Introducing a modified version of Jones's $\beta$-numbers for Carnot groups which we call stratified $\beta$-numbers. We show that an analogue of Jones's traveling salesman theorem on 1-rectifiability of sets holds for any Carnot group if we replace previous notions of $\beta$-numbers in Carnot groups with stratified $\beta$-numbers. As we generalize both directions of the traveling salesman theorem, we get a characterization of subsets of Carnot groups that lie on finite length rectifiable curves. Our proof expands upon previous analysis of the Hebisch–Sikora norm for Carnot groups. In particular, we find new estimates on the drift between almost parallel line segments that take advantage of the stratified functions $\beta$ and also develop a Taylor expansion technique of the norm. We also give an example of a Carnot group for which a traveling salesman theorem based on the unmodified $\beta$-numbers must exhibit a gap between the necessary and sufficient directions.

MSC 2020
28A75, 53C17 (primary), 22E25 (secondary)
depends heavily on the geometry of the metric space $X$. In 1990, Peter Jones gave a solution to the problem in $\mathbb{R}^2$ [11] via the introduction of the so-called $\beta$-numbers.

Given an arbitrary subset $E \subset \mathbb{R}^2$ and a ball $B(x, r)$, we define

$$
\beta_E(x, r) := \inf_L \sup_{z \in E \cap B(x, r)} \frac{d(z, L)}{r}.
$$

Here, the infimum is taken over affine lines $L \subset \mathbb{R}^2$. We see that $\beta_E(x, r)$ is a number in $[0, 1]$ that can be thought of as the minimal (rescaled) radius tube that contains $E$ in $B(x, r)$ and so measures how close $E$ lies to some affine line.

Given the notion of $\beta$-numbers, Jones proved the following theorem, which is now known as the traveling salesman theorem.

**Theorem 1.1** [11]. For a set $E$ in $\mathbb{R}^2$, we define

$$
\beta(E) := \int_0^\infty \int_{\mathbb{R}^2} \beta_E(x, r)^2 \, dx \frac{dr}{r^2}.
$$

Then $E \subset \mathbb{R}^2$ lies on a finite length rectifiable curve if and only if $\text{diam}(E) + \beta(E)$ is finite. More precisely, there exists a universal constant $C > 1$ so that

1. if $\Gamma$ is any curve containing $E$ then $\beta(E) + \text{diam}(E) \leq CH^1(\Gamma)$,
2. if $\beta(E) + \text{diam}(E) < \infty$, then there exists a curve $\Gamma \supset E$ for which $H^1(\Gamma) \leq C(\beta(E) + \text{diam}(E))$.

The original theorem in Jones’s paper has $\beta$ expressed in terms of a sum over dyadic cubes of $\mathbb{R}^2$ (and $\beta_E$ is also expressed in terms of cubes), but it is well-known that the sum over cubes is equivalent to the integral over balls up to absolute multiplicative constants. It is also important to note that the $2$ in the exponent of $\beta_E^2$ comes from the power type of the modulus of convexity of $\mathbb{R}^n$, whereas the $2$ in the exponent of the $r^2$ is simply the Hausdorff dimension of $\mathbb{R}^2$.

In order for the term $\int_0^\infty \int_{\mathbb{R}^2} \beta_E(x, r)^2 \, dx \frac{dr}{r^2}$ to be finite, one must have that $\beta_E(x, r)$ is small for ‘most’ balls. Thus, one can view this as a quantitative version of Rademacher’s theorem, which says rectifiable curves have tangents almost everywhere. Note that there are two directions for the first statement of the theorem. The ‘only if’ direction is the necessary direction and the ‘if’ direction is the sufficient.

Since Jones’s original paper, the traveling salesman theorem has been generalized to higher dimensional Euclidean spaces $\mathbb{R}^d$ [17] and even Hilbert space $\ell_2$ [19]. See [3, 16] for more context and additional applications of these theorems. Recently, researchers have even studied variants of the analyst’s traveling salesman problem for Hölder curves [2], although we still do not have a complete picture.

As for addressing the analyst’s traveling salesman problem in non-Euclidean spaces, much of the effort has been in the setting of Carnot groups [5], and in particular the Heisenberg group [8, 12, 14, 15], although there has also been work done in general metric spaces [6, 9, 18] and certain fractal spaces [7].

We will give an overview of Carnot groups in the following section. For now, we simply say that they are a special class of nilpotent Lie groups that are topologically just some $\mathbb{R}^n$ and the reader can view as (possibly) noncommutative versions of vector spaces. Importantly, they contain a distinguished set of lines called the horizontal lines that will be the analogue of affine lines in $\mathbb{R}^n$. 
Due to the presence of these horizontal lines, one can define a naïve analogue of the $\beta$ numbers by simply infimizing over horizontal, rather than affine, lines:

$$\beta_E(x, r) = \inf_L \sup_{z \in E \cap B(x, r)} \frac{d(z, L)}{r}.$$ 

As mentioned above, $L$ are horizontal lines and $d$ will be any metric induced by a homogeneous norm (to be defined later). The ball $B(x, r)$ is a ball as defined by the metric $d$.

Results in Carnot groups have been partial. For the Heisenberg group $\mathbb{H}$ (the simplest nonabelian Carnot group) there has been an almost tight characterization.

**Theorem 1.2** [14, 15]. Let $E \subset \mathbb{H}$ be a subset of the first Heisenberg group. Define

$$\beta_p(E) := \int_0^\infty \int_{\mathbb{H}} \beta_E(x, r)^p \frac{dx}{r^d} dr.$$ 

If $E$ lies on a finite length rectifiable curve, then $\text{diam}(E) + \beta_4(E) < \infty$. On the other hand, if there is any $p < 4$ so that $\text{diam}(E) + \beta_p(E) < \infty$, then $E$ lies on a finite length rectifiable curve.

As $\beta(E) \in [0, 2]$, having integrability of a lower power of $\beta(E)$ is a stronger condition. It is unknown whether finiteness of $\text{diam}(E) + \beta_4(E)$ is sufficient for $E$ to lie on a rectifiable curve.

For general Carnot groups, there is only a one-sided result.

**Theorem 1.3** [5]. Let $E \subset G$ be a subset of a step $s$ Carnot group $G$ of Hausdorff dimension $Q$. If $E$ lies on a finite length rectifiable curve, then

$$\text{diam}(E) + \int_0^\infty \int_G \beta_E(x, r)^{2s^2} \frac{dx}{r^Q} dr < \infty.$$ 

If $s = 2$, we can take the exponent of $\beta_E$ to be 4.

The challenge for the sufficient direction has been that the direction of movement for Carnot groups are far more heterogeneous than Euclidean spaces, which cannot be detected by the functions $\beta$. In the next section, we will exhibit a Carnot group for which the necessary and sufficient directions of the traveling salesman theorem with the original functions $\beta$ must have a gap.

**Proposition 1.4.** There exists a Carnot group $G$ with Hausdorff dimension 6 so that for every $p_1 < 4$ and $p_2 > 2$ there are compact curves $E_1, E_2 \subset G$ such that $E_1$ is rectifiable and $E_2$ is not, but

$$\int_G \int_0^\infty \beta_{E_1}(x, r)^{p_1} \frac{dr}{r^6} dx = \infty,$$

$$\int_G \int_0^\infty \beta_{E_2}(x, r)^{p_2} \frac{dr}{r^6} dx < \infty.$$ 

Thus, we see that we need a new kind of $\beta$-number if we are to have any hope of a sharp traveling salesman theorem that works for every Carnot group. Note that this proposition does not preclude
the possibility that the usual $\beta$-based traveling salesman theorem might be tight for certain Carnot groups (like $\mathbb{H}$).

Given a step $s$ Carnot group, we can form a projective system by quotienting out the center. We let $G_s = G$ and get projections $G_i \to G_{i-1}$ until $G_0 = \{0\}$. Let $\pi_i : G \to G_i$ be the composition of the projections from $G_s$ to $G_i$.

We now define a new type of $\beta$-number that we call stratified $\beta$-numbers. For a subset $E \subset G$ and a ball $B(x,r)$, define

$$\hat{\beta}_E(x,r)^{2i} := \inf_{L} \sum_{i=1}^{s} \sup_{z \in E \cap B(x,r)} \left( \frac{d(\pi_i(z), \pi_i(L))}{r} \right)^{2i}$$

where the infimum is over horizontal lines. We have chosen the exponent so that $\hat{\beta}_E \geq \beta_E$. One can replace the $\ell_1$-sum by an $\ell_p$ combination of the summands to get an equivalent notion of $\hat{\beta}_E$.

Note that we have abused notation here as $d$ is originally a metric on $G$, not the projected $\pi_i(G)$. For now the reader can interpret this as $d$ being different homogeneous norms for every $\pi_i(G)$ (which will still give a valid $\hat{\beta}_E$) but we will clarify what we mean in the following section.

The following is the main result of this paper.

**Theorem 1.5.** Let $G$ be a step $s$ Carnot group with Hausdorff dimension $Q$. For a set $E \subset G$, we define the quantity

$$\hat{\beta}(E) := \int_0^\infty \int_G \hat{\beta}_E(x,r)^{2i} \frac{dx dr}{r^Q}.$$

Then $E$ lies on a finite length rectifiable curve if and only if $\text{diam}(E) + \hat{\beta}(E)$ is finite. More precisely, there exists a constant $C > 1$ depending only on $d$ and the Euclidean structure $| \cdot |$ on $G$ so that:

1. if $\Gamma$ is any curve containing $E$ then $\hat{\beta}(E) + \text{diam}(E) \leq C \mathcal{H}^1(\Gamma)$,
2. if $\hat{\beta}(E) + \text{diam}(E) < \infty$, then there exists a curve $\Gamma \supset E$ for which $\mathcal{H}^1(\Gamma) \leq C(\hat{\beta}(E) + \text{diam}(E))$.

We conclude with some equivalent formulations of $\hat{\beta}_E$ that might be of use in the future. The original $\beta_E(x,r)$ was described as the rescaled radius of the thinnest tube containing $E \cap B(x,r)$. Thus, it can be rewritten as

$$\beta_E(x,r) = \inf_L \inf \{ \varepsilon > 0 : E \cap B_G(x,r) \subset L \cdot \delta_r(B_G(\varepsilon)) \}.$$
The normalization in (3) cannot be removed as $\rho$ does not behave well with regard to scaling. One sees that $\hat{\beta}$ gives asymptotically tighter bounds than $\beta$ the lower the layer we are considering. Proposition 1.6 will be proven in Section 7. The fact that the $B_{\mathbb{R}^n}(\varepsilon^s)$ is a Euclidean ball is not terribly important. The important thing is the way it scales with $\varepsilon$. It could be the ball of any norm on $\mathbb{R}^n$ of radius $\varepsilon^s$.

Proofs of the traveling salesman theorem tend to follow a standard path. The most important step is to establish the following Alexandrov-like curvature condition: if $p_1, p_2, p_3$ are in $B(x, r)$ and $d(p_i, p_j) \geq \alpha r$ for $i \neq j$, then there exist constants $c, C > 0$ depending on $\alpha$ so that

$$c \beta(x, r)^{\gamma} r \leq d(p_1, p_2) + d(p_2, p_3) - d(p_1, p_3) \leq C \beta(x, r)^{\gamma} r,$$

(3)

where $\beta^\gamma$ is the appropriate $\beta$ for the situation. Once the left-hand inequality is established, one can apply a clever telescoping argument to derive the necessary direction (although a little more work is needed, see Subsection 5.4). If the right-hand inequality is established, then one applies a farthest insertion algorithm (given in Jones's original paper) to derive the sufficient direction. Thus, most of this paper will be devoted to proving results in the vein of (3) for $\hat{\beta}_E^{\infty}$ (see Proposition 5.6 and Corollaries 5.7 and 6.8 for the relevant results). We will not go too in-depth for the subsequent steps, instead referring the reader to other papers where these steps are covered in detail.

2 | PRELIMINARIES

A Lie algebra $\mathfrak{g}$ is stratified if it is nilpotent and can be decomposed into a direct sum of subspaces $\mathfrak{g} = V_1 \oplus \cdots \oplus V_s$ for which $[V_1, V_j] = V_{j+1}$ for $j \geq 1$. It is understood that $V_k = 0$ for $k > s$. The layer $V_1$ is called the horizontal layer and the largest $s$ for which $V_s \neq 0$ is called the (nilpotency) step of $\mathfrak{g}$. We let $n_i = \dim(V_i)$ and $n = n_1 + \cdots + n_s$. We can thus identify $\mathfrak{g}$ with $\mathbb{R}^n$ where the direct sum $\mathfrak{g} = \bigoplus V_i$ is orthogonal. Note that our choice of basis for $\mathbb{R}^n \cong \mathbb{R}^{n_1+\cdots+n_s}$ has been chosen to be consistent with the grading of $\mathfrak{g}$.

A Carnot group $G$ is a simply connected Lie group whose Lie algebra is stratified. The exponential map is a diffeomorphism between simply connected Lie groups and their Lie algebras, and so we can use $\exp$ to identify elements of $G$ using vectors of $\mathfrak{g} \cong \mathbb{R}^n$. This allows us to push the coordinates of $\mathbb{R}^n$ to $G$ (the so-called exponential coordinates). Thus, we will use $(g_1, \ldots, g_s) \in \mathbb{R}^n$ to write coordinates for elements $g \in G$ where $g_i$ are the components of $\exp^{-1}(g)$ in $V_i$. The identity element of $G$ is still 0.

The subgroups $\exp(V_i \oplus V_{i+1} \oplus \cdots \oplus V_s)$ are all normal. For $i \in \{1, \ldots, s\}$, we then define the projection maps $\pi_i : G \to G/\exp(V_{i+1} \oplus \cdots \oplus V_s)$. In exponential coordinates, this amounts to the usual projection of $\mathbb{R}^n$ onto $\mathbb{R}^{n_1+\cdots+n_i}$. The image of projection $\pi_i$ is also a Carnot group and its Lie algebra is simply $\mathfrak{g}/\bigoplus_{j=i+1}^s V_j$.

We use $\| \cdot \|$ to denote the standard Euclidean norm on $G$ when viewed as $\mathbb{R}^n$ through exponential coordinates. This will depend upon an arbitrary choice of basis for the $V_i$'s. We can then also define centered Euclidean balls in $G$ as such:

$$B_{\mathbb{R}^n}(\varepsilon) = \{g \in G : \|g\| < \varepsilon\}.$$
The group multiplication in \( G \) can be expressed at the level of the Lie algebra via the Baker–Campbell–Hausdorff (BCH) formula:

\[
\log(\exp(U) \exp(V)) = \sum_{k>0} \frac{(-1)^{k-1}}{k} \sum_{r_1+s_1>0, \ldots, r_k+s_k>0, 1 \leq i \leq k} a(r_1, s_1, \ldots, r_k, s_k) [U^{r_1} V^{s_1} \ldots U^{r_k} V^{s_k}], \tag{4}
\]

where the bracket term denotes iterated Lie brackets

\[
[X^{r_1} Y^{s_1} \ldots X^{r_n} Y^{s_n}] = [X, [X, \ldots [X, Y, \ldots [Y, \ldots Y]]]]
\]

One can also translate the BCH formula into the exponential coordinates as

\[
(u_1, \ldots, u_s) \cdot (v_1, \ldots, v_s) = (u_1 + v_1, u_2 + v_2 + P_2, \ldots, u_s + v_s + P_s),
\]

where each \( P_i \) are polynomials of \( u_1, \ldots, u_{i-1}, v_1, \ldots, v_{i-1} \). Recall that each \( u_i, v_i \) are elements of \( \mathbb{R}^{n_i} \). We call these \( P_i \) the BCH polynomials.

For example, the Heisenberg group \( \mathbb{H} \) has a Lie algebra that is generated by \( X, Y, Z \) with the bracket relations \([X, Y] = Z, [X, Z] = [Y, Z] = 0 \). We identify each element of \( \mathbb{H} \) by \((a, b, c) = \exp(aX + bY + cZ)\). Thus, the BCH formula says that

\[
(a, b, c) \cdot (a', b', c') = \exp \left( aX + bY + cZ + a'X + b'Y + c'Z \right.
\]

\[
+ \frac{1}{2} \left[ aX + bY + cZ, a'X + b'Y + c'Z \right] \bigg) = \exp((a + a')X + (b + b')Y + (c + c' + (ab' - a'b)/2)Z) = (a + a', b + b', c + c' + (ab' - a'b)/2).
\]

In the first line, we used the fact that all the \( k > 2 \) terms of (4) are zero due to the nilpotency of the Lie algebra. Thus, \( P_2((a, b), (a', b')) = \frac{1}{2}(ab' - a'b) \) in this case.

An important fact about BCH polynomials is that \( P_k(u, v) \) are composed of monomials of iterated Lie brackets of terms of \( u = (u_1, \ldots, u_s) \) and \( v = (v_1, \ldots, v_s) \) whose layers add up to \( k \). For example, \( P_3 \) is a linear combination of the monomials

\[
[u_1, [u_1, v_1]], [v_1, [u_1, v_1]], [u_1, v_2], [v_1, u_2].
\]

Importantly, for every \( \lambda > 0 \), there exists an automorphism \( \delta_\lambda : G \to G \), which in exponential coordinates is given by

\[
\delta_\lambda(g_1, \ldots, g_s) = (\lambda^1 g_1, \lambda^2 g_2, \ldots, \lambda^s g_s).
\]

A homogeneous norm is a function \( N : G \to [0, \infty) \) for which

1. \( N(g^{-1}) = N(g) \),
2. \( N(\delta_\lambda(g)) = \lambda N(g) \), for all \( g \in G, \lambda > 0 \),
Property (4) says that the norm is subadditive and is not always standard in the definition of homogeneous norm. We include it because then one can define, for any homogeneous norm, a corresponding homogeneous metric \( d(g, h) = N(g^{-1}h) \), which will be a left-invariant metric. If subadditivity were not present, then one cannot guarantee the triangle inequality. When we want to stress the subadditivity property of the homogeneous norm, we will use the term *subadditive homogeneous norm*. Property (2) of homogeneous norms tells us that the \( \delta_\lambda \) automorphisms scale the metric:

\[
d(\delta_\lambda(g), \delta_\lambda(h)) = \lambda d(g, h).
\]

We remark that it is well-known that any two homogeneous norms (subadditive or not) on a Carnot group are bi-Lipschitz equivalent.

