INTRINSIC LIPSCHITZ GRAPHS AND VERTICAL β-NUMBERS IN THE HEISENBERG GROUP

VASILEIOS CHOUSIONIS, KATRIN FÄSSLER AND TUOMAS ORPONEN

ABSTRACT. The purpose of this paper is to introduce and study some basic concepts of quantitative rectifiability in the first Heisenberg group $\mathbb{H}$. In particular, we aim to demonstrate that new phenomena arise compared to the Euclidean theory, founded by G. David and S. Semmes in the 90's. The theory in $\mathbb{H}$ has an apparent connection to certain nonlinear PDEs, which do not play a role with similar questions in $\mathbb{R}^3$.

Our main object of study are the *intrinsic Lipschitz graphs* in $\mathbb{H}$, introduced by B. Franchi, R. Serapioni and F. Serra Cassano in 2006. We claim that these 3-dimensional sets in $\mathbb{H}$, if any, deserve to be called quantitatively 3-rectifiable. Our main result is that the intrinsic Lipschitz graphs satisfy a *weak geometric lemma* with respect to *vertical β-numbers*. Conversely, extending a result of David and Semmes from $\mathbb{R}^n$, we prove that a 3-Ahlfors-David regular subset in $\mathbb{H}$, which satisfies the weak geometric lemma and has *big vertical projections*, necessarily has *big pieces of intrinsic Lipschitz graphs*.

CONTENTS

2010 *Mathematics Subject Classification.* 28A75 (Primary), 28C10, 35R03 (Secondary).
V.C. is supported by the Simons Foundation via the Collaboration grant *Analysis and dynamics in Carnot groups*, grant no. 521845. K.F. is supported by the Academy of Finland through the grant *Sub-Riemannian manifolds from a quasiconformal viewpoint*, grant number 285159. T.O. is supported by the Academy of Finland through the grant *Restricted families of projections, and applications to Kakeya type problems*, grant number 274512. T.O. is also a member of the Finnish CoE in Analysis and Dynamics Research.

1. INTRODUCTION

Rectifiability is a fundamental concept in geometric measure theory. Rectifiable sets extend the class of surfaces considered in classical differential geometry; while admitting a few edges and sharp corners, they are still smooth enough to support a rich theory of local analysis. However, for certain questions of more global nature – the boundedness of singular integrals being the main example – the notion of rectifiability is too qualitative.

In a series of influential papers around the year 1990, [15, 19, 18, 16], G. David and S. Semmes developed an extensive theory of *quantitative rectifiability* in Euclidean spaces. One of their main objectives was to find geometric criteria to characterize the m-dimensional subsets of $\mathbb{R}^n$, $0 < m < n$, on which "nice" singular integral operators (SIO) are L^2-bounded. Here, "nice" refers to SIOs with smooth, odd Calderón-Zygmund kernels, the archetype of which is the Riesz kernel $x/|x|^{m+1}$, $x \in \mathbb{R}^n$. Notice that for $n = 2$ and $m = 1$, the Riesz kernel essentially coincides with the Cauchy kernel $1/z, z \in \mathbb{C}$, in the complex plane.

A motivation for the efforts of David and Semmes was the significance of SIOs for the problem of finding a geometric characterization of *removable sets* for bounded analytic functions and Lipschitz harmonic functions. Due to seminal papers by David [20], David and P. Mattila [21], and F. Nazarov, X. Tolsa and A. Volberg [40], [41], it is now known that these removable sets coincide with the purely $(n-1)$-unrectifiable sets in $\mathbb{R}^n$, i.e. the sets which intersect every $\mathcal{C}^1$ hypersurface in a set of vanishing $(n-1)$-dimensional Hausdorff measure. The geometric characterization of removability, and its connections to geometric measure theory and harmonic analysis, has a very interesting history; we refer to the excellent survey by Volberg and V. Eiderman [22], and to the recent book of Tolsa [43].

The problem of characterizing removable sets for Lipschitz harmonic functions has a natural analogue outside the Euclidean setting in certain non-commutative Lie groups, of which the Heisenberg group is the simplest example. In this group, the role of the Euclidean Laplace operator is played by the *sub-Laplacian* and harmonic functions are, by definition, the solutions to the sub-Laplacian equation. The question of characterizing removability for Lipschitz harmonic functions was considered in [12] for Heisenberg groups $\mathbb{H}^n$ endowed with a sub-Riemannian metric. It was shown that sets with Hausdorff dimension lower than $2n+1$ are removable, while those with dimension higher than $2n+1$ are not. Moreover, there exist both removable and non-removable sets with Hausdorff dimension equal to $2n+1$. Hence, as in the Euclidean case, the dimension threshold for removable sets is $\dim_{\mathrm{H}} \mathbb{H}^n - 1 = 2n+1$, where $\dim_{\mathrm{H}} \mathbb{H}^n$ denotes the Hausdorff dimension of $\mathbb{H}^n$. The results from [12] were extended in [13] from Heisenberg groups to a larger class of Lie groups, the *Carnot groups*. There exists a well developed theory of sub-Laplacians in this setting, see for instance the book [10] by A. Bonfiglioli, E. Lanconelli and F. Uguzzoni.

In order to characterize removable sets for Lipschitz harmonic functions in $\mathbb{R}^n$, one has to characterize the sets on which the SIO associated with the Riesz kernel $x/|x|^n$ is bounded in L^2. In $\mathbb{H}^1$, one would need to face a SIO with kernel

$$K(p) := \left(\frac{x|z|^2 + yt}{(|z|^4 + t^2)^{3/2}}, \frac{y|z|^2 - xt}{(|z|^4 + t^2)^{3/2}} \right)$$

for $p = (z,t), z = x + iy \in \mathbb{C}, t \in \mathbb{R}$. At this point, the knowledge about the action of this SIO on 3-dimensional subsets of $\mathbb{H}^1$ (i.e. subsets of co-dimension 1) is extremely limited. In the present paper, we will not address this question further, but it motivates the study of quantitative rectifiability in $\mathbb{H}^1$.

The main purpose of the present paper is to initiate this study, and to introduce some new, relevant concepts in $\mathbb{H} = \mathbb{H}^1$. Our aims are twofold. First, we demonstrate that some parts, at least, of the Euclidean theory of quantitative rectifiability carry over to $\mathbb{H}$. To us, this gives hope that – some day in the distant future – questions on the boundedness of SIOs on subsets of $\mathbb{H}$ may be understood as well as they currently are in $\mathbb{R}^n$. Our second aim is somewhat more philosophical: we want to demonstrate that building a theory of quantitative rectifiability in $\mathbb{H}$ is worth the effort. In particular, the proofs are not, merely, technically challenging replicas of their Euclidean counterparts. New phenomena appear. In particular, proving our main result, Theorem 1.1, involved studying non-smooth solutions of the (planar) non-linear PDE

$$\partial_y \phi + \phi \partial_t \phi = c, \qquad c \in \mathbb{R}, \tag{1.1}$$

known as the (or rather "a") *Burgers equation*. In proving the Euclidean counterpart of Theorem 1.1, such considerations are not required. At least to us, any connection between the innocent-looking statement of Theorem 1.1, and the PDE (1.1), was quite a surprise at first sight.

In the terminology of David and Semmes, Theorem 1.1 is the *weak geometric lemma* for certain subsets of $\mathbb{H}$, called *intrinsic Lipschitz graphs*. For now, we just briefly explain the meaning of these concepts; precise definitions are postponed to Sections 2 and 3. We consider two kinds of subgroups of $\mathbb{H}$: *horizontal* and *vertical*. Writing $\mathbb{H} = \mathbb{R}^2 \times \mathbb{R}$, the horizontal subgroups are lines through the origin inside $\mathbb{R}^2 \times \{0\}$, while the vertical subgroups are planes spanned by a horizontal subgroup and the "t-axis" $\{0\} \times \mathbb{R}$.

In the present paper, we are mainly concerned with *intrinsic graphs* over a fixed (but arbitrary) vertical subgroup $\mathbb{W}$, which we often take to be the "(y,t)-plane" $\mathbb{W}_{y,t} := \{(x,y,t) \in \mathbb{H} : x = 0\}$. Let $\mathbb{V}_x$ be the horizontal subgroup $\mathbb{V}_x = \{(x,0,0) : x \in \mathbb{R}\} \subset \mathbb{H}$, and consider a function $\phi \colon \mathbb{W}_{y,t} \to \mathbb{V}_x$. The intrinsic graph of ϕ (over $\mathbb{W}_{y,t}$) is the set

$$\Gamma^{\phi} := \{w \cdot \phi(w) : w \in \mathbb{W}_{y,t}\} \subset \mathbb{H},$$

where "$\cdot$" refers to the Heisenberg product (see Section 2). Note that Γ^{ϕ} is, in general different, from the "Euclidean graph" $\{(\phi(y,t), y, t) : (y,t) \in \mathbb{W}_{y,t}\}$, and in fact Γ^{ϕ} need not be representable as the Euclidean graph of any function (see Example 2.5 in [27]).

Recall that a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ is (Euclidean) Lipschitz, if and only if there exists a cone, which, when centered at any point x on the graph of f, only intersects the graph at x. The notion of *intrinsic Lipschitz function* in $\mathbb{H}$ is **defined** with this characterization in mind, with "graph" replaced by "intrinsic graph", and "cone" replaced by a natural $\mathbb{H}$-analogue, see (2.2). Intrinsic Lipschitz functions were introduced by B. Franchi, R. Serapioni and F. Serra Cassano in [27], and they turned out to be very influential in the evolution of geometric analysis in Heisenberg groups, see for instance [1, 3, 5, 9, 14, 29, 28, 39]. Curiously, the definition does not guarantee that an intrinsic Lipschitz function is (metrically) Lipschitz between the spaces $\mathbb{W}_{y,t}$ and $\mathbb{V}_x$.

If the reader is familiar with the theory of rectifiability in metric spaces, but not with that in $\mathbb{H}$, she may wonder why such "intrinsic" notions are necessary in the first place; why cannot one study (metrically) Lipschitz images $\mathbb{R}^k \to \mathbb{H}$? The reason is simple: a Lipschitz image $f(\mathbb{R}^k) \subset \mathbb{H}$ has vanishing k-dimensional measure for $k \in \{2, 3, 4\}$. This is a result of L. Ambrosio and B. Kirchheim [2]. The work of Mattila, Serapioni and Serra Cassano [39] and Franchi, Serapioni and Serra Cassano [26, 28] suggests that intrinsic Lipschitz graphs, instead, are the correct class of sets to consider in connection with 3-rectifiability in the Heisenberg group. We believe that this is true also in the quantitative setting.

Lipschitz graphs in $\mathbb{R}^n$ are, arguably, the most fundamental examples of quantitatively rectifiable sets in the sense of David and Semmes. In the present paper, we propose that intrinsic Lipschitz graphs play the same role in $\mathbb{H}$. In $\mathbb{R}^n$, the term "quantitatively rectifiable" has many meanings; the fundamental results of David and Semmes show that a certain class of sets enjoys – and can be characterized – by a wide variety of properties, both geometric and analytic, each of which could be taken as the definition of "quantitatively rectifiable". In $\mathbb{H}$, no such results are available, yet, so we have to specify our viewpoint to "quantitative rectifiability". It will be that of "quantitative affine approximation". Theorem 1.1, informally stated, says that intrinsic Lipschitz graphs admit good affine approximations "at most places and scales".

The traditional way to quantify such a statement is via the notion of β-numbers, introduced by P. Jones in [33] in order to control the Cauchy singular integral on 1-dimensional Lipschitz graphs. They were later used by Jones [34] and David and Semmes [19, 18] in order to characterize quantitative rectifiability. The same approach works in $\mathbb{H}$, if the β-numbers are defined correctly. In Definition 3.3 below, we introduce the *vertical β-numbers*. These nearly coincide with the usual (Euclidean) β-numbers, defined with respect to the metric in $\mathbb{H}$ of course: the single, crucial, difference is that approximating affine planes are restricted to sets of the form $z \cdot \mathbb{W}'$, where $z \in \mathbb{H}$ and $\mathbb{W}'$ is a vertical subgroup. Viewing $\mathbb{H}$ as $\mathbb{R}^3$ for a moment, these sets are simply (Euclidean) translates of the sets $\mathbb{W}'$. So, they are quite literally vertical planes.

Here is, finally, the main result:

Theorem 1.1. *An intrinsic Lipschitz graph in $\mathbb{H}$ satisfies the weak geometric lemma for the vertical β-numbers.*

For a more precise restatement see Theorem 4.2. In brief, the weak geometric lemma states that, for any fixed ε, the vertical β-numbers of the graph have size at most ε in all balls, except perhaps a family satisfying a Carleson packing condition (with constants depending on ε). This manner of quantifying the "smallness" of an exceptional family of balls is ubiquitous in the theory of David and Semmes.

Theorem 1.1 does not explain our need to define the vertical β-numbers; since the vertical β-numbers are, evidently, at least as large as the "usual" ones (with no restrictions on the approximating affine planes), the statement of Theorem 1.1 merely becomes a little weaker, if the word "vertical" is removed. However, it turns out that the weak geometric lemma for the vertical β-numbers, combined with a condition on *vertical projections*, essentially **characterizes** intrinsic Lipschitz graphs. This is the content of our second main result, a counterpart of a theorem of David and Semmes [19] from 1990:

Theorem 1.2. *Assume that a 3-regular set $E \subset \mathbb{H}$ satisfies the weak geometric lemma for the vertical β-numbers, and has big vertical projections. Then E has big pieces of intrinsic Lipschitz graphs.*

As before, we postpone explaining the notions of *big vertical projections* (Definition 3.1) and *big pieces of intrinsic Lipschitz graphs* (Definition 3.2). The latter condition does not guarantee that E is an intrinsic Lipschitz graph (such a statement would be false, rather obviously). Instead, $E \cap B$ contains a measure-theoretically big piece of an intrinsic Lipschitz for every ball B centered on E.

Finally, we mention that Theorem 1.2 admits a converse, which follows from Theorem 1.1 by standard considerations, outlined at the end of the paper:

Theorem 1.3. *Assume that a 3-regular set $E \subset \mathbb{H}$ has big pieces of intrinsic Lipschitz graphs. Then $E \subset \mathbb{H}$ satisfies the weak geometric lemma for the vertical β-numbers, and has big vertical projections.*

So, the property of having big pieces of intrinsic Lipschitz graphs is characterized by the combination of the weak geometric lemma for the vertical β-numbers, and the big vertical projections condition.

We remark that for 1-dimensional sets (in contrast to 3-dimensional sets in the present paper), quantitative rectifiability in $\mathbb{H}$ has been studied earlier. F. Ferrari, B. Franchi, and H. Pajot [24], N. Juillet [35] and recently S. Li and R. Schul [37], [36], considered the validity of the traveling salesman theorem of P. Jones [34]. Already in $\mathbb{R}^n$, there is a considerable difference in the techniques required to treat higher (than one) dimensional quantitative rectifiability. In $\mathbb{H}$, that difference is even more pronounced because of the result of Ambrosio and Kirchheim

[2] mentioned earlier: 1-dimensional rectifiable sets are, essentially, metric Lipschitz images of $\mathbb{R}$, whereas for 3-dimensional sets one needs another approach. Finding such an approach is well-motivated: the critical dimension for the removability problem in $\mathbb{H}$ is 3, and the development of quantitative rectifiability in this dimension is essential for making progress in that direction.

The paper is organized as follows. In Section 2 we lay down the necessary background in the Heisenberg group and we also discuss intrinsic Lipschitz graphs and some of their main properties. In Section 3, we give sufficient conditions for a 3-dimensional set in the first Heisenberg group to have big pieces of intrinsic Lipschitz graphs. Section 4 is devoted to the proof of Theorem 1.1 and is definitely the most technical part of the paper. To facilitate the reader's navigation through the somewhat lengthy Section 4 a second introductory part appears in Section 4.1.

Acknowledgements. Part of the research for this paper was completed while various subsets of the authors visited the Universities of Bern, Helsinki and Connecticut. The hospitality of these institutions is acknowledged.

2. Preliminaries

The *Heisenberg group* $\mathbb{H}$ is $\mathbb{R}^3$ endowed with the group law

$$(x,y,t)\cdot(x',y',t') = (x+x', y+y', t+t'+(xy'-yx')/2) \tag{2.1}$$

for $(x,y,t),(x',y',t')\in\mathbb{R}^3$. We will sometimes identify $\mathbb{R}^3$ with $\mathbb{C}\times\mathbb{R}$ and denote points in the Heisenberg group by (z,t) for $z = x+iy\in\mathbb{C}$ and $t\in\mathbb{R}$.

We use the following metric on $\mathbb{H}$:

$$d_{\mathbb{H}} : \mathbb{H}\times\mathbb{H}\to[0,\infty), \quad d_{\mathbb{H}}(p,q) := \|q^{-1}\cdot p\|,$$

where

$$\|(z,t)\| := \max\left\{|z|, |t|^{1/2}\right\}.$$

The closed balls in $(\mathbb{H}, d_{\mathbb{H}})$ will be denoted by $B(x,r)$. We will also denote by $\mathcal{H}^s$ the s-dimensional Hausdorff measure in $(\mathbb{H}, d_{\mathbb{H}})$. The reader who is not familiar with the notion of Hausdorff measures should have a look at [38]. For more information on the Heisenberg group, see for instance the book [11] by Capogna, Danielli, Pauls and Tyson. Here we just mention that $\dim_{\mathrm{H}}\mathbb{H} = 4$, and the usual Lebesgue measure on $\mathbb{R}^3$ coincides (up to a constant) with $\mathcal{H}^4$ on $\mathbb{H}$.

The distance $d_{\mathbb{H}}$ is invariant with respect to left translations

$$\tau_p : \mathbb{H}\to\mathbb{H}, \quad \tau_p(q) = p\cdot q, \quad (p\in\mathbb{H}),$$

and homogeneous with respect to dilations

$$\delta_r : \mathbb{H}\to\mathbb{H}, \quad \delta_r((z,t)) = (rz, r^2 t), \quad (r>0).$$

Recall that a closed set $E\subset\mathbb{H}$ is 3-*(Ahlfors-David)-regular*, if there exists a constant $1\le C<\infty$, the *regularity constant* of E, such that

$$C^{-1}r^3 \le \mathcal{H}^3(B(x,r)\cap E) \le Cr^3$$

for all $x \in E$ and $0 < r \leq \operatorname{diam}(E)$.

We stress once more that metric concepts in $\mathbb{H}$, such as "ball", "Hausdorff measure" or "Ahlfors-David regular" are always defined with respect to $d_{\mathbb{H}}$, unless explicitly stated otherwise.

Identify $\mathbb{H}$ with $\mathbb{C} \times \mathbb{R}$ for a moment. If $V \subset \mathbb{C}$ is a line through the origin, then $\mathbb{V} := V \times \{0\}$ is called a *horizontal subgroup of* $\mathbb{H}$. A *vertical subgroup of* $\mathbb{H}$ is a set of the form $\mathbb{W} = V \times \mathbb{R}$, where $V \subset \mathbb{C}$ is a line through the origin. Note that both $\mathbb{W}$ and $\mathbb{V}$ are subgroups of $\mathbb{H}$, and closed under the action of δ_r.

Under the identification of $\mathbb{W}$ with $\mathbb{R}^2$, the subgroup $\mathbb{W}$ can be endowed with the 2-dimensional Lebesgue measure $\mathcal{L}^2$. This turns out to be a Haar measure on $(\mathbb{W}, \cdot)$, and it agrees (up to a multiplicative constant) with $\mathcal{H}^3$ on $(\mathbb{W}, d_{\mathbb{H}})$, see [39, Proposition 2.20].

Definition 2.1 (Complementary subgroups)**.** Given a vertical subgroup $\mathbb{W} = V \times \mathbb{R}$ of $\mathbb{H}$, we define the *complementary horizontal subgroup*

$$\mathbb{V} := \mathbb{V}_{\mathbb{W}} := V^{\perp} \times \{0\},$$

where $V^{\perp}$ is the orthogonal complement of V in $\mathbb{C}$. Then every point $p \in \mathbb{H}$ can be written **uniquely** as $p = p_{\mathbb{W}} \cdot p_{\mathbb{V}}$ with $p_{\mathbb{W}} \in \mathbb{W}$ and $p_{\mathbb{V}} \in \mathbb{V}$.

One could also consider other splittings of the Heisenberg group, but in this paper we will always assume that the groups $\mathbb{W}$ and $\mathbb{V}_{\mathbb{W}}$ are *orthogonal*; by this we mean that V and $V^{\perp}$ are orthogonal as above.

Definition 2.2 (Horizontal and vertical projections)**.** Let $\mathbb{W} = V \times \mathbb{R}$ be a vertical subgroup with complementary horizontal subgroup $\mathbb{V} = V^{\perp} \times \{0\}$. As we observed, every point $p \in \mathbb{H}$ can be written uniquely as $p = p_{\mathbb{W}} \cdot p_{\mathbb{V}}$ with $p_{\mathbb{W}} \in \mathbb{W}$ and $p_{\mathbb{V}} \in \mathbb{V}$. This gives rise to the *vertical projection* $\pi_{\mathbb{W}}$ and *horizontal projection* $\pi_{\mathbb{V}}$, defined by

$$\pi_{\mathbb{W}}(p) := p_{\mathbb{W}} \quad \text{and} \quad \pi_{\mathbb{V}}(p) := p_{\mathbb{V}}.$$

The mappings $\pi_{\mathbb{W}}$ and $\pi_{\mathbb{V}}$ have the following explicit formulae:

$$\pi_{\mathbb{V}}(z,t) := (z,t)_{\mathbb{V}} = (\pi_{V^{\perp}}(z), 0),$$

and

$$\pi_{\mathbb{W}}(z,t) := (z,t)_{\mathbb{W}} = (\pi_V(z), t - 2\omega(\pi_V(z), \pi_{V^{\perp}}(z))),$$

for $(z,t) \in \mathbb{C} \times \mathbb{R} \cong \mathbb{H}$. Here π_V and $\pi_{V^{\perp}}$ are the usual orthogonal projections onto the lines V and $V^{\perp}$ in $\mathbb{C}$. We also used the abbreviating notation

$$\omega(z, z') = \tfrac{1}{4}\Im\left(\overline{z}z'\right) = \tfrac{1}{4}(xy' - yx')$$

for $z = x + \mathrm{i}y$ and $z' = x' + \mathrm{i}y'$.

The horizontal projections $\pi_{\mathbb{V}}$ are both (metrically) Lipschitz functions, and group homomorphisms. The vertical projections $\pi_{\mathbb{W}}$ are neither. However, as we will see many times in this paper, the projections $\pi_{\mathbb{W}}$ interact well with intrinsic Lipschitz graphs, defined below.

Definition 2.3 ($\mathbb{H}$-cones)**.** Let $\mathbb{W}$ be a vertical subgroup with complementary horizontal subgroup $\mathbb{V}$. An *$\mathbb{H}$-cone perpendicular to $\mathbb{W}$ and with aperture $\alpha > 0$* is the following set $C_{\mathbb{W}}(\alpha)$:

$$C_{\mathbb{W}}(\alpha) := \{p \in \mathbb{H} : \|p_{\mathbb{W}}\| \leq \alpha \|p_{\mathbb{V}}\|\}. \tag{2.2}$$

Definition 2.4 (Intrinsic Lipschitz graphs and functions)**.** A subset $\Gamma \subset \mathbb{H}$ is an *intrinsic L-Lipschitz graph over a vertical subgroup* $\mathbb{W}$, if

$$(x \cdot C_{\mathbb{W}}(\alpha)) \cap \Gamma = \{x\} \quad \text{for } x \in \Gamma \text{ and } 0 < \alpha < \tfrac{1}{L}. \tag{2.3}$$

If $A \subset \mathbb{W}$ is any set, and $\mathbb{V}$ is the complementary horizontal subgroup of $\mathbb{W}$, we say that a function $\phi \colon A \to \mathbb{V}$ is *an intrinsic L-Lipschitz function*, if the *intrinsic graph of ϕ*, namely

$$\Gamma^{\phi} := \{w \cdot \phi(w) : w \in A\} \subset \mathbb{H},$$

is an intrinsic L-Lipschitz graph.

The *intrinsic Lipschitz constant* of ϕ (or Γ) is defined as the infimum over all constants L for which ϕ (or Γ) is intrinsic L-Lipschitz.

For a nice picture of intrinsic Lipschitz graphs and the $\mathbb{H}$-cones, see Section 3 in [27].

Remark 2.5 (Parametrisation of intrinsic Lipschitz graphs)**.** An intrinsic Lipschitz graph can be uniquely parametrised by an intrinsic Lipschitz function. More precisely, given an intrinsic L-Lipschitz graph Γ over a vertical subgroup $\mathbb{W}$, there exists a unique intrinsic L-Lipschitz function $\phi_{\Gamma} \colon \pi_{\mathbb{W}}(\Gamma) \to \mathbb{V}$ such that $\Gamma = \Gamma^{\phi_{\Gamma}}$.

To see this, one first checks that the cone condition (2.3) implies the injectivity of $\pi_{\mathbb{W}}|_{\Gamma}$. Indeed, assume that $\pi_{\mathbb{W}}(x) = \pi_{\mathbb{W}}(y)$ for some $x, y \in \Gamma$. Writing (uniquely) $x = x_{\mathbb{W}} \cdot x_{\mathbb{V}}$ and $y = y_{\mathbb{W}} \cdot y_{\mathbb{V}}$, this gives

$$\pi_{\mathbb{W}}(x^{-1} \cdot y) = \pi_{\mathbb{W}}(x_{\mathbb{V}}^{-1} \cdot y_{\mathbb{V}}) = 0,$$

since $x_{\mathbb{V}}^{-1} \cdot y_{\mathbb{V}} \in \mathbb{V}$, and $\pi_{\mathbb{W}}$ annihilates $\mathbb{V}$. Hence, $x^{-1} \cdot y \in C_{\mathbb{W}}(\alpha)$ for any $\alpha > 0$, and so $y \in (x \cdot C_{\mathbb{W}}(\alpha)) \cap \Gamma$, implying $x = y$. Consequently, the following mapping $\phi_{\Gamma} \colon \pi_{\mathbb{W}}(\Gamma) \to \mathbb{V}$ is well-defined:

$$\phi_{\Gamma}(\pi_{\mathbb{W}}(x)) := \pi_{\mathbb{V}}(x).$$

The mapping ϕ_{Γ} clearly satisfies $\Gamma = \Gamma^{\phi_{\Gamma}}$, and thus ϕ_{Γ} is intrinsic L-Lipschitz by definition. The uniqueness of ϕ_{Γ} follows from the uniqueness of the representation $x = x_{\mathbb{W}} \cdot x_{\mathbb{V}}$, $x \in \mathbb{H}$. We refer to ϕ_{Γ} as the *parametrisation of* Γ.

A key property of intrinsic Lipschitz graphs is that they are invariant under left translations and dilations in $\mathbb{H}$; if Γ is an intrinsic L-Lipschitz graph, then $\delta_r(\tau_p(\Gamma))$ is also an intrinsic L-Lipschitz graph for any $p \in \mathbb{H}$ and $r > 0$. This is why these sets are called "intrinsic"!

