



HAL
open science

Measure contraction properties for two-step analytic sub-Riemannian structures and Lipschitz Carnot groups

Zeinab Badreddine, Ludovic Rifford

► **To cite this version:**

Zeinab Badreddine, Ludovic Rifford. Measure contraction properties for two-step analytic sub-Riemannian structures and Lipschitz Carnot groups. 2018. hal-01662544v1

HAL Id: hal-01662544

<https://hal.science/hal-01662544v1>

Preprint submitted on 25 Dec 2017 (v1), last revised 17 Dec 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Measure contraction properties for two-step sub-Riemannian structures and medium-fat Carnot groups

Z. Badreddine* L. Rifford†

Abstract

We prove that two-step sub-Riemannian structures on a compact manifold equipped with a smooth measure and medium-fat Carnot groups satisfy measure contraction properties.

1 Introduction

The aim of this paper is to provide new examples of sub-Riemannian structures satisfying measure contraction properties. Let M be a smooth manifold of dimension $n \geq 3$ equipped with a sub-Riemannian structure (Δ, g) of rank $m < n$, whose geodesic distance d_{SR} is supposed to be complete. We refer the reader to Appendix A for the notations used throughout the paper. As in the previous paper of the second author on the same subject [23], we restrict our attention to the notion of measure contraction properties in metric measured spaces with negligible cut loci (if $A \subset M$ is a Borel set then $\mathcal{L}^n(A) = 0$ means that A has vanishing n -dimensional Lebesgue measures in charts):

Definition 1. *We say that the sub-Riemannian structure (Δ, g) on M has negligible cut loci if for every $x \in M$, there is a measurable set $\mathcal{C}(x) \subset M$ with*

$$\mathcal{L}^n(\mathcal{C}(x)) = 0,$$

and a measurable map $\gamma_x : (M \setminus \mathcal{C}(x)) \times [0, 1] \rightarrow M$ such that for every $y \in M \setminus \mathcal{C}(x)$ the curve

$$s \in [0, 1] \mapsto \gamma_x(s, y)$$

is the unique minimizing horizontal path from x to y .

*Université Côte d'Azur, CNRS, Inria, Labo. J.-A. Dieudonné, UMR CNRS 7351, Parc Valrose 06108 Nice, Cedex 2, France (Zeinab.Badreddine@unice.fr)

†Université Côte d'Azur, CNRS, Inria, Labo. J.-A. Dieudonné, UMR CNRS 7351, Parc Valrose 06108 Nice, Cedex 2, France (Ludovic.Rifford@math.cnrs.fr)

Measure contraction properties consists in comparing the contraction of volumes along minimizing geodesics from a given point with what happens in classical model spaces of Riemannian geometry. We recall that for every $K \in \mathbb{R}$, the comparison function $s_K : [0, +\infty) \rightarrow [0, +\infty)$ ($s_K : [0, \pi/\sqrt{K}) \rightarrow [0, +\infty)$ if $K > 0$) is defined by

$$s_K(t) := \begin{cases} \frac{\sin(\sqrt{K}t)}{\sqrt{K}} & \text{if } K > 0 \\ t & \text{if } K = 0 \\ \frac{\sinh(\sqrt{-K}t)}{\sqrt{-K}} & \text{if } K < 0. \end{cases}$$

In our setting, the following definition is equivalent to the notion of measure contraction property introduced by Ohta in [21] for more general measured metric spaces (see also [28]).

Definition 2. *Let (Δ, g) be a sub-Riemannian structure on M with negligible cut loci, μ a measure absolutely continuous with respect to \mathcal{L}^n and $K \in \mathbb{R}, N > 1$ be fixed. We say that (Δ, g) equipped with μ satisfies MCP(K, N) if for every $x \in M$ and every measurable set $A \subset M \setminus \mathcal{C}(x)$ (provided that $A \subset B_{SR}(x, \pi\sqrt{N-1/K})$ if $K > 0$) with $0 < \mu(A) < \infty$,*

$$\mu(A_s) \geq \int_A s \left[\frac{s_K(sd_{SR}(x, z)/\sqrt{N-1})}{s_K(d_{SR}(x, z)/\sqrt{N-1})} \right]^{N-1} d\mu(z) \quad \forall s \in [0, 1],$$

where A_s is the s -interpolation of A from x defined by

$$A_s := \left\{ \gamma_x(s, y) \mid y \in A \setminus \mathcal{C}(x) \right\} \quad \forall s \in [0, 1].$$

In particular, (Δ, g) equipped with μ satisfies MCP($0, N$) if for every $x \in M$ and every measurable set $A \subset M \setminus \mathcal{C}(x)$ with $0 < \mu(A) < \infty$,

$$\mu(A_s) \geq s^N \mu(A) \quad \forall s \in [0, 1].$$

To our knowledge, the first study of measure contraction properties in the sub-Riemannian setting has been performed by Juillet in his thesis. In [16], Juillet proved that the n -th Heisenberg group \mathbb{H}^n (with $n \geq 1$) equipped with its sub-Riemannian distance and the Lebesgue measure \mathcal{L}^{2n+1} (in this case the ambient space is \mathbb{R}^{2n+1}) satisfies MCP($0, 2n+3$). This result is sharp for two reasons. First, Juillet proved that \mathbb{H}^n does not satisfy any other stronger notion of "Ricci curvature bounded from below" in metric measured spaces such as for example the so-called curvature dimension property (see [19, 27, 28, 29]). Secondly, Juillet showed that $2n+3$ is the optimal dimension for which \mathbb{H}^n satisfies MCP($0, N$), there is no $N < 2n+3$ such that \mathbb{H}^n

(equipped with d_{SR} and \mathcal{L}^{2n+1}) satisfies $MCP(0, N)$. The Juillet's Theorem, which settled the case of the simplest sub-Riemannian structures, paved the way to the study of measure contraction properties for more general sub-Riemannian structures. In [6], Agrachev and Lee investigated the case of sub-Riemannian structures associated with contact distributions in dimension 3. In [17, 18], Lee and Lee, Li and Zelenko studied the particular case of Sasakian manifolds. In [23], the second author proved that any ideal Carnot group satisfy $MCP(0, N)$ for some $N > 1$ (it has been shown later by Rizzi [26] that a Carnot group is ideal if and only if it is fat). In [26], Rizzi showed that any co-rank 1 Carnot group of dimension $k+1$ (equipped with the sub-Riemannian distance and a left-invariant measure) satisfies $MCP(0, k+3)$. Finally, more recently, Barilari and Rizzi [9] proved that H -type Carnot groups of rank k and dimension n satisfy $MCP(0, k+3(n-k))$. The purpose of the present paper is to pursue the qualitative approach initiated by the second author in [23]. We aim to show that some assumptions on the sub-Riemannian structure insure that the sub-Riemannian distance enjoys some properties which guarantee that some measure contraction property of the form $MCP(0, N)$ is satisfied. Our approach is purely qualitative, we do not compute any curvature type quantity in order to find the best exponents. Our results are concerned with two-step distributions and medium-fat Carnot groups.

We recall that a distribution Δ , or a sub-Riemannian structure (Δ, g) , is *two-step* if

$$[\Delta, \Delta](x) := \left\{ [X, Y](x) \mid X, Y \text{ smooth sections of } \Delta \right\} = T_x M \quad \forall x \in M.$$

A measure on M is called *smooth* if it is locally defined by a positive smooth density times the Lebesgue measure \mathcal{L}^n . Our first result is the following:

Theorem 3. *Every two-step sub-Riemannian structure on a compact manifold equipped with a smooth measure satisfies $MCP(0, N)$ for some $N > 0$.*

In the case of Carnot groups, the homogeneity allows us to extend the above result to left-invariant medium-fat distributions. A distribution Δ , or a sub-Riemannian structure with distribution Δ or a Carnot group whose first layer Δ is equipped with a left-invariant metric, is called *medium-fat* if, for every $x \in M$ and every smooth section X of Δ with $X(x) \neq 0$, there holds

$$T_x M = \Delta(x) + [\Delta, \Delta](x) + [X, [\Delta, \Delta]](x),$$

where

$$[X, [\Delta, \Delta]](x) := \left\{ [X, [Y, Z]](x) \mid Y, Z \text{ smooth sections of } \Delta \right\}.$$

The notion of medium-fat distribution has been introduced by Agrachev and Sarychev in [8]. The class of medium-fat distributions contain the class of two-step distributions.

An important feature of medium-fat distributions is that they do not admit non-trivial Goh paths (see Section 2.3). Of course, in the case of a Carnot group the property of being medium-fat depends only on the properties of its Lie algebra. Our second result is the following:

Theorem 4. *Any medium-fat Carnot group whose first layer is equipped with a left-invariant metric and equipped with Haar measure satisfies MCP(0, N) for some $N > 0$.*

The proofs of Theorem 3 and 4 are based on the fact that squared sub-Riemannian pointed distances $d_{SR}(x, \cdot)^2$ satisfy a certain property of horizontal semiconcavity. This property together with the Lipschitzness of $d_{SR}(x, \cdot)^2$ allows us to give an upper bound for divergence of horizontal gradients of f^x which implies the desired measure contraction property.

