

---

7-1-2016

## Regularity For Subelliptic Pde Through Uniform Estimates In Multi-scale Geometries

Luca Capogna  
*Worcester Polytechnic Institute, lcapogna@smith.edu*

Giovanna Citti  
*Alma Mater Studiorum Università di Bologna*

Follow this and additional works at: [https://scholarworks.smith.edu/mth\\_facpubs](https://scholarworks.smith.edu/mth_facpubs)



Part of the [Mathematics Commons](#)

---

### Recommended Citation

Capogna, Luca and Citti, Giovanna, "Regularity For Subelliptic Pde Through Uniform Estimates In Multi-scale Geometries" (2016). Mathematics and Statistics: Faculty Publications, Smith College, Northampton, MA.

[https://scholarworks.smith.edu/mth\\_facpubs/129](https://scholarworks.smith.edu/mth_facpubs/129)

This Article has been accepted for inclusion in Mathematics and Statistics: Faculty Publications by an authorized administrator of Smith ScholarWorks. For more information, please contact [scholarworks@smith.edu](mailto:scholarworks@smith.edu)

# Regularity for subelliptic PDE through uniform estimates in multi-scale geometries

Luca Capogna<sup>1</sup> · Giovanna Citti<sup>2</sup>

Received: 27 February 2015 / Revised: 26 September 2015 / Accepted: 29 September 2015 /  
Published online: 12 October 2015  
© The Author(s) 2015. This article is published with open access at Springerlink.com

**Abstract** We aim at reviewing and extending a number of recent results addressing stability of certain geometric and analytic estimates in the Riemannian approximation of subRiemannian structures. In particular we extend the recent work of the the authors with Rea (Math Ann 357(3):1175–1198, 2013) and Manfredini (Anal Geom Metric Spaces 1:255–275, 2013) concerning stability of doubling properties, Poincaré inequalities, Gaussian estimates on heat kernels and Schauder estimates from the Carnot group setting to the general case of Hörmander vector fields.

**Keywords** Subelliptic PDE · SubRiemannian geometry · Riemannian approximation · Heat kernels

---

Communicated by Nil Trudinger.

---

L. C. is partially supported by NSF awards DMS-1449143 and DMS-1503683.  
G. C. is partially funded by MAnET “Metric Analysis for Emergent Technologies”, Marie Curie Initial Training Network, Grant agreement No. 607643.

---

✉ Luca Capogna  
lcapogna@wpi.edu  
Giovanna Citti  
citti@unibo.it

<sup>1</sup> Department of Mathematical Sciences, Worcester Polytechnic Institute,  
Worcester, MA 01609, USA

<sup>2</sup> Dipartimento di Matematica, Università di Bologna,  
Bologna, Italy

## Contents

1	Introduction	174
2	Definitions and preliminary results	181
2.1	Carnot–Caratheodory distance	182
2.2	The approximating distances	183
2.3	A special case: the Heisenberg group $\mathbb{H}^1$	184
3	Stability of the homogenous structure	188
3.1	The Nagel–Stein–Wainger estimates	188
3.2	Uniform estimates as $\epsilon \rightarrow 0$	189
3.3	Equiregular subRiemannian structures and equivalent pseudo-distances	194
4	Stability of the Poincaré inequality	198
5	Stability of heat Kernel estimates	199
5.1	Hörmander type parabolic operators in non divergence form	199
5.2	The Rothschild–Stein freezing and lifting theorems	203
5.3	A lifting procedure uniform in $\epsilon$	205
5.4	Proof of the stability result	206
5.5	Differential of the integral operator associated to $\Gamma_\epsilon$	208
6	Stability of interior Schauder estimates	211
6.1	Uniform Schauder estimates	211
7	Application I: Harnack inequalities for degenerate parabolic quasilinear equations hold uniformly in $\epsilon$	216
7.1	Admissible ambient space geometry	217
7.2	Quasilinear degenerate parabolic PDE	219
8	Application II: regularity for quasilinear subelliptic PDE through Riemannian approximation	222
8.1	Interior gradient bounds	224
8.2	Global gradient bounds	225
8.3	Harnack inequalities and $C^{1,\alpha}$ estimates	226
8.4	Schauder estimates and higher order estimates	226
	References	227

## 1 Introduction

A subRiemannian manifold as a triplet  $(M, \Delta, g_0)$  where  $M$  is a connected, smooth manifold of dimension  $n \in \mathbb{N}$ ,  $\Delta$  denotes a subbundle of  $TM$  bracket-generating  $TM$ , and  $g_0$  is a positive definite smooth, bilinear form on  $\Delta$ , see for instance [66]. Similarly to the Riemannian setting, one endows  $(M, \Delta, g_0)$  with a metric space structure by defining the *Carnot–Caratheodory* (CC) control distance: For any pair  $x, y \in M$  set

$$d_0(x, y) = \inf\{\delta > 0 \text{ such that there exists a curve } \gamma \in C^\infty([0, 1]; M) \\ \text{with endpoints } x, y \text{ such that } \dot{\gamma} \in \Delta(\gamma) \text{ and } |\dot{\gamma}|_g \leq \delta\}.$$

Curves whose velocity vector lies in  $\Delta$  are called *horizontal*, their length is defined in an obvious way. Subriemannian metrics can be defined, by prescribing a smooth distributions of vector fields  $X = (X_1, \dots, X_m)$  in  $R^n$ , orthonormal with respect to  $g_0$ , and satisfying the Hörmander finite rank condition

$$\text{rank Lie}(X_1, \dots, X_m)(x) = n, \quad \forall x \in \Omega. \quad (1.1)$$

When attempting to extend known Riemannian results to the subRiemannian setting one naturally is led to approximating the sub-Riemannian metric (and the associated

distance function  $d_0(\cdot, \cdot)$ ) with a one-parameter family of degenerating Riemannian metric (associated to distance functions  $d_\epsilon(\cdot, \cdot)$ ), which converge in the Gromov–Hausdorff sense as  $\epsilon \rightarrow 0$  to the original one. This approximation is described in detail in from the point of view of the distance functions in Sect. 2.2 and from the point of view of the Riemannian setting in Definition 3.7. The approximating distance functions  $d_\epsilon$  can be defined in terms of an extended generating frame of smooth vector fields  $X_1^\epsilon, \dots, X_p^\epsilon$ , with  $p \geq n$  and  $X_i^\epsilon = X_i$  for  $i = 1, \dots, m$ , that converges/collapses uniformly on compact sets to the original family  $X_1, \dots, X_m$  as  $\epsilon \rightarrow 0$ . This frame includes all the higher order commutators needed to bracket generate the tangent space. When coupled with uniform estimates, this method provides a strategy to extend known Riemannian results to the subRiemannian setting. Such approximations have been widely used since the mid-80’s in a variety of contexts. As example we recall the work of Debiard [33], Koranyi [55,56], Ge [45], Rumin [77] as well as the references in [67] and [68]. More recently this technique has been used in the study of minimal surfaces and mean curvature flow in the Heisenberg group. Starting from the existence theorem of Pauls [71], and Cheng et al. [24], to the regularity results by Manfredini and the authors [15,16]. Our work is largely inspired to the results of Manfredini and one of us [27] where the Nagel et al. estimates for the fundamental solution of subLaplacians have been extended to the Riemannian approximants uniformly as  $\epsilon \rightarrow 0$ . In the following we list in more detail the nature of the stability estimates we investigate. Given a Riemannian manifold  $(M^n, g)$ , with a Riemannian smooth volume form expressed in local coordinates  $(x_1, \dots, x_n)$  as  $d \text{ vol} = \sqrt{g} dx_1 \dots dx_n$ , one can consider the corresponding heat operator acting on functions  $u : M \rightarrow \mathbb{R}$ ,

$$L_g u = \partial_t u - \frac{1}{\sqrt{g}} \sum_{i,j=1}^n \partial_i \left( \sqrt{g} g^{ij} \partial_j u \right).$$

The study of such operators is closely related to certain geometric and analytic estimates, namely: For  $K \subset\subset M$  and  $r_0 > 0$  there exists positive constants  $C_D, C_P, \dots$  below depending on  $K, r_0, g$  such that for all  $x \in K$  and  $0 < r < r_0$ , one has

- (Doubling property)

$$\text{vol}(B(x, r)) \geq C_D \text{vol}(B(x, 2r)); \tag{1.2}$$

- (Poincaré inequality)  $\int_{B(x,r)} |u - u_{B(x,r)}| d\text{vol} \leq C_P r \int_{B(x,2r)} |\nabla_g u| d\text{vol}$ ;
- (Gaussian estimates) If  $h_g$  denotes the heat kernel of  $L_g, x, y \in M$  and  $t > 0$  one has

$$\begin{aligned} & C_g^{-1} (\text{vol}(B(x, \sqrt{t})))^{-n/2} \exp\left(A_g \frac{d(x, y)^2}{t}\right) \\ & \leq |h(x, y, t)| \leq C_g (\text{vol}(B(x, \sqrt{t})))^{-n/2} \exp\left(B_g \frac{d(x, y)^2}{t}\right) \end{aligned} \tag{1.3}$$

and if appropriate curvature conditions hold

$$|\partial_t^s \partial_{t_1} \cdots \partial_{t_k} h(x, y, t, s) \leq C_{s,k,g} t^{-s-\frac{k}{2}} (\text{vol}(B(x, t-s)))^{-n/2} \exp\left(B_G \frac{d(x, y)^2}{t-s}\right); \tag{1.4}$$

- (Parabolic Harnack inequality) If  $L_g u = 0$  in  $Q = M \times (0, T)$  and  $u \geq 0$  then

$$\sup_{B(x,r) \times (t-r^2, t-r^2/2)} u \leq C_g \inf_{B(x,r) \times (t+r^2/2, t+r^2)} u. \tag{1.5}$$

The connections between such estimates was made evident in the work of Saloff-Coste [78] and Grigoryan [46], who independently established the equivalence

$$\begin{aligned} (\text{Poincare}) + (\text{Doubling}) &\iff \text{Gaussian estimates (1.3)} \\ &\iff \text{Parabolic Harnack inequality (1.5)}. \end{aligned}$$

See also related works by Biroli and Mosco [6], and Sturm [80].

*This paper aims at describing the behavior of such estimates along a sequence of metrics  $g_\epsilon$ , that collapse to a subRiemannian structure as  $\epsilon \rightarrow 0$ . We will prove that the estimates are stable as  $\epsilon \rightarrow 0$  and explore some of the consequences of this stability.* Although, thanks to the work of Jerison [52], Nagel et al. [70] and Jerison and Sanchez-Calle [53], the Poincarè inequality, the doubling property and the Gaussian bounds are well known for subRiemannian structures, it is not immediate that they continue to hold uniformly in the approximation as  $\epsilon \rightarrow 0$ . For one thing, the Riemannian curvature tensor is unbounded as  $\epsilon \rightarrow 0$ , thus preventing the use of Li-Yau’s estimates. Moreover, as  $\epsilon \rightarrow 0$  the Hausdorff dimension of the metric spaces  $(M, d_\epsilon)$ , where  $d_\epsilon$  denotes the distance function associated to  $g_\epsilon$ , typically does not remain constant and in fact increases at  $\epsilon = 0$  to the homogeneous dimension associated to the subRiemannian structure. The term *multiscale* from the title reflects the fact that the blow up of the metric as  $\epsilon \rightarrow 0$  is Riemannian at scales less than  $\epsilon$  and subRiemannian at larger scales.

To illustrate our work we introduce a prototype for the class of spaces we investigate, we consider the manifold  $M = \mathbb{R}^2 \times S^1$ , with coordinates  $(x_1, x_2, \theta)$ . The horizontal distribution is given by

$$\Delta = \text{span}\{X_1, X_2\}, \text{ with } X_1 = \cos \theta \partial_{x_1} + \sin \theta \partial_{x_2}, \text{ and } X_2 = \partial_\theta.$$

The subRiemannian metric  $g_0$  is defined so that  $X_1$  and  $X_2$  form a orthonormal basis. This is the group of Euclidean isometries defined below in Example 2.1. For each  $\epsilon > 0$  we also consider the Riemannian metric  $g_\epsilon$  on  $M$  uniquely defined by the requirement that  $X_1, X_2, \epsilon X_3$  is an orthonormal basis, with  $X_3 = -\sin \theta \partial_{x_1} + \cos \theta \partial_{x_2}$ . Denote by  $d_\epsilon$  the corresponding Riemannian distance, by  $X_i^*$  the adjoint of  $X_i$  with respect to Lebesgue measure and by  $\Gamma_\epsilon$  the fundamental solution of the Laplace-Beltrami operator  $L_\epsilon = \sum_{i=1}^3 X_i^* X_i$ . Since  $L_\epsilon$  is uniformly elliptic, then there exists  $C_\epsilon, R_\epsilon > 0$  such that for  $d_\epsilon(x, y) < R_\epsilon$  the fundamental solution will satisfy

$$C_\epsilon^{-1} d_\epsilon(x, y)^{-1} \leq \Gamma_\epsilon(x, y) \leq C_\epsilon d_\epsilon(x, y)^{-1}.$$

As  $\epsilon \rightarrow 0$  this estimate will degenerate in the following way:  $R_\epsilon \rightarrow 0$ ,  $C_\epsilon \rightarrow \infty$  and for  $\epsilon = 0$  one will eventually have

$$\Gamma_0(x, y) \approx d_0(x, y)^{-2}.$$

As a result of the work in [70] one has that for each  $\epsilon > 0$  there exists  $C_\epsilon > 0$  such that

$$C^{-1} \frac{d_\epsilon^2(x, y)}{|B_\epsilon(x, d(x, y))|} \leq \Gamma_\epsilon(x, y) \leq C \frac{d_\epsilon^2(x, y)}{|B_\epsilon(x, d(x, y))|}.$$

The main result of [27] was to provide stable bounds for the fundamental solution by proving that one can choose  $C_\epsilon$  independent of  $\epsilon$  as  $\epsilon \rightarrow 0$ . In this paper we extend such stable bounds to the degenerate parabolic setting and to the more general subRiemannian setting.

Since our results will be local in nature, unless explicitly stated we will always assume that  $M = \mathbb{R}^n$  and use as volume the Lebesgue measure. The first result we present is due to Rea and the authors [18] and concerns stability of the doubling property.

**Theorem 1.1** *For every  $\epsilon_0 > 0$ , and  $K \subset\subset \mathbb{R}^n$  there exist constants  $R, C > 0$  depending on  $K, \epsilon_0$  and on the subRiemannian structure, such that for every  $\epsilon \in [0, \epsilon_0]$ ,  $x \in K$  and  $0 < r < R$ ,*

$$|B_\epsilon(x, 2r)| \leq C|B_\epsilon(x, r)|.$$

Here we have denoted by  $B_\epsilon$  the balls related to the  $d_\epsilon$  distance function.

We present here a rather detailed proof of this result, amending some minor gaps in the exposition in [18]. If the subRiemannian structure is equiregular, as an original contribution of this paper, in Theorem 3.10 we also present a quantitative version of this result, by introducing an explicit quasi-norms equivalent to  $d_\epsilon$ . These families of quasi-norms play a role analogue to the one played by the Koranyi Gauge quasi-norm (2.5) in the Heisenberg group. We also sketch the proof of the stability of Jerison’s Poincare inequality from [18].

**Theorem 1.2** *Let  $K \subset\subset \mathbb{R}^n$  and  $\epsilon_0 > 0$ . The vector fields  $(X_i^\epsilon)_{i=1\dots p}$  satisfy the Poincare inequality*

$$\int_{B_\epsilon(x, R)} |u - u_{B_\epsilon(x, r)}| dx \leq C_P \int_{B_\epsilon(x, 2r)} |\nabla^\epsilon u| dx$$

with a constant  $C_P$  depending on  $K, \epsilon_0$  and the subRiemannian structure, but independent of  $\epsilon$ . Here we have denoted by  $\nabla^\epsilon u$  the gradient of  $u$  along the frame  $X_1^\epsilon, \dots, X_p^\epsilon$ .

Our next results concerns the stability, as  $\epsilon \rightarrow 0$ , of the Gaussian estimates for the heat kernels associated to the family of second order, sub-elliptic differential equations

in non divergence form

$$L_{\epsilon,A}u \equiv \partial_t u - \sum_{i,j=1}^p a_{ij}^\epsilon X_i^\epsilon X_j^\epsilon u = 0,$$

in a cylinder  $Q = \Omega \times (0, T)$ . Here  $\{a_{ij}^\epsilon\}_{i,j=1,\dots,p}$  is a constant real matrix such that

$$\frac{1}{2}\Lambda^{-1} \sum_{i=1}^p \xi_i^2 \leq \sum_{i,j=1}^p a_{ij}^\epsilon \xi_i \xi_j \leq 2\Lambda \sum_{i=1}^p \xi_i^2, \tag{1.6}$$

for all  $\xi \in \mathbb{R}^p$ , uniformly in  $\epsilon > 0$  and

$$\Lambda^{-1} \sum_{i=1}^m \xi_i^2 \leq \sum_{i,j=1}^m a_{ij}^\epsilon \xi_i \xi_j \leq \Lambda \sum_{i=1}^m \xi_i^2, \tag{1.7}$$

for all  $\xi \in \mathbb{R}^m$  and  $\epsilon > 0$ .

**Theorem 1.3** *Let  $K \subset\subset \mathbb{R}^n$ ,  $\Lambda > 0$  and  $\epsilon_0 > 0$ . The fundamental solution  $\Gamma_{\epsilon,A}$  of the operator  $L_{\epsilon,A}$ , is a kernel with exponential decay of order 2, uniform with respect to  $\epsilon \in [0, \epsilon_0]$  and for any coefficients matrix  $A$  satisfying the bounds above for the fixed  $\Lambda > 0$ . In particular, the following estimates hold:*

- For every  $K \subset\subset \Omega$  there exists a constant  $C_\Lambda > 0$  depending on  $\Lambda$  but independent of  $\epsilon \in [0, \epsilon_0]$ , and of the matrix  $A$  such that for each  $\epsilon \in [0, \epsilon_0]$ ,  $x, y \in K$  and  $t > 0$  one has

$$C_\Lambda^{-1} \frac{e^{-C_\Lambda \frac{d_\epsilon(x,y)^2}{t}}}{|B_\epsilon(x, \sqrt{t})|} \leq P_{\epsilon,A^\epsilon}(x, y, t) \leq C_\Lambda \frac{e^{-\frac{d_\epsilon(x,y)^2}{C_\Lambda t}}}{|B_\epsilon(x, \sqrt{t})|}. \tag{1.8}$$

- For  $s \in \mathbb{N}$  and  $k$ -tuple  $(i_1, \dots, i_k) \in \{1, \dots, m\}^k$  there exists a constant  $C_{s,k} > 0$  depending only on  $k, s, X_1, \dots, X_m, \Lambda$  such that

$$\left| (\partial_t^s X_{i_1} \cdots X_{i_k} P_{\epsilon,A^\epsilon})(x, y, t) \right| \leq C_{s,k} \frac{t^{-\frac{2s-k}{2}} e^{-\frac{d_\epsilon(x,y)^2}{C_\Lambda t}}}{|B_\epsilon(x, \sqrt{t})|} \tag{1.9}$$

for all  $x, y \in K$  and  $t > 0$ .

- For any  $A_1, A_2 \in M_\Lambda$ ,  $s \in \mathbb{N}$  and  $k$ -tuple  $(i_1, \dots, i_k) \in \{1, \dots, m\}^k$  there exists  $C_{s,k} > 0$  depending only on  $k, s, X_1, \dots, X_m, \Lambda$  such that

$$\begin{aligned} & \left| (\partial_t^s X_{i_1} \cdots X_{i_k} P_{\epsilon,A_1})(x, y, t) - \partial_t^s X_{i_1} \cdots X_{i_k} P_{\epsilon,A_2}(x, y, t) \right| \\ & \leq \|A_1 - A_2\| C_{s,k} \frac{t^{\frac{h-2s-k}{2}} e^{-\frac{d_\epsilon(x,y)^2}{C_\Lambda t}}}{|B_\epsilon(x, \sqrt{t})|}, \end{aligned} \tag{1.10}$$

where  $\|A\|^2 := \sum_{i,j=1}^n a_{ij}^2$ .

Moreover, if  $\Gamma_A$  denotes the fundamental solution of the operator  $L_A = \sum_{i,j=1}^m a_{ij}^0 X_i X_j$ , then one has

$$X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon \partial_t^\delta \Gamma_{\epsilon,A^\epsilon} \rightarrow X_{i_1} \cdots X_{i_k} \partial_t^\delta \Gamma_{A^0} \tag{1.11}$$

as  $\epsilon \rightarrow 0$  uniformly on compact sets and in a dominated way on subcompacts of  $\Omega$ .

This theorem extends to the general Hörmander vector fields setting analogue results proved by Manfredini and the authors in [17], in the setting of Carnot groups.

In a similar fashion, one of our main result in this paper is the extension to the Hörmander vector fields setting of the Carnot groups Schauder estimates established in previous work with Manfredini in [17]. To prove such extension we combine the Gaussian bounds above with a refined version of Rothschild and Stein [76] freezing and lifting scheme, adapted to the multi-scale setting, to establish Schauder type estimates which are uniform in  $\epsilon \in [0, \epsilon_0]$ , for the family of second order, sub-elliptic differential equations in non divergence form

$$L_{\epsilon,A^\epsilon} u \equiv \partial_t u - \sum_{i,j=1}^n a_{ij}^\epsilon(x,t) X_i^\epsilon X_j^\epsilon u = 0, \tag{1.12}$$

in a cylinder  $Q = \Omega \times (0, T)$ . Our standing assumption is that the coefficients of the operator satisfy (1.6), and (1.7) for some fixed  $\Lambda > 0$ .

**Theorem 1.4** *Let  $\alpha \in (0, 1)$ ,  $f \in C^\infty(Q)$  and  $w$  be a smooth solution of  $L_{\epsilon,A^\epsilon} w = f$  on  $Q$ . Let  $K$  be a compact sets such that  $K \subset\subset Q$ , set  $2\delta = d_0(K, \partial_p Q)$  and denote by  $K_\delta$  the  $\delta$ -tubular neighborhood of  $K$ . Assume that there exists a constant  $C > 0$  such that*

$$\|a_{ij}^\epsilon\|_{C_{\epsilon,X}^{k,\alpha}(K_\delta)} \leq C,$$

for some value  $k \in \mathbb{N}$  and for every  $\epsilon \in [0, \epsilon_0]$ . There exists a constant  $C_1 > 0$  depending on  $\alpha, C, \epsilon_0, \delta$ , and the constants in Proposition 5.2, but independent of  $\epsilon$ , such that

$$\|w\|_{C_{\epsilon,X}^{k+2,\alpha}(K)} \leq C_1 \left( \|f\|_{C_{\epsilon,X}^{k,\alpha}(K_\delta)} + \|w\|_{C_{\epsilon,X}^{k+1,\alpha}(K_\delta)} \right).$$

Here we have set

$$\|u\|_{C_{\epsilon,X}^\alpha(Q)} = \sup_{(x,t) \neq (x_0,t_0)} \frac{|u(x,t) - u(x_0,t_0)|}{\tilde{d}_\epsilon^\alpha((x,t), (x_0,t_0))} + \sup_Q |u|.$$

and if  $k \geq 1$  we have let  $u \in C_{\epsilon,X}^{k,\alpha}(Q)$  if for all  $i = 1, \dots, m$ , one has  $X_i \in C_{\epsilon,X}^{k-1,\alpha}(Q)$ .



Analogous estimates in the  $L^p$  spaces, for operators independent of  $\epsilon$  are well known (see for instance [76] for the constant coefficient case and [9] for the Carnot group setting). Our result yield a stable version, as  $\epsilon \rightarrow 0$ , of such estimates, which is valid for any family of Hörmander vector fields.

**Theorem 1.5** *Let  $\alpha \in (0, 1)$ ,  $f \in C^\infty(Q)$  and  $w$  be a smooth solution of  $L_{\epsilon,A}w = f$  on  $Q$ . Let  $K$  be a compact sets such that  $K \subset\subset Q$ , set  $2\delta = d_0(K, \partial_p Q)$  and denote by  $K_\delta$  the  $\delta$ -tubular neighborhood of  $K$ . Assume that there exists a constant  $C > 0$  such that*

$$\|a_{ij}^\epsilon\|_{C_{\epsilon,X}^{k,\alpha}(K_\delta)} \leq C,$$

for some value  $k \in \mathbb{N}$  and for every  $\epsilon \in [0, \epsilon_0]$ . For any  $p > 1$ , there exists a constant  $C_1 > 0$  depending on  $p, \alpha, C, \epsilon_0, \delta$ , and the constants in Proposition 5.2, but independent of  $\epsilon$ , such that

$$\|w\|_{W_{\epsilon,X}^{k+2,p}(K)} \leq C_1 \left( \|f\|_{W_{\epsilon,X}^{k,p}(K_\delta)} + \|w\|_{W_{\epsilon,X}^{k+1,p}(K_\delta)} \right).$$

Here we have set

$$\|w\|_{W_{\epsilon,X}^{k,p}} := \sum_{i=1}^k \sum_{|J|=i} \|X_{i_1}^\epsilon \dots X_{i_k}^\epsilon w\|_{L^p}.$$

We conclude the paper with two, related, groups of applications of our stability results. In the first we recall the notion of  $p$ -admissible structure (Definition 7.1), originally introduced by Hajlasz and Koskela in [48]. This class of spaces supports a rich analytic structure and allows for the development of a first-order (in the sense of derivatives up to order one) potential theory. We review some recent results by the authors and collaborators [1, 18] concerning Harnack inequalities for weak solutions of classes of quasilinear degenerate parabolic PDE in such spaces. The main point of the section is that in view of Theorems 1.1 and 1.2, the Riemannian approximations of a subRiemannian structure satisfy the hypothesis of  $p$ -admissible structure uniformly in  $\epsilon \geq 0$ . Consequently, the Harnack inequalities hold uniformly across all scales. This provides a powerful technique in the study of degenerate elliptic and parabolic problems through the process of regularization and approximation. To exemplify this observation in a simple case, we consider approximating Riemannian metrics  $g_\epsilon$  with generating frame  $X_1^\epsilon, \dots, X_n^\epsilon$  defined in an open set  $\Omega \subset \mathbb{R}^n$ , and a family of divergence form parabolic linear equations analogue to (1.12), i.e.

$$L_{\epsilon,A}u \equiv \partial_t u - \sum_{i,j=1}^n X_i^{\epsilon,*}(a_{ij}^\epsilon(u, x, t)X_j^\epsilon u) = 0, \tag{1.13}$$

in a cylinder  $Q = \Omega \times (0, T)$ . We assume that the coefficients of the operator depend smoothly on  $u$  and satisfy (1.6), and (1.7) for some fixed  $\Lambda > 0$ . Thanks to the stability

estimates one can prove that there exists positive constants  $C, R > 0$  depending only on the fixed subRiemannian structure, but independent of  $\epsilon$ , such that for any  $\epsilon \geq 0$  and any non-negative weak solution  $u \geq 0$  of (1.13), one has

$$\sup_{B_\epsilon(x,r)} u \leq C \inf_{B_\epsilon(x,r)} u$$

for any metric ball  $B_\epsilon(x, 4r) \subset \Omega$  and  $0 < r < R$ . Clearly this yields Hölder regularity for  $u$  that is stable as  $\epsilon \rightarrow 0$ . Applying the Schauder estimates from Theorem 1.4 one obtains higher order regularity, uniformly in  $\epsilon \rightarrow 0$  and so in particular we obtain smoothness of solutions in the case  $\epsilon = 0$ . For further details and for a more general version of this result, applied to weak solutions of quasilinear equations, we refer the reader to Theorem 7.5.

