Uniform Gaussian Bounds for Subelliptic Heat Kernels and an Application to the Total Variation Flow of Graphs Over Carnot Groups

Luca Capogna
Worcester Polytechnic Institute, lcapogna@smith.edu

Giovanna Citti
Alma Mater Studiorum Università di Bologna

Maria Manfredini
Alma Mater Studiorum Università di Bologna

Uniform Gaussian Bounds for Subelliptic Heat Kernels and an Application to the Total Variation Flow of Graphs over Carnot Groups

Abstract

In this paper we study heat kernels associated with a Carnot group G, endowed with a family of collapsing left-invariant Riemannian metrics σ_ϵ which converge in the Gromov-Hausdorff sense to a sub-Riemannian structure on G as $\epsilon \to 0$. The main new contribution are Gaussian-type bounds on the heat kernel for the σ_ϵ metrics which are stable as $\epsilon \to 0$ and extend the previous time-independent estimates in [16]. As an application we study well posedness of the total variation flow of graph surfaces over a bounded domain in a step two Carnot group (G, σ_ϵ). We establish interior and boundary gradient estimates, and develop a Schauder theory which are stable as $\epsilon \to 0$. As a consequence we obtain long time existence of smooth solutions of the sub-Riemannian flow $(\epsilon = 0)$, which in turn yield sub-Riemannian minimal surfaces as $t \to \infty$.

Keywords

Mean curvature flow • sub-Riemannian geometry • Carnot groups

MSC: 53C44, 53C17, 35R03

Luca Capogna[1,2*], Giovanna Citti[2†], Maria Manfredini[2‡]

1   Department of Mathematical Sciences,
     Worcester Polytechnic Institute, Worcester, MA 01609, USA

3   Dipartimento di Matematica, Piazza Porta S. Donato 5,
     40126 Bologna, Italy

1. Introduction

Models of image processing based on total variation flow

$$\partial_t u = \mathrm{div}(\nabla u / |\nabla u|)$$

have been first introduced in by Rudin, Osher, and Fatemi in [44], in order to perform edge preserving denoising. We refer the reader to the review paper [19] where many recent applications of total variation equation in the Euclidean setting are presented. The flow $t \to u(\cdot, t)$ represents the gradient descent associated with the total variation energy $\int |\nabla u|$ and as such has the property that both the total variation of the solution $t \to u(\cdot, t)$ and the perimeter measure of fixed level sets $\{u(\cdot, t) = const\}$ are non–increasing in time. Aside from its usefulness in image processing, the flow also arises in connection with the limit of solutions of the parabolic $p-$Laplacian $\partial_t u_p = \mathrm{div}(|\nabla u_p|^{p-2} \nabla u_p)$ as the parameter $p \to 1^+$. In the case where the evolution of graphs $S_t = \{(x, u(x, t))\}$ is considered, i.e. $\partial_t u = \mathrm{div}(\nabla u/\sqrt{1 + |\nabla u|^2})$, then in both the total variation flow and the closely related *mean curvature flow* $\partial_t u = \sqrt{1 + |\nabla u|^2}\,\mathrm{div}(\nabla u/\sqrt{1 + |\nabla u|^2})$, given appropriate boundary/initial conditions, global in time solutions asymptotically converge to minimal graphs. For further

*   *E–mail: lcapogna@wpi.edu (Corresponding Author)*

†   *E–mail: giovanna.citti@unibo.it*

‡   *E–mail: maria.manfredini@unibo.it*

results, references and applications of the Euclidean total variation flow we refer the reader to [1], [2], [5] and for an anisotropic version of the flow to [37].

In this paper we study long time existence of graph solutions of the total variation flow in a special class of degenerate Riemannian ambient spaces: The so-called (sub-Riemannian) Carnot groups [26], [47]. Such spaces are nilpotent Lie groups endowed with a metric structure (G, σ_0) that arises as limit of collapsing left-invariant Riemannian structures (G, σ_ϵ). The results in this paper are restricted to the study of the total variation flow in Carnot groups of step two. The restriction on the step is motivated by the lack of a sufficiently large set of barriers functions in higher step groups (see below). Our motivations are twofold:

(a) On the one hand, minimal surfaces and mean curvature flow in the Carnot group setting (and in particular in the special case of the Heisenberg group) have been the subject of numerous papers in recent years, leading to partial solutions of long-standing problems such as the Pansu conjecture[1]. Despite such advances little is known about explicit construction of minimal surfaces. Since the asymptotic limit of graphical mean curvature flow provide minimal surfaces in the Riemannian setting it is only natural to follow the same approach in the sub-Riemannian setting. However this avenue runs into considerable (so far unsolved) technical difficulties due to the combined effect of the degeneracy of the metric and the non-divergence form aspect of the relevant PDE. In the present work we provide the basis for the construction of sub-Riemannian minimal graphs through the asymptotic behavior as $t \to \infty$ of solutions of the sub-Riemannian total variation flow. For previous work on a similar theme see the work of Pauls [41] and Cheng, Hwang and Yang [14].

(b) A new class of image processing models based on the functionality of the visual cortex have been recently introduced in Lie groups. With a generalization of the classical Bargmann transform studied by Folland (see [27]) which lifts a L^2 function to a new one, defined on the phase space, an image can be lifted to a Lie group with a sub-Riemannian metric. The choice of the Lie group depends on the geometric property of the image to be studied: in [18], and [23], [24] 2D images are lifted to surfaces in the Lie group of Euclidean motions of the plane to study geometric properties of their level lines, in [25] cardiac images are lifted in the Heisenberg group to study their deformations, in [4] moving images are lifted in the Galilei group. The image processing is performed in these groups with algorithms expressed in terms of second order subelliptic differential equations. In particular sub-Riemannian mean curvature flows or total variation flows can be applied to perform image completion or impainting in this setting.

Our approach to the existence of global (in time) smooth solutions is based on a Riemannian approximation scheme: We study graph solutions of the total variation flow in the Riemannian spaces (G, σ_ϵ) where G is a Carnot group and σ_ϵ is a family of tame Riemannian metrics that 'collapse' as $\epsilon \to 0$ to a sub-Riemannian metric σ_0 in G. The *main technical novelties in the paper* are a series of different a-priori estimates, which are stable as $\epsilon \to 0$:

(1) Heat kernel estimates for sub-Laplacians and their elliptic regularizations (see Proposition 2.2). These estimates are established in the setting of any Carnot group, with no restriction on the step, and provide a *parabolic* counterpart to the time-independent estimates proved by two of the authors in [16].

(2) Uniform Schauder estimates for second order, non-divergence form subelliptic PDE and their *elliptic* regularizations (see Proposition 4.4). These estimates are established in the setting of any Carnot group, with no restriction on the step.

(3) Interior gradient estimates for solutions of the total variation flow (see Proposition 3.2). These estimates are established in the setting of any Carnot group, with no restriction on the step.

(4) Boundary gradient estimates for solutions of the total variation flow (see Proposition 3.4). These estimates are established only in the setting of Carnot group of step two.

We remark that (3) and (4) can also be proved, with similar arguments, for solutions of the sub-elliptic mean curvature flow[2]. We reiterate that the limitation to Carnot groups of step two in (4) is due to the fact that for step three and

---

[1] *See for instance [9, 10, 13–15, 21, 38, 41, 42, 46] and references therein*
[2] *For more references on the sub-Riemannian mean curvature flow see [8, 12, 22, 36]*

higher there no known suitable barrier functions, or in other words, the class of known explicit minimal graphs is not sufficiently large.

In order to state our results we need to introduce some notation.

1.1. Carnot group structure

Let G be an analytic and simply connected Lie group with topological dimension n and such that its Lie algebra $\mathcal{G}$ admits a stratification $\mathcal{G} = V^1 \oplus V^2 \oplus \ldots \oplus V^r$, where $[V^1, V^j] = V^{j+1}$, if $j = 1, \ldots, r-1$, and $[V^k, V^r] = 0$, $k = 1, \ldots, r$. Such groups are called in [26, 28, 47] stratified nilpotent Lie groups. Set $H(0) = V^1$, and for any $x \in G$ we let $H(x) = xH(0) = \text{span}(X_1, \ldots, X_m)(x)$. The distribution $x \to H(x)$ is called *the horizontal sub-bundle H*.

The metric structure is given by assuming that one has a left invariant positive definite form σ_0 defined in the sub-bundle H. We fix a orthonormal horizontal frame $X_1, \ldots, X_m$ which we complete to a basis $(X_1, \ldots, X_n)$ of $\mathcal{G}$ by choosing for every $k = 2, \ldots, r$ a basis of V_k. If X_i belongs to V_k, then we will define its *homogeneous degree* as

$$d(i) = k. \tag{1.1}$$

We will denote by $xX = \sum_{i=1}^{n} x_i X_i$ a generic element of $\mathcal{G}$. Since the exponential map $\exp : \mathcal{G} \to G$ is a global diffeomorphism we use exponential coordinates in G, and denote $x = (x_1, \ldots, x_n)$ the point $\exp(xX)$. We also set $x_H = (x_1, \ldots, x_m)$. Define non-isotropic dilations as $\delta_s(x) = (s^{d(i)} x_i)$, for $s > 0$.

We let $\nabla_0 = (X_1, \ldots, X_m)$ denote the *horizontal gradient* operator. If $\phi \in C^\infty(G)$ we set $\nabla_0 \phi = \sum_{i=1}^{m} X_i \phi X_i$ and $|\nabla_0 \phi|^2 = \sum_{i=1}^{m} (X_i \phi)^2$.

We denote by $(X_1, \ldots, X_n)$ (resp. $(X_1^r, \ldots, X_n^r)$) the left invariant (resp. right invariant) translation of the frames $(X_1, \ldots, X_n)$.

The vectors $X_1, \ldots, X_m$ and their commutators span all the Lie algebra $\mathcal{G}$, and consequently verify Hörmander's finite rank condition ([32]). This allows to use the results from [40], and define a control distance $d_{0,cc}(x, y)$ associated with the distribution $X_1, \ldots, X_m$, which is called the Carnot-Carathéodory metric (denote by $d_{r,0}$ the corresponding right invariant distance). We call the couple $(G, d_{0,cc})$ a *Carnot Group*.