We recall that all the notations used throughout the paper are listed in Appendix A. The material required for the proof of the two theorems above is worked out in Section 2. The proofs of Theorems 3 and 4 are respectively given in Sections 3 and 4.

2 Preliminaries

Throughout all this section, (Δ, g) denotes a complete sub-Riemannian structure on M of rank $m \leq n$.

2.1 The minimizing Sard conjecture

The minimizing Sard conjecture is concerned with the size of points that can be reached from a given point by singular minimizing geodesics. Following [24], given $x \in M$, we set

$$\mathcal{S}_{\Delta, \min^g}^x := \left\{ \gamma(1) \mid \gamma \in W_{\Delta}^{1,2}([0, 1], M), \gamma \text{ sing.}, d_{SR}(x, \gamma(1))^2 = \text{energy}_g(\gamma) \right\}.$$

Note that for every $x \in M$, the set $\mathcal{S}_{\Delta, \min^g}^x$ is closed and contains x (because $m < n$). Let us introduce the following definition.

Definition 5. *We say that (Δ, g) satisfies the minimizing Sard conjecture at $x \in M$ if the set $\mathcal{S}_{\Delta, \min^g}^x$ has Lebesgue measure zero in M . We say that it satisfies the minimizing Sard conjecture if this property holds for any $x \in M$.*

It is not known if all complete sub-Riemannian structures satisfy the minimizing Sard conjecture (see [3, 24]). The best general result is due to Agrachev who proved in [2] that all closed sets $\mathcal{S}_{\Delta, \min^g}^x$ have empty interior. As the next result shows,

the minimizing Sard conjecture is related to regularity properties of pointed distance functions. Following Agrachev [2], we call smooth point of the function $y \mapsto d_{SR}(x, y)$ (for a fixed $x \in M$) any $y \in M$ for which there is $p \in T_x^*M$ which is not a critical point of the exponential mapping \exp_x and such that the projection $\gamma_{x,p}$ of the normal extremal $\psi : [0, 1] \rightarrow T^*M$ starting at (x, p) is the unique minimizing geodesic from x to $y = \gamma_{x,p}(1)$. By Agrachev's Theorem, the set \mathcal{O}_x of smooth points is always open and dense in M . The following holds:

Proposition 6. *Let $x \in M$ be fixed, the following properties are equivalent:*

- (i) *the structure (Δ, g) satisfies the minimizing Sard conjecture at $x \in M$,*
- (ii) *the function $y \mapsto d_{SR}(x, y)$ is differentiable almost everywhere in M ,*
- (iii) *the set of smooth points \mathcal{O}_x is an open set with full measure in M .*

Furthermore, the function $y \mapsto d_{SR}(x, y)$ is smooth on \mathcal{O}_x and if M and (Δ, g) are real-analytic or if all singular minimizing geodesics are strictly abnormal, then the set \mathcal{O}_x is geodesically star-shaped at x , that is

$$\gamma(s, y) \in \mathcal{O}_x \quad \forall s \in (0, 1], \forall y \in \mathcal{O}_x, \quad (2.1)$$

where $\gamma_x(\cdot, y) \in W_{\Delta}^{1,2}([0, 1], M)$ is the unique minimizing geodesic from x to y .

Proof of Proposition 6. Let $x \in M$ be fixed. The part (iii) \Rightarrow (ii) is immediate. Let us prove that (ii) \Rightarrow (i). By assumption the set of differentiability D of $f := d_{SR}(x, \cdot)$ has full measure in M . Recall that for every $y \in D$, there is a unique minimizing geodesic from x to y which is given by the projection of the normal extremal $\psi : [0, 1] \rightarrow T^*M$ such that $\psi(1) = (y, d_{SR}(x, y)d_y f)$ (see [22, Lemma 2.15 p. 54]). By Sard's Theorem, the set S of $\exp_x(p)$ with $p \in T_x^*M$ critical has Lebesgue measure zero in M . Therefore, the set $D \setminus S$ has full measure and for every $y \in D \setminus S$ there is a unique minimizing geodesic from x to y and it is not singular, which shows that y does not belong to $\mathcal{S}_{\Delta, \text{min}^g}^x$. Let us now show that (i) \Rightarrow (iii). By definition of $\mathcal{S}_{\Delta, \text{min}^g}^x$, for every $y \notin \mathcal{S}_{\Delta, \text{min}^g}^x$ all minimizing horizontal paths between x and y are not singular. So repeating the proof of [22, Theorem 3.14 p. 98] (see also [12]), we can show that the function $f : y \mapsto d_{SR}(x, y)$ is locally semiconcave and so locally Lipschitz on the open set $U := M \setminus \mathcal{S}_{\Delta, \text{min}^g}^x$. Thus for every compact set $K \subset U$, there is a compact set $\mathcal{P}_K \subset T_x^*M$ such that for every $y \in K$, there is $p \in \mathcal{P}_K$ with $\exp_x(p) = y$ and $H(x, p) = d_{SR}(x, y)^2/2$ (in other words $\gamma_{x,p} : [0, 1] \rightarrow M$ is a minimizing geodesic from x to y). By Sard's Theorem, the set S_K of $\exp_x(p)$ with $p \in \mathcal{P}_K$ critical is a closed set of Lebesgue measure zero. For every positive integer k , set (here the diameter of the convex set $d_y^+ f$ is taken with respect to some geodesic distance on T^*M)

$$\Sigma^k(f) := \left\{ y \in U \mid \text{diam}(d_y^+ f) \geq 1/k \right\}.$$

By local semiconcavity of f in U , each set $\Sigma^k(f)$ is a closed set in U with Lebesgue measure zero (see [13, Proposition 4.1.3 p. 79]). We claim that

$$S'_K := K \cap \overline{\bigcup_{k>0} \Sigma^k(f)} \subset \left(K \cap \bigcup_{k>0} \Sigma^k(f) \right) \cup S_K.$$

As a matter of fact, if $y \in K$ belongs to $\overline{\bigcup_{k>0} \Sigma^k(f)} \setminus \bigcup_{k>0} \Sigma^k(f)$, then $d_y^+ f$ is a singleton and there is a sequence $\{y_l\}_l$ converging to y such that all $d_{y_l}^+ f$ have dimension at least one and tend to $d_y^+ f$. This implies that the covector p such that $\exp_x(p) = y$ and $H(x, p) = d_{SR}(x, y)^2/2$ is critical, which shows that y belongs to S_K . By construction, every point in $K \setminus S'_K$ is a smooth point. We conclude easily.

It remains to prove the second part. The smoothness of $f : y \mapsto d_{SR}(x, y)$ is an easy consequence of the inverse function theorem. As a matter of fact, we can show easily that for every $y \in \mathcal{O}_x$ such that $y = \exp_x(p)$ with $H(x, p) = d_{SR}(x, y)^2/2$ and $p \in T_x^*M$ non-critical, there is a neighborhood U of y in \mathcal{O}_x such that

$$f(z)^2 = 2H(x, \exp_x(z)^{-1}) \quad \forall z \in U,$$

where \exp_x^{-1} denotes a local inverse of the exponential mapping from a neighborhood of p to U . To prove (2.1), we argue by contradiction. If there are $x \in M$, $y \in \mathcal{O}_x$ and $s \in (0, 1)$ such that $z := \gamma(s, y) \in \mathcal{O}_x$ then either there are two distinct minimizing geodesics from x to z or there is only one minimizing geodesic from x to y which is singular. In the first case, we infer the existence of two distinct minimizing geodesics from x to y , which contradicts the smoothness of y . In the second case, we deduce that the minimizing geodesic $\gamma_x(\cdot, y)$ is the projection of a normal extremal which is regular and whose restriction to $[0, s]$ is singular. This cannot happen under the assumption of analyticity of the data or if all singular minimizing geodesics are strictly abnormal. \square

Remark 7. *By Proposition 6, any (complete) sub-Riemannian structure satisfying the minimizing Sard conjecture has negligible cut loci.*

Remark 8. *It can be shown by transversality techniques that sub-Riemannian structures all of whose singular minimizing geodesics are strictly abnormal are generic (see [11], [14, Proposition 2.7]).*

2.2 Two characterizations for MCP(0, N)

The following result was implicit in the previous paper [23] of the second author. The measure contraction property MCP(0, N) is equivalent to some upper bound on the divergence of the horizontal gradient of the squared pointed sub-Riemannian distance. This result holds at least whenever the horizontal gradient is well-defined and the sets \mathcal{O}_x are geodesically star-shaped.

Proposition 9. *Assume that (Δ, g) satisfies the minimizing Sard conjecture and that all its sets \mathcal{O}_x are geodesically star-shaped, and let μ be a smooth measure on M and $N > 0$ be fixed. Then (Δ, g) equipped with μ satisfies MCP(0, N) if and only if*

$$\operatorname{div}_y^\mu \left(\nabla^h f^x \right) \leq N \quad \forall y \in \mathcal{O}_x, \forall x \in M, \quad (2.2)$$

where $f^x : M \rightarrow \mathbb{R}$ is the function defined by $f^x(y) := d_{SR}(x, y)^2/2$.

