Ergod. Th. & Dynam. Sys., (2022), **42**, 3656–3715
doi:10.1017/etds.2021.110

Effective equidistribution for generalized higher-step nilflows

MINSUNG KIM

Department of Mathematics, University of Maryland, College Park, MD 20742, USA;
and
Faculty of Mathematics and Computer Science, Nicolaus Copernicus University, Chopina 12/18, 87-100 Toruń, Poland
(e-mail: mkim16@mat.umk.pl)

(*Received* 25 *January* 2020 *and accepted in revised form* 27 *August* 2021)

Abstract. In this paper we prove bounds for ergodic averages for nilflows on general higher-step nilmanifolds. Under Diophantine condition on the frequency of a toral projection of the flow, we prove that almost all orbits become equidistributed at polynomial speed. We analyze the rate of decay which is determined by the number of steps and structure of general nilpotent Lie algebras. Our main result follows from the technique of controlling scaling operators in irreducible representations and measure estimation on close return orbits on general nilmanifolds.

Key words: nilflows, cohomological equations, ergodic averages, coadjoint orbits
2020 Mathematics Subject Classification: 37A17 (Primary); 22E27, 37A25, 37A45, 37A46 (Secondary)

Contents

1. *Introduction*

By a general result of Green and Tao [**GT12**], all orbits of Diophantine flows on any nilmanifold become equidistributed at polynomial speed. Their approach is an extension of Weyl's method, based on induction on the number of steps, but the rate of decay in their theorem is not explicit and presumably far from optimal. Flaminio and Forni also established estimates for the *quadratic* polynomial speed of equidistribution of nilflows on *higher-step* nilmanifolds [**FF14**] called quasi-abelian (filiform). This is the simplest class of nilmanifolds of arbitrarily higher-step structures, and it has an application in proving the bound of an exponential sum called the *Weyl sum*. It is notable that the bound obtained from Flaminio and Forni is established only almost everywhere but comparable with the results by Wooley [**T15**] from number theory (see also [**BDG15**]).

In this paper we extend the result for quasi-abelian to a non-renormalizable class of nilflows on higher-step nilmanifolds under Diophantine conditions on the frequencies of their toral projections (see Definition 5.15).

For any $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$, for any $N \in \mathbb{N}$ and every $\delta > 0$, let

$$R_\alpha(N, \delta) = \{r \in [-N, N] \cap \mathbb{Z} \mid |r\alpha|_1 \le \delta^{1/n}, \ldots, |r\alpha|_n \le \delta^{1/n}\}.$$

For every $\nu > 1$, let $D_n(\nu) \subset (\mathbb{R}\backslash\mathbb{Q})^n$ be the subset defined as follows: the vector $\alpha \in D_n(\nu)$ if and only if there exists a constant $C(\alpha) > 0$ such that, for all $N \in \mathbb{N}$ for all $\delta > 0$,

$$\#R_\alpha(N, \delta) \le C(\alpha) \max\{N^{1-(1/\nu)}, N\delta\}. \tag{1}$$

The Diophantine condition $D_n(\nu)$ contains the set of simultaneously Diophantine vectors so that $D_n(\nu)$ has a full measure for sufficiently large $\nu \ge 1$.

For a set of generator $\mathfrak{G}_\alpha$ of $\mathfrak{n}$, there exist an element $X_\alpha \in \mathfrak{G}_\alpha$ and a codimension 1 ideal $\mathfrak{I}$ such that $X_\alpha \notin \mathfrak{I} \subset \mathfrak{n}$. We assume that the Lie algebra satisfies the *transversality*

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

condition if

$$\langle \mathfrak{G}_\alpha \rangle + \mathrm{Ran}(\mathrm{ad}_{X_\alpha}) + C_{\mathfrak{I}}(X_\alpha) = \mathfrak{n}$$

where $C_{\mathfrak{I}}(X_\alpha) = \{Y \in \mathfrak{I} \mid [Y, X_\alpha] = 0\}$ is the centralizer.

Under this hypothesis, the rate of convergence for ergodic average of nilflows $(\phi^t_{X_\alpha})$ under Diophantine conditions $\alpha \in D_n(\nu)$ is polynomial for *almost all* points, as a function of step size and total number of elements of Lie algebras.

THEOREM 1.1. *Let $(\phi^t_{X_\alpha})$ be a nilflow on a k-step nilmanifold M on $n+1$ generators such that the projected toral flow $(\bar{\phi}^t_{X_\alpha})$ is a linear flow with frequency vector $\alpha := (1, \alpha_1, \ldots, \alpha_n) \in \mathbb{R} \times \mathbb{R}^n$. Assume that the Lie algebra satisfies the transversality condition and $\alpha \in D_n(\nu)$ for some $1 \le \nu \le k/2$. Then there exists a Sobolev norm $\|\cdot\|$ on the space $C^\infty(M)$ of smooth function on M and for every $\epsilon > 0$ there exists a positive measurable function $K_\epsilon \in L^p(M)$ for all $p \in [1, 2)$, such that the following bound holds. For every smooth zero-average function $f \in C^\infty(M)$, for every $T \ge 1$, for almost all $x \in M$,*

$$\left| \frac{1}{T} \int_0^T f \circ \phi^t_{X_\alpha}(x)\, dt \right| \le K_\epsilon(x) T^{-(1/3S_{\mathfrak{n}}(k))+\epsilon} \|f\|$$

where $S_{\mathfrak{n}}(k)$ is a higher order polynomial determined by the structure of $\mathfrak{n}$. Specifically, if n_i is the number of elements in $\mathfrak{n}$ with step size i,

$$S_{\mathfrak{n}}(k) := (n_1 - 1)(k-1) + n_2(k-2) + \cdots + n_{k-1}. \tag{2}$$

In the general higher-step nilmanifold, no renormalization for nilflows is known. Instead, based on the theory of unitary representations for the nilpotent Lie group (Kirillov theory), it is possible to choose a proper scaling operator on the space of invariant distributions. Compared to earlier work on the quasi-abelian case [**FF14**], the main novelty of our results lies in generalization of the scaling method to the general Lie algebra satisfying a transversality condition.

The transversality condition enables the measure estimate (§5) for the return orbit. This condition is sufficient, and in principle there are no obstructions to a generalization to arbitrary nilflows with Diophantine frequencies and all points $x \in M$, except that this would require new approaches to estimation other than a Borel–Cantelli type argument. On the other hand, the necessity of the condition explains that the total number of elements in the basis cannot grow too fast as the step size gets larger: it grows almost linearly in the number of steps and generators.

We can view this phenomenon in the following way. If the growth of the number of elements in lower steps (generated by basis) is too large, then it lacks the dimensions to count the measure of the return orbit on the transverse manifold. For instance, we observe this phenomenon in free nilpotent Lie algebras. Even a small number of generators creates a large number of elements in the lower level under small steps of commutations, which behave in a completely different way than strictly triangular and quasi-abelian. We present such an example in the appendix to motivate our condition.

The above theorem is supplemented by its corollary on the *strictly triangular nilmanifold* $M_k^{(k)}$. Let $N_k^{(k)}$ denote a step-k nilpotent Lie group on k generators. Up to isomorphism, $N_k^{(k)}$ is the group of upper triangular unipotent matrices

$$[x_1X_1, \ldots x_nX_n, \ldots y_i^{(j)}Y_i^{(j)} \cdots zZ] := \begin{pmatrix} 1 & x_1 & \cdots & \cdots & z \\ 0 & 1 & x_2 & y_i^{(j)} & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & x_n \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad x_i, y_i^{(j)}, z \in \mathbb{R}, \tag{3}$$

with one-dimensional center. The next result states that the rate of equidistribution for nilflows on the triangular nilmanifold $M_k^{(k)}$ decays at a polynomial speed with exponent, and *cubically* as a function of number of steps.

COROLLARY 1.2. *Let $(\phi^t_{X_\alpha})$ be a nilflow on k-step strictly triangular nilmanifold M on k generators such that the projected toral flow $(\bar{\phi}^t_{X_\alpha})$ is a linear flow with frequency vector $\alpha := (1, \alpha_1, \ldots, \alpha_{k-1}) \in \mathbb{R} \times \mathbb{R}^{k-1}$. Under the condition that $\alpha \in D_n(\nu)$ for some $1 \le \nu \le k/2$, there exists a Sobolev norm $\|\cdot\|$ on the space $C^\infty(M_k^{(k)})$ of smooth function on $M_k^{(k)}$ and there exists a positive measurable function $K_\epsilon \in L^p(M_k^{(k)})$ for all $p \in [1, 2)$ and for every $\epsilon > 0$, such that the following bound holds. For every smooth zero-average function $f \in C^\infty(M_k^{(k)})$, for almost all $x \in M$ and for every $T \ge 1$,*

$$\left| \frac{1}{T} \int_0^T f \circ \phi^t_{X_\alpha}(x)\, dt \right| \le K_\epsilon(x) T^{-(1/(3(k-1)(k^2+k-3)))+\epsilon} \|f\|.$$

We also establish the uniform bound for the step-3 strictly triangular nilmanifold case. The result holds for *all points* by counting arguments for linearly divergent close return orbits under Roth-type Diophantine condition. The step-3 case (as well as the filiform case [**F16**]) is a good example for deriving a simplified proof.

THEOREM 1.3. *Let (ϕ^t_X) be a nilflow on the 3-step nilmanifold M on three generators such that the projected toral flow $(\bar{\phi}^t_X)$ is a linear flow with frequency vector $\upsilon := (1, \alpha, \beta)$ of Diophantine condition with exponent $\nu = 1 + \epsilon$ for all $\epsilon > 0$. For every $s > 26$, there exists a constant C_s such that for every zero-average function $f \in W^s(M)$, for all $(x, T) \in M \times \mathbb{R}$, we have*

$$\left| \frac{1}{T} \int_0^T f \circ \phi^t_X(x)\, dt \right| \le C_s T^{-1/12+\epsilon} \|f\|_s.$$

In the last section, we present exponential mixing of hyperbolic nilautomorphisms as a main application. Exponential mixing of ergodic automorphisms and its applications to the central limit theorem on compact nilmanifolds was proven by Gorodnik and Spatzier [**GS14**]. Their approach was based on the result of Green and Tao [**GT12**], and mixing follows from the equidistribution of the exponential map called *box map* satisfying certain Diophantine conditions. Our result also shows specific exponent of exponential mixing

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

depending on the structure of nilmanifolds, which follows from equidistribution results and a renormalization argument on partial hyperbolic automorphisms (hyperbolic on the projected torus). However, they are limited to a special class of nilautomorphisms due to lack of hyperbolicity on the group of automorphisms on general nilpotent Lie algebras (cf. triangular step 3 with three generators).

Our work leaves open questions.

Problem 1. Can we find *uniform* bounds on the average width under the transversality condition? That is, can we prove Theorem 1.1 with uniform bound for *all* points?

Problem 2. Give an effective bound on ergodic averages on any nilmanifold.

It is still conjectured that the result may hold for *all points on any nilmanifolds*. However, our argument for averaged width can not be improved to control the slow growth of displacement for the return orbit on general higher-step cases.

This paper is organized as follows. In §2 we define structures of nilmanifolds and nilflows. In §3 we find Sobolev estimates on solutions of the cohomological equation and on invariant distributions as an application of Kirillov theory of unitary representations of nilpotent groups. In §4 we introduce the notion of average width and prove a Sobolev trace theorem. In §5 we prove an effective equidistribution theorem for good points by a Borel–Cantelli argument. In §6 we prove bounds on the average width of an orbit segment of nilflows by gluing all the irreducible representations. In §7 we introduce a uniform width estimate under a Roth-type Diophantine condition based on counting return time directly without a good point argument for the width estimate in §5.4. Finally, in §8, as an application, we prove the mixing of nilautomorphisms.

2. *Nilflows on higher-step nilmanifolds*

In this section we review nilpotent Lie algebras, groups and basic structures. We recall Kirillov theory and representation theory.

2.1. *Background on nilpotent Lie groups and Lie algebras.*

Let N be a connected, simply connected k-step nilpotent Lie group with Lie algebra $\mathfrak{n}$ with $n+1$ generators. Let Γ be a cocompact lattice in N. The quotient $N/\Gamma := M$ is then a compact nilmanifold on which the left action of N is given by translations. Denote by μ the N-invariant measure on M.

For $j = 1, \ldots, k$, let $\mathfrak{n}_j$ denote the descending central series of $\mathfrak{n}$:

$$\mathfrak{n}_1 = \mathfrak{n}, \mathfrak{n}_2 = [\mathfrak{n}, \mathfrak{n}], \ldots, \mathfrak{n}_j = [\mathfrak{n}_{j-1}, \mathfrak{n}], \ldots, \mathfrak{n}_k \subset Z(\mathfrak{n}) \tag{4}$$

where $Z(\mathfrak{n})$ is the center of $\mathfrak{n}$. In this setting, there exists a strong Malcev basis through the filtration $(\mathfrak{n}_i)_{i=1}^k$ strongly based at the lattice Γ (see [**CG90**]). That is, given a basis

$$\mathscr{F} = \{\xi^{(1)}, \eta_1^{(1)}, \ldots, \eta_{n_1}^{(1)}, \ldots, \eta_1^{(k)}, \ldots \eta_{n_k}^{(k)}\}$$

with $\xi := \xi^{(1)} \in \mathfrak{n}_1 \backslash \mathfrak{n}_2$ and $\eta_i^{(l)} \in \mathfrak{n}_l \backslash \mathfrak{n}_{l+1}$ for $i = 1, \ldots, n_l$, we have the following assertions.

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

(1) If we drop the first l elements of the basis, we obtain a basis of a subalgebra $\mathfrak{n}$ of codimension l.
(2) For each j, the elements in order $\eta_1^{(j)}, \ldots \eta_{n_j}^{(j)}, \ldots \eta_1^{(k)}, \ldots \eta_{n_k}^{(k)}$ form a basis of an ideal $\mathfrak{n}_j$ of $\mathfrak{n}$.
(3) The lattice Γ is generated by $\{x, y_1^{(1)}, \ldots, y_{n_k}^{(k)}\}$ with

$$x := \exp(\xi), \quad y_{n_i}^{(j)} := \exp(\eta_{n_i}^{(j)}).$$

For any nilpotent Lie algebra $\mathfrak{n}$, there exists a codimension-1 subalgebra $\mathfrak{I}$ where

$$\mathfrak{n} = \mathbb{R}\xi \oplus \mathfrak{I}.$$

Then $\mathfrak{I}$ is an ideal and $[\mathfrak{n}, \mathfrak{n}] \subseteq \mathfrak{I}$ (see [**H73**, Ch. 3, p. 12] and [**CG90**, Lemma 1.1.8]). For convenience, we write dimension $a = \dim(\mathfrak{I}) = n_1 + \cdots + n_k$ and set $n = n_1$.

Definition 2.1. An *adapted basis* of the Lie algebra $\mathfrak{n}$ is an ordered basis $(X, Y) := (X, Y_1, \ldots, Y_a)$ of $\mathfrak{n}$ such that $X \notin \mathfrak{I}$ and $Y := (Y_1, \ldots, Y_a)$ is a basis of $\mathfrak{I}$.

A *strongly adapted* basis $(X, Y) := (X, Y_1, \ldots, Y_a)$ is an adapted basis such that the following assertions hold.
(1) The system $(X, Y_1, \ldots, Y_n)$ is a system of generators of $\mathfrak{n}$, hence its projection is a basis of the abelianization $\mathfrak{n}/[\mathfrak{n}, \mathfrak{n}]$ of the Lie algebra $\mathfrak{n}$.
(2) The system $(Y_{n+1}, \ldots, Y_a)$ is a basis of the ideal $[\mathfrak{n}, \mathfrak{n}]$.

2.2. *Nilmanifolds and nilflows.* Every nilmanifold M is a fiber bundle over a torus. In fact, the group $N^{ab} = N/[N, N]$ is abelian, connected and simply connected, hence isomorphic to $\mathbb{R}^{n+1}$, and $\Gamma^{ab} = \Gamma/[\Gamma, \Gamma]$ is a lattice in N^{ab}. Thus, we have a natural projection $pr_1 : M \to \mathbb{T}^{n+1}$.

We introduce two fibrations of the nilmanifold M. Let $M_2 \simeq N_2/\Gamma_2$ with $N_2 = \exp(\mathfrak{n}_2)$ and its lattice Γ_2. Then there exists an exact sequence

$$0 \to M_2 \to M \xrightarrow{pr_1} \mathbb{T}^{n+1} \to 0. \tag{5}$$

Another fibration arises from the canonical homomorphism $N \to N/N' \approx \langle \exp \xi \rangle$. For $\theta \in \mathbb{T}^1$, the fiber $M_\theta^a = {pr_2}^{-1}(\theta)$ is a local section of the nilflow on M:

$$0 \to M_\theta^a \to M \xrightarrow{pr_2} \mathbb{T}^1 \to 0. \tag{6}$$

On the nilmanifold M, the nilflow ϕ_X^t generated by $X \in \mathfrak{n}$ is the flow obtained by the restriction of this action to the one-parameter subgroup $(\exp tX)_{t\in\mathbb{R}}$ of N, with

$$\phi_X^t(x) = x \exp(tX), \quad x \in M,\ t \in \mathbb{R}.$$

The projection $\bar{X}$ of X is the generator of a linear flow $\psi_{\bar{X}} := \{\psi_{\bar{X}}^t\}_{t\in\mathbb{R}}$ on $\mathbb{T}^{n+1} \approx \mathbb{R}^{n+1}\backslash\bar{\Gamma}$ defined by

$$\psi_{\bar{X}}^t(x_1, \ldots, x_{n+1}) = (x_1 + tv_1, \ldots, x_{n+1} + tv_{n+1}).$$

The canonical projection $pr_1 : M \to \mathbb{T}^{n+1}$ intertwines the flows ϕ_X^t and $\psi_{\bar{X}}^t$.

We recall the following result.

THEOREM 2.2. [AGH63] *The following assertions are equivalent.*

(1) *The nilflow (ϕ^t_X) on M is ergodic.*
(2) *The nilflow (ϕ^t_X) on M is uniquely ergodic.*
(3) *The nilflow (ϕ^t_X) on M is minimal.*
(4) *The projected flow $(\psi^t_{\bar{X}})$ on $\overline{M} = N^{ab}/\Gamma^{ab} \simeq \mathbb{T}^{n+1}$ is an irrational linear flow.*

Consider the set of indices

$$\begin{aligned} J &:= \{(i,j) \mid 1 \le i \le n_j, 1 \le j \le k\}, \\ J^+ &:= \{(i,j) \mid 1 \le i \le n_1,\ j = 1\}, \\ J^- &:= \{(i,j) \mid 1 \le i \le n_j,\ j > 1\}, \\ J^{-2} &:= \{(i,j) \mid 1 \le i \le n_j,\ j > 2\}. \end{aligned}$$

Let $\alpha = (\alpha_i^{(j)}) \in \mathbb{R}^J$ and $X := X_\alpha$ be a vector field on M defined

$$X_\alpha := \log \left[x^{-1} \exp \left(\sum_{(i,j)\in J} \alpha_i^{(j)} \eta_i^{(j)} \right) \right], \quad x = \exp(\xi), \tag{7}$$

and equivalently we write

$$X_\alpha := -\xi + \sum_{(i,j)\in J} \alpha_i^{(j)} \eta_i^{(j)}. \tag{8}$$

For $\theta \in \mathbb{T}^1$, let $M^a_\theta = {pr_2}^{-1}(\theta)$ denote a fiber over $\theta \in \mathbb{T}^1$ of the fibration pr_2. The transverse section M^a_θ of the nilflow $\{\phi^t_{X_\alpha}\}_{t\in\mathbb{R}}$ corresponds to

$$\begin{aligned} &\left\{ \Gamma \exp(\theta\xi) \exp\left(\sum_{i=1}^a s_i \eta_i \right) \mid (s_i)_{i=1}^a \in \mathbb{R}^a \right\} \\ &\quad = \left\{ \Gamma \exp\left(e^{ad(\theta\xi)} \sum_{i=1}^a s_i \eta_i \right) \exp(\theta\xi) \mid (s_i)_{i=1}^a \in \mathbb{R}^a \right\}. \end{aligned}$$

LEMMA 2.3. *The flow $(\phi^t_{X_\alpha})_{t\in\mathbb{R}}$ on M is isomorphic to the suspension of its first return map $\Phi_{\alpha,\theta} : M^a_\theta \to M^a_\theta$. For every $(i,j) \in J$ and $r \ge 1$, there exists a polynomial $p^{(j)}_{i,r}(\alpha, \mathbf{s})$ for $\mathbf{s} = (s_i)^a_{i=1} \in \mathbb{R}^a$ such that the rth return map $\Phi^r_{\alpha,\theta}$ is written as follows.*

On the coordinates of $\mathbf{s} = (s_i^{(j)})$ for $\Gamma \exp(\theta\xi) \exp(\sum_{(i,j)\in J} s_i^{(j)} \eta_i^{(j)}) \in M^a_\theta$,

$$\begin{aligned} \Phi_{\alpha,\theta}(\mathbf{s}) = \Gamma \exp(\theta\xi) \exp\Bigg(&\sum_{(i,j)\in J} (s_i^{(j)} + \alpha_i^{(j)}) \eta_i^{(j)} \\ &+ \left[\sum_{(i,j)\in J} s_i^{(j)} \eta_i^{(j)}, X_\alpha \right] + \sum_{(i,j)\in J^{2-}} p_i^{(j)}(\alpha, \mathbf{s}) \eta_i^{(j)} \Bigg), \end{aligned} \tag{9}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

and for $r \in \mathbb{N}$*,*

$$\Phi^r_{\alpha,\theta}(\mathbf{s}) = \Gamma \exp(\theta\xi) \exp\bigg(\sum_{(i,j)\in J} (s_i^{(j)} + r\alpha_i^{(j)})\eta_i^{(j)} + \bigg[\sum_{(i,j)\in J} s_i^{(j)}\eta_i^{(j)}, rX_\alpha \bigg] + \sum_{(i,j)\in J^{2-}} p_{i,r}^{(j)}(\alpha, \mathbf{s})\eta_i^{(j)} \bigg). \quad (10)$$

Proof. By (7), we have

$$\begin{aligned}\exp\bigg(\sum_{(i,j)\in J} s_i^{(j)}\eta_i^{(j)} \bigg) \exp(X_\alpha) &= \exp\bigg(\sum_{(i,j)\in J} s_i^{(j)}\eta_i^{(j)} \bigg) x^{-1} \exp\bigg(\sum_{(i,j)\in J} \alpha_i^{(j)}\eta_i^{(j)} \bigg) \\ &= x^{-1} \exp(e^{\mathrm{ad}(\xi)} \sum_{(i,j)\in J} s_i^{(j)}\eta_i^{(j)}) \exp\bigg(\sum_{(i,j)\in J} \alpha_i^{(j)}\eta_i^{(j)} \bigg).\end{aligned}$$

By the Baker–Campbell–Hausdorff formula, there exist polynomials $p_i^{(j)}(\alpha, s)$ with

$$\exp\bigg(\sum_{(i,j)\in J} s_i^{(j)}\eta_i^{(j)} \bigg) \exp(X_\alpha) = x^{-1} \exp\bigg(\sum_{(i,j)\in J} (s_i^{(j)} + \alpha_i^{(j)})\eta_i^{(j)} + \bigg[\sum_{(i,j)\in J} s_i^{(j)}\eta_i^{(j)}, X_\alpha \bigg] + \sum_{(i,j)\in J^{2-}} p_i^{(j)}(\alpha, \mathbf{s})\eta_i^{(j)} \bigg).$$

Since $x \in \Gamma$, we conclude

$$\begin{aligned}&\Gamma \exp(\theta\xi) \exp\bigg(\sum_{(i,j)\in J} s_i^{(j)}\eta_i^{(j)} \bigg) \exp(X_\alpha) \\ &= \Gamma \exp(\theta\xi) \exp\bigg(\sum_{(i,j)\in J} (s_i^{(j)} + \alpha_i^{(j)})\eta_i^{(j)} + \bigg[\sum_{(i,j)\in J} s_i^{(j)}\eta_i^{(j)}, X_\alpha \bigg] \\ &\quad + \sum_{(i,j)\in J^{2-}} p_i^{(j)}(\alpha, \mathbf{s})\eta_i^{(j)} \bigg).\end{aligned}$$

The formula implies that $t = 1$ is a return time of the restriction of the flow to $M^a_\theta \subset M$. The formula for $r \in \mathbb{N}$ follows by induction. □

2.3. *Kirillov theory and classification.* Kirillov theory yields the complete classification of irreducible unitary representations of N. All the irreducible unitary representations of nilpotent Lie groups are parametrized by the coadjoint orbits $\mathcal{O} \subset \mathfrak{n}^*$. A *polarizing* (or maximal subordinate) subalgebra for $\Lambda \in \mathfrak{n}^*$ is a maximal isotropic subspace $\mathfrak{m} \subset \mathfrak{n}$ which is a subalgebra of $\mathfrak{n}$. It is well known that for any $\Lambda \in \mathfrak{n}^*$ there exists a polarizing subalgebra $\mathfrak{m}$ for a nilpotent Lie algebra $\mathfrak{n}$ (see [**CG90**, Theorem 1.3.3]). Let $\mathfrak{m}$ be a polarizing subalgebra for given linear form $\Lambda \in \mathfrak{n}^*$. Then the character $\chi_{\Lambda,\mathfrak{m}} : \exp \mathfrak{m} \to S^1$ is defined by

$$\chi_{\Lambda,\mathfrak{m}}(\exp Y) = e^{2\pi\iota\Lambda(Y)}.$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

Given the pair $(\Lambda, \mathfrak{m})$, we associate the unitary representation

$$\pi_\Lambda = \mathrm{Ind}_{\exp \mathfrak{m}}^{N}(\chi)$$

where the induced representation π_Λ is defined by

$$\pi_\Lambda(x) f(g) = f(g \cdot x) \quad \text{for } x \in N \text{ and } f \in H_{\pi_{\nu,\mathfrak{m}}}.$$

These unitary representations are irreducible up to equivalence, and all unitary irreducible representations are obtained in this way. It is known that Λ and Λ' belong to the same coadjoint orbit if and only if $\pi_{\Lambda,\mathfrak{m}}$ and $\pi_{\Lambda',\mathfrak{m}'}$ are unitarily equivalent and $\pi_{\Lambda,\mathfrak{m}}$ is irreducible whenever $\mathfrak{m}$ is maximal subordinate for Λ. We write $\pi_\Lambda \simeq \pi_{\Lambda'}$ if Λ and Λ' are in the same coadjoint orbit.

Since the action of N on M preserves the measure μ, we obtain a unitary representation π of N. The regular representation of $L^2(M)$ of N decomposes as a countable direct sum (or direct integral) of the irreducible, unitary representation H_π, which occurs with at most finite multiplicity

$$L^2(M, d\mu) = \bigoplus_{\pi} H_\pi. \tag{11}$$

The derived representation π_* of a unitary representation π of N on a Hilbert space H_π is a representation of the Lie algebra $\mathfrak{n}$ on H_π defined as follows. For every $X \in \mathfrak{n}$,

$$\pi_*(X) = \lim_{t\to 0}(\pi(\exp tX) - I)/t. \tag{12}$$

We recall that a vector $v \in H_\pi$ is of C^∞*-vectors* in H_π for representation π if the function $g \in N \mapsto \pi(g)v \in H_\pi$ is of class C^∞ as a function on N with values in a Hilbert space.

