Università degli Studi di Ferrara

DOTTORATO DI RICERCA IN
Matematica

CICLO XXXVIII

COORDINATORE

Prof. **Lorenzo Brasco**

Hypoelliptic Partial Differential Equations: regularity theory and applications to degenerate McKean-Vlasov processes

---

Settore Scientifico Disciplinare MATH-03/A

Supervisore
Prof. **Sergio Polidoro**

Co-Supervisore
Prof.ssa **Cecilia Vernia**

Dottorando
Dott. **Giulio Pecorella**

Anni Accademici 2022/2023 - 2024/2025

Abstract

This dissertation is devoted to the study of some degenerate Partial Differential Equations, characterized by an underlying hypoelliptic regularizing property compensating for the absence of uniform ellipticity.

The first part is concerned with regularity theory for a class of subelliptic operators on Carnot groups. In Chapter 3 we investigate some regularity properties of stationary operators with Hölder continuous coefficients. We first adapt Levi's parametrix method to establish the existence of a local fundamental solution. Building on this, we provide novel results on the strong maximum principle and the invariant Harnack inequality, employing the mean value formula technique originally developed for harmonic functions. In Chapter 4 we establish analogous results for the evolutive counterpart. In this setting, to overcome certain drawbacks of the mean value formula, we employ Kuptsov method to obtain an improved kernel, and we prove its boundedness by using Rothschild and Stein lifting technique. The results presented in these chapters heavily rely on the properties of the sub-Riemannian geometry relevant to these subelliptic operators, and in particular on the Carnot-Carathéodory distance they induce.

In the second part we analyse some McKean-Vlasov stochastic processes by studying the associated density, which satisfies a nonlinear, degenerate Kolmogorov equation. In Chapter 6 we study the mean field Kuramoto model with inertia and noise, addressing the corresponding Cauchy problems. We establish existence, uniqueness and some precise bounds, and we propose a finite difference scheme to numerically compute the density. In Chapter 7 we study a McKean-Vlasov optimal control problem. We first show that the mean field problem provides a suitable approximation of the finite size control problem, by proving the Γ-convergence of the corresponding cost functionals. We then employ the Lagrangian formalism to derive first order optimality conditions for the optimizers. Finally, these optimality conditions are implemented in a gradient descent algorithm based on the finite difference scheme introduced in Chapter 6.

Contents

Chapter 1

Introduction

In this thesis we investigate several problems related to a class of degenerate PDEs. These equations lack uniform ellipticity, a difficulty we address by exploiting their underlying sub-Riemannian hypoelliptic structure.
The first part of this dissertation is devoted to the study of regularity theory for Hörmander operators on Carnot groups: we prove the existence of their fundamental solution and derive several useful properties; these results, combined with some recent mean value formulas, are then employed to provide new proofs of maximum propagation results and Harnack inequalities for subelliptic operators.
In the second part we focus on applying the theory of degenerate Kolmogorov equations to some McKean-Vlasov stochastic processes: firstly, we establish existence, uniqueness, and estimates for the density of a particular process (a Kuramoto model), developing also a stable numerical scheme for its computation; subsequently, we provide theoretical and computational tools to handle a class of McKean-Vlasov optimal control problems.

1.1 Hypoelliptic operators on Lie groups

The progenitor of the equations studied in this thesis was introduced by Kolmogorov in the context of kinetic theory. In his seminal work [70], Kolmogorov analysed the following stochastic process: let $X_t \in \mathbb{R}$ and $V_t \in \mathbb{R}$ denote the position and the velocity of a particle at time $t \in [t_0, T]$, and suppose that the particle evolves according to the following Langevin dynamics

$$\begin{cases} X_t = x_{t_0} + \int_{t_0}^{t} V_s \, ds, \\ V_t = v_{t_0} + \sqrt{2}\, W_t, \end{cases}$$

where $(W_t)_{t\in[t_0,T]}$ is a standard Wiener process. The density f of the process satisfies the following *degenerate second order equation*

$$\mathcal{K}f(v,x,t) = \partial^2_{vv} f(v,x,t) - v\partial_x f(v,x,t) - \partial_t f(v,x,t) = 0. \tag{1.1.1}$$

This equation lacks diffusion in the position variable x, that is the *degenerate variable*. In his analysis, Kolmogorov derived the explicit expression of the density

$$\begin{aligned} &f\Big((v,x,t),(v_0,x_0,t_0)\Big)\\ &=\frac{\sqrt{3}}{2\pi(t-t_0)^2}\exp\left(-\frac{(v-v_0)^2}{t-t_0} - 3\frac{(v-v_0)(x-x_0-(t-t_0)v_0)}{(t-t_0)^2} - 3\frac{(x-x_0-(t-t_0)v_0)^2}{(t-t_0)^3}\right). \end{aligned}$$

Three notable features emerge upon examining this expression:

1. f is left-invariant w.r.t. the following *non abelian group law*

$$(v_0,x_0,t_0)^{-1}\circ(v,x,t) = \Big(v-v_0, x-x_0-(t-t_0)v_0, t-t_0\Big);$$

2. f is homogeneous of degree -4 w.r.t. the following *dilations law*

$$\delta_r(v,x,t) = \Big(rv, r^3x, r^2t\Big), \qquad r>0;$$

3. f is smooth in the set $\Big\{(v,x,t,v_0,x_0,t_0)\in\mathbb{R}^6, t>t_0\Big\}$.

Properties *(1.)-(2.)* highlights the underlying geometric structure of (1.1.1): indeed, $\mathcal{K}$ is left-invariant w.r.t. the composition law $\circ$, and homogeneous of degree 2 w.r.t. the family of dilations $(\delta_r)_{r>0}$, that means

$$\begin{aligned} (\mathcal{K}f)\Big((v_0,x_0,t_0)^{-1}\circ(v,x,t)\Big) &= \mathcal{K}\Big(f\Big((v_0,x_0,t_0)^{-1}\circ(v,x,t)\Big)\Big),\\ \mathcal{K}\Big(f(\delta_r(v,x,t)\Big) &= r^2(\mathcal{K}f)\Big(\delta_r(v,x,t)\Big). \end{aligned}$$

Equipped with the group law $\circ$ and the dilations law $(\delta_r)_{r>0}$, $\mathbb{R}^3$ is endowed with the structure of a *non abelian homogeneous Lie group*.

Definition (Homogeneous Lie group). *A homogeneous Lie group on $\mathbb{R}^N$ is a group $\mathbb{G} = \Big(\mathbb{R}^N,\circ,\delta_r\Big)$ with a smooth composition law $\circ$, and a dilations law $\{\delta_r\}_{r>0}$ that is an automorphism of the group,* i.e.

$$\delta_r(x\circ y) = \Big(\delta_r(x)\Big)\circ\Big(\delta_r(y)\Big), \qquad \text{for every } x,y\in\mathbb{R}^N,\ r>0.$$

The last property *(3.)* is perhaps the most striking upon initial inspection: although the operator being degenerate, the associated density turns out to be smooth. In other words, equation (1.1.1) exhibits the same regularizing effect of uniformly elliptic equations with smooth coefficients. This observation leads us to the following definition.

Definition (Hypoelliptic operators). *We say that a differential operator $\mathcal{L}$ is* hypoelliptic *if, for any $\varphi \in C^\infty$, every distributional solution to $\mathcal{L}u = \varphi$ is smooth.*

In his celebrated paper [63], Hörmander exploited Kolmogorov findings providing a sufficient condition for hypoellipticity. Hörmander focused on a broader class of differential operators, nowadays referred to as *Hörmander type I* and *Hörmander type II* operators

$$\mathcal{L}u(x) = \sum_{i=1}^{m_1} X_i^2 u(x), \qquad \mathcal{H}u(x,t) = \sum_{i=1}^{m_1} X_i^2 u(x,t) + Yu(x,t), \tag{1.1.2}$$

with $(x,t) \in \mathbb{R}^N \times \mathbb{R}$. Here, $\{X_i\}_{i=1}^{m_1}$ and Y are smooth vector fields on $\mathbb{R}^N$, namely differential operators of the form

$$X_i(x) = \sum_{j=1}^{N} \varphi_j^i(x)\partial_{x_j}, \qquad Y(x) = \sum_{j=1}^{N} \varphi_j^0(x)\partial_{x_j} + \varphi^0(x)\partial_t,$$

with $\varphi_j^i \in C^\infty(\mathbb{R}^N)$, and are assumed to be linearly independent. Hörmander's main result is the following sufficient condition for $\mathcal{L}$ and $\mathcal{H}$ to be hypoelliptic.

Theorem (Hörmander's rank condition). *Suppose that the Lie algebra generated by the vector fields $X_1, \dots, X_{m_1}$ spans $\mathbb{R}^N$, namely*

$$\operatorname{rank}\Big(\operatorname{Lie}\big(X_1, \dots, X_{m_1}\big)\Big)(x) = N, \qquad \forall x \in \mathbb{R}^N,$$

then $\mathcal{L}$ is hypoelliptic. Analogously, if the Lie algebra generated by $X_1, \dots, X_{m_1}, Y$ spans $\mathbb{R}^{N+1}$, then $\mathcal{H}$ is hypoelliptic.

If $m_1 = N$, then $\mathcal{L}$ is an uniformly elliptic operator, and the corresponding evolution operator $\mathcal{H}$ is uniformly parabolic. Conversely, when $m_1 < N$ both $\mathcal{L}$ and $\mathcal{H}$ are degenerate, in the sense that the associated characteristic form fails to be positive definite, and it is instead only positive semidefinite. The hypoelliptic regularizing property cannot, in general, be expected to hold for arbitrary degenerate operators: consider for instance the operator $\mathscr{L} = \partial_{xx}^2$ on $\mathbb{R}^2$; the function $u(x,y) = \operatorname{sign}(y)$ is not even continuous, but still satisfies pointwise $\mathscr{L}u = 0$. Chow–Rashevsky connectivity theorem provides a geometric interpretation of the rank condition. This result asserts that, under Hörmander's condition, any two points in the domain can be joined by a concatenation of integral curves of the vector fields. This unveils the intrinsic sub-Riemannian structure underlying these degenerate operators, providing a geometric interpretation of their regularizing property. Roughly speaking, $\mathcal{L}$ and $\mathcal{H}$ are degenerate since they lack some second order derivatives. Nevertheless, Hörmander's condition ensures that the missing derivatives are recovered through iterated Lie brackets.

In this dissertation we study differential operators with variable coefficients which are neither invariant under a Lie group structure nor hypoelliptic. However, we treat them as perturbations of simpler model operators (specifically, sublaplacians on Carnot groups and hypoelliptic Kolmogorov operators) that do possess these properties. This approach falls within the classical perturbative theory of differential operators (we quote for instance the popular Levi's parametrix method, that will be employ in Chapter 3 and Chapter 4), where the analytical structure of the model operator is used to approximate the behavior of more general operators.

1.2 Regularity theory for subelliptic operators on stratified groups

In the first part of this thesis we study some properties of a class of subelliptic operators. Let $\{X_i\}_{i=1}^{m_1}$ be a set of smooth vector fields on $\mathbb{R}^N$ satisfying Hörmander's rank condition, and generating a homogeneous Lie group $\mathbb{G}$ with stratified algebra (namely, a Carnot group $\mathbb{G}$). We focus on the analysis of classical solutions to the following class of differential operators

$$\mathscr{L}_{\mathbb{G}} u = \sum_{i,j=1}^{m_1} a_{ij} X_i X_j u + \sum_{i=1}^{m_1} b_i X_i u + cu, \qquad \mathscr{H}_{\mathbb{G}} u = \mathscr{L}_{\mathbb{G}} u - \partial_t u,$$

with particular emphasis on the regularity theory of divergence form operators

$$\mathscr{L}_{\mathbb{G}}^{\text{div}} u = \text{div}_{\mathbb{G}}(A \nabla_{\mathbb{G}} u) + \langle b, \nabla_{\mathbb{G}} u \rangle + cu, \qquad \mathscr{H}_{\mathbb{G}}^{\text{div}} u = \mathscr{L}^{\text{div}}{}_{\mathbb{G}} u - \partial_t u.$$

Here, $A = (a_{ij})$ is an $m_1 \times m_1$ symmetric, uniformly elliptic matrix, while $\nabla_{\mathbb{G}}$ and $\text{div}_{\mathbb{G}}$ denote the horizontal gradient and the divergence operator on $\mathbb{G}$. The regularity assumptions on the coefficients are their Hölder continuity in the $\mathbb{G}$ metric and their boundedness. These operators can be seen as the Hölder coefficients extension of the sublaplacian and the heat operator on $\mathbb{G}$. The study of these operators is not of purely theoretical interest, since these PDEs arise in various applications: geometric theory of several complex variables, curvature problems for CR-manifolds, human vision (see for instance [41, 42, 65, 87]). The regularity theory for subelliptic operators on stratified groups received a major impetus after the foundational work of Hörmander [63]. Among the most influential contributors we quote Folland [54], Rothschild and Stein [106], and Lanconelli [25–27, 34] who played a pivotal role in the development of this field, by extending potential theory and singular integral techniques to the setting of nilpotent Lie groups.

