Nonlocal diffusion equations in Carnot groups

Isolda E. Cardoso1 · Raúl E. Vidal2

Received: 9 March 2022 / Accepted: 11 June 2022 / Published online: 25 July 2022
© The Author(s), under exclusive licence to Springer-Verlag Italia S.r.l., part of Springer Nature 2022

Abstract
Let \( G \) be a Carnot group. We study nonlocal diffusion equations in a domain \( \Omega \) of \( G \) of the form

\[
u_t^\varepsilon(x, t) = \frac{1}{t} \int_G \frac{1}{\varepsilon^2} K_\varepsilon(x, y)(u_t^\varepsilon(y, t) - u_t^\varepsilon(x, t)) \, dy, \quad x \in \Omega
\]

with \( u^\varepsilon = g(x, t) \) for \( x \notin \Omega \). For an appropriated rescaled kernel \( K_\varepsilon \), we apply the Taylor series development in Carnot groups in order to prove that the solutions \( u^\varepsilon \) uniformly approximate the solution of a certain local Dirichlet problem in \( \Omega \), when \( \varepsilon \to 0 \).

Keywords Nonlocal diffusion equations · Carnot groups · Taylor expansion · Dirichlet problem

Mathematics Subject Classification 35R03 · 35K57 · 45L05

1 Introduction

Diffusion processes have been modeled in the Euclidean space \( \mathbb{R}^n \) by the so-called nonlocal diffusion problems. Indeed, in [8] they consider the probabilist density function of a single population at the point \( x \) at time \( t \), namely \( u(x, t) \), and denote by \( J \) a symmetric function such that \( \int_{\mathbb{R}^n} J(x) \, dx = 1 \) which they use to model the evolution in the following sense: they have that \( J(x - y) \) is the probability distribution of jumping from location \( y \) to location \( x \); \( J * u(x, t) = \int_{\mathbb{R}^n} J(y - x)u(y, t) \, dy \) is the rate at which individuals are arriving to position \( x \) from all other places, and \( -u(x, t) = \int_{\mathbb{R}^n} J(x - y)u(x, t) \, dy \) is the rate at which they are leaving location \( x \) to travel to all other sites. Thus, they have that \( u \) satisfies a nonlocal evolution equation of the form

---

\(1\) Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Pellegrini 250, Rosario, Santa Fe, Argentina

\(2\) Facultad de Matemática, Astronomía y Física, Medina Allende s/n, Córdoba, Córdoba, Argentina
The diffusion processes and this kind of models have been applied to very different contexts. For example in biology [4] and [13], image processing [12], particle systems [3], coagulation models [10], etc.

In the work [6] the authors prove that solutions of properly rescaled nonlocal Dirichlet problems of the Eq. (1.1) uniformly approximate the solution of the corresponding Dirichlet problem for the classical heat equation in $\mathbb{R}^n$. Also, some of these results have been generalized for kernels that are not convolution.

It is always interesting how this kind of problems translate to other settings. For example, in [17] the author considers a nonlocal diffusion problem on the Heisenberg group and analogous results to those obtained in the works [5] and [6]. In our work we will consider the following problems, originally set in $\mathbb{R}^n$, in the context of Carnot groups. This passage is not straightforward and needs to be handled carefully. Recall that both the Euclidean space and the Heisenberg group are examples of Carnot groups.

In the work [14] the authors prove that smooth solutions to the Dirichlet problem for the parabolic equation

$$u_t(x, t) = J * u(x, t) - u(x, t).$$

(1.1)

properly rescaled, where $\int_{\mathbb{R}^n} J(x) dx = 1$ and $u(x, t) = g(x, t)$ for $x \notin \Omega$.

The study of Carnot groups and PDE’s on them has been increasing in the last years, since the topology is similar to the Euclidean topology and the hypoelliptic equations are easily defined (see the fundamental work of Hörmander [11]). The study of regularity of the solutions is not trivial and constitutes an active topic of interest nowadays. There are examples of smooth domains where the problems with smooth initial data have solutions which are not smooth in certain points (see [1] and [2]).

A Carnot group is a simply connected and connected Lie group $G$, whose Lie algebra $\mathfrak{g}$ is stratified, this means that $\mathfrak{g}$ admits a vector space decomposition $\mathfrak{g} = V_1 \oplus \cdots \oplus V_m$. 
with grading \( |V_1, V_j| = V_{j+1} \), for \( j = 1, \ldots, m - 1 \), and has a family of dilations \( \{ \delta \} \) such that \( \delta X = e^t X \) if \( X \in V_j \).

Let be \( \{ X_1, \ldots, X_{n_i} \} \) a basis of \( V_1 \) and \( \{ X_{n_i+1}, \ldots, X_{n_i+n_2} \} \) a basis of \( V_2 \) and let \( \Omega \subset G \) be a bounded \( C^{2+\alpha} \), \( 0 < \alpha < 1 \), domain (that is, open and connected). The smoothness condition on \( \Omega \) means that the boundary is the graph of a \( C^{2+\alpha} \) function, its precise definition will be given in Sect. 2.2.1.

We consider the following second order local parabolic differential equation with Dirichlet boundary conditions

\[
\begin{cases}
    v_j(x, t) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} a_{ij}(x) X_i X_j v(x, t) + \sum_{i=1}^{n_1+n_2} b_i(x) X_i v(x, t), & x \in \Omega, t > 0, \\
    v(x, t) = g(x, t), & x \in \partial \Omega, t > 0, \\
    v(x, 0) = u_0(x), & x \in \Omega,
\end{cases}
\tag{1.2}
\]

where the coefficients \( a_{ij}(x), b_i(x) \) are smooth in \( \overline{\Omega} \) and \( (a_{ij}(x)) \) is a symmetric positive definite matrix, i.e., \( \sum_{ij} a_{ij}(x) \xi_i \xi_j \geq \nu |\xi|^2 \) for every real vector \( \xi = (\xi_1, \ldots, \xi_{n_1}) \neq 0 \) and for some \( \nu > 0 \). Also we have the following nonlocal rescaled Dirichlet problem

\[
\begin{cases}
    u^\epsilon(x, t) = K_\epsilon(u^\epsilon)(x, t), & x \in \Omega, t > 0, \\
    u^\epsilon(x, t) = g(x, t), & x \notin \Omega, t > 0, \\
    u^\epsilon(x, 0) = u_0(x), & x \in \Omega,
\end{cases}
\tag{1.3}
\]

where \( K_\epsilon \) is the nonlocal operator

\[
K_\epsilon u(x) = \int_\Omega K_\epsilon(x, y)(u(y) - u(x)) dy,
\]

defined by a rescaled \( K_\epsilon \) kernel of the form

\[
K_\epsilon(x, y) = \frac{c(x)}{e^{Q + \delta}} \hat{E}(x, y) \hat{L}(x, y),
\]

with the functions \( c, \hat{E} \) and \( \hat{L} \) suitable defined from the coefficients \( a_{ij} \) and \( b_i \) from the differential operator in (1.2). Also, \( \epsilon > 0 \) and \( Q \) is the homogeneous dimension of \( G \). For the full description of \( K_\epsilon \) we need more notation since we have to go back and forth from the Carnot group to its Lie algebra, hence we will postpone it until Sect. 3.2, that is, after we have set the notation.

By means of the nontrivial Taylor series development we will prove the next Theorem:

**Theorem 1.1** Assume that problem (1.2) has a solution \( v \) in \( C^{2+\alpha,1+\alpha}(\overline{\Omega} \times [0, T]) \). Let \( u^\epsilon \) be the solution of problem (1.3) where \( K_\epsilon \) is defined by formula (3.10), \( g \in C^{2+\alpha,1+\alpha}(\Omega^\epsilon \times [0, T]) \) and \( u_0 \in C^{2+\alpha}(\overline{\Omega}) \). Then there exists a positive constant \( c \) such that

\[
||u^\epsilon - v||_{L^{\infty}(\Omega^\epsilon \times [0, T])} \leq c \epsilon^\alpha.
\]

We also study the Fokker-Planck parabolic problem with Dirichlet condition.
\[
\begin{aligned}
v_t(x, t) &= \sum_{i=1}^{n_x} X_iX_i(a(\cdot)v(\cdot, t))(x), & x \in \Omega, & t > 0, \\
v(x, t) &= g(x, t), & x \in \partial \Omega, & t > 0, \\
v(x, 0) &= u_0(x), & x \in \Omega,
\end{aligned}
\] (1.4)

where the coefficient \( a \in C^\infty(G) \); and the nonlocal reescaled Dirichlet problem given by

\[
\begin{aligned}
u^\varepsilon_t(x, t) &= \mathcal{L}_c(u^\varepsilon)(x, t), & x \in \Omega, & t > 0, \\
u^\varepsilon(x, t) &= g(x, t), & x \notin \Omega, & t > 0, \\
u^\varepsilon(x, 0) &= u_0(x), & x \in \Omega,
\end{aligned}
\] (1.5)

where \( \mathcal{L}_c \) is defined as

\[
\mathcal{L}_c(u)(x) = \frac{2C(J)}{e^{\theta+2}} \int_{G} J(\delta_{y^{-1}}x)[a(y)u(y) - a(x)u(x)]dy,
\]

where all of its elements will be clear after we give the notation in the next section. We will also recall this definition in Sect. 3.3 for the benefit of the reader.

