Dimension estimates for the set of points with non-dense orbit in homogeneous spaces

Dmitry Kleinbock[1] · Shahriar Mirzadeh[2]

Abstract

Let $X = G/\Gamma$, where G is a Lie group and Γ is a lattice in G, and let U be a subset of X whose complement is compact. We use the exponential mixing results for diagonalizable flows on X to give upper estimates for the Hausdorff dimension of the set of points whose trajectories miss U. This extends a recent result of Kadyrov et al. (Dyn Syst 30(2):149–157, 2015) and produces new applications to Diophantine approximation, such as an upper bound for the Hausdorff dimension of the set of weighted uniformly badly approximable systems of linear forms, generalizing an estimate due to Broderick and Kleinbock (Int J Number Theory 11(7):2037–2054, 2015).

Mathematics Subject Classification Primary 37A17 · 37A25; Secondary 11J13

1 Introduction

Throughout the paper, we let G be a Lie group and Γ a lattice in G, denote by X the homogeneous space G/Γ and by μ the G-invariant probability measure on X. The notation

$$A \gg B \quad (\text{resp.}, A \underset{+}{\gg} B),$$

where A and B are quantities depending on certain parameters, will mean $A \geq CB$ (resp., $A \geq CB + D$), where C, D are constants dependent only on X and F. Let $F^+ := (g_t)_{t \geq 0}$ be a one-parameter subsemigroup of G. Following [10], for any subset U of X define the set

$$E(F^+, U) := \{x \in X : \overline{F^+ x} \cap U = \varnothing\} \tag{1.1}$$

✉ Dmitry Kleinbock
  kleinboc@brandeis.edu

  Shahriar Mirzadeh
  mirzade2@msu.edu

[1] Department of Mathematics, Brandeis University, Waltham, MA, USA

[2] Department of Mathematics, Michigan State University, East Lansing, MI, USA

of points in X whose F^+-orbits stay away from U. If the flow (X, μ, g_t) is ergodic, then the orbit $\{g_t x\}_{t \geq 0}$ is dense for μ-almost all $x \in X$; hence $\mu(E(F^+, U)) = 0$ whenever U is non-empty.

A natural question one can ask is: how large can this set of measure zero be? If the semigroup F^+ is *quasiunipotemt*, that is, all eigenvalues of Ad g_1 have absolute value 1, then, whenever the action is ergodic and U is non-empty, the set (1.1) is contained in a countable union of proper submanifolds of X – this follows from Ratner's Measure Classification Theorem and the work of Dani and Margulis, see [21, Lemma 21.2] and [6, Proposition 2.1]. On the other hand, if F^+ is not quasiunipotemt and $U = \{z\}$ for some $z \in X$, it is shown in [11] that the set (1.1) has full Hausdorff dimension.

Fix a right-invariant Riemannian structure on G, and denote by 'dist' the corresponding Riemannian metric, using the same notation for the induced metric on X. Also denote by $B(r)$ the open ball of radius r centered at the identity element of G, and by $B(z, r)$ the open ball of radius r centered at $z \in X$. The aforementioned result of [11] can thus be stated as

$$\dim E\big(F^+, B(z, r)\big) \to \dim X \text{ as } r \to 0. \tag{1.2}$$

Here and hereafter dim E means the Hausdorff dimension of the set E, and codim E will stand for its Hausdorff codimension, i.e. the difference between the dimension of the ambient set and the Hausdorff dimension of E. Until recently a problem of estimating the left hand side of (1.2), or more generally, the quantity $\dim E(F^+, U)$ where U is a non-empty open subset of X, has not been addressed. In [2] Broderick and the first named author considered the case

$$G = \mathrm{SL}_{m+n}(\mathbb{R}), \ \Gamma = \mathrm{SL}_{m+n}(\mathbb{Z}), \ X = G/\Gamma, \tag{1.3}$$

with the action of $F^+ = (g_t)_{t \geq 0}$ where

$$g_t = \mathrm{diag}(e^{t/m}, \ldots, e^{t/m}, e^{-t/n}, \ldots, e^{-t/n}), \tag{1.4}$$

This action is important because of its Diophantine applications. In particular, a system of linear forms is badly approximable if and only if (see [5]) the g_t-trajectory of a certain element of X does not enter the set

$$U(\varepsilon) := \big\{g\Gamma \in X : \delta(g\Gamma) < \varepsilon\big\} \tag{1.5}$$

for some $\varepsilon > 0$, where

$$\delta(g\Gamma) := \inf_{\mathbf{v} \in \mathbb{Z}^{m+n} \smallsetminus \{0\}} \|g\mathbf{v}\|. \tag{1.6}$$

It was essentially[1] shown there that for all $\varepsilon > 0$ one has

$$\mathrm{codim}\, E\big(F^+, U(\varepsilon)\big) \gg \frac{\varepsilon^{m+n}}{\log(1/\varepsilon)}. \tag{1.7}$$

The main ingredient of the proof in [2] was the exponential mixing of the g_t-action on X (see Sect. 2 for the definition). This theme was continued by Kadyrov in [10], where an estimate similar to (1.7) was proved for the Hausdorff dimension of $E(F^+, B(z, r))$ under the assumptions that $X = G/\Gamma$ is compact and the F^+-action on X is exponentially mixing.

---

[1] [2, Theorem 1.3] is stated in a number-theoretic language; however it readily implies (1.7).in view of [2, Lemma 3.1]. Note that recently a precise asymptotic formula for the left hand side of (1.7) was obtained by Simmons [20]: namely, that as $\varepsilon \to 0$, the ratio $\frac{\mathrm{codim}\, E(F^+, U(\varepsilon))}{\varepsilon^{m+n}}$ tends to a constant depending only on m, n.

Namely, it is shown there that there exist $r_0 > 0$ i such that for any $r \in (0, r_0)$ and any $z \in X$ one has

$$\operatorname{codim} E\left(F^+, B(z, r)\right) \gg \frac{r^{\dim X}}{\log(1/r)}. \tag{1.8}$$

In the present paper we strengthen Kadyrov's result in two ways: by considering more general open sets U in place of balls $B(z, r)$, and by relaxing the assumption of compactness of X to that of compactness of $X \smallsetminus U$. Our main theorem generalizes results from both [2,10] and can be used to produce new applications to Diophantine approximation.

We need to introduce the following notation: for a subset U of X and $r > 0$ denote by $\sigma_r U$ the *inner r-core* of U, defined as

$$\sigma_r U := \{x \in X : \operatorname{dist}(x, U^c) > r\},$$

and by $\partial_r U$ the *r-neighborhood* of U by

$$\partial_r U := \{x \in X : \operatorname{dist}(x, U) < r\}.$$

Also, for $x \in X$ denote by π_x the map $G \to X$ given by $\pi_x(g) := gx$, and by $r_0(x)$ the *injectivity radius* of x:

$$r_0(x) := \sup\{r > 0 : \pi_x \text{ is injective on } B(r)\}.$$

If $K \subset X$ is bounded, let us denote by $r_0(K)$ the *injectivity radius* of K:

$$r_0(K) := \inf_{x \in K} r_0(x) = \sup\{r > 0 : \pi_x \text{ is injective on } B(r) \ \forall x \in K\}.$$

Here is the main result of the paper:

Theorem 1.1 *Let G be a Lie group, Γ a lattice in G, $X = G/\Gamma$, and let F^+ be a one-parameter* Ad-*diagonalizable subsemigroup of G whose action on X is exponentially mixing. Then there exists $r' > 0$ such that for any $U \subset X$ such that U^c is compact and any $0 < r < \min(r_0(\partial_1 U^c), r')$ one has*

$$\operatorname{codim} E(F^+, U) \gg \frac{\mu(\sigma_r U)}{\log(1/r) + \log(1/\mu(\sigma_r U))}. \tag{1.9}$$

We note that in the above inequality, as well as in similar statements below, the implicit constant in $\gg$ is independent of U and r and is only dependent on X and F. Also note that the right hand side of (1.9) depends on r while the left hand side does not. Since the inequality holds for all sufficiently small values of r, in applications one needs to choose an optimal r to strengthen the result. In particular, it is not hard to see, by taking U to be an open ball of radius r centered at z and assuming that X is compact, that Kadyrov's result (1.8) is a special case of (1.9). Moreover one has the following generalization:

Corollary 1.2 *Let F^+ be as in Theorem 1.1. Assume that X is compact. Then there exists $r' > 0$ such that for any closed subset S of X and any $0 < r < r'$ one has*

$$\operatorname{codim} E(F^+, \partial_r S) \gg \frac{\mu(\partial_{r/2} S)}{\log(1/r)}.$$

Consequently, if $S \subset X$ is a k-dimensional compact embedded submanifold, then for some $C = C(S, F)$ and any $0 < r < r'$ one has

$$\operatorname{codim} E(F^+, \partial_r S) \geq C \frac{r^{\dim X - k}}{\log(1/r)}. \tag{1.10}$$

The case $k = 0$ and $S = \{z\}$ of (1.10) coincides with (1.8): it is easy to show, by looking at the proof, that $C(\{z\}, F)$ is independent on $z \in X$.

Similarly to the previous papers [2,10] on the subject, the main theorem is deduced from a result that estimates

$$\dim E(F^+, \sigma_r U) \cap Hx,$$

where $x \in X$ and H is the *unstable horospherical subgroup* with respect to F^+, defined as

$$H := \{g \in G : \text{dist}(g_t g g_{-t}, e) \to 0 \ as \ t \to -\infty\}. \tag{1.11}$$

More generally, in the following theorem we estimate

$$\dim E(F^+, \sigma_r U) \cap Px$$

for $x \in X$ and some proper subgroups P of H, namely those which have *Effective Equidistribution Property* (EEP, see Sect. 2 for the definition) with respect to the flow (X, F^+). Note that for $P = H$ this property follows from the exponential mixing of the action, as shown in [14].

Theorem 1.3 *Let G, Γ and X be as in Theorem 1.1, let F^+ be a one-parameter Ad-diagonalizable subsemigroup of G, and let P be a subgroup of H which has property* (EEP) *with respect to the flow (X, F^+). Then there exists $r'' > 0$ such that for any $x \in X$, any $U \subset X$ such that U^c is compact and any $0 < r < \min(r_0(\partial_{1/2} U^c), r'')$ one has*

$$\text{codim}\{g \in P : gx \in E(F^+, U)\} \gg \frac{\mu(\sigma_r U)}{\log \frac{1}{r} + \log \frac{1}{\mu(\sigma_r U)}} . \tag{1.12}$$

The general statement of Theorem 1.3 makes it possible to derive a corollary involving simultaneous Diophantine approximation with weights. Take

$$\mathbf{i} = (i_k : k = 1, \ldots, m) \text{ and } \mathbf{j} = (j_\ell : \ell = 1, \ldots, n)$$

with

$$i_k, j_\ell > 0 \ \ and \ \ \sum_{k=1}^{m} i_k = 1 = \sum_{\ell=1}^{n} j_\ell, \tag{1.13}$$

and define the **i**-*quasinorm* of $\mathbf{x} \in \mathbb{R}^m$ and the **j**-*quasinorm* of $\mathbf{y} \in \mathbb{R}^n$ by

$$\|\mathbf{x}\|_{\mathbf{i}} := \max_{1 \le k \le m} |x_k|^{1/i_k} \text{ and } \|\mathbf{y}\|_{\mathbf{j}} := \max_{1 \le \ell \le n} |y_\ell|^{1/j_\ell}.$$

A system of linear forms given by $A \in M_{m,n}(\mathbb{R})$ is said to be $(\mathbf{i}, \mathbf{j})$-*badly approximable* if

$$\inf_{\mathbf{p} \in \mathbb{Z}^m, \ \mathbf{q} \in \mathbb{Z}^n \smallsetminus \{0\}} \|A\mathbf{q} + \mathbf{p}\|_{\mathbf{i}} \|\mathbf{q}\|_{\mathbf{j}} > 0$$

This generalizes the notion of (unweighted) badly approximable systems of linear forms, which correspond to the choice of equal weights

$$\mathbf{i} = \mathbf{m} := (1/m, \ldots, 1/m), \quad \mathbf{j} = \mathbf{n} := (1/n, \ldots, 1/n). \tag{1.14}$$

Now for any $c > 0$ set

$$\text{Bad}_{\mathbf{i},\mathbf{j}}(c) := \{A \in M_{m,n} : \inf_{\mathbf{p} \in \mathbb{Z}^m, \ \mathbf{q} \in \mathbb{Z}^n \smallsetminus \{0\}} \|A\mathbf{q} + \mathbf{p}\|_{\mathbf{i}} \|\mathbf{q}\|_{\mathbf{j}} \ge c\}. \tag{1.15}$$

It is known, see [18, Theorem 2] and [17, Corollary 4.5], that for any $\mathbf{i}, \mathbf{j}$ as in (1.13) the set of $(\mathbf{i}, \mathbf{j})$-badly approximable systems of linear forms, which is the union of the sets $\text{Bad}_{\mathbf{i},\mathbf{j}}(c)$

over $c > 0$, has Hausdorff dimension mn. One can ask for an estimate for the Hausdorff dimension of $\text{Bad}_{\mathbf{i},\mathbf{j}}(c)$ for fixed $\mathbf{i}$, $\mathbf{j}$ and c. Our goal in Sect. 8 is to deduce the following theorem from Theorem 1.3:

Theorem 1.4 *There exists $c_0 > 0$ such that for any $\mathbf{i}, \mathbf{j}$ as in (1.13) and any $0 < c < c_0$ one has*

$$\text{codim Bad}_{\mathbf{i},\mathbf{j}}(c) \gg \frac{c}{\log \frac{1}{c}},$$

where the implicit constant in $\gg$ is independent of c but depends on $\mathbf{i}$, $\mathbf{j}$.

