VERTICAL VERSUS HORIZONTAL INEQUALITIES ON SIMPLY CONNECTED NILPOTENT LIE GROUPS AND GROUPS OF POLYNOMIAL GROWTH

SEUNG-YEON RYOO

ABSTRACT. We establish "vertical versus horizontal inequalities" for functions from nonabelian simply connected nilpotent Lie groups and not virtually abelian finitely generated groups of polynomial growth into uniformly convex Banach spaces using the vector-valued Littlewood–Paley–Stein theory approach of Lafforgue and Naor (2012). This is a quantitative nonembeddability statement that shows that any Lipschitz mapping from the aforementioned groups into a uniformly convex space must quantitatively collapse along certain subgroups. As a consequence, a ball of radius $r \geq 2$ in the aforementioned groups must incur bilipschitz distortion at least a constant multiple of $(\log r)^{1/q}$ into a $q(\geq 2)$-uniformly convex Banach space. This bound is sharp for the L^p $(1 < p < \infty)$ spaces.

In the special case of mappings of Carnot groups into the L^p $(1 < p < \infty)$ spaces, we prove that the quantitative collapse occurs on a larger subgroup that is the commutator subgroup; this is in line with the qualitative Pansu–Semmes nonembeddability argument given by Cheeger and Kleiner (2006) and Lee and Naor (2006). We prove this by establishing a version of the classical Dorronsoro theorem on Carnot groups. Previously, in the setting of Heisenberg groups, Fässler and Orponen (2019) established a one-sided Dorronsoro theorem with a restriction $0 < \alpha < 2$ on the range of exponents α of the Laplacian; this restriction does not appear in the commutative setting and is caused by their use of horizontal polynomials as approximants. We identify the correct class of approximant polynomials and prove the two-sided Dorronsoro theorem with the full range $0 < \alpha < \infty$ of exponents in the general setting of Carnot groups, thus strengthening and extending the work of Fässler and Orponen.

CONTENTS

1. INTRODUCTION

It is well known that nonabelian simply connected nilpotent Lie groups and not virtually abelian finitely generated groups of polynomial growth fail to embed bilipschitzly (or quasi-isometrically) into uniformly convex Banach spaces (or, more generally, Banach spaces with the Radon–Nikodým property). This is because nonabelian Carnot groups do not bilipschitzly embed into Banach spaces with the Radon–Nikodým property [CK06, LN06], and nonabelian simply connected nilpotent Lie groups have nonabelian Carnot groups as asymptotic cone and hence do not quasi-isometrically embed into Banach spaces with the Radon-Nikodým property (see [HS20] for example), while not virtually abelian finitely generated groups of polynomial growth are quasi-isometric to nonabelian simply connected nilpotent Lie groups [Gro81, Mal49] and hence fail to embed quasi-isometrically into Banach spaces with the Radon–Nikodým property. The aim of this paper is to provide quantitative counterparts to these qualitative nonembeddability statements.

Let G be a nonabelian simply connected nilpotent Lie group, endowed with a left-invariant Riemannian metric. We write

$$B_r(x) := \{y \in G : d_G(x, y) < r\}, \ B_r := B_r(e_G), \quad x \in G, \ r > 0,$$

where e_G is the identity element of G. Note that $B_r(x) = xB_r$, by left-invariance.

Let Γ be a finitely generated group of polynomial growth. This means that Γ has a finite generating set S, and that if we denote by $d_W(\cdot, \cdot)$ the left-invariant word metric on Γ induced by S and $B_n^\Gamma = \{x \in \Gamma : d_W(x, e_\Gamma) \le n\}$ the corresponding closed ball of radius $n \in \mathbb{N}$, where e_Γ is the identity element of Γ, then the cardinality $|B_n^\Gamma|$ grows at most polynomially in n. We assume in addition that Γ is not virtually abelian, i.e., that it has no finite index subgroup isomorphic to $\mathbb{Z}^n$; by [DCTV07, Corollary 1.5] Γ has no finite index subgroups that have abelian groups as quotients by finite normal subgroups.

Let the Banach space $(X, \|\cdot\|_X)$ be a uniformly convex Banach space, that is, for every $\varepsilon \in (0, 1)$ there exists $\delta \in (0, 1)$ such that every $x, y \in X$ with $\|x\|_X = \|y\|_X = 1$ and $\|x - y\|_X \ge \varepsilon$ satisfy $\|x + y\|_X \le 2(1 - \delta)$. By [BCL94, Fig76, Pis75], there is an equivalent norm on X such that if we renorm X using this norm, there is some exponent $q \in [2, \infty)$ for which the Banach space $(X, \|\cdot\|_X)$ is q-uniformly convex, which means that the q-uniform convexity constant of X defined by

$$K_q(X) := \inf\left\{ K > 0 : \forall x, y \in X \ \left(\|x\|_X^q + \frac{1}{K^q}\|y\|_X^q\right)^{1/q} \le \left(\frac{\|x + y\|_X^q + \|x - y\|_X^q}{2}\right)^{1/q}\right\}$$

is finite. We will assume this renorming has happened, because it does not affect our theorems, so that X is q-uniformly convex for some $q \ge 2$. By [Han56] and [BCL94], the L_p spaces for $1 < p < \infty$ are $\max\{p, 2\}$-uniformly convex with

$$K_2(L_p) \le \frac{1}{\sqrt{p-1}}, \ 1 < p \le 2, \quad K_p(L_p) \le 1, \ p \ge 2.$$

A separable metric space (M, d_M) is said to embed *(D-)bilipschitzly* into a normed space $(X, \|\cdot\|_X)$ if there is a finite $D \in [1, \infty)$ and a mapping $f : M \to X$ such that $d_M(x, y) \le \|f(x) - f(y)\| \le Dd_M(x, y)$ for all $x, y \in M$. The bilipschitz distortion $c_X(M, d_M)$ is defined to be the infimum over those D for which such a mapping exists. We write[1] $c_{L_p(\sigma)}(M, d_M) = c_{L_p([0,1])}(M, d_M) =: c_p(M, d_M)$ and call $c_2(M, d_M)$ the Euclidean distortion of (M, d_M). The space (M, d_M) is said to *quasi-isometrically embed* into $(X, \|\cdot\|_X)$ if

---

[1]This equality is due to the fact that any separable subspace of an $L^p(\sigma)$ space is isometric to a subspace of $L^p([0,1])$, see [Ost13, Fact 1.20].

there is a mapping $f : M \to X$, $D \in [1, \infty)$ and $C > 0$ such that $d_M(x, y) - C \le \| f(x) - f(y) \| \le D d_M(x, y) + C$ for all $x, y \in M$.

Now we have defined all the concepts used in the first sentence of the introduction (except for the Radon–Nikodým property, which will not appear in other parts of this paper), which, in other words, can be written as

$$c_X(G) = c_X(\Gamma) = \infty,$$

for G, Γ, and X given as above. In this paper, we quantify this fact by providing the following growth rates of the bilipschitz constant of balls in such groups.

Theorem 1. *Let G be a nonabelian simply connected nilpotent Lie group, and let X be a q-uniformly convex Banach space, $q \ge 2$. Then*[2]

$$c_X(B_r) \gtrsim_G \frac{(\log r)^{1/q}}{K_q(X)}, \quad r \ge 2.$$

Theorem 2. *Let Γ be a not virtually abelian finitely generated group of polynomial growth, and let X be a q-uniformly convex Banach space, $q \ge 2$. Then*

$$c_X(B_n^\Gamma) \gtrsim_\Gamma \frac{(\log n)^{1/q}}{K_q(X)}, \quad n \ge 2.$$

As a special case, we obtain sharp bounds on the L^p-distortions, the upper bounds following from the Assouad embedding theorem.

Corollary 3. *Let G be a nonabelian simply connected nilpotent Lie group. Then,*

$$c_p(B_r) \asymp_{G,p} (\log r)^{1/\max\{p,2\}}, \quad 1 < p < \infty, \; r \ge 2.$$

Corollary 4. *Let Γ be a not virtually abelian finitely generated group of polynomial growth. Then*

$$c_p(B_n^\Gamma) \asymp_{\Gamma,p} (\log n)^{1/\max\{p,2\}}, \quad 1 < p < \infty, \; n \ge 2.$$

We remark that Theorem 2 and Corollary 4 were proven for the discrete Heisenberg groups by Lafforgue and Naor [LN14], where the discrete Heisenberg groups $\mathbb{H}_{\mathbb{Z}}^{2k+1}$, $k \in \mathbb{Z}_{>0}$ are defined as the groups with word relations as follows:

$$\mathbb{H}_{\mathbb{Z}}^{2k+1} = \Big\langle a_1, \cdots, a_k, b_1, \cdots, b_k, c \mid \forall i \, (c = [a_i, b_i])$$
$$\wedge \, \forall i, j \Big(i \ne j \to [a_i, a_j] = [b_i, b_j] = [a_i, b_j] = [a_i, c] = [b_i, c] = e_{\mathbb{H}_{\mathbb{Z}}^{2k+1}} \Big) \Big\rangle.$$

A previously weaker result was given by Li [Li14, Theorem 1.4], who proved that if Γ is a finitely generated nonabelian torsion-free nilpotent group, and X is a uniformly convex Banach space, then there exists $c > 0$ depending on Γ and X such that

$$c_X(B_n^\Gamma) \gtrsim_{G,X} (\log n)^c, \quad n \ge 2.$$

A stronger bound was given for cocompact lattices Γ of Carnot groups by Gartland [Gar20, Corollary 1.6]: if G is a Carnot group of step s (to be defined below) and Γ is a cocompact lattice of G, then

$$c_{\mathbb{R}^d}(B_n^\Gamma) \gtrsim_{\Gamma,d} \frac{(\log n)^{\frac{1}{2} - \frac{1}{2s}}}{(\log\log n)^{\frac{1}{2} + \frac{1}{2s}}}, \quad n \ge 3,$$

---

[2]We will use the following (standard) asymptotic notation. For $P, Q > 0$, the notations $P \lesssim Q$, $Q \gtrsim P$, $P = O(Q)$, and $Q = \Omega(P)$ mean that $P \le KQ$ for a universal constant $K \in (0, \infty)$, and the notation $P \asymp Q$ means $P \lesssim Q$ and $Q \lesssim P$. If we need to allow for dependence on parameters, we indicate this by subscripts. For example, in the presence of auxiliary parameters ψ, ξ, the notations $P \lesssim_{\psi,\xi} Q$, $Q \gtrsim_{\psi,\xi} P$, $P = O_{\psi,\xi}(Q)$, $Q = \Omega_{\psi,\xi}(P)$ mean that $P \le K(\psi, \xi)Q$ where $K(\psi, \xi) \in (0, \infty)$ may depend only on ψ and ξ, and $P \asymp_{\psi,\xi} Q$ means that $P \lesssim_{\psi,\xi} Q$ and $Q \lesssim_{\psi,\xi} P$.

under the additional restriction that G contains a copy of the model filiform group $J^{s-1}(\mathbb{R})$ of step s (this requirement is automatically fulfilled if $s \le 3$), which is the Carnot group whose Lie algebra $\mathfrak{j}^{s-1}$ is spanned by elements $x, y_0, \cdots, y_{s-1}$ with the only nontrivial bracket relations being

$$[x, y_i] = y_{i+1}, \quad i = 0, \cdots, s-2.$$

For example, $J^0(\mathbb{R}) = \mathbb{R}^2$, $J^1(\mathbb{R}) =: \mathbb{H}^3$ is called the real Heisenberg group of dimension 3, and $J^2(\mathbb{R})$ is called the Engel group. Compared to this result, Corollary 4 removes the exponent gap of $\frac{1}{2s}$ and gets rid of the lower order factor.

We obtain Theorems 1 and 2 by proving the following "vertical-versus-horizontal inequalities" on the groups G and Γ.

We first state the inequality on nilpotent groups. In the following theorem, given left-invariant vector fields $X_1, \cdots, X_k$ that form a basis of the Lie algebra $\mathfrak{g}$ of G at e_G, we write the horizontal gradient of a function $f : G \to X$ as

$$\nabla f := (X_1 f, \cdots, X_k f),$$

if each $X_i f$ exists. Also, given $v \in G$ and $t \in \mathbb{R}$, we write $v^t := \exp(tw)$, where $w \in \mathfrak{g}$ is such that $v = \exp(w)$, and $\exp : \mathfrak{g} \to G$ is the Lie group exponential map. Let μ denote the bi-invariant Haar measure of G, which is the push-forward of the Lebesgue measure on $\mathfrak{g}$ by the exponential map.

Theorem 5. *Let G be a nonabelian simply connected nilpotent Lie group. Let $v \in Z(G)$ and $\rho \in \mathbb{N}$ with $\rho \ge 2$ and $d_G(v^t, e_G) \le t^{1/\rho}$, $t > 0$. Suppose that $p \in (1, \infty)$ and $q \in [2, \infty)$. Let $(X, \|\cdot\|_X)$ be a Banach space with $K_q(X) < \infty$, and let $f : G \to X$ be smooth and compactly supported. Then*

$$\left(\int_0^\infty \left(\int_G \left(\frac{\|f(hv^t) - f(h)\|_X}{t^{1/\rho}} \right)^p d\mu(h) \right)^{\max\{p,q\}/p} \frac{dt}{t} \right)^{1/\max\{p,q\}}$$
$$\lesssim \max\left\{ (p-1)^{1/q-1}, K_q(X) \right\} \left(\int_G \|\nabla f(h)\|_{\ell_2^k(X)}^p d\mu(h) \right)^{1/p}. \tag{1}$$

In particular, when $p = q$,

$$\left(\int_1^\infty \int_G \left(\frac{\|f(hv^t) - f(h)\|_X}{t^{1/\rho}} \right)^q d\mu(h) \frac{dt}{t} \right)^{1/q} \lesssim K_q(X) \left(\int_G \|\nabla f(h)\|_{\ell_2^k(X)}^q d\mu(h) \right)^{1/q}. \tag{2}$$

Note that the v described in this theorem exists, for example, if we choose $v \in \underbrace{[[G,G], \cdots, G]}_{s \text{ times}} \setminus \{e_G\}$ where s is the nilpotency step of G, i.e., it is the largest integer s such that $\underbrace{[[\mathfrak{g}, \mathfrak{g}], \cdots, \mathfrak{g}]}_{s \text{ times}} \ne 0$, then we may normalize v so that $d_G(v^t, e_G) \le t^{1/s}$ for $t > 0$. We remark that $\exp\left(\underbrace{[[\mathfrak{g}, \mathfrak{g}], \cdots, \mathfrak{g}]}_{j \text{ times}} \right) = \underbrace{[[G,G], \cdots, G]}_{j \text{ times}}$ for $j = 1, \cdots, s$, where $[\mathfrak{g}, \mathfrak{g}]$ denotes the Lie algebra bracket and $[G, G]$ denotes the commutator subgroup.

We next state the inequality on groups of polynomial growth.

Theorem 6. *Let Γ be a not virtually abelian finitely generated group of polynomial growth. There exist $v_\Gamma \in \Gamma$, $s \in \mathbb{N}$ with $s \ge 2$, and $c = c(G) \in \mathbb{N}$ such that the following is true. First, $d_W(v_\Gamma^n, e_\Gamma) \asymp_\Gamma n^{1/s}$ for $n \in \mathbb{N}$. Second, let $p \in (1, \infty)$ and $q \in [2, \infty)$. Suppose that $(X, \|\cdot\|_X)$ is a Banach space satisfying $K_q(X) < \infty$. Then*

4

for every $n \in \mathbb{N}$ and every $f : \Gamma \to X$ we have

$$\left(\sum_{k=1}^{n^s} \frac{1}{k^{1+\max\{p,q\}/s}} \left(\sum_{x \in B_n^\Gamma} \|f(xv_\Gamma^k) - f(x)\|_X^p \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}}$$

$$\lesssim_{\Gamma,v_\Gamma} \max\left\{(p-1)^{1/q-1}, K_q(X)\right\} \left(\sum_{x \in B_{cn}^\Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^p \right)^{1/p}. \tag{3}$$

In particular, when $p = q$,

$$\left(\sum_{k=1}^{n^s} \sum_{x \in B_n^\Gamma} \frac{\|f(xv_\Gamma^k) - f(x)\|_X^q}{k^{1+q/s}} \right)^{1/q} \lesssim_{\Gamma,v_\Gamma} K_q(X) \left(\sum_{x \in B_{cn}^\Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^q \right)^{1/q}. \tag{4}$$

Our choice of s and v_Γ is as follows. By [Gro81], Γ admits a subgroup Γ' of finite index that is nilpotent. Let T be the torsion subgroup of Γ' and consider the quotient subgroup $\Gamma'' = \Gamma'/T$. We may take any $v'' \in Z(\Gamma'') \setminus \{e_{\Gamma''}\}$ with the property that $d_W((v'')^n, e_{\Gamma''}) \asymp_{\Gamma''} n^{1/s}$, $n \geq 2$, for some integer $s \geq 2$, and take v_Γ to be any representative of v''. For example, we may take s to be the nilpotency step of Γ'' and $v'' \in \underbrace{[[\Gamma'',\Gamma''],\cdots,\Gamma'']}_{s \text{ times}} \setminus \{e_{\Gamma''}\}$. (See Section 4 for details.)

One can see that changing to another finite symmetric generating set S of Γ or changing the vertical element v_Γ affects only $c \in \mathbb{N}$ and the constant in the inequalities (3) and (4) up to constant factors depending on S and v_Γ.

Theorems 5 and 6 are extensions from the real Heisenberg groups $\mathbb{H}^{2k+1}$ to general simply connected nilpotent Lie groups and from the discrete Heisenberg groups $\mathbb{H}_\mathbb{Z}^{2k+1}$ to general finitely generated groups of polynomial growth, respectively, of the "vertical versus horizontal inequalities" established by Austin, Naor, and Tessera [ANT13] and Lafforgue and Naor [LN14]. See also Naor and Young [NY18, NY20] for the endpoint case $p = 1$ for the Heisenberg groups. Here, the real Heisenberg groups $\mathbb{H}^{2k+1}$, $k \in \mathbb{Z}_{>0}$, are defined to be the simply connected nilpotent Lie groups whose Lie algebras $\mathfrak{h}^{2k+1}$ are spanned by the $2k+1$ elements $x_1, \cdots, x_k, y_1, \cdots, y_k$, and z, the only nontrivial bracket relations among which are

$$[x_i, y_i] = z, \quad i = 1, \cdots, k.$$

The discrete Heisenberg groups $\mathbb{H}_\mathbb{Z}^{2k+1}$, $k \geq 1$, embed naturally in the real Heisenberg groups $\mathbb{H}^{2k+1}$.

The inequality (1) is proven following the argument of Lafforgue and Naor [LN14] by comparing the left-hand side of (1) against convolutions of ∇f with derivatives of the heat kernel on the real line $\mathbb{R}$, and then upper bounding this quantity using the vector-valued Littlewood–Paley–Stein theory of [HN19, MTX06, Xu20]. The inequality (3) follows from inequality (1) by a discretization argument.

The utility of Theorems 5 and 6 is that they prove quantitatively that nonabelian simply connected nilpotent Lie groups and not virtually abelian finitely generated groups of polynomial growth, respectively, do not bilipschitzly embed into uniformly convex spaces. More precisely, they give Theorems 1 and 2 as Corollaries; see Section 3 for proofs.

We will prove Theorem 5 in Section 2. Actually, Theorem 5 will be a special case of the more general Theorem 32 and Corollary 33, and the resulting Theorem 1 will be a special case of the more general Theorem 34 and Corollary 35.

It would appear that we may also derive similar constraints for other coarse embeddings (see [Ost13, Definition 1.36] for the definition). For example, we have the following theorem, where for a metric space (M, d_M) and $0 < \alpha < 1$, (M, d_M^α) is a metric space called the α-*snowflake* of (M, d_M).

Theorem 7. *Let Γ be a not virtually abelian finitely generated group of polynomial growth, and let X be a q-uniformly convex Banach space. Then*

$$c_X(\Gamma, d_W^{1-\varepsilon}) \gtrsim_\Gamma \frac{1}{K_q(X)\varepsilon^{1/q}}.$$

In particular,

$$c_p(\Gamma, d_W^{1-\varepsilon}) \asymp_{p,\Gamma} \varepsilon^{-1/\max\{p,2\}}.$$

Indeed, by (4), any D-bilipschitz embedding $(\Gamma, d_W^{1-\varepsilon}) \to X$ must satisfy for all n

$$|B_n^\Gamma|^{1/q} \left(\sum_{k=1}^{n^s} \frac{1}{k^{1+\varepsilon q/s}} \right)^{1/q} \lesssim_\Gamma K_q(X) D |S|^{1/q} |B_{cn}^\Gamma|^{1/q}$$

which gives the stated lower bound. The second assertion of Theorem 10, namely the estimate for $c_p(N_{r_1,r_2})$, follows from an optimized version [LMN05, Theorem 5.1] of the Assouad embedding theorem [Ass83]:

$$c_p(\Gamma, d_W^{1-\varepsilon}) \lesssim_{G,p} 1/\varepsilon^{1/\max\{p,2\}}, \quad 0 < \varepsilon < 1.$$

Using Theorem 6, we may also characterize the compression rate of not virtually abelian finitely generated groups of polynomial growth into L^p spaces. Here, the *compression rate* of a Lipschitz function $f : (M, d_M) \to (X, \|\cdot\|_X)$ from a metric space (M, d_M) into a Banach space $(X, \|\cdot\|_X)$ is the largest nondecreasing function $\omega_f : (0, \infty) \to [0, \infty)$ such that for all $x, y \in M$ we have $\|f(x) - f(y)\|_X \geq \omega_f(d_M(x, y))$.

Corollary 8. *Let Γ be a not virtually abelian finitely generated group of polynomial growth.*

(1) For $n \in \mathbb{N}$, $n \geq 2$, a nondecreasing function $\theta : (0, \infty) \to [0, \infty)$ satisfies $\theta(t) \lesssim \omega_f(t)$ for some 1-Lipschitz function $f : B_n^\Gamma \to L^p$, $p > 1$, if and only if

$$\int_1^{2n} \left(\frac{\theta(t)}{t} \right)^{\max\{p,2\}} \frac{dt}{t} \lesssim_\Gamma 1.$$

(2) A nondecreasing function $\theta : (0, \infty) \to [0, \infty)$ satisfies $\theta(t) \lesssim \omega_f(t)$ for some 1-Lipschitz function $f : \Gamma \to L^p$, $p > 1$, if and only if

$$\int_1^\infty \left(\frac{\theta(t)}{t} \right)^{\max\{p,2\}} \frac{dt}{t} \lesssim_\Gamma 1.$$

The if direction is due to [Tes11, Theorem 1]. The only if direction is given by Theorem 6, from which it follows that

$$s \int_1^{2n} \left(\frac{\theta(t)}{t} \right)^{\max\{p,2\}} \frac{dt}{t} = \int_1^{(2n)^s} \frac{\theta(\tau^{1/s})^{\max\{p\}}}{\tau^{1+\max\{p\}/s}} d\tau \lesssim \sum_{k=1}^{(2n)^s} \frac{\theta(k^{1/s})^{\max\{p\}}}{k^{1+\max\{p,2\}/s}} \lesssim_\Gamma 1.$$

Under a slightly more abstract setting, one can obtain the following statement about the compression function.

Theorem 9. *Let Γ be an amenable group with finite generating set S, with $v \in Z(\Gamma)$ and $\rho > 1$ such that $d_W(v^k, e_\Gamma) \asymp_\Gamma k^{1/\rho}$, $k \in \mathbb{N}$. Let $(X, \|\cdot\|_X)$ be a q-uniformly convex space, and let $f : \Gamma \to X$ be a 1-Lipschitz function. Then for every $t \geq 3$ there exists an integer $t \leq n \leq t^2$ such that*

$$\frac{\omega_f(n)}{n} \lesssim_\Gamma K_q(X) \left(\frac{\log\log n}{\log n} \right)^{1/q}.$$

(Here we consider ρ to be dependent on Γ, so that $\lesssim_\Gamma$ includes dependence on ρ.)

We prove Theorem 9 in Section 5 following the argument of [ANT13]. Theorem 9 is somewhat weaker than our distortion result of Theorem 2 and the vertical versus horizontal inequality of Theorem 6 under the setting of groups of polynomial growth, because the distortion bound one can obtain from Theorem 9 is of the weaker form $c_X(B_n) \gtrsim_\Gamma \frac{1}{K_q(X)} \left(\frac{\log n}{\log\log n} \right)^{1/q}$ (this deduction follows the argument of [ANT13, Section 6], and requires the assumption that Γ be of polynomial growth).

On a different note, we can investigate the distortion of nets in nonabelian simply connected nilpotent Lie groups. In the following, with $r_1, r_2 > 0$, an r_1-*covering* of a metric space (M, d_M) is a subset $N \subset M$ such that for any $x \in M$ there exists $n_x \in N$ such that $d_M(x, n_x) < r_1$, an r_2-*packing* of (M, d_M) is a subset

$N \subset M$ such that for all distinct $n_1, n_2 \in N$ we have $d_M(n_1, n_2) \geq r_2$, and an (r_1, r_2)-*net* of (M, d_M) is a subset which is both an r_1-covering and an r_2-packing.

Theorem 10. *Let $r_1, r_2 > 1$ with $r_1 \geq 2r_2$. Let G be a nonabelian simply connected nilpotent Lie group, and let N_{r_1, r_2} be an r_2-covering of B_{r_1}. Let X be a q-uniformly convex space, $q \geq 2$. Then*

$$c_X(N_{r_1, r_2}) \gtrsim \left(\frac{C_1 s^{1/q}}{\sqrt{k} K^{2/q} \log K} \right) \frac{(\log(r_1/r_2))^{1/q}}{K_q(X)},$$

where K is the largest of the doubling constant of G and the doubling constant of the measure μ, i.e., it is the smallest constant such that $\mu(B_{2r}) \leq K\mu(B_r)$ for all $r \geq 0$, and that for all $r > 0$ there exist $y_1, \cdots, y_K \in G$ such that $B_{2r} \subset \cup_{i=1}^K B_r(y_i)$, and $C_1 > 0$ denotes a constant such that $C_1 t^{1/s} \leq d(v^t, e_G) \leq t^{1/s}$ for $t \geq 1$.

In particular, if N_{r_1, r_2} is an $(r_2, \Omega(r_2))$-net of B_{r_1}, then

$$c_p(N_{r_1, r_2}) \asymp_{G,p} (\log(r_1/r_2))^{1/\max\{p,2\}}, \quad 1 < p < \infty.$$

We remark that the second assertion of Theorem 10, namely the estimate for $c_p(N_{r_1, r_2})$, follows from an optimized version [LMN05, Theorem 5.1] of the Assouad embedding theorem [Ass83]:

$$c_p(G, d_G^{1-\varepsilon}) \lesssim_{G,p} 1/\varepsilon^{1/\max\{p,2\}}, \quad 0 < \varepsilon < 1.$$

When N_{r_1, r_2} is an $\Omega(r_2)$-packing of B_{r_1}, $(N_{r_1, r_2}, d_G^{1-1/2\log(r_1/r_2)})$ is $O(1)$-bilipschitz equivalent to (N_{r_1, r_2}, d_G), so we have the second assertion of Theorem 10.

Thus, the bound of Theorem 10 is sharp for $X = L^p$, and therefore is the best result we can attain in terms of q-uniformly convex spaces. These distortion bounds generalizes those established for the Heisenberg group in [LN14]. The same conclusion holds with L^p replaced by the Schatten class S_p, whose modulus of uniform convexity was computed in [TJ74].

We may also derive the following variant of Theorem 10 for snowflakes.

Theorem 11. *Let N'_{c_1, c_2} be a (c_1, c_2)-net of a nonabelian simply connected nilpotent Lie group G, where $c_1, c_2 \geq 1$ with $c_1 \geq c_2/2$. Let X be a q-uniformly convex space. Then*

$$c_X(N'_{c_1, c_2}, d_G^{1-\varepsilon}) \gtrsim \left(\frac{C_1^{1-\varepsilon}(c_2/c_1)^\varepsilon s^{(1-\varepsilon)/q}}{k^{(1-\varepsilon)/2} K^{2(1-\varepsilon)/q} \log K} \right) \frac{1}{K_q(X)^{1-\varepsilon} \varepsilon^{1/q}}, \quad 0 < \varepsilon < 1,$$

where K and C_1 are as in Theorem 10. In particular,

$$c_p(N'_{c_1, c_2}, d_G^{1-\varepsilon}) \asymp_{G,p,c_1,c_2} \varepsilon^{-1/\max\{p,2\}}, \quad 1 < p < \infty.$$

The nonembeddability of nonabelian simply connected nilpotent Lie groups and not virtually abelian finitely generated groups of polynomial growth into uniformly convex Banach spaces is known from the fact that the asymptotic cone of these groups are nonabelian Carnot groups, and that nonabelian Carnot groups do not admit bilipschitz embeddings into uniformly convex Banach spaces. The definition of Carnot groups is as follows.

Definition 12 ([LD17]). *A (sub-Riemannian) Carnot group is a 5-tuple $(G, \delta_\lambda, B, |\cdot|, d_G)$, where:*

- *The simply connected nilpotent Lie group G is such that its Lie algebra $\mathfrak{g}$ admits an s-step stratification, $s \in \mathbb{Z}_{>0}$, i.e., a direct sum decomposition*

$$\mathfrak{g} = V_1 \oplus V_2 \oplus \cdots \oplus V_s,$$

  where $V_s \neq 0$, $V_{s+1} = 0$, and $V_{r+1} = [V_1, V_r]$ for $r = 1, \cdots, s$.
- *For each $\lambda \in \mathbb{R}^+$, the linear map $\delta_\lambda : \mathfrak{g} \to \mathfrak{g}$ is defined by*

$$\delta_\lambda|_{V_r} = \lambda^i \mathrm{id}_{V_r}, \quad r = 1, \cdots, s.$$

- *The bundle B over G is the extension of V_1 to a left-invariant subbundle:*

$$B_p := (dL_p)_e V_1, \quad p \in G.$$

- The inner product norm[3] $|\cdot|$ is initially defined on V_1, and is then extended to B as a left-invariant norm:
$$|(dL_p)_e(v)| := |v|, \quad p \in G, \ v \in V_1.$$

- The metric d_G on G is the Carnot–Carathéodory distance associated to B and $|\cdot|$, i.e.,
$$d_G(p,q) := \inf\left\{ \int_0^1 |\dot{\gamma}(t)| dt : \gamma \in C^\infty_{\mathrm{pw}}([0,1];G), \gamma(0) = p, \gamma(1) = q, \dot{\gamma} \in B \right\}, \quad p,q \in G,$$

where $C^\infty_{\mathrm{pw}}([0,1];G)$ consists of the piecewise smooth functions from $[0,1]$ to G.

For simplicity, we will call G the Carnot group. When we wish to emphasize the step size, we will call G an s-step Carnot group or a Carnot group of step s.

It is easy to see that commutative Carnot groups are precisely those of step 1 and are the Euclidean spaces $\mathbb{R}^d$. The nonabelian ones are those of step $s \geq 2$. An example of this are the Heisenberg groups $\mathbb{H}^{2k+1}$ defined above, which are Carnot groups of step 2 whose Lie algebras $\mathfrak{h}^{2k+1}$ admit the 2-step stratification
$$V_1(\mathfrak{h}) = \mathrm{span}\{x_1, \cdots, x_k, y_1, \cdots, y_k\}, \quad V_2(\mathfrak{h}) = \mathrm{span}\{z\}.$$
Another example of Carnot groups are the aforementioned model filiform groups $J^{s-1}(\mathbb{R})$, $s \geq 1$, which are Carnot groups of step s whose Lie algebras $\mathfrak{j}^{s-1}$ admit the s-step grading
$$V_1(\mathfrak{j}^{s-1}) = \mathrm{span}\{x, y_0\}, \ V_2(\mathfrak{j}^{s-1}) = \mathrm{span}\{y_1\}, \ \cdots, \ V_s(\mathfrak{j}^{s-1}) = \mathrm{span}\{y_{s-1}\}.$$

Let our choice of left-invariant vector fields $X_1, \cdots, X_k$, $k := \dim V_1$, be such that $(X_1)_{e_G}, \cdots, (X_k)_{e_G}$ form an orthonormal basis of $(V_1, |\cdot|)$. We have the center $Z(G) = \exp(V_s)$ and commutator subgroup $[G,G] = \exp(V_2 \oplus \cdots \oplus V_s)$. Also, the maps δ_λ act as scalings in the Carnot–Carathéodory distance, in the sense that $d_G(\delta_\lambda(p), \delta_\lambda(q)) = \lambda d_G(p,q)$, for $\lambda > 0$ and $p,q \in G$.

In an algebraic sense, nonabelian Carnot groups are nonabelian simply connected nilpotent Lie groups, but their metrics differ, since we endowed nilpotent Lie groups with *Riemannian* distances while we endowed Carnot groups with *sub-Riemannian* distances. Nevertheless, Theorem 5 holds for nonabelian Carnot groups G, stated as follows.

