SUB-RIEMANNIAN INTERPOLATION INEQUALITIES

DAVIDE BARILARI[♭] AND LUCA RIZZI[♯]

ABSTRACT. We prove that ideal sub-Riemannian manifolds (i.e., admitting no non-trivial abnormal minimizers) support interpolation inequalities for optimal transport. A key role is played by sub-Riemannian Jacobi fields and distortion coefficients, whose properties are remarkably different with respect to the Riemannian case. As a byproduct, we characterize the cut locus as the set of points where the squared sub-Riemannian distance fails to be semiconvex, answering to a question raised by Figalli and Rifford in [FR10].

As an application, we deduce sharp and intrinsic Borell-Brascamp-Lieb and geodesic Brunn-Minkowski inequalities in the aforementioned setting. For the case of the Heisenberg group, we recover in an intrinsic way the results recently obtained by Balogh, Kristály and Sipos in [BKS18], and we extend them to the class of generalized H-type Carnot groups. Our results do not require the distribution to have constant rank, yielding for the particular case of the Grushin plane a sharp measure contraction property and a sharp Brunn-Minkowski inequality.

CONTENTS

[♭] INSTITUT DE MATHÉMATIQUES DE JUSSIEU-PARIS RIVE GAUCHE, UMR CNRS 7586, UNIVERSITÉ PARIS-DIDEROT, BATIMENT SOPHIE GERMAIN, CASE 7012, 75205 PARIS CEDEX 13, FRANCE
[♯] UNIV. GRENOBLE ALPES, CNRS, INSTITUT FOURIER, 38000 GRENOBLE, FRANCE
E-mail addresses: davide.barilari@imj-prg.fr, luca.rizzi@univ-grenoble-alpes.fr.

2010 *Mathematics Subject Classification.* 53C17, 49J15, 49Q20.

1. Introduction

In the seminal paper [CEMS01] it is proved that some natural inequalities holding in the Euclidean space generalize to the Riemannian setting, provided that the geometry of the ambient space is taken into account through appropriate distortion coefficients. The prototype of these inequalities in $\mathbb{R}^n$ is the Brunn-Minkowski one, or its functional counterpart in the form of Borell-Brascamp-Lieb inequality.

The main results of [CEMS01], which are purely geometrical, were originally formulated in terms of optimal transport. The theory of optimal transport (with quadratic cost) is nowadays well understood in the Riemannian setting, thanks to the works of McCann [McC01], who adapted to manifolds the theory of Brenier in the Euclidean space [Bre99]. We refer to [Vil09] for references, including a complete historical account of the theory and its subsequent developments.

Let then μ_0 and μ_1 be two probability measures on an n-dimensional Riemannian manifold (M, g). We assume μ_0, μ_1 to be compactly supported, and absolutely continuous with respect to the Riemannian measure m_g, so that $\mu_i = \rho_i \mathsf{m}_g$ for some $\rho_i \in L^1(M, \mathsf{m}_g)$. Under these assumptions, there exists a unique optimal transport map $T : M \to M$, such that $T_\sharp \mu_0 = \mu_1$ and which solves the Monge problem:

$$\int_M d^2(x, T(x)) d\mathsf{m}_g(x) = \inf_{S_\sharp \mu_0 = \mu_1} \int_M d^2(x, S(x)) d\mathsf{m}_g(x).$$

Furthermore, for μ_0-a.e. $x \in M$, there exists a unique constant-speed geodesic $T_t(x)$, with $0 \leq t \leq 1$, such that $T_0(x) = x$ and $T_1(x) = T(x)$. The map $T_t : M \to M$ defines the dynamical interpolation $\mu_t = (T_t)_\sharp \mu_0$, a curve in the space of probability measures joining μ_0 with μ_1. More precisely, $(\mu_t)_{0 \leq t \leq 1}$ is the unique Wasserstein geodesic between μ_0 and μ_1, with respect to the quadratic transportation cost.

By a well-known regularity result, μ_t is absolutely continuous with respect to m_g, that is $\mu_t = \rho_t \mathsf{m}_g$ for some $\rho_t \in L^1(M, \mathsf{m}_g)$. The fundamental result of [CEMS01] is that the concentration $1/\rho_t$ during the transportation process can be estimated with respect to its initial and final values. More precisely, for all $t \in [0, 1]$, the following *interpolation inequality* holds:

$$\frac{1}{\rho_t(T_t(x))^{1/n}} \geq \frac{\beta_{1-t}(T(x), x)^{1/n}}{\rho_0(x)^{1/n}} + \frac{\beta_t(x, T(x))^{1/n}}{\rho_1(T(x))^{1/n}}, \qquad \mu_0 - \text{a.e. } x \in M. \tag{1}$$

Here, $\beta_t(x,y)$, for $t \in [0,1]$, are *distortion coefficients* which depend only on the geometry of the underlying Riemannian manifold, and can be computed once the Riemannian structure is given, see Definition 2. The distortion coefficients are in general difficult to compute but, if $\mathrm{Ric}_g(M) \geq Kg$, then $\beta_t(x,y)$ are controlled from below by their analogues on the Riemannian space forms of constant curvature equal to K and dimension n. More precisely, we have

$$\beta_t(x,y) \geq \beta_t^{(K,n)}(x,y) = \begin{cases} t\left(\frac{\sin(t\alpha)}{\sin(\alpha)}\right)^{n-1} & \text{if } K>0, \\ t^n & \text{if } K=0, \\ t\left(\frac{\sinh(t\alpha)}{\sinh(\alpha)}\right)^{n-1} & \text{if } K<0, \end{cases} \tag{2}$$

where

$$\alpha = \sqrt{\frac{|K|}{n-1}}d(x,y).$$

Inequality (1), when expressed in terms of the reference coefficients (2), is one of the incarnations of the so-called curvature-dimensions $\mathrm{CD}(K,N)$ condition, which allows to generalize the concept of Ricci curvature bounded from below and dimension bounded from above to more general metric measure spaces. This is the beginning of the synthetic approach propugnated by Lott-Villani and Sturm [LV09, Stu06a, Stu06b] and extensively developed subsequently.

The main tools used in [CEMS01] are Jacobi fields and the properties of the cut locus, the nature of which changes dramatically in the sub-Riemannian setting. For this reason the extension of the above inequalities to the sub-Riemannian world has remained elusive (see also [Oht09] for a discussion of the Finsler case). For example, it is now well-known that the Heisenberg group equipped with a left-invariant measure, which is the simplest sub-Riemannian structure, does not satisfy any form of $\mathrm{CD}(K,N)$, as proved in [Jui09].

On the other hand, it has been recently proved in [BKS18] that the Heisenberg group actually supports interpolation inequalities as (1), with distortion coefficients whose properties are quite different with respect to the Riemannian case. The techniques in [BKS18] consist in employing a one-parameter family of Riemannian extension of the Heisenberg structure, converging to the latter as $\varepsilon \to 0$. Starting from the Riemannian interpolation inequalities, a fine analysis is required to obtain a meaningful limit for $\varepsilon \to 0$. It is important to stress that the Ricci curvature of the Riemannian extensions is unbounded from below as $\varepsilon \to 0$.

The results of [BKS18] and the extension to the corank 1 case of [BKS17] suggest that a sub-Riemannian theory of interpolation inequalities which parallels the Riemannian one actually exists. In this paper, we answer to the following question:

> *Do sub-Riemannian manifolds support weighted interpolation inequalities à la* [CEMS01]*? How to recover the correct weights and what are their properties?*

We obtain a satisfying and positive answer, at least for the so-called *ideal* structures, that is admitting no non-trivial abnormal minimizing geodesics (this is a generic property, see Proposition 14). This is the most general setting in which the sub-Riemannian transportation problem is well posed (see Section 5.1).

1.1. **Interpolation inequalities.**

To introduce our results, let $(\mathcal{D}, g)$ be a sub-Riemannian structure on a smooth manifold M, and fix a smooth reference (outer) measure m. Let us introduce the (sub-)Riemannian distortion coefficients.

Definition 1. Let $A, B \subset M$ be measurable sets, and $t \in [0,1]$. The set $Z_t(A,B)$ of *t-intermediate points* is the set of all points $\gamma(t)$, where $\gamma : [0,1] \to M$ is a minimizing geodesic such that $\gamma(0) \in A$ and $\gamma(1) \in B$.

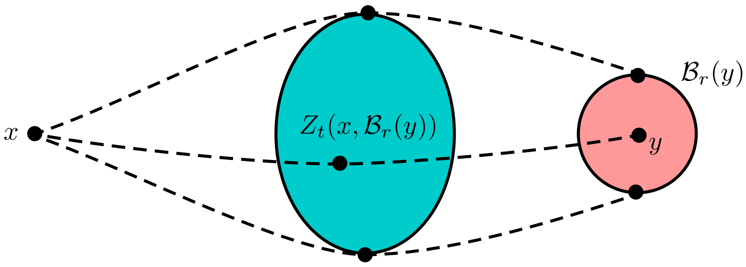

FIGURE 1. The distortion coefficient $\beta_t(x, y)$.

Let $\mathcal{B}_r(x)$ denote the sub-Riemannian ball of center $x \in M$ and radius $r > 0$.

Definition 2 (Distortion coefficient)**.** Let $x, y \in M$. The *distortion coefficient* from x to y at time $t \in [0, 1]$ is

$$\beta_t(x, y) := \limsup_{r \downarrow 0} \frac{\mathsf{m}(Z_t(x, \mathcal{B}_r(y)))}{\mathsf{m}(\mathcal{B}_r(y))}. \tag{3}$$

Notice that $\beta_0(x, y) = 0$ and $\beta_1(x, y) = 1$.

Remark 3. In the Riemannian case $\beta_t(x, y) \sim t^n$ for $t \to 0$. This universal asymptotics, valid in the Riemannian case, led [CEMS01] to extract a factor t^n in (1), expressing it in terms of the modified distortion coefficients $v_t(x, y) := \beta_t(x, y)/t^n$. The main difference is that here we do not extract a factor $1/t^n$, since the topological dimension does not describe the correct asymptotic behavior in the sub-Riemannian case (see Theorem 5). Compare also (3) with [Vil09, Def. 14.17, Prop. 14.18].

Despite the lack of a canonical Levi-Civita connection and curvature, in this paper we develop a suitable theory of sub-Riemannian (or rather Hamiltonian) Jacobi fields, which is powerful enough to derive interpolation inequalities. Our techniques are based on the approach initiated in [AZ02a, AZ02b, ZL09], and subsequently developed in a language that is closer to our presentation in [ABR13, BR17b, BR16]. Our first main result is the extension of (1) to the ideal sub-Riemannian setting.

Theorem 4 (Interpolation inequality)**.** *Let $(\mathcal{D}, g)$ be an ideal sub-Riemannian structure on M, and $\mu_0, \mu_1 \in \mathcal{P}_c^{ac}(M)$. Let $\rho_s = d\mu_s/d\mathsf{m}$. For all $t \in [0, 1]$, it holds*

$$\frac{1}{\rho_t(T_t(x))^{1/n}} \geq \frac{\beta_{1-t}(T(x), x)^{1/n}}{\rho_0(x)^{1/n}} + \frac{\beta_t(x, T(x))^{1/n}}{\rho_1(T(x))^{1/n}}, \qquad \mu_0 - \text{a.e.}\, x \in M. \tag{4}$$

If μ_1 is not absolutely continuous, an analogous result holds, provided that $t \in [0, 1)$, and that in (4) the second term on the right hand side is omitted.

Theorem 4 is proved in Section 5.3. A key role in the proof is played by a positivity lemma (cf. Lemma 29) inspired by [Vil09, Ch. 14, Appendix: Jacobi fields forever], which allows to overcome the non positive definiteness of the sub-Riemannian Hamiltonian. Moreover, with respect to previous approaches, we stress that we do not make use of any canonical frame, playing the role of a parallel transported frame.

Concerning the sub-Riemannian distortion coefficients, they can be explicitly computed in terms of sub-Riemannian Jacobi fields (cf. Lemma 44). This relation is then used in Section 7 to yield explicit formulas in different examples.

Even in the most basic examples, the distortion coefficients have some peculiar features that have no analogue in the Riemannian case. These are discussed in Section 8. Here we only give the following statement, which also allows us to introduce the important concept of geodesic dimension (for a proof see Section 8).

Theorem 5 (Asymptotics of sub-Riemannian distortion)**.** *Let $(\mathcal{D}, g)$ be a sub-Riemannian structure on M, not necessarily ideal. Let $x \in M$ and $y \notin \mathrm{Cut}(x)$. Then, there exists an integer $\mathcal{N}(x,y)$ and a constant $C(x,y) > 0$ such that*

$$\beta_t(x,y) \sim C(x,y) t^{\mathcal{N}(x,y)}, \qquad \textit{for } t \to 0^+.$$

Furthermore, for a.e. $y \notin \mathrm{Cut}(x)$, the exponent $\mathcal{N}(x,y)$ attains its minimal value

$$\mathcal{N}(x) := \min\{\mathcal{N}(x,z) \mid z \notin \mathrm{Cut}(x)\}.$$

The number $\mathcal{N}(x)$ is called the geodesic dimension *of the sub-Riemannian structure at x. Finally, the following inequality holds*

$$\mathcal{N}(x) \geq \dim(M),$$

with equality if and only if the structure is Riemannian at x, that is $\mathcal{D}_x = T_x M$.

We mention that there is an explicit formula for the geodesic dimension of a sub-Riemannian manifold of the form

$$\mathcal{N}(x) = \sum_{i=1}^{m} (2i-1)(\dim \mathcal{F}_x^i - \dim \mathcal{F}_x^{i-1}),$$

where $\mathcal{F}_x^1 \subset \cdots \subset \mathcal{F}_x^m = T_x M$ is a flag of subspaces associated to generic geodesics. This formula is reminiscent of Mitchell's formula [Mit85] for Hausdorff dimension for equiregular manifolds

$$Q = \sum_{j=1}^{r} j(\dim \mathcal{D}_x^j - \dim \mathcal{D}_x^{j-1}),$$

where $\mathcal{D}_x^1 \subset \cdots \subset \mathcal{D}_x^r = T_x M$ is the classical flag of the distribution. We stress that the two flags are different, in general. The geodesic dimension was initially discovered for sub-Riemannian structures in [ABR13, Sec. 5.6], and generalized to metric measure spaces in [Riz16], to which we refer for more details.

1.2. **Regularity of distance.** The proof of Theorem 4 is also related with the structure of the cut locus. In Riemannian geometry, it is well-known that for almost every geodesic γ involved in the transport, $\gamma(1) \notin \mathrm{Cut}(\gamma(0))$. In particular, this implies (in a non-trivial way), that the cut locus, which is defined as the set of points where the squared distance is not smooth, can be characterized actually as the set of points where the squared distance fails to be semiconvex [CEMS01].

Here, we extend the latter to the sub-Riemannian setting, answering affirmatively to the open problem raised by Figalli and Rifford in [FR10, Sec. 5.8], at least when non-trivial abnormal geodesics are not present.

Theorem 6 (Failure of semiconvexity at the cut locus)**.** *Let $(\mathcal{D}, g)$ be an ideal sub-Riemannian structure on M. Let $y \neq x$. Then $x \in \mathrm{Cut}(y)$ if and only if the squared sub-Riemannian distance from y fails to be semiconvex at x, that is, in local coordinates around x, we have*

$$\inf_{0<|v|<1} \frac{d_{SR}^2(x+v,y) + d_{SR}^2(x-v,y) - 2d_{SR}^2(x,y)}{|v|^2} = -\infty.$$

The characterization of Theorem 6 is false in the non-ideal case, as we discuss in Section 4.2. Some related open problems are proposed in Section 4.2.1.

1.3. **Geometric inequalities.** The classical consequences of interpolation inequalities follow from standard arguments. In Section 6 we discuss the the p-mean and the Borell-Brascamp-Lieb inequalities (Theorems 46 and 45, respectively). In this introduction we focus on their geometric counterpart: the Brunn-Minkowski inequality. Its classical version asserts that for measurable sets $A, B \subset \mathbb{R}^n$ one has

$$|(1-t)A + tB|^{1/n} \geq (1-t)|A|^{1/n} + t|B|^{1/n}, \qquad 0 \leq t \leq 1,$$

where $|\cdot|$ denotes the Lebesgue measure. The set $Z_t(A,B) = (1-t)A + tB$ consists of the locus of points $\gamma(t)$ as γ varies over all line segments $(1-t)x + ty$, for $0 \leq t \leq 1$ joining points $x \in A$ to points $y \in B$. We refer to [Gar02] for a comprehensive review in the Euclidean context.

To introduce the geodesic Brunn-Minkowski inequality, we define for any pair of Borel subsets $A, B \subset M$ the following quantity:

$$\beta_t(A,B) := \inf \{\beta_t(x,y) \mid (x,y) \in (A \times B) \setminus \mathrm{Cut}(M)\}, \tag{5}$$

with the convention that $\inf \emptyset = 0$. Notice that $0 \leq \beta_t(A,B) < +\infty$, cf. Lemma 44.

Theorem 7 (Sub-Riemannian Brunn-Minkowski inequality). *Let $(\mathcal{D}, g)$ be an ideal sub-Riemannian structure on a n-dimensional manifold M, equipped with a smooth measure m. Let $A, B \subset M$ be Borel subsets. Then we have*

$$\mathsf{m}(Z_t(A,B))^{1/n} \geq \beta_{1-t}(B,A)^{1/n}\mathsf{m}(A)^{1/n} + \beta_t(A,B)^{1/n}\mathsf{m}(B)^{1/n}. \tag{6}$$

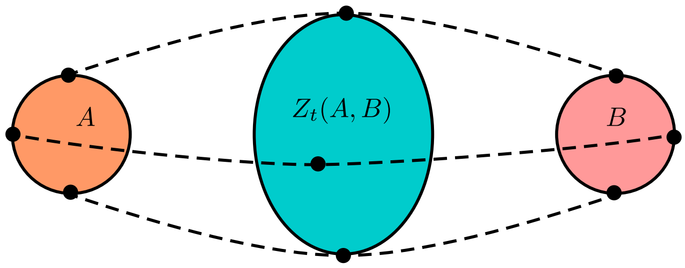

FIGURE 2. The set $Z_t(A,B)$.

Remark 8. A different generalization of the Euclidean Brunn-Minkowski inequality, at least for left-invariant structures on Lie groups, is the *multiplicative* Brunn-Minkowski inequality. The latter is defined by replacing the Minkowski sum $A + B$ with the group multiplication $A \star B$. For the Heisenberg group $\mathbb{H}_3$, with group law $\star$ and left-invariant measure m, the multiplicative Brunn-Minkowski inequality reads

$$\mathsf{m}(A \star B)^{1/d} \geq \mathsf{m}(A)^{1/d} + \mathsf{m}(B)^{1/d}, \qquad A, B \subset \mathbb{H}^3.$$

The above inequality is true for the topological dimension $d = 3$ [LM05], but false for the Hausdorff dimension $d = 4$ [Mon03].

A particular role in Theorem 7 is played by structures where $\beta_t(x,y) \geq t^N$ for some $N \in \mathbb{N}$, for all $t \in [0,1]$ and $(x,y) \notin \mathrm{Cut}(M)$. By Theorem 7, this implies the so-called *measure contraction property* $\mathrm{MCP}(0,N)$, first introduced in [Oht07] (see also [Stu06b] for a similar formulation). The MCP was first investigated in Carnot groups in [Jui09, Rif13a]. In our setting, we prove the following equivalence result.

Theorem 9 (Equivalence of inequalities). *Let $(\mathcal{D}, g)$ be an ideal sub-Riemannian structure on a n-dimensional manifold M, equipped with a smooth measure m. Let $N \geq 1$. Then, the following properties are equivalent:*

(i) $\beta_t(x,y) \geq t^N$*, for all* $(x,y) \notin \mathrm{Cut}(M)$ *and* $t \in [0,1]$*;*

(ii) the Brunn-Minkowski inequality holds: for all non-empty Borel sets A, B

$$\mathsf{m}(Z_t(A,B))^{1/n} \geq (1-t)^{N/n}\mathsf{m}(A)^{1/n} + t^{N/n}\mathsf{m}(B)^{1/n}, \qquad \forall\, t \in [0,1]; \tag{7}$$

(iii) the measure contraction property MCP$(0,N)$ *is satisfied: for all non-empty Borel sets B and $x \in M$*

$$\mathsf{m}(Z_t(x,B)) \geq t^N \mathsf{m}(B), \qquad \forall\, t \in [0,1].$$

We stress that on a n-dimensional sub-Riemannian manifold that is not Riemannian, the MCP$(0,n)$ is never satisfied (see [Riz16, Thm. 6]).

This clarifies the fact that an Euclidean Brunn-Minkowski inequality with linear weights (that is (7) with $N = n$), is not adapted for generalizations to genuine sub-Riemannian situations, as well as the classical curvature-dimension condition. We mention that generalized curvature-dimension type inequalities suitable for particular classes of sub-Riemannian structures have been developed in [BG17, BKW16].

1.4. **Old and new examples.** In Section 7, we discuss some examples, where the distortion coefficients are explicit. In particular, we consider:

- **The Heisenberg group** $\mathbb{H}_3$. In this case we recover, in an intrinsic way, the results of [BKS18], with the same distortion coefficients. See Section 7.1.
- **Generalized H-type groups**. This is a class of Carnot groups of arbitrary large corank, introduced in [BR17c], and which extends the class of Kaplan H-type groups. In the ideal case we obtain sharp interpolations inequalities for general measures (Corollary 57). In the general and possibly non-ideal case, we prove sharp Brunn-Minkowski inequalities (Corollary 59) and measure contraction properties. See Section 7.2.
- **Grushin plane** $\mathbb{G}_2$. Our techniques work also for sub-Riemannian distributions $\mathcal{D}$ whose rank is not constant. In this setting we are able to obtain for the first time interpolation inequalities (Corollary 64), sharp Brunn-Minkowski inequalities (Corollary 65), and sharp measure-contraction properties (Corollary 66). See Section 7.3.
- **Sasakian structures**. Sasakian manifolds are a particular class of contact sub-Riemannian structures. When endowed with their canonical volume, Sasakian manifolds satisfy a measure contraction property under suitable curvature lower bounds. Combining these results with Theorem 9, we get a sharp Brunn-Minkowski inequality (Corollary 67). See Section 7.4.

In all the above cases, we are able to prove that the distortion coefficients satisfy

$$\beta_t(x,y) \geq t^{\mathcal{N}}, \qquad \forall (x,y) \notin \mathrm{Cut}(M),\ \forall t \in [0,1],$$

for some minimal $\mathcal{N}$, given by the geodesic dimension of the sub-Riemannian structure (cf. Theorem 5). The interpolation inequalities take hence a very pleasant sharp form. For example in the case of the Brunn-Minkowski inequality, for all non-empty Borel sets A, B, we have

$$\mathsf{m}(Z_t(A,B))^{1/n} \geq (1-t)^{\mathcal{N}/n}\mathsf{m}(A)^{1/n} + t^{\mathcal{N}/n}\mathsf{m}(B)^{1/n}, \qquad \forall\, t \in [0,1], \tag{8}$$

In all these cases, (8) is sharp, in the sense that if one replaces the exponent $\mathcal{N}$ with a smaller one, the inequality fails for some choice of A, B.

