



# Zimmer’s conjecture for actions of $SL(m, \mathbb{Z})$

Aaron Brown<sup>1</sup> · David Fisher<sup>2</sup> ·  
Sebastian Hurtado<sup>1</sup>

Received: 21 November 2017 / Accepted: 27 February 2020 /  
Published online: 16 March 2020  
© Springer-Verlag GmbH Germany, part of Springer Nature 2020

**Abstract** We prove Zimmer’s conjecture for  $C^2$  actions by finite-index subgroups of  $SL(m, \mathbb{Z})$  provided  $m > 3$ . The method utilizes many ingredients from our earlier proof of the conjecture for actions by cocompact lattices in  $SL(m, \mathbb{R})$  (Brown et al. in Zimmer’s conjecture: subexponential growth, measure rigidity, and strong property (T), 2016. [arXiv:1608.04995](https://arxiv.org/abs/1608.04995)) but new ideas are needed to overcome the lack of compactness of the space  $(G \times M)/\Gamma$  (admitting the induced  $G$ -action). Non-compactness allows both measures and Lyapunov exponents to escape to infinity under averaging and a number of algebraic, geometric, and dynamical tools are used control this escape. New ideas are provided by the work of Lubotzky, Mozes, and Raghunathan on the structure of nonuniform lattices and, in particular, of  $SL(m, \mathbb{Z})$  providing a geometric decomposition of the cusp into rank one directions, whose geome-

---

David Fisher was partially supported by NSF Grants DMS-1308291 and DMS-1607041. David Fisher was also partially supported by the University of Chicago, and by NSF Grants DMS 1107452, 1107263, 1107367, “RNMS: Geometric Structures and Representation Varieties” (the GEAR Network) during a visit to the Isaac Newton Institute in Cambridge.

---

✉ David Fisher  
fisherdm@indiana.edu

Aaron Brown  
awb@northwestern.edu

Sebastian Hurtado  
shurtados@uchicago.edu

<sup>1</sup> University of Chicago, Chicago, IL 60637, USA

<sup>2</sup> Indiana University, Bloomington, Bloomington, IN 47401, USA

try is more easily controlled. The proof also makes use of a precise quantitative form of non-divergence of unipotent orbits by Kleinbock and Margulis, and an extension by de la Salle of strong property (T) to representations of nonuniform lattices.

## 1 Introduction

### 1.1 Statement of results

The main result of this paper is the following:

**Theorem A** *Let  $\Gamma$  be a finite-index subgroup of  $\mathrm{SL}(m, \mathbb{Z})$  and let  $M$  be a closed manifold of dimension  $\dim(M) \leq m - 2$ . If  $\alpha: \Gamma \rightarrow \mathrm{Diff}^2(M)$  is a group homomorphism then  $\alpha(\Gamma)$  is finite.<sup>1</sup> In addition, if  $\omega$  is a volume form on  $M$ ,  $m > 2$ , and if  $\dim(M) \leq m - 1$ , then if and  $\alpha: \Gamma \rightarrow \mathrm{Diff}^2(M, \omega)$  is a group homomorphism then  $\alpha(\Gamma)$  is finite.*

For  $m \geq 3$ , we remark that the conclusion of Theorem A is known for actions on the circle by results of Witte Morris [26] (see also [7, 13] for actions by more general lattices on the circle) and for volume-preserving actions on surfaces by results of Franks and Handel and of Polterovich [12, 21]. The proof in this paper requires that  $m \geq 4$  though we expect it can be modified to cover actions by  $\mathrm{SL}(3, \mathbb{Z})$ ; since these results are not new, we only present the case for  $m \geq 4$ . While this is a very special case of Zimmer's conjecture, it is a key example. For instance, the version of Zimmer's conjecture restated by Margulis in his problem list [20] is a special case of Theorem A.

Note that if  $\Gamma$  is a finite-index subgroup of  $\mathrm{SL}(m, \mathbb{Z})$  acting on compact manifold  $M$ , we may induce an action of  $\mathrm{SL}(m, \mathbb{Z})$  on a (possibly non-connected) compact manifold  $\tilde{M} = (\mathrm{SL}(m, \mathbb{R}) \times M) / \sim$  where  $(\gamma, x) \sim (\gamma', x')$  if there is  $\hat{\gamma} \in \Gamma$  with  $\gamma' = \gamma \hat{\gamma}$  and  $x' = \alpha(\hat{\gamma}^{-1})(x)$ . Connectedness of  $M$  is neither assumed nor is it used in either the proof of Theorem A or in [4]. Thus, for the remainder we will simply assume  $\Gamma = \mathrm{SL}(m, \mathbb{Z})$ .

This paper is a first step in extending the results in [4] to the case where  $\Gamma$  is a nonuniform lattice in a split simple Lie group  $G$  and the strategy of the proof of Theorem A relies strongly on the strategy used in [4]. In the remainder of the introduction, we recall the proof in the cocompact case, indicate where the difficulties arise in the nonuniform case, and outline the proof of Theorem A. At the end of the introduction we make some remarks on other approaches and difficulties we encountered.

We recall a key definition from [4]. Let  $\Gamma$  be a finitely generated group. Let  $\ell: \Gamma \rightarrow \mathbb{N}$  denote the word-length function with respect to some choice

<sup>1</sup> After this work was completed, Brown-Damjanovic-Zhang showed that some modifications of our arguments also give a proof for  $C^1$  diffeomorphisms [3].

of finite generating set for  $\Gamma$ . Given a  $C^1$  diffeomorphism  $f: M \rightarrow M$  let  $\|Df\| = \sup_{x \in M} \|D_x f\|$  (for some choice of norm on  $TM$ ).

**Definition 1.1** An action  $\alpha: \Gamma \rightarrow \text{Diff}^1(M)$  has *uniform subexponential growth of derivatives* if

$$\text{for every } \varepsilon > 0, \text{ there is } C_\varepsilon \text{ such that } \|D\alpha(\gamma)\| \leq C_\varepsilon e^{\varepsilon \ell(\gamma)} \text{ for all } \gamma \in \Gamma. \tag{1}$$

The main result of the paper is the following:

**Theorem B** *For  $m \geq 4$ , let  $\Gamma = SL(m, \mathbb{Z})$  and let  $M$  be a closed manifold.*

- (1) *If  $\dim(M) \leq m - 2$  then any action  $\alpha: \Gamma \rightarrow \text{Diff}^2(M)$  has uniform subexponential growth of derivatives;*
- (2) *if  $\omega$  is a volume form on  $M$  and  $\dim(M) \leq m - 1$  then any action  $\alpha: \Gamma \rightarrow \text{Diff}^2(M, \omega)$  has uniform subexponential growth of derivatives.*

To deduce Theorem A from Theorem B, we apply [4, Theorem 2.9] and de la Salle’s recent result establishing strong property (T) for nonuniform lattices [8, Theorem 1.2] and conclude that any action  $\alpha$  as in Theorem A preserves a continuous Riemannian metric. For clarity, we point out that we need de la Salle’s Theorem 1.2 and not his Theorem 1.1 because we need the measures converging to the projection to be positive measures. That Theorem [8, Theorem 1.2] provides positive measures where [8, Theorem 1.1] does not is further clarified in [8, Section 2.3]. Once a continuous invariant metric is preserved, the image of any homomorphism  $\alpha$  in Theorem A is contained in a compact Lie group  $K$ . All such homomorphisms necessarily have finite image due to the presence of unipotent elements in  $SL(m, \mathbb{Z})$ . We remark that while the finiteness of the image of  $\alpha$  was deduced using Margulis’s superrigidity theorem in [4], it is unnecessary in the setting of Theorem A since, as any unipotent element of  $SL(m, \mathbb{Z})$  lies in the center of some integral Heisenberg subgroup of  $SL(m, \mathbb{Z})$ , all unipotent elements have finite image in  $K$  and therefore so does  $SL(m, \mathbb{Z})$ .

### 1.2 Review of the cocompact case

To explain the proof of Theorem B, we briefly explain the difficulties in extending the arguments from [4] to the setting of actions by nonuniform lattices. We begin by recalling the proof in the cocompact setting.

In both [4] and the proof of Theorem B, we consider a fiber bundle

$$M \rightarrow M^\alpha := (G \times M) / \Gamma \xrightarrow{\pi} G / \Gamma$$

which allows us to replace the  $\Gamma$ -action on  $M$  with a  $G$ -action on  $M^\alpha$ . In the case that  $\Gamma$  is cocompact, showing subexponential growth of derivatives of the  $\Gamma$ -action is equivalent to showing subexponential growth of the fiberwise derivative cocycle for the  $G$ -action.

To prove such subexponential growth for the  $G$ -action on  $M^\alpha$  we argued by contradiction to obtain a sequence of points  $x_n \in M^\alpha$  and semisimple elements  $a_n$  in a Cartan subgroup  $A \subset G$  which satisfy  $\|D_{x_n} a_n|_F\| \geq e^{\lambda d(a_n, \text{Id})}$  for some  $\lambda > 0$ . Here  $D_x g$  denotes the derivative of translation by  $g$  at  $x \in M^\alpha$ ,  $F$  is the fiberwise tangent bundle of  $M^\alpha$ , and  $D_{x_n} a_n|_F$  is the restriction of  $D_{x_n} a_n$  to  $F(x_n)$ .

The pairs  $(x_n, a_n)$  determine empirical measures  $\mu_n$  on  $M^\alpha$  supported on the orbit  $\{a_n^s(x_n) : 0 \leq s \leq t_n\}$  which accumulate on a measure  $\mu$  that is  $a$ -invariant for some  $a \in A$  and has a positive Lyapunov exponent for the fiberwise derivative cocycle of size at least  $\lambda$ . Using classical results in homogeneous dynamics in conjunction with the key proposition from [6], we averaged the measure  $\mu$  to obtain a  $G$ -invariant measure  $\mu'$  on  $M^\alpha$  with a non-zero fiberwise Lyapunov exponent; the existence of such a measure  $\mu'$  contradicts Zimmer's cocycle superrigidity theorem.

### 1.3 Difficulties in the nonuniform setting

When  $\Gamma$  is nonuniform the space  $M^\alpha$  is not compact and the sequence of empirical measures  $\mu_n$  might diverge to infinity in  $M^\alpha$ ; that is, in the limit we might have a "loss of mass". Additionally, even if the measures  $\{\mu_n\}$  satisfy some tightness criteria so as to prevent escape of mass, one might have "escape of Lyapunov exponents:" for a limiting measure  $\mu$ , the Lyapunov exponents may be infinite or the value could drop below the value expected by the growth of fiberwise cocycles along the orbits  $\{a^s(x_n) : 0 \leq s \leq t_n\}$ . For instance, the contribution to the exponential growth of derivatives along the sequence of empirical measures could arise primarily from excursions of orbits deep into the cusp. If one makes naïve computations with the *return cocycle*  $\beta: G \times G/\Gamma \rightarrow \Gamma$  (measuring for  $x$  in a fundamental domain  $D$  the element of  $\Gamma$  needed to bring  $gx$  back to a  $D$ ) one in fact expects that the fiberwise derivative are very large for translations of points far out in the cusp since the orbits of such points cross a large number of fundamental domains. The weakest consequence of this observation is that subexponential growth of the fiberwise derivative of the induced  $G$ -action is much stronger than subexponential growth of derivatives of the  $\Gamma$ -action. While we still work with the induced  $G$ -action and the fiberwise derivative in many places, the arguments become more complicated than in the cocompact case.

In the homogeneous dynamics literature, there are many tools to study escape of mass. Controlling the escape of Lyapunov exponents seems to be

more novel. To rule out escape of mass, it suffices to prove tightness of family of measures  $\{\mu_n\}$ . To control Lyapunov exponents, we introduce a quantitative tightness condition: we construct measures  $\{\mu_n\}$  with *uniformly exponentially small mass in the cusps*. See Sect. 3.2. It is a standard computation to show the Haar measure on  $\mathrm{SL}(m, \mathbb{R})/\mathrm{SL}(m, \mathbb{Z})$  (or any  $G/\Gamma$  where  $G$  is semisimple and  $\Gamma$  is a lattice) has exponentially small mass in the cusps.

## 1.4 Outline of proof

With the above difficulties in mind, we outline the strategy of the proof of Theorem B. Lubotzky, Mozes and Raghunathan proved that  $\mathrm{SL}(m, \mathbb{Z})$  is quasi-isometrically embedded in  $\mathrm{SL}(m, \mathbb{R})$ . And in this special case, they give a proof that every element  $\gamma \in \mathrm{SL}(m, \mathbb{Z})$  can be written as a product of at most  $m^2$  elements  $\delta_i$  contained in canonical copies of  $\mathrm{SL}(2, \mathbb{Z})$  determined by pairs of standard basis vectors for  $\mathbb{R}^m$ ; moreover the word-length of each  $\delta_i$  is at most proportional to the word-length of  $\gamma$  [17, Corollary 3]. (We note however that such effective generation of  $\Gamma$  only holds for  $\mathrm{SL}(m, \mathbb{Z})$ ; for the general case, in [18] a weaker generation of  $\Gamma$  in terms of  $\mathbb{Q}$ -rank 1 subgroups is shown.) Thus, to show uniform subexponential growth of derivatives for the action of  $\mathrm{SL}(m, \mathbb{Z})$ , it suffices to show uniform subexponential growth of derivatives for the restriction of our action to each canonical copy of  $\mathrm{SL}(2, \mathbb{Z})$ .

We first obtain uniform subexponential growth of derivatives for the unipotent elements in  $\mathrm{SL}(2, \mathbb{Z})$  in Sect. 4. See Proposition 4.1. The strategy is to consider a subgroup of the form  $\mathrm{SL}(2, \mathbb{Z}) \times \mathbb{Z}^2 \subset \mathrm{SL}(m, \mathbb{Z})$ . We first prove that a large proportion of elements in  $\mathrm{SL}(2, \mathbb{Z})$  satisfy (1). To prove this, we use that if  $a^t := \mathrm{diag}(e^t, e^{-t})$  then a typical  $a^t$ -orbit in  $\mathrm{SL}(2, \mathbb{R})/\mathrm{SL}(2, \mathbb{Z})$  equidistributes to the Haar measure. In particular, for the empirical measures along such  $a$ -orbits we apply the techniques from [4] to show subexponential growth of fiberwise derivatives along such orbits and conclude that a large proportion of  $\mathrm{SL}(2, \mathbb{Z})$  satisfies (1). See Proposition 4.2. The proof of this fact repeats most of the ideas and techniques from [4] as well a quantitative non-divergence of unipotent averages following Kleinbock and Margulis. The precise averaging procedure is different here than in [4].

Having shown Proposition 4.2, we consider the  $\mathrm{SL}(2, \mathbb{Z})$ -action on the normal subgroup  $\mathbb{Z}^2$  of  $\mathrm{SL}(2, \mathbb{Z}) \times \mathbb{Z}^2$  to show that for every  $n \geq 0$ , the ball  $B_n$  of radius  $n$  in  $\mathbb{Z}^2$  contains a positive-density subset of unipotent elements satisfying (1). Taking iterated sumsets of such good unipotent elements of  $B_n(\mathbb{Z}^2)$  with a finite set one obtains uniform subexponential growth of derivatives for every element in  $B_n$ . This relies heavily on the fact that  $\mathbb{Z}^2$  is abelian. See Sect. 4.2.

It is worth noting that the subgroups of the form  $\mathrm{SL}(2, \mathbb{Z}) \times \mathbb{Z}^2 \subset \Gamma$  are also considered in the work of Lubotzky, Mozes, and Raghunathan in [17] as

well as in Margulis's early constructions of expander graphs and subsequent work on property (T) and expanders [19].

Having established Proposition 4.1, we assume for the sake of contradiction that the restriction of  $\alpha$  to  $\mathrm{SL}(2, \mathbb{Z})$  fails to exhibit uniform subexponential growth of derivatives. We obtain in Sect. 5.2 a sequence  $\zeta_n$  of  $a^t$ -orbit segments in  $\mathrm{SL}(2, \mathbb{R})/\mathrm{SL}(2, \mathbb{Z})$  which drift only a sub-linear distance into the cusp with respect to their length and accumulate exponential growth of the fiberwise derivative. Here we use that orbits deep in the cusp of  $\mathrm{SL}(2, \mathbb{R})/\mathrm{SL}(2, \mathbb{Z})$  correspond to unipotent deck transformations and that Proposition 4.1 implies that these do not contribute to the exponential growth of the fiberwise derivative. Here, we heavily use the structure of  $\mathrm{SL}(2, \mathbb{Z})$  subgroups.

We promote the family of orbit segments  $\zeta_n$  in  $M^\alpha$  to a family of measures  $\{\mu_n\}$  all of whose subsequential limits are  $A$ -invariant measures  $\mu$  on  $M^\alpha$  with non-zero fiberwise exponents. To construct  $\mu_n$ , we construct a Følner sequence  $F_n \subset G$  inside a solvable subgroup  $AN'$  where  $A$  is the full Cartan subgroup of  $\mathrm{SL}(m, \mathbb{R})$  and  $N'$  is a well-chosen abelian subgroup of unipotent elements. We average our orbit segments  $\zeta_n$  over  $F_n$  to obtain the sequence of measures  $\mu_n$  in  $M^\alpha$ . In general, Følner sets for  $AN'$  are subsets which are linearly large in the  $A$ -direction and exponentially large in the  $N'$  direction. In our case the  $N'$ -part will not affect the Lyapunov exponent because we work inside a subset where the return cocycle  $\beta$  restricted to  $N'$  takes unipotent values and we have already proven subexponential growth of the fiberwise derivatives for unipotent elements.

The fact that  $\mu_n$  behaves well in the cusp is due to two facts: First, the segments obtained in Sect. 5.2 do not drift too deep into the cusp of  $\mathrm{SL}(2, \mathbb{R})/\mathrm{SL}(2, \mathbb{Z})$ . Second, we choose our subgroup  $N'$  such that the  $N'$ -orbits of each point along each  $\zeta_n$  is a closed torus that is well-behaved when translated by  $A$ . The argument here is related to the fact closed horocycles in the cusp of  $\mathrm{SL}(2, \mathbb{R})/\mathrm{SL}(2, \mathbb{Z})$  equidistribute to the Haar measure when flowed backwards by the geodesic flow.

To finish the argument, we show that any  $AN'$ -invariant measure on  $M^\alpha$  projects to Haar measure on  $\mathrm{SL}(m, \mathbb{R})/\mathrm{SL}(m, \mathbb{Z})$  using Ratner's measure classification and equidistribution theorems. Then, as in [4], we can use [6, Proposition 5.1] and argue as in the cocompact case in [4] show that  $\mu$  is in fact  $G$ -invariant and thereby obtain a contradiction with Zimmer's cocycle superrigidity theorem.

## 1.5 A few remarks on other approaches

We close the introduction by making some remarks on other approaches, particularly other approaches for controlling the escape of mass. We emphasize here that one key difficulty for all approaches is that we are not able to con-

trol the “images” of the cocycle  $\beta: G \times G/\Gamma \rightarrow \Gamma$  in either our special case or in general. To understand this remark better, consider first the case where  $G = SL(2, \mathbb{R})$  and  $\Gamma = SL(2, \mathbb{Z})$ . If we take a one-parameter subgroup  $c(t) < SL(2, \mathbb{R})$  and take the trajectory  $c(t)x$  for  $t$  in some interval  $[0, T]$  and assume and assume the entire trajectory on  $G/\Gamma$  lies deep enough in the cusp, then  $\beta(a(t), x)$  is necessarily unipotent for all  $t$  in  $[0, T]$ . No similar statement is true for  $G = SL(m, \mathbb{R})$  and  $\Gamma = SL(m, \mathbb{Z})$ . In fact analogous statements are true if and only if  $\Gamma$  has  $\mathbb{Q}$ -rank one, this is closely related to the fact that higher  $\mathbb{Q}$ -rank locally symmetric spaces are 1-connected at infinity. This forces us to “factor” the action into actions of rank-one subgroups in order to control the growth of derivatives.

One might hope to obtain subexponential growth of derivatives more directly for all elements of  $SL(2, \mathbb{Z})$ , or even directly in  $SL(m, \mathbb{Z})$ , by proving better estimates on the size of the “generic” subsets of  $SL(2, \mathbb{R})$  (or  $SL(m, \mathbb{R})$ ) whose  $A$ -orbits define empirical measures satisfying some tightness condition. While one can get good estimates on the size of the sets in Proposition 4.2 using Margulis functions and large deviation estimates as in [1, 10], the resulting estimates are not sharp enough to allow us to prove subexponential growth of derivatives. One can compare with the conjectures in [14] about loss of mass.

An elementary related question is the following: Let  $B_n$  be a ball of radius  $n$  in a Lie group  $G$  (or a lattice  $\Gamma$ ) and suppose there exists subset  $S_n$  of  $B_n$  such that  $S_n$  and  $B_n$  have more or less equal mass, meaning that:

$$\frac{vol(B_n \setminus S_n)}{vol(B_n)} < \varepsilon_n$$

for a certain sequence  $\varepsilon_n$  of numbers converging to zero. Does there exists an integer  $k$  (independent of  $n$ ) such that for  $n$  large:

$$B_n \subset S_n * S_n * \dots * S_n \tag{2}$$

Observe that the question depends on how fast  $\varepsilon_n$  is decreasing and on the group  $G$ . For example if  $G$  abelian,  $\varepsilon_n$  can be a sufficiently small constant as a consequence of Proposition 4.9. Also, it is not hard to see that for any group  $G$  the existence of  $k$  is guaranteed if  $\varepsilon_n$  decreases exponentially quickly. So the real question is how fast  $\varepsilon_n$  has to decrease to zero in order for this statement to hold. Does (2) holds for  $G = SL_3(\mathbb{Z})$  and  $\varepsilon_n = 2^{-n^c}$  for some  $c < 1$ ? If the answer to this question is yes, then it would be possible to approach our results via Margulis functions and large deviation estimates.

## 2 Standing notation

We review the notation introduced in [4] and establish some standing notation and conventions as well as state some facts used in the remainder of the paper.

### 2.1 Lie theoretic and geometric notation

We write  $G = \mathrm{SL}(m, \mathbb{R})$  and  $\Gamma = \mathrm{SL}(m, \mathbb{Z})$ . Let  $\mathfrak{g}$  denote the Lie algebra of  $G$ . Let  $\mathrm{Id}$  denote the identity element of  $G$ . We fix the standard Cartan involution  $\theta: \mathfrak{g} \rightarrow \mathfrak{g}$  given by  $\theta(X) = -X^t$  and write  $\mathfrak{k}$  and  $\mathfrak{p}$ , respectively, for the  $+1$  and  $-1$  eigenspaces of  $\theta$ . Define  $\mathfrak{a}$  to be a maximal abelian subalgebra of  $\mathfrak{p}$ . Then  $\mathfrak{a}$  is the vector space of diagonal matrices.

The roots of  $\mathfrak{g}$  are the linear functionals  $\beta_{i,j} \in \mathfrak{a}^*$  defined as

$$\beta_{i,j}(\mathrm{diag}(t_1, \dots, t_m)) = t_i - t_j.$$

The simple positive roots are  $\alpha_j = \beta_{j,j+1}$  and the positive roots are the positive integral combinations of  $\{\alpha_j\}$  that are still roots.

For a root  $\beta$ , write  $\mathfrak{g}^\beta$  for the associated root space. Each root space  $\mathfrak{g}^\beta$  exponentiates to a 1-parameter unipotent subgroup  $U^\beta \subset G$ . The Lie subalgebra  $\mathfrak{n}$  generated by all root spaces  $\mathfrak{g}^\beta$  for positive roots  $\beta$ , coincides with the Lie algebra of all strictly upper-triangular matrices.

