

EFFECTIVE EQUIDISTRIBUTION FOR GENERALIZED HIGHER STEP NILFLOWS

MINSUNG KIM*

ABSTRACT. The main results of this paper are to 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 the polynomial speed. We analyze the rate of decay which is determined by the number of steps and structure of general nilpotent Lie algebras. Main result follows from the technique over controlling scaling operators in irreducible representations and measure estimation on close return orbit on general nilmanifolds.

CONTENTS

1. Introduction	2
2. Nilflows on higher step nilmanifold	5
2.1. Background of nilpotent Lie group and Lie algebras	5
2.2. Nilmanifold and nilflows	6
2.3. Kirillov theory and Classification	8
3. The cohomological equation	9
3.1. Distributions and Sobolev space	9
3.2. A priori estimates	10
3.3. Rescaling method	12
3.4. Scaling of invariant distribution	17
4. A Sobolev trace theorem	18
4.1. Sobolev a priori bounds	19
5. Average width estimate	21
5.1. Almost periodic points	22
5.2. Expected width bounds.	28
5.3. Diophantine estimates	33
5.4. Width estimates along orbit segments	35
6. Bounds on ergodic average	36
6.1. Coboundary estimates for rescaled basis	37
6.2. Bounds on ergodic averages in an irreducible subrepresentation.	39
6.3. General bounds on ergodic averages.	42
7. Uniform bound of the average width of step 3 case.	46
7.1. Average Width Function	46
8. Application : Mixing of nilautomorphism.	51
Appendix A.	52

Date: July 14, 2021.

2010 Mathematics Subject Classification. 37A17, 37A25, 37A44, 37A46.

Key words and phrases. Nilflows, Cohomological Equations, Ergodic averages.

A.1. Free group type of step 5 with 3 generators.	52
References	54

1. INTRODUCTION

By a general results of B. Green and T. 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. G. Forni and L. Flaminio also established estimates for the *quadratic* polynomial speed of equidistribution of nilflows on *higher step* nilmanifolds [FF14] called Quasi-abelian (Filiform). It 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 *Weyl sum*. It is notable that the bound obtained from Flaminio-Forni is established only almost everywhere but comparable with the results by T.D Wooley [T15] from the 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 of nilmanifolds under Diophantine conditions on the frequencies of their toral projections (see Definition 5.14).

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

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

For every $\nu > 1$, let $D_n(\nu) \subset (\mathbb{R} \setminus \mathbb{Q})^n$ 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 $\delta > 0$,

$$(1) \quad \#R_\alpha(N, \delta) \leq C(\alpha) \max\{N^{1-\frac{1}{\nu}}, N\delta\}.$$

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

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

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

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

Under this hypothesis, the speed of ergodic average of nilflows $(\phi_{X_\alpha}^t)$ 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 $(\phi_{X_\alpha}^t)$ be a nilflow on a k -step nilmanifold M on $n+1$ generators such that the projected toral flow $(\bar{\phi}_{X_\alpha}^t)$ is a linear flow with frequency vector $\alpha := (1, \alpha_1, \dots, \alpha_n) \in \mathbb{R} \times \mathbb{R}^n$. Assume that the Lie algebra satisfies the transversality condition and $\alpha \in D_n(\nu)$ for some $1 \leq \nu \leq \frac{k}{2}$. 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 $T \geq 1$,*

for almost all $x \in M$,

$$\left| \frac{1}{T} \int_0^T f \circ \phi_{X_\alpha}^t(x) dt \right| \leq K_\epsilon(x) T^{-\frac{1}{3S_{\mathfrak{n}}(k)} + \epsilon} \|f\|$$

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

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

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 the 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 transversality conditions.

The transversality condition enables the measure estimate (section 5) for the return orbit. This condition is sufficient, and in principle, there is no obstructions to a generalisation 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 linear in the number of steps and generators.

We can view this phenomena in the following way: if the growth of the number of elements in lower steps (generated by basis) are too large, then it lacks the dimensions to count the measure of 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 appreciated by its corollary on *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

$$(3) \quad [x_1 X_1, \dots, x_n X_n, \dots, y_i^{(j)} Y_i^{(j)} \dots z Z] := \begin{pmatrix} 1 & x_1 & \dots & \dots & z \\ 0 & 1 & x_2 & y_i^{(j)} & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 1 & x_n \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad x_i, y_i^{(j)}, z \in \mathbb{R}$$

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

Corollary 1.2. *Let $(\phi_{X_\alpha}^t)$ be a nilflow on k -step strictly triangular nilmanifold M on k generators such that the projected toral flow $(\bar{\phi}_{X_\alpha}^t)$ is a linear flow with frequency vector $\alpha := (1, \alpha_1, \dots, \alpha_{k-1}) \in \mathbb{R} \times \mathbb{R}^{k-1}$. Under the condition that $\alpha \in D_n(\nu)$ for some $1 \leq \nu \leq \frac{k}{2}$, 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 $T \geq 1$,

$$\left| \frac{1}{T} \int_0^T f \circ \phi_{X_\alpha}^t(x) dt \right| \leq K_\epsilon(x) T^{-\frac{1}{3(k-1)(k^2+k-3)} + \epsilon} \|f\|.$$

We also establish the uniform bound for step-3 strictly triangular nilmanifold case. The result holds for *all points* by estimating the width with counting close return time directly under Roth-type Diophantine condition. The step-3 case (as well as the filiform case, [F16]) is a good example to derive a simplified proof.

Theorem 1.3. *Let (ϕ_X^t) be a nilflow on 3-step nilmanifold M on 3 generators such that the projected toral flow $(\bar{\phi}_X^t)$ is a linear flow with frequency vector $v := (1, \alpha, \beta)$ 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*

$$\left| \frac{1}{T} \int_0^T f \circ \phi_X^t(x) dt \right| \leq C_s T^{-1/12 + \epsilon} \|f\|_s.$$

In the last section, we present exponential mixing of hyperbolic nilautomorphism as a main application. Exponential mixing of ergodic automorphism and its applications to the Central Limit Theorem on compact nilmanifolds was proven by A. Gorodnik and R. 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 depending on the structure of nilmanifolds, which follows from equidistribution results and renormalization argument of partial hyperbolic automorphism (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 3 generators.)

Open problem. Our work leaves open natural questions.

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

Problem 2. Give an effective bound of ergodic averages for any nilmanifolds.

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 under general higher step cases.

This paper is organized as follows. In section 2, we define structures of nilmanifolds and nilflows. In section 3, we carry out 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 section 4, we introduce the notion of average width and prove a Sobolev trace theorem. In section 5, we prove an effective equidistribution theorem for good points by a Borel-Cantelli argument. In section 6, we prove bounds on the average width of an orbit segment of nilflows

by gluing all the irreducible representations. In section 7, we introduce a uniform width estimate under a Roth-type Diophantine condition based on counting return time directly to avoid good points argument. Finally, in section 8, as an application, we prove the mixing of nilautomorphism.

Acknowledgement. The author deeply appreciates Giovanni Forni for his illuminating suggestions and guidance. He is grateful to Livio Flaminio and Rodrigo Treviño for giving several comments to improve the draft. Part of the paper was written when the author visited the Institut de Mathématiques de Jussieu-Paris Rive Gauche in Paris, France. He also acknowledges Anton Zorich for hospitality during the visit. He is grateful to Davi Obata, Davide Ravotti and Xuesen Na for helpful discussion and thanks Jacky Jia Chong, Lucia D. Simonelli, and J.T Rustad for their encouragement. Lastly, the author is thankful to the referee for his/her helpful comments and suggestions for improvement in the presentation of this work. This research was partially supported by the NSF grant DMS 1600687 and by the Centre of Excellence “Dynamics, mathematical analysis and artificial intelligence” at Nicolaus Copernicus University in Toruń.

2. NILFLOWS ON HIGHER STEP NILMANIFOLD

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

2.1. Background of nilpotent Lie group 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 co-compact lattice in N . The quotient N/Γ is then a compact nilmanifold M which N acts on the left by translations. Denote μ the N -invariant measure on M .

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

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

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 Theorem 1.1.13 and 5.16 of [CG90]). That is, given a basis

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

with $\xi := \xi^{(1)} \in \mathfrak{n}_1 \setminus \mathfrak{n}_2$ and $\eta_i^{(l)} \in \mathfrak{n}_l \setminus \mathfrak{n}_{l+1}$ for $i = 1, \dots, n_l$, we have

- (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 $\eta_1^{(j)}, \dots, \eta_{n_j}^{(j)}, \dots, \eta_1^{(k)}, \dots, \eta_{n_k}^{(k)}$ form a basis of an ideal \mathfrak{n}_j of \mathfrak{n} ;
- (3) The lattice Γ is generated by $\{x, y_1^{(1)}, \dots, y_{n_k}^{(k)}\}$ 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{J} where

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

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

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

A *strongly adapted basis* $(X, Y) := (X, Y_1, \dots, Y_a)$ is an adapted basis such that the following holds:

- (1) the system (X, Y_1, \dots, Y_n) is a system of generators of \mathfrak{n} , hence its projection is a basis of the Abelianisation $\mathfrak{n}/[\mathfrak{n}, \mathfrak{n}]$ of the Lie algebra \mathfrak{n} ;
- (2) The system (Y_{n+1}, \dots, Y_a) is a basis of the ideal $[\mathfrak{n}, \mathfrak{n}]$.

2.2. Nilmanifold 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 $pr_1 : M \rightarrow \mathbb{T}^{n+1}$.

We introduce two fibrations of 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

$$(5) \quad 0 \rightarrow M_2 \rightarrow M \xrightarrow{pr_1} \mathbb{T}^{n+1} \rightarrow 0.$$

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

$$(6) \quad 0 \rightarrow M_\theta^a \rightarrow M \xrightarrow{pr_2} \mathbb{T}^1 \rightarrow 0.$$

On 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

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

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

$$\psi_X^t(x_1, \dots, x_{n+1}) = (x_1 + tv_1, \dots, x_{n+1} + tv_{n+1}).$$

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

We recall the following:

Theorem 2.2. [AGH63] *The followings are equivalent.*

- (1) The nilflow (ϕ_X^t) on M is ergodic.
- (2) The nilflow (ϕ_X^t) on M is uniquely ergodic.
- (3) The nilflow (ϕ_X^t) on M is minimal.
- (4) The projected flow $(\psi_{\bar{X}}^t)$ on $\bar{M} = N^{ab}/\Gamma^{ab} \simeq \mathbb{T}^{n+1}$ is an irrational linear flow.

Notation. Consider the set of indices

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

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

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

and equivalently we write

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

For $\theta \in \mathbb{T}^1$ let $M_\theta^a = pr_2^{-1}(\theta)$ denote fiber over $\theta \in \mathbb{T}^1$ of the fibration pr_2 . Transverse section M_θ^a of the nilflow $\{\phi_{X_\alpha}^t\}_{t \in \mathbb{R}}$, $\mathbf{s} = (s_i)_{i=1}^a \in \mathbb{R}^a$ corresponds to

$$\{\Gamma \exp(\theta \xi) \exp(\sum_{i=1}^a s_i \eta_i) \mid (s_i) \in \mathbb{R}^a\} = \{\Gamma \exp(e^{ad(\theta \xi)} \sum_{i=1}^a s_i \eta_i) \exp(\theta \xi) \mid (s_i) \in \mathbb{R}^a\}.$$

Lemma 2.3. *The flow $(\phi_{X_\alpha}^t)_{t \in \mathbb{R}}$ on M is isomorphic to the suspension of its first return map $\Phi_{\alpha, \theta} : M_\theta^a \rightarrow M_\theta^a$. For every $(i, j) \in J$, there exists a polynomial $p_{i, N}^{(j)}(\alpha, \mathbf{s})$ for $\mathbf{s} \in \mathbb{R}^a$ such that return map $\Phi_{\alpha, \theta}$ is given by the following:*

In the coordinate of $\mathbf{s} = (s_i^{(j)})$ for $\Gamma \exp(\theta \xi) \exp(\sum_{(i,j) \in J} s_i^{(j)} \eta_i^{(j)}) \in M_\theta^a$,

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

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

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

Proof. By (7), we have

$$\exp\left(\sum_{(i,j) \in J} s_i^{(j)} \eta_i^{(j)}\right) \exp(X_\alpha) = \exp\left(\sum_{(i,j) \in J} s_i^{(j)} \eta_i^{(j)}\right) x^{-1} \exp\left(\sum_{(i,j) \in J} \alpha_i^{(j)} \eta_i^{(j)}\right) \\ = x^{-1} \exp(e^{ad(\xi)} \sum_{(i,j) \in J} s_i^{(j)} \eta_i^{(j)}) \exp\left(\sum_{(i,j) \in J} \alpha_i^{(j)} \eta_i^{(j)}\right).$$

By Baker-Campbell-Hausdorff formula, there exist polynomial $p_i^{(j)}(\alpha, \mathbf{s})$ with

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

Since $x \in \Gamma$, we conclude

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

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

2.3. Kirillov theory and Classification. Kirillov's theory yields the complete classification of irreducible unitary representation of N . All the irreducible unitary representation 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 polarizing subalgebra for a linear form $\Lambda \in \mathfrak{n}^*$. Then, the character $\chi_{\Lambda, \mathfrak{m}} : \exp \mathfrak{m} \rightarrow S^1$ is defined

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

Given pair (Λ, \mathfrak{m}) , we associate unitary representation

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

where induced representation π_Λ is defined by

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

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 irreducible, unitary representation H_π , which occur with at most finite multiplicity

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

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

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

We recall that a vector $v \in H_\pi$ is 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 semi-norms

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

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

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

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 $\mathfrak{n} = \mathbb{R}X \oplus \mathfrak{J}$ with its codimension 1 ideal \mathfrak{J} , 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 derived representation of π_* of the induced representation $\pi = \text{Ind}_{N'}^N(\pi')$ has the following description.

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 $Y \in \mathfrak{n}'$,

$$\begin{aligned} (\pi_*(Y)f)(x) &= \sum_{j=0}^{d_Y} \frac{1}{j!} \pi'_*(\text{ad}_X^j(Y)f)(x) \\ (13) \quad &= \iota P_Y(x)f(x) = \iota \sum_{j=0}^{d_Y} \frac{1}{j!} (\Lambda \circ \text{ad}_X^j Y) x^j f(x). \end{aligned}$$

We define its degree $d_Y \in \mathbb{N}$ with respect to the representation $\pi_*(Y)$ to be the degree of polynomial. Let (d_1, \dots, d_a) be the degrees of the elements (Y_1, \dots, Y_a) respectively. The degree of 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 a (priori) Sobolev estimate on the Green's operator for the cohomological equation $Xu = f$ of nilflow with generator X . We estimate bounds of Green's operator on Sobolev norm and on scaling of invariant distributions.

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

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

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

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

Equivalently,

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

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 $C^\infty(H_\pi)$ the space of C^∞ vectors of the irreducible unitary representation π . Following 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 operator $(I + \pi_*(\Delta_{\mathcal{F}}))^{\frac{\sigma}{2}}$. I.e, for every $f \in C^\infty(H_\pi)$ and $\sigma > 0$,

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

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 a 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

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

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

By Lemma 3.5 of [FF07], each invariant distribution D has a Sobolev order equal to $1/2$, i.e $D \in W^{-\sigma}(H_\pi)$ for any $\sigma > 1/2$.