Definition 2.4. The space of *Schwartz functions* on $\mathbb{R}$ with values of C^∞ vectors for the representation π' on H' is denoted $\mathcal{S}(\mathbb{R}, C^\infty(H'))$. It is endowed with the Fréchet topology induced by the family of seminorms

$$\{\|\cdot\|_{i,j,Y_1,Y_2,\ldots,Y_m} \mid i, j, m \in \mathbb{N} \text{ and } Y_1, \ldots, Y_m \in \mathfrak{n}'\}$$

and defined as follows: for all $f \in \mathcal{S}(\mathbb{R}, C^\infty(H'))$,

$$\|f\|_{i,j,Y_1,Y_2,\ldots,Y_m} := \sup_{t\in\mathbb{R}} \|(1+t^2)^{j/2}\pi'_*(Y_1)\ldots\pi'_*(Y_m) f^{(i)}(t)\|_{H'}, \quad t \in \mathbb{R}.$$

LEMMA 2.5. [**FF07**, Lemma 3.4] *As a topological vector space*

$$C^\infty(H_\pi) = \mathcal{S}(\mathbb{R}, C^\infty(H'))$$

where $\mathcal{S}(\mathbb{R}, C^\infty(H'))$ *is Schwartz space.*

Suppose that $\mathfrak{n} = \mathbb{R}X \oplus \mathfrak{I}$ with its codimension-1 ideal $\mathfrak{I}$, and $N = \mathbb{R} \ltimes N'$ with a normal subgroup N' of N. Let π' be a unitary irreducible representation of N' on a Hilbert space H'. Each irreducible representation H_π is unitarily equivalent to $L^2(\mathbb{R}, H', dx)$, and the derived representation of π_* of the induced representation $\pi = \mathrm{Ind}_{N'}^{N}(\pi')$ has the following description.

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

For $f \in L^2(\mathbb{R}, H', dx)$, the group $\mathbb{R}$ acts by translations and its representation is polynomial in the variable x. For any $Y \in \mathfrak{n}'$,

$$(\pi_*(Y)f)(x) = \sum_{j=0}^{d_Y} \frac{1}{j!}\pi'_*(\mathrm{ad}_X^j(Y)f(x)$$
$$= \iota P_Y(x)f(x) = \iota \sum_{j=0}^{d_Y} \frac{1}{j!}(\Lambda \circ \mathrm{ad}_X^j Y)x^j f(x). \tag{13}$$

We define its degree $d_Y \in \mathbb{N}$ with respect to the representation $\pi_*(Y)$ to be the degree of the polynomial. Let $(d_1, \ldots, d_a)$ be the degrees of the elements $(Y_1, \ldots, Y_a)$, respectively. The degree of the representation π is defined as the maximum of the degrees of the elements of any basis.

3. *The cohomological equation*

In this section we prove an *a priori* Sobolev estimate on Green's operator for the cohomological equation $Xu = f$ of nilflows with generator X. Then we estimate the bounds of Green's operator by invariant distributions with rescaled Sobolev norm.

3.1. *Distributions and Sobolev space.*

Let $L^2(M)$ be the space of complex-valued, square integrable functions on M. Given an ordered basis $\mathcal{F}$ of $\mathfrak{n}$, the *transverse Laplace–Beltrami* operator is a second-order differential operator defined by

$$\Delta_{\mathcal{F}} = -\sum_{i=1}^{a} Y_i^2, \quad Y_i \in \mathfrak{I}.$$

For any $\sigma \geq 0$, let $|\cdot|_{\sigma,\mathcal{F}}$ be the transverse Sobolev norm defined as follows: for all functions $f \in C^\infty(M)$, let

$$|f|_{\sigma,\mathcal{F}} := \|(I + \Delta_{\mathcal{F}})^{(\sigma/2)} f\|_{L^2(M)}.$$

Equivalently,

$$|f|_{\sigma,\mathcal{F}} = \bigg(\|f\|_2^2 + \sum_{1\leq m\leq \sigma} \|Y_{j_1}\cdots Y_{j_m} f\|_2^2\bigg)^{1/2}, \quad Y_{j_m} \in \mathfrak{I}.$$

The completion of $C^\infty(M)$ with respect to the norm $|\cdot|_{\sigma,\mathcal{F}}$ is denoted $W^\sigma(M,\mathcal{F})$ and the distributional dual space (as a space of functional with values in H') to $W^\sigma(M)$ is denoted

$$W^{-\sigma}(M,\mathcal{F}) := (W^\sigma(M,\mathcal{F}))'.$$

We denote by $C^\infty(H_\pi)$ the space of C^∞ vectors of the irreducible unitary representation π. Following the notation in (11), let $W^\sigma(H_\pi) \subset H_\pi$ be the Sobolev space of vectors, endowed with the Hilbert space norm in the maximal domain of the essential self-adjoint

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

operator $(I + \pi_*(\Delta_{\mathcal{F}}))^{\sigma/2}$. That is to say, for every $f \in C^\infty(H_\pi)$ and $\sigma > 0$,

$$|f|_{\sigma,\mathcal{F}} := \bigg(\int_{\mathbb{R}} \|(1 + \pi_*(\Delta_{\mathcal{F}}))^{\sigma/2} f(x)\|^2_{H'}\, dx \bigg)^{1/2}$$

where $\pi_*(\Delta_{\mathcal{F}})$ is determined by derived representations.

3.2. *A priori estimates.* The distributional obstruction to the existence of solutions of the cohomological equation

$$Xu = f, \quad f \in C^\infty(H_\pi),$$

in an irreducible unitary representation H_π is the normalized X-invariant distribution.

Definition 3.1. For any $X \in \mathfrak{n}$, the space of X-invariant distributions for the representation π is the space $\mathcal{I}_X(H_\pi)$ of all distributional solutions $D \in D'(H_\pi)$ of the equation $\pi_*(X)D = XD = 0$. Let

$$\mathcal{I}^\sigma_X(H_\pi) := \mathcal{I}_X(H_\pi) \cap W^{-\sigma}(H_\pi)$$

be the subspace of invariant distributions of order at most $\sigma \in \mathbb{R}^+$.

By [FF07], each invariant distribution D has Sobolev order equal to $1/2$, that is, $D \in W^{-\sigma}(H_\pi)$ for any $\sigma > 1/2$.

For all $\sigma > 1$, let $\mathcal{K}^\sigma(H_\pi) = \{f \in W^\sigma(H_\pi) \mid D(f) = 0 \in C^\infty(H_\pi),$ for any $D \in W^{-\sigma}(H_\pi)\}$ be the kernel of the X-invariant distribution on the Sobolev space $W^\sigma(H_\pi)$. Green's operator $G_X : C^\infty(H_\pi) \to C^\infty(H_\pi)$ with

$$G_X f(t) = \int_{-\infty}^t f(s)\, ds$$

is well defined on the kernel of distribution $\mathcal{K}^\infty(H_\pi)$ on $C^\infty(H_\pi)$. In fact, for any $f \in \mathcal{K}^\infty(H_\pi)$, we have $\int_{\mathbb{R}} f(t)\, dt = 0 \in C^\infty(H')$ and

$$G_X f(t) = \int_{-\infty}^t f(s)\, ds = -\int_t^\infty f(s)\, ds \in C^\infty(\mathbb{R}, H').$$

Now we define the generalized (complex-valued) invariant distribution on smooth vector $C^\infty(H_\pi)$.

LEMMA 3.2. *The invariant distribution is generalized in the following sense. For every invariant distribution D, there exists a linear functional $\ell : C^\infty(H_\pi) \to \mathbb{C}$ such that for every function $f \in C^\infty(H_\pi) \subset L^2(\mathbb{R}, H')$,*

$$D(f) = \int_{\mathbb{R}} \ell(f(t))\, dt. \tag{14}$$

Furthermore, $\ell \in W^{-s}(H_\pi)$ for $s > 1/2$ and

$$\int_{\mathbb{R}} \ell(f(t))\, dt = \ell\bigg(\int_{\mathbb{R}} f(t)\, dt \bigg). \tag{15}$$

Proof. We construct a linear functional ℓ in (14) as follows. Let $\chi \in C_0^\infty(\mathbb{R})$ be a smooth function with compact support with unit integral over $\mathbb{R}$. Given an invariant distribution $D \in \mathcal{I}_X(H_\pi)$, let us define $\ell(v) = D(f_v)$ for $f_v = \chi v$ and $v \in C^\infty(H')$.

Firstly, we prove that ℓ is well defined. Let $\chi_1 \neq \chi_2 \in C_0^\infty(\mathbb{R})$ be functions with compact support such that $\int_{\mathbb{R}} \chi_1(t)\, dt = \int_{\mathbb{R}} \chi_2(t)\, dt = 1$. Note that there exists $\psi \in C_0^\infty(\mathbb{R})$ such that $\chi_1 - \chi_2 = \psi'$ with $\psi(t) = \int_{-\infty}^t (\chi_1(x) - \chi_2(x))\, dx$.

Then we have

$$\chi_1(t)v - \chi_2(t)v = \frac{d}{dt}(\psi(t)v) = \pi_*(X)(\psi(t)v) \in C^\infty(H_\pi),$$

and $\chi_1(t)v - \chi_2(t)v$ is an X-coboundary for every $v \in C^\infty(H')$. Hence, $D(\chi_1(t)v - \chi_2(t)v) = 0$, which implies that $D(\chi_1(t)v) = D(\chi_2(t)v)$. Therefore, $\ell(v)$ does not depend on the choice of χ and the functional ℓ is well defined.

Next, we verify that ℓ is a distribution on $C^\infty(H')$. It suffices to prove that ℓ is bounded and continuous. For $v \in C^\infty(H')$ and $s > 1/2$,

$$|\ell(v)| = |D(\chi(t)v)| \leq \|D\|_{-s} \|\chi(t)(v)\|_{W^s(H')}.$$

In the representation, $\pi_*(Y_i)$ acts as multiplication of the polynomial $p_i(t)$ on $L^2(\mathbb{R}, H')$. By definition of the Sobolev norm in representation, there exists a non-zero constant

$$C := C(\chi, p_1, \ldots, p_s) = \max_{\substack{t\in\mathbb{R}, j_1+\cdots j_d = s \\ 0\leq j_i \leq s, 1\leq i\leq d\leq s}} \{\chi(t) p_1^{j_1}(t) \cdots p_d^{j_d}(t)\}$$

such that

$$\|\chi(t)(v)\|_{W^s(H_\pi)} \leq C\|v\|_{W^s(H')}.$$

Hence, for $\ell \in W^{-s}(H')$,

$$\|\ell\|_{-s} := \sup_{\|v\|=1} \frac{|\ell(v)|}{\|v\|_{W^s(H')}} \leq C\|D\|_{-s}.$$

Therefore, ℓ is continuous on $C^\infty(H')$.

To prove equality (15), for any $f \in C^\infty(H_\pi)$, we observe $f(t) - \chi(t)(\int_{\mathbb{R}} f(x)\, dx)$ has zero average, hence it is a coboundary with smooth transfer function. Since distribution D is invariant under translation, we obtain

$$D\bigg(f - \chi(t)\bigg(\int_{\mathbb{R}} f(x)\, dx\bigg)\bigg) = 0$$

and

$$D(f) = D\bigg(\chi(t)\int_{\mathbb{R}} f(x)\, dx\bigg) = \ell\bigg(\int_{\mathbb{R}} f(t)\, dt\bigg).$$

Furthermore,

$$\begin{aligned}\int_{\mathbb{R}} \ell(f(t))\, dt &= \int_{\mathbb{R}} D(f_\chi(t))\, dt = \int_{\mathbb{R}} \bigg(\int_{\mathbb{R}} \chi(x) f(t)\, dx\bigg)\, dt \\ &= \int_{\mathbb{R}} f(t)\bigg(\int_{\mathbb{R}} \chi(x)\, dx\bigg)\, dt.\end{aligned}$$

Since D is an invariant distribution for translation and $\int_{\mathbb{R}} \chi = 1$, we conclude

$$\int_{\mathbb{R}} \ell(f(t))\, dt = \int_{\mathbb{R}} f(t)\, dt = D(f) = \ell\bigg(\int_{\mathbb{R}} f(t)\, dt\bigg). \qquad \square$$

Let $\mathcal{O}$ be any coadjoint orbit of maximal rank. For all $(X, Y) \in \mathfrak{n} \times \mathfrak{n}_{k-1}$ and $\Lambda \in \mathcal{O}$, the skew-symmetric bilinear form

$$B_\Lambda(X, Y) = \Lambda([X, Y])$$

does not depend on the choice of linear form $\Lambda \in \mathcal{O}$.

Let

$$\delta_{\mathcal{O}}(X, Y) := |B_\Lambda(X, Y)| \quad \text{for any } \Lambda \in \mathcal{O},$$
$$\delta_{\mathcal{O}}(X) := \max\{\delta_{\mathcal{O}}(X, Y) \mid Y \in \mathfrak{n}_{k-1} \text{ and } \|Y\| = 1\}. \qquad (16)$$

Here we recall estimates for Green's operator.

LEMMA 3.3. [FF07, Lemma 2.5] *Let $X \in \mathfrak{n}$ and $Y \in \mathfrak{n}_{k-1}$ be any operator such that $B_\Lambda(X, Y) \neq 0$. The derived representation π_* of the Lie algebra $\mathfrak{n}$ satisfies*

$$\pi_*(X) = \frac{d}{dx}, \quad \pi_*(Y) = 2\pi\iota B_\Lambda(X, Y)x Id_{H'} \text{ on } L^2(\mathbb{R}, H', dx). \qquad (17)$$

THEOREM 3.4. [FF07, Theorem 3.6] *Let $\delta_{\mathcal{O}} := \delta_{\mathcal{O}}(X) > 0$, and let π be an irreducible representation of $\mathfrak{n}$ on a Hilbert space H_π. If $f \in W^s(H_\pi)$, $s > 1$, and $D(f) = 0$ for all $D \in \mathcal{I}_X(H_\pi)$, then $G_X f \in W^r(H_\pi)$, for all $r < (s-1)/k$, and there exists a constant $C := C(X, k, r, s)$ such that*

$$|G_X f|_{r,\mathcal{F}} \leq C \max\{1, \delta_{\mathcal{O}}^{-(k-1)r-1)}\}|f|_{s,\mathcal{F}}.$$

3.3. *Rescaling method.*

Definition 3.5. The *deformation space* of a k-step nilmanifold M is the space $T(M)$ of all adapted bases of the Lie algebra $\mathfrak{n}$ of the group N.

The renormalization dynamics is defined as the action of the diagonal subgroup of the Lie group on the deformation space. Let $\rho := (\rho_1, \ldots, \rho_a) \in (\mathbb{R}^+)^a$ be any vector with rescaling condition $\sum_{i=1}^a \rho_i = 1$. Then there exists a one-parameter subgroup $\{A_t^\rho\}$ of the Lie group of $SL(a+1, \mathbb{R})$ defined as follows:

$$A_t^\rho(X, Y_1, \ldots, Y_a) = (e^t X, e^{-\rho_1 t} Y_1, \ldots, e^{-\rho_a t} Y_a). \qquad (18)$$

The renormalization group $\{A_t^\rho\}$ preserves the set of all adapted bases. However, it is not a group of automorphisms of the Lie algebra since generated elements in the center $Z(\mathfrak{n})$ are not preserved in higher steps. Therefore, on higher-step nilmanifolds, the dynamics induced by the renormalization group on the deformation space is trivial (it has no recurrent orbits).

Definition 3.6. Given any adapted basis $\mathcal{F} = (X, Y)$, the *rescaled basis* $\mathcal{F}(t) = (X(t), Y(t)) = \{e^t X, e^{-\rho_1 t} Y_1, \ldots, e^{-\rho_a t} Y_a\}$ of $\mathcal{F}$ is a basis of Lie algebra $\mathfrak{n}$ satisfying (18).

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

Let $(d_1, \ldots, d_i)$ be the degrees of the elements $(Y_1, \ldots, Y_i)$, respectively. For any $\rho = (\rho_1, \ldots, \rho_a) \in \mathbb{R}^a$, let

$$\lambda_{\mathcal{F}}(\rho) := \min_{i:d_i \neq 0} \left(\frac{\rho_i}{d_i} \right). \tag{19}$$

Definition 3.7. The scaling factor $\rho = (\rho_1, \ldots, \rho_a)$ is called *homogeneous* if the vector ρ is proportional to the vector $d = (d_1, \ldots, d_a)$ of degree of $(Y_1, \ldots, Y_a)$. That is, under homogeneous scaling, $\lambda_{\mathcal{F}}(\rho) = \rho_i/d_i$ for all $1 \leq i \leq a$.

For all $i = 1, \ldots, a$, denote

$$\Lambda_i^{(j)}(\mathcal{F}) := (\Lambda \circ \mathrm{ad}^j(X))(Y_i). \tag{20}$$

This is the coefficient appearing in (13). Set

$$|\Lambda(\mathcal{F})| := \sup_{(i,j):1 \leq i \leq a, 0 \leq j \leq d_i} \left| \frac{\Lambda_i^{(j)}(\mathcal{F})}{j!} \right|. \tag{21}$$

Let $\mathfrak{U}(\mathfrak{n})$ be the enveloping algebras of $\mathfrak{n}$. The generator δ is the derivation on $\mathfrak{U}(\mathfrak{n}')$ obtained by extending the derivation $\mathrm{ad}(X)$ of $\mathfrak{n}'$ to $\mathfrak{U}(\mathfrak{n}')$. From the nilpotency of $\mathfrak{n}$ it follows that for any $L \in \mathfrak{U}(\mathfrak{n}')$ there exists a first integer $[L]$ such that $\delta^{[L]+1}L = 0$.

Recall that $\mathfrak{I}$ is a codimension-1 ideal of $\mathfrak{n}$ used in §2.3.

LEMMA 3.8. *For each element $L \in \mathfrak{I}$ with degree $[L] = i$, there exists $Q_j \in \mathfrak{U}(\mathfrak{n})$ such that $\pi_*(L) = \sum_{j=0}^{i} (1/j!)\pi'_*(Q_j)x^j$ and $[Q_j] = [L] + 1 - j$.*

Proof. Firstly, we fix elements X and Y as stated in Lemma 3.3. For convenience, we normalize the constant of $\pi_*(Y)$ by 1. That is, there exists $X, Y \in \mathfrak{n}$ such that

$$\pi_*(X) = \frac{d}{dx}, \quad \pi_*(Y) = x.$$

We now replace the expansion of $\pi_*(L) := \sum_{j=0}^{i} \frac{1}{j!} \Lambda_L^{(j)}(\mathcal{F})x^j$ by choosing elements Q_i in the enveloping algebra $\mathfrak{U}(\mathfrak{n})$.

For the coefficient of top degree, denote $Q_i = (1/i!)\mathrm{ad}_X^i(L) \in \mathfrak{n}$. For degree $i - 1$, we set $Q_{i-1} = \mathrm{ad}_X^{i-1}(L) - Q_i Y \in \mathfrak{U}(\mathfrak{n})$ such that

$$\pi_*(Q_{i-1}) = \pi_*(\mathrm{ad}_X^{i-1}(L)) - \pi_*(Q_i)\pi_*(Y) = \Lambda_L^{(i-1)}(\mathcal{F}).$$

Repeating this process up to degree 0, there exists $Q_l \in \mathfrak{U}(\mathfrak{n})$ such that for $0 < l < i$,

$$\pi_*(Q_l) = \pi_*(\mathrm{ad}_X^l(L)) - \frac{1}{l!} \sum_{j=l+1}^{i} \pi_*(Q_j)\pi_*(Y)^{j-l}$$

and

$$\pi_*(Q_0) = \pi_*(L) - \frac{1}{l!} \sum_{l=1}^{i} \pi_*(Q_l)\pi_*(Y)^l = \Lambda_L^{(j)}(\mathcal{F}). \qquad \square$$

Definition 3.9. If A is self-adjoint on a Hilbert space H and $\langle Au, u\rangle \geq 0$ for every $u \in H$, then A is called *positive*, denoted by $A \geq 0$.

Remark 3.10. For two self-adjoint operators A and B, $A \geq B$ if and only if $A - B \geq 0$. Suppose that A and B are bounded operators and commute. Then $A \geq 0$ and $B \geq 0$ implies that $AB \geq 0$. By the spectral theorem, there exists a unique, self-adjoint square root $A^{1/2} = \int_{\sigma(A)} x^{1/2} dE_A(x)$. Since A and B commute, $A^{1/2}$ commutes with B. Then

$$\langle ABu, u\rangle = \langle A^{1/2}Bu, A^{1/2}u\rangle = \langle BA^{1/2}u, A^{1/2}u\rangle \geq 0.$$

Also, if $0 \leq A \leq B$, then $A^2 \leq B^2$ ($B^2 - A^2 = B(B - A) + (B - A)A$).

Recall that for any positive operators A and B,

$$AB + BA \leq A^2 + B^2 \text{ and } (A + B)^2 \leq 2(A^2 + B^2). \tag{22}$$

LEMMA 3.11. *For any $r \geq 1$ and $a \geq 1$, there exists constant $C(a, r) > 0$ such that*

$$\|\pi_*(\Delta(t)^{2r})u\| \leq C(a, r)\bigg\| \sum_{i=1}^{a} \pi_*(Y_i(t)^{4r})u\bigg\|. \tag{23}$$

Proof. It suffices to prove that there exists $C = C(a, r) > 0$ such that

$$\Delta(t)^{2r} \leq C\bigg(\sum_{i=1}^{a} Y_i(t)^{4r}\bigg) \tag{24}$$

since this implies $\Delta(t)^{4r} \leq C^2(\sum_{i=1}^{a} Y_i(t)^{4r})^2$ by the remark.

We prove (24) by induction. If $r = 1$, then

$$\Delta(t)^2 = \bigg(\sum_{i=1}^{a} Y_i(t)^2\bigg)^2 = \bigg(\sum_{i=1}^{a} Y_i(t)^4 + \sum_{i\neq j} Y_i(t)^2 Y_j(t)^2\bigg).$$

By (22), for each i and j,

$$Y_i(t)^2 Y_j(t)^2 + Y_j(t)^2 Y_i(t)^2 \leq Y_i(t)^4 + Y_j(t)^4.$$

Then there exists $C_0 = (a + 1)$ such that

$$\Delta(t)^2 \leq C_0\bigg(\sum_{i=1}^{a} Y_i(t)^4\bigg). \tag{25}$$

Assume that the statement holds for large r. Then, since $\Delta(t)^2$ and $\sum_{i=1}^{a} Y_i(t)^4$ are positive, by (25), there exists $C_1(a, r)$ such that

$$\begin{aligned}\Delta(t)^{2(r+1)} &\leq C_1(a, r)\bigg(\sum_{i=1}^{a} Y_i(t)^{4r}\bigg)(\Delta(t)^2)\\ &\leq C_1(a, r)(a + 1)\bigg(\sum_{i=1}^{a} Y_i(t)^{4r}\bigg)\bigg(\sum_{i=1}^{a} Y_i(t)^4\bigg).\end{aligned}$$

Note that the following inequality holds: for any $r \geq 1$,

$$Y_i(t)^{4r} Y_j(t)^4 + Y_i(t)^4 Y_j(t)^{4r} \leq Y_i(t)^{4(r+1)} + Y_j(t)^{4(r+1)}. \tag{26}$$

This inequality is proved by showing that

$$(Y_i(t)^4 - Y_j(t)^4)(Y_i(t)^{4r} - Y_j(t)^{4r})$$
$$= (Y_i(t)^4 - Y_j(t)^4)^2\bigg(\sum_{l=0}^{r-1}(Y_i(t)^4)^l(Y_j(t)^4)^{r-1-l}\bigg) \geq 0.$$

Since $Y_i(t)^4$, $Y_j(t)^4$ are all positive, the last inequality holds.

Then, by (26),

$$\bigg(\sum_{i=1}^{a} Y_i(t)^{4r}\bigg)\bigg(\sum_{i=1}^{a} Y_i(t)^4\bigg)$$
$$= \bigg(\sum_{i=1}^{a} Y_i(t)^{4r+4} + \sum_{i<j}^{a}(Y_i(t)^{4r}Y_j(t)^4 + Y_j(t)^{4r}Y_i(t)^4)\bigg)$$
$$\leq (a+1)\bigg(\sum_{i=1}^{a} Y_i(t)^{4r+4}\bigg).$$

Setting $C_2(a, r) = C_1(a, r)(a + 1)^2$,

$$\Delta(t)^{4(r+1)} \leq C_2(a, r)\bigg(\sum_{i=1}^{a} Y_i(t)^{4(r+1)}\bigg).$$

Therefore, induction holds and we finish the proof. □

For the cohomological equation $X(t)u = f$, denote its Green's operator $G_{X(t)}$. The following theorem states an estimate for the rescaled version of Theorem 3.4.

THEOREM 3.12. *For $r > 1$, let $s > r(k + 1) + 1/4$. For any $f \in \mathcal{K}^s(M)$, there exists $C_{r,k,s} > 0$ such that the following holds: for any $t > 0$,*

$$|G_{X(t)} f|_{r,\mathcal{F}(t)} \leq C_{r,k,s}e^{-(1-\lambda_{\mathcal{F}})t} \max\{1, \delta_O^{-4r(k+1)}\}|f|_{s,\mathcal{F}(t)}.$$

Proof. First, we estimate the bound on Green's operator with the Sobolev norm.

By Lemma 3.3, there exists a rescaled operator $Y(t) \in \mathfrak{n}_{k-1}$ with

$$\pi_*(Y(t)) = 2\pi\iota\delta_O(t)x\mathrm{Id}_{H'}$$

where $\delta_O(t) = \delta_O e^{-\rho_Y t} > 0$.