This is a weighted generalization of [2, Theorem 1.3]. Note that in the paper [20], mentioned in the footnote before (1.7), it is shown that codim $\text{Bad}_{\mathbf{m},\mathbf{n}}(c)$ is asymptotic to a constant times c as $c \to 0$. However the methods of [20] do not seem to extend to the weighted case.

The structure of the paper is as follows. In the next section we define exponential mixing and property (EEP), and, following [14,16], show that the exponential mixing of the g_t-action on X implies (EEP) for the expanding horospherical subgroup relative to g_1. In Sect. 3 we deduce Theorem 1.1 and Corollary 1.2 from Theorem 1.3. The next three sections are devoted to proving Theorem 1.3. In Sect. 8 we prove Theorem 1.4 by reducing the problem to dynamics on the space G/Γ with G and Γ as in (1.3) and

$$g_t = g_t^{\mathbf{i},\mathbf{j}} := \text{diag}(e^{i_1 t}, \ldots, e^{i_m t}, e^{-j_1 t}, \ldots, e^{-j_n t}). \tag{1.16}$$

Theorem 1.3 is then applied to the subgroup

$$P = \left\{ \begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix} : A \in M_{m,n}(\mathbb{R}) \right\} \tag{1.17}$$

of G, which, following [16], is shown in Sect. 7 to satisfy property (EEP) relative to the $g_t^{\mathbf{i},\mathbf{j}}$-action. We conclude the paper with a few remarks and open questions.

2 Exponential mixing implies (EEP) for H

We start with the definition of Sobolev spaces on Lie groups and their homogeneous spaces. Let G be a Lie group and Γ a discrete subgroup of G. Denote by X the homogeneous space G/Γ and by N the dimension of G. In what follows, $\| \cdot \|_p$ will stand for the L^p norm, and $(\cdot, \cdot)$ for the inner product in $L^2(X, \mu)$, where μ is a (fixed) G-invariant measure on X. If Γ is a lattice in G, we will always take μ to be the probability measure. Note though that much of the set-up below applies to the case $\Gamma = \{e\}$ and $X = G$.

Fix a basis $\{Y_1, \ldots, Y_n\}$ for the Lie algebra $\mathfrak{g}$ of G, and, given a smooth function $h \in C^\infty(X)$ and $\ell \in \mathbb{Z}_+$, define the "L^p, order ℓ" *Sobolev norm* $\|h\|_{\ell,p}$ of h by

$$\|h\|_{\ell,p} \overset{\text{def}}{=} \sum_{|\alpha| \leq \ell} \|D^\alpha h\|_p,$$

where $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multiindex, $|\alpha| = \sum_{i=1}^n \alpha_i$, and D^α is a differential operator of order $|\alpha|$ which is a monomial in $Y_1, \ldots, Y_n$, namely $D^\alpha = Y_1^{\alpha_1} \ldots Y_n^{\alpha_n}$. This definition depends on the basis, however, a change of basis would only distort $\|h\|_{\ell,p}$ by a bounded factor. We also let

$$C_2^\infty(X) = \{h \in C^\infty(X) : \|h\|_{\ell,2} < \infty \text{ for any } \ell = \mathbb{Z}_+\}.$$

Clearly smooth compactly supported functions belong to $C_2^\infty(X)$. We will also use the operators D^α to define C^ℓ norms of smooth functions f on X:

$$\|f\|_{C^\ell} := \sup_{x \in X, \, |\alpha| \leq \ell} |D^\alpha f(x)|.$$

Definition 2.1 Let $F^+ = \{g_t : t \geq 0\}$ be a one-parameter subsemigroup of G, and let $X = G/\Gamma$ where Γ is a lattice in G. We say that a flow (X, F^+) is *exponentially mixing* if there exist $\gamma > 0$ and $\ell \in \mathbb{Z}_+$ such that for any $\varphi, \psi \in C_2^\infty(X)$ and for any $t \geq 0$ one has

$$\left| (g_t \varphi, \psi) - \int_X \varphi \, d\mu \int_X \psi \, d\mu \right| \ll e^{-\gamma t} \|\varphi\|_{\ell,2} \|\psi\|_{\ell,2}. \tag{2.1}$$

As is the case in many applications, we will use the exponential mixing to study expanding translates of pieces of certain subgroups of G. If $P \subset G$ is a subgroup with a fixed Haar measure ν, ψ a function on X, f a function on P, $x \in X$ and $t \geq 0$, let us define

$$I_{f,\psi}(g_t, x) := \int_P f(h) \psi(g_t h x) \, d\nu(h).$$

Definition 2.2 Say that a subgroup P of G has *Effective Equidistribution Property* (EEP) with respect to the flow (X, F^+) if P is normalized by F^+, and there exists $\lambda > 0$ and $\ell \in \mathbb{N}$ such that for any $x \in X$ and $t > 0$ with

$$t \gg_+ \log \frac{1}{r_0(x)}, \tag{2.2}$$

any $f \in C_{comp}^\infty(P)$ with supp $f \subset B^P(1)$ and any $\psi \in C_2^\infty(X)$ it holds that

$$\left| I_{f,\psi}(g_t, x) - \int_P f \, d\nu \int_X \psi \, d\mu \right| \ll \max(\|\psi\|_{C^1}, \|\psi\|_{\ell,2}) \cdot \|f\|_{C^\ell} \cdot e^{-\lambda t}. \tag{2.3}$$

Here ν stands for a Haar measure on P. Note that the implicit constants in both (2.2) and (2.3) are independent on f, ψ, t and x. This definition is quite involved but it is justified by the fact that in many special cases (2.3) can be derived from exponential mixing, for example when $P = H$, the unstable horospherical subgroup relative to F^+. This was essentially proved in [16], together with another important example of a proper subgroup of H with the same property, namely with P as in (1.17). We are going to revisit the argument from that paper and make the constants appearing there explicit.

Remark 2.3 Note that it suffices to establish (EEP) for functions ψ with $\int_X \psi \, d\mu = 0$: indeed, if $\psi_0 := \psi - \int_X \psi \, d\mu$, one clearly has

$$I_{f,\psi_0}(g_t, z) = I_{f,\psi}(g_t, z) - \int_H f \, d\nu \int_X \psi \, d\mu.$$

Let $\mathfrak{g}$ be a Lie algebra of G, $\mathfrak{g}_\mathbb{C}$ its complexification, and for $\lambda \in \mathbb{C}$, let E_λ be the eigenspace of Ad g_1 corresponding to λ. Let $\mathfrak{h}, \mathfrak{h}^\circ, \mathfrak{h}^-$ be the subalgebras of $\mathfrak{g}$ with complexifications:

$$\mathfrak{h}_\mathbb{C} = \mathrm{span}(E_\lambda : |\lambda| > 1), \quad \mathfrak{h}_\mathbb{C}^0 = \mathrm{span}(E_\lambda : |\lambda| = 1), \quad \mathfrak{h}_\mathbb{C}^- = \mathrm{span}(E_\lambda : |\lambda| < 1).$$

Let H, H^0, H^- be the corresponding subgroups of G. Note that H is precisely the unstable horospherical subgroup with respect to F^+ [defined in (1.11)] and H^- is the stable horospherical subgroup defined by:

$$H^- = \{h \in G : g_t h g_{-t} \to e \text{ as } t \to +\infty\}.$$

Since $\mathrm{Ad}\, g_1$ is assumed to be diagonalizable over $\mathbb{C}$, $\mathfrak{g}$ is the direct sum of $\mathfrak{h}$, $\mathfrak{h}^\circ$ and $\mathfrak{h}^-$. Therefore G is locally (at a neighborhood of identity) a direct product of the subgroups H, H^0 and H^-. In what follows, if P is a subgroup of G, we will denote by $B^P(r)$ the open ball of radius r centered at the identity element with respect to the metric on P corresponding to the Riemannian structure induced from G.

Denote the group H^-H^0 by $\tilde{H}$, and fix $0 < \rho < 1$ with the following properties:

$$\text{the multiplication map } \tilde{H} \times H \to G \text{ is one to one on } B^{\tilde{H}}(\rho) \times B^H(\rho), \qquad (2.4)$$

and

$$g_t B^{\tilde{H}}(r) g_{-t} \subset B^{\tilde{H}}(2r) \text{ for any } 0 < r < \rho \text{ and } t \geq 0 \qquad (2.5)$$

(the latter can be done since F is Ad-diagonalizable and the restriction of the map $g \to g_t g g_{-t}$, $t > 0$, to the subgroup $\tilde{H}$ is non-expanding).

Let μ^G be the Haar measure on G which locally projects to μ, and let us choose Haar measures ν^-, ν^0 and ν on H^-, H^0 and H respectively, normalized so that μ is locally almost the product of ν^-, ν^0 and ν. More precisely, see [4, Ch. VII, Sect. 9, Proposition 13], μ can be expressed via ν^-, ν^0 and ν in the following way: for any $\varphi \in L^1(G, \mu^G)$ supported on a small neighborhood of identity,

$$\int_G \varphi(g)\, d\mu(g) = \int_{H^- \times H^0 \times H} \varphi(h^- h^0 h)\Delta(h^0)\, d\nu^-(h^-)\, d\nu^0(h^0)\, d\nu(h), \qquad (2.6)$$

where Δ is the modular function of (the non-unimodular group) $\tilde{H}$.

Now we are going to show, following [16], that H, the unstable horospherical subgroup of G with respect to F^+, satisfies property (EEP). We will start with an auxiliary statement, essentially[2] established in [16, Theorem 2.3]:

Theorem 2.4 *Suppose that the flow (X, F^+) is exponentially mixing, and let γ and ℓ be as in (2.1). Then for any $f \in C^\infty_{comp}(H)$, $0 < r < \rho/2$ and $x \in X$, if*

(i) $\mathrm{supp}\, f \subset B^H(r)$, *and*
(ii) π_x *is injective on $B^G(2r)$,*

then for any $\psi \in C^\infty_2(X)$ with $\int_X \psi\, d\mu = 0$ and any $t \geq 0$ one has

$$\left| I_{f,\psi}(g_t, x) \right| \ll \max\left(\|\psi\|_{C^1}, \|\psi\|_{\ell,2} \right) \left(r\|f\|_1 + e^{-\gamma t} r^{-(\ell + \tilde{k}/2)} \|f\|_{\ell,2} \right),$$

where $\tilde{k} = \dim \tilde{H}$.

Using this and again following [16], we can establish

Theorem 2.5 *H satisfies property* (EEP) *with respect to the flow (X, F^+).*

For the proof and for later applications we will need the following lemma, which is a modification of [14, Lemma 2.4.7(b)] and [16, Lemma 2.2(a)]:

Lemma 2.6 *Let G be a Lie group of dimension N. Then for each $\ell \in \mathbb{Z}_+$ there exists M_ℓ (depending only on G) with the following property: for any $0 < \varepsilon < 1$ there exists a nonnegative smooth function φ_ε on G such that*

---

[2] The statement of [16, Theorem 2.3] featured a constant $E(\psi)$ in place of $\max(\|\psi\|_{C^1}, \|\psi\|_{\ell,2})$, but it is easy to see from the proof that E depends linearly on $\|\psi\|_{C^1}$ and $\|\psi\|_{\ell,2}$.

(1) *the support of φ_ε is inside the ball of radius ε centered at e;*
(2) $\|\varphi_\varepsilon\|_1 = 1$;
(3) $\|\varphi_\varepsilon\|_{C^\ell} \le M_\ell \cdot \varepsilon^{-(\ell+N)}$;
(4) $\|\varphi_\varepsilon\|_{\ell,p} \le M_\ell \cdot \varepsilon^{-(\ell+\frac{p-1}{p}N)}$.

Proof of Theorem 2.5 Suppose we are given $f \in C^\infty_{comp}(H)$ with supp $f \subset B^H(1)$, $\psi \in C^\infty_2(X)$ with $\int_X \psi \, d\mu = 0$, and $x \in X$. Put $r = e^{-\beta t}$, where β is to be specified later, and take ℓ as in (2.1). Then, using Lemma 2.6 with G replaced by H, take a non-negative smooth function θ supported on $B^H(r)$ such that

$$\int_H \theta \, d\nu = 1 \text{ and } \|\theta\|_{\ell,2} \ll r^{-(\ell+k/2)}, \qquad (2.7)$$

where $k = \dim H = N - \tilde{k}$. Since ν is translation-invariant, one can write

$$I_{f,\psi}(g_t, x) = \int_H f(h)\psi(g_t hx) \, d\nu(h) \int_H \theta(y) \, d\nu(y)$$

$$= \int_H \int_H f(yh)\theta(y)\psi(g_t yhx) \, d\nu(y) \, d\nu(h)$$

$$= \int_H \int_H f(yh)\theta(y)\psi(g_t yhx) \, d\nu(y) \, d\nu(h).$$

Note that, as long as $\theta(y) \ne 0$, the supports of all functions of the form $h \mapsto f(yh)$ are contained in $\tilde{B} := B^H(2)$. We would like to apply Theorem 2.4 with $r = e^{-\beta t}$, hx in place of x and

$$f_h(y) := f(yh)\theta(y)$$

in place of f. It is clear that supp $f_h \subset B^H(r)$ for any h, i.e. condition (i) of Theorem 2.4 is satisfied. For other conditions we need to require $e^{-\beta t} \le \min(r_0(hx)/2, \rho/2)$. Since $r_0(hx) \gg r_0(x)$ as long as $h \in \tilde{B}$, it amounts to assuming

$$2e^{-\beta t} \le a_0 \min(r_0(x), \rho) \qquad (2.8)$$

for some uniform constant $a_0 > 0$. Also, in view of [16, Lemma 2.2(b)] and (2.7), we have

$$\|f_h\|_{\ell,2} \ll \|f\|_{C^\ell}\|\theta\|_{\ell,2} \ll e^{(\ell+k/2)\beta t}\|f\|_{C^\ell}.$$

Then from Theorem 2.4 one gets

$$\left| I_{f,\psi}(g_t, x) \right| = \left| \int_{\tilde{B}} \int_H f(yh)\theta(y)\psi(g_t yhx) \, d\nu(y) \, d\nu(h) \right| \le \int_{\tilde{B}} \left| I_{f_h,\psi}(g_t, hx) \right| \, d\nu(h)$$

$$\ll \max\left(\|\psi\|_{C^1}, \|\psi\|_{\ell,2} \right) \left(e^{-\beta t} \int_H |f_h| \, d\nu(h) + e^{(\ell+\tilde{k}/2)\beta t}\|f_h\|_{\ell,2} \cdot e^{-\gamma t} \right) \nu(\tilde{B})$$

$$\ll \max\left(\|\psi\|_{C^1}, \|\psi\|_{\ell,2} \right) \left(\sup |f| \cdot e^{-\beta t} + \|f\|_{C^\ell} \cdot e^{-(\gamma-(2\ell+\frac{N}{2})\beta)t} \right).$$

An elementary computation shows that choosing β equalizing the two exponents above will produce

$$\beta = \lambda = \frac{\gamma}{1 + 2\ell + N/2},$$

and therefore (2.8) becomes equivalent to (2.2) with some uniform constants a, b. This shows that (2.2) implies (2.3), and finishes the proof. $\qquad \square$

3 Proving Theorem 1.1 and Corollary 1.2

We now assume Theorem 1.3 is true and give a proof of Theorem 1.1.