We define a family of left invariant Riemannian metrics σ_ϵ, $\epsilon > 0$ in $\mathcal{G}$ by requesting that

$$\{X_1^\epsilon, \ldots, X_n^\epsilon\} := \{X_1, \ldots, X_m, \epsilon X_{m+1}, \ldots, \epsilon X_n\}$$

is an orthonormal frame. We will denote by d_ϵ the corresponding distance functions, see definition in (2.2). Correspondingly we use ∇_ϵ, (resp. ∇_ϵ^r) to denote the left (resp. right) invariant gradients. In particular, if $\phi \in C^\infty(G)$ we set $\nabla_\epsilon \phi = \sum_{i=1}^{m} X_i^\epsilon \phi X_i^\epsilon$ and $|\nabla_\epsilon \phi|^2 = \sum_{i=1}^{n} (X_i^\epsilon \phi)^2$.

We conclude by recalling the expression of the left invariant vector fields in exponential coordinates (see [43])

$$X_i = \partial_i + \sum_{k=d(i)+1}^{r} \sum_{d(j)=k} p_{ik}^j(x) \partial_j, \tag{1.2}$$

where $p_{ik}^j(x)$ is an homogeneous polynomial of degree $k - d(i)$ and depends only on x_h, with $d(1) \leq d(h) \leq k - d(i)$.

1.2. The total variation

The total variation flow is characterized by the fact that each point of the evolving surface graph moves in the direction of the *upward* unit normal with speed equal to the mean curvature times the volume element. In the setting of the approximating Riemannian metrics (G, σ_ϵ) and in terms of the functions $t \to u(\cdot, t) : \Omega \subset G \to \mathbb{R}$ describing the evolving graphs, the relevant equation reads:

$$\frac{\partial u_\epsilon}{\partial t} = h_\epsilon = \sum_{i=1}^{n} X_i^\epsilon \left(\frac{X_i^\epsilon u_\epsilon}{W_\epsilon} \right) = \sum_{i,j=1}^{n} a_{ij}(\nabla_\epsilon u_\epsilon) X_i^\epsilon X_j^\epsilon u_\epsilon \tag{1.3}$$

for $x \in \Omega$ and $t > 0$, with $u_\epsilon(x, 0) = \varphi(x)$, h_ϵ is the mean curvature of the graph of $u_\epsilon(\cdot, t)$ and

$$W_\epsilon^2 = 1 + |\nabla_\epsilon u_\epsilon|^2 = 1 + \sum_{i=1}^{n}(X_i^\epsilon u_\epsilon)^2 \text{ and } a_{ij}(\xi) = \frac{1}{\sqrt{1 + |\xi|^2}}\left(\delta_{ij} - \frac{\xi_i \xi_j}{1 + |\xi|^2}\right), \tag{1.4}$$

for all $\xi \in \mathbb{R}^n$, where $|\cdot|$ denotes the Euclidean norm in $\mathbb{R}^n$.
In the sub-Riemannian limit $\epsilon = 0$ the equation reads

$$\frac{\partial u}{\partial t} = \sum_{i=1}^{m} X_i\left(\frac{X_i u}{\sqrt{1 + |\nabla_0 u|^2}}\right), \tag{1.5}$$

for $x \in \Omega$ and $t > 0$, with $u(x, 0) = \varphi(x)$.
We will be concerned with uniform (in the parameter ϵ as $\epsilon \to 0$) estimates and with the asymptotic behavior of solutions to the initial value problem for the mean curvature motion of graphs over bounded domains of a Carnot group G,

$$\begin{cases} \partial_t u_\epsilon = h_\epsilon & \text{in } Q = \Omega \times (0, T) \\ u_\epsilon = \varphi & \text{on } \partial_p Q. \end{cases} \tag{1.6}$$

Here $\partial_p Q = (\Omega \times \{t = 0\}) \cup (\partial\Omega \times (0, T))$ denotes the parabolic boundary of Q.
Given appropriate hypothesis on the data, for instance convexity of Ω and φ twice continuously differentiable, the classical parabolic theory yields local existence and uniqueness for smooth solutions u_ϵ of (1.6), see [34, Chapter 5]. However classical parabolic theory will only provide estimates involving constants that degenerate as $\epsilon \to 0$ (in the transition from parabolic to degenerate parabolic regime). Our main goal consists in proving stable estimates.
Our first result consists in showing that if the initial/boundary data is sufficiently smooth then the solutions of (1.6) are Euclidean Lipschitz function up to the boundary uniformly in $\epsilon > 0$.

Theorem 1.1.

(Global gradient bounds) Let G be a Carnot group of step two, $\Omega \subset G$ a bounded, open, convex[3] set and $\varphi \in C^2(\bar\Omega)$. For $1 \geq \epsilon > 0$ 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 (1.6). There exists $C = C(G, \|\varphi\|_{C^2(\bar\Omega)}) > 0$ such that

$$\sup_{\bar\Omega \times (0,T)} |\nabla_\epsilon u_\epsilon| \leq \sup_{\bar\Omega \times (0,T)} |\nabla_1 u_\epsilon| \leq C, \tag{1.7}$$

where ∇_1 is the full σ_1–Riemannian gradient.

Remark 1.2.

The hypothesis of the Theorem above can be slightly weakened by asking that G is a step two Carnot group, $\varphi \in C^2(\bar\Omega)$, $\Omega \subset G$ bounded open set, but that $\partial\Omega$ be required to be convex in the Euclidean sense only in a neighborhood of its characteristic locus $\Sigma(\partial\Omega)$ and that its mean curvature $h_\epsilon^{\partial\Omega} \leq -\delta < 0$ in $\partial\Omega \setminus \Sigma(\partial\Omega)$.

Having established Lipschitz bounds, it is easy to show that the right derivatives $X_i^r u_\epsilon$ of the solutions of (1.6) are themselves solutions of (3.1), a divergence form, degenerate parabolic PDE whose weak solutions satisfy a Harnack inequality (see [11] and Proposition 3.8 below). Consequently one obtains $C^{1,\alpha}$ interior estimates for the solution u_ϵ

---

[3] *We say that a set $\Omega \subset G$ is convex in the Euclidean sense if $\exp^{-1}(\Omega) \subset \mathcal{G}$ is convex in the Euclidean space $\mathcal{G}$. In a group of step two this condition is translation invariant.*

which are uniform in $\epsilon > 0$. At this point one rewrites the PDE in (1.6) in non–divergence form[4] of the total variation flow equation $\partial_t u_\epsilon = a_{ij}(\nabla_\epsilon u_\epsilon) X_i^\epsilon X_j^\epsilon u_\epsilon$ and invokes the Schauder estimates in Proposition 4.4 to prove local higher regularity and long time existence. Since all the estimates are stable as $\epsilon \to 0$ one immediately obtains smoothness and the consequent global in time existence of the solution for the sub–Riemannian case $\epsilon = 0$.

Theorem 1.3.

In the hypothesis of Theorem 1.1 (or Remark 1.2) one has that there exists a unique solution $u_\epsilon \in C^\infty(\Omega \times (0, \infty)) \cap L^\infty((0, \infty), C^1(\bar{\Omega}))$ of the initial value problem

$$\begin{cases} \partial_t u_\epsilon = h_\epsilon & in\ Q = \Omega \times (0, \infty) \\ u_\epsilon = \varphi & on\ \partial_p Q \end{cases} \tag{1.8}$$

and that for each $k \in \mathbb{N}$ there exists $C_k = C_k(G, \varphi, k, \Omega) > 0$ not depending on ϵ such that

$$\|u_\epsilon\|_{C^k(Q)} \leq C_k. \tag{1.9}$$

Since the estimates are uniform in ϵ and in time, and with respect to ϵ, we will deduce the following corollary:

Corollary 1.4.

Under the assumptions of the Theorem 1.1, as $\epsilon \to 0$ the solutions u_ϵ converge uniformly (with all its derivatives) on compact subsets of Q to the unique, smooth solution $u_0 \in C^\infty(\Omega \times (0, \infty)) \cap L^\infty((0, \infty), C^1(\bar{\Omega}))$ of the sub–Riemannian total variation flow (1.5) in $\Omega \times (0, \infty)$ with initial data φ.

Corollary 1.5.

Under the assumptions of Theorem 1.1, as $t \to \infty$ the solutions $u_\epsilon(\cdot, t)$ converge uniformly on compact subsets of Ω to the unique solution $\tilde{u}_\epsilon$ of the minimal surface equation

$$h_\epsilon = 0 \quad in\ \Omega$$

with boundary value φ, while $\tilde{u}_0 = \lim_{\epsilon \to 0} \tilde{u}_\epsilon \in C^\infty(\Omega) \cap Lip(\bar{\Omega})$ is the unique solution of the sub–Riemannian minimal surfaces equation $h_0 = 0$ in Ω, with boundary data φ.

2. Structure stability in the Riemannian limit

The Carnot–Carathéodory metric $d_{0,cc}$ is equivalent to a more explicitly defined pseudo-distance function, that we will call (improperly) gauge distance, defined as

$$|x|_G^{2r!} = \sum_{k=1}^{r} \sum_{i=1}^{m_k} |x_i|^{\frac{2r!}{d(i)}}, \ and\ d_0(x, y) = |y^{-1}x|_G. \tag{2.1}$$

---

[4] *This is possible since for $\epsilon > 0$ the solutions u_ϵ are sufficiently smooth*

The *ball-box* theorem in [40] states that there exists $A = A(G, \sigma_0)$ such that for each $x \in G$,

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

If $x \in G$ and $r > 0$, we will denote by

$$B(x, r) = \{y \in G \mid d_0(x, y) < r\}$$

the balls in the gauge distance. For each $\epsilon > 0$ we also define the distance function d_ϵ corresponding to the Riemannian metric σ_ϵ,

$$d_\epsilon(x, y) = \inf\{\int_0^1 |\gamma'|_{\sigma_\epsilon}(s)ds \text{ with } \gamma : [0, 1] \to G \quad \text{a Lipschitz curve s. t. } \gamma(0) = x, \gamma(1) = y\}, \tag{2.2}$$

where $|\cdot|_{\sigma_\epsilon}$ denotes the norm with respect the Riemannian metric σ_ϵ, as well as the pseudo–distance $d_{G,\epsilon}(x, y) = N_\epsilon(y^{-1}x)$ with

$$N_\epsilon^2(x) = \sum_{d(i)=1} x_i^2 + \min\left\{\sum_{i=2}^r (\sum_{d(k)=i} x_k^2)^{\frac{1}{i}}, \epsilon^{-2} \sum_{d(i)\geq 2} x_i^2\right\}. \tag{2.3}$$

Note that in the definition of d_ϵ, if the curve for which the infimum is achieved happens to be horizontal then $d_\epsilon(x, y) = d_{0,cc}(x, y)$. In general we have $\sup_{\epsilon>0} d_\epsilon(x, y) = d_{0,cc}(x, y)$ and it is well known[5] that (G, d_ϵ) converges in the Gromov–Hausdorff sense as $\epsilon \to 0$ to the sub-Riemannian space $(G, d_{0,cc})$. The *ball-box* theorem in [40] and elementary considerations yield that there exists $A = A(G, \sigma_0) > 0$ independent of ϵ such that for all $x, y \in G$

$$A^{-1}d_{G,\epsilon}(x, y) \leq d_\epsilon(x, y) \leq Ad_{G,\epsilon}(x, y) \tag{2.4}$$

(see for example [11]).