In the first part of this thesis we show the existence of the *fundamental solutions* of $\mathscr{L}_{\mathbb{G}}$ and $\mathscr{H}_{\mathbb{G}}$. To this end, we employ the well known Levi's parametrix method [82], allowing us to establish also some important properties of the fundamental solution (for instance *Gaussian estimates*). This is then employed to obtain some of the main novelties of this dissertation: a *strong maximum principle* and an *invariant Harnack inequality* for $\mathscr{L}_{\mathbb{G}}^{\text{div}}$ and $\mathscr{H}_{\mathbb{G}}^{\text{div}}$. The proofs are obtained by adapting an elegant and elementary argument originally developed for harmonic functions, based on Gauss mean value formula: let u be harmonic in $\Omega \subset \mathbb{R}^N$, then it satisfies

$$u(\xi) = \frac{|B_1|^{-1}}{r^N} \int_{B_r(\xi)} u(x)dx, \tag{1.2.1}$$

for every $B_r(\xi) \subset \Omega$. The strong maximum principle and the Harnack inequality for u follow as a simple consequence of (1.2.1).

1. Suppose u harmonic attains its maximum at $\xi \in \Omega$, then u is constant on every connected component of Ω containing ξ. Indeed, let $r > 0$ such that $B_r(\xi) \subset \Omega$, by (1.2.1) we get

$$0 = u(\xi) - \frac{|B_1|^{-1}}{r^N} \int_{B_r(\xi)} u(x)dx = \frac{|B_1|^{-1}}{r^N} \int_{B_r(\xi)} \Big(u(\xi) - u(x)\Big)dx. \tag{1.2.2}$$

Since ξ is a maximum point, from (1.2.2) we infer $u(x) = u(\xi)$ for every $x \in B_r(\xi)$, and the maximum propagation follows by a covering argument.

2. Let $u \geq 0$ be harmonic, then for every $\xi \in \Omega$ and $r > 0$ such that $B_{4r}(\xi) \subseteq \Omega$, it holds

$$\sup_{B_r(\xi)} u \leq 3^N \inf_{B_r(\xi)} u.$$

Indeed, for every $y, w \in B_r(\xi)$, equation (1.2.1) yields

$$u(y) = \frac{|B_1|^{-1}}{r^N} \int_{B_r(y)} u(x)dx \leq 3^N \frac{|B_1|^{-1}}{(3r)^N} \int_{B_{3r}(w)} u(x)dx = 3^N u(w),$$

concluding by the arbitrariness of y and w.

In Chapter 3 we adapt these arguments to $\mathscr{L}_{\mathbb{G}}^{\text{div}}$. Mean value formulas for $\mathscr{L}_{\mathbb{G}}^{\text{div}}$ have been recently proved by Pallara and Polidoro in [91], where the authors employed the tools of Geometric Measure Theory on Carnot groups to derive a mean value formula analogous to (1.2.1). The main difference lies in the appearance of a non constant kernel $M_{\mathbb{G}}$, depending on the horizontal gradient of the fundamental solution Γ^* of the adjoint equation. The proof

of the maximum principle requires $M_{\mathbb{G}} > 0$ (up to a set of zero Lebesgue measure), while the proof of the Harnack inequality needs the existence of two positive constants m^- and m^+ such that $m^- < M_{\mathbb{G}} < m^+$. In the constant coefficients case, the horizontal gradient vanishes at most on a negligible set (thanks to Sard's lemma, since hypoellipticity implies $\Gamma^* \in C^\infty$). Combined with the homogeneity of Γ^*, this allows to adapt the arguments employed in *(1.)* and *(2.)* (see for instance [24,40]). However, in the Hölder coefficients setting the fundamental solution Γ^* is generally not even C^1. As a consequence, Sard's lemma does not hold and we cannot infer anything regarding the set $\{M_{\mathbb{G}} = 0\}$. Furthermore, in the Hölder coefficients setting homogeneity may fail to hold. To overcome these two difficulties we develop an alternative strategy, relying on the sub-Riemannian framework of $\mathscr{L}_{\mathbb{G}}^{\mathrm{div}}$ and strictly tied on the connection between Hörmander's rank condition and Chow–Rashevsky connectivity theorem: we show that the kernel $M_{\mathbb{G}}$ remains strictly positive along the integral curves of the vector fields $\{X_i\}_{i=1}^{m_1}$, which allows us to prove a maximum principle propagation and a Harnack inequality along these trajectories; then, by using Hörmander condition (hence Chow–Rashevsky connectivity theorem) we propagate these results to the whole domain.
The results discussed above are contained in the paper [98].

In Chapter 4 we prove similar results for $\mathscr{H}_{\mathbb{G}}^{\mathrm{div}}$, by employing the mean value formula developed by Pallara and Polidoro in [92]. The extension of the arguments used in *(1.)* and *(2.)* to the setting of evolution operators presents similar difficulties, but requires different tools. In this context, an additional challenge arises due to the unboundedness of the kernel (an issue that is already present in the case of the heat operator on $\mathbb{R}^2$). To overcome this unboundedness, Kuptsov introduced in [73] a technique, based on Hadamard's descent method, to construct an improved kernel with the remarkable feature to be bounded. This technique gained significant popularity in recent years, and was later employed by Garofalo and Lanconelli in [59] for uniformly parabolic operators, and by Polidoro [102] and Rebucci [105] for Kolmogorov operators. This improved kernel is strictly positive, hence the maximum propagation follows by adapting the argument in *(1.)*. Proving a Harnack inequality as in *(2.)*, however, is more delicate, since boundedness is not straightforward in our setting. To address this issue, we rely on Polidoro's method [102] to derive the required local estimates to prove the boundedness. The most delicate aspect is obtaining sharp estimates for the fundamental solution of constant coefficients operators. To this end, we *lift* $\mathbb{G}$ to a *free Carnot group* $\widetilde{\mathbb{G}}$ by means of the Rothschild–Stein lifting technique [106], following the construction developed by Bonfiglioli, Lanconelli, and Uguzzoni in [25]. Within this framework, we exploit the equivalence of sublaplacians on free Carnot groups established by Bonfiglioli and Uguzzoni [28], which provides the required estimates. Proved the boundedness of the

improved kernel, we then establish a *parabolic* Harnack inequality in the lifted group, which in turn yields the corresponding inequality for $\mathscr{H}_{\mathbb{G}}^{\mathrm{div}}$. The parabolic Harnack inequality obtained in this thesis is analogous to the one established by Bramanti, Bonfiglioli, Uguzzoni and Lanconelli in [34] with the Krylov-Safonov method, although derived within a completely different setting. The main feature of our technique is that it allows us to establish an *invariant* Harnack inequality defined in terms of future and past cylinders, similarly to the classical parabolic case (see for instance [85]). This invariant Harnack inequality follows from the optimal control technique introduced by Boscain and Polidoro in [31], which is particularly well adapted to the vector field structure of $\mathscr{H}_{\mathbb{G}}^{\mathrm{div}}$.
These results are presented in the preprint [100] by the author and Rebucci.

1.3 PDEs and degenerate McKean–Vlasov stochastic processes

The second part of this thesis is concerned with the study of the following degenerate, nonlinear, Stochastic Differential Equations

$$\begin{cases} dV_t = \Big(b(V_t, X_t) + \int_{\mathbb{R}^2} K_t\big(v - V_t, x - X_t\big) dF_t(v, x)\Big)\, dt + \sqrt{2}\, dW_t, \\ dX_t = V_t\, dt, \end{cases} \tag{1.3.1}$$

usually referred to as degenerate McKean-Vlasov equations. Here, $X_t \in \mathbb{R}$ and $V_t \in \mathbb{R}$ denote the position and the velocity of a particle at time $t \in [0, T]$, $(W_t)_{t\in[0,T]}$ is a standard Wiener process, b is a drift term, and F denotes the law of the process. The study of such stochastic processes is primarily motivated by their wide range of applications in biology, engineering, social sciences, and physics (see for instance [37, 88, 111, 117]). We investigate these processes by studying the relative density: under some suitable assumptions on the drift function b and the kernel K, Itô formula implies that the density f of the law F solves the following *degenerate nonlinear Kolmogorov equation*

$$\begin{aligned} \partial^2_{vv} f(v, x, t) - \partial_v \Big(\Big(b(v, x) + \int_{\mathbb{R}^2} K(\eta - v, \xi - x, t) f(\eta, \xi, t) d\eta d\xi\Big) f(v, x, t)\Big) \\ - v\partial_x f(v, x, t) - \partial_t f(v, x, t) = 0, \end{aligned} \tag{1.3.2}$$

which structurally resembles the Kolmogorov equation (1.1.1), except for the presence of the drift term and the nonlinear interaction term. Equation (1.3.1) arises as *mean field limit* of N

interacting stochastic particles. More precisely, consider the system

$$\begin{cases} dV_{i,t} = \Big(b(V_{i,t}, X_{i,t}) + \dfrac{1}{N}\displaystyle\sum_{k=1}^{N} K_t\Big(V_{k,t} - V_{i,t}, X_{k,t} - X_{i,t}\Big)\Big)dt + \sqrt{2}\, dW_{i,t}, & i = 1, \dots, N, \\ dX_{i,t} = dV_{i,t}\; dt, & i = 1, \dots, N, \end{cases} \tag{1.3.3}$$

where $(W_{i,t})_{t\in[0,T]}$ are N independent, standard Wiener processes. As the number of particles N grows, the processes $(V_{i,t}, X_{i,t})_{i\leq N}$ get in some sense uncorrelated, and asymptotically behaves like the mean field process (V_t, X_t) that solves (1.3.2). This relation was formalized by Kac [67] and McKean [86] in the context of kinetic gases, by introducing the notion of *propagation of chaos* (see Chapter 7 for further details and a rigorous definition). The study of the mean field limit represents therefore a powerful tool, both from an analytical and practical point of view, since the system of N particles often becomes unmanageable as N grows. In addition, the mean field framework can be analysed through the associated PDE (1.3.2), which is the perspective we adopt here. This approach provides access to a wide range of analytical techniques, such as fundamental solutions, regularity estimates, and functional inequalities, which allow for a deeper understanding of the existence and the qualitative behavior of solutions (see for instance [64, 95, 96, 122]).

In Chapter 6 we study a particular McKean-Vlasov process, that is a Kuramoto-type model. These are very famous models, used to describe the collective behavior of a population viewed as a system of interacting oscillators (see for instance the survey [111]). We study the Kuramoto model with inertia and noise: consider a system of N identical coupled oscillators, and denote as $\bar{\vartheta}_{i,t}$ and $\omega_{i,t}$ their phase and frequency, respectively. Let K be a constant (modelling the intensity of the coupling between oscillators), m a positive constant (modelling inertial effects), and let $D > 0$ be the noise amplitude. The evolution of the system is given by

$$\begin{cases} d\omega_{i,t} = \dfrac{1}{m}\left(- \omega_{i,t} + \dfrac{K}{N}\displaystyle\sum_{j=1}^{N} \sin\left(\bar{\vartheta}_{j,t} - \bar{\vartheta}_{i,t}\right)\right)dt + \dfrac{D}{m}\sqrt{2}\, dW_{i,t}, & i = 1, \dots, N, \\ d\bar{\vartheta}_{i,t} = \omega_{i,t} dt, & i = 1, \dots, N. \end{cases} \tag{1.3.4}$$

Owing to the periodic nature of the Kuramoto model, we set $\vartheta_{i,t} := \bar{\vartheta}_{i,t} - 2\pi k\Big(\bar{\vartheta}_{i,t}\Big)$, where $k(x) := \max\{j \in \mathbb{Z} : 2\pi j < x\}$. The mean field process density $\varrho = \varrho(\omega, \vartheta, t)$ solves

$$\begin{aligned} \frac{D}{m^2}\partial^2_{\omega\omega}\varrho(\omega, \vartheta, t) - \frac{1}{m}\partial_\omega \left(\Big(- \omega + K \int_{\mathbb{R}}\int_0^{2\pi} \sin(\vartheta' - \vartheta)\varrho(\omega, \vartheta, t) d\omega d\vartheta \Big) \varrho_t(\omega, \vartheta) \right) & \\ - \omega\partial_\vartheta\varrho(\omega, \vartheta, t) - \partial_t\varrho(\omega, \vartheta, t) = 0. & \end{aligned} \tag{1.3.5}$$

We provide a characterization of the relevant Cauchy problems by employing the tools of degenerate Kolmogorov operators developed in recent years (see for instance the survey [9] by Anceschi and Polidoro). Our results include existence, uniqueness and estimates, and can be regarded as a particular instance of the broader theory of kinetic McKean–Vlasov equations. In this broader context, we mention for example [95, 96], where the authors considered general diffusion terms depending also on the law of the process, as well as [62, 64, 122], which address models in a low-regularity setting (*i.e.* allowing for distributional drift). Chapter 6 is developed within a different framework, namely the classical theory of degenerate Kolmogorov equations, and is entirely devoted to the analytical study of the PDE (1.3.5). This approach enables us to precisely characterize the intrinsic regularity of the solutions and to treat also the case of non-identical oscillators. Moreover, it enables us to introduce a *stable* finite difference scheme based on the following *Lie derivative approximation*

$$\omega \partial_\vartheta \varrho(\omega, \vartheta, t) + \partial_t \varrho(\omega, \vartheta, t) \approx \frac{\varrho(\omega, \vartheta - \omega \Delta t, t - \Delta t) - \varrho(\omega, \vartheta, t)}{\Delta t}. \tag{1.3.6}$$

The numerical results obtained here are consistent with those in [1], which were derived in a different, fully probabilistic setting. More precisely, the authors adopted a Monte Carlo scheme to simulate the N particles stochastic system (1.3.4), whereas in our study we focus on its mean field limit, computing the relevant density through the aforementioned finite difference scheme.