Proof of Theorem 1.1 assuming Theorem 1.3 Let r'' be as in Theorem 1.3, and define

$$r' := \min\left(1/4, r'', \rho\right) \tag{3.1}$$

where ρ is as in (2.4), (2.5). For any $r \le \rho$ choose s such that $B(s)$ is contained in the product $B^{\tilde{H}}(r/4)B^H(r/4)$. Now take $U \subset X$ such that U^c is compact, and for $x \in X$ denote

$$E_{x,s} := \{g \in B(s) : gx \in E(F^+, U)\}. \tag{3.2}$$

In view of the countable stability of Hausdorff dimension, in order to prove the theorem it suffices to prove that for any $x \in X$,

$$\dim E_{x,s} \le \dim X - C \frac{\mu(\sigma_r U)}{\log \frac{1}{r} + \log \frac{1}{\mu(\sigma_r U)}} \tag{3.3}$$

with the constant $C > 0$ only dependent on X and F. Indeed, $E(F^+, U)$ can be covered by countably many sets $\{gx : g \in E_{x,s}\}$, with the maps $\pi_x : E_{x,s} \to X$ being Lipschitz and at most finite-to-one.

Since every $g \in B(s)$ can be written as $g = h'h$, where $h' \in B^{\tilde{H}}(r/4)$ and $h \in B^H(r/4)$, for any $y \in X$ we can write

$$\begin{aligned}
\mathrm{dist}(g_t g x, y) &\le \mathrm{dist}(g_t h' h x, g_t h x) + \mathrm{dist}(g_t h x, y) \\
&= \mathrm{dist}\left(g_t h' g_{-t} g_t h x, g_t h x\right) + \mathrm{dist}(g_t h x, y).
\end{aligned} \tag{3.4}$$

Hence in view of (2.5), $g \in E_{x,s}$ implies that hx belongs to $E(F^+, \sigma_{r/2}U)$, and by using Wegmann's Product Theorem [23] we conclude that:

$$\begin{aligned}
\dim E_{x,s} &\le \dim\left(\{h \in B^H(r/4) : hx \in E(F^+, \sigma_{r/2}U)\} \times B^{\tilde{H}}(r/4)\right) \\
&\le \dim\left(\{h \in B^H(r/4) : hx \in E(F^+, \sigma_{r/2}U)\}\right) + \dim \tilde{H}.
\end{aligned} \tag{3.5}$$

Since $\partial_{1/2}(\sigma_{r/2}U)^c$ is contained in $\partial_1 U^c$, we have:

$$r_0(\partial_1 U^c) \le r_0\left(\partial_{1/2}(\sigma_{r/2}U)^c\right).$$

Therefore, by Theorems 2.5 and 1.3 applied to $P = H$ and U replaced by $\sigma_{r/2}U$, there exists a constant $C > 0$, only dependent on X and F, such that the set $\{h \in B^H(r/4) : hx \in E(F^+, \sigma_{r/2}U)\}$ has Hausdorff dimension at most

$$\begin{aligned}
\dim H - C \frac{\mu(\sigma_{r/4}\sigma_{r/2}U)}{\log \frac{4}{r} + \log \frac{1}{\mu(\sigma_r U)}} &\le \dim H - C \frac{\mu(\sigma_r U)}{\log \frac{4}{r} + \log \frac{1}{\mu(\sigma_r U)}} \\
&\le \dim H - C' \frac{\mu(\sigma_r U)}{\log \frac{1}{r} + \log \frac{1}{\mu(\sigma_r U)}},
\end{aligned} \tag{3.6}$$

where $C' = 2C$. (C' should be chosen so that we have

$$C' \ge C \cdot \frac{\log \frac{4}{r} + \log \frac{1}{\mu(\sigma_r U)}}{\log \frac{1}{r} + \log \frac{1}{\mu(\sigma_r U)}} = C \cdot \left(1 + \frac{\log 4}{\log \frac{1}{r} + \log \frac{1}{\mu(\sigma_r U)}}\right).$$

Since $r < 1/4$, we can choose $C' = 2C$.) It follows from (3.5) and (3.6) that

$$\dim E_{x,s} \leq \dim X - C' \frac{\mu(\sigma_r U)}{\log \frac{1}{r} + \log \frac{1}{\mu(\sigma_r U)}},$$

which finishes the proof. □

Proof of Corollary 1.2 Take r' as in (3.1). If $S = \varnothing$ there is nothing to prove. Otherwise, by Theorem 1.1 applied to $U = \partial_r S$ and with $r/2$ in place of r, there exists a constant $C > 0$ independent of S such that for any $0 < r < \min(r_0(X), r')$, the set $E(F^+, \partial_r S)$ has Hausdorff codimension at most

$$C \frac{\mu\left(\sigma_{r/2}(\partial_r S)\right)}{\log \frac{2}{r} + \log \frac{1}{\mu(\sigma_{r/2}(\partial_r S))}} \geq C \frac{\mu(\partial_{r/2} S)}{\log \frac{2}{r} + \log \frac{1}{\mu(\partial_{r/2} S)}}. \tag{3.7}$$

Since S is non-empty, $\partial_{r/2} S$ contains a ball of radius $r/2$, so there exists a constant d_0 independent of r such that for any $0 < r < r_0(X)$ we have:

$$\mu(\partial_{r/2} S) \geq d_0 r^N. \tag{3.8}$$

Since $r' < 1/4$, by combining (3.7) and (3.8) it is easy to see that the set $E(F^+, \partial_r S)$ has Hausdorff codimension at most

$$C \frac{\mu(\partial_{r/2} S)}{(N+1) \log \frac{1}{r} + \log 2 + \log \frac{1}{d_0}} \geq \frac{C \log 4}{(N+1) \log 4 + \log 2 + \log \frac{1}{d_0}} \cdot \frac{\mu(\partial_{r/2} S)}{\log \frac{1}{r}}.$$

This proves the main part of the corollary.

For the "consequently" part, if S is a k-dimensional compact embedded submanifold in X, then it is easy to see that for some constant d_1 dependent on S and for all $r < r_0(X)$ one has

$$\mu(\partial_{r/2} S) \geq d_1 r^{N-k}. \tag{3.9}$$

Therefore in this case, combining (3.7) and (3.9), it is easy to see that for any $0 < r < \min(r_0(X), r')$ one has

$$\operatorname{codim} E(F^+, \partial_r S) \geq \frac{C \log 4}{(N-k+1) \log 4 + \log 2 + \log \frac{1}{d_1}} \cdot \frac{r^{N-k}}{\log \frac{1}{r}}.$$

□

4 Reduction to a covering result

In the next three sections our goal is to prove Theorem 1.3. Fix a subgroup P of H that satisfies (EEP) relative to F^+, and fix a Haar measure ν on P. Put $L = \dim P$. Also take $0 < r'' < 1/8$ such that the exponential map from $\mathfrak{p} := \operatorname{Lie}(P)$ to P is 2-bi-Lipischitz on the ball of radius r'' centered at $0 \in \mathfrak{p}$, The latter implies that there exist constants $c_1, c_2, c_3 > 0$ such that for any $0 < r < r''$ one nas

$$c_1 r^L \leq \nu\left(B^P(r)\right) \leq c_2 r^L \tag{4.1}$$

and

$$\frac{d}{dr} \nu\left(B^P(r)\right) \leq c_3 r^{L-1}. \tag{4.2}$$

For $x \in X$, $t > 0$, $k \in \mathbb{N}$ and a subset S of X we define

$$A^P(t, r, S, k, x) := \left\{ h \in B^P(r) : g_{\ell t} h x \in S \; \forall \ell \in \{1, 2, \cdots, k\} \right\}. \tag{4.3}$$

Also, let us define

$$\lambda_{\max} := \max\{|\lambda| : \lambda \text{ is an eigenvalue of } \mathrm{ad}_{g_1}|_{\mathfrak{p}}\}.$$

One of our main goals in the next three sections will be to prove the following theorem:

Theorem 4.1 *Let F^+ be a one-parameter* Ad-*diagonalizable subsemigroup of G, and P a subgroup of G with property* (EEP). *Then there exist positive constants a, b, K_0, K_1, K_2 and λ_1 such that for any subset U of X whose complement is compact, any $0 < r < r_0$ where*

$$r_0 := \min\left(r_0(\partial_{1/2} U^c), r''\right), \tag{4.4}$$

any $x \in \partial_r U^c$, $k \in \mathbb{N}$ and any

$$t > a + b \log \frac{1}{r}, \tag{4.5}$$

the set $A^P(t, \frac{r}{16\sqrt{L}}, U^c, k, x)$ can be covered with at most

$$K_0 e^{Lk\lambda_{\max}t} \left(1 - K_1 \mu(\sigma_r U) + \frac{K_2 e^{-\lambda_1 t}}{r^L}\right)^k$$

balls in P of radius $re^{-k\lambda_{\max}t}$.

It is not hard to see a connection between the above theorem and Theorem 1.3: indeed, for any $x \in X$ the intersection of the set in the left hand side of (1.12) with $B^P(\frac{r}{16\sqrt{L}})$ is contained in $A^P(t, \frac{r}{16\sqrt{L}}, U^c, k, x)$ for any $t > 0$ and any $k \in \mathbb{N}$. Thus the covering constructed in Theorem 4.1 can be used to estimate the Hausdorff dimension of the intersection of the set $\pi_x^{-1}(E(F^+, U))$ with P from above.

Proof of Theorem 1.3 assuming Theorem 4.1 First note that the statement of Theorem 1.3 involves just the semigroup F^+ as a whole and does not depend on its parametrization. Thus, applying a linear time change to the flow g_t, without loss of generality for the proof of the theorem we can assume that $\lambda_{\max} = 1$.

Let $0 < r < r_0$. We are again going to use the notation $E_{x,s}$ introduced in (3.2). In view of the countable stability of Hausdorff dimension it suffices to find $s > 0$ such that for any $x \in X$,

$$\dim\left(E_{x,s} \cap P\right) \leq \dim X - C' \frac{\mu(\sigma_r U)}{\log \frac{1}{r} + \log \frac{1}{\mu(\sigma_r U)}} \tag{4.6}$$

with the constant $C' > 0$ only dependent on X and F.

Note that $E_{x,r/2} \cap P = \varnothing$ for any $x \notin \partial_r U^c$, so in this case (4.6) is clearly satisfied for $s = r/2$. So, let $x \in \partial_r U^c$ and take $s = \frac{r}{16\sqrt{L}}$.

Let $\underline{\dim}_B$ denote the lower box dimension. Since for any $t > 0$ we have

$$E_{x, \frac{r}{16\sqrt{L}}} \cap P \subset \bigcap_{k \in \mathbb{N}} A^P\left(t, \frac{r}{16\sqrt{L}}, U^c, k, x\right),$$

from Theorem 4.1, in view of the assumption $\lambda_{\max} = 1$, it follows that

$$
\underline{\dim}_B \left(E_{x, \frac{r}{16\sqrt{L}}} \cap P \right) \leq \liminf_{k \to \infty} \frac{\log \left(K_0 e^{Lkt} \left(1 - K_1 \mu(\sigma_r U) + \frac{K_2 e^{-\lambda_1 t}}{r^L} \right)^k \right)}{-\log(re^{-kt})}
$$

$$
= \liminf_{k \to \infty} \frac{\log K_0 + Lkt + k \log \left(1 - K_1 \mu(\sigma_r U) + \frac{K_2 e^{-\lambda_1 t}}{r^L} \right)}{-\log r + kt} \quad (4.7)
$$

$$
= L + \frac{\log \left(1 - K_1 \mu(\sigma_r U) + \frac{K_2 e^{-\lambda_1 t}}{r^L} \right)}{t}
$$

whenever t satisfies (4.5). It remains to choose an optimal t. Take q to be a natural number which satisfies the following conditions:

$$
\left(\tfrac{1}{8} \right)^q < \frac{K_1}{2K_2},
$$
$$
q > \lambda_1 b - L, \quad (4.8)
$$

and set

$$
t = a + \frac{L+q}{\lambda_1} \log \frac{1}{r\mu(\sigma_r U)}.
$$

It is easy to see that in view of (4.8), t as above satisfies (4.5), and we have

$$
\frac{K_2 e^{-\lambda_1 t}}{r^L} = K_2 r^{-L} e^{-\lambda_1 \left(a + \frac{L+q}{\lambda_1} \log \frac{1}{r\mu(\sigma_r U)} \right)}
$$

$$
= e^{-\lambda_1 a} K_2 r^{-L} r^{L+q} \mu(\sigma_r U)^{L+q} = e^{-\lambda_1 a} K_2 \cdot r^q \cdot \mu(\sigma_r U)^{L+q} \quad (4.9)
$$

$$
< e^{-\lambda_1 a} K_2 \cdot \left(\tfrac{1}{8} \right)^q \cdot \mu(\sigma_r U) < e^{-\lambda_1 a} K_2 \frac{K_1}{2K_2} \cdot \mu(\sigma_r U) \leq \frac{K_1}{2} \mu(\sigma_r U).
$$

Combining (4.7) and (4.9), we have:

$$
\dim \left(E_{x, \frac{r}{16\sqrt{L}}} \cap P \right) \leq L + \frac{\log \left(1 - \frac{K_1}{2} \mu(\sigma_r U) \right)}{t} \leq L - \frac{\frac{K_1}{2} \mu(\sigma_r U)}{t}
$$

$$
= L - \frac{\frac{K_1}{2} \cdot \mu(\sigma_r U)}{\frac{(L+q)}{\lambda_1} \cdot \log \frac{1}{r\mu(\sigma_r U)}} = L - C' \cdot \frac{\mu(\sigma_r U)}{\log \frac{1}{r} + \log \frac{1}{\mu(\sigma_r U)}},
$$

where $C' = \frac{K_1 \lambda_1}{2(L+q)}$. This finishes the proof. $\square$

5 A measure estimate

Our goal in this section is to prove the following proposition which gives a lower bound for the measure of sets

$$
A^P \left(t, \frac{r}{16\sqrt{L}}, \sigma_{r/2} U, 1, x \right) = \left\{ h \in B^P \left(\frac{r}{16\sqrt{L}} \right) : g_t h x \in \sigma_{r/2} U \right\} \quad (5.1)
$$

whenever t satisfies (4.5), and x belongs to $\partial_r U^c$.