Proof. Let $x \in M$ be fixed, the vector field $Z := -\nabla^h f^x$ is well-defined and smooth on \mathcal{O}_x . Moreover by assumption, every solution of $\dot{y}(t) = Z(y(t))$ with $y(0) \in \mathcal{O}_x$ remains in \mathcal{O}_x for all $t \geq 0$, we denote by $\{\varphi_t\}_{t \geq 0}$ the flow of Z on \mathcal{O}_x . For every $y \in \mathcal{O}_x$, the function $\theta : t \in [0, +\infty) \mapsto d_{SR}(\varphi_t(y), y)$ satisfies

$$\theta(0) = 0 \quad \text{and} \quad \theta(t) = \operatorname{length}^g(\varphi_{[0,t]}(y)) = \int_0^t |Z(\varphi_s(y))| ds.$$

So that, for all $t \geq 0$,

$$\begin{aligned} \dot{\theta}(t) &= |Z(\varphi_t(y))| = d_{SR}(x, \varphi_t(y)) = d_{SR}(x, y) - d_{SR}(y, \varphi_t(y)) = d_{SR}(x, y) - \theta(t), \end{aligned}$$

which yields

$$\theta(t) = d_{SR}(x, y) (1 - e^{-t}) \quad \forall t \geq 0.$$

Consequently, if $A \subset \mathcal{O}_x$ is a Borel set and $s \in (0, 1]$, then we have

$$A_s = \left\{ \gamma_x(s, y) \mid y \in A \right\} = \varphi_t(A) \quad \text{with} \quad t = -\ln(s).$$

Let us now assume that (2.2) is satisfied. By definition of $\operatorname{div}^\mu Z$, for every $x \in M$ and any measurable set $A \subset \mathcal{O}_x$, we have for every $t \geq 0$ (see for example, see [10, Proposition B.1]),

$$\mu(\varphi_t(A)) = \int_A \exp \left(\int_0^t \operatorname{div}_{\varphi_s(y)}^\mu(Z) ds \right) d\mu(y),$$

which by (2.2) implies with $s = e^{-t}$,

$$\mu(A_s) = \mu(\varphi_t(A)) \geq \int_A \exp(-Nt) d\mu(y) = s^N \mu(A).$$

This shows that (2.2) implies MCP(0, N). Conversely, if (Δ, g) equipped with μ satisfies MCP(0, N) then for every $x \in M$ and every small ball $B_\delta(y) \subset \mathcal{O}_x$ (say a Riemannian ball with respect to the Riemannian extension g), we have

$$\mu(\varphi_t(B_\delta(y))) = \int_{B_\delta(y)} \exp \left(\int_0^t \operatorname{div}_{\varphi_s(y)}^\mu(Z) ds \right) d\mu(y) \geq e^{-Nt} \mu(B_\delta(y)) \quad \forall t \geq 0.$$

For every $t \geq 0$, letting δ go to 0 yields

$$\exp\left(\int_0^t \operatorname{div}_{\varphi_s(y)}^\mu(Z) ds\right) \geq e^{-Nt}.$$

We infer (2.2) by dividing by t and letting t go to 0. \square

In the case of Carnot groups, the invariance of the divergence of $\nabla^h f^x$ by dilation allows us to characterize $\text{MCP}(0, N)$ in term of a control on the divergence over a compact set not containing the origin.

Proposition 10. *Let \mathbb{G} be a Carnot group whose first layer is equipped with a left-invariant metric satisfying the minimizing Sard conjecture and $N > 0$ fixed. Then the metric space (\mathbb{G}, d_{SR}) with Haar measure μ satisfies $\text{MCP}(0, N)$ if and only if*

$$\operatorname{div}_y^\mu\left(\nabla^h f^0\right) \leq N \quad \forall y \in \mathcal{O}_0 \cap S_{SR}(0, 1), \quad (2.3)$$

where $f^0 : M \rightarrow \mathbb{R}$ is the function defined by $f^0(y) := d_{SR}(0, y)^2/2$.

Proof. Since Carnot groups are indeed analytic, by the second part of Proposition 6 and Proposition 9, it is sufficient to show that (2.3) is equivalent to

$$\operatorname{div}_y^\mu\left(\nabla^h f^0\right) \leq N \quad \forall y \in \mathcal{O}_0. \quad (2.4)$$

Recall that by taking a set of exponential coordinates (x_1, \dots, x_n) , we can identify \mathbb{G} with its Lie algebra $\mathfrak{g} \simeq \mathbb{R}^n$ and indeed consider that we work with the Lebesgue measure in \mathbb{R}^n and that the sub-Riemannian structure is globally parametrized by an orthonormal family of analytic vector fields X^1, \dots, X^n in \mathbb{R}^n satisfying

$$X^i(\delta_\lambda(x)) = \lambda^{-1} \delta_\lambda(X^i(x)) \quad \forall x \in \mathbb{R}^n, \forall i = 1, \dots, n, \quad (2.5)$$

where $\{\delta_\lambda\}_{\lambda>0}$ is a family of dilations defined as (d_1, \dots, d_n) are positive integers)

$$\delta_\lambda(x_1, \dots, x_n) = \left(\lambda^{d_1} x_1, \lambda^{d_2} x_2, \dots, \lambda^{d_n} x_n\right) \quad \forall x \in \mathbb{R}^n.$$

By the homogeneity property, we have $d_{SR}(0, \delta_\lambda(x)) = \lambda d_{SR}(0, x)$ for all $x \in \mathbb{R}^n$ and $\lambda > 0$. Then we have

$$f^0(\delta_\lambda(x)) = \lambda^2 f^0(x) \quad \text{and} \quad d_{\delta_\lambda(x)} f^0 \circ \delta_\lambda = \lambda^2 d_x f^0 \quad \forall x \in \mathbb{R}^n, \forall \lambda > 0. \quad (2.6)$$

Recall that the horizontal gradient $\nabla^h f^0$ is given by

$$\nabla_x^h f^0 = \sum_{i=1}^m (X^i \cdot f^0)(x) X^i(x) \quad \forall x \in \mathbb{R}^n.$$

Therefore, by (2.5)-(2.6), we infer that for every $x \in \mathbb{R}^n$ and $\lambda > 0$,

$$\begin{aligned}\nabla_{\delta_\lambda(x)}^h f^0 &= \sum_{i=1}^m d_{\delta_\lambda(x)} f^0 (X^i(\delta_\lambda(x))) X^i(\delta_\lambda(x)) \\ &= \lambda^{-2} \sum_{i=1}^m d_{\delta_\lambda(x)} f^0 (\delta_\lambda(X^i(x))) \delta_\lambda(X^i(x)) \\ &= \sum_{i=1}^m d_x f^0 (X^i(x)) \delta_\lambda(X^i(x)) = \delta_\lambda(\nabla_x^h f^0).\end{aligned}$$

We deduce that

$$\operatorname{div}_y^\mu(\nabla_{\delta_\lambda(x)}^h f^0) = \operatorname{div}_y^\mu(\nabla_x^h f^0) \quad \forall x \in \mathbb{R}^n, \forall \lambda > 0,$$

which shows that (2.3) and (2.4) are equivalent and concludes the proof. \square

2.3 Nearly horizontally semiconcave functions

Recall that without loss of generality, we can assume that the metric g over Δ is the restriction of a global Riemannian metric on M . This metric allows us to define the C^2 -norms of functions from \mathbb{R}^m to M . In the following statement, (e_1, \dots, e_m) stands for the canonical basis in \mathbb{R}^m .

Definition 11. *Let $C > 0$, U an open subset of M and $K \subset M$, a function $f : U \rightarrow \mathbb{R}$ in an open set $U \subset M$ is said to be C -nearly horizontally semiconcave with respect to (Δ, g) in K if for every $y \in K$, there are an open neighborhood V^y of 0 in \mathbb{R}^m , a function $\varphi^y : V^y \subset \mathbb{R}^m \rightarrow U$ of class C^2 and a function $\psi^y : V^y \subset \mathbb{R}^m \rightarrow \mathbb{R}$ of class C^2 such that*

$$\varphi^y(0) = y, \quad \psi^y(0) = f(y), \quad f(\varphi^y(v)) \leq \psi^y(v) \quad \forall v \in V^y, \quad (2.7)$$

$$\left\{ d_0 \varphi^y(e_1), \dots, d_0 \varphi^y(e_m) \right\} \text{ is an orthonormal family of vectors in } \Delta(y), \quad (2.8)$$

and

$$\|\varphi^y\|_{C^2}, \|\psi^y\|_{C^2} \leq C, \quad (2.9)$$

where $\|\varphi^y\|_{C^2}, \|\psi^y\|_{C^2}$ denote the C^2 -norms of φ^y and ψ^y .