These results are presented in the paper [99] by the author, Polidoro, and Vernia.

In Chapter 7 we turn our attention to a McKean-Vlasov optimal control problem, where the kernel $K(v, x, t)$ in (1.3.2) is of the form $u(t)K(v, x)$, with $u(t)$ being the control function. The cost functional is

$$J(f, u) = \int_{\mathbb{R}^2} \int_0^T h(v, x, t) f(v, x, t) dt dv dx + \int_{\mathbb{R}^2} g(v, x) f(v, x, T) dv dx + \frac{\nu}{2} \int_0^T |u(t)|^2 dt, \tag{1.3.7}$$

where f is the density of the process, h represents the running cost, g the final cost, and $\nu \geq 0$ is a penalization parameter. We study this control problem as mean field approximation of the N particles control problem relevant to (1.3.3), with cost functional

$$J_N\Big(\big(V_{i,t}, X_{i,t}\big), u\Big) = \sum_{i=1}^{N} \frac{1}{N} \mathbb{E} \left[\int_0^T h\big(V_{i,t}, X_{i,t}, t\big) dt + g\big(V_{i,T}, X_{i,T}\big) \right] + \frac{\nu}{2} \int_0^T |u(t)|^2 dt.$$

The study of the relation between these two problems is a very recent area of research, first explored by Fischer and Livieri [53] in the context of a mean-variance portfolio optimization. Among the most significant subsequent contributions are the works of Lacker [75], Djete [49] and Djete, Possamaï and Tan [50]. These studies, using a fully probabilistic

approach, established the convergence of *approximate* optimal controls for the N particles problem to the optimal controls of the corresponding mean field problem. In this thesis we adopt a completely different approach: we prove the Γ-convergence of the N particles cost functional to its mean field counterpart in the $L^\infty(0,T)$ weak* topology, this implying the convergence of the optimal controls, and hence the validity of the mean field approach to compute the optimal controls. In particular, we recover the results of [75] under the uniqueness of minimizers. Moreover, given this relation, we provide some optimality conditions for the mean field optimal control by employing the relevant PDE. This method has been gaining popularity in recent years [10–12, 107], since it enables to employ the Lagrangian framework to minimize the mean field functional (1.3.7). Finally, we make use of these optimality conditions by implementing a gradient descent algorithm to numerically compute the mean field optimal controls.
These results are presented in the submitted paper [97].

Part I

Regularity theory

Chapter 2

Preliminaries: Hörmander operators on Carnot groups

In this preliminary chapter we introduce the analytical framework for the study of the differential operators

$$\mathscr{L}_{\mathbb{G}} = \sum_{i,j=1}^{m_1} a_{ij} X_i X_j + \sum_{i=1}^{m_1} b_i X_i + c, \qquad \mathscr{H}_{\mathbb{G}} = \mathscr{L}_{\mathbb{G}} - \partial_t,$$

where A is an $m_1 \times m_1$ symmetric matrix with continuous real entries that satisfies the uniform ellipticity condition, and the low order coefficients b and c are continuous and bounded. Here, $\{X_i\}_{i=1}^{m_1}$ are smooth vector fields on $\mathbb{R}^N$ verifying the following assumption

(H1) *Hörmander's rank condition holds*

$$\text{rank}\, \mathfrak{g}(x) = N, \qquad \forall x \in \mathbb{R}^N,$$

and for every $i = 1, \dots, m_1$ *we have* $X_i(0) = \partial_{x_i}$.

In this dissertation we employ the notations introduced in the monograph [19] concerning differential operators on stratified Lie groups, that we briefly recall. Let $\mathfrak{g}$ be the smallest Lie subalgebra of all smooth vector fields on $\mathbb{R}^N$ containing $\{X_i\}_{i=1}^{m_1}$. For every $x \in \mathbb{R}^N$ we identify every smooth vector field $X(x) = \sum_{j=1}^N \varphi_j(x) \partial_{x_j}$ with the vector valued function $\big(\varphi_1(x), \dots, \varphi_N(x)\big) \in \mathbb{R}^N$, and we set $\mathfrak{g}(x) := \big\{X(x), X \in \mathfrak{g}\big\}$.

In this thesis we study Hörmander operators by following the approach introduced by Folland in [54], which is based on the existence of an underlying Carnot group (that we recall to be a homogeneous Lie group with stratified algebra). This approach is motivated by several considerations, among which we point out the following. Constant coefficients elliptic operators are left-invariant w.r.t. the Euclidean translation, and are homogeneous of degree 2 w.r.t. the Euclidean dilations. Hence, these operators are linked to the *abelian* Euclidean group on $\mathbb{R}^N$. Regularity theory for *uniformly* elliptic operators can then be derived by means of perturbation arguments, based on such Euclidean structure. When dealing with degenerate operators, it is therefore natural to assume an underlying *non abelian* Lie group structure. More precisely, it is useful to assume that the *sublaplacian*

$$\Delta_{\mathbb{G}} = \sum_{i=1}^{m_1} X_i^2,$$

is left-invariant and homogeneous of degree 2 w.r.t. a non abelian Lie group on $\mathbb{R}^N$, whose Lie algebra reflects Hörmander's rank condition (namely, the behaviour of the commutators). Such a framework, in particular, makes it possible to exploit a homogeneous structure, thereby enabling many of the standard constructions available in the Euclidean setting, such as the analysis of singular integrals. In order to pursue this approach, we consider vector fields satisfying the following assumptions

(H2) $X_1, \dots, X_{m_1}$ *are linearly independent smooth vector fields, and*

$$\operatorname{rank} \mathfrak{g} = N.$$

(H3) *There exists a family of (non-isotropic) dilations* $\{\delta_r\}_{r>0}$ *on* $\mathbb{R}^N$ *of the kind*

$$\delta_r(x) = \Big(r^{\sigma_1} x_1, r^{\sigma_2} x_2, \dots, r^{\sigma_N} x_N\Big),$$

where $1 = \sigma_1 \le \sigma_2 \le \dots \le \sigma_N$ *are integers such that the vector fields* $X_1, \dots, X_{m_1}$ *are* δ_r*-homogeneous of degree* 1*,* i.e.

$$X_i\left(f\Big(\delta_r(x)\Big)\right) = r\Big(X_i f\Big)\Big(\delta_r(x)\Big), \quad \forall r>0,\ x \in \mathbb{R}^N,\ f \in C^\infty(\mathbb{R}^N),\ i = 1, \dots, m_1.$$

Although at first glance **(H1)** and **(H2)** may appear to be similar, we remark that they are in fact independent. For example, $\operatorname{rank} \mathfrak{g}(x) = \operatorname{rank} \mathfrak{g}$ does not necessarily holds for every $x \in \mathbb{R}^N$: consider on $\mathbb{R}^3$ the vector fields $X_1 = \partial_x$ and $X_2 = x\partial_y + y\partial_t$; one has $\operatorname{rank} \mathfrak{g}(0) = 3$ and $\operatorname{rank} \mathfrak{g} = 4$.

If assumptions **(H1)**, **(H2)** and **(H3)** hold, it is possible to explicitly construct a Carnot group $\mathbb{G} = \Big(\mathbb{R}^N, \circ, \delta_r\Big)$ whose Lie algebra $\mathfrak{g}$ is generated by the vector fields $\{X_i\}_{i=1}^{m_1}$.

Theorem 2.0.1 (Theorem 1.7 in [22])**.** *Assume that the vector fields* $\{X_i\}_{i=1}^{m_1}$ *verify **(H1)**, **(H2)** and **(H3)**. Then, there exists a Carnot group* $\mathbb{G} = \Big(\mathbb{R}^N, \circ, \delta_r\Big)$ *(where the dilations* $\{\delta_r\}_{r>0}$ *are the ones given by **(H3)**) such that the vector fields* $\{X_i\}_{i=1}^{m_1}$ *are left-invariant* *w.r.t.* $\circ$, *namely*

$$X_i\Big(f(x\circ\xi)\Big)=\Big(X_i f\Big)(x\circ\xi),\quad \forall x,\xi\in\mathbb{R}^N,\quad f\in C^\infty(\mathbb{R}^N),\quad i=1,\dots,m_1,$$

and the sum of squares operator

$$\Delta_{\mathbb{G}}=\sum_{i=1}^{m_1}X_i^2,$$

is the canonical sublaplacian *on* $\mathbb{G}$.

We now give some remarks relevant to this Carnot group. Assumption **(H3)** induces the following *stratification* on the Lie algebra $\mathfrak{g}$

$$\mathfrak{g}=V_1\oplus\ldots\oplus V_\kappa, \tag{2.0.1}$$

where $V_1=\text{span}\{X_i\}_{i=1}^m$, $V_{k+1}=\text{span}\Big\{[X,Y],X\in V_1,Y\in V_k\Big\}$ and $\Big\{[X,Y],X\in V_1,Y\in V_\kappa\Big\}=\emptyset$, with $[X,Y]=XY-YX$. We observe that the stratification (2.0.1) implies that the dilations $\{\delta_r\}_{r>0}$ take the following form (see Remark 1.4.2 in [27])

$$\delta_r(x)=\left(rx^{(1)},r^2x^{(2)},\dots,r^\kappa x^{(\kappa)}\right),\qquad x^{(k)}\in\mathbb{R}^{m_k},\ k=1,\dots,\kappa,$$

where m_k denotes the dimension of V_k, for every $k=1,\dots,\kappa$. The number

$$Q:=\sum_{k=1}^{\kappa}k\,m_k,$$

is usually called *homogeneous dimension* of $\mathbb{G}$. A group $\mathbb{G}$ with the properties above is often referred to as Carnot group of step κ and m_1 generators.

