

ISOMETRIC RIGIDITY OF THE WASSERSTEIN SPACE $\mathcal{W}_1(\mathbf{G})$ OVER CARNOT GROUPS

ZOLTÁN M. BALOGH, TAMÁS TITKOS, AND DÁNIEL VIROSZTEK

ABSTRACT. This paper aims to study isometries of the 1-Wasserstein space $\mathcal{W}_1(\mathbf{G})$ over Carnot groups endowed with horizontally strictly convex norms. Well-known examples of horizontally strictly convex norms on Carnot groups are the Heisenberg group $\mathbb{H}^n$ endowed with the Heisenberg-Korányi norm, or with the Naor-Lee norm; and H -type Iwasawa groups endowed with a Korányi-type norm. We prove that on a general Carnot group there always exists a horizontally strictly convex norm. The main result of the paper says that if $(\mathbf{G}, N_{\mathbf{G}})$ is a Carnot group where $N_{\mathbf{G}}$ is a horizontally strictly convex norm on $\mathbf{G}$, then the Wasserstein space $\mathcal{W}_1(\mathbf{G})$ is isometrically rigid. That is, for every isometry $\Phi : \mathcal{W}_1(\mathbf{G}) \rightarrow \mathcal{W}_1(\mathbf{G})$ there exists an isometry $\psi : \mathbf{G} \rightarrow \mathbf{G}$ such that $\Phi = \psi_{\#}$.

CONTENTS

1. Introduction: motivation and main results	1
2. Preliminaries on Wasserstein spaces and Carnot groups	3
3. Isometric rigidity of $\mathcal{W}_1(\mathbf{G})$: Proof of Theorem 1.1	8
4. Horizontal strict convexity of the Hebisch-Sikora norm	14
Acknowledgements	18
Declarations	19
References	19

1. INTRODUCTION: MOTIVATION AND MAIN RESULTS

As Hermann Weyl wrote in [40], whenever a structure-endowed entity is given, one can expect a deep insight into the constitution of that entity by determining its group of automorphisms. In this spirit, there has been recent activity to understand the structure of isometries of various metric spaces of measures, see e.g. [10, 11, 16, 18, 34]. One of the most

2020 *Mathematics Subject Classification.* 46E27 49Q22 54E40 .

Key words and phrases. Isometries, isometric embeddings, isometric rigidity, horizontal strict convexity, Wasserstein Space, Heisenberg group, Carnot group, Hebisch-Sikora norm, Heisenberg-Korányi norm, Naor-Lee norm.

Z. M. Balogh is supported by the Swiss National Science Foundation, Grant Nr. 200020_191978.

T. Titkos is supported by the Hungarian National Research, Development and Innovation Office (NKFIH) under grant agreements no. K134944 and no. Excellence_151232, and by the Momentum program of the Hungarian Academy of Sciences under grant agreement no. LP2021-15/2021.

D. Virosztek is supported by the Momentum program of the Hungarian Academy of Sciences under grant agreement no. LP2021-15/2021, by the Hungarian National Research, Development and Innovation Office (NKFIH) under grant agreement no. Excellence_151232, and partially supported by the ERC Synergy Grant No. 810115.

Email address of the corresponding author (Dániel Virosztek): virosztek.daniel@renyi.hu.

intensively studied metric structures of measures nowadays is the so-called p -Wasserstein space $\mathcal{W}_p(X)$, where the underlying space (X, ϱ) is a complete and separable metric space (see the precise definition later). The study of the geometric properties of these metric spaces is motivated by the theory of optimal mass transportation. Indeed, connections between the geometry of X and $\mathcal{W}_2(X)$ have been investigated by Lott and Villani in the groundbreaking paper [33]. The pioneering papers of Bertrand and Kloeckner [8, 9, 27, 28] started to explore fundamental geometric features of 2-Wasserstein spaces, including the description of complete geodesics and geodesic rays, determining their different type of ranks, and understanding the structure of their isometry group $\text{Isom}(\mathcal{W}_2(X))$. Isometries and isometric embeddings of p -Wasserstein spaces have been intensively studied in recent years for various underlying spaces X by Bertrand, Gehér, Kiss, Kloeckner, Santos-Rodriguez, and the authors; see [4, 8, 9, 19–23, 26–28, 37].

An important feature of p -Wasserstein spaces is that the underlying space X embeds isometrically into $\mathcal{W}_p(X)$, and similarly, the isometry group $\text{Isom}(X)$ of the underlying space X embeds into $\text{Isom}(\mathcal{W}_p(X))$ by push-forward of measures by isometries. In many cases, it is known that this embedding is surjective, i.e., $\text{Isom}(X)$ and $\text{Isom}(\mathcal{W}_p(X))$ are isomorphic. In such a case we call the Wasserstein space isometrically rigid. The value of p can play a distinctive role here. For example, in [20] it has been shown that $\mathcal{W}_p([0, 1])$ is isometrically rigid if and only if $p \neq 1$, and the results of [27] and [20] together imply that $\mathcal{W}_p(\mathbb{R})$ is isometrically rigid if and only if $p \neq 2$. These examples show that the isometric rigidity of $\mathcal{W}_1(X)$ does not imply the rigidity of $\mathcal{W}_p(X)$ for $p > 1$, and the reverse implication is also not true in general.

In our recent manuscript [4], we studied isometries and isometric embeddings of $\mathcal{W}_p(\mathbb{H}^n)$, the p -Wasserstein space over the Heisenberg group endowed with the Heisenberg-Korányi distance. We found that $\mathcal{W}_p(\mathbb{H}^n)$ is isometrically rigid if $p > 1$. Our method was based on a description of complete geodesics and geodesic rays (isometric embeddings of $\mathbb{R}$ and $\mathbb{R}_+$ into $\mathcal{W}_p(\mathbb{H}^n)$), which is not available if $p = 1$. Our aim is to overcome this difficulty and find a method that can be applied not only in the case of the Heisenberg group but to a much larger class of spaces. The main result of this paper is the following.

Theorem 1.1. *Let $(\mathbf{G}, N_{\mathbf{G}})$ be a Carnot group where $N_{\mathbf{G}}$ is a horizontally strictly convex norm on $\mathbf{G}$. Then the Wasserstein space $\mathcal{W}_1(\mathbf{G})$ is isometrically rigid. That is, for every isometry $\Phi : \mathcal{W}_1(\mathbf{G}) \rightarrow \mathcal{W}_1(\mathbf{G})$ there exists an isometry $\psi : \mathbf{G} \rightarrow \mathbf{G}$ such that $\Phi = \psi_{\#}$.*

The definition of horizontally strictly convex norms was introduced in the setting of the Heisenberg group in [2] in connection to isometric embeddings. This concept can easily be generalized to the setting of Carnot groups and it is done in the next section. We note here in advance that the horizontal strict convexity of a norm is a stronger property than the usual strict convexity, although the terminology may suggest it is weaker. Indeed, horizontal strict convexity of a norm means that the subadditivity of the norm is saturated only if the vectors involved are scalar multiples of each other *and horizontal* — see Definition 2.9. As an immediate consequence of this theorem, we get that the 1-Wasserstein space over the Heisenberg group endowed by the Heisenberg-Korányi norm or the Naor-Lee norm is isometrically rigid. Here, we use the fact that these norms are horizontally strictly convex, as proven in [2]. Other examples where our theorem applies are H -type Iwasawa groups endowed with a Korányi type norm (see [15]). The relevance of our result is underlined by the following existence theorem of horizontally strictly convex norms on general Carnot groups:

Theorem 1.2. *Let $(\mathbf{G}, \cdot)$ be a general Carnot group. Then there exists a horizontally strictly convex norm $N_{\mathbf{G}} : \mathbf{G} \rightarrow [0, \infty)$.*

The paper is organized as follows: in Section 2 we fix notation and recall preliminaries on Wasserstein spaces and on Carnot groups. In Section 3 we prove the main result: Theorem 1.1, and in Section 4 we prove Theorem 1.2.

2. PRELIMINARIES ON WASSERSTEIN SPACES AND CARNOT GROUPS

We start with notations that will be used in the sequel, for more details and references on p -Wasserstein spaces we refer the reader to any of the following comprehensive textbooks [1, 13, 36, 38, 39]. In this paper, we will focus on the case $p = 1$.

Let (X, ϱ) be a complete and separable metric space. We denote by $\mathcal{M}^+(X)$ the set of all non-negative finite Borel measures on X . For all $r > 0$ the symbol $\mathcal{M}_\varrho^r(X)$ stands for the set

$$\mathcal{M}_\varrho^r(X) = \left\{ \mu \in \mathcal{M}^+(X) \mid \mu(X) = r, \int_X \varrho(x, \hat{x}) \, d\mu(x) < \infty \text{ for some } \hat{x} \in X \right\}$$

Definition 2.1. *A non-negative Borel measure Π on $X \times X$ is called a coupling for $\mu, \nu \in \mathcal{M}_\varrho^r(X)$ if the marginals of Π are μ and ν , that is,*

$$\Pi(A \times X) = \mu(A) \quad \text{and} \quad \Pi(X \times B) = \nu(B)$$

for all Borel sets $A, B \subseteq X$. The set of all couplings is denoted by $C(\mu, \nu)$.

For two non-negative finite measures with the same total weight $\mu, \nu \in \mathcal{M}_\varrho^r(X)$ we can define the distance function $d_1^{(r)} : \mathcal{M}_\varrho^r(X) \times \mathcal{M}_\varrho^r(X) \rightarrow \mathbb{R}_{\geq 0}$ as follows:

$$d_1^{(r)}(\mu, \nu) := \inf_{\Pi \in C(\mu, \nu)} \iint_{X \times X} \varrho(x, y) \, d\Pi(x, y). \quad (2.1)$$

Definition 2.2. *The 1-Wasserstein space $\mathcal{W}_1(X)$ is the set $\mathcal{M}_\varrho^1(X)$ endowed with the so-called 1-Wasserstein distance $d_1^{(1)}$.*

It is known (see e.g. Theorem 1.5 in [1] with $c = \varrho$) that the infimum in (2.1) is, in fact, a minimum in this setting. Those couplings that minimize (2.1) are called optimal couplings, or optimal transport plans.

From now on, with a slight abuse of notation, we will write d_1 instead of $d_1^{(r)}$, no matter what the value of r is. However, it is important to emphasize that if we write $d_1(\mu, \nu)$ for some non-negative finite measures μ and ν , then it is always assumed that $\mu(X) = \nu(X)$. The advantage of this notation will be clear later in (2.3).