In the last section we discuss one of the motivating applications of our work. We outline how the structure stability results, the stability of the Schauder estimates and of the Harnack inequalities can be used to prove regularity and long time existence theorems for solutions of the subRiemannian mean curvature flow and the total curvature flow of graphs over bounded sets in step 2 Carnot groups and even in some non-nilpotent Lie groups. This is part of the work developed by the authors jointly with Maria Manfredini in [14, 17]. The notion of horizontal, or  $p$ -mean curvature has arisen in the last 10 years thanks to the work of many researchers. The two main motivations are Pansu conjecture, concerning the isoperimetric profile of the Heisenberg group [19, 26, 43, 49, 51, 69, 73–75]; and the existence, regularity and uniqueness of minimal surfaces [20–25, 31, 32, 50, 71], and [72]. The mean curvature flow and the total curvature flow arise in connection to gradient descent for the perimeter functional and as such can be used for both applications. Very little is known about both flows in the subRiemannian setting and as far as we know the results in [14, 17] are the first to establish existence of long time smooth flows. For other contributions to this topics, from different points of view, we recall the recent work in [35, 36].

## 2 Definitions and preliminary results

Let  $X = (X_1, \dots, X_m)$  denote a collection of smooth vector fields defined in an open subset  $\Omega \subset \mathbb{R}^n$  satisfying Hörmander’s finite rank condition (1.1), that is *there exists an integer  $s$  such that the set of all vector fields, along with their commutators up to order  $s$  spans  $\mathbb{R}^n$  for every point in  $\Omega$ ,*

$$\text{rank Lie}(X_1, \dots, X_m)(x) = n, \quad \text{for all } x \in \Omega. \tag{2.1}$$

*Example 2.1* The standard example for such families is the Heisenberg group  $\mathbb{H}^1$ . This is a Lie group whose underlying manifold is  $\mathbb{R}^3$  and is endowed with a group law  $(x_1, x_2, x_3)(y_1, y_2, y_3) = (x_1 + y_1, x_2 + y_2, x_3 + y_3 - (x_2y_1 - x_1y_2))$ . With respect to such law one has that the vector fields  $X_1 = \partial_{x_1} - x_2\partial_{x_3}$  and  $X_2 = \partial_{x_2} + x_1\partial_{x_3}$  are left-invariant. Together with their commutator  $[X_1, X_2] = 2\partial_{x_3}$  they yield a basis of  $\mathbb{R}^3$ . A second example is given by the classical group of rigid motions of the plane, also known as the *roto-translation* group  $\mathcal{RT}$ . This is a Lie group with underlying manifold

$\mathbb{R}^2 \times S^1$  and a group law  $(x_1, x_2, \theta_1)(y_1, y_2, \theta_2) = (x_1 + y_1 \cos \theta - y_2 \sin \theta, x_2 + y_1 \sin \theta + y_2 \cos \theta, \theta_1 + \theta_2)$ .

Following Nagel, Stein and Wainger, [70, page 104] we define

$$X^{(1)} = \{X_1, \dots, X_m\}, X^{(2)} = \{[X_1, X_2], \dots, [X_{m-1}, X_m]\}, \text{ etc.} \dots \tag{2.2}$$

letting  $X^{(k)}$  denote the set of all commutators of order  $k = 1, \dots, r$ . Indicate by  $Y_1, \dots, Y_p$  an enumeration of the components of  $X^{(1)}, X^{(2)}, \dots, X^{(r)}$  such that  $Y_i = X_i$  for every  $i \leq m$ . If  $Y_k \in X^{(i)}$  we say that  $Y_k$  has a formal degree  $d(Y_k) = d(k) = i$ . The collection of vector fields  $\{Y_1, \dots, Y_p\}$  spans  $\mathbb{R}^n$  at every point.

*Example 2.2* If we consider the Heisenberg group vector fields  $X_1 = \partial_{x_1} - x_2 \partial_{x_3}$  and  $X_2 = \partial_{x_2} + x_1 \partial_{x_3}$  with  $(x_1, x_2, x_3) \in \mathbb{R}^3$ , then  $X^{(1)} := \{X_1, X_2\}$  and  $X^{(2)} = \{2\partial_{x_3}\}$ . If we instead consider the vectors arising from the group of roto-translations one has  $X_1 = \cos \theta \partial_{x_1} + \sin \theta \partial_{x_2}$  and  $X_2 = \partial_\theta$  with  $(x_1, x_2, \theta) \in \mathbb{R}^2 \times S^1$  and  $X^{(1)} = \{X_1, X_2\}$  and  $X^{(2)} = \{\sin \theta \partial_{x_1} - \cos \theta \partial_{x_2}\}$ .

*Example 2.3* Note that the sets  $X^{(i)}$  may have non-trivial intersection. For instance, consider the vector fields

$$X_1 = \cos \theta \partial_{x_1} + \sin \theta \partial_{x_2}; X_2 = \partial_\theta; X_3 = \partial_{x_3}; \text{ and } X_4 = x_3^2 \partial_{x_4}$$

in  $(x_1, x_2, x_3, x_4, \theta) \in \mathbb{R}^4 \times S^1$ . In this case  $r = 3$  and

$$\begin{aligned} X^{(1)} &= \{X_1, X_2, X_3, X_4\}; X^{(2)} \\ &= \{\sin \theta \partial_{x_1} - \cos \theta \partial_{x_2}, 2x_3 \partial_{x_4}\}; \text{ and } X^{(3)} = \{\pm X_1, 2\partial_{x_4}\} \end{aligned}$$

with  $Y_1 = X_1, \dots, Y_4 = X_4, Y_5 = \sin \theta \partial_{x_1} - \cos \theta \partial_{x_2}, y_6 = 2x_3 \partial_{x_4}, Y_7 = X_1, Y_8 = -X_1$ , and  $Y_{10} = 2\partial_{x_4}$ .

### 2.1 Carnot–Caratheodory distance

For each  $x, y \in \Omega$  and  $\delta > 0$  denote by  $\Gamma(\delta)$  the space of all absolutely continuous curves  $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ , joining  $x$  to  $y$  (i.e.,  $\gamma(0) = x$  and  $\gamma(1) = y$ ) which are tangent a.e. to the horizontal distribution  $span\{X_1, \dots, X_m\}$ , and such that if we write

$$\gamma'(t) = \sum_{i=1}^m \alpha_i(t) X_i|_{\gamma(t)},$$

then  $\sum_{i=1}^m |\alpha_i(t)| \leq \delta$  a.e.  $t \in [0, 1]$ . The Carnot–Caratheodory distance between  $x$  and  $y$  is defined to be

$$d_0(x, y) := \inf_{\Gamma(\delta) \neq \text{Empty Set}} \delta. \tag{2.3}$$

In [70], the authors introduce several other distances that eventually are proved to be equivalent to  $d_0(x, y)$ . The equivalence itself yields new insight into the Carnot–Caratheodory distance. Because of this, we will remind the reader of one of these distances. For each  $x, y \in \Omega$  and  $\delta > 0$  denote by  $\hat{\Gamma}(\delta)$  the space of all absolutely continuous curves  $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ , joining  $x$  to  $y$  and such that if one writes

$$\gamma'(t) = \sum_{i=1}^p \beta_i(t) Y_i|_{\gamma(t)},$$

then  $|\beta_i(t)| \leq \delta^{d(i)}$ . One then sets

$$\hat{d}(x, y) := \inf_{\hat{\Gamma}(\delta) \neq \text{Empty Set}} \delta.$$

It is fairly straightforward (see [70, Proposition 1.1] to see that

**Proposition 2.4** *The function  $\hat{d}$  is a distance function in  $\Omega$  and for any  $K \subset\subset \Omega$  there exists  $C = C(X_1, \dots, X_m, K) > 0$  such that*

$$C^{-1}|x - y| \leq \hat{d}(x, y) \leq C|x - y|^{\max_i d(i)}.$$

It is far less trivial to prove the following (see [70, Theorem 4])

**Theorem 2.5** *The distance functions  $d_0$  and  $\hat{d}$  are equivalent.*

### 2.2 The approximating distances

There are several possible definitions for Riemannian distance functions which approximate a Carnot–Caratheodory metric in the Gromov-Hausdroff sense.

**Definition 2.6** Let  $\{Y_1, \dots, Y_p\}$  be a generating family of vector fields constructed as in (2.2) from a family of Hörmander vector fields  $X_1, \dots, X_m$ . For every  $\epsilon > 0$  denote by  $d_\epsilon(\cdot, \cdot)$  the Carnot–Caratheodory metric associated to the family of vector fields  $(X_1^\epsilon, \dots, X_p^\epsilon)$ , defined as

$$X_i^\epsilon = \begin{cases} Y_i & \text{if } i \leq m, \\ \epsilon^{d(i)-1} Y_i & \text{if } m + 1 \leq i \leq p, \\ Y_{i-p+m} & \text{if } p + 1 \leq i \leq 2p - m \end{cases} \quad (2.4)$$

We will also define an extension of the degree function, setting  $d_\epsilon(i) = 1$  for all  $i \leq p$ , and  $d_\epsilon(i) = d(i - p + m)$  if  $i \geq p + 1$ . In order to simplify notations we will denote  $X = X^0$ ,  $d_0 = d$  and use the same notation for both families of vector fields (dependent or independent of  $\epsilon$ ).

Note that for every  $\epsilon \in (0, \bar{\epsilon})$  the sets  $\{X_i^\epsilon\}$  extends the original family of vector fields  $(X_i)$  to a new families of vector fields satisfying assumption (I) on page 107 [70], i.e. there exist smooth functions  $c_{jk}^l$ , depending on  $\epsilon$ , such that

$$[X_j^\epsilon, X_k^\epsilon] = \sum_{d_\epsilon(l) \leq d_\epsilon(j) + d_\epsilon(k)} c_{jk}^l X_l^\epsilon$$

and

$$\{X_j^\epsilon\}_{j=1}^{2p-m} \text{ span } \mathbb{R}^n \text{ at every point .}$$

*Remark 2.7* Note that the coefficients  $c_{jk}^l$  will be unbounded as  $\epsilon \rightarrow 0$ . In principle this could be a problem as the doubling constant in the proof in [70] depends indirectly from the  $C^r$  norm of these functions. In this survey we will describe a result, originally proved in [18], showing that this is not the case.

*Remark 2.8* It follows immediately from the definition that for fixed  $x, y \in \Omega$  the function  $d_\epsilon(x, y)$  is decreasing in  $\epsilon$  and for every  $\epsilon \in (0, \bar{\epsilon})$ ,

$$d_0(x, y) \geq d_\epsilon(x, y)$$

*Remark 2.9* Let us consider a special case where  $\dim \text{span}(X_1, \dots, X_m)$  is constant and the vector fields  $X_1, \dots, X_p$  are chosen to be linearly independent in  $\Omega$ . In this case we can consider two positive defined symmetric quadratic forms  $g_0$ , and  $\lambda$  defined respectively on the distribution  $H(x) = \text{span}(X_1, \dots, X_m)(x)$ , for  $x \in \Omega$  and on  $H^\perp(x)$ . The product metric  $g_0 \oplus \lambda$  is then a Riemannian metric on all of  $T\Omega$ . The form  $g_0$  is called a *subRiemannian* metric on  $\Omega$ , corresponding to  $H$ . Next, for every  $\epsilon \in (0, \bar{\epsilon}]$  reconsider the rescaled metric  $g_\epsilon := g_0 \oplus \epsilon^{-1}\lambda$  and the corresponding Riemannian distance function  $d_\epsilon$  in  $\Omega$ . The latter is bi-Lipschitz equivalent to the distance  $d_\epsilon$  defined above. In [45, Theorem 1.1] Ge proved that that as metric spaces, the sequence  $(\Omega, d_\epsilon)$  converges to  $(\Omega, d_0)$  as  $\epsilon \rightarrow 0$  in the sense of Gromov–Hausdorff. In this limit the Hausdorff dimension of the space degenerates from coinciding with the topological dimension, for  $\epsilon > 0$ , to a value  $Q > n$  which may change from open set to open set. We will go more in detail about this point in the next section. In this sense the limiting approximation scheme we are using can be described by the collapsing of a family of Riemannian metric to a subRiemannian metric. See also [68, Theorem 1.2.1] for yet another related Riemannian approximation scheme.

*Remark 2.10* From different perspectives, note that the subLaplacian associated to the family  $X_1^\epsilon, \dots, X_m^\epsilon$  i.e.  $\mathcal{L}u = \sum_{i=1}^m X_i^{\epsilon,2}u$  is an elliptic operator for all  $\epsilon > 0$ , degenerating to a subelliptic operator for  $\epsilon = 0$ .

### 2.3 A special case: the Heisenberg group $\mathbb{H}^1$

In this section we describe the behavior of the distance  $d_\epsilon$  (and of the corresponding metric balls  $B_\epsilon(x, r)$ ) as  $\epsilon \rightarrow 0$ , by looking at the special case of the Heisenber group.

In this setting we will also provide an elementary argument showing that the doubling property holds uniformly as  $\epsilon \rightarrow 0$ .

Consider the vector fields from Example 2.1  $X_1 = \partial_{x_1} - x_2\partial_{x_3}$ ,  $X_2 = \partial_{x_2} + x_1\partial_{x_3}$  and  $X_3 = \partial_{x_3}$  with  $(x_1, x_2, x_3) \in \mathbb{R}^3$ . The Carnot–Carathéodory distance  $d_0$  associated to the subRiemannian metric defined by the orthonormal frame  $X_1, X_2$  is equivalent to a more explicitly defined pseudo-distance function, that we call *gauge distance*, defined as

$$|x|^4 = \left(x_1^2 + x_2^2\right)^2 + x_3^2, \text{ and } \rho(x, y) = |y^{-1}x|, \tag{2.5}$$

where  $y^{-1} = (-y_1, -y_2, -y_3)$  and  $y^{-1}x = (x_1 - y_1, x_2 - y_2, x_3 - y_3 - (y_1x_2 - x_1y_2))$  is the *Heisenberg group multiplication*.

**Lemma 2.11** *For each  $x \in \mathbb{R}^3$ ,*

$$A^{-1}|x| \leq d_0(x, 0) \leq A|x|, \tag{2.6}$$

for some constant  $A > 0$ .

*Proof* Observe that the 1-parameter family of diffeomorphisms

$$(x_1, x_2, x_3) \rightarrow \delta_\lambda(x_1, x_2, x_3) := (\lambda x_1, \lambda x_2, \lambda^2 x_3)$$

satisfies  $|\delta_\lambda(x)| = \lambda|x|$ , and  $d\delta_\lambda X_i = \lambda X_i \circ \delta_\lambda$  for  $i = 1, 2$ . Consequently  $d_0(\delta_\lambda(x), \delta_\lambda(y)) = \lambda d_0(x, y)$ , and  $\delta_\lambda(B(0, 1)) = B(0, \lambda)$ . Since the unit ball  $B(0, 1)$  is a bounded open neighborhood of the origin, it will contain a set of the form  $|x| \leq A^{-1}$  and will be contained in a set of the form  $|x| \leq A$ . By applying  $\delta_\lambda$  we then have that for any  $R > 0$ ,

$$\left\{x \in \mathbb{R}^3 \mid |x| \leq A^{-1}R\right\} \subset B(0, R) \subset \{x \in \mathbb{R}^3 \mid |x| \leq AR\}$$

concluding the proof of (2.6). □

*Remark 2.12* Since the Heisenberg group is a Lie group, then it is natural to use a left-invariant volume form to measure the size of sets, namely the Haar measure. It is not difficult to see [29] that the Haar measure coincides with the Lebesgue measure in  $\mathbb{R}^3$ . It follows immediately from the previous lemma that the corresponding volume of a ball  $B(x, r)$  is

$$|B(x, r)| = Cr^4. \tag{2.7}$$

As a consequence one can show that the Hausdorff dimension of the metric space  $(\mathbb{H}^1, d_0)$  is 4. The Hausdorff dimension of any horizontal curve (i.e. tangent to the distribution generated by  $X_1$  and  $X_2$ ) is 1, while the Hausdorff dimension of the vertical  $z$ -axis is 2.

Next we turn our attention to the balls in the metrics  $g_\epsilon$  and the associated distance functions  $d_\epsilon$ . To better describe the approximate shape of such balls we define the

pseudo-distance function  $d_{G,\epsilon}(x, y) = N_\epsilon(y^{-1}x)$  corresponding to the regularized gauge function

$$N_\epsilon^2(x) = x_1^2 + x_2^2 + \min \left\{ |x_3|, \epsilon^{-2}x_3^2 \right\}. \tag{2.8}$$

Our next goal is to show that the Riemannian distance function  $d_\epsilon$  is well approximated by the gauge pseudo-distance  $d_{G,\epsilon}$ .

**Lemma 2.13** *There exists  $A > 0$  independent of  $\epsilon$  such that for all  $x, y \in \mathbb{R}^3$*

$$A^{-1}d_{G,\epsilon}(x, y) \leq d_\epsilon(x, y) \leq Ad_{G,\epsilon}(x, y). \tag{2.9}$$

The estimate (2.9) yields immediately

**Corollary 2.14** *The doubling property holds uniformly in  $\epsilon > 0$ .*

*Remark 2.15* Before proving (2.9) it is useful to examine a specific example: compare two trajectories from the origin  $0 = (0, 0, 0)$  to the point  $x = (0, 0, x_3)$ . The first is the segment  $\gamma_1$  defined by  $s \rightarrow (0, 0, x_3s)$ , for  $s \in [0, 1]$ . The length of this segment in the Riemannian metric  $g^\epsilon$  given by the orthonormal frame  $X_1, X_2, \epsilon X_3$  is

$$\ell_\epsilon(\gamma_1) = \epsilon^{-1}|x_3|.$$

We also consider a second trajectory  $\gamma_2$  given by the subRiemannian geodesic between the two points. In view of (2.6) the length of this curve in the subRiemannian metric  $g^0$  defined by the orthonormal frame  $X_1, X_2$  is proportional to  $\sqrt{|x_3|}$  and coincides with the length in the Riemannian metric  $g^\epsilon$ , i.e.

$$\ell_\epsilon(\gamma_2) = \ell_0(\gamma_2) \approx \sqrt{|x_3|}.$$

Since  $d_\epsilon$  is computed by selecting the shortest path between two points in the  $g^\epsilon$  metric, then if  $\sqrt{|x_3|} > \epsilon$  one will have  $d_\epsilon(x, 0) \leq \sqrt{|x_3|} \approx N_\epsilon(x)$ , whereas at *small scales* (i.e. for  $d_0(x, 0) < \epsilon$ ) one will have  $d_\epsilon(x, 0) \leq \epsilon^{-1}|x_3|$ . By left translation invariance of  $d_{G,\epsilon}$  we have that for any two points  $x = (x_1, x_2, s)$  and  $x' = (x_1, x_2, t)$ ,

$$d_\epsilon(x, x') \leq C \min \left( \epsilon^{-1}|t - s|, \sqrt{|t - s|} \right). \tag{2.10}$$

From this simple example one can expect that at large scale (i.e. for points  $d^0(x, 0) > \epsilon$ ) the Riemannian and the subRiemannian distances are approximately the same  $d_\epsilon(x, 0) \approx d_0(x, 0)$ .

*Proof* From the invariance by left translations of both  $d_{G,\epsilon}$  and  $d_\epsilon$  it is sufficient to prove that  $d_\epsilon(x, 0)$  and  $N_\epsilon(x)$  are equivalent. We begin by establishing the first inequality in (2.9), i.e. we want to show that there exists a positive constant  $A$  such that

$$A^{-1}N_\epsilon(x) \leq d_\epsilon(0, x).$$

Consider a point  $x = (x_1, x_2, x_3) \in \mathbb{R}^3$  and three curves

- A length minimizing curve  $\gamma : [0, 1] \rightarrow \mathbb{R}^3$  for the metric  $g_\epsilon$ , such that

$$d_\epsilon(0, x) = \ell_\epsilon(\gamma) := \int_0^1 \sqrt{a_1^2(t) + a_2^2(t) + \epsilon^{-2}a_3^2(t)} dt,$$

where  $\gamma'(t) = \sum_{i=1,3} a_i(t)X_i|_{\gamma(t)}$ .

- An horizontal curve  $\gamma_1 : [0, 1] \rightarrow \mathbb{R}^3$  with one end-point at the origin ( $t = 0$ ) and such that  $\gamma_1'(t) = a_1(t)X_1|_{\gamma(t)} + a_2(t)X_2|_{\gamma(t)}$ . Denote by  $P = \gamma_1(1)$  and observe that  $P = (x_1, x_2, p_3)$  for some value of  $p_3$  such that  $\int_0^1 a_3(t)dt = x_3 - p_3$ .
- A vertical segment  $\gamma_2 : [0, 1] \rightarrow \mathbb{R}^3$  with endpoints  $P$  and  $x$ , such that  $\gamma_2'(t) = a_3(t)X_3|_{\gamma_2(t)}$ . Note that

$$\begin{aligned} \epsilon^{-1}|x_3 - p_3| &\leq \left| \epsilon^{-1} \int_0^1 a_3(t)dt \right| \leq \int_0^1 |a_3(t)|\epsilon^{-1} dt \\ &= \ell_\epsilon(\gamma_2) \leq \ell_\epsilon(\gamma) \leq d_\epsilon(x, p). \end{aligned}$$

Observe that in view of the equivalence (2.6),

$$C^{-1}\sqrt{x_1^2 + x_2^2} \leq d_0(P, 0) \leq \ell_0(\gamma_1) = \ell_\epsilon(\gamma_1) \leq \ell_\epsilon(\gamma),$$

for some constant  $C > 0$ . On the other hand one also has

$$\epsilon^{-1}|x_3 - p_3| \leq d_\epsilon(x, p) \leq d_\epsilon(x, 0) + d_\epsilon(0, p) \leq d_\epsilon(x, 0) + \ell_\epsilon(\gamma_1) \leq 2d_\epsilon(0, x).$$

Hence if  $|p_3| \leq \frac{1}{2}|x_3|$  then  $|x_3 - p_3| \geq \frac{1}{2}|x_3|$  and consequently

$$d_\epsilon(x, 0) = \ell_\epsilon(\gamma) \geq \epsilon^{-1}|x_3 - p_3| \geq \min(\epsilon^{-1}|x_3|, \sqrt{|x_3|}).$$

The latter yields immediately that  $d_\epsilon(x, 0) \geq C^{-1}N_\epsilon(x)$ , for some value of  $C > 0$  independent of  $\epsilon > 0$ . Next we consider the case  $|p_3| > \frac{1}{2}|x_3|$ . This yields

$$\begin{aligned} \min(\epsilon^{-1}|x_3|, \sqrt{|x_3|}) &\leq \frac{1}{2} \min(\epsilon^{-1}|p_3|, \sqrt{|p_3|}) \leq \sqrt{|p_3|} \leq |P| \leq C d_0(P, 0) \\ C \ell_0(\gamma_1) = C \ell_\epsilon(\gamma_1) &\leq C \ell_\epsilon(\gamma) = C d_\epsilon(x, 0), \end{aligned}$$

where  $|P|$  is defined as in (2.6). In summary, so far we have proved the first half of (2.9).

To prove the second half of the inequality we consider an horizontal segment  $\Gamma_1$  joining the origin to  $Q = (x_1, x_2, 0)$ . Note that  $d_0(0, Q) = d_\epsilon(0, Q) = \ell_0(\Gamma_1) = \ell_\epsilon(\Gamma_1)$ . In view of (2.10) one has

$$d_\epsilon(0, x) \leq d_\epsilon(0, Q) + d_\epsilon(Q, x) = d_0(0, Q) + C \min(\epsilon^{-1}|x_3|, \sqrt{|x_3|}) \leq C N_\epsilon(x).$$

The latter completes the proof of (2.9). □



*Remark 2.16* Similar arguments continue to hold more in general, in the setting of Carnot groups.

As a consequence of Lemma 2.13, one has that for  $\epsilon > 0$  the metric space  $(\mathbb{R}^3, d_\epsilon)$  is locally bi-Lipschitz to the Euclidean space, and hence its Hausdorff dimension will be 3. As  $\epsilon \rightarrow 0$  the non-horizontal directions are *penalized* causing a sharp *phase transition* between the regime at  $\epsilon > 0$  and  $\epsilon = 0$ .

The intuition developed through this example hints at the multiple scale aspect of the  $d_\epsilon$  metrics: At scales smaller than  $\epsilon > 0$  the local geometry of the metric space  $(\mathbb{R}^3, d_\epsilon)$  is roughly Euclidean; For scales larger than  $\epsilon > 0$  it is subRiemannian. This intuition will inform the proofs of the stability for the doubling property in the next section.

### 3 Stability of the homogenous structure

The volume of Carnot–Caratheodory balls, and its doubling property, has been studied in Nagel, Stein and Wainger’s seminal work [70]. In this section we recall the main results in this paper and show how to modify their proof so that the stability of the doubling constant as  $\epsilon \rightarrow 0$  becomes evident.