The study of Hörmander operators on Carnot groups is motivated not only by the rich theory these operators enjoy in the presence of this homogeneous structure, but also by their connection to general Hörmander operators lacking such a structure. Indeed, a remarkable fact is that *every Hörmander sum of squares operator can be approximated by a sublaplacian on a Carnot group*, a deep result stemming from the *lifting and approximation technique* of Rothschild and Stein [106]. While we shall not delve into the technical details of the approximation procedure (as it is quite involved and will not be used in this thesis), we focus on the lifting technique, that will be employed in Chapter 4. To give a concrete example, we present a glimpse of this technique in the context of the Grushin operator on $\mathbb{R}^2$

$$\mathscr{L}=\partial^2_{xx}+x^2\partial^2_{yy}=X_1^2+X_2^2.$$

It is clear that the vector fields $X_1=\partial_x$ and $X_2=x\partial_y$ satisfy Hörmander's rank condition on $\mathbb{R}^2$, since $[X_1, X_2] = \partial_y$. Nevertheless, these vector fields do not fulfil assumption **(H1)** since at the points $(0, y)$ they are not independent. In fact, there is no Lie group structure on $\mathbb{R}^2$ making X_1 and X_2 (hence $\mathscr{L}$) left-invariant. Indeed, if two left-invariant vector fields are independent at a given point, they are necessarily independent at every point. This follows from the fact that, if $c_1X_1 + c_2X_2 = 0$ at some point, with c_1 c_2 constants, then the vector field $Z := c_1X_1+c_2X_2$ is also left-invariant and vanishes at some point, hence it must vanish identically by invariance. In order to recover the group invariance, Rothschild and Stein idea was to add some variables (namely, to *lift* the vector fields), an idea that is somehow in the same spirit of Hadamard's descent method. Consider on $\mathbb{R}^3$ the differential operator

$$\widetilde{\mathscr{L}} = \partial_{xx}^2 + (x\partial_y + \partial_t)^2 = \widetilde{X}_1^2 + \widetilde{X}_2^2.$$

The vector fields $\widetilde{X_1} = X_1 = \partial_x$ and $\widetilde{X_2} = X_2 + \partial_t = x\partial_y + \partial_t$ satisfy both **(H1)** and **(H2)**. We notice that they are left-invariant w.r.t. the Kolmogorov group law on $\mathbb{R}^3$

$$(x, t, y) \circ (\xi, \tau, \eta) = (x + \xi, t + \tau, y + \eta + t\xi),$$

and are homogeneous of degree 1 w.r.t. the dilations $\delta_r(x, t, y) = \Big(rx, rt, r^2y\Big)$. The homogeneous Lie group $\widetilde{\mathbb{G}} = \Big(\mathbb{R}^3, \circ, \delta_r\Big)$ is a Carnot group of step $\kappa = 2$ and homogeneous dimension $Q = 4$ with generators $\widetilde{X_1}$ and $\widetilde{X_2}$, and $\widetilde{\mathscr{L}}$ is the canonical sublaplacian on $\widetilde{\mathbb{G}}$. A crucial point is that the *lifting property* holds: if the function $(x, y) \mapsto u(x, y)$ solves $\mathscr{L}u = 0$, then the function $(x, t, y) \mapsto \tilde{u}(x, t, y) := u(x, y)$ solves $\widetilde{\mathscr{L}}\tilde{u} = 0$. This link allows one to deduce certain regularity properties of $\mathscr{L}$ from the analogous properties of $\widetilde{\mathscr{L}}$. For instance, suppose that one needs an a priori estimate for X_1u, with u being a solution to $\mathscr{L}u = 0$. Such an estimate can be obtained by first proving the analogous bound for $\widetilde{X_1}\tilde{u}$ (often easier to prove thanks to the homogeneous structure), and then transferring it back to X_1u via the lifting property.

In the years immediately following the work of Rothschild and Stein, several authors proposed alternative proofs of the lifting technique, both with the aim of simplifying it and adapting it to different frameworks (see Chapter 10 in [33] for more details about Rothschild and Stein lifting and approximation technique). In Chapter 4, we adopt the lifting technique developed by Bonfiglioli, Lanconelli and Uguzzoni in [25], which provides a procedure to lift a general Carnot group to a *free* Carnot group.

Definition 2.0.2 (Free Carnot group)**.** *Let $\mathbb{G}$ be a Carnot group with Lie algebra generated by the vector fields $\{X_i\}_{i=1}^{m_1}$. We say that $\mathbb{G}$ is a* free Carnot group *if the X_i's and their commutators do not satisfy any linear relation other then those which hold automatically as a consequence of the antisymmetry of the Lie bracket and the Jacobi identity.*

We now provide two famous examples of Carnot groups, one being free and the other not.

- The Heisenberg group $\mathbb{H}^1$ on $\mathbb{R}^3$. The vector fields generating its Lie algebra are

$$X_1 = \partial_x + 2y\partial_t, \qquad X_2 = \partial_y - 2x\partial_t,$$

with commutator $[X_1, X_2] = -4\partial_t$. Clearly, no linear relation between X_1, X_2 and $[X_1, X_2]$ holds except for the antisymmetry of the Lie bracket and the Jacobi identity.

- The Heisenberg group $\mathbb{H}^2$ on $\mathbb{R}^5$. The vector fields generating its Lie algebra are

$$X_1 = \partial_{x_1} + 2y_1\partial_t, \quad X_2 = \partial_{x_2} + 2y_2\partial_t, \quad X_3 = \partial_{y_1} - 2x_1\partial_t, \quad X_4 = \partial_{y_2} - 2x_2\partial_t.$$

In this case the following relation holds

$$[X_1, X_3] = -4\partial_t = [X_2, X_4],$$

that is a nontrivial relation between two commutators of step 2.

As mentioned above, the main motivation behind Rothschild and Stein lifting technique is to recover a homogeneous group structure for vector fields that do not possess such homogeneity. This raises a natural question: why a lifting technique should be employed if a homogeneous group structure is already available? In the analysis of [25, 26], which is closely related to our work, the motivation is driven by the following purpose. The authors employed Levi's parametrix method to construct the fundamental solution of the following Hölder coefficients operator

$$\mathscr{H}_{\mathbb{G}} = \sum_{i,j=1}^{m_1} a_{ij}(x,t) X_i X_j - \partial_t.$$

This method provides us with a fundamental solution given by the sum of a *parametrix* and a perturbation term. The parametrix is the fundamental solution of a *constant coefficients* operator

$$\mathscr{H}_{\bar{A}} = \sum_{i,j=1}^{m_1} \bar{a}_{ij} X_i X_j - \partial_t, \tag{2.0.2}$$

with $\bar{A} = (\bar{a}_{ij})$ being a symmetric and positive definite $m_1 \times m_1$ matrix (see Section 3.3 and Section 4.3 for a precise definition of parametrix). To make Levi's method work, one needs a precise estimate on the dependence of the parametrix on the corresponding matrix $\bar{A}$. The estimate of this dependence is quite straightforward in the non degenerate case: consider for instance the parabolic operator

$$\mathscr{H} = \sum_{i,j=1}^{N} \bar{a}_{ij}\partial_{x_i}\partial_{x_j} - \partial_t.$$

The corresponding fundamental solution $\Gamma_{\bar{A}}$ (hence the parametrix) can be expressed in terms of the fundamental solution of the heat equation as follows

$$\Gamma_{\bar{A}}(x,t) = \frac{c}{t^{\frac{N}{2}}} \exp\left(- \frac{T_{\bar{A}}(x)^2}{4t}\right),$$

for some $c > 0$, with $T_{\bar{A}}$ denoting the diffeomorphism $x \mapsto T_{\bar{A}}(x) = \bar{A}^{-\frac{1}{2}}x$. A careful inspection of $T_{\bar{A}}$ yields precise estimates on the dependence of $\Gamma_{\bar{A}}$ on $\bar{A}$.
If $\mathbb{G}$ is a free group, it has been shown in [28] that there exists an analogous Lie group diffeomorphism $T_{\bar{A}}$ turning (2.0.2) into the canonical heat operator $\Delta_{\mathbb{G}} - \partial_t$, as in the non degenerate case. Moreover, the properties of such diffeomorphism allow one to obtain the required estimates. The natural question that now arises is obvious: does such diffeomorphism always exist if $\mathbb{G}$ is not free? Unfortunately, the authors of [28] provided a negative answer, showing that such diffeomorphism may not exist if $\mathbb{G}$ is not free. To overcome this difficulty, in [25] the authors developed the following lifting technique.

Theorem 2.0.3 (Theorem 8.3 in [25])**.** *Let $\{X_i\}_{i=1}^{m_1}$ be a system of vector fields generating a Carnot group $\mathbb{G} = \left(\mathbb{R}^N, \circ, \delta_r\right)$ of homogeneous dimension Q and step κ. Then, there exists $\widetilde{N} \in \mathbb{N}$ and a free Carnot group $\widetilde{\mathbb{G}} = \left(\mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N, \widetilde{\circ}, \widetilde{\delta}_r\right)$ with homogeneous dimension $\widetilde{Q}$, step κ, and generators $\left\{\widetilde{X}_i\right\}_{i=1}^{m_1}$, such that the following lifting property holds: let $\pi : \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \to \mathbb{R}^N$ denote the canonical projection onto the last N coordinates, then*

$$\widetilde{X}_i u\Big(\pi(\widetilde{x}, x)\Big) = X_i u(x), \qquad (\widetilde{x}, x) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N, \tag{2.0.3}$$

for every $u \in C^\infty(\mathbb{R}^N)$ and for every $i = 1, \dots, m_1$.

This lifting technique will also play a crucial role in our analysis for two reasons.
The first one is that we construct the fundamental solution of

$$\mathscr{H}_{\mathbb{G}} = \sum_{i,j=1}^{m_1} a_{ij}(x,t)X_iX_j + \sum_{i=1}^{m_1} b_i(x,t)X_i + c(x,t) - \partial_t,$$

by means of parametrix method, obtaining also several useful properties. To obtain these results we rely on the estimates derived in [25], which are based on this lifting technique.
The second main reason is the derivation of further precise estimates (in the same spirit of the ones mentioned above) to establish a parabolic Harnack inequality for the divergence

form operator

$$\mathscr{H}_{\mathbb{G}}^{\mathrm{div}} = \sum_{i,j=1}^{m_1} X_i\Big(a_{ij}(x,t)X_j\Big) + \sum_{i=1}^{m_1} b_i(x,t)X_i + c(x,t) - \partial_t.$$

As mentioned in the introduction, to derive such Harnack inequality we employ the method introduced by Polidoro in [102], which relies on these precise estimates. Therefore, we employ the lifting technique to lift $\mathscr{H}_{\mathbb{G}}^{\mathrm{div}}$ and obtain these estimates for the lifted operator $\mathscr{H}_{\widetilde{\mathbb{G}}}^{\mathrm{div}}$. This, in turn, yields a Harnack inequality for $\mathscr{H}_{\widetilde{\mathbb{G}}}^{\mathrm{div}}$. Finally, the lifting property (2.0.3) allows us to transfer this result and to obtain the desired Harnack inequality for $\mathscr{H}_{\mathbb{G}}^{\mathrm{div}}$.

Chapter 3

Fundamental solution, maximum principle and Harnack inequality for second order subelliptic operators

3.1 Introduction

In this chapter we focus on the analysis of the following class of subelliptic operators

$$\mathscr{L}_{\mathbb{G}} = \sum_{i,j=1}^{m_1} a_{ij}(x)X_iX_j + \sum_{i=1}^{m_1} b_i(x)X_i + c(x), \qquad x \in \Omega. \tag{3.1.1}$$

Here Ω is an open, connected and bounded subset of $\mathbb{R}^N$, while X_i's are smooth vector fields on $\mathbb{R}^N$

$$X_i(x) = \sum_{j=1}^{N} \varphi_j^i(x)\partial_{x_j}, \qquad i = 1, \ldots, m_1,$$

with $\varphi_j^i \in C^\infty(\mathbb{R}^N)$. In the sequel, $A(x) = \left(a_{i,j}(x)\right)_{i,j=1}^{m_1}$ will denote an $m_1 \times m_1$ symmetric matrix with continuous real entries that satisfies the usual *uniform ellipticity condition*, moreover the low order coefficients $b(x)$ and $c(x)$ are supposed to be continuous and bounded. We suppose that $m_1 < N$, this implying the strong degeneracy of $\mathscr{L}_{\mathbb{G}}$. However, we recover some regularity properties for $\mathscr{L}_{\mathbb{G}}$ by assuming that the vector fields $\{X_i\}_{i=1}^{m_1}$ verify the Hörmander's rank condition.

The vector fields are assumed to satisfy the following hypotheses

(H1) *Hörmander's rank condition holds*

$$\text{rank}\, \mathfrak{g}(x) = N, \qquad \forall x \in \mathbb{R}^N,$$

and for every $i = 1, \dots, m_1$ we have $X_i(0) = \partial_{x_i}$.