Remark 2.6. We will often "without loss of generality" assume that the intrinsic Lipschitz functions and graphs under consideration are defined over the specific vertical plane $\mathbb{W}_{y,t} := \{(0,y,t) : y,t \in \mathbb{R}\}$. This is legitimate, because the notions are invariant under the rotations R_θ around the t-axis, given by $R_\theta(z,t) := (e^{i\theta}z,t)$. The rotations are both group homomorphisms and isometries with respect to the metric $d_{\mathbb{H}}$. The homomorphism property implies that $\pi_{R_\theta\mathbb{W}} \circ R_\theta = R_\theta \circ \pi_{\mathbb{W}}$, and then the isometry property gives the following: if Γ is an intrinsic L-Lipschitz graph over $\mathbb{W}$, then $R_\theta\Gamma$ is an intrinsic L-Lipschitz graph over $R_\theta\mathbb{W}$.

The class of intrinsic Lipschitz functions has (in greater generality) been introduced and studied by Franchi, Serapioni and Serra Cassano. The following facts are special cases of the results in [28] and [30]:

- For $A \subset \mathbb{W}$, an intrinsic L-Lipschitz function $\phi : A \to \mathbb{V}$ can be extended to an intrinsic L'-Lipschitz function $\mathbb{W} \to \mathbb{V}$, where L' depends only on L.
- An intrinsic Lipschitz function $\phi : \mathbb{W} \to \mathbb{V}$ is intrinsically differentiable $\mathcal{L}^2$ almost everywhere on $\mathbb{W}$.
- An intrinsic L-Lipschitz graph over $\mathbb{W}$ is 3-regular with regularity constant depending only on L.

We will write more about intrinsic differentiability and the related notion of *intrinsic gradient* in Section 4.2.

If $A, B > 0$, we will use the notation $A \lesssim_p B$ to signify that there exists a constant $C \geq 1$ depending only on p such that $A \leq CB$. If the constant C is absolute, we write $A \lesssim B$. The two-sided inequality $A \lesssim_p B \lesssim_p A$ is abbreviated to $A \sim_p B$.

3. A SUFFICIENT CONDITION FOR BIG PIECES OF INTRINSIC LIPSCHITZ GRAPHS

In this section, we start proving our two main results. To warm up, we begin with the significantly easier Theorem 1.2: if a 3-regular set $E \subset \mathbb{H}$ satisfies the weak geometric lemma for vertical β-numbers, and has big vertical projections (see definitions below), then E has BPiLG. The argument is very similar to the Euclidean counterpart, due to David and Semmes [19]. In fact, the greatest surprise here is probably the similarity of the arguments itself: considering that the vertical projections $\pi_{\mathbb{W}} : \mathbb{H} \to \mathbb{W}$ are not Lipschitz, one might expect a rockier ride ahead.

We start with a few central definitions and auxiliary results.

Definition 3.1 (BVP)**.** A 3-regular set $E \subseteq \mathbb{H}$ is said to have *big vertical projections* (BVP in short) if there exists a constant $\delta > 0$ with the following property: for all $x \in E$ and for all $0 < R \leq \mathrm{diam}_{\mathbb{H}}(E)$ there exists a vertical subgroup $\mathbb{W}$ such that

$$\mathcal{L}^2(\pi_{\mathbb{W}}(E \cap B(x,R))) \geq \delta R^3.$$

Definition 3.2 (BPiLG)**.** A 3-regular set $E \subseteq \mathbb{H}$ has *big pieces of intrinsic Lipschitz graphs* (BPiLG in short) if there exist constants $L \geq 1$ and $\theta > 0$ with the following property. For all $x \in E$ and $0 < R \leq \mathrm{diam}_{\mathbb{H}}(E)$, there exists an intrinsic L-Lipschitz graph Γ over some vertical subgroup such that

$$\mathcal{H}^3(E \cap \Gamma \cap B(x,R)) \geq \theta R^3.$$

Definition 3.3 (Vertical β-numbers). Let $E \subset \mathbb{H}$ be a set, let $B \subset \mathbb{H}$ be a ball with radius $r(B) > 0$ centered on E, let $\mathbb{W}$ be a vertical subgroup, and let $z \in \mathbb{H}$. We write

$$\beta_E(B; z \cdot \mathbb{W}) := \sup_{y \in B \cap E} \frac{\mathrm{dist}_{\mathbb{H}}(y, z \cdot \mathbb{W})}{r(B)},$$

and then we define the *vertical β-number* as

$$\beta(B) := \beta_E(B) := \inf_{\mathbb{W}, z} \beta_E(B; z \cdot \mathbb{W}).$$

The infimum is taken over all vertical subgroups $\mathbb{W}$, and all points $z \in \mathbb{H}$.

Remark 3.4. The following observation is useful, and not quite as trivial as its Euclidean counterpart. Assume $B \subset \mathbb{H}$ is a ball centered on E, $\mathbb{W}$ is any vertical subgroup, and $z \in \mathbb{H}$. Then

$$\sup_{x,y \in B \cap E} \frac{\mathrm{dist}_{\mathbb{H}}(x, y \cdot \mathbb{W})}{r(B)} \leq 2\beta_E(B; z \cdot \mathbb{W}).$$

To prove this, observe that $p \cdot \mathbb{W} = \mathbb{W} \cdot p$ for any point $p \in \mathbb{H}$. In particular, if $y \in B \cap E$, we have

$$y \cdot \mathbb{W} = z \cdot \mathbb{W} \cdot z^{-1} \cdot y.$$

Hence, if further $x \in B \cap E$ and $w, w' \in \mathbb{W}$, we have

$$\mathrm{dist}_{\mathbb{H}}(x, y \cdot \mathbb{W}) \leq d_{\mathbb{H}}(x, z \cdot w' \cdot w^{-1} \cdot z^{-1} \cdot y) \leq d_{\mathbb{H}}(x, z \cdot w') + d_{\mathbb{H}}(y, z \cdot w).$$

Since this holds for all $w, w' \in \mathbb{W}$, we have

$$\frac{\mathrm{dist}_{\mathbb{H}}(x, y \cdot \mathbb{W})}{r(B)} \leq 2\beta_E(B; z \cdot \mathbb{W}),$$

as claimed.

Definition 3.5 (WGL). We say that a 3-regular set $E \subseteq \mathbb{H}$ satisfies the *weak geometric lemma for vertical β-numbers* (WGL in short), if

$$\int_0^R \int_{E \cap B(x,R)} \chi_{\{(y,s) \in E \times \mathbb{R}_+ :\, \beta(B(y,s)) > \varepsilon\}}(y, s)\, d\mathcal{H}^3(y) \frac{ds}{s} \lesssim_\varepsilon R^3$$

for all $\varepsilon > 0$, $x \in E$ and $R > 0$.

The following lemma shows that even if the vertical projections $\pi_{\mathbb{W}}$ are not Lipschitz, they still cannot increase $\mathcal{H}^3$-measure (too much). This is rather surprising, as it is easy to find less than three-dimensional sets $E \subset \mathbb{H}$ such that $\dim_{\mathrm{H}} \pi_{\mathbb{W}}(E) > \dim_{\mathrm{H}} E$, see Example 4.1 in [6].

Lemma 3.6. *Let $\mathbb{W}$ be a vertical subgroup in $\mathbb{H}$. Then there exists a constant $0 < C < \infty$ such that for all $A \subseteq \mathbb{H}$, one has*

$$\mathcal{L}^2(\pi_{\mathbb{W}}(A)) \leq C\mathcal{H}^3(A). \tag{3.1}$$

Proof. The lemma follows from [30, Lemma 2.20], which, when specialized to the Heisenberg group, states that there is a constant $C > 0$ such that $\mathcal{L}^2(\pi_{\mathbb{W}}(B(p,r)) = Cr^3$ for all $p \in \mathbb{H}$ and $r > 0$. See also [27, Lemma 3.14].

Given a set $A \subset \mathbb{H}$ of positive $\mathcal{H}^3$-measure, choose a covering of A by closed balls $B_i = B(p_i, r_i)$, $i \in \mathbb{N}$, such that

$$\sum_{i \in \mathbb{N}} r_i^3 \leq 2\mathcal{H}^3(A).$$

Then, we find that

$$\mathcal{L}^2(\pi_{\mathbb{W}}(A)) \leq \sum_{i \in \mathbb{N}} \mathcal{L}^2(\pi_{\mathbb{W}}(B_i)) = C \sum_{i \in \mathbb{N}} r_i^3 \leq 2C\mathcal{H}^3(A).$$

This completes the proof. □

3.0.1. *David cubes.* We recall the construction of *David cubes*, first introduced by David in [15], which can be defined on any regular set in a geometrically doubling metric space. Let $E \subset \mathbb{H}$ be a 3-regular set. Then, there exists a family of partitions Δ_j of E, $j \in \mathbb{Z}$, with the following properties:

(i) If $j \leq k$, $Q \in \Delta_j$ and $Q' \in \Delta_k$, then either $Q \cap Q' = \emptyset$, or $Q \subset Q'$.
(ii) If $Q \in \Delta_j$, then $\operatorname{diam}_{\mathbb{H}} Q \leq 2^j$.
(iii) Every set $Q \in \Delta_j$ contains a set of the form $B(z_Q, c2^j) \cap E$ for some $z_Q \in Q$, and some constant $c > 0$.

The sets in $\Delta := \cup \Delta_j$ are called David cubes, or just cubes, of E. For $Q \in \Delta_j$, we define $\ell(Q) := 2^j$. Thus, by (ii), we have $\operatorname{diam}_{\mathbb{H}}(Q) \leq \ell(Q)$ for $Q \in \Delta$. Given a fixed cube $Q_0 \in \Delta$, we write

$$\Delta(Q_0) := \{Q \in \Delta : Q \subset Q_0\}.$$

It follows from (ii), (iii), and the 3-regularity of E that $\mathcal{H}^3(Q) \sim \ell(Q)^3$ for $Q \in \Delta_j$. It is an easy fact, needed a bit later, that the following holds: if $x, y \in E$ are distinct points, there exists an index $j \in \mathbb{Z}$, and disjoint cubes $Q_x, Q_y \in \Delta_j$, containing x and y, respectively, with the properties that $2^j \sim d_{\mathbb{H}}(x,y)$,

$$Q_x \subset B(z_{Q_y}, 4\ell(Q_y)) \quad \text{and} \quad Q_y \subset B(z_{Q_x}, 4\ell(Q_x)).$$

Indeed, let $j \in \mathbb{Z}$ be the largest integer such that $2^j \leq d_{\mathbb{H}}(x,y)$, and let $Q_x, Q_y \in \Delta_j$ be the unique cubes containing x and y. Then $2^j \sim d_{\mathbb{H}}(x,y)$, and since $d_{\mathbb{H}}(x,y) \leq 2\ell(Q_x)$, we have $Q_y \subset B(z_{Q_x}, 4\ell(Q_x))$. The same holds with the roles of x and y reversed.

In the sequel, we write

$$B_Q := B(z_Q, 4\ell(Q)), \qquad Q \in \Delta.$$

The weak geometric lemma (Definition 3.5) implies the following reformulation in terms of David cubes. Write

$$\beta(Q) := \beta(B_Q) \qquad Q \in \Delta.$$

Then

$$\sum_{\{Q\in\Delta(Q_0):\beta(Q)\geq\varepsilon\}} \mathcal{H}^3(Q) \lesssim_\varepsilon \mathcal{H}^3(Q_0) \tag{3.2}$$

for any $\varepsilon > 0$ and $Q_0 \in \Delta$. Deriving this property from Definition 3.5 is an easy exercise using the properties of David cubes, and we omit the details.

3.1. **Proof of Theorem 1.2.** Here is the statement of Theorem 1.2 once more:

Theorem 3.7. *A 3-regular set in $\mathbb{H}$ with BVP and satisfying the WGL has BPiLG.*

Let Δ be a system of David cubes on E, let $c, \varepsilon > 0$ be constants, and let $\mathbb{W}$ be a vertical subgroup. Throughout this section, a cube $Q \in \Delta$ will be called *good* (more precisely $(c, \varepsilon, \mathbb{W})$-good), if

$$\mathcal{L}^2(\pi_{\mathbb{W}}(Q)) \geq c\mathcal{H}^3(Q), \tag{3.3}$$

and

$$\beta(Q) \leq \varepsilon. \tag{3.4}$$

We outline the proof of Theorem 3.7. The proof is divided into two parts, a geometric one, and an abstract one. The geometric part shows that good cubes $Q \in \Delta$ are already "almost" intrinsic Lipschitz graphs in the following sense: if $x \in Q$ and $y \in B_Q \cap E$ satisfy $d_{\mathbb{H}}(x, y) \sim \ell(Q)$, then $y \notin x \cdot C_{\mathbb{W}}(\alpha)$ for some small $\alpha > 0$.

The abstract part uses the WGL and BVP assumptions to infer (cutting a few corners here) that only a small fraction of $E \cap B(z, R)$, $z \in E$, meets B_Q for some non-good cube Q. Hence, a large set $F \subset E \cap B(z, R)$ meets B_Q only for good cubes Q. Unfortunately, this is not literally true, and additional (technical) considerations are needed. Ignoring these for now, we can complete the proof as follows. Fixing $x, y \in F$, we can use the discussion in Section 3.0.1 to find a cube $Q \in \Delta$ with $x \in Q$, $y \in B_Q$ and $\ell(Q) \sim d_{\mathbb{H}}(x, y)$. Since $x, y \in F$, we know that Q is a good cube, and it follows from the geometric part that $y \notin C_{\mathbb{W}}(\alpha)$. Consequently, $F \subset B(z, R) \cap E$ is an intrinsic Lipschitz graph.

3.1.1. *The geometric part.* The following lemma is our counterpart of Lemma 2.19 in David and Semmes' proof in [19].

Lemma 3.8. *Assume that Q is a $(c, \varepsilon, \mathbb{W})$-good cube, $x \in Q$ and $y \in B_Q \cap E$ with $d_{\mathbb{H}}(x, y) \sim \ell(Q)$. Then $y \notin x \cdot C_{\mathbb{W}}(\alpha)$, if ε is sufficiently small with respect to c, and $\alpha > 0$ is small enough (depending on the constants c, ε).*

Proof. We start with a reduction to "unit scale". Assume that the statement of the lemma fails for certain parameters $c, \varepsilon, \mathbb{W}, \alpha$, and a certain $(c, \varepsilon, \mathbb{W})$-good cube $Q \in \Delta$. By this, we mean that (3.3) and (3.4) hold for Q, yet $y \in x \cdot C_{\mathbb{W}}(\alpha)$ for some $x \in Q$ and $y \in B_Q \cap E$ with $d_{\mathbb{H}}(x, y) \sim \ell(Q) =: r$.

Consider the set $Q_{x,r} := \delta_{1/r}(x^{-1} \cdot Q)$, and observe that it, also, satisfies (3.3), since $\mathcal{L}^2(\pi_{\mathbb{W}}(Q_{x,r})) = r^{-3}\mathcal{L}^2(\pi_{\mathbb{W}}(Q))$ and $\mathcal{H}^3(Q_{x,r}) = r^{-3}\mathcal{H}^3(Q)$. The first equation is not altogether trivial, but it follows from the equations

$$\mathcal{L}^2(\pi_{\mathbb{W}}(Q_{x,r})) = r^{-3}\mathcal{L}^2(\pi_{\mathbb{W}}(x^{-1} \cdot Q)) = r^{-3}\mathcal{L}^2(\pi_{\mathbb{W}}(x^{-1} \cdot \pi_{\mathbb{W}}(Q))),$$

and the fact that the mapping $P_p : \mathbb{W} \to \mathbb{W}$, $P_p(w) = \pi_{\mathbb{W}}(p \cdot w)$, has unit Jacobian for any fixed $p \in \mathbb{H}$, see the proof of Lemma 2.20 in [30], or (4.9) below.

Further, if $B_{Q_{x,r}} := \delta_{1/r}(x^{-1} \cdot B_Q)$, then $\beta(Q_{x,r}) := \beta(B_{Q_{x,r}}) \leq \varepsilon$. Here, β denotes the vertical β-number associated with $E_{x,r} := \delta_{1/r}(x^{-1} \cdot E)$. Note that $E_{x,r}$ is 3-regular with the same constants as E. Finally, note that $0 \in Q_{x,r}$, and

$$y_{x,r} := \delta_{1/r}(x^{-1} \cdot y) \in B_{Q_{x,r}} \cap E_{x,r} \cap C_{\mathbb{W}}(\alpha)$$

with $d_{\mathbb{H}}(0, y_{x,r}) \sim 1$. To sum up, if the lemma fails for Q, then we can construct another 3-regular set $E_{x,r}$, and another good David cube $Q_{x,r}$ (for $E_{x,r}$) with $0 \in Q_{x,r}$ and $\ell(Q_{x,r}) = 1$, such that the lemma fails for $Q_{x,r}$. Thus, it suffices to prove the lemma for a David-cube Q with the additional properties $0 \in Q$ and $\ell(Q) = 1$.

To this end, assume to the contrary that $y \in C_{\mathbb{W}}(\alpha)$ with $d_{\mathbb{H}}(0, y) \sim 1$. We will use this to show that the entire projection $\pi_{\mathbb{W}}(Q)$ is contained in a small neighbourhood of the t-axis T. This will violate (3.3). Somewhat abusing notation, we write

$$\mathbb{W}^{\perp} := V^{\perp} \times \mathbb{R}.$$

Let $p \in \mathbb{H}$ and $\mathbb{W}'$ be such that

$$\beta_E(B_Q; p \cdot \mathbb{W}') \leq 2\beta(Q) \leq 2\varepsilon.$$

The first task is to show that the angle $\theta(\mathbb{W}', \mathbb{W}^{\perp})$ between $\mathbb{W}' = V' \times \mathbb{R}$ and $\mathbb{W}^{\perp}$ satisfies

$$\theta(\mathbb{W}', \mathbb{W}^{\perp}) \lesssim \alpha + \varepsilon. \tag{3.5}$$

Write $y = (y_H, y_t)$. The plan is to use the smallness of $\beta_E(B_Q; p \cdot \mathbb{W}')$ in order to find a point $w' =: (w'_H, w'_t)$ on $\mathbb{W}'$, but close to y, such that

$$|\pi_V(w'_H)| \lesssim (\alpha + \varepsilon)|w'_H|. \tag{3.6}$$

This proves that the angle between $V^{\perp}$ and V' is $\lesssim (\alpha + \varepsilon)$ and thus (3.5), as claimed. In order to show (3.6), we first observe that the assumption of $y \in C_{\mathbb{W}}(\alpha)$ implies

$$\max\left\{|\pi_V(y_H)|, |y_t - 2\omega(\pi_V(y_H), \pi_{V^{\perp}}(y_H))|^{1/2}\right\} \lesssim \alpha|\pi_{V^{\perp}}(y_H)|. \tag{3.7}$$

Recalling that $d_{\mathbb{H}}(0, y) \sim 1$, this is only possible if

$$|y_H| \sim 1. \tag{3.8}$$

Indeed, we even have $|y_H|^2 \geq |y_t|$; otherwise the left hand side of (3.7) can be bounded from below by $|y_t/2|^{1/2}$, while the upper bound is then $\lesssim \alpha|y_H| \leq \alpha|y_t|^{1/2}$. For small enough α, this is impossible. Thus we may suppose (3.8).

Further, by Remark 3.4 we find for all $y' \in B_Q \cap E$ that

$$\mathrm{dist}_{\mathbb{H}}(y', \mathbb{W}') \leq 8\beta_E(B_Q; p \cdot \mathbb{W}')\ell(Q) \leq 16\varepsilon, \tag{3.9}$$

by our choice of p and $\mathbb{W}'$. In particular, for $y' = y$, there exists a vector $w' = (w'_H, w'_t) \in \mathbb{W}'$ with

$$|w'_H - y_H| \leq d_{\mathbb{H}}(y, w') \leq 16\varepsilon.$$

By (3.8), this gives $|w'_H| \sim 1$, and finally, using $y \in C_{\mathbb{W}}(\alpha)$,

$$|\pi_V(w'_H)| \leq |w'_H - y_H| + |\pi_V(y_H)| \lesssim (\alpha + \varepsilon)|w'_H|.$$

This proves (3.6) and (3.5).

So, we know that

(i) Q is close to $\mathbb{W}'$ (by (3.9)),
(ii) $\mathbb{W}'$ is close to $\mathbb{W}^\perp$ (by (3.5)).

As we will next demonstrate, $\pi_{\mathbb{W}}(Q)$ is close to $\pi_{\mathbb{W}}(\mathbb{W}^\perp) = T$.

Indeed, since we do not care about the best constants here, we can finish the proof very quickly: let $\tau_{\alpha,\varepsilon} > 0$ be a number such that if $w' \in \mathbb{W}' \cap B(0,2)$, then $d_{\mathbb{H}}(w', w^\perp) \leq \tau_{\alpha,\varepsilon}$ for some $w^\perp \in \mathbb{W}^\perp \cap B(0,3)$. Recalling (3.5), we can pick $\tau_{\alpha,\varepsilon}$ arbitrarily small by choosing α, ε small enough. Now, if $y' \in Q$, then by (3.9) and the triangle inequality, we have $d_{\mathbb{H}}(y', w^\perp) \leq 16\varepsilon + \tau_{\alpha,\varepsilon}$ for some $w^\perp \in \mathbb{W}^\perp \cap B(0,3)$. Since $\pi_{\mathbb{W}}$ is locally $1/2$-Hölder continuous, it follows that

$$\mathrm{dist}_{\mathbb{H}}(\pi_{\mathbb{W}}(y'), T) \leq d_{\mathbb{H}}(\pi_{\mathbb{W}}(y'), \pi_{\mathbb{W}}(w^\perp)) \lesssim d_{\mathbb{H}}(y', w^\perp)^{1/2} \leq (16\varepsilon + \tau_{\alpha,\varepsilon})^{1/2}.$$

The same holds with T replaced by $\pi_{\mathbb{W}}(B(0,3))$. Finally, the $\mathcal{L}^2$-measure of the $C(16\varepsilon + \tau_{\alpha,\varepsilon})^{1/2}$-neighbourhood of $T \cap \pi_{\mathbb{W}}(B(0,3))$ is bounded by a constant depending only on α, ε, and this constant tends to zero as $\alpha, \varepsilon \to 0$. For sufficiently small values of α, ε, this violates (3.3), and the proof is complete. $\square$

3.1.2. *The abstract part.* In this section, we apply Lemma 3.8 to good cubes inside a set $E \subset \mathbb{H}$ satisfying the weak geometric lemma for vertical β-numbers. This is a counterpart for Theorem 2.11 in [19], which in turn is modelled on a result of P. Jones [32]. The proof below is very similar to that in [19]; given Lemmas 3.6 and 3.8, the argument does not really see the difference between $\mathbb{H}$ and $\mathbb{R}^n$. We still record the full details.

Theorem 3.9. *Assume that $E \subseteq \mathbb{H}$ is a 3-regular set satisfying the WGL and let $b > 0$. Then there exist numbers $\alpha > 0$ and $M \in \mathbb{N}$, depending only on b and the 3-regularity and WGL constants of E, such that the following holds:*

For every David cube Q_0 in E and for all vertical projections $\pi_{\mathbb{W}}$, there exist intrinsic $(1/\alpha)$-Lipschitz graphs $F_j \subset Q_0$, $1 \leq j \leq M$,

$$\mathcal{L}^2(\pi_{\mathbb{W}}(Q_0 \setminus \cup F_j)) \leq b\mathcal{H}^3(Q_0).$$

Proof. Let E and b be as in the assumptions of Theorem 3.9. Let further $\varepsilon > 0$ be a small number to be chosen later (based on Lemma 3.8). Fix an arbitrary cube $Q_0 \in \Delta$, and an arbitrary vertical subgroup $\mathbb{W}$.

First, we will group the cubes in $\triangle(Q_0)$ into "good" and "bad" cubes, and control the quantity of the bad cubes via the WGL assumption. Second, Lemma 3.8, coupled with a "coding argument", will be used to partition the complement of the "bad" cubes in Q_0 into the sets F_j.

The "good" cubes $\mathcal{G}$ are the familiar $(b/2, \varepsilon, \mathbb{W})$-good cubes defined right above Lemma 3.8. The class $\mathcal{B}_1$ consists of those maximal (hence disjoint) cubes in Q_0

that violate the first goodness condition, i.e.,

$$\mathcal{B}_1 := \left\{ Q \in \triangle(Q_0) : \mathcal{L}^2(\pi_{\mathbb{W}}(Q)) < \tfrac{b}{2}\mathcal{H}^3(Q) \right\}.$$

Let $\mathcal{B}_2$ be the class of (all, not maximal) cubes that violate the second goodness condition:

$$\mathcal{B}_2 := \{ Q \in \triangle(Q_0) : \beta(Q) > \varepsilon \}.$$

Then, clearly, $\mathcal{G} = \triangle(Q_0) \setminus \bigcup_{j=1}^{2} \mathcal{B}_j$. It is also clear that the projections of bad cubes from the first class have small measure: for $R_1 := \bigcup_{Q \in \mathcal{B}_1} Q$, we have

$$\mathcal{L}^2(\pi_{\mathbb{W}}(R_1)) \leq \tfrac{b}{2} \cdot \mathcal{H}^3(R_1) \leq \tfrac{b}{2} \cdot \mathcal{H}^3(Q_0). \tag{3.10}$$

On the other hand, for the second bad class, one can control the measure of the cubes directly by the variant of WGL formulated in (3.2):

$$\sum_{Q \in \mathcal{B}_2} \mathcal{H}^3(Q) \leq C(\varepsilon)\mathcal{H}^3(Q_0).$$

Since $\mathcal{L}^2(\pi_{\mathbb{W}}(A)) \leq C\mathcal{H}^3(A)$ for all $A \subset \mathbb{H}$ by Lemma 3.6, the inequality above shows that the $\pi_{\mathbb{W}}$-projection of $\bigcup_{Q \in \mathcal{B}_2} Q$ has measure no larger than $CC(\varepsilon)\mathcal{H}^3(Q_0)$. This is a little bit too weak for our purposes; in analogy with (3.10), we wish to replace $CC(\varepsilon)$ by a small constant. To this end, we set

$$R_2 = \left\{ x \in Q_0 : \sum_{Q \in \mathcal{B}_2} \chi_{B_Q}(x) \geq N \right\}$$

where $N = N_{b,\varepsilon}$ is so large that $\mathcal{H}^3(R_2) \leq \frac{b}{2C}\mathcal{H}^3(Q_0)$. This is possible:

$$N\mathcal{H}^3(R_2) \leq \int_{Q_0} \sum_{Q \in \mathcal{B}_2} \chi_{B_Q}(x)\, d\mathcal{H}^3(x) \lesssim \sum_{Q \in \mathcal{B}_2} \mathcal{H}^3(Q) \leq C(\varepsilon) \cdot \mathcal{H}^3(Q_0).$$

With this definition of R_2, Lemma 3.6 gives

$$\mathcal{L}^2(\pi_{\mathbb{W}}(R_1 \cup R_2)) \leq \tfrac{b}{2} \cdot \mathcal{H}^3(Q_0) + \tfrac{b}{2} \cdot \mathcal{H}^3(Q_0) \leq b\mathcal{H}^3(Q_0).$$

It remains to find subsets $F_1, \ldots, F_M$ such that $Q_0 \setminus (R_1 \cup R_2) = \bigcup F_j$ and for every $j = 1, \ldots, M$ and every pair $x, y \in F_j, x \neq y$, it holds that $y \notin x \cdot C_{\mathbb{W}}(\alpha)$ for α small enough (only depending on b and the 3-regularity and WGL constants of E). This is done via a "coding argument", which goes back to Jones, see [19, p.866-867]. The argument is also explained briefly in David's book [17], p. 81–82, but we present the full details.