In this paper, we will work primarily with homogeneous metrics, and specifically, with the metrics corresponding to two (classes of) subadditive homogeneous norms that exist for every Carnot group. The first norm was introduced by Hebisch and Sikora in [10]. For a fixed Euclidean ball \( B_{\mathbb{R}^n}(\eta) \subset \mathbb{R}^n \cong G \) centered at the origin, one can define the Minkowski gauge

\[
N_{HS}(g) = \inf \{ \lambda > 0 : \delta_{1/\lambda}(g) \in B_{\mathbb{R}^n}(\eta) \}.
\]

As the balls \( B_{\mathbb{R}^n}(\eta) \) are not homothetic under the dilation \( \delta_\lambda \), we get truly different norms for different \( \eta \). The main result in [10] is that for sufficiently small \( \eta \), \( N_{HS} \) is a subadditive homogeneous norm that is smooth on \( G \setminus \{0\} \). We use \( d_{HS} \) for the corresponding homogeneous metric. The balls of \( N_{HS} \) centered at the origin are then dilations of \( B_{\mathbb{R}^n}(\eta) \) and so are axis-aligned ellipsoids. Given a Carnot group \( G \) and one of its projections \( \pi_i(G) \), if we take \( \eta > 0 \) small enough, then \( B_{\mathbb{R}^n}(\eta) \) and \( B_{\mathbb{R}^n_1+\cdots+n_i}(\eta) \) both can form Hebisch–Sikora norms. In this way, we can assume that a \( N_{HS} \), defined for \( G \), is also a homogeneous norm for all of its projections \( \pi_i(G) \). This is how we make sense of the \( d \) in \( \hat{\beta} \).

Having fixed \( \eta \), we also take \( B_{\mathbb{R}^n_1+\cdots+n_i}(\eta) \) as the unit ball of a new Euclidean norm \(| \cdot |\) on each of the quotients \( \pi_i(G) \cong \mathbb{R}^{n_1+\cdots+n_i} \). This is equivalent to defining \(| \cdot | = \frac{1}{\eta} \| \cdot \|\) where \( \| \cdot \|\) is the standard Euclidean metric. Now \(| g \pm h |\) makes sense for \( g, h \in G \). As the unit ball of \( N_{HS} \) is now also the unit ball of \(| \cdot |\), we then have that

\[
d_{HS}(\pi_1(g), \pi_1(h)) = |\pi_1(g) - \pi_1(h)|, \quad \forall g, h \in G.
\]

From now on, all balls of the form \( B_{\mathbb{R}^n} \) will be taken with respect to \(| \cdot |\).

We remark that the maps \( \pi_i \) are 1-Lipschitz as we have the relationship \( N_{HS}(\pi_i(g)) \leq N_{HS}(g) \). This gives us the following relation between \( d_{HS} \) in \( G \) and in its projections under \( \pi_i \).

**Lemma 2.1.**

\[
d_{HS}(\pi_1(g), \pi_1(h)) = \min_{g' \in \pi_1^{-1}(\pi_1(g)), h' \in \pi_1^{-1}(\pi_1(h))} d_{HS}(g', h'), \quad \forall g, h \in G.
\]

**Proof.** In one direction, we have \( d_{HS}(g', h') \geq d_{HS}(\pi_1(g'), \pi_1(h')) = d_{HS}(\pi_1(g), \pi_1(h)) \).
In the other direction, let $u = \pi_i(g)^{-1}\pi_i(h) \in G_i \cong \mathbb{R}^{n_1+\cdots+n_i}$. We define the element $(u, 0) \in G$ where $0 \in \mathbb{R}^{n_1+\cdots+n_i}$. Given any $g' \in \pi_i^{-1}(g)$, we let $h' = g' \cdot (u, 0)$. Then

$$d_{HS}(g', h') = N_{HS}(u, 0) = N_{HS}(u) = d_{HS}(\pi_i(g), \pi_i(h)).$$

Here, the first $N_{HS}$ is in $G$ and the second is in $G_i$. □

The second norm is analogous to the $\ell_\infty$ norm and is defined as

$$N_\infty(x_1, \ldots, x_s) = \max_{1 \leq i \leq s} \lambda_i |x_i|^{1/i},$$

where $\lambda_i > 0$ are constants with $\lambda_1 = 1$. Lemma 2.5 of [4] says that one can choose the $\lambda_i$s so that $N_\infty$ is a subadditive homogeneous norm. The proof chooses $\lambda_i$ inductively based only on the BCH polynomial $P_i$ and $\{\lambda_j\}_{j=1}^{i-1}$. Thus, we may choose $\{\lambda_j\}_{j=2}^{s}$ so that each set of coefficients $\{\lambda_j\}_{j=1}^{i}$ defines a subadditive homogeneous norms on $\pi_i(G)$. We use $d_\infty$ for the corresponding homogeneous metric.

Both norms are known to be quasi-convex, that is, there is a constant $C_d > 0$ depending only on the metric so that for any two points $x, y \in G$, there is a rectifiable curve $\gamma$ going from $x$ to $y$ with

$$\text{Length}(\gamma) \leq C_d d(x, y).$$

Also, we can use left-translation to pushforward a given inner product on $\mathfrak{g} \cong T_0 G$ to get a left-invariant inner product on $TG$. This allows us to define a Riemannian metric $\rho$. It is known that given a $M > 0$, there exists a constant $C_0 > 0$ depending on $M$, $| \cdot |$, and $\rho$ so that origin centered Riemannian balls of radius at most $M$ are equivalent to origin centered Euclidean balls:

$$B_{\mathbb{R}^n}(\varepsilon/C_0) \subseteq B_{\rho}(\varepsilon) \subseteq B_{\mathbb{R}^n}(C_0 \varepsilon), \quad \forall \varepsilon \in (0, M].$$

Elements of the exponential image of the horizontal layer, $\exp(V_1)$, are called horizontal elements of $G$. In exponential coordinates, horizontal elements look like $(g_1, 0, \ldots, 0)$. Let $S^{n_1-1}$ be the unit sphere of $V_1 \subset \mathfrak{g}$ with respect to the Euclidean metric $| \cdot |$. Given any $v \in S^{n_1-1}$ and $g \in G$, we get an isometric embedding of $\mathbb{R}$ into $G$ by $t \mapsto g \exp(tv)$. We call the image of such mappings horizontal lines.

We also define the map

$$\bar{\pi} : G \to G$$

$$(g_1, \ldots, g_s) \mapsto (g_1, 0, \ldots, 0).$$

Thus, $\bar{\pi}$ maps $g$ to the horizontal element of $G$ that lies ‘under’ $g$. Note that this is not a homomorphism, nor it is Lipschitz. However, we do have

$$N(\bar{\pi}(g)) \leq N(g), \quad \forall g \in G$$

(8)
when $N \in \{N_{HS}, N_\infty\}$. We can now define
\[
NH(g) = d(g, \bar{\pi}(g)),
\]
which measures how non-horizontal an element of $g$ is. Here, $d$ can be any homogeneous metric.

Finally, it is not hard to see from the BCH polynomials that the Jacobians of all left translations have determinant 1. This tells us that the Lebesgue measure on the underlying manifold $\mathbb{R}^n$ of the Carnot group $G$ is also a Haar measure.

3 | THE INADEQUACY OF $\beta$

In this section, we show that the non-stratified functions $\beta$ cannot provide sharp two-sided traveling salesman inequalities for certain Carnot groups. This necessitates our use of a more complex $\beta$ in the sequel.

Consider the Carnot group $G = \mathbb{R}^2 \times \mathbb{H}$. This can be viewed as the group $(\mathbb{R}^5, \cdot)$ where
\[
(x, y, u, v, w) \cdot (x', y', u', v', w') = (x + x', y + y', u + u', v + v', w + w' + \frac{1}{2}(uw' - u'v)).
\]

We can write elements of $G$ as $(z, g)$ where $z \in \mathbb{R}^2$ and $g \in \mathbb{H}_1$. Then we endow $G$ with a product metric $d((z, g), (z', g')) = \max\{|z - z'|, d_\infty(g, g')\}$ where $d_\infty(g, g') = N_\infty(g^{-1}g')$ comes from the norm of (5). The Hausdorff dimension of $G$ is 6.

Let $G_1 = \mathbb{H}$ and $G_2 = \mathbb{R}^2$. We can also view these as subgroups $H_1 = \{0\} \times \mathbb{H}$ and $H_2 = \mathbb{R}^2 \times \{0\}$ of $G$. It is immediate from the construction of the metric on $G$ that $G_i$ is isometrically isomorphic to $H_i$. Let $\iota_1 : \mathbb{H} \to H_1$ and $\iota : \mathbb{R}^2 \to H_2$ be the natural isomorphisms. Since the $\beta$-based traveling salesman theorems for $\mathbb{H}$ and $\mathbb{R}^2$ have different exponents, we will use these two subgroups to exhibit a gap between the necessary and sufficient directions a $\beta$-based traveling salesman theorem for $G$.

Proposition 1.4 will follow from the following proposition.

**Proposition 3.1.** For every $p_1 < 4$, there exists a finite length rectifiable curve $E_1 \subset H_1$ such that
\[
\int_G \int_0^\infty \beta_{E_1}(x, r)^{p_1} \frac{dr}{r^6} dx = \infty.
\]

On the other hand, for every $p_2 > 2$, there exists a non-rectifiable curve $E_2 \subset H_2$ such that
\[
\int_G \int_0^\infty \beta_{E_2}(x, r)^{p_2} \frac{dr}{r^6} dx < \infty.
\]

Let $\bar{\pi}_i : G \to H_i$ be the natural projection maps. From the definition of the product metric $d$, we see that the $\bar{\pi}_i$ are all 1-Lipschitz. We first show that if a set $E$ lies in one of these subgroups, the best approximating horizontal line can also be taken to lie in the subgroup.
Lemma 3.2. Let $i \in \{1, 2\}$. If $E$ is a subset of $G_i$, then
\[
\inf_{L \subset G} \inf_{z \in \mathcal{E}(E)} d(z, L) = \inf_{L \subset H_i} \inf_{z \in E} d(z, L).
\]
Here, the $L$ in the infimums are horizontal lines in their respective groups.

Proof. As $G_i$ and $H_i$ are isometrically isomorphic via $\iota_i$, it suffices to prove that if $E \subseteq H_i$, then
\[
\inf_{L \subset G} \inf_{z \in \mathcal{E}(E)} d(z, L) = \inf_{L \subset H_i} \inf_{z \in E} d(z, L),
\]
where the infimum on the right-hand side is over horizontal lines lying in $H_i$. The $\leq$ inequality is clear. Thus, we only need to prove the $\geq$ inequality.

Let $L : \mathbb{R} \to G$ be the parameterization of a horizontal line in $G$ that we will also identify with its image. Now define the new map $L' = \tilde{\pi}_i \circ L$. This also parameterizes a horizontal line that is now a subset of $H_i$. It suffices to show
\[
d(z, L) \geq d(z, L'), \quad \forall z \in E.
\]
Let $z \in E$ and let $t \in \mathbb{R}$. We have by 1-Lipschitzness of $\tilde{\pi}_i$ that
\[
d(z, L(t)) \geq d(\tilde{\pi}_i(z), \tilde{\pi}_i(L(t))) = d(z, L'(t)),
\]
where we used the construction of $L'$ and the fact that $z \in H_i$ in the last equality. This proves that $d(z, L) \geq d(z, L')$, which proves the lemma.

Lemma 3.3. There are constants $c_1, c_2 > 0$ so that for any $p$,
\[
\int_G \int_0^\infty \beta_{1\mathbb{E}}(x, r)^p \frac{dr}{r^6} \, dx = c_1 \int_{G_1} \int_0^\infty \beta_{E}(x, r)^p \frac{dr}{r^4} \, dx, \quad \forall E \subset G_1,
\]
\[
\int_G \int_0^\infty \beta_{2\mathbb{E}}(x, r)^p \frac{dr}{r^6} \, dx = c_2 \int_{G_2} \int_0^\infty \beta_{E}(x, r)^p \frac{dr}{r^2} \, dx, \quad \forall E \subset G_2.
\]

Proof. We will prove the $G_1$ case of the lemma as the $i = 2$ case only requires trivial modifications of the forthcoming proof.

Let $(z, g) \in G$ here $z \in \mathbb{R}^2$ and $g \in \mathbb{H}$. We first claim that $B((z, g), r)$ intersects $H_1$ if and only if $|z| \leq r$ in which case
\[
B((z, g), r) \cap H_1 = \iota_1(B(g, r)).
\]
Note that $H_1 = \{(0, h) : h \in \mathbb{H}\}$. Thus, the first part of the claim follows as soon as we show that $d((z, g), H_1) = |z|$. The $\leq$ direction follows as $(0, g) \in H_1$ and $d((z, g), (0, g)) = |z|$. Now let
\((0, h) \in H_1\). Then
\[
d((z, g), (0, h)) = \max\{|z|, d_{\infty}(g, h)| \geq |z|
\]
which gives the \(\geq\) direction as desired.

As for the second part of the claim, we see that when \(|z| \leq r\), then \((0, h) \in B((z, g), r) \cap H_1\) precisely when \(d(g, h) \leq r\), that is \(h \in B(g, r)\). As \(\{0\} \times B(g, r) = t_1(B(g, r))\), we get the second part of the claim.

Now let \(E \subset G_1\) and let \(g \in H\). By the previous claim we have that \(\beta_{t_1(E)}((z, g), r) = 0\) unless \(|z| \leq r\) in which case
\[
\beta_{t_1(E)}((z, g), r) = \inf_{LCG} \inf_{y \in B((z, g), r) \cap t_1(E)} \frac{d(y, L)}{r} = \inf_{LCG} \inf_{y \in t_1(B(g, r) \cap E)} \frac{d(y, L)}{r} = \beta_{E}(g, r)
\]
where we used Lemma 3.2 in the penultimate equality.

We now integrate
\[
\int_0^\infty \int_G \beta_{t_1(E)}(x, r) \, dx \, dr \frac{dr}{r^6} = \int_0^\infty \int_H \int_{\mathbb{R}^2} \beta_{t_1(E)}((z, g), r) \, dz \, dy \, dr \frac{dr}{r^6}
\]
\[
= \int_0^\infty \int_H \int_{\mathbb{R}^2} \mathbf{1}_{|z| \leq r} \beta_E(g, r) \, dz \, dy \, dr \frac{dr}{r^6}
\]
\[
= c_1 \int_0^\infty \int_H \beta_E(g, r) \, dy \, dr \frac{dr}{r^4}.
\]
This proves the lemma. \(\square\)

We now prove Proposition 3.1.

\textbf{Proof of Proposition 3.1.} As mentioned at the beginning of this section, this will amount to recalling the corresponding \(E_1 \subset G_1\) so that
\[
\int_H \int_0^\infty \beta_{E_1}(x, r) p_1 \, dr \frac{dr}{r^4} \, dx = \infty,
\]
\[
\int_{\mathbb{R}^2} \int_0^\infty \beta_{E_2}(x, r) p_2 \, dr \frac{dr}{r^2} \, dx < \infty
\]
for the chosen \(p_1, p_2\).

The \(E_1\) will come from the counterexample from [12], although three remarks need to be made.

First, in [12], Juillet uses a special metric on \(H\), the Carnot–Carathéodory distance (that we will not introduce), whereas we are using the \(N_{\infty}\) group norm. However, as the Carnot–Carathéodory distance is induced by a homogeneous norm, the two metrics are bi-Lipschitz equivalent. As \(\beta\)-numbers change only by a nonzero multiplicative factor under bi-Lipschitz change of metrics, the finiteness of the Carleson integrals are the same regardless of which metric we use.
Second, Juillet sums functions \( \beta \) over balls whose centers come from a system of nets in \( \mathbb{H} \), but it is well-known that summing \( \beta \) over nets and Carleson integrating \( \beta \) give quantities that differ only by a multiple depending on the choice of nets. In particular, infiniteness of the quantity is preserved.

Finally, the result of Juillet states our result only for when \( p_1 \) is chosen to be 2 (although the statement also holds for any \( p_1 < 2 \)). But a simple modification of the construction by choosing \( \theta_n = \frac{C}{n^{2/p_1}} \) on \([12, p. 1046]\) allows us to extend the example to any \( p_1 < 4 \). The rest of the proof will remain unchanged except at the following spots. For \([12, p. 1047, Lemma 3.2]\), \( L \), the length of the curve, will now be bounded by a constant that depends on \( p_1 \). In the proof we have the modified bound:

\[
\sum_{n=1}^{N} \theta_n^2 \leq \sum_{n=1}^{N} \frac{C^2}{n^{4/p_1}} < \infty
\]

as \( p_1 > 4 \).

At the summation on \([12, p. 1052]\), we change \( \beta^2 \) to \( \beta^{p_1} \) and get the lower bound for the sum of functions \( \beta \):

\[
\sum_{k \in \mathbb{N}} 2^{-k} \sum_{x \in \Delta_k} \beta^{p_1}_k(x, A \cdot 2^{-k})(\omega([0, 1])) \geq \sum_{k \in \mathbb{N}} 2^{-k} 2^{k} \left( \frac{D_{\theta_k}[k/2]+1}{A} \right)^{p_1} \\
\geq C^{p_1/2} \sum_{k \in \mathbb{N}} \frac{1}{[k/2]+1} = \infty.
\]

For \( E_2 \), one can use the example of \([1, Proposition 4.1]\). Again, the quantity calculated there was a sum of \( \beta \) numbers over balls centered at a system of nets, but this is equivalent to our Carleson integral. \( \square \)

4 | LEMMAS

In this section, we prove a few results that will be useful in the following sections.