For all $\sigma > 1$, let $\mathcal{K}^\sigma(H_\pi) = \{f \in W^\sigma(H_\pi) \mid D(f) = 0 \in C^\infty(H_\pi), \text{ for any } D \in W^{-\sigma}(H_\pi)\}$ be the kernel of the X -invariant distribution on the Sobolev space $W^\sigma(H_\pi)$. The Green's operator $G_X : C^\infty(H_\pi) \rightarrow C^\infty(H_\pi)$ 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 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 $\ell : C^\infty(H_\pi) \rightarrow \mathbb{C}$ such that for every function $f \in C^\infty(H_\pi) \subset L^2(\mathbb{R}, H')$,*

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

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

$$(15) \quad \int_{\mathbb{R}} \ell(f(t)) dt = \ell \left(\int_{\mathbb{R}} f(t) dt \right).$$

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 a 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 a multiplication of polynomial $p_i(t)$ on $L^2(\mathbb{R}, H')$. By definition of Sobolev norm in representation, there exists a non-zero constant $C := C(\chi, p_1, \dots, p_s) = \max_{\substack{t \in \mathbb{R}, j_1 + \dots + j_d = s \\ 0 \leq j_i \leq s, 1 \leq i \leq d \leq s}} \{\chi(t)p_1^{j_1}(t) \cdots p_d^{j_d}(t)\}$ 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 $f(t) - \chi(t) \left(\int_{\mathbb{R}} f(x)dx \right)$ has zero average, hence it is a coboundary with smooth transfer function. Since distribution D is invariant under translation, we obtain

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

and

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

Furthermore,

$$\int_{\mathbb{R}} \ell(f(t))dt = \int_{\mathbb{R}} D(f_\chi(t))dt = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \chi(x)f(t)dx \right) dt = \int_{\mathbb{R}} f(t) \left(\int_{\mathbb{R}} \chi(x)dx \right) dt.$$

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

$$\int_{\mathbb{R}} \ell(f(t))dt = \int_{\mathbb{R}} f(t)dt = D(f) = \ell \left(\int_{\mathbb{R}} f(t)dt \right).$$

□

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

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

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*

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

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

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

The renormalization group $\{A_t^{\rho}\}$ preserves the set of all adapted basis. However, it is not a group of automorphism of the Lie algebra. 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 $\mathcal{F} = (X, Y)$, *rescaled basis* $\mathcal{F}(t) = (X(t), Y(t)) = \{e^t X, e^{-\rho_1 t} Y_1, \dots, e^{-\rho_a t} Y_a\}$ of \mathcal{F} is a basis of Lie algebra \mathfrak{n} satisfying (18).

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

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

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

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

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

It is coefficient appearing in (13) and set

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

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 $\text{ad}(X)$ of \mathfrak{n}' to $\mathfrak{U}(\mathfrak{n}')$. From 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{J} is a codimension 1 ideal of \mathfrak{n} used in §2.3.

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

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

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

Now, we will 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 enveloping algebra $\mathfrak{U}(\mathfrak{n})$.

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

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

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

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

and

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

□

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. 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 $AB \geq 0$.¹ Also, if $0 \leq A \leq B$, then $A^2 \leq B^2$.²

Recall that for any positive operators A and B ,

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

Lemma 3.10. *For any $r \geq 1$ and $a \geq 1$, there exists constant $C(a, r) > 0$ such that*

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

¹By 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,

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

² $B^2 - A^2 = B(B - A) + (B - A)A$.

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

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

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

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

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

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

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

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

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

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

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

This inequality is proved by showing that

$$\begin{aligned} &(Y_i(t)^4 - Y_j(t)^4) (Y_i(t)^{4r} - Y_j(t)^{4r}) \\ &= (Y_i(t)^4 - Y_j(t)^4)^2 \left(\sum_{l=0}^{r-1} (Y_i(t)^4)^l (Y_j(t)^4)^{r-1-l} \right) \geq 0. \end{aligned}$$

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

Then, by (26),

$$\begin{aligned} &\left(\sum_{i=1}^a Y_i(t)^{4r} \right) \left(\sum_{i=1}^a Y_i(t)^4 \right) \\ &= \left(\sum_{i=1}^a Y_i(t)^{4r+4} + \sum_{i < j} (Y_i(t)^{4r} Y_j(t)^4 + Y_j(t)^{4r} Y_i(t)^4) \right) \\ &\leq (a + 1) \left(\sum_{i=1}^a Y_i(t)^{4r+4} \right). \end{aligned}$$

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

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

Therefore, induction holds and we finish the proof. \square

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

Theorem 3.11. *For $r > 1$, let $s > 2r(k + 1) + 1/2$. For any $f \in \mathcal{K}^s(M)$, there exists $C_{r,k,s} > 0$ such that the following holds: for any $t > 0$,*

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

Proof. Firstly, we estimate the bound of Green's operator with Sobolev norm.

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

$$\pi_*(Y(t)) = 2\pi\iota\delta_{\mathcal{O}}(t)x\text{Id}_{H'}.$$

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

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

$$\begin{aligned} & \|\pi_*(Y(t))^l G_{X(t)}f\|^2 \\ & \leq \int_0^\infty \left(|2\pi\delta_{\mathcal{O}}(t)x|^l \int_x^\infty e^{-t} \|f(s)\|_{H'} ds \right)^2 dx + \int_{-\infty}^0 \left(|2\pi\delta_{\mathcal{O}}(t)x|^l \int_{-\infty}^x e^{-t} \|f(s)\|_{H'} ds \right)^2 dx \\ & \leq \int_0^\infty |2\pi\delta_{\mathcal{O}}(t)x|^{2l} \int_x^\infty e^{-2t} \|f(s)\|_{H'}^2 ds dx + \int_{-\infty}^0 |2\pi\delta_{\mathcal{O}}(t)x|^{2l} \int_{-\infty}^x e^{-2t} \|f(s)\|_{H'}^2 ds dx. \end{aligned}$$

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

$$(27) \quad \int_{-\infty}^x \|f(s)\|_{H'}^2 ds \leq \left(\int_{-\infty}^x (1 + 4\pi^2\delta_{\mathcal{O}}(t)^2 s^2)^{-\alpha} ds \right) \|(I - Y(t)^2)^{\frac{\alpha}{2}} f\|^2.$$

For all $\alpha > 1$, we set

$$(28) \quad \begin{aligned} C_{\alpha,l}^2 &= \int_0^\infty (2\pi x)^{2l} \left(\int_x^\infty (1 + (4\pi^2 s^2))^{-(l+\alpha)} ds \right) dx \\ &+ \int_{-\infty}^0 (2\pi x)^{2l} \left(\int_{-\infty}^x (1 + (4\pi^2 s^2))^{-(l+\alpha)} ds \right) dx < \infty. \end{aligned}$$

By Hölder's inequality and change of variables for $x' = \delta_{\mathcal{O}}(t)x$ and $s' = \delta_{\mathcal{O}}(t)s$

$$(29) \quad \begin{aligned} \|\pi_*(Y(t))^l G_{X(t)}f\| &\leq C_{\alpha,l} \left(\frac{e^{-t}}{\delta_{\mathcal{O}}(t)} \right) \left\| \pi_*((I - Y(t)^2)^{\frac{l+\alpha}{2}} f) \right\| \\ &= C_{\alpha,l} e^{-(1-\rho_Y)t} \left(\frac{1}{\delta_{\mathcal{O}}} \right) \left\| \pi_*((I - Y(t)^2)^{\frac{l+\alpha}{2}} f) \right\|. \end{aligned}$$

Let $E_x : C^\infty(H_\pi) \rightarrow C^\infty(H')$ be linear operator defined by $E_x f = f(x)$. Let $L(t)$ be a rescaled element of $L \in \mathcal{F}$. By definition of representation (13) and by Lemma 3.8, for any $L(t)$ there exists $Q_i(t) \in \mathfrak{U}(\mathfrak{n})$ such that

$$(30) \quad E_x \pi_*(L(t)) = \sum_{j=0}^{[L]} \frac{1}{j!} \pi_*(Q_j(t)) \pi_*(Y(t))^j E_x.$$

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

$$\begin{aligned}
 E_x \pi_*(L(t)) G_{X(t)} &= \sum_{j=0}^{[L]} \frac{1}{j!} \pi_*(Q_j(t)) \pi_*(Y(t))^j E_x \\
 &= \sum_{j=0}^{[L]} \frac{1}{j!} \pi_*(Y(t))^j \int_{-\infty}^x \pi_*(Q_j(t)) E_s ds \\
 &= \sum_{j=0}^{[L]} \frac{1}{j!} \pi_*(Y(t))^j \int_{-\infty}^x E_s \pi_*(Q_j(t)) ds \\
 &= \sum_{j=0}^{[L]} \frac{1}{j!} \pi_*(Y(t))^j E_x G_{X(t)} (\pi_*(Q_j(t))). \\
 &= E_x \sum_{j=0}^{[L]} \frac{1}{j!} \pi_*(Y(t))^j G_{X(t)} (\pi_*(Q_j(t))).
 \end{aligned}
 \tag{31}$$

Combining (29) and (31), there exists a constant $C'_{\alpha,j} > 0$ such that

$$\begin{aligned}
 \|\pi_*(L(t)) G_{X(t)}(f)\| &\leq \sum_{j=0}^{[L]} \frac{1}{j!} \|\pi_*(Y(t))^j G_{X(t)} (\pi_*(Q_j(t)) f)\| \\
 &\leq \sum_{j=0}^{[L]} C'_{\alpha,j} e^{-(1-\rho_Y)t} \left(\frac{1}{\delta_{\mathcal{O}}} \right)^{j+1} \left\| \pi_*((I - Y(t)^2)^{\frac{j+\alpha}{2}} \pi_*(Q_j(t)) f) \right\|.
 \end{aligned}$$

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

$$\pi_*(L(t))^{2r} = \left(\sum_{j=0}^{[L]} \frac{1}{j!} \pi_*(Y(t))^j \pi_*(Q_j(t)) \right)^{2r} := \sum_{j=0}^{2r[L]} \pi_*(Y(t))^j \pi_*(R_j(t)).
 \tag{32}$$

Especially, $R_j(t)$ is product of $Q_i(t)$'s and $[R_j(t)] = 2r([L] + 1) - j$. Specifically, to compute transverse Laplacian, here we assume that $L(t) = (Y_i(t))^{4r}$ for each element $Y_i \in \mathcal{F}$. Then, by (32), for any $\alpha > 1$

$$\begin{aligned}
 &\|\pi_*(Y_i(t)^{4r}) G_{X(t)}(f)\| \\
 &\leq \sum_{j=0}^{4r[Y_i]} C(\alpha, i, j, r) e^{-(1-\rho_Y)t} \left(\frac{1}{\delta_{\mathcal{O}}} \right)^{j+1} \left\| \pi_*((I - Y(t)^2)^{\frac{j+\alpha}{2}} \pi_*(R_j(t)) f) \right\| \\
 &\leq C(\alpha, i, r) e^{-(1-\rho_Y)t} \max\{1, \delta_{\mathcal{O}}^{-4r[Y_i]+1}\} |f|_{\alpha+4r([Y_i]+1), \mathcal{F}(t)}.
 \end{aligned}
 \tag{33}$$

Therefore, combining with Lemma 3.10

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

Since $[\Delta^r] \leq 2kr$, there exists $C' = C'(a, \alpha, r, X) > 0$ such that

$$|G_{X(t)}f|_{2r, \mathcal{F}(t)} \leq C' e^{-(1-\rho_Y)t} \max\{1, \delta_{\mathcal{O}}^{-1}\} |f|_{\alpha+4(k+1)r, \mathcal{F}(t)}.$$

By interpolation, for all $s > 2r(k+1) + 1/2$, there exists a constant $C_{r,s} := C_{r,s}(k, X) > 0$ such that

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

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

3.4. Scaling of invariant distribution. In this section, we introduce the Lyapunov norm and compare bounds between Sobolev dual norm and 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 $U_t : L^2(\mathbb{R}, H') \rightarrow L^2(\mathbb{R}, H')$ be the unitary operator defined as follows:

$$(34) \quad (U_t f)(x) = e^{-\frac{\lambda}{2}t} f(e^{-\lambda t} x).$$

We consider the comparison of the norm estimate on the scaling of invariant distributions

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

in terms of $|D|_{-r, \mathcal{F}}$.

Theorem 3.12. *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 3.11. By Lemma 3.8, there exists $(i-j+1)$ th order $Q_j \in \mathfrak{U}(\mathfrak{n})$ with

$$U_t^{-1} L(t) U_t = x^i Q_i + e^{-\lambda t} x^{i-1} Q_{i-1} + e^{-2\lambda t} x^{i-2} Q_{i-2} + \cdots + e^{-i\lambda t} Q_0.$$

Then there exists $C > 0$ such that

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

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

$$(35) \quad |U_t f|_{1, \mathcal{F}(t)} \leq C_1 |f|_{k+1, \mathcal{F}}.$$

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

$$(36) \quad |U_t f|_{r, \mathcal{F}(t)} \leq C_{r,s} |f|_{s, \mathcal{F}}.$$

\square

Theorem 3.13. *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*

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

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

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

Then by unitarity (36),

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

and

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

□

Definition 3.14 (Lyapunov norm). For any basis \mathcal{F} and all $\sigma > 1/2$, define Lyapunov norm

$$(37) \quad \|D\|_{-\sigma, \mathcal{F}} := \inf_{\tau \geq 0} e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}\tau} |D|_{-\sigma, \mathcal{F}(\tau)}.$$

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

Lemma 3.15. For all $t \geq 0$, we have

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

Proof. By definition of the norm,

$$\begin{aligned} \|D\|_{-\sigma, \mathcal{F}} &= \inf_{\tau \geq 0} e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}\tau} |D|_{-\sigma, \mathcal{F}(\tau)} \\ &= e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}t} \inf_{\tau+t \geq 0} e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}\tau} |D|_{-\sigma, \mathcal{F}(t+\tau)} \leq e^{-\frac{\lambda_{\mathcal{F}}(\rho)}{2}t} \|D\|_{-\sigma, \mathcal{F}(t)}. \end{aligned}$$

□

We conclude this section by introducing useful inequality that follows from the Theorem 3.13,

$$(38) \quad C_{r,s}^{-1} |D|_{-s, \mathcal{F}} \leq \|D\|_{-r, \mathcal{F}} \leq |D|_{-r, \mathcal{F}}.$$

4. A SOBOLEV TRACE THEOREM

In this section, the notion of the average width (39) 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, 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 rescaled basis. For any $x \in M$, let $\phi_{x,t} : \mathbb{R} \times \mathbb{R}^a \rightarrow M$ be the local embedding defined by

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

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

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

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

Proof. Let $\mathbf{s} + h_i$ denote sequence with $(\mathbf{s} + h)_i = s_i + h$ and $(\mathbf{s} + h_i)_j = s_j$ if $i \neq j$. By definitions,