We start with a brief informal overview. Write $F := Q_0 \setminus (R_1 \cup R_2)$. Why do we need a "coding argument"? Maybe we can show, directly, that if $x, y \in F, x \neq y$, then $y \notin x \cdot C_{\mathbb{W}}(\alpha)$? Pick two distinct points $x, y \in F$, and pick two disjoint cubes $Q_x, Q_y \subset Q_0$ of some common generation such that $x \in Q_x$, $y \in Q_y$, $Q_y \subset B_{Q_x}$ and $d_{\mathbb{H}}(x, y) \sim \ell(Q)$ (such cubes exist, as discussed in Section 3.0.1). Now, since $Q_x \not\subset R_1$, we know that $\mathcal{L}^2(\pi_{\mathbb{W}}(Q_x)) \geq \frac{b}{2} \cdot \mathcal{H}^3(Q_x)$. If – and this is the "big if" – we also knew that $\beta(Q_x) \leq \varepsilon$ for $\varepsilon > 0$ small enough, we could infer from Lemma 3.8 that $y \notin x \cdot C_{\mathbb{W}}(\alpha)$ (assuming also that α is small enough). Of course, we do not know that $\beta(Q_x) \leq \varepsilon$ for the particular cube Q_x we are interested in: even though $x \notin R_2$, there can still be up to N "exceptional" cubes $Q \ni x$ such that $\beta(Q) > \varepsilon$. The "coding argument" is needed to fix this issue. Essentially, if we declare that $x \in F_1$, say, we want to make sure that the following holds: whenever $Q \ni x$ with $\beta(Q) > \varepsilon$, then all the points $y \in B_Q \setminus Q$ are stored safely away in the other sets F_j, $j \geq 2$. Once that has been accomplished, the argument above works for F_1 (or any F_j) in place of F.

We turn to the details, which are repeated from [17] nearly verbatim. For each cube $Q \subset Q_0$, we will associate a certain finite sequence of 0's and 1's, denoted by $\sigma(Q)$. The length of such a sequence is denoted by $|\sigma(Q)|$. We declare $\sigma(Q_0)$ to be the empty sequence.

Next, assume inductively that the numbers $\sigma(Q)$ have been defined for the descendants of $Q \subset Q_0$ down to a certain generation, say k. We now aspire to make the definition for cubes Q of generation $k+1$. If Q is such a cube, and Q^* is its parent, we initially set $\sigma(Q) := \sigma(Q^*)$.

Assume that $Q \in \mathcal{B}_2$, that is, $\beta(Q) > \varepsilon$, and assume that there exists at least one other cube Q_1 of the same generation as Q such that $Q_1 \subset B_Q$ (if either of these assumptions fails, we do not alter $\sigma(Q)$ now). Note that $\sigma(Q)$ and $\sigma(Q_1)$ have both been initially defined.

There are two cases to consider.

- Case 1: $|\sigma(Q)| = |\sigma(Q_1)|$. If $\sigma(Q) \neq \sigma(Q_1)$, we do not alter $\sigma(Q)$ or $\sigma(Q_1)$. But if $\sigma(Q) = \sigma(Q_1)$, we re-define $\sigma(Q)$ by adding a "0", and we re-define $\sigma(Q_1)$ by adding a "1".
- Case 2: $|\sigma(Q)| \neq |\sigma(Q_1)|$. If, for instance, $|\sigma(Q)| > |\sigma(Q_1)|$, then we do not alter $\sigma(Q)$. But we re-define $\sigma(Q_1)$ by adding either "0" or "1" to it in such a fashion that the new $\sigma(Q_1)$ is **not** an initial segment of $\sigma(Q)$. Finally, if $|\sigma(Q)| < |\sigma(Q_1)|$, then we repeat the same step with the roles of Q and Q_1 reversed.

After this procedure is complete, we pick another cube $Q_2 \subset B_Q$ of generation $k+1$ (if it exists), and perform the previous case chase with the pair (Q, Q_1) replaced by (Q, Q_2). Once all the pairs (Q, Q_i) with $Q_i \subset B_Q$ have been processed, we move on to other pairs (Q', Q_i) with $Q' \in \mathcal{B}_2$ and $Q_i \subset B_{Q'}$, and give them the same treatment as above.

The algorithm terminates eventually (because there are only finitely many cube-pairs to consider), and, at the end, every cube Q of generation $k+1$ has an associated sequence $\sigma(Q)$. If $Q \not\subset B_{Q'}$ for all cubes $Q' \in \mathcal{B}_2$ of generation $k+1$, then $\sigma(Q)$ retains the initial value $\sigma(Q^*)$. Even if $Q \subset B_{Q'}$ for some $Q' \in \mathcal{B}_2$, this can occur only for a bounded number, say C', of alternatives $Q' \in \mathcal{B}_2$ of generation $k+1$. Consequently, $\sigma(Q)$ differs from $\sigma(Q^*)$ by a sequence of length $\leq C'$.

By applying the procedure at all generations k, every sub-cube of Q_0 gets associated with a (finite) sequence $\sigma(Q)$. Next, we wish to extend the definition of these sequences from cubes to points in $F = Q_0 \setminus (R_1 \cup R_2)$. Fix $x \in F$, and let $Q_0 \supset Q_1 \supset \ldots$ be the unique sequence of dyadic cubes converging to x. As discussed in the previous paragraph, $\sigma(Q_{j+1})$ can differ from $\sigma(Q_j)$ only in case Q_{j+1} is contained in $B_{Q'}$ for some $Q' \in \mathcal{B}_2$ of the same generation as Q_{j+1}. By definition of $x \notin R_2$, there can only be $< N$ such indices j. In particular, the sequences $\sigma(Q_j)$ converge to some finite sequence of 0's and 1's, denoted by $\sigma(x)$. Furthermore, for those $< N$ indices j, where $\sigma(Q_{j+1})$ possibly differs from $\sigma(Q_j)$, this difference is a sequence of length at most C'. Consequently, the possible values of $\sigma(x)$ form a finite set S, whose cardinality can be bounded from above in terms of the constants N and C'. Given an element $s \in S$, we now define

$$F_s := \{x \in F : \sigma(x) = s\}.$$

It remains to check that the sets F_s satisfy the useful property we hinted at in the informal overview. Assume that $x, y \in F$ belong to "nearby" two cubes Q_x, Q_y of the same generation, namely with $Q_x \subset B_{Q_y}$ and $Q_y \subset B_{Q_x}$, and assume that either $\beta(Q_x) > \varepsilon$ or $\beta(Q_y) > \varepsilon$. Then, we claim that x and y belong to two different sets of the form F_s. Consider the sequences $\sigma(Q_x)$ and $\sigma(Q_y)$ (which are initial sequences in $\sigma(x)$ and $\sigma(y)$, respectively). Assume, for instance, that $\beta(Q_x) > \varepsilon$. This means that $Q_x \in \mathcal{B}_2$, so the pair (Q_x, Q_y) is considered while defining the sequences $\sigma(Q_x)$ and $\sigma(Q_y)$. Then, inspecting Case 1 and Case 2, it is clear that neither of the sequences $\sigma(Q_x)$ and $\sigma(Q_y)$ can be an initial sequence of the other. This proves that $\sigma(x) \neq \sigma(y)$, as claimed.

Now, we can quickly prove that if $x, y \in F_s, x \neq y$, then $y \notin x \cdot C_{\mathbb{W}}(\alpha)$, if $\alpha > 0$ is small enough. Pick $x, y \in F_s$, and let Q_x and Q_y be sub-cubes of Q_0, containing x and y, respectively, with same generation, satisfying $Q_x \subset B_{Q_y}$ and $Q_y \subset B_{Q_x}$, and with $\ell(Q) \sim d_{\mathbb{H}}(x, y)$. It follows from the claim in the previous paragraph that $\beta(Q_x) \leq \varepsilon$ and $\beta(Q_y) \leq \varepsilon$. Consequently, by Lemma 3.8, we have $y \notin x \cdot C_{\mathbb{W}}(\alpha)$ for small enough α. This completes the proof of the theorem. $\square$

We are now prepared to prove Theorem 3.7. Again, the proof is very similar to the Euclidean argument, see Theorem 1.14 in [19].

Proof of Theorem 3.7. Let E be a 3-regular subset of $\mathbb{H}$ with BVP, and satisfying the WGL. The former property ensures that for fixed $x \in E$ and $0 < R \leq \operatorname{diam}_{\mathbb{H}}(E)$, there is a vertical subgroup $\mathbb{W}$ such that

$$\mathcal{L}^2(\pi_{\mathbb{W}}(E \cap B(x, \tfrac{R}{2}))) \geq \delta \left(\tfrac{R}{2}\right)^3,$$

where δ is a constant depending only on E. Then, there exists $b > 0$, depending only on δ and the 3-regularity constant of E, and a David cube $Q_0 \subset E \cap B(x, R)$ such that $\mathcal{H}^3(Q_0) \sim R^3$ and

$$\mathcal{L}^2(\pi_{\mathbb{W}}(Q_0)) \geq 2b\mathcal{H}^3(Q_0). \tag{3.11}$$

Now we apply Theorem 3.9 to this particular cube Q_0 and choice of $b > 0$. It follows that there exist numbers $\alpha > 0$ and $M \in \mathbb{N}$ (depending only on b, and the 3-regularity and WGL constants of E) with the following property: there exist intrinsic $(1/\alpha)$-Lipschitz graphs $F_j \subset Q_0$, $1 \leq j \leq M$, such that

$$\mathcal{L}^2(\pi_{\mathbb{W}}(Q_0 \setminus \cup F_j)) \leq b\mathcal{H}^3(Q_0).$$

This, together with (3.11), implies

$$\sum_{j=1}^{M} \mathcal{L}^2(\pi_{\mathbb{W}}(F_j)) \geq \mathcal{L}^2\left(\pi_{\mathbb{W}}\left(\bigcup_{j=1}^{M} F_j\right)\right) \geq \mathcal{L}^2(\pi_W(Q_0)) - b\mathcal{H}^3(Q_0) \geq b\mathcal{H}^3(Q_0).$$

Thus there must exist some $1 \leq j \leq M$ such that

$$\mathcal{H}^3(F_j) \gtrsim \mathcal{L}^2(\pi_{\mathbb{W}}(F_j)) \geq b\mathcal{H}^3(Q_0)/M \sim bR^3/M.$$

The proof is complete. $\square$

4. The weak geometric lemma for intrinsic Lipschitz graphs

4.1. **Introduction, part II.** In the first half of the paper, we saw that any 3-Ahlfors-David regular subset of $\mathbb{H}$ with big vertical projections (BVP), and satisfying the weak geometric lemma (WGL), has big pieces of intrinsic Lipschitz graphs (BPiLG). The second half of the paper is devoted to proving the converse. The fact that intrinsic Lipschitz graphs (and thus BPiLG as well) have BVP is almost trivial, see Remark 4.21. Hence in order to prove Theorem 1.3 it suffices to prove the WGL for sets which have BPiLG.

Theorem 4.1. *If $E \subset \mathbb{H}$ has big pieces of intrinsic Lipschitz graphs with constants $L \geq 1, \theta > 0$, then it satisfies the weak geometric lemma; i.e.*

$$\int_0^R \int_{E \cap B(x,R)} \chi_{\{(y,s) \in E \times \mathbb{R}_+ \colon \beta(B(y,s)) \geq \varepsilon\}}(y,s) \, d\mathcal{H}^3(y) \frac{ds}{s} \lesssim_{\varepsilon,L,\theta} R^3$$

for any $\varepsilon > 0$, $x \in E$ and $R > 0$.

Theorem 4.1 will follow using standard arguments, recalled at the end of the paper, once we have at our disposal the WGL for intrinsic Lipschitz graphs. Therefore the rest of the paper will be devoted to the proof of Theorem 1.1, which is precisely stated below.

Theorem 4.2. *Let Γ be an intrinsic L-Lipschitz graph. Then*

$$\int_0^R \int_{\Gamma \cap B(x,R)} \chi_{\{(y,s) \in \Gamma \times \mathbb{R}_+ \colon \beta(B(y,s)) \geq \varepsilon\}}(y,s) \, d\mathcal{H}^3(y) \frac{ds}{s} \lesssim_{\varepsilon,L} R^3$$

for any $\varepsilon > 0$, $x \in \Gamma$ and $R > 0$.

The proof of Theorem 4.2 in the Euclidean case is relatively straightforward, and can be carried out as follows. Assume that $f \colon \mathbb{R}^2 \to \mathbb{R}$ is an (entire) L-Lipschitz function, and let Δ be a system of dyadic squares on $\mathbb{R}^2$. Assume that $Q \in \Delta$ is such that f is far from affine in Q, in the sense that $\sup_{x \in Q} |f(x) - A(x)| \geq \varepsilon\ell(Q)$ for all affine functions $A \colon \mathbb{R}^2 \to \mathbb{R}$. Then, it is fairly easy to verify that the gradient of f must fluctuate significantly near Q: there exists a fairly large sub-cube $Q' \subset Q$ such that $|\mathbb{E}_Q(\nabla f) - \mathbb{E}_{Q'}(\nabla f)| \geq \delta$, where δ only depends on ε and L, and $\mathbb{E}_Q \nabla f$ denotes the average of ∇f over Q. The WGL follows from this observation, plus the fact that $\|\nabla f\|_{L^2(Q_0)} \lesssim_L |Q_0|$ for any fixed cube $Q_0 \in \Delta$ (for more details on this final step, see the argument after (4.37)).

It is a reasonable first thought that a similar argument should work for intrinsic Lipschitz functions $\phi \colon \mathbb{W} \to \mathbb{V}$. After all, for such a function, there is a concept of an *intrinsic gradient* $\nabla^{\phi}\phi$ (see Section 4.2 below), which is known to exist at almost every point on $\mathbb{W}$, and moreover $\nabla^{\phi}\phi \in L^{\infty}(\mathbb{W})$. So, if it were the case that the local "non-affinity" of ϕ forces $\nabla^{\phi}\phi$ to fluctuate noticeably, one could wrap up the argument in the fashion outlined above. However, this is simply not true: in any bounded domain $\Omega \subset \mathbb{W}$, the equation $\nabla^{\phi}\phi = 0$ admits (smooth) non-affine solutions! For instance, the function

$$\phi : (-1, +\infty) \times \mathbb{R} \to \mathbb{R}, \quad \phi(y,t) = \tfrac{t}{y+1}$$

satisfies $\nabla^{\phi}\phi \equiv 0$ on its domain. In fact, even non-smooth continuous solutions are possible: the function $\phi \colon (-1,1)^2 \to \mathbb{R}$, discussed in [7, Remark 4.4.2] and defined by

$$\phi(y,t) := \begin{cases} \frac{t}{y+1}, & t \geq 0, \\ \frac{t}{y-1}, & t < 0, \end{cases} \tag{4.1}$$

satisfies $\nabla^{\phi}\phi = 0$ on $(-1,1)^2$, but it is not $\mathcal{C}^1$. The intrinsic graph of ϕ over the $(-1,1)^2$ is depicted in Figure 1. Further examples of similar flavour have been

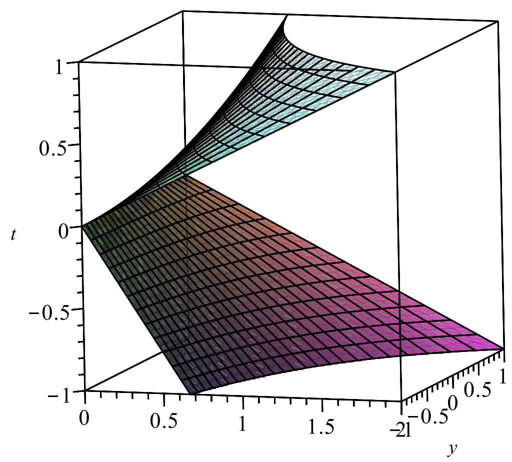

FIGURE 1. The intrinsic graph of the function in (4.1) over the (y,t)-plane.

discussed in [1] in connection with minimal surfaces in the Heisenberg group. We emphasise that the graphs of non-affine solutions to $\nabla^{\phi}\phi = 0$ have plenty of non-zero β-numbers, but this behaviour is not registered by the fluctuation of $\nabla^{\phi}\phi$.

What can be done? We still want to use the intrinsic gradient, so we need to invent a condition to replace "non-affinity", which forces $\nabla^{\phi}\phi$ to fluctuate locally. It turns out that the right notion is "being far from constant-gradient (CG) graphs". The following (informal) statement may sound almost tautological, but it requires a fair amount of work to verify: if ϕ deviates locally from all (locally defined) CG intrinsic Lipschitz functions, then $\nabla^{\phi}\phi$ must fluctuate noticeably. Following the Euclidean idea, this observation (made precise in Proposition 4.20) allows us to conclude that intrinsic Lipschitz graphs satisfy a "WGL for CG β-numbers", see Theorem 4.16.

Up to this point, the results can be accused of being abstract nonsense; as far as we know, locally defined CG graphs could be quite wild, and we do not even claim to understand them very well. What we can understand, however, are **globally defined** CG graphs. These turn out to be affine (see Proposition 4.8)! Using this fact, and a compactness argument, we can prove that even locally defined CG graphs have the following key property: if Γ is a CG graph "defined in the whole ball $B(x,r)$", then Γ is almost flat in all sufficiently small sub-balls of $B(x,r)$. This implies almost immediately that the usual WGL holds for CG-graphs, even if we do not record the argument separately. Instead, we use our "approximation by CG graphs" result to conclude directly that the WGL holds for all intrinsic Lipschitz graphs.

We wish to mention that the proof strategy above was influenced by X. Tolsa's proof [44] of the fact that the *weak constant density condition* implies *uniform rectifiability* in $\mathbb{R}^n$. Should the reader be familiar with that proof, she may wish to draw the following parallels in her mind: "weak constant density" is represented by "intrinsic Lipschitz", and "uniform measure" is represented by "constant gradient graph".

4.2. **The intrinsic gradient.** Our proof of the WGL for intrinsic Lipschitz graphs is based on the notion of *intrinsic gradient*. The present section serves the purpose of reviewing the definition and relevant properties.

4.2.1. *Definitions.* According to a well known theorem by Rademacher, Lipschitz maps between Euclidean spaces are differentiable almost everywhere. The same result, appropriately interpreted, holds true for the intrinsic Lipschitz functions appearing in this paper. Similarly as in the Euclidean setting, a function $\phi : \mathbb{W} \to \mathbb{V}$ shall be differentiable at a point $w_0 \in \mathbb{W}$, if its graph at $p_0 = w_0 \cdot \phi(w_0)$ can be well approximated by the graph of a "linear" function $L : \mathbb{W} \to \mathbb{V}$.

Following the terminology in [28], a function $L : \mathbb{W} \to \mathbb{V}$ between complementary homogeneous subgroups is said to be *intrinsic linear*, if its intrinsic graph $\{w \cdot L(w) : \ w \in \mathbb{W}\}$ is a homogeneous subgroup of $\mathbb{H}$. If $\mathbb{W}$ is a vertical subgroup and $\mathbb{V}$ is a complementary orthogonal horizontal subgroup in $\mathbb{H}$, a map $L : \mathbb{W} \to \mathbb{V}$ is intrinsic linear if and only if it is a *homogeneous homomorphism* (see for instance Proposition 3.26 in [5]). The latter means that L is a group homomorphism with the additional property that $L(\delta_r(w)) = rL(w)$ for all $w \in \mathbb{W}$ and

$r > 0$. To give an example, if $\mathbb{W}$ denotes the (y,t)-plane, it is not difficult to see that all intrinsic linear maps, or equivalently all homogeneous homomorphisms, are of the form $L(y,t) = cy$ for a constant $c \in \mathbb{R}$.

We are now ready to state the definition of *intrinsic differentiability*. Again we assume that $\mathbb{W}$ is a vertical subgroup with complementary horizontal subgroup $\mathbb{V}$. First let us consider a function $\phi : \Omega \to \mathbb{V}$ defined on an open subset of $\Omega \subset \mathbb{W}$ containing the origin, which we assume to be fixed under ϕ. We say that ϕ is intrinsic differentiable at the origin, if there exists an intrinsic linear map $L : \mathbb{W} \to \mathbb{V}$ such that, for all $w \in \mathbb{W}$,

$$|\phi(w) - L(w)| = o(\|w\|), \quad \text{as } \|w\| \to 0.$$

The map L is called the *intrinsic differential* of ϕ at 0 and it is denoted by $L = D_0\phi$.

Since the definition of differentiability is supposed to be *intrinsic*, we extend it in a left invariant fashion to arbitrary functions and points. To explain the definition, we consider a function $\phi : \Omega \to \mathbb{V}$ on an open set $\Omega \subset \mathbb{W}$ with intrinsic graph Γ. For $w_0 \in \Omega$, we write $p_0 = w_0 \cdot \phi(w_0)$ and let $\phi_{p_0^{-1}}$ be the uniquely defined function $\mathbb{W} \to \mathbb{V}$, which parametrizes the graph $p_0^{-1} \cdot \Gamma$, see Remark 2.5. Note that the definition of intrinsic graph ensures that any left translate of Γ by a point in Γ is an intrinsic graph passing through the origin. This uniquely determines the function $\phi_{p_0^{-1}}$ and it ensures that the origin is fixed under this function. We also formulate an equivalent definition of intrinsic Lipschitz functions using the maps ϕ_p. If $A \subset \mathbb{W}$, a function $\phi : A \to \mathbb{V}$ with intrinsic graph Γ is an intrinsic L-Lipschitz function if and only if for every $p \in \Gamma$

$$\|\phi_{p^{-1}}(w)\| \leq L\|w\| \tag{4.2}$$

for all w in the domain of $\phi_{p^{-1}}$. See also [42, Proposition 4.49] for other equivalent algebraic definitions for intrinsic Lipschitz functions. An explicit formula for $\phi_{p_0^{-1}}$ is given in Lemma 4.7 below.

Definition 4.3. We say that a function $\phi : \Omega \to \mathbb{V}$, defined on an open set $\Omega \subset \mathbb{W}$, is *intrinsic differentiable* at a point w_0, if the function $\phi_{p_0^{-1}}$ for $p_0 = w_0 \cdot \phi(w_0)$ is intrinsic differentiable at the origin. The *intrinsic differential* of ϕ at w_0 is given by $D_{w_0}\phi = D_0\phi_{p_0^{-1}}$.

Intrinsic differentiability can be characterized in various equivalent ways, see for instance the results in [27], [28], and [5]. One can also define intrinsic differentiability in terms of a "graph distance", see Definition 1.4 in [14], and for intrinsic Lipschitz maps this definition is equivalent to the one above.

The intrinsic differential is unique, and its action can be expressed in terms of a "gradient" similarly as in the Euclidean case. To explain this, we identify $\mathbb{W} \cong \mathbb{R}^2$ and $\mathbb{V} \cong \mathbb{R}$. Points in $\mathbb{W}$ are then denoted by (y,t). Assume that a map $\phi\colon \mathbb{W} \to \mathbb{V}$ is intrinsic differentiable at $w_0 = (y_0, t_0)$. Then, as pointed out earlier, the intrinsic differential $L := D_{w_0}\phi$ of ϕ at w_0 is a linear mapping of the form $L(y,t) = cy$ for some $c \in \mathbb{R}$.

Definition 4.4. The number c is called the *intrinsic gradient of ϕ at w_0*, and it will be denoted by

$$\nabla^{\phi}\phi(w_0) := c.$$

We next derive a few useful formulae for $\nabla^{\phi}\phi(w_0)$. Let $p_0 := w_0 \cdot \phi(w_0)$ be the point on the graph of ϕ. By definition of intrinsic differentiability,

$$|\phi_{p_0^{-1}}(h,0) - \nabla^{\phi}\phi(w_0)h| = |\phi_{p_0^{-1}}(h,0) - L(h,0)| = o(|h|), \quad h \in \mathbb{R} \setminus \{0\}.$$

Consequently, dividing by h, we have

$$\nabla^{\phi}\phi(w_0) = \lim_{h\to 0} \frac{\phi_{p_0^{-1}}(h,0)}{h}. \tag{4.3}$$

Moreover, wherever ϕ is differentiable in the usual (Euclidean) sense, the formula above, and

$$\lim_{h\to 0} \frac{\phi_{p_0^{-1}}(h,0)}{h} = \lim_{h\to 0} \frac{\phi(y_0+h, t_0+\phi(y_0,t_0)h) - \phi(y_0,t_0)}{h}$$

(this follows from the formula for $\phi_{p_0^{-1}}$ in Lemma 4.7) yields the following representation for $\nabla^{\phi}\phi$:

$$\nabla^{\phi}\phi = \partial_y\phi + \phi\partial_t\phi. \tag{4.4}$$

This was observed in Example 5.5 in [5].

A large class of almost everywhere intrinsic differentiable functions is provided by intrinsic Lipschitz functions whose target is a 1-dimensional horizontal subgroup. This result was first proved by Franchi, Serapioni and Serra Cassano in [28, Theorem 4.29] for Heisenberg groups, and later by Franchi, Marchi and Serapioni in [29] for certain more general Carnot groups. We state here the result relevant for the current paper:

Theorem 4.5 (Franchi, Serapioni, Serra Cassano)**.** *Let $\mathbb{W}$ be a vertical subgroup of $\mathbb{H}$ with complementary horizontal subgroup $\mathbb{V}$. Assume that Ω is an open subset of $\mathbb{W}$ and $\phi : \Omega \to \mathbb{V}$ is intrinsic Lipschitz. Then ϕ is intrinsic differentiable $\mathcal{L}^2$ almost everywhere in $\mathbb{W}$.*

It follows that the intrinsic gradient $\nabla^{\phi}\phi$ of an intrinsic Lipschitz function $\phi : \Omega \to \mathbb{V}$ exists almost everywhere in Ω, and it is an L^{∞} function, see for instance Proposition 4.4 in [14]. In our situation, one can say something more precise:

Lemma 4.6. *Let $\mathbb{W}$ be a vertical subgroup with complementary horizontal subgroup $\mathbb{V}$ in $\mathbb{H}$, and let $\phi : \Omega \to \mathbb{V}$ be an intrinsic L-Lipschitz function on an open set $\Omega \subset \mathbb{W}$. Then*

$$\|\nabla^{\phi}\phi\|_{L^{\infty}(\Omega)} \leq L. \tag{4.5}$$

Such a result was stated for *difference quotients* in [27, Proposition 3.9 (i)]. For the convenience of the reader, we spell out the argument for the intrinsic gradient.

