Invent. math. (2023) 231:1141–1237
https://doi.org/10.1007/s00222-022-01162-5

Polynomial effective density in quotients of $\mathbb{H}^3$ and $\mathbb{H}^2 \times \mathbb{H}^2$

E. Lindenstrauss[1] · **A. Mohammadi**[2]

Received: 5 April 2021 / Accepted: 1 September 2022 / Published online: 23 September 2022

Abstract We prove effective density theorems, with a polynomial error rate, for orbits of the upper triangular subgroup of $\mathrm{SL}_2(\mathbb{R})$ in arithmetic quotients of $\mathrm{SL}_2(\mathbb{C})$ and $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$. The proof is based on the use of a Margulis function, tools from incidence geometry, and the spectral gap of the ambient space.

Contents

E.L. acknowledges support by ERC 2020 Grant HomDyn (Grant No. 833423). A.M. acknowledges support by the NSF Grants DMS-1764246 and 2055122.

✉ A. Mohammadi
ammohammadi@ucsd.edu

E. Lindenstrauss
elon@math.huji.ac.il

1 The Einstein Institute of Mathematics, Edmond J. Safra Campus, Givat Ram, The Hebrew University of Jerusalem, Jerusalem 91904, Israel

2 Department of Mathematics, University of California, San Diego, CA 92093, USA

1 Introduction

The quantitative understanding of the behavior of orbits in homogeneous spaces is a fundamental problem. Let G be a connected Lie group and $\Gamma \subset G$ a lattice (a discrete subgroup with finite covolume). Let $L \subset G$ be a closed connected subgroup. Ratner's celebrated resolution of Raghunathan's conjectures, [51–53], provides a complete classification for the closure of *individual* L-orbits in G/Γ if L is unipotent, or more generally is generated by unipotent subgroups (this is true even if L is not assumed to be connected, see [59]). Prior to Ratner's work, some important special cases of this problem were studied by Margulis [44], and Dani and Margulis [14, 15].

These remarkable results all share the lacuna that they are not quantitative, e.g. they do not provide any rate at which the orbit fills up its closure. Indeed Ratner's work relies on the pointwise ergodic theorem which is hard to effectivize. The work of Dani and Margulis uses minimal sets, which though formally ineffective can be effectivized with some effort; a result in this spirit was obtained by Margulis and the first named author in [41], though the rates obtained there are of polylog form, and that too after significant effort. With Margulis and Shah, we have obtained a general effective orbit closure theorem for unipotent orbits on arithmetic quotients, the first piece of this being [42] and the continuation is in preparation; however the rates obtained are even worse than [41].

When G is a unipotent group, Green and Tao gave an effective equidistribution theorem for orbits of subgroups $L \subset G$ (that of course will also be unipotent) in [30] with polynomial error rates. When G is semisimple, however, not much seems to be known. A notable exception is the case where $L \subset G$ is a horospherical subgroups, that is to say if there is an element $a \in G$ so that

$$L = \{g \in G : a^n g a^{-n} \to 1 \text{ as } n \to \infty\},$$

for instance if L is the full group of strictly upper triangular matrices in $G = \mathrm{SL}_n(\mathbb{R})$. In this case, the behaviour of individual orbits can be related to decay of matrix coefficients, and hence effective equidistribution with polynomial error rate can be established. The first works in this direction we are aware of by Sarnak [54], Burger [10], and Kleinbock and Margulis [37] based on Margulis' thesis, as well as the more recent papers by Flaminio and Forni [26], Strömbergsson [60], and Sarnak and Ubis [55]. Quantitative horospheric equidistribution has now been established in much greater generality e.g. by Kleinbock and Margulis in [36], McAdam in [47] and by Asaf Katz [34]. Moreover a quantitative equidistribution estimate twisted by a character was proved by Venkatesh [64] and further developed by Tani and Vishe as well as Flaminio, Forni, and Tanis [27,63]; this was generalized to a disjointness result with a general nil-system by Asaf Katz in [34]. Closely related is the case of translates of periodic orbits of subgroups $L \subset G$ which are fixed by an involution by Duke, Rudnick and Sarnak, Eskin and McMullen, and Benoist and Oh in [2,16,23].

Beyond the horospherical case[1] (and the related case of groups fixed by an involution) equidistribution results with polynomial rates were known only for skew products by Strömbergsson [61], Strömbergsson and Vishe [62] and by Wooyeon Kim [35], for random walks by automorphisms of the torus (cf. [6] by Bourgain Furman, Mozes and the first named author and subsequent works in this direction, e.g. [32] by He and de Saxce), and for the special case of periodic orbits of increasing volume by Einsiedler, Margulis, Venkatesh

[1] Strictly speaking, the twisted horospherical averages considered in [27,34,63,64] can also be considered as a non-horospherical flow on a suitable product space, though they are closely related to the horospherical case.

and by these three authors with the second named author [17,18]. There are also some quantitative equidistribution results for particular types of unipotent orbits, e.g. [11] by Chow and Lei Yang.

In this paper, we prove an effective density theorem, with a *polynomial* error rate, for orbits of the upper triangular subgroup of $\mathrm{SL}_2(\mathbb{R})$ in arithmetic quotients of $\mathrm{SL}_2(\mathbb{C})$ and $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$. These are first results in the literature which provide a polynomial rate for general orbits in a homogeneous space of a semisimple group, beyond the aforementioned case of horospherical subgroups.

Let us now fix some notation in order to state the main theorems. Let

$$G = \mathrm{SL}_2(\mathbb{C}) \quad \text{or} \quad G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R}).$$

Let $\Gamma \subset G$ be a lattice, and put $X = G/\Gamma$.

Let d be the right invariant metric on G which is defined using the killing form. This metric induces a metric d_X on X, and natural volume forms on X and its submanifolds. The injectivity radius of a point $x \in X$ may be defined using this metric. For every $\eta > 0$, let

$$X_\eta = \{x \in X : \text{injectivity radius of } x \text{ is } \geq \eta\}.$$

Throughout the paper, H denotes $\mathrm{SL}_2(\mathbb{R})$ if $G = \mathrm{SL}_2(\mathbb{C})$ or the diagonally embedded copy of $\mathrm{SL}_2(\mathbb{R})$ in G if $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$. That is

$$\mathrm{SL}_2(\mathbb{R}) \subset \mathrm{SL}_2(\mathbb{C}) \quad \text{or} \quad \{(g, g) : g \in \mathrm{SL}_2(\mathbb{R})\} \subset \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R}).$$

Let $P \subset H$ denote the group of upper triangular matrices in H.

An orbit $Hx \subset X$ is periodic if $H \cap \mathrm{Stab}(x)$ is a lattice in H. For the semisimple group H, the orbit Hx is periodic iff it is closed.

Let $|\ |$ denote the absolute value on $\mathbb{C}$, and let $\|\ \|$ denote the maximum norm on $\mathrm{Mat}_2(\mathbb{C})$ or $\mathrm{Mat}_2(\mathbb{R}) \times \mathrm{Mat}_2(\mathbb{R})$ with respect to the standard basis. For every $T > 0$ and every subgroup $L \subset G$, let

$$B_L(e, T) = \{g \in L : \|g - I\| \leq T\}.$$

The following is the main theorem in this paper.

Theorem 1.1 *Assume that Γ is an arithmetic lattice. For every $0 < \delta < 1/2$, every $x_0 \in X$, and large enough T (depending explicitly on δ and the injectivity radius of x_0) at least one of the following holds.*

(1) For every $x \in X_{T^{-\delta\kappa_1}}$, we have

$$d_X\Big(x, B_P\big(e, T^A\big).x_0\Big) \leq C_1 T^{-\delta\kappa_1}.$$

(2) There exists $x' \in X$ such that Hx' is periodic with $\mathrm{vol}(Hx') \leq T^{\delta}$, *and*

$$d_X(x', x_0) \leq C_1 T^{-1}.$$

where A, κ_1, and C_1 are positive constants depending on X.

The proof of Theorem 1.1 has a similar flavor to [28] by Gamburd, Jakobson, and Sarnak as well as to the work of Bourgain and Gamburd [7,8] and the aforementioned work of Bourgain, Furman, Lindenstrauss, and Mozes [6]. Indeed in the first step, we use a Diophantine condition to produce some dimension at a certain scale (*initial dimension*). In the second step, we use a Margulis function to show that by passing to a larger scale and translating $B_P(e, T^{\delta}).x_0$ with a random element of controlled size, we obtain a set with *large dimension*. Margulis functions were introduced in the context of homogeneous dynamics in [21] by Eskin, Margulis, and Mozes, and have become an indispensable tool in homogeneous dynamics and beyond.

We then use a projection theorem to move this additional dimension to the direction of a horospherical subgroup of G. The projection theorem we use is an adaptation of the work of Käenmäki, Orponen, and Venieri [33] and is based on the works of Wolff and Schlag [56,65]. Finally, we use an argument due to Venkatesh [64] to conclude the proof.

The main proposition

Let $U \subset N$ denote the group of upper triangular unipotent matrices in $H \subset G$, respectively.

More explicitly, if $G = \mathrm{SL}_2(\mathbb{C})$, then

$$N = \left\{ n(r, s) = \begin{pmatrix} 1 & r + is \\ 0 & 1 \end{pmatrix} : (r, s) \in \mathbb{R}^2 \right\}$$

and $U = \{n(r, 0) : r \in \mathbb{R}\}$; we will often denote the elements in U by u_r, i.e., $n(r, 0)$ will often be denoted by u_r for $r \in \mathbb{R}$. Let

$$V = \{n(0, s) = v_s : s \in \mathbb{R}\}.$$

If $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$, then

$$N = \left\{ n(r, s) = \left(\begin{pmatrix} 1 & r + s \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix} \right) : (r, s) \in \mathbb{R}^2 \right\}$$

and $U = \{n(r, 0) : r \in \mathbb{R}\}$. As before, $n(r, 0)$ will be denoted by u_r for $r \in \mathbb{R}$. Let $V = \{n(0, s) = v_s : s \in \mathbb{R}\}$. In both cases, we have $N = UV$.

The following proposition is a crucial step in the proof. Roughly speaking, it states that for every $x_0 \in X$, we can find a subset of V with dimension almost 1 near $P.x_0$ unless x_0 is extremely close to a periodic H-orbit with small volume.

Proposition 1.1 (Main Proposition) *There exists some $\eta_0 > 0$ depending on X with the following property.*

Let $0 < \theta, \delta < 1/2$, $0 < \eta < \eta_0$, and $x_0 \in X$. There are κ_2 and A', depending on θ, and T_1 depending on δ, η, and the injectivity radius of x_0, so that for all $T > T_1$ at least one of the following holds.

(1) There exists a finite subset $I \subset [0, 1]$ so that both of the following are satisfied.
   (a) The set I supports a probability measure ρ which satisfies

$$\rho(J) \leq C_\theta |J|^{1-\theta}$$

   for every interval J with $|J| \geq T^{-\delta\kappa_2}$ where $C_\theta \geq 1$ depends on θ.
   (b) There is a point $y_0 \in X_\eta$ so that

$$d_X\Big(v_s.y_0, B_P\Big(e, T^{A'}\Big).x_0\Big) \leq C_2 T^{-\delta\kappa_2}$$

   for all $s \in I \cup \{0\}$.

(2) There exists $x' \in X$ so that Hx' is periodic with $\mathrm{vol}(Hx') \leq T^\delta$ and

$$d_X(x', x_0) \leq C_2 T^{-1}.$$

where C_2 depends on X.

The proof of this proposition will be completed in Sect. 8; it involves three main steps, which we now outline.

(1) Let us assume that the injectivity radius of x_0 is bounded below by some constant depending on X; we can always reduce to this case using certain non-divergence results which are discussed in Sect. 3.
   Since we are interested in information about how points approach each other transversal to H, we will work with a thickening of $P.x_0$ with B^H, a *small* neighborhood of the identity in H. In the first step, we use Proposition 6.1 (a closing lemma) to show that either Proposition 1.1(2) holds, or we can find some $x \in \Big(\mathsf{B}^H \cdot B_P\Big(e, T^{O(\delta)}\Big)\Big).x_0$, whose injectivity radius is bounded below depending on X, so that any two nearby points in $\Big(\mathsf{B}^H \cdot B_P\Big(e, T^\delta\Big)\Big).x$ have distance $> T^{-1}$ transversal to H.

(2) Assuming Proposition 1.1(2) does not hold, in the second step, we use a Margulis function to show that translations of the aforementioned thickening of $B_P\left(e, T^{\delta}\right).x$ by certain random elements in $B_P\left(e, T^{O_\theta(1)}\right)$ have dimension $1-\theta$ transversal to H at scale $T^{-0.1\delta}$. This step is carried out in Sect. 7.
The random elements we use in this step further have the property that translations of $\left(\mathsf{B}^H \cdot B_P\left(e, T^{\delta}\right)\right).x$ with them stay near $P.x$ — this property is reminiscent of Margulis' thickening technique, albeit unlike the latter we only thicken in H and not in G.
(3) In the third step, we use a projection theorem (Theorem 5.1) combined with some arguments in homogeneous dynamics, to project the aforementioned entropy to the direction of N. This is the content of Sect. 5.

Let us now elaborate on how Proposition 1.1 may be used to complete the proof of Theorem 1.1.

The argument is based on the quantitative decay of correlations for the ambient space X: There exists $\kappa_X > 0$ so that

$$\left|\int \varphi(gx)\psi(x)\,\mathrm{d}m_X - \int \varphi\,\mathrm{d}m_X \int \psi\,\mathrm{d}m_X\right| \ll_{\varphi,\psi} e^{-\kappa_X d(e,g)} \tag{1.1}$$

for all $\varphi, \psi \in C_c^\infty(X) + \mathbb{C}\cdot 1$, where m_X is the probability Haar measure on X and d is our fixed right G-invariant metric on G. See e.g. [37, §2.4] and references there for (1.1); we note that κ_X is absolute if Γ is a congruence subgroup, see [9,13,29].

As it is well studied, (1.1) implies quantitative equidistribution results for expanding pieces of the horospherical group N in X. Note, however, that we are only supplied with the set

$$B = \{u_r v_s : r \in [0,1], s \in I\}$$

where I is as in Proposition 1.1, i.e., we do not have the luxury of using an open subset of N. To remedy this issue, we use an argument due to Venkatesh [64] and show that so long as θ is small enough — this is quantified using (1.1) — expanding translations of B are already equidistributed in X, see Proposition 4.2.

Periodic orbits

The techniques we develop here allow us to prove an effective density theorem for periodic orbits of H as well. We will show in Lemma 3.3 that there exists some $\eta_X > 0$ so that for every periodic orbit Y, we have

$$\mu_Y(X_{\eta_X}) \geq 0.9 \tag{1.2}$$

where μ_Y denotes the H-invariant probability measure on Y.

Theorem 1.2 *Let $Y \subset X$ be a periodic H-orbit in X. Then for every $x \in X_{\mathrm{vol}(Y)^{-\kappa_3}}$ we have*

$$d_X(x, Y) \leq C_3 \mathrm{vol}(Y)^{-\kappa_3}.$$

where $\kappa_3 \geq \kappa_X^4/L$ (for an absolute constant L) and C_3 depends explicitly on κ_X, vol(X)*, and the minimum of the injectivity radius of points in X_{η_X}, see (9.14). If Γ is congruence, κ_3 is absolute.*

If Γ is an arithmetic lattice, Theorem 1.2 is a rather special case of a theorem of Einsiedler, Margulis, and Venkatesh [18] or (when the corresponding $\mathbb{Q}$-group has over $\mathbb{R}$ compact factors) the followup work by Einsiedler, Margulis, and Venkatesh and the second named author [17]. Note however that Theorem 1.2 does *not* require Γ to be arithmetic. In particular, unlike [17,18], our argument does not rely on property (τ).

By the arithmeticity theorems of Selberg and Margulis, irreducible lattices in $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ are arithmetic. Regarding reducible quotients of $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$, if such a quotient $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R}) / \Gamma_1 \times \Gamma_2$ contains infinitely many closed orbits of H, then Γ_2 is commensurable to Γ_1 (up to a conjugation) and moreover Γ_1 has infinite index in its commensurator. By a theorem of Margulis, it follows that Γ_1 is arithmetic, see [45, Ch. IX]. Moreover, it was recently shown, [1,46], that if $\mathrm{SL}_2(\mathbb{C}) / \Gamma$ contains infinitely many closed orbits of H, then Γ is arithmetic.

Thus in all cases covered by Theorem 1.2, either Γ is arithmetic hence [17,18] apply (though the proof we give here is very different) or there are only finitely many closed H-orbits. The key point of Theorem 1.2 is that the rate of equidistribution depends only on rather coarse properties of X namely the rate of mixing κ_X, the volume of X, and the injectivity radius of the compact core of X, suitably interpreted. This can be used in some special cases to give an effective version of the finiteness theorems of [1,46], as we discuss in the next subsection. It is interesting to note that the proofs in [1,46] rely on equidistribution results [49] which are in the spirit of Theorem 1.2, albeit in a qualitative form.

Totally geodesic planes in hybrid manifolds

Gromov and Piatetski-Shapiro [31] constructed examples of non-arithmetic hyperbolic manifolds by gluing together pieces of non-commensurable arithmetic manifolds. Let Γ_1 and Γ_2 be two torsion free lattices in $\mathrm{Isom}(\mathbb{H}^3)$ — recall that $\mathrm{Isom}(\mathbb{H}^3)$ is an index 2 subgroup of $\mathrm{O}(3, 1)$ and that $\mathrm{SL}_2(\mathbb{C})$ is locally isomorphic to $\mathrm{O}(3, 1)$. Let $M_i = \mathbb{H}^3/\Gamma_i$. Assume further that for $i = 1, 2$, there exists 3-dimensional submanifolds with boundary $N_i \subset M_i$ so that

- The Zariski closure of $\pi_1(N_i) \subset \Gamma_i$ contains $\mathrm{O}(3, 1)^\circ$ where $\mathrm{O}(3, 1)^\circ$ is the connected component of the identity in $\mathrm{O}(3, 1)$.
- Every connected component of ∂N_i is a totally geodesic embedded surface in M_i which separates M_i.
- ∂N_1 and ∂N_2 are isometric.

Let M be the manifold obtained by gluing N_1 and N_2 using the isometry between ∂N_1 and ∂N_2. Then M carries a complete hyperbolic metric, thus, we consider $\pi_1(M)$ as a lattice in $\mathrm{O}(3, 1)$. Let $\Gamma' = \pi_1(M) \cap \mathrm{O}(3, 1)^\circ$, and let Γ denote the inverse image of Γ' in $G = \mathrm{SL}_2(\mathbb{C})$. If Γ_1 and Γ_2 are arithmetic and non-commensurable, then M is non-arithmetic, i.e., Γ is a non-arithmetic lattice in G. A totally geodesic plane in M lifts to a periodic orbit of $H = \mathrm{SL}_2(\mathbb{R})$ in $X = G/\Gamma$.

The following finiteness theorem, in qualitative form, was proved by Fisher, Lafont, Miller, and Stover [25, Thm. 1.4], see also [3, §12].

Theorem 1.3 *Let M be a hyperbolic 3-manifold obtained by gluing the pieces N_1 and N_2 from non-commensurable arithmetic manifolds along $\Sigma = \partial N_1 = \partial N_2$ as described above. The number of totally geodesic planes in M is at most*

$$L\left(\mathrm{area}(\Sigma)\mathrm{vol}(X)\eta_X^{-1}\kappa_X^{-1}\right)^{L/\kappa_X^4}$$

where L is absolute and $X = G/\Gamma$ is as above.

2 Notation and preliminaries

Throughout the paper

$$G = \mathrm{SL}_2(\mathbb{C}) \quad \text{or} \quad G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R}).$$

Let $\Gamma \subset G$ be a lattice, and put $X = G/\Gamma$.

We define the subgroups H, N, U, and V as in the introduction.

Also let $U^- = \{u_r^- : r \in \mathbb{R}\}$ denote the group of lower triangular unipotent matrices in H.

For every $t \in \mathbb{R}$, let a_t denote the images of

$$\begin{pmatrix} e^{t/2} & 0 \\ 0 & e^{-t/2} \end{pmatrix} \tag{2.1}$$

in H. Note that $a_t n(r,s) a_{-t} = n(e^t(r,s))$ for all $t \in \mathbb{R}$ and all $(r,s) \in \mathbb{R}^2$.

Lie algebras and norms

Let $|\ |$ denote the usual absolute value on $\mathbb{C}$ (and on $\mathbb{R}$). Let $\|\ \|$ denotes the maximum norm on $\mathrm{Mat}_2(\mathbb{C})$ and $\mathrm{Mat}_2(\mathbb{R}) \times \mathrm{Mat}_2(\mathbb{R})$, with respect to the standard basis.

Let $\mathfrak{g} = \mathrm{Lie}(G)$, that is, $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ or $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{R}) \oplus \mathfrak{sl}_2(\mathbb{R})$. We write $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{r}$ where $\mathfrak{h} = \mathrm{Lie}(H) \simeq \mathfrak{sl}_2(\mathbb{R})$, $\mathfrak{r} = i\mathfrak{sl}_2(\mathbb{R})$ if $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{C})$ and $\mathfrak{r} = \mathfrak{sl}_2(\mathbb{R}) \oplus \{0\}$ if $\mathfrak{g} = \mathfrak{sl}_2(\mathbb{R}) \oplus \mathfrak{sl}_2(\mathbb{R})$.

Throughout the paper, we will use the uniform notation

$$w = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix}$$

for elements $w \in \mathfrak{r}$, where $w_{ij} \in i\mathbb{R}$ if $G = \mathrm{SL}_2(\mathbb{C})$ and $w_{ij} \in \mathbb{R}$ if $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$.

Note that $\mathfrak{r}$ is a *Lie algebra* in the case $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$, but not when $G = \mathrm{SL}_2(\mathbb{C})$.

We fix a norm on $\mathfrak{h}$ by taking the maximum norm where the coordinates are given by $\mathrm{Lie}(U)$, $\mathrm{Lie}(U^-)$, and $\mathrm{Lie}(A)$; similarly fix a norm on $\mathfrak{r}$. By taking maximum of these two norms we get a norm on $\mathfrak{g}$. These norms will also be denoted by $\|\ \|$.

Let $C_4 \geq 1$ be so that

$$\|hw\| \leq C_4 \|w\| \text{ for all } \|h - I\| \leq 2 \text{ and all } w \in \mathfrak{g}. \tag{2.2}$$

For all $\beta > 0$, we define

$$\mathsf{B}_\beta^H := \{u_s^- : |s| \leq \beta\} \cdot \{a_t : |t| \leq \beta\} \cdot \{u_r : |r| \leq \beta\} \tag{2.3}$$

for all $0 < \beta < 1$. Note that for all $h_i \in (\mathsf{B}_\beta^H)^{\pm 1}$, $i = 1, \ldots, 5$, we have

$$h_1 \cdots h_5 \in \mathsf{B}_{100\beta}^H. \tag{2.4}$$

We also define $\mathsf{B}^G_\beta := \mathsf{B}^H_\beta \cdot \exp(B_{\mathfrak{r}}(0,\beta))$ where $B_{\mathfrak{r}}(0,\beta)$ denotes the ball of radius β in $\mathfrak{r}$ with respect to $\|\ \|$.

We deviate slightly from the notation in the introduction, and define the injectivity radius of $x \in X$ using B^G_β instead of the metric d on G. Put

$$\mathrm{inj}(x) = \min\Big\{0.01, \sup\Big\{\beta : g \mapsto gx \text{ is injective on } \mathsf{B}^G_{10\beta}\Big\}\Big\}. \tag{2.5}$$

Taking a further minimum if necessary, we always assume that the injectivity radius of x defined using the metric d dominates $\mathrm{inj}(x)$.

For every $\eta > 0$, let

$$X_\eta = \Big\{x \in X : \mathrm{inj}(x) \geq \eta\Big\}.$$

Constants and the $\star$-notation

In our analysis, the dependence of the exponents on Γ are via the application of results in Sect. 4, see (4.1), and Sect. 6.

We will use the notation $A \asymp B$ when the ratio between the two lies in $[C^{-1}, C]$ for some constant $C \geq 1$ which depends at most on G and Γ in general. We write $A \ll B^\star$ (resp. $A \ll B$) to mean that $A \leq CB^\kappa$ (resp. $A \leq CB$) for some constant $C > 0$ depending on G and Γ, and $\kappa > 0$ which follows the above convention about exponents.

Lemma 2.1 *There exist absolute constants β_0 and $C_5 \geq 1$ so that the following holds. Let $0 < \beta \leq \beta_0$, and let $w_1, w_2 \in B_{\mathfrak{r}}(0,\beta)$. There are $h \in H$ and $w \in \mathfrak{r}$ which satisfy*

$$0.5\|w_1 - w_2\| \leq \|w\| \leq 2\|w_1 - w_2\| \quad \textit{and} \quad \|h - I\| \leq C_5\beta\|w\|$$

so that $\exp(w_1)\exp(-w_2) = h\exp(w)$.

Proof Using the Baker–Campbell–Hausdorff formula, we have

$$\exp(w_1)\exp(-w_2) = \exp(w_1 - w_2 + \bar{w})$$

where $\bar{w} \in \mathfrak{g}$ and $\|\bar{w}\| \ll \beta\|w_1 - w_2\|$.

Using the open mapping theorem and Baker–Campbell–Hausdorff formula again, for all small enough β, there is $(w_{\mathfrak{h}}, w_{\mathfrak{r}}) = B_{\mathfrak{h}}(0, C\beta) \times B_{\mathfrak{r}}(0, C\beta)$ and $w' \in \mathfrak{g}$ with $\|w'\| \ll \|w_{\mathfrak{h}}\|\|w_{\mathfrak{r}}\|$, so that

$$\exp(w_1 - w_2 + \bar{w}) = \exp(w_{\mathfrak{h}})\exp(w_{\mathfrak{r}}) = \exp(w_{\mathfrak{h}} + w_{\mathfrak{r}} + w') \tag{2.6}$$

where C and the implied constant are absolute.

We show that $h = \exp(w_{\mathfrak{h}})$ and $w = w_{\mathfrak{r}}$ satisfy the claims in the lemma. In view of (2.6), we need to verify the bounds on $\|h - I\|$ and $\|w_{\mathfrak{r}}\|$.

First note that if β is small enough, (2.6) implies that

$$w_1 - w_2 + \bar{w} = w_{\mathfrak{h}} + w_{\mathfrak{r}} + w'. \tag{2.7}$$

Recall that we are using the max norm with respect to $\mathfrak{r}$ and $\mathfrak{h}$ which are two orthogonal subspaces. Note also that $w_1, w_2, w_{\mathfrak{r}} \in \mathfrak{r}$ and $w_{\mathfrak{h}} \in \mathfrak{h}$. Thus, (2.7) implies that $\|w_{\mathfrak{h}}\| \ll \|\bar{w}\| + \|w'\|$. Recall now that $\|\bar{w}\| \ll \beta\|w_1 - w_2\|$ and $\|w'\| \ll \|w_{\mathfrak{h}}\|\|w_{\mathfrak{r}}\| \ll \beta\|w_{\mathfrak{h}}\|$. Thus assuming β is small enough, we conclude that $\|w_{\mathfrak{h}}\| \ll \beta\|w_1 - w_2\|$ as we wanted to show.

To see the estimate on $\|w_{\mathfrak{r}}\|$, we again use (2.7). Indeed $(w_1 - w_2) - w_{\mathfrak{r}} = w_{\mathfrak{h}} + w' - \bar{w}$; moreover, $\|\bar{w}\| \ll \beta\|w_1 - w_2\|$, $\|w_{\mathfrak{h}}\| \ll \beta\|w_1 - w_2\|$, and $\|w'\| \ll \|w_{\mathfrak{h}}\|\|w_{\mathfrak{r}}\| \ll \beta\|w_{\mathfrak{h}}\| \ll \beta^2\|w_1 - w_2\|$. Again assuming β is small enough, we conclude that

$$0.5\|w_1 - w_2\| \le \|w_{\mathfrak{r}}\| \le 2\|w_1 - w_2\|,$$

which finishes the proof. □

Lemma 2.2 *There exists β_0 so that the following holds for all $0 < \beta \le \beta_0$. Let $x \in X_{10\beta}$ and $w \in B_{\mathfrak{r}}(0, \beta)$. If there are $h, h' \in \mathsf{B}^H_{2\beta}$ so that $\exp(w')hx = h'\exp(w)x$, then*

$$h' = h \quad \textit{and} \quad w' = \mathrm{Ad}(h)w.$$

Moreover, we have $\|w'\| \le 2\|w\|$.

Proof Recall that $\mathfrak{r}$ is invariant under the adjoint action of H. We rewrite the equation $\exp(w')hx = h'\exp(w)x$ as follows

$$\exp(w')hx = \exp(\mathrm{Ad}(h')w)h'x. \tag{2.8}$$

Since $h' \in \mathsf{B}^H_{2\beta}$, we have $\mathrm{Ad}(h')w' = w' + \hat{w}$ where $\|\hat{w}\| \ll \beta\|w'\|$. Therefore, assuming β is small enough, we have $0.5\|w\| \le \|\mathrm{Ad}(h')w'\| \le 2\|w\|$. This estimate, (2.8), and the fact that $x \in X_{10\beta}$ imply that

$$\exp(w')h = \exp(\mathrm{Ad}(h')w)h'.$$

Moreover, the map $(\bar{w}, \bar{h}) \mapsto \exp(\bar{w})\bar{h}$ from $B_{\mathfrak{r}}(0, 2\beta) \times \mathsf{B}^H_{2\beta}$ to G is injective, for all small enough β. Therefore, $h = h'$ and $w' = \mathrm{Ad}(h')w$.

The final claim follows as $\|w'\| = \|\mathrm{Ad}(h')w\| \le 2\|w\|$. □

The set $\mathsf{E}_{\eta,t,\beta}$

For all $\eta, \beta > 0$ and $t \geq 0$, set

$$\mathsf{E}_{\eta,t,\beta} := \mathsf{B}^H_\beta \cdot a_t \cdot \left\{u_r : r \in [0,\eta]\right\} \subset H. \tag{2.9}$$

Then $m_H(\mathsf{E}_{\eta,t,\beta}) \asymp \eta\beta^2 e^t$ where m_H denotes our fixed Haar measure on H.

Throughout the paper, the notation $\mathsf{E}_{\eta,t,\beta}$ will be used only for $\eta, t, \beta > 0$ which satisfy $e^{-0.01t} < \beta < \eta^2$ even if this is not explicitly mentioned.

For all $\eta, \beta, m > 0$, put

$$\mathsf{Q}^H_{\eta,\beta,m} = \left\{u^-_s : |s| \leq \beta e^{-m}\right\} \cdot \{a_t : |t| \leq \beta\} \cdot \left\{u_r : |r| \leq \eta\right\}. \tag{2.10}$$

Roughly speaking, $\mathsf{Q}^H_{\eta,\beta,m}$ is a *small thickening* of the (β,η)-neighborhood of the identity in AU. We write $\mathsf{Q}^H_{\beta,m}$ for $\mathsf{Q}^H_{\beta,\beta,m}$

The following lemma will also be used in the sequel.

Lemma 2.3 *(1) Let* $m \geq 1$, *and let* $0 < \eta, \beta < 0.1$. *Then*

$$\left(\left(\mathsf{Q}^H_{0.01\eta,0.01\beta,m}\right)^{\pm 1}\right)^3 \subset \mathsf{Q}^H_{\eta,\beta,m}.$$

(2) For all $0 \leq \beta \leq \eta \leq 1$, $t, m > 0$, *and all* $|r| \leq 2$, *we have*

$$\left(\mathsf{Q}^H_{\beta^2,m}\right)^{\pm 1} \cdot a_m u_r \mathsf{E}_{\eta,t,\beta'} \subset a_m u_r \mathsf{E}_{\eta,t,\beta}, \tag{2.11}$$

where $\beta' = \beta - 100\beta^2$.

Proof Recall that for all a, b, c, d with $ad - bc = 1$ and $a \neq 0$, we have

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ c/a & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & 1/a \end{pmatrix} \begin{pmatrix} 1 & b/a \\ 0 & 1 \end{pmatrix}.$$

The claim in part (1) follows from this identity.

To see part (2), recall that

$$(u^-_s a u_{r'}) \cdot (a_m u_r) = a_m u_r u_r^{-1} u^-_{e^m s} a u_{e^{-m} r'} u_r$$

for all $u^-_s a u_{r'} \in \mathsf{Q}^H_{\beta^2,m}$.

Note that $e^m|s| \leq \beta^2$ and $e^{-m}|r'| \leq \beta^2$. Let now

$$(u^-_c a_d u_b) \cdot a_t \cdot u_{r''} \in \mathsf{E}_{\eta,t,\beta-100\beta^2},$$

where $|c|, |d|, |b| \le \beta - 100\beta^2$, $|r''| \le \eta$.

Then

$$(u_s^- a u_{r'})(a_m u_r)(u_c^- a_d u_b a_t u_{r''}) = a_m u_r (u_r^{-1} u_{e^m s}^- a u_{e^{-m} r'} u_r)(u_c^- a_d u_b) a_t u_{r''}.$$

Since $|r| \le 2$, we have $u_r \cdot \mathsf{B}^H_{\beta^2} \cdot u_{-r} \subset \mathsf{B}^H_{10\beta^2}$. Moreover, $\mathsf{B}^H_{10\beta^2} \cdot \mathsf{B}^H_{\beta} \subset \mathsf{B}^H_{\beta+100\beta^2}$. The claim follows. □

A linear algebra lemma

Note that both $\mathfrak{h}$ and $\mathfrak{r}$ are invariant under the adjoint representation of H on $\mathfrak{g}$; moreover, both of these representations are isomorphic to the adjoint representation of H on $\mathrm{Lie}(H)$.

We will use the following lemma in the sequel

Lemma 2.4 ([22], Lemma 5.1, and [20]) *Let* $1/3 < \alpha < 1$, $0 \neq w \in \mathfrak{g}$, *and* $t > 0$. *Then*

$$\int_0^1 \|a_t u_r w\|^{-\alpha}\,\mathrm{d}r \le \frac{C_6 e^{-\hat{\alpha} t}}{2 - 2^{\alpha}} \|w\|^{-\alpha};$$

where C_6 *is an absolute constant and* $\hat{\alpha} = \frac{1-\alpha}{4}$.

We will apply the above lemma with $t = \ell m_\alpha$, $\ell \in \mathbb{N}$, where m_α is defined by $\frac{C_6}{2-2^\alpha} e^{-\hat{\alpha} m_\alpha} = e^{-1}$. The choice of m_α and Lemma 2.4 imply

$$\int_0^1 \|a_{m_\alpha} u_r w\|^{-\alpha}\,\mathrm{d}r \le e^{-1} \|w\|^{-\alpha}. \tag{2.12}$$

3 Nondivergence results

In this section, we record some facts which will be used to deal with non-uniform lattices; the results in this section are known to the experts. Our goal here is to tailor these results to our applications in the paper.

Throughout this section, Γ is assumed to be non-uniform unless otherwise is explicated. *We do not assume* Γ *is arithmetic in this section.*

To deal with cases where Γ may not be arithmetic, we appeal to some facts from hyperbolic geometry, see Case 1 below. If Γ is a non-uniform irreducible lattice in $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$, i.e. Case 2 below, Γ is arithmetic by a theorem of Selberg — this is a special case of Margulis' arithmeticity theorem.

Proposition 3.1 *There exist $C_7 \geq 1$ with the following property. Let $0 < \varepsilon, \eta < 1$ and $x \in X$. Let $I \subset [-10, 10]$ be an interval with $|I| \geq \eta$. Then*

$$\left|\left\{r \in I : \operatorname{inj}(a_t u_r x) < \varepsilon^2\right\}\right| < C_7 \varepsilon |I|$$

so long as $t \geq |\log(\eta^2 \operatorname{inj}(x))| + C_7$.

Proposition 3.1 in particular implies that for all $t \geq \log\big(\eta^2 \operatorname{inj}(x)\big) + O(1)$ *most* points in $\{a_t u_r x : r \in I\}$ return to a fixed compact subset of X.

For the proof of the proposition, it is more convenient to investigate two separate cases as follows. These are:

Case 1: $G = \mathrm{SL}_2(\mathbb{C})$ or $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ and Γ is reducible.

Case 2: $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ and Γ is irreducible.

The proofs ultimately rely on non-divergence results of Margulis, Dani, and Kleinbock. To prepare the stage for such results to be applicable, in Case 1 we use the thick-thin decomposition from hyperbolic geometry. This will be completed in this section. In Case 2 thanks to Selberg's theorem Γ is an arithmetic lattice. The proof in this case uses explicit reduction theory of such lattices and and the aforementioned works of Margulis et al; this proof is given in Appendix A.

Let us thus assume $G = \mathrm{SL}_2(\mathbb{C})$ or $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ and Γ is reducible. Let $\mathbb{F}$ denote $\mathbb{R}$ or $\mathbb{C}$, and let $\Delta \subset \mathrm{SL}_2(\mathbb{F})$ be a lattice. Using the thick-thin decomposition of $\mathrm{SL}_2(\mathbb{F})/\Delta$, there exists a compact subset $\mathfrak{S} \subset \mathrm{SL}_2(\mathbb{F})/\Delta$ and a finite collection of disjoint cusps $\{\mathfrak{C}_j : 1 \leq j \leq \ell\}$ so that

$$\mathrm{SL}_2(\mathbb{F})/\Delta = \mathfrak{S} \bigsqcup (\sqcup_{j=1}^{\ell} \mathfrak{C}_j).$$

Each cusp $\mathfrak{C}_j$ corresponds to the Δ-orbit of a parabolic fixed point of Δ in $\partial\mathbb{H}^d$, $d = 2$ or 3 depending on $\mathbb{F}$; alternatively, $\mathfrak{C}_j$ corresponds to a tube of closed U-orbits

$$a_t \mathsf{N} \mathsf{g}_j \Delta \subset \mathrm{SL}_2(\mathbb{F}) \qquad t < 0,$$

where N denotes the group of upper triangular unipotent matrices in $\mathrm{SL}_2(\mathbb{F})$.