Theorem 13. *Let G be a nonabelian Carnot group of step $s \geq 2$. Let $v \in Z(G) \setminus \{e_G\}$ be normalized so that $d_G(v, e_G) = 1$. Suppose that $p \in (1, \infty)$ and $q \in [2, \infty)$. Let $(X, \|\cdot\|_X)$ be a Banach space with $K_q(X) < \infty$, and let $f : G \to X$ be smooth and compactly supported. Then*

$$\left(\int_0^\infty \left(\int_G \left(\frac{\|f(hv^t) - f(h)\|_X}{t^{1/s}} \right)^p d\mu(h) \right)^{\max\{p,q\}/p} \frac{dt}{t} \right)^{1/\max\{p,q\}}$$
$$\lesssim \max\left\{ (p-1)^{1/q-1}, K_q(X) \right\} \left(\int_G \|\nabla f(h)\|^p_{\ell_2^k(X)} d\mu(h) \right)^{1/p}. \tag{5}$$

In particular, when $p = q$,

$$\left(\int_0^\infty \int_G \left(\frac{\|f(hv^t) - f(h)\|_X}{t^{1/s}} \right)^q d\mu(h) \frac{dt}{t} \right)^{1/q} \lesssim K_q(X) \left(\int_G \|\nabla f(h)\|^q_{\ell_2^k(X)} d\mu(h) \right)^{1/q}. \tag{6}$$

For nilpotent Lie groups, we have $d_G(v^t, e_G) \asymp_G t$ for $0 < t < 1$ and $d_G(v^t, e_G) \asymp_G t^{1/s}$ for $t \geq 1$ for our choice of v, while for Carnot groups we have $d_G(v^t, e_G) = t^{1/s}$ for all $t > 0$. Thus, Theorem 5 only gives global nonembeddability of nonabelian simply connected nilpotent Lie groups into uniformly convex spaces, i.e., we only know that the entire group fails to embed into X, while Theorem 13 gives local

---

[3]For sub-Finsler Carnot groups we allow $|\cdot|$ to be more generally a norm. Then, compared to the sub-Riemannian case, the resulting distance d_G is then distorted by a factor of at most $\sqrt{\dim V_1}$ by the John ellipsoid theorem [Joh48], and the results of this paper follow up to multiplicative factors of $\sqrt{\dim V_1}$.

nonembeddability of Carnot groups into uniformly convex spaces, i.e., we know also that any nonempty open subset of the group fails to embed into X.

Although Theorems 5 and 13 have a formal difference in the sense that Theorem 5 considers Riemannian distances on nilpotent Lie groups while Theorem 13 considers sub-Riemannian distances on Carnot groups, it will turn out that they are instances of a more general theorem regarding sub-Riemannian distances on nilpotent Lie groups. This is stated as Corollary 33 in Section 2. See Corollary 35 for the resulting distortion results.

Theorem 6 applies to cocompact lattices Γ of a nonabelian Carnot group G, because they are not virtually abelian finitely generated groups of polynomial growth. We can take s to be the step of G and $v_\Gamma \in Z(G) \cap \Gamma \setminus \{e_G\}$; we will not repeat the statement of Theorem 6 here.

We may state simplified versions of Theorems 10 and 11 in the setting of Carnot groups, thanks to the scale-invariance. In the following, $n_h := \sum_{r=1}^{s} r k_r$ is the Hausdorff dimension of G

Theorem 14. *Let $r_1, r_2 > 0$ with $r_1 \geq 2r_2$. Let G be a nonabelian Carnot group, and let N_{r_1,r_2} be an r_2-covering of B_{r_1}. Let X be a q-uniformly convex space, $q \geq 2$. Then*

$$c_X(N_{r_1,r_2}) \gtrsim \left(\frac{C_1 s^{1/q}}{\sqrt{k} 4^{n_h/q} n_h} \right) \frac{(\log(r_1/r_2))^{1/q}}{K_q(X)}.$$

In particular, if N_{r_1,r_2} is an $(r_2, \Omega(r_2))$-net of B_{r_1}, then

$$c_p(N_{r_1,r_2}) \asymp_{G,p} (\log(r_1/r_2))^{1/\max\{p,2\}}, \quad 1 < p < \infty.$$

Theorem 15. *Let N'_{c_1,c_2} be a (c_1, c_2)-net of a nonabelian Carnot group G, where $c_1, c_2 > 0$ with $c_1 \geq c_2/2$. Let X be a q-uniformly convex space. Then*

$$c_X(N'_{c_1,c_2}, d_G^{1-\varepsilon}) \gtrsim \left(\frac{C_1^{1-\varepsilon} (c_2/c_1)^\varepsilon s^{(1-\varepsilon)/q}}{k^{(1-\varepsilon)/2} 4^{n_h(1-\varepsilon)/q} n_h} \right) \frac{1}{K_q(X)^{1-\varepsilon} \varepsilon^{1/q}}, \quad 0 < \varepsilon < 1.$$

In particular,

$$c_p(N'_{c_1,c_2}, d_G^{1-\varepsilon}) \asymp_{G,p,c_1,c_2} \varepsilon^{-1/\max\{p,2\}}, \quad 1 < p < \infty.$$

That a nonabelian Carnot group G fails to embed bilipschitzly into $\mathbb{R}^n$ was first proven by Semmes [Sem96] using the differentiation theorem of Pansu [Pan89]. Pansu's theorem states that any Lipschitz mapping $f : G \to \mathbb{R}^n$ is differentiable a.e. with the derivative being a Lie group homomorphism that commutes with the dilations δ_λ of G and $\mathbb{R}^n$. In other words, one can 'blow-up' f almost everywhere and the resulting 'blow-up' function is a Carnot group homomorphism. Semmes' observation was that if f were a *bilipschitz* mapping, then the blow-up of f would also be bilipschitz, resulting in a Carnot group *monomorphism*. This is impossible because a group homomorphism $G \to \mathbb{R}^n$ must send the nontrivial commutator subgroup $[G,G]$ to 0. Later this argument was generalized, independently by Cheeger and Kleiner [CK06] and Lee and Naor [LN06], to Banach space targets with the Radon–Nikodým property, which includes uniformly convex spaces.

From the standpoint of the Pansu–Semmes argument which simply outputs the qualitative nonexistence of a bilipschitz mapping $G \to \mathbb{R}^n$, the significance and advantage of Theorems 5, 6, and 13 is that they give quantitative nonembeddability statements, such as Theorems 1, 2, 10, 11, 14, and 15. However, the Pansu–Semmes argument tells us that any Lipschitz function $f : G \to X$ must 'collapse' along directions of the commutator group $[G,G]$, and from this point of view, Theorem 13 possesses a curious limitation, namely that it requires us to measure the collapse along *central* directions.

This limitation is caused by a technical requirement in the proof of Theorems 5 and 13, which is for the horizontal derivative and the convolution along the direction of v to commute (more precisely, equation (28) below). Since there is no such limitation in the original Pansu–Semmes proof, we pose the following question.

Question 16. *Let G be a nonabelian Carnot group. Is Theorem 13 true for all $v \in [G,G]$ with $d_G(v, e_G) = 1$?*

We answer this question in the affirmative for L^p targets, albeit with some loss of control on the constants.

Theorem 17. *Let G be a nonabelian Carnot group. For $1 < p \le 2$, $f \in L^p(G)$ with $\nabla f \in L^p(G; \ell_2^k)$, and $v \in [G, G]$ with $d_G(v, e) = 1$, we have*

$$\left(\int_0^\infty \left[\int_G \left(\frac{|f(h) - f(h\delta_r(v))|}{r} \right)^p d\mu(h) \right]^{2/p} \frac{dr}{r} \right)^{1/2} \lesssim_{G,p} \|\nabla f\|_{L^p(G;\ell_2^k)}. \tag{7}$$

Although Theorem 17 is stated for $\mathbb{R}$-targets, it is easy to tensorize inequality (7) using the triangle inequality when $f : G \to L^p(\sigma)$ is smooth and compactly supported.

Remark 18. *With q defined as in Theorem 13, Theorem 17 only considers the exponent $q = 2$. This is because the case $q = \infty$ formally holds, so that the inequality for $q = 2$ implies the inequalities for $q \ge 2$. See Section 2, page 19 for a proof of this simple fact. We cannot make this simplification in Theorem 13 or in Theorem 5 due to the stated dependence of the constants on q.*

However, we have not managed to prove a strengthening of Theorem 6 to directions in the commutator group for Carnot groups, formulated as follows.

Question 19. *Let Γ be a cocompact lattice of a nonabelian Carnot group G. For any $v_\Gamma \in [\Gamma, \Gamma] \setminus \{e_G\}$, does there exist $c \in \mathbb{N}$ such that the following is true? If we let $s' \ge 2$ be the largest integer such that $v_\Gamma \in \underbrace{[[G, G], \cdots, G]}_{s' \text{ times}}$, let $p \in (1, \infty)$ and $q \in [2, \infty)$, and suppose $(X, \|\cdot\|_X)$ is a Banach space satisfying $K_q(X) < \infty$, then for every finitely supported $f : \Gamma \to X$ we have*

$$\left(\sum_{k=1}^{n^{s'}} \frac{1}{k^{1 + \max\{p,q\}/s'}} \left(\sum_{x \in B_n^\Gamma} \|f(x v_\Gamma^k) - f(x)\|_X^p \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}}$$

$$\lesssim_{\Gamma, v_\Gamma} \max\left\{(p-1)^{1/q-1}, K_q(X)\right\} \left(\sum_{x \in B_{cn}^\Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^p \right)^{1/p}.$$

We will deduce Theorem 17 from a Dorronsoro theorem for Carnot groups following the argument of Fässler and Orponen [FO20a]. We will also prove the Carnot group Dorronsoro theorem itself, which states that the $L^p(G)$ norm of the L^p fractional Laplacian $\|(-\Delta_p)^\alpha f\|_{L^p(G)}$ ($1 < p < \infty$, $\alpha > 0$) of a function $f \in L^p(G)$, is equivalent up to constant factors to a singular integral that, roughly speaking, measures, at all points of G and at all scales, the deviation of f from being a polynomial of weighted degree $\le \lfloor \alpha \rfloor$. This was first proven for $G = \mathbb{R}^n$ for all $\alpha > 0$ by Dorronsoro [Dor85], and for the Heisenberg groups $\mathbb{H}^{2k+1}$ one side of the Dorronsoro statement (namely that the aforementioned singular integral is bounded above by $\|(-\Delta_p)^\alpha f\|_{L^p}$) was proven for the restricted exponent range $0 < \alpha < 2$ in [FO20a]. The reason for this restriction $0 < \alpha < 2$ in [FO20a] is that they considered deviations from the smaller class of "horizontal polynomials", which is inadequate for $\alpha \ge 2$ (see Remark 23 for a proof).

We will discuss the Dorronsoro theorem for Carnot groups in detail later in subsection 1.1. To summarize, the novelty of this paper in this direction is threefold. First, we recognize the correct class of polynomials to approximate by, namely polynomials of weighted degree that depend on the "full set of coordinates" as opposed to just the "horizontal coordinates". Second, we recover the full Dorronsoro theorem, i.e., we prove both directions of the equivalence. Third, we verify that generalizing to higher step Carnot groups introduces no serious problems. The proof method of Theorem 22 is based on Dorronsoro's original proof in [Dor85] and adds various ingredients used in Fässler and Orponen's proof in [FO20a].

Remark 20. *So far, we have mentioned two proof methods of the vertical versus horizontal inequality, namely the Littlewood–Paley–Stein theory for Theorems 5 and 13 and the Dorronsoro theorem for the variant Theorem 17. There is a third possible proof method for L^2 targets which follows the representation*

theoretic proof of [ANT13] *for the case* $G = \mathbb{H}^3$, $X = L^2$. *Specifically, one can reduce proving the vertical versus horizontal inequality to proving a certain inequality regarding* 1-*cocycles* $G \to L^2$ *of irreducible unitary representations of* G. *The irreducible unitary representations of* $\mathbb{H}^3$ *are given by the Stone–von Neumann theorem, while the irreducible unitary representations have been described by Dixmier* [Dix59, Dix57] *and Kirillov* [Kir62]. *We will not pursue this proof method since the anticipated results do not seem to provide an improvement or advantage over neither Theorem 5 nor Theorem 17.*

Focusing on Euclidean distortion for concreteness, Theorem 14 tells us that

$$c_2(N_R) \asymp_{n_h} \sqrt{\log R},$$

for any Carnot group G and $(1, \Omega(1))$-net N_R of B_R. One may have expected that Carnot groups with higher steps may have a 'hierarchy' of collapsing happening, with distortions being propagated and amplified as one passes from $[G, G]$, $[G, [G, G]]$, all the way up to $\underbrace{[G, [G, \cdots, G]]}_{s \text{ times}}$, with the center ending up "super-collapsed"; certainly if G is high-dimensional and V_1 is low dimensional, the collapsing happens along the subgroup $[G, G]$ of small codimension, and one may have enough room to have many more interesting phenomena such as collapsing happen. In this case, the dependency on s of the asymptotics of $c_2(N_R)$ would appear in the exponent of $\log R$, but actually the asymptotics is exactly that of $\sqrt{\log R}$. Thus, it seems that the collapsing happens "only once" and happens "uniformly across all of $[G, G]$". The algebraic structure of the Carnot group seems to only affect the constants involved while having no effect on the exponent. We then ask what is the correct dependence on the Lie group structure:[4]

Question 21. *In the case of Euclidean distortion, what are the precise asymptotics? More precisely, we ask the following.*

    (1) *Let G be a nonabelian simply connected nilpotent Lie group. Does there exist a constant c_G such that we have the following?*

$$c_2(B_r) \asymp c_G \sqrt{\log r}, \quad r \geq 2.$$

    Similarly, for $r_1, r_2 > 1$ with $r_1 \geq 2r_2$, let N_{r_1, r_2} be an $(r_2, \Omega(r_2))$-net of B_{r_1}. Does there exist a constant c'_G such that we have the following?

$$c_2(N_{r_1, r_2}) \asymp c'_G \sqrt{\log(r_1/r_2)} \quad \text{as } R \to \infty.$$

    If so, how are the values of c_G and c'_G determined by the algebraic structure of G?

    (2) *Let Γ be a not virtually abelian finitely generated group of polynomial growth. Does there exist a constant c_Γ such that we have the following?*

$$c_2(B_n^\Gamma) \asymp c_\Gamma \sqrt{\log n} \quad \text{as } n \to \infty.$$

    If so, how is the value of c_Γ determined by the algebraic structure of Γ?

We know that, for Carnot groups G, c_G is bounded above and below by functions of n_h, since the constants appearing in the upper bound due to the Assouad embedding theorem and in the lower bound due to Theorem 14 can all be made depending on n_h.

Note that the Carnot groups of lowest nonabelian step $s = 2$, such as the Heisenberg group $\mathbb{H}^3$, already give the worst dependence $\sqrt{\log R}$. We however remark that the nonembeddability statements, both qualitative and quantitative, for Carnot groups into uniformly convex spaces do not formally follow from those of the Heisenberg group $\mathbb{H}^3$, because it is not true that any nonabelian Carnot group contains a bilipschitz copy of $\mathbb{H}^3$.[5] For example, the aforementioned model filiform spaces $J^{s-1}(\mathbb{R})$ with $s \geq 3$ do

---

[4]I thank Professor Assaf Naor for discussion on this matter and asking this question.
[5]I thank Professors Assaf Naor and Robert Young for this comment.

not contain a bilipschitz copy of $\mathbb{H}^3$. Recalling the definition of the model filiform group, $J^{s-1}(\mathbb{R})$ admits the s-step grading $V_1(j^{s-1}) = \text{span}\{x, y_0\}$, $V_2(j^{s-1}) = \text{span}\{y_1\}$, $\cdots$, $V_s(j^{s-1}) = \text{span}\{y_{s-1}\}$. If $J^{s-1}(\mathbb{R})$ were to contain a bilipschitz copy of $\mathbb{H}^3$, then by the Pansu–Semmes differentiation theorem there would be a Carnot group monomorphism $\mathbb{H}^3 \to J^{s-1}(\mathbb{R})$, namely a Lie group monomorphism that commutes with the dilations. This would induce a Lie algebra monomorphism $\mathfrak{h}^3 \to j^{s-1}$ that sends $V_i(\mathfrak{h}) \to V_i(j^{s-1})$ for all i; by matching dimensions, we actually have vector space isomorphisms $V_1(\mathfrak{h}) \xrightarrow{\approx} V_1(j^{s-1})$ and $V_2(\mathfrak{h}) \xrightarrow{\approx} V_2(j^{s-1})$, but now we have a contradiction since this implies we have a surjection

$$0 = V_3(\mathfrak{h}) = [V_1(\mathfrak{h}), V_2(\mathfrak{h})] \to [V_1(j^{s-1}), V_2(j^{s-1})] = V_3(j^{s-1}) \neq 0.$$

Roadmap. The rest of the introduction is organized as follows. We first discuss in detail Dorronsoro's theorem on Carnot groups (Theorem 22) and describe the resulting more refined fractional vertical versus horizontal inequalities (Theorem 24) in subsection 1.1. In subsection 1.2, we present conjectural quantitative nonembeddability statements of nonabelian simply connected nilpotent Lie groups into L^1 (Conjecture 30) to suggest a candidate behavior for the L^1 distortion of balls in groups of polynomial growth (Question 31).

After the introduction, the rest of this paper is organized as follows. We begin by proving our main Theorems 5 and 13 in Section 2, and we then prove the distortion bounds of Theorems 1, 2, 10, 11, 14, and 15 and Corollaries 3 and 4 in Section 3. We derive the discretized inequalities of Theorem 6 in Section 4. In Section 5, we prove Theorem 9 by analyzing cocycles. We then, in Section 6, derive the fractional vertical versus horizontal inequalities of Theorem 24, and thereby also prove Theorem 17, from the Dorronsoro Theorem 22. Finally, we prove the Dorronsoro Theorem 22 in Section 7.

1.1. **Dorronsoro's theorem on Carnot groups.**
As mentioned earlier, a byproduct of this paper is a formulation and proof of a Dorronsoro theorem for Carnot groups, which is used to prove Theorem 17. In this subsection, we state the Dorronsoro theorem and obtain more refined vertical versus horizontal inequalities.

Let G be a Carnot group in this subsection. Following Folland [Fol75] and denoting by $\Delta = \sum_{i=1}^{k} X_i^2$ the sub-Laplacian and H_t the corresponding heat kernel, we define the operator $(-\Delta_p)^\alpha$ for $1 < p < \infty$ and $\text{Re}\,\alpha > 0$ by

$$(-\Delta_p)^\alpha f = \lim_{\varepsilon \to 0} \frac{1}{\Gamma(\lfloor \text{Re}\,\alpha \rfloor + 1 - \alpha)} \int_\varepsilon^\infty t^{\lfloor \text{Re}\,\alpha \rfloor - \alpha} (-\Delta)^{\lfloor \text{Re}\,\alpha \rfloor + 1} H_t f \, dt$$

for all $f \in L^p(G)$ such that the limit exists in the L^p norm. Then, for $1 < p < \infty$ and $\alpha \geq 0$, the Sobolev space S_α^p is the Banach space $\text{Dom}((-\Delta_p)^{\alpha/2})$ with norm

$$\| \cdot \|_{p,\alpha} := \| \cdot \|_{L^p(G)} + \|(-\Delta_p)^{\alpha/2}(\cdot)\|_{L^p(G)}.$$

Fix a basis $X_{r,1}, \cdots, X_{r,k_r}$, where $k_r = \dim V_r$, of each stratum V_r. As G is nilpotent and simply connected, the exponential map $\exp : \mathfrak{g} \to G$ is a diffeomorphism. Thus, each point $p \in G$ can be expressed in the coordinates $p = \exp\left(\sum_{r=1}^{s} \sum_{i=1}^{k_r} x_{r,i} X_{r,i}\right)$, $x_{r,i} \in \mathbb{R}$; let this be the single coordinate chart on G. For polynomials of the coordinates $x_{r,i}$, we assign weight r to the variable $x_{r,i}$, $r = 1, \cdots, s$, $i = 1, \cdots, k_r$.

Recall that $n_h = \sum_{r=1}^{s} r k_r$ is the Hausdorff dimension of G. We have $n_h \geq 4$, as we must have $s \geq 2$, $k_1 \geq 2$ and $k_2 \geq 1$, since G is nonabelian.

For $r > 0$, recall the notation for open balls

$$B_r = \{h \in G : d_G(h, e_G) < r\}, \; B_r(g) = \{h \in G : d_G(h, g) < r\} = g B_r, \quad g \in G, \; r > 0.$$

Recall that δ_λ is a scaling in the Carnot–Carathéodory metric:

$$d_G(\delta_\lambda(p), \delta_\lambda(p')) = \lambda d_G(p, p'), \quad p, p' \in G.$$

Thus $B_r = \delta_r(B_1)$, $r > 0$. Also, by compactness we have

$$d_G(y, e_G) \asymp_G \sum_{r=1}^{s} \sum_{i=1}^{k_r} |y_{r,i}|^{1/r}.$$

Let $d \in \mathbb{Z}_{\geq 0}$ and let $\mathscr{A}_d$ denote the family of polynomials $G \to \mathbb{R}$ of weighted degree d. (Note that this family is left-invariant. Indeed, one can express the group law in this coordinate system using the Baker–Campbell–Hausdorff formula

$$gh = \sum_{m=1}^{s} \frac{(-1)^{m-1}}{m} \sum_{\substack{r_1+s_1>0 \\ \cdots \\ r_m+s_m>0}} \frac{[g^{r_1} h^{s_1} g^{r_2} h^{s_2} \cdots g^{r_m} h^{s_m}]}{(\sum_{j=1}^{m}(r_j+s_j)) \cdot \prod_{i=1}^{m} r_i! s_i!},$$

where the sum is finite since G is of step s, and we have used the following notation:

$$[g^{r_1} h^{s_1} \cdots g^{r_m} h^{s_m}] = [\underbrace{g, [g, \cdots [g}_{r_1}, [\underbrace{h, [h, \cdots [h}_{s_1}, \cdots [\underbrace{g, [g, \cdots [g}_{r_m}, [\underbrace{h, [h, \cdots h}_{s_m}]] \cdots]].$$

Thus, we can see that

$$\left(\sum_{r=1}^{s} \sum_{i=1}^{k_r} x_{r,i}^0 X_{r,i} \right) \left(\sum_{r=1}^{s} \sum_{i=1}^{k_r} x_{r,i}^1 X_{r,i} \right) = \left(\sum_{r=1}^{s} \sum_{i=1}^{k_r} x_{r,i}^2 X_{r,i} \right)$$

where

$$x_{r,i}^2 = x_{r,i}^0 + x_{r,i}^1 + \text{(homogeneous polynomial of } \{x_{r',i'}^0\}_{r'<r}, \{x_{r',i'}^1\}_{r'<r} \text{ of weighted degree } r).$$

Therefore a polynomial of weighted degree d precomposed with a left translation is still of weighted degree d.) This definition of $\mathscr{A}_d$ is in contrast from [FO20a], where they only considered horizontal polynomials, i.e., polynomials that depend only on the 'horizontal coordinates' $x_{1,1}, \cdots, x_{1,k_1}$; here we are allowing for non-horizontal coordinates, provided they satisfy the weighted degree condition.

For a locally integrable function $f : G \to \mathbb{R}$, $x \in G$ and $r > 0$, let $A_{x,r}^d f$ denote the unique element of $\mathscr{A}_d$ such that

$$\int_{B_r(x)} (f(y) - A_{x,r}^d f(y)) A(y) \, dy = 0, \quad \forall A \in \mathscr{A}_d.$$

For example, $A_{x,r}^0 f = \langle f \rangle_{B_r(x)}$, the average of f on $B_r(x)$, and a formula for $A_{x,r}^1 f$ is given below in (54). We measure how well $A_{x,r}^d f$ approximates f in the ball $B_r(x)$ by the following quantity:

$$\beta_{f,d,q}(B_r(x)) = \left(\fint_{B_r(x)} |f(y) - A_{x,r}^d f(y)|^q dy \right)^{1/q}, \quad 1 \leq q < \infty.$$

Theorem 22 (Dorronsoro's theorem for Carnot groups). *Let $1 < p < \infty$, $\alpha > 0$, and*

$$1 \leq q < \frac{\min\{p,2\} n_h}{n_h - \min\{p,2\}}.$$

Then for all $f \in L^p(G)$,

$$\left(\int_G \left(\int_0^\infty \left[\frac{1}{r^\alpha} \beta_{f,\lfloor \alpha \rfloor, q}(B_r(x)) \right]^2 \frac{dr}{r} \right)^{p/2} d\mu(x) \right)^{1/p} \asymp_{G,\alpha,p,q} \| (-\Delta_p)^{\alpha/2} f \|_{L^p(G)}. \tag{8}$$

in the sense that $f \in S_\alpha^p(G)$ if and only if the left-hand side of (8) is finite, in which case the above relation holds.

Our proof of Theorem 22 is based on Dorronsoro's original proof [Dor85] and adds on ingredients from Fässler and Orponen's proof [FO20a]. Actually, in the case of $\alpha = 1$ and $G = \mathbb{H}^{2k+1}$ there is a simpler proof of Dorronsoro's theorem involving the Fourier transform [Azz16, Subsection 7.3].[6] It seems likely that

---

[6]I thank Ian Fleschler for pointing out this reference.

there would be a similar simpler proof of Theorem 22 for $\alpha = 1$, but we have not pursued this direction since we can obtain the full range $\alpha > 0$ with our proof.

Remark 23. *Fässler and Orponen's version* [FO20a] *of the Dorronsoro theorem* [Dor85] *for Heisenberg groups states the $\lesssim$ portion of the above inequality* (8) *for $f \in S_\alpha^p(G)$ in the case $G = \mathbb{H}^{2k+1}$ and $0 < \alpha < 2$. (They did not prove the $\gtrsim$ portion because it is not needed when proving the vertical versus horizontal inequalities.) The restriction $0 < \alpha < 2$ was necessary because they were approximating by horizontal polynomials, i.e., polynomials that depend only on the horizontal coordinates; in the range $0 < \alpha < 2$ our formulation of Theorem 22 agrees with that of* [FO20a] *because then $\mathscr{A}_1$ is precisely the family of horizontal polynomials. To see why the restriction $0 < \alpha < 2$ is necessary when we consider only horizontal polynomials, let $\alpha \geq 2$, and suppose $\mathscr{A}_d$ were defined as the set of horizontal polynomials of degree at most d. In the case of the Heisenberg group $\mathbb{H}^{2k+1}$ the group structure is given by*

$$\exp\left(\sum_{i=1}^{k}(x_i^0 \partial_{x_i} + y_i^0 \partial_{y_i}) + z^0 \partial_z\right) \cdot \exp\left(\sum_{i=1}^{k}(x_i^1 \partial_{x_i} + y_i^1 \partial_{y_i}) + z^1 \partial_z\right)$$

$$= \exp\left(\sum_{i=1}^{k}\left((x_i^0 + x_i^1)\partial_{x_i} + (y_i^0 + y_i^1)\partial_{y_i}\right) + \left(z^0 + z^1 + \frac{1}{2}\sum_{i=1}^{k}(x_i^0 y_i^1 - x_i^1 y_i^0)\right)\partial_z\right)$$

and thus the left-invariant horizontal vector fields are given by

$$X_i = \frac{\partial}{\partial x_i} + \frac{y_i}{2}\frac{\partial}{\partial z}, \quad Y_i = \frac{\partial}{\partial y_i} - \frac{x_i}{2}\frac{\partial}{\partial z}, \quad i = 1, \cdots, k,$$

so that the Laplacian is given by

$$\Delta = \sum_{i=1}^{k}(X_i^2 + Y_i^2) = \sum_{i=1}^{k}\left(\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + y_i\frac{\partial^2}{\partial x_i \partial z} - x_i\frac{\partial^2}{\partial y_i \partial z}\right) + \frac{1}{4}\sum_{i=1}^{k}(x_i^2 + y_i^2)\frac{\partial^2}{\partial z^2}.$$

Let $f : \mathbb{H}^{2k+1} \to \mathbb{R}$ be a function which is smooth, supported on B_2 and agrees with the function z on B_1. Clearly $\|(-\Delta)^{\alpha/2}f\|_{L^p(G)}$ is finite. However, for $r \in (0, \frac{1}{2})$, $A_{0,r}^d f = A_{0,r}^d z = 0$ because B_r is symmetric about reflection with respect to the plane $z = 0$; thus

$$\beta_{f,d,q}(B_r) = \left(\fint_{B_r}|z|^q d\mu\right)^{1/q} \asymp_{G,q} r^2.$$

Since $\mathscr{A}_d$ is invariant under left-translation and left-translation is measure-preserving, at

$$p = \exp\left(\sum_{i=1}^{k}(x_i^0 \partial_{x_i} + y_i^0 \partial_{y_i}) + z^0 \partial_z\right) \in B_{1/2}$$

we have

$$A_{p,r}^d f = z^0 + \frac{1}{2}\sum_{i=1}^{k}(x_i^0 y_i - x_i y_i^0), \quad \beta_{f,d,q}(B_r(x)) = \beta_{f,d,q}(B_r) \asymp_{G,q} r^2.$$

Therefore, the left-hand side of (8) *is bounded below by*

$$\left(\int_{B_{1/2}}\left(\int_0^{1/2}\left[\frac{1}{r^\alpha}\beta_{f,\lfloor\alpha\rfloor,q}(B_r(x))\right]^2 \frac{dr}{r}\right)^{p/2} d\mu(x)\right)^{1/p} \asymp_{G,q} \mu(B_{1/2})^{1/p}\left(\int_0^{1/2}\frac{dr}{r^{1+2(\alpha-2)}}\right)^{1/2} = \infty.$$

Thus (8) *fails to hold when $\alpha \geq 2$ when $\mathscr{A}_d$ is restricted to horizontal polynomials.*

Of course, the purpose of the Dorronsoro Theorem 22 in this paper is to prove the vertical versus horizontal inequality of Theorem 17. The proof of Theorem 17 from Theorem 22 is given in Section 6, and uses the special case of $\alpha = 1$ of Theorem 22 along with the fact that $\|(-\Delta_p)^{1/2}f\|_{L^p(G)} \asymp \|\nabla f\|_{L^p(G;\ell_2^k)}$ [CRTN01, (52)] to see that it is enough to upper bound the left-hand side of (7) by the left-hand side of (8).

However, one may wonder what vertical versus horizontal inequalities emerge when we don't specialize to $\alpha = 1$. In this case, we obtain fractional order generalizations of Theorem 5.

Theorem 24. *Let $1 < p \leq 2$, $\alpha > 0$, $n \in \mathbb{N}$, and let $f \in S_\alpha^p(G)$. Let $v \in \exp(V_{\lfloor \alpha/n \rfloor + 1} \oplus \cdots \oplus V_s)$ with $d_G(v, e_G) = 1$. Then*

$$\left(\int_0^\infty \left[\int_G \left(\frac{1}{r^\alpha} \left| \Delta_{\delta_r(v)}^n f(h) \right| \right)^p d\mu(h) \right]^{2/p} \frac{dr}{r} \right)^{1/2} \lesssim_{G,p,\alpha,n} \| (-\Delta_p)^{\alpha/2} f \|_{L^p(G)}, \tag{9}$$

where for $g \in G$ and $F : G \to \mathbb{R}$, $\Delta_g F(x) := F(xg) - F(x)$ denotes the finite difference.

The simple example $n = 1$ is given as follows.

Example 25. *Let $1 < p \leq 2$, $\alpha > 0$, and let $f \in S_\alpha^p(G)$. Let $v \in \exp(V_{\lfloor \alpha \rfloor + 1} \oplus \cdots \oplus V_s)$ with $d_G(v, e_G) = 1$. Then*

$$\left(\int_0^\infty \left[\int_G \left(\frac{|f(h) - f(h\delta_r(v))|}{r^\alpha} \right)^p d\mu(h) \right]^{2/p} \frac{dr}{r} \right)^{1/2} \lesssim_{G,p,\alpha} \| (-\Delta_p)^{\alpha/2} f \|_{L^p(G)}.$$

The case $G = \mathbb{H}^{2k+1}$, $n = 1$, $0 < \alpha < 2$ of Example 25 has been obtained in [FO20b].

In the above, the Dorronsoro Theorem 22 was used for nonembeddings, but there are also other applications. For example, the case of the Heisenberg groups $\mathbb{H}^{2k+1}$, $k \geq 1$, due to [FO20a], is used in the work [CLY22];[7] it stands to reason that our more general result will have similar applications, but we defer this to future investigations.

1.2. L^1-distortion: vertical perimeter versus horizontal perimeter.

Given the discussion so far, one may ask whether the results of this paper follows for $L^1(\sigma)$-targets. Since ℓ^1 is not uniformly convex, our results in this paper do not apply; in fact, that general nonabelian simply connected nilpotent Lie groups G do not embed bilipschitzly into L^1 spaces was proven only recently by Eriksson-Bique, Gartland, Le Donne, Naples, and Nicolussi-Golo [EBGLD+21]. For the Heisenberg group this was proven previously by Cheeger and Kleiner [CK10].

The results of this paper give quantitative nonembeddability of nonabelian simply connected nilpotent Lie groups into uniformly convex spaces. In this subsection, which closely follows Section 4 of [LN14], we will present hypothetical analogues of the results of this paper for quantitative nonembeddability into L^1.

Theorem 5 states in the case $X = \mathbb{R}$ and $p \in (1, q]$ that for every smooth and compactly supported $f : G \to \mathbb{R}$,

$$\left(\int_0^\infty \left(\int_G |f(hv^t) - f(h)|^p d\mu(h) \right)^{q/p} \frac{dt}{t^{1+q/s}} \right)^{1/q} \lesssim_G (p-1)^{1/q-1} \left(\int_G \| \nabla f(h) \|_{\ell_2^k}^p d\mu(h) \right)^{1/p}. \tag{10}$$

The constant $(p-1)^{1/q-1}$ is unbounded as $p \to 1$; nevertheless, we ask whether the endpoint case $p = 1$ of (10) does hold true.