Proof. By Theorem 4.5, ϕ is intrinsic differentiable in almost every point of Ω. We let w_0 be such a point, and we write $p_0 = w_0 \cdot \phi(w_0)$ for the corresponding point on the graph. Recall that the function $\phi_{p_0{}^{-1}}$ in (4.3) is defined as the function whose graph is the left translate of the graph of ϕ by p_0^{-1}. We denote the domain of $\phi_{p_0{}^{-1}}$ by $\Omega_{p_0^{-1}}$.

In order to prove (4.5), it suffices to find a bound for the limit in (4.3). To this end, let h be small enough such that $(h, 0) \in \Omega_{p_0^{-1}}$. Then by (4.2)

$$\left| \frac{\phi_{p_0{}^{-1}}(h, 0))}{h} \right| \leq h^{-1} L \|(h, 0)\| = L.$$

Here we have used (4.2). Thus $|\nabla^\phi \phi(w_0)| \leq L$. Since w_0 was an arbitrary point of intrinsic differentiability, (4.5) follows by Theorem 4.5. □

In the converse direction, Proposition 1.8 in [14] provides local upper bounds for the Lipschitz constant in terms of the L^∞-norm of the intrinsic gradient. While these results connect $\nabla^\phi \phi$ to the geometry of intrinsic Lipschitz graphs, the intrinsic gradient has a life of its own outside the world of intrinsic Lipschitz maps. To see this, it is best to express $\nabla^\phi \phi$ as in (4.4). The equation

$$\partial_y \phi + \phi \partial_t \phi = 0 \tag{4.6}$$

is well known in PDE theory as the *inviscid Burgers equation*. This will be discussed further in Section 4.2.3 below.

In which sense are intrinsic Lipschitz functions solutions to an equation of Burgers' type? By Theorem 4.5, the intrinsic gradient of an intrinsic Lipschitz function exists pointwise almost everywhere. In connection with PDE theory, it is useful to know that the pointwise intrinsic gradient of an intrinsic Lipschitz function is also a *distributional* gradient. This is the content of Proposition 4.7 in [14]. Precisely, if $\phi : \mathbb{W} \to \mathbb{V}$ is intrinsic Lipschitz with intrinsic gradient $\nabla^\phi \phi$, defined almost everywhere by Theorem 4.5, then (4.4) holds in a distributional sense:

$$\int_{\mathbb{W}} \phi \partial_y \psi + \tfrac{1}{2} \phi^2 \partial_t \psi \, d\mathcal{L}^2 = - \int_{\mathbb{W}} \nabla^\phi \phi \psi \, d\mathcal{L}^2 \tag{4.7}$$

for all $\psi \in \mathcal{C}_c^1(\mathbb{W})$.

We emphasize that the intrinsic Lipschitz functions as in Definition 2.4 coincide with the intrinsic Lipschitz functions in the sense of [14], see Theorem 4.60 in [42] and Remark 3.6 in [30]. Note further that the formula (4.7) looks slightly different from [14] due to a different model for the Heisenberg group, see Definition 3.1 in [9].

4.2.2. *Translated and dilated graphs.* While the definition of intrinsic Lipschitz continuity is tailored so that the class of intrinsic L-Lipschitz graphs is preserved under dilations and translations in the Heisenberg group, the explicit formula for the parametrization of a translated graph becomes in general slightly complicated due to the non-commutativity of the group law. In the case we consider in the present paper: functions from vertical to horizontal subgroups in $\mathbb{H}$, the computations are straightforward.

For convenience, given a point $p \in \mathbb{H}$, we define the map

$$P_p : \mathbb{W} \to \mathbb{W}, \quad P_p(w) := \pi_{\mathbb{W}}(p \cdot w).$$

We note that P_q is a diffeomorphism with Jacobian determinant constant equal to 1 under the obvious identification of $\mathbb{W}$ with $\mathbb{R}^2$ (see [30, Lemma 2.20]), and with inverse map $(P_p)^{-1} = P_{p^{-1}}$. The latter claim follows from the fact that

$$P_p(w) = p \cdot w \cdot \pi_{\mathbb{V}}(p)^{-1}, \quad \text{for all } w \in \mathbb{W}.$$

Lemma 4.7. *Let Γ be the intrinsic graph of a function $\phi : \mathbb{W} \to \mathbb{V}$ on a vertical subgroup $\mathbb{W}$, and let Ω be a domain in $\mathbb{H}$. Then, for $q \in \mathbb{H}$, the set $\tau_q(\Omega \cap \Gamma)$ is the intrinsic graph of the function*

$$\phi_q : \pi_{\mathbb{W}}(\tau_q(\Omega \cap \Gamma)) \to \mathbb{V}, \quad \phi_q(w) := \pi_{\mathbb{V}}(q) \cdot \phi(P_{q^{-1}}(w)).$$

If ϕ is intrinsic L-Lipschitz, then so is ϕ_q with

$$\nabla^{\phi_q}\phi_q = \nabla^{\phi}\phi \circ P_{q^{-1}}, \quad \textit{almost everywhere}. \tag{4.8}$$

Analogously, for $r > 0$, the set $\delta_r(\Omega \cap \Gamma)$ is the intrinsic graph of the function

$$\phi_r : \pi_{\mathbb{W}}(\delta_r(\Omega \cap \Gamma)) \to \mathbb{V}, \quad \phi_r(w) = \delta_r(\phi(\delta_{1/r}(w))).$$

If ϕ is intrinsic L-Lipschitz, then so is ϕ_r with

$$\nabla^{\phi_r}\phi_r = \nabla^{\phi}\phi \circ \delta_{r^{-1}}, \quad \textit{almost everywhere}.$$

Directly from the definition of intrinsic differential and intrinsic gradient, it follows that

$$\nabla^{\phi_q}\phi_q(0) = \nabla^{\phi}\phi(\pi_{\mathbb{W}}(q^{-1})).$$

In (4.8), we show how $\nabla^{\phi_q}\phi_q$ and $\nabla^{\phi}\phi$ are related in a generic point of $\mathbb{W}$.

Proof. We concentrate on proving the statement for left translations. In [27, Proposition 2.7] it has been shown that if Γ is the intrinsic graph of a function ϕ over a domain in $\mathbb{V}$, then $\tau_q(\Gamma)$ is an intrinsic graph parametrized by the function ϕ_q. The domain of the new function ϕ_q is simply the image of $\tau_q(\Gamma)$ under the projection onto $\mathbb{W}$. The intrinsic Lipschitz property of ϕ_q – assuming the corresponding property for ϕ – is the content of [27, Theorem 3.2].

To compute the intrinsic gradient, we may assume without loss of generality that $\mathbb{W}$ agrees with the (y,t)-plane and that ϕ is defined on the entire plane $\mathbb{W}$. In this case, for $q = (x_0, y_0, t_0)$, we have that

$$\phi_q(y,t) = x_0 + \phi(P_{q^{-1}}(y,t)).$$

We then use the fact, proved in [14], that the intrinsic gradient of an intrinsic Lipschitz function is also a distributional gradient; see the discussion in Section 4.2. Let now ψ be an arbitrary test function, that is, a compactly supported $\mathcal{C}^1$ function on $\mathbb{W}$. Since

$$DP_q = \begin{pmatrix} 1 & 0 \\ x_0 & 1 \end{pmatrix}, \tag{4.9}$$

we find that

$$\frac{\partial \psi \circ P_q}{\partial y} = \frac{\partial \psi}{\partial y} \circ P_q + x_0 \frac{\partial \psi}{\partial t} \circ P_q \quad \text{and} \quad \frac{\partial \psi \circ P_q}{\partial t} = \frac{\partial \psi}{\partial t} \circ P_q.$$

This, together with the facts that $\det DP_q = 1$ and $\nabla^{\phi}\phi$ is a distributional gradient,recall (4.7), gives

$$\begin{aligned}
\int \nabla^{\phi_q}\phi_q \psi \, d\mathcal{L}^2 &= -\int \phi_q \tfrac{\partial \psi}{\partial y} + \tfrac{1}{2}\phi_q^2 \tfrac{\partial \psi}{\partial t} \, d\mathcal{L}^2 \\
&= -\int (\phi \circ P_{q^{-1}}) \tfrac{\partial \psi}{\partial y} + x_0 (\phi \circ P_{q^{-1}}) \tfrac{\partial \psi}{\partial t} + \tfrac{1}{2}(\phi^2 \circ P_{q^{-1}}) \tfrac{\partial \psi}{\partial t} \, d\mathcal{L}^2 \\
&= -\int \phi (\tfrac{\partial \psi}{\partial y} \circ P_q) + x_0 \phi (\tfrac{\partial \psi}{\partial t} \circ P_q) + \tfrac{1}{2}\phi^2 (\tfrac{\partial \psi}{\partial t} \circ P_q) \, d\mathcal{L}^2 \\
&= -\int \phi \tfrac{\partial \psi \circ P_q}{\partial y} + \tfrac{1}{2}\phi^2 \tfrac{\partial \psi \circ P_q}{\partial t} \, d\mathcal{L}^2 \\
&= \int \nabla^{\phi}\phi \, (\psi \circ P_q) \, d\mathcal{L}^2 \\
&= \int (\nabla^{\phi}\phi) \circ P_{q^{-1}} \, \psi \, d\mathcal{L}^2.
\end{aligned}$$

As this computation is valid for arbitrary test functions ψ, the claim (4.8) follows. □

4.2.3. *Graphs with constant gradient.* In this subsection, we prove that "entire" intrinsic Lipschitz functions with almost surely constant gradient are affine. As mentioned in Section 4.2, if $\mathbb{W}$ is identified with $\mathbb{R}^2$, the differential equation $\nabla^{\phi}\phi = 0$ is known as the inviscid Burgers equation and it is not difficult to see by the *method of characteristics* (see [23, Proposition 5.1]) that the only global $\mathcal{C}^1$ solutions are constant functions. If the right-hand side of the equation is replaced by some other constant c, one can show in the same vein that the only $\mathcal{C}^1$ solutions are affine functions of the form $\phi(y,t) = cy + d$; see [1, Remark 4.3].

Our task is to establish the same result for functions ϕ that are merely assumed to be intrinsic Lipschitz with intrinsic gradient constant almost everywhere.

Proposition 4.8. Let $\phi : \mathbb{W} \to \mathbb{V}$ be an intrinsic Lipschitz function. If there exists a constant $c \in \mathbb{R}$ such that $\nabla^{\phi}\phi = c$ almost everywhere in $\mathbb{W}$, then the graph of ϕ is the left translate of some vertical plane $\mathbb{W}' = \mathbb{W}'_{\mathbb{W},c}$.

Remark 4.9. We thank Enrico Le Donne and the anonymous referee for pointing out that Proposition 4.8 also follows directly from existing results in the literature. Indeed, by [28, Theorem 4.17], the subgraph Γ^{ϕ}_{-} of an intrinsic Lipschitz graph Γ^{ϕ} over a vertical plane is a set with locally finite $\mathbb{H}$-perimeter. If the intrinsic gradient $\nabla^{\phi}\phi$ is constant, then the horizontal normal to the boundary of Γ^{ϕ}_{-} is constant. According to the proof of [26, Claim 3] (see also [4, Proposition 5.4]), this implies that Γ^{ϕ}_{-} is the left translate of a vertical halfspace, and hence Γ^{ϕ} is the left translate of a vertical plane.

Proof of Proposition 4.8. Throughout the proof, we identify $\mathbb{W}$ with $\mathbb{R}^2$, using coordinates (y,t). We will prove that $\phi(y,t) = cy + d$ for some $d \in \mathbb{R}$. Since ϕ is intrinsic Lipschitz, it is continuous, see for instance [27, Proposition 3.4]. We start by observing that the almost sure constancy of the intrinsic gradient leads to improved regularity for ϕ. By [14, Proposition 4.7], the function ϕ is a distributional solution to the equation $\nabla^{\phi}\phi = g$ for $g(y,t) \equiv c$. Since g is constant, it is in particular Lipschitz continuous in the Euclidean sense and [8, Corollary 1.4] implies that ϕ is locally Lipschitz on $\mathbb{W}$ with respect to the Euclidean metric; see also Theorem 4.2.1 and Theorem 4.4.1 in [7]. Hence, almost every point w of $\mathbb{W}$ is "good" in the sense that the function ϕ is differentiable at w both in the usual Euclidean sense and in the intrinsic sense with $\nabla^{\phi}\phi(w) = c$. We denote by G the set of such good points in $\mathbb{W}$, so that $\mathcal{L}^2(\mathbb{W} \setminus G) = 0$.

For every $t \in \mathbb{R}$, we define a curve $\gamma_t : \mathbb{R} \to \mathbb{W}$, by setting

$$\gamma_t(s) := \left(s, \tfrac{c}{2}s^2 + \phi(0,t)s + t\right).$$

We will prove for almost every $t \in \mathbb{R}$ that

$$\phi(\gamma_t(s)) = cs + \phi(0,t) \quad \text{for all } s \in \mathbb{R}. \tag{4.10}$$

Let us assume for a moment that $t \in \mathbb{R}$ is such that $\gamma_t(s) \in G$ for almost every $s \in \mathbb{R}$. Towards a proof of (4.10), we define the function

$$z : \mathbb{R} \to \mathbb{R}, \quad z(s) := \phi(\gamma_t(s)) - (cs + \phi(0,t)).$$

We note that z is locally Lipschitz continuous, so $z'(s)$ exists for almost all s. Further, by the assumption that $\gamma_t(s) \in G$ for almost every s, we have for such points that

$$\begin{aligned} z'(s) &= \partial_y\phi(\gamma_t(s)) + (cs + \phi(0,t))\partial_t\phi(\gamma_t(s)) - c \\ &= c - \phi(\gamma_t(s))\partial_t\phi(\gamma_t(s)) + (cs + \phi(0,t))\partial_t\phi(\gamma_t(s)) - c \\ &= -\left[\phi(\gamma_t(s)) - (cs + \phi(0,t))\right]\partial_t\phi(\gamma_t(s)) \\ &= -z(s)\partial_t\phi(\gamma_t(s)). \end{aligned}$$

Here we have used that $\nabla^{\phi}\phi = c = \partial_y\phi + \phi\partial_t\phi$ on G, see (4.4) and the subsequent discussion.

Thus z solves an ODE of the form

$$\begin{cases} z'(s) = a(s)z(s), & \text{almost everywhere,} \\ z(0) = 0. \end{cases} \tag{4.11}$$

Clearly, $z \equiv 0$ is a solution, but we have to argue that it is the *only* solution. Here we are interested in *Carathéodory solutions* $z : \mathbb{R} \to \mathbb{R}$, that is, in functions which are absolutely continuous on every closed interval $[\alpha,\beta] \subset \mathbb{R}$ and which fullfill the differential equation pointwise almost everywhere; see for instance [25, Chapter 1] for a thorough discussion of Carathéodory differential equations.

Since ϕ is locally Lipschitz as a function on the Euclidean plane and γ_t is a smooth curve with $\gamma_t(s) \in G$ for almost every $s \in \mathbb{R}$, the function

$$a : \mathbb{R} \to \mathbb{R}, \quad a(s) := \begin{cases} -\partial_t \phi(\gamma_t(s)), & \gamma_t(s) \in G, \\ 0, & \text{else}, \end{cases}$$

is locally integrable on every interval $[\alpha, \beta] \subset \mathbb{R}$. By Theorem 3 in [25, Chapter 1], this suffices to ensure that the ODE (4.11) has a unique Carathéodory solution on $\mathbb{R}$. Hence, $z \equiv 0$ and (4.10) follows for this particular choice of t.

Next, we would like to show that almost every $t \in \mathbb{R}$ has the crucial property that $\gamma_t(s) \in G$ for almost every s. This is the content of Lemma 4.12 below. The statement would be immediate if we knew that the curves γ_t foliated the plane $\mathbb{W}$, or even a large portion thereof, but there is no such *a priori* information available. In fact, this foliation property is part of the statement we want to prove.

So, we have to work a bit harder, and we are essentially rescued by the local Lipschitz regularity of $t \mapsto \phi(0, t)$. In the proof of Lemma 4.12 we need a sharpened version of the "easy implication" in the Besicovitch projection theorem. This result may be known to some experts, but we did not find it in the literature:

Lemma 4.10. *Let $K \subset \mathbb{R}^2$ be a rectifiable set with $0 < \mathcal{H}^1(K) < \infty$. Then, there exists a set of unit vectors $G \subset S^1$, depending only on K, with the following properties:*

(i) $\mathcal{H}^1(S^1 \setminus G) = 0$.
(ii) *If $F \subset K$ is any $\mathcal{H}^1$-measurable subset with $\mathcal{H}^1(F) > 0$ and $e \in G$, then $\mathcal{H}^1(\pi_e(F)) > 0$. Here π_e is the orthogonal projection $\pi_e(x) = x \cdot e$.*

Remark 4.11. The lemma immediately extends to rectifiable sets with σ-finite $\mathcal{H}^1$-measure.

Proof of Lemma 4.10. We start with a series of reductions. Without loss of generality we can assume that K is bounded. Moreover it is enough to prove the lemma for $K = f([-R, R])$ for $f : \mathbb{R} \to \mathbb{R}$ Lipschitz and $R > 0$. To see this, first recall that if K is a bounded rectifiable set there exist countably many Lipschitz maps $f_n : \mathbb{R} \to \mathbb{R}$ such that $\mathcal{H}^1(K \setminus \bigcup_n f_n([-R, R])) = 0$ for some $R > 0$. Applying the lemma to each of the sets $\gamma_n := f_n([-R, R])$ we obtain sets $G_n \subset S^1$ satisfying (i) and (ii). Now let $G = \bigcap_{n \in \mathbb{N}} G_n$. Trivially $\mathcal{H}^1(S^1 \setminus G) = 0$ and if $F \subset K$ is any $\mathcal{H}^1$-measurable subset with $\mathcal{H}^1(F) > 0$, then there exists some γ_n such that $\mathcal{H}^1(\gamma_n \cap F) > 0$. Since γ_n satisfies (ii), if $e \in G \subset G_n$ then $\mathcal{H}^1(\pi_e(F)) > 0$.

Fix $\varepsilon > 0$. It is enough to find a subset $G_\varepsilon \subset S^1$ such that $\mathcal{H}^1(S^1 \setminus G_\varepsilon) = 0$ and the following property holds: if $F \subset K$ is measurable with $\mathcal{H}^1(F) > \varepsilon$ and $e \in G_\varepsilon$, then $\mathcal{H}^1(\pi_e(F)) > 0$. Then, we can complete the proof by setting $G := \bigcap_j G_{1/j}$.

By [38, Theorem 7.4] there exists a compact $\mathcal{C}^1$-curve $\Gamma = \Gamma_{\varepsilon,K}$ such that $\mathcal{H}^1(K \setminus \Gamma) \leq \varepsilon/2$. Then, if $F \subset K$ is measurable with $\mathcal{H}^1(F) > \varepsilon$, we have $\mathcal{H}^1(F \cap \Gamma) > \varepsilon/2$. Thus, it actually suffices to construct G_ε so that the following holds: if $F \subset \Gamma$ is measurable with $\mathcal{H}^1(F) > \varepsilon/2$ and $e \in G_\varepsilon$, then $\mathcal{H}^1(\pi_e(F)) > 0$. One final reduction: for fixed $\delta > 0$, we construct a set G_ε^δ with the properties that (a)

$\mathcal{H}^1(S^1 \setminus G_\varepsilon^\delta) < \delta$, and (b) if $F \subset \Gamma$ is measurable with $\mathcal{H}^1(F) > \varepsilon/2$ and $e \in G_\varepsilon^\delta$, then $\mathcal{H}^1(\pi_e(F)) > 0$. This suffices, since $G_\varepsilon := \bigcup_j G_\varepsilon^{1/j}$ is then the set we are after.

To construct G_ε^δ, we fix a number $m = m_{\delta,\varepsilon} \in \mathbb{N}$, to be specified later, and cover S^1 by a a collection $\mathcal{J} := \{J_1, \ldots, J_m\}$ of disjoint arcs of length between $1/m$ and $10/m$. Next, for some $n \in \mathbb{N}$ depending on m, we partition Γ into short, connected sub-curves $\mathcal{F} := \{\Gamma_1, \ldots, \Gamma_n\}$ such that the following holds: for every fixed $\Gamma_j \in \mathcal{F}$, the restriction $\pi_e|_{\Gamma_j}$ is bi-Lipschitz for all $e \in S^1$, except possibly those e in the union of four arcs in $\mathcal{J}$ (depending only on Γ_j). Such a partition $\mathcal{F}$ exists, because Γ is compact and $\mathcal{C}^1$.

Consider a bi-partite graph with vertex set $\mathcal{F} \cup \mathcal{J}$ and the following edge set E: draw an edge between $\Gamma_j \in \mathcal{F}$ and $J_k \in \mathcal{J}$, if and only if $\pi_e|_{\Gamma_j}$ is bi-Lipschitz for **all** $e \in J_k$. Thus, every vertex $\mathcal{F}$ is adjacent to at least $(m-4)$ vertices in $\mathcal{J}$. For an edge $(\Gamma_j, J_k) \in E$, define the weight

$$w(\Gamma_j, J_k) := \mathcal{H}^1(\Gamma_j).$$

Thus, if $w(E)$ is the sum of all the weights of edges in E, we have

$$w(E) = \sum_j \sum_{k:(\Gamma_j, J_k) \in E} w(\Gamma_j, J_k) \geq \sum_j (m-4)\mathcal{H}^1(\Gamma_j) = (m-4)\mathcal{H}^1(\Gamma).$$

Now, write

$$\tau := \min\{\delta/20, \varepsilon/(2\mathcal{H}^1(\Gamma))\},$$

and call a vertex $J_k \in \mathcal{J}$ *light*, if the total weight of edges emanating from J_k is at most $(1-\tau)\mathcal{H}^1(\Gamma)$. Other vertices in $\mathcal{J}$ are *heavy*. Denoting the light and heavy vertices in $\mathcal{J}$ by $\mathcal{J}_{light}$ and $\mathcal{J}_{heavy}$, respectively, we have

$$\begin{aligned}(m-4)\mathcal{H}^1(\Gamma) \leq w(E) &\leq (1-\tau)\mathcal{H}^1(\Gamma)|\mathcal{J}_{light}| + \mathcal{H}^1(\Gamma)|\mathcal{J}_{heavy}| \\ &= m\mathcal{H}^1(\Gamma) - \tau\mathcal{H}^1(\Gamma)|\mathcal{J}_{light}|,\end{aligned}$$

which simplifies to $|\mathcal{J}_{light}| \leq 4/\tau$. We now fix m so large that $m \geq 4/\tau^2$, which gives $|\mathcal{J}_{light}| \leq \tau m$. Then, let

$$G_\varepsilon^\delta := \bigcup_{J_k \in \mathcal{J}_{heavy}} J_k.$$

The set G_ε^δ satisfies the correct length estimate:

$$\mathcal{H}^1\left(S^1 \setminus G_\varepsilon^\delta\right) \leq \sum_{J_k \in \mathcal{J}_{light}} \mathcal{H}^1(J_k) \leq \frac{10\tau m}{m} < \delta$$

by the choice of τ.

Finally, we want to show that $\mathcal{H}^1(\pi_e(F)) > 0$, whenever $F \subset \Gamma$ is measurable with $\mathcal{H}^1(F) > \varepsilon/2$, and $e \in G_\varepsilon^\delta$. So, fix $F \subset \Gamma$ with $\mathcal{H}^1(F) > \varepsilon/2$, and write $\mathcal{F}_F := \{\Gamma_j \in \mathcal{F} : \mathcal{H}^1(F \cap \Gamma_j) > 0\}$. Then

$$\varepsilon/2 < \mathcal{H}^1(F) \leq \sum_{\Gamma_j \in \mathcal{F}_F} \mathcal{H}^1(\Gamma_j),$$

Then, fix $e \in G_{\varepsilon}^{\delta}$, so that $e \in J_k$ for some $\mathcal{J}_{heavy}$. This implies that J_k is adjacent to at least one vertex $\Gamma_j \in \mathcal{F}_F$; otherwise, recalling that $\tau < \varepsilon/(2\mathcal{H}^1(\Gamma))$, we have

$$\sum_{j:(\Gamma_j,J_k)\in E} w(\Gamma_j, J_k) \leq \sum_{j:\Gamma_j\notin\mathcal{F}_F} w(\Gamma_j, J_k) \leq \mathcal{H}^1(\Gamma) - \varepsilon/2 < (1-\tau)\mathcal{H}^1(\Gamma)$$

which contradicts $J_k \in \mathcal{J}_{heavy}$. Now, pick $\Gamma_j \in \mathcal{F}_F$ such that $(\Gamma_j, J_k) \in E$. By definition of E, this means that $\pi_e|_{\Gamma_j}$ is bi-Lipschitz, and consequently

$$\mathcal{H}^1(\pi_e(F)) \geq \mathcal{H}^1(\pi_e(F \cap \Gamma_j)) > 0.$$

The proof is complete. $\square$

We are ready to prove that the curves γ_t mostly avoid the set $\mathbb{W} \setminus G$:

Lemma 4.12. *Let $B \subset \mathbb{W}$ be a set with $\mathcal{L}^2(B) = 0$. Then, for almost every t, we have $\gamma_t(s) \in \mathbb{W} \setminus B$ for almost every s.*

Proof. With $B_s := \{t : (s,t) \in B\}$, we may re-write the claim as follows:

$$\begin{aligned} 0 = \mathcal{L}^2(\{(s,t) : \gamma_t(s) \in B\}) &= \mathcal{L}^2(\{(s,t) : (s, \tfrac{c}{2}s^2 + \phi(0,t)s + t) \in B\}) \\ &= \mathcal{L}^2(\{(s,t) : \tfrac{c}{2}s^2 + \phi(0,t)s + t \in B_s\}). \end{aligned}$$

So, it suffices to show that for almost every $s \in \mathbb{R}$, we have $\mathcal{H}^1(E_s) = 0$, where

$$E_s := \{t : \tfrac{c}{2}s^2 + \phi(0,t)s + t \in B_s\}.$$

Assume that this claim is false: there exists a positive measure set S of parameters s such that $\mathcal{H}^1(E_s) > 0$. Observe that $\mathcal{H}^1(\Gamma_s) \geq \mathcal{H}^1(E_s) > 0$ for $s \in S$, where

$$\Gamma_s := \{(\phi(0,t), t) : t \in E_s\} \subset \{(\phi(0,t), t) : t \in \mathbb{R}\} =: \Gamma.$$

Next, write $\pi_s(y,t) := (y,t) \cdot (s,1)$ for $(y,t) \in \mathbb{W}$; then, up to scaling, π_s is the orthogonal projection onto the line spanned by $(s,1)$ in the (y,t)-plane. Recalling Remark 4.11 we can apply Lemma 4.10 to the Lipschitz graph Γ and obtain a set of parameters $G \subset \mathbb{R}$ with $\mathcal{H}^1(\mathbb{R} \setminus G) = 0$, with the property that $\mathcal{H}^1(\pi_s(\Gamma_s)) > 0$, whenever $s \in G$ and $\mathcal{H}^1(\Gamma_s) > 0$. In particular, $\mathcal{H}^1(\pi_s(\Gamma_s)) > 0$ for almost all $s \in S$. Observing that $\frac{c}{2}s^2 + \pi_s(\Gamma_s) \subset B_s$ for every s, this forces $\mathcal{H}^1(B_s) > 0$ for almost all $s \in S$, which contradicts $\mathcal{L}^2(B) = 0$. The proof of the lemma is complete. $\square$

We have now established that (4.10) holds for almost every $t \in \mathbb{R}$, and the rest of the proof of Proposition 4.8 is easy. First, notice that

$$\gamma_t(s) = \gamma_{t'}(s') \quad \text{if and only if} \quad (s = s' \text{ and } \phi(0,t)s + t = \phi(0,t')s + t').$$

Now, if $\phi(0,t) \neq \phi(0,t')$ for some t, t', then $\phi(0,t)s_0 + t = \phi(0,t')s_0 + t'$ for some $s_0 \in \mathbb{R}$. It follows that for such t, t', the curves γ_t and $\gamma_{t'}$ intersect at $\gamma_t(s_0) = \gamma_{t'}(s_0)$.