We will also consider a linearized version of the thick-thin decomposition. It is more convenient to identify $\mathrm{SL}_2(\mathbb{F})/\{\pm I\}$ with $\mathrm{SO}(\mathsf{Q})^\circ$ where $\mathsf{Q}(v_1, v_2, v_3) = 2v_1 v_3 + v_2^2$ if $d = 2$, and $\mathsf{Q}(v_1, v_2, v_3, v_4) = 2v_1 v_4 + v_2^2 + v_3^2$ if $d = 3$. We choose this identification so that N fixes e_1 where $\{\mathsf{e}_j\}$ is the standard basis for $\mathbb{R}^{d+1}$.

If $d = 2$, that is $\mathbb{F} = \mathbb{R}$, we let $L = \mathrm{SO}(\mathsf{Q})^\circ$ and write $W = \mathbb{R}^3$. If $d = 3$, that is: $\mathbb{F} = \mathbb{C}$, we let L be the isometry group of the restriction of Q to the subspace W spanned by $\{\mathsf{e}_1, \mathsf{e}_3, \mathsf{e}_4\}$ — in the latter case $L \simeq \mathrm{PSL}_2(\mathbb{R})$ and

$h\mathsf{e}_2 = \mathsf{e}_2$ for all $h \in L$. Note that in both cases the adjoint action of H on $\mathfrak{sl}_2(\mathbb{R})$ factors through the action of L on W.

Set $v_j := g_j^{-1}\mathsf{e}_1$ for $1 \le j \le \ell$ where e_1 is the first coordinate vector in $\mathbb{R}^{d+1}$ and $g_j \in \mathrm{SL}_2(\mathbb{F})$. Note that $\Delta v_j \subset \mathbb{R}^{d+1}$ is a closed (and hence discrete) subset of $\mathbb{R}^{d+1}$, see e.g. [48, Lemma 6.2].

Given a point $g\Delta \in \mathrm{SL}_2(\mathbb{F})/\Delta$ we define

$$\omega_\Delta(g\Delta) = \max\Big\{2, \max\Big\{\|g\delta v_j\|^{-1} : \delta \in \Delta, 1 \le j \le \ell\Big\}\Big\}.$$

For the following see e.g. [48, §6].

Lemma 3.1 *Let $\Delta \subset \mathrm{SL}_2(\mathbb{F})$ be a lattice. There exists some $C = C(\Delta) > 2$ so that the following holds. Assume that $\omega_\Delta(g\Delta) \ge C$ for some $g\Delta \in \mathrm{SL}_2(\mathbb{F})/\Delta$. Then there exists some $1 \le j_0 \le \ell$ and some $\delta_0 \in \Delta$ so that $\|g\delta_0 v_{j_0}\|^{-1} = \omega_\Delta(g\Delta)$ and*

$$\|g\delta v_j\| > 1/C, \quad \textit{for all}\ \ (\delta, j) \ne (\delta_0, j_0).$$

We will also use the following elementary lemma.

Lemma 3.2 *Let $\eta > 0$, and let I be an interval of length at least η. There exists some C_8 so that the following holds. Let $\varrho > 0$, and let $v \in \mathrm{SO}(\mathsf{Q})^\circ.\mathsf{e}_1$. Then*

$$\Big|\Big\{r \in I : \|a_t u_r v\| \le e^t \eta \|v\| \varrho^2\Big\}\Big| \le C_8 \varrho |I|.$$

Proof Note that we may assume ϱ is small compared to absolute constants.

Let us consider the case $d = 3$, the other case, i.e., $d = 2$, is contained in this case. Recall that W denotes the $\mathbb{R}$-span of $\{\mathsf{e}_1, \mathsf{e}_3, \mathsf{e}_4\}$; write $v = c_v\mathsf{e}_2 + w_v$ where $w_v \in W$ and $c_v \in \mathbb{R}$. Since $\mathsf{Q}(v) = 0$, we have $\|w_v\| \ge c\|v\|$ for some absolute constant $0 < c < 1$. Moreover, for every $h \in L = H$

$$hv = c_v\mathsf{e}_2 + hw_v. \tag{3.1}$$

Identifying W with the adjoint representation of H, for every $w \in W$ and every $0 < \delta < 1$, let

$$I(w, \delta) = \Big\{r \in I : |(\mathrm{Ad}(u_r)w)_{12}| \le 0.01\delta\eta^2\|w\|\Big\},$$

where w_{ij} is the (i, j)-th entry of $w \in \mathfrak{sl}_2(\mathbb{R})$.

A direct computation gives

$$\Big(\mathrm{Ad}(u_r)w\Big)_{12} = -w_{21}r^2 - 2w_{11}r + w_{12}. \tag{3.2}$$

Therefore, $\sup_I |(\mathrm{Ad}(u_r)w)_{12}| \geq 0.01\eta^2\|w\|$ — recall that $|I| \geq \eta$. We conclude that $|I(w,\delta)| \leq C\delta^{1/2}|I|$ for some $C > 0$, see e.g. [38, §3].

Let $\delta = 100c^{-1}\varrho^2$, where we assume ϱ is small enough so that $\delta < 1$. Let v be as in the statement, and define w_v as above. Then $\|w_v\| \geq c\|v\|$ and $|I(w_v,\delta)| \leq 10Cc^{-1/2}\varrho|I|$.

Let $r \in I \setminus I(w_v,\delta)$, then

$$\|(\mathrm{Ad}(u_r)w_v)_{12}\| \geq c^{-1}\eta^2\|w_v\|\varrho^2.$$

Since a_t expands the $(1,2)$-entry by a factor of e^t, we conclude

$$\begin{aligned}\|a_t u_r v\| &\geq \|a_t u_r w_v\| && \text{by (3.1)}\\ &\geq e^t|(\mathrm{Ad}(u_r)w_v)_{12}| \geq c^{-1}e^t\eta^2\|w_v\|\varrho^2\\ &\geq e^t\eta^2\|v\|\varrho^2.\end{aligned}$$

The claim thus holds with $C_8 = 10Cc^{-1/2}$. □

Proof of Proposition 3.1: Case 1 Let us first consider $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$. Since Γ is reducible, there exists a finite index subgroup $\Gamma' \subset \Gamma$ so that $\Gamma' = \Gamma_1 \times \Gamma_2$. The constant C_7 in Proposition 3.1 is allowed to depend on the index of Γ' in Γ, thus, abusing the notation, we replace Γ by Γ' in the remaining parts of the argument. In particular,

$$X = X_1 \times X_2 = \mathrm{SL}_2(\mathbb{R})/\Gamma_1 \times \mathrm{SL}_2(\mathbb{R})/\Gamma_2.$$

Let us write ω_i for ω_{Γ_i}, for $i = 1, 2$. Define

$$\omega(x) := \max\{\omega_1(x_1), \omega_2(x_2)\} \tag{3.3}$$

for all $x = (x_1, x_2) \in X$.

We denote the corresponding vectors for Γ_1 by v_{1j}, $1 \leq j \leq \ell_1$, and for Γ_2 by v_{2k}, $1 \leq k \leq \ell_2$.

Note that $\omega(x) \asymp \mathrm{inj}(x)^{-1}$, see e.g. [48, Prop. 6.7]. Therefore, it suffices to prove the proposition with $\mathrm{inj}(x)$ replaced by $\omega(x)$.

Let $(g_1, g_2) \in G$, $(\gamma_1, \gamma_2) \in \Gamma$, $1 \leq j \leq \ell_1$, and $1 \leq k \leq \ell_2$. By Lemma 3.2 applied with $g_1\gamma_1 v_{1j}$ and $g_2\gamma_2 v_{2k}$, we conclude

$$\Big|\Big\{r \in I : \|a_t u_r(g_1\gamma_1 v_{1j}, g_2\gamma_2 v_{2k})\| \leq e^t\eta^2\|(g_1 v_{1j}, g_2 v_{2k})\|\varrho^2\Big\}\Big| \leq 2C_8\varrho|I|$$

for every $0 < \varrho < 1$.

Let $\varrho_0 = 0.1C_8^{-1}$, and choose $(g_1, g_2) \in G$ so that $x = (g_1\Gamma, g_2\Gamma)$. Then the above implies that for all $(\gamma_1, \gamma_2) \in \Gamma$, all $1 \le j \le \ell_1$, and all $1 \le k \le \ell_2$, there exists some $r \in I$ so that

$$\begin{aligned}\|a_t u_r (g_1\gamma_1 v_{1j}, g_2\gamma_2 v_{2k})\| &\ge e^t \eta^2 \|(g\gamma_1 v_{1j}, g_2\gamma_2 v_{2k})\|\varepsilon^2 \\ &\ge e^t \eta^2 \omega(x)^{-1} \varrho_0^2. \end{aligned} \tag{3.4}$$

In view of (3.4), and by choosing C_7 large enough to account for the implicit constant in $\omega(x) \asymp \mathrm{inj}(x)^{-1}$, we have

$$\sup\{\|a_t u_r (g_1\gamma_1 v_{1j}, g_2\gamma_2 v_{2k})\| : r \in I\} \ge \varrho_0^2$$

so long as $t \ge |\log(\eta^2 \,\mathrm{inj}(x))| + C_7$.

Therefore, we may apply [38, Thm. 4.1] and the proposition follows in this case. The argument in the case $G = \mathrm{SL}_2(\mathbb{C})$ is similar — in light of Lemma 3.1, the use of [38, Thm. 4.1] simplifies significantly. □

As we mentioned the proof in Case 2 is given in Appendix A.

Proposition 3.2 *There exists $0 < \eta_X < 1$, depending on X, so that the following holds. Let $0 < \eta < 1$ and let $x \in X$. Let $I \subset [-10, 10]$ be an interval with length at least η. Then*

$$|\{r \in I : a_t u_r x \in X_{\eta_X}\}| \ge 0.99|I|$$

for all $t \ge |\log(\eta^2 \,\mathrm{inj}(x))| + C_7$.

Proof Apply Proposition 3.1 with $\varepsilon = 0.01C_7^{-1}$. The claim thus holds with $\eta_X = \varepsilon^2$. □

3.1 The subsets X_{cpt} and $\mathfrak{S}_{\mathrm{cpt}}$

Decreasing η_X if necessary we always assume that $X \setminus X_{\eta_X}$ is a disjoint union (possibly empty) of finitely many cusps.

If X is compact, let $X_{\mathrm{cpt}} = X$; otherwise, let $X_{\mathrm{cpt}} = \{gx : x \in X_{\eta_X}, \|g - I\| \le 2\}$ where X_{η_X} is given by Proposition 3.2.

We also fix once and for all a compact subset with piecewise smooth boundary $\mathfrak{S}_{\mathrm{cpt}} \subset G$ which projects onto X_{cpt}.

We end this section with the following

Lemma 3.3 *Let Y be a periodic H-orbit. Then $\mu_Y(X_{\eta_X}) \ge 0.9$ where μ_Y denotes the H-invariant probability measure on Y.*

Proof Let $\varphi = \mathbb{1}_{X_\eta}$, and let $y \in Y$. Then by [37, §2.2.2] we have

$$\lim_{t\to\infty} \int_0^1 \varphi(a_t u_r y)\,\mathrm{d}r = \int \varphi\,\mathrm{d}\mu_Y.$$

The lemma thus follows from Proposition 3.2. □

4 From large dimension to effective density

In this section we use the exponential decay of correlations for the ambient space X to prove Proposition 4.2, which says that expanding translations of subsets of N which are foliated by local U orbits and have dimension close but not necessarily equal to 2 are equidistributed in X.

This proposition will be used in the proofs of Theorems 1.1 and 1.2, but it is also of independent interest. The proof is similar to an argument in [64, §3].

Recall our notation from Sect. 2: $n(r, s) = u_r v_s$ where $v_s = n(0, s)$ and $u_r = n(r, 0) \in U$. Recall also that $a_t n(r, s)a_{-t} = n(e^t(r, s))$ for all $t \in \mathbb{R}$ and all $(r, s) \in \mathbb{R}^2$.

We need the following estimate on the decay of correlations in X. There exists κ_X depending on X so that

$$\left|\int \varphi(gx)\psi(x)\,\mathrm{d}m_X - \int \varphi\,\mathrm{d}m_X \int \psi\,\mathrm{d}m_X\right| \ll e^{-\kappa_X d(e,g)}\mathcal{S}(\varphi)\mathcal{S}(\psi) \quad (4.1)$$

for all $\varphi, \psi \in C_c^\infty(X) + \mathbb{C}\cdot 1$ where the implied constant is absolute and d is our fixed right G-invariant on G, see e.g. [37, §2.4] and references there. We note that κ_X is absolute if Γ is a congruence subgroup, see [9,13,29].

Here $\mathcal{S}(\cdot)$ is a certain Sobolev norm on $C_c^\infty(X) + \mathbb{C}\cdot 1$ which is assumed to dominate $\|\cdot\|_\infty$ and the Lipschitz norm $\|\cdot\|_{\mathrm{Lip}}$. Moreover, $\mathcal{S}(g.f) \ll \|g\|^\star \mathcal{S}(f)$ where the implied constants are absolute.

Let us put

$$\bar{\mathsf{C}}_X = \eta_X^{-1}\mathrm{vol}(G/\Gamma) \quad (4.2)$$

where η_X is as in Proposition 3.2 and $\mathrm{vol}(G/\Gamma)$ is computed using the Riemannian metric d.

We also need the following statement.

Proposition 4.1 ([37], Prop. 2.4.8) *There exists $\kappa_4 \gg \kappa_X$ (where the implied constant is absolute) and an absolute constant κ_5 so that the following holds.*

Let $0 < \eta < 1$, $t > 0$, *and* $x \in X_\eta$. *Then for every* $f \in C_c^\infty(X) + \mathbb{C} \cdot 1$,

$$\left| \int_{B_N(0,1)} f(a_t n.x)\,\mathrm{d}n - \int f\,\mathrm{d}m_X \right| \leq C_9 \eta^{-1/\kappa_5} \mathcal{S}(f) e^{-\kappa_4 t}$$

where $B_N(0,1) = \left\{ u_r v_s : 0 \leq r, s \leq 1 \right\}$, *the measure on* N *is normalized so that* $B_N(0,1)$ *has measure* 1, *and* $C_9 \leq L\bar{\mathsf{C}}_X^L$ *for an absolute constant* L *and* $\bar{\mathsf{C}}_X$ *as in (4.2).*

Proof This statement is well known to the experts, see e.g. [34,36,37,47]; we reproduce the argument for the convenience of the reader.

Throughout the argument, the implied exponents are absolute and implied multiplicative constants are $\leq L\bar{\mathsf{C}}_X^L$ for an absolute L. Let $0 \leq \varphi^+ \leq 1$ be a smooth function supported on $B_N(0,1)$ so that $\int_{B_N(0,1)}(1 - \varphi^+)\,\mathrm{d}n \leq e^{-\kappa t}$ and $\mathcal{S}(\varphi^+) \ll e^{\star\kappa t}$ for some κ which will be optimized later. Then

$$\left| \int_{B_N(0,1)} f(a_t n.x)\,\mathrm{d}n - \int_N f(a_t n.x)\varphi^+(n)\,\mathrm{d}n \right| \ll \|f\|_\infty e^{-\kappa t}. \tag{4.3}$$

Recall that $B_N(0,1)X_\eta \subset X_{0.1\eta}$; using a smooth partition of unity argument, we can write $\varphi^+ = \sum_{j=1}^M \varphi_j^+$ so that $M \ll \eta^{-\star}$, $\mathcal{S}(\varphi_j^+) \ll \eta^{-\star} e^{\star\kappa t}$, and the map $g \mapsto gy$ is injective on $\mathrm{supp}(\varphi_j^+)$ for all $y \in B_N(0,1).X_\eta$ and all j.

In consequence, we may fix one φ_j^+ for the rest of the argument. Arguing as in [37, Prop. 2.4.8], see also [36, Thm. 2.3], there exists a compactly supported smooth function φ (an $e^{-\kappa t}$-*thickening* of φ_j^+ along the weak-stable directions in G) so that $\mathcal{S}(\varphi) \ll_X \eta^{-\star} e^{\star\kappa t}$ and

$$\left| \int_N f(a_t n.x)\varphi_j^+(n)\,\mathrm{d}n - \int_X f(a_t y)\varphi(y)\,\mathrm{d}m_X(y) \right| \ll \|f\|_{\mathrm{Lip}} e^{-\kappa t}, \tag{4.4}$$

where $\|f\|_{\mathrm{Lip}}$ is the Lipschitz constant of f.

Finally in view of (4.1), we have

$$\begin{aligned} \left| \int f(a_t y)\varphi(y)\,\mathrm{d}m_X(y) - \int f\,\mathrm{d}m_X \int \varphi\,\mathrm{d}m_X \right| &\ll \mathcal{S}(f)\mathcal{S}(\varphi)e^{-\kappa_X t} \\ &\ll \eta^{-\star} e^{\star\kappa t} \mathcal{S}(f) e^{-\kappa_X t}. \end{aligned} \tag{4.5}$$

The claim follows from (4.3), (4.4), and (4.5) by optimizing κ. □

The following is a generalization of Proposition 4.1 where one replaces the average over $B_N(0,1)$ with an average over certain subsets of dimension close to 2, but not necessarily equal to 2.

Proposition 4.2 *There exist κ_6 and ε_0 (both $\gg \kappa_X^2$ with an absolute implied constant) so that the following holds. Let $0 \le \varepsilon \le \varepsilon_0$ and $0 < b \le 0.1$. Let ρ be a probability measure on $[0,1]$ which satisfies*

$$\rho(J) \le Cb^{1-\varepsilon} \tag{4.6}$$

for every interval J of length b and a constant $C \ge 1$.

Let $0 < \eta < 1$, $x \in X_\eta$, then

$$\left| \int_0^1 \int_0^1 f(a_t u_r v_s.x)\, \mathrm{d}r\, \mathrm{d}\rho(s) - \int f\, \mathrm{d}m_X \right| \le C_{10} C \eta^{-\frac{1}{2\kappa_5}} \mathcal{S}(f) e^{-\kappa_6 t}$$

for all $|\log b|/4 \le t \le |\log b|/2$ and all $f \in C_c^\infty(X) + \mathbb{C}\cdot 1$, where $C_{10} \le L\bar{\mathsf{C}}_X^L$ for an absolute constant L and $\bar{\mathsf{C}}_X$ as in (4.2).

Proof We will prove this for the case $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$; the proof in the case $G = \mathrm{SL}_2(\mathbb{C})$ is similar.

Throughout the argument, the implicit multiplicative constants are $\le L\bar{\mathsf{C}}_X^L$ for some absolute L.

Without loss of generality, we may assume $\int_X f\, \mathrm{d}m_X = 0$.

Let $M \in \mathbb{N}$ be so that $1/M \le b \le 1/(M-1)$. For every $1 \le j \le M$, let $I_j = \left[\frac{j-1}{M}, \frac{j}{M}\right)$; also put $s_j = \frac{2j-1}{2M}$ and $c_j = \rho(I_j)$ for all j. Since I_j's are disjoint, we have $\sum_j c_j = 1$.

For all such j, let

$$\mathsf{B}_j = \left\{ u_r v_s : r \in [0,1],\, s \in (s_j - \tfrac{b}{4}, s_j + \tfrac{b}{4}) \right\}.$$

In view of the choice of M, we have $\mathsf{B}_j \cap \mathsf{B}_{j'} = \emptyset$ for all $j \ne j'$. Let $\varphi = \sum_j 2b^{-1} c_j \mathbb{1}_{\mathsf{B}_j}$. Then $\int_N \varphi(n(r,s))\, \mathrm{d}r\, \mathrm{d}s = 1$.

We make the following observation. Using (4.6), we have $c_j \le Cb^{1-\varepsilon}$ for all j. This and the fact that B_j's are disjoint imply that

$$\varphi(n(z)) \le \max\{2b^{-1}c_j : 1 \le j \le M\} \le 2Cb^{-\varepsilon} \tag{4.7}$$

for all $n(z) \in N$; here and in what follows, $z = (r,s)$ and $\mathrm{d}z = \mathrm{d}r\, \mathrm{d}s$.

Using the fact that I_j's are disjoint, we have

$$\int_0^1 \int_0^1 f(a_t u_r v_s.x)\,\mathrm{d}r\,\mathrm{d}\rho(s) = \sum_j \int_{I_j} \int f(a_t u_r v_s.x)\,\mathrm{d}r\,\mathrm{d}\rho(s);$$

thus, we conclude that

$$\begin{aligned} &\left| \int_0^1 \int_0^1 f(a_t u_r v_s.x)\,\mathrm{d}r\,\mathrm{d}\rho(s) - \sum_j c_j \int f(a_t u_r v_{s_j}.x)\,\mathrm{d}r \right| \\ &\leq \sum_j \int_{I_j} \int \left| f(a_t u_r v_s.x) - f(a_t u_r v_{s_j}.x) \right| \mathrm{d}r\,\mathrm{d}\rho(s) \ll \mathcal{S}(f) b^{1/2} \end{aligned} \tag{4.8}$$

where we used the facts that $|s - s_j| \leq b$ and $t \leq |\log b|/2$ in the last inequality.

In view of (4.8), thus, we need to bound $\sum_j c_j \int f(a_t u_r v_{s_j} x)\,\mathrm{d}r$. Similar to (4.8), we can now make the following computation.

$$\begin{aligned} &\left| \sum_j \int_0^1 c_j f(a_t n(r, s_j).x)\,\mathrm{d}r - \int_N \varphi(n(z)) f(a_t n(z).x)\,\mathrm{d}z \right| \\ &\quad \leq \sum_j \int_0^1 2b^{-1} c_j \int_{s_j - \frac{b}{4}}^{s_j + \frac{b}{4}} \left| f(a_t n(r, s_j).x) - f(a_t n(r, s).x)\,\mathrm{d}s \right| \mathrm{d}r \\ &\quad \ll \mathcal{S}(f) b^{1/2} \end{aligned} \tag{4.9}$$

where again we used the facts that $|s - s_j| \leq b$ and $t \leq |\log b|/2$.

Thus, it suffices to investigate

$$A_1 = \int \varphi(n(z)) f(a_t n(z).x)\,\mathrm{d}z.$$

To that end, let $\ell \geq 2$ be a parameter which will be optimized later. Set $\tau = e^{\frac{1-\ell}{\ell} t} = e^{-t + \frac{t}{\ell}}$, and define

$$A_2 := \frac{1}{\tau} \int_0^\tau \int \varphi(n(z)) f(a_t u_r n(z).x)\,\mathrm{d}z\,\mathrm{d}r;$$

roughly speaking, we introduce an extra averaging in the direction of U.

For every $0 \le r \le \tau$, we have $|(\mathsf{B}_j + r)\Delta \mathsf{B}_j| \ll |\mathsf{B}_j|\tau$. Hence,

$$\begin{aligned}\left|\int \varphi(z) f(a_t u_r n(z).x)\,\mathrm{d}z - \int \varphi(z) f(a_t n(z).x)\,\mathrm{d}z\right| \\ &\le \sum_j 2b^{-1}c_j \int_{(\mathsf{B}_j+r)\Delta\mathsf{B}_j} |f(a_t n(z)x)|\,\mathrm{d}z \\ &\le \sum_j 2b^{-1}c_j|\mathsf{B}_j|\tau\|f\|_\infty \\ &\le \|f\|_\infty\tau \ll \mathcal{S}(f)\tau;\end{aligned}$$

we used $|\mathsf{B}_j| = b/2$ for every j and $\sum c_j = 1$, in the penultimate inequality. Averaging the above over $[0, \tau]$, we conclude that

$$|A_1 - A_2| \ll \mathcal{S}(f)\tau \le \mathcal{S}(f)e^{-t/2} \ll \mathcal{S}(f)b^{1/8}; \tag{4.10}$$

recall that $\tau = e^{\frac{1-\ell}{\ell}t}$, $\ell \ge 2$, and $t \ge |\log b|/4$.

In consequence, we have reduced to the study of A_2 to which we now turn. By the Cauchy-Schwarz inequality, we have

$$|A_2|^2 \le \int \left(\frac{1}{\tau}\int_0^\tau f(a_t u_r n(z).x)\,\mathrm{d}r\right)^2 \varphi(n(z))\,\mathrm{d}z.$$

Now using $\left(\frac{1}{\tau}\int_0^\tau f(a_t n(r+z).x)\,\mathrm{d}r\right)^2 \ge 0$, (4.7), and the above estimate, we conclude

$$\begin{aligned}|A_2|^2 &\le 2Cb^{-\varepsilon}\int_{B(0,1)}\left(\frac{1}{\tau}\int_0^\tau f(a_t n(z)u_r.x)\,\mathrm{d}r\right)^2\mathrm{d}z \\ &= \frac{1}{\tau^2}\int_0^\tau\int_0^\tau\int_{B(0,1)} 2Cb^{-\varepsilon}\hat{f}_{r_1,r_2}(a_t n(z).x)\,\mathrm{d}z\,\mathrm{d}r_1\,\mathrm{d}r_2\end{aligned} \tag{4.11}$$

where $B(0,1) = B_N(0,1) = \{u_r v_s : 0 \le r, s \le 1\}$ has measure 1 with respect to $\mathrm{d}z$, and for all $r_1, r_2 \in [0, \tau]$ we put

$$\hat{f}_{r_1,r_2}(y) = f(a_t u(r_1)a_{-t}.y)f(a_t u(r_2)a_{-t}.y).$$

Note that $\mathcal{S}(\hat{f}_{r_1,r_2}) \ll \mathcal{S}(f)^2(e^t\tau)^\star \ll \mathcal{S}(f)^2e^{\star t/\ell}$. We now choose $\ell \ll 1/\kappa_4$ large enough so that

$$\mathcal{S}(\hat{f}_{r_1,r_2}) \ll \mathcal{S}(f)^2e^{\kappa_4 t/2}. \tag{4.12}$$

By Proposition 4.1, we have

$$\left| b^{-\varepsilon} \int_{B(0,1)} \hat{f}_{r_1,r_2}(a_t n(z)x)\, \mathrm{d}z \right| = b^{-\varepsilon} \int_X \hat{f}_{r_1,r_2}\, \mathrm{d}m_X \\ + b^{-\varepsilon} \eta^{-1/\kappa_5} O(\mathcal{S}(\hat{f}_{r_1,r_2}) e^{-\kappa_4 t}).$$

Recall from (4.12) that $\mathcal{S}(\hat{f}_{r_1,r_2})e^{-\kappa_4 t} \leq \mathcal{S}(f)^2 e^{-\kappa_4 t/2}$. Moreover, since $t \geq |\log b|/4$ if we assume $\varepsilon \leq \kappa_4/16$, then $e^{-\kappa_4 t/2} b^{-\varepsilon} \leq b^{\kappa_4/16}$. Altogether, we conclude that

$$\left| b^{-\varepsilon} \int_{B(0,1)} \hat{f}_{r_1,r_2}(a_t n(z)x)\, \mathrm{d}z \right| = b^{-\varepsilon} \int_X \hat{f}_{r_1,r_2}\, \mathrm{d}m_X \\ + \mathcal{S}(f)^2 \eta^{-1/\kappa_5} b^{\kappa_4/16}. \tag{4.13}$$

We now use estimates on the decay of matrix coefficients, (4.1), together with the fact that $d(e, u_t) \geq |t|$, and obtain the following bound.

$$\left| \int_X \hat{f}_{r_1,r_2}(x)\, \mathrm{d}m_X \right| \ll \mathcal{S}(f)^2 e^{-\frac{\kappa_X}{2\ell} t} \quad \text{if } |r_1 - r_2| > e^{-t + \frac{t}{2\ell}}. \tag{4.14}$$

Divide now the integral $\int_0^\tau \int_0^\tau$ in (4.11) into terms: one with $|r_1 - r_2| > e^{-t+\frac{t}{2\ell}} = \tau e^{-\frac{t}{2\ell}}$ and the other its complement. We thus get from (4.11), (4.13), and (4.14) that

$$|A_2|^2 \ll C \eta^{-\frac{1}{\kappa_5}} \mathcal{S}(f)^2 \left(b^{-\varepsilon} \left(e^{\frac{-\kappa_X}{2\ell} t} + e^{\frac{-1}{2\ell} t} \right) + b^{\kappa_4/16} \right).$$

Recall that $\ell \ll 1/\kappa_4$ and $\kappa_4 \gg \kappa_X$. Thus if $\varepsilon \leq \kappa_4^2/L$ for a large enough L, the above, together with (4.8), (4.9), and (4.10), finishes the proof. □

5 A Marstrand type projection theorem

In this section, we combine a certain projection theorem with some arguments in homogeneous dynamics to prove Proposition 5.1. The outcome of this proposition will serve as an input when we apply Proposition 4.2.

Proposition 5.1 *Let* $0 < \eta < 0.01\eta_X$, *and let* $0 < 100\varepsilon < \alpha < 1$. *Suppose there exist* $x_1 \in X_\eta$ *and* $F \subset B_{\mathfrak{r}}(0, \eta^2)$, *containing* 0, *so that*

$$\mathcal{F} := \{\exp(w)x_1 : w \in F\} \subset X_\eta \qquad \textit{and} \\ \sum_{w' \in F \setminus \{w\}} \|w - w'\|^{-\alpha} \leq D \cdot (\#F)^{1+\varepsilon} \qquad \textit{for all } w \in F, \tag{5.1}$$

for some $D \geq 1$.

Assume further that $\#F$ is large enough, depending explicitly on η and ε. Then exists a finite subset $I \subset [0, 1]$, some $b_1 > 0$ with

$$(\#F)^{-\frac{3-\alpha+5\varepsilon}{3-\alpha+20\varepsilon}} \leq b_1 \leq (\#F)^{-\varepsilon}, \tag{5.2}$$

and some $x_2 \in X_\eta \cap \Big(a_{|\log(b_1)|} \cdot \{u_r : |r| \leq 2\}\Big).\mathcal{F}$ so that both of the following statements hold true.

(1) The set I supports a probability measure ρ which satisfies

$$\rho(J) \leq C'_\varepsilon \cdot |J|^{\alpha-30\varepsilon}$$

for all intervals J with $|J| \geq (\#F)^{\frac{-15\varepsilon}{3-\alpha+20\varepsilon}}$, where $C'_\varepsilon \ll \varepsilon^{-\star}$ (with absolute implied constants).

(2) There is an absolute constant C, so that for all $s \in I$, we have

$$v_s x_2 \in \Big(\mathsf{B}^G_{Cb_1} \cdot a_{|\log(b_1)|} \cdot \{u_r : |r| \leq 2\}\Big).\mathcal{F}.$$

The proof of Proposition 5.1 is based on the following projection theorem. This theorem may be thought of as a finitary version of the work of Käenmäki, Orponen, and Venieri, [33]. Its proof, which is given in Appendix B, is based on the works of Wolff and Schlag, [56,65] which in turn relies on a cell decomposition theorem of Clarkson, Edelsbrunner, Guibas, Sharir, and Welzl [12].

Theorem 5.1 *Let $0 < \alpha, b_0, b_1 < 1$ (α should be thought of fixed, and $b_0 < b_1$ as small). Let $E \subset B_{\mathfrak{r}}(0, b_1)$ be so that*

$$\tfrac{\#(E \cap B_{\mathfrak{r}}(w,b))}{\#E} \leq D' \cdot (b/b_1)^\alpha$$

for all $w \in \mathfrak{r}$ and all $b \geq b_0$, and some $D' \geq 1$. Let $0 < \kappa < 0.1$, and let $J \subset \mathbb{R}$ be an interval. There exists $J' \subset J$ with $|J'| \geq 0.9|J|$ satisfying the following. Let $r \in J'$, then there exists a subset $E_r \subset E$ with

$$\#E_r \geq 0.9 \cdot (\#E)$$

such that for all $w \in E_r$ and all $b \geq b_0$, we have

$$\tfrac{\#\{w' \in E : |\xi_r(w') - \xi_r(w)| \leq b\}}{\#E} \leq C_\kappa \cdot (b/b_1)^{\alpha-7\kappa},$$

where C_κ is a constant which depends polynomially on κ, $|J|$, and D', and

$$\xi_r(w) = (\mathrm{Ad}(u_r)w)_{12} = -w_{21}r^2 - 2w_{11}r + w_{12}. \tag{5.3}$$

with w_{ij} denoting the (i, j)-th entry of $w \in \mathfrak{r}$.

The proof of Proposition 5.1 will also use the following version of [6, Lemma 5.2], see also [5]. We reproduce the argument in Appendix C.

Lemma 5.1 *Let $F \subset B_{\mathfrak{r}}(0, 1)$ be a subset which satisfies (5.1). Then there exist $w_0 \in F$, $b_1 > 0$, with*

$$(\#F)^{-\frac{3-\alpha+5\varepsilon}{3-\alpha+20\varepsilon}} \le b_1 \le (\#F)^{-\varepsilon},$$

and a subset $F' \subset B_{\mathfrak{r}}(w_0, b_1) \cap F$ so that the following holds. Let $w \in \mathfrak{r}$, and let $b \ge (\#F)^{-1}$. Then

$$\tfrac{\#(F' \cap B(w,b))}{\#F'} \le C' \cdot (b/b_1)^{\alpha - 20\varepsilon},$$

where $C' \ll_D \varepsilon^{-\star}$ with absolute implied constants.

We now begin the proof of the proposition.

Proof of Proposition 5.1 The general strategy is straightforward. First we apply Lemma 5.1 to replace the set F with a local version of it, i.e., we replace F with $F' \subset B_{\mathfrak{r}}(w_0, b_1) \cap F$. Then using Theorem 5.1, we project the *discretized dimension* in $\mathfrak{r}$ to the direction of $\mathrm{Lie}(V) = \mathfrak{r} \cap \mathrm{Lie}(N)$. Finally, we use the action of A to expand this subset of V to size 1.

The details however are a bit more involved, in particular, we need to carefully control the size of various elements; we also need to use Proposition 3.1 (when X is not compact) to ensure returns to X_η.

Throughout the proof, we will assume $\#F$ is large enough so that

$$(\#F)^{-\varepsilon} \le (2C_5C_7)^{-1}\eta^3, \tag{5.4}$$

see Lemma 2.1 and Proposition 3.1. □

Localizing the entropy

Apply Lemma 5.1 with F as in the proposition. Let $w_0 \in F$, $b_1 > 0$, and $F' \subset B_{\mathfrak{r}}(w_0, b_1) \cap F$ be given by that lemma; in particular, we have

$$(\#F)^{-\frac{3-\alpha+5\varepsilon}{3-\alpha+20\varepsilon}} \le b_1 \le (\#F)^{-\varepsilon}. \tag{5.5}$$

Replacing w_0 with a different point in F and increasing C' if necessary, we will assume that $F' \subset B_{\mathfrak{r}}(w_0, b_1/(6C_5)) \cap F$. In view of Lemma 2.1, for all $w' \in F'$, there exist $h \in H$ and $w \in \mathfrak{r}$ so that

$$\begin{aligned} &h \exp(w) = \exp(w') \exp(-w_0) \\ &\|h - I\| \le b_1^2/3 \quad \text{and} \quad \|w\| \le 2\|w_0 - w'\| \le b_1/(3C_5). \end{aligned} \tag{5.6}$$

Set

$$E = \Big\{ w \in \mathfrak{r} : \exists h \in H,\, w' \in F' \text{ so that } h, w, w_0, w' \text{ satisfy (5.6)} \Big\}. \tag{5.7}$$

Lemma 5.2 *Let the notation be as above. Then*

$$\tfrac{\#(E \cap B(w,b))}{\#E} \le \hat{C} \cdot (b/b_1)^{\alpha - 20\varepsilon} \tag{5.8}$$

for all $w \in \mathfrak{r}$ and $b \ge (\#F)^{-1}$ where $\hat{C} \le 2C'$.

This lemma is proved after the completion of the proof of the proposition. Let $x_2' := \exp(w_0)x_1$, and let $w' \in F'$. Then if h and w are as in (5.6),

$$h \exp(w) x_2' = \exp(w') \exp(-w_0) \exp(w_0) x_1 = \exp(w') x_1 \in \mathcal{F}. \tag{5.9}$$

We also need the following elementary lemma whose proof will be given after the completion of the proof of the proposition.

Lemma 5.3 *There exists $r_0 \in [0, 1]$ and a subset*

$$\bar{E} \subset \mathrm{Ad}(u_{r_0}) E \cap \Big\{ w \in B_{\mathfrak{r}}(0, \eta) : |w_{12}| \ge 10^{-3}\|w\| \Big\}$$

so that $\#\bar{E} \ge \#E/4$.

Thanks to Lemma 5.3, we may replace x_2' by $u_{r_0} x_2'$ for some $r_0 \in [0, 1]$ and E by a subset $\bar{E}$ with $\#\bar{E} \ge \#E/4$ (which we continue to denote by E), to ensure that

$$E \subset \Big\{ w \in B_{\mathfrak{r}}(0, \eta) : |w_{12}| \ge 10^{-3}\|w\| \Big\}, \tag{5.10}$$

where w_{12} denotes the $(1, 2)$-th entry of $w \in \mathfrak{r}$, see (5.3). Note that (5.8) holds for the new E with $4\hat{C}$, we suppress the factor 4.