Again we apply the Taylor series development to prove the next Theorem:

**Theorem 1.2** Assume that problem (1.4) has a solution \( v \) in \( C^{2+\alpha, 1+\alpha}([\Omega \times [0, T]) \). Let \( u^\varepsilon \) be the solution of problem (1.5) where \( \mathcal{L}_c \) is defined by formula (3.13), \( g \in C^{2+\alpha, 1+\alpha}([\Omega^c \times [0, T]) \) and \( u_0 \in C^{2+\alpha}(\Omega) \). Then there exists a positive constant \( c \) such that

\[
||u^\varepsilon - v||_{L^\infty([\Omega \times [0, T])} \leq ce^{a}.
\]

The rest of the paper is organized as follows. In Sect. 2 we recall some definitions and results on Carnot groups and set the notation to be used later. In Sect. 3 we define and study the operators \( \mathcal{K}_c \) and \( \mathcal{L}_c \). In Sect. 4 we study the existence, uniqueness and other properties of the solutions of the problems (1.3) and (1.5). In Sect. 5 we prove the Main Theorems 1.1 and 1.2.

## 2 Preliminaries

### 2.1 Carnot groups

Let \( G \) be a Carnot group and let \( \mathfrak{g} = V_1 \oplus \cdots \oplus V_m \) denote its (stratified) Lie algebra, as stated in the introduction. The grading of the Lie algebra is closely related to the dilations. Indeed, we can choose a basis \( \beta = \{X_1, \ldots, X_n\} \) of \( \mathfrak{g} \) which is adapted to the grading as we describe as follows. For each \( r > 0 \) the dilation \( \delta_r \) is the automorphism defined by \( \delta_r = \text{Exp}(A \log r) \), where \( \text{Exp} \) denotes the matrix exponential function, and \( A \) is a diagonalizable linear transformation on \( \mathfrak{g} \) with positive eigenvalues \( 1 = \lambda_1 \leq \cdots \leq \lambda_n \).

We also have that...
\[ [A]_\beta = \begin{pmatrix} \lambda_1 & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \lambda_n \end{pmatrix}, \quad \text{and} \quad [\delta_r]_\beta = \begin{pmatrix} r^{\lambda_1} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & r^{\lambda_n} \end{pmatrix}. \]

and moreover,

- the set of eigenvalues of \( A \) is \( \lambda = \{1, 2, \ldots, m\} \),
- the set of eigenvalues for each \( \delta_r \), \( r > 0 \), is \( \{r^1, \ldots, r^m\} \),
- if \( \dim(V_k) = n_k \) for \( 1 \leq k \leq m \), then \( n = n_1 + \cdots + n_m \), and
  - \( \beta_1 = \{X_1, \ldots, X_{n_1}\} \) is a basis of \( V_1 \) of eigenvectors associated to the eigenvalue \( \lambda = 1 \),
  - \( \beta_2 = \{X_{n_1+1}, \ldots, X_{n_1+n_2}\} \) is a basis of \( V_2 \) of eigenvectors associated to the eigenvalue \( \lambda = 2 \),
  - \( \ldots \)
  - \( \beta_m = \{X_{n_1+\cdots+n_{m-1}+1}, \ldots, X_n\} \) is a basis of \( V_m \) of eigenvectors associated to the eigenvalue \( \lambda = m \).
- the homogeneous dimension is \( Q = \text{trace}(A) = \sum_{k=1}^{m} kn_k \).

Since \( g \) is nilpotent, we can identify \( \mathbb{R}^n \) with the group \( G \) via the exponential map: let \( \varphi : \mathbb{R}^n \to G \) be such that every \( (x_1, \ldots, x_n) \in \mathbb{R}^n \) is identified with \( \varphi(t_1, \ldots, t_n) = \exp(t_1X_1 + \cdots + t_nX_n) \); thus \( \varphi^{-1} \) defines a global chart on the Carnot group \( G \). Also, we may resort at some point to the identification \( \varphi = (\phi_1, \ldots, \phi_n) : g \to \mathbb{R}^n \) such that if \( X = t_1X_1 + \cdots + t_nX_n \in g \), then \( \phi(X) = (t_1, \ldots, t_n) \), and \( \phi_j(X) = t_j \). Observe that if \( x \in G \), \( \phi_j(\exp^{-1}(x)) = \pi_j(\varphi^{-1}(x)) \), where \( \pi_j : \mathbb{R}^n \to \mathbb{R} \) denotes the projection.

The Lebesgue measure on \( g \) induces a biinvariant Haar measure on \( G \), and we fix the normalization of Haar measure on \( G \) by requiring that the measure of the unitary ball to be 1. We shall denote with \( |E| \) the measure of a measurable set \( E \) and with \( \int f = \int fd\mu \) the integral of a function \( f \) with respect to this measure. Hence, \( |\delta_r(E)| = r^Q |E| \) and \( d(rx) = r^Q dx \). In particular, \( |B(r, s)| = r^Q \) for all \( r > 0 \) and \( x \in G \).

We will also use the following multiindex notation: if \( I = (i_1, \ldots, i_n) \in \mathbb{N}_0^n \), we set \( x^I = X_1^{i_1}X_2^{i_2} \cdots X_n^{i_n} \). The operators \( X^I \) form a basis for the algebra of left invariant differential operators on \( G \), by the Poincaré-Birkhoff-Witt Theorem. The order of the differential operators \( X^I \) is \( |I| = i_1 + i_2 + \cdots + i_n \) and its homogeneous degree is \( d(I) = \lambda_1 i_1 + \lambda_2 i_2 + \cdots + \lambda_n i_n \). The additive subgroup of \( \mathbb{R} \) generated by \( 0, \lambda_1, \ldots, \lambda_n \) is \( \triangle \). Observe that \( \triangle = \{d(I) : I \in \mathbb{N}_0^n \} \supset \mathbb{N} \) (since \( \lambda_1 = 1 \), which is no other than \( \triangle = \mathbb{N} \). Finally, \( X \in g \) is homogeneous of degree \( k \) if and only if \( X \in V_k \). Let us now define \( J = \sum_{j=1}^{m} X_j^2 \), thus \( -J \) is a left invariant second order differential operator which is homogeneous of degree 2 called the subLaplacian of \( G \) (relative to the stratification and the basis). Its role on \( G \) is analogous to (minus) the ordinary Laplacian in \( \mathbb{R}^n \).

### 2.2 Taylor polynomials in homogeneous Lie groups

Now we are going to recall some concepts and notations on the definition of Taylor polynomials for homogeneous Lie groups from [9]. We say that a function \( P \) on \( G \) is a polynomial if \( P \circ \exp \) is a polynomial on \( g \). Let \( \{\xi_1, \ldots, \xi_n\} \) be the basis for the linear forms on \( g \) dual to the basis \( \{X_1, \ldots, X_n\} \) on \( g \). Let us consider \( \eta_1, \ldots, \eta_n \) are polynomials on \( G \) which form a global coordinate system on \( G \), and generate the algebra of polynomials on \( G \). Thus, every polynomial on \( G \) can be written uniquely as \( P = \sum a_j \eta_j^q \), for
If \( \eta^j \) is homogeneous of degree \( d(I) \), the set of possible degrees of homogeneity for polynomials is the set \( \triangle \). We call the degree of a polynomial \( \max \{ |I| : a_I \neq 0 \} \) the isotropic degree, and its homogeneous degree is \( \max \{ d(I) : a_I \neq 0 \} \). For each \( N \in \mathbb{N} \) we define the space \( \mathcal{P}^N \) of polynomials of isotropic degree \( \leq N \), and for each \( j \in \triangle \) we define the space \( \mathcal{P}_j \) of polynomials of homogeneous degree \( \leq j \). It follows that \( \mathcal{P}_N \subset \mathcal{P}^N \subset \mathcal{P}_{2N} \). The space \( \mathcal{P}_j \) is invariant under left and right translations (see Proposition 1.25 of [9]), but the space \( \mathcal{P}^N \) is not (unless \( N = 0 \) or \( G \) is abelian). For a function \( f \) whose derivatives \( X^l f \) are continuous functions on a neighbourhood of a point \( x \in G \), and for \( j \in \triangle \) such that \( d(I) \leq j \) we define the left Taylor polynomial of \( f \) at \( x \) of homogeneous degree \( j \) to be the unique polynomial \( P \in \mathcal{P}_j \) such that

\[
X^j P(e) = X^j f(x).
\]

Here, we have that \( X^j f(x) = \frac{\partial^{d(I)}}{\partial^1 t_1 \cdots \partial^N t_N} f \left( x \exp \sum_{j=1}^n t_j X_j \right) \bigg|_{t_1 = \cdots = t_N = 0} \). From now on we are going to consider every Taylor polynomial as a left Taylor polynomial, hence we will drop the word left.