Recall that we aim to show that $\phi(y,t) = cy + d$ for some $d \in \mathbb{R}$. We first show that $t \mapsto \phi(0,t)$ is constant. Pick t and t' satisfying (4.10). If $\phi(0,t) \neq \phi(0,t')$, then by the discussion in the previous paragraph, $\gamma_t(s_0) = \gamma_{t'}(s_0)$ for some $s_0 \in \mathbb{R}$. Consequently,

$$cs_0 + \phi(0,t) = \phi(\gamma_t(s_0)) = \phi(\gamma_{t'}(s_0)) = cs_0 + \phi(0,t')$$

by (4.10), which contradicts $\phi(0,t) \neq \phi(0,t')$. So, $t \mapsto \phi(0,t)$ is constant, say d, on the set where (4.10) holds. Referring again to (4.10), we find that, for $\mathcal{L}^2$ almost all $(s,t) \in \mathbb{R} \times \mathbb{R}$, we have

$$\phi(\gamma_t(s)) = \phi(s, \tfrac{c}{2}s^2 + ds + t) = cs + d.$$

Since ϕ is continuous, this is in fact true for all pairs (s,t), and hence $\phi(y,t) = \phi(y, \frac{c}{2}y^2 + dy + (t - \frac{c}{2}y^2 - dy)) = cy + d$ for all $(y,t) \in \mathbb{R}^2$. The proof is complete. $\square$

4.3. **A weak geometric lemma for constant gradient β-numbers.** In this section, we start to implement the plan outlined in Section 4.1: we define a variant of β-numbers, the *constant gradient β-numbers*, and prove that intrinsic Lipschitz graphs satisfy a weak geometric lemma with respect to this new definition.

If Γ is an intrinsic Lipschitz graph over $\mathbb{W}$, the constant gradient β-number (with parameter L) of a ball $B(x,r)$ is designed to described how well $\Gamma \cap B(x,r)$ can be approximated by the graph of an intrinsic L-Lipschitz function whose gradient is constant almost everywhere in $\pi_{\mathbb{W}}(B(x, b_L r))$. Here b_L is a small constant, given by the following lemma, and $B(x,r)$ denotes a closed ball with radius r centred at x.

Lemma 4.13. *For every $L > 0$ there exists a constant b_L such that if $\phi : \mathbb{W} \to \mathbb{V}$ is an intrinsic L-Lipschitz function on a vertical subgroup $\mathbb{W}$, then*

$$\pi_{\mathbb{W}}(B(x, b_L r)) \subseteq \pi_{\mathbb{W}}(B(x,r) \cap \Gamma) \subseteq \pi_{\mathbb{W}}(B(x,r))$$

for all x on the graph Γ of ϕ and for all $r > 0$.

For a proof of this lemma, see (44) in [30]. We are now ready to state our definition of constant gradient β-number.

Definition 4.14. Let $\Gamma = \{w \cdot \phi(w) : w \in \mathbb{W}\}$ be an intrinsic graph, where $\phi \colon \mathbb{W} \to \mathbb{V}$ is an intrinsic Lipschitz function. Fix a point $x \in \Gamma$ and a radius $r > 0$. Then, for $L \geq 1$, define

$$\beta_{\mathrm{CG}}(B(x,r)) := \beta_{\mathrm{CG},\Gamma,L}(B(x,r)) := \inf_{\psi} \sup_{w \in \pi_{\mathbb{W}}(B(x,r) \cap \Gamma)} \frac{|\phi(w) - \psi(w)|}{r}.$$

The infimum is taken over all intrinsic L-Lipschitz functions $\psi : \mathbb{W} \to \mathbb{V}$ which have intrinsic gradient constant almost everywhere on the set $\pi_{\mathbb{W}}(B(x, b_L r))$. The class of such "admissible" functions ψ will be denoted by

$$\mathrm{Adm}(B(x,r)) := \mathrm{Adm}_{\mathrm{CG},L}(B(x,r)).$$

If the Lipschitz constant L is clear from the context, we omit the subscript L for the constant b_L. Note that $\|\nabla^{\phi}\phi\|_{\infty} \leq L$ for $\phi \in \mathrm{Adm}(B(x,r))$ by Lemma 4.6.

Remark 4.15. Observe that

$$|\phi(w) - \psi(w)| = \|\Psi(w)^{-1} \cdot \Phi(w)\| = d_{\mathbb{H}}(\Psi(w), \Phi(w))$$

for $w \in \pi_{\mathbb{W}}(B(x,r) \cap \Gamma)$, where Ψ and Φ are the graph mappings

$$\Psi(w) := w \cdot \psi(w) \quad \text{and} \quad \Phi(w) = w \cdot \phi(w) \in \Gamma.$$

Thus, if $\beta_{\mathrm{CG}}(B(x,r)) < \varepsilon$, there exists $\psi \in \mathrm{Adm}(B(x,r))$ with graph Γ^{ψ} such that

$$\sup_{y \in \Gamma \cap B(x,r)} \frac{\mathrm{dist}_{\mathbb{H}}(y, \Gamma^{\psi})}{r} \leq \varepsilon.$$

The aim of this section is to prove the following weak geometric lemma for the constant gradient β-numbers:

Theorem 4.16. *Let Γ be an intrinsic L-Lipschitz graph over a vertical subgroup. Then*

$$\int_0^R \int_{\Gamma \cap B(x,R)} \chi_{\{(y,s) \in \Gamma \times \mathbb{R}_+ :\, \beta_{\mathrm{CG}}(B(y,s)) \geq \varepsilon\}}(y,s)\, d\mathcal{H}^3(y)\, \frac{ds}{s} \lesssim_{\varepsilon} R^3$$

for any $\varepsilon > 0$, $x \in \Gamma$ and $R > 0$. Here $\beta_{\mathrm{CG}}(B(y,s)) := \beta_{\mathrm{CG},\Gamma,L}(B(y,s))$.

As explained in Section 4.1, a large β_{CG} number implies that $\nabla^{\phi}\phi$ fluctuates locally. More precisely, Proposition 4.20 will show that if $\beta_{\mathrm{CG}}(B(x,r)) \geq \varepsilon$ for some $x \in \Gamma$ and $r > 0$, then there exists another ball $B(y,s) \subset B(x,r)$ such $\mathrm{dist}_{\mathbb{H}}(y,\Gamma) \leq s/10$, $s \geq \delta_{\varepsilon,L} r$ and $|\mathbb{E}_{\pi_{\mathbb{W}}(B(x,r)\cap\Gamma)} \nabla^{\phi}\phi - \mathbb{E}_{\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)} \nabla^{\phi}\phi| \geq \delta_{\varepsilon,L}$. In Section 4.4, we use this to prove Theorem 4.16.

4.3.1. *Auxiliary results.* Before stating Proposition 4.20, we record a few lemmas. The first one gives an upper bound on how much $\mathbb{E}_{\pi_{\mathbb{W}}(B(x,r)\cap\Gamma)} \nabla^{\phi}\phi$ can change as a function of r. Here and in the following, we employ the notation

$$\mathbb{E}_A f = \frac{1}{\mathcal{L}^2(A)} \int_A f \, d\mathcal{L}^2$$

for the average of a function f over a set A in the plane.

Lemma 4.17. *Assume that Γ is an intrinsic Lipschitz graph defined over $\mathbb{W}$, and that $f \in L^{\infty}(\mathbb{W})$. Further, let $x \in \mathbb{H}$, and $0 < s_1 \leq s_2 < \infty$. Then,*

$$|\mathbb{E}_{\pi_{\mathbb{W}}(B(x,s_1)\cap\Gamma)} f - \mathbb{E}_{\pi_{\mathbb{W}}(B(x,s_2)\cap\Gamma)} f| \lesssim \frac{\mathcal{H}^3(A(x,s_1,s_2) \cap \Gamma)}{\mathcal{L}^2(\pi_{\mathbb{W}}(B(x,s_2) \cap \Gamma))} \cdot \|f\|_{\infty},$$

where $A(x,s_1,s_2)$ is the annulus $\{y \in \mathbb{H} : s_1 \leq d_{\mathbb{H}}(x,y) \leq s_2\}$.

Proof. Write $B_1 := \pi_{\mathbb{W}}(B(x,s_1) \cap \Gamma)$ and $B_2 := \pi_{\mathbb{W}}(B(x,s_2) \cap \Gamma)$. In this proof, let $|U| := \mathcal{L}^2(U)$ for $U \subset \mathbb{W}$. Then,

$$\begin{aligned}|\mathbb{E}_{B_1} f - \mathbb{E}_{B_2} f| &= \frac{1}{|B_1|} \left| \int_{B_1} f \, d\mathcal{L}^2 - \frac{|B_1|}{|B_2|} \int_{B_2} f \, d\mathcal{L}^2 \right| \\ &= \frac{1}{|B_1|} \left| \left(1 - \frac{|B_1|}{|B_2|}\right) \int_{B_1} f \, d\mathcal{L}^2 - \frac{|B_1|}{|B_2|} \int_{B_2 \setminus B_1} f \, d\mathcal{L}^2 \right| \\ &\leq \frac{|B_2 \setminus B_1|}{|B_2|} \cdot \|f\|_\infty + \frac{|B_2 \setminus B_1|}{|B_2|} \cdot \|f\|_\infty\end{aligned}$$

by the triangle inequality. Next, we observe that

$$B_2 \setminus B_1 = \pi_{\mathbb{W}}([B(x,s_2) \setminus B(x,s_1)] \cap \Gamma) \subset \pi_{\mathbb{W}}(A(x,s_1,s_2) \cap \Gamma)$$

by the injectivity of $\pi_{\mathbb{W}}$ restricted to Γ. Finally, we use the fact that $\mathcal{H}^3(\pi_{\mathbb{W}}(A)) \leq C\mathcal{H}^3(A)$, recall Lemma 3.6. $\square$

It is desirable to have quantitative control on the upper bound appearing in Lemma 4.17. This motivates the following definition:

Definition 4.18. Let (X,d,μ) be a metric measure space. A ball $B(x,r) \subset X$ has *A-thin boundary* (with respect to μ) if the following holds:

$$\mu\left(B(x,2r) \cap A(x,(1-\lambda)r,(1+\lambda)r)\right) \leq A\,\lambda\,\mu(B(x,2r)), \qquad \lambda > 0.$$

Here $A(x,s,t) := \{y : s \leq d(x,y) \leq t\}$.

Balls with thin boundary are abundant:

Lemma 4.19. *Let (X,d,μ) be metric measure space. For any $0 < \delta < 1/4$, there exists a constant $A = A_\delta < \infty$ with the following property: for any ball $B(x,r)$ with $x \in X$ and $r > 0$, there exists a radius $s \in [r,(1+\delta)r]$ such that $B(x,s)$ has A-thin boundary.*

Proof. For a fixed $x \in X$, let $\pi \colon X \to [0,\infty)$ be the mapping $\pi(y) := d(x,y)$, and consider the push-forward measure $\nu := \pi_\sharp[\mu|_{B(x,2r)}]$. Let M be the usual centred Hardy-Littlewood maximal operator on $\mathbb{R}$, namely

$$Mf(s) := \sup_{t>0} \frac{1}{2t} \int_{B(s,t)} |f(y)| \, dy.$$

We extend the definition from f to ν in the obvious way. It is well-known (see for instance [38, Theorem 2.19]) that M is weakly bounded in the sense that

$$\mathcal{H}^1(\{s : M\nu(s) > A\}) \lesssim \frac{\|\nu\|}{A} = \frac{\mu(B(x,2r))}{A}.$$

In particular,

$$\mathcal{H}^1\left(\left\{s : M\nu(s) > A\frac{\mu(B(2x,r))}{r}\right\}\right) \lesssim \frac{r}{A}.$$

For $A = A_\delta \geq 1$ large enough, this implies that there is some $s \in [r, (1+\delta)r]$ such that

$$\sup_{\lambda>0} \frac{\nu(B(s,\lambda s))}{2\lambda s} = M\nu(s) \leq A\frac{\mu(B(2x,r))}{r}.$$

Recalling the definition of ν this means precisely that

$$\mu(A(x,(1-\lambda)s,(1+\lambda)s)) \leq \frac{As}{r} \cdot \mu(B(x,2r)) \leq 2A\mu(B(x,2r))$$

for all such λ that $A(x,(1-\lambda)s,(1+\lambda)s) \subset B(x,2r)$. Since $\delta < 1/4$, this covers all $0 < \lambda < 1/2$. For $\lambda > 1/2$, the thin boundaries condition is trivial. $\square$

4.3.2. *Fluctuation of the intrinsic gradient.* We are now ready to state our main technical milestone on the way to Theorem 4.16:

Proposition 4.20. For every $\varepsilon > 0$ and $L \geq 1$ there exist constants $A = A_{\varepsilon,L} \geq 1$ and $\delta = \delta_{\varepsilon,L} > 0$ with the following property. Assume that $\phi\colon \mathbb{W} \to \mathbb{V}$ is an intrinsic L-Lipschitz function with graph Γ. Assume that $x \in \Gamma$ and $r > 0$ are such that

$$\beta_{\mathrm{CG}}(B(x,r)) := \beta_{\mathrm{CG},\Gamma,L}(B(x,r)) \geq \varepsilon,$$

Then, there exists a ball $B(y,s) \subset B(x,r)$ with A-thin boundary (with respect to $\mathcal{H}^3|_\Gamma$) such that $s \geq \delta r$, $\mathrm{dist}_{\mathbb{H}}(y,\Gamma) \leq s/10$, and

$$|\mathbb{E}_{\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)} \nabla^\phi \phi - \mathbb{E}_{\pi_{\mathbb{W}}(B(x,r)\cap\Gamma)} \nabla^\phi \phi| \geq \delta > 0. \tag{4.12}$$

In particular,

$$\mathcal{L}^2(\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)) \gtrsim_{\varepsilon,L} r^3. \tag{4.13}$$

Remark 4.21. Observe that (4.13) is an immediate consequence of y being contained in the $s/10$-neighborhood of Γ; this implies that $B(y,s)\cap\Gamma$ contains a set of the form $B(y',s')\cap\Gamma$ with $y' \in \Gamma$ and $s' \sim s$. Further, for such balls $B(y',s')$ centred on Γ, we can apply Lemma 4.13 to find

$$\pi_{\mathbb{W}}(B(y',s')\cap\Gamma) \supset \pi_{\mathbb{W}}(B(y',s''))$$

for some $s'' \sim_L s'$. Finally,

$$\mathcal{L}^2(\pi_{\mathbb{W}}(B(y',s''))) = c(s'')^3$$

with $c = \mathcal{L}^2(\pi_{\mathbb{W}}(B(0,1)) > 0$, recalling that the mapping $P_p\colon w \mapsto \pi_{\mathbb{W}}(p\cdot w)$, has unit Jacobian for any $p \in \mathbb{H}$, see (4.9). See also [27, Lemma 3.14]. This concludes the proof of (4.13)

We also explain here, why a set $E \subset \mathbb{H}$ with BPiLG has big vertical projections. This is an immediate consequence of Lemma 4.13 if E is an intrinsic Lipschitz graph over an (entire) vertical plane $\mathbb{W}$. In the general case, one can easily deduce the BVP property from the area formula for intrinsic Lipschitz functions (Theorem 1.6 in [14]).

In the following subsections, we proceed with proving the remaining statements of Proposition 4.20. The outline is the following:

4.3.3 We formulate a counter assumption to the main claim in Proposition 4.20. Assuming the validity of this assumption, we find $L \geq 1$ and a sequence of intrinsic L-Lipschitz functions $(\phi_j)_j$ and associated graphs $(\Gamma_j)_j$, such that Γ_j has large β_{CG} number in a ball $B(x_j, r_j)$ centred on Γ_j, yet $\nabla^{\phi_j}\phi_j$ does not fluctuate much in that ball. We use a blow-up procedure to normalize so that we may assume $B(x_j, r_j) = B(0,1)$ for all j.

4.3.4 We show that a subsequence of $(\phi_j)_j$ converges locally uniformly to an intrinsic L-Lipschitz function ϕ with graph Γ such that, roughly speaking,
(i) $\beta_{\mathrm{CG}}(B(0,1))$ is large,
(ii) $\mathbb{E}_{\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)}\nabla^{\phi}\phi$ is independent of $B(y,s) \subset B(0,1)$, $y \in \Gamma$.

4.3.5 We show that the conditions (i) and (ii) are incompatible, which concludes the proof of Proposition 4.20.

4.3.3. *The counter assumption.* Denote by $E(\delta)$ the metric δ-neighborhood of a set E. Given an intrinsic Lipschitz graph Γ, a ball $B(x,r)$ with $x \in \Gamma$, and $j \in \mathbb{N}$, define the following collection of "good" balls $\mathcal{G}_j = \mathcal{G}_j(\Gamma, B(x,r))$. A ball $B(y, sr) \subset B(x,r)$ is in $\mathcal{G}_j$, if

(a) $y \in \Gamma(sr/10)$,
(b) $s \geq 2^{-j}$, and
(c) $B(y, sr)$ has 2^j-thin boundary with respect to $\mathcal{H}^3|_{\Gamma}$.

Then, Proposition 4.20 follows, if we can prove the next statement:

Claim 4.22. For every $\varepsilon > 0$, $L \geq 1$, there exists $j = j_{\varepsilon,L} \in \mathbb{N}$ with the following property. If Γ is any intrinsic L-Lipschitz graph and $B(x,r)$ is centred on Γ with $\beta_{\mathrm{CG}}(B(x,r)) \geq \varepsilon$, then there exists a ball $B = B(y, sr) \in \mathcal{G}_j(\Gamma, B(x,r))$ such that

$$|\mathbb{E}_{\pi_{\mathbb{W}}(B\cap\Gamma)}\nabla^{\phi}\phi - \mathbb{E}_{\pi_{\mathbb{W}}(B(x,r)\cap\Gamma)}\nabla^{\phi}\phi| > \tfrac{1}{j}.$$

If the claim fails, then it also fails with $B(x,r) = B(0,1)$. This reduction is the content of the next lemma:

Lemma 4.23. *If Claim 4.22 fails, then there there exist $\varepsilon > 0$ and $L \geq 1$ such that for every $j \in \mathbb{N}$ we can find an intrinsic L-Lipschitz graph Γ_j, parametrised by $\phi_j \colon \mathbb{W} \to \mathbb{V}$ such that*

$$\beta_{\mathrm{CG},\Gamma_j}(B(0,1)) \geq \varepsilon,$$

yet

$$\sup_{B\in\mathcal{G}_j(\Gamma_j,B(0,1))} |\mathbb{E}_{\pi_{\mathbb{W}}(B\cap\Gamma_j)}\nabla^{\phi_j}\phi_j - \mathbb{E}_{\pi_{\mathbb{W}}(B(0,1)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j| \leq \tfrac{1}{j}. \tag{4.14}$$

Proof. By definition, if Claim 4.22 fails, then there exist $\varepsilon > 0$ and $L \geq 1$ such that for every $j \in \mathbb{N}$ we can find an intrinsic L-Lipschitz graph Γ_j, parametrised by $\phi_j \colon \mathbb{W} \to \mathbb{V}$, and *some ball* $B(x_j, r_j)$ centred on Γ_j such that

$$\beta_{\mathrm{CG},\Gamma_j}(B(x_j,r_j)) \geq \varepsilon,$$

yet

$$\sup_{B\in\mathcal{G}_j(\Gamma_j,B(x_j,r_j))} |\mathbb{E}_{\pi_{\mathbb{W}}(B\cap\Gamma_j)}\nabla^{\phi_j}\phi_j - \mathbb{E}_{\pi_{\mathbb{W}}(B(x_j,r_j)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j| \leq \tfrac{1}{j}. \tag{4.15}$$

In order to prove Lemma 4.23, we left-translate Γ_j by x_j^{-1} and dilate it by $\delta_{r_j^{-1}}$. The resulting set $\widetilde{\Gamma}_j$:

(i) is again an intrinsic L-Lipschitz graph (Lemma 4.7),
(ii) has β_{CG}-number at least ε on $B(0,1)$ (Lemma 4.24 below),
(iii) is parametrised by a function $\widetilde{\phi}_j : \mathbb{W} \to \mathbb{V}$ so that (4.15) holds with $\widetilde{\phi}_j$ in place of ϕ_j. (Lemma 4.25 below).

□

We now proceed to establish the two auxiliary results, needed in the proof of Lemma 4.23.

Lemma 4.24. *Assume that $\phi : \mathbb{W} \to \mathbb{V}$ is an intrinsic Lipschitz function with graph Γ, x is a point on Γ, and $r > 0$. Then*

$$\beta_{\mathrm{CG},L,\Gamma}(B(x,r)) = \beta_{\mathrm{CG},L,\widetilde{\Gamma}}(B(0,1))$$

for $\widetilde{\Gamma} = \delta_{r^{-1}}(\tau_{x^{-1}}(\Gamma))$.

Proof. According to Lemma 4.7 and Remark 2.5 there exists a uniquely defined intrinsic Lipschitz function $\widetilde{\phi}$ that parametrizes $\widetilde{\Gamma}$. By definition,

$$\beta_{\mathrm{CG},L,\widetilde{\Gamma}}(B(0,1)) = \inf_{\sigma \in \mathrm{Adm}_{\mathrm{CG},L}(B(0,1))} \sup_{w \in \pi_{\mathbb{W}}(B(0,1)\cap\widetilde{\Gamma})} |\widetilde{\phi}(w) - \sigma(w)|.$$

We first aim to prove that the family $\mathrm{Adm}_{\mathrm{CG},L}(B(0,1)) =: \mathrm{Adm}(B(0,1))$ is in 1-to-1 correspondence with the family $\mathrm{Adm}_{\mathrm{CG},L}(B(x,r)) =: \mathrm{Adm}(B(x,r))$. Assume that $\psi \in \mathrm{Adm}(B(x,r))$. Define

$$\widetilde{\psi} : \mathbb{W} \to \mathbb{V}, \quad \widetilde{\psi}(w) := \delta_{\frac{1}{r}}\left(\pi_{\mathbb{V}}(x)^{-1} \cdot \psi\left(P_x(\delta_r(w))\right)\right), \tag{4.16}$$

and note that $\tilde{\psi} = (\psi_{x^{-1}})_{r^{-1}}$ in the notation of Lemma 4.7. Let us prove that $\tilde{\psi} \in \mathrm{Adm}(B(0,1))$. To this end, we observe first that

$$\pi_{\mathbb{W}}\left(x \cdot \delta_r(\pi_{\mathbb{W}}(p))\right) = \pi_{\mathbb{W}}\left(x \cdot \delta_r(p)\right), \quad \text{for all } p \in \mathbb{H},$$

see for instance [39, Proposition 2.15]. By homogeneity and left invariance of the distance $d_{\mathbb{H}}$, it follows that

$$P_x\left(\delta_r\left(\pi_{\mathbb{W}}(B(0,1)\cap\widetilde{\Gamma})\right)\right) = \pi_{\mathbb{W}}(B(x,r)\cap\Gamma) \tag{4.17}$$

and

$$P_x\left(\delta_r\left(\pi_{\mathbb{W}}(B(0,b_L))\right)\right) = \pi_{\mathbb{W}}(B(x,b_L r)).$$

Lemma 4.7 then implies that $\tilde{\psi} \in \mathrm{Adm}(B(0,1))$.

Conversely, if $\sigma \in \mathrm{Adm}(B(0,1))$, then $(\sigma_r)_x \in \mathrm{Adm}(B(x,r))$. Thus, by (4.17),

$$\begin{aligned}\beta_{\mathrm{CG},L,\widetilde{\Gamma}}(B(0,1)) &= \inf_{\psi \in \mathrm{Adm}(B(x,r))} \sup_{w \in \pi_{\mathbb{W}}(B(0,1)\cap\widetilde{\Gamma})} |\widetilde{\phi}(w) - \widetilde{\psi}(w)| \\ &= \inf_{\psi \in \mathrm{Adm}(B(x,r))} \sup_{w \in \pi_{\mathbb{W}}(B(0,1)\cap\widetilde{\Gamma})} r^{-1}|\phi(P_x(\delta_r(w))) - \psi(P_x(\delta_r(w)))| \\ &= \inf_{\psi \in \mathrm{Adm}(B(x,r))} \sup_{w \in \pi_{\mathbb{W}}(B(x,r)\cap\Gamma)} r^{-1}|\phi(w) - \psi(w)| \\ &= \beta_{\mathrm{CG},L,\Gamma}(B(x,r)).\end{aligned}$$

This completes the proof. $\square$

Lemma 4.25. *Assume that $\phi : \mathbb{W} \to \mathbb{V}$ is an intrinsic Lipschitz function with graph Γ, x is a point on Γ, and $r > 0$. Then*

$$\delta_{r^{-1}}\tau_{x^{-1}}\mathcal{G}_j(\Gamma, B(x,r)) = \mathcal{G}_j(\widetilde{\Gamma}, B(0,1)) \tag{4.18}$$

and

$$\mathbb{E}_{\pi_{\mathbb{W}}(B(y,sr)\cap\Gamma)}\nabla^{\phi}\phi = \mathbb{E}_{\pi_{\mathbb{W}}(B(\delta_{r^{-1}}(x^{-1}\cdot y),s)\cap\widetilde{\Gamma})}\nabla^{\widetilde{\phi}}\widetilde{\phi}. \tag{4.19}$$

Here $\widetilde{\Gamma} = \delta_{r^{-1}}(\tau_{x^{-1}}(\Gamma))$ is the graph parametrized by $\widetilde{\phi}$ (defined as in (4.16)).

Proof. We start with the first claim. Since the Heisenberg distance is left invariant with respect to the group law, and homogeneous with respect to the dilations $(\delta_r)_{r>0}$, if a ball $B(y,sr)$ has A-thin boundary, then so does the ball $\delta_{r^{-1}}\tau_{x^{-1}}B(y,sr) = B(\delta_{r^{-1}}(x^{-1}\cdot y),s)$. The remaining conditions that one has to verify in order to prove (4.18) are also immediate.