1.5. **Carnot groups.** Another class of examples is given by ideal Carnot groups. In [Rif13a] it was proved that for any ideal Carnot group there exists $N \geq \mathcal{N}$ such that the MCP$(0,N)$ property is satisfied (see [BR17a] for the generalization to the medium-fat case). Hence we have the following result.

Corollary 10. *For any ideal Carnot group G there exists $N \geq \mathcal{N}$ such that G, equipped with its left-invariant sub-Riemannian structure and the Haar measure, satisfies the inequalities of Theorem 9.*

Open questions. Let G be a Carnot group equipped with a left-invariant sub-Riemannian structure and the Haar measure.

(i) Is it true that the Brunn-Minkowski type inequality (7) holds for some $N \in \mathbb{N}$?
(ii) Is the optimal N such that (7) holds equal to the geodesic dimension?

In question (i), if such N exists, then it is greater or equal than the geodesic dimension $\mathcal{N}$ of the Carnot group, as a consequence of the asymptotics of Theorem 5 as $t \to 0$. Indeed, a proof of the Brunn-Minkowski inequality would require a control on the distortion coefficients for all $t \in [0, 1]$.

1.6. **Afterwords.** In this work we focused in laying the groundwork for interpolation inequalities in sub-Riemannian geometry. It remains to understand which is the correct class of models whose distortion coefficients are the reference ones, playing the role of Riemannian space forms in Riemannian geometry. This will be the object of a subsequent work. We anticipate here that the natural reference spaces *do not belong* to the category of sub-Riemannian structures. The unifying framework that we propose is the one of optimal control problems. This class of variational problems is large enough to include infinitesimal models for all of the three great classes of geometries: Riemannian, sub-Riemannian, Finslerian, providing the first step of the "great unification" auspicated in [Vil17, Sec. 9]. In the spirit of [BR16], linear-quadratic optimal control problems play the role of constant curvature spaces.

Another challenging problem is to understand how to include abnormal minimizers in this picture. Abnormal geodesics, as [BKS17] suggests for the case of corank 1 Carnot groups, are not *a priori* an obstacle to interpolation inequalities. These remarkable results are the consequence of the special structure of corank 1 Carnot groups, which are the metric product of an (ideal) contact Carnot group and a suitable copy of a flat $\mathbb{R}^n$. In general, an organic theory of transport and Jacobi fields along abnormal geodesics is still lacking. In this paper, we discuss some aspects of the non-ideal case and some open problems in Section 4.2.

2. Preliminaries

We start by recalling some basic facts in sub-Riemannian geometry. For a comprehensive introduction, we refer to [ABB16b, Rif14, Mon02].

2.1. **Sub-Riemannian geometry.** A sub-Riemannian structure on a smooth, connected n-dimensional manifold M, where $n \geq 3$, is defined by a set of m global smooth vector fields $X_1, \dots, X_m$, called a *generating frame*. The *distribution* is the family of subspaces of the tangent spaces spanned by the vector fields at each point

$$\mathcal{D}_x = \operatorname{span}\{X_1(x), \dots, X_m(x)\} \subseteq T_x M, \qquad \forall\, x \in M.$$

The generating frame induces an inner product g_x on $\mathcal{D}_x$ as follows: given $v, w \in \mathcal{D}_x$ the inner product $g_x(v, w)$ is defined by the polarization formula, letting

$$g_x(v, v) := \inf \left\{ \sum_{i=1}^m u_i^2 \mid \sum_{i=1}^m u_i X_i(x) = v \right\}.$$

We assume that the distribution is *bracket-generating*, i.e., the tangent space $T_x M$ is spanned by the vector fields $X_1, \dots, X_m$ and their iterated Lie brackets evaluated at x. A *horizontal curve* $\gamma : [0,1] \to M$ is an absolutely continuous path such that there exists $u \in L^2([0,1],\mathbb{R}^m)$ satisfying

$$\dot\gamma(t) = \sum_{i=1}^m u_i(t) X_i(\gamma(t)), \qquad \text{a.e.}\, t \in [0,1].$$

This implies that $\dot\gamma(t) \in \mathcal{D}_{\gamma(t)}$ for almost every t. If γ is horizontal, the map $t \mapsto \sqrt{g(\dot\gamma(t),\dot\gamma(t))}$ is measurable on $[0,1]$, hence integrable [ABB16a, Lemma 3.11]. We define the *length* of an horizontal curve as follows

$$\ell(\gamma) = \int_0^1 \sqrt{g(\dot\gamma(t),\dot\gamma(t))}dt.$$

The *sub-Riemannian distance* is defined by:

$$d_{SR}(x,y) = \inf\{\ell(\gamma) \mid \gamma(0) = x,\, \gamma(1) = y,\, \gamma \text{ horizontal}\}. \tag{9}$$

Remark 11. The above definition includes rank-varying sub-Riemannian structures on M, see [ABB16a]. When $\dim \mathcal{D}_x$ is constant, then $\mathcal{D}$ is a vector distribution in the classical sense. If $m \leq n$ and the vector fields $X_1,\dots,X_m$ are linearly independent everywhere, they form a basis of $\mathcal{D}$ and g coincides with the inner product on $\mathcal{D}$ for which $X_1,\dots,X_m$ is an orthonormal frame.

By Chow-Rashevskii theorem, the bracket-generating condition implies that $d_{SR} : M \times M \to \mathbb{R}$ is finite and continuous. If the metric space (M, d_{SR}) is complete, then for any $x, y \in M$ the infimum in (9) is attained. In place of the length ℓ, it is often convenient to consider the *energy functional*

$$J(\gamma) = \frac{1}{2}\int_0^1 g(\dot\gamma(t),\dot\gamma(t))dt.$$

On the space of horizontal curves defined on a fixed interval and with fixed endpoints, the minimizers of J coincide with the minimizers of ℓ parametrized with constant speed. Since ℓ is invariant by reparametrization, and every horizontal curve is the reparametrization of a constant-speed one, we define *geodesics* as horizontal curves that locally minimize the energy between their endpoints.

The *Hamiltonian* of the sub-Riemannian structure $H : T^*M \to \mathbb{R}$ is defined by

$$H(\lambda) = \frac{1}{2}\sum_{i=1}^m \langle \lambda, X_i\rangle^2, \qquad \lambda \in T^*M,$$

where $X_1,\dots,X_m$ is the generating frame. Here $\langle\lambda,\cdot\rangle$ denotes the dual action of covectors on vectors. Different generating frames defining the same distribution and scalar product at each point yield the same Hamiltonian. The Hamiltonian vector field $\vec H$ is the unique vector field such that $\sigma(\cdot,\vec H) = dH$, where σ is the canonical symplectic form on T^*M. In particular, the *Hamilton equations* are

$$\dot\lambda(t) = \vec H(\lambda(t)), \qquad \lambda(t) \in T^*M. \tag{10}$$

If (M, d_{SR}) is complete, solutions of (10) are defined for all times.

2.2. **End-point map and Lagrange multipliers.** Let $\gamma_u : [0,1] \to M$ be an horizontal curve joining x and y, where $u \in L^2([0,1],\mathbb{R}^m)$ is a *control* such that

$$\dot\gamma_u(t) = \sum_{i=1}^m u_i(t) X_i(\gamma_u(t)), \qquad \text{a.e. } t \in [0,1].$$

Let $\mathcal{U} \subset L^2([0,1],\mathbb{R}^m)$ be the neighbourhood of u such that, for $v \in \mathcal{U}$, the equation

$$\dot\gamma_v(t) = \sum_{i=1}^m v_i(t) X_i(\gamma_v(t)), \qquad \gamma_v(0) = x,$$

has a well defined solution for a.e. $t \in [0,1]$. We define the *end-point map* with base point x as the smooth map $E_x : \mathcal{U} \to M$, which sends v to $\gamma_v(1)$.

We can consider $J : \mathcal{U} \to \mathbb{R}$ as a smooth functional on $\mathcal{U}$. Let γ_u be a minimizing geodesic, that is a solution of the constrained minimum problem

$$\min\{J(v) \mid v \in \mathcal{U},\, E_x(v) = y\}.$$

By the Lagrange multipliers rule, there exists a non-trivial pair (λ_1, ν), such that

$$\lambda_1 \circ D_u E_x = \nu D_u J, \qquad \lambda_1 \in T_y^* M, \qquad \nu \in \{0,1\}, \tag{11}$$

where $\circ$ denotes the composition of linear maps and D the (Fréchet) differential. If $\gamma_u : [0,1] \to M$ with control $u \in \mathcal{U}$ is an horizontal curve (not necessarily minimizing), we say that a non-zero pair $(\lambda_1, \nu) \in T_y^* M \times \{0,1\}$ is a *Lagrange multiplier* for γ_u if (11) is satisfied. The multiplier (λ_1, ν) and the associated curve γ_u are called *normal* if $\nu = 1$ and *abnormal* if $\nu = 0$. Observe that Lagrange multipliers are not unique, and a horizontal curve may be both normal *and* abnormal. Observe also that γ_u is an abnormal curve if and only if u is a critical point for E_x. In this case, γ_u is also called a *singular curve*. The following characterization is a consequence of the Lagrange multipliers rule, and can also be seen as a specification of the Pontryagin Maximum Principle to the sub-Riemannian length minimization problem.

Theorem 12. *Let $\gamma_u : [0,1] \to M$ be an horizontal curve joining x with y. A non-zero pair $(\lambda_1, \nu) \in T_y^* M \times \{0,1\}$ is a Lagrange multiplier for γ_u if and only if there exists a Lipschitz curve $\lambda(t) \in T^*_{\gamma_u(t)} M$ with $\lambda(1) = \lambda_1$, such that*

- *(N) if $\nu = 1$ then $\dot\lambda(t) = \vec{H}(\lambda(t))$,*
- *(A) if $\nu = 0$ then $\sigma(\dot\lambda(t), T_{\lambda(t)}\mathcal{D}^\perp) = 0$,*

where $\mathcal{D}^\perp \subset T^*M$ is the set of covectors that annihilate the distribution.

In the first (resp. second) case, $\lambda(t)$ is called a *normal* (resp. *abnormal*) *extremal*. Normal extremals are integral curves $\lambda(t)$ of $\vec{H}$. As such, they are smooth, and characterized by their *initial covector* $\lambda = \lambda(0)$. A geodesic is normal (resp. abnormal) if admits a normal (resp. abnormal) extremal. On the other hand, it is well-known that the projection $\gamma_\lambda(t) = \pi(\lambda(t))$ of a normal extremal is locally minimizing, hence it is a normal geodesic. The *exponential map* at $x \in M$ is the map $\exp_x : T_x^* M \to M$, which assigns to $\lambda \in T_x^* M$ the final point $\pi(\lambda(1))$ of the corresponding normal geodesic. The curve $\gamma_\lambda(t) = \exp_x(t\lambda)$, for $t \in [0,1]$, is the normal geodesic corresponding to λ, which has constant speed $\|\dot\gamma_\lambda(t)\| = \sqrt{2H(\lambda)}$ and length $\ell(\gamma|_{[t_1,t_2]}) = \sqrt{2H(\lambda)}(t_2 - t_1)$.

Definition 13. A sub-Riemannian structure $(\mathcal{D}, g)$ on M is *ideal* if the metric space (M, d_{SR}) is complete and there exists no non-trivial abnormal length minimizers.[1]

The above terminology was introduced in [Rif13b, Rif14]. By [Mon02, Sec. 5.6], all complete fat structures are ideal. Moreover, the ideal assumption is generic, when the rank of the distribution is at least 3, in the following sense.

Proposition 14 ([CJT06, Thm. 2.8])**.** *Let $k \geq 3$ be a positive integer, and $\mathcal{G}_k$ be the set of sub-Riemannian structures $(\mathcal{D}, g)$ on M with $\operatorname{rank}\mathcal{D} = k$, endowed with the Whitney C^∞ topology. There exists an open dense subset W_k of $\mathcal{G}_k$ such that every element of W_k does not admit non-trivial abnormal minimizers.*

Next, we recall the definition of conjugate points.

[1] This means that the only possible abnormal length minimizers are constant curves.

Definition 15. Let $\gamma : [0,1] \to M$ be a normal geodesic with initial covector $\lambda \in T_x^*M$, that is $\gamma(t) = \exp_x(t\lambda)$. We say that $y = \exp_x(\bar{t}\lambda)$ is a *conjugate point* to x along γ if $\bar{t}\lambda$ is a critical point for $\exp_x$.

Given a normal geodesic $\gamma : [0,1] \to M$ and $0 \leq s < t \leq 1$, we say that $\gamma(s)$ and $\gamma(t)$ are *conjugate* if $\gamma(t)$ is conjugate to $\gamma(s)$ along $\gamma|_{[s,t]}$.

In the Riemannian setting, conjugate points along a geodesic are isolated, and geodesics cease to be minimizers after the first conjugate point. In the general sub-Riemannian setting, the picture is more complicated, but this result remains valid for geodesics that do not contain abnormal segments.

Definition 16. A normal geodesic $\gamma : [0,1] \to M$ *contains no abnormal segments* if for every $0 \leq s_1 < s_2 \leq 1$ the restriction $\gamma|_{[s_1,s_2]}$ is not abnormal.

Theorem 17 (Conjugate points and minimality)**.** *Let $\gamma : [0,1] \to M$ be a minimizing geodesic which does not contain abnormal segments. Then $\gamma(s)$ is not conjugate to $\gamma(s')$ for every $s, s' \in [0,1]$ with $|s - s'| < 1$.*

Theorem 17 is not new, and well-known to experts. We provide a self-contained proof in Appendix A, following the arguments of [ABB16b] (see also [Sar80]). Notice that, as in the Riemannian case, $\gamma(1)$ can be conjugate to $\gamma(0)$ along γ.

2.3. **Regularity of the sub-Riemannian distance.** We recall now some basic regularity properties of the sub-Riemannian distance.

Definition 18. Let $(\mathcal{D}, g)$ be a complete sub-Riemannian structure on M, and $x \in M$. We say that $y \in M$ is a *smooth point* (with respect to x) if there exists a unique minimizing geodesic joining x with y, which is not abnormal, and the two points are not conjugate along such a curve. The *cut locus* $\mathrm{Cut}(x)$ is the complement of the set of smooth points with respect to x. The *global cut-locus* of M is

$$\mathrm{Cut}(M) := \{(x, y) \in M \times M \mid y \in \mathrm{Cut}(x)\}.$$

We have the following fundamental result [Agr09, RT05].

Theorem 19. *The set of smooth points is open and dense in M, and the squared sub-Riemannian distance is smooth on $M \times M \setminus \mathrm{Cut}(M)$.*

3. Jacobi fields and second differential

If $x \in M$ is a critical point for $f \in C^\infty(M)$, one can define the Hessian of f as

$$\mathrm{Hess}(f)|_x : T_xM \times T_xM \to \mathbb{R}, \qquad \mathrm{Hess}(f)|_x(v, w) = V(W(f))(x),$$

where V, W are local vector fields such that $V(x) = v$ and $W(x) = w$. Since x is a critical point, the definition is well posed, and $\mathrm{Hess}(f)|_x$ is a symmetric bilinear map. The quadratic form associated with the second differential of f at x which, for simplicity, we denote by the same symbol $\mathrm{Hess}(f)|_x : T_xM \to \mathbb{R}$, is

$$\mathrm{Hess}(f)|_x(v) = \left.\frac{d^2}{dt^2}\right|_{t=0} f(\gamma(t)), \qquad \gamma(0) = x, \quad \dot{\gamma}(0) = v.$$

When $x \in M$ is not a critical point, we define the second differential of f as the differential of df, thought as a smooth section of T^*M.

Definition 20 (Second differential at non-critical points)**.** Let $f \in C^\infty(M)$, and

$$df : M \to T^*M, \qquad df : x \mapsto d_xf.$$

The *second differential* of f at $x \in M$ is the linear map

$$d_x^2 f := d_x(df) : T_xM \to T_\lambda(T^*M),$$

where $\lambda = d_x f \in T^*M$.

The image of $df : M \to T^*M$ is a Lagrangian[2] submanifold of T^*M. Thus, the image of the second differential $d_x^2 f(T_xM)$ at a point x is the tangent space of $df(M)$ at $\lambda = d_x f$, which is an n-dimensional Lagrangian subspace of $T_\lambda(T^*M)$ transverse to $T_\lambda(T_x^*M)$. Letting $\pi : T^*M \to M$ be the cotangent bundle projection, and since $\pi \circ df = \mathrm{id}_M$, we have that $\pi_* \circ d_x^2 f = \mathrm{id}_{T_xM}$.

Lemma 21. *Let $\lambda \in T_x^*M$. The set $\mathcal{L}_\lambda := \{d_x^2 f \mid f \in C^\infty(M),\, d_x f = \lambda\}$ is an affine space over the vector space of quadratic forms on T_xM.*

The above lemma follows from the fact that if $f_1, f_2 \in C^\infty(M)$ satisfy $d_x f_1 = d_x f_2 = \lambda$, then x is a critical point for $f_1 - f_2$ and one can define the difference between $d_x^2 f_1$ and $d_x^2 f_2$ as the quadratic form $\mathrm{Hess}(f_1 - f_2)|_x$.

Remark 22. Definition 20 can be extended to any $f : M \to \mathbb{R}$ twice differentiable at x. In this case, fix local coordinates around x, and let $b(x) \in \mathbb{R}^n$ and $A(x) \in \mathrm{Sym}(n \times n)$ such that

$$\lim_{t \downarrow 0} \frac{f(x+tv) - f(x) - tb(x) \cdot v - \frac{t^2}{2} v \cdot A(x) v}{t^2} = 0, \qquad \forall v \in \mathbb{R}^n.$$

Letting $(q,p) \in \mathbb{R}^{2n}$ denote canonical coordinates around $d_x f \in T^*M$, we define

$$d_x^2 f\,(\partial_{q_i}) := \partial_{q_i}|_{d_x f} + \sum_{j=1}^n A_{ij} \partial_{p_j}|_{d_x f}, \qquad \forall i = 1, \dots, n.$$

This definition is well posed, i.e., it does not depend on the choice of coordinates.

3.1. **Sub-Riemannian Jacobi fields.** Let $\lambda_t = e^{t\vec{H}}(\lambda_0)$, $t \in [0,1]$ be an integral curve of the Hamiltonian flow. For any smooth vector field $\xi(t)$ along λ_t, the dot denotes the Lie derivative in the direction of $\vec{H}$, namely

$$\dot{\xi}(t) := \left.\frac{d}{d\varepsilon}\right|_{\varepsilon=0} e_*^{-\varepsilon \vec{H}} \xi(t+\varepsilon).$$

A vector field $\mathcal{J}(t)$ along λ_t is a *Jacobi field* if it satisfies the equation

$$\dot{\mathcal{J}} = 0. \tag{12}$$

Jacobi fields along λ_t are of the form $\mathcal{J}(t) = e_*^{t\vec{H}} \mathcal{J}(0)$, for some unique initial condition $\mathcal{J}(0) \in T_{\lambda_0}(T^*M)$, and the space of solutions of (12) is a $2n$-dimensional vector space. On T^*M we define the smooth sub-bundle with Lagrangian fibers:

$$\mathcal{V}_\lambda := \ker \pi_*|_\lambda = T_\lambda(T^*_{\pi(\lambda)}M) \subset T_\lambda(T^*M), \qquad \lambda \in T^*M,$$

which we call the *vertical subspace.* In this formalism, letting

$$\gamma(t) = \exp_x(t\lambda_0) = \pi \circ e^{t\vec{H}}(\lambda_0), \qquad t \in [0,1],$$

we have that $\gamma(s)$ is conjugate with $\gamma(0)$ along the normal geodesic γ if and only if the Lagrangian subspace $e_*^{s\vec{H}} \mathcal{V}_{\lambda_0} \subset T_{\lambda_s}(T^*M)$ intersects $\mathcal{V}_{\lambda_s}$ non-trivially.

The next statement generalizes the well known Riemannian fact that, in absence of conjugate points, Jacobi fields are either determined by their value and the value of the covariant derivative in the direction of the given geodesic at the initial time, or by their value at the final and initial times.

[2] A Lagrangian submanifold of T^*M is a submanifold such that its tangent space is a Lagrangian subspace of the symplectic space $T_\lambda(T^*M)$. A subspace $L \subset \Sigma$ of a symplectic vector space (Σ, σ) is Lagrangian if $\dim L = \dim \Sigma/2$ and $\sigma|_L = 0$.

Lemma 23. *Assume that, for $s \in (0,1]$, $\gamma(0)$ is not conjugate to $\gamma(s)$ along γ. Let $\mathcal{H}_{\lambda_i} \subset T_{\lambda_i}(T^*M)$ be transverse to $\mathcal{V}_{\lambda_i}$, for $i = 0, s$. Then for any pair $(J_0, J_s) \in \mathcal{H}_{\lambda_0} \times \mathcal{H}_{\lambda_s}$, there exists a unique Jacobi field $\mathcal{J}(t)$ along λ_t, $t \in [0,1]$, such that the projection of $\mathcal{J}(i)$ on $\mathcal{H}_{\lambda_i}$ is equal to J_i, for $i = 0, s$.*

Proof. The condition at $t = 0$ implies that $\mathcal{J}(0) \in J_0 + \mathcal{V}_{\lambda_0}$ (an affine space). By definition of Jacobi field, $\mathcal{J}(t) = e^{t\vec{H}}_* \mathcal{J}(0)$, in particular $\mathcal{J}(s) \in e^{s\vec{H}}_* J_0 + e^{s\vec{H}}_* \mathcal{V}_{\lambda_0}$. By the non-conjugate assumption and since $T_{\lambda_s}(T^*M) = \mathcal{V}_{\lambda_s} + \mathcal{H}_{\lambda_s}$, the projection of the affine space $e^{s\vec{H}}_* J_0 + e^{s\vec{H}}_* \mathcal{V}_{\lambda_0}$ on $\mathcal{H}_{\lambda_s}$ is a bijection, yielding the statement. □

3.2. **Jacobi matrices.** We introduce a formalism to describe families of subspaces generated by Jacobi fields. Let $\gamma : [0,1] \to M$ be a normal geodesic, projection of $\lambda_t = e^{t\vec{H}}(\lambda_0)$, for some $\lambda_0 \in T^*M$. Consider the family of n-dimensional subspaces generated by a set of independent Jacobi fields $\mathcal{J}_1(t), \dots, \mathcal{J}_n(t)$ along λ_t, that is

$$\mathcal{L}_t = \operatorname{span}\{\mathcal{J}_1(t), \dots, \mathcal{J}_n(t)\} \subset T_{\lambda_t}(T^*M).$$

Since $\mathcal{L}_t = e^{t\vec{H}}_* \mathcal{L}_0$, then $\mathcal{L}_t$ is Lagrangian if and only if it is Lagrangian at time $t = 0$.

