QUANTITATIVE BEHAVIOR OF UNIPOTENT FLOWS AND AN EFFECTIVE AVOIDANCE PRINCIPLE

By

ELON LINDENSTRAUSS,* GREGORII MARGULIS,† AMIR MOHAMMADI‡ AND NIMISH A. SHAH§

Abstract. We give an effective bound on how much time orbits of a unipotent group U on an arithmetic quotient G/Γ can stay near homogeneous subvarieties of G/Γ corresponding to $\mathbb{Q}$-subgroups of G. In particular, we show that if such a U-orbit is moderately near a proper homogeneous subvariety of G/Γ for a long time, it is very near a different homogeneous subvariety. Our work builds upon the linearization method of Dani and Margulis.

Our motivation in developing these bounds is in order to prove quantitative density statements about unipotent orbits, which we plan to pursue in a subsequent paper. New qualitative implications of our effective bounds are also given.

1 Introduction

A basic challenge in homogeneous dynamics is the quantitative understanding of behavior of orbits, in particular of unipotent orbits. In this paper, we give a sharper form of the Dani–Margulis linearization method [17], that allows to control the amount of time a unipotent trajectory spends near invariant subvarieties of a homogeneous space; related techniques were also considered by Shah in [46].

One important use of this technique is to be able to relate the behavior of individual unipotent (or unipotent-like, see, e.g., [20]) orbits with Ratner's landmark measure classification result [42]. This result says that any measure invariant and ergodic under a connected unipotent group U on a homogeneous space G/Γ has to be in one of countably many families; for the cases of G/Γ and unipotent group U we will consider, the group U acts ergodically on G/Γ with respect to the uniform

*E. L. acknowledges support by the ISF 891/15. Parts of this work was written when E. L. was a member at the IAS in Princeton, supported by NSF grant DMS-1638352

†G. M. acknowledges support by the NSF grant DMS-1265695

‡A. M. acknowledges support by the NSF grants DMS-1724316, 1764246, 1128155, and Alfred P. Sloan Research Fellowship.

§N. S. acknowledges support by the NSF grant DMS-1700394.

JOURNAL D'ANALYSE MATHÉMATIQUE, Vol. TBD (2023)
DOI 10.1007/s11854-023-0308-xPi10.1007/s11854-023-0309-9

measure on G/Γ, hence this uniform measure is one of the countably many possibilities. All other ergodic measures will be supported on proper homogeneous subvarieties of G/Γ. If one is able to show, using linearization or a different technique, that a given collection of orbits under consideration of increasing size do not spend much time next to any one of these countably many families of not fully supported invariant measures, then one is able to conclude using the measure classification that this collection of orbits tends to become equidistributed in G/Γ. We note that for the special case when one looks at a single orbit of a one-parameter unipotent group Ratner was able to establish such an avoidance by a different argument in [41].

However, the linearization technique of [17] is interesting in its own sake, and in fact originated in work of Dani and Margulis before the proof of measure classification such as [15] in order to give a purely topological proof of the Raghunathan Conjecture for the action of a generic one-parameter unipotent group on $\mathrm{SL}_3(\mathbb{R})/\mathrm{SL}_3(\mathbb{Z})$. Notably, unlike the techniques of Ratner used to prove the measure classification result in [42, 43] or the techniques used to give a related but different proof of this result by Margulis and Tomanov in [35], which in particular rely on results such as the pointwise ergodic theorem and Luzin's theorem which are hard to make effective, the linearization technique relies essentially only on the polynomial nature of the action: not only are the elements of the unipotent group (considered as a subgroup of some SL_N) polynomial, but the same holds for any linear representation of G.

In a subsequent paper we plan to make essential use of the results of this paper in order to provide a fully effective orbit closure classification theorem for unipotent flows on arithmetic homogenous spaces (albeit with very slow rates). We provide some other applications of independent interest here.

Somewhat surprisingly, many of the most striking applications of the theory of unipotent flows to number theory require working in the S-arithmetic context, i.e., for products of real and p-adic groups (here we prefer to use Σ for the set of places instead of the more traditional S, so we refer to this case as the Σ-arithmetic case). Ratner's measure classification result was generalized to this context by Ratner [43] and by Margulis and Tomanov [35]; the linearization techniques of Dani and Margulis were adapted to this context by Tomanov and by Gorodnik and Oh in [50, 25]. With a view to potential applications, our paper is written for Σ-arithmetic quotients. For simplicity we state here in the introduction the main results in the special case where we consider the action of a one-parameter unipotent group and consider only real algebraic groups, deferring stating the slightly more technical general statements to §3. We emphasize that in order to get a fully explicit and effective result, we assume that the lattice is arithmetic. By the Margulis Arithmeticity Theorem this assumption automatically holds for a large class of groups G, and in any case arithmetic quotients are the only type of quotients G/Γ that seem to appear in number theoretic applications.

The non-divergence result of Margulis [32], which were sharpened by Dani in [13], are effective and have been given a very explicit and effective form by Kleinbock and Margulis in [29]. The technique of linearization is related, but we are not aware of an effective treatment of the main results in [17], and doing so in this paper relies on employing an effective Nullstellensatz by Masser and Wüstholz [36, Thm. IV] as well as some local non-vanishing theorems related to Lojasiewicz inequality by Brownawell and Greenberg [11, 23, 24]. Moreover, since we are not content with analyzing what happens in the limit, we need to be able to analyze trajectories that are somewhat near a subvariety for a long time, which is an issue that has not been discussed in previous works on the linearization method.

Let $\mathbf{G}$ be a connected $\mathbb{Q}$-group and put $G = \mathbf{G}(\mathbb{R})$. We assume Γ is an arithmetic lattice in G. More specifically, fix an embedding $\iota : \mathbf{G} \to \mathrm{SL}_N$, defined over $\mathbb{Q}$ so that $\iota(\Gamma) \subset \mathrm{SL}_N(\mathbb{Z})$. Using ι, we identify $\mathbf{G}$ with $\iota(\mathbf{G}) \subset \mathrm{SL}_N$ and hence $G \subset \mathrm{SL}_N(\mathbb{R})$. Note that using the restriction of scalars from number fields to $\mathbb{Q}$, our results are applicable also in the case of groups defined over a general number field.

Let $U = \{u(t) : t \in \mathbb{R}\} \subset G$ be a one parameter unipotent subgroup of G, and put $X = G/\Gamma$.

Define the following family

$$\mathcal{H} = \{\mathbf{H} \subset \mathbf{G} : \mathbf{H} \text{ is a connected } \mathbb{Q}\text{-subgroup and } \mathrm{R}(\mathbf{H}) = \mathrm{R}_u(\mathbf{H})\}$$

where $\mathrm{R}(\mathbf{H})$ (resp. $\mathrm{R}_u(\mathbf{H})$) denotes the solvable (resp. unipotent) radical of $\mathbf{H}$. Alternatively, $\mathbf{H} \in \mathcal{H}$ if and only if $\mathbf{H}$ is a connected $\mathbb{Q}$-subgroup which is generated by unipotent subgroups over the algebraic closure of $\mathbb{Q}$. By a theorem of Borel and Harish-Chandra, $\mathbf{H}(\mathbb{R}) \cap \Gamma$ is a lattice in $\mathbf{H}(\mathbb{R})$ for any $\mathbf{H} \in \mathcal{H}$.

Our standing assumption is that $\mathbf{G} \in \mathcal{H}$ and
that U is not contained in $\mathbf{H}(\mathbb{R})$ for any proper normal $\mathbf{H} \lhd \mathbf{G}$.

For any $\mathbf{H} \in \mathcal{H}$ put $H = \mathbf{H}(\mathbb{R})$, we also write $H \in \mathcal{H}$. Define

$$N_G(U, H) := \{g \in G : Ug \subset gH\}.$$

Note that $N_G(U, H)$ is an $\mathbb{R}$-subvariety of G. Moreover, if $H \lhd G$ and $U \subset H$, then $N_G(U, H) = G$.

Put

$$\mathcal{S}(U) = \Big(\bigcup_{\substack{H\in\mathcal{H}\\ H\neq G}} N_G(U,H)\Big)/\Gamma \quad \text{and} \quad \mathcal{G}(U) = X \setminus \mathcal{S}(U).$$

Following Dani and Margulis, [17], points in $\mathcal{S}(U)$ are called *singular* with respect to U, and points in $\mathcal{G}(U)$ are called **generic** with respect to U—these are, a priori, different from the measure theoretically generic points in the sense of Furstenberg for the action of U on X equipped with the G-invariant probability measure (see, e.g., [22, p. 98] for a definition); however, any measure theoretically generic point is generic in this explicit sense as well. In the early 1990's Ratner proved the remarkable result [41], previously conjectured by Raghunathan, that for every $x \in \mathcal{G}(U)$ we have $\overline{Ux} = X$. Prior to Ratner's proof of the general case of Raghunathan's Conjecture in [41], important cases of Raghunathan's Conjecture were proven in [33, 14, 15].

Roughly speaking our main theorems guarantee that unless there is an explicit obstruction, most points on a unipotent orbit are generic. We begin with the following statement which follows from our main effective theorems in this paper.

1.1 Theorem. *Let $\eta > 0$. Let $\{H_i : 1 \leq i \leq r\} \subset \mathcal{H}$ be a finite subset consisting of proper subgroups, and for each $1 \leq i \leq r$ let $\mathcal{C}_i \subset N_G(U, H_i)$ be a compact subset. There exists an open set $\mathcal{O} = \mathcal{O}(\eta, \{H_i\}, \{\mathcal{C}_i\})$ so that $X \setminus \mathcal{O}$ is compact and disjoint from $\bigcup_i \mathcal{C}_i\Gamma/\Gamma$, and so that for every $x \in \mathcal{G}(U)$ there exists some $T_0 = T_0(\eta, \{H_i\}, \{\mathcal{C}_i\}, x)$ so that for all $T \geq T_0$ we have*

$$|\{t \in [-T, T] : u(t)x \in \mathcal{O}\}| < \eta T$$

We note that this theorem can also be deduced by combining Ratner's measure classification theorem, [42], and results in [17]; however this would only give a non-effective proof of the above statement. Without appealing to [42] and only utilizing statements in [17] (where the proof is essentially effective), one does not get uniformity as in Theorem 1.1: indeed, from the argument in [17] the set $\mathcal{O}$ above will depend on the initial point x. This distinction is similar to the difference between the non-divergence statement given by Dani in [13] and the dependence on the base point in Margulis' [32].

1.2 Effective versions of linearization. The main theorems in this paper yield a more precise and effective information about the compact set $X \setminus \mathcal{O}$ appearing above, with a polynomial dependence on the relevant parameters. We need some preliminary notation before we can state our main results.

Let $\| \ \|_\infty$ (or simply $\| \ \|$) denote the max norm on $\mathfrak{sl}_N(\mathbb{R})$ with respect to the standard basis; this induces a family of norms on $\wedge\mathfrak{sl}_N(\mathbb{R})$, which we continue to denote by $\| \ \|_\infty$ (or simply $\| \ \|$). We also let $\| \ \|$ be a norm on $\mathrm{SL}_N(\mathbb{R})$ fixed once and for all. For every $g \in \mathrm{SL}_N(\mathbb{R})$, in particular for any $g \in G$, we let

$$|g| = \max\{\|g\|, \|g^{-1}\|\}.$$

Let $\mathfrak{g} = \mathrm{Lie}(G)$ and put $\mathfrak{g}(\mathbb{Z}) := \mathfrak{g} \cap \mathfrak{sl}_N(\mathbb{Z})$.

For every $\eta > 0$, we define

$$X_\eta = \{g\Gamma \in X : \min_{0 \neq v \in \mathfrak{g}(\mathbb{Z})} \|\mathrm{Ad}(g)v\| \geq \eta\}.$$

For every $\eta > 0$ the space X_η is compact (see §2.7 and Lemma 2.8), and

$$\bigcup_{\eta > 0} X_\eta = G/\Gamma.$$

Recall that U is a one parameter unipotent subgroup of G. Fix a $z \in \mathfrak{g}$ with $\|z\| = 1$ so that

$$U = \{u(t) = \exp(tz) : t \in \mathbb{R}\}. \tag{1.1}$$

Let $\mathbf{H} \in \mathcal{H}$ be a nontrivial proper subgroup of $\mathbf{G}$ and put

$$\rho_H := \wedge^{\dim \mathbf{H}} \mathrm{Ad} \quad \text{ and } \quad V_H := \wedge^{\dim \mathbf{H}} \mathfrak{g}.$$

The representation ρ_H is defined over $\mathbb{Q}$.

Let $\mathbf{v_H}$ be a primitive integral vector in $\wedge^{\dim \mathbf{H}} \mathrm{Lie}(\mathbf{G})$ corresponding to the Lie algebra of $\mathbf{H}$, i.e., we fix a $\mathbb{Z}$-basis for $\mathrm{Lie}(H) \cap \mathfrak{sl}_N(\mathbb{Z})$, and let $\mathbf{v_H}$ be the corresponding wedge product.

We also view $\mathbf{v_H}$ as an element in $\wedge^{\dim \mathbf{H}} \mathfrak{g}$; in order to put an emphasis on the local nature of this vector, we will denote it by v_H. Define

$$\eta_H(g) := \rho_H(g) v_H \quad \text{for every } g \in G.$$

With this notation, for an element $\mathbf{H} \in \mathcal{H}$, we have

$$N_G(U, H) = \{g \in G : z \wedge \eta_H(g) = 0\}.$$

Note that $N_G(U, H)$ is a variety and could change drastically under small perturbations of U. However, effective notions must be stable under small perturbations. We will use the above finite-dimensional representations to give an effective notion of generic points. The integer vector $\mathbf{v_H}$ also allows us to give a notion of arithmetic complexity for subgroups in $\mathcal{H}$ by defining the height of the group $\mathbf{H}$ to be

$$\mathrm{ht}(\mathbf{H}) := \|\mathbf{v_H}\|_\infty. \tag{1.2}$$

Thus the height of a $\mathbb{Q}$-group $\mathbf{H}$ is given by the height of the corresponding point in the Grassmanian of Lie($\mathbf{G}$); cf. [4, §1.5].

The following definition will play a crucial role in this paper.

1.3 Definition. Let $\varepsilon : \mathbb{R}^+ \to (0, 1)$ be a monotone decreasing function, and let $t \in \mathbb{R}^+$. Let z be as in (1.1). A point $g\Gamma$ is called **(ε, t)-Diophantine** for the action of U if for all $\mathbf{H} \in \mathcal{H}$ with $\{e\} \neq \mathbf{H} \neq \mathbf{G}$

$$\|z \wedge \eta_H(g)\| \geq \varepsilon(\|\eta_H(g)\|) \quad \text{if } \|\eta_H(g)\| < e^t. \tag{1.3}$$

A point is **ε-Diophantine** if it is (ε, t)-Diophantine for all $t > 0$.

Note that this is a condition on the pair $(U, g\Gamma)$. Unless $U < \mathbf{H}(\mathbb{R})$ for some (proper) $\mathbf{H} \lhd \mathbf{G}$, the set $\mathcal{G}(U)$ is non-empty, and moreover any $x \in \mathcal{G}(U)$ is ε-Diophantine for some ε as above. In most interesting examples the singular set $\mathcal{S}(U)$ is a dense subset of X. Therefore, $\mathcal{G}(U)$ is usually a G_δ-set without any interior points. For any $t \in \mathbb{R}^+$, on the other hand, the set of (ε, t)-Diophantine points in Definition 1.3 is a nice closed set with interior points (indeed, is the closure of its interior points).

We can now state our main theorem (in slightly simplified form, see Theorem 3.2 below for the full version with all the features):

1.4 Theorem. *There are constants $A, D > 1$ depending only on N, and $E_1 > 1$ depending on N, G and Γ, so that the following holds. Let $g \in G$, $t > 0$, $k \geq 1$, and $0 < \eta < 1/2$. Assume $\varepsilon : \mathbb{R}^+ \to (0, 1)$ satisfies for any $s > 0$ that*

$$\varepsilon(s) \leq \eta^A s^{-A}/E_1.$$

Then at least one of the following three possibilities holds.

(1) $|\{\xi \in [-1, 1] : u(e^k\xi)g\Gamma \notin X_\eta$ *or*

$$u(e^k\xi)g\Gamma \text{ is not } (\varepsilon, t)\text{-Diophantine}\}| < E_1\eta^{1/D}.$$

(2) *There exist a nontrivial proper subgroup $\mathbf{H} \in \mathcal{H}$ of*

$$\mathrm{ht}(\mathbf{H}) \leq E_1(|g|^A + e^{At})\eta^{-A}$$

so that the following hold for all $\xi \in [-1, 1]$:

$$\|\eta_H(u(e^k\xi)g)\| \leq E_1(|g|^A + e^{At})\,\eta^{-A},$$
$$\|z \wedge \eta_H(u(e^k\xi)g)\| \leq E_1 e^{-k/D}(|g|^A + e^{At})\,\eta^{-A},$$

where z is as in (1.1).

(3) *There exist a nontrivial proper normal subgroup* $\mathbf{H} \lhd \mathbf{G}$ *of*

$$\mathrm{ht}(\mathbf{H}) \le E_1 e^{At} \eta^{-A}$$

so that

$$\| z \wedge v_H \| \le \varepsilon (\mathrm{ht}(\mathbf{H})^{1/A} \eta / E_1)^{1/A}.$$

Similar to [17], the proof of Theorem 1.4, and its Σ-arithmetic analogue, relies on the polynomial like behavior of the unipotent orbits. However, in addition to being polynomially effective, our results here also differ from [17] in the following sense. They provide a compact subset of $\mathcal{G}(U)$ which is independent of the base point and to which a unipotent orbit returns unless there is an algebraic obstruction; this uniformity is used essentially in Theorem 1.1 and Theorem 1.5. Regarding non-divergence properties of unipotent orbits, such uniformity is well known and is due to Dani (see [13, 16]), but in this context it is new.

These features have been made possible using two main ingredients. First is the use of an effective notion of a generic point, Definition 1.3. The second ingredient is the use of a group $\mathbf{M}_{\mathbf{H}}$, see §4.7, to control the speed of unipotent orbits in the representation space V_H; this group does not feature in the analysis in [17].

Using Theorem 1.4 one can give a topological analogue of a result of Mozes and Shah [39]. To deal with groups with infinitely many normal $\mathbb{Q}$-subgroups we need the following definition:

For any $T > 0$, put

$$\sigma(T) = \min(\{1\} \cup \{ \| z \wedge v_H \| : \mathbf{H} \in \mathcal{H}, \mathbf{H} \lhd \mathbf{G}, \mathrm{ht}(\mathbf{H}) \le T, \{1\} \ne \mathbf{H} \ne \mathbf{G}\}).$$

1.5 Theorem. *There exists some* $D > 1$ *depending on* N *and* $E_1 > 0$ *depending on* N, G, *and* Γ *so that the following holds. Let* $0 < \eta < 1/2$.

Let $\{x_m\}$ *be a sequence of points in* X, *and let* $T_m \to \infty$ *be a sequence of real numbers. For each* m *let* $I_m \subset [-T_m, T_m]$ *be a measurable set with measure* $> \eta T_m$. *Let*

$$Y = \bigcap_{k \ge 1} \overline{\bigcup_{m \ge k} \{ u(t) x_m : t \in I_m \}}.$$

Then exactly one of the following holds.

(1) Y *contains an* ε*-Diophantine point for*

$$\varepsilon(s) = (\eta s^{-1} \sigma(E_1^A \eta^{-A} s^A)/2E_1)^A.$$

(2) *There exists a countable (or finite) collection*

$$\mathcal{F} = \{(\mathbb{H}_i, L_i) : i \in I\} \subset \mathcal{H} \times \mathbb{R}^+$$

so that if

$$Y_i = \{g \in N(U, H_i) : \|\eta_{H_i}(g)\| \le L_i\}\Gamma/\Gamma$$

then

(a) $Y \subset \bigcup_{i\in I} Y_i$

(b) *for any* $\beta > 0$

$$\#\{i \in I : Y \cap X_\beta \cap Y_i \ne \emptyset\} < \infty.$$

As we shall see in Corollary 4.10.1 below, for any $\mathbf{H} \in \mathcal{H}$ and $L > 0$ the set

$$Y = \{g \in N(U, H) : \|\eta_H(g)\| \le L\}\Gamma/\Gamma$$

is a closed (though in general not compact) subset of X. For instance, for $G = \mathrm{SL}_2(\mathbb{R})$ and $\Gamma = \mathrm{SL}_2(\mathbb{Z})$, if we take $\mathbf{H}$ to be the stabilizer of the vector $\binom{0}{1} \in \mathbb{Z}^2$, $U = H = \mathbf{H}(\mathbb{R})$ and define Y as above, then Y is the union of all periodic U-orbits of period $\le L$.

Theorem 1.5 is related to [17, Thm. 4]. Specifically in that paper it is proved that if one assumes that sequence $\{x_m\}$ converges to a point in $\mathcal{G}(U)$, then a less precise form of (1) of Theorem 1.5, namely that Y contains a point in $\mathcal{G}(U)$, holds.

1.6 Friendly measures. In this section we discuss generalizations of Theorem 1.5 to the class of friendly measures which were studied in [28].

Let (Y, d) be a σ-compact metric space; for every $y \in Y$ and $r > 0$, let $B(y, r)$ denote the open ball of radius r centered at y. Let μ be a locally finite Borel measure on Y. If $A > 0$ and $O \subset Y$ is an open subset, the measure μ is called A-**Federer** on O if for all $y \in \mathrm{supp}(\mu) \cap O$ one has

$$\frac{\mu(B(y, 3r))}{\mu(B(y, r))} < A$$

whenever $B(y, 3r) \subset O$.

Let $Y = \mathbb{R}$ be equipped with the standard metric. Given a point $a \in \mathbb{R}$ and $\delta > 0$, we let $I_\delta(a) = (a - \delta, a + \delta)$. Given $c, \alpha > 0$ and an open subset $O \subset \mathbb{R}$, we say μ is (c, α)-**absolutely decaying** on O if for every non-empty open interval $J \subset O$ centered in $\mathrm{supp}(\mu)$, every point $a \in \mathbb{R}$, and every $\delta > 0$ we have

$$\mu(J \cap I_\delta(a)) < c\left(\frac{\delta}{r}\right)^\alpha \mu(J) \tag{1.4}$$

where J has length $2r$, see [28, Lemma 2.2].

We will say a measure μ on $\mathbb{R}$ is **uniformly friendly** if μ is A-Federer and (c, α)-absolutely decaying for some $A, c, \alpha > 0$.

Let the notation be as in §1.2; in particular,

$$U = \{u(t) = \exp(tz) : t \in \mathbb{R}\}$$

for some nilpotent element $z \in \mathfrak{g}$ with $\|z\| = 1$.

1.7 Theorem. *Let μ be a uniformly friendly measure on $\mathbb{R}$. There exists some $D > 1$ depending on N and μ, and $E_1 > 0$ depending on N, G, Γ, and μ so that the following holds.*

Let $0 < \eta < 1/2$. Let $\{x_m\}$ be a sequence of points in X, $0 < \eta < 1/2$, and let $k_m \to \infty$ be a sequence of real numbers. For each m let $I_m \subset [-1, 1]$ be a measurable set with $\mu(I_m) > \eta\, \mu([-1, 1])$. Let

$$Y = \bigcap_{\ell \geq 1} \overline{\bigcup_{m \geq \ell} \{u(e^{k_m} t)x_m : t \in I_m\}}.$$

Then exactly one of the following holds:

(1) *Y contains an ε-Diophantine point for*

$$\varepsilon(s) = (\eta s^{-1} \sigma(E_1^A \eta^{-A} s^A)/2E_1)^A.$$

(2) *There exists a countable (or finite) collection*

$$\mathcal{F} = \{(\mathcal{H}_i, L_i) : i \in I\} \subset \mathcal{H} \times \mathbb{R}^+$$

so that if

$$Y_i = \{g \in N(U, H) : \|\eta_{H_i}(g)\| \leq L_i\}\Gamma/\Gamma,$$

then

(a) *$Y \subset \bigcup_{i \in I} Y_i$,*

(b) *for any $\beta > 0$*

$$\#\{i \in I : Y \cap X_\beta \cap Y_i \neq \emptyset\} < \infty.$$

See §9 for a more detailed discussion of this generalization.

Acknowledgements. We would like to thank M. Einsiedler, H. Oh, and A. Wieser for their helpful comments on earlier drafts of this paper.

2 Notation

2.1 Let $\mathcal{S} = \{\infty\} \cup \{p : p \text{ is a prime}\}$ denote the set of places of $\mathbb{Q}$. We let $\mathcal{S}_f = \mathcal{S} \setminus \{\infty\}$ denote the set of finite places in $\mathcal{S}$. For every $v \in \mathcal{S}$ let $\mathbb{Q}_v$ be the completion of $\mathbb{Q}$ at v ; we often write $\mathbb{R}$ for $\mathbb{Q}_\infty$.

For every $p \in \mathcal{S}_f$, we let $\mathbb{C}_p$ be the completion of the algebraic closure, $\overline{\mathbb{Q}}_p$, of $\mathbb{Q}_p$ with respect to the p-adic norm. The field $\mathbb{C}_p$ is a complete and algebraically closed field.

Given a finite subset $\Sigma \subset \mathcal{S}$, we put $\Sigma_f := \Sigma \setminus \{\infty\}$; also set $\mathbb{Q}_\Sigma = \prod_{v\in\Sigma}\mathbb{Q}_v$. Given an element $r \in \mathbb{Q}_\Sigma$, we put $|r|_\Sigma = \max_{v\in\Sigma}|r|_v$.

For any $p \in \mathcal{S}_f$, let $\mathbb{Z}_p$ denote the ring of p-adic integers in $\mathbb{Q}_p$. The ring of Σ-integers in $\mathbb{Q}$ is as usual denoted by $\mathbb{Z}_\Sigma$.

Given a $\mathbb{Q}$-variety $\mathbf{Y}$, we put $Y_v = \mathbf{Y}(\mathbb{Q}_v)$; given a finite subset $\Sigma \subset \mathcal{S}$, we also write Y_Σ, or simply Y if there is no confusion, for $\prod_\Sigma Y_v$.