The following lemma will allow us to bound BCH polynomials.

**Lemma 4.1.** There exists some constant \( C > 0 \) depending only on \( G \) and the Euclidean structure \( \| \cdot \| \) so that if \( |y_i| \leq \eta \) and \( |x_i| \leq 1 \) for all \( i \in \{1, \ldots, k-1\} \) and any \( \eta \in (0, 1) \), then

\[
|P_k(x_1, \ldots, x_{k-1}, y_1, \ldots, y_{k-1})| \leq C \eta.
\]

**Proof.** Indeed, as group multiplication is smooth, we get that \( P_k \) are smooth functions. Furthermore, as \( P_k \) come from sums of nested Lie bracket monomials, we have that \( P_k(x_1, \ldots, x_{k-1}, 0, \ldots, 0) = 0 \). One can then use Taylor’s approximation to get that when \( \max_{1 \leq i \leq k-1} |y_i| \leq 1 \), we have

\[
|P_k(x_1, \ldots, x_{k-1}, y_1, \ldots, y_{k-1})| \leq C_x \max_{1 \leq i \leq k-1} |y_i|,
\]
where $C_x$ is a constant that depends smoothly on $x = (x_1, \ldots, x_{k-1})$. Taking

$$C = \sup\{C_x : \max_{1 \leq i \leq k-1} |x_i| \leq 1\}$$

gives the lemma.

The next few lemmas are the main results of this section. They will allow us to conclude the following fact: if $g \in h \cdot B_{\mathbb{R}^n}(\varepsilon)$ and $u, v \in S^{n_1-1}$ are so that $|u - v| < \varepsilon$, then $g \exp(u) \in h \exp(v) \cdot B_{\mathbb{R}^n}(C\varepsilon)$ for some constant $C$ depending only on the metric $d$ and the Euclidean metric on $G$. Had we only had $d(g, h) \leq \varepsilon$, then we could only guarantee $d(g \exp(u), h \exp(v)) \leq C\varepsilon^{1/s}$ (see [13, Lemmas 3.6, 3.7, and 3.8]).

This essentially comes from the fact that we are group multiplying an element $g^{-1}h$ with entries of order $\varepsilon$ by displacements of order $1$. One can see this by considering a simple case in the Heisenberg group. Let $h = (0, 0, 0)$, $g = (0, \varepsilon, 0)$, and $u = v = (1, 0, 0)$.

Thus, while $|g^{-1}h|$ and $N(g^{-1}h)$ are both of order $\varepsilon$, we have that $|((g \exp(u))^{-1}h \exp(v)|$ is of order $\varepsilon$ while $N(((g \exp(u))^{-1}h \exp(v))$ is of order $\varepsilon^{1/2}$.

This behavior will be why stratified $\beta$-numbers are needed.

We first begin with a smoothness argument on products of elements.

**Lemma 4.2.** There exists a $C > 0$ depending only on $G$ and the Euclidean structure $|\cdot|$ so that if $\eta \in (0, 1)$, $p \in B_{\mathbb{R}^n}(\eta)$, and $u, v \in B_{\mathbb{R}^n}(1)$ are such that $|u - v| \leq \eta$, then

$$v^{-1}pu \in B_{\mathbb{R}^n}(C\eta).$$

**Proof.** Define the function $\varphi(v, p, u) = v^{-1}pu$. This is a smooth function on $G \times G \times G$. In particular, it is $L$-Lipschitz with respect to the Euclidean metric on $B_{\mathbb{R}^n}(2) \times B_{\mathbb{R}^n}(2) \times B_{\mathbb{R}^n}(2)$ for some $L > 0$ (depending on $|\cdot|$) and so

$$|\varphi(v, p, u)| \leq |\varphi(v, p, u) - \varphi(u, p, u)| + |\varphi(u, p, u) - \varphi(u, 0, u)| \leq L|v - u| + L|p| \leq 2L\eta. \quad \square$$

We can now prove the main result of this section.

**Lemma 4.3.** There exists a $C > 0$ depending only on $G$ and the Euclidean structure $|\cdot|$ so that if $g, h \in G$ and $u, v \in B_{\mathbb{R}^n}(1)$ are such that $g \in h \cdot \delta_{\varepsilon}(B_{\mathbb{R}^n}(\eta))$ and $|u - v| \leq \eta$ for some $\eta \in (0, 1)$ and $\varepsilon > 0$, then

$$g\delta_t(u) \in h\delta_t(v) \cdot \delta_{\varepsilon}(B_{\mathbb{R}^n}(C\eta)), \quad \forall t \in [0, \varepsilon].$$

**Proof.** We may dilate the setting to reduce to the case that $\varepsilon = 1$. Thus, let $t \in [0, 1]$ and $u' = \delta_t(u)$ and $v' = \delta_t(v)$. We need to prove that

$$(v')^{-1}h^{-1}gu' \in B_{\mathbb{R}^n}(C\eta)$$

for some $C$. 

As $u, v$ lie in the first layer $\mathbb{R}^{n_1}$, we get that $u' = tu$ and $v' = tv$ and so we still have $|u' - v'| = t|u - v| \leq \eta$. Note that $h^{-1}g \in B_{\mathbb{R}^n}(\eta)$ and $u', v' \in B_{\mathbb{R}^{n_1}}(1)$. Thus, the result follows from applying Lemma 4.2 with $v', h^{-1}g$, and $u'$.

We now relate the relation between two points of $G$ in terms of containment by Euclidean balls with a $\hat{\beta}$-style bound on their distance.

**Lemma 4.4.** For $\ell > 0$, $p, q \in G$, and homogeneous metric $d$, define

$$E_{\ell}(p, q) = \inf \{ t > 0 : p \in q\delta_{\mathbb{R}^n}(t) \},$$

$$P_{d, \ell}(p, q) = \max_{1 \leq i \leq s} \left( \frac{d(\pi_i(p), \pi_i(q))}{\ell} \right)^i.$$

Then for every $M > 0$ and homogeneous metric $d$, there exists a constant $C \geq 1$ depending only on $M, d$, and the Euclidean structure $| \cdot |$ so that for all $\ell > 0$, if $E_{\ell}(p, q) \leq M$, then

$$\frac{1}{C} P_{d, \ell}(p, q) \leq E_{\ell}(p, q) \leq C P_{d, \ell}(p, q).$$

**Proof.** As the $\pi_i$ are homomorphisms, we may translate and dilate to reduce to the case $\ell = 1$ and $q = 0$. We will prove the lemma for $d_{\infty}$ as that allows us to conclude the same statement for any homogeneous metric by a bi-Lipschitz change of metric (for a different $C$). We now suppose $E_1(p, 0) \leq M$.

Let $\eta = E_1(p, 0) = |p|$. Then $|p_i| \leq C_0 \eta$ for all $i$ for some $C_0$ depending on the Euclidean metric. As $\eta \leq M$, we have for any $i$ that

$$P_{d_{\infty}, 1}(p, 0) = \max_{1 \leq i \leq s} N_{\infty}(\pi_i(p))^i \leq \max_{1 \leq i \leq s} \max_{1 \leq j \leq i} \lambda_i^j C_0^i/j \eta^i/j \leq \eta \max_{1 \leq i \leq s} \max_{1 \leq j \leq i} \lambda_i^j C_0^i/j M^i/j - 1.$$

The lower bound on $E_1(p, 0)$ now holds for some $C$ depending on the $\lambda_i, M$, and $C_0$.

Now assume $P_{d_{\infty}, 1}(p, 0) = \eta$ and so $N_{\infty}(\pi_i(p))^i \leq \eta$ for all $i$. Then for any $i$

$$\lambda_i^j |p_i| \leq N_{\infty}(\pi_i(p))^i \leq \eta$$

and so $|p_i| \leq C_1 \eta$ for some $C_1 > 0$ depending on the $\lambda_i$. The upper bound on $E_1(p, 0) = |p|$ now holds for some $C$ depending on the $\lambda_i$ and the Euclidean metric. \(\square\)

The following corollary is an immediate consequence of the previous two lemmas.

**Corollary 4.5.** There exists a constant $C > 0$ depending only on $d_{HS}$ and the Euclidean structure $| \cdot |$ so that if $g, h \in G$ and $u, v \in B_{\mathbb{R}^{n_1}}(1)$ are such that $g \in h \delta_{\mathbb{R}^n}(\eta)$ and $|u - v| \leq \eta$ for some $\eta \in (0, 1)$ and $\ell > 0$, then

$$\sup_{i \in [0, \ell]} \left( \frac{d_{HS}(\pi_i(g \cdot (tu, 0, ..., 0)), \pi_i(h \cdot (tv, 0, ..., 0)))}{\ell} \right)^i \leq C \eta, \quad \forall i \in \{1, ..., s\}.$$
5 | NECESSITY

We begin by proving the necessary direction of Theorem 1.5. This will be achieved by proving the following theorem.

**Theorem 5.1.** Let $G$ be a step $s$ Carnot group with Hausdorff dimension $Q$ and $\Gamma \subset G$, a finite length rectifiable curve. Then there is some constant depending only on $G$ and its Euclidean structure $\cdot \cdot \cdot$ so that

$$\text{diam}(\Gamma) + \int_0^\infty \int_G \hat{\beta}^r(x,r)^{2s} \, dx \frac{dr}{r^Q} \leq C \ell(\Gamma).$$

Note that the bound on the $\text{diam}(\Gamma)$ term is obvious. In fact, we may assume without loss of generality $\text{diam}(\Gamma) = 1$ as $\hat{\beta}_{\delta}(E)(\delta(x), \lambda r) = \hat{\beta}_E(x, r)$. Indeed, we have that $z \in E \cap B(x, r)$ if and only if $\delta(z) \in \delta_1(E) \cap B(\delta_1(x), \lambda r)$. Now given any horizontal line $L$, we get

$$\frac{d(\pi_i(\delta_1(z)), \pi_i(\delta_1(L)))}{\lambda r} = \frac{d(\delta_1(\pi_i(z)), \delta_1(\pi_i(L)))}{\lambda r} = \frac{d(\pi_i(z), \pi_i(L))}{r}.$$

For $g, h \in G$ and $t \in [0,1]$, define

$$L_{g,h}(t) := g \delta_1(\pi(g^{-1}h)).$$

This is a horizontal line segment that starts from $g$ and goes in the horizontal direction towards $h$ (although it may not hit $h$). We also write

$$\overline{g,h} = \{L_{g,h}(t)\}_{t \in [0,1]}.$$

We first prove the following simple bound.

**Lemma 5.2.** Let $p_1, p_2, \ldots \in G$ and $\ell > 0$ be so that $d_{HS}(p_i, p_j) \leq \ell$ for all $i, j$. Then for any $i \in \{1, \ldots, s\}$, we have

$$\sup_{t \in [0,1]} d_{HS}(\pi_i(L_{p_j, p_k}(t)), \pi_i(\overline{p_m, p_n})) \leq 2\ell, \quad \forall j, k, m, n.$$

**Proof.** As $\pi_i$ is $1$-Lipschitz, it suffices to prove the inequality without the $\pi_i$ present. As $N_{HS}(\pi(g)) \leq N_{HS}(g)$, we get for any $i, j$ that

$$d_{HS}(p_j, L_{p_j, p_k}(t)) = d_{HS}(0, \delta_1(\pi(p^{-1}_j p_k))) = tN_{HS}(\pi(p^{-1}_j p_k)) \leq t \ell \leq \ell$$

when $t \in [0,1]$. As $p_m \in \overline{p_m, p_n}$, we get

$$d_{HS}(L_{p_j, p_k}(t), \overline{p_m, p_n}) \leq d_{HS}(p_j, L_{p_j, p_k}(t)) + d_{HS}(p_j, p_m) \leq 2\ell.$$  \[\square\]
5.1 Stratified curvature

As mentioned in the introduction, we will be interested in proving inequalities for the triangle inequality excess:

$$\Delta_i(p_1, \ldots, p_k) = \sum_{j=1}^{k-1} d_{HS}(\pi_i(p_j), \pi_i(p_{j+1})) - d_{HS}(\pi_i(p_1), \pi_i(p_k)).$$

It is important to note that this quantity is not preserved in any good sense under a change of bi-Lipschitz equivalent metrics. Thus, most of the results in this subsection will only be valid for $d_{HS}$. However, these results are used to prove the necessary direction of the main theorem which is bi-Lipschitz invariant.

Given $\lambda \in (0,1)$, $\ell > 0$, and $i \in \{1, \ldots, s\}$, we define

$$A(\lambda, \ell, k) = \{(g, h) \in G \times G : N_{HS}(\pi_k(g)), N_{HS}(\pi_k(h)), N_{HS}(\pi_k(gh)) \in (\lambda \ell, \ell)\},$$

$$\overline{A}(\lambda, \ell, k) = \{(g, h) \in G \times G : N_{HS}(\pi_k(g)), N_{HS}(\pi_k(h)), N_{HS}(\pi_k(gh)) \in [\lambda \ell, \ell]\}.$$ 

Note that $A(\lambda, \ell, k)$ and $\overline{A}(\lambda, \ell, k)$ are open and compact in $G \times G$ respectively.

We will first need the following lemma from [5] that bounds $\Delta_s(p_1, p_2, p_3)$ from below when the points $p_i$ are sufficiently well-separated.

**Lemma 5.3.** For every $\lambda \in (0,1)$, there exists a constant $C > 0$ depending only on $d_{HS}$ and $\lambda$ so that if $p_1, p_2, p_3 \in G$ are so that $\lambda \ell \leq d_{HS}(p_i, p_j) \leq \ell$ when $i \neq j$ for some $\ell > 0$, then

$$\frac{d_{HS}(\pi_1(p_2), \pi_1(p_1, p_3))}{\ell} + \frac{NH(p_1^{-1} p_3)^{2s}}{\ell^{2s-1}} \leq C \Delta_s(p_1, p_2, p_3).$$

**Proof.** This is essentially [5, Lemma 3.4], but we will explain the conversion, as [5] uses slightly different notation. In [5], $r$ is the step of $G$, the constant $m \in [0,1]$ can be chosen arbitrarily as stated right before Lemma 3.4, and $\alpha$ is a constant depending only on the group and metric structure of $G$ as defined in [5, (3.1)]. The $\mathbb{R}^n$ is the first layer and so corresponds to our $\mathbb{R}^{n_1}$. The $x_1, y_1$ corresponds to our projections $\pi_1(x)$ and $\pi_1(y)$ in $\mathbb{R}^{n_1}$ and the $\| \cdot \|$ denotes the Hebisch–Sikora norm in the Carnot group. The $NH(g)$ in [5] is just the group term $\overline{\pi}(g)^{-1}g$, and so $\|NH(g)\|$ is equivalent to our $NH(g)$.

For $u \in \mathbb{R}^{n_1}$ the $\ell_u$ is the line segment in $\mathbb{R}^{n_1}$ from the origin to $u$. Note that this is not the same as our $\pi_1(p, q)$ as this line segment starts from $\pi_1(p)$ and goes to $\pi_1(q)$. However, we do have that

$$\pi_1(p, q) = \pi_1(p) + \ell \pi_1(p^{-1} q).$$

One now simply sets $x = p_1^{-1} p_2$ and $y = p_2^{-1} p_3$ and uses the fact that $x_1 + y_1 = \pi_1(p_1^{-1} p_3)$ (the left-hand side uses the projection notation of [5] but the right-hand side does not) to get that

$$\ell_{x_1 + y_1} - x_1 = (\pi_1(p_1) + \ell \pi_1(p_1^{-1} p_3)) - (\pi_1(p_1) + \pi_1(p_1^{-1} p_2)) = \pi_1(p_1) - \pi_1(p_2).$$
where we are interpreting $\ell_{x_1+y_1}$ as a set. Thus, we have that

$$d_{\mathbb{R}^n}(x_1, \ell_{x_1+y_1}) = d_{HS}(\pi_1(p_2), \pi_1(\overline{p_1}, \overline{p_2}))$$

where the left-hand side is written in the notation of [5].

To get our lemma, we first bound

$$NH(p_1^{-1} p_3) \leq N_{HS}(p_1^{-1} p_3) + N_{HS}(\tilde{\pi}(p_1^{-1} p_3)) \leq 2\ell,$$  \hspace{1cm} (8)

$$d_{HS}(\pi_1(p_2), \pi_1(\overline{p_1}, \overline{p_3})) \leq d_{HS}(\pi_1(p_2), \pi_1(p_1)) \leq \ell.$$  \hspace{1cm} (10)

As $d_{HS}(p_i, p_j) \geq \lambda \ell$, we get that the $A$ term in [5] is bounded by

$$A \leq \frac{m^2 s}{16} \max\{\alpha, 1\} (2\lambda^{1-2s} + (2\lambda)^{-1}) \ell$$

Thus, by choosing $m$ sufficiently small depending on $\alpha$ and $\lambda$, we get that $d_{HS}(p_i, p_j) \geq \lambda \ell \geq 4A$ for all $i \neq j$ for the hypothesis of the lemma. The conclusion of [5, Lemma 3.4] now gives our lemma (remembering again that $d(p_i, p_j) \geq \lambda \ell$).

We will need to augment this bound as it is not sufficient for our goals.

First, we will prove a lemma that says if the projection under $\pi_k$ of well-separated elements of $G$ are no longer well-separated, then the points must have a large $\Delta_s$.