Notice that $\mathcal{L}_t$ can be regarded as a smooth curve in a suitable (Lagrange) Grassmannian bundle over T^*M. We do not pursue this approach here, and we opt for an extrinsic formulation based on Darboux frames. To this purpose, and in order to exploit the symplectic structure of T^*M, fix a Darboux moving frame along λ_t, that is a collection of smooth vector fields $E_1(t), \dots, E_n(t), F_1(t), \dots, F_n(t)$ such that

$$\sigma(E_i, F_j) - \delta_{ij} = \sigma(E_i, E_j) = \sigma(F_i, F_j) = 0, \qquad \forall i, j = 1, \dots, n,$$

and such that the $E_1(t), \dots, E_n(t)$ generate the vertical subspace $\mathcal{V}_{\lambda_t} = \ker \pi_*|_{\lambda_t}$:

$$\mathcal{V}_{\lambda_t} = \operatorname{span}\{E_1(t), \dots, E_n(t)\}.$$

We also denote with $X_i(t) := \pi_* F_i(t)$ the corresponding moving frame along the geodesic γ. In this case, we say that $E_i(t), F_i(t)$ is a *Darboux lift* of $X_i(t)$. Notice that any smooth moving frame along a normal geodesic admits a Darboux lift along a corresponding normal extremal.

We identify $\mathcal{L}_t = \operatorname{span}\{\mathcal{J}_1(t), \dots, \mathcal{J}_n(t)\}$ with a smooth family of $2n \times n$ matrices

$$\mathbf{J}(t) = \begin{pmatrix} M(t) \\ N(t) \end{pmatrix}, \qquad t \in [0,1],$$

such that, with respect to the given Darboux frame, we have

$$\mathcal{J}_i(t) = \sum_{j=1}^n E_j(t) M_{ji}(t) + F_j(t) N_{ji}(t), \qquad \forall i = 1, \dots, n. \tag{13}$$

We call $\mathbf{J}(t)$ a *Jacobi matrix*, while the $n \times n$ matrices $M(t)$ and $N(t)$ represent respectively its "vertical" and "horizontal" components with respect to the decomposition induced by the Darboux moving frame

$$T_{\lambda_t}(T^*M) = \mathcal{H}_{\lambda_t} \oplus \mathcal{V}_{\lambda_t}, \qquad \text{with} \qquad \mathcal{H}_{\lambda_t} := \operatorname{span}\{F_1(t), \dots, F_n(t)\}.$$

The following property is fundamental for the following.

Lemma 24. *There exist smooth families of matrices $A(t), B(t), R(t)$, $t \in [0,1]$, with $B(t), R(t)$ symmetric and $B(t) \geq 0$, such that for any Jacobi matrix $\mathbf{J}(t)$, we have*

$$\frac{d}{dt} \begin{pmatrix} M \\ N \end{pmatrix} = \begin{pmatrix} -A(t) & -R(t) \\ B(t) & A(t)^* \end{pmatrix} \begin{pmatrix} M \\ N \end{pmatrix}. \tag{14}$$

On any interval $I \subseteq [0,1]$ such that $M(t)$ is non-degenerate, the matrix $W(t) := N(t)M(t)^{-1}$ satisfies the Riccati equation

$$\dot{W} = B(t) + A(t)^* W + W A(t) + W R(t) W.$$

The associated family of subspaces $\mathcal{L}_t$ is Lagrangian if and only if $W(t)$ is symmetric.

Proof. By completeness of the frame, there exist smooth matrices $A(t), B(t), C(t)$ and $R(t)$ such that, for all $t \in [0,1]$, it holds

$$\dot{E} = E \cdot A(t) - F \cdot B(t), \qquad \dot{F} = E \cdot R(t) - F \cdot C(t)^*. \tag{15}$$

The notation in (15) means that $\dot{E}_i = \sum_{j=1}^n E_j A(t)_{ji} - F_j B(t)_{ji}$, and similarly for $\dot{F}_i$. For n-tuples V, W, the pairing $\sigma(V,W)$ denotes the matrix $\sigma(V_i, W_j)$. In this notation, $\sigma(V,W)^* = -\sigma(W,V)$. Thanks to the Darboux condition, we obtain

$$C(t) = \sigma_{\lambda_t}(\dot{F}, E) = -\sigma_{\lambda_t}(F, \dot{E}) = A(t).$$

The symmetry of $R(t)$ and $B(t)$ follows similarly. Moreover, we have

$$B(t) = \sigma_{\lambda_t}(\dot{E}, E) = 2H(E,E) \geq 0. \tag{16}$$

Here, H is the Hamiltonian seen as a fiber-wise bilinear form on T^*M, and we identify $T^*_{\gamma_t}M \simeq T_{\lambda_t}(T^*_{\gamma_t}M)$. The second equality in (16) follows from a direct computation in canonical coordinates on T^*M. Observe that $B(t)$ has a non-trivial kernel if and only if the structure is not Riemannian. Finally, equation (14) follows from (15), (13) and the Jacobi equation $\dot{\mathcal{J}}_i(t) = 0$. The claim about Riccati equation is proved by direct verification.

Using (13), the Jacobi fields $\mathcal{J}_1(t), \dots, \mathcal{J}_n(t)$ associated with the Jacobi matrix $\mathbf{J}(t)$ generate a family of Lagrangian subspaces if and only if

$$0 = \sigma_{\lambda_t}(\mathcal{J}, \mathcal{J}) = M(t)^* N(t) - N(t)^* M(t).$$

The above identity is equivalent to the symmetry of $W(t)$. $\square$

In Riemannian geometry, the forthcoming manipulations are greatly simplified thanks to the existence of a particular Darboux frame, such that $A(t) = \mathbb{0}$, $B(t) = \mathbb{1}$ and where $R(t)$ represents the Riemannian sectional curvature of all 2-planes containing $\dot{\gamma}(t)$. In the sub-Riemannian case, such a convenient frame is not available in full generality. To circumvent this problem we "lift" the problem on the cotangent bundle and avoid to pick some particular frame. See [BR16, BR17b] for details.

3.3. Special Jacobi matrices.

3.3. Special Jacobi matrices. Fix a normal geodesic $\gamma : [0,1] \to M$ and a smooth moving frame $E_1(t), \dots, E_n(t), F_1(t), \dots, F_n(t)$ along the corresponding extremal. Any Jacobi matrix is uniquely defined by its value at some intermediate time $\mathbf{J}(s)$, for $s \in [0,1]$. The following special Jacobi matrices will play a prominent role:

$$\mathbf{J}_s^{\mathrm{V}}(t) = \begin{pmatrix} M_s^{\mathrm{V}}(t) \\ N_s^{\mathrm{V}}(t) \end{pmatrix}, \quad \text{such that} \quad \mathbf{J}_s^{\mathrm{V}}(s) = \begin{pmatrix} \mathbb{1} \\ \mathbb{0} \end{pmatrix}, \qquad (\text{"vertical" at time } s),$$

$$\mathbf{J}_s^{\mathrm{H}}(t) = \begin{pmatrix} M_s^{\mathrm{H}}(t) \\ N_s^{\mathrm{H}}(t) \end{pmatrix}, \quad \text{such that} \quad \mathbf{J}_s^{\mathrm{H}}(s) = \begin{pmatrix} \mathbb{0} \\ \mathbb{1} \end{pmatrix}, \qquad (\text{"horizontal" at time } s),$$

representing, respectively, the families of Lagrange subspaces

$$e_*^{(t-s)\vec{H}} \operatorname{span}\{E_1(s), \dots, E_n(s)\} \qquad \text{and} \qquad e_*^{(t-s)\vec{H}} \operatorname{span}\{F_1(s), \dots, F_n(s)\}.$$

Remark 25. Let $s_1, s_2 \in [0,1]$. Then $\gamma(s_1)$ is conjugate to $\gamma(s_2)$ along γ if and only if at least one (and then both) the matrices $N_{s_1}^{\mathrm{V}}(s_2)$ and $N_{s_2}^{\mathrm{V}}(s_1)$ are degenerate.

4. Main Jacobian estimate

We now state our two main technical results, which will be proved together.

Theorem 26. *Let $(\mathcal{D}, g)$ be a sub-Riemannian structure on M. Let $x \neq y \in M$ and assume that there exists a function $\phi : M \to \mathbb{R}$, twice differentiable at x, such that*

$$\frac{1}{2} d^2_{SR}(x,y) = -\phi(x), \qquad \text{and} \qquad \frac{1}{2} d^2_{SR}(z,y) \geq -\phi(z), \quad \forall z \in M. \tag{17}$$

Assume moreover that the minimizing curve joining x with y, which is unique and given by $\gamma(t) = \exp_x(t d_x \phi)$, does not contain abnormal segments. Then $x \notin \mathrm{Cut}(y)$.

We will usually apply Theorem 26 to situations in which ϕ is twice differentiable almost everywhere, in such a way that the map

$$T_t(z) = \exp_z(t d_z \phi), \qquad \mathsf{m}-\text{a.e. } z \in M,$$

is well defined. The next result is an estimate for its Jacobian determinant at x.

Theorem 27 (Main Jacobian estimate)**.** *Under the same hypotheses of Theorem 26, let $\gamma(t) = \exp_x(t d_x \phi)$, with $t \in [0,1]$, be the unique minimizing curve joining x with y, which does not contain abnormal segments. Then, the linear maps*

$$d_x T_t : T_x M \to T_{\gamma(t)} M, \qquad d_x T_t := \pi_* \circ e^{t\vec{H}}_* \circ d^2_x \phi,$$

satisfy the following estimate, for all fixed $s \in (0,1]$:

$$\det(d_x T_t)^{1/n} \geq \left(\frac{\det N^{\mathrm{v}}_s(t)}{\det N^{\mathrm{v}}_s(0)} \right)^{1/n} + \left(\frac{\det N^{\mathrm{v}}_0(t)}{\det N^{\mathrm{v}}_0(s)} \right)^{1/n} \det(d_x T_s)^{1/n}, \quad \forall t \in [0,s], \tag{18}$$

where the determinants are computed with respect to some smooth moving frame along γ, and the matrices $N^{\mathrm{v}}_s(t)$ are defined as in Section 3.3, with respect to some Darboux lift along the corresponding extremal.

Both terms in the right hand side of (18) are non-negative for $t \in [0,s]$ and, for $t \in [0,s)$, the first one is positive. In particular $\det(d_x T_t) > 0$ for all $t \in [0,1)$.

Remark 28. As a matter of fact, $\det(d_x T_1)$ can be zero. For example, fix $x \notin \mathrm{Cut}(y)$. The assumptions of Theorem 27 are satisfied by any smooth function ϕ such that $\phi(z) = -d^2_{SR}(z,y)/2$ for all z in a neighbourhood $\mathcal{O}_x$ of x. In particular $\exp_z(d_z\phi) = \pi \circ e^{\vec{H}}(d_z \phi) = y$ for all $z \in \mathcal{O}_x$, and thus $d_x T_1 = 0$.

We first discuss the strategy of the proof of Theorems 26 and 27. It is well known that, if (17) holds and ϕ is differentiable at x, there exists a unique minimizing curve joining x with y, which is the normal geodesic $\gamma(t) = \exp_x(t d_x \phi)$, $t \in [0,1]$, see e.g. [Rif14, Lemma 2.15]. By Theorem 17, there are no conjugate points along γ, except possibly the pair $\gamma(0)$ and $\gamma(1)$. Thanks to this observation, we first prove that (18) holds for all $s < 1$. Then, we prove that if $\gamma(1)$ is conjugate to $\gamma(0)$, the right hand side of (18) tends to $+\infty$ for $s \uparrow 1$ and any fixed $t > 0$, hence $\det(d_x T_t)^{1/n} = +\infty$, leading to a contradiction. This implies that $\gamma(1)$ is not conjugate to $\gamma(0)$, yields the validity of (18) for all $s \in (0,1]$, and proves that $y \notin \mathrm{Cut}(x)$.

Proof of Theorems 26 and 27. Let $\lambda(t) := e^{t\vec{H}}(d_x\phi)$, and $\gamma(t) = \pi(\lambda(t))$ the corresponding minimizing geodesic, with $t \in [0,1]$. Let $E_1(t), \dots, E_n(t), F_1(t), \dots, F_n(t)$ be a Darboux lift along $\lambda(t)$ of a smooth moving frame $X_1(t), \dots, X_n(t)$ along $\gamma(t)$, that is satisfying

$$\sigma(E_i, F_j) - \delta_{ij} = \sigma(E_i, E_j) = \sigma(F_i, F_j) = 0, \qquad \forall i,j = 1, \dots, n,$$

with $X_i(t) = \pi_* F_i(t)$ and $\pi_* E_i(t) = 0$ for $i = 1, \dots, n$ and $t \in [0,1]$.

Since ϕ is twice differentiable at x, the family of Lagrangian subspaces $e^{t\vec{H}}_* \circ d^2_x\phi(T_xM) \subset T_{\lambda(t)}(T^*M)$ is well defined for all $t \in [0,1]$, and is associated via the given Darboux frame to the Jacobi matrix

$$\mathbf{J}(t) = \begin{pmatrix} M(t) \\ N(t) \end{pmatrix}, \qquad \text{such that} \qquad e^{t\vec{H}}_* \circ d^2_x\phi(X(0)) = E(t)\cdot M(t) + F(t) \cdot N(t).$$

In particular, $d_xT_t(X(0)) = X(t)\cdot N(t)$. Let now $s \in (0,1)$, and consider the Jacobi matrices $\mathbf{J}^{\mathrm{V}}_0$ and $\mathbf{J}^{\mathrm{V}}_s$ of Section 3.3. Since $\gamma(0)$ is not conjugate to $\gamma(s)$, we have

$$e^{s\vec{H}}_*\mathcal{V}_{\lambda(0)} \cap \mathcal{V}_{\lambda(s)} = \{0\}, \qquad \forall s \in (0,1). \tag{19}$$

Equivalently, $N^{\mathrm{V}}_s(0)$ and $N^{\mathrm{V}}_0(s)$ are invertible. By Lemma 23, a Jacobi matrix is uniquely specified by its horizontal components $N(0)$ and $N(s)$, hence we have

$$\mathbf{J}(t) = \mathbf{J}^{\mathrm{V}}_s(t)N^{\mathrm{V}}_s(0)^{-1}N(0) + \mathbf{J}^{\mathrm{V}}_0(t)N^{\mathrm{V}}_0(s)^{-1}N(s), \qquad t\in[0,1], \quad s \in (0,1). \tag{20}$$

By construction $N(0) = \mathbb{1}$, and the horizontal component of (20) reads

$$N(t) = N^{\mathrm{V}}_s(t)N^{\mathrm{V}}_s(0)^{-1} + N^{\mathrm{V}}_0(t)N^{\mathrm{V}}_0(s)^{-1}N(s), \qquad t\in[0,1], \quad s \in (0,1). \tag{21}$$

The next crucial lemma is a consequence of two facts: the non-negativity of the Hamiltonian, and assumption (17). We postpone its proof to Appendix B.

Lemma 29 (Positivity)**.** *Under the assumptions of Theorem 26, there exists a smooth family of $n\times n$ matrices $K(t) = N^{\mathrm{V}}_0(t)^{-1}$, defined for $t \in (0,1)$, such that, for all $s \in (0,1)$, we have*

- *(a)* $\det K(t) > 0$ *for all $t \in (0,1)$,*
- *(b)* $K(t)N^{\mathrm{V}}_s(t)N^{\mathrm{V}}_s(0)^{-1} \geq 0$*, for all $t \in (0,s]$,*
- *(c)* $K(t)N^{\mathrm{V}}_0(t)N^{\mathrm{V}}_0(s)^{-1}N(s) \geq 0$*, for all $t \in (0,1)$.*

Furthermore, if $\gamma(1)$ is not conjugate to $\gamma(0)$, the above properties hold for all $s \in (0,1]$ and $t \in (0,1]$.

Minkowski determinant theorem [MM92, 4.1.8] states that the function $A \mapsto (\det A)^{1/n}$ is concave on the set of $n \times n$ non-negative symmetric matrices. Thus, by multiplying from the left (21) by the matrix $K(t)$ of Lemma 29, we obtain

$$\det(d_xT_t)^{1/n} \geq \left(\frac{\det N^{\mathrm{V}}_s(t)}{\det N^{\mathrm{V}}_s(0)}\right)^{1/n} + \left(\frac{\det N^{\mathrm{V}}_0(t)}{\det N^{\mathrm{V}}_0(s)}\right)^{1/n}\det(d_xT_s)^{1/n}, \quad t \in [0,s]. \tag{22}$$

Notice that we do not use Lemma 29 to prove (22) for $t = 0$, but in this case the inequality holds since $d_xT_0 = \mathrm{id}|_{T_xM}$ and $N^{\mathrm{V}}_0(0) = \mathbb{0}$. Hence, we obtain (22) for all $t \in [0,s]$, $s \in (0,1)$ and, if $\gamma(0)$ is not conjugate with $\gamma(1)$, also for $s = 1$. We claim that, under the assumptions of Theorem 26, the latter is always the case.

By contradiction, assume that $\gamma(1)$ is conjugate to $\gamma(0)$. As we already remarked, $N^{\mathrm{V}}_0(s)$ and $N^{\mathrm{V}}_s(0)$ are non-degenerate for $s \in (0,1)$, but now $\det N^{\mathrm{V}}_0(1) = \det N^{\mathrm{V}}_1(0) = 0$. We claim that, for fixed $t\in(0,1)$, the right hand side of (22) tends to $+\infty$ for $s \uparrow 1$. To prove this claim, notice that both terms in the right hand side of (22) are non-negative thanks to Lemma 29, and therefore

$$\det(d_xT_t)^{1/n} \geq \left(\frac{\det N^{\mathrm{V}}_s(t)}{\det N^{\mathrm{V}}_s(0)}\right)^{1/n} \geq 0, \qquad t \in [0,s]. \tag{23}$$

By Theorem 17, $\gamma(t)$ is not conjugate to $\gamma(1)$ for any fixed $t \in (0,1)$. Hence $N^{\mathrm{V}}_1(t)$ is not degenerate. On the other hand $\gamma(0)$ and $\gamma(1)$ are conjugate by our assumption, and $N^{\mathrm{V}}_1(0)$ is degenerate. Taking the limit for $s\uparrow 1$, and since the left hand side of (23) does not depend on s, we obtain $\det(d_xT_t)^{1/n} = +\infty$, leading to a contradiction.

We have so far proved that there is a unique minimizing geodesic joining x with y, which is not abnormal, and y is not conjugate to x. This means that $y \notin \mathrm{Cut}(x)$, and concludes the proof of Theorem 26. Moreover, (19)-(21) hold for all $s \in (0,1]$ and $t \in [0,1]$. Therefore we can apply Lemma 29 also for $s=1$, which completes the proof of (18) for all $s \in (0,1]$ and $t \in [0,s]$.

We already proved that both terms in the right hand side of (18) are non-negative for $t \in [0,s]$ (actually, the second one is non-negative for all $t \in [0,1]$, by part (c) of Lemma 29). Now that we also proved that $\gamma(t)$ is not conjugate to $\gamma(s)$ for all possible $0 < |s-t| \leq 1$, we obtain that the first term cannot be zero except for $t = s$, and hence it is strictly positive for all $t \in [0,s)$. □

4.1. **Failure of semiconvexity at the cut locus: proof of Theorem 6.** We say that a continuous function $f : M \to \mathbb{R}$ fails to be semiconvex at $x \in M$ if, in any set of local coordinates around x, we have

$$\inf_{0<|v|<1} \frac{f(x+v)+f(x-v)-2f(x)}{|v|^2} = -\infty. \tag{24}$$

Similarly, we say that f fails to be semiconcave at $x \in M$ if

$$\sup_{0<|v|<1} \frac{f(x+v)+f(x-v)-2f(x)}{|v|^2} = +\infty.$$

For background on locally semiconvex and semiconcave function we refer to [CS04].

Let $\mathsf{d}_y^2 : M \to \mathbb{R}$ be the sub-Riemannian squared distance from $y \in M$, that is

$$\mathsf{d}_y^2(z) := d_{SR}^2(z,y), \qquad \forall\, z \in M.$$

The following result is a consequence of Theorem 26. For ideal structures, one can take $\Omega = M \setminus \{y\}$ in Corollary 30, yielding Theorem 6 presented in Section 1.

Corollary 30. *Let $(\mathcal{D}, g)$ be a sub-Riemannian structure on M. Let $y \neq x$. Assume that there exists a neighbourhood Ω of x such that all minimizing geodesics joining y with points of Ω do not contain abnormal segments. Then $x \in \mathrm{Cut}(y)$ if and only if the function d_y^2 fails to be semiconvex at x.*

Proof. We prove the contrapositive of the above statement. First, if $x \notin \mathrm{Cut}(y)$ then $f(z) := d_{SR}^2(z,y)$ is smooth in a neighbourhood of x by Theorem 19, and hence the infimum in (24) is finite.

To prove the opposite implication, observe that that by [CR08, Thm. 1, Thm. 5] f is locally semiconcave in a neighbourhood of x (for this property it suffices that no minimizing geodesic joining y with points of Ω is abnormal). By standard properties of semiconcave functions (see [CR08, Prop. 3.3.1]), there exist local charts around x, and $p \in \mathbb{R}^n$, $C \in \mathbb{R}$ such that

$$f(x+v) - f(x) \leq p \cdot v + C|v|^2, \qquad \forall\, |v| < 1. \tag{25}$$

Hence, assume that the infimum in (24) is finite, that is there exists $K \in \mathbb{R}$ such that, in local charts around x, we have

$$f(x+v) + f(x-v) - 2f(x) \geq K|v|^2, \qquad \forall\, |v| < 1. \tag{26}$$

Equations (25)–(26) yield that there exists a function $\phi : M \to \mathbb{R}$, twice differentiable at x, such that $f(z) \geq -\phi(z)$ for all $z \in M$, and $f(x) = -\phi(x)$. Our assumptions imply that the unique geodesic joining x with y does not contain abnormal segments. Hence, by Theorem 26, $y \notin \mathrm{Cut}(x)$. □

Remark 31. We observe the following general fact. For any sub-Riemannian structure, the infimum in (24) for $f = \mathsf{d}_y^2$ is always finite for $x = y$. On the other hand (when the structure is not Riemannian) d_y^2 fails to be semiconcave at y. The first statement follows trivially observing that $\mathsf{d}_y^2(y) = 0$ and $\mathsf{d}_y^2(z) \geq 0$. The second statement is a classical consequence of the Ball-Box theorem [Jea14].

4.2. **Regularity versus optimality: the non-ideal case.** Theorem 6 is false in the non-ideal case. In fact, consider the standard left-invariant sub-Riemannian structure on the product $\mathbb{H} \times \mathbb{R}$ of the three-dimensional Heisenberg group and the Euclidean line. Denoting points $x = (q, s) \in \mathbb{H} \times \mathbb{R}$, one has

$$d_{SR}^2((q,s),(q',s')) = d_{\mathbb{H}}^2(q,q') + |s-s'|^2.$$

Without loss of generality, fix $(q', s') = (0, 0)$. The set of points reached by abnormal minimizers is $\mathrm{Abn}(0) = \{(0, s) \mid s \in \mathbb{R}\}$. Here, the squared distance $\mathsf{d}_0^2(q, s) := d_{SR}^2((q, s), (0, 0))$ is not smooth, but the infimum in (24) is finite. In fact, the loss of smoothness is due to the failure of semiconcavity, see Remark 31.