For any $\mathbb{Q}$-variety $\mathbf{Y}$, we denote by $\dim \mathbf{Y}$ the dimension in the algebro-geometric sense. In particular $\dim \mathbf{G}$ is the dimension of $\mathbf{G}$ as an algebraic group. Note that $\dim_{\mathbb{Q}_v} Y_v$, the dimension of Y_v as a $\mathbb{Q}_v$-manifold, equals $\dim \mathbf{Y}$; see, e.g., [34, Ch. I, §2.5]. We also put

$$\dim Y_\Sigma := \sum_\Sigma \dim_{\mathbb{Q}_p} Y_p = (\#\Sigma)\dim \mathbf{Y}.$$

Given a $\mathbb{Q}$-group, $\mathbf{H}$, we denote by $\mathrm{Lie}(\mathbf{H})$ the Lie algebra of $\mathbf{H}$. We will use lower case gothic letters to denote the Lie algebra of $\mathbf{H}$ over various local fields, e.g., $\mathfrak{h}_p = \mathrm{Lie}(\mathbf{H}(\mathbb{Q}_p))$; similarly, we write $\mathfrak{h}_\Sigma$, or simply $\mathfrak{h}$, for $\oplus_\Sigma \mathfrak{h}_p$.

The space $\mathfrak{h}_\Sigma$ is a $\mathbb{Q}_\Sigma$-module; and the notation rw for $r \in \mathbb{Q}_\Sigma$ and $w \in \mathfrak{h}_\Sigma$ in the sequel refers to this module structure.

Given a natural number $m \le \dim(\mathrm{Lie}(\mathbf{H}))$, we write $\wedge^m\mathfrak{h}$ or $\wedge^m\mathfrak{h}_\Sigma$ to denote $\oplus_\Sigma(\wedge^m\mathfrak{h}_v)$.

For any (compact) subset $\mathsf{K} \subset \mathbb{Q}_v^m$ and any $\delta > 0$, we let $\mathcal{N}_\delta(\mathsf{K})$ denote the δ-neighborhood of K. Also let $|\mathsf{K}|$ denote the Haar measure of K.

Let $\mathbf{H}$ be a $\mathbb{Q}_\Sigma$-group and put $H = \mathbf{H}(\mathbb{Q}_\Sigma)$. Given a subset $B \subset H$, we define

$$Z_H(B) = \{g \in H : gb = bg \text{ for all } b \in B\}.$$

Given two subsets $B_1, B_2 \subset H$, we define

$$N_H(B_1, B_2) = \{g \in H : g^{-1}B_1 g \subset B_2\},$$

and put $N_H(B) := N_H(B, B)$ for any $B \subset H$.

2.2 For any place $v \in \mathcal{S}$, let $\|\ \|_v$ denote the max norm, with respect to the standard basis, on $\mathfrak{sl}_N(\mathbb{Q}_v)$ and on $\wedge\mathfrak{sl}_N(\mathbb{Q}_v)$. Given a finite subset $\Sigma \subset \mathcal{S}$, the norm $\|\ \|_\Sigma$ (or simply $\|\ \|$) is defined by

$$\|\ \|_\Sigma = \max_{v\in\Sigma}\|\ \|_v.$$

We let d denote the induced metric on the exterior algebra $\wedge\mathfrak{sl}_N(\mathbb{Q}_\Sigma)$ induced from $\|\ \|$.

We also fix norms, which we continue to denote by $\|\ \|_v$, on $\mathrm{SL}_N(\mathbb{Q}_v)$ for all v; and put $\|\ \| = \max_\Sigma \|\,\|_v$. For every $g \in \mathrm{SL}_N(\mathbb{Q}_\Sigma)$, in particular for every $g \in G$, we set

$$|g| := \max\{\|g\|, \|g^{-1}\|\}. \tag{2.1}$$

Note that

$$|g| = |g^{-1}| \quad \text{and} \quad |g_1 g_2| \ll |g_1||g_2|. \tag{2.2}$$

Fix A_1 and A_2 both depending only on N so that

$$\| \wedge^r \mathrm{Ad}(g) z\| \leq A_2 |g|^{A_1} \|z\| \quad \text{for all } z \in \mathfrak{sl}_N(\mathbb{Q}_\Sigma) \text{ and } 1 \leq r \leq N^2. \tag{2.3}$$

2.3 Let $\mathbf{W} \subset \mathfrak{sl}_N$ be a rational subspace, then $\wedge^{\dim \mathbf{W}}\mathbf{W}$ defines a rational line in $\wedge^{\dim \mathbf{W}}\mathfrak{sl}_N$. This line is diagonally embedded in $\wedge^{\dim \mathbf{W}}\mathfrak{sl}_N(\mathbb{Q}_\Sigma)$, and we do not distinguish between this diagonal embedding and the line.

Fix a $\mathbb{Z}$-basis for $\mathbf{W}(\mathbb{R}) \cap \mathfrak{sl}_N(\mathbb{Z})$. Let $\mathbf{v}_\mathbf{W}$ denote the corresponding primitive integral vector on $\wedge^{\dim \mathbf{W}}\mathbf{W}$, and define

$$\mathrm{ht}(\mathbf{W}) = \|\mathbf{v}_\mathbf{W}\|.$$

Note that we used the max norm in the above definition, in particular, we have: $\mathrm{ht}(\mathbf{W})$ is an integer.

Alternatively, $\mathrm{ht}(\mathbf{W})$ may be defined as follows. Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_{\dim \mathbf{W}}\}$ be a $\mathbb{Q}$-basis for $\mathbf{W}$. Then

$$\mathrm{ht}(\mathbf{W}) = \prod_{v \in \mathcal{S}} \|\mathbf{e}_1 \wedge \cdots \wedge \mathbf{e}_{\dim \mathbf{W}}\|_v.$$

In view of the product formula, the above is independent of our choice of the rational basis for $\mathbf{W}$, see [4, §1.5].

Given a $\mathbb{Q}$-subgroup $\mathbf{H}$ of SL_N, we put $\mathbf{v}_\mathbf{H} := \mathbf{v}_{\mathrm{Lie}(\mathbf{H})}$ and define

$$\mathrm{ht}(\mathbf{H}) := \mathrm{ht}(\mathrm{Lie}(\mathbf{H})) = \|\mathbf{v}_\mathbf{H}\|. \tag{2.4}$$

2.4 For the rest of this paper, fix a finite subset $\Sigma \subset \mathcal{S}$ containing ∞.

The exponents in this paper are denoted by A with numerical indices. These constants depend only on N. The understanding is that $A_. > 1$.

Similarly, the constants C, D, and F in the sequel depend only on N, and are implicitly assumed to be > 1.

We use the notation $T \ll R$ to denote $T \le cR$ where the multiplicative constant c is allowed to depend on N, the number of places $\#\Sigma$, and polynomially on the finite primes in Σ and on $\mathrm{ht}(G)$. Similarly we define $T \gg R$.

It will also be convenient to use $\star$ to denote a constant. More precisely, we write $T \ll R^\star$ if $T \le cR^A$ or $T \le cR^\alpha$ where c is allowed to depend on N, $\#\Sigma$, polynomially on the finite primes in Σ and $\mathrm{ht}(\mathbf{G})$, and the exponent is either a "big enough" constant or a "small enough" constant depending only on N; hopefully the context will make it clear if the exponent needs to be large or small.

2.5 For all $v \in \Sigma$, let $\| \ \|_v$ denote the max norm with respect to the standard basis on $\mathbb{Q}_v^m$; we put $\|z\| = \|z\|_\Sigma = \max \|z_v\|_v$ for all $z = (z_v) \in \mathbb{Q}_\Sigma^m$.

Define

$$\mathsf{c}(z) = \prod_{v\in\Sigma} \|z\|_v \quad \text{for all } z \in \mathbb{Q}_\Sigma^m. \tag{2.5}$$

Note that $\mathsf{c}(rz) = \mathsf{c}(z)$ for all $r \in \mathbb{Z}_\Sigma^\times$ and all $z \in \mathbb{Q}_\Sigma^m$.

2.6 Lemma. *There exists A_3 and some $C_{m,\Sigma} \ge 1$ so that the following holds. Let $z \in \mathbb{Q}_\Sigma^m$ be a vector so that $\mathsf{c}(z) \ne 0$.*

(1) *There exists some $r_0 \in \mathbb{Z}_\Sigma^\times$ so that*

$$C_{m,\Sigma}^{-1} \|r_0 z\|_\Sigma \le \|r_0 z\|_v \le C_{m,\Sigma} \|r_0 z\|_\Sigma$$

for all $v \in \Sigma$, in particular, we have

$$\min_{r\in\mathbb{Z}_\Sigma^\times} \|rz\|_\Sigma \le C_{m,\Sigma} \|r_0 z\|_\Sigma \le C_{m,\Sigma} \mathsf{c}(r_0 z)^{1/\#\Sigma}. \tag{2.6}$$

(2) *Let $\|z\|_\Sigma = 1$, and let $T > 0$. Then*

$$\#\{r \in \mathbb{Z}_\Sigma^\times : \|rz\|_\Sigma \le T\} \le C_{m,\Sigma} \Big(\log \frac{T}{\mathsf{c}(z)}\Big)^{A_3}. \tag{2.7}$$

Proof. The claim in part (1) is proved in [30, Lemma 8.6].

We now turn to the proof of part (2). Let $\ell = \#\Sigma$ and for every $a > 0$ put

$$\mathcal{E}_a = \Big\{(w_1, \ldots, w_\ell) \in \mathbb{R}_+^\ell : \prod w_i = a\Big\}.$$

Note that $\mathcal{E}_a$ is invariant under multiplication by positive diagonal matrices in $\mathrm{SL}_\ell(\mathbb{R})$.

Let D_Σ denote the group of positive diagonal matrices in $\mathrm{SL}_\ell(\mathbb{Q})$ whose entries are in $\mathbb{Z}_\Sigma$. Let $a = \mathsf{c}(z)$. Then $(|z|_v) \in \mathcal{E}_a$, and for every $r \in \mathbb{Z}_\Sigma^\times$ we have $\mathrm{Diag}(|r|_v) \in D_\Sigma$.

Let $\| \ \|_{\mathrm{op}}$ denote the operator norm on $\mathrm{SL}_\ell(\mathbb{R})$ and let $\| \ \|_{\mathrm{m}}$ denote the max norm on $\mathbb{R}^\ell$. We have

$$\#\{A \in D_\Sigma : \|A\|_{\mathrm{op}} \leq S\} \ll_{m,\Sigma} (\log S)^\star. \tag{2.8}$$

Further, if $w = (w_1, \ldots, w_\ell) \in \mathcal{E}_a$ is so that $|w_i| \ll_{m,\Sigma} \|w\|_{\mathrm{m}} \ll_{m,\Sigma} |w_i|$ for all i, then $\|A\|_{\mathrm{op}}\|w\|_{\mathrm{m}} \ll_{m,\Sigma} \|wA\|_{\mathrm{m}} \ll_{m,\Sigma} \|A\|_{\mathrm{op}}\|w\|_{\mathrm{m}}$.

Hence, the claim follows from (2.8) if we replace z by $r_0 z$ so that

$$\|r_0 z\|_\Sigma \ll_{m,\Sigma} \|r_0 z\|_v \ll_{m,\Sigma} \|r_0 z\|_\Sigma$$

for all $v \in \Sigma$. □

Similar to (2.5), we define $\mathsf{c}(w) = \prod \|w\|_v$ for all $w \in \wedge\mathfrak{sl}_N(\mathbb{Q}_\Sigma)$.

2.7 Let $\mathbf{G}$ be a connected $\mathbb{Q}$-group of class $\mathcal{H}$. Fix an embedding $\iota : \mathbf{G} \to \mathrm{SL}_N$ defined over $\mathbb{Q}$. Put $G = G_\Sigma$ and $\mathfrak{g} = \mathfrak{g}_\Sigma$.

We identify $\mathbf{G}$ with $\iota(\mathbf{G}) \subset \mathrm{SL}_N$, hence, $G \subset \mathrm{SL}_N(\mathbb{Q}_\Sigma)$. Let $\mathfrak{g}(\mathbb{Z}_\Sigma) := \mathfrak{g} \cap \mathfrak{sl}_N(\mathbb{Z}_\Sigma)$, then $[\mathfrak{g}(\mathbb{Z}_\Sigma), \mathfrak{g}(\mathbb{Z}_\Sigma)] \subset \mathfrak{g}(\mathbb{Z}_\Sigma)$.

We fix a $\mathbb{Z}$-basis, $\mathcal{B}_{\mathbf{G}} = \{z_1, \ldots, z_d\}$, for $\mathfrak{g} \cap \mathrm{SL}_N(\mathbb{Z})$ so that $\mathsf{c}(z_i) \ll \mathrm{ht}(\mathbf{G})^\star$ for all $1 \leq i \leq d$. Using this basis, we identify $\mathrm{Lie}(\mathbf{G})$ with a d-dimensional vector space with a $\mathbb{Q}$-structure. We also identify the $\mathbb{Z}$-span of $\mathcal{B}_{\mathbf{G}}$ with $\mathbb{Z}^d$; hence we get a representation $\mathrm{Ad} : \mathbf{G} \to \mathrm{SL}_d$.

Let $\Gamma \subset G \cap \mathrm{SL}_N(\mathbb{Z}_\Sigma)$ be a lattice. Then Γ fixes $\mathfrak{g}(\mathbb{Z}_\Sigma)$, which implies that $\mathrm{Ad}(\Gamma) \subset \mathrm{SL}_d(\mathbb{Z}_\Sigma)$—recall that we are using $\mathcal{B}_G$ to define Ad over the ring $\mathbb{Z}_\Sigma$ (and not just as a $\mathbb{Q}$-representation).

Let $X := G/\Gamma$. For any $\eta > 0$, set

$$X_\eta := \{g\Gamma \in X : \min_{0 \neq v \in \mathfrak{g}(\mathbb{Z}_\Sigma)} \mathsf{c}(gv) \geq \eta\} \tag{2.9}$$

where here and in what follows we often simply write gv for $\mathrm{Ad}(g)v$; similarly, for $w \in \wedge\mathfrak{g}$, we simply write gw to denote the corresponding wedge power of the adjoint representation.

For any $\eta > 0$, the set X_η is a compact subset of G/Γ, and $G/\Gamma = \bigcup_{\eta > 0} X_\eta$.

We will need a quantitative version of the former statement:

2.8 Lemma. *There exist some $E_{\mathbf{G}}$ (depending on the geometry of G/Γ) and F (depending only on N) so that the following holds. Let $g \in G$ be so that $g\Gamma \in X_\eta$. There exists some $\gamma \in \Gamma$ so that*

$$|g\gamma| \leq E_{\mathbf{G}} \eta^{-F}.$$

Remark. For a point $g\Gamma$ to be X_η essentially means that the local injectivity radius for G/Γ at $g\Gamma$ is $\gg \eta^\star$. Thus Lemma 2.8 can be viewed as an estimate of the diameter of the part of G/Γ which has injectivity radius greater than $\eta^\star$. In particular, for G/Γ compact, F is essentially meaningless, and $E_{\mathbf{G}}$ is the diameter of the smallest norm ball in G needed to cover G/Γ; see [37, Thm. 1.7 and Thm. 6.9] for more explicit estimates.

For future convenience, we set $\tilde{E}_{\mathbf{G}} = \mathrm{ht}(\mathbf{G}) \cdot E_{\mathbf{G}}$. In the proof of this lemma implicit constants are allowed to depend on G and Γ where indicated.

Proof. By [37, Prop. 3.1] there exists a Levi subgroup $\mathbf{L}$ so that $\mathrm{ht}(\mathbf{L})$ is bounded by $\mathrm{ht}(\mathbf{G})^\star$, in particular, $L \cap \Gamma$ is a lattice in L.[1]

For any $g \in G$, define

$$\alpha_{\mathfrak{g}}(g) := \max\{\mathsf{c}(\mathrm{Ad}(g)z)^{-1} : 0 \neq z \in \mathfrak{g}(\mathbb{Z}_\Sigma)\};$$

define similarly $\alpha_{\mathfrak{l}}$ for any $g \in L = \mathbf{L}(\mathbb{Q}_\Sigma)$.

Let $g \in G$ and write $g = g^0 g^u$ where $g^0 \in L$ and $g^u \in \mathrm{R}_u(G)$. Then

$$\alpha_{\mathfrak{l}}(g^0) \leq \alpha_{\mathfrak{g}}(g^0)$$

and by [37, Lemmas 4.4 and 6.8] we have $\alpha_{\mathfrak{g}}(g^0) \ll_{\mathbf{G},\Gamma} \alpha_{\mathfrak{g}}(g)^\star$. It follows from reduction theory for L ([40, Thm. 4.8] and [40, Thm. 4.17]) that there exists some $\gamma_0 \in L \cap \Gamma$ so that

$$|g^0\gamma_0| \ll_{\mathbf{G},\Gamma} \alpha_{\mathfrak{l}}(g^0)^\star,$$

and combining the above we get $|g^0\gamma_0| \ll \alpha_{\mathfrak{g}}(g)^\star$. Moreover, $\gamma_0^{-1} g^u \gamma_0 \in \mathrm{R}_u(G)$ and $\mathrm{ht}(\mathrm{R}_u(\mathbf{G})) \ll 1$, see Lemma 4.2. Therefore, there exists some $\gamma_1 \in \mathrm{R}_u(G) \cap \Gamma$ so that

$$|\gamma_0^{-1} g^u \gamma_0 \gamma_1| \ll_{\mathbf{G},\Gamma} 1,$$

see, e.g., [37, Lemma 5.6].

Put $\gamma = \gamma_0\gamma_1$. Then

$$\begin{aligned}|g\gamma| &= |g^0 g^u \gamma_0\gamma_1|\\ &\ll_{\mathbf{G},\Gamma} |g^0\gamma_0||\gamma_0^{-1} g^u \gamma_0\gamma_1| \ll_{\mathbf{G},\Gamma} \alpha_{\mathfrak{g}}(g)^\star,\end{aligned}$$

as was claimed. □

[1] Note that without the estimate on the height, the existence of a Levi subgroup so that $L \cap \Gamma$ is a lattice in L is a theorem of Mostow [38].

2.9 Let $U = \prod_{v\in\Sigma} U_v \subset G$ where for all $v \in \Sigma$ we have U_v is a (possibly trivial) unipotent subgroup of G. We will refer to such groups as a $\mathbb{Q}_\Sigma$-unipotent subgroup of G. Define

$$\Sigma' = \{v \in \Sigma : U_v \neq \{e\}\}. \tag{2.10}$$

Let $\mathfrak{u}$ (resp. $\mathfrak{u}_v$) denote the Lie algebra of U (resp. U_v). The exponential map defines an isomorphism of $\mathbb{Q}_\Sigma$-varieties from $\mathfrak{u}$ onto U. We fix once and for all a basis $\mathcal{B}_U$ for $\mathfrak{u}$ consisting of elements which are nontrivial at only one place.

For $\delta > 0$, let

$$\mathsf{B}_\mathfrak{u}(0,\delta) = \left\{ \sum_{z\in\mathcal{B}_U} r_z z : |r_z|_\Sigma \le \delta \right\}$$

where $r_z \in \mathbb{Q}_\Sigma$ and $|r|_\Sigma := \max_\Sigma |r_v|_v$; put

$$\mathsf{B}_U(e) := \exp(\mathsf{B}_\mathfrak{u}(0,1)).$$

Thus $\mathsf{B}_U(e)$ is a product of neighborhoods of 1 in U_v for $v \in \Sigma$. A subset $\mathsf{B} \subset \mathsf{B}_U(e)$ will be called a **ball** if it is the image, under the exponential map, of a norm ball in $\mathfrak{u}$.

Let $\lambda : \mathfrak{u} \to \mathfrak{u}$ be a $\mathbb{Q}_\Sigma$-diagonalizable expanding linear map, and for all $k \in \mathbb{N}$ let $\lambda_k : \mathfrak{u} \to \mathfrak{u}$ denote the k-fold composition of λ with itself, i.e., $\lambda_k = \lambda \circ \cdots \circ \lambda$, k-times. We will throughout make the assumption that for some fixed $\kappa > 0$ and all $k \ge 1$

$$\exp(\lambda_{k-\kappa}(\mathsf{B}_\mathfrak{u}(0,1))) \cdot \exp(\lambda_{k-1}(\mathsf{B}_\mathfrak{u}(0,1))) \subset \exp(\lambda_k(\mathsf{B}_\mathfrak{u}(0,1))). \tag{2.11}$$

We now explicate two examples of λ which satisfy the required conditions. One may take λ to be an expanding automorphism of the Lie algebra $\mathfrak{u}$ as Margulis and Tomanov did in [35]; more explicitly, we may embed $\mathbf{G}$ in a larger group in which one can find an element h so that $\lambda = \mathrm{Ad}(h)$ expands $\mathfrak{u}$.

The following is an alternative construction for a λ which satisfies the required assumptions: Let $v \in \Sigma$, and consider the lower central series for $\mathfrak{u}_v$. That is:

$$\mathfrak{u}_v = \mathfrak{u}_{v,0} \supset \mathfrak{u}_{v,1} \supset \cdots \supset \mathfrak{u}_{v,n_v} = \{0\}$$

where $\mathfrak{u}_{v,i+1} = [\mathfrak{u}_v, \mathfrak{u}_{v,i}]$ for all $0 \le i < n_v$.

For each i, let $\mathfrak{u}_v^i$ denote an orthogonal complement of $\mathfrak{u}_{v,i+1}$ in $\mathfrak{u}_{v,i}$.[2] In particular, we have

$$\mathfrak{u}_{v,i} = \mathfrak{u}_{v,i+1} \oplus \mathfrak{u}_v^i .$$

[2] Recall that for a finite prime p, a set of unit vectors in $\mathbb{Q}_p^m$ is called orthonormal if it can be extended to a $\mathbb{Z}_p$-basis for $\mathbb{Z}_p^m$.

Fixing an orthonormal basis of $\mathfrak{u}_v^i$ for all $0 \le i < n_v$, we obtain an orthonormal basis of $\mathfrak{u}_v$. Let $a_p = p^{-3}$ if p is a finite prime and $a_\infty = e^3$. For each $k \in \mathbb{N}$, define $\lambda_k : \mathfrak{u}_v \to \mathfrak{u}_v$ by

$$\lambda_k(z) = a_v^{(i+1)k} z \quad \text{for all } z \in \mathfrak{u}_v^i .$$

We leave the verification that this example does indeed satisfy (2.11) to the reader.

Abusing the notation, for an element $u = \exp(z) \in U$ we set $\lambda(u) := \exp(\lambda(z))$; that is: λ and λ_k are also considered as functions on U.

In this paper, we assume that a linear expanding map λ satisfying (2.11) is fixed; moreover, we assume that the parameters κ, $\frac{|\lambda_1(\mathsf{B}_U(e))|}{|\mathsf{B}_U(e))|}$, etc. depend only on N and polynomially on $\mathrm{ht}(\mathbf{G})$. For instance, the examples above satisfy these properties, and the reader may take λ to be one of these examples.

However, if for some reason, the reader is keen on taking some particularly wild expanding linear map λ satisfying (2.11), the only adverse effect would be that the implicit multiplicative constants need to be allowed to depend polynomially on the parameters of λ.

3 Statements of the main theorems: Σ-arithmetic

Let $\mathbf{G} \subset \mathrm{SL}_N$ be a $\mathbb{Q}$-group. Recall the family

$$\mathcal{H} = \{\mathbf{H} \subset \mathbf{G} : \mathbf{H} \text{ is a connected } \mathbb{Q}\text{-subgroup and } \mathrm{R}(\mathbf{H}) = \mathrm{R}_u(\mathbf{H})\},$$

where $\mathrm{R}(\mathbf{H})$ (resp. $\mathrm{R}_u(\mathbf{H})$) denotes the solvable (resp. unipotent) radical of $\mathbf{H}$. We always assume that $\mathbf{G} \in \mathcal{H}$. Recall also our notation $G_v = \mathbf{G}(\mathbb{Q}_v)$ for all $v \in \Sigma$, and $G = \prod_{v\in\Sigma} G_v$.

Let $\mathbf{H} \in \mathcal{H}$ be a proper subgroup and put

$$\rho_H := \oplus_\Sigma(\wedge^{\dim \mathbf{H}} \mathrm{Ad}) \quad \text{and} \quad V_H := \wedge^{\dim \mathbf{H}} \mathfrak{g} = \oplus_\Sigma(\wedge^{\dim \mathbf{H}} \mathfrak{g}_v).$$

We shall identify between the representation ρ_H of G and the $\mathbb{Q}$-representaiton $\wedge^{\dim \mathbf{H}} \mathrm{Ad}$ of $\mathbf{G}$.

Let $\mathbf{v_H}$ be a primitive integral vector in $\wedge^{\dim \mathbf{H}}\mathrm{Lie}(\mathbf{G})$ corresponding to the Lie algebra of $\mathbf{H}$. Recall from (2.4) that

$$\mathrm{ht}(\mathbf{H}) = \|\mathbf{v_H}\|_\Sigma = \|\mathbf{v_H}\|. \tag{3.1}$$

The vector $\mathbf{v_H}$ is diagonally embedded in V_H (which is a product of local factors); in order to put an emphasis on the local nature of this diagonally embedded vector, we will denote it by v_H. Define

$$\eta_H(g) := \rho_H(g) v_H \quad \text{for every } g \in G.$$

Throughout, $U = \prod_{v\in\Sigma} U_v \subset G$ is a $\mathbb{Q}_\Sigma$-unipotent subgroup. We will use the notation from §2.9. In particular, $\mathcal{B}_U$ is an orthonormal basis for $\mathfrak{u}$, and

$$\mathsf{B}_U(e) = \exp\Big(\Big\{\sum_{z\in\mathcal{B}_U} r_z z : |r_z|_\Sigma \leq 1\Big\}\Big).$$

Recall also from §2.9 the notion of an admissible expanding map $\lambda_k : U \to U$ for all $k \in \mathbb{N}$.

The following is a Σ-arithmetic version of Definition 1.3, and plays a crucial role in this paper.

3.1 Definition. Let $\varepsilon : \mathbb{R}^+ \to (0, 1)$ be a monotone decreasing function, $t \in \mathbb{R}^+$, and $\mathcal{F} \subset \mathcal{H}$ a subcollection that is Γ-invariant with respect to conjugation. A point $g\Gamma$ is called $(\varepsilon, t, \mathcal{F})$**-Diophantine** for the action of U if for all $\mathbf{H} \in \mathcal{F}$ with $\{e\} \neq \mathbf{H} \neq \mathbf{G}$ and $\mathsf{c}(\eta_H(g)) < e^t$

$$\max_{z\in\mathcal{B}_U} \|z \wedge \eta_H(g)\| \geq \varepsilon(\mathsf{c}(\eta_H(g))). \tag{3.2}$$

A point is (ε, t)**-Diophantine** if it is $(\varepsilon, t, \mathcal{H})$-Diophantine. A point is ε**-Diophantine** if it is (ε, t)-Diophantine for all $t > 0$.

