METHODS AND APPLICATIONS OF ANALYSIS.
Vol. 28, No. 1, pp. 061–076, March 2021

005

# MULTI-MARGINAL OPTIMAL TRANSPORT ON THE HEISENBERG GROUP*

BRENDAN PASS†, ANDREA PINAMONTI‡, AND MATTIA VEDOVATO§

**Abstract.** We consider the multi-marginal optimal transport of aligning several compactly supported marginals on the Heisenberg group to minimize the total cost, which we take to be the sum of the squared Carnot-Carathéodory distances from the marginal points to their barycenter. Under certain technical hypotheses, we prove existence and uniqueness of optimal maps. We also point out several related open questions.

**Key words.** optimal transport, multi-marginal problems, Heisenberg group, sub-Riemannian geometry, Wasserstein barycenters.

**Mathematics Subject Classification.** 35R3, 53C17, 49Q22.

**1. Introduction.** Given Borel probability measures $\mu_1, \mu_2, ..., \mu_m$ on a metric space $X$, and a cost function $c : X^m \to \mathbb{R}$, Monge's multi-marginal optimal transport problem is to minimize

$$\int_X c(x_1, T_2(x_1)..., T_m(x_1))d\mu_1(x_1) \tag{MP}$$

among $(m-1)$-tuples of mappings $(T_2, ..., T_m)$ where each $T_i : X \to X$ pushes $\mu_1$ forward to $\mu_i$, $(T_i)_\#\mu_1 = \mu_i$; that is, $\mu_1(T_i^{-1}(A)) = \mu_i(A)$ for each measurable $A \subset X$. When $m = 2$, (MP) reduces to the well known classical optimal transport problem of Monge, which has many applications both within and outside of mathematics; see [Vil03, Vil09, San15] for comprehensive surveys. Of particular interest are cost functions reflecting the underlying geometry of $X$; the most thoroughly studied (and most important in applications) case arises when $c(x_1, x_2) = d^2(x_1, x_2)$ is the metric distance squared. For this cost, when $X = \mathbb{R}^n$, a seminal theorem of Brenier [Bre87, Bre91] asserts that there exists a unique minimizer to (MP). This result has been extended to much more exotic geometrical settings, beginning with the work of McCann when $X$ is a Riemannian manifold [McC01]. Of special interest in the present paper will be the case where $X = \mathbb{H}^n$ is the Heisenberg group equipped with the Carnot-Carathéodory distance, for which the existence and uniqueness of optimal maps was established by Ambrosio and Rigot in [AR04]. Although it is not directly relevant here, we mention that these properties have been extended to a wide class of general cost functions (those satisfying the *twist* condition, injectivity of $x_2 \mapsto D_{x_1}c(x_1, x_2)$; see, for example, [San15]). Moreover, it's worth noting that analogous results are available in more general subriemannian spaces (again taking the squared Carnot-Carathéodory distance $d_c^2$ as a cost function, see [FR10]), and existence holds on the Heisenberg group with the distance cost function ($c = d_c$, see [DR11]).

---

*Received March 21, 2020; accepted for publication June 11, 2020.

†Department of Mathematical and Statistical Sciences, University of Alberta, Edmonton, Alberta, Canada (pass@ualberta.ca). BP is pleased to acknowledge support from Natural Sciences and Engineering Research Council of Canada Grant 04658-2018.

‡Dipartimento di Matematica, Università degli Studi di Trento, Trento, Italy (andrea.pinamonti@unitn.it). AP is partially supported by Gruppo Nazionale per l'Analisi Matematica, la Probabilitá e le loro Applicazioni (GNAMPA) of the Istituto Nazionale di Alta Matematica (INDAM).

§Dipartimento di Matematica, Università degli Studi di Trento, Trento, Italy (mattia.vedovato@unitn.it).

Multi-marginal problems (ie, (MP) with $m \geq 3$) have received increasing attention in recent years, due to their own fairly wide variety of applications. In particular, we note that for a natural extension of the quadratic cost on $\mathbb{R}^n$, the existence of unique solutions was proven in a pioneering paper by Gangbo and Swiech [GŚ98]. These results were extended to Riemannian manifolds by Kim and Pass [KP15], but have so far not been established on many other spaces (an exception is Alexandrov spaces; see [Jia17]). More generally, the multi-marginal problem is much more delicate than its two marginal counterpart; existence and uniqueness results have only been established for very special general cost functions, and in fact many examples of multi-marginal cost functions have been exhibited for which solutions to the relaxed, Kantorovich version of the problem ((KP) below) do not induce solutions to (MP) and are not unique (see [Pas15] for an overview). The purpose of the present short paper is to show that the techniques of Ambrosio-Rigot can be combined with those of Kim-Pass to yield the existence of unique solutions to the multi-marginal optimal transport problem on the Heisenberg group; see Theorem 23 and Corollary 29 below.

Our proof requires two technical assumptions, not present in either [AR04] or [KP15], which may seem surprising to experts. First, we require in Theorem 23 that the mapping from optimally coupled points to their barycenter is injective; in Corollary 29 this follows from the sufficient condition that each $\mu_i$ assigns mass 0 to the set of barycenters of points in the supports on the marginals. This fact can be proven in the Riemannian setting, but we were unable to prove it on the Heisenberg group without additional assumptions, essentially because semi-concavity of the distance squared fails along the diagonal. In the Riemannian case, global semi-concavity allows one to use a now quite standard argument, originally introduced in [McC01], to show that differentiability of Kantorovich potentials implies differentiability of the squared distance for optimally coupled points. The second additional assumption (present in Corollary 29 but not Theorem 23) is that *all*, rather than only the first, marginals must be absolutely continuous with respect to Lebesgue measure. This is because we use almost everywhere differentiability of each Kantorovich potential to ensure differentability of the distance squared from each marginal point $x_i$ to the barycenter $y$ (under the assumption above ensuring $y \neq x_i$), allowing us to invert the relationship $(x_1, ..., x_m) \mapsto y$. We do not know at this time whether either of these assumptions can be removed. However, we note that somewhat surprisingly, the semi-concavity of the cost function issue vanishes when we replace the quadratic cost with the $p$ power-cost for $p > 2$; see Section 4.1.

The manuscript is organized as follows: in section 2, we recall basic facts about both the Heisenberg group and multi-marginal optimal transport. In section 3, we introduce the cost function we will work with and establish some preliminary properties of it. Section 4 is devoted to the precise statement and proof of our main results: Theorem 23 and Corollary 29.

## 2. Preliminaries: multimarginal optimal transport on $\mathbb{H}^n$.

**2.1. Heisenberg group.** In this paper, we will represent the Heisenberg group $\mathbb{H}^n$ as the set $\mathbb{C}^n \times \mathbb{R} \equiv \mathbb{R}^{2n+1}$; thus its points will be described as $x = [z,t] = [\zeta + i\eta, t] = (\zeta, \eta, t)$, with $z \in \mathbb{C}^n$, $\zeta, \eta \in \mathbb{R}^n$, $t \in \mathbb{R}$. The group operation on $\mathbb{H}^n$ will be defined as follows: whenever $x = [z,t] \in \mathbb{H}^n$ and $x' = [z', t'] \in \mathbb{H}^n$,

$$x \cdot x' \doteq \big[z + z', t + t' + 2\,\mathrm{Im}\big(\langle z, \bar{z'}\rangle\big)\big]. \tag{1}$$

As a consequence, it is easy to verify that the group identity is the origin 0 and the inverse of a point is given by $[z,t]^{-1} = [-z, -t]$. Even though they will only play a role "in the background" in the current work, we also introduce for the sake of completeness the following family of *non-isotropic dilations*, since they are fundamental in the geometry of the Heisenberg group: if $x = [z, t] \in \mathbb{H}^n$ and $\lambda > 0$,

$$\delta_\lambda(x) \doteq [\lambda z, \lambda^2 t]. \tag{2}$$

The Heisenberg group $\mathbb{H}^n$ admits the structure of a Lie group of topological dimension $2n + 1$. Its Lie algebra $\mathfrak{h}_n$ of left invariant vector fields is (linearly) generated by

$$X_j = \frac{\partial}{\partial x_j} + 2y_j \frac{\partial}{\partial t}, \quad Y_j = \frac{\partial}{\partial y_j} - 2x_j \frac{\partial}{\partial t}, \quad \text{for } j = 1, \dots, n; \quad \text{and } Z = \frac{\partial}{\partial t}; \tag{3}$$

notice that the only non-trivial commutator relation is

$$[X_j, Y_j] = -4Z, \tag{4}$$

valid for any $j = 1, \dots, n$. The vector fields $X_1, \dots, X_n, Y_1, \dots, Y_n$ will be called *horizontal vector fields*: the group $\mathbb{H}^n$ endowed with this Lie algebra has the structure of a Carnot group.