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

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

For fixed i and $j > i$,

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

By Campbell-Hausdorff formula, we set

$$= \exp(s_i Y_j(t)) \exp(s_i Y_j(t)) \exp(h(Y_i(t) + [Y_i(t), s_j Y_j(t)])) + O(h^2)).$$

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

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

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

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

It concludes the proof by choosing $S_i = Y_i(t) + \sum_{l>i}^a q_l(s, t) Y_l(t)$. Also, commutation in rescaled elements $[Y_i(t), s_j Y_j(t)] = s_j e^{-\rho t} Y_k(t)$ implies $q_l(s, t)$ include exponential term with negative exponent such that it decreases on $t \geq 0$. \square

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

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

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

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

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 $w_\Omega : \mathbb{R} \rightarrow [0, 1]$ defined as follows:

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

Consider the family $\mathcal{O}_{x,t,T}$ of open sets $\Omega \subset \mathbb{R} \times \mathbb{R}^a$ satisfying two conditions

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

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

$$(39) \quad w_{\mathcal{F}(t)}(x, T) = \sup_{\Omega \in \mathcal{O}_{x,t,T}} \left(\frac{1}{T} \int_0^T \frac{ds}{w_\Omega(s)} \right)^{-1}.$$

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

Lemma 4.2. [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 $C_\sigma > 0$ such that for all functions $F \in C^\infty(\Omega)$ and all $\tau \in I$, we have*

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

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

Theorem 4.3. *For all $\sigma > a/2$, there is a constant $C_\sigma > 0$ such that the following holds.*

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

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

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

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

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

Operators on both sides are essentially self-adjoint,

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

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

$$(40) \quad \|(I - \Delta_{\mathbb{R}^a})^{\frac{\sigma}{2}} f \circ \phi_{x,t}\|_{L^2(\Omega)}^2 \leq C_\sigma \|(I - \Delta_{\mathcal{F}(t)})^{\frac{\sigma}{2}} f\|_{L^2(M)}^2.$$

By Lemma 4.2, we can see that for $\sigma > a/2$, setting $F(\tau, 0) = f \circ \phi_{X(t)}^\tau(x)$

$$\begin{aligned} \left| \frac{1}{T} \int_0^T f \circ \phi_{X(t)}^\tau(x) dt \right|^2 &= \left(\frac{1}{T} \int_0^T |F(\tau, 0)|^2 d\tau \right)^2 \\ &\leq C_\sigma \frac{1}{T} \left(\frac{1}{T} \int_0^T \frac{ds}{w_\Omega(s)} \right) \int_\Omega |(I - \Delta_{\mathbb{R}^a})^{\frac{\sigma}{2}} F(\tau, s)| d\tau ds \\ &\leq C_\sigma T^{-1} w_{\mathcal{F}(t)}(x, T)^{-1} \|(I - \Delta_{\mathcal{F}(t)})^{\frac{\sigma}{2}} f\|_{L^2(M)}^2. \end{aligned}$$

□

5. AVERAGE WIDTH ESTIMATE

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

Strategy.

- (1) In section 5.1, we introduce basic settings. Since the flow commutes with centralizer, we take quotient map to obtain local diffeomorphism. It is remarkable to see that we set tubular neighborhood to consider close return orbits on the quotient space. This contributes calculating the measure of close return so called almost periodic set. (See (5.3) and Lemma 5.3.)
- (2) The range of the differential of displacement map coincides with the range of adjoint map ad_{X_α} . This is one main reason that it requires a necessity of transverse condition. Without this condition, there could be a direction that return orbit that does not reach on the transverse manifold, which fails the idea that the set of almost periodic point should have a small measure up to rescaling vector ρ . (See Lemma 5.7.)
- (3) In section 5.2, we prove a bound of average width by ergodic averages of cut-off functions. By definition 5.8, we classify the type of close return orbits by growth of local coordinates. This leads that the width of function does not vanish on such set and the map is injective under the restricted domain. (See Lemma 5.9 and 5.10.)
- (4) Finally, in section 5.3 and 5.4, we follow the known estimate from [FF14], which are necessary for proving bounds of ergodic averages in the section 6. In particular, Definition 5.20 of good point means the width along transverse direction to the flow cannot be too small and we prove the complement of the set of good points has a small measure. (See Lemma 5.21.)

5.1. Almost periodic points. Let X_α be the vector field on M defined in (7). Recall the 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 estimate.

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

$$(41) \quad \langle \mathfrak{G}_\alpha \rangle + \text{Ran}(\text{ad}_{X_\alpha}) + C_{\mathfrak{J}}(X_\alpha) = \mathfrak{n}$$

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

It is clear that the set of generators are neither included in the range of ad_{X_α} , nor in the centralizer $C_{\mathfrak{J}}(X_\alpha)$. We will restrict \mathfrak{n} satisfying the condition (41) in the rest of sections.

Remark. The transversality condition implies that displacement (or distance between x and $\Phi_{\alpha,\theta}^r(x)$), induced by return map $\Phi_{\alpha,\theta}$, should intersect the set of centralizer transversally. I.e the measure of the set of close return orbit in transverse manifold M_θ^a should not be invariant under the action of flow. This condition is crucial in estimating the almost periodic orbit (53) under rescaling of basis in the Lemma 5.7.

Recall that M_θ^a denotes the fiber at $\theta \in \mathbb{T}^1$ of the fibration $pr_2 : M \rightarrow \mathbb{T}^1$. $\Phi_{\alpha,\theta}$ denote the first return map of nilflow $\{\phi_{X_\alpha}^t\}$ to the transverse section M_θ^a and $\Phi_{\alpha,\theta}^r$ denote r -th iterate of the map $\Phi_{\alpha,\theta}$. Let G denote nilpotent Lie group with its lattice Γ defining $M_\theta^a = \Gamma \backslash G$. G acts on M_θ^a by right action and action of G extends on $M_\theta^a \times M_\theta^a$.

Define a map $\psi_{\alpha,\theta}^{(r)} : M_\theta^a \rightarrow M_\theta^a \times M_\theta^a$ given by $\psi_{\alpha,\theta}^{(r)}(x) = (x, \Phi_{\alpha,\theta}^r(x))$. By its definition, the map $\Phi_{\alpha,\theta}^r$ commutes with the action of the centralizer $C_G = \exp(C_{\mathfrak{J}}(X_\alpha)) \subset G$ and its action on product $M_\theta^a \times M_\theta^a$ commutes with $\psi_{\alpha,\theta}^{(r)}$. That is, for $c \in C_G$ and $x = \Gamma g$,

$$(42) \quad \psi_{\alpha,\theta}^{(r)}(xc) = (xc, \Phi_{\alpha,\theta}^r(xc)) = (xc, \Phi_{\alpha,\theta}^r(x)c) = \psi_{\alpha,\theta}^{(r)}(x)c.$$

Then quotient map is well-defined on

$$(43) \quad \Psi_{\alpha,\theta}^{(r)} := M_\theta^a / C_G \longrightarrow M_\theta^a \times M_\theta^a / C_G.$$

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

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

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

$$(44) \quad (w_1, w_2) = 1/2(w_1 + w_2, w_1 + w_2) + 1/2(w_1 - w_2, -(w_1 - w_2)).$$

(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

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

Then tangent space of 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 (44), for $w_1 = v$ and $w_2 = -d_x \Phi_{\alpha, \theta}^r(v)$, we write

$$(45) \quad (T_{(x, \Phi_{\alpha, \theta}^r(x))}\Delta_{(x, \Phi_{\alpha, \theta}^r(x))})^\perp = \{(1/2(v - d_x \Phi_{\alpha, \theta}^r(v)), -1/2(v - d_x \Phi_{\alpha, \theta}^r(v)) \mid v \in T_x M_\theta^a\}.$$

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

$$(46) \quad T_{\pi(x, y)}\Delta_{\pi(x, y)} = T_{(x, y)}\Delta_{(x, y)}, \quad T_{\pi(x, y)}\Delta_{\pi(x, y)}^\perp = T_{(x, y)}\Delta_{(x, y)}^\perp.$$

Define a map $F^{(r)} : M_\theta^a \rightarrow M_\theta^a \times M_\theta^a$ given by $F^{(r)} = \pi \circ \psi_{\alpha, \theta}^{(r)}$. In the local coordinate, by identification (45) and (46),

$$(47) \quad d_x F^{(r)}(v) = (1/2(v - d_x \Phi_{\alpha, \theta}^r(v)), -1/2(v - d_x \Phi_{\alpha, \theta}^r(v)), \quad v \in T_x M_\theta^a.$$

By (42) and 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 $F_C^{(r)} : M_\theta^a/C_G \rightarrow M_\theta^a \times M_\theta^a/C_G$. From (47), the range of differential $DF_C^{(r)}$ is determined by $I - D\Phi_{\alpha, \theta}^r$.

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

Lemma 5.2. *For all $r \in \mathbb{Z} \setminus \{0\}$, range of $I - D\Phi_{\alpha, \theta}^r$ on $\mathfrak{I}/C_3(X_\alpha)$ coincides with $\text{Ran}(\text{ad}_{X_\alpha})$ and Jacobian of $F_C^{(r)}$ is non-zero constant.*

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

$$\begin{aligned} \exp(tY_i^{(j)}) \exp(rX_\alpha) &= \exp(rX_\alpha) \exp(-rX_\alpha) \exp(tY_i^{(j)}) \exp(rX_\alpha) \\ &= \exp(rX_\alpha) \exp(e^{-r(\text{ad}_{X_\alpha})}(tY_i^{(j)})) \end{aligned}$$

and

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

By definition, $\frac{\partial \Phi_{\alpha, \theta}^r}{\partial s_i^{(j)}}(x) = \frac{d}{dt}(\gamma_{i, j}^x(t)) \big|_{t=0}$ and we have $I - D\Phi_{\alpha, \theta}^r = I - \sum_{(i, j) \in J} \frac{\partial \Phi_{\alpha, \theta}^r}{\partial s_i^{(j)}} \frac{\partial}{\partial s_i^{(j)}}$. Then,

$$(48) \quad (I - D\Phi_{\alpha, \theta}^r) \left(\sum_{(i, j) \in J} s_i^{(j)} Y_i^{(j)} \right) = [r(\text{ad}_{X_\alpha}) \left(\sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} (\text{ad}_{X_\alpha})^k \right)] \left(\sum_{(i, j) \in J} s_i^{(j)} Y_i^{(j)} \right).$$

Therefore, range of $I - D\Phi_{\alpha, \theta}^r$ is contained in $\text{Ran}(\text{ad}_{X_\alpha})$.

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

$$(49) \quad (I - D\Phi_{\alpha,\theta}^r) \left(\left(\frac{1 - e^{-\text{ad}_{X_\alpha}}}{\text{ad}_{X_\alpha}} \right)^{-1} \left(\sum_{(i,j) \in J} s_i^{(j)} Y_i^{(j)} \right) \right) = r(\text{ad}_{X_\alpha}) \left(\sum_{(i,j) \in J} s_i^{(j)} Y_i^{(j)} \right).$$

Therefore, we conclude that range of $I - D\Phi_{\alpha,\theta}^r$ is $\text{Ran}(\text{ad}_{X_\alpha})$.

If $\sum_{(i,j) \in J} s_i^{(j)} Y_i^{(j)} \in C_{\mathfrak{J}}(X_\alpha)$, then $(I - D\Phi_{\alpha,\theta}^r)(\sum_{(i,j) \in J} s_i^{(j)} Y_i^{(j)}) = 0$ and kernel of $I - D\Phi_{\alpha,\theta}^r$ is $C_{\mathfrak{J}}(X_\alpha)$. I.e, $I - D\Phi_{\alpha,\theta}^r$ is bijective on $\mathfrak{J}/C_{\mathfrak{J}}(X_\alpha)$. Thus, by (48) Jacobian of $I - D\Phi_{\alpha,\theta}^r$ is non-zero constant and it concludes the statement. \square

Setting (continued). (iv) Set submanifold $\mathcal{S} \subset M_\theta^a \times M_\theta^a$ that consists of diagonal Δ and coordinates of generators in normal (transverse) directions. Denote its quotient $\mathcal{S}_C = \mathcal{S}/C_G \subset M_\theta^a \times M_\theta^a/C_G$. Then, following Lemma 5.2, we obtain transversality of $F_C^{(r)}$ to \mathcal{S}_C . For every $p \in (F_C^{(r)})^{-1}(\mathcal{S}_C)$, the transversality holds on tangent space:

$$(50) \quad T_{F_C^{(r)}(p)} \mathcal{S}_C + DF_C^{(r)}(T_p M_\theta^a/C_G) = T_{F_C^{(r)}(p)}(M_\theta^a \times M_\theta^a/C_G).$$

(v) Denote Lebesgue measure $\mathcal{L}^{a+1}(= \text{vol}_M)$ on nilmanifold M and conditional measure $\mathcal{L}_\theta^a(= \text{vol}_{M_\theta^a})$ on transverse manifold M_θ^a . On quotient space M_θ^a/C_G , we write measure $\mathcal{L}_\theta^c(= \text{vol}_{M_\theta^a/C_G})$. Similarly, we set conditional measure $\mu_\theta^a(= \text{vol}_{M_\theta^a \times M_\theta^a})$ on product manifold and $\mu_\theta^c(= \text{vol}_{M_\theta^a \times M_\theta^a/C_G})$ on its quotient space.

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

For any open set $U_{\mathcal{S}_C} \subset M_\theta^a \times M_\theta^a/C_G$, we write push-forward measure $(F_C^{(r)})_* \mathcal{L}_\theta^c$

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

By compactness of M_θ^a (or M_θ^a/C_G), the above expression is finite. By Lemma 5.2, Jacobian of $F_C^{(r)}$ is constant and $(F_C^{(r)})_* \mathcal{L}_\theta^c = \mu_{\theta,r}^c$ is Lebesgue.

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

$$(51) \quad \mu_{\theta,r}^c(U_{\mathcal{S}_C} \cap M_{\theta,r,C}^a) = \mu_{\theta,r}^a(U_{\mathcal{S}} \cap M_{\theta,r}^a)$$

and by definition of conditional measure,

$$(52) \quad \mu_{\theta,r}^a(U_{\mathcal{S}} \cap M_{\theta,r}^a) = \mu_\theta^a(U_{\mathcal{S}}).$$

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

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

$$(53) \quad AP^r(U_\delta) := \{x \in M_\theta^a \mid d(F^{(r)}(x), \mathcal{S}) < \delta\}.$$

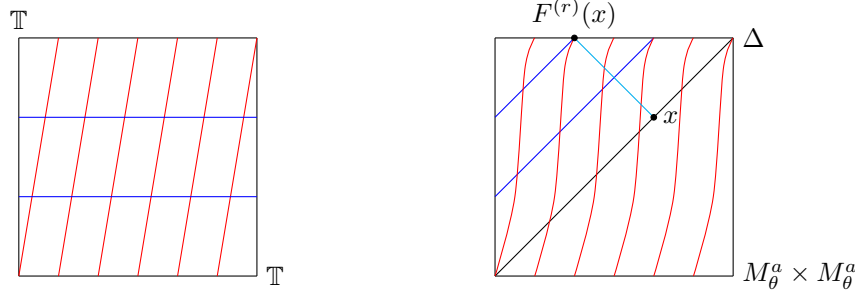


FIGURE 5.1. Illustration of displacement $F^{(r)}$ in product $M_\theta^a \times M_\theta^a$ and comparison with uniform expanding map.