Next, we highlight two special features of the 1-Wasserstein metric. The first special feature is that there is a duality phenomenon, the so-called Kantorovich-Rubinstein duality. For given $\mu, \nu \in \mathcal{W}_1(X)$ the 1-Wasserstein distance can be computed as a supremum

$$d_1(\mu, \nu) = \sup_{\|f\|_L=1} \int_X f \, d(\nu - \mu) \quad (2.2)$$

where $\|f\|_L$ denotes the Lipschitz norm of f , i.e. the smallest constant C such that

$$|f(q) - f(q')| \leq C\varrho(q, q') \quad \text{for all } q, q' \in X.$$

The second special feature is a consequence of the duality: the 1-Wasserstein distance is translation invariant in the following sense:

$$d_1(\mu + \xi, \nu + \xi) = d_1(\mu, \nu). \quad (2.3)$$

For more details see Theorem 4.4 in [12].

As the terminology suggests, all notions introduced above are strongly related to the theory of optimal transportation. Indeed, for given sets A and B the quantity $\Pi(A \times B)$ is the weight of mass that is transported from A to B as μ is transported to ν along the transport plan Π .

The symbol $\text{supp} \mu$ stands for the support of $\mu \in \mathcal{P}(X)$. The set of Dirac measures will be denoted by $\Delta(X) = \{\delta_x \mid x \in X\}$. For $\mu, \nu \in \mathcal{W}_1(X)$ and $\lambda \in (0, 1)$ let us denote $M_\lambda(\mu, \nu)$ the metric λ -ratio set

$$M_\lambda(\mu, \nu) = \{\xi \in \mathcal{W}_1(X) \mid d_1(\mu, \xi) = \lambda d_1(\mu, \nu), d_1(\xi, \nu) = (1 - \lambda) d_1(\mu, \nu)\}.$$

Definition 2.3. *Given two metric spaces (Y, d_Y) and (Z, d_Z) , a map $f : Y \rightarrow Z$ is an isometric embedding if $d_Z(f(y), f(y')) = d_Y(y, y')$ for all $y, y' \in Y$. A self-map $\psi : Y \rightarrow Y$ is called an isometry if it is a surjective isometric embedding of Y onto itself. The symbol $\text{Isom}(\cdot)$ will stand for the group of all isometries.*

For an isometry $\psi \in \text{Isom}(X)$ the induced push-forward map is

$$\psi_\# : \mathcal{P}(X) \rightarrow \mathcal{P}(X); \quad (\psi_\#(\mu))(A) = \mu(\psi^{-1}[A])$$

for all Borel sets $A \subseteq X$ and $\mu \in \mathcal{P}(X)$, where $\psi^{-1}[A] = \{x \in X \mid \psi(x) \in A\}$. We call $\psi_\#(\mu)$ the *push-forward* of μ with ψ .

It is important to remark that $\mathcal{W}_1(X)$ contains an isometric copy of X . Indeed, since $C(\delta_q, \delta_{q'})$ has only one element (the Dirac measure $\delta_{(q, q')}$), we have that

$$d_1(\delta_q, \delta_{q'}) = \iint_{X \times X} \varrho(u, v) \, d\delta_{(q, q')}(u, v) = \varrho(q, q'),$$

and thus the embedding

$$X \hookrightarrow \mathcal{W}_p(X); \quad x \mapsto \delta_x \quad (2.4)$$

is distance preserving. Furthermore, the set of finitely supported measures (in other words, the collection of all finite convex combinations of Dirac measures)

$$\mathcal{F}(X) = \left\{ \sum_{j=1}^k \lambda_j \delta_{x_j} \mid k \in \mathbb{N}, x_j \in X, \lambda_j \geq 0 (1 \leq j \leq k), \sum_{j=1}^k \lambda_j = 1 \right\}$$

is dense in $\mathcal{W}_1(X)$, see e.g. Example 6.3 and Theorem 6.18 in [39]. Another important feature is that isometries of X appear in $\text{Isom}(\mathcal{W}_p(X))$ by means of a natural group homomorphism.

Definition 2.4. *Let $\#$ denote the map which embeds $\text{Isom}(X)$ into $\text{Isom}(\mathcal{W}_p(X))$*

$$\# : \text{Isom}(X) \rightarrow \text{Isom}(\mathcal{W}_p(X)), \quad \psi \mapsto \psi_\#. \quad (2.5)$$

Isometries that belong to the image of $\#$ are called trivial isometries. If $\#$ is surjective, i.e., if every isometry is trivial, then we say that $\mathcal{W}_p(X)$ is isometrically rigid.

Definition 2.5. A geodesic segment (or shortly: geodesic) is a curve $\gamma : [a, b] \rightarrow \mathcal{W}_1(X)$ such that

$$d_1(\gamma(t), \gamma(s)) = C|t - s|$$

for all $t, s \in [a, b]$. Note, that by reparametrising the curve γ we can always achieve that $C = 1$. Geodesics with $C = 1$ will be called unit-speed geodesics.

In the main result of this paper, the underlying metric space X will be a Carnot group. There are several recent studies on the metric structure of these groups (see e.g. [3, 5, 32]), as they serve as natural examples of sub-Riemannian geometries where the theory of optimal mass transport was recently developed [6, 14, 35]. In this paper, we refer to [7] for background and we shall be using the same notations. A Carnot group $(\mathbf{G}, \cdot)$ of step r is a connected, simply connected, nilpotent Lie group whose Lie algebra $\mathfrak{g}$ of left-invariant vector fields admits a stratification, i.e. there exist non-zero subspaces $\{V_j\}_1^r$ such that

$$\mathfrak{g} = \bigoplus_{i=1}^r V_i, \quad [V_1, V_j] = V_{j+1} \neq 0 \quad \text{for } j = 1, \dots, r-1, \quad [V_1, V_r] = 0.$$

We assume that a scalar product is given on $\mathfrak{g}$ for which the subspaces V_j are mutually orthogonal. The first layer V_1 of the Lie algebra plays a key role: its elements are called horizontal vectors. We denote by n_i the dimension of the vector space V_i . Let $X = \{X_1, X_2, \dots, X_n\}$ be an orthonormal basis of $\mathfrak{g}$ such that for every j (with $1 \leq j \leq r$) the set

$$\{X_i; \text{ with } n_0 + \dots + n_{j-1} < i \leq n_1 + \dots + n_j\}$$

is a basis for V_j (here we put $n_0 = 0$ and $N_j = n_0 + \dots + n_j$ and hence $N_r = n$).

The exponential map $\exp : \mathfrak{g} \rightarrow \mathbf{G}$ is defined by

$$\exp(x_1 X_1 + \dots + x_n X_n) = (x_1, \dots, x_n),$$

and it is a global diffeomorphism; we denote by $\xi = (\xi_1, \xi_2, \dots, \xi_r)$ the inverse of $\exp$, where $\xi_j : \mathbf{G} \rightarrow V_j$. This mapping defines the so-called coordinates of the first kind on $\mathbf{G}$, that is identified as a point set with $\mathbb{R}^n$. The group $\mathbf{G}$ is somewhat similar to the usual Euclidean space in the sense that it has a natural family of non-isotropic dilations and the group law has polynomial expressions. To be more precise, we note that the natural family of non-isotropic dilations for $\lambda > 0$ on $\mathfrak{g}$ associated with its grading is given by

$$\tilde{\delta}_\lambda(v_1 + v_2 + \dots + v_r) = \lambda v_1 + \lambda^2 v_2 + \dots + \lambda^r v_r,$$

if $v_i \in V_i$, $1 \leq i \leq r$. By means of the exponential map, one lifts these dilations to the family of the automorphisms of $\mathbf{G} \simeq \mathbb{R}^n$ $\delta_\lambda(x) = \exp(\tilde{\delta}_\lambda(\xi(x)))$, i.e.

$$\delta_\lambda(x) = (\lambda^{\sigma_1} x_1, \dots, \lambda^{\sigma_n} x_n),$$

with $x \in \mathbb{R}^n$ and $\lambda > 0$, where $\sigma_i = j$ for $N_{j-1} < i \leq N_j$.

The Baker–Campbell–Dynkin–Hausdorff formula gives the relation between the bracket relations in the Lie algebra $\mathfrak{g}$ and the expressions of the group operations on $\mathbf{G}$. Taking into consideration the symmetries encoded in the bracket relations in $\mathfrak{g}$, the group law will also possess some useful symmetry relations (see [25]): According to this, we can write an element in $\mathbf{G}$ as $p = (x, y)$ where $x \in \mathbb{R}^{n_1}$ and $y \in \mathbb{R}^{n_2 + \dots + n_s}$. In this notation the group operation in $\mathbf{G}$ can be written as

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 + x_2, y_1 + y_2 + P_1(x_1, x_2) + P_2((x_1, y_1), (x_2, y_2))),$$

where $P_1 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{n_2+\dots+n_s}$ and $P_2 : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n_2+\dots+n_s}$ are mappings with polynomial components in the variables x_1, x_2 resp. $(x_1, y_1), (x_2, y_2)$. We denote by $e = 0 \in \mathbb{R}^n$ the null element in $\mathbf{G}$ and note that the inverse of the element $q = (x_1, \dots, x_n)$ is simply $q^{-1} = -q = (-x_1, \dots, -x_n)$.

In what follows we turn to the metric structure of our Carnot group. We consider left-invariant homogeneous metrics on $\mathbf{G}$ defined as follows. The first step is to define a norm on $\mathbf{G}$ that is compatible with the group's law:

Definition 2.6. *Let $(\mathbf{G}, \cdot)$ be a Carnot group with neutral element e . We say that a map $N : \mathbf{G} \rightarrow \mathbb{R}_{\geq 0}$ is a norm on $\mathbf{G}$ if it satisfies*

- i) $N(q) = 0 \Leftrightarrow q = e, \forall q \in \mathbf{G},$
- ii) $N(q^{-1}) = N(q), \forall q \in \mathbf{G},$
- iii) $N(q \cdot q') \leq N(q) + N(q'), \forall q, q' \in \mathbf{G}.$

Definition 2.7. *Let $(\mathbf{G}, \cdot)$ be a Carnot group. A metric $d : \mathbf{G} \times \mathbf{G} \rightarrow \mathbb{R}_{\geq 0}$ is called left-invariant, if for every $q_o \in \mathbf{G}$, the map $L_{q_o} : (\mathbf{G}, d) \rightarrow (\mathbf{G}, d), q \mapsto q_o \cdot q$ is an isometry, that is, $d(q_o \cdot q, q_o \cdot q') = d(q, q')$, for all $q, q' \in \mathbf{G}$.*

Every norm $N : \mathbf{G} \rightarrow \mathbb{R}_{\geq 0}$ induces a left-invariant metric $d_N : \mathbf{G} \times \mathbf{G} \rightarrow \mathbb{R}_{\geq 0}$, by the formula $d_N(q, q') := N(q^{-1} \cdot q')$. Conversely, given a left-invariant metric $d : \mathbf{G} \times \mathbf{G} \rightarrow \mathbb{R}_+$, the map $N(q) = d(q, e)$ defines a norm on $\mathbf{G}$.