Notice that abnormal geodesics joining the origin with points in $\mathrm{Abn}(0)$ are straight lines $t \mapsto (0, t)$, which are optimal for all times. Hence it seems likely that the failure of semiconvexity is related with the loss of optimality, while the failure of semiconcavity is related with the presence of abnormal minimizers. In the conclusion of this section, we formalize this latter statement.

4.2.1. *On the definition of cut locus.* In this paper, following [FR10], we defined the cut locus $\mathrm{Cut}(x)$ as the set of points y where the squared distance from x is not smooth. Classically, the cut locus is related with the loss of optimality of geodesics. To give a precise definition, we proceed as follows. First, we say that a geodesic $\gamma : [0, T] \to M$ (a horizontal curve which locally minimizes the energy between its endpoints) is maximal if it is not the restriction of a geodesic defined on a larger interval $[0, T']$. The cut time of a maximal geodesic is

$$t_{\mathrm{cut}}(\gamma) := \sup\{t > 0 \mid \gamma|_{[0,t]} \text{ is a minimizing geodesic}\}.$$

Assuming (M, d_{SR}) to be complete, we define the *optimal cut locus* of $x \in M$ as

$$\mathrm{CutOpt}(x) := \{\gamma(t_{\mathrm{cut}}(\gamma)) \mid \gamma \text{ is a maximal geodesic starting at } x\}.$$

In the ideal case, which includes the Riemannian case, it is well known that

$$\mathrm{CutOpt}(x) = \mathrm{Cut}(x) \setminus \{x\}. \tag{27}$$

For a general, complete sub-Riemannian structure $(\mathcal{D}, g)$ on M, let $x \in M$ and define the following sets:

$$\begin{aligned}
\mathrm{SC}^-(x) &:= \{y \in M \mid \mathsf{d}_x^2 \text{ fails to be semiconcave at } y\},\\
\mathrm{SC}^+(x) &:= \{y \in M \mid \mathsf{d}_x^2 \text{ fails to be semiconvex at } y\},\\
\mathrm{Abn}(x) &:= \{y \in M \mid \exists \text{ abnormal minimizing geodesic joining } x \text{ to } y\}.
\end{aligned}$$

Open questions. Are the following equalities true?

$$\mathrm{CutOpt}(x) = \mathrm{SC}^+(x), \tag{28}$$

$$\mathrm{Abn}(x) = \mathrm{SC}^-(x). \tag{29}$$

In the ideal case, (27) holds and $\mathrm{Abn}(x) = \{x\}$. Hence (28) follows from Corollary 30, (29) follows from the results of [CR08] (where the general inclusion $\mathrm{Abn}(x) \supseteq \mathrm{SC}^-(x)$ is proved). In particular, the following identities are true in the ideal case:

$$\mathrm{Cut}(x) = \mathrm{CutOpt}(x) \cup \mathrm{Abn}(x), \qquad \mathrm{Cut}(x) = \mathrm{SC}^+(x) \cup \mathrm{SC}^-(x). \tag{30}$$

We do not know whether (30) remain true in general. Notice that the first union in (30), in general, is not disjoint [RS17, BM16].

We mention that (29) holds true for the Martinet flat structure (a rank-varying structure on $\mathbb{R}^3$). In fact, as proved in [ABCK97], Martinet spheres possess outward corners in correspondence of points reached by abnormal minimizing geodesics, and this implies the loss of semiconcavity. The same characterization holds for the Engel group (a step 3 and rank 2 Carnot structure on $\mathbb{R}^4$), and for all free Carnot group of step 2, as proved in [MM16]. Finally, in [MM17], the authors proved the inclusion $\mathrm{CutOpt}(x) \subseteq \mathrm{SC}^+(x)$ for the free Carnot group of step 2 and rank 3.

5. Optimal transport and interpolation inequalities

The study of the Monge optimal transportation problem in sub-Riemannian geometry has been initiated in [AR04, FJ08] for the Heisenberg group and subsequently developed in [AL09, FR10, Lee13] for more general structures.

5.1. **Sub-Riemannian optimal transport.** Let us fix a smooth (outer) measure m on M, that is, defined by a positive tensor density. In particular m is Borel regular and locally finite, hence Radon [EG15].

The space of compactly supported probability measures on M is denoted by $\mathcal{P}_c(M)$, while $\mathcal{P}_c^{ac}(M)$ is the subset of the absolutely continuous ones w.r.t. m. We denote by $\pi_i : M \times M \to M$, for $i = 1, 2$, the projection on the i-th factor. Furthermore, let $D = \{(x, y) \in M \times M \mid x = y\}$.

Given two probability measures μ_0, μ_1 on M, we look for *transport maps* between μ_0 and μ_1, that is measurable maps $T : M \to M$, such that $T_\sharp \mu_0 = \mu_1$. Furthermore, for a given cost function $c : M \times M \to \mathbb{R}$, we want to minimize the transport cost among all transport maps, that is solve the Monge problem:

$$C_{\mathcal{M}}(\mu_0, \mu_1) = \min_{T_\sharp \mu_0 = \mu_1} \int_M c(x, T(x)) d\mathsf{m}(x). \tag{31}$$

Solutions of (31), when they exist, are called *optimal transport maps.*

We take from [FR10, Thm. 3.2] the main result about well-posedness of the Monge problem, and we specify it to the ideal setting. We give here a simplified but equivalent statement, which does not require a background in optimal transportation.

Theorem 32 (Well posedness of Monge problem)**.** *Let $(\mathcal{D}, g)$ be an ideal sub-Riemannian structure on M, and $\mu_0 \in \mathcal{P}_c^{ac}(M)$, $\mu_1 \in \mathcal{P}_c(M)$. Then there exists a unique transport map $T : M \to M$ such that $T_\sharp \mu_0 = \mu_1$, optimal w.r.t. the cost*

$$c(x, y) = \frac{1}{2} d_{SR}^2(x, y),$$

The map T is characterized as follows. There exist a closed set $\mathcal{S}^\psi$ (the static set), an open set $\mathcal{M}^\psi = M \setminus \mathcal{S}^\psi$ (the moving set), and a function $\psi : M \to \mathbb{R}$ locally semiconvex in a neighbourhood of $\mathcal{M}^\psi \cap \mathrm{supp}(\mu_0)$, such that

(i) For μ_0-a.e. $x \in \mathcal{S}^\psi$ then $T(x) = x$;
(ii) For μ_0-a.e. $x \in \mathcal{M}^\psi$ then $T(x) = y$ if and only if

$$\psi(x) + c(x, y) \leq \psi(z) + c(z, y), \qquad \forall z \in M. \tag{32}$$

In particular for μ_0-a.e. $x \in M$ there exists a unique minimizing geodesic between x and $T(x)$ given by

$$T_t(x) := \begin{cases} \exp_x(t d_x \psi) & x \in \mathcal{M}^\psi \cap \mathrm{supp}(\mu_0), \\ x & x \in \mathcal{S}^\psi \cap \mathrm{supp}(\mu_0), \end{cases} \qquad t \in [0, 1].$$

Remark 33. In the language of optimal transport, ψ is a Kantorovich potential associated with the problem and (32) means that the pair (x, y) belongs to the c-subdifferential $\partial_c \psi$.

Remark 34. All the results proved in this section hold if, for given $\mu_0 \in \mathcal{P}_c^{ac}(M)$, $\mu_1 \in \mathcal{P}_c(M)$, we replace the ideal hypothesis with the assumption that (M, d_{SR}) is complete and that there exists an open set $\Omega \subset M \times M$ such that $\mathrm{supp}(\mu_0 \times \mu_1) \subset \Omega$ and all minimizing geodesics with endpoints in $\Omega\backslash D$ contain no abnormal segments. This assumption is crucial for our Jacobian estimates, and it is used through Theorem 17.

It is currently unknown whether the Monge problem is well posed for general structures satisfying the so-called minimizing Sard property, i.e., where abnormal minimizers from any fixed point reach a set of measure zero. The reason is technical, since in this case one is not able to deduce enough regularity for the Kantorovich potentials associated with the optimal transport problem (see for example [Rif14]).

For what concerns the issue of regularity of T, in [FR10, Thm. 3.7], Figalli and Rifford obtained a formula for the differential of the transport map akin the classical one of [CEMS01] in terms of the Hessian of the distance, under additional hypothesis on the sub-Riemannian cut locus. More precisely, they assume that if $x \in \mathrm{Cut}(y)$, then there exist at least two distinct minimizing geodesics joining x with y. It turns out that the statement of [FR10, Thm. 3.7] holds with no assumptions on the cut locus in the ideal case. The differentiability result is most easily expressed in terms of approximate differential (see e.g. [AGS08, Sec. 5.5]).

Definition 35 (Approximate differential)**.** We say that $f : M \to \mathbb{R}$ has an approximate differential at $x \in M$ if there exists a function $g : M \to \mathbb{R}$ differentiable at x such that the set $\{f = g\}$ has density 1 at x with respect to m.[3] In this case, the approximate value of f at x is defined as $\tilde{f}(x) = g(x)$, and the approximate differential of f at x is defined as $\tilde{d}_x f := d_x g : T_xM \to T_{g(x)}M$.

Theorem 36 (Regularity of transport)**.** *Under the same assumptions of Theorem 32, the map T_t is differentiable μ_0-a.e. on $\mathcal{M}^\psi \cap \mathrm{supp}(\mu_0)$, and it is approximately differentiable μ_0-a.e. For μ_0-a.e. $x \in M$ its approximate differential is*

$$\tilde{d}_x T_t = \begin{cases} \pi_* \circ e^{t\vec{H}}_* \circ d^2_x\psi & x \in \mathcal{M}^\psi \cap \mathrm{supp}(\mu_0), \\ \mathrm{id}|_{T_xM} & x \in \mathcal{S}^\psi \cap \mathrm{supp}(\mu_0). \end{cases} \tag{33}$$

Finally, $\det(\tilde{d}_x T_t) > 0$ *for all $t \in [0, 1)$ and μ_0-a.e. $x \in M$.*

Remark 37. If $\mathcal{S}^\psi$ is empty, which is the case for example when $\mathrm{supp}(\mu_0) \cap \mathrm{supp}(\mu_1)$ is empty, then T_t is differentiable, and not only approximately differentiable, μ_0-a.e.

Proof. The closed set $\mathcal{S}^\psi$ is measurable, $\mu_0 \ll \mathsf{m}$, and m is smooth. Then by applying Lebesgue density theorem we obtain that T_t is approximately differentiable μ_0-a.e. on $\mathcal{S}^\psi \cap \mathrm{supp}(\mu_0)$, with approximate differential given by the identity map. Furthermore, $\det(\tilde{d}_x T_t) = 1$ for μ_0-a.e. $x \in \mathcal{S}^\psi \cap \mathrm{supp}(\mu_0)$.

We consider now the case $x \in \mathcal{M}^\psi$. Since local semiconvexity is invariant by diffeomorphisms, and since m is smooth, Alexandrov theorem in $\mathbb{R}^n$ (see e.g. [FR10, Thm. A.5]) yields that ψ is twice differentiable at m-a.e. point $x \in \mathcal{M}^\psi$, and its second differential can be computed according to Definition 20 and Remark 22. Hence, since $\mu_0 \ll \mathsf{m}$, the map

$$x \mapsto T_t(x) = \exp_x(t d_x\psi) = \pi \circ e^{t\vec{H}} \circ d_x\psi$$

is differentiable for μ_0-a.e. $x \in \mathcal{M}^\psi$, and its differential is computed as in (33).

By Theorem 32, $y = T_1(x)$ if and only if $\psi(z) + c(z, y) - \psi(x) - c(x, y) \geq 0$ for all $z \in M$. In particular, one can apply Theorems 26 and 27 with the function $\phi(z) := \psi(z) - \psi(x) - c(x, T(x))$, at any point x where ψ is twice differentiable, i.e. μ_0-almost everywhere on $\mathcal{M}^\psi$. In particular, for all such points $T(x) \notin \mathrm{Cut}(x)$ and $\det(d_x T_t) > 0$ for all $t \in [0, 1)$. □

[3]We compute the density using Euclidean balls in local charts around x. Since m is smooth, it has positive density with respect to the Lebesgue measure in charts, hence the concept of density does not depend on the choice of charts. In particular, Lebesgue density theorem holds.

Corollary 38. *Under the same assumptions of Theorem 32, for μ_0-a.e. $x \in \mathcal{M}^\psi$ we have $T(x) \notin \mathrm{Cut}(x)$.*

As a consequence of Theorem 36 and the estimate of Theorem 27, we obtain an independent proof of [FR10, Thm. 3.5] about the absolute continuity of the Wasserstein geodesic between μ_0 and μ_1.

Theorem 39 (Absolute continuity of Wasserstein geodesic)**.** *Under the same assumptions of Theorem 32, there exists a unique Wasserstein geodesic joining μ_0 with μ_1, given by $\mu_t = (T_t)_\sharp\mu_0$, for $t \in [0,1]$. Moreover, $\mu_t \in \mathcal{P}_c^{ac}(M)$ for all $t \in [0,1)$.*

Proof. The existence and uniqueness part is standard and is done as in [FR10, Sec. 6.3, first paragraph]. Since m is a smooth positive measure, the absolute continuity statement is equivalent to the absolute continuity of $\mu_t = (T_t)_\sharp\mu_0$ with respect to Lebesgue measure $\mathcal{L}^d$ in all local coordinate charts, where $\mu_0 = \rho\mathcal{L}^d$, for some $\rho \in L^1(\mathbb{R}^d)$. Thanks to Theorem 36, in charts $T_t : \mathbb{R}^d \to \mathbb{R}^d$ is approximately differentiable $\rho\mathcal{L}^d$-almost everywhere. Hence, the statement follows from the next lemma, which is a relaxed version of [AGS08, Lemma 5.5.3]. Its proof for completeness is in Appendix C.

Lemma 40. *Let $\rho \in L^1(\mathbb{R}^d)$ be a non-negative function. Let $f : \mathbb{R}^d \to \mathbb{R}^d$ be a measurable function and let Σ_f be the set where it is approximately differentiable. Assume there exists a Borel set $\Sigma \subseteq \Sigma_f$ such that the difference $\{\rho > 0\} \setminus \Sigma$ is $\mathcal{L}^d$-negligible. Then $f_\sharp(\rho\mathcal{L}^d) \ll \mathcal{L}^d$ if and only if $|\det(\tilde{d}_x f)| > 0$ for $\mathcal{L}^d$-a.e. $x \in \Sigma$.*

In this case, letting $f_\sharp(\rho\mathcal{L}^d) = \rho_f\mathcal{L}^d$, we have

$$\rho_f(y) = \sum_{x \in \tilde{f}^{-1}(y)\cap\Sigma} \frac{\rho(x)}{|\det(\tilde{d}_x f)|}, \qquad y \in \mathbb{R}^n,$$

with the convention that the r.h.s. is zero if $y \notin \tilde{f}(\Sigma)$. In particular, if we further assume that $\tilde{f}|_\Sigma$ is injective, then we have

$$\rho_f(\tilde{f}(x)) = \frac{\rho(x)}{|\det(\tilde{d}_x f)|}, \qquad \forall x \in \Sigma.$$

Notice that, in order to prove Theorem 39, we need only the first implication of Lemma 40, that is if $|\det(\tilde{d}_x f)| > 0$ for $\mathcal{L}^d$-a.e. $x \in \Sigma$, then $f_\sharp(\rho\mathcal{L}^d) \ll \mathcal{L}^d$. ☐

Thanks to Theorem 39, for all $t \in [0,1)$, and also $t = 1$ if $\mu_1 \in \mathcal{P}_c^{ac}(M)$, let $\rho_t := d\mu_t/d\mathsf{m}$. Then we have the following Jacobian identity.

Theorem 41 (Jacobian identity)**.** *Under the same assumptions of Theorem 32, let $\rho_t = d\mu_t/d\mathsf{m}$ for all $t \in [0,1)$, and also $t = 1$ if $\mu_1 \in \mathcal{P}_c^{ac}(M)$. Then, for μ_0-a.e. $x \in M$, we have*

$$\frac{\rho_0(x)}{\rho_t(T_t(x))} = \begin{cases} \det(d_xT_t)\frac{\mathsf{m}(X_1(t),\dots,X_n(t))}{\mathsf{m}(X_1(0),\dots,X_n(0))} > 0 & x \in \mathcal{M}^\psi \cap \mathrm{supp}(\mu_0), \\ 1 & x \in \mathcal{S}^\psi \cap \mathrm{supp}(\mu_0), \end{cases} \tag{34}$$

where $X_1(t),\dots,X_n(t)$ is some smooth moving frame along the geodesic $t \mapsto T_t(x)$, and the determinant of the linear map $d_xT_t : T_xM \to T_{T_t(x)}M$ is computed with respect to the given frame, that is

$$d_xT_t(X_i(0)) = \sum_{j=1}^n N_{ij}(t)X_j(t), \qquad \det(d_xT_t) := \det N(t).$$

Remark 42. In the Riemannian case, when $\mathsf{m} = \mathsf{m}_g$ is the Riemannian volume, one can compute the determinant in (34) with respect to orthonormal frames, eliminating any dependence on the frame and obtaining the classical Monge-Ampère equation.

Proof. By Theorem 39, $\mu_t = (T_t)_\sharp \mu_0 \ll \mu_0$, hence one can repeat the arguments in the last paragraph of [FR10, Sec. 6.4]. Since $\mu_t \in \mathcal{P}_c^{ac}(M)$, there are optimal transport maps T_t, S_t such that $(T_t)_\sharp \mu_0 = \mu_t$ and $(S_t)_\sharp \mu_t = \mu_0$. By uniqueness of the transport map, we obtain that T_t is μ_0-a.e. injective. Hence we can use the second part of Lemma 40, and in particular for μ_0-a.e. $x \in \mathcal{M}^\psi$ we have

$$\frac{\rho_0(x)}{\rho_t(T_t(x))} = \det(d_x T_t) \frac{\mathsf{m}(X_1(t), \dots, X_n(t))}{\mathsf{m}(X_1(0), \dots, X_n(0))}.$$

The extra term in the right hand side is due to the fact that we are not computing $d_x T_t$ in a set of local coordinates, but with respect to a smooth frame. □

5.2. **Distortion coefficients.** Let $(\mathcal{D}, g)$ be a sub-Riemannian structure on M, not necessarily ideal. The next lemma provides a general bound for the distortion coefficient (see definitions 1 and 2).

Lemma 43 (On-diagonal distortion bound)**.** *Let* m *be a smooth measure on* M*. Then, for any* $x \in M$*, there exists* $Q(x) \geq \dim(M)$ *such that*

$$\beta_t(x,x) \leq t^{Q(x)}, \qquad \forall t \in [0,1].$$

Proof. The proof is based on privileged coordinates and dilations in sub-Riemannian geometry, see [Jea14] for reference. Fix $x \in M$, and let z denote a system of privileged coordinates on a neighbourhood $\mathcal{O}$ of x (which we identify from now on with a relatively compact open set of $\mathbb{R}^n$, where x corresponds to the origin). We claim that there exists $Q(x) \geq \dim(M)$ and a constant $C(x) > 0$ such that, for sufficiently small ε, we have

$$\mathsf{m}(\mathcal{B}_\varepsilon(x)) = \varepsilon^{Q(x)} C(x) \left(1 + O(\varepsilon)\right).$$

This claim, together with the observation that $Z_t(x, \mathcal{B}_r(x)) \subseteq \mathcal{B}_{tr}(x)$, implies the statement. In order to prove the claim, in the given set of privileged coordinates, let $\mathsf{m} = \mathsf{m}(z) d\mathcal{L}(z)$ for some smooth, strictly positive function m. Assume ε to be sufficiently small such that $\mathcal{B}_\varepsilon(x) \subset \mathcal{O}$. Let δ_ε be the non-homogeneous dilation defined by the given system of privileged coordinates at x, with non-holonomic weights $w_i(x)$, for $i = 1, \dots, n$, that is $\delta_\varepsilon(z_1, \dots, z_n) = (\varepsilon^{w_1} z_1, \dots, \varepsilon^{w_n} z_n)$. Notice that, in coordinates, $\det(d_z \delta_\varepsilon) = \varepsilon^{Q(x)}$, where $Q(x) = \sum_{i=1}^n w_i(x)$.

As a consequence of the distance estimates in [Jea14, Thm. 2.2], for all sufficiently small $\varepsilon > 0$ there exists $\alpha(\varepsilon) \downarrow 0$ such that

$$\widehat{\mathcal{B}}_{(1-\alpha(\varepsilon))\varepsilon} \subseteq \mathcal{B}_\varepsilon(x) \subseteq \widehat{\mathcal{B}}_{(1+\alpha(\varepsilon))\varepsilon},$$

where $\widehat{\mathcal{B}}$ denotes the ball of the nilpotent structure, centered at the origin, in this set of privileged coordinates. By the homogeneity with respect to δ_ε, we have

$$\widehat{\mathcal{B}}_{1-\alpha(\varepsilon)} \subseteq \delta_{1/\varepsilon}(\mathcal{B}_\varepsilon(x)) \subseteq \widehat{\mathcal{B}}_{1+\alpha(\varepsilon)}.$$

The above relation, and the monotonicity of the Lebesgue measure as a function of the domain, imply that there exists a constant $B(x) = \mathcal{L}^n(\widehat{\mathcal{B}_1}) > 0$, such that

$$\lim_{\varepsilon \to 0} \mathcal{L}^n(\delta_{1/\varepsilon}(\mathcal{B}_\varepsilon)(x)) = B(x).$$

Hence, since δ_ε and m are smooth, we have

$$\begin{aligned} \mathsf{m}(\mathcal{B}_\varepsilon) &= \int_{\mathcal{B}_\varepsilon} \mathsf{m}(z) \mathcal{L}^n(dz) \\ &= \int_{\delta_{1/\varepsilon}(\mathcal{B}_\varepsilon)} \mathsf{m}(\delta_\varepsilon(z)) \det(d_z \delta_\varepsilon) \mathcal{L}^n(dz) \end{aligned}$$

$$= \varepsilon^{Q(x)}\mathsf{m}(0)\int_{\delta_{1/\varepsilon}(\mathcal{B}_\varepsilon)} (1+\varepsilon R(\delta_\varepsilon(z)))\,\mathcal{L}^n(dz)$$
$$= \varepsilon^{Q(x)}\mathsf{m}(0)B(x)\,(1+O(\varepsilon))\,,$$

where $R(z)$ is a smooth remainder, and the remainder term $O(\varepsilon)$ possibly depends on x. This concludes the proof of the claim. □

Lemma 44 (Computation of the distortion coefficients)**.** *Let $x, y \in M$, with $x \notin \mathrm{Cut}(y)$. Let $X_1, \dots, X_n$ be a smooth frame along the unique geodesic from x to y. Then, in terms of the Jacobi matrices defined in Section 3.3, we have*

$$\beta_t(x,y) = \frac{\det N_0^{\mathrm{V}}(t)}{\det N_0^{\mathrm{V}}(1)}\frac{\mathsf{m}(X_1(t),\dots,X_n(t))}{\mathsf{m}(X_1(1),\dots,X_n(1))}, \qquad \forall\, t \in [0,1], \tag{35}$$

$$\beta_{1-t}(y,x) = \frac{\det N_1^{\mathrm{V}}(t)}{\det N_1^{\mathrm{V}}(0)}\frac{\mathsf{m}(X_1(t),\dots,X_n(t))}{\mathsf{m}(X_1(0),\dots,X_n(0))}, \qquad \forall\, t \in [0,1]. \tag{36}$$

Moreover, $\beta_t(x,y) > 0$, for all $t \in (0,1]$.