Estimates on the size of elements

Let $t = |\log(b_1)|$. By (5.9), for all $r \in [0, 1]$, we have

$$a_t u_r h \exp(w).x_2' \in a_t \cdot \{u_r : r \in [0, 1]\}.\mathcal{F}, \tag{5.11}$$

where $w \in E$, i.e, $h \exp(w) = \exp(w') \exp(-w_0)$.

We now investigate properties of the element $a_t u_r h \exp(w) u_{-r} a_{-t}$. In view of (5.6) and the definition of t, for all $r \in [0, 1]$, we have

$$\|\operatorname{Ad}(a_t u_r) w\| \leq 1, \quad \text{and} \tag{5.12a}$$

$$\|a_t u_r h u_{-r} a_{-t} - I\| \leq b_1; \tag{5.12b}$$

note, moreover, that $a_t u_r h u_{-r} a_{-t} \in H$.

In view of (5.10), for all $|r| \leq 10^{-4}$ we have

$$|(\operatorname{Ad}(u_r) w)_{12}| \geq 10^{-4} \|w\|.$$

Therefore, for all $|r| \leq 10^{-4}$, we have

$$\operatorname{Ad}(a_t u_r) w = \begin{pmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{pmatrix}$$

where $|v_{11}|, |v_{22}| \leq 10^4 e^{-t} |v_{12}|$ and $|v_{21}| \leq 10^4 e^{-2t} |v_{12}|$. Hence for $|r| \leq 10^{-4}$, we have

$$a_t u_r h \exp(w).x_2' = (a_t u_r h u_{-r} a_{-t}) \cdot g \cdot \exp\Big(e^t (\operatorname{Ad}(a_t u_r) w)_{12} E_{12}\Big).a_t u_r x_2';$$

for some $g \in G$ which in view of the estimate in (5.12a) satisfies

$$\|g - I\| \ll b_1 \tag{5.13}$$

with an absolute implied constant.

Using (5.11) and (5.12b), we conclude that

$$\begin{aligned} &\exp\Big(e^t (\operatorname{Ad}(a_t u_r) w)_{12} E_{12}\Big).a_t u_r x_2' \\ &\quad \in \Big(\mathsf{B}^G_{Cb_1} \cdot \mathsf{B}^H_{Cb_1} \cdot a_t \cdot \{u_r : r \in [0, 1]\}\Big).\mathcal{F}, \end{aligned} \tag{5.14}$$

where C is an absolute constant.

Applying Theorem 5.1

We now choose a particular $|r| \leq 10^{-4}$ in order to the define the set I in Proposition 5.1. This choice is based on Proposition 3.1 and Theorem 5.1.

Recall that $t = |\log(b_1)|$ and

$$b_1 \leq (\#F)^{-\varepsilon} \leq (2C_5C_7)^{-1}\eta^3. \tag{5.15}$$

Apply Proposition 3.1 with t, $x_2' = \exp(w_0)x_1 \in X_\eta$, and the interval $J = [-10^{-4}, 10^{-4}]$. Then if we set

$$J'' = \{r : |r| \leq 10^{-4}, a_t u_r.x_2' \in X_\eta\} \tag{5.16}$$

by the proposition $|J''| > 0.9 \cdot 2 \cdot 10^{-4}$.

We also apply Theorem 5.1 with E, $J = [-10^{-4}, 10^{-4}]$, $\alpha - 20\varepsilon$, and $\kappa = \varepsilon$. Let J' be given by that Theorem. Fix some $r \in J' \cap J''$ for the remainder of the argument.

Put $x_2 := a_t u_r.x_2'$. By definition of J'' in (5.16), $x_2 \in X_\eta$, and by (5.14)

$$\exp\Big(e^t(\mathrm{Ad}(u_r w)_{12})\Big).x_2 \in \Big(\mathsf{B}^G_{Cb_1} \cdot \mathsf{B}^H_{Cb_1} \cdot a_t \cdot \{u_r : r \in [0,1]\}\Big).\mathcal{F}. \tag{5.17}$$

In the notation of Theorem 5.1, put

$$I := \{e^t \xi_r(w) : w \in E_r\};$$

recall that $\xi_r(w) = (\mathrm{Ad}(a_\tau \mathsf{r}_\theta)w)_{12}$. We will show that the proposition holds with x_2, I, and b_1. First note that the claimed bound (5.2) on b_1 in the statement of the proposition holds in view of (5.5). The assertion in part (2) of the proposition also holds by (5.17).

Thus it only remains to establish (1) of the proposition. Let ρ be the pushforward of the normalized counting measure on E_r under the map $w \mapsto e^t\xi_r(w)$. That is,

$$\rho(K) = \tfrac{\#\{w \in E_r \,:\, e^t\xi_r(w) \in K\}}{\#E_r}$$

for any interval $K \subset \mathbb{R}$.

Recall again that $e^{-t} = b_1$. Let $w \in E_r$, and put $s = e^t\xi_r(w)$. By Theorem 5.1, and in view of the fact that $\#E_r \geq 0.9 \cdot (\#E)$, for every $b \geq e^t \cdot (\#F)^{-1}$, we have that

$$\rho\Big(\{s'\in I : |s-s'|\le b\}\Big) = \frac{\#\Big\{w'\in E_r : |\xi_r(w')-\xi_r(w)|\le e^{-t}b\Big\}}{\#E_r}$$
$$\le \bar{C}_\varepsilon\cdot(e^{-t}b/b_1)^{\alpha-27\varepsilon} = \bar{C}_\varepsilon b^{\alpha-27\varepsilon}, \tag{5.18}$$

where $\bar{C}_\varepsilon \ll \varepsilon^{-\star}$.

Using the estimate in (5.5), we have

$$e^t\cdot(\#F)^{-1}\le(\#F)^{\frac{-15\varepsilon}{3-\alpha+20\varepsilon}};$$

this estimate and (5.18) finish the proof of part (1). □

Proof of Lemma 5.2 Let $\bar\eta\le 0.01$, and let $w_0\in B_{\mathfrak{r}}(0,\bar\eta)$. Define the map $f : B_{\mathfrak{r}}(0,\bar\eta)\to B_{\mathfrak{r}}(0,2\bar\eta)$ by $f(w')=w$ where

$$h\exp(w)=\exp(w')\exp(-w_0)\quad\text{with } h\in \mathsf{B}^H_{2C_5\bar\eta^2}\text{ and } w\in B_{\mathfrak{r}}(0,2\bar\eta).$$

By the Baker–Campel–Hausdorff formula, see Lemma 2.1, f is a diffeomorphism. Moreover, we have

$$\Big\|\mathrm{D}_{w'}\Big(f^{\pm1}\Big)-I\Big\|\le 0.1$$

for all $w'\in B_{\mathfrak{r}}(0,\bar\eta)$, in particular, $\mathrm{D}_{w'}(f^{\pm1})$ is invertible for all $w'\in B_{\mathfrak{r}}(0,\bar\eta)$.

We conclude that $\#f(E)=\#E$, and

$$\#\Big(B_{\mathfrak{r}}(\bar w,b)\cap f(E)\Big)\le\#\Big(B_{\mathfrak{r}}(f^{-1}(\bar w),2b)\cap E\Big)$$

for all $b\le\bar\eta$. The claim follows. □

Proof of Lemma 5.3 This is a consequence of the fact that the adjoint action of H on $\mathfrak{r}$ is irreducible; the argument below is based on explicit computations.

Recall that $\|w\|=\max\{|w_{12}|,|w_{21}|,|w|_{21}\}$; moreover, recall that

$$\Big(\mathrm{Ad}(u_r)w\Big)_{12}=-w_{21}r^2-2w_{11}r+w_{12}. \tag{5.19}$$

Now if

$$\#\{w\in E : |w_{12}|\ge 0.001\|w\|\}\ge \#E/4,$$

then the claim holds with $r_0=0$.

Therefore, we assume $\#\hat{E} \geq \frac{3\cdot(\#E)}{4}$ where $\hat{E} = \{w \in E : |w_{12}| \leq 0.001\|w\|\}$. If

$$\#\{w \in \hat{E} : |w_{11}| \geq 0.1\|w\|\} \geq \#E/4,$$

then the claim holds with $r_0 = 0.1$ and the set on the left side of the above.

Therefore, we may assume

$$\#\{w \in \hat{E} : |w_{11}| \leq 0.1\|w\|\} \geq \#E/2.$$

For every w in the set on the left side of the above, $\|w\| = |w_{21}|$. The claim now holds with $r_0 = 0.9$ and the set on the left side of the above. $\square$

6 A closing lemma

For the proof of Theorem 1.1, one needs to guarantee that a certain initial separation is satisfied. This is the task in this section. This initial separation estimate is then bootstrapped in Sect. 7 to give a better (finitary) dimension estimate that is used to conclude the theorem. Throughout this section, Γ *is assumed to be arithmetic*. Indeed, this section is the only place where arithmeticity of Γ is used in this paper, more specifically Lemma 6.1. Superficially arithmeticity is also used Lemma 6.2, but there the usage of arithmeticity is rather mild — by local rigidity a lattice Γ in $\mathrm{SL}(2, \mathbb{C})$ or an irreducible lattice in $\mathrm{SL}(2, \mathbb{R}) \times \mathrm{SL}(2, \mathbb{R})$ can be conjugated to have algebraic entries in some number field, which is good enough for our (relatively coarse) purposes.

Recall from (2.9) the definition

$$\mathsf{E}_{\eta,t,\beta} = \mathsf{B}^H_\beta \cdot a_t \cdot \big\{u_r : r \in [0, \eta]\big\} \subset H;$$

recall also that we always assume $e^{-0.01t} < \beta < 1$, and in this section we will be mainly interested in the case $\eta = 1$; to simplify the notation, we will write E_t for $\mathsf{E}_{1,t,\beta}$.

Let $x \in X$ and $t > 0$. For every $z \in \mathsf{E}_t.x$, put

$$I_t(z) := \Big\{w \in \mathfrak{r} : 0 < \|w\| < \mathrm{inj}(z),\ \exp(w)z \in \mathsf{E}_t.x\Big\}. \tag{6.1}$$

Note that this is a finite subset of $\mathfrak{r}$. In (7.3), we will define $I_{\mathcal{E}}(h, z)$ for all $h \in H$ and more general sets $\mathcal{E}$.

Let $0 < \alpha < 1$. Define the function $f_{t,\alpha} : \mathsf{E}_t.x \to [2, \infty)$ (which we will later use as a Margulis function in the bootstrap phase of the proof) as follows

$$f_{t,\alpha}(z) = \begin{cases} \sum_{w \in I_t(z)} \|w\|^{-\alpha} & \text{if } I_t(z) \neq \emptyset \\ \mathrm{inj}(z)^{-\alpha} & \text{otherwise} \end{cases}.$$

The following is the main result of this section.

Proposition 6.1 *There exists D_0 (which depends explicitly on Γ) satisfying the following. Let $D \geq D_0 + 1$, and let $x_0 \in X$. Then for all large enough t (depending explicitly on* $\mathrm{inj}(x_0)$ *and X) at least one of the following holds.*

(1) There is some $x \in X_{\mathrm{cpt}} \cap \{a_{8t} u_r.x_0 : r \in [0, 1]\}$ such that
  (a) $\mathsf{h} \mapsto \mathsf{h}x$ is injective over E_t.
  (b) For all $z \in \mathsf{E}_t.x$, we have

$$f_{t,\alpha}(z) \leq e^{Dt}$$

  for all $0 < \alpha < 1$.

(2) There is $x' \in X$ such that $H.x'$ is periodic with

$$\mathrm{vol}(H.x') \leq e^{D_0 t} \quad \textit{and} \quad d_X(x', x_0) \leq e^{(-D+D_0)t}.$$

The proof we give here is similar to that of Margulis and the first named author in [41, Lemma 5.2]. A certain Diophantine condition (namely, *inheritable boundedness condition*) is used in the formulation of loc. cit. to guarantee in particular that our initial point is not close to a periodic U orbit. We do not need such a condition here since we consider essentially translations of local U orbits by expanding elements in A, and not long orbits of U (this is reminiscent of a result of Nimish Shah [58, Thm. 1.1]). As in [41] the argument is elementary; a result of similar spirit to our Proposition 6.1 is proved by Einsiedler, Margulis, and Venkatesh in [18, Prop. 13.1] using property-τ, i.e. a uniform spectral gap.

Let us begin with some preliminary statements. In Proposition 6.1, we are allowed to choose t large depending on Γ. Therefore, by passing to a finite index subgroup, we will assume that both of the following hold: Γ *is torsion free* and if $\Gamma \subset \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ is reducible, then $\Gamma = \Gamma_1 \times \Gamma_2$

It is more convenient to consider G as the set of $\mathbb{R}$-points of an algebraic group defined over $\mathbb{R}$ — this way H can be realized of as an algebraic subgroup of G. To that end, we let $\mathbf{G} = \mathrm{SL}_2 \times \mathrm{SL}_2$ if $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$. If $G = \mathrm{SL}_2(\mathbb{C})$, we let $\mathbf{G} = \mathrm{Res}_{\mathbb{C}/\mathbb{R}}(\mathrm{SL}_2)$. In either case, $\mathbf{G}$ is defined over $\mathbb{R}$ and $G = \mathbf{G}(\mathbb{R})$.

Recall that Γ is assumed to be arithmetic. Therefore, there exists a semisimple $\mathbb{Q}$-group $\tilde{\mathbf{G}} \subset \mathrm{SL}_M$, for some M, and an epimorphism $\rho : \tilde{\mathbf{G}}(\mathbb{R}) \to \mathbf{G}(\mathbb{R}) = G$ of $\mathbb{R}$-groups with compact kernel so that

$$\Gamma \text{ is commensurable with } \rho(\tilde{\mathbf{G}}(\mathbb{Z})), \tag{6.2}$$

where $\tilde{\mathbf{G}}(\mathbb{Z}) = \tilde{\mathbf{G}}(\mathbb{R}) \cap \mathrm{SL}_M(\mathbb{Z})$. Note that $\tilde{\mathbf{G}}$ can be chosen to be $\mathbb{Q}$-almost simple unless $\Gamma \subset \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ is a reducible lattice, in which case $\tilde{\mathbf{G}}$ can be chosen to have two $\mathbb{Q}$-almost simple factors.

Let $\tilde{\mathfrak{g}} = \mathrm{Lie}(\tilde{\mathbf{G}}(\mathbb{R}))$, this Lie algebra has a natural $\mathbb{Q}$-structure. Moreover, $\tilde{\mathfrak{g}}_{\mathbb{Z}} := \tilde{\mathfrak{g}} \cap \mathfrak{sl}_M(\mathbb{Z})$ is a $\tilde{\mathbf{G}}(\mathbb{Z})$-stable lattice in $\tilde{\mathfrak{g}}$.

We continue to write $\mathrm{Lie}(G) = \mathfrak{g}$ and $\mathrm{Lie}(H) = \mathfrak{h}$; these are considered as 6-dimensional (resp. 3-dimensional) $\mathbb{R}$-vector spaces.

Let v_H be a unit vector on the line $\wedge^3\mathfrak{h}$. Note that

$$N_G(H) = \{g \in G : g v_H = v_H\}$$

which contains H as a subgroup of index two.

Recall also that we fixed a compact subset $\mathfrak{S}_{\mathrm{cpt}} \subset G$ which projects onto X_{cpt}, see Sect. 3.1 for the notation.

Lemma 6.1 *There exist C_{11} and κ_7 depending on M and $\mathfrak{S}_{\mathrm{cpt}}$, so that the following holds. Let $\gamma_1, \gamma_2 \in \Gamma$ be two non-commuting elements. If $g \in \mathfrak{S}_{\mathrm{cpt}}$ is so that $\gamma_i g^{-1} v_H = g^{-1} v_H$ for $i = 1, 2$, then $Hg\Gamma$ is a closed orbit with*

$$\mathrm{vol}(Hg\Gamma) \le C_{11}\Big(\max\{\|\gamma_1^{\pm 1}\|, \|\gamma_2^{\pm 1}\|\}\Big)^{\kappa_7}.$$

Proof In view of our assumption in the lemma, we have

$$\langle \gamma_1, \gamma_2 \rangle \subset \mathrm{Stab}_G(g^{-1} v_H) = N_G(g^{-1} H g).$$

Let $\Lambda_1 := \langle g\gamma_1 g^{-1}, g\gamma_2 g^{-1} \rangle$. We claim that $\Lambda := \Lambda_1 \cap H$ is Zariski dense in H. Indeed since $\langle \gamma_1, \gamma_2 \rangle$ is a torsion free, non-commutative, discrete subgroup of $N_G(g^{-1} H g)$, we have Λ is discrete and torsion free. This and the fact that $H \simeq \mathrm{SL}_2(\mathbb{R})$ imply that if Λ is non-commutative, then it is Zariski dense in H. Assume thus that Λ is commutative, which implies that $\Lambda \simeq \mathbb{Z}$ and that $\Lambda \subsetneq \Lambda_1$ (recall that Λ_1 is non-commutative). Since $N_G(H) = HC$ where C is the center of G if $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ and $C = \langle \mathrm{diag}(i, -i) \rangle$ if $G = \mathrm{SL}_2(\mathbb{C})$, we have $N_G(H)/H \simeq \mathbb{Z}/2\mathbb{Z}$; thus $\Lambda_1/\Lambda \simeq \mathbb{Z}/2\mathbb{Z}$. This implies that Λ_1 is isomorphic to $\mathbb{Z}$ or $\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ or $\mathbb{Z}/2\mathbb{Z} \ltimes \mathbb{Z}$. Either possibility leads to a contradiction to Λ_1 being non-commutative and torsion free.

Let $\mathbf{L}$ be the Zariski closure of $\langle\gamma_1,\gamma_2\rangle$. In view of the above discussion,

$$g^{-1}Hg \subset \mathbf{L}(\mathbb{R}) \subset N_G(g^{-1}Hg). \tag{6.3}$$

Since $N_G(H)/H \simeq \mathbb{Z}/2\mathbb{Z}$, replacing γ_i by γ_i^2 if necessary we assume that $\mathbf{L}(\mathbb{R}) = g^{-1}Hg$.

Let $\tilde\gamma_i \in \tilde{\mathbf{G}}(\mathbb{Z})$ be so that $\rho(\tilde\gamma_i) = \gamma_i$. Then the Zariski closure $\tilde{\mathbf{L}}$ of $\langle\tilde\gamma_1,\tilde\gamma_2\rangle$ is semisimple and $\rho(\tilde{\mathbf{L}}(\mathbb{R})) = \mathbf{L}(\mathbb{R})$. Therefore, in view of a theorem of Borel and Harish-Chandra [4, Thm. 7.8], we have $\tilde{\mathbf{L}}(\mathbb{R}) \cap \tilde{\mathbf{G}}(\mathbb{Z})$ is a lattice in $\tilde{\mathbf{L}}(\mathbb{R})$.

This implies that $\mathbf{L}(\mathbb{R})\Gamma$ is a periodic orbit, which in view of (6.3) implies that $Hg\Gamma$ is a periodic orbit.

We now turn to the proof of the second claim. Let $\tilde{\mathfrak{l}} = \mathrm{Lie}(\tilde{\mathbf{L}}(\mathbb{R})) \subset \tilde{\mathfrak{g}}$. Then $\tilde{\mathfrak{l}}$ is a rational subspace of $\tilde{\mathfrak{g}}$; we will show that the height of this subspace is $\ll \Theta^\star$ where $\Theta := \max\{\|\gamma_1^{\pm1}\|, \|\gamma_2^{\pm1}\|\}$. That is to say: $\tilde{\mathfrak{l}}$ has a basis consisting of vectors in $\tilde{\mathfrak{g}}_{\mathbb{Z}} \cap \tilde{\mathfrak{l}}$ with norm $\ll \Theta^\star$, e.g., by Minkowski's second theorem.

Indeed by Chevalley's theorem and the fact that $\tilde{\mathbf{L}}(\mathbb{R})$ is semisimple (hence it has no character), there exists a finite dimensional $\mathbb{Q}$-representation of $\tilde{\mathbf{G}}$ on a space Φ with the following property. Let Φ^0 denote the vectors in $\Phi_{\mathbb{R}}$ which are fixed by $\tilde{\mathbf{L}}(\mathbb{R})$, then

$$\tilde{\mathbf{L}}(\mathbb{R}) = \{g \in \tilde{\mathbf{G}}(\mathbb{R}) : g.q = q, \text{ for all } q \in \Phi^0\};$$

in terms of the Lie algebras, this is $\tilde{\mathfrak{l}} = \{w \in \tilde{\mathfrak{g}} : w.\Phi^0 = 0\}$.

Since $\langle\tilde\gamma_1,\tilde\gamma_2\rangle$ is Zariski dense in $\tilde{\mathbf{L}}$, we conclude that Φ^0 is a rational subspace with height $\ll (\max\{\|\tilde\gamma_1^{\pm1}\|, \|\tilde\gamma_2^{\pm1}\|\})^\star \ll \Theta^\star$; we used the fact that $\rho(\tilde\gamma_i) = \gamma_i$ to bound $\|\tilde\gamma_i^{\pm1}\|$ from above by $\|\gamma_i^{\pm1}\|^\star$ for $i = 1, 2$.

Using this and the fact that $\tilde{\mathfrak{l}} = \{w \in \tilde{\mathfrak{g}} : w.\Phi^0 = 0\}$, we conclude that height of $\tilde{\mathfrak{l}}$ is $\ll \Theta^\star$ as we claimed. This height bound implies that

$$\mathrm{vol}\Big(\tilde{\mathbf{L}}(\mathbb{R})\tilde{\mathbf{G}}(\mathbb{Z})\Big) \ll \Theta^\star.$$

see e.g. [18, §17], or [17, App. B] (see also [19, §2], which treats the case of tori; the proof there works for the semisimple case as well).

We deduce that $\mathrm{vol}(\mathbf{L}(\mathbb{R})\Gamma) \ll \Theta^\star$; recall that the kernel of ρ is compact and $\mathbf{L}(\mathbb{R}) = \rho(\tilde{\mathbf{L}}(\mathbb{R}))$. The claimed bound on $\mathrm{vol}(Hg\Gamma)$ now follows in view of (6.3) and the fact that $g \in \mathfrak{S}_{\mathrm{cpt}}$. □

We also need the following lemma.

Lemma 6.2 *There exist κ_8, κ_9, and C_{12} so that the following holds. Let $\gamma_1, \gamma_2 \in \Gamma$ be two non-commuting elements, and let*

$$\delta \leq C_{12}^{-1}\Big(\max\{\|\gamma_1^{\pm 1}\|, \|\gamma_2^{\pm 1}\|\}\Big)^{-\kappa_8}.$$

Suppose there exists some $g \in \mathfrak{S}_{\mathrm{cpt}}$ so that $\gamma_i g^{-1} v_H = \epsilon_i g^{-1} v_H$ for $i = 1, 2$ where $\|\epsilon_i - I\| \leq \delta$. Then, there is some $g' \in G$ such that

$$\|g' - g^{-1}\| \leq C_{12}\delta\Big(\max\{\|\gamma_1^{\pm 1}\|, \|\gamma_2^{\pm 1}\|\}\Big)^{\kappa_9}$$

and $\gamma_i g' v_H = g' v_H$ for $i = 1, 2$.

Proof This is essentially proved in [18, §13.3, §13.4], we recall parts of the argument for the convenience of the reader.

With a slight change in the notation from the proof of the previous lemma, let $\tilde{\mathbf{L}}$ be the $\mathbb{R}$-group defined by $\tilde{\mathbf{L}}(\mathbb{R}) = \rho^{-1}(g^{-1}Hg) \subset \tilde{\mathbf{G}}(\mathbb{R})$, and let $d = \dim(\tilde{\mathbf{L}}(\mathbb{R}))$. Fix a unit vector v_0 on the line $\wedge^d(\mathrm{Lie}(\tilde{\mathbf{L}}(\mathbb{R})))$.

Let also $\tilde{\gamma}_i \in \tilde{\mathbf{G}}(\mathbb{Z})$ be so that $\rho(\tilde{\gamma}_i) = \gamma_i$, for $i = 1, 2$. Then [18, Lemma 13.1] holds true for linear transformation

$$A = (\tilde{\gamma}_1 - I) \oplus (\tilde{\gamma}_2 - I)$$

from $\wedge^d \tilde{\mathfrak{g}}$ to $\wedge^d \tilde{\mathfrak{g}} \oplus \wedge^d \tilde{\mathfrak{g}}$. Therefore, there exists a vector $w \in \wedge^d \tilde{\mathfrak{g}}$, with

$$\|w - v_0\| \leq C\Theta^{\kappa}\delta \tag{6.4}$$

so that $Aw = 0$, where $\Theta := \max\{\|\gamma_1^{\pm 1}\|, \|\gamma_2^{\pm 1}\|\}$, C depends on $\tilde{\mathbf{G}}$ and κ depends on $\dim \tilde{\mathbf{G}}$. We again used $\rho(\tilde{\gamma}_i) = \gamma_i$ to bound $\|\tilde{\gamma}_i^{\pm 1}\|$ by a power of $\|\gamma_i^{\pm 1}\|$.

This implies that $\tilde{\gamma}_i w = w$ for $i = 1, 2$. By [18, Lemma 13.2], there exist $\bar{C}$ and $\bar{\kappa} \geq 1$ so that if

$$\|w - v_0\| \leq \bar{C}^{-1}\Theta^{-\bar{\kappa}},$$

then there exists $\tilde{g} \in \tilde{\mathbf{G}}(\mathbb{R})$ satisfying that $\|\tilde{g} - I\| \leq C'\|w - v_0\|$ and

$$\tilde{\gamma}_i \tilde{g} v_0 = \tilde{g} v_0 \text{ for } i = 1, 2,$$

see [43] for sharper results concerning equivariant projections.

Let now δ satisfy

$$0 < \delta \leq (C\bar{C})^{-1}\Theta^{-\kappa' - \kappa}.$$

Then (6.4) implies that there exists some $\tilde{g} \in \tilde{\mathbf{G}}(\mathbb{R})$ with $\|\tilde{g} - I\| \le C' C \Theta^\kappa \delta$ so that $\tilde{\gamma}_i \tilde{g} \tilde{v}_0 = \tilde{g} v_0$ for $i = 1, 2$. This estimate implies that

$$\|\rho(\tilde{g}) g^{-1} - g^{-1}\| \le C'' \Theta^\kappa \delta$$

for some C'' depending on $\tilde{\mathbf{G}}$.

Let $g' = \rho(\tilde{g}) g^{-1}$. Then $\gamma_i g' v_H = g' v_H$ and the claim holds for $g' v_H$. □

We need the following lemma, see Lemma 7.2 in the sequel for a more general statement.

Lemma 6.3 *Let $x \in X_{\mathrm{cpt}}$. Then for every $z \in \mathsf{E}_t.x$, we have*

$$\# I_t(z) \ll e^{4t}.$$

For the convenience of the reader, we recall from (6.1) that

$$I_t(z) := \Big\{ w \in \mathfrak{r} : 0 < \|w\| < \mathrm{inj}(z),\ \exp(w) z \in \mathsf{E}_t.x \Big\}.$$

Proof Recall from (2.5) that

$$\mathrm{inj}(z) = \min \Big\{ 0.01, \sup \Big\{ \varepsilon : g \mapsto gz \text{ is injective on } \mathsf{B}^G_{10\varepsilon} \Big\} \Big\},$$

where for every $0 < \varepsilon \le 0.1$, we put $\mathsf{B}^G_\varepsilon := \mathsf{B}^H_\varepsilon \cdot \exp(B_{\mathfrak{r}}(0, \varepsilon))$.

Note that since $x \in X_{\mathrm{cpt}}$, we have

$$\mathrm{inj}(\mathsf{h}x) > 10ce^{-t} \quad \text{for all } \mathsf{h} \in \mathsf{E}_t, \tag{6.5}$$

where c depends only on X.

Let $z \in \mathsf{E}_t.x$ and $w \in I_t(z)$ (hence $\exp(w) z \in \mathsf{E}_t.x$). Therefore,

$$\mathsf{B}^H_{ce^{-t}} \exp(w) z \subset \mathsf{E}_{t+}.x,$$

where we define $\mathsf{E}_{t+} = \mathsf{B}^H_{\beta + 2ce^{-t}} \cdot \mathsf{E}_t$.

In view of (6.5) and the definition of $\mathrm{inj}(z)$, the map $(\mathsf{h}, w) \mapsto \mathsf{h} \exp(w) z$ is injective over $\mathsf{B}^H_{ce^{-t}} \times \exp(B_{\mathfrak{r}}(0, \mathrm{inj}(z)))$. Hence we have

$$\mathsf{B}^H_{ce^{-t}} \exp(w) z \cap \mathsf{B}^H_{ce^{-t}} \exp(w') z = \emptyset \qquad \text{for all distinct } w, w' \in I_t(z).$$

Since $m_H(\mathsf{E}_{t+}) \ll e^t$ and $m_H(\mathsf{B}^H_{ce^{-t}}) \gg e^{-3t}$, the claim follows. □

Proof of Proposition 6.1 By Proposition 3.2 if $d \geq |\log(10^{-6}\,\mathrm{inj}(y))| + C_7$, then

$$|\{r \in J : a_d u_r y \in X_{\mathrm{cpt}}\}| \geq 0.99|J| \tag{6.6}$$

for all $J \subset [0, 1]$ with $|J| \geq 10^{-3}$.

Let $t \geq |\log(10^{-6}\,\mathrm{inj}(x_0))| + C_7$ for the rest of the argument. Let $r_0 \in [0, 1/2]$ be so that $x_1 = a_t u_{r_0} x_0$ satisfies both of the following: $x_1 \in X_{\mathrm{cpt}}$ and $a_{7t} x_1 \in X_{\mathrm{cpt}}$. Write $x_1 = g_1 \Gamma$ where $g_1 \in \mathfrak{S}_{\mathrm{cpt}}$.

We introduce the shorthand notation $h_r := a_{7t} u_r$, for any $r \in [0, 1]$. Note that for all $r \in [0, 1]$, we have $h_r x_1 \in \{a_{8t} u_{r'} x_0 : r' \in [0, 1]\}$. Assume now the claim in part (1) fails for all $r \in [0, 1]$ so that $h_r x_1 \in X_{\mathrm{cpt}}$. That is: for all $r \in [0, 1]$ so that $h_r x_1 \in X_{\mathrm{cpt}}$

- either there exists $z \in \mathsf{E}_t.h_r x_1$ so that $f_{t,\alpha}(z) > e^{Dt}$,
- or the map $\mathsf{h} \mapsto \mathsf{h} h_r x_1$ is not injective on E_t.

In what follows all the implied multiplicative constants depend only on X.

Finding lattice elements γ_r

Let us first investigate the former situation. That is: fix $r \in [0, 1]$ so that $h_r x_1 \in X_{\mathrm{cpt}}$ and suppose that for some $z = \mathsf{h}_1 h_r x_1 \in \mathsf{E}_t.h_r x_1$, it holds that $f_{t,\alpha}(z) > e^{Dt}$. Since $h_r x_1 \in X_{\mathrm{cpt}}$, we have

$$\mathrm{inj}(\mathsf{h} h_r x_1) \gg e^{-t}, \quad \text{for all } \mathsf{h} \in \mathsf{E}_t. \tag{6.7}$$

Using the definition of $f_{t,\alpha}$, thus, we conclude that if $I_t(z) = \emptyset$, then $f_{t,\alpha}(z) \ll e^t$. Hence, assuming t is large enough, $I_t(z) \neq \emptyset$; recall also from Lemma 6.3 that $\# I_t(z) \ll e^{4t}$.

Altogether, if $D \geq 5$ and t is large enough, there exists some $w \in I_t(z)$ with

$$0 < \|w\| \leq e^{(-D+5)t}.$$

The above implies that for some $w \in \mathfrak{r}$ with $\|w\| \leq e^{(-D+5)t}$ and $\mathsf{h}_1 \neq \mathsf{h}_2 \in \mathsf{E}_t$, we have $\exp(w)\mathsf{h}_1 h_r x_1 = \mathsf{h}_2 h_r x_1$. Thus

$$\exp(w_r) h_r^{-1} \mathsf{s}_r h_r x_1 = x_1, \tag{6.8}$$

where $\mathsf{s}_r = \mathsf{h}_2^{-1}\mathsf{h}_1$, $w_r = \mathrm{Ad}(h_r^{-1}\mathsf{h}_2^{-1}) w$. In particular, $\|w_r\| \ll e^{(-D+13)t}$. Assuming t is large enough compared to the implied multiplicative constant,

$$0 < \|w_r\| \leq e^{(-D+14)t}. \tag{6.9}$$

Recall that $x_1 = g_1\Gamma$ where $g_1 \in \mathfrak{S}_{\mathrm{cpt}}$, thus, (6.8) implies

$$\exp(w_r)h_r^{-1}\mathsf{s}_r h_r = g_1\gamma_r g_1^{-1}, \tag{6.10}$$

where $1 \neq \mathsf{s}_r \in H$ with $\|\mathsf{s}_r\| \ll e^t$ and $e \neq \gamma_r \in \Gamma$.

Similarly, if $\mathsf{h} \mapsto \mathsf{h}h_r x_1$ is not injective, we conclude that

$$h_r^{-1}\mathsf{s}_r h_r = g_1\gamma_r g_1^{-1} \neq e.$$

In this case we actually have $e \neq \gamma_r \in g_1^{-1}Hg_1$ — we will not use this extra information in what follows.

Some properties of the elements γ_r

Note that, in either case, we have

$$\|\gamma_r^{\pm 1}\| \leq e^{9t} \tag{6.11}$$

again we assumed t is large compared to $\|g_1\|$ hence the estimate $\ll e^{8t}$ is replaced by $\leq e^{9t}$.

Let $\xi > 0$ be so that $\|g\gamma g^{-1} - I\| \geq 20\xi$ for all $\gamma \in \Gamma \setminus \{1\}$ and $g \in \mathfrak{S}_{\mathrm{cpt}}$. Write $\mathsf{s}_r = \begin{pmatrix} a_1 & a_2 \\ a_3 & a_4 \end{pmatrix} \in H$ where $|a_i| \leq 10e^t$. Then by (6.10), we have

$$\|h_r^{-1}\mathsf{s}_r h_r - I\| = \left\| u_{-r}\begin{pmatrix} a_1 & e^{-7t}a_2 \\ e^{7t}a_3 & a_4 \end{pmatrix} u_r - I \right\| \geq 10\xi,$$

which implies that

$$\max\{e^{7t}|a_3|, |a_1 - 1|, |a_4 - 1|\} \geq \xi \gg 1. \tag{6.12}$$

Note also that if $e^{7t}|a_3| < \xi$, then $|a_2a_3| \leq 10\xi e^{-6t}$, thus $|a_1a_4 - 1| \ll e^{-6t}$. We conclude from (6.12) that $|a_1 - a_4| \gg 1$. Altogether,

$$\max\{e^{7t}|a_3|, |a_1 - a_4|\} \gg 1. \tag{6.13}$$

Let $I_{\mathrm{cpt}} = \{r \in [0,1] : h_r x_1 \in X_{\mathrm{cpt}}\}$ and $J_{\mathrm{cpt}} = \{r \in [1/2,1] : h_r x_1 \in X_{\mathrm{cpt}}\}$.

Claim

There are $\gg e^{3t}$ distinct elements in $\{\gamma_r : r \in J_{\mathrm{cpt}}\}$.

By (6.6) applied with $y = x_1$, $d = 7t$, and $J = [1/2, 1]$ we have $|J_{\mathrm{cpt}}| \geq 1/4$ (assuming t is large enough). Fix $r \in J_{\mathrm{cpt}}$ as above, and consider the set of $r' \in J_{\mathrm{cpt}}$ so that and $\gamma_r = \gamma_{r'}$. Then for each such r',

$$\begin{aligned} h_r^{-1}\mathsf{s}_r h_r &= \exp(-w_r) g_1 \gamma_r g_1^{-1} = \exp(-w_r)\exp(w_{r'}) h_{r'}^{-1}\mathsf{s}_{r'} h_{r'} \\ &= \exp(w_{rr'}) h_{r'}^{-1}\mathsf{s}_{r'} h_{r'}, \end{aligned}$$

where $w_{rr'} \in \mathfrak{g}$ and $\|w_{rr'}\| \ll e^{(-D+14)t}$.

Set $\tau = e^{7t}(r' - r)$. Assuming $D \geq 30$, we conclude that

$$u_\tau \mathsf{s}_r u_{-\tau} = h_{r'} h_r^{-1}\, \mathsf{s}_r\, h_r h_{r'}^{-1} = \exp(\hat{w}_{rr'})\mathsf{s}_{r'}, \tag{6.14}$$

where $\|\hat{w}_{rr'}\| = \|\operatorname{Ad}(h_{r'}) w_{rr'}\| \ll e^{(-D+21)t}$.

Finally, we compute

$$u_\tau \mathsf{s}_r u_{-\tau} = \begin{pmatrix} a_1 + a_3\tau & a_2 + (a_4 - a_1)\tau - a_3\tau^2 \\ a_3 & a_4 - a_3\tau \end{pmatrix}.$$

In view of (6.13), for every $r \in J_{\mathrm{cpt}}$ the set of $r' \in J_{\mathrm{cpt}}$ so that

$$|a_2 e^{-7t} + (a_4 - a_1)(r' - r) - a_3 e^{7t}(r' - r)^2| \leq 10^4 e^{-6t} \tag{6.15}$$

has measure $\ll e^{-3t}$ since at least one of the coefficients of this quadratic polynomial is of size $\gg 1$. Let J_r be the set of $r' \in J_{\mathrm{cpt}}$ for which (6.15) holds.