Note that if there exists some nontrivial $\mathbf{H} \lhd \mathbf{G}$ so that $U \subset \mathbf{H}(\mathbb{Q}_\Sigma)$, then for any $\varepsilon : \mathbb{R}^+ \to (0, 1)$ the set of ε-Diophantine points is empty.

We now state the main result of this paper.

3.2 Theorem. *There exist constants A and D depending only on N, and constants E depending on N, $\#\Sigma$ and polynomially on* $\mathrm{ht}(\mathbf{G})$ *and the primes in Σ, and E_1 depending in addition also (polynomially) on $E_{\mathbf{G}}$, so that the following holds. Let $g \in G$, $t > 0$, $k \geq 1$, and $0 < \eta < 1/2$. Assume $\varepsilon : \mathbb{R}^+ \to (0, 1)$ satisfies for any $s \in \mathbb{R}^+$ that*

$$\varepsilon(s) \leq \eta^A s^{-A}/E_1. \tag{3.3}$$

Then at least one of the following three possibilities holds.

(1) $|\{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \notin X_\eta$ *or* $\lambda_k(u)g\Gamma$ *is not* (ε, t)*-Diophantine*$\}| < E_1\eta^{1/D}$.

(2) *There exists a nontrivial proper subgroup* $\mathbf{H} \in \mathcal{H}$ *with*

$$\mathrm{ht}(\mathbf{H}) \leq (E|g|^A + E_1 e^{At})\eta^{-A}$$

so that the following hold for all $u \in \mathsf{B}_U(e)$:

$$\mathsf{c}(\eta_H(\lambda_k(u)g)) \leq (E|g|^A + E_1 e^{At})\eta^{-A},$$
$$\max_{z\in\mathcal{B}_U} \|z \wedge \eta_H(\lambda_k(u)g)\| \leq e^{-k/D}(E|g|^A + E_1 e^{At})\eta^{-A}.$$

(3) *There exists a nontrivial proper normal subgroup* $\mathbf{H} \lhd \mathbf{G}$ *with*

$$\mathrm{ht}(\mathbf{H}) \leq E_1(e^t \eta^{-1})^A$$

so that

$$\max_{z \in \mathcal{B}_U} \|z \wedge v_H\| \leq \varepsilon(\mathrm{ht}(\mathbf{H})^{1/A}\eta/E_1)^{1/A}.$$

Of course if $\mathbf{G}$ is $\mathbb{Q}$-simple, possibility (3) cannot hold. A typical example where there are infinitely many normal subgroups is

$$\mathbf{G} = \mathrm{SL}_n \ltimes ((\mathrm{G_a})^n \times (\mathrm{G_a})^n)$$

with $\mathrm{G_a}$ denoting the one-dimensional additive group (the simplest possible algebraic group!). The group $\mathbf{G}$ is a perfect group, and for any $l, k \in \mathbb{Z}$ the subgroup

$$\mathbf{H}_{l,k} = \{(g, lv, kv) : g \in \mathrm{SL}_n, v \in \mathrm{G}_\mathrm{a}^n\}$$

is a normal subgroup of $\mathbf{G}$.

We note the following interesting corollary of Theorem 3.2. For simplicity we state it in the case where $\mathbf{G}$ has only finitely many normal $\mathbb{Q}$-subgroups (it is fairly easy to adjust the statement and the proof to accommodate general $\mathbf{G}$, but they become a bit messier). Results of a similar flavor were given by Lindenstrauss and Margulis in [31, Prop. 4.4].

3.3 Corollary. *Let* $\mathbf{G}, G, \Gamma, U$ *be as above, with* $\mathbf{G}$ *having only finitely many normal* $\mathbb{Q}$*-subgroups, and* $U \not\subset \mathbf{H}(\mathbb{Q}_\Sigma)$ *for all* $\mathbf{H} \lhd \mathbf{G}$. *There are* A_4, A_5 *depending only on* N *and* ϵ_1, ϵ_2 *depending on* $N, \#\Sigma$ *and polynomially on* $\mathrm{ht}(\mathbf{G})$*, the primes in* Σ*, and* $E_{\mathbf{G}}$*, and* t_0 *that depend in addition also on* U *and how far it is from lying in any* $\mathbf{H} \lhd \mathbf{G}$*, so that if* $\varepsilon(s) = \epsilon_1 \eta^{A_4} s^{-A_4}$ *then if* $t > t_0$*, and if* $t', k \geq A_5(t + \log(1/\eta) + \log(1/\epsilon_2))$*, then for any* (ε, t')*-Diophantine* $g\Gamma \in X_\eta$*,*

$$\left| \left\{ u \in \mathsf{B}_U(e) : \begin{array}{c} \lambda_k(u)g\Gamma \notin X_\eta \text{ or} \\ \lambda_k(u)g\Gamma \text{ is not } (\varepsilon, t)\text{-Diophantine} \end{array} \right\} \right| < E_1 \eta^{1/D}.$$

Proof of Corollary 3.3 assuming Theorem 3.2.

Assuming that the constants in Corollary 3.3 were appropriately chosen, ε satisfies (3.3) and we may apply Theorem 3.2.

If (1) of that theorem holds there is nothing to prove. Otherwise either (2) or (3) of that theorem holds. (3) is ruled out by our assumption that $U \not\subset \mathbf{H}(\mathbb{Q}_\Sigma)$ for all $\mathbf{H} \lhd \mathbf{G}$ if t_0 is large enough.

Suppose then we are in case (2). As $g\Gamma \in X_\eta$ it follows that there is a nontrivial subgroup $\mathbf{H} \in \mathcal{H}$ for which in particular[3]

$$\begin{aligned} \mathsf{c}(\eta_H(g)) &\le 2E_1 e^{At}\eta^{-A}, \\ \max_{z\in\mathcal{B}_U} \|z \wedge \eta_H((u)g)\| &\le 2E_1 e^{-k/D} e^{At}\eta^{-A}. \end{aligned} \tag{3.4}$$

Choose A_5, ϵ_2 so that in particular $t' > \log(2E_1 e^{At}\eta^{-A})$. Then since $g\Gamma$ is (ε, t')-Diophantine

$$\max_{z\in\mathcal{B}_U} \|z \wedge \eta_H((u)g)\| \ge \epsilon_1 \eta^{A_4}(2E_1 e^{At}\eta^{-A})^{-A_4} \gg \eta^{2AA_4} e^{-AA_4 t}.$$

But this contradicts (3.4) if $k \ge A_5(t + \log(1/\eta) + \log(1/\epsilon_2))$ for sufficiently large A_5. □

4 The family $\mathcal{H}$ and the Diophantine condition

4.1 Recall the family

$$\mathcal{H} = \{\mathbf{H} \subset \mathbf{G} : \mathbf{H} \text{ is a connected } \mathbb{Q}\text{-subgroup and } \mathrm{R}(\mathbf{H}) = \mathrm{R}_u(\mathbf{H})\}$$

where $\mathrm{R}(\mathbf{H})$ (resp. $\mathrm{R}_u(\mathbf{H})$) denotes the solvable (resp. unipotent) radical of $\mathbf{H}$.

For any subgroup $\mathbf{H} \in \mathcal{H}$, we put $H = \mathbf{H}(\mathbb{Q}_\Sigma)$. Sometimes we write $H \in \mathcal{H}$.

4.2 Lemma. *There exists some A_6 so that the following holds. Let $\mathbf{L} \subset \mathrm{SL}_N$ be a connected algebraic group defined over $\mathbb{Q}$. Then*

$$\mathrm{ht}([\mathbf{L},\mathbf{L}]) \ll \mathrm{ht}(\mathbf{L})^{A_6}; \quad \mathrm{ht}(\mathrm{R}(\mathbf{L})) \ll \mathrm{ht}(\mathbf{L})^{A_6}; \quad \textit{and} \quad \mathrm{ht}(\mathrm{R}_u(\mathbf{L})) \ll \mathrm{ht}(\mathbf{L})^{A_6}.$$

Proof. Let $\mathcal{B}$ be a $\mathbb{Z}$-basis for $\mathrm{Lie}(\mathbf{L}) \cap \mathrm{SL}_N(\mathbb{Z})$ so that $\|z\| \ll \mathrm{ht}(\mathbf{L})^\star$ for all $z \in \mathcal{B}$. Then $\{[z, z'] : z, z' \in \mathcal{B}\}$ generates $[\mathrm{Lie}(\mathbf{L}), \mathrm{Lie}(\mathbf{L})]$. Hence,

$$\mathrm{ht}([\mathrm{Lie}(\mathbf{L}), \mathrm{Lie}(\mathbf{L})]) \ll \mathrm{ht}(\mathbf{L})^\star.$$

It remains to bound $\mathrm{ht}(\mathrm{R}_u(\mathbf{L}))$. To that end, first note that

$$\mathrm{R}(\mathrm{Lie}(\mathbf{L})) = \{z \in \mathrm{Lie}(\mathbf{L}) : \mathsf{k}_{\mathbf{L}}(z, [w, w']) = 0, \forall w, w' \in \mathrm{Lie}(\mathbf{L})\}$$

where $\mathsf{k}_{\mathbf{L}}$ is the killing form of $\mathrm{Lie}(\mathbf{L})$. Therefore, $\mathrm{ht}(\mathrm{R}(\mathrm{Lie}(\mathbf{L}))) \ll \mathrm{ht}(\mathbf{L})^\star$.

Now let $\mathcal{B}'$ be a $\mathbb{Z}$-basis for $\mathrm{R}(\mathrm{Lie}(\mathbf{L})) \cap \mathrm{SL}_N(\mathbb{Z})$ so that $\|z\| \ll \mathrm{ht}(\mathbf{L})^\star$ for all $z \in \mathcal{B}'$. Then

$$\mathrm{R}_u(\mathrm{Lie}(\mathbf{L})) = \{z \in \mathrm{R}(\mathrm{Lie}(\mathbf{L})) : \mathrm{tr}(w_1 \cdots w_s z) = 0, \forall 1 \le s \le N, w_i \in \mathcal{B}'\}.$$

Hence, $\mathrm{ht}(\mathrm{R}_u(\mathrm{Lie}(\mathbf{L}))) \ll \mathrm{ht}(\mathbf{L})^\star$. □

---

[3] Possibly for a slightly larger A than in the theorem.

4.3 Algebraic properties of subgroups in class $\mathcal{H}$. A quantitative notion of a point satisfying a Diophantine condition was given in Definition 3.1. This definition is formulated in terms of certain representations whose constructions and basic properties we now recall.

Let $\mathbf{H}\in\mathcal{H}$ be a proper subgroup. Recall that $\mathfrak{g}=\oplus_\Sigma\,\mathfrak{g}_v$ where $\mathfrak{g}_v=\mathrm{Lie}(G_v)$. Put

$$\rho_H := \wedge^{\dim \mathbf{H}}\mathrm{Ad} \quad\text{and}\quad V_H := \oplus_\Sigma \wedge^{\dim\mathbf{H}}\mathfrak{g}_v.$$

The representation ρ_H is defined over $\mathbb{Q}$.

Let $\mathbf{v_H}$ be a primitive integral vector in $\wedge^{\dim\mathbf{H}}\mathrm{Lie}(\mathbf{G})$ (or $\wedge^{\dim\mathbf{H}}\mathfrak{sl}_N$) corresponding to the Lie algebra of $\mathbf{H}$, see §2.3. We embed $\mathbf{v_H}$ diagonally in V_H and denote this vector by v_H. Let $\eta_H : G\to V_H$ denote the orbit map, that is

$$\eta_H(g)=\rho_H(g)v_H \quad\text{for all } g\in G.$$

Note that ρ_H and V_H depend only on $\dim\mathbf{H}$, however, $\mathbf{v_H}$ (similarly v_H) uniquely determines $\mathrm{Lie}(\mathbf{H})$ and hence $\mathbf{H}$.

4.4 Lemma. (1) *$N_G(H)=\{g\in G:\rho_H(g)v_H=(\chi_H(g_p))_{v\in\Sigma}v_H\}$, where χ_H is a rational character.*

(2) *The orbit $\eta_H(\Gamma)$ is discrete and closed in V_H.*

Proof. Property (1) is a consequences of the definition.

In light of our assumption that Γ is arithmetic, property (2) also follows from the definitions. We note, however, that this qualitative result does not require arithmeticity of Γ, see [17, Thm. 3.4]. □

4.5 Lemma. *There exists some constant A_7 so that the following holds.*

$$\#\{\mathbf{H}\in\mathcal{H}:\mathrm{ht}(\mathbf{H})\le T\}\ll T^{A_7}. \tag{4.1}$$

Proof. This follows from the definitions of $\mathbf{v_H}$ and $\mathrm{ht}(\mathbf{H})$. □

4.6 Lemma. *There exists some $A_8>0$ so that the following holds. Given any $\mathbb{Q}$-group $\mathbf{L}\subset\mathbf{G}$, there exists a normal subgroup $\mathbf{L}^{\mathcal{H}}\subset\mathbf{L}$ which is maximal among all subgroups of $\mathbf{L}$ which belong to class $\mathcal{H}$; moreover,*

$$\mathrm{ht}(\mathbf{L}^{\mathcal{H}})\ll\mathrm{ht}(\mathbf{L})^{A_8}. \tag{4.2}$$

Proof. Since $\mathbf{L}/\,\mathrm{R}_u(\mathbf{L})$ is a reductive group and unipotent subgroups in $\mathbf{L}$ map to unipotent subgroups in $\mathbf{L}/\,\mathrm{R}_u(\mathbf{L})$, we have

$$\mathrm{Lie}(\mathbf{L}^{\mathcal{H}})=[\mathrm{Lie}(\mathbf{L}),\mathrm{Lie}(\mathbf{L})]+\mathrm{Lie}(\mathrm{R}_u(\mathbf{L})); \tag{4.3}$$

in particular, $\mathbf{L}^{\mathcal{H}}$ exists.

By Lemma 4.2 we have $\mathrm{ht}([\mathbf{L},\mathbf{L}])\ll\mathrm{ht}(\mathbf{L})^\star$ and $\mathrm{ht}(\mathrm{R}_u(\mathbf{L}))\ll\mathrm{ht}(\mathbf{L})^\star$. The claim thus follows from (4.3). □

4.7 Let
$$\mathbf{L}_{\mathbf{H}} = \{g \in \mathbf{G} : \wedge^{\dim \mathbf{H}} \mathrm{Ad}(g)\mathbf{v}_{\mathbf{H}} = \mathbf{v}_{\mathbf{H}}\}.$$
Then $\mathbf{L}_{\mathbf{H}}$ is a $\mathbb{Q}$-group. The subgroup $\mathbf{L}_{\mathbf{H}}$ is not necessarily in $\mathcal{H}$. Define
$$\mathbf{M}_{\mathbf{H}} := \mathbf{L}_{\mathbf{H}}^{\mathcal{H}};$$
see (4.2) for the notation.

Put $L_H = \mathbf{L}_{\mathbf{H}}(\mathbb{Q}_\Sigma)$ and $M_H = \mathbf{M}_{\mathbf{H}}(\mathbb{Q}_\Sigma)$. Note that
$$L_H = \{g \in G : \rho_H(g)v_H = v_H\}.$$

We will simply denote these groups by $\mathbf{L}$, L, $\mathbf{M}$, and M when there is no confusion.

4.8 Lemma. *There exist A_9 with the following property. For any $\mathbf{H} \in \mathcal{H}$ we have*
$$\mathrm{ht}(\mathbf{L}_{\mathbf{H}}) \ll \mathrm{ht}(\mathbf{H})^{A_9} \quad \textit{and} \quad \mathrm{ht}(\mathbf{M}_{\mathbf{H}}) \ll \mathrm{ht}(\mathbf{H})^{A_9}. \tag{4.4}$$

Proof. Since $\mathbf{M}_{\mathbf{H}} := \mathbf{L}_{\mathbf{H}}^{\mathcal{H}}$, the second inequality is a consequence of the first inequality and (4.2).

Recall now that
$$\mathrm{Lie}(\mathbf{L}_{\mathbf{H}}) = \{z \in \mathrm{Lie}(\mathbf{G}) : \wedge^{\dim \mathbf{H}} \mathrm{ad}(z)\mathbf{v}_{\mathbf{H}} = 0\},$$
and that $\mathbf{v}_{\mathbf{H}}$ is an integral vector with $\|\mathbf{v}_{\mathbf{H}}\| = \mathrm{ht}(\mathbf{H})$.

The first inequality thus follows, and the proof is complete. □

4.9 Lemma. (1) *For any $\gamma \in \Gamma$ and any $\mathbf{H} \in \mathcal{H}$, we have*
$$1 \le \mathrm{ht}(\gamma \mathbf{H} \gamma^{-1}) = \mathsf{c}(\eta_H(\gamma)).$$

(2) *Let $r > 1$ and suppose $\gamma \in \Gamma$ is so that $\mathsf{c}(\eta_H(\gamma)) \le r$. Then*
  (a) $\mathrm{ht}(\gamma \mathbf{L}_{\mathbf{H}} \gamma^{-1}) \ll r^\star$;
  (b) $\mathrm{ht}(\gamma \mathbf{M}_{\mathbf{H}} \gamma^{-1}) \ll r^\star$.

Proof. Recall that $\mathrm{Ad}(\Gamma) \subset \mathrm{SL}_d(\mathbb{Z}_\Sigma)$. Recall that $\mathbf{v}_{\mathbf{H}}$ is primitive, in particular, $\|\eta_H(\gamma)\|_p = 1$ for all $p \notin \Sigma$. Part (1) of the lemma thus follows from the definition of $\mathrm{ht}(\gamma \mathbf{H} \gamma^{-1})$.

To see parts (2)a and (2)b, note that
$$\gamma \mathbf{L}_{\mathbf{H}} \gamma^{-1} = \mathbf{L}_{\gamma \mathbf{H} \gamma^{-1}} \quad \text{and} \quad \gamma \mathbf{M}_{\mathbf{H}} \gamma^{-1} = \mathbf{M}_{\gamma \mathbf{H} \gamma^{-1}}.$$
Hence, the claim follows from part (1) and (4.4). □

Let $\mathbf{H} \in \mathcal{H}$. For any $g \in G$ and any $r > 1$, put

$$m_H(g, r) := \lceil \log(R_H(g, r)) \rceil, \tag{4.5}$$

where $R_H(g, r) := \max\{\mathsf{c}(\eta_{M_H}(g\gamma)) : \gamma \in \Gamma, \mathsf{c}(\eta_H(\gamma)) \le r\}$.

4.9.1 Corollary. (1) $R_H(g, r) \ll |g|^\star r^\star$.
(2) $\#(\eta_H(\Gamma) \cap B_{V_H}(v_H, r)) \ll r^\star$.

Proof. We first prove part (1). For any $\gamma \in \Gamma$ so that $\mathsf{c}(\eta_H(\gamma)) \le r$, we have $\mathrm{ht}(\gamma \mathbf{M_H} \gamma^{-1}) \ll r^\star$, see Lemma 4.9(2)(b). Moreover, by Lemma 4.9(1), we have

$$1 \le \mathrm{ht}(\gamma \mathbf{M_H} \gamma^{-1}) = \mathsf{c}(\eta_{M_H}(\gamma)).$$

Using (2.3) to control the effect of g, the above implies the claim in part (1).

The second claim follows from the fact that $\mathrm{Ad}(\Gamma) \subset \mathrm{SL}_d(\mathbb{Z}_\Sigma)$. □

4.10 Lemma. *Let $\mathbf{H} \in \mathcal{H}$. Assume there exist an $L > 0$, a sequence $\ell_n \to 0$, and a sequence $g_n\Gamma \to g\Gamma$ satisfying the following:*
(1) *$\mathsf{c}(\eta_H(g_n)) \le L$ for all n, and*
(2) *$\max_{z \in \mathcal{B}_U} \|z \wedge \eta_H(g_n)\| \le \ell_n$ for all n.*
Then $g \in \{g' \in N_G(U, H) : \mathsf{c}(\eta_H(g')) \le L\}\Gamma$.

Proof. In view of our assumption, there exists a sequence $\{\gamma_n\}$ so that

$$g_n \gamma_n^{-1} \to g.$$

Hence, using the assumption in (1), we get that

$$\mathsf{c}(\eta_H(g_n \gamma_n^{-1} \gamma_n)) = \mathsf{c}(\eta_H(g_n)) \le L. \tag{4.6}$$

Moreover, since $g_n \gamma_n^{-1} \to g$, we have $|g_n \gamma_n^{-1}| \le 1 + |g|$ for all large enough n. This and the above imply that for some constant A' depending only on N, we have

$$\mathsf{c}(\eta_H(\gamma_n)) = \mathrm{ht}(\gamma_n \mathbf{H} \gamma_n^{-1}) \le L(2 + |g|)^{A'}$$

for all large enough n.

Using (4.1) and passing to a subsequence, we assume that $\gamma_n \mathbf{H} \gamma_n^{-1} = \gamma \mathbf{H} \gamma^{-1}$ for all n, or equivalently that $\eta_H(\gamma_n) = \eta_H(\gamma)$. Then for any $z \in \mathcal{B}_U$

$$\begin{aligned} z \wedge \eta_{\gamma H \gamma^{-1}}(g_n \gamma_n^{-1}) &= z \wedge \eta_H(g_n \gamma_n^{-1} \gamma_n) \\ &= z \wedge \eta_H(g_n). \end{aligned}$$

This computation and the assumption in (2) now imply that

$$\|z \wedge \eta_{\gamma H\gamma^{-1}}(g\gamma_n^{-1})\| \le \ell_n \quad \text{for all } z \in \mathcal{B}_U.$$

Passing to the limit, we get that $z \wedge \eta_{\gamma H\gamma^{-1}}(g) = 0$ for all $z \in \mathcal{B}_U$. That is

$$z \wedge \eta_H(g\gamma) = 0 \quad \text{for all } z \in \mathcal{B}_U. \tag{4.7}$$

Similarly, using the fact that $\eta_H(\gamma_n) = \eta_H(\gamma)$ for all n and passing to the limit in (4.6) we get that

$$\mathsf{c}(\eta_H(g\gamma)) \le L. \tag{4.8}$$

In view of (4.7) and (4.8) we obtain

$$g\gamma \in \{g' \in N_G(U, H) : \mathsf{c}(\eta_H(g')) \le L\},$$

as we claimed.

4.10.1 Corollary. *Let* $\mathbf{H} \in \mathcal{H}$ *and let* $L > 0$*. The set*

$$\{g \in N_G(U, H) : \mathsf{c}(\eta_H(g)) \le L\}\Gamma/\Gamma$$

is a closed subset of G/Γ.

Proof. Recall that $N_G(U, H) = \{g \in G : z \wedge \eta_H(g\gamma) = 0 \text{ for all } z \in \mathcal{B}_U\}$. The claim thus follows from Lemma 4.10. □

4.11 Theorems A and B below will be used in the proof of Lemma 4.12. We begin by recalling an effective versions of Hilbert's Nullstellensatz theorem; the statement presented here is due to D. Masser, G. Wüstholz, [36, Thm. IV]; see also [45, 26] and references there.

Theorem (Effective Nullstellensatz)**.** *Assume* $f, f_1, \ldots, f_n \in \mathbb{Z}[t_1, \ldots, t_m]$ *have total degree at most* D_0 *and logarithmic height at most* h*. Suppose* f *vanishes at all the common zeros (if any) of* $\{f_i\}$ *in* $\mathbb{C}^m$*.*

Put $M = 2^{m-1}$*. Then there exist*

- *some* $b \in \mathbb{N}$ *with* $b \ll (8D_0)^{2M}$*,*
- $q_1, \ldots, q_n \in \mathbb{Z}[t_1, \ldots, t_m]$ *of total degree at most* $(8D_0)^{2M+1}$ *and logarithmic height at most* $(8D_0)^{4M-1}(\mathsf{h} + 8D_0\log(8D_0))$*, and*
- *some* $a \in \mathbb{Z}$ *with* $\log|a| \le (8D_0)^{4M-1}(\mathsf{h} + 8D_0\log(8D_0))$

so that

$$af^b = \sum_i q_i f_i.$$

We need the following theorem of W. Brownawell which can be thought of as a local version of the above theorem.

Theorem A (cf. [11])**.** *Let $f_1, \dots, f_n \in \mathbb{Z}[t_1, \dots, t_m]$ have total degree at most D_0 and logarithmic height at most h. If $f_1, \dots, f_n$ have no common zero within $0 < \mathcal{b} \le 1$ of some $w \in \mathbb{C}^m$, then*

$$\max\{|f_j(w)| : 1 \le j \le n\} \ge C_1 e^{-A_{10}\mathsf{h}}\Big(\frac{\|w\|^2}{\mathcal{b}}\Big)^{-A_{10}}$$

where C_1 and A_{10} are explicit constants depending only on n, m, and D_0.

In the p-adic setting, we have the following theorem. This theorem is proved by M. Greenberg; we reconstruct Greenberg's proof in Appendix Appendix A to make the dependence on the height of the polynomials in question explicit.

Theorem B (cf. [23] and [24])**.** *Let $f_1, \dots, f_n \in \mathbb{Z}[t_1, \dots, t_m]$ have total degree at most D_0 and logarithmic height at most h. There exists A_{11} depending only on m, n, and D_0 so that the following holds.*

Suppose $w_1, \dots, w_m \in \mathbb{Z}_p$ and $C_2 > 2A_{11}\mathsf{h}$ are such that

$$f_j(w_1, \dots, w_m) \equiv 0 \pmod{p^{C_2}} \quad \textit{for all } j.$$

Then, there exist $y_1, \dots, y_m \in \mathbb{Z}_p$ such that

$$y_i \equiv w_i \quad \Big(\mathrm{mod}\ \mathrm{p}^{\lceil \frac{\mathrm{C}_2 - \mathrm{A}_{11}\mathsf{h}}{\mathrm{A}_{11}} \rceil}\Big)$$

and $f_j(y_1, \dots, y_m) = 0$ for all j.

The following lemma is a crucial ingredient for our inductive argument in the proof of Theorem 3.2.