Proof of Lemma 44. We prove first (35). For $t = 0$ both sides are zero, hence let $t \in (0,1]$. Let λ_0 be the initial covector of the unique minimizing geodesic such that $\exp_x(\lambda_0) = y$. Since $x \notin \mathrm{Cut}(y)$, there exists an open neighbourhood $\mathcal{O}$ of y and $O \subset T_x^*M$ such that $\exp_x : O \to \mathcal{O}$ is a smooth diffeomorphism, and for all $\lambda' \in O$, the geodesic $t \mapsto \exp_x(t\lambda')$ is the unique minimizing geodesic joining x with $y' = \exp_x(\lambda')$, and y' is not conjugate with x along such a geodesic. Assuming r sufficiently small such that $\mathcal{B}_r(y) \subset \mathcal{O}$, let $A_r \subset O$ be the relatively compact set such that $\exp_x(A_r) = \mathcal{B}_r(y)$. The map $\exp_x^t(\cdot) = \exp_x(t\cdot)$ is a smooth diffeomorphism from A_r onto $Z_t(x, \mathcal{B}_r(y))$. In particular, we have

$$\beta_t(x,y) = \lim_{r\downarrow 0}\frac{\int_{A_r}\exp_x^{t*}\mathsf{m}}{\int_{A_r}\exp_x^{1*}\mathsf{m}} = \frac{(\exp_x^{t*}\mathsf{m})(\lambda_0)}{(\exp_x^{1*}\mathsf{m})(\lambda_0)}. \tag{37}$$

The right hand side of (37) is the ratio of two smooth tensor densities computed at λ_0. To compute it, we evaluate both factors on a n-tuple of independent vectors of T_x^*M. Thus, pick a Darboux frame $E_1(t), \dots, E_n(t), F_1(t), \dots, F_n(t) \in T_{\lambda(t)}(T^*M)$ such that $\pi_* E_i(t) = 0$ and $\pi_* F_i(t) = X_i(t)$ for all $t \in [0,1]$, $i = 1, \dots, n$. Then,

$$(\exp_x^{t*}\mathsf{m})(E_1(0),\dots,E_n(0)) = \mathsf{m}(\pi_*\circ e_*^{t\vec{H}}E_1(0),\dots,\pi_*\circ e_*^{t\vec{H}}E_1(0)).$$

The n-tuple $\mathcal{J}_i(t) = e_*^{t\vec{H}}E_i(0)$, $i = 1, \dots, n$, corresponds to the Jacobi matrix $\mathbf{J}_0^{\mathrm{V}}(t)$, such that $M_0^{\mathrm{V}}(0) = \mathbb{1}$ and $N_0^{\mathrm{V}}(0) = \mathbb{0}$, defined in Section 3.3. Thus,

$$(\exp_x^{t*}\mathsf{m})(E_1(0),\dots,E_n(0)) = \det N_0^{\mathrm{V}}(t)\,\mathsf{m}(X_1(t),\dots,X_n(t)).$$

By replacing the above formula in (37), we obtain $\beta_t(x,y)$. Since $\gamma(t)$ is not conjugate to $\gamma(0)$ for all $t \in (0,1]$, we have $\beta_t(x,y) > 0$ on that interval.

Formula (36) for $\beta_{1-t}(y,x)$ is deduced in a similar way and with some additional care, following the geodesic backwards starting from the final point. We sketch the proof for this case. Let $\gamma : [0,1] \to M$ be the unique minimizing geodesic from x to y, with extremal $\lambda : [0,1] \to T^*M$. Of course, the unique minimizing geodesic from y to x is $\tilde{\gamma}(t) = \gamma(1-t)$. The corresponding normal extremal is $\tilde{\lambda}(t) = -\lambda(1-t)$. Consider the inversion map $\iota : T^*M \to T^*M$, such that $\iota(\lambda) = -\lambda$. In particular if $E_i(t), F_i(t)$ are a Darboux frame along $\lambda(t)$, then $\tilde{E}_i(t) := -\iota_* E_i(1-t)$ and $\tilde{F}_i(t) := -\iota_* F_i(1-t)$ are a Darboux frame along $\tilde{\lambda}(t)$. Hence, we have

$$\beta_{1-t}(y,x) = \frac{\mathsf{m}(\pi_*\circ e_*^{(1-t)\vec{H}}\tilde{E}_1(0),\dots,\pi_*\circ e_*^{(1-t)\vec{H}}\tilde{E}_n(0))}{\mathsf{m}(\pi_*\circ e_*^{\vec{H}}\tilde{E}_1(0),\dots,\pi_*\circ e_*^{\vec{H}}\tilde{E}_n(0))}, \qquad \forall t \in [0,1].$$

We conclude the proof by observing that the n-tuple

$$\tilde{\mathcal{J}}_i(t) = e_*^{(1-t)\vec{H}}\tilde{E}_i(0) = e_*^{(t-1)\vec{H}}E_i(1), \qquad i = 1,\dots,n,$$

corresponds to the Jacobi matrix $\mathbf{J}_1^{\mathrm{v}}(t) = \left(\begin{smallmatrix} M_1^{\mathrm{v}}(t) \\ N_1^{\mathrm{v}}(t)\end{smallmatrix}\right)$ in terms of $E_1(t),\dots,F_n(t)$. $\square$

5.3. **Interpolation inequalities: proof of Theorem 4.** For μ_0-a.e. $x \in \mathcal{S}^\psi$, by Theorems 32-36 we have $T_t(x) = x$ and $\rho_t(x) = \rho_0(x)$ for all $t \in [0,1]$. In this case the inequality follows from Lemma 43, which implies that

$$\beta_{1-t}(x,x)^{1/n} + \beta_t(x,x)^{1/n} \leq (1-t)^{Q(x)/n} + t^{Q(x)/n} \leq 1, \qquad \forall t \in [0,1].$$

Fix now $x \in \mathcal{M}^\psi$, such that (i) $\psi : M \to \mathbb{R}$ is twice differentiable, (ii) the Jacobian identity of Theorem 36 holds. By the absolute continuity of μ_0 w.r.t. m, properties (i)-(ii) are satisfied μ_0-a.e. in $\mathcal{M}^\psi$. Letting $X_i(t)$ be a moving frame along the geodesic $T_t(x) = \exp_x(t d_x\psi(x))$, we have

$$\frac{\rho_0(x)}{\rho_t(T_t(x))} = \det(d_x T_t)\frac{\mathsf{m}(X_1(t),\dots,X_n(t))}{\mathsf{m}(X_1(0),\dots,X_n(0))} > 0. \tag{38}$$

Recall that, according to Theorem 32, $y = T_1(x)$ if and only if for all $z \in M$ one has $\psi(z) + c(z,y) - \psi(x) - c(x,y) \geq 0$. One can apply Theorem 26 with $\phi(z) := \psi(z) - \psi(x) - c(x,T(x))$ at the point x, where ψ is twice differentiable. Hence, Theorem 27 yields an estimate for the determinant of the linear map

$$d_x T_t = \pi_* \circ e_*^{t\vec{H}} \circ d_x^2\phi = \pi_* \circ e_*^{t\vec{H}} \circ d_x^2\psi : T_xM \to T_{T_t(x)}M,$$

given by (18) for $s = 1$. Since $T(x) \notin \mathrm{Cut}(x)$ we can use the expressions for the distortion coefficients $\beta_t(x,T(x))$ of Lemma 44, and we obtain

$$\frac{\rho_0(x)^{1/n}}{\rho_t(T_t(x))^{1/n}} \geq \beta_{1-t}(y,x)^{1/n} + \beta_t(x,y)^{1/n}\left(\det(d_xT_1)\frac{\mathsf{m}(X_1(1),\dots,X_n(1))}{\mathsf{m}(X_1(0),\dots,X_n(0))}\right)^{1/n}.$$

If $\mu_1 \in \mathcal{P}_c^{ac}(M)$ we can again use (38) for $t = 1$ to replace the term containing $\det(d_xT_1)$, thus proving Theorem 4 for $t \in [0,1]$. If $\mu_1 \in \mathcal{P}_c(M) \setminus \mathcal{P}_c^{ac}(M)$ then (38) holds only for $t \in [0,1)$. In this case, in the first step, we simply omit the second term in (18) (which is non-negative). $\square$

6. Geometric and functional inequalities

The first consequence of the interpolation inequalities proved so far is a sub-Riemannian Borell-Brascamp-Lieb inequality, that is Theorem 45. Its proof follows, without any modification, as in [CEMS01, Sec. 6], and makes use of the assumption (39) for triple of points (x,y,z) satisfying $y = T_1(x)$ and $z = T_t(x)$, for some transport map T. This justifies removing $\mathrm{Cut}(M)$ from $A \times B$.

Theorem 45 (Sub-Riemannian Borell-Brascamp-Lieb inequality). *Let $(\mathcal{D},g)$ be an ideal sub-Riemannian structure on a n-dimensional manifold M, equipped with a smooth measure m. Fix $t \in [0,1]$. Let $f,g,h : M \to \mathbb{R}$ be non-negative and $A,B \subset M$ Borel subsets such that $\int_A f\, d\mathsf{m} = \int_B g\, d\mathsf{m} = 1$. Assume that for every $(x,y) \in (A\times B)\setminus \mathrm{Cut}(M)$ and $z \in Z_t(x,y)$,*

$$\frac{1}{h(z)^{1/n}} \leq \left(\frac{\beta_{1-t}(y,x)}{f(x)}\right)^{1/n} + \left(\frac{\beta_t(x,y)}{g(y)}\right)^{1/n}. \tag{39}$$

Then $\int_M h\, d\mathsf{m} \geq 1$.

Let $t \in [0,1]$ and $a, b \geq 0$. We introduce the p-mean for $p \in \mathbb{R} \cup \{\pm\infty\}$. When $p \neq 0, \pm\infty$ then

$$\mathcal{M}_t^p(a,b) := \begin{cases} ((1-t)a^p + tb^p)^{1/p} & \text{if } ab \neq 0, \\ 0 & \text{if } ab = 0. \end{cases}$$

The limit cases are defined as follows

$$\mathcal{M}_t^0(a,b) := a^{1-t}b^t, \quad \mathcal{M}_t^{+\infty}(a,b) := \max\{a,b\}, \quad \mathcal{M}_t^{-\infty}(a,b) := \min\{a,b\}.$$

The next result follows in a standard way from Theorem 45, by elementary properties of $\mathcal{M}_t^p$. Theorem 45 can be recovered from Theorem 46 by setting $p = -1/n$. The case $p = 0$ is the so-called Prékopa-Leindler inequality.

Theorem 46 (Sub-Riemannian p-mean inequality)**.** *Let $(\mathcal{D}, g)$ be an ideal sub-Riemannian structure on a n-dimensional manifold M, equipped with a smooth measure m. Fix $p \geq -1/n$ and $t \in [0,1]$. Let $f, g, h : M \to \mathbb{R}$ be non-negative and $A, B \subset M$ be Borel subsets such that $\int_A f \, d\mathsf{m} = \|f\|_{L^1(M)}$ and $\int_B g \, d\mathsf{m} = \|g\|_{L^1(M)}$. Assume that for every $(x,y) \in (A \times B) \setminus \mathrm{Cut}(M)$ and $z \in Z_t(x,y)$,*

$$h(z) \geq \mathcal{M}_t^p\left(\frac{(1-t)^n f(x)}{\beta_{1-t}(y,x)}, \frac{t^n g(y)}{\beta_t(x,y)}\right).$$

Then,

$$\int_M h \, d\mathsf{m} \geq \mathcal{M}_t^{p/(1+np)}\left(\int_M f \, d\mathsf{m}, \int_M g \, d\mathsf{m}\right),$$

with the convention that if $p = +\infty$ then $p/(1+np) = 1/n$, and if $p = -1/n$ then $p/(1+np) = -\infty$.

6.1. **Brunn-Minkowski inequality: proof of Theorem 7.** For $t = 0$ or $t = 1$, inequality (6) is trivially verified. Hence let $t \in (0,1)$. Assume first that $Z_t(A,B)$ is measurable, and set

$$f = \frac{\beta_{1-t}(B,A)}{(1-t)^n}\chi_A, \qquad g = \frac{\beta_t(A,B)}{t^n}\chi_B, \qquad h = \chi_{Z_t(A,B)}, \tag{40}$$

where χ_S is the characteristic function of a set $S \subset M$ and $\beta_t(A,B)$ is defined in (5). The assumption in Theorem 46 is satisfied with $p = +\infty$ since for every $(x,y) \in (A \times B) \setminus \mathrm{Cut}(M)$ and $z \in Z_t(x,y)$,

$$1 = h(z) \geq \max\left\{\frac{(1-t)^n f(x)}{\beta_{1-t}(y,x)}, \frac{t^n g(y)}{\beta_t(x,y)}\right\} = \max\left\{\frac{\beta_{1-t}(B,A)}{\beta_{1-t}(y,x)}, \frac{\beta_t(A,B)}{\beta_t(x,y)}\right\}.$$

Then, we have (when $p = +\infty$ it is understood that $p/(1+np) = 1/n$)

$$\begin{aligned} \mathsf{m}(Z_t(A,B)) = \int_M h \, d\mathsf{m} &\geq \mathcal{M}_t^{1/n}\left(\int_M f \, d\mathsf{m}, \int_M g \, d\mathsf{m}\right) \\ &= \left(\beta_{1-t}(B,A)^{1/n}\mathsf{m}(A)^{1/n} + \beta_t(A,B)^{1/n}\mathsf{m}(B)^{1/n}\right)^n, \end{aligned}$$

which proves the required inequality.

Assume now that $Z_t(A,B)$ is not measurable. Since m is Borel regular, there exists a measurable set C such that $Z_t(A,B) \subset C$, with $\mathsf{m}(Z_t(A,B)) = \mathsf{m}(C)$. We have clearly that $\chi_C \geq \chi_{Z_t(A,B)}$ and χ_C is measurable. The conclusion follows repeating the argument above replacing $h_{Z_t(A,B)}$ with h_C in (40). $\square$

6.2. **Equivalence of inequalities: proof of Theorem 9.** Let $(\mathcal{D}, g)$ be an ideal sub-Riemannian structure on a n-dimensional manifold M, equipped with a smooth measure m, and $N \geq 1$. We prove that (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i). First, by plugging (i) in the result of Theorem 7 we obtain (ii). Furthermore, (iii) is a particular case of (ii) by considering only sets of the form $A = \{x\}$. Finally, (iii) implies (i) by choosing in the former $B = \mathcal{B}_r(y)$, and recalling Definition 2 of $\beta_t(x, y)$. $\square$

7. Examples

In this section we discuss the distortion coefficients for some examples. The first one is the Heisenberg group.

7.1. **Heisenberg group.** The Heisenberg group $\mathbb{H}_3$ is the sub-Riemannian structure on $M = \mathbb{R}^3$ defined by the global generating vector fields

$$X_1 = \partial_x - \frac{y}{2}\partial_z, \qquad X_2 = \partial_y + \frac{x}{2}\partial_z.$$

The distribution has constant rank equal to two, and the sub-Riemannian structure is left-invariant with respect to the group product

$$(x, y, z) \star (x', y', z') = \left(x + x', y + y', z + z' + \frac{1}{2}(xy' - yx')\right).$$

The Heisenberg group is hence a Lie group and we equip it with the Lebesgue measure $\mathsf{m} = \mathcal{L}^3$, which is a Haar measure. Thanks to the left-invariance of the sub-Riemannian structure, it is enough to compute the distortion coefficients when one of the two points is the origin.

In dual coordinates (u, v, w, x, y, z) on $T^*\mathbb{R}^3$, the corresponding Hamiltonian is

$$H(u, v, w, x, y, z) = \frac{1}{2}\left(\left(u - \frac{y}{2}w\right)^2 + \left(v + \frac{x}{2}w\right)^2\right).$$

Hamilton equations can be explicitly integrated. In particular for an initial covector $\lambda_0 = u_0 dx + v_0 dy + w_0 dz \in T_0^*\mathbb{R}^3$, the exponential map from the origin $\exp_0^t(\lambda_0) := \exp_0(t\lambda_0)$ reads $\exp_0^t(u_0, v_0, w_0) = (x(t), y(t), z(t))$, where

$$\begin{aligned}
x(t) &= \frac{u_0 \sin(w_0 t) + v_0(\cos(w_0 t) - 1)}{w_0}, \\
y(t) &= \frac{u_0(1 - \cos(w_0 t)) + v_0 \sin(w_0 t)}{w_0}, \\
z(t) &= (u_0^2 + v_0^2)\frac{w_0 t - \sin(w_0 t)}{2w_0^2}.
\end{aligned}$$

Remark 47. The geodesic flow is an analytic function of the initial data, and if $w_0 = 0$ the above equations are understood by taking the limit $w_0 \to 0$. We always adopt this convention in this section.

In order to use Lemma 44 for the computation of the distortion coefficient, we choose the global Darboux frame induced by the global sections of $T(T^*\mathbb{R}^3)$:

$$E_1 = \partial_u, \quad E_2 = \partial_v, \quad E_3 = \partial_w, \quad F_1 = \partial_x, \quad F_2 = \partial_y, \quad F_3 = \partial_z.$$

In particular, the horizontal part of the Jacobi matrix $N_0^{\mathrm{V}}(t)$ is simply the Jacobian of the exponential map $(u, v, w) \mapsto \exp_0^t(u, v, w)$ computed at (u_0, v_0, w_0) in these coordinates. A straightforward computation and Lemma 44 yield the following.

Proposition 48 (Heisenberg distortion coefficient)**.** *Let $q \notin \mathrm{Cut}(0)$. Then*

$$\beta_t(0, q) = t\frac{\sin\left(\frac{tw_0}{2}\right)}{\sin\left(\frac{w_0}{2}\right)} \frac{\sin\left(\frac{tw_0}{2}\right) - \frac{tw_0}{2}\cos\left(\frac{tw_0}{2}\right)}{\sin\left(\frac{w_0}{2}\right) - \frac{w_0}{2}\cos\left(\frac{w_0}{2}\right)}, \qquad \forall t \in [0, 1],$$

where (u_0, v_0, w_0) *is the initial covector of the unique geodesic joining* 0 *with* q*.*

For the Heisenberg group, it is well-known that $t_{\mathrm{cut}}(u_0, v_0, w_0) = 2\pi/|w_0|$ (see e.g. [ABB12, Lemma 37]). Hence, since $q \notin \mathrm{Cut}(0)$, in the above formula it is understood that $|w_0| < 2\pi$, in which case one can check that $\beta_t(0, q) > 0$ for all $t \in (0, 1]$.

Remark 49. In the above notation, $d_{SR}^2(0, q) = \|\lambda\|^2 = u_0^2 + v_0^2$. We observe that the Heisenberg distortion coefficient *does not* depend on the distance $d_{SR}(0, q)$, but rather on the "vertical part" w_0 of the covector λ. See Section 8.

The following lemma is a consequence of the inequalities of [Riz16, Lemma 18].

Lemma 50 (Sharp bound to Heisenberg distortion)**.** *Let* $N \in \mathbb{R}$*. The inequality*

$$\beta_t(q_0, q) \geq t^N, \qquad \forall t \in [0, 1],$$

holds for all points $q_0, q \in \mathbb{H}_3$ *with* $q \notin \mathrm{Cut}(q_0)$*, if and only if* $N \geq 5$*.*

One can verify that for $q \in \mathrm{Cut}(q_0)$ one has $\beta_t(q_0, q) = +\infty$ for every $t \in (0, 1)$. For a proof see [BKS18, Cor. 2.1]. We recover the following results from [BKS18].

Corollary 51. *Let* $\mu_0 \in \mathcal{P}_c^{ac}(\mathbb{H}_3)$*, and* $\mu_1 \in \mathcal{P}_c(\mathbb{H}_3)$*. Let* $\mu_t = (T_t)_\sharp \mu_0 = \rho_t \mathcal{L}^3$ *be the unique Wasserstein geodesic joining* μ_0 *with* μ_1*. Then,*

$$\frac{1}{\rho_t(T_t(x))^{1/3}} \geq \frac{(1-t)^{5/3}}{\rho_0(x)^{1/3}} + \frac{t^{5/3}}{\rho_1(T(x))^{1/3}}, \qquad \mathcal{L}^3 - \text{a.e.}, \forall\, t \in [0, 1].$$

The above inequality is sharp, in the sense that if one replaces the exponent 5 *with a smaller one, the inequality fails for some choice of* μ_0, μ_1*.*

Corollary 52. *For all non-empty Borel sets* $A, B \subset \mathbb{H}_3$*, we have*

$$\mathcal{L}^3(Z_t(A, B))^{1/3} \geq (1-t)^{5/3} \mathcal{L}^3(A)^{1/3} + t^{5/3} \mathcal{L}^3(B)^{1/3}, \qquad \forall\, t \in [0, 1].$$

The above inequality is sharp, in the sense that if one replaces the exponent 5 *with a smaller one, the inequality fails for some choice of* A, B*.*

Notice that, as a consequence of Theorem 9, we recover also the following result originally obtained in [Jui09]: the Heisenberg group $\mathbb{H}_3$, equipped with the Lebesgue measure, satisfies the $\mathrm{MCP}(K, N)$ if and only if $N \geq 5$ and $K \leq 0$.

7.2. **Generalized H-type groups.** These structures were introduced in [BR17c], and constitute a large class of Carnot groups where the optimal synthesis is known. This class contains Kaplan H-type groups, and some of these structures admit non-trivial abnormal minimizing geodesics.

We take the definitions directly from [BR17c], to which we refer for more details. Let $(G, \star)$ be a step 2 Carnot group, with Lie algebra $\mathfrak{g}$ of rank k, dimension n satisfying $\dim \mathfrak{g}_1 = k$, $\dim \mathfrak{g}_2 = n - k$ and

$$[\mathfrak{g}_1, \mathfrak{g}_1] = \mathfrak{g}_2, \qquad [\mathfrak{g}_i, \mathfrak{g}_2] = 0, \qquad i = 1, 2.$$

Any choice g of a scalar product on $\mathfrak{g}_1$ induces a left-invariant sub-Riemannian structure $(\mathcal{D}, g)$ on G, such that $\mathcal{D}(p) = \mathfrak{g}_1(p)$ for all $p \in G$. We extend the scalar product g on $\mathfrak{g}_1$ to the whole $\mathfrak{g}$, which we denote with the same symbol.