Proposition 5.1 *Let F^+ be a one-parameter Ad-diagonalizable subsemigroup of G, and P a subgroup of G with property* (EEP). *Then there exist positive constants a, b, E', λ' such that for any $U \subset X$ such that U^c is compact, any $x \in \partial_r U^c$, any $0 < r < r_0$ where r_0 is as in* (4.4), *and any t satisfying* (4.5) *one has*

$$\inf_{x \in \partial_r U^c} \nu \left(A^P \left(t, \frac{r}{16\sqrt{L}}, \sigma_{r/2} U, 1, x \right) \right) \geq \nu \left(B^P \left(\frac{r}{16\sqrt{L}} \right) \right) \mu(\sigma_r U) - E' e^{-\lambda' t}. \quad (5.2)$$

To prove (5.2) we will apply (EEP) to smooth approximations of $1_{B^P \left(\frac{r}{16\sqrt{L}} \right)}$ and $1_{\sigma_{r/2} U}$. In order to extract useful information from (EEP) we will need to bound the norms of the derivatives of those approximations. The next two lemmas will be used to approximate $1_{\sigma_{r/2} U}$ and $1_{B^P \left(\frac{r}{16\sqrt{L}} \right)}$ respectively.

Lemma 5.2 *Let O be a nonempty open subset of X, and let $0 < \varepsilon_0 < 1$, $\delta < 1$ be such that*

$$\delta \mu(O) \leq \mu(\sigma_{\varepsilon_0} O) < \mu(O). \quad (5.3)$$

Then for any $0 < \varepsilon \leq \varepsilon_0$ one can find a nonnegative function $\psi_\varepsilon \in C^\infty_{comp}(X)$ such that:

(1) $\psi_\varepsilon \leq 1_O$;
(2) $\delta \mu(O) \leq \int_X \psi_\varepsilon \, d\mu$;
(3) $\|\psi_\varepsilon\|_{\ell,2} \leq 4^\ell M_\ell \varepsilon^{-\ell}$;
(4) $\|\psi_\varepsilon\|_{C^\ell} \leq 4^\ell M_\ell \varepsilon^{-\ell}$,

where M_ℓ is as in Lemma 2.6.

Proof Let O be a nonempty open subset of X, and let $0 < \varepsilon_0 < 1$ and $\delta < 1$ be such that (5.3) holds. Since O is open and the function $x \mapsto \text{dist}(x, O^c)$ is continuous, for any $0 < \varepsilon < \varepsilon_0$ we have:

$$\delta \mu(O) < \mu(\sigma_\varepsilon O) < \mu(O).$$

By the inner regularity of μ we can find a compact subset $A_\varepsilon \subset \sigma_\varepsilon O$ such that:

$$\delta \mu(O) \leq \mu(A_\varepsilon) \leq \mu(\sigma_\varepsilon O) < \mu(O).$$

Denote by A_ε^+, A_ε^{++} the closed $\frac{\varepsilon}{4}$ and $\frac{\varepsilon}{2}$ neighborhoods of A_ε. Since A_ε is compact, these sets are compact as well. Now take $\psi_\varepsilon = \varphi_{\varepsilon/4} * 1_{A_\varepsilon^+}$, where $\varphi_{\varepsilon/4}$ is as in Lemma 2.6. Sincet $\varphi_{\varepsilon/4}$ is supported on $B^G(\varepsilon/4)$, the support of the function ψ_ε is contained in $A_\varepsilon^{++} \subset O$, so property (1) holds. Furthermore, $\psi_\varepsilon = 1$ on A_ε, therefore:

$$\mu(O) \geq \int_X \psi_\varepsilon \, d\mu \geq \mu(A_\varepsilon) \geq \delta \mu(O),$$

which gives us property (2). Let $\alpha = (\alpha_1, \ldots, \alpha_N)$ be such that $|\alpha| \leq \ell$. For any $x \in X$ we have

$$\left| D^\alpha \psi_\varepsilon(x) \right| = \left| D^\alpha (\varphi_{\varepsilon/4} * 1_{A_\varepsilon^+})(x) \right| = \left| D^\alpha \varphi_{\varepsilon/4} * 1_{A_\varepsilon^+}(x) \right|$$
$$\leq \left\| D^\alpha \varphi_{\varepsilon/4} \right\|_1 \leq \left\| \varphi_{\varepsilon/4} \right\|_{\ell,1} \leq M_\ell (\tfrac{\varepsilon}{4})^{-\ell},$$

and likewise, by Young's inequality,

$$\left\| D^\alpha \psi_\varepsilon \right\|_2 \leq \| D^\alpha \varphi_{\varepsilon/4} * 1_{A_\varepsilon^+} \|_2 \leq \left\| D^\alpha \varphi_{\varepsilon/4} \right\|_1 \cdot \left\| 1_{A_\varepsilon^+} \right\|_2 \leq \left\| D^\alpha \varphi_{\varepsilon/4} \right\|_1 \leq M_\ell (\tfrac{\varepsilon}{4})^{-\ell},$$

which implies (3) and (4). $\qquad \square$

Similarly to the proof of the above lemma, one can get the smooth estimations for characteristic functions of small balls in P (we omit the proof for brevity):

Lemma 5.3 *For any $\ell \in \mathbb{Z}_+$ there exist constants $M'_\ell > 0$ (depending only on P) such that the following holds: for any $\varepsilon, r > 0$ there exist functions $f_\varepsilon : P \to [0, 1]$ such that*

(1) $f_\varepsilon = 1$ on $B^P(r)$;
(2) $f_\varepsilon = 0$ on $\left(B^P(r + \varepsilon)\right)^c$;
(3) $\|f_\varepsilon\|_{\ell,2} \le M'_\ell \varepsilon^{-\ell}$;
(4) $\|f_\varepsilon\|_{C^\ell} \le M'_\ell \varepsilon^{-\ell}$.

Proof of Proposition 5.1 Let ℓ and λ be as in Definition 2.2, and let a, b, E_1 be the implicit constants in (2.2) and (2.3) such that $t > a + b \log \frac{1}{r_0(x)}$ implies

$$\left| I_{f,\psi}(g_t, x) - \int_P f \, d\nu \int_X \psi \, d\mu \right| \le E_1 \max(\|\psi\|_{C^1}, \|\psi\|_{\ell,2}) \cdot \|f\|_{C^\ell} \cdot e^{-\lambda t} \quad (5.4)$$

for any f and ψ as in Definition 2.2. Then choose $\lambda' > 0$ such that

$$\lambda - 2\ell\lambda' > \lambda' \quad \text{and} \quad 1/\lambda' > b. \quad (5.5)$$

Now let $U \subset X$ be such that U^c is compact, and take $0 < r < r_0$ and $x \in \partial_r U^c$. If $\mu(\sigma_r U) = 0$, (5.2) is trivially satisfied; thus let us assume that $\mu(\sigma_r U) > 0$. Then put

$$O := \sigma_{r/2} U$$

and take

$$\delta := \frac{\mu(\sigma_r U)}{\mu(\sigma_{r/2} U)}.$$

Note that (5.3) holds with $\varepsilon_0 = r/2$. Also, since U is open, the function $x \mapsto \text{dist}(x, U^c)$ is continuous, which implies that $\delta < 1$.

Now set $f = 1_{B^P\left(\frac{r}{16\sqrt{L}}\right)}$ and take

$$t \ge a + \frac{1}{\lambda'} \log \frac{2}{r} > a + b \log \frac{2}{r} > a + b \log \frac{1}{r_0(x)} \quad (5.6)$$

(the last inequality holds since $x \in \partial_r U^c$). Also define

$$\varepsilon := e^{-\lambda' t}.$$

Note that $\varepsilon < r/2$ in view of (5.6). So let us apply Lemma 5.2 with $\varepsilon_0 = r/2$, and Lemma 5.3 with $\frac{r}{16\sqrt{L}}$ in place of r. Let ψ_ε and f_ε be the corresponding functions. Then we have

$$\begin{aligned}
\max(\|\psi_\varepsilon\|_{C^1}, \|\psi_\varepsilon\|_{\ell,2}) \cdot \|f_\varepsilon\|_{C^\ell} \cdot e^{-\lambda t} &\le \max(\|\psi_\varepsilon\|_{C^\ell}, \|\psi_\varepsilon\|_{\ell,2}) \cdot \|f_\varepsilon\|_{C^\ell} \cdot e^{-\lambda t} \\
&\le 4^\ell M_\ell \varepsilon^{-\ell} M'_\ell \varepsilon^{-\ell} e^{-\lambda t} \quad (5.7) \\
&= 4^\ell M_\ell M'_\ell e^{2\ell\lambda' - \lambda t} \le 4^\ell M_\ell M'_\ell e^{-\lambda' t}.
\end{aligned}$$

Note also that supp $f_\varepsilon \subset B^P\left(\frac{r}{16\sqrt{L}} + r/2\right) \subset B^P(1)$. In view of (5.6) and (5.7), the estimate (5.4) can be applied to $\psi_\varepsilon, f_\varepsilon, x$ and t, and yields

$$\int_P f_\varepsilon(h)\psi_\varepsilon(g_t h x) \, d\nu(h) \ge \int_P f_\varepsilon \, d\nu \int_X \psi_\varepsilon \, d\mu - 4^\ell M_\ell M'_\ell E_1 e^{-\lambda' t}.$$

In view of (5.1) we have:

$$
v\left(A^P\left(t, \frac{r}{16\sqrt{L}}, \sigma_{r/2}U, 1, x\right)\right) = \int_P f(h) 1_{\sigma_{r/2}U}(g_t hx)\, dv(h)
$$

$$
\geq \int_P f(h)\psi_\varepsilon(g_t hx)\,dv(h) \geq \int_P f_\varepsilon(h)\psi_\varepsilon(g_t hx)\, dv(h) - \int_P |f_\varepsilon - f|\, dv
$$

$$
\geq \int_P f_\varepsilon(h)\psi_\varepsilon(g_t hx)\, dv(h) - v\left(B^P\left(\frac{r}{16\sqrt{L}} + e^{-\lambda' t}\right) \smallsetminus B^P\left(\frac{r}{16\sqrt{L}}\right)\right).
$$

By the mean-value theorem and (4.2), for some $\frac{r}{16\sqrt{L}} < s < \frac{r}{16\sqrt{L}} + e^{-\lambda' t}$ it holds that

$$
v\left(B^P\left(\frac{r}{16\sqrt{L}} + e^{-\lambda' t}\right) \smallsetminus B^P\left(\frac{r}{16\sqrt{L}}\right)\right) = v\left(B\left(\frac{r}{16\sqrt{L}} + e^{-\lambda' t}\right)\right) - v\left(B\left(\frac{r}{16\sqrt{L}}\right)\right)
$$

$$
\leq c_3 e^{-\lambda' t} s^{L-1} \leq c_3 e^{-\lambda' t}\left(\frac{r}{16\sqrt{L}} + \frac{r}{2}\right)^{L-1} \leq c_3 e^{-\lambda' t}.
$$

Combining the above computations, we obtain

$$
v\left(A^P\left(t, \frac{r}{16\sqrt{L}}, \sigma_{r/2}U, 1, x\right)\right) \geq \int_P f_\varepsilon(h)\psi_\varepsilon(g_t hx)\, dv(h) - c_3 e^{-\lambda' t}
$$

$$
\geq \int_P f_\varepsilon\, dv \int_X \psi_\varepsilon\, d\mu - 4^\ell M_\ell M'_\ell E_1 e^{-\lambda' t} - c_3 e^{-\lambda' t}
$$

$$
\geq v\left(B^P\left(\frac{r}{16\sqrt{L}}\right)\right)\frac{\mu(\sigma_r U)}{\mu(\sigma_{r/2}U)} \cdot \mu(\sigma_{r/2}U) - (4^\ell M_\ell M'_\ell E_1 + c_3)e^{-\lambda' t}
$$

$$
= v\left(B^P\left(\frac{r}{16\sqrt{L}}\right)\right)\mu(\sigma_r U) - E' e^{-\lambda' t}
$$

where $E' := 4^\ell M_\ell M'_\ell E_1 + c_3$. $\qquad\square$

6 Tessellations of P and Bowen boxes: proof of Theorem 4.1

In order to prove Theorem 4.1 it will be instrumental to use a technique of *tessellations* of nilpotent Lie groups, as developed in [14]. It allows one to cover subsets of P with objects that behave like non-overlapping cubes in a Euclidean space. In this aspect our method differs from the one by Kadyrov [10]: using *Bowen boxes* defined below, as opposed to Bowen balls considered in [10], turns out to be a more efficient way to cover P [see (6.8) below and the subsequent footnote for explanation]. We are going to revisit the construction in [14] and then use it to find efficient coverings of sets of the form $A^P(t, \frac{r}{16\sqrt{L}}, U^c, k, x)$.