By the Cauchy–Schwarz inequality, for any $l \in \mathbb{N}$,

$$\|\pi_*(Y(t))^l G_{X(t)} f\|^2$$
$$\leq \int_0^\infty \bigg(|2\pi\delta_O(t)x|^l \int_x^\infty e^{-t}\|f(s)\|_{H'}\, ds\bigg)^2 dx$$
$$+ \int_{-\infty}^0 \bigg(|2\pi\delta_O(t)x|^l \int_{-\infty}^x e^{-t}\|f(s)\|_{H'}\, ds\bigg)^2 dx$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

$$\leq \int_0^\infty |2\pi\delta_O(t)x|^{2l} \int_x^\infty e^{-2t}\|f(s)\|_{H'}^2 \, ds dx$$
$$+ \int_{-\infty}^0 |2\pi\delta_O(t)x|^{2l} \int_{-\infty}^x e^{-2t}\|f(s)\|_{H'}^2 \, ds dx.$$

If $\alpha > 1$, then by Hölder's inequality,

$$\int_{-\infty}^x \|f(s)\|_{H'}^2 \, ds \leq \bigg(\int_{-\infty}^x (1 + 4\pi^2\delta_O(t)^2 s^2)^{-\alpha} \, ds \bigg) \|(I - Y(t)^2)^{\alpha/2} f\|^2. \tag{27}$$

For all $\alpha > 1$, we set

$$C_{\alpha,l}^2 = \int_0^\infty (2\pi x)^{2l} \bigg(\int_x^\infty (1 + (4\pi^2 s^2))^{-(l+\alpha)} \, ds \bigg) \, dx$$
$$+ \int_{-\infty}^0 (2\pi x)^{2l} \bigg(\int_{-\infty}^x (1 + (4\pi^2 s^2))^{-(l+\alpha)} \, ds \bigg) \, dx < \infty. \tag{28}$$

By Hölder's inequality and change of variables for $x' = \delta_O(t)x$ and $s' = \delta_O(t)s$,

$$\begin{aligned} \|\pi_*(Y(t))^l G_{X(t)} f\| &\leq C_{\alpha,l} \bigg(\frac{e^{-t}}{\delta_O(t)} \bigg) \|\pi_*(I - Y(t)^2)^{(l+\alpha)/2} f\| \\ &= C_{\alpha,l} e^{-(1-\rho_Y)t} \bigg(\frac{1}{\delta_O} \bigg) \|\pi_*(I - Y(t)^2)^{(l+\alpha)/2} f\|. \end{aligned} \tag{29}$$

Let $E_x : C^\infty(H_\pi) \to C^\infty(H')$ be the linear operator defined by $E_x f = f(x)$ (cf. Lemma 2.5). Then the action of $\pi_*(Y(t))$ on $C^\infty(H_\pi)$ can be rewritten as

$$E_x \pi_*(Y(t)) = (2\pi\iota x \delta_O(t))^j E_x \quad \text{for all } j \in \mathbb{N} \text{ and } x \in \mathbb{R}. \tag{30}$$

Let $L(t)$ be a rescaled element of $L \in \mathcal{F}$. By definition of representation (13) and by Lemma 3.8 for rescaled basis $\mathcal{F}(t)$, for any $L(t)$ there exists $Q_j(t) \in \mathfrak{U}(\mathfrak{n})$ such that

$$E_x \pi_*(L(t)) = \sum_{j=0}^{[L]} \frac{(2\pi\iota x \delta_O(t))^j}{j!} \pi'_*(Q_j(t)) E_x. \tag{31}$$

For Green's operator $G_{X(t)}$, since $E_x G_{X(t)} = \int_{-\infty}^x E_s \, ds$ for all $x \in \mathbb{R}$,

$$\begin{aligned} E_x \pi_*(L(t)) G_{X(t)} &= \sum_{j=0}^{[L]} \frac{(2\pi\iota x \delta_O(t))^j}{j!} \pi'_*(Q_j(t)) E_x G_{X(t)} \\ &= \sum_{j=0}^{[L]} \frac{(2\pi\iota x \delta_O(t))^j}{j!} \int_{-\infty}^x \pi'_*(Q_j(t)) E_s \, ds \\ &= \sum_{j=0}^{[L]} \frac{(2\pi\iota x \delta_O(t))^j}{j!} \int_{-\infty}^x E_s \pi_*(Q_j(t)) \, ds \end{aligned}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

$$= \sum_{j=0}^{[L]} \frac{(2\pi \iota x \delta_O(t))^j}{j!} E_x G_{X(t)}(\pi_*(Q_j(t)))$$
$$= E_x \sum_{j=0}^{[L]} \frac{1}{j!} \pi_*(Y(t))^j G_{X(t)}(\pi_*(Q_j(t))). \tag{32}$$

Combining (29) and (32), there exists a constant $C'_{\alpha,j} > 0$ such that

$$\|\pi_*(L(t))G_{X(t)}(f)\| \leq \sum_{j=0}^{[L]} \frac{1}{j!} \|\pi_*(Y(t))^j G_{X(t)}(\pi_*(Q_j(t))f)\|$$
$$\leq \sum_{j=0}^{[L]} C'_{\alpha,j} e^{-(1-\rho_Y)t} \bigg(\frac{1}{\delta_O}\bigg)^{j+1} \|\pi_*(I - Y(t)^2)^{((j+\alpha)/2)} \pi_*(Q_j(t)f)\|.$$

Note that by binomial formula, there exists $R_j(t) \in \mathfrak{U}(\mathfrak{n})$ such that

$$\pi_*(L(t))^{2r} = \bigg(\sum_{j=0}^{[L]} \frac{1}{j!} \pi_*(Y(t)^j) \pi_*(Q_j(t))\bigg)^{2r} := \sum_{j=0}^{2r[L]} \pi_*(Y(t)^j) \pi_*(R_j(t)). \tag{33}$$

In particular, $R_j(t)$ is product of $Q_i(t)$s and $[R_j(t)] = 2r([L]+1) - j$. Specifically, to compute the transverse Laplacian, here we assume that $L(t) = (Y_i(t))^{4r}$ for each element $Y_i \in \mathcal{F}$. Then, by (33), for any $\alpha > 1$,

$$\|\pi_*(Y_i(t)^{4r})G_{X(t)}(f)\|$$
$$\leq \sum_{j=0}^{4r[Y_i]} C(\alpha, i, j, r) e^{-(1-\rho_Y)t} \bigg(\frac{1}{\delta_O}\bigg)^{j+1} \|\pi_*(I - Y(t)^2)^{((j+\alpha)/2)} \pi_*(R_j(t)) f\|$$
$$\leq C(\alpha, i, r) e^{-(1-\rho_Y)t} \max\{1, {\delta_O}^{-4r[Y_i]+1}\} |f|_{\alpha+4r([Y_i]+1),\mathcal{F}(t)}. \tag{34}$$

Therefore, combining with Lemma 3.11,

$$\|\pi_*(\Delta(t)^{2r})G_{X(t)}(f)\|$$
$$\leq C(a, r) \sum_{i=1}^{a} \|\pi_*(Y_i(t)^{4r})G_{X(t)}(f)\|$$
$$\leq C(a, r, \alpha) \sum_{i=1}^{a} e^{-(1-\rho_Y)t} \max\{1, {\delta_O}^{-4r([Y_i]+1)}\} |f|_{\alpha+4r([Y_i]+1),\mathcal{F}(t)}$$
$$\leq C(a, r, \alpha) e^{-(1-\rho_Y)t} \max\{1, {\delta_O}^{-4r(k+1)}\} |f|_{\alpha+4r(k+1),\mathcal{F}(t)}.$$

Then, there exists $C' = C'(a, \alpha, r, X) > 0$ such that

$$|G_{X(t)} f|_{4r,\mathcal{F}(t)} \leq C' e^{-(1-\rho_Y)t} \max\{1, \delta_O^{-4r(k+1)}\} |f|_{\alpha+4(k+1)r,\mathcal{F}(t)}.$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

By interpolation, for all $s > r(k+1) + 1/4$, there exists a constant $C_{r,s} := C_{r,s}(k, X) > 0$ such that

$$|G_{X(t)} f|_{r,\mathcal{F}(t)} \leq C_{r,s} e^{-(1-\rho_Y)t} \max\{1, \delta_O^{-4r(k+1)}\}|f|_{s,\mathcal{F}(t)}.$$

By the choice of Y, we obtain $\rho_Y = \lambda_{\mathcal{F}}$, which finishes the proof. □

3.4. *Scaling of invariant distributions.* We introduce the Lyapunov norm and compare bounds between the Sobolev dual norm and the Sobolev Lyapunov norm of invariant distribution in every irreducible, unitary representation.

For all $t \in \mathbb{R}$ and $\lambda := \lambda_{\mathcal{F}}(\rho)$ defined in (19), let the operator $U_t : L^2(\mathbb{R}, H') \to L^2(\mathbb{R}, H')$ be the unitary operator defined as follows:

$$(U_t f)(x) = e^{-(\lambda/2)t} f(e^{-\lambda t} x). \tag{35}$$

We will compare the norm estimate of invariant distributions with respect to scaled basis

$$|D|_{-r,\mathcal{F}(t)} = \sup_{f \in W^r(H_\pi)} \{|D(f)| : |f|_{r,\mathcal{F}(t)} = 1\}$$

by the unscaled norm $|D|_{-r,\mathcal{F}}$.

THEOREM 3.13. *For $r \geq 1$ and $s > r(k+1)$, there exists a constant $C_{r,s} > 0$ such that for all $t > 0$, the following bound holds:*

$$|U_t f|_{r,\mathcal{F}(t)} \leq C_{r,s} |f|_{s,\mathcal{F}}.$$

Proof. Assume the same hypothesis for $L \in \mathcal{F}$ and $L(t)$ in the proof of Theorem 3.12. By Lemma 3.8, there exists an $(i - j + 1)$th-order $Q_j \in \mathfrak{U}(\mathfrak{n})$ with

$$U_t^{-1} L(t) U_t = x^i Q_i + e^{-\lambda t} x^{i-1} Q_{i-1} + e^{-2\lambda t} x^{i-2} Q_{i-2} + \cdots + e^{-i\lambda t} Q_o.$$

Then there exists $C > 0$ such that

$$\begin{aligned} \|U_t^{-1} L(t) U_t f\| &\leq \sum_{j=0}^{i} \|e^{-(i-j)\lambda t} x^j Q_j f\| \\ &\leq C \max_{0 \leq j \leq i} \|Y^j Q_j f\| \\ &\leq C |f|_{[L]+1,\mathcal{F}}. \end{aligned}$$

Since $[L] \leq k$, by unitarity

$$|U_t f|_{1,\mathcal{F}(t)} \leq C_1 |f|_{k+1,\mathcal{F}}. \tag{36}$$

Hence, for any $s > r(k+1)$,

$$|U_t f|_{r,\mathcal{F}(t)} \leq C_{r,s} |f|_{s,\mathcal{F}}. \tag{37}$$

□

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

THEOREM 3.14. *For $r \geq 1$ and $s > r(k+1)$, there exists a constant $C_{r,s} > 0$ such that for any $\lambda > 0$ and $t > 0$, the invariant distribution defined in (14) satisfies*

$$|D|_{-s,\mathcal{F}} \leq C_{r,s} e^{-(\lambda/2)t} |D|_{-r,\mathcal{F}(t)}.$$

Proof. Recall the functional ℓ defined in the Lemma 3.2. For $f \in C^\infty(H_\pi)$,

$$\begin{aligned} D(U_t f) &= \int_{\mathbb{R}} \ell(e^{-(\lambda/2)t} f(e^{-\lambda t} x))\, dx \\ &= \int_{\mathbb{R}} e^{-(\lambda/2)t} \ell(f(e^{-\lambda t} x))\, dx \\ &= e^{(\lambda/2)t} \int_{\mathbb{R}} \ell(f(y))\, dy \\ &= e^{(\lambda/2)t} D(f). \end{aligned}$$

Then, by unitarity (37),

$$|D|_{-r,\mathcal{F}(t)} = \sup_{f\neq 0} \frac{|D(f)|}{|f|_{r,\mathcal{F}(t)}} = \sup_{f\neq 0} \frac{|D(U_t f)|}{|U_t f|_{r,\mathcal{F}(t)}} \geq \sup_{f\neq 0} \frac{e^{(\lambda/2)t}|D(f)|}{C_{r,s}|f|_{s,\mathcal{F}}} = C_{r,s}^{-1} e^{(\lambda/2)t} |D|_{-s,\mathcal{F}},$$

and

$$|D|_{-s,\mathcal{F}} \leq C_{r,s} e^{-(\lambda/2)t} |D|_{-r,\mathcal{F}(t)}.$$

□

Definition 3.15. (Lyapunov norm) For any basis $\mathcal{F}$ and all $\sigma > 1/2$, define the Lyapunov norm

$$\|D\|_{-\sigma,\mathcal{F}} := \inf_{\tau \geq 0} e^{-(\lambda_{\mathcal{F}}(\rho)/2)\tau} |D|_{-\sigma,\mathcal{F}(\tau)}. \tag{38}$$

The following lemma is immediate from the definition of the norm.

LEMMA 3.16. [**FF14**, Lemma 4.15] *For all $t \geq 0$ and $\sigma > 1/2$, we have*

$$\|D\|_{-\sigma,\mathcal{F}} \leq e^{-(\lambda_{\mathcal{F}}(\rho)/2)t} \|D\|_{-\sigma,\mathcal{F}(t)}.$$

We conclude this section by introducing a useful inequality that follows from the Theorem 3.14:

$$C_{r,s}^{-1} |D|_{-s,\mathcal{F}} \leq \|D\|_{-r,\mathcal{F}} \leq |D|_{-r,\mathcal{F}}. \tag{39}$$

4. *A Sobolev trace theorem*

In this section the notion of the average width (40), which is an average measure of close returns along an orbit, is introduced and we prove a Sobolev trace theorem for nilpotent orbits. According to this theorem, the uniform norm of an ergodic integral is bounded in terms of the average width of the orbit segment times the transverse Sobolev norms of the function, with respect to a given basis of the Lie algebra.

4.1. *Sobolev a priori bounds.*

Assume $\mathcal{F}(t) = (X(t), Y(t))$ is a rescaled basis. For any $x \in M$, let $\phi_{x,t} : \mathbb{R} \times \mathbb{R}^a \to M$ be the local embedding defined by

$$\phi_{x,t}(\tau, \mathbf{s}) = x \exp(\tau X(t)) \prod_{i=1}^{a} \exp(s_i Y_i(t)), \quad \mathbf{s} = (s_i)_{i=1}^a.$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

LEMMA 4.1. *For any $x \in M$, $t \geq 0$, and $f \in C^\infty(M)$, we have*

$$\partial_{s_i} f \circ \phi_{x,t}(\tau, \mathbf{s}) = S_i f \circ \phi_{x,t}(\tau, \mathbf{s}), \qquad S_i = Y_i(t) + \sum_{l>i}^{a} q_l(\mathbf{s}, t) Y_l(t) \in \mathfrak{n},$$

where q is polynomial in $\mathbf{s}$ of degree at most $k-1$ and $|q_l(\mathbf{s}, t)| \leq |q_l(\mathbf{s}, 0)|$ for all $t \geq 0$.

Proof. For $h \in \mathbb{R}$ and $1 \leq i \leq a$, let $\mathbf{s} + h_i$ denote a sequence with $(\mathbf{s} + h_i)_i = s_i + h$ and $(\mathbf{s} + h_i)_j = s_j$ if $i \neq j$. By definition,

$$\partial_{s_i} f \circ \phi_{x,t}(\tau, \mathbf{s}) = \lim_{h \to 0} \frac{f \circ \phi_{x,t}(\tau, \mathbf{s} + h_i) - f \circ \phi_{x,t}(\tau, \mathbf{s})}{h},$$

and we plan to rewrite $f \circ \phi_{x,t}(\tau, \mathbf{s} + h_i)$ in a suitable manner for differentiation.

Fix i. Then for $j > i$,

$$\begin{aligned}
&\exp((s_i + h) Y_i(t)) \exp(s_j Y_j(t)) \\
&\quad = \exp(s_i Y_i(t)) \exp(h Y_i(t)) \exp(s_j Y_j(t)) \\
&\quad = \exp(s_i Y_i(t)) \exp(e^{ad(hY_i(t))} s_j Y_j(t)) \exp(h Y_i(t)) \\
&\quad = \exp(s_i Y_i(t)) \exp(s_j Y_j(t)) \exp\left(\sum_{n=1}^{\infty} \frac{1}{n!} ad^n_{hY_i(t)} s_j Y_j(t)\right) \exp(h Y_i(t)).
\end{aligned}$$

By the Baker–Campbell–Hausdorff formula, we set

$$\begin{aligned}
&\exp((s_i + h) Y_i(t)) \exp(s_j Y_j(t)) \\
&\quad = \exp(s_i Y_i(t)) \exp(s_j Y_j(t)) \exp\Big(h(Y_i(t) + [Y_i(t), s_j Y_j(t)]) + O(h^2) \sum_{m>j} Y_m\Big).
\end{aligned}$$

Choose $j = i + 1$ and observe that all the terms of h are combined into an exponential. Iteratively, we will repeat this process from $j = i + 1$ to a until all the terms of h are pushed back. That is, we conclude

$$\begin{aligned}
\phi_{x,t}(\tau, \mathbf{s} + h_i) = \phi_{x,t}(\tau, \mathbf{s}) \exp(h(&Y_i(t) + [Y_i(t), s_{i+1} Y_{i+1}(t)] \\
&+ [[Y_i(t), s_{i+1} Y_{i+1}(t)], s_{i+2} Y_{i+2}(t)] \\
&+ \cdots + [\cdots [Y_i(t), s_{i+1} Y_{i+1}(t)], \cdots], s_a Y_a(t)]]) + O(h^2) \sum_{m>i+1}^{a} Y_m.
\end{aligned}$$

For convenience, we write coefficient function $q_l(\mathbf{s}, t)$ in polynomial degree at most k for $\mathbf{s}$ such that

$$\phi_{x,t}(\tau, \mathbf{s} + h_i) = \phi_{x,t}(\tau, \mathbf{s}) \exp\left(h\left(Y_i(t) + \sum_{l>i}^{a} q_l(\mathbf{s}, t) Y_l(t)\right) + O(h^2) \sum_{m>i+1}^{a} Y_m\right).$$

We conclude the proof by choosing $S_i = Y_i(t) + \sum_{l>i}^{a} q_l(\mathbf{s}, t) Y_l(t)$. Also, commutation between rescaled elements yields $[Y_i(t), s_j Y_j(t)] = s_j e^{-\rho_{ij} t} Y_k(t)$ for some $\rho_{ij} > 0$. This implies that the term $q_l(\mathbf{s}, t)$ for each $l > i$ includes exponential terms with negative exponent so that it decreases for $t \geq 0$. □

Let $\triangle_{\mathbb{R}^a}$ be the Laplacian operator on $\mathbb{R}^a$ given by

$$\triangle_{\mathbb{R}^a} = -\sum_{i=1}^{a} \frac{\partial^2}{\partial s_i^2}.$$

Given an open set $O \subset \mathbb{R}^a$ containing the origin, let $\mathcal{R}_O$ be the family of all a-dimensional symmetric rectangles $R \subset [-(1/2), 1/2]^a \cap O$ that are centered at the origin. The *inner width* of the set $O \subset \mathbb{R}^a$ is the positive number

$$w(O) = \sup\{\mathrm{Leb}(R) \mid R \in \mathcal{R}_O\},$$

where Leb is Lebesgue measure on R. The width function of a set $\Omega \subset \mathbb{R} \times \mathbb{R}^a$ containing the line $\mathbb{R} \times \{0\}$ is the function $w_\Omega : \mathbb{R} \to [0, 1]$ defined as follows:

$$w_\Omega(\tau) := w(\{\mathbf{s} \in \mathbb{R}^a \mid (\tau, \mathbf{s}) \in \Omega\}), \quad \text{for all } \tau \in \mathbb{R}.$$

Definition 4.2. Consider the family $O_{x,t,T}$ of open sets $\Omega \subset \mathbb{R} \times \mathbb{R}^a$ satisfying the two conditions

$$[0, T] \times \{0\} \subset \Omega \subset \mathbb{R} \times [-\tfrac{1}{2}, \tfrac{1}{2}]^a$$

and $\phi_{x,t}$ is injective on the open set $\Omega \subset \mathbb{R}^a$. The *average width* of the orbit segment of the rescaled nilflow $\{\phi_{x,t}(\tau, 0) \mid 0 \le t \le T\}$,

$$w_{\mathcal{F}(t)}(x, T) := \sup_{\Omega \in O_{x,t,T}} \left(\frac{1}{T} \int_0^T \frac{ds}{w_\Omega(s)} \right)^{-1}, \tag{40}$$

is a positive number.

The following lemma is derived from the standard Sobolev embedding theorem under the rescaling argument.

LEMMA 4.3. [**FF14**, Lemma 3.7] *Let $I \subset \mathbb{R}$ be an interval, and let $\Omega \subset \mathbb{R} \times \mathbb{R}^a$ be a Borel set containing the segment $I \times \{0\} \subset \mathbb{R} \times \mathbb{R}^a$. For every $\sigma > a/2$, there is a constant $C_s > 0$ such that for all functions $F \in C^\infty(\Omega)$ and all $\tau \in I$, we have*

$$\left(\int_I |F(\tau, 0)|\, d\tau \right)^2 \le C_\sigma \left(\int_I \frac{d\tau}{w_\Omega(\tau)} \right) \int_\Omega |(I - \triangle_{\mathbb{R}^a})^{\sigma/2} F(\tau, s)|\, d\tau ds.$$

The following theorem indicates the bound of the ergodic average of the scaled nilflow $\phi^\tau_{X(t)}$ with width function on general nilmanifolds (see also [**FFT16**, Theorem 5.2] for twisted horocycle flows).

THEOREM 4.4. *For all $\sigma > a/2$, there is a constant $C_\sigma > 0$ such that*

$$\left| \frac{1}{T} \int_0^T f \circ \phi^\tau_{X(t)}(x)\, d\tau \right| \le C_\sigma T^{-1/2} w_{\mathcal{F}(t)}(x, T)^{-1/2} |f|_{\sigma, \mathcal{F}(t)}.$$

Proof. Recall that for any self-adjoint operators A and B,

$$(A + B)^2 \le 2(A^2 + B^2).$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

Since $|s_i| \leq \frac{1}{2}$ and $t \geq 1$, by Lemma 4.1, each polynomial q_j is bounded in $\mathbf{s}$ and t. Then, essentially by skew-adjointness of $Y_i(t)$, there exists a large constant $C > 1$ with

$$-S_i^2 = -\left(Y_i(t) + \sum_{l>i}^{a} q_l(\mathbf{s}, t) Y_l(t)\right)^2$$
$$\leq -C \sum_{j=i}^{a} Y_j(t)^2.$$

Since operators on both sides are essentially self-adjoint,

$$\left(I - \sum_{i=1}^{a} S_i^2\right)^{\sigma/2} \leq C^{(\sigma/2)}\left(I - \sum_{i=1}^{a} Y_i(t)^2\right)^{\sigma/2}.$$

Thus, there is a constant $C_\sigma > 0$ such that

$$\|(I - \triangle_{\mathbb{R}^a})^{\sigma/2} f \circ \phi_{x,t}\|^2_{L^2(\Omega)} \leq C_\sigma \|(I - \Delta_{\mathcal{F}(t)})^{\sigma/2} f\|^2_{L^2(M)}. \tag{41}$$

By Lemma 4.3, we can see that for $\sigma > a/2$, setting $F(\tau, 0) = f \circ \phi^\tau_{X(t)}(x)$,

$$\left|\frac{1}{T}\int_0^T f \circ \phi^\tau_{X(t)}(x)\, dt\right|^2 = \left(\frac{1}{T}\int_0^T |F(\tau, 0)\, d\tau|^2\right)^2$$
$$\leq C_\sigma \frac{1}{T}\left(\frac{1}{T}\int_0^T \frac{ds}{w_\Omega(s)}\right)\int_\Omega |(I - \triangle_{\mathbb{R}^a})^{\sigma/2} F(\tau, \mathbf{s})|\, d\tau d\mathbf{s}$$
$$\leq C_\sigma T^{-1} w_{\mathcal{F}(t)}(x, T)^{-1} \|(I - \Delta_{\mathcal{F}(t)})^{\sigma/2} f\|^2_{L^2(M)}. \qquad \square$$

5. *Average width estimate*

This section is devoted to the proof of estimates on the averaged width of orbits of nilflows. Compared with the quasi-abelian case (see [**FF14**, Lemma 2.4]), there are no explicit expressions for the return map on transverse higher-step nilmanifolds. Instead, we calculate the differential of displacement and estimate the measure of close return orbits with respect to the rescaled vector fields.

Strategy.

(1) In §5.1 we introduce basic settings. Since the flow commutes with the centralizer, we take the quotient map to obtain a local diffeomorphism. It is remarkable to see that we set a tubular neighborhood to consider close return orbits on the quotient space. This calculates the measure of the set of close return orbits so-called almost periodic set. (See (54) and Lemma 5.4.)

(2) The range of the differential of the displacement map coincides with the range of the adjoint map ad_{X_α}. This is one important reason why a transverse condition is needed. Without this condition, there could be a direction that the return orbit that does not reach on the transverse manifold, which contradicts the idea that the set of almost periodic points should have a small measure up to rescaling vector ρ. (See Lemma 5.8.)

(3) In §5.2 we prove a bound on the average width by ergodic averages of cut-off functions. By Definition 5.9, we classify the type of close return orbits by growth

of local coordinates. The width of the function does not vanish on such a set and it is injective under the restricted domain. (See Lemmas 5.10 and 5.11.)

(4) Finally, in §5.3 and 5.4, we follow the known estimates from [**FF14**], which are necessary for proving bounds of ergodic averages in §6. In particular, the definition of a good point (Definition 5.21) means the set of points whose the width along the direction transverse to the flow cannot be too small and we prove the complement of the set of good points has a small measure. (See Lemma 5.22.)

5.1. *Almost periodic points.* Let X_α be the vector field on M defined in (7). Recall formula (8):

$$X_\alpha := \xi + \sum_{(i,j)\in J} \alpha_i^{(j)} \eta_i^{(j)}.$$

Let us introduce special type of condition for the Lie algebra $\mathfrak{n}$ required for width estimates.

Definition 5.1. The nilpotent Lie algebra $\mathfrak{n}$ satisfies the *transversality condition* if there exists a basis (X_α, Y) of $\mathfrak{n}$ such that

$$\langle \mathfrak{G}_\alpha \rangle + \mathrm{Ran}(\mathrm{ad}_{X_\alpha}) + C_{\mathfrak{I}}(X_\alpha) = \mathfrak{n} \tag{42}$$

where $\mathfrak{G}_\alpha = (X_\alpha, Y_i^{(1)})_{1\le i\le n}$ is a set of generator, $\mathrm{Ran}(\mathrm{ad}_{X_\alpha}) = \{Y \in \mathfrak{I} \mid Y = \mathrm{ad}_{X_\alpha}(W), W \in \mathfrak{I}\}$ and $C_{\mathfrak{I}}(X_\alpha) = \{Y \in \mathfrak{I} \mid [Y, X_\alpha] = 0\}$ is the centralizer.

It is clear that the set of generators is included neither in the range of ad_{X_α} nor in the centralizer $C_{\mathfrak{I}}(X_\alpha)$. We will restrict $\mathfrak{n}$ to satisfying the condition (42) in the rest of sections.

Remark 5.2. The transversality condition implies that the displacement (or distance between x and $\Phi^r_{\alpha,\theta}(x)$), induced by return map $\Phi_{\alpha,\theta}$, should intersect the set of centralizers transversally. That is, the measure of the set of close return orbits in the transverse manifold M^a_θ should not be invariant under the action of flow. This condition is crucial in estimating the almost periodic orbit (54) under rescaling of basis in Lemma 5.8.

Recall that M^a_θ denotes the fiber at $\theta \in \mathbb{T}^1$ of the fibration $pr_2 : M \to \mathbb{T}^1$. $\Phi_{\alpha,\theta}$ denotes the first return map of the nilflow $\{\phi^t_{X_\alpha}\}$ to the transverse section M^a_θ, and $\Phi^r_{\alpha,\theta}$ denotes the rth iterate of the map $\Phi_{\alpha,\theta}$. Let G denote the nilpotent Lie group with its lattice Γ defining $M^a_\theta = \Gamma\backslash G$. G acts on M^a_θ by right action and the action of G extends to $M^a_\theta \times M^a_\theta$.