In our main Monge-type result, we will need the following geometric result about the structure of the Heisenberg group: the space $\mathbb{H}^n$ can be parametrized by a system of "adapted spherical coordinates".

Proposition 1 (Spherical coordinates, [AR04, Proposition 3.9]). *Let*

$$\mathbb{S} \doteq \{a + ib \in \mathbb{C}^n \mid |a + ib| = 1\}, \tag{5}$$

*and let* $\mathbf{D} \doteq \mathbb{S} \times (-2\pi, 2\pi) \times (0, +\infty)$. *Define the following map:*

$$\begin{aligned} \Upsilon : \qquad \mathbf{D} &\longrightarrow \mathbb{H}^n \\ (a + ib, v, r) &\longmapsto (\xi_1, \dots, \xi_n, \eta_1, \dots, \eta_n, t) \end{aligned} \tag{6}$$

*with*

$$\begin{aligned} \xi_j &\doteq \frac{b_j(1 - \cos v) + a_j \sin v}{v} r \\ \eta_j &\doteq \frac{-a_j(1 - \cos v) + b_j \sin v}{v} r \\ t &\doteq 2\frac{v - \sin v}{v^2} r^2. \end{aligned} \tag{7}$$

*Then* $\Upsilon$ *is a* $C^1$ *diffeomorphism from* $\mathbf{D}$ *onto* $\mathbb{H}^n \setminus L$, *where* $L = \{0\} \times \mathbb{R}$ *is the vertical axis.*

Thanks to Proposition 1, one can define a generalization of the exponential map to the Heisenberg group:

Definition 2 (Exponential map on $\mathbb{H}^n$). Let $\Upsilon$ be the diffeomorphism introduced in Proposition 1. We define the *exponential map* $\exp_{\mathbb{H}} : \mathbb{C}^n \times [-\frac{\pi}{2}, \frac{\pi}{2}] \to \mathbb{H}^n$ as

$$\exp_{\mathbb{H}}(A + iB, w) \doteq \Upsilon\left(\frac{A + iB}{|A + iB|}, 4w, |A + iB|\right) \tag{8}$$

if $A + iB \neq 0$, and $\exp_{\mathbb{H}}(0, w) = 0$ for all $w$.

As shown in [AR04, Lemma 6.8], this definition is consistent with the classical Riemannian one: one can define in a standard way a Riemannian approximation to the subriemannian structure of $\mathbb{H}^n$; and now $\exp_{\mathbb{H}}$ can be recovered as a limit of Riemannian exponential coordinates.

Other important properties of the Heisenberg group (such as the metric structure induced by the Carnot-Carathéodory distance) will be introduced in Section 3; we refer the reader to [Ser16] for a more complete introduction to the Heisenberg group and Carnot groups in general.

**2.2. Multimarginal optimal transport.** In what follows, $\mu_1, \dots, \mu_m \in \mathcal{P}(\mathbb{H}^n)$ will be Borel probability measures on the Heisenberg group $\mathbb{H}^n$. We'll also assume they are absolutely continuous with respect to the Lebesgue measure $\mathscr{L}^{2n+1}$ (which is, up to a constant, the Haar measure of $\mathbb{H}^n$). We will call **Kantorovič problem** the following minimization problem:

$$\inf\left\{ \int_{(\mathbb{H}^n)^m} c(x_1, \dots, x_m)\, d\gamma(x_1, \dots, x_m) \;\middle|\; \gamma \in \Pi(\mu_1, \dots, \mu_m) \right\} \tag{KP}$$

where $c : (\mathbb{H}^n)^m \to \mathbb{R}$ is a suitable cost function, $\Pi(\mu_1, \dots, \mu_m)$ is the following family of **transport plans**:

$$\Pi(\mu_1, \dots, \mu_m) \doteq \left\{ \mu \in \mathcal{P}((\mathbb{H}^n)^m) \;\middle|\; \pi_{i\sharp}\mu = \mu_i \text{ for all } i = 1, \dots, m \right\}, \tag{9}$$

and $\pi_{i\sharp}\mu$ denotes the push-forward measure of $\mu$ with respect to the projection on the $i^{\text{th}}$ component.

We also consider the following **dual problem**:

$$\sup\left\{ \sum_{i=1}^m \int_{\mathbb{H}^n} u_i(x_i)\, d\mu_i(x_i) \;\middle|\; \begin{array}{c} u_i \in L^1_{\mu_i}(\mathbb{H}^n) \text{ for all } i = 1, \dots, m, \\ \sum_{i=1}^m u_i(x_i) \le c(x_1, \dots, x_m) \end{array} \right\}. \tag{DP}$$

In what follows, we will often denote by $\mathbf{x}$ the $m$-tuple $(x_1, \dots, x_m) \in (\mathbb{H}^n)^m$.

*Restriction to compact subsets.* From now on, we will always assume $\mu_1, \dots, \mu_m$ are compactly supported. If we define the compact sets

$$K_i \doteq \operatorname{spt}(\mu_i) \subset \mathbb{H}^n \quad \text{and} \quad \mathbf{K} \doteq \prod_i \operatorname{spt}(\mu_i) \subset (\mathbb{H}^n)^m, \tag{10}$$

then the Kantorovič problem (KP) is actually equivalent to

$$\inf\left\{ \int_{\mathbf{K}} c(x_1, \dots, x_m)\, d\gamma(x_1, \dots, x_m) \;\middle|\; \gamma \in \Pi(\mu_1, \dots, \mu_m) \right\}. \tag{KP$_c$}$$

Indeed, for any $\mu \in \Pi(\mu_1, \dots, \mu_m)$, we have $\pi_i(\operatorname{spt}\mu) \subset \operatorname{spt}(\pi_i\sharp\mu) = \operatorname{spt}(\mu_i)$ (by elementary properties of measures and by continuity of $\pi_i$), thus any transport plan is supported in $\prod_i \operatorname{spt}(\mu_i) = \mathbf{K}$. The dual problem (DP), on the other hand, is actually equivalent to the following:

$$\sup\left\{ \sum_{i=1}^m \int_{K_i} u_i(x_i)\, d\mu_i(x_i) \;\middle|\; \begin{array}{c} u_i \in L^1_{\mu_i}(K_i) \text{ for all } i = 1, \dots, m, \\ \sum_{i=1}^m u_i(x_i) \le c(x_1, \dots, x_m) \end{array} \right\}, \tag{DP$_c$}$$

in the sense that any solution of (DP) restricts to a solution of (DP$_c$) and vice versa, a solution of (DP$_c$) extends to a solution of (DP), for example by defining the value of $u_i$ to be $-\infty$ out of $K_i$.