Let us say that an open subset V of P is a *tessellation domain* for P relative to a countable subset Λ of P if

- $v(\partial V) = 0$.
- $V\gamma_1 \cap V\gamma_2 = \varnothing$ for different $\gamma_1, \gamma_2 \in \Lambda$.
- $P = \bigcup_{\gamma \in \Lambda} \overline{V}\gamma$.

Note that P is a connected simply connected nilpotent Lie group. Let $I_P \subset \mathfrak{p} = \mathrm{Lie}(P)$ be the cube centered at 0 with side length 1 with respect to a suitably chosen basis of $\mathfrak{p}$. For any $r > 0$ let us define $V_r := \exp(\frac{r}{4\sqrt{L}} I_P)$. Then, as shown in [14, Proposition 3.3], V_r is a tessellation domain for P relative to some discrete subset Λ_r of P. Since the exponential

map is 2-bi-Lipschitz on $\frac{r}{4\sqrt{L}}I_P$ for $r < r''$, we have

$$B^P\left(\frac{r}{16\sqrt{L}}\right) \subset V_r \subset B^P(r/4) \tag{6.1}$$

Also it is easy to see that there exists $K_3 > 0$ such that for any $\delta \leq 1$

$$\nu\left(\{h \in P : \text{dist}(h, \partial V_r) < \delta\}\right) < K_3\delta. \tag{6.2}$$

Define

$$\lambda_0 := \min\{|\lambda| : \lambda \text{ is an eigenvalue of } \text{ad}_{g_1}|_{\mathfrak{p}}\}. \tag{6.3}$$

Again using the bi-Lipschitz property of exp, we can conclude that for any $0 < r < r''$ and any $t > 0$ one has

$$\text{diam}(g_{-t}V_r g_t) < 2re^{-\lambda_0 t}. \tag{6.4}$$

Let us now define a *Bowen (t, r)-box* in P to be a set of the form $g_{-t}V_r\gamma g_t$ for some $\gamma \in P$ and $t > 0$. Also define

$$S_{r,t} := \{\gamma \in \Lambda_r : g_{-t}V_r\gamma g_t \cap V_r \neq \varnothing\}.$$

Note that V_r can be covered with at most $\#S_{r,t}$ Bowen (t, r)-boxes in P. The following lemma gives an upper bound for $\#S_{r,t}$:

Lemma 6.1 *For any $0 < r < r''$ and any $t > 0$*

$$\#S_{r,t} \leq \frac{\nu(V_r)}{\nu(g_{-t}V_r g_t)}\left(1 + \frac{K_3 e^{-\lambda_0 t}}{\nu(V_r)}\right).$$

Proof Let $0 < r < r''$ and $t > 0$. One has:

$$\#S_{r,t} = \#\{\gamma \in \Lambda_r : g_{-t}V_r\gamma g_t \subset V_r\} + \#\{\gamma \in \Lambda_r : g_{-t}V_r\gamma g_t \cap \partial V_r \neq \varnothing\}.$$

Since V_r is a tessellation domain of P relative to Λ_r, the first term in the above sum is not greater than $\frac{\nu(V_r)}{\nu(g_{-t}V_r g_t)}$, while in view of (6.2) and (6.4), the second term is not greater than:

$$\frac{\nu(\{p \in P : \text{dist}(p, \partial V_r) < \text{diam}(g_{-t}V_r g_t)\})}{\nu(g_{-t}V_r g_t)} < \frac{2rK_3 e^{-\lambda_0 t}}{\nu(g_{-t}V_r g_t)} < \frac{K_3 e^{-\lambda_0 t}}{\nu(g_{-t}V_r g_t)}.$$

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

Now let U be an arbitrary subset of X. The next lemma can be used to turn the measure estimate from Sect. 5 into a covering result.

Lemma 6.2 *For any $x \in X$, any $U \subset X$, any $0 < r < r''$ and any $t > 0$ we have*

$$A^P\left(t, \frac{r}{16\sqrt{L}}, \sigma_{r/2}U, 1, x\right) \subset \bigcup_{\substack{\gamma \in S_{r,t} \\ V_r\gamma g_t x \subset U}} g_{-t}V_r\gamma g_t.$$

Proof For any $\gamma \in P$ and any $p_1, p_2 \in V_r$ we have:

$$\text{dist}(p_1 \gamma g_t x, p_2 \gamma g_t x) \leq \text{dist}(p_1, p_2) \leq \text{diam}(V_r) < r/2. \tag{6.5}$$

Hence, if

$$A^P\left(t, \frac{r}{16\sqrt{L}}, \sigma_{r/2} U, 1, x\right) \cap g_{-t} V_r \gamma g_t \neq \varnothing$$

for $\gamma \in \Lambda_r$, then for some $p \in B^P(\frac{r}{16\sqrt{L}}) \subset V_r$ one has $g_t p x \in \sigma_{r/2} U \cap V_r \gamma g_t x$, and in view of (6.5) and $\partial_{r/2} \sigma_{r/2} U \subset U$, we can conclude that $V_r \gamma g_t x \subset U$. $\qquad \square$

The next corollary follows immediately from Lemma 6.2:

Corollary 6.3 *For any $x \in X$, $U \subset X$, $0 < r < r''$ and $t > 0$ we have*

$$\#\{\gamma \in S_{r,t} : V_r \gamma g_t x \subset U\} \geq \frac{\nu\left(A^P(t, \frac{r}{16\sqrt{L}}, \sigma_{r/2} U, 1, x)\right)}{\nu(g_{-t} V_r g_t)}.$$

For the proof of Theorem 4.1 we will also need to cover Bowen boxes by small balls. The next lemma provides a bound for the number of balls of radius $re^{-\lambda_{\max} t}$ needed to cover a Bowen (t, r)-box.

Lemma 6.4 *There exists $K_4 > 0$ such that for any $0 < r < r''$ and any $t > 0$, any Bowen (t, r)-box in P can be covered with at most $K_4 \frac{\nu(g_{-t} V_r g_t)}{\nu(B^P(re^{-\lambda_{\max} t}))}$ balls in P of radius $re^{-\lambda_{\max} t}$.*

Proof Let $B = g_{-t} V_r \gamma g_t$ be a Bowen (t, r)-box. In view of the Besicovitch covering property of P, any covering of B by balls in P of radius $re^{-\lambda_{\max} t}$ has a subcovering of index uniformly bounded from above by a fixed constant (the Besicovitch constant of P). The union of those balls is contained in the $re^{-\lambda_{\max} t}$-neighborhood of B. But since B is a translate of the exponential image of a box in $\mathfrak{p}$ whose smallest sidelength is $re^{-\lambda_{\max} t}$, it follows that the measure of the $re^{-\lambda_{\max} t}$-neighborhood of B is bounded by a uniform constant times $\nu(B)$, and the lemma follows. $\qquad \square$

We are now ready to begin the

Proof of Theorem 4.1 Take a, b, E', λ' be as in Proposition 5.1, K_3 as in (6.2), K_4 as in Lemma 6.4 and λ_0 as in (6.3). Fix $U \subset X$ such that U^c is compact, and take $0 < r < r_0$, $x \in \partial_r U^c$, and $t > a + b \log \frac{1}{r}$. Define for any $k \in \mathbb{N}$

$$E_{V_r}(t, k, x) := \left\{p \in V_r : g_{\ell t} p x \notin U \ \forall \ell \in \{1, 2, \cdots, k\}\right\}.$$

Recall that our goal is to construct a covering of the set $A^P(t, \frac{r}{16\sqrt{L}}, U^c, k, x)$ for any $k \in \mathbb{N}$, which is a subset of $E_{V_r}(t, k, x)$ in view of (6.1). Note that for $\gamma \in P$, the Bowen (t, r)-box $g_{-t} V_r \gamma g_t$ does not intersect $E_{V_r}(t, 1, x)$ if and only if $V_r \gamma g_t x \subset U$. Combining Lemma 6.1 with Corollary 6.3 and then with Proposition 5.1, we conclude that $E_{V_r}(t, 1, x)$ can be covered with at most

$$\#S_{r,t} - \#\{\gamma \in S_{r,t} : V_r \gamma g_t x \subset U\}$$

$$\leq \frac{\nu(V_r)}{\nu(g_{-t} V_r g_t)}\left(1 + \frac{K_3 e^{-\lambda_0 t}}{\nu(V_r)}\right) - \frac{\nu\left(A^P(t, \frac{r}{16\sqrt{L}}, \sigma_{r/2} U, 1, x)\right)}{\nu(g_{-t} V_r g_t)}$$

$$\leq \frac{\nu(V_r)}{\nu(g_{-t} V_r g_t)} \cdot \left(1 + \frac{K_3 e^{-\lambda_0 t} - \nu\left(B^P(\frac{r}{16\sqrt{L}})\right)\mu(\sigma_r U) + E' e^{-\lambda' t}}{\nu(V_r)}\right) \tag{6.6}$$

$$=: N(r, t)$$

Bowen (t, r)-boxes in P.

Now let $g_{-t} V_r \gamma g_t$ be one of the Bowen (t, r)-boxes in the above cover which has non-empty intersection with $E_{V_r}(t, 1, x)$. Take any $q = g_{-t} h \gamma g_t \in g_{-t} V_r \gamma g_t$; then $g_t q x = h \gamma g_t x$, hence $\{g_t q x : q \in g_{-t} V_r \gamma g_t\} = \{h \gamma g_t x : h \in V_r\}$. Consequently,

$$\{q \in g_{-t} V_r \gamma g_t : g_{2t} q x \notin U\} = g_{-t} E_{V_r}(t, 1, x) \gamma g_t. \tag{6.7}$$

Note that since $\operatorname{diam}(V_r) < r$ and $g_{-t} V_r \gamma g_t \cap E_{V_r}(t, 1, x) \neq \varnothing$, we have $\gamma g_t x \in \partial_r U^c$. Hence, by going through the same procedure, this time using $\gamma g_t x$ in place of x, we can cover the set in the left hand side of (6.7) with at most $N(r, t)$ Bowen $(2t, r)$-boxes in P. Therefore, we conclude that the set $E_{V_r}(t, 2, x)$ can be covered with at most $N(r, t)^2$ Bowen $(2t, r)$-boxes in P. By doing this procedure inductively, we can see that for any $k \in \mathbb{N}$, the set $E_{V_r}(t, k, x)$ can be covered with at most $N(r, t)^k$ Bowen (tk, r)-boxes in P. Thus, in view of Lemma 6.4, the set $E_{V_r}(t, k, x)$ can be covered with at most

$$K_4 \frac{\nu(g_{-tk} V_r g_{tk})}{\nu\big(B^P(re^{-k\lambda_{\max} t})\big)} N(r, t)^k$$

balls of radius $re^{-k\lambda_{\max} k}$ in P.

Now observe that for any $r > 0$ and any $k \in \mathbb{N}$ one has

$$\left(\frac{\nu(V_r)}{\nu(g_{-t} V_r g_t)}\right)^k = \frac{\nu(V_r)}{\nu(g_{-kt} V_r g_{tk})}. \tag{6.8}$$

Here it is crucially important[3] that the translates of V_r form a tessellation of P. Using (4.1) and (6.8) we get

$$\frac{\nu(g_{-tk} V_r g_{tk})}{\nu\big(B^P(re^{-k\lambda_{\max} t})\big)} \left(\frac{\nu(V_r)}{\nu(g_{-t} V_r g_t)}\right)^k = \frac{\nu(V_r)}{\nu\big(B^P(re^{-k\lambda_{\max} t})\big)}$$

$$\leq \frac{c_2(r/4)^L}{c_1 r^L e^{-Lk\lambda_{\max} t}} = \frac{c_2}{4^L c_1} e^{Lk\lambda_{\max} t},$$

which, in view of (6.1), (4.1) and the definition (6.6) of $N(r, t)$, implies that

$$A^P\big(t, \frac{r}{16\sqrt{L}}, U^c, k, x\big) \subset E_{V_r}(t, k, x)$$

can be covered with at most

$$\frac{K_4 c_2}{4^L c_1} e^{Lk\lambda_{\max} t} \cdot \left(1 + \frac{K_3 e^{-\lambda_0 t} - \nu\big(B^P\big(\frac{r}{16\sqrt{L}}\big)\big) \mu(\sigma_r U) + E' e^{-\lambda' t}}{\nu(V_r)}\right)^k$$

$$\leq \frac{K_4 c_2}{4^L c_1} e^{Lk\lambda_{\max} t} \left(1 + \frac{K_3 (16\sqrt{L})^L e^{-\lambda_0 t}}{c_1 r^L} - \frac{c_1}{c_2(4\sqrt{L})^L} \mu(\sigma_r U) + \frac{4^L E' e^{-\lambda' t}}{c_2 r^L}\right)^k$$

$$\leq K_0 e^{Lk\lambda_{\max} t} \left(1 - K_1 \mu(\sigma_r U) + \frac{K_2 e^{-\lambda_1 t}}{r^L}\right)^k$$

[3] We note that a similar step in the proof of [10, Theorem 3.1] uses balls instead of boxes, and the boundary effects make it difficult to justify the corresponding equality.

balls in P of radius $re^{-k\lambda_{\max}t}$, where

$$K_0 = \frac{K_4 c_2}{4^L c_1}, \quad K_1 = \frac{c_1}{c_2(4\sqrt{L})^L}, \quad K_2 = \frac{K_3(16\sqrt{L})^L}{c_1} + \frac{4^L E'}{c_2},$$

and $\lambda_1 = \min(\lambda_0, \lambda')$. □

7 (EEP) for the group P as in (1.17)

In the last two sections of the paper we prove Theorem 1.4. Namely we fix two positive integers m, n, take $X = G/\Gamma$ as in (1.3) and consider $F = \{g_t\} = g_t^{\mathbf{i},\mathbf{j}}$ as in (1.16), where $\mathbf{i}$ and $\mathbf{j}$ are as in (1.13). We also define

$$\alpha = \min\{i_1, \ldots, i_m, j_1, \ldots, j_n\}. \tag{7.1}$$

Let us denote $m + n$ by d. In what follows, constants $C_1, C_2, \ldots$ will only depend on m and n.