2.1. Stability of the homogenous structure as $\epsilon \to 0$

If G is a Carnot group, d_ϵ is the distance function associated with σ_ϵ, we will denote

$$B_\epsilon(x, r) = \{y \in G | d_\epsilon(x, y) < r\}.$$

If we denote by dx the Lebesgue measure and by $|\Omega|$ the corresponding measure of a subset Ω, then Rea and two of the authors have recently proved in [11] that

Proposition 2.1.
There is a constant C independent of ϵ such that for every $x \in G$ and $r > 0$,

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

Having this property the spaces (G, d_ϵ, dx) are called *homogenous* with constant $C > 0$ independent of ϵ (see [20]). Let $\tau > 0$ and consider the space $\tilde{G} = G \times (0, \tau)$ with its product Lebesgue measure $dxdt$. In $\tilde{G}$ define the pseudo-distance function

$$\tilde{d}_\epsilon((x, t), (y, s)) = \max(d_\epsilon(x, y), \sqrt{|t - s|}). \tag{2.5}$$

Proposition 2.1 tells us that $(\tilde{G}, \tilde{d}_\epsilon, dxdt)$ is a homogeneous space with constant independent of $\epsilon \geq 0$.
In the paper [11] it is also shown that a Poincaré inequality holds with a choice of a constant which is stable as $\epsilon \to 0$. The stability of the homogenous space structure and of the constant in the Poincaré inequality are two of the key factors in the proof of Proposition 3.8.

---

[5] *See for instance [30] and references therein*

2.2. Stability in the estimates of the Heat Kernel

The results in this section are of independent interest and concern uniform Gaussian estimates for the heat kernel associated with elliptic regularizations of the Carnot group sub–Laplacians. They will be used in this paper in connection with the Schauder estimates and the higher regularity of the total variation flow. We will deal with non–divergence form operators similar to those in (1.3), but with constant coefficients. More precisely we will consider the operator

$$L_{\epsilon,A} = \partial_t - \sum_{i,j=1}^{n} a_{ij} X_i^\epsilon X_j^\epsilon \tag{2.6}$$

where $A = (a_{ij})_{ij=1,\ldots n}$ is a symmetric, real–valued $n \times n$ matrix, such that for some choice of constants $\Lambda, C_1, C_2 > 0$ and for all $\xi \in \mathbb{R}^n$ one has

$$\Lambda^{-1} \sum_{d(i)=1} \xi_i^2 + C_1 \sum_{d(i)>1} \xi_i^2 \le \sum_{i,j=1}^{n} a_{ij}\xi_i\xi_j \le \Lambda \sum_{d(i)=1} \xi_i^2 + C_2 \sum_{d(i)>1} \xi_i^2. \tag{2.7}$$

The constants C_1, C_2, Λ (independent of ϵ) provide $\Lambda-$*uniform coercivity of* (2.6) *in the horizontal directions*. Formally, in the sub–Riemannian limit $\epsilon \to 0$ the equation becomes

$$L_A = \partial_t - \sum_{i,j=1}^{m} a_{ij} X_i X_j. \tag{2.8}$$

In order to ensure that the operator $L_{\epsilon,A}$ (resp. L_A) is uniformly elliptic (reps. subelliptic), we will assume that the matrix $A = (a_{ij})$ belongs to a set of the form

$$M_\Lambda = \{A : A \text{ is a } \text{ symmetric } n \times n, \text{ real valued constant matrix, } \text{ satisfying (2.7) for some choice of } C_1 \text{ and } C_2\}$$

for some fixed $\Lambda > 0$.

Heat kernel estimates in Nilpotent groups are well known (see for instance [33] and references therein). We also refer to [6] where a self contained proof is provided. In our work we will need estimates which are uniform in the variable ϵ as $\epsilon \to 0$, in the same spirit as the results in [16]. We will denote $\Gamma_{\epsilon,A}(x, t)$ the fundamental solution of (2.6), with matrix (a_{ij}) in M_Λ, and $\Gamma_A(x, t)$ the fundamental solution of (2.8).

Proposition 2.2.

There exists constants $C_\Lambda > 0$ depending on G, σ_0, Λ but independent of ϵ such that for each $\epsilon > 0$, $x \in G$ and $t > 0$ one has

$$C_\Lambda^{-1} \frac{e^{-C_\Lambda \frac{d_\epsilon(x,0)^2}{t}}}{|B_\epsilon(0, \sqrt{t})|} \le \Gamma_{\epsilon,A}(x, t) \le C_\Lambda \frac{e^{-\frac{d_\epsilon(x,0)^2}{C_\Lambda t}}}{|B_\epsilon(0, \sqrt{t})|}. \tag{2.9}$$

For $s \in \mathbb{N}$ and $k-$tuple $(i_1,\ldots,i_k) \in \{1,\ldots,n\}^k$ there exists $C_{s,k} > 0$ depending only on $k, s, G, \sigma_0, \Lambda$ such that

$$|(\partial_t^s X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon \Gamma_{\epsilon,A})(x, t)| \le C_{s,k} t^{-s-k/2} \frac{e^{-\frac{d_\epsilon(x,0)^2}{C_\Lambda t}}}{|B_\epsilon(0, \sqrt{t})|} \tag{2.10}$$

for all $x \in G$ and $t > 0$. For any $A_1, A_2 \in M_\Lambda$, $s \in \mathbb{N}$ and $k-$tuple $(i_1,\ldots,i_k) \in \{1,\ldots,m\}^k$ there exists $C_{s,k} > 0$ depending only on $k, s, G, \sigma_0, \Lambda$ such that

$$|(\partial_t^s X_{i_1} \cdots X_{i_k} \Gamma_{\epsilon,A_1})(x, t) - (\partial_t^s X_{i_1} \cdots X_{i_k} \Gamma_{\epsilon,A_2})(x, t)| \le ||A_1 - A_2|| C_{s,k} t^{-s-k/2} \frac{e^{-\frac{d_\epsilon(x,0)^2}{C_\Lambda t}}}{|B_\epsilon(0, \sqrt{t})|}, \tag{2.11}$$

where $||A||^2 := \sum_{i,j=1}^{n} a_{ij}^2$.
Moreover, for $s \in \mathbb{N}$ and $k-$tuple $(i_1,\ldots,i_k) \in \{1,\ldots,m\}^k$ as $\epsilon \to 0$ one has

$$X_{i_1} \cdots X_{i_k} \partial_t^s \Gamma_{\epsilon,A} \to X_{i_1} \cdots X_{i_k} \partial_t^s \Gamma_A \tag{2.12}$$

uniformly on compact sets and in a dominated way on all G.

The proof of our result is directly inspired by the proof in [16, Theorem 1.1], where the time independent case was studied by two of us, and where a general procedure was introduced for handling the dependence on ϵ and obtaining independent estimates in the sub-Laplacian case.

Let us consider the group $\bar{G} = G \times G$ defined in terms of n new coordinates $y \in G$, so that points of $\bar{G}$ will be denoted $\bar{x} = (x, y)$. Denote by $Y_1, \ldots, Y_n$ a copy of the vectors $X_1, \ldots, X_n$, defined in terms of new variables y. The vector fields (X_i) and (Y_i) are defined in the product algebra $\bar{\mathcal{G}} = \mathcal{G} \times \mathcal{G}$.

In $\bar{G}$ we will consider two different families of sub-Riemannian structures:

- A sub-Riemannian structure determined by the choice of horizontal vector fields given by

$$(\bar{X}_1^0, \cdots, \bar{X}_{m+n}^0) = (X_1, \ldots, X_m, Y_1, \ldots, Y_n). \tag{2.13}$$

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

$$\bar{L}_{0,A} = \partial_t - \sum_{i,j=1}^{m+n} \bar{a}_{ij} \bar{X}_i^0 \bar{X}_j^0 = \partial_t - \sum_{i,j=1}^{m} a_{ij} X_i X_j - \sum_{i,j=1}^{n} a_{ij} Y_i Y_j. \tag{2.14}$$

  We complete the horizontal frame to a basis of the whole Lie algebra by adding the commutators of the vectors (X_i): $(X_{m+1}, \ldots X_n)$. We denote the exponential coordinates of a point $\bar{x}$ around a point $\bar{x}_0$ in terms of the full frame through the coefficients (v_i^0, w_i^0) which are defined by

$$\bar{x} = \exp\left(\sum_{i=1}^{m} v_i^0 X_i + \sum_{i=1}^{n} w_i^0 Y_i + \sum_{i=m+1}^{n} v_i^0 X_i \right)(\bar{x}_0).$$

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

$$(\bar{X}_1^\epsilon, \cdots, \bar{X}_{m+n}^\epsilon) = (X_1, \ldots, X_m, Y_1, \ldots, Y_m, X_{m+1}^\epsilon + Y_{m+1}, \ldots, X_n^\epsilon + Y_n). \tag{2.15}$$

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

$$\bar{L}_{\epsilon,A} = \partial_t - \sum_{i,j=1}^{m} a_{ij} X_i X_j - \sum_{i,j=1}^{m} a_{ij} Y_i Y_j - \sum_{i>m \text{ or } j>m} a_{ij}(X_i^\epsilon + Y_i)(X_j^\epsilon + Y_j).$$

  Analogously, the exponential coordinates associated with $\bar{L}_{\epsilon,A}$ will depend on the given horizontal frame (2.15) and the family $(X_{m+1}, \ldots X_n)$. The coordinates of a point $\bar{x}$ are the coefficients v_i^ϵ and w_i^ϵ satisfying

$$\bar{x} = \exp\left(\sum_{i=1}^{m} v_i^\epsilon X_i + \sum_{i=1}^{m} w_i^\epsilon Y_i + \sum_{i=m+1}^{n} w_i^\epsilon(X_i^\epsilon + Y_i) + \sum_{i=m+1}^{n} v_i^\epsilon X_i \right)(\bar{x}_0). \tag{2.16}$$

We denote by $\bar{\Gamma}_{0,A}$ and $\bar{\Gamma}_{\epsilon,A}$ the heat kernels of the corresponding heat operators. In both structures we define associated (pseudo)distance functions $\bar{d}_0(\bar{x}, \bar{y})$ and $\bar{d}_\epsilon(\bar{x}, \bar{y})$ that are equivalent to those defined in (2.1) and that assign unit weight to the corresponding horizontal vectors while $(X_{m+1}, \ldots, X_n)$ will be weighted according to their degree $d(i)$.