Let  $A$ ,  $N$ , and  $K$  be the analytic subgroups of  $G$  corresponding to  $\mathfrak{a}$ ,  $\mathfrak{n}$  and  $\mathfrak{k}$ . Then

- (1)  $A = \exp(\mathfrak{a})$  is the group of all diagonal matrices with positive entries.  $A$  is an abelian group and we identify linear functionals on  $\mathfrak{a}$  with linear functionals on  $A$  via the exponential map  $\exp: \mathfrak{a} \rightarrow A$ ;
- (2)  $N = \exp(\mathfrak{n})$  is the group of upper-triangular matrices with 1s on the diagonal;
- (3)  $K = \mathrm{SO}(m)$ .

The Weyl group of  $G$  is the group of permutation matrices. This acts transitively on the set of all roots  $\Sigma$ .

For  $1 \leq i, j \leq m$ , the subgroup of  $G$  generated by  $U^{\beta_{i,j}}$  and  $U^{\beta_{j,i}}$  is isomorphic to  $\mathrm{SL}(2, \mathbb{R})$ . We denote this subgroup by  $H_{i,j} = \mathrm{SL}_{i,j}(2, \mathbb{R})$ . Then  $\Lambda_{i,j} := H_{i,j} \cap \Gamma$  is a lattice in  $\mathrm{SL}_{i,j}(2, \mathbb{R})$  isomorphic to  $\mathrm{SL}(2, \mathbb{Z})$ . Note then that  $X_{i,j} := H_{i,j}/\Lambda_{i,j}$  is the unit tangent bundle to the modular surface. We will use the standard notation  $E_{i,j}$  for an elementary matrix with 1s on the diagonal and in the  $(i, j)$ -place and 0s everywhere else. Note that  $E_{i,j}$  and  $E_{j,i}$  generate  $\Lambda_{i,j}$ .

We equip  $G$  with a left- $K$ -invariant and right- $G$ -invariant metric. Such a metric is unique up to scaling. Let  $d$  denote be the induced distance on  $G$ . With



respect to this metric and distance  $d$ , each  $H_{i,j}$  is geodesically embedded. By rescaling the metric, we may assume the restriction of  $d$  to each  $H_{i,j}$  coincides with the standard metric of constant curvature  $-1$  on the upper half plane  $SO(2)\backslash SL(2, \mathbb{R})$ . This metric has the following properties that we exploit throughout.

- (1) For any matrix norm  $\| \cdot \|$  on  $H_{i,j} \simeq SL(2, \mathbb{R})$  there is a  $C_1$  such that

$$2 \log \|A\| - C_1 \leq d(A, \text{Id}) \leq 2 \log \|A\| + C_1 \tag{3}$$

for all  $A \in H_{i,j}$ .

- (2) Let  $B(\text{Id}, r)$  denote the metric ball of radius  $r$  in  $H_{i,j}$  centered at  $\text{Id}$ . Then with respect to the induced Riemannian volume on  $H_{i,j}$  we have

$$\text{vol}(B(\text{Id}, r)) = 4\pi(\cosh(r) - 1) \leq 4\pi e^r$$

and for all sufficiently large  $r > 0$

$$\text{vol}(B(x, r)) \geq e^r. \tag{4}$$

- (3) For any matrix norm  $\| \cdot \|$  on  $SL(m, \mathbb{R})$ , there are constants  $C_0 > 1$  and  $\kappa > 1$  such that for any matrix  $A \in SL(m, \mathbb{R})$  we have

$$\kappa^{-1} \log \|A\| - C_0 \leq d(A, \text{Id}) \leq \kappa \log \|A\| + C_0. \tag{5}$$

- (4) In particular, there are  $C_2$  and  $C_3$  so that if  $E_{i,j} \in SL(m, \mathbb{Z})$  is an elementary unipotent matrix then

$$d(E_{i,j}^k, \text{Id}) \leq C_2 \log k + C_3. \tag{6}$$

## 2.2 Suspension space and induced $G$ -action

Let  $M^\alpha = (G \times M)/\Gamma$  be the fiber-bundle over  $SL(m, \mathbb{R})/SL(m, \mathbb{Z})$  obtained as follows: on  $G \times M$  let  $\Gamma$  act as

$$(g, x) \cdot \gamma = (g\gamma, \alpha(\gamma^{-1})(x))$$

and let  $G$  act as

$$g' \cdot (g, x) = (g'g, x).$$

The  $G$ -action on  $G \times M$  descends to a  $G$ -action on the quotient  $M^\alpha = (G \times M)/\Gamma$ . Let  $\pi : M^\alpha \rightarrow SL(m, \mathbb{R})/SL(m, \mathbb{Z})$  be the canonical projection. As in

[4], we write  $F = \ker D\pi$  for the fiberwise tangent bundle to  $M^\alpha$ . Write  $\mathbb{P}F$  for the projectivization of the fiberwise tangent bundle. We write  $D_xg|_F : F(x) \rightarrow F(gx)$  for the fiberwise derivative as in [4]. For  $(x, [v]) \in \mathbb{P}F$  and  $g \in G$ , write

$$g \cdot (x, [v]) := (g \cdot x, [D_xg|_{F(x)}v])$$

for the action of  $g$  on  $\mathbb{P}F$  induced by  $D_xg|_F$ .

We follow [6, Section 2.1] and equip  $G \times M$  with a  $C^1$  Riemannian metric  $\langle \cdot, \cdot \rangle$  with the following properties:

- (1)  $\langle \cdot, \cdot \rangle$  is  $\Gamma$ -invariant.
- (2) for  $x \in M$  and  $g \in G$ , under the canonical identification of the  $G$ -orbit of  $(g, x)$  with  $G$ , the restriction of  $\langle \cdot, \cdot \rangle$  to the  $G$ -orbit of  $(g, x)$  coincides with the fixed right-invariant metric on  $G$ .
- (3) There is a Siegel fundamental set  $D \subset G$  and  $C > 1$  such that for any  $g_1, g_2 \in D$ , the map  $(g_1, x) \mapsto (g_2, x)$  distorts the restrictions of  $\langle \cdot, \cdot \rangle$  to  $\{g_1\} \times M$  and  $\{g_2\} \times M$  by at most  $C$ .

The metric then descends to a  $C^1$  Riemannian metric on  $M^\alpha$ . Note that by averaging the metric over the left action of  $K$ , we may also assume that the metric on  $M^\alpha$  is left- $K$ -invariant. This, in particular, implies the right-invariant metric on  $G$  in (2) above is chosen to be left- $K$ -invariant.

To analyze the coarse dynamics of the suspension action, it is often useful to consider the *return cocycle*  $\beta : G \times G/\Gamma \rightarrow \Gamma$ . This cocycle is defined relative to a fundamental domain  $\mathcal{F}$  for the right  $\Gamma$ -action on  $G$ . For any  $x \in G/\Gamma$ , take  $\tilde{x}$  to be the unique lift of  $x$  in  $\mathcal{F}$  and define  $\beta(g, x)$  to be the unique element of  $\gamma \in \Gamma$  such that  $g\tilde{x}\gamma^{-1} \in \mathcal{F}$ . Any two choices of fundamental domain for  $\Gamma$  define cohomologous cocycles but we require a choice of well-controlled fundamental domains  $\mathcal{F}$ . Namely, we choose  $\mathcal{F}$  to either be contained in a Siegel fundamental set or to be a Dirichlet domain for the identity. With these choices, we have the following.

Let  $\tilde{\mathcal{D}} \subset \text{SL}(m, \mathbb{R})$  denote the Dirichlet domain of the identity for the  $\text{SL}(m, \mathbb{Z})$  action on  $\text{SL}(m, \mathbb{R})$ ; that is

$$\tilde{\mathcal{D}} := \{g \in \text{SL}(n, \mathbb{R}) : d(g, \text{Id}) \leq d(g\gamma, \text{Id}) \text{ for all } \gamma \in \text{SL}(m, \mathbb{Z})\}.$$

Since each  $H_{i,j}$  is geodesically embedded in  $\text{SL}(m, \mathbb{R})$  and since  $\Lambda_{i,j} = H_{i,j} \cap \text{SL}(m, \mathbb{Z})$ , it follows

$$\mathcal{D} := H_{1,2} \cap \tilde{\mathcal{D}} \tag{7}$$

is a Dirichlet domain of the identity for the  $\Lambda_{1,2}$ -action on  $H_{1,2}$ . Viewing  $H_{1,2} \simeq \text{SL}(2, \mathbb{R})$  acting on the upper half-plane model of hyperbolic space  $\mathbb{H}^2 = \text{SO}(2)\backslash\text{SL}(2, \mathbb{R})$  by Möbius transformations  $\text{SO}(2)\backslash\mathcal{D}$  is the standard

Dirichlet domain for the modular surface, the hyperbolic triangle with endpoints at  $1/2 + i\sqrt{3}/2, -1/2 + i\sqrt{3}/2,$  and  $\infty$ .

**Lemma 2.1** *If  $\mathcal{F}$  is either contained in either a Siegel fundamental set or a Dirichlet domain for the identity then there is a constant  $C$  such that for all  $g \in G$  and  $x \in G/\Gamma$*

$$\ell(\beta(g, x)) \leq Cd(g, e) + Cd(x, \Gamma) + C.$$

In the above lemma,  $\ell$  is the word-length of  $\beta(g, x), d(g, e)$  is the distance from  $g$  to  $e$  in  $G,$  and  $d(x, \Gamma)$  is the distance from  $x \in G/\Gamma$  to the identity coset  $\Gamma$  in  $G/\Gamma.$  For a Dirichlet domain for the identity, the Lemma is shown in [25, §2]; for fundamental domains contained in Siegel fundamental sets, the estimate follows from [11, Corollary 3.19] and the fact that the distance to the identity in a Siegel domain is quasi-Lipschitz equivalent to the distance to the identity in the quotient  $G/\Gamma.$  Both estimates heavily use the main theorem of Lubotzky, Mozes, and Raghunathan [17, 18] to compare the word-length of  $\beta(g, x) \in SL(m, \mathbb{Z})$  with  $\log(\|\beta(g, x)\|).$

Fix once and for all a fundamental domain  $\mathcal{F} \subset \tilde{\mathcal{D}} \subset SL(m, \mathbb{R}).$

The estimates in Lemma 2.1 is often used to obtain integrability properties of  $\beta$  and related cocycles with respect to the Haar measure on  $G/\Gamma.$  As the function  $x \mapsto d(x, \Gamma)$  is in  $L^p(G/\Gamma, \text{Haar})$  for any compact set  $K \subset G$  we have that

$$x \mapsto \sup_{g \in K} \ell(\beta(g, x))$$

is in  $L^p(G/\Gamma, \text{Haar})$  for all  $p \geq 1.$  In the sequel, we typically do not directly use the integrability properties (since we work with measures other than Haar) but rather the estimate in Lemma 2.1.

### 3 Preliminaries on measures, averaging, and Lyapunov exponents

We present a number of technical facts regarding invariant measures, equidistribution, averaging, and Lyapunov exponents that will be used in the remainder of the paper.

#### 3.1 Ratner’s measure classification and equidistribution theorems

We recall Ratner’s theorems on equidistribution of unipotent flows. Let  $U = \{u(t) = \exp_{\mathfrak{g}}(tX)\}$  be a 1-parameter unipotent subgroup in  $G.$  Given any Borel

probability measure  $\mu$  on  $G/\Gamma$  let

$$U^T * \mu := \frac{1}{T} \int_0^T u(t)_* \mu \, dt.$$

**Theorem 3.1** (Ratner) *Let  $U = \{u(t) = \exp_{\mathfrak{g}}(tX)\}$  be a 1-parameter unipotent subgroup and consider the action on  $G/\Gamma$ . The following hold:*

- Every ergodic,  $U$ -invariant probability measure on  $G/\Gamma$  is homogeneous [22, Theorem 1].
- The orbit closure  $\mathcal{O}_x := \overline{\{u \cdot x : u \in U\}}$  is homogeneous for every  $x \in G/\Gamma$  [22, Theorem 3].
- The orbit  $U \cdot x$  equidistributes in  $\mathcal{O}_x$ ; that is  $U^T * \delta_x$  converges to the Haar measure on  $\mathcal{O}_x$  as  $T \rightarrow \infty$ .
- Let  $\beta$  be a root of  $\mathfrak{g}$  and let  $\mathfrak{sl}_{\beta}(2) \subset \mathfrak{g}$  be the Lie subalgebra generated by  $\mathfrak{g}^{\beta}$  and  $\mathfrak{g}^{-\beta}$ . Let  $e, f, h \subset \mathfrak{sl}(2, \mathbb{R})$  triple with  $e \in \mathfrak{g}^{\beta}$  and  $f \in \mathfrak{g}^{-\beta}$  and let  $\mathfrak{h}^{\beta} = \text{span}(h)$ . Let  $H^{\beta} = \exp \mathfrak{h}^{\beta}$ . Let  $\mu$  be a  $U^{\beta}$ -invariant Borel probability measure on  $G/\Gamma$ . If  $\mu$  is  $H^{\beta}$ -invariant, then  $\mu$  is  $U^{-\beta}$ -invariant.

Conclusion (d) follows from [23, Proposition 2.1] and the structure of  $\mathfrak{sl}(2, \mathbb{R})$ -triples. See also the discussion in the paragraph preceding [22, Theorem 9]. In our earlier work on cocompact lattices [4], we averaged over higher-dimensional unipotent subgroups and required a variant of (c) due to Nimish Shah [24]. Here we only average over one-dimensional root subgroups and can use the earlier version due to Ratner.

From Theorem 3.1, for any probability measure  $\mu$  on  $G/\Gamma$  it follows that the weak-\* limit

$$U * \mu := \lim_{T \rightarrow \infty} U^T * \mu$$

exists and that the  $U$ -ergodic components of  $U * \mu$  are homogeneous.

### 3.2 Measures with exponentially small mass in the cusps

We now define precisely the notion of measures with exponentially small mass in the cusps from the introduction. Let  $(X, d)$  be a complete, second countable, metric space. Then  $X$  is Polish. Let  $\mu$  be a finite Borel (and hence Radon) measure on  $X$ . We say that  $\mu$  has *exponentially small mass in the cusps with exponent  $\eta_{\mu}$*  if for all  $0 < \eta < \eta_{\mu}$

$$\int_X e^{\eta d(x_0, x)} \, d\mu(x) < \infty \tag{8}$$

for some (and hence any) choice of base point  $x_0 \in X$ . We say that a collection  $\mathcal{M} = \{\mu_\zeta\}$  of probability measures on  $X$  has *uniformly exponentially small mass in the cusps with exponent  $\eta_0$*  if for all  $0 < \eta < \eta_0$

$$\sup_{\mu_\zeta \in \mathcal{M}} \left\{ \int e^{\eta d(x_0, x)} d\mu_\zeta(x) \right\} < \infty.$$

Below, we often work in the setting  $X = G/\Gamma$  where  $G = SL(m, \mathbb{R})$  and  $\Gamma = SL(m, \mathbb{Z})$  and where  $d$  the distance induced from a right-invariant metric on  $G$ . When  $X = SL(m, \mathbb{R})/SL(m, \mathbb{Z})$  we interpret a point  $x = g\Gamma \in G/\Gamma$  as a unimodular lattice  $\Lambda_g = g \cdot \mathbb{Z}^m$ . Fix any norm on  $\mathbb{R}^m$  and define the systole of a lattice  $\Lambda \subset \mathbb{R}^m$  to be

$$\delta(\Lambda) := \inf \{ \|v\| : v \in \Lambda \setminus \{0\} \}.$$

We have that

$$c_1 \leq \frac{1 - \log(\delta(\Lambda_g))}{1 + (d(g\Gamma, e\Gamma))} \leq c_2 \tag{9}$$

for some constants whence

$$C_1 e^{c_1 d(g\Gamma, e\Gamma)} \leq \frac{1}{\delta(\Lambda_g)} \leq C_2 e^{c_2 d(g\Gamma, e\Gamma)}.$$

Thus, if we only care about finding a positive exponent  $\eta_\mu > 0$  such that (8) holds for all  $\eta < \eta_\mu$ , it suffices to find  $\eta$  such that

$$\int \delta(\Lambda_g)^{-\eta} d\mu(g\Gamma) < \infty. \tag{10}$$

We define the *systolic exponent*  $\eta_\mu^S$  to be the supremum of all  $\eta$  satisfying (10).

In the sequel, we will frequently use the following proposition to avoid escape of mass into the cusps of  $G/\Gamma$  when averaging a measure along a unipotent flow.

**Proposition 3.2** *Let  $U$  be a 1-parameter unipotent subgroup of  $G$ . Let  $\mu$  be a probability measure on  $X = SL(m, \mathbb{R})/SL(m, \mathbb{Z})$  with exponentially small mass in the cusps. Then the family of measures*

$$\{U^T * \mu : T \in \mathbb{R}\} \cup \{U * \mu\}$$

*has uniformly exponentially small mass in the cusps.*

### 3.3 Proof of Proposition 3.2

We first show that the family of averaged measures

$$\{U^T * \mu : T \in \mathbb{R}\}$$

has uniformly exponentially small mass in the cusps. The key idea is to use the quantitative non-divergence of unipotent orbits following Kleinbock and Margulis.

**Lemma 3.3** *Let  $\mu$  be a probability measure on  $X = \text{SL}(m, \mathbb{R})/\text{SL}(m, \mathbb{Z})$  with exponentially small mass in the cusps and systolic exponent  $\eta_\mu^S$ .*

*Then the family of measures  $\{U^T * \mu : T \in \mathbb{R}\}$  has uniformly exponentially small mass in the cusps with systolic exponent  $\min\{\eta_\mu^S, \frac{1}{m^2}\}$ .*

*Proof* Let  $\Delta \subset \mathbb{R}^m$  be a discrete subgroup. Let  $\|\Delta\|$  denote the volume of  $\Delta_{\mathbb{R}}/\Delta$  where  $\Delta_{\mathbb{R}}$  denotes the  $\mathbb{R}$ -span of  $\Delta$ . It follows from Minkowski’s lemma that there is a constant  $c_m$  (depending only on  $m$ ) such that if

$$\|\Delta\| \leq (\rho')^{\text{rk}(\Delta)}$$

then there is a non-zero vector  $v \in \Delta$  with  $\|v\| \leq c_m \rho'$ . In particular, if  $\delta(\Delta) \geq \rho$  then for some constant  $c'_m$  we have

$$\|\Delta\| \geq (c'_m \rho)^{\text{rk}(\Delta)}$$

for all discrete subgroups  $\Delta \subset \Lambda$ .

From [15, Theorem 5.3] as extended in [16, Theorem 0.1], there is a  $C > 1$  such that for every  $\Lambda_g \in G/\Gamma$  and  $\varepsilon > 0$ , if  $\delta(\Lambda_g) \geq \rho$  then, since  $\|\Delta\| \geq (c'_m \rho)^{\text{rk}(\Delta)}$  for every discrete subgroup  $\Delta \subset \Lambda_g$ , we have

$$m\{t \in [0, T] : \delta(\Lambda_{u_t g}) \leq \varepsilon\} \leq C \left(\frac{\varepsilon}{(c'_n)^{-1} \rho}\right)^{\frac{1}{m^2}} T = \hat{C} \left(\frac{\varepsilon}{\rho}\right)^{\frac{1}{m^2}} T \tag{11}$$

where  $m(A)$  is the Lebesgue measure of the set  $A \subset \mathbb{R}$ . Note that (11) still holds even in the case  $\varepsilon \geq \rho$ . Note that if  $\beta < \frac{1}{m^2}$  then for  $\varepsilon < \rho$  we have

$$\left(\frac{\varepsilon}{\rho}\right)^{\frac{1}{m^2}} T < \left(\frac{\varepsilon}{\rho}\right)^\beta T.$$

In particular, when  $\beta < \frac{1}{m^2}$  we have (for all  $\varepsilon > 0$  including  $\varepsilon > \delta(\Lambda_g)$ ) that

$$m\{t \in [0, T] : \delta(\Lambda_{u_t g}) \leq \varepsilon\} \leq \hat{C} \left(\frac{\varepsilon}{\delta(\Lambda_g)}\right)^\beta T.$$

Then for  $\eta > 0$  and  $\beta < \frac{1}{m^2}$  we have

$$\begin{aligned} & \int [\delta(\Lambda_g)]^{-\eta} dU^T * \mu(g) \\ &= \int_M \frac{1}{T} \int_0^T [\delta(\Lambda_{u_t g})]^{-\eta} dt d\mu(g) \\ &= \int_M \frac{1}{T} \int_0^\infty m\{t \in [0, T] : [\delta(\Lambda_{u_t g})]^{-\eta} \geq \ell\} d\ell d\mu(g) \\ &\leq \int_M \frac{1}{T} \left[ T + \int_1^\infty m\{t \in [0, T] : [\delta(\Lambda_{u_t g})]^{-\eta} \geq \ell\} d\ell \right] d\mu(g) \\ &= 1 + \int_M \frac{1}{T} \int_1^\infty m\{t \in [0, T] : [\delta(\Lambda_{u_t g})] \leq \ell^{-\frac{1}{\eta}}\} d\ell d\mu(g) \\ &\leq 1 + \int_M \frac{1}{T} \int_1^\infty \left[ \hat{C} \left( \frac{1}{\ell^{\frac{1}{\eta}} \delta(\Lambda_g)} \right)^\beta T \right] d\ell d\mu(g) \\ &= 1 + \hat{C} \left( \int_M \left( \frac{1}{\delta(\Lambda_g)} \right)^\beta d\mu(g) \right) \left( \int_1^\infty \left( \frac{1}{\ell^{\frac{1}{\eta}}} \right)^\beta d\ell \right) \end{aligned}$$

which is uniformly bounded in  $T$  as long as  $\eta < \beta < \min\{\eta_\mu^S, \frac{1}{m^2}\}$ . □

For the limit measure  $U * \mu = \lim_{T \rightarrow \infty} U^T * \mu$  we have the following which holds in full generality.

**Lemma 3.4** *Let  $(X, d)$  be a complete, second countable, metric space. Let  $\nu_j$  be a sequence of Borel probability measures on  $X$  converging in the weak- $*$  topology to a measure  $\nu$ . If the family  $\{\nu_j\}$  has uniformly exponentially small mass in the cusps with exponent  $\eta_0$  then the limit  $\nu$  has exponentially small mass in the cusps with exponent  $\eta_0$ .*

*Proof* We have that  $\nu_j \rightarrow \nu$  in the weak- $*$  topology. In particular, for any closed set  $C \subset X$  and open set  $U \subset X$  we have

$$\limsup_{j \rightarrow \infty} \nu_j(C) \leq \nu(C) \quad \text{and} \quad \liminf_{j \rightarrow \infty} \nu_j(U) \geq \nu(U).$$

Fix  $0 < \eta' < \eta < \eta_0$  and take  $\delta := \frac{\eta}{\eta'} - 1$ . Fix  $N$  with

$$\int e^{\eta d(x, x_0)} d\nu_j(x) < N$$

for all  $j$ . Using Markov’s inequality, for all  $M > 0$  and every  $j$  we have

$$\nu_j\{x : e^{\eta d(x,x_0)} > M\} \leq N/M$$

so

$$\nu\{x : e^{\eta d(x_0,x)} > M\} \leq N/M.$$

Then, for the limit measure  $\nu$ , we have

$$\begin{aligned} \int_{G/\Gamma} e^{\eta' d(x_0,x)} d\nu(x) &= \int_0^\infty \nu\{x : e^{\eta' d(x_0,x)} \geq M\} dM \\ &= \int_0^\infty \nu\{x : (e^{\eta d(x_0,x)})^{1/(1+\delta)} \geq M\} dM \\ &= \int_0^\infty \nu\{x : e^{\eta d(x_0,x)} \geq M^{1+\delta}\} dM \\ &\leq 1 + \int_1^\infty \frac{N}{M^{1+\delta}} dM. \end{aligned}$$

### 3.4 Averaging certain measures on $SL(m, \mathbb{R})/SL(m, \mathbb{Z})$

Take  $\{\alpha_1, \dots, \alpha_m\}$  to be the standard set of simple positive roots of  $SL(m, \mathbb{R})$ :

$$\alpha_j(\text{diag}(e^{t_1}, \dots, e^{t_m})) = t_j - t_{j+1}.$$

Let  $H_1$  be the analytic subgroup of  $SL(m, \mathbb{R})$  whose Lie algebra is generated by roots spaces associated to  $\{\pm\alpha_1\}$  and let  $H_2$  be the analytic subgroup of  $SL(m, \mathbb{R})$  whose Lie algebra is generated by roots spaces associated to  $\{\pm\alpha_3, \dots, \pm\alpha_n\}$ . We have  $H_1 \cong SL(2, \mathbb{R})$  and  $H_2 \cong SL(m - 2, \mathbb{R})$ . Then  $H = H_1 \times H_2 \subset SL(m, \mathbb{R})$  is the subgroup of all matrices of the form

$$\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$$

where  $\det(B) = \det(C) = 1$ .