**Lemma 5.4.** For each $\lambda \in (0, 1)$ and $k \in \{1, \ldots, s-1\}$ there exists a constant $C > 0$ depending also on $d_{HS}$ so that if $(u, v) \in A(\lambda, \ell', s) \setminus A(\lambda/2, \ell', k)$ then

$$N_{HS}(u) + N_{HS}(v) - N_{HS}(uv) \geq C\ell'.$$

**Proof.** Assume such a $C$ does not exist. Then one can find a sequence $(u_n, v_n)$ for which

$$N_{HS}(u_n) + N_{HS}(v_n) - N_{HS}(u_nv_n) \to 0.$$

As $A(\lambda, \ell', s) \setminus A(\lambda/2, \ell', k)$ is compact, we can pass to a limit pair $(u, v) \in \overline{A}(\lambda, \ell', s)$ for which

$$N_{HS}(u) + N_{HS}(v) - N_{HS}(uv) = 0.$$  \hspace{1cm} (11)

By Lemma 5.3, we then have that $NH(uv) = 0$. This means $uv$ is precisely the horizontal element $uv = (\pi_1(u) + \pi_1(v), 0, \ldots, 0)$ and that $N_{HS}(uv) = |\pi_1(u) + \pi_1(v)|$.

As the Hebisch–Sikora norm is 1-Lipschitz under $\pi_1$ and isometric only for horizontal elements, we have that $u$ and $v$ must be horizontal as otherwise (11) will be positive. However, this would mean that $N_{HS}(\pi_k(u)) = N_{HS}(u)$ and $N_{HS}(\pi_k(v)) = N_{HS}(v)$, which would mean $(u, v) \in A(\lambda/2, \ell', k)$, a contradiction.

We now start controlling the triangle inequality excess from below. In the following lemma, we seek to apply Lemma 5.3 to projections of points under $\pi_k$. However, if the projected points are
no longer well-separated, one of the requirements of Lemma 5.3, then we can apply Lemma 5.4 instead.

Lemma 5.5. For every \( \lambda \in (0,1) \) and \( k \in \{2, \ldots, s\} \), there exists a constant \( C > 0 \) depending only on \( d_{HS} \) and \( \lambda \) so that if \( p_1, p_2, p_3 \in G \) are so that \( \lambda \ell \leq d_{HS}(p_i, p_j) \leq \ell \) when \( i \neq j \) for some \( \ell > 0 \), then

\[
\frac{d_{HS}(\pi_1(p_2), \pi_1(p_1, p_3))^2}{\ell} + \frac{NH(\pi_k(p_1^{-1}p_3))^{2k}}{\ell^{2k-1}} \leq C(\Delta_k(p_1, p_2, p_3) + \Delta_s(p_1, p_2, p_3)).
\]

Proof. Let \( u = p_1^{-1}p_2 \) and \( v = p_2^{-1}p_3 \). By assumption, \((u, v) \in \overline{A}(\lambda, \ell, s)\). If \((u, v) \in A(\lambda/2, \ell, k)\), then we apply Lemma 5.3 to the points \( \pi_k(p_j) \) in the Carnot group \( \pi_k(G) \) to get an upper bound of \( C\Delta_k(p_1, p_2, p_3) \) for the left-hand side of the desired inequality.

Thus, we may now assume that \((u, v) \in \overline{A}(\lambda, \ell, s) \setminus A(\lambda/2, \ell, k)\). Then Lemma 5.4 tells us there is a constant \( C_0 > 0 \) so that

\[
\Delta_s(p_1, p_2, p_3) \geq C_0 \ell.
\]

If we show

\[
\frac{d_{HS}(\pi_1(p_2), \pi_1(p_1, p_3))^2}{\ell} + \frac{NH(\pi_k(p_1^{-1}p_3))^{2k}}{\ell^{2k-1}} \leq 2^{2k+1} \ell
\]

then the lemma follows with \( C = 2^{2k+1}C_0^{-1} \). This follows from (10) and a bound similar to (9), but using the fact that \( \pi_k \) is 1-Lipschitz. \( \square \)

We now combine the curvature bounds of Lemma 5.5 across all layers to get the following new stratified curvature bound. One can now see where the \( \hat{\beta} \) starts appearing.

Proposition 5.6. For any \( \lambda > 0 \), there exists a constant \( C > 0 \) depending only on \( d_{HS} \), \( \lambda \), and Euclidean structure \( | \cdot | \) so that the following holds. Let \( p_1, p_2, p_3, p_4 \in G \) be points so that \( \lambda \ell \leq d_{HS}(p_i, p_j) \leq \ell \) if \( i \neq j \) for some \( \ell > 0 \). Then

\[
\sum_{i=1}^{s} \sup_{t \in [0,1]} \frac{d_{HS}(\pi_1(L_{p_1,p_3}(t)), \pi_1(p_i, p_4))^{2i}}{\ell^{2i-1}} + \sup_{t \in [0,1]} \frac{d_{HS}(\pi_1(L_{p_3,p_4}(t)), \pi_1(p_1, p_4))^{2i}}{\ell^{2i-1}} \leq C\Delta,
\]

where \( \Delta = \sum_{i=2}^{s} \Delta_i(p_1, p_2, p_3, p_4) \).

It may seem strange that \( p_2 \) does not appear in the left-hand side of the conclusion but does for the right. The \( p_2 \) will come from applying Lemma 5.5 to the points \( p_1, p_2, p_3 \) to get a bound on \( NH(p_1^{-1}p_3) \). This, combined with bounds on \( d_{HS}(\pi_1(p_3), \pi_1(p_1, p_4)) \) that we get from applying Lemma 5.5 to \( p_1, p_3, p_4 \) allows us to bound how far \( p_3 \) is to the horizontal line segment \( p_1, p_4 \) so that we can get the second term bound on \( L_{p_3,p_4} \) via Corollary 4.5. This second term bound is required in the proof of Corollary 5.7 that follows this proposition.
Note that it is impossible to derive such a bound using Lemma 5.5 directly on \( p_1, p_3, p_4 \) as it only bounds the behavior of the projection \( \pi_1(p_3) \). To achieve this, we would need a stronger version of Lemma 5.3, but we do not seek to do so as it is not necessary.

**Proof.** We may translate and dilate the setting so that \( p_1 = 0 \) and \( \ell' = 1 \). Lemma 5.2 gives that the left-hand side is then bounded by \( s^{2l+1} \). Fix a \( \mu > 0 \) to be determined. If \( \Delta \geq \mu \), then we have that the inequality holds for \( C = s^{2l+1}/\mu \). Thus, we may suppose \( \Delta < \mu \).

We first bound the \( L_{p_1,p_3} \) term. We start by applying Lemma 5.3 to the triple \( p_1, p_3, p_4 \) to get (with some overestimating) that

\[
d_{HS}(\pi_1(p_3), \pi_1(0, p_4))^2 \leq C_0 \Delta.
\]

Choose \( u, v \in B_{\mathbb{R}^n}(1) \) so that \( \pi_1(p_3) = u \) and \( \pi_1(p_4) = v \). Recall that we assumed that \( \Delta < \mu \) for some yet to be determined \( \mu \). We now choose \( \mu \) sufficiently small so that \( \sqrt{C_0\mu} < 1 \). Recalling that \( d_{HS} \) on \( G_1 \) is equivalent to the Euclidean metric \( | \cdot | \), we see that there is a \( \gamma \in [0, 1] \) so that \( |u - \gamma v| \leq \sqrt{C_0\Delta} < 1 \). Applying Corollary 4.5 to \( g = h = p_1 \) with vectors \( u \) and \( \gamma v \), we get that there exists \( C > 0 \) so that

\[
\sum_{i=1}^{s} \sup_{t \in [0,1]} d_{HS}(\pi_i(L_{0,p_3}(t)), \pi_i(L_{0,p_4}(\gamma t)))^{2i} \leq C \Delta. \quad (13)
\]

As \( \gamma \in [0, 1] \), we get that \( L_{0,p_4}(\gamma t) \) is contained in \( 0, p_4 \) and, remembering that \( p_1 = 0 \), we get

\[
\sum_{i=1}^{s} \sup_{t \in [0,1]} d_{HS}(\pi_i(L_{p_1,p_3}(t)), \pi_i(0, p_4))^{2i} \leq C \Delta, \quad (14)
\]

which takes care of the \( L_{p_1,p_3} \) term in the summand.

As for the \( L_{p_3,p_4} \) term, we again seek to apply Corollary 4.5, but we first need to verify that \( p_3 \) lies close to \((\gamma v, 0, \ldots, 0) \in p_1, p_4 \). To do so, we first take \( t = 1 \) in (13) to get that \( \pi(p_3) = L_{p_1,p_3}(1) \) is close to \((\gamma v, 0, \ldots, 0) \) in the following way:

\[
\sum_{i=1}^{s} d_{HS}(\pi_i(\pi(p_3)), \pi_i(\gamma v, 0, \ldots, 0))^{2i} \leq C \Delta. \quad (15)
\]

Here, we used the fact that \( \pi_1(p_3) = v \) and the fact that \( p_1 = 0, \ell' = 1 \). Applying Lemma 5.5 (for many \( i \)) on the triple \( p_1, p_2, p_3 \), we get

\[
\sum_{i=2}^{s} NH(\pi_i(p_3))^{2i} \leq sC_0 \Delta.
\]

Using the definition of \( NH \) and the fact that \( \pi(p_3) = \pi_i(\pi_3(p)) \) for all \( i \), we then get that \( p_3 \) is close to \( \pi(p_3) \) in the following way:

\[
\sum_{i=1}^{s} d_{HS}(\pi_i(p_3), \pi_i(\pi(p_3)))^{2i} \leq sC_0 \Delta. \quad (16)
\]
We now combine (15) and (16) to get that \( p_3 \) is close to \((\gamma v, 0, \ldots, 0)\) in the following way:

\[
\sum_{i=1}^{s} d_{HS}(\pi_i(p_3), \pi_i(\gamma v, 0, \ldots, 0))^{2i} \leq \sqrt{s} \sum_{i=1}^{s} d_{HS}(\pi_i(p_3), \pi_i(\gamma v, 0, \ldots, 0))^{2i} \leq \sqrt{s} 2^{s-1}(sC_0 + C_1)\Delta.
\]

Now an application of Lemma 4.4 gives that \( p_3 \in (\gamma v, 0, \ldots, 0) \cdot B_{\mathbb{R}^n}(C_2\sqrt{\Delta}) \) for some \( C_2 \) depending only on \( G \).

We can finally apply Corollary 4.5 with points \( p_3 \) and \((\gamma v, 0, \ldots, 0)\) and vectors \( v - u \) and \((1 - \gamma)v\) as we did to get (13) and (14) to get that

\[
\sum_{i=1}^{s} \sup_{t \in [0,1]} d_{HS}(\pi_i(L_{p_3, p_4}(t)), \pi_i(p_1, p_4))^{2i} \leq C_3\Delta
\]

for some \( C_3 > 0 \). The proposition now follows from (14) and (17). \( \square \)

Note that Proposition 5.6 can only bound the drift between horizontal line segments that share some endpoint (namely either \( p_1 \) or \( p_4 \)). In the sequel, we will also be considering pairs of line segments that do not necessarily share an endpoint. We now combine Lemma 5.5 and Proposition 5.6 to derive the following useful corollary, which allows us to bound horizontal line segments that do not share any endpoints.

**Corollary 5.7.** For any \( \lambda > 0 \), there exists a constant \( C > 0 \) depending only on \( d_{HS}, \lambda, \) and the Euclidean structure \( | \cdot | \) so that the following holds. Let \( p_1, p_2, p_3, p_4, p_5 \in G \) be points so that \( \lambda \ell \leq d_{HS}(p_i, p_j) \leq \ell \) if \( i \neq j \) for some \( \ell > 0 \). Then

\[
\sum_{k=1}^{s} \sup_{t \in [0,1]} \frac{d_{HS}(\pi_k(L_{p_3, p_4}(t)), \pi_k(p_1, p_5))^{2k}}{\ell^{2k-1}} \leq C\Delta,
\]

where \( \Delta = \sum_{k=2}^{s} \Delta_k(p_1, p_2, p_3, p_4, p_5) \).

Like Proposition 5.6, the presence of \( p_2 \) on the right-hand side cannot be removed without a stronger version of Lemma 5.3.

**Proof.** We dilate the setting so that \( \ell' = 1 \). By Lemma 5.2, the left-hand side of our desired inequality is then bounded by \( s2^{2s} \). Define

\[
\Delta^{(1)} = \sum_{k=2}^{s} \Delta_k(p_1, p_4, p_5),
\]

\[
\Delta^{(2)} = \sum_{k=2}^{s} \Delta_k(p_1, p_2, p_3, p_4).
\]

Then \( \Delta^{(1)}, \Delta^{(2)} \geq 0 \) and \( \Delta^{(1)} + \Delta^{(2)} = \Delta \). Let \( C_0 \) be the constant from Lemma 5.5. If \( C_0 \Delta^{(1)} \geq 1 \), then the corollary is satisfied by taking \( C = s2^{2s}C_0 \). Thus, we may assume \( C_0 \Delta^{(1)} < 1 \).
We begin by showing $L_{p_1,p_4}$ is close to $p_1, p_5$. We first apply Lemma 5.5 to $p_1, p_4, p_5$ and use the fact that $\ell' = 1$ to get

$$d_{HS}(\pi_1(p_4), \pi_1(p_1, p_5))^2 \leq C_0 \Delta^{(1)}. \tag{18}$$

Let $u = \pi_1(p_1^{-1} p_4)$ and $v = \pi_1(p_1^{-1} p_5)$. By (18), there is some $\gamma \in [0, 1]$ so that $|u - \gamma v| \leq \sqrt{C_0 \Delta^{(1)}} < 1$. We now apply Corollary 4.5 to $g = h = p_1$ and vectors $u, \gamma v$ with $\ell = 1$ so that

$$\max_{1 \leq k \leq s} \sup_{t \in [0,1]} d_{HS}(\pi_k(L_{p_1,p_4}(t)), \pi_k(p_1, p_5))^{2k} \leq C_1 \Delta^{(1)} \leq C_1 \Delta,$$

which gives

$$\sum_{k=1}^{s} \sup_{t \in [0,1]} d_{HS}(\pi_k(L_{p_1,p_4}(t)), \pi_k(p_1, p_5))^{2k} \leq sC_1 \Delta.$$

We now show that $L_{p_3,p_4}$ is close to $p_1, p_4$ by applying Proposition 5.6 to $p_1, p_2, p_3, p_4$ to get

$$\sum_{k=1}^{s} \sup_{t \in [0,1]} d_{HS}(\pi_k(L_{p_3,p_4}(t)), \pi_k(p_1, p_4))^{2k} \leq C_2 \Delta^{(2)} \leq C_2 \Delta.$$

The last two inequalities, together with the triangle inequality and the inequality $(a + b)^p \leq 2^{p-1}(a^p + b^p)$ give the corollary.

\[\square\]

### 5.2 Balls, filtrations, and multiresolutions

In this subsection, we will discretize the Carleson integral and reduce to a parametric setting. Most of the setup will resemble [14, sections 2.2–2.5] which go over the construction at much greater detail, including choices of parameters which still work in our setting. We thus refer the reader to those sections for complete details.

In this section, we will endow $G$ with $d = d_\infty$. Thus, all balls and diameters will be defined in terms of this metric. The $d_\infty$ gives us a way to prove Lemma 5.15 but is otherwise inconsequential. We do not know if Lemma 5.15 can be proven using $d_{HS}$. Given a ball $B = B(x, r)$ and $\lambda > 0$, we define $\lambda B := B(x, \lambda r)$.

Let $\Gamma$ be a rectifiable curve in $G$ with $\text{diam}(\Gamma) = 1$. We may parameterize $\Gamma$ via a surjective 1-Lipschitz function $\gamma : \mathbb{T} \to \Gamma$ where $\mathbb{T}$ is a circle in $\mathbb{R}^2$ with circumference $CH^1(\Gamma)$ where $C > 0$ depends on the metric $d$. Indeed, this follows from [19, the appendix] (see especially Lemma 3.7 of the appendix). Note that although [19] is proven for a geodesic metric on the Hilbert space, the proof also works for quasi-convex metrics on generic metric spaces by using (6) to connect points. The Lipschitz bound on the parameterization will change depending on the quasi-convexity constant.

We fix a direction of flow along $\mathbb{T}$. By a subarc $\tau$ of $\gamma$, we mean a restriction of $\gamma$ onto a subinterval of $\mathbb{T}$. We define set relations between two subarcs $\tau$ and $\zeta$ by using the set relation for their defining intervals. However, the diameter of a subarc is the diameter of the image.
For each $n \in \mathbb{Z}$, let $Y_n \subset \Gamma$ be a maximal $2^{-n}$ separated net\footnote{Recall an $\alpha$-separated net of $X$ is a subset $A \subset X$ for which $d(x, y) \geq \alpha$ for all $x, y \in A$.} and define the following multiresolution families of $\Gamma$

$$
\hat{\mathcal{G}} = \{B(x, 10 \cdot 2^{-n}) : x \in Y_n, n \in \mathbb{Z}\},
$$

$$
\mathcal{G} = \{B(x, 10 \cdot 2^{-n}) : x \in Y_n, n \geq 0\}.
$$

As $\text{diam}(\Gamma) = 1$, each $Y_n$ is then a singleton set when $n < 0$. Given a ball $B = B(x, r)$, we define $\hat{\beta}_I(B) := \hat{\beta}_I(x, r)$.