Define a map $\psi^{(r)}_{\alpha,\theta} : M^a_\theta \to M^a_\theta \times M^a_\theta$ given by $\psi^{(r)}_{\alpha,\theta}(x) = (x, \Phi^r_{\alpha,\theta}(x))$. By its definition, the map $\Phi^r_{\alpha,\theta}$ commutes with the action of the centralizer $C_G = \exp(C_{\mathfrak{I}}(X_\alpha)) \subset G$ and its action on the product $M^a_\theta \times M^a_\theta$ commutes with $\psi^{(r)}_{\alpha,\theta}$. That is, for $c \in C_G$ and $x = \Gamma g$,

$$\psi^{(r)}_{\alpha,\theta}(xc) = (xc, \Phi^r_{\alpha,\theta}(xc)) = (xc, \Phi^r_{\alpha,\theta}(x)c) = \psi^{(r)}_{\alpha,\theta}(x)c. \tag{43}$$

Then the quotient map is well defined on

$$\Psi^{(r)}_{\alpha,\theta} := M^a_\theta / C_G \longrightarrow M^a_\theta \times M^a_\theta / C_G. \tag{44}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

Setting. (i) In $M_\theta^a \times M_\theta^a$, we set the diagonal $\Delta = \{(x, x) \mid x \in M_\theta^a\}$ which is isomorphic to M_θ^a by identifying (x, x) with $x \in M_\theta^a$. Given $(x, x) \in \Delta$, the tangent space of the diagonal is $T_{(x,x)}\Delta := \{(v, v) \mid v \in T_x M_\theta^a\}$, and its normal space is defined by $(T_{(x,x)}\Delta)^\perp = \{(v, -v) \mid v \in T_x M_\theta^a\} = T_{(x,x)}\Delta^\perp$. On the tangent space at $(x, x) \in M_\theta^a \times M_\theta^a$, it splits as

$$T_{(x,x)}(M_\theta^a \times M_\theta^a) = T_{(x,x)}\Delta \oplus (T_{(x,x)}\Delta)^\perp.$$

For any $w_1, w_2 \in T_x M_\theta^a$,

$$(w_1, w_2) = (1/2)(w_1 + w_2, w_1 + w_2) + (1/2)(w_1 - w_2, -(w_1 - w_2)). \tag{45}$$

(ii) Given $x = \Gamma h_1$, $y = \Gamma h_2 \in M_\theta^a$, define a set $\Delta_{(x,y)} = \{(xg, yg) \mid g \in G\} \subset M_\theta^a \times M_\theta^a$ for $(xg, yg) = (\Gamma h_1 g, \Gamma h_2 g)$ and $\Delta_{(x,y)}^\perp = \{(xg, yg^{-1}) \mid g \in G\}$ that contains (x, y). For $\psi_{\alpha,\theta}^{(r)}(x) = (x, \Phi_{\alpha,\theta}^r(x))$, its tangent space in $M_\theta^a \times M_\theta^a$ is decomposed as

$$T_{(x,\Phi_{\alpha,\theta}^r(x))}(M_\theta^a \times M_\theta^a) = T_{(x,\Phi_{\alpha,\theta}^r(x))}\Delta_{(x,\Phi_{\alpha,\theta}^r(x))} \oplus (T_{(x,\Phi_{\alpha,\theta}^r(x))}\Delta_{(x,\Phi_{\alpha,\theta}^r(x))})^\perp.$$

Then the tangent space of the diagonal is $T_{(x,\Phi_{\alpha,\theta}^r(x))}\Delta_{(x,\Phi_{\alpha,\theta}^r(x))} = \{(v, d_x\Phi_{\alpha,\theta}^r(v)) \mid v \in T_x M_\theta^a\}$ and its normal space is identified as

$$(T_{(x,\Phi_{\alpha,\theta}^r(x))}\Delta_{(x,\Phi_{\alpha,\theta}^r(x))})^\perp = T_{(x,\Phi_{\alpha,\theta}^r(x))}\Delta_{(x,\Phi_{\alpha,\theta}^r(x))}^\perp.$$

By identification in (45), for $w_1 = v$ and $w_2 = -d_x\Phi_{\alpha,\theta}^r(v)$, we write

$$\begin{aligned}&(T_{(x,\Phi_{\alpha,\theta}^r(x))}\Delta_{(x,\Phi_{\alpha,\theta}^r(x))})^\perp \\ &\quad = \{(1/2)(v - d_x\Phi_{\alpha,\theta}^r(v)), -(1/2)(v - d_x\Phi_{\alpha,\theta}^r(v)) \mid v \in T_x M_\theta^a\}.\end{aligned} \tag{46}$$

(iii) Now define the orthogonal projection $\pi : M_\theta^a \times M_\theta^a \to M_\theta^a \times M_\theta^a$ along the direction of the diagonal. That is, for $(x, y) \in M_\theta^a \times M_\theta^a$, there exists (x', y') such that $\pi(x, y) = (x', y') \in \Delta_{(x,y)} \cap \Delta_{(x,x)}^\perp$. Then

$$T_{\pi(x,y)}\Delta_{\pi(x,y)} = T_{(x,y)}\Delta_{(x,y)}, \quad T_{\pi(x,y)}\Delta_{\pi(x,y)}^\perp = T_{(x,y)}\Delta_{(x,y)}^\perp. \tag{47}$$

Define a map $F^{(r)} : M_\theta^a \to M_\theta^a \times M_\theta^a$ given by $F^{(r)} = \pi \circ \psi_{\alpha,\theta}^{(r)}$. In the local coordinate, by identification (46) and (47),

$$d_x F^{(r)}(v) = (1/2)(v - d_x\Phi_{\alpha,\theta}^r(v)), -(1/2)(v - d_x\Phi_{\alpha,\theta}^r(v)), \quad v \in T_x M_\theta^a. \tag{48}$$

By (43) and the definition of $F^{(r)}$, we have $F^{(r)}(xc) = F^{(r)}(x)c$ for $c \in C_G$. Then for all $r \in \mathbb{Z}$, $F^{(r)}$ induces a quotient map $F_C^{(r)} : M_\theta^a/C_G \to M_\theta^a \times M_\theta^a/C_G$. From (48), the range of the differential $DF_C^{(r)}$ is determined by $I - D\Phi_{\alpha,\theta}^r$ (refer to Figure 1).

In the next lemma, we verify the range of the differential map $DF_C^{(r)}$.

LEMMA 5.3. *For all $r \in \mathbb{Z}\backslash\{0\}$, the range of $I - D\Phi_{\alpha,\theta}^r$ on $\mathfrak{I}/C_{\mathfrak{I}}(X_\alpha)$ coincides with $Ran(ad_{X_\alpha})$ and the Jacobian of $F_C^{(r)}$ is non-zero constant.*

Proof. Recall that $\Phi_{\alpha,\theta}^r$ is the rth return map on M_θ^a. We find the differential in the direction of each Y_i^j for fixed i and j. For $x \in M$, set a curve $\gamma_{i,j}^x(t) = x\exp(tY_i^{(j)})\exp(rX_\alpha)$.

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

Note that

$$\exp(tY_i^{(j)})\exp(rX_\alpha) = \exp(rX_\alpha)\exp(-rX_\alpha)\exp(tY_i^{(j)})\exp(rX_\alpha)$$
$$= \exp(rX_\alpha)\exp(e^{-r(\mathrm{ad}_{X_\alpha})}(tY_i^{(j)}))$$

and

$$\frac{d}{dt}\gamma_{i,j}^x(t)\mid_{t=0} = e^{-r(\mathrm{ad}_{X_\alpha})}(Y_i^{(j)}).$$

By definition, $(\partial\Phi^r_{\alpha,\theta}/\partial s_i^{(j)})(x) = d/dt(\gamma^x_{i,j}(t))\mid_{t=0}$ and we have $I - D\Phi^r_{\alpha,\theta} = I - \sum_{(i,j)\in J}(\partial\Phi^r_{\alpha,\theta}/\partial s_i^{(j)})$. Then

$$(I - D\Phi^r_{\alpha,\theta})\bigg(\sum_{(i,j)\in J} s_i^{(j)}Y_i^{(j)}\bigg) = \bigg[r(\mathrm{ad}_{X_\alpha})\bigg(\sum_{k=0}^{\infty}\frac{(-1)^k}{(k+1)!}(\mathrm{ad}_{X_\alpha})^k\bigg)\bigg]\bigg(\sum_{(i,j)\in J} s_i^{(j)}Y_i^{(j)}\bigg). \tag{49}$$

Therefore, the range of $I - D\Phi^r_{\alpha,\theta}$ is contained in $\mathrm{Ran}(\mathrm{ad}_{X_\alpha})$.

Conversely, $(1 - e^{-\mathrm{ad}_{X_\alpha}})/\mathrm{ad}_{X_\alpha} = \sum_{k=0}^{\infty}((-1)^k)/((k+1)!)(\mathrm{ad}_{X_\alpha})^k$ is invertible and

$$(I - D\Phi^r_{\alpha,\theta})\bigg(\bigg(\frac{1 - e^{-\mathrm{ad}_{X_\alpha}}}{\mathrm{ad}_{X_\alpha}}\bigg)^{-1}\bigg(\sum_{(i,j)\in J} s_i^{(j)}Y_i^{(j)}\bigg)\bigg) = r(\mathrm{ad}_{X_\alpha})\bigg(\sum_{(i,j)\in J} s_i^{(j)}Y_i^{(j)}\bigg). \tag{50}$$

Therefore, we conclude that the range of $I - D\Phi^r_{\alpha,\theta}$ is $\mathrm{Ran}(\mathrm{ad}_{X_\alpha})$.

If $\sum_{(i,j)\in J} s_i^{(j)}Y_i^{(j)} \in C_{\mathfrak{I}}(X_\alpha)$, then $(I - D\Phi^r_{\alpha,\theta})(\sum_{(i,j)\in J} s_i^{(j)}Y_i^{(j)}) = 0$ and the kernel of $I - D\Phi^r_{\alpha,\theta}$ is $C_{\mathfrak{I}}(X_\alpha)$. That is, $I - D\Phi^r_{\alpha,\theta}$ is bijective on $\mathfrak{I}/C_{\mathfrak{I}}(X_\alpha)$. Thus, by (49), the Jacobian of $I - D\Phi^r_{\alpha,\theta}$ is non-zero constant, and this concludes the proof. □

Setting (continued). (iv) Set the submanifold $\mathcal{S} \subset M^a_\theta \times M^a_\theta$ that consists of the diagonal Δ and the coordinates of generators in normal (transverse) directions. Denote its quotient $\mathcal{S}_C = \mathcal{S}/C_G \subset M^a_\theta \times M^a_\theta/C_G$. Then, following Lemma 5.3, we obtain the transversality of $F_C^{(r)}$ to $\mathcal{S}_C$. For every $p \in (F_C^{(r)})^{-1}(\mathcal{S}_C)$, the transversality holds on the tangent space:

$$T_{F_C^{(r)}(p)}\mathcal{S}_C + DF_C^{(r)}(T_pM^a_\theta/C_G) = T_{F_C^{(r)}(p)}(M^a_\theta \times M^a_\theta/C_G). \tag{51}$$

(v) Denote the Lebesgue measure $\mathcal{L}^{a+1}(=\mathrm{vol}_M)$ on the nilmanifold M and the conditional measure $\mathcal{L}^a_\theta(=\mathrm{vol}_{M^a_\theta})$ on the transverse manifold M^a_θ. On the quotient space M^a_θ/C_G, we write the measure $\mathcal{L}^c_\theta(=\mathrm{vol}_{M^a_\theta/C_G})$. Similarly, we set the conditional measure $\mu^a_\theta(=\mathrm{vol}_{M^a_\theta\times M^a_\theta})$ on the product manifold and $\mu^c_\theta(=\mathrm{vol}_{M^a_\theta\times M^a_\theta/C_G})$ on its quotient space.

Denote the image of $F^{(r)}$ by $M^a_{\theta,r} := F^{(r)}(M^a_\theta) \subset M^a_\theta \times M^a_\theta$ and $M^a_{\theta,r,C} := F_C^{(r)}(M^a_\theta/C_G)$. We write its conditional Lebesgue measure $\mu^a_{\theta,r} := \mu^a_\theta|_{M^a_{\theta,r}}$ and $\mu^c_{\theta,r} := \mu^c_\theta|_{M^a_{\theta,r,c}}$, respectively.

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

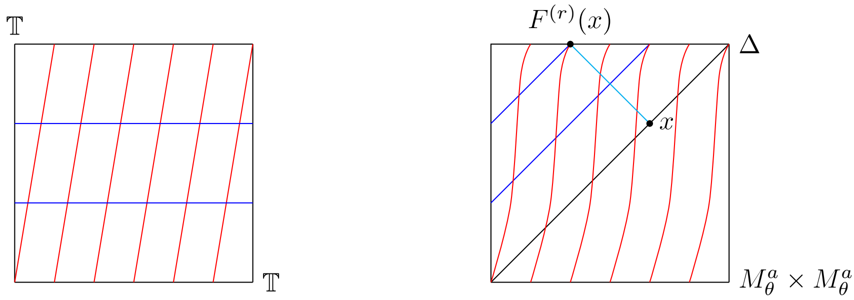

FIGURE 1. Illustration of displacement $F^{(r)}$ in the product $M^a_\theta \times M^a_\theta$ and comparison with the uniform expanding map.

For any open set $U_{\mathcal{S}_C} \subset M^a_\theta \times M^a_\theta / C_G$, we write the pushforward measure $(F_C^{(r)})_* \mathcal{L}^c_\theta$ as

$$\begin{aligned}(F_C^{(r)})_* \mathcal{L}^c_\theta (U_{\mathcal{S}_C} \cap M^a_{\theta,r,C}) &= \mathcal{L}^c_\theta ((F_C^{(r)})^{-1}(U_{\mathcal{S}_C} \cap M^a_{\theta,r,C})) \\ &= \int_{U_{\mathcal{S}_C}} \sum_{x \in (F_C^{(r)})^{-1}(\{z\}), z \in U_{\mathcal{S}_C}} \frac{1}{\mathrm{Jac}(F_C^{(r)}(x))} d\mathrm{vol}_{M^a_{\theta,r,c}}(z).\end{aligned}$$

By the compactness of M^a_θ (or M^a_θ / C_G), the above expression is finite. By Lemma 5.3, the Jacobian of $F_C^{(r)}$ is constant and $(F_C^{(r)})_* \mathcal{L}^c_\theta = \mu^c_{\theta,r}$ is Lebesgue.

By invariance of the action of the centralizer, for any neighborhood $U_{\mathcal{S}} \in M^a_\theta \times M^a_\theta$ with $U_{\mathcal{S}_C} = U_{\mathcal{S}} / C_G$,

$$\mu^c_{\theta,r}(U_{\mathcal{S}_C} \cap M^a_{\theta,r,C}) = \mu^a_{\theta,r}(U_{\mathcal{S}} \cap M^a_{\theta,r}), \tag{52}$$

and by definition of conditional measure,

$$\mu^a_{\theta,r}(U_{\mathcal{S}} \cap M^a_{\theta,r}) = \mu^a_\theta(U_{\mathcal{S}}). \tag{53}$$

Let d be a distance function in $M^a_\theta \times M^a_\theta$, and we abuse notation d for the induced distance on $M^a_\theta \times M^a_\theta / C_G$. Let $U_\delta = \{z \in M^a_\theta \times M^a_\theta \mid d(z, \mathcal{S}) < \delta\}$ be a δ-tubular neighborhood of $\mathcal{S}$ and $U_{\delta,C} = \{z \in M^a_\theta \times M^a_\theta / C_G \mid d(z, \mathcal{S}_C) < \delta\}$ be its quotient.

Define the *almost-periodic set* (set of rth close return) on the diagonal

$$AP^r(\mathcal{U}_\delta) := \{x \in M^a_\theta \mid d(F^{(r)}(x), \mathcal{S}) < \delta\}. \tag{54}$$

Since $F^{(r)}$ commutes with C_G, $AP^r(\mathcal{U}_\delta)/C_G = \{x \in M^a_\theta / C_G \mid d(F_C^{(r)}(x), \mathcal{S}_C) < \delta\}$ and $\mathcal{L}^a_\theta(AP^r(\mathcal{U}_\delta)) = \mathcal{L}^c_\theta(AP^r(\mathcal{U}_\delta)/C_G)$.

The following volume estimate of *almost-periodic set* holds.

LEMMA 5.4. *Let $U_{\delta,C}$ be any tubular neighborhood of $\mathcal{S}_C$ in $M^a_\theta \times M^a_\theta / C_G$. For all $r \in \mathbb{Z} \backslash \{0\}$, the conditional measure $\mathrm{vol}_{M^a_\theta}$ of $AP^r(\mathcal{U}_\delta)$ is given as follows:*

$$\mathcal{L}^a_\theta(AP^r(\mathcal{U}_\delta)) = \mu^c_{\theta,r}(U_{\delta,C} \cap M^a_{\theta,r,C}).$$

Proof. By setting (v), it suffices to prove $\mathcal{L}^c_\theta(AP^r(\mathcal{U}_\delta)/C_G) = \mu^c_{\theta,r}(U_\delta \cap M^a_{\theta,r,C})$. Note that $(F^{(r)}_C)^{-1}(AP^r(\mathcal{U}_\delta)/C_G) = \{x \in M^a_\theta/C_G \mid d(z, \mathcal{S}_C) < \delta\}$ if $z = F^{(r)}_C(x)$ for some $x \in M^a_\theta/C_G$, otherwise it is an empty set.

Then $(F^{(r)}_C)^{-1}(AP^r(\mathcal{U}_\delta)/C_G) = (U_{\delta,C} \cap M^a_{\theta,r,C})$. Thus, by definition of the pushforward measure, the equality holds. □

Recall that $F^{(r)}_C : M^a_\theta/C_G \to M^a_{\theta,r,C}$ has non-zero constant Jacobian if $r \neq 0$ by Lemma 5.3 and it is a local diffeomorphism. Thus, by the transversality of $F^{(r)}_C$, in a small tubular neighborhood U, $F^{(r)}_C$ is covering.

LEMMA 5.5. *For any $z \in U \cap M^a_{\theta,r,C}$, there exist a finite number of pre-images of $F^{(r)}_C$.*

Proof. If we suppose that $(F^{(r)}_C)^{-1}(z)$ contains infinitely many different points, then since the manifold M^a_θ is compact (and M^a_θ/C_G is compact), there exists a sequence of pairwise different points $x_i \in (F^{(r)}_C)^{-1}(z)$, which converges to x_0. We have $(F^{(r)}_C)(x_0) = z$ and, by the inverse function theorem, the point x_0 has a neighborhood U' in which $F^{(r)}_C$ is a homeomorphism. In particular, $U' \backslash \{x_0\} \cap (F^{(r)}_C)^{-1}(z) = \emptyset$, which is a contradiction. □

Set $N_r(z) = \#\{x \in M^a_\theta/C_G \mid F^{(r)}_C(x) = z\}$ as the number of pre-images of $F^{(r)}_C$. The number $N_r(z)$ is independent of the choice of $z \in U \cap M^a_{\theta,r,C}$ since Jacobian is constant and the degree of the map is invariant (see [**DAS**, §3]).

We now introduce the volume estimate of the δ-neighborhood $U_{\delta,C}$.

PROPOSITION 5.6. *The following volume estimate holds: for any $r \neq 0$, there exists $C := C(M^a_\theta) > 0$ such that*

$$\mu^c_{\theta,r}(U_{\delta,C} \cap M^a_{\theta,r,C}) < C\delta.$$

Proof. Let $U \subset M^a_\theta \times M^a_\theta/C_G$ be a tubular neighborhood of $\mathcal{S}_C$ that contains $U_{\delta,C}$ with the following condition:

$$\mathrm{vol}_{M^a_\theta \times M^a_\theta/C_G}(U_{\delta,C}) = \delta \mathrm{vol}_{M^a_\theta \times M^a_\theta/C_G}(U). \tag{55}$$

If $U \cap M^a_{\theta,r,C} = \emptyset$, then there is nothing to prove since $U_{\delta,C} \cap M^a_{\theta,r,C} = \emptyset$. Assume $z \in U \cap M^a_{\theta,r,C}$ and let $\{J_k\}_{k\geq 1}$ be connected components of $(F^{(r)}_C)^{-1}(U \cap M^a_{\theta,r,C})$. We first claim that $F^{(r)}_C|_{J_k}$ is injective.

Given $z \in U \cap M^a_{\theta,r,C}$, assume that there exist $x_1 \neq x_2 \in J_k$ for some k such that $z = F^{(r)}_C|_{J_k}(x_1) = F^{(r)}_C|_{J_k}(x_2)$. Let $\gamma : [0, 1] \to J_k$ be a path that connects $\gamma(0) = x_1$ and $\gamma(1) = x_2$. Set the lift of the path $\tilde{\gamma} = F^{(r)}_C|_{J_k} \circ \gamma : [0, 1] \to U$. Then $\tilde{\gamma}(0) = \tilde{\gamma}(1) = z$ and $\tilde{\gamma}$ is a loop in U. Since U is simply connected, $\tilde{\gamma}$ is contractible and there exists a homotopy of the path $g_s : [0, 1] \to U$ such that $g_0 = \tilde{\gamma}$ is homotopic to a constant loop $g_1 = c$ by fixing two end points $F^{(r)}_C|_{J_k}(x_1) = F^{(r)}_C|_{J_k}(x_2) = z$ for $s \in [0, 1]$.

Note that $F^{(r)}_C|^{-1}_{J_k} \circ g_s$ is a lift of homotopy g_s, and a lift of g_0 is $\gamma = F^{(r)}_C|^{-1}_{J_k}(\tilde{\gamma})$ with fixed end points x_1 and x_2. By homotopy continuity, g_s also keeps the same end points x_1 and x_2 fixed for all $s \in [0, 1]$. Since g_1 is constant loop and its lift should be a single point,

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

γ is homotopic to a constant. Since the end points of γ are fixed, they have to be constant, but this leads a contradiction. Therefore, we have $x_1 = x_2$.

By injectivity of $F_C^{(r)}|_{J_k}$, we obtain

$$(F_C^{(r)})^{-1}(U) = (F_C^{(r)})^{-1}(U \cap M_{\theta,r,C}^a) = \bigcup_{k=1}^{N_r} J_k.$$

Furthermore, we obtain the following equality:

$$\mathrm{vol}_{M_\theta^a/C_G}(J_k) = \frac{\mathrm{vol}_{M_{\theta,r,C}^a}(U \cap M_{\theta,r,C}^a)}{\mathrm{Jac}(F_C^{(r)}|_{J_k})} = \frac{\mathrm{vol}_{M_{\theta,r,C}^a}(U \cap M_{\theta,r,C}^a)}{\mathrm{Jac}(F_C^{(r)})}. \tag{56}$$

Since the volume of M_θ^a/C_G is a finite,

$$N_r\left(\frac{\mathrm{vol}_{M_\theta^a \times M_\theta^a/C_G}(U)}{\mathrm{Jac}(F_C^{(r)})}\right) = N_r\left(\frac{\mathrm{vol}_{M_{\theta,r,C}^a}(U \cap M_{\theta,r,C}^a)}{\mathrm{Jac}(F_C^{(r)})}\right) \tag{57}$$

$$= \sum_{k=1}^{N_r} \mathrm{vol}_{M_\theta^a/C_G}(J_k) < \infty. \tag{58}$$

Assume that $(F_C^{(r)})^{-1}(U_{\delta,C}) = \bigcup_{k=1}^{N_r}(F_C^{(r)}|_{J_k})^{-1}(U_{\delta,C})$. Then by (56) and definition of conditional measure,

$$\begin{aligned}\mathrm{vol}_{M_\theta^a/C_G}((F_C^{(r)})^{-1}(U_{\delta,C})) &= \sum_{k=1}^{N_r} \mathrm{vol}_{M_\theta^a/C_G}((F_C^{(r)}|_{J_k})^{-1}(U_{\delta,C}))\\ &= N_r\left(\frac{\mathrm{vol}_{M_{\theta,r,C}^a}(U_{\delta,C} \cap M_{\theta,r,C}^a)}{\mathrm{Jac}(F_C^{(r)})}\right)\\ &= N_r\left(\frac{\mathrm{vol}_{M_\theta^a \times M_\theta^a/C_G}(U_{\delta,C})}{\mathrm{Jac}(F_C^{(r)})}\right).\end{aligned}$$

By the last equality, together with condition (55),

$$\mathrm{vol}_{M_\theta^a/C_G}((F_C^{(r)})^{-1}(U_{\delta,C})) = N_r\left(\frac{\delta \mathrm{vol}_y M_\theta^a \times M_\theta^a/C_G(U)}{\mathrm{Jac}(F_C^{(r)})}\right). \tag{59}$$

Therefore, combining (57) and (59), there exists $C > 0$ such that

$$\mu_{\theta,r}^c(U_{\delta,C}) = (F_C^{(r)})_* \mathrm{vol}_{M_\theta^a/C_G}(U_{\delta,C}) = \mathrm{vol}_{M_\theta^a/C_G}((F_C^{(r)})^{-1}(U_{\delta,C})) < C\delta. \qquad \square$$

Definition 5.7. For any basis $Y = \{Y_1, \ldots Y_a\}$ of codimension-1 ideal $\mathfrak{I}$ of $\mathfrak{n}$, let $I := I(Y)$ be the supremum of all constants $I' \in (0, \frac{1}{2})$ such that, for any $x \in M$, the map

$$\phi_x^Y : (s_1, \ldots, s_a) \mapsto x \exp\left(\sum_{i=1}^{a} s_i Y_i\right) \in M \tag{60}$$

is a local embedding (injective) on the domain

$$\{\mathbf{s} \in \mathbb{R}^a \mid |s_i| < I' \text{ for all } i = 1, \ldots, a\}.$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

For any $x, x' \in M$, set local distance d_* (measured locally in the Lie algebra) on the transverse section M_θ^a along the Y_i direction by $d_{Y_i}(x, x') = |s_i|$ if there is $\mathbf{s} := (s_1, \ldots, s_a) \in [-I/2, I/2]^a$ such that

$$x' = x \exp\left(\sum_{i=1}^{a} s_i Y_i\right),$$

otherwise $d_{Y_i}(x, x') = I$.

Recall the projection map $pr_1 : M \to \mathbb{T}^{n+1}$ onto the base torus. On the transverse manifold, for all $\theta \in \mathbb{T}^1$, let $pr_\theta : M_\theta^a \to \mathbb{T}^n$ be the restriction to M_θ^a. Then, by applying formula (10) to projection to the base torus,

$$d_{Y_i}(pr_\theta(\Phi^r_{\alpha,\theta}(x)), pr_\theta(x)) = r\alpha_i, \quad 1 \le i \le n. \tag{61}$$

For any $L \ge 1, r \in \mathbb{Z}, x \in M_\theta^a$ and given scaling factor $\rho = (\rho_1, \ldots, \rho_a) \in [0, 1)^a$, we define

$$\begin{aligned} \epsilon_{r,L} &:= \max_{1 \le i \le n} \min\{I, L^{\rho_i} d_{Y_i}(\Phi^r_{\alpha,\theta}(x), x)\}, \\ \delta_{r,L}(x) &:= \max_{n+1 \le i \le a} \min\{I, L^{\rho_i} d_{Y_i}(\Phi^r_{\alpha,\theta}(x), x)\}. \end{aligned} \tag{62}$$

We note that the distance $d_{Y_i}(\Phi^r_{\alpha,\theta}(x), x)$ on the generator level does not depend on the choice of x. For this reason, we split the cases $\epsilon_{r,L}$ and $\delta_{r,L}(x)$ to control the distance separately on higher steps.

The condition $\epsilon_{r,L} < \epsilon < I$ and $\delta' < \delta_{r,L}(x) < \delta < I$ are equivalent to saying

$$\Phi^r_{\alpha,\theta}(x) = x \exp\left(\sum_{i=1}^{a} s_i Y_i\right) \tag{63}$$

for some vectors $\mathbf{s} := (s_1, \ldots, s_a) \in [-I/2, I/2]^a$ such that

$$\begin{aligned} &|s_i| < \epsilon L^{-\rho_i} \quad \text{for all } i \in \{1, \ldots, n\}, \\ &|s_i| < \delta L^{-\rho_i} \quad \text{for all } i \in \{n+1, \ldots, a\}, \\ &|s_j| > \delta' L^{-\rho_j} \quad \text{for some } j \in \{n+1, \ldots, a\}. \end{aligned}$$

For every $r \in \mathbb{Z} \backslash \{0\}$ and $j \ge 0$, let $AP^r_{j,L} \subset M$ be a set defined as follows:

$$AP^r_{j,L} = \begin{cases} \emptyset & \text{if } \epsilon_{r,L} > \frac{I}{2}, \\ (\delta_{r,L})^{-1}\Big((2^{-(j+1)} I, 2^{-j} I]\Big) & \text{otherwise.} \end{cases} \tag{64}$$

In the next lemma, the Lebesgue measure of the set of almost-periodic points $AP^r_{j,L}$ on M is estimated by the volume of theδ-neighborhood $\mathcal{U}_\delta$.

LEMMA 5.8. *For all $r \in \mathbb{Z}\backslash\{0\}$, $j \in \mathbb{N}$, $L \ge 1$, the $(a+1)$ dimensional Lebesgue measure of the set $AP^r_{j,L}$ can be estimated as follows: there exists $C > 0$ such that*

$$\mathcal{L}^{a+1}(AP^r_{j,L}) \le \frac{C I^{a-n}}{2^{j(a-n)}} L^{-\sum_{i=n+1}^{a} \rho_i}.$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

Proof. Without loss of generality, we assume that $AP^r_{j,L} \neq \emptyset$.