If m were equal to n that is if we were in the Riemannian case, the above definition would coincide with the classical definitions of semiconcave functions (see [13, 22]). Here, in the case $m < n$, the definition tells that at each point, there is a support function from above of class C^2 which bounds the function along a C^2 submanifold which is tangent to the distribution. This type of mild horizontal ssemiconcavity will allow us, at least in certain cases, to bound the divergence of the horizontal gradient of squared pointed sub-Riemannian distance functions.

Before stating the main result of this section, we recall that a minimizing geodesic or more generally an horizontal path $\gamma : [0, 1] \rightarrow M$ is called a Goh path if it admits an abnormal lift $\psi : [0, 1] \rightarrow T^*M$ which annihilates $[\Delta, \Delta]$, that is to say an abnormal lift $\psi = (x, p)$ in local coordinates such that for every local parametrization of Δ by smooth vector fields X^1, \dots, X^m in a neighborhood of $\gamma([0, 1])$, we have

$$p(t) \cdot [X^i, X^j](\gamma(t)) = 0 \quad \forall t \in [0, 1], \forall i, j = 1, \dots, m. \quad (2.10)$$

As we shall see in the proof of Theorem 3, the absence of minimizing Goh paths implies the Lipschitz regularity of the distance function. The proof of this result, due to Agrachev and Lee [5], follows by a study at second order of possibly abnormal minimizing geodesics. Second order considerations, in [5] as in the proof of Proposition 12 below, owes a lot to techniques and material introduced by Agrachev and his co-authors (see [7, Chapter 20] and [22]).

Proposition 12. *Assume that any minimizing geodesic for (Δ, g) joining two points in M is not a Goh path, then for every compact set $K \subset M$ there is $C > 0$ such that for every $x \in K$, the function $y \mapsto d_{SR}(x, y)^2$ is C -nearly horizontally semiconcave in K . In particular, if M is compact and Δ is two-step then there is $C > 0$ such that all functions $d_{SR}(x, \cdot)^2$ are C -nearly horizontally semiconcave in M .*

Proof of Proposition 12. Let K be a compact subset of M and $x \in K$ fixed, let us first show how to construct functions φ^y, ψ^y of class C^2 satisfying (2.7)-(2.8) for some $y \in K$. Pick a minimizing geodesic $\bar{\gamma} : [0, 1] \rightarrow M$ from x to $y = \bar{\gamma}(1)$. There is an open neighborhood $U_{\bar{\gamma}}$ of $\bar{\gamma}([0, 1])$ and a family $\mathcal{F}_{\bar{\gamma}}$ of m smooth vector fields $X_{\bar{\gamma}}^1, \dots, X_{\bar{\gamma}}^m$ in M such that for every $z \in U_{\bar{\gamma}}$ the family $\{X_{\bar{\gamma}}^1(z), \dots, X_{\bar{\gamma}}^m(z)\}$ is orthonormal with respect to g and parametrize Δ (that is $\text{Span}\{X_{\bar{\gamma}}^1(z), \dots, X_{\bar{\gamma}}^m(z)\} = \Delta(z)$) and for every $z \in M \setminus U_{\bar{\gamma}}$, $X_{\bar{\gamma}}^1(z), \dots, X_{\bar{\gamma}}^m(z)$ belongs to $\Delta(z)$. Consider the End-Point mapping from x in time 1 associated with the family $\mathcal{F}_{\bar{\gamma}} = \{X_{\bar{\gamma}}^1, \dots, X_{\bar{\gamma}}^m\}$, it is defined by

$$E_{\mathcal{F}_{\bar{\gamma}}}^{x,1} : L^2([0, 1]; \mathbb{R}^m) \mapsto \gamma_u^{\mathcal{F}_{\bar{\gamma}}}(1),$$

where $\gamma_u^{\mathcal{F}_{\bar{\gamma}}}(1) : [0, 1] \rightarrow M$ is the solution to the Cauchy problem

$$\dot{\gamma}(t) = \sum_{i=1}^m u_i(t) X_{\bar{\gamma}}^i(\gamma(t)) \quad \text{for a.e. } t \in [0, 1], \quad \gamma(0) = x. \quad (2.11)$$

Note that taking the vector fields $X_{\bar{\gamma}}^1, \dots, X_{\bar{\gamma}}^m$ equal to zero outside an neighborhood of $\bar{U}_{\bar{\gamma}}$, we may assume without loss of generality that $E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}$ is well-defined on $L^2([0, 1]; \mathbb{R}^m)$. Recall that the function $E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}$ is smooth and satisfies (see [22, Proposition 1.10 p. 19])

$$\Delta(y) \subset \text{Im} \left(d_{\bar{u}} E_{\mathcal{F}_{\bar{\gamma}}}^{x,1} \right),$$

where $u_{\bar{\gamma}}$ is the unique control $u \in L^2([0, 1], \mathbb{R}^m)$ such that $\gamma_u^{\mathcal{F}_{\bar{\gamma}}} = \bar{\gamma}$. Therefore, there are $v_{\bar{\gamma}}^1, \dots, v_{\bar{\gamma}}^m \in L^2([0, 1], \mathbb{R}^m)$ such that

$$d_{u_{\bar{\gamma}}} E_{\mathcal{F}_{\bar{\gamma}}}^{x,1} (v_{\bar{\gamma}}^i) = X_{\bar{\gamma}}^i(y) \quad \forall i = 1, \dots, m. \quad (2.12)$$

Define $\varphi_{\bar{\gamma}} : \mathbb{R}^m \rightarrow M$ by

$$\varphi_{\bar{\gamma}}(v) := E_{\mathcal{F}_{\bar{\gamma}}}^{x,1} \left(u_{\bar{\gamma}} + \sum_{i=1}^m v_i v_{\bar{\gamma}}^i \right) \quad \forall v = (v_1, \dots, v_m) \in \mathbb{R}^m.$$

By construction, $\varphi_{\bar{\gamma}}$ is smooth and satisfies

$$\varphi_{\bar{\gamma}}(0) = E_{\mathcal{F}_{\bar{\gamma}}}^{x,1} (u_{\bar{\gamma}}) = \bar{\gamma}(1) = y,$$

and

$$d_0 \varphi_{\bar{\gamma}}(e_i) = d_{u_{\bar{\gamma}}} E_{\mathcal{F}_{\bar{\gamma}}}^{x,1} (v_{\bar{\gamma}}^i) = X_{\bar{\gamma}}^i(y) \quad \forall i = 1, \dots, m.$$

Moreover, for every $v \in \mathbb{R}^m$ such that the solution to (2.11) associated with the control $u_{\bar{\gamma}} + \sum_{i=1}^m v_i v_{\bar{\gamma}}^i$ remains in $U_{\bar{\gamma}}$, we have

$$d_{SR}(x, \varphi_{\bar{\gamma}}(v))^2 \leq \left\| u_{\bar{\gamma}} + \sum_{i=1}^m v_i v_{\bar{\gamma}}^i \right\|_{L^2}^2 =: \psi_{\bar{\gamma}}(v).$$

By construction, $\varphi^y := \varphi_{\bar{\gamma}}$ and $\psi^y := \psi_{\bar{\gamma}}$ are smooth, defined in a neighborhood of $0 \in \mathbb{R}^m$ and satisfy (2.7)-(2.8). Furthermore, we observe that there is $C_{\bar{\gamma}} > 0$ and a neighborhood $\mathcal{U}_{\bar{\gamma}}$ of $u_{\bar{\gamma}}$ in $L^2([0, 1], \mathbb{R}^m)$ such that for every $u \in \mathcal{U}_{\bar{\gamma}}$ there are $v_u^1, \dots, v_u^m \in L^2([0, 1], \mathbb{R}^m)$ satisfying (2.12) at u , that is

$$d_u E_{\mathcal{F}_{\bar{\gamma}}}^{x,1} (v_u^i) = X_{\bar{\gamma}}^i \left(E_{\mathcal{F}_{\bar{\gamma}}}^{x,1} (u) \right) \quad \forall i = 1, \dots, m,$$

and such that the smooth functions $\varphi_u : \mathbb{R}^m \rightarrow M$, $\psi_u : \mathbb{R}^m \rightarrow \mathbb{R}$ defined by

$$\varphi_u(v) := E_{\mathcal{F}_{\bar{\gamma}}}^{x,1} \left(u + \sum_{i=1}^m v_i v_u^i \right) \text{ and } \psi_u(v) := \left\| u + \sum_{i=1}^m v_i v_u^i \right\|_{L^2}^2 \quad \forall v \in \mathbb{R}^m$$

have C^2 -norms less than $C_{\bar{\gamma}}$ on a neighborhood of $0 \in \mathbb{R}^m$. This claim follows readily from the following lemma whose proof is based on the fact that there are no minimizing Goh paths.

