Pure and Applied Mathematics Quarterly
Volume 19, Number 2, 529–573, 2023

Genericity on submanifolds and application to universal hitting time statistics

Han Zhang

Abstract: We investigate Birkhoff genericity on submanifolds of homogeneous space $X = SL_d(\mathbb{R}) \ltimes (\mathbb{R}^d)^k / SL_d(\mathbb{Z}) \ltimes (\mathbb{Z}^d)^k$, where $d \geq 2$ and $k \geq 1$ are fixed integers. The submanifolds we consider are parameterized by unstable horospherical subgroup U of a diagonal flow a_t in $SL_d(\mathbb{R})$. As long as the intersection of the submanifold with any affine rational subspace has Lebesgue measure zero, we show that the trajectory of a_t along Lebesgue almost every point on the submanifold gets equidistributed on X. This generalizes the previous work of Frączek, Shi and Ulcigrai in [8].

Following the scheme developed by Dettmann, Marklof and Strömbergsson in [3], we then deduce an application to universal hitting time statistics for integrable flows.

Keywords: Homogeneous dynamics, ergodic theory, equidistribution, diagonal flow.

1. Introduction

Let $(X, \mathcal{B}, \mu, R)$ be a probability measure preserving system, where $(X, \mathcal{B})$ is a Borel measurable space with probability measure μ, and $R^t : X \to X$ is an $\mathbb{R}$-action (or $\mathbb{Z}$-action) preserving μ. Assume that R is ergodic, then Birkhoff's ergodic theorem (cf. [4]) asserts that, for any $f \in L^1_\mu(X)$,

$$(1.1) \quad \frac{1}{T}\int_0^T f(R^t x)dt \xrightarrow{T\to\infty} \int_X f d\mu, \left(\text{or } \frac{1}{N}\sum_{n=1}^{N} f(R^n x) \xrightarrow{N\to\infty} \int_X f d\mu\right),$$

for μ-almost every $x \in X$. In particular, (1.1) holds for any $f \in C_c(X)$, where $C_c(X)$ denotes the collection of all compactly supported and continuous

Received April 29, 2022.
2010 Mathematics Subject Classification: Primary 37A17, 37A50; secondary 37J35.

functions on X. Therefore, (1.1) implies that for μ-almost every $x \in X$,

$$\frac{1}{T}\int_0^T \delta_{R^t x}dt \xrightarrow{T\to\infty} \mu, \left(\text{or } \frac{1}{N}\sum_{n=1}^{N} \delta_{R^n x} \xrightarrow{N\to\infty} \mu\right), \tag{1.2}$$

in the weak* topology on the set of all probability measures on X. Here δ is the Dirac measure on X.

For $x \in X$, we say that x is **Birkhoff generic** with respect to (μ, R) if x satisfies (1.2). Given a Radon measure ν on X (possibly singular to μ), if ν-almost every $x \in X$ is Birkhoff generic with respect to (μ, R), we say that ν is Birkhoff generic with respect to (μ, R). It is then natural to ask the following.

Question 1.1. Under what conditions the measure ν is Birkhoff generic with respect to (μ, R)?

This question had been previously studied in the case of $X = \mathbb{R}/\mathbb{Z}$, μ_X is the Lebesgue measure on X and $R^n = \times n \bmod \mathbb{Z}$ in [10]. It was shown that for any $m, n \in \mathbb{N}$, any R^m invariant ergodic probability measure ν is Birkhoff generic with respect to (R^n, μ_X). This result was strengthened later in [9]. An analogous question was also studied in the context of moduli space of translation surfaces in [29].

We consider Question 1.1 in the setting of homogeneous dynamics. Let $X = G/\Gamma$, where G is a Lie group and Γ is a lattice in G. Here and hereafter let $\mu = \mu_X$ be the G-invariant probability measure on X. Let $\{R^t\}_{t\in\mathbb{R}}$ be a one-parameter flow in X. Assume that $R^t = u(t)$, where $\{u(t)\}_{t\in\mathbb{R}}$ is a one-parameter unipotent flow, that is, the adjoint action of $u(t)$ on the Lie algebra of G is unipotent. In this case, Ratner's uniform distribution theorem [20] says that for any $x \in X$, x is Birkhoff generic with respect to $(\nu, u(t))$, where ν is the $u(t)$-invariant probability measure supported on the orbit closure $\overline{\{u(t)x : t \in \mathbb{R}\}}$. By Ratner's orbit closure theorem [19], these orbit closures are all homogeneous. Thus this provides a satisfactory answer to Question 1.1.

On the other hand, when $R^t = a_t$, here and hereafter $\{a_t\}_{t\in\mathbb{R}}$ is a one-parameter diagonal flow on G, that is, the adjoint action of a_t on Lie algebra of G is semisimple, a description of Birkhoff genericity of $x \in X$ under the flow a_t is much harder. Indeed, the question of describing the orbit closures of diagonal action remains open (cf. [15][27, Conjecture 1]).

Nevertheless, there is a natural class of probability measures ν on X that are interesting to study with respect to Question 1.1. This class of measures is given as follows. We define the unstable horospherical subgroup U^+ with respect to a_t by

$$U^+ := \{g \in G : a_{-t} g a_t \xrightarrow{t\to\infty} Id\},$$

where Id is the identity element of G. Let $Y \subset U^+$ be a submanifold and ν be a normalized bounded supported volume measure of Y (here and hereafter by normalized measure, we mean that ν is renormalized to be a probability measure). It has been proved in [21][22][23][28] that if Y satisfies certain algebraic conditions, then the translation of the measure ν under a_t converges weakly to μ as $t \to \infty$. That is, for any $f \in C_c(X)$,

$$\int f(a_t x) d\nu(x) \xrightarrow{t\to\infty} \int f d\mu.$$

As in Question 1.1, it is curious to ask the following

Question 1.2. Assume that a bounded supported normalized volume measure ν of a submanifold $Y \subset U^+$ is such that the translate of ν under a_t is equidistributed with respect to μ, is it true that ν is also Birkhoff generic with respect to (μ, a_t)?

Roughly speaking, Question 1.2 is answered when the manifold Y is considerably "large" compared to the unstable horosphere subgroup of a_t. In [25], Shi considered the situation where G is a semisimple Lie group, $Y = U$ is the a_t expanding subgroup (cf. [24]) of U^+ and ν is a normalized bounded supported Haar measure on U. By [21], ν satisfies the assumption of Question 1.2. Shi showed that ν is also Birkhoff generic with respect to (μ, a_t), and thus gave an affirmative answer to Question 1.2. In the special case where $G = SL_d(\mathbb{R})$ and $\Gamma = SL_d(\mathbb{Z})$, authors in [13] also obtained the effective convergence rate of (1.2).

In [8], the authors considered the setting where $G = SL_2(\mathbb{R}) \ltimes \mathbb{R}^2$, $\Gamma = SL_2(\mathbb{Z}) \ltimes \mathbb{Z}^2$ and $X = G/\Gamma$. One of the main results in [8] asserts that if Y is a C^1 curve in U^+ that intersects any affine rational line in a Lebesgue null set, then a normalized bounded supported volume measure ν on Y is Birkhoff generic with respect to (μ, a_t). By the equidistribution result of translation of such ν under a_t in [3], the result in [8] also gives an affirmative answer to Question 1.2 in the case where Y is a curve.

Let $X = SL_3(\mathbb{R})/SL_3(\mathbb{Z})$. In a recent preprint [12, Theorem 1.4], it is shown that when the natural measure ν on the planar line $L \subset U^+$ gets equidistributed under a_t, then for ν almost every point x in L, the orbit $\{a_t x\}_{t\geq 0}$ is dense in X. This supports an affirmative answer to Question 1.2.

The aim of this paper is to generalize the genericity results in [8] to $X = SL_d(\mathbb{R}) \ltimes (\mathbb{R}^d)^k / SL_d(\mathbb{Z}) \ltimes (\mathbb{Z}^d)^k$ and gives an affirmative answer to

Question 1.2 when the manifold Y satisfies certain diophantine condition. We also deduce an application of our results to the statistics of universal hitting time for integrable flows.

1.1. Notations

From now on, any vector in the Euclidean space will be taken to be a column vector, and we will use boldface letters to denote vectors and matrices. Also, a.e. will be the shorthand for Lebesgue almost everywhere. $|\cdot|$ will denote Lebesgue measure of measurable subsets of Euclidean space or absolute value of real numbers. $\|\cdot\|$ will denote the standard Euclidean norm and $\|\cdot\|_\infty$ the sup norm of a vector or matrix. Throughout this article, for two matrices A and B, $A \cdot B$ will denote **matrix multiplication.**

For $m, n \in \mathbb{N}$, $Mat_{m\times n}(\mathbb{R})$ will denote the space of m by n real matrices. $(\mathbb{R}^n)^m$ is the direct product of m copies of $\mathbb{R}^n$.

Fix integers $d \geq 2$, $k \geq 1$. Let $G' = SL_d(\mathbb{R})$, $G = SL_d(\mathbb{R}) \ltimes (\mathbb{R}^d)^k$, $\Gamma' = SL_d(\mathbb{Z})$ and $\Gamma = SL_d(\mathbb{Z}) \ltimes (\mathbb{Z}^d)^k$. It is well known that Γ' is a lattice in G' and Γ is a lattice in G.

Let $X = G/\Gamma$. Denote μ_X the G-invariant probability measure on G/Γ. Note that the action of G' on $(\mathbb{R}^d)^k$ is given by

$$g \cdot \mathbf{v} = (g \cdot \mathbf{v_1}, \cdots, g \cdot \mathbf{v_k}),$$

where $g \in G'$ and $\mathbf{v} = (\mathbf{v_1}, \cdots, \mathbf{v}_k)$ with $\mathbf{v_i} \in \mathbb{R}^d$. The multiplication law in G is given by

$$(g, \mathbf{v}) \cdot (g', \mathbf{v'}) = (g \cdot g', \mathbf{v} + g \cdot \mathbf{v'}).$$

G' is naturally embedded into G by

$$G' \cong (G', \mathbf{0}) \leq G.$$

Fix an $r \in \{1, \cdots, d-1\}$. For $t \in \mathbb{R}$ and $\mathbf{s} \in Mat_{r\times(d-r)}(\mathbb{R})$, denote

$$a_t = diag[e^{(d-r)t}, \cdots, e^{(d-r)t}, e^{-rt}, \cdots, e^{-rt}], \tag{1.3}$$

$$u(\mathbf{s}) = \begin{bmatrix} \mathbf{1}_r & \mathbf{s} \\ \mathbf{0}_{d-r,r} & \mathbf{1}_{d-r} \end{bmatrix}, \tag{1.4}$$

$$U := \{u(\mathbf{s}) : \mathbf{s} \in Mat_{r\times(d-r)}(\mathbb{R})\} \cong Mat_{r\times(d-r)}(\mathbb{R}). \tag{1.5}$$

For a column vector or a matrix $\mathbf{v}$, let $(\mathbf{v})_{\le r}$ (or $(\mathbf{v})_{>r}$) be the first r rows (or last $d-r$ rows) of $\mathbf{v}$. For example, if $\mathbf{v} \in (\mathbb{R}^d)^k$, then

$$(\mathbf{v})_{\le r} \in (\mathbb{R}^r)^k, \ (\mathbf{v})_{>r} \in (\mathbb{R}^{d-r})^k.$$

With the above notations, the unstable horospherical subgroup U^+ of a_t in G is

$$U^+ = U \cdot \left\{ \left(Id, \begin{bmatrix} (\mathbf{v})_{\le r} \\ \mathbf{0} \end{bmatrix} \right) : \mathbf{v} \in (\mathbb{R}^d)^k \right\}.$$

Lastly, for a map $\boldsymbol{\varphi} : Mat_{r\times(d-r)}(\mathbb{R}) \to (\mathbb{R}^d)^k$, we write $\boldsymbol{\varphi}(\mathbf{s}) = (\varphi_{ij}(\mathbf{s}))_{1\le i\le d, 1\le j\le k}$.

We also write

$$u_{\boldsymbol{\varphi}}(\mathbf{s}) := u(\mathbf{s}) \cdot (Id, \boldsymbol{\varphi}(\mathbf{s})). \tag{1.6}$$

1.2. Main results

For any $\mathbf{s} \in Mat_{r\times(d-r)}(\mathbb{R})$ and $T > 0$, define the probability measure

$$\mu_{\mathbf{s},T} = \frac{1}{T}\int_0^T \delta_{a_t u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma} dt. \tag{1.7}$$

As before, we say that $u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_t) if $\mu_{\mathbf{s},T}$ converges to μ_X in the weak*-topology as $T \to \infty$.

One of our main results is the following:

Theorem 1.3. *Let $\mathcal{U} \subset Mat_{r\times(d-r)}(\mathbb{R})$ be a bounded open subset. Let $\boldsymbol{\varphi} : \mathcal{U} \to (\mathbb{R}^d)^k$ be a C^1-map satisfying $(\boldsymbol{\varphi}(\mathbf{s}))_{>r} \equiv \mathbf{0}$ for any $\mathbf{s} \in \mathcal{U}$. Assume that for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$,*

$$|\{\mathbf{s} \in \mathcal{U} : (\boldsymbol{\varphi}(\mathbf{s}))_{\le r} \cdot \mathbf{m} \in \mathbf{s} \cdot \mathbb{Z}^{d-r} + \mathbb{Z}^r\}| = 0, \tag{1.8}$$

then for Lebesgue a.e. $\mathbf{s} \in \mathcal{U}$, $u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_t).

Using the observation that if a trajectory equidistributes with respect to μ_X, then all other parallel trajectories will also equidistribute with respect to μ_X (see Lemma 2.3), we can remove the assumption that $(\boldsymbol{\varphi}(\mathbf{s}))_{>r} \equiv \mathbf{0}$ and strengthen Theorem 1.3 to the following.

Corollary 1.4. *Let $\mathcal{U}$ be a bounded open subset of $Mat_{r\times(d-r)}(\mathbb{R})$. Let $\boldsymbol{\varphi} : \mathcal{U} \to (\mathbb{R}^d)^k$ be a C^1 map. If for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$,*

$$|\{\mathbf{s} \in \mathcal{U} : (\boldsymbol{\varphi}(\mathbf{s}))_{\leq r} \cdot \mathbf{m} \in \mathbf{s} \cdot \mathbb{Z}^{d-r} + \mathbb{Z}^r\}| = 0,$$

then for Lebesgue a.e. $\mathbf{s} \in \mathcal{U}$, $u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_t).

We also obtain the following variants of Theorem 1.3.

Theorem 1.5. *Let $\mathcal{U}$ be a bounded open subset of $Mat_{r\times(d-r)}(\mathbb{R})$. Let $\boldsymbol{\varphi} : \mathcal{U} \to (\mathbb{R}^d)^k$ be a C^1 map. Let $\mathbf{M} \in SL_d(\mathbb{R})$. If for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$,*

$$\left|\left\{\mathbf{s} \in \mathcal{U} : \boldsymbol{\varphi}(\mathbf{s}) \cdot \mathbf{m} \in \mathbf{M}^{-1} u(-\mathbf{s}) \cdot \begin{bmatrix} \mathbf{0} \\ \mathbb{R}^{d-r} \end{bmatrix} + \mathbb{Z}^d \right\}\right| = 0. \tag{1.9}$$

Then for Lebesgue a.e. $\mathbf{s} \in \mathcal{U}$, $u(\mathbf{s})\mathbf{M}(Id, \boldsymbol{\varphi}(\mathbf{s}))\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_t).

Remark 1.6. By the equidistribution result in [3], Theorem 1.5 gives an affirmative answer to Question 1.2.

Remark 1.7. The conditions (1.8) and (1.9) are indeed necessary. For example in Theorem 1.3, suppose that $(\boldsymbol{\varphi}(\mathbf{s}))_{\leq r} \cdot \mathbf{m} \in \mathbf{s} \cdot \mathbb{Z}^{d-r} + \mathbb{Z}^r$ for some $\mathbf{s} \in \mathcal{U}$ and $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, then the a_t trajectory along $u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma$ will concentrate on a proper submanifold of G/Γ.

Corollary 1.8. *Let $\mathcal{U}$ be a bounded open subset of $Mat_{r\times(d-r)}(\mathbb{R})$. Let $\boldsymbol{\varphi} : \mathcal{U} \to (\mathbb{R}^d)^k$ be a C^1 map. Let $(\mathbf{M}, \mathbf{v}) \in G$. If for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$,*

$$\left|\left\{\mathbf{s} \in \mathcal{U} : (\boldsymbol{\varphi}(\mathbf{s}) + \mathbf{v}) \cdot \mathbf{m} \in u(-\mathbf{s}) \cdot \begin{bmatrix} \mathbf{0} \\ \mathbb{R}^{d-r} \end{bmatrix} + \mathbf{M} \cdot \mathbb{Z}^d \right\}\right| = 0,$$

then for Lebesgue a.e. $\mathbf{s} \in \mathcal{U}$, $u_{\boldsymbol{\varphi}}(\mathbf{s})(\mathbf{M}, \mathbf{v})\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_t).

For $1 \leq i \leq d$, let $\mathbf{e}_i \in \mathbb{R}^d$ be the column vector such that i-th row of $\mathbf{e}_i$ is 1 and others are 0.

Corollary 1.9. *Let $\mathcal{U}$ be a bounded open subset of $Mat_{r\times(d-r)}(\mathbb{R})$. Let $\mathbf{E_1} : \mathcal{U} \to SO_d(\mathbb{R})$ be a smooth map such that the map $\mathbf{s} \mapsto \mathbf{E_1}(\mathbf{s})^{-1} \cdot [\mathbf{e}_{r+1}, \cdots, \mathbf{e}_d]$ has a nonsingular differential at Lebesgue almost every $\mathbf{s} \in \mathcal{U}$. Let $\boldsymbol{\varphi} : \mathcal{U} \to (\mathbb{R}^d)^k$ be a C^1 map. Assume that for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$,*

$$\left|\left\{\mathbf{s} \in \mathcal{U} : \boldsymbol{\varphi}(\mathbf{s}) \cdot \mathbf{m} \in \mathbf{E}_1(\mathbf{s})^{-1} \cdot \begin{bmatrix} \mathbf{0} \\ \mathbb{R}^{d-r} \end{bmatrix} + \mathbb{Z}^d \right\}\right| = 0,$$

then for Lebesgue a.e. $\mathbf{s} \in \mathcal{U}$, $\mathbf{E}_1(\mathbf{s})(Id, \boldsymbol{\varphi}(\mathbf{s}))\Gamma$ *is Birkhoff generic with respect to* (X, μ_X, a_t).

Corollary 1.9 will allow us to deduce an application to universal hitting time for integrable flows in d-torus $\mathbb{T}^d$ (see Theorem 9.5).

1.3. Ingredients of the proof

Proof of Theorem 1.3 follows the similar strategy as in [8]. However, some new ingredients are required. We need the description of orbit closures of G' in X. This can be done using Ratner's orbit closure theorem following the approach in [3].

We need to construct a suitable mixed height function in our situation, which measures the distance of point to the cusp and singular submanifolds.

Also due to higher rank, some technical difficulties arise in the proof of uniform contraction property of the mixed height function. To overcome these difficulties, we apply a linear algebra lemma (see Lemma 6.13) which is inspired by the proof of [11, Proposition 3.4].

1.4. Overview

In Section 2, we make some reductions and give a proof of Theorem 1.3 and Corollary 1.4.

In Section 3, we will investigate the orbit closure of G' in X using Ratner's orbit closure theorem.

In Section 4, we prove that for a.e. $\mathbf{s} \in \mathcal{U}$, the limit measure is invariant under the unipotent group U. This enables us to apply Ratner's measure classification theorem.

In Section 5 and Section 6, we will construct mixed height function $\beta_{\mathbf{m}}$ for $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, and give a proof of its uniform contraction property.

In Section 7, we prove Proposition 2.1 using mixed height function.

In Section 8, we deduce variants of Theorem 1.3.

In Section 9, we deduce an application to universal hitting time statistics.

2. Reductions and proof of Theorem 1.3

In this section, assuming several Theorems/Propositions/Lemmas that will be proved later, we give a proof of Theorem 1.3.

By Proposition 4.1, for a.e. $\mathbf{s} \in \mathcal{U}$, after possibly passing to a subsequence the weak* limit $\mu_{\mathbf{s}}$ of $\mu_{\mathbf{s},T}$ is U-invariant. From the definition of $\mu_{\mathbf{s},T}$ (see

(1.7)), it follows that $\mu_{\mathbf{s}}$ is also $D = \{a_t : t \in \mathbb{R}\}$-invariant. Hence for a.e. $\mathbf{s} \in \mathcal{U}$, $\mu_{\mathbf{s}}$ is DU-invariant. Note that DU is an epimorphic subgroup of $G' = SL_d(\mathbb{R})$. By [17], as $\mu_{\mathbf{s}}$ is a probability measure invariant under DU, $\mu_{\mathbf{s}}$ is G'-invariant. By Ratner's measure classification theorem, any G' invariant and ergodic probability measure is supported on an orbit closure of G' on X.