We will be using the Taylor polynomial of a function \( f \) at a point \( x \in G \) of homogeneous degree 2, hence we will explicitly show its form with the notation we have presented. Let us call it \( P = P_{x,G} = \sum_{I \in \mathbb{N}^n : d(I) \leq 2} a_I \eta^I \). If the multiindex \( I \in \mathbb{N}^n \) has homogeneous degree \( d(I) = 0 \), \( I = \emptyset \). If \( d(I) = 1 \) then \( I = \bar{e}_j \) for \( 1 \leq j \leq n_1 \) (where \( \bar{e}_j \) denotes the canonical unitary vectors of \( \mathbb{R}^n \), whose \( i \)-component is defined as \( \delta_{ij} \)), and if \( d(I) = 2 \) then either \( I = \bar{e}_i + \bar{e}_j \) for \( 1 \leq i, j \leq n_1 \) or \( I = \bar{e}_j \) for \( n_1 + 1 \leq j \leq n_1 + n_2 \). Hence, for \( y \in G \),

\[
P(y) = a_0 + \sum_{j=1}^{n_1+n_2} a_{\bar{e}_j} \eta^{\bar{e}_j}(y) + \sum_{i,j=1}^{n_1} a_{\bar{e}_i+\bar{e}_j} \eta^{\bar{e}_i+\bar{e}_j}(y).
\]

In order to explicitly state the coefficients we perform some straightforward computations, namely:

- If \( I = \emptyset \) then since \( X^0 P(e) = X^0 f(x) \) we have that \( a_0 = f(x) \).
- Let us consider \( I = \bar{e}_j \) for \( 1 \leq j \leq n_1 + n_2 \). Then for \( y = \exp \sum_{i=1}^n t_i X_i \) we have that \( \eta_{\bar{e}_j}(y) = (\xi_j \circ \exp^{-1})(y) = t_j = (\pi_j \circ \varphi^{-1})(y) \). Since \( X^j P(e) = X^j f(x) \) it follows that \( a_{\bar{e}_j} = X^j f(x) \).
- Similarly, if we consider \( I = \bar{e}_i + \bar{e}_j \) for \( 1 \leq i, j \leq n_1 \) we have that \( \eta^{\bar{e}_i+\bar{e}_j}(y) = t_i t_j = (\pi_i \circ \varphi^{-1})(y)(\pi_j \circ \varphi^{-1})(y) \). And from the equality \( X^{\bar{e}_i+\bar{e}_j} P(e) = X^{\bar{e}_i+\bar{e}_j} f(x) \) it follows that \( a_{\bar{e}_i+\bar{e}_j} + a_{\bar{e}_i+\bar{e}_j} = X^j X^j f(x) \).

We are now able to present the Taylor polynomial \( P \) in a more familiar form:

\[
P(y) = f(x) + \sum_{j=1}^{n_1+n_2} (\pi_j \circ \varphi^{-1})(y) X^j f(x) + \frac{1}{2} \sum_{i,j=1}^{n_1} (\pi_i \circ \varphi^{-1})(y)(\pi_j \circ \varphi^{-1})(y) X^j X^j f(x).
\]

And if we are considering coordinates,

\[
P \left( \exp \left( \sum_{i=1}^n t_i X_i \right) \right) = f(x) + \sum_{j=1}^{n_1+n_2} t_j X^j f(x) + \frac{1}{2} \sum_{i,j=1}^{n_1} t_i t_j X^j X^j f(x). \tag{2.1}
\]
2.2.1 Stratified Taylor inequality

Throughout this section we will consider a fixed stratified group $G$ with the notation described previously. We will regard the elements of the basis of $\mathfrak{g}$ adapted to the gradation as left invariant differential operators on $G$.

Since $V_i$ generates $\mathfrak{g}$ as a Lie algebra, we have that $\exp(V_i)$ generates $G$. More precisely:

**Lemma 2.1** (Lemma 1.40 of [9]) If $G$ is stratified there exist $C > 0$ and $N \in \mathbb{N}$ such that any $x \in G$ can be expressed as $x = x_1 \ldots x_N$ with $x_j \in \exp(V_i)$ and $|x_j| \leq C|x|$, for all $j$.

For $k \in \mathbb{N}$ we define $C^k(G)$ to be the space of continuous functions $f$ on $G$ whose derivatives $X^I f$ are continuous functions on $G$ for $d(I) \leq k$. Also, for $0 < \alpha < 1$ we define the space $C^{k+\alpha}(G)$ as the function $f$ in $C^k(G)$ where

$$\sup_{x,y \in G, d(I) = k} |X^I f(xy) - X^I f(x)| < C|y|^{\alpha},$$

with $C$ independent of $x$ and $y$.

Let us also define the space $C^k(\Omega)$ of those functions $f$ defined on $\Omega$ such that $Df$ is continuous for every differential operator $D$ of homogeneous degree less or equal to $k$.

Another important consequence of the fact that $V_i$ generates $\mathfrak{g}$ is that the set of left invariant differential operators which are homogeneous of degree $k$ (which is the linear span of $\{X^I : d(I) = k\}$) is precisely the linear span of the operators $X_{i_1} \ldots X_{i_k}$ with $1 \leq i_j \leq n_i$ for $j = 1, \ldots, k$. We thus have the following results:

**Theorem 2.2** (Theorem 1.41, Stratified Mean Value Theorem, [9])

Suppose $G$ is stratified. There exist $C > 0$ and $b > 0$ such that for all $f \in C^1$ and all $x, y \in G$,

$$|f(xy) - f(x)| \leq C|y| \sup_{|z| \leq b|y|, 1 \leq k \leq n_1} |X_k f(xz)|.$$

**Theorem 2.3** (Theorem 1.42, Stratified Taylor Inequality, of [9])

Suppose $G$ is stratified. For each positive integer $k$ there is a constant $C_k$ such that for all $f \in C^k$ and all $x, y \in G$,

$$|f(xy) - P_x(y)| \leq C_k|y|^k \eta(x, b^k|y|),$$

where $P_x$ is the Taylor polynomial of $f$ at $x$ of homogeneous degree $k$, $b$ is as in the Stratified Mean Value Theorem, and for $r > 0$,

$$\eta(x, r) = \sup_{|z| \leq r, d(I) = k} |X^I f(xz) - X^I f(x)|.$$

For a function $f \in C^k(\Omega)$ and $x \in \Omega$ let $P = P_{f,x,k}$ denote the Taylor polynomial of $f$ at $x$ of homogeneous degree $k$. By Theorem 2.3 we have that for $\epsilon > 0$,

$$\frac{1}{\epsilon^k} |f(x + \epsilon \delta_x y) - P(y)| \leq C_{\epsilon,k}|\delta_x y|^k \eta(x, b^k|\delta_x y|) = c_{\epsilon,k}|y|^k \eta(x, b^k|\delta_x y|), \quad (2.2)$$

Springer
which goes to 0 as $\epsilon$ does.

Hence, if $f$ in $C^{2+\alpha}(G)$, with $0 < \alpha < 1$ we have the following Taylor expansion of $f$ at $x$ of homogeneous degree $k = 2$: for

$$
f\left(x\exp\left(\sum_{i=1}^{n} t_i X_i\right)\right) = f(x) + \sum_{j=1}^{n_1+n_2} t_j X_j f(x) + \frac{1}{2} \sum_{i,j=1}^{n} t_i t_j X_i X_j f(x) + o(|t_1 X_1 + \cdots + t_n X_n|^2),
$$

in the sense that

$$
\lim_{\epsilon \to 0} \frac{o(|\delta_\epsilon(t_1 X_1 + \cdots + t_n X_n)|^2)}{\epsilon^2} = 0.
$$

Indeed, by (2.1), (2.2) and the fact that $f \in C^{2+\alpha}(G)$, we get

$$
\left| \frac{o(|\delta_\epsilon(t_1 X_1 + \cdots + t_n X_n)|^2)}{\epsilon^2} \right| \\
\leq c \exp(t_1 X_1 + \cdots + t_n X_n)^2 |\eta(x, b^2| \exp(\delta_\epsilon(t_1 X_1 + \cdots + t_n X_n))|) \\
= c |\exp(t_1 X_1 + \cdots + t_n X_n)|^2 \sup_{|z| \leq b^2} |X^i f(x) - X^i f(x)| \\
\leq c |\exp(t_1 X_1 + \cdots + t_n X_n)|^2 b^2 \exp(\delta_\epsilon(t_1 X_1 + \cdots + t_n X_n))|^\alpha \\
= c b^2 |\exp(t_1 X_1 + \cdots + t_n X_n)|^{2+\alpha} e^\alpha.
$$

Before we move forward, let us remark that throughout the work we will denote with $c$ a positive constant that may vary from line to line.