We let  $A'$  be the the co-rank-1 subgroup  $A' \subset A$  of the Cartan subgroup  $A$  given by  $A' = A \cap H$ . Let  $\delta = \alpha_1 + \dots + \alpha_n$  be the highest positive root.

**Proposition 3.5** *Let  $\mu$  be any  $H$ -invariant probability on  $SL(m, \mathbb{R})/SL(m, \mathbb{Z})$ . Let  $\beta' = \alpha_2$  or  $\beta' = \delta$  and let  $\hat{\beta} = -\alpha_2$  or  $\hat{\beta} = -\delta$ .*

*Then  $U^{\beta'} * \mu$  is  $H$ -invariant and*

$$U^{\hat{\beta}} * U^{\beta'} * \mu$$



is the Haar measure on  $G/\Gamma$ .

*Proof* We have that  $\mu$  is  $A'$ -invariant. Let  $\mu' = U^{\beta'} * \mu$  and note that  $\mu'$  remains  $H$ - and  $A'$ -invariant.

**Case 1(a)**  $\beta' = \alpha_2$ . Consider first the case that  $\beta' = \alpha_2$ . Then  $\mu'$  remains invariant under  $U^{-\alpha_1}$  and  $U^{-\alpha_j}$  for all  $3 \leq j \leq n$  since these roots commute with  $\beta'$ . By Theorem 3.1(d) we have that  $\mu'$  is also invariant under  $U^{\alpha_1}$  and  $U^{\alpha_j}$  for all  $3 \leq j \leq n$ . Taking brackets,  $\mu'$  is invariant under  $U^\beta$  for every positive root  $\beta \in \Sigma_+$ .

**Case 1(b)**  $\beta' = \delta$ . Consider now the case that  $\beta' = \delta$ . Then  $\mu'$  remains invariant under  $U^{\alpha_1}$  and  $U^{\alpha_j}$  for all  $3 \leq j \leq n$  since these roots commute with  $\delta$ . By Theorem 3.1(d) we have that  $\mu'$  is also invariant under  $U^{-\alpha_1}$  and  $U^{-\alpha_j}$  for all  $3 \leq j \leq n$ . Taking brackets,  $\mu'$  is invariant under  $U^\beta$  for every positive root  $\beta$  of the form  $\delta - \alpha_n - \alpha_{n-1} - \dots - \alpha_j = \alpha_1 + \dots + \alpha_{j-1}$  for each  $j \geq 3$ . In particular,  $\mu'$  is invariant under  $U^{\alpha_1 + \alpha_2}$  and hence also invariant under  $U^{\alpha_2}$ . In particular  $\mu'$  is invariant under  $U^\beta$  for every positive root  $\beta \in \Sigma_+$ .

Note that in either case, we have that  $\mu'$  is invariant under  $U^\beta$  for every positive root  $\beta \in \Sigma_+$ .

Let  $\hat{\mu} = U^{\beta'} * \mu'$ .

**Case 2(a)**  $\hat{\beta} = -\alpha_2$ . If  $\hat{\beta} = -\alpha_2$ , then  $\hat{\mu}$  remains invariant under  $U^{\alpha_1}$  and  $U^{\alpha_j}$  for all  $3 \leq j \leq n$ . Note additionally  $\hat{\mu}$  remains invariant under the highest-root group  $U^\delta$ . Again, by Theorem 3.1(d) we have that  $\hat{\mu}$  is also invariant under  $U^{-\alpha_1}$  and  $U^{-\alpha_j}$  for all  $3 \leq j \leq n$ . In particular  $\hat{\mu}$  is also invariant under  $U^\beta$  for every negative root  $\beta \in \Sigma_-$ . It follows as in Case 1(b) that  $\hat{\mu}$  is invariant under  $U^{\alpha_2}$  and hence invariant under  $U^\beta$  for every positive root  $\beta \in \Sigma_+$ . Thus  $\mu$  is  $G$ -invariant.

**Case 2(b)**  $\hat{\beta} = -\delta$ . If  $\hat{\beta} = -\delta$ , then  $\hat{\mu}$  remains invariant under  $U^{-\alpha_1}$  and  $U^{-\alpha_j}$  for all  $3 \leq j \leq n$ . Note additionally  $\hat{\mu}$  remains invariant under  $U^{\alpha_2}$ . Again, we have that  $\hat{\mu}$  is also invariant under  $U^{\alpha_1}$  and  $U^{\alpha_j}$  for all  $3 \leq j \leq n$ . In particular  $\hat{\mu}$  is also invariant under  $U^\beta$  for every positive root  $\beta \in \Sigma_+$ . As in Case 1(b) that  $\hat{\mu}$  is invariant under  $U^{-\alpha_2}$  and hence invariant under  $U^\beta$  for every negative root  $\beta \in \Sigma_-$ . Thus  $\mu$  is  $G$ -invariant. □

### 3.5 Lyapunov exponents for unbounded cocycles

Let  $(X, d)$  be a second countable, complete metric space. We moreover assume the metric  $d$  is proper. Let  $G$  act continuously on  $X$ .

Let  $\mathcal{E} \rightarrow X$  be a continuous, finite-dimensional vector bundle equipped with a norm  $\|\cdot\|$ . A linear cocycle over the  $G$ -action on  $X$  is an action  $\mathcal{A}: G \times \mathcal{E} \rightarrow \mathcal{E}$

by vector-bundle automorphisms that projects to the  $G$ -action on  $X$ . We write  $\mathcal{A}(g, x)$  for the linear map between Banach spaces  $\mathcal{E}_x$  and  $\mathcal{E}_{g \cdot x}$ . By the norm of  $\mathcal{A}(g, x)$  we mean the operator norm and the conorm is  $m(\mathcal{A}(g, x)) = \|\mathcal{A}(g, x)^{-1}\|^{-1}$ . We say that  $\mathcal{A}$  is *tempered* with respect to the metric  $d$  if there is a  $k \geq 0$  such that for any compact set  $K \subset G$  and base point  $x_0 \in X$  there is  $C > 1$  so that

$$\sup_{g \in K} \|\mathcal{A}(g, x)\| \leq C e^{kd(x, x_0)}$$

and

$$\inf_{g \in K} m(\mathcal{A}(g, x)) \geq \frac{1}{C} e^{-kd(x, x_0)}$$

where  $\|\cdot\|$  denotes the operator norm and  $m(\cdot)$  denotes the operator conorm applied to linear maps between Banach spaces  $\mathcal{E}_x$  and  $\mathcal{E}_{g \cdot x}$ .

If  $\mu$  is a probability measure on  $(X, d)$  with exponentially small mass in the cusps, it follows that the function  $x \mapsto d(x, x_0)$  is  $L^1(\mu)$  whence we immediately obtain the following.

**Claim 3.6** *Let  $\mu$  a probability measure on  $X$  with exponentially small mass in the cusps. Suppose that  $\mathcal{A}$  is tempered. Then for any compact  $K \subset G$ , the functions*

$$x \mapsto \sup_{s \in K} \log \|\mathcal{A}(s, x)\|, \quad x \mapsto \inf_{s \in K} \log m(\mathcal{A}(s, x))$$

are  $L^1(\mu)$ .

Given  $s \in G$  and an  $s$ -invariant Borel probability measure  $\mu$  on  $X$  we define the *average leading* (or *top*) *Lyapunov exponent* of  $\mathcal{A}$  to be

$$\lambda_{\text{top}, s, \mu, \mathcal{A}} := \inf_{n \rightarrow \infty} \frac{1}{n} \int \log \|\mathcal{A}(s^n, x)\| d\mu(x). \quad (12)$$

From the integrability of the function  $x \mapsto \log \|\mathcal{A}(s, x)\|$  we obtain the finiteness of Lyapunov exponents.

**Corollary 3.7** *For  $s \in G$  and  $\mu$  an  $s$ -invariant probability measure on  $X$  with exponentially small mass in the cusps, if  $\mathcal{A}$  is tempered then the average leading Lyapunov exponent  $\lambda_{\text{top}, s, \mu, \mathcal{A}}$  of  $\mathcal{A}$  is finite.*

Note that for an  $s$ -invariant measure  $\mu$ , the sequence  $\int \log \|\mathcal{A}(s^n, x)\| d\mu(x)$  is subadditive whence the infimum in (12) maybe replaced by a limit.

As in the case of bounded continuous linear cocycles, we obtain upper-semicontinuity of leading Lyapunov exponents for continuous tempered

cocycles when restricted to families of measures with uniformly exponentially small measure in the cusp.

**Lemma 3.8** *Let  $A$  be a tempered cocycle. Given  $s \in G$  suppose the restriction of the cocycle  $A: G \times \mathcal{E} \rightarrow \mathcal{E}$  to the action of  $s$  is continuous.*

*Then—when restricted to a set of  $s$ -invariant Borel probability measures with uniformly exponentially small mass in the cusps—the function*

$$\mu \mapsto \lambda_{\text{top},s,\mu,A}$$

*is upper-semicontinuous with respect to the weak- $*$  topology.*

*Proof* Let  $\mathcal{M} = \{\mu_\zeta\}_{\zeta \in \mathcal{I}}$  be a family of  $s$ -invariant Borel probability measures with uniformly exponentially small mass in the cusps. As the pointwise infimum of continuous functions is upper-semicontinuous, is enough to show that the function

$$\mathcal{M} \rightarrow \mathbb{R}, \quad \mu \mapsto \int \log \|\mathcal{A}(s^n, x)\| \, d\mu(x)$$

is continuous with respect to the weak- $*$  topology for each  $n$ . As the weak- $*$  topology is first countable, it is enough to show  $\mu \mapsto \int \log \|\mathcal{A}(s^n, x)\| \, d\mu(x)$  is sequentially continuous.

Let  $\mu_j \rightarrow \mu_\infty$  in  $\mathcal{M}$ . Given  $M > 0$ , fix a continuous  $\psi_M: X \rightarrow [0, 1]$  with

$$\psi_M(x) = 1 \text{ if } d(x, x_0) \leq M \text{ and } \psi_M(x) = 0 \text{ if } d(x, x_0) \geq M + 1.$$

As we assume our metric is proper,  $x \mapsto \psi_M(x) \log \|\mathcal{A}(s^n, x)\|$  is a bounded continuous function whence

$$\int \log \psi_M(x) \log \|\mathcal{A}(s^n, x)\| \, d\mu_j(x) \rightarrow \int \psi_M(x) \log \|\mathcal{A}(s^n, x)\| \, d\mu_\infty(x).$$

Moreover, there are  $C > 1, k \geq 1$ , and  $\eta > 0$  such that for all  $x \in X$  and  $\mu_\zeta \in \mathcal{M}$

$$-\log C - kd(x, x_0) \leq \log \|\mathcal{A}(s^n, x)\| \leq \log C + kd(x, x_0),$$

and

$$\int e^{\eta d(x, x_0)} \, d\mu_\zeta(x) \leq C.$$

In particular,

$$\mu_\zeta(\{x : d(x, x_0) \geq M\}) \leq Ce^{-\eta M}.$$

Thus for any  $\mu_\zeta \in \mathcal{M}$ , we have

$$\begin{aligned} & \int |\log \|\mathcal{A}(s^n, x)\| - \psi_M(x) \log \|\mathcal{A}(s^n, x)\|| \, d\mu_\zeta(x) \\ & \leq \int_{\{x:d(x,x_0)\geq M\}} |\log \|\mathcal{A}(s^n, x)\| - \psi_M(x) \log \|\mathcal{A}(s^n, x)\|| \, d\mu_\zeta(x) \\ & \leq \int_{\{x:d(x,x_0)\geq M\}} |\log \|\mathcal{A}(s^n, x)\|| \, d\mu_\zeta(x) \\ & \leq \int_{\{x:d(x,x_0)\geq M\}} \log C + kd(x, x_0) \, d\mu_\zeta(x) \\ & \leq (\log C)Ce^{-\eta M} + k \int_{\{x:d(x,x_0)\geq M\}} d(x, x_0) \, d\mu_\zeta(x) \\ & \leq (\log C + kM)Ce^{-\eta M} + k \int_{\ell=M}^\infty \mu_\zeta\{x : d(x, x_0) \geq \ell\} \, d\ell \\ & \leq (\log C + kM)Ce^{-\eta M} + k \int_{\ell=M}^\infty Ce^{-\eta\ell} \, d\ell \\ & \leq (\log C + kM)Ce^{-\eta M} + k \frac{Ce^{\eta(-M)}}{\eta}. \end{aligned}$$

It follows that given  $\varepsilon > 0$  there is  $M$  so that

$$\int |\log \|\mathcal{A}(s^n, x)\| - \psi_M(x) \log \|\mathcal{A}(s^n, x)\|| \, d\mu_\zeta(x) \leq \varepsilon$$

for all  $\mu_\zeta \in \mathcal{M}$ .

In particular, taking  $M$  and  $j$  sufficiently large we have

$$\begin{aligned} & \left| \int \log \|\mathcal{A}(s^n, \cdot)\| \, d\mu_\infty - \int \log \|\mathcal{A}(s^n, \cdot)\| \, d\mu_j \right| \\ & \leq \int |\log \|\mathcal{A}(s^n, \cdot)\| - \psi_M \log \|\mathcal{A}(s^n, \cdot)\|| \, d\mu_\infty \\ & \quad + \left| \int \psi_M \log \|\mathcal{A}(s^n, \cdot)\| \, d\mu_j - \int \psi_M \log \|\mathcal{A}(s^n, \cdot)\| \, d\mu_\infty \right| \\ & \quad + \int |\log \|\mathcal{A}(s^n, \cdot)\| - \psi_M \log \|\mathcal{A}(s^n, \cdot)\|| \, d\mu_j \\ & \leq 3\varepsilon. \end{aligned}$$

Sequential continuity then follows. □

### 3.6 Lyapunov exponents under averaging and limits

We now consider the behavior of the top Lyapunov exponent  $\lambda_{\text{top},s,\mu,\mathcal{A}}$  as we average an  $s$ -invariant probability measure  $\mu$  over an amenable subgroup of  $G$  contained in the centralizer of  $s$ .

**Lemma 3.9** *Let  $s \in G$  and let  $\mu$  be an  $s$ -invariant probability measure on  $X$  with exponentially small mass in the cusps. Let  $\mathcal{A}: G \times \mathcal{E} \rightarrow \mathcal{E}$  be a tempered continuous cocycle.*

*For any amenable subgroup  $H \subset C_G(s)$  and any Følner sequence of precompact sets  $F_n$  in  $H$ , if the family  $\{F_n * \mu\}$  has uniformly exponentially small mass in the cusps then for any subsequential limit  $\mu'$  of  $\{F_n * \mu\}$  we have*

$$\lambda_{\text{top},s,\mu,\mathcal{A}} \leq \lambda_{\text{top},s,\mu',\mathcal{A}}.$$

*Proof* First note that Lemma 3.4 implies the family  $\{F_n * \mu\} \cup \{\mu'\}$  has uniformly exponentially small mass in the cusps. Note also that for every  $m$ , the measure  $F_m * \mu$  is  $s$ -invariant.

We first claim that  $\lambda_{\text{top},s,F_m*\mu,\mathcal{A}} = \lambda_{\text{top},s,\mu,\mathcal{A}}$  for every  $m$ .

For  $t \in H$  define  $c_t(x) = \sup\{\|\mathcal{A}(t, x)\|, m(\mathcal{A}(t, x))^{-1}\}$  and let  $c_m(x) = \sup_{t \in F_m} c_t(x)$ . As  $F_m$  is precompact, from Claim 3.6 we have that  $\log c_m \in L^1(\mu)$ .

For  $x \in M$  and  $t \in F_m$ , the cocycle property and subadditivity of norms yields

$$\begin{aligned} \log \|\mathcal{A}(s^n, tx)\| &\leq \log \|\mathcal{A}(t^{-1}, tx)\| + \log \|\mathcal{A}(s^n, x)\| + \log \|\mathcal{A}(t, s^n x)\| \\ &= \log \|\mathcal{A}(t, x)^{-1}\| + \log \|\mathcal{A}(s^n, x)\| + \log \|\mathcal{A}(t, s^n x)\| \\ &\leq \log c_m(x) + \log c_m(s^n(x)) + \log \|\mathcal{A}(s^n, x)\|. \end{aligned}$$

Using that  $\mu$  is  $s$ -invariant, we have for every  $n$  that

$$\begin{aligned} \int \log \|\mathcal{A}(s^n, x)\| d(F_m * \mu)(x) &= \frac{1}{|F_m|} \int_{F_m} \int \log \|\mathcal{A}(s^n, x)\| dt * \mu(x) dt \\ &= \frac{1}{|F_m|} \int_{F_m} \int \log \|\mathcal{A}(s^n, tx)\| d\mu(x) dt \\ &\leq \frac{1}{|F_m|} \int_{F_m} \left( \int \log c_m(x) + \log c_m(s^n(x)) + \log \|\mathcal{A}(s^n, x)\| d\mu(x) \right) dt \\ &\leq 2 \int \log c_m(x) d\mu(x) + \int \log \|\mathcal{A}(s^n, x)\| d\mu(x) \end{aligned}$$

Dividing by  $n$  yields  $\lambda_{\text{top},s,F_n*\mu,\mathcal{A}} \leq \lambda_{\text{top},s,\mu,\mathcal{A}}$ . The reverse inequality is similar.

The inequality then follows from the upper-semicontinuity in Lemma 3.8. □

Consider now any  $Y \in \mathfrak{g}$  with  $\|Y\| = 1$ , a point  $x \in X$ , and  $t > 0$ . The empirical measure  $\eta(Y, t, x)$  along the orbit  $\exp(sY)x$  until time  $t$  is the measure defined as follows: given a bounded continuous  $\phi: X \rightarrow \mathbb{R}$ , the integral of  $\phi$  with respect to the empirical measure  $\eta(Y, t, x)$  is

$$\int \phi \, d\eta(Y, t, x) := \frac{1}{t} \int_0^t \phi(\exp(sY) \cdot x) \, ds.$$

Similarly, given a probability measure  $\mu$  on  $X$ , the empirical distribution  $\eta(Y, t, \mu)$  of  $\mu$  along the orbit of  $\exp(sY)$  until time  $t$  is defined as

$$\int \phi \, d\eta(Y, t, \mu) := \frac{1}{t} \int_X \int_0^t \phi(\exp(sY) \cdot x) \, ds \, d\mu(x).$$

Consider now sequences  $Y_n \in \mathfrak{g}$  with  $\|Y_n\| = 1$  and  $t_n > 0$ . For part (c) of the following lemma, we add an additional assumption that the action of  $G$  on  $(X, d)$  has *uniform displacement*: for any compact  $K \subset G$  there is  $C'$  such that for all  $x \in X$  and  $g \in K$ ,

$$d(g \cdot x, x) \leq C'.$$

**Lemma 3.10** *Suppose the action of  $G$  on  $(X, d)$  has uniform displacement and let  $\mathcal{A}: G \times \mathcal{E} \rightarrow \mathcal{E}$  be a tempered continuous cocycle.*

*Let  $Y_n \in \mathfrak{g}$  and  $t_n \geq 0$  be sequences with  $\|Y_n\| = 1$  for all  $n$  and  $t_n \rightarrow \infty$ . Let  $\mu_n$  be a sequence of Borel probability measures on  $X$  and define  $\eta_n := \eta(Y_n, t_n, \mu_n)$  to be the empirical distribution of  $\mu_n$  along the orbit of  $\exp(sY_n)$  for  $0 \leq s \leq t_n$ . Assume that*

- (1) *the family of empirical distributions  $\{\eta_n\}$  defined above has uniformly exponentially small mass in the cusps; and*
- (2)  $\int \log \|\mathcal{A}(\exp(t_n Y_n), x)\| \, d\mu_n(x) \geq \varepsilon t_n$ .

*Then*

- (a) *the family  $\{\eta_n\}$  is pre-compact;*
- (b) *for any subsequential limit  $Y_\infty = \lim_{j \rightarrow \infty} Y_{n_j}$ , any subsequential limit  $\eta_\infty$  of  $\{\eta_{n_j}\}$  is invariant under the 1-parameter subgroup  $\{\exp(tY_\infty) : t \in \mathbb{R}\}$ ;*
- (c)  $\lambda_{\text{top},\exp(Y_\infty),\eta_\infty,\mathcal{A}} \geq \varepsilon > 0$ .

*Proof of Lemma 3.10(a) and (b)* As in the proof of Lemma 3.8, from the assumption that  $\{\eta_n\}$  has uniformly exponentially small mass in the cusps we obtain uniform bounds

$$\eta_n(\{x : d(x, x_0) \geq \ell\}) \leq Ce^{-\eta\ell}$$

for all  $n$ . Combined with the properness of  $d$ , this establishes uniform tightness of the family of measures  $\{\eta_n\}$  and (a) follows.

For (b), let  $\phi : X \rightarrow \mathbb{R}$  be a compactly supported continuous function. Then for any  $s > 0$

$$\begin{aligned} \int_X \phi \circ \exp(sY_\infty) - \phi \, d\eta_n &= \int_X \phi \circ \exp(sY_\infty) - \phi \circ \exp(sY_n) \, d\eta_n \\ &\quad + \int_X \phi \circ \exp(sY_n) - \phi \, d\eta_n \end{aligned}$$

The first integral converges to zero as the functions  $\phi \circ \exp(wY_\infty) - \phi \circ \exp(wY_n)$  converges uniformly to zero in  $n$  for fixed  $w$ . The second integral clearly converges to zero since for  $t_n \geq s$  we have

$$\begin{aligned} &\int_X \phi \circ \exp(sY_n) - \phi \, d\eta_n \\ &= \frac{1}{t_n} \int_0^{t_n} \int_X \phi(\exp((s+t)Y_n)x) - \phi(\exp(tY_n)x) \, d\mu_n(x) \, dt \\ &= \frac{1}{t_n} \left[ - \int_0^s \int_X \phi(\exp(tY_n)x) \, \mu_n(x) \, dt \right. \\ &\quad \left. + \int_{t_n}^{t_n+s} \int_X \phi(\exp(tY_n)x) \, d\mu_n(x) \, dt \right] \end{aligned}$$

which converges to 0 as  $t_n \rightarrow \infty$  as  $\phi$  is bounded. □

The proof of Lemma 3.10(c) is quite involved. It is the analogue in the non-compact setting of [4, Lemma 3.6]; we recommend the reader read the proof of [4, Lemma 3.6] first. Two technical complications arise in the proof of Lemma 3.10(c). First, we must control for “escape of Lyapunov exponent” as our cocycle is unbounded. Second, in [4] it was sufficient to consider the average of Dirac masses  $\delta_{x_n}$  along a single orbit  $\exp(sY_n)x_n$ ; here we average measures  $\mu_n$  along an orbit of  $\exp(sY_n)$ .