If $r' \in J_{\mathrm{cpt}} \setminus J_r$, then $|a_2 + (a_4 - a_1)\tau - a_3\tau^2| > 10^4 e^t$ (recall that $\tau = e^{7t}(r' - r)$), thus for all $r' \in J_{\mathrm{cpt}} \setminus J_r$, we have

$$\|u_\tau \mathsf{s}_r u_{-\tau}\| > 10^4 e^t > \|\exp(\hat{w}_{rr'})\mathsf{s}_{r'}\|,$$

in contradiction to (6.14).

In other words, for each $\gamma \in \Gamma$ the set of $r \in J_{\mathrm{cpt}}$ for which $\gamma_r = \gamma$ has measure $\ll e^{-3t}$ and so the set $\{\gamma_r : r \in J_{\mathrm{cpt}}\}$ has at least $\gg e^{3t}$ distinct elements, establishing the claim.

Zariski closure of the group generated by $\{\gamma_r : r \in I_{\mathrm{cpt}}\}$

We now consider two possibilities for the elements $\{\gamma_r : r \in I_{\mathrm{cpt}}\}$.

Case 1

The family $\{\gamma_r : r \in I_{\mathrm{cpt}}\}$ is commutative.

Let $\mathbf{L}$ denote the Zariski closure of $\langle \gamma_r : r \in I_{\rm cpt}\rangle$. Since $\langle \gamma_r\rangle$ is commutative, so is $\mathbf{L}$. Let $C_{\mathbf{G}}$ denote the center of $\mathbf{G}$. We claim that $\mathbf{L} = \mathbf{L}'\mathbf{C}'$ where $\mathbf{C}' \subset C_{\mathbf{G}}$ and $\mathbf{L}'$ is either a unipotent group or a torus. Indeed since $\mathbf{L}$ is commutative, we have $\mathbf{L} = \mathbf{TV}$ where $\mathbf{T}$ is a (possibly finite) algebraic subgroup of a torus, $\mathbf{V}$ is a unipotent group and $\mathbf{T}$ and $\mathbf{V}$ commute. Therefore, if both $\mathbf{T}$ and $\mathbf{V}$ are non-central, then $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ and $\Gamma = \Gamma_1 \times \Gamma_2$ is reducible. Moreover, $\mathbf{T} \subset \mathbf{T}'C_{\mathbf{G}}$ where $\mathbf{T}'$ is an algebraic subgroup of a torus, and $\mathbf{T}'$ and $\mathbf{V}$ belong to different $\mathrm{SL}_2(\mathbb{R})$ factors in G. Let us assume $\mathbf{V}$ belongs to the second factor. Recall from (6.8) that

$$\exp(w_r)h_r^{-1}\mathsf{s}_r h_r = g_1\gamma_r g_1^{-1}, \tag{6.16}$$

where $\|w_r\| \le e^{(-D+14)t}$ with $D \ge 30$ and $h_r^{-1}\mathsf{s}_r h_r \in H = \{(h,h) : h \in \mathrm{SL}_2(\mathbb{R})\}$. Now if $\gamma_r = (\gamma_r^1, \gamma_r^2)$, then (6.16) together with the bound $\|h_r^{-1}\mathsf{s}_r h_r\| \ll e^{8t}$ implies that $|\mathrm{tr}(\gamma_r^1) - \mathrm{tr}(\gamma_r^2)| \ll e^{(-D+22)t}$; moreover, since $\gamma_r^2 \in \mathbf{V}C_{\mathbf{G}}$, we have $|\mathrm{tr}(\gamma_r^2)| = 2$. This and the fact that the length of closed geodesics in (finite volume) hyperbolic surfaces is bounded away from zero imply that $|\mathrm{tr}(\gamma_r^1)| = 2$ if t is large enough. This contradicts the fact that $\mathbf{T}$ is a non-central subgroup of a torus. Hence, the claim holds.

We now show that $\mathbf{L}'$ is indeed a unipotent group. In view of the above discussion, $\#\{\gamma_r : r \in J_{\rm cpt}\} \ge e^{3t}$. Note also that that for every torus $T \subset G$, we have

$$\#(B_T(e,R)\cap\Gamma) \ll (\log R)^2,$$

where the implied constant is absolute. These, in view of the bound $\|\gamma_r\| \le e^{9t}$, see (6.11), imply that $\mathbf{L}'$ is unipotent.

Since $\mathbf{L}'$ is a unipotent subgroup of $\mathbf{G}$, we have that

$$\#\{\gamma_r : \|\gamma_r\| \le e^{4t/3}\} \ll e^{8t/3}.$$

Furthermore, there are $\gg e^{3t}$ distinct elements γ_r with $r \in J_{\rm cpt}$. Thus

$$\#\{\gamma_r : \|\gamma_r\| > 100e^{4t/3} \text{ and } r \in J_{\rm cpt}\} \gg e^{3t}.$$

For every $r \in I_{\rm cpt}$, write

$$\mathsf{s}_r = \begin{pmatrix} a_{1,r} & a_{2,r} \\ a_{3,r} & a_{4,r}\end{pmatrix} \in H,$$

where $|a_{j,r}| \le 10e^t$.

We will obtain an improvement of (6.12). Let $\xi \le \Upsilon \le e^{4t/3}$ and assume that $\|g_1\gamma_r g_1^{-1} - I\| \ge 20\Upsilon$ — by definition of ξ, this holds with $\Upsilon = \xi$ for all $r \in I_{\mathrm{cpt}}$ and as we have just seen this also holds for with $\Upsilon = e^{4t/3}$ for many choices of $r \in J_{\mathrm{cpt}}$. We claim

$$|a_{3,r}| \geq \Upsilon e^{-7t}. \tag{6.17}$$

Indeed by (6.10), we have

$$\|h_r^{-1}\mathsf{s}_r h_r - I\| = \left\| u_{-r}\begin{pmatrix} a_{1,r} & e^{-7t}a_{2,r} \\ e^{7t}a_{3,r} & a_{4,r}\end{pmatrix} u_r - I\right\| \geq 10\Upsilon.$$

This implies that $\max\{e^{7t}|a_{3,r}|, |a_{1,r} - 1|, |a_{3,r} - 1|\} \geq \Upsilon$. Assume contrary to our claim that $|a_{3,r}| < \Upsilon e^{-7t}$. Then

$$\max\{|a_{1,r} - 1|, |a_{4,r} - 1|\} \geq \Upsilon; \tag{6.18}$$

furthermore, we get $|a_{2,r}a_{3,r}| \ll \Upsilon e^{-6t}$. Thus,

$$|a_{1,r}a_{4,r} - 1| \ll \Upsilon e^{-6t} \ll e^{-14t/3}. \tag{6.19}$$

Moreover, since $h_r^{-1}\mathsf{s}_r h_r$ is very nearly $g_1\gamma_r g_1^{-1}$, and the latter is either a unipotent element or its minus, we conclude that

$$\min(|a_{1,r} + a_{4,r} - 2|, |a_{1,r} + a_{4,r} + 2|) \ll e^{(-D+22)t}. \tag{6.20}$$

Equations (6.19) and (6.20) contradict (6.18) if t is large enough, hence necessarily $|a_{3,r}| \geq \Upsilon e^{-7t}$.

Using this, we now show that Case 1 cannot occur. Since $\mathbf{L}'$ is unipotent, there exists some g so that $\mathbf{L}'(\mathbb{R}) \subset gNg^{-1}$; moreover g can be chosen to be in the maximal compact subgroup of G — for our purposes, we only need to know that the size of g can be bounded by an absolute constant.

It follows that

$$u_{-r}\begin{pmatrix} a_{1,r} & e^{-7t}a_{2,r} \\ e^{7t}a_{3,r} & a_{4,r}\end{pmatrix} u_r \in \exp(-w_r)(gNg^{-1}) \cdot C_{\mathbf{G}} \tag{6.21}$$

for all $r \in I_{\mathrm{cpt}}$. We show that this leads to a contradiction when $G = \mathrm{SL}_2(\mathbb{C})$, the proof in the other case is similar by considering first and second coordinates.

Let us write $g = \begin{pmatrix} a & b \\ c & d\end{pmatrix}$, then for all $z \in \mathbb{C}$ we have

$$g\begin{pmatrix} 1 & z \\ 0 & 1\end{pmatrix} g^{-1} = \begin{pmatrix} 1 - acz & a^2 z \\ -c^2 z & 1 + acz\end{pmatrix}.$$

Recall from the beginning of the proof that $h_0 x_1 \in X_{\mathrm{cpt}}$, i.e., $0 \in I_{\mathrm{cpt}}$. It follows that for some $z_0 \in \mathbb{C}$,

$$\begin{pmatrix} a_{1,0} & e^{-7t}a_{2,0} \\ e^{7t}a_{3,0} & a_{4,0} \end{pmatrix} = \pm\exp(-w_r)\begin{pmatrix} 1-acz_0 & a^2z_0 \\ -c^2z_0 & 1+acz_0 \end{pmatrix}.$$

By (6.17) applied with $\Upsilon = \xi$, $|a_{3,0}| \ge \xi e^{-7t}$. Since $|a|, |b|, |c|, |d| \ll 1$, comparing the bottom left entries of the matrices we get $|z_0| \gg 1$. Now, since $|a_{2,0}| \le 10e^t$, comparing the top right entries we conclude that $|a| \ll e^{-3t}$. Since $\det(g) = 1$, it follows that $|c|$ is also $\gg 1$.

Let now $r \in J_{\mathrm{cpt}}$ be so that $\|\gamma_r\| \ge 100e^{4t/3}$. We write $a'_{2,r} = e^{-7t}a_{2,r}$ and $a'_{3,r} = e^{7t}a_{3,r}$. By (6.17), applied this time with $\Upsilon = e^{4t/3}$, we have that $|a'_{3,r}| \ge e^{4t/3}$; note also that $|a'_{2,r}| \ll e^{-6t}$. In view of (6.21), there exists $z_r \in \mathbb{C}$ so that

$$\begin{aligned} u_{-r}\begin{pmatrix} a_{1,r} & a'_{2,r} \\ a'_{3,r} & a_{4,r} \end{pmatrix} u_r &= \begin{pmatrix} a_{1,r} - ra'_{3,r} & a'_{2,r} + (a_{4,r} - a_{1,r})r - a'_{3,r}r^2 \\ a'_{3,r} & a_{4,r} + ra'_{3,r} \end{pmatrix} \\ &= \pm\exp(-w_r)\begin{pmatrix} 1-acz_r & a^2z_r \\ -c^2z_r & 1+acz_r \end{pmatrix}. \end{aligned}$$

Since $|a'_{3,r}| \ge e^{4t/3}$, $|a_{1,r}|$ and $|a_{4,r}|$ are $\ll e^t$, and $|a'_{2,r}| \ll e^{-6t}$, and since $r \in [\frac{1}{2}, 1]$, we have that

$$|a'_{3,r}|/10 \le |a'_{2,r} + (a_{4,r} - a_{1,r})r - a'_{3,r}r^2| \le 2|a'_{3,r}|;$$

hence, since w_r is small, a^2z_r and c^2z_r should be comparable in size. On the other hand, using $r = 0$ we already established $|a| \ll e^{-3t}$ and $|c| \gg 1$, thus $|a^2z_r| \ll e^{-3t}|c^2z_r|$, in contradiction.

Altogether, we conclude that Case 1 cannot occur.

Case 2

There are $r, r' \in I_{\mathrm{cpt}}$ so that γ_r and $\gamma_{r'}$ do not commute.

Let v_H be as in Lemma 6.2. Then since $\exp(w_r)h_r^{-1}\mathsf{s}_r h_r = g_1\gamma_r g_1^{-1}$

$$\gamma_r.g_1^{-1}v_H = \exp(\mathrm{Ad}(g_1^{-1})w_r).g_1^{-1}v_H.$$

Moreover, since $\|w_r\| \le e^{(-D+14)t}$,

$$\|\mathrm{Ad}(g_1^{-1})w_r\| \ll e^{(-D+14)t};$$

similar statements also hold for r'.

Therefore, if D is large enough, we may apply Lemma 6.2 to conclude that there exists some $g_2 \in G$ with

$$\|g_1 - g_2\| \le C_{12} e^{(-D+14+9\kappa_9)t},$$

so that $\gamma_r . g_2^{-1} v_H = g_2^{-1} v_H$ and $\gamma_{r'} . g_2^{-1} v_H = g_2^{-1} v_H$.

In view of Lemma 6.1, thus, we have $Hg_2\Gamma$ is periodic and

$$\mathrm{vol}(Hg_2\Gamma) \le C_{11}\Big(\max\{\|\gamma_r^{\pm 1}\|, \|\gamma_{r'}^{\pm 1}\|\}\Big)^{\kappa_7} \le C_{11} e^{9\kappa_7 t},$$

where we used $\|\gamma_r^{\pm 1}\|, \|\gamma_{r'}^{\pm 1}\| \le e^{9t}$.

Then for t large enough, $\mathrm{vol}(Hg_2\Gamma) \le e^{D_0' t}$ and $d_X(g_1\Gamma, g_2\Gamma) \ll e^{(-D+D_0')t}$ for $D_0' = 9\max\{\kappa_7, \kappa_9\} + 14$.

Since $g_1\Gamma = x_1 = a_t u_{r_0} x_0$, part (2) in the proposition holds with $x' = (a_t u_{r_0})^{-1} g_2\Gamma$ and $D_0 = \max\{D_0' + 2, 30\}$ if t is large enough (recall that we already assumed in several places that $D \ge 30$). □

7 Margulis functions and random walks

As was mentioned earlier, the proof of Proposition 1.1 relies on two main ingredients: evolutions of Margulis functions under a certain random walk, and the (finitary) projection theorem, specifically Proposition 5.1, proved in Sect. 5. In this section we develop the necessary Margulis function techniques and show how to combine them with the results of Sect. 5 to prove Theorem 1.1 in Sect. 8.

The following is the main proposition encapsulating what is obtained using Margulis function techniques (and then input into Proposition 5.1).

Proposition 7.1 *Let* $0 < \eta < 0.01\eta_X$, $D \ge D_0 + 1$, *and* $x_0 \in X$, *where* D_0 *is as in Proposition 6.1, and* η_X *as in Proposition 3.2. Then there exists* t_0, *depending on* η, $\mathrm{inj}(x_0)$, *and* X, *so that if* $t \ge t_0$, *then at least one of the following holds:*

(1) Let $0 < \varepsilon < 0.1$ *and* $0 < \alpha < 1$. *Then there exist* $x_1 \in X_\eta$, *some* τ *with* $9t \le \tau \le 9t + 2m_0 Dt$ *(for* m_0 *depending on* α *— see (7.1)), and a subset* $F \subset B_{\mathfrak{r}}(0, 1)$ *containing* 0 *with*

$$e^{t/2} \le \#F \le e^{5t},$$

so that both of the following properties are satisfied:

- $\Big\{\exp(w)x_1 : w \in F\Big\} \subset \Big(\mathsf{B}^H_{e^{-t/R}} \cdot a_\tau \cdot \{u_r x_0 : |r| \le 4\}\Big) \cap X_\eta$, *where* $R > 0$ *depends on* D, ε, *and* α,
- $\sum_{w' \ne w} \|w - w'\|^{-\alpha} \le C \cdot (\#F)^{1+\varepsilon}$ *for all* $w \in F$ *(where the summation is over* $w' \in F$ *and* C *is an absolute constant).*

(2) There is $x' \in X$ *such that* Hx' *is periodic with*

$$\mathrm{vol}(Hx') \le e^{D_0 t} \quad \textit{and} \quad d_X(x', x_0) \le e^{(-D+D_0)t}.$$

Explicitly, m_0 is equal to m_α of (2.12), chosen so that for all $w \in \mathfrak{g}$, we have

$$\int_0^1 \|a_{m_0} u_r w\|^{-\alpha}\,\mathrm{d}r \le e^{-1}\|w\|^{-\alpha}. \tag{7.1}$$

7.1 The definition of a Margulis function

Throughout this section, $\mathcal{E} \subset X$ denotes a Borel set which is a disjoint finite union of local H orbits. More precisely, there is a finite set F and for every $w \in F$, there exist $x_w \in X$ and a bounded Borel set $\mathsf{E}_w \subset H$ satisfying the following

- the map $\mathsf{h} \mapsto \mathsf{h}.x_w$ is injective over E_w for all $w \in F$, and
- $\mathsf{E}_w.x_w \cap \mathsf{E}_{w'}.x_{w'} = \emptyset$ for all $w \ne w'$,

so that $\mathcal{E} = \bigcup_{w \in F} \mathsf{E}_w.x_w$.

For every $w \in F$, let μ_{E_w} denote the pushforward of the Haar measure $m_H|_{\mathsf{E}_w}$ under the map $h \mapsto h.x_w$. Put

$$\mu_{\mathcal{E}} = \frac{1}{\sum_w m_H(\mathsf{E}_w)} \sum_w \mu_{\mathsf{E}_w}. \tag{7.2}$$

For every $(h, z) \in H \times \mathcal{E}$, define

$$I_{\mathcal{E}}(h, z) := \Big\{ w \in \mathfrak{r} : 0 < \|w\| < \mathrm{inj}(hz),\ \exp(w)hz \in h\mathcal{E} \Big\}. \tag{7.3}$$

Since E_w is bounded for every w and F is finite, $I_{\mathcal{E}}(h, z)$ is a finite set for all $(h, z) \in H \times \mathcal{E}$.

Fix some $0 < \alpha < 1$. Define the Margulis function $f_{\mathcal{E}} = f_{\mathcal{E},\alpha} : H \times \mathcal{E} \to [1, \infty)$ as follows:

$$f_{\mathcal{E}}(h, z) = \begin{cases} \sum_{w \in I_{\mathcal{E}}(h,z)} \|w\|^{-\alpha} & \text{if } I_{\mathcal{E}}(h, z) \ne \emptyset \\ \mathrm{inj}(hz)^{-\alpha} & \text{otherwise} \end{cases}. \tag{7.4}$$

Let $\nu = \nu(\alpha)$ be the probability measure on H defined by

$$\nu(\varphi) = \int_0^1 \varphi(a_{m_0} u_r)\,\mathrm{d}r \qquad \text{for all } \varphi \in C_c(H), \tag{7.5}$$

where m_0 is as in (7.1).

Define $\psi_{\mathcal{E}}$ on $H \times \mathcal{E}$ by

$$\psi_{\mathcal{E}}(h, z) := \Big(\max\big\{\# I_{\mathcal{E}}(h, z), 1\big\}\Big) \cdot \mathrm{inj}(hz)^{-\alpha}. \tag{7.6}$$

We will use the following lemma to increase the *transversal* dimension inductively.

Lemma 7.1 *There exists some $C_{13} = C_{13}(\nu)$ so that for all $\ell \in \mathbb{N}$ and all $z \in \mathcal{E}$, we have*

$$\int f_{\mathcal{E}}(h, z)\,\mathrm{d}\nu^{(\ell)}(h) \le e^{-\ell} f_{\mathcal{E}}(e, z) + C_{13} \sum_{j=1}^{\ell} e^{j-\ell} \int \psi_{\mathcal{E}}(h, z)\,\mathrm{d}\nu^{(j)}(h),$$

where $\nu^{(j)}$ denotes the j-fold convolution of ν for every $j \in \mathbb{N}$.

Proof Throughout the argument, the set $\mathcal{E}$ is fixed; thus, we drop it from the indices in the notation. Note that $\mathrm{supp}(\nu) \subset \{h \in H : \|h\| \le e^{2m_0+1}\}$. Let $C \ge 1$ be so that

$$\|\mathrm{Ad}(h)w\| \le C\|w\|$$

for all h with $\|h\| \le e^{2m_0+1}$ and all $w \in \mathfrak{g}$. Increasing C if necessary, we also assume that $\mathrm{inj}(z)/C \le \mathrm{inj}(hz) \le C\,\mathrm{inj}(z)$ for all such h and all $z \in X$.

Let $h = a_{m_0} u_r$ for some $r \in [0, 1]$. Let $z \in \mathcal{E}$, and let $h' \in H$. First, let us assume that there exists some $w \in I(hh', z)$ with $\|w\| < \mathrm{inj}(hh'z)/C^2$. In view of the choice of C, this in particular implies that both $I(hh', z)$ and

$I(h', z)$ are non-empty. Hence, we have

$$\begin{aligned} f(hh', z) &= \sum_{w \in I(hh',z)} \|w\|^{-\alpha} \\ &= \sum_{\|w\| < \mathrm{inj}(hh'z)/C^2} \|w\|^{-\alpha} + \sum_{\|w\| \geq \mathrm{inj}(hh'z)/C^2} \|w\|^{-\alpha} \\ &\leq \sum_{w \in I(h',z)} \|\operatorname{Ad}(h)w\|^{-\alpha} + C^{2\alpha} \cdot \Big(\# I(hh', z) \Big) \cdot \mathrm{inj}(hh'z)^{-\alpha} \\ &= \sum_{w \in I(h',z)} \|\operatorname{Ad}(h)w\|^{-\alpha} + C^{2\alpha} \psi(hh', z). \end{aligned} \tag{7.7}$$

Note also that if $\|w\| \geq \mathrm{inj}(hh'z)/C^2$ for all $w \in I(hh', z)$ (which in view of the choice of C includes the case $I(h', z) = \emptyset$) or if $I(hh', z) = \emptyset$, then

$$\begin{aligned} f(hh', z) &\leq C^{2\alpha} \cdot \Big(\max\{\# I(hh', z), 1\} \Big) \cdot \mathrm{inj}(hh'z)^{-\alpha} \\ &= C^{2\alpha} \psi(hh', z). \end{aligned} \tag{7.8}$$

We now average (7.7) and (7.8) over $[0, 1]$ and conclude hat

$$\begin{aligned} \int_0^1 f(a_{m_0} u_r h', z)\, \mathrm{d}r \leq & \sum_{w \in I(h',z)} \int_0^1 \|a_{m_0} u_r w\|^{-\alpha}\, \mathrm{d}r \\ & + C^{2\alpha} \int_0^1 \psi(a_{m_0} u_r h', z)\, \mathrm{d}r, \end{aligned}$$

where we replace the summation on the right by 0 if $I(h', z) = \emptyset$. Thus by (7.1) we may conclude that

$$\int f(hh', z)\, \mathrm{d}\nu(h) \leq e^{-1} \cdot f(h', z) + C^{2\alpha} \int \psi(hh', z)\, \mathrm{d}\nu(h)$$

for all $h' \in H$. Iterating this estimate, we have

$$\int f(h, z)\, \mathrm{d}\nu^{(\ell)}(h) \leq e^{-1} \int f(h', z)\, \mathrm{d}\nu^{(\ell-1)}(h') + C^{2\alpha} \int \psi(h, z)\, \mathrm{d}\nu^{(\ell)}(h).$$

The claim in the lemma thus follows from the above by induction if we let $C_{13} = C^2$ and sum the geometric series. $\square$

7.2 Incremental dimension increase

Let $0 < \eta \leq 0.01\eta_X$ and $0 < \beta \leq \eta^2$. Define

$$\mathsf{E} = \mathsf{B}^H_\beta \cdot \left\{u_r : |r| \leq 0.1\eta\right\}.$$

Let $F \subset B_{\mathfrak{r}}(0, \beta)$ be a finite set, and let $y_0 \in X_{2\eta}$. Then for all $w \in F$ $\exp(w)y_0 \in X_\eta$, and $h \mapsto h\exp(w)y_0$ is injective on E. Put

$$\mathcal{E} = \mathsf{E}.\{\exp(w)y_0 : w \in F\}. \tag{7.9}$$

Let us begin with the following two elementary lemmas.

Lemma 7.2 *There exists $C_{14} > 0$ so that the following holds. For every $m \in \mathbb{N}$, every $|r| \leq 2$, and every $z \in \mathcal{E}$, we have*

$$\#I_{\mathcal{E}}(a_m u_r, z) \leq C_{14}\beta^{-6}e^{4m} \cdot (\#F)$$

Moreover, we have

$$\psi_{\mathcal{E}}(a_m u_r, z) \leq C_{14}\beta^{-7}e^{5m} \cdot (\#F).$$

Proof Let $z \in \mathcal{E}$, and let $w \in I_{\mathcal{E}}(a_m u_r, z)$. Then $\exp(w)a_m u_r z \in a_m u_r \mathcal{E}$. Therefore, using Lemma 2.3(2), we have

$$\mathsf{Q}^H_{\beta^2,m}.\exp(w)a_m u_r z \subset a_m u_r \mathcal{E}_+$$

where $\mathcal{E}_+ = \mathsf{B}^H_{\beta+100\beta^2}\left\{u_r \exp(w)y_0 : |r| \leq 0.1\eta,\ w \in F\right\}$ and

$$\mathsf{Q}^H_{\beta^2,m} = \left\{u^-_s : |s| \leq \beta^2 e^{-m}\right\} \cdot \{a_t : |t| \leq \beta^2\} \cdot \left\{u_r : |r| \leq \beta^2\right\}.$$

Note that the map $(\mathsf{h}, w') \mapsto \mathsf{h}\exp(w')a_m u_r z$ is injective over

$$\mathsf{Q}^H_{\mathrm{inj}(a_m u_r z)} \times \exp(B_{\mathfrak{r}}(0, \mathrm{inj}(a_m u_r z))),$$

and let $\mu_{\mathcal{E}_+}$ is the probability measure on $\mathcal{E}_+$ defined as in (7.2). Then

$$a_m u_r.\mu_{\mathcal{E}_+}\left(\mathsf{Q}^H_{\beta^2,m}\exp(w).a_m u_r z\right) \gg (\min\{\beta^2, \mathrm{inj}(a_m u_r z)\})^3 e^{-m}(\#F)^{-1},$$

where the implied constant is absolute.

Recall now that $\mathcal{E} \subset X_\eta$. Thus, $\mathrm{inj}(a_m u_r z) \gg e^{-m}\eta$. Recall also that $\beta \leq \eta^2$, this implies the first claim.

We now show the second claim. The above estimate and the definition of $\psi_{\mathcal{E}}(h, z)$ thus imply that

$$\psi_{\mathcal{E}}(a_m u_r, z) \ll \left(\beta^{-6} e^{4m} \cdot (\#F)\right) \cdot \mathrm{inj}(a_m u_r z)^{-1};$$

we also used $0 < \alpha < 1$ in the above upper bound. The second claim in the lemma follows. □

Lemma 7.3 *Let the notation be as above. In particular,* $y_0 \in X_{2\eta}$ *and*

$$\mathcal{E} = \mathsf{E}.\{\exp(w)y_0 : w \in F\},$$

where $F \subset B_{\mathfrak{r}}(0, \beta)$. *Let* $w_0 \in F$, *then*

$$\textstyle\sum_{w \neq w_0} \|w - w_0\|^{-\alpha} \leq 2 f_{\mathcal{E}}(e, z),$$

where $z = \exp(w_0)y_0$ *and the summation is over* $w \in F$.

Proof By the definition of $f_{\mathcal{E}}$, we have

$$f_{\mathcal{E}}(e, z) = \textstyle\sum_{v \in I_{\mathcal{E}}(e,z)} \|v\|^{-\alpha}.$$

Let $w_0 \neq w \in F$. We will find a unique vector $v_w \in I_{\mathcal{E}}(e, z)$ whose length is comparable to $\|w - w_0\|$. Let us begin with the following computation.

$$\begin{aligned} \exp(w)y &= \exp(w)\exp(-w_0)\exp(w_0)y_0 \\ &= h_w \exp(v_w)\exp(w_0)y_0 \\ &= h_w \exp(v_w)z, \end{aligned}$$

where $h_w \in H$, $v_w \in \mathfrak{r}$, $\|h_w - I\| \leq C_5\beta\|v_w\|$, and

$$0.5\|w - w_0\| \leq \|v_w\| \leq 2\|w - w_0\|, \tag{7.10}$$

see Lemma 2.1.

In particular, we have $\|h_w - I\| \ll \beta^2$; assuming $\beta \leq \eta^2$ is small enough, we conclude that $h_w^{\pm 1} \in \mathsf{B}^H_\beta$. Hence,

$$\exp(v_w)z = h_w^{-1}\exp(w)y_0 \in \mathcal{E}.$$

Moreover, using (7.10), we have $\|v_w\| \leq 2\beta \leq \mathrm{inj}(z)$. We thus conclude that $v_w \in I_{\mathcal{E}}(e, z)$.

Since $\exp(w)y_0 \neq \exp(w')y_0$ for $w \neq w' \in F \subset B_{\mathfrak{r}}(0, \beta)$, the map $w \mapsto v_w$ is well-defined and one-to-one. Altogether, we deduce that

$$\sum_{w \neq w_0} \|w - w_0\|^{-\alpha} \leq 2 \sum_{v \in I_{\mathcal{E}}(e,z)} \|v\|^{-\alpha} = 2 f_{\mathcal{E}}(e, z),$$

as was claimed. □

Lemma 7.4 *There exist* $0 < \kappa_{10} = \kappa_{10}(\nu) \leq \frac{1}{4m_0}$ *and* n_0 *depending on* X *so that the following holds. Let* $\mathcal{E}$ *be defined as in (7.9). Assume further that*

$$f_{\mathcal{E}}(e, z) \leq e^{Mn} \qquad \text{for all } z \in \mathcal{E} \tag{7.11}$$

for some $M > 0$ *and an integer* $n \geq n_0$.

Then for all $0 < \varepsilon < 0.1$ *and all* $\beta \geq e^{-0.01\varepsilon n}$ *at least one of the following holds.*

(1) $e^{Mn} < e^{\varepsilon n/2} \cdot (\#F)$*, or*
(2) For all integers $0 < \ell \leq \kappa_{10}\varepsilon n$ *and all* $z \in \mathcal{E}$*, we have*

$$\int f_{\mathcal{E}}(h, z)\, \mathrm{d}\nu^{(\ell)}(h) \leq 2e^{Mn-\ell}.$$

Proof By Lemma 7.1, applied with $f_{\mathcal{E}}$, we have

$$\int f_{\mathcal{E}}(h, z)\, \mathrm{d}\nu^{(\ell)}(h) \leq e^{-\ell} f_{\mathcal{E}}(e, z) + C_{13} \sum_{j=1}^{\ell} e^{j-\ell} \int \psi_{\mathcal{E}}(h, z)\, \mathrm{d}\nu^{(j)}(h).$$

Assuming n is large enough, Lemma 7.2 implies that there exists a constant C depending only on ν so that if $j \leq \varepsilon n/C$, then

$$\psi_{\mathcal{E}}(h, z) \leq (2C_{13})^{-1} e^{\varepsilon n/4} \cdot (\#F),$$

for all $h \in \mathrm{supp}(\nu^{(j)})$ — we used $\beta \geq e^{-0.01\varepsilon n}$ and assumed n is large enough to account for the factor $C_{14}\beta^{-7}$ in Lemma 7.2.

Let $\kappa_{10} = (2C)^{-1}$, and let $\ell \leq \kappa_{10}\varepsilon n$. Then

$$\int f_{\mathcal{E}}(h, z)\, \mathrm{d}\nu^{(\ell)}(h) \leq e^{-\ell} f_{\mathcal{E}}(e, z) + e^{\varepsilon n/4} \cdot (\#F) \leq e^{Mn-\ell} + e^{\varepsilon n/4} \cdot (\#F).$$

Therefore, either part (1) holds or $e^{Mn-\ell} \geq e^{(0.5-\kappa_{10})\varepsilon n} \cdot (\#F) \geq e^{\varepsilon n/4} \cdot (\#F)$. In the latter case, the above implies that

$$\int f_{\mathcal{E}}(h, z)\, \mathrm{d}\nu^{(\ell)}(h) \leq 2e^{Mn-\ell}$$

as we claimed in part (2). ◻

From this point until the Lemma 7.8, we fix some $0 < \varepsilon < 0.1$, and let $\beta = e^{-\kappa n/2}$ where $0 < \kappa \leq 0.02\kappa_{10}\varepsilon$ will be explicated later.

The following lemma will convert the estimate we obtained on average in Lemma 7.4 into pointwise information at most points. This is done in a fairly straightforward way essentially by using the Chebyshev inequality. Recall from Proposition 3.1 that for any interval $I \subset \mathbb{R}$ of length at least η and $t \geq |\log(\eta^2 \operatorname{inj}(x))| + C_7$

$$\left|\left\{r \in I : \operatorname{inj}(a_t u_r x) < \varepsilon^2\right\}\right| < C_7\varepsilon|I|.$$

Lemma 7.5 *Let the notation be as in Lemma 7.4. Let* $0 < \varepsilon < 0.1$, *and assume that*

$$\ell = \lfloor \kappa_{10}\varepsilon n \rfloor \geq 3|\log \eta| + C_7 + 6.$$

Further assume that Lemma 7.4(2) holds for these choices.

There exists a subset $L_{\mathcal{E}} \subset \operatorname{supp}(\nu^{(\ell)})$ *with* $\nu^{(\ell)}(L_{\mathcal{E}}) \geq 1 - 2e^{-\ell/8}$ *so that both of the following hold.*

(1) For all $h_0 \in L_{\mathcal{E}}$ *we have*

$$\int f_{\mathcal{E}}(h_0, z)\, \mathrm{d}\mu_{\mathcal{E}}(z) \leq e^{Mn - \frac{7\ell}{8}}.$$

(2) For all $h_0 \in L_{\mathcal{E}}$, *there exists* $\mathcal{E}(h_0) \subset \mathcal{E}$ *with* $\mu_{\mathcal{E}}(\mathcal{E}(h_0)) \geq 1 - O(\eta^{1/2})$, *so that for all* $z \in \mathcal{E}(h_0)$ *we have*

$$\mathsf{B}^H_{100\beta^2}.z \subset \mathcal{E} \tag{7.12a}$$

$$h_0 z \in X_{2\eta} \tag{7.12b}$$

$$f(h_0, z) \leq e^{Mn - \frac{3\ell}{4}}. \tag{7.12c}$$

Proof Let us begin by finding $L_{\mathcal{E}}$ which satisfies part (1). Apply Lemma 7.4 with $\ell = \lfloor \kappa_{10}\varepsilon n \rfloor$. Since Lemma 7.4(2) holds, we have

$$\iint f_{\mathcal{E}}(h, z)\, \mathrm{d}\mu_{\mathcal{E}}(z)\, \mathrm{d}\nu^{(\ell)}(h) \leq 2e^{Mn - \ell}.$$

Using this estimate and Chebyshev's inequality, we have

$$\nu^{(\ell)}\left\{h \in \operatorname{supp}(\nu^{(\ell)}) : \textstyle\int f(h, z)\, \mathrm{d}\mu_{\mathcal{E}}(z) > e^{Mn - \frac{7\ell}{8}}\right\} < 2e^{-\ell/8}. \tag{7.13}$$

Let $L_{\mathcal{E}}$ be the complement in $\mathrm{supp}(\nu^{(\ell)})$ of the set on the left side of (7.13), and let $h_0 \in L_{\mathcal{E}}$. Then

$$\int f(h_0, z)\, \mathrm{d}\mu_{\mathcal{E}}(z) \le e^{Mn - \frac{7\ell}{8}}. \tag{7.14}$$

The claim in part (1) thus holds with $L_{\mathcal{E}}$.

Let us now turn to the proof of (2). Let $h \in \mathrm{supp}(\nu^{(\ell)})$. Then $h = a_{\ell m_0} u_{\hat{r}}$ where $\hat{r} = \sum_{j=0}^{\ell-1} e^{-jm_0} r_{j+1}$ for some $r_1, \ldots, r_\ell \in [0, 1]$.

For every $z = u_s^- a u_{r'} u_r \exp(w).y_0 \in \mathcal{E}$, we have

$$hz = (a_{\ell m_0} u_{\hat{r}}) u_s^- a u_{r'} u_r \exp(w).y_0 = h' a_{\ell m_0} u_{r_s' + \hat{r} + r} \exp(w).y_0$$

where $h' \in \mathsf{B}_\beta^H$ and $|r_s'| \ll \beta$ for an absolute implied constant. Therefore, if $a_{\ell m_0} u_{r_s' + \hat{r} + r} \exp(w) y_0 \in X_{4\eta}$, then $hz \in X_{2\eta}$.

Apply Proposition 3.1 with $\exp(w) y_0 \in \mathcal{E} \subset X_\eta$ and the interval $I = [r_s' + \hat{r} - 0.1\eta, r_s' + \hat{r} + 0.1\eta]$. Since $\ell \ge 3|\log \eta| + C_7 + 6$, we conclude

$$|\{r \in [-0.1\eta, 0.1\eta] : a_{\ell m_0} u_{r_s' + \hat{r} + r} \exp(w) y_0 \notin X_{4\eta}\}| \le 0.4 C_7 \eta \sqrt{\eta}.$$

This estimate, the above observation, and the definition of $\mu_{\mathcal{E}}$ imply that

$$\mu_{\mathcal{E}}\{z \in \mathcal{E} : hz \notin X_{2\eta}\} \le 2C_7 \sqrt{\eta}, \tag{7.15}$$

for every $h \in \mathrm{supp}(\nu^{(\ell)})$.

Put

$$\mathcal{E}_- = \mathsf{B}^H_{\beta - 200\beta^2}\{u_r \exp(w) y_0 : |r| \le 0.1\eta, w \in F\};$$

then $\mu_{\mathcal{E}}(\mathcal{E}_-) \ge 1 - O(\beta)$.

Let now $h_0 \in L_{\mathcal{E}}$. Recall also that $0 < \beta < \eta^2$. Then (7.15), implies that there is a subset $\mathcal{E}'(h_0) \subset \mathcal{E}_-$ with

$$\mu_{\mathcal{E}}(\mathcal{E}'(h_0)) \ge 1 - O(\eta^{1/2}),$$

so that for all $z \in \mathcal{E}'(h_0)$ we have $h_0 z \in X_{2\eta}$. Hence all points in $\mathcal{E}'(h_0)$ satisfy (7.12a) and (7.12b).