### 3 Some nonlocal diffusion problems

Throughout this section we let $G$ be a Carnot group with Lie algebra $\mathfrak{g}$ and let $\Omega$ be an open, bounded and connected subset of $G$. The aim of this section is to properly define the operators $\mathcal{K}_\epsilon$ and $\mathcal{L}_\epsilon$ from the introduction. In order to understand the techniques involved, we will first work with an evolution operator of a much simpler form (namely the operator given in (1.1)), in the context of the Carnot group $G$. We will see that the solutions to the nonlocal Dirichlet rescaled problems uniformly approximate the solution of the classical heat equation with Dirichlet conditions.

Let us consider a positive and radial function $J \in L^1(G)$ with compact support $F$, normalized such that $\int_G J dx = 1$, whence for $i = 1, \ldots, n$

$$
\int_{\mathbb{R}^n} J(\exp(t_1 X_1 + \cdots + t_n X_n)) t_1 dt_1 \cdots dt_n = 0;
$$

and also

$$
\sum_{i} t_i X_i f(x) + o(|t_1 X_1 + \cdots + t_n X_n|^2),
$$
for a constant \( C(J) > 0 \), \( i = 1, \ldots, n \). From both properties it follows that for \( i, j = 1, \ldots, n \),
\[
\int_{\mathbb{R}^n} J(\exp(t_1 X_1 + \cdots + t_n X_n))t_i^2 dt_1 \cdots dt_n = C(J),
\]
(3.2)

for each \( \epsilon > 0 \) we define the rescaled operator
\[
\mathcal{E}_\epsilon u(x) = \frac{1}{\epsilon^2} [(u \ast J_\epsilon)(x) - u(x)],
\]
(3.6)

we have that
\[
\mathcal{E}_\epsilon u(x) = \frac{1}{\epsilon^2} \left[ (u \ast J_\epsilon)(x) - u(x) \right] = \frac{1}{\epsilon^2} \int_G u(xy^{-1})J_\epsilon(y)dy - u(x)
\]
(3.7)

\[= \frac{1}{\epsilon^2} \int_G u(xy^{-1}) \frac{1}{\epsilon Q} J_{\delta_{\epsilon} y} dy - u(x)\]
\[= \frac{1}{\epsilon^2} \int_G u(x(\delta_{\epsilon}(y))^{-1}) \frac{1}{\epsilon Q} J(y)dy - \int_G J(y)u(x)dy\]
\[= \frac{1}{\epsilon^2} \int_G \left[ u(x(\delta_{\epsilon}(y)^{-1})) - u(x) \right] J(y)dy.\]

\[E\]

\[u(x) = \int_G K_\epsilon(x, y)(u(y) - u(x))dy,\]
for \( G = \mathbb{R}^n \), where the kernel \( K_\epsilon(x, y) \) is a positive function with compact support in \( \Omega \times \Omega \) for \( \Omega \subset G \) a bounded domain such that \( 0 < \sup_{y \in \Omega} K_\epsilon(x, y) = c_\epsilon(x) \in L^\infty(\Omega)\).
Following the ideas of Molino and Rossi, let us consider:

- A function $J$ as defined in the beginning of the section.
- A $n_1 \times n_1$ symmetric and positive definite matrix $\tilde{A}(x) = (a_{ij}(x))$, where the coefficients are smooth in $\bar{\Omega}$ with $\tilde{A}(x) = \tilde{L}(x)\tilde{L}'(x)$ its Cholesky factorization, with $\tilde{L}(x) = (l_{ij}(x))$ and $\tilde{L}^{-1}(x) = (l_{ij}^{-1}(x))$. Also, let $A(x)$ be the $n \times n$ matrix defined by blocks as follows:

$$A(x) = \begin{pmatrix} A(x) & 0 \\ 0 & 1 \\ 0 & \ddots \\ 0 & 1 \end{pmatrix}.$$ 

That is, $A(x)$ is the matrix $\tilde{A}(x)$ extended by the identity to size $n \times n$, and let $L(x)$ and $L'(x)$ be similarly defined.

- A $n_1 \times n_1$ symmetric and positive definite matrix $\tilde{A}(x) = (a_{ij}(x))$, where the coefficients are smooth in $\bar{\Omega}$ with $\tilde{A}(x) = \tilde{L}(x)\tilde{L}'(x)$ its Cholesky factorization, with $\tilde{L}(x) = (l_{ij}(x))$ and $\tilde{L}^{-1}(x) = (l_{ij}^{-1}(x))$. Also, let $A(x)$ be the $n \times n$ matrix defined by blocks as follows:

$$A(x) = \begin{pmatrix} A(x) & 0 \\ 0 & 1 \\ 0 & \ddots \\ 0 & 1 \end{pmatrix}.$$ 

That is, $A(x)$ is the matrix $\tilde{A}(x)$ extended by the identity to size $n \times n$, and let $L(x)$ and $L'(x)$ be similarly defined.

- A $n \times n$ diagonal matrix $W(x) = \text{diag}(\tilde{b}_1(x), \ldots, \tilde{b}_n(x))$ where $\tilde{b}_i(x) = b_i(x)$ if $1 \leq i \leq n_1$, $\tilde{b}_i(x) = \frac{b_i(x)}{c^2}$ if $n_1 < i \leq n_1 + n_2$ and $\tilde{b}_i(x) = 1$ if $n_1 + n_2 < i \leq n$.
- A function $a : G \rightarrow \mathbb{R}$ defined by $a(x) = \sum_{i=1}^n \phi_i(x)\exp^{-1}(x^{-1}) + M$, where $M > 0$ is large enough to ensure $a(x) \geq \beta > 0$ for $x$ belonging to an appropriate set $F'$ defined as

$$F' = \{ x \in G : x = y \exp \delta x L(y) \exp^{-1}(x^{-1}), \forall y \in \Omega, \forall z \in F \},$$

where $F$ is the support of $J$.

Thus, we will work with the scaled kernels defined for each $c > 0$ by

$$K_c(x, y) = \frac{c(x)}{c^{\frac{n+2}{2}}} a((\exp(E(x)\exp^{-1}(y^{-1}x)))^{-1}) \times J(\exp(L^{-1}(x)\exp^{-1}(\delta^{-1}y^{-1}x))),$$

where for $x \in G$, $c(x) = \frac{2}{c(J)^{\frac{n+2}{2}}}$ and $E(x) = \frac{M}{2} W(x) A(x)^{-1}$. Let us remark that we understand the action of a $n \times n$ matrix $M$ on $q$ via the identification $\phi$ with $\mathbb{R}^n$ (with respect to the basis $\beta$) as follows: if $M = (m_{ij})$ and $X = \sum x_i X_i \in \mathfrak{g}$,

$$M X = \sum_{i=1}^n \left( \sum_{k=1}^n m_{ik} x_k \right) X_i.$$ 

Also, since the matrix $A(x)$, $L(x)$ and $W(x)$ are defined by blocks (with corresponding blocks of the same size), and the matrix which defines $\delta_c$ is also defined by blocks (again, of the same corresponding sizes) as a constant times the identity on each block, we have that $\delta_c$ commutes with all of them.
Hence, for these rescaled kernels we will study the integral operators

\[ K_\epsilon u(x) = \frac{c(x)}{e^{Q+2}} \int_G a((\exp(E(x) \exp^{-1}(y^{-1}x)))^{-1}) \times J(\exp(L^{-1}(x) \exp^{-1}(\delta_{x^{-1}}y^{-1}x))(u(y) - u(x))dy. \]

(3.10)

More precisely, we will prove that \( K_\epsilon u \) approximates \( Kv \) where \( K \) is the second order operator defined by

\[ K(v)(x) = \sum_{i,j=1}^{n_1} a_{ij}(x)X_iX_jv(x) + \sum_{i=1}^{n_1+n_2} b_i(x)X_iv(x). \]

(3.11)

### 3.3 A reaction-diffusion equation

In [15] the authors work in the same spirit as Molino and Rossi to approximate the solutions of the Fokker-Planck equation by solutions of operators defined by reescaled kernels which in our present context assume the form, respectively:

\[ L(v)(x) = \sum_{i=1}^n X_i(a(x)v(x)), \]

(3.12)

\[ L_\epsilon(u)(x) = \frac{2C(J)}{e^{Q+2}} \int_G J(\delta_{x^{-1}}y^{-1}x)[a(y)u(y) - a(x)u(x)]dy, \]

(3.13)

with the coefficient \( a \in C^\infty(G) \).