By Tonelli's theorem,

$$\mathcal{L}^{a+1}(AP^r_{j,L}) = \int_0^1 \mathcal{L}^a_\theta(AP^r_{j,L} \cap M^a_\theta)d\theta. \tag{65}$$

Recall the definition $AP^r(\mathcal{U}_\delta)$ in (54). Choose $\delta = I^{a-n}/2^{j(a-n)}L^{-\sum_{i=n+1}^a \rho_i}$ and set $\mathcal{U}^{L,j}_\delta := \mathcal{U}_\delta$. Then we claim that $AP^r_{j,L} \cap M^a_\theta \subset AP^r(\mathcal{U}^{L,j}_\delta)$.

By identification of $x \in M^a_\theta$ with (x, x) in the diagonal $\Delta \subset M^a_\theta \times M^a_\theta$, the local distance $d_{Y_i}(\Phi^r_{\alpha,\theta}(x), x)$ is identified by the distance function d in the product $M^a_\theta \times M^a_\theta$. Thus, $x \in AP^r_{j,L} \cap M^a_\theta$ implies that $d(F^{(r)}(x), \mathcal{S}) < \delta$. That is, $AP^r_{j,L} \cap M^a_\theta \subset AP^r(\mathcal{U}^{L,j}_\delta)$. By Lemma 5.4, the volume estimate follows

$$\mathcal{L}^a_\theta(AP^r_{j,L} \cap M^a_\theta) \leq \mathcal{L}^a_\theta(AP^r(\mathcal{U}^{L,j}_\delta)) = \mu^c_{\theta,r}(\mathcal{U}^{L,j}_{\delta,C} \cap M^a_{\theta,r,C}).$$

Finally, by Proposition 5.6,

$$\mu^c_{\theta,r}(\mathcal{U}^{L,j}_{\delta,C} \cap M^a_{\theta,r,C}) \leq \frac{CI^{a-n}}{2^{j(a-n)}}L^{-\sum_{i=n+1}^a \rho_i}.$$

Thus the proof follows from formula (65). □

5.2. *Expected width bounds.* We prove a bound on the average width of an orbit on a nilmanifold with respect to a scaled basis. This subsection follows along the lines of [**FF14**, §5.2]. To complete of the proof, we repeat the similar arguments in nilmanifolds under transverse conditions.

For $L \geq 1$ and $r \in \mathbb{Z}\backslash\{0\}$, let us consider the function

$$h_{r,L} = \sum_{j=1}^{\infty} \min\left\{2^{j(a-n)}, \left(\frac{2}{\epsilon_{r,L}}\right)^n\right\}\chi_{AP^r_{j,L}}. \tag{66}$$

Define the cut-off function $J_{r,L} \in \mathbb{N}$ by the formula

$$J_{r,L} := \max\left\{j \in \mathbb{N} \mid 2^{j(a-n)} \leq \left(\frac{2}{\epsilon_{r,L}}\right)^n\right\}. \tag{67}$$

The function $h_{r,L}$ is

$$h_{r,L} = \sum_{j=1}^{J_{r,L}} 2^{j(a-n)}\chi_{AP^r_{j,L}} + \sum_{j>J_{r,L}} \left(\frac{2}{\epsilon_{r,L}}\right)^n \chi_{AP^r_{j,L}}. \tag{68}$$

For every $L \geq 1$, let $\mathcal{F}^{(L)}_\alpha$ be the rescaled strongly adapted basis

$$\mathcal{F}^{(L)}_\alpha = (X^{(L)}_\alpha, Y^{(L)}_1, \ldots, Y^{(L)}_a) = (LX_\alpha, L^{-\rho_1}Y_1, \ldots, L^{-\rho_a}Y_a). \tag{69}$$

For $(x, T) \in M \times \mathbb{R}$, let $w_{\mathcal{F}^{(L)}_\alpha}(x, T)$ denote the *average width* of the (scaled) orbit segment

$$\gamma^T_{X^{(L)}_\alpha}(x) := \{\phi^t_{X^{(L)}_\alpha}(x) \mid 0 \leq t \leq T\}.$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

We prove a bound for the average width of the orbit arc in terms of the following function:

$$H_L^T := 1 + \sum_{|r|=1}^{[TL]} h_{r,L}. \tag{70}$$

Definition 5.9. For $t \in [0, T]$, we define a set of points $\Omega(t) \subset \{t\} \times \mathbb{R}^a$ as follows.

Case 1. If $\phi^t_{X^{(L)}_\alpha}(x) \notin \bigcup_{|r|=1}^{[TL]} \bigcup_{j>0} AP^r_{j,L}$, let $\Omega(t)$ be the set of all points $(t, s_1, \ldots, s_a)$ such that

$$|s_i| < I/4, \quad i \in \{1, \ldots, a\}.$$

If $\phi^t_{X^{(L)}_\alpha}(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j>0} AP^r_{j,L}$, then we consider two subcases.

Case 2-1. If $\phi^t_{X^{(L)}_\alpha}(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j>J_{r,L}} AP^r_{j,L}$, let $\Omega(t)$ be the set of all points $(t, \mathbf{s})$ such that

$$\begin{aligned} |s_i| &< \tfrac{1}{4} \min_{1\le |r| \le [TL]} \min_{j > J_{r,L}} \{\epsilon_{r,L} : \phi^t_{X^{(L)}_\alpha}(x) \in AP^r_{j,L}\} \quad \text{for } i \in \{1, \ldots, n\}, \\ |s_i| &< \frac{I}{4} \quad \text{for } i \in \{n+1, \ldots, a\}. \end{aligned}$$

Case 2-2. If $\phi^t_{X^{(L)}_\alpha}(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j\le J_{r,L}} AP^r_{j,L} \setminus \bigcup_{|r|=1}^{[TL]} \bigcup_{j>J_{r,L}} AP^r_{j,L}$, let l be the largest integer such that

$$\phi^t_{X^{(L)}_\alpha}(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{l\le j\le J_{r,L}} AP^r_{j,L} \setminus \bigcup_{|r|=1}^{[TL]} \bigcup_{j>J_{r,L}} AP^r_{j,L},$$

and let $\Omega(t)$ be the set of all points $(t, \mathbf{s})$ such that

$$\begin{aligned} |s_i| &< \frac{I}{4} \quad \text{for } i \in \{1, \ldots, n\}, \\ |s_i| &< \frac{I}{4}\frac{1}{2^{l+1}} \quad \text{for } i \in \{n+1, \ldots, a\}. \end{aligned}$$

We set

$$\Omega := \bigcup_{t\in[0,T]} \Omega(t) \subset [0, T] \times [-I/4, I/4]^a \subset [0, T] \times \mathbb{R}^a.$$

LEMMA 5.10. *The restriction to Ω of the map*

$$(t, \mathbf{s}) \in \Omega \mapsto x \exp(tX^{(L)}_\alpha) \exp\bigg(\sum_{i=1}^a s_i Y^{(L)}_i\bigg) \tag{71}$$

is injective.

Proof. Assume there exist $t' \ge t$ and $(t, s_1, \ldots, s_n) \neq (t', s'_1, \ldots, s'_n)$ such that

$$\phi^t_{X^{(L)}_\alpha}(x) \exp\bigg(\sum_{i=1}^a s_i Y^{(L)}_i\bigg) = \phi^{t'}_{X^{(L)}_\alpha}(x) \exp\bigg(\sum_{i=1}^a s'_i Y^{(L)}_i\bigg). \tag{72}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

By considering the projection on the base torus, we have the following identity:

$$(t, s_1, \ldots, s_n) \bmod \mathbb{Z}^{n+1} = pr_1(\phi^{t}_{X^{(L)}_\alpha}(x)) \\ = pr_1(\phi^{t'}_{X^{(L)}_\alpha}(x)) = (t', s'_1, \ldots, s'_n) \bmod \mathbb{Z}^{n+1}, \qquad (73)$$

which implies $t \equiv t'$ modulo $\mathbb{Z}$. As $\phi^{tL}_{X_\alpha} = \phi^{t}_{X^{(L)}_\alpha}$, the number $r_0 = t' - t$ is a non-negative integer satisfying $r_0 \leq TL$; hence $r_0 \leq [TL]$.

If $r_0 = 0$, then $t' = t$ and $s'_i = s_i$ for all $1 \leq i \leq a$. Then injectivity is obtained by definition of I. Assume that $r_0 \neq 0$. Let $p, q \in M^a_\theta$. Then we have

$$p := \phi^{t}_{X^{(L)}_\alpha}(x), \quad q := \phi^{t'}_{X^{(L)}_\alpha}(x) \Longrightarrow q = \Phi^{r_0}_{\alpha,\theta}(p).$$

From identity (72) we have

$$q = p \exp\bigg(\sum_{i=1}^{a} s_i Y^{(L)}_i\bigg) \exp\bigg(-\sum_{i=1}^{a} s'_i Y^{(L)}_i\bigg) \\ = p \exp\bigg(\sum_{i=1}^{a}(s'_i - s_i + P_i(s_i, s'_i))L^{-\rho_i} Y_i\bigg) \qquad (74)$$

where P_i is polynomial expression following from the Baker–Cambell–Hausdorff formula.

Note that $P_i(s_i, s'_i) = 0$ if $i = 1, \ldots, n$ and $|P_i(s_i, s'_i)| \leq \sum_{l=1}^{\infty} 1/2|s_l s'_l|^l$ for $i > n$. Since $|s_i|, |s'_i| \leq I/4 \ll 1$,

$$q = p \exp\bigg(\sum_{i=1}^{a}(s'_i - s_i + \epsilon_i)L^{-\rho_i} Y_i\bigg) \quad \text{for some } \epsilon_i \in [0, I_\epsilon)$$

where $I_\epsilon = \sum_{l=1}(I/4)^l = I/(4 - I) < I/3$. Thus for all $i \in \{1, \ldots, a\}$,

$$L^{\rho_i} d_{Y_i}(p, \Phi^{r_0}_{\alpha,\theta}(p)) = L^{\rho_i}|(s'_i - s_i + \epsilon_i)L^{-\rho_i}| \leq |s'_i| + |s_i| + |\epsilon_i| \qquad (75)$$

and

$$\epsilon_{r_0,L} = \max_{1\leq i\leq n} L^{\rho_i} d_{Y_i}(p, \Phi^{r_0}_{\alpha,\theta}(p)) \leq \max_{1\leq i\leq n}\{|s_i| + |s'_i| + |\epsilon_i|\} \leq \tfrac{5}{6} I.$$

For the same reason, from formula (74) we also obtain that

$$\delta_{r_0,L}(p) = \delta_{-r_0,L}(q) < I/2.$$

By defining $j_0 \in \mathbb{N}$ as the unique non-negative integer such that

$$\frac{I}{2^{j_0+1}} \leq \delta_{r_0,L}(p) \leq \frac{I}{2^{j_0}}$$

and by the Definition 5.7, we have $p \in AP^{r_0}_{j_0,L}$ and $q \in AP^{-r_0}_{j_0,L}$.

If $j_0 > J_{r_0,L} = J_{-r_0,L}$, then $p, q \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j \geq J_{r,L}}$. It follows that the sets $\Omega(t)$ and $\Omega(t')$ are both defined in case 2-1. Hence by (75),

$$\epsilon_{r_0,L} \leq \max_{1\leq i\leq n}\{|s_i| + |s'_i| + |\epsilon_i|\} \leq \tfrac{5}{6}\epsilon_{r_0,L},$$

which is a contradiction.

If the map in formula (71) fails to be injective at points (t, s) and (t', s') with $t \geq t'$, then there are integers $r_0 \in [1, TL]$, $j_0 \in [1, J(|r_0|)]$ and $\theta \in \mathbb{T}^1$ such that the points p and q satisfy

$$q = \Phi^{r_0}_{\alpha,\theta}(p), \quad p, q \notin \bigcup_{|r|=1}^{[TL]} \bigcup_{j > J_{r,L}} AP^r_{j,L}.$$

In this case, the sets $\Omega(t)$ and $\Omega(t')$ are both defined according to case (2-2). Let l_1 and l_2 as the largest integers such that

$$p \in \bigcup_{|r|=1}^{[TL]} \bigcup_{l_1 \leq j \leq J_{r,L}} AP^r_{j,L} \quad \text{and} \quad q \in \bigcup_{|r|=1}^{[TL]} \bigcup_{l_2 \leq j \leq J_{r,L}} AP^r_{j,L}.$$

In case (2-2), we have

$$|s_i| < \frac{I}{4}\frac{1}{2^{l_1+1}}, \quad |s'_i| < \frac{I}{4}\frac{1}{2^{l_2+1}}, \quad \text{for all } i \in \{n+1, \ldots, a\},$$

which also leads to a contradiction because $l_1, l_2 > j_0$ deduce the contradiction

$$\begin{aligned} \frac{I}{2^{j_0+1}} \leq \delta_{r_0,L}(p) &\leq \max_{i \geq n+1} \{|s_i| + |s'_i| + |\epsilon_i|\} \\ &< \frac{1}{4}\frac{I}{2^{l_1+1}} + \frac{1}{4}\frac{I}{2^{l_2+1}} + \left(\frac{1}{32}\right) I^2/2^{l_1+l_2+2} \leq \frac{3}{4}\frac{I}{2^{j_0+1}}. \end{aligned}$$

Hence, the injectivity is proved. □

We simply re-prove the bound on the averaged width (see [**FF14**, Lemma 5.5]) in the general settings (under transversality conditions) by combining with Lemma 5.10.

LEMMA 5.11. *For all $x \in M$ and for all $T, L \geq 1$ we have*

$$\frac{1}{w_{\mathcal{F}^{(L)}_\alpha}(x, T)} \leq \left(\frac{2}{I}\right)^a \frac{1}{T} \int_0^T H^T_L \circ \phi^t_{X^{(L)}_\alpha}(x)\, dt.$$

Proof. The width function w_Ω of the set Ω is given by

$$w_\Omega(t) = \begin{cases} \left(\dfrac{I}{2}\right)^a & \text{case 1,} \\ \left(\dfrac{I}{2}\right)^a \left(\dfrac{\min_{1 \leq |r| \leq TL}\{\epsilon_{r,L}\}}{2}\right)^n & \text{case 2-1,} \\ \left(\dfrac{I}{2}\right)^a 2^{-(a-n)(l+1)} & \text{case 2-2,} \end{cases} \tag{76}$$

and implies that

$$\frac{1}{w_\Omega(t)} \leq \begin{cases} \left(\frac{2}{I}\right)^{a-n} & \text{case 1,} \\ \left(\frac{2}{I}\right)^{a-n} \sum_{|r|=1}^{[TL]} \sum_{j > J_{r,L}} \frac{2^n \chi_{AP^r_{j,L}}(\phi^t_{X^{(L)}_\alpha}(x))}{(\epsilon_{r,L})^n} & \text{case 2-1,} \\ \left(\frac{2}{I}\right)^{a} \sum_{|r|=1}^{[TL]} \sum_{j > J_{r,L}} 2^{(j+1)(a-n)} \chi_{AP^r_{j,L}}(\phi^t_{X^{(L)}_\alpha}(x)) & \text{case 2-2.} \end{cases} \tag{77}$$

By the definition of the function H^T_L in formula (70), we have

$$\frac{1}{w_\Omega(t)} \leq \left(\frac{2}{I}\right)^a H^T_L \circ \phi^t_{X^{(L)}_\alpha}(x) \quad \text{for all } t \in [0, T]. \tag{78}$$

From the definition (40) of the average width of the orbit segment $\{x \exp(tX^{(L)}_\alpha) \mid 0 \leq t \leq T\}$, we have the estimate

$$\frac{1}{w_{\mathcal{F}^{(L)}_\alpha}(x, T)} \leq \frac{1}{T} \int_0^T \frac{dt}{w_\Omega(t)} \leq \left(\frac{2}{I}\right)^a \frac{1}{T} \int_0^T H^T_L \circ \phi^t_{X^{(L)}_\alpha}(x)\, dt. \qquad \square$$

LEMMA 5.12. *For all $r \in \mathbb{Z}\setminus\{0\}$ and for all $L \geq 1$, the following estimate holds:*

$$\left| \int_M h_{r,L}(x)\, dx \right| \leq C I^{a-n}(1 + J_{r,L}) L^{-\sum_{i=n+1}^a \rho_i}.$$

Proof. It follows from Lemma 5.8 that for $r \neq 0$ and for all $j \geq 0$, the Lebesgue measure of the set $AP^r_{j,L}$ satisfies the following bound:

$$\mathcal{L}^{a+1}(AP^r_{j,L}) \leq \frac{C I^{a-n}}{2^{j(a-n)}} L^{-\sum_{i=n+1}^a \rho_i}. \tag{79}$$

From formula (68), it follows that

$$\int_M h_{r,L}(x)\, dx \leq 1 + \sum_{j=1}^{J_{r,L}} 2^{j(a-n)} \mathcal{L}^{a+1}(AP^r_{j,L}) + \sum_{j > J_{r,L}} \frac{2^n \mathcal{L}^{a+1}(AP^r_{j,L})}{(\epsilon_{r,L})^n}.$$

By the estimate in formula (79), we immediately have that

$$\sum_{j=1}^{J_{r,L}} 2^{j(a-n)} \mathcal{L}^{a+1}(AP^r_{j,L}) \leq C I^{a-n} J_{r,L} L^{-\sum_{i=n+1}^a \rho_i}.$$

By the definition of the cut-off in formula (67) we have the bound

$$\frac{2^{n-(J_{r,L}+1)(a-n)}}{(\epsilon_{r,L})^n} \leq 1,$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

and by an estimate on a geometric sum

$$\sum_{j>J_{r,L}} \frac{2^n \mathcal{L}^{a+1}(AP^r_{j,L})}{(\epsilon_{r,L})^n} \le \frac{2^{n-(J_{r,L}+1)(a-n)}}{(\epsilon_{r,L})^n} C I^{a-n} L^{-\sum_{i=n+1}^a \rho_i}$$
$$\le C I^{a-n} L^{-\sum_{i=n+1}^a \rho_i}. \qquad \square$$

5.3. *Diophantine estimates.* We review the simultaneous Diophantine condition.

Definition 5.13. A vector $\alpha \in \mathbb{R}^n \backslash \mathbb{Q}^n$ is simultaneously Diophantine of exponent $\nu \ge 1$, say $\alpha \in DC_{n,\nu}$ if there exists a constant $c(\alpha) > 0$ such that, for all $r \in \mathbb{N}\backslash\{0\}$,

$$\min_i \|r\alpha_i\| = d(r\alpha, \mathbb{Z}^n) = \|r\alpha\| \ge \frac{c(\alpha)}{r^{(\nu/n)}}.$$

Definition 5.14. For any basis $\bar{Y} := \{\bar{Y}_1, \ldots, \bar{Y}_n\} \subset \mathbb{R}^n$, let $\bar{I} := \bar{I}(\bar{Y})$ be the supremum of all constants $\bar{I}' > 0$ such that the map

$$(s_1, \ldots, s_n) \to \exp\left(\sum_{i=1}^n s_i \bar{Y}_i\right) \in \mathbb{T}^n$$

is a local embedding on the domain

$$\{\mathbf{s} \in \mathbb{R}^n \mid |s_i| < \bar{I}' \text{ for all } i = 1, \ldots, n\}.$$

For any $\theta \in \mathbb{R}^n$, let $[\theta] \in \mathbb{T}^n$ be its projection onto the torus $\mathbb{T}^n := \mathbb{R}^n/\mathbb{Z}^n$ and let

$$|\theta|_1 = |s_1|, \ldots, |\theta|_i = |s_i|, \ldots, |\theta|_n = |s_n|,$$

if there is $\mathbf{s} := (s_1, \ldots, s_n) \in [-\bar{I}/2, \bar{I}/2]^n$ such that

$$[\theta] = \exp\left(\sum_{i=1}^n s_i \bar{Y}_i\right) \in \mathbb{T}^n;$$

otherwise we set $|\theta|_1 = \cdots = |\theta|_n = \bar{I}$.

Definition 5.15. (Counting for close return time) Let $\sigma = (\sigma_1, \ldots, \sigma_n) \in (0, 1)^n$ be such that $\sigma_1 + \cdots + \sigma_n = 1$. For any $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n$, for any $N \in \mathbb{N}$ and every $\delta > 0$, let

$$R_\alpha(N, \delta) = \{r \in [-N, N] \cap \mathbb{Z} \mid |r\alpha|_1 \le \delta^{\sigma_1}, \ldots, |r\alpha|_n \le \delta^{\sigma_n}\}.$$

For every $\nu > 1$, let $D_n(\bar{Y}, \sigma, \nu) \subset (\mathbb{R}\backslash\mathbb{Q})^n$ be the subset defined as follows: the vector $\alpha \in D_n(\bar{Y}, \sigma, \nu)$ if and only if there exists a constant $C(\bar{Y}, \sigma, \alpha) > 0$ such that for all $N \in \mathbb{N}$ for all $\delta > 0$,

$$\#R_\alpha(N, \delta) \le C(\bar{Y}, \sigma, \alpha) \max\{N^{1-(1/\nu)}, N\delta\}. \tag{80}$$

The Diophantine condition for counting return time implies a standard simultaneous Diophantine condition.

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

LEMMA 5.16. [**FF14**, Lemma 5.9] *Let $\alpha \in D_n$. For all $r \in \mathbb{Z}\backslash\{0\}$, we have*

$$\max\{|r\alpha|_1, \ldots, |r\alpha|_n\} \geq \min\left\{\frac{\bar{I}^2}{4}, \frac{1}{[1 + C(\bar{Y}, \sigma, \alpha)]^{2\nu}}\right\}\frac{1}{|r|^{\nu}}.$$

For any vector $\sigma = (\sigma_1, \ldots, \sigma_n) \in (0, 1)^n$ such that $\sigma_1 + \cdots + \sigma_n = 1$, let

$$m(\sigma) = \min\{\sigma_1, \ldots, \sigma_n\} \quad \text{and} \quad M(\sigma) = \max\{\sigma_1, \ldots, \sigma_n\}.$$

LEMMA 5.17. [**FF14**, Lemma 5.12] *For all bases $\bar{Y} \subset \mathbb{R}^n$, for all $\sigma = (\sigma_1, \ldots, \sigma_n) \in (0, 1)^n$ such that $\sigma_1 + \cdots + \sigma_n = 1$ and for all $\nu \geq 1$, the inclusion*

$$DC_{n,\nu} \subset D_n(\bar{Y}, \sigma, \nu)$$

holds under the assumption

$$\mu \leq \min\left\{\nu, \left[\frac{M(\sigma)}{\nu} + 1 - \frac{1}{n}\right]^{-1}, \left[\frac{1}{\nu} + \left(1 - \frac{2}{n}\right)\left(1 - \frac{m(\sigma)}{M(\sigma)}\right)\right]^{-1}\right\}.$$

The set $D_n(\bar{Y}, \sigma, \nu)$ has full measure if

$$\frac{1}{\nu} < \min\left\{[M(\sigma)n]^{-1}, 1 - \left(1 - \frac{2}{n}\right)\left(1 - \frac{m(\sigma)}{M(\sigma)}\right)\right\}. \tag{81}$$

In dimension 1, the vector space has unique basis up to scaling. The following result is immediate.

LEMMA 5.18. [**FF14**, Lemma 5.13] *For all $\nu \geq 1$ the following identity holds:*

$$DC_{1,\nu} = D_1(\nu).$$

Let $\mathcal{F}_\alpha := (X_\alpha, Y)$ be a basis and let $\bar{Y} = \{\bar{Y}_1, \ldots, \bar{Y}_n\} \in \mathbb{R}$ denote the projection of the basis of codimension-1 ideal $\mathfrak{I}$ onto the abelianized Lie algebra $\bar{\mathfrak{n}} := \mathfrak{n}/[\mathfrak{n}, \mathfrak{n}] \approx \mathbb{R}^n$. For $\rho = (\rho_1, \ldots, \rho_a) \in [0, 1)^a$, we write a vector of scaling exponents

$$\bar{\rho} = (\rho_1, \ldots, \rho_n), \quad |\bar{\rho}| = \rho_1 + \cdots + \rho_n.$$

Let $\alpha_1 = (\alpha_1^{(1)}, \ldots, \alpha_n^{(1)}) \in D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$. For brevity, let $C(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \alpha_1)$ denote the constant appearing in (80) for $\alpha_1 \in D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$. Let

$$C(\alpha_1) = 1 + C(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \alpha_1). \tag{82}$$

We prove the upper bound on the cut-off function in formula (67). Let $I = I(Y)$ and $\bar{I} = \bar{I}(\bar{Y})$ be the positive constant introduced in Definitions 5.7 and 5.14. We observe that $I \leq \bar{I}$ since the basis $\bar{Y}$ is the projection of the basis $Y \subset \mathfrak{n}'$ and the canonical projection commutes with the exponential map. Then the following logarithmic upper bound holds.

LEMMA 5.19. [**FF14**, Lemma 5.14] *For every $\rho \in [0, 1)^a$, for every $\nu \leq 1/|\bar{\rho}|$ and for every $\alpha \in D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \alpha)$, there exists a constant $K > 0$ such that for all $T \geq 1$ and for all $r \in \mathbb{Z}\backslash\{0\}$, the following bound holds:*

$$J_{r,L} \leq K\{1 + \log^+[I(Y)^{-1}] + \log C(\alpha_1)\}(1 + \log|r|).$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

Proof. By Lemma 5.16 and by the definition of $\epsilon_{r,L}$ in formula (62), it follows that, for all $T > 0$, $L \geq 1$ and for all $r \in \mathbb{Z}\backslash\{0\}$, we have

$$\epsilon_{r,L} \geq \max_{1\leq i\leq n} \min\{I, |r\alpha_1|_i\} \geq \min\left\{I, \frac{\bar{I}^2}{4}, \frac{1}{[1+C(\alpha_1)]^{2\nu}}\right\}\frac{1}{|r|^{\nu}}.$$

It follows by the bound above and by the definition of the cut-off function (67) that

$$J_{r,L} \leq \frac{n}{a-n}(3\log 2 + 3\log^+(1/I) + 2\nu\log[1+C(\alpha_1)] + \nu\log|r|). \qquad \square$$

Assume that there exists $\nu \in 1/|\bar{\rho}|$ such that $\alpha_1 \in D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$. For brevity, we introduce the following notation:

$$\mathcal{H}(Y, \rho, \alpha) = 1 + I(Y)^{a-n}C(\alpha_1)\{1 + \log^+[I(Y)^{-1}] + \log C(\alpha_1)\}. \qquad ((83)$$

THEOREM 5.20. [**FF14**, Theorem 5.15] *For every $\rho \in [0,1)^a$, for every $\nu \leq 1/|\bar{\rho}|$ such that $\alpha_1 = \alpha_i^{(1)} \in D_n(\bar{Y}, \bar{\rho}, \nu)$ there exists a constant $K' > 0$ such that for all $T > 0$ and for all $L \geq 1$, the following bound holds:*

$$\left|\int_M H_L^T(x)\,dx\right| \leq K'\mathcal{H}(Y, \rho, \alpha)(1+T)(1+\log^+ T + \log L)L^{1-\sum_{i=1}^a \rho_i}.$$

Proof. By the definition of H_L^T in formula (70), the statement follows from Lemmas 5.12 and 5.19. In fact, for all $r \in \mathbb{Z}\backslash\{0\}$ and $j \geq 0$, by definition (64) the set $AP_{j,L}^r$ is non-empty only if $\epsilon_{r,L} < (I/2)$. Since $\nu \leq 1/|\bar{\rho}|$, it follows from the definition of the Diophantine class D_n that

$$\#\{r \in [-TL, TL] \cap \mathbb{Z}\backslash\{0\} \mid AP_{j,L}^r \neq \emptyset\} \leq C(\bar{Y}, \sigma, \alpha_1)(1+T)L^{1-|\bar{\rho}|}.$$

Hence, the statement follows from Lemmas 5.12 and 5.19. $\square$

The main idea of the proof above follows from Lemma 5.12 which is based on the estimate to the upper bound of the measure of the almost periodic set $AP_{j,L}^r$. In Lemma 5.8, this bound is independent of choice of transverse section M_θ^a.