#### 3.1 The Nagel–Stein–Wainger estimates

Consider the Carnot–Caratheodory metric  $d_\epsilon(\cdot, \cdot)$  associated to the family of vector fields  $(X_1^\epsilon, \dots, X_p^\epsilon)$ , defined in (2.4). Denote by  $B_\epsilon(x, r) = \{y \mid d_\epsilon(x, y) < r\}$  the corresponding metric balls.

For every  $n$ -tuple  $I = (i_1, \dots, i_n) \in \{1, \dots, 2p - m\}^n$ , and for  $\bar{\epsilon} \geq \epsilon \geq 0$  define the coefficient

$$\lambda_I^\epsilon(x) = \det(X_{i_1}^\epsilon(x), \dots, X_{i_n}^\epsilon(x)).$$

For a fixed  $0 \leq \epsilon \leq \bar{\epsilon}$  and for a fixed constant  $0 < C_{2,\epsilon} < 1$ , choose  $I_\epsilon = (i_{\epsilon 1}, \dots, i_{\epsilon n})$  such that

$$|\lambda_{I_\epsilon}^\epsilon(x)| r^{d_\epsilon(I_\epsilon)} \geq C_{2,\epsilon} \max_J |\lambda_J^\epsilon(x)| r^{d_\epsilon(J)}, \tag{3.1}$$

where the maximum ranges over all  $n$ -tuples. Denote  $J_\epsilon$  the family of remaining indices, so that  $\{X_{i_{\epsilon,j}}^\epsilon : i_{\epsilon,j} \in I_\epsilon\} \cup \{X_{i_{\epsilon,k}}^\epsilon : i_{\epsilon,k} \in J_\epsilon\}$  is the complete list  $X_1^\epsilon, \dots, X_{2p-m}^\epsilon$ . When  $\epsilon = 0$  we will refer to  $I_0$  as a choice corresponding to the  $n$ -tuple  $X_{i_{01}}^0, \dots, X_{i_{0n}}^0$  realizing (3.1). One of the main contributions in Nagel, Stein and Wainger’s seminal work [70], consists in the proof that for a  $v$  and a  $x$  fixed, and letting

$$Q_\epsilon(r) = \left\{ u \in \mathbb{R}^n : |u_j| \leq r^{d_\epsilon(i_{\epsilon j})} \right\}$$

denote a *weighted* cube in  $\mathbb{R}^n$ , then the quantity  $|\lambda_{I_\epsilon}^\epsilon(x)|$  provides an estimates of the Jacobian of the exponential mapping  $u \rightarrow \Phi_{\epsilon,v,x}(u)$  defined for  $u \in Q(r)$  as

$$\Phi_{\epsilon,v,x}(u) = \exp \left( \sum_{i_\epsilon,j \in I_\epsilon} u_j X_{i_\epsilon,j}^\epsilon + \sum_{i_\epsilon,k \in J_\epsilon} v_k X_{i_\epsilon,k}^\epsilon \right) (x). \tag{3.2}$$

More precisely, for  $\epsilon \geq 0$  and fixed one has

**Theorem 3.1** [70, Theorem 7] *For every  $\epsilon \geq 0$ , and  $K \subset\subset \mathbb{R}^n$  there exist  $R_\epsilon > 0$  and constants  $0 < C_{1,\epsilon}, C_{2,\epsilon} < 1$  such that for every  $x \in K$  and  $0 < r < R_\epsilon$ , if  $I_\epsilon$  is such that (3.1) holds, then*

- (i) *if  $|v_k| \leq C_{2\epsilon} r^{d(i_{\epsilon k})}$ ,  $\Phi_{\epsilon,v,x}$  is one to one on the box  $Q_\epsilon(C_{1,\epsilon}r)$*
- (ii) *if  $|v_k| \leq C_{2\epsilon} r^{d(i_{\epsilon k})}$  the Jacobian matrix of  $\Phi_{\epsilon,v,x}$  satisfies on the cube  $Q_\epsilon(C_{1,\epsilon}r)$*

$$\frac{1}{4} |\lambda_{I_\epsilon}^\epsilon(x)| \leq |J\Phi_{\epsilon,v,x}| \leq 4 |\lambda_{I_\epsilon}^\epsilon(x)|$$

(iii)  $\Phi_{\epsilon,v,x}(Q_\epsilon(C_{1,\epsilon}r)) \subset B_\epsilon(x,r) \subset \Phi_{\epsilon,v,x}(Q_\epsilon(C_{1,\epsilon}r/C_{2,\epsilon}))$

As a corollary one has that the volume of a Carnot–Caratheodory ball centered in  $x$  can be estimated by the measure of the corresponding cube and the Jacobian determinant of  $\Phi_{\epsilon,v,x}$ .

**Corollary 3.2** ([70, Theorem 1]) *For every  $\epsilon \geq 0$ , and  $K \subset\subset \mathbb{R}^n$  and for  $R_\epsilon > 0$  as in Theorem 3.1, there exist constants  $C_{3\epsilon}, C_{4\epsilon} > 0$  depending on  $K, R_\epsilon, C_{1,\epsilon}$  and  $C_{2\epsilon}$  such that for all  $x \in K$  and  $0 < r < R_\epsilon$  one has*

$$C_{3\epsilon} \sum_I |\lambda_I^\epsilon(x)| r^{d(I)} \leq |B_\epsilon(x,r)| \leq C_{4\epsilon} \sum_I |\lambda_I^\epsilon(x)| r^{d(I)}, \tag{3.3}$$

Estimates (3.3) in turn implies the doubling condition (1.2) with constants depending eventually on  $R_\epsilon, C_{1\epsilon}$  and  $C_{2\epsilon}$ .

### 3.2 Uniform estimates as $\epsilon \rightarrow 0$

Having already proved the stability of the doubling property in the special case of the Heisenberg group, in this section we turn to the general case of Hörmander’s vector fields and describe in some details results from [18] establishing that the constants  $C_{1\epsilon} C_{2\epsilon}$  do not vanish as  $\epsilon \rightarrow 0$ . Without loss of generality one may assume that both constants are non-decreasing in  $\epsilon$ . In fact, if that is not the case one may consider a new pair of constants  $\tilde{C}_{i,\epsilon} = \inf_{s \in [\epsilon, \bar{\epsilon}]} C_{i,s}$ , for  $i = 1, 2$ .

**Proposition 3.3** *For every  $\epsilon \in [0, \bar{\epsilon}]$ , the constants  $R_\epsilon, C_{1,\epsilon}$  and  $C_{2,\epsilon}$  in Theorem 3.1 may be chosen to be independent of  $\epsilon$ , depending only on the  $C^{r+1}$  norm of the vector fields, on the number  $\bar{\epsilon}$ , and on the compact  $K$ .*

*Proof* The proof is split in two cases: first we study the range  $\epsilon < r < R_0$  which roughly corresponds to the balls of radius  $r$  having a sub-Riemannian shape. In this range we show that one can select the constants  $C_{i,\epsilon}$  to be approximately  $C_{i,0}$ . The second case consists in the analysis of the range  $r < \epsilon < \bar{\epsilon}$ . In this regime the balls are roughly of Euclidean shape and we show that the constants  $C_{i,\epsilon}$  can be approximately chosen to be  $C_{i,\bar{\epsilon}}$ .

Let us fix  $\epsilon \in (0, \bar{\epsilon}]$ ,  $R = R_0$  and  $r < R_0$ . We can start by describing the family  $I_\epsilon$  defined in (3.1), which maximize  $\lambda_j^\epsilon(x)$ . We first note that for every  $\epsilon > 0$  and for every  $i$ ,  $m + 1 \leq i \leq p$  we have

$$X_i^\epsilon r^{d_\epsilon(i)} = \epsilon^{d(i)-1} r Y_i, \quad X_{i+p-m}^\epsilon r^{d_\epsilon(i+p-m)} = r^{d(i)} Y_i. \tag{3.4}$$

In the range  $0 < r < \epsilon < \bar{\epsilon}$  one can assume without loss of generality that the  $n$ -tuple satisfying the maximality condition (3.1) will include only vectors of the form  $\{\epsilon^{d(i_{\epsilon 1})-1} Y_{i_{\epsilon 1}}, \dots, \epsilon^{d(i_{\epsilon n})-1} Y_{i_{\epsilon n}}\}$  for some  $n$ -index  $I_\epsilon = (i_{\epsilon 1}, \dots, i_{\epsilon n})$ , with  $1 \leq i_{\epsilon k} \leq p$ . In fact, if this were not the case and the  $n$ -tuple were to include a vector of the form  $X_j^\epsilon = Y_{j-p+m}$  for some  $p < j$ , then we could substitute such vector with  $X_{j-p+m}^\epsilon = Y_{j-p+m} \epsilon^{d(j-p+m)-1}$  and from (3.4) infer that the value of the corresponding term  $|\lambda_{I_\epsilon}^\epsilon(x)| r^{d_\epsilon(I_\epsilon)}$  would increase.

Similarly, in the range  $0 < \epsilon < r < \bar{\epsilon}$  one can assume that the  $n$ -tuple satisfying the maximality condition (3.1) will include only vectors of the form  $\{Y_{i_{\epsilon 1}}, \dots, Y_{i_{\epsilon n}}\}$  for some  $n$ -index  $I_\epsilon = (i_{\epsilon 1}, \dots, i_{\epsilon n})$ , with  $1 \leq i_{\epsilon k} \leq p$ . Note that the corresponding expression

$$|\lambda_{I_\epsilon}^\epsilon(x)| r^{d_\epsilon(I_\epsilon)-1} = |\det(Y_{i_{\epsilon 1}}, \dots, Y_{i_{\epsilon n}})(x)| r^{\sum_{I_\epsilon} d(i_{\epsilon k})}$$

would then be one of the terms in the left hand side of (3.1) for  $\epsilon = 0$ , and thus is maximized by  $C_{2,0}^{-1} |\lambda_{I_0}^0(x)| r^{d(I_0)-1}$ .

**Case 1:** In view of the argument above, for every  $\epsilon < r < R_0$  the indices  $I_\epsilon$  defined by the maximality condition (3.1) can be chosen to coincide with indices of the family  $I_0$  and do not depend on  $\epsilon$ . On the other hand the vector excluded from  $I_\epsilon$  will be not only those in  $J_0$  but also the ones that have been added with a weight factor of a power of  $\epsilon$ ,

$$\begin{aligned} \{X_k^\epsilon : k \in J_\epsilon\} &= \left\{ X_{i_{0,k}}^0 : i_{0,k} \in J_0 \right\} \cup \left\{ \epsilon^{d(i_{0,k})-1} X_{i_{0,k}}^0 : i_{0,k} \in I_0, i_{0,k} > m \right\} \\ &\cup \left\{ \epsilon^{d(i_{0,k})-1} X_{i_{0,k}}^0 : i_{0,k} \in J_0, i_{0,k} > m \right\}. \end{aligned}$$

In correspondence with this decomposition of the set of indices we define a splitting in the  $v$ -variables in (5.14) as

$$v = (\hat{v}, \tilde{v}, \bar{v}).$$

Consequently for every  $\epsilon < r$  the function  $\Phi_{\epsilon, v, x}(u)$  can be written as

$$\begin{aligned}
 \Phi_{\epsilon, v, x}(u) &= \exp \left( \sum_{i_{\epsilon j} \in I_{\epsilon}} u_j X_{i_{\epsilon j}}^{\epsilon} + \sum_{i_{\epsilon k} \in J_{\epsilon}} v_k X_{i_{\epsilon k}}^{\epsilon} \right) (x) \\
 &= \exp \left( \sum_{i_{0j} \in I_0} u_j X_{i_{0j}}^0 + \sum_{i_{\epsilon k} \in J_{\epsilon}} v_k X_{i_{\epsilon k}}^{\epsilon} \right) (x) \\
 &= \exp \left( \sum_{i_{0j} \in I_0} u_j X_{i_{0j}}^0 + \sum_{i_{0k} \in J_0} \hat{v}_k Y_{i_{0k}} + \sum_{i_{0k} \in I_0, i > m} \tilde{v}_k \epsilon^{d(i_{0k})-1} Y_{i_{0k}} \right. \\
 &\quad \left. + \sum_{i_{0k} \in J_0, i_{0, k} > m} \bar{v}_k \epsilon^{d(i_{0k})-1} Y_{i_{0k}} \right) (x) \\
 &= \Phi_{0, \hat{v}_k + \bar{v}_k \epsilon^{d(i_{0k})-1}, x} \left( u_1, \dots, u_m, u_{m+1} + \tilde{v}_{m+1} \epsilon^{d(i_{0m+1})-1}, \dots, u_n \right. \\
 &\quad \left. + \tilde{v}_n \epsilon^{d(i_{0n})-1} \right). \tag{3.5}
 \end{aligned}$$

Let us define mappings

$$F_{1, \epsilon, v}(u) = \left( u_1, \dots, u_m, u_{m+1} + \tilde{v}_{m+1} \epsilon^{d(i_{0m+1})-1}, \dots, u_n + \tilde{v}_n \epsilon^{d(i_{0n})-1} \right),$$

and

$$F_{2, \epsilon}(v) = \left( \hat{v}_1 + \bar{v}_1 \epsilon^{d(i_{01})-1}, \dots, \hat{v}_{2p-m} + \bar{v}_{2p-m} \epsilon^{d(i_{0, 2p-m})-1} \right).$$

In view of (3.5) we can write

$$\Phi_{\epsilon, v, x}(u) = \Phi_{0, F_{2, \epsilon}(v), x}(F_{1, \epsilon, v}(u)). \tag{3.6}$$

Note that for any  $\epsilon \geq 0$  and for a fixed  $v$ , the mapping  $u \rightarrow F_{1, \epsilon, v}(u)$  is invertible and volume preserving in all  $\mathbb{R}^n$ . Moreover  $J\Phi_{\epsilon, v, x}(u) = J\Phi_{0, F_{2, \epsilon}(v), x}(F_{1, \epsilon, v}(u))$ . In view of (3.6) and of Theorem 3.1, as a function of  $u$ , the mapping  $\Phi_{\epsilon, v, x}(u)$  is defined, invertible, and satisfies the Jacobian estimates in Theorem 3.1 (ii)

$$\frac{1}{4} \left| \lambda_{I_0}^0(x) \right| \leq \left| J\Phi_{0, F_{2, \epsilon}(v), x}(F_{1, \epsilon, v}(u)) \right| = \left| J\Phi_{\epsilon, v, x}(u) \right| \leq 4 \left| \lambda_{I_0}^0(x) \right|$$

for all  $u$  such that  $F_{1, \epsilon, v}(u) \in Q_0(C_{1,0}r)$  and for  $v$  such that

$$\begin{aligned}
 \left| F_{2, \epsilon}^k(v) \right| &= \left| \hat{v}_k + \bar{v}_k \epsilon^{d(i_{0k})-1} \right| \leq C_{2,0} r^{d(i_{0k})}, \\
 |u_1| \leq C_{1,0} r^{d(i_{01})} \dots |u_m| &\leq C_{1,0} r^{d(i_{0m})}, \quad \left| u_{m+1} + \tilde{v}_{m+1} \epsilon^{d(i_{0m+1})-1} \right| \leq C_{1,0} r^{d(i_{0m+1})},
 \end{aligned}$$

when  $k = 1, \dots, 2p - m$ . □

The completion of the proof of Case 1 rests on the following two claims:

**Claim 1** let  $\epsilon < r < R_0$ . There exists  $C_6 > 0$ , independent of  $\epsilon$ , such that for all  $v$  satisfying  $|v_k| \leq C_6 r^{d(i_{\epsilon k})}$  one has  $|F_{2,\epsilon}^k(v)| = |\hat{v}_k + \bar{v}_k \epsilon^{d(i_{0k})-1}| \leq C_{2,0} r^{d(i_{0k})}$ .

*Proof of the claim* If we choose  $C_6 < \min\{C_{1,0}, C_{2,0}\}$  and

$$|\hat{v}_k|, |\tilde{v}_k|, |\bar{v}_k| \leq \min\{C_{1,0}, C_{2,0}\} \frac{r^{d(i_{\epsilon k})}}{4}, \quad |u_j| \leq C_{1,0} \frac{r^{d(i_{\epsilon j})}}{4},$$

it follows that

$$|\hat{v}_k| \leq C_{2,0} \frac{r^{d(i_{0k})}}{4}, \quad |\tilde{v}_k|, |\bar{v}_k| \leq C_{1,0} \frac{r}{4}, \quad |u_j| \leq C_{1,0} \frac{r^{d(i_{\epsilon j})}}{4}.$$

So that

$$|\hat{v}_k| \leq C_{2,0} \frac{r^{d(i_{0k})}}{4}, \quad \epsilon^{d(i_{0k})-1} |\tilde{v}_k|, \quad \epsilon^{d(i_{0k})-1} |\bar{v}_k| \leq C_{1,0} \frac{r^{d(i_{0k})}}{4}, \quad |u_j| \leq C_{1,0} \frac{r^{d(i_{0j})}}{4},$$

completing the proof of the claim. □

**Claim 2** Let  $\epsilon < r < R_0$  and  $v$  fixed such that  $|v_k| \leq C_6 r^{d(i_{\epsilon k})}$  for  $k = 1, \dots, 2p - m$ . One has that

$$Q_\epsilon \left( C_5^{-1} r \right) \subset F_{1,\epsilon,v}^{-1} \left( Q_0(C_{1,0} r) \right) \subset Q_\epsilon(C_5 r)$$

for some constant  $C_5 > 0$  independent of  $\epsilon \geq 0$ .

*Proof of the claim* Choose  $C_5$  sufficiently large so that  $2 \max\{C_5^{-1}, C_6\} \leq C_{1,0}$  and observe that if  $u \in Q_\epsilon(C_5^{-1} r)$  then for  $k = 1, \dots, m$  we have  $|u_k| \leq C_{1,0} r^{d(i_{\epsilon,k})} = C_{1,0} r^{d(i_{0,k})}$  while for  $k = m + 1, \dots, n$  we have  $|F_{1,\epsilon,v}^k(u)| = |u_k + \tilde{v}_k \epsilon^{d(i_{0k})-1}| \leq \max\{C_5^{-1}, C_6\} r^{d(i_{0k})} (1 + \bar{\epsilon}^{d(i_{0k})-1}) \leq C_{1,0} r^{d(i_{0k})}$ . This proves the first inclusion in the claim. To establish the second inclusion we choose  $C_5$  large enough so that  $2(C_{1,0} + C_{2,\bar{\epsilon}}) \leq C_5$  and observe that if  $F_{1,\epsilon,v}(u) \in Q_0(C_{1,0} r)$  then for  $k = m + 1, \dots, n$  one has  $|u_k| \leq |u_k + \tilde{v}_k \epsilon^{d(i_{0k})-1}| + |\tilde{v}_k| \epsilon^{d(i_{0k})-1} \leq 2(C_{1,0} + C_{2,\bar{\epsilon}}) r^{d(i_{0k})} \leq C_5 r^{d(i_{0k})}$ . The corresponding estimate for the range  $k = 1, \dots, m$  is immediate.

In view of Claims 1 and 2, and of Theorem 3.1 It follows that for  $\epsilon < r$  and these choices of constants (independent of  $\epsilon$ )<sup>1</sup> the function  $\Phi_{\epsilon,v,x}(u)$  is invertible on  $Q_0(C_{1,0} r)$  and (i), (ii) and (iii) are satisfied.

Case 2: As remarked above, in the range  $0 < r < \epsilon < \bar{\epsilon}$  one can assume that the  $n$ -tuple satisfying the maximality condition (3.1) will include only vectors of the form  $\{\epsilon^{d(i_{\epsilon 1})-1} Y_{i_{\epsilon 1}}, \dots, \epsilon^{d(i_{\epsilon n})-1} Y_{i_{\epsilon n}}\}$  for some  $n$ -index  $I_\epsilon = (i_{\epsilon 1}, \dots, i_{\epsilon n})$ , with

---

<sup>1</sup>  $R_0$  in place of  $R_\epsilon$ ,  $C_5$  in place of  $C_{1,\epsilon}$  and  $C_6$  in place of  $C_{2,\epsilon}$ .

$1 \leq i_{\epsilon k} \leq p$ . Note that in view of (3.4) and the maximality condition (3.1) the corresponding term

$$|\lambda_{I_\epsilon}^\epsilon(x)| r^{d_\epsilon(I_\epsilon)}$$

can be rewritten and estimated as follows

$$|\lambda_{I_\epsilon}^\epsilon(x)| r^{d_\epsilon(I_\epsilon)} = \epsilon^{d(I_\epsilon)-n} r^n |\det(Y_{i_{\epsilon 1}}, \dots, Y_{i_{\epsilon n}})(x)|.$$

It is then clear that the maximizing  $n$ -tuple  $I_\epsilon$  in (3.1) will be identified by the lowest degree  $d(I_\epsilon)$  among all  $n$ -tuples corresponding to non-vanishing determinants  $\det(Y_{i_{\epsilon 1}}, \dots, Y_{i_{\epsilon n}})$  in a neighborhood of the point  $x$ . Since this choice does not depend on  $\epsilon > r$ , then one has that  $I_\epsilon = I_{\bar{\epsilon}}$ . In other words, if we denote

$$(X^{\bar{\epsilon}})_{i_{\bar{\epsilon},k} \in I_{\bar{\epsilon}}} = \left\{ \bar{\epsilon}^{d(i_{\bar{\epsilon},1})-1} Y_{i_{\bar{\epsilon},1}}, \dots, \bar{\epsilon}^{d(i_{\bar{\epsilon},n})-1} Y_{i_{\bar{\epsilon},n}} \right\}$$

then the maximality condition (3.1) in the range  $0 < r < \epsilon < \bar{\epsilon}$  can be satisfied independently from  $\epsilon$  by selecting the family of vector fields:

$$(X^\epsilon)_{i_{\epsilon,k} \in I_\epsilon} = \left\{ \epsilon^{d(i_{\bar{\epsilon},1})-1} Y_{i_{\bar{\epsilon},1}}, \dots, \epsilon^{d(i_{\bar{\epsilon},n})-1} Y_{i_{\bar{\epsilon},n}} \right\}$$

The complementary family  $J_\epsilon$  becomes

$$\begin{aligned} \left\{ Y_{i_{\epsilon k}}^\epsilon : i_{\epsilon k} \in J_\epsilon \right\} &= \left\{ \epsilon^{d(i_{\bar{\epsilon},k})-1} Y_{i_{\bar{\epsilon},k}} : i_{0,k} \in J_{\bar{\epsilon}}, \text{ with } i_{\bar{\epsilon},k} \leq p \right\} \\ &\cup \left\{ Y_{i_{\bar{\epsilon},k-p+m}} : i_{\bar{\epsilon},k} \in J_{\bar{\epsilon}}, \text{ with } i_{\bar{\epsilon},k} > p \right\} \end{aligned} \tag{3.7}$$

If we denote  $A_\epsilon$ , and  $B_\epsilon$  these three sets, and split the  $v$ -variable from (5.14) as  $v = (\hat{v}, \tilde{v})$ , then it is clear that

$$Y \in A_\epsilon \text{ iff } \frac{\bar{\epsilon}^{d(i_{\bar{\epsilon},k})-1}}{\epsilon^{d(i_{\bar{\epsilon},k})-1}} Y \in A_{\bar{\epsilon}},$$

and in this case the values of  $d_\epsilon$  and  $d_{\bar{\epsilon}}$  are the same on the corresponding indices. Analogously  $Y \in B_\epsilon$  iff  $Y \in B_{\bar{\epsilon}}$  and the degrees are the same.

For every  $\epsilon > r$  the map  $\Phi_{\epsilon, v, x}(u)$  then can be written as

$$\begin{aligned} \Phi_{\epsilon, v, x}(u) &= \exp \left( \sum_{i_{\epsilon j} \in I_\epsilon} u_j X_{i_{\epsilon j}}^\epsilon + \sum_{i_{\epsilon k} \in J_\epsilon} v_k X_{i_{\epsilon k}}^\epsilon \right) (x) \\ &= \exp \left( \sum_{i_{\bar{\epsilon} j} \in I_{\bar{\epsilon}}} u_j X_{i_{\bar{\epsilon} j}}^\epsilon + \sum_{i_{\bar{\epsilon} k} \in J_{\bar{\epsilon}}} v_k X_{i_{\bar{\epsilon} k}}^\epsilon \right) (x) \end{aligned}$$

$$\begin{aligned}
 &= \exp \left( \sum_{i_{\bar{\epsilon}j} \in I_{\bar{\epsilon}}} u_j \frac{\epsilon^{d(i_{\bar{\epsilon},k})-1}}{\bar{\epsilon}^{d(i_{\bar{\epsilon},k})-1}} X_{i_{\bar{\epsilon}j}}^{\bar{\epsilon}} + \sum_{i_{\bar{\epsilon}k} \in J_{\bar{\epsilon}} \text{ and } i_{\bar{\epsilon}j} \leq p} \hat{v}_k \frac{\epsilon^{d(i_{0,k})-1}}{\bar{\epsilon}^{d(i_{0,k})-1}} X_{i_{0k}}^{\bar{\epsilon}} \right. \\
 &\quad \left. + \sum_{i_{\bar{\epsilon}k} \in J_{\bar{\epsilon}} \text{ and } i_{\bar{\epsilon}j} > p} \tilde{v}_k X_{i_{\bar{\epsilon}k}}^{\bar{\epsilon}} \right) (x)
 \end{aligned}$$

This function is defined and invertible for

$$|\tilde{v}_k|, |\hat{v}_k| \frac{\epsilon^{d(i_{\bar{\epsilon},k})-1}}{\bar{\epsilon}^{d(i_{\bar{\epsilon},k})-1}} \leq C_{2,\bar{\epsilon}} r^{d_{\bar{\epsilon}}(i_{\bar{\epsilon}k})}, |u_j| \frac{\epsilon^{d(i_{0,j})-1}}{\bar{\epsilon}^{d(i_{\bar{\epsilon},j})-1}} \leq C_{1,\bar{\epsilon}} r^{d_{\bar{\epsilon}}(i_{\bar{\epsilon}j})}.$$

Recall that with the present choice of  $r < \epsilon < \bar{\epsilon}$ , we have  $C_{1,\bar{\epsilon}} r^{d_{\bar{\epsilon}}(i_{\bar{\epsilon}j})} = C_{1,\bar{\epsilon}} r^{d_{\epsilon}(i_{\bar{\epsilon}j})} = C_{1,\bar{\epsilon}} r^{d_{\epsilon}(i_{\epsilon j})}$ . If we set

$$\begin{aligned}
 |\hat{v}_k|, |\tilde{v}_k| &\leq C_{2,\bar{\epsilon}} r^{d_{\bar{\epsilon}}(i_{\bar{\epsilon}k})}, \\
 |u_j| &\leq C_{1,\bar{\epsilon}} r^{d_{\bar{\epsilon}}(i_{\bar{\epsilon}j})},
 \end{aligned}$$

and argue similarly to Case 1, then the function  $\Phi_{\epsilon,v,x}$  will satisfy conditions i), ii), and iii) on  $Q(C_{1,\bar{\epsilon}}r)$  and hence on  $Q(C_{1,\epsilon}r)$ , with constants independent of  $\epsilon$ .  $\square$

### 3.3 Equiregular subRiemannian structures and equivalent pseudo-distances

The intrinsic definition, based on a minimizing choice, of the Carnot–Caratheodory metric is not convenient when one needs to produce quantitative estimates, as we will do in the following sections. It is then advantageous to use equivalent pseudo-distances which are explicitly defined in terms of certain system of coordinates. In the last section we have already encountered two special cases, i.e. the norms  $|\cdot|$  defined in (2.5) and its Riemannian approximation (2.8). In this section we extend this construction to a all equi-regular subRiemannian structures. For  $\Omega \subset \mathbb{R}^n$  consider the subRiemannian manifold  $(\Omega, \Delta, g)$  and iteratively set  $\Delta^1 := \Delta$ , and  $\Delta^{i+1} = \Delta^i + [\Delta^i, \Delta]$  for  $i \in \mathbb{N}$ . The bracket generating condition is expressed by saying that there exists an integer  $s \in \mathbb{N}$  such that  $\Delta_p^s = \mathbb{R}^n$  for all  $p \in M$ .