Our goal in this section is to prove that P as in (1.17) satisfies (EEP) with respect to the F^+-action on X. Note that, unless $\mathbf{i} = \mathbf{m}$ and $\mathbf{j} = \mathbf{n}$, P is a proper subgroup of the expanding horospherical subgroup relative to g_1, hence Theorem 2.5 is not applicable. In [16], the proof of effective equidistribution of g_t-translates of orbits of P used the observation that P is an expanding horospherical subgroup relative to another element of G. We are going to work out an explicit estimate for the constant in [16, Theorem 1.3]; namely, establish

Theorem 7.1 *Let P be as in (1.17), $F = \{g_t\}$ as in (1.16), and X as in (1.3). Then P satisfies* (EEP) *relative to the F^+-action on X.*

Recall that X can be identified with the space of unimodular lattices in $\mathbb{R}^d$ via $g\Gamma \mapsto g\mathbb{Z}^d$. It will be useful to relate the injectivity radius $r_0(x)$ of an element $x = g\Gamma \in X$ with the function

$$\delta(g\Gamma) := \inf_{\mathbf{v} \in \mathbb{Z}^d \smallsetminus \{0\}} \|g\mathbf{v}\|. \tag{7.2}$$

Here $\|\cdot\|$ stands for some norm on $\mathbb{R}^d$; the implicit constants in the statements below will depend on the choice of the norm.

Lemma 7.2 *There exist $C_1, C_2 > 0$ such that for any $x \in X$ one has*

$$C_1\delta(x)^d \leq r_0(x) \leq C_2\delta(x)^{\frac{d}{d-1}}.$$

Proof The lower estimate can be found in [16, Proposition 3.5] or [2, Lemma 3.6]. To prove the upper estimate, take $\|\cdot\|$ to be the Euclidean norm, suppose $\delta(x) = \varepsilon$, and let $\lambda_1, \ldots, \lambda_d$ be the successive minima of the lattice x. Let $\mathbf{v}_1, \ldots, \mathbf{v}_d$ be vectors realizing the first and the last minimum of x respectively, and take g to be an element of G which fixes $\mathbf{v}_1, \ldots, \mathbf{v}_{d-1}$ and sends $\mathbf{v}_d$ to $\mathbf{v}_d + \mathbf{v}_1$. Then $gx = x$, and, since $\|\mathbf{v}_1\| = \varepsilon$ and $\|\mathbf{v}_d\| \geq \varepsilon^{-\frac{1}{d-1}}$, it follows that

$$\text{dist}(g, e) \ll \|g - I\|_{op} \ll \varepsilon^{1+\frac{1}{d-1}} = \delta(x)^{\frac{d}{d-1}},$$

(here and hereafter $\|\cdot\|_{op}$ refers to the operator norm as a linear transformation of $\mathbb{R}^d$), finishing the proof. □

The next ingredient of the proof is quantitative nondivergence of translates of P-orbits. Let us denote by $\mathfrak{a}^+$ the set of d-tuples $\mathbf{t} = (t_1, \ldots, t_d) \in \mathbb{R}^d$ such that

$$t_1, \ldots, t_d > 0 \quad \text{and} \quad \sum_{i=1}^{m} t_i = \sum_{j=1}^{n} t_{m+j},$$

and for $\mathbf{t} \in \mathfrak{a}_+$ define

$$g_{\mathbf{t}} := \operatorname{diag}(e^{t_1}, \ldots, e^{t_m}, e^{-t_{m+1}}, \ldots, e^{-t_d}) \in G$$

and

$$\lfloor \mathbf{t} \rfloor := \min_{i=1,\ldots,d} t_i.$$

The following statement about quantitative non-divergence of $g_{\mathbf{t}}$-translates of P orbits in X was proved in [16, Corollary 3.4]: for any compact $L \subset X$ and any ball $B \subset P$ centered at e there exist constants $T = T(B, L)$ and $C = C(B, L)$ such that for every $0 < \varepsilon < 1$, any $x \in L$ and any $\mathbf{t} \in \mathfrak{a}^+$ with $\lfloor \mathbf{t} \rfloor \geq T$ one has

$$\nu\left(\left\{h \in B : \delta(g_{\mathbf{t}} hx) < \varepsilon\right\}\right) \leq C\varepsilon^{\frac{1}{mn(d-1)}} \nu(B).$$

For our purposes we need an effective version:

Proposition 7.3 *There exist constants C_3, C_4, C_5 such that for every $0 < \varepsilon < 1$, any $x \in X$ and any $\mathbf{t} \in \mathfrak{a}^+$ with $\lfloor \mathbf{t} \rfloor \geq C_3 + C_4 \log \frac{1}{r_0(x)}$ it holds that*

$$\nu\left(\{h \in B^P(2) : \delta(g_{\mathbf{t}} hx) < \varepsilon\}\right) \leq C_5\varepsilon^{\frac{1}{mn(d-1)}}. \tag{7.3}$$

Proof According to [16, Theorem 3.1], which is a special case of general quantitative non-divergence result [3, Theorem 6.2], there exists an explicit constant $C_6 > 0$, depending only on m and n, such that for every ball $B \subset P$, any $x = g\mathbb{Z}^d \in X$, any $\mathbf{t} \in \mathfrak{a}^+$ and any $0 < \varepsilon < 1$ not greater than

$$c := \inf_{\substack{\mathbf{w} \in \bigwedge^k(\mathbb{Z}^d) \smallsetminus \{0\} \\ k=1,\ldots,d-1}} \sup_{h \in B} \|g_{\mathbf{t}} hg\mathbf{w}\|, \tag{7.4}$$

it holds that

$$\nu\left(\{h \in B : \delta(g_{\mathbf{t}} hx) < \varepsilon\}\right) \leq C_6(\varepsilon/c)^{\frac{1}{mn(d-1)}} \nu(B).$$

On the other hand, [16, Lemma 3.2] asserts the existence of $C_7 > 0$ and, for each ball $B \subset P$, a constant C_B such that for any $\mathbf{t} \in \mathfrak{a}^+$ and any $\mathbf{w} \in \bigwedge^k(\mathbb{R}^d)$, $k = 1, \ldots, d-1$, one has

$$\sup_{h \in B} \|g_{\mathbf{t}} h\mathbf{w}\| \geq C_B e^{C_7 \lfloor \mathbf{t} \rfloor} \|\mathbf{w}\|.$$

Also, by Minkowski's Lemma there exists $C_8 > 0$ such that

$$\inf_{\mathbf{w} \in \bigwedge^k(\mathbb{Z}^d) \smallsetminus \{0\}} \|g\mathbf{w}\| \geq C_8 \delta(x)^k.$$

Therefore c as in (7.4) is not less than

$$C_B e^{C_7 \lfloor \mathbf{t} \rfloor} C_8 \delta(x)^{d-1} \geq C_8 C_B e^{C_7 \lfloor \mathbf{t} \rfloor} \left(\frac{r_0(x)}{C_2}\right)^{\frac{(d-1)^2}{d}} \tag{7.5}$$

(the last inequality holds in view of Lemma 7.2). Now take $B = B^P(2)$ and choose $\mathbf{t}$ so that the right hand side of (7.5) is not less than 1; equivalently, such that

$$\lfloor \mathbf{t} \rfloor \geq \frac{1}{C_7} \log \frac{C_2^{\frac{(d-1)^2}{d}}}{C_8 C_{B^P(2)}} + \frac{(d-1)^2}{dC_7} \log \frac{1}{r_0(x)}.$$

Then (7.3) will hold for any $0 < \varepsilon < 1$ with $C_5 = C_6 \cdot \nu\big(B^P(2)\big)$. $\qquad\square$

Proof of Theorem 7.1 Write $g_t = a_t b_t$, where

$$a_t = \mathrm{diag}(e^{(i_1 - \frac{\alpha}{2m})t}, \ldots, e^{(i_m - \frac{\alpha}{2m})t}, e^{(-j_1 + \frac{\alpha}{2n})t}, \ldots, e^{(-j_n + \frac{\alpha}{2n})t})$$

and

$$b_t = \mathrm{diag}(e^{\alpha t/2m}, \ldots, e^{\alpha t/2m}, e^{-\alpha t/2n}, \ldots, e^{-\alpha t/2n}),$$

where α is as in (7.1). Suppose we are given $f \in C^\infty_{comp}(P)$ with $\mathrm{supp}\, f \subset B^P(1)$, $\psi \in C^\infty_2(X)$ with $\int_X \psi \, d\mu = 0$, and $x \in X$. Put $r = e^{-\frac{\beta\alpha}{2}t}$, where β is to be specified later, and, again using [16, Lemma 2.2(a)], take a non-negative function θ supported on $B^P(r)$ such that (2.7) holds. Since ν is translation-invariant, one can write

$$\begin{aligned} I_{f,\psi}(g_t, x) &= \int_P f(h)\psi(g_t hx)\, d\nu(h) \int_P \theta(y)\, d\nu(y) \\ &= \int_P \int_P f(a_{-t}ya_t h)\theta(y)\psi(a_t b_t a_{-t} ya_t hx)\, d\nu(y)\, d\nu(h) \\ &= \int_P \int_P f(a_{-t}ya_t h)\theta(y)\psi(b_t ya_t hx)\, d\nu(y)\, d\nu(h). \end{aligned}$$

Note that

$$\min\left(i_1 - \frac{\alpha}{2m}, \ldots, i_m - \frac{\alpha}{2m}, j_1 - \frac{\alpha}{2n}, \ldots, j_n - \frac{\alpha}{2n}\right) \geq \alpha/2, \tag{7.6}$$

therefore

$$\mathrm{dist}\left(e, a_{-t}ha_t\right) \leq e^{-\alpha t} \, \mathrm{dist}(e, h)$$

for any $h \in P$. Also, as long as $\theta(y) \neq 0$, the supports of all functions of the form $h \mapsto f(a_{-t}ya_t h)$ are contained in

$$B^P\left(1 + e^{-(\alpha + \frac{\beta\alpha}{2})t}\right) \subset \tilde{B} := B^P(2).$$

Define

$$\varepsilon := \left(\frac{2}{C_1} e^{-\frac{\beta\alpha}{2}t}\right)^{1/d}, \tag{7.7}$$

where C_1 is as in Lemma 7.2, and let

$$A(x, t) := \left\{h \in \tilde{B} \mid \delta(a_t hx) < \varepsilon\right\}.$$

So, in view of (7.6) and Proposition 7.3, for any $x \in X$ and any

$$t \geq \frac{2}{\alpha}\left(C_3 + C_4 \log \frac{1}{r_0(x)}\right) \tag{7.8}$$

one has

$$\nu\big(A(x,t)\big) \le C_5 \varepsilon^{\frac{1}{mn(d-1)}}.$$

Hence, assuming (7.8), the absolute value of

$$\int_{A(x,t)} \int_P f\big(a_{-t} y a_t h\big)\theta(y)\psi\big(b_t y a_t h x\big)\, d\nu(y)\, d\nu(h)$$

is

$$\ll \varepsilon^{\frac{1}{mn(d-1)}} \nu(\tilde{B}) \sup |f| \sup |\psi| \int_P \theta\, d\nu \ll \sup |f| \sup |\psi| \cdot e^{-\frac{\beta\alpha}{2mnd(d-1)}t}.$$

Next, let us assume that $h \in \tilde{B} \setminus A(x,t)$. We are going to apply Theorem 2.4 with b_t in place of g_t, $r = e^{-\frac{\beta\alpha}{2}t}$, $a_t h x$ in place of x and

$$f_h(y) := f\big(a_{-t} y a_t h\big)\theta(y)$$

in place of f. It is clear that supp $f_h \subset B^P(r)$ for any h, i.e. condition (i) of Theorem 2.4 is satisfied. Since $\delta(a_t h x) < \varepsilon$ whenever $h \notin A(x,t)$, condition (ii) is satisfied in view of Lemma 7.2 and (7.7). So we only need to require that $e^{-\frac{\beta\alpha}{2}t}$ is less than $\rho/2$. Also, in view of [16, Lemma 2.2(b)] and (2.7), for any $\ell \in \mathbb{Z}_+$ we have

$$\|f_h\|_{\ell,2} \ll \|f\|_{C^\ell}\|\theta\|_{\ell,2} \ll e^{(\ell+\frac{mn}{2})\frac{\beta\alpha}{2}t}\|f\|_{C^\ell}.$$