We will find a subset $\mathcal{E}(h_0) \subset \mathcal{E}'(h_0)$ which satisfies (7.12c). Let

$$\mathcal{E}'' = \left\{ z \in \mathcal{E}'(h_0) : f(h_0, z) > e^{Mn - \frac{3\ell}{4}} \right\}.$$

Then

$$\mu_{\mathcal{E}}(\mathcal{E}'')e^{Mn-\frac{3\ell}{4}} \le \int_{\mathcal{E}''} f(h_0, z)\,\mathrm{d}\mu_{\mathcal{E}}(z)$$
$$\le \int_{\mathcal{E}} f(h_0, z)\,\mathrm{d}\mu_{\mathcal{E}}(z) \le e^{Mn-\frac{7\ell}{8}} \qquad \text{by (7.14).}$$

We conclude from the above that $\mu_{\mathcal{E}}(\mathcal{E}'') \ll e^{-\ell/8}$. Recall that $\beta = e^{-\kappa n/2}$ where $0 < \kappa \le 0.02\kappa_{10}\varepsilon$, thus we conclude that $\mu_{\mathcal{E}}(\mathcal{E}'') \ll \eta$.

Put $\mathcal{E}(h_0) := \mathcal{E}'(h_0) \setminus \mathcal{E}''$. Then $\mu_{\mathcal{E}}(\mathcal{E}(h_0)) \ge 1 - O(\eta^{1/2})$ and (7.12c) holds for every $z \in \mathcal{E}(h_0)$. The proof is complete. □

In the remaining parts of this section, we will write Q^H for

$$\mathsf{Q}^H_{\beta^2,\ell m_0} = \left\{u^-_s : |s| \le \beta^2 e^{-\ell m_0}\right\} \cdot \{a_t : |t| \le \beta^2\} \cdot \left\{u_r : |r| \le \beta^2\right\}, \tag{7.16}$$

where $\ell = \lfloor \kappa_{10}\varepsilon n \rfloor$, see (2.10).

Let us also define a subset in G by thickening Q^H in the transversal direction as follows. Put

$$\mathsf{Q}^G := \mathsf{Q}^H \cdot \exp(B_{\mathfrak{r}}(0, 2\beta^2)). \tag{7.17}$$

Lemma 7.6 *There exists a covering* $\left\{\mathsf{Q}^G.y_j : j \in \mathcal{J},\, y_j \in X_\eta\right\}$ *of* $X_{2\eta}$ *where* $\#\mathcal{J} \ll \beta^{-12} e^{\ell m_0}$ *and the implied constant depends on* X.

Moreover, if for every $h_0 \in L_{\mathcal{E}}$ *we let*

$$\mathcal{J}(h_0) = \left\{ j \in \mathcal{J} : h_0.\mu_{\mathcal{E}}\Big(h_0\mathcal{E}(h_0) \cap \mathsf{Q}^G.y_j\Big) \ge \beta^{13} e^{-\ell m_0} \right\} \tag{7.18}$$

and define $\hat{\mathcal{E}}(h_0) \subset \mathcal{E}(h_0)$ *by*

$$h_0\hat{\mathcal{E}}(h_0) = h_0\mathcal{E}(h_0) \bigcap \Big(\textstyle\bigcup_{j\in\mathcal{J}(h_0)} \mathsf{Q}^G.y_j \Big),$$

then $\mu_{\mathcal{E}}(\hat{\mathcal{E}}(h_0)) \ge 1 - O(\sqrt{\eta})$ *where the implied constant depends on* X. *In particular,* $\mathcal{J}(h_0) \ne \emptyset$.

Proof For simplicity in the notation, let us write B^G for

$$\mathsf{B}^G_{\beta^2} = \mathsf{B}^H_{\beta^2} \cdot \exp(B_{\mathfrak{r}}(0, \beta^2)).$$

We begin by constructing a covering of B^G. First recall that

$$m_H(\mathsf{Q}^H_{0.01\beta^2,\ell m_0}) \asymp e^{-\ell m_0} m_H(\exp(B_{\mathfrak{h}}(0, \beta^2))), \tag{7.19}$$

where the implied constant is absolute, see (2.10). Moreover, by Lemma 2.3 we have

$$\mathsf{Q}^H_{0.01\beta^2,\ell m_0} \cdot (\mathsf{Q}^H_{0.01\beta^2,\ell m_0})^{\pm 1} \subset \mathsf{Q}^H_{\beta^2,\ell m_0}. \tag{7.20}$$

Fix a maximal subset $\mathcal{H} \subset \mathsf{B}^H_{\beta^2}$ so that

$$\mathsf{Q}^H_{0.01\beta^2,\ell m_0} h \cap \mathsf{Q}^H_{0.01\beta^2,\ell m_0} h' = \emptyset,$$

for all $h \neq h' \in \mathcal{H}$. In view of (7.19), we have $\#\mathcal{H} \ll e^{\ell m_0}$ where the implied constant is absolute. Then using (7.20), we conclude that $\{\mathsf{Q}^H h_j : h_j \in \mathcal{H}\}$ covers $\mathsf{B}^H_{\beta^2}$ and $\#\mathcal{H} \asymp e^{\ell m_0}$.

Taking the product with $\exp(B_{\mathfrak{r}}(0,\beta^2))$, we thus obtain a covering

$$\{\mathsf{Q}^H h_j \exp(B_{\mathfrak{r}}(0,\beta^2)) : h_j \in \mathcal{H}\}$$

of the set B^G.

Recall that $\beta \leq \eta^2$, and that by Lemma 2.1, we have $(\mathsf{B}^G_\delta)^{-1} \cdot \mathsf{B}^G_\delta \subset \mathsf{B}^G_{c\delta}$ for all $\delta > 0$, where c is an absolute constant. Hence, arguing as above, there exists a covering

$$\{\mathsf{B}^G.\hat{y}_k : k \in \mathcal{K}, \hat{y}_k \in X_{2\eta}\},$$

of $X_{2\eta}$ which satisfies $\#\mathcal{K} \asymp \beta^{-12}$ for an implied constant depending on X.

Combining these two coverings, we obtain a covering

$$\{\mathsf{Q}^H h_j \exp(B_{\mathfrak{r}}(0,\beta^2)).\hat{y}_k : h_j \in \mathcal{H}, k \in \mathcal{K}\}.$$

of $X_{2\eta}$. Note further that

$$\mathsf{Q}^H h_j \exp(B_{\mathfrak{r}}(0,\beta^2)) = \mathsf{Q}^H \exp\Big(\mathrm{Ad}(h_j) B_{\mathfrak{r}}(0,\beta^2)\Big) h_j \subset \mathsf{Q}^G h_j;$$

where we used the fact that $\mathrm{Ad}(h_j)B_{\mathfrak{r}}(0,\beta^2) \subset B_{\mathfrak{r}}(0,2\beta^2)$ in the final inclusion above — this holds since $\|h_j - I\| \leq 2\beta^2$ and β is small.

Finally note that since $\hat{y}_k \in X_{2\eta}$ and $\|h_j - I\| \leq 2\beta^2$, we have $h_j\hat{y}_k \in X_\eta$, for every j, k. Altogether, we obtain a covering

$$\{\mathsf{Q}^G.y_j : j \in \mathcal{J}, y_j \in X_\eta\} = \{\mathsf{Q}^G.h_j\hat{y}_k : h_j \in \mathcal{H}, k \in \mathcal{K}\}$$

of $X_{2\eta}$ where $\#\mathcal{J} \ll \beta^{-12} e^{\ell m_0}$. This finishes the proof of the first claim.

To see the other claims, let $h_0 \in L_{\mathcal{E}}$, and define $\mathcal{J}(h_0)$ as in the statement. Then for every $j \notin \mathcal{J}(h_0)$, we have

$$h_0.\mu_{\mathcal{E}}\Big(h_0\mathcal{E}(h_0) \cap \mathsf{Q}^G.y_j\Big) < \beta^{13}e^{-\ell m_0}.$$

This estimate and the bound on $\#\mathcal{J}$ yield

$$h_0.\mu_{\mathcal{E}}\Big(h_0\mathcal{E}(h_0) \cap (\cup_{j\notin\mathcal{J}(h_0)}\mathsf{Q}^G.y_j)\Big) \ll \beta,$$

where the implied constant depends on X. The desired bound on the measure of $h_0\hat{\mathcal{E}}(h_0)$ thus follows since $h_0.\mu_{\mathcal{E}}\Big(h_0\hat{\mathcal{E}}(h_0)) \geq 1 - O(\sqrt{\eta})$.

The fact that $\mathcal{J}(h_0) \neq \emptyset$ is a consequence of the fact that $\hat{\mathcal{E}}(h_0) \neq \emptyset$, which is immediate from the above bound. □

The following lemma yields a set $\mathcal{E}_1$ defined as in (7.9), for some y_1 and F_1, but with an improved bound for $f_{\mathcal{E}_1}(e, z)$. This lemma will serve as our main tool for incremental dimension increase in the proof of Proposition 7.1.

Lemma 7.7 *There exists n_0 so that the following holds for all $n \geq n_0$. Let the notation be as in Lemmas 7.5 and 7.6. In particular, $0 < \varepsilon \leq 0.1$ and*

$$\ell = \lfloor \kappa_{10}\varepsilon n \rfloor \geq 3|\log\eta| + C_7 + 6;$$

assume further that $\#F \geq e^{n/2}$ and that Lemma 7.4(2) holds.

Let $h_0 \in L_{\mathcal{E}}$, and let $y = y_j$ for some $j \in \mathcal{J}(h_0)$. There exists some

$$h_0 z_1 \in h_0\mathcal{E}(h_0) \cap \mathsf{Q}^G.y$$

and a subset

$$F_1 \subset B_{\mathfrak{r}}(0, \beta) \quad \textit{with} \quad \#F_1 = \lceil \beta^{10}\cdot(\#F)\rceil$$

containing 0, so that both of the following are satisfied.

(1) For all $w \in F_1$, we have

$$\exp(w)h_0z_1 \in \mathsf{B}^H_{100\beta^2}.h_0\mathcal{E}(h_0).$$

(2) If we define $\mathcal{E}_1 = \mathsf{E}.\{\exp(w)h_0z_1 : w \in F_1\}$, then at least one of the following two possibilities hold

$$f_{\mathcal{E}_1}(e, z) \leq 2\cdot(\#F_1)^{1+\varepsilon} \qquad \textit{for all } z \in \mathcal{E}_1,\ \textit{or} \tag{7.21a}$$

$$f_{\mathcal{E}_1}(e, z) \leq e^{(M-\frac{2\kappa_{10}\varepsilon}{3})n} \qquad \textit{for all } z \in \mathcal{E}_1. \tag{7.21b}$$

Proof Let $h_0 \in L_{\mathcal{E}}$ and $y = y_j$ be as in the statement of the lemma.

The set $h_0\mathcal{E}(h_0) \cap \mathsf{Q}^G.y$ is contained in a finite union of local H-orbits. Let $\mathsf{M} \in \mathbb{N}$ be minimal so that

$$h_0\mathcal{E}(h_0) \cap \mathsf{Q}^G.y \subset \bigcup_{i=1}^{\mathsf{M}} \mathsf{Q}^H.\exp(w_i)y, \tag{7.22}$$

where $w_i \in B_{\mathfrak{r}}(0, 2\beta^2)$.

For each $1 \leq i \leq \mathsf{M}$, fix some $z_i \in \mathcal{E}(h_0)$ so that $h_0 z_i \in \mathsf{Q}^G.y$ and write

$$h_0 z_i = \mathsf{h}_i \exp(w_i)y \qquad \text{for some } \mathsf{h}_i \in \mathsf{Q}^H. \tag{7.23}$$

We claim that both of the following properties are satisfied

$$\mathsf{Q}^H.h_0 z_i \cap \mathsf{Q}^H.h_0 z_j = \emptyset \qquad 1 \leq i \neq j \leq M. \tag{7.24a}$$

$$h_0\mathcal{E}(h_0) \cap \mathsf{Q}^G.y \subset \bigcup_{i=1}^{\mathsf{M}} \mathsf{Q}^H \cdot (\mathsf{Q}^H)^{-1}.h_0 z_i. \tag{7.24b}$$

Assume contrary to (7.24a) that $\mathsf{h}h_0 z_i = \mathsf{h}' h_0 z_j$ for $i \neq j$. Then

$$\begin{aligned} \mathsf{h}^{-1}\mathsf{h}'\mathsf{h}_j \exp(w_j)y &= \mathsf{h}^{-1}\mathsf{h}' h_0 z_j \\ &= h_0 z_i = \mathsf{h}_i \exp(w_i)y. \end{aligned}$$

That is $\exp(-w_i)\hat{\mathsf{h}}\exp(w_j)y = y$ where $\hat{\mathsf{h}} = \mathsf{h}_i^{-1}\mathsf{h}^{-1}\mathsf{h}'\mathsf{h}_j$. Note moreover that $\hat{\mathsf{h}} \in \mathsf{B}^H_{100\beta^2}$, see (2.4), and $w_i \neq w_j \in B_{\mathfrak{r}}(0, 2\beta^2)$. Therefore $I \neq \exp(-w_i)\hat{\mathsf{h}}\exp(w_j) \in \mathsf{B}^G_{200\beta^2}$. Recall however that $\beta \leq \eta^2$ and $y \in X_{2\eta}$, thus, $g \mapsto g.h_0 z_i$ is injective on $\mathsf{B}^G_{1000\beta^2}$ for all small enough β. This contradiction implies that (7.24a) holds.

We now show (7.24b). Let $h_0 z \in h_0\mathcal{E}(h_0) \cap \mathsf{Q}^G.y$, then $h_0 z = \mathsf{h}\exp(w_i)y$ for $1 \leq i \leq \mathsf{M}$ and $\mathsf{h} \in \mathsf{Q}^H$. Moreover, we have $h_0 z_i = \mathsf{h}_i \exp(w_i)y$, thus $h_0 z = \mathsf{h}\mathsf{h}_i^{-1} h_0 z_i$ as claimed in (7.24b).

Recall now that $\mathcal{E} = \mathsf{E}.\{\exp(w)x : w \in F\}$ where $\mathsf{E} \subset H$ with $m_H(\mathsf{E}) \asymp \beta^2\eta$. In view of the definition of $\mu_{\mathcal{E}}$, see (7.2), we conclude that

$$h_0\mu_{\mathcal{E}}(\mathsf{Q}^H.h_0 z_i) \ll \beta^6 e^{-\ell m_0}\beta^{-2}\eta^{-1}(\#F)^{-1} \ll \beta^{3.5}e^{-\ell m_0}(\#F)^{-1};$$

recall that $\beta \leq \eta^2$.

Using (7.24a) and the definition of $\mathcal{J}(h_0)$ in (7.18), we deduce from the above that $\mathsf{M} \gg \beta^{9.5} \cdot (\#F)$. Assuming β is small so to account for the implied multiplicative constant (which depends only on G and Γ), we get

$$\mathsf{M} \geq \beta^{10} \cdot (\#F). \tag{7.25}$$

Let $1 \leq i, j \leq \mathsf{M}$, then using (7.23) we have

$$\begin{aligned} h_0 z_i &= \mathsf{h}_i \exp(w_i) y = \mathsf{h}_i \exp(w_i) \exp(-w_j) \mathsf{h}_j^{-1} h_0 z_j \\ &= \mathsf{h}_i \mathsf{h}_j^{-1} \exp(\mathrm{Ad}(\mathsf{h}_j) w_i) \exp(-\mathrm{Ad}(\mathsf{h}_j) w_j) h_0 z_j \\ &= \mathsf{h}_i \mathsf{h}_j^{-1} \mathsf{h}_{ij} \exp(w_{ij}) h_0 z_j, \end{aligned} \tag{7.26}$$

where $\mathsf{h}_{ij} \in H$ and $w_{ij} \in \mathfrak{r}$, $\mathsf{h}_{ii} = I$, $w_{ii} = 0$ for all i, j; moreover, we have

$$\|\mathsf{h}_{ij} - I\| \leq C_5 \beta^2 \|w_{ij}\| \quad \text{and} \tag{7.27a}$$

$$0.5\| \mathrm{Ad}(\mathsf{h}_j)(w_i - w_j)\| \leq \|w_{ij}\| \leq 2\| \mathrm{Ad}(\mathsf{h}_j)(w_i - w_j)\|, \tag{7.27b}$$

for all i, j, see Lemma 2.1.

Let $\{w_{i1}\}$ be defined as in (7.26), and let

$$F_1 \subset \{w_{i1} : 1 \leq i \leq \mathsf{M}\} \quad \text{with} \quad \#F_1 = \lceil \beta^{10} \cdot (\#F) \rceil; \tag{7.28}$$

this is possible thanks to (7.25). We will show that the claims in the lemma hold with z_1 and F_1.

First note that $h_0 z_1 \in h_0 \mathcal{E}(h_0) \cap \mathsf{Q}^G.y$ by its definition, and that F_1 satisfies the claimed properties by its definition and (7.28). Let us now show that part (1) in the statement of the lemma holds. Indeed by (7.26), we have

$$h_0 z_i = \mathsf{h}_i \mathsf{h}_1^{-1} \mathsf{h}_{i1} \exp(w_{i1}) h_0 z_1 \in \left(\mathsf{B}^H_{10\beta^2}\right). \exp(w_{i1}) h_0 z_1 \cap h_0 \mathcal{E}(h_0).$$

Therefore, $\exp(w_{i1}) h_0 z_1 \in (\mathsf{B}^H_{10\beta^2})^{-1} h_0 \mathcal{E}(h_0) \subset \mathsf{B}^H_{100\beta^2} h_0 \mathcal{E}(h_0)$, see (2.4) for the last inclusion. This establishes the claim in part (1) of the lemma.

For the proof of part (2) in the statement of the lemma, we need the following.

Sublemma *Let*

$$\mathcal{E}_1 = \mathsf{E}.\{\exp(w) h_0 z_1 : w \in F_1\}.$$

Let $z \in \mathcal{E}_1$, and write $z = \mathsf{h} u_r \exp(w_{i1}) h_0 z_1$ where $\mathsf{h} \in \mathsf{B}^H_\beta$, $|r| \leq 0.1\eta$, and $w_{i1} \in F_1$. Then

$$f_{\mathcal{E}_1}(e, z) \leq 2 f_{\mathcal{E}}(h_0, z_i) + \beta^{-2} e^{\ell m_0} \cdot (\#F_1),$$

where $z_i \in \mathcal{E}(h_0)$ is defined as in (7.23), in particular it satisfies

$$h_0 z_i = \mathsf{h}_i \mathsf{h}_1^{-1} \mathsf{h}_{i1} \exp(w_{i1}) h_0 z_1,$$

see (7.26), and $\ell = \lfloor \kappa_{10} \varepsilon n \rfloor$.

Let us first assume the sublemma, and finish the proof of the lemma.
Recall that $\beta = e^{-\kappa n/2}$ where

$$0 < \kappa \leq 0.02 \kappa_{10} \varepsilon. \tag{7.29}$$

In view of (7.25), we have

$$\# F_1 = \mathsf{M} \geq \beta^{10} \cdot (\# F) \geq e^{(1-10\kappa)n/2}, \tag{7.30}$$

where we used the bound $\# F \geq e^{n/2}$.
Recall also that $\kappa_{10} m_0 \leq 1/4$; this estimate and (7.29) imply that

$$\kappa_{10} \varepsilon m_0 + \kappa \leq (1 - 10\kappa)\varepsilon/2.$$

Using this and (7.30), we conclude that

$$e^{(\kappa_{10}\varepsilon m_0 + \kappa)n} \cdot (\# F_1) \leq e^{(1-10\kappa)\varepsilon n/2} \cdot (\# F_1) \leq (\# F_1)^{1+\varepsilon}. \tag{7.31}$$

Let $z \in \mathcal{E}_1$, and let $z_i \in \mathcal{E}(h_0)$ be as in the sublemma. Then, by (7.12c) we have

$$f_{\mathcal{E}}(h_0, z_i) \leq e^{Mn - \frac{3\ell}{4}},$$

where $\ell = \lfloor \kappa_{10} \varepsilon n \rfloor$. Thus, using the sublemma and (7.31) we deduce that

$$\begin{aligned} f_{\mathcal{E}_1}(e, z) &\leq (2e) \cdot e^{(M - \frac{3\kappa_{10}\varepsilon}{4})n} + e^{(\kappa_{10}\varepsilon m_0 + \kappa)n} \cdot (\# F_1) \\ &\leq 6 e^{(M - \frac{3\kappa_{10}\varepsilon}{4})n} + (\# F_1)^{1+\varepsilon}. \end{aligned}$$

We now consider two possibilities. Indeed, if $(\# F_1)^{1+\varepsilon} \geq 6 e^{(M - \frac{3\kappa_{10}\varepsilon}{4})n}$, then the above bound implies that

$$f_{\mathcal{E}_1}(e, z) \leq 2 (\# F_1)^{1+\varepsilon},$$

hence, (7.21a) holds.

Alternatively, if $(\#F_1)^{1+\varepsilon} < 6e^{(M-\frac{3\kappa 10^{\varepsilon}}{4})n}$, then

$$f_{\mathcal{E}_1}(e,z) \le 7e^{(M-\frac{3\kappa 10^{\varepsilon}}{4})n} \le e^{(M-\frac{2\kappa 10^{\varepsilon}}{3})n},$$

assuming $n \ge n_0$ is large enough. In consequence, (7.21b) holds.

These estimate finish the proof of part (2) and of the lemma, assuming the sublemma. □

Proof of the Sublemma The proof is similar to the proof of Lemma 7.3.

Let $z \in \mathcal{E}_1$. Then

$$\begin{aligned} f_{\mathcal{E}_1}(e,z) &= \textstyle\sum_{w\in I_{\mathcal{E}_1}(e,z)} \|w\|^{-\alpha} \\ &= \textstyle\sum_{\|w\|\le e^{-\ell m_0}\beta^2} \|w\|^{-\alpha} + \sum_{\|w\|> e^{-\ell m_0}\beta^2} \|w\|^{-\alpha} \\ &\le \textstyle\sum_{\|w\|\le e^{-\ell m_0}\beta^2} \|w\|^{-\alpha} + e^{\ell m_0}\beta^{-2}\cdot(\#F_1). \end{aligned} \tag{7.32}$$

In consequence, we need to investigate the first summation in (7.32). Let $w \in I_{\mathcal{E}_1}(e,z)$, then $z, \exp(w)z \in \mathcal{E}_1$. In view of the definition of $\mathcal{E}_1$ and (7.26), we may write

$$z = \mathsf{h}u_r\exp(w_{i1})h_0z_1 = \mathsf{h}u_r\mathsf{h}_{i1}^{-1}\mathsf{h}_1\mathsf{h}_i^{-1}h_0z_i = \bar{\mathsf{h}}h_0z_i$$

similarly, $\exp(w)z = \bar{\mathsf{h}}'h_0z_j$ where $1 \le i, j \le M$ and $\bar{\mathsf{h}}, \bar{\mathsf{h}}' \in \mathsf{B}^H_{0.15\eta}$, see (2.4).

Recall also from (7.26), that

$$h_0z_j = \mathsf{h}_j\mathsf{h}_i^{-1}\mathsf{h}_{ji}\exp(w_{ji})h_0z_i$$

where h_{ji} and w_{ji} satisfy (7.27a) and (7.27b). Hence we may apply Lemma 2.2, recall that $\beta^2 \le 0.1\eta$, and conclude

$$\|w_{ji}\| \le 2\|w\|. \tag{7.33}$$

Moreover, since h_0z_k's belong to different local H-orbits, see (7.23), $w \mapsto w_{ji}$ is well-defined and is one-to-one.

Assume now that $\|w\| \le e^{-\ell m_0}\beta^2$, then $\|w_{ji}\| \le 2e^{-\ell m_0}\beta^2$. This estimate and (7.27a) imply that

$$\|\mathsf{h}_{ji} - I\| \le 2C_5\beta^2\|w_{ji}\| \le e^{-\ell m_0}\beta^2$$

assuming β is small enough.

Recall also that $\mathsf{h}_i, \mathsf{h}_j \in \mathsf{Q}^H$ and that (7.12a) holds for z_j. Therefore, as $h_0 \in \operatorname{supp}(\nu^{(\ell)})$, in particular it is of the form $h_0 = a_{\ell m_0}u_r$ for $|r| < 2$, we have by (2.11) and (7.12a) that $\mathsf{h}_{ji}^{-1}\mathsf{h}_i\mathsf{h}_j^{-1}h_0z_i \in h_0\mathcal{E}$. That yields

$$\exp(w_{ji})h_0z_i = \mathsf{h}_{ji}^{-1}\mathsf{h}_i\mathsf{h}_j^{-1}h_0z_i \in h_0\mathcal{E}$$

which implies $w_{ji} \in I_{\mathcal{E}}(h_0, z_i)$ — recall that $\|w_{ji}\| \leq 2e^{-\ell m_0}\beta^2 < \mathrm{inj}(h_0z_i)$. This, (7.33), and the fact that $w \mapsto w_{ji}$ is one-to-one imply that

$$\textstyle\sum_{\|w\|\leq e^{-\ell m_0}\beta^2} \|w\|^{-\alpha} \leq 2f_{\mathcal{E}}(h_0, z_i).$$

This estimate and (7.32) finish the proof of the sublemma. □

We also need a lemma which is based on Proposition 6.1 and will provide the base case for our inductive argument in the proof Proposition 7.1.

Lemma 7.8 *Let the notation be as in Proposition 7.1. In particular, let $0 < \eta < 0.01\eta_X$, $D \geq D_0$, and $x_0 \in X$. There exists t_1, depending on η, D, and the injectivity radius of x_0, so that the following holds for all $t \geq t_1$.*

Let $0 < \varepsilon < 0.1$, and let $\beta = e^{-\kappa(t+1)/2}$ where $0 < \kappa \leq 0.02\kappa_{10}\varepsilon$. Then at least one of the following holds.

(1) There exists a subset $F \subset B_{\mathfrak{r}}(0, \beta)$ with

$$e^{t-5\kappa(t+1)} \leq \#F \leq e^{4t+0.5\kappa(t+1)}$$

and some $y \in X_{2\eta} \cap \left(\mathsf{B}^H_\beta \cdot a_{9t}\right).\{u_rx_0 : r \in [0, 1.05]\}$ so that if we put

$$\mathcal{E} = \mathsf{E}.\{\exp(w)y : w \in F\},$$

then $\mathcal{E} \subset \left(\mathsf{B}^H_{10\beta} \cdot a_{9t}\right).\{u_rx_0 : r \in [0, 1.1]\}$ and

$$f_{\mathcal{E}}(e, z) \leq e^{D(t+1)} \qquad \text{for all } z \in \mathcal{E}.$$

(2) There is $x' \in X$ such that Hx' is periodic with

$$\mathrm{vol}(Hx') \leq e^{D_0t} \quad \text{and} \quad d_X(x_0, x') \leq e^{(-D+D_0)t}.$$

Proof Put $\mathcal{C}_0 = \{a_{8t}u_rx_0 : r \in [0, 1]\}$. Apply Proposition 6.1 with x_0 and t. If part (2) in that proposition holds, then part (2) above holds and the proof is complete. Therefore, let us assume that Proposition 6.1(1) holds.

Let $x \in X_{\mathrm{cpt}} \cap \mathcal{C}_0$ be a point given by Proposition 6.1(1); put

$$\mathcal{C} = \left(\mathsf{B}^H_\beta \cdot a_t\right).\{u_rx : r \in [0, 1]\} \subset X;$$

and let $\mathcal{C}_- = \mathsf{B}^H_{\beta-100\beta^2} \cdot a_t \cdot \left\{u_rx : r \in [100e^{-t}, 1 - 100e^{-t}]\right\}$.

Let $\mu_{\mathcal{C}}$ denote the pushforward to $\mathcal{C}$ of the normalized restriction of the Haar measure on H to $\mathsf{C} := \mathsf{B}^H_\beta \cdot a_t \cdot \{u_r : r \in [0,1]\} \subset H$ — the set C was denoted by $\mathsf{E}_{1,t,\beta}$ in (2.9), we will use the notation C in this proof to avoid confusion with $\mathsf{E} = \mathsf{B}^H_\beta \cdot \{u_r : |r| \le 0.1\eta\}$ from Sect. 7.2.

We now use arguments similar to, and simpler than, the ones used in Lemmas 7.6 and 7.7 to construct the set $\mathcal{E}$ as in part (1).

First note that by Proposition 3.1, if $t > |\log\eta| + C$ (where C depends on X) we have

$$\mu_{\mathcal{C}}(\mathcal{C}_- \cap X_{4\eta}) \ge 1 - O(\sqrt{\eta}), \tag{7.34}$$

where the implied constant depends on G and Γ.

Let $\{\mathsf{B}^G_{\beta^2}.\hat{y}_j : j \in J\}$ be a covering of $X_{4\eta}$ so that $J \asymp \beta^{-12}$ where the implied constant depends on G and Γ, see Lemma 7.6. Let J' be the set of those $j \in J$ so that

$$\mu_{\mathcal{C}}(\mathcal{C}_- \cap X_{4\eta} \cap \mathsf{B}^G_{\beta^2}.\hat{y}_j) \ge \beta^{13}. \tag{7.35}$$

This definition, the fact that $\mu_{\mathcal{C}}$ is a probability measure (and moreover by (7.34) a probability measure giving large measure to $\mathcal{C}_- \cap X_{4\eta}$) and the estimate $J \asymp \beta^{-12}$ imply that

$$\mu_{\mathcal{C}}\Big(\mathcal{C}_- \bigcap \Big(\bigcup_{j\in J'} \mathsf{B}^G_{\beta^2}.\hat{y}_j\Big)\Big) \ge 1 - O(\sqrt{\eta}),$$

where the implied constant depends on X. Moreover, (7.35) implies that for any $j \in J'$, $\mathsf{B}^G_{\beta^2}.\hat{y}_j \subset X_{3\eta}$.

Let $j \in J'$; put $\hat{y} = \hat{y}_j$ and $\hat{\mathcal{C}} = \mathcal{C}_- \cap \mathsf{B}^G_{\beta^2}.\hat{y}$. Then, there are $w_i \in B_{\mathfrak{r}}(0, \beta^2)$ and $\mathsf{h}_i \in \mathsf{B}^H_{\beta^2}$, $i = 1, \ldots, \mathsf{M}$, so that $\mathsf{h}_i \exp(w_i)\hat{y} \in \mathcal{C}_-$ and

$$\hat{\mathcal{C}} = \bigcup_{i=1}^{\mathsf{M}} \mathsf{C}_i \mathsf{h}_i \exp(w_i)\hat{y},$$

where $\mathsf{C}_i \subset \mathsf{B}^H_{10\beta^2}$.

Recall that $\beta \le \eta^2$ and that $m_H(\mathsf{C}) \asymp e^t\beta^2$. In consequence, we have

$$\mu_{\mathcal{C}}(\mathsf{B}^H_{10\beta^2}) \ll \beta^6 \cdot (e^t\beta^2)^{-1} = \beta^4 e^{-t}.$$

This and (7.35) imply that $\mathsf{M} \gg \beta^9 e^t$. Assuming that β is small enough, to account for the implicit constant, we have

$$\mathsf{M} \geq \beta^{10} e^t. \tag{7.36}$$

We now use $\hat{\mathcal{C}}$ to define $\mathcal{E}$ which satisfies the desired properties in part (1). To that end, note that for every i and j we have

$$\begin{aligned} \mathsf{h}_i \exp(w_i)\hat{y} &= \mathsf{h}_i \exp(w_i)\exp(-w_j)\mathsf{h}_j^{-1}\mathsf{h}_j \exp(w_j)\hat{y} \\ &= \mathsf{h}_i\mathsf{h}_j^{-1}\mathsf{h}_{ij}\exp(w_{ij})\mathsf{h}_j \exp(w_j)\hat{y}, \end{aligned} \tag{7.37}$$

where $\mathsf{h}_{ij} \in H$ and $w_{ij} \in \mathfrak{r}$, $\mathsf{h}_{ii} = 1$, $w_{ii} = 0$ for all i, j; moreover, we have

$$\|\mathsf{h}_{ij} - I\| \leq C_5\beta^2\|w_{ij}\| \qquad \text{and} \tag{7.38a}$$

$$0.5\|\operatorname{Ad}(\mathsf{h}_j)(w_i - w_j)\| \leq \|w_{ij}\| \leq 2\|\operatorname{Ad}(\mathsf{h}_j)(w_i - w_j)\|, \tag{7.38b}$$

for all i, j, see Lemma 2.1. In particular, for all i, j we have

$$\|\mathsf{h}_{ij} - I\| \ll \beta^4 \tag{7.39}$$

for an absolute implied constant.

Thus, assuming β is small enough, we have $\mathsf{h}_i\mathsf{h}_j^{-1}\mathsf{h}_{ij} \in \mathsf{B}^H_{10\beta^2}$, for all i, j. This and the fact that $\mathsf{h}_i \exp(w_i)\hat{y} \in \mathcal{C}_-$ imply that

$$\begin{aligned} \exp(w_{ij})\mathsf{h}_j \exp(w_j)\hat{y} &= (\mathsf{h}_i\mathsf{h}_j^{-1}\mathsf{h}_{ij})^{-1}\mathsf{h}_i \exp(w_i)\hat{y} \\ &\in \mathsf{B}^H_{10\beta^2}.\mathcal{C}_- \subset \mathcal{C}, \end{aligned} \tag{7.40}$$

for all i and j.

Let $y := \mathsf{h}_1 \exp(w_1)\hat{y} \in \mathcal{C}_- \cap X_{2\eta}$ and $F = \{w_{i1} : i = 1, \ldots, \mathsf{M}\}$. First note that by (7.40) and Lemma 6.3, we have

$$\#F \ll e^{4t} \leq \beta^{-1}e^{4t},$$

where in the last inequality we assume β is small to account for the implied constant. This and (7.36) imply that

$$e^{t-5\kappa(t+1)} = \beta^{10}e^t \leq \#F = \mathsf{M} \leq \beta^{-1}e^{4t} = e^{4t+0.5\kappa(t+1)}, \tag{7.41}$$

which is the bound we claimed in part (1).

Define $\mathcal{E} = \mathsf{E}.\{\exp(w_{i1})y : w_{i1} \in F\}$. By (7.40), we have $\{\exp(w_{i1})y : w_{i1} \in F\} \subset \mathsf{B}^H_{10\beta^2}.\mathcal{C}_-$. Recall also that $\mathsf{E} = \mathsf{B}^H_\beta \cdot \{u_r : |r| \leq 0.1\eta\}$ and

$$u_r \cdot \mathsf{B}^H_\beta \cdot a_t \subset \mathsf{B}^H_{2\beta} \cdot a_t \cdot u_{e^{-t}r}, \tag{7.42}$$

for all $|r| \leq 0.1\eta$. Thus

$$\begin{aligned}\mathcal{E} &= \mathsf{B}^H_\beta \cdot \{u_r : |r| \leq 0.1\eta\}.\{\exp(w_{i1})y : w_{i1} \in F\}\\ &\subset \mathsf{B}^H_\beta \cdot \mathsf{B}^H_{2\beta} \cdot a_t.\{u_r x : r \in [0,1]\}\\ &\subset \mathsf{B}^H_{5\beta} \cdot a_t.\{u_r x : r \in [0,1]\}\\ &\subset \Big(\mathsf{B}^H_{5\beta} \cdot a_t \cdot \{u_r : r \in [0,1]\}\Big) \cdot a_{8t}.\{u_r x_0 : r \in [0,1]\}\\ &\subset \mathsf{B}^H_{5\beta} \cdot a_t \cdot \mathsf{B}^s_{5\beta} \cdot \{u_r : |r| \leq 2\} \cdot a_{8t}.\{u_r x_0 : r \in [0,1]\},\end{aligned}$$

where $\mathsf{B}^s_\varrho = \{u^-_s : |s| \leq \varrho\} \cdot \{a_d : |d| \leq \varrho\}$ and we use $x \in \mathcal{C}_0$ in the third line. Using $u_r a_{8t} = a_{8t} u_{e^{-8t}r}$, which holds for all r and t, we conclude

$$\mathcal{E} \subset \mathsf{B}^H_{5\beta} \cdot a_t \cdot \mathsf{B}^s_{5\beta} \cdot a_{8t}.\{u_r x_0 : r \in [0, 1.1]\},$$

so long as $t \geq 1$.

Finally note that $a_t \mathsf{B}^s_{2\beta} a_{-t} = \{u^-_s : |s| \leq 2e^{-t}\beta\} \cdot \{a_\ell : |l| \leq 2\beta\}$ for all t. Thus assuming t is large enough, we have

$$\mathcal{E} \subset \mathsf{B}^H_{10\beta} \cdot a_{9t} \cdot \{u_r x_0 : r \in [0, 1.1]\}.$$

We claim

$$f_{\mathcal{E}}(e, z) \leq 2e^{Dt} \leq e^{D(t+1)} \qquad \text{for all } z \in \mathcal{E}. \tag{7.43}$$

In view of the above discussion, this estimate finishes the proof of part (1) and of the lemma modulo (7.43).