**Definition 3.4** A subRiemannian manifold  $(\Omega, \Delta, g)$  is *equiregular* if, for all  $i \in \mathbb{N}$ , the dimension of  $\Delta_p^i$  is constant in  $p \in \Omega$ . The *homogenous dimension*

$$Q = \sum_{i=1}^{s-1} \left[ \dim(\Delta_p^{i+1}) - \dim(\Delta_p^i) \right], \tag{3.8}$$

coincides with the Hausdorff dimension with respect to the Carnot–Caratheodory distance.

This class is *generic* as any subRiemannian manifold has a dense open subset on which the restriction of the subRiemannian metric is equiregular.

*Example 3.5* Systems of free vector fields, as defined in Definition 5.4, yield a distribution  $\Delta$  that supports an equiregular subRiemannian structure for any choice of the horizontal metric  $g$ .

*Example 3.6* An analytic Lie group  $G$  is called a *homogenous stratified Lie group* if its Lie algebra admits a stratification  $\mathcal{G} = V^1 \oplus \dots \oplus V^r$  with  $[V^i, V^j] = V^{i+j}$  and  $[V^i, V^r] = 0$ . Given a positive definite bilinear form  $g_0$  on  $V^1$  we call the pair  $(G, g_0)$  a *Carnot group* and the corresponding left invariant metric  $g_0$  is a equiregular sub-Riemannian metric.

Next we assume we have a equiregular subRiemannian manifold  $(\Omega, \Delta, g)$  and consider an orthonormal horizontal basis  $X_1, \dots, X_m$  of  $\Delta$ . Following the process in (2.2) one can construct a frame  $Y_1, \dots, Y_n$  for  $\mathbb{R}^n$  where  $Y_1, \dots, Y_m$  is the original horizontal frame and  $Y_{m+1}, \dots, Y_n$  are commutators such that  $(Y_1, \dots, Y_{m_k})|_p$  spans  $\Delta_p^k$ , for  $k = 1, \dots, s$ . The degree  $d(i)$  of  $Y_i$  is the order of commutators needed to generate  $Y_i$  out of the horizontal span, i.e.  $d(i) = k$  if  $Y_i \in \Delta_p^k$  but  $Y_i \notin \Delta_p^{k-1}$ . In particular one has  $d(i) = 1$  for  $i = 1, \dots, m$ . The equiregularity hypothesis allows one to choose  $Y_1, \dots, Y_n$  linearly independent. Next we extend  $g$  to a Riemannian metric  $g_1$  on all of  $T\Omega$  by imposing that  $Y_1, \dots, Y_n$  is an orthonormal basis.

**Definition 3.7** For any  $\epsilon \in (0, \bar{\epsilon}]$  we define the Riemannian metric  $g_\epsilon$  by setting that  $\{\epsilon^{d(i)-1}Y_i, i = 1, \dots, n\}$  is an orthonormal frame. Denote by  $d_\epsilon(x, y)$  the corresponding Riemannian distance function.

*Remark 3.8* Repeating the proof of [70, Theorem 4] one immediately sees that  $d_\epsilon$  as defined here is comparable to the distance  $d_\epsilon$  defined in Sect. 2.2, with equivalence constants independent of  $\epsilon > 0$ .

We define canonical coordinates around a point  $x_0 \in \Omega$  as follows. Since  $Y_1, \dots, Y_n$  is a generating frame for  $T\Omega$  then for any point  $x$  in a neighborhood  $\omega$  of  $x_0$  one has that there exists a unique  $n$ -tuple  $(x_1, \dots, x_n)$  such that

$$\exp\left(\sum_{i=1}^n x_i Y_i\right)(x_0) = x. \tag{3.9}$$

We will set  $x = (x_1, \dots, x_n)$  and use this  $n$ -tuple as local coordinates in  $\omega$ .

**Definition 3.9** For every  $x = (x_1, \dots, x_n) \in \omega$  we define a pseudo-distance  $d_{G,\epsilon}(x, x_0) := N_\epsilon(x_1, \dots, x_n)$  with

$$N_\epsilon(x_1, \dots, x_n) := \sqrt{\sum_{i=1}^m x_i^2 + \sum_{i=m+1}^n \min\left(\epsilon^{-(d(i)-1)}|x_i|, |x_i|^{1/d(i)}\right)}. \tag{3.10}$$

For  $\epsilon = 0$  we set

$$N_0(x_1, \dots, x_n) := \sqrt{\sum_{i=1}^m x_i^2 + \sum_{i=m+1}^n |x_i|^{1/d(i)}}.$$



**Theorem 3.10** *For every compact  $x_0 \in K \subset \omega$  there exists  $C = C(K, \Delta, g, \omega) > 0$ , independent of  $\epsilon \in (0, \bar{\epsilon}]$ , such that*

$$C^{-1}d_{G,\epsilon}(x, x_0) \leq d_\epsilon(x, x_0) \leq Cd_{G,\epsilon}(x, x_0)$$

for all  $x \in K$ .

*Remark 3.11* Note that for  $\epsilon = 0$  the equivalence is a direct consequence of the *Ball-Box theorem* proved by Nagel et al. [70] or Mitchell [65, Lemma 3.4]. This observation replaces the estimates (2.6) from the Heisenberg group setting.

The proof of Theorem 3.10 follows as a corollary of the following

**Proposition 3.12** *In the hypothesis of Theorem 3.10 one has that there exists  $R = R(K, \Delta, g, \omega) > 0$ ,  $C = C(K, \Delta, g, \omega) > 0$ , independent of  $\epsilon \in (0, \bar{\epsilon}]$ , such that for all  $x \in K$  and  $r \in (0, R)$ ,*

$$B_{G,\epsilon}(x_0, C^{-1}r) \subset B_\epsilon(x_0, r) \subset B_{G,\epsilon}(x_0, Cr),$$

where

$$B_{G,\epsilon}(x_0, r) := \left\{ x \in \mathbb{R}^n \text{ such that } \max_{i=1,\dots,s} \left[ \min \left( \epsilon^{-d(i)-1} |x_i|, |x_i|^{1/d(i)} \right) \right] < r \right\}.$$

*Proof* The proof follows closely the arguments in the previous section and is based on the results in [70]. In view of the equiregularity hypothesis note that  $Y_1, \dots, Y_n$  are linearly independent and the construction in (2.4) yields the distribution  $X_1^\epsilon, \dots, X_{2n-m}^\epsilon$  over  $\Omega$ . Recall from (5.14), Proposition 3.3 and Theorem 3.1 that if  $I_\epsilon, J_\epsilon$  are chosen as in (3.1) and for any  $v = (v_1, \dots, v_{n-m})$  such that  $|v_k| \leq C_{2\epsilon} r^{d(i_{\epsilon k})}$ , one has

$$B_\epsilon(x_0, r) \approx \Phi_{\epsilon, v, x_0}(Q_\epsilon(r)), \tag{3.11}$$

with constants independent from  $\epsilon \geq 0$ , where  $Q_\epsilon = \{u \in \mathbb{R}^n : |u_j| \leq r^{d_\epsilon(i_{\epsilon j})}\}$ , and

$$\Phi_{\epsilon, v, x}(u) = \exp \left( \sum_{i_{\epsilon, j} \in I_\epsilon} u_j X_{i_{\epsilon, j}}^\epsilon + \sum_{i_{\epsilon, k} \in J_\epsilon} v_k X_{i_{\epsilon, k}}^\epsilon \right) (x).$$

The  $n$ -tuple  $I_\epsilon$  contains  $n$  indexes related either to the horizontal vector fields  $X_1^\epsilon, \dots, X_m^\epsilon$  or to the commutators  $X_{m+1}^\epsilon, \dots, X_n^\epsilon$ . The latter may consist of *weighted* versions  $X_{m+1}^\epsilon, \dots, X_n^\epsilon$  or *unweighted versions*  $X_{n+1}^\epsilon, \dots, X_{2n-m}^\epsilon$ . In either case the same vector will appear both in the weighted and in the unweighted version (either among the  $I_\epsilon$  indexes or in the complement  $J_\epsilon$ ). Comparing the representation  $\Phi_{\epsilon, v, x_0}$  with the  $x$ -coordinates representation (3.9) one has

$$\exp \left( \sum_{i=1}^n x_i Y_i \right) (x_0) = \exp \left( \sum_{i_{\epsilon, j} \in I_\epsilon} u_j X_{i_{\epsilon, j}}^\epsilon + \sum_{i_{\epsilon, k} \in J_\epsilon} v_k X_{i_{\epsilon, k}}^\epsilon \right) (x_0),$$

and we let for each  $k = 1, \dots, n$

$$x_k = \begin{cases} \epsilon^{d(k)-1} u_{i_k} + v_{j_k} & \text{if } i_k \leq n \\ u_{i_k} + \epsilon^{d(k)-1} v_{j_k} & \text{if } i_k > n \end{cases}.$$

From the latter we obtain that for all  $k = 1, \dots, n$

$$|x_k| \leq C \left( \epsilon^{d(k)-1} r + r^{d(k)} \right).$$

If  $x \in B_\epsilon(x_0, r)$  then  $|u_{i_k}|, |v_{j_k}| \leq Cr^{d(k)}$ . Consequently,

$$\begin{aligned} & \min(\epsilon^{-(d(k)-1)} |x_k|, |x_k|^{1/d(k)}) \\ & \leq C \min \left( \epsilon^{-(d(k)-1)} |\epsilon^{d(k)-1} r + r^{d(k)}|, \left[ \epsilon^{d(k)-1} r + r^{d(k)} \right]^{1/d(k)} \right) \\ & \leq C \min \left( r \left[ 1 + \left( \frac{r}{\epsilon} \right)^{d(k-1)} \right], r \left[ \left( \frac{\epsilon}{r} \right)^{d(k)-1} + 1 \right]^{1/d(k)} \right) \leq 2Cr. \end{aligned}$$

This shows that for  $r > 0$  sufficiently small, and for some choice of  $C > 0$  independent of  $\epsilon \geq 0$ , we have  $B_\epsilon(x_0, r) \subset B_{G,\epsilon}(x_0, Cr)$ .

To prove the reverse inclusion we consider a point  $x = \exp(\sum_{i=1}^n x_i Y_i)(x_0) \in B_{G,\epsilon}(x_0, Cr)$ . Select  $I_\epsilon$  as in (3.1) and set  $v = 0$  to represent  $x$  in the basis  $X_{i_1}, \dots, X_{i_n}$  as

$$x = \exp \left( \sum_{i_\epsilon, j \in I_\epsilon} u_j X_{i_\epsilon, j}^\epsilon \right) (x_0).$$

In view of Theorem 3.1, and (3.11), to prove the proposition it suffices to show that there exists a constant  $C > 0$  independent of  $\epsilon > 0$  such that for each  $j = 1, \dots, n$  one has  $|u_j| \leq Cr^{d_\epsilon(i_\epsilon j)}$ .

We distinguish two cases: In the range  $\epsilon \geq 2r$  one can argue as in (3.4) to deduce that for each  $j = 1, \dots, n$  we may assume without loss of generality that the contribution due to  $u_j X_{i_\epsilon, j}^\epsilon$  follows from the choice of a weighted vector, and hence is of the form  $u_j \epsilon^{d(k)-1} Y_k$  for some  $k > m$ . Consequently one has  $d_\epsilon(i_\epsilon, j) = 1$  and  $x_k = u_j \epsilon^{d(k)-1}$ .

On the other hand, since  $\epsilon \geq 2r$  then one must also have that

$$\min \left( \epsilon^{-(d(k)-1)} |x_k|, |x_k|^{1/d(k)} \right) = \epsilon^{-(d(k)-1)} |x_k| < r.$$

Consequently one has

$$|u_j| = |x_k| \epsilon^{1-d(k)} \leq r = r^{d_\epsilon(i_j)}.$$

In the range  $\epsilon < 2r$  we observe that one must have  $|x_k| \leq Cr^d(k)$ . Arguing as in (3.4) we see that without loss of generality, or each  $j = 1, \dots, n$ , the contribution due to  $u_j X_{i_{\epsilon,j}}^\epsilon$  follows from the choice of a un-weighted vector, and hence is of the form  $u_j Y_k$  for some  $k > m$ . Consequently one has  $d_\epsilon(i_j) = d(k) > 1$  and  $x_k = u_j$ , concluding the proof.  $\square$

### 4 Stability of the Poincaré inequality

In this section we will focus on the Poincaré inequality and prove that it holds with a choice of a constant which is stable as  $\epsilon \rightarrow 0$ . Our argument rests on results of Lanconelli and Morbidelli [59] whose proof, in some respects, simplifies the method used by Jerison in [52]. Using some Jacobian estimates from [44] or [40] we will establish that the assumptions required in the key result [59, Theorem 2.1] are satisfied independently from  $\epsilon \geq 0$ . We start by recalling

**Theorem 4.1** [59, Theorem 2.1] *Assume that the doubling condition (D) is satisfied and there exist a sphere  $B_\epsilon(x_0, r)$ , a cube  $Q_\epsilon \subset \mathbb{R}^n$  and a map  $E : B_\epsilon(x_0, r) \times Q_\epsilon \rightarrow \mathbb{R}^n$  satisfying the following conditions:*

- (i)  $B_\epsilon(x_0, 2r) \subset E(x, Q_\epsilon)$  for every  $x \in B_\epsilon(x_0, r)$
- (ii) the function  $u \mapsto E(x, u)$  is one to one on the box  $Q_\epsilon$  as a function of the variable  $u$  and there exists a constant  $\alpha_1 > 0$  such that

$$\frac{1}{\alpha_1} |JE(x, 0)| \leq |JE(x, u)| \leq \alpha_1 |JE(x, 0)| \quad \text{for every } u \in Q_\epsilon$$

Also assume that there exists a positive constant  $\alpha_2$ , and a function  $\gamma : B_\epsilon(x_0, r) \times Q_\epsilon \times [0, \alpha_2 r] \rightarrow \mathbb{R}^n$  satisfying the following conditions

- (iii) For every  $(x, u) \in B_\epsilon(x_0, r) \times Q_\epsilon$  the function  $t \mapsto \gamma(x, u, t)$  is a subunit path connecting  $x$  and  $E(x, u)$
- (iv) For every  $(h, t) \in B_\epsilon(x_0, r) \times Q_\epsilon$  the function  $x \mapsto \gamma(x, u, t)$  is a one-to-one map and there exists a constant  $\alpha_3 > 0$  such that

$$\inf_{B_\epsilon(x_0, r) \times Q_\epsilon} \left| \det \frac{\partial \gamma}{\partial x} \right| \geq \alpha_3$$

Then there exists a constant  $C_P$  depending only on the constants  $\alpha_1, \alpha_2, \alpha_3$  and the doubling constant  $C_D$  such that (P) is satisfied.

We are now ready to prove Theorem 1.2

*Proof* All one needs to establish is that the assumptions of Theorem 4.1 are satisfied uniformly in  $\epsilon$  on a metric ball. Apply Proposition 3.3 and Theorem 3.1 with  $K = B_\epsilon(x_0, r)$  and choose the constants  $C_i$  produced by these results. Set  $Q_\epsilon = Q_\epsilon(\frac{3C_1}{C_2}r)$  and let

$$E(x, u) = \Phi_{\epsilon, 0, x}(u), \text{ defined on } K \times Q_\epsilon \rightarrow \mathbb{R}^n.$$

To establish assumption (i) of Theorem 4.1 it suffices to note that by virtue of condition (iii) in Theorem 3.1 one has that for  $x \in B_\epsilon(x_0, r)$ ,

$$B_\epsilon(x_0, 2r) \subset B_\epsilon(x, 3r) \subset E(x, Q_\epsilon).$$

Assumption (ii) in Theorem 4.1 is a direct consequence of condition (ii) in Theorem 3.1, with  $\alpha_1 = 16$ . Chow’s connectivity theorem implies that  $E(x, u)$  satisfies assumption (iii), with a function  $\gamma$ , piecewise expressed as exponential mappings of vector fields of  $\epsilon$ -degree one. Let us denote  $(X_i^\epsilon)_{i \in I_\epsilon}$  the required vector fields. With this choice of path, it is known (see for example [44, Lemma 2.2] or [40, pp 99–101]) that  $x \rightarrow \gamma(x, u, t)$  is a  $C^1$  path, with Jacobian determinant

$$\left| \det \frac{\partial \gamma}{\partial x}(x, u, t) \right| = 1 + \psi(x, u, t),$$

for a suitable function  $\psi(x, u, t)$  satisfying

$$|\psi(x, u, t)| \leq cr, \text{ on } K \times Q_\epsilon \times [0, cr].$$

Since the constant  $c$  depends solely on the Lipschitz constant of the vector fields  $(X_i^\epsilon)_{i \in I_\epsilon}$  then it can be chosen independently of  $\epsilon$ . As a consequence condition (iv) is satisfied and the proof is concluded. □

## 5 Stability of heat Kernel estimates

### 5.1 Hörmander type parabolic operators in non divergence form

The results in this section concern uniform Gaussian estimates for the heat kernel of certain degenerate parabolic differential equations, and their parabolic regularizations. We will consider a collection of smooth vector fields  $X = (X_1, \dots, X_m)$  satisfying Hörmander’s finite rank condition (1.1) in an open set  $\Omega \subset \mathbb{R}^n$ . We will use throughout the section the definition of degree  $d(i)$  relative to the stratification (2.2).

A second order, non-divergence form, ultra-parabolic operator with constant coefficients  $a_{ij}$  can be expressed as:

$$L_A = \partial_t - \sum_{i,j=1}^m a_{ij} X_i X_j, \tag{5.1}$$

where  $A = (a_{ij})_{i,j=1,\dots,m}$  is a symmetric, real-valued, positive definite  $m \times m$  matrix satisfying

$$\Lambda^{-1} \sum_{d(i)=1} \xi_i^2 \leq \sum_{i,j=1}^m a_{ij} \xi_i \xi_j \leq \Lambda \sum_{d(i)=1} \xi_i^2 \tag{5.2}$$

for a suitable constant  $\Lambda$ . We will also call

$$M_{m,\Lambda} \text{ the set of symmetric } m \times m \text{ real valued matrix, satisfying (5.2)} \tag{5.3}$$

If  $A$  is the identity matrix then the existence of a heat kernel for the operator  $L_A$  is a by now classical result due to Folland [37] and Rothschild and Stein [76]. Gaussian estimates have been provided by Jerison and Sanchez-Calle [53], and by Kusuoka and Strook [58]. There is a broad, more recent literature dealing with Gaussian estimates for non divergence form operators with Hölder continuous coefficients  $a_{ij}$ . Such estimates have been systematically studied in [7–9, 11] where a self-contained proof is provided.

A natural technique for studying the properties of the operator  $L_A$  is to consider a parabolic regularization induced by the vector fields  $X_i^\epsilon$  defined in (2.4). More precisely, we will define the operator

$$L_{\epsilon,A} = \partial_t - \sum_{i,j=1}^p a_{ij}^\epsilon X_i^\epsilon X_j^\epsilon \tag{5.4}$$

where  $a_{i,j}^\epsilon$  is any  $p \times p$  positive definite matrix belonging to  $M_{p,2\Lambda}$  and such that

$$a_{i,j}^\epsilon = a_{i,j} \quad \text{for } i, j = 1, \dots, m.$$

We will denote

$$M_{p,2\Lambda}^\epsilon \tag{5.5}$$

the set of such matrices. Formally, the operator  $L_A$  can be recovered as a limit as  $\epsilon \rightarrow 0$  of operator  $L_{\epsilon,A}$ . Here we are interested in understanding which are the properties of solutions of  $L_{\epsilon,A}$  which are preserved in the limit.

For  $\epsilon > 0$  consider a Riemannian metric  $g_\epsilon$  defined as in Remark 2.9, such that the vector fields  $X_i^\epsilon$  are orthonormal. The induced distance function  $d_\epsilon$  is biLipschitz equivalent to the Euclidean norm  $\| \cdot \|_E$ . Consequently, the operator  $L_{\epsilon,A}$  has a fundamental solution  $\Gamma_{\epsilon,A}$ , which can be estimated as

$$\Gamma_{\epsilon,A}(x) \leq C_\epsilon \frac{e^{-\frac{|x|_E^2}{C_\epsilon t}}}{t^{n/2}} \tag{5.6}$$

for some positive constant  $C_\epsilon$  depending on  $A, \epsilon$  and  $X_1, \dots, X_m$ .

Unfortunately the constant  $C_\epsilon$  blows up as  $\epsilon$  approaches 0, so the Riemannian estimate (5.7) alone does not provide Gaussian bounds of the fundamental solution  $\Gamma_A$  of the limit operator (5.1) as  $\epsilon$  goes to 0. In [57] the elliptic regularization technique has been used to obtain  $L^p$  and  $C^\alpha$  regularity of the solutions, which however are far from being optimal. In [27], new estimates uniform in  $\epsilon$  have been provided, in the time independent setting which are optimal with respect to the decay of the limit operator. In [17] the result has been extended to the parabolic operators, in the special case of Carnot groups.

In order to further extend these estimates, we need to formulate the following definition:

**Definition 5.1** (*Definition of the  $\mathcal{E}(2, d_\epsilon, M_{2\Lambda}^\epsilon)$  spaces*) We define  $\mathcal{E}(2+h, d_\epsilon, M_{p,2\Lambda}^\epsilon)$  to be the set of all kernels  $(P_{\epsilon,A})_{\epsilon>0, A \in M_{p,2\Lambda}^\epsilon}$ , defined on  $\mathbb{R}^{2n} \times ]0, \infty[$  that have an exponential decay of order  $2+h$ , uniformly with respect to a family of distances  $(d_\epsilon)_\epsilon$  and of matrices  $A \in M_{p,2\Lambda}^\epsilon$  (see definition 5.5), on any compact sets of an open set  $\Omega$ . More precisely, we will say that  $P_{\epsilon,A} \in \mathcal{E}(2+h, d_\epsilon, M_{p,2\Lambda}^\epsilon)$  if the following three conditions hold:

- For every  $K \subset\subset \Omega$  there exists a constant  $C_\Lambda > 0$  depending on  $\Lambda$  but independent of  $\epsilon > 0$ , and of the matrix  $A \in M_{p,2\Lambda}^\epsilon$  such that for each  $\epsilon > 0, x, y \in K$  and  $t > 0$  one has

$$C_\Lambda^{-1} t^{\frac{h}{2}} e^{-C_\Lambda \frac{d_\epsilon(x,y)^2}{t}} \leq P_{\epsilon,A}(x, y, t) \leq C_\Lambda \frac{t^{\frac{h}{2}} e^{-\frac{d_\epsilon(x,y)^2}{C_\Lambda t}}}{|B_\epsilon(x, \sqrt{t})|}. \tag{5.7}$$

- For  $s \in \mathbb{N}$  and  $k$ -tuple  $(i_1, \dots, i_k) \in \{1, \dots, m\}^k$  there exists a constant  $C_{s,k} > 0$  depending only on  $k, s, X_1, \dots, X_m, \Lambda$  such that

$$\left| (\partial_t^s X_{i_1} \cdots X_{i_k} P_{\epsilon,A})(x, y, t) \right| \leq C_{s,k} \frac{t^{\frac{h-2s-k}{2}} e^{-\frac{d_\epsilon(x,y)^2}{C_\Lambda t}}}{|B_\epsilon(x, \sqrt{t})|} \tag{5.8}$$

for all  $x, y \in K$  and  $t > 0$ .

- For any  $A_1, A_2 \in M_\Lambda, s \in \mathbb{N}$  and  $k$ -tuple  $(i_1, \dots, i_k) \in \{1, \dots, m\}^k$  there exists  $C_{s,k} > 0$  depending only on  $k, s, X_1, \dots, X_m, \Lambda$  such that

$$\begin{aligned} & \left| (\partial_t^s X_{i_1} \cdots X_{i_k} P_{\epsilon,A_1})(x, y, t) - \partial_t^s X_{i_1} \cdots X_{i_k} P_{\epsilon,A_2}(x, y, t) \right| \\ & \leq \|A_1 - A_2\| C_{s,k} \frac{t^{\frac{h-2s-k}{2}} e^{-\frac{d_\epsilon(x,y)^2}{C_\Lambda t}}}{|B_\epsilon(x, \sqrt{t})|}, \end{aligned} \tag{5.9}$$

where  $\|A\|^2 := \sum_{i,j=1}^n a_{ij}^2$ .