4.12 Lemma. *There exist A_{12}, A_{13}, and C_0 where C_0 depends on N, the number of places $\#\Sigma$, and polynomially on the finite primes in Σ and on $\mathrm{ht}(G)$ so that the following holds.*

Let $r > 1$, $\epsilon > 1$, and $g \in G$ be fixed. Suppose $\mathbf{H}_1, \mathbf{H}_2 < \mathbf{G}$ are two $\mathbb{Q}$-subgroups of class-$\mathcal{H}$ with $\mathsf{c}(\eta_{H_i}(g)) \le r$ for $i = 1, 2$. Assume that

$$\max_{z \in \mathcal{B}_U} \|z \wedge \eta_{H_i}(g)\| \le \epsilon \quad \textit{for } i = 1, 2. \tag{4.9}$$

Let $\mathbf{H}_{1,2} := \big(\mathbf{H}_1 \cap \mathbf{H}_2\big)^{\mathcal{H}}$. Then if $\epsilon \le C_0|g|^{-A_{12}}r^{-A_{12}}$, the group $\mathbf{H}_{1,2}$ is not trivial, $\mathrm{ht}(\mathbf{H}_{1,2}) \ll |g|^{\star}r^{\star}$, and

$$\max_{z \in \mathcal{B}_U} \|z \wedge \eta_{H_{1,2}}(g)\| \ll |g|^{A_{13}}r^{A_{13}}\epsilon^{1/A_{13}}. \tag{4.10}$$

Proof. First note that (4.9) and (2.3) imply the following:

$$\|\operatorname{Ad}(g^{-1})z \wedge v_{H_i}\| \ll |g|^\star \epsilon \quad \text{for } i = 1, 2 \text{ and all } z \in \mathcal{B}_U.$$

Rewriting this at the level of the Lie algebra, we have

$$\mathsf{d}(\operatorname{Ad}(g^{-1})z, \mathfrak{h}_i) \ll |g|^\star r^\star \epsilon \quad \text{for } i = 1, 2 \text{ and all } z \in \mathcal{B}_U, \tag{4.11}$$

where $\mathfrak{h}_i$ denotes the Lie algebra of $H_i = \mathbf{H}_i(\mathbb{Q}_\Sigma)$.

Now (2.3) and $\mathsf{c}(\eta_{H_i}(g)) \le r$ imply

$$\operatorname{ht}(\mathbf{H}_i) = \mathsf{c}(v_{H_i}) \ll r|g|^\star \quad \text{for } i = 1, 2.$$

Therefore, $\operatorname{ht}(\mathbf{H}_1 \cap \mathbf{H}_2) \ll \operatorname{ht}(\mathbf{H}_1) \cdot \operatorname{ht}(\mathbf{H}_2)$ and hence by (4.2) we have

$$\operatorname{ht}(\mathbf{H}_{1,2}) \ll |g|^\star r^\star.$$

As $\mathfrak{h}_1$ and $\mathfrak{h}_2$ are rational subspaces of $\mathfrak{g}$ with height $\ll |g|^\star r^\star$, the estimates (4.11) thus imply that

$$\mathsf{d}(\operatorname{Ad}(g^{-1})z, \mathfrak{h}_1 \cap \mathfrak{h}_2) \ll |g|^\star r^\star \epsilon \quad \text{for all } z \in \mathcal{B}_U,; \tag{4.12}$$

see, e.g., [18, §13.4].

For every finite place $p \in \Sigma$ let $\Omega_p = \mathbb{Q}_p$, and let $\Omega_\infty = \mathbb{C}$. Set $\Omega_\Sigma = \prod_\Sigma \Omega_p$. Let $\mathcal{N}$ denote the cone of ad-nilpotent elements in $\mathfrak{g} \otimes \Omega_\Sigma$. Then

$$\operatorname{Ad}(g^{-1})z \in \mathcal{N}. \tag{4.13}$$

There are $n, m \gg 1$ so that the subspace $\mathfrak{h}_1 \cap \mathfrak{h}_2$ and the cone $\mathcal{N}$ are $\mathbb{Q}$-varieties defined by

$$\{f_{\mathrm{sp},j} : 1 \le j \le n\} \subset \mathbb{Z}[t_1, \ldots, t_m] \quad \text{and} \quad \{f_{\mathrm{cn},j} : 1 \le j \le n\} \subset \mathbb{Z}[t_1, \ldots, t_m],$$

respectively;[4] further, the logarithmic heights h of these polynomials are bounded by

$$B_0 + \log r$$

for some B_0 depending on N, the number of places $\#\Sigma$, and polynomially on the finite primes in Σ and on $\operatorname{ht}(G)$.

In particular, conditions of Theorems A and B are satisfied for $\{f_{\mathrm{sp},j}\} \cup \{f_{\mathrm{cn},j}\}$. In view of Theorems A and B, thus, (4.12) and (4.13) imply the following estimate:

$$\mathsf{d}(\operatorname{Ad}(g^{-1})z, \mathcal{N} \cap ((\mathfrak{h}_1 \cap \mathfrak{h}_2) \otimes \Omega_\Sigma)) \ll |g|^\star r^\star \epsilon^\star.$$

[4]The subscript sp stands for subspace and cp stands for cone.

Let $\mathfrak{h}_{1,2} = \mathrm{Lie}(\mathbf{H}_{1,2})$. By the definition of the family $\mathcal{H}$, see §4.1, we have $\mathfrak{h}_{1,2}$ contains the Lie algebra generated by $\mathcal{N} \cap ((\mathfrak{h}_1 \cap \mathfrak{h}_2) \otimes \mathbf{\Omega}_\Sigma)$. Therefore, the above estimate implies that

$$\mathrm{d}(\mathrm{Ad}(g^{-1})z, \mathfrak{h}_{1,2} \otimes \mathbf{\Omega}_\Sigma) \ll |g|^\star r^\star \epsilon^\star \quad \text{for all } z \in \mathcal{B}_U. \tag{4.14}$$

Now since $\mathrm{Ad}(g^{-1})z \in \mathfrak{g}$, we get the following from (4.14):

$$\mathrm{d}(\mathrm{Ad}(g^{-1})z, \mathfrak{h}_{1,2}) \ll |g|^\star r^\star \epsilon^\star \quad \text{for all } z \in \mathcal{B}_U. \tag{4.15}$$

Equation (4.15) implies $\mathfrak{h}_{1,2} \neq \{0\}$ so long as the right-hand side of (4.15) is a sufficiently high power of $|g|^{-1}$; this is satisfied if $\epsilon \ll |g|^\star r^\star$. Equation (4.10) is now an immediate consequence of (4.15). □

5 Non-divergence of unipotent flows in $\mathrm{SL}_N(\mathbb{Q}_\Sigma)/\mathrm{SL}_N(\mathbb{Z}_\Sigma)$ with an application to almost invariant Lie algebras

In this section we recall the basic non-divergence results regarding the action of unipotent groups on $\mathrm{SL}_N(\mathbb{Q}_\Sigma)/\mathrm{SL}_N(\mathbb{Z}_\Sigma)$, and deduce some important corollaries that will play a central role in the following sections. The basic reference for this section is the paper [30] by Kleinbock and Tomanov, which can be viewed as a Σ-arithmetic adaptation of [29] by Kleinbock and Margulis (which itself relies on the non-divergence result of Margulis [32], perhaps the first general result regarding dynamics of unipotent groups on arithmetic quotients, and Dani [13]).

Some of the implicit multiplicative constants in this section satisfy a stricter requirement, i.e., they depend on N, $\#\Sigma$, and polynomially on the finite primes in Σ, but not on $\mathrm{ht}(\mathbf{G})$. We will explicate these by an index, i.e., we write $\ll_{N,\Sigma}$ or $\gg_{N,\Sigma}$ for these implicit multiplicative constants.

5.1 Let $\mathrm{GL}^1_N(\mathbb{Q}_\Sigma)$ denote the group

$$\mathrm{GL}^1_N(\mathbb{Q}_\Sigma) = \left\{ (g_v) \in \mathrm{GL}(\mathbb{Q}_\Sigma) : \prod_{v \in \Sigma} \det(g_v) = 1 \right\}.$$

Then we can identify $\mathrm{GL}^1_N(\mathbb{Q}_\Sigma)/\mathrm{GL}(\mathbb{Z}_\Sigma)$ with the space of discrete $\mathbb{Z}_\Sigma$-modules in $\mathbb{Q}_\Sigma^N$ of covolume 1, and there is a natural injective proper map from $\mathrm{SL}_N(\mathbb{Q}_\Sigma)/\mathrm{SL}_N(\mathbb{Z}_\Sigma)$ to $\mathrm{GL}^1_N(\mathbb{Q}_\Sigma)/\mathrm{GL}(\mathbb{Z}_\Sigma)$ obtained by assigning to $(g_v)_{v\in\Sigma}$ the $\mathbb{Q}_\Sigma$-module spanned by the elements in $\mathbb{Q}_\Sigma^N$ formed by taking the ith column of all g_v for $i = 1, \ldots, N$. In view of this, we will view $\mathrm{SL}_N(\mathbb{Q}_\Sigma)/\mathrm{SL}_N(\mathbb{Z}_\Sigma)$ as embedded in $\mathrm{GL}^1_N(\mathbb{Q}_\Sigma)/\mathrm{GL}(\mathbb{Z}_\Sigma)$.

Let $\Gamma_1 = \mathrm{GL}_N(\mathbb{Z}_\Sigma)$ and $G_1 = \mathrm{GL}^1_N(\mathbb{Q}_\Sigma)$. For $x = g\Gamma_1/\Gamma_1 \in G_1/\Gamma_1$, let

$$\alpha(x) = \max\{\, 1/\mathsf{c}(z) : z \in g\mathbb{Z}_\Sigma^N \setminus \{0\}\};$$

this function is a proper map from G_1/Γ_1 to $\mathbb{R}^+$ (as well as from the quotient space $\mathrm{SL}_N(\mathbb{Q}_\Sigma)/\mathrm{SL}_N(\mathbb{Z}_\Sigma)$ to $\mathbb{R}^+$) and any compact subset of G_1/Γ_1 is contained in the compact subset of the form $\{x : \alpha(x) < M\}$ for some $M > 0$. Let Δ be a $\mathbb{Z}_\Sigma$-submodule of rank k in a discrete $\mathbb{Z}_\Sigma$-module $g\mathbb{Z}_\Sigma^N$, say generated over $\mathbb{Z}_\Sigma$ by $v_1, \ldots, v_k \in \mathbb{Q}_\Sigma^N$. Then while $v_1, \ldots, v_k$ are not uniquely defined, the wedge $v_1 \wedge \cdots \wedge v_k$ in $\wedge^k \mathbb{Q}_\Sigma^N$ is, and we define $\mathsf{c}(\Delta) = \mathsf{c}(v_1 \wedge \cdots \wedge v_k)$. A $\mathbb{Z}_\Sigma$-submodule Δ of $g\mathbb{Z}_\Sigma^N$ is said to be **primitive** in $g\mathbb{Z}_\Sigma^N$ if it is maximal with respect to finite-index extensions, i.e., it is not a proper $\mathbb{Z}_\Sigma$-submodule of finite index in any $\mathbb{Z}_\Sigma$-submodule of $g\mathbb{Z}_\Sigma^N$.

The results of [30] are more general in that they deal with general "(c, α)-good" maps from a convex B in a product of parameter spaces over $\mathbb{Q}_v$ for v in some subset of Σ to G_1/Γ_1, but the basic non-divergence estimate of the paper [30, Thm. 9.4] gives the following:

5.2 Theorem (cf. [30])**.** *Let $U = \prod_{v\in\Sigma} U_v$ be a $\mathbb{Q}_\Sigma$-unipotent subgroup, $\mathsf{B}_U(e)$ an open ball in U and λ_k as in §2.9. Let $g \in \mathrm{GL}^1_N(\mathbb{Q}_\Sigma)$ and assume that for every primitive $\mathbb{Z}_\Sigma$-submodule Δ of $g\mathbb{Z}_\Sigma^N$ of rank $1 \le k \le N-1$*

$$\max_{u\in \mathsf{B}_U(e)} \mathsf{c}(\lambda_k(u)\Delta) \ge \eta. \tag{5.1}$$

Then

$$|\{\, u \in \mathsf{B}_U(e) : \alpha(\lambda_k(u)g\Gamma_1) > \epsilon^{-1}\}| < E\Big(\frac{\epsilon}{\eta}\Big)^{1/D} |\mathsf{B}_U(e)|,$$

with D depending only on N and E depending on N, $\#\Sigma$, and polynomially on finite primes in Σ.

In fact, the basic inductive argument used to prove Theorem 5.2, specifically [30, Thm. 6.1] can be used to provide a more precise result that would be important for us in the sequel. This result does not seem to appear in the literature. One can view Kleinbock's [27, Thm. 0.2] as a step in this direction, and a result very close to what we give below can be found in a draft by Breuillard and de Saxce [10].

For $g\Gamma_1 \in G_1/\Gamma_1$ and $1 \le i \le N-1$ let

$$\alpha_i(g\Gamma_1) = 1/\min\left\{\mathsf{c}(\Delta) : \Delta \text{ is a primitive } \mathbb{Z}_\Sigma\text{-submodule of } g\mathbb{Z}_\Sigma^N \text{ of rank } i\right\}.$$

5.3 Theorem. *With the notations of Theorem 5.2 (but without the assumption* (5.1)*), there are $0 = k_0 < k_1 < k_2 < \cdots < k_\ell < k_{\ell+1} = N$, and primitive*

$\mathbb{Z}_\Sigma$-submodules $\Delta_{k_1} < \Delta_{k_2} < \cdots < \Delta_{k_\ell}$ of $g\mathbb{Z}_\Sigma^N$ of rank corresponding to their index so that if $\eta(k_0), \ldots, \eta(k_{\ell+1}) \in (0, 1]$ is defined by

$$\begin{aligned} &\eta(0) = \eta(N) = 1, \\ &\eta(k_i) = \max_{u \in \mathsf{B}_U(e)} \mathsf{c}(\lambda_k(u)\Delta_{k_i}) \quad \textit{for } 1 \le i \le \ell, \end{aligned} \tag{5.2}$$

then $\eta(\bullet)$ can be extended to a function $[1, N] \to (0, 1]$ so that $-\log \eta : [1, N] \to \mathbb{R}^+$ is concave and linear on each interval $[k_0, k_1], \ldots, [k_\ell, k_{\ell+1}]$ and

$$\left| \left\{ u \in \mathsf{B}_U(e) : \exists i \textit{ s.t. } \frac{\alpha_i(\lambda_k(u)g\Gamma_1)^{-1}}{\eta(i)} < \epsilon^i \right\} \right| < E\epsilon^{1/D} |\mathsf{B}_U(e)|,$$

with D depending only on N and E depending on N, $\#\Sigma$, and polynomially on finite primes in Σ. Moreover, given a primitive $\mathbb{Z}_\Sigma$-submodule $\tilde{\Delta} < g\mathbb{Z}_\Sigma^N$, we can choose $\Delta_{k_1} < \Delta_{k_2} < \cdots < \Delta_{k_\ell}$ so that

$$\eta(\mathrm{rk}(\tilde{\Delta})) \le \max_{u \in \mathsf{B}_U(e)} \mathsf{c}(\lambda_k(u)\tilde{\Delta}).$$

Note that it easily follows from the Σ-arithmetic version of Minkowski's second theorem, [4, §C.2, specifically Thm. C.2.11], that under the assumption (5.2) for any $u \in \mathsf{B}_U(e)$ one can complete the partial flag $\Delta_{k_1} < \Delta_{k_2} < \cdots < \Delta_{k_\ell}$ of submodules of $g\mathbb{Z}_\Sigma^N$to a full flag of primitive $\mathbb{Z}_\Sigma$-modules $\Delta_1 < \cdots < \Delta_{N-1}$ so that if $k_i < r < k_{i+1}$ and $\tau = (k_{i+1} - r)/(k_{i+1} - k_i)$ then

$$\begin{aligned} \mathsf{c}(\lambda_k(u)\Delta_r) &< A\mathsf{c}(\lambda_k(u)\Delta_{k_i})^\tau \mathsf{c}(\lambda_k(u)\Delta_{k_{i+1}})^{1-\tau} \\ &\le A\eta(k_i)^\tau \eta(k_{i+1})^{1-\tau} = A\eta(r), \end{aligned} \tag{5.3}$$

with A depending only on N and Σ. Hence for all $u \in \mathsf{B}_U(e)$

$$\frac{\alpha_r(\lambda_k(u)g\Gamma_1)^{-1}}{\eta(r)} < A.$$

Proof. Consider the (finite) collection of all primitive $\mathbb{Z}_\Sigma$-submodules $\Delta < g\mathbb{Z}_\Sigma^N$ so that

$$\max_{u \in \mathsf{B}_U(e)} \mathsf{c}(\lambda_k(u)\Delta) < 1, \tag{5.4}$$

and for each such Δ, let

$$\eta_\Delta = \max_{u \in \mathsf{B}_U(e)} \mathsf{c}(\lambda_k(u)\Delta).$$

From all the possible partial flags of primitive $\mathbb{Z}_\Sigma$-submodules $\Delta_{k_1} < \Delta_{k_2} < \cdots < \Delta_{k_\ell}$ with all Δ_{k_i} in this subcollection, choose one for which the convex hull of the pairs of points

$$\{(0, 0), (k_1, -\log \eta_{\Delta_{k_1}}), \ldots, (k_1, -\log \eta_{\Delta_{k_\ell}}), (N, 0)\} \tag{5.5}$$

is maximal (with respect to the usual partial order by inclusion on subsets of $\mathbb{R}^2$). There could be more than one possible choice, but any one of these choices would be good enough for us, and if $\tilde{\Delta}$ satisfies (5.4) we can choose such a $\Delta_{k_1} < \Delta_{k_2} < \cdots < \Delta_{k_\ell}$ so that the convex hull of the points in (5.5) contains the point $(\mathrm{rk}(\tilde{\Delta}), -\log \eta_{\tilde{\Delta}})$.

Fix the choice of primitive $\mathbb{Z}_\Sigma$-submodules $\Delta_{k_1} < \Delta_{k_2} < \cdots < \Delta_{k_\ell}$ and let $\eta : [0, N] \to \mathbb{R}^+$ be as in the statement of the theorem. Then the graph of $-\log \eta(\bullet)$ forms the upper half of the boundary of the convex hull of the set in (5.5), and $\eta(\mathrm{rk}(\tilde{\Delta})) \leq \eta_{\tilde{\Delta}}$.

By the choice of the Δ_{k_i} and definition of $\eta(\bullet)$, it follows that for any $1 \leq r \leq N-1$ and any $\mathbb{Z}_\Sigma$-primitive submodule Δ of rank r of $g\mathbb{Z}_\Sigma^N$ compatible with $\Delta_{k_1} < \Delta_{k_2} < \cdots < \Delta_{k_\ell}$,

$$\max_{u \in \mathsf{B}_U(e)} \mathsf{c}(\lambda_k(u)\Delta) \geq \eta(r).$$

Applying [30, Thm. 6.1] similarly to the way it is used to prove [30, Thm. 9.3], but with the poset used in [30, Thm. 6.1] being the collection of $\mathbb{Z}_\Sigma$-submodules of $g\mathbb{Z}_\Sigma^N$ compatible with the chosen partial flag $\Delta_{k_1} < \Delta_{k_2} < \cdots < \Delta_{k_\ell}$ one obtains that outside a subset $\mathsf{C} \subset \mathsf{B}_U(e)$ of measure $|\mathsf{C}| \ll_{N,\Sigma} \epsilon^\star$ we can find for every $u \in \mathsf{B}_U(e) \setminus \mathsf{C}$ a completion $\Delta_1 < \cdots < \Delta_{N-1}$ (depending on u) of the fixed partial flag $\Delta_{k_1} < \cdots < \Delta_{k_\ell}$ so that for every i

$$\epsilon\eta(i) \leq \mathsf{c}(\lambda_k(u)\Delta_i) \leq A'\eta(i), \tag{5.6}$$

with A' depending only on N and Σ. To be precise, we apply a variant of [30, Thm. 6.1] where the marking equations (M1) and (M2) on [30, p. 540] for a partial flag $\mathfrak{G}_u$ (compatible with our fixed flag $\Delta_{k_1} < \cdots < \Delta_{k_\ell}$) are replaced by (in the notations of this paper)

(M1) $\eta(\mathrm{rk}\Delta) \geq \mathsf{c}(\lambda_k(u)\Delta) \geq \epsilon\eta(\mathrm{rk}\Delta)$ for every $\Delta \in \mathfrak{G}_u$

(M2) $\mathsf{c}(\lambda_k(u)\Delta) \geq \eta(\mathrm{rk}\Delta)$ for every Δ compatible with $\mathfrak{G}_u$ and $\Delta_{k_1} < \cdots < \Delta_{k_\ell}$ but not in $\mathfrak{G}_u$.

The argument of [30, Thm. 6.1] would give us that for u outside the set C as above there exists a partial flag $\mathfrak{G}_u$ for which (M1), (M2) hold. Subsequently applying Minkowski's 2nd theorem (cf. note following the statement of Theorem 5.3, particularly (5.3)) we can complete the flag $\mathfrak{G}_u$ to a full flag so that (5.6) holds.

Such a marking was used in [30] (and [29]) to show that there is no primitive $v \in \lambda_k(u)g\mathbb{Z}_\Sigma^N$ with small $\mathsf{c}(v)$, i.e., to control $\alpha(g\Gamma_1) = \alpha_1(g\Gamma_1)$, but in fact can be used to show $\alpha_i(\lambda_k(u)g\Gamma_1) \ll_{N,\Sigma} \epsilon^{-i}\eta(i)^{-1}$, as we now show.

The proof is by induction on the rank of the submodule $\Delta < g\mathbb{Z}_\Sigma^N$, and all implicit constants may depend on the step in the induction. Note that since $-\log \eta(i)$ is a concave function,

$$\frac{\eta(i)}{\eta(i-1)} \le \frac{\eta(i+1)}{\eta(i)} \quad \text{for all } 1 \le i \le N-1. \tag{5.7}$$

We also recall the following important inequality for any primitive $\Delta,\, \Delta' < g\mathbb{Z}_\Sigma^N$ and any $u \in U$,

$$\mathsf{c}(u\Delta)\,\mathsf{c}(u\Delta') \ge \mathsf{c}(u\Delta \cap u\Delta')\,\mathsf{c}(u\Delta' + u\Delta)/A, \tag{5.8}$$

with A depending only on N, Σ.

We start induction with rank one primitive submodules $\mathbb{Z}_\Sigma v < g\mathbb{Z}_\Sigma^N$. Let i be such that $v \in \Delta_{i+1}$ but not in Δ_i (where for this purpose we take $\Delta_0 = \{0\}$ and $\Delta_N = g\mathbb{Z}_\Sigma^N$). Then by (5.8) and (5.7),

$$\mathsf{c}(\lambda_k(u)\mathbb{Z}_\Sigma v) \ge \frac{\mathsf{c}(\lambda_k(u)\Delta_{i+1})}{A\mathsf{c}(\lambda_k(u)\Delta_i)} \ge \frac{\epsilon\eta(i+1)}{AA'\eta(i)} \ge \frac{\epsilon\eta(1)}{AA'}.$$

Consider now a rank-r primitive submodule $\Delta < g\mathbb{Z}_\Sigma^N$, let i be such that $\Delta < \Delta_{i+1}$ and i is minimal such (clearly, $i+1 \ge r$). Applying (5.8) once again, we obtain

$$\mathsf{c}(\lambda_k(u)\Delta)\,\mathsf{c}(\lambda_k(u)\Delta_i) \ge \mathsf{c}(\lambda_k(u)(\Delta_i \cap \Delta))\,\mathsf{c}(\lambda_k(u)\Delta_{i+1})/A.$$

By induction $\mathsf{c}(\lambda_k(u)(\Delta_i \cap \Delta)) \gg_{N,\Sigma} \epsilon^{r-1}\eta(r-1)$ hence

$$\begin{aligned}
\mathsf{c}(\lambda_k(u)\Delta) &\gg_{N,\Sigma} \epsilon^{r-1}\eta(r-1)\frac{\mathsf{c}(\lambda_k(u)\Delta_{i+1})}{\mathsf{c}(\lambda_k(u)\Delta_i)} \\
&\gg_{N,\Sigma} \epsilon^{r}\eta(r-1)\frac{\eta(i+1)}{\eta(i)} \\
&\gg_{N,\Sigma} \epsilon^{r}\eta(r-1)\frac{\eta(r)}{\eta(r-1)} = \epsilon^r\eta(r),
\end{aligned}$$

and we are done. □

A key ingredient in the works of Margulis, Dani, Kleinbock–Margulis, and Kleinbock–Tomanov quoted above is an estimate on the size of the set where a polynomial function is small. The result needed, at least for the real case (i.e., $\Sigma = \{\infty\}$) is known as Remez inequality, and is used in [29] and [30] to verify the "(C, α)-good" property. Since we will also use it in the sequel, we quote it below (in a slightly sharper form, though this is not relevant to us; cf., e.g., [29, Prop. 3.2]).

5.4 Lemma. *Let $\mathfrak{F}$ be a local field with absolute value $|\ |$. Let B be a compact convex subset of $\mathfrak{F}^r$, and let $f \in \mathfrak{F}[t_1, \cdots, t_r]$ be a nonzero polynomial of degree d. Then for any $\delta > 0$ we have*

$$\left|\left\{z \in \mathsf{B} : |f(z)| < \delta \sup_{z\in\mathsf{B}} |f(z)|\right\}\right| \leq c\delta^{1/d}|\mathsf{B}|, \tag{5.9}$$

where $|\mathsf{K}|$ denotes the Haar measure of K for any subset $\mathsf{K} \subset \mathfrak{F}^r$, with c depending only on d and r.

See [12] for a proof for $k = \mathbb{R}$; the general case is essentially identical.