A consequence of Ratner's orbit closure theorem (Theorem 3.1) shows that any orbit closure of G' is either

(1) the whole X, or

(2) concentrated in a proper closed submanifold $X_{\mathbf{m}}$ for some $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, where

$$X_{\mathbf{m}} = \left\{(g, g\mathbf{v})\Gamma : g \in G', \mathbf{v} \cdot \mathbf{m} \in \mathbb{Z}^d\right\}.$$

Therefore, it remains to show that for a.e. $\mathbf{s} \in \mathcal{U}$, $\mu_{\mathbf{s}}$ is a probability measure on X and $\mu_{\mathbf{s}}(X_{\mathbf{m}}) = 0$ for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$.

Let

$$M_1 := N_1 \cdot \left(\max_{1\le i\le r, 1\le j\le d-r} \sup_{s\in\mathcal{U}} \|\partial_{ij}\boldsymbol{\varphi}(\mathbf{s})\|_\infty\right) + 1, \tag{2.1}$$

where $N_1 = 8r^2k^{1/2}(d-r)$, and $\partial_{ij}\boldsymbol{\varphi}$ is a d by k matrix whose (p,q)-th entry is $\partial\varphi_{pq}/\partial s_{ij}$. Here the choice of N_1 is flexible, we just choose a value for N_1 that is convenient for us.

By assumption (1.8) of Theorem 1.3, for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, the set

$$\begin{aligned} Bad_{\mathbf{m}} = \Big\{\mathbf{s} \in \mathcal{U} : \exists \mathbf{a} \in \mathbb{Z}^{d-r}, \mathbf{b} \in \mathbb{Z}^r \text{ such that } (\boldsymbol{\varphi}(\mathbf{s}))_{\le r} \cdot \mathbf{m} = \mathbf{s} \cdot \mathbf{a} + \mathbf{b}, \\ \text{and } \|\mathbf{a}\|_\infty \le M_1 \|\mathbf{m}\| \Big\} \end{aligned} \tag{2.2}$$

has Lebesgue measure zero.

Since there are only finitely many $\mathbf{a} \in \mathbb{Z}^{d-r}$ such that $\|\mathbf{a}\|_\infty \le M_1 \cdot \|\mathbf{m}\|$, $Bad_{\mathbf{m}}$ is a closed set with Lebesgue measure zero. Thus to prove Theorem 1.3, it suffices to prove it for a closed cube contained in $\mathcal{U} \setminus Bad_{\mathbf{m}}$. Now let's fix a closed cube $I \subset \mathcal{U} \setminus Bad_{\boldsymbol{m}}$.

Let K be a measurable subset of X. For any $T > 0$, we define the average operator $\mathcal{A}_K^T : \mathcal{U} \to [0,1]$ by

$$\mathcal{A}_K^T(\mathbf{s}) = \frac{1}{T}\int_0^T \chi_K(a_t u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma)dt = \mu_{\mathbf{s},T}(K),$$

where χ_K is the characteristic function of K.

The key proposition, which ensures that $\mu_{\mathbf{s}}$ is a probability measure putting zero mass on $X_{\mathbf{m}}$ for a.e. $\mathbf{s} \in I$, is the following:

Proposition 2.1. *Let* $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$. *Let* $\boldsymbol{\varphi} : I \to (\mathbb{R}^d)^k$ *be a* C^1 *map satisfying* $(\boldsymbol{\varphi}(\mathbf{s}))_{>r} \equiv \mathbf{0}$ *for any* $\mathbf{s} \in I$. *Suppose that*

$$\inf_{\mathbf{s}\in I}\{\|(\boldsymbol{\varphi}(\mathbf{s}))_{\le r}\cdot\mathbf{m} - \mathbf{s}\cdot\mathbf{a} - \mathbf{b}\|_\infty : \|\mathbf{a}\|_\infty \le M_1\|\mathbf{m}\|, \mathbf{a}\in\mathbb{Z}^{d-r}, \mathbf{b}\in\mathbb{Z}^r\} > 0. \tag{2.3}$$

Then for any $\epsilon > 0$, *there exists a compact subset* $K_\epsilon \subset X \setminus X_{\mathbf{m}}$ *and* $\mathfrak{v} > 0$ *such that for any* $T > 0$,

$$|\{\mathbf{s} \in I : \mathcal{A}^T_{K_\epsilon}(\mathbf{s}) \le 1-\epsilon\}| \le e^{-\mathfrak{v}T}|I|. \tag{2.4}$$

It will be proved in Lemma 6.1 that condition (2.3) in Proposition 2.1 follows from condition (1.8) in Theorem 1.3.

Proposition 2.1 will be proved in Section 6. Combining Borel-Cantelli lemma, a direct consequence of Proposition 2.1 is the following:

Proposition 2.2. *Under the assumptions of Theorem 1.3, for a.e.* $\mathbf{s} \in \mathcal{U}$, *by possibly passing to a subsequence,* $\mu_{\mathbf{s},T}$ *converges to a probability measure* $\mu_{\mathbf{s}}$ *on* X *in weak*-topology as* $T \to \infty$, *and* $\mu_{\mathbf{s}}(X_{\mathbf{m}}) = 0$ *for any* $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$.

Proof. Fix an $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$ and $\epsilon > 0$. By Proposition 2.1, we can choose a compact subset K_ϵ of $X \setminus X_{\mathbf{m}}$ such that (2.4) holds for any $T > 0$. Let $T = n \in \mathbb{N}$ and apply Borel-Cantelli lemma to the collection of the sets

$$\{\mathbf{s} \in I : \mathcal{A}^n_{K_\epsilon}(\mathbf{s}) \le 1-\epsilon\}, n \in \mathbb{N}.$$

We can find a measurable subset $I^\epsilon_{\mathbf{m}}$ of I with full measure such that for any $\mathbf{s} \in I^\epsilon_{\mathbf{m}}$, $\mathcal{A}^n_{K_\epsilon}(\mathbf{s}) > 1-\epsilon$ for all sufficiently large $n \in \mathbb{N}$. Therefore, for any $\mathbf{s} \in I^\epsilon_{\mathbf{m}}$, $\mu_{\mathbf{s}}(X) \ge 1-\epsilon$ and $\mu_{\mathbf{s}}(X_{\mathbf{m}}) \le \epsilon$. Let $I_{\mathbf{m}} = \cap_{n=1}^\infty I^{\frac{1}{n}}_{\mathbf{m}}$, then $I_{\mathbf{m}}$ has full Lebesgue measure in I, and for any $\mathbf{s} \in I_{\mathbf{m}}$, $\mu_{\mathbf{s}}(X) = 1$ and $\mu_{\mathbf{s}}(X_{\mathbf{m}}) = 0$.

To complete the proof, we let $I' = \bigcap_{\mathbf{m}\in\mathbb{Z}^k\setminus\{\mathbf{0}\}} I_{\mathbf{m}}$. Then I' has full Lebesgue measure in I and the proposition holds for all $\mathbf{s} \in I'$. ☐

Assuming Proposition 2.2, Proposition 4.1 and Theorem 3.1, we are ready to prove Theorem 1.3:

Proof of Theorem 1.3. By Proposition 2.2 and Proposition 4.1, we conclude that for a.e. $\mathbf{s} \in \mathcal{U}$, the weak* limit $\mu_{\mathbf{s}}$ of $\mu_{\mathbf{s},T}$ as $T \to \infty$ is

(1) a probability measure on X, and $\mu_{\mathbf{s}}(X_{\mathbf{m}}) = 0$ for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$;

(2) DU-invariant.

Since DU is an epimorphic subgroup of $G' = SL_d(\mathbb{R})$, and $\mu_{\mathbf{s}}$ is a DU-invariant probability measure on X, $\mu_{\mathbf{s}}$ is G'-invariant by [17, Theorem 1].

By Ratner's measure classification theorem [19], any ergodic component of such $\mu_{\mathbf{s}}$ is supported on an orbit closure of G' on X. Theorem 3.1 describes all the possible orbit closures of G' on X: either it is X or it is concentrated on $X_{\mathbf{m}}$ for some $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$.

Since $\mu_{\mathbf{s}}(X_{\mathbf{m}}) = 0$ for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, we conclude that for a.e. $\mathbf{s} \in \mathcal{U}$, $\mu_{\mathbf{s}} = \mu_X$. □

We note the following

Lemma 2.3. *Assume that for some* $x \in X = G/\Gamma$,

$$\frac{1}{T}\int_0^T \delta_{a_t x} dt \xrightarrow{T\to\infty} \mu_{G/\Gamma}, \text{ in weak*-topology,}$$

and for some $g \in G$, $a_t g a_{-t} \to Id \in G$ *as* $t \to \infty$, *then*

$$\frac{1}{T}\int_0^T \delta_{a_t g x} dt \xrightarrow{T\to\infty} \mu_{G/\Gamma}, \text{ in weak*-topology.}$$

For any $\boldsymbol{\varphi} : Mat_{r\times(d-r)}(\mathbb{R}) \to (\mathbb{R}^d)^k$, and any $\mathbf{s} \in Mat_{r\times(d-r)}(\mathbb{R})$, we can write

$$\begin{aligned} a_t u_{\boldsymbol{\varphi}}(\mathbf{s}) &= a_t(u(\mathbf{s}), \boldsymbol{\varphi}(\mathbf{s})) \\ &= a_t(Id, \begin{bmatrix} \mathbf{0} \\ (\boldsymbol{\varphi}(\mathbf{s}))_{>r} \end{bmatrix}) a_{-t} \cdot a_t(u(\mathbf{s}), \begin{bmatrix} (\boldsymbol{\varphi}(\mathbf{s}))_{\le r} \\ \mathbf{0} \end{bmatrix}), \end{aligned}$$

where

$$\boldsymbol{\varphi}(\mathbf{s}) = \begin{bmatrix} \mathbf{0} \\ (\boldsymbol{\varphi}(\mathbf{s}))_{>r} \end{bmatrix} + \begin{bmatrix} (\boldsymbol{\varphi}(\mathbf{s}))_{\le r} \\ \mathbf{0} \end{bmatrix}.$$

Since

$$a_t(Id, \begin{bmatrix} \mathbf{0} \\ (\boldsymbol{\varphi}(\mathbf{s}))_{>r} \end{bmatrix}) a_{-t} \to (Id, \mathbf{0}),$$

by Lemma 2.3 and Theorem 1.3, Corollary 1.4 is proven.

3. Orbit closure

In this section, we will classify all orbit closures of G' in X following [3]. Recall that $G' = SL_d(\mathbb{R})$ and $G = SL_d(\mathbb{R}) \ltimes (\mathbb{R}^d)^k$.

Consider a base point $(Id, \boldsymbol{\xi}) \in G$. Since G' is a simple Lie group, an application of Ratner's orbit closure theorem gives the following theorem describing the orbit closure of $G' \cdot (Id, \boldsymbol{\xi})\Gamma/\Gamma$ in G/Γ:

Theorem 3.1. *The orbit closure* $\overline{G' \cdot (Id, \boldsymbol{\xi})\Gamma/\Gamma}$ *is* G/Γ *if and only if for any* $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, $\boldsymbol{\xi} \cdot \mathbf{m} \notin \mathbb{Z}^d$.

By Ratner's orbit closure theorem ([20]), for any $\boldsymbol{\xi} \in (\mathbb{R}^d)^k$, there exists a closed subgroup H of G containing G' such that

$$\overline{G' \cdot (Id, \boldsymbol{\xi})\Gamma/\Gamma} = H \cdot (Id, \boldsymbol{\xi})\Gamma/\Gamma,$$

and $H \cdot (Id, \boldsymbol{\xi})\Gamma/\Gamma$ admits an H-invariant probability measure.

It can be checked that if there exists $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$ such that $\boldsymbol{\xi} \cdot \mathbf{m} \in \mathbb{Z}^d$, then $G' \cdot (Id, \boldsymbol{\xi})\Gamma \subset X_{\mathbf{m}}$, where

$$X_{\mathbf{m}} = \{(g, g\mathbf{v})\Gamma : g \in G', \mathbf{v} \in (\mathbb{R}^d)^k \text{ such that } \mathbf{v} \cdot \mathbf{m} \in \mathbb{Z}^d\}. \tag{3.1}$$

This is a closed submanifold of X of codimension d. In this case, the orbit $G' \cdot (Id, \boldsymbol{\xi})\Gamma/\Gamma$ does not equidistribute in X.

The converse of Theorem 3.1 will follow from Lemmas 3.2-3.4. We will follow the proof strategy of [3, Theorem 3]. Let $\boldsymbol{\xi}$, H be as above.

Lemma 3.2. *There is a linear subspace* $\boldsymbol{U} \subset \mathbb{R}^k$ *such that* $H = SL_d(\mathbb{R}) \ltimes L(\boldsymbol{U})$, *where* $L(\boldsymbol{U})$ *is a subset of* $Mat_{d\times k}(\mathbb{R})$ *such that for any element* $\mathbf{v}$ *of* $L(\boldsymbol{U})$, *each row vector of* $\mathbf{v}$ *is a vector in* $\boldsymbol{U}$.

Proof. Let $\boldsymbol{L} = \{\mathbf{v} \in (\mathbb{R}^d)^k : (Id, \mathbf{v}) \in H\}$. Because $G' \subset H$, for any $\mathbf{v} \in \boldsymbol{L}$, we have $(g, \mathbf{0}) \cdot (Id, \mathbf{v}) \cdot (g, \mathbf{0})^{-1} = (Id, g\mathbf{v}) \in H$. It follows that $\boldsymbol{L}$ is G'-invariant and $SL_d(\mathbb{R}) \ltimes \boldsymbol{L} \subset H$. For any $(g, \mathbf{v}) \in H$, we have $(g^{-1}, \mathbf{0}) \cdot (g, \mathbf{v}) = (Id, g^{-1}\mathbf{v}) \in H$, so $g^{-1}\mathbf{v} \in \boldsymbol{L}$. Since $\boldsymbol{L}$ is G'-invariant, $\mathbf{v} \in \boldsymbol{L}$. Therefore $H = SL_d(\mathbb{R}) \ltimes \boldsymbol{L}$.

Let $A \in \mathfrak{sl}_d = Lie(G')$, then for any $t \in \mathbb{R}$, and any $\mathbf{v} \in \boldsymbol{L}$

$$\frac{exp(tA)\mathbf{v} - \mathbf{v}}{t} \in \boldsymbol{L}.$$

Let $t \to 0$, we obtain $A \cdot \mathbf{v} \in \boldsymbol{L}$. Recall that $\mathfrak{sl}_d$ consists of all trace zero $d \times d$ matrices. Let $\mathbf{E}_{ij}$ be the $d \times d$ matrix with 1 in the (i, j)-th entry and zero for all other entries. Then for any $i \neq j$, $\mathbf{E}_{ij}\mathbf{v} \in \boldsymbol{L}$. Since $\mathbf{E}_{ij} \cdot \mathbf{E}_{ji} = \mathbf{E}_{ii}$, for any i we have $\mathbf{E}_{ii}\mathbf{v} \in \boldsymbol{L}$ as well. Therefore, $\boldsymbol{L}$ is invariant under left multiplication of all $d \times d$ real matrices. Since left multiplication is row operation, there is a linear subspace $\boldsymbol{U} \subset \mathbb{R}^k$ such that $\boldsymbol{L} = L(\boldsymbol{U})$. ☐

Let $\pi_1 : G \to G'$ be the natural projection map and $\Gamma_L = L(\boldsymbol{U}) \cap \Gamma$.

Lemma 3.3. *Let $\boldsymbol{U}$ be the linear subspace of $\mathbb{R}^k$ obtained by the Lemma 3.2. Then $\boldsymbol{U} \cap \mathbb{Z}^k$ is a lattice in $\boldsymbol{U}$ and $\boldsymbol{\xi} \in (\mathbb{Q}^d)^k + L(\boldsymbol{U})$.*

Proof. By Lemma 3.2, $H = SL_d(\mathbb{R}) \ltimes L(\boldsymbol{U})$ and $H \cdot (Id, \boldsymbol{\xi})\Gamma/\Gamma$ is closed and admits an H-invariant probability measure, therefore $\Gamma_H = (Id, \boldsymbol{\xi})\Gamma(Id, -\boldsymbol{\xi}) \cap H$ is a lattice in H.

By [18, Corollary 8.28], Γ_L is a lattice in $L(\boldsymbol{U})$, that is, $(\mathbb{Z}^d)^k \cap L(\boldsymbol{U})$ is a lattice in $L(\boldsymbol{U})$. Thus $\boldsymbol{U}$ has a basis belonging to $\mathbb{Z}^k$, and it follows that $\mathbb{Z}^k \cap \boldsymbol{U}$ is a lattice in $\boldsymbol{U}$.

Recall that $\Gamma' = SL_d(\mathbb{Z})$. Now consider $\pi_1(\Gamma_H) = \{\gamma \in \Gamma' : \boldsymbol{\xi} - \gamma \cdot \boldsymbol{\xi} \in (\mathbb{Z}^d)^k + L(\boldsymbol{U})\}$. Again by [18, Corollary 8.28], $\pi_1(\Gamma_H)$ is a lattice in G'. Therefore $\pi_1(\Gamma_H)$ is a finite index subgroup of Γ'. Pick a $\gamma \in \pi_1(\Gamma_H)$ such that $Id - \gamma$ is invertible, then $\boldsymbol{\xi} \in (\mathbb{Q}^d)^k + L(\boldsymbol{U})$. ☐

Lemma 3.4. *Let $\boldsymbol{U}$ be the linear subspace of $\mathbb{R}^k$ obtained by Lemma 3.2. If for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, $\boldsymbol{\xi} \cdot \mathbf{m} \notin \mathbb{Z}^d$. Then $\boldsymbol{U} = \mathbb{R}^k$ and hence, $H = G$.*

Proof. Suppose $\boldsymbol{U} \neq \mathbb{R}^k$, then $\dim \boldsymbol{U} < k$. Since $\boldsymbol{U} \cap \mathbb{Z}^k$ is a lattice in $\boldsymbol{U}$, there exists a nonzero $\mathbf{v} \in \mathbb{Z}^k \cap \boldsymbol{U}^\perp$. Since $\boldsymbol{\xi} \cdot \mathbf{v} \in (\mathbb{Q}^d)^k \cdot \mathbf{v} + L(\boldsymbol{U}) \cdot \mathbf{v} = (\mathbb{Q}^d)^k \cdot \mathbf{v}$, we can choose $\mathbf{m}$ to be a suitable integral multiple of $\mathbf{v}$ such that $\boldsymbol{\xi} \cdot \mathbf{m} \in \mathbb{Z}^d$, this contradicts to the assumption of the lemma. ☐

4. Unipotent invariance

The collection of all probability measures on the one point compactification X^* of X is a compact space in weak*-topology. Therefore, for any $\mathbf{s} \in \mathcal{U}$, after possibly passing to a subsequence, we have

$$\frac{1}{T}\int_0^T \delta_{a_t u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma} dt \xrightarrow{T\to\infty} \mu_{\mathbf{s}} \text{ in weak* topology,}$$

for some probability measure $\mu_{\mathbf{s}}$ on X^*. Throughout this section, the function $\boldsymbol{\varphi}$ is assumed to be C^1 and satisfy $(\boldsymbol{\varphi}(\mathbf{s}))_{>r} \equiv \mathbf{0}$.

Proposition 4.1. *For a.e. $\mathbf{s} \in \mathcal{U}$, $\mu_{\mathbf{s}}$ is U-invariant.*

Proof. Since $\mathcal{U}$ is a bounded open subset of $Mat_{r\times(d-r)}(\mathbb{R})$, it is enough to prove the proposition for a.e. $\mathbf{s}$ in an open cube of $\mathcal{U}$.

We may choose an open interval $\mathbb{I}\subset\mathbb{R}$ such that $\mathbb{I}^{r(d-r)}\subset\mathcal{U}$. For $1\le i\le r, 1\le j\le d-r$, let $\mathbf{E}_{ij}\in Mat_{r\times(d-r)}(\mathbb{R})$ be the matrix with 1 in (i,j)-th entry and zero otherwise.

If s_1, s_2 are two real numbers linearly independent over $\mathbb{Q}$, then the closure of the subgroup generated by $\{u(s_1\mathbf{E}_{ij}), u(s_2\mathbf{E}_{ij}) : 1\le i\le r, 1\le j\le d-r\}$ is U.