Definition 2.3.
For every $\epsilon > 0$ and $\bar{x}, \bar{x}_0 \in \bar{G}$ define

$$\bar{d}_\epsilon(\bar{x}, \bar{x}_0) = \sum_{i=1}^{m} |v_i^\epsilon| + \sum_{i=1}^{m} |w_i^\epsilon| + \sum_{i=m+1}^{n} \min\left(|w_i^\epsilon|, |w_i^\epsilon|^{1/d(i)} \right) + \sum_{i=m+1}^{n} |v_i^\epsilon|^{1/d(i)}.$$

For $\epsilon = 0$ and $\bar{x}, \bar{x}_0 \in \bar{G}$ define

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

We will denote by $\bar{B}_\epsilon$ and $\bar{B}_0$ the corresponding metric balls.

Lemma 2.4.

There exists constants $C_\Lambda > 0$ depending on G, σ_0, Λ but independent of ϵ such that for each $\epsilon > 0$, $\bar{x} \in \bar{G}$ and $t > 0$ one has

$$C_\Lambda^{-1} \frac{e^{-C_\Lambda \frac{\bar{d}_\epsilon(\bar{x},0)^2}{t}}}{|\bar{B}_\epsilon(0, \sqrt{t})|} \leq \bar{\Gamma}_{\epsilon,A}(\bar{x}, t) \leq C_\Lambda \frac{e^{-\frac{\bar{d}_\epsilon(\bar{x},0)^2}{C_\Lambda t}}}{|\bar{B}_\epsilon(0, \sqrt{t})|}. \tag{2.17}$$

For $s \in \mathbb{N}$ and $k-$tuple $(i_1, \ldots, i_k) \in \{1, \ldots, n+m\}^k$ there exists $C_{s,k} > 0$ depending only on $k, s, G, \sigma_0, \Lambda$ such that

$$|(\partial_t^s \bar{X}_{i_1}^\epsilon \cdots \bar{X}_{i_k}^\epsilon \bar{\Gamma}_{\epsilon,A})(\bar{x}, t)| \leq C_{s,k} \, t^{-s-k/2} \frac{e^{-\frac{\bar{d}_\epsilon(\bar{x},0)^2}{C_\Lambda t}}}{|\bar{B}_\epsilon(0, \sqrt{t})|} \tag{2.18}$$

for all $\bar{x} \in \bar{G}$ and $t > 0$. Moreover, for $s \in \mathbb{N}$ and $k-$tuple $(i_1, \ldots, i_k) \in \{1, \ldots, m\}^k$ as $\epsilon \to 0$ one has

$$\bar{X}_{i_1}^\epsilon \cdots \bar{X}_{i_k}^\epsilon \partial_t^s \bar{\Gamma}_{\epsilon,A} \to \bar{X}_{i_1}^0 \cdots \bar{X}_{i_k}^0 \partial_t^s \bar{\Gamma}_{0,A} \tag{2.19}$$

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

Proof.　In order to estimate the fundamental solution of the operators $\bar{L}_{\epsilon,A}$ in terms of $\bar{L}_{0,A}$, we define a volume preserving change of variables on the Lie algebra and on G: Define a Lie algebra automorphism $\bar{F}_\epsilon : \bar{G} \to \bar{G}$ as $\bar{F}_\epsilon(\bar{x}) = \exp(T_\epsilon(\log(\bar{x}))$ with

$$T_\epsilon(\bar{X}_i^0) = \bar{X}_i^\epsilon \text{ for } i = 1, \ldots, m+n.$$

By definition the relation between the distances $\bar{d}_0$ and $\bar{d}_\epsilon$ is expressed by the formula

$$\bar{d}_\epsilon(\bar{x}, \bar{x}_0) = \bar{d}_0(\bar{F}_\epsilon(\bar{x}), \bar{F}_\epsilon(\bar{x}_0)). \tag{2.20}$$

Analogously we also have

$$\bar{\Gamma}_{\epsilon,A}(\bar{x}, t) = \bar{\Gamma}_{0,A}(\bar{F}_\epsilon(\bar{x}), t), \tag{2.21}$$

$$X_i \bar{\Gamma}_{\epsilon,A}(\bar{x}, t) = X_i \bar{\Gamma}_{0,A}(\bar{F}_\epsilon(\bar{x}), t), \text{ for } i = 1, \cdots, m$$

and

$$(X_i^\epsilon + Y_i) \bar{\Gamma}_{\epsilon,A}(\bar{x}, t) = Y_i \bar{\Gamma}_{0,A}(\bar{F}_\epsilon(\bar{x}), t), \text{ for } i = m+1, \cdots, n,$$

with similar identities holding for the iterated derivatives.

In view of the latter, assertions (2.17) and (2.18) immediately follow from the well known estimates of $\bar{\Gamma}_{0,A}$ (see for instance the references cited above or [6, Theorem 2.5]). The pointwise convergence (2.19) is also an immediate consequence of (2.20) and (2.21). In order to prove the dominated convergence result we need to relate the distances $\bar{d}_0$ and $\bar{d}_\epsilon$: Expressing the exponential coordinates $v_i^\epsilon, w_i^\epsilon$ in terms of v_i^0, w_i^0 one easily obtains

$$\bar{d}_\epsilon(\bar{x}, \bar{x}_0) = \sum_{i=1}^m (|v_i^0| + |w_i^0|) + \sum_{i=m+1}^n \left(|v_i^0 - \epsilon w_i^0|^{1/d(i)} + \min(|w_i^0|, |w_i^0|^{1/d(i)}) \right)$$

so that for all[6] $\bar{x}, \bar{x}_0 \in \bar{G}$

$$\bar{d}_0(\bar{x}, \bar{x}_0) - C_0 \leq \bar{d}_\epsilon(\bar{x}, \bar{x}_0) \leq \bar{d}_0(\bar{x}, \bar{x}_0) + C_0 \tag{2.22}$$

where C_0 is independent of ϵ. The latter and (2.18) imply dominated convergence on $\bar{G}$.　□

---

[6] *This estimate indicates the well-known fact that at large scale the Riemannian approximating distances are equivalent to the sub-Riemannian distance*

Lemma 2.5.

For $s \in \mathbb{N}$ and $k-$tuple $(i_1, \ldots, i_k) \in \{1, \ldots, m\}^k$ there exists $C_{s,k} > 0$ depending only on $k, s, G, \sigma_0, \Lambda$ but independent of ϵ such that for each $A_1, A_2 \in M_\Lambda$, $\epsilon > 0$, $\bar{x} \in \bar{G}$ and $t > 0$ one has

$$|(\partial_t^s \bar{X}_{i_1}^\epsilon \cdots \bar{X}_{i_k}^\epsilon \bar{\Gamma}_{\epsilon, A_1})(\bar{x}, t) - (\partial_t^s \bar{X}_{i_1}^\epsilon \cdots \bar{X}_{i_k}^\epsilon \bar{\Gamma}_{\epsilon, A_2})(\bar{x}, t)| \leq ||A_1 - A_2|| C_{s,k} t^{-s-k/2} \frac{e^{-\frac{\bar{d}_\epsilon(\bar{x}, 0)^2}{C_\Lambda t}}}{|\bar{B}_\epsilon(0, \sqrt{t})|}. \tag{2.23}$$

Proof. In view of (2.20) and (2.21) it is sufficient to establish the result for $\bar{\Gamma}_{0, A_1}$ and $\bar{\Gamma}_{0, A_2}$, thus eliminating the dependence on ϵ.

Although the vector fields $\bar{X}_i^0$, $i = 1, ..., m + n$ are not free, we can invoke the Rothschild and Stein lifting theorem [43] and lift them to a new family of free vector fields in a free Carnot group. For the sake of simplicity we will continue to use the same notation $\bar{X}_i^0$, $i = 1, ..., m + n$ to denote such family of free vector fields. To recover the desired estimate (2.23) for the original vector fields one needs to argue through projections down to the original space, exactly as done in [40] and [16]. A standard argument (see for instance [6, Theorem A]) yields that for every $A \in M_\Lambda$, if we set $\bar{A}$ as in (2.14) then there exists a Lie group authomorphism F_A such that

$$(\sum_{j=1}^{m+n} (\bar{A}^{1/2})_{ij} \bar{X}_j^0)(u \circ F_A) = (\bar{X}_i u) \circ F_A,$$

and

$$\bar{\Gamma}_{0,A}(\bar{x}, t) = |\det \bar{A}^{1/2}| \bar{\Gamma}_{0,I}(F_A(\bar{x}), t)$$

where I denotes the identity matrix. The authomorphism F_A is defined by

$$F_A(\bar{x}) = \exp(T_A(\log(\bar{x}))), \quad \text{with } T_A(\bar{X}_i^0) = \sum_{j=1}^{m+n} (A^{1/2})_{ij} \bar{X}_j^0 \text{ for } i = 1, \ldots, m$$

and is extended to the whole Lie algebra as morphism (see the Appendix for more details). As in (2.16), we denote by (v_i, w_i) the canonical coordinates of $F_{A_1}(\bar{x})$ around $F_{A_2}(\bar{x})$, then the Mean Value Theorem yields

$$(\partial_t^s \bar{X}_{i_1}^0 \cdots \bar{X}_{i_k}^0 \bar{\Gamma}_I)(F_{A_1}(\bar{x}), t) - (\partial_t^s \bar{X}_{i_1}^0 \cdots \bar{X}_{i_k}^0 \bar{\Gamma}_I)(F_{A_2}(\bar{x}), t) = \sum_{i=1}^{n} (v_i X_i + w_i Y_i)(\partial_t^s \bar{X}_{i_1}^0 \cdots \bar{X}_{i_k}^0 \bar{\Gamma}_I)(\bar{y}, t).$$

By the result in the Appendix, the operators $v_i X_i + w_i Y_i$ are zero order differential operators whose coefficients can be estimated by $||A_1 - A_2||$. The conclusion follows by virtue of Proposition 4.5. $\square$

We conclude this section with the proof of the main result Proposition 2.2.