Question 26. *Let G be a nonabelian simply connected nilpotent Lie group, and let $v \in Z(G)$ be as in Theorem 5. For which exponents $q \geq 1$ does every smooth and compactly supported $f : G \to \mathbb{R}$ satisfy*

$$\left(\int_0^\infty \left(\int_G |f(hv^t) - f(h)| d\mu(h) \right)^q \frac{dt}{t^{1+q/s}} \right)^{1/q} \lesssim_{G,q} \int_G \| \nabla f(h) \|_{\ell_2^k} d\mu(h). \tag{11}$$

Similarly, we may ask the following question.

---

[7]I thank Professor Tuomas Orponen for pointing out this fact.

Question 27. *Let Γ be a not virtually abelian finitely generated group of polynomial growth. Choose v_Γ and $s > 0$ as in Theorem 6. For which exponents $q \geq 1$ do there exist $c = c(\Gamma) \in \mathbb{N}$ such that for any $n \in \mathbb{N}$ and $f : \Gamma \to L^1$ we have the following?*

$$\left(\sum_{k=1}^{n^s} \frac{1}{k^{1+q/s}} \sum_{x \in B_n^\Gamma} \| f(xv_\Gamma^k) - f(x) \|_{L^1}^q \right)^{1/q} \lesssim_{\Gamma, q} \sum_{x \in B_{cn}^\Gamma} \sum_{a \in S} \| f(xa) - f(x) \|_{L^1}.$$

By the discretization argument of Section 4, an exponent q that answers Question 26 positively also answers Question 27 positively.

Because the case $q = \infty$ formally holds for Question 26 (see Remark 18), it is clear that the set of q that satisfies Question 26 is either of the form (q_G, ∞) or $[q_G, \infty)$ for some $q_G \in [1, \infty]$ depending on G. The exponent range is $[4, \infty)$ for $G = \mathbb{H}^3$ [NY20] and is $[2, \infty)$ for $G = \mathbb{H}^k$, $k \geq 5$ [NY18]. We wish to know whether there are finite exponents q satisfying (11), since then not only would we have a quantitative proof of [EBGLD$^+$21], but we would also have quantitative nonembeddability statements analogous to Theorem 10, namely that if N_{r_1, r_2} is a r_2-covering of B_{r_1}, $r_1 \geq 2r_2$, $r_1, r_2 > 1$, then

$$c_1(N_{r_1, r_2}) \gtrsim_{G, q} (\log(r_1/r_2))^{1/q}.$$

Also, if Γ is a not virtually abelian finitely generated group of polynomial growth that is quasi-isometric to G, we would have

$$c_1(B_n^\Gamma) \gtrsim_{\Gamma, q} (\log n)^{1/q}, \quad n \geq 2.$$

By the co-area formula it suffices to prove (11) when f is an indicator of a measurable set $A \subseteq G$, in which case the right-hand side of (11) is interpreted as the horizontal perimeter PER(A) of A (see [Amb01] and [CK10, Section 2] for a precise definition).

Definition 28 (Vertical perimeter at scale t [LN14, NY18]). *Let $v \in Z(G)$ be as in Theorem 5. Let $A \subseteq G$ be measurable and $t \in (0, \infty)$. The vertical perimeter $v_t(A)$ of A at scale t is defined as the quantity*

$$v_t(A) := \mu\left(\{ h \in A : hv^t \notin A \quad \text{or} \quad hv^{-t} \notin A \} \right). \tag{12}$$

By this definition, we may reformulate Question 26 into an isoperimetric inequality.

Question 29. *For which exponents q is it true that for every measurable $A \subseteq G$ one has*

$$\left(\int_0^\infty \frac{v_t(A)^q}{t^{1+q/s}} \, dt \right)^{1/q} \lesssim_{G, q} \text{PER}(A). \tag{13}$$

Of course, the set of q that satisfies Question 29 is the same as that of Question 26. In light of the nonembeddability result [EBGLD$^+$21] and the situation in the Heisenberg group, we make the following conjecture.

Conjecture 30. *There exist finite exponents q that answer Questions 26 and 29 positively, and the infimum q_G among such q is attained.*

Conditioned on Conjecture 30, we would have for a $(r_2, \Omega(r_2))$-net N_{r_1, r_2} of B_{r_1}, where $r_1, r_2 > 1$ with $r_1 \geq 2r_2$, that $c_1(N_{r_1, r_2}) \gtrsim_G (\log(r_1/r_2))^{1/q_G}$, and for a not virtually abelian finitely generated group of polynomial growth Γ we would have $c_1(B_n^\Gamma) \gtrsim_\Gamma (\log n)^{1/q_G}$, $n \geq 2$. However, for the 5 or higher dimensional Heisenberg groups $\mathbb{H}^{2k+1}$, $k \geq 2$, with $q_{\mathbb{H}^{2k+1}} = 2$ we have the matching upper bounds $c_1(N_{r_1, r_2}) \lesssim \sqrt{\log(r_1/r_2)}$ and $c_1(B_n^{\mathbb{H}_\mathbb{Z}^{2k+1}}) \lesssim \sqrt{\log n}$ by the Assouad embedding theorem, while for the 3-dimensional Heisenberg group $\mathbb{H}^3$ we have the matching upper bound $c_1(N_{r_1, r_2}) \lesssim (\log(r_1/r_2))^{1/4}$ and $c_1(B_n^{\mathbb{H}_\mathbb{Z}^3}) \lesssim (\log n)^{1/4}$ by [NY20, Theorem 3.1]. We thus pose the following question as well.

Question 31. *For a $(r_2, \Omega(r_2))$-net N_{r_1, r_2} of B_{r_1}, does $c_1(N_{r_1, r_2}) \asymp_G (\log(r_1/r_2))^{1/q_G}$, where $r_1, r_2 > 1$ with $r_1 \geq 2r_2$? Furthermore, let Γ be a not virtually abelian finitely generated group of polynomial growth which is quasi-isometric to G. Then does $c_1(B_n^\Gamma, d_W) \asymp_G (\log n)^{1/q_G}$, $n \geq 2$?*

2. Proof of Theorem 5: differentiation along random paths

We will prove a theorem slightly more general than Theorems 5 and 13, namely Theorem 32 below. We first explain the terminology behind the statements.

A Lie group G is said to be unimodular if its left-invariant Haar measure μ is also right-invariant. Given left-invariant vector fields $X_1, \cdots, X_k$, the pointwise span of $X_1, \cdots, X_k$ forms a left-invariant vector sub-bundle B over G, and on each fibre of the vector bundle B we may define a left-invariant Euclidean norm[8] $|\cdot|$ that has $X_1, \cdots, X_k$ as an orthonormal basis (this generalizes our earlier choice of $X_1, \cdots, X_k$, which we took to be a basis at e_G in the setting of nilpotent Lie groups and a basis of V_1 in the case of Carnot groups). We may define the associated *sub-Riemannian distance* as the Carnot-Carathéodory distance associated to B to $|\cdot|$, i.e.,

$$d_G(p,q) := \inf\left\{ \int_0^1 |\dot{\gamma}(t)| dt : \gamma \in C_{\mathrm{pw}}^\infty([0,1];G), \gamma(0) = p, \gamma(1) = q, \dot{\gamma} \in B \right\}, \quad p, q \in G,$$

If $X_1, \cdots, X_k$ and their brackets generate $\mathfrak{g}$, and if G is connected, then $d_G(\cdot, \cdot)$ is finite everywhere. The center $Z(\mathfrak{g})$ of the Lie algebra $(\mathfrak{g}, [,])$ of G consists of elements $v \in \mathfrak{g}$ such that $[g, h] = 0$ for all $h \in \mathfrak{g}$, and has the property that if $v \in Z(\mathfrak{g})$ then $\exp(v) \in Z(G)$, the group theoretic center of G.

The following is an extension of Theorems 5 and 13; it is clear that Theorem 32 implies Theorems 5 and 13.

Theorem 32. *Let G be a unimodular Lie group with Haar measure μ, left-invariant vector fields $X_1, \cdots, X_k$, and associated sub-Riemannian distance $d_G(\cdot, \cdot)$. Suppose there is an element $v \in Z(\mathfrak{g})$ such that*

$$d_G(\exp(tv), e_G) \le t^{1/\rho} \quad \forall t > 0$$

for some real number $\rho > 1$ and such that there is a subset S of G with measure ν such that the push-forward of the product measure of ν and the Lebesgue measure of $\mathbb{R}$ under the map $S \times \mathbb{R} \to G$, $(s,t) \mapsto s\exp(tv)$ is the Haar measure μ.

Suppose that $p \in (1, \infty)$ and $q \in [2, \infty)$. Let $(X, \|\cdot\|_X)$ be a Banach space with $K_q(X) < \infty$. If $p \ge q$, then every smooth and compactly supported $f : G \to X$ satisfies

$$\left(\int_0^\infty \int_G \left(\frac{\|f(h\exp(tv)) - f(h)\|_X}{t^{1/\rho}} \right)^p d\mu(h) \frac{dt}{t} \right)^{1/p} \lesssim \frac{\rho}{\rho - 1} K_q(X) \left(\int_G \|\nabla f(h)\|_{\ell_2^k(X)}^p d\mu(h) \right)^{1/p}. \tag{14}$$

If $p < q$, and if for any $t > 0$ we have $\lim_{r \to \infty} \mu(B_{r+t})/\mu(B_r) = 1$, then

$$\left(\int_0^\infty \left(\int_G \left(\frac{\|f(h\exp(tv)) - f(h)\|_X}{t^{1/\rho}} \right)^p d\mu(h) \right)^{q/p} \frac{dt}{t} \right)^{1/q}$$

$$\lesssim \frac{\rho}{\rho - 1} \max\left\{ (p-1)^{1/q-1}, K_q(X) \right\} \left(\int_G \|\nabla f(h)\|_{\ell_2^k(X)}^p d\mu(h) \right)^{1/p}. \tag{15}$$

The hypothesis for Theorem 32 is satisfied when G is a simply connected nilpotent group. The product decomposition of the Haar measure easily follows from that of the Lebesgue measure on $\mathbb{R}^n$. The distance requirement follows by [BLD12, Proposition 2.13] and [Jea14, Theorem 2.1], which tells us that if $z \in Z(\mathfrak{g}) \setminus \{0\}$ then

$$d_G(\exp(tv), e_G) \asymp_G \begin{cases} t^{1/s}, & t \ge 1, \\ t^{1/r}, & 0 \le t < 1, \end{cases} \tag{16}$$

---

[8]Again, we could also take any general left-invariant norm on B, in which case the distance is sub-Finsler. Compared to the sub-Riemannian case, the resulting distance d_G is then distorted by a factor of at most $\sqrt{\dim V_1}$ by the John ellipsoid theorem [Joh48], and the results of this paper follow up to multiplicative factors of $\sqrt{\dim V_1}$.

where $s \geq 2$ is the nilpotency step of G and r is the smallest integer that $v \in \underbrace{[B, [B, \cdots, B]]}_{r \text{ times}}$ (recall B is the left-invariant bundle spanned by $X_1, \cdots, X_k$). By applying Theorem 32 with $\rho = s$ and $\rho = \max\{r, 1 + \varepsilon\}$, we obtain the following.

Corollary 33. *Let G be a nonabelian simply connected nilpotent Lie group with left-invariant vector fields $X_1, \cdots, X_k$ such that the X_i's and their brackets generate $\mathfrak{g}$. Let $v \in Z(\mathfrak{g}) \setminus \{0\}$. Suppose that $p \in (1, \infty)$ and $q \in [2, \infty)$. Let $(X, \|\cdot\|_X)$ be a Banach space with $K_q(X) < \infty$, and let $f : G \to X$ be smooth and compactly supported.*

(1) *Let $v \in Z(\mathfrak{g}) \setminus \{0\}$ be normalized so that $d_G(v^t, e_G) \leq t^{1/s}$ for $t > 0$, where s is the nilpotency step of G. Then*

$$\left(\int_1^\infty \left(\int_G \left(\frac{\|f(hv^t) - f(h)\|_X}{t^{1/s}} \right)^p d\mu(h) \right)^{\max\{p,q\}/p} \frac{dt}{t} \right)^{1/\max\{p,q\}}$$

$$\lesssim \max\left\{ (p-1)^{1/q-1}, K_q(X) \right\} \left(\int_G \|\nabla f(h)\|_{\ell_2^k(X)}^p d\mu(h) \right)^{1/p}. \tag{17}$$

(2) *Let r be the smallest integer that $v \in \underbrace{[B, [B, \cdots, B]]}_{r \text{ times}}$. If $r \geq 2$, and if $v \in Z(\mathfrak{g}) \setminus \{0\}$ is normalized so that $d_G(v^t, e_G) \leq t^{1/r}$ for $t > 0$. Then*

$$\left(\int_0^1 \left(\int_G \left(\frac{\|f(hv^t) - f(h)\|_X}{t^{1/r}} \right)^p d\mu(h) \right)^{\max\{p,q\}/p} \frac{dt}{t} \right)^{1/\max\{p,q\}}$$

$$\lesssim \max\left\{ (p-1)^{1/q-1}, K_q(X) \right\} \left(\int_G \|\nabla f(h)\|_{\ell_2^k(X)}^p d\mu(h) \right)^{1/p}. \tag{18}$$

If $r = 1$, then for any $\varepsilon > 0$, if $v \in Z(\mathfrak{g}) \setminus \{0\}$ is normalized so that $d_G(v^t, e_G) \leq t^{1/(1+\varepsilon)}$ for $t > 0$, then

$$\left(\int_0^1 \left(\int_G \left(\frac{\|f(hv^t) - f(h)\|_X}{t^{1/(1+\varepsilon)}} \right)^p d\mu(h) \right)^{\max\{p,q\}/p} \frac{dt}{t} \right)^{1/\max\{p,q\}}$$

$$\lesssim \frac{1+\varepsilon}{\varepsilon} \max\left\{ (p-1)^{1/q-1}, K_q(X) \right\} \left(\int_G \|\nabla f(h)\|_{\ell_2^k(X)}^p d\mu(h) \right)^{1/p}. \tag{19}$$

Based on these theorems, we have the following nonembeddability statements.

Theorem 34. *Let the Lie group G, left-invariant vector fields $X_1, \cdots, X_k$, and $v \in Z(\mathfrak{g})$ satisfy the hypotheses of Theorem 32. Suppose $X_1, \cdots, X_k$ and their brackets generate $\mathfrak{g}$. Let X be a q-uniformly convex Banach space, with $q \geq 2$.*

(1) *If $d_G(\exp(tv), e_G) \asymp_G t^{1/\rho}$ for $0 < t < 1$ for some $\rho > 1$, then any nonempty open subset U of G fails to embed bilipschitzly into X.*

(2) *If $d_G(\exp(tv), e_G) \asymp_G t^{1/\rho}$ for $t \geq 1$ for some $\rho > 1$, and if the Haar measure μ of G is close to being Ahlfors regular at large scales in the sense that for any constant $c' \geq 1$*

$$\mu(B_{c'r})/\mu(B_r) \asymp_G 1, \quad r \geq 1,$$

then $c_X(B_r) \gtrsim_G \frac{\rho-1}{\rho} \cdot \frac{(\rho \log r)^{1/q}}{K_q(X)}$ for $r = \Omega_G(1)$. In particular, G fails to bilipschitzly embed into X. We have $c_p(B_r) \asymp (\log r)^{1/\max\{p,2\}}$ for $1 < p < \infty$.

Indeed, for (1), we may assume by translation that $B_r \subset U$ for some $r > 0$. If $f : U \to X$ were a bilipschitz embedding, $d_G(x, y) \leq \|f(x) - f(y)\|_X \leq D d_G(x, y)$, $x, y \in U$, we may assume by translation that $f(e_G) = 0$. Multiplying f by a smooth cutoff function, we may construct $F : G \to X$ which is Lipschitz, agrees with f on $B_{r/3}$, and is supported on $B_{2r/3}$. By a smooth approximation argument, we may apply (14) to F. Let

$c_1, c_2 > 0$ be such that $c_1 t^{1/\rho} \le d_G(\exp(tv), e_G) \le c_2 t^{1/\rho}$ for $0 < t < 1$. Then the right-hand side of (14) is a finite quantity, whereas the left-hand side is at least

$$\left(\int_0^{r^\rho/c_2^\rho 6^\rho} \int_{B_{r/6}} \left(\frac{\|f(h\exp(tv)) - f(h)\|_X}{t^{1/\rho}} \right)^q d\mu(h) \frac{dt}{t} \right)^{1/q} = \left(\mu(B_{r/6}) \int_0^{r^\rho/c_2^\rho 6^\rho} c_1^q \frac{dt}{t} \right)^{1/q} = \infty,$$

giving a contradiction.

For (2), if $f : B_r \to X$ satisfied $d_G(x, y) \le \|f(x) - f(y)\|_X \le D d_G(x, y)$ for $x, y \in B_r$ and, then translating and multiplying f by a cutoff function we may construct $F : G \to X$ which is $3D$-Lipschitz, agrees with f on $B_{r/2}$, and is supported on B_r. Again, let $c_1, c_2 > 0$ be such that $c_1 t^{1/\rho} \le d_G(\exp(tv), e_G) \le c_2 t^{1/\rho}$ for $t > 1$. Applying (14), we have

$$c_1 \mu(B_{r/4})^{1/q} \log(r^\rho/c_2^\rho 4^\rho)^{1/q} \lesssim_G \left(\int_1^{r^\rho/c_2^\rho 4^\rho} \int_{B_{r/4}} \left(\frac{\|f(h\exp(tv)) - f(h)\|_X}{t^{1/\rho}} \right)^q d\mu(h) \frac{dt}{t} \right)^{1/q}$$

$$\lesssim \frac{\rho}{\rho-1} K_q(X) \left(\int_{B_r} \|\nabla F\|_{\ell_2^k(X)}^q d\mu \right)^{1/q} \lesssim_G \frac{\rho}{\rho-1} K_q(X) \mu(B_r)^{1/q} D.$$

This gives the stated estimate.

By (16), we have the following.

Corollary 35. *Let G be a nonabelian simply connected nilpotent Lie group with left-invariant vector fields $X_1, \cdots, X_k$ such that the X_i's and their brackets generate $\mathfrak{g}$. Suppose that $q \in [2, \infty)$ and let $(X, \|\cdot\|_X)$ be a Banach space with $K_q(X) < \infty$.*

(1) If $Z(\mathfrak{g}) \not\subseteq \mathrm{span}\{X_1, \cdots, X_k\}$, then any nonempty open subset U of G fails to embed bilipschitzly into X.

(2) If $Z(\mathfrak{g}) \subseteq \mathrm{span}\{X_1, \cdots, X_k\}$, we have $c_X(B_r) \gtrsim_G \frac{(\log r)^{1/q}}{K_q(X)}$ for $r \ge 2$. In particular, G fails to bilipschitzly embed into X.

Before we begin the proof of Theorem 32, for motivation we begin by proving Remark 18 which states that the case $q = \infty$ of Theorem 17 formally holds.

Proof of Remark 18. It is enough to prove that, for each $r > 0$,

$$\left[\int_G \left(\frac{|f(h) - f(h\delta_r(v))|}{r} \right)^p d\mu(h) \right]^{1/p} \le \|\nabla f\|_{L^p(G; \ell_2^k)}.$$

Let $\gamma : [0, r] \to G$ be a piecewise smooth horizontal curve of unit speed connecting e_G to $\delta_r(v)$. Then

$$|f(h) - f(h\delta_r(v))| \le \int_0^r |\nabla f(h\gamma(s)) \cdot dL_h \dot\gamma(s)| ds \le \int_0^r |\nabla f(h\gamma(s))| ds.$$

Thus, by Jensen's inequality and right-invariance of the Haar measure μ,

$$\left[\int_G \left(\frac{|f(h) - f(h\delta_r(v))|}{r} \right)^p d\mu(h) \right]^{1/p} \le \left[\frac{1}{r} \int_G \int_0^r |\nabla f(h\gamma(s))|^p ds d\mu(h) \right]^{1/p} \le \|\nabla f\|_{L^p(G; \ell_2^k)},$$

and we are done. $\square$

Likewise, it is not hard to obtain a proof of the $q = \infty$ case of Theorem 32 by the above proof. This proof basically just takes a path from h to $h\delta_r(v)$ and differentiates f along that path; however, the resulting inequality is weak. To make up for this weakness, we will choose over a random distribution of paths, the randomness arising from a heat flow on the line spanned by v. We will then apply the above proof method to each of those random paths (in Lemma 37). This will result in an upper bound given in terms of a convolution of ∇f with a derivative of the heat kernel, which we treat using Littlewood–Paley–Stein g-function estimates. This proof method is borrowed from and extends that of [LN14].

We set some notation and terminology.

Recall there is $v \in Z(\mathfrak{g})$ and a subset S of G with measure ν such that the push-forward of the product measure of ν and the Lebesgue measure of $\mathbb{R}$ under the map $S \times \mathbb{R} \to G$, $(s,t) \mapsto s\exp(tv)$ is the Haar measure μ. We denote $v^t := \exp(tv) \in G$. For $p \in [1,\infty)$, the Lebesgue–Bochner spaces

$$L^p(\mathbb{R}, \text{Lebesgue}; X) = L^p(\mathbb{R}; X), \quad L^p(S, \nu; X) = L^p(S; X), \quad L^p(G, \mu; X) = L^p(G; X)$$

are defined. For $f \in L^p(G; X)$ define $f^v : \mathbb{R} \to L^p(S; X)$ by

$$f^v(z)(x) = f(xv^z), \quad x \in S, \ z \in \mathbb{R}. \tag{20}$$

Given $\psi \in L^1(\mathbb{R})$, define the convolution $\psi * f := \psi * f^v \in L^p(G; X)$,

$$\psi * f(xv^z) := \int_{\mathbb{R}} \psi(u) f(xv^{z-u}) \, du \in X.$$

So $\psi * f$ is the usual group convolution of f with the measure supported on $\exp(\text{span}(v))$ whose density is ψ.

We will take ψ to be the heat kernel or its derivatives. The heat kernel on $\mathbb{R}$ is defined for $t > 0$, $x \in \mathbb{R}$ as

$$h_t(x) := \frac{1}{2\sqrt{\pi t}} e^{-\frac{x^2}{4t}}, \quad x \in \mathbb{R},$$

and has derivatives

$$\dot{h}_t(x) = \partial_t h_t(x) = \frac{x^2 - 2t}{8\sqrt{\pi} t^{5/2}} e^{-\frac{x^2}{4t}}, \quad \partial_x h_t(x) = -\frac{x}{4\sqrt{\pi} t^{3/2}} e^{-\frac{x^2}{4t}}.$$

Then for a universal constant C,

$$\int_0^\infty x^{1/\rho} h_t(x) \, dx = \Gamma\left(\frac{1}{2} + \frac{1}{2\rho}\right) \frac{(4t)^{1/\rho}}{2\sqrt{\pi}}, \quad \|\dot{h}_t\|_{L^1(\mathbb{R})} = \frac{C}{t}, \quad \|\partial_x h_t\|_{L^1(\mathbb{R})} = \frac{1}{\sqrt{\pi t}} \qquad t > 0. \tag{21}$$

We now begin the proof of Theorem 32. First, note that the case $p > q$ of Theorem 32 follows from that of $p = q$, because $K_q(X) \geq K_p(X)$. Thus, it is enough to consider the cases $p = q$ and $p < q$. The case $p < q$ will also be shown to follow from that of $p = q$, albeit in a more complicated way. We prove the case $p = q$ first.

2.1. **The case $p = q$.**

2.1.1. *Step 1.* The first step is to bound the vertical variations $f(gv^t) - f(g)$ by the quantities $\dot{h}_t * f$. This corresponds to Lemma 2.6 of [LN14], which is a modification of the corresponding statement in [Ste16].

Lemma 36 (Analogue of [LN14, Lemma 2.6])**.** *Fix $p \in [1,\infty)$ and let $(X, \|\cdot\|_X)$ be a Banach space. Every smooth and compactly supported $f : G \to X$ satisfies*

$$\left(\int_0^\infty \int_G \|f(gv^t) - f(g)\|_X^p \, dg \, \frac{dt}{t^{1+p/\rho}}\right)^{1/p} \lesssim \frac{\rho}{\rho - 1} \left(\int_0^\infty t^{p\frac{2\rho-1}{2\rho}-1} \|\dot{h}_t * f\|_{L^p(G;X)}^p \, dt\right)^{1/p}. \tag{22}$$

This is the part of the argument that uses that $\rho > 1$. In the nilpotent setting, this means that our group G has nilpotency step $s > 1$ and hence is nonabelian.

Proof. For every $g \in G$ and $t > 0$ we have

$$f(gv^t) - f(g) = [f(gv^t) - h_{t^2} * f(gv^t)] + [h_{t^2} * f(gv^t) - h_{t^2} * f(g)] + [h_{t^2} * f(g) - f(g)]$$

so by the triangle inequality

$$\left(\int_0^\infty \int_G \|f(gv^t) - f(g)\|_X^p\, dg\, \frac{dt}{t^{1+p/\rho}}\right)^{1/p} \leq \left(\int_0^\infty \int_G \|f(gv^t) - h_{t^2} * f(gv^t)\|_X^p\, dg\, \frac{dt}{t^{1+p/\rho}}\right)^{1/p}$$

$$+ \left(\int_0^\infty \int_G \|h_{t^2} * f(gv^t) - h_{t^2} * f(g)\|_X^p\, dg\, \frac{dt}{t^{1+p/\rho}}\right)^{1/p}$$

$$+ \left(\int_0^\infty \int_G \|h_{t^2} * f(g) - f(g)\|_X^p\, dg\, \frac{dt}{t^{1+p/\rho}}\right)^{1/p}.$$

We bound each term by the right-hand side of (22). First,

$$\left(\int_0^\infty \int_G \|f(gv^t) - h_{t^2} * f(gv^t)\|_X^p\, dg\, \frac{dt}{t^{1+p/\rho}}\right)^{1/p}$$

$$= \left(\int_0^\infty \int_G \left\|\int_0^{t^2} \dot{h}_\tau * f(gv^t)\, d\tau\right\|_X^p\, dg\, \frac{dt}{t^{1+p/\rho}}\right)^{1/p}$$

$$\leq \left(\int_0^\infty \left(\int_0^{t^2} \|\dot{h}_\tau * f\|_{L^p(G;X)}\, d\tau\right)^p\, \frac{dt}{t^{1+p/\rho}}\right)^{1/p}$$

$$\leq 2^{1-1/p}\rho\left(\int_0^\infty t^{p\frac{2\rho-1}{2\rho}-1}\|\dot{h}_t * f\|_{L^p(G;X)}^p\, dt\right)^{1/p},$$

where in the first inequality we used the triangle inequality in $L^p(G;X)$ and the fact that the Haar measure on G is right-invariant, and in the second inequality we used Hardy's inequality [HLP$^+$52, Theorem 330]

$$\left(\int_0^\infty \left[x^{-\nu}\int_0^x f(t)\, dt\right]^p\, dx\right)^{1/p} \leq \frac{1}{\nu - 1/p}\left(\int_0^\infty x^{(1-\nu)p}f(x)^p\, dx\right)^{1/p}, \quad \nu > \frac{1}{p}, \tag{23}$$

with $\nu = \frac{1}{p} + \frac{1}{2\rho}$. The third term of the right-hand side of (22) is bounded using the same method.

We now bound the second term of the right-hand side of (22). By the semigroup property we have $h_t = h_{t/2} * h_{t/2}$ for $t > 0$, hence $\dot{h}_t = h_{t/2} * \dot{h}_{t/2}$ and so

$$\partial_t \partial_x h_t = \partial_x \dot{h}_t = \partial_x h_{t/2} * \dot{h}_{t/2}.$$

Since by Young's inequality

$$\|\partial_x h_t * f\|_{L^p(G;X)} \leq \|\partial_x h_t\|_{L^1(\mathbb{R})}\|f\|_{L^p(G;X)} = \frac{1}{\sqrt{\pi t}}\|f\|_{L^p(G;X)}, \tag{24}$$

we have $\lim_{t\to\infty}\partial_x h_t * f = 0$ in $L^p(G;X)$, and thus

$$\partial_x h_{t^2} * f = -\int_{t^2}^\infty \partial_\tau(\partial_x h_\tau * f)\, d\tau = -\int_{t^2}^\infty \partial_x h_{\tau/2} * \dot{h}_{\tau/2} * f\, d\tau,$$

and we may write

$$h_{t^2} * f(gv^t) - h_{t^2} * f(g) = \int_0^t \partial_x h_{t^2} * f(gv^u)\, du = -\int_0^t \int_{t^2}^\infty \partial_x h_{\tau/2} * \dot{h}_{\tau/2} * f(gv^u)\, d\tau\, du.$$

21

Observe that for every $t \in (0, \infty)$ we have

$$\left(\int_0^\infty \int_G \| h_{t^2} * f(gv^t) - h_{t^2} * f(g) \|_X^p \, dg \frac{dt}{t^{1+p/\rho}} \right)^{1/p}$$

$$\leq \left(\int_0^\infty \int_G \left\| \int_0^t \int_{t^2}^\infty \partial_x h_{\tau/2} * \dot{h}_{\tau/2} * f(gv^u) \, d\tau \, du \right\|_X^p \, dg \frac{dt}{t^{1+p/\rho}} \right)^{1/p}$$

$$\leq \left(\int_0^\infty \left(\int_0^t \int_{t^2}^\infty \left(\int_G \| \partial_x h_{\tau/2} * \dot{h}_{\tau/2} * f(gv^u) \|_X^p \, dg \right)^{1/p} \, d\tau \, du \right)^p \frac{dt}{t^{1+p/\rho}} \right)^{1/p}$$

$$= \left(\int_0^\infty \left(t \int_{t^2}^\infty \| \partial_x h_{\tau/2} * \dot{h}_{\tau/2} * f) \|_{L^p(G;X)} \, d\tau \right)^p \frac{dt}{t^{1+p/\rho}} \right)^{1/p}$$

$$\leq \left(\int_0^\infty t^{p\frac{\rho-1}{\rho}-1} \left(\int_{t^2}^\infty \| \partial_x h_{\tau/2} * \dot{h}_{\tau/2} * f \|_{L^p(G;X)} \, d\tau \right)^p \, dt \right)^{1/p}$$

$$\leq \frac{2^{1-1/p}\rho}{\rho-1} \left(\int_0^\infty t^{p\frac{3\rho-1}{2\rho}-1} \| \partial_x h_{t/2} * \dot{h}_{t/2} * f \|_{L^p(G;X)}^p \, dt \right)^{1/p},$$

where the second inequality uses the triangle inequality in $L^p(G; X)$, the first equality uses the right-invariance of the Haar measure μ, and the fourth inequality uses the second form of Hardy's inequality [HLP+52, Theorem 330]

$$\left(\int_0^\infty \left[t^\nu \int_t^\infty f(u) \, du \right]^p \, dt \right)^{1/p} \leq \frac{1}{\nu+1/p} \left(\int_0^\infty x^{p(1+\nu)} f(x)^p \, dx \right)^{1/p}, \quad \nu > -1/p \qquad (25)$$

with $\nu = \frac{\rho-1}{2\rho} - \frac{1}{p}$. (This is the step where we use $\rho > 1$.) By Young's inequality,

$$\| \partial_x h_{t/2} * \dot{h}_{t/2} * f \|_{L^p(G;X)} \leq \| \partial_x h_{t/2} \|_{L^1(\mathbb{R})} \| \dot{h}_{t/2} * f \|_{L^p(G;X)} = \sqrt{\frac{2}{\pi t}} \| \dot{h}_{t/2} * f \|_{L^p(G;X)}.$$

Therefore

$$\left(\int_0^\infty \int_G \| h_{t^2} * f(gv^t) - h_{t^2} * f(g) \|_X^p \, dg \frac{dt}{t^{1+p/\rho}} \right)^{1/p} \leq \frac{2^{3/2-1/p-1/2\rho}\rho}{\sqrt{\pi}(\rho-1)} \left(\int_0^\infty t^{p\frac{2\rho-1}{2\rho}-1} \| \dot{h}_t * f \|_{L^p(G;X)}^p \, dt \right)^{1/p}.$$

This completes the proof. $\qquad \square$

2.1.2. *Step 2.* We will next bound $\dot{h}_t * f$ using $\dot{h}_t * \nabla f$. We first prove the following lemma.

