250(1):1–20, 2007.
- [22] Tushar Das, Lior Fishman, David Simmons, and Mariusz Urbański. Extremality and dynamically defined measures, part II: measures from conformal dynamical systems. *Ergodic Theory Dynam. Systems*, 41(8):2311–2348, 2021.
- [23] H. Davenport, P. Erdős, and W. J. LeVeque. On Weyl’s criterion for uniform distribution. *Michigan Math. J.*, 10:311–314, 1963.
- [24] Yiftach Dayan, Arijit Ganguly, and Barak Weiss. Random walks on tori and normal numbers in self-similar sets. *American Journal of Mathematics*, 2020.
- [25] Dmitry Dolgopyat. On decay of correlations in Anosov flows. *Ann. of Math. (2)*, 147(2):357–390, 1998.
- [26] Semyon Dyatlov. An introduction to fractal uncertainty principle. *J. Math. Phys.*, 60(8):081505, 31, 2019.
- [27] Semyon Dyatlov and Long Jin. Dolgopyat’s method and the fractal uncertainty principle. *Analysis & PDE*, 11(6):1457 – 1485, 2018.
- [28] Semyon Dyatlov and Joshua Zahl. Spectral gaps, additive energy, and a fractal uncertainty principle. *Geom. Funct. Anal.*, 26(4):1011–1094, 2016.
- [29] Semyon Dyatlov and Maciej Zworski. Fractal uncertainty for transfer operators. *Int. Math. Res. Not. IMRN*, 3:781–812, 2020.
- [30] Sam Edwards, Minju Lee, and Hee Oh. Torus counting and self-joinings of Kleinian groups. *arXiv e-prints*, page arXiv:2210.10229, October 2022.
- [31] Kenneth Falconer and Xiong Jin. Exact dimensionality and projection properties of Gaussian multiplicative chaos measures. *Trans. Am. Math. Soc.*, 372(4):2921–2957, 2019.
- [32] Daniel Galicer, Santiago Saglietti, Pablo Shmerkin, and Alexia Yavicoli.  $L^q$  dimensions and projections of random measures. *Nonlinearity*, 29(9):2609–2640, 2016.
- [33] Christophe Garban and Vincent Vargas. Harmonic analysis of gaussian multiplicative chaos on the circle, 2024.
- [34] Rui Han and Wilhelm Schlag. A higher-dimensional Bourgain-Dyatlov fractal uncertainty principle. *Anal. PDE*, 13(3):813–863, 2020.
- [35] Xiaolong Han and Pouria Salekani. Fractal uncertainty principle for random cantor sets, 2024.
- [36] Michael Hochman. A short proof of Host’s equidistribution theorem. *Israel J. Math.*, 251(2):527–539, 2022.
- [37] Michael Hochman and Pablo Shmerkin. Equidistribution from fractal measures. *Invent. Math.*, 202(1):427–479, 2015.
- [38] Wassily Hoeffding. Probability inequalities for sums of bounded random variables. *J. Amer. Statist. Assoc.*, 58:13–30, 1963.
- [39] Bernard Host. Nombres normaux, entropie, translations. *Israel J. Math.*, 91(1-3):419–428, 1995.
- [40] Tadeusz Iwaniec and Gaven Martin. Quasiregular semigroups. *Annales Fennici Mathematici*, 21(2):241–254, Aug. 1996.
- [41] Tadeusz Iwaniec and Gaven Martin. *Geometric function theory and non-linear analysis*. Oxford Mathematical Monographs. The Clarendon Press, Oxford University Press, New York, 2001.
- [42] Thomas Jordan and Tuomas Sahlsten. Fourier transforms of Gibbs measures for the Gauss map. *Math. Ann.*, 364(3-4):983–1023, 2016.
- [43] Thomas Jordan and Tuomas Sahlsten. Fourier transforms of Gibbs measures for the Gauss map. *Math. Ann.*, 364(3-4):983–1023, 2016.
- [44] Antti Käenmäki and Tuomas Orponen. Absolute continuity in families of parametrised non-homogeneous self-similar measures. *J. Fractal Geom.*, 10(1-2):169–207, 2023.
- [45] Ilmari Kangasniemi. Sharp cohomological bound for uniformly quasiregularly elliptic manifolds. *American Journal of Mathematics*, 143:1079 – 1113, 2017.