With these notations we will now extend all these previous results to vector fields which only satisfy the Hörmander condition, establishing estimates which are uniform in the variable  $\epsilon$  as  $\epsilon \rightarrow 0$ , and in the choice of the matrix  $A \in M_{2\Lambda}^\epsilon$  for the fundamental solutions  $\Gamma_{\epsilon,A}$  of the operators  $L_{\epsilon,A}$ . To be more specific, we will prove:

**Proposition 5.2** *The fundamental solution  $\Gamma_{\epsilon,A}$  of the operator  $L_{\epsilon,A}$ , is a kernel with exponential decay of order 2, uniform with respect to  $\epsilon > 0$  and to  $A \in M_{m,\Lambda}^\epsilon$ , according to definition (5.1). Hence it belongs to the set  $\mathcal{E}(2, d_\epsilon, M_{2\Lambda}^\epsilon)$ . Moreover, if  $\Gamma_A$  is the fundamental solution of the operator  $L_A$  defined in (5.1) one has*

$$X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon \partial_t^s \Gamma_{\epsilon,A} \rightarrow X_{i_1} \cdots X_{i_k} \partial_t^s \Gamma_A \tag{5.10}$$

as  $\epsilon \rightarrow 0$  uniformly on compact sets and in a dominated way on subcompacts of  $\Omega$ .

Our main contribution is that all the constants are independent of  $\epsilon$ . The proof of this assertion is based on a lifting procedure, which allows to express the fundamental solution of the operator  $L_{A,\epsilon}$  in terms of the fundamental solution of a new operator  $\bar{L}_A$  independent of  $\epsilon$ . The lifting procedure is composed by a first step in which we apply the delicate Rothschild and Stein lifting technique [76]. After that, when the vector fields are free up to a specific step, we apply a second lifting which has been introduced in [27], where the time independent case was studied, and from [17] where the Carnot group setting is considered.

The simplest example of such an equation is the Heat equation associated to the Kohn Laplacian in the Heisenberg group,  $\partial_t - X_1^2 - X_2^2$ , where the vector fields  $X_1$  and  $X_2$  have been expressed on coordinates in Example 2.1. In order to present our approach we will give an outline of the proof in this special setting.

*Example 5.3* Denote by  $(x_1, x_2, x_3)$  points of  $\mathbb{R}^3$ , let  $X_1, X_2, X_3$  be the vector fields defined in Example 2.1, and let  $I$  denote the identity matrix: Consider the parabolic operator

$$L_{\epsilon,I} = -\partial_t + X_1^2 + X_2^2 + \epsilon^2 X_3^2,$$

and note that it becomes degenerate parabolic as  $\epsilon \rightarrow 0$ . Let  $d_\epsilon$  denote the Carnot–Caratheodory distance associated to the distribution  $X_1, X_2, \epsilon X_3$ .

In order to handle such degeneracy we introduce new variables  $(z_1, z_2, z_3)$  and a new set of vector fields replicating the same structure of the initial ones, i.e.,

$$\hat{Z}_1 = \partial_{z_1} + z_2 \partial_{z_3}, \quad \hat{Z}_2 = \partial_{z_2} - z_1 \partial_{z_3}, \quad \hat{Z}_3 = \partial_{z_3}$$

with  $(x_1, x_2, x_3, z_1, z_2, z_3) \in \mathbb{H}^1 \times \mathbb{H}^1$ . The next step consists in lifting  $L_{\epsilon,I}$  to an operator

$$\bar{L}_\epsilon = \partial_t + X_1^2 + X_2^2 + Z_1^2 + Z_2^2 + (Z_3 + \epsilon X_3)^2,$$

defined on  $\mathbb{H}^1 \times \mathbb{H}^1$ , and denote by  $\bar{\Gamma}_\epsilon$  its fundamental solution. Let  $\bar{d}_\epsilon$  denote the Carnot–Caratheodory distance generated by  $X_1, X_2, Z_1, Z_2, (Z_3 + \epsilon X_3)$  and arguing as in (5.22) note that  $\bar{d}_\epsilon((x, z), (y, z)) \geq d_\epsilon(x, y) - C_0$ , for some constant  $C_0$  independent of  $\epsilon$ . Consider the change of variables on the Lie algebra of  $\mathbb{H}^1 \times \mathbb{H}^1$ ,

$$X_i \rightarrow X_i, Z_i \rightarrow Z_i, \quad \text{for } i = 1, 2, Z_3 + \epsilon X_3 \rightarrow Z_3.$$

Note that the Jacobian of such change of variables does not depend on  $\epsilon$  and that it reduces the operator  $\bar{L}_\epsilon$  to

$$\bar{L} = \partial_t + X_1^2 + X_2^2 + Z_1^2 + Z_2^2 + Z_3^2$$

whose fundamental solution we denote by  $\bar{\Gamma}$ . Note that this operator is parabolic with respect to the vector fields  $Z_i$  and degenerate parabolic with respect to the vector fields

$X_i$ . It is clear that the operator  $\bar{L}$  is independent of  $\epsilon$ , and consequently its fundamental solution  $\bar{\Gamma}$  satisfies standard Gaussian estimates with constants independent of  $\epsilon$

$$\bar{\Gamma}(x, t) \leq C_\Lambda \frac{e^{-\frac{\bar{d}(x,0)^2}{C_\Lambda t}}}{|\bar{B}(0, \sqrt{t})|},$$

where  $\bar{d}$  denotes the Carnot–Carathéodory distance in  $\mathbb{H}^1 \times \mathbb{H}^1$  generated by the distribution of vector fields  $X_1, X_2, Z_1, Z_2, Z_3$ . Changing back to the original variable we see that also  $\bar{\Gamma}_\epsilon$  satisfies analogous estimates with the same constants, with the distance  $\bar{d}$  replaced by the distance  $\bar{d}_\epsilon$  naturally associated to the operator  $\bar{L}_\epsilon$ . Finally, integrating with respect to the added variable  $(z_1, z_2, z_3)$ , we obtain an uniform bound for the fundamental solution of the operator  $L_{\epsilon,l}$  in terms of the distance  $d_\epsilon$ .

### 5.2 The Rothschild–Stein freezing and lifting theorems

Let us first recall a local lifting procedure introduced by Rothschild and Stein in [76] which, starting from a family  $(X_i)_{i=1,\dots,m}$  of Hörmander type vector fields of step  $s$  in a neighborhood of  $\mathbb{R}^n$ , leads to the construction of a new family of vector fields which are free, and of Hörmander type with the same step  $s$ , in a neighborhood of a larger space. The projection of the new free vector fields on  $\mathbb{R}^n$  yields the original vector fields, and that is why they are called *liftings*.

Let us start with some definitions:

**Definition 5.4** Denote by  $n_{m,s}$  the dimension (as a vector space) of the free nilpotent Lie algebra with  $m$  generators and step  $s$ . Let  $X_1, \dots, X_m$  be a set of smooth vector fields defined in an open neighborhood of a point  $x_0 \in \mathbb{R}^n$ , and let

$$V^{(s)} = \text{span} \{X^{(1)}, \dots, X^{(r)}\},$$

where the sets  $X^j$  are as defined in (2.2). We shall say that  $X_1, \dots, X_m$  are *free up to step  $s$*  if for any  $1 \leq r \leq s$  we have  $n_{m,s} = \dim(V^{(s)})$ .

If a point  $x_0 \in \mathbb{R}^n$  is fixed, the lifting procedure of Rothschild–Stein locally introduces new variables  $\tilde{z}$  and new vector fields  $(\tilde{Z}_i)$  expressed in terms of the new variables such that in a neighborhood  $U$  of  $x_0$  the vector fields  $\tilde{X}_i = (X_i + \tilde{Z}_i)_{i=1,\dots,m}$  are free at step  $s$ . More precisely, one has [76, Theorem 4]

**Theorem 5.5** Let  $X_1, \dots, X_m$  be a system of smooth vector fields, satisfying (1.1) in an open set  $U \subset \mathbb{R}^n$ . For any  $x \in U$  there exists a connected open neighborhood of the origin  $V \subset \mathbb{R}^{v-n}$ , and smooth functions  $\lambda_{ij}(x, \tilde{z})$ , with  $x \in \mathbb{R}^n$  and  $\tilde{z} = (z_{n+1}, \dots, z_v) \in V$ , defined in a neighborhood  $\tilde{U}$  of  $\tilde{x} = (x, 0) \in U \times V \subset \mathbb{R}^v$ , such that the vector fields  $\tilde{X}_1, \dots, \tilde{X}_m$  given by

$$\tilde{X}_i = X_i + \tilde{Z}_i, \quad \tilde{Z}_i = \sum_{j=n+1}^v \lambda_{ij}(x, \tilde{z}) \partial_{z_j}$$



are free up to step  $r$  at every point in  $\tilde{U}$ .

*Remark 5.6* In the literature the lifting procedure described above is often coupled with another key result introduced in [76], a nilpotent approximation which is akin to the classical *freezing* technique for elliptic operators. Let us explicitly note that in Sect. 5.3 we need only to apply the lifting theorem mentioned above, and not the freezing procedure. In particular, in the example of the so called Grushin vector fields

$$X_3 = \partial_{x_1} \text{ and } X_4 = x_1 \partial_{x_2}$$

they would need to be lifted through this procedure to the Heisenberg group structure

$$X_3 = \partial_{x_1} \text{ and } X_4 = \partial_{x_3} + x_1 \partial_{x_2}.$$

On the other hand the vector fields

$$X_1 = \cos \theta \partial_{x_1} + \sin \theta \partial_{x_2} \text{ and } X_2 = \partial_\theta$$

will be unchanged by the lifting process, since they are already free up to step 2.

Later on, In Sect. 5.4 we will apply Rothschild and Stein’s freezing theorem to a family of vector fields  $X_1, \dots, X_m$  free up to step  $r$ . This will allow to approximate a given family of vector fields with homogeneous ones. Note that in this case the function  $\Phi$  in (5.14) is independent of  $v$  and its expression reduces to:

$$\Phi_x(u) = \exp\left(\sum_i u_i X_i\right)(x). \tag{5.11}$$

The pertinent theorem from [76] is the following,

**Theorem 5.7** *Let  $X_1, \dots, X_m$  be a family of vector fields are free up to rank  $r$  at every point. Then for every  $x$  there exists a neighborhood  $V$  of  $x$  and a neighborhood  $U$  of the identity in  $G_{m,r}$ , such that:*

- (a) *the map  $\Phi_x : U \rightarrow V$  is a diffeomorphism onto its image. We will call  $\Theta_x$  its inverse map*
- (b) *we have*

$$d\Theta_x(X_i) = Y_i + R_i, \quad i = 1, \dots, m \tag{5.12}$$

where  $R_i$  is a vector field of local degree less or equal than zero, depending smoothly on  $x$ .

Hence the operator  $R_i$  will be represented in the form:

$$R_i = \sum_{jh} \sigma_i(u) X_j,$$

where  $\sigma$  is an homogeneous polynomial of degree  $d(X_i) - 1$ .

### 5.3 A lifting procedure uniform in $\epsilon$

So far we have started with a set of Hörmander vector fields  $X_1, \dots, X_m$  in  $\Omega \subset \mathbb{R}^n$  and we have lifted them through Theorem 5.5 to a set  $\tilde{X}_1, \dots, \tilde{X}_m$  of Hörmander vector fields that are free up to a step  $s$  in a neighborhood  $\tilde{\Omega} \subset \mathbb{R}^v$ . Next, we perform a second lifting inspired by the work in [17]. We will consider the augmented space  $\mathbb{R}^v \times \mathbb{R}^v$  defined in terms of  $v$  new coordinates  $\hat{z} = (\hat{z}_1, \dots, \hat{z}_v)$ . Set  $z = (\tilde{z}, \hat{z})$  and denote points of  $\mathbb{R}^v \times \mathbb{R}^v$  by  $\bar{x} = (x, \tilde{z}, \hat{z}) = (x, z)$ . Denote by  $\hat{Z}_1, \dots, \hat{Z}_m$  a family of vector fields free up to step  $s$ .  $\tilde{X}_1, \dots, \tilde{X}_m$ , i.e. a family of vector fields free of step  $s$  in the variables  $\hat{z}$ , and let

$$\hat{Z}_{m+1}, \dots, \hat{Z}_v$$

denote the complete set of their commutators, as we did in (2.2). Note that the subRiemannian structure generated by  $\hat{Z}_1, \dots, \hat{Z}_m$  coincides with the structure generated by the family  $\tilde{X}_i$ , but are defined in terms of new variables  $\hat{z}$ .

For every  $\epsilon \in [0, 1)$  consider a sub-Riemannian structure determined by the choice of horizontal vector fields given by

$$(\bar{X}_1^\epsilon, \dots, \bar{X}_{m+v}^\epsilon) = (\tilde{X}_1, \dots, \tilde{X}_m, \hat{Z}_1, \dots, \hat{Z}_m, \tilde{X}_{m+1}^\epsilon + \hat{Z}_{m+1}, \dots, \tilde{X}_v^\epsilon + \hat{Z}_v). \tag{5.13}$$

Since the space is free up to step  $r$  the function  $\Phi$  in (5.14) is independent of  $v$  and its expression reduces to:

$$\Phi_{\epsilon, \bar{x}}(u) = \exp\left(\sum_i u_i^\epsilon \bar{X}_i^\epsilon\right)(\bar{x}). \tag{5.14}$$

In the sequel, when we will need to explicitly indicate the vector fields defining  $\Phi$  we will also use the notation:

$$\Phi_{\epsilon, \bar{x}, \bar{X}^\epsilon}(u) = \Phi_{\epsilon, \bar{x}}(u), \text{ and } \Phi_{\epsilon, \bar{x}, \bar{X}^\epsilon_i}(u) \text{ its components,} \tag{5.15}$$

and analogous notations will be used for the inverse map  $\Theta_{\epsilon, \bar{x}, \bar{X}^\epsilon}$

For every  $\epsilon > 0$  and  $\bar{x}, \bar{x}_0$ , in view of Theorem 3.10 the associated ball box distances reduce to:

$$\bar{d}_\epsilon(\bar{x}, \bar{x}_0) = \sum_{i=1}^{2m} |u_i^\epsilon| + \sum_{i=2m+1}^{v+m} \min(|u_i^\epsilon|, |u_i^\epsilon|^{1/d(i)}) + \sum_{i=v+m+1}^{2v} |u_i^\epsilon|^{1/d(i)}$$

For  $\epsilon = 0$  and  $\bar{x}, \bar{x}_0$  we have

$$\bar{d}_0(\bar{x}, \bar{x}_0) = \sum_{i=1}^n |u_i^0|^{1/d(i)}$$

### 5.4 Proof of the stability result

The sub-Laplacian/heat operator associated to this structure is

$$\bar{L}_{\epsilon,A} = \partial_t - \sum_{i=1}^{m+\nu} \bar{a}_{ij} \bar{X}_i^\epsilon \bar{X}_j^\epsilon,$$

where

$$\bar{A} = A \oplus \lambda I$$

and  $I$  is the identity matrix of dimension  $\nu \times \nu$ . We denote by  $\bar{\Gamma}_{\epsilon,A}$  the heat kernels of the corresponding heat operators, and prove a lemma analogous to Lemma 5.2 for the lifted operator:

**Lemma 5.8** *The fundamental solution  $\bar{\Gamma}_{\epsilon,A}$  of the operator  $\bar{L}_{\epsilon,A}$ , is a kernel with local uniform exponential decay of order 2 with respect to  $\epsilon > 0$  and  $A \in M_{m+\nu,\Lambda}^\epsilon$ , according to definition (5.1). Hence it belongs to the set  $\mathcal{E}(2, \bar{d}_\epsilon, M_{m+\nu,\Lambda}^\epsilon)$ . Moreover, as  $\epsilon \rightarrow 0$  one has*

$$X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon \partial_t^s \bar{\Gamma}_{\epsilon,A} \rightarrow X_{i_1} \cdots X_{i_k} \partial_t^s \bar{\Gamma}_A \tag{5.16}$$

uniformly on compact sets, in a dominated way on all  $\bar{G}$ .

*Proof* The result for the limit operator  $\bar{L}_{0,A}$  is well known and contained for example in [11]. Hence we only have to estimate the fundamental solution of the operators  $\bar{L}_{\epsilon,A}$  in terms of the one of  $\bar{L}_{0,A}$ . In order to do so, we first define a change of variable on the Lie algebra:

$$T_\epsilon(\bar{X}_i^\epsilon) = \bar{X}_i^0 \text{ for } i = 1, \dots, \nu + m \tag{5.17}$$

Then from a fixed point  $\bar{z}$  we apply the exponential map to induce on the Lie group a volume preserving change of variables. Using the notation introduced in (5.15), we will denote

$$\bar{F}_{\epsilon,\bar{z}} : \bar{G} \rightarrow \bar{G}, \quad \bar{F}_\epsilon(\bar{x}) = \exp(\Phi_{\epsilon,\bar{z},T_\epsilon(\bar{X}_i^0)}^{-1}(\bar{x})\bar{X}_i^0)(\bar{z})$$

Since the distances are defined in terms of the exponential maps, this change of variables induces a relation between the distances  $\bar{d}_0$  and  $\bar{d}_\epsilon$ :

$$\bar{d}_\epsilon(\bar{x}, \bar{x}_0) = \bar{d}_0(\bar{F}_\epsilon(\bar{x}), \bar{F}_\epsilon(\bar{x}_0)). \tag{5.18}$$

Analogously we also have

$$\bar{\Gamma}_{\epsilon,A}(\bar{x}, \bar{y}, t) = \bar{\Gamma}_{0,A}(\bar{F}_\epsilon(\bar{x}), \bar{F}_\epsilon(\bar{y}), t), \tag{5.19}$$

Hence assertions (5.7) follow from the estimates of  $\bar{\Gamma}_{0,A}$  contained for instance in [53]. Indeed the second inequality can be established as follows:

$$\bar{\Gamma}_{\epsilon,A}(\bar{x}, \bar{y}, t) = \bar{\Gamma}_{0,A}(\bar{F}_\epsilon(\bar{x}), \bar{F}_\epsilon(\bar{y}), t) \leq C_\Lambda \frac{e^{-\frac{\bar{d}_0(\bar{F}_\epsilon(\bar{x}), \bar{F}_\epsilon(\bar{y}))^2}{C_\Lambda t}}}{|\bar{B}_0(\bar{F}_\epsilon(\bar{x}), \sqrt{t})|} = C_\Lambda \frac{e^{-\frac{\bar{d}_\epsilon(\bar{x}, \bar{y})^2}{C_\Lambda t}}}{|\bar{B}_\epsilon(\bar{x}, \sqrt{t})|}.$$

The proof of the first inequality in (5.7) and (5.8) is analogous, while (5.9) follows from the estimates of the fundamental solution contained in ([11]).

The pointwise convergence (5.16) is also an immediate consequence of (5.18) and (5.19). In order to prove the dominated convergence result we need to relate the distances  $\bar{d}_0$  and  $\bar{d}_\epsilon$ . On the other side, the change of variable (5.17) allows to express exponential coordinates  $u_i^\epsilon$ , in terms of  $u_i^0$  as follows:

$$\bar{d}_\epsilon(\bar{x}, \bar{x}_0) = \sum_{i=1}^{2m} |u_i^0| + \sum_{i=2m+1}^v \left( |u_i^0 - \epsilon w_{i+\nu}^0|^{1/d(i)} + \min(|u_i^0|, |u_i^0|^{1/d(i)}) \right)$$

so that for all<sup>2</sup>  $\bar{x}, \bar{x}_0 \in \bar{G}$

$$\bar{d}_0(\bar{x}, \bar{x}_0) - C_0 \leq \bar{d}_\epsilon(\bar{x}, \bar{x}_0) \leq \bar{d}_0(\bar{x}, \bar{x}_0) + C_0 \tag{5.20}$$

where  $C_0$  is independent of  $\epsilon$ . The latter and (5.8) imply that there is a constant  $\tilde{C}_{s,k}$  independent of  $\epsilon$  such that

$$\left| (\partial_t^s X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon \bar{\Gamma}_{\epsilon,A})(\bar{x}, \bar{y}, t) \right| \leq \tilde{C}_{s,k} t^{-s-k/2} \frac{e^{-\frac{\bar{d}_0(\bar{x}, \bar{y})^2}{C_\Lambda t}}}{|\bar{B}_0(\bar{x}, \sqrt{t})|}$$

and this imply dominated convergence with respect to the  $\epsilon$  variable. □

In order to be able to conclude the proof of Proposition 5.2, we need to study the relation between the fundamental solutions  $\Gamma_A(x, y, t)$  and its lifting  $\bar{\Gamma}_{0,A}((x, 0), (y, z), t)$ , as well as the relation between  $\Gamma_{\epsilon A}(x, y, t)$  and  $\bar{\Gamma}_{\epsilon,A}((x, 0), (y, z), t)$ ,

*Remark 5.9* We first note that for every  $f \in C_0^\infty(\mathbb{R}^n \times \mathbb{R}^+)$   $f$  can be identified with a  $C^\infty$  and bounded function defined on  $\mathbb{R}^{n+\nu} \times \mathbb{R}^+$  and constant in the  $z$ -variables. Hence

$$L_A f = \bar{L}_A f, \quad L_{\epsilon,A} f = \bar{L}_{\epsilon,A} f,$$

Consequently:

$$f(x, t) = \int \int \left( \int \bar{\Gamma}_{\epsilon,A}((x, 0, s), (y, z, t)) dz \right) L_{\epsilon,A} f(y, s) dy ds$$

---

<sup>2</sup> This estimate indicates the well known fact that at large scale the Riemannian approximating distances are equivalent to the sub-Riemannian distance.

From the definition of fundamental solution we can deduce that

$$\begin{aligned} \Gamma_A(x, y, t) &= \int_G \bar{\Gamma}_{0,A}((x, 0), (y, z), t) dz, \quad \text{and} \\ \Gamma_{\epsilon A}(x, y, t) &= \int_G \bar{\Gamma}_{\epsilon,A}((x, 0), (y, z), t) dz, \end{aligned} \tag{5.21}$$

for any  $x \in G$  and  $t > 0$ .

We conclude this section with the proof of the main result Proposition 5.2.

*Proof* In view of the previous remark and (global) dominated convergence of the derivatives of  $\bar{\Gamma}_{\epsilon,A}$  to the corresponding derivatives of  $\bar{\Gamma}_{0,A}$  as  $\epsilon \rightarrow 0$ , we deduce that

$$\int_G \bar{\Gamma}_{\epsilon,A}((x, 0), (y, z), t) dz \rightarrow \int_G \bar{\Gamma}_{0,A}((x, 0), (y, z), t) dz$$

as  $\epsilon \rightarrow 0$ . The Gaussian estimates of  $\Gamma_{\epsilon,A}$  follow from the corresponding estimates on  $\bar{\Gamma}_{\epsilon,A}$  and the fact that in view of (5.20),

$$\begin{aligned} \bar{d}_\epsilon((x, z), (x_0, z_0)) &\geq \bar{d}_0((x, z), (x_0, z_0)) - C_0 \geq d_0(x, x_0) + d_0(z, z_0) - C_0 \\ &\geq d_\epsilon(x, x_0) + d_\epsilon(z, z_0) - 3C_0 \end{aligned} \tag{5.22}$$

Indeed the latter shows that there exists a constant  $C > 0$  depending only on  $G, \sigma_0$  such that for every  $x \in G$ ,

$$\int_G e^{-\frac{d_\epsilon^2((x,z),(x_0,z_0))}{t}} dz \leq C e^{-\frac{d_\epsilon^2(x,x_0)}{t}} \int_G e^{-\frac{d_\epsilon^2(z,z_0)}{t}} dz \leq C e^{-\frac{d_\epsilon^2(x,x_0)}{t}}.$$

The conclusion follows at once. □

### 5.5 Differential of the integral operator associated to $\Gamma_\epsilon$

In this subsection we will show how to differentiate a functional  $F$  expressed as follows:

$$F(f)(x, t) = \int \Gamma_{\epsilon,A}(x, y, t) f(y, s) dy ds.$$

In order to do so, we will need to differentiate both with respect to  $x$  and to  $y$ , so that we will denote  $X_i^{\epsilon,x} \Gamma_{\epsilon,A}(x, y, t)$  the derivative with respect to the variable  $x$  and  $X_i^{\epsilon,y} \Gamma_{\epsilon,A}(x, y, t)$  the derivative with respect to the variable  $y$ .

Analogously, we will denote the derivative with the first variable of the lifted fundamental solution

$$\bar{X}_i^{\epsilon,\bar{x}} \bar{\Gamma}_{\epsilon,A}((x, w), (y, z), t).$$

For  $\epsilon = 0$ , we will have by definition

$$\tilde{X}_i^{0,\bar{x}} \bar{\Gamma}_{\epsilon,A}((x, w), (y, z), t) = (X_i^{0,x} + \tilde{Z}_i^w) \bar{\Gamma}_{\epsilon,A}((x, w), (y, z), t).$$

The derivative with respect to the second variable will be denoted  $\tilde{X}_i^{0,\bar{y}}$ . If  $\Gamma$  is the Euclidean heat kernel, there is a simple relation between the derivative with respect to the two variables, indeed in this case  $\Gamma_{\epsilon,A}(x, y, t) = \Gamma_{\epsilon,A}(x - y, 0, t)$ , so that

$$X_i^{\epsilon,x} \Gamma_{\epsilon,A}(x, y, t) = -X_i^{\epsilon,y} \Gamma_{\epsilon,A}(x, y, t). \tag{5.23}$$

Consequently for every function  $f \in C_0^\infty$

$$\partial_{x_i} F(f)(x, t) = \int \Gamma_{\epsilon,A}(x, y, t) \partial_{y_i} f(y) dy.$$

This is no more the case in general Lie groups, or for Hörmander vector fields. However we will see that there is a relation between the two derivatives, which allows to prove the following:

**Proposition 5.10** *Assume that  $f \in C_0^\infty(\Omega \times ]0, T[)$  in an open set  $\Omega \times ]0, T[$ . For every  $x \in K \subset \subset \Omega$ , for every  $i = 1, \dots, m$  there exists the derivative  $X_i^\epsilon F(f)(x, t)$ . Precisely there exist kernels  $P_{\epsilon,i,h}(x, y, t), R_{\epsilon,i}(x, y, t) \in \mathcal{E}(2, d_\epsilon, M_{m,\Lambda}^\epsilon)$  such that*

$$X_i^\epsilon F(f) = - \int \sum_{h=1}^m X_h^{\epsilon,y,*} P_{\epsilon,i,h}(x, y, t) f(y) dy - \int R_{\epsilon,i}(x, y, t) f(y) dy.$$

(Let us note explicitly that the term  $R_{\epsilon,i,h}(x, y, t)$  plays the role of an error term).