5.4. *Width estimates along orbit segments.* In this subsection we introduce the definition of *good points*, which is crucial in controlling the average width estimate.

Definition 5.21. For any increasing sequence (T_i) of positive real numbers, let $h_i \in [1,2]$ denote the ratio $\log T_i/[\log T_i]$ for every $T_i \geq 1$. Set $N_i = [\log T_i]$ and $T_{j,i} = e^{jh_i}$ for integer $j \in [0, N_i]$.

Let $\zeta > 0$ and $w > 0$. A point $x \in M$ is (w, T_i, ζ)-*good* for the basis $\mathcal{F}_\alpha$ if, setting $y_i = \phi_{X_\alpha}^{T_i}(x)$, then for all $i \in \mathbb{N}$ and for all $0 \leq j \leq N_i$,

$$w_{\mathcal{F}_\alpha^{(T_{j,i})}}(x, 1) \geq w/T_i^{\zeta}, \quad w_{\mathcal{F}_\alpha^{(T_{j,i})}}(y_i, 1) \geq w/T_i^{\zeta}.$$

Under the transversality condition, the following proof is still valid and extendable from the quasi-abelian case. For the completion of the subsection, we repeat the proof of the following lemma.

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

LEMMA 5.22. [**FF14**, Lemma 5.18] *Let $\zeta > 0$ be fixed and let (T_i) be an increasing sequence of positive real numbers satisfying the condition*

$$\Sigma((T_i), \zeta) := \sum_{i \in \mathbb{N}} (\log T_i)^2 T_i^{-\zeta} < \infty. \tag{84}$$

Let $\rho \in [0, 1)$ with $\sum \rho_i = 1$. Then the Lebesgue measure of the complement of the set $\mathcal{G}(w, (T_i), \zeta)$ of $(w, (T_i), \zeta)-$good points is bounded above. That is, there exists $K > 0$ such that

$$\text{meas}(\mathcal{G}(w, (T_i), \zeta)^c) \leq K \Sigma((T_i), \zeta)[1/I(Y)]^a \mathcal{H}(Y, \rho, \alpha) w.$$

Proof. For all $i \in \mathbb{N}$ and for all $j = 0, \ldots, N_i$, let

$$\mathfrak{S}_{j,i} = \{z \in M : w_{\mathcal{F}_\alpha^{(T_{j,i})}}(z, 1) < T_i^\zeta / w\}.$$

By definition we have

$$\mathcal{G}(w, (T_i), \zeta)^c = \bigcup_{i \in \mathbb{N}} \bigcup_{j=0}^{N_i} (\mathfrak{S}_{j,i} \cup \phi_{X_\alpha}^{-T_i}(\mathfrak{S}_{j,i})). \tag{85}$$

By Lemma 5.11 for all $z \in \mathfrak{S}_{j,i}$ we have

$$(I/2)^a T_i^\zeta / w < \int_0^1 H_{T_{j,i}}^1 \circ \phi^\tau_{X_\alpha^{(T_{j,i})}}(z)\, d\tau = \frac{1}{T_{j,i}} \int_0^{T_{j,i}} H_{T_{j,i}}^1 \circ \phi_{X_\alpha}^\tau(z)\, d\tau.$$

It follows that

$$\mathfrak{S}_{j,i} \subset \mathfrak{S}(j, i) := \left\{ z \in M : \sup_{J > 0} \frac{1}{J} \int_0^J H_{T_{j,i}}^1 \circ \phi_{X_\alpha}^\tau(z) d\tau > (I/2)^a T_i^\zeta / w \right\}.$$

By the maximal ergodic theorem, the Lebesgue measure $\text{meas}(\mathfrak{S}_{j,i})$ of the set $\mathfrak{S}(j, i)$ satisfies the inequality

$$\text{meas}(\mathfrak{S}_{j,i}) \leq \text{meas}(\mathfrak{S}(j, i)) = (2/I)^a (w/T_i^\zeta) \int_M H_{T_{j,i}}^1(z)\, dz.$$

Let $\mathcal{H} = \mathcal{H}(Y, \rho, \nu)$ denote the constant defined in formula (83). By Theorem 5.20, since by hypothesis $\nu \leq 1/|\bar{\rho}|$ and $\alpha \in D_n(\bar{\rho}/|\bar{\rho}|, \nu)$, there exists a constant $K' := K'(a, n, \nu) > 0$ such that the following bound holds:

$$\left| \int_M H_{T_{j,i}}^1(z)\, dz \right| \leq K' \mathcal{H}(1 + \log T_{j,i}).$$

Hence, by the definition of the $T_{j,i}$, we have

$$N_i \leq \log T_i \leq N_i + 1, \quad \log T_{j,i} \leq 2j. \tag{86}$$

Thus, for some constant K'', we have

$$\text{meas}(\mathfrak{S}_{j,i}) \leq K''(2/I)^a \mathcal{H} w(1 + j) T_i^{-\zeta}.$$

By (86), for some constant $K''' > 0$,

$$\mathrm{meas}\bigg(\bigcup_{j=0}^{N_i} \mathfrak{S}_{j,i} \cup \phi_{X_\alpha}^{-T_i}(\mathfrak{S}_{j,i})\bigg) \leq K'''(2/I)^a \mathcal{H}w(\log T_i)^2 T_i^{-\zeta}.$$

By sub-additivity of the Lebesgue measure, we derive the bound

$$\mathrm{meas}\bigg(\bigcup_{i\in\mathbb{N}} \bigcup_{j=0}^{N_i} \mathfrak{S}_{j,i} \cup \phi_{X_\alpha}^{-T_i}(\mathfrak{S}_{j,i})\bigg) \leq K'''\Sigma((T_i), \zeta)\mathcal{H}w.$$

By formula (85), the above estimate concludes the proof. □

Remark 5.23. Lemma 5.22 proves that the Lebesgue measure of the complement of a good set is bounded along the sequence $(T_i^{-\zeta})_{i\in\mathbb{N}}$ with a certain rate of decay. This explains that the Lebesgue measure of the set of $x \in M$ with a small average width along unit time (or slow divergence along close return orbits) is bounded by the ergodic integral of cut-off functions for length 1. In particular, this lemma is necessary to handle bounds of ergodic averages of flows under rescaled time in §6.2, which is a counterpart argument to the use of the Sobolev constant that is only valid for the proof of (renormalizable) Heisenberg nilflows (cf. [**FF06**]).

6. *Bounds on ergodic average*

In this section, we prove the bound for deviation of ergodic averages of nilflows, for each function on irreducible representations. This bound will be derived from the calculations for the coboundary in §3 and the average width estimates for good points in §5.

We shall introduce assumptions on coadjoint orbits $\mathcal{O} \subset \mathfrak{n}^*$.

Definition 6.1. A linear form $\Lambda \in \mathcal{O}$ is *integral* if the coefficients $\Lambda(\eta_i^{(m)})$, $(i, m) \in J$, are integer multiples of 2π. Denote by $\widehat{M}$ the set of coadjoint orbits $\mathcal{O}$ of integral linear forms Λ.

There exist coadjoint orbits $\mathcal{O} \subset \mathfrak{n}^*$ that correspond to unitary representations which do not factor through the quotient $N/\exp \mathfrak{n}_k$, $\mathfrak{n}_k \subset Z(\mathfrak{n})$. Such coadjoint orbits and unitary representations are called *maximal* (see [**FF07**, §2.2]).

Let $\widehat{M}_0$ be subset of all coadjoint orbits of forms Λ such that $\Lambda(\eta_i^{(m)}) \neq 0$ for $m = k$. This space has maximal rank and $\Lambda(\eta_i^{(m)}) \neq 0$, for all $(i, m) \in J$.

Definition 6.2. Given a coadjoint orbit $\mathcal{O} \in \widehat{M}_0$ and a linear functional $\Lambda \in \mathcal{O}$, let us denote by $\mathcal{F}_{\alpha,\Lambda}$ the completed basis obtained by completion of $(X_\alpha, \eta_*^{(1)}, \ldots, \eta_*^{(k)})$. For all $t \in \mathbb{R}$, we write the scaled basis $\mathcal{F}_{\alpha,\Lambda}(t)$ as

$$\mathcal{F}_{\alpha,\Lambda}(t) = (X_\alpha(t), Y_\Lambda(t)) = A_t^\rho(X_\alpha, Y_\Lambda).$$

For any $\mathcal{O} \in \widehat{M}_0$, let $H_\mathcal{O}$ denote the primary subspace of $L^2(M)$ which is a direct sum of subrepresentations equivalent to $\mathrm{Ind}_{N'}^N(\Lambda)$. For the adapted basis $\mathcal{F}$, set

$$W^r(H_\mathcal{O}, \mathcal{F}) = H_\mathcal{O} \cap W^r(M, \mathcal{F}).$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

6.1. *Coboundary estimates for rescaled basis.* Recall the definition of *degree* of $Y_i^{(m)}$ and

$$d_i^{(m)} = \begin{cases} k-m & \text{for all } 1 \leq m \leq k-1, \\ 0, & m = k. \end{cases}$$

For any linear functional Λ, the degree of the representation π_Λ only depends on its coadjoint orbits. We denote the scaling vector $\rho \in (\mathbb{R}^+)^J$ such that

$$\sum_{(i,j)\in J} \rho_i^{(j)} = 1 \quad \text{and} \quad \rho_i^{(j)} = 0 \quad \text{for } \deg(Y_i^{(j)}) = 0.$$

Assume that the number of bases of $\mathfrak{n}$ with degree $k-m$ is n_m. Define

$$S_\mathfrak{n}(k) := (n_1 - 1)(k-1) + n_2(k-2) + \cdots + n_{k-1}, \tag{87}$$

$$\delta(\rho) := \min_{\substack{1\leq m\leq k-1 \\ 1\leq i,j\leq n_m}} \{\rho_i^{(m)} - \rho_j^{(m+1)},\ \rho_i^{(m)} - \rho_i^{(m+1)}\}. \tag{88}$$

We have $\delta(\rho) \leq \lambda(\rho)$. This inequality is strict unless one has homogeneous scaling

$$\rho_i^{(j)} = \frac{d_j}{S_\mathfrak{n}} \quad \text{for } j \leq k-1.$$

LEMMA 6.3. *There exists a constant $C > 0$ such that, for all $r \in \mathbb{R}^+$ and for any function $f \in W^r(H_O)$, we have*

$$\sum_{(m,i)\in J} |[X_\alpha(t), Y_i^{(m)}(t)]f|_{r,\mathcal{F}_{\alpha,\Lambda}(t)} \leq Ce^{t(1-\delta(\rho))}|f|_{r+1,\mathcal{F}_{\alpha,\Lambda}(t)}.$$

Proof. For all $(m,i) \in J$, we have

$$[X_\alpha(t), Y_i^{(m)}(t)] = \sum_{l\geq 1} c_l^{(m+1)} e^{t(1-\rho_i^{(m)}+\rho_l^{(m+1)})} Y_l^{(m+1)}(t).$$

We note that $c_l^{(j)} = 0$ for $j = k$ and for some l, which is determined by the commutation relation. Setting $C = \max_{(i,j)\in J^+}\{|c_i^{(j)}|\}$,

$$\sum_{(m,i)\in J} |[X_\alpha(t), Y_i^{(m)}(t)]f|_{r,\mathcal{F}_{\alpha,\Lambda}(t)} \leq Ce^{t(1-\delta(\rho))} \sum_{(m,i)\in J^+} |Y_i^{(m)}(t)f|_{r,\mathcal{F}_{\alpha,\Lambda}(t)}. \qquad \square$$

For $x \in M$, let γ_x be the *Birkhoff average operator*

$$\gamma_x^T(f) = \frac{1}{T}\int_0^T f \circ \phi_{X_\alpha}^s(x)\, ds.$$

Consider the decomposition of the restriction of linear functional γ_x to $W_0^r(H_O, \mathcal{F}_{\alpha,\Lambda}(t))$ as an orthogonal sum $\gamma_x = D(t) + R(t) \in W_0^{-r}(H_O, \mathcal{F}_{\alpha,\Lambda}(t))$ of an X_α-invariant distribution $D(t)$ and an orthogonal complement $R(t)$.

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

THEOREM 6.4. *Let $r > (k+1)(a/2+1)+1/4$. For $g \in W^r(H_O, \mathcal{F}_{\alpha,\Lambda}(t))$ and for all $t \geq 0$, there exists a constant $C_r^{(1)} > 0$ such that*

$$|R(t)(g)| \leq C_r^{(1)} e^{(1-\delta(\rho)-(1-\lambda))t} \max\{1, \delta_O^{-4(r-1)}\} \times T^{-1}(w_{\mathcal{F}_{\alpha,\Lambda}(t)}(x,1)^{-(1/2)} + w_{\mathcal{F}_{\alpha,\Lambda}(t)}(\phi^T_{X_\alpha}(x),1)^{-(1/2)})|g|_{r,\mathcal{F}_{\alpha,\Lambda}(t)}. \tag{89}$$

Proof. Fix $t \geq 0$ and set $D = D(t)$, $R = R(t)$ for convenience. Let $g \in W^r(H_O, \mathcal{F}_{\alpha,\Lambda}(t))$. We write $g = g_D + g_R$, where g_R is the kernel of X_α-invariant distributions and g_D is orthogonal to g_R in $W^r(H_O, \mathcal{F}_{\alpha,\Lambda}(t))$. Then g_R is a coboundary and $R(g_D) = 0$. Let $f = G^{X_\alpha(t)}_{X_\alpha,\Lambda}(g_R)$. From $D(g_R) = 0$,

$$|R(g)| = |R(g_D + g_R)| = |R(g_R)| = |\gamma_x(g_R) - D(g_R)| = |\gamma_x(g_R)|. \tag{90}$$

By the Gottschalk–Hedlund argument,

$$\begin{aligned} |\gamma_x(g_R)| &= \left|\frac{1}{T}\int_0^T g \circ \phi^s_{X_\alpha}(x)\, ds\right| = \frac{1}{T}|f \circ \phi^T_{X_\alpha}(x) - f(x)| \\ &\leq \frac{1}{T}(|f(x)| + |f \circ \phi^T_{X_\alpha}(x)|). \end{aligned} \tag{91}$$

By Theorem 4.4 and Lemma 6.3, for any $\tau > a/2 + 1$, there exists a positive constant C_r such that for any $z \in M$,

$$|f(z)| \leq \frac{C_r}{w_{\mathcal{F}_{\alpha,\Lambda}(t)}(z,1)^{1/2}}(Ce^{(1-\delta(\rho))t}|f|_{\tau,\mathcal{F}_{\alpha,\Lambda}(t)} + |g|_{\tau-1,\mathcal{F}_{\alpha,\Lambda}(t)}). \tag{92}$$

By Theorem 3.12, if $r > (k+1)\tau + 1/4$, then

$$|f|_{\tau,\mathcal{F}_{\alpha,\Lambda}(t)} \leq C_{r,k,\tau} e^{-(1-\lambda)t} \max\{1, \delta_O^{-4\tau(k+1)}\}|g_R|_{r,\mathcal{F}_{\alpha,\Lambda}(t)}.$$

By orthogonality, we have $|g_R|_{r,\mathcal{F}_{\alpha,\Lambda}(t)} \leq |g|_{r,\mathcal{F}_{\alpha,\Lambda}(t)}$. □

COROLLARY 6.5. *For every $r > (k+1)(a/2+1)+1/4$, there is a constant $C_r^{(2)} > 0$ such that the following holds for every $O \in \widehat{I_0}$ and every $x \in M$:*

$$|R|_{-r,\mathcal{F}_{\alpha,\Lambda}} \leq C_r^{(2)}[1/I(Y_\Lambda)]^{a/2} \max\{1, \delta_O^{-4(r-1)}\}T^{-1}.$$

Proof. For all $x \in M$, we have

$$w_{\mathcal{F}_{\alpha,\Lambda}}(x,1) \geq \left(\frac{I(Y_\Lambda)}{2}\right)^a.$$

It follows from Theorem 6.4 applied to the orthogonal decomposition of $\gamma_x = D(0) + R(0)$. □

6.2. *Bounds on ergodic averages in an irreducible subrepresentation.* In this subsection we derive the bounds on ergodic averages of nilflows for functions in a single irreducible subrepresentation.

For brevity, let us set

$$C_r(O) = (1 + \delta_O^{-4(r-1)}). \tag{93}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

PROPOSITION 6.6. *Assume $(T_i)_{i\in\mathbb{N}}$ is an increasing sequence of positive real numbers ≥ 1. Let $0 < w < I(Y)^a$ and $\zeta > 0$. Given $r > (k+1)(a/2+1)+1/4$, there exists a constant $C_r(\rho)$ such that for every $\mathcal{G}(w,(T_i),\zeta)$-good point $x \in M$ and all $f \in W^r(H_O,\mathcal{F})$, we have*

$$\left|\frac{1}{T_i}\int_0^{T_i} f\circ\phi^t_{X_\alpha}(x)\,dt\right| \leq C_r(\rho)C_r(O)w^{-1/2}T_i^{-\delta(\rho)+\zeta+\lambda/2}|f|_{r,\mathcal{F}_\alpha}. \tag{94}$$

Proof. By group action for scaling (18), a sequence of frames $\mathcal{F}(t_j) = A^{t_j}_\rho\mathcal{F}$ is chosen with other scaling factors $\rho_i t_j$ on elements of Lie algebras Y_i. Then, as j increases from 0 to N, the scaling parameter t_j becomes larger, while the scaled length of the arc becomes shorter, approaching 1. Let $\phi^s_{X_j}$ denote the flow of the scaled vector field $e^{t_j}X = X(t_j)$.

For each $j = 0,\ldots,N$, let $\gamma = D_j + R_j$ be the orthogonal decomposition of γ in the Hilbert space $W^{-r}(H_\pi,\mathcal{F}(t_j))$ into an X_α-invariant distribution D_j and an orthogonal complement R_j. For convenience, we denote by $|\cdot|_{r,j}$ and $\|\cdot\|_{r,j}$ respectively the transversal Sobolev norm $|\cdot|_{r,\mathcal{F}(t_j)}$ and Lyapunov Sobolev norm $\|\cdot\|_{r,\mathcal{F}(t_j)}$ relative to the rescaled basis $\mathcal{F}(t_j)$.

Let us set $N_i = [\log T_i]$ and $t_{j,i} := T_{j,i} = \log T_i^{j/N_i}$ for integer $j \in [0, N_i]$. We observe that $N_i < \log T_i < N_i + 1$. For simplicity, we will omit the index $i \in \mathbb{N}$ and set $T = T_i$, $N = N_i$ for a while within the proof and lemmas of this subsection.

Our goal is to estimate $|\gamma|_{-r,\mathcal{F}_\alpha} = |\gamma|_{-r,0}$ (the norm of the distribution of unscaled bases). By the triangle inequality and Corollary 6.5,

$$\begin{aligned} |\gamma|_{-r,0} &\leq |D_0|_{-r,0} + |R_0|_{-r,0} \\ &\leq |D_0|_{-r,0} + C_r^{(2)}[1/I(Y)]^{a/2}C_r(O)T^{-1}. \end{aligned} \tag{95}$$

We now estimate $|D_0|_{-r,0}$. By definition of the Lyapunov norm and its bound (39), for $-s < -r < 0$,

$$|D_0|_{-s,0} \leq C_{r,s}\|D_0\|_{-r,0}. \tag{96}$$

Since $D_j + R_j = D_{j-1} + R_{j-1}$, observe that $D_{j-1} = D_j + R'_j$, where R'_j denotes the orthogonal projection of R_j on the space of invariant distributions $W^{-r}(H_O,\mathcal{F}(t_{j-1}))$. By definition of the Lyapunov norm,

$$\begin{aligned} \|D_{j-1}\|_{-r,j-1} &\leq \|D_j\|_{-r,j-1} + \|R'_j\|_{-r,j-1} \\ &\leq \|D_j\|_{-r,j-1} + |R'_j|_{-r,j-1} \\ &\leq \|D_j\|_{-r,j-1} + |R_j|_{-r,j-1}. \end{aligned}$$

By Lemma 6.9, equivalence of norms gives

$$\|D_{j-1}\|_{-r,j-1} \leq \|D_j\|_{-r,j-1} + C|R_j|_{-r,j}. \tag{97}$$

By Lemma 3.16, for any X_α-invariant distribution D and for all $t_j \geq t_{j-1}$,

$$\|D\|_{-r,\mathcal{F}(t_{j-1})} \leq e^{-\lambda(\rho)(t_j-t_{j-1})/2}\|D\|_{-r,\mathcal{F}(t_j)}.$$

Since $\mathcal{F}(t_j) = A_\rho^{t_j - t_{j-1}}\mathcal{F}(t_{j-1})$ and $t_j - t_{j-1} = \log T/N$, we obtain

$$\|D_j\|_{-r,j-1} \leq T^{-\lambda(\rho)/2N}\|D_j\|_{-r,j}.$$

From (97) we conclude by induction

$$\|D_0\|_{-r,0} \leq T^{-\lambda(\rho)/2}\bigg(\|D_N\|_{-r,N} + C\sum_{l=0}^{N-1} T^{(l+1)\lambda(\rho)/2N}|R_{N-l}|_{-r,N-l}\bigg). \tag{98}$$

By Lemma 6.7 and 6.8,

$$\|D_0\|_{-r,0} \leq C_r^1(\rho)C_r(O)w^{-1/2}T^{1-\delta(\rho)+\zeta/2-(1-\lambda)-\lambda(\rho)/2}. \tag{99}$$

From (95) and the above, we conclude that there exists a constant $C_r(\rho)$ such that

$$|\gamma|_{-r,\mathcal{F}} \leq C_r(\rho)C_r(O)w^{-1/2}T^{-\delta(\rho)+\zeta/2+\lambda/2}. \qquad \square$$

Here we provide the proofs of our supplementary lemmas.

LEMMA 6.7. *For any $r > a/2$, there exists a constant $C_r > 0$ such that for all good points $x \in \mathcal{G}(w, (T_i), \zeta)$, we have*

$$\|D_N\|_{-r,N} \leq C_r T^{\zeta/2}/w^{1/2}.$$

Proof. By definition of norm,

$$\|D_N\|_{-r,N} \leq |D_N|_{-r,N} \leq |\gamma|_{-r,N}.$$

It suffices to find the bound of orbit segment with respect to rescaled bases.

For all $i \in \mathbb{N}$, set $t_j = t_{j,i}$. By Definition 5.21, for $x \in \mathcal{G}(w, (T_i), \zeta)$ and $y_i = \phi_{X_\alpha}^{T_i}(x)$

$$\frac{1}{w_{\mathcal{F}_\alpha^{(t_j)}}(x, 1)} \leq T_i^\zeta/w \quad \text{and} \quad \frac{1}{w_{\mathcal{F}_\alpha^{(t_j)}}(y_i, 1)} \leq T_i^\zeta/w. \tag{100}$$

Note that the orbit segment $(\phi_{X_\alpha}^t(x))_{0\leq t\leq T}$ coincides with the orbit segment $(\phi_{X_\alpha(t_N)}^\tau(x))_{0\leq\tau\leq 1}$ of length 1 since $X_\alpha(t_N) = X_\alpha(\log T) = TX_\alpha$. Then by Theorem 4.4,

$$|\gamma|_{-r,N} \leq C_r w_{\mathcal{F}_\alpha^{(t_N)}}(x, 1)^{-1/2}.$$

Therefore, by the inequality (100),

$$w_{\mathcal{F}_\alpha^{(t_N)}}(x, 1)^{-1/2} \leq T^{\zeta/2}/w^{1/2}. \qquad \square$$

LEMMA 6.8. *For every $r > 2(k+1)(a/2+1)+1/2$, there is a constant $C_r(\rho) > 0$ such that for every good point $x \in \mathcal{G}(w, (T_i), \zeta)$, we have*

$$\sum_{l=0}^{N-1} T^{(l+1)\rho_Y/2N}|R_{N-l}|_{-r,N-l} \leq C_r^{(1)}(\rho)C_r(O)w^{-1/2}T^{1-\delta(\rho)-(1-\lambda)+\zeta/2}. \tag{101}$$

Proof. The orbit segment $(\phi^t_{X_\alpha}(x))_{0\le t\le T}$ has length $T^{l/N}$ with respect to the generator $X_\alpha(t_{N-l}) = X_\alpha((1-l/N)\log T) = T^{1-l/N}X_\alpha$. Thus, by Theorem 6.4 with $e^{(1-\delta(\rho))t_{N-l}} = T^{(1-l/N)(1-\delta(\rho))}$, we obtain

$$\begin{aligned}|R_{N-l}|_{-r,N-l} &\le C_r^{(1)}C_r(O)T^{(1-l/N)(1-\delta(\rho)-(1-\lambda))-l/N)}\\ &\quad\times\bigg(\frac{1}{w_{\mathcal{F}_\alpha^{(t_{N-l})}}(x,1)^{1/2}}+\frac{1}{w_{\mathcal{F}_\alpha^{(t_{N-l})}}(y,1)^{1/2}}\bigg)\\ &\le 2C_r^{(1)}C_r(O)w^{-1/2}T^{(1-l/N)(\lambda-\delta(\rho))-l/N+\zeta/2}.\end{aligned}$$

Let $C = 2C_r^{(1)}C_r(O)w^{-1/2}$. Since $N_i = [\log T_i]$ and $N_i \le \log T_i \le N_i + 1$, we have ${T_i}^{1/(N_i+1)} \le e \le {T_i}^{1/N_i}$. By setting $T = T_i$, we obtain

$$\begin{aligned}&\sum_{l=0}^{N-1} T^{(l+1)\rho_Y/2N}|R_{N-l}|_{-r,N-l}\\ &\le CT^{1-\delta(\rho)-(1-\lambda)+\zeta/2}\sum_{l=0}^{N-1}T^{(l+1)\rho_Y/2N}T^{-l/N(1-\delta(\rho)-(1-\lambda))-l/N}\\ &\le CT^{1-\delta(\rho)-(1-\lambda)+\zeta/2+\rho_Y/2N}\sum_{l=0}^{N-1}T^{-l/N(2-\delta(\rho)-(1-\lambda)-\rho_Y/2)}\\ &\le e^rCT^{1-\delta(\rho)-(1-\lambda)+\zeta/2}\sum_{l=0}^{\infty}e^{-l(1+\lambda-\delta(\rho)-\rho_Y/2)}.\end{aligned}$$

By (88), we have $1+\lambda-\delta(\rho)-\rho_Y/2 \ge 1-\rho_Y/2 > 1/2$, thus the geometric series converges. □

LEMMA 6.9. *There exists a constant $C := C(r) > 0$ such that, for all $j = 0, \ldots, N$,*

$$C^{-1}|\cdot|_{-r,j} \le |\cdot|_{-r,j-1} \le C|\cdot|_{-r,j}.$$

Proof. From (86), $t_j - t_{j-1} \le 2$ and observe that $\mathcal{F}(t_j) = A^{t_j-t_{j-1}}\mathcal{F}(t_{j-1})$. Passing from the frame $\mathcal{F}(t_{j-1})$ to $\mathcal{F}(t_j)$, it can be verified that distortion of the corresponding transversal Sobolev norm is uniformly bounded. □

Let

$$\widetilde{M}_0 = \bigcup_{O\in\hat{M}_0}\{\Lambda \in O \mid \Lambda \text{ integral}\} \tag{102}$$

be the collection of maximal integral coadjoint orbits.