Definition 3. Let $c : (\mathbb{H}^n)^m \to \mathbb{R}$ be a cost function. We say that the $m$-tuple of functions $\mathbf{u} = (u_1, \ldots, u_m) : \mathbb{H}^n \to (\mathbb{R} \cup \{-\infty\})^m$ is *c-conjugate* if, for some $U \subset \mathbb{H}^n$ and for every fixed $1 \le i \le m$ and $x_i \in \mathbb{H}^n$, it holds:

$$u_i(x_i) = \inf\left\{ c(x_1, \ldots, x_m) - \sum_{j \neq i} u_j(x_j) \;\middle|\; x_j \in U \text{ for any } j \neq i \right\}. \tag{11}$$

Moreover, when $\mathbf{u} = (u_1, \ldots, u_m)$ is such an $m$-tuple, we denote by $\Gamma_{\mathbf{u}}$ the set

$$\Gamma_{\mathbf{u}} \doteq \left\{ \mathbf{x} \in U^m \;\middle|\; \sum_{i=1}^{m} u_i(x_i) = c(\mathbf{x}) \right\}. \tag{12}$$

Definition 4 (Cyclical monotonicity). Let $\Gamma \subset (\mathbb{H}^n)^m$. We say that $\Gamma$ is **$c$-cyclically monotone** if the following holds: for any $N \in \mathbb{N}$, for any $m$-tuple $(\sigma_1, \ldots, \sigma_m)$ of permutations of $N$ elements, and for any subset

$$\left\{ \mathbf{x}^j = (x_1^j, \ldots, x_m^j) \in (\mathbb{H}^n)^m \;\middle|\; j = 1, \ldots, N \right\} \subset \Gamma \tag{13}$$

the following inequality holds true:

$$\sum_{j=1}^{N} c(x_1^j, \ldots, x_m^j) \le \sum_{j=1}^{N} c\Big(x_1^{\sigma_1(j)}, \ldots x_m^{\sigma_m(j)}\Big). \tag{14}$$

Proposition 5. *Let $c : (\mathbb{H}^n)^m \to \mathbb{R}$ be a cost function, and let $(u_1, \ldots, u_m) : \mathbb{H}^n \to \mathbb{R}^m$ be a c-conjugate m-tuple. Then the set $\Gamma_{\mathbf{u}}$ defined in Equation* (12) *is c-cyclically monotone.*

*Proof.* Fix $N \in \mathbb{N}$, a $m$-tuple $(\sigma_1, \ldots, \sigma_m)$ of $N$-permutations and a subset $\{\mathbf{x}^j\}_{j=1}^N \subset \Gamma_{\mathbf{u}}$. Then

$$\begin{aligned}\sum_{j=1}^{N} c\Big(x_1^{\sigma_1(j)}, \ldots x_m^{\sigma_m(j)}\Big) &\ge \sum_{j=1}^{N} \sum_{i=1}^{m} u_i(x_i^{\sigma_i(j)}) = \sum_{i=1}^{m} \sum_{j=1}^{N} u_i(x_i^{\sigma_i(j)}) = \\ &= \sum_{i=1}^{m} \sum_{j=1}^{N} u_i(x_i^j) = \sum_{j=1}^{N} c(x_1^j, \ldots x_m^j).\end{aligned} \tag{15}$$

□

The following Existence and Duality Theorem is a classical result in optimal transportation: even though it's not present in literature at this level of generality (multimarginal and subriemannian), the classical proof can be easily adapted to this context (see for example [AG13, Theorem 1.17 and Remark 1.18]).

Theorem 6 (Existence and Duality). *Let $\mu_1, \ldots, \mu_m \in \mathcal{P}(\mathbb{H}^n)$ be compactly supported, Borel probability measures. Let $c : (\mathbb{H}^n)^m \to \mathbb{R}$ be a continuous cost function. Then:*

*(i)* Existence for (KP$_c$): *there exists $\gamma \in \Pi(\mu_1, \ldots, \mu_m)$ which achieves the minimum in the Kantorovič problem* (KP$_c$)*;*

*(ii)* Existence for (DP$_c$): *there exists a c-conjugate m-tuple* $(u_1, \dots, u_m)$ *which achieves the maximum in the dual problem* (DP$_c$)*;*
*(iii)* Duality: *the minimum in* (KP$_c$) *and the maximum in* (DP$_c$) *coincide;*
*(iv)* *If* $\gamma$ *is an optimal plan and* $\mathbf{u} = (u_1, \dots, u_m)$ *is a c-conjugate solution to* (DP$_c$)*, then* $\operatorname{spt}(\gamma) \subset \Gamma_{\mathbf{u}}$.

Remark 7. If $\mathbf{u}$ is a $c$-conjugate $m$-tuple which maximizes (DP$_c$), we can assume

$$u_i(x_i) = \inf \left\{ c(x_1, \dots, x_m) - \sum_{j \neq i} u_j(x_j) \;\middle|\; x_j \in K_j \text{ for any } j \neq i \right\}. \tag{16}$$

for $\mu_i$-almost every $x_i \in \mathbb{H}^n$.

**3. Cost functions associated to distances.** We now specify what kind of cost functions we are interested in: namely, we adapt the idea of "cost functions which grow quadratically with the distance" to the multimarginal problem. The same problem was approached in [GŚ98] in the Euclidean space $\mathbb{R}^n$ with the standard distance: here the cost $c(\mathbf{x}) = \sum_{i \neq j} |x_i - x_j|^2$ was considered. When moving to the Riemannian case (see [KP15]), a natural way to generalize this is to take the "barycentric cost": for $x_i, y \in M$, with $(M, g)$ Riemannian manifold and $d$ distance associated to the metric $g$, one takes the cost

$$c(\mathbf{x}) = \inf_y \sum_i d^2(x_i, y), \tag{17}$$

which in $\mathbb{R}^n$ is actually equivalent to the Gangbo-Święch one. The same choice is made here for the Heisenberg group.

Definition 8 (Barycentric cost associated to a distance). For a given distance $d$ on $\mathbb{H}^n$, we define the cost function $c_d : (\mathbb{H}^n)^m \to \mathbb{R}$ associated to $d$ as the function

$$c_d(x_1, \dots, x_m) \doteq \inf_{y \in \mathbb{H}^n} \left\{ \sum_{i=1}^m d^2(x_i, y) \right\}. \tag{18}$$

We'll drop the subscript $d$ when the metric is clear from the context.

The infimum in (18) is actually a minimum, and the minimum points will play a fundamental role in our work:

Definition 9 (Barycenters). We'll say that $y \in \mathbb{H}^n$ is a **barycenter** for the $m$-tuple $(x_1, \dots, x_m)$ if $y$ realizes the infimum in Equation (18). We denote by $\mathbf{b} : (\mathbb{H}^n)^m \to \mathscr{P}(\mathbb{H}^n)$ the multivalued map which associates an $m$-tuple with its (never empty) set of barycenters. Here $\mathscr{P}(\mathbb{H}^n)$ is the power set of $\mathbb{H}^n$.

Remark 10. For any compact set $K \subset (\mathbb{H}^n)^m$ we can find a larger compact set $K'$ which contains all the barycenters of $m$-tuples $(x_1, \dots, x_m) \in K$: indeed, $c_d$ is upper semi-continuous, thus bounded from above by a value $M > 0$ in $K$; hence the set

$$H \doteq \bigcup_{\mathbf{x} \in K} \left\{ y \in \mathbb{H}^n \;\middle|\; \sum_{i=1}^m d^2(x_i, y) \leq M \right\} \tag{19}$$

contains all the barycenters and is easily seen to be bounded; it suffices to take $K' = \bar{H}$. In particular, if $(x_1, \ldots, x_m) \in K^m$, then

$$c_d(x_1, \ldots, x_m) \doteq \inf_{y \in K'} \left\{ \sum_{i=1}^{m} d^2(x_i, y) \right\}. \tag{20}$$

Notice that with the same argument one can easily prove that the cost function $c_d$ is actually continuous.