Regarding (4.19), we first recall that Lemma 4.7 yields that

$$\nabla^{\widetilde{\phi}}\widetilde{\phi} = \nabla^{\phi}\phi \circ P_x \circ \delta_r.$$

Since δ_r restricted to $\mathbb{W}$ has Jacobian determinant constant equal to r^3, and P_x has Jacobian determinant equal to 1, it follows by the usual transformation formula for functions on $\mathbb{R}^2$ that

$$\mathcal{L}^2(\pi_{\mathbb{W}}(B(y,sr)\cap\Gamma)) = r^3\mathcal{L}^2(\pi_{\mathbb{W}}(B(\delta_{r^{-1}}(x^{-1}\cdot y),s)\cap\widetilde{\Gamma}))$$

and

$$\int_{\pi_{\mathbb{W}}(B(y,sr)\cap\Gamma)} \nabla^{\phi}\phi \, d\mathcal{L}^2 = r^3 \int_{\pi_{\mathbb{W}}(B(\delta_{r^{-1}}(x^{-1}\cdot y),s)\cap\widetilde{\Gamma})} \nabla^{\widetilde{\phi}}\widetilde{\phi}\, d\mathcal{L}^2.$$

This establishes (4.19). $\square$

4.3.4. *Limiting procedure.* In this section, we work under the standing (counter) assumption to Proposition 4.20. In particular, we may assume by Lemma 4.23 that there exists $\varepsilon, L > 0$ and a sequence $(\phi_j)_{j\in\mathbb{N}}$ of intrinsic L-Lipschitz functions with graphs $(\Gamma_j)_j$ such that $\beta_{\mathrm{CG},\Gamma_j,L}(B(0,1)) \geq \varepsilon$, yet the intrinsic gradient $\nabla^{\phi_j}\phi_j$ fluctuates only little in $B(0,1)\cap\Gamma_j$ as quantified in (4.14).

The main goal of this section is to consider an "accumulation point" ϕ of the sequence $(\phi_j)_j$, and to discuss how the properties of the maps ϕ_j carry over to ϕ.

Lemma 4.26. *The sequence $(\phi_j)_j$ defined above contains a subsequence that converges locally uniformly on $\mathbb{W}$ to an intrinsic L-Lipschitz function $\phi : \mathbb{W} \to \mathbb{V}$ with graph Γ such that*

$$\beta_{\mathrm{CG},\Gamma,L}(B(0,1)) \geq \varepsilon. \tag{4.20}$$

Proof. Since each graph Γ_j by construction contains the origin, we have $\phi_j(0) = 0$ for every j. Therefore 4.2 implies that the family $(\phi_j)_{j\in\mathbb{N}}$ is locally equibounded. Hence, by Proposition 3.10 in [30], there exists a subsequence which converges locally uniformly to an intrinsic L-Lipschitz function ϕ on $\mathbb{W}$ with intrinsic graph Γ. For simplicity, we also denote this subsequence by $(\phi_j)_{j\in\mathbb{N}}$.

Let $\delta > 0$. In order to prove that $\beta_{\mathrm{CG},\Gamma,L}(B(0,1)) \geq \varepsilon$, it suffices to fix $\psi \in \mathrm{Adm}_{\mathrm{CG},L}(B(0,1))$ and find a point $w_\psi \in \pi_{\mathbb{W}}(\Gamma \cap B(0,1))$ with $|\psi(w_\psi) - \phi(w_\psi)| \geq (1-\delta)\varepsilon$. To this end, the assumption

$$\beta_{\mathrm{CG},\Gamma_j,L}(B(0,1)) \geq \varepsilon, \quad \text{for all } j \in \mathbb{N},$$

implies that, for each $j \in \mathbb{N}$, there exists a point $w_\psi^j \in \pi_{\mathbb{W}}(\Gamma_j \cap B(0,1))$ such that

$$|\psi(w_\psi^j) - \phi_j(w_\psi^j)| \geq (1-\delta)\varepsilon.$$

Write $p_\psi^j := w_\psi^j \cdot \phi_j(w_\psi^j)$. The sequence $(p_\psi^j)_{j\in\mathbb{N}}$ has a subsequence $(p_\psi^{j_k})_{k\in\mathbb{N}}$ convergent to a point $p_\psi \in \Gamma \cap B(0,1)$. Since $\pi_{\mathbb{W}}$ is continuous, the points $w_\psi^{j_k} = \pi_{\mathbb{W}}(p_\psi^{j_k})$ converge to $w_\psi := \pi_{\mathbb{W}}(p_\psi) \in \pi_{\mathbb{W}}(\Gamma \cap B(0,1))$.

$$|\psi(w_\psi) - \phi(w_\psi)| = \lim_{k\to\infty} |\psi(w_\psi^{j_k}) - \phi_{j_k}(w_\psi^{j_k})| \geq (1-\delta)\varepsilon$$

by the continuity of ψ, and the locally uniform convergence $\phi_j \to \phi$. The proof is complete. □

Without loss of generality, we assume in the following that the whole sequence $(\phi_j)_{j\in\mathbb{N}}$ converges locally uniformly to ϕ. Our next goal is to prove the following convergence result for the corresponding intrinsic gradients.

Lemma 4.27. *Let $(\phi_j)_j : \mathbb{W} \to \mathbb{V}$ be a sequence of intrinsic L-Lipschitz functions converging locally uniformly to an L-Lipschitz function $\phi : \mathbb{W} \to \mathbb{V}$. Then*

$$\int_{\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)} \nabla^\phi \phi \, d\mathcal{L}^2 = \lim_{j\to\infty} \int_{\pi_{\mathbb{W}}(B(y,s)\cap\Gamma_j)} \nabla^{\phi_j} \phi_j \, d\mathcal{L}^2 \tag{4.21}$$

for all balls $B(y,s) \subset B(0,1)$ with $y \in \Gamma$, and such that $\mathcal{H}^3(\partial B(y,s) \cap \Gamma) = 0$.

The subtle point here is, of course, that the domain of integration is different on both sides of the above equation. Therefore the following auxiliary result will be useful in the proof of Lemma 4.27.

Lemma 4.28. *Assume that Γ_j is a sequence of intrinsic L-Lipschitz graphs, which converges locally in the Hausdorff metric in $\mathbb{H}$ to an intrinsic Lipschitz graph Γ. Then, for any $x \in \Gamma$ and $0 < r < s < \infty$, we have*

$$\limsup_{j\to\infty} \mathcal{H}^3(A(x,r,s) \cap \Gamma_j) \lesssim_L \mathcal{H}^3(A(x,r,s) \cap \Gamma).$$

Proof. Fix $\delta > 0$. First, pick $r_1 < r$ and $s_1 > s$ so that

$$\mathcal{H}^3(A(x,r_1,s_1) \cap \Gamma) \leq \mathcal{H}^3(A(x,r,s) \cap \Gamma) + \delta.$$

Then, let $\varepsilon > 0$ be so small that if $y \in A(x,r,s)$, then $B(y,20\varepsilon) \subset A(x,r_1,s_1)$. Let $\{y_1,\ldots,y_N\}$ be an ε-net in $A(x,r,s)$, and, finally, let Y_j be the subset of points y in this net with the property that $B(y,5\varepsilon)$ contains a point in $A(x,r,s) \cap \Gamma_j$. Then

$$\mathcal{H}^3(A(x,r,s) \cap \Gamma_j) \leq \sum_{y \in Y_j} \mathcal{H}^3(B(y,5\varepsilon) \cap \Gamma_j) \lesssim_L |Y_j| \cdot \varepsilon^3,$$

by the 3-regularity of intrinsic L-Lipschitz graphs. For j large enough, every ball $B(y,10\varepsilon)$ with $y \in Y_j$ also contains a point in Γ, whence $\varepsilon^3 \lesssim_L \mathcal{H}^3(B(y,20\varepsilon) \cap \Gamma)$, $y \in Y_j$. Since the balls $B(y,20\varepsilon) \subset A(x,r_1,s_1)$ have bounded overlap (independent of $\varepsilon > 0$), we conclude that

$$\mathcal{H}^3(A(x,r,s) \cap \Gamma_j) \lesssim_L \mathcal{H}^3(A(x,r_1,s_1) \cap \Gamma) \leq \mathcal{H}^3(A(x,r,s) \cap \Gamma) + \delta$$

for all large enough j. This completes the proof. □

Proof of Lemma 4.27. We employ the fact proven in [14, Proposition 4.7] that the intrinsic gradient is also a distributional gradient for intrinsic Lipschitz functions: if ϕ is intrinsic Lipschitz and defined on $\mathbb{W}$, then

$$\int_{\mathbb{W}} (\phi \partial_y \psi + \tfrac{1}{2}\phi^2 \partial_t \psi) \, d\mathcal{L}^2 = -\int_{\mathbb{W}} [\nabla^\phi \phi] \psi \, d\mathcal{L}^2$$

for all compactly supported $\mathcal{C}^1$-functions ψ on $\mathbb{W}$ (here we assume that $\mathbb{W}$ is the (y,t)-plane, as we may). Since uniform convergence implies weak convergence, we infer that

$$\int_{\mathbb{W}} [\nabla^\phi \phi] \psi \, d\mathcal{L}^2 = \lim_{j \to \infty} \int_{\mathbb{W}} [\nabla^{\phi_j} \phi_j] \psi \, d\mathcal{L}^2 \tag{4.22}$$

for all compactly supported $\mathcal{C}^1$ functions $\psi \colon \mathbb{W} \to \mathbb{V}$. It remains to deduce from this statement the claim (4.21).

To achieve this, we recall from Lemma 4.6 that

$$\|\nabla^{\phi_j} \phi_j\|_\infty \leq L, \quad j \in \mathbb{N}, \quad \text{and} \quad \|\nabla^\phi \phi\|_\infty \leq L. \tag{4.23}$$

Then, given $B(y,s) \subset B(0,1)$ with $y \in \Gamma$, for $\varepsilon > 0$, choose an open set U in $\pi_{\mathbb{W}}(B(0,2))$ so that

$$\overline{\pi_{\mathbb{W}}(B(y,s) \cap \Gamma)} \subset U,$$

and $\mathcal{L}^2(U \setminus \pi_{\mathbb{W}}(B(y,s) \cap \Gamma)) \leq \varepsilon/4L$. In particular,

$$\left| \int_{U \setminus \pi_{\mathbb{W}}(B(y,s) \cap \Gamma)} \left(\nabla^{\phi_j} \phi_j - \nabla^\phi \phi \right) \psi \, d\mathcal{L}^2 \right| \leq \tfrac{\varepsilon}{2} \tag{4.24}$$

for all test functions ψ with $0 \leq \psi \leq 1$ and all $j \in \mathbb{N}$. Now, let ψ be a smooth cut-off function with compact support in U, $0 \leq \psi \leq 1$ on U and such that $\psi \equiv 1$

on $\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)$. Then, by (4.22), there exists $j(\varepsilon)\in\mathbb{N}$ such that for $j\geq j(\varepsilon)$, one has

$$\left|\int_{\mathbb{W}}\left(\nabla^{\phi_j}\phi_j-\nabla^{\phi}\phi\right)\psi\,d\mathcal{L}^2\right|\leq\tfrac{\varepsilon}{2}. \tag{4.25}$$

Let us denote $B:=\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)$. Combining (4.24) and (4.25), and recalling that ψ is supported on U, we find that

$$\left|\int_B\nabla^{\phi_j}\phi_j-\nabla^{\phi}\phi d\mathcal{L}^2\right|=\left|\int_U\left(\nabla^{\phi_j}\phi_j-\nabla^{\phi}\phi\right)\psi d\mathcal{L}^2-\int_{U\setminus B}\left(\nabla^{\phi_j}\phi_j-\nabla^{\phi}\phi\right)\psi d\mathcal{L}^2\right|\leq\varepsilon$$

for $j\geq j(\varepsilon)$. This yields

$$\int_{\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)}\nabla^{\phi}\phi\,d\mathcal{L}^2=\lim_{j\to\infty}\int_{\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)}\nabla^{\phi_j}\phi_j\,d\mathcal{L}^2.$$

To establish (4.21), it remains to show that Γ on the right hand side of the above equation can be replaced by Γ_j. This follows immediately from the uniform bound for $\nabla^{\phi}\phi$ and $\nabla^{\phi_j}\phi_j$, see (4.23), provided we can show the following:

$$\mathcal{L}^2(\pi_{\mathbb{W}}(B(y,s)\cap\Gamma_j)\triangle\pi_{\mathbb{W}}(B(y,s)\cap\Gamma))\to 0,\quad\text{as } j\to\infty. \tag{4.26}$$

We now prove (4.26). Here we need to assume that $\mathcal{H}^3(\partial B(y,s)\cap\Gamma)=0$. Since

$$\partial B(y,s)\cap\Gamma=\bigcap_{n\in\mathbb{N}}A(y,(1-\tfrac{1}{n})s,s)\cap\Gamma,$$

there exists for every $\delta>0$ a number $n\in\mathbb{N}$ such that

$$\mathcal{H}^3\left(A(y,(1-\tfrac{1}{n})s,s)\cap\Gamma\right)\leq\delta. \tag{4.27}$$

Let $n=n(\delta)$ be such and observe that

$$\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)\subset\pi_{\mathbb{W}}(B(y,(1-\tfrac{1}{n})s)\cap\Gamma)\cup\pi_{\mathbb{W}}(A(y,(1-\tfrac{1}{n})s,s)\cap\Gamma)$$

and

$$\pi_{\mathbb{W}}(B(y,(1-\tfrac{1}{n})s)\cap\Gamma)\subset\pi_{\mathbb{W}}(B(y,s)\cap\Gamma_j)$$

for j larger than some $j(\delta)$ given by uniform convergence. So

$$\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)\setminus\pi_{\mathbb{W}}(B(y,s)\cap\Gamma_j)\subset\pi_{\mathbb{W}}(A(y,(1-\tfrac{1}{n})s,s)\cap\Gamma)$$

for $j\geq j(\delta)$. Since

$$\mathcal{L}^2(\pi_{\mathbb{W}}(A(y,(1-\tfrac{1}{n})s,s)\cap\Gamma))\lesssim\mathcal{H}^3(A(y,(1-\tfrac{1}{n})s,s)\cap\Gamma))\leq\delta,$$

by the general inequality $\mathcal{L}^2(\pi_{\mathbb{W}}(B))\lesssim\mathcal{H}^3(B)$ (Lemma 3.6) and by (4.27), we infer that

$$\mathcal{L}^2\left(\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)\setminus\pi_{\mathbb{W}}(B(y,s)\cap\Gamma_j)\right)\to 0,$$

as $j\to\infty$. It remains to prove the same with the roles of Γ and Γ_j reversed. Repeating the argument above, it suffices to estimate the measure of the projection

$\pi_{\mathbb{W}}(A(y,(1-\frac{1}{n})s,s)\cap\Gamma_j)$. This is not altogether trivial, since the assumption on $\partial B(y,s)$ concerns $\mathcal{H}^3|_{\Gamma}$, not $\mathcal{H}^3|_{\Gamma_j}$. However, Lemma 4.28 still implies that

$$\limsup_{j\to\infty}\mathcal{H}^3(A(y,(1-\tfrac{1}{n})s,s)\cap\Gamma_j)\lesssim_L \mathcal{H}^3(A(y,(1-\tfrac{1}{n})s,s)\cap\Gamma)\leq\delta,$$

which proves that

$$\limsup_{j\to\infty}\mathcal{L}^2(\pi_{\mathbb{W}}(B(y,s)\cap\Gamma_j)\setminus\pi_{\mathbb{W}}(B(y,s)\cap\Gamma))\lesssim\delta.$$

Since $\delta>0$ is arbitrary, (4.26), and hence also (4.21), follows. This concludes the proof of Lemma 4.27. $\square$

We now return to the specific sequence $(\phi_j)_{j\in\mathbb{N}}$ and limit function ϕ (with graph Γ) given by Lemma 4.26. In particular, we work under the standing assumptions that $\beta_{\mathrm{CG},\Gamma_j}(B(0,1))\geq\varepsilon$ and $\nabla^{\phi_j}\phi_j$ fluctuates only little on $\pi_{\mathbb{W}}(B(0,1)\cap\Gamma_j)$, as made precise in Lemma 4.23.

Lemma 4.29. *Let ϕ be as in Lemma 4.26. Then $\mathbb{E}_{\pi_{\mathbb{W}}(B(y,s))}\nabla^{\phi}\phi$ is a constant independent of $B(y,s)\subset B(0,1)$, for all balls $B(y,s)$ such that $y\in\Gamma$ and $\mathcal{H}^3(\partial B(y,s)\cap\Gamma)=0$.*

Proof. Pick two balls $B(y_1,s_1)\subset B(0,1)$ and $B(y_2,s_2)\subset B(0,1)$ with $y_1,y_2\in\Gamma$ satisfying the assumption stated in the lemma, and let $\delta>0$ be arbitrary. Our goal is to show that

$$|\mathbb{E}_{\pi_{\mathbb{W}}(B(y_1,s_1)\cap\Gamma)}\nabla^{\phi}\phi-\mathbb{E}_{\pi_{\mathbb{W}}(B(y_2,s_2)\cap\Gamma)}\nabla^{\phi}\phi|\lesssim\delta.$$

We start by applying the triangle inequality:

$$|\mathbb{E}_{\pi_{\mathbb{W}}(B(y_1,s_1)\cap\Gamma)}\nabla^{\phi}\phi-\mathbb{E}_{\pi_{\mathbb{W}}(B(y_2,s_2)\cap\Gamma)}\nabla^{\phi}\phi|\leq D_1+D_2+D_3,$$

where

$$\begin{aligned}
D_1 &:= |\mathbb{E}_{\pi_{\mathbb{W}}(B(y_1,s_1)\cap\Gamma)}\nabla^{\phi}\phi-\mathbb{E}_{\pi_{\mathbb{W}}(B(y_1,s_1)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j|,\\
D_2 &:= |\mathbb{E}_{\pi_{\mathbb{W}}(B(y_2,s_2)\cap\Gamma)}\nabla^{\phi}\phi-\mathbb{E}_{\pi_{\mathbb{W}}(B(y_2,s_2)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j|,\\
D_3 &:= |\mathbb{E}_{\pi_{\mathbb{W}}(B(y_1,s_1)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j-\mathbb{E}_{\pi_{\mathbb{W}}(B(y_2,s_2)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j|.
\end{aligned}$$

To see that $D_1+D_2\leq\delta$ for large enough j, writing temporarily $|U|:=\mathcal{L}^2(U)$ for $U\subset\mathbb{W}$, and

$$B:=\pi_{\mathbb{W}}(B(y_1,s_1)\cap\Gamma)\quad\text{and}\quad B_j:=\pi_{\mathbb{W}}(B(y_1,s_1)\cap\Gamma_j),$$

we perform the following estimate:

$$\begin{aligned}
D_1 &= |\mathbb{E}_B\nabla^{\phi}\phi-\mathbb{E}_{B_j}\nabla^{\phi_j}\phi_j|\\
&= \frac{1}{|B|}\left|\int_B\nabla^{\phi}\phi\,d\mathcal{L}^2-\int_{B_j}\nabla^{\phi_j}\phi_j\,d\mathcal{L}^2\right|+\frac{||B|-|B_j||}{|B||B_j|}\left|\int_{B_j}\nabla^{\phi_j}\phi_j\,d\mathcal{L}^2\right|\\
&\leq \frac{1}{|B|}\left|\int_B\nabla^{\phi}\phi d\mathcal{L}^2-\int_{B_j}\nabla^{\phi_j}\phi_j d\mathcal{L}^2\right|+\frac{|B_j\triangle B|}{|B|}\|\nabla^{\phi_j}\phi_j\|_{\infty}.
\end{aligned}$$

The same is true with "1" replaced by "2". By (4.21) and (4.26), and the uniform bound $\|\nabla^{\phi_j}\phi_j\|_\infty \leq L$, both terms above tend to zero as $j \to \infty$. So, it suffices to deal with D_3.

We first make the subsequent estimate for certain $s_1^j \geq s_1$ and $s_2^j \geq s_2$ to be fixed soon:

$$D_3 \leq E_1 + E_2 + E_3,$$

where

$$\begin{aligned} E_1 &= |\mathbb{E}_{\pi_{\mathbb{W}}(B(y_1,s_1)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j - \mathbb{E}_{\pi_{\mathbb{W}}(B(y_1,s_1^j)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j|, \\ E_2 &= |\mathbb{E}_{\pi_{\mathbb{W}}(B(y_2,s_2)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j - \mathbb{E}_{\pi_{\mathbb{W}}(B(y_2,s_2^j)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j|, \\ E_3 &= |\mathbb{E}_{\pi_{\mathbb{W}}(B(y_1,s_1^j)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j - \mathbb{E}_{\pi_{\mathbb{W}}(B(y_2,s_2^j)\cap\Gamma_j)}\nabla^{\phi_j}\phi_j|. \end{aligned}$$

It follows from the key (counter) assumption (4.14) that $E_3 \leq 2/j$ if the balls $B(y_1, s_1^j)$ and $B(y_2, s_2^j)$ belong to $\mathcal{G}_j(\Gamma_j, B(0,1))$. This requires that

(a) $\operatorname{dist}(y_1, \Gamma_j) \leq s_1/10$ and $\operatorname{dist}(y_2, \Gamma_j) \leq s_2/10$,
(b) $s_1^j \geq 2^{-j}$ and $s_2^j \geq 2^{-j}$, and
(c) $B(y_1, s_1^j)$ and $B(y_2, s_2^j)$ have 2^j-thin boundary with respect to $\mathcal{H}^3|_{\Gamma_j}$.

Condition (a) is automatically satisfied for large enough j, since $y_1, y_2 \in \Gamma$. Condition (b) is also trivially satisfied for large enough j, since $s_1^j \geq s_1$ and $s_2^j \geq s_2$. To achieve condition (c), fix an auxiliary parameter $\tau > 0$, which will depend on $\delta, y_1, y_2, s_1, s_2, \Gamma$ and L. Then, for large enough j, Lemma 4.19 guarantees the existence of $s_1^j \in [s_1, (1+\tau)s_1]$ and $s_2^j \in [s_2, (1+\tau)s_2]$ such that (c) is satisfied for $B(y_1, s_1^j)$ and $B(y_2, s_2^j)$. We choose s_1^j, s_2^j accordingly, and then $E_3 \leq 2/j \leq \delta$ (for $j \geq 2/\delta$).

It remains to estimate E_1 and E_2; by symmetry, we may concentrate on E_1. Recalling Lemma 4.17, we first have

$$E_1 \lesssim \frac{\mathcal{H}^3(A(y_1, s_1, s_1^j) \cap \Gamma_j)}{\mathcal{L}^2(\pi_{\mathbb{W}}(B(y_1, s_1^j) \cap \Gamma_j))} \lesssim_L \frac{\mathcal{H}^3(A(y_1, s_1, (1+\tau)s_1) \cap \Gamma_j)}{s_1^3}$$

since $s_1^j \leq (1+\tau)s_1$, and $\mathcal{L}^2(\pi_{\mathbb{W}}(B(y_1, s_1^j) \cap \Gamma_j)) \gtrsim_L s_1^3$, see Remark 4.21. Lemma 4.28 then tells us that

$$\limsup_{j\to\infty} \mathcal{H}^3(A(y_1, s_1, (1+\tau)s_1) \cap \Gamma_j) \lesssim_L \mathcal{H}^3(A(y_1, s_1, (1+\tau)s_1) \cap \Gamma),$$

and now we choose $\tau = \tau(\delta, y_1, y_2, s_1, s_2, \Gamma, L)$ so small that the right hand side above is smaller than $c_L\delta$ for some small constant c_L depending only on L, and the same holds with "1" replaced by "2". If c_L is small enough, this implies that $E_1 + E_2 \leq \delta$.

Combining the estimates for D_i and E_i above, we conclude that

$$|\mathbb{E}_{\pi_{\mathbb{W}}(B(y_1,s_1)\cap\Gamma)}\nabla^{\phi}\phi - \mathbb{E}_{\pi_{\mathbb{W}}(B(y_2,s_2)\cap\Gamma)}\nabla^{\phi}\phi| \leq 3\delta.$$

Since $\delta > 0$ was arbitrary, this proves that $\mathbb{E}_{\pi_{\mathbb{W}}(B(y,s)\cap\Gamma)}\nabla^{\phi}\phi$ is independent of the ball $B(y,s) \subset B(0,1)$ with $y \in \Gamma$ and satisfying $\mathcal{H}^3(\partial B(y,s) \cap \Gamma) = 0$. $\square$

4.3.5. *Conclusion.* In this section, we conclude the proof of Proposition 4.20. Under the standing counter assumption derived in Lemma 4.23, we will infer from Lemma 4.29 that $\nabla^{\phi}\phi$ is a.e. constant on $\pi_{\mathbb{W}}(B(0,1)\cap\Gamma) \supseteq \pi_{\mathbb{W}}(B(0,b_L))$. This will contradict the conclusion of Lemma 4.26, which gave $\beta_{\mathrm{CG},\Gamma,L}(B(0,1)) \geq \varepsilon$.

Lemma 4.30. *Let ϕ be as in Lemma 4.26, with graph Γ. Then $\phi \in \mathrm{Adm}_{\mathrm{CG},L}(B(0,1))$.*

Recall that we have shown in Lemma 4.29 that the averages of $\nabla^{\phi}\phi$ are the same for a large class of sets in $\mathbb{W}$. In order to conclude that then $\nabla^{\phi}\phi$ must be constant almost everywhere on $\pi_{\mathbb{W}}(B(0,1)\cap\Gamma)$, and in particular on $\pi_{\mathbb{W}}(B(0,b_L))$, we apply a Lebesgue differentiation theorem for the measure space $(\mathbb{W},\mathcal{L}^2)$ endowed with the *graph distance* d_Γ. The latter can be defined for an arbitrary intrinsic Lipschitz function $\phi : \mathbb{W} \to \mathbb{V}$ with graph Γ by setting

$$d_\Gamma(w,w') = d_{\mathbb{H}}(w\cdot\phi(w), w'\cdot\phi(w')), \quad w,w' \in \mathbb{W}.$$

Proof of Lemma 4.30. The triple $(\mathbb{W}, d_\Gamma, \mathcal{L}^2)$ is a metric space of homogeneous type, so the usual Lebesgue differentiation theorem is valid:

$$\lim_{r\to 0+} \frac{1}{\mathcal{L}^2(B_{d_\Gamma}(w_0,r))} \int_{B_{d_\Gamma}(y,s)} f(w)\, d\mathcal{L}^2 = f(w_0), \quad \mathcal{L}^2 \text{ a.e. } w_0 \in \mathbb{W}, \tag{4.28}$$

holds for all $f \in L^1_{\mathrm{loc}}(\mathbb{W},\mathcal{L}^2)$. See for instance [31].

We apply this theorem for $f = \nabla^{\phi}\phi$. More precisely, we choose for every Lebesgue point w_0 of $f = \nabla^{\phi}\phi$ a sequence of radii $(s_k)_{k\in\mathbb{N}}$ such that $s_k \to 0$ and $\mathcal{H}^3(\partial B(y_0,s_k)\cap\Gamma) = 0$ for $\pi_{\mathbb{W}}(y_0) = w_0$. Then we apply (4.28) to conclude that

$$\lim_{k\to\infty} \mathbb{E}_{B_{d_\Gamma}(w_0,s_k)} \nabla^{\phi}\phi = \nabla^{\phi}\phi(w_0). \tag{4.29}$$

Without loss of generality we may assume here that w_0 has been chosen so that the pointwise intrinsic gradient $\nabla^{\phi}\phi$ exists in w_0.