Definition 2.8. *A norm $N : \mathbf{G} \rightarrow \mathbb{R}_{\geq 0}$ on a Carnot group is called homogeneous if*

$$N(\delta_\lambda(p)) = \lambda N(p), \text{ for all } \lambda > 0, \text{ for all } p \in \mathbf{G}.$$

It is easy to see that a norm N on $\mathbf{G}$ is homogeneous if and only if its associated left-invariant metric is homogeneous in the sense that $d_N(\delta_\lambda(p), \delta_\lambda(q)) = \lambda d_N(p, q)$. Every left-invariant distance on $\mathbf{G}$ induced by a homogeneous norm is a homogeneous distance. From now on, we will use the expression *homogeneous distance on $\mathbf{G}$* to talk about the left-invariant metric induced by a homogeneous norm. The topology induced by any homogeneous distance on $\mathbf{G}$ coincides with the Euclidean topology on the base space $\mathbb{R}^n$.

In fact, once the homogeneous distances are known to be continuous with respect to the standard topology on $\mathbb{R}^n$ one can show by a standard argument the even stronger fact that all such norms are bi-Lipschitz equivalent. On the other hand, the metric structure induced by a homogeneous norm N on $\mathbf{G}$ is very different from $\mathbb{R}^n$ endowed with the Euclidean distance d_E . The two distances d_N and d_E are not bi-Lipschitz equivalent for any choice of homogeneous norm N on $\mathbf{G}$. For a comparison of the Hausdorff dimension of sets with respect to these two norms we refer to [5].

The simplest non-trivial Carnot group of step 2 is the Heisenberg group $\mathbb{H}^n \simeq \mathbb{R}^{2n+1}$ where the group operation is given by

$$q \cdot q' = (x, y, z) \cdot (x', y', z') = \left(x + x', y + y', z + z' + 2 \sum_{i=1}^n (x'_i y_i - x_i y'_i) \right),$$

and the action of the non-isotropic dilations is defined as $\delta_\lambda : \mathbb{H}^n \rightarrow \mathbb{H}^n, \lambda > 0$

$$\delta_\lambda(x, y, z) = (\lambda x, \lambda y, \lambda^2 z).$$

The left-invariant, homogeneous Heisenberg-Korányi metric d_H is defined by the homogeneous Korányi norm on $\mathbb{H}^n$

$$\|(x, y, z)\|_K = \left(\left(\sum_{i=1}^n (x_i^2 + y_i^2) \right)^2 + z^2 \right)^{\frac{1}{4}}.$$

Another interesting norm on $\mathbb{H}^n$ is the norm $N_{LN} : \mathbb{H}^n \rightarrow \mathbb{R}_+$ introduced by Lee and Naor in [31] and is given by:

$$N_{LN}(x, y, z) = \sqrt{\|(x, y, z)\|_K^2 + \|(x, y)\|_E^2},$$

where $\|\cdot\|_E$ is the usual Euclidean norm. We leave it as an exercise for the interested reader to verify that the above formula satisfies the conditions of Definition 2.6. We refer to [2] for more details, and many additional examples of homogeneous norms defined on the Heisenberg group. In that paper, the class of *horizontally strictly convex* norms has been introduced in the setting of the Heisenberg group. We can define this property in more general Carnot groups as follows:

Definition 2.9. *A homogeneous norm $N : \mathbf{G} \rightarrow \mathbb{R}_+$ is horizontally strictly convex if and only if*

$$q, q' \in \mathbf{G}, q, q' \neq e, \quad N(q \cdot q') = N(q) + N(q') \implies$$

q and q' lie on the same horizontal line through the origin, i.e., there exists $z \in \mathbb{R}^{n_1}$ such that

$$q = (sz, 0), q' = (s'z, 0), \quad \text{for some } s, s' \in \mathbb{R}.$$

It is easy to check that a homogenous norm N on $\mathbf{G}$ is horizontally strictly convex if and only if for all $q_1, q_2, q \in \mathbf{G}$ with $q_1 \neq q$ and $q_2 \neq q$ the equality

$$d_N(q_1, q_2) = d_N(q, q_1) + d_N(q, q_2)$$

implies that q_1, q_2 belong to the same horizontal line through q :

$$l = \{q \cdot (sz, 0) : s \in \mathbb{R}\},$$

for some non-zero $z \in \mathbb{R}^{n_1}$.

It was shown in [2] that the Korányi norm and the Lee-Naor norm in the Heisenberg group are horizontally strictly convex. On more general H-type Iwasawa groups of step two one can still define a norm of Korányi type that is horizontally strictly convex [15]. On the other hand, we note, that many homogeneous norms on Carnot groups are *not* horizontally strictly convex. A prominent example is the Carnot-Carathéodory norm coming from the sub-Riemannian Carnot-Carathéodory metric. Here *any pair of points* can be joined by a geodesic and therefore there are many points saturating the triangle inequality for any given pair of points. The same situation occurs in the case of more general sub-Finsler metrics [2].

In the following, we consider another class of norms on the Heisenberg group $\mathbb{H}^n$ that are *not* horizontally strictly convex. Let $1 \leq p \leq \infty$ and $0 < a \leq n^{-1/2}$, and consider the function

$$N_{p,a} : \mathbb{H}^n \rightarrow [0, \infty); (x, y, z) \mapsto \max \left\{ \|(x, y)\|_p, a\sqrt{|z|} \right\}$$

where $\|\cdot\|_p$ denotes the l^p norm on $\mathbb{R}^{2n}$. Then $N_{p,a}$ is a homogeneous norm on $\mathbb{H}^n$ — see [2, Example 5.4]. Now we check that for any $1 \leq p \leq \infty$ this norm is not horizontally

strictly convex. Indeed, for $|z| < 1/a$ consider the two points $q = (1, 0, \dots, 0, z)$, $q' = (1, 0, \dots, 0, -z) \in \mathbb{H}^n$. Then

$$2 = N_{p,a}(q \cdot q') = N_{p,a}(q) + N_{p,a}(q')$$

but clearly q and q' do not lie on the same horizontal line, and thus $N_{p,a}$ is not horizontally strictly convex.

Motivated by the statement of Theorem 1.1 we intend to find horizontally strictly convex norms in general Carnot groups. In this general setting, this could be a non-trivial matter. However, a prominent candidate seems to be the class of norms introduced by Hebisch and Sikora in [25]. Let us recall, that the Hebisch-Sikora norm on $N_{HS} : \mathbf{G} \rightarrow \mathbb{R}_+$ is defined as:

$$N_{HS}(q) := \inf\{t > 0 : \delta_{\frac{1}{t}}(q) \in B(0, r)\}, \quad (2.6)$$

where $B(0, r) \subseteq \mathbb{R}^n$ is the usual Euclidean ball in $\mathbb{R}^n$ centered at the origin, with radius r .

In [25] Hebisch and Sikora proved, that there exists $r_0 > 0$ such that for all $0 < r < r_0$ the function N_{HS} from (2.6) is a homogeneous norm. One can verify by a direct calculation that the above Lee-Naor norm is a particular instance of the Hebisch-Sikora norms in the setting of the Heisenberg group.

The Hebisch-Sikora norm N_{HS} was already proven to be useful in several studies: Lee and Naor [31] used N_{LN} on the Heisenberg group to give a counterexample to the Goemans-Linial conjecture. Le Donne and Rigot proved that in the setting of the Heisenberg group (see [29]) and also in the setting of Carnot groups of step two (see [30]) the Hebisch-Sikora norm N_{HS} satisfies the so-called Besicovitch covering property used in differentiation of measures. In a recent paper [32] Li used the norm N_{HS} in connection to the analyst's traveling salesman problem in Carnot groups. As we shall see in Theorem 4.1: the Hebisch-Sikora norm is horizontally strictly convex, verifying the statement of Theorem 1.2.

3. ISOMETRIC RIGIDITY OF $\mathcal{W}_1(\mathbf{G})$: PROOF OF THEOREM 1.1

First, let us introduce some notations. Let (X, ϱ) be a general metric space. For $q, q' \in X$ the symbol $[q, q']$ stands for the metric segment between q and q'

$$[q, q'] = \{r \in X \mid \varrho(q, r) + \varrho(r, q') = \varrho(q, q')\}.$$

The property that $[q, q'] = \{q, q'\}$ i.e. that the interior of the metric segment $[q, q']$ is empty will be important in the sequel. In fact, in the proof of Theorem 1.1 we need an even stronger property: we will use the symbol $q \sim q'$ for the relation

$$q \sim q' \iff \begin{cases} \forall r \in X : \varrho(q, q') = \varrho(q, r) + \varrho(r, q') \implies r \in \{q, q'\} \\ \forall r \in X : \varrho(r, q) = \varrho(r, q') + \varrho(q', q) \implies r \in \{q, q'\} \\ \forall r \in X : \varrho(q', r) = \varrho(q', q) + \varrho(q, r) \implies r \in \{q, q'\} \end{cases} \quad (3.1)$$

Observe that if X has at least two points then $q \sim q'$ implies $q \neq q'$. And obviously, $q \sim q'$ implies $[q, q'] = \{q, q'\}$.

Our aim in the next Lemma is to characterize those pairs of measures such that there is only one unit-speed geodesic connecting them in $\mathcal{W}_1(X)$, namely the linear interpolation. Although this might be known, or even folklore, we did not find it in the literature, so we include it for the sake of completeness.

Lemma 3.1. *Let (X, ϱ) be a complete and separable metric space, and let μ and ν be two different elements of $\mathcal{W}_1(X)$ with $d_1(\mu, \nu) := T > 0$. Then the following statements are equivalent.*

(i) The signed measure $\mu - \nu$ can be written as

$$\mu - \nu = c\delta_q - c\delta_{q'}$$

for some positive real number c and $q, q' \in X$ such that $[q, q'] = \{q, q'\}$. In other words, μ and ν can be written as

$$\mu = \eta + c\delta_q \quad \text{and} \quad \nu = \eta + c\delta_{q'}$$

for some non-negative measure η , $c > 0$, and $q, q' \in X$ such that $[q, q'] = \{q, q'\}$.