To prove Lemma 3.10(c) we first introduce a number of standard auxiliary objects. Let  $\mathbb{P}\mathcal{E} \rightarrow X$  denote the projectivization of the tangent bundle  $\mathcal{E} \rightarrow X$ . We represent a point in  $\mathbb{P}\mathcal{E}$  as  $(x, [v])$  where  $[v]$  is an equivalence class of non-

zero vectors in the fiber  $\mathcal{E}(x)$ . For each  $n$ , let  $\sigma_n: X \rightarrow \mathcal{E} \setminus \{0\}$  be a nowhere vanishing Borel section such that

$$\|\mathcal{A}(\exp(t_n Y_n), x)(\sigma_n(x))\| \|(\sigma_n(x))\|^{-1} = \|\mathcal{A}(\exp(t_n Y_n), x)\|$$

for every  $x \in X$ . The  $G$ -action on  $\mathcal{E}$  by vector-bundle automorphisms induces a natural  $G$ -action on  $\mathbb{P}\mathcal{E}$  which restricts to projective transformations between each fiber and its image. For each  $n$ , let  $\tilde{\eta}_n$  be the probability measure on  $\mathbb{P}\mathcal{E}$  given as follows: given a bounded continuous  $\phi: \mathbb{P}\mathcal{E} \rightarrow \mathbb{R}$  define

$$\int_{\mathbb{P}\mathcal{E}} \phi \, d\tilde{\eta}_n := \frac{1}{t_n} \int_0^{t_n} \int_X \phi(\exp(t Y_n) \cdot (x, [\sigma_n(x)])) \, d\mu_n(x) \, dt.$$

We have that  $\tilde{\eta}_n$  projects to  $\eta_n$  under the natural projection  $\mathbb{P}\mathcal{E} \rightarrow X$ ; moreover, if  $\eta_{j_k}$  is a subsequence converging to  $\eta_\infty$ , then any weak- $*$  subsequential limit  $\tilde{\eta}_\infty$  of  $\{\tilde{\eta}_{n_{j_k}}\}$  projects to  $\eta_\infty$ .

Define  $\Phi: \mathfrak{g} \times \mathbb{P}\mathcal{E} \rightarrow \mathbb{R}$  by

$$\Phi(Y, (x, [v])) := \log(\|\mathcal{A}(\exp(Y), x)v\| \|v\|^{-1}).$$

Note for each fixed  $Y \in \mathfrak{g}$  that  $\Phi$  satisfies a cocycle property:

$$\Phi((s + t)Y, (x, [v])) = \Phi(tY, (x, [v])) + \Phi(sY, \exp(tY) \cdot (x, [v])) \tag{13}$$

By hypothesis, there are  $C > 1, k \geq 1$ , and  $\eta > 0$  such that

$$\int e^{\eta d(x, x_0)} \, d\eta_n \leq C$$

for all  $n$  and

$$\frac{1}{C} e^{-kd(x, x_0)} \leq \|\mathcal{A}(\exp(Y), x)v\| \|v\|^{-1} \leq C e^{kd(x, x_0)}$$

for all  $(x, [v]) \in \mathbb{P}\mathcal{E}$  and  $Y \in \mathfrak{g}$  with  $\|Y\| \leq 1$ .

For each  $n$ , let

$$M_n(x) = \sup_{0 \leq t \leq t_n} \{d((\exp(t Y_n)x), x_0)\}.$$

As we assume the  $G$ -action on  $(X, d)$  has uniform displacement, take

$$C_1 = \sup_{\|Y\| \leq 1, x \in X} \{d(\exp(Y) \cdot x, x)\}.$$



We have

$$\frac{1}{t_n} \int_0^{t_n} \int_X e^{\eta d((\exp(tY_n)x, x_0))} d\mu_n(x) dt = \int e^{\eta d(x, x_0)} d\eta_n \leq C.$$

If  $t_n \geq 1$  then for every  $x$  there is an interval  $I_x \subset [0, t_n]$  of length 1 on which

$$d((\exp(tY_n)x, x_0)) \geq (M_n(x) - C_1)$$

for all  $t \in I_x$ . It follows that

$$\int_X e^{\eta(M_n(x) - C_1)} d\mu_n(x) \leq \int_X \int_{I_x} e^{\eta d((\exp(tY_n)x, x_0))} dt d\mu_n(x) \leq Ct_n.$$

By Jensen’s inequality we have

$$\int_X \eta(M_n(x) - C_1) d\mu_n(x) \leq \log \int_X e^{\eta(M_n(x) - C_1)} d\mu_n(x)$$

whence

$$\int M_n(x) d\mu_n(x) \leq \eta^{-1}(\log C + \log t_n) + C_1 =: \eta^{-1} \log t_n + C_2.$$

Since  $\|Y_n\| = 1$ , we have

$$\begin{aligned} \sup_{0 \leq t \leq t_n, 0 \leq s \leq 1} \int_X |\Phi(sY_n, \exp(tY_n) \cdot (x, [\sigma_n(x)]))| d\mu_n(x) \\ \leq \int_X \sup_{0 \leq t \leq t_n, 0 \leq s \leq 1} |\Phi(sY_n, \exp(tY_n) \cdot (x, [\sigma_n(x)]))| d\mu_n(x) \\ \leq \int |\log C| + kM_n(x) d\mu_n(x) \\ \leq |\log C| + k(\eta^{-1} \log t_n + C_2) \\ =: k\eta^{-1} \log t_n + C_3. \end{aligned} \tag{14}$$

In particular, we have

$$\begin{aligned} \frac{1}{t_n} \int_X \log \|\mathcal{A}(\exp(t_n Y_n), x)\| d\mu_n(x) \\ = \frac{1}{t_n} \int_X \Phi(t_n Y_n, (x, [\sigma_n(x)])) d\mu_n(x) \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{t_n} \int_X \Phi(\lfloor t_n \rfloor Y_n, (x, [\sigma_n(x)])) \, d\mu_n(x) \\
 &\quad + \frac{1}{t_n} \int_X \Phi((t - \lfloor t_n \rfloor) Y_n, \exp(\lfloor t_n \rfloor Y_n) \cdot (x, [\sigma_n(x)])) \, d\mu_n(x).
 \end{aligned}$$

Since

$$\begin{aligned}
 &\left| \frac{1}{t_n} \int_X \Phi((t - \lfloor t_n \rfloor) Y_n, \exp(\lfloor t_n \rfloor Y_n) \cdot (x, [\sigma_n(x)])) \, d\mu_n(x) \right| \\
 &\leq \frac{1}{t_n} (k\eta^{-1} \log t_n + C_3)
 \end{aligned}$$

goes to 0 as  $t_n \rightarrow \infty$  it follows that

$$\begin{aligned}
 \liminf_{n \rightarrow \infty} \int_X \frac{1}{t_n} \Phi(\lfloor t_n \rfloor Y_n, (x, [\sigma_n(x)])) \, d\mu_n(x) \\
 &= \liminf_{n \rightarrow \infty} \frac{1}{t_n} \int_X \log \|\mathcal{A}(\exp(t_n Y_n), x)\| \, d\mu_n(x) \tag{15} \\
 &\geq \varepsilon > 0.
 \end{aligned}$$

With the above objects and estimates we complete the proof of Lemma 3.10.

*Proof of Lemma 3.10(c)* Consider first the expression  $\int \Phi(Y_n, \cdot) \, d\tilde{\eta}_n$ . We have

$$\begin{aligned}
 &\int \Phi(Y_n, \cdot) \, d\tilde{\eta}_n \\
 &= \frac{1}{t_n} \int_0^{t_n} \int_X \Phi(Y_n, \exp(t Y_n) \cdot (x, [\sigma_n(x)])) \, d\mu_n(x) \, dt \\
 &= \frac{1}{t_n} \int_0^{\lfloor t_n \rfloor} \int_X \Phi(Y_n, \exp(t Y_n) \cdot (x, [\sigma_n(x)])) \, d\mu_n(x) \, dt \\
 &\quad + \frac{1}{t_n} \int_{\lfloor t_n \rfloor}^{t_n} \int_X \Phi(Y_n, \exp(t Y_n) \cdot (x, [\sigma_n(x)])) \, d\mu_n(x) \, dt
 \end{aligned}$$

Note that the contribution of the second integral is bounded by

$$\begin{aligned}
 &\left| \frac{1}{t_n} \int_{\lfloor t_n \rfloor}^{t_n} \int_X \Phi(Y_n, \exp(t Y_n) \cdot (x, [\sigma_n(x)])) \, d\mu_n(x) \, dt \right| \\
 &\leq \frac{1}{t_n} (k\eta^{-1} \log t_n + C_3)
 \end{aligned}$$

which goes to zero as  $t_n \rightarrow \infty$ .

Repeatedly applying the cocycle property (13) of  $\Phi(Y_n, \cdot)$  we have for  $t_n \geq 1$  that

$$\begin{aligned} & \frac{1}{t_n} \int_X \int_0^{\lfloor t_n \rfloor} \Phi(Y_n, \exp(tY_n) \cdot (x, [\sigma_n(x)])) dt d\mu_n(x) \\ &= \frac{1}{t_n} \int_X \int_0^1 \Phi(\lfloor t_n \rfloor Y_n, \exp(tY_n) \cdot (x, [\sigma_n(x)])) dt d\mu_n(x) \\ &= \frac{1}{t_n} \int_X \int_0^1 \left( \Phi(\lfloor t_n \rfloor Y_n, (x, [\sigma_n(x)])) - \Phi(tY_n, (x, [\sigma_n(x)])) \right. \\ &\quad \left. + \Phi(tY_n, \exp(\lfloor t_n \rfloor Y_n) \cdot (x, [\sigma_n(x)])) \right) dt d\mu_n(x) \\ &= \frac{1}{t_n} \int_X \Phi(\lfloor t_n \rfloor Y_n, (x, [\sigma_n(x)])) d\mu_n(x) \\ &\quad + \frac{1}{t_n} \int_X \int_0^1 \left( -\Phi(tY_n, (x, [\sigma_n(x)])) \right. \\ &\quad \left. + \Phi(tY_n, \exp(\lfloor t_n \rfloor Y_n) \cdot (x, [\sigma_n(x)])) \right) dt d\mu_n(x) \end{aligned}$$

From (14), the contribution of the second and third integrals is bounded by

$$\begin{aligned} & \left| \frac{1}{t_n} \int_X \int_0^1 \left( -\Phi(tY_n, (x, [\sigma_n(x)])) + \Phi(tY_n, \exp(\lfloor t_n \rfloor Y_n) \cdot (x, [\sigma_n(x)])) \right) dt d\mu_n(x) \right| \\ & \leq \frac{1}{t_n} \int_0^1 2(k\eta^{-1} \log t_n + C_3) dt \\ & = \frac{1}{t_n} 2(k\eta^{-1} \log t_n + C_3) \end{aligned}$$

which tend to zero as  $t_n \rightarrow \infty$ . We then conclude from (15) that

$$\begin{aligned} \liminf_{n \rightarrow \infty} \int \Phi(Y_n, \cdot) d\tilde{\eta}_n &= \liminf_{n \rightarrow \infty} \frac{1}{t_n} \int_X \Phi(\lfloor t_n \rfloor Y_n, (x, [\sigma_n(x)])) d\mu_n(x) \\ &\geq \varepsilon > 0. \end{aligned} \tag{16}$$

To complete the proof of (c), for  $M > 0$  take  $\psi_M : X \rightarrow [0, 1]$  continuous with

$$\psi_M(x) = 1 \text{ if } d(x, x_0) \leq M \text{ and } \psi_M(x) = 0 \text{ if } d(x, x_0) \geq M + 1.$$

Let  $\Psi_M : \mathbb{P}\mathcal{E} \rightarrow [0, 1]$  be

$$\Psi_M(x, [v]) = \psi_M(x).$$

and define  $\Phi_M : \mathfrak{g} \times \mathbb{P}\mathcal{E} \rightarrow \mathbb{R}$  to be

$$\Phi_M(Y, (x, [v])) := \Psi_M(x, [v])\Phi(Y, (x, [v])).$$

As the family

$$\mathcal{N} = \{\eta_n\} \cup \{\eta_\infty\}$$

has uniformly exponentially small mass in the cusps we have

$$\int e^{\eta d(x, x_0)} d\hat{\eta} < C$$

and hence  $\hat{\eta}\{x : d(x, x_0) \geq \ell\} \leq Ce^{-\eta\ell}$  for all  $\hat{\eta} \in \mathcal{N}$ . It follows for all  $\tilde{\eta} \in \{\tilde{\eta}_n\} \cup \{\tilde{\eta}_\infty\}$  that—letting  $\hat{\eta} \in \mathcal{N}$  denote the image of  $\tilde{\eta}$  in  $X$ —we have for any  $Y \in \mathfrak{g}$  with  $\|Y\| \leq 1$  that

$$\begin{aligned} & \int_{\mathbb{P}\mathcal{E}} |\Phi(Y, \cdot) - \Phi_M(Y, \cdot)| d\tilde{\eta} \\ &= \int_{\{(x, [v]) \in \mathbb{P}\mathcal{E} : d(x, x_0) \geq M\}} |\Phi(Y, \cdot) - \Phi_M(Y, \cdot)| d\tilde{\eta} \\ &\leq \int_{\{(x, [v]) \in \mathbb{P}\mathcal{E} : d(x, x_0) \geq M\}} |\Phi(Y, \cdot)| d\tilde{\eta} \\ &\leq \int_{\{x \in X : d(x, x_0) \geq M\}} \log(C) + kd(x, x_0) d\hat{\eta} \\ &\leq (\log C + kM)Ce^{-\eta M} + k \int_M^\infty \hat{\eta}\{x : d(x, x_0) \geq \ell\} d\ell \\ &\leq (\log C + kM)Ce^{-\eta M} + k \frac{Ce^{\eta(-M)}}{\eta}. \end{aligned}$$

In particular, given any  $\delta > 0$ , by taking  $M > 0$  sufficiently large we may ensure that

$$\int_{\mathbb{P}\mathcal{E}} |\Phi(Y, \cdot) - \Phi_M(Y, \cdot)| d\tilde{\eta} \leq \delta$$

for any

$$\tilde{\eta} \in \{\tilde{\eta}_n\} \cup \{\tilde{\eta}_\infty\}.$$

Since the restriction of  $\Phi_M$  to  $\{Y \in \mathfrak{g} : \|Y\| \leq 1\} \times \mathbb{P}\mathcal{E}$  is compactly supported, it is uniformly continuous whence

$$\int \Phi_M(Y_n, \cdot) d\tilde{\eta}_n - \Phi_M(Y_\infty, \cdot) d\tilde{\eta}_\infty \rightarrow 0$$

as  $n \rightarrow \infty$ . In particular given  $\delta > 0$  we may take  $M$  and  $n$  sufficiently large so that

$$\begin{aligned} & \left| \int_{\mathbb{P}\mathcal{E}} \Phi(Y_n, \cdot) d\tilde{\eta}_n - \int_{\mathbb{P}\mathcal{E}} \Phi(Y_\infty, \cdot) d\tilde{\eta}_\infty \right| \\ & \leq \int_{\mathbb{P}\mathcal{E}} |\Phi(Y_n, \cdot) - \Phi_M(Y_n, \cdot)| d\tilde{\eta}_n \\ & \quad + \int_{\mathbb{P}\mathcal{E}} |\Phi_M(Y_n, \cdot) - \Phi_M(Y_\infty, \cdot)| d\tilde{\eta}_n \\ & \quad + \int_{\mathbb{P}\mathcal{E}} |\Phi(Y_\infty, \cdot) - \Phi_M(Y_\infty, \cdot)| d\tilde{\eta}_\infty \\ & \leq 3\delta. \end{aligned}$$

Let  $g_\infty = \exp(Y_\infty)$ . Note for each  $n$  that

$$\int_X \log \|\mathcal{A}(g_\infty^n, x)\| d\eta_\infty(x) \geq \int_{\mathbb{P}\mathcal{E}} \log(\|\mathcal{A}(g_\infty^n, x)v\| \|v\|^{-1}) d\tilde{\eta}_\infty(x, [v]).$$

It then follows for any  $\delta > 0$

$$\begin{aligned} \lambda_{\text{top}, g_\infty, \eta, \mathcal{A}} &= \lim_{n \rightarrow \infty} \frac{1}{n} \int_X \log \|\mathcal{A}(g_\infty^n, x)\| d\eta_\infty(x) \\ &\geq \liminf_{n \rightarrow \infty} \frac{1}{n} \int_{\mathbb{P}\mathcal{E}} \log(\|\mathcal{A}(g_\infty^n, x)v\| \|v\|^{-1}) d\tilde{\eta}_\infty(x, [v]) \\ &= \liminf_{n \rightarrow \infty} \frac{1}{n} \int_{\mathbb{P}\mathcal{E}} \Phi(nY_\infty, (x, [v])) d\tilde{\eta}_\infty(x, [v]) \\ &= \int_{\mathbb{P}\mathcal{E}} \Phi(Y_\infty, (x, [v])) d\tilde{\eta}_\infty(x, [v]) \\ &\geq \liminf_{n \rightarrow \infty} \int_{\mathbb{P}\mathcal{E}} \Phi(Y_n, \cdot) d\tilde{\eta}_n - 3\delta. \end{aligned}$$

where the third equality follows from the invariance of  $\tilde{\eta}_\infty$  and the cocycle property of  $\Phi$ . Since

$$\liminf_{n \rightarrow \infty} \int_{\mathbb{P}\mathcal{E}} \Phi(Y_n, \cdot) d\tilde{\eta}_n \geq \varepsilon$$

we conclude that

$$\lambda_{\text{top},g_\infty,\eta,\mathcal{A}} \geq \varepsilon - 3\delta$$

for any  $\delta > 0$  whence the result follows. □

### 3.7 Oseledec’s theorem for cocycles over actions by higher-rank abelian groups

Let  $A \subset G$  be a split Cartan subgroup. Then  $A \simeq \mathbb{R}^d$  where  $d$  is the rank of  $G$ . We have the following consequence of the higher-rank Oseledec’s multiplicative ergodic theorem (c.f. [5, Theorem 2.4]).

Fix any norm  $|\cdot|$  on  $A \simeq \mathbb{R}^d$  and let  $\eta: X \rightarrow \mathbb{R}$  be

$$\eta(x) := \sup_{|a| \leq 1} \log \|\mathcal{A}(a, x)\|.$$

**Proposition 3.11** *Let  $\mu$  be an ergodic,  $A$ -invariant Borel probability measure on  $X$  and suppose  $\eta \in L^{d,1}(\mu)$ . Then there are*

- (1) *an  $\alpha$ -invariant subset  $\Lambda_0 \subset X$  with  $\mu(\Lambda_0) = 1$ ;*
- (2) *linear functionals  $\lambda_i: A \rightarrow \mathbb{R}$  for  $1 \leq i \leq p$ ;*
- (3) *and splittings  $\mathcal{E}(x) = \bigoplus_{i=1}^p E_{\lambda_i}(x)$  into families of mutually transverse,  $\mu$ -measurable subbundles  $E_{\lambda_i}(x) \subset \mathcal{E}(x)$  defined for  $x \in \Lambda_0$*

such that

- (a)  $\mathcal{A}(s, x)E_{\lambda_i}(x) = E_{\lambda_i}(s \cdot x)$  and
- (b)  $\lim_{|s| \rightarrow \infty} \frac{\log \|\mathcal{A}(s, x)(v)\| - \lambda_i(s)}{|s|} = 0$

for all  $x \in \Lambda_0$  and all  $v \in E_{\lambda_i}(p) \setminus \{0\}$ .

Note that (b) implies for  $v \in E_{\lambda_i}(x)$  the weaker result that for  $s \in A$ ,

$$\lim_{k \rightarrow \pm\infty} \frac{1}{k} \log \|\mathcal{A}(s^k, x)(v)\| = \lambda_i(s).$$

Also note that for  $s \in A$ , and  $\mu$  an  $A$ -invariant,  $A$ -ergodic measure that

$$\lambda_{\text{top},s,\mu,\mathcal{A}} = \max_i \lambda_i(s). \tag{17}$$

If  $\mu$  is not  $A$ -ergodic, we have the following.

**Claim 3.12** *Let  $\mu$  be an  $A$ -invariant measure with  $\eta \in L^{d,1}(\mu)$  and  $\lambda_{\text{top},s,\mu,\mathcal{A}} > 0$  for some  $s \in A$ . Then there is an  $A$ -ergodic component  $\mu'$  of  $\mu$  with*

- (1)  $\eta \in L^{d,1}(\mu')$ ;
- (2) there is non-zero Lyapunov exponent  $\lambda_j \neq 0$  for the  $A$ -action on  $(X, \mu')$ .

We have the following which follows from the above definitions.

**Lemma 3.13** *Let  $\mu$  be an  $A$ -invariant probability measure on  $X$  with exponentially small mass in the cusps. Suppose that  $\mathcal{A}$  is a tempered cocycle. Then  $\eta \in L^q(\mu)$  for all  $q \geq 1$ . In particular,  $\eta \in L^{d,1}(\mu)$ .*

### 3.8 Applications to the suspension action

We summarize the previous discussion in the setting in which we will apply the above results in the sequel. Recall we work with in a fiber bundle with compact fiber

$$M \rightarrow M^\alpha = (G \times M) / \Gamma \xrightarrow{\pi} G / \Gamma$$

over non-compact base  $G / \Gamma$ . From the discussion in [6, Section 2.1], we may equip  $G \times M$  with a  $C^1$  metric that is

- (1)  $\Gamma$ -invariant;
- (2) the restriction to  $G$ -orbits coincides with the fixed right-invariant metric on  $G$ ;
- (3) there is a Siegel fundamental set  $D \subset G$  on which the restrictions to the fibers of the metrics are uniformly comparable.

The metric then descends to a  $C^1$  Riemannian metric on  $M^\alpha$ . We fix this metric for the remainder. It follows that the diameter of any fiber of  $M^\alpha$  is uniformly bounded. It then follows that if  $\mu$  is a measure on  $M^\alpha$  then the image  $\nu = \pi_*\mu$  in  $G / \Gamma$  has exponentially small mass in the cusps if and only if  $\mu$  does; moreover, a family  $\{\mu_\zeta\}$  of probability measures on  $M^\alpha$  has uniformly exponentially small mass in the cusps if and only if the family of projected measures  $\{\pi_*\mu_\zeta\}$  on  $G / \Gamma$  does. Note that by averaging the metric over the left-action of  $K$ , we may also assume that the metric is left- $K$ -invariant. This, in particular, implies the right-invariant metric on  $G$  in (2) above is left- $K$ -invariant.

For the remainder, the cocycle of interest will be the fiberwise derivative cocycle on the fiberwise tangent bundle,

$$\mathcal{A}(g, x): F \rightarrow F, \quad \mathcal{A}(g, x) = D_x g|_F.$$

Given  $g \in G$  and a  $g$ -invariant probability measure on  $M^\alpha$ , the average leading Lyapunov exponent for the fiberwise derivative cocycle for translation by  $g$  is written either as  $\lambda_{\text{top}, \mu, g}^F$  or as  $\lambda_{\text{top}, \mu, g, \mathcal{A}}$ .

The next observation we need is a variant of a fairly standard observation about cocycle over the suspension action.

**Lemma 3.14** *The fiberwise derivative cocycle  $D_x g|_F$  is tempered.*

*Proof* Write  $\pi : M^\alpha \rightarrow G/\Gamma$ . By the construction of the metric in the fibers of  $M^\alpha$  there is a  $C > 0$  with the following properties: given  $x \in M^\alpha$  and  $g \in G$ , writing  $\bar{x} = \pi(x) \in G/\Gamma$  we have

$$\|D_x g|_F\| \leq C^{\beta(g, \bar{x})+1}$$

and

$$m(D_x g|_F) \geq C^{-\beta(g, \bar{x})-1}.$$

The conclusion is then an immediate consequence of Lemma 2.1. □

We now assemble the consequences of the results in this section in the form we will use them below in a pair of lemmas. The first is just a special case of Corollary 3.7.

**Lemma 3.15** *Let  $s \in A$  and let  $\nu$  be an  $s$ -invariant measure on  $G/\Gamma$  with exponentially small mass in the cusps. Let  $\mu$  be an  $s$ -invariant measure on  $M^\alpha$  projecting to  $\nu$ . Then the average leading Lyapunov exponent for the fiberwise derivative cocycle,  $\lambda_{\text{top}, \mu, s}^F$ , is finite.*

The second lemma summarizes the above abstract results in the setting of  $G$  acting on  $M^\alpha$ .