Remark 6.10. Let $\sigma = (\sigma_1, \ldots, \sigma_n) \in (0,1)^n$ be such that $\sigma_1 + \cdots + \sigma_n = 1$. For simplicity, we choose $\sigma_i = 1/n$ from now on (see Definition 5.13 or Lemma 5.17).

THEOREM 6.11. *For any $\Lambda \in \widetilde{M}_0$, let $\nu \in [1, 1+(k/2-1)(1/n)]$. Then, for any $r > (k+1)(a/2+1)+1/4$, there exists a constant $C(\sigma,\nu)$ satisfying the following condition. For every $\epsilon > 0$ there exists a constant $K_\epsilon(\sigma,\nu) > 0$ such that, for every*

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

$\alpha_1 = (\alpha_1^{(1)}, \ldots, \alpha_n^{(1)}) \in D_n(\sigma, \nu)$ *and for every* $w \in (0, I(Y)^a]$, *there exists a measurable set* $\mathcal{G}_\Lambda(\sigma, \epsilon, w)$ *satisfying the estimate*

$$\text{meas}(\mathcal{G}_\Lambda(\sigma, \epsilon, w)^c) \leq K_\epsilon(\sigma, \nu)\bigg(\frac{w}{I(Y_\Lambda)^a}\bigg)\mathcal{H}(Y_\Lambda, \rho, \alpha). \tag{103}$$

Then for every $x \in \mathcal{G}_\Lambda(\sigma, \epsilon, w)$, *for every* $f \in W^r(H_O, \mathcal{F})$ *and* $T \geq 1$ *we have*

$$\bigg|\frac{1}{T}\int_0^T f \circ \phi^t_{X_\alpha}(x)\, dt\bigg| \leq \frac{C_r(\sigma, \nu)C_r(O)}{w^{1/2}} T^{-(1-\epsilon)(1/3S_{\mathfrak{n}}(k))}|f|_{r,\mathcal{F}_{\alpha,\Lambda}}.$$

Proof. If the coadjoint orbit O is integral and maximal with full rank, then the optimal exponent of $\rho = (\rho_i^{(m)})$ will be attained by the following homogeneous scaling:

$$\rho_i^{(j)} = \frac{d_j}{S_{\mathfrak{n}}} \quad \text{for } i \leq k.$$

Let us set $\zeta = 2\delta(\rho)/3 - \lambda/3$. Given $\epsilon > 0$ and for all $i \in \mathbb{N}$, set $T_i = i^{(1+\epsilon)\zeta^{-1}}$. Then there exists a constant $K_\epsilon(\rho) > 0$ such that

$$\Sigma(w, (T_i), \zeta) = \sum_{i\in\mathbb{N}}(\log T_i)^2 T_i^{-\zeta} \leq K_\epsilon(\rho).$$

Let $\mathcal{G} = \mathcal{G}_\Lambda(\sigma, \epsilon, w) = \mathcal{G}(w, (T_i), \zeta)$ be the set of $(w, (T_i), \zeta)$-good points for the basis $\mathcal{F}_\alpha$. Then the estimate in formula (103) follows from Lemma 5.22 and the definition of good points. By Proposition 6.6, for all $x \in \mathcal{G}$ and for every $f \in W^r(H_O, \mathcal{F})$, estimate (94) holds true. Given $T \in [T_i, T_{i+1}]$,

$$\int_0^T f \circ \phi^t_{X_\alpha}(x)\, dt = \int_0^{T_i} f \circ \phi^t_{X_\alpha}(x)\, dt + \int_{T_i}^T f \circ \phi^t_{X_\alpha}(x)\, dt = (I) + (II).$$

Let $C = C_r(\rho)C_r(O)/w^{1/2}$. The first term is estimated by formula (94):

$$(I) \leq C T_i^{1-\delta(\rho)+\zeta/2+\lambda/2}|f|_{r,\mathcal{F}_{\alpha,\Lambda}} = C T_i^{1-2\delta(\rho)/3+\lambda/3}|f|_{r,\mathcal{F}_{\alpha,\Lambda}}.$$

For the second term, let us set $\gamma = (1+\epsilon)\zeta^{-1}$ and observe that $\gamma^{-1} = \zeta(1+\epsilon)^{-1} \geq (1-\epsilon)\zeta$. We have

$$\begin{aligned}(II) &\leq (T - T_i)\|f\|_\infty \leq \beta 2^{\gamma-1}T^{1-\gamma^{-1}}\|f\|_\infty \\ &\leq C'(\rho)T^{1-(1-\epsilon)(-2\delta(\rho)/3+\lambda/3)}|f|_{r,\mathcal{F}_{\alpha,\Lambda}}.\end{aligned}$$

By the estimates on the terms (I) and (II), the proof is complete. □

Remark 6.12. If O is integral but not maximal, then the restriction of Λ factors through an irreducible representation of the $(k-1)$-step nilpotent group $N/\exp n'_k$. Then $\mathfrak{n}/\mathfrak{n}_k$ is a polarizing subalgebra for subrepresentation and reduces to the maximal integral case. Since the growth rate is determined by the scaling factors and the exponent λ is determined by the step size and number of elements, the highest exponent is obtained by the integral maximal full rank case.

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

6.3. *General bounds on ergodic averages.* In this subsection the bounds on ergodic averages for a function on a single irreducible subrepresentation obtained in §6.2 are extended to all sufficiently smooth functions.

Definition 6.13. For every $\mathcal{O} \in \widehat{M}_0$, we define $|\mathcal{O}| = \max_{\eta_i \in \mathfrak{n}_k} |\Lambda(\eta_i^{(k)})|$.

Note that $|\mathcal{O}|$ does not depend on the choice of Λ and $|\mathcal{O}| \neq 0$ by maximality. We specifically choose an element $\eta_*^{(k)}$ of degree k such that

$$|\mathcal{O}| = |\Lambda(\eta_*^{(k)})|.$$

LEMMA 6.14. *For every $\mathcal{O} \in \widehat{M}_0$ and for every $\Lambda \in \mathcal{O}$, we have*

$$I(Y_\Lambda)^{-a}\mathcal{H}(Y_\Lambda) \leq C(\alpha_1)(1 + \log C(\alpha_1))2^{a+1}.$$

Proof. The return time of the flow X_α to any orbit of the codimension-1 subgroup $N' \subset N$ is 1. Hence, by Definition 5.7, we have $I(Y_\Lambda) = 1/2$ for the basis. By (82), we have $C(\alpha_1) \geq 1$. Then, from the definition of the constant $H(Y, \rho, \alpha)$, we obtain

$$\begin{aligned} I(Y_\Lambda)^{-a}\mathcal{H}(Y) &\leq I(Y_\Lambda)^{-a} + I(Y)^{-n}C(\alpha_1)(1 + \log^+[I(Y)^{-1}] + \log C(\alpha_1)) \\ &\leq C(\alpha_1)(1 + \log C(\alpha_1))(I(Y_\Lambda)^{-a} + I(Y_\Lambda)^{-n} \log^+[I(Y)^{-1}]) \\ &\leq 2C(\alpha_1)(1 + \log C(\alpha_1))I(Y_\Lambda)^{-a}. \end{aligned}$$

□

COROLLARY 6.15. *For every $\mathcal{O} \in \widehat{M}_0$, $\Lambda \in \mathcal{O}$, $w > 0$ and $\epsilon > 0$, let*

$$w_\Lambda = w|\Lambda(\mathcal{F})|^{-2a-\epsilon}. \tag{104}$$

Then, for every $w > 0$ and $\epsilon > 0$, the set

$$\mathcal{G}(\sigma, \epsilon, w) = \bigcap_{\Lambda \in \widetilde{M}_0} \mathcal{G}_\Lambda(\sigma, \epsilon, w_\Lambda)$$

has measure greater than $1 - Cw\epsilon^{-1}$, with $C = 2^{-a+1}K_\epsilon(\sigma, \nu)C(\alpha_1)(1 + \log C(\alpha_1))$. Furthermore, if $\epsilon' < \epsilon$ we have $\mathcal{G}(\sigma, \epsilon, w) \subset \mathcal{G}(\sigma, \epsilon', w)$.

Proof. Recall that $|\Lambda(\mathcal{F})|$ is an integral multiple of 2π. By Lemma 6.14, inequality (103) and definition of w_Λ, we have

$$\begin{aligned} \mathrm{meas}(\mathcal{G}_\Lambda(\sigma, \epsilon, w_\Lambda)^c) &\leq K_\epsilon(\sigma, \nu)\left(\frac{w_\Lambda}{I(Y)^a}\right)\mathcal{H}(Y_\Lambda, \rho, \alpha) \\ &\leq C'|\Lambda(\mathcal{F})|^{-2a-\epsilon}w, \end{aligned}$$

where $C' = 2^{a+1}K_\epsilon(\sigma, \nu)C(\alpha_1)(1 + \log C(\alpha_1))$. Since $|\Lambda(\mathcal{F})| = 2\pi l$ is bounded by $(2l)^{a-1}$,

$$\begin{aligned} \sum_{\Lambda \in \tilde{M}_0} \mathrm{meas}(\mathcal{G}_\Lambda(\sigma, \epsilon, w_\Lambda)^c) &\leq 2^{-2a}wC' \sum_{l>0} \sum_{\Lambda \in \tilde{M}_0 : |\Lambda| = 2\pi l} l^{-a-\epsilon} \\ &\leq Cw \sum_{l>0} l^{-1-\epsilon} < Cw\epsilon^{-1}. \end{aligned}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

The last statement on the monotonicity of the set follows from the analogous statement in Theorem 6.11. □

In every coadjoint orbit, we will make a particular choice of a linear form to sum up estimates of the bound for each irreducible subrepresentation in terms of higher regularity of norms.

Definition 6.16. For every $\mathcal{O} \in \widehat{M}_0$, we define $\Lambda_{\mathcal{O}}$ as the unique integral linear form $\Lambda \in \mathcal{O}$ such that

$$0 \leq \Lambda(\eta_*^{(k-1)}) < |\mathcal{O}|.$$

The existence and uniqueness of $\Lambda_{\mathcal{O}}$ follow from

$$\Lambda \circ \mathrm{Ad}(\exp(tX_\alpha))(\eta_*^{(k-1)}) = \Lambda(\eta_*^{(k-1)}) + t|\mathcal{O}|,$$

and the form $\Lambda \circ \mathrm{Ad}(\exp(tX_\alpha)$ is integral for all integer values of $t \in \mathbb{R}$.

LEMMA 6.17. *There exists a constant $C(\Lambda) > 0$ such that the following holds on the primary subspace $C^\infty(H_{\mathcal{O}})$. Given a basis $\mathcal{F}_\alpha = \mathcal{F}_{\alpha,\eta} = (X_\alpha, \eta)$,*

$$|\Lambda_{\mathcal{O}}(\mathcal{F}_\alpha)|\mathrm{Id} \leq C(\Lambda)(1 + \Delta_{\mathcal{F}_\alpha})^{k/2}.$$

Proof. Let $x_0 = -\Lambda_{\mathcal{O}}(\eta_*^{(k-1)})/|\mathcal{O}|$. Then there exists a unique $\Lambda' \in \mathcal{O}$ such that $\Lambda'(\eta_*^{(k-1)}) = 0$ given by $\Lambda' = \Lambda \circ \mathrm{Ad}(e^{x_0 X_\alpha})$. The element $W \in \mathfrak{I}$ is represented in the representation as multiplication operators by the polynomials

$$P(\Lambda, W)(x) = \Lambda(\mathrm{Ad}(e^{xX_\alpha})W). \tag{105}$$

By the definition of the linear form, the identity $[X_\alpha, \eta_*^{(k-1)}] = \eta_*^{(k)}$ implies

$$P(\Lambda', \eta_*^{(k-1)})(x) = |\mathcal{O}|x.$$

From (105), we have

$$\sum_{j=1}^{k} \frac{(-x)^j}{j!} P(\Lambda', \mathrm{ad}(X_\alpha)^j W) = \Lambda'(W) \quad \text{for all } W \in \mathfrak{I}.$$

Then we obtain

$$\begin{aligned}
\Lambda'(W) &= \sum_{j=1}^{k} \frac{(-x)^j}{j!} P(\Lambda', \mathrm{ad}(X_\alpha)^j W) \\
&= \sum_{j=1}^{k} \frac{(-1)^j}{j!} \left(\frac{P(\Lambda', \eta_*^{(k-1)})}{|\mathcal{O}|} \right)^j P(\Lambda', \mathrm{ad}(X_\alpha)^j W) \\
&= |\mathcal{O}|^{1-k} \sum_{j=1}^{k} \frac{(-1)^j}{j!} P(\Lambda', \eta_*^{(k-1)}) P(\Lambda', \eta_*^{(k)})^{k-1-j} P(\Lambda', \mathrm{ad}(X_\alpha)^j W).
\end{aligned} \tag{106}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

For any $\Lambda \in \mathfrak{n}^*$ the transversal Laplacian for a basis $\mathcal{F}$ in the representation π_Λ is the operator of multiplication by the polynomial and derivative operators

$$\Delta_{\Lambda,\mathcal{F}} = \sum_{W\in\mathcal{F}} \pi_\Lambda^{X_\alpha}(W)^2 = \sum_{W\in\mathcal{F}} P(\Lambda, W)^2.$$

Hence, for any $(m, j) \in J$,

$$|P(\Lambda', \eta_j^{(m)})| \leq (1 + \Delta_{\Lambda',\mathcal{F}_{\alpha,\eta}})^{1/2}.$$

In view of (106), the constant operators $\Lambda'(\eta_j^{(m)})$ are given by polynomial and derivative expressions of degree k in the operators $P(\Lambda', \eta_j^{(m)})$. Then we obtain the estimate

$$|\Lambda'(\mathcal{F}_{\alpha,\eta})|\mathrm{Id} \leq C_1(\Lambda)(1 + \Delta_{\Lambda',\mathcal{F}_{\alpha,\eta}})^{k/2}.$$

Since the representations $\pi_{\Lambda'}$ and π_{Λ_O} are unitarily intertwined by the translation operator by x_0, and since constant operators commute with translations, we also have

$$|\Lambda'(\mathcal{F}_{\alpha,\eta})|\mathrm{Id} \leq C_1(\Lambda)(1 + \Delta_{\Lambda_O,\mathcal{F}_{\alpha,\eta}})^{k/2}.$$

Since x_0 is bounded by a constant depending only the step size k, the norms of the linear maps $\mathrm{Ad}(\exp(\pm x_0 X_\alpha))$ are bounded by a constant depending only on k. Therefore, $|\Lambda_O(\mathcal{F}_{\alpha,\eta})| \leq C_2(k)|\Lambda'(\mathcal{F}_{\alpha,\eta})|$ and we finish the proof. □

COROLLARY 6.18. *There exists a constant $C'(\Lambda)$ such that for all $O \in \widehat{M}_0$ and for any sufficiently smooth function $f \in H_O$,*

$$C_r(O) w_{\Lambda_O}^{-1/2} |f|_{r,\mathcal{F}_{\alpha,\Lambda}} \leq C'(\Lambda) w^{-1/2} |f|_{r+l,\mathcal{F}_\alpha}$$

where $l = 2k(r-1) + (1/2)ak(k+1)$.

Proof. From definition (93) we have $C_r(O) \leq (1 + |\Lambda_O(\mathcal{F}_{\alpha,\Lambda_O})|)^{l_1}$ with $l_1 = 4(r-1)$. By formula (104) and inequality $|\Lambda_O(\mathcal{F}_{\alpha,\Lambda_O})| \leq |\Lambda_O(\mathcal{F}_\alpha)|$, we have

$$C_r(O) w_{\Lambda_O}^{-1/2} \leq w^{-1/2}(1 + |\Lambda_O(\mathcal{F}_\alpha)|)^{l_2}$$

with $l_2 = l_1 + a(k+1)$. By Lemma 6.17 we have

$$(1 + |\Lambda_O(\mathcal{F}_\alpha)|)^{l_2} \leq C'(\Lambda)(1 + \Delta_{\mathcal{F}_\alpha})^{l_2 k/2}.$$ □

PROPOSITION 6.19. *Let $r > (k+1)(a+1) + 1/4$. Let $\sigma = (1/n, \ldots, 1/n) \in (0,1)^n$ be a positive vector. Let us assume that $\nu \in [1, 1 + (k/2 - 1)1/n]$ and let $\alpha \in D_n(\sigma, \nu)$. For every $\epsilon > 0$ and $w > 0$, there exists a measurable set $\mathcal{G}(\sigma, \epsilon, w)$ satisfying*

$$\mathrm{meas}(\mathcal{G}(\sigma, \epsilon, w)^c) \leq C w \epsilon^{-1} \quad \textit{with } C = 2^{-a+1} K_\epsilon(\sigma, \nu) C(\alpha_1)(1 + \log C(\alpha_1)),$$

such that for every $x \in \mathcal{G}(\sigma, \epsilon, w)$, $f \in W^r(M)$ and any $T \geq 1$ we have

$$\left| \frac{1}{T} \int_0^T f \circ \phi_{X_\alpha}^t(x)\, dt \right| \leq C w^{-1/2} T^{-(1-\epsilon)(1/3 S_{\mathfrak{n}}(k))} |f|_{r,\mathcal{F}_\alpha}. \tag{107}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

Proof. Let $\tau := r - ak/2 > (a/2+1)(k+1) + 1/4$. Let $f \in W^{\tau}(M, \mathcal{F})$ and let $f = \sum_{O \in \widehat{M}_0} f_O$ be its orthogonal decomposition onto the primary subspace H_O. Recall that for each $O \in \widehat{M}_0$, the constant w_{Λ_O} is provided in (104) and by equivalence from definition (102) the set

$$\mathcal{G}(\sigma, \epsilon, w) = \bigcap_{O \in \widehat{M}_0} \mathcal{G}_{\Lambda_O}(\sigma, \epsilon, w_{\Lambda_O})$$

has measure greater than $1 - Cw\epsilon^{-1}$ by Corollary 6.15.

If $x \in \mathcal{G}(\sigma, \epsilon, w)$, then by Theorem 6.11 and Corollary 6.18, for every $O \in \widehat{M}_0$ and all $T \geq 1$,

$$\left| \frac{1}{T} \int_0^T f_O \circ \phi^t_{X_\alpha}(x)\, dt \right| \leq C_r(\sigma, \nu) w^{-1/2} T^{-(1-\epsilon)(1/3S_{\mathfrak{n}}(k))} |f_O|_{r, \mathcal{F}_\alpha}.$$

For any $\tau > 0$ and any $\epsilon' > 0$, by Lemma 6.17 and orthogonal splitting of H_O we have

$$\begin{aligned} \left| \sum_{O \in \widehat{M}_0} |f_O|_{\tau, \mathcal{F}_\alpha} \right|^2 &\leq \sum_{O \in \widehat{M}_0} (1 + |\Lambda_O(\mathcal{F}_\alpha)|)^{-a-\epsilon'} \sum_{O \in \widehat{M}_0} (1 + |\Lambda_O(\mathcal{F}_\alpha)|)^{a+\epsilon'} |f_O|^2_{\tau, \mathcal{F}_\alpha} \\ &\leq C(a) |f|^2_{\tau + (a+\epsilon')k/2, \mathcal{F}_\alpha}. \end{aligned}$$

Thus the proof ends by linearity. □

Proof of Theorem 1.1. Under the same hypothesis of Proposition 6.19, for $i \in \mathbb{N}$ let $w_i = 1/2^i C$ and $\mathcal{G}_i = \mathcal{G}(\sigma, \epsilon, w_i)$. Set $K_\epsilon(x) = 1/{w_i}^{1/2}$ if $x \in \mathcal{G}_i \backslash \mathcal{G}_{i-1}$. By Proposition 6.19, the sets $\mathcal{G}_i$ are increasing and satisfy $\text{meas}(\mathcal{G}_i^c) \leq 1/2^i \epsilon$. Hence, the set $\mathcal{G}(\sigma, \epsilon) = \bigcup_{i \in \mathbb{N}} \mathcal{G}_i$ has full measure and the function K_ϵ is in $L^p(M)$ for every $p \in [1, 2)$. □

Proof of Corollary 1.2. Recall that the k step strictly triangular nilpotent Lie algebra $\mathfrak{n}$ has dimension $\frac{1}{2}k(k+1)$ with one-dimensional center. If the coadjoint orbit O is integral and maximal, then the optimal exponent will be attained by formula (87). Recall that the scaling vector $\rho = (\ldots, \rho_i^{(m)}, \ldots)$ is obtained by

$$S_{\mathfrak{n}}(k) = \left[(k-1)^2 + \sum_{n=1}^{k-2} n(n+1) \right] = (k-1)(k^2 + k - 3),$$

and, in particular, by choosing homogeneous scaling,

$$\rho_i^{(j)} = \frac{d_i^{(j)}}{S_{\mathfrak{n}}(k)} = \frac{k-j}{(k-1)(k^2+k-3)} \quad \text{for } i \leq j.$$

Then we verify

$$\lambda(\rho) = \delta(\rho) = \frac{1}{(k-1)(k^2+k-3)}. \tag{108}$$

By repeating the same argument for proof of Theorem 1.1, this concludes the proof of Corollary 1.2. □

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

7. *Uniform bound of the average width in the step-3 case*

In this section we prove Theorem 1.3, uniform bound of the effective equidistribution of nilflows on a strictly triangular step-3 nilmanifold. On this structure it is possible to derive a uniform bound of ergodic averages under the Roth-type Diophantine condition due to the *linear divergence* of nearby orbits. This argument is based on counting principles of close return times which substitutes the necessity of good points (see also [**F16**] for the 3-step filiform case).

7.1. *Average width function.* Let N be a step-3 nilpotent Lie group on three generators as introduced in (3). We denote its Lie algebra $\mathfrak{n}$, with its basis $\{X_1, X_2, X_3, Y_1, Y_2, Z\}$ satisfying the commutation relations

$$[X_1, X_2] = Y_1, \ [X_2, X_3] = Y_2, \ [X_1, Y_2] = [Y_1, X_3] = Z. \tag{109}$$

As introduced in §2, $\{\phi^t_V\}_{t\in\mathbb{R}}$ is a measure preserving flow generated by $V := X_1 + \alpha X_2 + \beta X_3$ and $(1, \alpha, \beta)$ satisfies the standard simultaneous Diophantine condition (Definition 5.13).

By definition of the average width (see Definition 4.2), for any $t \geq 0$ and for any $(x, T) \in M \times [1, +\infty)$ we will construct an open set $\Omega_t(x, T) \subset \mathbb{R}^6$ which contains the segment $\{(s, 0, \ldots, 0) \mid 0 \leq s \leq T\}$ such that the map

$$\begin{aligned}&\phi_x(s, x_2, x_3, y_1, y_2, z)\\ &= \Gamma x \exp(se^t V) \exp(e^{-(1/3)t} x_2 X_2 + e^{-(1/3)t} x_3 X_3 + e^{-(1/6)t} y_1 Y_1 + e^{-(1/6)t} y_2 Y_2 + zZ)\end{aligned}$$

is injective on $\Omega_t(x, T)$. Injectivity fails if and only if there exist vectors

$$(s, x_2, x_3, y_1, y_2, z) \neq (s', x'_2, x'_3, y'_1, y'_2, z')$$

such that

$$\begin{aligned}&\Gamma x \exp(s'e^t V) \exp(e^{-(1/3)t} x'_2 X_2 + e^{-(1/3)t} x'_3 X_3 + e^{-(1/6)t} y'_1 Y_1 + e^{-(1/6)t} y'_2 Y_2 + z'Z)\\ &= \Gamma x \exp(se^t V) \exp(e^{-(1/3)t} x_2 X_2 + e^{-(1/3)t} x_3 X_3 + e^{-(1/6)t} y_1 Y_1 + e^{-(1/6)t} y_2 Y_2 + zZ).\end{aligned} \tag{110}$$

Let us denote $r = s' - s$ and $\tilde{x}_i = x_i{}' - x_i$, $\tilde{y}_i = y_i{}' - y_i$ and $\tilde{z} = z' - z$. Let $c_\Gamma > 0$ denote the distance from the identity of the smallest non-zero element of the lattice Γ. Let us assume that

$$|x_i|, |x'_i|, |y_i|, |y'_i| \leq c_\Gamma/4 \tag{111}$$

so that $\tilde{x}_i, \tilde{y}_i \in [-(c_\Gamma/2), (c_\Gamma/2)]$.

For all $t \geq 0$ and $s \in [0, T]$, let us adopt the notation

$$\begin{aligned}&\tilde{x_2}(t, s) = \tilde{x_2}, \quad \tilde{x_3}(t, s) = \tilde{x_3}, \quad \tilde{y_1}(t, s) = \tilde{y_1} + e^{(5/6)t} s\tilde{x_2},\\ &\tilde{y_2}(t, s) = \tilde{y_2} + \alpha e^{(5/6)t} s\tilde{x_3} + 1/2e^{-1/2t}(x_2 x'_3 - x'_2 x_3).\end{aligned}$$

From the identity in formula (110), we derive the identity

$$\exp(re^t V) \exp(e^{-(1/3)t} \tilde{x}_2 \bar{X}_2 + e^{-(1/3)t} \tilde{x}_3 \bar{X}_3 + e^{-(1/6)t} \tilde{y}_1 \bar{Y}_1 + e^{-(1/6)t} \tilde{y}_2 \bar{Y}_2) \in x^{-1}\Gamma x.$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

Projecting the above identity on the base torus, we obtain

$$\exp(re^t\bar{V})\exp(e^{-(1/3)t}\tilde{x}_2\bar{X}_2 + e^{-(1/3)t}\tilde{x}_3\bar{X}_3) \in \overline{\Gamma}, \tag{112}$$

which implies that re^t is return time for the projected toral linear flow at most distant from $e^{-t/3}c_\Gamma/2$.

Let $R_t(x, T)$ denote the set of $r \in [-T, T]$ such that equation (112) on the projected torus has a solution $\tilde{x}_2, \tilde{x}_3 \in [-(c_\Gamma/2), (c_\Gamma/2)]$. Then for every $r \in R_t(x, T)$, the solution $\tilde{x}_i := \tilde{x}_i(r)$ of the identity in formula (110) is unique. Given $r \in R_t(x, T)$, let $\mathcal{S}(r)$ be the set of $s \in [0, T]$ such that there exists a solution of identity (110) satisfying (111).

Recall that $w_{\Omega_t(r)}(s)$ is the (inner) width function along the orbit $\phi_x(s, \cdot)$ for $s \in [0, T]$.

LEMMA 7.1. *The following average width estimate holds: for every $T > 1$, there exists a constant $C_\alpha > 0$ such that*

$$\frac{1}{T}\int_0^T \frac{ds}{w_{\Omega_t(r)}(s)} \le \frac{1024}{c_\Gamma^2}\frac{C_\alpha}{e^{(2/3)t}\|(\tilde{x}_2(r), \tilde{x}_3(r))\|}. \tag{113}$$

Proof. We approximate the average width by estimating counting close return orbits. By definition $\mathcal{S}(r)$ is a union of intervals I^* of length at most

$$\max\{c_\Gamma|\tilde{x}_2(r)|^{-1}e^{-5t/6}/2, c_\Gamma|\alpha\tilde{x}_3(r)|^{-1}e^{-5t/6}/2\}.$$

To count the number of such intervals, we will choose certain points where the distance is minimized. As long as $|\tilde{x}_2(r)| \ge e^{-5t/6}$, for each component I^* of $\mathcal{S}(r)$, there exists $s^* \in I^*$ solution of the equation $\tilde{y}_1(t, s) = \tilde{y}_1(r) + e^{(5/6)t}s\tilde{x}_2(r)$ (see Figure 2). The same argument holds for $|\tilde{x}_3(r)|$. Let $\mathcal{S}^*(r)$ be the set of all such solutions. Its cardinality can be estimated by counting points.