(ii) For every $\lambda \in (0, 1)$ the metric λ -ratio set is a singleton, that is,

$$M_\lambda(\mu, \nu) = \{(1 - \lambda)\mu + \lambda\nu\}.$$

(iii) There is only one unit-speed geodesic $\gamma : [0, T] \rightarrow \mathcal{W}_1(X)$ connecting μ and ν ($\gamma(0) = \mu$, $\gamma(T) = \nu$). That geodesic is the linear interpolation

$$\gamma(t) = \left(1 - \frac{t}{T}\right)\mu + \frac{t}{T}\nu \quad (t \in [0, T]). \quad (3.2)$$

Moreover, if (i) holds with the stronger assumption $q \sim q'$, then the unique unit-speed geodesic segment between μ and ν cannot be extended in any direction if and only if $\eta = 0$.

Proof. To prove (i) $\implies$ (ii) assume that $\mu = \eta + c\delta_q$, $\nu = \eta + c\delta_{q'}$, and $[q, q'] = \{q, q'\}$. If f is any 1-Lipschitz function (i.e., $|f(v) - f(u)| \leq \varrho(v, u)$ for all $u, v \in X$) then for any $\Pi \in C(\mu, \nu)$ we have

$$\begin{aligned} \iint_{X \times X} f(v) - f(u) \, d\Pi(u, v) &\leq \iint_{X \times X} |f(v) - f(u)| \, d\Pi(u, v) \\ &\leq \iint_{X \times X} \varrho(v, u) \, d\Pi(u, v). \end{aligned}$$

The left-hand side of the above inequality can be written as

$$\begin{aligned} \iint_{X \times X} f(v) - f(u) \, d\Pi(u, v) &= \int_X f(v) \, d\nu(v) - \int_X f(u) \, d\mu(u) \\ &= \int_X f \, d(\nu - \mu), \end{aligned} \quad (3.3)$$

so we see that if a function f solves the dual side and Π solves the primal side of the Kantorovich-Rubinstein duality then

$$f(v) - f(u) = \varrho(v, u) \quad \text{for } \Pi\text{-almost every pair } (u, v). \quad (3.4)$$

According to the assumption, $\nu - \mu = c(\delta_{q'} - \delta_q)$, so $f_q(z) := \varrho(q, z)$ solves the dual problem 2.2. Indeed, $\int_X f d(\nu - \mu) = c(f(q') - f(q))$ for any (Borel measurable) function $f : X \rightarrow \mathbb{R}$, and if f is furthermore 1-Lipschitz, that is, $\|f\|_L = 1$, then $c(f(q') - f(q)) \leq c\varrho(q, q')$. On the other hand, for the above-defined function f_q we clearly have $\int_X f_q d(\nu - \mu) = c\varrho(q, q')$. Consequently, for any optimal coupling $\Pi \in C(\mu, \nu)$ we know that if $(w, q') \in \text{supp}\Pi$ then

$$\varrho(q', w) = f_q(q') - f_q(w) = \varrho(q, q') - \varrho(q, w), \quad (3.5)$$

or equivalently, $\varrho(q, q') = \varrho(q, w) + \varrho(q', w)$. Since $[q, q'] = \{q, q'\}$, this implies that $w \in \{q, q'\}$. The assumption $\nu - \mu = c(\delta_{q'} - \delta_q)$ implies that the mass to be transported from $X \setminus \{q'\}$ to $\{q'\}$ by the optimal transport plan Π is at least c . That is, $\Pi((X \setminus \{q'\}) \times \{q'\}) \geq$

c , which in turn implies that $\Pi(\{q\} \times \{q'\}) = \Pi(\{(q, q')\}) \geq c$. However, by translation invariance we have $d_1(\mu, \nu) = c\varrho(q, q')$ and hence

$$c\varrho(q, q') = d_1(\mu, \nu) = \iint_{X \times X} \varrho(u, v) \, d\Pi(u, v) \geq \Pi(\{(q, q')\}) \varrho(q, q') \quad (3.6)$$

which implies that $\Pi(\{(q, q')\}) \leq c$ and consequently $\Pi(\{(q, q')\}) = c$. Moreover, (3.6) implies that

$$\iint_{(X \times X) \setminus \{(q, q')\}} \varrho(u, v) \, d\Pi(u, v) = 0.$$

Therefore, we conclude that there is only one optimal transport plan: the one which moves mass of magnitude c from q to q' and leaves everything else intact. In mathematical terms, the only optimal coupling Π_* is given by

$$\Pi_* := c\delta_{(q, q')} + (\text{Id} \times \text{Id})_{\#}\eta. \quad (3.7)$$

Now fix a $\lambda \in (0, 1)$ and take a $\xi \in M_\lambda(\mu, \nu)$. We have to show that $\xi = (1 - \lambda)\mu + \lambda\nu$. Since $d_1(\mu, \xi) + d_1(\xi, \nu) = d_1(\mu, \nu)$, we know on the one hand that if $\Pi_{\mu, \xi} \in C(\mu, \xi)$ and $\Pi_{\xi, \nu} \in C(\xi, \nu)$ are optimal couplings, then gluing them together we get an optimal coupling between μ and ν , and that optimal coupling must be Π_* . On the other hand,

$$\begin{aligned} d_1(\mu, \nu) &= \int_X f_q(z) \, d(\nu - \mu)(z) \\ &= \int_X f_q(z) \, d([\nu - \xi] + [\xi - \mu])(z) \\ &= \int_X f_q(z) \, d(\nu - \xi)(z) + \int_X f_q(z) \, d(\xi - \mu)(z) \\ &\leq d_1(\nu, \xi) + d_1(\xi, \mu) \\ &= d_1(\mu, \nu), \end{aligned} \quad (3.8)$$

which means that $f_q(z)$ solves the dual problem for both pairs (ν, ξ) and (ξ, μ) . The same calculation as in (3.4)-(3.5) ensures that if $(w, q') \in \text{supp}\Pi_{\xi, \nu}$ then $w \in \{q, q'\}$. Combining these two facts we get that ξ can be written for some $0 \leq \alpha \leq c$ as

$$\xi = \eta + \alpha\delta_q + (c - \alpha)\delta_{q'}. \quad (3.9)$$

Again, using the translation invariance of the distance, we have

$$\begin{aligned} d_1(\mu, \xi) &= d_1(\eta + c\delta_q, \eta + \alpha\delta_q + (c - \alpha)\delta_{q'}) \\ &= d_1(c\delta_q, \alpha\delta_q + (c - \alpha)\delta_{q'}) = (c - \alpha)\varrho(q, q'). \end{aligned} \quad (3.10)$$

This implies that

$$(1 - \lambda)c\varrho(q, q') = (1 - \lambda)d_1(\mu, \nu) = (c - \alpha)\varrho(q, q'), \quad (3.11)$$

or equivalently, $\alpha = \lambda c$. Now we see from (3.9) that

$$\xi = \eta + \lambda c\delta_q + (c - \lambda c)\delta_{q'} = (1 - \lambda)\mu + \lambda\nu.$$

To prove (ii) $\implies$ (iii), assume by contradiction that there exist at least two different unit-speed geodesics γ_1 and γ_2 , connecting μ and ν . Since the two curves are not identical, there exists at least one $t \in (0, T)$ such that $\gamma_1(t) \neq \gamma_2(t)$. For that $t \in (0, T)$ we have

$\{\gamma_1(t), \gamma_2(t)\} \subseteq M_{1-\frac{t}{T}}(\mu, \nu)$, a contradiction.

To prove (iii) $\implies$ (i), assume by contradiction that (i) does not hold. This is the case if either (a) $\mu - \nu = c\delta_q - c\delta_{q'}$ for some $q, q' \in X$ and $c > 0$, but $[q, q'] \neq \{q, q'\}$; or (b) $\mu - \nu \neq c\delta_q - c\delta_{q'}$ for any $q, q' \in X$ and $c > 0$.

Let us handle case (a) first. Since $\mu - \nu = c\delta_q - c\delta_{q'}$ by assumption, we can write μ and ν as $\mu = \eta + c\delta_q$ and $\nu = \eta + c\delta_{q'}$ for some non-negative measure η and $c > 0$. On the other hand, since $[q, q'] \neq \{q, q'\}$, we can choose a $z \notin \{q, q'\}$ such that $\varrho(q, q') = \varrho(q, z) + \varrho(z, q')$. Using this $z \in X$ we can construct a unit-speed geodesic $\tilde{\gamma} : [0, T] \rightarrow \mathcal{W}_1(X)$ between μ and ν , passing through $\xi := \eta + c\delta_z$. Set $T' = d_1(\mu, \xi)$ and consider

$$\tilde{\gamma}(t) := \begin{cases} \eta + \left(1 - \frac{t}{T'}\right)c\delta_q + \frac{t}{T'}c\delta_z & \text{if } t \in [0, T'], \\ \eta + \left(1 - \frac{t-T'}{T-T'}\right)c\delta_z + \frac{t-T'}{T-T'}c\delta_{q'} & \text{if } t \in (T', T]. \end{cases}$$

An elementary calculation shows that $\tilde{\gamma}$ is a unit-speed geodesic between μ and ν (such that $\tilde{\gamma}(T') = \xi$), which is of course different from the linear interpolation (3.2).

To handle case (b) assume that the signed measure $\mu - \nu$ is not of the form $c\delta_q - c\delta_{q'}$ for any $c > 0$ and $q, q' \in X$. Let us denote by $\mu' = (\mu - \nu)_+$ and $\nu' = (\mu - \nu)_-$ its positive and negative part, respectively. Note, that since $(\mu - \nu)(X) = 0$ we have that $\mu'(X) = \nu'(X)$. According to the assumption, $\text{supp}\mu' \cap \text{supp}\nu' = \emptyset$ and at least one of them (by symmetry, let's say μ') is not a constant multiple of a Dirac measure. Then we can write μ' as a sum $\mu' = \mu'_1 + \mu'_2$ such that μ'_1 is not a constant multiple of μ'_2 . Indeed, choose a set S such that $0 < \mu'(S) < \mu'(X)$ and define μ'_1 and μ'_2 as follows: $\mu'_1(A) := \mu'(A \cap S)$ and $\mu'_2(A) := \mu'(A \setminus S)$ for all Borel sets $A \subseteq X$. This decomposition of μ' also induces a decomposition $\nu' = \nu'_1 + \nu'_2$. To see this, let us choose an optimal coupling $\Pi \in C(\mu', \nu')$ and set $\nu'_1(B) = \Pi(S \times B)$, $\nu'_2(B) := \Pi((X \setminus S) \times B)$ for all Borel sets $B \subseteq X$. These decompositions guarantee that we can construct at least two different unit-speed geodesics between μ and ν . (In words, we can transport the whole μ'_1 first, and then μ'_2 , or vice versa.) Set $T' = d_1(\mu'_1, \nu'_1)$, and observe that

$$T = d_1(\mu, \nu) = d_1(\mu'_1, \nu'_1) + d_1(\mu'_2, \nu'_2)$$

according to (2.3) and the optimality of Π . Now we leave it to the interested reader, to verify that the curves γ_1 and γ_2 defined by

$$\gamma_1(t) = \begin{cases} \eta + \mu'_2 + \left(1 - \frac{t}{T'}\right)\mu'_1 + \frac{t}{T'}\nu'_1 & \text{if } t \in [0, T'], \\ \eta + \nu'_1 + \left(1 - \frac{t-T'}{T-T'}\right)\mu'_2 + \frac{t-T'}{T-T'}\nu'_2 & \text{if } t \in (T', T] \end{cases}$$

and

$$\gamma_2(t) = \begin{cases} \eta + \mu'_1 + \left(1 - \frac{t}{T-T'}\right)\mu'_2 + \frac{t}{T-T'}\nu'_2 & \text{if } t \in [0, T-T'], \\ \eta + \nu'_2 + \left(1 - \frac{t-(T-T')}{T'}\right)\mu'_1 + \frac{t-(T-T')}{T'}\nu'_1 & \text{if } t \in (T-T', T] \end{cases}$$

are two different unit-speed geodesics connecting μ and ν , a contradiction. This finishes the proof of (i) $\iff$ (ii) $\iff$ (iii).