Lemma 37. *Suppose that $p \in [1, \infty)$ and $t \in (0, \infty)$. Then for every Banach space $(X, \| \cdot \|_X)$ and every smooth and compactly supported $f : G \to X$ we have*

$$\| \dot{h}_t * f - \dot{h}_{2t} * f \|_{L^p(G;X)} \lesssim t^{1/2\rho} \| \dot{h}_t * \nabla f \|_{L^p(G;\ell_2^k(X))}. \qquad (26)$$

Proof. Recalling the semigroup property $\dot{h}_{2t} = h_t * \dot{h}_t$, we have

$$\dot{h}_t * f(g) - \dot{h}_{2t} * f(g) = \dot{h}_t * f(g) - h_t * \dot{h}_t * f(g) = \int_{\mathbb{R}} h_t(u) \left(\dot{h}_t * f(g) - \dot{h}_t * f(gv^{-u}) \right) du. \qquad (27)$$

For $u \in [0, \infty)$, let $\gamma_u : [0, d_G(v^u, e_G)] \to G$ be a measurable family of geodesics parametrized by arclength joining e_G to v^u. (Such a measurable family exists by the Aumann measurable selection theorem [Bog, Theorem 6.9.13]). For every $u \in [0, \infty)$ and $g \in G$,

$$\dot{h}_t * f(g) - \dot{h}_t * f(gv^{-u}) = \dot{h}_t * f(gv^{-u}\gamma_u(d_G(v^u, e_G))) - \dot{h}_t * f(gv^{-u})$$

$$= \int_0^{d_G(v^u, e_G)} \frac{d}{d\theta} \dot{h}_t * f(gv^{-u}\gamma_u(\theta)) d\theta.$$

Because γ_u is horizontal and parametrized by arclength, and because convolution with $\dot{h}_t$ commutes with ∇ (this is because v is in the center of G), we have for every $\theta \in [0, d_G(v^u, e_G)]$

$$\left\| \frac{d}{d\theta} \dot{h}_t * f(gv^{-u}\gamma_u(\theta)) \right\|_X \lesssim \left\| \dot{h}_t * \nabla f(gv^{-u}\gamma_u(\theta)) \right\|_{\ell_2^k(X)}. \tag{28}$$

We thus obtain

$$\left(\int_G \left\| \int_0^\infty h_t(u) \left(\dot{h}_t * f(g) - \dot{h}_t * f(gv^{-u}) \right) du \right\|_X^p dg \right)^{1/p}$$

$$\lesssim_G \int_0^\infty \int_0^{d_G(v^u, e_G)} h_t(u) \left(\int_G \left\| \dot{h}_t * \nabla f(gv^{-u}\gamma_u(\theta)) \right\|_{\ell_2^k(X)}^p dg \right)^{1/p} d\theta \, du$$

$$\leq \left(\int_0^\infty u^{1/2\rho} h_t(u) du \right) \left\| \dot{h}_t * \nabla f \right\|_{L^p(G;\ell_2^k(X))}$$

$$\overset{(21)-(i)}{\lesssim} t^{1/2\rho} \left\| \dot{h}_t * \nabla f \right\|_{L^p(G;\ell_2^k(X))},$$

where in the first equality we used the right-invariance of the Haar measure on G.

Similarly, for $u \in (-\infty, 0]$,

$$\dot{h}_t * f(g) - \dot{h}_t * f(gv^{-u}) = -\int_0^{d_G(v^u, e_G)} \frac{d}{d\theta} \dot{h}_t * f(g\gamma_{-u}(\theta)) d\theta,$$

and by the same reasoning,

$$\left(\int_G \left\| \int_{-\infty}^0 h_t(u) \left(\dot{h}_t * f(g) - \dot{h}_t * f(gv^{-u}) \right) du \right\|_X^p dg \right)^{1/p} \lesssim t^{1/2\rho} \left\| \dot{h}_t * \nabla f \right\|_{L^p(G;\ell_2^k(X))}.$$

These two estimates along with (27) gives the stated inequality. $\qquad\square$

Now we bound $\dot{h}_t * f$ using $\dot{h}_t * \nabla f$.

Lemma 38. *Fix $p \in [1, \infty)$. For every Banach space $(X, \|\cdot\|_X)$, every smooth and compactly supported $f : G \to X$ satisfies*

$$\left(\int_0^\infty t^{p \frac{2\rho-1}{2\rho} - 1} \left\| \dot{h}_t * f \right\|_{L^p(G;X)}^p dt \right)^{1/p} \lesssim_G \left(\int_0^\infty t^{p-1} \left\| \dot{h}_t * \nabla f \right\|_{L^p(G;\ell_2^k(X))}^p dt \right)^{1/p}.$$

Proof. By Young's inequality,

$$\| \dot{h}_t * f \|_{L^p(G;X)} \leq \| \dot{h}_t \|_{L^1(\mathbb{R})} \| f \|_{L^p(G;X)} \overset{(21)-(ii)}{\asymp} \frac{1}{t} \| f \|_{L^p(G;X)},$$

so $\lim_{t \to \infty} \dot{h}_t * f = 0$ in $L^p(G;X)$. Therefore

$$\dot{h}_t * f = \sum_{m=1}^\infty \left(\dot{h}_{2^{m-1}t} * f - \dot{h}_{2^m t} * f \right),$$

23

from which it follows that

$$\left(\int_0^\infty t^{p\frac{2\rho-1}{2\rho}-1}\|\dot{h}_t * f\|_{L^p(G;X)}^p \, dt\right)^{1/p}$$

$$\le \sum_{m=1}^\infty \left(\int_0^\infty t^{p\frac{2\rho-1}{2\rho}-1}\|\dot{h}_{2^{m-1}t} * f - \dot{h}_{2^m t} * f\|_{L^p(G;X)}^p \, dt\right)^{1/p}$$

$$\lesssim \sum_{m=1}^\infty \left(\int_0^\infty t^{p\frac{2\rho-1}{2\rho}-1}(2^{m-1}t)^{p/\rho}\|\dot{h}_{2^{m-1}t} * \nabla f\|_{L^p(G;\ell_2^k(X))}^p \, dt\right)^{1/p}$$

$$= \left(\sum_{m=1}^\infty \frac{1}{2^{(m-1)(2\rho-1)/2\rho}}\right)\left(\int_0^\infty t^{p-1}\|\dot{h}_t * \nabla f\|_{L^p(G;\ell_2^k(X))}^p \, dt\right)^{1/p},$$

where the second inequality uses Lemma 37. $\qquad\square$

2.1.3. *Step 3.* We are now ready to prove the $p = q$ case of Theorem 32.

Given a Banach space $(\mathbb{B}, \|\cdot\|_\mathbb{B})$, for every function $\phi \in L^p(\mathbb{R};\mathbb{B})$ its generalized Hardy–Littlewood g-function $\mathfrak{G}_p(\phi):\mathbb{R}\to[0,\infty]$ is defined as follows.

$$\mathfrak{G}_p(\phi)(x) := \left(\int_0^\infty t^{p-1}\|\dot{h}_t * \phi(x)\|_\mathbb{B}^p \, dt\right)^{1/p}. \tag{29}$$

By [HN19, Theorem 17], which is a quantitative version of[9] [MTX06, Theorem 2.1] and [Xu20, Theorem 2]

$$\phi \in L^p(\mathbb{R};\mathbb{B}) \implies \|\mathfrak{G}_p(\phi)\|_{L^p(\mathbb{R};\mathbb{B})} \lesssim K_p(\mathbb{B})\|\phi\|_{L^p(\mathbb{R};\mathbb{B})}. \tag{30}$$

We will apply (30) to $\mathbb{B} = L^p(S;\ell_2^k(X))$. By [MN14, Corollary 6.4],

$$K_p(\mathbb{B}) \le \left(\frac{5p^2}{(p-1)^2}\right)^{1-1/p} K_p(\ell_2^k(X)) \le \left(\frac{50p^3}{(p-1)^3}\right)^{1-1/p} K_p(X) \lesssim K_p(X),$$

since $p \ge 2$.

Recalling (20), we choose $\phi = (\nabla f)^\nu : \mathbb{R} \to \mathbb{B}$. Then,

$$\left(\int_0^\infty t^{p-1}\|\dot{h}_t * \nabla f\|_{L^p(G;\ell_2^k(X))}^p \, dt\right)^{1/p} = \left(\int_0^\infty \int_\mathbb{R} \left(t\|\dot{h}_t * \phi(z)\|_\mathbb{B}\right)^p \, dz\frac{dt}{t}\right)^{1/p}$$

$$\le \left(\int_\mathbb{R} \int_0^\infty \left(t\|\dot{h}_t * \phi(z)\|_\mathbb{B}\right)^p \frac{dt}{t}dz\right)^{1/p}$$

$$= \|\mathfrak{G}_p(\phi)\|_{L^p(\mathbb{R};\mathbb{B})}$$

$$\lesssim K_p(X)\|\phi\|_{L^p(\mathbb{R};\mathbb{B})} = K_p(X)\|\nabla f\|_{L^p(G;\ell_2^k(X))}.$$

By this discussion and Lemmas 37 and 38, the proof for the case $p = q$ is complete.

---

[9]More precisely, [MTX06] proves inequality (30) for 'subordinated' semigroups, such as the Poisson semigroup subordinated by the heat semigroup, and [Xu20] proves the inequality for general symmetric diffusion semigroups. The work [HN19] has the advantage that it obtains, for the heat semigroup, the explicit constant $m_p(\mathbb{B})$, the martingale cotype p constant of $\mathbb{B}$, which in turn is bounded by $K_p(\mathbb{B})$ by [Pis75]. Since [HN19] and [Xu20] were unavailable when [LN14] was written, [LN14] worked with the Poisson semigroup. We could have also started with the Poisson semigroup P_t, which gives the same constant in (30) due to semigroup subordination. See [NY20, Appendix A] for a detailed account of this discussion.

2.2. **The case** $p < q$. The following is an imitation of the argument of [NY20, Appendix A]. Recall that for all $t > 0$, $\lim_{r\to\infty} \mu(B_{r+t})/\mu(B_r) = 1$.

For $M > 1$, let $\beta_M : G \to [0,1]$ be an $O(1)$-Lipschitz smooth bump function with $\beta_M = 1$ on B_M and $\operatorname{supp}\beta_M \subset B_{M+1}$. For $f : G \to X$, define $F_M : G \to L^p(G;E)$ by

$$F_M(h)(g) = \beta_M(h)f(gh), \quad g,h \in G.$$

Then by F_M is compactly supported and smooth, so by the $p = q$ case,

$$\left(\int_0^\infty \|D_\nu^t F_M\|^q_{L^q(G;L^p(G;X))} \frac{dt}{t}\right)^{1/q} \lesssim \frac{\rho}{\rho-1} K_q(L^p(G;X)) \|\nabla F_M\|_{L^q(G;\ell_q^k(L^p(G;X)))},$$

where D_ν^t is shorthand for $D_\nu^t f(g) = \frac{f(g\nu^t)-f(g)}{t^{1/\rho}}$.

For $q \in (1,2]$, the q-uniform smoothness constant of X is defined by

$$S_q(X) := \inf\left\{K > 0 : \forall x,y \in E \ \left(\|x\|_X^q + K^q\|y\|_X^q\right)^{1/q} \geq \left(\frac{\|x+y\|_X^q + \|x-y\|_X^q}{2}\right)^{1/q}\right\}.$$

Then we have

$$K_q(L^p(G;X)) = S_{q/(q-1)}(L^{p/(p-1)}(G;X^*)) \leq \max\left\{\left(\frac{p}{p-1}\right)^{(q-1)/q}, S_{q/(q-1)}(X^*)\right\} \asymp \max\left\{(p-1)^{\frac{1}{q}-1}, K_q(X)\right\}$$

Where the equality and $\asymp$ follow from [BCL94, Lemma 5] and the inequality follows from equation (4.4) of [Nao14]. Therefore we have

$$\left(\int_0^\infty \|D_\nu^t F_M\|^q_{L^q(G;L^p(G;X))} \frac{dt}{t}\right)^{1/q} \lesssim \frac{\rho}{\rho-1}\max\{(p-1)^{1/q-1}, K_q(E)\}\|\nabla F_M\|_{L^q(G;\ell_q^k(L^p(G;X)))}. \tag{31}$$

As

$$\nabla F_M(h)(g) = f(gh)\nabla\beta_M(h) + \beta_M(h)\nabla f(gh)$$

from left-invariance of ∇, we have (with the normalization $\mu(B_1) = 1$)

$$\|\nabla F_M(h)\|_{\ell_2^k(L^p(G;X))} \lesssim \|f\|_{L^p(G;X)} \mathbb{1}_{B_{M+1}\setminus B_M}(h) + \|\nabla f\|_{L^p(G;X)} \mathbb{1}_{B_{M+1}}(h), \quad h \in G,$$

and thus (with the normalization $\mu(B_1) = 1$)

$$\|\nabla F_M\|_{L^q(G;\ell_2^k(L^p(G;X)))} \lesssim \mu(B_{M+1}\setminus B_M)^{1/q}\|f\|_{L^p(G;X)} + \mu(B_M)^{1/q}\|\nabla f\|_{L^p(G;\ell_2^k(X))}. \tag{32}$$

If $0 < t < M^s$ and $h \in B_{M-t^{1/s}}$, then $h\nu^t \in B_M$ so $\beta_M(h) = \beta_M(h\nu^t) = 1$. Thus, for any $h \in B_{M-t^{1/s}}$, $\|D_\nu^t F_M(h)\|_{L^p(G;X)} = \|D_\nu^t f\|_{L^p(G;X)}$. We have

$$\|D_\nu^t F_M\|_{L^q(G;L^p(G;X))} \gtrsim \mu(B_{M-t^{1/s}})^{1/q}\|D_\nu^t f\|_{L^p(G;X)},$$

and for $0 < T < M$,

$$\left(\int_0^{T^s} \|D_\nu^t F_M\|^q_{L^q(G;L^p(G;X))} \frac{dt}{t}\right)^{1/q} \gtrsim \mu(B_{M-T})^{1/q}\left(\int_0^{T^s} \|D_\nu^t f\|^q_{L^p(G;X)} \frac{dt}{t}\right)^{1/q}. \tag{33}$$

Combining (31), (32), and (33), we are left with

$$\mu(B_{M-T})^{1/q}\left(\int_0^{T^s} \|D_\nu^t f\|^q_{L^p(G;X)} \frac{dt}{t}\right)^{1/q}$$

$$\lesssim \frac{\rho}{\rho-1}\max\{(p-1)^{1/q-1}, K_q(X)\}\left(\mu(B_{M+1}\setminus B_M)^{1/q}\|f\|_{L^p(G;X)} + \mu(B_M)^{1/q}\|\nabla f\|_{L^p(G;\ell_2^k(X))}\right)$$

for all $0 < T < M$. Taking $M \to \infty$ and then $T \to \infty$, we obtain the stated inequality. This completes the proof of Theorem 32.

3. Proof of Theorems 1, 2, 10, 11, 14, and 15 and Corollaries 3 and 4: distortion bounds

In this section we prove Theorems 1 and 2, Corollaries 3 and 4, and Theorems 10, 11, 14 and 15. Theorems 1, 10, 11, 14 and 15 will follow from Theorem 5, Theorem 2 will follow from Theorem 6, and Corollaries 3 and 4 will follow from Theorems 1 and 2, respectively.

First we show the proof of Theorem 1 from Theorem 5.

Proof of Theorem 1 from Theorem 5. Suppose $f : B_r \to X$ satisfies $d_G(x, y) \le \|f(x) - f(y)\|_X \le D d_G(x, y)$ for $x, y \in B_r$. Translating and multiplying f by a cutoff function we may construct $F : G \to X$ which is $3D$-Lipschitz, agrees with f on $B_{r/2}$, and is supported on B_r. Supposing Theorem 5 and applying (2), we have, using $\|f(hv^t) - f(h)\|_X \ge d_G(v^t, e_G) \gtrsim_G t^{1/\rho}$ for $t \ge 1$,

$$\mu(B_{r/4})^{1/q} \log(r^s/4^s)^{1/q} \lesssim_G \left(\int_1^{r^s/4^s} \int_{B_{r/4}} \left(\frac{\|f(hv^t) - f(h)\|_X}{t^{1/s}} \right)^q d\mu(h) \frac{dt}{t} \right)^{1/q}$$

$$\lesssim K_q(X) \left(\int_{B_r} \|\nabla F\|_{\ell_2^k(X)}^q d\mu \right)^{1/q} \lesssim_G K_q(X) \mu(B_r)^{1/q} D.$$

But by [Kar94, Corollary 4.11], we have $\mu(B_r) \asymp_G r^{n_h}$ for $r = \Omega(1)$ for a fixed positive integer n_h. We therefore obtain Theorem 1. $\qquad\square$

The proof of Theorem 2 from Theorem 6 is similar.

Proof of Theorem 2 from Theorem 6. Let c be as in Theorem 6. It is enough to show that $c_X(B_{(c+1)n}) \gtrsim_\Gamma \frac{(\log n)^{1/q}}{K_q(X)}$. Suppose that $f : \Gamma \to X$ satisfies $d_W(x, y) \le \|f(x) - f(y)\|_X \le D d_W(x, y)$ for all $x, y \in B_{(c+1)n}^\Gamma$. Recall that

$$d_W(v_\Gamma^k, e_\Gamma) \asymp_\Gamma k^{1/s}, \quad k \in \mathbb{N},$$

and also that $|B_m| \asymp_\Gamma m^{n_h}$ for every $m \in \mathbb{N}$ (see [BLD12, Theorem 1.1]). Thus, Theorem 6 applied to f yields the following estimate.

$$n^{n_h/q} (\log n)^{1/q} \lesssim_\Gamma \left(\sum_{k=1}^{n^s} n^{n_h} \frac{k^{q/s}}{k^{1+q/s}} \right)^{1/q} \lesssim_\Gamma K_q(X) n^{n_h/q} D, \qquad (34)$$

which gives Theorem 2. $\qquad\square$

Proof of Corollary 4. The lower bound of Corollary 4 follows from Theorem 2. The upper bound follows from a version [LMN05, Theorem 5.1] of the Assouad embedding theorem [Ass83]:

$$c_p(\Gamma, d_\Gamma^{1-\varepsilon}) \lesssim_{\Gamma, p} 1/\varepsilon^{1/\max\{p, 2\}}, \quad 0 < \varepsilon < 1,$$

Also, $(B_n^\Gamma, d_W^{1-1/2\log n})$ is $O(1)$-bilipschitz equivalent to (B_n^Γ, d_W), so we have the upper bound. $\qquad\square$

Proof of Corollary 3. The lower bound of Corollary 3 follows from Theorem 2. To show the upper bound, we note that the proof of the aforementioned version [LMN05, Theorem 5.1] of the Assouad embedding theorem [Ass83], with $\varepsilon = \frac{1}{2\log R}$, produces a mapping $\psi : G \to L^p$ such that ψ is $O_G\left((\log R)^{1/\max\{p, 2\}}\right)$-Lipschitz and satisfies

$$\|\psi(g) - \psi(h)\|_{L^p} \gtrsim_G d_G(g, h)^{1-1/2\log R}, \quad g, h \in G \text{ such that } d_G(g, h) = \Omega(1).$$

It is thus enough to construct a mapping $\phi : G \to \mathbb{R}^d$ such that ϕ is $O(1)$-Lipschitz and

$$|\phi(g) - \phi(h)| \gtrsim_G d_G(g, h), \quad g, h \in G \text{ such that } d_G(g, h) \le 1,$$

for then $\psi \oplus \phi : G \to L^p \oplus \mathbb{R}^d = L^p$ gives the desired mapping.

Since G is Riemannian and $\exp : \mathfrak{g} \to G$ is a diffeomorphism, $\exp$ is $O(1)$-bilipschitz on balls of radius $O(1)$. The idea is to 'glue' these mappings together to produce the mapping ϕ.

Let N be a $(1,1)$-net of G. Since G is doubling, there is a constant K such that $N = \cup_{i=1}^{K} N_i$, where N_i are 8-packings of G. For each $i = 1, \cdots, K$ extend the mapping

$$\cup_{g \in N_i} g B_2 \to \mathfrak{g}, \quad h \in g B_2 \mapsto \exp^{-1}(g^{-1} h)$$

to an $O(\log K)$-Lipschitz mapping $\phi_i : G \to \mathfrak{g}$ by [LN05], and set $\phi = \oplus_{i=1}^{K} \phi_i$. Then ϕ has the desired properties. $\qquad\square$

Proof of Theorem 10. Let N_{r_1, r_2} be an r_2-covering of B_{r_1} and let $f : N_{r_1, r_2} \to X$ be a mapping with

$$d_G(x, y) \leq \|f(x) - f(y)\| \leq D d_G(x, y), \quad x, y \in N_{r_1, r_2}.$$

By [LN05] there is an extension $F : G \to X$ of f that is $CD \log K$-Lipschitz for a universal constant $C > 0$. Let ξ be a smooth bump function that equals 1 on B_{r_1}, is supported on B_{2r_1} and is $O(1)$-Lipschitz. Then $\phi := \xi F$ is $O(D \log K)$-Lipschitz, equals F on B_{r_1} and is supported on B_{2r_1}. To apply Theorem 5 with $p = q$, we note that

$$\left(\int_G \|\nabla \phi(h)\|_{\ell_2^k(X)}^q \, d\mu(h) \right)^{1/q} \lesssim \sqrt{k} D(\log K) \mu(B_{2r_1})^{1/q}.$$

On the other hand, for $x \in B_{r_1/2}$ and $t \in ((\frac{(4 + 4CD \log K) r_2}{C_1})^s, (r_1/2)^s]$ (we may assume $(\frac{(4 + 4CD \log K) r_2}{C_1})^s \leq (r_1/2)^s$ since otherwise $D > \frac{C_1(r_1/r_2)}{16C \log K}$ and the proof is complete), we have $\phi(x) = F(x)$ and $\phi(xv^t) = F(xv^t)$. Choosing $n_1, n_2 \in N$ such that $d_G(x, n_1) < r_2$, $d_G(xv^t, n_2) < r_2$, we have

$$\|F(xv^t) - F(x)\| \geq \|f(n_1) - f(n_2)\| - 2CDr_2 \log K \geq d(n_1, n_2) - 2CDr_2 \log K \geq (C_1 t^{1/s} - 2r_2) - 2CDr_2 \log K > \frac{C_1}{2} t^{1/s}.$$

Thus

$$\left(\int_0^\infty \int_G \left(\frac{\|\phi(hv^t) - \phi(h)\|_X}{t^{1/s}} \right)^q d\mu(h) \frac{dt}{t} \right)^{1/q} \geq \left(\int_{(\frac{(4 + 4CD \log K) r_2}{C_1})^s}^{(r_1/2)^s} \int_{B_{r_1/2}} \frac{C_1^q}{2^q} d\mu(h) \frac{dt}{t} \right)^{1/q}$$

$$\geq \frac{C_1 s^{1/q}}{2} \left(\log(r_1/2) - \log(\frac{(4 + 4CD \log K) r_2}{C_1}) \right)^{1/q} \mu(B_{r_1/2})^{1/q}.$$

Thus by Theorem 5 with $p = q$,

$$\frac{C_1 s^{1/q}}{2} \left(\log(r_1/2) - \log(\frac{(4 + 4CD \log K) r_2}{C_1}) \right)^{1/q} \mu(B_{r_1/2})^{1/q} \lesssim K_q(X) \sqrt{k} D(\log K) \mu(B_{2r_1})^{1/q}.$$

But we may also assume $\log(r_1/2) - \log(\frac{(4 + 4CD \log K) r_2}{C_1}) \geq \frac{1}{2} \log(r_1/r_2)$, for otherwise $D > \frac{C_1 \sqrt{r_1/r_2}}{16C \log K}$ and the proof is complete. Thus

$$C_1 s^{1/q} (\log(r_1/r_2))^{1/q} \lesssim K_q(X) \sqrt{k} D K^{2/q} \log K,$$

from which it follows that

$$D \gtrsim \left(\frac{C_1 s^{1/q}}{\sqrt{k} K^{2/q} \log K} \right) \frac{(\log(r_1/r_2))^{1/q}}{K_q(X)}.$$

$\qquad\square$

Proof of Theorem 11. Let $N := N'_{c_1, c_2}$ be a (c_1, c_2)-net of G and let $f : N \to X$ be a mapping with

$$d_G(x, y)^{1-\varepsilon} \leq \|f(x) - f(y)\| \leq D d_G(x, y)^{1-\varepsilon}, \quad x, y \in N.$$

Then f is $c_2^{-\varepsilon} D$-Lipschitz, so by [LN05] there is an extension $F : G \to X$ of f that is $C c_2^{-\varepsilon} D \log K$-Lipschitz for a universal constant $C > 0$. Fix $R > 0$ sufficiently large, let ξ be a smooth bump function that equals 1 on B_R, is supported on B_{2R} and is $O(1)$-Lipschitz. Then $\phi := \xi F$ is $O(c_2^{-\varepsilon} D \log K)$-Lipschitz, equals F on B_R and is supported on B_{2R}. To apply Theorem 5 with $p = q$, we note that

$$\left(\int_G \|\nabla \phi(h)\|_{\ell_2^k(X)}^q \, d\mu(h) \right)^{1/q} \lesssim \sqrt{k} c_2^{-\varepsilon} D(\log K) \mu(B_{2R})^{1/q}.$$

On the other hand, for $x \in B_{R/2}$ and $t \in ((8c_1 c_2^{-\varepsilon} D \log K)^{s/(1-\varepsilon)}/C_1^s, (R/2)^s]$ (this interval is nonempty since we took $R > 0$ sufficiently large) $(8c_1 c_2^{-\varepsilon} D \log K)^{s/(1-\varepsilon)}/C_1^s \geq 2^s c_1^s C_1^s$, we have $\phi(x) = F(x)$ and $\phi(xv^t) = F(xv^t)$. Choosing $n_1, n_2 \in N$ such that $d_G(x, n_1) < c_1$, $d_G(xv^t, n_2) < c_1$, we have

$$\|F(xv^t) - F(x)\| \geq \|f(n_1) - f(n_2)\| - 2c_1 c_2^{-\varepsilon} CD \log K \geq d(n_1, n_2)^{1-\varepsilon} - 2c_1 c_2^{-\varepsilon} CD \log K$$

$$\geq (C_1 t^{1/s} - 2c_1)^{1-\varepsilon} - 2c_1 c_2^{-\varepsilon} CD \log K > \frac{C_1^{1-\varepsilon}}{4} t^{(1-\varepsilon)/s}.$$

Thus

$$\left(\int_0^\infty \int_G \left(\frac{\|\phi(hv^t) - \phi(h)\|_X}{t^{1/s}} \right)^q d\mu(h) \frac{dt}{t} \right)^{1/q} \geq \left(\int_{(8c_1 c_2^{-\varepsilon} D \log K)^{s/(1-\varepsilon)}/C_1^s}^{(R/2)^s} \int_{B_{R/2}} \frac{C_1^{q(1-\varepsilon)}}{4^q} d\mu(h) \frac{dt}{t^{1+\varepsilon q/s}} \right)^{1/q}$$

$$\geq \frac{C_1^{1-\varepsilon}}{4} \left(\frac{s}{\varepsilon q} \left[(8c_1 c_2^{-\varepsilon} D \log K)^{-\varepsilon q/(1-\varepsilon)}/C_1^{-\varepsilon q} - (R/2)^{-\varepsilon q} \right] \right)^{1/q} \mu(B_{R/2})^{1/q}.$$

Thus by Theorem 5 with $p = q$,

$$\frac{C_1^{1-\varepsilon}}{4} \left(\frac{s}{\varepsilon q} \left[(8c_1 c_2^{-\varepsilon} D \log K)^{-\varepsilon q/(1-\varepsilon)}/C_1^{-\varepsilon q} - (R/2)^{-\varepsilon q} \right] \right)^{1/q} \mu(B_{R/2})^{1/q} \lesssim K_q(X) \sqrt{k} c_2^{-\varepsilon} D(\log K) \mu(B_{2R})^{1/q}.$$

Since $R > 0$ was arbitrarily large,

$$\frac{C_1^{1-\varepsilon} s^{1/q}}{\varepsilon^{1/q}} (8c_1 c_2^{-\varepsilon} D \log K)^{-\varepsilon/(1-\varepsilon)}/C_1^{-\varepsilon} \lesssim K_q(X) \sqrt{k} c_2^{-\varepsilon} DK^{2/q} \log K.$$

It follows that

$$D \gtrsim \left(\frac{C_1^{1-\varepsilon} (c_2/c_1)^\varepsilon s^{(1-\varepsilon)/q}}{k^{(1-\varepsilon)/2} K^{2(1-\varepsilon)/q} \log K} \right) \frac{1}{K_q(X)^{1-\varepsilon} \varepsilon^{1/q}}.$$

$\square$

The proof of Theorem 14 is similar to that of Theorem 10 and the proof of Theorem 15 is similar to that of Theorem 11 and we thus omit them. To see the dependence on the constants, we only need to distinguish when K is used as the metric doubling constant or the measure doubling constant.

4. Proof of Theorem 6: discretization

In this section, we prove Theorem 6 by discretizing the inequality of Theorem 5.

Finitely generated groups of polynomial growth have nilpotent subgroups of finite index [Gro81], so let $\Gamma' \leq \Gamma$ be a finite index nilpotent subgroup. We first can see that it is enough to prove Theorem 6 for Γ', i.e., we may assume Γ is a not virtually abelian finitely generated nilpotent group. Indeed, the distance requirement $d_W(v_\Gamma^n, e_\Gamma) \asymp_G n^{1/s}$ carries over from Γ' to Γ because a finitely generated word metric on a group and its finite index subgroup are known to be quasi-isometric. Also, once we know (3) for Γ' and $S' \subset \Gamma'$, it automatically follows that it holds for Γ and $S \supset S'$. (Observe that changing the generating set only affects the right-hand side of (3) up to universal constant factors.)

Let $T := \{x \in \Gamma' : \exists n \in \mathbb{Z}_{>0} \text{ such that } x^n = e_{\Gamma'}\}$ be the torsion subgroup of Γ'. It is well-known that T is a finite normal subgroup of Γ', so that $\Gamma'' := \Gamma'/T$ is a torsion-free finitely generated nilpotent group. We now see that it is enough to prove Theorem 6 for Γ'', i.e., we may assume Γ is a torsion-free finitely generated nonabelian nilpotent group. Indeed, again the distance requirement $d_W(v_\Gamma^n, e_\Gamma) \asymp_G n^{1/s}$ carries over from Γ'' to Γ' because a finitely generated word metric on a group and its quotient by a finite normal subgroup are known to be quasi-isometric. Also, suppose we know (3) for Γ'' and $S'' \subset \Gamma''$. Let S' be the union of T and a set of representatives for S'' in Γ', and let $v_{\Gamma'}$ be a representative for $v_{\Gamma''}$ in Γ'. Let $\pi : \Gamma' \to \Gamma''$ denote the quotient map, and let $C > 0$ be such that $\pi(B_k^{\Gamma'}(x)) \subset B_{Ck}^{\Gamma''}(\pi(x))$ for $k \in \mathbb{Z}_{>0}$ and

$x \in \Gamma'$ and $\pi^{-1}(B_k^{\Gamma''}(y)) \subset B_{Ck}^{\Gamma'}(x)$ for $k \in \mathbb{Z}_{>0}$, $y \in \Gamma''$ and $x \in \pi^{-1}(y)$. Given $f : \Gamma' \to X$, let $[f] : \Gamma'' \to X$ be given as

$$[f](y) := \frac{1}{|T|} \sum_{x \in \pi^{-1}(y)} f(x), \quad y \in \Gamma''.$$

Then

$$\left(\sum_{k=1}^{n^s} \frac{1}{k^{1+\max\{p,q\}/s}} \left(\sum_{x \in B_n^{\Gamma'}} \| f(x v_{\Gamma'}^k) - f(x) \|_X^p \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}}$$

$$\leq \left(\sum_{k=1}^{n^s} \frac{1}{k^{1+\max\{p,q\}/s}} \left(\sum_{x \in B_n^{\Gamma'}} \left\| \left(f(x v_{\Gamma'}^k) - [f]([x v_{\Gamma'}^k]) \right) - \left(f(x) - [f]([x]) \right) \right\|_X^p \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}} \tag{35}$$

$$+ \left(\sum_{k=1}^{n^s} \frac{1}{k^{1+\max\{p,q\}/s}} \left(\sum_{x \in B_n^{\Gamma'}} \left\| [f]([x v_{\Gamma'}^k]) - [f]([x]) \right\|_X^p \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}}.$$

The second term is at most

$$\left(\sum_{k=1}^{n^s} \frac{1}{k^{1+\max\{p,q\}/s}} \left(|T| \sum_{y \in B_{Cn}^{\Gamma''}} \| [f](y v_{\Gamma''}^k) - [f](y) \|_X^p \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}}$$

$$\lesssim_\Gamma \max\left\{ (p-1)^{1/q-1}, K_q(X) \right\} \left(\sum_{y \in B_{cCn}^{\Gamma'}} \sum_{a \in S''} \| [f](ya) - [f](y) \|_X^p \right)^{1/p}$$

$$= \max\left\{ (p-1)^{1/q-1}, K_q(X) \right\} \left(\sum_{y \in B_{cCn}^{\Gamma'}} \sum_{a \in S''} \left\| \frac{1}{|T|} \sum_{x \in \pi^{-1}(y)} (f(xa') - f(x)) \right\|_X^p \right)^{1/p}$$

$$\leq \max\left\{ (p-1)^{1/q-1}, K_q(X) \right\} \left(\sum_{y \in B_{cCn}^{\Gamma'}} \sum_{a \in S''} \frac{1}{|T|} \sum_{x \in \pi^{-1}(y)} \| f(xa') - f(x) \|_X^p \right)^{1/p}$$

$$\leq \max\left\{ (p-1)^{1/q-1}, K_q(X) \right\} \left(\sum_{x \in B_{cC^2 n}^{\Gamma'}} \sum_{a \in S''} \frac{1}{|T|} \| f(xa') - f(x) \|_X^p \right)^{1/p},$$

and is thus bounded by the right-hand side of (3), where for $a \in S''$, $a' \in S'$ denotes an element of $\pi^{-1}(a)$.