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

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

Lemma 5.3. *Let $U_{\delta,C}$ be any tubular neighborhood of \mathcal{S}_C in $M_\theta^a \times M_\theta^a/C_G$. For all $r \in \mathbb{Z} \setminus \{0\}$, the conditional measure $\text{vol}_{M_\theta^a}$ of $AP^r(\mathcal{U}_\delta)$ is given as follows:*

$$\mathcal{L}_\theta^a(AP^r(\mathcal{U}_\delta)) = \mu_{\theta,r}^c(U_{\delta,C} \cap M_{\theta,r,C}^a).$$

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

Then, $(F_C^{(r)})^{-1}(AP^r(\mathcal{U}_\delta)/C_G) = (U_{\delta,C} \cap M_{\theta,r,C}^a)$. Thus, by definition of push-forward measure, the equality holds. \square

Recall that $F_C^{(r)} : M_\theta^a/C_G \rightarrow M_{\theta,r,C}^a$ has non-zero constant Jacobian if $r \neq 0$ by Lemma 5.2 and it is a local diffeomorphism. Thus, by transversality of $F_C^{(r)}$, in a small tubular neighborhood U , $F_C^{(r)}$ is covering.

Lemma 5.4. *For any $z \in U \cap M_{\theta,r,C}^a$, there exist finite number of pre-images of $F_C^{(r)}$.*

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

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

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

Proposition 5.5. *The following volume estimate holds: for any $r \neq 0$, there exists $C := C(M_\theta^a) > 0$ such that*

$$\mu_{\theta,r}^c(U_{\delta,C} \cap M_{\theta,r,C}^a) < C\delta.$$

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

$$(54) \quad \text{vol}_{M_\theta^a \times M_\theta^a / C_G}(U_{\delta,C}) = \delta \text{vol}_{M_\theta^a \times M_\theta^a / C_G}(U).$$

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

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

Note that $F_C^{(r)}|_{J_k}^{-1} \circ g_s$ is a lift of homotopy g_s , and lift of g_0 is $\gamma = F_C^{(r)}|_{J_k}^{-1}(\tilde{\gamma})$ with fixed end points x_1 and x_2 . By continuity of homotopy, 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, γ is homotopic to a constant. Since end points of γ is fixed, it has to be a constant but it 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:

$$(55) \quad \text{vol}_{M_\theta^a / C_G}(J_k) = \frac{\text{vol}_{M_{\theta,r,C}^a}(U \cap M_{\theta,r,C}^a)}{\text{Jac}(F_C^{(r)}|_{J_k})} = \frac{\text{vol}_{M_{\theta,r,C}^a}(U \cap M_{\theta,r,C}^a)}{\text{Jac}(F_C^{(r)})}.$$

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

$$(56) \quad N_r \left(\frac{\text{vol}_{M_\theta^a \times M_\theta^a / C_G}(U)}{\text{Jac}(F_C^{(r)})} \right) = N_r \left(\frac{\text{vol}_{M_{\theta,r,C}^a}(U \cap M_{\theta,r,C}^a)}{\text{Jac}(F_C^{(r)})} \right)$$

$$(57) \quad = \sum_{k=1}^{N_r} \text{vol}_{M_\theta^a / C_G}(J_k) < \infty.$$

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 (55) and definition of conditional measure,

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

$$= N_r \left(\frac{\text{vol}_{M_\theta^a \times M_\theta^a / C_G}(U_{\delta, C})}{\text{Jac}(F_C^{(r)})} \right).$$

By previous last equality with condition (54),

$$(58) \quad \text{vol}_{M_\theta^a / C_G}((F_C^{(r)})^{-1}(U_{\delta, C})) = N_r \left(\frac{\delta \text{vol}_{M_\theta^a \times M_\theta^a / C_G}(U)}{\text{Jac}(F_C^{(r)})} \right).$$

Therefore, combining (56) and (58), there exists $C > 0$ such that

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

□

Definition 5.6. For any basis $Y = \{Y_1, \dots, Y_a\}$ of codimension 1 ideal \mathfrak{J} of \mathfrak{n} , let $I := I(Y)$ be the supremum of all constant $I' \in (0, \frac{1}{2})$ such that for any $x \in M$ the map

$$(59) \quad \phi_x^Y : (s_1, \dots, s_a) \mapsto x \exp\left(\sum_{i=1}^a s_i Y_i\right) \in M$$

is local embedding (injective) on the domain

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

For any $x, x' \in M$, set local distance d_* (measured locally in the Lie algebra) on transverse section M_θ^a along Y_i direction by $d_{Y_i}(x, x') = |s_i|$ if there is $\mathbf{s} := (s_1, \dots, s_a) \in [-I/2, I/2]^a$ 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 $pr_1 : M \rightarrow \mathbb{T}^{n+1}$ onto the base torus. On transverse manifold, for all $\theta \in \mathbb{T}^1$, let $pr_\theta : M_\theta^a \rightarrow \mathbb{T}^n$ be the restriction to M_θ^a . Then, by applying formula (10) to projection to base torus,

$$(60) \quad d_{Y_i}(pr_\theta(\Phi_{\alpha, \theta}^r(x)), pr_\theta(x)) = r\alpha_i, \quad 1 \leq i \leq n.$$

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

$$(61) \quad \begin{aligned} \epsilon_{r, L} &:= \max_{1 \leq i \leq n} \min\{I, L^{\rho_i} d_{Y_i}(\Phi_{\alpha, \theta}^r(x), x)\} \\ \delta_{r, L}(x) &:= \max_{n+1 \leq i \leq a} \min\{I, L^{\rho_i} d_{Y_i}(\Phi_{\alpha, \theta}^r(x), x)\}. \end{aligned}$$

We note the distance $d_{Y_i}(\Phi_{\alpha, \theta}^r(x), x)$ on the generator level does not depend on choice of x . For this reason, we split the cases $\epsilon_{r, L}$ and $\delta_{r, L}(x)$ for step ≥ 2 .

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

$$(62) \quad \Phi_{\alpha, \theta}^r(x) = x \exp\left(\sum_{i=1}^a s_i Y_i\right)$$

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

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

For every $r \in \mathbb{Z} \setminus \{0\}$ and $j \geq 0$, let $AP_{j,L}^r \subset M$ be sets defined as follows

$$(63) \quad AP_{j,L}^r = \begin{cases} \emptyset & \text{if } \epsilon_{r,L} > \frac{I}{2}; \\ (\delta_{r,L})^{-1} ((2^{-(j+1)}I, 2^{-j}I]) & \text{otherwise.} \end{cases}$$

In the next lemma, Lebesgue measure of the set of almost-periodic points $AP_{j,L}^r$ on M is estimated by the volume of δ -neighborhood \mathcal{U}_δ .

Lemma 5.7. *For all $r \in \mathbb{Z} \setminus \{0\}$, $j \in \mathbb{N}$, $L \geq 1$, the $(a+1)$ dimensional Lebesgue measure of the set $AP_{j,L}^r$ can be estimated as follows: there exists $C > 0$ such that*

$$\mathcal{L}^{a+1}(AP_{j,L}^r) \leq \frac{CI^{a-n}}{2^{j(a-n)}} L^{-\sum_{i=n+1}^a \rho_i}.$$

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

By Tonelli's theorem,

$$(64) \quad \mathcal{L}^{a+1}(AP_{j,L}^r) = \int_0^1 \mathcal{L}_\theta^a(AP_{j,L}^r \cap M_\theta^a) d\theta.$$

Recall the definition $AP^r(\mathcal{U}_\delta)$ in (53). Choose $\delta = \frac{I^{a-n}}{2^{j(a-n)}} L^{-\sum_{i=n+1}^a \rho_i}$ and set $\mathcal{U}_\delta^{L,j} := \mathcal{U}_\delta$. Then we claim that $AP_{j,L}^r \cap M_\theta^a \subset AP^r(\mathcal{U}_\delta^{L,j})$.

By identification of $x \in M_\theta^a$ to (x, x) in diagonal $\Delta \subset M_\theta^a \times M_\theta^a$, local distance $d_{Y_i}(\Phi_{\alpha,\theta}^r(x), x)$ is identified by the distance function d in the product $M_\theta^a \times M_\theta^a$. Thus, $x \in AP_{j,L}^r \cap M_\theta^a$ implies that $d(F^{(r)}(x), \mathcal{S}) < \delta$. That is, $AP_{j,L}^r \cap M_\theta^a \subset AP^r(\mathcal{U}_\delta^{L,j})$. By Lemma 5.3, the volume estimate follows

$$\mathcal{L}_\theta^a(AP_{j,L}^r \cap M_\theta^a) \leq \mathcal{L}_\theta^a(AP^r(\mathcal{U}_\delta^{L,j})) = \mu_{\theta,r}^c(\mathcal{U}_{\delta,C}^{L,j} \cap M_{\theta,r,C}^a).$$

Finally, by Proposition 5.5,

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

Thus the proof follows from formula (64). \square

5.2. Expected width bounds. We prove a bound on the average width of a orbit on nilmanifold with respect to scaled basis. This section follows in the same way of [FF14, §5.2]. For completion of the proof, we repeat the similar arguments in nilmanifolds under transverse conditions.

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

$$(65) \quad h_{r,L} = \sum_{j=1}^{\infty} \min\{2^{j(a-n)}, (\frac{2}{\epsilon_{r,L}})^n\} \chi_{AP_{j,L}^r}.$$

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

$$(66) \quad J_{r,L} := \max\{j \in \mathbb{N} \mid 2^{j(a-n)} \leq (\frac{2}{\epsilon_{r,L}})^n\}.$$

The function $h_{r,L}$ is

$$(67) \quad h_{r,L} = \sum_{j=1}^{J_{r,L}} 2^{j(a-n)} \chi_{AP_{j,L}^r} + \sum_{j > J_{r,L}} (\frac{2}{\epsilon_{r,L}})^n \chi_{AP_{j,L}^r}.$$

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

$$(68) \quad \mathcal{F}_\alpha^{(L)} = (X_\alpha^{(L)}, Y_1^{(L)}, \dots, Y_a^{(L)}) = (LX_\alpha, L^{-\rho_1}Y_1, \dots, L^{-\rho_a}Y_a).$$

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

$$\gamma_{X_\alpha^{(L)}}^T(x) := \{\phi_{X_\alpha^{(L)}}^t(x) \mid 0 \leq t \leq T\}.$$

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

$$(69) \quad H_L^T := 1 + \sum_{|r|=1}^{[TL]} h_{r,L}.$$

Definition 5.8. For $t \in [0, T]$, we define a set of points $\Omega(t) \subset \{t\} \times \mathbb{R}^a$ as follows:

Case 1. If $\phi_{X_\alpha^{(L)}}^t(x) \notin \bigcup_{|r|=1}^{[TL]} \bigcup_{j>0} AP_{j,L}^r$, let $\Omega(t)$ be the set of all points (t, s_1, \dots, s_a) such that

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

If $\phi_{X_\alpha^{(L)}}^t(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j>0} AP_{j,L}^r$, then we consider two subcases.

Case 2-1. if $\phi_{X_\alpha^{(L)}}^t(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j>J_{r,L}} AP_{j,L}^r$, let $\Omega(t)$ be the set of all points (t, \mathbf{s}) such that

$$\begin{aligned} |s_i| &< \frac{1}{4} \min_{1 \leq |r| \leq [TL]} \min_{j>J_{r,L}} \{\epsilon_{r,L} : \phi_{X_\alpha^{(L)}}^t(x) \in AP_{j,L}^r\}, \text{ for } i \in \{1, \dots, n\} \\ |s_i| &< \frac{I}{4} \quad \text{for } i \in \{n+1, \dots, a\}. \end{aligned}$$

Case 2-2. if $\phi_{X_\alpha^{(L)}}^t(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{j \leq J_{r,L}} AP_{j,L}^r \setminus \bigcup_{|r|=1}^{[TL]} \bigcup_{j>J_{r,L}} AP_{j,L}^r$, let l be the largest integer such that

$$\phi_{X_\alpha^{(L)}}^t(x) \in \bigcup_{|r|=1}^{[TL]} \bigcup_{l \leq j \leq J_{r,L}} AP_{j,L}^r \setminus \bigcup_{|r|=1}^{[TL]} \bigcup_{j>J_{r,L}} AP_{j,L}^r,$$

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

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

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 5.9. *The restriction to Ω of the map*

$$(70) \quad (t, \mathbf{s}) \in \Omega \mapsto x \exp(tX_\alpha^{(L)}) \exp\left(\sum_{i=1}^a s_i Y_i^{(L)}\right)$$

is injective.

Proof. For every $t \in [0, T]$, we define a set $\Omega(t) \subset \{t\} \times \mathbb{R}^a$ as follows:

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

Then we set

$$(71) \quad \phi_{X_\alpha}^t(x) \exp\left(\sum_{i=1}^a s_i Y_i^{(L)}\right) = \phi_{X_\alpha}^{t'}(x) \exp\left(\sum_{i=1}^a s'_i Y_i^{(L)}\right).$$

Let us assume $t' \geq t$. By considering the projection on the base torus, we have the following identity:

$$(72) \quad (t, s_1, \dots, s_n) \bmod \mathbb{Z}^{n+1} = pr_1(\phi_{X_\alpha}^t(x)) \\ = pr_1(\phi_{X_\alpha}^{t'}(x)) = (t', s'_1, \dots, s'_n) \bmod \mathbb{Z}^{n+1},$$

which implies $t \equiv t'$ modulo \mathbb{Z} . As $\phi_{X_\alpha}^t = \phi_{X_\alpha}^{t'}$, 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 $s'_i = s_i$. Then injectivity is obtained by definition of I. Assume that $r_0 \neq 0$. Let $p, q \in M_\theta^a$ and then we have

$$p := \phi_{X_\alpha}^t(x), \quad q := \phi_{X_\alpha}^{t'}(x) \implies q = \Phi_{\alpha, \theta}^{r_0}(p).$$

From identity (71) we have

$$(73) \quad q = p \exp\left(\sum_{i=1}^a s_i Y_i^{(L)}\right) \exp\left(-\sum_{i=1}^a s'_i Y_i^{(L)}\right) \\ = p \exp\left(\sum_{i=1}^a (s'_i - s_i + P_i(s_i, s'_i)) L^{-\rho_i} Y_i\right)$$

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

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

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

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

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

and

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

For the same reason, from formula (73) 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 5.6, we have $p \in AP_{j_0, L}^{r_0}$ and $q \in AP_{j_0, L}^{-r_0}$.

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 on case 2-1. Hence,

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

which is a contradiction.