In the following Theorem, we show that if $\mathbf{u}$ is $c_d$-conjugate and $y$ is a barycenter for two different $m$-tuples $\mathbf{x}, \bar{\mathbf{x}}$ of $\Gamma_{\mathbf{u}}$, then it is also a barycenter for the $m$-tuples obtained by swapping the $i^{\text{th}}$-components of $\mathbf{x}$ and $\bar{\mathbf{x}}$.

Lemma 11. *Let:*

- $c = c_d : (\mathbb{H}^n)^m \to \mathbb{R}$ *be a cost function associated to a distance* $d$ *as in Definition 8;*
- $\mathbf{u} = (u_1, \ldots, u_m) : \mathbb{H}^n \to \mathbb{R}^m$ *be a* $c_d$*-conjugate* $m$*-tuple in* $U \subset \mathbb{H}^n$*;*
- $\mathbf{x}, \bar{\mathbf{x}}$ *belong to the set* $\Gamma_{\mathbf{u}}$ *defined in Equation* (12).

*Assume the point* $y \in \mathbb{H}^n$ *is a barycenter for both* $\mathbf{x}$ *and* $\bar{\mathbf{x}}$*. Then* $y$ *is also a barycenter for the* $m$*-tuple* $\mathbf{x}' \doteq (\bar{x}_1, x_2, \ldots, x_m)$ *and for the* $m$*-tuple* $\bar{\mathbf{x}}' \doteq (x_1, \bar{x}_2, \ldots, \bar{x}_m)$*.*

*Proof.* By $c$-cyclical monotonicity,

$$\begin{aligned} c(\mathbf{x}) + c(\bar{\mathbf{x}}) &\le c(\mathbf{x}') + c(\bar{\mathbf{x}}') \\ &\le d_{\mathrm{c}}^2(\bar{x}_1, y) + \sum_{i=2}^{m} d_{\mathrm{c}}^2(x_i, y) + d_{\mathrm{c}}^2(x_1, y) + \sum_{i=2}^{m} d_{\mathrm{c}}^2(\bar{x}_i, y) \\ &= \sum_{i=1}^{m} d_{\mathrm{c}}^2(x_i, y) + \sum_{i=1}^{m} d_{\mathrm{c}}^2(\bar{x}_i, y) \\ &= c(\mathbf{x}) + c(\bar{\mathbf{x}}). \end{aligned} \tag{21}$$

We must therefore have equality throughout the above string of inequalities, meaning that

$$c(\mathbf{x}') = d_{\mathrm{c}}^2(\bar{x}_1, y) + \sum_{i=2}^{m} d_{\mathrm{c}}^2(x_i, y) \text{ and } c(\bar{\mathbf{x}}') = d_{\mathrm{c}}^2(x_1, y) + \sum_{i=2}^{m} d_{\mathrm{c}}^2(\bar{x}_i, y), \tag{22}$$

as desired. □

Our main result concerns the cost function associated to a special distance on $\mathbb{H}^n$, *i.e.* the subriemannian Carnot-Carathéodory distance, which we introduce here:

Definition 12 (Carnot-Carathéodory distance). We call a *subunit curve* a Lipschitz curve $\gamma : [0, T] \longrightarrow \mathbb{H}^n$ such that for a.e. $t \in [0, T]$,

$$\begin{gathered} \dot{\gamma}(t) = \sum_{j=1}^{n} a_j(t) X_j(\gamma(t)) + b_j(t) Y_j(\gamma(t)) \\ \text{and } \sum_{j=1}^{n} a_j^2(t) + b_j^2(t) \le 1 \end{gathered} \tag{23}$$

with $a_1, ..., a_n, b_1, \dots, b_n$ measurable coefficients. The Carnot-Carathéodory distance between the points $x, y \in \mathbb{H}^n$ is defined as

$$d_c(x,y) \doteq \inf\left\{ T \geq 0 \;\middle|\; \begin{array}{c} \text{there exists a subunit curve } \gamma : [0,T] \to \mathbb{H}^n \\ \text{such that } \gamma(0) = x \text{ and } \gamma(T) = y \end{array} \right\}. \tag{24}$$

Remark 13. The Carnot-Carathéodory distance on $\mathbb{H}^n$ is an example of a *left invariant* and *homogeneous* metric, *i.e.*, whenever $x, \bar{x} \in \mathbb{H}^n$, $p \in \mathbb{H}^n$ and $\lambda > 0$, the following hold:

$$d_c(p \cdot x, p \cdot \bar{x}) = d_c(x, \bar{x}) \quad \text{and} \quad d_c(\delta_\lambda(x), \delta_\lambda(\bar{x})) = d(x, \bar{x}), \tag{25}$$

where $\delta_\lambda([z,t]) = [\lambda z, \lambda^2 t]$ for any $[z,t] \in \mathbb{C}^n \times \mathbb{R} \simeq \mathbb{H}^n$. It's easy to prove that any left invariant and homogeneous metric is equivalent to $d_c$ (see for example [BLU07, Corollary 5.1.5]). This also means that the notion of "Lipschitz function" does not depend on the particular metric chosen, as long as it is left invariant and homogeneous.

We collect here some properties of the (squared) distance function, which are mostly proved in [AR04]; some of them will turn out to be useful later, the others are stated for the sake of completeness.

Proposition 14 (Differentiability properties of $d_{\mathrm{c}}^2$). *Let us denote, with an abuse of notation,* $d_{\mathrm{c}}(z) \doteq d_{\mathrm{c}}(0,z)$ *for any* $z \in \mathbb{H}^n$. *Let* $L = \{s e_{2n+1} \mid s \in \mathbb{R}\}$ *be the vertical axis. Then:*

*(a) If* $z \notin L$, *then* $d_{\mathrm{c}}^2$ *is differentiable at* $z$.

*(b)* $X_j d_{\mathrm{c}}^2(0) = 0$ *and* $Y_j d_{\mathrm{c}}^2(0) = 0$*; moreover,* $Z^{\pm} d_{\mathrm{c}}^2(0) = \pm\pi$.

*(c) If* $z = [0, \dots, 0, t] \in L \setminus \{0\}$, *then* $d_{\mathrm{c}}^2(z) = \pi|t|$, *thus,* $Z d_{\mathrm{c}}^2(z) = \operatorname{sgn}(t)\pi$*; however,* $d_{\mathrm{c}}^2$ *is not differentiable at* $z$ *along the directions* $X_j, Y_j$ *with* $j = 1, \dots, n$.

We now introduce the notion of $d^2$-concavity, which plays a key role in the two-marginals optimal transport theory, and is closely related to the (multi-marginal) notion of $c_d$-conjugacy (see Lemma 19).

Definition 15. Let $u : \mathbb{H}^n \to \mathbb{R}$. Let $d$ be a metric on $\mathbb{H}^n$. We say that $u$ is $d^2$-concave if

$$u(x) = \inf_{y \in U} \{ d^2(x,y) - \phi(y) \} \qquad \forall x \in \mathbb{H}^n \tag{26}$$

for some non-empty set $U \in \mathbb{H}^n$ and $\phi : U \to \mathbb{R} \cup \{-\infty\}$, $\phi \not\equiv -\infty$ (see [AR04, Definition 4.1]).

Remark 16. If $u$ is $d^2$-concave, then Equation (26) also holds with $U = \mathbb{H}^n$, up to defining $\phi \equiv -\infty$ on $\mathbb{H}^n \setminus U$.