Finally, in order to prove the second part of the statement, we have to show that if (i) holds with the stronger assumption $q \sim q'$, then the unique unit-speed geodesic segment between μ and ν cannot be extended in any direction if and only if $\eta = 0$.

Assume first by contradiction that $\mu = \eta + c\delta_q$ and $\nu = \eta + c\delta_{q'}$ where $q \sim q'$, but $\eta \neq 0$. Then at least one of μ and ν is not a Dirac measure. By symmetry, we can assume that

ν is not a Dirac measure. Then the geodesic can be extended towards $\delta_{q'}$. Indeed, set $T' = d_1(\mu, \delta_{q'})$, and observe that the curve

$$\tilde{\gamma}(t) = \begin{cases} \eta + c\left((1 - \frac{t}{T})\delta_q + \frac{t}{T}\delta_{q'}\right) & \text{if } t \in [0, T], \\ c\delta_{q'} + \left(1 - \frac{t-T}{T'-T}\right)\eta + \frac{t-T}{T'-T}(1-c)\delta_{q'} & \text{if } t \in (T, T'] \end{cases}$$

is a unit-speed geodesic segment connecting μ and $\delta_{q'}$, such that $\tilde{\gamma}(T) = \nu$, a contradiction. To see the converse direction, assume that $\eta = 0$, or equivalently, $\mu = \delta_q$ and $\nu = \delta_{q'}$ where $q \sim q'$. Then we have $T = d_1(\mu, \nu) = d_1(\delta_q, \delta_{q'}) = \varrho(q, q')$, and the unique geodesic segment between δ_q and $\delta_{q'}$ is

$$\gamma : [0, T] \rightarrow \mathcal{W}_1(X), \quad \gamma(t) = \left(1 - \frac{t}{T}\right)\delta_q + \frac{t}{T}\delta_{q'}. \quad (3.12)$$

We have to show that γ cannot be extended in any direction. That is, for all $\varepsilon > 0$ there is no unit-speed geodesic segment $\gamma_1 : [-\varepsilon, T] \rightarrow \mathcal{W}_1(\mathbb{H}^n)$ such that $\gamma_1|_{[0, T]} = \gamma$, and similarly, there is no geodesic segment $\gamma_2 : [0, T + \varepsilon] \rightarrow \mathcal{W}_1(\mathbb{H}^n)$ such that $\gamma_2|_{[0, T]} = \gamma$. We present the argument for γ_1 , the case of γ_2 is similar. Set $\xi := \gamma_1(-\varepsilon)$ and recall that the triple $\gamma_1(-\varepsilon) = \xi$, $\gamma_1(0) = \delta_q$, $\gamma_1(T) = \delta_{q'}$ satisfies the triangle inequality with equality:

$$d_1(\xi, \delta_{q'}) = d_1(\xi, \delta_q) + d_1(\delta_q, \delta_{q'}). \quad (3.13)$$

Using that $1 = \int_X 1 \, d\xi(r)$, we have $d_1(\delta_q, \delta_{q'}) = \varrho(q, q') = \int_X \varrho(q, q') \, d\xi(r)$, and thus we can write the above equality as

$$\int_X \varrho(r, q') \, d\xi(r) = \int_X \varrho(r, q) \, d\xi(r) + \int_X \varrho(q, q') \, d\xi(r), \quad (3.14)$$

or equivalently,

$$\int_X (\varrho(r, q) + \varrho(q, q') - \varrho(r, q')) \, d\xi(r) = 0. \quad (3.15)$$

As $\varrho(r, q) + \varrho(q, q') - \varrho(r, q') \geq 0$ by the triangle inequality, this means that for ξ -almost every $r \in X$ we have

$$\varrho(r, q') = \varrho(r, q) + \varrho(q, q'). \quad (3.16)$$

According to $q \sim q'$ this implies that $\text{supp}(\xi) \subseteq \{q, q'\}$, and thus $\gamma_1(-\varepsilon) = \gamma_1(t)$ for some $t \in [0, T]$, a contradiction. $\square$

Now we are ready to prove the main result of the paper.

Proof of Theorem 1.1. Let G be a Carnot group with a horizontally strictly convex norm $N_{\mathbf{G}}$, and denote by d the metric induced by $N_{\mathbf{G}}$. We want to prove that the Wasserstein space $\mathcal{W}_1(\mathbf{G})$ over $(\mathbf{G}, d)$ is isometrically rigid. The first step is to show that the Φ -image of pairs of the form $\{\delta_q, \delta_{q'}\}$ where $q \sim q'$ is again a pair of Dirac measures.

We note here that for elements of a Carnot group with a horizontally strictly convex norm, the relation $q \sim q'$ defined by (3.1) is equivalent to the a priori weaker condition $[q, q'] = \{q, q'\}$. Moreover, these equivalent conditions have clear and transparent characterization: $q \sim q'$ if and only if $[q, q'] = \{q, q'\}$ if and only if $q^{-1} \cdot q'$ is *not* horizontal. Indeed, if $q^{-1} \cdot q'$ is horizontal, then

$$\{q, q'\} \subsetneq \{q \cdot \delta_\alpha(q^{-1} \cdot q') \mid \alpha \in [0, 1]\} \subseteq [q, q']$$

as

$$d(q, q \cdot \delta_\alpha(q^{-1} \cdot q')) = \alpha N_{\mathbf{G}}(q^{-1} \cdot q')$$

and

$$d(q', q \cdot \delta_\alpha(q^{-1} \cdot q')) = (1 - \alpha)N_{\mathbf{G}}(q^{-1} \cdot q'),$$

and consequently, $q \approx q'$. On the other hand, if $q \approx q'$ then there exists an $r \notin \{q, q'\}$ such that r, q and q' saturate the triangle inequality — that is, one of the three equations in (3.1) holds true. We can assume without too much loss of generality that $d(r, q') = d(r, q) + d(q, q')$ as the two other cases are very similar. In this case we get that $N_{\mathbf{G}}(r^{-1} \cdot q) + N_{\mathbf{G}}(q^{-1} \cdot q') = N_{\mathbf{G}}(r^{-1} \cdot q)$. By the horizontal strict convexity of $N_{\mathbf{G}}$, see Definition 2.9, we immediately get that $r^{-1} \cdot q$ and $q^{-1} \cdot q'$ lie on the same horizontal line through the origin, in particular, $q^{-1} \cdot q'$ is a horizontal vector.

Step 1. Let us take a pair $\{\delta_q, \delta_{q'}\}$ where $q \sim q'$, and set $T := d(q, q') > 0$. According to Lemma 3.1, there is only one unit-speed geodesic $\gamma : [0, T] \rightarrow \mathcal{W}_1(\mathbf{G})$ connecting δ_q and $\delta_{q'}$, and this geodesic cannot be extended in any directions. Since an isometry is a bijection, we obtain that there is only one unit-speed geodesic between $\Phi(\delta_q)$ and $\Phi(\delta_{q'})$ as well. Again, using Lemma 3.1, this means that there exists a non-negative measure η , a constant $0 < c \leq 1$ and two points $r, r' \in X$ satisfying $[r, r'] = \{r, r'\}$ such that $\Phi(\delta_q) = \eta + c\delta_r$ and $\Phi(\delta_{q'}) = \eta + c\delta_{r'}$. In fact, since the norm is horizontally strictly convex, we also have $r \sim r'$, and thus we know that $\eta = 0$. Indeed, assume indirectly that $\eta \neq 0$. Then the geodesic segment between $\Phi(\delta_q)$ and $\Phi(\delta_{q'})$ could be extended according to the second part of Lemma 3.1, and the pre-image of this geodesic segment would be an extension of γ , a contradiction.

Step 2. In the second step we show that there exists an isometry $\psi : \mathbf{G} \rightarrow \mathbf{G}$ such that $\Phi(\delta_q) = \delta_{\psi(q)}$ for all $q \in \mathbf{G}$. Since $N_{\mathbf{G}}$ is horizontally strictly convex, with a non-horizontal perturbation for any $q \in \mathbf{G}$ we can find a $q' \in \mathbf{G}$ such that $q \sim q'$. Since Φ is bijective, this implies that Φ maps the set of Dirac masses onto itself. Indeed, it is clear by Step 1 that $\Phi(\delta_q)$ is a Dirac mass for any $q \in \mathbf{G}$, and for any $r \in \mathbf{G}$ there is a $\mu \in \mathcal{W}_1(\mathbf{G})$ such that $\Phi(\mu) = \delta_r$ as Φ is surjective. Moreover, μ is a Dirac mass because $\mu = \Phi^{-1}(\delta_r)$, and Φ^{-1} is an isometry of $\mathcal{W}_1(\mathbf{G})$ which sends Dirac masses to Dirac masses by Step 1. Setting ψ as $\Phi(\delta_q) := \delta_{\psi(q)}$ we obtain a map $\psi : \mathbf{G} \rightarrow \mathbf{G}$, which is in fact an isometry of $\mathbf{G}$. Indeed, the distance preserving property follows from the fact that $d_1(\delta_q, \delta_r) = d(q, r)$ for all $q, r \in \mathbf{G}$, and the surjectivity of ψ follows from the fact that Φ maps the set of Dirac masses onto itself.