If the map in formula (70) fails 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

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

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

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

On case (2-2), we have

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

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

$$\frac{I}{2^{j_0+1}} \leq \delta_{r_0, L}(p) \leq \max_{i \geq n+1} |s_i| + |s'_i| + |\epsilon_i| < \frac{I}{2^{l_1+1}} + \frac{I}{2^{l_2+1}} \leq \frac{I}{2} \frac{1}{2^{j_0+1}}.$$

Hence, the injectivity is proved. \square

We simply reprove the bound of averaged width (see [FF14, Lemma 5.5]) in the general settings (under transversality conditions) by combining with Lemma 5.9.

Lemma 5.10. *For all $x \in M$ and for all $T, L \geq 1$ we have*

$$\frac{1}{w_{\mathcal{F}_\alpha^{(L)}}(x, T)} \leq \left(\frac{2}{I}\right)^a \frac{1}{T} \int_0^T H_L^T \circ \phi_{X_\alpha^{(L)}}^t(x) dt.$$

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

$$(74) \quad w_\Omega(t) = \begin{cases} \left(\frac{I}{2}\right)^a & \text{case 1} \\ \left(\frac{I}{2}\right)^a \left(\frac{\min\{\epsilon_{r, L}\}}{2}\right)^n & \text{case 2-1} \\ \left(\frac{I}{2}\right)^a 2^{-(a-n)(l+1)} & \text{case 2-2,} \end{cases}$$

and it implies that

$$(75) \quad \frac{1}{w_\Omega(t)} \leq \begin{cases} \left(\frac{2}{I}\right)^{a-n} & \text{case 1} \\ \left(\frac{2}{I}\right)^{a-n} \sum_{|r|=1}^{[TL]} \sum_{j>J_{r,L}} \frac{2^n \chi_{AP_{j,L}^r}(\phi_{X_\alpha^{(L)}}^t(x))}{(\epsilon_{r,L})^n} & \text{case 2-1} \\ \left(\frac{2}{I}\right)^a \sum_{|r|=1}^{[TL]} \sum_{j>J_{r,L}} 2^{(j+1)(a-n)} \chi_{AP_{j,L}^r}(\phi_{X_\alpha^{(L)}}^t(x)) & \text{case 2-2.} \end{cases}$$

By the definition of the function H_L^T in formula (69), we have

$$(76) \quad \frac{1}{w_\Omega(t)} \leq \left(\frac{2}{I}\right)^a H_L^T \circ \phi_{X_\alpha^{(L)}}^t(x), \text{ for all } t \in [0, T].$$

From the definition (39) of the average width of the orbit segment $\{x \exp(tX_\alpha^{(L)}) \mid 0 \leq t \leq T\}$, we have the estimate

$$\frac{1}{w_{\mathcal{F}_\alpha^{(L)}}(x, T)} \leq \frac{1}{T} \int_0^T \frac{dt}{w_\Omega(t)} \leq \left(\frac{2}{I}\right)^a \frac{1}{T} \int_0^T H_L^T \circ \phi_{X_\alpha^{(L)}}^t(x) dt.$$

□

Lemma 5.11. *For all $r \in \mathbb{Z} \setminus \{0\}$ and for all $L \geq 1$, the following estimate holds:*

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

Proof. It follows from the Lemma 5.7 that for $r \neq 0$ and for all $j \geq 0$, the Lebesgue measure of the set $AP_{j,L}^r$ satisfies the following bound:

$$(77) \quad \mathcal{L}^{a+1}(AP_{j,L}^r) \leq \frac{CI^{a-n}}{2^{j(a-n)}} L^{-\sum_{i=n+1}^a \rho_i}.$$

From the formula (67), it follows that

$$\begin{aligned} \int_M h_{r,L}(x) dx &\leq 1 + \sum_{j=1}^{J_{r,L}} 2^{j(a-n)} \mathcal{L}^{a+1}(AP_{j,L}^r) \\ &\quad + \sum_{j>J_{r,L}} \frac{2^n \mathcal{L}^{a+1}(AP_{j,L}^r)}{(\epsilon_{r,L})^n}. \end{aligned}$$

By estimate in the formula (77), we immediately have that

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

By the definition of the cut-off in formula (66) we have the bound

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

and by an estimate on a geometric sum

$$\begin{aligned} \sum_{j > J_{r,L}} \frac{2^n \mathcal{L}^{a+1}(AP_{j,L}^r)}{(\epsilon_{r,L})^n} &\leq \frac{2^{n-(J_{r,L}+1)(a-n)}}{(\epsilon_{r,L})^n} CI^{a-n} L^{-\sum_{i=n+1}^a \rho_i} \\ &\leq CI^{a-n} L^{-\sum_{i=n+1}^a \rho_i}. \end{aligned}$$

□

5.3. Diophantine estimates. In this section we review the concept of simultaneous Diophantine condition. The bounds on the expected average width is estimated under Diophantine conditions.

Definition 5.12. For any basis $\bar{Y} := \{\bar{Y}_1, \dots, \bar{Y}_n\} \subset \mathbb{R}^n$, let $\bar{I} := \bar{I}(\bar{Y})$ be the supremum of all constants $\bar{I}' > 0$ such that the map

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

is a local embedding on the domain

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

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

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

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

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

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

Here is Simultaneous Diophantine condition used in [FF14, Def. 5.8].

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

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

Definition 5.14. Let $\sigma = (\sigma_1, \dots, \sigma_n) \in (0, 1)^n$ be such that $\sigma_1 + \dots + \sigma_n = 1$. For any $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{R}^n$, for any $N \in \mathbb{N}$ and every $\delta > 0$, let

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

For every $\nu > 1$, let $D_n(\bar{Y}, \sigma, \nu) \subset (\mathbb{R} \setminus \mathbb{Q})^n$ 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 $\delta > 0$,

$$(78) \quad \#R_\alpha(N, \delta) \leq C(\bar{Y}, \sigma, \alpha) \max\{N^{1-\frac{1}{\nu}}, N\delta\}.$$

The Diophantine condition implies a standard simultaneous Diophantine condition. We quote following Lemmas proved in [FF14].

Lemma 5.15. [FF14, Lemma 5.9] Let $\alpha \in D_n$. For all $r \in \mathbb{Z} \setminus \{0\}$, we have

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

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

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

Lemma 5.16. [FF14, Lemma 5.12] *For all bases $\bar{Y} \subset \mathbb{R}^n$, for all $\sigma = (\sigma_1, \dots, \sigma_n) \in (0, 1)^n$ such that $\sigma_1 + \dots + \sigma_n = 1$ 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

$$(79) \quad \frac{1}{\nu} < \min\{[M(\sigma)n]^{-1}, 1 - \left(1 - \frac{2}{n}\right)\left(1 - \frac{m(\sigma)}{M(\sigma)}\right)\}.$$

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

Lemma 5.17. [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 $\bar{Y} = \{\bar{Y}_1, \dots, \bar{Y}_n\} \in \mathbb{R}$ denote the projection of the basis of codimension 1 ideal \mathfrak{J} onto the Abelianized Lie algebra $\bar{\mathfrak{n}} := \mathfrak{n}/[\mathfrak{n}, \mathfrak{n}] \approx \mathbb{R}^n$. For $\rho = (\rho_1, \dots, \rho_n) \in [0, 1]^n$, we write a vector of scaling exponents

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

Let $\alpha_1 = (\alpha_1^{(1)}, \dots, \alpha_n^{(1)}) \in D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$. For brevity, let $C(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \alpha_1)$ denote the constant appeared in (78) for $\alpha_1 \in D_n(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \nu)$. Let

$$(80) \quad C(\alpha_1) = 1 + C(\bar{Y}, \bar{\rho}/|\bar{\rho}|, \alpha_1).$$

We prove the upper bound on the cut-off function in the formula (66). Let $I = I(Y)$ and $\bar{I} = \bar{I}(\bar{Y})$ be the positive constant introduced in the Definition 5.6 and 5.12. 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 exponential map. Then the following logarithmic upper bound holds.

Lemma 5.18. [FF14, Lemma 5.14] *For every $\rho \in [0, 1]^n$, 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 $r \in \mathbb{Z} \setminus \{0\}$, the following bound holds:*

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

Proof. By Lemma 5.15 and by the definition of $\epsilon_{r,L}$ in formula (61), it follows that, for all $T > 0, L \geq 1$ and for all $r \in \mathbb{Z} \setminus \{0\}$, 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 (66),

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

□

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:

$$(81) \quad \mathcal{H}(Y, \rho, \alpha) = 1 + I(Y)^{a-n} C(\alpha_1) \{1 + \log^+ [I(Y)^{-1}] + \log C(\alpha_1)\}.$$

Theorem 5.19. [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 the formula (69), the statement follows from the Lemma 5.11 and Lemma 5.18. In fact, for all $r \in \mathbb{Z} \setminus \{0\}$ and $j \geq 0$, by definition (63) the set $AP_{j,L}^r$ is nonempty only if $\epsilon_{r,L} < \frac{1}{2}$. Since $\nu \leq 1/|\bar{\rho}|$, it follows from the definition of the Diophantine class D_n

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

Hence, the statement follows from the Lemma 5.11 and 5.18. \square

Remark. *Main idea of the proof above follows from the Lemma 5.11 which is based on the estimate to the upper bound of the measure of almost periodic set $AP_{j,L}^r$. In the Lemma 5.7, this bound is independent of choice of transverse section M_θ^a .*

5.4. Width estimates along orbit segments. We introduce a definition of good points, that is, points on the nilmanifold for which we can prove bounds on the width of sufficiently many orbit segments to derive by our method bounds on ergodic averages.

Definition 5.20. 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 $T_i \geq 1$. Let's say $N_i = [\log T_i]$ 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}_α if having set $y_i = \phi_{X_\alpha}^{T_i}(x)$, for all $i \in \mathbb{N}$ and for all $0 \leq j \leq N_i$, we have

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

Lemma 5.21. [FF14, Lemma 5.18] *Let $\zeta > 0$ be fixed and let (T_i) be an increasing sequence of positive real numbers satisfying the condition*

$$(82) \quad \Sigma((T_i), \zeta) := \sum_{i \in \mathbb{N}} (\log T_i)^2 (T_i)^{-\zeta} < \infty.$$

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, $\exists K > 0$ 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 $j = 0, \dots, N_i$, 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

$$(83) \quad \mathcal{G}(w, (T_i), \zeta)^c = \bigcup_{i \in \mathbb{N}} \bigcup_{j=0}^{N_i} (\mathfrak{S}_{j,i} \cup \phi_{X_\alpha}^{-T_i}(\mathfrak{S}_{j,i})).$$

By Lemma 5.10 for all $z \in \mathfrak{S}_{j,i}$ we have

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

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

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

Let $\mathcal{H} = \mathcal{H}(Y, \rho, \nu)$ denote the constant defined in the formula (81). By Theorem 5.19, since by hypothesis $\nu \leq 1/|\bar{\rho}|$ and $\alpha \in D_n(\bar{\rho}/|\bar{\rho}|, \nu)$, there exists a constant $K'(a, n, \nu) > 0$ 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

$$(84) \quad N_i \leq \log T_i \leq N_i + 1, \quad \log T_{j,i} \leq 2j.$$

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 (84), for some constant $K''' > 0$,

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

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

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

By formula (83), the above estimate concludes the proof. \square

6. BOUNDS ON ERGODIC AVERAGE

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 \widehat{M} the set of coadjoint orbits of \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 representation are called *maximal*. (See [FF07, Lemma 2.3])

Definition 6.2. Given a coadjoint orbit $\mathcal{O} \in \widehat{M}_0$ and a linear functional $\Lambda \in \mathcal{O}$, let us denote $\mathcal{F}_{\alpha, \Lambda}$ the completed basis $\mathcal{F}_{\alpha, \Lambda} = (X_\alpha, Y_\Lambda)$. For all $t \in \mathbb{R}$, we write scaled basis $\mathcal{F}_{\alpha, \Lambda}(t)$ by

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

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 $\Lambda(\eta_i^{(m)}) \neq 0, \forall (i, m) \in J$. 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 sub-representations equivalent to $\text{Ind}_{N'}^N(\Lambda)$. For adapted basis \mathcal{F} , set

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

6.1. Coboundary estimates for rescaled basis. Recall 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 scaling vector $\rho \in (\mathbb{R}^+)^J$ such that

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

for any $Y_i^{(j)}$ with $\deg(Y) = 0$.

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

$$(85) \quad S_{\mathfrak{n}}(k) := (n_1 - 1)(k - 1) + n_2(k - 2) + \dots + n_{k-1}$$

and

$$(86) \quad \delta(\rho) := \min_{\substack{1 \leq m \leq k-1 \\ 1 \leq i, j \leq n_m}} \{\rho_i^{(m)} - \rho_j^{(m+1)}, \rho_i^{(m)} - \rho_i^{(m+1)}\}.$$

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

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

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

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

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 commutation relation. Setting $C = \max_{(i,j) \in J^+} \{|c_i^{(j)}|\}$,

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

□

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

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

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

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

$$(87) \quad |R(t)(g)| \leq C_r^{(1)} e^{(1-\delta(\rho)-(1-\lambda))t} \max\{1, \delta_{\mathcal{O}}^{-1}\} \\ \times T^{-1} \left(w_{\mathcal{F}_{\alpha, \Lambda}(t)}(x, 1)^{-\frac{1}{2}} + w_{\mathcal{F}_{\alpha, \Lambda}(t)}(\phi_{X_{\alpha}}^T(x), 1)^{-\frac{1}{2}} \right) |g|_{r, \mathcal{F}_{\alpha, \Lambda}(t)}.$$

Proof. Fix $t \geq 0$ and set $D = D(t)$, $R = R(t)$ for convenience. Let $g \in W_{\alpha, \Lambda}^r(H_{\mathcal{O}}, \mathcal{F}(t))$. 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 . Then, g_R is a coboundary and $R(g_D) = 0$. Let $f = G_{X_{\alpha}, \Lambda}^{X_{\alpha}(t)}$. From $|D(g_R)| = 0$,

$$(88) \quad |R(g)| = |R(g_D + g_R)| = |R(g_R)| = |\gamma_x(g_R) - D(g_R)| = |\gamma_x(g_R)|.$$

By the Gottschalk-Hedlund argument,

$$(89) \quad |\gamma_x(g_R)| = \left| \frac{1}{T} \int_0^T g \circ \phi_{X_{\alpha}}^s(x) ds \right| \\ = \frac{1}{T} |f \circ \phi_{X_{\alpha}}^T(x) - f(x)| \\ \leq \frac{1}{T} (|f(x)| + |f \circ \phi_{X_{\alpha}}^T(x)|).$$

By Theorem 4.3 and Lemma 6.3, for any $\tau > a/2 + 1$, there exists a positive constant C_r such that for any $z \in M$

$$(90) \quad |f(z)| \leq \frac{C_r}{w_{\mathcal{F}_{\alpha, \Lambda}(t)}(z, 1)^{\frac{1}{2}}} \left(C e^{(1-\delta(\rho))t} |f|_{\tau, \mathcal{F}_{\alpha, \Lambda}(t)} + |g|_{\tau-1, \mathcal{F}_{\alpha, \Lambda}(t)} \right).$$

By Theorem 3.11, if $r > 2(k+1)\tau + 1/2$, then

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

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 $r > 2(k+1)(a/2+1) + 1/2$, there is a constant $C_r^{(2)}$ such that the following holds for every $\mathcal{O} \in \widehat{I}_0$ and every $x \in M$. Then,*

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

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 section, we derive the bounds on ergodic averages of nilflows for function in a single irreducible sub-representation.