Therefore, given $s'\in\mathbb{R}$, without loss of generality, it suffices to prove that for a.e. $\mathbf{s}\in\mathbb{I}^{r(d-r)}$, the limit measure $\mu_{\mathbf{s}}$ is invariant under $u(s'\mathbf{E}_{11})$.

Note that there exists a countable dense subset of $C_c(G/\Gamma)$ consisting of smooth functions. Let $\psi\in C_c^\infty(G/\Gamma)$. For $t>0$ and $\mathbf{w}\in Mat_{r\times(d-r)}(\mathbb{R})$, define

$$\psi_t(\mathbf{w}) = \psi(a_t u_{\boldsymbol{\varphi}}(\mathbf{w})\Gamma) - \psi(u(s'\mathbf{E}_{11})a_t u_{\boldsymbol{\varphi}}(\mathbf{w})\Gamma).$$

Hence, we only need to show that for this ψ, for a.e. $\mathbf{w}\in\mathbb{I}^{r(d-r)}$,

$$\frac{1}{T}\int_0^T \psi_t(\mathbf{w})dt \xrightarrow{T\to\infty} 0.$$

This follows from Theorem 4.2 and Lemma 4.3 as follows. □

Theorem 4.2. *[13, Theorem 3.1] Let (Y,μ) be a probability space. Let $F : Y\times\mathbb{R}^+\to\mathbb{R}$ be a bounded measurable function. Suppose that there exist $\delta>0$ and $c>0$ such that for any $l\ge t\ge 0$,*

$$|\int_Y F(x,t)F(x,l)d\mu(x)| \le c\cdot e^{-\delta\min(t,l-t)}, \tag{4.1}$$

then given any $\epsilon>0$, for μ-a.e. $y\in Y$,

$$\frac{1}{T}\int_0^T F(y,t)dt = o(T^{-\frac{1}{2}}\cdot log^{\frac{2}{3}+\epsilon}T).$$

Lemma 4.3. *There exist $c>0$ such that for any $t,l>0$,*

$$|\int_{\mathbb{I}^{r(d-r)}} \psi_t(\mathbf{w})\psi_l(\mathbf{w})d\mathbf{w}| \le c\cdot e^{-|l-t|}.$$

Proof. In the following proof, for positive valued functions f,g, we write $f=O(g)$ if there exists a positive constant C depending only on $\psi, \boldsymbol{\varphi}, s'$ and $|\mathbb{I}|$ (these are fixed throughout the proof) such that $f\le Cg$. Also, for any positive number $\epsilon > 0$, we let $O_G(\epsilon)$ denote a group element in a $O(\epsilon)$-neighborhood of Id in G.

Without loss of generality, we assume that $l \geq t$. For $s_0 \in \mathbb{I}$, consider the interval

$$\mathbb{I}(s_0) = (s_0 - |\mathbb{I}|e^{-\frac{d(l+t)}{2}}, s_0 + |\mathbb{I}|e^{-\frac{d(l+t)}{2}})$$

such that $\mathbb{I}(s_0) \subset \mathbb{I}$. For any $s \in \mathbb{I}(s_0)$, and any $\mathbf{w} \in \{0\} \times \mathbb{I}^{r(d-r)-1}$,

$$\begin{aligned}&|\psi_t(s\mathbf{E}_{11} + \mathbf{w}) - \psi_t(s_0\mathbf{E}_{11} + \mathbf{w})| \leq |\psi(a_t u_{\boldsymbol{\varphi}}(s\mathbf{E}_{11} + \mathbf{w})\Gamma)\\&- \psi(a_t u_{\boldsymbol{\varphi}}(s_0\mathbf{E}_{11} + \mathbf{w})\Gamma)| + |\psi(u(s'\mathbf{E_{11}})a_t u_{\boldsymbol{\varphi}}(s\mathbf{E_{11}} + \mathbf{w})\Gamma)\\&- \psi(u(s'\mathbf{E_{11}})a_t u_{\boldsymbol{\varphi}}(s_0\mathbf{E_{11}} + \mathbf{w})\Gamma)|.\end{aligned}$$

Note that

$$\begin{aligned}&a_t u_{\boldsymbol{\varphi}}(s\mathbf{E_{11}} + \mathbf{w}) = a_t(u(s\mathbf{E}_{11} + \mathbf{w}), \boldsymbol{\varphi}(s\mathbf{E}_{11} + \mathbf{w}))\\&= a_t(u(s_0\mathbf{E}_{11} + \mathbf{w} + (s - s_0)\mathbf{E}_{11}), \boldsymbol{\varphi}(s_0\mathbf{E}_{11} + \mathbf{w})\\&\quad + \boldsymbol{\varphi}(s\mathbf{E}_{11} + \mathbf{w}) - \boldsymbol{\varphi}(s_0\mathbf{E}_{11} + \mathbf{w}))\\&= (u(e^{dt}(s - s_0)\mathbf{E_{11}}), e^{(d-r)t}(s - s_0)\partial_{11}\boldsymbol{\varphi})a_t u_{\boldsymbol{\varphi}}(s_0\mathbf{E_{11}} + \mathbf{w}).\end{aligned}$$

where the last equality follows by mean value theorem (for simplicity of notations, by uniform boundedness of $||\partial_{11}\varphi||_\infty$ on $\mathcal{U}$, we write $\partial_{11}\boldsymbol{\varphi}$ for $\partial_{11}\boldsymbol{\varphi}(\tilde{s}\mathbf{E_{11}} + \mathbf{w})$ with arbitrary $\tilde{s}$). Since

$$e^{(d-r)t}|s - s_0| \leq e^{dt}|s - s_0| \leq e^{-\frac{d(l+t)}{2}} \cdot e^{dt}|\mathbb{I}| = e^{-\frac{d(l-t)}{2}}|\mathbb{I}|,$$

we have

$$\begin{aligned}a_t u_{\boldsymbol{\varphi}}(s\mathbf{E_{11}} + \mathbf{w}) &= O_G(e^{-\frac{d(l-t)}{2}}|\mathbb{I}|)a_t u_{\boldsymbol{\varphi}}(s_0\mathbf{E_{11}} + \mathbf{w})\\&= O_G(e^{-\frac{d(l-t)}{2}})a_t u_{\boldsymbol{\varphi}}(s_0\mathbf{E_{11}} + \mathbf{w}).\end{aligned}$$

Likewise,

$$u(s'\mathbf{E_{11}})a_t u_{\boldsymbol{\varphi}}(s\mathbf{E_{11}} + \mathbf{w}) = O_G(e^{-\frac{d(l-t)}{2}})u(s'\mathbf{E_{11}})a_t u_{\boldsymbol{\varphi}}(s_0\mathbf{E_{11}} + \mathbf{w}).$$

Since $\psi \in C_c^\infty(G/\Gamma)$, ψ is Lipschitz, and hence

$$|\psi_t(s\mathbf{E_{11}} + \mathbf{w}) - \psi_t(s_0\mathbf{E_{11}} + \mathbf{w})| = O(e^{-\frac{d(l-t)}{2}}).$$

Therefore,

$$\begin{aligned}
&\int_{\mathbb{I}(s_0)} \psi_t(s\mathbf{E_{11}}+\mathbf{w})\psi_l(s\mathbf{E_{11}}+\mathbf{w})ds\\
&= \int_{\mathbb{I}(s_0)} (\psi_t(s_0\mathbf{E_{11}}+\mathbf{w}) + O(e^{-\frac{d(l-t)}{2}}))\cdot \psi_l(s\mathbf{E_{11}}+\mathbf{w})ds\\
&= \psi_t(s_0\mathbf{E_{11}}+\mathbf{w}) \int_{\mathbb{I}(s_0)} \psi_l(s\mathbf{E_{11}}+\mathbf{w})ds + O(e^{-\frac{d(l-t)}{2}})\cdot |\mathbb{I}(s_0)|. \qquad (4.2)
\end{aligned}$$

Now we estimate $\int_{\mathbb{I}(s_0)} \psi_l(s\mathbf{E_{11}}+\mathbf{w})ds$. Note that

$$\begin{aligned}
&u(s'\mathbf{E}_{11})a_l u_{\boldsymbol{\varphi}}(s\mathbf{E_{11}}+\mathbf{w})\\
&= a_l(u((s+e^{-dl}s')\mathbf{E_{11}}+\mathbf{w}), \boldsymbol{\varphi}(s\mathbf{E_{11}}+\mathbf{w}))\\
&= a_l(Id, \boldsymbol{\varphi}(s\mathbf{E_{11}}+\mathbf{w}) - \boldsymbol{\varphi}((e^{-dl}s'+s)\mathbf{E_{11}}+\mathbf{w}))\\
&\quad \cdot a_{-l}a_l u_{\boldsymbol{\varphi}}((s+e^{-dl}s')\mathbf{E_{11}}+\mathbf{w})\\
&= a_l(Id, e^{-dl}s'\partial_{11}\boldsymbol{\varphi})a_{-l}a_l u_{\boldsymbol{\varphi}}((s+e^{-dl}s')\mathbf{E_{11}}+\mathbf{w})\\
&= O_G(e^{-rl})a_l u_{\boldsymbol{\varphi}}((s+e^{-dl}s')\mathbf{E_{11}}+\mathbf{w}).
\end{aligned}$$

As ψ is Lipschitz,

$$\begin{aligned}
&\psi_l(s\mathbf{E_{11}}+\mathbf{w})\\
&= \psi(a_l u_{\boldsymbol{\varphi}}(s\mathbf{E_{11}}+\mathbf{w})\Gamma) - \psi(a_l u_{\boldsymbol{\varphi}}((s+e^{-dl}s')\mathbf{E_{11}}+\mathbf{w})\Gamma) + O(e^{-rl}).
\end{aligned}$$

Since $\mathbb{I}(s_0)$ and $\mathbb{I}(s_0)+e^{-dl}s'$ overlap except for a length of $O(e^{-dl})$, we have

$$\begin{aligned}
&\int_{\mathbb{I}(s_0)} \psi_l(s\mathbf{E_{11}}+\mathbf{w})ds\\
&= \int_{\mathbb{I}(s_0)} \psi(a_l u_{\boldsymbol{\varphi}}(s\mathbf{E_{11}}+\mathbf{w})\Gamma) - \psi(a_l u_{\boldsymbol{\varphi}}((s+e^{-dl}s')\mathbf{E_{11}}+\mathbf{w})\Gamma) + O(e^{-rl})ds\\
&= O(e^{-dl}) + O(e^{-rl})|\mathbb{I}(s_0)|\\
&= O(e^{-\frac{d(l-t)}{2}})|\mathbb{I}(s_0)| + O(e^{-(l-t)})|\mathbb{I}(s_0)|\\
&= O(e^{-(l-t)})|\mathbb{I}(s_0)|.
\end{aligned}$$

Now we consider the partition $\mathbb{I} = \bigcup_{j=1}^{p} \mathbb{I}_j$ such that $\mathbb{I}_j = [s_{j-1}, s_j]$ with $s_j - s_{j-1} = 2e^{-\frac{d(l+t)}{2}}|\mathbb{I}|$ for $1 \le j \le p-1$, and $s_p - s_{p-1} \le 2e^{-\frac{d(l+t)}{2}}|\mathbb{I}|$. By (4.2), we have

$$\int_{\mathbb{I}} \psi_t(s\mathbf{E_{11}}+\mathbf{w})\psi_l(s\mathbf{E_{11}}+\mathbf{w})ds$$

$$= \sum_{j=1}^{p} \int_{\mathbb{I}_j} \psi_t(s\mathbf{E_{11}} + \mathbf{w})\psi_l(s\mathbf{e_{11}} + \mathbf{w})ds$$
$$= \sum_{j=1}^{p-1} \int_{\mathbb{I}_j} \psi_t(s\mathbf{E_{11}} + \mathbf{w})\psi_l(s\mathbf{E_{11}} + \mathbf{w})ds + O(e^{-\frac{d(l+t)}{2}})|\mathbb{I}|$$
$$= \sum_{j=1}^{p-1} O(e^{-(l-t)})|\mathbb{I}_j| + O(e^{-\frac{d(l+t)}{2}})|\mathbb{I}|$$
$$= O(e^{-(l-t)})|\mathbb{I}| = O(e^{-(l-t)}).$$

The above estimate holds for any $\mathbf{w} \in \{\mathbf{0}\} \times \mathbb{I}^{r(d-r)-1}$. Now the lemma follows from the above estimate and Fubini's theorem. □

5. Margulis' height function

In this section, we will recall the definition of Margulis' height function on $SL_d(\mathbb{R})/SL_d(\mathbb{Z})$ and its uniform contraction property.

Margulis' height function was first introduced in [6] and later developed in several papers (see for example [1][25]). It measures the depth of elements of X into the cusps. It has been used to study equidistribution problem for certain unbounded functions (cf. [6][7][16]) and random walks on homogeneous spaces (cf. [1][5]).

We start with the vector space $\mathbf{V} = \wedge^*\mathbb{R}^d = \bigoplus_{0\le i\le d} \wedge^i\mathbb{R}^d$, where $G' = SL_d(\mathbb{R})$ acts on $\mathbf{V}$ naturally.

Let Δ be a lattice in $\mathbb{R}^d$. We say that a subspace L of $\mathbb{R}^d$ is Δ-rational if $L \cap \Delta$ is a lattice in L. For any Δ-rational subspace L, denote $d(L)$ or $d_\Delta(L)$ the volume of $L/L\cap\Delta$. Note that $d(L)$ is the norm of $u_1 \wedge u_2 \wedge \cdots \wedge u_l$ in $\mathbf{V}$, where $\{u_i\}_{1\le i\le l}$ is a $\mathbb{Z}$-basis of $L\cap\Delta$. If $L = \{\mathbf{0}\}$, we set $d(L) = 1$.

For any lattice Δ, we define for $0 \le i \le d$,

$$\alpha_i(\Delta) := \sup\left\{\frac{1}{d(L)} : L \text{ is a } \Delta\text{-rational subspace of dimension } i\right\}.$$

Proposition 5.1. *There exists a continuous map $\tilde{\alpha} : SL_d(\mathbb{R})/SL_d(\mathbb{Z}) \to [1,\infty]$ and $b_1 > 0$ such that for B a bounded open box in $Mat_{r\times(d-r)}(\mathbb{R})$, for all $t > 0$ large enough and for any unimodular lattice Λ of $\mathbb{R}^d$,*

$$\frac{1}{|B|}\int_B \tilde{\alpha}(a_t u(\mathbf{s})\Lambda)d\mathbf{s} < 2^{-r(d-r)-2}\tilde{\alpha}(\Lambda) + b_1,$$

and there exists $\nu > 0$ *such that*

$$\alpha_1(\Lambda)^\nu \leq \tilde{\alpha}(\Lambda).$$

Moreover, a measurable subset K *of* $SL_d(\mathbb{R})/SL_d(\mathbb{Z})$ *is precompact if there exists* $N > 0$ *such that*

$$K \subset \{x \in X : \tilde{\alpha}(x) \leq N\}.$$

Proof. Define $\tilde{\alpha} = \epsilon^{-q(1)} \cdot \sum_{i=0}^{d} \epsilon^{q(i)} \cdot \alpha_i^\nu$, where $\epsilon, \nu > 0$ are sufficiently small numbers and $q(i) = i(d-i)$. The fact that $\tilde{\alpha}$ satisfies conclusion of Proposition 5.1 will follow from [25, Lemma 4.1]. □

The function $\tilde{\alpha}$ above is the Margulis' height function that we need in our setting.

Remark 5.2. The function $\tilde{\alpha}$ satisfies Lipschitz property as follows: For any bounded neighborhood $\mathcal{V}$ of e of $SL_d(\mathbb{R})$, there exists $\overline{M} > 0$ such that for any $x \in SL_d(\mathbb{R})/SL_d(\mathbb{Z})$, any $g \in \mathcal{V}$,

$$\tilde{\alpha}(gx) \leq \overline{M}\tilde{\alpha}(x).$$

Indeed, $\overline{M} > 0$ is the maximum of operator norms of elements in $\mathcal{V}$ acting on $\mathbf{V}$.

6. Mixed height function

In this section, we will construct a mixed height function, which is crucial for us to prove Proposition 2.1. The main result of this section is the following:

Proposition 6.1. *Let* $\boldsymbol{\varphi}$ *be a* C^1 *map from* $\mathcal{U}$ *to* $(\mathbb{R}^d)^k$ *satisfying* $(\boldsymbol{\varphi}(\mathbf{s}))_{>r} \equiv \mathbf{0}$ *for any* $\mathbf{s} \in \mathcal{U}$. *For any* $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, *any closed cube* $I \subset \mathcal{U} \setminus Bad_{\mathbf{m}}$ *(for* $Bad_{\mathbf{m}}$, *see (2.2)), there are* $t > 0$ *sufficiently large (depending on* I, $\mathbf{m}$*) and measurable function* $\beta_{\mathbf{m}} : G/\Gamma \to (0, \infty]$ *such that the following hold:*

(1) For any $l > 0$, $\{x \in G/\Gamma : \beta_{\mathbf{m}}(x) \leq l\}$ *is compact;*

(2) For any $x \in G/\Gamma$, $\beta_{\mathbf{m}}(x) = \infty$ *if and only if* $x \in X_{\mathbf{m}}$*;*

(3) Given any $n \in \mathbb{Z}_{\geq 0}$, *a box* $J \subset I$ *with* $J = \prod_{i=1}^{r(d-r)} J_i$, *where* $J_i \subset \mathbb{R}$ *and* $|J_i| \leq 2e^{-dnt}$ *for all* i. *There exists* $\tilde{M}_1 > 0$ *such that for any* $\mathbf{s}, \tilde{\mathbf{s}} \in J$, *one has*

$$\beta_{\mathbf{m}}(a_{nt}u_{\boldsymbol{\varphi}}(\tilde{\mathbf{s}})\Gamma) \leq \tilde{M}_1\beta_{\mathbf{m}}(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma);$$

(4) There exists $\tilde{M}_2 > 0$ depending on t, for any $n \in \mathbb{Z}_{\geq 0}$, any $\mathbf{s} \in I$ and any $\tau \in \mathbb{R}$ with $|\tau| \leq t$, one has

$$\beta_{\mathbf{m}}(a_\tau a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) \leq \tilde{M}_2 \beta_{\mathbf{m}}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma);$$

(5) There exists $b > 0$ such that the following holds: for any $n \in \mathbb{Z}_{\geq 0}$ and any box $J \subset I$ with $J = \prod_{i=1}^{r(d-r)} J_i$ satisfying either $n \geq 1$ and $|J_i| \geq e^{-dnt}$ for all $1 \leq i \leq r(d-r)$, or $n = 0$ and $J = I$, one has

$$\int_J \beta_{\mathbf{m}}(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) d\mathbf{s} \leq \frac{1}{2} \int_J \beta_{\mathbf{m}}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) d\mathbf{s} + b|J|.$$

Remark 6.2. The function $\beta_{\mathbf{m}}$ in Proposition 6.1 is the desired mixed height function.

From now on until the end of this section, we will fix a closed cube $I \subset \mathcal{U} \setminus Bad_{\mathbf{m}}$. Recall that the finite number M_1 is defined as in (2.1). By the choice of I, and the fact that there are only finitely many $\mathbf{a} \in \mathbb{Z}^{d-r}$ satisfying $\|\mathbf{a}\|_\infty \leq M_1 \|\mathbf{m}\|$, we obtain $\sigma > 0$ such that

(6.1)
$$\inf_{\mathbf{s} \in I} \{ \|(\boldsymbol{\varphi}(\mathbf{s}))_{\leq r} \cdot \mathbf{m} - \mathbf{s} \cdot \mathbf{a} - \mathbf{b}\|_\infty : \|\mathbf{a}\|_\infty \leq M_1 \|\mathbf{m}\|, \mathbf{a} \in \mathbb{Z}^{d-r}, \mathbf{b} \in \mathbb{Z}^r \} = \sigma.$$

Remark 6.3. By (6.1), we can choose a closed neighborhood I' of I such that I' is a closed cube contained in $\mathcal{U}$ and satisfies

$$\inf_{\mathbf{s} \in I'} \{ \|(\boldsymbol{\varphi}(\mathbf{s}))_{\leq r} \cdot \mathbf{m} - \mathbf{s} \cdot \mathbf{a} - \mathbf{b}\|_\infty : \|\mathbf{a}\|_\infty \leq M_1 \|\mathbf{m}\|, \mathbf{a} \in \mathbb{Z}^{d-r}, \mathbf{b} \in \mathbb{Z}^r \} = \frac{\sigma}{2}.$$

Next we construct a suitable function measuring the distance to the closed submanifold $X_{\mathbf{m}}$. For $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$, consider the quotient space

$$(\mathbb{R}^d)^k_{\mathbf{m}} := \{\mathbf{v} \in (\mathbb{R}^d)^k : \mathbf{v} \cdot \mathbf{m} \in \mathbb{Z}^d\}/\sim,$$

where $\mathbf{v} \sim \mathbf{v}'$ if and only if $\mathbf{v} \cdot \mathbf{m} = \mathbf{v}' \cdot \mathbf{m}$. One can directly verify that $\sim$ is an equivalence relation.