Proof of Proposition 2.2. From the definition of fundamental solution we have that

$$\Gamma_{0,A}(x, t) = \int_G \bar{\Gamma}_{0,A}((x, y), t) dy, \quad \text{and} \quad \Gamma_{\epsilon, A}(x, t) = \int_G \bar{\Gamma}_{\epsilon, A}((x, y), t) dy,$$

for any $x \in G$ and $t > 0$. In view of the (global) dominated convergence of the derivatives of $\bar{\Gamma}_{\epsilon, A}$ to the corresponding derivatives of $\bar{\Gamma}_{0,A}$ as $\epsilon \to 0$, we deduce that

$$\int_G \bar{\Gamma}_{\epsilon, A}((x, y), t) dy \to \int_G \bar{\Gamma}_{0,A}((x, y), t) dy$$

as $\epsilon \to 0$. The Gaussian estimates of $\Gamma_{\epsilon,A}$ follow from the corresponding estimates on $\bar{\Gamma}_{\epsilon,A}$ and the fact that in view of (2.22),

$$\bar{d}_\epsilon((x, y), (x_0, y_0)) \geq \bar{d}_0((x, y), (x_0, y_0)) - C_0 \geq d_0(x, x_0) + d_0(y, y_0) - C_0 \geq d_\epsilon(x, x_0) + d_\epsilon(y, y_0) - 3C_0.$$

Indeed the latter shows that there exists a constant $C > 0$ depending only on G, σ_0 such that for every $x \in G$,

$$\int_G e^{-\frac{d_\epsilon^2((x,y),(x_0,y_0))}{t}} \, dy \leq C e^{-\frac{d_\epsilon^2(x,x_0)}{t}} \int_G e^{-\frac{d_\epsilon^2(y,y_0)}{t}} \, dy \leq C e^{-\frac{d_\epsilon^2(x,x_0)}{t}}.$$

The conclusion follows at once. □

3. Gradient estimates

In this section we prove Theorem 1.1. The proof is carried out in two steps: First we use the maximum principle to establish interior L^∞ bounds for the full gradient of the solution $\nabla_1 u$ of (1.6) with respect to the Lipschitz norm of u on the parabolic boundary. Next, we construct appropriate barriers and invoke the comparison principle established in [8] to prove boundary gradient estimates. The combination of the two will yield the uniform global L^∞ gradient bounds.

3.1. Interior gradient estimates

In order to prove L^∞ bounds on the horizontal gradient of solutions of (1.6) one *cannot* differentiate equation (1.3) with respect to the left inariant frame, because these vector fields do not commute. On the other hand the right invariant vector fields X_j^r commute with the left invariant frame X_i, $i = 1, \ldots, n$. Hence it is easy to show through a direct computation the following result.

Lemma 3.1.
Let $u_\epsilon \in C^3(Q)$ be a solution to (1.3) and denote $v_0 = \partial_t u$, $v_i = X_i^r u$ for $i = i, \ldots, n$. Then for every $h = 0, \ldots, n$ one has that v_h is a solution of

$$\partial_t v_h = X_i^\epsilon(a_{ij} X_j^\epsilon v_h) = a_{ij}(\nabla_\epsilon u_\epsilon) X_i^\epsilon X_j^\epsilon v_h + \partial_{\xi_k} a_{ij}(\nabla_\epsilon u) X_i^\epsilon X_j^\epsilon u_\epsilon X_k^\epsilon v_h, \tag{3.1}$$

where a_{ij} is defined in (1.4).

From here we can deduce the following

Proposition 3.2.
Let $u_\epsilon \in C^3(Q)$ be a solution to (1.6) with Ω bounded. There exists $C = C(G, ||\varphi||_{C^2(\Omega)}) > 0$ such that for every compact subset $K \subset\subset \Omega$ one has

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

where ∇_1 is the full σ_1-Riemannian gradient.

Proof. The proof is similar to the argument in [8, Proposition 5.1]. By lemma 3.1 any right derivative $X_i^r u$ of u is solution of equation (3.1). By the weak maximum principle applied to (3.1) we obtain

$$\sup_{K \times [0,T)} |X_i^r u| \leq \sup_{\partial_p Q} |X_i^r u| \leq \sup_{\partial_p Q}(|\nabla_1 u| + |\partial_t u|), \tag{3.2}$$

since the right hand side contains a complete basis of the Riemannian tangent space. Again using the fact that the left invariant basis of vector fields can be expressed in terms of the right invarian one, we obtain:

$$\sup_{K \times [0,T)} |\nabla_1 u| \leq C \sup_{K \times [0,T)} \sum_{i=1}^n |X_i^r u| \leq C \sup_{\partial_p Q}(|\nabla_1 u| + |\partial_t u|).$$

□

3.2. Linear barrier functions

In [8, Section 4.2] it is shown that, in a step two Carnot group, coordinate hyperplanes (i.e. images under the exponential of level sets of the form $x_k = 0$) solve the minimal surface equation $h_0 = 0$. In the same paper it is also shown that this may fail for step three or higher. In the construction of the barrier function we will need the following slight refinement of this result,

Lemma 3.3.

Let G be a step two Carnot group. If $f : G \to \mathbb{R}$ is linear (in exponential coordinates) then for every $\epsilon \geq 0$, the matrix with entries $X_i^\epsilon X_j^\epsilon f$ is anti-symmetric, in particular every level set of f satisfies $h_\epsilon = 0$.

Proof.   We need to show that for $f(x) = \sum_{i=1}^n a_i x_i$,

$$\sum_{i,j=1}^n \left(\delta_{ij} - \frac{X_i^\epsilon f X_j^\epsilon f}{1 + |\nabla_\epsilon f|^2} \right) X_i^\epsilon X_j^\epsilon f = 0. \tag{3.3}$$

We recall the expression (1.2) for the vector fields X_i, $d(i) = 1$ in terms of exponential coordinates $X_i = \partial_{x_i} + \sum_{d(j)=1, d(h)=2} c_{ij}^h x_j \partial_{x_h}$. The Campbell-Hausdorff formula implies the anti-symmetry relation $c_{ij}^h = -c_{ji}^h$. It is immediate to observe that, if $d(k) = 2$ one has

$$X_i(x_k) = \sum_{d(j)=1} c_{ij}^k x_j, \quad \text{and} \quad X_i X_j(x_k) = c_{ji}^k, \quad \text{for } d(i) = d(j) = 1, \tag{3.4}$$

if either $d(i) = 2$ or $d(j) = 2$ it is easy to check $X_i^\epsilon X_j^\epsilon(x_k) = 0$ for all $k = 1, \dots, n$. Since $\left(\delta_{ij} - \frac{X_i^\epsilon f X_j^\epsilon f}{1 + |\nabla_\epsilon f|^2} \right)$ is symmetric in i, j it follows that

$$h_\epsilon = \sum_{k=1}^n a_k \sum_{i,j=1}^n \left(\delta_{ij} - \frac{X_i^\epsilon f X_j^\epsilon f}{1 + |\nabla_\epsilon f|^2} \right) X_i^\epsilon X_j^\epsilon(x_k) = 0. \tag{3.5}$$

□

3.3. Boundary gradient estimates

We say that a set $\Omega \subset G$ is convex in the Euclidean sense if $\exp^{-1}(\Omega) \subset \mathcal{G}$ is convex in the Euclidean space $\mathcal{G}$. In a group of step two this condition is translation invariant.

Proposition 3.4.

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 (1.6). There exists $C = C(G, \|\varphi\|_{C^2(\bar{\Omega})}) > 0$ such that

$$\sup_{\partial\Omega \times (0,T)} |\nabla_\epsilon u_\epsilon| \leq \sup_{\partial\Omega \times (0,T)} |\nabla_1 u_\epsilon| \leq C. \tag{3.6}$$

We start by recalling an immediate consequence of the proof of [8, Theorem 3.3].

Lemma 3.5.

For each $\epsilon \geq 0$, if U_ϵ is a bounded subsolution and V_ϵ is a bounded supersolution of (1.6) then $U_\epsilon(x, t) \leq V_\epsilon(x, t)$ for all $(x, t) \in Q$.

Let $u_\epsilon \in C^2(Q)$ be a solution of (1.6), and express the evolution PDE in non–divergence form

$$\partial_t u_\epsilon = h_\epsilon = \sum_{i,j=1}^{n} a_{ij}(\nabla_\epsilon u)X_i^\epsilon X_j^\epsilon u_\epsilon. \tag{3.7}$$

Set $v_\epsilon = u_\epsilon - \varphi$ so that v_ϵ solves the homogenous 'boundary' value problem

$$\begin{cases} \partial_t v_\epsilon = a_{ij}(\nabla_\epsilon v_\epsilon + \nabla_\epsilon \varphi)X_i^\epsilon X_j^\epsilon v_\epsilon + b^\epsilon & \text{in } Q = \Omega \times (0, T) \\ v_\epsilon = 0 & \text{on } \partial_p Q, \end{cases} \tag{3.8}$$

with $b^\epsilon(x) = a_{ij}e(\nabla_\epsilon v_\epsilon(x) + \nabla_\epsilon \varphi(x))X_i^\epsilon X_j^\epsilon \varphi(x)$. We define our (weakly) parabolic operator for which the function v_ϵ is a solution

$$Q(v) = a_{ij}(\nabla_\epsilon v_\epsilon + \nabla_\epsilon \varphi)X_i^\epsilon X_j^\epsilon v_\epsilon + b^\epsilon - \partial_t v. \tag{3.9}$$

In the following we construct for each point $p_0 = (x_0, t_0) \in \partial\Omega \times (0, T)$ a *barrier function* for Q, v_ϵ: i.e.,

Lemma 3.6.