For any $V \in \mathfrak{g}_2$, the skew-symmetric operator $J_V : \mathfrak{g}_1 \to \mathfrak{g}_1$ is defined by

$$g(X, J_V Y) = g(V, [X, Y]), \qquad \forall X, Y \in \mathfrak{g}_1.$$

Definition 53. We say that a step 2 Carnot group is of *generalized H-type* if there exists a symmetric, non-zero and non-negative operator $S : \mathfrak{g}_1 \to \mathfrak{g}_1$ such that

$$J_V J_W + J_W J_V = -2g(V, W) S^2, \qquad \forall V, W \in \mathfrak{g}_2. \tag{41}$$

Remark 54. The above definition is well posed and does not depend on the choice of the extension of g. More precisely, if (41) is verified for the operators J_V defined by a choice of an extension of g, then the operators $\tilde{J}_V$ defined by a different extension $\tilde{g}$ will verify (41), with the *same* operator S.

Remark 55. A generalized H-type group does not admit non-trivial abnormal geodesics, and is thus ideal, if and only if S is invertible. When $n = k + 1$, we are in the case of corank 1 Carnot groups. If S is also non-degenerate (and thus $k = 2d$ is even and $S > 0$), we are in the case of contact Carnot groups. The case $S = \mathrm{Id}_{\mathfrak{g}_1}$ and $k = 2d$ corresponds to classical Kaplan H-type groups.

The next result follows from the explicit expression for the Jacobian determinant of generalized H-type groups [BR17c, Lemma 20], which in turn allows to compute explicit distortion coefficients. The latter, in turn, can be bounded by a power law thanks to [BR17c, Cor. 27]. In particular, we have the following.

Lemma 56 (Sharp bound to generalized H-type distortion)**.** *Let $(G, \mathcal{D}, g)$ be a generalized H-type group, with dimension n and rank k, equipped with a left-invariant measure m. Let $N \in \mathbb{R}$. The inequality*

$$\beta_t(x, y) \geq t^N, \qquad \forall t \in [0, 1],$$

holds for all points $x, y \in G$ with $y \notin \mathrm{Cut}(x)$, if and only if $N \geq k + 3(n - k)$, the latter number being the geodesic dimension of the Carnot group.

As a consequence of 56 and Theorem 4 we have the following result, in the ideal setting.

Corollary 57. *Let $(G, \mathcal{D}, g)$ be an ideal generalized H-type group, with dimension n and rank k, equipped with a left-invariant measure m. Let $\mu_0, \mu_1 \in \mathcal{P}_c^{ac}(G)$. Let $\mu_t = (T_t)_\sharp \mu_0 = \rho_t \mathsf{m}$ be the unique Wasserstein geodesic joining μ_0 with μ_1. Then,*

$$\frac{1}{\rho_t(T_t(x))^{1/n}} \geq \frac{(1-t)^{\frac{k+3(n-k)}{n}}}{\rho_0(x)^{1/n}} + \frac{t^{\frac{k+3(n-k)}{n}}}{\rho_1(T(x))^{1/n}}, \qquad \mathsf{m}-\text{a.e.}, \forall t \in [0, 1].$$

The above inequality is sharp, in the sense that if one replaces the exponent $k + 3(n - k)$ with a smaller one, the inequality fails for some choice of μ_0, μ_1.

Remark 58. The restriction to ideal structures in the above corollary arises from the requirements of the general theory leading to Theorem 4, while this assumption is not necessary in Lemma 56. However, we remark that abnormal geodesics of generalized H-type groups are very docile (they consists in straight lines, and never lose minimality). Thus, we expect all the above results to hold also for non-ideal generalized H-type groups. This is supported by the positive results obtained for corank 1 Carnot groups obtained in [BKS17] and the forthcoming Corollary 59.

The sharp Brunn-Minkowski inequality for ideal generalized H-type groups follows from Theorem 9 and Lemma 56. However, thanks to the results of [RYN18] for product structures, we are able to remove the ideal assumption.

Corollary 59. *Let $(G, \mathcal{D}, g)$ be a generalized H-type group, with dimension n and rank k, equipped with a left-invariant measure m. For all non-empty Borel sets $A, B \subset G$, we have*

$$\mathsf{m}(Z_t(A, B))^{\frac{1}{n}} \geq (1-t)^{\frac{k+3(n-k)}{n}} \mathsf{m}(A)^{\frac{1}{n}} + t^{\frac{k+3(n-k)}{n}} \mathsf{m}(B)^{\frac{1}{n}}, \quad \forall t \in [0, 1].$$

The above inequality is sharp, in the sense that if one replaces the exponent $k + 3(n - k)$ with a smaller one the inequality fails for some choice of A, B.

Proof. In this proof, given $N, n \in \mathbb{N}$, we denote $\mathrm{BM}(N,n)$ the following property: for all non-empty Borel sets A, B, we have

$$\mathsf{m}(Z_t(A,B))^{1/n} \geq (1-t)^{N/n}\mathsf{m}(A)^{1/n} + t^{N/n}\mathsf{m}(B)^{1/n}, \qquad \forall\, t \in [0,1].$$

A generalized H-type group G is the product of an ideal one G_0 with dimension $n_0 = n-d$ and rank $k_0 = k-d$, and d copies of the Euclidean $\mathbb{R}$, for a unique $d \geq 0$ (see [BR17c, Prop. 19]). Furthermore, the left-invariant measure m of G is the product of the a left-invariant measure m_0 of G_0 and d copies of the Lebesgue measures $\mathcal{L}$ of each factor $\mathbb{R}$. It follows immediately from Lemma 56 and Theorem 9 that G_0, equipped with the measure m_0, satisfies $\mathrm{BM}(N_0, n_0)$, with $N_0 = k_0 + 3(n_0 - k_0)$. Furthermore, each copy of $\mathbb{R}$, equipped with the Lebesgue measure, satisfies the standard linear Brunn-Minkowski inequality $\mathrm{BM}(1,1)$. It follows from [RYN18, Thm. 3.3] that the product $G = G_0 \times \mathbb{R}^d$ equipped with the left-invariant measure $\mathsf{m} = \mathsf{m}_0 \times \mathcal{L}^d$ satisfies $\mathrm{BM}(N_0 + d, n_0 + d) = \mathrm{BM}(k + 3(n-k), n)$, which is the desired inequality.

Assume that G satisfies $\mathrm{BM}(k + 3(n-k) - \varepsilon, n)$ for some $\varepsilon > 0$. Let $x \notin \mathrm{Cut}(y)$. Letting $A = x \in G$ and $B = \mathcal{B}_r(y) \subset G$, and taking the limit for $r \downarrow 0$, we obtain that $\beta_t(x,y) \geq t^{k+3(n-k)-\varepsilon}$, contradicting the results of Lemma 56. □

We can also easily recover the following result proved in [BR17c]: a generalized H-type group with dimension n and rank k, equipped with a left-invariant measure m, satisfies the $\mathrm{MCP}(K,N)$ if and only if $N \geq k + 3(n-k)$ and $K \leq 0$.

7.3. **Grushin plane.** The Grushin plane $\mathbb{G}_2$ is the sub-Riemannian structure on $\mathbb{R}^2$ defined by the global generating vector fields

$$X_1 = \partial_x, \qquad X_2 = x\partial_y.$$

We stress that the rank of $\mathcal{D} = \mathrm{span}\{X_1, X_2\}$ is not constant. More precisely, the structure is Riemannian on $\{x \neq 0\}$, and it is singular otherwise. We equip the Grushin plane with the Lebesgue measure $\mathsf{m} = \mathcal{L}^2$ of $\mathbb{R}^2$. In canonical coordinates (u,v,x,y) on $T^*\mathbb{R}^2$, the corresponding Hamiltonian is

$$H(u,v,x,y) = \frac{1}{2}(u^2 + x^2v^2).$$

Hamilton equations are easily integrated, and the Hamiltonian flow

$$e^{t\vec{H}}(u_0,v_0,x_0,y_0) = (u(t),v(t),x(t),y(t))$$

with initial covector $\lambda_0 = u_0dx + v_0dy \in T^*_{(x_0,y_0)}\mathbb{R}^2$ reads

$$\begin{aligned}
u(t) &= u_0\cos(tv_0) - x_0v_0\sin(tv_0), \\
v(t) &= v_0, \\
x(t) &= x_0\cos(tv_0) + u_0\frac{\sin(tv_0)}{v_0}, \\
y(t) &= y_0 + \frac{\sin(2tv_0)\left(v_0^2x_0^2 - u_0^2\right) + 2v_0\left(t\left(v_0^2x_0^2 + u_0^2\right) + u_0x_0 - u_0x_0\cos(2tv_0)\right)}{4v_0^2}.
\end{aligned}$$

In particular, $\exp^t_{(x_0,y_0)}(u_0,v_0) = (x(t),y(t))$.

Remark 60. Notice that the geodesic flow is an analytic function of the initial data, and if $v_0 = 0$ the above equations are understood by taking the limit $v_0 \to 0$. We always adopt this convention in this section.

To compute the distortion coefficients, fix $q_0 = (x_0,y_0) \in \mathbb{R}^2$, let $q \notin \mathrm{Cut}(q_0)$, and let $\lambda_0 = u_0dx + v_0dy \in T^*_{(x_0,y_0)}\mathbb{R}^2$ the covector of the unique minimizing geodesic $\gamma : [0,1] \to \mathbb{R}^2$ joining q_0 with q.

In order to use Lemma 44 for the computation of the distortion coefficient, we choose the global Darboux frame induced by the global sections of $T(T^*\mathbb{R}^2)$:

$$E_1 = \partial_u, \qquad E_2 = \partial_v, \qquad F_1 = \partial_x, \qquad F_2 = \partial_y.$$

In particular, the horizontal part of the Jacobi matrix $N_0^{\mathrm{V}}(t)$ is simply the Jacobian of the exponential map $(u,v) \mapsto \exp^t_{(x_0,y_0)}(u,v)$ in these coordinates, computed at (u_0,v_0). A straightforward computation and Lemma 44 yield the following.

Proposition 61 (Grushin distortion coefficient)**.** *Let $q \notin \mathrm{Cut}(q_0)$. Then*

$$\beta_t(q_0,q) = t\frac{(u_0^2 + tu_0v_0^2x_0 + v_0^2x_0^2)\sin(tv_0) - tu_0^2v_0\cos(tv_0)}{(u_0^2 + u_0v_0^2x_0 + v_0^2x_0^2)\sin(v_0) - u_0^2v_0\cos(v_0)}, \qquad \forall t \in [0,1], \tag{42}$$

where (u_0,v_0) is the initial covector of the unique geodesic joining q_0 with q.

For the Grushin plane, $t_{\mathrm{cut}}(u_0,v_0) = \pi/|v_0|$ (see [ABS08, Sec. 3.2] or [ABB16b, Ch. 13]). Hence, since $q \notin \mathrm{Cut}(q_0)$, in the above formula it is understood that $|v_0| < \pi$, in which case one can check directly that $\beta_t(q_0,q) > 0$ for all $t \in (0,1]$.

We have the following non-trivial estimate.

Proposition 62. *Let $N \in \mathbb{R}$. The inequality*

$$\beta_t(q_0,q) \geq t^N, \qquad \forall t \in [0,1],$$

holds for all points $q_0, q \in \mathbb{G}_2$ with $q \notin \mathrm{Cut}(q_0)$, if and only if $N \geq 5$.

Remark 63. Even though the Grushin plane is a quotient of the Heisenberg group, it is not clear how to deduce a bound for distortion coefficients of $\mathbb{G}_2$ starting from the knowledge of the corresponding inequality for $\mathbb{H}_3$. Actually, the most surprising aspect of Proposition 62 is its sharpness. As it is clear from the proof, the necessity of the condition $N \geq 5$ is due to pairs of points q_0, q located on opposite sides of the singular set $\{x = 0\}$.

Proof. Let $q = \exp_{q_0}(u_0,v_0)$, with $|v_0| < \pi$, and $q_0 = (x_0,y_0) \in \mathbb{R}^2$. If $x_0 = 0$,

$$\beta_t(q_0,q) = t \times \frac{\sin(tv_0) - tv_0\cos(tv_0)}{\sin(v_0) - v_0\cos(v_0)} \geq t^4, \qquad \forall t \in [0,1],$$

which follows from the inequality of [Riz16, Lemma 18], and $|v_0| < \pi$. We now proceed by assuming $x_0 \neq 0$ (by symmetry we actually assume $x_0 > 0$).

Case $v_0 = 0$. This case, corresponding to straight horizontal lines possibly crossing the singular region, is the one which yields the "only if" part of the theorem, and we will settle it first. In this case the trigonometric terms disappear, and

$$\beta_t(q_0,q) = t^2 \times \frac{t^2u_0^2 + 3tu_0x_0 + 3x_0^2}{u_0^2 + 3u_0x_0 + 3x_0^2}.$$

We want to find the best $N \in \mathbb{R}$, such that for all $x_0 > 0$ and $u_0 \in \mathbb{R}$, it holds

$$\frac{t^2u_0^2 + 3tu_0x_0 + 3x_0^2}{u_0^2 + 3u_0x_0 + 3x_0^2} \geq t^{N-2}, \qquad \forall t \in [0,1].$$

Since both sides are strictly positive for all $t \in (0,1]$, we can take the logarithms and the above inequality is equivalent to

$$\int_{tu}^{u} \frac{d}{dz}\log f_{x_0}(z)\,dz \leq (N-2)\int_{tu}^{u} \frac{d}{dz}\log|z|\,dz, \qquad \forall t \in (0,1),\ u \in \mathbb{R},$$

where $f_{x_0}(z) := z^2 + 3zx_0 + 3x_0^2$. This inequality is equivalent to the corresponding inequality for the integrands. After some computations, we obtain the condition

$$(N-4)z^2 + 3x_0(N-3)z + 3x_0^2(N-2) \geq 0, \qquad \forall x_0 > 0,\ z \in \mathbb{R}.$$

One easily checks that the above holds if and only if $N \geq 5$. This proves the "only if" part of the statement.

Case $v_0 \neq 0$. By symmetry, we actually assume $v_0 > 0$. If $u_0 = 0$, then

$$\beta_t(q_0, q) = t\frac{\sin(tv_0)}{\sin(v_0)} \geq t^2 \geq t^5, \qquad \forall t \in (0,1).$$

Hence in the following we consider $u_0 \neq 0$. We recall the assumptions made so far:

$$x_0 > 0, \qquad v_0 > 0, \qquad u_0 \neq 0.$$

In this case we rewrite (42) as

$$\beta_t(q_0, q) = t \times \frac{f_a(tv_0)}{f_a(v_0)}, \qquad a := \frac{v_0 x_0}{u_0} \in \mathbb{R}_0 = \mathbb{R} \setminus \{0\},$$

where, for all $a \in \mathbb{R}_0$, we defined

$$f_a(\xi) := (1 + a\xi + a^2)\sin(\xi) - \xi\cos(\xi).$$

It remains to prove that for all $a \in \mathbb{R}_0$ and $N \geq 5$ it holds

$$\frac{f_a(tv_0)}{f_a(v_0)} \geq t^{N-1}, \qquad \forall t \in [0,1].$$

In particular, it is sufficient to prove the case $N = 5$, which we assume from now on. Since both sides are strictly positive on $t \in (0,1]$, we can take the logarithms and the inequality is equivalent to

$$\int_{tv_0}^{v_0} \frac{d}{dz}\log f_a(z)\,dz \leq 4\int_{tv_0}^{v_0} \frac{d}{dz}\log|z|\,dz, \qquad \forall t \in (0,1),\ a \in \mathbb{R}_0.$$

The above inequality is equivalent to the corresponding one for the integrands. After some computation, we obtain the equivalent inequality

$$W_a(z) := Q_a(z)\sin(z) - zP_a(z)\cos(z) \geq 0, \qquad \forall z \in (0,\pi),\ a \in \mathbb{R}_0, \tag{43}$$

where we defined the two polynomials:

$$P_a(z) = a(a+z) + 4, \qquad Q_a(z) = (a+z)(4a - z) + 4.$$

Consider $a \mapsto W_a(z)$. It is easy to check that for all fixed $z \in (0,\pi)$, we have

$$\lim_{a \to \pm\infty} W_a(z) = +\infty.$$

Moreover, $\partial_a W_a(z)$ is linear, hence the function $a \mapsto W_a(z)$ has a unique minimum. Then (43) is equivalent to the fact that this minimum is non-negative for all $z \in (0,\pi)$. Setting $\partial_a W_a(z) = 0$, we obtain

$$a_{\min} = -\frac{z}{2} \times \frac{3\sin(z) - z\cos(z)}{4\sin(z) - z\cos(z)} \leq 0.$$

Hence, (43) is equivalent to $W_{a_{\min}}(z) \geq 0$ for all $z \in (0,\pi)$. Replacing, and after some computations, we have that such a condition is equivalent to

$$\bar{W}(z) := \alpha(z)\sin(z)^2 + \beta(z)\cos(z)\sin(z) + \gamma(z)\cos(z)^2 \geq 0, \qquad \forall z \in (0,\pi), \tag{44}$$

where we have defined the following polynomials

$$\begin{aligned}
\alpha(z) &= (64 - 25z^2),\\
\beta(z) &= 10z(z^2 - 8),\\
\gamma(z) &= z^2(16 - z^2) > 0.
\end{aligned}$$

By looking to the graph of $W(z)$, for $z \in (0,\pi)$, one notices that the inequality (44) is extremely sharp for z close to 0, while it is easier to prove for larger z. Hence, we split the proof of (44) into two parts.

(i) Proof of (44) on $(0, 2.67)$. Notice that $\bar{W}^{(i)}(0) = 0$ for all $i = 0, 1, 2, 3, 4$, and

$$\bar{W}^{(5)}(z) = 8\left(2z^4 + 8z^2 + 3\right)\sin(2z) + 80z\cos(2z).$$

This is the first i-th derivative whose polynomial factors multiplying the trigonometric functions are all non-negative. Furthermore, recall that

$$\sin(x) \geq x - \frac{x^3}{6}, \qquad x \in [0, +\infty),$$

$$\cos(x) \geq 1 - \frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720}, \qquad x \in [0, +\infty).$$

Hence, using the explicit form of $\bar{W}^{(5)}$, and the fact that the polynomial factors are non-negative, we obtain

$$\begin{aligned} \bar{W}^{(5)}(z) &\geq 8\left(2z^4 + 8z^2 + 3\right)\left(2z - \frac{4z^3}{3}\right) + 80z\left(1 - 2z^2 + \frac{2z^4}{3} - \frac{4z^6}{45}\right) \\ &= \frac{64}{9}z\left(18 - 9z^2 - 4z^6\right), \qquad \forall z \in (0, \pi). \end{aligned}$$

Integrating five times the above inequality on the interval $[0, z]$, and since $\bar{W}^{(i)}(0) = 0$ for all $i = 0, 1, 2, 3, 4$, we obtain

$$\bar{W}(z) \geq z^6\left(-\frac{4z^6}{13365} - \frac{z^2}{105} + \frac{8}{45}\right), \qquad \forall z \in (0, \pi).$$

The term in parenthesis in right hand side of the above is a third order polynomial, for which we can compute explicitly the roots. The first positive root occurs at $z_* \simeq 2.67491$. Hence the inequality $\bar{W}(z) \geq 0$ is proved for $z \in (0, 2.67)$.
(ii) Proof of (44) on $[2.67, \pi)$. On this interval the inequality is easier to verify with rough estimates. Indeed, $\bar{W}(z)$ in (44) is the sum of three terms, and it is sufficient to bound each one of them with the corresponding minimum. The first term

$$(64 - 25z^2)\sin(z)^2 \quad \text{attains its minimum on } [2.67, \pi) \text{ at } z = 2.67,$$

where its value is approximately -23.57. The second term

$$10z(z^2 - 8)\sin(z)\cos(z) \quad \text{attains it minimum on } [2.67, \pi) \text{ at } z \simeq 3,$$

where its value is approximately -4.20. The third term

$$z^2(16 - z^2)\cos(z)^2 \quad \text{attains its minimum on } [2.67, \pi) \text{ at } z = 2.67.$$

where its value is approximately 50.19. Thus $\bar{W}(z)$ is larger than the sum of the three aforementioned values, which is positive. This proves (44), and concludes the proof of the proposition. □

Since the Grushin structure is ideal, one obtains the following consequences.

Corollary 64. *Let $\mu_0, \mu_1 \in \mathcal{P}_c^{ac}(\mathbb{G}_2)$. Let $\mu_t = (T_t)_\sharp \mu_0 = \rho_t \mathcal{L}^2$ be the unique Wasserstein geodesic joining μ_0 with μ_1. Then,*

$$\frac{1}{\rho_t(T_t(x))^{1/2}} \geq \frac{(1-t)^{5/2}}{\rho_0(x)^{1/2}} + \frac{t^{5/2}}{\rho_1(T(x))^{1/2}}, \qquad \mathcal{L}^2 - \text{a.e.}, \forall t \in [0, 1]. \tag{45}$$

The above inequality is sharp, in the sense that if one replaces the exponent 5 with a smaller one, the inequality fails for some choice of μ_0, μ_1. If μ_1 is not absolutely continuous, an analogous result holds, provided that $t \in [0, 1)$, and that in (45) the second term on the right hand side is omitted.

Corollary 65. *For all non-empty Borel sets $A, B \subset \mathbb{G}_2$, we have*

$$\mathcal{L}^2(Z_t(A,B))^{1/2} \geq (1-t)^{5/2}\mathcal{L}^2(A)^{1/2} + t^{5/2}\mathcal{L}^2(B)^{1/2}, \qquad \forall\, t \in [0,1].$$

The above inequality is sharp, in the sense that if one replaces the exponent 5 with a smaller one, the inequality fails for some choice of A, B.

Finally, by taking $A = y \in \mathbb{G}_2$, and using the fact that the Grushin plane admits a one-parameter group of metric dilations, we obtain the following result.

Corollary 66. *The Grushin plane, equipped with the Lebesgue measure, satisfies the* $\mathrm{MCP}(K,N)$ *if and only if $N \geq 5$ and $K \leq 0$.*

On the other hand, the Grushin *half*-planes satisfy the $\mathrm{MCP}(K,N)$ if and only if $N \geq 4$ and $K \leq 0$, see [Riz17].

7.4. **Sasakian manifolds.** Let $(\mathcal{D}, g)$ be a contact sub-Riemannian structure on a $2d+1$-dimensional manifold M, that is $\mathcal{D} = \ker \omega$ where $\omega \in \Lambda^1 M$ is a one-form such that and $d\omega|_{\mathcal{D}}$ is non-degenerate. In particular, $(\mathcal{D}, g)$ is ideal.

The Reeb vector field X_0 is the unique vector field satisfying $\omega(X_0) = 1$ and $d\omega(X_0, \cdot) = 0$. We extend the sub-Riemannian metric to a global Riemannian structure (that we denote with the same symbol g) by declaring X_0 to an unit vector orthogonal to $\mathcal{D}$. We denote by m the corresponding canonical measure. We define the *contact endomorphism* $J : TM \to TM$ by:

$$g(X, JY) = d\omega(X,Y), \qquad \forall X, Y \in \Gamma(TM).$$

The structure is called *Sasakian* if the $(1,1)$ tensor defined on $M \times \mathbb{R}$

$$\mathbf{J}(X, f\partial_t) = (JX - fX_0, \omega(X)\partial_t),$$

defines a complex structure. We denote with R^* and Ric^* the Riemann and Ricci tensors associated with the Tanaka-Webster connection ∇^* (see references below for details and precise definitions).