The proof of (7.43) is similar to the proof of Lemma 7.3. For every $1 \leq i \leq \mathsf{M}$, put $z_i = \mathsf{h}_i \exp(w_i)\hat{y}$. Let $w \in I_{\mathcal{E}}(e, z)$, then $z, \exp(w)z \in \mathcal{E}$. In view of the definition of $\mathcal{E}$ and (7.37), we may write

$$z = \mathsf{h} u_r \exp(w_{i1})y = \mathsf{h} u_r (\mathsf{h}_i \mathsf{h}_1^{-1} \mathsf{h}_{i1})^{-1} z_i = \bar{\mathsf{h}} z_i$$

similarly, $\exp(w)z = \bar{\mathsf{h}}' z_j$ where $1 \leq i, j \leq M$ and $\bar{\mathsf{h}}, \bar{\mathsf{h}}' \in \mathsf{B}^H_{0.15\eta}$, see (7.39) and (2.4). Recall also from (7.37) again that

$$z_j = \mathsf{h}_j \mathsf{h}_i^{-1} \mathsf{h}_{ji} \exp(w_{ji}) z_i$$

where h_{ji} and w_{ji} satisfy (7.38a) and (7.38b). Hence we may apply Lemma 2.2, recall that $\beta \leq \eta^2$, and conclude

$$\|w_{ji}\| \leq 2\|w\|. \tag{7.44}$$

Moreover, since $\mathsf{h}_k \exp(w_k)\hat{y}$'s belong to different local H-orbits, $w \mapsto w_{ji}$ is well-defined and one-to-one. Recall also from (7.40) that

$$(\mathsf{h}_j \mathsf{h}_i^{-1} \mathsf{h}_{ji})^{-1} z_j = \exp(w_{ji}) z_i \in \mathcal{C},$$

for all i, j. Moreover by (7.38b), we have $\|w_{ji}\| \ll \beta^2 \leq \mathrm{inj}(z_i)$. Altogether, we conclude that $w_{ji} \in I_{\mathcal{C}}(e, z_i)$.

This, (7.44), and the fact that $w \mapsto \hat{w}_{ji}$ is one-to-one imply that

$$\begin{aligned} f_{\mathcal{E}}(e, z) &= \textstyle\sum_{w \in I_{\mathcal{E}}(e,z)} \|w\|^{-\alpha} \\ &\leq 2 \textstyle\sum_{w \in I_{\mathcal{C}}(e,z_i)} \|w\|^{-\alpha} \\ &= 2 f_{\mathcal{C}}(e, z_i) \leq 2e^{Dt}, \end{aligned}$$

where the last inequality is a consequence of Proposition 6.1(1). □

Proof of Proposition 7.1

We now complete the proof of Proposition 7.1. Roughly speaking, the proof is based on repeatedly applying Lemma 7.7 to improve the bound on the corresponding Margulis function.

Let $0 < \eta < 0.01\eta_X$, $D \geq D_0 + 1$ (for D_0 as in Proposition 6.1), $x_0 \in X$, and $t > 0$ (large) be as in the statement of Proposition 7.1.

Fix some κ satisfying

$$0 < \kappa \leq \frac{\kappa_{10}\varepsilon}{100D}, \tag{7.45}$$

and put $\beta = e^{-\kappa(t+1)/2}$.

We assume t is large enough so that $\beta \leq \eta^2$; assume further that $t \geq t_1$ where t_1 is as in Lemma 7.8.

Base of the induction

Apply Lemma 7.8 with η, β, D, x_0, and t. If Lemma 7.8(2) holds, then Proposition 7.1(2) holds and the proof is complete. Therefore, we assume that Lemma 7.8(1) holds. Let

$$\mathcal{E} = \mathsf{E}.\{\exp(w)y : w \in F\} \subset \mathsf{B}^H_{10\beta} \cdot a_{9t} \cdot \{u_r x_0 : r \in [0, 1.1]\} \qquad (7.46)$$

be as in Lemma 7.8(1). Put $n = t + 1$, $M = D$, $y_0 = y$, $F_0 = F$, and $\mathcal{E}_0 = \mathcal{E}$. We further assume $t + 1 \geq 4n_0$ where n_0 is as in Lemma 7.4.

Apply Lemma 7.4 with this $\mathcal{E}_0$. If Lemma 7.4(1) holds, then $e^{Mn} \leq e^{\varepsilon n/2} \cdot (\#F_0)$. Since $\#F_0 \geq e^{t-5\kappa(t+1)} \geq e^{n/2}$, we have

$$f_{\mathcal{E}_0}(e, z) \leq e^{Mn} \leq e^{\varepsilon n/2} \cdot (\#F_0) \leq (\#F_0)^{1+\varepsilon}.$$

Hence by Lemma 7.3, for all $w \in F_0$,

$$\textstyle\sum_{w \neq w'} \|w - w'\|^{-\alpha} \leq 4 \cdot (\#F_0)^{1+\varepsilon}.$$

This estimate together with (7.46) implies that part (1) in the proposition holds with $\tau = 9t$, $x_1 = y$ and $F = F_0$ if we choose R large enough so that $e^{-t/R} \geq 10\beta$.

The inductive step

In view of the above discussion, let us assume that Lemma 7.4(2) holds for $\mathcal{E}_0$. Let $L_{\mathcal{E}_0}$ be as in Lemma 7.5. Let $h_0 \in L_{\mathcal{E}_0}$, and let y_j for some $j \in \mathcal{J}(h_0)$ be as in Lemma 7.6. Moreover, note that

$$e^{n/2} \leq e^{t-5\kappa(t+1)} \leq \#F_0 \leq e^{4t+0.5\kappa(t+1)} = \beta^{-1} e^{4t},$$

and $n > n_0$. Therefore, we may apply Lemma 7.7. By that lemma, there exist z_1 with

$$h_0 z_1 \in h_0 \mathcal{E}_0(h_0) \cap \mathsf{Q}^G.y_j$$

and a subset $F_1 \subset B_{\mathfrak{r}}(0, \beta)$, containing 0, with

$$\#F_1 = \lceil \beta^{10} \cdot (\#F_0) \rceil$$

so that both of the following are satisfied.

(I-1) For all $w \in F_1$, we have

$$\exp(w) h_0 z_1 \in \mathsf{B}^H_{100\beta^2}.h_0 \mathcal{E}_0(h_0).$$

(I-2) If we put $\mathcal{E}_1 = \mathsf{E}.\{\exp(w)h_0 z_1 : w \in F_1\}$, then at least one of the following properties hold:

$$f_{\mathcal{E}_1}(e, z) \leq 2 \cdot (\#F_1)^{1+\varepsilon} \qquad \text{for all } z \in \mathcal{E}_1, \text{ or} \tag{7.47a}$$

$$f_{\mathcal{E}_1}(e, z) \leq e^{(M - \frac{2\kappa_{10\varepsilon}}{3})n} \qquad \text{for all } z \in \mathcal{E}_1. \tag{7.47b}$$

If (7.47a) holds, we set $\mathcal{E}_{\text{fin}} = \mathcal{E}_1$. Otherwise, we repeat the above construction to define sets $F_2, \ldots$ and the corresponding $\mathcal{E}_2, \ldots$.

Let $i_{\max} := \lfloor \frac{6M-3}{4\kappa_{10\varepsilon}} \rfloor + 1$, then by the choice of κ in (7.45), we have

$$M - \frac{2\kappa_{10\varepsilon}}{3} i_{\max} \leq 1/2 \quad \text{and} \quad 5\kappa(i_{\max} + 1) \leq 1/4 \tag{7.48}$$

Suppose now that $i \leq i_{\max}$, and we have constructed $\mathcal{E}_0, \ldots, \mathcal{E}_i$ so that (7.47a) does *not* hold for $\mathcal{E}_k$, for all $0 \leq k \leq i$. Then (7.47b) holds and we have

$$f_{\mathcal{E}_k}(e, z) \leq e^{(M - \frac{2\kappa_{10\varepsilon}}{3}k)n} \qquad \text{for all } 0 \leq k \leq i \text{ and all } z \in \mathcal{E}_k. \tag{7.49}$$

By the second estimate in (7.48), for all $0 \leq k \leq i$, we have

$$\begin{aligned} \#F_k \geq \beta^{10k} \cdot (\#F_0) &\geq e^{t - 5\kappa(k+1)(t+1)} \\ &\geq e^{(3t-1)/4} \geq e^{2n/3}. \end{aligned}$$

Since (7.47a) does not hold for $\mathcal{E}_k$, but (7.47b) holds, we have

$$e^{\varepsilon n/2} \cdot (\#F_k) \leq (\#F_k)^{1+\varepsilon} \leq e^{(M - \frac{2\kappa_{10\varepsilon}}{3}k)n}$$

for all $0 \leq k \leq i$.

Thus we are in case Lemma 7.4(2) for all these k, moreover, we have the bound $\#F_k \geq e^{2n/3}$. In consequence, Lemma 7.7 is applicable in every step, and we can define F_{i+1} and $\mathcal{E}_{i+1}$.

The conclusion of the proof

We now show that in at most $i_{\max}$ many steps, we obtain a set $\mathcal{E}$ which satisfies (I-1) above and (7.47a). Indeed, in view of the first estimate in (7.48),

$$e^{(M - \frac{2\kappa_{10\varepsilon}}{3} i_{\max})n} < e^{n/2}.$$

As $\#F_k \geq e^{2n/3}$ for all F_k's which are constructed, this observation together with (7.49) implies that in at most $i_{\max}$ number of steps, (7.47a) holds.

In consequence, we get some $i_{\text{fin}} \le i_{\max}$, so that if we put $F_{\text{fin}} := F_{i_{\text{fin}}} \subset B_{\mathfrak{r}}(0, \beta)$, then $\#F_{\text{fin}} \ge e^{2n/3}$, and the set

$$\mathcal{E}_{\text{fin}} = \mathsf{E}.\{\exp(w)y_{\text{fin}} : w \in F_{\text{fin}}\}$$

satisfies

$$f_{\mathcal{E}_{\text{fin}}}(e, z) \le 2 \cdot (\#F_{\text{fin}})^{1+\varepsilon} \tag{7.50}$$

for all $z \in \mathcal{E}_{\text{fin}}$ (cf. (7.47a)).

We claim that F_{fin} and y_{fin} also satisfy

$$\{\exp(w)y_{\text{fin}} : w \in F_{\text{fin}}\} \subset \left(\mathsf{B}^H_{100(i_{\text{fin}}+10)\beta} \cdot a_\tau \cdot \{u_r : |r| \le 4\}\right).x_0 \cap X_\eta, \tag{7.51}$$

with τ satisfying

$$9t \le \tau = 9t + i_{\text{fin}}\kappa_{10}\varepsilon m_0(t+1) \le 9t + 2m_0Mt = 9t + 2m_0Dt. \tag{7.52}$$

Let us first assume (7.51) and finish the proof of the proposition.

First note that using the above definitions, we have

$$e^{t/2} \le \#F_{\text{fin}} \le \#F_0 \le \beta^{-1}e^{4t} \le e^{5t}.$$

The assertion (7.50) and Lemma 7.3 imply that for all $w \in F_{\text{fin}}$,

$$\textstyle\sum_{w \neq w'} \|w - w'\|^{-\alpha} \le 4 \cdot (\#F_{\text{fin}})^{1+\varepsilon}.$$

This estimate together with (7.51) implies that part (1) in the proposition holds with $x_1 = y_{\text{fin}}$ and $F = F_{\text{fin}}$ if we choose R large enough so that $e^{-t/R} \ge 100(i_{\text{fin}} + 10)\beta$. This concludes the proof of Proposition 7.1 modulo the proof of (7.51).

To see that (7.51) holds, note that at every step, the element h_0 is of the form $a_{m_0\ell}u_{r_k}$ where $r_k \in [0, 1]$ and $\ell = \lfloor \kappa_{10}\varepsilon(t+1)\rfloor$. Now for all $0 \le k < i_{\text{fin}}$, we have

$$\mathcal{E}_{k+1} \subset \mathsf{B}^s_{2\beta} \cdot a_{m_0\ell}u_{r_k} \cdot \{u_{\bar r} : |\bar r| \le 2e^{-m_0\ell}\} \cdot \mathcal{E}_k, \tag{7.53}$$

where $\mathsf{B}^s_\varrho = \{u^-_s : |s| \le \varrho\} \cdot \{a_d : |d| \le \varrho\}$. To see this note that by (I-1), we have

$$\{\exp(w)x_1 : w \in F_{k+1}\} \subset \mathsf{B}^H_{100\beta^2} \cdot a_{m_0\ell}u_{r_k}.\mathcal{E}_k.$$

Now for every $|r| \le 1$, $\hat{\mathsf{h}} \in \mathsf{B}^H_\beta$ and $\mathsf{h} \in \mathsf{B}^H_{100\beta^2}$, we have $\hat{\mathsf{h}} u_r \mathsf{h} = \mathsf{h}' u_{r'}$ where $\mathsf{h}' \in \mathsf{B}^s_{2\beta}$ and $|r'| \le 2$; moreover, $u_{r'} a_{m_0\ell} = a_{m_0\ell} u_{e^{-m_0\ell} r'}$. Assuming $\ell \ge 5$, which may be guaranteed by taking t large, and using the definition

$$\mathcal{E}_{i+1} = \mathsf{E}.\{\exp(w)x_1 : w \in F_{i+1}\},$$

the inclusion in (7.53) follows.

Arguing similarly, (7.46) implies that

$$\mathcal{E}_0 \subset \mathsf{B}^s_{10\beta} \cdot a_{9t} \cdot \{u_r x_0 : r \in [0, 1.15]\}.$$

Using the fact that $a_{m_0\ell} \mathsf{B}^s_\varrho a_{-m_0\ell} \subset \mathsf{B}^s_\varrho$ and arguing inductively,

$$\mathcal{E}_{i+1} \subset \mathsf{B}^H_{100(i_{\mathrm{fin}}+10)\beta} \cdot (a_{m_0\ell} u_{\hat{r}_{i+1}} \mathsf{U}_{i+1}) \cdots (a_{m_0\ell} u_{\hat{r}_1} \mathsf{U}_1) \cdot \{a_{9t} u_r : |r| \le 2\}.x_0,$$

where $\hat{r}_k \in [0, 1]$ and $\mathsf{U}_k = \{u_{\bar{r}} : |\bar{r}| \le 100(k+10)\beta\}$. Moreover, for every $i \le i_{\max}$,

$$(a_{m_0\ell} u_{\hat{r}_{i+1}} \mathsf{U}_{i+1}) \cdots (a_{m_0\ell} u_{\hat{r}_1} \mathsf{U}_1) \subset a_{m_0(i+1)\ell} \cdot u_{\hat{r}} \cdot \{u_{\bar{r}} : |\bar{r}| \le 10^4 \beta\},$$

where $\hat{r} = \sum e^{-m_0(k-1)\ell} \hat{r}_k \in [0, 1.5]$.

This implies (7.51) except for the bound (7.52) on τ. To see the claimed bound on τ, note that

$$i_{\max}\ell \le \left(\tfrac{6M-3}{4\kappa_{10}\varepsilon} + 1\right)\kappa_{10}\varepsilon(t+1) \le 2Mt$$

which implies the bound on τ. □

8 Proof of the main theorem

In this section we will complete the proofs of Proposition 1.1 and Theorem 1.1.

8.1 Proof of Proposition 1.1

Let D_0 be as in Proposition 6.1, and choose $D \ge 2D_0$ so that $\delta/2 \le D_0/(D - D_0) \le \delta$.

Let $\eta_0 = 0.01\eta_X$, and let $0 < \eta < \eta_0$. Let $x_1 \in X_\eta$, and let t_0 be as in Proposition 7.1 applied with D and η.

Define t by $T = e^{(D-D_0)t}$, and let T_1 be so that $T \ge T_1$ implies $t \ge t_0$.

We may assume that Proposition 7.1(1) holds. Indeed, if Proposition 7.1(2) holds, then since $e^{D_0 t} = T^{D_0/(D-D_0)}$ and $\delta/2 \le D_0/(D-D_0) \le \delta$, Proposition 1.1(2) holds and the proof is complete.

Let $0 < \theta < 1/2$ be arbitrary. Apply Proposition 7.1(1) with $\varepsilon = 0.01\theta$ and $\alpha = 1 - \varepsilon$. Without loss of generality, we will further assume that T_1 is large enough so that $e^{-\varepsilon t/2} \le (2C_5C_7)^{-1}\eta^3$, this is motivated by (5.4).

By Proposition 7.1(1), there exists $R > 0$, depending on D and θ, so that the following holds. There exist $x_1 \in X_\eta$, some $9t \le \tau \le 9t + 2m_0Dt$ (where m_0 depends on θ as in (7.1)), and a subset $F \subset B_{\mathfrak{r}}(0, 1)$, containing 0, with $e^{t/2} \le \#F \le e^{5t}$, so that both of the following properties are satisfied.

$$\{\exp(w)x_1 : w \in F\} \subset \left(\mathsf{B}^H_{e^{-t/R}} \cdot a_\tau.\{u_r x_0 : |r| \le 4\}\right) \cap X_\eta \quad \text{and} \tag{8.1a}$$

$$\textstyle\sum_{w' \ne w} \|w - w'\|^{-\alpha} \ll (\#F)^{1+\varepsilon} \qquad \text{for all } w \in F, \tag{8.1b}$$

where the implied constant depends on X.

Now apply Proposition 5.1 with η, ε, $\alpha = 1 - \varepsilon$, x_1, and F; note that (5.4) is satisfied since $\#F \ge e^{t/2}$. Let

$$x_2 \in X_\eta \cap a_{|\log b_1|}.\{u_r \exp(w)x_1 : |r| \le 2, w \in F\}, \tag{8.2}$$

$I \subset [0, 1]$, $b_1 > 0$, and the probability measure ρ on I be as in that proposition. In particular, we have

$$e^{-5t} \le (\#F)^{-\frac{2+6\varepsilon}{2+21\varepsilon}} \le b_1 \le (\#F)^{-\varepsilon}, \tag{8.3}$$

and the following hold

$$\rho(J) \le C'_\varepsilon |J|^{\alpha - 30\varepsilon} \qquad \text{for all } |J| \ge (\#F)^{\frac{-15\varepsilon}{2+21\varepsilon}} \tag{8.4a}$$

$$v_s x_2 \in \mathsf{B}^G_{Cb_1} \cdot a_{|\log b_1|}.\{u_r \exp(w)x_1 : |r| \le 2, w \in F\} \quad \text{for all } s \in I, \tag{8.4b}$$

where C is an absolute constant.

Set $\kappa := \frac{\varepsilon}{4D_0} = \frac{\theta}{400D_0}$. Since $\#F \ge e^{t/2}$, we have

$$(\#F)^{\frac{-15\varepsilon}{2+21\varepsilon}} \le (\#F)^{-\varepsilon} \le e^{-\varepsilon t/2} \le T^{-\delta\varepsilon/4D_0} = T^{-\delta\kappa}; \tag{8.5}$$

recall that $\delta/2 \le D_0/(D - D_0) \le \delta$ and $T = e^{(D-D_0)t}$.

Combining (8.5) and equation (8.4a), we conclude that

$$\rho(J) \le C'_\varepsilon |J|^{\alpha - 30\varepsilon} \le C'_\varepsilon |J|^{1-\theta}, \quad \text{for all intervals } J \text{ with } |J| \ge T^{-\delta\kappa}. \tag{8.6}$$

This establishes Proposition 1.1(1)(a) if we put $C_\theta = C'_\varepsilon$.

Let us now turn to the proof of Proposition 1.1(1)(b). We first claim that

$$\{u_r \exp(w)x_1 : |r| \le 2, w \in F\} \subset \mathsf{B}^s_{10\varrho} \cdot a_\tau \cdot \{u_r x_0 : |r| \le 9/2\}, \tag{8.7}$$

where $\varrho = e^{-t/R}$ and $\mathsf{B}^s_\varrho = \{u^-_d : |d| \le \varrho\} \cdot \{a_\ell : |\ell| \le \varrho\}$. To see this, first note that using (8.1a), we have

$$\{\exp(w)x_1 : w \in F\} \subset \mathsf{B}^H_\varrho \cdot a_\tau \cdot \{u_r x_0 : |r| \le 4\}.$$

Now for every $|r| \le 2$ and $h \in \mathsf{B}^H_\varrho$, we have $u_r h = h' u_{r'}$ where $h' \in \mathsf{B}^s_{10\varrho}$ and $|r'| \le 3$; moreover, $u_{r'} a_\tau = a_\tau u_{e^{-\tau} r'}$. The claim follows as $\tau \ge 2$.

Combining (8.7), (8.4b), and (8.2) for all $s \in I \cup \{0\}$ we have

$$\begin{aligned} v_s x_2 &\in \mathsf{B}^G_{Cb_1} \cdot a_{|\log b_1|} \cdot \{u_r \exp(w)x_1 : |r| \le 2, w \in F\} \\ &\in \mathsf{B}^G_{Cb_1} \cdot a_{|\log b_1|} \cdot \mathsf{B}^s_{10\varrho} \cdot a_\tau \cdot \{u_r x_0 : |r| \le 9/2\}. \end{aligned} \tag{8.8}$$

By the definition of $\mathsf{B}^s_{10\varrho}$ above, we conclude that

$$a_{|\log b_1|} \mathsf{B}^s_{10\varrho} a_{-|\log b_1|} \subset \{u^-_d : |d| \le b_1\} \cdot \{a_\ell : |\ell| \le 10\varrho\}.$$

This and (8.8) imply that

$$v_s x_2 \in \mathsf{B}^G_{C'b_1} \cdot \Big(\{a_\ell : |\ell| \le 10\varrho\} \cdot a_{\tau + |\log b_1|} \cdot \{u_r : |r| \le 9/2\}\Big).x_0. \tag{8.9}$$

Recall that $b_1 \le (\#F)^{-\varepsilon} \le e^{-\varepsilon t/2} \le T^{-\varepsilon\delta/4D_0}$ and $\varrho = e^{-t/R}$. Moreover, note that the bound $e^{-6t} \le b_1$ in (8.3) and $\tau \le 9t + 2m_0 Dt$ imply

$$e^{(\tau + |\log b_1|)/2} \le e^\tau \le e^{9t + 2m_0 Dt} \le T^{A'-1},$$

for A' depending only on θ. Hence, in view of (8.9), we have

$$d_X\Big(v_s x_2, B_P\big(e, T^{A'}\big).x_0\Big) \ll_X T^{-\delta\varepsilon/4D_0},$$

for all $s \in I \cup \{0\}$.

The above and (8.5) finish the proof of the proposition if we let $y_0 = x_2$ and $\kappa_2 = \frac{\varepsilon}{4D_0} = \frac{\theta}{400D_0}$.

8.2 Proof of Theorem 1.1

Let $\theta = \varepsilon_0/2$ where ε_0 is given by Proposition 4.2.

Apply Proposition 1.1 with $x_0, \theta, \eta = 10^{-4}\eta_X$, and the given δ. Let $T > T_1$ where T_1 is as in Proposition 1.1.

If Proposition 1.1(2) holds, then Theorem 1.1(2) holds and we are done. Therefore, let us assume that Proposition 1.1(1) holds. Let y_0, I, and ρ be as in Proposition 1.1(1).

Let $0 < \varrho < 0.1\eta_X$, and let $z \in X_\varrho$. There is a function $f_{\varrho,z}$ supported on $\mathsf{B}^G_{0.1\varrho}.z$ with $\int f_{\varrho,z}\,\mathrm{d}m_X = 1$ and $\mathcal{S}(f_{\varrho,z}) \le \varrho^{-N}$, where N is absolute.

Let $b = T^{-\delta\kappa_2}$, and let $t = |\log b|/4$. In view of Proposition 1.1(1), ρ satisfies (4.6) with C_θ.

Apply Proposition 4.2, with $f = f_{\varrho,z}$ for $\varrho = e^{-\kappa_6 t/2N}$. Then

$$\left| \iint f(a_t u_r v_s.y_0)\,\mathrm{d}\rho(s)\,\mathrm{d}r - 1 \right| \ll_{C_\theta} \mathcal{S}(f)e^{-\kappa_6 t} \ll_{C_\theta} e^{-\kappa_6 t/2};$$

where we used $\eta = 10^{-4}\eta_X$, hence the dependence on η in Proposition 4.2 can be absorbed in the implicit constant.

Assuming T is large enough, depending on θ, the right side of the above is $< 1/2$. Thus $a_t u_r v_s.y_0 \in \mathrm{supp}(f)$ for some $r \in [0, 1]$ and $s \in I$.

Let $\kappa_{11} = \kappa_6/8N$. The above thus implies that

$$d_X\Big(z, a_t.\Big\{u_r v_s y_0 : r \in [0,1], s \in I\Big\}\Big) \ll b^{\kappa_{11}} \tag{8.10}$$

for all $z \in X_{b^{\kappa_{11}}}$.

Moreover, by Proposition 1.1(1), we have

$$d_X\Big(u_r v_s.y_0, \Big(u_r \cdot B_P(e, T^{A'})\Big).x_0\Big) \le C_2' b,$$

for all $s \in I \cup \{0\}$ and $r \in [0, 1]$. Note also that if $z, z' \in X$ satisfy, $d(z, z') \le C_2' b$, then $d_X(a_t z, a_t z') \ll b^{1/2}$. In consequence,

$$d_X\Big(a_t.\Big\{u_r v_s y_0 : r \in [0,1], s \in I\Big\}, B_P(e, T^{A'+1}).x_0\Big) \ll b^{1/2}, \tag{8.11}$$

where we used

$$a_t \cdot \{u_r : r \in [0,1]\} \cdot B_P(e, T^{A'}) \subset B_P(e, T^{A'+1}),$$

which in turn follows from $t = |\log b|/4$ and $b = T^{-\delta\kappa_2}$.

Combining (8.10) and (8.11), we conclude that

$$d_X(z, B_P(e, T^{A'+1}).x_0) \ll b^{\kappa_{11}} = T^{-\delta\kappa_2\kappa_{11}}$$

for all $z \in X_{b^{\kappa_{11}}}$, where the implied constant depends on X. This implies Theorem 1.1(1) with $\kappa_1 = \kappa_2\kappa_{11}$.

As was remarked in Sect. 4, κ_X in (4.1) is absolute if Γ is a congruence subgroup, see [9,13,29]. Hence, if Γ is assumed to be a congruence subgroup, then A and κ_1 only depend on Γ via (6.2). $\square$

9 Proof of Theorem 1.2

Let η_X be as in Proposition 3.2 and C_7 as in Proposition 3.1. Define

$$\mathsf{C}_X = \eta_X^{-1}\,\mathrm{vol}(G/\Gamma)\,e^{C_7}, \tag{9.1}$$

where $\mathrm{vol}(G/\Gamma)$ is computed using the Riemannian metric d, see also (4.2).

For $0 < \alpha < 1$ choose an $m_\alpha > 0$ as in (2.12), i.e., m_α satisfies that

$$\int_0^1 \|a_{m_\alpha} u_r w\|^{-\alpha}\,\mathrm{d}r \le e^{-1}\|w\|^{-\alpha} \qquad \text{for all } w \in \mathfrak{g}. \tag{9.2}$$

In this section, the notation $a \ll_X b$ means $a \le L\mathsf{C}_X^L b$ where L is an absolute constant. Similarly, $a \ll_{X,\alpha} b$ means

$$a \le L\mathsf{C}_X^L e^{Lm_\alpha}\,b, \tag{9.3}$$

where L is an absolute constant. Define $a \gg_X b$ and $a \gg_{X,\alpha}$ accordingly.

Throughout this section, $Y = Hx$ is a periodic orbit. Let μ_{Hx} denote the probability H-invariant measure on Hx. We put $\mathrm{vol}(Y) = \mathsf{v}$. In view of Lemma 3.3, we have $\mathsf{v} \gg_X 1$. The following proposition is our replacement for Proposition 7.1 in the setting at hand.

Proposition 9.1 *Let $0 < \alpha < 1$. There exists $y_0 \in Y$ and a subset $F \subset B_{\mathfrak{r}}(0, 1)$, containing 0, with $\#F \gg_X \mathsf{v}$ so that both of the following properties are satisfied:*

(9.1-a) $\left\{\exp(w)y_0 : w \in F\right\} \subset Y \cap X_{\mathrm{cpt}}$, *see Sect. 3.1 for the definition of* X_{cpt}.

(9.1-b) $\sum_{w' \neq w} \|w - w'\|^{-\alpha} \ll_{X,\alpha} \#F$ *for all $w \in F$ where the summation is over $w' \in F$.*

The general strategy in proving Proposition 9.1 is similar to the strategy we used to prove Proposition 7.1. However, the argument simplifies significantly thanks to the fact that Y is equipped with an H-invariant probability measure. In particular, we do not require Proposition 6.1, hence Γ is *not* assumed to be an arithmetic lattice in this section, see Proposition 9.2.

For every $0 < \delta \leq 1$ and every $y \in Y$, put

$$I(y, \delta) = \Big\{ w \in \mathfrak{r} : 0 < \|w\| < \delta \,\mathrm{inj}(y) \text{ and } \exp(w)y \in Y \Big\},$$

see also (7.3). We will write $I(y) = I(y, \delta_0)$ where

$$\delta_0 = e^{-3-C_7} \min\{\mathrm{inj}(x) : x \in X_{\mathrm{cpt}}\}, \tag{9.4}$$

see (9.1); recall also that $\mathrm{inj}(x) \leq 1$ for all $x \in X$.

We need the following lemma.

Lemma 9.1 *There exists $C_{15} \ll_X 1$ so that*

$$\#I(y) \leq C_{15}\mathsf{v}$$

for every $y \in Y$.

Proof This is proved for $G = \mathrm{SL}_2(\mathbb{C})$ in [48, Lemma 8.13], see also [24, §8].

The same argument applies in the case of $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ if we replace [48, Lemma 8.4] by Proposition 3.2. We sketch the proof for the sake of completeness.

By virtue of Lemma 7.2, for all $y \in X_{\mathrm{cpt}}$, we have

$$\#I(y, 1) \ll_X \mathsf{v}.$$

Suppose now that $y \in Y \setminus X_{\mathrm{cpt}}$, and let $t = |\log \mathrm{inj}(y)| + C_7$. By Proposition 3.2, there exists $|r| \leq 1$ so that $a_t u_r y \in X_{\mathrm{cpt}}$. Moreover, for all $\|w\| < \delta_0 \,\mathrm{inj}(y)$, see (9.4), we have

$$\|a_t u_r w\| \leq 3e^t \|w\| = 3e^{C_7} \mathrm{inj}(y)^{-1} \|w\| < 0.5 \,\mathrm{inj}(a_t u_r y).$$

This and the fact that Y is invariant under H imply that if $w \in I(y) = I(y, \delta_0)$, then $a_t u_r w \in I(a_t u_r y, 1)$.

The above estimate also implies that the map $w \mapsto a_t u_r w$ is an injective map from $I(y)$ into $I(a_t u_r y, 1)$. Consequently,

$$\#I(y) \leq \#I(a_t u_r y, 1) \ll_X \mathsf{v}.$$

The proof is complete. $\square$

Let $0 < \alpha < 1$, and define a Margulis function $f_Y : Y \to [2, \infty)$ by

$$f_Y(y) = \begin{cases} \sum_{w \in I(y)} \|w\|^{-\alpha} & \text{if } I(y) \neq \emptyset \\ \mathrm{inj}(y)^{-\alpha} & \text{otherwise} \end{cases}.$$

Let m_α be as in (9.2). Define the probability measure ν on H by the property that for every $\varphi \in C_c(X)$

$$\nu * \varphi(y) = \int_0^1 \varphi(a_{m_\alpha} u_r y)\, \mathrm{d}r.$$

The following proposition may be thought of as our replacement for Proposition 6.1.

Proposition 9.2 *There exists $C_{16} \ll_{X,\alpha} 1$ so that*

$$\int f_Y(y)\, \mathrm{d}\mu_Y(y) \leq C_{16} \cdot \mathsf{v}.$$

The following lemma is analogue of Lemma 7.1, and will be used in the proof of Proposition 9.2.

Lemma 9.2 *There exists $C_{17} \ll_{X,\alpha} 1$ so that for all $\ell \in \mathbb{N}$ and all $y \in Y$, we have*

$$\nu^{(\ell)} * f_Y(y) \leq e^{-\ell} f_Y(y) + C_{17}\mathsf{v} \textstyle\sum_j^{\ell} e^{j-\ell} \nu^{(j)} * \mathrm{inj}(y)^{-\alpha}. \tag{9.5}$$

Proof Note that $\mathrm{supp}(\nu) \subset \{h \in H : \|h\| \leq e^{2m_\alpha+1}\}$. Let $C \geq 1$ be so that

$$\|\operatorname{Ad}(h)w\| \leq C\|w\|$$

for all h with $\|h\| \leq e^{2m_\alpha+1}$ and all $w \in \mathfrak{g}$. Increasing C if necessary, we also assume that $\mathrm{inj}(z)/C \leq \mathrm{inj}(hz) \leq C\,\mathrm{inj}(z)$ for all such h and all $z \in X$. Arguing as in the proof of Lemma 7.1, there exists some C so that

$$\nu * f_Y(y) \leq e^{-1} \cdot f_Y(y) + C \cdot \nu * \psi(y)$$

for all $y \in Y$, where $\psi(y) = \max\{1, \#I(y)\} \cdot \mathrm{inj}(y)^{-\alpha}$. This and Lemma 9.1 imply that

$$\nu * f_Y(y) \leq e^{-1} \cdot f_Y(y) + C_{17}\mathsf{v} \cdot \left(\nu * \mathrm{inj}(y)^{-\alpha}\right) \tag{9.6}$$

with $C_{17} = CC_{15}$. Iterating (9.6), we get (9.5). □

Proof of Proposition 9.2 The fact that estimates similar to Lemma 9.2 imply integrability is by now a standard fact, see e.g. [21, §5] or [24, Lemma 11.1]; we recall the argument. In view of Proposition A.3, we have

$$\int_H \mathrm{inj}(hx)^{-\alpha}\, \mathrm{d}\nu^{(n)}(h) \leq e^{-n}\, \mathrm{inj}^{-\alpha}(x) + B$$

for all $n \in \mathbb{N}$ where $B \ll_X 1$. This and Lemma 9.2 imply that

$$\limsup \nu^{(n)} * f_Y(y) \le 1 + 2C_{17}\mathsf{v}B. \tag{9.7}$$

Note that $\mathrm{supp}(\nu^{(n)}) \subset \{a_{m_\alpha n} u_r : |r| \le 4\}$. This, together with the fact that (H, μ_Y) is mixing, implies that μ_Y is ν-ergodic. Thus by Chacon-Ornstein theorem, for every $\varphi \in L^1(Y, \mu_Y)$ and μ_Y-a.e. $y \in Y$, we have $\frac{1}{N+1}\sum_{n=0}^N \nu^{(n)} * \varphi(y) \to \int \varphi \, \mathrm{d}\mu_Y$.

For every $k \in \mathbb{N}$, put $\varphi_k = \min\{f_Y, k\}$. There exists a full measure set Y_0 so that for every $y \in Y_0$ and every k, there exists some $N_{k,y}$ so that if $N \ge N_{k,y}$, then $\frac{1}{N+1}\sum_{n=0}^N \nu^{(n)} * \varphi_k(y) \ge 0.5 \int \varphi_k \, \mathrm{d}\mu_Y$.

Let $y \in Y_0$, then the above estimate and (9.7), applied with y, imply that $\int \varphi_k \, \mathrm{d}\mu_Y \le 2(1+2C_{17}\mathsf{v}B)$ for all k. Using Lebesgue's monotone convergence theorem, we conclude that

$$\int f_Y \, \mathrm{d}\mu_Y \le 2(1 + 2C_{17}\mathsf{v}B).$$

The claim follows as $\mathsf{v} \gg_X 1$. $\square$

Proof of Proposition 9.1 Put $\eta = 0.1\eta_X$ where η_X is as in Proposition 3.2. Recall from Lemma 3.3 that

$$\mu_Y(X_{2\eta}) \ge 0.9. \tag{9.8}$$

As was done in Lemma 7.5, we will first convert the information in Proposition 9.2 into a pointwise estimate at most points. Let

$$Y'' = \{y \in Y : f_Y(y) \le 100C_{16}\mathsf{v}\}. \tag{9.9}$$

Then by Proposition 9.2, we have $\mu_Y(Y \setminus Y'') \le 0.01$.

Let $Y' = Y'' \cap X_{2\eta}$, and let $\beta = \eta^2 = 0.01\eta_X^2$. The above and (9.8) imply that $\mu_Y(Y') \ge 0.9$. Let $\{\mathsf{B}^G_{\beta^2}.z_j : z_j \in X_{2\eta}, j \in \mathcal{J}\}$ be a covering of $X_{2\eta}$ so that $\#\mathcal{J} \ll_X 1$. Then there exists some $c \gg_X 1$ and some j_0 so that

$$\mu_Y(\mathsf{B}^G_{\beta^2}.z_{j_0} \cap Y') \ge c. \tag{9.10}$$

Recall that Y is H-invariant and $gz_j \in X_{\mathrm{cpt}}$ for all j and $\|g - I\| \le 2$, see Sect. 3.1 where X_{cpt} is defined. Let $y_0 \in \mathsf{B}^G_{\beta^2}.z_{j_0} \cap Y'$. As was done in Lemma 7.7, let $F_1 \subset B_{\mathfrak{r}}(0, 2\beta^2)$ be so that

$$\mathsf{B}^G_{\beta^2}.z_{j_0} \cap Y' \subset \bigcup_{w \in F_1} \mathsf{B}^H_\beta.\exp(w)y_0.$$

Then $\#F_1 \geq c\eta^{-3}\mathsf{v}$. Put

$$\mathcal{E}_1 = \mathsf{E} \cdot \{\exp(w)y_0 : w \in F_1\} \subset Y \cap X_{\mathrm{cpt}};$$

recall that $\mathsf{E} = \mathsf{B}^H_\beta \cdot \Big\{u_r : |r| \leq 0.1\eta\Big\}$.