(H2) $X_1, \dots, X_{m_1}$ *are linearly independent smooth vector fields, and*

$$\operatorname{rank} \mathfrak{g} = N.$$

(H3) *There exists a family of (non-isotropic) dilations $\{\delta_r\}_{r>0}$ on $\mathbb{R}^N$ of the kind*

$$\delta_r(x) = \Big(r^{\sigma_1}x_1, r^{\sigma_2}x_2, \dots, r^{\sigma_N}x_N\Big),$$

where $1 = \sigma_1 \leq \sigma_2 \leq \ldots \leq \sigma_N$ are integers such that the vector fields $X_1, \dots, X_{m_1}$ are δ_r-homogeneous of degree 1, i.e.

$$X_i\left(f\Big(\delta_r(x)\Big)\right) = r\Big(X_i f\Big)\Big(\delta_r(x)\Big), \qquad \forall r > 0,\ x \in \mathbb{R}^N,\ f \in C^\infty(\mathbb{R}^N),\ i = 1, \dots, m_1.$$

We recall now Theorem 2.0.1, that asserts that the assumptions **(H1)**, **(H2)** and **(H3)** yield the existence of a Carnot group $\mathbb{G} = \Big(\mathbb{R}^N, \circ, \delta_r\Big)$ (where the dilations $\{\delta_r\}_{r>0}$ are the ones given by **(H3)**) such that $\{X_i\}_{i=1}^{m_1}$ are left invariant w.r.t. $\circ$, and the operator

$$\Delta_{\mathbb{G}} = \sum_{i=1}^{m_1} X_i^2, \tag{3.1.2}$$

is the canonical sublaplacian on $\mathbb{G}$. Despite being degenerate, this is an hypoelliptic operator, meaning that every distributional solution u to $\Delta_{\mathbb{G}} u = f$ is smooth whenever the right-hand side term f is smooth.

We associate to the smooth vector fields some function spaces, that are the natural spaces in which to look for classical solutions to $\mathscr{L}_{\mathbb{G}} u = f$.

Definition 3.1.1 (Lie derivative). *Let $X = \sum_{j=1}^N \varphi_j \partial_{x_j}$ be a smooth vector field on $\mathbb{R}^N$. For every $x \in \mathbb{R}^N$ we define as* integral path of X starting at x *the solution to the following Cauchy problem*

$$\begin{cases} \dot\gamma_{X,x}(s) = X\Big(\gamma_{X,x}(s)\Big), \\ \gamma_{X,x}(0) = x. \end{cases}$$

We say that a function u is Lie differentiable at x w.r.t. X if the following limit exists and is finite

$$Xu(x) = \lim_{s \to 0} \frac{u\Big(\gamma_{X,x}(s)\Big) - u(x)}{s}.$$

We observe that if $u : \Omega \to \mathbb{R}$ is such that $\partial_{x_j} u \in C(\Omega)$ for every $j = 1, \dots, N$, then it is Lie differentiable w.r.t. the vector field X, and its Lie derivative is given by $Xu = \sum_{j=1}^N \varphi_j \partial_{x_j} u$

considered as linear combination of $\partial_{x_1}u, \ldots, \partial_{x_N}u$. On the contrary, the reverse implication does not necessarily hold.
We introduce the functional space naturally associated with this setting

$$C^2_{\mathbb{G}}(\Omega) := \Big\{u \in C(\Omega) : X_i u, X_i X_j u \in C(\Omega), \quad \text{for every} \ \ i, j = 1, \ldots, m_1\Big\},$$

where the derivatives $X_i u$ and $X_i X_j u$ are meant as first and second order Lie derivatives.
We are now ready to give the definition of classical solutions to (3.1.1).

Definition 3.1.2 (Classical solutions for $\mathscr{L}_{\mathbb{G}}$)**.** *Let $f \in C(\Omega)$. We say that a function u is a classical solution to*

$$\mathscr{L}_{\mathbb{G}} u(x) = \sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j u(x) + \sum_{i=1}^{m_1} b_i(x) X_i u(x) + c(x)u(x) = f(x), \qquad x \in \Omega,$$

if $u \in C^2_{\mathbb{G}}(\Omega)$ and the above equation is satisfied at every point $x \in \Omega$.

The first main result of this chapter concerns the existence of the fundamental solution of (3.1.1). We begin by recalling its definition (from now on for every set Ω we denote as $D_\Omega = \Big\{(x, x) \in \Omega \times \Omega\Big\}$ its diagonal).

Definition 3.1.3 (Fundamental solution of $\mathscr{L}_{\mathbb{G}}$)**.** *Let Ω be a bounded open subset of $\mathbb{R}^N$. We say that a function $\Gamma : (\Omega \times \Omega) \setminus D_\Omega \to \mathbb{R}$ is a fundamental solution for* (3.1.1) *if:*

1. *for every $\xi \in \Omega$ we have that $x \mapsto \Gamma(x, \xi) \in C^2_{\mathbb{G}}(\Omega \setminus \{\xi\}) \cap L^1(\Omega)$ and is a classical solution to $\mathscr{L}_{\mathbb{G}}\Gamma(x, \xi) = 0$ for every $x \in \Omega \setminus \{\xi\}$;*

2. *for every $\varphi \in C^\infty_c(\Omega)$ the function*

$$w(x) := \int_\Omega \Gamma(x, \xi)\varphi(\xi) d\xi, \qquad x \in \Omega,$$

belongs to $C^2_{\mathbb{G}}(\Omega)$ and is a classical solution to $\mathscr{L}_{\mathbb{G}} w = -\varphi$.

We prove the existence of the fundamental solution under the assumption that the coefficients belong to the Hölder space $C^\alpha_{\mathbb{G}}(\Omega)$, defined w.r.t. a distance d related to the Carnot group structure (see (3.2.3) and Definition 3.2.7 in Section 3.2). In particular we prove that, as in the non degenerate case, the fundamental solution Γ belongs to the space $C^{2+\alpha}_{\mathbb{G},\mathrm{loc}}\Big((\Omega \times \Omega) \setminus D_\Omega\Big)$, namely the space of functions that, for every $\Omega_0 \subset\subset \Big((\Omega \times \Omega) \setminus D_\Omega\Big)$, belong $C^\alpha_{\mathbb{G}}(\Omega_0)$ with first and second order Lie derivatives in $C^\alpha_{\mathbb{G}}(\Omega_0)$.
We summarise below all the assumptions on the coefficients

(H4) *For every $i, j = 1, \ldots, m_1$ we have $a_{ij}, b_i, c \in C\left(\overline{\Omega}\right)$, and there exist $\alpha \in (0, 1]$ and*

$M_1, M_2 > 0$ *such that*

$$|a_{ij}(x)|, |b_i(x)|, |c(x)| \leq M_1, \qquad \forall x \in \Omega,$$
$$|a_{ij}(x) - a_{ij}(y)|, |b_i(x) - b_i(y)|, |c(x) - c(y)| \leq M_2\, d(x,y)^\alpha, \qquad \forall x, y \in \Omega.$$

Moreover, there exists $\lambda > 0$ *such that* A *belongs to the uniform ellipticity class*

$$\boldsymbol{M}_\lambda := \left\{ M \in \mathit{Sym}_{m_1}(\mathbb{R}) : \ \frac{1}{\lambda}|\xi|^2 \leq \langle M(x)\xi, \xi\rangle \leq \lambda |\xi|^2, \quad \forall x \in \Omega, \, \forall \xi \in \mathbb{R}^{m_1} \right\}.$$

We present the first main result relevant to the fundamental solution. In the following statement, $|\Omega|$ denotes the Lebesgue measure of Ω.

Theorem 3.1.4. *Let* $\mathscr{L}_{\mathbb{G}}$ *be a differential operator as in* (3.1.1) *verifying **(H1),(H2),(H3)** and **(H4)**. Then, there exists* $\delta_0 > 0$ *(depending on* M_1, M_2 *and* λ*) such that, if* $|\Omega| < \delta_0$*, then there exists a fundamental solution* Γ *to* $\mathscr{L}_{\mathbb{G}}$*. Moreover, for every* $\xi \in \Omega$ *the map* $x \mapsto \Gamma(x,\xi)$ *belongs to* $C^{2+\alpha}_{\mathbb{G},\mathrm{loc}}(\Omega \setminus \{\xi\})$*. Finally, for every* $f \in C^{\alpha}_{\mathbb{G}}(\Omega) \cap L^1(\Omega)$ *the function*

$$u(x) := \int_\Omega \Gamma(x,\xi) f(\xi) d\xi, \qquad x \in \Omega,$$

belongs to $C^{2+\alpha}_{\mathbb{G}}(\Omega)$ *and is a classical solution to* $\mathscr{L}_{\mathbb{G}} u = -f$*.*

We prove this result by employing the parametrix method introduced by Levi in [82] and [83], where he considered *non degenerate* uniformly elliptic operators

$$\mathscr{L} = \sum_{i,j=1}^{N} a_{ij}(x)\partial_{x_i}\partial_{x_j} + \sum_{i=1}^{N} b_i(x)\partial_{x_i} + c(x).$$

This method was subsequentially improved by several authors, we quote Kalf [68] and the bibliography therein among them. Let us briefly describe Levi's method: we consider the fundamental solution $\Gamma_{A_\xi} = \Gamma_{A_\xi}(w,w')$ of the constant coefficient operator $\mathscr{L}_{A_\xi} = \sum a_{ij}(\xi)\partial_{x_i}\partial_{x_j}$, obtained by *freezing* at ξ the coefficients of the principal part of $\mathscr{L}$; the parametrix is defined as $Z(x,\xi) := \Gamma_\xi(x,\xi)$, and the fundamental solution of $\mathscr{L}$ is then obtained from Z by a perturbative argument. A key point in the non degenerate case is that the explicit expression of the parametrix is available. By using the Hölder continuity of the coefficients and the explicit expression of the derivatives of the parametrix, we can rely on the singular integrals theory to prove the existence of the fundamental solution to $\mathscr{L}$. Even though in the degenerate case we do not have an explicit expression for the parametrix, the estimates developed in [25], alongside a thorough investigation of the convolution of some singular functions, provide us with the tools we need to adapt the method to the degenerate case.

The second main contribution of this chapter is a novel proof of a strong maximum principle and a Harnack inequality, which we obtain as a direct application of Theorem 3.1.4. These results rely on the mean value formula developed by Pallara and Polidoro in [91], valid for divergence form operators and built upon the fundamental solution of the corresponding adjoint. Thus, we need some additional assumptions on the coefficients

(H5) *For every $i, j = 1, \dots, m_1$ we have $X_j a_{ij}, X_i b_i, \in C\left(\overline{\Omega}\right)$, and there exist $M_1, M_2 > 0$ and $\alpha \in (0, 1]$ such that*

$$
\begin{aligned}
&|X_j a_{ij}(x)|, |X_i b_i(x)| \leq M_1, && \forall x \in \Omega,\\
&|X_j a_{ij}(x) - X_j a_{ij}(y)|, |X_i b_i(x) - X_i b_i(y)|, |c(x) - c(y)| \leq M_2\, d(x, y)^{\alpha}, && \forall x, y \in \Omega.
\end{aligned}
$$

We introduce the operators $\mathrm{div}_{\mathbb{G}}$ and $\nabla_{\mathbb{G}}$, that are the extension of the divergence and gradient operator in the non-Euclidean setting of $\mathbb{G}$. In order to give the right meaning to such operators, we agree to identify a *horizontal section* $F = \sum_{i=1}^{m_1} F_i X_i$ with its *canonical coordinates* $F = (F_1, \dots, F_{m_1})$. Then, for every $F \in C^1_{\mathbb{G}}\left(\mathbb{R}^N, \mathbb{R}^{m_1}\right)$ and for every real valued $f \in C^1_{\mathbb{G}}(\mathbb{R}^N)$ we set

$$
\mathrm{div}_{\mathbb{G}} F := \sum_{i=1}^{m_1} X_i F_i, \qquad \nabla_{\mathbb{G}} f := \sum_{i=1}^{m_1} (X_i f) X_i. \tag{3.1.3}
$$

If **(H1)**, **(H2)**, **(H3)**, **(H4)** and **(H5)** hold, both the divergence form operator and its adjoint are well defined

$$
\begin{aligned}
\mathscr{L}^{\mathrm{div}}_{\mathbb{G}} u(x) &= \sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j u(x) + \sum_{i=1}^{m_1} \left(\sum_{j=1}^{m_1} X_j a_{ij}(x) + b_i(x) \right) X_i u(x) + c(x) u(x),\\
&= \mathrm{div}_{\mathbb{G}}(A(x) \nabla_{\mathbb{G}} u(x)) + \langle b(x), \nabla_{\mathbb{G}} u(x) \rangle + c(x) u(x),
\end{aligned} \tag{3.1.4}
$$

$$
\mathscr{L}^{\mathrm{div},*}_{\mathbb{G}} u(x) = \mathrm{div}_{\mathbb{G}}(A(x) \nabla_{\mathbb{G}} u(x)) - \langle b(x), \nabla_{\mathbb{G}} u(x) \rangle + (c(x) - \mathrm{div}_{\mathbb{G}} b(x)) u(x). \tag{3.1.5}
$$

With a slight abuse of notation, we shall denote by $\mathscr{L}_{\mathbb{G}}$ the divergence form operator (3.1.4) and by $\mathscr{L}^*_{\mathbb{G}}$ its adjoint (3.1.5), so as to keep the notation as light as possible.
We employ the mean value formula in [91] to prove our second main contribution, formulated in the following two theorems.