Lemma 6.4. *For any $(g, \mathbf{v}) \in G$, there exists at most one $\mathbf{v}_0 \in (\mathbb{R}^d)^k_{\mathbf{m}}$ such that*

$$\|(\mathbf{v} - g\mathbf{v}_0) \cdot \mathbf{m}\| < \frac{1}{2} \inf_{\mathbf{w} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}} \|g\mathbf{w}\|. \tag{6.2}$$

Proof. Suppose there are two vectors $\mathbf{v}_0$ and $\mathbf{v}_0'$ in $(\mathbb{R}^d)^k_{\mathbf{m}}$ satisfying (6.2) such that $\mathbf{v}_0 \not\sim \mathbf{v}_0'$. Then

$$\|g(\mathbf{v}_0 - \mathbf{v}_0') \cdot \mathbf{m}\| \leq \|(\mathbf{v} - g\mathbf{v}_0) \cdot \mathbf{m}\| + \|(\mathbf{v} - g\mathbf{v}_0') \cdot \mathbf{m}\| < \inf_{\mathbf{w} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}} \|g\mathbf{w}\| .$$

But since $\mathbf{v}_0 \not\sim \mathbf{v}_0'$, $(\mathbf{v}_0 - \mathbf{v}_0') \cdot \mathbf{m} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}$. This is a contradiction. □

Definition 6.5. Let $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$. For any $(g, \mathbf{v}) \in G$, we say that $\boldsymbol{\xi}_{g,v} \in (\mathbb{R}^d)^k_{\mathbf{m}}$ exists if $\boldsymbol{\xi}_{g,v}$ satisfies (6.2) in the place of $\mathbf{v}_0$. By convention, we set $\boldsymbol{\xi}_{g,\mathbf{v}} = \infty$ if it does not exist. Define the function

$$\alpha_{\mathbf{m}}(g, \mathbf{v}) = \begin{cases} \left\|(\mathbf{v} - g\boldsymbol{\xi}_{g,\mathbf{v}}) \cdot \mathbf{m}\right\|^{-1}, & \text{If } \boldsymbol{\xi}_{g,\mathbf{v}} \text{ exists;} \\ 1, & \text{if } \boldsymbol{\xi}_{g,\mathbf{v}} = \infty. \end{cases} \tag{6.3}$$

Remark 6.6. By Lemma 6.4, $\alpha_{\mathbf{m}}$ is a well-defined function on G. By Minkowski's first theorem, there is a constant $0 < \mu_d \leq 1$ such that if $\boldsymbol{\xi}_{g,\mathbf{v}}$ exists for $(g, \mathbf{v})$, then $\alpha_{\mathbf{m}}(g, \mathbf{v}) > \mu_d$. Therefore, by definition of $\alpha_{\mathbf{m}}$, we have $\alpha_{\mathbf{m}}(g, \mathbf{v}) > \mu_d$ for any $(g, \mathbf{v}) \in G$. Moreover, by (6.1), for any $\mathbf{s} \in I$, $\alpha_{\mathbf{m}}(u_{\boldsymbol{\varphi}}(\mathbf{s})) \leq \sigma^{-1}$.

Lemma 6.7. *$\alpha_{\mathbf{m}}$ is a well-defined function on G/Γ. Moreover, $\alpha_{\mathbf{m}}$ is lower semi-continuous.*

Proof. Take any $(g, \mathbf{v}) \in G$ and any $(\gamma, \mathbf{v}') \in \Gamma$.

If $\boldsymbol{\xi}_{g,\mathbf{v}}$ exists, then $\gamma^{-1}(\boldsymbol{\xi}_{g,\mathbf{v}} + \mathbf{v}') \in (\mathbb{R}^d)^k_{\mathbf{m}}$. Note that

$$\left\|(\mathbf{v} + g\mathbf{v}' - g\gamma \cdot \gamma^{-1}(\boldsymbol{\xi}_{g,\mathbf{v}} + \mathbf{v}')) \cdot \mathbf{m}\right\| = \left\|(\mathbf{v} - g\boldsymbol{\xi}_{g,\mathbf{v}}) \cdot \mathbf{m}\right\| < \frac{1}{2} \inf_{\mathbf{w} \in \mathbb{Z}^d \setminus \{\mathbf{0}\}} \|g\mathbf{w}\| .$$

Thus $\boldsymbol{\xi}_{g\gamma,\mathbf{v}+g\mathbf{v}'} = \gamma(\boldsymbol{\xi}_{g,\mathbf{v}} + \mathbf{v}')$ exists and $\alpha_{\mathbf{m}}(g, \mathbf{v}) = \alpha_{\mathbf{m}}(g\gamma, \mathbf{v} + g\mathbf{v}')$.

If $\boldsymbol{\xi}_{g,\mathbf{v}}$ does not exist, same argument as above shows that $\boldsymbol{\xi}_{g\gamma,\mathbf{v}+g\mathbf{v}'}$ does not exist neither.

If $\boldsymbol{\xi}_{g,\mathbf{v}}$ exists, it is locally constant. Therefore, $\alpha_{\mathbf{m}}$ is lower semi-continuous. □

Let $\nu \in (0, \frac{1}{r(d-r)})$ be a number satisfying Proposition 5.1. Let $c = 4 \cdot (10r^2d)^\nu \cdot 2^{r(d-r)}$ and $t > 0$ be a sufficiently large number (to be specified later). Define

$$\beta_{\mathbf{m}} = \alpha_{\mathbf{m}}^\nu + ce^{\nu rt}\tilde{\alpha}. \tag{6.4}$$

We will prove that $\beta_{\mathbf{m}}$ satisfies properties (1)-(5) of Proposition 6.1.

Proof of Proposition 6.1 (1). Since

$$\{x \in X : \beta_{\mathbf{m}}(x) \leq l\} \subset \{x \in X : \tilde{\alpha}(x) \leq lc^{-1}e^{-\nu rt}\},$$

$\{x \in X : \beta_{\mathbf{m}}(x) \leq l\}$ is precompact. As $\tilde{\alpha}$ is continuous and $\alpha_{\mathbf{m}}$ is lower semicontinuous, $\{x \in X : \beta_{\mathbf{m}}(x) \leq l\}$ is closed and thus compact. □

Proof of Proposition 6.1 (2). If for some $x \in G/\Gamma$, $\beta_{\mathbf{m}}(x) = \infty$, then $\alpha_{\mathbf{m}}(x) = \infty$ as $\tilde{\alpha}(x) < \infty$. Let $x = (g, \mathbf{v})\Gamma/\Gamma$. By definition of $\alpha_{\mathbf{m}}$, we have $(\mathbf{v} - g\boldsymbol{\xi}) \cdot \mathbf{m} = \mathbf{0}$ for some $\boldsymbol{\xi} \in (\mathbb{R}^d)^k_{\mathbf{m}}$. Note that $g^{-1}\mathbf{v} \cdot \mathbf{m} = \boldsymbol{\xi} \cdot \mathbf{m} \in \mathbb{Z}^d$. Hence $(g, \mathbf{v})\Gamma = (g, gg^{-1}\mathbf{v})\Gamma \in X_{\mathbf{m}}$.

Conversely, if $x = (g, \mathbf{v})\Gamma/\Gamma \in X_{\mathbf{m}}$, then by definition of $X_{\mathbf{m}}$, we have $g^{-1}\mathbf{v} \cdot \mathbf{m} \in \mathbb{Z}^d$. Choose any $\boldsymbol{\xi} \in (\mathbb{R}^d)^k$ such that $g^{-1}\mathbf{v} \cdot \mathbf{m} = \boldsymbol{\xi} \cdot \mathbf{m}$, then $(\mathbf{v} - g\boldsymbol{\xi}) \cdot \mathbf{m} = g(g^{-1}\mathbf{v} \cdot \mathbf{m} - \boldsymbol{\xi} \cdot \mathbf{m}) = \mathbf{0}$. Therefore, $\alpha_{\mathbf{m}}(x) = \infty$. □

Notations. Let's fix some simplified notations for the rest of the proof. We will fix an $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$ till the end of this section. In the following, $t > 0$ is a sufficiently large number.

- For any $n \in \mathbb{N}$, any $\mathbf{s} \in \mathcal{U}$, denote $a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s}) = (g_n(\mathbf{s}), \mathbf{v}_n(\mathbf{s}))$.
- If $\boldsymbol{\xi}_{g_n(\mathbf{s}),\mathbf{v}_n(\mathbf{s})}$ exists, denote $\boldsymbol{\xi}_{g_n(\mathbf{s}),\mathbf{v}_n(\mathbf{s})} = \boldsymbol{\xi}_{n,\mathbf{s}}$.
- For any $\mathbf{v} \in (\mathbb{R}^d)^k_{\mathbf{m}}$, $n \in \mathbb{N}$, $\mathbf{s} \in \mathcal{U}$, let

$$\begin{aligned} w(n, \mathbf{s}, \mathbf{v}) &= (\mathbf{v}_n(\mathbf{s}) - g_n(\mathbf{s})\mathbf{v}) \cdot \mathbf{m} \\ &= \begin{bmatrix} e^{(d-r)nt}[(\boldsymbol{\varphi}(\mathbf{s}))_{\leq r} - (\mathbf{v})_{\leq r} - \mathbf{s} \cdot (\mathbf{v})_{>r}] \cdot \mathbf{m} \\ e^{-rnt}(\mathbf{v})_{>r} \cdot \mathbf{m} \end{bmatrix} = \begin{bmatrix} w_1(n, \mathbf{s}, \mathbf{v}) \\ \vdots \\ w_d(n, \mathbf{s}, \mathbf{v}) \end{bmatrix}. \end{aligned}$$

  We note that if $\boldsymbol{\xi}_{n,\mathbf{s}}$ exists, then $\alpha_{\mathbf{m}}(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})) = \left\| w(n, \mathbf{s}, \boldsymbol{\xi}_{n,\mathbf{s}}) \right\|^{-1}$.
- For any differentiable function $\psi : Mat_{r\times(d-r)}(\mathbb{R}) \to \mathbb{R}$, by mean value theorem in several variables, for any $\mathbf{s}, \tilde{\mathbf{s}} \in Mat_{r\times(d-r)}(\mathbb{R})$, there is a $\hat{\mathbf{s}}$ such that

$$\psi(\tilde{\mathbf{s}}) - \psi(\mathbf{s}) = \sum_{i=1}^{r} \sum_{j=1}^{d-r} \frac{\partial \psi}{\partial s_{ij}}(\hat{\mathbf{s}}) \cdot (\tilde{s}_{ij} - s_{ij}).$$

  Since the functions that we consider have bounded first derivative on a bounded set, we will omit this $\hat{\mathbf{s}}$ for simplicity.
- For $1 \leq i \leq d - r$, let $\mathbf{e}_i$ denote the column vector in $\mathbb{R}^{d-r}$ with 1 in i-th row and 0 elsewhere. Let $<, >$ denote the usual inner product of column vectors.

Lemma 6.8. *Let $n \geq 1$ be an integer and $t > log(2\mu_d^{-1}\sigma^{-1})$, where $\mu_d > 0$ is the constant given as in Remark 6.6. For any $\mathbf{s} \in I'$, where I' is given as in Remark 6.3, if $\boldsymbol{\xi}_{n,\mathbf{s}}$ exists, then $\left\|(\boldsymbol{\xi}_{n,\mathbf{s}})_{>r} \cdot \mathbf{m}\right\|_\infty > M_1 \|\mathbf{m}\|$.*

Proof. Suppose $\left\|(\boldsymbol{\xi}_{n,\mathbf{s}})_{>r} \cdot \mathbf{m}\right\|_\infty \leq M_1 \|\mathbf{m}\|$. By definition of I',

$$\begin{aligned} \left\|(\mathbf{v}_n(\mathbf{s}) - g_n(\mathbf{s})\boldsymbol{\xi}_{n,\mathbf{s}}) \cdot \mathbf{m}\right\| &\geq \left\|(\mathbf{v}_n(\mathbf{s}) - g_n(\mathbf{s})\boldsymbol{\xi}_{n,\mathbf{s}})_{\leq r} \cdot \mathbf{m}\right\|_\infty \\ &\geq e^{(d-r)nt} \left\|[(\boldsymbol{\varphi}(\mathbf{s}))_{\leq r} - (\boldsymbol{\xi}_{n,\mathbf{s}})_{\leq r} - \mathbf{s}(\boldsymbol{\xi}_{n,\mathbf{s}})_{>r}] \cdot \mathbf{m}\right\|_\infty \\ &\geq e^t \frac{\sigma}{2} > \mu_d^{-1}. \end{aligned}$$

By Remark 6.6, this contradicts the existence of $\boldsymbol{\xi}_{n,\mathbf{s}}$. □

Proof of Proposition 6.1 (3). If $n = 0$, by Remark 6.6, we have $\mu_d^\nu \leq \alpha_{\mathbf{m}}^\nu(u_{\boldsymbol{\varphi}}(\tilde{\mathbf{s}})) \leq \sigma^{-\nu}$. Since $\tilde{\alpha}$ is continuous and bounded on compact sets, there exists $0 < m < M$ such that for any $\mathbf{s} \in I$,

$$m \leq \tilde{\alpha}(u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) \leq M.$$

Therefore,

$$\beta_{\mathbf{m}}(u_{\boldsymbol{\varphi}}(\tilde{\mathbf{s}})\Gamma) \leq \sigma^{-\nu} + ce^{\nu rt} \cdot M \leq \frac{\sigma^{-\nu} + ce^{\nu rt}M}{\mu_d^\nu + ce^{\nu rt}m} \beta_{\mathbf{m}}(u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma).$$

Assume $n \geq 1$. If $\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}$ does not exist, then by definition, $\alpha_{\mathbf{m}}^\nu(a_{nt}u_{\boldsymbol{\varphi}}(\tilde{\mathbf{s}})\Gamma) = 1$. Now we suppose that $\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}$ exists.

Case 1: $\left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{\leq r}\right\|_\infty \geq N_2 \left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{>r}\right\|_\infty$, where $N_2 = 2(r+1)(d-r)$.

Choose $t > log(2\mu_d^{-1}\sigma^{-1})$, by Lemma 6.8, the Lipschitz continuity of $\boldsymbol{\varphi}$, and the choices of M_1 and J, we have

$$\begin{aligned} \left\|(w(n,\tilde{\mathbf{s}},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{\leq r} - (w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{\leq r}\right\|_\infty &\leq 2e^{-rnt}(r+1)(d-r) \left\|(\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}})_{>r} \cdot \mathbf{m}\right\|_\infty \\ &= 2(r+1)(d-r) \left\|(w(n,\tilde{\mathbf{s}},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{>r}\right\|_\infty. \end{aligned}$$

Hence, by the assumption of Case 1,

$$\begin{aligned} \left\|w(n,\tilde{\mathbf{s}},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}})\right\| &\geq \left\|(w(n,\tilde{\mathbf{s}},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{\leq r}\right\|_\infty \\ &\geq \left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{\leq r}\right\|_\infty - 2(r+1)(d-r) \left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{>r}\right\|_\infty \end{aligned}$$

$$\geq \frac{1}{d^2}\left\|w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}})\right\|.$$

Case 2: $\left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{\leq r}\right\|_\infty < N_2\left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{>r}\right\|_\infty$. Then

$$\begin{aligned}\left\|w(n,\tilde{\mathbf{s}},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}})\right\| &\geq \left\|(w(n,\tilde{\mathbf{s}},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{>r}\right\|_\infty = \left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}}))_{>r}\right\|_\infty \\ &\geq \frac{1}{\sqrt{rN_2^2+d-r}}\left\|w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}})\right\| \geq \frac{1}{5dr^2}\left\|w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}})\right\|.\end{aligned}$$

By construction of $\boldsymbol{\xi}_{n,\mathbf{s}}$, we have $\left\|w(n,\mathbf{s},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}})\right\| \geq \left\|w(n,\mathbf{s},\boldsymbol{\xi}_{n,\mathbf{s}})\right\|$. Combining Case 1 and Case 2, we have

$$\left\|w(n,\tilde{\mathbf{s}},\boldsymbol{\xi}_{n,\tilde{\mathbf{s}}})\right\| \geq \min\{\frac{1}{5dr^2},\frac{1}{d^2}\}\left\|w(n,\mathbf{s},\boldsymbol{\xi}_{n,\mathbf{s}})\right\|.$$

Therefore,

$$\alpha^\nu_{\mathbf{m}}(a_{nt}u_{\boldsymbol{\varphi}}(\tilde{\mathbf{s}})\Gamma) \leq \max(d^{2\nu},(5dr^2)^\nu)\cdot\alpha^\nu_{\mathbf{m}}(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma).$$

For $\tilde{\alpha}$, by Remark 5.2, we have for large enough $\overline{M}>0$,

$$\tilde{\alpha}(a_{nt}u_{\boldsymbol{\varphi}}(\tilde{\mathbf{s}})\Gamma) = \tilde{\alpha}(u(e^{dnt}(\tilde{\mathbf{s}}-\mathbf{s}))a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) \leq \overline{M}\tilde{\alpha}(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma).$$

By the above, let $\tilde{M}_1 = 2\max\{\overline{M},d^{2\nu},(5dr^2)^\nu,\frac{\sigma^{-\nu}+ce^{\nu rt}M}{\mu_d^\nu+ce^{\nu rt}m}\}$, then

$$\beta_{\mathbf{m}}(a_{nt}u_{\boldsymbol{\varphi}}(\tilde{\mathbf{s}})\Gamma) \leq \tilde{M}_1\beta_{\mathbf{m}}(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma), \forall n\in\mathbb{Z}_{\geq 0}.$$ □

Proof of Proposition 6.1 (4). By Remark 5.2, since $|\tau|\leq t$, we have

$$\tilde{\alpha}(a_\tau a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) \leq e^{r(d-r)\nu t}\tilde{\alpha}(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma).$$

Let $a_\tau a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s}) = (a_\tau g_n(\mathbf{s}), a_\tau\mathbf{v}_n(\mathbf{s}))$.

If $\boldsymbol{\xi}_{a_\tau g_n(\mathbf{s}),a_\tau\mathbf{v}_n(\mathbf{s})}$ does not exist, then $\alpha_{\mathbf{m}}(a_\tau a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma)=1$.

If $\boldsymbol{\xi}_{a_\tau g_n(\mathbf{s}),a_\tau\mathbf{v}_n(\mathbf{s})}=\mathbf{v}$ exists, then

$$\begin{aligned}\alpha^\nu_{\mathbf{m}}(a_\tau a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) &= \left\|(a_\tau(\mathbf{v}_n(\mathbf{s})-g_n(\mathbf{s})\mathbf{v})\cdot\mathbf{m}\right\|^{-\nu} \\ &\leq e^{(d-r)\nu t}\left\|(\mathbf{v}_n(\mathbf{s})-g_n(\mathbf{s})\mathbf{v})\cdot\mathbf{m}\right\|^{-\nu} \\ &\leq \begin{cases} e^{(d-r)\nu t}\alpha^\nu_{\mathbf{m}}(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) & \text{If } \mathbf{v}=\boldsymbol{\xi}_{g_n(\mathbf{s}),\mathbf{v}_n(\mathbf{s})} \\ e^{(d-r)\nu t}2^\nu\tilde{\alpha}(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) & \text{If } \mathbf{v}\neq\boldsymbol{\xi}_{g_n(\mathbf{s}),\mathbf{v}_n(\mathbf{s})}\end{cases}\end{aligned}$$

Let

$$\tilde{M}_2 = \max\left(e^{(d-r)\nu t}, \frac{2^\nu e^{(d-r)\nu t} + c \cdot e^{r(d-r)\nu t} e^{\nu r t}}{c \cdot e^{\nu r t}}\right),$$

then by the above estimates, Proposition 6.1 (4) is proved. ☐

To prove property (5) of Proposition 6.1, we record the following lemmas:

Lemma 6.9. *[25, Lemma 4.8] Let $n \in \mathbb{Z}_{\geq 0}$ and $t > 0$. Let $I_0 = [-1,1]^{r(d-r)}$, $J = \prod_{i=1}^{r(d-r)} J_i$, where J_i is an interval with $|J_i| \geq e^{-dnt}$ for each i. Let $\Psi : G/\Gamma \to \mathbb{R}_+$ be a measurable function. Then*

$$\int_J \Psi(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) d\mathbf{s} \leq \int_J \int_{I_0} \Psi(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s} + \tilde{\mathbf{s}} e^{-dnt})\Gamma) d\tilde{\mathbf{s}} d\mathbf{s}. \tag{6.5}$$

Lemma 6.10. *Let $\kappa : I_0 = [-1,1]^{r(d-r)} \to \mathbb{R}_+$ be a measurable function. Suppose that there exists $C > 0$ such that for any $\epsilon > 0$,*

$$|\{\mathbf{s} \in I_0 : \kappa(\mathbf{s}) < \epsilon\}| \leq C \cdot \epsilon^{\frac{1}{r(d-r)}}.$$

Then for any $0 < \nu < \frac{1}{r(d-r)}$, there exists $c_\nu > 0$ such that

$$\int_{I_0} \kappa(\mathbf{s})^{-\nu} d\mathbf{s} \leq C^{\nu \cdot r(d-r)} \cdot c_\nu.$$

Proof. This is a direct generalization of [8, Lemma 6.10]. ☐

The following lemma is a special case of [11, Lemma 3.3].