A straightforward modification of [14, Lemma 2.6] shows that to prove Theorem 5.1, it suffices to prove

$$
\sum_{B \in \hat{\mathcal{G}}} \hat{\beta}_I(B)^{2s} \text{diam}(B) \leq C H^1(\Gamma).
$$

**Lemma 5.8.** There exists a constant $C > 0$ depending only on the metric structure of $G$ so that

$$
\sum_{B \in \hat{\mathcal{G}} \setminus \mathcal{G}} \hat{\beta}_I(B)^{2s} \text{diam}(B) \leq C H^1(\Gamma).
$$

**Proof.** By picking any horizontal line $L$ that intersects $\Gamma$, we have for $B = B(x, 10 \cdot 2^{-n}) \in \hat{\mathcal{G}} \setminus \mathcal{G}$ that

$$
\hat{\beta}_I(B)^{2s} \leq \sum_{i=1}^{s} \left(\frac{\text{diam}(\Gamma)}{10 \cdot 2^{-n}}\right)^{2i} \leq s 2^{2n}.
$$

Recall that $\#Y_n = 1$ when $n < 0$. Thus,

$$
\sum_{B \in \hat{\mathcal{G}} \setminus \mathcal{G}} \hat{\beta}_I(B)^{2s} \text{diam}(B) \leq \sum_{n < 0} s 2^{2n} \cdot 20 \cdot 2^{-n} \leq 40s \text{diam}(\Gamma) \leq 40s H^1(\Gamma).
$$

By the previous lemma, we see that it now suffices to prove that

$$
\sum_{B \in \mathcal{G}} \hat{\beta}_I(B)^{2s} \text{diam}(B) \leq C H^1(\Gamma). \quad (19)
$$

This is our goal for the remainder of the section.

Fix $J \in \mathbb{N}$. We decompose $B := \{2B : B \in \mathcal{G}\}$ into $D'$ families $\{B_i\}$ where for any two distinct balls $B_1, B_2 \in B_i$,

1. $r(B_1)/r(B_2) \in 2^J \mathbb{Z}$,
2. if $r(B_1) = r(B_2) = r$, then $d(B_1, B_2) > 3r$.

In the second condition, we use the usual notion of distance between sets:

$$
d(B_1, B_2) = \inf_{x \in B_1, y \in B_2} d(x, y).
$$

By [14, Lemma 2.11] (where we fix $C = 20$ and $\kappa = 3$), we have that $D'$ is bounded.
Lemma 5.9 [14, Lemma 2.11]. We may take $D' = D \cdot J$ where $D$ is some constant depending only on $G$.

We will do the decomposition with $J = 100$ so that $D'$ is a constant depending only on $G$. Let $\mathcal{C}' \cup ... \cup \mathcal{C}'_{D'}$ denote the decomposition of $\mathcal{C}$ that follows from the decomposition $B_1 \cup ... \cup B_{D'}$. We then apply [14, Lemma 2.12] with $J$ and $k = 3$ to each $B_i$, to produce a family of open ‘dyadic cubes’ $\{\Delta(B,i)\}_{i=1}^{D'}$ so that

1. For all $B \in \mathcal{C}'_i$, there exists $Q = Q(B) \in \Delta(B,i)$ for which $2B \subset Q \subset 2(1 + 2^{-J+2})B$,
2. If $B, B' \in \mathcal{C}'_i$ is so that $r(B) < r(B')$ and $Q(B) \cap Q(B') \neq \emptyset$, then $Q(B') \supseteq Q(B)$,
3. If $B, B' \in \mathcal{C}'_i$ is so that $r = r(B) = r(B')$ and $B \neq B'$, then $d(Q(B), Q(B')) > 2r$.

Note that properties (1) and (3), together with the $2^J$ separation of scales of $\mathcal{C}'_i$, show that the map $B \mapsto Q(B)$ is injective.

For each $B \in \mathcal{C}'_i$, we define

$$\Lambda(B) = \{ \tau = \gamma|_I : I \subset \mathbb{T} \text{ is a connected component of } \gamma^{-1}(\Gamma \cap Q(B)) \text{ and } \gamma(I) \cap B \neq \emptyset \}.$$ 

Thus, $\Lambda(B)$ is a collection of subarcs $\gamma|_I$.

Lemma 2.17 of [14] tells us that each $\mathcal{F}^{0,i} = \bigcup_{B \in \mathcal{C}'_i} \Lambda(B)$ is a prefiltration, that is, for each $i$ there exists some $L_i > 0$ so that one has the decomposition $\mathcal{F}^{0,i} = \bigcup_{n \geq 1} \mathcal{F}^{0,i}_n$ into collections of arcs such that

1. For any $\tau \in \mathcal{F}^{0,i}_n$, $L_i 2^{-nJ} \leq \text{diam}(\tau) \leq L_i 2^{-nJ+3}$,
2. For distinct $\tau_1, \tau_2 \in \mathcal{F}^{0,i}_n$, we must have $\tau_1 \cap \tau_2 = \emptyset$,
3. If $\tau \in \mathcal{F}^{0,i}_n$ and $\zeta \in \mathcal{F}^{0,i}_{n+k}$ for $k > 0$, then either $\tau \cap \zeta = \emptyset$ or $\zeta \subseteq \tau$.

Recall that $J = 100$ is the same constant as before and set relation for arcs is defined by the set relation of their defining intervals. We will not really deal with prefiltrations. Instead, we use [14, Lemma 2.13] to construct, for each prefiltration $\mathcal{F}^0$ in $\{\mathcal{F}^{0,i}\}_i$, a collection of arcs $\mathcal{F} = \bigcup_n \mathcal{F}_n$ so that

1. For any $\tau \in \mathcal{F}_n$, $L_i 2^{-100n-10} \leq \text{diam}(\tau) \leq L_i 2^{-100n+4}$,
2. For $\tau_1, \tau_2 \in \mathcal{F}_n$, they are either disjoint, identical, or intersect in one or both of their endpoints,
3. For $\zeta \in \mathcal{F}_{n+1}$, there is a unique element $\tau \in \mathcal{F}_n$ so that $\zeta \subseteq \tau$,
4. For all $n$, $\bigcup_{\tau \in \mathcal{F}_n} \tau = \mathbb{T}$,
5. For each $\tau^0 \in \mathcal{F}^0_n$, there is an element $\tau \in \mathcal{F}_n$ so that $\tau \supseteq \tau^0$, and moreover if $I_0$ and $I$ are domains of $\tau^0$ and $\tau$, respectively, then the image under $\gamma$ of any connected component of $I \setminus I_0$ has diameter no more than $L_i 2^{-100n-10}$. Recall that $I, I_0$ are intervals,
6. Different $\tau_0, \tau_1 \in \mathcal{F}^0_n$ give rise to different arcs in $\mathcal{F}_n$.

By property (4), we are saying that the union of the domains of the subarcs is the entirety of $\mathbb{T}$.

Such collections of arcs are called filtrations, so we get $D'$ different filtrations. Given a subarc $\tau$, we let $I_{\tau} \subseteq \mathbb{T}$ denote its parameterizing space.

To recap, we first construct our multiresolution family $\mathcal{C}$ from the maximal $2^{-n}$-separated nets $\mathcal{Y}_n$. We then decomposed $\mathcal{C}$ into $D'$ families of balls $\mathcal{C}'_i$, each with good scale separation. For each $\mathcal{C}'_i$, we constructed families of dyadic cubes $\Delta(B,i)$ which we used to get our prefiltrations $\mathcal{F}^{0,i}$. Finally, we used each $\mathcal{F}^{0,i}$ to construct a family of filtrations $\mathcal{F}_i$. 
As properties (2) and (5) of filtrations tells us that each $\tau \in F^0_j$ gives rise to some unique $\tau' \in F^i$ of the same ‘generation,’ we define

$$\Lambda'(B) = \{\tau' : \tau \in \Lambda(B)\}.$$  

We record the following lemma which controls the diameter of arcs in filtrations relative to the ball they come from.

**Lemma 5.10.** If $\tau \in \Lambda'(B)$, then

$$\frac{1}{2} \text{diam}(B) \leq \text{diam}(\tau) \leq \text{diam}(3B) = 3 \text{diam}(B). \quad (20)$$

**Proof.** As $\lambda B = x\delta_\lambda(x^{-1}B)$ where $x$ is the center of $B$, we get that $\text{diam}(\lambda B) = \lambda \text{diam}(B)$ when $\text{diam}$ is computed with any metric that scales with dilation.

Let $\tau^0 \in \Lambda(B)$ be the subarc that gave rise to $\tau$ so that $\tau \supseteq \tau^0$. We have

$$\frac{1}{2} \text{diam}(B) \leq \text{diam}(\tau^0) \leq 2(1 + 2^{-98}) \text{diam}(B).$$

The upper bound comes from the fact that the image of $\tau^0$ lies in $Q(B)$ and property (1) of cubes. The lower bound comes from the fact that $\tau^0$ goes from inside $B$ to the boundary of $Q(B) \supset 2B$. The lower bound on $\text{diam}(\tau)$ now follows from the fact that $\tau$ contains $\tau^0$.

We now derive the upper bound. Let $B \in G^j$. We first note that property (1) of prefiltrations and our upper bound on $\tau^0$ gives for some minimal $n$ that

$$L_i2^{-100n} \leq \text{diam}(\tau^0) \leq 2(1 + 2^{-98}) \text{diam}(B)$$

Now let $I, I^0$ be the domains of $\tau$ and $\tau^0$, and let $I'$ be a connected interval of $I \setminus I^0$. By property (5) of filtrations, we have

$$\text{diam}(\tau\vert_{I'}) \leq L_i2^{-100n-10} < \frac{1}{4} \text{diam}(B).$$

Finally, as the image of $\tau\vert_{I^0} = \tau^0$ lies in $Q(B) \subseteq 2(1 + 2^{-98})B$, we get the upper bound of $\text{diam}(3B)$ on $\text{diam}(\tau)$. \qed

Given a curve $\tau$, we let $a(\tau)$ and $b(\tau)$ denote the endpoints of the domain where $a(\tau) < b(\tau)$. We also define

$$L_\tau := \gamma(a(\tau)), \gamma(b(\tau)),$$

and

$$\beta_i(\tau) := \sup_{t \in L_\tau} \frac{d(\pi_i(\gamma(t)), \pi_i(L_\tau))}{\text{diam}(\tau)}, \quad i \in \{1, \ldots, s\},$$

$$\hat{\beta}(\tau)^{2s} := \sum_{i=1}^{s} \beta_i(\tau)^{2i}.$$
We define
\[ G_2 := \{ B \in G : \tilde{\beta}(\tau) < \varepsilon_0 \tilde{\beta}_1(B) \forall \tau \in \Lambda'(B) \}. \]
\[ G_1 := G \setminus G_2. \]

Here \( \varepsilon_0 > 0 \) will be a sufficiently small constant that we choose in Subsection 5.4.

Balls of \( G_1 \) have one curve whose ‘wiggyness’ captures most of \( \tilde{\beta}_1(B) \) whereas curves of \( G_2 \) are mostly flat, but it is the combination of two or more of these curves that gives \( \tilde{\beta}_1(B) \). See Figure 1 for representative images. We call the families \( G_1 \) and \( G_2 \) non-flat and flat balls, respectively.

To prove (19), it now suffices to prove
\[ \sum_{B \in G_1} \tilde{\beta}_1(B)^{2s} \text{diam}(B) \leq C \mathcal{H}^1(\Gamma), \quad (21) \]
\[ \sum_{B \in G_2} \tilde{\beta}_1(B)^{2s} \text{diam}(B) \leq C \mathcal{H}^1(\Gamma). \quad (22) \]

We will prove (21) in the following subsection and (22) in the subsection after.

### 5.3 Non-flat balls

The main goal of this subsection will be to prove the following proposition.

**Proposition 5.11.** There exists a \( C > 0 \) depending only on \( d \) so that for any filtration \( F \) and \( i \in \{1, \ldots, s\} \), we have
\[ \sum_{\tau \in F} \beta_i(\tau)^{2i} \text{diam}(\tau) \leq C \mathcal{H}^1(\Gamma). \]

Assuming this proposition, one can now derive (21).
Corollary 5.12.  
\[ \sum_{B \in G_1} \hat{\beta}_1(B)^{2s} \text{diam}(B) \leq C \mathcal{H}^1(\Gamma). \]

Proof. For each \( B \in G_1 \), let \( \tau_B \in \Lambda_1(B) \) denote a curve so that \( \hat{\beta}(\tau_B) \geq \varepsilon_0 \hat{\beta}_1(B) \). Recall that \( G \) was decomposed into \( D' \) families \( \{G'_1\}_{i=1}^{D'} \), which thus also partitions \( G_1 \) into \( \{G'_1\}_{i=1}^{D'} \). Let \( F_i \) denote the filtration constructed using \( G'_i \). By (20), \( \text{diam}(\tau_B) \geq \frac{1}{2} \text{diam}(B) \). As \( D' \) is a constant depending only on \( G \), we now get

\[ \sum_{B \in G_1} \hat{\beta}_1(B)^{2s} \text{diam}(B) = \sum_{i=1}^{D'} \sum_{B \in G'_i} \hat{\beta}_1(B)^{2s} \text{diam}(B) \leq \frac{2}{\varepsilon_0} \sum_{i=1}^{D'} \sum_{B \in G'_i} \hat{\beta}(\tau_B)^{2s} \text{diam}(\tau_B) \]
\[ \leq \frac{2}{\varepsilon_0^{2s}} \sum_{i=1}^{D'} \sum_{\tau \in F_i} \hat{\beta}(\tau)^{2s} \text{diam}(\tau) = \frac{2}{\varepsilon_0^{2s}} \sum_{i=1}^{D'} \sum_{\tau \in F_i} \sum_{j=1}^{s} \hat{\beta}_j(\tau)^{2j} \text{diam}(\tau) \leq C \mathcal{H}^1(\Gamma). \]

We now start working on Proposition 5.11. We will now view \( i \in \{1, \ldots, s\} \) and the filtration \( F \) as fixed. For simplicity, we will assume for this filtration \( F \) that \( L = 1 \) in property (1) of filtrations. Indeed, one could just first scale everything by \( L^{-1} \), prove the results of this subsection, and then rescale back.

We define for \( \tau \in F_n \) and \( k \in \mathbb{N} \) the families

\[ F_{\tau, k} := \{\tau' \in F_{n+k} : \tau' \subset \tau\}. \]

Given \( \tau \in F \), we further define

\[ d_\tau := \max_{\tau' \in F_{\tau, 1}} \sup_{z \in \pi_i(L_{\tau'})} d(z, \pi_i(L_{\tau})). \]

Lemma 5.13.

\[ \sum_{\tau \in F} \frac{d_\tau^{2i}}{\text{diam}(\tau)^{2i-1}} \leq C \mathcal{H}^1(\Gamma). \]  \( (23) \)

Proof. Although the quantities \( d_\tau \), \( \text{diam} \), and \( \mathcal{H}^1 \) are all calculated with respect to \( d = d_\infty \), it suffices to prove (23) with \( d = d_{HS} \) instead. Indeed, a bi-Lipschitz change of metric affects all quantities by only a multiplicative constant. Thus, for the rest of this proof, we will assume \( \text{diam} \), \( d_\tau \), and \( \mathcal{H}^1 \) are with respect to \( d_{HS} \).

As \( d_{HS} \) and \( d \) are bi-Lipschitz, property (1) of filtrations gives that there exists a \( C_0 > 1 \) so that

\[ \frac{1}{C_0} 2^{-100k} \leq \text{diam}(\tau) \leq C_0 2^{-100k}, \quad \forall \tau \in F_k. \]  \( (24) \)
Define \( M = \lceil 3 + \log_2 C_0 \rceil \). We have chosen \( M \) so that for all \( k \)

\[
C_0 2^{-100(k+M)} \leq C_0^{-1} 2^{-202-100k}. \tag{25}
\]

For \( \ell' \in \{2, \ldots, s\} \), define

\[
\Delta_{\ell'}(\tau) := \left( \sum_{\tau' \in \mathcal{P}_{\tau,M}} d_{HS}(\pi_{\ell'}(\gamma(a(\tau'))), \pi_{\ell'}(\gamma(b(\tau')))) \right) - d_{HS}(\pi_{\ell'}(\gamma(a(\tau))), \pi_{\ell'}(\gamma(b(\tau)))).
\]

It suffices to prove that

\[
\frac{d_{2i}^{2i}}{\text{diam}(\tau)^{2i-1}} \leq C \sum_{\ell=2}^{s} \Delta_{\ell'}(\tau). \tag{26}
\]

Indeed, summing both sides over \( \tau \in \mathcal{P} \) gives

\[
\sum_{\tau \in \mathcal{P}} \frac{d_{2i}^{2i}}{\text{diam}(\tau)^{2i-1}} \leq C \sum_{\ell=2}^{s} \sum_{\tau \in \mathcal{P}} \Delta_{\ell'}(\tau) \leq sCMH^1(\Gamma)
\]

where we used the fact that \( \sum_{\ell} \Delta_{\ell'}(\tau) \leq MH^1(\Gamma) \) for each \( \ell' \) as \( \pi_{\ell'} \) is 1-Lipschitz and the summation of differences telescopes.

Fix a \( \tau \in \mathcal{P}_k \) and let \( \{\tau_j\}_{j=1}^m \) denote the subarcs of \( \mathcal{P}_{\tau,1} \) in order of the flow along \( \bar{\tau} \) so that \( a(\tau_1) = a(\tau) \) and \( b(\tau_m) = b(\tau) \). Let \( P \) denote the images under \( \gamma \) of all the endpoints of subarcs of \( \mathcal{P}_{\tau,2} \). We will want to apply the results of the last subsection to points from \( P \). It may seem strange that we are taking the endpoints of \( \mathcal{P}_{\tau,2} \) instead of endpoints of \( \mathcal{P}_{\tau,1} \), but this is necessary as we will indicate below.

Let us first assume \( P \) is a \( C_0^{-1} 2^{-1000-100k} \)-separated set (in \( d_{HS} \)). Note that we also have

\[
d_{HS}(p, q) \leq \text{diam}(\tau), \quad \forall p, q \in P. \tag{27}
\]

We will bound

\[
\sup_{z \in \pi_\ell(L_{\tau_j})} \frac{d_{HS}(z, \pi_\ell(L_{\tau_j}))^{2i}}{\text{diam}(\tau)^{2i-1}}
\]

for different \( j \)s and take the maximal of these bounds to get our needed result. There will be three cases, \( j = 1, j = m \), and the rest.