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

$$Q(w) \le 0 \text{ in } V \cap Q \text{ with } V \text{ a parabolic neighborhood of } p_0, \tag{3.10}$$

$$w(p_0) = 0 \text{ and } w \ge v_\epsilon \text{ in } \partial_p V \cap Q.$$

Proof. In the hypothesis that Ω is convex in the Euclidean sense we have that every $x_0 \in \partial\Omega$ supports a tangent hyperplane P defined by an equation of the form $\Pi(x) = \sum_{i=1}^{n} a_i x_i = 0$ with $\Pi > 0$ in Ω, $\Pi(x_0) = 0$, and normalized as $\sum_{d(i)=1,2} a_i^2 = 1$. Following the standard argument (see for instance [35, Chapter 10]) we select the barrier at $(x_0, t_0) \in \partial\Omega \times (0, T)$ independent of time with

$$w = \Phi(\Pi) \tag{3.11}$$

with Φ solution of

$$\Phi'' + v(\Phi')^2 = 0, \tag{3.12}$$

in particular

$$\Phi(s) = \frac{1}{v} \log(1 + ks), \tag{3.13}$$

with k and v chosen appropriately so that conditions (3.10) will hold. We choose a neighborhood $V = O \times (0, T)$ such that $P \cap O \cap \partial\Omega = \{x_0\}$. By an appropriate choice of k sufficiently large we can easily obtain $w_\epsilon(p_0) = 0$ and $w_\epsilon \ge v_\epsilon$ in $\partial_p V \cap Q$.
To verify $Q(w_\epsilon) \le 0$ we begin by observing that w satisfies

$$Q(w) = \Phi' a_{ij}(\nabla_\epsilon w + \nabla_\epsilon \varphi)X_i^\epsilon X_j^\epsilon \Pi + \frac{\Phi''}{(\Phi')^2}\mathcal{F} + b_\epsilon, \tag{3.14}$$

with $\mathcal{F} = a_{ij}(\nabla_\epsilon w + \nabla_\epsilon \varphi)X_i^\epsilon w X_j^\epsilon w$.
We will show:

$$a_{ij}(\nabla_\epsilon w + \nabla_\epsilon \varphi)X_i^\epsilon X_j^\epsilon \Pi \le 0 \tag{Claim 1}$$

$$\frac{\Phi''}{(\Phi')^2}\mathcal{F} + b_\epsilon \le 0 \tag{Claim 2}$$

in a parabolic neighborhood of p_0.

The first claim holds with an equality as a_{ij} is symmetric and $X_i^\epsilon X_j^\epsilon \Pi$ is anti-symmetric in view of Lemma 3.3. To establish (Claim 2) we first note that Lemma 3.3 implies

$$\frac{\epsilon^2}{2} \leq \max(\sum_{d(i)=1} a_i^2, \epsilon^2 \sum_{d(k)=1} a_k^2) \leq |\nabla_\epsilon \Pi| =$$

$$= \sum_{d(i)=1} \left(a_i + \sum_{d(k)=2, d(j)=1} c_{ij}^k a_k x_j \right)^2 + \epsilon^2 \sum_{d(k)=2} a_k^2 \leq C(G)(1+\epsilon^2),$$

for some constant $C(G) > 0$. Consequently, for $\Phi' >> 1$ sufficiently large one finds

$$\mathcal{F} \geq \frac{|\nabla_\epsilon w|^2}{(1+|\nabla_\epsilon w + \nabla_\epsilon \varphi|^2)^{3/2}} \geq C(G) \frac{|\nabla_\epsilon w|^2}{(1+|\Phi'|^2+|\nabla_\epsilon \varphi|^2)^{3/2}} \geq C(G)\epsilon^2 > 0, \tag{3.15}$$

with $C(G) > 0$ a constant depending only on G (not always the same along the chain of inequalities). In view of the definition of b_ϵ and (3.12) with an appropriate choice of $v = v(G, \epsilon, \phi) > 0$ and $k = k(G, \phi) >> 1$ in (3.13), we conclude

$$\frac{\Phi''}{(\Phi')^2} \mathcal{F} + b_\epsilon \leq \left(\frac{\Phi''}{(\Phi')^2} + v \right) \mathcal{F} = 0. \tag{3.16}$$

$\square$

In view of Lemma 3.5, a comparison with the barrier constructed above yields that

$$0 \leq \frac{v_\epsilon(x, t)}{dist_{\sigma_1}(x, x_0)} \leq \frac{w(x, t)}{dist_{\sigma_1}(x, x_0)} \leq C(k, v), \tag{3.17}$$

in $V \cap Q$, with $dist_{\sigma_1}(x, x_0)$ being the distance between x and x_0 in the Riemannian metric σ_1, concluding the proof of the boundary gradient estimates.

The proof of Theorem 1.1 now follows immediately from Proposition 3.2 and Proposition 3.4.

Having established uniform global Lipschitz bounds one now notes that equation (3.1) satisfies *horizontal* coercivity conditions uniformly in $\epsilon > 0$. Such conditions are among the main hypothesis of the Harnack inequality in [11]. In this paper, G. Rea and two of the authors have proved that given a homogenous structure and a Poincaré inequality, then a sub-elliptic analogue of Aronson and Serrin's Harnack inequality [3] for quasilinear parabolic equations holds. As a consequence one obtains for some $\alpha \in (0, 1)$ that the solutions u_ϵ to (1.6) satisfy $C^{1,\alpha}$ Hölder estimates, uniform in $\epsilon \in (0, 1)$.

Definition 3.7.

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 $(x, t), (x_0, t_0) \in Q$

$$|u(x, t) - u(x_0, t_0)| \leq M \tilde{d}_\epsilon^\alpha((x, t), (x_0, t_0)). \tag{3.18}$$

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 $i = 1, \dots, m$ $X_i u \in C_{\epsilon,X}^{k-1,\alpha}(Q)$, where we have set $C_{\epsilon,X}^{0,\alpha}(Q) = C_{\epsilon,X}^\alpha(Q)$.

Corollary 3.8.

(Interior $C_X^{1,\alpha}$ estimates) In the hypothesis of the previous results, 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 $i = 1, \ldots, n$ one has that $v = X_i^\epsilon u$ satisfies

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

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

4. Regularity properties in the $C^{k,\alpha}$ spaces

In this section we will prove uniform estimates for solution of (1.3) in the $C_{\epsilon,X}^{k,\alpha}$ Hölder spaces. This is accomplished by using the uniform Gaussian bounds established in Section 2.2 to develop new uniform Schauder estimates for solutions of second order sub-elliptic differential equations in non divergence form

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

in a cylinder $Q = \Omega \times (0,T)$ that are stable as $\epsilon \to 0$ As usual we will make use of the associated linear, constant coefficient *frozen* operator:

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

where $(x_0, t_0) \in Q$.

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

Lemma 4.1.

Let w be a smooth solution to $L_\epsilon w = f$ in $Q \subset$. For every $\phi \in C_0^\infty(Q)$,

$$(w\phi)(x,t) = \int_Q \Gamma_{(x_0,t_0)}^\epsilon((x,t),(y,\tau)) \left(L_{\epsilon,(x_0,t_0)} - L_\epsilon \right)(w\phi)(y,\tau) dy d\tau + \tag{4.1}$$

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

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

We explicitly note that for $\epsilon > 0$ fixed the operator $L_{\epsilon,(x_0,t_0)}$ is uniformly parabolic. Its heat kernel can be studied through standard singular integrals theory in the corresponding Riemannian balls. Hence, as noted in [29, Chapter 4], one is allowed to differentiate twice the kernels defined in (4.1) with respect to any right or left invariant vector field.

Proposition 4.2.

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_δ the $\delta-$tubular neighborhood of K. Assume that there exists a constant $C > 0$ such that for every $\epsilon \in (0,1)$

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

There exists a constant $C_1 > 0$ depending on δ, α, C and the constants in Proposition 2.2 such that

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

Proof.   The proof follows the outline of the standard case, as in [29, Theorem 4, Chapter 3], and rests cru-cially on the Gaussian estimates proved in Proposition 2.2, and the fact that the functions $\left(L_{\epsilon,(x_0,t_0)} - L_\epsilon\right)(w\,\phi)$, and $\left(f\phi + wL_\epsilon\phi + 2a_{ij}^\epsilon X_i^\epsilon wX_j^\epsilon\phi\right)$ are Hölder continuous.