Definition 17. Let $u : \mathbb{H}^n \to \mathbb{R}$ be a function, and let $d$ be a metric on $\mathbb{H}^n$. We define the $d^2$-superdifferential of $u$ at $x \in \mathbb{H}^n$ as

$$\partial_{d^2} u(x) \doteq \{ y \in \mathbb{H}^n \mid d^2(x,y) - u(x) \leq d^2(z,y) - u(z) \; \forall z \in \mathbb{H}^n \}. \tag{27}$$

We denote by $\partial_{d^2} u$ the graph of the $d^2$-superdifferential:

$$\partial_{d^2} u \doteq \{ (x,y) \in \mathbb{H}^n \times \mathbb{H}^n \mid y \in \partial_{d^2} u(x) \}. \tag{28}$$

Remark 18. If $u$ is $d^2$-concave, then $\partial_{d^2} u$ coincides with the set

$$\big\{(x,y)\in\mathbb{H}^n\times\mathbb{H}^n \;\big|\; u(x)+\phi(y)=d^2(x,y)\big\}. \tag{29}$$

In the following Theorem, we specify the relation between $c_d$-conjugacy and $d^2$-concavity: in particular, we show that each component of a $c_d$-conjugate solution $\mathbf{u}$ to the dual problem $(\mathrm{DP}_c)$ is $d^2$-concave.

Lemma 19. *Let $c_d : (\mathbb{H}^n)^m \to \mathbb{R}$ be a cost function associated to a homogeneous and left invariant distance $d$, let $\mu_1,\dots,\mu_m$ be compactly supported, absolutely continuous probability measures on $\mathbb{H}^n$ and let $(u_1,\dots,u_m) : \mathbb{H}^n \to \mathbb{R}^m$ be a $c_d$-conjugate solution to the dual problem $(\mathrm{DP}_c)$. Then each $u_j$ is $d^2$-concave. Moreover, if $\bar{\mathbf{x}} \in \Gamma_{\mathbf{u}}$ and $y\in\mathbb{H}^n$ is a barycenter for $\bar{\mathbf{x}}$, then $y\in\partial_{d^2}u_i(\bar{x}_i)$ for all $i\in\{1,\dots,m\}$.*

*Proof.* Notice that for all $x_i\in\mathbb{H}^n$ it holds:

$$u_i(x_i) = \inf_{x_j\in K_j, j\neq i}\left\{\inf_{y\in K'}\left\{\sum_{j=1}^m d^2(x_j,y)\right\} - \sum_{j\neq i} u_j(x_j)\right\}, \tag{30}$$

where $K_j$ is defined in Equation (10) and $K'$ is a suitable compact set (see Remark 10); thus

$$u_i(x_i) = \inf_{y\in K'}\left\{d^2(x_i,y) + \inf_{x_j\in K_j, j\neq i}\left\{\sum_{j\neq i}(d^2(x_j,y)-u_j(x_j))\right\}\right\}. \tag{31}$$

This proves the first statement. Let now $\bar{\mathbf{x}}\in\Gamma_{\mathbf{u}}$, and $y$ be a barycenter. Then, taking for example $i=1$, we have

$$d^2(\bar{x}_1,y) - u_1(\bar{x}_1) = \sum_{j>1}(u_j(\bar{x}_j) - d^2(\bar{x}_j,y)); \tag{32}$$

since $\mathbf{u}$ is $c_d$-conjugate, we then have for any $z\in\mathbb{H}^n$

$$d^2(\bar{x}_1,y) - u_1(\bar{x}_1) \le c_d(z,\bar{x}_2,\dots,\bar{x}_m) - \sum_{j>1} d^2(\bar{x}_j,y) - u_1(z) \le d^2(z,y) - u_1(z), \tag{33}$$

which is what we needed to prove $y\in\partial_{d^2}u_1(\bar{x}_1)$. □

Corollary 20. *Let $c : (\mathbb{H}^n)^m \to \mathbb{R}$ be the cost function associated to the Carnot-Carathéodory distance $d_c$; let $\mu_1,\dots,\mu_m$ be compactly supported, absolutely continuous probability measures on $\mathbb{H}^n$ and let $(u_1,\dots,u_m) : \mathbb{H}^n\to\mathbb{R}^m$ be a c-conjugate solution to the dual problem $(\mathrm{DP}_c)$. Then each $u_j$ is Lipschitz, and thus Pansu differentiable almost everywhere; moreover, $Zu_j$ also exists almost everywhere in $\mathbb{H}^n$.*

*Proof.* This is a consequence of [AR04, Lemma 4.5 and Lemma 4.6]. □

Remark 21. We refer to [AR04, Paragraph 2.2] for a definition of Pansu differentiability and related issues, such as the Rademacher-type Theorem which is needed for Corollary 20. What is important to us is that each $u_j$ is differentiable along $X_1,\dots,X_n$, $Y_1,\dots,Y_n$ and $Z$ for a.e. $x\in\mathbb{H}^n$, in the following sense: for any $k=1,\dots,2n+1$, the map $s\mapsto u_j(x\cdot\delta_s(e_k))$ is differentiable at $s=0$.

Now fix $d_c$ to be the Carnot-Carathéodory distance, and denote by $c$ the associated cost function. Let $\mathbf{u} = (u_1, \dots, u_m)$ be a $c$-conjugate solution to the dual problem ($\mathrm{DP}_c$), and let now $E_i \subset \mathbb{H}^n$ be the ($\mathscr{L}^{2n+1}$-negligible) set where $u_i$ fails to be differentiable along some of the directions $X_j, Y_j, Z$. Define the set

$$\Omega_{\mathbf{u}} \doteq \prod_{i=1}^{m} (\mathbb{H}^n \setminus E_i). \tag{34}$$

Then, if $\mu_i \ll \mathscr{L}^{2n+1}$ for all $i = 1, \dots, m$ and $\mu \in \Pi(\mu_1, \dots, \mu_m)$ is a transport plan, one has:

$$\mu((\mathbb{H}^n)^m \setminus \Omega_{\mathbf{u}}) \leq \sum_{i=1}^{m} \mu(\mathbb{H}^n \times \cdots \times E_i \times \cdots \times \mathbb{H}^n) = \mu_i(E_i) = 0, \tag{35}$$

thus $\Omega_{\mathbf{u}}$ has full $\mu$-measure in $(\mathbb{H}^n)^m$.

**4. Main results.** In the following definition, we introduce a condition which will turn out to be useful in our main result: we assume that in a suitable subset of $\mathbb{H}^n$ the barycenter map can be inverted.

Definition 22. Let $c_d : (\mathbb{H}^n)^m \to \mathbb{R}$ be a cost function associated to a distance $d$, and let $\Gamma \subset (\mathbb{H}^n)^m$. We say that $\Gamma$ satisfies the condition (C1) if:

$$\text{there exists } \Gamma_1 \subset \Gamma \text{ such that } \mathscr{L}^{2n+1}(\pi_1(\Gamma \setminus \Gamma_1)) = 0 \text{ and the restricted barycenter map } \mathbf{b}\big|_{\Gamma_1} \text{ is injective.} \tag{C1}$$

Equivalently,

$$\text{there exists } \Gamma_1 \subset \Gamma \text{ such that } \mathscr{L}^{2n+1}(\pi_1(\Gamma \setminus \Gamma_1)) = 0, \text{ and whenever } \mathbf{x},\, \bar{\mathbf{x}} \in \Gamma_1 \text{ have a common barycenter } y \in \mathbb{H}^n,\ \mathbf{x} \text{ and } \bar{\mathbf{x}} \text{ must coincide.} \tag{C1'}$$

We are now ready to prove a general Monge-type Theorem under the assumption (C1) on the invertibility of the barycenter map. We will then state a condition on the marginals which ensures (C1).