Theorem 3.1.5. *Let $\mathscr{L}_{\mathbb{G}}$ be a differential operator as in* (3.1.4) *verifying **(H1)**,**(H2)**,**(H3)**, **(H4)** and **(H5)**, and let u be a classical solution to $\mathscr{L}_{\mathbb{G}} u = f$ in Ω. Suppose that $c \leq 0$, $f \geq 0$ and $\mathrm{div}_{\mathbb{G}} b - c \geq 0$. If there exists $\xi \in \Omega$ such that $u(\xi) = \max_{\Omega} u \geq 0$, then*

$$
u(x) = u(\xi) \quad \textit{and} \quad f(x) = u(\xi) c(x),
$$

for every $x \in \Omega$. An analogous result holds if $u(\xi) = \min_{\Omega} u \leq 0$ and $f \leq 0$. Finally, the *assumption on the sign of* $u(\xi)$ *can be dropped if* $c = 0$.

Theorem 3.1.6. *Let* $\mathscr{L}_{\mathbb{G}}$ *be a differential operator as in* (3.1.4) *verifying **(H1)**,**(H2)**,**(H3)**, **(H4)** and **(H5)**, and assume that* $\mathrm{div}_{\mathbb{G}} b - c \geq 0$. *Then, for every* H *compact and connected subset of* Ω *there exists* $C_H > 0$ *(depending on* $M_1, M_2, \lambda, \Omega$ *and* H*) such that*

$$\sup_H u \leq C_H \inf_H u,$$

for every $u \geq 0$ *classical solution to* $\mathscr{L}_{\mathbb{G}} u = 0$.

The proof relies on arguments similar to those usually employed for harmonic functions (see Chapter 1). We point out the main difficulties encountered when dealing with the degenerate operator (3.1.4).

Let us consider first harmonic functions. The Gauss mean value formula states that

$$\begin{aligned} \Delta u = 0 \iff u(\xi) &= \frac{1}{r^N} \int_{B_r(\xi)} M(x,\xi) u(x) dx, \\ M(x,\xi) &= \omega_N^{-1}, \\ B_r(\xi) &:= \left\{ x \in \Omega : |\xi - x|^2 < r \right\}, \end{aligned} \tag{3.1.6}$$

where ω_N denotes the measure of the unit ball in $\mathbb{R}^N$. Analogous mean value formula for uniformly elliptic operators have been proved in Theorem 2.1 in [91]. In this case the kernel M is not constant, but is strictly positive and bounded for sufficiently small r, see (2.5), (2.6) and (2.8) in [91]. These two properties of M allow to prove a maximum principle and a Harnack inequality for uniformly elliptic operators, following the same elementary argument for harmonic functions (see for instance Theorem 2.2 and Theorem 2.5 in [61]). A similar mean value formula for the degenerate operator $\mathscr{L}_{\mathbb{G}}$ in (3.1.4) has recently been proved (see Theorem 3.3 in [91]), unfortunately in this case the kernel is merely nonnegative, as the following basic example shows. Let us consider on $\mathbb{R}^3$ the Kohn Laplacian on the Heisenberg group

$$\mathscr{L}_{\mathbb{H}} := (\partial_{x_1} + 2x_2\partial_{x_3})^2 + (\partial_{x_2} - 2x_1\partial_{x_3})^2 = X_1^2 + X_2^2. \tag{3.1.7}$$

As in (3.1.6), we can represent any classical solution to $\mathscr{L}_{\mathbb{H}} u = 0$ as

$$\begin{aligned} u(\xi) &= \frac{C}{r^4} \int_{B_r(\xi)} M_{\mathbb{H}}(x,\xi) u(x) dx, \\ M_{\mathbb{H}}(x,\xi) &= \frac{|(x_1 - \xi_1, x_2 - \xi_2)|^2}{\sqrt{((x_1-\xi_1)^2 + (x_2-\xi_2)^2)^2 + 16(x_3 - \xi_3 + x_1\xi_2 - \xi_1 x_2)^2)}}, \\ B_r(x) &= \left\{ \xi \in \Omega : \left((x_1-\xi_1)^2 + (x_2-\xi_2)^2)^2 + 16(x_3 - \xi_3 + x_1\xi_2 - \xi_1 x_2)^2 \right)^{\frac{1}{4}} < r \right\}, \end{aligned}$$

for a suitable positive constant C. We note that $M_{\mathbb{H}}(x,\xi) = 0$ in the line $\Big\{(x_1,x_2) = (\xi_1,\xi_2)\Big\}$. Even though we cannot rely on the fact that $\inf_{B_r(\xi)} M(\cdot,\xi) > 0$, we rely on a geometric argument that allows us to stay far away from the set $\{M_{\mathbb{H}}(x,\xi) = 0\}$ in our proof of the strong maximum principle and the Harnack inequality. Specifically, we consider the integral paths of the vector fields X_1 and X_2 starting at ξ

$$\gamma_{X_1,\xi}(s) = \Big(s+\xi_1, \xi_2, -s\xi_2+\xi_3\Big), \quad \gamma_{X_2,\xi}(s) = \Big(\xi_1, s+\xi_2, -\xi_1+\xi_3\Big),$$

and we note that $M_{\mathbb{H}}\Big(\gamma_{X_1,\xi}(s),\xi\Big) = M_{\mathbb{H}}\Big(\gamma_{X_2,\xi}(s),\xi\Big) = 1$. This implies that, if we move along the integral paths of X_1 and X_2, we preserve a positive lower bound for $M_{\mathbb{H}}$, as well as ensuring that it remains uniformly bounded from above. Consequently, we can establish a local maximum principle and a Harnack inequality by following these paths. Then, this local result extends globally by using Hörmander's condition **(H1)**, that allows us to reach every point by following the integral paths of X_1 and X_2. This argument does not depend on the particular choice of the Kohn Laplacian operator (3.1.7), but applies to every operator (3.1.4) satisfying **(H1)**,**(H2)**,**(H3)**, **(H4)** and **(H5)**.

Before outlining the chapter's plan, let us give some historical background on the topic, alongside discussing the motivation behind it.
The fundamental solution for operators defined in a sub-Riemannian setting has been studied since the pioneering works of Folland [54], Rothschild and Stein [106], and Sanchéz-Calle [109], who proved the existence of the fundamental solution of the equation (3.1.2), together with its asymptotic behaviour. Later on, the theory of sublaplacians on Carnot group has been widely extended, and the potential theory for subelliptic operators has been setup (see [23] for a more extensive discussion). Regarding the fundamental solution for the constant coefficient operator $\mathscr{L} = \sum a_{i,j} X_i X_j$, its existence has been proved by Bonfiglioli, Lanconelli and Uguzzoni in [28] and [25], employing the lifting techniques developed in [106]. Thereafter, the extension to the Hölder continuous coefficients case (3.1.1) whenever $b = c = 0$ has been extended by the previous authors in [26]. We also quote the work [35] by Bramanti, Brandolini, Manfredini and Pedroni, where the authors weaken the regularity assumption on the vector fields $\{X_i\}_{i=1}^{m_1}$.
Provided the existence of the fundamental solution, we give a simple proof of a maximum principle and a Harnack inequality for the divergence form operator (3.1.4), based on the mean value formula recently developed by Pallara and Polidoro in [91]. We emphasize that the authors of [91] extended to the Hölder coefficients case the results previously known only for C^∞ smooth coefficients. Such a reduction of regularity turns in a decrease of regularity of the fundamental solution: indeed, by the results of Hörmander the fundamental solution is smooth if the coefficients are smooth, but is not even C^1 when we consider Hölder continu-

ous coefficients (as we have seen in Definition 3.1.3). In particular, this low regularity setting requires the use of the tools developed in the last years in the field of Geometric Measure Theory on Carnot groups (see [110] for an overview of this topic). We point out that the mean value formula is based on the fundamental solution of the adjoint of (3.1.4). In order to rely on the results of [26] the authors had to require the strict condition

$$b_i = 2\sum_{j=1}^{m_1} X_i a_{ij}, \qquad c = \sum_{i,j=1}^{m_1} X_i X_j a_{ij}.$$

The proof of the existence of the fundamental solution with low order terms allows us to remove such assumption. We emphasize that our approach is different from that used in [26]: there, the authors firstly proved the existence of the fundamental solution for the evolution operator

$$\mathscr{H}_{\mathbb{G}} = \sum_{i,j=1}^{m_1} a_{ij}(x,t) X_i X_j - \partial_t, \qquad (x,t) \in \mathbb{R}^{N+1},$$

alongside suitable long time estimates. Then, by t-saturation and approximation arguments, the authors proved the existence of a fundamental solution for $\mathscr{L}_{\mathbb{G}}$. We also quote the article [34], were Bramanti, Brandolini, Lanconelli and Uguzzoni extended some of the results of [26], to more general evolution operator of the form

$$\mathscr{H}_{\mathbb{G}} = \sum_{i,j=1}^{m_1} a_{ij}(x,t) X_i X_j + \sum_{i=1}^{m_1} b_i(x,t) X_i + c(x,t) - \partial_t, \qquad (x,t) \in \Omega \times (T_1, T_2), \tag{3.1.8}$$

without any underlying Carnot group structure. They proved the existence of a fundamental solution for (3.1.8), leaving as an open problem the existence of a fundamental solution to the stationary operator corresponding to (3.1.8), due to the lack of suitable long time estimates.

Let's briefly discuss about the maximum propagation principle and the Harnack inequality. We recall that a maximum principle for (3.1.1) with C^∞ smooth coefficients was first obtained by Bony in [30] with a barrier argument, and then extended by Amano in [5] for C^1 coefficients. We also quote Bonfiglioli and Uguzzoni [29] where, assumed the existence of a Picone function, a maximum principle for (3.1.1) with continuous coefficients is proved. We point out that Theorem 3.1.5 reduces the regularity assumption of [5], and does not require the existence of a Picone function.

Since Bony's work [30], several statements of the Harnack inequality for operators with smooth coefficients are available. Concerning operators with Hölder continuous coefficients, we quote Bramanti, Brandolini, Lanconelli and Uguzzoni [34], were they proved a Harnack inequality for positive solutions to (3.1.8) with $c = 0$, adapting the method introduced by Krylov and Safonov in [72]. Clearly, this result implies a Harnack inequality for the stationary operator (3.1.1) with $c = 0$. Our proof of the Harnack inequality for divergence form operators relies on a different technique, which, however, requires the existence of the fundamental solution for $\mathscr{L}_{\mathbb{G}}^*$ defined in (3.1.5).

This chapter is organized as follows: in Section 3.2 we recall some properties of Carnot groups; in Section 3.3 we define the tools we need to construct the fundamental solution via parametrix method; in Section 3.4 we give the proof of Theorem 3.1.4; in Section 3.5, as an application of Theorem 3.1.4, we give the proof of Theorem 3.1.5 and Theorem 3.1.6.

We give some notational remarks. In this chapter we consistently refrain from restating the phrase 'for $i, j = 1, \dots, m_1$', as it is always implied. Moreover, for every measurable set $A \subset \mathbb{R}^N$, $|A|$ will denote its N-dimensional Lebesgue measure, $\mathcal{H}^{N-1}$ will denote the $N-1$ Hausdorff measure, ν the Euclidean inward normal unit vector and $\langle \cdot, \cdot \rangle$ the usual inner product.

3.2 Preliminaries on Carnot group

In this section we recall some properties concerning Carnot groups. For a more detailed exposition we refer to the monograph [27].

Let $\mathbb{G} = \left(\mathbb{R}^N, \circ, \delta_r\right)$ be the Carnot group induced by the smooth vector fields $\{X_i\}_{i=1}^{m_1}$. We recall that a Carnot group is a homogeneous Lie group, namely a group with a smooth composition law $\circ$ and a dilation law $\{\delta_r\}_{r>0}$ that is an automorphism of the group

$$\delta_r(x \circ y) = \left(\delta_r(x)\right) \circ \left(\delta_r(y)\right), \qquad \text{for every } x, y \in \mathbb{R}^N,\ r > 0.$$

By assumptions **(H1)**, **(H2)** and **(H3)**, $\mathfrak{g}$ is the Lie algebra of the group $\mathbb{G}$. Moreover, **(H3)** induces a stratification on $\mathfrak{g}$

$$\mathfrak{g} = V_1 \oplus \ldots \oplus V_\kappa, \tag{3.2.1}$$

where $V_1 = \text{span}\{X_i\}_{i=1}^m$, $V_{k+1} = \text{span}\Big\{[X,Y], X \in V_1, Y \in V_k\Big\}$ and $\Big\{[X,Y], X \in V_1, Y \in V_\kappa\Big\} = \emptyset$, with $[X,Y] = XY - YX$. We observe that the stratification (3.2.1) implies that the dilations $\{\delta_r\}_{r>0}$ take the following form (see Remark 1.4.2 in [27])

$$\delta_r(x) = \left(rx^{(1)}, r^2x^{(2)}, \dots, r^\kappa x^{(\nu)}\right), \qquad x^{(k)} \in \mathbb{R}^{m_k},\ k = 1, \dots, \kappa, \tag{3.2.2}$$

where m_k denotes the dimension of V_k, for every $k = 1, \dots, \kappa$. In particular, the number

$$Q := \sum_{k=1}^{\kappa} k\, m_k,$$

is usually called homogeneous dimension of $\mathbb{G}$. A group $\mathbb{G}$ with the properties above is often referred to as Carnot group of step κ and m_1 generators.