**Lemma 3.16** *Let  $s \in A$  and let  $\nu$  be an  $s$ -invariant measure on  $G/\Gamma$  with exponentially small mass in the cusps. Let  $\mu$  be an  $s$ -invariant measure on  $M^\alpha$  projecting to  $\nu$ .*

- (1) *For any amenable subgroup  $H \subset C_G(s)$ , if  $\nu$  is  $H$ -invariant then*
  - (a) *for any Følner sequence of precompact sets  $F_n$  in  $H$ , the family  $\{F_n * \mu\}$  has uniformly exponentially small mass in the cusps; and*
  - (b) *for any subsequential limit  $\mu'$  of  $\{F_n * \mu\}$  we have*

$$\lambda_{\text{top}, s, \mu}^F \leq \lambda_{\text{top}, s, \mu'}^F.$$

- (2) *For any one-parameter unipotent subgroup  $U$  centralized by  $s$* 
  - (a) *the family  $\{U^T * \mu\}$  has uniformly exponentially small mass in the cusps; and*
  - (b) *for any accumulation point  $\mu'$  of  $\{U^T * \mu\}$  as  $T \rightarrow \infty$  we have*

$$\lambda_{\text{top}, s, \mu}^F \leq \lambda_{\text{top}, s, \mu'}^F.$$



*Proof* Part (a) of the first conclusion is immediate since  $H$ -invariance of  $\nu$  implies  $\nu = \pi_*(F_n * \mu)$  for all  $n$ ; part (b) then follows from Lemma 3.9. The second conclusion follows from Proposition 3.2 and Lemma 3.9.  $\square$

We remark that we will also use Lemma 3.10 in the proof of the main theorem, but we do not reformulate a special case of it here since the reformulation adds little clarity.

### 4 Subexponential growth of derivatives for unipotent elements

In this section we show that the restriction of the action  $\alpha$  to certain unipotent elements in each copy  $\Lambda_{i,j} \cong SL(2, \mathbb{Z})$  have uniform subexponential growth of derivatives with respect to a right-invariant distance on  $SL(2, \mathbb{R})$ . Note that each  $SL(2, \mathbb{R})$  is geodesically embedded whence the  $SL(2, \mathbb{R})$  distance is the same as the  $SL(m, \mathbb{R})$  distance. By [17, 18], the  $SL(m, \mathbb{R})$  distance is quasi-isometric to the word-length in  $SL(m, \mathbb{Z})$ . Recall that  $d(\cdot, \cdot)$  denotes a right-invariant distance on  $SL(m, \mathbb{R})$  and that  $\text{Id}$  is the identity in  $SL(m, \mathbb{R})$ .

For  $1 \leq i < j \neq n$ , let  $\Lambda_{i,j} \cong SL(2, \mathbb{Z})$  be the copy of  $SL(2, \mathbb{Z})$  in  $SL(m, \mathbb{Z})$  corresponding to the elements in  $SL(m, \mathbb{Z})$  which acts only on the lattice  $\mathbb{Z}^2 < \mathbb{Z}^m$  generated by  $\{e_i, e_j\}$ . Note that as all  $\Lambda_{i,j}$  are conjugate under the Weyl group, it suffices to work with one of them.

Define the unipotent element  $u := \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$  viewed as an element of  $\Lambda_{i,j}$ .

Note that any upper or lower triangular unipotent element of  $\Lambda_{i,j}$  is conjugate to a power of  $u$  under the Weyl group.

**Proposition 4.1** (Subexponential growth of derivatives for unipotent elements) *For any  $\Lambda_{i,j}$  and any  $\varepsilon > 0$ , there exists  $N_\varepsilon > 0$  such that for any  $n \geq N_\varepsilon$ :*

$$\|D(\alpha(u^n))\| \leq e^{\varepsilon d(u^n, \text{Id})}$$

To establish Proposition 4.1, we first show that generic elements in  $SL(2, \mathbb{Z})$  have uniform subexponential growth of derivatives. This first part requires reusing most of the key arguments from [4] in a slightly modified form. We encourage the reader to read that paper first.

#### 4.1 Slow growth for “most” elements in $SL(2, \mathbb{Z})$

For  $\varepsilon > 0, k > 0$ , and  $x \in SL(2, \mathbb{R})$ , we make the following definitions:

- (1) For  $S \subset SL(2, \mathbb{R})$  let  $|S|$  denote the Haar-volume of  $S$ .
- (2) Let  $K = SO(2) \subset SL(2, \mathbb{R})$ . For  $S \subset K$  let  $|S|$  denote the Haar-volume of  $S$ .

- (3) Let  $B_k(x)$  denote the ball of radius  $k$  centered at  $x$  in  $SL(2, \mathbb{R})$ .
- (4) Let  $T_k := B_k(\text{Id}) \cap SL(2, \mathbb{Z})$ . Given  $S \subset SL(2, \mathbb{Z})$  write  $|S|$  for the cardinality of  $S$ .
- (5) Define the set of  $\varepsilon$ -bad elements to be

$$M_{\varepsilon,k} := \{\gamma \in T_k \text{ such that } \|D(\alpha(\gamma))\| \geq e^{\varepsilon k}\}.$$

- (6) Define the set of  $\varepsilon$ -good elements to be

$$G_{\varepsilon,k} := T_k \setminus M_{\varepsilon,k}.$$

To establish Proposition 4.1, we first show that the set  $G_{\varepsilon,k}$  contains a positive proportion of  $T_k$  when  $k$  is large enough.

**Proposition 4.2** *For any  $\delta > 0$ , the set  $G_{\varepsilon,k}$  has at least  $(1 - \delta)|T_k|$  elements for every sufficiently large  $k$ .*

We have the following well-known fact. See for instance [9, Section 2].

**Lemma 4.3** *There exist positive constants  $c, C$  such that for any  $k \geq 0$ :*

$$c|B_k| \leq |T_k| \leq C|B_k|.$$

For an element  $x \in SL(2, \mathbb{R})$ , let  $\bar{x}$  denote the projection in  $SL(2, \mathbb{R})/SL(2, \mathbb{Z})$ . Define

$$\|D_{\bar{x}}g\|_{\text{Fiber}} = \sup\{\|D_yg|_F\| : y \in M^\alpha, \pi(y) = \bar{x}\}.$$

Let

$$G'_{\varepsilon,k}(x) := \{g \in B_k(x) \text{ such that } \|D_{\bar{x}}g\|_{\text{Fiber}} \leq e^{\varepsilon d(g, \text{Id})}\}.$$

**Lemma 4.4** *For almost every  $x \in SL(2, \mathbb{R})$  and any  $\delta > 0$  we have*

$$|G'_{\varepsilon,k}(x)| > (1 - \delta)|B_k|$$

for all  $k$  sufficiently large.

*Proof* Let  $a^t \in SL(2, \mathbb{R})$  be the matrix

$$a^t := \begin{bmatrix} e^t & 0 \\ 0 & e^{-t} \end{bmatrix}.$$

Recall that the action of the one-parameter diagonal subgroup  $\{a^t\}$  on  $SL(2, \mathbb{R})/SL(2, \mathbb{Z})$  is ergodic with respect to Haar measure.

Let  $\mathcal{M}$  denote the set of Borel probability measures on  $SL(2, \mathbb{R})/SL(2, \mathbb{Z})$  equipped with the standard topology (dual to bounded continuous functions). The topology on  $\mathcal{M}$  is metrizable (see [2, Theorem 6.8]); fix a metric on  $\rho_{\mathcal{M}}$  on  $\mathcal{M}$ .

Consider the function  $\psi : SL(2, \mathbb{R})/SL(2, \mathbb{Z}) \rightarrow \mathbb{R}$  given by  $\psi(x) := e^{\eta d(x, x_0)}$  where  $x_0 = SL(2, \mathbb{Z})$  is the identity coset and  $\eta > 0$  is chosen sufficiently small so that  $\psi$  is  $L^1$  with respect to the Haar measure. By the pointwise ergodic theorem, for almost every  $x \in SL(2, \mathbb{R})$  and almost every  $k_1 \in SO(2)$  we have

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \psi(a^t k_1 \bar{x}) dt = \int_{SL(2, \mathbb{R})/SL(2, \mathbb{Z})} \psi d\text{Haar} < \infty. \tag{18}$$

Similarly, for almost every  $x \in SL(2, \mathbb{R})$  and almost every  $k_1 \in SO(2)$  we have

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \delta_{a^t k_1 \bar{x}} dt = \text{Haar}. \tag{19}$$

Let  $S \subset SL(2, \mathbb{R})$  be the set of  $x \in SL(2, \mathbb{R})$  such that (18) and (19) hold for almost every  $k_1 \in SO(2)$ . The set  $S$  is  $SL(2, \mathbb{Z})$ -invariant and co-null. We show any  $x \in S$  satisfies the conclusion of the lemma.

For fixed  $x \in S$  and fixed  $\delta > 0$ , there exist  $T_\delta = T_\delta(x)$ , a sequence  $T_j = T_j(x)$  for  $j \in \mathbb{N}$ , and a set  $K_\delta = K_\delta(x) \subset SO(2)$  such that  $|K_\delta| \geq (1 - \delta/2)|SO(2)|$  with the property that for any  $k_1 \in K_\delta$  and any  $T \geq T_\delta$  we have

$$\frac{1}{T} \int_0^T \psi(a^t k_1 \bar{x}) dt < 2 \int_{SL(2, \mathbb{R})/SL(2, \mathbb{Z})} \psi d\text{Haar} \tag{20}$$

and for each  $1 \leq j$

$$\rho_{\mathcal{M}} \left( \frac{1}{T} \int_0^T \delta_{a^t k_1 \bar{x}} dt, \text{Haar} \right) < \frac{1}{j}. \tag{21}$$

for all  $T \geq T_j$ . To finish the proof of the lemma, define the set

$$G''_k(x) := \{k_1 a^t k_2 \text{ where } k_1 \in SO(2), k_2 \in K_\delta(x) \text{ and } (\delta/2)k < t < k\}.$$

For  $k$  large enough, we have that  $|G''_k(x)| \geq (1 - \delta)|B_k|$ . We claim that

$$G''_k(x) \subset G'_{\varepsilon, k}(x) \tag{22}$$

for  $k$  sufficiently large. For the sake of contradiction, suppose (22) fails. Using that the norm on  $F$  is chosen to be  $K$ -invariant, there exists  $x_n \in \text{SL}(2, \mathbb{R})$  with each  $x_n$  in the  $K_\delta(x)$ -orbit of  $x$  such that  $\|D_{x_n}(a^{t_n})\|_{\text{Fiber}} \geq e^{\varepsilon t_n}$  for some sequence  $t_n \rightarrow \infty$ . Moreover, the corresponding empirical measures

$$\eta_n := \frac{1}{t_n} \int_0^{t_n} \delta_{a^t \bar{x}_n} dt$$

have uniformly exponentially small mass in the cusps by equation (20).

By Lemma 3.10 and (21), a subsequence of the measures  $\eta_n$  converge to an  $a^t$ -invariant measure  $\mu_0$  on  $M^\alpha$  whose projection to  $\text{SL}(m, \mathbb{R})/\text{SL}(m, \mathbb{Z})$  is Haar measure on the embedded modular surface  $\text{SL}(2, \mathbb{R})/\text{SL}(2, \mathbb{Z})$  and has positive fiberwise Lyapunov exponent for the action of  $a^1$ . Since  $a^t$  is ergodic on  $\text{SL}(2, \mathbb{R})/\text{SL}(2, \mathbb{Z})$ , we can assume  $\mu_0$  is ergodic by taking an ergodic component without changing any other properties.

We average as in [4] to improve  $\mu_0$  to a measure whose projection is the Haar measure on  $\text{SL}(m, \mathbb{R})/\text{SL}(m, \mathbb{Z})$ . Difficulties related to escape of mass are handled by the preliminaries in Sect. 3.

As above, we note that there is a canonical copy of  $H_2 = \text{SL}(m - 2, \mathbb{R})$  in  $\text{SL}(m, \mathbb{R})$  commuting with our chosen  $H_1 = \text{SL}(2, \mathbb{R})$ . Recall  $A$  is the Cartan subgroup of  $\text{SL}(m, \mathbb{R})$  of positive diagonal matrices. The subgroup  $A$  contains the one-parameter group  $\{a^t\}$  and a Cartan subgroup of  $H_2$ . Let

- $A_1 = A \cap H_1 = \{a^t\}$ ,
- $A_2 = A \cap H_2$ , and
- $A' = A \cap (H_1 \times H_2)$ .

Note that  $A' < A$  has codimension one. Our chosen modular surface  $\text{SL}(2, \mathbb{R})/\text{SL}(2, \mathbb{Z}) \subset \text{SL}(m, \mathbb{R})/\text{SL}(m, \mathbb{Z})$  is such that

$$\begin{aligned} \text{SL}(2, \mathbb{R})/\text{SL}(2, \mathbb{Z}) &\subset \text{SL}(2, \mathbb{R})/\text{SL}(2, \mathbb{Z}) \times \text{SL}(m - 2, \mathbb{R})/\text{SL}(m - 2, \mathbb{Z}) \\ &\subset \text{SL}(m, \mathbb{R})/\text{SL}(m, \mathbb{Z}). \end{aligned}$$

Define an  $A'$ -ergodic,  $A'$ -invariant measure  $\mu_1$  on  $M^\alpha$  that projects to Haar measure on  $\text{SL}(2, \mathbb{R})/\text{SL}(2, \mathbb{Z}) \times \text{SL}(m - 2, \mathbb{R})/\text{SL}(m - 2, \mathbb{Z})$  as follows: Let  $M_{2,m-2}^\alpha$  denote the restriction of the fiber-bundle  $M^\alpha$  to  $\text{SL}(2, \mathbb{R})/\text{SL}(2, \mathbb{Z}) \times \text{SL}(m - 2, \mathbb{R})/\text{SL}(m - 2, \mathbb{Z})$ . Pick point  $y$  in  $\text{SL}(m - 2, \mathbb{R})/\text{SL}(m - 2, \mathbb{Z})$  that equidistributes to the Haar measure on  $\text{SL}(m - 2, \mathbb{R})/\text{SL}(m - 2, \mathbb{Z})$  under a Følner sequence in  $A_2$ . Consider  $\mu_0$  as a measure on the restriction of  $M^\alpha$  to  $\text{SL}(2, \mathbb{R})/\text{SL}(2, \mathbb{Z}) \times \{y\}$ . Now average  $\mu_0$  over a Følner sequence in  $A_2$  and take a limit  $\hat{\mu}_1$ . Note that  $\hat{\mu}_1$  has positive fiberwise Lyapunov exponent  $\lambda_{\text{top}, a^1, \mu_1}^F > 0$ . This can be seen by mimicking the proof of Lemma 3.9. Let  $\mu_1$  be an  $A'$  ergodic component of  $\hat{\mu}_1$ , then the measure  $\mu_1$  has the desired

properties and is supported on the subset of  $M^\alpha$  defined by restricting the bundle to  $SL(2, \mathbb{R})/SL(2, \mathbb{Z}) \times SL(m - 2, \mathbb{R})/SL(m - 2, \mathbb{Z})$ .

We consider the  $A'$ -action on  $(M^\alpha, \mu_1)$  and the fiberwise derivative cocycle  $\mathcal{A}(g, y) = D_y g|_F$ . By (17), there is a non-zero Lyapunov exponent  $\lambda_{\mu_1, A'}^F: A' \rightarrow \mathbb{R}$  for this action. We apply the averaging procedure in Proposition 3.5 to this measure. Take  $\beta'$  to be either  $\alpha_2$  or  $\delta$  so that  $\beta': A' \rightarrow \mathbb{R}$  is not proportional to  $\lambda_{\mu_1, A'}^F$ . Choose  $a_0 \in A'$  such that  $a_0 \in \ker(\beta')$  and  $\lambda_{\mu_1, A'}^F(a_0) > 0$ . Let  $U = U^{\beta'}$  and let  $\mu_2$  be any subsequential limit of  $U^T * \mu_1$  as  $T \rightarrow \infty$ . Then  $\mu_2$  is  $a_0$ -invariant, and has positive fiberwise Lyapunov exponent  $\lambda_{\text{top}, a_0, \mu_2}^F > 0$ . Moreover,  $\pi_* \mu_2$  is  $H$ -invariant. By Lemma 3.16 and Proposition 3.5,  $\mu_2$  has exponentially small mass in the cusps. We may also assume  $\mu_2$  is ergodic by passing to an ergodic component and by Claim 3.12 assume  $\mu_2$  has a non-zero fiberwise Lyapunov exponent  $\lambda_{\mu_2, A'}^F$  for the  $A'$ -action.

We now average  $\mu_2$  over  $A'$  to obtain  $\mu_3$ . Then  $\mu_3$  has a non-zero fiberwise Lyapunov exponent  $\lambda_{\mu_3, A'}^F$  and has exponentially small mass in the cusps by Lemma 3.16(1). Since  $\pi_* \mu_2$  was  $A'$ -invariant, we have  $\pi_* \mu_2 = \pi_* \mu_3$ . Once again, we may pass to an  $A'$ -ergodic component of  $\mu_3$  that retains the desired properties.

Take  $\hat{\beta}$  to be either  $-\alpha_2$  or  $-\delta$  so that  $\hat{\beta}$  is not proportional to  $\lambda_{\mu_3, A'}^F$  on  $A'$ . Select  $a_1$  with  $\lambda_{\mu_3, A'}^F(a_1) > 0$  and  $\hat{\beta}(a_1) = 0$ . By Proposition 3.5 and Lemma 3.16, we obtain a new measure  $\mu_4$  with  $\pi_* \mu_4$  the Haar measure on  $SL(m, \mathbb{R})/SL(m, \mathbb{Z})$ . We have  $\lambda_{\text{top}, a_1, \mu_4}^F > 0$ . Finally, average  $\mu_4$  over all of  $A$  to obtain  $\mu_5$ . Since  $\pi_* \mu_4$  is the Haar measure and thus  $A$ -invariant, we have that  $\pi_* \mu_4 = \pi_* \mu_5$ . By Lemma 3.16,  $\mu_5$  has a non-zero fiberwise Lyapunov exponent  $\lambda_{\mu_5, A}^F$  for the action of  $A$ . Replace  $\mu_5$  by an ergodic component with positive fiberwise Lyapunov exponent.

Exactly as in [4, Section 5.5], we apply [6, Proposition 5.1] and conclude that  $\mu_5$  is a  $G$ -invariant measure on  $M^\alpha$ . We then obtain a contradiction with Zimmer’s cocycle superrigidity theorem. To conclude that  $\mu_5$  is a  $G$ -invariant, note that [6, Proposition 5.1] holds for actions induced from actions of any lattice in  $SL(m, \mathbb{R})$  and shows that  $\mu_5$  is invariant under root subgroups corresponding to *non-resonant* roots. Dimension counting exactly as in [4, Section 5.5] shows that the non-resonant roots of  $SL(m, \mathbb{R})$  generate all of  $G$  if the dimension of  $M$  is at most  $m - 2$  or if the dimension of  $M$  is  $m - 1$  and the action preserves a volume. □

We derive Proposition 4.2 from Lemma 4.4.

*Proof of Proposition 4.2* Fix  $0 < c < 1$  sufficiently small so that if  $d(\text{Id}, g) < c$  then  $\|D_\Gamma g\|_{\text{Fiber}} \leq e^{\varepsilon/4}$ . Fix a point  $x \in SL(2, \mathbb{R})$  as in Lemma 4.4 with  $d(\text{Id}, x) < c$ . Observe that if  $k \geq 1$  and  $g \in G'_{\varepsilon/4, k}(x)$ , then

$gx \in G'_{\varepsilon/2, k+c}(\text{Id})$ . In particular, for any  $\delta > 0$  we have for all  $k$  sufficiently large that

$$|B_{k+c} \setminus G'_{\varepsilon/2, k+c}(\text{Id})| < \delta \hat{C} |B_k| \tag{23}$$

where  $\hat{C}$  is a constant depending on  $c$ . Take  $U$  to be the ball of radius  $c$  centered at the identity coset in  $\text{SL}(2, \mathbb{R})/\text{SL}(2, \mathbb{Z})$  and consider lifts of  $U$  to  $\text{SL}(2, \mathbb{R})$  intersecting the ball  $B_k$ . If a lift of  $U$  intersects  $G'_{\varepsilon/2, k+c}(\text{Id})$ , then the corresponding element of the deck group  $\text{SL}(2, \mathbb{Z})$  belongs to  $G'_{3\varepsilon/4, k}(\text{Id})$ .

Let  $\tilde{U}$  be the set of lifts of  $U$ . From Lemma 4.3 and (23), it follows that ratio of the measure of  $\tilde{U} \cap B_k \cap G'_{\varepsilon/2, k}(\text{Id})$  to the measure of  $\tilde{U} \cap B_k$  goes to one as  $k \rightarrow \infty$ . Finally, since the norms on the fiber of  $M^\alpha$  above the identity coset and the original norm on  $M$  are uniformly comparable, the result follows.  $\square$

*Remark 4.5* Using large deviations, one can make  $\delta$  to be decreasing with  $k$ , roughly as  $\delta_k = e^{-k^{1/1000}}$ . See [1, 10]. This is not necessary for our argument.

### 4.2 Subexponential growth of derivatives for unipotent elements in $\text{SL}(2, \mathbb{Z})$

We work here with a specific copy of the group  $\text{SL}(2, \mathbb{R}) \times \mathbb{R}^2$  embedded in  $\text{SL}(m, \mathbb{R})$  and its intersection with the lattice  $\Gamma$ ; the copy of  $\text{SL}(2, \mathbb{R}) \times \mathbb{R}^2$  corresponds to the elements of  $\text{SL}_m(\mathbb{R})$  which differ from the identity matrix only in the first two rows and first three columns. Any unipotent element of any  $\Lambda_{i,j} \subset \Gamma$  considered in the statement of Proposition 4.1 is conjugate by an element of the Weyl group to a power of the elementary matrix  $E_{1,3}$ . Thus, after conjugation, any such element is contained in the distinguished copy of  $\text{SL}(2, \mathbb{Z}) \times \mathbb{Z}^2$  generated by  $\text{SL}(2, \mathbb{Z}) = \Lambda_{1,2}$  and the normal subgroup  $\mathbb{Z}^2$  generated by  $E_{1,3}$  and  $E_{2,3}$ .

For the remainder of this subsection, we work with this fixed group. Identify  $H_{1,2}$  with  $\text{SL}(2, \mathbb{R})$ . Let  $U_{1,2} := \{u_{a,b}\}$  denote the abelian subgroup of  $\text{SL}(m, \mathbb{R})$  consisting of unipotent elements of the form

$$u_{a,b} := \begin{pmatrix} 1 & 0 & a & & & \\ 0 & 1 & b & & & \\ 0 & 0 & 1 & & & \\ & & & \ddots & & \\ & & & & & 1 \end{pmatrix}$$

Clearly,  $U_{1,2}$  is normalized by  $H_{1,2}$  and  $H_{1,2} \times U_{1,2} \cong SL(2, \mathbb{R}) \times \mathbb{R}^2$ . We have an embedding

$$SL(2, \mathbb{R}) \times \mathbb{R}^2 / SL(2, \mathbb{Z}) \times \mathbb{Z}^2 \rightarrow SL(m, \mathbb{R}) / SL(m, \mathbb{Z})$$

where  $\mathbb{Z}^2$  is identified with the subgroup generated by the unipotent elements  $u_{1,0}$  and  $u_{0,1}$ . Note that  $SL(2, \mathbb{R}) \times \mathbb{R}^2 / SL(2, \mathbb{Z}) \times \mathbb{Z}^2$  is a torus bundle over the unit-tangent bundle of the modular surface.