Lemma 13. *There are a neighborhood $\mathcal{U}_{\bar{\gamma}}$ of $u_{\bar{\gamma}}$ in $L^2([0, 1], \mathbb{R}^m)$ and $C_{\bar{\gamma}} > 0$ such that for every $u \in \mathcal{U}_{\bar{\gamma}}$, there are $v_u^1, \dots, v_u^m \in L^2([0, 1], \mathbb{R}^m)$ such that*

$$d_u E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}(v_u^i) = X_{\bar{\gamma}}^i \left(E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}(u) \right) \quad \forall i = 1, \dots, m$$

and

$$\|v_u^i\|_{L^2} \leq c_{\bar{\gamma}} \quad \forall i = 1, \dots, m.$$

Proof of Lemma 13. Firstly, taking a chart on a neighborhood of $y = \bar{\gamma}(1)$ we may assume that the restriction of $E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}$ to a neighborhood of $u_{\bar{\gamma}}$ is valued in \mathbb{R}^n . Define the function $F : L^2([0, 1], \mathbb{R}^m) \times L^2([0, 1], \mathbb{R}^m) \rightarrow \mathbb{R}^n$ by (this function is indeed well-defined only in that neighbourhood of $u_{\bar{\gamma}}$)

$$F(u, v) := d_u E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}(v) \quad \forall u, v \in L^2([0, 1], \mathbb{R}^m).$$

Let $i \in \{1, \dots, m\}$ be fixed and $V^i \in \mathbb{R}^n$ defined by

$$V^i := X_{\bar{\gamma}}^i \left(E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}(u_{\bar{\gamma}}) \right),$$

we claim that there is $v^i \in L^2([0, 1], \mathbb{R}^m)$ such that $F(u_{\bar{\gamma}}, v^i) = V^i$ and F is a submersion at $(u_{\bar{\gamma}}, v^i)$. Two cases may appear.

First case: $u_{\bar{\gamma}}$ is not singular.

In this case, by definition the linear mapping $v \mapsto d_{u_{\bar{\gamma}}} E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}(v)$ is surjective, so the result holds, because the image of $d_{(u_{\bar{\gamma}}, v)} F$ contains the image of $d_{u_{\bar{\gamma}}} E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}$.

Second case: $u_{\bar{\gamma}}$ is singular.

Let $\bar{u} := u_{\bar{\gamma}}$ and $E := E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}$, by assumption $\bar{\gamma}$ is not a Goh path, therefore by [22, Theorem 2.20 p. 61], for every $\bar{p} \in (\mathbb{R}^n)^* \setminus \{0\}$ such that $\bar{p} \cdot d_{\bar{u}} E = 0$, there holds (we refer the reader to [7, 22] for the definition of the negative index of a quadratic form)

$$\text{ind}_- \left(\bar{p} \cdot (d_{\bar{u}}^2 E)_{|\text{Ker}(d_{\bar{u}} E)} \right) = +\infty.$$

Consequently, by compactness of the set of $\bar{p} \in (\mathbb{R}^n)^*$ with $|\bar{p}| = 1$ and $\bar{p} \cdot d_{\bar{u}} E = 0$, there is space $X \subset L^2([0, 1], \mathbb{R}^m)$ of finite dimension such that the restriction \tilde{E} of E to the affine space $\bar{u} + X$ satisfies

$$\text{Im} \left(d_{\bar{u}} \tilde{E} \right) = \text{Im} (d_{\bar{u}} E) \tag{2.13}$$

and

$$\text{ind}_- \left(\bar{p} \cdot (d_{\bar{u}}^2 \tilde{E})_{|\text{Ker}(d_{\bar{u}} \tilde{E})} \right) \geq r \quad \forall \bar{p} \in \left(\text{Im} \left(d_{\bar{u}} \tilde{E} \right) \right)^\perp \setminus \{0\} \tag{2.14}$$

with

$$r := n - \dim(\operatorname{Im}(d_{\bar{u}}E)).$$

Let $\mathcal{K} := \operatorname{Im}(d_0\tilde{E})^\perp$ of dimension r , $\operatorname{Proj}_{\mathcal{K}} : \mathbb{R}^n \rightarrow \mathcal{K}$ be the orthogonal projection onto \mathcal{K} , and $Q : \operatorname{Ker}(d_{\bar{u}}\tilde{E}) \rightarrow \mathcal{K}$ the quadratic mapping defined by

$$Q(v) := \operatorname{Proj}_{\mathcal{K}} \left[\left(d_{\bar{u}}^2 \tilde{E} \right) \cdot (v, v) \right] \quad \forall v \in \operatorname{Ker}(d_{\bar{u}}\tilde{E}).$$

By [22, Lemma B.6 p. 130], there is $\bar{v} \in \operatorname{Ker}(d_{\bar{u}}\tilde{E})$ such that $d_{\bar{v}}Q$ is surjective which means that the linear mapping

$$w \in \operatorname{Ker}(d_{\bar{u}}\tilde{E}) \mapsto \operatorname{Proj}_{\mathcal{K}} \left[\left(d_{\bar{u}}^2 \tilde{E} \right) \cdot (\bar{v}, w) \right] \in \mathcal{K}$$

is surjective. Thus means that there are $w^1, \dots, w^r \in L^2([0, 1], \mathbb{R}^m)$ such that the linear mapping

$$\lambda \in \mathbb{R}^r \mapsto \operatorname{Proj}_{\mathcal{K}} \left[d_{\bar{u}}^2 \tilde{E} \left(\bar{v}, \sum_{j=1}^r \lambda_j w^j \right) \right]$$

is bijective. By [22, Proposition 1.10 p. 19] and (2.13), we know that there is $\bar{v}^i \in L^2([0, 1], \mathbb{R}^m)$ such that $d_{\bar{u}}\tilde{E}(\bar{v}^i) = V^i$ (without loss of generality we may assume that \bar{v}^i belongs to X). Moreover, by bilinearity we have for every $\epsilon > 0$,

$$\begin{aligned} \operatorname{Proj}_{\mathcal{K}} \left[d_{\bar{u}}^2 \tilde{E} \left(\bar{v}^i + \frac{1}{\epsilon} \bar{v}, \sum_{j=1}^m \lambda_j w^j \right) \right] = \\ \operatorname{Proj}_{\mathcal{K}} \left[d_{\bar{u}}^2 \tilde{E} \left(\bar{v}^i, \sum_{j=1}^m \lambda_j w^j \right) \right] + \frac{1}{\epsilon} \operatorname{Proj}_{\mathcal{K}} \left[d_{\bar{u}}^2 \tilde{E} \left(\bar{v}, \sum_{j=1}^m \lambda_j w^j \right) \right]. \end{aligned}$$

So, we infer that for $\epsilon > 0$ large enough the linear mapping

$$\lambda \in \mathbb{R}^r \mapsto \operatorname{Proj}_{\mathcal{K}} \left[d_{\bar{u}}^2 \tilde{E} \left(\bar{v}^i + \frac{1}{\epsilon} \bar{v}, \sum_{j=1}^m \lambda_j w^j \right) \right]$$

is bijective. We conclude easily because $v^i := \bar{v}^i + \bar{v}/\epsilon \in L^2([0, 1], \mathbb{R}^m)$ satisfies

$$F(\bar{u}, v^i) = V^i$$

and the differential of F at (\bar{u}, v^i) contains the sum of the images of the linear mappings $d_{\bar{u}}E$ and $d_{\bar{u}}^2E(v^i, \cdot)$ which by the above construction contains $\mathcal{K} = \operatorname{Im}(d_0E)^\perp$.

Returning to the proof of Lemma 13, since $F(u_{\bar{\gamma}}, v^i) = V^i$ and F is a submersion at $(u_{\bar{\gamma}}, v^i)$, by the Inverse Function Theorem, for every $u \in L^2([0, 1], \mathbb{R}^m)$ sufficiently close to $u_{\bar{\gamma}}$ there is v_u^i such that $d_u E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}(v_u^i) = X_{\bar{\gamma}}^i \left(E_{\mathcal{F}_{\bar{\gamma}}}^{x,1}(u) \right)$ with a control of $\|v_u^i\|_{L^2}$. We conclude easily. \square

It remains to explain why (2.9) holds for all $y \in K$ for some uniform constant C . It is a consequence of the following compactness result (see[1] or [15, Proposition 5.8]).

Lemma 14. *The set*

$$\Gamma := \left\{ \gamma \in W_{\Delta}^{1,2}([0, 1], M) \mid \gamma(0), \gamma(1) \in K \text{ and } d_{SR}(\gamma(0), \gamma(1))^2 = \text{energy}_g(\gamma) \right\}$$

is a compact set in the $W^{1,2}$ topology.

By compactness (Lemma 14), there is a finite family of minimizing geodesics $\{\gamma_l\}_{l \in L}$ such that

$$\Gamma \subset \bigcup_{l \in L} \mathcal{U}_{\gamma_l}.$$

We let the reader to show that (2.9) holds for all $y \in K$ with C equal to the maximum of all C_{γ_l} .

The second part of Proposition 12 follows by compactness and the fact that a two-step distribution does not admit minimizing Goh paths (even constant curves cannot be Goh paths). \square

We state now the result that will be used to prove Theorem 4.