29

It remains to show that the first term of the right-hand side of (35) is bounded by the right-hand side of (3). Indeed, it is equal to

$$
\left(\sum_{k=1}^{n^s} \frac{1}{k^{1+\max\{p,q\}/s}} \left(\sum_{x\in B_n^{\Gamma'}} \left\| \frac{1}{|T|} \sum_{b\in T} \left(f(xv_{\Gamma'}^k) - f(xv_{\Gamma'}^k b) - f(x) + f(xb) \right) \right\|_X^p \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}}
$$

$$
\leq \left(\sum_{k=1}^{n^s} \frac{1}{k^{1+\max\{p,q\}/s}} \left(\frac{2^{p-1}}{|T|} \sum_{x\in B_n^{\Gamma'}} \sum_{b\in T} \left(\left\| f(xv_{\Gamma'}^k) - f(xv_{\Gamma'}^k b) \right\|_X^p + \left\| f(x) - f(xb) \right\|_X^p \right) \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}}
$$

$$
\leq \left(\sum_{k=1}^{n^s} \frac{1}{k^{1+\max\{p,q\}/s}} \left(\frac{2^p}{|T|} \sum_{x\in B_{c'n}^{\Gamma'}} \sum_{b\in T} \left\| f(x) - f(xb) \right\|_X^p \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}}
$$

$$
\lesssim_\Gamma \left(\sum_{x\in B_{c'n}^{\Gamma'}} \sum_{b\in T} \left\| f(x) - f(xb) \right\|_X^p \right)^{1/p}
$$

where $c' > 0$ is such that $B_n^{\Gamma'} v_{\Gamma'}^k \subset B_{c'n}^{\Gamma'}$ for $n\in\mathbb{N}$ and $k = 1,\cdots,n^s$.

This completes our reduction of Theorem 6 to proving it for Γ being a torsion-free finitely generated nonabelian nilpotent group. By the Malcev embedding theorem [Mal49], there exists a simply connected nilpotent Lie group G, called the Malcev completion of Γ, such that Γ embeds as a cocompact subgroup of G. Because Γ is nonabelian, G is nonabelian, so Theorem 5 applies to G.

There exist choices of s and v_Γ as stated. It is enough to take $v\in Z(\Gamma)\setminus\{e_\Gamma\}$ with $d_W(v^n,e_\Gamma)\asymp_\Gamma n^{1/s}$ for some integer $s\geq 2$. For example, let s denote the nilpotency step of G, i.e., the largest integer so that $\underbrace{[[G,G]\cdots,G]}_{s \text{ times}}$ is a nontrivial subgroup. If we choose $v_\Gamma\in\underbrace{[[\Gamma,\Gamma]\cdots,\Gamma]}_{s \text{ times}}\setminus\{e_\Gamma\}\subset Z(\Gamma')\setminus\{e_\Gamma\}$, then

$$
d_W(v_\Gamma^k, e_\Gamma) \asymp_{\Gamma,v_\Gamma} k^{1/s}, \quad k\in\mathbb{N}
$$

because by [BLD12, Theorem 1.3], $|d_W(v_\Gamma^k,e_\Gamma) - d_G(v_\Gamma^k,e_G)| = O_{\Gamma,v_\Gamma}(d_G(v_\Gamma^k,e_G)^{1-\alpha})$ for some fixed $\alpha = \alpha_G$, and $d_G(v^k,e_G)\asymp_G k^{1/s}$ by [BLD12, Proposition 2.13].

Also, $|B_m^\Gamma|\asymp_\Gamma m^{n_h}$ for every $m\in\mathbb{N}$ [BLD12, Theorem 1.1], where n_h is given by the Bass–Guivarc'h formula

$$
n_h := \sum_{k=1}^s k \dim\left(\underbrace{[[G,G]\cdots,G]}_{k \text{ times}} / \underbrace{[[G,G]\cdots,G]}_{k+1 \text{ times}} \right).
$$

The rest of this section follows closely the argument of Section 3 of [LN14].

Before we begin the proof of Theorem 6, we prove two metric-space valued local Poincaré inequalities on Γ for preparation. Lemma 39 is an extension of the local Poincaré inequality of Kleiner [Kle10, Thm. 2.2] and was essentially proven in [LN14, Lemma 3.2] for all finitely generated groups. We repeat the proof for completeness.

Lemma 39 (Analogue of [LN14, Lemma 3.2]). *Fix $p\in[1,\infty)$ and $n\in\mathbb{N}$. Let (M,d_M) be a metric space. For every $f:\Gamma\to M$,*

$$
\sum_{x,y\in B_n^\Gamma} d_M(f(x),f(y))^p \leq (2n)^p |B_{2n}^\Gamma| \sum_{x\in B_{3n}^\Gamma} \sum_{a\in S} d_M(f(xa),f(x))^p.
$$

Proof. For every $z\in B_{2n}^\Gamma$ choose $s_1(z),\ldots,s_{2n}(z)\in S\cup\{e_\Gamma\}$ such that $z = s_1(z)\cdots s_{2n}(z)$. For $i\in\{1,\ldots,2n\}$ write $w_i(z) = s_1(z)\cdots s_i(z)$ and set $w_0(z) = e_\Gamma$. By the triangle inequality and Hölder's inequality, for every

$x, y \in B_n^\Gamma$ we have

$$d_M(f(x), f(y))^p \le (2n)^{p-1} \sum_{i=0}^{2n-1} d_M\left(f\left(x w_i\left(x^{-1} y\right)\right), f\left(x w_i\left(x^{-1} y\right) s_{i+1}\left(x^{-1} y\right)\right)\right)^p.$$

Consequently,

$$\sum_{x,y \in B_n^\Gamma} d_M(f(x), f(y))^p \le (2n)^{p-1} \sum_{z \in B_{2n}^\Gamma} \sum_{i=1}^{2n-1} \sum_{x \in B_n^\Gamma} d_M\left(f\left(x w_i(z)\right), f\left(x w_i(z) s_{i+1}(z)\right)\right)^p$$

$$= (2n)^{p-1} \sum_{z \in B_{2n}^\Gamma} \sum_{i=0}^{2n-1} \sum_{g \in B_n^\Gamma w_i(z)} d_M\left(f(g), f(g s_{i+1}(z))\right)^p$$

$$\le (2n)^{p-1} \cdot |B_{2n}^\Gamma| \cdot 2n \sum_{x \in B_{3n}^\Gamma} \sum_{a \in S} d_M(f(xa), f(x))^p.$$

$\square$

The following lemma is an analogue of [LN14, Lemma 3.4], which was again proven essentially for all finitely generated groups. A mapping $f : \Gamma \to M$ is said to be finitely supported if there exists $m_0 \in M$ such that $|f^{-1}(M \setminus \{m_0\})| < \infty$.

Lemma 40 (Analogue of [LN14, Lemma 3.4]). *Fix $p \in [1, \infty)$ and $n \in \mathbb{N}$. Let (M, d_M) be a metric space. For every finitely supported $f : \Gamma \to M$,*

$$\sum_{x \in \Gamma} \sum_{z \in B_n^\Gamma} d_M(f(xz), f(x))^p \le n^p |B_n^\Gamma| \sum_{x \in \Gamma} \sum_{a \in S} d_M(f(xa), f(x))^p.$$

Proof. Let s_i and w_i be as in the proof of Lemma 39. By the triangle inequality and Hölder's inequality,

$$\sum_{x \in \Gamma} \sum_{z \in B_n^\Gamma} d_M(f(xz), f(x))^p \le n^{p-1} \sum_{x \in \Gamma} \sum_{z \in B_n^\Gamma} \sum_{i=0}^{n-1} d_M\left(f\left(x w_i(z)\right), f\left(x w_i(z) s_{i+1}(z)\right)\right)^p.$$

$$\le n^p |B_n^\Gamma| \sum_{x \in \Gamma} \sum_{a \in S} d_M(f(xa), f(x))^p.$$

$\square$

With Lemmas 39 and 40, we begin by discretizing the inequality of Theorem 5, as in the following theorem. After this, we will localize the statement into that of Theorem 6, which states a local version of Theorem 41 on balls in Γ.

The following theorem corresponds to Theorem 3.3 of [LN14], whose proof is in turn a variant of the proof of Claim 7.3 in [ANT13]. We may state it in terms of discrete groups of polynomial growth.

Theorem 41 (Analogue of [LN14, Theorem 3.3]). *Let Γ be a not virtually abelian finitely generated group of polynomial growth. There exist $v_\Gamma \in \Gamma$, $s \in \mathbb{N}$ with $s \ge 2$, and $c = c(\Gamma) \in \mathbb{N}$ such that the following is true. First, $d_W(v_\Gamma^n, e_\Gamma) \asymp_\Gamma n^{1/s}$ for $n \in \mathbb{N}$. Second, let $p \in (1, \infty)$ and $q \in [2, \infty)$. Suppose that $(X, \|\cdot\|_X)$ is a Banach space satisfying $K_q(X) < \infty$. Then for every finitely supported $f : \Gamma \to X$ we have*

$$\left(\sum_{k=1}^{n^s} \frac{1}{k^{1+\max\{p,q\}/s}} \left(\sum_{x \in \Gamma} \|f(x v_\Gamma^k) - f(x)\|_X^p \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}}$$

$$\lesssim_{\Gamma, v_\Gamma} \max\left\{(p-1)^{1/q-1}, K_q(X)\right\} \left(\sum_{x \in \Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^p \right)^{1/p}.$$

Proof. With the same argument as in the beginning of this section, we may assume Γ is a torsion-free nonabelian finitely generated nilpotent group, and take v_Γ and s as before.

The idea of the proof is that given a finitely supported function $f : \Gamma \to X$, we extend it to a global function $F : G \to X$ via a partition of unity, and then Theorem 5 for F will give Theorem 41 for f.

Let G be the Malcev completion of Γ. Since Γ is a co-compact lattice of G, there exists a compactly supported smooth function $\chi : G \to [0, 1]$ with

$$\forall h \in G, \qquad \sum_{x \in \Gamma} \chi_x(h) = 1, \tag{36}$$

where for each $x \in \Gamma$, $\chi_x : G \to X$ is given by $\chi_x(h) = \chi(x^{-1}h)$, $h \in G$. Let $A = \operatorname{supp} \chi$; we may assume $A^{-1} = A$. Note that $\operatorname{supp} \chi_x = xA$ and

$$\bigcup_{x \in \Gamma} xA = G.$$

As A is compact, we may fix $m \in \mathbb{N}$ for which $A^2 \cap \Gamma \subseteq B_m^\Gamma$.

Let $f : \Gamma \to X$ be finitely supported. Define $F : G \to X$ by

$$F(h) := \sum_{x \in \Gamma} \chi_x(h) f(x). \tag{37}$$

For a fixed $x \in \Gamma$ and $h \in xA$, we have

$$\sum_{y \in \Gamma} \nabla \chi_y(h) = 0,$$

with $\nabla \chi_y(h) \neq 0$ implying $y^{-1}h \in \operatorname{supp} \nabla \chi \subset A$, which implies $y \in hA^{-1} \subset xA^2$, hence $y \in xB_m^\Gamma$. We thus have the bound

$$\|\nabla F(h)\|_{\ell_2^k(X)} = \left\| \sum_{y \in xB_m^\Gamma} \nabla \chi_y(h)(f(y) - f(x)) \right\|_{\ell_2^k(X)}$$

$$\leq \|\nabla \chi\|_{L^\infty(G; \ell_2^k)} \sum_{z \in B_m^\Gamma} \|f(xz) - f(x)\|_X$$

$$\lesssim_\Gamma |B_m^\Gamma|^{1-1/p} \left(\sum_{z \in B_m^\Gamma} \|f(yz) - f(y)\|_X^p \right)^{1/p}.$$

By integrating over yA and summing over $y \in \Gamma$,

$$\left(\int_G \|\nabla F(h)\|_{\ell_2^k(X)}^p \, dh \right)^{1/p} \leq \left(\sum_{y \in \Gamma} \int_{yA} \|\nabla F(h)\|_{\ell_2^k(X)}^p \, dh \right)^{1/p} \lesssim_\Gamma \left(\sum_{y \in \Gamma} \sum_{z \in B_m^\Gamma} \|f(yz) - f(y)\|_X^p \right)^{1/p}$$

$$\overset{\text{Lemma 40}}{\lesssim_\Gamma} \left(\sum_{x \in \Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^p \right)^{1/p}. \tag{38}$$

Since Γ induces a covering space action on G, we may find a bounded open neighborhood $U \subseteq G$ of e_G such that $U \cap (xU) = \emptyset$ for all $x \in \Gamma \setminus \{e_\Gamma\}$. As $v_\Gamma^{[0,1]} := \{v_\Gamma^\tau : \tau \in [0,1]\}$ and U are bounded in G, we can choose $r \in \mathbb{N}$ such that $(Uv_\Gamma^{[0,1]}A) \cap \Gamma \subseteq B_r^\Gamma$. For $x \in \Gamma$, $h \in xU$, $k \in \mathbb{N}$, and $t \in [k, k+1]$, we have by (36) and (37)

$$F(hv_\Gamma^t) - f(xv_\Gamma^k) = \sum_{w \in \Gamma} \chi_w(hv_\Gamma^t)(f(w) - f(xv_\Gamma^k)).$$

Every $w \in \Gamma$ that satisfies $\chi_w(hv_\Gamma^t) \neq 0$ should have $hv_\Gamma^t \in wA$, and since $A = A^{-1}$ and $v_\Gamma \in Z(G)$, we have $w \in hv_\Gamma^t A \subseteq xv_\Gamma^k Uv_\Gamma^{[0,1]}A$, so that $w \in xv_\Gamma^k B_r^\Gamma$. Therefore

$$\|F(hv_\Gamma^t) - f(xv_\Gamma^k)\|_X \lesssim_\Gamma \left(\sum_{z \in B_r^\Gamma} \|f(xv_\Gamma^k z) - f(xv_\Gamma^k)\|_X^p \right)^{1/p}.$$

32

The case $k = t = 0$ gives

$$\|F(h) - f(x)\|_X \lesssim_\Gamma \left(\sum_{z \in B_r^\Gamma} \|f(xz) - f(x)\|_X^p \right)^{1/p}.$$

Therefore by the triangle inequality,

$$\|f(xv_\Gamma^k) - f(x)\|_X \lesssim_\Gamma \|F(hv_\Gamma^t) - F(h)\|_X + \left(\sum_{z \in B_r^\Gamma} \left(\|f(xv_\Gamma^k z) - f(xv_\Gamma^k)\|_X^p + \|f(xz) - f(x)\|_X^p \right) \right)^{1/p}.$$

Integration over $h \in xU$ gives

$$\|f(xv_\Gamma^k) - f(x)\|_X \lesssim_\Gamma \left(\int_{xU} \|F(hv_\Gamma^t) - F(h)\|_X^p dh \right)^{1/p} + \left(\sum_{z \in B_r^\Gamma} \left(\|f(xv_\Gamma^k z) - f(xv_\Gamma^k)\|_X^p + \|f(xz) - f(x)\|_X^p \right) \right)^{1/p}.$$

Summing over $x \in \Gamma$ and using that the sets $\{xU\}_{x \in \Gamma}$ are pairwise disjoint,

$$\left(\sum_{x \in \Gamma} \|f(xv_\Gamma^k) - f(x)\|_X^p \right)^{1/p} \lesssim_\Gamma \left(\int_G \|F(hv_\Gamma^t) - F(h)\|_X^p dh \right)^{1/p}$$

$$+ \left(\sum_{x \in \Gamma} \sum_{z \in B_r^\Gamma} \left(\|f(xv_\Gamma^k z) - f(xv_\Gamma^k)\|_X^p + \|f(xz) - f(x)\|_X^p \right) \right)^{1/p}$$

$$\overset{\text{Lemma } 40}{\lesssim_\Gamma} \left(\int_G \|F(hv_\Gamma^t) - F(h)\|_X^p dh \right)^{1/p} + \left(\sum_{x \in \Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^p \right)^{1/p}.$$

Integrating over $t \in [k, k+1]$ yields

$$\frac{1}{k^{1+\max\{p,q\}/s}} \left(\sum_{x \in \Gamma} \|f(xv_\Gamma^k) - f(x)\|_X^p \right)^{\max\{p,q\}/p}$$

$$\leq C^{\max\{p,q\}} \int_k^{k+1} \left(\int_G \|F(hv_\Gamma^t) - F(h)\|_X^p dh \right)^{\max\{p,q\}/p} \frac{dt}{t^{1+\max\{p,q\}/s}}$$

$$+ \frac{C^{\max\{p,q\}}}{k^{1+\max\{p,q\}/s}} \left(\sum_{x \in \Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^p \right)^{\max\{p,q\}/p},$$

where $C = C_G \in (0, \infty)$ is a constant depending on G. Summing over $k \in \mathbb{N}$,

$$\left(\sum_{k=1}^\infty \frac{1}{k^{1+\max\{p,q\}/s}} \left(\sum_{x \in \Gamma} \|f(xv_\Gamma^k) - f(x)\|_X^p \right)^{\max\{p,q\}/p} \right)^{1/\max\{p,q\}}$$

$$\lesssim_\Gamma \left(\int_0^\infty \left(\int_G \|F(hv_\Gamma^t) - F(h)\|_X^p dh \right)^{\max\{p,q\}/p} \frac{dt}{t^{1+\max\{p,q\}/s}} \right)^{1/\max\{p,q\}} + \left(\sum_{x \in \Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^p \right)^{1/p}.$$

$$\tag{39}$$

The desired inequality follows from (38) and (39) along with Theorem 5 for F (although it may be the case that $d_G(v_\Gamma^t, e_G) \not\lesssim t^{1/s}$, a simple rescaling argument gives the same inequality up to constant factors depending on v_Γ). $\qquad\square$

Now we are ready to prove Theorem 6. Of course, assume Γ is torsion-free nonabelian finitely generated nilpotent, and choose v_Γ and s as before.

Proof of Theorem 6. The argument follows the proof of Theorem 3.1 of [LN14], which in turn follows the proof of Claim 7.2 in [ANT13]. Recall c is such that

$$d_W(e_\Gamma, v_\Gamma^k) \leq ck^{1/s}.$$

Fix $n \in \mathbb{N}$ and an $f : \Gamma \to X$. By translation, we may assume $\sum_{x \in B_{(c+3)n}^\Gamma} f(x) = 0$. Then

$$
\left(\sum_{x \in B_{(c+3)n}^\Gamma} \|f(x)\|_X^p \right)^{1/p} = \left(\sum_{x \in B_{(c+3)n}^\Gamma} \left\| \frac{1}{|B_{(c+3)n}^\Gamma|} \sum_{y \in B_{(c+3)n}^\Gamma} (f(x) - f(y)) \right\|_X^p \right)^{1/p}
$$

$$
\leq \left(\frac{1}{|B_{(c+3)n}^\Gamma|} \sum_{x,y \in B_{(c+3)n}^\Gamma} \|f(x) - f(y)\|_X^p \right)^{1/p} \tag{40}
$$

$$
\overset{\text{Lemma 39}}{\lesssim_\Gamma} n \left(\sum_{x \in B_{3(c+3)n}^\Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^p \right)^{1/p}.
$$

Define the cutoff function $\xi : \Gamma \to [0, 1]$ as

$$
\xi(x) := \begin{cases} 1 & x \in B_{(c+1)n}^\Gamma, \\ c + 2 - \frac{d_W(x,e)}{n} & x \in B_{(c+2)n}^\Gamma \setminus B_{(c+1)n}^\Gamma, \\ 0 & x \in \Gamma \setminus B_{(c+2)n}^\Gamma. \end{cases}
$$

Then $\phi := \xi f$ is supported on $B_{(c+2)n}^\Gamma$ and we may apply Theorem 41. As ξ is $\frac{1}{n}$-Lipschitz and takes values in $[0, 1]$, we have for all $a \in S$ and $x \in \Gamma$

$$
\|\phi(x) - \phi(xa)\|_X \leq |\xi(x) - \xi(xa)| \cdot \|f(x)\|_X + |\xi(xa)| \cdot \|f(x) - f(xa)\|_X
$$

$$
\leq \frac{1}{n} \|f(x)\|_X + \|f(x) - f(xa)\|_X. \tag{41}
$$

We have $d_W(e_\Gamma, v_\Gamma^k) \leq cn$ for all $k \in \{1, \dots, n^s\}$. Thus, for every $x \in B_n^\Gamma$ we have $xv_\Gamma^k \in B_{(c+1)n}^\Gamma$ and thus $\phi(x) = f(x)$ and $\phi(xv_\Gamma^k) = f(xv_\Gamma^k)$, so

$$
\left(\sum_{k=1}^{n^s} \frac{1}{k^{1+q/s}} \left(\sum_{x \in B_n^\Gamma} \|f(xv_\Gamma^k) - f(x)\|_X^p \right)^{q/p} \right)^{1/q} \leq \left(\sum_{k=1}^{\infty} \frac{1}{k^{1+q/s}} \left(\sum_{x \in \Gamma} \|\phi(xv_\Gamma^k) - \phi(x)\|_X^p \right)^{q/p} \right)^{1/q}. \tag{42}
$$

On the other hand,

$$
\left(\sum_{x \in \Gamma} \sum_{a \in S} \|\phi(xa) - \phi(x)\|_X^p \right)^{1/p} = \left(\sum_{x \in B_{(c+3)n}^\Gamma} \sum_{a \in S} \|\phi(xa) - \phi(x)\|_X^p \right)^{1/p}
$$

$$
\overset{(41)}{\leq} \frac{|S|^{1/p}}{n} \left(\sum_{x \in B_{(c+3)n}^\Gamma} \|f(x)\|_X^p \right)^{1/p} + \left(\sum_{x \in B_{(c+3)n}^\Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^p \right)^{1/p}
$$

$$
\overset{(40)}{\lesssim_\Gamma} \left(\sum_{x \in B_{3(c+3)n}^\Gamma} \sum_{a \in S} \|f(xa) - f(x)\|_X^p \right)^{1/p},
$$

where the first equality holds since ϕ is supported on $B_{(c+2)n}^\Gamma$. The desired inequality now follows from the above two inequalities combined with Theorem 41 for ϕ. $\qquad\square$

5. Proof of Theorem 9: sublinear growth of cocycles

Our goal in this section is to prove Theorem 9. The proof follows closely that of [ANT13].

Let $(X, \|\cdot\|_X)$ be a q-uniformly convex space. We will see by [NP11] that is enough to prove Theorem 9 for 1-Lipschitz 1-cocycles. Here, $f \in Z^1(\pi)$ be a 1-cocycle with respect to an action $\pi : \Gamma \to \mathrm{Aut}(X)$ of Γ on X by linear isometric automorphisms if the function $f : \Gamma \to X$ satisfies $f(xy) = \pi(x)f(y) + f(x)$ for all $x, y \in \Gamma$. The 1-Lipschitz requirement for f is equivalent to $\|f(a)\|_X \leq 1$ for all $a \in S$.

We first recall the following property of q-uniformly convex spaces.

Lemma 42 ([ANT13, Lemma 3.1]). *Let $(X, \|\cdot\|_X)$ be q-uniformly convex. For a fixed $z \in X$ and linear operator $T: X \to X$ with $\|T\| \le 1$, define*

$$s_n := \frac{1}{2^n} \sum_{j=0}^{2^n-1} T^j z, \quad n \ge 0.$$

Then we have

$$\sum_{i=0}^{\infty} \frac{1}{2^l} \sum_{j=0}^{2^l-1} \left\| s_{(i+1)l} - T^{j2^{il}} s_{il} \right\|_X^q \le (2K_q(X))^q \|z\|_X^q, \quad l \in \mathbb{N}. \tag{43}$$

On the other hand, as X is reflexive, X is ergodic [DS88, p.662], i.e., for every linear isometry $T: X \to X$ and $x \in X$ the sequence $\left\{ \frac{1}{n} \sum_{j=0}^{n-1} T^j x \right\}_{n=1}^{\infty}$ converges in norm. Thus, the operator $P: X \to X$ defined by

$$Px := \lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \pi(v)^n x, \quad x \in X,$$

is well-defined, has norm ≤ 1, is a contraction onto the subspace X_0 of $\pi(v)$-invariant vectors, and is idempotent. As v is a central element of Γ, P commutes with $\pi(g)$ for all $g \in \Gamma$, and so the maps $Pf, (I - P)f: \Gamma \to X$ are both Lipschitz and are 1-cocycles.

Now define linear operators $P_n: X \to X$, $n \in \mathbb{N}$, by

$$P_n := \frac{1}{2^n} \sum_{j=0}^{2^n-1} \pi(v)^j.$$

Of course, $\|P_n\| \le 1$ and P_n commutes with $\pi(g)$ for all $g \in \Gamma$.

Lemma 43 (Analogue of [ANT13, Lemma 4.1]). *Let $(X, \|\cdot\|_X)$ be a q-uniformly convex space. Then for every $l, k, m \in \mathbb{N}$ there exist integers $i \in [k+1, k+m]$ and $j \in [0, 2^l - 1]$ such that*

$$\left\| \pi\left(v^{-j2^{il}} \right) P_{(i+1)l} f\left(v^{\lfloor n^\rho \rfloor} \right) - P_{il} f\left(v^{\lfloor n^\rho \rfloor} \right) \right\|_X \lesssim_G \frac{K_q(X) n}{m^{1/p}}. \tag{44}$$

Proof. By Lemma 42 and the fact that $\|f(a)\|_X \le 1$, we have for each $a \in S$

$$\sum_{i=k+1}^{k+m} \frac{1}{2^l} \sum_{j=0}^{2^l-1} \left\| \pi\left(v^{-j2^{il}} \right) P_{(i+1)l} f(a) - P_{il} f(a) \right\|_X^q = \sum_{i=k+1}^{k+m} \frac{1}{2^l} \sum_{j=0}^{2^l-1} \left\| P_{(i+1)l} f(a) - \pi\left(v^{j2^{il}} \right) P_{il} f(a) \right\|_X^q \le (2K_q(X))^q.$$

Thus

$$|S|(2K_q(X))^q \ge \sum_{i=k+1}^{k+m} \frac{1}{2^l} \sum_{j=0}^{2^l-1} \sum_{a \in S} \left\| \pi\left(v^{-j2^{il}} \right) P_{(i+1)l} f(a) - P_{il} f(a) \right\|_X^q$$

$$\ge m \min_{\substack{k+1 \le i \le k+m \\ 1 \le j \le 2^l-1}} \sum_{a \in S} \left\| \pi\left(v^{-j2^{il}} \right) P_{(i+1)l} f(a) - P_{il} f(a) \right\|_X^q,$$

and so there exist $i \in [k+1, k+m]$, $j \in [0, 2^l - 1]$ such that

$$\max_{a \in S} \left\| \pi\left(v^{-j2^{il}} \right) P_{(i+1)l} f(a) - P_{il} f(a) \right\|_X \lesssim_\Gamma \frac{K_q(X)}{m^{1/q}}.$$

As $v^{\lfloor n^\rho \rfloor} = a_1 \cdots a_b$ for some $a_i \in S$, where $b = O_\Gamma(n)$, we have by the cocycle identity

$$f(v^{\lfloor n^\rho \rfloor}) = \sum_{i=1}^{b} \pi(a_1 \cdots a_{i-1}) f(a_i).$$

35

Since P_r, $r \geq 0$, and $\pi(g)$, $g \in \Gamma$, commute, and v and $g \in \Gamma$ commute, and $\pi(g)$ is an isometry for $g \in \Gamma$, we have

$$\left\| \pi\left(v^{-j2^{il}}\right) P_{(i+1)l} f\left(v^{\lfloor n^\rho \rfloor}\right) - P_{il} f\left(v^{\lfloor n^\rho \rfloor}\right) \right\|_X \leq \sum_{i=1}^{b} \left\| \pi\left(v^{-j2^{il}}\right) P_{(i+1)l} f(a_i) - P_{il} f(a_i) \right\|_X \lesssim_\Gamma \frac{K_q(X) n}{m^{1/q}},$$

as claimed. $\qquad\square$

Lemma 44 (Analogue of [ANT13, Lemma 4.2]). *For every* $m, n \in \mathbb{N}$,

$$\left\| P_m f\left(v^{\lfloor n^\rho \rfloor}\right) \right\|_X \lesssim_\Gamma \frac{n^{(\rho^2 + \rho - 1)/(2\rho - 1)}}{2^{m(\rho-1)/(2\rho-1)}}.$$

Proof. If $\frac{n^\rho}{2^m} \geq \rho - 1$ then we have the obvious bound

$$\left\| P_m f\left(v^{\lfloor n^\rho \rfloor}\right) \right\|_X \leq \left\| f\left(v^{\lfloor n^\rho \rfloor}\right) \right\|_X \leq d_W\left(e_\Gamma, v^{\lfloor n^\rho \rfloor}\right) \lesssim_\Gamma n \left(\frac{n^\rho}{2^m}\right)^{(\rho-1)/(2\rho-1)}.$$

Thus we assume $\frac{n^\rho}{2^m} < \rho - 1$.

Because

$$P_m - \pi(v^k) P_m = \frac{1}{2^m} \sum_{j=0}^{k-1} \pi(v)^j - \frac{1}{2^m} \sum_{j=2^m}^{2^m + k - 1} \pi(v)^j,$$

we have

$$\left\| P_m - \pi(v^k) P_m \right\| \leq \frac{2k}{2^m}.$$

By the cocycle identity,

$$f\left(v^{\lfloor k^\rho \rfloor \lfloor n^\rho \rfloor}\right) = \sum_{j=0}^{\lfloor k^\rho \rfloor - 1} \pi\left(v^{j \lfloor n^\rho \rfloor}\right) f\left(v^{\lfloor n^\rho \rfloor}\right)$$

Because f is 1-Lipschitz, $d_W\left(e_\Gamma, v^{\lfloor k^\rho \rfloor \lfloor n^\rho \rfloor}\right) \lesssim_\Gamma kn$, and $\|P_m\| \leq 1$,

$$kn \gtrsim_\Gamma \left\| P_m f\left(v^{\lfloor k^\rho \rfloor \lfloor n^\rho \rfloor}\right) \right\|_X = \left\| \lfloor k^\rho \rfloor P_m f\left(v^{\lfloor n^\rho \rfloor}\right) - \sum_{j=0}^{\lfloor k^\rho \rfloor - 1} \left(P_m - \pi\left(v^{j \lfloor n^\rho \rfloor}\right) P_m\right) f\left(v^{\lfloor n^\rho \rfloor}\right) \right\|_X$$

$$\geq \lfloor k^\rho \rfloor \left\| P_m f\left(v^{\lfloor n^\rho \rfloor}\right) \right\|_X - \sum_{j=0}^{\lfloor k^\rho \rfloor - 1} \left\| P_m - \pi\left(v^{j \lfloor n^\rho \rfloor}\right) P_m \right\| \cdot \left\| f\left(v^{\lfloor n^\rho \rfloor}\right) \right\|_X$$

$$\geq \lfloor k^\rho \rfloor \left\| P_m f\left(v^{\lfloor n^\rho \rfloor}\right) \right\|_X - \sum_{j=0}^{\lfloor k^\rho \rfloor - 1} \frac{2j \lfloor n^\rho \rfloor}{2^m} d_W\left(e_\Gamma, v^{\lfloor n^\rho \rfloor}\right)$$

and rearranging terms,

$$\left\| P_m f\left(v^{\lfloor n^\rho \rfloor}\right) \right\|_X \lesssim_\Gamma \frac{n}{k^{\rho-1}} + \frac{n^{\rho+1} k^\rho}{2^m}.$$

Since this is true for all k, choosing $k = \left\lceil \left(\frac{(\rho-1)2^m}{\rho n^\rho}\right)^{1/(2\rho-1)} \right\rceil$ gives the stated bound. Indeed, by $\frac{n^\rho}{2^m} < \rho - 1$ we have $k \asymp \left(\frac{(\rho-1)2^m}{\rho n^\rho}\right)^{1/(2\rho-1)}$ and plugging in gives the desired bound. $\qquad\square$

Lemma 45 (Analogue of [ANT13, Lemma 4.3]). *For every* $m, n \in \mathbb{N}$,

$$\left\| f\left(v^{\lfloor n^\rho \rfloor}\right) - P_m f\left(v^{\lfloor n^\rho \rfloor}\right) \right\|_X \lesssim_\Gamma 2^{m/(\rho+1)} n^{1/(\rho+1)}.$$

Proof. If we define $\tilde{f} : \Gamma \to X$ by

$$\tilde{f}(h) := f(h) - P_m f(h) = (I - P_m) f(h),$$

36

then $\tilde{f} \in Z^1(\pi)$. Let $k \geq 1$ be an integer to be determined later. If we set

$$w := -\frac{1}{k}\sum_{j=0}^{k-1}\tilde{f}(v^j),$$

then

$$\|w\|_X \lesssim_\Gamma \frac{1}{k}\sum_{j=0}^{k-1} j^{1/\rho} \lesssim k^{1/\rho}.$$

For every $h \in \Gamma$ we have the following identity:

$$-\pi(h)w + \tilde{f}(h) = \frac{1}{k}\sum_{j=0}^{k-1}\left(\pi(h)\tilde{f}(v^j) + \tilde{f}(h)\right) = \frac{1}{k}\sum_{j=0}^{k-1}\tilde{f}(hv^j) = \frac{1}{k}\sum_{j=0}^{k-1}\tilde{f}(v^j h)$$

$$= \frac{1}{k}\sum_{j=0}^{k-1}\left(\pi(v^j)\tilde{f}(h) + \tilde{f}(v^j)\right) = \frac{1}{k}\sum_{j=0}^{k-1}\pi(v^j)\tilde{f}(h) - w.$$

(45)

But

$$\frac{1}{k}\sum_{j=0}^{k-1}\pi(v^j)\tilde{f}(h) = \frac{1}{k}\sum_{j=0}^{k-1}\left(\pi(v^j)f(h) - \frac{1}{2^m}\sum_{i=0}^{2^m-1}\pi(v^{j+i})f(h)\right) = \frac{1}{2^m}\sum_{i=0}^{2^m-1}\left(\frac{1}{k}\sum_{j=0}^{k-1}\pi(v^j) - \frac{1}{k}\sum_{j=i}^{i+k-1}\pi(v^j)\right)f(h),$$

so that

$$\left\|\frac{1}{k}\sum_{j=0}^{k-1}\pi(v^j)\tilde{f}(h)\right\|_X \leq \frac{d_W(h,e_\Gamma)}{2^m}\sum_{i=0}^{2^m-1}\frac{2i}{k} \leq \frac{2^m}{k}d_W(h,e_\Gamma).$$

Because of (45), $\tilde{f}$ is close to a coboundary in the following sense:

$$\left\|\tilde{f}(h) - (\pi(h)w - w)\right\| \leq \frac{2^m}{k}d_W(h,e_\Gamma).$$

Writing $v^{\lfloor n^\rho \rfloor} = a_1\cdots a_b$ for some $a_i \in S$, where $b = O_\Gamma(n)$, we have

$$\left\|\tilde{f}\left(v^{\lfloor n^\rho \rfloor}\right)\right\|_X = \left\|\sum_{i=1}^b \pi(a_1\cdots a_{i-1})\tilde{f}(a_i)\right\|_X \lesssim_\Gamma \left\|\sum_{i=1}^b \pi(a_1\cdots a_{i-1})(\pi(a_i)w - w)\right\|_X + \frac{n2^m}{k}$$

$$= \left\|\pi\left(v^{\lfloor n^\rho \rfloor}\right)w - w\right\|_X + \frac{n2^m}{k} \lesssim_\Gamma k^{1/\rho} + \frac{n2^m}{k}.$$

With $k = \lceil n^{\frac{\rho}{\rho+1}}2^{\frac{m\rho}{\rho+1}}\rceil$, we have the stated bound. $\qquad\square$

We now prove theorem 9.