Lemma 6.11. *Let V be a bounded open subset of $Mat_{r\times(d-r)}(\mathbb{R})$, and let $f \in C^1(V)$ be such that for some constants $A_1, A_2 > 0$, one has*

$$A_2 \leq |\partial_{ij} f(\mathbf{s})| \leq A_1, \forall 1 \leq i \leq r, 1 \leq j \leq d-r, \forall \mathbf{s} \in V,$$
$$\text{and } \|f\|_V \leq A_1,$$

where $\|\cdot\|_V$ denote the sup norm of a function on V. Then for any box (or ball) $B \subset V$, any $\epsilon > 0$, one has

$$|\{\mathbf{s} \in B : |f(\mathbf{s})| < \epsilon\}| \leq r(d-r) \cdot C_{A_1,A_2} \left(\frac{\epsilon}{\|f\|_B}\right)^{\frac{1}{r(d-r)}} |B|,$$

with $C_{A_1,A_2} = \frac{12A_1}{A_2}$.

Lemma 6.12. *Let $N \geq 1$ be an integer. Let $f : \mathbb{R}^N \to \mathbb{R}$ be a C^1 map. Given $\mathbf{B} \in SO_N(\mathbb{R})$, for any $\mathbf{s} \in \mathbb{R}^N$, let $\mathbf{s}' = \mathbf{B}\mathbf{s}$. Then*

$$\left(\frac{\partial f}{\partial s_1'}, \cdots, \frac{\partial f}{\partial s_N'}\right)^t = \mathbf{B}\left(\frac{\partial f}{\partial s_1}, \cdots, \frac{\partial f}{\partial s_N}\right)^t,$$

where superscript t denotes the transpose of the vector.

Proof. Since $\mathbf{s}' = \mathbf{B}\mathbf{s}$, $\mathbf{s} = \mathbf{B}^{-1}\mathbf{s}'$. Using chain rule, it can be verified that

$$\left(\frac{\partial f}{\partial s_1'}, \cdots, \frac{\partial f}{\partial s_N'}\right)^t = (\mathbf{B}^{-1})^t \cdot \left(\frac{\partial f}{\partial s_1}, \cdots, \frac{\partial f}{\partial s_N}\right)^t.$$

As $\mathbf{B} \in SO_N(\mathbb{R})$, we have $(\mathbf{B}^{-1})^t = \mathbf{B}$. This proves the lemma. □

Lemma 6.13. *Let $N \geq 1$ be an integer. Given real numbers $0 < c_2 < C_2 < c_1 < C_1$, and a partition $\{\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3\}$ of $\{1, \cdots, N\}$ such that $\mathcal{I}_1 \neq \emptyset$. There is $\mathbf{B} \in SO_N(\mathbb{R})$ (depending only on the partition) such that the following holds: For any vector $\mathbf{v} = (v_1, \cdots, v_N)^t \in \mathbb{R}^N$ (here superscript t denote the transpose of the corresponding vector) satisfying*

- *For any $i \in \{1, \cdots, N\}$, $|v_i| \leq C_1$;*
- *For any $i \in \mathcal{I}_1$, $|v_i| \geq c_1$;*
- *For any $i \in \mathcal{I}_2$, $|v_i| \geq C_2$;*
- *For any $i \in \mathcal{I}_3$, $|v_i| \leq c_2$.*

If we denote $\mathbf{v}' = (v_1', \cdots, v_N')^t = \mathbf{B}\mathbf{v}$, then for any $i = 1, \cdots, N$,

$$\min\left\{\frac{c_1}{\sqrt{N}} - \sqrt{N}c_2, C_2\right\} \leq |v_i'| \leq C_1 + \sqrt{N}c_2.$$

Proof. If $\mathcal{I}_3 = \emptyset$, then the lemma is trivial. Now we assume that $\mathcal{I}_3 \neq \emptyset$. Let $p \in \mathbb{N}$ be such that $p - 1 = \#\mathcal{I}_3$, then $2 \leq p \leq N$. Without loss of generality, we may assume that $1 \in \mathcal{I}_1$ and $\mathcal{I}_3 = \{2, \cdots, p\}$. Choose a $\mathbf{B} = (b_{ij}) \in SO_N(\mathbb{R})$ satisfying

- $b_{i1} = \frac{1}{\sqrt{p}}$, for $1 \leq i \leq p$;
- $b_{ij} = 0$ if $j \leq p < i$, or $i \leq p < j$;
- $b_{ij} = \delta_{ij}$ if $i \geq p+1$ and $j \geq p+1$.

Note that for any $i = 1, \cdots, p$, $v_i' = 1/\sqrt{p} \cdot v_1 + \sum_{j=2}^{p} b_{ij}v_j$. Therefore, for any

$1 \le i \le p$, we have the lower bound

$$|v_i'| \ge \frac{1}{\sqrt{p}}|v_1| - \sum_{j=2}^{p} |b_{ij}||v_j| \ge \frac{c_1}{\sqrt{p}} - \left(\sum_{j=2}^{p} |c_2|^2\right)^{\frac{1}{2}} \ge \frac{c_1}{\sqrt{N}} - \sqrt{N}c_2,$$

where in the second inequality we apply Cauchy-Schwartz inequality. Also for any $1 \le i \le p$, we have the upper bound

$$|v_i'| \le \frac{1}{\sqrt{p}}|v_1| + \sum_{j=2}^{p} |b_{ij}||v_j| \le C_1 + \sqrt{N}c_2.$$

On the other hand, for any $p+1 \le i \le N$, we have $v_i' = v_i$. Therefore, for any $1 \le i \le N$,

$$\min\left\{\frac{c_1}{\sqrt{N}} - \sqrt{N}c_2, C_2\right\} \le |v_i'| \le C_1 + \sqrt{N}c_2. \qquad \square$$

Roughly speaking, Lemma 6.13 says that one can find a suitable rotation $\mathbf{B} \in SO_N(\mathbb{R})$ depending **only** on the partition of $\{1, \cdots, N\}$ such that for any vector $\mathbf{v} \in \mathbb{R}^N$, as long as there is a coordinate of $\mathbf{v}$ with large enough absolute value, the absolute value of all coordinates of the new vector $\mathbf{Bv}$ are bounded below by a suitable constant.

Proof of (5) of Proposition 6.1. If $n = 0$ and $J = I$. Then for any $\mathbf{s} \in I$, by Proposition 6.1 (4),

$$\beta_{\mathbf{m}}(a_t u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) \le \tilde{M}_2 \beta_{\mathbf{m}}(u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) \le \tilde{M}_2(\sigma^{-\nu} + M).$$

Then for any $b \in \mathbb{R}$ such that $b > \tilde{M}_2(\sigma^{-\nu} + M)$,

$$\int_I \beta_{\mathbf{m}}(a_t u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) d\mathbf{s} \le \frac{1}{2}\int_I \beta_{\mathbf{m}}(u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) d\mathbf{s} + b|I|.$$

Now we assume $n \ge 1$ and let $t > 0$ be a sufficiently large number (to be specified later). By Lemma 6.9,

$$\int_J \beta_{\mathbf{m}}(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) d\mathbf{s} \le \int_J \int_{I_0} \beta_{\mathbf{m}}(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s} + \tilde{\mathbf{s}} e^{-dnt})\Gamma) d\tilde{\mathbf{s}} d\mathbf{s}.$$

By Proposition 5.1, for $t > 0$ sufficiently large, there exists $b_1 > 0$ such that for any $\mathbf{s} \in J$,

$$\int_{I_0} \tilde{\alpha}(a_t u(\tilde{\mathbf{s}}) a_{nt} u(\mathbf{s})\Gamma) d\tilde{\mathbf{s}} \le \frac{1}{4}\tilde{\alpha}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) + b_1. \tag{6.6}$$

Note that since

$$\begin{aligned}\int_{I_0} \tilde{\alpha}(a_t u(\tilde{\mathbf{s}}) a_{nt} u(\mathbf{s})\Gamma) d\tilde{\mathbf{s}} &= \int_{I_0} \tilde{\alpha}(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s} + \tilde{\mathbf{s}} e^{-dnt})\Gamma) d\tilde{\mathbf{s}} \\ &\leq \frac{1}{4}\tilde{\alpha}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) + b_1,\end{aligned}$$

by definition of $\beta_{\mathbf{m}}$, it remains to estimate the following integral for any $\mathbf{s} \in J$:

$$\int_{I_0} \alpha_{\mathbf{m}}^{\nu}(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s} + \tilde{\mathbf{s}} e^{-dnt})\Gamma) d\tilde{\mathbf{s}}. \tag{6.7}$$

Define for any $\mathbf{s} \in J$,

$$\begin{aligned} I_{01}(\mathbf{s}) &:= \{\tilde{\mathbf{s}} \in I_0 : \hat{\mathbf{s}} = \mathbf{s} + \tilde{\mathbf{s}} e^{-dnt}, \boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}} \text{ exists}, \boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}} \neq \boldsymbol{\xi}_{n,\mathbf{s}}\}, \\ I_{02}(\mathbf{s}) &:= \{\tilde{\mathbf{s}} \in I_0 : \hat{\mathbf{s}} = \mathbf{s} + \tilde{\mathbf{s}} e^{-dnt}, \boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}} \text{ exists}, \boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}} = \boldsymbol{\xi}_{n,\mathbf{s}}\}, \\ I_{03}(\mathbf{s}) &:= \{\tilde{\mathbf{s}} \in I_0 : \hat{\mathbf{s}} = \mathbf{s} + \tilde{\mathbf{s}} e^{-dnt}, \boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}} \text{ does not exist}\}. \end{aligned}$$

Since for $\tilde{\mathbf{s}} \in I_{03}$, $\alpha_{\mathbf{m}}^{\nu}$ is dominated by $\tilde{\alpha}$, by Lemmas 6.14, 6.15 given as follows, we have

$$\begin{aligned}&\int_{I_0} \alpha_{\mathbf{m}}^{\nu}(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s} + \tilde{\mathbf{s}} e^{-dnt})\Gamma) d\tilde{\mathbf{s}} = \int_{I_{01}(\mathbf{s})} \alpha_{\mathbf{m}}^{\nu} d\tilde{\mathbf{s}} + \int_{I_{02}(\mathbf{s})} \alpha_{\mathbf{m}}^{\nu} d\tilde{\mathbf{s}} + \int_{I_{03}(\mathbf{s})} \alpha_{\mathbf{m}}^{\nu} d\tilde{\mathbf{s}} \\ &\leq (10r^2 d)^{\nu} e^{\nu r t} \cdot 2^{r(d-r)} \cdot \tilde{\alpha}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) + \frac{1}{4}\alpha_{\mathbf{m}}^{\nu}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) + 2^{\nu + r(d-r)}.\end{aligned}$$

As we choose $b > 2^{\nu + r(d-r)} + c b_1 e^{r\nu t}$, recall that $c = 4 \cdot (10r^2 d)^{\nu} \cdot 2^{r(d-r)}$, we have

$$\begin{aligned}&\int_J \beta_{\mathbf{m}}(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) d\mathbf{s} \leq \int_J \int_{I_0} \beta_{\mathbf{m}}(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s} + \tilde{\mathbf{s}} e^{-dnt})\Gamma) d\tilde{\mathbf{s}} d\mathbf{s} \\ &\leq \int_J [\frac{1}{4} c \cdot e^{r\nu t} \tilde{\alpha}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) + \frac{1}{4}\alpha_{\mathbf{m}}^{\nu}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) + 2^{\nu + r(d-r)} + c b_1 e^{r\nu t}] d\mathbf{s}, \\ &\leq \frac{1}{2}\int_J \beta_{\mathbf{m}}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma) d\mathbf{s} + b|J|.\end{aligned}$$

This finishes the proof of property (5) of Proposition 6.1, modulo Lemmas 6.14, 6.15. □

Lemma 6.14. *Let J be the box as in Proposition 6.1 (5). There is $t > 0$ sufficiently large such that for any $\mathbf{s} \in J$,*

$$\int_{I_{01}(\mathbf{s})} \alpha_{\mathbf{m}}^{\nu}(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s} + \tilde{\mathbf{s}} e^{-dnt})\Gamma) d\tilde{\mathbf{s}} \leq e^{\nu r t}(10r^2 d)^{\nu} 2^{r(d-r)} \cdot \tilde{\alpha}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma).$$

Proof. We will prove that for $t>0$ sufficiently large, for any $\tilde{\mathbf{s}} \in I_{01}(\mathbf{s})$,

$$\alpha_{\mathbf{m}}^{\nu}\left(a_{(n+1)t}u_{\boldsymbol{\varphi}}(\mathbf{s}+\tilde{\mathbf{s}}e^{-dnt})\Gamma\right) \le e^{\nu rt}\left(10r^2d\right)^{\nu}\cdot\tilde{\alpha}\left(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma\right).$$

For $\tilde{\mathbf{s}} \in I_0$, denote $\hat{\mathbf{s}} = \mathbf{s}+\tilde{\mathbf{s}}e^{-dnt}$.

Case 1: $\left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}}))_{\le r}\right\|_\infty \ge N_2\left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}}))_{>r}\right\|_\infty$, where $N_2 = 2(r+1)(d-r)$.

By definition of M_1 (cf.(2.1)), the choice of N_1 and Lipschitz continuity of $\boldsymbol{\varphi}$, we have for any i,j,p,q,

$$\left|\sum_{q=1}^{k}\frac{\partial\boldsymbol{\varphi}_{pq}}{\partial s_{ij}}\cdot m_q\right| \le \|\mathbf{m}\|\cdot\left(\sum_{q=1}^{k}|\frac{\partial\boldsymbol{\varphi}_{pq}}{\partial s_{ij}}|^2\right)^{\frac{1}{2}} \le \|\mathbf{m}\|\,k^{\frac{1}{2}}\frac{M_1}{N_1} \le \left\|(\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})_{>r}\cdot\mathbf{m}\right\|_\infty.$$

By Lemma 6.8, the choices of the sidelength of the box and N_2,

$$\begin{aligned}
&\left\|w(n+1,\hat{\mathbf{s}},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})\right\| \ge \left\|(w(n+1,\hat{\mathbf{s}},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})_{\le r}\right\|_\infty \\
&= e^{(d-r)(n+1)t}\left\|[(\boldsymbol{\varphi}(\hat{\mathbf{s}}))_{\le r}-(\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})_{\le r}-\hat{\mathbf{s}}\cdot(\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})_{>r}]\cdot\mathbf{m}\right\|_\infty \\
&\ge e^{(d-r)t}\left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}}))_{\le r}\right\|_\infty - \\
&\quad (1-\frac{(r+1)(d-r)}{N_2})e^{(d-r)t}\left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}}))_{\le r}\right\|_\infty \\
&\ge e^{(d-r)t}\frac{1}{2d}\left\|w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})\right\|.
\end{aligned}$$

Choose $t>0$ large enough such that $e^{(d-r)t}\frac{1}{2d}>1$, we obtain

$$\begin{aligned}
\alpha_{\mathbf{m}}^{\nu}(a_{(n+1)t}u_{\boldsymbol{\varphi}}(\mathbf{s}+\tilde{\mathbf{s}}e^{-dnt})\Gamma) &= \left\|w(n+1,\hat{\mathbf{s}},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})\right\|^{-\nu} \\
&\le e^{-\nu(d-r)t}(2d)^{\nu}\left\|w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})\right\|^{-\nu} \\
&\le e^{-\nu(d-r)t}(2d)^{\nu}\cdot 2^{\nu}\sup_{\mathbf{w}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}}\|a_{nt}u(\mathbf{s})\mathbf{w}\|^{-\nu} \\
&\le \tilde{\alpha}(a_{nt}u_{\boldsymbol{\varphi}(\mathbf{s})}\Gamma).
\end{aligned}$$

Case 2: $\left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}}))_{\le r}\right\|_\infty < N_2\left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}}))_{>r}\right\|_\infty$.
Then by the choice of N_2, we have

$$\left\|w(n+1,\hat{\mathbf{s}},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})\right\| \ge e^{-rt}\left\|(w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}}))_{>r}\right\|_\infty$$

$$\geq e^{-rt}\frac{1}{5r^2d}\left\|w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})\right\|.$$

Therefore,

$$\begin{aligned}\alpha_{\mathbf{m}}^{\nu}(a_{(n+1)t}u_{\boldsymbol{\varphi}}(\mathbf{s}+\tilde{\mathbf{s}}e^{-dnt})\Gamma) &= \left\|w(n+1,\hat{\mathbf{s}},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})\right\|^{-\nu}\\ &\leq e^{\nu rt}(5r^2d)^{\nu}\left\|w(n,\mathbf{s},\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}})\right\|^{-\nu}\\ &\leq e^{\nu rt}(10r^2d)^{\nu}\sup_{\mathbf{w}\in\mathbb{Z}^d\setminus\{\mathbf{0}\}}\|a_{nt}u(\mathbf{s})\mathbf{w}\|^{-\nu}\\ &\leq e^{\nu rt}(10r^2d)^{\nu}\tilde{\alpha}(a_{nt}u_{\boldsymbol{\varphi}(\mathbf{s})}\Gamma).\end{aligned}$$

Combining cases 1 and 2, the lemma is proven. □

Lemma 6.15. *There exists $t>0$ sufficiently large such that for any $\mathbf{s}\in J$,*

$$\int_{I_{02}(\mathbf{s})}\alpha_{\mathbf{m}}^{\nu}(a_{(n+1)t}u_{\boldsymbol{\varphi}}(\mathbf{s}+\tilde{\mathbf{s}}e^{-dnt})\Gamma)d\tilde{\mathbf{s}}\leq\frac{1}{4}\alpha_{\mathbf{m}}^{\nu}(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma).$$

Proof. We fix $\mathbf{s}\in J$ for the rest of the proof. For $\tilde{\mathbf{s}}\in I_0$, denote $\hat{\mathbf{s}}=\mathbf{s}+\tilde{\mathbf{s}}e^{-dnt}$. Since $\tilde{\mathbf{s}}\in I_{02}(\mathbf{s})$, for simplicity we denote $\boldsymbol{\xi}_{n+1,\hat{\mathbf{s}}}=\boldsymbol{\xi}_{n,\mathbf{s}}=\mathbf{v}$.

Case 1: $\|(w(n,\mathbf{s},\mathbf{v}))_{\leq r}\|_{\infty}\geq N_2\|(w(n,\mathbf{s},\mathbf{v}))_{>r}\|_{\infty}$, where $N_2=2(r+1)(d-r)$.

Then we have by **Case 1** of Lemma 6.14,

$$\|w(n+1,\hat{\mathbf{s}},\mathbf{v})\|\geq e^{(d-r)t}\frac{1}{2d}\|w(n,\mathbf{s},\mathbf{v})\|.$$

Hence, for $t>0$ sufficiently large such that $e^{-(d-r)\nu t}\cdot(2d)^{\nu}\leq\frac{1}{4}$, we obtain

$$\alpha_{\mathbf{m}}^{\nu}\left(a_{(n+1)t}u_{\boldsymbol{\varphi}}(\mathbf{s}+\tilde{\mathbf{s}}e^{-dnt})\Gamma\right)\leq\frac{1}{4}\alpha_{\mathbf{m}}^{\nu}\left(a_{nt}u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma\right).$$

Case 2: $\|(w(n,\mathbf{s},\mathbf{v}))_{\leq r}\|_{\infty}<N_2\|(w(n,\mathbf{s},\mathbf{v}))_{>r}\|_{\infty}$.