Step 3. Since $\Phi = \psi_{\#}$ if and only if $\Phi \circ (\psi_{\#})^{-1} = \text{Id}$, we can assume without loss of generality that $\psi(q) = q$ for all $q \in \mathbf{G}$. Our aim under this assumption will be to show that $\Phi(\mu) = \mu$ for a dense subset of $\mathcal{W}_1(\mathbf{G})$. As Φ is continuous, this will finish the proof. Let us introduce the notation

$$\mathcal{F}_{\sim}(\mathbf{G}) := \left\{ \sum_{j=1}^n \lambda_j \delta_{q_j} \mid n \in \mathbb{N}, q_j \sim q_l \text{ for all } 1 \leq j < l \leq n \right\} \quad (3.17)$$

and show that $\mathcal{F}_{\sim}(\mathbf{G})$ is dense in $\mathcal{W}_1(\mathbf{G})$. Here again, we can take advantage of the horizontal strict convexity of the norm. Namely that for all finite subsets $\{q_i\}_{i \in I} \subseteq \mathbf{G}$ ($|I| < \infty$) and for all $\varepsilon > 0$ with small non-horizontal perturbations taken step-by-step we can find a finite subset $\{q_i^\varepsilon\}_{i \in I} \subseteq \mathbf{G}$ such that

$$d(q_i, q_i^\varepsilon) < \varepsilon \quad \text{for all } i \in I, \quad (3.18)$$

and

$$q_i^\varepsilon \sim q_j^\varepsilon \quad \text{for all } i, j \in I. \quad (3.19)$$

Indeed, these assumptions imply that for any $\tilde{\varepsilon} > 0$ and for any $\xi = \sum_{j=1}^n \lambda_j \delta_{q_j}$ we can find an $\varepsilon > 0$ and $\{q_i^\varepsilon\}_{i=1}^n$ satisfying (3.18) and (3.19) such that the measure $\xi' := \sum_{j=1}^n \lambda_j \delta_{q_j^\varepsilon}$ belongs to $\mathcal{F}_\sim(\mathbf{G})$ and $d_1(\xi, \xi') < \tilde{\varepsilon}$. Therefore, $\mathcal{F}_\sim(\mathbf{G})$ is dense in $\mathcal{F}(\mathbf{G})$ and thus in $\mathcal{W}_1(\mathbf{G})$.

What remains to be proven is that Φ fixes all elements of $\mathcal{F}_\sim(\mathbf{G})$. We already know that Φ fixes all Dirac measures. Fix a $k \in \mathbb{N}$ and suppose we proved the statement for all elements of $\mathcal{F}_\sim(\mathbf{G})$ that are supported on at most k points. Take a measure $\xi = \sum_{j=1}^{k+1} \lambda_j \delta_{q_j} \in \mathcal{F}_\sim(\mathbf{G})$ supported on $k+1$ points. We can write ξ as

$$\xi = (1 - \lambda) \cdot \xi_1 + \lambda \cdot \xi_2$$

with some $0 < \lambda < 1$ and $\xi_1, \xi_2 \in \mathcal{F}_\sim(\mathbf{G})$ whose supports are sets with k elements. Indeed,

$$\xi_1 := \sum_{j=1}^{k-1} \lambda_j \delta_{q_j} + (\lambda_k + \lambda_{k+1}) \delta_{q_k}, \quad \xi_2 := \sum_{j=1}^{k-1} \lambda_j \delta_{q_j} + (\lambda_k + \lambda_{k+1}) \delta_{q_{k+1}},$$

and $\lambda := \frac{\lambda_{k+1}}{\lambda_k + \lambda_{k+1}}$. Observe that ξ_1 and ξ_2 satisfies the conditions of Lemma 3.1 (i) with

$$\eta = \sum_{j=1}^{k-1} \lambda_j \delta_{q_j}, \quad c = \lambda_k + \lambda_{k+1}, \quad q = q_k, \quad q' = q_{k+1}.$$

Using Lemma 3.1 we know on the one hand that $M_\lambda(\xi_1, \xi_2) = \{\xi\}$, on the other hand the set $M_\lambda(\Phi(\xi_1), \Phi(\xi_2))$ is just the Φ -image of $M_\lambda(\xi_1, \xi_2)$. Combining these observations with $\Phi(\xi_1) = \xi_1$ and $\Phi(\xi_2) = \xi_2$ we get $\{\xi\} = M_\lambda(\xi_1, \xi_2) = M_\lambda(\Phi(\xi_1), \Phi(\xi_2)) = \{\Phi(\xi)\}$. The proof is complete.

We close this section with a brief remark on the proof.

Remark 3.1. *The scheme of the proof of Theorem 1.1 works in the following abstract setting as well. Assume that (X, ϱ) is metric space where (1) for every $q \in X$ there exists a $q' \in X$ such that $q \sim q'$, and (2) for all finite subsets $\{q_i\}_{i \in I} \subseteq X$ ($|I| < \infty$) and for all $\varepsilon > 0$ there exists a $\{q_i^\varepsilon\}_{i \in I} \subseteq X$ such that (2/a) $\varrho(q_i, q_i^\varepsilon) < \varepsilon$ for all $i \in I$, and (2/b) $q_i^\varepsilon \sim q_j^\varepsilon$ for all $i, j \in I$. A prominent example of such a metric space is the unit sphere $\mathbb{S}^n$ endowed with the distance inherited from the ambient Euclidean space $\mathbb{R}^{n+1}$ (we note that the isometric rigidity of the quadratic Wasserstein space $\mathcal{W}_2(\mathbb{S}^n)$ has recently been proven in [17] for any finite dimension n .) Indeed, we have $x \sim x'$ for any distinct points $x, x' \in \mathbb{S}^n$. However, it is important to remark that these conditions are only sufficient, and are very far from being necessary. For example, the Wasserstein space $\mathcal{W}_1(\mathbb{R})$ is isometrically rigid, but there are no points $q, q' \in \mathbb{R}$ such that $q \sim q'$.*

4. HORIZONTAL STRICT CONVEXITY OF THE HEBISCH-SIKORA NORM

In this section, we prove Theorem 1.2 by verifying the horizontally strict convexity property of the Hebisch-Sikora norm. We state this result as the following:

Theorem 4.1. *Let $(\mathbf{G}, \cdot)$ be a Carnot group. Then there exists $r_0 > 0$ such that for all $0 < r < r_0$ the function N_{HS} defined by (2.6) is a horizontally strictly convex norm.*

Proof. Let us note, that the homogeneity of d_{HS} as well as properties i) and ii) in Definition 2.6 follow immediately from the definition of d_{HS} . The triangle inequality has been proven

in [25], and our non-trivial task here is to characterize the case of equality in the triangle inequality iii)

$$N_{HS}(q_1 \cdot q_2) \leq N_{HS}(q_1) + N_{HS}(q_2), \quad (4.1)$$

in order to prove the horizontal strict convexity of this norm. This needs a refined analysis. Although our proof is an adaptation of the main idea of the proof of Theorem 2 from [25], we need to make a more careful analysis such that in addition to the proof of the triangle inequality we can study the equality case in (4.1) to prove horizontal strict convexity.

Let us notice first that by the homogeneity of N_{HS} (4.1) can be written equivalently as

$$N_{HS}\left(\delta_{\frac{1}{N_{HS}(q_1)+N_{HS}(q_2)}}(q_1 \cdot q_2)\right) \leq 1.$$

Using the relation $\delta_{t_1 t_2} = \delta_{t_1} \delta_{t_2}$ we can write the above inequality as

$$N_{HS}\left(\delta_{\frac{N_{HS}(q_1)}{N_{HS}(q_1)+N_{HS}(q_2)}}\left(\delta_{\frac{1}{N_{HS}(q_1)}}(q_1)\right) \cdot \delta_{\frac{N_{HS}(q_2)}{N_{HS}(q_1)+N_{HS}(q_2)}}\left(\delta_{\frac{1}{N_{HS}(q_2)}}(q_2)\right)\right) \leq 1.$$

Denoting by $p_1 = \delta_{\frac{1}{N_{HS}(q_1)}}(q_1)$ and $p_2 = \delta_{\frac{1}{N_{HS}(q_2)}}(q_2)$ and $t = \frac{N_{HS}(q_1)}{N_{HS}(q_1)+N_{HS}(q_2)}$ we see that $t \in (0, 1)$ and

$$N_{HS}(p_1) = N_{HS}(p_2) = 1 \quad (4.2)$$

and the above inequality is written equivalently as

$$N_{HS}(\delta_t(p_1) \cdot \delta_{1-t}(p_2)) \leq 1. \quad (4.3)$$

In conclusion we see that (4.1) becomes equivalent to (4.3) subjected to (4.2). Moreover, the equality case in (4.1) becomes equivalent to the equality case in (4.3) (subject to (4.2)).

Let us notice that by definition the unit ball centered in $0 \in \mathbf{G}$ with respect to d_{HS} coincides as a set with the Euclidean ball $B(0, r) \subset \mathbb{R}^n$ and also, that the unit sphere with respect to d_{HS} coincides with $\partial B(0, r)$. By this observation (4.3) (subjected to (4.2)) can be formulated equivalently as follows: We have to prove that there exists $r_0 > 0$ such that for $0 < r < r_0$, it follows that if $p_1, p_2 \in \mathbb{R}^n$ satisfy the property that $\|p_1\|^2 = \|p_2\|^2 = r^2$ then

$$\|\delta_t(p_1) \cdot \delta_{1-t}(p_2)\|^2 \leq r^2, \quad (4.4)$$

for any $t \in (0, 1)$. Moreover, we have to show that equality in (4.4) can happen if and only if

$$p_1 = p_2 = (x, 0) \text{ for some } x \in \mathbb{R}^{n_1}, \|x\|^2 = r^2. \quad (4.5)$$

It is clear that the inequality (4.4) is saturated if p_1 and p_2 are given as in (4.5). From now on, we work on proving that equality in (4.4) can happen only if (4.5) holds. Here, and in the following we use the same notation $\|\cdot\|$ for the Euclidean norm of vectors in various Euclidean spaces of (possibly) different dimensions.