Sketch of proof. Let $\delta' = \delta \sup_{z\in\mathsf{B}} |f(z)|$. For $r = 1$ this follows from Lagrange's interpolation formula. For higher dimension, let $x \in \mathsf{B}$ be such that $f(x) = \sup_{z\in\mathsf{B}} |f(z)|$. Then there is a line ℓ through x where

$$\frac{|\{z \in \mathsf{B} : |f(z)| < \delta'\} \cap \ell|}{|\mathsf{B} \cap \ell|} > c_1 \frac{|\{z \in \mathsf{B} : |f(z)| < \delta'\}|}{|\mathsf{B}|}.$$

Since $x \in \ell$ by the choice of x we have

$$\sup_{z\in\mathsf{B}} |f(z)| = \sup_{z\in\mathsf{B}\cap\ell} |f(z)|;$$

now apply the one dimensional result. □

5.5 Lemma. *Let $\Sigma' \subset \Sigma$. For all positive integers r and d, there exists explicit constant $c = c(r, d, \Sigma')$ with the following property. For every $\upsilon \in \Sigma'$ and every $1 \leq j \leq r'_\upsilon$ let $f_{\upsilon,j} \in \mathbb{Q}_\upsilon[t_1, \cdots, t_{r_\upsilon}]$ be a nonzero polynomial of degree $\leq d$. Define*

$$f_\upsilon(t_1, \cdots, t_{r_\upsilon}) = \|(f_{\upsilon,1}(t), \ldots, f_{\upsilon,r'_\upsilon}(t))\|_\upsilon = \max\{|f_{\upsilon,j}(t)|_\upsilon : 1 \leq j \leq r'_\upsilon\}.$$

Let $\mathsf{B} = \prod_{\upsilon\in\Sigma'} \mathsf{B}_\upsilon$ where B_υ is a convex set in $\mathbb{Q}_\upsilon^{r_\upsilon}$ for each υ, and set

$$F(t_{\upsilon i} : \upsilon \in \Sigma', 1 \leq i \leq r_\upsilon) = \prod_{\upsilon\in\Sigma'} f_\upsilon(t_1, \cdots, t_{r_\upsilon}).$$

Then for any $\delta > 0$ we have

$$\left|\left\{z \in \mathsf{B} : F(z) < \delta \sup_{z\in\mathsf{B}} F(z)\right\}\right| \leq c|\log \delta|^{\#\Sigma'-1}\delta^{1/d}|\mathsf{B}|.$$

Similarly, if we put $F(t_{\upsilon i}) = \max_{\Sigma'} f_\upsilon(t_{\upsilon i})$, then

$$\left|\left\{z \in \mathsf{B} : F(z) < \delta \sup_{z\in\mathsf{B}} F(z)\right\}\right| \leq c\delta^{1/d}|\mathsf{B}|.$$

Proof. We first prove the first claim. Hence, let $F = \prod_{v\in\Sigma'} f_v$ be as in that statement; note that $\max F = \prod \max f_v$. Moreover, (5.9) holds true for f_v in place of $|f|$; see, e.g., [30, Lemma 3.1].

Note also that it suffices to prove the lemma for $\delta = 2^{-m}$ where m is a non-negative integer. For all non-negative integers m' and any $v \in \Sigma'$, put

$$\mathsf{B}_v^{f_v, m'} = \Big\{ z \in \mathsf{B}_v \ : f_v(z) \leq 2^{-m'} \max_{\mathsf{B}_v} f_v \Big\}.$$

Then we have

$$\Big\{ z \in \mathsf{B} : F(z) < 2^{-m} \sup_{z\in\mathsf{B}} F(z) \Big\} = \bigcup \prod_{\Sigma'} \mathsf{B}_v^{f_v, m_v}$$

where the union is taken over all partitions $m = \sum_{\Sigma'} m_v$ with m_v non-negative integer for all $v \in \Sigma$.

Now by (5.9) applied for f_v implies that $|\mathsf{B}_v^{f_v, m_v}| \leq C 2^{-m_v/d} |\mathsf{B}_v|$ for all $v \in \Sigma'$ and m_v. The claim follows from this as the number partitions $m = \sum_{\Sigma'} m_v$ is $\leq m^{\#\Sigma'-1}$.

To see the second claim, let v be so that $\max_{\mathsf{B}_v} f_v = \max F$. The claim then follows from the fact that (5.9) holds for f_v. $\square$

Note that replacing $\frac{1}{d}$ with $\frac{1}{d} - \epsilon$, for a small enough ϵ depending only on d and the constant c by a bigger constant depending on Σ' if necessary, we have the following. There exists some $\alpha = \alpha(d)$ so that for all F as in Lemma 5.5 we have

$$\Big| \Big\{ z \in \mathsf{B} : F(z) < \delta \sup_{z\in\mathsf{B}} F(z) \Big\} \Big| \leq c\delta^{\alpha} |\mathsf{B}| \tag{5.10}$$

where $c = c(r, d, \Sigma')$.

In the sequel, we will deal with functions defined on U of the form $u \mapsto \mathsf{c}(\eta_H(\lambda_k(u)g))$ and $u \mapsto \|z \wedge \eta_H(\lambda_k(u)g)\|$, see §4 for the notation. We let α be so that (5.10) holds true for all of these functions; note that α depends only on N.

5.6 Lemma. *There exists some A_{14} so that the following holds. Let $\mathbf{H} \in \mathcal{H}$ and $g \in G$. Put $\epsilon_g = \max\{ \|z \wedge \eta_{M_H}(g)\| : z \in \mathcal{B}_U \}$. Assume $\epsilon_g > 0$, i.e. that $g^{-1}Ug$ does not normalize H. Then*

$$|\{ u \in \mathsf{B}_U(e) : \mathsf{c}(\eta_H(\lambda_k(u)g)) \leq R \}| \ll (R|g| \,\mathrm{ht}(\mathbf{H})/\epsilon_g)^{A_{14}} e^{-k/A_{14}}.$$

We need the following lemma for the proof of Lemma 5.6.

5.7 Lemma. *There exists some A_{15} so that the following holds. Let the notation and assumptions be as in Lemma 5.6. Moreover, let* $\mathrm{Der}\rho_H$ *denote the derivative of ρ_H. Then*

$$\max\{\,\|\mathrm{Der}\rho_H(z)\eta_H(g)\| : z \in \mathcal{B}_U\} \gg \delta$$

where $\delta = (\epsilon_g\, \mathrm{ht}(\mathbf{H})^{-1}|g|^{-1})^{A_{15}}$.

Proof. Let $b > 0$ and assume that $\max\{\,\|\mathrm{Der}\rho_H(z)\eta_H(g)\| : z \in \mathcal{B}_U\} \le b$. Using (2.3), then we have

$$\max\{\,\|\mathrm{Der}\rho_H(\mathrm{Ad}(g^{-1})z)v_H\| : z \in \mathcal{B}_U\} \ll |g|^\star b. \tag{5.11}$$

Recall from the definition of $\mathbf{L_H}$ that

$$\mathrm{Lie}(L_H) = \{\,w \in \mathfrak{g} : \mathrm{Der}\rho_H(w)v_H = 0\}.$$

That is: $\mathrm{Lie}(L_H)$ is the kernel of the linear map $w \mapsto \mathrm{Der}\rho_H(w)v_H$ from $\mathfrak{g}$ to V_H. The vector v_H is an integral vector of size $\mathrm{ht}(\mathbf{H})$. Therefore, the map

$$w \mapsto \mathrm{Der}\rho_H(w)v_H$$

can be realized by an integral matrix whose entries are bounded by $\mathrm{ht}(\mathbf{H})^\star$.

Now by (5.11), for all $z \in \mathfrak{u}$ with $\|z\| = 1$ the vector $\mathrm{Ad}(g^{-1})z$ almost belongs to the kernel of this map; in view of the above bound we get that

$$\mathrm{d}(\mathrm{Ad}(g^{-1})z, \mathrm{Lie}(L_H)) \ll |g|^\star\, \mathrm{ht}(\mathbf{H})^\star b^\star, \tag{5.12}$$

see, e.g., [18, §13.4].

Recall that $\mathfrak{u}$ is a nilpotent Lie algebra and $\mathbf{M_H} = \mathbf{L_H^{\mathcal{H}}}$. Hence, arguing as in the proof of Lemma 4.12, i.e., using Theorems A and B, we get the following from (5.12):

$$\mathrm{d}(\mathrm{Ad}(g^{-1})z, \mathrm{Lie}(M_H)) \ll |g|^\star\, \mathrm{ht}(\mathbf{H})^\star b^\star \quad \text{for all } z \in \mathcal{B}_U.$$

The above estimate thus implies that

$$\|z \wedge \eta_{M_H}(g)\| \ll |g|^\star r^\star b^\star \quad \text{for all } z \in \mathcal{B}_U,$$

as we wanted to show. □

Proof of Lemma 5.6. In view of Lemma 5.5, it suffices to prove that

$$\max\{\,\mathsf{c}(\eta_H(\lambda_k(u)g)) : u \in \mathsf{B}\} \gg \Big(\frac{\epsilon_g}{\mathrm{ht}(\mathbf{H})|g|}\Big)^\star e^{\star k}. \tag{5.13}$$

To see this, for any $z \in \mathcal{B}_U$ define

$$f_z(t) = \rho_H(\exp(tz))\eta_H(g).$$

Then f_z is a polynomial map from $\mathbb{Q}_v$ into V_H. Let us write $f_z = c_{z,0} + \hat{f}_z$ where $c_{z,0} \in V_H$ and $\hat{f}_z(0) = 0$.

Let δ be as in the previous lemma. By the conclusion of that lemma, there exists some $z_0 \in \mathcal{B}_U$ so that

$$\max\{\,|c|_v\ : c \text{ is a coefficient of } \hat{f}_{z_0}\} \gg \delta^\star. \tag{5.14}$$

For any nonzero $T \in \mathbb{Q}_v$, define the renormalized polynomial

$$\hat{f}_{z_0,T}(t) := \tfrac{1}{T}\hat{f}_{z_0}(Tt).$$

Then by (5.14), we have $\sup_{|t|_v \le 1} \|\hat{f}_{z_0,T}(t)\| \gg \delta^\star$.

Hence, there exists some $v \in \Sigma$ so that

$$\max\{u_v \in \mathsf{B}_{U_w}(e) : \|\eta_H(\lambda_k(u_v)g)\|_v\} \gg \delta^\star e^{\star k};$$

we also used the fact that for all $w \in \Sigma$ we have $\|\eta_H(g)\|_w \gg |g|^{-\star}\,\mathrm{ht}(\mathbf{H})^{-\star}$; this lower bound follows as v_H is an integral vector whose ∞-norm is $\mathrm{ht}(\mathbf{H})$.

Altogether, we get that

$$\begin{aligned}
&\max\{\mathsf{c}(\eta_H(\lambda_k(u)g) : u \in \mathsf{B}_U(e)\}\\
&\qquad \ge \max\{\|\eta_H(\lambda_k(u_v)g\|_v \prod_{w \ne v} \|\eta_H(g)\|_w : u_v \in \mathsf{B}_{U_v}(e)\}\\
&\qquad \ge \delta^\star |g|^{-\star}\,\mathrm{ht}(\mathbf{H})^{-\star} e^{\star k};
\end{aligned}$$

this completes the proof of (5.13) and hence the lemma. □

5.8 Proposition. *There is a constant D' depending only on N so that the following holds. Let $\mathbf{H} \in \mathcal{H}$ and $r > 1$. Suppose that $k > 1$ and*

$$\mathsf{c}(\eta_H(\lambda_k(u)g)) < r \quad \textit{for all } u \in \mathsf{B}_U(e). \tag{5.15}$$

Then

$$\mathsf{c}(\eta_{M_H}(\lambda_k(u)g)) \ll r^\star |g|^\star \quad \textit{for all } u \in \mathsf{B}_U(e),$$

moreover, for all $u \in \mathsf{B}_U(e)$ and $z \in \mathcal{B}_U$ we have

$$\|z \wedge \eta_{M_H}(\lambda_k(u)g)\| \ll r^\star |g|^\star e^{-k/D'}.$$

Proof. Let $l = \dim(\mathbf{H})$. Recall that v_H denotes the integer vector corresponding to Lie($\mathbf{H}$) in $\wedge^l \mathfrak{g} \subset \wedge^l \mathfrak{sl}_N(\mathbb{Q}_\Sigma)$ — here and in what follows we view $\wedge^r \mathfrak{g}$ as a rational subspace of $\wedge^r \mathfrak{sl}_N(\mathbb{Q}_\Sigma)$ of height $\ll 1$ (recall from §2.4 that the implicit constants for $\ll$ and $\gg$ are allowed to depend polynomialy on ht($\mathbf{G}$)).

In the notations of §2.9, let $\mathsf{B} = \mathsf{B}_U(e)$ and set

$$\vartheta = 0.1\frac{|\mathsf{B}|}{|\lambda_1(\mathsf{B})|}.$$

For any primitive $\mathbb{Z}_\Sigma$-submodule Δ of $g\mathbb{Z}_\Sigma^N$, it holds that $\mathsf{c}(\Delta) \gg |g|^{-\mathrm{rk}\Delta}$, hence by Theorem 5.2 there exists a subset $\mathsf{B}_g \subset \lambda_k(\mathsf{B})$ with

$$|\lambda_k(\mathsf{B}) \setminus \mathsf{B}_g| < \vartheta|\lambda_k(\mathsf{B})|$$

so that

$$\alpha(ug\Gamma) \ll |g|^\star \quad \text{for all } u \in \mathsf{B}_g.$$

This implies that for every $u \in \mathsf{B}_g$ there exists some $\gamma_u \in \mathrm{SL}_N(\mathbb{Z}_\Sigma)$ so that

$$|ug\gamma_u^{-1}| \ll |g|^\star. \tag{5.16}$$

Now (5.16) and (5.15) imply that

$$\mathsf{c}(\gamma_u v_H) \ll |ug\gamma_u^{-1}|^\star \cdot \mathsf{c}(ugv_H) \ll |g|^\star r. \tag{5.17}$$

Applying a similar argument to the integral vector $w \in \wedge^{\dim \mathbf{G}} \mathfrak{sl}_N(\mathbb{Q}_\Sigma)$ corresponding to $\wedge^{\dim\mathbf{G}}\mathrm{Lie}(\mathbf{G})$ and using the fact that

$$w = ugw = ug\gamma_u^{-1}\gamma_u w,$$

we have that $\mathrm{Ad}(\gamma_u)\mathfrak{g}$ is a rational subspace of $\mathfrak{sl}_N(\mathbb{Q}_\Sigma)$ of height $\ll |g|^\star$.

Define $\mathbf{L}' = \{g \in \mathrm{SL}_N(\mathbb{Q}_\Sigma) : gv_H = v_H\}$. It follows from (5.15) applied with $u = e$ that $\mathsf{c}(v_H) \ll |g|^\star r$; hence, $\mathrm{ht}(\mathbf{L}') \ll |g|^\star r$.

Moreover, the definitions imply that $\mathbf{L_H} = \mathbf{G} \cap \mathbf{L}'$ and that $\mathbf{M_H} = \mathbf{L_H}^{\mathcal{H}}$, see §4.7. Further, in view of (4.4) we have $\mathrm{ht}(\mathbf{M_H}) \ll |g|^\star r^\star$.

Similarly, for each $u \in \mathsf{B}_g$ define

$$\mathbf{L}'_u = \{g \in \mathrm{SL}_N(\mathbb{Q}_\Sigma) : g\gamma_u v_H = \gamma_u v_H\} = \gamma_u \mathbf{L}' \gamma_u^{-1};$$

then $\mathrm{ht}(\mathbf{L}'_u) \ll |g|^\star r$. Put $\mathbf{L}_u = \gamma_u \mathbf{L_H} \gamma_u^{-1}$, and let $\mathbf{M}_u = \mathbf{L}_u^{\mathcal{H}} = \gamma_u \mathbf{M_H} \gamma_u^{-1}$. Then $\mathrm{ht}(\mathbf{M}_u) = \mathsf{c}(\gamma_u v_{M_H}) \ll |g|^\star r^\star$.

For every $u \in \mathsf{B}_g$ we have

$$\mathsf{c}(\eta_{M_H}(ug)) \ll |ug\gamma_u^{-1}|^\star \mathsf{c}(\gamma_u v_{M_H}) \ll |g|^\star r^\star. \tag{5.18}$$

Since $u \mapsto \eta_{M_H}(ug\gamma)$ is a polynomial, the estimate in (5.18) and Lemma 5.5 imply that

$$\mathsf{c}(\eta_{M_H}(\lambda_k(u)g)) \ll |g|^\star r^\star \quad \text{for all } u \in \mathsf{B}. \tag{5.19}$$

In particular, the first claim in the proposition holds.

We now turn to the proof of the second claim. Let $u \in \mathsf{B}_g \cap \lambda_{k-1}(\mathsf{B})$. By the choice of ϑ, this set has measure $\geq 0.9|\lambda_{k-1}(\mathsf{B})|$, in particular is non-empty. Let $\gamma_u \in \mathrm{SL}_N(\mathbb{Z}_\Sigma)$ be as in (5.16).

By (2.11), $\lambda_{k-\kappa}(\mathsf{B})\lambda_{k-1}(\mathsf{B}) \subset \lambda_k(\mathsf{B})$; hence by (5.15) we have

$$\mathsf{c}(\eta_H(\lambda_{k-\kappa}(v)ug)) < r \quad \text{for all } v \in \mathsf{B}.$$

Therefore, by Lemma 5.6 for every $u \in \mathsf{B}_g \cap \lambda_{k-1}(\mathsf{B})$,

$$\max\{\, \|z \wedge \eta_{M_H}(ug)\| : z \in \mathcal{B}_U\} \ll r|ug\gamma_u^{-1}| \cdot \mathrm{ht}(\gamma_u \mathbf{H}\gamma_u^{-1})e^{-k/A_{14}{}^2}.$$

For $u \in \mathsf{B}_g$, $|ug\gamma_u^{-1}| \ll |g|^\star$ and $\mathrm{ht}(\gamma_u \mathbf{H}\gamma_u^{-1}) = \mathsf{c}(\gamma_u v_H) \ll |g|^\star r$ hence for $u \in \mathsf{B}_g \cap \lambda_{k-1}(\mathsf{B})$

$$\max\{\, \|z \wedge \eta_{M_H}(ug)\| : z \in \mathcal{B}_U\} \ll |g|^\star r^\star e^{-k/\star}.$$

Since $u \mapsto z \wedge \eta_{M_H}(\lambda_{k-1}(u)g)$ is a polynomial, the above estimate together with Lemma 5.5 implies that

$$\max\{\, \|z \wedge \eta_{M_H}(\lambda_k(u)g)\| : z \in \mathcal{B}_U\} \ll |g|^\star r^\star e^{-k/\star} \quad \text{for all } u \in \mathsf{B}.$$

This finishes the proof of the second claim and the proposition. □

6 Non-divergence of unipotent flows for general algebraic groups

Consider now $\mathbf{G}$ a $\mathbb{Q}$-group of class-$\mathcal{H}$ and $G = \mathbf{G}(\mathbb{Q}_\Sigma)$ as in §2. Let $d = \dim \mathbf{G}$. Recall that $\mathfrak{g}(\mathbb{Z}_\Sigma) = \mathfrak{g} \cap \mathfrak{sl}_N(\mathbb{Z}_\Sigma)$, see §2.7. Let $U = \prod_{v\in\Sigma} U_v \subset G$ be a $\mathbb{Q}_\Sigma$-unipotent group as in §2.9. By assumption $\mathbf{G}$ is equipped with an embedding $\iota : \mathbf{G} \to \mathrm{SL}_N$, and a lattice Γ commensurable to $G \cap \mathrm{SL}_N(\mathbb{Z}_\Sigma)$. Taking a finite index subgroup if necessary, we assume that $\Gamma < G \cap \mathrm{SL}_N(\mathbb{Z}_\Sigma)$, and that $\mathrm{Ad}(\Gamma)$ preserves $\mathfrak{g}(\mathbb{Z}_\Sigma)$. Hence we get a finite to one map

$$G/\Gamma \to \mathrm{SL}_N(\mathbb{Q}_\Sigma)/\mathrm{SL}_N(\mathbb{Z}_\Sigma),$$

or using the adjoint representation a different map

$$G/\Gamma \to \mathrm{SL}_d(\mathbb{Q}_\Sigma)/\mathrm{SL}_d(\mathbb{Z}_\Sigma).$$

This latter map is in general not finite to one, but has compact fibers, since if $\mathbf{C}_\mathbf{G}$ denotes the connected component of the center of G (necessarily a unipotent group as $\mathbf{G}$ is of class-$\mathcal{H}$) the fibers would be a finite-to-one extension of the compact space $\mathbf{C}_\mathbf{G}(\mathbb{Q}_\Sigma)/\mathbf{C}_\mathbf{G}(\mathbb{Z}_\Sigma)$. Therefore we may apply Theorem 5.2 to the image of G/Γ to either of these quotient spaces to deduce that for every $g \in G$ and $\delta > 0$, there is a compact $K \subset G/\Gamma$ so that for every k for all $u \in \mathsf{B}_U(e)$ outside a set of measure $< \delta$ we have that $\lambda_k(u)g\Gamma \in K$.

Theorem 5.2 gives more: it also says that for a compact set K that does not depend on the point $g\Gamma$, if $\lambda_k(u)g\Gamma \notin K$ for a large set of $u \in \mathsf{B}_U(e)$ then there would be a $\mathbb{Z}_\Sigma$-submodule in $g\mathrm{SL}_N(\mathbb{Z}_\Sigma)$ which is not changed much by the action of U, at least not when we act by $\lambda_k(\mathsf{B}_U(e))$, and Theorem 5.3 gives somewhat finer information. However both of these theorems relate the properties of U orbits in G/Γ to the structure of the ambient $\mathrm{SL}_\bullet(\mathbb{Q}_\Sigma)/\mathrm{SL}_\bullet(\mathbb{Z}_\Sigma)$ and not some intrinsic algebraic structure of G/Γ.

In [16] Dani and Margulis prove (in the real case) that given a one-parameter unipotent subgroup u_t of G and $\delta > 0$, one can find a (fixed) compact subset $K \subset G/\Gamma$ so that if a trajectory of the one-parameter unipotent group u_t starting from $g\Gamma$ does not eventually spend $1-\delta$ of its time in K, then there is a $\mathbb{Q}$-parabolic subgroup $\mathbf{P} < \mathbf{G}$ so that $g \in \mathbf{P}(\mathbb{Q}_\Sigma)$. This information is intrinsic for G/Γ.

The purpose of this section is to provide an effective version of [16], where the existence of many $u \in \mathsf{B}_U(e)$ for which $\lambda_k(u)g\Gamma$ is outside a suitable fixed compact region is used to imply some Diophantine conditions at appropriate scale for $g\Gamma$. We note that in addition to [16], understanding intrinsically behavior of orbits near the cusp in arithmetic quotients G/Γ, this time for certain diagonalizable groups, was studied by Tomanov and Weiss in [51].

Recall from §2.7 the definition

$$X_\eta = \Big\{g\Gamma \in X : \min_{0 \neq z \in \mathfrak{g}(\mathbb{Z}_\Sigma)} \mathsf{c}(\mathrm{Ad}(g)z) \geq \eta\Big\}.$$

It follows from the discussion at the beginning of this section that for any $\eta > 0$, the set X_η is a compact subset of G/Γ.

6.1 Lemma. *There exists some* $0 < \kappa(N, \Sigma) < 1$ *with the following property. Let* $w \in \mathfrak{g}(\mathbb{Z}_\Sigma)$ *and suppose that there exists some* $g \in G$ *so that*

$$\mathsf{c}(\mathrm{Ad}(g)w) \leq \kappa(N, \Sigma).$$

Then w *is a nilpotent element.*

Proof. Let $\bar{\sigma}(w) = \prod \sigma$ where the product is taken over all the nonzero eigenvalues of w; if the product is empty, i.e., w is nilpotent, put $\bar{\sigma}(w) = 0$. Then $\bar{\sigma}(w) \in \mathbb{Q}$—indeed $\bar{\sigma}(w)$ is invariant under the Galois group of the splitting field of w. Further, since $w \in \mathfrak{g}(\mathbb{Z}_\Sigma)$, the product formula implies that either $\mathsf{c}(\bar{\sigma}(w)) \geq 1$ or $\bar{\sigma}(w) = 0$.

Let $\kappa > 0$ and assume that $\mathsf{c}(\mathrm{Ad}(g)w) \leq \kappa$ for some $g \in G$. Using Lemma 2.6, there exist some $r \in \mathbb{Z}_\Sigma^\times$ and a constant $A = A(N, \Sigma)$ so that

$$A^{-1}\mathsf{c}(\mathrm{Ad}(g)w)^{1/\#\Sigma} \leq \|r\,\mathrm{Ad}(g)w\|_v \leq A\mathsf{c}(\mathrm{Ad}(g)w)^{1/\#\Sigma}$$

for all $v \in \Sigma$. Therefore, all the eigenvalues of $r\,\mathrm{Ad}(g)w$ have v -norm $\ll_{N,\Sigma} \kappa^{\star/\#\Sigma}$.

Since $\mathsf{c}(r) = 1$ and $\mathrm{Ad}(g)w$ has the same eigenvalues as w, we get that $\mathsf{c}(\bar{\sigma}(w)) \geq 1$ cannot hold for small enough κ; hence, w is nilpotent. □

6.2 Lemma. *There exists some $\kappa'(N, \Sigma)$ with the following property. Let $V \subset \mathfrak{g}$ be a nonzero rational subspace, and let $v \in \wedge\mathfrak{g}(\mathbb{Z}_\Sigma)$ be a primitive integral vector corresponding to V. Assume that there is some $g \in G$ so that*

$$\max_{u \in \mathsf{B}_U(e)} \mathsf{c}(\lambda_k(u)gv) \leq \rho < \kappa'(N, \Sigma).$$

Then there exists a unipotent $\mathbb{Q}$-group $\mathbf{W} < \mathbf{G}$ so that

$$\max_{u \in \mathsf{B}_U(e)} \mathsf{c}(\lambda_k(u)gv_W) \ll \rho^{\dim(W)/\dim(V)} \tag{6.1}$$

where $v_W \in \wedge^{\dim(\mathbf{W})}\mathfrak{g}(\mathbb{Z}_\Sigma)$ is the primitive integer vector corresponding to $\mathbf{W}$ as in §2.3.

Proof. Let $d = \dim(\mathbf{G})$. We apply Theorem 5.3 on the image of G/Γ in $\mathrm{SL}_d(\mathbb{Q}_\Sigma)/\mathrm{SL}_d(\mathbb{Z}_\Sigma)$ obtained via Ad, with $\tilde{\Delta}$ the $\mathbb{Z}_\Sigma$-submodule of $\mathrm{Ad}(g)\mathfrak{g}(\mathbb{Z}_\Sigma)$ corresponding to V (or more precisely $gV(\mathbb{Q})$); let r denote the dimension of the $\mathbb{Q}$-subspace V (equivalently, $r = \mathrm{rk}(\tilde{\Delta})$).