Choose a parabolic sphere[7] $B_{\epsilon,\delta} \subset\subset K$ 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 clearly implies that for some constant $C > 0$ depending only on G and σ_0, $|\nabla_\epsilon\phi| \leq C\delta^{-1}$, $|L^\epsilon\phi| \leq C\delta^{-2}$, in Q. Next, invoking (4.1) one has that for every multi-index $I = (i_1, i_2) \in \{1, \dots, m\}^2$ and for every $(x_0, t_0) \in B_{\epsilon,\delta}$

$$X_{i_1}^\epsilon X_{i_2}^\epsilon(w\phi)(x_0, t_0) = \int_Q X_{i_1}^\epsilon X_{i_2}^\epsilon \Gamma_{(x_0,t_0)}^\epsilon(\cdot, (y, \tau))|_{(x_0,t_0)} \left(L_{\epsilon,(x_0,t_0)} - L_\epsilon\right)(w\,\phi)(y, \tau)dyd\tau+ \tag{4.2}$$

$$+ \int_Q X_{i_1}^\epsilon X_{i_2}^\epsilon \Gamma_{(x_0,t_0)}^\epsilon(\cdot, (y, \tau)) \left(f\phi + wL_\epsilon\phi + 2\sum_{i,j=1}^n a_{ij}^\epsilon X_i^\epsilon wX_j^\epsilon\phi\right)(y, \tau)dyd\tau.$$

The uniform Hölder continuity of a_{ij}^ϵ, with Proposition 2.2 and Lemma 2.5 yield

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

with $C > 0$ independent of ϵ. In view of the latter, using basic singular integral properties (see [26, Theorem 6.1]) and proceeding as in [29, Theorem 2, Chapter 4], we obtain

$$\left|\left| \int X_{i_1}^\epsilon X_{i_2}^\epsilon \Gamma_{(x_0,t_0)}^\epsilon(\cdot, (y, \tau))(L_\epsilon - L_{\epsilon,(x_0,t_0)})(w\,\phi)(y, \tau)dyd\tau \right|\right|_{C_{\epsilon,X}^\alpha(B_{\epsilon,\delta})} \leq C_1 \left|\left|(L_\epsilon - L_{\epsilon,(x_0,t_0)})(w\,\phi)\right|\right|_{C_{\epsilon,X}^\alpha(B_{\epsilon,\delta})} \tag{4.3}$$

$$= C_1 \sum_{i,j} \left|\left|(a_{ij}^\epsilon(x_0, t_0) - a_{ij}^\epsilon(\cdot))X_j^\epsilon X_j^\epsilon(w\,\phi)\right|\right|_{C_{\epsilon,X}^\alpha(B_{\epsilon,\delta})}$$

$$\leq \tilde{C}_1\delta^\alpha||a_{ij}^\epsilon||_{C_{\epsilon,X}^\alpha(B_{\epsilon,\delta})}||w\phi||_{C_{\epsilon,X}^{2,\alpha}(B_{\epsilon,\delta})},$$

where C_1, and $\tilde{C}_1$ are stable as $\epsilon \to 0$. Similarly, if ϕ is fixed, the Hölder norm of the second term in the representation formula (4.2) is bounded by

$$\left|\left| \int X_{i_1}^\epsilon X_{i_2}^\epsilon \Gamma_{(x_0,t_0)}^\epsilon((x_0, t_0), (y, \tau))\left(f\phi(y, \tau) + wL\phi(y, \tau) + 2a_{ij}^\epsilon X_i^\epsilon wX_j^\epsilon\phi\right)dyd\tau \right|\right|_{C_{\epsilon,X}^\alpha(B_{\epsilon,\delta})}$$

$$\leq C_2\left(||f||_{C_{\epsilon,X}^\alpha(K_\delta)} + \frac{C}{\delta^2}||w||_{C_{\epsilon,X}^{1,\alpha}(K_\delta)}\right). \tag{4.4}$$

From (4.2), (4.3) and (4.4) 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 δ sufficiently small we prove the assertion on the fixed sphere $B_{\epsilon,\delta}$. The conclusion follows from a standard covering argument.

$\square$

---

[7] *That is a sphere in the group $\tilde{G} = G \times \mathbb{R}$ in the pseudo-metric $\tilde{d}_\epsilon$ defined in (2.5).*

The Hölder estimate of the second order derivative of a solution u obtained in the previous lemma, are a direct generalization of the analogous procedure in the Euclidean setting. In that classical case any derivative $v = \partial_i u$ is a solution of a second order equation, so that the Hölder estimates of the second order derivatives of v provide an estimate of the third order derivatives of the solution u, and so on. As we noted at the beginning of Section 3.1, in our setting the vector fields do not commute, and the gradient of the solution is not a solution of a new second order differential equation, expressed solely in terms of left invariant vector fields. Even the method of differentiating along right invariant derivatives, as we did in Section 3.1, does not work here. In fact in this instance we are not evaluating the full Riemannian gradient, but only the left derivatives belonging in the first layer, which can not be represented in terms of right derivatives lying in the first layer. These difficulties call for a different technique to establish higher order estimates for solutions. We observe that a new representation formula for derivatives of any order can be derived directly from (4.1). Specifically, differentiating the representation formula (4.1) yields

Lemma 4.3.

Let w be a smooth solution to $L_\epsilon w = f$ in $Q \subset$. For every $\phi \in C_0^\infty(Q)$, for any $k-$tuple $(i_1, \dots, i_k) \in \{1, \dots, m\}^k$,

$$
X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon (w\phi)(x,t) = \int_Q \Gamma_{(x_0,t_0)}^\epsilon((x,t),(y,\tau)) \left(a_{ij}^\epsilon - a_{ij}^\epsilon(x_0,t_0) \right) X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon X_i^\epsilon X_j^\epsilon (w\phi)(y,\tau) dy d\tau
$$

$$
+ \int_Q \Gamma_{(x_0,t_0)}^\epsilon((x,t),(y,\tau)) B(w\phi)(y,\tau) dy d\tau
$$

$$
+ \int_Q \Gamma_{(x_0,t_0)}^\epsilon((x,t),(y,\tau)) X_{i_1}^\epsilon \cdots X_{i_k}^\epsilon \left(f\phi(y,\tau) + w L_\epsilon \phi(y,\tau) + 2 \sum_{i,j=1}^n a_{ij}^\epsilon X_i^\epsilon w X_j^\epsilon \phi \right) dy d\tau,
$$

where B is a differential operator of order $k + 1$, depending on horizontal derivatives of a_{ij}^ϵ of order at most k, such that

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

Consequently, applying to this representation formula the same procedure as in Lemma 4.2 we obtain

Proposition 4.4.

Let w be a smooth solution of $L^\epsilon w = f$ 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_δ 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 α, C, δ, and the constants in Proposition 2.2, but independent of ϵ, such that

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

Proof of Theorem 1.3. We will 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

$$
\|u_\epsilon\|_{C_{\epsilon,X}^{k,\alpha}(K)} \leq C, \tag{4.5}
$$

for every $\epsilon > 0$. The assertion is true if $k = 2$, by Corollary 3.8 and Proposition 4.2. If (4.5) is true for $k + 1$, then the coefficients in (1.3) satisfy $a_{ij}^\epsilon \in C_{\epsilon,X}^{k,\alpha}$ uniformly as $\epsilon \in (0,1)$ and (4.5) thus holds at the level $k + 2$ by virtue of Proposition 4.4. $\qquad \square$

Appendix

Let us assume that G is a free nilpotent Lie group, and let $\mathcal{G} = V^1 \oplus \ldots \oplus V^r$ be its associated Lie algebra. Denote by $X_1, \ldots, X_m$ a basis of the first layer of the Lie algebra. and by $X_{m+1}, \ldots, X_n$ a list of vectors which complete a basis of the tangent space. If $u(x)$ is an homogeneous polynomial of order p and X is a differential operator of degree q we will call $u(x)X$ differential operator of degree $q - p$. If $a_{ij} = (A)_{ij}$ is a real valued, constant coefficient, $m \times m$ positive definite matrix , one can define a Lie algebra automorphism

$$T_A : V^1 \to V^1, \text{ with } T_A(X_i) = \sum_{j=1}^m a_{ij} X_j$$

for every $i = 1, \ldots, m$ and extend it to the whole Lie algebra as a morphism

$$T_A[X_i, X_j] = [T_A(X_i), T_A(X_j)] = \sum_{h,k=1}^m a_{ih} a_{jk} [X_h, X_k].$$

Note that T_A can be represented as a block matrix $\bar{A} = A_1 \oplus A_2 \cdots \oplus A_r$, with $A_1 = A$ and where the block A_j act on vectors of degree j, and its coefficients are polynomial of order j of the coefficients of A_1. In particular $T_A(X)$ and X have the same degree. Via the exponential map T_A induces a group automorphism on the whole group

$$F_A : G \to G, \text{ defined by } F_A(x) = \exp(T_A(\log(x)).$$

In terms of exponential coordinates (*around* the origin)

$$x = \exp\left(\sum_{i=1}^n v_i X_i\right)(0) \tag{A1}$$

one has

$$F_A(x) = \exp\left(\sum_{i=1}^n v_i T_A(X_i)\right)(0) = \exp\left(\sum_{s=1}^r \sum_{d(i)=d(j)=s} v_i (A_s)_{ij} X_j\right)(0),$$

where $d(i)$ denotes the homogenous degree defined as in (1.1). Since $v_i X_i$ is a zero order homogeneous operator then also $v_i T_A(X_i)$ is a differential operator of order zero.

Proposition 4.5.

Let $(w_i)_{i=1,\ldots,n}$ denote the canonical coordinates of $F_A(x)$ around $F_B(x)$, and $(v_j)_{j=1,\ldots,n}$ the coordinates of x around 0, as in (A1). There exists $M \in \mathbb{N}$ depending only on the group structure; constants $c_1, \ldots, c_M$ depending only on the group structure; zero-order differential operators $Y_1, \ldots, Y_M$ depending only on the group structure, on the coefficients $(v_j)_{j=1,\ldots,n}$ and their derivatives along vector fields up to order $r - 1$ (but independent on A and B), and a constant $C = C(||A||, ||B||, G) > 0$ such that

$$\sum_{i=1}^n w_i X_i = \sum_{l=1}^M c_l Y_l \quad \text{and } |c_l| \leq C||A - B|| \text{ for } l = 1, \ldots, M. \tag{A2}$$

Proof. By definition

$$\sum_{j=1}^n w_j X_j = \sum_{i=1}^n v_i T_A(X_i) * \left(-\sum_{i=1}^n v_i T_B(X_i)\right)$$

where $*$ is the Baker–Campbell–Hausdorff operation (see for instance [43] and references therein). This formula allows to represent the left-hand side as a finite[8] sum of terms involving commutators (up to order r) of the two terms on the right–hand side. Let us consider separately the terms in this representation starting with those involving no commutators, i.e.

$$\sum_{i=1}^{n} v_i T_A(X_i) - \sum_{i=1}^{n} v_i T_B(X_i) = \sum_{s=1}^{r} \sum_{d(i)=d(j)=s} v_i \Big((A_s)_{ij} - (B_s)_{ij} \Big) X_j.$$

Clearly this expression is in the same form as (A2). Next we consider the second term in the Baker–Campbell–Hausdorff sum, i.e. that involving commutators of the form

$$[\sum_{i=1}^{n} v_i T_A(X_i), \sum_{h=1}^{n} v_h T_B(X_h)] = \sum_{p,s=1}^{r} \sum_{d(i)=d(j)=d(h)=d(k)=s} (A_s)_{ij}(B_p)_{hk} \Big[v_i X_j, v_h X_k \Big] =$$

$$\sum_{p,s=1}^{r} \sum_{d(i)=d(j)=d(h)=d(k)=s} \frac{1}{2} \Big((A_s)_{ij}(B_p)_{hk} - (A_p)_{hk}(B_s)_{ij} \Big) \Big[v_i X_j, v_h X_k \Big].$$

It is evident that the coefficients of the commutators above are Lipschitz functions of A and B, vanishing for $A = B$, and consequently the whole sum satisfies (A2). To conclude the proof we note that all further terms of the Baker–Campbell–Hausdorff sum involve operators as in (A2) and zero order operators of the form $\sum_{i=1}^{n} v_i T_A(X_i)$ or $-\sum_{i=1}^{n} v_i T_B(X_i)$. Such commutators give rise to differential operators of zero order whose coefficients are polynomials in A and B and vanish when $A = B$. $\square$

Acknowledgments

References

[1] F. Andreu, C. Ballester, V. Caselles and J. M. Mazón, The Dirichlet problem for the total variation flow. J. Funct. Anal. 180 (2001), no. 2, 347-403.