For brevity, let us set

$$(91) \quad C_r(\mathcal{O}) = (1 + \delta_{\mathcal{O}}^{-1}).$$

Proposition 6.6. *Let $r > 2(k+1)(a/2+1) + 1/2$. Let (T_i) be an increasing sequence of positive real numbers ≥ 1 and let $0 < w < I(Y)^a$. Let $\zeta > 0$. There exists a constant $C_r(\rho)$ such that for every $\mathcal{G}(w, (T_i), \zeta)$ -good points $x \in M$ and all $f \in W^r(H_{\mathcal{O}}, \mathcal{F})$, we have*

$$(92) \quad \left| \frac{1}{T_i} \int_0^{T_i} f \circ \phi_{X_{\alpha}}^t(x) dt \right| \leq C_r(\rho) C_r(\mathcal{O}) w^{-1/2} T_i^{-\delta(\rho) + \zeta + \lambda/2} |f|_{r, \mathcal{F}_{\alpha}}.$$

Proof. By group action of scaling (18), a sequence of frame is chosen $\mathcal{F}(t_j) = A_{\rho}^{t_j} \mathcal{F}$ 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 to 1. Let $\phi_{X_j}^s(x)$ denote the flow of the scaled vector field $e^{t_j} X = X(t_j)$.

For each $j = 0, \dots, N$, 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 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 $N_i = \lfloor \log T_i \rfloor$ and $t_{j,i} := T_{j,i} = \log T_i^{j/N_i}$ for integer $j \in [0, N_i]$. We observe $N_i < \log T_i < N_i + 1$. For simplicity, we will omit 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 distribution of unscaled basis). By triangle inequality and Corollary 6.5,

$$(93) \quad \begin{aligned} |\gamma|_{-r, 0} &\leq |D_0|_{-r, 0} + |R_0|_{-r, 0} \\ &\leq |D_0|_{-r, 0} + C_r^{(2)} [1/I(Y)]^{a/2} C_r(\Lambda_{\mathcal{O}}) T^{-1}. \end{aligned}$$

We now estimate $|D_0|_{-r, 0}$. By definition of the Lyapunov norm and its bound (38), for $-s < -r < 0$

$$(94) \quad |D_0|_{-s, 0} \leq C_{r,s} \|D_0\|_{-r, 0}.$$

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

$$(95) \quad \|D_{j-1}\|_{-r, j-1} \leq \|D_j\|_{-r, j-1} + C |R_j|_{-r, j}.$$

By Lemma 3.15, 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 $t_j - t_{j-1} = \log T/N$ implies

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

From (95) we conclude by induction

$$(96) \quad \|D_0\|_{-r, 0} \leq T^{-\lambda(\rho)/2} \left(\|D_N\|_{-r, N} + C \sum_{l=0}^{N-1} T^{(l+1)\lambda(\rho)/2N} |R_{N-l}|_{-r, N-l} \right).$$

By Lemma 6.7 and 6.8,

$$(97) \quad \|D_0\|_{-r, 0} \leq C_r^1(\rho) C_r(\mathcal{O}) w^{-1/2} T^{1-\delta(\rho)+\zeta/2-(1-\lambda)-\lambda(\rho)/2}.$$

From (93) and the above, we conclude that there exists a constant $C_r(\rho)$ such that

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

□

Here we introduce the proof of 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. For all $i \in \mathbb{N}$, set $t_j = t_{j,i}$. By definition 5.20, for $x \in \mathcal{G}(w, (T_i), \zeta)$ and $y_i = \phi_{X_\alpha}^{T_i}(x)$

$$(98) \quad \frac{1}{w_{\mathcal{F}_\alpha^{(t_j)}}(x, 1)} \leq T_i^\zeta / w \quad \text{and} \quad \frac{1}{w_{\mathcal{F}_\alpha^{(t_j)}}(y_i, 1)} \leq T_i^\zeta / w.$$

By definition of norm, we obtain $\|D_N\|_{-r, N} \leq |D_N|_{-r, N} \leq |\gamma|_{-r, N}$. The orbit segment $(\phi_{X_\alpha}^t(x))_{0 \leq t \leq T}$ coincides with the orbit segment $(\phi_{X_\alpha}^\tau(x))_{0 \leq \tau \leq 1}$ of length 1 since $X_\alpha(t_N) = X_\alpha(\log T) = T X_\alpha$. By Theorem 4.3,

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

By the inequality (98),

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

□

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

$$(99) \quad \sum_{l=0}^{N-1} T^{(l+1)\rho_Y/2N} |R_{N-l}|_{-r, N-l} \leq C_r^{(1)}(\rho) C_r(\mathcal{O}) w^{-1/2} T^{1-\delta(\rho)-(1-\lambda)+\zeta/2}.$$

Proof. The orbit segment $(\phi_{X_\alpha}^t(x))_{0 \leq t \leq T}$ 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 $e^{(1-\delta(\rho))t_{N-l}} = T^{(1-l/N)(1-\delta(\rho))}$. Then,

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

$$\leq 2C_r^{(1)}C_r(\mathcal{O})w^{-1/2}T^{(1-l/N)(\lambda-\delta(\rho))-l/N+\zeta/2}.$$

Let $C = 2C_r^{(1)}C_r(\mathcal{O})w^{-1/2}$. Since $N_i = \lfloor \log T_i \rfloor$ and $N_i \leq \log T_i \leq N_i + 1$, we have $T_i^{1/(N_i+1)} \leq e \leq T_i^{1/N_i}$. By setting $T = T_i$,

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

By (86), we have $1 + \lambda - \delta(\rho) - \rho_Y/2 \geq 1 - \rho_Y/2 > 1/2$, thus geometric series converges. \square

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

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

Proof. From (84), $t_j - t_{j-1} \leq 2$ and observe $\mathcal{F}(t_j) = A^{t_j-t_{j-1}}\mathcal{F}(t_{j-1})$. Passing from the frame $\mathcal{F}(t_{j-1})$ to \mathcal{F}_t , it can be verified that distortion of the corresponding transversal Sobolev norm is uniformly bounded. \square

Let

$$\widetilde{M}_0 = \bigcup_{\mathcal{O} \in \widetilde{M}_0} \{\Lambda \in \mathcal{O} \mid \Lambda \text{ integral}\}$$

be collection of maximal integral coadjoint orbits.

Remark. Let $\sigma = (\sigma_1, \dots, \sigma_n) \in (0, 1)^n$ be such that $\sigma_1 + \dots + \sigma_n = 1$. For simplicity, we choose $\sigma_i = 1/n$ from now on. (See Definition 5.13 or Lemma 5.16)

Theorem 6.10. *For any $\Lambda \in \widetilde{M}_0$, let $\nu \in [1, 1 + (k/2 - 1)\frac{1}{n}]$. Then, for any $r > (k+1)(a/2 + 1) + 1/2$, there exists a constant $C(\sigma, \nu)$ satisfying the following. For every $\epsilon > 0$ there exists a constant $K_\epsilon(\sigma, \nu) > 0$ such that, for every $\alpha_1 = (\alpha_1^{(1)}, \dots, \alpha_n^{(1)}) \in D_n(\sigma, \nu)$ and for every $w \in (0, I(Y)^a]$ there exists a measurable set $\mathcal{G}_\Lambda(\sigma, \epsilon, w)$ satisfying the estimate*

$$(100) \quad \text{meas}(\mathcal{G}_\Lambda(\sigma, \epsilon, w)^c) \leq K_\epsilon(\sigma, \nu) \left(\frac{w}{I(Y_\Lambda)^a} \right) \mathcal{H}(Y_\Lambda, \rho, \alpha).$$

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

$$\left| \frac{1}{T} \int_0^T f \circ \phi_{X_\alpha}^t(x) dt \right| \leq \frac{C_r(\sigma, \nu)C_r(\Lambda)}{w^{\frac{1}{2}}} T^{-(1-\epsilon)\frac{1}{3S_n(k)}} |f|_{r, \mathcal{F}_{\alpha, \Lambda}}.$$

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

$$\rho_i^{(j)} = \frac{d_j}{S_n} \text{ for } i \leq k.$$

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

$$\Sigma(w, (T_i), \zeta) \leq \sum_i (\log T_i)^2 T_i^{-\zeta} \leq K_\epsilon(\rho).$$

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}_α . The estimate in the formula (100) follows from the Lemma 5.21 and definition of good points. By Proposition 6.6, for all $x \in \mathcal{G}$ and for every $f \in W^r(H_\mathcal{O}, \mathcal{F})$, the estimate (92) holds true. Given $T \in [T_i, T_{i+1}]$,

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

Let $C = C_r(\rho)C_r(\mathcal{O})/w^{1/2}$. The first term is estimated by the formula (92):

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

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. \square

Remark. If \mathcal{O} is integral but not maximal, then the restriction of Λ factors through an irreducible representation of the $k-1$ step nilpotent group $N/\exp \mathfrak{n}'_k$. Then, $\mathfrak{n}/\mathfrak{n}_k$ is polarizing subalgebra for subrepresentation and it reduces to the case of maximal integral. 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 integral maximal full rank case.

6.3. General bounds on ergodic averages. Finally, in order to solve cohomological equation on nilmanifold, we glue the solutions constructed in every irreducible sub-representation of N . The main idea is to increase extra regularity of the Sobolev norm to obtain the estimates that are uniformly bounded across all irreducible subrepresentation.

Definition 6.11. 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 $\eta_*^{(k)}$ whose degree k such that

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

Lemma 6.12. 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 one subgroup $N' \subset N$ is 1. Hence, by Definition 5.6, we have $I(Y_\Lambda) = 1/2$ for the basis. By (80), we have and $C(\alpha_1) \geq 1$. Then, from the definition of the constant $H(Y, \rho, \alpha)$, we obtain

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

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

□

Corollary 6.13. *For every $\mathcal{O} \in \widehat{M}_0$, $\Lambda \in \mathcal{O}$, $w > 0$ and $\epsilon > 0$, let*

$$(101) \quad w_\Lambda = w |\Lambda(\mathcal{F})|^{-2a-\epsilon}.$$

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

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

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 integral multiples of 2π . By Lemma 6.12, inequality (100) and definition of w_Λ , we have

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

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

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

The last statement on the monotonicity of the set follows from the analogous statement in Theorem 6.10. □

In every coadjoint orbit, we will make a particular choice of a linear form to accomplish the estimates of the bound for each irreducible sub-representation in terms of higher norms.

Definition 6.14. 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 $\Lambda_{\mathcal{O}}$ follows from

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

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

Lemma 6.15. *There exists a constant $C(\Gamma) > 0$ such that the following holds on the primary subspace $C^\infty(H_{\mathcal{O}})$ the following holds:*

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

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 $\Lambda' = \Lambda \circ \text{Ad}(e^{x_0 X_\alpha})$. The element $W \in \mathfrak{J}$ is represented in the representation as multiplication operators by the polynomials

$$(102) \quad P(\Lambda, W)(x) = \Lambda(\text{Ad}(e^{x X_\alpha})W).$$

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 (102), we have

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

Then we obtain

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

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,

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

By above identity in formula, the constant operators $\Lambda'(\eta_j^{(m)})$ are given by polynomial and derivative expressions of degree k in the operators $P(\Lambda', \eta_j^{(m)})$ we obtain the estimate

$$|\Lambda'(\mathcal{F})| \text{Id} \leq C_1(\Gamma)(1 + \Delta_{\Lambda', \mathcal{F}})^{k/2}.$$

Since the representation $\pi_{\Lambda'}$ and $\pi_{\Lambda_{\mathcal{O}}}$ are unitarily intertwined by the translation operator by x_0 , and since constant operators commute with translations, we also have

$$|\Lambda'(\mathcal{F})| \text{Id} \leq C_1(\Gamma)(1 + \Delta_{\Lambda_{\mathcal{O}}, \mathcal{F}})^{k/2}.$$

Since x_0 is bounded by a constant depending only step size k , the norms of the linear maps $\text{Ad}(\exp(\pm x_0 X_\alpha))$ are bounded by a constant depending only on k . Therefore, $|\Lambda_{\mathcal{O}}(\mathcal{F})| \leq C_2(k)|\Lambda'(\mathcal{F})|$ and the statements of the lemma follows. \square

Corollary 6.16. *There exists a constant $C'(\Gamma)$ such that for all $\mathcal{O} \in \widehat{M}_0$ and for any sufficiently smooth function $f \in H_{\mathcal{O}}$,*

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

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

Proof. From the definition (91) we have $C_r(\Lambda_{\mathcal{O}}) = (1 + |\Lambda_{\mathcal{O}}(\mathcal{F})|)^{l_1}$ with $l_1 = kr + 1$. By the formula (101),

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

with $l_2 = l_1 + 2a$. By Lemma 6.15 we have

$$(1 + |\Lambda_{\mathcal{O}}(\mathcal{F})|)^{l_2} \leq C'(\Gamma)(1 + \Delta_{\mathcal{F}})^{l_2 k/2}.$$

□

Proposition 6.17. *Let $r > (k+1)(3a/4+2)+1/2$. Let $\sigma = (1/n, \dots, 1/n) \in (0, 1)^n$ be a positive vector. Let us assume that $\nu \in [1, 1 + (k/2 - 1)\frac{1}{n}]$ 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*

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

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

$$(103) \quad \left| \frac{1}{T} \int_0^T f \circ \phi_{X_{\alpha}}^t(x) dt \right| \leq Cw^{-1/2}T^{-(1-\epsilon)\frac{1}{3S_n(k)}} |f|_{r, \mathcal{F}_{\alpha}}.$$

Proof. Let $\tau := r - ak/2 > (a+2)(k+1) + 1/2$. Let $f \in W^{\tau}(M, \mathcal{F})$ and let $f = \sum_{\mathcal{O} \in \widehat{M}_0} f_{\mathcal{O}}$ be its orthogonal decomposition onto the primary subspace $H_{\mathcal{O}}$. For each $\mathcal{O} \in \widehat{M}_0$, the constant $w_{\mathcal{O}}$ is given and the set

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

has measure greater than $1 - Cw\epsilon^{-1}$ as proved in Corollary 6.13.