Equip  $\mathbb{Z}^2$  with the  $L_\infty$  norm with respect to the generating set  $\{u_{1,0}, u_{0,1}\}$  and let  $B_n(\mathbb{Z}^2)$  denote the closed ball of radius  $n$  in  $\mathbb{Z}^2$  centered at 0 with respect to this norm. Given  $S \subset \mathbb{Z}^2$  let  $|S|$  denote the cardinality of the set  $S$ .

Define the set of “ $\varepsilon$ -good unipotent elements” of  $\mathbb{Z}^2 \subset \Gamma$ , denoted by  $GU_{\varepsilon,n}$ , to be the following subset of  $\mathbb{Z}^2$ :

$$GU_{\varepsilon,n} := \left\{ u_{a,b} \in B_n(\mathbb{Z}^2) \text{ such that } \|D(\alpha(u_{a,b}^{\pm 1}))\| \leq e^{\varepsilon \log(n)} \right\}. \tag{24}$$

The main results of this subsection is the following.

**Proposition 4.6** *For any  $\varepsilon > 0$ , there exists  $N_\varepsilon > 0$  such that if  $n \geq N_\varepsilon$ , then  $GU_{\varepsilon,n} = B_n(\mathbb{Z}^2)$*

Proposition 4.1 follows from Proposition 4.6 using that any subgroup  $\langle u^n \rangle$  in Proposition 4.1 is conjugate to a subgroup of the group  $\mathbb{Z}^2$  and the fact that  $d(u^n, \text{Id}) = O(\log(n))$  from (6). The proof of Proposition 4.6 consists of conjugating elements of  $U_{1,2}$  by elements of  $G_{\varepsilon,n}$  in order to obtain a subset of  $G_{\varepsilon,n}$  that contains a positive density of elements of  $B_n(\mathbb{Z}^2)$ . Then, using the fact that  $\mathbb{Z}^2$  is abelian, we promote such a subset to all of  $B_n(\mathbb{Z}^2)$  by taking sufficiently large sumsets in Proposition 4.9.

**Lemma 4.7** *There exists  $\delta' > 0$  with the following properties: for any  $\varepsilon > 0$  there is an  $N'_\varepsilon > 0$  such that for any  $n \geq N'_\varepsilon$  we have*

$$|GU_{\varepsilon,n}| \geq \delta' |B_n(\mathbb{Z}^2)|.$$

*Proof* Recall that  $T_k$  denotes the intersection of the ball of radius  $k$  in  $SL(2, \mathbb{R}) \simeq H_{1,2}$  with  $SL(2, \mathbb{Z}) = \Lambda_{1,2}$  and  $|T_k|$  denotes the cardinality of  $T_k$ . As  $|T_k|$  grows exponentially in  $k$ , we may take  $s$  fixed so that  $|T_{k-s}| < \frac{1}{2}|T_k|$  for all  $k$  sufficiently large. Given  $\varepsilon' > 0$ , define the subset  $S_k \subset SL(2, \mathbb{Z})$  to be

$$S_k := G_{\varepsilon',k} \cap G_{\varepsilon',k}^{-1} \cap (T_k \setminus T_{k-s}).$$

From Proposition 4.2, we may assume that

$$|S_k| \geq \frac{1}{2}|T_k|.$$

From (3), there exists  $C_1 > 0$  such that if  $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$  belongs to  $S_k$  then either

$$\|(a, b)\|_\infty \geq C_1 e^{\frac{1}{2}(k-s)} \text{ or } \|(c, d)\|_\infty \geq C_1 e^{\frac{1}{2}(k-s)}.$$

Without loss of generality, we assume that at least half of the elements in  $S_k$  satisfy  $\|(a, b)\|_\infty \geq C_1 e^{\frac{1}{2}(k-s)}$ .

Consider the map  $P : S_k \rightarrow \mathbb{Z}^2$  that assigns  $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$  to  $(a, b)$ . By (3), there is  $C_2 > 1$  such that the image  $P(S_k)$  of  $S_k$  lies in the norm-ball  $B_{C_2 e^{\frac{k}{2}}}(\mathbb{Z}^2)$  for all  $k$ .

Let  $k(n) = 2 \log(n) - \log C_2$ . Then  $P(S_{k(n)}) \subset B_n(\mathbb{Z}^2)$ . If  $n$  is sufficiently large and  $A = \begin{bmatrix} a & c \\ b & d \end{bmatrix} \in S_{k(n)}$  then we have  $u_{a,b} \in GU_{(5\varepsilon', n)}$ ; indeed

$$\alpha(u_{a,b}) = \alpha(A) \circ \alpha(u_{1,0}) \circ \alpha(A^{-1})$$

whence

$$\|D\alpha(u_{a,b})\| \leq \|D\alpha(u_{1,0})\| e^{2\varepsilon'k(n)}.$$

We have  $|B_n(\mathbb{Z}^2)| \leq D_1 n^2$  for some  $D_1 \geq 1$ . Also, from (4) and Lemma 4.3 we have  $|S_{k(n)}| \geq \frac{1}{2}|T_{k(n)}| \geq \frac{1}{2}e^{k(n)} = \frac{1}{D_2}n^2$  for some  $D_2 \geq 1$ .

To to complete the proof, we show that the preimage  $P^{-1}((a, b))$  in  $S_k$  of any  $(a, b) \in \mathbb{Z}^2$  satisfying  $\|(a, b)\|_\infty \geq C_1 e^{\frac{1}{2}(k-s)}$  has uniformly bounded cardinality depending only on  $s$ . Observe that if  $A, A' \in \text{SL}(2, \mathbb{Z})$  satisfy  $P(A) = P(A')$ , then  $A' = AU$ , where  $U = \begin{bmatrix} 1 & m \\ 0 & 1 \end{bmatrix}$  for some  $m \in \mathbb{Z}$  and we have

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}, \text{ and } A' = \begin{bmatrix} a & am + c \\ b & bm + d \end{bmatrix}.$$

If  $A'$  belongs to  $T_k$  then  $\|(am + c, bm + d)\|_\infty \leq C_2 e^{\frac{k}{2}}$  and if  $A$  belongs to  $T_k$  then  $\|(c, d)\|_\infty \leq C_2 e^{\frac{k}{2}}$ . We thus have that  $|am| \leq 2C_2 e^{\frac{k}{2}}$  and  $|bm| \leq 2C_2 e^{\frac{k}{2}}$ . As we assume that

$$\|(a, b)\|_\infty \geq C_1 e^{\frac{k-s}{2}}$$

we have that  $|m| \leq 2\frac{C_2}{C_1}e^{\frac{s}{2}}$ . Thus, the preimage  $P^{-1}((a, b))$  has at most  $4\frac{C_2}{C_1}e^{\frac{s}{2}} + 1$  elements in  $S_k$ .



With  $\varepsilon' = \frac{1}{5}\varepsilon$ , having taken  $n$  sufficiently large, we thus have

$$\frac{|GU_{\varepsilon,n}|}{|B_n(\mathbb{Z}^2)|} \geq \frac{1}{D_1 n^2} \frac{\frac{1}{2}|S_{k(n)}|}{4 \frac{C_2}{C_1} e^{s/2} + 1} \geq \frac{1}{2} \frac{\frac{1}{D_2} n^2}{4 \frac{C_2}{C_1} e^{s/2} + 1} \frac{1}{D_1 n^2} =: \delta'$$

which completes the proof. □

To complete the proof of Proposition 4.6, we show that any element in  $B_n(\mathbb{Z}^2)$  can be written as a product of a bounded number of elements in  $GU_{\varepsilon,n}$  independent of  $\varepsilon$ . This follows from the structure of sumsets of abelian groups.

From the chain rule and submultiplicativity of norms, we have the following.

**Claim 4.8** *For any positive integers  $n, m$  and  $\varepsilon_1, \varepsilon_2 > 0$ , if  $u_{a,b} \in GU_{\varepsilon_1,n}$  and  $u_{c,d} \in GU_{\varepsilon_2,m}$  then the product  $u_{a,b}u_{c,d} \in GU_{\max\{\varepsilon_1, \varepsilon_2\}, n+m}$*

For subsets  $A, B \subset \mathbb{Z}^2$  we denote by  $A + B$  the sumset of  $A, B$ .

**Claim 4.9** *For any  $0 < \delta < 1$ , there exists a positive integer  $k_\delta$  and a finite set  $F_\delta \subset \mathbb{Z}^2$  such that for any  $n$  and any symmetric set  $S_n \subset B_n(\mathbb{Z}^2)$  with  $|S_n| > \delta|B_n|$ , we have that*

$$B_n \subset F_\delta + \underbrace{S_n + S_n + \dots + S_n}_{k_\delta \text{ times}}$$

*Proof* Fix  $M \in \mathbb{Z}_+$  with  $\frac{1}{M} < \delta$ . Take  $N_\delta := (M + 1)!, k_\delta = 4N_\delta$ , and  $F_\delta := B_{N_\delta}(\mathbb{Z}^2)$ . Consider a symmetric set  $S_n \subset B_n(\mathbb{Z}^2)$  with  $|S_n| > \delta|B_n(\mathbb{Z}^2)|$ .

If  $n \leq N_\delta$  then  $B_n(\mathbb{Z}^2) \subset F_\delta$  and we are done. Thus, consider  $n \geq N_\delta$ . To complete the proof the claim, we argue that the set

$$\sum_{k_\delta} S_n := \underbrace{S_n + S_n + \dots + S_n}_{k_\delta \text{ times}}$$

contains the intersection of the sublattice  $N_\delta \mathbb{Z}^2$  with  $B_n(\mathbb{Z}^2)$ . Adding  $F_\delta$  to the sumset then implies the claim. Consider any non-zero vector  $\tilde{v} \in N_\delta \mathbb{Z}^2 \cap B_n(\mathbb{Z}^2)$  of the form  $(\tilde{\ell}, 0)$  for some  $\tilde{\ell} \in [-n, n] \cap N_\delta \mathbb{Z}$ . Then  $\tilde{v} = N_\delta v$  where  $v = (\ell, 0)$  is such that  $0 < |\ell| \leq \lfloor n N_\delta^{-1} \rfloor$ .

Consider the equivalence relation in  $B_n(\mathbb{Z}^2)$  defined by declaring that two elements  $x, y \in R(n)$  are equivalent if  $x - y$  is an integer multiple of  $v$ . Each equivalence class is of the form

$$C_x = \{ \dots, x - v, x, x + v, x + 2v, \dots \}.$$

As  $|S_n| \geq \frac{1}{M}|B_n(\mathbb{Z}^2)|$ , there exists one equivalence class  $C_x$  such that  $|C_x \cap S_n| \geq \frac{1}{M}|C_x|$ . Since  $0 < |\ell| \leq \lfloor n N_\delta^{-1} \rfloor$ , each equivalence class contains at

least  $M + 1$  elements and hence  $C_x \cap S_n$  contains at least two elements  $a, b$  with  $b = a + iv$  for  $|i| \leq M$ . In particular, since  $a - b = iv$ , we have  $iv \in S_n + S_n$ . As  $i$  divides  $N_\delta$ , we have that  $\tilde{v} = N_\delta v \in \sum_{2N_\delta} S_n$ .

Similarly, for  $n \geq N_\delta$  and any  $\tilde{u} \in N_\delta \mathbb{Z}^2 \cap B_n(\mathbb{Z}^2)$  of the form  $(0, \tilde{v})$  we have  $\tilde{u} \in \sum_{2N_\delta} S_n$ . Then

$$\tilde{u} + \tilde{v} \in \sum_{4N_\delta} S_n$$

completing the proof. □

*Proof of Proposition 4.6* Given  $\varepsilon' > 0$ , let  $\delta'$  and  $N'_{\varepsilon'}$  be given by Lemma 4.7. Let  $S_n := GU_{\varepsilon',n}$  be as in (24) and take  $k'_{\delta'}$  and  $F_{\delta'}$  as in Lemma 4.9. Note that  $GU_{\varepsilon',n}$  is symmetric by definition. Take  $N \geq N'_{\varepsilon'}$ , such that  $F_{\delta'} \in GU_{\varepsilon',n}$  whenever  $n \geq N$ . For  $n \geq N$  and any  $u_{a,b} \in B_n(\mathbb{Z}^2)$  we have that  $u_{a,b} \in F_{\delta'} + S_n + S_n + \dots + S_n$  ( $k_{\delta'}$  times) by Proposition 4.9. Proposition 4.8 then implies that  $u_{a,b} \in GU_{\varepsilon',(k_{\delta'}+1)n}$  so  $\|D(u_{a,b})^{\pm 1}\| \leq e^{\varepsilon' \log((k_{\delta'}+1)n)}$ . With  $\varepsilon' = \varepsilon/2$ , take  $N_\varepsilon \geq \max\{N, (k_{\delta'} + 1)\}$ . Then for all  $n \geq N_\varepsilon$  we have

$$\varepsilon' \log((k_{\delta'} + 1)n) \leq \varepsilon \log(n)$$

whence

$$\|D(u_{a,b})^{\pm 1}\| \leq e^{\varepsilon \log(n)}$$

and for  $u_{a,b} \in B_n(\mathbb{Z}^2)$  with  $n \geq N_\varepsilon$ . □

## 5 Proof of Theorem B

### 5.1 Reduction to the restriction of an action by $\Lambda_{i,j}$

We recall the work of Lubotzky, Mozes, and Raghunathan, namely [17] and [18], which establishes quasi-isometry between the word and Riemannian metrics on lattices in higher-rank semisimple Lie groups. In the special case of  $\Gamma = \text{SL}(m, \mathbb{Z})$  for  $m \geq 3$ , in [17, Corollary 3] it is shown that any element  $\gamma$  of  $\text{SL}(m, \mathbb{Z})$  is written as a product of at most  $m^2$  elements  $\gamma_i$ . Moreover each  $\gamma_i$  is contained some  $\Lambda_{i,j} \simeq \text{SL}(2, \mathbb{Z})$  and the word-length of each  $\gamma_i$  is proportional to the word-length of  $\gamma$ .

Thus, to establish that an action  $\alpha : \Gamma \rightarrow \text{Diff}^1(M)$  has uniform subexponential growth of derivatives in Theorem B, it is sufficient to show that the restriction  $\alpha|_{\Lambda_{i,j}} : \Gamma \rightarrow \text{Diff}^1(M)$  has uniform subexponential growth of derivatives for each  $1 \leq i \neq j \leq m$ . We emphasize that to measure

subexponential growth of derivatives, the word-length on  $\Lambda_{i,j}$  is measured as the word-length as embedded in  $SL(m, \mathbb{Z})$  (which is quasi-isometric to the Riemannian metric on  $SL(m, \mathbb{R})$ ) rather than the intrinsic word-length in  $\Lambda_{i,j} \simeq SL(2, \mathbb{Z})$  (which is not quasi-isometric to the Riemannian metric on  $SL(2, \mathbb{R})$ ).

As the Weyl group acts transitively on the set of all  $\Lambda_{i,j}$ , it is sufficient to consider a fixed  $\Lambda_{i,j}$ . Thus to deduce Theorem B, in the remainder of this section we establish the following, which is the main proposition of the paper.

**Proposition 5.1** *For any action  $\alpha : \Gamma \rightarrow \text{Diff}^1(M)$  as in Theorem B, the restricted action  $\alpha|_{\Lambda_{1,2}} : \Gamma \rightarrow \text{Diff}^1(M)$  has uniform subexponential growth of derivatives.*

### 5.2 Orbits with large fiber growth yet low depth in the cusp

To prove Proposition 5.1, as in Sect. 4.2 we consider a canonical embedding  $X = H_{1,2}/\Lambda_{1,2}$  of  $SL(2, \mathbb{R})/SL(2, \mathbb{Z})$  in  $SL(m, \mathbb{R})/SL(m, \mathbb{Z})$ . Write

$$a^t := \text{diag}(e^{t/2}, e^{-t/2}) \in SL(2, \mathbb{R})$$

for the geodesic flow on  $X$ . Let  $X_{\text{thick}}$  be a fixed compact  $SO(2)$ -invariant “thick part” of  $X$ ; that is, relative to the Dirichlet domain  $\mathcal{D}$  in (7), points in  $SO(2) \backslash X_{\text{thick}}$  corresponds to the points in  $SO(2) \backslash \mathcal{D}$  whose imaginary part is bounded above, say, by 17.

A geodesic curve in the modular surface of length  $t$  corresponds to the image of an orbit  $\zeta = \{a^s(x)\}_{0 \leq s \leq t}$  where  $x \in X$  and  $t \geq 0$ . Denote the length of such a curve by  $l(\zeta)$ . For an orbit  $\zeta = \{a^s(x)\}_{0 \leq s \leq t}$  of  $\{a^t\}$  in  $X$  we define

$$c(\zeta) := \log(\|D_x(a^t)\|_{\text{Fiber}}).$$

The following claim is straightforward from the compactness  $X_{\text{thick}}$  and the quasi-isometry between the word and Riemannian metrics on  $\Gamma$ .

**Claim 5.2** *For an action  $\alpha : SL(m, \mathbb{Z}) \rightarrow \text{Diff}^1(M)$ , the following statements are equivalent:*

- (1) *the restriction  $\alpha|_{\Lambda_{1,2}} : \Lambda_{1,2} \rightarrow \text{Diff}^1(M)$  has uniform subexponential growth of derivatives;*
- (2) *for any  $\varepsilon > 0$  there is a  $t_\varepsilon > 0$  such that for any orbit  $\zeta = \{a^s(x)\}_{0 \leq s \leq t}$  with  $x \in X_{\text{thick}}$ ,  $a^t(x) \in X_{\text{thick}}$ , and  $l(\zeta) = t \geq t_\varepsilon$  we have*

$$c(\zeta) \leq \varepsilon l(\zeta).$$

Define the maximal fiberwise growth rate of orbits starting and returning to  $X_{\text{thick}}$  to be

$$\chi_{\max} := \limsup_{t>0} \left\{ \sup \left\{ \frac{\log \|D_x(a^t)|_{\text{Fiber}}\|}{t} : x \in X_{\text{thick}}, a^t(x) \in X_{\text{thick}} \right\} \right\}. \tag{25}$$

Using Claim 5.2, to establish Proposition 5.1 it is sufficient to show that  $\chi_{\max} = 0$ .

For an orbit  $\zeta = \{a^s(x)\}_{0 \leq s \leq t}$ , define the following function which measures the depth of  $\zeta$  into the cusp:

$$d(\zeta) = \max_{0 \leq s \leq t} \text{dist}(a^s(x), X_{\text{thick}}).$$

The following lemma is the main result of this subsection.

**Lemma 5.3** *If  $\chi_{\max} > 0$  then there exists a sequence of orbits  $\zeta_n = \{a^s(x_n)\}_{0 \leq s \leq t_n}$  with  $x_n \in X_{\text{thick}}, a^{t_n}(x_n) \in X_{\text{thick}}$ , and  $t_n = l(\zeta_n) \rightarrow \infty$  such that*

- (1)  $c(\zeta_n) \geq \frac{\chi_{\max}}{2} t_n$ ;
- (2)  $\lim_{n \rightarrow \infty} \frac{d(\zeta_n)}{t_n} = 0$ .

We first have the following claim.

**Claim 5.4** *For any  $\varepsilon > 0$  there exists  $t_\varepsilon$  with the following properties: for any  $x \in \partial X_{\text{thick}}$  and  $t \geq t_\varepsilon$  such that  $a^s(x) \in X \setminus X_{\text{thick}}$  for all  $0 < s < t$  and  $a^t(x) \in \partial X_{\text{thick}}$  then, for the orbit  $\zeta = \{a^s(x)\}_{0 \leq s \leq t}$ , we have*

$$c(\zeta) \leq \varepsilon t = \varepsilon l(\zeta).$$

Indeed, the claim follows from the fact that the value of the return cocycle  $\beta(a^s, x)$  is defined by geodesic in the cusp of  $X$  is given by a unipotent matrix of the form  $\begin{pmatrix} 1 & n \\ 0 & 1 \end{pmatrix} \in \Lambda_{1,2} \subset \text{SL}(m, \mathbb{Z})$  and Proposition 4.1.

*Proof of Lemma 5.3* Let  $\zeta_n := \{a^s(x_n)\}_{0 \leq s \leq t_n}$  be a sequence of orbits with  $x_n \in X_{\text{thick}}, a^{t_n}(x_n) \in X_{\text{thick}}, t_n \rightarrow \infty$ , and such that

$$\chi_{\max} = \lim_{n \rightarrow \infty} \frac{c(\zeta_n)}{t_n}.$$

Replacing  $\zeta_n$  with a subsequence, we may assume the following limit exists:

$$\beta := \lim_{n \rightarrow \infty} \frac{d(\zeta_n)}{t_n}.$$

We aim to prove that  $\beta = 0$ . Arguing by contradiction, suppose  $0 < \beta \leq 1$ . We decompose the orbit

$$\zeta_n = \alpha_{k_n} \omega_{k_n-1} \alpha_{k_n-1} \cdots \omega_1 \alpha_1$$

as a concatenation of smaller orbit segments  $\alpha_i, \omega_i$  with the following properties:

- (1) each orbit  $\alpha_i$  is such that  $d(\alpha_i) \leq \frac{\beta}{2}t_n$ ;
- (2) the endpoints of each orbit  $\alpha_i$  are contained in  $X_{\text{thick}}$ ;
- (3) each orbit  $\omega_i$  is contained entirely in  $(X \setminus X_{\text{thick}}) \cup \partial X_{\text{thick}}$  with endpoints contained in  $\partial X_{\text{thick}}$ ;
- (4) each orbit  $\omega_i$  satisfies  $d(\omega_i) \geq \frac{\beta}{2}t_n$  whence  $l(\omega_i) \geq \frac{\beta}{2}t_n$  for  $t_n$  sufficiently large.

Note for each  $n$ , that  $k_n \leq \lfloor \frac{2}{\beta} \rfloor + 1$  and thus  $k_n$  is bounded by some  $k$  independent of  $n$ . Additionally, since  $SL(m, \mathbb{Z})$  is finitely generated and (equipped with the word metric) is quasi-isometrically embedded in  $SL(m, \mathbb{R})$ , there exists a constant  $K$  such that for any orbit segment  $\zeta$  whose endpoints are contained in  $X_{\text{thick}}$ , we have  $c(\zeta) \leq Kl(\zeta)$ . By the definition of  $\chi_{\text{max}}$ , for any  $\varepsilon > 0$  there is a positive constant  $M_\varepsilon$  such that for any orbit sub-segment  $\alpha_i$

- (1)  $c(\alpha_i) \leq (\chi_{\text{max}} + \varepsilon)l(\alpha_i)$  whenever  $l(\alpha_i) > M_\varepsilon$
- (2)  $c(\alpha_i) \leq KM_\varepsilon$  whenever  $l(\alpha_i) \leq M_\varepsilon$ .

From Claim 5.4, for any  $\varepsilon > 0$  we have, assuming that  $n$  and hence  $t_n$  are sufficiently large, that

$$c(\omega_i) < \varepsilon l(\omega_i)$$

for all orbit sub-segments  $\omega_i$ .