In [AL14, LLZ16] it is proved that a $2d+1$-dimensional Sasakian structure endowed with its canonical volume satisfies the $\mathrm{MCP}(0, 2d+3)$, under some curvature bounds. Thanks to Theorem 9 and [LLZ16, Thm. 1.1], we have the following consequence.

Corollary 67. *Let $(\mathcal{D}, g)$ a sub-Riemannian Sasakian structure on a $2d+1$-dimensional manifold M, equipped with its canonical measure m. Assume that*

(i) $\mathrm{R}^*(X, JX, JX, X) \geq 0$ *for every* $X \in \Gamma(\mathcal{D})$,
(ii) $\mathrm{Ric}^*(X,X) - \mathrm{R}^*(X, JX, JX, X) \geq 0$ *for every* $X \in \Gamma(\mathcal{D})$.

Then, for all non-empty Borel sets $A, B \subset M$, we have

$$\mathsf{m}(Z_t(A,B))^{\frac{1}{2d+1}} \geq (1-t)^{\frac{2d+3}{2d+1}}\mathsf{m}(A)^{\frac{1}{2d+1}} + t^{\frac{2d+3}{2d+1}}\mathsf{m}(B)^{\frac{1}{2d+1}}, \quad \forall\, t \in [0,1].$$

The above inequality is sharp, in the sense that if one replaces the exponent $2d+3$ with a smaller one the inequality fails for some choice of A, B.

8. Properties of the distortion coefficients

As we have discussed in Section 7, sub-Riemannian distortion coefficients present major differences with respects to the Riemannian case. In this section, we discuss some of their general properties. Henceforth, let $(\mathcal{D}, g)$ be a fixed ideal sub-Riemannian structure on M, and let $x, y \in M$, with $y \notin \mathrm{Cut}(x)$.

8.1. **Dependence on the distance.** Under the above assumptions, $y = \exp_x(\lambda)$ for a unique $\lambda \in T_x^*M$ such that $\|\lambda\| = \sqrt{2H(\lambda)} = d_{SR}(x, y)$. In particular, one can regard the sub-Riemannian distortion coefficients as a one-parameter family of functions depending on the initial covector $\lambda \in T^*M$ of a minimizing geodesic joining a pair of points $(x, y) \in M \times M \setminus \mathrm{Cut}(M)$. Loosely speaking:

$$\beta_t(x, \exp_x(\lambda)) = f_t(\lambda).$$

The basic Riemannian examples where the β_t's are explicit are space forms, where they depend on λ *only* through its (dual) norm $\|\lambda\| = d(x, y)$, see (2). As we discussed in Section 7.1, in the simplest sub-Riemannian structure, the Heisenberg group, the dependence on λ is fundamentally more complicated, and $\beta_t(x, y)$ is *not* a function of the sub-Riemannian distance between x and y. A similar phenomenon occurs in the case of the Grushin plane, treated in Section 7.3.

8.2. **Small time asymptotics: proof of Theorem 5.** For Riemannian structures, it is well known that

$$\beta_t(x, y) \sim C(x, y)t^n,$$

with $n = \dim(M)$. This is the reason for the presence of a normalization factor t^{-n} in the standard Riemannian distortion coefficients, which we did not include in Definition 2. In fact, in the genuinely sub-Riemannian case, the asymptotic is remarkably different, as stated in Theorem 5. This result follows directly from [ABR13, Sec. 5.6, Sec. 6.5]. We sketch a proof here for completeness.

Proof of Theorem 5. Let $\Sigma_x := M \setminus \mathrm{Cut}(x)$. The function $\mathfrak{f} : M \to \mathbb{R}$ defined by $\mathfrak{f}(y) = \frac{1}{2}d_{SR}^2(x, y)$ is smooth on Σ_x. Furthermore, $d_y\mathfrak{f} \in T_y^*M$ is the final covector of the unique geodesic joining x with y. In particular, define the homothety $\phi_t : \Sigma_x \to M$ of ratio $t \in [0, 1]$ and center $x \in M$ by the formula

$$\phi_t(y) := \pi \circ e^{(t-1)\vec{H}}(d_y\mathfrak{f}), \qquad \forall y \in \Sigma_x.$$

For all $\Omega \subset \Sigma_x$ we have $Z_t(x, \Omega) = \phi_t(\Omega)$. Since ϕ_t is a local diffeomorphism, and Σ_x is open, we have that

$$\beta_t(x, y) \sim \frac{\det(d_y\phi_t)}{\det(d_y\phi_1)}, \qquad \forall y \in \Sigma_x, \quad t \in [0, 1].$$

By [ABR13, Lemma 6.24], there exists $C(x, y) > 0$ such that

$$\beta_t(x, y) \sim C(x, y)t^{\mathcal{N}_\lambda},$$

where $\mathcal{N}_\lambda \in \mathbb{N} \cup \{+\infty\}$ is defined for all $\lambda \in T_x^*M$ in [ABR13, Def. 5.44].

As a consequence of [ABR13, Prop. 5.46], the function $\lambda \mapsto \mathcal{N}_\lambda$ is constant on an open Zariski set $A_x \subseteq T_x^*M$, where it attains its minimal value. In particular, this implies that the points $y \in \Sigma_x$ where $\mathcal{N}(x, y) > \min \mathcal{N}_\lambda$ has zero measure in Σ_x. The fact that $\mathcal{N}(x) \geq \dim(M)$ is [ABR13, Prop. 5.49]. □

We stress that the set of initial covectors $\lambda \in T_x^*M$ such that $\mathcal{N}_\lambda > \mathcal{N}(x)$ can be non-empty. This is the case for all Carnot groups with Goursat-type distribution and dimension $n \geq 4$, such as the Cartan and the Engel groups. See [Mun17].

Appendix A. Conjugate times and optimality: proof of Theorem 17

The aim of this appendix is to give a self-contained proof of the fact that geodesics not containing abnormal segments lose minimality after their first conjugate point, following [ABB16b, Sar80]. The main difference with respect to the proof of the analogue statement in the Riemannian setting is that the explicit formula for the second variation of energy is usually expressed in terms of Levi-Civita connection and curvature, which are not available in the sub-Riemannian setting. Hence one has to work with a suitable generalization of the index form on the space of controls. Here, the sub-Riemannian structure is not required to be ideal.

A minimizing geodesic is a horizontal curve γ_u, parametrized with constant speed, such that its control u is a solution of the constrained minimum problem:

$$\min\{J(v) \mid v \in \mathcal{U},\, E_x(v) = y\}. \tag{46}$$

Here $J : \mathcal{U} \to \mathbb{R}$ is the energy functional and $E_x : \mathcal{U} \to M$ is the end-point map based at x, where $\mathcal{U} \subseteq L^2([0,1],\mathbb{R}^m)$, cf. Section 2.2.

The Lagrange multipliers rule, in the normal case, gives the first order necessary condition for a control (and the corresponding horizontal curve) to be a minimizer: there exists $\lambda \in T_y^*M$, such that

$$\lambda \circ D_u E_x = D_u J. \tag{47}$$

Hence a solution of (46) is a pair (u,λ) satisfying (47). Higher order conditions for the minimality of γ_u are given by the second variation of J on the level sets of E_x. The second differential of the restriction to the level set is not in general the restriction of the second differential to the tangent space to the level set $T_u E_x^{-1}(y) = \ker D_u E_x$. The following formulas hold (see [ABB16b, Ch. 8] and [Rif14, Sec. 2.4]).

Proposition 68 (Second variation of the energy)**.** *Let $\gamma_u : [0,1] \to M$ be a normal geodesic containing no abnormal segments joining x with y satisfying (47) for some $\lambda \in T_y^*M$. Then, we have*

$$\mathrm{Hess}_u\, J|_{E_x^{-1}(y)}(v) = D_u^2 J(v) - \lambda \circ D_u^2 E_x(v), \qquad \forall\, v \in \ker D_u E_x.$$

Moreover we have

$$D_u^2 J(v) = \|v\|_{L^2}, \quad D_u^2 E_x(v) = \iint\limits_{0\leq\tau\leq t\leq 1} [(P_{\tau,1})_* X_{v(\tau)}, (P_{t,1})_* X_{v(t)}](y)\, d\tau dt,$$

where $X_{v(t)} := \sum_{i=1}^m v_i(t) X_i$ and $P_{\tau,t}$ denotes the flow of the non-autonomous vector field $X_{u(t)}$, with initial datum at time τ and final time t.

Given a pair (u,λ) such that γ_u is a normal geodesic satisfying the first order condition (47), we denote by $u^s(t) := su(st)$ the reparametrized control associated with the reparametrized trajectory $\gamma_{u^s}(t) = \gamma_u(st)$, both defined for $t \in [0,1]$. The covector $\lambda^s = s(P_{s,1}^*)\lambda \in T_{\gamma_u(s)}^*M$, is a Lagrange multiplier associated with u^s.

For normal geodesics containing no abnormal segments (see definition 16), conjugate points (in the sense of definition 15) can be characterized by the second variation of the energy, as in the Riemannian case, cf. [ABB16b, Ch. 8].

Proposition 69. *Assume that a normal geodesic $\gamma_u : [0,1] \to M$ contains no abnormal segments. Then $\gamma_u(s)$ is conjugate to $\gamma_u(0)$ if and only if* $\mathrm{Hess}_{u^s}\, J|_{E_x^{-1}(\gamma_u(s))}$ *is a degenerate quadratic form.*

The next two lemmas, proved in [ABB16b, Ch. 8], are crucial. For the reader's convenience, we provide a sketch of the proof.

Lemma 70. *Assume that a normal geodesic $\gamma_u : [0,1] \to M$ contains no abnormal segments. Define the function $\alpha : (0,1] \to \mathbb{R}$ as follows*

$$\alpha(s) := \inf\left\{\|v\|_{L^2}^2 - \lambda_s \circ D_{u^s}^2 E_x(v) \mid \|v\|_{L^2}^2 = 1,\ v \in \ker D_{u^s} E_x\right\}. \tag{48}$$

Then α is continuous and has the following properties:

(a) $\alpha(0) := \lim_{s\to 0} \alpha(s) = 1$;

(b) $\alpha(s) = 0$ *implies that* $\mathrm{Hess}_{u^s} J|_{E_x^{-1}(\gamma_u^s(1))}$ *is degenerate;*
(c) α *is monotone decreasing;*
(d) if $\alpha(\bar{s}) = 0$ *for some* $\bar{s} > 0$*, then* $\alpha(s) < 0$ *for* $s > \bar{s}$*.*

Sketch of the proof. Notice that one can write

$$\|v\|_{L^2}^2 - \lambda_s \circ D_{u^s}^2 E_x(v) = \langle (I - Q_s)(v)|v\rangle_{L^2},$$

where $Q_s : L^2([0,1],\mathbb{R}^m) \to L^2([0,1],\mathbb{R}^m)$ is a compact and symmetric operator As a consequence, one can prove that the infimum in (48) is attained.

Since every restriction $\gamma_u|_{[0,s]}$ is not abnormal, the rank of $D_{u^s}E_x$ is maximal, equal to n, for all $s \in (0,1]$. Then, by Riesz representation Theorem, we find a continuous orthonormal basis $\{v_i^s\}_{i\in\mathbb{N}}$ for $\ker D_{u^s}E_x$, yielding a continuous one-parameter family of isometries $\phi_s : \ker D_{u^s}E_x \to \mathcal{H}$ on a fixed Hilbert space $\mathcal{H}$. Since also $s \mapsto Q_s$ is continuous (in the norm topology), we reduce (48) to

$$\alpha(s) = 1 - \sup\{\langle \phi_s \circ Q_s \circ \phi_s^{-1}(w)|w\rangle_{\mathcal{H}} \mid w \in \mathcal{H},\ \|w\|_{\mathcal{H}} = 1\},$$

where the composition $\tilde{Q}_s := \phi_s \circ Q_s \circ \phi_s^{-1}$ is a continuous one-parameter family of symmetric and compact operators on a fixed Hilbert space $\mathcal{H}$. The supremum coincides with the largest eigenvalue of $\tilde{Q}_s$, which is well known to be continuous as a function of s if $\tilde{Q}_s$ is (see [Kat95, V Thm. 4.10]). This proves that α is continuous.

By a rescaling one can see that

$$D_{u^s}^2 E_x(v) = s^2 \iint\limits_{0\le\tau\le t\le 1} [(P_{s\tau,1})_* X_{v(s\tau)}, (P_{st,1})_* X_{v(st)}]|_{\gamma_u(s)} d\tau dt.$$

Taking the limit $s \to 0$, one can show that $Q_s \to 0$, hence $\tilde{Q}_s \to 0$, proving (a).

To prove (b), notice that $\alpha(\bar{s}) = 0$ means that $I - Q_{\bar{s}} \ge 0$, and that there exists a sequence $v_n \in \ker D_{u^{\bar{s}}}E_x$ of controls with $\|v_n\|_{L^2} = 1$ and such that $\|v_n\|_{L^2}^2 - \langle Q_{\bar{s}}(v_n)|v_n\rangle_{L^2} \to 0$ for $n \to \infty$. Up to extraction of a sub-sequence, we have that v_n is weakly convergent to some $\bar{v}$. By compactness of $Q_{\bar{s}}$, we deduce that $\langle Q_{\bar{s}}(\bar{v})|\bar{v}\rangle_{L^2} = 1$. Since $\|\bar{v}\|_{L^2}^2 \le 1$, we have $\langle (I - Q_{\bar{s}})(\bar{v})|\bar{v}\rangle_{L^2} = 0$. Being $I - Q_{\bar{s}}$ a bounded, non-negative symmetric operator, and since $\bar{v} \neq 0$, this implies that $A_{\bar{s}}$ is degenerate.

To prove (c) let us fix $0 \le s \le s' \le 1$ and $v \in \ker D_{u^s}E_x$. Define

$$\widehat{v}(t) := \begin{cases} \sqrt{\dfrac{s'}{s}}\, v\left(\dfrac{s'}{s}t\right), & 0 \le t \le \dfrac{s}{s'}, \\ 0, & \dfrac{s}{s'} < t \le 1. \end{cases}$$

It follows that $\|\widehat{v}\|_{L^2}^2 = \|v\|_{L^2}^2$, $\widehat{v} \in \ker D_{u^{s'}}E_x$, and $D_{u^s}^2 E_x(v) = D_{u^{s'}}^2 E_x(\widehat{v})$. As a consequence, $\alpha(s) \ge \alpha(s')$.

To prove (d), assume by contradiction that there exists $s_1 > \bar{s}$ such that $\alpha(s_1) = 0$. By monotonicity of point (c), $\alpha(s) = 0$ for every $\bar{s} \le s \le s_1$. This implies that every point in the image of $\gamma_u|_{[\bar{s},s_1]}$ is conjugate to $\gamma_u(0)$. Thanks to Lemma 71, the segment $\gamma_u|_{[\bar{s},s_1]}$ is abnormal, contradicting the assumption on γ_u. $\square$

Lemma 71. *Let* $\gamma : [0,1] \to M$ *be a normal geodesic that does not contain abnormal segments. Then the set* $\mathcal{T}_c := \{t > 0 \mid \gamma(t)$ *is conjugate to* $\gamma(0)\}$ *is discrete.*

Sketch of the proof. Let $\lambda(t)$ be a normal extremal associated with the geodesic $\gamma(t)$, satisfying condition (N) of Theorem 12. Assume that the set $\mathcal{T}_c$ has an accumulation point $\gamma(\bar{t})$. The fact that the Hamiltonian is non-negative yields the existence of a segment $\gamma|_{[\bar{t},\bar{t}+\varepsilon]}$ all of whose points are conjugate to $\gamma(0)$. A computation in local coordinates on T^*M shows that $\gamma|_{[\bar{t},\bar{t}+\varepsilon]}$ is an abnormal extremal, namely satisfies characterization (A) of Theorem 12. $\square$

We can now prove the following fundamental result.

Theorem 72. *Let $\gamma : [0,1] \to M$ be a normal geodesic that does not contain abnormal segments. Then,*

(i) $t_c := \inf\{t > 0 \mid \gamma(t)$ is conjugate to $\gamma(0)\} > 0$;
(ii) For every $s > t_c$ the curve $\gamma|_{[0,s]}$ is not a minimizer.

Proof. Claim (i) follows directly from Proposition 69 and (a)–(b) of Lemma 70 (or also, independently, from Lemma 71). Using also (d) of Lemma 70, one obtains claim (ii). Indeed, since the Hessian has a negative eigenvalue, we can find a variation joining the same end-points and shorter than the original geodesic, contradicting the minimality assumption. □

By applying Theorem 72 to every restriction $\gamma|_{[s_1,s_2]}$ with $0 \leq s_1 < s_2 < 1$, we obtain Theorem 17 stated in Section 2.

Appendix B. Positivity: proof of Lemma 29

Proof of Lemma 29. Recall that $E_1(t),\dots,E_n(t),F_1(t),\dots,F_n(t)$ is a fixed Darboux frame along the normal extremal $\lambda : [0,1] \to T^*M$, with initial covector $\lambda(0) = d_x\phi$. For all $s \in [0,1]$, consider the Jacobi matrices $\mathbf{J}_s^{\mathrm{V}}(t)$ and $\mathbf{J}_s^{\mathrm{H}}(t)$ defined in Section 3.3, representing the family of Lagrange subspaces

$$e_*^{(t-s)\vec{H}} \operatorname{span}\{E_1(s),\dots,E_n(s)\}, \qquad \text{and} \qquad e_*^{(t-s)\vec{H}} \operatorname{span}\{F_1(s),\dots,F_n(s)\},$$

respectively. Notice that $N_0^{\mathrm{V}}(0) = \mathbb{0}$ and, by the assumption of the Lemma, $N_0^{\mathrm{V}}(t)$ is non-degenerate for all $t \in (0,1)$. We define $K(t) := N_0^{\mathrm{V}}(t)^{-1}$.

We prove (a) for $t \in (0,1)$. Since no point $\gamma(t)$ is conjugate to $\gamma(0)$ for $t \in (0,1)$, it is sufficient to prove that $\det K(t) > 0$ for small $t > 0$. By applying Lemma 24 to the Jacobi matrix $\mathbf{J}_0^{\mathrm{V}}(t)$, we obtain that $W(t) := N_0^{\mathrm{V}}(t) M_0^{\mathrm{V}}(t)^{-1}$ is symmetric and satisfies the Riccati equation

$$\dot{W} = B(t) + A(t)^* W + W A(t) + W R(t) W, \qquad W(0) = \mathbb{0}. \tag{49}$$

Equation (49) holds provided that $M_0^{\mathrm{V}}(t)$ is non-degenerate which, since $M_0^{\mathrm{V}}(0) = \mathbb{1}$, holds true for sufficiently small $t > 0$. Again by Lemma 24, $B(t) \geq 0$. Hence, a direct application of the matrix Riccati comparison theorem [BR16, Appendix A] yields that $W(t) \geq 0$ for $t \in [0,\varepsilon]$. Moreover, since $M_0^{\mathrm{V}}(0) = \mathbb{1}$, and $N_0^{\mathrm{V}}(t)$ is non-degenerate for $t \in (0,1)$, we have that $W(t) = N_0^{\mathrm{V}}(t) M_0^{\mathrm{V}}(t)^{-1} > 0$ for small t. In particular $\det N_0^{\mathrm{V}}(t) M_0^{\mathrm{V}}(t)^{-1} > 0$, which implies $\det K(t) > 0$, yielding (a).

To prove (b) and (c), we introduce a change of basis lemma, and the S matrix.

Lemma 73 (Change of basis)**.** *For all $s \in (0,1)$ and $t \in [0,1]$, we have*

$$\mathbf{J}_s^{\mathrm{V}}(t) = -\mathbf{J}_0^{\mathrm{V}}(t) N_0^{\mathrm{V}}(s)^{-1} N_0^{\mathrm{H}}(s) N_s^{\mathrm{V}}(0) + \mathbf{J}_0^{\mathrm{H}}(t) N_s^{\mathrm{V}}(0). \tag{50}$$

Moreover, for the original Jacobi matrix $\mathbf{J}(t)$, we have

$$\mathbf{J}(t) = \mathbf{J}_0^{\mathrm{V}}(t) M(0) + \mathbf{J}_0^{\mathrm{H}}(t). \tag{51}$$

Proof. To prove (50), let $\mathbf{J}_s^{\mathrm{V}}(t) = \mathbf{J}_0^{\mathrm{V}}(t) C_{\mathrm{V}} + \mathbf{J}_0^{\mathrm{H}}(t) C_{\mathrm{H}}$ for unique $n \times n$ matrices $C_{\mathrm{H}}, C_{\mathrm{V}}$. These can be computed by evaluating the horizontal component of both sides at times $t = s$ and $t = 0$. To prove (51), let $\mathbf{J}(t) = \mathbf{J}_0^{\mathrm{V}}(t) D_{\mathrm{V}} + \mathbf{J}_0^{\mathrm{H}}(t) D_{\mathrm{H}}$ for unique $n \times n$ matrices $D_{\mathrm{H}}, D_{\mathrm{V}}$. The latter can computed by evaluating both the horizontal and vertical components of $\mathbf{J}(t) = \left(\begin{smallmatrix} M(t) \\ N(t) \end{smallmatrix}\right)$ at time $t = 0$. □

Lemma 74 (S matrix)**.** *Consider the smooth family of $n \times n$ matrices*

$$S(t) := N_0^{\mathrm{V}}(t)^{-1} N_0^{\mathrm{H}}(t), \qquad \forall\, t \in (0,1).$$

Such a matrix is symmetric and $\dot{S}(t) \leq 0$.

Proof. In order to prove the lemma, we start by clarifying the geometric interpretation of $S(t)$. Indeed, observe that, letting

$$Z(t) := E(t) \cdot (M_0^{\mathrm{V}}(t)S(t) - M_0^{\mathrm{H}}(t)),$$

we have

$$E(0) \cdot S(t) - F(0) = e_*^{-t\vec{H}} Z(t). \tag{52}$$

In particular, $Z(t)$ represents a n-tuple of vertical vector fields along $\lambda(t)$, and the left hand side of (52) generates the smooth curve of Lagrange subspaces $\Lambda(t) := e_*^{-t\vec{H}} \mathcal{V}_{\lambda(t)} \subset T_{\lambda(0)}(T^*M)$. In particular $S(t)$ is symmetric, since[4]

$$0 = \sigma\left(E(0) \cdot S(t) - F(0), E(0) \cdot S(t) - F(0)\right) = S(t) - S(t)^*,$$

and $S(t)$ is non-increasing:

$$\dot{S}(t) = \sigma_{\lambda(0)}\left(e_*^{-t\vec{H}} Z(t), \frac{d}{dt} e_*^{-t\vec{H}} Z(t)\right) = -2H(Z(t)) \leq 0, \qquad \forall t \in (0, s],$$

where in the last equality we identified $Z(t) \in \mathcal{V}_{\lambda(t)} \simeq T^*_{\gamma(t)}M$. The above inequality holds for any smooth family of vertical vector fields $Z(t)$ along $\lambda(t)$, and follows from a straightforward computation in local coordinates around $\lambda(t)$. □

Using Lemma 73, one can check that (b) and (c) are equivalent to

(b') $S(t) - S(s) \geq 0$, for all $t \in (0, s]$,
(c') $M(0) + S(s) \geq 0$, for all $t \in (0, 1)$.