If $x \in \mathcal{G}(\sigma, \epsilon, w)$, then by Theorem 6.10 and Corollary 6.16, the following estimate holds true for every $\mathcal{O} \in \widehat{M}_0$ and all $T \geq 1$:

$$\left| \frac{1}{T} \int_0^T f_{\mathcal{O}} \circ \phi_{X_{\alpha}}^t(x) dt \right| \leq C_r(\sigma, \nu)w^{-1/2}T^{-(1-\epsilon)\frac{1}{3S_n(k)}} |f_{\mathcal{O}}|_{r, \mathcal{F}_{\alpha}}.$$

For any $\tau > 0$ and any $\epsilon' > 0$, by Lemma 6.15 and orthogonal splitting of $H_{\mathcal{O}}$ we have

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

and the theorem follows after renaming the constant. □

Proof of Theorem 1.1. Under same hypothesis of Proposition 6.17, 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 $x \in \mathcal{G}_i \setminus \mathcal{G}_{i-1}$. By Proposition 6.17, the set \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. For step- k strictly triangular nilpotent Lie algebra \mathfrak{n} has dimension $\frac{1}{2}k(k+1)$ with 1 dimensional center. If the coadjoint orbit \mathcal{O} is integral

and maximal, then the optimal exponent will be attained by the formula (85). Let $\rho = (\cdots, \rho_i^{(m)}, \cdots)$ be the rescaling factor given :

$$S_n(k) = [(k-1)^2 + \sum_{n=1}^{k-2} n(n+1)] = (k-1)(k^2 + k - 3),$$

and we choose homogeneous scaling

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

Then, we can verify that

$$(104) \quad \lambda(\rho) = \delta(\rho) = \frac{1}{(k-1)(k^2 + k - 3)}.$$

By inductive argument with rescaling again, the exponent is obtained which proves Corollary 1.2. \square

7. UNIFORM BOUND OF THE AVERAGE WIDTH OF STEP 3 CASE.

In this section, we prove Theorem 1.3 on the effective equidistribution of nilflow on strictly triangular step 3 nilmanifold. On its structure, it is possible to derive uniform bound under Roth-type Diophantine condition due to linear divergence of orbit. This argument is based on counting principles of close return times which substitute the necessity of good point.

7.1. Average Width Function. Let N be a step 3 nilpotent Lie group on 3 generators 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 following commutation relations

$$(105) \quad [X_1, X_2] = Y_1, [X_2, X_3] = Y_2, [X_1, Y_2] = [Y_1, X_3] = Z.$$

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

By definition of the average width (74), for any $t \geq 0$ and for any $(x, T) \in M \times [1, +\infty)$ we construct an open set $\Omega_t(x, T) \subset \mathbb{R}^6$ which contains the segment $\{(s, 0, \cdots, 0) \mid 0 \leq s \leq T\}$ such that the map

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

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

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

such that

$$\begin{aligned} (106) \quad \Gamma x \exp(s'e^t V) \exp(e^{-\frac{1}{3}t} x'_2 X_2 + e^{-\frac{1}{3}t} x'_3 X_3 + e^{-\frac{1}{6}t} y'_1 Y_1 + e^{-\frac{1}{6}t} y'_2 Y_2 + z'Z) \\ = \Gamma x \exp(se^t V) \exp(e^{-\frac{1}{3}t} x_2 X_2 + e^{-\frac{1}{3}t} x_3 X_3 + e^{-\frac{1}{6}t} y_1 Y_1 + e^{-\frac{1}{6}t} y_2 Y_2 + zZ). \end{aligned}$$

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 .

Lemma 7.1. *Under equality (106), we obtain followings:*

$$\begin{aligned}\tilde{x}_2(t, s) &= \tilde{x}_2, & \tilde{x}_3(t, s) &= \tilde{x}_3, & \tilde{y}_1(t, s) &= \tilde{y}_1 + e^{\frac{5}{6}t} s \tilde{x}_2 \\ \tilde{y}_2(t, s) &= \tilde{y}_2 + \alpha e^{\frac{5}{6}t} s \tilde{x}_3 + 1/2 e^{-1/2t} (x_2 x'_3 - x'_2 x_3).\end{aligned}$$

By Definition 5.6, let $c_\Gamma (= I) > 0$ denote the supremum of all constants $c_\Gamma \in (0, 1/2)$ such that for all the map $\phi_x(0, \mathbf{s})$ is local embedding on the domain

$$\{\mathbf{s} \in \mathbb{R}^5 \mid |x_i|, |y_i| < c_\Gamma \text{ for all } i\}.$$

Let us assume that $|x_i|', |x_i|, |y_i|, |y_i'| \leq c_\Gamma/4$ and $\tilde{x}_i, \tilde{y}_i, \tilde{z} \in [-\frac{c_\Gamma}{2}, \frac{c_\Gamma}{2}]$. Projecting (106) on the base torus, we obtain

$$(107) \quad \exp(re^t \bar{V}) \exp(e^{-\frac{1}{3}t} \tilde{x}_2 \bar{X}_2 + e^{-\frac{1}{3}t} \tilde{x}_3 \bar{X}_3) \in \bar{\Gamma}$$

where 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 the equation (107) on projected torus has a solution $\tilde{x}_2, \tilde{x}_3 \in [-\frac{c_\Gamma}{2}, \frac{c_\Gamma}{2}]$. Then for every $r \in R_t(x, T)$, the solution $\tilde{x}_i := \tilde{x}_i(r)$ of the identity in formula (106) is unique.

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 7.2. *The following average-width estimate holds: for every $T > 1$,*

$$(108) \quad \frac{1}{T} \int_0^T \frac{ds}{w_{\Omega_t(r)}(s)} \leq \frac{1024}{c_\Gamma^2} \frac{C_\alpha}{e^{\frac{2}{3}t} \|(\tilde{x}_2(r), \tilde{x}_3(r))\|}.$$

Proof. 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 which fails injectivity. Here we approximate concrete width estimates with counting principles. By its 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 $|\tilde{x}_2(r)| \geq e^{-5t/6}$, there exists solution s^* of the equation $\tilde{y}_1(t, s) = \tilde{y}_1 + e^{\frac{5}{6}t} s \tilde{x}_2$. The same 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.

$$(109) \quad \#\mathcal{S}^*(r) \leq c_\Gamma^{-1} C_\alpha \|(\tilde{x}_2(r), \tilde{x}_3(r))\| e^{\frac{1}{6}t} T.$$

Proof. Let us say that s^* is a point on the manifold M , which minimizes the distance between orbit and its return. Denote

$$s^* = \min_s \max\{|\tilde{y}_1(t, s)|, |\tilde{y}_2(t, s)|\}$$

with

$$\tilde{y}_1(t, s) = \tilde{y}_1 + e^{\frac{5}{6}t} s \tilde{x}_2, \quad \tilde{y}_2(t, s) = \tilde{y}_2 + \alpha e^{\frac{5}{6}t} s \tilde{x}_3 + 1/2 e^{-\frac{1}{2}t} (x_2 x'_3 - x'_2 x_3).$$

If either distance $|\tilde{y}_1(t, s)|$ or $|\tilde{y}_2(t, s)|$ dominates another, then it reduces to simply finding a solution to single equation. For other case, we assume $|\tilde{y}_1(t, s)| = |\tilde{y}_2(t, s)|$. We distinguish following two cases but in either case, we can restrict either $\tilde{y}_1(r) = 0$ or $\tilde{y}_2(r) = 0$ for convenience.

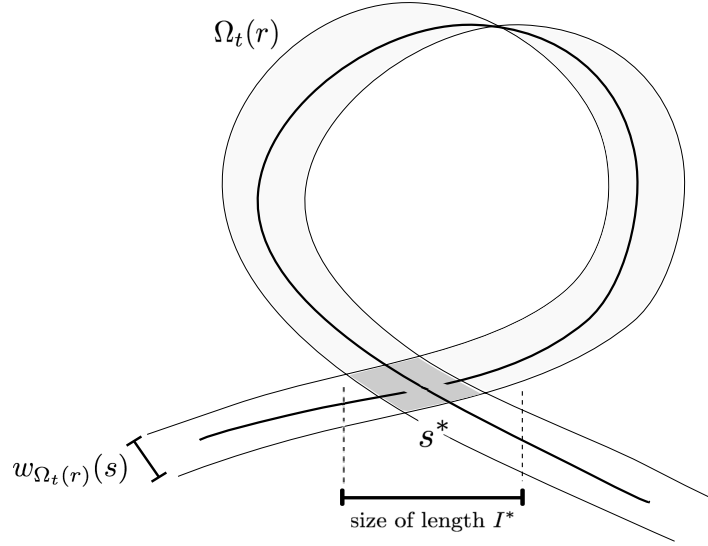


FIGURE 7.1. Illustration of width function and related quantities

If $\tilde{y}_1(t, s) = \tilde{y}_2(t, s)$, then assuming $\tilde{y}_1(r) = 0$, we obtain

$$s = \frac{e^{-\frac{5}{6}t}(\tilde{y}_2 + 1/2e^{-\frac{1}{2}t}(x_2x'_3 - x'_2x_3))}{\tilde{x}_2(r) - \alpha\tilde{x}_3(r)}.$$

If $\tilde{y}_1(t, s) = -\tilde{y}_2(t, s)$, then

$$s = \frac{e^{-\frac{5}{6}t}(\tilde{y}_2 + 1/2e^{-\frac{1}{2}t}(x_2x'_3 - x'_2x_3))}{\tilde{x}_2(r) + \alpha\tilde{x}_3(r)}.$$

From the bound $|\tilde{y}_2 + 1/2e^{-\frac{1}{2}t}(x_2x'_3 - x'_2x_3)| \geq |2/c_\gamma - 16/c_\gamma^2|$, we count

$$\#\mathcal{S}^*(r) \leq \begin{cases} (2c_\gamma - 16)/c_\gamma^{-2}|\tilde{x}_2(r) - \alpha\tilde{x}_3(r)|e^{\frac{1}{6}t}T & \text{if } \tilde{x}_2(r)\tilde{x}_3(r) < 0 \\ (2c_\gamma - 16)/c_\gamma^{-2}|\tilde{x}_2(r) + \alpha\tilde{x}_3(r)|e^{\frac{1}{6}t}T & \text{if } \tilde{x}_2(r)\tilde{x}_3(r) > 0 \end{cases}$$

thus, we prove the claim. \square

Since $\#\mathcal{S}^*(r)$ counts specific subspace on whole components, it suffices to conclude the number of interval has same bounds with $\#\mathcal{S}^*(r)$. Let

$$\Omega_t(r) := \{(s, x_2, x_3, y_1, y_2, z) \mid \max\{|x_2|, |x_3|, |y_1|, |y_2|\} < \delta_r(t, s), |z| < c_\Gamma/16\}$$

and define *the set of narrow width* by

$$\Omega_t(x, T) := \bigcap_{r \in R_t(x, T)} \Omega_t(r).$$

Under 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 given by return times r and $-r$.

For every $r \in R_t(x, T)$ and every $s \in [0, T]$, we define function

$$\delta_r(t, s) = \begin{cases} \frac{1}{16} \|(\tilde{x}_2(r), \tilde{x}_3(r))\| |(s - s^*)e^{\frac{5t}{6}}| & \text{for } s \in I^* \text{ with } |s - s^*| \geq e^{-\frac{5}{6}t} \\ \frac{1}{16} \|(\tilde{x}_2(r), \tilde{x}_3(r))\| & \text{for } s \in I^* \text{ with } |s - s^*| \leq e^{-\frac{5}{6}t} \\ \frac{c_\Gamma}{16} & \text{for all } s \in [0, T] \setminus \mathcal{S}(r) \end{cases}$$

associated with $\Omega_t(r)$.

By the definition of inner width and by construction of the set $\Omega_t(r)$ we have that

$$w_{\Omega_t(r)}(s) = \delta_r(t, s)^2, \quad \forall s \in [0, T].$$

It follows that for every subinterval $I^* \subset S(r)$ we have (using definition of δ_r)

$$\int_{I^*} \frac{ds}{w_{\Omega_t(r)}(s)} \leq \frac{512}{e^{\frac{5t}{6}} \|(\tilde{x}_2(r), \tilde{x}_3(r))\|^2}.$$

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. \square

Recall from Definition 5.13, we choose simultaneously Diophantine number $\alpha \in \mathbb{R}^2 \setminus \mathbb{Q}^2$ of exponent $\nu \geq 1$.

Lemma 7.3. *Given Diophantine condition of exponent $\nu \geq 1$, there exists a constant $C(\alpha) > 0$ such that all solutions of formula (107) satisfy the following lower bound*

$$\|(\tilde{x}_2, \tilde{x}_3)\|_{\mathbb{Z}^2} \geq C_\alpha e^{(\frac{1}{3} - \frac{\nu}{2})t} r^{-\frac{\nu}{2}}.$$

Proof. From projected identity (107) on base 3-torus,

$$(re^t, re^t \alpha + e^{-t/3} \tilde{x}_2, re^t \beta + e^{-t/3} \tilde{x}_3) \in \mathbb{Z}^3.$$

If it holds, 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

$$\begin{aligned} \|(\tilde{x}_2, \tilde{x}_3)\|_{\mathbb{Z}^2} &= e^{\frac{1}{3}t} \|(p_1 - q\alpha, p_2 - q\beta)\|_{\mathbb{Z}^2} \\ &= e^{\frac{1}{3}t} \|(q\alpha, q\beta)\|_{\mathbb{Z}^2} \\ &= e^{\frac{1}{3}t} \|q(\alpha)\|_{\mathbb{Z}^2} \\ &\geq C_\alpha e^{\frac{1}{3}t} q^{-\frac{\nu}{2}} \end{aligned}$$

which proves the statement. \square

For every $n \in \mathbb{N}$, let $R_t^{(n)}(x, T) \subset R_t(x, T)$ characterized by

$$\max(\tilde{x}_2(r), \tilde{x}_3(r)) \in (\frac{c_\Gamma}{2^{n+1}}, \frac{c_\Gamma}{2^n}].$$

Lemma 7.4. *For all $\epsilon > 0$, if the frequency of the projected linear flow satisfies Diophantine condition of exponent $\nu = \sqrt{2} + \epsilon$, then there exists $C_\epsilon > 0$ such that*

$$(110) \quad \#R_t^{(n)}(x, T) \leq C_\epsilon(\bar{V}) T \frac{c_\Gamma}{2^n} e^{\frac{2}{3}t + \frac{\epsilon t}{2}}.$$

Proof. Under a Diophantine condition of exponent $\nu \geq 1$, from inequality (78) and definition of $R_t^{(n)}(x, T)$, we have following:

$$(111) \quad \#R_t^{(n)}(x, T) \leq C_\nu(V) \max\{(Te^t)^{1-\frac{1}{\nu}}, Te^t \frac{c_\Gamma}{2^n} e^{-\frac{5t}{6}}\}.$$

It suffices to show $(Te^t)^{1-\frac{1}{\nu}}$ is less than or equal to the desired bound. From Lemma 7.3,

$$\begin{aligned} (Te^t)^{1-\frac{1}{\nu}} / \|(\tilde{x}_2, \tilde{x}_3)\| &\leq (Te^t)^{1-\frac{1}{\nu}} C_\alpha e^{(-\frac{1}{3}+\frac{\nu}{2})t} r^{\frac{\nu}{2}} \\ &\leq T^{1+\frac{\nu}{2}-\frac{1}{\nu}} C_\alpha e^{(\frac{2}{3}+\frac{\nu}{2}-\frac{1}{\nu})t}. \end{aligned}$$

Limiting $\nu \rightarrow \sqrt{2}$,

$$(Te^t)^{1-\frac{1}{\nu}} / \|(\tilde{x}_2, \tilde{x}_3)\| \leq C_\alpha T e^{(\frac{2}{3}+\frac{\epsilon}{2})t}.$$

Approximating $\|(x_2, x_3)\| \sim 1/2^n$,

$$(Te^t)^{1-\frac{1}{\nu}} \leq C_\epsilon(\bar{V}) T^{\frac{C_\Gamma}{2^n}} e^{(\frac{2}{3}+\frac{\epsilon}{2})t}.$$

□

By combining counting return time and width estimates, we obtain uniform bound.