Finally, it follows immediately from the definition of the graph distance that

$$B_{d_\Gamma}(\pi_{\mathbb{W}}(y), s) = \pi_{\mathbb{W}}(B(y,s)\cap\Gamma) \tag{4.30}$$

for all y on the graph of ϕ and $s > 0$, and hence the claim of the lemma follows from (4.29) and Lemma 4.29. □

Proof of Proposition 4.20. According to Remark 4.21, it suffices prove (4.12). Assume that this is not true. Then, by Lemma 4.26 and Lemma 4.30, there exists an intrinsic Lipschitz function $\phi\colon \mathbb{W} \to \mathbb{V}$ with graph Γ such that

$$\phi \in \mathrm{Adm}_{\mathrm{CG},L}(B(0,1)) \quad \text{and} \quad \beta_{\mathrm{CG},\Gamma,L}(B(0,1)) \geq \varepsilon.$$

This contradiction proves Proposition 4.20. □

4.4. **Proof of the weak geometric lemma for constant gradient β-numbers.** In this section, we apply Proposition 4.20 to prove the weak geometric lemma for the β_{CG}-numbers, Theorem 4.16. In brief, if the integral appearing in Theorem 4.16 is large, then we can find many balls centered on the graph Γ with large β_{CG}-number (this is quantified in Lemma 4.32). Then, Proposition 4.20 implies that $\nabla^{\phi}\phi$ fluctuates strongly inside (the projections of) such balls, which can be used to find a lower bound for $\|\nabla^{\phi}\phi\|^2_{L^2(\pi_{\mathbb{W}}(B(x,R)\cap\Gamma)}$ with $x \in \Gamma$ and $R > 0$. Since we also have the upper bound $\lesssim LR^3$ for this quantity, we will eventually find the correct upper bound for the integral in Theorem 4.16.

Definition 4.31. A collection $\mathcal{B}$ of balls in a metric space (X,d) is *pre-dyadic*, if

$$\sum_{\substack{B\in\mathcal{B}\\ r/N\leq \mathrm{diam}(B)<r}} \chi_B(x) \lesssim_N 1, \quad x \in X,\ r > 0,$$

for any $N > 1$. A collection of balls (or sets in general) in (X,d) is *dyadic*, if for every pair of sets B, B' in the family, either $B\cap B' = \emptyset$, or $B \subset B'$, or $B' \subset B$.

Lemma 4.32. *Let Γ be an intrinsic Lipschitz graph, and let $x \in \Gamma$, $R > 0$. Write*

$$\int_0^R \int_{\Gamma\cap B(x,R)} \chi_{\{(y,s)\in\Gamma\times\mathbb{R}_+ :\, \beta_{\mathrm{CG}}(B(y,s))\geq\varepsilon\}}(y,s)\, d\mathcal{H}^3(y)\,\frac{ds}{s} =: C.$$

Then, there exists a pre-dyadic family of balls $B(x_i,r_i) \subset B(x,2R)$ with $x_i \in \Gamma$ such that $\beta_{\mathrm{CG}}(B(x_i,r_i)) \geq \varepsilon/2$ for every i, every ball $B(x_i,r_i)$ has A_0-thin boundary with respect to $\mathcal{H}^3|_{\Gamma}$ for some $A_0 \geq 1$, and

$$\sum_i \mathcal{H}^3(B(x_i,r_i)\cap\Gamma) \gtrsim C.$$

Proof. Write $\Gamma_R := \Gamma \cap B(x,R)$. For $j \geq 1$, let

$$E_j := \{y \in \Gamma_R : \beta_{\mathrm{CG}}(B(y,s)) \geq \varepsilon \text{ for some } s \in (2^{-j}R, 2^{-j+1}R]\}.$$

Let $P_j \subset E_j$ be a maximal $2^{-j+5}R$-separated subset, and for each $y \in P_j$, choose a radius $s_y \in (2^{-j}R, 2^{-j+1}R]$ such that $\beta_{\mathrm{CG}}(B(y,s_y)) \geq \varepsilon$. Then, any pair of balls $B(y_1, 2s_{y_1}), B(y_2, 2s_{y_2})$ with distinct $y_1, y_2 \in P_j$ is disjoint. Further, since the sets $B(y, 2^{-j+7}R)\cap\Gamma$ with $y \in P_j$ cover E_j, we have

$$\mathcal{H}^3(E_j) \leq \sum_{y\in P_j} \mathcal{H}^3(B(y,2^{-j+7}R)\cap\Gamma) \sim \sum_{y\in P_j} (2^{-j}R)^3 \sim \sum_{y\in P_j} \mathcal{H}^3(B(y,s_y)\cap\Gamma).$$

Next, observe that for any $y \in \Gamma_R$, we have

$$\int_{2^{-j}R}^{2^{-j+1}R} \chi_{\{(y,s)\in\Gamma\times\mathbb{R}_+ : \beta_{\mathrm{CG}}(y,s))\geq\varepsilon\}}(y,s)\,\frac{ds}{s} \lesssim \chi_{E_j}(y),$$

since the left hand side is always bounded by $\lesssim 1$, and it is evidently zero for $y \notin E_j$. It follows that

$$\begin{aligned}\sum_{j=1}^{\infty} \mathcal{H}^3(E_j) &= \sum_{j=1}^{\infty} \int_{\Gamma_R} \chi_{E_j}(y)\, d\mathcal{H}^3(y) \\ &\gtrsim \sum_{j=1}^{\infty} \int_{\Gamma_R} \int_{2^{-j}R}^{2^{-j+1}R} \chi_{\{(y,s)\in\Gamma\times\mathbb{R}_+ : \beta_{\mathrm{CG}}(B(y,s))\geq\varepsilon\}}(y,s)\, d\mathcal{H}^3(y)\, \frac{ds}{s} \\ &= \int_0^R \int_{\Gamma_R} \chi_{\{(y,s)\in\Gamma\times\mathbb{R}_+ : \beta_{\mathrm{CG}}(B(y,s))\geq\varepsilon\}}(y,s)\, d\mathcal{H}^3(y)\, \frac{ds}{s} = C,\end{aligned}$$

and consequently

$$\sum_{j=1}^{\infty} \sum_{y\in P_j} \mathcal{H}^3(B(y,s_y)\cap\Gamma) \gtrsim C. \tag{4.31}$$

Next, observe that if $\beta_{\mathrm{CG}}(B(y,s)) \geq \varepsilon$ and $s' \in [s, \frac{9}{8}s]$, then $\beta_{\mathrm{CG}}(B(y,s')) \geq \varepsilon/2$. Then, for every ball $B(y,s_y)$, apply Lemma 4.19 with $\delta = 1/8$ and $\mu = \mathcal{H}^3|_\Gamma$ to find a radius $s_y' \in [s_y, \frac{9}{8}s_y]$ such that $B(y,s_y')$ has $A_{1/8}$-thin boundary with respect to $\mathcal{H}^3|_\Gamma$, and $\beta_{\mathrm{CG}}(B(y,s_y')) \geq \varepsilon/2$. Then the 3-regularity of Γ implies that (4.31) holds for the balls $B(y,s_y')$ as well. Finally, since the balls $B(y,s_y')$ with y in a fixed set P_j are disjoint, the collection

$$\{B(y,s_y') : y \in P_j, j \in \mathbb{N}\}$$

is pre-dyadic. Indeed, if $x \in \mathbb{H}$, $N > 1$ and $r > 0$ are given, there are clearly at most $\lesssim_N 1$ indices j such that $x \in B(y,s_y')$ and $r/N \leq \mathrm{diam}(B(y,s_y')) < r$ for some $y \in P_j$. The proof is complete. □

The next lemma "refines" a pre-dyadic family of balls into a dyadic one.

Lemma 4.33. *Let (X,d,μ) be a metric measure space, and let $\delta > 0$ and $A \geq 1$. Assume that $\mathcal{B}$ is a pre-dyadic family of balls with $\sup\{\mathrm{diam}(B) : B \in \mathcal{B}\} < \infty$, and with $\sum_{B\in\mathcal{B}} \mu(B) = C$. Moreover, assume that every ball $B \in \mathcal{B}$ contains a (possibly empty) family of disjoint sub-balls $\mathcal{F}(B)$ such that that $\mathrm{diam}(B') \geq \delta\, \mathrm{diam}(B)$ for $B' \in \mathcal{F}(B)$. Finally, assume that all the balls B' in the families $\mathcal{B}$ and $\mathcal{F}(B)$, $B \in \mathcal{B}$, have A-thin boundary, and the balls $B \in \mathcal{B}$ have the doubling property*

$$\mu(5B) \leq A\mu(B).$$

Then, there exists a dyadic sub-collection $\{B_j\} \subset \mathcal{B}$ such that

(i) *$\sum \mu(B_j) \gtrsim_{A,\delta} C$, and*
(ii) *if $B_j \subset B_{j'}$ for $j \neq j'$ then $B_j \cap B' = \emptyset$ for all $B' \in \mathcal{F}(B_{j'})$ or $B \subset B'$ for some $B' \in \mathcal{F}(B_{j'})$.*

Proof. Split the balls in $\mathcal{B}$ into families $\mathcal{B}_j := \{B \in \mathcal{B} : N^{j-1} \leq \mathrm{diam}(B) < N^j\}$, where $N = N_\delta \geq 2/\delta$. By the assumption of the family $\mathcal{B}$ being pre-dyadic, and the doubling property of the balls $B \in \mathcal{B}$, we can (by applying the $5r$-covering lemma to the balls in $\mathcal{B}_j$) choose a subfamily of $\mathcal{B}_j$ **disjoint** balls (still denoted by $\mathcal{B}_j$) with

$$\sum_j \sum_{B \in \mathcal{B}_j} \mu(B) \gtrsim_{A,\delta} C.$$

Next, observe that either

$$\sum_j \sum_{B \in \mathcal{B}_{2j}} \mu(B) \gtrsim_{A,\delta} C \quad \text{or} \quad \sum_j \sum_{B \in \mathcal{B}_{2j+1}} \mu(B) \gtrsim_{A,\delta} C.$$

We assume that the former option holds, and from now on we will only consider balls $B \in \mathcal{B}_{2j}$, $j \in \mathbb{Z}$. Note that if $j < i$, $B_1 \in \mathcal{B}_{2j}$, $B_2 \in \mathcal{B}_{2i}$ and $B_2' \in \mathcal{F}(B_2)$, then

$$\frac{\operatorname{diam}(B_1)}{\operatorname{diam}(B_2')} \leq \frac{N^{2j}}{\delta N^{2i-1}} = \frac{N^{2(j-i)+1}}{\delta} \leq \frac{N^{j-i}}{\delta} \leq 2^{j-i}, \tag{4.32}$$

recalling that $N \geq 2/\delta$.

Since $\sup\{\operatorname{diam}(B) : B \in \mathcal{B}\} < \infty$, the collection $\mathcal{B}_{2j}$ is empty for large enough j. Let j_0 be the index of the largest non-empty collection $\mathcal{B}_{2j_0}$, and set

$$\mathcal{G}_{j_0} := \mathcal{B}_{2j_0} \cup \bigcup_{B \in \mathcal{B}_{2j_0}} \mathcal{F}(B).$$

Inductively, for $j < i \leq j_0$, let $\mathcal{R}_j^i$ be the collection of balls in $\mathcal{B}_{2j}$, which meet the boundary of one of the balls in $\mathcal{G}_i$, let

$$\mathcal{G}_j^0 := \mathcal{B}_{2j} \setminus \bigcup_{j < i \leq j_0} \mathcal{R}_j^i,$$

and finally

$$\mathcal{G}_j := \mathcal{G}_j^0 \cup \bigcup_{B \in \mathcal{G}_j^0} \mathcal{F}(B).$$

Note that if $B \in \mathcal{R}_j^i$, $j < i \leq j_0$, then there is some ball $B' = B(y, r) \in \mathcal{G}_i$ such that

$$\begin{aligned} B \subset \{x : \operatorname{dist}(x, \partial B') \leq \operatorname{diam}(B)\} &\subset \{x : \operatorname{dist}(x, \partial B') \leq 2^{j-i} \operatorname{diam}(B')\} \\ &\subset B(y, 2r) \cap A(y, (1 - 2^{j-i-1})r, (1 + 2^{j-i+1})r). \end{aligned}$$

(The second inclusion follows from (4.32)). Consequently, by the disjointness of the balls in $\mathcal{B}_{2j} \supset \mathcal{R}_j^i$, and the A-thinness of boundaries of the balls in $\mathcal{B} \supset \mathcal{G}_i$, we infer that

$$\begin{aligned}\sum_{B \in \mathcal{R}_j^i} \mu(B) &= \mu\left(\bigcup_{B \in \mathcal{R}_j^i} B\right)\\ &\leq \mu\left(\bigcup_{B(y,r) \in \mathcal{G}_i} B(y,2r) \cap A(y,(1-2^{j-i-1})r,(1+2^{j-i+1})r)\right)\\ &\leq \sum_{B(y,r) \in \mathcal{G}_i} \mu(B(y,2r) \cap A(y,(1-2^{j-i-1})r,(1+2^{j-i+1})r))\\ &\lesssim A2^{j-i} \sum_{B' \in \mathcal{G}_i} \mu(B').\end{aligned}$$

Finally, since every ball in $\mathcal{B}_{2j}$ either belongs to one of the collections $\mathcal{R}_j^i$, with $j < i \leq j_0$, or to $\mathcal{G}_j^0 \subset \mathcal{G}_j$, we have

$$\begin{aligned}C \lesssim_{A,\delta} \sum_{j \leq j_0} \sum_{B \in \mathcal{B}_{2j}} \mu(B) &\leq \sum_{j \leq j_0} \left[\sum_{j < i \leq j_0} \sum_{B \in \mathcal{R}_j^i} \mu(B) + \sum_{B \in \mathcal{G}_j} \mu(B)\right]\\ &\lesssim A \sum_{j \leq j_0} \left[\sum_{j<i\leq j_0} 2^{j-i} \sum_{B' \in \mathcal{G}_i} \mu(B') + \sum_{B \in \mathcal{G}_j} \mu(B)\right]\\ &\sim A \sum_{j \leq j_0} \sum_{B \in \mathcal{G}_j} \mu(B). \qquad (4.33)\end{aligned}$$

Define

$$\mathcal{G} := \bigcup_{j \leq j_0} \mathcal{G}_j.$$

Inspecting the definition of $\mathcal{G}_j$, it is clear that $\mathcal{G}$ can indeed be written as

$$\mathcal{G} = \{B_j\} \cup \bigcup_j \mathcal{F}(B_j)$$

for a certain sub-collection $\{B_j\} \subset \mathcal{B}$. Since $\sum_{B' \in \mathcal{F}(B)} \mu(B') \leq \mu(B)$ for all balls $B \in \mathcal{B}$, we infer from (4.33) that $\sum \mu(B_j) \gtrsim_{A,\delta} C$, as required by condition (i). The fact that $\{B_j\}$ is a dyadic family of balls follows immediately from the inductive definition. In fact, the construction shows that even the family $\mathcal{G}$ is dyadic, and this combined with (4.32) implies condition (ii). The proof is complete. □

4.4.1. *Proof of the Theorem 4.16.* During the proof, we will abbreviate $\lesssim_{\varepsilon,L}$ and $\gtrsim_{\varepsilon,L}$ to simply $\lesssim$ and $\gtrsim$. We apply Lemma 4.32 to infer that if

$$\int_0^R \int_{\Gamma \cap B(x,R)} \chi_{\{(y,s) \in \Gamma \times \mathbb{R}_+ : \, \beta_{\mathrm{CG}}(B(y,s)) \geq \varepsilon\}}(y,s) \, d\mathcal{H}^3(y) \, \frac{ds}{s} =: C, \qquad (4.34)$$

then there exists a pre-dyadic family of balls $\{B_j\}$, contained in $B(x, 2R)$ and centred on Γ, with A_0-thin boundaries, with $\beta_{\mathrm{CG}}(B_j) \geq \varepsilon/2$, and such that

$$\sum_j \operatorname{diam}(B_j)^3 \gtrsim \sum_j \mathcal{H}^3(B_j \cap \Gamma) \gtrsim C. \tag{4.35}$$

For each such ball B_j, we use Proposition 4.20 to find another ball $\hat{B}_j \subset B_j$ with A_ε-thin boundary, such that $\operatorname{diam}(\hat{B}_j) \geq \delta_{\varepsilon,L} \operatorname{diam}(B)$,

$$\mathcal{L}^2(\pi_{\mathbb{W}}(\hat{B}_j \cap \Gamma)) \sim \operatorname{diam}(B_j)^3, \tag{4.36}$$

and

$$|\mathbb{E}_{\pi_{\mathbb{W}}(\hat{B}_j \cap \Gamma)} \nabla^\phi \phi - \mathbb{E}_{\pi_{\mathbb{W}}(B_j \cap \Gamma)} \nabla^\phi \phi| \geq \delta_{\varepsilon,L}. \tag{4.37}$$

Note that every ball in the family $\{B_j, \hat{B}_j\}$ has $\max\{A_0, A_\varepsilon\}$-thin boundary. Hence, using Lemma 4.33 on the metric space $(\mathbb{H}, d_{\mathbb{H}}, \mathcal{H}^3|_\Gamma)$, with $\mathcal{B} := \{B_j\}$ and $\mathcal{F}(B_j) := \{\hat{B}_j\}$, we find a dyadic sub-collection $\{B_{j_i}, \hat{B}_{j_i}\}$, which still satisfies $\sum_i \mathcal{H}^3(B_{j_i} \cap \Gamma) \gtrsim C$. To avoid the double indices, we assume that the family $\{B_j, \hat{B}_j\}$ is dyadic itself. Since $\pi_{\mathbb{W}}$ is injective on Γ, it follows that the family $\{A^j, B^j\}$ with $A^j := \pi_{\mathbb{W}}(\hat{B}_j \cap \Gamma)$ and $B^j := \pi_{\mathbb{W}}(B_j \cap \Gamma)$ is also dyadic. Moreover, by (ii) of Lemma 4.33, the family $\{A^j, B^j\}$ even satisfies slightly more, namely if $j \neq j'$, then

$$B^j \subset B^{j'} \quad \Longrightarrow \quad B^j \subset A^{j'} \text{ or } B^j \cap A^{j'} = \emptyset. \tag{4.38}$$

Holding that thought, we now make some remarks of more abstract nature. Given two sets $A, B \subset \mathbb{W}$ with $A \subset B$, let $V_{A,B}$ be the subspace of $L^2(\mathbb{W})$ consisting of those functions which are zero outside B, are constant on both $B \setminus A$ and A, and have integral zero. Let $\pi_{A,B} := \pi_{V_{A,B}}$ be the orthogonal projection onto this subspace. Finding an explicit formula for $\pi_{A,B}$ is simple, as $V_{A,B}$ is one-dimensional, and spanned by the unit vector

$$e_{A,B} := \left(\frac{|B \setminus A|}{|A||B|}\right)^{1/2} \chi_A - \left(\frac{|A|}{|B||B \setminus A|}\right)^{1/2} \chi_{B \setminus A}.$$

Here, and for the rest of the proof, we write $|U| := \mathcal{L}^2(U)$ for $U \subset \mathbb{W}$. It follows that $\pi_{A,B}(f) = (f \cdot e_{A,B}) e_{A,B}$, and in particular

$$\begin{aligned}
\|\pi_{A,B}(f)\|_2^2 = |f \cdot e_{A,B}|^2 &= \frac{|B \setminus A|}{|A||B|} \left(\int_A f \, d\mathcal{L}^2 - \frac{|A|}{|B \setminus A|} \int_{B \setminus A} f d\mathcal{L}^2\right)^2 \\
&= \frac{|B \setminus A|}{|A||B|} \left(\frac{|B|}{|B \setminus A|} \int_A f \, d\mathcal{L}^2 - \frac{|A|}{|B \setminus A|} \int_B f \, d\mathcal{L}^2\right)^2 \\
&= \frac{|A||B|}{|B \setminus A|} \left(\frac{1}{|A|} \int_A f \, d\mathcal{L}^2 - \frac{1}{|B|} \int_B f \, d\mathcal{L}^2\right)^2
\end{aligned} \tag{4.39}$$

We apply these observations to the subspaces V_{A^j,B^j} with the family $\{A^j, B^j\}$ defined above. Since this family is dyadic and satisfies (4.38), it is easy to check that the subspaces V_{A^j,B^j} are pairwise orthogonal. Now, let

$$f := \nabla^{\phi}\phi\chi_{\pi_{\mathbb{W}}(B(x,2R)\cap\Gamma)}$$

and recall that $\|f\|_{L^\infty} \lesssim_L 1$. Since $B^j \subset \pi_{\mathbb{W}}(B(x,2R)\cap\Gamma)$ for all $j \in \mathbb{N}$, it follows from Bessel's inequality, (4.39), (4.37), (4.36) and (4.35) (in this order) that

$$R^3 \gtrsim \|f\|_2^2 \geq \sum_j \|\pi_{A^j,B^j}(f)\|_2^2 = \sum_j \frac{|A^j||B^j|}{|B^j \setminus A^j|}\left[|\mathbb{E}_{A^j}(f) - \mathbb{E}_{B^j}(f)|^2\right]$$
$$\gtrsim \delta_{\varepsilon,L}^2 \sum_j |A^j| \sim \delta_{\varepsilon,L}^2 \sum_j \operatorname{diam}(B_j)^3 \gtrsim \delta_{\varepsilon,L}^2 C.$$

Recalling the definition of C from (4.34), this completes the proof of Theorem 4.16.

4.5. **Proof of the weak geometric lemma for intrinsic Lipschitz graphs.** Armed with Theorem 4.16, we are nearly ready to prove the WGL for intrinsic Lipschitz graphs (Theorem 4.2). First, we need a few short lemmas. The first one states that any intrinsic Lipschitz graph Γ has a "big flat piece" inside any ball $B(x,r)$ with $x \in \Gamma$:

Lemma 4.34. *For every $\varepsilon > 0$ and $L \geq 1$, there exists $\delta = \delta_{\varepsilon,L} > 0$ with the following property. If Γ is an intrinsic L-Lipschitz graph, $x \in \Gamma$ and $r > 0$, then there exists a ball $B(y,s) \subset B(x,r)$ with $y \in \Gamma$ and $s \geq \delta r$ such that $\beta_\Gamma(B(y,s)) \leq \varepsilon$.*

Proof. The proof is based on the compactness of intrinsic L-Lipschitz graphs, and the existence of tangent subgroups. We make a counter assumption that there exist $\varepsilon > 0$ and $L \geq 1$ with the following properties. For every $j \in \mathbb{N}$, there exists an intrinsic Lipschitz graph $\Gamma_j = \{w \cdot \phi_j(w) : w \in \mathbb{W}\}$, and a ball $B(x_j, r_j)$ with $x_j \in \Gamma_j$ such that

$$\beta_{\Gamma_j}(B(y, sr_j)) \geq \varepsilon$$

for all balls $B(y, sr_j) \subset B(x_j, r_j)$ with $y \in \Gamma$ and $s \geq 2^{-j}$. Without loss of generality (or recalling a similar reduction from the previous section), we may assume that $x_j = \phi_j(0) = 0$, and $r_j = 1$ for all $j \in \mathbb{N}$. Further, by the compactness results for intrinsic L-Lipschitz functions, already employed in the previous section, we may assume without loss of generality that the functions ϕ_j converge locally uniformly to an intrinsic L-Lipschitz function $\phi \colon \mathbb{W} \to \mathbb{V}$ with graph Γ. One can then easily check (emulating the argument for Lemma 4.26) that

$$\beta_\Gamma(B(y,s)) \geq \varepsilon$$

for any ball $B(y,s) \subset B(0,1)$ with $y \in \Gamma$. But this is absurd: the function ϕ is intrinsically differentiable at almost every point of $w \in \mathbb{W}$, which means precisely that $\mathcal{H}^3$-almost every point $y \in \Gamma$ has a vertical tangent subgroup to Γ (Theorem 4.15 in [28]). We record that a subgroup $\mathbb{G} \subset \mathbb{H}$ is a tangent group to Γ in $y \in \Gamma$ if

$$\lim_{h\to\infty} \delta_h(y^{-1}\cdot\Gamma) = \mathbb{G}$$

with respect to the Hausdorff convergence of compact subsets of $\mathbb{H}$. But this is evidently impossible for such $y \in \Gamma$ that $\beta_\Gamma(B(y,s)) \geq \varepsilon$ for all sufficiently small $s > 0$. All the points $y \in \Gamma \cap \operatorname{int} B(0,1)$ have this property by the previous discussion, and we have reached a contradiction. □

Recall that the graph of an entire intrinsic Lipschitz function $\phi : \mathbb{W} \to \mathbb{V}$ with constant gradient almost everywhere is a left translate of a vertical plane (Proposition 4.8). The next lemma, essentially a corollary of this fact, gives a substantial improvement over the previous lemma for constant gradient graphs.