*Proof* We can apply the lifting procedure described in Sects. 5.2 and 5.3, and representing the fundamental solution as in (5.19) and (5.21), we obtain the following expression for  $F_\epsilon$ :

$$\begin{aligned} F_\epsilon(f) &= \int \int_G \bar{\Gamma}_{\epsilon,A}((x, 0), (y, z), t) dz f(y) dy \\ &= \int \int \bar{\Gamma}_{0,A}(\bar{F}_\epsilon(x, 0), \bar{F}_\epsilon(y, z), t) dz f(y) dy. \end{aligned}$$

By differentiating with respect to  $X_i^\epsilon$  we get:

$$X_i^\epsilon F_\epsilon(f)(x) = \int \int (\tilde{X}_i^{0,x} - \tilde{Z}_i^w) \bar{\Gamma}_{0,A}(\bar{F}_\epsilon(x, 0), \bar{F}_\epsilon(y, z), t) dz f(y) dy. \tag{5.24}$$

Note that the family of vectors  $\tilde{X}_i^0$  is independent of  $\epsilon$  and free of step  $r$ . Hence, in view of [76, page 295, line 3 from below] one has that for every  $i, j = 1, \dots, m$ , there

exist families of indices  $I_{i,j}$ , and polynomials  $\bar{p}_{ih}$  homogeneous of degree  $\geq h$  such that:

$$\begin{aligned} \bar{X}_i^{0,\bar{x}} \bar{\Gamma}_{0,A}(\bar{x}, \bar{y}, t) &= \sum_{j=1}^m \left( \bar{X}_j^{0,\bar{y}} \right)^* \sum_{h \in I_{i,j}} \bar{X}_h^{0,\bar{y}} \left( \bar{p}_{ih}(\Theta_{\bar{x}}(\bar{y})) \right) \bar{\Gamma}_{0,A}(\bar{x}, \bar{y}, t) \\ &\quad - \left( \sum_{j=1}^m \left( \bar{X}_j^{0,\bar{y}} \right)^* \sum_{h \in I_{i,j}} \bar{X}_h^{0,\bar{y}} \right) \left( \bar{p}_{ih}(\Theta_{\bar{x}}(\bar{y})) \right) \bar{\Gamma}_{0,A}(\bar{x}, \bar{y}, t). \end{aligned}$$

In particular using this expression in the variable  $z$  alone, and integrating by parts we deduce

$$\int \int \tilde{Z}_i^w \bar{\Gamma}_{0,A}(\bar{F}_\epsilon(x, 0), \bar{F}_\epsilon(y, z), t) dz = 0$$

We now call

$$\bar{R}_{0,i}(\bar{x}, \bar{y}, t) = \left( \sum_{j=1}^m \left( \bar{X}_j^{0,\bar{y}} \right)^* \sum_{h \in I_{i,j}} \bar{X}_h^{0,\bar{y}} \right) \left( \bar{p}_{ih}(\Theta_{\bar{x}}(\bar{y})) \right) \bar{\Gamma}_{0,A}(\bar{x}, \bar{y}, t)$$

This kernel, being obtained by multiplication of  $\Gamma_{0,A}(\bar{x}, \bar{y}, t)$  by a polynomial, has locally the same decay as  $\Gamma_{0,A}(\bar{x}, \bar{y}, t)$ . In particular it is clear that the conditions 5.7, 5.8, 5.9 are satisfied uniformly with respect to  $\epsilon$ , since there is no dependence on  $\epsilon$ . As a consequence, if we set

$$R_{\epsilon,i}(x, y, t) = \int \bar{R}_{0,i}(\bar{F}_\epsilon(x, 0), \bar{F}_\epsilon(y, z), t) dz$$

then  $R_{\epsilon,i}(x, y, t) \in \mathcal{E}(2, d_\epsilon, M_{m,\Lambda}^\epsilon)$  Similarly we call

$$\bar{P}_{\epsilon,i,h}(\bar{x}, \bar{y}, t) = \sum_{h \in I_{i,j}} \bar{X}_h^{0,\bar{y}} \left( \bar{p}_{ih}(\Theta_{\bar{x}}(\bar{y})) \right) \bar{\Gamma}_{0,A}(\bar{x}, \bar{y}, t)$$

Now we use the fact that  $\bar{\Gamma}_{0,A} \in \mathcal{E}(2, \bar{d}, M_{m+\nu,\Lambda}^\epsilon)$  together with the fact that  $\bar{p}_{ih}$  is a polynomial of the degree equal of the order of  $\bar{X}_h^{0,\bar{y}}$  to conclude that

$$\bar{P}_{\epsilon,i,h}(\bar{x}, \bar{y}, t) \in \mathcal{E}(2, \bar{d}, M_{m+\nu,\Lambda}^\epsilon)$$

Setting

$$P_{\epsilon,i,h}(x, y, t) = \int \bar{P}_{0,i,h}(\bar{F}_\epsilon(x, 0), \bar{F}_\epsilon(y, z), t) dz$$

it follows that  $P_{\epsilon,i,h}(x, y, t) \in \mathcal{E}(2, d_\epsilon, M_{m,\Lambda}^\epsilon)$

Substituting this expression into equation (5.24) we get

$$X_i^\epsilon F_\epsilon(f)(x) = - \int \sum_{j=1}^m \left(\tilde{X}_j^{0,y}\right)^* P_{\epsilon,i,h}(x, y, t) f(y) dy - \int R_{\epsilon,i}(x, y, t) f(y) dy. \tag{5.25}$$

□

### 6 Stability of interior Schauder estimates

In this section we will prove uniform estimates in spaces of Hölder continuous functions and in Sobolev spaces for solutions of second order sub-elliptic differential equations in non divergence form

$$L_{\epsilon,A}u \equiv \partial_t u - \sum_{i,j=1}^n a_{ij}^\epsilon(x, t) X_i^\epsilon X_j^\epsilon u = 0,$$

in a cylinder  $Q = \Omega \times (0, T)$  that are stable as  $\epsilon \rightarrow 0$ . The proof of both estimates is largely based on the heat kernel estimates established above. Internal Schauder estimates for these type of operators are well known. We recall the results of Xu [83], Bramanti and Brandolini [10] for heat-type operators, and the results of Lunardi [62], and Polidoro and Di Francesco [34], and Gutierrez and Lanconelli [47], which apply to a large class of squares of vector fields plus a drift term. We also recall [64] where uniform Schauder estimates for a particular elliptic approximation of subLaplacians are proved.

Here the novelty is due to the uniform condition with respect to  $\epsilon$ . This is accomplished by using the uniform Gaussian bounds established in in the previous section. This result extends to Hörmander type operators the analogous assertion proved by Manfredini and the authors in [14] in the setting of Carnot Groups.

#### 6.1 Uniform Schauder estimates

Let us start with the definition of classes of Hölder continuous functions in this setting

**Definition 6.1** Let  $0 < \alpha < 1$ ,  $Q \subset \mathbb{R}^{n+1}$  and  $u$  be defined on  $Q$ . We say that  $u \in C_{\epsilon,X}^\alpha(Q)$  if there exists a positive constant  $M$  such that for every  $(x, t), (x_0, t_0) \in Q$

$$|u(x, t) - u(x_0, t_0)| \leq M \tilde{d}_\epsilon^\alpha((x, t), (x_0, t_0)). \tag{6.1}$$

We put

$$\|u\|_{C_{\epsilon,X}^\alpha(Q)} = \sup_{(x,t) \neq (x_0,t_0)} \frac{|u(x, t) - u(x_0, t_0)|}{\tilde{d}_\epsilon^\alpha((x, t), (x_0, t_0))} + \sup_Q |u|.$$



Iterating this definition, if  $k \geq 1$  we say that  $u \in C_{\epsilon, X}^{k, \alpha}(Q)$  if for all  $i = 1, \dots, m$   $X_i \in C_{\epsilon, X}^{k-1, \alpha}(Q)$ . Where we have set  $C_{\epsilon, X}^{0, \alpha}(Q) = C_{\epsilon, X}^{\alpha}(Q)$ .

The main results of this section, which generalizes to the Hörmander vector fields setting our previous result with Manfredini in [14] is

**Proposition 6.2** *Let  $w$  be a smooth solution of  $L_{\epsilon, A} w = f$  on  $Q$ . Let  $K$  be a compact sets such that  $K \subset\subset Q$ , set  $2\delta = d_0(K, \partial_p Q)$  and denote by  $K_\delta$  the  $\delta$ -tubular neighborhood of  $K$ . Assume that there exists a constant  $C > 0$  such that*

$$\|a_{ij}^\epsilon\|_{C_{\epsilon, X}^{k, \alpha}(K_\delta)} \leq C,$$

for any  $\epsilon \in (0, 1)$ . There exists a constant  $C_1 > 0$  depending on  $\alpha, C, \delta$ , and the constants in Proposition 5.2, but independent of  $\epsilon$ , such that

$$\|w\|_{C_{\epsilon, X}^{k+2, \alpha}(K)} \leq C_1 \left( \|f\|_{C_{\epsilon, X}^{k, \alpha}(K_\delta)} + \|w\|_{C_{\epsilon, X}^{k+1, \alpha}(K_\delta)} \right).$$

We will first consider to a constant coefficient operator, for which we will obtain a representation formula, then we will show how to obtain from this the claimed result.

Precisely we will consider the constant coefficient *frozen* operator:

$$L_{\epsilon, (x_0, t_0)} \equiv \partial_t - \sum_{i, j=1}^n a_{ij}^\epsilon(x_0, t_0) X_i^\epsilon X_j^\epsilon,$$

where  $(x_0, t_0) \in Q$ . We explicitly note that for  $\epsilon > 0$  fixed the operator  $L_{\epsilon, (x_0, t_0)}$  is uniformly parabolic, so that its heat kernel can be studied through standard singular integrals theory in the corresponding Riemannian balls.

As a direct consequence of the definition of fundamental solution one has the following representation formula

**Lemma 6.3** *Let  $w$  be a smooth solution to  $L_\epsilon w = f$  in  $Q \subset \mathbb{R}^{n+1}$ . For every  $\phi \in C_0^\infty(Q)$ ,*

$$\begin{aligned} (w\phi)(x, t) &= \int_Q \Gamma_{(x_0, t_0)}^\epsilon((x, t), (y, \tau)) (L_{\epsilon, (x_0, t_0)} - L_\epsilon) (w\phi)(y, \tau) dy d\tau \\ &+ \int_Q \Gamma_{(x_0, t_0)}^\epsilon((x, t), (y, \tau)) \left( f\phi + wL_\epsilon\phi + 2 \sum_{i, j=1}^n a_{ij}^\epsilon(y, \tau) X_i^\epsilon w X_j^\epsilon \phi \right) (y, \tau) dy d\tau, \end{aligned} \tag{6.2}$$

where we have denoted by  $\Gamma_{(x_0, t_0)}^\epsilon$  the heat kernel for of  $L_{\epsilon, (x_0, t_0)}$ .

Iterating the previous lemma we get the following

**Lemma 6.4** *Let  $k \in \mathbb{N}$  and consider a  $k$ -tuple  $(i_1, \dots, i_k) \in \{1, \dots, m\}^k$ . There exists a differential operator  $B$  of order  $k + 1$ , depending on horizontal derivatives of*

$a_{ij}^\epsilon$  of order at most  $k$ , such that

$$X_{i_k}^\epsilon \cdots X_{i_1}^\epsilon (L_{\epsilon, (x_0, t_0)} - L_\epsilon) = \sum_{i, j=1}^n (a_{ij}^\epsilon - a_{ij}^\epsilon(x_0, t_0)) X_{i_k}^\epsilon \cdots X_{i_1}^\epsilon X_i^\epsilon X_j^\epsilon + B.$$

*Proof* The proof can be made by induction. Indeed it is true with  $B = 0$  by definition if  $k = 0$ :

$$L_{\epsilon, (x_0, t_0)} - L_\epsilon = \sum_{i, j=1}^n (a_{ij}^\epsilon - a_{ij}^\epsilon(x_0, t_0)) X_i^\epsilon X_j^\epsilon.$$

if it true for a fixed value of  $k$  then we have

$$\begin{aligned} & X_{i_{k+1}}^\epsilon X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon (L_{\epsilon, (x_0, t_0)} - L_\epsilon) \\ &= \sum_{i, j=1}^n (a_{ij}^\epsilon - a_{ij}^\epsilon(x_0, t_0)) X_{i_{k+1}}^\epsilon X_{i_k}^\epsilon \cdots X_{i_1}^\epsilon X_i^\epsilon X_j^\epsilon + \tilde{B} \end{aligned}$$

where

$$\tilde{B} = X_{i_{k+1}}^\epsilon (a_{ij}^\epsilon - a_{ij}^\epsilon(x_0, t_0)) X_{i_k}^\epsilon \cdots X_{i_1}^\epsilon X_i^\epsilon X_j^\epsilon + X_{i_{k+1}}^\epsilon B.$$

By the properties of  $B$  it follows that  $\tilde{B}$  is a differential operator of order  $k + 2$ , depending on horizontal derivatives of  $a_{ij}^\epsilon$  of order at most  $k + 1$ . This concludes the proof.  $\square$

We can go back to our operator  $L$  and establish the following regularity results, differentiating twice the representation formula:

**Proposition 6.5** *Let  $0 < \alpha < 1$  and  $w$  be a smooth solution of  $L_\epsilon w = f \in C_{\epsilon, X}^\alpha(Q)$  in the cylinder  $Q$ . Let  $K$  be a compact sets such that  $K \subset\subset Q$ , set  $2\delta = d_0(K, \partial_p Q)$  and denote by  $K_\delta$  the  $\delta$ -tubular neighborhood of  $K$ . Assume that there exists a constant  $C > 0$  such that for every  $\epsilon \in (0, 1)$*

$$\|a_{ij}^\epsilon\|_{C_{\epsilon, X}^\alpha(K_\delta)} \leq C.$$

*There exists a constant  $C_1 > 0$  depending on  $\delta, \alpha, C$  and the constants in Proposition 5.2 such that*

$$\|w\|_{C_{\epsilon, X}^{2, \alpha}(K)} \leq C_1 \left( \|f\|_{C_{\epsilon, X}^\alpha(K_\delta)} + \|w\|_{C_{\epsilon, X}^{1, \alpha}(K_\delta)} \right).$$

*Proof* The proof follows the outline of the standard case, as in [41], and rests crucially on the Gaussian estimates proved in Proposition 5.2. Choose a parabolic sphere<sup>3</sup>

<sup>3</sup> That is a sphere in the group  $\tilde{G} = G \times \mathbb{R}$  in the pseudo-metric  $\tilde{d}_\epsilon((x, t), (y, s)) = \max(d_\epsilon(x, y), \sqrt{|t - s|})$ .

$B_{\epsilon,\delta} \subset\subset K$  where  $\delta > 0$  will be fixed later and a cut-off function  $\phi \in C_0^\infty(\mathbb{R}^{n+1})$  identically 1 on  $B_{\epsilon,\delta/2}$  and compactly supported in  $B_{\epsilon,\delta}$ . This implies that for some constant  $C > 0$  depending only on  $G$  and  $\sigma_0$ ,

$$|\nabla_\epsilon \phi| \leq C\delta^{-1}, \quad |L^\epsilon \phi| \leq C\delta^{-2},$$

in  $\mathcal{Q}$ . Now we represent the function  $w\phi$  through the formula 6.3 and take two derivatives in the direction of the vector fields. We remark once more that the operator is uniformly elliptic due to the  $\epsilon$ -regularization, hence the differentiation under the integral can be considered standard. As a consequence for every multi-index  $I = (i_1, i_2) \in \{1, \dots, m\}^2$  and for every  $(x_0, t_0) \in B_{\epsilon,\delta}$  one has:

$$X_{i_1}^\epsilon X_{i_2}^\epsilon (w\phi)(x_0, t_0) \tag{6.3}$$

$$\begin{aligned} &= \int_{\mathcal{Q}} X_{i_1}^\epsilon X_{i_2}^\epsilon \Gamma_{(x_0,t_0)}^\epsilon(\cdot, (y, \tau))|_{(x_0,t_0)} (L_{\epsilon,(x_0,t_0)} - L_\epsilon) (w\phi)(y, \tau) dy d\tau \\ &+ \int_{\mathcal{Q}} X_{i_1}^\epsilon X_{i_2}^\epsilon \Gamma_{(x_0,t_0)}^\epsilon(\cdot, (y, \tau))_{(x_0,t_0)} \left( f\phi + wL_\epsilon\phi \right. \\ &\left. + 2 \sum_{i,j=1}^n a_{ij}^\epsilon X_i^\epsilon w X_j^\epsilon \phi \right) (y, \tau) dy d\tau. \end{aligned} \tag{6.4}$$

In order to study the Hölder continuity of the second derivatives, we note that the uniform Hölder continuity of  $a_{ij}^\epsilon$ , and Proposition 5.2 ensure that the kernel satisfy the classical singular integral properties (see [37]):

$$\begin{aligned} &|X_{i_1}^\epsilon X_{i_2}^\epsilon \Gamma_{(x,t)}^\epsilon((x, t), (y, \tau)) - X_{i_1}^\epsilon X_{i_2}^\epsilon \Gamma_{(x_0,t_0)}^\epsilon((x_0, t_0), (y, \tau))| \\ &\leq C \tilde{d}_\epsilon^\alpha((x, t), (x_0, t_0)) \frac{(\tau - t_0)^{-1} e^{-\frac{d_\epsilon(x_0,y)^2}{C_\Lambda(\tau-t_0)}}}{|B_\epsilon(0, \sqrt{\tau - t_0})|}, \end{aligned}$$

with  $C > 0$  independent of  $\epsilon$ . From here, proceeding as in [41, Theorem 2, Chapter 4], the first term in the right hand side of formula (6.3) can be estimated as follows:

$$\begin{aligned} &\left\| \int X_{i_1}^\epsilon X_{i_2}^\epsilon \Gamma_{(x_0,t_0)}^\epsilon(\cdot, (y, \tau))(L_\epsilon - L_{\epsilon,(x_0,t_0)})(w\phi)(y, \tau) dy d\tau \right\|_{C_{\epsilon,X}^\alpha(B_{\epsilon,\delta})} \\ &\leq C_1 \left\| (L_\epsilon - L_{\epsilon,(x_0,t_0)})(w\phi) \right\|_{C_{\epsilon,X}^\alpha(B_{\epsilon,\delta})} \\ &= C_1 \sum_{i,j} \left\| (a_{ij}^\epsilon(x_0, t_0) - a_{ij}^\epsilon(\cdot)) X_j^\epsilon X_j^\epsilon (w\phi) \right\|_{C_{\epsilon,X}^\alpha(B_{\epsilon,\delta})} \\ &\leq \tilde{C}_1 \delta^\alpha \|a_{ij}^\epsilon\|_{C_{\epsilon,X}^\alpha(B_{\epsilon,\delta})} \|w\phi\|_{C_{\epsilon,X}^{2,\alpha}(B_{\epsilon,\delta})}, \end{aligned} \tag{6.5}$$

where  $C_1$ , and  $\tilde{C}_1$  are stable as  $\epsilon \rightarrow 0$ . Similarly, if  $\phi$  is fixed, the Hölder norm of the second term in the representation formula (6.3) is bounded by

$$\begin{aligned} & \left\| \int X_{i_1}^\epsilon X_{i_2}^\epsilon \Gamma_{(x_0, t_0)}^\epsilon((x_0, t_0), (y, \tau))(f\phi(y, \tau) + wL\phi(y, \tau) \right. \\ & \quad \left. + 2a_{ij}^\epsilon X_i^\epsilon wX_j^\epsilon \phi) dy d\tau \right\|_{C_{\epsilon, X}^\alpha(B_{\epsilon, \delta})} \\ & \leq C_2 \left( \|f\|_{C_{\epsilon, X}^\alpha(K_\delta)} + \frac{C}{\delta^2} \|w\|_{C_{\epsilon, X}^{1, \alpha}(K_\delta)} \right). \end{aligned} \tag{6.6}$$

From (6.3), (6.5) and (6.6) we deduce that

$$\|w\phi\|_{C_{\epsilon, X}^{2, \alpha}(B_\delta)} \leq \tilde{C}_2 \delta^\alpha \|w\phi\|_{C_{\epsilon, X}^{2, \alpha}(B_\delta)} + C_2 \left( \|f\|_{C_{\epsilon, X}^\alpha(K_\delta)} + \frac{C}{\delta^2} \|w\|_{C_{\epsilon, X}^{1, \alpha}(K_\delta)} \right).$$

Choosing  $\delta$  sufficiently small we prove the assertion on the fixed sphere  $B_{\epsilon, \delta}$ . The conclusion follows from a standard covering argument.  $\square$

We can now conclude the proof of Proposition 6.2:

*Proof* The proof is similar to the previous one for  $k = 1$ . We start by differentiating the representation formula (6.2) along an arbitrary direction  $X_{i_1}$

$$\begin{aligned} & X_{i_1}^\epsilon (w\phi)(x, t) \tag{6.7} \\ & = \int_Q X_{i_1}^\epsilon \Gamma_{(x_0, t_0)}^\epsilon(\cdot, (y, \tau)) (L_{\epsilon, (x_0, t_0)} - L_\epsilon) (w\phi)(y, \tau) dy d\tau \\ & \quad + \int_Q X_{i_1}^\epsilon \Gamma_{(x_0, t_0)}^\epsilon(\cdot, (y, \tau)) \left( f\phi + wL_\epsilon\phi + 2 \sum_{i, j=1}^n a_{ij}^\epsilon X_i^\epsilon wX_j^\epsilon \phi \right) (y, \tau) dy d\tau. \end{aligned} \tag{6.8}$$

Now we apply Theorem 5.10 and deduce that there exist kernels

$$P_{e, i_1, h, (x_0, t_0)}((x, t), (y, \tau)), R_{e, i_1, (x_0, t_0)}((x, t), (y, \tau)),$$

with the same decay of the fundamental solution such that

$$\begin{aligned} & X_{i_1}^\epsilon (w\phi)(x, t) \\ & = - \int \sum_{h=1}^m P_{e, i_1, h, (x_0, t_0)}((x, t), (y, \tau)) X_h^{\epsilon, y} (L_{\epsilon, (x_0, t_0)} - L_\epsilon) (w\phi)(y, \tau) dy d\tau \\ & \quad - \int R_{e, i_1, (x_0, t_0)}((x, t), (y, \tau)) (L_{\epsilon, (x_0, t_0)} - L_\epsilon) (w\phi)(y, \tau) dy d\tau dy \\ & \quad - \sum_{h=1}^m \int P_{e, i_1, h, (x_0, t_0)}((x, t), (y, \tau)) X_h^{\epsilon, y} (f\phi + wL_\epsilon\phi \end{aligned}$$

$$\begin{aligned}
 & + 2 \sum_{i,j=1}^n a_{ij}^\epsilon X_i^\epsilon w X_j^\epsilon \phi \Big) (y, \tau) dy d\tau - \int R_{\epsilon, i_1, (x_0, t_0)}((x, t), (y, \tau)) \\
 & \times \left( f\phi + wL_\epsilon\phi + 2 \sum_{i,j=1}^n a_{ij}^\epsilon X_i^\epsilon w X_j^\epsilon \phi \right) (y, \tau) dy d\tau. \tag{6.9}
 \end{aligned}$$

Using Lemma 6.4, this yields the existence of new kernels  $P_{\epsilon, h_1, \dots, h_k, (x_0, t_0)}^{i_1, \dots, i_k}((x, t), (y, \tau))$  with the behavior of a fundamental solution (and the same dependence on  $\epsilon$ ) such that

$$\begin{aligned}
 X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon (w\phi)(x, t) &= \int_{\epsilon, h_1, \dots, h_k, (x_0, t_0)}^{i_1, \dots, i_k} ((x, t), (y, \tau)) \\
 & \left( a_{ij}^\epsilon - a_{ij}^\epsilon(x_0, t_0) \right) X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon X_i^\epsilon X_j^\epsilon (w\phi)(y, \tau) dy d\tau \\
 & + \int_{\epsilon, h_1, \dots, h_k, (x_0, t_0)}^{i_1, \dots, i_k} ((x, t), (y, \tau)) B(w\phi)(y, \tau) dy d\tau \\
 & + \int_{\epsilon, h_1, \dots, h_k, (x_0, t_0)}^{i_1, \dots, i_k} ((x, t), (y, \tau)) X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon \left( f\phi(y, \tau) + wL_\epsilon\phi(y, \tau) \right. \\
 & \left. + 2 \sum_{i,j=1}^n a_{ij}^\epsilon X_i^\epsilon w X_j^\epsilon \phi \right) dy d\tau \\
 & + \text{lower order terms}
 \end{aligned}$$

where  $\phi$  is as in the proof of Proposition 6.5 and  $B$  is a differential operator of order  $k + 1$ . The conclusion follows by further differentiating the previous representation formula along two horizontal directions  $X_{j_1}^\epsilon X_{j_2}^\epsilon$  and arguing as in the proof of Proposition 6.5. □

### 7 Application I: Harnack inequalities for degenerate parabolic quasilinear equations hold uniformly in $\epsilon$

The results we have presented so far show that for any  $\epsilon_0 > 0$ , the 1-parameter family of metric spaces  $(\mathcal{M}, d_\epsilon)$  associated to the Riemannian approximations of a subRiemannian metric space  $(\mathcal{M}, d_0)$ , satisfy uniformly in  $\epsilon \in [0, \epsilon_0]$  the hypothesis in the definition of *p-admissible structure* in the sense of [48, Theorem 13.1]. This class of metric measure spaces has a very rich analytic structure (Sobolev-Poincaré inequalities, John–Nirenberg lemma, ...) that allows for the development of a basic regularity theory for weak solutions of classes of nonlinear degenerate parabolic and elliptic PDE. In the following we will remind the reader of the definition and basic properties of *p-admissible structures* and sketch some of the main regularity results from the recent papers [1] and [18]. We will conclude the section with a sample

application of these techniques to the global (in time) existence of solutions for a class of evolution equations that include the subRiemannian total variation flow [14].