Taking  $n$  sufficiently large we have

$$(\chi_{\text{max}} - \varepsilon)t_n < c(\zeta_n) \leq \sum_i c(\omega_i) + \sum_i c(\alpha_i). \tag{26}$$

As we assume  $\beta > 0$ , for all sufficiently large  $n$  there exists at least one orbit sub-segment  $\omega_i$  and thus for such  $n$

$$\begin{aligned} \sum_i c(\alpha_i) &\leq kKM_\varepsilon + (\chi_{\text{max}} + \varepsilon) \\ &\times \sum_i l(\alpha_i) \leq kKM_\varepsilon + (\chi_{\text{max}} + \varepsilon)(1 - \beta/2)t_n. \end{aligned} \tag{27}$$

From (26) and (27) we obtain that

$$(\chi_{\max} - \varepsilon)t_n \leq (\varepsilon t_n) + \left(kKM_\varepsilon + (\chi_{\max} + \varepsilon)(1 - \beta/2)t_n\right). \tag{28}$$

Dividing by  $t_n$  and taking  $n \rightarrow \infty$  obtain

$$\chi_{\max} - \varepsilon \leq \varepsilon + (\chi_{\max} + \varepsilon)(1 - \beta/2).$$

As we assumed  $\chi_{\max} > 0$  and  $\beta > 0$ , we obtain a contradiction by taking  $\varepsilon > 0$  sufficiently small. □

### 5.3 Construction of a Følner sequence and family averaged measures

Assuming that  $\chi_{\max}$  in (25) is non-zero, we start from the orbit segments constructed in Lemma 5.3 and perform an averaging procedure to obtain a family of measures  $\{\mu_n\}$  on  $M^\alpha$  whose properties lead to a contradiction. In particular, the projection of any weak- $*$  limit  $\mu_\infty$  of  $\mu_n$  to  $M^\alpha$  will be  $A$ -invariant, well behaved at the cusps, and have non-zero Lyapunov exponents. These measures on  $M^\alpha$  are obtained by averaging certain Dirac measures against Følner sequences in a certain amenable subgroup of  $G$ .

Consider the copy of  $SL(m - 1, \mathbb{R}) \subset SL(m, \mathbb{R})$  as the subgroup of matrices that differ from the identity away from the  $m$ th row and  $m$ th column. Let  $N' \simeq \mathbb{R}^{m-1}$  be the abelian subgroup of unipotent elements that differ from the identity only in the  $m$ th column; that is given a vector  $r = (r_1, r_2, \dots, r_{m-1}) \in \mathbb{R}^{m-1}$  define  $u^r$  to be the unipotent element

$$u^r = \begin{pmatrix} 1 & 0 & 0 & \dots & r_1 \\ & 1 & 0 & \dots & r_2 \\ & & \ddots & & \vdots \\ & & & 1 & r_{m-1} \\ & & & & 1 \end{pmatrix} \tag{29}$$

and let  $N' = \{u^r\}$ .  $N'$  is normalized by  $SL(m - 1, \mathbb{R})$ .

Identifying  $N'$  with  $\mathbb{R}^{m-1}$  we have an embedding  $SL(m - 1, \mathbb{R}) \ltimes \mathbb{R}^{m-1} \subset SL(m, \mathbb{R})$ . The subgroup  $SL(m - 1, \mathbb{R}) \ltimes \mathbb{R}^{m-1}$  has as a lattice the subgroup

$$SL(m - 1, \mathbb{Z}) \ltimes \mathbb{Z}^{m-1} := \Gamma \cap (SL(m - 1, \mathbb{R}) \ltimes \mathbb{R}^{m-1})$$

and there is a natural embedding given by the inclusion

$$(SL(m - 1, \mathbb{R}) \ltimes \mathbb{R}^{m-1}) / (SL(m - 1, \mathbb{Z}^{-1}) \ltimes \mathbb{Z}^{m-1}) \subset SL(m, \mathbb{R}) / SL(m, \mathbb{Z}).$$

Recall  $A$  is the group of diagonal matrices with positive entries. Let  $a^t, b^s \in A$  denote matrices

$$\begin{aligned}
 a^t &= \text{diag}(e^{t/2}, e^{-t/2}, 1, 1, \dots, 1) \\
 b^s &= \text{diag}(e^s, e^s, e^s, \dots, e^s, e^{-s(m-1)}).
 \end{aligned}$$

Complete the set  $\{a, b\}$  to a spanning set  $\{a, b, c_1, c_2, \dots, c_{m-3}\}$  of  $A$  viewed as vector space where the  $c_i$  are diagonal matrices whose  $(m, m)$ -entry is equal to 1.

Let  $F_n \subset AN'$  be the subset of  $G$  consisting of all the elements of the form

$$a^t b^s \prod_{c=1}^{m-3} c_i^{s_i} u^r \tag{30}$$

where, for some  $\delta > 0$  to be determined later (in the proof of Proposition 5.10 below),

- (1)  $0 < t < t_n$ ;
- (2)  $\delta t_n/2 < s < \delta t_n$ ;
- (3)  $0 < s_i < \sqrt{t_n}$ ;
- (4)  $r \in B_{\mathbb{R}^{m-1}}(e^{200t_n})$ .

**Claim 5.5**  $\{F_n\}$  is Følner sequence in  $AN'$ .

Observe that  $F_n$  is linearly-long in the  $a$ -direction and exponentially-long in the  $N'$ -direction. From conditions (2) and (4), the  $A$ -component of  $F_n$  is much longer in the  $a^t$ -direction than in the other directions. The condition (2) that  $\delta t_n/2 < s$  is fundamental in our estimates in Sect. 5.4 that ensure the measures constructed below  $\{\mu_n\}$  have uniformly exponentially small mass in the cusps. These estimates are related to the fact that orbits of  $N'$  correspond to the unstable manifolds for the flow defined by  $b_s$  in  $SL(m, \mathbb{R})/SL(m, \mathbb{Z})$  and open subsets of unstable manifolds equidistribute to the Haar measure on  $SL(m, \mathbb{R})/SL(m, \mathbb{Z})$  under the flow  $b^s$ .

Recall we have a sequence of fiber bundles

$$F \rightarrow M^\alpha \rightarrow G/\Gamma$$

and may consider  $F$  as a fiber bundle over  $G/\Gamma$ . Given  $x \in G/\Gamma$ , let  $F(x) \simeq TM$  denote the fiber of  $F$  over  $x$ . An element  $v \in F(x)$  is a pair  $v = (y, \xi)$  where, identifying the fiber of  $M^\alpha$  through  $x$  with  $M$ , we have  $y \in M$  and  $\xi \in T_y M$ . Given  $v = (y, \xi) \in F(x)$ , we write  $\|v\| = \|\xi\|$  using our chosen norm on  $F$ . Given  $v = (y, \xi) \in F(x)$ , let  $p(v) = y$  denote the footpoint of  $v$  in the fiber of  $M^\alpha$  through  $x$ .

If uniform subexponential growth of derivatives fails for the restriction of the  $\alpha$  to  $\Lambda_{1,2}$ , then there exist sequences  $x_n \in X_{\text{thick}}$ ,  $v_n \in F(x_n)$  with  $\|v_n\| = 1$ , and  $t_n \in \mathbb{R}$  as in Lemma 5.3 and Claim 5.2 with  $t_n \rightarrow \infty$ , such that

$$\|D_{x_n} a_n^{t_n}(v_n)\| \geq e^{\lambda t_n} \tag{31}$$

for some  $\lambda > 0$ .

Note that  $AN'$  is a solvable group. We may equip  $AN'$  with any left-invariant Haar measure. Note that the ambient Riemannian metric induces a right-invariant Haar measure on  $AN'$  but as  $AN'$  is not unimodular these measures do not coincide.

For each  $n$ , take  $\mu_n$  to be the measure on  $M^\alpha$  obtained by averaging the Dirac measure  $\delta_{(x_n, p(v_n))}$  over the set  $F_n$ :

$$\mu_n := \frac{1}{|F_n|_\ell} \int_{F_n} g \cdot \delta_{(x_n, p(v_n))} dg$$

where  $|F_n|_\ell$  is the volume of  $F_n$  and  $dg$  indicates integration with respect to left-invariant Haar measure on  $AN'$ .

We expand the above integral in our coordinates introduced above. Then for any bounded continuous function  $f: M^\alpha \rightarrow \mathbb{R}$ , integrating against our Euclidean parameters  $t, s, s_i$ , and  $r$  we have

$$\int_{M^\alpha} f d\mu_n = \frac{2 \int_0^{t_n} \int_{\delta t_n/2}^{\delta t_n} \int_{[0, \sqrt{t_n}]^{m-3}} \int_{B_{\mathbb{R}^{m-1}}(e^{200t_n})} f \left( a^t b^s \prod_{c=1}^{m-3} c_i^{s_i} u^r \cdot (x_n, p(v_n)) \right) dr ds_i ds dt}{t_n \delta t_n \sqrt{t_n}^{m-3} |B_{\mathbb{R}^{m-1}}(e^{200t_n})|} \tag{32}$$

where  $|B_{\mathbb{R}^{m-1}}(e^{200t_n})|$  denotes the volume of

$$B_{\mathbb{R}^{m-1}}(e^{200t_n}) = N'_{t_n} \subset N'$$

with respect to the Euclidean parameters  $r$ .

For each  $n$ , let  $\nu_n$  denote the image of the measure  $\mu_n$  under the canonical projection from  $M^\alpha$  to  $G/\Gamma$ . The following proposition is shown in the next subsection.

**Proposition 5.6** *There exists  $\eta > 0$  such that the sequence of measures  $\{\nu_n\}$  has uniformly exponentially small mass in the cusp with exponent  $\eta$ .*

By the uniform comparability of distances in fibers of  $M^\alpha$ , this implies the family of measures  $\{\mu_n\}$  has uniformly exponentially small measure in the cusp.



By Lemma 3.10(a) the families of measures  $\{\mu_n\}$  and  $\{\nu_n\}$  are precompact families. As  $F_n$  is a Følner sequence in a solvable group, we have that any weak- $*$  subsequential limit of  $\{\mu_n\}$  or  $\{\nu_n\}$  is  $AN'$ -invariant. Moreover, from Theorem 3.1(d), it follows that any weak- $*$  subsequential limit  $\nu_\infty$  of  $\{\nu_n\}$  is invariant under the group  $-N'$  generated by the root groups  $U^{m,j}$  for each  $1 \leq j \leq m - 1$ . Since  $N'$  and  $-N'$  generate all of  $G$ , we have that  $\nu_\infty$  is a  $G$ -invariant measure on  $G/\Gamma$ .

### 5.4 Proof of Proposition 5.6

#### 5.4.1 Heuristics of the proof

The heuristic of the proof is the following. Observe that for a fixed choice of  $t$  and  $s_i$  as given by the choice of Følner set  $F_n$ , the point

$$a^t \prod_{i=1}^{m-3} c_i^{s_i}(x_n)$$

lies at sub-linear distance to the thick part of  $G/\Gamma$  with respect to  $t_n$ . Observe that the  $N'$ -orbit of such point is an embedded  $(m - 1)$ -dimensional torus in  $G/\Gamma$ . As the range of points in  $N'$  in the Følner set  $F_n$  is quite large, averaging a Dirac measure of the point  $a^t \prod_{i=1}^{m-3} c_i^{s_i}(x_n)$  in the  $N'$ -direction in  $F_n$  yields a measure quite close to Haar measure on the  $N'$ -orbit.

Observe that  $N'$ -orbits correspond to unstable manifolds for the action of the flow  $b^s$  on  $SL(m, \mathbb{R})/SL(m, \mathbb{Z})$ . As the action of  $b_s$  is known to be mixing, we expect that if  $s$  is sufficiently large, flowing by  $b_s$  the  $N'$ -orbit of  $a^t \prod_{i=1}^{m-3} c_i^{s_i}(x_n)$  will become equidistributed and in particular it will intersect non-trivially the thick part of  $G/\Gamma$ . This is the reason why the condition  $s > \delta/2t_n$  is assumed.

While intuition about mixing motivates the proof, we do not use it explicitly. Instead we use that for large enough  $s$ , the action of  $b_s$  expands the  $N'$ -orbits in a way that forces them to hit the thick part. We verify this fact by explicit matrix multiplication.

As  $b^s$  normalizes  $N'$ , the image under  $b^s$  of the  $N'$ -orbit of  $a^t \prod_{i=1}^{m-3} c_i^{s_i}(x_n)$  is the  $N'$ -orbit of a point  $y_n$  in the thick part of  $G/\Gamma$ . Having in mind the quantitative non-divergence of unipotent flows as in the proof Proposition 3.2, the  $N'$ -orbits have uniformly (over all  $n, s_i$ , and  $t$ ) exponentially small mass in the cusps whence so do the measures  $\nu_n$ .

The following proof of Proposition 5.6 uses explicit matrix calculations and estimates to verify these heuristics.

5.4.2 Proof of Proposition 5.6

Recall that we identify each coset

$$g\text{SL}(m, \mathbb{Z}) \in \text{SL}(m, \mathbb{R})/\text{SL}(m, \mathbb{Z})$$

with a unimodular lattice  $\Lambda_g := g \cdot \mathbb{Z}^m$  in  $\mathbb{R}^m$ . We define the systole of a unimodular lattice  $\Lambda \subset \mathbb{R}^m$  to be

$$\delta(\Lambda) := \min_{v \in \Lambda \setminus \{0\}} \|v\|$$

and for an element  $g \in \text{SL}(m, \mathbb{R})$ , we denote by  $\delta(g)$  the systole

$$\delta(g) = \delta(g \cdot \mathbb{Z}^m).$$

From (9), to prove Proposition 5.6 it is sufficient to find  $\eta > 0$  so that the integrals

$$\int_{G/\Gamma} \delta(g)^{-\eta} d\nu_n(g\Gamma)$$

are uniformly bounded in  $n$ .

As discussed in the above heuristic, from (32) to bound the integrals  $\int_{G/\Gamma} \delta^{-\eta}(g) d\nu_n(g\Gamma)$  it is sufficient to show each integral

$$\frac{1}{|B(e^{200t_n})|} \int_{B(e^{200t_n})} \delta(a^t b^s (\Pi c_i^{s_i}) u^r x_n)^{-\eta} dr$$

is uniformly bounded in  $n$  and in all parameters  $t, s, s_i$  for  $0 < t < t_n$ ,  $\delta t_n/2 < s < \delta t_n$ , and  $0 < s_i < \sqrt{t_n}$ . Recall here that  $x_n \in G/\Gamma$  are the points  $x_n \in X_{\text{thick}} \subset H_{1,2}/\Lambda_{1,2}$  satisfying (31) used in the construction of the measures  $\mu_n$ .

We have  $H_{1,2}$  is canonically embedded in  $\text{SL}(m, \mathbb{R})$ . Given  $x_n \in H_{1,2}/\Lambda_{1,2}$ , let

$$\tilde{x}_n \in H_{1,2} \subset \text{SL}(m, \mathbb{R})$$

denote the element mapping to  $x_n$  under the map  $H_{1,2} \rightarrow H_{1,2}/\Lambda_{1,2}$  which is contained in a fundamental domain contained in the Dirichlet domain  $\mathcal{D} \subset \text{SL}(2, \mathbb{R})$  in (7) in Sect. 2.2. Let  $\|\cdot\|$  denote the operator norm on  $\text{SL}(m, \mathbb{R})$  and  $m(\cdot)$  the associated conorm.

**Claim 5.7** *For every  $n, t \leq t_n$ , and  $0 \leq s_i \leq \sqrt{t_n}$  as above, there exist*

$$A_n = A_{n,t,s_1,\dots,s_{m-3}} \in SL(m - 1, \mathbb{R}) \text{ and } \gamma_n = \gamma_{n,t,s_1,\dots,s_{m-3}} \in SL(m - 1, \mathbb{Z})$$

such that:

- (1)  $a^t \prod_{i=1}^{m-3} c_i^{s_i} \tilde{x}_n = \begin{pmatrix} A_n \gamma_n & 0_{m-1 \times 1} \\ 0_{1 \times m-1} & 1 \end{pmatrix}$
- (2)  $\lim_{n \rightarrow \infty} \sup_{t \leq t_n, 0 \leq s_i \leq \sqrt{t_n}} \frac{\log \|A_n\|}{t_n} = 0$  and  $\lim_{n \rightarrow \infty} \inf_{t \leq t_n, 0 \leq s_i \leq \sqrt{t_n}} \frac{\log(m(A_n))}{t_n} = 0$

*Proof* (1) is immediate from construction. The uniform limit in (2) follows from Lemma 5.3(2), equation (5), and the fact that the  $s_i$  are chosen so that  $0 \leq s_i \leq \sqrt{t_n}$  whence

$$\frac{d(x_n, a^t(\prod c_i^{s_i}) \cdot x_n)}{t_n} \rightarrow 0$$

uniformly in  $t, s_i$ . □

In the remainder, we will suppress the dependence of choices on  $t, s, s_i$ . We take  $K_n \in SL(m - 1, \mathbb{R})$  be such that

$$\tilde{x}_n = \begin{pmatrix} K_n & 0_{m-1 \times 1} \\ 0_{1 \times m-1} & 1 \end{pmatrix}.$$

Note that  $K_n$  differs from the identity only in the first two rows and columns. Since each  $x_n$  is contained in  $X_{\text{thick}}$ , we have that the matrix norm and conorm  $\|K_n\|$  and  $m(K_n)$  are bounded above and below, respectively, by constants  $M_1$  and  $\frac{1}{M_1}$  independent of  $n$ .

Recall  $r$  denotes a vector in  $\mathbb{R}^{m-1}$  and  $u^r \in SL(m, \mathbb{R})$  is the unipotent element given by (29). Matrix computation yields

$$a^t(\prod c_i^{s_i})u^r \tilde{x}_n = \begin{pmatrix} A_n \gamma_n & A_n \gamma_n K_n^{-1} r \\ 0_{1 \times m-1} & 1 \end{pmatrix}$$

whence

$$b^s(\prod c_i^{s_i})a^t u^r \tilde{x}_n = \begin{pmatrix} e^s A_n \gamma_n & e^s A_n \gamma_n K_n^{-1} r \\ 0_{1 \times m-1} & e^{-(m-1)s} \end{pmatrix}.$$

We have

$$\begin{aligned} \delta(b^s (\prod c_i^{s_i}) a^t u^r \tilde{x}_n) &= \delta(b^s (\prod c_i^{s_i}) a^t u^r x_n) \\ &= \inf_{z \in \mathbb{Z}^m \setminus \{0\}} \left\| \begin{pmatrix} e^s A_n \gamma_n & e^s A_n \gamma_n K_n^{-1} r \\ 0_{1 \times m-1} & e^{-(m-1)s} \end{pmatrix} z \right\| \end{aligned} \tag{33}$$

To reduce notation, for fixed  $t, s,$  and  $s_i$  define

$$\beta(r) := -\log \delta(a^t b^s (\prod c_i^{s_i}) u^r \tilde{x}_n).$$

We aim to find an upper bound of

$$\frac{1}{|B(e^{200t_n})|} \int_{B(e^{200t_n})} e^{\eta\beta(r)} dr$$

that is independent of  $n$  and  $t, s,$  and  $s_i.$

Observe that if  $r - r'$  differ by an element of the unimodular lattice  $K_n \mathbb{Z}^{m-1} \subset \mathbb{R}^{m-1},$  then  $\beta(r) = \beta(r').$  Indeed, if  $r' = r + K_n z'$  for some  $z' = (z'_1, \dots, z'_{m-1}) \in \mathbb{Z}^{m-1}$  and if  $z \in \mathbb{Z}^m \setminus \{0\}$  is  $z = (z_1, \dots, z_m)$  then

$$\begin{pmatrix} e^s A_n \gamma_n & e^s A_n \gamma_n K_n^{-1} r' \\ 0_{1 \times m-1} & e^{-(m-1)s} \end{pmatrix} z = \begin{pmatrix} e^s A_n \gamma_n & e^s A_n \gamma_n K_n^{-1} r \\ 0_{1 \times m-1} & e^{-(m-1)s} \end{pmatrix} \tilde{z}$$

where  $\tilde{z} = (z_1 + z_m z'_1, \dots, z_{m-1} + z_m z'_{m-1}, z_m) \in \mathbb{Z}^m \setminus \{0\}.$  Thus we have that  $\beta: \mathbb{R}^{m-1} \rightarrow (0, \infty)$  descends to a function on the torus  $\mathbb{R}^{m-1} / (K_n \mathbb{Z}^{m-1}).$

Let  $D_n = K_n \cdot ([-1/2, 1/2]^{m-1})$  be a fundamental domain for this torus in  $\mathbb{R}^{m-1}$  centered at 0. Let  $c_n$  denote the number of  $(K_n \mathbb{Z}^{m-1})$ -translates of  $D_n$  that intersect  $B(e^{200t_n}).$  Then, if  $t_n$  is sufficiently large we have that

$$\frac{1}{|B(e^{200t_n})|} \int_{B(e^{200t_n})} e^{\eta\beta(r)} dr \leq \frac{1}{|B(e^{200t_n})|} c_n \int_{D_n} e^{\eta\beta(r)} dr \leq 2 \int_{D_n} e^{\eta\beta(r)} dr$$

The first inequality follows from inclusion. The second inequality follows from the fact that the perimeter of  $B(q)$  grows like  $q^{m-2},$  the volume of  $B(q)$  grows like  $q^{m-1},$  and the domains  $D_n = K_n \cdot ([-1/2, 1/2]^{m-1})$  have uniformly comparable geometry over  $n.$

It remains to estimate  $\int_{D_n} e^{\eta\beta(r)} dr.$  Given  $c > 0$  and fixed  $n, t, s_i,$  and  $s$  we define

$$T_c = \{r \in D_n : \beta(r) > c\}.$$

Proposition 5.6 follows immediately from the estimate in the following lemma.

**Lemma 5.8** *There exists constants  $M_3, M_4 > 0$ , independent of  $n, t, s_i$ , and  $s$ , such that*

$$|T_c| \leq M_3 e^{-cM_4}.$$

Indeed, if  $\eta^{-1} > M_4$  then

$$\begin{aligned} \int_{D_n} e^{\eta\beta(r)} dr &= \int_0^\infty |\{r \in D_n : e^{\eta\beta(r)} \geq \tau\}| d\tau \\ &\leq 1 + \int_1^\infty |\{r \in D_n : e^{\eta\beta(r)} \geq \tau\}| d\tau \\ &= 1 + \int_1^\infty |\{r \in D_n : \beta(r) \geq \log(\tau^{\frac{1}{\eta}})\}| d\tau \\ &= 1 + \int_1^\infty |T_{\log(\tau^{\frac{1}{\eta}})}| d\tau \\ &= 1 + \int_1^\infty M_3 \tau^{-\frac{M_4}{\eta}} d\tau < \infty \end{aligned}$$

and Proposition 5.6 follows.

*Proof of Lemma 5.8* From (33), given any  $r \in \mathbb{R}^{m-1}$ , if  $\beta(r) > c$  then there exists a non-zero  $z = (z_1, z_2, z_3, \dots, z_m) \in \mathbb{Z}^m$  such that

$$e^s \|A_n \gamma_n(z_1, \dots, z_{m-1}) + z_m A_n \gamma_n K_n^{-1} r\| < e^{-c} \quad \text{and} \quad |z_m| < e^{-c} e^{(m-1)s}$$

which (as  $\gamma_n \in SL(m-1, \mathbb{Z})$ ) holds if and only if there is a non-zero  $z = (z_1, z_2, z_3, \dots, z_m) \in \mathbb{Z}^m$

$$\begin{aligned} e^s \|A_n((z_1, \dots, z_{m-1}) + z_m K_n^{-1} (K_n \gamma_n K_n^{-1}) r)\| &< e^{-c} \quad \text{and} \\ |z_m| &< e^{-c} e^{(m-1)s} \end{aligned} \tag{34}$$

As  $K_n \gamma_n K_n^{-1}$  induces a volume-preserving automorphism of  $\mathbb{R}^{m-1} / (K_n \mathbb{Z}^{m-1})$ , the set of  $r \in D_n$  satisfying (34) for some  $z \in \mathbb{Z}^m$  has the same measure as the set of  $r \in D_n$  satisfying

$$e^s \|A_n((z_1, \dots, z_{m-1}) + z_m K_n^{-1} r)\| < e^{-c} \quad \text{and} \quad |z_m| < e^{-c} e^{(m-1)s}$$

for some  $z \in \mathbb{Z}^m$ .