Lemma 4.35. *For every $\varepsilon > 0$ and $L \geq 1$, there exists $\delta = \delta_{\varepsilon,L} > 0$ with the following property. Assume that $\Gamma = \{w \cdot \phi(w) : w \in \mathbb{W}\}$ is an intrinsic L-Lipschitz graph, and $x \in \Gamma$, $r > 0$ are such that $\phi \in \mathrm{Adm}_{\mathrm{CG},L}(B(x,r))$. Then, if $y \in \Gamma$, $0 < s < r$, and $B(y, s/\delta) \subset B(x, b_L r)$, we have*

$$\beta(B(y,s)) \leq \varepsilon.$$

Proof. We make a counter assumption: there exist $\varepsilon > 0$ and $L \geq 1$ such that the following holds for every $\delta_j = 1/j$, $j \in \mathbb{N}$. There exists an intrinsic L-Lipschitz graph $\Gamma_j = \{w \cdot \phi_j(w) : w \in \mathbb{W}\}$ and two balls $B(y_j, js_j) \subset B(x_j, b_L r_j)$ with $x_j, y_j \in \Gamma_j$, such that

$$\phi_j \in \mathrm{Adm}_{\mathrm{CG},L}(B(x_j,r_j)) \quad \text{and} \quad \beta_{\Gamma_j}(B(y_j,s_j)) \geq \varepsilon.$$

We may assume that $B(y_j,s_j) = B(0,1)$. Thus, $0 = \phi_j(0) \in \Gamma_j$ and $\beta_{\Gamma_j}(B(0,1)) \geq \varepsilon$, and $\phi_j \in \mathrm{Adm}_{\mathrm{CG},L}(B(z_j, r_j/s_j))$ for a certain point $z_j \in \Gamma_j$. The assumption $B(y_j, js_j) \subset B(x_j, b_L r_j)$ implies that the balls $B(z_j, b_L[r_j/s_j])$ eventually cover any ball $B(0,R)$, $R > 0$. We infer that a subsequence of the functions ϕ_j converge locally uniformly to an intrinsic L-Lipschitz function $\phi_R \colon \mathbb{W} \to \mathbb{V}$, which has a.e. constant gradient on $\pi_{\mathbb{W}}(B(0,R))$, and such that $\beta_{\Gamma_R}(B(0,1)) \geq \varepsilon$ (here Γ_R is the graph of ϕ_R). Finally, using a diagonal procedure, a further subsequence of the functions ϕ_j converges locally uniformly to an intrinsic L-Lipschitz function ϕ, which has a.e. constant gradient on $\mathbb{W}$, and $\beta_\Gamma(B(0,1)) \geq \varepsilon$. But the existence of such a ϕ was ruled out in Proposition 4.8, and we have reached a contradiction. The proof of the lemma is complete. □

Let Γ be an intrinsic Lipschitz graph, and let Δ be a system of David cubes on Γ; recall that every cube $Q \in \Delta$ contains a set of the form $B(z_Q, c\ell(Q)) \cap \Gamma$ for some point $z_Q \in Q$ and some constant $c > 0$. In this section, where our task is to prove the weak geometric lemma for intrinsic L-Lipschitz graphs, we write

$$B_Q := B(z_Q, C\ell(Q)),$$

where $C = C_L \geq 2$ is a constant depending on L only. The precise requirement for C is that $b_L C \geq 8$, where $b_L > 0$ is the constant from the definition of $\beta_{\mathrm{CG},\Gamma,L}$. With this notation for B_Q, we still write $\beta(Q) := \beta(B_Q)$, and $\beta_{\mathrm{CG},\Gamma,L}(Q) = \beta_{\mathrm{CG},\Gamma,L}(B_Q)$. The reason for choosing C_L so large is the following: if $Q, R \in \Delta$ with $Q \subset R$, and $\delta > 0$ is any constant, then

$$B(z_Q, C\ell(Q)/\delta) \subset (b_L/4)B_R := B(z_R, (b_L/4)C\ell(R)) \tag{4.40}$$

as soon as $\ell(Q)/\ell(R) \leq b_L\delta/8$. This is easy to check: if $\ell(Q)/\ell(R) \leq b_L\delta/8$ and $y \in B(z_Q, C\ell(Q)/\delta)$, then

$$d_{\mathbb{H}}(y, z_R) \leq d_{\mathbb{H}}(y, z_Q) + d_{\mathbb{H}}(z_Q, z_R) \leq (C\ell(Q)/(\delta\ell(R)) + 1)\ell(R) \leq b_L C\ell(R)/4,$$

We also mention that $C \geq 2$ implies

$$Q \subset R \quad \Longrightarrow \quad B_Q \subset B_R,$$

which is verified in the similar manner as above. In this section, we abbreviate $\mathcal{H}^3(A \cap \Gamma) =: |A|$. We choose to prove the weak geometric lemma in the following form: for any fixed cube $Q_0 \in \Delta$,

$$\sum_{Q \in \mathcal{B}_\varepsilon(Q_0)} |Q| \lesssim_{\varepsilon,L} |Q_0|, \tag{4.41}$$

where $\mathcal{B}_\varepsilon(Q_0) := \{Q \in \Delta(Q_0) : \beta(Q) > \varepsilon\}$; the integral formulation is an easy corollary. To this end, we note that the following inequality is a simple consequence (essentially a reformulation) of Theorem 4.16:

$$\sum_{Q \in \mathcal{B}_{\mathrm{CG},\varepsilon}(Q_0)} |Q| \lesssim_{\varepsilon,L} |Q_0|, \tag{4.42}$$

where $\mathcal{B}_{\mathrm{CG},\varepsilon}(Q_0) := \{Q \in \Delta(Q_0) : \beta_{\mathrm{CG}}(Q) > \varepsilon\}$.

Let $\eta > 0$ be a constant, which will depend only on ε and L. Now, by (4.42), the weak geometric lemma (4.41) will follow, if we are able to prove that

$$\sum_{Q \in \mathcal{B}_\varepsilon(Q_0) \setminus \mathcal{B}_{\mathrm{CG},\eta}(Q_0)} |Q| \lesssim_{\varepsilon,L} |Q_0|. \tag{4.43}$$

The following corollary of Lemma 4.35 will be crucial:

Corollary 4.36. *For every $\varepsilon > 0$ and $L \geq 1$, there exists $\eta = \eta_{\varepsilon,L}$ with the following property. Assume that $Q, R \in \Delta$ are cubes such that $Q \subset R$, $\beta(R) \leq \eta$, and $\beta_{\mathrm{CG}}(P) \leq \eta$ for all $P \in \Delta$ with $Q \subset P \subset R$. Then $\beta(Q) \leq \varepsilon$.*

Proof. Write $\delta := \delta_{\varepsilon/8,L}$, where $\delta_{\varepsilon/8,L} > 0$ is the constant from Lemma 4.35. Write $Q^{(n)}$ for the n^{th} dyadic ancestor of Q, with $Q^{(0)} := Q$. Let $n_{\varepsilon,L} \in \mathbb{N}$ be a large number depending on ε and L; we will eventually pick $n_{\varepsilon,L}$ first, and then require that η is sufficiently small depending on $n_{\varepsilon,L}$. If $R = Q^{(n)}$ for some $n < n_{\varepsilon,L}$, we have

$$\beta(Q) \leq \beta(R) \cdot \frac{\ell(R)}{\ell(Q)} = 2^n \beta(R) \leq 2^{n_{\varepsilon,L}}\eta \leq \varepsilon,$$

assuming $\eta \leq 2^{-n_{\varepsilon,L}}\varepsilon$. This is the first restriction we place on η.

Now, assume that $R = Q^{(n)}$ for some $n \geq n_{\varepsilon,L}$. It follows that $P := Q^{(n_{\varepsilon,L})}$ satisfies $Q \subset P \subset R$, hence $\beta_{\mathrm{CG}}(P) \leq \eta$. Recalling Remark 4.15 this means that there exists an intrinsic L-Lipschitz function $\psi\colon \mathbb{W} \to \mathbb{V}$ with graph Γ_P such that ψ has a.e. constant gradient on $\pi_{\mathbb{W}}(b_L B_P) = \pi_{\mathbb{W}}(B(z_P, b_L C\ell(P)))$, and

$$\sup_{y \in B_P \cap \Gamma} \frac{\operatorname{dist}_{\mathbb{H}}(y, \Gamma_P)}{C\ell(P)} \leq 2\eta. \tag{4.44}$$

Let x_P, x_Q be the nearest points to $z_P, z_Q \in B_P \cap \Gamma$ on Γ_P, so that

$$d_{\mathbb{H}}(x_P, z_P) \leq 2C\eta\ell(P) \quad \text{and} \quad d_{\mathbb{H}}(x_Q, z_Q) \leq 2C\eta\ell(P).$$

Write $r := C\ell(P)/2$, and observe that

$$(b_L/4)B_P \subset B(x_P, b_L r) \subset b_L B_P \tag{4.45}$$

for η small enough. Then, write $s := 2C\ell(Q)$, and observe that (4.40) gives the following chain of inclusions, assuming $2^{n_{\varepsilon,L}} = \ell(P)/\ell(Q)$ to be large enough, and $\eta_{\varepsilon,L} > 0$ to be small enough:

$$B(x_Q, s/\delta) \subset B(z_Q, 2s/\delta) \subset (b_L/4)B_P \subset B(x_P, b_L r). \tag{4.46}$$

From (4.45) we first conclude that ψ has a.e. constant gradient on $\pi_{\mathbb{W}}(B(x_P, b_L r))$, which means that

$$\psi \in \mathrm{Adm}_{\mathrm{CG},L}(B(x_P, r)).$$

Then, from (4.46) we infer that Lemma 4.35 is applicable to the ball $B(x_Q, s)$. Consequently,

$$\beta_{\Gamma_P}(B(x_Q, s)) \leq \varepsilon/8.$$

Thus, there exists a set of the form $z \cdot \mathbb{W}'$, where $z \in \mathbb{H}$ and $\mathbb{W}'$ is a vertical subgroup, such that $\operatorname{dist}_{\mathbb{H}}(y, z \cdot \mathbb{W}') \leq C(\varepsilon/2)\ell(Q)$ for every $y \in B(x_Q, s) \cap \Gamma_P$. Finally, recalling that we wish to prove $\beta_\Gamma(Q) \leq \varepsilon$, we fix

$$x \in B_Q \cap \Gamma \subset B_P \cap \Gamma.$$

By (4.44) we can pick $y \in \Gamma_P$ with $d_{\mathbb{H}}(x, y) \leq 2C\eta\ell(P)$. Then, for $\eta > 0$ small enough, we have $y \in B(x_Q, s) \cap \Gamma_P$, and $\operatorname{dist}_{\mathbb{H}}(x, z \cdot \mathbb{W}') \leq 2C\eta\ell(P) + C(\varepsilon/2)\ell(Q) \leq \varepsilon C\ell(Q)$. This proves that $\beta_\Gamma(Q) \leq \varepsilon$, as required. □

We also record a version of the "big flat piece lemma", Lemma 4.34, for cubes:

Lemma 4.37. *For every $\eta > 0$ and $L \geq 1$, there exists $\delta = \delta_{\eta,L} > 0$ with the following property. For any cube $R \in \Delta$, there exists a cube $Q \in \Delta(R)$ such that $\ell(Q) \geq \delta\ell(R)$ and $\beta(Q) \leq \eta$.*

Proof. Let $c := c_L > 0$ be a constant depending only on L, and let $\varepsilon := c_L\eta$. By Lemma 4.34 applied to the ball $B(z_R, c\ell(R))$ with $B(z_R, c\ell(Q)) \cap \Gamma \subset R$, there exists a point $y \in B(z_R, c\ell(R)) \cap \Gamma$ and a radius $s \gtrsim_{\varepsilon,L} \ell(R)$ such that $B(y, s) \subset B(z_R, c\ell(R))$ and $\beta_\Gamma(B(y, s)) \leq \varepsilon$. Now, it suffices to note that $B(y, s)$ contains a ball of the form B_Q for some $Q \subset R$ with $\ell(Q) \sim_L s$. Hence $\beta(Q) \lesssim_L \varepsilon = c_L\eta$. Choosing $c_L > 0$ small enough, the lemma follows. □

We are prepared to prove the weak geometric lemma. As discussed, it remains to verify (4.43). The proof is nearly verbatim the same as on the last pages of [44], but we record the details for completeness.

4.5.1. *Proof of the weak geometric lemma for intrinsic Lipschitz graphs.* We write $\mathcal{G}_\eta := \mathcal{G}_{\mathrm{CG},\eta}(Q_0) := \Delta(Q_0) \backslash \mathcal{B}_{\mathrm{CG},\eta}(Q_0)$, where $\eta = \eta_{\varepsilon,L}$ is the constant from Corollary 4.36. Then, we partition the collection $\mathcal{G}_\eta$ into *trees*. A family $\mathcal{T} \subset \Delta$ is called a tree, if

- $\mathcal{T}$ has a maximal element $Q(\mathcal{T})$, called the *root* of $\mathcal{T}$.
- If $Q, Q' \in \mathcal{T}$ with $Q \subset Q'$, and $Q \subset P \subset Q'$, then also $P \in \mathcal{T}$.
- If $Q \in \mathcal{T}$, then either all, or none, of the children of Q belong to $\mathcal{T}$.
- Those cubes $Q \in \mathcal{T}$ with no children in $\mathcal{T}$ are called the *stopping cubes* of $\mathcal{T}$, and they are denoted by $Stop(\mathcal{T})$.

We now partition $\mathcal{G}_\eta$ into trees by an inductive procedure. The roots of the initial trees are simply the maximal cubes in $\mathcal{G}_\eta$. Then, if $\mathcal{T}_1$ is one of these trees, and $Q \in \mathcal{T}_1$, we add all the children of Q into $\mathcal{T}_1$ if all of them belong to $\mathcal{G}_\eta$; if not, we declare that $Q \in Stop(\mathcal{T})$ and stop building $\mathcal{T}$ along this branch.

After the initial trees have been constructed, and in case some cubes in $\mathcal{G}_\eta$ still remain outside them, we repeat the previous procedure: we pick the maximal cubes in $\mathcal{G}_\eta$, which are not, yet, contained in a tree, and we declare these to be the roots of new trees. These trees are, then, constructed by the rule described above.

Iterating this algorithm produces a partition of $\mathcal{G}_\eta$ into a countable number of trees $\{\mathcal{T}_1, \mathcal{T}_2, \ldots\}$. If $Q(\mathcal{T}_i) = Q_0$ for some (unique) $i = i_0$, we set $Q(\mathcal{T}_i)' := Q(\mathcal{T}_i)$. In the opposite case, $Q(\mathcal{T}_i) \subsetneq Q_0$, it follows from the construction that either the parent or one of the siblings of $Q(\mathcal{T}_i)$, say Q' is in $\Delta(Q_0) \setminus \mathcal{G}_\eta = \mathcal{B}_{\mathrm{CG},\eta}(Q_0)$. In this case, we set $Q(\mathcal{T}_i)' := Q'$.

With this notation in hand, and observing that any cube Q' can only serve as $Q(\mathcal{T}_i)'$ for boundedly many indices i, we are prepared to show that the roots $Q(\mathcal{T}_i)$ satisfy a Carleson property:

$$\sum_i |Q(\mathcal{T}_i)| \lesssim \sum_i |Q(\mathcal{T}_i)'| \lesssim |Q_0| + \sum_{Q \in \mathcal{B}_{\mathrm{CG},\eta}(Q_0)} |Q| \lesssim_\varepsilon |Q_0|. \tag{4.47}$$

The last inequality is, of course, (4.42).

Next, recalling our objective (4.43), and writing $\mathcal{B}_\varepsilon := \mathcal{B}_\varepsilon(Q_0)$, we make the following natural splitting:

$$\sum_{Q \in \mathcal{B}_\varepsilon \cap \mathcal{G}_\eta} |Q| = \sum_i \sum_{Q \in \mathcal{B}_\varepsilon \cap \mathcal{T}_i} |Q|.$$

By (4.47), it remains to prove that

$$\sum_{Q \in \mathcal{B}_\varepsilon \cap \mathcal{T}_i} |Q| \lesssim_{\varepsilon,L} |Q(\mathcal{T}_i)| \tag{4.48}$$

for any fixed i. To this end, let $\mathcal{F}_i$ be the family of maximal cubes $Q \in \mathcal{T}_i$ with the property that $\beta(Q) \leq \eta$. By Corollary 4.36, if $Q' \in \mathcal{T}_i$ is contained in a cube in $\mathcal{F}_i$, then $\beta(Q') \leq \varepsilon$, and hence $Q' \notin \mathcal{B}_\varepsilon$. So, in fact the summation in (4.48) runs over at most those cubes which are not contained in any cube in $\mathcal{F}_i$; let those cubes be called $\mathcal{H}_i$.

Fix $R \in \mathcal{H}_i$, and let $Q \in \Delta(R)$ be the maximal cube with the property that $\beta(Q) \leq \eta$. Lemma 4.37 promises that there exists such a cube Q with $\ell(Q) \geq \delta_{\eta,L}\ell(R)$. If $Q \in \mathcal{T}_i$, then obviously $Q \in \mathcal{F}_i$, and we set $f(R) := Q$. If, on the other hand, Q already lies outside $\mathcal{T}_i$, then there is some cube $Q' \in Stop(\mathcal{T}_i)$ with $Q \subset Q' \subset R$, and we set $f(R) := Q'$. In either case $|f(R)| \gtrsim_{\eta,L} |R|$. Taking into account that both families $\mathcal{F}_i$ and $Stop(\mathcal{T}_i)$ consist of disjoint cubes, and that each cube $Q \in \mathcal{F}_i \cup Stop(\mathcal{T}_i)$ can only be assigned as $f(R)$ for boundedly many cubes R (with bounds depending only on η, L), we may conclude that

$$\sum_{R\in\mathcal{H}_i} |R| \lesssim_{\eta,L} \sum_{R\in\mathcal{H}_i} |f(R)| \lesssim_{\eta,L} \sum_{Q\in\mathcal{F}_i} |Q| + \sum_{Q\in Stop(\mathcal{T}_i)} |Q| \lesssim |Q(\mathcal{T}_i)|.$$

Since $\eta = \eta_{\varepsilon,L}$ only depends on ε and L, we have proven (4.48), and the weak geometric lemma (4.41) (Theorem 4.2) for intrinsic Lipschitz graphs.

4.5.2. *BPiLG implies WGL.* We will conclude the paper with a brief discussion of Theorem 4.1. The analogous result in the Euclidean setting is due to David and Semmes and it follows by Theorem 1.8 [18, Part IV, Chapter 1]. Before restating it in our setting, we record a definition.

Definition 4.38. Let $C_0 > 0$ and $\gamma : (0,1] \to (0,\infty)$. We define WGL$(C_0,\gamma)$ to be the collection of 3-regular sets $E \subset \mathbb{H}$ with regularity constant at most C_0 such that

$$\int_0^R \int_{E\cap B(x,R)} \chi_{\{(y,s)\in E\times\mathbb{R}_+ :\, \beta(B(y,s))\geq\varepsilon\}}(y,s)\, d\mathcal{H}^3(y)\, \frac{ds}{s} \leq \gamma(\varepsilon)\, R^3$$

for $\varepsilon > 0$, $x \in E$ and $R > 0$.

We can now provide the restatement of Theorem 1.8 [18, Part IV, Chapter 1] in the Heisenberg group.

Theorem 4.39. *Let $E \subset \mathbb{H}$ be a 3-regular set. Suppose that there exist $\theta > 0, C_0 > 0$ and $\gamma : (0,1] \to (0,\infty)$ such that for each $x \in E$ and $0 < R \leq \mathrm{diam}_{\mathbb{H}}(E)$ there exists $\tilde{E} \in$ WGL(C_0,γ) satisfying*

$$\mathcal{H}^3(E \cap \tilde{E} \cap B(x,R)) \geq \theta R^3.$$

Then E satisfies the WGL.

As it happens, the proof of Theorem 4.39 follows exactly as the proof of Theorem 1.8 [18, Part IV, Chapter 1], modulo notational changes. Therefore we skip the details. Now observe that Theorem 4.1 follows from Theorem 4.2 and Theorem 4.39, since the 3-regularity and WGL constants (for any fixed $\varepsilon > 0$) of an intrinsic L-Lipschitz graph are bounded above by constants depending only on ε and L.

References

[1] V. B. Adesi, F. Serra Cassano, D. Vittone: *The Bernstein problem for intrinsic graphs in Heisenberg groups and calibrations*, Calc. Var. PDE. **30** (2007), Issue 1, 17–49

[2] L. Ambrosio and B. Kirchheim: *Rectifiable sets in metric and Banach spaces*, Math. Ann. **318** (2000), 527–555

[3] L. Ambrosio, F. Serra Cassano, D. Vittone: *Intrinsic Regular Hypersurfaces in Heisenberg Groups*, J. Geom. Anal. **16** (2) (2006), 187–232

[4] L. Ambrosio, B. Kleiner, and E. Le Donne: *Rectifiability of Sets of Finite Perimeter in Carnot Groups: Existence of a Tangent Hyperplane*, J. Geom. Anal. **19** (2009), 509–540

[5] G. Arena and R. Serapioni: *Intrinsic regular submanifolds in Heisenberg groups are differentiable graphs*, Calc. Var. PDE. **35** (4) (2009), 517 – 536

[6] Z. M. Balogh, E. Durand Cartagena, K. Fässler, P. Mattila and J. T. Tyson: *The effect of projections on dimension in the Heisenberg group*, Rev. Mat. Iberoam. **29** (2013), no. 2, 381–432

[7] F. Bigolin: *Intrinsic regular hypersurfaces in Heisenberg groups and weak solutions of non linear first-order PDEs*, PhD thesis, Trento, (2009)

[8] F. Bigolin and F. Serra Cassano: *Distributional solutions of Burgers' equation and intrinsic regular graphs in Heisenberg groups*, J. Math. Anal. Appl., **366** (2) (2010), 561–568

[9] F. Bigolin, L. Caravenna and F. Serra Cassano: *Intrinsic Lipschitz graphs in Heisenberg groups and continuous solutions of a Balance equation*, Ann. Inst. H. Poincaré Anal. Non Linéaire **32** (5) (2015), 925–963

[10] A. Bonfiglioli, E. Lanconelli and F. Uguzzoni: *Stratified Lie Groups and Potential Theory for Their Sub-Laplacians*, Springer Monographs in Mathematics 2007

[11] L. Capogna, D. Danielli, S.D. Pauls and J. Tyson: *An Introduction to the Heisenberg Group and the Sub-Riemannian Isoperimetric Problem*, Progress in Mathematics (2007), Birkhäuser

[12] V. Chousionis and P. Mattila: *Singular integrals on self-similar sets and removability for Lipschitz harmonic functions in Heisenberg groups*, J. Reine Angew. Math. **691** (2014), 29–60

[13] V. Chousionis, V. Magnani and J. T. Tyson *Removable sets for Lipschitz harmonic functions on Carnot groups*, Calc. Var. PDE. **53** (2015), no. 3-4, 755–780

[14] G. Citti, M. Manfredini, A. Pinamonti and F. Serra Cassano: *Smooth approximation for intrinsic Lipschitz functions in the Heisenberg group*, Cal. Var. PDE **49** (3) (2014), 1279–1308

[15] G. David: *Morceaux de graphes lipschitziennes et intégrales singuliéres sur un surface*, Rev. Mat. Iberoamericana **4** (1988), 73–114

[16] G. David and S. Semmes: *Singular integrals and rectifiable sets in* $\mathbf{R}^n$*: Beyond Lipschitz graphs.* Astérisque **193**, 152 pp. (1991)

[17] G. David: *Wavelets and Singular Integrals on Curves and Surfaces*, Springer-Verlag Berlin Heidelberg 1991

[18] G. David and S. Semmes: *Analysis of and on uniformly rectifiable sets.* In: Mathematical Surveys and Monographs 38, Amer. Math. Soc., Providence, RI (1993)

[19] G. David and S. Semmes: *Quantitative rectifiability and Lipschitz mappings*, Trans. Amer. Math. Soc. **337** (2) (1993), 855–889

[20] G. David: *Unrectifiable* 1*-sets have vanishing analytic capacity*. Rev. Mat. Iberoamericana **14** (1998), no. 2, 369–479

[21] G. David and P. Mattila: *Removable sets for Lipschitz harmonic functions in the plane*. Rev. Mat. Iberoamericana **16**, (2000), no. 1, 137–215

[22] V. Eiderman and A. Volberg *Non-homogeneous harmonic analysis: 16 years of development*, Russ. Math. Surv. **68** 973, (2013)

[23] L. ERDŐS: *Notes on the Burgers equation*, available at https://www.mathematik.uni-muenchen.de/ lerdos/SS06/Pde/burg.pdf

[24] 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

[25] A. F. FILIPPOV: *Differential Equations with Discontinuous Righthand Sides: Control Systems*, Springer 1988

[26] B. FRANCHI, R. SERAPIONI AND F. SERRA CASSANO: *Rectifiability and perimeter in the Heisenberg group*, Math. Ann. **321** (2001), 479–531

[27] B. FRANCHI, R. SERAPIONI AND F. SERRA CASSANO: *Intrinsic Lipschitz graphs in Heisenberg groups*, J. Nonlinear Convex Anal. **7** (3) (2006), 423–441

[28] B. FRANCHI, R. SERAPIONI AND F. SERRA CASSANO: *Differentiability of Intrinsic Lipschitz Functions Within Heisenberg Groups*, J. Geom. Anal., **21** (2011), 1044–1084

[29] B. FRANCHI, M. MARCHI AND R. SERAPIONI: *Differentiability and Approximate Differentiability for Intrinsic Lipschitz Functions in Carnot Groups and a Rademacher Theorem*, Anal. Geom. Metr. Spaces **2** (1) (2014), 2299–3274

[30] B. FRANCHI AND R. SERAPIONI: *Intrinsic graphs within Carnot groups*, J. Geom. Anal. **26**, (2016), Issue 3, 1946–1994

[31] J. HEINONEN, P. KOSKELA, N. SHANMUGALINGAM AND J. TYSON: *Sobolev Spaces on Metric Measure Spaces*, Cambridge University Press, 2015

[32] P. JONES: *Lipschitz and bi-Lipschitz functions*, Rev. Mat. Iberoamericana **4** (1988), 115–122

[33] P. JONES: *Square functions, Cauchy integrals, analytic capacity, and harmonic measure,* Harmonic analysis and partial differential equations (El Escorial, 1987), Lecture Notes in Math., vol. 1384, Springer, Berlin, 1989, 24–68

[34] P. JONES: *Rectifiable sets and the traveling salesman problem.* Invent. Math. **102** (1990), 1–15

[35] N. JUILLET: *A counterexample for the geometric traveling salesman problem in the Heisenberg group.* Rev. Mat. Iberoam. **26** (2010), no. 3, 1035–1056

[36] 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

[37] S. LI AND R. SCHUL: *The traveling salesman problem in the Heisenberg group: upper curvature bound.* Trans. Amer. Math. Soc. **368**, (2016), no 7, 4585–4620

[38] P. MATTILA: *Geometry of sets and measures in Euclidean spaces*, Cambridge University Press, 1995

[39] P. MATTILA, R. SERAPIONI AND F. SERRA CASSANO: *Characterizations of intrinsic rectifiability in Heisenberg groups*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **5** Vol. IX (2010), 687–723

[40] F. NAZAROV, X. TOLSA AND A. VOLBERG: *On the uniform rectifiability of AD-regular measures with bounded Riesz transform operator: the case of codimension 1*, Acta Math. **213:2** (2014), 237–321

[41] F. NAZAROV, X. TOLSA AND A. VOLBERG: *The Riesz transform, rectifiability, and removability for Lipschitz harmonic functions.* Publ. Mat. *58:2* (2014), 517–532

[42] F. SERRA CASSANO: *Some topics of geometric measure theory in Carnot groups.* In Geometry, Analysis and Dynamics on sub-Riemannian Manifolds, vol. I. Edited by D. Barilari, U. Boscain, M. Sigalotti. EMS Ser.Lect. Notes in Math. EMS, Zurich, 2016

[43] X. TOLSA: *Analytic capacity, the Cauchy transform, and non-homogeneous Calderón-Zygmund theory*, volume 307 of *Progress in Mathematics*. Birkhäuser Verlag, Basel, 2014

[44] X. TOLSA: *Uniform measures and uniform rectifiability*, J. London Math. Soc. (2) **92** (2015), 1–18

UNIVERSITY OF CONNECTICUT, DEPARTMENT OF MATHEMATICS

UNIVERSITY OF JYVÄSKYLÄ, DEPARTMENT OF MATHEMATICS AND STATISTICS
(CURRENT AFFILIATION: UNIVERSITY OF FRIBOURG, DEPARTMENT OF MATHEMATICS)

UNIVERSITY OF HELSINKI, DEPT. OF MATHEMATICS AND STATISTICS
E-mail address: vasileios.chousionis@uconn.edu
E-mail address: katrin.faessler@unifr.ch
E-mail address: tuomas.orponen@helsinki.fi