By Lemma 74, $\dot{S}(t) \leq 0$ for $t \in (0, 1)$, thus proving assertion (b'). To prove (c'), which a fortiori does not depend on t, recall that by the assumptions of Theorem 26,

$$\frac{1}{2} d_{SR}^2(z, y) + \phi(z) \geq 0, \qquad \forall z \in M, \tag{53}$$

with equality at $z = x$. Moreover

$$\frac{1}{2s} d_{SR}^2(z, \gamma(s)) \geq \frac{1}{2} d_{SR}^2(z, y) - \frac{1-s}{2} d_{SR}^2(x, y) \qquad \forall z \in M, \quad s \in (0, 1]. \tag{54}$$

See [CEMS01, Claim 2.4] for a proof of (54) in the Riemannian setting. The proof carries over to the sub-Riemannian case, and is solely a consequence of the triangular and the arithmetic-geometric inequalities. In particular, a property similar to (53) holds replacing y with any midpoint $\gamma(s) \in Z_s(x, y)$, that is

$$\frac{1}{2s} d_{SR}^2(z, \gamma(s)) + \phi(z) \geq \mathrm{const}(s, x, y), \qquad \forall z \in M, \quad s \in (0, 1),$$

with equality when $z = x$. By Theorem 17, $\gamma(s)$ is not conjugate to $x = \gamma(0)$ along the unique minimizing curve γ joining x with y, which is not abnormal. Hence $\gamma(s) \notin \mathrm{Cut}(x)$, and $c_s(z) := d_{SR}^2(z, \gamma(s))/2s$ is smooth at $z = x$. Furthermore, ϕ is two times differentiable at x by the assumptions of Theorem 26. Hence, $z \mapsto c_s(z) + \phi(z)$ has a critical point at $z = x$ and a well defined non-negative Hessian

$$\mathrm{Hess}(c_s + \phi)|_x \geq 0, \tag{55}$$

as a quadratic form on T_xM. We claim that (55) is equivalent to (c'). To prove this claim, we use the next Lemma, which is essentially, a rewording of Lemma 21.

[4] Here, for n-tuples V, W, the pairing $\sigma(V, W)$ denotes the matrix $\sigma(V_i, W_j)$. Moreover, if A is an $n \times n$ matrix, the notation $W = V \cdot A$ denotes the n-tuple W whose i-th element is $W_i = \sum_{j=1}^n A_{ji} V_j$.

Lemma 75 (Second differential and Hessian)**.** *Let $f, g : M \to \mathbb{R}$, twice differentiable at $x \in M$, and such that x is a critical point for $f+g$. Then*

$$d_x^2 f - d_x^2(-g) = \mathrm{Hess}(f+g)|_x, \tag{56}$$

where we used the fact that the space of second differentials at $\lambda = d_x f = d_x(-g)$ is an affine space on the space of quadratic forms on T_xM.

Remark 76. The difference of second differentials in the left hand side of (56) is a linear map $T_xM \to T_\lambda(T^*M)$, with values in $\mathcal{V}_\lambda = T_\lambda(T_x^*M) \simeq T_x^*M$, and it is identified with the quadratic form $\mathrm{Hess}(f+g)|_x : T_xM \times T_xM \to \mathbb{R}$, i.e. the Hessian of $f+g$ at the critical point x.

We intend to apply Lemma 75 to $\phi + c_s$, which has a minimum point at x. Since $d_x(-c_s) = d_x\phi$, both $e_*^{t\vec{H}} \circ d_x^2\phi(X(0))$ and $e_*^{t\vec{H}} \circ d_x^2(-c_s)(X(0))$ are n-tuples of Jacobi fields along the same extremal $\lambda(t) = e^{t\vec{H}}(d_x\phi)$. We exploit the relation with Jacobi matrices to compute both second differentials of ϕ and $-c_s$ separately.

Since c_s is smooth in a neighbourhood $\mathcal{O}_x$ of x, by [Rif14, Lemma 2.15], we have that $\exp_z(sd_z(-c_s)) = \gamma(s)$ for all $z \in \mathcal{O}_x$ and $s \in (0,1)$. Thus,

$$\pi \circ e^{s\vec{H}}(d_z(-c_s)) = \gamma(s), \quad \forall z \in \mathcal{O}_x \qquad \Rightarrow \qquad e_*^{s\vec{H}} \circ d_x^2(-c_s)(T_xM) = \mathcal{V}_{\lambda(s)}. \tag{57}$$

Equation (57) implies that the n-tuple of Jacobi fields $e_*^{t\vec{H}} \circ d_x^2(-c_s)(X(0))$ is associated with the Jacobi matrix $\mathbf{J}_s^{\mathrm{V}}(t)L_s$, for some $n\times n$ matrix L_s. Evaluating at $t=0$, we obtain $L_s = N_s^{\mathrm{V}}(0)^{-1}$. More precisely, for all $s \in (0,1)$ we have

$$e_*^{t\vec{H}} \circ d_x^2(-c_s)(X(0)) = E(t)\cdot M_s^{\mathrm{V}}(t)N_s^{\mathrm{V}}(0)^{-1} + F(t)\cdot N_s^{\mathrm{V}}(t)N_s^{\mathrm{V}}(0)^{-1}, \quad t \in [0,1]. \tag{58}$$

Furthermore, by definition of the Jacobi matrix $\mathbf{J}(t) = \begin{pmatrix} M(t) \\ N(t) \end{pmatrix}$, we have

$$e_*^{t\vec{H}} \circ d_x^2\phi(X(0)) = E(t)\cdot M(t) + F(t)\cdot N(t), \qquad t \in [0,1]. \tag{59}$$

By evaluating (58) and (59) at $t=0$, we obtain

$$\begin{aligned} d_x^2(-c_s)(X(0)) &= E(0)\cdot M_s^{\mathrm{V}}(0)N_s^{\mathrm{V}}(0)^{-1} + F(0), \\ d_x^2\phi(X(0)) &= E(0)\cdot M(0) + F(0). \end{aligned}$$

In particular, from (55) and Lemma 75 we finally prove (c'), since

$$\begin{aligned} 0 \le \mathrm{Hess}(\phi + c_s)|_x &= M(0) - M_s^{\mathrm{V}}(0)N_s^{\mathrm{V}}(0)^{-1} \\ &= M(0) + \underbrace{M_0^{\mathrm{V}}(0)}_{=\mathbb{1}}\underbrace{N_0^{\mathrm{V}}(s)^{-1}N_0^{\mathrm{H}}(s)}_{=S(s)} - \underbrace{M_0^{\mathrm{H}}(0)}_{=\mathbb{0}}, \end{aligned}$$

where, in the second line, we used Lemma 73 to eliminate $M_s^{\mathrm{V}}(0)$. □

Appendix C. Density formula: proof of Lemma 40

In the proof of Theorem 39 we used a slightly more general reformulation of [AGS08, Lemma 5.5.3] for non-injective maps. The proof is essentially the same as in the aforementioned reference, and we report it here for completeness.

Proof of Lemma 40. We start by proving that if $\det(\tilde{d}_x f) > 0$ for $\mathcal{L}^d$-a.e. $x \in \Sigma$, then $f_\sharp(\rho\mathcal{L}^d) \ll \mathcal{L}^d$. For any Borel function $h : \mathbb{R}^d \to [0,+\infty]$, we have the area formula for approximately differentiable maps [AGS08, Eq. 5.5.2], that is

$$\int_{\Sigma_f} h(x)|\det(\tilde{d}_x f)|dx = \int_{\mathbb{R}^d} \sum_{x \in \tilde{f}^{-1}(y)\cap\Sigma_f} h(x)dy. \tag{60}$$

Since f is measurable, then $f(x) = \tilde{f}(x)$ up to a $\mathcal{L}^d$-negligible set (see [AGS08, Remark 5.5.2]), and hence $f_\sharp(\rho\mathcal{L})^d = \tilde{f}_\sharp(\mathcal{L}^d)$. Since $|\det(\tilde{d}_x f)| > 0$ for $\mathcal{L}^d$-a.e. $x \in \Sigma \subseteq \Sigma_f$, the function $h : \mathbb{R}^d \to [0, +\infty]$ given by

$$h(x) := \begin{cases} \frac{\rho(x)\chi_{\tilde{f}^{-1}(B)\cap\Sigma}(x)}{|\det(\tilde{d}_x f)|} & x \in \Sigma, \\ 0 & \text{otherwise}, \end{cases}$$

is Borel and well defined. Hence, for any Borel set $B \subset \mathbb{R}^d$, we obtain

$$\begin{aligned} f_\sharp(\rho\mathcal{L}^d)(B) = (\rho\mathcal{L}^d)(\tilde{f}^{-1}(B)\cap\Sigma) &= \int_{\tilde{f}^{-1}(B)\cap\Sigma} \rho(x)\, dx \\ &= \int_{\Sigma_f} \rho(x)\chi_{\tilde{f}^{-1}(B)\cap\Sigma}(x)\, dx \\ &= \int_{\mathbb{R}^d} \sum_{x\in\tilde{f}^{-1}(y)\cap\Sigma_f} \frac{\rho(x)\chi_{\tilde{f}^{-1}(B)\cap\Sigma}(x)}{|\det(\tilde{d}_x f)|}\, dy \\ &= \int_B \sum_{x\in\tilde{f}^{-1}(y)\cap\Sigma} \frac{\rho(x)}{|\det(\tilde{d}_x f)|}\, dy, \end{aligned}$$

where in the last line we used (60). In particular if $\mathcal{L}^d(B) = 0$, then also $f_\sharp(\rho\mathcal{L}^d)(B)$.

The inverse implication is proved by contradiction. Assume that there exists a Borel set $B \subset \Sigma$ with $\mathcal{L}^d(B) > 0$ and $\det(\tilde{d}_x f) = 0$ on B. Then, the area formula (60) with $h = \chi_{B\cap\Sigma}$ yields

$$0 = \int_{\mathbb{R}^d} \sum_{x\in\Sigma_f\cap\tilde{f}^{-1}(y)} \chi_{B\cap\Sigma}(x)\, dy \geq \mathcal{L}^d(\tilde{f}(B)).$$

On the other hand, since $f_\sharp(\rho\mathcal{L}^d) = \tilde{f}_\sharp(\rho\mathcal{L}^d)$, we have

$$f_\sharp(\rho\mathcal{L}^d)(\tilde{f}(B)) = \int_{\tilde{f}^{-1}(\tilde{f}(B))} \rho(x)\, dx \geq \int_B \rho(x)\, dx = \mathcal{L}^d(B) > 0.$$

Thus $f_\sharp(\rho\mathcal{L}^d)$ cannot be absolutely continuous w.r.t. $\mathcal{L}^d$. □

Acknowledgments

This work was supported by the Grant ANR-15-CE40-0018 of the ANR, and by a public grant as part of the Investissement d'avenir project, reference ANR-11-LABX-0056-LMH, LabEx LMH, in a joint call with the "FMJH Program Gaspard Monge in optimization and operation research". This work has been supported by the ANR project ANR-15-IDEX-02.

We wish to thank Andrei Agrachev, Luigi Ambrosio, Zoltan Balogh, Ludovic Rifford and Kinga Sipos for stimulating discussions. We also thank the anonymous referees for their comments.

References

[ABB12] A. Agrachev, D. Barilari, and U. Boscain, *On the Hausdorff volume in sub-Riemannian geometry*, Calc. Var. Partial Differential Equations **43** (2012), no. 3-4, 355–388.

[ABB16a] A. Agrachev, D. Barilari, and U. Boscain, *Introduction to geodesics in sub-Riemannian geometry.*, Geometry, analysis and dynamics on sub-Riemannian manifolds. Volume II, 2016, pp. 1–83 (English).

[ABB16b] A. Agrachev, D. Barilari, and U. Boscain, *Introduction to Riemannian and sub-Riemannian geometry (Lecture Notes). http://webusers.imj-prg.fr/ davide.barilari/notes.php* (2016). v20/11/16.

[ABCK97] A. Agrachev, B. Bonnard, M. Chyba, and I. Kupka, *Sub-Riemannian sphere in Martinet flat case*, ESAIM Control Optim. Calc. Var. **2** (1997), 377–448.

[ABR13] A. Agrachev, D. Barilari, and L. Rizzi, *Curvature: a variational approach*, Memoirs of the AMS (in press) (June 2013), available at 1306.5318.

[ABS08] A. Agrachev, U. Boscain, and M. Sigalotti, *A Gauss-Bonnet-like formula on two-dimensional almost-Riemannian manifolds*, Discrete Contin. Dyn. Syst. **20** (2008), no. 4, 801–822.

[Agr09] A. Agrachev, *Any sub-Riemannian metric has points of smoothness*, Dokl. Akad. Nauk **424** (2009), no. 3, 295–298.

[AGS08] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Second, Lectures in Mathematics ETH Zürich, Birkhäuser Verlag, Basel, 2008.

[AL09] A. Agrachev and P. W. Y. Lee, *Optimal transportation under nonholonomic constraints*, Trans. Amer. Math. Soc. **361** (2009), no. 11, 6019–6047. MR2529923

[AL14] ———, *Generalized Ricci curvature bounds for three dimensional contact subriemannian manifolds*, Math. Ann. **360** (2014), no. 1-2, 209–253. MR3263162

[AR04] L. Ambrosio and S. Rigot, *Optimal mass transportation in the Heisenberg group*, J. Funct. Anal. **208** (2004), no. 2, 261–301. MR2035027

[AZ02a] A. Agrachev and I. Zelenko, *Geometry of Jacobi curves. I*, J. Dynam. Control Systems **8** (2002), no. 1, 93–140.

[AZ02b] ———, *Geometry of Jacobi curves. II*, J. Dynam. Control Systems **8** (2002), no. 2, 167–215.

[BG17] F. Baudoin and N. Garofalo, *Curvature-dimension inequalities and Ricci lower bounds for sub-Riemannian manifolds with transverse symmetries*, J. Eur. Math. Soc. (JEMS) **19** (2017), no. 1, 151–219. MR3584561

[BKS17] Z. M. Balogh, A. Kristály, and K. Sipos, *Jacobian determinant inequality on Corank 1 Carnot groups with applications*, ArXiv e-prints (Jan. 2017), available at 1701.08831.

[BKS18] Z. M. Balogh, A. Kristály, and K. Sipos, *Geometric inequalities on Heisenberg groups*, Calc. Var. Partial Differential Equations **57** (2018), no. 2, Art. 61, 41. MR3774461

[BKW16] F. Baudoin, B. Kim, and J. Wang, *Transverse Weitzenböck formulas and curvature dimension inequalities on Riemannian foliations with totally geodesic leaves*, Comm. Anal. Geom. **24** (2016), no. 5, 913–937. MR3622309

[BM16] I. Beschastnyi and A. Medvedev, *Left-invariant Sub-Riemannian Engel structures: abnormal geodesics and integrability*, ArXiv e-prints (Nov. 2016), available at 1611.03634.

[BR16] D. Barilari and L. Rizzi, *Comparison theorems for conjugate points in sub-Riemannian geometry*, ESAIM Control Optim. Calc. Var. **22** (2016), no. 2, 439–472.

[BR17a] Z. Badreddine and L. Rifford, *Measure contraction properties for two-step sub-Riemannian structures and medium-fat Carnot groups*, ArXiv e-prints (Dec. 2017), available at 1712.09900.

[BR17b] D. Barilari and L. Rizzi, *On Jacobi fields and a canonical connection in sub-Riemannian geometry*, Archivum Mathematicum **53** (2017), no. 2, 77–92.

[BR17c] ———, *Sharp measure contraction property for generalized H-type Carnot groups*, Commun. Contemp. Math. (in press) (Feb. 2017), available at 1702.04401.

[Bre99] Y. Brenier, *Minimal geodesics on groups of volume-preserving maps and generalized solutions of the Euler equations*, Comm. Pure Appl. Math. **52** (1999), no. 4, 411–452.

[CEMS01] D. Cordero-Erausquin, R. J. McCann, and M. Schmuckenschläger, *A Riemannian interpolation inequality à la Borell, Brascamp and Lieb*, Invent. Math. **146** (2001), no. 2, 219–257.

[CJT06] Y. Chitour, F. Jean, and E. Trélat, *Genericity results for singular curves*, J. Differential Geom. **73** (2006), no. 1, 45–73.

[CR08] P. Cannarsa and L. Rifford, *Semiconcavity results for optimal control problems admitting no singular minimizing controls*, Ann. Inst. H. Poincaré Anal. Non Linéaire **25** (2008), no. 4, 773–802.

[CS04] P. Cannarsa and C. Sinestrari, *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*, Progress in Nonlinear Differential Equations and their Applications, vol. 58, Birkhäuser Boston, Inc., Boston, MA, 2004.

[EG15] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*, Revised, Textbooks in Mathematics, CRC Press, Boca Raton, FL, 2015.

[FJ08] A. Figalli and N. Juillet, *Absolute continuity of Wasserstein geodesics in the Heisenberg group*, J. Funct. Anal. **255** (2008), no. 1, 133–141. MR2417812

[FR10] A. Figalli and L. Rifford, *Mass transportation on sub-Riemannian manifolds*, Geom. Funct. Anal. **20** (2010), no. 1, 124–159.

[Gar02] R. J. Gardner, *The Brunn-Minkowski inequality*, Bull. Amer. Math. Soc. (N.S.) **39** (2002), no. 3, 355–405. MR1898210

[Jea14] F. Jean, *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*, SpringerBriefs in Mathematics, Springer, Cham, 2014.

[Jui09] N. Juillet, *Geometric inequalities and generalized Ricci bounds in the Heisenberg group*, Int. Math. Res. Not. IMRN **13** (2009), 2347–2373.

[Kat95] T. Kato, *Perturbation theory for linear operators*, Classics in Mathematics, Springer-Verlag, Berlin, 1995. Reprint of the 1980 edition.

[Lee13] P. W. Y. Lee, *Displacement interpolations from a Hamiltonian point of view*, J. Funct. Anal. **265** (2013), no. 12, 3163–3203. MR3110498

[LLZ16] P. W. Y. Lee, C. Li, and I. Zelenko, *Ricci curvature type lower bounds for sub-Riemannian structures on Sasakian manifolds*, Discrete Contin. Dyn. Syst. **36** (2016), no. 1, 303–321. MR3369223

[LM05] G. P. Leonardi and S. Masnou, *On the isoperimetric problem in the Heisenberg group* $\mathbb{H}^n$, Ann. Mat. Pura Appl. (4) **184** (2005), no. 4, 533–553.

[LV09] J. Lott and C. Villani, *Ricci curvature for metric-measure spaces via optimal transport*, Ann. of Math. (2) **169** (2009), no. 3, 903–991.

[McC01] R. J. McCann, *Polar factorization of maps on Riemannian manifolds*, Geom. Funct. Anal. **11** (2001), no. 3, 589–608.

[Mit85] J. Mitchell, *On Carnot-Carathéodory metrics*, J. Differential Geom. **21** (1985), no. 1, 35–45. MR806700

[MM16] A. Montanari and D. Morbidelli, *On the lack of semiconcavity of the subRiemannian distance in a class of Carnot groups*, J. Math. Anal. Appl. **444** (2016), no. 2, 1652–1674.

[MM17] ———, *On the subRiemannian cut locus in a model of free two-step Carnot group*, Calc. Var. Partial Differential Equations **56** (2017), no. 2, Paper No. 36, 26.

[MM92] M. Marcus and H. Minc, *A survey of matrix theory and matrix inequalities*, Dover Publications, Inc., New York, 1992. Reprint of the 1969 edition.

[Mon02] R. Montgomery, *A tour of subriemannian geometries, their geodesics and applications*, Mathematical Surveys and Monographs, vol. 91, American Mathematical Society, Providence, RI, 2002.

[Mon03] R. Monti, *Brunn-Minkowski and isoperimetric inequality in the Heisenberg group*, Ann. Acad. Sci. Fenn. Math. **28** (2003), no. 1, 99–109.

[Mun17] I. H. Munive, *Sub-Riemannian curvature of Carnot groups with rank-two distributions*, J. Dyn. Control Syst. **23** (2017), no. 4, 779–814. MR3688894

[Oht07] S.-i. Ohta, *On the measure contraction property of metric measure spaces*, Comment. Math. Helv. **82** (2007), no. 4, 805–828. MR2341840

[Oht09] ———, *Finsler interpolation inequalities*, Calc. Var. Partial Differential Equations **36** (2009), no. 2, 211–249. MR2546027

[Rif13a] L. Rifford, *Ricci curvatures in Carnot groups*, Math. Control Relat. Fields **3** (2013), no. 4, 467–487.

[Rif13b] ———, *Ricci curvatures in Carnot groups*, Math. Control Relat. Fields **3** (2013), no. 4, 467–487.

[Rif14] ———, *Sub-Riemannian geometry and optimal transport*, Springer Briefs in Mathematics, Springer, Cham, 2014.

[Riz16] L. Rizzi, *Measure contraction properties of Carnot groups*, Calc. Var. Partial Differential Equations **55** (2016), no. 3, Art. 60, 20.

[Riz17] L. Rizzi, *A counterexample to gluing theorems for MCP metric measure spaces*, ArXiv e-prints (Nov. 2017), available at 1711.04499.

[RS17] L. Rizzi and U. Serres, *On the cut locus of free, step two Carnot groups*, Proc. Amer. Math. Soc. **145** (2017), no. 12, 5341–5357. MR3717961

[RT05] L. Rifford and E. Trélat, *Morse-Sard type results in sub-Riemannian geometry*, Math. Ann. **332** (2005), no. 1, 145–159.

[RYN18] M. Ritoré and J. Yepes Nicolás, *Brunn-Minkowski inequalities in product metric measure spaces*, Adv. Math. **325** (2018), 824–863. MR3742604

[Sar80] A. V. Sarychev, *Index of second variation of a control system*, Mat. Sb. (N.S.) **113(155)** (1980), no. 3(11), 464–486, 496.

[Stu06a] K.-T. Sturm, *On the geometry of metric measure spaces. I*, Acta Math. **196** (2006), no. 1, 65–131.

[Stu06b] ———, *On the geometry of metric measure spaces. II*, Acta Math. **196** (2006), no. 1, 133–177.

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

[Vil17] C. Villani, *Séminaire Bourbaki, Volume 2016/2017, exposé 1127*, Astérisque (2017).

[ZL09] I. Zelenko and C. Li, *Differential geometry of curves in Lagrange Grassmannians with given Young diagram*, Differential Geom. Appl. **27** (2009), no. 6, 723–742.