This way, by using Theorem 2.4 we get, for some $\gamma > 0$ and $\ell \in \mathbb{Z}_+$,

$$\left|\int_{\tilde{B}\setminus A(x,t)} \int_P f\big(a_{-t} y a_t h\big)\theta(y)\psi\big(b_t y a_t h x\big)\, d\nu(y)\, d\nu(h)\right|$$

$$\le \int_{\tilde{B}\setminus A(x,t)} \left|I_{f_h,\psi}(b_t, a_t h x)\right| d\nu(h)$$

$$\ll \max\big(\|\psi\|_{C^1}, \|\psi\|_{\ell,2}\big) \left(e^{-\frac{\beta\alpha}{2}t}\|f_h\|_1 + e^{\left(\ell+\frac{d^2-1-mn}{2}\right)\frac{\beta\alpha}{2}t}\|f_h\|_{\ell,2}\, e^{-\gamma t} \nu(\tilde{B})\right)$$

$$\ll \max\big(\|\psi\|_{C^1}, \|\psi\|_{\ell,2}\big) \left(\sup |f| \cdot e^{-\frac{\beta\alpha}{2}t} + \|f\|_{C^\ell} \cdot e^{(2\ell+\frac{d^2-1}{2})\frac{\beta\alpha}{2}t-\gamma t}\right).$$

By combining the two estimates above, we get that, as long as $t \gg_+ \log\frac{1}{r_0(x)}$,

$$\left|I_{f,\psi}(g_t, x)\right| \ll \sup |f| \sup |\psi| e^{-\frac{\beta\alpha}{2mnd(d-1)}t}$$

$$+ \max\big(\|\psi\|_{C^1}, \|\psi\|_{\ell,2}\big) \left(\sup |f| \cdot e^{-\frac{\beta\alpha}{2}t} + \|f\|_{C^\ell} \cdot e^{(2\ell+\frac{d^2-1}{2})\frac{\beta\alpha}{2}t-\gamma t}\right)$$

$$\ll \max\big(\|\psi\|_{C^1}, \|\psi\|_{\ell,2}\big) \|f\|_{C^\ell} \cdot \max\left(e^{-\frac{\beta\alpha}{2mnd(d-1)}t}, e^{-\left(\gamma-(2\ell+\frac{d^2-1}{2})\frac{\beta\alpha}{2}\right)t}\right).$$

Choosing β equalizing the two exponents above, that is

$$\beta = \frac{2\gamma/\alpha}{\frac{1}{mnd(d-1)} + 2\ell + \frac{d^2-1}{2}},$$

will satisfy (2.3) with

$$\lambda = \frac{\beta\alpha}{2mnd(d-1)} = \frac{\gamma}{1 + mnd(d-1)(2\ell + \frac{d^2-1}{2})},$$

which finishes the proof.   □

8 Weighted badly approximable matrices

Now let us recall a connection between Diophantine approximation with weights and the action of $F = \{g_t\}$ as in (1.16) on the space X. It is shown in [12, Theorem 6.2] that $A \in M_{m,n}(\mathbb{R})$ is $(\mathbf{i}, \mathbf{j})$-badly approximable iff the orbit $\{g_t u_A \mathbb{Z}^k : t > 0\}$ is bounded in X, where $u_A = \begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix}$. We want to make this equivalence quantitative. Recall that for $\mathbf{p} = (p_1, \ldots, p_m)$ and $\mathbf{q} = (q_1, \ldots, q_n)$ we defined

$$\|\mathbf{p}\|_{\mathbf{i}} = \max\left(|p_1|^{1/i_1}, \ldots, |p_m|^{1/i_m}\right) \text{ and } \|\mathbf{q}\|_{\mathbf{j}} = \max\left(|q_1|^{1/j_1}, \ldots, |q_n|^{1/j_n}\right).$$

Now, for $\mathbf{p} \in \mathbb{R}^m$ and $\mathbf{q} \in \mathbb{R}^n$, if $\mathbf{v} = (\mathbf{p}, \mathbf{q})$ let us define the $(\mathbf{i}, \mathbf{j})$-*quasinorm* $\|\mathbf{v}\|_{\mathbf{i},\mathbf{j}}$ of $\mathbf{v}$ by

$$\|\mathbf{v}\|_{\mathbf{i},\mathbf{j}} := \max(\|\mathbf{p}\|_{\mathbf{i}}^{1/m}, \|\mathbf{q}\|_{\mathbf{j}}^{1/n}).$$

Then for $x \in X$ let

$$\delta_{\mathbf{i},\mathbf{j}}(x) := \inf_{\mathbf{v} \in x \smallsetminus \{0\}} \|\mathbf{v}\|_{\mathbf{i},\mathbf{j}},$$

and for $\varepsilon > 0$ let us consider

$$U_{\mathbf{i},\mathbf{j}}(\varepsilon) := \left\{x \in X : \delta_{\mathbf{i},\mathbf{j}}(x) < \varepsilon\right\}. \tag{8.1}$$

Mahler's Compactness Criterion implies that a subset K of X is relatively compact if and only if the restriction of $\delta_{\mathbf{i},\mathbf{j}}$ to K is bounded away from zero (that is, K is contained in the complement of $U_{\mathbf{i},\mathbf{j}}(\varepsilon)$ for some $\varepsilon > 0$).

Note that in the case $\mathbf{i} = \mathbf{m}$ and $\mathbf{j} = \mathbf{n}$, the $(\mathbf{m}, \mathbf{n})$-quasinorm is simply the sup norm on $\mathbb{R}^d$, $\delta_{\mathbf{m},\mathbf{n}}(x) = \delta(x)$, and $U_{\mathbf{m},\mathbf{n}}(\varepsilon)$ is the same as $U(\varepsilon)$ defined in (1.5). Also it is easy to check that for arbitrary $\mathbf{i}, \mathbf{j}$ and any $x \in X$ one has

$$\delta(x) \geq \delta_{\mathbf{i},\mathbf{j}}(x)^{\max(m,n)}. \tag{8.2}$$

Now we can state a quantitative form of [12, Theorem 6.2], which is also a weighted version of [2, Lemma 3.1]:

Lemma 8.1 *For any $0 < c < 1$, $A \in \mathrm{Bad}_{\mathbf{i},\mathbf{j}}(c)$ if and only if*

$$\{g_t u_A \mathbb{Z}^d : t \geq 0\} \cap U_{\mathbf{i},\mathbf{j}}(\varepsilon) = \varnothing, \tag{8.3}$$

where $\varepsilon = c^{1/d}$.

Proof First note that $g_t u_A \mathbb{Z}^d$ consists of vectors of the form

$$
\begin{pmatrix}
e^{i_1 t}(p_1 + A_1 \mathbf{q}) \\
\vdots \\
e^{i_m t}(p_m + A_m \mathbf{q}) \\
e^{-j_1 t} q_1 \\
\vdots \\
e^{-j_n t} q_n
\end{pmatrix},
$$

where $A_1, \ldots, A_m$ are the rows of A. Suppose that

$$
\|A\mathbf{q} + \mathbf{p}\|_{\mathbf{i}} \|\mathbf{q}\|_{\mathbf{j}} \geq c \tag{8.4}
$$

for all $\mathbf{p} \in \mathbb{Z}^m$ and $\mathbf{q} \in \mathbb{Z}^n \smallsetminus \{0\}$. Take an arbitrary $t \geq 0$. If $\left|e^{-j_\ell t} q_j\right|^{1/j_k} \geq \varepsilon^n$ for some $1 \leq k \leq n$, it follows that

$$
\left\| g_t \begin{pmatrix} A\mathbf{q} + \mathbf{p} \\ \mathbf{q} \end{pmatrix} \right\|_{\mathbf{i},\mathbf{j}} \geq \varepsilon,
$$

and we are done. So suppose that $\left|e^{-j_k t} q_k\right|^{1/j_k} = e^{-t}\left|q_j\right|^{1/j_k} < \varepsilon^n$ for all k. Then we have $\|\mathbf{q}\|_{\mathbf{j}} < \varepsilon^n e^t$. In view of (8.4), there exists $1 \leq k \leq m$ such that

$$
|A_k \mathbf{q} + p_k|^{1/i_k} \geq \frac{c}{\|\mathbf{q}\|_s} > \frac{ce^{-t}}{\varepsilon^n},
$$

hence

$$
\left|e^{i_k t}(A_k \mathbf{q} + p_k)\right|^{1/i_k} = e^t |A_k q + p_k|^{1/i_k} \geq \frac{c}{\varepsilon^n} = \varepsilon^m.
$$

This proves that if $\mathbf{q} \neq 0$, then $g_t u_A \mathbb{Z}^d \notin U_{\mathbf{i},\mathbf{j}}(\varepsilon)$. And if $\mathbf{q} = 0$ and $\mathbf{p} \neq 0$, then

$$
\left\| g_t \begin{pmatrix} A\mathbf{q} + \mathbf{p} \\ \mathbf{q} \end{pmatrix} \right\|_{\mathbf{i},\mathbf{j}} = \left\| g_t \begin{pmatrix} \mathbf{p} \\ \mathbf{0} \end{pmatrix} \right\|_{\mathbf{i},\mathbf{j}} \geq e^{t/m} \|\mathbf{p}\|_{\mathbf{i},\mathbf{j}} \geq 1 \geq \varepsilon.
$$

So $g_t u_A \mathbb{Z}^d \notin U_{\mathbf{i},\mathbf{j}}(\varepsilon)$ holds in this case as well, and we are done.

Vice versa, assume (8.3); that is, suppose that for any nonzero $(\mathbf{p}, \mathbf{q}) \in \mathbb{Z}^{m+n}$ and $t \geq 0$ we have

$$
\left\| \begin{pmatrix}
e^{i_1 t}(p_1 + A_1 \mathbf{q}) \\
\vdots \\
e^{i_m t}(p_m + A_m \mathbf{q}) \\
e^{-j_1 t} q_1 \\
\vdots \\
e^{-j_n t} q_n
\end{pmatrix} \right\|_{\mathbf{i},\mathbf{j}} \geq \varepsilon \tag{8.5}
$$

Fix such $\mathbf{p}$ and $\mathbf{q}$, take an arbitrary $0 < \varepsilon_1 < \varepsilon$, and choose $t \geq 0$ so that

$$
\left\| \begin{pmatrix}
e^{-j_1 t} q_1 \\
\vdots \\
e^{-j_n t} q_n
\end{pmatrix} \right\|_{\mathbf{j}} = e^{-t} \|\mathbf{q}\|_{\mathbf{j}} = \varepsilon_1^n.
$$

Then by (8.5) for some $1 \leq k \leq m$ we must have

$$\left| e^{i_k t} (A_k q + p_k) \right|^{1/i_k} = e^t |A_k q + p_k|^{1/i_k} \geq \varepsilon^m.$$

Consequently $\|Aq + \mathbf{p}\|_{\mathbf{i}} \|\mathbf{q}\|_{\mathbf{j}} \geq \varepsilon^m \varepsilon_1^n$, which, since $\varepsilon_1 < \varepsilon$ was arbitrary, implies that $\|A\mathbf{q} + \mathbf{p}\|_{\mathbf{i}} \|\mathbf{q}\|_{\mathbf{j}} \geq c$. Since $\mathbf{p}$ and $\mathbf{q}$ were arbitrary, $A \in \mathrm{Bad}_{\mathbf{i},\mathbf{j}}(c)$, which finishes the proof of the lemma. $\square$

We will also need a lower bound for the Haar measure of the inner r-core of the set $U_{\mathbf{i},\mathbf{j}}(\varepsilon)$, where $0 < \varepsilon < 1$ and r is small enough. The first step is a weighted version of [15, Proposition 7.1]:

Proposition 8.2 *There exist $C_9, C_{10} > 0$ depending only on d such that for all $0 < \varepsilon < 1$ one has*

$$C_9 \varepsilon^d \geq \mu\big(U_{\mathbf{i},\mathbf{j}}(\varepsilon)\big) \geq C_9 \varepsilon^d - C_{10} \varepsilon^{2d}. \tag{8.6}$$

Proof For $x \in X$ and $1 \leq k \leq d$, denote by $P^k(x)$ the set of all *primitive* (i.e. extendable to a basis of x) ordered k-tuples $(\mathbf{v}_1, \ldots, \mathbf{v}_k)$ of vectors in x. Then, given a function φ on $\mathbb{R}^d$, for any $k = 1, \ldots, d - 1$ define a function $\overset{\wedge k}{\varphi}$ on X by

$$\overset{\wedge k}{\varphi}(x) := \sum_{(\mathbf{v}_1,\ldots,\mathbf{v}_k) \in P^k(x)} \varphi(\mathbf{v}_1, \ldots, \mathbf{v}_k).$$

According to a generalized Siegel's summation formula [15, Theorem 7.3], for any $1 \leq k \leq d$ there exists a constant c_k dependent on k and d such that for any $\varphi \in L^1(\mathbb{R}^{kd})$,

$$\int_X \overset{\wedge k}{\varphi}(\mathbf{v}_1, \ldots \mathbf{v}_k)\, dx = c_k \int_{(\mathbb{R}^d)^k} \varphi\, d\mathbf{v}_1 \ldots d\mathbf{v}_k. \tag{8.7}$$

The case $k = 1$ corresponds to the classical Siegel transform,

$$\widehat{\varphi}(x) := \overset{\wedge 1}{\varphi}(x) = \sum_{\mathbf{v} \in P^1(x)} \varphi(\mathbf{v}),$$

and Siegel's summation formula [19], $\int_X \widehat{\varphi}\, d\mu = c_1 \int_{\mathbb{R}^d} \varphi(\mathbf{v})\, d\mathbf{v}$.