[2] F. Andreu, V. Caselles and J. M. Mazón Ruiz, Parabolic quasilinear equations minimizing linear growth functionals, Progr. Math., 223, Birkhauser, Basel, 2004

[3] D. G. Aronson. and J. Serrin, Local behavior of solutions of quasilinear parabolic equations, Arch. Rational Mech. Anal., 25 (1967), 81-122.

[4] D. Barbieri. G. Citti, G. Cocci and A. Sarti, A cortical–inspired geometry for contour perception and motion integration, submitted paper.

[5] G. Bellettini, V. Caselles and M. Novaga, The total variation flow in $\mathbb{R}^N$, J. Differential Equations 184 (2002), no. 2, 475-525.

[6] A. Bonfiglioli, E. Lanconelli and F. Uguzzoni, Uniform Gaussian estimates of the fundamental solutions for heat operators on Carnot groups, Adv. Differential Equations, 7 (2002), 1153-1192.

[7] A. Bonfiglioli, E. Lanconelli and F. Uguzzoni, Fundamental solutions for non–divergence form operators on stratified groups, Trans. Amer. Math. Soc. 356 (2004), no. 7, 2709-2737.

[8] L. Capogna and G. Citti, Generalized mean curvature flow in Carnot groups, Comm. Partial Differential Equations 34 (2009), no. 7-9, 937-956.

[9] L. Capogna, G. Citti and M. Manfredini, Regularity of non–characteristic minimal graphs in the Heisenberg group H^1, Indiana Univ. Math. J. 58 (5) (2009) 2115-2160.

[10] L. Capogna, G. Citti and M. Manfredini, Smoothness of Lipschitz intrinsic minimal graphs in the Heisenberg group H^n, $n > 1$ J. Reine Angew. Math. (Crelle's Journal) (2010), 648, 75-110.

---

[8] *since the group is nilpotent*

[11] L. Capogna, G. Citti and G. Rea, A sub-elliptic analogue of Aronson-Serrin Harnack inequality, to appear in Mathematische Annalen, (2013).

[12] L. Capogna, G. Citti and C. Senni Guidotti Magnani,Sub-Riemannian Heat Kernels and Mean Curvature of graphs, Jour. of Functional Analysis, 264 (2013), no. 8 1899–1928.

[13] L. Capogna, D. Danielli, J. Tyson and S. Pauls, An Introduction to the Heisenberg Group and the Sub-Riemannian Isoperimetric Problem, Birkhauser, 2007.

[14] J.-H. Cheng, J.-F. Hwang and P. Yang, Regularity of C^1 smooth surfaces with prescribed p-mean curvature in the Heisenberg group, Math. Ann. 344 (2009), no. 1, 1-35.

[15] J.H. Cheng and J.F. Hwang, A. Malchiodi and P. Yang, Minimal surfaces in pseudohermitian geometry, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 4 (2005), 129-177.

[16] G. Citti and M. Manfredini, Uniform estimates of the fundamental solution for a family of hypoelliptic operators, Potential Anal. 25 (2006), no. 2, 147-164.

[17] G. Citti, A. Pascucci and S. Polidoro, On the regularity of solutions to a nonlinear ultraparabolic equation arising in mathematical finance. Differential Integral Equations 14 (2001), no. 6, 701-738.

[18] G. Citti and A. Sarti, A cortical based model of perceptual completion in the Roto-Translation space, Journal of Mathematical Analysis and Vision, (2006), 24, 307 – 326.

[19] T. Chan, S. Esedoglu, F. Park and A. Yip, Total Variation Image Restoration: Overview and Recent Developments Handbook of Mathematical Models in Computer Vision, (2006), 17-31.

[20] R. Coifman and G. Weiss, Analyse harmonique non-commutative sur certains espaces homognes. (french) Etude de certaines intégrales singulières, Lecture Notes in Mathematics, Vol. 242, Springer-Verlag, 1971.

[21] D. Danielli, N. Garofalo, D.M. Nhieu and S.D. Pauls, Instability of graphical strips and a positive answer to the Bernstein problem in the Heisenberg group H^1, J. Differ. Geom, (2009), 81(2), 251-295.

[22] N. Dirr, F. Dragoni and M. von Renesse, Evolution by mean curvature flow in sub-Riemannian geometries: A stochastic approach, Communications on Pure and Applied Mathematics, 9 (2), (2010) pp. 307-326.

[23] R. Duits and E. M. Franken, Left invariant parabolic evolution equations on SE(2) and contour enhancement via invertible orientation scores, part I: Linear left-invariant diffusion equations on SE(2), Quarterly of Applied Mathematics, AMS, (2010), 68, 255-292.

[24] R. Duits and E. Franken, Left invariant parabolic evolution equations on SE(2) and contour enhancement via invertible orientation scores, part II: Nonlinear left-invariant diffusion equations on invertible orientation scores,Quarterly of Applied Mathematics, AMS, (2010), 68, 293-331.

[25] R. Duits, H. Fuhr, B.J. Janssen, M. Bruurmijn, L.M.J. Florack and H.A.C.van Assen, Evolution Equations on Gabor Transforms and their Applications, Applied and Computational Harmonic Analysis, accepted for publication, to appear

[26] G. B. Folland, Subelliptic estimates and function spaces on nilpotent Lie groups, Ark. Mat. 2 (1975), no. 13, 161-207.

[27] G. B. Folland, Harmonic Analysis in Phase Space, Annals of Mathematics Studies, 122, Princeton Univ. Press, Princeton, 1989.

[28] G. B. Folland and E. M. Stein, Hardy spaces on homogeneous groups, Princeton University Press, Princeton, New Jersey, 1982.

[29] A. Friedman, Partial differential equations of parabolic type, Prentice-Hall, Inc., Englewood Cliffs, N.J. 1964 xiv+347 pp.

[30] M. Gromov, Metric structures for Riemannian and non-Riemannian spaces, Birkhauser Boston Inc., Boston, MA, 1999, Edited by J. LaFontaine and P. Pansu. Based on the 1981. French original, Translated from the French by Sean Michael Bates.

[31] A. A. Grigor'yan, The heat equation on noncompact Riemannian manifolds. (Russian) Mat. Sb. 182 (1991), no. 1, 55-87; translation in Math. USSR-Sb. 72 (1992), no. 1, 47-77.

[32] L. Hormander, Hypoelliptic second order differential equations, Acta Math. (1967), no. 119, 147-171.

[33] D. S. Jerison and A. Sanchez-Calle, Estimates for the heat kernel for a sum of squares of vector fields, Indiana Univ. Math. J. 35 (1986), no. 4, 835-854.

[34] O. A. Ladyzenskaja, V. A. Solonnikov, and N. N. Uraltseva, Linear and quasilinear equations of parabolic type, Translated from the Russian by S. Smith. Translations of Mathematical Monographs, Vol. 23, American Mathematical Society, Providence, R.I., 1968 (Russian).

[35] G. M. Lieberman, Second order parabolic differential equations, World Scientific Publishing Co. Inc., River Edge,

NJ, 1996.

[36] F. Ferrari, J. Manfredi and Q. Liu, On the horizontal Mean Curvature Flow for Axisymmetric surfaces in the Heisenberg Group, preprint (2012).

[37] J. S. Moll, The anisotropic total variation flow. (English summary) Math. Ann. 332 (2005), no. 1, 177–218.

[38] R. Monti and M. Rickly, Convex isoperimetric sets in the Heisenberg group, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 8 (2009), no. 2, 391–415.

[39] J. Moser, A Harnack inequality for parabolic differential equations, Comm. Pure Appl. Math. 17 (1964), 101–134.

[40] A. Nagel, E. M. Stein, and S. Wainger, Balls and metrics defined by vector fields. I. Basic properties, Acta Math. 155 (1985), no. 1–2, 103–147.

[41] S.D. Pauls: Minimal surfaces in the Heisenberg group, Geom. Dedic., (2004), 104, 201–231.

[42] M. Ritoré and C. Rosales, Area–stationary surfaces in the Heisenberg group H^1, Adv. Math. 219 (2008), 633–671.

[43] L. P. Rothschild and E. M. Stein, Hypoelliptic differential operators and nilpotent groups, Acta Math. 137 (1976), no. 3–4, 247–320.

[44] L. I. Rudin, S. Osher and E. Fatemi, Nonlinear total variation based noise removal algorithms, Physica D 60 (1992), 259–268.

[45] L. Saloff-Coste, A note on Poincaré, Sobolev, and Harnack inequalities, Internat. Math. Res. Notices 1992 (1992), no. 2, 27–38.

[46] F. Serra Cassano and D. Vittone, Graphs of bounded variation, existence and local boundedness of non–parametric minimal surfaces in Heisenberg groups, (accepted for publication in Advances in Calculus of Variations).

[47] E. M. Stein, Harmonic analysis: real–variable methods, orthogonality, and oscillatory integrals, Princeton Mathematical Series, vol. 43, Princeton University Press, Princeton, NJ, 1993, With the assistance of Timothy S. Murphy, Monographs in Harmonic Analysis, III.