Theorem 23. *Let $\mu_1, \dots, \mu_m$ be compactly supported, absolutely continuous probability measures on $\mathbb{H}^n$, and let $c$ be the cost associated to $d_c$. Let $\mathbf{u} = (u_1, \dots, u_m)$ be a $c$-conjugate solution to the dual problem* ($\mathrm{DP}_c$). *Assume that the set $\Gamma_{\mathbf{u}}$ defined in* (12) *satisfies the condition* (C1) *for some subset $\Gamma_1 \subset \Gamma_{\mathbf{u}}$. Then:*

*(i) Any optimal transport plan $\gamma$ which achieves the minimum in* ($\mathrm{KP}_c$) *is induced by a transport map $\Psi$ over the first variable $x_1$.*

*(ii) The optimal transport plan $\gamma$ and map $\Psi$ are both unique.*

*(iii) Let $\exp_{\mathbb{H}}$ denote the subriemannian exponential map defined in Definition 2; let*

$$\mathbf{b}_{\Gamma_1}^{-1} : \mathbf{b}(\Gamma_1) \to \Gamma_1 \subset (\mathbb{H}^n)^m \tag{36}$$

*denote the inverse of the barycenter map, and let $\mathbf{f} : \mathbf{b}(\Gamma_1) \to (\mathbb{H}^n)^{m-1}$ be defined by $\mathbf{f}_i = (\mathbf{b}_{\Gamma_1}^{-1})_i$ for $i = 2, \dots, m$. Then $\Psi$ can be represented ($\mu_1$-a.e.) as*

$$\Psi(x_1) = \mathbf{f} \circ \psi\,(x_1), \tag{37}$$

*where*

$$\psi(x_1) = x_1 \cdot \exp_{\mathbb{H}}(-Xu_1(x_1) - iYu_1(x_1), -Zu_1(x_1)) \tag{38}$$

*Proof.* Recall that a transport plan $\mu \in \Pi(\mu_1, \dots, \mu_m)$ is induced by a transport map if and only if there exists a $\mu$-measurable set $\tilde{\Gamma} \subset (\mathbb{H}^n)^m$ where $\mu$ is concentrated, such that for $\mu_1$-a.e. $x_1$ there exists only one $(x_2, \dots, x_m) = \Psi(x_1) \in (\mathbb{H}^n)^{m-1}$ such that $(x_1, x_2, \dots, x_m) \in \tilde{\Gamma}$; in this case $\mu$ is induced by the map $\Psi$ (see [AG13, Lemma 1.20]).

We already know from Theorem 6 that, if $\gamma$ is an optimal Kantorovič plan, then it is concentrated on the set $\Gamma_{\mathbf{u}}$ for some $c$-conjugate solution $\mathbf{u} = (u_1, ..., u_m)$ to the dual problem (DP). Let now $\Omega_{\mathbf{u}}$ be as defined in (34); by condition (C1) and by absolute continuity of $\mu_1$, the projection $\pi_1(\Omega_{\mathbf{u}} \cap \Gamma_1)$ has full $\mu_1$-measure. It is enough to show that for any $x_1 \in \pi_1(\Omega_{\mathbf{u}} \cap \Gamma_1)$ there exists exactly one $\mathbf{x} = (x_1, x_2, \dots, x_m)$ which belongs to $\Omega_{\mathbf{u}} \cap \Gamma_1$.

Fix $x_1 \in \pi_1(\Omega_{\mathbf{u}} \cap \Gamma_1)$, and let $\mathbf{x} = (x_1, \dots, x_m) \in \Omega_{\mathbf{u}} \cap \Gamma_1$. Letting $y$ be a barycenter for $\mathbf{x}$, we have that $y \in \partial_{d_c^2} u_1(x_1)$ by Lemma 19. Then [AR04, Lemma 4.8] implies that $\partial_{d_c^2} u_1(x_1)$ is a singleton, so that $y$ is the *unique* barycenter of $\mathbf{x}$; here we used the differentiability of $u_1$ along each direction $X_j$, $Y_j$, $Z$. By our assumption (C1), then, $\mathbf{x}$ is uniquely determined by $\mathbf{b}_{\Gamma_1}^{-1}(y)$: this proves that for any $x_1 \in \pi_1(\Omega_{\mathbf{u}} \cap \Gamma_1)$ there exists exactly one such $\mathbf{x}$. Moreover, by [AR04, Lemma 4.8 and Theorem 5.1], the unique barycenter $y$ can be expressed as

$$x_1 \cdot \exp_{\mathbb{H}}(-Xu_1(x_1) - iYu_1(x_1), -Z(x_1)), \tag{39}$$

which proves our last statement.

Uniqueness of the optimal transport plan $\gamma$ and map $\Psi$ then follows by a standard argument; if $\gamma_0$ and $\gamma_1$ are both optimal, then so is $\gamma_{1/2} := \frac{1}{2}(\gamma_0 + \gamma_1)$, since (KP) is a linear minimization over a convex set. The argument above implies that $\gamma_i$ is induced by a transport map $\Psi_i$, for $i = 0, 1$; $\gamma_{1/2}$ is then concentrated on the union of these two graphs. Another application of the argument above means that $\gamma_{1/2}$ must also be concentrated on a graph; this is a contradiction unless $\Psi_0 = \Psi_1$ almost everywhere, in which case $\gamma_0 = \gamma_1$. □

REMARK 24. By tracing back our argument, one can see that only the absolute continuity of $\mu_1$ is actually required; thus Theorem 23 could be stated under this weaker assumption. However, we will soon need the absolute continuity of $\mu_2, \dots, \mu_m$ as well: for this reason, we decided to introduce this assumption already at this stage.

The real nature of the maps $\psi$ and $\mathbf{f}$ appearing in Theorem 23 is clarified by the following Theorem. Recall that the *Wasserstein barycenter* of the measures $\mu_i$ a probability measure which minimizes the functional

$$\nu \mapsto \sum_{i=1}^{m} W_2^2(\mu_i, \nu) \tag{40}$$

among all probability measures $\nu$ on $\mathbb{H}^n$, where $W_2^2(\mu_i, \nu)$ is the squared Wasserstein distance between $\mu_i$ and $\nu$:

$$W_2^2(\mu_i, \nu) := \inf \left\{ \int_{(\mathbb{H}^n)^2} d_c^2(x_i, y)\, d\gamma(x_1, y) \;\middle|\; \gamma \in \Pi(\mu_1, \nu) \right\}. \tag{41}$$

Existence of a Wasserstein barycenter follows immediately by basic continuity-compactness arguments. Absolute continuity of $\mu_1$, together with Theorem 5.1 in [AR04] imply via a standard argument (see Proposition 7.19 in [San15]) that the functional (40) is strictly convex, and so the Wasserstein barycenter is unique. The next result asserts an equivalence between the Wasserstein barycenter and multi-marginal optimal transport problems, analogous to the relationship on $\mathbb{R}^n$ due to [AC11] and its extension to Riemannian manifolds established in [KP17]. Stated simply, each optimal component mapping in (MP) pushing $\mu_1$ forward to $\mu_i$ is the composition of the two marginal optimal mapping for the quadratic cost $d_c^2$ between $\mu_1$ and the Wasserstein barycenter $\nu$, and the (two marginal quadratic cost) optimal mapping from $\nu$ to $\mu_i$.

Proposition 25. *Let the assumptions and the notations be the same as in Theorem 23. Let $\nu$ be the Wasserstein barycenter of the measures $\mu_i$ (*i.e.*, $\nu$ minimizes* (40)*). Then:*

*(i) The map $\psi : \mathbb{H}^n \to \mathbb{H}^n$ in* (37) *is the optimal map carrying $\mu_1$ to $\nu$;*
*(ii) For each $i = 2, \dots, m$, the map $\mathbf{f}_i : \mathbf{b}(\Gamma_1) \to \mathbb{H}^n$ in* (37) *is the optimal map carrying $\nu$ to $\mu_i$.*

*Proof.* By Lemma 19, $u_1$ is a $d_{\mathrm{c}}^2$ concave map; moreover, by assumption, it is differentiable along the directions $X_j$, $Y_j$, $Z$ for $\mu$-a.e. $x_1 \in \mathbb{H}^n$. The last part of [AR04, Theorem 5.1] then implies that $\psi$ is the optimal transport map between $\mu_1$ and $\psi_\sharp\mu_1$: thus we only need to show that $\psi_\sharp\mu_1 = \nu$. On the other hand, we can represent $\psi$ as $\mathbf{b}_{\Gamma_1} \circ (\mathrm{Id}, \Psi)$; thus

$$\psi_\sharp\mu_1 = (\mathbf{b}_{\Gamma_1})_\sharp\gamma, \tag{42}$$

where $\gamma$ is the optimal plan induced by $\Psi$; by a result of Carlier and Ekeland [CE10], the push-forward of the optimal plan through the barycenter map is the Wasserstein barycenter itself, thus proving our statement.