For every integer  $k$  satisfying  $|k| < e^{-c} e^{(m-1)s}$ , let  $T_{c,k}$  be the subset of  $r \in D_n$  such that there exists  $(z_1, z_2, \dots, z_{m-1}) \in \mathbb{Z}^{m-1}$  satisfying

$$e^s \|A_n((z_1, \dots, z_{m-1}) + k K_n^{-1} r)\| < e^{-c}.$$

Then  $|T_c| \leq \sum_{|k| < e^{-c} e^{(m-1)s}} |T_{c,k}|$ . Thus the estimate reduces to the following.

**Claim 5.9** *There exists  $M_5 \geq 0$  such that  $|T_{c,k}| < M_5 e^{-(m-1)(s+c)}$  for all  $n$  sufficiently large.*

*Proof* Recall that  $\delta t_n/2 < s$ . If  $k = 0$  then, for any non-zero  $(z_1, \dots, z_{m-1}) \in \mathbb{Z}^{m-1}$ , we have

$$e^s \|A_n(z_1, \dots, z_{m-1})\| > e^{\delta t_n/2} m(A_n).$$

From Claim 5.7(2), if  $n$  is large enough then  $m(A_n) \geq e^{-\delta t_n/4}$  and so the term in the left hand side above is greater than one, therefore  $T_{c,0} = \emptyset$  for  $n$  sufficiently large.

If  $k \neq 0$ , observe that the map  $M_k: \mathbb{R}^{m-1}/K_n\mathbb{Z}^{m-1} \rightarrow \mathbb{R}^{m-1}/K_n\mathbb{Z}^{m-1}$  given by

$$r + K_n\mathbb{Z}^{m-1} \mapsto kr + K_n\mathbb{Z}^{m-1}$$

preserves the Lebesgue measure on  $\mathbb{R}^{m-1}/K_n\mathbb{Z}^{m-1}$ . In particular, this implies that  $T_{c,k}$  and  $T_{c,1}$  have the same volume.

We thus take  $k = 1$ . Note that  $K_n^{-1}D_n = [-1/2, 1/2]$ . There is a  $L \geq 1$ , depending only on  $m - 1$ , such that the set

$$Q = \{z' \in \mathbb{Z}^{m-1} : |z' + r| \leq 1 \text{ for some } r \in K_n^{-1}D_n\}$$

has cardinality at most  $L$ . From Claim 5.7(2), if  $n$  is large enough then  $m(A_n) \geq e^{-\delta t_n/4}$  whence for all  $(z_1, \dots, z_{m-1}) \in \mathbb{Z}^{m-1} \setminus Q$  and all  $r \in D_n$ ,

$$e^s \left\| A_n \left( (z_1, \dots, z_{m-1}) + K_n^{-1}r \right) \right\| \geq 1.$$

We thus need only consider  $(z_1, \dots, z_{m-1}) \in Q$ .

Given a fixed  $z = (z_1, \dots, z_{m-1}) \in Q \subset \mathbb{Z}^{m-1}$ , using that  $K_n \in \text{SL}(m - 1, \mathbb{R})$  we have

$$\left| \{r \in \mathbb{R}^{m-1} : \|z + K_n^{-1}r\| \leq \ell\} \right| \leq (2\ell)^{m-1}$$

whence

$$\left| \{r \in \mathbb{R}^{m-1} : \|(z_1, \dots, z_{m-1}) + K_n^{-1}r\| \leq e^{-c}\} \right| \leq 2^{m-1} e^{-c(m-1)}.$$

If  $r \in T_{c,1}$  so that

$$e^s \left\| A_n \left( (z_1, \dots, z_{m-1}) + K_n^{-1}r \right) \right\| \leq e^{-c}$$

then

$$\|A_n((z_1, \dots, z_{m-1}) + K_n^{-1}r)\| \leq e^{-c-s}. \tag{35}$$

Since  $A_n \in SL(m - 1, \mathbb{R})$  the set of  $r \in \mathbb{R}^{m-1}$  satisfying (35) has the same volume as the set of  $r \in \mathbb{R}^{m-1}$  satisfying

$$\|(z_1, \dots, z_{m-1}) + K_n^{-1}r\| \leq e^{-c-s}.$$

It follows that  $|T_{c,1}| \leq 2^{m-1}Le^{-(s+c)(m-1)}$ . □

To finish the proof of Lemma 5.8, from Claim 5.9 we have

$$|T_c| \leq \sum_{|k| < e^{-c}e^{(m-1)s}} |T_{c,k}| \leq (2e^{-c}e^{(m-1)s} + 1)M_5e^{-(m-1)(s+c)} \leq M_3e^{-cM_4}$$

for some constants  $M_3, M_4$  independent of  $n$ . □

### 5.5 Positive Lyapunov exponents for limit measures

To deduce Proposition 5.1, having assumed that  $\chi_{\max}$  in (25) is non-zero, we show that any weak-\* subsequential limit of the sequence of measures  $\{\mu_n\}$  has a positive Lyapunov exponent from which we derive a contradiction.

Recall from Sect. 5.3 that we fixed sequences  $x_n, v_n, t_n$  such that  $\|D_{x_n}a^{t_n}(v_n)\| \geq e^{\lambda t_n}$  for some fixed  $\lambda > 0$ . Let  $\mathcal{A}: G \times F \rightarrow F$  be the fiberwise derivative cocycle over the action of  $G$  on  $M^\alpha$ .

Our main result is the following.

**Proposition 5.10** *For any weak-\* subsequential limit  $\mu_\infty$  of  $\{\mu_n\}$  we have*

$$\lambda_{\text{top},a,\mu_\infty,\mathcal{A}} \geq \lambda/2 > 0.$$

We first show that averaging over  $N'$  does not change the Lyapunov exponents of the cocycle.

**Claim 5.11** *Given any  $\varepsilon > 0$  there is  $t_\varepsilon > 0$  such that for any  $t \geq t_\varepsilon$  and any  $r \in B_{\mathbb{R}^{m-1}}(e^t)$  we have*

$$\|D_x u^r\|_{\text{Fiber}} \leq e^{\varepsilon t}$$

for any  $x \in X_{\text{thick}}$ .

*Proof* Recall that the  $N'$ -orbit of any  $x \in X := H_{1,2}/\Lambda_{1,2} \subset \mathrm{SL}(m, \mathbb{R})/\mathrm{SL}(m, \mathbb{Z})$  is a closed torus. Then the  $N'$ -orbit of  $X_{\text{thick}}$  is compact. Recall our fixed fundamental domain  $\mathcal{F} \subset \tilde{\mathcal{D}}$  contained in the Dirichlet domain  $\tilde{\mathcal{D}}$  of the identity for  $\mathrm{SL}(m, \mathbb{R})/\mathrm{SL}(m, \mathbb{Z})$  as discussed in Sect. 2.2. Given  $x \in \mathrm{SL}(m, \mathbb{R})/\mathrm{SL}(m, \mathbb{Z})$ , let  $\tilde{x}$  be the lift of  $x$  in  $\mathcal{F}$ . Let  $\tilde{X}_{\text{thick}} \subset H_{1,2} \cap \mathcal{F}$  denote the lift of  $X_{\text{thick}}$  to  $\mathcal{F}$  and let  $\hat{X}_{\text{thick}}$  be the lift of the orbit  $N'X_{\text{thick}}$  to  $\mathcal{F}$ . As discussed in Sect. 2.2, we have that  $\tilde{X}_{\text{thick}}$  is contained in the Dirichlet domain  $\mathcal{D}$  of the identity for the  $\Lambda_{1,2}$ -action on  $H_{1,2}$ . Moreover,  $\hat{X}_{\text{thick}}$  is precompact in  $\mathrm{SL}(m, \mathbb{R})$ .

Fix  $r \in \mathbb{R}^{m-1}$  and  $x \in X_{\text{thick}}$ . Write

$$\tilde{x} = \begin{pmatrix} K & 0 \\ 0 & 1 \end{pmatrix}$$

for some  $K \in \mathrm{SL}(m - 1, \mathbb{R})$ ; we have  $\|K\| \leq M_1$  and  $m(K) \geq \frac{1}{M_1}$  for all  $x \in X_{\text{thick}}$ . The deck group of the orbit  $N'\tilde{x}$  is

$$\tilde{x}\{u^z : z \in \mathbb{Z}^{m-1}\}\tilde{x}^{-1} = \{u^{K \cdot z} : z \in \mathbb{Z}^{m-1}\}.$$

Thus, there is  $z \in \mathbb{Z}^{m-1}$  and  $r' \in \mathbb{R}^{m-1}$  such that

$$\begin{aligned} u^r \tilde{x} &= \begin{pmatrix} K & r \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} K & r' + Kz \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & r' \\ 0 & 1 \end{pmatrix} \begin{pmatrix} K & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & z \\ 0 & 1 \end{pmatrix} \\ &= u^{r'} \tilde{x} u^z \end{aligned}$$

and  $u^{r'} \tilde{x} \in \hat{X}_{\text{thick}}$ . Then

$$\|D_x u^r\| \leq \|D_x \tilde{x}^{-1}\|_{\text{Fiber}} \cdot \|D_{\text{Id}} \Gamma u^z\|_{\text{Fiber}} \cdot \|D_{\text{Id}} \Gamma u^{r'} \tilde{x}\|_{\text{Fiber}}.$$

Since  $\tilde{x}$  and  $u^{r'} \tilde{x}$  are in precompact sets, the first and last terms of the right hand side are uniformly bounded in  $r$  and  $x \in X_{\text{thick}}$ .

There exists some  $C$  such that

$$\|D_{\text{Id}} \Gamma u^z\|_{\text{Fiber}} \leq C \|D\alpha(u^z)\|.$$

Since  $r \in B_{\mathbb{R}^{m-1}}(e^t)$  we have  $z \in B_{\mathbb{R}^{m-1}}(M_1 e^t)$  whence  $d(u^z, \text{Id}) \leq C_2 t + C_3$  for some constants  $C_2$  and  $C_3$ . Proposition 4.1 implies for any  $\varepsilon'$  that

$$\|D\alpha(u^z)\| \leq e^{\varepsilon'(C_2 t + C_3)}$$

and taking  $\varepsilon' > 0$  sufficiently small, the claim follows. □



By Lemma 2.1, the fact that  $SL(m, \mathbb{Z})$  is finitely generated, and the uniform comparability of the fibers of  $M^\alpha$ , we also have the following.

**Claim 5.12** *There are uniform constants  $C_5$  and  $C_6$  with the following property: Let  $x \in G/\Gamma$ . Then for any  $X \in \mathfrak{g}$  with  $\|X\| \leq 1$  we have*

$$\left\| (D_x \exp(tX))^{\pm 1} \right\|_{\text{Fiber}} \leq e^{C_5 t + C_5 d(x, \text{Id}) + C_6}.$$

We now prove Proposition 5.10.

*Proof of Proposition 5.10* Recall we take  $x_n \in X_{\text{thick}}$ ,  $t_n \rightarrow \infty$ , and  $v_n \in F(x_n)$  with  $\|v_n\| = 1$  such that  $\|D_{x_n} a^{t_n}(v_n)\| \geq e^{\lambda t_n}$  for some fixed  $\lambda > 0$  in (31) in Sect. 5.3. We also write  $\mathcal{A}: G \times F \rightarrow F$  for the fiberwise derivative cocycle.

The measures  $\mu_n$  constructed in Sect. 5.3 are defined by averaging last along the orbit  $a^t$ ,  $0 \leq t \leq t_n$ . Let  $\xi_n$  be the measure on  $M^\alpha$  given by

$$\begin{aligned} \int_{M^\alpha} f \, d\xi_n &= \frac{2}{\delta t_n} \left( \frac{1}{\sqrt{t_n}} \right)^{m-3} \frac{1}{|B_{\mathbb{R}^{m-1}}(e^{200t_n})|} \int_{\delta t_n/2}^{\delta t_n} \int_{[0, \sqrt{t_n}]^{m-3}} \int_{B_{\mathbb{R}^{m-1}}(e^{200t_n})} f \\ &\quad \times \left( a^t b^s \prod_{c=1}^{m-3} c_i^{s_i} u^r \cdot (x_n, p(v_n)) \right) \, dr \, ds_i \, ds. \end{aligned}$$

In the context of Lemma 3.10, the measures  $\mu_n = \int_0^{t_n} (a_*^t \xi_n) \, dt$  constructed in Sect. 5.3 correspond to the empirical measures  $\eta_n = \eta(\log a, t_n, \xi_n)$  appearing in the proof of Lemma 3.10. From Lemma 3.10, to establish Proposition 5.10 it is sufficient to show that

$$\int \log \|\mathcal{A}(a^{t_n}, \cdot)\| \, d\xi_n \geq \frac{\lambda}{2} t_n.$$

We have

$$\begin{aligned} &\int_{M^\alpha} \log \|\mathcal{A}(a^{t_n}, \cdot)\| \, d\xi_n \\ &= \frac{2}{\delta t_n} \left( \frac{1}{\sqrt{t_n}} \right)^{m-3} \frac{1}{|B_{\mathbb{R}^{m-1}}(e^{200t_n})|} \int_{\delta t_n/2}^{\delta t_n} \int_{[0, \sqrt{t_n}]^{m-3}} \int_{B_{\mathbb{R}^{m-1}}(e^{200t_n})} \\ &\quad \times \log \|\mathcal{A}(a^{t_n}, b^s \Pi c_i^{s_i} u^r \cdot (x_n, p(v_n)))\| \, dr \, ds_i \, ds \\ &\geq \frac{2}{\delta t_n} \left( \frac{1}{\sqrt{t_n}} \right)^{m-3} \frac{1}{|B_{\mathbb{R}^{m-1}}(e^{200t_n})|} \int_{\delta t_n/2}^{\delta t_n} \int_{[0, \sqrt{t_n}]^{m-3}} \int_{B_{\mathbb{R}^{m-1}}(e^{200t_n})} \\ &\quad \times \log \frac{\|D_{x_n}(a^{t_n} b^s \Pi c_i^{s_i} u^r)(v_n)\|}{\|D_{x_n}(b^s \Pi c_i^{s_i} u^r)(v_n)\|} \, dr \, ds_i \, ds. \end{aligned}$$

Consider fixed  $r, s,$  and  $s_i$ . Take  $r' \in \mathbb{R}^{m-1}$  such that  $a^{t_n} u^r = u^{r'} a^{t_n}$ . Then

$$\begin{aligned} \log \frac{\|D_{x_n}(a^{t_n} b^s \Pi c_i^{s_i} u^r)(v_n)\|}{\|D_{x_n}(b^s \Pi c_i^{s_i} u^r)(v_n)\|} &= \log \frac{\|D_{a^{t_n} \cdot x_n}(b^s \Pi c_i^{s_i} u^{r'}) \circ D_{x_n} a^{t_n}(v_n)\|}{\|D_{x_n}(b^s \Pi c_i^{s_i} u^r)(v_n)\|} \\ &\geq \log \|D_{x_n} a^{t_n}(v_n)\| - \log \|(D_{x_n} b^s \Pi c_i^{s_i} u^r)\|_{\text{Fiber}} \\ &\quad - \log \|(D_{a^{t_n} \cdot x_n}(b^s \Pi c_i^{s_i} u^{r'}))^{-1}\|_{\text{Fiber}} \\ &\geq \log \|D_{x_n} a^{t_n}(v_n)\| - \log \|D_{x_n}(u^r)\|_{\text{Fiber}} - \log \|D_{u^r \cdot x_n}(b^s \Pi c_i^{s_i})\|_{\text{Fiber}} \\ &\quad - \log \|(D_{u^{r'} a^{t_n} \cdot x_n}(b^s \Pi c_i^{s_i}))^{-1}\|_{\text{Fiber}} - \log \|D_{a^{t_n} \cdot x_n}(u^{-r'})\|_{\text{Fiber}}. \end{aligned}$$

Observe that both  $u^r \cdot x_n$  and  $u^{r'} a^{t_n} \cdot x_n$  are contained in a fixed compact subset of  $G/\Gamma$  and hence, by Claim 5.12, having taken  $\delta > 0$  sufficiently small in the construction of the Følner sequence, from the constraints on  $s_i$  and  $s$  we have  $\|D_{u^r \cdot x_n} \Pi c_i^{s_i} b^s\|_{\text{Fiber}} \leq e^{\lambda t_n/100}$  and  $\|(D_{u^{r'} a^{t_n} \cdot x_n} \Pi c_i^{s_i} b^s)^{-1}\|_{\text{Fiber}} \leq e^{\lambda t_n/100}$  for all  $n$  sufficiently large.

Moreover, from Claim 5.11, we have  $\|D_{x_n} u^r\|_{\text{Fiber}} \leq e^{\lambda t_n/100}$  for all  $n$  sufficiently large.

Finally, there exists  $\kappa > 0$  such that  $\|r'\| \leq e^{\kappa t_n} \|r\|$  whence  $r' \in B_{\mathbb{R}^{m-1}}(e^{(200+\kappa)t_n})$ . Again from Claim 5.11, we have  $\|D_{a^{t_n} \cdot x_n} u^{-r'}\|_{\text{Fiber}} \leq e^{\lambda t_n/100}$  for  $n$  sufficiently large. Combined with (31) we then have

$$\frac{1}{t_n} \int_{M^\alpha} \log \|\mathcal{A}(a^{t_n}, \cdot)\| d\xi_n \geq \lambda - \frac{4}{100} \lambda.$$

Proposition 5.10 then follows from Lemma 3.10. □

### 5.6 Proof of Proposition 5.1

Having assumed that  $\chi_{\max}$  in (25) is non-zero, we arrive at a contradiction. Take any weak- $*$  subsequential limit  $\mu_\infty$  of the sequence of measure  $\{\mu_n\}$  on  $M^\alpha$ . We have that  $\mu_\infty$  is  $A$ -invariant and has a non-zero fiberwise Lyapunov exponent for the fiberwise derivative over the action of  $a^t$ . Moreover, we have that  $\mu_\infty$  projects to  $\nu_\infty$  on  $G/\Gamma$  which, as discussed above after Proposition 5.6, is the Haar measure on  $G/\Gamma$ . We may replace  $\mu_\infty$  with an  $A$ -ergodic component  $\mu'$  with the same properties as above. Then  $\mu$  is  $A$ -ergodic, projects to Haar, and the fiberwise derivative cocycle over the  $A$ -action on  $(M^\alpha, \mu)$  has a non-zero Lyapunov exponent functional  $\lambda_i : A \rightarrow \mathbb{R}$ .

As in the conclusion of Lemma 4.4, the arguments of [4, Section 5.5] using [6, Proposition 5.1] imply that the measure  $\mu$  is, in fact,  $\text{SL}(m, \mathbb{R})$ -invariant. As before, we note that [6, Proposition 5.1] does not assume  $\Gamma$  is cocompact, so the algebraic argument applying that proposition in [4, Section 5.5] goes through

verbatim. For a more self-contained proof that applies since we only consider the case of  $SL(m, \mathbb{R})$  see [3, Proposition 4]. We then obtain a contradiction with Zimmer's cocycle superrigidity by constraints on the dimension of the fibers of  $M^\alpha$ . Thus we must have  $\chi_{\max} = 0$  and Proposition 5.1 follows.

**Acknowledgements** We thank Dave Witte Morris for his generous willingness to answer questions of all sorts throughout the production of this paper and [4]. We also thank to Shirali Kadyrov, Jayadev Athreya and Alex Eskin for helpful conversations, particularly on the material in Sect. 1.5 and Mikael de la Salle for many helpful conversations regarding strong property (T). We also thank the anonymous referee for a very careful reading and numerous comments which helped improve the exposition.

## References

1. Athreya, J.S.: Quantitative recurrence and large deviations for Teichmüller geodesic flow. *Geom. Dedic.* **119**, 121–140 (2006)
2. Billingsley, P.: *Convergence of Probability Measures*. Wiley Series in Probability and Statistics: Probability and Statistics, 2nd edn. Wiley, New York (1999)
3. Brown, A., Damjanovic, D., Zhang, Z.:  $C^1$  actions on manifolds by lattices in Lie groups. Preprint (2018). [arXiv:1801.04009](https://arxiv.org/abs/1801.04009)
4. Brown, A., Fisher, D., Hurtado, S.: Zimmer's conjecture: subexponential growth, measure rigidity, and strong property (T). Preprint (2016). [arXiv:1608.04995](https://arxiv.org/abs/1608.04995)
5. Brown, A., Rodriguez Hertz, F.: Smooth ergodic theory of  $\mathbb{Z}^d$ -actions part 1: Lyapunov exponents, dynamical charts, and coarse Lyapunov manifolds. Preprint (2016). [arXiv:1610.09997](https://arxiv.org/abs/1610.09997)
6. Brown, A., Rodriguez Hertz, F., Wang, Z.: Invariant measures and measurable projective factors for actions of higher-rank lattices on manifolds. Preprint (2016). [arXiv:1609.05565](https://arxiv.org/abs/1609.05565)
7. Burger, M., Monod, N.: Continuous bounded cohomology and applications to rigidity theory. *Geom. Funct. Anal.* **12**, 219–280 (2002)
8. de la Salle, M.: Strong (T) for higher rank lattices (2017). [arXiv:1711.01900](https://arxiv.org/abs/1711.01900)
9. Eskin, A., McMullen, C.: Mixing, counting, and equidistribution in Lie groups. *Duke Math. J.* **71**, 181–209 (1993)
10. Eskin, A., Mirzakhani, M.: Counting closed geodesics in moduli space. *J. Mod. Dyn.* **5**, 71–105 (2011)
11. Fisher, D., Margulis, G.A.: Local rigidity for cocycles. In: *Surveys in Differential Geometry*, Vol. VIII (Boston, MA, 2002), Volume 8 of *Surveys in Differential Geometry*, pp. 191–234. International Press, Somerville (2003)
12. Franks, J., Handel, M.: Distortion elements in group actions on surfaces. *Duke Math. J.* **131**, 441–468 (2006)
13. Ghys, É.: Actions de réseaux sur le cercle. *Invent. Math.* **137**, 199–231 (1999)
14. Kadyrov, S., Kleinbock, D.Y., Lindenstrauss, E., Margulis, G.A.: Singular systems of linear forms and non-escape of mass in the space of lattices. Preprint (2016). [arXiv:1407.5310](https://arxiv.org/abs/1407.5310)
15. Kleinbock, D.Y., Margulis, G.A.: Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. Math. (2)* **148**, 339–360 (1998)
16. Kleinbock, D.: An extension of quantitative nondivergence and applications to Diophantine exponents. *Trans. Am. Math. Soc.* **360**, 6497–6523 (2008)
17. Lubotzky, A., Mozes, S., Raghunathan, M.S.: Cyclic subgroups of exponential growth and metrics on discrete groups. *Comptes Rendus Acad. Sci. Paris Sér. I Math.* **317**, 735–740 (1993)

18. Lubotzky, A., Mozes, S., Raghunathan, M.S.: The word and Riemannian metrics on lattices of semisimple groups. *Inst. Hautes Études Sci. Publ. Math.* **2000**, 5–53 (2001)
19. Margulis, G.A.: Explicit constructions of expanders. *Probl Peredači Inf.* **9**, 71–80 (1973)
20. Margulis, G.: Problems and conjectures in rigidity theory. In: *Mathematics: Frontiers and Perspectives*, pp. 161–174. American Mathematical Society, Providence (2000)
21. Polterovich, L.: Growth of maps, distortion in groups and symplectic geometry. *Invent. Math.* **150**, 655–686 (2002)
22. Ratner, M.: Invariant measures and orbit closures for unipotent actions on homogeneous spaces. *Geom. Funct. Anal.* **4**, 236–257 (1994)
23. Ratner, M.: On Raghunathan’s measure conjecture. *Ann. Math. (2)* **134**, 545–607 (1991)
24. Shah, N.A.: Limit distributions of polynomial trajectories on homogeneous spaces. *Duke Math. J.* **75**, 711–732 (1994)
25. Shalom, Y.: Rigidity of commensurators and irreducible lattices. *Invent. Math.* **141**, 1–54 (2000)
26. Witte, D.: Arithmetic groups of higher  $Q$ -rank cannot act on 1-manifolds. *Proc. Am. Math. Soc.* **122**, 333–340 (1994)

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.