Recall the definition $f_{\mathcal{E}_1}$ from (7.4). There exists $C' \ll_{X,\alpha} 1$ so that

$$f_{\mathcal{E}_1}(e, z) \leq f_Y(z) \leq C'\mathsf{v} \quad \text{for all } z \in \mathcal{E}_1 \tag{9.11}$$

To see this, note that by the definition of f_Y, for every $h \in H$ with $\|h - I\| \leq 1$ and all $y \in X_\eta \cap Y$, we have $f_Y(hy) \leq f_Y(y) + \bar{C}\mathsf{v}$ where $\bar{C} \ll_X 1$. Now for every $z \in \mathcal{E}_1$, there exists $y \in Y' \subset Y''$ and some $h \in H$ with $\|h - I\| \leq 10\eta^2$ so that $z = hy$. This implies the claim in view of the definition of Y'' in (9.9). Alternatively, (9.11) can be seen by letting $\ell = 0$ in the proof of the sublemma in Lemma 7.7, see in particular (7.32).

Now (9.11) and Lemma 7.3 imply that

$$\sum_{w' \neq w} \|w - w'\|^{-\alpha} \leq C\mathsf{v},$$

where the summation is over $w' \in F_1$ and $C \ll_{X,\alpha} 1$.

The proposition holds with y_0 and $F = F_1$. □

9.1 Proof of Theorem 1.2

The proof goes along the same lines as the proof of Theorem 1.1 if we replace Proposition 7.1 with Proposition 9.1 as we now explicate.

Let $\varepsilon = 0.0005\varepsilon_0$ and $\alpha = 1 - \varepsilon$ where ε_0 is given by Proposition 4.2. By Proposition 9.1, the conditions in Proposition 5.1 holds with $y_0 \in Y \cap X_{\mathrm{cpt}}$, F, α, and $\eta = 0.1\eta_X$.

Recall that $\#F \gg_X \mathsf{v}$. We assume v is large enough so that

$$(\#F)^{-\varepsilon} \leq (2C_5C_7)^{-1}\eta^3.$$

Then by Proposition 5.1, there exist $y_1 \in X_\eta$, a finite subset $I \subset [0, 1]$, and some $b_1 > 0$ with

$$\mathsf{v}^{-\frac{2+6\varepsilon}{2+21\varepsilon}} \ll_X (\#F)^{-\frac{2+6\varepsilon}{2+21\varepsilon}} \leq b_1 \leq (\#F)^{-\varepsilon} \ll_X \mathsf{v}^{-\varepsilon}, \tag{9.12}$$

so that both of the following two statements hold true:

(1) The set I supports a probability measure ρ which satisfies

$$\rho(J) \le C'_\varepsilon \cdot |J|^{\alpha - 30\varepsilon}$$

for all intervals J with $|J| \ge (\#F)^{\frac{-15\varepsilon}{2+21\varepsilon}}$, where $C'_\varepsilon \ll \varepsilon^{-\star}$ for absolute implied constants.

(2) There is an absolute constant $C \ll_X 1$, so that for all $s \in I$, we have

$$\begin{aligned} v_s y_1 &\in \mathsf{B}^G_{Cb_1} \cdot \Big(a_{|\log b_1|} \cdot \{u_r : |r| \le 2\} \Big).\{\exp(w) y_0 : w \in F\} \\ &\subset \mathsf{B}^G_{Cb_1}.Y. \end{aligned}$$

For the last inclusion in (2) we used (9.1-a) and the H-invariance of Y.

In particular, part (2) and $b_1 \le (\#F)^{-\varepsilon}$ imply that

$$d_X\Big(v(s)y_1, Y\Big) \le C' \mathsf{v}^{-\varepsilon} \qquad \text{for all } s \in I, \tag{9.13}$$

where $C' \ll_{X,\alpha} 1$.

The proof of Theorem 1.2 is now completed as the proof of Theorem 1.1 if we replace Proposition 1.1 with part (1) above and (9.13), see Sect. 8.2.

We note that

$$C_3 \ll_{X,\alpha} 1 \quad \text{and} \quad \kappa_3 = c\kappa_6\varepsilon, \tag{9.14}$$

where the notation $\ll_{X,\alpha}$ is defined in (9.3), c is an absolute constant, and κ_6 is as in Proposition 4.2; we also used the fact that $C_{10} \ll_X 1$, see Proposition 4.2.

Note that κ_X in (4.1), and hence κ_3, is absolute if Γ is congruence. □

9.2 Proof of Theorem 1.3

Let $\Gamma \subset \mathrm{SL}_2(\mathbb{C})$ be as in the statement. As was mentioned prior to Theorem 1.3, a totally geodesic plane in M lifts to a periodic orbit of $H = \mathrm{SL}_2(\mathbb{R})$ in $X = G/\Gamma$.

Recall from Sect. 3.1 that $X \setminus X_{\eta_X}$ is a disjoint union of finitely many cusps. Let $M_0 \subset M$ denote the image of X_{η_X} in M. Then $M \setminus M_0$ is a disjoint union of finitely many (possibly none) cusps.

Let $\eta_1 > 0$ be so that for $i = 1, 2$ there exists $x_i \in X_{\eta_0}$ such that $\mathsf{B}^G_{\eta_1}.x_i$ projects into the interior of $N_i \cap M_0$. In view of [48, Thm. 1.5], applied with $s = 1/2$, we have $\eta_1 \gg_X \mathrm{area}(\Sigma)^{-4}$ where $\Sigma = \partial N_1 = \partial N_2$.

Thus, Theorem 1.2 implies that if Hx is a periodic orbit which satisfies

$$C_3 \mathrm{vol}(Hx)^{-\kappa_3} \le 0.5 \min\{\eta_1, \eta_X\}, \tag{9.15}$$

then $Hx \cap \mathsf{B}^G_{\eta_1}.x_i \neq \emptyset$, for $i = 1, 2$. Therefore, the corresponding plane crosses Σ.

Let us now assume that S is a plane which crosses Σ. By [25, Thm. 4.1], see also [3, Prop. 12.1], S intersects Σ orthogonally. It is shown in [25, Prop 5.1] that one can construct an explicit open set O of the unit tangent bundle of M which projects into the 1-neighborhood of M_0 and does not intersect such an S — indeed this set is constructed using a tubular neighborhood of $\Sigma \cap M_0$.

Let η_2 and $x \in X$ be so that $\mathsf{B}^G_{\eta_2}.x$ projects into O. In view of [48, Thm. 1.5], applied with $s = 1/2$, and the construction in [25, Prop 5.1], we have $\eta_2 \gg_X \mathrm{area}(\Sigma)^{-4}$.

Note that $Hx \cap \mathsf{B}^G_{\eta_2}.x = \emptyset$. However, by Theorem 1.2 again, if

$$C_3 \mathrm{vol}(Y)^{-\kappa_3} \leq 0.5\eta_2,$$

then $Hx \cap \mathsf{B}^G_{\eta_2}.x \neq \emptyset$.

This and (9.15) thus imply that

$$\mathrm{vol}(Hx) \leq \left(\frac{2C_3}{\min\{\eta_X, \eta_1, \eta_2\}}\right)^{1/\kappa_3} \ll_X \mathrm{area}(\Sigma)^{4/\kappa_3} C_3^{1/\kappa_3}.$$

Moreover, in view of [48, Cor. 10.7], the number of periodic H-orbits with $\mathrm{vol}(Hx) \leq T$ is $\ll_X T^6$.

When $G = \mathrm{SL}_2(\mathbb{C})$ (which is the case here), $C_7 \ll |\log \eta_X|$ for an absolute implied constant; see the proof of Proposition 3.1. Moreover, in view of Lemma 2.4 and the fact that $\alpha = 1 - 0.0005\varepsilon_0$, we have $e^{m_\alpha} \ll \kappa_X^{-\star}$ for absolute implied constants (see Proposition 4.2).

The proof is thus complete in view of the above, (9.14), and (9.3). $\square$

Acknowledgements We would like to thank the Hausdorff Institute for its hospitality during the winter of 2020. A.M. would like to thank the Institute for Advanced Study for its hospitality during the fall of 2019 where parts of this project were carried out. The authors would like to thank Gregory Margulis and Nimish Shah for many discussions about effective density, and Joshua Zahl for helpful communications regarding projections theorems. We would also like to thank Zhiren Wang with whom we discussed related questions. We thank the anonymous referees for their helpful comments.

Appendix A: Proof of Proposition 3.1, Case 2

In this section we complete the proof of Proposition 3.1. Recall that we are left with the case where $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ and Γ is irreducible.

By a theorem of Selberg [57], we have the following: up to automorphisms of G, irreducible non-uniform lattices in $\mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ are commensurable to $\mathrm{SL}_2(\mathcal{O})$ where $\mathcal{O}$ is the ring of integers in a totally real quadratic extension $L/\mathbb{Q}$.

Passing to a finite index subgroup, we may assume that $\Gamma \subset \mathrm{SL}_2(\mathcal{O})$. Since the statement of Proposition 3.1 is insensitive to passing to a finite index subgroup we may (and will) assume $\Gamma = \mathrm{SL}_2(\mathcal{O})$. By fixing a $\mathbb{Z}$-basis for $\mathcal{O}$ one can now identify

$$G = \mathbf{G}(\mathbb{R}) \quad \text{and} \quad \Gamma = \mathbf{G}(\mathbb{Z}),$$

where $\mathbf{G} = \mathrm{Res}_{L/\mathbb{Q}}(\mathrm{SL}_2)$, the restriction of scalars from L to $\mathbb{Q}$. This choice of $\mathbb{Z}$ basis induces a canonical identification between $\mathbf{G}(\mathbb{Q})$ and $\mathrm{SL}_2(L)$ and in the sequel we shall implicitly identify these two groups.

Let $\mathbf{B} \subset \mathrm{SL}_2$ denote the group of upper triangular matrices in SL_2 and put $\mathbf{P} = \mathrm{Res}_{L/\mathbb{Q}}(\mathbf{B})$. Then $\mathbf{P}$ is a minimal and maximal $\mathbb{Q}$-parabolic subgroup of $\mathbf{G}$. By a theorem of Borel and Harish-Chandra, he action of Γ on $\mathbf{P}(\mathbb{Q})\backslash\mathbf{G}(\mathbb{Q})$ has finitely many orbits; let $\Xi \subset \mathbf{G}(\mathbb{Q})$ be a finite subset which contains exactly one representative for each orbit (we always assume Ξ contains the identity element). Then

$$\mathbf{G}(\mathbb{Q}) = \mathbf{P}(\mathbb{Q})\,\Xi\Gamma, \tag{A.1}$$

and if $\gamma\xi_1\mathbf{P}(\mathbb{Q})\xi_1^{-1}\gamma^{-1} = \xi_2\mathbf{P}(\mathbb{Q})\xi_2^{-1}$ where $\gamma \in \Gamma$ and $\xi_i^{-1} \in \Xi$, then $\xi_1 = \xi_2$.

In the case at hand, $\mathfrak{g} = \mathrm{Lie}(G) = \mathfrak{sl}_2(\mathbb{R}) \oplus \mathfrak{sl}_2(\mathbb{R})$, moreover, $\mathfrak{g}$ is equipped with the $\mathbb{Q}$-structure:

$$\mathfrak{g}_{\mathbb{Q}} = \mathfrak{sl}_2(L) \subset \mathfrak{g}.$$

We will also write $\mathfrak{g}_{\mathbb{Z}}$ for $\mathfrak{sl}_2(\mathcal{O})$; then $\mathfrak{g}_{\mathbb{Z}}$ is a lattice in $\mathfrak{g}$.

Note that $\mathcal{O}^\times\mathfrak{g}_{\mathbb{Z}} = \mathfrak{g}_{\mathbb{Z}}$. Recall the following elementary fact: there exists some $c = c_L$ so that the following holds. For every $w = (w_1, w_2) \in \mathfrak{g}$ with $\|w_1\|\|w_2\| \neq 0$, there exists some $\mathsf{s} \in \mathcal{O}^\times$ so that

$$c^{-1}\Big(\|w_1\|\|w_2\|\Big)^{1/2} \le \|p_i(\mathsf{s}w)\| \le c\Big(\|w_1\|\|w_2\|\Big)^{1/2}, \tag{A.2}$$

for $i = 1, 2$, where p_i denotes the projection onto the i-th components, see e.g. [39, Lemma 8.6].

Let $N = R_u(\mathbf{P}(\mathbb{R}))$, i.e. N is the unipotent radical of $\mathbf{P}(\mathbb{R})$. We fix a basis $\{v_1, v_2\}$ for $\mathrm{Lie}(N)$ consisting of primitive integral vectors as follows. Write $L = \mathbb{Q}[\sqrt{\beta}]$; put $v_1 = \Big(E_{12}, E_{12}\Big)$ and $v_2 = \Big(\sqrt{\beta}E_{12}, -\sqrt{\beta}E_{12}\Big)$ where E_{12} denotes the elementary matrix with 1 at the $(1, 2)$-entry, and define

$$v := v_1 \wedge v_2 \in \wedge^2\mathfrak{g}.$$

Since $v \in \wedge^2\mathfrak{g}_{\mathbb{Z}}$, for any $g \in \mathbf{G}(\mathbb{Q})$, we have $\Gamma g.v$ is contained in the set of rational vectors in $\wedge^2\mathfrak{g}$ whose denominators (with respect to the $\mathbb{Z}$-structure given by $\mathfrak{g}_{\mathbb{Z}}$) are bounded in terms of g. In particular, $\Gamma g.v$ is a discrete and closed subset of $\wedge^2\mathfrak{g}$.

Note that for any $g = (g_1, g_2) \in G$, we have

$$\begin{aligned} gv &= (gv_1) \wedge (gv_2) \\ &= -2\sqrt{\beta}\Big(g_1E_{12}, 0\Big) \wedge \Big(0, g_2E_{12}\Big). \end{aligned} \tag{A.3}$$

Define $\omega : G/\Gamma \to [2, \infty)$ as follows:

$$\omega(g\Gamma) = \max\Big\{2, \max\Big\{\|g\gamma\xi.v\|^{-1} : \xi \in \Xi^{-1}, \gamma \in \Gamma\Big\}\Big\}. \tag{A.4}$$

We have the following analogue of Lemmas 3.1 and 3.2. In the case at hand, this result is a consequence of the fact that the $\mathbb{Q}$-rank of $\mathbf{G}$ is 1 — recall that $\mathbf{P}$ is a minimal and maximal $\mathbb{Q}$-parabolic subgroup of $\mathbf{G}$.

Lemma A.1 *Let the notation be as above.*

(1) There exists $C = C(\Gamma) \geq 2$ so that the following holds. Let $g\Gamma \in X$. If $\omega(g\Gamma) \geq C$, then there is $\xi_0 \in \Xi^{-1}$ and $\gamma_0 \in \Gamma$ so that $\|g\gamma_0\xi_0.v\|^{-1} = \omega(g\Gamma)$ and

$$\|g\gamma\xi.v\| > 1/C, \quad \textit{for all} \ \ (\xi, \gamma) \ \ \textit{so that} \ \ \gamma\xi.v \neq \gamma_0\xi_0.v.$$

(2) There exists C_{18} so that the following holds. Let $0 < \varrho, \eta < 1$, $t > 0$, and $g \in G$. Let $I \subset \mathbb{R}$ be an interval of length at least η. Then

$$\Big|\Big\{r \in I : \|a_tu_rg.v\| \leq e^{2t}\eta^4\varrho^4\|gv\|\Big\}\Big| \leq C_{18}\varrho|I|.$$

Proof As we mentioned above, there is some $M \in \mathbb{N}$ so that $\Gamma\Xi^{-1}.v_i \subset \frac{1}{M}\mathfrak{g}_{\mathbb{Z}}$.

Let $0 < \delta < 1$ be a small number which will be explicated later. Suppose there are $\gamma\xi.v \neq \gamma'\xi'.v$ so that

$$\|g\gamma\xi.v\| < \delta \quad \text{and} \quad \|g\gamma'\xi'.v\| < \delta. \tag{A.5}$$

We first show that $\gamma\xi.v \notin \mathbb{R}.\gamma'\xi'v$. Assume contrary to this claim that $\gamma\xi.v = \lambda\gamma'\xi'v$ for some $\lambda \in \mathbb{R}$. Then since $\mathbf{P}(\mathbb{R})$ is the projective stabilizer of v, we conclude that

$$\gamma\xi\mathbf{P}(\mathbb{R})\xi^{-1}\gamma^{-1} = \gamma'\xi'\mathbf{P}(\mathbb{R})\xi'^{-1}\gamma'^{-1}.$$

This in view of the choice of Ξ, see the discussion following (A.1), implies that $\xi = \xi'$. Thus, since $\mathbf{P}(\mathbb{R})$ is its own normalizer in $\mathbf{G}(\mathbb{R})$, $\gamma^{-1}\gamma' \in \xi\mathbf{P}(\mathbb{R})\xi^{-1}$. We conclude that $\lambda = N_{L/\mathbb{Q}}(\mathsf{s}^2)$ for a unit in $\mathsf{s} \in \mathcal{O}^\times$ (recall that $\mathbf{G} = R_{L/\mathbb{Q}}(\mathrm{SL}_2)$). Hence, $\lambda = 1$ which contradicts our assumption.

Recall that $v = v_1 \wedge v_2$ where $v_1 = \Big(E_{12}, E_{12}\Big)$ and $v_2 = \Big(\sqrt{\beta}E_{12}, -\sqrt{\beta}E_{12}\Big)$. Since $\gamma\xi.v \notin \mathbb{R}.\gamma'\xi' v$ the subspace generated by the four vectors $w_i = g\gamma\xi.v_i$ $w_i' = g\gamma'\xi'.v_i$, for $i = 1, 2$ has dimension ≥ 3. We claim this subspace also generates a nilpotent subalgebra of $\mathfrak{g}$. This contradicts the fact that the dimension of any maximal nilpotent subalgebra in $\mathfrak{g}$ is 2 and finishes the proof of part (1).

To see the claim, note that (A.5) and the identity in (A.3) imply

$$\|p_1(w_1)\| \cdot \|p_2(w_2)\| \leq \delta/2,$$

similarly for w_1' and w_2'. In view of the definition of v_i (and w_i), therefore, $\|p_1(w_i)\| \cdot \|p_2(w_i)\| \ll_\beta \delta$ for $i = 1, 2$. Similarly, we have w_1' and w_2'.

We now apply (A.2) to the four vectors w_1, w_2, w_1', w_2'. In consequence, there are $\mathsf{s}_i, \mathsf{s}_i' \in \mathcal{O}^\times$ so that $\|\mathsf{s}_i w_i\| \ll_\beta \delta^{1/2}$ and $\|\mathsf{s}_i' w_i'\| \ll_\beta \delta^{1/2}$ for $i = 1, 2$.

Moreover, $\{\mathsf{s}_1 w_1, \mathsf{s}_2 w_2, \mathsf{s}_1' w_1', \mathsf{s}_2' w_2'\}$ are nilpotent elements in $\frac{1}{M}\,\mathrm{Ad}(g)\mathfrak{g}_\mathbb{Z}$. Since $\Big\|[w, w']\Big\| \leq \|w\|\|w'\|$, we get from the discreteness of $\mathrm{Ad}(g)\mathfrak{g}_\mathbb{Z}$ that if δ is small enough, then $\{\mathsf{s}_1 w_1, \mathsf{s}_2 w_2, \mathsf{s}_1' w_1', \mathsf{s}_2' w_2'\}$ generates a nilpotent Lie algebra as we claimed.

The argument for part (2) is similar to the proof of Lemma 3.2 as we now explain. For every $g \in G$ and every $\delta > 0$, put

$$I(g, \delta) = \Big\{r \in I : \|p_i^+(u_r g.v_i)\| \leq 0.01\delta\eta^2\|p_i(g.v_i)\| \text{ for } i = 1 \text{ or } i = 2\Big\},$$

where p_1^+ denotes the projection from $\mathfrak{g}$ onto $\mathbb{R}(E_{12}, 0)$ and p_2^+ denotes the projection from $\mathfrak{g}$ onto $\mathbb{R}(0, E_{12})$; recall also that p_i denotes projection onto the i-th component. As it was observed in Lemma 3.2, we have

$$|I(g, \delta)| \leq 2C'\delta^{1/2}|I|.$$

Let $\delta = 100\varrho^2$, and let $r \in I \setminus I(g, \delta)$. Then

$$\|p_i^+(u_r g.v_i)\| \geq \eta^2\|p_i(g.v_i)\|\varrho^2 \quad \text{for } i = 1, 2. \tag{A.6}$$

Using (A.3), we have $\|g.v\| = 2\|p_1(g.v_1)\| \cdot \|p_2(g.v_2)\|$. Since $a_t.w = e^t w$ for any $w \in \operatorname{span}\Big\{(E_{12}, 0), (0, E_{12})\Big\}$, using (A.3) and (A.6), we conclude that

$$\begin{aligned} e^{2t}\eta^4\|g.v\|\varrho^4 &= 2e^{2t}\eta^4\|p_1(g.v_1)\| \cdot \|p_2(g.v_2)\|\varrho^4 \\ &\le 2e^{2t}\|p_1^+(u_r g.v_1)\| \cdot \|p_2^+(u_r g.v_2)\| \\ &= \Big\|a_t\Big((p_1^+(u_r g.v_1), 0) \wedge (0, p_2^+(u_r g.v_2))\Big)\Big\| \le \|a_t u_r g.v\|. \end{aligned}$$

The claim thus holds with $C_{18} = 20C'$. □

Lemma A.2 *Let the notation be as above. There exists C_{19} so that*

$$C_{19}^{-1}\omega(x)^{-1} \le \operatorname{inj}(x)^2 \le C_{19}\omega(x)^{-1}$$

for all $x \in X$.

Proof Let $g \in G$ and assume that $\operatorname{inj}(g\Gamma) < \delta$. Then

$$g\Gamma g^{-1} \cap \mathsf{B}^G_{C\delta} \ne \{e\}$$

where C is an absolute constant.

If δ is small enough, then $g\Gamma g^{-1} \cap \mathsf{B}^G_{C\delta}$ consists only of unipotent elements. Therefore, there exists some nilpotent element $w \in \mathfrak{g}_{\mathbb{Z}}$ so that

$$\|gw\| \ll \delta,$$

where the implied constant is absolute.

Since all minimal $\mathbb{Q}$-parabolic subgroups of $\mathbf{G}$ are conjugate to each other by elements in $\mathbf{G}(\mathbb{Q})$, it follows from (A.1) that there exists some $\gamma \in \Gamma$ and some $\xi \in \Xi$ so that $w \in \gamma^{-1}\xi^{-1}.\mathrm{Lie}(N)$. Therefore, we may write

$$w = \gamma^{-1}\xi^{-1}.\Big((b + c\sqrt{\beta})E_{12}, (b - c\sqrt{\beta})E_{12}\Big)$$

where $b, c \in \frac{1}{M}\mathbb{Z}$ for some M depending on Ξ.

Using the Iwasawa decomposition, we write $g\gamma^{-1}\xi^{-1} = kan$ where $k \in \mathrm{SO}(2) \times \mathrm{SO}(2)$, $n \in N$, and $a = (a_{t_1}, a_{t_2})$ is diagonal. Therefore,

$$e^{t_1+t_2}(b^2 + c^2\beta) \ll \delta^2,$$

where the implied constant is absolute.

Now since $b, c \in \frac{1}{M}\mathbb{Z}$ are non-zero, we have $b^2 + c^2\beta \gg_M 1$. Altogether, we conclude that

$$\begin{aligned} \|g\gamma^{-1}\xi^{-1}.v\| &= 2\|p_1(a_{t_1}.v_1)\|\|p_2(a_{t_2}.v_2)\| \\ &\le 2\sqrt{\beta}e^{t_1+t_2} \le 2\sqrt{\beta}\hat{C}\delta^2 \end{aligned}$$

where $\hat{C}$ depends on Γ. Since $\omega(g\Gamma)^{-1} \leq \|g\gamma^{-1}\xi^{-1}.v\|$, the lower bound in the lemma follows.

We now turn to the proof of the upper bound. Using the reduction theory for arithmetic groups, see e.g. [50, Ch. 4], there exist $t_0, r_0 > 0$ so that

$$\Big(\mathrm{SO}(2) \times \mathrm{SO}(2)\Big) \cdot \Big\{(a_t, a_{t'}\Big) : t + t' \leq t_0\Big\} \cdot \{n(r,s); |r|, |s| \leq r_0\} \cdot \Xi$$

is a (generalized) fundamental domain for Γ in G.

In particular, using Lemma A.1(1), there exists $t_1 \leq t_0$ so that if $g = k(a_t, a_{t'})n(r,s)\xi_0\gamma_0$ for $t + t' \leq t_1$, then

$$\begin{aligned}\omega(g\Gamma) &= \max\Big\{\|g\gamma\xi^{-1}.v\|^{-1} : (\xi,\gamma) \in \Xi \times \Gamma\Big\} = \|g\gamma_0^{-1}\xi_0^{-1}.v\|^{-1}\\ &= \|k(a_t, a_{t'})n(r,s).v\|^{-1} = e^{-t-t'}\|v\|^{-1}.\end{aligned}$$

Moreover, using (A.3) and (A.2) we conclude that $g\gamma_0^{-1}\xi_0^{-1}(N\cap\Gamma)\xi_0\gamma_0 g^{-1}$ contains elements of size $e^{(-t-t')/2}$. The upper bound estimate follows. $\square$

Proof of Proposition 3.1: Case 2 By Lemma A.2, $t \geq |\log(\eta^2\,\mathrm{inj}(g\Gamma))| + C_7$ implies $2t \geq \log(\omega(g\Gamma)/\eta^4)$ if we assume C_7 is large enough.

Let $\varrho_0 = 0.1C_{18}^{-1}$. In view of Lemma A.1(2) we have

$$\sup\Big\{\|a_t u_r g\gamma\xi^{-1}.v\| : r \in I\Big\} \geq \varrho_0^4 \qquad \text{for all } \gamma \in \Gamma \text{ and } \xi \in \Xi$$

so long as $2t \geq |\log(\omega(g\Gamma)/\eta^4)|$.

Altogether, the conditions in [38, Thm. 4.1] are satisfied so long as $t \geq |\log(\eta^2\,\mathrm{inj}(g\Gamma))| + C_7$. Hence, similar to the previous case, the conclusion of the proposition in this case also holds true in view of [38, Thm. 4.1] — in light of Lemma A.1(1), the proof simplifies significantly. $\square$

We also record the following which is a special case of the results and techniques developed in [22] and [20] tailored to our setup here.

Proposition A.3 *Let $0 < \alpha < 1$ and let m_α be as in (2.12). There exists some $B = B(X, \alpha) \geq 1$ satisfying the following. For every $x \in X$ and every $n \in \mathbb{N}$ we have*

$$\int_H \mathrm{inj}(hx)^{-\alpha}\,\mathrm{d}\nu^{(n)}(h) \leq e^{-n}\,\mathrm{inj}^{-\alpha}(x) + B,$$

where $\nu(\varphi) = \int_0^1 \varphi(a_{m_\alpha}u_r)\,\mathrm{d}r$ for every $\varphi \in C_c(H)$ and $\nu^{(n)}$ denotes the n-fold convolution of ν.

Proof If X is compact, then $\mathrm{inj} : X \to \mathbb{R}$ is a bounded function and the result is clear.

Therefore, we may assume X is not compact. If $G = \mathrm{SL}_2(\mathbb{C})$, the claim in the proposition is proved in [48].

We now consider $G = \mathrm{SL}_2(\mathbb{R}) \times \mathrm{SL}_2(\mathbb{R})$ and consider separately the cases where Γ is a reducible lattice and Γ is irreducible.

Case 1. Let use first assume that Γ is reducible. As was done before, passing to a finite index subgroup, we may assume $\Gamma = \Gamma_1 \times \Gamma_2$.

Let ω be defined as in (3.3). That is:

$$\omega(x) = \max\{\omega_1(x_1), \omega_2(x_2)\},$$

where $x = (x_1, x_2)$.

By [48, Prop. 6.7] we have $\omega(x) \asymp \mathrm{inj}(x)^{-1}$. Therefore, it suffices to prove the proposition with $\mathrm{inj}(x)$ replaced by $\omega(x)$. The result for ω_1 and ω_2 is well-known, see e.g. [20,22,48].

The result for ω thus follows as $\omega^\alpha \le \omega_1^\alpha + \omega_2^\alpha \le 2\omega^\alpha$.

Case 2. Assume now that Γ is irreducible. We will use the notation which we fixed in the beginning of this appendix. In particular, as was done in (A.4), define

$$\omega(g\Gamma) = \max\left\{2, \max\left\{\|g\gamma\xi.v\|^{-1} : \xi \in \Xi^{-1}, \gamma \in \Gamma\right\}\right\}.$$

In view of Lemma A.2, we have $\omega(x) \asymp \mathrm{inj}(x)^{-2}$ for all $x \in X$. Therefore, it suffices to prove the claim for $\omega^{1/2}$ instead if inj.

Let us recall from (A.3) that

$$\begin{aligned} gv &= -2(p_1(g.v), 0) \wedge (0, p_2(g.v)) \\ &= -2\sqrt{\beta}(g_1 E_{12}, 0) \wedge (0, g_2 E_{12}) \end{aligned} \tag{A.7}$$

for any $g = (g_1, g_2)$.

Let $x = g\Gamma$. Fix $\gamma \in \Gamma$ and $\xi \in \Xi^{-1}$; for all $r \in [0, 1]$ and $\ell \in \mathbb{N}$ put $h_r = a_\ell u_r \gamma\xi$. In view of the Cauchy-Schwarz inequality and (A.7), applied with $h_r g$, we have

$$\begin{aligned} &\left(\int_0^1 \|h_r g v\|^{-\alpha/2}\,\mathrm{d}r\right)^2 \\ &\quad \le 2\sqrt{\beta}\int_0^1 \|h_{r1} g_1 E_{12}\|^{-\alpha}\,\mathrm{d}r \int_0^1 \|h_{r2} g_2 E_{12}\|^{-\alpha}\,\mathrm{d}r. \end{aligned} \tag{A.8}$$

Then for $i = 1, 2$, by choice of m_α, we have

$$\int_0^1 \|a_{m_\alpha} u_r g_i \gamma_i \xi_i E_{12}\|^{-\alpha} \,\mathrm{d}r < e^{-1} \|g_i \gamma_i \xi_i E_{12}\|^{-\alpha},$$

see (2.12).

Using (A.7) in reverse order and (A.8), we conclude from the above two estimates that

$$\int_0^1 \|a_{m_\alpha} u_r g \gamma \xi v\|^{-\alpha/2} \,\mathrm{d}r \leq e^{-1} \|g \gamma \xi v\|^{-\alpha/2}. \tag{A.9}$$

Let $C(\Gamma)$ be as in Lemma A.1. Then there exists some $B'_{m_\alpha} > 0$ so that if $\omega(g\Gamma) = \|g\gamma\xi v\|^{-1} \geq C(\Gamma) \cdot B'_{m_\alpha}$, then

$$\omega(a_{m_\alpha} u_r g\Gamma) = \|a_{m_\alpha} u_r g \gamma \xi v\|^{-1} \geq C(\Gamma)$$

for all $r \in [0, 1]$.

This and (A.9) imply that for all $x \in X$, we have

$$\int \omega^{\alpha/2}(hx) \,\mathrm{d}\nu(h) = \int_0^1 \omega^{\alpha/2}(a_{m_\alpha} u_r x) \,\mathrm{d}r \leq e^{-1} \omega^{\alpha/2}(x) + B'',$$

where $B'' = \max\{\omega(a_{m_\alpha} u_r g\Gamma) : r \in [0, 1], \omega(g\Gamma) \leq C(\Gamma) \cdot B'_{m_\alpha}\}$.

Iterating this estimate and summing the geometric sum, we conclude that

$$\int \omega^{\alpha/2}(hx) \,\mathrm{d}\nu^{(n)}(h) \leq e^{-n} \omega^{\alpha/2}(x) + B \tag{A.10}$$

for all $n \in \mathbb{N}$ where $B = 2B''$. The proof is complete. □

Appendix B: Proof of Theorem 5.1

Recall that $\mathfrak{r} \subset \mathrm{Lie}(G)$ is identified with $\mathfrak{sl}_2(\mathbb{R})$ equipped with the adjoint action of $\mathrm{SL}_2(\mathbb{R})$.

Theorem B.1 *Let $0 < \alpha \leq 1$, and let $0 < b_0 < b_1 < 1$. Let $E \subset B_{\mathfrak{r}}(0, b_1)$ be a finite set, and let ρ denote the uniform measure on E. Assume that*

$$\rho(B_{\mathfrak{r}}(w, b)) \leq \Upsilon \cdot (b/b_1)^\alpha \qquad \textit{for all } w \textit{ and all } b \geq b_0, \tag{B.1}$$

where $\Upsilon \geq 1$.

Let $0 < \varepsilon < 0.01\alpha$, *and let* $J \subset [0, 1]$ *be an interval with* $|J| \geq 10^{-6}$. *For every* $b \geq b_0$, *there exists a subset* $J_b \subset J$ *with* $|J \setminus J_b| \leq C_\varepsilon (b/b_1)^\varepsilon$ *so that the following holds. Let* $r \in J_b$, *then there exists a subset* $E_{b,r} \subset E$ *with*

$$\rho(E \setminus E_{b,r}) \leq C_\varepsilon (b/b_1)^\varepsilon$$

such that for all $w \in E_{b,r}$, *we have*

$$\rho\Big(\{w' \in E : |\xi_r(w') - \xi_r(w)| \leq b\}\Big) \leq C_\varepsilon (b/b_1)^{\alpha - 7\varepsilon}$$

where $C_\varepsilon \ll \varepsilon^{-\star}\Upsilon^\star$ *(implied constants are absolute) and*

$$\xi_r(w) = (\mathrm{Ad}(u_r)w)_{12} = -w_{21}r^2 - 2w_{11}r + w_{12}.$$

We need some more notation for the proof. First note that the assumption and the conclusion in the theorem are invariant under scaling. Thus replacing E by $b_1^{-1} \cdot E$ and b_0 by b_0/b_1, we may assume $b_1 = 1$. We prove the theorem for $J = [0, 1]$, the proof in general is similar.

Let

$$\Xi(w) = \Big\{(r, \xi_r(w)) : r \in [0, 1]\Big\}$$

for every $w \in E$, and let $\Xi = \bigcup_w \Xi(w)$.

For every $b > 0$ and every $w \in E$, let

$$\Xi^b(w) = \Big\{(q_1, q_2) \in [0, 1] \times \mathbb{R} : |q_2 - \xi_{q_1}(w)| \leq b\Big\}.$$

Finally, for all $q \in \mathbb{R}^2$ and $b > 0$, define

$$m_\rho^b(q) := \rho\Big(\{w' \in \mathfrak{r} : q \in \Xi^b(w')\}\Big). \tag{B.2}$$

The assertion in the theorem may be rewritten in terms of the multiplicity function m_ρ^b as follows. We seek the set $J_b \subset [0, 1]$, and for every $r \in J_b$, the set $E_{b,r} \subset E$ so that

$$m_\rho^b\Big((r, \xi_r(w))\Big) \leq C_\varepsilon b^{\alpha - 7\varepsilon} \quad \text{for all } w \in E_{b,r}. \tag{B.3}$$

The following lemma plays a crucial role in the proof of Theorem B.1. This is a more detailed version of [56, Lemma 8] in the setting at hand, see also [65, Lemma 1.4] and [66, Lemma 2.1]. Indeed, Lemma B.2 is a restatement of [33, Lemma 5.1] for a family of parabolas; similar to loc. cit., the regularity of the measure ρ, (B.1), is used as a replacement for the assumption in [56, Lemma 8] that the family has separated radii.

Lemma B.2 *Let the notation be as in Theorem B.1 with $b_1 = 1$. In particular, $E \subset B_{\mathfrak{r}}(0,1)$ and (B.1) is satisfied. For every $0 < \varepsilon \le 0.01\alpha$, there exists $0 < D \ll \varepsilon^{-\star}\Upsilon^{\star}$ (implied constants are absolute) so that the following holds. Let $b \ge b_0$. Then there exists a subset $\hat{E} = \hat{E}_b \subset E$ with $\#(E \setminus \hat{E}) \le b^{\varepsilon} \cdot (\#E)$ so that for every $w \in \hat{E}$, we have*

$$\left|\Xi^b(w) \cap \left\{q \in \mathbb{R}^2 : m^b_\rho(q) \ge Db^{\alpha - 7\varepsilon}\right\}\right| \le b^{2\varepsilon/\alpha}|\Xi^b(w)|.$$

The proof of this lemma is mutatis mutandis of the argument in [33, Lemma 5.1] where one replaces the use of [65, Lemma 1.4] with [66, Lemma 5.18]. We explicate the notation and the main steps for the convenience of the reader.

Define $\Phi : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ by

$$\Phi(x, y) = y_2 + 2x_1 y_1 + x_2 y_1^2.$$

Given $x_0 \in \mathbb{R}^2$ and $r_0 \in \mathbb{R}$, the set $\{y \in \mathbb{R}^2 : \Phi(x_0, y) = r_0\}$ is a special example of a Φ-circle in [40,66].

Note that $\Xi(w) = \left\{y \in \mathbb{R}^2 : y_1 \in [0,1], \Phi\Big((w_{11}, w_{21}), y\Big) = w_{12}\right\}$. The family Ξ satisfies the *cinematic* curvature conditions [66, Eq. (1.5) and (1.6)]. Indeed in the case at hand, these conditions follow from the following estimate

$$\tfrac{1}{3}\max\{|x_1|, |x_2|\} \le |\tfrac{\partial \Phi}{\partial y_1}| + |\tfrac{\partial^2 \Phi}{\partial y_1^2}| \le 3\max\{|x_1|, |x_2|\}; \tag{B.4}$$

we remark that when $\Phi(0, y) = y_2$, as is the case here, (B.4) (with 3 replaced by a constant C) may be taken as the definition of the cinematic curvature conditions, see [40, Eq. (21)].