Proof of Theorem 9. By [NP11, Theorem 9.1], by amenability of Γ, if X is q-uniformly convex and $f : \Gamma \to X$ is 1-Lipschitz, then there exists a Banach space $(Y, \|\cdot\|_Y)$ that is also q-uniformly convex with $K_q(Y) = K_q(X)$, an action π of Γ on Y by linear isometric automorphisms, and a 1-cocycle $F : \Gamma \to Y$ such that $\omega_F = \omega_f$, where $F : \Gamma \to Y$ is a 1-cocycle if $F(xy) = \pi(x)F(y) + F(x)$ for all $x, y \in \Gamma$. Thus, we may assume without loss of generality that $f \in Z^1(\pi)$ for some action π of Γ on X by linear isometric automorphisms.

Let $C_1 < C_2$ be constants depending on Γ such that

$$C_1 n \leq d_W\left(v^{\lfloor n^\rho \rfloor}, e_\Gamma\right) \leq C_2 n, \quad n \in \mathbb{N}.$$

We may assume t is sufficiently large so that if m is the largest integer such that

$$\frac{2C_1}{C_2}2^{\frac{m}{\rho}}m^{\frac{m(\rho^2+2\rho-2)}{\rho\rho(\rho-1)}} \leq \sqrt{t},$$

(46)

then

$$\frac{p(\rho-1)}{\rho^2+2\rho-2} \le \frac{1}{2}\log_2 m. \tag{47}$$

Given m, let k be the smallest integer such that

$$m^{\frac{\rho+1}{\rho p}+\frac{(\rho^2+2\rho-2)(k+1)}{p\rho(\rho-1)}} \ge t, \tag{48}$$

and define

$$\ell := \left\lceil \frac{\rho^2+2\rho-2}{p(\rho-1)}\log_2 m \right\rceil. \tag{49}$$

Using Lemma 43, we may find integers $i \in [k+1, k+m]$ and $j \in [0, 2^\ell - 1]$ such that for all $n \in \mathbb{N}$,

$$\left\| \pi\left(v^{-j2^{i l}}\right)P_{(i+1)l}f\left(v^{\lfloor n^\rho\rfloor}\right) - P_{il}f\left(v^{\lfloor n^\rho\rfloor}\right) \right\|_X \lesssim_\Gamma \frac{K_q(X)n}{m^{1/p}}. \tag{50}$$

Finally, define

$$n := \left\lceil \frac{1}{C_1}m^{\frac{\rho+1}{\rho p}}2^{\frac{i\ell}{\rho}} \right\rceil, \tag{51}$$

Expanding, we have

$$f\left(v^{\lfloor n^\rho\rfloor}\right) = \pi\left(v^{-j2^{i\ell}}\right)P_{(i+1)\ell}f\left(v^{\lfloor n^\rho\rfloor}\right)+\left(P_{i\ell}f\left(v^{\lfloor n^\rho\rfloor}\right)-\pi\left(v^{-j2^{i\ell}}\right)P_{(i+1)\ell}f\left(v^{\lfloor n^\rho\rfloor}\right)\right)+\left(f\left(v^{\lfloor n^\rho\rfloor}\right)-P_{i\ell}f\left(v^{\lfloor n^\rho\rfloor}\right)\right).$$

Thus by Lemmas 44 and 45 and inequality (50), we obtain:

$$\omega_f\left(d_W\left(v^{\lfloor n^\rho\rfloor}, e_\Gamma\right)\right) \le \left\|f\left(v^{\lfloor n^\rho\rfloor}\right)\right\|_X \lesssim_\Gamma \frac{n^{(\rho^2+\rho-1)/(2\rho-1)}}{2^{(i+1)l(\rho-1)/(2\rho-1)}} + \frac{K_q(X)n}{m^{1/p}} + 2^{il/(\rho+1)}n^{1/(\rho+1)} \overset{(49)\wedge(51)}{\lesssim_\Gamma} \frac{K_q(X)n}{m^{1/p}}. \tag{52}$$

We compute

$$d_W\left(v^{\lfloor n^\rho\rfloor}, e_\Gamma\right) \ge C_1 n \overset{(51)}{\ge} m^{\frac{\rho+1}{\rho p}}2^{\frac{i\ell}{\rho}} \ge m^{\frac{\rho+1}{\rho p}}2^{\frac{(k+1)\ell}{\rho}} \overset{(49)}{\ge} m^{\frac{\rho+1}{\rho p}+\frac{(\rho^2+2\rho-2)(k+1)}{p\rho(\rho-1)}} \overset{(48)}{\ge} t$$

and

$$d_W\left(v^{\lfloor n^\rho\rfloor}, e_\Gamma\right) \le C_2 n \overset{(51)}{\le} \frac{2C_2}{C_1}m^{\frac{\rho+1}{\rho p}}2^{\frac{i\ell}{\rho}} \le \frac{2C_2}{C_1}m^{\frac{\rho+1}{\rho p}}2^{\frac{(k+m)\ell}{\rho}} \overset{(49)}{\le} \frac{2C_1}{C_2}2^{\frac{k+m}{\rho}}m^{\frac{\rho+1}{\rho p}+\frac{k(\rho^2+2\rho-2)}{p\rho(\rho-1)}}m^{\frac{m(\rho^2+2\rho-2)}{p\rho(\rho-1)}}$$

$$\overset{(48)}{<} \frac{2C_1}{C_2}2^{\frac{k+m}{\rho}}tm^{\frac{m(\rho^2+2\rho-2)}{p\rho(\rho-1)}} \overset{(48)}{\le} \frac{2C_1}{C_2}2^{\frac{p(\rho-1)}{\rho^2+2\rho-2}\log_m t}2^{\frac{m}{\rho}}tm^{\frac{m(\rho^2+2\rho-2)}{p\rho(\rho-1)}} \overset{(46)}{\le} t^{3/2}2^{\frac{p(\rho-1)}{\rho^2+2\rho-2}\log_m t} \overset{(47)}{\le} t^2.$$

Therefore $t \le d_W\left(v^{\lfloor n^\rho\rfloor}, e_\Gamma\right) \le t^2$. The definition of m implies $m \gtrsim_\Gamma \min\left\{\log n, \frac{p\log n}{\log\log n}\right\}$, and thus by (52), we have

$$\frac{\omega_f\left(d_W\left(v^{\lfloor n^\rho\rfloor}, e_\Gamma\right)\right)}{n} \lesssim K_q(X)\left(\frac{\log\log n}{\log n}\right)^{1/p}.$$

The proof of Theorem 9 is complete. $\qquad\square$

6. Derivation of Theorem 24 from Dorronsoro's Theorem 22

In this section, we prove Theorem 24, which proves Theorem 17 since it is a special case. This section follows Section 7 of [FO20a] closely.

Using an argument similar to the Vitali covering lemma, we may find a collection $\mathscr{B}_r$ of balls B of radius r whose union covers G, and such that the concentric balls $\hat{B}$ of radius $(n+1)r$ have bounded

overlap. Then

$$\int_G \left(\frac{1}{r^\alpha} \left| \sum_{j=0}^{n} (-1)^j \binom{n}{j} f(x(\delta_r(v))^j) \right| \right)^p dx$$

$$\leq \sum_{B \in \mathcal{B}_r} \int_B \left(\frac{1}{r^\alpha} \left| \sum_{j=0}^{n} (-1)^j \binom{n}{j} f(x(\delta_r(v))^j) \right| \right)^p dx$$

$$\lesssim_p \sum_{j=0}^{n} \binom{n}{j}^p \sum_{B \in \mathcal{B}_r} \int_B \left(\frac{1}{r^\alpha} \left| (f(x(\delta_r(v))^j) - A_{\hat{B}}^{\lfloor \alpha \rfloor} f(x(\delta_r(v))^j) \right| \right)^p dx$$

$$+ \int_B \left(\frac{1}{r^\alpha} \left| \sum_{j=0}^{n} (-1)^j \binom{n}{j} A_{\hat{B}}^{\lfloor \alpha \rfloor} f(x(\delta_r(v))^j) \right| \right)^p dx$$

$$\lesssim_{n,p} \sum_{B \in \mathcal{B}_r} \int_{\hat{B}} \left(\frac{1}{r^\alpha} \left| f(x) - A_{\hat{B}}^{\lfloor \alpha \rfloor} f(x) \right| \right)^p dx + 0$$

$$= \sum_{B \in \mathcal{B}_r} \left[\frac{1}{r^\alpha} \beta_{f,\lfloor \alpha \rfloor, p}(\hat{B}) \right]^p |\hat{B}| \lesssim_{G,\alpha} \int_G \left[\frac{1}{r^\alpha} \cdot \beta_{f,\lfloor \alpha \rfloor, p}(B_{2(n+1)r}(x)) \right]^p dx.$$

We have used in the penultimate inequality the fact that since $A_{\hat{B}}^{\lfloor \alpha \rfloor} f$ is a polynomial of weighted degree at most $\lfloor \alpha \rfloor$, and since $\delta_r(v) \in V_{\lfloor \alpha/n \rfloor + 1} \oplus \cdots \oplus V_s$, $A_{\hat{B}}^{\lfloor \alpha \rfloor} f(x(\delta_r(v))^j)$ is a polynomial in j of degree at most $n - 1$, and hence

$$\Delta_{\delta_r(v)}^n f(x) = \sum_{j=0}^{n} (-1)^{n-j} \binom{n}{j} A_{\hat{B}}^{\lfloor \alpha \rfloor} f(x(\delta_r(v))^j) = 0.$$

In the last inequality we used Corollary 47. Now, the inequality (9) follows from Minkowski's integral inequality and Theorem 22 with $q = p$:

$$\left(\int_0^\infty \left[\int_G \left(\frac{1}{r^\alpha} \left| \sum_{j=0}^{n} (-1)^j \binom{n}{j} f(x(\delta_r(v))^j) \right| \right)^p dx \right]^{2/p} \frac{dr}{r} \right)^{1/2}$$

$$\lesssim_{G,\alpha,n,p} \left(\int_0^\infty \left[\int_G \left[\frac{1}{r^\alpha} \beta_{f,\lfloor \alpha \rfloor, p}(B_{2(n+1)r}(x)) \right]^p dx \right]^{2/p} \frac{dr}{r} \right)^{1/2}$$

$$\leq \left(\int_G \left[\int_0^\infty \left[\frac{1}{r^\alpha} \beta_{f,\lfloor \alpha \rfloor, p}(B_{2(n+1)r}(x)) \right]^2 \frac{dr}{r} \right]^{p/2} dx \right)^{1/p} \lesssim_{G,\alpha,n,p} \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}.$$

The proof of Theorem 24 is complete.

7. Proof of Dorronsoro's Theorem 22 for Carnot groups

We need to prove two directions. In one direction, we assume $f \in S_\alpha^p(G)$ and prove the $\lesssim$ direction of (8). In the other direction, we assume $f \in L^p(G)$ with the left-hand side of (8) finite, and prove $f \in S_\alpha^p(G)$ along with the $\gtrsim$ direction of (8). Our proof will be based on Dorronsoro's original proof of the $G = \mathbb{R}^n$ case in [Dor85] and will borrow modifications inspired by the proof of $G = \mathbb{H}^k$ case in [FO20a] to deal with general Carnot groups.

We will first prove Theorem 22 for $q = 1$, the $\lesssim$ direction in subsection 7.2 and the $\gtrsim$ direction in subsection 7.3. Since the left-hand side of (8) is minimized when $q = 1$, this will finish the proof of the $\gtrsim$ direction; for the $\lesssim$ direction, we will see in subsection 7.4 that the $\lesssim$ inequality for $q = 1$ implies the $\lesssim$ inequality for $1 \leq q < \frac{\min\{p,2\} n_h}{n_h - \min\{p,2\}}$.

For the $\lesssim$ direction, we will first see by an approximation argument that it is enough to look at smooth functions f (subsubsection 7.2.1), for which it turns out that we may as well approximate at all scales

39

simultaneously by Taylor polynomials (subsubsection 7.2.2). Then, for $0 < \alpha < 1$, it will turn out that the desired inequality has already been proven in [CRTN01, Theorem 5] (subsubsection 7.2.4), and for $\alpha > 1$ nonintegral we may use an induction argument on α (subsubsection 7.2.5). We finish off the case of $\alpha \geq 1$ integral by an interpolation argument (subsubsection 7.2.6), with the extra terms arising in this process having been taken care of by a homogenization argument (subsubsection 7.2.3).

For the $\gtrsim$ direction, we assume $f \in L^p(G)$ and the finiteness of a certain singular integral and need to derive $f \in S_\alpha^p(G)$. We first see again that we may approximate by a fixed polynomial for all scales (which we suspect to be the Taylor polynomial in the distributive sense but we can only prove this up to first derivatives) in Proposition 54. For $0 < \alpha \leq 1$ we again use characterizations of the fractional Laplacian given by [CRTN01, Theorem 5] for $0 < \alpha < 1$ and [DNM21, Theorem 1.4] for $\alpha = 1$ (subsubsection 7.3.1). We then prove the case $\alpha > 1$ by induction (subsubsection 7.3.2).

For simplicity, we define the L^1-beta numbers

$$\beta_{f,d}(B_r(x)) := \beta_{f,d,1}(B_r(x)).$$

If we define the function

$$\mathfrak{G}_\alpha f(x) := \left(\int_0^\infty \left(\frac{\beta_{f,\lfloor \alpha \rfloor}(B_r(x))}{r^\alpha} \right)^2 \frac{dr}{r} \right)^{1/2}, \quad x \in G,$$

then Theorem 22 states that

$$\|\mathfrak{G}_\alpha f\|_{L^p(G)} \asymp_{G,\alpha,p} \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}, \quad f \in S_\alpha^p. \tag{53}$$

Before we begin the proof, we briefly remark on coordinate and multi-index notation on G. Recall the coordinate system $x = \exp\left(\sum_{r=1}^s \sum_{i=1}^{k_r} x_{r,i} X_{r,i} \right)$. A multi-index $\gamma = ((\gamma_{r,i})_{i=1}^{k_r})_{r=1}^s$ is a multi-index on $\sum_{r=1}^s k_r$ entries, and we denote

$$|\gamma| := \sum_{r=1}^s \sum_{i=1}^{k_r} r |\gamma_{r,i}| \in \mathbb{N}, \quad \gamma! := \prod_{r=1}^s \prod_{i=1}^{k_r} \gamma_{r,i}!, \quad x^\gamma := \prod_{r=1}^s \prod_{i=1}^{k_r} x_{r,i}^{\gamma_{r,i}} \in \mathbb{R}.$$

The latter should not be confused with the previous notation $x^t = \exp\left(\sum_{r=1}^s \sum_{i=1}^{k_r} t x_{r,i} X_{r,i} \right)$ for $t \in \mathbb{R}$.

We now define the weighted degree of linear combinations of 'polynomial differential operators', that is, linear combinations of differential operators of the form

$$x^{\gamma_1} \left(\frac{\partial}{\partial x} \right)^{\gamma_2},$$

by assigning weight $|\gamma_2| - |\gamma_1|$ to the above term. Then, a polynomial differential operator of homogeneous weighted degree d_1 will act on a polynomial of homogeneous weighted degree d_2 to produce a polynomial of homogeneous weighted degree $d_2 - d_1$. We observe that the left-invariant differential operators $X_{r,i}$ is a polynomial differential operator of homogeneous weighted degree r for $r = 1, \cdots, s$ and $i = 1, \cdots, k_r$.

Recall from subsection 1.1 that the polynomials of weighted degree $\leq r$ are left-invariant.

7.1. L^1-optimality of the $A_{x,r}^d f$'s.

We begin the proof by discussing some basic properties of $A_{x,r}^d f$ and $\beta_{f,d}(B_r(x))$.

It is well-known that $A_{x,r}^d f$ is the optimal $L^2(B_r(x))$-approximation of f in $\mathscr{A}_d$. It turns out that it is also an optimal $L^1(B_r(x))$-approximation up to constants. To see this, we first observe that for all $d \geq 0$, there exists an integral formula for the coefficients of $A_{x,r}^d f$, given by the Gram–Schmidt process. For example in $d = 1$,

$$A_{x,r}^1 f(xy) = \langle f \rangle_{B_r(x)} + \sum_{j=1}^k \frac{\fint_{B_r(x)} f(z)(z_{1,j} - x_{1,j}) dz}{\fint_{B_r(x)} (z_{1,j} - x_{1,j})^2 dz} y_j, \quad y \in G. \tag{54}$$

More generally, for each $d \geq 0$, there exists a family $\{p_\gamma^d\}_{|\gamma| \leq d}$ of polynomials such that

$$A_{0,1}^d f(y) = \sum_{|\gamma| \leq d} \left(\int_{B_1} f p_\gamma^d \right) y^\gamma, \quad y \in G,$$

and by rescaling we have for $x \in G$ and $r > 0$ that

$$A_{x,r}^d f(xy) = \sum_{|\gamma| \leq d} r^{-n_h - |\gamma|} \left(\int_{B_r(x)} f(z) p_\gamma^d (\delta_{1/r}(x^{-1}z)) dz \right) y^\gamma.$$

Therefore

$$\|A_{x,r}^d f\|_{L^\infty(B_r(x))} \lesssim_{G,d} \fint_{B_r(x)} |f(y)| dy. \tag{55}$$

In other words, $A_{x,r}^d$ is a linear projection of $L^1(B_r(x))$ onto $\mathscr{A}_d|_{B_r(x)}$ which is bounded in the $L^1 \to L^\infty$ norm.

We now prove the L^1-optimality of the $A_{x,r}^d f$'s.

Lemma 46. *For $x \in G$, $r > 0$, $f \in L_{\mathrm{loc}}^1(B_r(x))$, and $d \in \mathbb{Z}_{\geq 0}$,*

$$\fint_{B_r(x)} |f(y) - A_{x,r}^d f(y)| dy \lesssim_{G,d} \fint_{B_r(x)} |f(y) - A(y)| dy, \quad A \in \mathscr{A}_d.$$

Proof. From (55),

$$\fint_{B_r(x)} |f - A_{x,r}^d f| \leq \fint_{B_r(x)} |f - A| + \fint_{B_r(x)} |A_{x,r}^d (f - A)| \lesssim_{G,d} \fint_{B_r(x)} |f - A|.$$

$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

It follows that β possesses a weak monotonicity property.

Corollary 47. *Let $x, y \in G$, $0 < r < s < \infty$, $d \in \mathbb{Z}_{\geq 0}$ be such that $B_r(x) \subset B_s(y)$. Then*

$$\beta_{f,d}(B_r(x)) \lesssim_{G,d} \left(\frac{s}{r} \right)^{n_h} \beta_{f,d}(B_s(y)).$$

By Corollary 47, we have that

$$\beta_{f,d}(x, r) \lesssim_{G,d} \beta_{f,d}(x, s) \lesssim_{G,d} \beta_{f,d}(x, 2r), \quad r \leq s \leq 2r.$$

Thus $\mathfrak{G}_\alpha f(x)$ is comparable to the series

$$\left(\sum_{i=-\infty}^{\infty} \left[2^{-i\alpha} \beta_{f,d}(x, 2^i) \right]^2 \right)^{1/2}.$$

One can also see that, if X is a left-invariant differential operator of weighted order n, then because $A_{x,r}^d f$ is a polynomial of weighted degree d,

$$\|X A_{x,r}^d f\|_{L^\infty(B_r(x))} \lesssim_{G,n,d,X} r^{-n} \|A_{x,r}^d f\|_{L^\infty(B_r(x))} \lesssim_{G,d} r^{-n} \fint_{B_r(x)} |f|. \tag{56}$$

7.2. Proof of the $\lesssim$ direction, $q = 1$.

7.2.1. *Approximation by smooth compactly supported functions.*

It is enough to prove the $\lesssim$ statement of Theorem 22 for f in the space $\mathscr{D}$ of smooth compactly supported functions on G. Indeed, by [Fol75, Theorem 4.5], $\mathscr{D}$ is dense in S_α^p for all $1 < p < \infty$ and $\alpha \geq 0$. Given $f \in S_\alpha^p$, choose a sequence $\{f_j\}_{j=1}^\infty \subset \mathscr{D}$ that converges to f in S_α^p. Then $f_j \to f$ in $L^p(G)$, so by the contraction property (55) we have

$$\beta_{f_j,\lfloor\alpha\rfloor}(B_r(x)) \to \beta_{f,\lfloor\alpha\rfloor}(B_r(x)) \text{ as } j \to \infty, \quad x \in G, \ r > 0.$$

Now, by two applications of Fatou's lemma, it follows that

$$\|\mathfrak{G}_\alpha f\|_{L^p(G)} \leq \liminf_{j\to\infty} \|\mathfrak{G}_\alpha f_j\|_{L^p(G)} \lesssim_{G,\alpha,p} \liminf_{j\to\infty} \|(-\Delta_p)^{\alpha/2} f_j\|_{L^p(G)} = \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}.$$

This completes the proof of our claim.

7.2.2. *Taylor polynomials.*

Recall the coordinate system on G where each $y \in G$ is uniquely expressed as $y = \exp\left(\sum_{r=1}^s \sum_{i=1}^{k_r} y_{r,i} X_{r,i}\right)$, $y_{r,i} \in \mathbb{R}$. Let $f \in \mathscr{D}$ and $x \in G$. The well-known Taylor formula tells us that for $n \in \mathbb{Z}_{\geq 0}$,

$$f(xy) = f(x) + \sum_{j=1}^n \frac{1}{j!} \frac{d^j}{dt^j}\bigg|_{t=0} f(xy^t) + \frac{1}{n!} \int_0^1 (1-t)^n \frac{d^{n+1}}{dt^{n+1}} f(xy^t) dt.$$

As y^t is the one-parameter subgroup generated by $\sum_{r=1}^s \sum_{i=1}^{k_r} y_{r,i} X_{r,i}$, we have

$$f(xy) = f(x) + \sum_{j=1}^n \frac{1}{j!} \left(\sum_{r=1}^s \sum_{i=1}^{k_r} y_{r,i} X_{r,i}\right)^j f(x) + \frac{1}{n!} \int_0^1 (1-t)^n \left(\sum_{r=1}^s \sum_{i=1}^{k_r} y_{r,i} X_{r,i}\right)^{n+1} f(xy^t) dt.$$

For each multi-index γ, if we denote by $\mathrm{Sym}(X^\gamma)$ the 'coefficient operator' of y^γ in $\frac{\gamma!}{j!}\left(\sum_{r=1}^s \sum_{j=1}^{k_r} y_{r,j} X_{r,j}\right)^{|\gamma|}$ where $j = \sum_{r=1}^s \sum_{i=1}^{k_r} \gamma_{r,i}$, then the above expression can be rewritten as

$$f(xy) = \sum_{\substack{\gamma \text{ multi-index} \\ \sum_{r=1}^s \sum_{i=1}^{k_r} \gamma_{r,i} \leq n}} \frac{1}{\gamma!} \left[\mathrm{Sym}(X^\gamma) f(x)\right] y^\gamma + (n+1) \int_0^1 \sum_{\substack{\gamma \text{ multi-index} \\ \sum_{r=1}^s \sum_{i=1}^{k_r} \gamma_{r,i} = n+1}} \frac{1}{\gamma!} \left[\mathrm{Sym}(X^\gamma) f(xy^t)\right] y^\gamma dt.$$

By isolating the terms with multi-index γ such that $|\gamma| \leq n$, we see that there exist for each multi-index γ with $|\gamma| \geq n+1$ and $\sum_{r=1}^s \sum_{i=1}^{k_r} \gamma_{r,i} \leq n+1$ a function $q_{x,\gamma} f \in \mathscr{D}$ such that

$$f(xy) = \underbrace{\sum_{|\gamma| \leq n} \frac{1}{\gamma!} \left[\mathrm{Sym}(X^\gamma) f(x)\right] y^\gamma}_{=:T_x^n f(xy)} + \underbrace{\sum_{\substack{|\gamma| \geq n+1 \\ \sum_{r=1}^s \sum_{i=1}^{k_r} \gamma_{r,i} \leq n+1}} q_{x,\gamma} f(y) y^\gamma}_{=o(d_G(y,e_G)^n) \text{ as } d_G(y,e_G)\to 0}, \quad y \in G,$$

where we denote the first sum to be the *Taylor polynomial of f of degree n at $x \in G$*, while we observe the second term to be $o(d_G(y, e_G)^n)$ as $d_G(y, e_G) \to 0$ because $d_G(y, e_G) \asymp_G \sum_{r=1}^s \sum_{i=1}^{k_r} |y_{r,i}|^{1/r}$.

We may observe that if Y is a left-invariant homogeneous differential operator of degree $d \leq n$, then

$$Yf(xy) = \underbrace{\sum_{|\gamma| \leq n} \frac{1}{\gamma!} \left[\mathrm{Sym}(X^\gamma) f(x)\right] Y y^\gamma}_{=T_x^{n-d} Yf(xy)} + \underbrace{\sum_{\substack{|\gamma| \geq n+1 \\ \sum_{r=1}^s \sum_{i=1}^{k_r} \gamma_{r,i} \leq n+1}} Y(q_{x,\gamma} f(y) y^\gamma)}_{=o(d_G(y,e_G)^{n-d}) \text{ as } d_G(y,e_G)\to 0}, \quad y \in G,$$

with the second sum being $o(d_G(y, e_G)^{n-d})$ and the first sum being a polynomial of weighted degree at most $n - d$ and hence is the Taylor polynomial of Yf of degree $n - d$ at $x \in G$. This proves that

$$Y T_x^n f = T_x^{n-d} Y f.$$

In particular, we have for the left-invariant horizontal vector fields that with $d \geq 1$,

$$X_i T_x^d f = T_x^{d-1} X_i f, \quad i = 1, \cdots, k.$$

We may define the β-numbers and $\mathfrak{G}$-function using the Taylor polynomial instead.

$$\tilde{\beta}_{f,d}(B_r(x)) := \fint_{B_r(x)} |f - T_{x,r}^{\lfloor \alpha \rfloor} f|,$$

and

$$\tilde{\mathfrak{G}}_\alpha f(x) := \left(\int_0^\infty \left(\frac{1}{r^\alpha} \tilde{\beta}_{f,\lfloor \alpha \rfloor}(B_r(x)) \right)^2 \frac{dr}{r} \right)^{1/2}, \quad x \in G.$$

Proposition 48. *If $f \in \mathcal{D}$ and α is nonintegral, then $\mathfrak{G}_\alpha f(x) \asymp_{G,\alpha} \tilde{\mathfrak{G}}_\alpha f(x)$ for all $x \in G$.*

Proof. That $\mathfrak{G}_\alpha f(x) \lesssim_{G,\alpha} \tilde{\mathfrak{G}}_\alpha f(x)$ follows from Lemma 46. Denote $d = \lfloor \alpha \rfloor$ and

$$A_{x,r}^d f(xy) = \sum_{|\gamma| \leq d} f_\gamma(x,r) y^\gamma,$$

then

$$f_\gamma(x,r) = \frac{1}{\gamma!} \operatorname{Sym}(X^\gamma) A_{x,r}^d f(x).$$

Also denote

$$T_x^d f(xy) = \sum_{|\gamma| \leq d} f_\gamma(x) y^\gamma,$$

then, in the case $f \in \mathcal{D}$, $f_\gamma(x,r) \to f_\gamma(x)$ as $r \to 0$ by Taylor expansion. Indeed, recalling that $f(xy) = T_x^d f(xy) + O_f(d_G(y, e_G)^{d+1})$,

$$\|A_{x,r}^d f - T_x^d f\|_{L^\infty(B_r(x))} = \|A_{x,r}^d[f - T_x^d f]\|_{L^\infty(B_r(x))} \overset{(55)}{\underset{G,d}{\lesssim}} \fint_{B_r(x)} |f - T_x^d f| = O_{G,d,f}(r^{d+1}) \quad \text{as } r \to 0.$$

By the above expansions for $A_{x,r}^d f$ and $T_x^d f$, we have

$$\| \sum_{|\gamma| \leq d} (f_\gamma(x,r) - f_\gamma(x)) r^{|\gamma|} y^\gamma \|_{L^\infty(B_1)} = \| \sum_{|\gamma| \leq d} (f_\gamma(x,r) - f_\gamma(x)) y^\gamma \|_{L^\infty(B_r)} = O_{G,d,f}(r^{d+1}).$$

But because the space $\mathscr{A}^d$ of polynomials of weighted degree $\leq d$ normed by $L^\infty(B_1)$ is finite dimensional, all norms are equivalent, and hence

$$\max_{|\gamma| \leq d} r^{|\gamma|} |f_\gamma(x,r) - f_\gamma(x)| = O_{G,d,f}(r^{d+1}).$$

This completes the proof of the convergence $f_\gamma(x,r) \to f_\gamma(x)$ as $r \to 0$.

With this, we provide a different bound on $|f_\gamma(x,r) - f_\gamma(x)|$ using the β numbers. For each $x \in G$,

$$
\begin{aligned}
|f_\gamma(x,r) - f_\gamma(x)| &\leq \sum_{i=-\infty}^{0} |f_\gamma(x, 2^i r) - f_\gamma(x, 2^{i-1} r)| \\
&= \frac{1}{\gamma!} \sum_{i=-\infty}^{0} \left| \operatorname{Sym}(X^\gamma)(A_{x,2^i r}^d f - A_{x,2^{i-1} r}^d f)(x) \right| \\
&\overset{(56)}{\underset{G,|\gamma|}{\lesssim}} \sum_{i=-\infty}^{0} 2^{-i|\gamma|} r^{-|\gamma|} \fint_{B_{2^{i-1} r}(x)} |f - A_{x,2^i r}^d f| \\
&= \sum_{i=-\infty}^{0} 2^{-i|\gamma|} r^{-|\gamma|} \beta_{f,d}(B_{2^i r}(x)) \\
&\overset{\text{Corollary 47}}{\underset{G,d}{\lesssim}} \int_0^{2r} \beta_{f,d}(B_u(x)) u^{-|\gamma|} \frac{du}{u}.
\end{aligned}
\tag{57}
$$

43

(With the Cauchy–Schwartz inequality, we may prove that $|f_\gamma(x,r) - f_\gamma(x)| \lesssim_{G,\alpha} r^{\alpha-|\gamma|}\mathfrak{G}_\alpha f(x)$ for $|\gamma| < \alpha$; see Proposition 54 for the computation. This is weaker than the $O_{G,d,f}(r^{d+1-|\gamma|})$ bound we get using Taylor approximation. This is because the Taylor approximation argument leverages on the fact that f has $(d+1)$ and higher derivatives, whereas the $\mathfrak{G}_\alpha f$ (in principle) only measures the "α-th derivative" of f at x; however, the latter approach is more natural since we have to work with $\mathfrak{G}_\alpha f$.)