- [46] Ilmari Kangasniemi, Yusuke Okuyama, Pekka Pankka, and Tuomas Sahlsten. Entropy in uniformly quasiregular dynamics. *Ergodic Theory and Dynamical Systems*, 41(8):2397–2427, 2021.
- [47] R. Kaufman. Continued fractions and Fourier transforms. *Mathematika*, 27(2):262–267, 1980.
- [48] Robert Kaufman. On Bernoulli convolutions. In *Conference in modern analysis and probability (New Haven, Conn., 1982)*, volume 26 of *Contemp. Math.*, pages 217–222. Amer. Math. Soc., Providence, RI, 1984.
- [49] Osama Khalil. Exponential Mixing Via Additive Combinatorics. *arXiv e-prints*, page arXiv:2305.00527, April 2023.
- [50] Dmitry Kleinbock, Elon Lindenstrauss, and Barak Weiss. On fractal measures and Diophantine approximation. *Selecta Math. (N.S.)*, 10(4):479–523, 2004.
- [51] Gaétan Leclerc. Julia sets of hyperbolic rational maps have positive Fourier dimension. *Comm. Math. Phys.*, 397(2):503–546, 2023.
- [52] Jialun Li. Decrease of Fourier coefficients of stationary measures. *Math. Ann.*, 372(3-4):1189–1238, 2018.
- [53] Jialun Li, Frédéric Naud, and Wenyu Pan. Kleinian Schottky groups, Patterson-Sullivan measures, and Fourier decay. *Duke Math. J.*, 170(4):775–825, 2021. With an appendix by Li.
- [54] Jialun Li and Wenyu Pan. Exponential mixing of geodesic flows for geometrically finite hyperbolic manifolds with cusps. *Invent. Math.*, 231(3):931–1021, 2023.
- [55] Jialun Li and Tuomas Sahlsten. Fourier transform of self-affine measures. *Adv. Math.*, 374:107349, 35, 2020.
- [56] Jialun Li and Tuomas Sahlsten. Trigonometric series and self-similar sets. *J. Eur. Math. Soc. (JEMS)*, 24(1):341–368, 2022.
- [57] Zhiqiang Li and Tianyi Zheng. Prime orbit theorems for expanding thurston maps: Lattès maps and split ruelle operators. *Advances in Mathematics*, 449:109723, 2024.
- [58] Elon Lindenstrauss and Peter Varju. Random walks in the group of euclidean isometries and self-similar measures. *Duke Mathematical Journal*, 165, 05 2014.
- [59] Pertti Mattila. *Fourier Analysis and Hausdorff Dimension*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, 2015.
- [60] Volker Meyer. Uniformly quasiregular mappings of Lattès type. *Conf. Geom. Dyn.*, 1:104 – 111, 1997.
- [61] Carolina A. Mosquera and Pablo S. Shmerkin. Self-similar measures: asymptotic bounds for the dimension and Fourier decay of smooth images. *Ann. Acad. Sci. Fenn. Math.*, 43(2):823–834, 2018.
- [62] Frédéric Naud. Expanding maps on Cantor sets and analytic continuation of zeta functions. *Ann. Sci. École Norm. Sup. (4)*, 38(1):116–153, 2005.
- [63] Hee Oh and Dale Winter. Prime number theorems and holonomies for hyperbolic rational maps. *Invent. Math.*, 208(2):401–440, 2017.
- [64] Yūsuke Okuyama and Pekka Pankka. Equilibrium measures for uniformly quasiregular dynamics. *Journal of the London Mathematical Society*, 89, 04 2012.
- [65] William Parry and Mark Pollicott. Zeta functions and the periodic orbit structure of hyperbolic dynamics. *Astérisque*, 187-188:268, 1990.
- [66] S. J. Patterson. The limit set of a Fuchsian group. *Acta Math.*, 136(3-4):241–273, 1976.
- [67] Norbert Patzschke. Self-conformal multifractal measures. *Adv. in Appl. Math.*, 19(4):486–513, 1997.
- [68] Frédéric Paulin, Mark Pollicott, and Barbara Schapira. Equilibrium states in negative curvature. *Astérisque*, 373:viii+281, 2015.
- [69] Andrew Pollington, Sanju Velani, Agamemnon Zafeiropoulos, and Evgeniy Zorin. Inhomogeneous diophantine approximation on  $m_0$ -sets with restricted denominators. *International Mathematics Research Notices*, 2022, 12 2020.
- [70] Martine Queffélec and Olivier Ramaré. Analyse de Fourier des fractions continues à quotients restreints. *Enseign. Math. (2)*, 49(3-4):335–356, 2003.
- [71] Ariel Rapaport. On the Rajchman property for self-similar measures on  $\mathbb{R}^d$ . *Adv. Math.*, 403:Paper No. 108375, 53, 2022.
- [72] Thomas Roblin. Ergodicité et équidistribution en courbure négative. *Mém. Soc. Math. Fr. (N.S.)*, 95:vi+96, 2003.
- [73] Thomas Roblin. Un théorème de Fatou pour les densités conformes avec applications aux revêtements galoisiens en courbure négative. *Israel J. Math.*, 147:333–357, 2005.
- [74] Eino Rossi and Pablo Shmerkin. On measures that improve  $L^q$  dimension under convolution. *Rev. Mat. Iberoam.*, 36(7):2217–2236, 2020.
- [75] Santiago Saglietti, Pablo Shmerkin, and Boris Solomyak. Absolute continuity of non-homogeneous self-similar measures. *Adv. Math.*, 335:60–110, 2018.
- [76] Tuomas Sahlsten. Fourier transforms and iterated function systems. *arXiv e-prints*, page arXiv:2311.00585, November 2023.
- [77] Tuomas Sahlsten and Connor Stevens. Fourier transform and expanding maps on cantor sets. *American Journal of Mathematics*, pages 1–32, 2023.