Proposition 15. *Let $x \in M$ be fixed, assume that any minimizing geodesic for (Δ, g) joining x to any $y \in M \setminus \{x\}$ is not a Goh path, then for every compact set $K \subset M \setminus \{x\}$ there is $C > 0$ such that the function $y \mapsto d_{SR}(x, y)^2$ is C -nearly horizontally semiconcave in K . In particular, if Δ is a medium-fat distribution, then for every compact set $K \subset M \setminus \{x\}$ there is $C > 0$ such that the function $d_{SR}(x, \cdot)^2$ is C -nearly horizontally semiconcave in $S_{SR}(0, 1)$.*

Proof. The first part follows exactly by the same arguments as in the proof of Proposition 12. The second part is a consequence of the following result.

Lemma 16. *If Δ is medium-fat, it does not admit non-trivial Goh paths.*

Proof of Lemma 16. Argue by contradiction and assume that $\gamma : [0, 1] \rightarrow M$ is a non-trivial horizontal path which admits an abnormal lift $\psi = (\gamma, p) : [0, 1] \rightarrow T^*M$ satisfying the Goh condition, then we have

$$p(t) \cdot [X^i, X^j](\gamma(t)) = 0 \quad \forall i, j = 1, \dots, m, \quad (2.15)$$

for every t in a small interval $I \subset [0, 1]$ such that $\gamma(t)$ is in a local chart of M and Δ is parametrized by a family $\mathcal{F} = \{X^1, \dots, X^m\}$ of smooth vector fields. Then if we denote by u the control which is associated to γ through \mathcal{F} , derivating the previous equality yields for any $i, j = 1, \dots, m$,

$$p(t) \cdot \left[\sum_{k=1}^m u_k(t) X^k, [X^i, X^j] \right] (\gamma(t)) = 0 \quad \text{for a.e. } t \in I. \quad (2.16)$$

Since $\psi = (\gamma, p)$ is an abnormal lift, we also have $p \cdot X^i = 0$ along γ . Moreover since γ is non-trivial, we may assume that

$$\dot{\gamma}(t) = \sum_{k=1}^m u_k(t) X^k \neq 0.$$

By (??), (2.15)-(2.16) we get a contradiction. □

The proof of Proposition 15 is complete. □

3 Proof of Theorem 3

By the second part of Proposition 12, there is $C > 0$ such that for every $x \in M$ the function $f^x : y \mapsto d_{SR}(x, y)^2/2$ is C -nearly horizontally semiconcave in M .

Lemma 17. *There is $B > 0$ such that for every $x \in M$ the following property holds: for every $y \in \mathcal{O}_x$, there is a neighborhood $U^y \subset \mathcal{O}_x$ of y along with an orthonormal family of smooth vector fields X^1, \dots, X^m which parametrize Δ in U^y such that*

$$\|X^i\|_{C^1} \leq B \quad \forall i = 1, \dots, m, \quad (3.1)$$

and

$$[X^i \cdot (X^i \cdot f^x)](z) \leq B |\nabla_z f^x| + B \quad \forall z \in U^y, \forall i = 1, \dots, m. \quad (3.2)$$

Proof of Lemma 17. First of all, we notice that there is $A > 0$ such that if v^1, \dots, v^m is an orthonormal family of tangent vectors in $\Delta(z)$ for some $z \in M$ then there is an orthonormal family of smooth vector fields X^1, \dots, X^m which generates the distribution Δ in a neighborhood of z and such that $\|X^i\|_{C^1}$ is bounded by A for all $i = 1, \dots, m$. Let $x \in M$ be fixed, then by C -nearly horizontal semiconcavity of f^x , for every $y \in M$, there are an open neighborhood V^y of 0 in \mathbb{R}^m , a function $\varphi^y : V^y \subset \mathbb{R}^m \rightarrow U$ of class C^2 and a function $\psi^y : V^x \subset \mathbb{R}^m \rightarrow \mathbb{R}$ of class C^2 such that (2.7) (with $f = f^x$),

(2.8) and (2.9) are satisfied. Fix $y \in \mathcal{O}_x$ and define the function $F^y : U^y \rightarrow \mathbb{R}$ by $F^y := f^x \circ \varphi^y - \psi^y$, it is of class C^2 and satisfies

$$d_0 F^y = 0 \quad \text{and} \quad \text{Hess}_0 F^y \leq 0.$$

Taking a chart near y we can assume that we work in \mathbb{R}^n . Let $\varphi^y = (\varphi_1^y, \dots, \varphi_n^y)$ and (x_1, \dots, x_n) and (v_1, \dots, v_m) the coordinates respectively in \mathbb{R}^n and \mathbb{R}^m . Then, we have

$$\frac{\partial F^y}{\partial v_i}(0) = \left(\sum_{k=1}^n \frac{\partial f^x}{\partial x_k}(y) \frac{\partial \varphi_k^y}{\partial v_i}(0) \right) - \frac{\partial \psi^y}{\partial v_i}(0) = 0 \quad \forall i = 1, \dots, m$$

and for every $i = 1, \dots, m$,

$$\begin{aligned} \frac{\partial^2 F^y}{\partial v_i^2}(0) &= \left(\sum_{k,l=1}^n \frac{\partial^2 f^x}{\partial x_l \partial x_k}(y) \frac{\partial \varphi_k^y}{\partial v_i}(0) \frac{\partial \varphi_l^y}{\partial v_i}(0) \right) \\ &\quad + \left(\sum_{k=1}^n \frac{\partial f^x}{\partial x_k}(y) \frac{\partial^2 \varphi_k^y}{\partial v_i^2}(0) \right) - \frac{\partial^2 \psi^y}{\partial v_i^2}(0) \leq 0, \end{aligned}$$

which yields

$$\begin{aligned} \sum_{k,l=1}^n \frac{\partial^2 f^x}{\partial x_l \partial x_k}(y) \frac{\partial \varphi_k^y}{\partial v_i}(0) \frac{\partial \varphi_l^y}{\partial v_i}(0) &\leq \frac{\partial^2 \psi^y}{\partial v_i^2}(0) - \sum_{k=1}^n \frac{\partial f^x}{\partial x_k}(y) \frac{\partial^2 \varphi_k^y}{\partial v_i^2}(0) \\ &\leq C + C |\nabla_y f^x|. \end{aligned} \quad (3.3)$$

By (2.8) and the observation made at the very beginning of this proof, there is an orthonormal family of smooth vector fields X^1, \dots, X^m which generates the distribution Δ in a neighborhood of z and such that

$$\|X^i\|_{C^1} \leq A \quad \text{and} \quad d_0 \varphi^y(e_i) = \frac{\partial \varphi^y}{\partial v_i}(0) = X^i(y) \quad \forall i = 1, \dots, m. \quad (3.4)$$

Setting $X^i = \sum_{k=1}^n a_k^i \partial_k$, we check easily that

$$X^i \cdot f^x = \sum_{k=1}^n a_k^i \frac{\partial f^x}{\partial x_k}$$

and

$$X^i \cdot (X^i \cdot f^x) = \sum_{k=1}^n \left(\sum_{l=1}^n a_l^i \frac{\partial a_k^i}{\partial x_l} \right) \frac{\partial f^x}{\partial x_k} + \sum_{k=1}^n a_k^i \left(\sum_{l=1}^n a_l^i \frac{\partial^2 f^x}{\partial x_l \partial x_k} \right).$$

The last expression at y yields, thanks to (3.3) and (3.4) (which implies $a_k^i(y) = \frac{\partial \varphi_k^y}{\partial v_i}(0)$ for all $i = 1, \dots, m$ and $k = 1, \dots, m$)

$$[X^i \cdot (X^i \cdot f^x)](y) \leq A^2 |\nabla_y f^x| + C + C |\nabla_y f^x| \quad \forall i = 1, \dots, m.$$

We conclude easily by smoothness of f^x in \mathcal{O}_x with U^y sufficiently small and $B > 0$ sufficiently large. \square

The following result, due to Agrachev and Lee [5] (see also [22]), is a consequence of the fact that Δ is two-step (and the compactness of M). We refer the reader to [5, 22] for the proof.