Recall that $I_0=[-1,1]^{r(d-r)}$. For any $\mathbf{B}\in SO_{r(d-r)}(\mathbb{R})$, we have $\mathbf{B}\cdot I_0\subset I_0'$, where I_0' is the unit ball in $\mathbb{R}^{r(d-r)}$. We may choose $t>0$ large enough such that for any $\mathbf{s}\in I$, any $\tilde{\mathbf{s}}\in I_0'$, we have $\mathbf{s}+e^{-dnt}\tilde{\mathbf{s}}\in I'$, where I' is given as in Remark 6.3. Define a function S on I_0' by

$$S(\tilde{\mathbf{s}})=\sum_{i=1}^{r}w_i\left(n+1,\mathbf{s}+\tilde{\mathbf{s}}e^{-dnt},\mathbf{v}\right),\forall\tilde{\mathbf{s}}\in I_0'.$$

Note that $\|w(n+1,\hat{\mathbf{s}},\mathbf{v})\|\geq\frac{1}{r}|S(\tilde{\mathbf{s}})|$.

We will apply Lemma 6.13 to find $\mathbf{B} \in SO_{r(d-r)}(\mathbb{R})$ such that after the change of basis $\tilde{\mathbf{s}}' = \mathbf{B}\tilde{\mathbf{s}}$, for

$$A_1 = e^{(d-r)t} \max\left\{4\sqrt{r(d-r)}, 2(r(d-r))^{3/2} + r\right\} \|w(n, \mathbf{s}, \mathbf{v})\|,$$

$$A_2 = e^{(d-r)t} \frac{1}{40r^3 d(d-r)} \|w(n, \mathbf{s}, \mathbf{v})\|, \tag{6.8}$$

we have

$$A_2 \leq \|S\|_{I_0'} \leq A_1;\ A_2 \leq \left|\frac{\partial S}{\partial \tilde{s}_{ij}'}(\tilde{\mathbf{s}}')\right| \leq A_1, \forall i, j, \forall \tilde{\mathbf{s}}' \in I_0'. \tag{6.9}$$

Applying Lemma 6.11 to $S(\tilde{\mathbf{s}}')$, since $\mathbf{B}$ preserves Lebesgue measure, we obtain that for any $\epsilon > 0$,

$$\begin{aligned}
|\{\tilde{\mathbf{s}} \in I_0 : |S(\tilde{\mathbf{s}})| \leq \epsilon\}| &\leq |\{\tilde{\mathbf{s}}' \in I_0' : |S(\tilde{\mathbf{s}}')| \leq \epsilon\}| \\
&\leq r(d-r)\frac{12A_1}{A_2} \cdot \left(\frac{\epsilon}{\|S\|_{I_0'}}\right)^{\frac{1}{r(d-r)}} \cdot |I_0'| \\
&\leq \tilde{C} \cdot e^{-\frac{1}{r}t} \cdot \epsilon^{\frac{1}{r(d-r)}} \cdot \|w(n, \mathbf{s}, \mathbf{v})\|^{-\frac{1}{r(d-r)}},
\end{aligned}$$

where $\tilde{C}$ is a constant depending only on r and d. Choose $t > 0$ large enough, by Lemma 6.10, with $0 < \nu < \frac{1}{r(d-r)}$,

$$\begin{aligned}
\int_{I_{02}(\mathbf{s})} \alpha_{\mathbf{m}}^{\nu}(a_{(n+1)t} u_{\boldsymbol{\varphi}}(\mathbf{s} + \tilde{\mathbf{s}}e^{-dnt})\Gamma) d\tilde{\mathbf{s}} &\leq \int_{I_0} \frac{1}{\|w(n+1, \mathbf{s} + e^{-dnt}\tilde{\mathbf{s}}, \mathbf{v})\|^{\nu}} d\tilde{\mathbf{s}} \\
&\leq r^{\nu} \cdot \int_{I_0} \frac{1}{|S(\tilde{\mathbf{s}})|^{\nu}} d\tilde{\mathbf{s}} \\
&\leq r^{\nu} c_{\nu} \tilde{C}^{\nu r(d-r)} \cdot e^{-\nu(d-r)t} \|w(n, \mathbf{s}, \mathbf{v})\|^{-\nu} \\
&\leq \frac{1}{4} \alpha_{\mathbf{m}}^{\nu}(a_{nt} u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma).
\end{aligned}$$

This prove the lemma. Therefore, it remains to achieve (6.9). Consider the function $\Psi = \sum_{q=1}^{r} \sum_{p=1}^{k} m_p \boldsymbol{\varphi}_{qp}$, where $\mathbf{m} = (m_1, \cdots, m_k)^t$. We have

$$\begin{aligned}
S(\tilde{\mathbf{s}}) = e^{(d-r)(n+1)t} \Bigg[&\Psi(\mathbf{s} + e^{-dnt}\tilde{\mathbf{s}}) - \Psi(\mathbf{s}) - \sum_{i=1}^{r} \sum_{j=1}^{d-r} e^{-dnt} \tilde{s}_{ij} < (\mathbf{v})_{>r} \cdot \mathbf{m}, \mathbf{e}_j > \\
&+ e^{-(d-r)nt} \sum_{i=1}^{r} w_i(n, \mathbf{s}, \mathbf{v}) \Bigg].
\end{aligned} \tag{6.10}$$

Let $N_3 = 4r(d-r)$, define the partition $\{\mathcal{I}_1(\mathbf{s}), \mathcal{I}_2(\mathbf{s}), \mathcal{I}_3(\mathbf{s})\}$ of $\{(i,j) : 1 \leq i \leq r, 1 \leq j \leq d-r\}$ by

$$\mathcal{I}_1(\mathbf{s}) := \{(i,j) : |< (\mathbf{v})_{>r} \cdot \mathbf{m}, \mathbf{e}_j >| = \|(\mathbf{v})_{>r} \cdot \mathbf{m}\|_\infty\};$$
$$\mathcal{I}_2(\mathbf{s}) := \{(i,j) : \|(\mathbf{v})_{>r} \cdot \mathbf{m}\|_\infty > |< (\mathbf{v})_{>r} \cdot \mathbf{m}, \mathbf{e}_j >| \geq \frac{1}{N_3} \|(\mathbf{v})_{>r} \cdot \mathbf{m}\|_\infty\};$$
$$\mathcal{I}_3(\mathbf{s}) := \{(i,j) : |< (\mathbf{v})_{>r} \cdot \mathbf{m}, \mathbf{e}_j >| < \frac{1}{N_3} \|(\mathbf{v})_{>r} \cdot \mathbf{m}\|_\infty\}.$$

Note that by definition, $\mathcal{I}_1(\mathbf{s}) \neq \emptyset$. Using (6.10), the choice of N_3, and the estimate

$$\left|\frac{\partial \Psi}{\partial \tilde{s}_{ij}}(\mathbf{s} + e^{-dnt}\tilde{\mathbf{s}})\right| \leq \left(\sum_{p=1}^{k} m_p^2\right)^{\frac{1}{2}} \cdot \left(\sum_{p=1}^{k}(\sum_{q=1}^{r} \frac{\partial \varphi_{qp}}{\partial s_{ij}})^2)\right)^{\frac{1}{2}}$$
$$\leq \|\mathbf{m}\| \frac{k^{\frac{1}{2}} \cdot rM_1}{N_1}, \forall \tilde{\mathbf{s}} \in I_0',$$

the following holds for any $\tilde{\mathbf{s}} \in I_0'$:

- For any (i,j), where $1 \leq i \leq r$, $1 \leq j \leq d-r$,
$$\left|\frac{\partial S}{\partial \tilde{s}_{ij}}(\tilde{\mathbf{s}})\right| \leq C_1 := 2e^{(d-r)t} \|(w(n,\mathbf{s},\mathbf{v}))_{>r}\|_\infty ;$$
- For any $(i,j) \in \mathcal{I}_1(\mathbf{s})$,
$$\left|\frac{\partial S}{\partial \tilde{s}_{ij}}(\tilde{\mathbf{s}})\right| \geq c_1 := e^{(d-r)t} \left(1 - \frac{k^{1/2} r}{N_1}\right) \|(w(n,\mathbf{s},\mathbf{v}))_{>r}\|_\infty ;$$
- For any $(i,j) \in \mathcal{I}_2(\mathbf{s})$,
$$\left|\frac{\partial S}{\partial \tilde{s}_{ij}}(\tilde{\mathbf{s}})\right| > C_2 := e^{(d-r)t} \left(\frac{1}{N_3} - \frac{k^{1/2} r}{N_1}\right) \|(w(n,\mathbf{s},\mathbf{v}))_{>r}\|_\infty ;$$
- For any $(i,j) \in \mathcal{I}_3(\mathbf{s})$,
$$\left|\frac{\partial S}{\partial \tilde{s}_{ij}}(\tilde{\mathbf{s}})\right| < c_2 := e^{(d-r)t} \left(\frac{1}{N_3} + \frac{k^{1/2} r}{N_1}\right) \|(w(n,\mathbf{s},\mathbf{v}))_{>r}\|_\infty .$$

Applying Lemma 6.13 with $N = r(d-r)$, and C_1, c_1, C_2, c_2 given as above, we obtain $\mathbf{B} = \mathbf{B}(\mathbf{s}) \in SO_{r(d-r)}(\mathbb{R})$ (Since the partition depends only on $\mathbf{s}$,

$\mathbf{B}$ depends only on $\mathbf{s}$), such that after the change of basis $\tilde{\mathbf{s}}' = \mathbf{B}\tilde{\mathbf{s}}$, the vector $\mathbf{v}(\tilde{\mathbf{s}}') = (\frac{\partial S}{\partial \tilde{s}'_{ij}}(\tilde{\mathbf{s}}'))_{ij}$ satisfies the following: For any $1 \leq i \leq r, 1 \leq j \leq d-r$,

$$\min\left\{\frac{c_1}{\sqrt{N}} - \sqrt{N}c_2, C_2\right\} \leq \left|\frac{\partial S}{\partial \tilde{s}'_{ij}}(\tilde{\mathbf{s}}')\right| \leq C_1 + \sqrt{N}c_2, \forall \tilde{\mathbf{s}}' \in I'_0. \tag{6.11}$$

By the assumption of Case 2, and the choice of N_2, it is elementary to verify that

$$\min\left\{\frac{c_1}{\sqrt{N}} - \sqrt{N}c_2, C_2\right\} \geq e^{(d-r)t}\frac{1}{40r^3 d(d-r)}\left\|w(n,\mathbf{s},\mathbf{v})\right\|, \text{ and}$$

$$C_1 + \sqrt{N}c_2 \leq 4\sqrt{r(d-r)}e^{(d-r)t}\left\|w(n,\mathbf{s},\mathbf{v})\right\|. \tag{6.12}$$

Moreover, using the expression (6.10), we obtain

$$\|S\|_{I'_0} \leq e^{(d-r)t}(2(r(d-r))^{3/2} + r)\left\|w(n,\mathbf{s},\mathbf{v})\right\|. \tag{6.13}$$

Also, note that

$$\|S\|_{I'_0} \geq \inf_{\tilde{\mathbf{s}} \in I_0, (i,j) \in \mathcal{I}_1}\left|\frac{\partial S}{\partial \tilde{s}_{ij}}(\tilde{\mathbf{s}})\right| \geq e^{(d-r)t}\frac{1}{10r^2 d}\left\|w(n,\mathbf{s},\mathbf{v})\right\|. \tag{6.14}$$

Now we choose A_1, A_2 as in (6.8), by (6.11)(6.12)(6.13)(6.14), (6.9) is achieved. This finishes the proof of the lemma. □

7. Proof of Proposition 2.1

Following a general strategy developed in [8, Section 6.6], we derive Proposition 2.1 from Proposition 6.1.

Let Y be a locally compact, second countable Hausdorff topological space. Let B be a compact box in $Mat_{r\times(d-r)}(\mathbb{R})$. Let $\phi : Mat_{r\times(d-r)}(\mathbb{R}) \to Y$ be a continuous map. Let $f : \mathbb{R} \times Y \to Y$ be a continuous map and we write $f(t, y)$ as $f^t(y)$ for $(t, y) \in \mathbb{R} \times Y$.

Let $\mathcal{F}_0 = \{B\}$. For every $n \in \mathbb{N}$, let $\mathcal{F}_n$ be a partition of elements in $\mathcal{F}_{n-1}$ into countably many subboxes with positive Lebesgue measure. By construction, $\{\mathcal{F}_n\}_{n\in\mathbb{Z}_{\geq 0}}$ is a filtration. For any $\mathbf{s} \in B$, let $I_n(\mathbf{s})$ denote the atom in $\mathcal{F}_n$ containing $\mathbf{s}$.

Let $\beta : Y \to [1, \infty]$ be a measurable map. Assume that β satisfies the following conditions:

(1) β satisfies contraction hypothesis: There exist $0 < a < 1$ and $b > 0$ such that for any $n \in \mathbb{Z}_{\geq 0}$ and any atom I_n in $\mathcal{F}_n$,

$$\int_{I_n} \beta(f^{n+1}\phi(\mathbf{s}))d\mathbf{s} < a \int_{I_n} \beta(f^n\phi(\mathbf{s}))d\mathbf{s} + b|I_n|; \tag{7.1}$$

(2) β satisfies Lipschitz property: There exists a constant $M > 0$ such that for any $\mathbf{s} \in B$, any $n \in \mathbb{Z}_{\geq 0}$, and any $\tilde{\mathbf{s}} \in I_n(\mathbf{s})$,

$$\begin{aligned} \beta(f^n\phi(\tilde{\mathbf{s}})) &\leq M\beta(f^n\phi(\mathbf{s})), \\ \beta(f^{n+1}\phi(\mathbf{s})) &\leq M\beta(f^n\phi(\mathbf{s})); \end{aligned} \tag{7.2}$$

(3) β is bounded on $\phi(B)$, that is, there exists $l > 0$ such that

$$\{\phi(\mathbf{s}) : \mathbf{s} \in B\} \subset Y_l = \{y \in Y : \beta(y) < l\}. \tag{7.3}$$

For any $T > 0$ and a measurable subset K of Y, define

$$\mathcal{A}_K^T(\mathbf{s}) := \frac{1}{T}\int_0^T \chi_K(f^t\phi(\mathbf{s}))dt,$$

where χ_K is the indicator function of K.

Lemma 7.1. *[8, Lemma 6.20] For any $\epsilon > 0$, there exist $0 < l_1 < \infty$ and $0 < c_1 < 1$ such that for $K = Y_{l_1}$, and any $T > 1$,*

$$|\{\mathbf{s} \in B : \mathcal{A}_K^T(\mathbf{s}) \leq 1 - \epsilon\}| \leq c_1^T|B|.$$

proof of Proposition 2.1. We will apply Lemma 7.1 to $Y = X$, $B = I$, $\beta = \beta_{\mathbf{m}}$, $\phi(\mathbf{s}) = u_{\boldsymbol{\varphi}}(\mathbf{s})\Gamma$ and $f^t = a_t$ for $t > 0$ sufficiently large so that Proposition 6.1 holds.

Recall that I is a closed cube in $Mat_{r\times(d-r)}(\mathbb{R})$. We may assume that $t > 0$ is large enough such that e^{-dt} is less than the length of each side of I.

We construct a filtration $\{\mathcal{F}_n\}_{n\in\mathbb{N}}$ on I as follows. Let $\mathcal{F}_0 = \{\emptyset, I\}$. Suppose that we have already constructed $\mathcal{F}_{n-1}$. We divide each box J of $\mathcal{F}_{n-1}$ consecutively into cubes and boxes such that cubes have side length e^{-dnt} and boxes have side length between e^{-dnt} and $2e^{-dnt}$. Then conditions (7.1)(7.2)(7.3) follows from Proposition 6.1.

Therefore, applying Lemma 7.1, we obtain $l_1 > 0$ such that the set K defined by

$$K := \{x \in X : \beta_{\mathbf{m}}(x) < l_1\}$$

is a compact subset of $X \setminus X_{\mathbf{m}}$, and (2.4) holds. □

8. Proof of variants of Theorem 1.3

Given $\mathbf{M} \in SL_d(\mathbb{R})$, we may write

$$\mathbf{M} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix}, \tag{8.1}$$

where $\mathbf{A} \in Mat_{r\times r}(\mathbb{R})$, $\mathbf{B} \in Mat_{r\times(d-r)}(\mathbb{R})$, $\mathbf{C} \in Mat_{(d-r)\times r}(\mathbb{R})$, $\mathbf{D} \in Mat_{(d-r)\times(d-r)}(\mathbb{R})$. For $\mathbf{s} \in \mathcal{U}$, we can write

$$u(\mathbf{s})\mathbf{M} = \begin{bmatrix} \mathbf{A} + \mathbf{s}\cdot\mathbf{C} & \mathbf{B} + \mathbf{s}\cdot\mathbf{D} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} = \begin{bmatrix} \mathbf{A}(\mathbf{s}) & \mathbf{B}(\mathbf{s}) \\ \mathbf{C} & \mathbf{D} \end{bmatrix}. \tag{8.2}$$

Since $\mathbf{M} \in SL_d(\mathbb{R})$, it is clear that the set of $\mathbf{s} \in \mathcal{U}$ such that $\det \mathbf{A}(\mathbf{s}) = 0$ is a proper algebraic subvariety of $\mathcal{U}$ and hence, it has Lebesgue measure zero.

Therefore, we can assume that $\det \mathbf{A}(\mathbf{s}) \neq 0$, and

$$u(\mathbf{s})\mathbf{M} = \begin{bmatrix} \mathbf{A}(\mathbf{s}) & \mathbf{0} \\ \mathbf{C} & \mathbf{D} - \mathbf{C}\mathbf{A}(\mathbf{s})^{-1}\mathbf{B}(\mathbf{s}) \end{bmatrix} \cdot \begin{bmatrix} \mathbf{1}_r & \mathbf{A}(\mathbf{s})^{-1}\mathbf{B}(\mathbf{s}) \\ \mathbf{0}_{d-r,r} & \mathbf{1}_{d-r} \end{bmatrix}.$$

We may write

$$u(\mathbf{s})\mathbf{M}(Id, \boldsymbol{\varphi}(\mathbf{s})) = (u(\mathbf{s})\mathbf{M}(Id, \boldsymbol{\varphi}(\mathbf{s})))_- \cdot (u(\mathbf{s})\mathbf{M}(Id, \boldsymbol{\varphi}(\mathbf{s})))_+,$$

where

$$(u(\mathbf{s})\mathbf{M}(Id, \boldsymbol{\varphi}(\mathbf{s})))_- = \begin{bmatrix} \mathbf{A}(\mathbf{s}) & \mathbf{0} \\ \mathbf{C} & \mathbf{D} - \mathbf{C}\mathbf{A}^{-1}(\mathbf{s})\mathbf{B}(\mathbf{s}) \end{bmatrix} \cdot \left(Id, \begin{bmatrix} \mathbf{0} \\ (\boldsymbol{\varphi}(\mathbf{s}))_{>r} \end{bmatrix} \right),$$

$$\begin{aligned}&(u(\mathbf{s})\mathbf{M}(Id, \boldsymbol{\varphi}(\mathbf{s})))_+ \\ &= \left(\begin{bmatrix} \mathbf{1}_r & \mathbf{A}(\mathbf{s})^{-1}\mathbf{B}(\mathbf{s}) \\ \mathbf{0}_{d-r,r} & \mathbf{1}_{d-r} \end{bmatrix}, \begin{bmatrix} (\boldsymbol{\varphi}(\mathbf{s}))_{\leq r} + \mathbf{A}(\mathbf{s})^{-1}\mathbf{B}(\mathbf{s}) \cdot (\boldsymbol{\varphi}(\mathbf{s}))_{>r} \\ \mathbf{0} \end{bmatrix} \right).\end{aligned}$$

Lemma 8.1. *For a.e. $\mathbf{s}_0 \in \mathcal{U}$, there is an open neighborhood $\mathcal{V}$ of $\mathbf{s}_0$ contained in $\mathcal{U}$ and an open subset $\tilde{\mathcal{V}}$ of $Mat_{r\times(d-r)}(\mathbb{R})$, such that the map $\phi : \mathcal{V} \to \tilde{\mathcal{V}}$ defined by*

$$\phi(\mathbf{s}) = \mathbf{A}(\mathbf{s})^{-1}\mathbf{B}(\mathbf{s}) \tag{8.3}$$

is a diffeomorphism.