- [78] Pratyush Sarkar and Dale Winter. Exponential mixing of frame flows for convex cocompact hyperbolic manifolds. *arXiv e-prints*, page arXiv:2004.14551, April 2020.
- [79] Pablo Shmerkin. On the exceptional set for absolute continuity of Bernoulli convolutions. *Geom. Funct. Anal.*, 24(3):946–958, 2014.
- [80] Pablo Shmerkin and Ville Suomala. Spatially independent martingales, intersections, and applications. *Memoirs of the American Mathematical Society*, 251, 09 2014.
- [81] David Simmons and Barak Weiss. Random walks on homogeneous spaces and diophantine approximation on fractals. *Inventiones mathematicae*, 216, 05 2019.
- [82] Boris Solomyak. Fourier decay for self-similar measures. *Proc. Amer. Math. Soc.*, 149(8):3277–3291, 2021.
- [83] Boris Solomyak. Fourier decay for homogeneous self-affine measures. *J. Fractal Geom.*, 9(1-2):193–206, 2022.
- [84] Luchezar Stoyanov. Spectra of Ruelle transfer operators for axiom A flows. *Nonlinearity*, 24(4):1089–1120, 2011.
- [85] Lauritz Streck. On absolute continuity and maximal Garsia entropy for self-similar measures with algebraic contraction ratio. *arXiv e-prints*, page arXiv:2303.07785, March 2023.
- [86] Dennis Sullivan. The density at infinity of a discrete group of hyperbolic motions. *Inst. Hautes Études Sci. Publ. Math.*, 50:171–202, 1979.
- [87] Terence Tao. *Higher order Fourier analysis*, volume 142 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2012.
- [88] Masato Tsujii. On the Fourier transforms of self-similar measures. *Dyn. Syst.*, 30(4):468–484, 2015.
- [89] Masato Tsujii and Zhiyuan Zhang. Smooth mixing Anosov flows in dimension three are exponentially mixing. *Ann. of Math. (2)*, 197(1):65–158, 2023.
- [90] Mariusz Urbański. Diophantine approximation and self-conformal measures. *J. Number Theory*, 110(2):219–235, 2005.
- [91] Péter P. Varjú and Han Yu. Fourier decay of self-similar measures and self-similar sets of uniqueness. *Anal. PDE*, 15(3):843–858, 2022.
- [92] Peter Walters. *An introduction to ergodic theory*, volume 79 of *Graduate Texts in Mathematics*. Springer-Verlag, New York-Berlin, 1982.

DEPARTMENT OF MATHEMATICAL SCIENCES, LOUGHBOROUGH UNIVERSITY, LOUGHBOROUGH, LE11 3TU, UK  
*Email address:* `simonbaker412@gmail.com`

DEPARTMENT OF MATHEMATICS, STATISTICS, AND COMPUTER SCIENCE, UNIVERSITY OF ILLINOIS CHICAGO, IL  
*Email address:* `okhalil@uic.edu`

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF HELSINKI, FINLAND  
*Email address:* `tuomas.sahlsten@helsinki.fi`