### 4 Existence and uniqueness of solutions

We shall now derive the existence and uniqueness of solutions of

\[
\begin{align*}
  u^\epsilon_t(x,t) &= \int_G K_\epsilon(x,y)(u(y,t) - u(x,t))dy \quad \text{for} \quad (x,t) \in \Omega \times [0,T], \\
  u^\epsilon(x,t) &= g(x,t) \quad \text{for} \quad x \notin \Omega, t \in [0,T], \\
  u^\epsilon(x,0) &= u_0(x) \quad \text{for} \quad x \in \Omega,
\end{align*}
\]

(4.1)

which is a consequence of Banach’s fixed point theorem. The main arguments are basically the same of [5] or [6], but we write them here to make the paper self-contained. Let us also remark that the analogous results for operator \( L_\epsilon \) holds and the proofs are completely similar.

Recall the definition of the set \( F' \) (3.8).

**Theorem 4.1** Let \( u_0 \in L^1(\Omega) \) and let \( J \) and \( K_\epsilon \) defined as in Sect. 3.2, with \( K_\epsilon(x,y) \leq C_\epsilon(x) \in L^\infty(\Omega) \) for \( (x,y) \in \Omega \times F' \). Then there exists a unique solution \( u \) of problem (4.1) such that \( u \in C([0,\infty), L^1(\Omega)) \).
\textbf{Proof} We will use the Banach's Fixed Point Theorem. For \( t_0 > 0 \) let us consider the Banach space

\[ X_{t_0} := \{ w \in C([0, t_0]; L^1(\Omega)) \}, \]

with the norm

\[ \|\|w\|\| := \max_{0 \leq t \leq t_0} \|w(\cdot, t)\|_{L^1(\Omega)}. \]

Our aim is to obtain the solution of (4.1) as a fixed point of the operator \( \mathcal{F} : X_{t_0} \rightarrow X_{t_0} \) defined by

\[ \mathcal{F}(w)(x, t) := \begin{cases} w_0(x) + \int_0^t \int_G K_e(x, y)(w(y, r) - w(x, r)) \, dy \, dr & \text{if } x \in \Omega, \\ g(x, t) & \text{if } x \notin \Omega, \end{cases} \]

where \( w_0(x) = w(x, 0) \).

Let \( w, v \in X_{t_0} \). Then there exists a constant \( C \) depending on \( K_e \) and \( \Omega \) such that

\[ \|\|\mathcal{F}(w) - \mathcal{F}(v)\|\| \leq C t_0 \|\|w - v\|\| + \|w_0 - v_0\|_{L^1(\Omega)}. \quad (4.2) \]

Indeed, since if \( x \notin \Omega \) then \( (w - v)(x, t) = 0 \), it follows that

\[
\int_{\Omega} \left| \mathcal{F}(w) - \mathcal{F}(v) \right|(x, t) \, dx \
\leq \int_{\Omega} |w_0 - v_0|(x) \, dx \\
+ \int_{\Omega} \left| \int_0^t \int_G K_e(x, y)((w - v)(y, r) - (w - v)(x, r)) \, dy \, dr \right| \, dx \\
\leq \|w_0 - v_0\|_{L^1(\Omega)} + t \|C_e(x)\|_{L^\infty(\Omega)} \|\|w - v\|\|. 
\]

Taking the maximum in \( t \) (4.2) follows.

Now, taking \( v_0 := v \equiv 0 \) in (4.2) we get that \( \mathcal{F}(w) \in C([0, t_0]; L^1(\Omega)) \) and this says that \( \mathcal{F} \) maps \( X_{t_0} \) into \( X_{t_0} \).

Finally, we will consider \( X_{t_0,t_0} = \{ u \in X_{t_0} : u(x, 0) = u_0(x) \} \). \( \mathcal{F} \) maps \( X_{t_0,t_0} \) into \( X_{t_0,t_0} \) and taking \( t_0 \) such that \( 2 \|C_e(x)\|_{L^\infty(\Omega)} |\Omega| t_0 < 1 \),

we can apply the Banach's fixed point theorem in the interval \([0, t_0]\) because \( \mathcal{F} \) is a strict contraction in \( X_{t_0,t_0} \). From this we get the existence and uniqueness of the solution in \([0, t_0]\). To extend the solution to \([0, \infty)\) we may take as initial data \( u(x, t_0) \in L^1(\Omega) \) and obtain a solution up to \([0, 2t_0]\). Iterating this procedure we get a solution defined in \([0, \infty)\). \( \square \)

In order to prove a comparison principle of the problem given in (4.1) we need to introduce the definition of sub and super solutions.

\textbf{Definition 4.2} A function \( u \in C([0, T]; L^1(\Omega)) \) is a supersolution of (4.1) if

\[
\begin{cases}
  u_r(x, t) \geq \int_G K_e(x, y)(u(y, t) - u(x, t)) \, dy, & \text{for } x \in \Omega \text{ and } t > 0, \\
  u_r(x, t) \geq g(x, t), & \text{for } x \notin \Omega \text{ and } t > 0, \\
  u(x, 0) \geq u_0(x), & \text{for } x \in \Omega.
\end{cases} \quad (4.3)
\]

As usual, subsolutions are defined analogously by reversing the inequalities.
Lemma 4.3 Let \( u_0 \in C(\overline{\Omega}) \), \( u_0 \geq 0 \), and let \( u \in C(\overline{\Omega} \times [0, T]) \) be a supersolution of (4.1) with \( g \geq 0 \). Then, \( u \geq 0 \).

**Proof** Assume to the contrary that \( u(x, t) \) is negative in some point. Let \( v(x, t) = u(x, t) + \gamma t \) with \( \gamma > 0 \) small such that \( v \) is still negative somewhere. Then, if \((x_0, t_0)\) is a point where \( v \) attains its negative minimum, there it holds that \( t_0 > 0 \) and

\[
v_t(x_0, t_0) = u_t(x_0, t_0) + \gamma > \int_G K_\varepsilon(x_0, y)(u(y, t_0) - u(x_0, t_0)) dy
= \int_G K_\varepsilon(x_0, y)(v(y, t_0) - v(x_0, t_0)) dy \geq 0.
\]

This contradicts that \((x_0, t_0)\) is a minimum of \( v \). Thus, \( u \geq 0 \). \( \square \)

Let \( f(x, t) \) a function in \( G \times (0, \infty) \), we consider next problem

\[
\begin{align*}
  u^c(x, t) &= \int_G K_\varepsilon(x, y)(u(y, t) - u(x, t)) dy + f(x, t) \quad \text{for } \ (x, t) \in \Omega \times [0, T], \\
  u^c(x, t) &= g(x, t) \quad \text{for } \ x \notin \Omega, \ t \in [0, T], \quad (4.4) \\
  u^c(x, 0) &= u_0(x) \quad \text{for } x \in \Omega,
\end{align*}
\]

**Corollary 4.4** Let \( K_\varepsilon \in L^\infty(G) \). Let \( u_0 \) and \( v_0 \) in \( L^1(\Omega) \) with \( u_0 \geq v_0 \) and let the functions \( g, h \in L^\infty((0, T); L^1(G \setminus \Omega)) \) with \( g \geq h \). Let \( u \) be a solution of (4.4) with \( u(x, 0) = u_0(x) \) and Dirichlet datum \( g \), and let \( v \) be a solution of (4.4) with \( v(x, 0) = v_0(x) \) and datum \( h \). Then, \( u \geq v \) a.e. \( \Omega \).

**Proof** Let \( w = u - v \). Then, \( w \) is a supersolution of (4.1) with initial datum \( u_0 - v_0 \geq 0 \) and datum \( g - h \geq 0 \). Using the continuity of the solutions with respect to the data and the fact that \( K_\varepsilon \in L^\infty(G) \), we may assume that \( u, v \in C(\Omega \times [0, T]) \). By Lemma 4.3 we obtain that \( w = u - v \geq 0 \). So the corollary is proved. \( \square \)

**Corollary 4.5** Let \( u, v \in C(\Omega \times [0, T]) \). If \( u \) is a supersolution of (4.4) and \( v \) is a subsolution of (4.4), then \( u \geq v \).