CLAIM. *There exists a constant $C_\alpha > 0$ such that*

$$\#\mathcal{S}^*(r) \le c_\Gamma^{-1}C_\alpha\|(\tilde{x}_2(r), \tilde{x}_3(r))\|e^{(1/6)t}T. \tag{114}$$

Proof. Note that s^* is a point which minimizes the distance between an orbit and its close return, say,

$$\min_s \max\{|\tilde{y}_1(t, s)|, |\tilde{y}_2(t, s)|\}.$$

If either distance $|\tilde{y}_1(t, s)|$ or $|\tilde{y}_2(t, s)|$ dominates the other, then it reduces to simply finding a solution to single equation. Otherwise, we assume $|\tilde{y}_1(t, s)| = |\tilde{y}_2(t, s)|$ and solve the equation for s. We distinguish the following two cases, but in either case we can restrict either $\tilde{y}_1(r) = 0$ or $\tilde{y}_2(r) = 0$ for convenience.

By solving the equation $|\tilde{y}_1(t, s)| = |\tilde{y}_2(t, s)|$, there exists $s^* \in I^*$ such that

$$s^* = \begin{cases} \dfrac{e^{-(5/6)t}(\tilde{y}_2 + 1/2e^{-(1/2)t}(x_2x_3' - x_2'x_3))}{\tilde{x}_2(r) - \alpha\tilde{x}_3(r)} & \text{if } \tilde{y}_1(t, s) = \tilde{y}_2(t, s), \\ \dfrac{e^{-(5/6)t}(\tilde{y}_2 + 1/2e^{-(1/2)t}(x_2x_3' - x_2'x_3))}{\tilde{x}_2(r) + \alpha\tilde{x}_3(r)} & \text{if } \tilde{y}_1(t, s) = -\tilde{y}_2(t, s). \end{cases}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

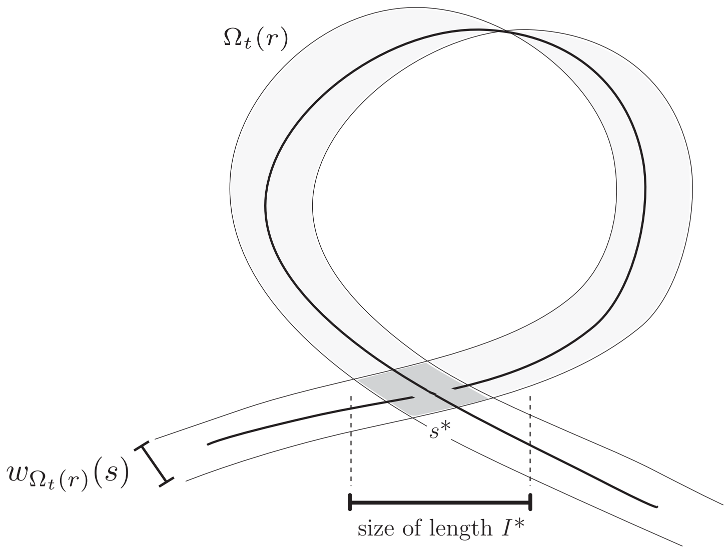

FIGURE 2. Illustration of width function and related quantities.

From the bound

$$|\tilde{y}_2 + 1/2e^{-(1/2)t}(x_2x_3' - x_2'x_3)| \leq c_\Gamma/2 + c_\Gamma^2/16,$$

we obtain

$$\#\mathcal{S}^*(r) \leq \begin{cases} 2c_\Gamma^{-1}|\tilde{x}_2(r) - \alpha\tilde{x}_3(r)|e^{(1/6)t}T & \text{if } \tilde{x}_2(r)\tilde{x}_3(r) < 0, \\ 2c_\Gamma^{-1}|\tilde{x}_2(r) + \alpha\tilde{x}_3(r)|e^{(1/6)t}T & \text{if } \tilde{x}_2(r)\tilde{x}_3(r) > 0. \end{cases}$$

Thus we prove the claim. □

For every $r \in R_t(x, T)$ and every $s \in [0, T]$, we define the function

$$\delta_r(t, s) = \begin{cases} \dfrac{1}{16}\|(\tilde{x}_2(r), \tilde{x}_3(r))\||(s - s^*)e^{5t/6}| & \text{for s} \in I^* \text{ with } |s - s^*| \geq e^{-(5/6)t}, \\ \dfrac{1}{16}\|(\tilde{x}_2(r), \tilde{x}_3(r))\| & \text{for s} \in I^* \text{ with } |s - s^*| \leq e^{-(5/6)t}, \\ \dfrac{c_\Gamma}{16} & \text{for all s} \in [0, T]\backslash\mathcal{S}(r), \end{cases}$$

and set

$$\Omega_t(r) := \{(s, x_2, x_3, y_1, y_2, z) \mid \max\{|x_2|, |x_3|, |y_1|, |y_2|\} < \delta_r(t, s), |z| < c_\Gamma/16\}.$$

Now we define the *set of narrow width* by

$$\Omega_t(x, T) := \bigcap_{r \in R_t(x,T)} \Omega_t(r).$$

Under the above construction, the map ϕ_x is injective on $\Omega_t(x, T)$. The open set $\Omega_t(r) \cap \Omega_t(-r)$ are narrowed near both endpoints of the return time r so that their images in M have no self-intersections under return times r and $-r$.

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

By the definition of inner width and by construction of the set $\Omega_t(r)$ we have that

$$w_{\Omega_t(r)}(s) = c_\Gamma \delta_r(t, s)^2 \quad \text{for all } s \in [0, T].$$

It follows that for every subinterval $I^* \subset S(r)$ we have (using the definition of δ_r)

$$\int_{I^*} \frac{ds}{w_{\Omega_t(r)}(s)} \le \frac{512 c_\Gamma^{-1}}{e^{(5t/6)} \|(\tilde{x}_2(r), \tilde{x}_3(r))\|^2}.$$

By the upper bound on the length of interval I^* and on the cardinality of the set $\mathcal{S}^*(r)$ we finally derive the conclusion. □ □

Recall from Definition 5.13 that we choose simultaneously Diophantine number $\alpha \in \mathbb{R}^2 \backslash \mathbb{Q}^2$ of exponent $\nu \ge 1$.

LEMMA 7.2. *Given the Diophantine condition of exponent $\nu \ge 1$, there exists a constant $C_\alpha := C(\alpha) > 0$ such that all solutions of formula (112) satisfy the lower bound*

$$\|(\tilde{x}_2, \tilde{x}_3)\|_{\mathbb{Z}^2} \ge C_\alpha e^{((1/3)-(\nu/2))t} r^{-(\nu/2)}.$$

Proof. By projected identity (112) on the base 3-torus, assume

$$(re^t, re^t\alpha + e^{-t/3}\tilde{x}_2, re^t\beta + e^{-t/3}\tilde{x}_3) \in \mathbb{Z}^3.$$

Then we set $re^t = q \in \mathbb{Z}$ and there exists $(p_1, p_2) \in \mathbb{Z}^2$ such that $p_1 - q\alpha = e^{-t/3}\tilde{x}_2,\ p_2 - q\beta = e^{-t/3}\tilde{x}_3$. By the Diophantine condition, there exists a constant $C(\alpha)$ such that

$$\begin{aligned} \|(\tilde{x}_2, \tilde{x}_3)\|_{\mathbb{Z}^2} &= e^{(1/3)t} \|(p_1 - q\alpha, p_2 - q\beta)\|_{\mathbb{Z}^2} \\ &= e^{(1/3)t} \|(q\alpha, q\beta)\|_{\mathbb{Z}^2} \\ &= e^{(1/3)t} \|q(\alpha)\|_{\mathbb{Z}^2} \\ &\ge C_\alpha e^{(1/3)t} q^{-(\nu/2)} \end{aligned}$$

which proves the statement. □

For every $n \in \mathbb{N}$, let $R_t^{(n)}(x, T) \subset R_t(x, T)$ be characterized by

$$\max\{|\tilde{x}_2(r)|, |\tilde{x}_3(r)|\} \in \left(\frac{c_\Gamma}{2^{n+1}}, \frac{c_\Gamma}{2^n} \right].$$

LEMMA 7.3. *For all $\epsilon > 0$, if the frequency of the projected linear flow satisfies the Diophantine condition of exponent $\nu = \sqrt{2} + \epsilon$, then there exists $C_\epsilon > 0$ such that*

$$\# R_t^{(n)}(x, T) \le C_\epsilon(\bar{V}) T \frac{c_\Gamma}{2^n} e^{(2/3)t + (\epsilon t/2)}. \tag{115}$$

Proof. Under a Diophantine condition of exponent $\nu \ge 1$, from inequality (80) and the definition of $R_t^{(n)}(x, T)$, we have

$$\# R_t^{(n)}(x, T) \le C_\nu(V) \max \left\{ (Te^t)^{1-(1/\nu)}, Te^t \frac{c_\Gamma}{2^n} e^{-(5t/6)} \right\}. \tag{116}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

It suffices to show $(Te^t)^{1-(1/\nu)}$ is less than or equal to the desired bound. From Lemma 7.2,

$$\begin{aligned}(Te^t)^{1-(1/\nu)}/\|(\tilde{x}_2,\tilde{x}_3)\| &\le (Te^t)^{1-(1/\nu)}C_\alpha e^{(-(1/3)+(\nu/2))t}r^{\nu/2}\\ &\le T^{1+(\nu/2)-(1/\nu)}C_\alpha e^{((2/3)+(\nu/2)-(1/\nu))t}.\end{aligned}$$

In the limit as $\nu \to \sqrt{2}$,

$$(Te^t)^{1-(1/\nu)}/\|(\tilde{x}_2,\tilde{x}_3)\| \le C_\alpha T e^{((2/3)+(\epsilon/2))t}.$$

Approximating $\|(x_2,x_3)\| \sim 1/2^n$,

$$(Te^t)^{1-(1/\nu)} \le C_\epsilon(\bar{V})T\frac{c_\Gamma}{2^n}e^{((2/3)+(\epsilon/2))t}. \qquad \square$$

By combining counting return time and width estimates, we obtain a uniform bound.

PROPOSITION 7.4. *There exists a constant $C_\epsilon(V) > 0$ such that*

$$\frac{1}{T}\int_0^T \frac{ds}{w_{\Omega_t(x,T)}(s)} \le C_\epsilon(V)e^{\epsilon t}.$$

Proof. By Lemma 7.1 and 7.3,

$$\begin{aligned}\frac{1}{T}\int_0^T \frac{ds}{w_{\Omega_t(x,T)}(s)} &\le \frac{1}{T}\sum_{r\in R_t(x,T)}\int_0^T \frac{ds}{w_{\Omega_t(r)}(s)}\\ &\le \frac{1}{T}\sum_{r\in R_t(x,T)}\left(\frac{1024}{c_\Gamma^2}\frac{C_\alpha}{e^{(2/3)t}\|(\tilde{x}_2(r),\tilde{x}_3(r))\|}\right)\\ &\le C_\epsilon(V)e^{\epsilon t}.\end{aligned} \qquad \square$$

Denote the *average width* of the orbit segment with length 1 by

$$w_{\mathcal{F}(t)}(x) := \sup\{w_{\mathcal{F}(t)}(y,1) \mid x \in \{y\exp(tV) \mid t\in[0,1]\}\}. \qquad (117)$$

COROLLARY 7.5. *Let ϕ_t^V be a nilflow on step-3 strictly triangular M generated by V such that the projected flow on $\mathbb{T}^3$ satisfies the Diophantine condition of Roth type with $\nu \in [1,\sqrt{2}+\epsilon]$. For every $\epsilon > 0$ there exists a constant $C_\epsilon(V) > 0$ such that*

$$w_{\mathcal{F}(t)}(x) \ge C_\epsilon(V)^{-1}e^{-\epsilon t} \quad \textit{for all } (x,t)\in M\times\mathbb{R}^+.$$

Proof of Theorem 1.3. By Corollary 7.5, we do not rely on the good points technique and Lyapunov norm. An improved bound of remainder term R in Theorem 6.4 can be obtained:

$$|R(g)|_{-r} \le C_r^{(1)}C_r(O)T^{-1} \qquad (118)$$

We revisit the backward iteration scheme introduced in the proof of Theorem 6.6. We have

$$\begin{aligned}|\gamma|_{-r,0} &\le |D_0|_{-r,0} + |R_0|_{-r,0}\\ &\le |D_0|_{-r,0} + C_r^{(1)}C_r(O)T^{-1}\end{aligned} \qquad (119)$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

and

$$|D_0|_{-r,0} \le |D_N|_{-r,0} + \sum_{j=1}^{N} |R'_{j-1}|_{-r,0}. \tag{120}$$

Changing the length to 1 and by uniform width bound from Corollary 7.5,

$$|D_N|_{-r,0} \le C_r T^{-1/12} |D_N|_{-s,\mathcal{F}(t_N)} \le C w_{\mathcal{F}(t_N)}(x,1)^{-1/2} \le C_\epsilon T^{\epsilon} \tag{121}$$

Then, by inductive argument resembling (98),

$$|D_0|_{-s,0} \le C_{r,s} T^{-1/12} \bigg(|D_N|_{-r,\mathcal{F}(t_N)} + \sum_{j=1}^{N} C_j T_j^{1/12} |R_{j-1}|_{-r,\mathcal{F}(t_{j-1})} \bigg). \tag{122}$$

Therefore

$$|\gamma|_{-r,0} \le C' C_r(O) T^{-1/12+\epsilon}.$$

Finally, we sum up all the functions on irreducible representation H_O, which only increase the regularity accordingly. □

8. *Application: mixing of nilautomorphisms*

In this section, as a further application of main equidistribution results, we verify an explicit bound for the rate of exponential mixing of hyperbolic automorphisms relying on a renormalization argument.

Let $\mathfrak{F}_{2,3} = \{X_1, X_2, Y_1, Z_1, Z_2\}$ be a step-3 free nilpotent Lie algebra with two generators with commutation relations

$$[X_1, X_2] = Y_1, \quad [X_1, Y_1] = Z_1, \quad [X_2, Y_1] = Z_2.$$

The group of automorphisms on Lie algebras induces an automorphism on the nilmanifold

$$\mathrm{Aut}(\mathfrak{n}) = \left\{ \begin{bmatrix} A & & \\ & 1 & \\ & & A \end{bmatrix}, \ A \in SL(2,\mathbb{Z}) \right\},$$

and we consider a hyperbolic automorphism T with an eigenvalue $\lambda > 1$ with corresponding eigenvector $V = X_1 + \alpha X_2$ satisfying Diophantine condition $(1, \alpha)$ on base torus $\mathbb{T}^2$. By direct computation, the following renormalization holds:

$$T \circ \exp(tV) = \exp(t\lambda V) \circ T.$$

THEOREM 8.1. *Let (ϕ_V^t) be a nilflow on a 3-step nilmanifold $M = \mathfrak{F}_{2,3}/\Gamma$ such that the projected toral flow $(\bar{\phi}_V^t)$ is a linear flow with frequency vector $v := (1, \alpha)$ in the Roth-type Diophantine condition (with exponent $\nu = 1 + \epsilon$ for all $\epsilon > 0$). For every $s > 12$, there exists a constant C_s such that for every zero-average function $f \in W^s(M)$, for all $(x, T) \in M \times \mathbb{R}$, we have*

$$\left| \frac{1}{T} \int_0^T f \circ \phi_V^t(x)\, dt \right| \le C_s T^{-1/6+\epsilon} \|f\|_s. \tag{123}$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

The detailed computation follows similarly to §7. The only difference with respect to the step-3 filiform case [**F16**] is that it has an extra element in the center which is redundant in the actual calculation on width, only raising required regularity of zero-average function.

The proposition below was first proved by Gorodnik and Spatzier in [**GS14**].

PROPOSITION 8.2. *The hyperbolic nilautomorphism T is exponential mixing.*

Proof. Let $f, g \in C^1(M)$ be smooth. Define $\langle f, g\rangle = \int_M fg d\mu$. Since the Haar measure is invariant under ϕ_V^t,

$$\langle f \circ T^n, g\rangle = \int_0^1 \langle f \circ T^n \circ \phi_V^t, g \circ \phi_V^t\rangle \, dt.$$

By integration by parts,

$$\langle f \circ T^n, g\rangle = \left\langle \int_0^1 f \circ T^n \circ \phi_V^t \, dt, g \circ \phi_V^t \right\rangle - \int_0^1 \left\langle \int_0^t f \circ T^n \circ \phi_V^s \, ds, Vg \circ \phi_V^t \right\rangle dt.$$

Therefore,

$$\langle f \circ T^n, g\rangle = (\|g\|_\infty + \|Vg\|_\infty) \int_M \sup_{s\in[0,1]} \left| \int_0^s f \circ T^n \circ \phi_V^t \, dt \right| d\mu. \tag{124}$$

By renormalizing the flow,

$$T^n \circ \phi_V^t = \phi_V^{\lambda^n t} \circ T^n$$

and

$$\int_0^s f \circ T^n \circ \phi_V^t(x) \, dt = \int_0^s f \circ \phi_V^{\lambda^n t} \circ T^n(x) \, dt = \frac{1}{\lambda^n} \int_0^{\lambda^n s} f \circ \phi_V^t \circ T^n(x) \, dt.$$

Therefore, by the result of equidistribution (123),

$$\langle f \circ T^n, g\rangle \leq \lambda^{(-1/6+\epsilon)n} \|f\|_s (\|g\|_\infty + \|Vg\|_\infty) \to 0. \tag{125}$$

□

Acknowledgement. The author deeply appreciates Giovanni Forni for his illuminating suggestions and guidance. He is grateful to Livio Flaminio and Rodrigo Treviño for making several comments which improved an earlier draft of this paper. He further thanks Krzysztof Frączek for careful advice on the revision of this work. Part of the paper was written when the author visited the Institut de Mathematiques de Jussieu-Paris Rive Gauche in Paris, France. He also acknowledges Anton Zorich for hospitality during the visit. He is grateful to Xuesen Na, Davi Obata, and Davide Ravotti for helpful discussions and thanks Jacky Jia Chong, J.T. Rustad and Lucia D. Simonelli for their encouragement. Lastly, the author is grateful to the referee for helpful comments and suggestions for improvement in the presentation of this work. This research was partially supported by

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

the NSF grant DMS 1600687 and by the Center of Excellence 'Dynamics, mathematical analysis and artificial intelligence' at Nicolaus Copernicus University in Toruń.

A. *Appendix*

In this appendix we introduce a specific example of a nilpotent Lie algebra which goes beyond our approach introduced in §5.

A.1. *Free group type of step 5 with three generators.*

In this example, we will show the failure of transversality condition. This only means that we cannot apply our theorem but we do not know whether the conclusion holds or not.

Let $\mathfrak{F}_n$ be a free nilpotent Lie algebra with n generators and $(\mathfrak{F}_n)_{k+1}$ be the $(k+1)$th subalgebra in the central series, following the notation in (4). Denote by $\mathfrak{F}_{n,k} := \mathfrak{F}_n/(\mathfrak{F}_n)_{k+1}$ the quotient of the free algebra with n generators $\mathfrak{F}_n$, which is finite-dimensional.

Definition A.1. Let $\mathfrak{n}$ be a nilpotent Lie algebra satisfying the *generalized transversality condition* if there exists basis (X_α, Y_Λ) of $\mathfrak{n}$ for each irreducible representation $\pi_\Lambda^{X_\alpha}$ such that

$$\langle \mathfrak{G}_\alpha \rangle \oplus \mathrm{Ran}(\mathrm{ad}_{X_\alpha}) + C_{\mathfrak{I}}(\pi_\Lambda^{X_\alpha}) = \mathfrak{n} \tag{A.1}$$

where $C_{\mathfrak{I}}(\pi_\Lambda^{X_\alpha}) = \{Y \in \mathfrak{I} \mid \Lambda([Y, X_\alpha]) = 0\}$.

The generalized transversality condition implies the existence of a completed basis for each irreducible representation $\pi_\Lambda^{X_\alpha}$ of non-zero degree. That is, given the adapted basis $\mathcal{F} = (X, Y_1, \ldots, Y_a)$, there exists reduced system $\bar{\mathcal{F}} = (X, Y_1', \ldots, Y_{a'}')$ satisfying transversality condition (42) and $\pi_\Lambda^X(Y_m') = 0$ for all $a' \leq m \leq a$.

Now we will investigate an example that fails the transversality condition as well as that in the sense of representation.

Let $\mathcal{F} = (X, Y_i^{(j)})$ be basis of $\mathfrak{F}_{5,3}$ with generators $\{X_1, X_2, X_3\}$ with the following relations:

$$\begin{array}{ccccc} & X_1 & X_2 & X_3 & \\ & Y_1 & Y_2 & Y_3 & \\ Z_1 & Z_2 & \cdots & Z_8 & Z_9 \end{array}$$

with

$$\begin{aligned} &[X_1, X_2] = Y_1, \quad [X_2, X_3] = Y_2, \quad [X_1, X_3] = Y_3,\\ &[X_1, Y_1] = Z_1, \quad [X_1, Y_2] = Z_2, \quad [X_1, Y_3] = Z_3,\\ &[X_2, Y_1] = Z_4, \quad [X_2, Y_2] = Z_5, \quad [X_2, Y_3] = Z_6,\\ &[X_3, Y_1] = Z_7, \quad [X_3, Y_2] = Z_8, \quad [X_3, Y_3] = Z_9. \end{aligned}$$

The rest of the elements in $\mathcal{F}_{5,3}$ are generated by some commutation relations but these are not listed here. In general, we write elements $Y_j^{(i)} \in \mathfrak{n}_i \backslash \mathfrak{n}_{i+1}$ and $Y_i^{(5)} \in Z(\mathfrak{n})$ for all i. By the Jacobi identity

$$[X_1, [X_2, X_3]] + [X_2, [X_3, X_1]] + [X_3, [X_1, X_2]] = 0 \iff Z_2 - Z_6 + Z_7 = 0.$$

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

For fixed α_i and β_i, let

$$V = X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \beta_1 Y_1 + \beta_2 Y_2 + \beta_3 Y_3$$

and set to be the $\mathfrak{I}$ ideal of $\mathfrak{F}_{5,3}$ codimension 1, not containing V.

PROPOSITION A.2. *$\mathfrak{F}_{5,3}$ does not satisfy the generalized transversality condition for some irreducible representation.*

Proof. To find the centralizer in the Lie algebra, for $a_i, b_i \in \mathbb{R}$, set

$$[V, X] = 0 \\ \iff X = a_1X_1 + a_2X_2 + a_3X_3 + b_1Y_1 + b_2Y_2 + b_3Y_3 + c_1Z_1 + \cdots + c_8Z_8.$$

Then it contains

$$(a_2 - \alpha_2 a_1)Y_1 + (\alpha_2 a_3 - \alpha_3 a_2)Y_2 + (a_3 - \alpha_3 a_1)Y_3 \\ + (b_1 - \beta_1 a_1)Z_1 + (b_2 - \beta_2 a_1)Z_2 + (b_3 - \beta_3 a_1)Z_3 + \cdots = 0.$$

By linear independence, all the coefficients vanish and we are left with

$$a_1X_1 + a_2X_2 + a_3X_3 = a_1(X_1 + \alpha_2X_2 + \alpha_3X_3),$$

$$b_1Y_1 + b_2Y_2 + b_3Y_3 = a_1(\beta_1Y_1 + \beta_2Y_2 + \beta_3Y_3).$$

Therefore, there is no non-trivial element in $C_{\mathfrak{I}}(V) \cap \mathfrak{n}_2 \backslash \mathfrak{n}_3$. Since the range of ad_V has rank 2, this model does not satisfy the transversality condition.

We now verify that the *generalized transversality condition* is not satisfied on some irreducible representation. By Schur's lemma, an irreducible representation π_Λ^V acts as a constant on the center $Z(\mathfrak{n})$.

Assume $\pi_*(W_i) = s_i I \neq 0$ for some $W_i \in Z(\mathfrak{n})$. Then it is possible to choose an element $L_i \in \mathfrak{n}_2 \backslash \mathfrak{n}_3$ such that

$$\begin{cases} \pi_*([V, L_1]) = (a_1t^2 + a_2t + a_3), \\ \pi_*([V, L_2]) = (b_1t^2 + b_2t + b_3), \\ \pi_*([V, L_3]) = (c_1t^2 + c_2t + c_3), \end{cases}$$

where (a_i, b_i, c_i) are non-proportional for each i, and

$$\pi_*(\mathrm{ad}_V^3(L_i)) = \pi_*(W_i) \neq 0.$$

However, on the given irreducible representation, any linear combination of L_1, L_2 and L_3 does not give a trivial relation. If $s_1L_1 + s_2L_2 + s_3L_3 \in C_{\mathfrak{I}}(\pi_\Lambda^V)$, then

$$\pi_*([V, s_1L_1 + s_2L_2 + s_3L_3]) \\ = s_1(a_1t^2 + a_2t + a_3) + s_2(b_1t^2 + b_2t + b_3) + s_3(c_1t^2 + c_2t + c_3) \\ = (s_1a_1 + s_2b_1 + s_3c_1)t^2 + (s_1a_2 + s_2b_2 + s_3c_2)t + (s_1a_3 + s_2b_3 + s_3c_3) = 0.$$

The system of equations has trivial solution ($t = 0$) by linear independence of each coefficient. Then there does not exist any element of $\mathfrak{n}_2 \backslash \mathfrak{n}_3$ that has degree 0. However,

https://doi.org/10.1017/etds.2021.110 Published online by Cambridge University Press

the range of ad_V has rank 2 and the generalized transversality condition cannot be satisfied in this example. □

REFERENCES

[AGH63] L. Auslander, L. Green and F. Hahn. *Flows on Homogeneous Spaces (Annals of Mathematics Studies, 53)*. Princeton University Press, Princeton, NJ, 1963, with the assistance of L. Markus and W. Massey, and an appendix by L. Greenberg.

[BDG15] J. Bourgain, C. Demeter and L. Guth. Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three. *Ann. of Math.* **184** (2016), 633–682.

[CG90] L. Corwin and F. P. Greenleaf. *Representations of Nilpotent Lie Groups and Their Applications. Part 1. Basic Theory and Examples (Cambridge Studies in Advanced Mathematics, 18)*. Cambridge University Press, Cambridge, 1990.

[DAS] B. A. Dubrovin, A. T. Fomenko and S. P. Novikov. *Modern Geometry—Methods and Applications. Part II: The Geometry and Topology of Manifolds (Graduate Texts in Mathematics, 104)*. Springer, New York, 1985.

[F16] G. Forni. Effective equidistribution of nilflows and bounds on Weyl sums. *Dyn. Anal. Number Theory* **437** (2016), 136–188.

[FF06] L. Flaminio and G. Forni. Equidistribution of nilflows and applications to theta sums. *Ergod. Th. & Dynam. Sys.* **26**(2) (2006), 409–433.

[FF07] L. Flaminio and G. Forni. On the cohomological equation for nilflows. *J. Mod. Dyn.* **1**(1) (2007), 37–60.

[FF14] L. Flaminio and G. Forni. On effective equidistribution for higher step nilflows. *Preprint*, 2014, arXiv:1407.3640.

[FFT16] L. Flaminio, G. Forni and J. Tanis. Effective equidistribution of twisted horocycle flows and horocycle maps. *Geom. Funct. Anal.* **26**(5) (2016), 1359–1448.

[GS14] A. Gorodnik and R. Spatzier. Exponential mixing of nilmanifold automorphisms. *J. Anal. Math.* **123** (2014), 355–396.

[GT12] B. Green and T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math.* **175** (2012), 465–540.

[H73] J. Humphreys. *Introduction to Lie Algebras and Representation Theory (Graduate Texts in Mathematics, 9)*. Springer, New York, 1972.

[T15] T. D. Wooley. Perturbations of Weyl sums. *Int. Math. Res. Not. IMRN* **2016**(9) (2015), 2632–2646.