Now, assumption (C1) implies that $\mathbf{b}_{\Gamma_1}$ is invertible almost everywhere. Since this mapping is optimal between $\mu_1$ and the Wasserstein barycenter $\nu$, its inverse $\mathbf{f}_1$ is the optimal map from $\nu$ to $\mu_1$. An identical argument implies that each $f_i$ is optimal between $\nu$ and $\mu_i$. □

We next identify a simple condition on the measures $\mu_1, \dots, \mu_m$ which guarantees that the (rather complicated) assumption (C1) on $\Gamma_{\mathbf{u}}$ is satisfied. Unfortunately, the condition we impose is quite strict and leaves out some interesting cases. Though the condition is needed for technical reasons here, we suspect it can at least be weakened in some way.

The following Theorem encodes the following geometric fact: let $\mathbf{x}$ be a $m$-tuple of points in $(\mathbb{H}^n)^m$; under a suitable non-verticality assumption, if we move one of the points $x_i$ and leave the others fixed, then the barycenters also move.

Lemma 26. *Let $\mathbf{x} = (x_1, x_2, \dots, x_m)$ and $\mathbf{x}' = (\bar{x}_1, x_2, \dots, x_m)$ be two $m$-tuples having a common barycenter $y \in \mathbb{H}^n$ with respect to the CC-distance $d_{\mathrm{c}}$. If $x_i \notin y \cdot L$ for all $i \geq 2$, then $x_1 = \bar{x}_1$.*

Here we denote by $L$ the vertical axis $\{[0, s] \in \mathbb{H}^n \mid s \in \mathbb{R}\}$.

*Proof.* Let us define, for $z \in \mathbb{H}^n$:

$$\phi_i(z) \doteq d_{\mathrm{c}}^2(x_i, z) \qquad \forall i = 2, \dots, m \tag{43}$$

$$\Phi(z) = -\sum_{i=2}^{m} \phi_i(z). \tag{44}$$

Since $y$ is a barycenter for $\mathbf{x}$ and $\mathbf{x}'$, by definition of $d_{\mathrm{c}}^2$-superdifferential we have that both $x_1$ and $\bar{x}_1$ belong to $\partial_{d_{\mathrm{c}}^2}\Phi(y)$. But now each $\phi_i$ is differentiable at $y$ (by [AR04, Lemma 3.11] and by the fact that $x_i \notin y \cdot L$), thus $\Phi$ is also differentiable at $y$. By [AR04, Lemma 4.8], this implies that $\partial_{d_c^2}\Phi(y)$ is a singleton. □

Finally, we introduce a second assumption on the mesures $\mu_1, \dots, \mu_1$; one can think of it as a request for the supports of the measures to be "sufficiently far" from one another:

Definition 27. Let $\mu_1, \dots, \mu_m$ be compactly supported, absolutely continuous probability measures on $\mathbb{H}^n$. We say that $\mu_1, \dots, \mu_m$ satisfy the condition (C2) if:

$$\text{\textit{the barycenters set} } \mathbf{B} \doteq \mathbf{b}\left(\textstyle\prod_i \operatorname{spt}(\mu_i)\right) \text{ \textit{has zero} } \mu_i\text{\textit{-measure for all} } i = 1, \dots, m. \tag{C2}$$

Here

$$\mathbf{B} = \mathbf{b}\left(\prod_{i=1}^{m} \operatorname{spt}(\mu_i)\right) \doteq \left\{ y \in \mathbb{H}^n \;\middle|\; \begin{array}{c} y \text{ is a barycenter for } (x_1, \dots, x_m), \\ \text{with } x_i \in \operatorname{spt}(\mu_i) \text{ for all } i \end{array} \right\}. \tag{45}$$

In this case, we denote by $F_i \subset \mathbb{H}^n$ the $\mu_i$-zero measure set

$$F_i \doteq \mathbf{B} \cap \operatorname{spt}(\mu_i) \tag{46}$$

and by $\Omega_{\mathrm{bc}} \subset (\mathbb{H}^n)^m$ the set

$$\Omega_{\mathrm{bc}} \doteq \prod_{i=1}^{m} (\mathbb{H}^n \setminus F_i), \tag{47}$$

which has full $\mu$-measure for any transport plan $\mu \in \Pi(\mu_1, \dots, \mu_m)$.

Lemma 28. *Let $\mu_1, \dots, \mu_m$ be compactly supported, absolutely continuous probability measures on $\mathbb{H}^n$, and let $c$ be the cost associated to $d_{\mathrm{c}}$. Assume that the $\mu_i$'s satisfy condition* (C2). *Let $\mathbf{u} = (u_1, \dots, u_m)$ be a $c$-conjugate solution to the dual problem* ($\mathrm{DP}_c$). *Let $\mathbf{x}, \bar{\mathbf{x}} \in \Omega_{\mathbf{u}} \cap \Omega_{\mathrm{bc}} \cap \Gamma_{\mathbf{u}}$, where $\Omega_{\mathbf{u}}$ is the set defined in Equation* (34) *and $\Omega_{\mathrm{bc}}$ is the set defined in Equation* (47). *If $y \in \mathbb{H}^n$ is a barycenter for both $\mathbf{x}$ and $\bar{\mathbf{x}}$, then $\mathbf{x} = \bar{\mathbf{x}}$.*

*Proof.* We show that, for example, $x_1 = \bar{x}_1$ (the same argument then applies to $x_2, \dots, x_m$). As a first observation, notice that by Lemma 11 $y$ is also a barycenter for $\mathbf{x}' = (\bar{x}_1, x_2, \dots, x_m)$. Furthermore, by our assumption (C2) and our definition of $\mathbf{x}$, we have that $y \neq x_i$ for any $i$. By Lemma 26, we reach the conclusion if we can also show that $x_i \notin y \cdot L^*$ for all $i \geq 2$, where $L^* = L \setminus \{0\}$. By Lemma 19, $u_i$ is $d_{\mathrm{c}}^2$-concave for each $i$, with $y \in \partial_{d_c^2} u_i(x_i)$. Since we assumed $u_i$ to be (horizontally) differentiable at $x_i$, then we can apply [AR04, Lemma 4.7], which ensures $\partial_{d_c^2} u_i(x_i) \cap (x_i \cdot L^*) = \emptyset$. □

As a consequence, the following result can be stated:

Corollary 29. *Let $\mu_1, \dots, \mu_m$ be compactly supported, absolutely continuous probability measures on $\mathbb{H}^n$, and let $c$ be the cost associated to $d_{\mathrm{c}}$. Assume that the $\mu_i$'s satisfy condition* (C2). *Let $\mathbf{u} = (u_1, \dots, u_m)$ be a $c$-conjugate solution to the dual problem* ($\mathrm{DP}_c$). *Then the set $\Gamma_{\mathbf{u}}$ satisfies condition* (C1); *in particular, Theorem 23 holds in this case.*

Remark 30 (Gauge distance). Up to a slight modification in the definition of the set $\Omega_{\mathbf{u}}$ (Equation (34)), the proof of Theorem 23 is still valid if we consider the Gauge distance $d_{\mathrm{g}}$ on $\mathbb{H}^n$ instead of $d_{\mathrm{c}}$:

$$d_{\mathrm{g}}([\zeta,t],[\zeta',t']) \doteq \sqrt[4]{|\zeta-\zeta'|^4+(t-t')^2}. \tag{48}$$

Indeed, using the same notations and assumptions of Theorem 23, one still needs to exclude the points $x_1 \in \mathbb{H}^n$ at which $u_1$ is not differentiable, but also the points at which $X_j u_1 = Y_j u_1 = 0$ for every $j$ and $Zu_1 \neq 0$: thanks to [AR04, Lemma 4.13], the set we are excluding is still $\mathscr{L}^{2n+1}$-negligible; moreover, thanks to [AR04, Lemma 4.11], this new extended assumption again ensures that $\partial_{d_{\mathrm{g}}^2} u_1(x_1)$ is a singleton; the rest of the proof follows in the same way.