**Proof** It follows from the proof of the previous corollary.

## 5 Proof of the Main Theorems

The following Lemmas are the key for the proof of Theorems 1.1 and 1.2. To illustrate the technique we first prove a result which refers to the evolution problem stated in section 3.1.

**Lemma 5.1** Let \( \Omega \subset G \) be a bounded domain, and let \( v \in C^{2+\alpha}(G) \) for some \( 0 < \alpha < 1 \). Then there exist constants \( c \) and \( c' \) that depends only of \( v \), \( J \) and \( \Omega \) such that for all \( \varepsilon > 0 \)

\[
\left\| \mathcal{E}_\varepsilon(v) - \frac{c'}{2} \mathcal{J}(v) \right\|_{L^\infty(\Omega)} \leq c\varepsilon^\alpha.
\]
where \( J(v)(x) = \sum_{i=1}^{n_1} X_i^2 v(x) \) denotes minus the subLaplacian.

**Proof** Let us begin by writing the formula that defines \( \mathcal{E}_\varepsilon \) by means of the global chart given by the fixed basis of the stratified Lie algebra \( \mathfrak{g} \): for \( x \in \Omega \), since for the coordinates \( (t_1, \ldots, t_n) \in \mathbb{R}^n \) adapted to the basis, we can write

\[
\mathcal{E}_\varepsilon(v)(x) = \frac{1}{\varepsilon^2} \int_G \left[ v(x(\delta_\varepsilon(y^{-1}))) - v(x) \right] J(y) dy
\]

\[
= \frac{1}{\varepsilon^2} \int_{\mathbb{R}^n} \left( v(x \exp(-\varepsilon t_1 X_1 - \cdots - \varepsilon^n t_n X_n)) - v(x) \right) J(\exp(t_1 X_1 + \cdots + t_n X_n)) dt_1 \cdots dt_n.
\]

Thus, from the Taylor expansion (2.3) discussed in Sect. 2,

\[
v(x \exp(-\varepsilon t_1 X_1 - \cdots - \varepsilon^n t_n X_n)) - v(x) = -\varepsilon \sum_{i=1}^{n_1+n_2} t_i X_i v(x) + \frac{\varepsilon^2}{2} \sum_{i,j=1}^{n_1} t_i t_j X_i X_j v(x) + o(\left| \delta_\varepsilon (t_1 X_1 + \cdots + t_n X_n) \right|^2).
\]

Therefore,

\[
\mathcal{E}_\varepsilon v(x) = \frac{1}{\varepsilon^2} I + \frac{1}{\varepsilon^2} II + \frac{1}{\varepsilon^2} III
\]

used (2.4), we have

\[
\left| \frac{1}{\varepsilon^2} III \right| = \left| \int_{\mathbb{R}^n} \frac{o(\left| \delta_\varepsilon (t_1 X_1 + \cdots + t_n X_n) \right|^2)}{\varepsilon^2} J(\exp(t_1 X_1 + \cdots + t_n X_n)) dt_1 \cdots dt_n \right|
\]

\[
\leq c \int_{\mathbb{R}^n} e^{a \left| \exp(t_1 X_1 + \cdots + t_n X_n) \right|^2} \left| J(\exp(t_1 X_1 + \cdots + t_n X_n)) \right| dt_1 \cdots dt_n
\]

\[
\leq c e^{a}. \]

Where from properties (3.1), (3.2) and (3.3) we can compute
\[
I = \frac{1}{\epsilon^2} \sum_{j=1}^{n_1+n_2} e^{t_j}X_j(x) \int_{\mathbb{R}^n} J(\exp(t_1X_1 + \cdots + t_nX_n)) t_1 dt_1 \cdots dt_n = 0,
\]
\[
II = \frac{1}{\epsilon^2} \sum_{j=1}^{n_1} X_j(x) \int_{\mathbb{R}^n} J(\exp(t_1X_1 + \cdots + t_nX_n)) t_1 dt_1 \cdots dt_n
= c' \frac{1}{2} \sum_{j=1}^{n_1} X_j^2(x).
\]

Finally,
\[
\left\| E_\epsilon v(x) - \frac{c'}{2} Jv(x) \right\|_{L^\infty(\Omega)} \leq c\epsilon^a.
\]

\[\square\]

**Lemma 5.2** Let \( \Omega \subset G \) be a bounded domain, and let \( v \in C^{2+a}(G) \) for some \( 0 < a < 1 \). Then, there exists a constant \( c \) that depends only of \( v \), the matrix \( A \), the vector \( b \), \( J \) and \( \Omega \) such that for all \( \epsilon > 0 \)
\[
\left\| K_\epsilon(v) - K(v) \right\|_{L^\infty(\Omega)} \leq c\epsilon^a,
\]
where \( K \) is the operator defined in (3.11).

**Proof** By changing variables via \( z = \exp(L^{-1}(x) \exp^{-1}(\delta_\epsilon\cdot(y^{-1}x))) \), since thus we have that
\[
y = x \exp(\delta_\epsilon L(x) \exp^{-1}(z^{-1})) \text{ and } dy = \epsilon^d \det(L(x))dz
\]
for \( \epsilon > 0 \), the reescaled kernel operator becomes
\[
K_\epsilon v(x) = \frac{c(x) \det(L(x))}{\epsilon^2} \int_G a \left( \left( \exp \frac{M}{2} \delta_\epsilon W(x)(L'(x))^{-1} \exp^{-1} z \right)^{-1} \right) J(z)
\]
\[
(\nu(x \exp(\delta_\epsilon L(x) \exp^{-1}(z^{-1}))) - v(x)) dz,
\]
and by definition of the function \( a \) it finally assumes the form
\[
K_\epsilon v(x) = \frac{2}{\epsilon^2 C(J) M} \int_G \left( -\frac{M}{2} \sum_{j=1}^{n} e^{t_j} \bar{b}_j(x) \sum_{h=1}^{n} t_h \phi_h(\exp^{-1} z) + M \right) \]
\[
(\nu(x \exp(\delta_\epsilon L(x) \exp^{-1}(z^{-1}))) - v(x)) dz.
\]

Now let us write the formula in terms of the global chart as we did before (recall the proof of Lemma 5.1):
\[
K_\epsilon v(x) = \frac{2}{\epsilon^2 C(J) M} \int_{\mathbb{R}^n} \left( -\frac{M}{2} \sum_{j=1}^{n} e^{t_j} \bar{b}_j(x) \sum_{h=1}^{n} t_h \phi_h(\exp^{-1} z) + M \right) J \left( \exp \sum_{i=1}^{n} t_i X_i \right)
\]
\[
\left( \nu \left( x \exp \left( -\sum_{i=1}^{n} e^{t_i} \sum_{k=1}^{n} t_k (x) t_k X_i \right) \right) - v(x) \right) dt_1 \cdots dt_n,
\]
where \( t_h = \phi_h(\exp^{-1} z) \).
For the last factor we apply the Taylor expansion of homogeneous degree 2 (recall formula (2.3))

\[
v \left( x \exp \left( - \sum_{i=1}^{n} e^{t_i} \sum_{k=1}^{n} l_{ik}(x) t_k X_i \right) \right) - v(x)
= - \sum_{i=1}^{n_1+n_2} e^{t_i} \sum_{k=1}^{n_1} l_{ik}(x) t_k X_i v(x) + \frac{e^2}{2} \sum_{i=1}^{n_1} \left( \sum_{k=1}^{n_1} l_{ik}(x) t_k \right) \left( \sum_{h=1}^{n} l_{ih}(x) t_h \right) X_i X_j v(x)
+ o \left( \left| \sum_{i=1}^{n_1} e^{t_i} \sum_{k=1}^{n_1} l_{ik}(x) t_k \right|^2 \right).
\]

Then we can split as follows

\[
K_{\epsilon} v(x) = K_{\epsilon,1} v(x) + K_{\epsilon,2} v(x) + E_{\epsilon}.
\]

By (2.4)

\[
|E_{\epsilon}| \leq \frac{1}{C(J)} \int_{\mathbb{R}^n} \left| \left( - \frac{M}{2} \sum_{j=1}^{n} e^{\tilde{b}_j(x)} \sum_{h=1}^{n} l_{hj}(x) t_h + M \right) J \left( \exp \sum_{r=1}^{n} t_r X_r \right) \right| ^\alpha dt_1 \cdots dt_n
\]

\[
\leq \frac{1}{C(J)} \int_{\mathbb{R}^n} \left| \left( - \frac{M}{2} \sum_{j=1}^{n} e^{\tilde{b}_j(x)} \sum_{h=1}^{n} l_{hj}(x) t_h + M \right) J' \left( \exp \sum_{r=1}^{n} t_r X_r \right) \right| ^\alpha \exp \left( \sum_{i=1}^{n_1} \sum_{k=1}^{n_1} l_{ik}(x) t_k X_i \right) \left( 2 + \alpha \right) dt_1 \cdots dt_n
\]

\[
= c e^\alpha.
\]

Now, for \( K_{\epsilon,1} v(x) \) and \( K_{\epsilon,2} v(x) \), by extensive use of properties (3.1), (3.2) and (3.3), we have
Thus the proof ends. □

**Lemma 5.3** Let $\Omega \subset G$ be a bounded domain, and let $u \in C^{2+\alpha}(G)$ for some $0 < \alpha < 1$. Then, there exists a constant $c$ that depends only of $v$, the function $a$, $J$ and $\Omega$ such that for all $\epsilon > 0$

$$||L_\epsilon(v) - L(v)||_{L^\infty(\Omega)} \leq c \epsilon^\alpha,$$

where $L$ is the operator defined in (3.12).