Now we bound

$$\tilde{\beta}_{f,d}(B_r(x)) = \fint_{B_r(x)} |f - T_x^d f|$$

$$\leq \fint_{B_r(x)} |f - A_{x,r}^d f| + \fint_{B_r(x)} |T_x^d f - A_{x,r}^d f|$$

$$\leq \beta_{f,d}(B_r(x)) + C \sum_{|\gamma| \leq d} |f_\gamma(x,r) - f_\gamma(x)| r^{|\gamma|}$$

$$\leq \beta_{f,d}(B_r(x)) + C \sum_{i=0}^{d} r^i \int_0^{2r} \beta_{f,d}(B_u(x)) u^{-i} \frac{du}{u},$$

where C is a constant depending on G and d. By plugging into the definition of $\tilde{\mathfrak{G}}$ and using the triangle inequality,

$$\tilde{\mathfrak{G}}_\alpha f(x) \leq \mathfrak{G}_\alpha f(x) + C \sum_{i=0}^{d} \left(\int_0^\infty \left[r^{i-\alpha} \int_0^{2r} \beta_{f,d}(B_u(x)) u^{-i} \frac{du}{u} \right]^2 \frac{dr}{r} \right)^{1/2}.$$

But by Hardy's inequality (23), we have

$$\tilde{\mathfrak{G}}_\alpha f(x) \leq \mathfrak{G}_\alpha f(x) + C \sum_{i=0}^{d} \frac{1}{\alpha - i} \left(\int_0^\infty \left[\frac{\beta_{f,d}(B_r(x))}{r^\alpha} \right]^2 \frac{dr}{r} \right)^{1/2}$$

$$\lesssim_{G,\alpha} \mathfrak{G}_\alpha f(x).$$

This is where we use the fact that α is nonintegral. $\qquad\square$

7.2.3. *Inhomogeneous estimates to homogeneous estimates.* We provide one more reduction in the proof of Theorem 22. It is enough to prove the weaker estimate

$$\|\mathfrak{G}_\alpha f\|_{L^p(G)} \lesssim_{G,p,\alpha} \|f\|_{p,\alpha} = \|f\|_{L^p(G)} + \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}, \quad f \in \mathcal{D}.$$

(In fact, the inequality Dorronsoro proves in [Dor85, Theorem 2] is

$$\|f\|_{L^p(G)} + \|\mathfrak{G}_\alpha f\|_{L^p(G)} \asymp_{\alpha,p} \|f\|_{p,\alpha}$$

when $G = \mathbb{R}^n$.) This is because a dimensional analysis reveals that the above inequality is inhomogeneous, so by exploiting scaling properties we may remove the inhomogeneous term $\|f\|_{L^p(G)}$.

Lemma 49. *For $1 < p < \infty$ and $\alpha > 0$, if we have the estimate,*

$$\|\mathfrak{G}_\alpha f\|_{L^p(G)} \leq C_{G,p,\alpha} \|f\|_{p,\alpha}, \quad f \in \mathcal{D}, \tag{58}$$

then with the same constant $C_{G,p,\alpha}$,

$$\|\mathfrak{G}_\alpha f\|_{L^p(G)} \leq C_{G,p,\alpha} \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}, \quad f \in \mathcal{D}. \tag{59}$$

Proof. Given $f \in \mathcal{D}$, set $f_s := f \circ \delta_s \in \mathcal{D}$, for $s > 0$, where δ_s is the Carnot group dilation. It is easy to see that

$$\begin{cases} \|f_s\|_{L^p(G)} = s^{-n_h/p} \|f\|_{L^p(G)}, \\ \|\mathfrak{G}_\alpha f_s\|_{L^p(G)} = s^{\alpha - n_h/p} \|\mathfrak{G}_\alpha f\|_{L^p(G)}, \end{cases} \quad s > 0.$$

Also, it is morally clear that we should have

$$\|(-\Delta_p)^{\alpha/2} f_s\|_{L^p(G)} = s^{\alpha - n_h/p} \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}, \quad s > 0,$$

and one may indeed verify this using the definition of $(-\Delta_p)^{\alpha/2}$; see the proof of [FO20a, Lemma 2.6] for details.

We now obtain (59) for f from (58) for f_s and taking $s \to \infty$. $\qquad\qquad\qquad\qquad\square$

7.2.4. *The case* $0 < \alpha < 1$. We now begin proving Theorem 22.

Fix $1 < p < \infty$ and $0 < \alpha < 1$. Note that we have

$$\tilde{\mathfrak{G}}_\alpha f(x) = \left(\int_0^\infty \left[\frac{1}{r^\alpha} \fint_{B_r} |f(xy) - f(x)| dy \right]^2 \frac{dr}{r} \right)^{1/2}, \quad f \in \mathscr{D}, \ x \in G,$$

then by Proposition 48

$$\mathfrak{G}_\alpha f(x) \asymp_{G,\alpha} \tilde{\mathfrak{G}}_\alpha f(x), \quad x \in G.$$

But by [CRTN01, Theorem 5],

$$\|\tilde{\mathfrak{G}}_\alpha f\|_{L^p(G)} \asymp_{G,\alpha,p} \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}, \quad f \in \mathscr{D}.$$

Therefore $\lesssim$ of Theorem 22 follows in this case. Note that we haven't proven both directions of the inequalities yet due to the restriction $f \in \mathscr{D}$.

7.2.5. *The case α nonintegral, $\alpha > 1$.* We will reduce to the case $0 < \alpha < 1$.

For induction, we need the following result. Although we are currently working with $f \in \mathscr{D}$, we state the proposition below for $f \in S_\alpha^p(G)$, because we will need it later again in the proof of the $\gtrsim$ direction of Dorronsoro's theorem.

Proposition 50 ([Fol75, Theorem 4.10]). *Let $1 < p < \infty$ and $\alpha > 1$, and let $f \in L^p(G)$. Then $f \in S_\alpha^p(G)$ if and only if $f \in S_{\alpha-1}^p(G)$ and the distributional derivatives $X_i f \in S_{\alpha-1}^p(G)$ for $i = 1, \cdots, k$, in which case*

$$\sum_{j=1}^k \|(-\Delta_p)^{(\alpha-1)/2} X_j f\|_{L^p(G)} \asymp_{G,\alpha,p} \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}, \quad f \in S_\alpha^p(G). \tag{60}$$

Remark 51. *To be precise, Theorem 4.10 of [Fol75] states the inhomogeneous estimate*

$$\|f\|_{L^p(G)} + \|(-\Delta_p)^{(\alpha-1)/2} f\|_{L^p(G)} + \sum_{j=1}^k \|X_j f\|_{L^p(G)} + \|(-\Delta_p)^{(\alpha-1)/2} X_j f\|_{L^p(G)}$$

$$\asymp_{G,\alpha,p} \|f\|_{L^p(G)} + \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}, \quad f \in S_\alpha^p(G).$$

But now (60) *easily follows using the homogenization argument of subsubsection 7.2.3.*

The following is an analogue of [Dor85, Theorem 5] and [FO20a, Proposition 4.2].

Proposition 52 (Analogue of [Dor85, Theorem 5] and [FO20a, Proposition 4.2]). *Let $1 < p < \infty$ and $\alpha > 0$. Then,*

$$\|\mathfrak{G}_{\alpha+1} f\|_{L^p(G)} \lesssim_{G,\alpha} \sum_{j=1}^k \|\mathfrak{G}_\alpha(X_j f)\|_{L^p(G)}, \qquad f \in \mathscr{D}.$$

Proof. Let $d = \lfloor \alpha \rfloor$. Define a function $\tilde{A}_{x,r}^{d+1} f \in \mathscr{A}_{d+1}$ for $x \in G$ and $r > 0$ as

$$\tilde{A}_{x,r}^{d+1} f(xy) := T_x^{d+1} f(xy) - \langle T_x^{d+1} f \rangle_{B_r(x)} + \langle f \rangle_{B_r(x)}.$$

We choose $C \geq 1$ depending on G such that the weak 1-Poincaré inequality due to Jerison [Jer86] holds:

$$\fint_{B_s(y)} |g - \langle g \rangle_{B_s(y)}| \lesssim_G s \fint_{B_{Cs}(y)} |\nabla g|, \quad y \in G, \ s > 0, \ g \in \mathscr{D}.$$

As $\langle f - \tilde{A}^{d+1}_{x,r} f \rangle_{B_r(x)} = 0$, we have

$$\beta_{f,d+1}(B_r(x)) \overset{\text{Lemma } 46}{\lesssim_{G,d}} \fint_{B_r(x)} |f - \tilde{A}^{d+1}_{x,r} f| \overset{\text{weak 1-Poincaré}}{\lesssim_G} r \fint_{B_{Cr}(0)} |\nabla f(xy) - \nabla T^{d+1}_x f(xy)| dy$$

$$\asymp_G r \sum_{j=1}^{k} \fint_{B_{Cr}(0)} |X_j f(xy) - X_j T^{d+1}_x f(xy)| = r \sum_{j=1}^{k} \tilde{\beta}_{X_j f, d}(B_{Cr}(x)).$$

because $X_j T^{d+1}_x f = T^d_x X_j f$. Now, by definition,

$$\mathfrak{G}_{\alpha+1} f(x) = \left(\int_0^\infty \left[\frac{1}{r^{\alpha+1}} \beta_{f,d+1}(B_r(x)) \right]^2 \frac{dr}{r} \right)^{1/2} \lesssim_{G,d} \sum_{j=1}^{k} \left(\int_0^\infty \left[\frac{1}{r^\alpha} \tilde{\beta}_{X_j f, d}(B_{Cr}(x)) \right]^2 \frac{dr}{r} \right)^{1/2}$$

$$= C^\alpha \sum_{j=1}^{k} \tilde{\mathfrak{G}}_\alpha (X_j f)(x), \quad x \in G.$$

By Proposition 48, the proof is complete. $\qquad\square$

If we suppose $\lesssim$ of Theorem 22 for $a < d$ nonintegral, $d \in \mathbb{N}$, then from the above two Propositions, for $d < \alpha < d+1$ and $f \in \mathscr{D}$,

$$\|\mathfrak{G}_\alpha f\|_{L^p(G)} \lesssim_{G,\alpha} \sum_{j=1}^{k} \|\mathfrak{G}_{\alpha-1}(X_j f)\|_{L^p(G)} \lesssim_{G,\alpha,p} \sum_{j=1}^{k} \|(-\Delta_p)^{(\alpha-1)/2} X_j f\|_{L^p(G)} \asymp_{G,\alpha,p} \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}.$$

This completes the proof of Theorem 22 for α nonintegral.

7.2.6. *Interpolation and the case α integral.* The case $\alpha = d \in \mathbb{Z}$ is proven by complex interpolation. This subsubsection follows Section 5 of [FO20a] closely.

For $1 < p < \infty$ and $0 < \alpha_0 < \alpha_1 < \infty$, the pair $(S^p_{\alpha_0}, S^p_{\alpha_1})$ of Banach spaces is an interpolation pair in the sense of Calderón, as they embed continuously in the space $\mathscr{S}'(G)$ of tempered distributions on G. Thus, we may define the complex interpolation space $[S^p_{\alpha_0}, S^p_{\alpha_1}]_\theta$, $\theta \in (0,1)$.

The following lemma asserts that we have a continuous embedding

$$S^p_{(1-\theta)\alpha_0 + \theta\alpha_1} \hookrightarrow [S^p_{\alpha_0}, S^p_{\alpha_1}]_\theta.$$

Lemma 53 ([FO20a, Lemma 5.1]). *Let $1 < p < \infty$, $0 < \alpha_0 < \alpha_1 < \infty$, and $\theta \in (0,1)$. Then*

$$\|f\|_{[S^p_{\alpha_0}, S^p_{\alpha_1}]_\theta} \lesssim_{G,\alpha_0,\alpha_1,p} \|f\|_{p,(1-\theta)\alpha_0 + \theta\alpha_1}.$$

The proof is exactly the same as in [FO20a, Lemma 5.1].

Next, we define, for $\alpha > 0$, the Banach space H_α of functions $F : B_1 \times (0,\infty) \to \mathbb{R}$ with norm

$$\|F\|_{H_\alpha} := \left(\int_0^\infty \left[\frac{1}{r^\alpha} \fint_{B_1} |F(y,r)| dy \right]^2 \frac{dr}{r} \right)^{1/2} < \infty.$$

Then, for $1 < p < \infty$, we may define the function space $L^p(G, H_\alpha)$. Now, for $0 < \alpha_0 < \alpha_1 < \infty$,

$$(L^p(G, H_{\alpha_0}), L^p(G, H_{\alpha_1}))$$

is a compatible couple, as both embed continuously into $L^1_{\text{loc}}(G \times B_1 \times (0,\infty))$. As in [Dor85, FO20a], we apply the results of [BL12, p. 107, 121] so that for $\theta \in (0,1)$,

$$[L^p(G, H_{\alpha_0}), L^p(G, H_{\alpha_1})]_\theta = L^p(G, [H_{\alpha_0}, H_{\alpha_1}]_\theta) = L^p(G, H_{(1-\theta)\alpha_0 + \theta\alpha_1}).$$

Now, for $d-1 < \alpha_0 < d < \alpha_1 < d+1$ (recall $d = \alpha$) and $\theta \in (0,1)$, consider the linear map $T : S^p_{\alpha_0} + S^p_{\alpha_1} \to L^p(G, H_{\alpha_0}) + L^p(G, H_{\alpha_1})$ given by

$$f \mapsto Tf(x,r,y) = f(x\delta_r(y)) - A^d_{x,r}(x\delta_r(y)).$$

46

By Lemma 46 and Theorem 22 for $\alpha \in (d-1,d) \cup (d, d+1)$, we have

$$\|Tf\|_{L^p(G,H_\alpha)} \lesssim_{G,d} \|\mathfrak{G}_\alpha f\|_{L^p(G)} \lesssim_{G,\alpha_0,\alpha_1,p} \|f\|_{p,\alpha}, \quad f \in S_\alpha^p, \ \alpha = \alpha_0, \alpha_1,$$

i.e., $T : S_{\alpha_0}^p \to L^p(G, H_{\alpha_0})$ and $T : S_{\alpha_1}^p \to L^p(G, H_{\alpha_1})$ are bounded. By complex interpolation,

$$T : [S_{\alpha_0}^p, S_{\alpha_1}^p]_\theta \to [L^p(G, H_{\alpha_0}), L^p(G, H_{\alpha_1})]_\theta = L^p(G, H_{(1-\theta)\alpha_0 + \theta\alpha_1})$$

is bounded.

Now set $\alpha_0 = d - 1/2$, $\alpha_1 = d + 1/2$, and $\theta = 1/2$. Then

$$\|\mathfrak{G}_d f\|_{L^p(G)} = \|Tf\|_{L^p(G,H_d)} \lesssim_{G,d} \|f\|_{[S_{d-1/2}^p, S_{d+1/2}^p]_{1/2}} \overset{\text{Lemma 53}}{\lesssim_{G,d,p}} \|f\|_{p,d}, \quad f \in \mathcal{D}.$$

This completes the proof of Theorem 22 for $\alpha = d \in \mathbb{Z}$.

7.3. **Proof of the $\gtrsim$ direction.**

Now, given a $f \in L^p(G)$ and $\alpha > 0$ with $\mathfrak{G}_\alpha f \in L^p(G)$, we need to prove that $f \in S_\alpha^p(G)$ and that the $\gtrsim$ direction of (8) holds.

We first prove the following preparatory result.

Proposition 54. *Let $f \in L^p(G)$ and $\alpha > 0$ with $\mathfrak{G}_\alpha f \in L^p(G)$. Denote*

$$A_{x,r}^{\lfloor \alpha \rfloor} f(xy) = \sum_{|\gamma| \le \lfloor \alpha \rfloor} f_\gamma(x,r) y^\gamma,$$

Then

(1) For each multi-index γ with $|\gamma| < \alpha$, $|f_\gamma(x,r) - f_\gamma(x,s)| \lesssim_{G,\alpha} r^{\alpha - |\gamma|} \mathfrak{G}_\alpha f(x)$ for $x \in G$ and $0 < s < r$, and $|f_\gamma(x,r)| \lesssim_{G,\alpha} \mathfrak{G}_\alpha f(x) + Mf(x)$ for $x \in G$ and $0 < r < 1$, where M denotes the Hardy–Littlewood maximal operator. In particular, $f_\gamma(\cdot, r) \in L^p(G)$ with $\|f_\gamma(\cdot, r)\|_{L^p(G)} \lesssim_{G,\alpha} \|\mathfrak{G}_\alpha f\|_{L^p(G)} + \|f\|_{L^p(G)}$.

(2) For each multi-index γ with $|\gamma| < \alpha$, there exists $f_\gamma \in L^p(G)$ such that $f_\gamma(\cdot, r) \to f_\gamma(\cdot)$ as $r \to 0$, satisfying

$$|f_\gamma(x,r) - f_\gamma(x)| \lesssim_{G,|\gamma|} \int_0^{2r} \beta_{f,d}(B_u(x)) u^{-|\gamma|} \frac{du}{u} \lesssim_{G,\alpha} r^{\alpha - |\gamma|} \mathfrak{G}_\alpha f(x)$$

for a.e. $x \in G$ and $r > 0$.

(3) If α is nonintegral, then

$$\left(\int_0^\infty \left(\frac{1}{r^\alpha} \fint_{B_r} \Big| f(xy) - \sum_{|\gamma| \le \lfloor \alpha \rfloor} f_\gamma(x) y^\gamma \Big| dy \right)^2 \frac{dr}{r} \right)^{1/2} \asymp_{G,\alpha} \mathfrak{G}_\alpha f(x), \quad \text{a.e. } x \in G,$$

(4) If α is integral and $|\gamma| = \alpha$, then $|f_\gamma(x,r)| \lesssim_{G,\alpha} \log(2/r) \mathfrak{G}_\alpha f(x) + Mf(x)$, $0 < r < 1$. Thus, $f_\gamma(\cdot, r) \in L^p(G)$ with $\|f_\gamma(\cdot, r)\|_{L^p(G)} \lesssim_{G,\alpha} \log(2/r) \|\mathfrak{G}_\alpha f\|_{L^p(G)} + \|f\|_{L^p(G)}$.

(5) If $0 < \alpha < 1$, then $f_0(x) = f(x)$ for a.e. $x \in G$.

Proof. (1) By repeating the computation (57) made in the proof of Proposition 48, it follows that

$$|f_\gamma(x,r) - f_\gamma(x,s)| \lesssim_{G,|\gamma|} \int_s^{2r} \beta_{f,d}(B_u(x)) u^{-|\gamma|} \frac{du}{u}, \quad x \in G, \ 0 < s < r.$$

In particular, by Cauchy–Schwartz,

$$|f_\gamma(x,r) - f_\gamma(x,s)| \lesssim_{G,|\gamma|} \left(\int_s^{2r} \left[\frac{\beta_{f,d}(B_u(x))}{u^\alpha} \right]^2 \frac{du}{u} \right)^{1/2} \left(\int_s^{2r} (u^{\alpha - |\gamma|})^2 \frac{du}{u} \right)^{1/2} \lesssim_\alpha r^{\alpha - |\gamma|} \mathfrak{G}_\alpha f(x), \quad x \in G, \ 0 < s < r.$$

Also, for $0 < r < 1$, $|f_\gamma(x,r)| \le |f_\gamma(x,r) - f_\gamma(x,1)| + |f_\gamma(x,1)| \lesssim_{G,\alpha} \mathfrak{G}_\alpha f(x) + Mf(x)$, where the latter inequality used the fact that $f_\gamma(x,1)$ is given as the integral of f multiplied by a polynomial determined by γ and $\lfloor \alpha \rfloor$. In particular, this implies that $f_\gamma(\cdot, r) \in L^p$.

(2) By (1) and the fact that $\mathfrak{G}_\alpha f \in L^p(G)$, $\{f_\gamma(\cdot, r)\}_{r>0}$ is a Cauchy sequence in L^p as $r \to 0$. Thus, there exists $f_\gamma(\cdot) \in L^p(G)$ so that $f_\gamma(\cdot, r) \to f_\gamma(\cdot)$ in $L^p(G)$ as $r \to 0$.

47

(3) The $\gtrsim$ follows from Lemma 47. For $\lesssim$, we bound

$$\fint_{B_r} |f(xy) - \sum_{|\gamma| \le \lfloor \alpha \rfloor} f_\gamma(x) y^\gamma| \le \fint_{B_r(x)} |f - A_{x,r}^{\lfloor \alpha \rfloor} f| + \fint_{B_r} |\sum_{|\gamma| \le \lfloor \alpha \rfloor} f_\gamma(x) y^\gamma - A_{x,r}^{\lfloor \alpha \rfloor} f(xy)| dy$$

$$\lesssim_{G,\alpha} \beta_{f,\lfloor \alpha \rfloor}(B_r(x)) + \sum_{|\gamma| \le \lfloor \alpha \rfloor} |f_\gamma(x,r) - f_\gamma(x)| r^{|\gamma|}$$

$$\lesssim_{G,\alpha} \beta_{f,\lfloor \alpha \rfloor}(B_r(x)) + \sum_{i=0}^{\lfloor \alpha \rfloor} r^i \int_0^{2r} \beta_{f,\lfloor \alpha \rfloor}(B_u(x)) u^{-i} \frac{du}{u}.$$

By plugging into the definition of $\tilde{\mathfrak{G}}$ and using the triangle inequality,

$$\left(\int_0^\infty \left(\frac{1}{r^\alpha} \fint_{B_r} |f(xy) - \sum_{|\gamma| \le \lfloor \alpha \rfloor} f_\gamma(x) y^\gamma| dy \right)^2 \frac{dr}{r} \right)^{1/2}$$

$$\lesssim_{G,\alpha} \mathfrak{G}_\alpha f(x) + \sum_{i=0}^{\lfloor \alpha \rfloor} \left(\int_0^\infty \left[r^{i-\alpha} \int_0^{2r} \beta_{f,\lfloor \alpha \rfloor}(B_u(x)) u^{-i} \frac{du}{u} \right]^2 \frac{dr}{r} \right)^{1/2}$$

$$\lesssim_{G,\alpha} \mathfrak{G}_\alpha f(x) + \sum_{i=0}^{\lfloor \alpha \rfloor} \frac{1}{\alpha - i} \left(\int_0^\infty \left[\frac{\beta_{f,\lfloor \alpha \rfloor}(B_r(x))}{r^\alpha} \right]^2 \frac{dr}{r} \right)^{1/2}$$

$$\lesssim_\alpha \mathfrak{G}_\alpha f(x).$$

where in the second inequality we used Hardy's inequality (23) with $v = \alpha - i + \frac{1}{2}$ and $p = 2$ along with the fact that α is nonintegral.

(4) By Cauchy–Schwartz,

$$|f_\gamma(x,r) - f_\gamma(x,1)| \lesssim_{G,|\gamma|} \left(\int_r^2 \left[\frac{\beta_{f,d}(B_u(x))}{u^\alpha} \right]^2 \frac{du}{u} \right)^{1/2} \left(\int_r^2 \frac{du}{u} \right)^{1/2} \lesssim_\alpha \log(2/r) \mathfrak{G}_\alpha f(x).$$

(5) If $0 < \alpha < 1$, then $f_0(x,r) = \fint_{B_r(x)} f$. Taking $r \to 0$, the claim follows by the Lebesgue differentiation theorem for doubling metric measure spaces [HKST15, (3.4.10)]. $\qquad \square$

Remark 55. *It seems likely that for $|\gamma| < \alpha$ we have $f_\gamma = \frac{1}{\gamma!} \mathrm{Sym}(X^\gamma) f$, with the latter derivative existing in the distributive sense. This is evident when $0 < \alpha < 1$ and $\gamma = 0$, and we will prove this holds when $\alpha > 1$ and $|\gamma| = 1$ in Lemma 57. We will not prove this statement in full generality since we do not need it in this paper.*

7.3.1. *The case $0 < \alpha \le 1$.* The case $0 < \alpha < 1$ is a repetition and strengthening of subsubsection 7.2.4. If we denote

$$\tilde{\mathfrak{G}}_\alpha f(x) = \left(\int_0^\infty \left[\frac{1}{r^\alpha} \fint_{B_r} |f(xy) - f(x)| dy \right]^2 \frac{dr}{r} \right)^{1/2}, \quad x \in G,$$

then by Proposition 54 (3) and (5)

$$\mathfrak{G}_\alpha f(x) \asymp_{G,\alpha} \tilde{\mathfrak{G}}_\alpha f(x), \quad \text{for a.e. } x \in G.$$

But by [CRTN01, Theorem 5],

$$\|\tilde{\mathfrak{G}}_\alpha f\|_{L^p(G)} \asymp_{G,p} \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}.$$

For $\alpha = 1$, it is known from [DNM21, Theorem 1.4] that for $0 < \alpha < 2$, if we define

$$S_\alpha f(x) = \left(\int_0^\infty \left[r^{-\alpha} \fint_{B_r(0)} |\Delta_y^{2,\mathrm{sym}} f(x)| dy \right]^2 \frac{dr}{r} \right)^{1/2},$$

48

where we denote the symmetric difference

$$\Delta_y^{2,\text{sym}} f(x) = f(xy) + f(xy^{-1}) - 2f(x),$$

then

$$\|f\|_{p,\alpha} \asymp_{G,p,\alpha} \|f\|_{L^p(G)} + \|S_\alpha f\|_{L^p(G)}.$$

By the scaling argument of subsubsection 7.2.3,

$$\|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)} \asymp_{G,p,\alpha} \|S_\alpha f\|_{L^p(G)}.$$

Let f be a locally integrable function, $d \in \mathbb{N}$, $x = z_0 \in G$ and $r > 0$. For any $y \in B_r(z_0)$ and ball $B_{2^{-k}r}(z_k)$ with $y \in B_{2^{-k}r}(z_k) \subset B_r(z_0)$, let $B_{2^{-k}r}(z_k) \subset B_{2^{-k+1}r}(z_{k-1}) \subset \cdots \subset B_{2^{-1}r}(z_1) \subset B_r(x)$ be a sequence of balls. We have

$$\fint_{B_{2^{-k}r}(z_k)} |f - A_{z_0,r}^d f| \leq \fint_{B_{2^{-k}r}(z_k)} |f - A_{z_k,2^{-k}r}^d f| + \sum_{i=0}^{k-1} \fint_{B_{2^{-k}r}(z_k)} |A_{z_i,2^{-i}r}^d f - A_{z_{i+1},2^{-i-1}r}^d f|$$

$$\leq \fint_{B_{2^{-k}r}(z_k)} |f - A_{z_k,2^{-k}r}^d f| + \sum_{i=0}^{k-1} \left\| A_{z_{i+1},2^{-i-1}r}^d (f - A_{z_i,2^{-i}r}^d f) \right\|_{L^\infty(B_{2^{-i-1}r}(z_{i+1}))}$$

$$\overset{(55)}{\lesssim_{G,d}} \fint_{B_{2^{-k}r}(z_k)} |f - A_{z_k,2^{-k}r}^d f| + \sum_{i=0}^{k-1} \fint_{B_{2^{-i-1}r}(z_{i+1})} |f - A_{z_i,2^{-i}r}^d f|$$

$$\lesssim_G \sum_{i=0}^{k} \fint_{B_{2^{-i}r}(z_i)} |f - A_{z_i,2^{-i}r}^d f| \overset{\text{Corollary 47}}{\lesssim_{G,d}} \int_0^{2r} \beta_{f,d}(B_s(y)) \frac{ds}{s}$$

so by Lebesgue's differentiation theorem for doubling metric measure spaces, for a.e. $y \in B_r(x)$,

$$|f(y) - A_{x,r}^d f(y)| = \lim_{B_{2^{-k}r}(z_k) \to y} \fint_{B_{2^{-k}r}(z_k)} |f - A_{z_0,r}^d f| \lesssim_{G,d} \int_0^{2r} \beta_{f,d}(B_s(y)) \frac{ds}{s}.$$

As the operator $\Delta_h^{2,\text{sym}}$ annihilates polynomials of weighted degree 1, we obtain for a.e. $x \in G$

$$\fint_{B_r} |\Delta_h^{2,\text{sym}} f(x)| dh = \fint_{B_r} |\Delta_h^{2,\text{sym}} (f - A_{x,2r}^1 f)(x)| dh$$

$$= \fint_{B_r} |(f - A_{x,2r}^1 f)(xh) + (f - A_{x,2r}^1 f)(xh^{-1}) - 2(f - A_{x,2r}^1 f)(x)| dh$$

$$\lesssim_{G,d} \fint_{B_r} \left(\int_0^{2r} \beta_{f,d}(B_s(xh)) \frac{ds}{s} + \int_0^{2r} \beta_{f,d}(B_s(xh^{-1})) \frac{ds}{s} + 2 \int_0^{2r} \beta_{f,d}(B_s(x)) \frac{ds}{s} \right) dh$$

$$\lesssim \int_0^{2r} M\beta_{f,1}(B_s(\cdot))(x) \frac{ds}{s}.$$

By Hardy's inequality (23) with $\nu = \frac{3}{2}$ and $p = 2$, we obtain

$$S_1 f(x) = \left(\int_0^\infty \left[r^{-1} \fint_{B_r} |\Delta_y^{2,\text{sym}} f(x)| dy \right]^2 \frac{dr}{r} \right)^{1/2} \lesssim_{G,d} \left(\int_0^\infty [M\beta_{f,1}(B_r(\cdot))(x)]^2 \frac{dr}{r} \right)^{1/2}.$$

Now we argue that

$$\left(\int_G \left(\int_0^\infty [M\beta_{f,1}(B_r(\cdot))(x)]^2 \frac{dr}{r} \right)^{p/2} dx \right)^{1/p} \lesssim_G \max\{p, \frac{1}{p-1}\} \left(\int_G \left(\int_0^\infty [\beta_{f,1}(B_r(\cdot))(x)]^2 \frac{dr}{r} \right)^{p/2} dx \right)^{1/p}$$

This is an instance of the Fefferman–Stein vector-valued maximal function inequality

$$\left(\int_G \left[\sum_{j \in \mathbb{Z}} M(g_j)^u(x) \right]^{v/u} dx \right)^{1/v} \lesssim_G \max\{1, \frac{1}{u-1}\} \max\{v, \frac{1}{v-1}\} \left(\int_G \left[\sum_{j \in \mathbb{Z}} g_j(x)^u \right]^{v/u} dx \right)^{1/v} \tag{61}$$

49

on n_h-regular metric measure spaces for measurable functions $(g_j : G \to \mathbb{R})_{j \in \mathbb{Z}}$ and $u, v \in (1, \infty)$, due to [GLY09, Theorem 1.2]. Indeed, we apply (61) to the functions

$$g_j(x) = \beta_{f,1}(B_{2^{-j}}(x))$$

and exponents $u = 2 > 1$ and $v = p > 1$, noting by Corollary 47 that

$$M(\beta_{f,1}(B_r(\cdot))) \lesssim_G M(g_j), \quad r \in [2^{-j-1}, 2^{-j}],$$

and

$$g_j \lesssim_G \beta_{f,1}(B_r(\cdot)), \quad r \in [2^{-j}, 2^{-j+1}],$$

Thus, we have $\|S_1 f\|_{L^p(G)} \lesssim_{G,p} \|\mathfrak{G}_1 f\|_{L^p(G)}$, $1 < p < \infty$, as desired.

7.3.2. *The case $\alpha > 1$.* We use induction on α.

The following lemma is an analogue of [Dor85, Lemma 1], and the proof proceeds in the same manner.

Lemma 56 (Analogue of [Dor85, Lemma 1]). *Let $f \in L^p(G)$ such that $\mathfrak{G}_\alpha f \in L^p(G)$, and γ be a multi-index with $|\gamma| < \alpha$. With f_γ as in Proposition 54, $\|\mathfrak{G}_{\alpha - |\gamma|} f_\gamma\|_{L^p(G)} \lesssim_{G,\alpha} \|\mathfrak{G}_\alpha f\|_{L^p(G)}$.*

Proof. Fix $x \in G$ and $r > 0$. For $y \in B_r(x)$, by Proposition 54 and (56),

$$|f_\gamma(y) - \frac{1}{\gamma!} \mathrm{Sym}(X^\gamma) A^d_{x,2r} f(y)| \le |f_\gamma(y) - f_\gamma(y, r)| + |\frac{1}{\gamma!} \mathrm{Sym}(X^\gamma)(A^d_{y,r} - A^d_{x,2r} f)(y)|$$

$$\lesssim_{G,\alpha} \int_0^r \beta_{f,d}(B_s(y)) s^{-|\gamma|-1} ds + r^{-|\gamma|} \beta_{f,d}(B_{2r}(x)).$$

Thus, by Lemma 46,

$$\beta_{f_\gamma, d - |\gamma|}(B_r(x)) \lesssim_{G,\alpha} \fint_{B_r(x)} \left(\int_0^r \beta_{f,d}(B_s(y)) s^{-|\gamma|-1} ds dy + r^{-|\gamma|} \beta_{f,d}(B_{2r}(x)) \right) dy$$

$$\le \int_0^r M(\beta_{f,d}(B_s(\cdot)))(x) s^{-|\gamma|-1} ds + r^{-|\gamma|} \beta_{f,d}(B_{2r}(x)).$$

By Hardy's inequality with $\nu = \alpha - |\gamma| + \frac{1}{2}$ and $p = 2$,

$$G_{\alpha - |\gamma|} f_\gamma(x) \lesssim_{G,\alpha} \mathfrak{G}_\alpha f(x) + \left(\int_0^\infty \left[r^{-\alpha} M \beta_{f,d}(B_r(\cdot))(x) \right]^2 \frac{dr}{r} \right)^{1/2},$$

and the lemma follows by the aforementioned Fefferman–Stein inequality (61) on vector-valued maximal operators with $u = v = 2 > 1$ and

$$g_j(x) = 2^{\alpha j} \beta_{f,d}(B_{2^{-j+1}}(x)).$$

$\square$

The following lemma is an analogue of [Dor85, Lemma 2], and the proof proceeds in the same manner.