Proof. For any $\mathbf{s}_0 \in \mathcal{U}$ such that $\det \mathbf{A}(\mathbf{s}_0) \neq 0$, there is a neighborhood $\mathcal{V}$ of $\mathbf{s}_0$, for any $\mathbf{s} \in \mathcal{V}$, the map $\phi(\mathbf{s})$ is well defined and differentiable. Therefore $\tilde{\mathcal{V}} = \phi(\mathcal{V}) \subset Mat_{r\times(d-r)}(\mathbb{R})$ is an open subset. Let

$$\phi^{-1}(\tilde{\mathbf{s}}) = (\mathbf{A}\tilde{\mathbf{s}} - \mathbf{B})(\mathbf{D} - \mathbf{C}\tilde{\mathbf{s}})^{-1}$$

for any $\tilde{\mathbf{s}}$ such that $\det(\mathbf{D} - \mathbf{C}\tilde{\mathbf{s}}) \neq 0$. Now we verify that ϕ^{-1} is the inverse of ϕ, that is, we need to verify that for $\tilde{\mathbf{s}} = \phi(\mathbf{s})$,

$$\mathbf{A}\tilde{\mathbf{s}} - \mathbf{B} = \mathbf{s}(\mathbf{D} - \mathbf{C}\tilde{\mathbf{s}}). \tag{8.4}$$

Note that as $\tilde{\mathbf{s}} = \phi(\mathbf{s})$, we have $(\mathbf{A} + \mathbf{s}\mathbf{C})\tilde{\mathbf{s}} = \mathbf{B} + \mathbf{s}\mathbf{D}$. Therefore, left hand side of (8.4) is $\mathbf{A}\tilde{\mathbf{s}} - \mathbf{B} = (\mathbf{A} + \mathbf{s}\mathbf{C} - \mathbf{s}\mathbf{C})\tilde{\mathbf{s}} - \mathbf{B} = \mathbf{s}\mathbf{D} - \mathbf{s}\mathbf{C}\tilde{\mathbf{s}}$, which is equal to the right hand side of (8.4). □

Proof of Theorem 1.5. Choose $\mathbf{s}_0, \mathcal{V}, \tilde{\mathcal{V}}$ satisfying Lemma 8.1. By Lemma 2.3, it suffices to prove that for a.e. $\mathbf{s} \in \mathcal{V}$, the point

$$(u(\mathbf{s})\mathbf{M}(Id, \boldsymbol{\varphi}(\mathbf{s})))_+\Gamma \tag{8.5}$$

is Birkhoff generic with respect to (X, μ_X, a_t). For $\tilde{\mathbf{s}} \in \tilde{\mathcal{V}}$, define

$$\tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{s}}) := \begin{bmatrix} (\boldsymbol{\varphi}(\phi^{-1}(\tilde{\mathbf{s}})))_{\leq r} + \tilde{\mathbf{s}} \cdot (\boldsymbol{\varphi}(\phi^{-1}(\tilde{\mathbf{s}})))_{>r} \\ \mathbf{0} \end{bmatrix}.$$

Applying Corollary 1.4 to $u_{\tilde{\boldsymbol{\varphi}}}(\tilde{\mathbf{s}})\Gamma$ for $\tilde{\mathbf{s}} \in \tilde{\mathcal{V}}$, we obtain that if for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$,

$$|\{\tilde{\mathbf{s}} \in \tilde{\mathcal{V}} : (\tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{s}}))_{\leq r} \cdot \mathbf{m} \in \tilde{\mathbf{s}} \cdot \mathbb{Z}^{d-r} + \mathbb{Z}^r\}| = 0, \tag{8.6}$$

then for a.e. $\tilde{\mathbf{s}} \in \tilde{\mathcal{V}}$, $u_{\tilde{\boldsymbol{\varphi}}}(\tilde{\mathbf{s}})\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_t).

Suppose for some $\tilde{\mathbf{s}} \in \tilde{\mathcal{V}}$, and some $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$,

$$(\tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{s}}))_{\leq r} \cdot \mathbf{m} \in \tilde{\mathbf{s}} \cdot \mathbb{Z}^{d-r} + \mathbb{Z}^r. \tag{8.7}$$

Let $\mathbf{s} = \phi^{-1}(\tilde{\mathbf{s}})$. By definition of ϕ and $\tilde{\boldsymbol{\varphi}}$, (8.7) implies that there exist $\mathbf{a} \in \mathbb{Z}^{d-r}$ and $\mathbf{b} \in \mathbb{Z}^r$ such that

$$(\mathbf{A}(\mathbf{s}) \cdot (\boldsymbol{\varphi}(\mathbf{s}))_{\leq r} + \mathbf{B}(\mathbf{s}) \cdot (\boldsymbol{\varphi}(\mathbf{s}))_{>r}) \cdot \mathbf{m} = \mathbf{B}(\mathbf{s}) \cdot \mathbf{a} + \mathbf{A}(\mathbf{s}) \cdot \mathbf{b},$$

then (8.7) implies that

$$\begin{bmatrix} \mathbf{A}(\mathbf{s}) & \mathbf{B}(\mathbf{s}) \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \cdot \boldsymbol{\varphi}(\mathbf{s}) \cdot \mathbf{m} = \begin{bmatrix} \mathbf{B}(\mathbf{s}) \cdot \mathbf{a} + \mathbf{A}(\mathbf{s}) \cdot \mathbf{b} \\ (\mathbf{C} \cdot (\boldsymbol{\varphi}(\mathbf{s}))_{\leq r} + \mathbf{D} \cdot (\boldsymbol{\varphi}(\mathbf{s}))_{>r}) \cdot \mathbf{m} \end{bmatrix}. \tag{8.8}$$

By (8.2), (8.8) further implies that

$$\begin{aligned} \boldsymbol{\varphi}(\mathbf{s}) \cdot \mathbf{m} &= \mathbf{M}^{-1} u(-\mathbf{s}) \cdot \begin{bmatrix} \mathbf{B}(\mathbf{s}) \cdot \mathbf{a} + \mathbf{A}(\mathbf{s}) \cdot \mathbf{b} \\ (\mathbf{C} \cdot (\boldsymbol{\varphi}(\mathbf{s}))_{\leq r} + \mathbf{D} \cdot (\boldsymbol{\varphi}(\mathbf{s}))_{>r}) \cdot \mathbf{m} \end{bmatrix} \\ &\in \mathbb{Z}^d + \mathbf{M}^{-1} u(-\mathbf{s}) \cdot \begin{bmatrix} \mathbf{0} \\ \mathbb{R}^{d-r} \end{bmatrix}. \end{aligned}$$

Therefore, the condition (1.9) implies (8.6). By definition, for any $\mathbf{s} \in \mathcal{V}$,

$$(u(\mathbf{s})\mathbf{M}(Id, \boldsymbol{\varphi}(\mathbf{s})))_+ = u_{\tilde{\boldsymbol{\varphi}}}(\tilde{\mathbf{s}}), \text{ where } \tilde{\mathbf{s}} = \phi(\mathbf{s}).$$

This finishes the proof. □

Proof of Corollary 1.8. Note that $u_{\boldsymbol{\varphi}}(\mathbf{s}) \cdot (\mathbf{M}, \mathbf{v}) = u(\mathbf{s})\mathbf{M}(Id, \tilde{\boldsymbol{\varphi}}(\mathbf{s}))$, where $\tilde{\boldsymbol{\varphi}}(\mathbf{s}) = \mathbf{M}^{-\mathbf{1}}(\boldsymbol{\varphi}(\mathbf{s}) + \mathbf{v})$. Applying Theorem 1.5 to $u(\mathbf{s})\mathbf{M}(Id, \tilde{\boldsymbol{\varphi}}(\mathbf{s}))\Gamma$, we obtain that if for any $\mathbf{m} \in \mathbb{Z} \setminus \{\mathbf{0}\}$,

$$|\{\mathbf{s} \in \mathcal{U} : \tilde{\boldsymbol{\varphi}}(\mathbf{s}) \cdot \mathbf{m} \in \mathbf{M}^{-1} u(-\mathbf{s}) \cdot \begin{bmatrix} \mathbf{0} \\ \mathbb{R}^{d-r} \end{bmatrix} + \mathbb{Z}^d\}| = 0, \tag{8.9}$$

then for Lebesgue a.e. $\mathbf{s} \in \mathcal{U}$, $u(\mathbf{s})\mathbf{M}(Id, \tilde{\boldsymbol{\varphi}}(\mathbf{s}))\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_t). By definition of $\tilde{\boldsymbol{\varphi}}$, (8.9) is equivalent to

$$|\{\mathbf{s} \in \mathcal{U} : (\boldsymbol{\varphi}(\mathbf{s}) + \mathbf{v}) \cdot \mathbf{m} \in u(-\mathbf{s}) \cdot \begin{bmatrix} \mathbf{0} \\ \mathbb{R}^{d-r} \end{bmatrix} + \mathbf{M} \cdot \mathbb{Z}^d\}| = 0.$$

The corollary is proven. □

Proof of Corollary 1.9. Fix an $\mathbf{s}_0 \in \mathcal{U}$ at which the map $\mathbf{s} \mapsto \mathbf{E_1}(\mathbf{s})^{-1} \cdot [\mathbf{e}_{r+1}, \cdots, \mathbf{e}_d]$ has a nonsingular differential. It is enough to prove the corollary for a.e. $\mathbf{s}$ in a neighborhood of $\mathbf{s}_0$. Choose a neighborhood $\mathcal{V}$ of $\mathbf{s}_0$ such that for any $\mathbf{s} \in \mathcal{V}$, as in (8.1) we can write

$$\mathbf{E}_2(\mathbf{s}) = \mathbf{E}_1(\mathbf{s}) \cdot \mathbf{E}_1(\mathbf{s}_0)^{-1} = \begin{bmatrix} \mathbf{A}(\mathbf{s}) & \mathbf{B}(\mathbf{s}) \\ \mathbf{C}(\mathbf{s}) & \mathbf{D}(\mathbf{s}) \end{bmatrix},$$

where $\det \mathbf{A}(\mathbf{s}) \neq 0$ and $\det \mathbf{D}(\mathbf{s}) \neq 0$. This can be done by smoothness of $\mathbf{E}_1$.

Since $\mathbf{E}_2(\mathbf{s}) \in SO_d(\mathbb{R})$, $\mathbf{E}_2(\mathbf{s}) \cdot \mathbf{E}_2(\mathbf{s})^t = Id$, that is,

$$\begin{bmatrix} \mathbf{A}(\mathbf{s}) & \mathbf{B}(\mathbf{s}) \\ \mathbf{C}(\mathbf{s}) & \mathbf{D}(\mathbf{s}). \end{bmatrix} \cdot \begin{bmatrix} \mathbf{A}(\mathbf{s})^t & \mathbf{C}(\mathbf{s})^t \\ \mathbf{B}(\mathbf{s})^t & \mathbf{D}(\mathbf{s})^t \end{bmatrix} = Id.$$

In particular, we have

$$\mathbf{A}(\mathbf{s}) \cdot \mathbf{C}(\mathbf{s})^t + \mathbf{B}(\mathbf{s}) \cdot \mathbf{D}(\mathbf{s})^t = \mathbf{0}.$$

We may write

$$\mathbf{E}_2(\mathbf{s}) = \mathbf{E}_2(\mathbf{s})_- \cdot u(-\mathbf{C}(\mathbf{s})^t \cdot (\mathbf{D}(\mathbf{s})^t)^{-1}),$$

where

$$\mathbf{E}_2(\mathbf{s})_- = \begin{bmatrix} \mathbf{A}(\mathbf{s}) & \mathbf{0} \\ \mathbf{C}(\mathbf{s}) & \mathbf{D}(\mathbf{s}) - \mathbf{C}\mathbf{A}(\mathbf{s})^{-1}\mathbf{B}(\mathbf{s}) \end{bmatrix}.$$

By Lemma 2.3, for any $\mathbf{s} \in \mathcal{U}$, $\mathbf{E}_1(\mathbf{s})(Id, \boldsymbol{\varphi}(\mathbf{s}))\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_t) if and only if

$$u(-\mathbf{C}(\mathbf{s})^t \cdot (\mathbf{D}(\mathbf{s})^t)^{-1}) \cdot \mathbf{E}_1(\mathbf{s}_0)(Id, \boldsymbol{\varphi}(\mathbf{s}))\Gamma$$

is Birkhoff generic with respect to (X, μ_X, a_t).

By assumption, the map

$$\mathbf{s} \mapsto \mathbf{E}_2(\mathbf{s})^{-1} \cdot [\mathbf{e}_{r+1}, \cdots, \mathbf{e}_d] = \begin{bmatrix} \mathbf{C}(\mathbf{s})^t \\ \mathbf{D}(\mathbf{s})^t \end{bmatrix}$$

has nonsingular differential at $\mathbf{s}_0$. Thus the map

$$\phi : \mathbf{s} \mapsto -\mathbf{C}(\mathbf{s})^t \cdot (\mathbf{D}(\mathbf{s})^t)^{-1} \tag{8.10}$$

also has nonsingular differential at $\mathbf{s}_0$.

Shrink the neighborhood $\mathcal{V}$ of $\mathbf{s}_0$ if necessary, we can assume that there exists an open subset $\tilde{\mathcal{V}}$ of $Mat_{r\times(d-r)}(\mathbb{R})$ such that $\phi : \mathcal{V} \to \tilde{\mathcal{V}}$ is a diffeomorphism. Denote ϕ^{-1} the inverse of ϕ.

Let $\tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{s}}) = \boldsymbol{\varphi}(\phi^{-1}(\tilde{\mathbf{s}}))$. Applying Theorem 1.5 to

$$\{u(\tilde{\mathbf{s}})\mathbf{E}_1(\mathbf{s}_0)(Id, \tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{s}}))\Gamma : \tilde{\mathbf{s}} \in \tilde{\mathcal{V}}\},$$

we obtain that if for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$,

$$\left|\left\{\tilde{\mathbf{s}} \in \tilde{\mathcal{V}} : \tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{s}}) \cdot \mathbf{m} \in \mathbf{E}_1(\mathbf{s}_0)^{-1} \cdot u(-\tilde{\mathbf{s}}) \cdot \begin{bmatrix} \mathbf{0} \\ \mathbb{R}^{d-r} \end{bmatrix} + \mathbb{Z}^d \right\}\right| = 0, \tag{8.11}$$

then for a.e. $\tilde{\mathbf{s}} \in \tilde{\mathcal{V}}$, $u(\tilde{\mathbf{s}})\mathbf{E}_1(\mathbf{s}_0)(Id, \tilde{\boldsymbol{\varphi}}(\tilde{\mathbf{s}}))\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_t). Since ϕ is a diffeomorphism, (8.11) is equivalent to

$$\left|\left\{\mathbf{s} \in \mathcal{V} : \boldsymbol{\varphi}(\mathbf{s}) \cdot \mathbf{m} \in \mathbf{E}_1(\mathbf{s})^{-1} \cdot \begin{bmatrix} \mathbf{0} \\ \mathbb{R}^{d-r} \end{bmatrix} + \mathbb{Z}^d \right\}\right| = 0.$$

This completes the proof. □

9. Application to universal hitting time statistics for integrable flows

9.1. An adapted form of Corollary 1.9

Following notations of [3], for $l > 0$, let

$$D(e^{-l}) = diag[e^{-(d-1)l}, e^l, \cdots, e^l].$$

Theorem 9.1. *Let $\mathcal{U}$ be a bounded open subset of $\mathbb{R}^{d-1}$ and $\varphi : \mathcal{U} \to (\mathbb{R}^d)^k$ be a C^1 map. Let $\mathbf{E_1} : \mathcal{U} \to SO_d(\mathbb{R})$ be a smooth map such that the map $\mathbf{s} \mapsto \mathbf{E_1}(\mathbf{s})^{-1} \cdot \mathbf{e}_1$ has a nonsingular differential at Lebesgue almost every $\mathbf{s} \in \mathcal{U}$. Assume that for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$,*

$$|\{\mathbf{s} \in \mathcal{U} : \varphi(\mathbf{s}) \cdot \mathbf{m} \in \mathbb{R}\mathbf{E}_1(\mathbf{s})^{-1} \cdot \mathbf{e}_1 + \mathbb{Z}^d\}| = 0.$$

Then for Lebesgue a.e. $\mathbf{s} \in \mathcal{U}$, $\mathbf{E}_1(\mathbf{s})(Id, \varphi(\mathbf{s}))\Gamma$ is Birkhoff generic with respect to $(X, \mu_X, D(e^{-l}))$.

Proof. For any $l > 0$, denote $a_l = diag[e^l, \cdots, e^l, e^{-(d-1)l}]$. Choose $\omega \in SO_d(\mathbb{R}) \cap SL_d(\mathbb{Z})$ such that for any

$$\omega^{-1} \cdot D(e^{-l}) \cdot \omega = a_l.$$

Note that since $\omega \in SL_d(\mathbb{Z})$,

$$\omega^{-1} D(e^{-l}) E_1(\mathbf{s})(Id, \varphi_(\mathbf{s}))\Gamma = a_l \omega^{-1} E_1(\mathbf{s})\omega(Id, \omega^{-1}\varphi(\mathbf{s}))\Gamma.$$

Applying Corollary 1.9, we obtain that for a.e. $\mathbf{s} \in \mathcal{U}$, $\omega^{-1}E_1(\mathbf{s})\omega(Id, \omega^{-1}\varphi(\mathbf{s}))\Gamma$ is Birkhoff generic with respect to (X, μ_X, a_l), and the theorem follows. □

9.2. Universal hitting time

Let $(\mathcal{M}, \mathcal{B}, \nu)$ be a measurable space with probability measure ν, and $\varphi^t : \mathcal{M} \to \mathcal{M}$ be a measure-preserving dynamical system. Given some target set $\mathcal{D} \subset \mathcal{M}$, it is natural to study how often a φ-trajectory along random initial data $x \in \mathcal{M}$ intersects this target set. On the other hand, another question is to consider a sequence of randomized target sets whose "size" shrink to zero, and study the distribution of intersection times of a random φ-trajectory with these shrinking targets. The interested reader is referred to [3] and the references therein for a survey of the history of aforementioned questions.

In the setting of universal hitting time statistics for integrable flows (cf. [3]), the above question is studied when the measurable space is a d dimensional torus $\mathbb{T}^d$ and φ^t is a linear flow on the torus. Now let $d \geq 2$ and $k \geq 1$ be fixed integers. In this article, the sequence of target sets we consider is a sequence of union of k many bounded codimensional one balls in $\mathbb{T}^d$, whose radius shrink to zero. More precisely, let $\mathcal{U}$ be a bounded open subset of $\mathbb{R}^{d-1}$. Consider the following smooth functions:

$$\boldsymbol{\theta}, \boldsymbol{\phi}_j : \mathcal{U} \to \mathbb{T}^d, 1 \leq j \leq k,$$
$$\mathbf{f}, \mathbf{u}_j : \mathcal{U} \to \mathbf{S}_1^{d-1}, 1 \leq j \leq k,$$

where $\mathbf{S}_1^{d-1}$ is the unit one sphere in $\mathbb{R}^d$. For the functions above, we assign to any $\mathbf{s} \in \mathcal{U}$ the following:

- $\boldsymbol{\theta}(\mathbf{s})$ =initial position of the flow;
- $\mathbf{f}(\mathbf{s})$ =direction of the flow;
- $\mathbf{u}_j(\mathbf{s})$ =direction of the j-th target ball;
- $\boldsymbol{\phi}_j(\mathbf{s})$ =center of the j-th target ball.