Lemma 18. *There is $L > 0$ such that $|\nabla_y f^x| \leq L$ for all $x, y \in M$.*

Let μ be a smooth measure on M , in order to prove Theorem 3, we need to bound from above the divergence of f^x over \mathcal{O}_x for all x in M . The following holds:

Lemma 19. *There is $N > 0$ such that the following property holds:*

$$\operatorname{div}_y^\mu (\nabla^h f^x) \leq N \quad \forall y \in \mathcal{O}_x, \forall x \in M. \quad (3.5)$$

Proof of Lemma 19. Let $x \in M$ and $y \in \mathcal{O}_x$ be fixed, by Lemma 17 there is a neighborhood $U^y \subset \mathcal{O}_x$ of y along with an orthonormal family of smooth vector fields X^1, \dots, X^m which parametrize Δ in U^y such that (3.2) holds. The horizontal gradient of f^x in U^y is given by

$$\nabla_y^h f^x = \sum_{i=1}^m (X^i \cdot f^x)(y) X^i(y).$$

So, we have

$$\operatorname{div}_y^\mu (\nabla^h f^x) = \sum_{i=1}^m (X^i \cdot f^x)(y) \operatorname{div}_y^\mu (X^i) + \sum_{i=1}^m [X^i \cdot (X^i \cdot f)](y).$$

The second term above (in the right-hand side) is bounded thanks to (3.2) and Lemma 18 and the first term is bounded by (3.1) and Lemma 18 (the quantities $(X^i \cdot f^x)(y)$ are indeed bounded by the fact that $d_{SR}(x, \cdot)$ is solution to the horizontal eikonal equation, see [15]). The proof of Lemma 19 is complete. \square

Let us now conclude the proof of Theorem 3. If we can show that (Δ, g) satisfies the minimizing Sard conjecture and that all the sets \mathcal{O}_x are geodesically star-shaped then we are done thanks to Lemma 19 together with Proposition 9. As we said above, since the distribution is two-step the functions f^x are Lipschitz on M , hence the functions $y \mapsto d_{SR}(x, y)$ are locally Lipschitz on $M \setminus \{x\}$ and by Proposition 6, the minimizing

Sard conjecture is satisfied. It remains to show that all the sets \mathcal{O}_x are geodesically star-shaped. In the case where all singular minimizing geodesics are strictly abnormal, then we get the result by the second part of Proposition 6. If not, we proceed by approximation of (Δ, g) by a sequence of sub-Riemannian structures all of whose singular minimizing geodesics are strictly abnormal (see Remark 8).

We can check that, without loss of generality, we may assume that all bounds obtained in Lemma 17, Lemma 18 and Lemma 19 remain valid for small perturbations of (Δ, g) in smooth topology. So by the above proof and Remark 8, there is $N > 0$ and a sequence of two-step sub-Riemannian structures $\{(\Delta_l, g_l)\}_l$ converging to (Δ, g) which satisfy MCP(0, N) for the measure μ . The property MCP(0, N) (for μ) passes to the limit. As a matter of fact, let $x \in M$ be fixed and $A \subset \mathcal{O}_x$ a measurable set, removing a set of measure zero in A we may assume that A is contained in all sets \mathcal{O}_x^l (which stand for the smooth sets with respect to (Δ_l, g_l)). Then, denoting for every l by A_s^l the s -interpolation of A from x with respect to (Δ_l, g_l) , we have

$$\mu(A_s^l) \geq s^N \mu(A) \quad \forall s \in [0, 1], \forall l.$$

For each $y \in A$, there is a unique minimizing geodesic γ from x to y . Moreover, since $A \subset \cap_l \mathcal{O}_x^l$ there is as well a unique minimizing geodesic γ_l from x to y with respect to each (Δ_l, g_l) . Hence, the sequence of curves $\{\gamma_l\}_l$ converges to γ in C^0 -topology. This shows that for every $s \in [0, 1]$, the characteristic functions of A_l converge pointwise to the characteristic function of A . We conclude by the Lebesgue's Dominated Convergence Theorem.

4 Proof of Theorem 4

By Proposition 10, it is sufficient to show that (2.3) holds. By Proposition 15, we know that the function $f^0 : y \rightarrow d_{SR}(0, y)^2/2$ is C -nearly horizontally semiconcave in K . Furthermore, the function f^0 is locally Lipschitz in $\mathbb{G} \setminus \{0\}$ (see [22, Theorem 3.15 p. 100]). So we can repeat the arguments used in the proof of Theorem 3 for $y \in \mathcal{O}_0 \cap S_{SR}(0, 1)$.

A Notations

We list below the notations used throughout this paper, we refer the reader to the monographs [4, 13, 20, 22] for further details:

- M is a smooth manifold of dimension $n \geq 3$.
- Δ is a smooth totally nonholonomic distribution of rank $m < n$.

- g is a smooth metric over Δ . Sometimes, we see g as the restriction of a global Riemannian metric g on M . We use the notation $\langle \cdot, \cdot \rangle$ instead of $g_x(\cdot, \cdot)$ and we denote the norm associated with g by $|\cdot|$ (instead of $|\cdot|_x = g_x(\cdot, \cdot)^{1/2}$). $B_r(x)$ stands for the open geodesic ball of radius $r > 0$ centered at x .
- We call horizontal path any $\gamma : [0, 1] \rightarrow M$ in $W^{1,2}$ which is almost everywhere tangent to Δ . We denote by $W_{\Delta}^{1,2}([0, 1], M)$ the set of horizontal paths $\gamma : [0, 1] \rightarrow M$ endowed with the $W^{1,2}$ -topology.
- For every $\gamma \in W_{\Delta}^{1,2}([0, 1], M)$, we define the length of γ (w.r.t. g) by $\text{length}^g(\gamma) = \int_0^1 |\dot{\gamma}(t)| dt$ and its energy (w.r.t. g) by $\text{energy}^g(\gamma) = \int_0^1 |\dot{\gamma}(t)|^2 dt$.
- For any $x, y \in M$, we denote by $d_{SR}(x, y)$ (resp. $e_{SR}(x, y)$) the infimum of lengths (resp. energies) of horizontal paths joining x to y . We note that $e_{SR} = d_{SR}^2$. We denote the open ball and the sphere centered at x with radius $r > 0$ respectively by $B_{SR}(x, r)$ and $S_{SR}(x, r)$.
- We call minimizing geodesic from x to y any $\gamma \in W_{\Delta}^{1,2}([0, 1], M)$ with $\gamma(0) = x, \gamma(1) = y$ which minimizes the energy $e_{SR}(x, y)$ (and so the distance $d_{SR}(x, y)$), that is such that $\text{energy}^g(\gamma) = e_{SR}(x, y)$.
- For every $x \in M$, we denote by $W_{\Delta, x}^{1,2}([0, 1], M)$ the set of paths in $W_{\Delta}^{1,2}([0, 1], M)$ starting at x (that is $\gamma(0) = x$) and we define the end-point map

$$E_{\Delta}^x : W_{\Delta, x}^{1,2}([0, 1], M) \longrightarrow M$$

by $E_{\Delta}^x(\gamma) = \gamma(1)$. The infinite dimensional space $W_{\Delta, x}^{1,2}([0, 1], M)$ has a smooth manifold structure and the end-point map E_{Δ}^x is smooth.

- An horizontal path $\gamma \in W_{\Delta, x}^{1,2}([0, 1], M)$ is called singular if it is singular with respect to the end-point map E_{Δ}^x , that is if the differential $d_{\gamma} E_{\Delta}^x$ is not surjective. It is convenient to rewrite the definition of singular curves in term of singular controls. If the distribution Δ is parametrized by a family \mathcal{F} of k smooth vector fields X^1, \dots, X^k in a open neighborhood of $\gamma([0, 1])$ and if $u \in L^2([0, 1], \mathbb{R}^k)$ satisfies

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

then γ is singular if and only if the control u is a singular point of the smooth mapping (well-defined in an open set \mathcal{U})

$$E_{\mathcal{F}}^{x,1} : \mathcal{U} \subset L^2([0, 1], \mathbb{R}^k) \longrightarrow M$$

defined by

$$E_{\mathcal{F}}^{x,1}(v) := \gamma_v(1) \quad \forall v \in L^2([0,1], \mathbb{R}^k),$$

where γ_v is the curve in $W_{\Delta,x}^{1,2}([0,1], M)$ solution to the Cauchy problem

$$\dot{\gamma}_v(t) = \sum_{i=1}^k v_i(t) X^i(\gamma_v(t)) \quad \text{for a.e. } t \in [0,1], \quad \gamma_v(0) = x.$$

- An horizontal path $\gamma \in W_{\Delta,x}^{1,2}([0,1], M)$ is singular if and only if it is the projection of an abnormal extremal $\psi : [0,1] \rightarrow T^*M$ that never intersects the zero section of T^*M , such that

$$\dot{\psi}(t) = \sum_{i=1}^k u_i(t) \vec{h}^i(\psi(t)) \quad \text{for a.e. } t \in [0,1],$$

where \mathcal{F} is a family of k smooth vector fields X^1, \dots, X^k which parametrizes Δ in a open neighborhood of $\gamma([0,1])$ and $\vec{h}^1, \dots, \vec{h}^k$ are the Hamiltonian vector fields associated canonically with $h^i(x, p) = p \cdot X^i(x)$ in T^*M . The curve ψ is called an abnormal lift of γ and γ is said to be abnormal.