To continue the proof let us recall from the introduction that we can write an element $p \in \mathbf{G}$ as $p = (x, y)$ where $x \in \mathbb{R}^{n_1}$ and $y \in \mathbb{R}^{n_2+\dots+n_s}$. In this notation the group operation in $\mathbf{G}$ can be written as

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 + x_2, y_1 + y_2 + P_1(x_1, x_2) + P_2((x_1, y_1), (x_2, y_2))),$$

where $P_1 : \mathbb{R}^{n_1} \times \mathbb{R}^{n_1} \rightarrow \mathbb{R}^{n_2+\dots+n_s}$ and $P_2 : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^{n_2+\dots+n_s}$ are mappings with polynomial components in the variables x_1, x_2 resp. $(x_1, y_1), (x_2, y_2)$. We recall from [25] that this representation uses the correspondence between the bracket relations in the Lie algebra and the group operation due to the Baker–Campbell–Dynkin–Hausdorff formula. Accordingly, P_1 corresponds to the contribution of vector fields in the first layer of the Lie

algebra and P_2 corresponds to those vector fields with at least one of them coming from higher layers. Based on this, we have two crucial properties (see [25]) of P_1 and P_2 that we are going to use in the following: There exists a constant $C_1 > 0$ such that

$$\|P_1(x_1, x_2)\| \leq C_1 \|x_1\| \cdot \|x_2\| \cdot \left\| \frac{x_1}{\|x_1\|} - \frac{x_2}{\|x_2\|} \right\| \quad (4.6)$$

for all $x_1, x_2 \in \mathbb{R}^{n_1}$ with $\|x_1\| \leq 1$ and $\|x_2\| \leq 1$. There exists a constant $C_2 > 0$ such that

$$\|P_2(x_1, y_1), (x_2, y_2)\| \leq C_2 (\|x_1\| \cdot \|y_2\| + \|x_2\| \cdot \|y_1\| + \|y_1\| \cdot \|y_2\|), \quad (4.7)$$

for all $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^n$ with $\|(x_1, y_1)\| \leq 1$ and $\|(x_2, y_2)\| \leq 1$.

The following Euclidean norm identity for non-zero vectors u and v will also be used in the course of the proof:

$$\|u + v\|^2 = (\|u\| + \|v\|)^2 - \|u\| \cdot \|v\| \cdot \left\| \frac{u}{\|u\|} - \frac{v}{\|v\|} \right\|^2. \quad (4.8)$$

After this preparation we are ready to address the proof of (4.4). Namely, we shall prove that if $r > 0$ is small enough, then for given $p_1 = (x_1, y_1)$ and $p_2 = (x_2, y_2)$ satisfying

$$\|x_1\|^2 + \|y_1\|^2 = \|x_2\|^2 + \|y_2\|^2 = r^2$$

it follows that

$$\|\delta_t(x_1, y_1) \cdot \delta_{1-t}(x_2, y_2)\|^2 \leq r^2. \quad (4.9)$$

Moreover, we shall show that equality in (4.9) holds if and only if $p_1 = p_2 = (x, 0)$ for some $x \in \mathbb{R}^{n_1}$ with $\|x\|^2 = r^2$.

Using the expression of the group law with P_1 and P_2 we can write (4.9) in the form

$$\|(tx_1 + (1-t)x_2, \delta_t(y_1) + \delta_{1-t}(y_2) + P_2(\delta_t(x_1, y_1), \delta_{1-t}(x_2, y_2)) + P_1(tx_1, (1-t)x_2)\|^2 \leq r^2.$$

Introducing the notation

$$w = \delta_t(y_1) + \delta_{1-t}(y_2) + P_2(\delta_t(x_1, y_1), \delta_{1-t}(x_2, y_2)),$$

the above relation can be written as

$$\|tx_1 + (1-t)x_2\|^2 + \|P_1(tx_1, (1-t)x_2)\|^2 + 2\langle w, P_1(tx_1, (1-t)x_2) \rangle + \|w\|^2 \leq r^2, \quad (4.10)$$

where the notation $\langle \cdot, \cdot \rangle$ stands for the usual scalar product. By Young's inequality we have

$$2\langle w, P_1(tx_1, (1-t)x_2) \rangle = t(1-t)2\langle w, P_1(x_1, x_2) \rangle \leq t(1-t)(\|w\|^2 + \|P_1(x_1, x_2)\|^2).$$

Therefore the inequality (4.10) (and thus (4.9)) will follow from

$$\|tx_1 + (1-t)x_2\|^2 + t(1-t)(1+t(1-t))\|P_1(x_1, x_2)\|^2 + (1+t(1-t))\|w\|^2 \leq r^2. \quad (4.11)$$

We are now ready to split (4.11) into two distinct parts:

$$\begin{aligned} \|tx_1 + (1-t)x_2\|^2 + t(1-t)(1+t(1-t))\|P_1(x_1, x_2)\|^2 &\leq (t\|x_1\| + (1-t)\|x_2\|)^2 \\ &\leq t\|x_1\|^2 + (1-t)\|x_2\|^2 \end{aligned} \quad (4.12)$$

and

$$(1+t(1-t))\|w\|^2 \leq (t\|y_1\| + (1-t)\|y_2\|)^2 \leq t\|y_1\|^2 + (1-t)\|y_2\|^2. \quad (4.13)$$

Let us assume for the moment that both (4.12) and (4.13) hold true. By adding up these two inequalities we obtain the following estimate for the left-hand side of (4.11)

$$\text{LHS}(4.11) \leq t(\|x_1\|^2 + \|y_1\|^2) + (1-t)(\|x_2\|^2 + \|y_2\|^2) = r^2, \quad (4.14)$$

proving the inequality (4.11) which also concludes the proof of (4.9).

Let us also note that in order to have equality in (4.11) (or (4.9)) we need that equalities should hold true throughout (4.12) and also (4.13).

In what follows we shall prove both (4.12) and (4.13) and also analyse the equality cases. In doing so, relations (4.6), (4.7) and (4.8) will be used in an essential way. Let us start with the first inequality in (4.12). (The second inequality in (4.12) follows trivially from Jensen's inequality applied to the convex function $x \rightarrow \|x\|^2$.) We can write the first inequality in (4.12) as

$$t(1-t)(1+t(1-t))\|P_1(x_1, x_2)\|^2 \leq (t\|x_1\| + (1-t)\|x_2\|)^2 - \|tx_1 + (1-t)x_2\|^2.$$

Applying (4.8) to $u = tx_1$ and $v = (1-t)x_2$ we see that the above inequality is equivalent to

$$(1+t(1-t))\|P_1(x_1, x_2)\|^2 \leq \|x_1\| \cdot \|x_2\| \cdot \left\| \frac{x_1}{\|x_1\|} - \frac{x_2}{\|x_2\|} \right\|^2.$$

Using (4.6) we can estimate

$$\|P_1(x_1, x_2)\|^2 \leq C_1^2 \|x_1\|^2 \cdot \|x_2\|^2 \cdot \left\| \frac{x_1}{\|x_1\|} - \frac{x_2}{\|x_2\|} \right\|^2.$$

Now choose $r < \frac{2}{\sqrt{5}C_1}$. Then for all $t \in (0, 1)$ we have

$$(1+t(1-t))C_1^2 r^2 < 1.$$

Since $\|x_1\| \leq r$ and $\|x_2\| \leq r$ this implies

$$(1+t(1-t))C_1^2 \|x_1\| \cdot \|x_2\| < 1,$$

and thus the first part of the inequality (4.12).

Note that the necessary and sufficient condition to have equality in this inequality is

$$\frac{x_1}{\|x_1\|} - \frac{x_2}{\|x_2\|} = 0. \quad (4.15)$$

On the other hand, the necessary and sufficient condition for the equality in the second inequality of (4.12) is $\|x_1\| = \|x_2\|$. We conclude, that the necessary and sufficient condition for equalities in (4.12) is $x_1 = x_2$.

We can turn now to the proof of (4.13). Here the estimate (4.7) will be used in an essential manner. We note, that combining the triangle inequality, (4.7) and inequalities $\|\delta_t y_1\| \leq t^2 \|y_1\|$ and $\|\delta_{1-t} y_2\| \leq (1-t)^2 \|y_2\|$ we obtain the following estimate for $\|w\|$:

$$\|w\| \leq t^2 \|y_1\| + (1-t)^2 \|y_2\| + t(1-t)C_2(\|x_1\| \cdot \|y_2\| + \|y_1\| \cdot \|x_2\| + \|y_1\| \cdot \|y_2\|).$$

Choosing $r < \frac{1}{4C_2}$ and using $\|x_1\| \leq r$, $\|x_2\| \leq r$ and also $\|y_1\| \leq r$, $\|y_2\| \leq r$ we obtain

$$\begin{aligned} \|w\| &\leq t^2 \|y_1\| + (1-t)^2 \|y_2\| + \frac{t(1-t)}{2}(\|y_1\| + \|y_2\|) \\ &= t\|y_1\| + (1-t)\|y_2\| - \frac{t(1-t)}{2}(\|y_1\| + \|y_2\|). \end{aligned}$$

This inequality can also be written as

$$\|w\| + \frac{t(1-t)}{2}(\|y_1\| + \|y_2\|) \leq t\|y_1\| + (1-t)\|y_2\| \quad (4.16)$$

Observe, that in particular, this implies that

$$\|w\| \leq \|y_1\| + \|y_2\|.$$

Thus we infer that

$$\begin{aligned}
\|w\|^2(1+t(1-t)) &= \|w\|^2 + t(1-t)\|w\|^2 \leq \|w\|^2 + t(1-t)\|w\|(\|y_1\| + \|y_2\|) \\
&\leq \|w\|^2 + t(1-t)\|w\|(\|y_1\| + \|y_2\|) + \left(\frac{t(1-t)}{2}(\|y_1\| + \|y_2\|)\right)^2 \\
&= \left(\|w\| + \frac{t(1-t)}{2}(\|y_1\| + \|y_2\|)\right)^2 \leq (t\|y_1\| + (1-t)\|y_2\|)^2.
\end{aligned}$$

Note that in the last inequality of the above chain we made use of (4.16). This proves the second inequality (4.13). Let us observe that in order to have equality in (4.13) it is necessary to have equalities all throughout in the above chain of estimates. In particular, we must have that

$$\frac{t(1-t)}{2}(\|y_1\| + \|y_2\|) = 0,$$

as the square of this term has been added to obtain the second inequality. But the only way that this can happen is that $y_1 = y_2 = 0$.

Let us now recall, that the two necessary conditions that we obtained for the equality cases were $x_1 = x_2$ and $y_1 = y_2 = 0$. Denoting by $x = x_1 = x_2$ this gives that in the equality case we must have $p_1 = p_2 = (x, 0)$ where $\|x\|^2 = r^2$ as required. This proves the strict horizontal convexity of the norm N_{HS} .

So we obtained that the condition $p_1 = p_2 = (x, 0)$ is a necessary and sufficient condition for the equality in the triangle inequality. If we translate this condition for the starting points q_1, q_2 we see that the necessary and sufficient condition for the equality

$$N_{HS}(q_1 \cdot q_2) = N_{HS}(q_1) + N_{HS}(q_2),$$

is that $q_1 = (s_1x, 0)$ and $q_2 = (s_2x, 0)$ for some $s_1, s_2 > 0$. That is in fact the original definition of the horizontal strict convexity of N_{HS} .