Let $\eta(\bullet)$ and $\Delta_{k_1} < \Delta_{k_2} < \cdots < \Delta_{k_\ell}$ be as in that theorem. Then $\eta(r) \leq \rho$, hence by concavity of $-\log \eta(\bullet)$ we have that $\eta(1) \leq \rho^{1/r}$. Let s be maximal so that $\frac{\eta(s)}{\eta(s-1)} \leq \rho^{1/r}$; clearly $1 \leq r \leq d-1$, and because $-\log \eta(\bullet)$ changes its slope at s this implies that there is some $1 \leq j \leq \ell$ for which $k_j = \mathrm{rk}(\Delta_{k_j}) = s$.

We claim that (assuming $\kappa'(N, \Sigma)$ is small enough) the rational subspace of $\mathfrak{g}$ corresponding to $\mathrm{rk}(\Delta_{k_j})$ is the Lie algebra of a unipotent $\mathbb{Q}$-subgroup $\mathbf{W} < \mathbf{G}$. Let us denote by v_W the vector in $\wedge^{\mathrm{rk}(\Delta_{k_j})}\mathfrak{g}(\mathbb{Z}_\Sigma)$ corresponding to this rational subspace. By the choice of Δ_{k_j}, we have

$$\mathsf{c}(\lambda_k(u)gv_W) \equiv \mathsf{c}(\lambda_k(u)\Delta_{k_j}) = \eta(s) \leq \rho^{s/r}$$

for all $u \in \mathsf{B}_U(e)$; so (6.1) is satisfied.

It remains to show that Δ_{k_j} does indeed correspond to a rational nilpotent Lie algebra. Fix an $\epsilon > 0$ (depending only on N) so that for a set of $u \in \mathsf{B}_U(e)$ of size $\geq 0.5|\mathsf{B}_U(e)|$ we can find a completion $\Delta_1 < \cdots < \Delta_{N-1}$ (depending on u) of the fixed partial flag $\Delta_{k_1} < \cdots < \Delta_{k_\ell}$ so that for every i

$$\epsilon\eta(i) \leq \mathsf{c}(\lambda_k(u)\Delta_i) \ll_{N,\Sigma} \eta(i). \tag{6.2}$$

Indeed, we will only use the existence of one such u.

Using (5.8), we can deduce from (6.2) that for every $v \in \lambda_k(u)\Delta_{i+1}$ which is not in $\lambda_k(u)\Delta_i$

$$\mathsf{c}(v) \gg \frac{\mathsf{c}(\lambda_k(u)\Delta_{i+1})}{\mathsf{c}(\lambda_k(u)\Delta_i)} \gg \frac{\eta(i+1)}{\eta(i)}$$

(since ϵ depends only on N, we absorbed it in the implicit constant).

Moreover, by induction one easily shows that we can pick $v_i \in \lambda_k(u)\Delta_i$ (in particular, $v_i \in \mathrm{Ad}(\lambda_k(u)g)\mathfrak{g}(\mathbb{Z}_\Sigma)$) so that $v_1, \ldots, v_i$ generate $\lambda_k(u)\Delta_i$ and $\mathsf{c}(v_i) \ll \frac{\eta(i)}{\eta(i-1)}$.

We conclude that there is some A depending only on N and Σ so that if $v \in \lambda_k(u)\Delta_i$ but not in $\lambda_k(u)\Delta_{i-1}$ then $\mathsf{c}(v) \geq \frac{\eta(i)}{A\eta(i-1)}$ and $\mathsf{c}(v_i) \leq \frac{A\eta(i)}{\eta(i-1)}$.

Recall that for $i \leq s$ we have that $\frac{\eta(i)}{\eta(i-1)} \leq \rho^{1/r}$. As $\mathsf{c}([z, z']) \ll \mathsf{c}(z)\mathsf{c}(z')$, it follows that if $\kappa'(N, \Sigma)$ (hence also ρ) is small enough, for $i < i' \leq s$ we have that $\mathsf{c}([v_i, v_{i'}])$ is so small it forces $[v_i, v_{i'}]$ to belong to Δ_{i-1}.

It follows that Δ_{k_j} is closed under $[\cdot, \cdot]$. Since by Lemma 6.1 if $\kappa'(N, \Sigma)$ is small enough, all the v_i for $i \leq s = k_j$ are nilpotent, it follows that all $v \in \Delta_{k_j}$ are nilpotent. Hence Δ_{k_j} corresponds to the Lie algebra of a unipotent $\mathbb{Q}$-subgroup of $\mathbf{G}$. □

6.3 Theorem. *There exists a constant F depending on N and a constant E depending on N, $\#\Sigma$ and polynomialy on $\mathrm{ht}(\mathbf{G})$ and the primes in Σ so that for any $g \in G$, $k \geq 1$, and any $0 < \eta \leq 1/2$ at least one of the following holds:*

(1)
$$|\{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \notin X_\eta\}| \leq E\eta^{1/F}.$$

(2) *There exists a unipotent $\mathbb{Q}$-subgroup $\mathbf{W}$ of height $\mathrm{ht}(\mathbf{W}) \leq E|g|^F\eta^{1/F}$ so that*

$$\mathsf{c}(\eta_W(\lambda_k(u)g)) \leq E\eta^{1/F} \quad \textit{for all } u \in \mathsf{B}_U(e). \tag{6.3}$$

Moreover, if we put $\mathbf{M} = \mathbf{M}_{\mathbf{W}}$, then $\mathbf{M} \neq \mathbf{G}$,

$$\mathrm{ht}(\mathbf{M}) \leq E|g|^F\eta^{1/F},$$

and we have:

(a) *For all $u \in \mathsf{B}_U(e)$ we have*

$$\mathsf{c}(\eta_M(\lambda_k(u)g)) \leq E|g|^F\eta^{1/F}.$$

(b) *For all $u \in \mathsf{B}_U(e)$ we have*

$$\max_{z\in\mathcal{B}_U} \|z \wedge \eta_M(\lambda_k(u)g)\| \leq E|g|^F \eta^{1/F} e^{-k/F}.$$

Proof. We may assume $\eta < \kappa'(N, \Sigma)$ with $\kappa'(N, \Sigma)$ as in Lemma 6.2 since otherwise, for a sufficiently large implicit constant, alternative (1) in the statement of this theorem becomes vacuous.

Apply Theorem 5.2. Then either alternative (1) in the theorem holds, or there exists some primitive integral vector $v \in \wedge^r \mathfrak{g}(\mathbb{Z}_\Sigma)$ so that

$$\max_{u\in\mathsf{B}_U(e)} \mathsf{c}(\lambda_k(u)gv) \leq \eta.$$

By Lemma 6.2, we conclude that there is some unipotent $\mathbb{Q}$-group $\mathbf{W} < \mathbf{G}$ so that

$$\max_{u\in\mathsf{B}_U(e)} \mathsf{c}(\lambda_k(u)gv_W) \ll \eta^{\dim(W)/r}. \tag{6.4}$$

Applying (6.4) with $u = e$ we get that $\mathrm{ht}(\mathbf{W}) \ll |g|^\star \eta^{\dim(W)/r}$. Let $\mathbf{M}$ be as in (2) in the statement of this theorem. Then $\mathrm{ht}(\mathbf{M}) \ll \mathrm{ht}(\mathbf{W})^\star \ll |g|^\star \eta^\star$, see (4.2). Moreover, if η is small enough, then (6.4) (say for $u = e$) implies that

$$\mathsf{c}(gv_W) < 1/2.$$

As $\mathsf{c}(v_W) \geq 1$ this means that g does not fix v_W and so (since $\mathbf{G}$ is of class-$\mathcal{H}$, hence fixes v_H for any normal subgroup $\mathbf{H} \lhd \mathbf{G}$) the group $\mathbf{W}$ is not a normal subgroup of $\mathbf{G}$. In particular, $\mathbf{M} \neq \mathbf{G}$.

Applying Proposition 5.8 for $\mathbf{H} = \mathbf{W}$, we have that parts (a) and (b) in (2) of the statement of the theorem hold, concluding the proof of this theorem. □

Theorem 6.3 allows us to give a new, and arguably more elementary, proof to the main result of [16] (though the main ingredients are similar):

6.4 Corollary. *Suppose $\mathbf{G}$ is semisimple, and $g \in G$ is such that*

$$|\{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \notin X_\eta\}| > E\eta^{1/F} \quad \textit{for infinitely many } k. \tag{6.5}$$

Then $Ug \subset g\mathbf{P}(\mathbb{Q}_\Sigma)$ for some parabolic proper $\mathbb{Q}$-subgroup of $\mathbf{G}$.

6.5 Lemma. *Assume $\mathbf{G}$ is semisimple. Let $W \subset \mathfrak{g}$ be a rational subspace which generates a unipotent subalgebra. There exists a $\mathbb{Q}$-parabolic subgroup $\mathbf{P}(W)$ so that* $\mathrm{ht}(\mathbf{P}(W)) \ll \mathrm{ht}(W)^\star$ *and* $W \subset \mathrm{Lie}(\mathrm{R}_u(\mathbf{P}(W)))$.

Proof. Let $\hat{W}$ be the algebra generated by W and let $U_0 = \{\exp(\hat{W})\}$. Let $\mathbf{U}_0$ denote the corresponding algebraic group. Define inductively $\mathbf{U}_i = \mathrm{R}_u(N_{\mathbf{G}}(\mathbf{U}_{i-1}))$. Then

$$\mathbf{U}_i \subset N_{\mathbf{G}}(\mathbf{U}_{i-1}) \quad \text{and} \quad \mathbf{U}_{i-1} \subset \mathbf{U}_i.$$

This process terminates after $d \le \dim \mathbf{G}$ number of steps and gives a unipotent subgroup $\mathbf{U}_d$ so that $\mathbf{U}_d = \mathrm{R}_u(N_{\mathbf{G}}(\mathbf{U}_d))$. Therefore, $N_{\mathbf{G}}(\mathbf{U}_d)$ is a parabolic subgroup, see [8]; the claim holds with $\mathbf{P}(W) = N_{\mathbf{G}}(\mathbf{U}_d)$. □

Proof of Corollary 6.4. Suppose (6.5) holds along some sequence, say $\ell_1, \ell_2, \ldots$ of ks (to avoid confusion with k_i of Theorem 5.3 we use ℓ rather than k). Then by Theorem 6.3 there exists for every j a unipotent $\mathbb{Q}$-subgroup $\mathbf{W}_j$ and $\mathbf{M}_j = \mathbf{M}_{\mathbf{W}_j}$ with $\mathbf{M}_j \neq \mathbf{G}$ so that the heights of $\mathbf{W}_j$ and $\mathbf{M}_j$ are bounded uniformly in j and so that

$$\max_{z \in \mathcal{B}_U} \|z \wedge \eta_M(g)\| \le E|g|^F \eta^{1/F} e^{-\ell_j/F}. \tag{6.6}$$

Since there are only finitely many $\mathbb{Q}$-subgroups are of a given height, passing to a subsequence if necessary, we may assume that $\mathbf{W}_j = \mathbf{W}$ and $\mathbf{M}_j = \mathbf{M}$ for all j, hence from (6.6) it follows that $z \wedge \eta_M(g) = 0$ for all $z \in \mathcal{B}_U$, hence $Ug \subset g\mathbf{M}(\mathbb{Q}_\Sigma)$. By Lemma 6.5, $\mathbf{M}$ is contained in some nontrivial $\mathbb{Q}$-parabolic subgroup $\mathbf{P} < \mathbf{G}$ (indeed, with $\mathrm{ht}(\mathbf{P}) \ll |g|^\star$). □

7 Proof of Theorem 3.2

For every Γ-invariant subcollection $\mathcal{F} \subset \mathcal{H}$, $t \in \mathbb{R}^+$, and $\varepsilon : \mathbb{R}^+ \to (0,1)$, let $\mathrm{Exc}(\varepsilon, t, \mathcal{F})$ be the set

$$\mathrm{Exc}(\varepsilon, t, \mathcal{F}) = \{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \text{ is not } (\varepsilon, t, \mathcal{F})\text{-Diophantine}\}.$$

For every $\epsilon \in (0,1)$, we let $\underline{\epsilon}$ denote the constant function $\underline{\epsilon} : s \mapsto \epsilon$. For every $1 \le r \le \dim \mathbf{G}$, we let $\mathcal{F}_r$ denote the collection of class-$\mathcal{H}$ subgroups of $\mathbf{G}$ of dimension $\le r$. For notational simplicity, let $\mathcal{F}_0 = \emptyset$.

The bulk of the proof of Theorem 3.2 is the following estimate:

7.1 Lemma. *There are A_{16}, $A_{17} > 2$, and $D \ge 1$ depending on N, and $0 < c_0 \le 1$ depending on N, $\#\Sigma$, and polynomially on the primes in Σ so that the following holds. For $1 \le r \le \dim \mathbf{G}$ and $\eta, \beta, \tau \in (0,1)$, $n \in \mathbb{R}^+$ with*

$$\eta^{1/A_{17}} \le c_0 \cdot (\tau\beta e^{-n}\tilde{E}_{\mathbf{G}}^{-1})^{A_{16}} \tag{7.1}$$

at least one of the following holds:

(1) $|(\mathrm{Exc}(\underline{\eta}, n, \mathcal{F}_r) \setminus \mathrm{Exc}(\underline{\eta}', n', \mathcal{F}_{r-1})) \cap \{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \in X_\tau\}| \ll \beta^{1/D}$ *for*

$$n' = A_{16}(n + \log(1/\tau) + \log(1/\beta) + \log \tilde{E}_{\mathbf{G}}), \quad \eta' = e^{n'}\eta^{1/A_{17}},$$

or

(2) *for some* $\mathbf{H}$ *of dimension* r*:*
   (a) *For all* $u \in \mathsf{B}_U(e)$ *we have*

$$\mathsf{c}(\eta_H(\lambda_k(u)g)) \le \beta^{-1}e^n.$$

   (b) *For all* $u \in \mathsf{B}_U(e)$ *we have*

$$\max_{z \in \mathcal{B}_U} \|z \wedge \eta_H(\lambda_k(u)g)\| \le \eta^{1/2}.$$

Proof. Recall from §2.9 the definition

$$\mathsf{B}_{\mathfrak{u}}(0, \delta) = \left\{ \sum_{z \in \mathcal{B}_U} r_z z : |r_z|_\Sigma \le \delta \right\}$$

so that $\mathsf{B}_U(e) := \exp(\mathsf{B}_{\mathfrak{u}}(0, 1))$.

We will cover the set

$$\log(\mathrm{Exc}(\underline{\eta}, n, \mathcal{F}_r)) \subset \mathsf{B} = \mathsf{B}_{\mathfrak{u}}(0, 1)$$

by a collection of balls $\mathcal{E} = \{\mathsf{B}_i = \mathsf{B}_{\mathfrak{u}}(0, \rho_i) + u_i\}_{i \in I}$ and for each such ball attach a class-$\mathcal{H}$ group $\mathbf{H}_i \in \mathcal{F}_r$ so that:

($\mathcal{E}1$) $\sum_{i \in I} |\mathsf{B}_i| \ll 1$.

($\mathcal{E}2$) For every $u \in \mathsf{B}_i$ and $u = \exp(u)$

$$\mathsf{c}(\eta_{H_i}(\lambda_k(u)g)) \le \beta^{-1}e^n, \tag{7.2}$$

$$\max_{\mathcal{B}_U} \|z \wedge \eta_{H_i}(\lambda_k(u)g)\| \le \eta^{1/2}. \tag{7.3}$$

($\mathcal{E}3$) For every $i \in I$, for some $u \in \mathsf{B}_i$, equality holds in at least one of (7.2) or (7.3).

More precisely, we will try to construct a cover $\mathcal{E}$ with these properties, and if we fail this will establish that part (2) of Lemma 7.1 holds.

Assuming we succeed, we will show that these properties imply, for a suitable choice of constants A_{16}, A_{17}, κ, that for n', η' as above

$$\begin{aligned}&(\mathrm{Exc}(\underline{\eta}, n, \mathcal{F}_r) \setminus \mathrm{Exc}(\underline{\eta}', n', \mathcal{F}_{r-1})) \cap \{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \in X_\tau\} \\ &\qquad \subset \bigcup_{i \in I} \{u \in \exp(\mathsf{B}_i) : \lambda_k(u)g \text{ satisfies } (7.2)' \text{ and } (7.3)'\}\end{aligned} \tag{7.4}$$

where (7.2)′ and (7.3)′ denote inequalities (7.2) and (7.3) but with e^n and η instead of $\beta^{-1}e^n$ and $\eta^{1/2}$, respectively. Once (7.4) is established, we get from Lemma 5.5 that

$$|(\mathrm{Exc}(\underline{\eta}, n, \mathcal{F}_r) \setminus \mathrm{Exc}(\underline{\eta}', n', \mathcal{F}_{r-1})) \cap \{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \in X_\tau\}| \ll \max(\eta^{1/2}, \beta)^{1/D}.$$

The desired estimate in part (1) follows from this bound and (7.1).

The construction of the open cover is straightforward. For every $u \in \mathsf{B}_{\mathfrak{u}}(0, 1)$ for which $u = \exp(u) \in \mathrm{Exc}(\underline{\eta}, n, \mathcal{F}_r)$, there is (by definition) a $\mathbb{Q}$-group $\mathbf{H}_u \in \mathcal{H}$ of dimension $\leq r$ so that (7.2)′ and (7.3)′ hold (with $\mathbf{H}_u$ replacing the yet undefined $\mathbf{H}_i$).

For each such u, let $\mathsf{B}(u)$ denote the set $\mathsf{B}(u) = \mathsf{B}_{\mathfrak{u}}(0, \rho_u) + u$ with ρ_u chosen to be as small as possible so that for some $v \in \mathsf{B}(u)$ either $\mathsf{c}(\eta_{H_u}(\lambda_k(\exp v)g)) = \beta^{-1}e^n$ or $\max_{\mathcal{B}_U} \|z \wedge \eta_{H_u}(\lambda_k(\exp v)g)\| = \eta^{1/2}$. Unless the estimates (2)(a) and (2)(b) of the statement of this theorem hold for some r-dimensional $\mathbf{H} \in \mathcal{H}$, for any $u \in \mathsf{B}_{\mathfrak{u}}(0, 1)$ it holds that $\mathsf{B}(u) \subset \mathsf{B}_{\mathfrak{u}}(0, 3)$. Note that the estimate of (2)(a) together with (2.3) gives that $\mathrm{ht}(\mathbf{H}) \ll |g|^\star e^{2n}$.

Assuming there is no such $\mathbf{H}$, the Vitali covering argument allows us to find a subcollection $\mathcal{E} = \{\mathsf{B}_i = \mathsf{B}_{\mathfrak{u}}(0, \rho_i) + u_i\}_{i\in I}$ of $\{\mathsf{B}(u)\}$ so that the collection of smaller balls $\{\mathsf{B}_{\mathfrak{u}}(0, \rho_i/3) + u_i\}_{i\in I}$ is a disjoint collection of subsets of $\mathsf{B}_{\mathfrak{u}}(0, 3)$ but

$$\bigcup_{i\in I} \mathsf{B}_i \supset \log(\mathrm{Exc}(\underline{\eta}, n, \mathcal{F}_r)). \tag{7.5}$$

The resulting collection $\mathcal{E}$ clearly satisfies ($\mathcal{E}$1)–($\mathcal{E}$3).

It remains to establish (7.4). Fix some $u \in \mathrm{Exc}(\underline{\eta}, n, \mathcal{F}_r) \cap X_\tau$. Then by (7.5) there is an $i \in I$ so that $u \in \exp(\mathsf{B}_i)$; put $\mathbf{H}_i = \mathbf{H}_{u_i}$. By the definition of B_i estimates (7.2) and (7.3) hold, while by the definition of $\mathrm{Exc}(\underline{\eta}, n, \mathcal{F}_r)$ there is an $\mathbf{H} \in \mathcal{H}$ of dimension $\leq r$ so that $\mathsf{c}(\eta_H(\lambda_k(u)g)) \leq e^n$ and $\max_{\mathcal{B}_U} \|z \wedge \eta_H(\lambda_k(u)g)\| \leq \eta$. There are now two possibilities: either $\mathbf{H} = \mathbf{H}_i$, in which case u is contained in the set on the right-hand side of (7.4), or $\mathbf{H} \neq \mathbf{H}_i$.

Thus suppose that $\mathbf{H} \neq \mathbf{H}_i$. By Lemma 2.8 there is a $\gamma \in \Gamma$ so that

$$|\lambda_k(u)g\gamma| \leq E_{\mathbf{G}}\tau^{-F}.$$

Since $\eta_{\gamma^{-1}\bullet\gamma}(\lambda_k(u)g\gamma) = \eta_\bullet(\lambda_k(u)g)$, we have

$$\left.\begin{aligned} \mathsf{c}(\eta_L(\lambda_k(u)g\gamma)) &\leq \beta^{-1}e^n \\ \max_{\mathcal{B}_U} \|z \wedge \eta_L(\lambda_k(u)g\gamma)\| &\leq \eta^{1/2} \end{aligned}\right. \quad \text{for } \mathbf{L} = \gamma^{-1}\mathbf{H}\gamma, \gamma^{-1}\mathbf{H}_i\gamma.$$

Since $\mathbf{H} \neq \mathbf{H}_i$, we have that $\tilde{\mathbf{H}} = (\gamma^{-1}\mathbf{H}\gamma \cap \gamma^{-1}\mathbf{H}_i\gamma)^{\mathcal{H}}$ is of dimension $\leq r - 1$. Applying Lemma 4.12 with $\epsilon = \eta^{1/2}$ and $r = \beta^{-1}e^n$ we get that if $\eta \ll E_{\mathbf{G}}^\star\beta^{-\star}\tau^{-\star}e^{\star n}$, the group $\tilde{\mathbf{H}}$ is nontrivial and

$$\mathsf{c}(\eta_{\tilde{H}}(\lambda_k(u)g\gamma)) \ll (E_{\mathbf{G}}\beta^{-1}\tau^{-1}e^n)^{A_{16}},$$
$$\max_{\mathcal{B}_U} \|z \wedge \eta_{\tilde{H}}(\lambda_k(u)g\gamma)\| \ll (E_{\mathbf{G}}\beta^{-1}\tau^{-1}e^n)^{A_{16}}\eta^{1/A_{17}}.$$

Therefore (in view of our convention regarding implicit constants, and recalling that $\tilde{E}_{\mathbf{G}} = E_{\mathbf{G}}\,\mathrm{ht}(\mathbf{G})$) we have that $u \in \mathrm{Exc}(\underline{\eta'}, n', \mathcal{F}_{r-1})$ for

$$n' = A_{16}(n + \log(1/\beta) + \log(1/\tau) + \log \tilde{E}_{\mathbf{G}})$$

and $\eta' = e^{n'}\eta^{1/A_{17}}$. □

Proof of Theorem 3.2. We need to estimate the size of the set

$$\mathsf{Bad} = \{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \not\in X_\eta \text{ or is not } (\varepsilon, t)\text{-Diophantine}\}.$$

Set

$$\mathsf{Bad}_\eta = \{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \not\in X_\eta\},$$
$$\mathsf{Bad}_{\mathrm{Exc}} = \{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \text{ is not } (\varepsilon, t)\text{-Diophantine but is in } X_\eta\}.$$

Clearly $\mathsf{Bad} = \mathsf{Bad}_\eta \cup \mathsf{Bad}_{\mathrm{Exc}}$.

We can bound the size of Bad_η using Theorem 6.3, obtaining that

$$|\mathsf{Bad}_\eta| \le E\eta^{1/F} \tag{7.6}$$

unless there exists a group $\mathbf{M} \neq \mathbf{G}$,

$$\mathrm{ht}(\mathbf{M}) \le E|g|^F\eta^{1/F},$$

so that for all $u \in \mathsf{B}_U(e)$

$$\mathsf{c}(\eta_M(\lambda_k(u)g)) \le E|g|^F\eta^{1/F},$$
$$\max_{z\in\mathcal{B}_U} \|z \wedge \eta_M(\lambda_{k-1}(u)g)\| \le E|g|^F\eta^{1/F}e^{-k/F}.$$

This clearly implies that (2) of Theorem 3.2 holds (if we choose A large enough).

Assume therefore for the remainder of the proof that (7.6) holds. Let $d = \dim \mathbf{G}$. It follows from Definition 3.1 that in the notations of Lemma 7.1

$$\mathsf{Bad}_{\mathrm{Exc}} \subset \left(\bigcup_{\ell=1}^{\lceil t\rceil} \mathrm{Exc}(\underline{\varepsilon(e^\ell)}, \ell+1, \mathcal{F}_{d-1}) \cup \mathrm{Exc}(\underline{\varepsilon(0)}, 1, \mathcal{F}_{d-1})\right) \cap (\mathsf{B}_U(e) \setminus \mathsf{Bad}_\eta). \tag{7.7}$$

Fix ℓ. To estimate $|\,\mathrm{Exc}(\underline{\varepsilon(e^\ell)}, \ell + 1, \mathcal{F}_{d-1}) \cap (\mathsf{B}_U(e) \setminus \mathsf{Bad}_\eta)|$, define iteratively, starting with $r = d - 1$, $n_{d-1} = \ell + 1$ and $\eta_{d-1} = \varepsilon(e^\ell)$. Proceed by induction to define

$$n_{r-1} = A_{16}(n_r + \ell + 2\log(1/\eta) + \log \tilde{E}_{\mathbf{G}}), \quad \eta_{r-1} = e^{n_{r-1}}\eta_r^{1/A_{17}}.$$

Then

$$
\begin{aligned}
&|\operatorname{Exc}(\underline{\varepsilon(e^{\ell})}, \ell+1, \mathcal{F}_{\dim \mathbf{G}}) \cap (\mathsf{B}_U(e) \setminus \mathsf{Bad}_\eta)| \\
&\qquad \leq \sum_{r=1}^{d-1} |(\operatorname{Exc}(\underline{\eta_r}, n_r, \mathcal{F}_r) \setminus (\operatorname{Exc}(\underline{\eta_{r-1}}, n_{r-1}, \mathcal{F}_{r-1})) \cap (\mathsf{B}_U(e) \setminus \mathsf{Bad}_\eta)|.
\end{aligned} \tag{7.8}
$$

We want to apply Lemma 7.1 (with $\tau = \eta$ and $\beta = \eta e^{-\ell}$ where η is as in this theorem). In order to apply Lemma 7.1, our choice of parameters needs to satisfy condition (7.1), with the critical case being that of $r = 1$. In this case (7.1) becomes

$$
\eta_1^{1/A_{17}} \leq c_0 \cdot (\eta^2 e^{-\ell} e^{-n_1} \tilde{E}_{\mathbf{G}}^{-1})^{A_{16}}. \tag{7.9}
$$

Iteratively working through the constants, there are $A_{18} > A_{16}$, $A_{19} > A_{17}$, easily explicated in terms of d and A_{16} and A_{17} so that

$$
e^{n_1} \leq \left(\frac{2\tilde{E}_{\mathbf{G}} e^{\ell}}{\eta^2 e^{-\ell}}\right)^{A_{18}}, \quad \eta_1 \leq e^{2n_1} \varepsilon(e^{\ell})^{1/A_{19}}.
$$

Then assuming (3.3) with A large enough and suitable choice of constant E_1 we can ensure that $\eta_1 \leq \varepsilon(e^{\ell})^{1/2A_{19}}$ and that (7.9) holds.