**Proof** Let us rewrite the operators as follows:

$$L_\epsilon(v)(x) = \frac{2C(J)}{\epsilon^{Q+2}} \int_G J(\delta_{\epsilon^{-1}}y^{-1}x)a(y)[v(y) - v(x)]dy$$

$$+ \frac{2C(J)}{\epsilon^{Q+2}} \int_G J(\delta_{\epsilon^{-1}}y^{-1}x)[a(y) - a(x)]v(x)dy.$$

As usual, let us first change variables according to $z = \delta_{\epsilon^{-1}}y^{-1}x$, hence $y = \delta_{\epsilon}xz^{-1}$ and $dz = -\epsilon^Q dy$ and then write it in coordinates:
\[
\mathcal{L}_c(v)(x) = \frac{2C(J)}{e^2} \int_G J(z) a(\delta_c x z^{-1}) [v(\delta_c x z^{-1}) - v(x)] dz
\]

\[
= \frac{2C(J)}{e^2} \int_G J(z) [a(\delta_c x z^{-1}) - a(x)] v(x) dz
\]

\[
= \frac{2C(J)}{e^2} \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^n t_r x_r \right) a \left( x \exp \left( - \sum_{k=1}^n e^{\lambda_k t_k} x_k \right) \right)
\]

\[
\left[ v \left( x \exp \left( - \sum_{i=1}^n e^{\lambda_i t_i} x_i \right) \right) - v(x) \right] dt_1 \ldots dt_n
\]

\[
+ \frac{2C(J)}{e^2} \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^n t_r x_r \right) \left[ a \left( x \exp \left( - \sum_{i=1}^n e^{\lambda_i t_i} x_i \right) \right) - a(x) \right]
\]

\[
\times v(x) dt_1 \ldots dt_n
\]

\[= I_1 + I_2,
\]

where \( t_k = \phi_k(\exp^{-1} z) \).

The next step is to apply Taylor decomposition of homogeneous degree 2 (recall formula (2.3)) to \( v \) in \( I \) and to \( a \) in \( I_2 \):

\[
v \left( x \exp \left( - \sum_{i=1}^n e^{\lambda_i t_i} x_i \right) \right) - v(x) = - \sum_{i=1}^{n_1+n_2} e^{\lambda_i t_i} x_i v(x)
\]

\[
+ \frac{e^2}{2} \sum_{i=1}^{n_1} t_i t_j x_i x_j v(x) + o \left( \delta_c \left( \sum_{i=1}^n t_i X_i \right) \right)^2,
\]

hence, by (2.4)

\[
I_1 = \frac{2C(J)}{e^2} \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^n t_r x_r \right) a \left( x \exp \left( - \sum_{k=1}^n e^{\lambda_k t_k} x_k \right) \right)
\]

\[
\times \left( - \sum_{i=1}^{n_1+n_2} e^{\lambda_i t_i} x_i v(x) \right) dt_1 \ldots dt_n
\]

\[
+ C(J) \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^n t_r x_r \right) a \left( x \exp \left( - \sum_{k=1}^n e^{\lambda_k t_k} x_k \right) \right)
\]

\[
\times \left( \sum_{i=1}^{n_1} t_i t_j x_i x_j v(x) \right) dt_1 \ldots dt_n
\]

\[
+ e^a c \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^n t_r x_r \right) a \left( x \exp \left( - \sum_{k=1}^n e^{\lambda_k t_k} x_k \right) \right) dt_1 \ldots dt_n
\]

\[= I_1 + I_2 + e^a c,
\]
and by applying Taylor formula again to \(a\), but this time of homogeneous degree 1, and extensive use of formulas (3.1), (3.2) and (3.3) it follows that

\[
I_1 = \frac{2C(J)}{e^2} \int_{\mathbb{R}^n} J\left(\exp\sum_{r=1}^{n} t_r X_r\right) \left(a(x) - \sum_{k=1}^{n_1} e t_k X_k(a)(x) + o\left(\delta \left(\sum_{r=1}^{n} t_r X_r\right)\right)\right)
\times \left(-\sum_{i=1}^{n_1+n_2} e^{\lambda_i} t_i X_i(v)(x)\right) dt_1 \ldots dt_n
\]

\[
I_1 = \frac{2C(J)}{e^2} \sum_{i=1}^{n_1} \sum_{k=1}^{n_1} e^{\lambda_i} t_i X_i(a)(x) X_i(v)(x) \int_{\mathbb{R}^n} J\left(\exp\sum_{r=1}^{n} t_r X_r\right) t_k t_1 dt_1 \ldots dt_n

+ \frac{2C(J)}{e^2} \int_{\mathbb{R}^n} J\left(\exp\sum_{r=1}^{n} t_r X_r\right) o\left(\delta \left(\sum_{r=1}^{n} t_r X_r\right)\right) \left(-\sum_{i=1}^{n_1} e^{\lambda_i} t_i X_i(v)(x)\right) dt_1 \ldots dt_n
\]

\[
= 2 \sum_{i=1}^{n_1} X_i(a)(x) X_i(v)(x),
\]

\[
+ \frac{C(J)}{e} \int_{\mathbb{R}^n} J\left(\exp\sum_{r=1}^{n} t_r X_r\right) o\left(\delta \left(\sum_{r=1}^{n} t_r X_r\right)\right) \left(-\sum_{i=1}^{n_1} e^{\lambda_i} t_i X_i(v)(x)\right) dt_1 \ldots dt_n
\]

Now, by (2.2) and Theorem 2.3, we get

\[
\int_{\mathbb{R}^n} J\left(\exp\sum_{r=1}^{n} t_r X_r\right) o\left(\delta \left(\sum_{r=1}^{n} t_r X_r\right)\right) \left(-\sum_{i=1}^{n_1} e^{\lambda_i} t_i X_i(v)(x)\right) dt_1 \ldots dt_n
\]

\[
\leq ec \int_{\mathbb{R}^n} J\left(\exp\sum_{r=1}^{n} t_r X_r\right) \left(-\sum_{i=1}^{n_1} e^{\lambda_i} t_i X_i(v)(x)\right) dt_1 \ldots dt_n
\]

\[
\leq ec.
\]

For \(I_2\) we have
Finally, by applying Taylor decomposition of homogeneous degree 2 again to $a$,

$$I_2 = C(J) \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^{n} t_r X_r \right) \left( a(x) - \sum_{k=1}^{n_1} e t_k X_k(a)(x) + o \left( \delta \left( \sum_{r=1}^{n} t_r X_r \right) \right) \right) \times \left( \sum_{i,j=1}^{n_1} t_i t_j X_i X_j v(x) \right) dt_1 \ldots dt_n$$

$$= C(J) \sum_{i,j=1}^{n_1} a(x) X_i X_j(v)(x) \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^{n} t_r X_r \right) t_i t_j dt_1 \ldots dt_n$$

$$+ C(J) \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^{n} t_r X_r \right) o \left( \delta \left( \sum_{r=1}^{n} t_r X_r \right) \right) \left( \sum_{i,j=1}^{n_1} t_i t_j X_i X_j v(x) \right) dt_1 \ldots dt_n$$

$$= \sum_{i=1}^{n_1} a(x) X_i X_i(v)(x)$$

$$+ C(J) \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^{n} t_r X_r \right) o \left( \delta \left( \sum_{r=1}^{n} t_r X_r \right) \right) \left( \sum_{i,j=1}^{n_1} t_i t_j X_i X_j v(x) \right) dt_1 \ldots dt_n.$$