Lemma 57 (Analogue of [Dor85, Lemma 2]). *Let $f \in L^p(G)$ with $\mathfrak{G}_\alpha f \in L^p(G)$, $\alpha > 1$. Then the weak partials $X_i f$ exist and coincide with $f_{(1,i)}$ a.e., where $(1,i)$ stands for the multi-index γ with $\gamma_{r,j} = \delta_{(r,j),(1,i)}$.*

Proof. Denote $e_{r,i} = \exp(x_{r,i})$ for $r = 1, \cdots, s$ and $i = 1, \cdots, k_r$. It is enough to show that $(f(xe^r_{1,i}) - f(x))/r$ tends to $f_{(1,i)}$ in L^p as $r \to 0$. First,

$$\left| \frac{f(xe^r_{1,i}) - f(x)}{r} - f_{(1,i)}(x) \right|$$

$$\le \frac{|f(xe^r_{1,i}) - f_0(xe^r_{1,i}, 2r)|}{r} + \frac{|f(x) - f_0(x, 2r)|}{r} + \left| \frac{f_0(xe^r_{1,i}, 2r) - f_0(x, 2r)}{r} - f_{(1,i)}(x) \right|$$

$$\overset{\text{Proposition 54(2)}}{\lesssim_{G,\alpha}} r^{\alpha-1} \left(\mathfrak{G}_\alpha f(xe^r_{1,i}) + \mathfrak{G}_\alpha f(x) \right) + \left| \frac{f_0(xe^r_{1,i}, 2r) - f_0(x, 2r)}{r} - f_{(1,i)}(x) \right|.$$

50

Next, since $f_0(y, r) = A^{\lfloor \alpha \rfloor}_{y,r} f(y)$,

$$f_0(xe^r_{1,i}, 2r) - f_0(x, 2r) = A^{\lfloor \alpha \rfloor}_{xe^r_{1,i}, 2r} f(xe^r_{1,i}) - A^{\lfloor \alpha \rfloor}_{x,2r} f(xe^r_{1,i}) + A^{\lfloor \alpha \rfloor}_{x,2r} f(xe^r_{1,i}) - A^{\lfloor \alpha \rfloor}_{x,2r} f(x)$$

$$= A^{\lfloor \alpha \rfloor}_{xe^r_{1,i}, 2r} f(xe^r_{1,i}) - A^{\lfloor \alpha \rfloor}_{x,2r} f(xe^r_{1,i}) + \sum_{j=1}^{\lfloor \alpha \rfloor} f_{j(1,i)}(x, 2r) r^j / j!.$$

Thus

$$\left| \frac{f_0(xe^r_{1,i}, 2r) - f_0(x, 2r)}{r} - f_{(1,i)}(x) \right|$$

$$\leq \left| \frac{f_0(xe^r_{1,i}, 2r) - f_0(x, 2r)}{r} - f_{(1,i)}(x, 2r) \right| + |f_{(1,i)}(x, 2r) - f_{(1,i)}(x)|$$

$$\overset{\text{Proposition } 54(2)}{\lesssim_{G,\alpha}} \frac{|A^{\lfloor \alpha \rfloor}_{xe^r_{1,i}, 2r} f(xe^r_{1,i}) - A^{\lfloor \alpha \rfloor}_{x,2r} f(xe^r_{1,i})|}{r} + \sum_{j=2}^{\lfloor \alpha \rfloor} |f_{j(1,i)}(x, 2r)| r^{j-1} + r^{\alpha-1} \mathfrak{G}_\alpha f(x)$$

$$\lesssim_{G,\alpha} r^{-1} \fint_{B_r(xe^r_{1,i})} |A^{\lfloor \alpha \rfloor}_{xe^r_{1,i}, 2r} f(y) - A^{\lfloor \alpha \rfloor}_{x,2r} f(y)| dy + \sum_{j=2}^{\lfloor \alpha \rfloor} |f_{j(1,i)}(x, 2r)| r^{j-1} + r^{\alpha-1} \mathfrak{G}_\alpha f(x)$$

$$\leq r^{-1} \fint_{B_r(xe^r_{1,i})} |f - A^{\lfloor \alpha \rfloor}_{xe^r_{1,i}, 2r} f(y)| dy + r^{-1} \fint_{B_r(xe^r_{1,i})} |f - A^{\lfloor \alpha \rfloor}_{x,2r} f(y)| dy + \sum_{j=2}^{\lfloor \alpha \rfloor} |f_{j(1,i)}(x, 2r)| r^{j-1} + r^{\alpha-1} \mathfrak{G}_\alpha f(x)$$

$$= r^{-1} \beta_{f, \lfloor \alpha \rfloor}(B_{2r}(xe^r_{1,i})) + r^{-1} \beta_{f, \lfloor \alpha \rfloor}(B_{2r}(x)) + \sum_{j=2}^{\lfloor \alpha \rfloor} |f_{j(1,i)}(x, 2r)| r^{j-1} + r^{\alpha-1} \mathfrak{G}_\alpha f(x),$$

where in the penultimate inequality we used the fact that the $L^1(B_1)$ and $L^\infty(B_1)$ norms on the space of polynomials of weighted degree $\leq \lfloor \alpha \rfloor$ are equivalent. Taking L^p norms,

$$\left\| \frac{f(\cdot e^r_{1,i}) - f(\cdot)}{r} - f_{(1,i)}(\cdot) \right\|_{L^p(G)} \lesssim_{G,\alpha} r^{-1} \| \beta_{f, \lfloor \alpha \rfloor}(B_{2r}(\cdot)) \|_{L^p(G)} + \sum_{j=2}^{\lfloor \alpha \rfloor} r^{j-1} \| f_{j(1,i)}(\cdot, 2r) \|_{L^p(G)} + r^{\alpha-1} \| \mathfrak{G}_\alpha f \|_{L^p(G)}.$$

By Corollary 47, $\| \beta_{f, \lfloor \alpha \rfloor}(B_{2r}(\cdot)) \|_{L^p(G)} \lesssim_{G,\alpha} r^\alpha \| \mathfrak{G}_\alpha f \|_{L^p(G)}$. By Proposition 54, if $j < \lfloor \alpha \rfloor$, then

$$\| f_{j(1,i)}(\cdot, 2r) \|_{L^p(G)} \lesssim_{G,\alpha} \| \mathfrak{G}_\alpha f \|_{L^p(G)} + \| f \|_{L^p(G)},$$

and if $j = \lfloor \alpha \rfloor = \alpha$, then

$$\| f_{j(1,i)}(\cdot, 2r) \|_{L^p(G)} \lesssim_{G,\alpha} (\log r^{-1})(\| \mathfrak{G}_\alpha f \|_{L^p(G)} + \| f \|_{L^p(G)}).$$

Thus $\left\| \frac{f(\cdot e^r_{1,i}) - f(\cdot)}{r} - f_{(1,i)}(\cdot) \right\|_{L^p(G)} \to 0$ as $r \to 0$. $\qquad \square$

The following lemma is an analogue of [Dor85, Lemma 3], and the proof proceeds in the same manner.

Lemma 58 (Analogue of [Dor85, Lemma 3]). *Let $f \in L^p(G)$ with $\mathfrak{G}_\alpha f \in L^p(G)$, $\alpha > 1$. Then $\mathfrak{G}_{\alpha-1} f \in L^p(G)$ with*

$$\| \mathfrak{G}_{\alpha-1} f \|_{L^p(G)} \lesssim_{G,\alpha} \| \mathfrak{G}_\alpha f \|_{L^p(G)} + \| f \|_{L^p(G)}.$$

Proof. Write

$$A^{\lfloor \alpha \rfloor}_{x,r} f(xy) = \underbrace{\sum_{|\gamma| < \lfloor \alpha \rfloor} f_\gamma(x, r) y^\gamma}_{=: A^{< \lfloor \alpha \rfloor}_{x,r} f(xy)} + \underbrace{\sum_{|\gamma| = \lfloor \alpha \rfloor} f_\gamma(x, r) y^\gamma}_{=: A^{= \lfloor \alpha \rfloor}_{x,r} f(xy)}.$$

Then by Lemma 46,

$$\beta_{f, \lfloor \alpha \rfloor - 1}(B_r(x)) \lesssim_{G,\alpha} \fint_{B_r(x)} |f - A^{< \lfloor \alpha \rfloor}_{x,r} f| \leq \beta_{f, \lfloor \alpha \rfloor}(B_r(x)) + \| A^{= \lfloor \alpha \rfloor}_{x,r} f \|_{L^\infty(B_r(x))}. \qquad (62)$$

By (56), for $i \in \mathbb{Z}_{\geq 0}$ and $|\gamma| = \lfloor \alpha \rfloor$,

$$|f_\gamma(x, 2^i r)| \lesssim_{G,\alpha} (2^i r)^{-\lfloor \alpha \rfloor} \fint_{B_{2^i r}(x)} |f| \lesssim_G (2^i r)^{-\lfloor \alpha \rfloor - n_h/p} \|f\|_{L^p(G)} \to 0 \quad \text{as } i \to \infty$$

and thus

$$|f_\gamma(x, r)| \leq \sum_{i=0}^\infty |f_\gamma(x, 2^i r) - f_\gamma(x, 2^{i+1} r)| \overset{(56)}{\lesssim}_{G,\alpha} \sum_{i=0}^\infty (2^i r)^{-\lfloor \alpha \rfloor} \|A_{x,2^i r}^{\lfloor \alpha \rfloor} f - A_{x,2^{i+1} r}^{\lfloor \alpha \rfloor} f\|_{L^\infty(B_{2^i r}(x))}$$

$$\overset{(55)}{\lesssim}_{G,\alpha} \sum_{i=0}^\infty (2^i r)^{-\lfloor \alpha \rfloor} \fint_{B_{2^i r}(x)} |f - A_{x,2^{i+1} r}^{\lfloor \alpha \rfloor} f| \lesssim_{G,\alpha} \sum_{i=0}^\infty (2^i r)^{-\lfloor \alpha \rfloor} \beta_{f,\lfloor \alpha \rfloor}(B_{2^{i+1} r}(x)).$$

$$\overset{\text{Corollary } 47}{\lesssim}_{G,\alpha} \int_r^\infty \beta_{f,\lfloor \alpha \rfloor}(B_\rho(x)) \rho^{-\lfloor \alpha \rfloor - 1} d\rho.$$

Plugging these results into (62), we have

$$\beta_{f,\lfloor \alpha \rfloor - 1}(B_r(x)) \lesssim_{G,\alpha} \beta_{f,\lfloor \alpha \rfloor}(B_r(x)) + r^{\lfloor \alpha \rfloor} \int_r^\infty \beta_{f,\lfloor \alpha \rfloor}(B_\rho(x)) \rho^{-\lfloor \alpha \rfloor - 1} d\rho,$$

and

$$\mathfrak{G}_{\alpha-1} f(x) = \left(\int_0^\infty \left[\frac{\beta_{f,\lfloor \alpha \rfloor - 1}(B_r(x))}{r^{\alpha-1}} \right]^2 \frac{dr}{r} \right)^{1/2}$$

$$\lesssim_{G,\alpha} \left(\int_0^\infty \left[\frac{\beta_{f,\lfloor \alpha \rfloor}(B_r(x))}{r^{\alpha-1}} \right]^2 \frac{dr}{r} \right)^{1/2} + \left(\int_0^\infty \left[r^{\lfloor \alpha \rfloor - \alpha + 1/2} \int_r^\infty \beta_{f,\lfloor \alpha \rfloor}(B_\rho(x)) \rho^{-\lfloor \alpha \rfloor - 1} d\rho \right]^2 dr \right)^{1/2}$$

$$\lesssim \left(\int_0^\infty \left[\frac{\beta_{f,\lfloor \alpha \rfloor}(B_r(x))}{r^{\alpha-1}} \right]^2 \frac{dr}{r} \right)^{1/2}$$

$$\leq \left(\int_0^1 \left[\frac{\beta_{f,\lfloor \alpha \rfloor}(B_r(x))}{r^{\alpha-1}} \right]^2 \frac{dr}{r} \right)^{1/2} + \left(\int_1^\infty \left[\frac{\beta_{f,\lfloor \alpha \rfloor}(B_r(x))}{r^{\alpha-1}} \right]^2 \frac{dr}{r} \right)^{1/2}$$

$$\overset{(55)}{\lesssim}_{G,\alpha} \left(\int_0^1 \left[\frac{\beta_{f,\lfloor \alpha \rfloor}(B_r(x))}{r^\alpha} \right]^2 \frac{dr}{r} \right)^{1/2} + \left(\int_1^\infty \left[\frac{Mf(x)}{r^{\alpha-1}} \right]^2 \frac{dr}{r} \right)^{1/2}$$

$$\lesssim_\alpha \mathfrak{G}_\alpha f(x) + Mf(x),$$

where the second inequality uses the second form of Hardy's inequality (25) with $p = 2$ and $\nu = \lfloor \alpha \rfloor - \alpha + 1/2$. This completes the proof. $\qquad \square$

To complete the induction, suppose the $\gtrsim$ direction of Theorem 8 holds for $\alpha - 1$. If $f \in L^p(G)$ with $\mathfrak{G}_\alpha f \in L^p(G)$, by Lemma 58, $\mathfrak{G}_{\alpha-1} f \in L^p(G)$, so that $f \in S_{\alpha-1}^p$ by induction. On the other hand, by Proposition 54 (2) and Lemmas 56 and 57, for $i = 1, \cdots, k$, $X_i f \in L^p(G)$ with $\mathfrak{G}_{\alpha-1}(X_i f) \in L^p(G)$, so again by induction $X_i f \in S_{\alpha-1}^p(G)$. Now, by Proposition 50, $f \in S_\alpha^p(G)$. Tracking the norms involved in this argument, we have that

$$\|\mathfrak{G}_\alpha f\|_{L^p(G)} \overset{\text{Lemma } 56}{\gtrsim}_{G,\alpha} \sum_{i=1}^k \|\mathfrak{G}_{\alpha-1} f_{(1,i)}\|_{L^p(G)} \overset{\text{Lemma } 57}{=} \sum_{i=1}^k \|\mathfrak{G}_{\alpha-1}(X_i f)\|_{L^p(G)}$$

$$\overset{\text{induction}}{\gtrsim}_{G,\alpha,p} \sum_{i=1}^k \|(-\Delta_p)^{(\alpha-1)/2} X_i f\|_{L^p(G)} \overset{\text{Proposition } 50}{\gtrsim}_{G,\alpha,p} \|(-\Delta_p)^{\alpha/2} f\|_{L^p(G)}.$$

This completes the $\gtrsim$ direction of Theorem 8.

7.4. L^q β-numbers, $q > 1$. The extension of Theorem 22 to β numbers defined in the L^q sense are discussed in [Dor85, Section 5] and [FO20a, Section 6]. We will follow their presentation.

By the discussion so far, it is enough to show that the left-hand side of (8) is maximized up to constants when $q = 1$.

Denote $d = \lfloor \alpha \rfloor$. We compute that for all $x \in G$, $r > 0$, $y \in B_r(x)$ and $n \in \mathbb{N}$ so that $B_{2^{-n}r}(y) \subset B_r(x)$, we have

$$
\fint_{B_{2^{-n}r}(y)} |f(z) - A_{x,r}^d f(z)| dz
$$

$$
\leq \fint_{B_{2^{-n}r}(y)} |f(z) - A_{y,2^{-n}r}^d f(z)| dz + \sum_{k=1}^{n+1} \fint_{B_{2^{-n}r}(y)} |A_{y,2^{k-1-n}r}^d f(z) - A_{y,2^{k-n}r}^d f(z)| dz
$$

$$
+ \fint_{B_{2^{-n}r}(y)} |A_{y,2r}^d f(z) - A_{x,r}^d f(z)| dz
$$

$$
= \beta_{f,d}(B_{2^{-n}r}(y)) + \sum_{k=1}^{n+1} \fint_{B_{2^{-n}r}(y)} |A_{y,2^{k-1-n}r}^d [f - A_{y,2^{k-n}r}^d f]| + \fint_{B_{2^{-n}r}(y)} |A_{x,r}^d [A_{y,2r}^d f - f]|
$$

$$
\leq \beta_{f,d}(B_{2^{-n}r}(y)) + \sum_{k=1}^{n+1} \left\| A_{y,2^{k-1-n}r}^d [f - A_{y,2^{k-n}r}^d f] \right\|_{L^\infty(B_{2^{k-1-n}r}(y))} + \left\| A_{x,r}^d [A_{y,2r}^d f - f] \right\|_{L^\infty(B_r(x))}
$$

$$
\overset{(55)}{\underset{G,d}{\lesssim}} \beta_{f,d}(B_{2^{-n}r}(y)) + \sum_{k=1}^{n+1} \fint_{B_{2^{k-1-n}r}(y)} |f - A_{y,2^{k-n}r}^d f| + \fint_{B_r(x)} |A_{y,2r}^d f - f|
$$

$$
\underset{G,d}{\lesssim} \beta_{f,d}(B_{2^{-n}r}(y)) + \sum_{k=1}^{n+1} \beta_{f,d}(B_{2^{k-n}r}(y)) + \beta_{f,d}(B_{2r}(y))
$$

$$
\overset{\text{Corollary 47}}{\underset{G,d}{\lesssim}} \int_{2^{-n}r}^{2^{-n+1}r} \beta_{f,d}(B_s(y)) \frac{ds}{s} + \sum_{k=1}^{n+1} \int_{2^{k-n}r}^{2^{k+1-n}r} \beta_{f,d}(B_s(y)) \frac{ds}{s} + \int_{2r}^{4r} \beta_{f,d}(B_s(y)) \frac{ds}{s}
$$

$$
\lesssim \int_{2^{-n}r}^{4r} \beta_{f,d}(B_s(y)) \frac{ds}{s}.
$$

Thus by Lebesgue's differentiation theorem for doubling metric measure spaces and Corollary 47, we infer that for $x \in G$ and $r > 0$,

$$
|f(y) - A_{x,r}^d f(y)| \underset{G,d}{\lesssim} \int_0^{4r} \beta_{f,d}(B_s(y)) \frac{ds}{s} \underset{G,d}{\lesssim} \int_0^{4r} \fint_{B_s(y)} \beta_{f,d}(B_{2s}(z)) dz \frac{ds}{s}, \quad \text{a.e. } y \in B_r(x). \tag{63}
$$

Now fix

$$
p > 1 \quad \text{and} \quad 1 \leq q < \frac{\min\{p,2\} n_h}{n_h - \min\{p,2\}}.
$$

Then, choose some $1 < w < \min\{p,2\}$ and $0 < \eta < \min\{n_h / w, \alpha\}$ such that

$$
q = \frac{w n_h}{n_h - \eta w}. \tag{64}
$$

Given this choice of $q, w,$ and η, [HLNT13, Theorem 4.1] tells us that in n_h-regular metric measure spaces, the following fractional Hardy–Littlewood maximal function is a bounded operator $L^w(G) \to L^q(G)$:

$$
M_\eta g(y) := \sup_{s>0} \left\{ s^\eta \fint_{B_s(y)} |g(z)| dz \right\}
$$

It follows that

$$r^{n_h/q}\beta_{f,d,q}(B_r(x)) \overset{(63)}{\underset{G,\alpha}{\lesssim}} \left(\fint_{B_r(x)} \left[\int_0^{4r} \fint_{B_s(y)} \beta_{f,d}(B_{2s}(z))\,dz \frac{ds}{s} \right]^q dy \right)^{1/q}$$

$$\leq \int_0^{4r} \left(\fint_{B_r(x)} \left[\fint_{B_s(y)} \beta_{f,d}(B_{2s}(z))\,dz \right]^q dy \right)^{1/q} \frac{ds}{s}$$

$$\leq \int_0^{4r} s^{-\eta} \left(\fint_{B_r(x)} M_\eta(\beta_{f,d}(B_{2s}(\cdot))\chi_{B_{5r}(x)}(\cdot))(y)^q\,dy \right)^{1/q} \frac{ds}{s}$$

$$\underset{p,q}{\lesssim} \int_0^{4r} s^{-\eta} \left(\fint_{B_{5r}(x)} \beta_{f,d}(B_{2s}(z))^w\,dz \right)^{1/w} \frac{ds}{s}$$

$$\underset{G}{\lesssim} \int_0^{4r} s^{-\eta} r^{n_h/w} \left[M(\beta_{f,d}(B_{2s}(\cdot))^w)(x) \right]^{1/w} \frac{ds}{s},$$

where in the second inequality we used Minkowski's inequality and in the fourth inequality we used the boundedness of $M_\eta : L^w(G) \to L^q(G)$. By (64), we have

$$\beta_{f,d,q}(B_r(x)) \underset{G,\alpha,p,q}{\lesssim} r^\eta \int_0^{4r} s^{-\eta-1} \left[M(\beta_{f,d}(B_{2s}(\cdot))^w)(x) \right]^{1/w}\,ds.$$

Next, noting that $\alpha > \eta$ and using Hardy's inequality (23) with $\nu = \alpha - \eta + 1/2$, we obtain

$$\left(\int_0^\infty \left[\frac{1}{r^\alpha} \beta_{f,d,q}(B_r(x)) \right]^2 \frac{dr}{r} \right)^{1/2} \underset{G,\alpha,p,q}{\lesssim} \left(\int_0^\infty r^{2\eta-2\alpha-1} \left[\int_0^{4r} s^{-\eta-1} \left[M(\beta_{f,d}(B_{2s}(\cdot))^w)(x) \right]^{1/w}\,ds \right]^2 dr \right)^{1/2}$$

$$\underset{\alpha,q}{\lesssim} \left(\int_0^\infty \left[M(r^{-w\alpha}\beta_{f,d}(B_{2r}(\cdot))^w)(x) \right]^{2/w} \frac{dr}{r} \right)^{1/2}.$$

We conclude that

$$\left(\int_G \left(\int_0^\infty \left[\frac{1}{r^\alpha} \beta_{f,d,q}(B_r(x)) \right]^2 \frac{dr}{r} \right)^{p/2} dx \right)^{1/p} \underset{G,\alpha,p,q}{\lesssim} \left(\int_G \left(\int_0^\infty \left[M(r^{-w\alpha}\beta_{f,d}(B_{2r}(\cdot))^w)(x) \right]^{2/w} \frac{dr}{r} \right)^{p/2} dx \right)^{1/p}$$

$$\underset{G,\alpha,p,q}{\lesssim} \left(\int_G \left(\int_0^\infty \left[\frac{1}{r^\alpha} \beta_{f,d}(B_r(x)) \right]^2 \frac{dr}{r} \right)^{p/2} dx \right)^{1/p},$$

where in the second inequality we applied the Fefferman–Stein vector-valued maximal function inequality (61) with functions

$$g_j(x) = 2^{\alpha w j} \beta_{f,d}(B_{2^{-j+1}}(x))$$

and exponents $u = 2/w > 1$ and $v = p/w > 1$, along with Corollary 47.

This completes the proof of Theorem 22.

References

[Amb01] L. Ambrosio. Some fine properties of sets of finite perimeter in Ahlfors regular metric measure spaces. *Advances in Mathematics*, 159(1):51–67, 2001.

[ANT13] T. Austin, A. Naor, and R. Tessera. Sharp quantitative nonembeddability of the Heisenberg group into superreflexive Banach spaces. *Groups, Geometry, and Dynamics*, 7(3):497–522, 2013.

[Ass83] P. Assouad. Plongements lipschitziens dans $\mathbb{R}^n$. *Bulletin de la Société Mathématique de France*, 111:429–448, 1983.

[Azz16] J. Azzam. Bi-Lipschitz parts of quasisymmetric mappings. *Revista Matemática Iberoamericana*, 32(2):589–648, 2016.

[BCL94] K. Ball, E. A. Carlen, and E. H. Lieb. Sharp uniform convexity and smoothness inequalities for trace norms. *Inventiones mathematicae*, 115(1):463–482, 1994.

[BL12] J. Bergh and J. Löfström. *Interpolation spaces: an introduction*, volume 223. Springer Science & Business Media, 2012.

[BLD12]   E. Breuillard and E. Le Donne. Nilpotent groups, asymptotic cones and subfinsler geometry. *Preprint*, 2012.

[Bog]   V. I. Bogachev. *Measure theory*, volume 2. Springer.

[CK06]   J. Cheeger and B. Kleiner. On the differentiability of Lipschitz maps from metric measure spaces to Banach spaces. In *Inspired By SS Chern: A Memorial Volume in Honor of A Great Mathematician*, pages 129–152. World Scientific, 2006.

[CK10]   J. Cheeger and B. Kleiner. Differentiating maps into L^1, and the geometry of BV functions. *Annals of mathematics*, pages 1347–1385, 2010.

[CLY22]   V. Chousionis, S. Li, and R. Young. The strong geometric lemma for intrinsic lipschitz graphs in heisenberg groups. *Journal für die reine und angewandte Mathematik (Crelles Journal)*, 2022(784):251–274, 2022.

[CRTN01]   T. Coulhon, E. Russ, and V. Tardivel-Nachef. Sobolev algebras on Lie groups and Riemannian manifolds. *American Journal of Mathematics*, 123(2):283–342, 2001.

[DCTV07]   Y. De Cornulier, R. Tessera, and A. Valette. Isometric group actions on Hilbert spaces: growth of cocycles. *GAFA Geometric And Functional Analysis*, 17(3):770–792, 2007.

[Dix57]   J. Dixmier. Sur les représentations unitaires des groupes de Lie nilpotents. ii. *Bulletin de la Société mathématique de France*, 85:325–388, 1957.

[Dix59]   J. Dixmier. Sur les représentations unitaires des groupes de Lie nilpotents. i. *American Journal of Mathematics*, 81(1):160–170, 1959.

[DNM21]   P. De Nápoli and R. D. Martín. Mean value inequalities and characterizations of Sobolev spaces on graded groups. *arXiv preprint arXiv:2109.05064*, 2021.

[Dor85]   J. R. Dorronsoro. A characterization of potential spaces. *Proceedings of the American Mathematical Society*, 95(1):21–31, 1985.

[DS88]   N. Dunford and J. T. Schwartz. *Linear operators, part 1: general theory*, volume 10. John Wiley & Sons, 1988.

[EBGLD+21] S. Eriksson-Bique, C. Gartland, E. Le Donne, L. Naples, and S. Nicolussi-Golo. Nilpotent groups and biLipschitz embeddings into L^1. *arXiv preprint arXiv:2112.11402*, 2021.

[Fig76]   T. Figiel. On the moduli of convexity and smoothness. *Studia Mathematica*, 56(2):121–155, 1976.

[FO20a]   K. Fässler and T. Orponen. Dorronsoro's theorem in Heisenberg groups. *Bulletin of the London Mathematical Society*, 52(3):472–488, 2020.

[FO20b]   K. Fässler and T. Orponen. Vertical versus horizontal Sobolev spaces. *Journal of Functional Analysis*, 279(2):108517, 2020.

[Fol75]   G. B. Folland. Subelliptic estimates and function spaces on nilpotent Lie groups. *Arkiv för matematik*, 13(1):161–207, 1975.

[Gar20]   C. Gartland. Estimates on the markov convexity of carnot groups and quantitative nonembeddability. *Journal of Functional Analysis*, 279(8):108697, 2020.

[GLY09]   L. Grafakos, L. Liu, and D. Yang. Vector-valued singular integrals and maximal functions on spaces of homogeneous type. *Mathematica Scandinavica*, pages 296–310, 2009.

[Gro81]   M. Gromov. Groups of polynomial growth and expanding maps (with an appendix by Jacques Tits). *Publications Mathématiques de l'IHÉS*, 53:53–78, 1981.

[Han56]   O. Hanner. On the uniform convexity of l^p and ℓ^p. *Arkiv för Matematik*, 3(3):239–244, 1956.

[HKST15]   J. Heinonen, P. Koskela, N. Shanmugalingam, and J. T. Tyson. *Sobolev spaces on metric measure spaces*. Number 27. Cambridge University Press, 2015.

[HLNT13]   T. Heikkinen, J. Lehrbäck, J. Nuutinen, and H. Tuominen. Fractional maximal functions in metric measure spaces. *Analysis and Geometry in Metric Spaces*, 1(2013):147–162, 2013.

[HLP+52]   G. H. Hardy, J. E. Littlewood, G. Pólya, G. Pólya, et al. *Inequalities*. Cambridge university press, 1952.

[HN19]   T. Hytönen and A. Naor. Heat flow and quantitative differentiation. *Journal of the European Mathematical Society*, 21(11):3415–3466, 2019.

[HS20]   Y. Huang and S. Sun. Non-embedding theorems of nilpotent Lie groups and sub-Riemannian manifolds. *Frontiers of Mathematics in China*, 15(1):91–114, 2020.

[Jea14]   F. Jean. *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*. Springer, 2014.

[Jer86]   D. Jerison. The Poincaré inequality for vector fields satisfying Hörmander's condition. *Duke Mathematical Journal*, 53(2):503–523, 1986.

[Joh48]   F. John. Extremum problems with inequalities as subsidiary conditions, Studies and Essays Presented to R. Courant on his 60th Birthday, January 8, 1948, 1948.

[Kar94]   R. Karidi. Geometry of balls in nilpotent Lie groups. *Duke Mathematical Journal*, 74(2):301–317, 1994.

[Kir62]   A. A. Kirillov. Unitary representations of nilpotent Lie groups. *Russian mathematical surveys*, 17(4):53–104, 1962.

[Kle10]   B. Kleiner. A new proof of Gromov's theorem on groups of polynomial growth. *Journal of the American Mathematical Society*, 23(3):815–829, 2010.

[LD17]   E. Le Donne. A primer on Carnot groups: homogenous groups, Carnot-Carathéodory spaces, and regularity of their isometries. *Analysis and Geometry in Metric Spaces*, 5(1):116–137, 2017.

[Li14]   S. Li. Coarse differentiation and quantitative nonembeddability for Carnot groups. *Journal of Functional Analysis*, 266(7):4616–4704, 2014.

[LMN05]   J. R. Lee, M. Mendel, and A. Naor. Metric structures in L^1: dimension, snowflakes, and average distortion. *European Journal of Combinatorics*, 26(8):1180–1190, 2005.

[LN05]   J. R. Lee and A. Naor. Extending Lipschitz functions via random metric partitions. *Inventiones mathematicae*, 160(1):59–95, 2005.

[LN06]   J. R. Lee and A. Naor. L^p metrics on the Heisenberg group and the Goemans-Linial conjecture. In *2006 47th Annual IEEE Symposium on Foundations of Computer Science (FOCS'06)*, pages 99–108. IEEE, 2006.

[LN14]   V. Lafforgue and A. Naor. Vertical versus horizontal Poincaré inequalities on the Heisenberg group. *Israel Journal of Mathematics*, 203(1):309–339, 2014.

[Mal49]   A. Malcev. On a class of homogeneous spaces, izv. *Akad. Nauk. Armyan. SSSR Ser. Mat*, 13:201–212, 1949.

[MN14]   M. Mendel and A. Naor. Nonlinear spectral calculus and super-expanders. *Publications mathématiques de l'IHÉS*, 119:1–95, 2014.

[MTX06]   T. Martínez, J. L. Torrea, and Q. Xu. Vector-valued Littlewood–Paley–Stein theory for semigroups. *Advances in Mathematics*, 203(2):430–475, 2006.

[Nao14]   A. Naor. Comparison of metric spectral gaps. *Analysis and Geometry in Metric Spaces*, 2(1), 2014.

[NP11]   A. Naor and Y. Peres. L^p compression, traveling salesmen, and stable walks. *Duke Mathematical Journal*, 157(1):53–108, 2011.

[NY18]   A. Naor and R. Young. Vertical perimeter versus horizontal perimeter. *Annals of Mathematics*, 188(1):171–279, 2018.

[NY20]   A. Naor and R. Young. Foliated corona decompositions. *arXiv preprint arXiv:2004.12522*, 2020.

[Ost13]   M. I. Ostrovskii. Metric embeddings. In *Metric Embeddings*. de Gruyter, 2013.

[Pan89]   P. Pansu. Métriques de Carnot-Carathéodory et quasiisométries des espaces symétriques de rang un. *Annals of Mathematics*, pages 1–60, 1989.

[Pis75]   G. Pisier. Martingales with values in uniformly convex spaces. *Israel Journal of Mathematics*, 20(3):326–350, 1975.

[Sem96]   S. Semmes. On the nonexistence of bilipschitz parameterizations and geometric problems about A_∞-weights. *Revista Matemática Iberoamericana*, 12(2):337–410, 1996.

[Ste16]   E. M. Stein. *Singular Integrals and Differentiability Properties of Functions (PMS-30), Volume 30*. Princeton university press, 2016.

[Tes11]   R. Tessera. Asymptotic isoperimetry on groups and uniform embeddings into Banach spaces. *Commentarii Mathematici Helvetici*, 86(3):499–535, 2011.

[TJ74]   N. Tomczak-Jaegermann. The moduli of smoothness and convexity and the Rademacher averages of the trace classes S_p ($1 \le p < \infty$). *Studia Mathematica*, 50(2):163–182, 1974.

[Xu20]   Q. Xu. Vector-valued Littlewood-Paley-Stein theory for semigroups II. *International Mathematics Research Notices*, 2020(21):7769–7791, 2020.

(Seung-Yeon Ryoo)

MATHEMATICS DEPARTMENT, PRINCETON UNIVERSITY,
PRINCETON, NEW JERSEY 08544-1000, UNITED STATES

Email address: sryoo@math.princeton.edu