Remark 31 (Open problem: more general Carnot groups). It is possible that the techniques we exploited can be generalized to include a wider class of Carnot groups. Specifically, it seems a promising idea to adapt our argument to the case of groups of type $H$, making use of the two-marginals results contained in [Rig05].

Remark 32 (Open problem: regularity of Wasserstein barycenters). As in [AC11] and [KP17] (and as we already saw in Theorem 25), under the conditions in Theorem 23, the measure

$$\nu := \Big(x_1 \mapsto x_1 \cdot \exp_{\mathbb{H}}(-Xu_1(x_1) - iYu_1(x_1), -Zu_1(x_1))\Big)_{\#}\mu_1 \tag{49}$$

is the Wasserstein barycenter of the $\mu_i$ (with equal weights); When $m = 2$, a result of [FJ08] establishes that the barycenter is in fact absolutely continuous with respect to Lebesgue measure. The strategy of proof there has been used to establish absolute continuity of Wasserstein barycenters on Riemannian manifolds for $m \geq 3$ in [KP17]. It relies in a crucial way on the Measure Contraction Property (or, in [KP17] on a barycentric version of it). This property (for two marginals) was established on the Heisenberg group in [Jui09]; whether a barycentric version of it holds, as it does on Riemannian manifolds, is an interesting open question. If so, it would presumably allow one to prove absolute continuity of the Wasserstein barycenter for $m \geq 3$.

### 4.1. Extension to higher power costs.

Ambrosio-Rigot noted that their two marginal results extend to the case when the cost function $c = d^2$ is replaced by $d^p$ for $p \geq 2$. Similarly, in our multi-marginal case if we choose an exponent $p \geq 2$ in the definition of the cost function, *i.e.*

$$c_{d;p}(x_1,\dots,x_m) \doteq \inf_{y\in\mathbb{H}^n}\left\{\sum_{i=1}^m d^p(x_i,y)\right\}, \tag{50}$$

then everything we proved still works. Moreover, if $p > 2$, the function $y \mapsto d^p(x_i, y)$ is differentiable as long as $x_i \notin y \cdot L^*$ (that is, unlike in the $p = 2$ case, differentiability holds at $y = x_i$); this is because $d^p([0,0],[0,t]) = \pi|t|^{p/2}$, which is differentiable at $t = 0$ for $p > 2$ (the distance itself is differentiable in the horizontal directions). Therefore, $\Phi$ defined in Lemma 26 is differentiable at $y$ even if $x_i = y$ for some $i$: this means that the Theorem holds true with the weaker assumption $x_i \notin y \cdot L^*$. In particular, in this modified setting, Lemma 28 and Corollary 29 hold *without* the need of assumption (C2).

In this case, our multi-marginal optimal maps in (MP) are compositions of the two marginal optimal maps characterized in section 7 of [AR04]; the first factor pushes $\mu_1$ forward to the $p$-th barycenter of the $\mu_i$ (that is, the $\nu$ which minimizes $\nu \mapsto \sum_{i=1}^m W_p^p(\mu_i, \nu)$, where $W_p$ is the $p$-Wasserstein distance); the second pushes this $\nu$ forward to $\mu_i$.

## REFERENCES

[AC11] M. Agueh and G. Carlier, *Barycenters in the Wasserstein space*, SIAM J. Math. Anal., 43:2 (2011), pp. 904–924. MR 2801182

[AR04] L. Ambrosio and S. Rigot, *Optimal mass transportation in the Heisenberg group*, J. Funct. Anal., 208:2 (2004), pp. 261–301. MR 2035027

[AG13] L. Ambrosio and N. Gigli, *A user's guide to optimal transport*, Modelling and optimisation of flows on networks, Lecture Notes in Math., vol. 2062, Springer, Heidelberg, 2013, pp. 1–155. MR 3050280

[BLU07] A. Bonfiglioli, E. Lanconelli and F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007. MR 2363343

[Bre87] Y. Brenier, *Décomposition polaire et réarrangement monotone des champs de vecteurs*, C. R. Acad. Sci. Paris Sér. I Math., 305:19 (1987), pp. 805–808. MR 923203

[Bre91] ———, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math., 44:4 (1991), pp. 375–417. MR 1100809

[CE10] G. Carlier and I. Ekeland, *Matching for teams*, Econom. Theory, 42:2 (2010), pp. 397–418. MR 2564442

[DR11] L. De Pascale and S. Rigot, *Monge's transport problem in the Heisenberg group*, Adv. Calc. Var., 4:2 (2011), pp. 195–227. MR 2793837

[FJ08] A. Figalli and N. Juillet, *Absolute continuity of Wasserstein geodesics in the Heisenberg group*, J. Funct. Anal., 255:1 (2008), pp. 133–141. MR 2417812

[FR10] A. Figalli and L. Rifford, *Mass transportation on sub-Riemannian manifolds*, Geom. Funct. Anal., 20:1 (2010), pp. 124–159. MR 2647137

[GŚ98] W. Gangbo and A. Święch, *Optimal maps for the multidimensional Monge-Kantorovich problem*, Comm. Pure Appl. Math., 51:1 (1998), pp. 23–45. MR 1486630

[Jia17] Y. Jiang, *Absolute continuity of Wasserstein barycenters over Alexandrov spaces*, Canad. J. Math., 69:5 (2017), pp. 1087–1108. MR 3693149

[Jui09] N. Juillet, *Geometric inequalities and generalized Ricci bounds in the Heisenberg group*, Int. Math. Res. Not. IMRN, (2009), no. 13, pp. 2347–2373. MR 2520783

[KP15] Y.-H. Kim and B. Pass, *Multi-marginal optimal transport on Riemannian manifolds*, Amer. J. Math., 137:4 (2015), no. 4, pp. 1045–1060. MR 3372314

[KP17] ———, *Wasserstein barycenters over Riemannian manifolds*, Adv. Math., 307 (2017), pp. 640–683. MR 3590527

[McC01] R. J. McCann, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal., 11:3 (2001), pp. 589–608. MR 1844080

[Pas15] B. Pass, *Multi-marginal optimal transport: theory and applications*, ESAIM Math. Model. Numer. Anal., 49:6 (2015), pp. 1771–1790. MR 3423275

[Rig05] S. Rigot, *Mass transportation in groups of type H*, Commun. Contemp. Math., 7:4 (2005), pp. 509–537. MR 2166663

[San15] F. Santambrogio, *Optimal transport for applied mathematicians*, Progress in Nonlinear Differential Equations and their Applications, vol. 87, BirkhäuserSpringer, Cham, 2015, Calculus of variations, PDEs, and modeling. MR 3409718

[Ser16] F. Serra Cassano, *Some topics of geometric measure theory in Carnot groups*, Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. 1, EMS Ser. Lect. Math., Eur. Math. Soc., Zürich, 2016, pp. 1–121. MR 3587666

[Vil03] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003. MR 1964483

[Vil09] ———, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009, Old and new. MR 2459454