First we consider when \( j \in \{2, \ldots, m-1\} \). Pick \( u \), any endpoint of a subarc of \( \mathcal{P}_{\tau_1,1} \) that is not also an endpoint of \( \tau_1 \). It then follows that \( u \) lies in \( P \). Note that it is here that we used the fact that \( P \) contains endpoints of \( \mathcal{P}_{\tau,2} \). We apply Corollary 5.7 to the collection of points \( \gamma(a(\tau)), \gamma(u), \gamma(a(\tau_j)), \gamma(b(\tau_j)), \gamma(b(\tau)) \in P \) with \( \ell = \text{diam}(\tau) \leq C_0 2^{-100k} \) and \( \lambda = C_0^{-2} 2^{-1000} \) to get

\[
\sup_{z \in \pi_\ell(L_{\tau_j})} \frac{d_{HS}(z, \pi_\ell(L_{\tau_j}))^{2i}}{\text{diam}(\tau)^{2i-1}} \leq C \sum_{\ell=2}^{s} \Delta_{\ell'}(\tau), \tag{28}
\]

where \( u \) plays the role of \( p_2 \) in Corollary 5.7.
For $j = 1$, we apply Proposition 5.6 to the points $\gamma(a(\tau)), \gamma(u), \gamma(b(\tau_1)), \gamma(b(\tau))$ using the same $u, \ell, \lambda$ as before to get

$$
\sup_{z \in \pi_i(L_\tau)} \frac{d_{HS}(z, \pi_i(L_\tau))^{2i}}{\text{diam}(\tau)^{2i-1}} \leq C \sum_{\ell = 2}^{s} \Delta_\ell(\tau).
$$

(29)

For $j = m$, we apply Proposition 5.6 to the points $\gamma(a(\tau)), \gamma(a(\tau_{m-1})), \gamma(a(\tau_m)), \gamma(b(\tau_m))$ to get

$$
\sup_{z \in \pi_i(L_{\tau_m})} \frac{d_{HS}(z, \pi_i(L_{\tau}))^{2i}}{\text{diam}(\tau)^{2i-1}} \leq C \sum_{\ell = 2}^{s} \Delta_\ell(\tau).
$$

(30)

The inequalities (28), (29), (30) give (26) in our current case where $P$ is $C^{-1} 2^{-1000-100k}$-separated.

Thus, we now assume that $P$ is not $C^{-1} 2^{-1000-100k}$-separated, and let $u, v$ be the endpoint of two subarcs of $F_{\tau,2}$ (where $u < v$ based on the flow of $\mathbb{T}$) so that $d_{HS}(\gamma(u), \gamma(v)) < C^{-1} 2^{-1000-100k}$. Let $\xi$ be any subarc of $F_{\tau,2} \subset F_{k+2}$ that lies between $u$ and $v$. As $\text{diam}(\xi) \geq C^{-1} 2^{-100(k+2)}$, there then exists some $w \in \xi$ so that

$$
d_{HS}(\gamma(w), \{\gamma(u), \gamma(v)\}) \geq \text{diam}(\xi) - d_{HS}(\gamma(w), \gamma(v)) \geq C^{-1} 2^{-201-100k}.
$$

This $w$ must lie in some subarc $\zeta \in F_{\xi, M-2} \subset F_{\tau, M}$ and as $\text{diam}(\xi) \leq C_0 2^{-100(k+M)}$, we get that if $z$ is an endpoint of $\zeta$, then

$$
d_{HS}(\gamma(z), \{\gamma(u), \gamma(v)\}) \geq d_{HS}(\gamma(w), \{\gamma(u), \gamma(v)\}) - \text{diam}(\zeta)
\geq C_0 2^{-201-100k} - C_0 2^{-100(k+M)} \geq C_0 2^{-202-100k}.
$$

Thus,

$$
d_{HS}(\gamma(u), \gamma(z)) + d_{HS}(\gamma(z), \gamma(v)) - d_{HS}(\gamma(u), \gamma(v)) > C_0^{-1} 2^{-203-100k}.
$$

As $z$ is an endpoint of a subarc of $F_{\tau, M}$ and lies between $u, v$ which are endpoints of subarcs of $F_{\tau,2}$, we get that by a simple triangle inequality argument that

$$
C_0^{-1} 2^{-203-100k} \leq d_{HS}(\gamma(u), \gamma(z)) + d_{HS}(\gamma(z), \gamma(v)) - d_{HS}(\gamma(u), \gamma(v)) \leq \Delta_\ell(\tau).
$$

Finally, Lemma 5.2 and (27) gives us that

$$
\frac{d_{HS}(\gamma(u), \gamma(z))}{\text{diam}(\tau)^{2i+1}} \leq 2^{2i} \text{diam}(\tau) \leq C_0 2^{28-100k} \text{ which, with the previous inequality, gives our result.}
$$

Given $\tau \in F$, we define $\{\tau_j\}_{j=0}^\infty$ inductively as $\tau_0 = \tau$ and $\tau_j \in F_{\tau, j}$ so that $d_{\tau_k}$ is maximal for all arcs in $F_{\tau, j}$. We now show that the $d_{\tau_k}$ controls the $\beta_\ell(\tau)$. The proof of the following lemma is exactly like [14, Lemma 3.6].
Lemma 5.14 [14, Lemma 3.6]. Let \( \tau \in \mathcal{P} \). Then

\[
\beta_i(\tau) \operatorname{diam}(\tau) \leq \sum_{k=0}^{\infty} d_{\tau_k}.\tag{31}
\]

Proof of Proposition 5.11. We compute in \( \ell^2 \)-fashion:

\[
\left( \sum_{\tau \in \mathcal{P}} \beta_i(\tau)^2 \operatorname{diam}(\tau) \right)^{1/2i} \leq \left( \sum_{\tau \in \mathcal{P}} \left( \sum_{k=0}^{\infty} d_{\tau_k} \right)^{2i} \operatorname{diam}(\tau)^{2i-1} \right)^{1/2i} \leq \sum_{k=0}^{\infty} \left( \sum_{\tau \in \mathcal{P}} \frac{d_{\tau_k}^{2i}}{\operatorname{diam}(\tau)^{2i-1}} \right)^{1/2i} \leq \sum_{k=0}^{\infty} 2^{-J(k-1)/2} \left( \sum_{\tau \in \mathcal{P}} \frac{d_{\tau_k}^{2i}}{\operatorname{diam}(\tau_k)^{2i-1}} \right)^{1/2i} \leq \sum_{k=0}^{\infty} 2^{-J(k-1)/2} C H^1(\Gamma)^{1/2i} \leq C H^1(\Gamma)^{1/2i}. \]

\[
\square
\]

5.4 | Flat balls

We are back to using just the \( d_\infty \) metric. One can compare the following lemma with [14, Lemma 4.3]. The crucial difference is that we are setting \( h = \beta_1(\tau') \operatorname{diam}(\tau') \) rather than \( \beta_s(\tau') \operatorname{diam}(\tau') \).

Lemma 5.15. There exists \( \eta_0 \in (0, 1) \) so that the following holds. Let \( B \in G_2 \) be a ball of radius \( r \) and \( Q = Q(B) \) for which \( 2B \subset Q \subset 2(1 + 2^{-98})B \). If \( \tau' \in \Lambda'(B) \) is such that \( \operatorname{Center}(B) \in \tau' \), \( \beta_s(\tau') < \eta_0 \), and

\[
h := \beta_1(\tau') \operatorname{diam}(\tau') < \frac{1}{10} r,
\]

then there is a subarc \( \overline{\tau} \subset \tau' \) with image in \( 2B \) such that \( \operatorname{diam}(\overline{\tau}) \geq 4r - 20h \).

Proof. We translate the setting so that \( \operatorname{Center}(B) = 0 \) and \( r = 1 \). Let \( k = \beta_3(\tau') \operatorname{diam}(\tau') \), \( L = L_{\tau'} \) and \( C(L, k) = \{ p \in G : d(p, L) \leq k \} \). Then \( \tau' \subset C(L, k) \). As \( 0 \in \tau' \subset C(L, k) \), one gets that \( d(0, L) \leq k \). Recall that horizontal lines going through 0 are lines of \( \mathbb{R}^{n_1} \times \{0\} \) that go through the origin. It thus follows from continuity of group multiplication that if \( \eta_0 \) (and so \( k \) also) is
sufficiently small, then $L$ is close in Hausdorff distance to a line of $\mathbb{R}^{n_1} \times \{0\}$ and so
\[ C(L, k) \cap 2B = C(L, k) \cap (B_{\mathbb{R}^{n_1}}(2) \times \mathbb{R}^{n_2+\cdots+n_s}). \] (32)

Here, we have used the fact that $B = \{x : N_\infty(x) \leq 1\}$ has a product structure. This is where we have used the fact that our metric is $d_\infty$. We will not use it anywhere else.

As $\tau' \subset C(L, k)$, we get from (32) that
\[ \pi_1(\tau' \cap 2B) = \pi_1(\tau' \cap (B_{\mathbb{R}^{n_1}}(2) \times \mathbb{R}^{n_2+\cdots+n_s})) = \pi_1(\tau') \cap B_{\mathbb{R}^{n_1}}(2). \]

Let $\xi' = \pi_1(\tau')$, which is now a curve in $\pi_1(G)$. As $\pi_1 : G \to \pi_1(G)$ is 1-Lipschitz, it suffices then to find a subcurve $\tilde{\xi} \subset \xi'$ contained in $B_{\pi_1(G)}(2)$ so that $\text{diam}(\tilde{\xi}) \geq 4 - 20h$. We now have the fact that $\beta_\infty(\tau') \leq \hat{\beta}(\tau') < \epsilon_0 \hat{\beta}_\Gamma(B) < 1 - \frac{1}{1000}$.

Recall that the metric $d_{HS}$ on $\pi_1(G) \cong \mathbb{R}^{n_1}$ is equivalent to the Euclidean metric $|\cdot|$. Thus, we have now reduced to the following problem in $\mathbb{R}^{n_1}$: We have a cylinder $\pi_1(C(L, h))$ of radius $h$ containing both 0 and a curve $\xi'$. Because the central axis of the cylinder $\pi_1(L)$ goes from $\xi'(a(\tau')) \in B_{\mathbb{R}^{n_1}}(2)^c$ to $\xi'(b(\tau')) \in B_{\mathbb{R}^{n_1}}(2)^c$ but the cylinder also contains 0, these endpoints must lie at opposite ends of the cylinder and so are almost antipodal with respect to the ball. We now want to find a subcurve $\tilde{\xi} \subset \xi'$ contained in $B_{\mathbb{R}^{n_1}}(2)$ with $\text{diam}(\tilde{\xi}) > 4 - 20h$. The existence of such a subcurve is a simple (but tedious) Euclidean exercise that involves finding the maximum subcurve of $\xi'$ in $B_{\mathbb{R}^{n_1}}(2)$. We leave the details for the reader.

We now choose once and for all $\epsilon_0 < \min\{\frac{\eta_0}{1000}, \frac{1}{6000^{\frac{1}{s}} - 1}\}$. Note that if $\tau' \in \Lambda'(B)$ is such that $B \in G_2$ and $\tau' \ni \text{Center}(B)$, then the hypothesis of Lemma 5.15 is automatically satisfied as
\[ \beta_\infty(\tau') \leq \hat{\beta}(\tau') < \epsilon_0 \hat{\beta}_\Gamma(B) < \frac{\eta_0}{1000} < \frac{1}{1000}. \] (33)

The following lemma will allow us to use the definition of flat balls to find a large subcurve $\hat{\xi}$ that is far away from $\tau'$.

**Lemma 5.16.** Let $B \in G_2$ be a ball of radius $r$. Let $\tau' \in \Lambda'(B)$ be such that $\text{Center}(B) \in \tau'$. Then there is a $\xi \in \Lambda(B)$ with a subarc $\tilde{\xi} \subset \xi$ with image inside $2B$ of diameter
\[ \text{diam}(\tilde{\xi}) > 100\epsilon_0^s \hat{\beta}_\Gamma(B)^s \text{diam}(B) \]
so that
\[ d(\tilde{\xi}, \tau') > 100\epsilon_0^s \hat{\beta}_\Gamma(B)^s \text{diam}(B). \]

**Proof.** Define $L = L_{\tau'}$. By definition, we have
\[ \sum_{i=1}^{s} \sup_{x \in B \cap \Gamma} \left( \frac{d(\pi_i(x), \pi_i(L))}{r} \right)^{2i} \geq \hat{\beta}_\Gamma(B)^{2s}. \]
Thus, there must exist an $i$ and $x \in B \cap \Gamma$ for which

$$\left( \frac{d(\pi_i(x), \pi_i(L))}{r} \right)^{2i} \geq \frac{\hat{\beta}_i(B)_{2s}}{2s}.$$ 

We can now bound

$$\varepsilon_0 \lesssim (20) \frac{\text{diam}(B)}{2000s^{-1/2s} \text{diam}(\tau')} \leq \frac{\text{diam}(B)/2 \text{diam}(\tau')}{1000s^{-1/2s}} \lesssim (20) \frac{\text{diam}(B)/2 \text{diam}(\tau')^{1/s}}{1000s^{-1/2s}} \leq \frac{r/\text{diam}(\tau')^{1/s}}{1000s^{-1/2s}}$$

and so

$$\left( \frac{d(\pi_i(x), \pi_i(L))}{\text{diam}(\tau')} \right)^{2i} \geq (1000\varepsilon_0)^{2s/s} \left( \frac{d(\pi_i(x), \pi_i(L))}{r} \right)^{2i} \geq (500\varepsilon_0\hat{\beta}_i(B))^{2s} > \hat{\beta}(\tau')^{2s}.$$ 

Thus, $x \notin \tau'$ and so $x \in \xi$ for some other $\xi \in \Lambda(B)$ for which

$$d(x, L) \geq d(\pi_i(x), \pi_i(L)) \geq 500(\varepsilon_0\hat{\beta}_i(B))^{s/i} \text{diam}(\tau').$$

On the other hand, we have

$$\sup_{i \in I_{\tau'}} d(\pi_i(y(t)), \pi_i(L)) \leq \hat{\beta}(\tau')^{s/i} \text{diam}(\tau') < (\varepsilon_0\hat{\beta}_i(B))^{s/i} \text{diam}(\tau').$$

Altogether we get

$$d(x, \tau') \geq d(\pi_i(x), \pi_i(\tau')) \geq 499(\varepsilon_0\hat{\beta}_i(B))^{s/i} \text{diam}(\tau') \geq 499(\varepsilon_0\hat{\beta}_i(B))^{s} \text{diam}(\tau').$$

The lemma now follows by taking $\hat{\xi}$ to be a maximal subarc of $\xi$ in $B(x, 100\varepsilon_0\hat{\beta}_i(B)^s \text{diam}(B))$ containing $x$ and remembering that $x \in B$ and $100\varepsilon_0\hat{\beta}_i(B)^s \leq 1/2$. □

**Proposition 5.17.** Let $B \in \mathcal{G}_2$ and $E = \Gamma \cap 2B$. If we cover $E$ with balls $\{B_i\}$ so that $\text{diam}(B_i) < 10\varepsilon_0\hat{\beta}_i(B)^s \text{diam}(B)$, then

$$\sum_i \text{diam}(B_i) \geq [2 + (\varepsilon_0\hat{\beta}_i(B)^s)] \text{diam}(B).$$

**Proof.** Let $\tau' \in \Lambda'(B)$ contain Center($B$), $\overline{\tau} \subset \tau'$ be the subcurve from Lemma 5.15, and $\hat{\xi}$ be the curve from Lemma 5.16 (applied to $\tau'$). Note that any $B_i$ may intersect at most one of the images of $\overline{\tau}$ and $\hat{\xi}$. Thus, as $\beta_1(\tau') \leq \hat{\beta}(\tau')^{s} \leq \varepsilon_0^s\hat{\beta}_i(B)^s$, we get

$$\sum_i \text{diam}(B_i) \geq \sum_{B_i \cap \overline{\tau} \neq \emptyset} \text{diam}(B_i) + \sum_{B_i \cap \hat{\xi} \neq \emptyset} \text{diam}(B_i) \geq \text{diam}(\overline{\tau}) + \text{diam}(\hat{\xi}).$$
\[ \geq 2 \text{diam}(B) - 20\varepsilon_0\hat{\beta}(B)^s \text{diam}(\tau') + 100\varepsilon_0\hat{\beta}(B)^s \text{diam}(B) \]

\[(20) \geq (2 + \varepsilon_0\hat{\beta}(B)^s) \text{diam}(B). \]

Now fix an integer \( M \geq \lfloor -\log_2(s) \rfloor =: M_0. \) As \( \hat{\beta}_E \leq s \) always, we can then define
\[ B^M := \{ B \in \mathcal{G}_2 : \hat{\beta}(B) \in [2^{-M-1}, 2^{-M}] \}. \]

Set \( J_M = \lfloor s(M - M_0) - s \log(10\varepsilon_0) + 10 \rfloor \) and we apply Lemma 5.9 to \( B^M \) with \( J = J_M \) to get \( \{ B^M_i \}_{i=1}^{D_M} \) where \( D_M = D \cdot J_M \). Note that this is a second application of Lemma 5.9 with a new value of \( J \) (as opposed to \( J = 100 \)). For each \( i \in \{1, \ldots, D_M\} \), apply the dyadic cube construction following Lemma 5.9 to \( B^M_i \) to get dyadic cubes \( \Delta(B^M_i, i) \). Fix one such \( \Delta = \Delta(B^M, i) \).

**Proposition 5.18.** Let \( B \in \mathcal{G}_2 \) and suppose \( Q = Q(B) \in \Delta \) is decomposed as
\[ Q = \left( \bigcup_i Q_i \right) \cup R_Q, \]
where \( Q_i = Q(B_i) \in \Delta \) are maximal so that \( Q_i \subsetneq Q \) and \( R_Q \) is chosen so that the union is disjoint. Then
\[ \sum_i \text{diam}(Q_i) + H^i(R_Q \cap \Gamma) \geq \text{diam}(Q) \left( 1 + \frac{1}{10\varepsilon_0\hat{\beta}(B)^s} \right). \]

**Proof.** The proof is essentially the same as [14, Proposition 4.7] except we will use our (34) instead of their Equation (28). \( \square \)

We can now prove the following proposition.