By Lemma 7.1 and (3.3) (for A large enough), for every r

$$
\begin{aligned}
&|(\operatorname{Exc}(\underline{\eta_r}, n_r, \mathcal{F}_r) \setminus \operatorname{Exc}(\underline{\eta_{r-1}}, n_{r-1}, \mathcal{F}_{r-1})) \cap (\mathsf{B}_U(e) \setminus \mathsf{Bad}_\eta)| \\
&\qquad\qquad \ll \eta^{1/D} e^{-\ell/D}
\end{aligned} \tag{7.10}
$$

unless for some $\mathbf{H}$ of dimension r, for all $u \in \mathsf{B}_U(e)$ and some constant F depending only on N and E_1' depending polynomially on $\tilde{E}_{\mathbf{G}}$ and on N,

$$
\mathsf{c}(\eta_H(\lambda_k(u)g)) \leq \eta^{-1} e^{\ell+n_r} \leq E_1' e^{F\ell} \eta^{-F}, \tag{7.11}
$$

$$
\max_{z \in \mathcal{B}_U} \|z \wedge \eta_H(\lambda_k(u)g)\| \leq \eta_r^{1/2} \leq E_1' e^{F\ell} \eta^{-F} \varepsilon(e^{\ell})^{1/F}. \tag{7.12}
$$

If equations (7.11) and (7.12) hold for all $u \in \mathsf{B}_U(e)$, there are two cases. Firstly, it may happen that $\mathbf{H} \lhd \mathbf{G}$ in which case

$$
\operatorname{ht}(\mathbf{H}) = \mathsf{c}(\eta_H(\lambda_k(u)g)) \leq E_1' e^{F\ell} \eta^{-F}.
$$

Then as we assumed $\varepsilon(s) \leq \eta^A s^{-A}/E_1$ (for the constants A and E_1 of the theorem we are proving, which are yet to be fixed) if A was chosen large enough, by (7.12)

$$
\max_{z \in \mathcal{B}_U} \|z \wedge \eta_H(\lambda_k(u)g)\| \leq \varepsilon(e^{\ell})^{1/2F}. \tag{7.13}
$$

For a given $\mathbf{H}$, the value of ℓ has to be large enough so that (7.11) holds, namely

$$
e^{\ell} \geq (\operatorname{ht}(\mathbf{H}) \eta^F / E_1')^{1/F},
$$

so

$$\max_{z\in\mathcal{B}_U} \|z \wedge \eta_H(\lambda_k(u)g)\| \le \varepsilon(\mathrm{ht}(\mathbf{H})^{1/F}\eta/E_1')^{1/2F}$$

and (3) of the statement of Theorem 3.2 is satisfied.

If equations (7.11) and (7.12) hold for all $u \in \mathsf{B}_U(e)$, but $\mathbf{H}$ is not a normal subgroup of $\mathbf{G}$ we apply Proposition 5.8 and conclude that $\mathbf{M} = \mathbf{M_H}$ satisfies for all $u \in \mathsf{B}_U(e)$

$$\begin{aligned} &\mathsf{c}(\eta_M(\lambda_k(u)g)) \ll E_{\mathbf{G}}{}^\star e^{\star\ell}\eta^{-\star} \\ &\max_{z\in\mathcal{B}_U} \|z \wedge \eta_M(\lambda_k(u)g)\| \ll E_{\mathbf{G}}{}^\star e^{\star\ell}\eta^{-\star}e^{-k/D'}. \end{aligned} \tag{7.14}$$

(The dependence of the upper bounds in (7.14) on $|g|$ can be eliminated as in the proof of Proposition 5.8 by using the fact that by (7.6) for most $u \in \mathsf{B}_U(e)$ there is a $\gamma_u \in \Gamma$ so that $|\lambda_k(u)g\gamma_u| \ll \eta^{-\star}$.) In this case, (2) of Theorem 3.2 holds.

The only remaining case is if (7.10) holds for every r and ℓ (as well as the analogous estimates for $\mathrm{Exc}(\underline{\varepsilon(0)}, 1, \mathcal{F}_{d-1})$, for which we omit the details, but is handled similarly), in which case it follows from (7.7) and (7.8) that $|\mathsf{Bad}_{\mathrm{Exc}}| \ll \eta^{1/D}$, establishing (1) of Theorem 3.2. □

Let us record the following corollary of the proofs of Lemma 7.1 and Theorem 3.2.

7.2 Corollary. *Let the notation be as in Theorem 3.2. In particular, A, D and E_1 are as in loc. cit. Let $g \in G$, $t > 0$, $k \geq 1$, and $0 < \eta < 1/2$. Assume*

$$0 < \varepsilon \le \eta^A e^{-tA}/E_1.$$

Suppose there exists

$$\mathrm{Exc}_r \subset \{u \in \mathsf{B}_U(e) : \lambda_k(u)g\Gamma \in X_\eta\}$$

with $|\,\mathrm{Exc}_r\,| > E_1\eta^{1/D}$ so that and for every $u \in \mathrm{Exc}_r$ there is a $\mathbb{Q}$-group $\mathbf{H}_u \in \mathcal{H}$ of dimension $\le r$ satisfying both of the following

$$\begin{aligned} &\mathsf{c}(\eta_{H_u}(\lambda_k(u)g)) \le e^t \\ &\max_{\mathcal{B}_U} \|z \wedge \eta_{H_u}(\lambda_k(u)g)\| \le \varepsilon. \end{aligned}$$

Then Theorem 3.2(2) holds with a subgroup $\mathbf{H} \in \mathcal{H}$ which is also contained in $\gamma\mathbf{H}_u\gamma^{-1}$ for some $u \in \mathrm{Exc}_r$ and some $\gamma \in \Gamma$.

8 Some corollaries of Theorem 3.2

In this section we discuss some of the consequences of Theorem 3.2. Recall that for any $T > 0$, we put

$$(8.1)\quad \sigma(T) = \min(\{1\} \cup \{\|z \wedge v_H\| : \mathbf{H} \in \mathcal{H}, \mathbf{H} \lhd \mathbf{G}, \mathrm{ht}(\mathbf{H}) \le T, \{1\} \ne \mathbf{H} \ne \mathbf{G}\}).$$

8.1 Theorem. *Let A, D, and E_1 be as in Theorem 3.2. There exists some ϑ depending only on N so that the following holds. Let $0 < \eta < 1/2$. Let $\{x_m\}$ be a sequence in X and let $k_m \to \infty$ be a sequence of natural numbers. For each m let $\mathsf{V}_m \subset \mathsf{B}_U(e)$ be a measurable set with measure $> \vartheta E_1 \eta^{1/D}$. Let*

$$Y = \bigcap_{\ell \ge 1} \overline{\bigcup_{m \ge \ell} \{\lambda_{k_m}(u)x_m : u \in \mathsf{V}_m\}}.$$

Then exactly one of the following holds.

(1) *Y contains an ε-Diophantine point for*

$$\varepsilon(s) = (\eta s^{-1} \sigma(E_1^A \eta^{-A} s^A)/2E_1)^A.$$

(2) *There exists*

(a) *a finite collection $\{(\mathbf{H}_i, L_i) : 1 \le i \le \ell\} \subset \mathcal{H} \times \mathbb{R}^+$, and*

(b) *a countable (possibly finite) collection*

$$\mathcal{W} = \{(\mathbf{W}_j, R_j, r_j) : j \in J\} \subset \mathcal{H} \times \mathbb{R}^+ \times \mathbb{R}^+$$

where $\mathbf{W}_j$ is a non-normal unipotent subgroup for all $j \in J$, and $r_j \to 0$ so that if we put

$$Y_i = \{g \in N_G(U, H_i) : \mathsf{c}(\eta_{H_i}(g)) \le L_i\}\Gamma/\Gamma$$

and

$$Z_j = \{g \in N_G(U, M_{W_j}) : \mathsf{c}(\eta_{M_{W_j}}(g)) \le R_j \;\&\; \mathsf{c}(\eta_{W_j}(g)) \le r_j\}\Gamma/\Gamma,$$

then

$$Y \subset \Big(\bigcup_{i=1}^{\ell} Y_i\Big) \bigcup \Big(\bigcup_{j \in J} Z_j\Big).$$

This theorem implies Theorem 1.5 since each Z_j is contained in

$$\{g \in N_G(U, M_{W_j}) : \mathsf{c}(\eta_{M_{W_j}}(g)) \le R_j\}\,\Gamma/\Gamma,$$

and as $r_j \to 0$ for any β only finitely many of the Z_j can interset X_β. Recall that for every i the sets Y_i above are closed subsets of X (see Corollary 4.10.1), and the same proof gives that the sets Z_j are closed as well.

We first prove a special case of Theorem 8.1.

8.2 Lemma. *Let A, D, and E_1 be as in Theorem 3.2. There exists some ϑ depending only on N so that the following holds. Let $0 < \eta < 1/2$ and let $t \in \mathbb{R}^+$. Let $\{x_m\}$ be a sequence in X and let $k_m \to \infty$ be a sequence of natural numbers. For each m let $\mathsf{V}_m \subset \mathsf{B}_U(e)$ be a measurable set with measure $> \vartheta E_1 \eta^{1/D}$. Let*

$$Y = \bigcap_{\ell \geq 1} \overline{\bigcup_{m \geq \ell} \{\lambda_{k_m}(u) x_m : u \in \mathsf{V}_m\}}.$$

Then at least one of the following holds.

(1) *$Y \cap X_\eta$ contains an (ε, t)-Diophantine point for*

$$\varepsilon(s) = (\eta s^{-1} \sigma(E_1^A \eta^{-A} s^A)/2E_1)^A.$$

(2) *There exists*
   (a) *a finite collection $\{(\mathbf{H}_i, L_i) : 1 \leq i \leq \ell\} \subset \mathcal{H} \times \mathbb{R}^+$, and*
   (b) *a countable (possibly finite) collection*

$$\mathcal{W} = \{(\mathbf{W}_j, R_j, r_j) : j \in J\} \subset \mathcal{H} \times \mathbb{R}^+ \times \mathbb{R}^+$$

where $\mathbf{W}_j$ is a non-normal unipotent subgroup for all $j \in J$, and $r_j \to 0$ so that if we put

$$Y_i = \{g \in N_G(U, H_i) : \mathsf{c}(\eta_{H_i}(g)) \leq L_i\}\Gamma/\Gamma$$

and

$$Z_j = \{g \in N_G(U, M_{W_j}) : \mathsf{c}(\eta_{M_{W_j}}(g)) \leq R_j \ \&\ \mathsf{c}(\eta_{W_j}(g)) \leq r_j\}\Gamma/\Gamma,$$

then

$$Y \subset \Big(\bigcup_{i=1}^{\ell} Y_i\Big) \bigcup \Big(\bigcup_{j \in J} Z_j\Big).$$

We need the following lemma.

8.3 Lemma. *Let E and F be as in Theorem 6.3. Let the notation be as in Theorem 8.1. Then one of the following holds:*

(1) *There exists some $\beta_0 > 0$ and subsequence $m_i \to \infty$ so that*

$$|\{u \in \mathsf{B}_U(e) : \lambda_{k_{m_i}-1}(u) x_{m_i} \notin X_\beta\}| \leq E\beta^{1/F},$$

for all $\beta \leq \beta_0$, or

(2) *$Y = \emptyset$.*

Proof. In view of Theorem 6.3 it suffices to show that there exists some β' and a subsequence $\{m_i\}$ so that

$$\{u \in \mathsf{B}_U(e) : \lambda_{k_{m_i}-1}(u)x_{m_i}\} \cap X_{\beta'} \neq \emptyset.$$

Indeed if this is established, then Theorem 6.3(2) cannot hold for any $\eta = \beta \leq \beta'^F/2E$, x_{m_i} and $k_{m_i}-1$, hence Theorem 6.3(1) holds which implies part (1) in this lemma with $\beta_0 = \beta'^F/2E$.

Assume contrary to the above claim that for every β there exists some m_β so that for all $m \geq m_\beta$

$$\{u \in \mathsf{B}_U(e) : \lambda_{k_m-1}(u)x_m\} \cap X_\beta = \emptyset.$$

Then by Theorem 6.3 applied with $\eta = \beta$ and the point $x_m = h_m\Gamma$, we thus get that there exists a unipotent $\mathbb{Q}$-subgroup $\mathbf{W}$ which is not normal in $\mathbf{G}$ so that

$$\mathsf{c}(\eta_W(\lambda_{k_m-1}(u)h_m)) \leq E\beta^{1/F} \quad \text{for all } u \in \mathsf{B}_U(e);$$

see (6.3), also Lemma 6.2. This and Lemma 5.5 imply that

$$\mathsf{c}(\eta_W(\lambda_{k_m}(u)h_m)) \leq E'\beta^{1/F} \quad \text{for all } u \in \mathsf{B}_U(e). \tag{8.2}$$

Therefore, we get that

$$\{\lambda_{k_m}(u)x_m : u \in \mathsf{B}_U(e)\} \cap X_{E'\beta^{1/F}} = \emptyset \quad \text{for all } m > m_\beta. \tag{8.3}$$

Hence the claim in part (2) holds. □

Proof of Lemma 8.2. The proof is based on applying Theorem 3.2 to the pieces of the orbits

$$\{\lambda_{k_m}(u)x_m : u \in \mathsf{V}_m\}.$$

We show that Theorem 3.2(3) cannot hold for the choice of ε we made in the lemma. Further, we show that if there are infinitely many m so that Theorem 3.2(1) holds, then part (1) in the lemma holds. In consequence, we are reduced to the case that for all but finitely many m, Theorem 3.2(2) holds. In this case we use Lemma 4.10 to conclude that part (2) above holds.

We begin by replacing x_m with a possibly different point in the orbit which is chosen to have a representative of a controlled size.

Assuming $Y \neq \emptyset$ and repeatedly applying Lemma 8.3, we may find natural numbers $\{n_i : i \in I\}$ with $|\log\beta_0| < n_1 < n_2 < \cdots$ so that if we put

$$J_i := \{m \in \mathbb{N} \setminus J_{i-1} : |\{u \in \mathsf{B}_U(e) : \lambda_{k_m-1}(u)x_m \notin X_{2^{-n_i}}| \leq E2^{-n_i/F}\},$$

for all $i \in I$ and $J_0 = \emptyset$, then the following hold:

- for all $i \in I$, J_i is an infinite set, and
- for every $x \in Y$, there exists an $i \in I$, a sequence $m_p \to \infty$ in J_i, and for any m_p there is some $u_{m_p} \in \mathsf{B}_U(e)$ so that

$$\lambda_{k_{m_p}}(u_{m_p})x_{m_p} \to x.$$

We remark that by Lemma 8.3 if $Y \neq \emptyset$, then $I \neq \emptyset$, but it may well be finite: for instance, if $\{x_m\}$ is a bounded sequence, then we may choose n_1 large enough so that $I = \{1\}$.

Recall the constants $E_{\mathbf{G}}$ and F from Lemma 2.8. For every $i \in I$ and $m \in J_i$ fix some $g_m \in \mathbf{G}$ so that

$$|g_m| \le E_{\mathbf{G}} 2^{n_i F} =: T_i \tag{8.4}$$

and $g_m\Gamma = \lambda_{k_m-1}(u_m)x_m \in X_{2^{-n_i}}$ for some $u_m \in \mathsf{B}_U(e)$.

Recall also from (2.11) that for some $\kappa > 0$ we have

$$\lambda_{k_m-\kappa}(\mathsf{B}_U(e))\lambda_{k_m-1}(u_m) \subset \lambda_{k_m}(\mathsf{B}_U(e)). \tag{8.5}$$

Let

$$\vartheta = 2|\mathsf{B}_U(e)|/|\lambda_{-\kappa}(\mathsf{B}_U(e))|.$$

Apply Theorem 3.2 with g_m, $k_m - \kappa$, η, t, and

$$\varepsilon(s) = (\eta s^{-1}\sigma(E_1^A\eta^{-A}s^A)/2E_1)^A; \tag{8.6}$$

note that ε satisfies the condition in (3.3).

We first argue that Theorem 3.2(3) cannot hold. Indeed, assume contrary to this claim that there exists some $\mathbf{H} \lhd \mathbf{G}$ satisfying Theorem 3.2(3). That is: $\mathrm{ht}(\mathbf{H}) \le E_1(e^t\eta^{-1})^A$ and

$$\max_{z\in\mathcal{B}_U} \|z \wedge v_H\| \le \varepsilon(\mathrm{ht}(\mathbf{H})^{1/A}\eta/E_1)^{1/A}.$$

In view of (8.6) we thus get that

$$\max_{z\in\mathcal{B}_U} \|z \wedge v_H\| \le \mathrm{ht}(\mathbf{H})^{-1/A}\sigma(\mathrm{ht}(\mathbf{H}))/2 < \sigma(\mathrm{ht}(\mathbf{H})).$$

However, this contradicts the definition of σ, see (8.1).

Assume now that the conclusion in Theorem 3.2(1) holds for a subsequence $m_i \to \infty$. Then, since

$$\lambda_{k_{m_i}-\kappa}(\mathsf{B}_U(e))\lambda_{k_{m_i}-1}(u_{m_i}) \subset \lambda_{k_{m_i}}(\mathsf{B}_U(e)),$$

$|\lambda_{k_{m_i}-\kappa}(\mathsf{B}_U(e))|/|\lambda_{k_{m_i}}(\mathsf{B}_U(e))| \ge 2/\vartheta$, and $|\mathsf{V}_{m_i}| > \vartheta E_1\eta^{1/D}$, we have that

$$\{\lambda_{k_{m_i}}(u)x_{m_i} : u \in \mathsf{V}_m\} \cap \{x \in X_\eta : x \text{ is } (\varepsilon, t)\text{-Diophantine}\} \neq \emptyset$$

for all m_i. Hence $Y \cap \{x \in X_\eta : x \text{ is } (\varepsilon, t)\text{-Diophantine}\} \neq \emptyset$ and part (1) in the lemma holds.

Altogether, we are reduced to the case that Theorem 3.2(2) holds for all but finitely many m. Dropping the first few terms, which does not affect Y, we assume that Theorem 3.2(2) holds for all m, or more precisely that Theorem 3.2(2) holds for g_m, $k_m - \kappa$, η, and (ε, t). In particular, we have the following: For every $m \in J_1$ there exists a nontrivial proper subgroup $\mathbf{H}_m \in \mathcal{H}$ with

$$\mathrm{ht}(\mathbf{H}_m) \leq (E|g_m|^A + E_1 e^{At})\eta^{-A} \leq (ET_1^A + E_1 e^{At})\eta^{-A} =: L,$$

so that the following hold:

$(\dagger)_1$ For all $u \in \mathsf{B}_U(e)$ we have

$$\mathsf{c}(\eta_{H_m}(\lambda_{k_m-\kappa}(u)g_m)) \leq L.$$

$(\ddagger)_1$ For every $u \in \mathsf{B}_U(e)$ we have

$$\max_{z \in \mathcal{B}_U} \|z \wedge \eta_{H_m}(\lambda_{k_m-\kappa}(u)g_m)\| \leq Le^{-k_m+\kappa/D}.$$

Let $\mathcal{F} = \{(\mathbf{H}, L) : \mathrm{ht}(\mathbf{H}) \leq L\}$. In view of (4.1), $\mathcal{F}$ is a finite family.

Let now $i \in I$ and $i \geq 2$—we note again that it is possible that $I = \{1\}$ and this case is empty.

Arguing as in Lemma 8.3, see in particular (8.3), for all $m \in J_i$ we have

$$\{\lambda_{k_m}(u)x_m : u \in \mathsf{B}_U(e)\} \cap X_{\theta_i} = \emptyset,$$

where $\theta_i = E' 2^{-n_{i-1}/F}$. This, in view of (8.5), implies that

$$\{\lambda_{k_m-\kappa}(u)g_m\Gamma : u \in \mathsf{B}_U(e)\} \cap X_{\theta_i} = \emptyset. \tag{8.7}$$

Therefore, by Theorem 6.3, for every $m \in J_i$ there exists some unipotent subgroup $\mathbf{W}_m$ with

$$\mathrm{ht}(\mathbf{W}_m) \leq ET_i^F \theta_i^{1/F} =: S_i$$

so that

$$\mathsf{c}(\eta_{W_m}(\lambda_{k_m-\kappa}(u)g)) \leq E\theta_i^{1/F} =: s_i \quad \text{for all } u \in \mathsf{B}_U(e). \tag{8.8}$$

Moreover, if we put $\mathbf{M}_m = \mathbf{M}_{\mathbf{W}_m}$, then $\mathbf{M}_m \neq \mathbf{G}$,

$$\mathrm{ht}(\mathbf{M}_m) \leq S_i,$$

and the following hold:

$(\dagger)_i$ For all $u \in \mathsf{B}_U(e)$ we have

$$\mathsf{c}(\eta_M(\lambda_{k_m-\kappa}(u)g)) \leq S_i.$$

$(\ddagger)_i$ For all $u \in \mathsf{B}_U(e)$ we have

$$\max_{z \in \mathcal{B}_U} \|z \wedge \eta_M(\lambda_{k_m-\kappa}(u)g)\| \leq S_i e^{-k_m+\kappa/F}.$$

Let

$$\mathcal{E}_i = \{(\mathbf{W}, S_i, s_i) : \mathrm{ht}(\mathbf{W}) \leq S_i, \mathbf{W} \in \mathcal{H} \text{ is unipotent and not normal}\}.$$

Then $\mathcal{E}_i$ is a finite family for each i.

We now show that the claim in part (2) holds with $\mathcal{F}$ and $\mathcal{E}_i$, $i \geq 2$. Let $x = g\Gamma \in Y$. Then there exists an $i \in I$ and a sequence $m_p \to \infty$ in J_i so that the following holds. For any m_p there is some $u_{m_p} \in \mathsf{V}_{m_p}$ so that

$$\lambda_{k_{m_p}}(u_{m_p})g_{m_p}\Gamma \to g\Gamma.$$

Assume first that $i = 1$. Then passing to a subsequence we may assume that $(\dagger)_1$ and $(\ddagger)_1$ hold with $\mathbf{H}_{m_p} = \mathbf{H}$ for all p. Hence by Lemma 4.10 we have

$$g \in \{g' \in N_G(U, M_W) : \mathsf{c}(\eta_H(g')) \leq L\}\Gamma.$$

Similarly, if $i \geq 2$ we may pass to a subsequence and assume that $\mathbf{W}_{m_p} = \mathbf{W}_m$ for all p. One then argues as in Lemma 4.10 and gets that

$$g \in \{g' \in N_G(U, M_W) : \mathsf{c}(\eta_{M_W}(g')) \leq S_i, \mathsf{c}(\eta_W(g')) \leq s_i\}.$$

The proof is complete. □

Proof of Theorem 8.1. Let $0 < \eta < 1/2$ and define ε as in part (1). Recall from Definition 3.1 that

$$\{x \in X : x \text{ is } \varepsilon\text{-Diophantine}\} = \bigcap_t \{x \in X : x \text{ is } (\varepsilon, t)\text{-Diophantine}\}.$$

Moreover, $\{x \in X_\eta : x \text{ is } (\varepsilon, t)\text{-Diophantine}\}$ is a nested family of compact sets. Therefore, if Lemma 8.2(1) holds for all t, then Theorem 8.1(1) holds. Therefore, we may assume there exists some t so that Lemma 8.2(2) holds. This implies that Theorem 8.1(2) holds and completes the proof. □

We now state and prove an analogue of Theorem 1.1 in the more general Σ-arithmetic setting.

8.4 Theorem. *Let $\alpha > 0$. Let $\{H_i : 1 \le i \le r\} \subset \mathcal{H}$ be a finite subset consisting of proper subgroups, and for each $1 \le i \le r$ let $\mathcal{C}_i \subset N_G(U, H_i)$ be a compact subset. There exists an open neighborhood $\mathcal{O} = \mathcal{O}(\alpha, \{H_i\}, \{\mathcal{C}_i\})$ so that $X \setminus \mathcal{O}$ is compact and disjoint from $\cup_i \mathcal{C}_i\Gamma/\Gamma$ so that the following holds. For every $x \in \mathcal{G}(U)$ there exists some $k_0 = k_0(\alpha, \{H_i\}, \{\mathcal{C}_i\}, x)$ so that for all $k \ge k_0$ we have*

$$|\{u \in \mathsf{B}_U(e) : \lambda_k(u)x \in \mathcal{O}\}| < \alpha$$

Proof. Let $\eta = (\alpha/E_1)^D$ where D and E_1 are as in Theorem 3.2.

Let $x \in \mathcal{G}(U)$ and let $g \in G$ be so that $x = g\Gamma$. Define

$$\varepsilon(s) = (\eta s^{-1}\sigma(E_1^A\eta^{-A}s^A)/2E_1)^A \tag{8.9}$$

where $\sigma(T)$ is defined as in (8.1).

Let $t \in \mathbb{R}^+$ be so that $\mathrm{ht}(\mathbf{H}_i) \le e^t$ and $\mathsf{c}(\eta_{H_i}(h)) < e^t$ for all $1 \le i \le r$ and all $h \in \mathcal{C}_i$. We will show that the theorem holds with

$$\mathcal{O} = \{x \in X : x \notin X_\eta \text{ or } x \text{ is not } (\varepsilon, t)\text{-Diophantine}\}.$$

First note that for any i and any $h \in \mathcal{C}_i$ we have $\mathsf{c}(\eta_{H_i}(h)) < e^t$ and $z \wedge \eta_{H_i}(h) = 0$ for all $z \in \mathcal{B}_U$. Therefore,

$$\cup_i \mathcal{C}_i\Gamma/\Gamma \subset \mathcal{O}.$$

We claim there exists some k_0 so that for all $k \ge k_0$, Theorem 3.2(1) holds for g, k, and (ε, t). First note that this claim in view of the assertion in Theorem 3.2(1) implies that

$$|\{u \in \mathsf{B} : \lambda_k(u)x \in \mathcal{O}\}| \le E_1\eta^{1/D} = \alpha$$

and the theorem follows.