Let us introduce the distance we will use in the sequel. We call *homogeneous norm* on $\mathbb{G}$ every continuous function $\|\cdot\|_{\mathbb{G}} : \mathbb{R}^N \to [0, +\infty)$ that is positive whenever $x \neq 0$, and δ_r-homogeneous of degree one

$$\|\delta_r(x)\|_{\mathbb{G}} = r\|x\|_{\mathbb{G}}, \qquad \text{for every } r > 0,\ x \in \mathbb{R}^N.$$

The corresponding distance is

$$d(x, y) = \|y^{-1} \circ x\|_{\mathbb{G}}, \qquad \forall x, y \in \mathbb{R}^N. \tag{3.2.3}$$

We recall two fundamental properties:

1. there exists $k_1 \geq 1$ such that

$$d(x, y) \leq k_1\Big(d(x, \xi) + d(\xi, y)\Big), \qquad \text{for every } x, y, \xi \in \mathbb{R}^N; \tag{3.2.4}$$

2. there exists $k_2 > 0$ such that

$$d(x, y) \leq k_2\, d(y, x), \qquad \text{for every } x, y \in \mathbb{R}^N. \tag{3.2.5}$$

We notice that, if $k_1 = 1$, (3.2.4) is the classical triangular inequality; for such reason this property is usual referred to as *pseudo-triangular inequality*. From now on in this chapter k_1 and k_2 will be uniquely associated with these constants.

We are in position now to give the definition of a suitable Hölder space for $\mathbb{G}$.

Definition 3.2.7 (Hölder continuous function in $\mathbb{G}$)**.** *Let $\alpha \in (0, 1]$ and $\Omega \subseteq \mathbb{G}$. We say that $u : \Omega \to \mathbb{R}$ is an α-Hölder continuous function with exponent α in Ω (in short $u \in C^{\alpha}_{\mathbb{G}}(\Omega)$) if there exists a positive constant $c > 0$ such that*

$$|u(x) - u(y)| \leq c\, d(x, y)^{\alpha}, \qquad \forall x, y \in \Omega.$$

By the equivalence of every two homogeneous norms in $\mathbb{R}^N$, it will be convenient for us to use a regular one: we denote as $|\cdot|$ the usual Euclidean norm, then we set

$$\|x\|_{\mathbb{G}} := \left(\sum_{k=1}^{\kappa} |x^{(k)}|^{\frac{2\kappa!}{k}} \right)^{\frac{1}{2\kappa!}}, \qquad x^{(k)} \in \mathbb{R}^{m_k}.$$

We observe that the corresponding distance d is smooth out of the origin, and this is useful when we need to integrate d-radial functions over d-radially symmetric domains.

Proposition 3.2.8. *Let $w \in \mathbb{R}^N$, denote as $B_r(w)$ the d-ball of radius r centred at w, and let $f \in L^1(B_r(w))$. If there exists some function φ such that $f(x) = \varphi(d(x,w))$, then we have*

$$\int_{B_r(w)} f(x)dx = Q\omega_Q \int_0^r s^{Q-1}\varphi(s)ds,$$

where ω_Q denotes the Lebesgue measure of the set $B_1(0)$. In particular, if $f(x) = d(x,w)^{\beta-Q}$ for some $\beta > 0$, then

$$\int_{B_r(w)} d(x,w)^{\beta-Q}dx = \frac{Q\omega_Q}{\beta}r^\beta. \tag{3.2.6}$$

Proof. First we note that the boundary $\partial B_r(w)$ is a smooth $N-1$ manifold, as d is smooth. Indeed, by Sard's lemma this holds true for almost every $r > 0$. The assertion then follows for every $r > 0$, since $\partial B_r(w)$ is diffeomorphic to $\partial B_1(w)$ via the dilation δ_r.
As second, we recall that the Jacobian matrix of the left-translation on $\mathbb{G}$ has a lower triangular form with entries equal to one in the diagonal (see Theorem 1.3.15 in [27]).
Bearing these observations in mind, the classical coarea formula gives us

$$\int_{B_r(w)} \varphi(d(x,w))dx = \int_{B_r(0)} \varphi(d(z))dz = \int_0^r \varphi(s)\left(\int_{d(y)=s} \frac{1}{|\nabla d(y)|}d\mathcal{H}^{N-1}(y)\right)ds.$$

Taking $\varphi = 1$, by (3.2.2) we get

$$\omega_Q\, r^Q = \int_0^r \left(\int_{d(y)=s} \frac{1}{|\nabla d(y)|}d\mathcal{H}^{N-1}(y)\right)ds.$$

Then, by differentiating this identity w.r.t. r we obtain

$$\int_{d(y)=s} \frac{1}{|\nabla d(y)|}d\mathcal{H}^{N-1}(y) = Q\,\omega_Q s^{Q-1}.$$

This proves the first part of the thesis. The second one follows by choosing $\varphi(s) = s^{\beta-Q}$. □

Remark 3.2.9. *As a consequence, for every $w \in \mathbb{R}^N$ the function $x \mapsto d(x,w)^\vartheta$ belongs to $L^1_{\text{loc}}(\mathbb{R}^N)$ if and only if $\vartheta > -Q$. This condition is the counterpart of the integrability condition for radial functions in the Euclidean setting, where we have $Q = N$.*

We conclude this section with the following *Stratified Mean Value Theorem*.

Theorem 3.2.10 (Theorem 20.3.1 in [27])**.** *There exist constants $c > 0$ and $b \geq 1$ (depending* *only on the Carnot group* $\mathbb{G}$*) such that*

$$|f(y)-f(x)| \leq c\, d(x,y) \sup_{\substack{z:d(x,z)\leq b\, d(x,y)\\ i=1,\ldots,m_1}} |X_i f(z)|\,,$$

for all real valued function $f \in C^1(\mathbb{R}^N)$ *and every* $x, y \in \mathbb{R}^N$*.*

3.3 Parametrix method

In this section we describe Levi's method for subelliptic operators, and we define some tools that we use to make it work. As said in the introduction, for every $\xi \in \mathbb{R}^N$ we denote by $\Gamma_{A_\xi} = \Gamma_{A_\xi}(w, w')$ the fundamental solution of the following constant coefficients operator

$$\mathscr{L}_{\mathbb{G},A_\xi} = \sum_{i,j=1}^{m_1} a_{ij}(\xi) X_i X_j,$$

obtained by computing the coefficients of the principal part of (3.1.1) at the point ξ. We call parametrix the function $Z(x,\xi) := \Gamma_{A_\xi}(x,\xi)$. We seek for a fundamental solution Γ by using the parametrix

$$\Gamma(x,\xi) = Z(x,\xi) + J(x,\xi),$$

where J is a *correction term* that we assume of the following form

$$J(x,\xi) = \int_\Omega Z(x,y) G(y,\xi) dy, \tag{3.3.1}$$

for some unknown function G. We seek a G such that Γ is a classical solution to

$$0 = \mathscr{L}_{\mathbb{G}}\Gamma(x,\xi) = \mathscr{L}_{\mathbb{G}} Z(x,\xi) + \mathscr{L}_{\mathbb{G}} J(x,\xi), \qquad \text{for every } x \neq \xi. \tag{3.3.2}$$

We look for a G such that we can differentiate J as follows

$$\mathscr{L}_{\mathbb{G}} J(x,\xi) = \int_\Omega \mathscr{L}_{\mathbb{G}} Z(x,y) G(y,\xi) dy - G(x,\xi). \tag{3.3.3}$$

Then, combining (3.3.2) and (3.3.3) we get the following equation for G

$$G(x,\xi) = \mathscr{L}_{\mathbb{G}} Z(x,\xi) + \int_\Omega \mathscr{L}_{\mathbb{G}} Z(x,y) G(y,\xi) dy, \tag{3.3.4}$$

which is a Fredholm integral equation of the second kind. If Ω is such that

$$\sup_{x\in\Omega} \int_\Omega |\mathscr{L}_{\mathbb{G}} Z(x,\xi)| d\xi < 1,$$

we solve (3.3.4) by means of successive approximation method, obtaining

$$G(x,\xi) = \sum_{k=1}^{\infty} (\mathscr{L}_{\mathbb{G}} Z)_k(x,\xi), \tag{3.3.5}$$

where $(\mathscr{L}_{\mathbb{G}} Z)_1 := \mathscr{L}_{\mathbb{G}} Z$ and

$$(\mathscr{L}_{\mathbb{G}} Z)_{k+1}(x,\xi) := \int_\Omega (\mathscr{L}_{\mathbb{G}} Z)_1(x,y)(\mathscr{L}_{\mathbb{G}} Z)_k(y,\xi)dy.$$

Therefore, our aim is to prove that the function G in (3.3.5) solves (3.3.4) and that (3.3.3) holds. Having shown this, it is straightforward to prove that Γ is a fundamental solution. To this end, we need to study the regularity properties of (3.3.1), which in particular depend on the regularity of the parametrix Z. With the purpose of taking the Lie derivatives under the integral sign, we need accurate estimates for the parametrix, that we recall to be a fundamental solution for an operator with constant coefficients.

Theorem 3.3.11 (Theorem 2.5 in [25]). *Let $A=(a_{ij})_{i,j=1}^{m_1}$ be a constant coefficient matrix in the class $\boldsymbol{M}_\lambda$. Then, there exists a fundamental solution $\Gamma_A \in C^\infty\left((\mathbb{R}^N \times \mathbb{R}^N) \setminus D_{\mathbb{R}^N}\right)$ for*

$$\mathscr{L}_{\mathbb{G},A} = \sum_{j=1}^{m_1} a_{ij} X_i X_j.$$

Moreover, for every integer p there exist two positive constants $c_\lambda, c_{\lambda,p}$ such that for every $i_1, \dots, i_p \in \{1, \dots, m_1\}$ and $x \neq \xi$ we have

$$\frac{1}{c_\lambda}\, d(x,\xi)^{2-Q} \le \Gamma_A(x,\xi) \le c_\lambda\, d(x,\xi)^{2-Q}, \tag{3.3.6}$$

$$\left| X_{i_1} \dots X_{i_p} \Gamma_A(x,\xi) \right| \le c_{\lambda,p}\, d(x,\xi)^{2-Q-p}. \tag{3.3.7}$$

In particular, there exists a homogeneous distance d_A on $\mathbb{G}$ such that $\Gamma_A = d_A^{2-Q}$.
Lastly, if A_1 and A_2 are constant matrices belonging to the class $\boldsymbol{M}_\lambda$, the following bounds hold for the fundamental solutions of the corresponding operators $\mathscr{L}_{\mathbb{G},A_1}$ and $\mathscr{L}_{\mathbb{G},A_2}$

$$\left| X_{i_1} \dots X_{i_p} \Gamma_{A_1}(x,\xi) - X_{i_1} \dots X_{i_p} \Gamma_{A_2}(x,\xi) \right| \le c_{\lambda,p}\, \|A_1 - A_2\|^{\frac{1}{\kappa}}\, d(x,\xi)^{2-Q-p}, \tag{3.3.8}$$

for some $c_{\lambda,p} > 0$. Here $\|A\| = \max_{\xi=1} |A\,\xi|$, and we recall that κ denotes the step of $\mathbb{G}$ and Q its homogeneous dimension.

Note that, in view of (3.3.6) and (3.3.7), we have to deal with convolutions of homogeneous functions in order to guarantee the convergence of (3.3.5) and the validity of (3.3.3). To this end, we give four auxiliary results concerning integrability of singular functions.