Take $0 < \varepsilon < 1$, denote by D the region in $\mathbb{R}^d$ defined by the following system of inequalities:

$$\begin{aligned} |x_\ell| &< \varepsilon^{m i_\ell} & 1 \leq \ell \leq m, \\ |x_{m+\ell}| &< \varepsilon^{n j_\ell} & 1 \leq \ell \leq n, \end{aligned}$$

and by φ the characteristic function of D. Note that the volume of D is equal to ε^d, and that

$$x \in U_{\mathbf{i},\mathbf{j}}(\varepsilon) \Leftrightarrow x \cap D \neq \{0\}.$$

The latter condition clearly implies that D contains at least two primitive vectors in x. Therefore in view of Siegel's formula we have

$$\mu\big(U_{\mathbf{i},\mathbf{j}}(\varepsilon)\big) \leq \frac{1}{2} \int_X \widehat{\varphi}\, d\mu = \frac{1}{2} c_1 \int_{\mathbb{R}^d} \varphi\, d\mathbf{v} = \frac{1}{2} c_1 \varepsilon^d.$$

For getting the lower bound, note that if two linearly independent primitive vectors $\mathbf{v}_1$ and $\mathbf{v}_2$ in $x \cap D$ do not form a primitive pair, then the line segment between $\mathbf{v}_1$ and $\mathbf{v}_2$ must contain another lattice point; and since D is convex, this lattice point must be in D. So one can easily see that whenever there exist at least two linearly independent vectors in $x \cap D$, for any

$\mathbf{v}_1 \in P^1(x) \cap D$ one can find $\mathbf{v}_2 \in x \cap D$ such that $(\mathbf{v}_1, \mathbf{v}_2)$, as well as $(\mathbf{v}_1, -\mathbf{v}_2)$, belong to $P^2(x)$. Therefore, if $\widehat{\varphi}(x) > 2$, one has

$$\widehat{\varphi}(x) = \#(P^1(x) \cap D) \leq \frac{1}{2} \#\big(P^2(x) \cap (D \times D)\big) = \frac{1}{2} \overset{\wedge_2}{\psi}(x),$$

where ψ is the characteristic function of $D \times D$ in $\mathbb{R}^{2k}$. Hence,

$$\int_X \widehat{\varphi} \, d\mu = \int_{\{x : \widehat{\varphi}(x) \leq 2\}} \widehat{\varphi} \, d\mu + \int_{\{x : \widehat{\varphi}(x) > 2\}} \widehat{\varphi} \, d\mu$$
$$\leq 2\mu(\{x : \widehat{\varphi}(x) = 2\}) + \tfrac{1}{2} \int_{\{x : \widehat{\varphi}(x) > 2\}} \overset{\wedge_2}{\psi} \, d\mu \leq 2\mu\big(U_{\mathbf{i},\mathbf{j}}(\varepsilon)\big) + \tfrac{1}{2} \int_X \overset{\wedge_2}{\psi} \, d\mu,$$

which implies that

$$2\mu\big(U_{\mathbf{i},\mathbf{j}}(\varepsilon)\big) \geq \int_X \widehat{\varphi} \, d\mu - \frac{1}{2} \int_X \overset{\wedge_2}{\psi} \, d\mu = c_1 \varepsilon^d - \frac{1}{2} \int_X \overset{\wedge_2}{\psi} \, d\mu.$$

Using the $k = 2$ case of (8.7) yields $\int_X \overset{\wedge_2}{\psi} \, d\mu = c_2 \varepsilon^{2d}$. Hence (8.6) holds with $C_9 = \frac{1}{2} c_1$ and $C_{10} = \frac{1}{4} c_2$. $\qquad\square$

Finally let us choose $C_{11} > 0$ such that for any $0 < r < 1$,

$$\max\big(\|g - I_d\|_{op}, \|g^{-1} - I_d\|_{op}\big) < C_{11} r \text{ whenever } g \in B^G(r).$$

Lemma 8.3 *Let* $0 < \varepsilon < 1$ *and*

$$0 < r < \frac{2^\alpha - 1}{d C_{11}} \varepsilon^{\max(m,n)}. \tag{8.8}$$

Then

$$U_{\mathbf{i},\mathbf{j}}(\varepsilon/2) \subset \sigma_r\big(U_{\mathbf{i},\mathbf{j}}(\varepsilon)\big).$$

Proof Take $x \in U_{\mathbf{i},\mathbf{j}}(\varepsilon/2)$ and $g \in B^G(r)$. We know that there exists $\mathbf{v} \in x \setminus \{0\}$ such that one of the following two conditions holds:

(1) $|v_k| < (\varepsilon/2)^{m i_k}$ for some $1 \leq k \leq m$;
(2) $|v_{m+k}| < (\varepsilon/2)^{n j_k}$ for some $1 \leq k \leq n$.

Assuming (1) and writing $g = (a_{k\ell})_{k,\ell=1,\ldots,d}$, one has

$$|(g\mathbf{v})_k| = \left| a_{kk} v_k + \sum_{\ell \neq k} a_{k\ell} v_\ell \right| \leq (1 + C_{11} r)(\varepsilon/2)^{m i_k} + (d-1) C_{11} r (\varepsilon/2)^{m\alpha}$$

$$\leq (\varepsilon/2)^{m i_k} + d C_{11} r (\varepsilon/2)^{m\alpha} \leq \frac{\varepsilon^{m i_k}}{2^{m\alpha}} \big(1 + d C_{11} r \varepsilon^{-m}\big),$$

which is smaller than $\varepsilon^{m i_k}$ in view of (8.8); hence $gx \in U_{\mathbf{i},\mathbf{j}}(\varepsilon)$. The argument in case of (2) is similar. $\qquad\square$

Now we can finish the

Proof of Theorem 1.4 In view of Theorem 7.1 and Lemma 8.1, one can apply Theorem 1.3 to P as in (1.17) and conclude that for any $c > 0$ and any $0 < r < \min\big(r_0\big(\partial_{1/2}(X \setminus U_{\mathbf{i},\mathbf{j}}(\varepsilon))\big), r''\big)$, it holds that

$$\text{codim Bad}_{\mathbf{i},\mathbf{j}}(c) \gg \frac{\mu(\sigma_r U_{\mathbf{i},\mathbf{j}}(\varepsilon))}{\log \frac{1}{r} + \log \frac{1}{\mu(\sigma_r U_{\mathbf{i},\mathbf{j}}(\varepsilon))}} \qquad (8.9)$$

where $\varepsilon = c^{1/d}$ and the implicit constant in (8.9) is independent of c but depends on $\mathbf{i}, \mathbf{j}$. Note that in view of (8.2) we have $X \smallsetminus U_{\mathbf{i},\mathbf{j}}(\varepsilon) \subset X \smallsetminus U(\varepsilon^{\max(m,n)})$, thus

$$r_0\Big(\partial_{1/2}\big(X \smallsetminus U_{\mathbf{i},\mathbf{j}}(\varepsilon)\big)\Big) \geq r_0\Big(\partial_{1/2}(X \smallsetminus U(\varepsilon^{\max(m,n)})\Big)$$
$$\gg r_0\left(X \smallsetminus U\Big(\frac{1}{1+C_{11}/2}\varepsilon^{\max(m,n)}\Big)\right) \geq \frac{C_1}{1+C_{11}/2}\varepsilon^{d \cdot \max(m,n)},$$

the last inequality being a consequence of Lemma 7.2. It follows that (8.9) holds whenever

$$r < \frac{C_1}{1+C_{11}/2}\varepsilon^{d \cdot \max(m,n)} \leq r''. \qquad (8.10)$$

Now define

$$c_0 := \min\left(\Big(\frac{1+C_{11}/2}{C_1}r''\Big)^{1/\max(m,n)}, C_9/2C_{10}\right),$$

take $\varepsilon < c_0^{1/d}$ and consider

$$r = \frac{1}{2}\min\left(\frac{2^\alpha - 1}{dC_{11}}, \frac{C_1}{1+C_{11}/2}\right)\varepsilon^{d \cdot \max(m,n)}.$$

Then both (8.8) and (8.10) will hold, and thus the right hand side of (8.9) is not less than

$$\frac{\mu(U_{\mathbf{i},\mathbf{j}}(\varepsilon/2))}{\log \frac{1}{r} + \log \frac{1}{\mu(U_{\mathbf{i},\mathbf{j}}(\varepsilon/2))}} \geq \frac{\frac{1}{2}C_9(\varepsilon/2)^d}{\log \frac{1}{r} + \log \frac{1}{\frac{1}{2}C_9(\varepsilon/2)^d}} \gg \frac{\varepsilon^d}{\log \frac{1}{\varepsilon}},$$

which finishes the proof. $\qquad\qquad\square$

9 Concluding remarks

9.1 Precise estimates for the Hausdorff dimension

Note that in view of the aforementioned result of Simmons [20] and similar results for other dynamical systems (see e.g, [8]), it is natural to expect that when U is either a small ball or the complement of a large compact subset of X, the codimension of $E(F^+, U)$ is, as U shrinks, asymptotic to a constant times the measure of U. That is, conjecturally there should not be a logarithmic term in the right hand side of (1.9). However it is not clear how to improve our upper bound, as well as how to establish a complimentary lower estimate for dim $E(F^+, U)$, using the exponential mixing of the action. Such questions can be asked in other contexts, such as for expanding maps on manifolds, see e.g. [1] for a lower estimate improving on [22].

9.2 A dimension drop problem

Another interesting question is whether the conclusion of Theorem 1.1 holds without the assumption of compactness of U^c. It fact, it is not even known in general that the dimension

of $E(F^+, U)$ is strictly smaller than the dimension of X as long as U is non-empty. In [7] it was established in the case when G is a connected semisimple Lie group of real rank 1. One possible approach to this problem for non-compact homogeneous spaces of higher rank is to combine the methods of the present paper with estimates on the escape of mass for translates of measures on horospherical subgroups, as developed in [13]. This is a work in progress. Recently in [9], by generalizing the methods used in [13] to arbitrary homogeneous spaces, it was shown that for any one parameter subgroup action on a homogeneous space, the Hausdorff dimension of the set of points with divergent trajectories is not full.

References

1. Abercrombie, A.G., Nair, R.: An exceptional set in the ergodic theory of expanding maps on manifolds. Monatsh. Math. **148**, 1–17 (2006)
2. Broderick, R., Kleinbock, D.: Dimension estimates for sets of uniformly badly approximable systems of linear forms. Int. J. Number Theory **11**(7), 2037–2054 (2015)
3. Bernik, V., Kleinbock, D., Margulis, G.A.: Khintchine-type theorems on manifolds: the convergence case for standard and multiplicative versions. Int. Math. Res. Not. **9**, 453–486 (2001)
4. Bourbaki, N.: Eléments de Mathematique, Livre VI: Integration, Chapitre 7: Mesure de Haar, Chapitre 8: Convolution et Representations. Hermann, Paris (1963)
5. Dani, S.G.: Divergent trajectories of flows on homogeneous spaces and Diophantine approximation. J. Reine Angew. Math. **359**, 55–89 (1985)
6. Dani, S.G., Margulis, G.A.: Limit distributions of orbits of unipotent flows and values of quadratic forms. In: I. M. Gelfand Seminar, pp. 91–137, Adv. Soviet Math., Vol. 16, Part 1. American Mathematical Society, Providence, RI (1993)
7. Einsiedler, M., Kadyrov, S., Pohl, A.: Escape of mass and entropy for diagonal flows in real rank one situations. Isr. J. Math. **210**(1), 245–295 (2015)
8. Ferguson, A., Pollicott, M.: Escape rates for Gibbs measures. Ergodic Theory Dyn. Syst. **32**, 961–988 (2012)
9. Guan, L., Shi, R.: Hausdorff dimension of divergent trajectories on homogeneous space (2018). arXiv:1805.07444 (preprint)
10. Kadyrov, S.: Exceptional sets in homogeneous spaces and Hausdorff dimension. Dyn. Syst. **30**(2), 149–157 (2015)
11. Kleinbock, D.: Nondense orbits of flows on homogeneous spaces. Ergodic Theory Dyn. Syst. **18**, 373–396 (1998)
12. Kleinbock, D.: Flows on homogeneous spaces and Diophantine properties of matrices. Duke Math. J. **95**, 107–124 (1998)
13. Kadyrov, S., Kleinbock, D., Lindenstrauss, E., Margulis, G.A.: Singular systems of linear forms and non-escape of mass in the space of lattices. J. Anal. Math. **133**, 253–277 (2017)
14. Kleinbock, D., Margulis, G.A.: Bounded orbits of nonquasiunipotent flows on homogeneous spaces. Sinaí's Moscow Seminar on Dynamical Systems, 141–172, Amer. Math. Soc. Trans. Ser. 2. American Mathematical Society, Providence, RI (1996)
15. Kleinbock, D., Margulis, G.A.: Logarithm laws for flows on homogeneous spaces. Invent. Math. **138**(3), 451–494 (1999)
16. Kleinbock, D., Margulis, G.A.: On effective equidistribution of expanding translates of certain orbits in the space of lattices. Number Theory, Analysis and Geometry, pp. 385–396. Springer, New York (2012)
17. Kleinbock, D., Weiss, B.: Dirichlet's theorem on Diophantine approximation and homogeneous flows. J. Mod. Dyn. **4**, 43–62 (2008)
18. Pollington, A., Velani, S.: On simultaneously badly approximable numbers. J. Lond. Math. Soc. (2) **66**(1), 29–40 (2002)
19. Siegel, C.L.: A mean value theorem in geometry of numbers. Ann. Math. **46**, 340–347 (1945)
20. Simmons, D.: A Hausdorff measure version of the Jarník–Schmidt theorem in Diophantine approximation. Math. Proc. Camb. Philos. Soc. **164**(3), 413–459 (2018)

21. Starkov, A.: Dynamical Systems on Homogeneous Spaces, Translations of Mathematical Monographs, vol. 190. American Mathematical Society, Providence (2000)
22. Urbański, M.: The Hausdorff dimension of the set of points with nondense orbit under a hyperbolic dynamical system. Nonlinearity **4**, 385–397 (1991)
23. Wegmann, H.: Die Hausdorff–Dimension von kartesischen Produktmengen in metrischen Räumen. J. Reine Angew. Math. **234**, 163–171 (1969)

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