Proposition 7.5. *There exists a constant*

$$\frac{1}{T} \int_0^T \frac{ds}{w_{\Omega_t(r)}(s)} \leq C_\epsilon(V) e^{\epsilon t}.$$

Proof. By inequality (108) and Lemma 7.4

$$\begin{aligned} \frac{1}{T} \int_0^T \frac{ds}{w_{\Omega_t(r)}(s)} &\leq \sum_{r \in R_t} \left(\frac{1024}{c_\Gamma^2} \frac{C_\alpha}{e^{\frac{2}{3}t} \|(\tilde{x}_2(r), \tilde{x}_3(r))\|} \right) \\ &\leq C_\epsilon(V) e^{\epsilon t}. \end{aligned}$$

□

Corollary 7.6. *For every nilflow generated with V such that projected flow on \mathbb{T}^3 is Diophantine condition of $\nu \in [1, \sqrt{2} + \epsilon]$. For every $\epsilon > 0$, there exists a constant $C_\epsilon(V) > 0$ such that, for all $t \geq 0$ and for all $(x, T) \in M \times \mathbb{R}$ we have*

$$w_{\mathcal{F}(t)}(x, T) \geq C_\epsilon(V)^{-1} e^{-\epsilon t}.$$

Proof of Theorem 1.3. By Corollary 7.6, it goes without quoting Good points technique and Lyapunov norm. Improved bound of remainder term R in Theorem 6.4 can be obtained.

$$(112) \quad |R(g)|_{-r} \leq C_r (1 + \delta_{\mathcal{O}}^{-1})^{r-2} T^{-1}.$$

We revisit backward iteration scheme introduced in proof of Theorem 6.6. We have

$$\begin{aligned} (113) \quad |\gamma|_{-r,0} &\leq |D_0|_{-r,0} + |R_0|_{-r,0} \\ &\leq |D_0|_{-r,0} + C_r (1 + \delta_{\mathcal{O}}^{-1})^{r-2} T^{-1} \end{aligned}$$

and

$$(114) \quad |D_0|_{-r,0} \leq |D_N|_{-r,0} + \sum_{j=1}^N |R'_{j-1}|_{-r,0}.$$

Changing the length to 1 and by uniform width bound from Corollary 7.6,

$$(115) \quad |D_N|_{-r,0} \leq C_r T^{-1/12} |D_N|_{-s, \mathcal{F}(t_N)} \leq C w_{\mathcal{F}(t_N)}(x, 1)^{-1/2} \leq C_\epsilon T^\epsilon$$

Then, by inductive argument resembling (96),

$$(116) \quad |D_0|_{-s,0} \leq C_{r,s} T^{-1/12} \left(|D_N|_{-r,\mathcal{F}(t_N)} + \sum_{j=1}^N C_j T_j^{1/12} |R_{j-1}|_{-r,\mathcal{F}(t_{j-1})} \right).$$

Therefore

$$|\gamma|_{-r,0} \leq C_r(\mathcal{O}) T^{-1/12+\epsilon}.$$

Finally, we glue all the function on irreducible representation $H_{\mathcal{O}}$, which only increases the regularity accordingly. \square

8. APPLICATION : MIXING OF NILAUTOMORPHISM.

In this section, as a further application of main equidistribution results, we can verify an explicit bound for the rate of exponential mixing of hyperbolic automorphism relying on renormalization argument.

Let $\mathfrak{F}_{2,3} = \{X_1, X_2, Y_1, Z_1, Z_2\}$ be 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 automorphism on Lie algebras induces automorphism on the nilmanifold

$$Aut(\mathfrak{n}) = \left\{ \begin{bmatrix} A & & \\ & 1 & \\ & & A \end{bmatrix}, A \in SL(2, \mathbb{Z}) \right\}$$

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 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 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*

$$(117) \quad \left| \frac{1}{T} \int_0^T f \circ \phi_V^t(x) dt \right| \leq C_s T^{-1/6+\epsilon} \|f\|_s.$$

The detailed computation follows from the section 7 and it is similar to the case of step 3 filiform [F16]. The only difference with filiform is that it has an extra element in center which is redundant in actual calculation on width, only raising required regularity of zero-average function.

The proposition below is firstly proved by A. Gorodnik and R. Spatzier in [GS14].

Proposition 8.2. *Hyperbolic nilautomorphism T is exponential mixing.*

Proof. Let $f, g \in C^1(M)$ be smooth. Define $\langle f, g \rangle = \int_M f g d\mu$. Since 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,

$$(118) \quad \langle f \circ T^n, g \rangle = \left\langle \int_0^1 f \circ T^n \circ \phi_V^t dt, g \circ \phi_V^t \right\rangle$$

$$(119) \quad - \int_0^1 \left\langle \int_0^t f \circ T^n \circ \phi_V^s ds, Vg \circ \phi_V^t \right\rangle dt.$$

Therefore,

$$(120) \quad \langle f \circ T^n, g \rangle = (\|g\|_\infty + \|Vg\|_\infty) \int_M \sup_{s \in [0,1]} \left| \int_0^s f \circ T^n \circ \phi_V^t dt \right| d\mu.$$

By renormalizing the flow,

$$T^n \circ \phi_V^t = \phi_V^{\lambda^n t} \circ T^n$$

and

$$\begin{aligned} \int_0^s f \circ T^n \circ \phi_V^t(x) dt &= \int_0^s f \circ \phi_V^{\lambda^n t} \circ T^n(x) dt \\ &= \frac{1}{\lambda^n} \int_0^{\lambda^n s} f \circ \phi_V^t \circ T^n(x) dt. \end{aligned}$$

Therefore, by the result of equidistribution (117),

$$(121) \quad \langle f \circ T^n, g \rangle \leq \lambda^{(-1/6+\epsilon)n} \|f\|_s (\|g\|_\infty + \|Vg\|_\infty) \rightarrow 0.$$

□

APPENDIX A.

In this appendix, we introduce specific example of nilpotent Lie algebra which goes beyond our approach introduced in the section 5.

A.1. Free group type of step 5 with 3 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 free nilpotent Lie algebra with n generators and $(\mathfrak{F}_n)_{k+1}$ be $k+1$ th subalgebra in central series, following notation in (4). Denote $\mathfrak{F}_{n,k} := \mathfrak{F}_n / (\mathfrak{F}_n)_{k+1}$ quotient of free algebra with n generators \mathfrak{F}_n and it is finite dimensional.

Definition A.1. Let \mathfrak{n} be nilpotent Lie algebra satisfying *generalized transversality condition* if there exists basis (X_α, Y_Λ) of \mathfrak{n} for each irreducible representation $\pi_\Lambda^{X_\alpha}$ such that

$$(122) \quad \langle \mathfrak{G}_\alpha \rangle \oplus \text{Ran}(\text{ad}_{X_\alpha}) + C_{\mathfrak{J}}(\pi_\Lambda^{X_\alpha}) = \mathfrak{n}$$

where $C_{\mathfrak{J}}(\pi_\Lambda^{X_\alpha}) = \{Y \in \mathfrak{J} \mid \Lambda([Y, X_\alpha]) = 0\}$.

Generalized transversality condition implies existence of completed basis for each irreducible representation $\pi_\Lambda^{X_\alpha}$ of non-zero degree. That is, given adapted basis $\mathcal{F} = (X, Y_1, \dots, Y_a)$, there exists reduced system $\bar{\mathcal{F}} = (X, Y'_1, \dots, Y'_{a'})$ satisfying transversality condition (41) and $\pi_\Lambda^X(Y'_m) = 0$ for all $a' \leq m \leq a$.

Now we will investigate an example that fails 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

$$\begin{aligned} [X_1, X_2] &= Y_1, & [X_2, X_3] &= Y_2, & [X_1, X_3] &= Y_3 \\ [X_1, Y_1] &= Z_1, & [X_1, Y_2] &= Z_2, & [X_1, Y_3] &= Z_3 \\ [X_2, Y_1] &= Z_4, & [X_2, Y_2] &= Z_5, & [X_2, Y_3] &= Z_6 \\ [X_3, Y_1] &= Z_7, & [X_3, Y_2] &= Z_8, & [X_3, Y_3] &= Z_9 \end{aligned}$$

and rest of elements are generated commutation relations with these. In general, we write elements $Y_j^{(i)} \in \mathfrak{n}_i \setminus \mathfrak{n}_{i+1}$ and $Y_i^{(5)} \in \mathfrak{Z}(\mathfrak{n})$ for all i . By Jacobi-identity

$$[X_1, [X_2, X_3]] + [X_2, [X_3, X_1]] + [X_3, [X_1, X_2]] = 0 \iff Z_2 - Z_6 + Z_7 = 0.$$

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 \mathfrak{I} ideal of $\mathfrak{F}_{5,3}$ codimension 1, not containing V .

Proposition A.2. *$\mathfrak{F}_{5,3}$ does not satisfy generalized transversality condition for some irreducible representation.*

Proof. To find centralizer in Lie algebra, for $a_i, b_i \in \mathbb{R}$, set

$$[V, X] = 0 \iff X = a_1 X_1 + a_2 X_2 + a_3 X_3 + b_1 Y_1 + b_2 Y_2 + b_3 Y_3 + c_1 Z_1 + \cdots + c_8 Z_8.$$

Then, it contains

$$\begin{aligned} (a_2 - \alpha_2 a_1) Y_1 + (\alpha_2 a_3 - \alpha_3 a_2) Y_2 + (a_3 - \alpha_3 a_1) Y_3 \\ + (b_1 - \beta_1 a_1) Z_1 + (b_2 - \beta_2 a_1) Z_2 + (b_3 - \beta_3 a_1) Z_3 + \cdots = 0. \end{aligned}$$

By linear independence, all the coefficients vanish and it remains

$$\begin{aligned} a_1 X_1 + a_2 X_2 + a_3 X_3 &= a_1 (X_1 + \alpha_2 X_2 + \alpha_3 X_3) \\ b_1 Y_1 + b_2 Y_2 + b_3 Y_3 &= a_1 (\beta_1 Y_1 + \beta_2 Y_2 + \beta_3 Y_3) \end{aligned}$$

Therefore, there is no non-trivial element in $C_{\mathfrak{I}}(V) \cap \mathfrak{n}_2 \setminus \mathfrak{n}_3$. Since range of ad_V has rank 2, this model does not satisfy transversality condition in the Lie algebra level.

Now, we verify *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 center.

Assume $\pi_*(W_i) = s_i I \neq 0$ for some $W_i \in \mathfrak{Z}(\mathfrak{n})$. Then, it is possible to choose element $L_i \in \mathfrak{n}_2 \setminus \mathfrak{n}_3$ such that

$$\begin{cases} \pi_*([V, L_1]) = (a_1 t^2 + a_2 t + a_3) \\ \pi_*([V, L_2]) = (b_1 t^2 + b_2 t + b_3) \\ \pi_*([V, L_3]) = (c_1 t^2 + c_2 t + c_3) \end{cases}$$

with (a_i, b_i, c_i) are non-proportional for each i , and

$$\pi_*(\text{ad}_V^3(L_i)) = \pi_*(W_i) \neq 0.$$

However, on given irreducible representation, any linear combination of L_1, L_2 and L_3 does not give any trivial relation. If $s_1 L_1 + s_2 L_2 + s_3 L_3 \in C_{\mathfrak{J}}(\pi_{\Lambda}^V)$, then

$$\begin{aligned} & \pi_*([V, s_1 L_1 + s_2 L_2 + s_3 L_3]) \\ &= s_1(a_1 t^2 + a_2 t + a_3) + s_2(b_1 t^2 + b_2 t + b_3) + s_3(c_1 t^2 + c_2 t + c_3) \\ &= (s_1 a_1 + s_2 b_1 + s_3 c_1)t^2 + (s_1 a_2 + s_2 b_2 + s_3 c_2)t + (s_1 a_3 + s_2 b_3 + s_3 c_3) = 0. \end{aligned}$$

The system of equations has trivial solution ($t = 0$) by linear independence of each coefficients. Then, there does not exist any element of $\mathfrak{n}_2 \setminus \mathfrak{n}_3$ that has degree 0. However, range of ad_V has rank 2 and generalized transversality condition cannot be satisfied in this example. \square

REFERENCES

- [AGH63] L. Auslander, L. Green, and F. Hahn, Flows on homogeneous spaces, with the assistance of L. Markus and W. Massey, and an appendix by L. Greenberg. *Annals of Mathematics Studies*, No. 53, Princeton University Press, Princeton, N.J., (1963).
- [BDG15] J. Bourgain, C. Demeter and L. Guth, Proof of the main conjecture in Vinogradov's mean value theorem for degrees higher than three, *Annals of Mathematics*, (2016) 633–682
- [CF15] S. Cosentino and L. Flaminio, Equidistribution for higher-rank Abelian actions on Heisenberg nilmanifolds. *J. Mod. Dyn.* 9 (2015), 305 - 353.
- [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, vol. 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. Vol. 104. *Springer Science & Business Media*, (2012).
- [FF03] L. Flaminio and G. Forni, Invariant distributions and time averages for horocycle flows, *Duke Math. J.* 119 (2003), no. 3, 465–526.
- [FF06] ———, Equidistribution of nilflows and applications to theta sums. *Ergodic Theory Dynam. Systems* 26 (2006), no. 2, 409–433.
- [FF07] ———, On the cohomological equation for nilflows *J. Mod. Dyn.* 1 (2007), no. 1, 37–60.
- [FF14] ———, On effective equidistribution for higher step nilflows. Preprint: arXiv:1407.3640.
- [FFT16] L. Flaminio, G. Forni, and J. Tanis. Effective equidistribution of twisted horocycle flows and horocycle maps. *Geom. Funct. Anal.* 26 (2016), no. 5, 1359–1448
- [F16] G. Forni. Effective Equidistribution of Nilflows and Bounds on Weyl Sums. *Dynamics and Analytic Number Theory* 437 (2016): 136.
- [GT12] B. Green and T. Tao, The quantitative behaviour of polynomial orbits on nilmanifolds, *Annals of Math.* 175 (2012), 465 - 540.
- [GS14] A. Gorodnik and R. Spatzier, Exponential mixing of nilmanifold automorphisms. *J. Anal. Math.* 123 (2014), 355–396.
- [H73] J. Humphreys. Introduction to Lie algebras and representation theory. Vol. 9. *Springer Science and Business Media*, 1973.
- [R72] M. S. Raghunathan, Discrete subgroups of Lie groups, *Springer-Verlag*, New York, Heidelberg, 1972, *Ergebnisse der Mathematik und ihrer Grenzgebiete*, Band 68.
- [T15] T.D Wooley. Perturbations of Weyl sums. *International Mathematics Research Notices* 2016.9 (2015): 2632–2646.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF MARYLAND, COLLEGE PARK, MD 20742, USA

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE, NICOLAUS COPERNICUS UNIVERSITY,
CHOPINA 12/18, 87-100 TORUŃ, POLAND
Email address: mkim16@mat.umk.pl