### 7.1 Admissible ambient space geometry

Consider a smooth real manifold  $M$  and let  $\mu$  be a locally finite Borel measure on  $M$  which is absolutely continuous with respect the Lebesgue measure when represented in local charts. Let  $d(\cdot, \cdot) : M \times M \rightarrow \mathbb{R}^+$  denote the control distance generated by a system of bounded,  $\mu$ -measurable, Lipschitz vector fields  $\Xi = (X_1, \dots, X_m)$  on  $M$ . As in [3] and [44] one needs to assume as basic hypothesis

$$\text{the inclusion } i : (M, \text{chart}) \rightarrow (M, d) \text{ is continuous,} \tag{7.1}$$

where we have denoted by  $(M, \text{chart})$  the topology on  $M$  induced by the Euclidean topology in  $\mathbb{R}^n$  via coordinate charts. For  $x \in M$  and  $r > 0$ , set  $B(x, r) = \{y \in M : d(x, y) < r\}$  and let  $|B(x, r)|$  denote the  $\mu$  measure of  $B(x, r)$ . In general, given a function  $u$  and a ball  $B = B(x, r)$  then  $u_B$  denotes the  $\mu$ -average of  $u$  on the ball  $B = B(x, r)$ . In view of (7.1) the closed metric ball  $\bar{B}$  is a compact set.

**Definition 7.1** Assume hypothesis (7.1) holds. Given  $1 \leq p < \infty$ , the triplet  $(M, \mu, d)$  is said to define a  $p$ -admissible structure (in the sense of [48, Theorem 13.1]) if for every compact subset  $K$  of  $M$  there exist constants  $C_D = C_D(\Xi, K)$ ,  $C_P = C_P(\Xi, K) > 0$ , and  $R = R(\Xi, K) > 0$ , such that the following hold.

(1) *Doubling property:*

$$|B(x, 2r)| \leq C_D |B(x, r)| \text{ whenever } x \in K \text{ and } 0 < r < R. \tag{D}$$

(2) *Weak (1, p)-Poincaré inequality:*

$$\int_{B(x,r)} |u - u_B| d\mu \leq C_P r \left( \int_{B(x,2r)} |\Xi u|^p d\mu \right)^{1/p}, \tag{P}$$

whenever  $x \in K$ ,  $0 < r < R$ ,  $u \in W_{\Xi}^{1,p}(B(x, 2r))$ .

Theorems 1.1 and 1.2 yield the following

**Theorem 7.2** Let  $X_1, \dots, X_m$  be a family of Hörmander vector fields in  $\Omega \subset \mathbb{R}^n$  and denote by  $\mu$  Lebesgue measure. For each  $\epsilon_0 > 0$  and  $\epsilon \in [0, \epsilon_0]$  denote by  $d_\epsilon$  the distance functions defined in Definition 2.6. For all  $\epsilon \in [0, \epsilon_0]$  and  $p \geq 1$ , the space  $(\Omega, \mu, d_\epsilon)$  is  $p$ -admissible, with constants  $C_D$  and  $C_P$  independent of  $\epsilon$ .

Other examples of  $p$ -admissible spaces are:

- The classical setting:  $\mathcal{M} = \mathbb{R}^n$ ,  $d\mu$  equals the  $n$ -dimensional Lebesgue measure, and  $\Xi = (X_1, \dots, X_m) = (\partial_{x_1}, \dots, \partial_{x_n})$ .

- Our setting is also sufficiently broad to include some non-smooth vector fields such as the Baouendi-Grushin frames, e.g., consider, for  $\gamma \geq 1$  and  $(x, y) \in \mathbb{R}^2$ , the vector fields  $X_1 = \partial_x$  and  $X_2 = |x|^\gamma \partial_y$ . Unless  $\gamma$  is a positive even integer these vector fields fail to satisfy Hörmander’s finite rank hypothesis. However, the doubling inequality as well as the Poincaré inequality hold and have been used in the work of Franchi and Lanconelli [38] to establish Harnack inequalities for linear equations.
- Consider a smooth manifold  $\mathcal{M}$  endowed with a complete Riemannian metric  $g$ . Let  $\mu$  denote the Riemann volume measure, and by  $\Xi$  denote a  $g$ -orthonormal frame. If the Ricci curvature is bounded from below ( $Ricci \geq -Kg$ ) then one has a 2-admissible structure. In fact, in this setting the Poincaré inequality follows from Buser’s inequality while the doubling condition is a consequence of the Bishop–Gromov comparison principle. If  $K = 0$ , i.e. the Ricci tensor is non-negative, then these assumptions holds globally and so does the Harnack inequality.

Spaces with a  $p$ -admissible structure support a homogenous structure in the sense of Coifman and Weiss [28].

**Lemma 7.3** *Let  $(\mathcal{M}, \mu, d)$  be a  $p$ -admissible structure for some  $p \geq 1$ ,  $\Omega$  a bounded open set in  $\mathcal{M}$  and set  $K = \widehat{\Omega}$ . If  $x \in K$  and  $0 < s < r < R$ , then the following holds.*

- (1) *There exists a constant  $N = N(C_D) > 0$ , called homogeneous dimension of  $K$  with respect to  $(\Xi, d, \mu)$ , such that  $|B(x, r)| \leq C_D \tau^{-N} |B(x, \tau r)|$ , for all  $0 < \tau \leq 1$ .*
- (2) *There exists a continuous function  $\phi \in C_0(B(x, r)) \cap W_\Xi^{1,\infty}(B(x, r))$  and a constant  $C = C(\Xi, K) > 0$ , such that  $\phi = 1$  in  $B(x, s)$  and  $|\Xi\phi| \leq C/(r - s)$ ,  $0 \leq \phi \leq 1$ .*
- (3) *Metric balls have the so called  $\hat{\delta}$ -annular decay property, i.e., there exists  $\hat{\delta} = \hat{\delta}(C_D) \in (0, 1]$ , such that*

$$|B(x, r) \setminus B(x, (1 - \epsilon)r)| \leq C\epsilon^{\hat{\delta}} |B(x, r)|,$$

whenever  $0 < \epsilon < 1$ .

*Proof* Statement (1) follows from (D) by a standard iteration argument. Statement (2) is proved in [44, Theorem 1.5]. Statement (3) follows from [12, Corollary 2.2], since we have a Carnot–Carathéodory space. Furthermore,  $\hat{\delta}$  depends only on  $C_D$ .  $\square$

Given  $\Omega \subset M$ , open, and  $1 \leq p \leq \infty$ , we let  $W_\Xi^{1,p}(\Omega) = \{u \in L^p(\Omega, \mu) : X_i u \in L^p(\Omega, \mu), i = 1, \dots, m\}$  denote the horizontal Sobolev space, and we let  $W_{\Xi,0}^{1,p} \subset W_\Xi^{1,p}$  be the closure<sup>4</sup> of the space of  $W_\Xi^{1,p}$  functions with compact (distributional) support in the norm  $\|u\|_{1,p}^p = \|u\|_p^p + \|\Xi u\|_p^p$  with respect to  $\mu$ . In the following we will omit  $\mu$  in the notation for Lebesgue and Sobolev spaces. Given  $t_1 < t_2$ , and  $1 \leq p \leq \infty$ , we let  $\Omega_{t_1,t_2} \equiv \Omega \times (t_1, t_2)$  and we let  $L^p(t_1, t_2; W_\Xi^{1,p}(\Omega))$ ,  $t_1 < t_2$ ,

<sup>4</sup> Here we avoid the issue “ $H = W$ ”. This is studied in detail in [39,40,42,44,54] and [79].

denote the parabolic Sobolev space of real-valued functions defined on  $\Omega_{t_1, t_2}$  such that for almost every  $t$ ,  $t_1 < t < t_2$ , the function  $x \rightarrow u(x, t)$  belongs to  $W_{\Xi}^{1,p}(\Omega)$  and

$$\|u\|_{L^p(t_1, t_2; W_{\Xi}^{1,p}(\Omega))} = \left( \int_{t_1}^{t_2} \int_{\Omega} (|u(x, t)|^p + |\Xi u(x, t)|^p) d\mu dt \right)^{1/p} < \infty.$$

The spaces  $L^p(t_1, t_2; W_{\Xi, 0}^{1,p}(\Omega))$  is defined analogously. We let  $W^{1,p}(t_1, t_2; L^p(\Omega))$  consist of real-valued functions  $\eta \in L^p(t_1, t_2; L^p(\Omega))$  such that the weak derivative  $\partial_t \eta(x, t)$  exists and belongs to  $L^p(t_1, t_2; L^p(\Omega))$ . Consider the set of functions  $\phi, \phi \in W^{1,p}(t_1, t_2; L^p(\Omega))$ , such that the functions

$$t \rightarrow \int_{\Omega} |\phi(x, t)|^p d\mu(x) \text{ and } t \rightarrow \int_{\Omega} |\partial_t \phi(x, t)|^p d\mu(x),$$

have compact support in  $(t_1, t_2)$ . We let  $W_0^{1,p}(t_1, t_2; L^p(\Omega))$  denote the closure of this space under the norm in  $W^{1,p}(t_1, t_2; L^p(\Omega))$ .

From [48, Corollary 9.5] one can see that the metric balls  $B(x_0, r)$  are John domains. Consequently, (D), (P), and [48, Theorem 9.7] yield Sobolev-Poincaré inequality,

**Lemma 7.4** *Let  $B(x_0, r) \subset \Omega$ ,  $0 < r < R$ ,  $1 \leq p < \infty$ . There exists a constant  $C = C(C_D, C_P, p) \geq 1$  such that for every  $u \in W_{\Xi}^{1,p}(B(x_0, r))$ ,*

$$\left( \int_{B(x_0, r)} |u - u_B|^{\kappa p} d\mu \right)^{1/\kappa} \leq Cr^p \int_{B(x_0, r)} |\Xi u|^p d\mu,$$

where  $u_B$  denotes the  $\mu$  average of  $u$  over  $B(x_0, r)$ , and where  $1 \leq \kappa \leq N/(N - p)$ , if  $1 \leq p < N$ , and  $1 \leq \kappa < \infty$ , if  $p \geq N$ . Moreover,

$$\left( \int_{B(x_0, r)} |u|^{\kappa p} d\mu \right)^{1/\kappa} \leq Cr^p \int_{B(x_0, r)} |\Xi u|^p d\mu,$$

whenever  $u \in W_{\Xi, 0}^{1,p}(B(x_0, r))$ .

### 7.2 Quasilinear degenerate parabolic PDE

In this section we list some recent results concerning regularity of weak solutions of certain nonlinear, degenerate parabolic PDE in spaces  $(M, \mu, d)$  that are  $p$ -admissible for some  $p \in [2, \infty)$ . If  $p = 2$  we can allow lower order terms, but at present this is not yet established for  $p > 2$ . Given a domain (i.e., an open, connected set)  $\Omega \subset M$ , and  $T > 0$  we set  $\Omega_T = \Omega \times (0, T)$ . For a function  $u : \Omega_T \rightarrow \mathbb{R}$ , and  $1 \leq p, q$  we define the norms

$$\|u\|_{p, q}^q = \left( \int_0^T \left( \int_{\Omega} |u|^p dx \right)^{\frac{q}{p}} dt \right)^{\frac{1}{q}}, \tag{7.2}$$



and the corresponding Lebesgue spaces  $L^{p,q}(\Omega_T) = L^q([0, T], L^p(\Omega))$ . We will say that  $\mathcal{A}, \mathcal{B}$  are *admissible symbols* (in  $\Omega_T$ ) if the following holds:

- (i)  $(x, t) \rightarrow \mathcal{A}(x, t, u, \xi), \mathcal{B}(x, t, u, \xi)$  are measurable for every  $(u, \xi) \in \mathbb{R} \times \mathbb{R}^m$ ,
- (ii)  $(u, \xi) \rightarrow \mathcal{A}(x, t, u, \xi), \mathcal{B}(x, t, u, \xi)$  are continuous for almost every  $(x, t) \in \Omega_T$ ,
- (iii) • For  $p = 2$ : There exist constants  $a, \bar{a} > 0$  and functions  $b, c, e, f, h \in L^{\tilde{p},q}(Q)$  with  $\tilde{p} > 2$ , and  $q$  given by  $\frac{N}{2\tilde{p}} + \frac{1}{q} < \frac{1}{2}$  and functions  $d, g \in L^{\alpha,\beta}(Q)$  with  $1 < \alpha$  and  $\beta$  given by  $\frac{N}{2\alpha} + \frac{1}{\beta} < 1$  such that for a.e.  $(x, t) \in \Omega_T$  and  $\xi \in \mathbb{R}^m$  one has

$$\begin{aligned} \sum_{i=1}^m \mathcal{A}_i(x, t, u, \xi) \xi_i &\geq a|\xi|^2 - b^2 u^2 - f^2, \\ |\mathcal{A}(x, t, u, \xi)| &\leq \bar{a}|\xi| + e|u| + h, \\ |\mathcal{B}(x, t, u, \xi)| &\leq c|\xi| + d|u| + g. \end{aligned} \tag{7.3}$$

In view of the conditions on  $\tilde{p}, q, \alpha, \beta$  there exists  $\theta > 0$  such that

$$\begin{aligned} \tilde{p} &\geq \frac{2}{1-\theta} \quad \text{and} \quad \frac{N}{2\tilde{p}} + \frac{1}{q} \leq \frac{1-\theta}{2} \\ \alpha &\geq \frac{1}{1-\theta} \quad \text{and} \quad \frac{N}{2\alpha} + \frac{1}{\beta} \leq 1-\theta. \end{aligned} \tag{7.4}$$

We say that *a constant depends on the structure conditions (7.3)*, if it depends only on <sup>5</sup>

$$a, \bar{a}, \|b\|, \|c\|, \|d\|, \|e\|, \|f\|, \|g\|, \|h\|, N, \theta,$$

and is uniformly bounded if these quantities are so.

- For  $p > 2$  we will only consider  $\mathcal{B} = 0$  and ask that the following bounds

$$\mathcal{A}(x, t, u, \xi) \cdot \xi \geq \mathcal{A}_0 |\xi|^p, \quad |\mathcal{A}(x, t, u, \xi)| \leq \mathcal{A}_1 |\xi|^{p-1}, \tag{7.5}$$

hold for every  $(u, \xi) \in \mathbb{R} \times \mathbb{R}^m$  and almost every  $(x, t) \in \Omega_T$ .

$\mathcal{A}_0$  and  $\mathcal{A}_1$  are called the structural constants of  $\mathcal{A}$ . If  $\mathcal{A}$  and  $\tilde{\mathcal{A}}$  are both admissible symbols, with the same structural constants  $\mathcal{A}_0$  and  $\mathcal{A}_1$ , then we say that the symbols are structurally similar.

Let  $E$  be a domain in  $\mathcal{M} \times \mathbb{R}$ . We say that the function  $u : E \rightarrow \mathbb{R}$  is a weak solution to

$$\partial_t u(x, t) = L_{A,p} u \equiv - \sum_{i=1}^m X_i^* \mathcal{A}_i(x, t, u, \Xi u) + \mathcal{B}(x, t, u, \Xi u), \tag{7.6}$$

---

<sup>5</sup> The  $\|\cdot\|$  norms are in the appropriate  $L^{\tilde{p},q}$  or  $L^{\alpha,\beta}$  classes.

in  $E$ , where  $X_t^*$  is the formal adjoint w.r.t.  $d\mu$ , if whenever  $\Omega_{t_1, t_2} \Subset E$  for some domain  $\Omega \subset \mathcal{M}$ ,  $u \in L^p(t_1, t_2; W_{\Xi}^{1,p}(\Omega))$  and

$$\begin{aligned} & \int_{t_1}^{t_2} \int_{\Omega} u \frac{\partial \eta}{\partial t} d\mu dt - \int_{t_1}^{t_2} \int_{\Omega} \mathcal{A}(x, t, u, \Xi u) \cdot \Xi \eta d\mu dt \\ & + \int_{t_1}^{t_2} \int_{\Omega} \mathcal{B}(x, t, u, \Xi u) \eta d\mu dt = 0, \end{aligned} \tag{7.7}$$

for every test function

$$\eta \in W_0^{1,2}(t_1, t_2; L^2(\Omega)) \cap L^p(t_1, t_2; W_{\Xi,0}^{1,p}(\Omega)).$$

A function  $u$  is a weak super-solution (sub-solution) to (7.6) in  $E$  if whenever  $\Omega_{t_1, t_2} \Subset E$  for some domain  $\Omega \subset \mathcal{M}$ , we have  $u \in L^p(t_1, t_2; W^{1,p}(\Omega))$ , and the left hand side of (7.7) is non-negative (non-positive) for all non-negative test functions  $W_0^{1,2}(t_1, t_2; L^2(\Omega)) \cap L^p(t_1, t_2; W_{\Xi,0}^{1,p}(\Omega))$ .

The main results in [1] and [18] can be summarized in the following theorem.

**Theorem 7.5** *Let  $(\mathcal{M}, \mu, d)$  be a  $p$ -admissible structure for some fixed  $p \in [2, \infty)$ . For a bounded open subset  $\Omega \subset \mathcal{M}$ , let  $u$  be a non-negative, weak solution to (7.6) in an open set containing the cylinder  $\Omega \times [0, T_0]$  and assume that the structure conditions (7.5) are satisfied.*

- For  $p = 2$  and for any subcylinder  $Q_{3\rho} = B(\bar{x}, 3\rho) \times (\bar{t} - 9\rho^2, \bar{t}) \subset Q$  there exists a constant  $C > 0$  depending on  $C_D, C_L, C_P$ , the structure conditions (7.3) and on  $\rho$  such that

$$\sup_{Q^-} u \leq C \inf_{Q^+} (u + \rho^\theta k), \tag{7.8}$$

where

$$Q^+ = B(x, \rho) \times (\bar{t} - \rho^2, \bar{t}) \text{ and } Q^- = B(x, \rho) \times (\bar{t} - 8\rho^2, \bar{t} - 7\rho^2) \tag{7.9}$$

$\theta > 0$  is defined as in (7.4), and we have let  $k = \|f\| + \|g\| + \|h\|$ .

- For  $p > 2$ : Assuming  $\mathcal{B} = 0$ , there exist constants  $C_1, C_2, C_3 \geq 1$ , depending only on  $\Xi, C_D, C_P, \mathcal{A}_0, \mathcal{A}_1, p$ , such that for almost all  $(x_0, t_0) \in \Omega \times [0, T_0]$ , the following holds: If  $u(x_0, t_0) > 0$ , and if  $0 < r \leq R(\Xi, \bar{\Omega})$  (from Definition 7.1) is sufficiently small so that

$$B(x_0, 8r) \subset \Omega \quad \text{and} \quad (t_0 - C_1 u(x_0, t_0)^{2-p} r^p, t_0 + C_1 u(x_0, t_0)^{2-p} r^p) \subset (0, T_0),$$

then

$$u(x_0, t_0) \leq C_2 \inf_Q u,$$

where

$$Q = B(x_0, r) \times \left( t_0 + \frac{1}{2} C_3 u(x_0, t_0)^{2-p} r^p, t_0 + C_3 u(x_0, t_0)^{2-p} r^p \right).$$

Furthermore, the constants  $C_1, C_2, C_3$  can be chosen independently of  $p$  as  $p \rightarrow 2$ .

We conclude this section with a corollary of the proof in [18] [Lemma 3.6], a weak Harnack inequality that plays an important role in the proof of the regularity of the mean curvature flow for graphs over certain Lie groups established in [14]. Consider a weak supersolution  $w \in L^p(t_1, t_2; W_{\Xi}^{1,p}(\Omega))$  of the linear equation

$$-\partial_t w - \sum_{i=1}^m X_i^*(a_{ij}(x, t)X_j w) = g(x, t), \tag{7.10}$$

with  $t_1, t_2, \Omega$  as defined above. Assume the coercivity hypothesis

$$\Lambda^{-1} \sum_{d(i)=1} \xi_i^2 \leq \sum_{i,j=1}^m a_{ij}(x, t)\xi_i\xi_j \leq \Lambda \sum_{d(i)=1} \xi_i^2 \tag{7.11}$$

for a.e.  $(x, t)$  and all  $\xi \in \mathbb{R}^m$ , for a suitable constant  $\Lambda$ .

**Proposition 7.6** *Let  $(\mathcal{M}, \mu, d)$  be a 2-admissible structure. For a bounded open subset  $\Omega \subset \mathcal{M}$ , let  $w$  be a non-negative, weak supersolution to (7.10) in an open set containing the cylinder  $\Omega \times [0, T_0]$  and assume that conditions (7.11) are satisfied. For any subcylinder  $Q_{3\rho} = B(\bar{x}, 3\rho) \times (\bar{t} - 9\rho^2, \bar{t}) \subset Q$  there exists a constant  $C > 0$  depending on  $C_D, C_L, C_P$ , the structure conditions (7.3) and on  $\rho$  such that*

$$\frac{1}{|Q^-|} \int_{Q^-} w \, dxdt \leq C \left( \inf_{Q^+} w + \sup_{Q^+} |g|\rho^2 \right), \tag{7.12}$$

with  $Q^+, Q^-$  as defined in (7.9).

## 8 Application II: regularity for quasilinear subelliptic PDE through Riemannian approximation

Let  $G$  be a Carnot group of step two. We denote by  $\mathfrak{g}$  its Lie algebra and by  $\mathfrak{g} = V^1 \oplus V^2$  its stratification. If  $g_0$  is a subRiemannian metric on  $V^1$  we let  $d_0$  denote its corresponding CC metric, and  $X_1, \dots, X_m$  a left-invariant orthonormal basis of  $V^1$ . Consider a smooth surface  $M \subset G$ . A point  $p \in M$  is characteristic if the horizontal distribution given by left translation of  $V^1$  is entirely contained in the tangent plane  $T_p M$ . Equivalently, if  $M$  is represented as a 0-level set of a function  $f$ , the points where the horizontal gradient of the defining function does vanish are called *characteristic*. At non-characteristic points several equivalent definitions for the notion of horizontal mean curvature  $h_0$  have been proposed. To quote a few:  $h_0$  can be defined in terms of the legendrian foliation [22], first variation of the area functional [19, 22, 30, 49, 75], as horizontal divergence of the horizontal unit normal or as limit of the mean curvatures  $h_\epsilon$  of suitable Riemannian approximating metrics

$\sigma_\epsilon$  [19]. If the surface is not regular, the notion of curvature can be expressed in the viscosity sense (we refer to [2,4,5,13,61,63,81,82] for viscosity solutions of PDE in the sub-Riemannian setting). The formulation we use here is the following, at every non-characteristic point  $p$  in the level set of  $f$  we set

$$h_0(p) = \sum_{i=1}^m X_i (X_i f / |\nabla_0 f|_0).$$

The mean curvature flow is the motion of a surface where each points is moving in the direction of the normal with speed equal to the mean curvature. In the total variation flow the speed is the mean curvature, scaled by the modulus of the gradient. Both flows play key roles in digital image processing as well as provide prototypes for modeling motion of interfaces in a variety of physical settings.

As an illustration of the usefulness of the uniform estimates established above, in this section we want to briefly sketch the strategy used in [14] and [17], where the Riemannian approximation scheme is used to establish regularity for the graph solutions of the Total Variation flow

$$\frac{\partial u}{\partial t} = \sum_{i=1}^m X_i \left( \frac{X_i u}{\sqrt{1 + |\nabla_0 u|^2}} \right), \tag{8.1}$$

and for the graphical solutions of the mean curvature flow

$$\frac{\partial u}{\partial t} = \sqrt{1 + |\nabla_0 u|^2} \sum_{i=1}^m X_i \left( \frac{X_i u}{\sqrt{1 + |\nabla_0 u|^2}} \right). \tag{8.2}$$

In both cases  $\Omega \subset G$  is a bounded open set, with  $G$  is a Lie group, free up to step two, but not necessarily nilpotent. These equations describe the motions of the (non-characteristic) hypersurfaces given by the graphs of the solutions in  $G \times \mathbb{R}$ .

We will consider solutions arising as limits of solutions of the analogue Riemannian flows, i.e.

$$\frac{\partial u_\epsilon}{\partial t} = h_\epsilon = \sum_{i=1}^n X_i^\epsilon \left( \frac{X_i^\epsilon u_\epsilon}{W_\epsilon} \right) \quad \text{for } x \in \Omega, t > 0, \tag{8.3}$$

and

$$\frac{\partial u_\epsilon}{\partial t} = W_\epsilon h_\epsilon = W_\epsilon \sum_{i=1}^n X_i^\epsilon \left( \frac{X_i^\epsilon u_\epsilon}{W_\epsilon} \right) = \sum_{i,j=1}^n a_{ij}^\epsilon (\nabla_\epsilon u_\epsilon) X_i^\epsilon X_j^\epsilon u_\epsilon \quad \text{for } x \in \Omega, t > 0, \tag{8.4}$$

where,  $h_\epsilon$  is the mean curvature of the graph of  $u_\epsilon(\cdot, t)$  and

$$W_\epsilon^2 = 1 + |\nabla_\epsilon u_\epsilon|^2 = 1 + \sum_{i=1}^n (X_i^\epsilon u_\epsilon)^2 \quad \text{and } a_{ij}^\epsilon(\xi) = \delta_{ij} - \frac{\xi_i \xi_j}{1 + |\xi|^2}, \tag{8.5}$$

for all  $\xi \in \mathbb{R}^n$ .

The main results in [14] and [17] concern long time existence of solutions of the initial value problems

$$\begin{cases} \partial_t u_\epsilon = h_\epsilon W_\epsilon & \text{in } Q = \Omega \times (0, T) \\ u_\epsilon = \varphi & \text{on } \partial_p Q, \end{cases} \quad \text{and} \quad \begin{cases} \partial_t u_\epsilon = h_\epsilon & \text{in } Q = \Omega \times (0, T) \\ u_\epsilon = \varphi & \text{on } \partial_p Q, \end{cases} \tag{8.6}$$

with  $\partial_p Q = (\Omega \times \{t = 0\}) \cup (\partial\Omega \times (0, T))$  denoting the parabolic boundary of  $Q$ .