Now, by (2.2) and Theorem 2.3, we get

$$\left| \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^{n} t_r X_r \right) o \left( \delta \left( \sum_{r=1}^{n} t_r X_r \right) \right) \left( \sum_{i,j=1}^{n_1} t_i t_j X_i X_j v(x) \right) dt_1 \ldots dt_n \right| \leq \epsilon C(J, v, a).$$

Finally, by applying Taylor decomposition of homogeneous degree 2 again to $a$,

$$II = \frac{2C(J)}{\epsilon^2} \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^{n} t_r X_r \right)$$

$$\times \left[ - \sum_{i=1}^{n_1} e^2 t_i X_i(a)(x) + \frac{\epsilon^2}{2} \sum_{i,j=1}^{n_1} t_i t_j X_i X_j(a)(x) + o \left( \delta \left( \sum_{r=1}^{n} t_r X_r \right) \right) \right]$$

$$\times v(x) dt_1 \ldots dt_n$$

$$= C(J) \sum_{j=1}^{n_1} X_j X_j(a)(x) v(x) \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^{n} t_r X_r \right) t_i t_j dt_1 \ldots dt_n$$

$$+ \frac{2C(J)}{\epsilon^2} \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^{n} t_r X_r \right) o \left( \delta \left( \sum_{r=1}^{n} t_r X_r \right) \right) v(x) dt_1 \ldots dt_n$$

$$= \sum_{i=1}^{n_1} X_i X_i(a)(x) v(x)$$

$$+ \frac{2C(J)}{\epsilon^2} v(x) \int_{\mathbb{R}^n} J \left( \exp \sum_{r=1}^{n} t_r X_r \right) o \left( \delta \left( \sum_{r=1}^{n} t_r X_r \right) \right) dt_1 \ldots dt_n,$$

by (2.4)
Next we turn to the proof of Theorem 1.1. The proof of Theorem 1.2 follows the same lines.

**Proof of Theorem 1.1** Let $v(\cdot, t) \in C^{2+\alpha}(\Omega)$ be a solution of problem (1.2), and define an extension $\tilde{v}$ of $v$ to the space $C^{2+\alpha,1+\alpha}(G \times [0, T])$ such that

\[
\begin{align*}
\tilde{v}_j(x, t) &= K(\tilde{v})(x, t), & x \in \Omega, & t > 0, \\
\tilde{v}(x, t) &= \tilde{g}(x, t), & x \not\in \Omega, & t > 0, \\
\tilde{v}(x, 0) &= u_0(x), & x \in \Omega,
\end{align*}
\]

(5.1)

where $\tilde{g}$ is a smooth function which satisfies $\tilde{g}(x, t) = g(x, t)$ if $x \in \partial \Omega$ and $\tilde{g}(x, t) = g(x, t) + O(\epsilon)$ if $x \approx \partial \Omega$, (in the sense $\lim_{\epsilon \to 0} O(\epsilon) = 0$). Let us define now the difference $w^\epsilon(x, t) = \tilde{v}(x, t) - u^\epsilon(x, t)$. Thus defined, $w^\epsilon$ satisfies

\[
\begin{align*}
w^\epsilon_j(x, t) &= K(\tilde{v})(x, t) - K_{\epsilon}\tilde{v}(x, t) + K_{\epsilon}w^\epsilon(x, t), & x \in \Omega, & t > 0, \\
w^\epsilon(x, t) &= g(x, t) - \tilde{g}(x, t), & x \not\in \Omega, & t > 0, \\
w^\epsilon(x, 0) &= 0, & x \in \Omega.
\end{align*}
\]

(5.2)

From Lemma 5.2 exists a constant $K_1$ dependent only of $\tilde{v}$ and the differential operator $K$ such that for all $\epsilon > 0$

\[|K\tilde{v}(x, t) - K_{\epsilon}\tilde{v}(x, t)| \leq K_1 \epsilon^\alpha.\]

Let $\overline{w}(x, t) = K_1 \epsilon^\alpha t + K_2 \epsilon$, where $K_2 > 0$ is a constant independent of $\epsilon$ to be chosen later. Now we see that $\overline{w}(x, t)$ is a supersolution of the problem (5.2). Since $\overline{w}(x, t)$ does not depend on $x$, we have

\[K_{\epsilon}\overline{w}(x, t) = \int_G K_{\epsilon}(x, y)(\overline{w}(y, t) - \overline{w}(x, t)) dy = 0,\]

and follows that

\[\overline{w}_j(x, t) = K_1 \epsilon^\alpha \geq K\tilde{v}(x, t) - K_{\epsilon}\tilde{v}(x, t) + K_{\epsilon}\overline{w}(x, t).\]

Also, $\overline{w}(x, 0) > 0$ and by the definition of $\tilde{g}$, we can choose $K_2 > 0$ such that

\[\overline{w}(x, t) \geq K_2 \epsilon \geq O(\epsilon),\]

for $x \in \Omega^c, x \approx \partial \Omega, t > 0$. Hence $\overline{w}$ is indeed a supersolution of (5.2).

From the comparison principle (Corollary 4.5) we get that $\overline{v} - u^\epsilon \leq \overline{w}(x, t) = K_1 \epsilon^\alpha t + K_2 \epsilon$. Applying the same arguments for $w(x, t) = -\overline{w}(x, t)$ we obtain that $w(x, t)$ is a subsolution of problem (5.2) and again by the comparison principle,
\[-K_1 \epsilon^t - K_2 \epsilon \leq \bar{v} - u^\epsilon \leq K_1 \epsilon^t + K_2 \epsilon.\]

Therefore,
\[
\|\bar{v} - u^\epsilon\|_{L^\infty(\Omega \times [0,T])} \leq K_1 \epsilon^T + K_2 \epsilon \to 0.
\]

\[\square\]

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

1. Baldi, A., Citti, G., Cupini, G.: Schauder estimates at the boundary for sub-laplacians in Carnot groups. Calc. Var. 58, 204 (2019). https://doi.org/10.1007/s00526-019-1628-7
2. Banerjee, A., Garofalo, N., Munive, I.H.: Compactness methods for $C^{1,\alpha}$-boundary Schauder estimates in Carnot groups. Calc. Var. 58, 97 (2019). https://doi.org/10.1007/s00526-019-1531-2
3. Bodnar, M., Velazquez, J.J.L.: An integro-differential equation arising as a limit of individual cell-based models. J. Differ. Equ. 222, 341–380 (2006)
4. Carrillo, C., Fife, P.: Spatial effects in discrete generation population models. J. Math. Biol. 50(2), 161–188 (2005)
5. Chasseigne, E., Chaves, M., Rossi, J.D.: Asymptotic behavior for nonlocal diffusion equations. J. de mathématiques pures et appliquées 86(3), 271–291 (2006)
6. Cortazar, C., Elgueta, M., Rossi, J.D.: Nonlocal diffusion problems that approximate the heat equation with Dirichlet boundary conditions. Israel J. Math. 170(1), 53–60 (2009)
7. Dyer, J.L.: A nilpotent Lie algebra with nilpotent automorphism group. Proc. Symp. Pure Math. 4, 33–49 (1961)
8. Fife, P.: Some Nonclassical Trends in Parabolic and Parabolic-like Evolutions. Trends in Nonlinear Analysis, Springer, Berlin, Heidelberg (2003)
9. Folland, G.B., Stein, E.M.: Hardy spaces on homogeneous groups, Princeton University Press, (1982)
10. Fournier, N., Laurençot, P.: Well-posedness of Smoluchowski’s coagulation equation for a class of homogeneous kernels. J. Funct. Anal. 233, 351–379 (2006)
11. Hörmander, L.: Hypoelliptic second order differential equations. Acta Math. 119, 147–171 (1967)
12. Kindermann, S., Osher, S., Jones, P.W.: Deblurring and denoising of images by nonlocal functionals. Multiscale Model. Simul. 4, 1091–1115 (2005)
13. Mogilner, A., Edelstein-Keshet, L.: A non-local model for a swarm. J. Math. Biol. 38, 534–570 (1999)
14. Molino, A., Rossi, J.D.: Nonlocal diffusion problems that approximate a parabolic equation with spatial dependence. Z. Angew. Math. Phys. 67(3), 1–4 (2016)
15. Sun, J.W., Li, W.T., Yang, F.I.: Approximate the Fokker-Planck equation by a class of nonlocal dispersive problems. Nonlinear Anal. Theory Methods Appl. 74, 3501–3509 (2011)
16. Varadarajan, V.S.: Lie Groups, Lie Algebras, and Their Representations. Springer-Verlag, New York (1984)
17. Vidal, R.E.: Nonlocal heat equations in Heisenberg group. Nonlinear Differ. Equ. Appl. 24(5), 1–21 (2017)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.