With these functions, we define the flow

$$\varphi^t : \mathcal{U} \to \mathbb{T}^d \times \mathcal{U}, \mathbf{s} \mapsto (\boldsymbol{\theta}(\mathbf{s}) + t\mathbf{f}(\mathbf{s}), \mathbf{s}). \tag{9.1}$$

From now on, we fix a map $\mathbf{v} \mapsto \mathbf{R}_{\mathbf{v}}$ from $\mathbf{S}_1^{d-1}$ to $SO_d(\mathbb{R})$ such that for all $\mathbf{v} \in \mathbf{S}_1^{d-1}$,

$$\mathbf{R}_{\mathbf{v}} \cdot \mathbf{v} = \mathbf{e}_1, \tag{9.2}$$

and $\mathbf{v} \mapsto \mathbf{R}_{\mathbf{v}}$ is smooth on $\mathbf{S}_1^{d-1} \setminus \{\mathbf{v}_0\}$ for a singular point $\mathbf{v}_0 \in \mathbf{S}_1^{d-1}$. For $1 \leq j \leq k$, fix a bounded open subset $\Omega_j \subset \mathbb{R}^{d-1} \times \mathcal{U}$. For any $l > 0$, denote the l-level target set with to be

$$\mathcal{D}_l = \bigcup_{j=1}^{k} \mathcal{D}_l(\mathbf{u}_j, \boldsymbol{\phi}_j, \Omega_j),$$

where

$$\mathcal{D}_l(\mathbf{u}_j, \boldsymbol{\phi}_j, \Omega_j) = \left\{ \left(\boldsymbol{\phi}_j(\mathbf{s}) + e^{-l} \mathbf{R}_{\mathbf{u}_j(\mathbf{s})}^{-1} \cdot \begin{bmatrix} 0 \\ \mathbf{x} \end{bmatrix}, \mathbf{s} \right) \in \mathbb{T}^d \times \mathcal{U} : (\mathbf{x}, \mathbf{s}) \in \Omega_j \right\}.$$

Note that $l > 0$ parameterizes the size of the target. For any $\mathbf{s} \in \mathcal{U}$, let $\mathcal{T}(\mathbf{s}, \mathcal{D}_l)$ be the set of hitting times defined by

$$\mathcal{T}(\mathbf{s}, \mathcal{D}_l) := \{t > 0 : \varphi^t(\mathbf{s}) \in \mathcal{D}_l\}.$$

This is a discrete subset of $\mathbb{R}_{>0}$, and we label its elements by

$$0 < t_1(\mathbf{s}, \mathcal{D}_l) < t_2(\mathbf{s}, \mathcal{D}_l) < \cdots .$$

By Santalo's formula (cf. [2]), if $\mathbf{s} \in \mathcal{U}$ is such that the components of $\mathbf{f}(\mathbf{s})$ are not rationally related, then for any $n \in \mathbb{N}$, the normalized n-th return time to target $\mathcal{D}_l$ is

$$\frac{t_n(\mathbf{s}, \mathcal{D}_l)}{e^{(d-1)l} \cdot \overline{\sigma}(\mathbf{s})},$$

where $e^{(d-1)l} \cdot \overline{\sigma}(\mathbf{s})$ is the mean return time (cf. [3, Section 2]), and

$$\overline{\sigma}(\mathbf{s}) = \frac{1}{\sum_{j=1}^{k} |\Omega_j(\mathbf{s})| \mathbf{u}_j(\mathbf{s}) \cdot \mathbf{f}(\mathbf{s})}, \;\; \Omega_j(\mathbf{s}) = \{\mathbf{x} \in \mathbb{R}^{d-1} : (\mathbf{x}, \mathbf{s}) \in \Omega_j\}.$$

Definition 9.2. The smooth map $\mathbf{f} : \mathcal{U} \to \mathbf{S}_1^{d-1}$ is regular if the push forward of Lebesgue measure on $\mathcal{U}$ under $\mathbf{f}$ is absolutely continuous with respect to the Haar measure on $\mathbf{S}_1^{d-1}$.

The following is a corollary of Theorem 9.1:

Corollary 9.3. *Let $\mathcal{U}$ be a bounded open subset of $\mathbb{R}^{d-1}$. Let $\mathbf{f} : \mathcal{U} \to \mathbf{S}_1^{d-1}$ be a regular smooth map. Let $\boldsymbol{\varphi} : \mathcal{U} \to (\mathbb{R}^d)^k$ be a C^1 map. If for any $\mathbf{m} \in \mathbb{Z}^k \setminus \{\mathbf{0}\}$,*

$$|\{\mathbf{s} \in \mathcal{U} : \boldsymbol{\varphi}(\mathbf{s}) \cdot \mathbf{m} \in \mathbb{R}\mathbf{f}(\mathbf{s}) + \mathbb{Z}^d\}| = 0, \tag{9.3}$$

then for Lebesgue a.e. $\mathbf{s}\in\mathcal{U}$, $\mathbf{R}_{\mathbf{f}(\mathbf{s})}(Id,\boldsymbol{\varphi}(\mathbf{s}))\Gamma$ *is Birkhoff generic with respect to* $(X,\mu_X,D(e^{-l}))$.

Proof. By assumption, $\mathbf{f}$ is a regular smooth map from $\mathcal{U}$ to $\mathbf{S}_1^{d-1}$. By Sard's theorem, the set of critical points of $\mathbf{f}$ has Lebesgue measure 0. For any $\mathbf{s}_0\in\mathcal{U}$ which is not a critical point of $\mathbf{f}$, there exists an open neighborhood $\mathcal{V}$ of $\mathbf{s}_0$ such that $\mathbf{f}$ is a diffeomorphism of $\mathcal{V}$ to some open subset of $\mathbf{S}_1^{d-1}$.

Therefore, the map $\mathbf{s}\mapsto\mathbf{R}_{\mathbf{f}(\mathbf{s})}^{-1}\cdot\mathbf{e}_1=\mathbf{f}(\mathbf{s})$ has nonsingular differentials for all $\mathbf{s}\in\mathcal{V}$. Then we apply Theorem 9.1 to $\mathbf{E}_1(\mathbf{s})=\mathbf{R}_{\mathbf{f}(\mathbf{s})}$ and the corollary follows. ☐

Definition 9.4. The k-tuple of smooth functions $\boldsymbol{\phi}_1,\cdots\boldsymbol{\phi}_k:\mathcal{U}\to\mathbb{T}^d$ is $\boldsymbol{\theta}$-generic if for any $\mathbf{m}=(m_1,\cdots,m_k)\in\mathbb{Z}^k\setminus\{\mathbf{0}\}$, we have

$$\left|\left\{\mathbf{s}\in\mathcal{U}:\sum_{j=1}^k m_j(\boldsymbol{\phi}_j(\mathbf{s})-\boldsymbol{\theta}(\mathbf{s}))\in\mathbb{R}\mathbf{f}(\mathbf{s})+\mathbb{Z}^d\right\}\right|=0.$$

To state our theorem precisely, we need some preparations. Our notations follows from [3, Section 6]. Given $N\in\mathbb{N}$, denote $\overline{N}=\{1,\cdots,N\}$. For $j\in\{1,\cdots,k\}$ and $\mathbf{s}\in\mathcal{U}$, define $\mathfrak{R}_j(\mathbf{s})=\mathbf{R}_{\mathbf{f}(\mathbf{s})}\mathbf{R}_{\mathbf{u}_j(\mathbf{s})}^{-1}$. Let $\tilde{\mathfrak{R}}_j(\mathbf{s})$ be the matrix of the linear transformation

$$\mathbf{x}\mapsto\left(\mathfrak{R}_j(\mathbf{s})\begin{bmatrix}0\\ \mathbf{x}\end{bmatrix}\right)_\perp\in\mathbb{R}^{d-1}, \tag{9.4}$$

where $\mathbf{u}_\perp=(u_2,\cdots,u_d)^{tr}\in\mathbb{R}^{d-1}$ for $\mathbf{u}=(u_1,\cdots,u_d)^{tr}\in\mathbb{R}^d$.

For any $\mathbf{s}\in\mathcal{U}$, we define

$$\tilde{\Omega}_j(\mathbf{s}):=\overline{\sigma}(\mathbf{s})^{\frac{1}{d-1}}\tilde{\mathfrak{R}}_j(\mathbf{s})\Omega_j(\mathbf{s})\subset\mathbb{R}^{d-1}. \tag{9.5}$$

Let $G_1=SL_d(\mathbb{R})\ltimes\mathbb{R}^d$. For $g=(g',(\boldsymbol{\xi}_1,\cdots,\boldsymbol{\xi}_k))\in G$ and $j\in\{1,\cdots,k\}$, write $g^{[j]}=(g',\boldsymbol{\xi}_j)\in G_1$. Our main theorem of this section is the following.

Theorem 9.5. *Let* $\mathcal{U}$ *be a bounded open subset of* $\mathbb{R}^{d-1}$. *For* $1\le j\le k$, *let* $\mathbf{f},\boldsymbol{\theta},\mathbf{u}_j,\boldsymbol{\phi}_j$ *be given as in the beginning of this section. Let* Ω_j *be a bounded open subset of* $\mathbb{R}^{d-1}\times\mathcal{U}$. *For each* $j=1,\cdots,k$, *assume that*

(1) $|\mathbf{u}_j^{-1}(\{\mathbf{v}_0\})|=0$,
(2) $\mathbf{u}_j(\mathbf{s})\cdot\mathbf{f}(\mathbf{s})>0$ *for all* $\mathbf{s}\in\mathcal{U}$,
(3) for a.e. $\mathbf{s}\in\mathcal{U}$, *the boundary* $\partial\Omega_j(\mathbf{s})$ *has Lebesgue measure* 0.
(4) $|\Omega_j(\mathbf{s})|$ *is a smooth positive function of* $\mathbf{s}\in\mathcal{U}$.

Also assume that $\mathbf{f}$ *is regular and* $(\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_k)$ *is* $\boldsymbol{\theta}$*-generic. Then for any* $N \in \mathbb{N}$*, any* $T_n > 0$ *for* $n \in \overline{N}$*, the following holds: For a.e.* $\mathbf{s} \in \mathcal{U}$*,*

$$\lim_{L\to\infty} \frac{1}{L} \left| \left\{ l \in [0, L] : \frac{t_n(\mathbf{s}, \mathcal{D}_l)}{e^{(d-1)l} \cdot \overline{\sigma}(\mathbf{s})} \le T_n, \forall n \in \overline{N} \right\} \right|$$
$$= \mu_X \left(\left\{ g\Gamma \in X : \sum_{j=1}^{k} \# \left\{ \begin{bmatrix} t \\ \mathbf{x} \end{bmatrix} \in g^{[j]}\mathbb{Z}^d : 0 < t < T_n, \mathbf{x} \in -\tilde{\Omega}_j(\mathbf{s}) \right\} \ge n, \right.\right.$$
$$\left.\left. \forall n \in \overline{N} \right\} \right).$$

Note that in [3, Theorem 2], the authors proved that for each $n \in \mathbb{N}$, there is some random variable τ_n in $\mathbb{R}_{>0}$ such that the n-th normalized hitting time $t_n(\cdot, \mathcal{D}_l)/(e^{(d-1)l} \cdot \overline{\sigma}(\cdot))$ converges to τ_n in distribution as $l \to \infty$. Unlike [3], in Theorem 9.5 we are interested in the question that given a **fixed** initial $\mathbf{s}$, when the target is shrinking, how often the n-th normalized hitting time is bounded by some given constant.

For any $j \in \{1, \cdots, k\}$, any real numbers $Y < Z$, following [3, Eq. (8.9)], we define

$$\tilde{\mathbf{A}}_{j,Y,Z} = \left\{ \left(\begin{bmatrix} t \\ -\tilde{\mathfrak{R}}_j(\mathbf{s})\mathbf{x} \end{bmatrix}, \mathbf{s} \right) : (\mathbf{x}, \mathbf{s}) \in \Omega_j, \overline{\sigma}(\mathbf{s})Y < t \le \overline{\sigma}(\mathbf{s})Z \right\}.$$

Given any real numbers $Y_n < Z_n$ for $n \in \overline{N}$, following [3, Eq.(8.10)], we define

$$B[(Y_n), (Z_n)] = \{(g\Gamma, \mathbf{s}) \in G/\Gamma \times \mathcal{U} : \sum_{j=1}^{k} \#(\tilde{A}_{j,Y_n,Z_n}(\mathbf{s}) \cap g^{[j]} \cdot \mathbb{Z}^d) \ge n, \forall n \in \overline{N}\}.$$

where

$$\tilde{A}_{j,Y_n,Z_n}(\mathbf{s}) = \{\mathbf{x} \in \mathbb{R}^d : (\mathbf{x}, \mathbf{s}) \in \tilde{A}_{j,Y_n,Z_n}\}.$$

For any $\mathbf{s} \in \mathcal{U}$, denote

$$B[(Y_n), (Z_n)](\mathbf{s}) := \{g\Gamma \in G/\Gamma : (g\Gamma, \mathbf{s}) \in B[(Y_n), (Z_n)]\}.$$

Lemma 9.6. *[3, Lemma 17] For every* $\mathbf{s} \in \mathcal{U}$*, and* $B = B[(Y_n), (Z_n)]$*,* $\mu_X(\partial B(\mathbf{s})) = 0$*.*

Proof. By [3, Lemma 14,16], it suffices to prove that for every $j \in \{1, \cdots, k\}$ and $n \in \{1, \cdots, N\}$, $\partial \tilde{A}_{j,Y_n,Z_n}(\mathbf{s})$ has Lebesgue measure zero. Now since for any $Y < Z$, we have

$$\partial \tilde{A}_{j,Y,Z}(\mathbf{s}) = \left\{ \begin{bmatrix} t \\ -\tilde{\mathfrak{R}}_j(\mathbf{s})x \end{bmatrix} : \mathbf{x} \in \partial\Omega_j(\mathbf{s}), \overline{\sigma}(\mathbf{s})Y < t \leq \overline{\sigma}(\mathbf{s})Z \right\}$$
$$\bigcup \left\{ \begin{bmatrix} t \\ -\tilde{\mathfrak{R}}_j(\mathbf{s})x \end{bmatrix} : \mathbf{x} \in \overline{\Omega_j(\mathbf{s})}, t \in \{\overline{\sigma}(\mathbf{s})Y, \overline{\sigma}(\mathbf{s})Z\} \right\}.$$

By assumption, for a.e. $\mathbf{s} \in \mathcal{U}$, $|\partial\Omega_j(\mathbf{s})| = 0$, the lemma follows. □

We are now ready to prove Theorem 9.5.

Proof of Theorem 9.5. The proof of Theorem 9.5 is almost the same as the proof of [3, Theorem 2], except that here we use the equidistribution result for the average along a_t trajectory, while in [3], the equidistribution of a_t translation of the average over a bounded open subset in horospherical subgroup is used.

Let $\tilde{\boldsymbol{\varphi}} : \mathcal{U} \to (\mathbb{R}^d)^k$ be a map given by

$$\tilde{\boldsymbol{\varphi}}(\mathbf{s}) = (\boldsymbol{\phi}_1(\mathbf{s}) - \boldsymbol{\theta}(\mathbf{s}), \cdots, \boldsymbol{\phi}_k(\mathbf{s}) - \boldsymbol{\theta}(\mathbf{s})).$$

Since $(\boldsymbol{\phi}_1, \cdots, \boldsymbol{\phi}_k)$ is $\boldsymbol{\theta}$-generic, $\mathbf{f}$ is regular, assumption (9.3) in Corollary 9.3 are satisfied for the maps $\tilde{\boldsymbol{\varphi}}$ and $\mathbf{f}$, thus Corollary 9.3 applies.

Let $B = B[(Y_n), (Z_n)]$ for $Y_n, Z_n \in \mathbb{R}$ and $n \in \overline{N}$. Since by Lemma 9.6, $\mu_X(\partial B(\mathbf{s})) = 0$, for all $\mathbf{s} \in \mathcal{U}$, we have for a.e. $\mathbf{s} \in \mathcal{U}$,

$$\lim_{L\to\infty} \frac{1}{L} \int_0^L \chi_{B(\mathbf{s})}(D(e^{-l})\mathbf{R}_{\mathbf{f}(\mathbf{s})}(Id, \tilde{\varphi}(\mathbf{s})))dl = \mu_X(B(\mathbf{s})).$$

Then the rest of the proof follows from the proof of [3, Theorem 2]. □

Acknowledgment

I would like to thank Yitwah Cheung, Nimish Shah, Ronggang Shi and Barak Weiss for helpful discussions and comments. I thank Jim Cogdell for pointing out some inaccuracies in the draft. I also thank anonymous referee for pointing out many inaccuracies of the paper, as well as providing a lot of helpful suggestions to improve this article.

References

[1] Yves Benoist and Jean-Francois Quint. Random walks on finite volume homogeneous spaces. *Invent. Math.*, **187**(1):37–59, 2012. MR2874934

[2] N. Chernov. Entropy, Lyapunov exponents, and mean free path for billiards. *J. Statist. Phys.*, **88**(1-2):1–29, 1997. MR1468377

[3] Carl P. Dettmann, Jens Marklof, and Andreas Strömbergsson. Universal hitting time statistics for integrable flows. *J. Stat. Phys.*, **166**(3-4):714–749, 2017. MR3607587

[4] Manfred Einsiedler and Thomas Ward. *Ergodic theory with a view towards number theory*, volume 259 of *Graduate Texts in Mathematics*. Springer-Verlag London, Ltd., London, 2011. MR2723325

[5] Alex Eskin and Gregory Margulis. Recurrence properties of random walks on finite volume homogeneous manifolds. In *Random walks and geometry*, pages 431–444. Walter de Gruyter, Berlin, 2004. MR2087794

[6] Alex Eskin, Gregory Margulis, and Shahar Mozes. Upper bounds and asymptotics in a quantitative version of the Oppenheim conjecture. *Ann. of Math. (2)*, **147**(1):93–141, 1998. MR1609447

[7] Alex Eskin, Gregory Margulis, and Shahar Mozes. Quadratic forms of signature $(2,2)$ and eigenvalue spacings on rectangular 2-tori. *Ann. of Math. (2)*, **161**(2):679–725, 2005. MR2153398

[8] Krzysztof Frączek, Ronggang Shi, and Corinna Ulcigrai. Genericity on curves and applications: pseudo-integrable billiards, Eaton lenses and gap distributions. *J. Mod. Dyn.*, **12**:55–122, 2018. MR3808209

[9] Michael Hochman and Pablo Shmerkin. Equidistribution from fractal measures. *Invent. Math.*, **202**(1):427–479, 2015. MR3402802

[10] Bernard Host. Nombres normaux, entropie, translations. *Israel J. Math.*, **91**(1-3):419–428, 1995. MR1348326

[11] D. Y. Kleinbock and G. A. Margulis. Flows on homogeneous spaces and Diophantine approximation on manifolds. *Ann. of Math. (2)*, **148**(1):339–360, 1998. MR1652916

[12] Dmitry Kleinbock, Nicolas de Saxcé, Nimish A Shah, and Pengyu Yang. Equidistribution in the space of 3-lattices and dirichlet-

improvable vectors on planar lines. *arXiv preprint arXiv:2106.08860*, 2021.

[13] Dmitry Kleinbock, Ronggang Shi, and Barak Weiss. Pointwise equidistribution with an error rate and with respect to unbounded functions. *Math. Ann.*, **367**(1-2):857–879, 2017. MR3606456

[14] A. Malcev. On the representation of an algebra as a direct sum of the radical and a semi-simple subalgebra. *C. R. (Doklady) Acad. Sci. URSS (N.S.)*, **36**:42–45, 1942. MR0007397

[15] Gregory Margulis. Problems and conjectures in rigidity theory. In *Mathematics: frontiers and perspectives*, pages 161–174. Amer. Math. Soc., Providence, RI, 2000. MR1754775

[16] Gregory Margulis and Amir Mohammadi. Quantitative version of the Oppenheim conjecture for inhomogeneous quadratic forms. *Duke Math. J.*, **158**(1):121–160, 2011. MR2794370

[17] Shahar Mozes. Epimorphic subgroups and invariant measures. *Ergodic Theory Dynam. Systems*, **15**(6):1207–1210, 1995. MR1366316

[18] M. S. Raghunathan. *Discrete subgroups of Lie groups*. Springer-Verlag, New York-Heidelberg, 1972. Ergebnisse der Mathematik und ihrer Grenzgebiete, Band 68. MR0507234

[19] Marina Ratner. On Raghunathan's measure conjecture. *Ann. of Math. (2)*, **134**(3):545–607, 1991. MR1135878

[20] Marina Ratner. Raghunathan's topological conjecture and distributions of unipotent flows. *Duke Math. J.*, **63**(1):235–280, 1991. MR1106945

[21] Nimish A. Shah. Limit distributions of expanding translates of certain orbits on homogeneous spaces. *Proc. Indian Acad. Sci. Math. Sci.*, **106**(2):105–125, 1996. MR1403756

[22] Nimish A. Shah. Equidistribution of expanding translates of curves and Dirichlet's theorem on Diophantine approximation. *Invent. Math.*, **177**(3):509–532, 2009. MR2534098

[23] Nimish A. Shah. Limiting distributions of curves under geodesic flow on hyperbolic manifolds. *Duke Math. J.*, **148**(2):251–279, 2009. MR2524496

[24] Ronggang Shi. Expanding cone and applications to homogeneous dynamics. *International Mathematics Research Notices*, 2015. MR4105521

[25] Ronggang Shi. Pointwise equidistribution for one parameter diagonalizable group action on homogeneous space. *Trans. Amer. Math. Soc.*, **373**(6):4189–4221, 2020. MR4251297

[26] Ronggang Shi and Barak Weiss. Invariant measures for solvable groups and Diophantine approximation. *Israel J. Math.*, **219**(1):479–505, 2017. MR3642031

[27] George Tomanov. Actions of maximal tori on homogeneous spaces. In *Rigidity in dynamics and geometry (Cambridge, 2000)*, pages 407–424. Springer, Berlin, 2002. MR1919414

[28] Pengyu Yang. Equidistribution of expanding translates of curves and Diophantine approximation on matrices. *Invent. Math.*, **220**(3):909–948, 2020. MR4094972

[29] Jon Chaika and Alex Eskin. Every flat surface is Birkhoff and Oseledets generic in almost every direction. *J. Mod. Dyn.*, **9**:1–23, 2015.

Han Zhang
Yau Mathematical Sciences Center
Tsinghua University
100084 Beijing
China
E-mail: hanzhang3906@tsinghua.edu.cn