Let us now prove the claim. Assume contrary to the claim that Theorem 3.2(2) or (3) holds for g, a sequence $k_n \to \infty$, and (ε, t). We first show that Theorem 3.2(3) cannot hold. Indeed, if Theorem 3.2(3) holds, then there is some $\mathbf{H} \lhd \mathbf{G}$ with $\mathrm{ht}(\mathbf{H}) \le E_1(e^t\eta^{-1})^A$ so that

$$\max_{z\in\mathcal{B}_U} \|z \wedge v_H\| \le E_1\eta^{-A}\varepsilon(\mathrm{ht}(\mathbf{H})^{1/A}\eta^A/E_1)^{1/A}.$$

In view of (8.9) we thus get that

$$\max_{z\in\mathcal{B}_U} \|z \wedge v_H\| \le \varepsilon(\mathrm{ht}(\mathbf{H})^{1/A})^{1/A}\sigma(\mathrm{ht}(\mathbf{H}))/2 < \sigma(\mathrm{ht}(\mathbf{H})).$$

This contradicts the definition of σ in (8.1).

Hence we may reduce to the case that Theorem 3.2(2) holds for g, $k_n \to \infty$, and (ε, t). Let

$$L := (E|g|^A + E_1 e^{At})\eta^{-A}.$$

Then in view of Theorem 3.2(2), for every n there exists a proper subgroup $\mathbf{H}_n \in \mathcal{H}$ with $\mathrm{ht}(\mathbf{H}_n) \le L$ so that for all $u \in \mathsf{B}_U(e)$ we have

$$\max_{z \in \mathcal{B}_U} \|z \wedge \eta_{H_n}(\lambda_{k_n}(u)g)\| \le e^{-k_n/D}L.$$

Since there are only finitely many subgroups $\mathbf{H} \in \mathcal{H}$ with $\mathrm{ht}(\mathbf{H}) \le L$, see (4.1), passing to a subsequence we assume $\mathbf{H}_n = \mathbf{H}$ for all n. Hence

$$\max_{z \in \mathcal{B}_U} \|z \wedge \eta_H(\lambda_{k_n}(u)g)\| \le e^{-k_n/D}L.$$

Applying this with $u = e$ and passing to the limit we get that

$$z \wedge \eta_H(g) = 0 \quad \text{for all } z \in \mathcal{B}_U.$$

This contradicts the fact that $g\Gamma \in \mathcal{G}(U)$ and completes the proof. □

9 Friendly measures

In this section we discuss generalizations of our main theorems to the class of friendly measures which were studied in [28], see §1.6 for the definition.

Let the notation be as in §1.2; in particular,

$$U = \{u(t) = \exp(tz) : t \in \mathbb{R}\}$$

for some nilpotent element $z \in \mathfrak{g}$ with $\|z\| = 1$.

In [28], an extension of Theorem 5.2 for $\Sigma = \{\infty\}$ was presented where the Haar measure on U is replaced by a (uniformly) friendly measure μ. While for simplicity of notation we keep our treatment of friendly measures to this case, Kleinbock and Tomanov wrote in [30] the Σ-arithmetic non-divergence results also for the case of friendly measures. The only difference between the statement of Theorem 5.2 and the analogous statement for uniformly friendly measures (other than the obvious difference of how the size of subsets of $\lambda_k(\mathsf{B}_U)$ are measured) is that the exponent $1/D$ of the theorem is allowed to depend on the doubling constant for μ. Theorem 5.3 also holds for uniformly friendly measures with the same modification. We also note (and use below) that in view of [28, Prop. 7.33], an analogue of Lemma 5.5 holds true for μ in place of the Haar measure on U (with a different c and exponent).

Repeating the proof of Theorem 6.3 but with the (uniformly) friendly versions of Theorems 5.2 and 5.3, we obtain the following:

9.1 Theorem. *Let μ be a uniformly friendly measure on $\mathbb{R}$. There exists a constant F depending on N and μ so that for any $g \in G$, $k \geq 1$, and any $0 < \eta \leq 1/2$ small enough at least one of the following holds:*

(1)
$$\mu(\{t \in [-1, 1] : u(e^k t)g\Gamma \notin X_\eta\}) \ll \eta^{1/F}.$$

(2) *There exists a unipotent $\mathbb{Q}$-subgroup $\mathbf{W}$ so that*
$$\|(\eta_W(u(e^k t)g)\| \ll \eta^{1/F} \text{ for all } t \in [-1, 1].$$

Moreover, if we put $\mathbf{M} = \mathbf{M}_\mathbf{W}$, then $\mathbf{M} \neq \mathbf{G}$,
$$\mathrm{ht}(\mathbf{M}) \ll |g|^F \eta^{1/F},$$

and we have:

(a) *For all $t \in [-1, 1]$ we have*
$$\|\eta_M(u(e^k t)g)\| \ll |g|^F \eta^{1/F}.$$

(b) *For all $t \in [-1, 1]$ we have*
$$\max_{z \in \mathcal{B}_U} \|z \wedge \eta_M(u(e^k t)g)\| \ll |g|^F \eta^{1/F} e^{-k/F}.$$

Similarly, the proof of Theorem 1.4 is easily adapted to the friendly case, giving:

9.2 Theorem. *Let μ be a uniformly friendly measure on $\mathbb{R}$. There are constants A, D depending only on N and μ, and E_1 depending on N, G, Γ, and μ so that the following holds. Let $g \in G$, $k \geq 1$, and $0 < \eta < 1/2$. Assume $\varepsilon : \mathbb{R}^+ \to (0, 1)$ satisfies for any $s \in \mathbb{R}^+$ that*
$$\varepsilon(s) \leq \eta^A s^{-A}/E_1.$$

Then at least one of the following three possibilities holds:

(1) $\mu(\{\xi \in [-1, 1] : u(e^k\xi)g\Gamma \notin X_\eta$ *or* $u(e^k\xi)g\Gamma$ *is not* (ε, t)*-Diophantine*$\}) < E_1\eta^{1/D}$.

(2) *There exist a nontrivial proper subgroup $\mathbf{H} \in \mathcal{H}$ of*
$$\mathrm{ht}(\mathbf{H}) \leq E_1(|g|^A + e^{At})\eta^{-A}$$

so that the following hold for all $\xi \in [-1, 1]$:
$$\|\eta_H(u(e^k\xi)g)\| \leq E_1(|g|^A + e^{At})\eta^{-A},$$
$$\|z \wedge \eta_H(u(e^k\xi)g)\| \leq E_1 e^{-k/D}(|g|^A + e^{At})\eta^{-A},$$

where z is as in (1.1).

(3) *There exists a nontrivial proper normal subgroup* $\mathbf{H} \lhd \mathbf{G}$ *of*

$$\mathrm{ht}(\mathbf{H}) \le E_1 e^{At} \eta^{-A}$$

so that

$$\|z \wedge v_H\| \le E_1 \eta^{-A} \varepsilon (\mathrm{ht}(\mathbf{H})^{1/A} \eta^A / E_1)^{1/A}.$$

As a consequence Theorem 1.7 follows; see the proof of Theorem 8.1. We also get the following analogues of Theorem 1.1 whose proof is mutatis mutandis the same as the proof of Theorem 8.4.

9.3 Theorem. *Let* μ *be a uniformly friendly measure on* $\mathbb{R}$. *Let* $\eta > 0$. *Let*

$$\{H_i : 1 \le i \le r\} \subset \mathcal{H}$$

be a finite subset consisting of proper subgroups, and for each $1 \le i \le r$ *let* $\mathcal{C}_i \subset N_G(U, H_i)$ *be a compact subset. There exists an open neighborhood* $\mathcal{O} = \mathcal{O}(\alpha, \{H_i\}, \{\mathcal{C}_i\})$ *so that* $X \setminus \mathcal{O}$ *is compact and disjoint from* $\cup_i \mathcal{C}_i \Gamma/\Gamma$ *so that the following holds. For every* $x \in \mathcal{G}(U)$ *there exists some* $k_0 = k_0(\mu, \eta, \{H_i\}, \{\mathcal{C}_i\}, x)$ *so that for all* $k \ge k_0$ *we have*

$$\mu(\{t \in [-1, 1] : u(e^k t)x \in \mathcal{O}\}) < \eta.$$

Appendix A Proof of Theorem B

In this section we prove Theorem B. In qualitative form, this is proved by Greenberg in [23] and [24]. We reproduce the argument here to make the estimates explicit.

Proof of Theorem B. Let $\mathbb{C}_p$ denote the completion of the algebraic closure of $\mathbb{Q}_p$ for all $p \in \Sigma_f$; as abstract fields, $\mathbb{C}$ and $\mathbb{C}_p$, for any $p \in \Sigma_f$, are isomorphic. Therefore, $\mathbb{C}^m$ in Effective Nullstellensatz theorem of §4.11 may be replaced by $\mathbb{C}_p^m$ for any $p \in \Sigma_f$.

As in [23, pp. 59–60] and [24, Steps 1 and 2] we begin with some reductions.

Let $\mathcal{I} \subset \mathbb{Z}[t_1, \ldots, t_m]$ be the ideal generated by $\{f_i\}$, and let $\mathbf{Y}$ be the variety defined by $\mathcal{I}$ in $\mathbb{C}_p^m$.

Put $\mathcal{J} := \mathcal{I}\mathbb{Q}[t_1, \ldots, t_m]$. The radical and the primary decomposition of $\mathcal{J}$ in $\mathbb{Q}[t_1, \ldots, t_m]$ can be computed, see [2, Chap. 8.7]; this computation uses the Gröbner basis and yields the following. There exists a computable constant $s = s(m, n, D_0)$ so that

- $(\sqrt{\mathcal{J}})^s \subset \mathcal{J}$,
- $\sqrt{\mathcal{J}} = \cap_1^b \mathcal{P}_j$ where $\mathcal{P}_j$ is a prime ideal for all $1 \le j \le b \le s$, and
- $\mathcal{P}_j$ is generated by $\{f_{j,\ell} : 1 \le \ell \le s\}$ where the total degree of $\{f_{j,\ell}\}$ is bounded by s and the logarithmic height of the numerators and denominators of each $f_{j,\ell}$ is controlled by $s\mathsf{h}$.

Moreover, by [21, Cor. 3.8], we may replace s with $s' \geq s$, which is again computable and depends only on m, n, and D_0, so that the following holds. For every $1 \leq j \leq b$ the ideal $\mathcal{P}_j \cap \mathbb{Z}[t_1, \ldots, t_m]$ is generated by $\{g_{j,k} : 1 \leq k \leq s'\}$; furthermore, the total degree of $\{g_{j,k}\}$ is bounded by s' and for every j and k the logarithmic height of $g_{j,k}$ is bounded by $s'\mathsf{h}$.

Altogether, we may assume that $\mathcal{I}$ is a prime ideal, i.e., $\mathbf{Y}$ is $\mathbb{Q}$-irreducible.

We now use induction on $u := \dim \mathbf{Y}$ to prove the claim, see [23, Case 1, p. 60] and [24, Step 3].

The base case is when $u = -1$, that is: when $\mathcal{I}$ contains a nonzero constant. In this case we use the effective nullstellensatz theorem above and find some $a \in \mathbb{Z}$ with

$$\log |a| \leq (8D_0)^{4M-1}(\mathsf{h} + 8D_0 \log(8D_0))$$

where $M = 2^{m-1}$ so that $a = \sum_i q_i f_i$.

This implies the claim in the theorem when $u = -1$.

Assume now that $\mathbf{Y}$ is non-empty and that the theorem is established in dimensions less than u. Let Jac be the Jacobian matrix of $\{f_i\}$ and let Δ be the system of minors of order $m - u$ taken from Jac. Since $\mathrm{char}(\mathbb{Q}) = 0$, the locus of common zeros of $\{\Delta, \{f_i\}\}$ is a proper $\mathbb{Q}$-subvariety of $\mathbf{Y}$. By inductive hypothesis, thus, there exists some d' depending on m, n, and D which satisfies the claim in the theorem for $\{\Delta, \{f_i\}\}$.

For any $1 \leq \alpha_1 < \cdots < \alpha_{m-u} \leq n$, put

$$(\alpha) = (\alpha_1, \ldots, \alpha_{m-u})$$

and set $f_{(\alpha)} = \{f_{\alpha_1}, \ldots, f_{\alpha_{m-u}}\}$. Let $\mathbf{Y}_{(\alpha)}$ be the variety defined by $f_{(\alpha)}$. Let

$$\mathbf{Z}_{(\alpha)} = \bigcup_{j=1}^{c} \mathbf{Z}_{(\alpha),j},$$

where for all $1 \leq j \leq b$, we have $\mathbf{Z}_{(\alpha),j} \subset \mathbf{Y}_{(\alpha)}$, $\mathbf{Z}_{(\alpha),j}$ is $\mathbb{Q}$-irreducible with

$$\dim \mathbf{Z}_{(\alpha),j} = u,$$

and $\mathbf{Z}_{(\alpha),j} \neq \mathbf{Y}$.

Let $\mathcal{I}_{(\alpha),j} \subset \mathbb{Z}[t_1, \ldots, t_m]$ be the ideal corresponding to $\mathbf{Z}_{(\alpha),j}$. Since $\mathbf{Y}_{(\alpha)}$ is defined by $f_{(\alpha)}$, a similar argument as above implies that there exists a computable constant $r = r(m, n, D_0)$ so that

- $c \leq r$, and
- for every $1 \leq j \leq c$, there exists $\{g_{(\alpha),j,k} : 1 \leq k \leq r\}$ so that $\mathcal{I}_{(\alpha),j}$ is generated by $\{g_{(\alpha),j,k}\}$, further, the total degree of $\{g_{(\alpha),j,k}\}$ is bounded by r and for every j and k the logarithmic height of $g_{(\alpha),j,k}$ is controlled by $r\mathsf{h}$

Since $\mathbf{Z}_{(\alpha),j} \neq \mathbf{Y}$ for all j, by inductive hypothesis, there exists $\mathscr{d}'_{(\alpha)}$ depending on m, n, and D_0 which satisfies the claim in the theorem for

$$\{\{g_{((\alpha),j,k)} : 1 \leq k \leq r\}, \{f_i : 1 \leq i \leq n\}\},$$

for all $1 \leq j \leq c$.

Given $(\alpha) = (\alpha_1, \dots, \alpha_{m-u})$ and $(\beta) = (\beta_1, \dots, \beta_{m-u})$ let $\Delta_{(\alpha),(\beta)}$ denote the corresponding minor from Jac. By the implicit function theorem, if $z \in \mathbf{Y}_{(\alpha)}$ is such that $\Delta_{(\alpha),(\beta)}(z) \neq 0$ for some (β), then z lies on exactly one component of $\mathbf{Y}_{(\alpha)}$, moreover, that component has dimension u.

Let $r' = r'(m, n, D_0)$ be so that the logarithmic height of $\{\Delta, \{f_i\}\}$ is bounded by $r'\mathsf{h}$. Define

$$\mathscr{d} := 2r'\mathscr{d}' + r\sum_{(\alpha)} \mathscr{d}'_{(\alpha)}.$$

We claim that the theorem holds with $A_{11} = \mathscr{d}$.

Let $w = (w_1, \dots, w_m)$ be as in the statement of the theorem. If either

(1) $v_p(\Delta_{(\alpha,(\beta))}(w)) > 2r'\mathscr{d}'\mathsf{h}$ for all (α) and (β), or

(2) $v_p(g_{(\alpha),j,k}(w)) > 2r\mathscr{d}'_{(\alpha)}\mathsf{h}$ for some (α), some j, and all k,

then we get the claim from the inductive hypothesis.

Therefore, we may assume that there are (α) and (β) so that

$$v_p(\Delta_{(\alpha),(\beta)}(w)) \leq 2r'\mathscr{d}'\mathsf{h}, \tag{A.1}$$

and for every (θ) and every j there exists some k so that

$$v_p(g_{(\theta),j,k}(w)) \leq 2r\mathscr{d}'_{(\theta)}\mathsf{h}. \tag{A.2}$$

Now a suitable version of Hensel's Lemma, see [24, Note 1], implies that there exists some $y \in \mathbb{Z}_p^m$ so that $f_{(\alpha)}(y) = 0$ and

$$v_p(y - w) > C_2 - 2r'\mathscr{d}'\mathsf{h}. \tag{A.3}$$

The theorem follows if we show that $y \in \mathbf{Y}$.

Let us recall that

$$C_2 > 4r'\mathscr{d}'\mathsf{h} + 2r(\max\{\mathscr{d}'_{(\theta)} : (\theta)\})\mathsf{h}. \tag{A.4}$$

Then, (A.2), (A.3), and (A.4), imply that $v_p(g_{(\theta),j,k}(w)) = v_p(g_{(\theta),j,k}(y))$. In particular, $y \notin \mathbf{Z}_{(\theta),j}$ for all (θ) and all j. Similarly, (A.1), (A.3), and (A.4) imply that $\Delta_{(\alpha),(\beta)}(y) \neq 0$.

Thus, the implicit function theorem implies that y belongs to $\mathbf{Y}$. $\square$

REFERENCES

[1] M. Artin, *Algebraic approximation of structures over complete local rings*, Inst. Hautes Études Sci. Publ. Math. **36** (1969), 23–58.

[2] T. Becker and V. Weispfenning, *Gröbner Basis, A Computational Approach to Commutative Algebra*, Springer, New York, 1993.

[3] F. Bien, A. Borel and J. Kollár, *Rationally connected homogeneous spaces*, Invent. Math. **124** (1996), 103–127.

[4] E. Bombieri and W. Gubler, *Heights in Diophantine Geometry*, Cambridge University Press, Cambridge, 2006.

[5] A. Borel, *Linear Algebraic Groups*, Springer, New York, 1991.

[6] A. Borel and Harish-Chandra, *Arithmetic subgroups of algebraic groups*, Ann. of Math. (2) **75** (1962), 485–535.

[7] A. Borel and T. Springer, *Rationality properties of linear algebraic groups. II*, Tôhoku Math. J. (2) **20** (1968), 443–497.

[8] A. Borel and J. Tits, *Éléments unipotents et sous-groupes paraboliques de groupes réductifs, I*, Invent. Math. **12** (1971), 95–104.

[9] N. Bourbaki, Groupes et Algébre de Lie, Chapitres 7 et 8, Hermann, Paris, 1975.

[10] E. Breuillard and N. de Saxce, in preparation.

[11] W. Brownawell, *Local diophantine Nullstellen inequalities*, J. Amer. Math. Soc. **1** (1988), 311–322.

[12] Ju. A. Brudnyi and M. I. Ganzburg. *A certain extremal problem for polynomials in n variables*, Izv. Akad. Nauk SSSR Ser. Mat. **37** (1973), 344–355.

[13] S. Dani, *On orbits of unipotent flows on homogeneous spaces, II*, Ergodic Theory Dynam. Systems **6** (1986), 167–182.

[14] S. Dani and G. Margulis, *Values of quadratic forms at primitive integral points*, Invent. Math. **98** (1989), 405–424.

[15] S. Dani and G. Margulis *Orbit closures of generic unipotent flows on homogeneous spaces of* $\mathrm{SL}(3,\mathbb{R})$, Math. Ann. **286** (1990), 101–128.

[16] S. Dani and G. Margulis, *Asymptotic behaviour of trajectories of unipotent flows on homogeneous spaces*, Proc. Indian Acad. Sci. Math. Sci. **101** (1991), 1–17.

[17] S. Dani and G. Margulis, *Limit distributions of orbits of unipotent flows and values of quadratic forms*, in *I. M. Gelfand Seminar*, American Mathematical Society, Providence, RI, 1993, pp. 91–137.

[18] M. Einsiedler, G. Margulis and A. Venkatesh, *Effective equidistribution for closed orbits of semisimple groups on homogeneous spaces*, Invent. Math. **177** (2009), 137–212.

[19] M. Einsiedler, G. Margulis, A. Mohammadi and A. Venkatesh, *Effective equidistribution and property* (τ), J. Amer. Math. Soc. **33** (2020), 223–289.

[20] A. Eskin, S. Mozes and N. Shah, *Unipotent flows and counting lattice points on homogeneous varieties*, Ann. of Math. (2) **143** (1996), 253–299.

[21] P. Gianni, B. Trager and G. Zacharias, *Gröbner basis and primary decomposition of polynomial ideals*, J. Symb. Comp. **6** (1988), 149–167.

[22] E. Glasner, *Ergodic Theory via Joinings*, American Mathematical Society, Providence, RI, 2003.

[23] M. Greenberg, *Rational points in Henselian discrete valuation rings*, Inst. Hautes Études Sci. Publ. Math. **31** (1964), 59–64.

[24] M. Greenberg, *Strictly local solutions of Diophantine equations*, Pacific J. Math. **51** (1974), 143–153, 1974.

[25] A. Gorodnik and H. Oh, *Rational points on homogeneous varieties and equidistribution of adelic periods*, Geom. Funct. Anal. **21** (2011), 319–392.

[26] S. Ji, J. Kollár and B. Shiffman, *A global Lojasiewicz inequality for algebraic varieties*, Trans. Amer. Math. Soc. **329** (1992), 813–818.

[27] D. Kleinbock, *An extension of quantitative nondivergence and applications to Diophantine exponents*, Trans. Amer. Math. Soc. **360** (2008), 6497–6523.

[28] D. Kleinbock, E. Lindenstrauss and B. Weiss, *On Fractal Measures and Diophantine Approximation*, Selecta Math. **10** (2004), 479–523.

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

[30] D. Kleinbock and G. Tomanov, *Flows on S-arithmetic homogenous spaces and application to metric Diophantine approximation*, Comment. Math. Helv. **82** (2007), 519–581.

[31] E. Lindenstrauss and G. Margulis, *Effective estimates on indefinite ternary forms*, Israel J. of Math. **203** (2014), 445–499.

[32] G. Margulis. *On the action of unipotent group in the space of lattices*, in *Lie Groups and their Representations (Proc. Summer School, Bolyai János Math. Soc., Budapest, 1971)*, Halsted, New York, 1975, pp. 365–370.

[33] G. Margulis, *Indefinite quadratic forms and unipotent flows on homogeneous spaces*, in *Dynamical Systems and Ergodic Theory (Warsaw, 1986)*, PWN, Warsaw, 1989, pp. 399–409.

[34] G. Margulis, *Discrete Subgroups of Semisimple Lie Groups*, Springer, Berlin, 1990.

[35] G. Margulis and G. Tomanov, *Invariant measures for actions of unipotent groups over local fields on homogeneous spaces*, Invent. Math. *116* (1994), 347–392.

[36] D. Masser and G. Wüstholz, *Fields of large transcendence degree generated by values of elliptic functions*, Invent. Math. **72** (1983), 407–464.

[37] A. Mohammadi, A. Salehi Golsefidy and F. Thilmany, Diameter of homogeneous spaces: An effective account, Math. Ann. **385** (2023), 1973–2012.

[38] G. Mostow, *Fully Reducible Subgroups of Algebraic Groups.* American J. of Math. **78** (1956), 200–221.

[39] S. Mozes and N. Shah, *On the space of ergodic invariant measures of unipotent flows*, Ergodic Theory Dynam. Systems **15** (1995), 149–159.

[40] V. Platonov and A. Rapinchuk, *Algebraic Groups and Number Theory*, Academic press, Boston, MA,1994.

[41] M. Ratner, *Raghunathan's topological conjecture and distributions of unipotent flows*, Duke Math. J. **63** (1991), 235–280.

[42] M. Ratner, On Raghunathan's measure conjecture, Ann. Math. (2) **134** (1992), 545–607.

[43] M. Ratner, *Raghunathan's conjectures for Cartesian products of real and p-adic Lie groups*, Duke Math. J. **77** (1995), 275–382.

[44] A. Schinzel, *Polynomials with Special Regard to Reducibility*, Cambridge University Press, Cambridge, 2000.

[45] A. Seidenberg, *Constructions in algebra*, Trans. Amer. Math. Soc. **197** (1974), 273–313.

[46] N. Shah, *Uniformly distributed orbits of certain flows on homogeneous spaces*, Math. Ann. **289** (1991), 315–334.

[47] N. Shah, *Limit distributions of polynomials on homogeneous spaces*, Duke Math. J. **73** (1994), 711–731.

[48] N. Shah, *Limit distributions of expanding translates of certain orbits on homogeneous spaces*, Proc. Indian Acad. Sci. Math. Sci. **106** (1996), 105–125.

[49] J. Tits, *Reductive groups over local fields*, in *Automorphic Forms, Representations and L-Functions*, American Mathematical Society, Providence, RI, 1979, pp. 29–69.

[50] G. Tomanov, *Orbits on homogeneous spaces of arithmetic origin and approximations*, in *Analysis on Homogeneous Spaces and Representation Theory of Lie Groups, Okayama–Kyoto (1997)*, The Mathematical Society of Japan, Tokyo, 2000, pp. 265–297.

[51] G. Tomanov and B. Weiss, *Closed orbits for actions of maximal tori on homogeneous spaces*, Duke Math. J. **199** (2003), 367–392.

Elon Lindenstrauss
THE EINSTEIN INSTITUTE OF MATHEMATICS
THE HEBREW UNIVERSITY OF JERUSALEM
EDMOND J. SAFRA CAMPUS
GIVAT RAM, JERUSALEM 91904, ISRAEL
email: elon@math.huji.ac.il

Gregorii Margulis
MATHEMATICS DEPARTMENT
YALE UNIVERSITY
NEW HAVEN, CT 06511, USA
email: gregorii.margulis@yale.edu

Amir Mohammadi
DEPARTMENT OF MATHEMATICS
UNIVERSITY OF CALIFORNIA AT SAN DIEGO
LA JOLLA, CA 92093-0021, USA
email: ammohammadi@ucsd.edu

Nimish Shah
DEPARTMENT OF MATHEMATICS
THE OHIO STATE UNIVERSITY
COLUMBUS, OH 43210, USA
email: shah@math.osu.edu

(Received March 31, 2019 and in revised form July 20, 2022)