**Theorem 8.1** *Let  $G$  be a Lie group of step two,  $\Omega \subset G$  a bounded, open, convex set (in a sense to be defined later) and  $\varphi \in C^2(\bar{\Omega})$ . There exists unique solutions  $u_\epsilon \in C^\infty(\Omega \times (0, \infty)) \cap L^\infty((0, \infty), C^1(\bar{\Omega}))$  of the two initial value problems in (8.6), and for each  $k \in \mathbb{N}$  and  $K \subset\subset Q$ , there exists  $C_k = C_k(G, \varphi, k, K, \Omega) > 0$  not depending on  $\epsilon$  such that*

$$\|u_\epsilon\|_{C^k(K)} \leq C_k. \tag{8.7}$$

**Corollary 8.2** *Under the assumptions of Theorem 8.1, as  $\epsilon \rightarrow 0$  the solutions  $u_\epsilon$  of either flow converge uniformly (with all their derivatives) on compact subsets of  $Q$  to the unique, smooth solution of the corresponding sub-Riemannian flow in  $\Omega \times (0, \infty)$  with initial data  $\varphi$ .*

The proof of this result rests crucially on the estimates established in this paper. In the following we list the main steps. First of all we note that in view of the short time existence result in the Riemannian setting we can assume that locally  $u_\epsilon$  are smooth both in time and space.

### 8.1 Interior gradient bounds

Denote by right  $X_i^r$  the left invariant frame corresponding to  $X_i^l$ s and observe that these two frames commute. For both flows, consider solutions  $u_\epsilon \in C^3(Q)$  and denote  $v_0 = \partial_t u_\epsilon$ ,  $v_i = X_i^r u_\epsilon$  for  $i = 1, \dots, n$ . Then for every  $h = 0, \dots, n$  one has that  $v_h$  is a solution of

$$\partial_t v_h = X_i^\epsilon (a_{ij} X_j v_h) = a_{ij}^\epsilon (\nabla_\epsilon u_\epsilon) X_i^\epsilon X_j^\epsilon v_h + a^{i,j,h} (\nabla_\epsilon u) X_i^\epsilon X_j^\epsilon u_\epsilon X_k^\epsilon v_h, \tag{8.8}$$

where

$$a^{i,j,h}(p) = \frac{\partial a_{ij}^\epsilon}{\partial p_h}, \quad a_{ij}^\epsilon(p) = \frac{1}{\sqrt{1 + |p|^2}} \left( \delta_{ij} - \frac{p_i p_j}{1 + |p|^2} \right),$$

for the total variation flow, while

$$a^{i,j,h}(p) = \frac{\partial a_{ij}^\epsilon}{\partial p_h} - \frac{\partial a_{ih}^\epsilon}{\partial p_j}, \quad a_{ij}^\epsilon(p) = \delta_{ij} - \frac{p_i p_j}{1 + |p|^2},$$

for the mean curvature flow. The weak parabolic maximum principle yields that there exists  $C = C(G, \|\varphi\|_{C^2(\Omega)}) > 0$  such that for every compact subset  $K \subset\subset \Omega$  one has

$$\sup_{K \times [0, T]} |\nabla_1 u_\epsilon| \leq \sup_{\partial_p Q} (|\nabla_1 u_\epsilon| + |\partial_t u_\epsilon|),$$

where  $\nabla_1$  is the full  $g_1$ –Riemannian gradient. This yields the desired uniform interior gradient bounds. This argument works in any Lie group, with no restrictions on the step.

### 8.2 Global gradient bounds

The proof of the boundary gradient estimates is more delicate and depends crucially on the geometry of the space. In particular the argument we outline here only holds in step two groups  $G$  and for domains  $\Omega \subset G$  that are locally Euclidean convex when expressed in the Rothschild–Stein preferred coordinates. As usual we note that the function  $v_\epsilon = u_\epsilon - \varphi$  solves the homogenous ‘boundary’ value problem

$$\begin{cases} P(v_\epsilon) = 0 & \text{in } Q = \Omega \times (0, T) \\ v_\epsilon = 0 & \text{on } \partial_p Q, \end{cases} \tag{8.9}$$

with  $b^\epsilon(x) = a_{ij}^\epsilon(\nabla_\epsilon v_\epsilon(x) + \nabla_\epsilon \varphi(x))X_i^\epsilon X_j^\epsilon \varphi(x)$ . and

$$P(v) = a_{ij}^\epsilon(\nabla_\epsilon v_\epsilon + \nabla_\epsilon \varphi)X_i^\epsilon X_j^\epsilon v_\epsilon + b^\epsilon - \partial_t v. \tag{8.10}$$

Then we construct for each point  $p_0 = (x_0, t_0) \in \partial\Omega \times (0, T)$  a *barrier function*:

**Lemma 8.3** *Let  $G$  be a Carnot group of step two and  $\Omega \subset G$  convex in the Euclidean sense. For each point  $p_0 = (x_0, t_0) \in \partial\Omega \times (0, T)$  one can construct a positive function  $w \in C^2(Q)$  such that*

$$\begin{aligned} Q(w) &\leq 0 \text{ in } V \cap Q \text{ with } V \text{ a parabolic neighborhood of } p_0, \\ w(p_0) &= 0 \text{ and } w \geq v_\epsilon \text{ in } \partial_p V \cap Q. \end{aligned} \tag{8.11}$$

*Proof* In the hypothesis that  $\Omega$  is convex in the Euclidean sense we have that every  $x_0 \in \partial\Omega$  supports a tangent hyperplane  $\Pi$  defined by an equation of the form  $\Pi(x) = \sum_{i=1}^n a_i x_i = 0$  with  $\Pi > 0$  in  $\Omega$ ,  $\Pi(x_0) = 0$ , and normalized as  $\sum_{d(i)=1,2} a_i^2 = 1$ . Following the standard argument (see for instance [60, Chapter 10]) we prove that there exists a function  $\Phi$  such that the barrier at  $(x_0, t_0) \in \partial\Omega \times (0, T)$  can be expressed in the form  $w = \Phi(\Pi)$ . □

Now comparison with the barrier constructed above yields that

**Proposition 8.4** *Let  $G$  be a Carnot group of step two,  $\Omega \subset G$  a bounded, open, convex (in the Euclidean sense) set and  $\varphi \in C^2(\bar{\Omega})$ . For  $\epsilon > 0$  denote by  $u_\epsilon \in$*

$C^2(\Omega \times (0, T)) \cap C^1(\bar{\Omega} \times (0, T))$  the non-negative unique solution of the initial value problem (8.6). There exists  $C = C(G, \|\varphi\|_{C^2(\bar{\Omega})}) > 0$  such that

$$\sup_{\partial\Omega \times (0, T)} |\nabla_\epsilon u_\epsilon| \leq \sup_{\partial\Omega \times (0, T)} |\nabla_1 u_\epsilon| \leq C. \tag{8.12}$$

*Proof*

$$0 \leq \frac{v_\epsilon(x, t)}{\text{dist}_{\sigma_1}(x, x_0)} \leq \frac{w(x, t)}{\text{dist}_{\sigma_1}(x, x_0)} \leq C(k, v), \tag{8.13}$$

in  $V \cap Q$ , with  $\text{dist}_{\sigma_1}(x, x_0)$  being the distance between  $x$  and  $x_0$  in the Riemannian metric  $\sigma_1$ , concluding the proof of the boundary gradient estimates.  $\square$

### 8.3 Harnack inequalities and $C^{1,\alpha}$ estimates

Let us first recall that  $(G, d_\epsilon)$  is a 2-admissible geometry in the sense of Definition 7.1, with Doubling and Poincare constants uniform in  $\epsilon \geq 0$ , as we proved in Theorems 1.1 and 1.2.

The total variation equation is expressed in divergence form, hence also the Eq. (8.8) satisfied by the right derivatives  $v_h = X_h^r u_\epsilon$  is in the same form. The mean curvature flow is not in divergence form. However, arguing as in [60] it is possible to show that there exists a regular, positive and strictly monotone function  $\Phi$  such that  $\Phi(v_h)$  satisfies a divergence form equation. As a consequence we can apply the Harnack inequalities in Theorem 7.5 and Proposition 7.6 to the bounded solutions  $v_h$  (or  $\Phi(v_h)$ ), thus yielding the  $C^{1,\alpha}$  uniform interior estimates.

**Corollary 8.5** *Letting  $K$  be a compact set  $K \subset\subset Q$ , there exist constants  $\alpha \in (0, 1)$  and  $C = C(K, \alpha) > 0$  such that for all  $i = 1, \dots, n$  one has that  $v = X_i^\epsilon u$  satisfies*

$$\|v\|_{C_{\epsilon, X}^\alpha(K)} + \|\nabla_\epsilon v\|_{L^2(K)} \leq C,$$

*uniformly in  $\epsilon \in (0, 1)$ .*

### 8.4 Schauder estimates and higher order estimates

The uniform Gaussian estimates and Schauder estimates in Theorem 1.4 applied to (8.8) yield the higher order estimates and conclude the proof. Once obtained the interior  $C^{1,\alpha}$  estimate of the solution uniform in  $\epsilon$ , we write the mean curvature flow equation in non divergence form:

$$\partial_t u_\epsilon - \sum_{i,j=1}^n a_{ij}^\epsilon(x, t) X_i^\epsilon X_j^\epsilon u_\epsilon = 0.$$

Applying Schauder estimates in Proposition 6.2 we immediately deduce the proof of Theorem 8.1.

*Proof of Theorem 8.1* Since the solution is of class  $C^{1,\alpha}$ , and the norm is bounded uniformly in  $\epsilon$  then  $u_\epsilon$  it is a solution of a divergence form equation

$$\partial_t u_\epsilon - \sum_{i,j=1}^n a_{ij}^\epsilon(x, t) X_i^\epsilon X_j^\epsilon u_\epsilon = 0,$$

with  $a_{ij}^\epsilon$  of class  $C^\alpha$  such that for every  $K$  be a compact sets such that  $K \subset\subset Q$  and  $2\delta = d_0(K, \partial_p Q)$  there exists a positive constant  $C_0$  such that

$$\|a_{ij}^\epsilon\|_{C_{\epsilon,X}^\alpha(K_\delta)} \leq C_0,$$

for every  $\epsilon \in (0, 1)$ . Consequently, by Proposition 6.2 there exists a constant  $C_2$  such that

$$\|u_\epsilon\|_{C^2(K)} \leq C_2.$$

We now prove by induction that for every  $k \in N$  and for every compact set  $K \subset\subset Q$  there exists a positive constant  $C$  such that

$$\|u_\epsilon\|_{C_{\epsilon,X}^{k,\alpha}(K)} \leq C, \tag{8.14}$$

for every  $\epsilon > 0$ . The assertion is true if  $k = 2$ , by Proposition 6.5. If (8.14) is true for  $k + 1$ , then the coefficients  $a_{ij}^\epsilon$  belong to  $C_{\epsilon,X}^{k,\alpha}$  uniformly as  $\epsilon \in (0, 1)$  and (8.14) holds at the level  $k + 2$  by virtue of Proposition 6.2.  $\square$

**Acknowledgments** The authors would like to acknowledge Maria Manfredini, who collaborated with us in developing and applying large part of the stability results presented in this paper. We are also grateful to Paul Yang for many conversations on these topics over the years.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

1. Avelin, B., Capogna, L., Citti, G., Nyström, K.: Harnack estimates for degenerate parabolic equations modeled on the subelliptic  $p$ -Laplacian. *Adv. Math.* **257**, 25–65 (2014)
2. Balogh, Z.M., Rickly, M.: Regularity of convex functions on Heisenberg groups. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **24**, 847–868 (2003)
3. Bellaïche, A., Risler, Eds.: *Sub-Riemannian Geometry*, vol. 144 of *Progress in Mathematics*. Birkhäuser, Basel (1996)
4. Bieske, T.: On  $\infty$ -harmonic functions on the Heisenberg group. *Commun. Partial Differ. Equ.* **3–4**(27), 727–761 (2002)
5. Bieske, T.: Comparison principle for parabolic equations in the Heisenberg group. *Electron. J. Differ. Equ.* **95**, 11 (2005). (electronic)
6. Biroli, M., Mosco, U.: Sobolev and isoperimetric inequalities for Dirichlet forms on homogeneous spaces. *Atti Accad. Naz. Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl.* **61**, 37–44 (1995)



7. Bonfiglioli, A., Lanconelli, E., Uguzzoni, F.: Uniform Gaussian estimates for the fundamental solutions for heat operators on Carnot groups. *Adv. Differ. Equ.* **7**, 1153–1192 (2002)
8. Bonfiglioli, A., Lanconelli, E., Uguzzoni, F.: Fundamental solutions for non-divergence form operators on stratified groups. *Trans. Am. Math. Soc.* **356**(7), 2709–2737 (2004)
9. Bonfiglioli, A., Lanconelli, E., Uguzzoni, F.: *Stratified Lie Groups and Potential Theory for their Sub-Laplacians*. Springer, New York (2007)
10. Bramanti, M., Brandolini, L.: Schauder estimates for parabolic nondivergence operators of Hörmander type. *J. Differ. Equ.* **234**(1), 177–245 (2007)
11. Bramanti, M., Brandolini, L., Lanconelli, E., Uguzzoni, F.: Non-divergence equations structured on Hörmander vector fields: heat kernels and harnack inequalities. *Mem. Am. Math. Soc.* **204**, 961 (2010)
12. Buser, P.: Isospectral riemann surfaces. *Ann. Inst. Fourier (Grenoble)* **36**, 167–192 (1986)
13. Capogna, L., Citti, G.: Generalized mean curvature flow in Carnot groups. *Commun. Partial Differ. Equ.* **34**(7–9), 937–956 (2009)
14. Capogna, L., Citti, G., Manfredini, M.: Regularity of mean curvature flow of graphs on Lie groups free up to step 2. *Nonlinear Anal.* **126**, 437–450 (2015)
15. Capogna, L., Citti, G., Manfredini, M.: Regularity of non-characteristic minimal graphs in the Heisenberg group  $\mathbb{H}^1$ . *Indiana Univ. Math. J.* **58**(5), 2115–2160 (2009)
16. Capogna, L., Citti, G., Manfredini, M.: Smoothness of lipschitz minimal intrinsic graphs in heisenberg groups  $\mathbb{H}^n$ ,  $n \geq 1$ . *Crelle's J.* **648**, 75110 (2010)
17. Capogna, L., Citti, G., Manfredini, M.: Uniform Gaussian bounds for subelliptic heat kernels and an application to the total variation flow of graphs over Carnot groups. *Anal. Geom. Metric. Spaces* **1**, 255–275 (2013)
18. Capogna, L., Citti, G., Rea, G.: A subelliptic analogue of Aronson–Serrin’s Harnack inequality. *Math. Ann.* **357**(3), 1175–1198 (2013)
19. Capogna, L., Danielli, D., Pauls, S., Tyson, J.: *An Introduction to the Heisenberg Group and the sub-Riemannian Isoperimetric Problem*, vol. 259 of *Progress in Mathematics*. Birkhäuser Verlag, Basel (2007)
20. Cheng, J.-H., Hwang, J.-F.: Properly embedded and immersed minimal surfaces in the Heisenberg group. *Bull. Aust. Math. Soc.* **70**(3), 507–520 (2004)
21. Cheng, J.-H., Hwang, J.-F.: Uniqueness of generalized  $p$ -area minimizers and integrability of a horizontal normal in the Heisenberg group. *Calc. Var. Partial Differ. Equ.* **50**(3–4), 579–597 (2014)
22. Cheng, J.-H., Hwang, J.-F., Malchiodi, A., Yang, P.: Minimal surfaces in pseudohermitian geometry. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **4**, 129–177 (2005)
23. Cheng, J.-H., Hwang, J.-F., Malchiodi, A., Yang, P.: A Codazzi-like equation and the singular set for  $C^1$  smooth surfaces in the Heisenberg group. *J. Reine Angew. Math.* **671**, 131–198 (2012)
24. Cheng, J.-H., Hwang, J.-F., Yang, P.: Existence and uniqueness for  $p$ -area minimizers in the Heisenberg group. *Math. Ann.* **337**(2), 253–293 (2007)
25. Cheng, J.-H., Hwang, J.-F., Yang, P.: Regularity of  $C^1$  smooth surfaces with prescribed  $p$ -mean curvature in the Heisenberg group. *Math. Ann.* **344**(1), 1–35 (2009)
26. Cheng, J.-H., Malchiodi, A., Yang, P.: Isoperimetric domains in homogeneous three-manifolds and the isoperimetric constant of the heisenberg group  $h^1$ . Preprint (2015)
27. Citti, G., Manfredini, M.: Uniform estimates of the fundamental solution for a family of hypoelliptic operators. *Potential Anal.* **25**(2), 147–164 (2006)
28. Coifman, R.R., Weiss, G.: *Analyse harmonique non-commutative sur certains espaces homogènes*. (French) Étude de certaines intégrales singulières. *Lecture Notes in Mathematics*, Vol. 242. Springer, Verlag (1971)
29. Corwin, L., Greenleaf, F.P.: *Representations of Nilpotent Lie Groups and Their Applications, Part I: Basic Theory and Examples*. Cambridge Studies in Advanced Mathematics. Cambridge University Press, Cambridge (1990)
30. Danielli, D., Garofalo, N., Nhieu, D.-M.: Sub-Riemannian calculus on hypersurfaces in Carnot groups. *Adv. Math.* **215**(1), 292–378 (2007)
31. Danielli, D., Garofalo, N., Nhieu, D.-M., Pauls, S.D.: Instability of graphical strips and a positive answer to the Bernstein problem in the Heisenberg group  $\mathbb{H}^1$ . *J. Differ. Geom.* **81**(2), 251–295 (2009)
32. Danielli, D., Garofalo, N., Nhieu, D.-M., Pauls, S.D.: The Bernstein problem for embedded surfaces in the Heisenberg group  $\mathbb{H}^1$ . *Indiana Univ. Math. J.* **59**(2), 563–594 (2010)
33. Debiard, A.: Espaces  $h^p$  au dessus de l’espace hermitien hyperbolique de  $e^n$  ( $n \geq 1$ ). *J. Funct. Anal.* **40**, 185–265 (1981)

34. Di Francesco, M., Polidoro, S.: Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form. *Adv. Differ. Equ.* **11**(11), 1261–1320 (2006)
35. Dirr, N., Dragoni, F., Von Renesse, M.: Evolution by mean curvature flow in sub-riemannian geometries. *Commun. Pure Appl. Anal.* **9**(2), 307–326 (2010)
36. Ferrari, F., Manfredi, J., Liu, Q.: On the horizontal mean curvature flow for axisymmetric surfaces in the heisenberg group. *Commun. Contemp. Math.* **16**(3), 1350027, 41 (2014)
37. Folland, G.B.: Subelliptic estimates and function spaces on nilpotent Lie groups. *Ark. Mat.* **2**(13), 161–207 (1975)
38. Franchi, B., Lanconelli, E.: Hölder regularity theorem for a class of linear nonuniformly elliptic operators with measurable coefficients. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4)* **10** **4**, 523–541 (1983)
39. Franchi, B., Serapioni, R., Cassano, F.S.: Meyers-Serrin type theorems and relaxation of variational integrals depending on vector fields. *Houston J. Math.* **22**(4), 859–890 (1996)
40. Franchi, B., Serapioni, R., Serra-Cassano, F.: Approximation and imbedding theorems for weighted sobolev spaces associated with lipschitz continuous vector fields. *Boll. Un. Mat. Ital. B* **7**(11), 83–117 (1997)
41. Friedman, A.: *Partial Differential Equations of Parabolic Type*. Prentice-Hall Inc, Englewood Cliffs, N.J. (1964)
42. Friedrichs, K.O.: The identity of weak and strong extensions of differential operators. *Trans. Am. Math. Soc.* 132–151 (1944)
43. Galli, M., Ritoré, M.: Existence of isoperimetric regions in contact sub-Riemannian manifolds. *J. Math. Anal. Appl.* **397**(2), 697–714 (2013)
44. Garofalo, N., Nhieu, D.M.: Lipschitz continuity, global smooth approximations and extension theorems for sobolev functions in Carnot–Carathodory spaces. *J. Anal. Math.* **74**, 67–97 (1998)
45. Ge, Z.: Collapsing Riemannian metrics to Carnot–Carathéodory metrics and Laplacians to sub-Laplacians. *Can. J. Math.* **45**(3), 537–553 (1993)
46. Grigor'yan, A.A.: The heat equation on non-compact riemannian manifolds. *Mat. Sb. (1)* 182 (1991)
47. Gutiérrez, C.E., Lanconelli, E.: Maximum principle, nonhomogeneous Harnack inequality, and Liouville theorems for  $X$ -elliptic operators. *Commun. Partial Differ. Equ.* **28**(11–12), 1833–1862 (2003)
48. Hajlasz, P., Koskela, P.: Sobolev met Poincaré. *Mem. Am. Math. Soc.* **145**, 688 (2000)
49. Hladky, R.K., Pauls, S.D.: Constant mean curvature surfaces in sub-Riemannian geometry. *J. Differ. Geom.* **79**(1), 111–139 (2008)
50. Hladky, R.K., Pauls, S.D.: Minimal surfaces in the roto-translation group with applications to a neurobiological image completion model. *J. Math. Imaging Vision* **36**(1), 1–27 (2010)
51. Hurtado, A., Ritoré, M., Rosales, C.: The classification of complete stable area-stationary surfaces in the Heisenberg group  $\mathbb{H}^1$ . *Adv. Math.* **224**(2), 561–600 (2010)
52. Jerison, D.: The Poincaré inequality for vector fields satisfying Hörmander's condition. *Duke Math. J.* **53**, 503–523 (1986)
53. Jerison, D.S., Sánchez-Calle, A.: Estimates for the heat kernel for a sum of squares of vector fields. *Indiana Univ. Math. J.* **35**(4), 835–854 (1986)
54. Kilpeläinen, T.: Smooth approximation in weighted Sobolev spaces. *Comment. Math. Univ. Carolin.* **38**(1), 29–35 (1997)
55. Korányi, A.: Geometric aspects of analysis on the Heisenberg group. *Ist. Naz. Alta Mat. Francesco Severi*, pp. 209–258 (1983)
56. Korányi, A.: Geometric properties of Heisenberg-type groups. *Adv. Math.* **56**, 28–38 (1985)
57. Krylov, N.V.: Hölder continuity and  $L_p$  estimates for elliptic equations under general Hörmander's condition. *Topol. Methods Nonlinear Anal.* **9**(2), 249–258 (1997)
58. Kusuoka, S., Stroock, D.: Long time estimates for the heat kernel associated with a uniformly subelliptic symmetric second order operator. *Ann. Math.* **127**(2), 165–189 (1988)
59. Lanconelli, E., Morbidelli, D.: On the poincare inequality for vector fields. *Ark. Mat.* **38**, 327–342 (2000)
60. Lieberman, G.M.: *Second order Parabolic Differential Equations*. World Scientific Publishing Co., Inc., River Edge, NJ (1996)
61. Lu, G., Manfredi, J.J., Stroffolini, B.: Convex functions on the Heisenberg group. *Calc. Var. Partial Differ. Equ.* **19**(1), 1–22 (2004)
62. Lunardi, A.: *Analytic Semigroups and Optimal Regularity in Parabolic Problems*. Modern Birkhäuser Classics. Birkhäuser/Springer Basel AG, Basel, (1995). [2013 reprint of the 1995 original] [MR1329547]

63. Magnani, V.: Lipschitz continuity, Aleksandrov theorem, and characterizations for H-convex functions. *Math. Ann.* **334**, 199–233 (2006)
64. Manfredini, M.: Uniform Schauder estimates for regularized hypoelliptic equations. *Ann. Mat. Pura Appl.* (4) **188** **3**, 417–428 (2009)
65. Mitchell, J.: On Carnot–Carathéodory metrics. *J. Differ. Geom.* **21**(1), 35–45 (1985)
66. Montgomery, R.: Survey of singular geodesics. Vol. 144 of *Bellaïche and Risler* [3], pp. 325–339 (1996)
67. Montgomery, R.: *A Tour of sub-Riemannian Geometries, Their Geodesics and Applications*. No. 91 in *Mathematical Surveys and Monographs*. American Mathematical Society, New York (2002)
68. Monti, R.: Distances, boundaries and surface measures in Carnot–Caratheodory spaces. Ph.D Thesis, Università degli studi di Trento (2001)
69. Monti, R., Rickly, M.: Geodetically convex sets in the Heisenberg group. *J. Convex Anal.* **12**(1), 187–196 (2005)
70. Nagel, A., Stein, E.M., Wainger, S.: Balls and metrics defined by vector fields. I. Basic properties. *Acta Math.* **155**(1–2), 103–147 (1985)
71. Pauls, S.D.: Minimal surfaces in the Heisenberg group. *Geom. Dedic.* **104**, 201–231 (2004)
72. Pauls, S.D.:  $H$ -minimal graphs of low regularity in  $\mathbb{H}^1$ . *Comment. Math. Helv.* **81**(2), 337–381 (2006)
73. Ritoré, M.: A proof by calibration of an isoperimetric inequality in the Heisenberg group  $\mathbb{H}^n$ . *Calc. Var. Partial Differ. Equ.* **44**(1–2), 47–60 (2012)
74. Ritoré, M., Rosales, C.: Rotationally invariant hypersurfaces with constant mean curvature in the Heisenberg group  $\mathbb{H}^n$ . *J. Geom. Anal.* **16**(4), 703–720 (2006)
75. Ritoré, M., Rosales, C.: Area-stationary surfaces in the Heisenberg group  $\mathbb{H}^1$ . *Adv. Math.* **219**(2), 633–671 (2008)
76. Rothschild, L.P., Stein, E.M.: Hypoelliptic differential operators and nilpotent groups. *Acta Math.* **137**(3–4), 247–320 (1976)
77. Rumin, M.: Sub-Riemannian limit of the differential form spectrum of contact manifolds. *Geom. Funct. Anal.* **10**(2), 407–452 (2000)
78. Saloff-Coste, L.: A note on Poincaré, Sobolev, and Harnack inequalities. *Internat. Math. Res. Notices* **1992**(2), 27–38 (1992)
79. Serra Cassano, F.: On the local boundedness of certain solutions for a class of degenerate elliptic equations. *Boll. Un. Mat. Ital. B* (7) **10** **3**, 651–680 (1996)
80. Sturm, K.T.: Analysis on local Dirichlet spaces. III. The parabolic Harnack inequality. *J. Math. Pure. Appl.* (9) **75** **3**, 273–297 (1996)
81. Wang, C.-Y.: Aronsson’s equations on Carnot–Carathéodory spaces. *Illinois J. Math.* **52**(3), 757–772 (2008)
82. Wang, C.Y.: Viscosity convex functions on Carnot groups. *Proc. Am. Math. Soc.* **133** **4**, 1247–1253 (2005). (electronic)
83. Xu, C.J.: Regularity for quasilinear second-order subelliptic equations. *Commun. Pure. Appl. Math.* **45**(1), 77–96 (1992)