**Proposition 5.19.**
\[ \sum_{Q \in \Delta} \text{diam}(Q) \leq \frac{10}{\varepsilon_0} 2^s H^1(\Gamma). \]

**Proof.** The proof is essentially the same as [14, Proposition 4.8] except we use our Proposition 5.18 instead of their Proposition 4.7. Note that this changes their quantity \( q \) to \( q = (1 + \varepsilon_0^s 2^{-sM}/10)^{-1} \) and so we get
\[ \sum q^n = \frac{1}{1 - q} \leq \frac{10}{\varepsilon_0^s} 2^{-sM}. \]

\( \square \)

We can finally finish by proving the following proposition, which is stronger than (22), the goal of this subsection.
Proposition 5.20.

\[ \sum_{B \in \mathcal{G}_2} \hat{\beta}(B)^{s+1} \operatorname{diam}(B) \leq C \mathcal{H}^1(\Gamma). \]

Proof. The proof is essentially the same as the proof of [14, eq. (13)] (right after the proof of their Proposition 4.8). Instead of raising by 2, we raise by \( s + 1 \). See also the discussion after the proof of [5, Lemma 4.7] for a small correction of the argument in [14].

Note that the exponent of \( \hat{\beta} \) in Proposition 5.20 is much better than the \( 2s \) that is needed. This does not mean that Theorem 5.1 can be improved as it is the non-flat balls that present the bottleneck. Indeed, even in Hilbert spaces, where this flat ball argument was first presented, one gets the better bound \( \sum_{B \in \mathcal{G}_2} \beta_B \operatorname{diam}(B) \leq C \ell(\Gamma) \) (see [19, Corollary 3.26]).

6 | SUFFICIENCY

We will prove the following theorem which will take care of the sufficient direction.

Theorem 6.1. Let \( G \) be a step \( s \) Carnot group with Hausdorff dimension \( Q \) and \( E \subset G \) a subset. Define

\[ \hat{\beta}(E) = \int_0^{\infty} \int_G \hat{\beta}_E(x, r)^{2s} \, dx \, dr r^Q. \]

There exists some constant \( C > 0 \) depending on \( G \) and its Euclidean structure \( | \cdot | \) so that if \( \operatorname{diam}(E) + \hat{\beta}(E) \) is finite then \( E \) lies on a rectifiable curve of length no more than \( C(\operatorname{diam}(E) + \hat{\beta}(E)) \).

One sees that Theorems 5.1 and 6.1 together imply Theorem 1.5.

The following lemma is an improvement on Lemma 4.4. Note that it does not follow from Lemma 4.4 because the closest point in \( L \) to \( p \) is chosen after the level \( i \) is fixed. As with Lemma 4.4, it holds for both \( d_\infty \) and \( d_{HS} \).

Lemma 6.2. There exists a constant \( C > 0 \) depending only on \( d, \eta_0 < 1 \), and the Euclidean structure \( | \cdot | \) so that if \( p \in G \) and \( L \subset G \) is a horizontal line for which \( \max_i (d(\pi_i(p), \pi_i(L)) / \ell')^{2i} = \eta < \eta_0 \) for some \( \ell' > 0 \), then \( p \in L \cdot \delta_{\epsilon'}(B_{\mathbb{R}^n}(C\eta^{1/2})). \)

Proof. As usual, we dilate so that \( \ell' = 1 \). Choose \( C_0 > 0 \) so that \( N(\pi_k(g)) \geq C_0 \sum_{1 \leq i \leq k} |g_i|^{1/i} \) for all \( k \). By choosing \( \eta_0 \) sufficiently small, we may suppose \( \eta^{1/2s} < \min\{1, C_0/2\} \).

We will prove for any \( k \) that there exists \( \eta_0 < 1 \) and \( c_1, ..., c_k > 0 \) so that if

\[ \max_{1 \leq i \leq k} d(\pi_i(p), \pi_i(L))^{2i} \leq \eta < \eta_0, \]

then

\[ p \in L \cdot \left( \prod_{i=1}^k [-c_i \eta^{1/2}, c_i \eta^{1/2}]^{n_i} \times \prod_{i=k+1}^s \mathbb{R}^{n_i} \right). \]

The lemma follows from this statement for \( k = s \).
The case of \( k = 1 \) is obvious. Assume we have the statement up to \( k - 1 \) and now assume 
\[
\max_{1 \leq i \leq k} d(\pi_i(p), \pi_i(L))^{2i} \leq \eta.
\]
Then by the induction hypothesis, we have 
\[
p \in L \cdot \left( \prod_{i=1}^{k-1} \left[-c_i\eta^{1/2}, c_i\eta^{1/2}\right]^{n_i} \times \prod_{i=k+1}^{s} \mathbb{R}^{n_i} \right).
\]
We choose \( \eta_0 \) so that 
\[
c_i\eta_0^{1/2} < 1/2 \quad \text{for all} \quad i \in \{1, \ldots, k - 1\}.
\]
Seeking a contradiction, let us suppose that 
\[
p \in L \cdot \left( \prod_{i=1}^{k-1} \left[-c_i\eta^{1/2}, c_i\eta^{1/2}\right]^{n_i} \times \left([-C\eta^{1/2}, C\eta^{1/2}]^{n_i}\right)^c \times \prod_{i=k+1}^{s} \mathbb{R}^{n_i} \right),
\]
for some sufficiently large \( C \) to be chosen (that will not depend on \( p \)). We translate the setting 
so that \( L = \{(vt,0,\ldots,0)\}_{t \in \mathbb{R}} \) where \( v \in S^{n_1-1} \) and \( p = (x_1, \ldots, x_s) \) where \( |x_i| \leq c_i\eta^{1/2} < 1/2 \) for 
\( 1 \leq i \leq k - 1 \) and \( |x_k| \geq C\eta^{1/2} \). Define 
\[
f(t) = d(\pi_k(p), \pi_k(vt,0,\ldots,0)) \geq C_0 \left(|x_1 - vt| + |x_k + P_k(-vt,0,\ldots,0,x_1,\ldots,x_{k-1})|^{1/k}\right),
\]
where we used the definition of \( C_0 \). We will prove \( f(t) > \eta^{1/2k} \) for all \( t \), which will be our contra-
diction.

First suppose \( |t| \geq 1 \). Recall that \( |v| = 1 \) so that 
\[
C_0|x_1 - vt| \geq \left(1 - |x_1|\right)C_0 \geq C_0/2 > \eta^{1/2k}.
\]
Thus, assume \( |t| \leq 1 \). Then by Lemma 4.1, we get that \( |P_k(-vt,0,\ldots,0,x_1,\ldots,x_{k-1})| \leq C_1\eta^{1/2} \) for
some \( C_1 \) depending only on \( G \). Thus, if we choose \( C > C_1 + C_0^{-k} \), then 
\[
C_0|x_k + P_k(-vt,0,\ldots,0,x_1,\ldots,x_{k-1})|^{1/k} \geq C_0(\eta^{1/2} - |P_k(-vt,0,\ldots,0,x_1,\ldots,x_{k-1})|^{1/k}) > \eta^{1/2k}.
\]

Following [8], we say that an ordered triple of points \( p_1, p_2, p_3 \) in a metric space is \textit{orderable} if 
\[
d(p_1, p_3) \geq \max\{d(p_1, p_2), d(p_2, p_3)\}.
\]
Note that \( t_1, t_2, t_3 \in \mathbb{R} \) is orderable if and only if \( t_2 \) is between \( t_1 \) and \( t_3 \).

The next lemma says that the triangle inequality excess of orderable points in \( \mathbb{R}^n \) can be 
bounded from above by \( \frac{h^2}{\ell} \) where \( h \) is the height of the triangle and \( \ell \) is the diameter.

**Lemma 6.3.** Let \( \alpha > 0 \) and \( P_1, P_2, P_3 \in \mathbb{R}^n \) be orderable points such that \( \alpha \ell \leq |P_i - P_j| \leq \ell \) for
some \( \ell > 0 \). Let \( L \) be any affine line. There exists a constant \( C > 0 \) depending only on \( \alpha \) so that 
\[
|P_1 - P_2| + |P_2 - P_3| - |P_1 - P_3| \leq C \max_{1 \leq i \leq 3} \frac{|P_i - L|^2}{\ell}.
\]

**Proof.** The three points lie on a Euclidean plane. As the right-hand side of (36) can only increase 
for lines not in the plane, we may assume \( \mathbb{R}^n = \mathbb{R}^2 \). We rotate and dilate so that \( \ell = 1, P_1 = 0, \) and
\[ P_3 = (\gamma, 0) \] where \( \gamma \in [\alpha, 1] \). Let \( P_2 = (x, y) \). As \( P_2 \notin \{P_1, P_3\} \), we must have \( x \in (0, \gamma) \) as otherwise the orderability condition is violated.

Let \( L' \) be the affine line going through \( P_1, P_3 \). We first claim that

\[ h := |P_2 - L'| \leq 2 \max_{1 \leq i \leq 3} |P_i - L|. \]

Indeed, if \( L \) is within \( h/2 \) of \( P_1, P_3 \) then \( |P_2 - L| \geq h/2 \), which gives the claim. Otherwise, \( L \) is at least \( h/2 \) away from one of \( P_1, P_3 \), which again gives the claim. Thus, it now suffices to prove that

\[ |P_1 - P_2| + |P_2 - P_3| - |P_1 - P_3| \leq C |P_2 - L'|^2. \tag{37} \]

As the left-hand side of (37) is bounded by 2, if \( |P_2 - L'| > \epsilon_0 \) for some \( \epsilon_0 > 0 \) sufficiently small to be determined, then the lemma immediately follows. Thus, we may assume \(|y| = h \leq \epsilon_0\).

By choosing \( \epsilon_0 < \alpha/2 \) we get that \( x \in (\alpha/2, \gamma - \alpha/2) \). Note then that \( P_2 = (x, y) \). We have

\[
|P_1 - P_2| + |P_2 - P_3| - |P_1 - P_3| = \sqrt{x^2 + h^2} + \sqrt{(\gamma - x)^2 + h^2} - x - (\gamma - x) \\
= \left( \frac{1}{\sqrt{x^2 + h^2 + x}} + \frac{1}{\sqrt{(\gamma - x)^2 + h^2 + y - x}} \right) h^2 \\
\leq \left( \frac{1}{2x} + \frac{1}{2(y - x)} \right) h^2 \\
\leq \frac{2}{\alpha} h^2,
\]

where we used the fact that \( x \in (\alpha/2, \gamma - \alpha/2) \) in the last inequality. \( \square \)

The next lemma says that in a metric space, the ordering of points is the same as the ordering of sufficiently close points on a subset isometric to \( \mathbb{R} \).

**Lemma 6.4.** Let \( L \subseteq X \) be a subset isometric to \( \mathbb{R} \) and \( p_1, p_2, p_3 \in X \) be points so that \( d(p_i, p_j) \geq \alpha \ell \) for some \( \ell > 0 \) and \( p_i \in B(t_i, \alpha \ell /8) \) for points \( t_i \in L \). Then \( p_1, p_2, p_3 \) are orderable if and only if \( t_1, t_2, t_3 \) are orderable.

**Proof.** By passing to the rescaled metric \( \frac{1}{\ell} d \) (which preserves orderability in the original metric), we may assume \( \ell = 1 \). By a triangle inequality argument, we have that

\[ d(t_i, t_j) \geq d(p_i, p_j) - \alpha \geq \frac{3}{4} \alpha. \]

First assume \( t_1, t_2, t_3 \) are orderable. Then

\[ d(t_1, t_3) \geq \max\{d(t_1, t_2), d(t_2, t_3)\} + \frac{3}{4} \alpha. \]
Here, we used the fact that $t_1, t_2, t_3$ lie in a subset isometric to $\mathbb{R}$. We have again by triangle inequality that

$$d(p_1, p_3) \geq d(t_1, t_3) - \alpha/4,$$
$$d(t_1, t_2) \geq d(p_1, p_2) - \alpha/4,$$
$$d(t_2, t_3) \geq d(p_2, p_3) - \alpha/4.$$

Thus,

$$d(p_1, p_3) \geq \max\{d(t_1, t_2), d(t_2, t_3)\} + \frac{\alpha}{2} \geq \max\{d(p_1, p_2), d(p_2, p_3)\}.$$

Now assume $p_1, p_2, p_3$ are orderable. As

$$d(t_1, t_3) \geq d(p_1, p_3) - \alpha/4,$$
$$d(p_1, p_2) \geq d(t_1, t_2) - \alpha/4,$$
$$d(p_2, p_3) \geq d(t_2, t_3) - \alpha/4,$$

we get that $d(t_1, t_3) \geq \max\{d(t_1, t_2), d(t_2, t_3)\} - \alpha/2$. As $d(t_i, t_j) \geq 3\alpha/4$, we must have that $t_2$ is between $t_1$ and $t_3$. \hfill \Box

**Lemma 6.5.** Let $\alpha > 0$ and $p_1, p_2, p_3 \in G$ be orderable points so that $\alpha \ell' \leq d_{HS}(p_i, p_j) \leq \ell'$ for some $\ell' > 0$ whenever $i \neq j$. Let $L$ be any horizontal line. If

$$\max_{i \in \{1, 2, 3\}} \frac{d_{HS}(p_i, L)}{\ell'} < \frac{\alpha}{8},$$

then $\frac{\alpha}{2} \ell' \leq |\pi_1(p_i) - \pi_1(p_j)| \leq \ell'$ whenever $i \neq j$ and $\pi_1(p_1), \pi_1(p_2), \pi_1(p_3)$ are also orderable in $\mathbb{R}^{n_1}$.

**Proof.** As usual, we dilate so that $\ell' = 1$. We translate so that $L = \{(tv, 0, \ldots, 0) : t \in \mathbb{R}\}$ for some $v \in S^{n_1-1}$. Then by assumption, there exists $t_1, t_2, t_3 \in \mathbb{R}$ so that $p_i \in (tv, 0, \ldots, 0) \cdot B_G(\alpha/8)$. We have by a triangle inequality argument that

$$|t_i - t_j| \geq d_{HS}(p_i, p_j) - \frac{\alpha}{4} \geq \frac{3}{4} \alpha$$

whenever $i \neq j$. Here, we used the fact that $N_{HS}(tv, 0, \ldots, 0) = |t|$. We first prove that $\frac{\alpha}{2} \leq |\pi_1(p_i) - \pi_1(p_j)| \leq 1$. The upper bound follows from the fact that $\pi_1$ is 1-Lipschitz. Thus, it suffices to prove the lower bound. We have by 1-Lipschitzness of $\pi_1$ that

$$\max_{i \in \{1, 2, 3\}} d_{HS}(\pi_1(p_i), \pi_1(t_i v, 0, \ldots, 0)) < \frac{\alpha}{8},$$
and so a triangle inequality argument gives
\[ |\pi_1(p_i) - \pi_1(p_j)| \geq |t_i v - t_j v| - \frac{\alpha}{4} \geq \frac{\alpha}{2}. \]

We now prove orderability. As \( L \) is isometric to \( \mathbb{R} \), we have by Lemma 6.4 that \( t_1, t_2, t_3 \) are orderable. As \( \pi_1 \) is 1-Lipschitz and an isometry when restricted on \( L \), we get by another application of Lemma 6.4 that \( \pi_1(p_1), \pi_1(p_2), \pi_1(p_3) \) is also orderable. 

The following lemma will allow us to Taylor expand the Hebisch–Sikora norm.

**Lemma 6.6.** For any Carnot group \( G \) and \( \alpha \in (0, 1) \), there exists a constant \( C > 0 \) so that if \( \alpha \leq N_{HS}(x_1, ..., x_{s-1}, 0) \leq 1 \) and \( |y| \leq 1/C \), then
\[ 0 \leq N_{HS}(x_1, ..., x_{s-1}, y) - N_{HS}(x_1, ..., x_{s-1}, 0) \leq C|y|^2. \]

**Proof.** As balls of the Hebisch–Sikora norm centered at the origin are axis-aligned ellipsoids, the left-hand inequality is obvious.

Define \( K = \{ (x_1, ..., x_{s-1}, 0) : \alpha \leq N_{HS}(x_1, ..., x_{s-1}, 0) \leq 1 \} \), which is a compact set. Recall that \( N_{HS} \) is smooth on \( G \setminus \{0\} \). For \( x = (x_1, ..., x_{s-1}) \), we define the function
\[ f_x(y) = N_{HS}(x_1, ..., x_{s-1}, y) - N_{HS}(x_1, ..., x_{s-1}, 0). \]

We have already proven that \( f_x(y) \geq 0 \) always. Furthermore, \( f_x(0) = 0 \) and so \( Df_x(0) = 0 \). Thus, by smoothness of \( N_{HS} \), we have that \( f_x(y) \leq \|D^2f_x(0)\|\|y\|^2 \) when \( |y| \leq \varepsilon_x \) where \( \|D^2f_x(0)\| \) is the operator norm of the Hessian of \( f_x \) at 0. Here, \( \varepsilon_x \) is positive for all \( x \in K \) and can be chosen in a continuous way. By compactness of \( K \), we get the lemma by choosing
\[ C = \max \left\{ \max_{x \in K} \varepsilon_x^{-1}, \max_{x \in K} \|D^2f_x(0)\| \right\}. \]

The following proposition is the main result of this section.

**Proposition 6.7.** For any \( \alpha \in (0, 1) \) there exists \( C > 0 \) depending on \( d_{HS} \) and the Euclidean structure \( |\cdot| \) so that if \( p_1, p_2, p_3 \in E \) are orderable points such that \( \alpha \ell \leq d_{HS}(p_i, p_j) \leq \ell \) for \( i \neq j \) for some \( \ell > 0 \), then
\[ d_{HS}(p_1, p_2) + d_{HS}(p_2, p_3) - d_{HS}(p_1, p_3) \leq C \max_{1 \leq i \leq 3} \max_{1 \leq k \leq s} d_{HS}(\pi_k(p_i), \pi_k(L))^{2k} \frac{d_{HS}(\pi_k(p_i), \pi_k(L))^{2k}}{\ell^{2k-1}} \]
where \( L \) is any horizontal line.