- The Hamiltonian $H : T^*M \rightarrow \mathbb{R}$ associated with (Δ, g) is defined by

$$H(x, p) := \frac{1}{2} \max \left\{ \frac{p(v)^2}{g_x(v, v)} \mid v \in \Delta(x) \setminus \{0\} \right\} \quad \forall (x, p) \in T^*M,$$

which coincides with

$$\frac{1}{2} \sum_{i=1}^m (p \cdot X^i(x))^2,$$

if Δ is parametrized locally by an orthonormal family X^1, \dots, X^m . The Hamiltonian vector field \vec{H} associated with (Δ, g) is the Hamiltonian vector field given by H with respect to the canonical symplectic form on T^*M . In local coordinates (x, p) the trajectories $\psi = (x, p)$ of \vec{H} are solution to

$$\dot{x} = \frac{\partial H}{\partial p}(x, p), \quad \dot{p} = -\frac{\partial H}{\partial x}(x, p),$$

we call them normal extremals. Any projection of a normal extremal is an horizontal path that is said to be normal.

- An horizontal path γ is called strictly abnormal if it is abnormal (singular) and not normal.

- For every $x \in M$, the exponential mapping $\exp_x : T_x^*M \rightarrow M$ associated with (Δ, g) at x is defined by $\exp_x := \pi(\psi_{x,p}(1))$ where $\psi_{x,p}$ is the trajectory of \vec{H} starting at (x, p) and $\pi : T_x^*M \rightarrow M$ is the canonical projection.
- A Carnot group (\mathbb{G}, \star) of step s is a simply connected Lie group whose Lie algebra $\mathfrak{g} = T_0\mathbb{G}$ (we denote by 0 the identity element of \mathbb{G}) admits a nilpotent stratification of step s , *i.e.*

$$\mathfrak{g} = V_1 \oplus \cdots \oplus V_s, \quad (\text{A.1})$$

with

$$[V_1, V_j] = V_{j+1} \quad \forall 1 \leq j \leq s, \quad V_s \neq \{0\}, \quad V_{s+1} = \{0\}. \quad (\text{A.2})$$

By simple-connectedness of \mathbb{G} and nilpotency of \mathfrak{g} , $\exp_{\mathbb{G}}$ is a smooth diffeomorphism, which allows to identify \mathbb{G} with its Lie algebra $\mathfrak{g} \simeq \mathbb{R}^n$. If the first layer V_1 of \mathbb{G} is equipped with a left-invariant metric, then there is a set of coordinates (x_1, \dots, x_n) , a one-parameter family of dilations $\{\delta_\lambda\}_{\lambda>0}$ of the form

$$\delta_\lambda(x_1, \dots, x_n) = (\lambda^{d_1}x_1, \lambda^{d_2}x_2, \dots, \lambda^{d_n}x_n) \quad \forall x \in \mathbb{R}^n,$$

and a orthonormal family of left-invariant vector fields generating V_1 satisfying

$$X^i(\delta_\lambda(x)) = \lambda^{-1} \delta_\lambda(X^i(x)) \quad \forall \lambda > 0, x \in \mathbb{R}^n.$$

- A function $f : U \rightarrow \mathbb{R}$ on a open set $U \subset M$ is called locally semiconcave if for every $x \in U$ there are a open neighborhood $V \subset U$ of x and $C > 0$ such that for any $y \in V$ there is a function $\psi : M \rightarrow \mathbb{R}$ with $\|\psi\|_{C^2} \leq C$ such that $f \leq \psi$ on M and $f(y) = \psi(y)$. For every $y \in U$, $d_y^+ f$ denotes the set of super-differentials of f at y , it is the set of $\alpha \in T_x^*M$ for which there is a function of class C^1 , $\psi : M \rightarrow \mathbb{R}$ such that $\psi \geq f$ on M , $\psi(y) = f(y)$ and $d_y\psi = \alpha$.
- If $f : U \rightarrow \mathbb{R}$ is smooth on the open set $U \subset M$, $\nabla^h f$ denotes its horizontal gradient with respect to (Δ, g) . For every $y \in U$, $\nabla_y^h f$ is defined as the unique $v \in \Delta(y)$ such that $d_y f(w) = \langle v, w \rangle$ for all $w \in \Delta(y)$. If $\Delta(y)$ is generated by an orthonormal family $X^1(y), \dots, X^m(y)$, then $\nabla_y^h f = \sum_{i=1}^m (X^i \cdot f)(y) X^i(y)$.

References

- [1] A. Agrachev. Compactness for Sub-Riemannian length-minimizers and subanalyticity. Control theory and its applications (Grado, 1998). *Rend. Sem. Mat. Univ. Politec. Torino*, 56(4):1–12, 2001.

- [2] A. Agrachev. Any sub-Riemannian metric has points of smoothness. *Dokl. Akad. Nauk*, 424(3): 295–298, 2009 Translation in *Dokl. Math.*, 79(1):45–47, 2009.
- [3] A. Agrachev. Some open problems. *Geometric control theory and sub-Riemannian geometry*, 1-13. Springer INdAM, Ser. 5, Springer, Cham, 2014.
- [4] A. Agrachev, D. Barilari and U. Boscain. *Introduction to Riemannian and sub-Riemannian geometry*. Book in preparation, 2017.
- [5] A. Agrachev and P. Lee. Optimal transportation under nonholonomic constraints. *Trans. Amer. Math. Soc.*, 361(11):6019–6047, 2009.
- [6] A. Agrachev and P. Lee. Generalized Ricci curvature bounds for three dimensional contact subriemannian manifolds. *Math. Ann.*, 360(1-2):209–253, 2014.
- [7] A.A. Agrachev and Y.L. Sachkov. *Control Theory from the Geometric Viewpoint*. Encyclopaedia of Mathematical Sciences, vol. 87, Springer, Heidelberg, 2004.
- [8] A. Agrachev and A. Sarychev. Sub-Riemannian metrics: minimality of singular geodesics versus subanalyticity. *ESAIM Control Optim. Calc. Var.*, 4:377–403, 1999.
- [9] D. Barilari and L. Rizzi. Sharp measure contraction property for generalized H -type Carnot groups. *Commun. Contemp. Math.*, to appear.
- [10] A. Belotto da Silva and L. Rifford. The Sard conjecture on Martinet surfaces. Preprint, 2016.
- [11] B. Bonnard and H. Heutte. La propriété de stricte anormalité est générique. *Preprint de l'Université de Bourgogne*, no. 77, 1995.
- [12] P. Cannarsa and L. Rifford. Semiconcavity results for optimal control problems admitting no singular minimizing controls. *Ann. Inst. H. Poincaré Non Linéaire*, 25(4):773–802, 2008.
- [13] P. Cannarsa and C. Sinestrari. *Semiconcave functions, Hamilton-Jacobi equations, and optimal control*. Progress in Nonlinear Differential Equations and their Applications, 58. Birkhäuser Boston Inc., Boston, MA, 2004.
- [14] Y. Chitour, F. Jean and E. Trélat. Genericity results for singular curves. *J. Diff. Geom.*, 73(1):45–73, 2006.
- [15] A. Figalli and L. Rifford. Mass Transportation on sub-Riemannian Manifolds. *Geom. Funct. Anal.*, 20(1):124–159, 2010.

- [16] N. Juillet. Geometric inequalities and generalized Ricci bounds in the Heisenberg group. *Int. Math. Res. Not. IMRN*, 13:2347–2373, 2009.
- [17] P.W.Y. Lee. On measure contraction property without Ricci curvature lower bound. *Potential Anal.*, 44:(1):27–41, 2016.
- [18] 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:(1):303–321, 2016.
- [19] J. Lott and C. Villani. Weak curvature conditions and functional inequalities. *J. Funct. Anal.*, 245(1):311–333, 2007.
- [20] R. Montgomery. *A tour of sub-Riemannian geometries, their geodesics and applications. Mathematical Surveys and Monographs*, Vol. 91. American Mathematical Society, Providence, RI, 2002.
- [21] S. Ohta. On the measure contraction property of metric measure spaces. *Comment. Math. Helv.*, 82(4):805–828, 2007.
- [22] L. Rifford. *Sub-Riemannian Geometry and Optimal Transport*. Springer Briefs in Mathematics, Springer, New York, 2014.
- [23] L. Rifford. Ricci curvatures in Carnot groups. *Math. Control Relat. Fields*, 3(4):467–487, 2013.
- [24] L. Rifford. Singulières minimisantes en géométrie sous-Riemannienne. Séminaire Bourbaki, Vol. 2015/2016. Exposés 1101-1119. *Astérisque*, Vol. 390, Exp. No. 1179: 277-301, 2017.
- [25] L. Rifford and E. Trélat. Morse-Sard type results in sub-Riemannian geometry. *Math. Ann.*, 332(1):145–159, 2005.
- [26] L. Rizzi. Measure contraction properties of Carnot groups. *Calc. Var. Partial Differential Equations*, 55(3), Art. 60, 20pp., 2016.
- [27] K. -T. Sturm. On the geometry of metric measure spaces. I. *Acta Math.*, 196(1):65–131, 2006.
- [28] K. -T. Sturm. On the geometry of metric measure spaces. II. *Acta Math.*, 196(1):133–177, 2006.
- [29] C. Villani. *Optimal transport, Old and New*. Grundlehren der Mathematischen Wissenschaften, 338. Springer-Verlag, Berlin, 2009.