Let $w, w' \in B_{\mathfrak{r}}(0,1)$; define

$$\Delta(w - w') = \Big|\det(w - w')\Big|.$$

The function Δ may be used to quantitatively measure the tangency of $\Xi(w)$ and $\Xi(w')$. Our choice of Δ is different from $\Delta_{B_{\mathfrak{r}}(0,2)}$ which is defined in [66, Def. 2.2], however, in the case at hand $\Delta \asymp \Delta_{B_{\mathfrak{r}}(0,2)}$ — indeed, the (reduced) discriminant of $\xi_r(w) - \xi_r(w')$ equals $-\det(w - w')$.

By [40, Lemma 3.1], for all $0 < \delta < 0.1$ and all $w, w' \in B_{\mathfrak{r}}(0,1)$, we have

$$\operatorname{diam}\Big(\Xi^{\delta}(w) \cap \Xi^{\delta}(w')\Big) \ll \frac{\sqrt{\Delta(w-w')+\delta}}{\sqrt{\|w-w'\|+\delta}}, \tag{B.5a}$$

$$|\Xi^{\delta}(w) \cap \Xi^{\delta}(w')| \ll \frac{\delta^2}{\sqrt{(\|w-w'\|+\delta)(\Delta(w-w')+\delta)}}, \tag{B.5b}$$

here and in the remaining parts of the argument, the implied constants are absolute unless otherwise is stated explicitly.

Let $\mathcal{W}, \mathcal{B} \subset B_{\mathfrak{r}}(0,1)$. We say $(\mathcal{W}, \mathcal{B})$ is t-bipartite if

$$\max\{\operatorname{diam}(\mathcal{W}), \operatorname{diam}(\mathcal{B})\} \le t \le d(\mathcal{W}, \mathcal{B}). \tag{B.6}$$

Let $0 < \delta \le t \le 1$. A (δ, t)-rectangle $R \subset \mathbb{R}^2$ is a δ-neighborhood of a piece of a parabola $\Xi(w)$, $w \in B_{\mathfrak{r}}(0,1)$, with length $\sqrt{\delta/t}$. We say that two (δ, t)-rectangles are C-comparable if there is a $(C\delta, t)$-rectangle which contains both of them. Otherwise, they are C-incomparable. Let $w \in B_{\mathfrak{r}}(0,1)$, the parabola $\Xi(w)$ is C-tangent to a (δ, t)-rectangle R, if $\Xi^{C\delta}(w)$ contains R. Finally, fixing some large absolute constant $\hat{C} \ge 1$, we say that two rectangles are comparable, if they are $\hat{C}$-comparable. Similarly, $\Xi(w)$ is said to be tangent to a rectangle R if $\Xi(w)$ is $\hat{C}$-tangent to R.

Let $0 < \delta \le t \le 1$, and let $(\mathcal{W}, \mathcal{B})$ be t-bipartite. Let R be a (δ, t)-rectangle. Put $\mathcal{W}_R = \{w \in \mathcal{W} : \Xi(w) \text{ is tangent to } R\}$; define $\mathcal{B}_R$ analogously. We say R is of type $(\ge \mu, \ge \nu)$ with respect to ρ, $\mathcal{W}$, and $\mathcal{B}$ if

$$\rho(\mathcal{W}_R) \ge \mu \quad \text{and} \quad \rho(\mathcal{B}_R) \ge \nu.$$

We say R is of type (μ, ν) if $\mu \le \rho(\mathcal{W}_R) < 2\mu$ and $\nu \le \rho(\mathcal{B}_R) < 2\nu$.

The following is an analogue of [65, Lemma 1.4] tailored to our setting here; see also [66, Lemma 5.18] and [33, Lemma 4.4].

Lemma B.3 *Let $0 < \delta \le t \le 1$, and let $(\mathcal{W}, \mathcal{B})$ be t-bipartite. Let $\varepsilon > 0$. Then the number of pairwise incomparable (δ, t)-rectangles of type $(\ge \mu, \ge \nu)$ with respect to ρ, $\mathcal{W}$, and $\mathcal{B}$ is at most*

$$D_\varepsilon(\mu\nu\delta)^{-\varepsilon}\left(\left(\frac{\rho(\mathcal{W})\rho(\mathcal{B})}{\mu\nu}\right)^{3/4} + \frac{\rho(\mathcal{W})}{\mu} + \frac{\rho(\mathcal{B})}{\nu}\right)$$

where $D_\varepsilon \ll \varepsilon^{-\star}$ and the implied constants are absolute.

Proof Replacing the use of [65, Lemma 1.4] with [66, Lemma 5.18], the same proof as in [33, Lemma 4.4] applies here. The argument is standard:

given $(\mathcal{W}, \mathcal{B})$ and a collection $\mathcal{R}$ of incomparable (δ, t)-rectangles, one uses a dyadic decomposition argument to find $i, j \in \mathbb{N}$ with

$$2^i/i^2 \leq \delta^{-3}\mu^{-1} \quad \text{and} \quad 2^j/j^2 \leq \delta^{-3}\nu^{-1},$$

a subset $\mathcal{R}' \subset \mathcal{R}$ with $\#\mathcal{R}' \gg \varepsilon^{-\star}(\#\mathcal{R})\delta^{6\varepsilon}\mu^\varepsilon\nu^\varepsilon$, and a t-bipartite $(\mathcal{W}', \mathcal{B}')$ where $\mathcal{W}', \mathcal{B}' \subset B_{\mathfrak{r}}(0, 1)$ are δ-separated with $\#\mathcal{W}' \ll 2^i\rho(\mathcal{W})$ and $\#\mathcal{B}' \ll 2^j\rho(\mathcal{B})$, so that every $R \in \mathcal{R}'$ is of type

$$(\geq D'_\varepsilon 2^i\mu^{1+\varepsilon}\delta^{3\varepsilon}, \geq D'_\varepsilon 2^j\nu^{1+\varepsilon}\delta^{3\varepsilon})$$

with respect to the counting measure, $\mathcal{W}'$, and $\mathcal{B}'$ for some $D'_\varepsilon \ll \varepsilon^{-\star}$. One then applies [66, Lemma 5.18] to $(\mathcal{W}', \mathcal{B}')$ and $\mathcal{R}'$ and obtains a bound for $\#\mathcal{R}'$ which implies the desired bound for $\#\mathcal{R}$. We note that the definition of a t-bipartite family in [66] requires the radii are δ-separated, [66, Def. 2.3]; this assumption however is not used in the proof of [66, Lemma 5.18]. Indeed as in [65, Lemma 1.4], one only needs δ-separation is the parameter space, i.e. $\|w - w'\| \geq \delta$ in the case at hand.

The final estimate $D_\varepsilon \ll \varepsilon^{-\star}$ follows from $D'_\varepsilon \ll \varepsilon^{-\star}$ and the fact that the implied constant in [66, Lemma 5.18] is $\ll \varepsilon^{-\star}$. This follows from the proof of [66, Lemma 5.18], see in particular [65, pp. 1252–1253]. □

Proof of Lemma B.2 Throughout the argument, D will be assumed to be a large constant which is allowed to depend (polynomially) on $1/\varepsilon$ and Υ.

Let $b \geq b_0$ be the largest dyadic number where the lemma fails; taking D large enough, we assume that b is small compared to absolute constants whenever necessary. Let $A = (Db^{-3\varepsilon})^{1/\alpha}$ and $\lambda = b^{2\varepsilon/\alpha}$. By the choice of b, there exists $\mu \geq Db^{\alpha-7\varepsilon} = A^\alpha\lambda^{-2\alpha}b^\alpha$ and a subset $E' \subset E$ with $\#E' > b^\varepsilon \cdot (\#E) = D^{1/3}A^{-\alpha/3} \cdot (\#E)$ so that for all $w \in E'$, we have

$$\left|\Xi^b(w) \cap \left\{q \in \mathbb{R}^2 : m_\rho^b(q) \geq \mu\right\}\right| \geq \lambda|\Xi^b(w)|.$$

For every $w \in \mathfrak{r}$ and dyadic numbers $t, \delta \in (b, 1]$, define

$$E_{\delta,t}(w) = \left\{w' \in E : \Xi^b(w) \cap \Xi^b(w') \neq \emptyset, \begin{array}{l} t \leq \|w - w'\| < 2t \\ \delta \leq \Delta(w - w') < 2\delta \end{array}\right\}.$$

Define $E_{b,t}(w)$ similarly, except in this case no lower bound is assumed for Δ, that is, we only assume $\Delta(w - w') < 2b$.

For every $F \subset E$, define $m_\rho^\bullet(q|F) = \rho\Big(\{w' \in F : q \in \Xi^\bullet(w')\}\Big)$. Replacing the use of [33, Lemma 3.6] with (B.5a) and (B.5b), one may argue as in the proof of [33, Eq. (5.4)] and conclude the following. There exist absolute constants $C, C_1 \geq 1$, $\bar{E} \subset E'$ with $\#\bar{E} \geq C^{-1}|\log b|^{-C} \cdot (\#E')$, and some dyadic number $n \in \{1, \ldots, \delta/b\}$, so that if we put

$$\lambda_\delta = |\log b|^{-C} \cdot \frac{\lambda\delta}{Cnb}, \qquad A_\delta = C|\log b|^{C} \cdot \frac{A\delta}{nb}, \tag{B.7}$$

and $\mu_\delta = |\log b|^{-C} \cdot \frac{n\mu}{C}$, then for all $w \in \bar{E}$ we have

$$|\Xi^\delta(w) \cap \{q \in \mathbb{R}^2 : m_\rho^{C_1\delta}(q|E_{\delta,t}(w)) \geq \mu_\delta\}| \geq 2\lambda_\delta |\Xi^\delta(w)|, \tag{B.8}$$

see [33, Eq. (5.12)]. Note also that $\mu_\delta \gg |\log b|^{-\star} A_\delta^\alpha \lambda_\delta^{-2\alpha} \delta^\alpha$.

Fix a large dyadic number $N \geq 2$, in particular, $N\delta \geq 2b$. Now (B.8) and the inductive hypothesis (recall the choice of b), imply that there exists a subset $\bar{E}' \subset \bar{E}$ with $\#\bar{E}' \gg \#\bar{E}$ so that for all $w \in \bar{E}'$, we have

$$\begin{aligned} &\left|\Xi^\delta(w) \cap \left\{q \in \mathbb{R}^2 : \mu_\delta \leq m_\rho^{C_1\delta}(q|E_{\delta,t}(w)) \leq m_\rho^{N\delta}(q) \leq M_\delta\right\}\right| \\ &\quad \geq \lambda_\delta \left|\Xi^\delta(w)\right|, \end{aligned} \tag{B.9}$$

where $M_\delta = A_\delta^\alpha (\lambda_\delta / CN)^{-2\alpha} \delta^\alpha \ll |\log b|^\star \mu_\delta$, see [33, Eq. (5.14)].

Let $\{B_{\mathfrak{r}}(w_i, 0.1t)\}$ be a covering of $\bar{E}'$ chosen so that $\{B_{\mathfrak{r}}(w_i, 2.1t)\}$ has bounded multiplicity. Replacing $\bar{E}'$ with a subset whose ρ measure is $\geq 0.5\rho(\bar{E}')$, we assume that $\rho(B_{\mathfrak{r}}(w_i, 0.1t) \cap \bar{E}') \gg t^3 \rho(\bar{E}')$ for all $w_i \in \bar{E}'$.

Let i_0 be so that $\rho(B_{\mathfrak{r}}(w_{i_0}, 0.1t) \cap \bar{E}')/\rho(B_{\mathfrak{r}}(w_{i_0}, 2.1t))$ is maximized. Put $\mathcal{W}' := B_{\mathfrak{r}}(w_{i_0}, 0.1t) \cap \bar{E}'$ and $\mathcal{B} := B_{\mathfrak{r}}(w_{i_0}, 2.1t) \setminus B_{\mathfrak{r}}(w_{i_0}, 0.9t)$.

Replacing $\mathcal{W}'$ by a subset $\mathcal{W} \subset \mathcal{W}'$ with $\rho(\mathcal{W}) \geq 0.5\rho(\mathcal{W}')$, we may assume that for all $z \in \mathcal{W}$, there is a dyadic cube $Q(z)$ of side-length δ which contains z and $\rho(Q(z) \cap \mathcal{W}) \gg (\delta/t)^3 \rho(\mathcal{W}) \gg |\log b|^{-\star} A^{-\alpha/3} \delta^3$. Note also that since the covering $\{B_{\mathfrak{r}}(w_{i_0}, 2.1t)\}$ has bounded multiplicity, we have

$$\rho(\mathcal{W}) \geq 0.5\rho(\mathcal{W}') \gg |\log b|^{-\star} A^{-\alpha/3} \rho(\mathcal{B}).$$

By the definition, $(\mathcal{W}, \mathcal{B})$ is t-bipartite, see (B.6). Moreover, for all $w \in \mathcal{W}$, we have $E_{\delta,t}(w) \subset \mathcal{B}$. Hence,

$$m_\rho^{C_1\delta}(q|E_{\delta,t}(w) \cap \mathcal{B}) = m_\rho^{C_1\delta}(q|E_{\delta,t}(w)), \tag{B.10}$$

for all $w \in \mathcal{W}$ and $q \in \mathbb{R}^2$. We conclude from (B.10), (B.9), and (B.1) that

$$|\log b|^{-\star} A_\delta^\alpha \lambda_\delta^{-2\alpha} \delta^\alpha \ll \mu_\delta \leq m_\rho^{C_1\delta}(q|E_{\delta,t}(w) \cap \mathcal{B}) \leq \rho(\mathcal{B}) \ll t^\alpha;$$

therefore, δ is much smaller than t if D is large enough, see (B.7) and recall that $A = (Db^{-3\varepsilon})^{1/\alpha}$ and $0 < \lambda_\delta \leq 1$.

Since $\mathcal{W} \subset \bar{E}'$, (B.9) and (B.10) imply that for all $w \in \mathcal{W}$, we have

$$\left|\Xi^\delta(w) \cap \left\{q \in \mathbb{R}^2 : \mu_\delta \le m_\rho^{C_1\delta}(q|E_{\delta,t}(w) \cap \mathcal{B}) \le m_\rho^{N\delta}(q) \le M_\delta\right\}\right| \ge \lambda_\delta \left|\Xi^\delta(w)\right|. \tag{B.11}$$

Assuming N is large enough, depending on C_1, (B.11) implies that every $w \in \mathcal{W}$ supplies $\gg \lambda_\delta\sqrt{t/\delta}$ incomparable (δ, t)-rectangles each of which is $N/2$-tangent to $\Xi(w)$ and has type $\ge \mu_\delta$ with respect to $\mathcal{B}$ where the type refers to N-tangency. From this, we conclude that there are

$$\gg |\log b|^{-\star}\rho(\mathcal{W})\lambda_\delta\sqrt{t/\delta}/\nu_\delta$$

incomparable (δ, t)-rectangles of type $(\ge \nu_\delta, \ge \mu_\delta)$ with respect to ρ, $\mathcal{W}$, and $\mathcal{B}$ where $b^4 \le \nu_\delta \le M_\delta$ is a dyadic number and type refers to N-tangency. Comparing this bound with the bound given by Lemma B.3 yields a contradiction and finishes the proof, see [33, pp. 20–21].

The assertion $D \ll \varepsilon^{-\star}\Upsilon^{-\star}$ follows from the above outline, together with the fact D_ε in Lemma B.3 is $\ll \varepsilon^{-\star}$. □

We now turn to the proof of Theorem B.1. The argument is a slight modification of the proof of [33, Thm. 7.2].

Proof of Theorem B.1 Assume that the conclusion of the theorem fails for some C. That is, there exists a subset $\bar{J} \subset [0, 1]$ with $|\bar{J}| > Cb^\varepsilon$ so that for all $r \in \bar{J}$ we have

$$\rho(E'_r) \ge Cb^\varepsilon, \tag{B.12}$$

where $E'_r = \left\{w \in E : m_\rho^b\Big((r, \xi_r(w))\Big) > Cb^{\alpha-7\varepsilon}\right\}$. We will get a contradiction if C is large enough.

Let $\hat{E}$ be as in Lemma B.2 applied with $8b$, then $\rho(\hat{E}) \ge 1 - (8b)^\varepsilon$. This and (B.12) now imply that for every $r \in \bar{J}$, we have $\rho(\hat{E} \cap E'_r) \ge Cb^\varepsilon/2$ so long as $C \ge 16$.

We conclude that

$$\begin{aligned} 0.5C^2b^{2\varepsilon} &\le \int_{\bar{J}} \rho(\hat{E} \cap E'_r)\,\mathrm{d}r \\ &\le \int_{\hat{E}} |\{r : m_\rho^b(r, \xi_r(w)) > Cb^{\alpha-7\varepsilon}\}|\,\mathrm{d}\rho. \end{aligned}$$

Therefore, there exists some $w_0 \in \hat{E}$ so that

$$\left|\left\{r \in [0,1] : m_\rho^b\Big((r, \xi_r(w_0))\Big) > Cb^{\alpha-7\varepsilon}\right\}\right| \geq 0.5C^2 b^{2\varepsilon}. \tag{B.13}$$

For every $r \in [0,1]$, let $L_r := \{(r,s) : s \in \mathbb{R}\}$ be a vertical line, and let $I \subset L_r$ be an interval of length b containing $(r, \xi_r(w_0))$. Put

$$I_{+,b} = \Big\{(q_1,q_2) \in [r-b, r+b] \times \mathbb{R} : \exists (r,s) \in I, |q_2 - s| \leq b\Big\}.$$

If $(q_1,q_2) \in I_{+,b}$, then $|q_1 - r| \leq b$ and $|q_2 - \xi_r(w_0)| \leq 2b$. Therefore,

$$|q_2 - \xi_{q_1}(w_0)| \leq |q_2 - \xi_r(w_0)| + |\xi_r(w_0) - \xi_{q_1}(w_0)| \leq 8b.$$

We conclude that $(q_1,q_2) \in \Xi^{8b}(w_0)$. This and $m_\rho^b\Big((r, \xi_r(w_0))\Big) > Cb^{\alpha-7\varepsilon}$ imply that for every $q \in I_{+,b}$, we have

$$m_\rho^{8b}(q) \geq \rho\Big(\{w' \in E : (r, \xi_r(w')) \in I\}\Big) \geq Cb^{\alpha-7\varepsilon}. \tag{B.14}$$

Combining (B.13) and (B.14), we obtain that

$$\begin{aligned}\Big|\Xi^{8b}(w_0) \cap \{q \in \mathbb{R}^2 : m_\rho^{8b}(q) \geq Cb^{\alpha-7\varepsilon}\}\Big| &\gg C^2 b^{1+2\varepsilon} \\ &\gg C^2 b^{2\varepsilon}|\Xi^{8b}(w_0)| > b^{2\varepsilon/\alpha}|\Xi^{8b}(w_0)|,\end{aligned}$$

where the implied constant is absolute, and we assume C is large enough so that the final estimate holds — recall that $0 < \alpha \leq 1$.

This contradicts the fact that $w_0 \in \hat{E}$ and finishes the proof. □

Proof of Theorem 5.1 Fix some κ. We may assume b's are dyadic numbers, in particular $b_i = 2^{-\ell_i}$, for $i = 0, 1$. Let ℓ_2 be so that

$$\sum_{\ell=\ell_2}^{\infty} C_\kappa 2^{-\kappa\ell} < 0.1 \min\{|J|, 1\}.$$

Let $J' = \bigcap_{\ell=\ell_2}^{\ell_0} J_{2^{-\ell}}$. Then the choice of ℓ_2 and Theorem B.1 imply that $|J'| \geq 0.9|J|$.

For every $r \in J'$, let $E_r = \bigcap_{\ell=\ell_2}^{\ell_0} E_{2^{-\ell},r}$. Then by Theorem B.1, $\rho(E_r) \geq 0.9$. Moreover, for all $w \in E_r$ and all $\ell_2 \leq \ell \leq \ell_0$ we have

$$\rho(\{w' \in E : |\xi_r(w') - \xi_r(w)| \leq 2^{-\ell}\}) \leq C_\kappa 2^{(\alpha-7\kappa)(\ell_1-\ell)}.$$

The above implies that Theorem 5.1 holds true with J' and E_r if we increase C_κ to account for all $b \geq 2^{-\ell_2}$. □

Appendix C: Proof of Lemma 5.1

We will prove Lemma 5.1 in this section. As was mentioned before, the proof is taken from [6, Lemma 5.2], see also [5]; we reproduce the argument to explicate the stated bounds on b_1.

Proof of Lemma 5.1 We identify $\mathfrak{r}$ with $\mathbb{R}^3$. By a dyadic cube we mean a cube

$$[\tfrac{n_1}{2^k}, \tfrac{n_1+1}{2^k}) \times [\tfrac{n_2}{2^k}, \tfrac{n_2+1}{2^k}) \times [\tfrac{n_3}{2^k}, \tfrac{n_3+1}{2^k})$$

for an integer $k \geq 0$ and $0 \leq n_i < 2^k$.

Let ρ denote the uniform measure on F. Let $b \geq (\#F)^{-(1+\varepsilon)/\alpha}$ and $w \in \mathbb{R}^3$, then

$$\begin{aligned} b^{-\alpha}\rho\Big(B(w,b)\Big) &\leq \frac{1}{\#F}\Big(b^{-\alpha} + \sum_{w' \in B(w,b), w' \neq w} \|w - w'\|^{-\alpha}\Big) \\ &\leq \frac{1}{\#F}\Big(b^{-\alpha} + D(\#F)^{(1+\varepsilon)}\Big) \\ &\leq (D+1)\cdot(\#F)^{\varepsilon}. \end{aligned} \tag{C.1}$$

We will absorb the constant D using the notation $\gg$ and $\ll$ in what follows. Let $b_0 = (\#F)^{-1}$. Using the Besicovitch covering lemma and the fact that ρ is probability measure, we conclude from (C.1) that F contains a subset $\hat{F}$ of b_0-separated points with

$$\#\hat{F} \gg b_0^{\varepsilon-\alpha},$$

where the implied constant is absolute.

Arguing as in the proof [6, Lemma 5.2], see also [5], with $\hat{F}$ and $\alpha - \varepsilon$, there exists some T, depending on ε, and a subset $F_1 \subset \hat{F}$, with

$$\#F_1 \geq \hat{C} b_0^{2\varepsilon-\alpha} \tag{C.2}$$

so that the following holds. Let $k_1 = \lceil -\log_2(b_0)/T \rceil$, then there exist integers $R_1, \ldots, R_{k_1}$ with $1 \leq R_\ell \leq 2^{3T}$ so that every $2^{-\ell T}$-cube which intersects F_1 contains exactly $R_{\ell+1}$, $2^{-(\ell+1)T}$-cubes which intersect F_1.

Since each remaining $2^{-k_1 T}$-cube contains exactly one point, we have

$$\sum_{\ell=1}^{k_1} \log_2 R_\ell = \log_2(\#F_1) \geq (\alpha - 2\varepsilon)T(k_1 - 2), \tag{C.3}$$

where we assume T is large enough to account for the constant $\hat{C}$.

For every $k > \lfloor k_1 \varepsilon \rfloor =: k_0$, let

$$M_k = \min_{k<\ell\leq k_1} \frac{1}{\ell - k} \sum_{k+1}^{\ell} \log_2 R_i.$$

Let k_2 be the smallest integer so that $M_{k_2} \geq (\alpha - 20\varepsilon)T$ if such exists, else let $k_2 = k_1$. We claim

$$\varepsilon k_1 \leq k_2 \leq \tfrac{3-\alpha+5\varepsilon}{3-\alpha+20\varepsilon} k_1. \tag{C.4}$$

The lower bound follows from the definition of k_2, we show the upper bound. First note that if $k_2 = k_0 + 1$, there is nothing to prove; suppose thus that $k_2 > k_0 + 1$. Then for every $k_0 < i < k_2$, there is some $i < i' \leq k_1$ so that $\sum_{\ell=i}^{i'} \log_2 R_\ell \leq (\alpha - s + \varepsilon)T(i - i')$; thus there is $k_2 \leq k \leq k_1$, so that

$$\sum_{\ell=k_0+1}^{k} \log_2 R_\ell \leq (\alpha - 20\varepsilon)T(k - k_0).$$

This, (C.3), and the fact that $\log_2 R_\ell \leq 3T$ for all ℓ imply that

$$3Tk_0 + (\alpha - 20\varepsilon)T(k - k_0) + 3T(k_1 - k) \geq 3Tk_0 + \sum_{\ell=k_0+1}^{k} \log_2 R_\ell$$
$$+ 3T(k_1 - k) \geq \sum_{\ell=1}^{k_1} \log_2 R_\ell \geq (\alpha - 2\varepsilon)T(k_1 - 2);$$

we conclude that $k(3 - \alpha + 20\varepsilon) \leq k_1\Big(3 - \alpha + 5\varepsilon\Big)$. This finishes the proof of (C.4) as $k_2 \leq k$.

Let now D be any $2^{-k_2 T}$-cube which intersects F_1. Let $k_2 < \ell \leq k_1$, and let $D' \subset D$ be a $2^{-\ell T}$-cube. Then

$$\#(D' \cap F_1) \leq \Big(\#(D \cap F_1)\Big) \cdot \prod_{i=k_2+1}^{\ell} R_i^{-1}.$$

Since $\sum_{k_2}^{\ell} \log_2 R_i \geq (\alpha - 20\varepsilon)T(\ell - k_2)$, we conclude that

$$\tfrac{\#(B(w,b)\cap D\cap F_1)}{\#(D\cap F_1)} \leq C'\Big(b/2^{-Tk_2}\Big)^{\alpha-20\varepsilon}$$

for all $b \geq (\#F)^{-1}$ where $C' \ll \varepsilon^{-\star}$ with absolute implied constants.

Let $F' = D \cap F_1$, and let $w_0 \in D \cap F_1$. The lemma holds with w_0, $b_1 = 2^{1-Tk_2}$, and $F' = D \cap F_1 \subset B(w_0, b_1)$. □

References

1. Bader, U., Fisher, D., Miller, N., Stover, M.: Arithmeticity, Superrigidity, and Totally Geodesic Submanifolds (2019)
2. Benoist, Y., Hee, O.: Effective equidistribution of S-integral points on symmetric varieties. Ann. Inst. Fourier (Grenoble) **62**(5), 1889–1942 (2012)
3. Benoist, Y., Oh, H.: Geodesic Planes in Geometrically Finite Acylindrical 3-Manifolds (2018)
4. Borel, A., Harish-Chandra: Arithmetic subgroups of algebraic groups. Ann. Math. **75**(2), 485–535 (1962)
5. Bourgain, J.: The discretized sum-product and projection theorems. J. Anal. Math. **112**, 193–236 (2010)
6. Bourgain, J., Furman, A., Lindenstrauss, E., Mozes, S.: Stationary measures and equidistribution for orbits of nonabelian semigroups on the torus. J. Am. Math. Soc. **24**(1), 231–280 (2011)
7. Bourgain, J., Gamburd, A.: On the spectral gap for finitely-generated subgroups of SU(2). Invent. Math. **171**(1), 83–121 (2008)
8. Bourgain, J., Gamburd, A.: Uniform expansion bounds for Cayley graphs of $\mathrm{SL}_2(\mathbb{F}_p)$. Ann. Math. (2) **167**(2), 625–642 (2008)
9. Burger, M., Sarnak, P.: Ramanujan duals. II. Invent. Math. **106**(1), 1–11 (1991)
10. Burger, M.: Horocycle flow on geometrically finite surfaces. Duke Math. J. **61**(3), 779–803 (1990)
11. Chow, S., Yang, L.: An Effective Ratner Equidistribution Theorem for Multiplicative Diophantine Approximation on Planar Lines (2019)
12. Clarkson, K.L., Edelsbrunner, H., Guibas, L.J., Sharir, M., Welzl, E.: Combinatorial complexity bounds for arrangements of curves and spheres. Discrete Comput. Geom. **5**(2), 99–160 (1990)
13. Clozel, L.: Démonstration de la conjecture τ. Invent. Math. **151**(2), 297–328 (2003)
14. Dani, S.G., Margulis, G.A.: Values of quadratic forms at primitive integral points. Invent. Math. **98**(2), 405–424 (1989)
15. Dani, S.G., Margulis, G.A.: Orbit closures of generic unipotent flows on homogeneous spaces of SL(3, R). Math. Ann. **286**(1–3), 101–128 (1990)
16. Duke, W., Rudnick, Z., Sarnak, P.: Density of integer points on affine homogeneous varieties. Duke Math. J. **71**(1), 143–179 (1993)
17. Einsiedler, M., Margulis, G., Mohammadi, A., Venkatesh, A.: Effective equidistribution and property (τ). J. Am. Math. Soc. **33**(1), 223–289 (2020)
18. Einsiedler, M., Margulis, G., Venkatesh, A.: Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces. Invent. Math. **177**(1), 137–212 (2009)
19. Einsiedler, M., Lindenstrauss, E., Michel, P., Venkatesh, A.: Distribution of periodic torus orbits on homogeneous spaces. Duke Math. J. **148**(1), 119–174 (2009)

20. Eskin, A., Margulis, G.: Recurrence properties of random walks on finite volume homogeneous manifolds. In: Random Walks and Geometry, pp. 431–444. Walter de Gruyter, Berlin (2004)
21. Eskin, A., Margulis, G., Mozes, S.: On a quantitative version of the Oppenheim conjecture. Electron. Res. Announc. Am. Math. Soc. **1**(3), 124–130 (1995)
22. Eskin, A., Margulis, G., Mozes, S.: Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture. Ann. Math. (2) **147**(1), 93–141 (1998)
23. Eskin, A., McMullen, C.: Mixing, counting, and equidistribution in lie groups. Duke Math. J. **71**(1), 181–209 (1993)
24. Eskin, A., Mirzakhani, M., Mohammadi, A.: Isolation, equidistribution, and orbit closures for the $\mathrm{SL}(2, \mathbb{R})$ action on moduli space. Ann. Math. (2) **182**(2), 673–721 (2015)
25. Fisher, D., Lafont, J.-F., Miller, N., Stover, M.: Finiteness of Maximal Geodesic Submanifolds in Hyperbolic Hybrids (2018)
26. Flaminio, L., Forni, G.: Invariant distributions and time averages for horocycle flows. Duke Math. J. **119**(3), 465–526 (2003)
27. Flaminio, L., Forni, G., Tanis, J.: Effective equidistribution of twisted horocycle flows and horocycle maps. Geom. Funct. Anal. **26**(5), 1359–1448 (2016)
28. Gamburd, A., Jakobson, D., Sarnak, P.: Spectra of elements in the group ring of $\mathrm{SU}(2)$. J. Eur. Math. Soc. **1**(1), 51–85 (1999)
29. Gorodnik, A., Maucourant, F., Hee, O.: Manin's and Peyre's conjectures on rational points and adelic mixing. Ann. Sci. Éc. Norm. Supér. (4) **41**(3), 383–435 (2008)
30. Green, B., Tao, T.: The quantitative behaviour of polynomial orbits on nilmanifolds. Ann. Math. (2) **175**(2), 465–540 (2012)
31. Gromov, M., Piatetski-Shapiro, I.I.: Non-arithmetic groups in Lobachevsky spaces. Publications Mathématiques de l'IHÉS **66**, 93–103 (1987); **(en)**
32. He, W., de Saxcé, N.: Linear Random Walks on the Torus (2019)
33. Käenmäki, A., Orponen, T., Venieri, L.: A Marstrand-Type Restricted Projection Theorem in $\mathbb{R}^3$ (2017)
34. Katz, A.: Quantitative Disjointness of Nilflows from Horospherical Flows (2019)
35. Kim, W.: Effective Equidistribution of Expanding Translates in the Space of Affine Lattices (2021)
36. Kleinbock, D., Margulis, G.: On effective equidistribution of expanding translates of certain orbits in the space of lattices. Dyn. Syst. **66**, 385–396 (2012)
37. Kleinbock, D.Y., Margulis, G.A.: Bounded orbits of nonquasiunipotent flows on homogeneous spaces. In: SinaUı's Moscow Seminar on Dynamical Systems, American Mathematical Society Translations Series 2, vol. 171, pp. 141–172. American Mathematical Society, Providence (1996)
38. Kleinbock, D.Y., Margulis, G.A.: Flows on homogeneous spaces and Diophantine approximation on manifolds. Ann. Math. (2) **148**(1), 339–360 (1998)
39. Kleinbock, D., Tomanov, G.M.: Flows on s-arithmetic homogeneous spaces and applications to metric diophantine approximation. Comment. Math. Helv. **82**(3), 519–581 (2007)
40. Kolasa, L., Wolff, T.: On some variants of the Kakeya problem. Pac. J. Math. **190**(1), 111–154 (1999)
41. Lindenstrauss, E., Margulis, G.: Effective estimates on indefinite ternary forms. Isr. J. Math. **203**(1), 445–499 (2014)
42. Lindenstrauss, E., Mohammadi, A., Margulis, G., Shah, N.: Quantitative Behavior of Unipotent Flows and an Effective Avoidance Principle (2019)
43. Luna, D.: Sur certaines operations differentiables des groupes de lie. Am. J. Math. **97**, 172–181 (1975)
44. Margulis, G.A.: Indefinite quadratic forms and unipotent flows on homogeneous spaces. In: Dynamical Systems and Ergodic Theory (Warsaw, 1986), vol. 23, pp. 399–409 (1989)

45. Margulis, G.A.: Discrete subgroups of semisimple Lie groups, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) [Results in Mathematics and Related Areas (3)], vol. 17. Springer, Berlin (1991). (92h:22021)
46. Margulis, G., Mohammadi, A.: Arithmeticity of Hyperbolic 3-Manifolds Containing Infinitely Many Totally Geodesic Surfaces (2019)
47. McAdam, T.: Almost-prime times in horospherical flows on the space of lattices. J. Mod. Dyn. **15**, 277–327 (2019)
48. Mohammadi, A., Oh, H.: Isolations of Geodesic Planes in the Frame Bundle of a Hyperbolic 3-Manifold (2020)
49. Mozes, S., Shah, N.: On the space of ergodic invariant measures of unipotent flows. Ergod. Theory Dyn. Syst. **15**(1), 149–159 (1995)
50. Platonov, V., Rapinchuk, A.: Algebraic Groups and Number Theory, Pure and Applied Mathematics, vol. 139. Academic Press, Boston (1994); Translated from the 1991 Russian original by Rachel Rowen
51. Ratner, M.: On measure rigidity of unipotent subgroups of semisimple groups. Acta Math. **165**(3–4), 229–309 (1990)
52. Ratner, M.: On Raghunathan's measure conjecture. Ann. Math. (2) **134**(3), 545–607 (1991)
53. Ratner, M.: Raghunathan's topological conjecture and distributions of unipotent flows. Duke Math. J. **63**(1), 235–280 (1991)
54. Sarnak, P.: Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series. Comm. Pure Appl. Math. **34**(6), 719–739 (1981)
55. Sarnak, P., Ubis, A.: The horocycle flow at prime times. J. Math. Pures Appl. (9) **103**(2), 575–618 (2015)
56. Schlag, W.: On continuum incidence problems related to harmonic analysis. J. Funct. Anal. **201**, 480–521 (2003)
57. Selberg, A.: Recent developments in the theory of discontinuous groups of motions of symmetric spaces. In: Proceedings of the Fifteenth Scandinavian Congress (Oslo, 1968) Lecture Notes in Mathematics, vol. 118, pp. 99–120. Springer, Berlin (1970)
58. Shah, N.A.: Limit distributions of expanding translates of certain orbits on homogeneous spaces. Proc. Indian Acad. Sci. Math. Sci. **106**(2), 105–125 (1996)
59. Shah, N.A.: Invariant measures and orbit closures on homogeneous spaces for actions of subgroups generated by unipotent elements. In: Lie Groups and Ergodic Theory (Mumbai, TATA Institute of Fundamental Research Studies in Mathematics, vol. 14, pp. 229–271. TATA Institute of Fundamental Research, Bombay (1996)
60. Strömbergsson, A.: On the uniform equidistribution of long closed horocycles. Duke Math. J. **123**(3), 507–547 (2004)
61. Strömbergsson, A.: An effective Ratner equidistribution result for $\mathrm{SL}(2,\mathbb{R}) \ltimes \mathbb{R}^2$. Duke Math. J. **164**(5), 843–902 (2015)
62. Strömbergsson, A., Vishe, P.: An effective equidistribution result for $\mathrm{SL}(2,\mathbb{R}) \ltimes (\mathbb{R}^2)^{\oplus k}$ and application to inhomogeneous quadratic forms. J. Lond. Math. Soc. (2) **102**(1), 143–204 (2020)
63. Tanis, J., Vishe, P.: Uniform bounds for period integrals and sparse equidistribution. Int. Math. Res. Not. **24**, 13728–13756 (2015)
64. Venkatesh, A.: Sparse equidistribution problems, period bounds and subconvexity. Ann. Math. (2) **172**(2), 989–1094 (2010)
65. Wolff, T.: Local smoothing type estimates on L^p for large p. Geom. Funct. Anal. **10**(5), 1237–1288 (2000)
66. Zahl, J.: L^3 estimates for an algebraic variable coefficient Wolff circular maximal function. Rev. Mat. Iberoam. **28**(4), 1061–1090 (2012)