Let us finally mention that from the proof it is seen that the value of $r_0 > 0$ in the statement of the theorem can be taken to be

$$r_0 = \min\left\{1, \frac{2}{\sqrt{5}C_1}, \frac{1}{4C_2}\right\}.$$

□

Using Proposition 6.2 from [24] we obtain as a by-product, the following consequence of Theorem 4.1:

Corollary 4.1. *Let $(\mathbf{H}, d_{\mathbf{H}})$ be a Carnot group with an arbitrary left-invariant homogeneous metric and $(\mathbf{G}, d_{N_{HS}})$ be a Carnot group endowed with the left-invariant metric induced by the Hebisch-Sikora norm N_{HS} defined by (2.6) for some $r < r_0$ and hence being horizontally strictly convex. Then all isometric embeddings $\phi : (\mathbf{H}, d_{\mathbf{H}}) \rightarrow (\mathbf{G}, d_{N_{HS}})$ are affine, i.e. they are compositions of left translations and group homomorphisms.*

ACKNOWLEDGEMENTS

This work was initiated at the thematic semester “Optimal transport on quantum structures” (Fall 2022) at the Erdős Center, Rényi Institute, Budapest. Zoltán M. Balogh would like to thank the Rényi Institute for the invitation to participate in this event and for the kind hospitality. We are grateful to Nicolas Juillet and Katrin Fässler for their inputs and for several discussions on the topic. We thank the anonymous reviewer for his/her detailed

comments and insightful suggestions which improved both the content and the presentation of this paper.

DECLARATIONS

Ethical Approval. Not applicable.

Funding. Z. M. Balogh is supported by the Swiss National Science Foundation, Grant Nr. 200020_191978 and 200021-228012. T. Titkos is supported by the Hungarian National Research, Development and Innovation Office (NKFIH) under grant agreements no. K134944 and no. Excellence.151232, and by the Momentum program of the Hungarian Academy of Sciences under grant agreement no. LP2021-15/2021. D. Virosztek is supported by the Momentum program of the Hungarian Academy of Sciences under grant agreement no. LP2021-15/2021, by the Hungarian National Research, Development and Innovation Office (NKFIH) under grant agreement no. Excellence.151232, and partially supported by the ERC Synergy Grant No. 810115.

Availability of data and materials. Not applicable.

REFERENCES

- [1] L. Ambrosio, N. Gigli, A user’s guide to optimal transport. Modelling and optimisation of flows on networks, *Lecture Notes in Math.*, 2062, Fond. CIME/CIME Found. Subser., Springer, Heidelberg, 2013.
- [2] Z. M. Balogh, K. Fässler, H. Sobrino, Isometric embeddings into Heisenberg groups, *Geom. Dedicata*, 185 (2018), 163–192.
- [3] Z. M. Balogh, A. Kristály, K. Sipos, Jacobian determinant inequality on corank-1 Carnot groups with applications, *J. Funct. Anal.* (277), 12, 108293 (2019).
- [4] Z. M. Balogh, T. Titkos, D. Virosztek, Isometries and isometric embeddings of Wasserstein spaces over the Heisenberg group, *Revista Matemática Iberoamericana*, DOI 10.4171/RMI/1576 (online first), arXiv:2303.15095 (2025).
- [5] Z. M. Balogh, J. T. Tyson, B. Warhurst, Sub-Riemannian vs. Euclidean dimension comparison and fractal geometry on Carnot groups, *Adv. Math.*, 220, (2009), 560–619.
- [6] D. Barilari and L. Rizzi, Sub-Riemannian interpolation inequalities, *Invent. Math.* 215 (2019), no. 3, 977–1038.
- [7] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni, Stratified Lie Groups and Potential Theory of their Sub-Laplacians, A. Bonfiglioli, E. Lanconelli, F. Uguzzoni, *Stratified Lie Groups and Potential Theory of their Sub-Laplacians* (2017).
- [8] J. Bertrand, B. Kloeckner, A geometric study of Wasserstein spaces: isometric rigidity in negative curvature, *Int. Math. Res. Notices*, 5 (2016), 1368–1386.
- [9] J. Bertrand, B. Kloeckner, A geometric study of Wasserstein spaces: Hadamard spaces, *J. Top. Anal.* Vol. 04 (2012), 515–542.
- [10] G. Dolinar, B. Kuzma, D. Mitrović, Isometries of probability measures with respect to the total variation distance, *J. Math. Anal. Appl.* (2021), Article No. 125829.
- [11] G. Dolinar, L. Molnár, Isometries of the space of distribution functions with respect to the Kolmogorov–Smirnov metric, *J. Math. Anal. Appl.* 348 (2008), 494–498.
- [12] D. A. Edwards, On the Kantorovich-Rubinstein theorem, *Expositiones Mathematicae*, Volume 29, Issue 4 (2011), 387–398.
- [13] A. Figalli, F. Glaudo, An Invitation to Optimal Transport, Wasserstein Distances, and Gradient Flows, *EMS Textbooks in Mathematics*, Volume 23, 2021.
- [14] A. Figalli and L. Rifford, Mass transportation on sub-Riemannian manifolds, *Geom. Funct. Anal.* 20 (2010), no. 1, 124–159.
- [15] A. Fotiadis, The Ptolemaean Inequality in H -Type Groups, *J. Geom. Symmetry Phys.* 34 (2014), 27–33.
- [16] Gy. P. Gehér, Surjective Kuiper isometries, *Houston J. Math.* 44 (2018), 263–281.
- [17] Gy. P. Gehér, A. Hruskova, T. Titkos, D. Virosztek, Isometric rigidity of Wasserstein spaces over Euclidean spheres, *J. Math. Anal. Appl.* 542 (2025), 128810.

- [18] Gy. P. Gehér, and T. Titkos, A characterisation of isometries with respect to the Lévy-Prokhorov metric, *Ann. Sc. Norm. Super. Pisa Cl. Sci.* Vol. XIX (2019), 655–677.
- [19] Gy. P. Gehér, T. Titkos, D. Virosztek, On isometric embeddings of Wasserstein spaces – the discrete case, *J. Math. Anal. Appl.* 480 (2019), Article No. 123435.
- [20] Gy. P. Gehér, T. Titkos, D. Virosztek, Isometric study of Wasserstein spaces-the real line, *Trans. Amer. Math. Soc.* 373 (2020), 5855–5883.
- [21] Gy. P. Gehér, T. Titkos, D. Virosztek, Isometry group of Wasserstein spaces: the Hilbertian case, *J. London Math. Soc.* 106 (2022), 3865–3894.
- [22] Gy. P. Gehér, T. Titkos, D. Virosztek, Isometric rigidity of Wasserstein tori and spheres, *Mathematika* 69 (2023), 20–32.
- [23] Gy. P. Gehér, T. Titkos, D. Virosztek, On the exotic isometry flow of the quadratic Wasserstein space over the real line, *Linear Algebra Appl.*, Vol. 693 (2024), 39–48.
- [24] E. Hakavouri, Infinite geodesics and Isometric embeddings in Carnot groups of step 2, *SIAM J. on Control and Optimization*, Vol. 58 (2020), 447–461.
- [25] W. Hebisch, A. Sikora, A smooth subadditive homogeneous norm on a homogeneous group, *Studia Math.* (96) 231–236, (1990).
- [26] G. Kiss, T. Titkos, Isometric rigidity of Wasserstein spaces: the graph metric case, *Proc. Amer. Math. Soc.* 150 (2022), 4083–4097.
- [27] B. Kloeckner, A geometric study of Wasserstein spaces: Euclidean spaces. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* 5 (2010), 297–323.
- [28] B. Kloeckner, A generalization of Hausdorff dimension applied to Hilbert cubes and Wasserstein spaces, *J. Top. Anal.* Vol. 04 (2012), 203–235.
- [29] E. Le Donne, S. Rigot, Besicovitch covering property for homogeneous distances on the Heisenberg groups, *J. Eur. Math. Soc.* 19 (2017), 1589–1617.
- [30] E. Le Donne, S. Rigot, Besicovitch Covering Property on graded groups and applications to measure differentiation *J. für die Reine und Angew. Math.* 750 (2019), 241–297.
- [31] J. R. Lee, A. Naor L^p metrics on the Heisenberg group and the Goemans-Linial conjecture. *Focs, IEEE Computer Society*, pp. 99–108 (2006)
- [32] S. Li, Stratified β -numbers and traveling salesman in Carnot groups, *J. London Math. Soc.* 106 (2022), 662–703.
- [33] J. Lott, C. Villani, Ricci curvature for metric measure spaces via optimal transport. *Ann. of Math.* 169 (2009), 903–991.
- [34] L. Molnár, Lévy isometries of the space of probability distribution functions, *J. Math. Anal. Appl.* 380 (2011), 847–852.
- [35] L. Rizzi, Measure contraction properties of Carnot groups, *Calc. Var. Partial Differ. Equ.* 55, (2016), no. 3, Art. 60, 20 pp.
- [36] F. Santambrogio, Optimal Transport for Applied Mathematicians, *Progress in Nonlinear Differential Equations and Their Applications* 87, Birkhauser Basel (2015).
- [37] J. Santos-Rodríguez, On isometries of compact L^p -Wasserstein spaces, *Adv. Math.* 409 (2022), Article No. 108632.
- [38] C. Villani, Topics in Optimal Transportation, *Graduate Studies in Mathematics.* vol. 58 of, Amer. Math. Soc. Providence, RI, 2003.
- [39] C. Villani, Optimal Transport: Old and New, *Grundlehren der mathematischen Wissenschaften* 338, Springer-Verlag Berlin Heidelberg, 2009.
- [40] H. Weyl, Symmetry, *Princeton University Press*, Princeton, NJ, 1989.

ZOLTÁN M. BALOGH, UNIVERSITÄT BERN, MATHEMATISCHES INSTITUT (MAI), SIDLERSTRASSE 12,
3012 BERN, SCHWEIZ

Email address: `zoltan.balogh@unibe.ch`

TAMÁS TITKOS, CORVINUS UNIVERSITY OF BUDAPEST, DEPARTMENT OF MATHEMATICS, FŐVÁM TÉR
13-15, BUDAPEST 1093, HUNGARY, AND HUN-REN ALFRÉD RÉNYI INSTITUTE OF MATHEMATICS, REÁLTANODA
U. 13-15., BUDAPEST 1053, HUNGARY

Email address: `titkos.tamas@renyi.hu`

DÁNIEL VIROSZTEK, HUN-REN ALFRÉD RÉNYI INSTITUTE OF MATHEMATICS, REÁLTANODA U. 13-15.,
BUDAPEST 1053, HUNGARY

Email address: `virosztek.daniel@renyi.hu`