**Proof.** We dilate and translate the setting so that \( \ell = 1 \) and \( L = \{ (vt, 0, ..., 0) : t \in \mathbb{R} \} \) where \( v \in S^{n_1-1} \). Let \( \mu > 0 \) be a constant to be determined. Note that the left-hand side of the desired inequality is bounded from above by \( 2 \). If the non-constant term of the right-hand side is bounded from below by \( \mu \), then the inequality holds with \( C = 2/\mu \). Thus, we may suppose that
\[ \eta := \max_{1 \leq i \leq 3} \max_{1 \leq k \leq s} d_{HS}(\pi_k(p_i), \pi_k(L))^{2k} \leq \mu \]

(38)
for some \( \mu \) to be determined. Our goal then is to bound the left-hand side of the desired inequality by some multiple of \( \eta \).

Let \( P_i = \pi_1(p_i) \). We will first let \( \mu \leq (\alpha/8)^2 \) so that we can apply Lemma 6.5 to get that \( \alpha/2 \leq |P_i - P_j| \leq 1 \) whenever \( i \neq j \) and \( P_1, P_2, P_3 \) are orderable in \( \mathbb{R}^n \).

We then let \( \mu < \eta_0 \) of Lemma 6.2 to get that there exists some \( C_0 > 0 \) so that \( p_i = (t_i v, 0, ..., 0) \cdot u_i \) where \( u_i \in B_{\mathbb{R}^n}(C_0 \eta_0^{1/2}) \) and \( t_1, t_2, t_3 \in \mathbb{R} \). Thus, \( p_i^{-1} p_j = u_i^{-1} \cdot ((t_j - t_i) v, 0, ..., 0) \cdot u_j \). Note that \( P_i = t_i v + \pi_1(u_i) \in \mathbb{R}^n \). By repeated applications of Lemma 4.1, we get that

\[
 p_i^{-1} p_j = (P_j - P_i, q_{i,j}^{(2)}, ..., q_{i,j}^{(s)}),
\]

where \( \max_{2 \leq k \leq s} |q_{i,j}^{(k)}| \leq C_1 \eta_0^{1/2} \) for some \( C_1 > 0 \).

Let \( (x_1, ..., x_s) = p_i^{-1} p_j \). As \( |P_1 - P_2| \geq \alpha/2 \), we have that \( \alpha/2 \leq N_{HS}(x_1, ..., x_{s-1}, 0) \leq 1 \). Let \( C \) be the constant from Lemma 6.6. Recall that

\[
 |x_s| = |q_{1,2}^{(s)}| \leq C_1 \eta_0^{1/2} \leq C_1 \mu^{1/2}.
\]

Thus, if we take \( \mu \) sufficiently small, we get that \( |x_s| < 1/C \). We can then apply Lemma 6.6 to get

\[
 N_{HS}(x_1, ..., x_s) - N_{HS}(x_1, ..., x_{s-1}, 0) \leq C |x_s|^2 \leq C C_1^2 \eta.
\]

We repeat this process of bounding

\[
 d_{HS}(\pi_i(p_1), \pi_i(p_2)) - d_{HS}(\pi_{i-1}(p_1), \pi_{i-1}(p_2)) = N_{HS}(x_1, ..., x_i) - N_{HS}(x_1, ..., x_{i-1}, 0)
\]

in \( \pi_i(G) \) by some multiple of \( \eta \) and then telescope the resulting bounds to get

\[
 d_{HS}(p_1, p_2) = d_{HS}(\pi_s(p_1), \pi_s(p_2)) \leq d_{HS}(\pi_1(p_1), \pi_1(p_2)) + C_2 \eta = |P_1 - P_2| + C_2 \eta
\]

for some \( C_2 > 0 \). Likewise,

\[
 d_{HS}(p_2, p_3) \leq |P_2 - P_3| + C_2 \eta.
\]

Also, we get from repeated applications of the other inequality of Lemma 6.6 that

\[
 d_{HS}(p_1, p_3) \geq |P_1 - P_3|.
\]

Altogether, we get that

\[
 d_{HS}(p_1, p_2) + d_{HS}(p_2, p_3) - d_{HS}(p_1, p_3) \leq |P_1 - P_2| + |P_2 - P_3| - |P_1 - P_3| + 2C_2 \eta.
\]

Now applying Lemma 6.3 gives us that \( |P_1 - P_2| + |P_2 - P_3| - |P_1 - P_3| \leq C_3 \eta \) for some \( C_3 \) depending only on \( \alpha \). Recalling the definition of \( \eta \) in (38), we get the proposition. \( \square \)

We immediately get the following corollary.
Corollary 6.8. For any $\alpha \in (0, 1)$ there exists $C > 0$ depending on $\alpha$, $d_{HS}$, and the Euclidean structure so that the following holds. If $p_1, p_2, p_3 \in E$ are orderable points such that $\alpha \ell' \leq d_{HS}(p_i, p_j) \leq \ell'$ for some $\ell' > 0$ when $i \neq j$, then

$$d_{HS}(p_1, p_2) + d_{HS}(p_2, p_3) - d_{HS}(p_1, p_3) \leq C\hat{\beta}_E(B(p_i, 2\alpha \ell'))^{2s} \ell'$$

for any $i \in \{1, 2, 3\}$.

Proof of Theorem 6.1. We may assume $E$ is bounded as otherwise there is nothing to prove. To prove the theorem, we will use the farthest insertion algorithm of [8, section 3], which we assume the reader is familiar with. As a refresher, their construction starts with $\{\Delta_k\}_{k=0}^{\infty}$, a nested set of finite maximal $2^{-k}$-nets of $E$. They then build a sequence of connected piecewise geodesic curves $\{\Gamma_k\}_{k=0}^{\infty}$ where $\Delta_k$ is contained at the endpoints of the geodesic segments (see the bottom of [8, p. 467]) and pass to a limit to get $\Gamma$.

The construction of [8] is in the Heisenberg group and uses the usual $\beta$ (which they call $\beta_{\mathbb{H}}$). However, they make no reference to the Heisenberg group except through their Proposition 3.2 and Corollary 2.16 (there is a typo as a square is missing in their $\beta_{\mathbb{H}}(P_i, C_0 t)$). By their Remark 3.1, their Proposition 3.2 holds for all Carnot groups. As for their Corollary 2.16, we will replace that with our Corollary 6.8. Additionally, in Case A, we use the following bound instead of their (3.3):

$$\ell(\Gamma_j) - \ell(\Gamma_{j-1}) \leq C 2^{-j} \leq \frac{C}{\varepsilon_0^2} 2^{-j} \beta_E(P, C_1 2^{-j})^{2s} \leq \frac{C}{\varepsilon_0^2} 2^{-j} \hat{\beta}_E(P, C_1 2^{-j})^{2s},$$

This, along Corollary 6.8, replaces the presence of $\beta_{\mathbb{H}}^2$ in their upper bound with $\hat{\beta}_E^{2s}$.

One important difference is that [8] uses the Carnot–Carathéodory metric, but we use $d_{HS}$. Thus, all nets, balls, and so on, will be done with respect to $d_{HS}$. As $d_{HS}$ is not geodesic, we must modify the method of [8]. We will actually apply the algorithm of [8] to inductively build abstract connected graphs $\Gamma_k$ whose vertices contain $\Delta_k$. The ‘length’ of an abstract graph $\ell(\Gamma_k)$ is the sum of the $d_{HS}$ length of all abstract edges in the graph, and all references to geodesic segments in the algorithm will be replaced by abstract edges.

The results of [8] show that their $\ell(\Gamma_k)$ is bounded their $\beta_{\mathbb{H}}(E) + \text{diam}(E)$ where $\beta_{\mathbb{H}}(E)$ is the Carleson integral of $\beta_{\mathbb{H}}^2$. Indeed, let $\mathcal{A}_k$ denote all geodesic segments of $\mathcal{A}$ (as defined right before Case B2(i)) that appear in $\Gamma_k$. Then, using the notation of [8], we can conclude using their equations (3.3), (3.4), (3.10), (3.11) (there is a typo in (3.11) as it should be $\tau_2([Q_1, R_0])$ on the left-hand side) that

$$\ell(\Gamma_k) = \sum_{j=0}^{k-1} (\ell(\Gamma_{k+1}) - \ell(\Gamma_k)) + \ell(\Gamma_0) \leq C \sum_{j \in \mathbb{N}} \sum_{P \in \Delta_j} \beta_{\mathbb{H}}(P, C_3 2^{-j})^2 2^{-j} + \sum_{I \in \mathcal{A}_k} \tau_1(I) + \sum_{I \in \mathcal{A}_k} \tau_2(I) + C \text{diam}(E) \leq C \sum_{j \in \mathbb{N}} \sum_{P \in \Delta_j} \beta_{\mathbb{H}}(P, C_3 2^{-j})^2 2^{-j} + \frac{1}{10^6} \ell(\Gamma_k) + \frac{1}{10^3} \ell(\Gamma_k) + C \text{diam}(E),$$

for some $\alpha \in (0, 1)$.
which implies that $\ell'(\Gamma_k) \leq C(\beta_\gamma(E) + \text{diam}(E))$ for all $k$ as the sum of the $\beta_\gamma^2$ is just a discretized Carleson integral. This is similar to their estimate at the top of p. 468. With the modifications we outlined above, we would get the following uniform bound for the abstract graphs

$$\ell'(\Gamma_k) \leq C(\text{diam}(E) + \hat{\beta}(E)).$$

Once the abstract graphs are built, we replace each abstract edge in $\Gamma_k$ with quasi-convex curves using (6) to get a sequence of actual continua $\{\Gamma_k\}_{k=0}^\infty$ such that $\Delta_k \subset \Gamma_k$. Note that the lengths are still uniformly bounded

$$\mathcal{H}^1(\Gamma_k) \leq C\ell'(\Gamma_k) \leq C(\text{diam}(E) + \hat{\beta}(E)).$$

Finally, as each $\Delta_k$ is finite and so compact and lies in $\Gamma_j$ for $j \geq k$, we get from their Theorem 5.1 and a diagonalization argument that the limiting curve $\Gamma$ of (a subsequence of) the $\Gamma_k$ contains each $\Delta_k$ and $\mathcal{H}^1(\Gamma) \leq C(\text{diam}(E) + \hat{\beta}(E))$. As $\Gamma$ is compact and $\bigcup_{k=1}^\infty \Delta_k \supseteq E$, we then get that $E \subseteq \Gamma$. This proves the theorem. □

7 PROOF OF PROPOSITION 1.6

We first prove the equivalence of (1) and (2). We prove $(2) \leq (1)$ (up to constants) first. Let $L$ be a horizontal line so that

$$\sum_{i=1}^s \sup_{z \in E \cap B(x, r)} \left( \frac{d_{HS}(\pi_i(z), \pi_i(L))}{r} \right)^{2i} \leq 2\hat{\beta}_E(x, r)^{2s}.$$

Then we also have

$$\sup_{z \in E \cap B(x, r)} \max_{i \in \{1, \ldots, s\}} \left( \frac{d_{HS}(\pi_i(z), \pi_i(L))}{r} \right)^{2i} \leq 2\hat{\beta}_E(x, r)^{2s}.$$

and so Lemma 6.2 gives that

$$E \cap B_G(x, r) \subseteq L \cdot \delta_r(B_{\mathbb{R}^n}(C\sqrt{2\hat{\beta}_E(x, r)^s}))$$

which gives the needed bound.

We now prove the converse bound. Suppose that

$$E \cap B_G(x, r) \subseteq L \cdot \delta_r(B_{\mathbb{R}^n}(\varepsilon^s)).$$

Note that we can take $\varepsilon \leq M$ where $M$ is a constant depending only on $d$ and $|\cdot|$ so that $B_G(0, 1) \subseteq B_{\mathbb{R}^n}(M^s)$.

Let $z \in E \cap B(x, r)$. Then there is a $p \in L$ so that $z \in p \cdot \delta_r(B_{\mathbb{R}^n}(\varepsilon^s))$. Then Lemma 4.4 gives

$$\left( \frac{d_{HS}(\pi_i(z), \pi_i(p))}{r} \right)^i \leq C\varepsilon^s, \quad \forall i \in \{1, \ldots, s\}.$$
As $p \in L$ and $z \in E \cap B(x, r)$ was arbitrary, we get for each $i$ that

$$\sup_{z \in E \cap B(x, r)} \left( \frac{d_{HS}(\pi_i(z), \pi_i(L))}{r} \right)^{2i} \leq C^2 \varepsilon^{2s}.$$ 

which allows us to conclude that

$$\hat{\beta}_E(x, r)^{2s} = \sum_{i=1}^{s} \sup_{z \in B(x, r) \cap E} \left( \frac{d_{HS}(\pi_i(z), \pi_i(L))}{r} \right)^{2i} \leq C^2 s \varepsilon^{2s}$$

as needed. This proves the equivalence between (1) and (2).

We now establish the equivalence between (2) and (3). Note that the $\varepsilon$ of (2) can be bounded by the same $M$ as above. There is also a similar upper bound for the $\varepsilon$ of (3) which depends only on $d$ and $\rho$. Now, by translating and dilating, we can rewrite (2) as

$$\inf \inf \{ \varepsilon > 0 : \delta_{1/r}(x^{-1}E) \cap B(0, 1) \subset L \cdot B_{\mathbb{R}^n}(\varepsilon^s) \}.$$ 

The left-invariance of $\rho$ along with (7) gives that there exists a $C_0 > 1$ so that

$$L \cdot B_{\mathbb{R}^n}(\varepsilon^s/C_0) \subseteq \{ z \in G : \rho(z, L) < \varepsilon^s \} \subseteq L \cdot B_{\mathbb{R}^n}(C_0 \varepsilon^s).$$

The equivalence of (2) and (3) now easily follows.

**ACKNOWLEDGEMENTS**

The research presented here was supported by NSF, Grant number: DMS-1812879. We would also like to thank the anonymous referees for their detailed review that caught many errors and helped improve the presentation of the paper, including substantial simplifications of many proofs.

**JOURNAL INFORMATION**

The *Journal of the London Mathematical Society* is wholly owned and managed by the London Mathematical Society, a not-for-profit Charity registered with the UK Charity Commission. All surplus income from its publishing programme is used to support mathematicians and mathematics research in the form of research grants, conference grants, prizes, initiatives for early career researchers and the promotion of mathematics.

**REFERENCES**

1. M. Badger and S. McCurdy, *Subsets of rectifiable curves in Banach spaces: sharp exponents in Schul-type theorems*, Preprint, 2020.
2. M. Badger, L. Naples, and V. Vellis, *Hölder curves and parameterizations in the Analyst’s Traveling Salesman theorem*, Adv. Math. 349 (2019), 564–647.
3. C. J. Bishop and Y. Peres, *Fractals in probability and analysis*, Cambridge Studies in Advanced Mathematics, Cambridge University Press, Cambridge, 2016.
4. E. Breuillard, *Geometry of locally compact groups of polynomial growth and shape of large balls*, Groups Geom. Dyn. 8 (2014), no. 3, 669–732.
5. V. Chousionis, S. Li, and S. Zimmerman, *The traveling salesman theorem on Carnot groups*, Calc. Var. PDE 58 (2019), no. 14.
6. G. C. David and R. Schul, *A sharp necessary condition for rectifiable curves in metric spaces*, Rev. Mat. Iberoamericana **37** (2021), no. 3, 1007–1044.

7. G. C. David and R. Schul, *The Analyst’s traveling salesman theorem in graph inverse limits*, Ann. Acad. Sci. Fenn. Math. **42** (2017), 649–692.

8. F. Ferrari, B. Franchi, and H. Pajot, *The geometric traveling salesman problem in the Heisenberg group*, Rev. Mat. Iberoam. **23** (2007), no. 2, 437–480.

9. I. Hahlomaa, *Menger curvature and rectifiability in metric spaces*, Adv. Math. **219** (2008), no. 6, 1894–1915.

10. W. Hebisch and A. Sikora, *A smooth subadditive homogeneous norm on a homogeneous group*, Studia Math. **96** (1990), no. 3, 231–236.

11. P. W. Jones, *Rectifiable sets and the traveling salesman problem*, Invent. Math. **102** (1990), no. 1, 1–15.

12. N. Juillet, *A counterexample for the geometric traveling salesman problem in the Heisenberg group*, Rev. Mat. Iberoam. **26** (2010), no. 3, 1035–1056.

13. S. Li, *Coarse differentiation and quantitative nonembeddability for Carnot groups*, J. Funct. Anal. **266** (2014) 4616–4704.

14. S. Li and R. Schul, *The traveling salesman problem in the Heisenberg group: upper bounding curvature*, Trans. Amer. Math. Soc. **386** (2016), no. 7, 4585–4620.

15. S. Li and R. Schul, *An upper bound for the length of a traveling salesman path in the Heisenberg group*, Rev. Mat. Iberoam. **32** (2016), no. 2, 391–417.

16. H. Pajot, *Analytic capacity, rectifiability, Menger curvature and cauchy integral*, Lecture Notes in Mathematics, vol. 1799, Springer, Berlin, 2002.

17. K. Okikiolu, *Characterizations of subsets of rectifiable curves in \( \mathbb{R}^n \)*, J. Lond. Math. Soc. (2) **46** (1992), 336–348.

18. R. Schul, *Ahlfors-regular curves in metric spaces*, Ann. Acad. Sci. Fenn. Mat. **32** (2007), 437–460.

19. R. Schul, *Subsets of rectifiable curves in Hilbert space—the Analyst’s TSP*, J. Anal. Math. **103** (2007), 331–375.