Proposition 3.3.12. *Let Ω_1 be a bounded subset of $\mathbb{R}^N$. Let $g, h \in C\left(\left(\overline{\Omega}_1 \times \overline{\Omega}_1\right) \setminus D_{\overline{\Omega}_1}\right)$ such that there exist $\beta, \gamma, c > 0$ such that*

$$|g(x,\xi)| \le c\, d(x,\xi)^{\beta-Q}, \qquad |h(x,\xi)| \le c\, d(x,\xi)^{\gamma-Q}, \tag{3.3.9}$$

for every $x, \xi \in \overline{\Omega}_1$, $x \neq \xi$. Then the function

$$f(x,\xi) := \int_{\Omega_1} g(x,y)h(y,\xi)dy, \qquad (x,\xi) \in \left(\overline{\Omega}_1 \times \overline{\Omega}_1\right) \setminus D_{\overline{\Omega}_1}, \tag{3.3.10}$$

is continuous in $\left(\overline{\Omega}_1 \times \overline{\Omega}_1\right) \setminus D_{\overline{\Omega}_1}$. Moreover, f can be extended to a continuous function in $\left(\overline{\Omega}_1 \times \overline{\Omega}_1\right)$ if $\beta + \gamma > Q$, while if $\beta + \gamma < Q$ there exists $C > 0$ such that

$$|f(x,\xi)| \le C\, d(x,\xi)^{\beta+\gamma-Q}. \tag{3.3.11}$$

Proof. We preliminarily remark that, by (3.3.9), $y \mapsto g(x,y)$ and $y \mapsto h(y,\xi)$ belong to $L^1(\Omega_1)$ for every $x, \xi \in \overline{\Omega}_1$, as follows from Remark 3.2.9. The function (3.3.10) is well defined then, and the continuity w.r.t. one variable when the other one is fixed is straightforward. In order to prove the joint continuity, we fix $(x,\xi), (x_0,\xi_0) \in \left(\overline{\Omega}_1 \times \overline{\Omega}_1\right) \setminus D_{\overline{\Omega}_1}$ and define $a := d(x_0,\xi_0)$. Then we recall (3.2.4) and (3.2.5), and we set $k := \max\left\{k_1^2, k_1k_2\right\}$. We take $\varepsilon > 0$, $\bar{b} > 0$ such that $0 < k^{-1} - 2\bar{b} < 1$, and $\eta > 0$ such that $\eta < \bar{b}\, a$. We have

$$\begin{aligned} f(x,\xi) - f(x_0,\xi_0) = &\int_{\Omega_1} \left[g(x,y) - g(x_0,y)\right] h(y,\xi)dy \\ &+ \int_{\Omega_1} \left[h(y,\xi) - h(y,\xi_0)\right] g(x_0,y)dy =: I_1 + I_2. \end{aligned} \tag{3.3.12}$$

Let us focus on I_1, that we split as

$$\begin{aligned} I_1 = &\int_{\Omega_1 \cap B_\eta(x_0)} \left[g(x,y) - g(x_0,y)\right] h(y,\xi)dy \\ &+ \int_{\Omega_1 \setminus \overline{B_\eta(x_0)}} \left[g(x,y) - g(x_0,y)\right] h(y,\xi)dy =: J_1 + J_2. \end{aligned}$$

We choose without loss of generality x and ξ such that $d(x,x_0) < \eta$, $d(\xi,\xi_0) < \bar{b}\, a$. Then, for all $y \in B_\eta(x_0)$, we have

$$d(x_0,\xi_0) \le k_1\Big(d(x_0,y) + d(y,\xi_0)\Big) \le k_1\Big(k_2 d(y,x_0) + k_1 d(y,\xi) + k_1 d(\xi,\xi_0)\Big).$$

From this estimate it follows $d(y,\xi) > a\left(k^{-1} - 2\bar{b}\right)$, which implies that the function h is continuous and bounded in the set $\Omega_1 \cap B_\eta(x_0)$. Moreover

$$d(y,x) \le k_1(d(y,x_0) + k_2 d(x,x_0)) \le \overline{k}\left(d(y,x_0) + d(x,x_0)\right) < 2\overline{k}\,\eta,$$

where $\overline{k} := \max\left\{k_1, k_1 k_2\right\} \geq 1$. Therefore, by (3.3.9) there exists $c_1 > 0$ such that

$$\begin{aligned}
|J_1| &\leq c^2 \int_{\Omega_1 \cap B_\eta(x_0)} d(y,\xi)^{\gamma-Q} d(x,y)^{\beta-Q} dy \\
&\quad + c^2 \int_{\Omega_1 \cap B_\eta(x_0)} d(y,\xi)^{\gamma-Q} d(x_0,y)^{\beta-Q} dy \\
&\leq c_1 \left(a\left(\frac{1}{k} - 2\bar{b}\right)\right)^{\gamma-Q} \left(\int_{B_{2\overline{k}\eta}(x)} d(x,y)^{\beta-Q} dy + \int_{B_\eta(x_0)} d(x_0,y)^{\beta-Q} dy\right) \\
&\qquad\qquad = c_1 \left(a\left(\frac{1}{k} - 2\bar{b}\right)\right)^{\gamma-Q} k_2^{\beta-Q} \frac{\omega_Q^\beta Q^\beta}{\beta} \left[(2\overline{k}\eta)^\beta + \eta^\beta\right],
\end{aligned}$$

as follows from (3.2.6). We conclude that $|J_1| \leq \varepsilon$ for η sufficiently small.
We examine J_2 now: let us choose $R > 0$ such that for all $\xi \in \Omega_1$ we have $\overline{\Omega}_1 \subset B_R(\xi)$, and set $\sigma := \frac{c\omega_Q Q R^\gamma}{\gamma}$. Since g is uniformly continuous in $\left(\overline{\Omega}_1 \cap \overline{B_{\frac{\eta}{2}}(x_0)}\right) \times \left(\overline{\Omega}_1 \setminus B_\eta(x_0)\right)$, there exists a positive constant c_η and $\delta_1 \in \left(0, \frac{\eta}{2}\right)$ such that

$$d(x,x_0) < \delta_1 \Longrightarrow |g(x,y) - g(x_0,y)| < \frac{\varepsilon}{\sigma},$$

for every $y \in \left(\overline{\Omega}_1 \setminus B_\eta(x_0)\right)$. As a consequence, by (3.2.6) we find

$$|J_2| \leq \frac{\varepsilon}{\sigma} c \int_{\Omega_1 \setminus \overline{B_\eta(x_0)}} d(y,\xi)^{\gamma-Q} dy \leq \frac{\varepsilon}{\sigma} c \int_{B_R(\xi)} d(y,\xi)^{\gamma-Q} dy = \varepsilon.$$

We estimate I_2 by the same argument, so we conclude that $f \in C\left(\left(\overline{\Omega}_1 \times \overline{\Omega}_1\right) \setminus D_{\overline{\Omega}_1}\right)$.

Let us show that $\beta + \gamma > Q$ yields $f \in C\left(\overline{\Omega}_1 \times \overline{\Omega}_1\right)$. We note that if $\beta + \gamma > Q$, then by Proposition 3.2.9 the integral (3.3.10) converges in $\overline{\Omega}_1 \times \overline{\Omega}_1$. Therefore we just need to prove that, when $\xi_0 = x_0$, I_1 and I_2 in (3.3.12) are bounded by ε when $d(x,x_0)$ and $d(\xi,x_0)$ are sufficiently small. By (3.3.9) we get

$$\begin{aligned}
|I_1| &\leq c^2 \int_{\Omega_1} d(x,y)^{\beta-Q} d(y,\xi)^{\gamma-Q} dy \\
&\quad + c^2 \int_{\Omega_1} d(x_0,y)^{\beta-Q} d(y,\xi)^{\gamma-Q} dy := J_1 + J_2.
\end{aligned} \tag{3.3.13}$$

Since Ω_1 is bounded, there exists $j_0 > 0$ such that

$$\Omega_{z,\xi} := \left\{w \in \mathbb{R}^N : \xi \circ \delta_{d(z,\xi)} w \in \Omega_1\right\} \subset B_{j_0}(0), \qquad \text{for every } z, \xi \in \Omega_1.$$

Let us choose $j \geq j_0$ and $(x,\xi) \in \left(\overline{\Omega}_1 \times \overline{\Omega}_1\right)$ such that $d(x_0,\xi), d(x,\xi) \leq \frac{1}{jj_0}$, $x \neq \xi$. In the first integral J_1 in (3.3.13) we set $u := \delta_{d(x,\xi)^{-1}}(\xi^{-1} \circ x)$, $w := \delta_{d(x,\xi)^{-1}}(\xi^{-1} \circ y)$ and $r_x := \frac{1}{j_0 d(x,\xi)}$, while in the second integral J_2 we make the same change of variable with x_0

instead of x. We obtain

$$
\begin{aligned}
|I_1| \leq \frac{c^2}{(jj_0)^{\beta+\gamma-Q}} \Bigg(& \int_{B_{r_x}(0)} d(w)^{\gamma-Q} d(u,w)^{\beta-Q} dw \\
& + \int_{B_{r_{x_0}}(0)} d(w)^{\gamma-Q} d(u,w)^{\beta-Q} dw \Bigg).
\end{aligned} \tag{3.3.14}
$$

We estimate the first integral in three steps. By (3.2.4) and (3.2.6), since $d(u) = 1$ we get

$$
\begin{aligned}
&\int_{B_{r_x}(0)\setminus B_2(0)} d(w)^{\gamma-Q} d(u,w)^{\beta-Q} dw \\
&\qquad \leq c_1 \int_{B_{r_x}(0)\setminus B_2(0)} d(w)^{\gamma-Q} d(w)^{\beta-Q} dw \\
&\qquad\qquad \leq c_2 \int_{B_{r_x}(0)} d(w)^{\beta+\gamma-2Q} dw \leq \frac{c_2\, Q\, \omega_Q}{j^{\beta+\gamma-Q}(\beta+\gamma-Q)},
\end{aligned}
$$

for some $c_1, c_2 > 0$. By the same argument we get

$$
\begin{aligned}
&\int_{B_2(0)\setminus B_{\frac{1}{2}}(0)} d(w)^{\gamma-Q} d(u,w)^{\beta-Q} dw \leq 2^{\gamma-Q} \int_{B_2(0)} d(u,w)^{\beta-Q} dw \\
&\leq c_3 \int_{B_2(0)} (1 + d(w))^{\beta-Q} dw \leq c_4 \int_{B_1(0)} 2^{\beta-Q} dw + 2^{\beta-Q} c_4 \int_{B_2(0)\setminus B_1(0)} d(w)^{\beta-Q} dw = c_5,
\end{aligned}
$$

for some $c_3, c_4, c_5 > 0$. Finally, applying the same reasoning yields

$$
\int_{B_{\frac{1}{2}}(0)} d(w)^{\gamma-Q} d(u,w)^{\beta-Q} dw \leq c_6 \int_{B_{\frac{1}{2}}(0)} d(w)^{\gamma-Q} dw = c_7.
$$

for some $c_6, c_7 > 0$. Therefore, since the same arguments hold for the second integral in (3.3.14), we infer the existence of some $\widetilde{j}, C > 0$ such that

$$
|I_1| \leq \frac{c^2}{(jj_0)^{\beta+\gamma-Q}} \left(C + \frac{C}{j^{\beta+\gamma-Q} d(x,y)^{\beta+\gamma-Q}} \right) \leq \varepsilon,
$$

if $j \geq \widetilde{j}$. Similarly we estimate I_2, concluding that $f \in C\left(\left(\overline{\Omega}_1 \times \overline{\Omega}_1\right) \setminus D_{\overline{\Omega}_1}\right)$.

It remains to establish (3.3.11). From (3.3.9) we obtain

$$
|f(x,\xi)| \leq c^2 \int_{\Omega_1} d(x,y)^{\beta-Q} d(y,\xi)^{\gamma-Q} dy := I.
$$

By setting $z = \xi^{-1} \circ y$ we get

$$
I = c^2 \int_{\Omega_2} \frac{1}{d(z)^{Q-\gamma} d(x, \xi \circ z)^{Q-\beta}} dz,
$$

where we have set $\Omega_2 := \Big\{z \in \mathbb{R}^N : \xi \circ z \in \Omega_1\Big\}$.
The change of variable $z = \delta_{d(x,\xi)}(w)$ yields

$$I = c^2 \int_{\Omega_3} \frac{d(x,\xi)^Q}{d(x,\xi)^{2Q-\gamma-\beta} d(w)^{Q-\gamma} d\left(\delta_{d(x,\xi)^{-1}}\left(\xi^{-1} \circ x\right), w\right)^{Q-\beta}} dw,$$

where $\Omega_3 := \Big\{w \in \mathbb{R}^N : \delta_{d(x,\xi)}(w) \in \Omega_2\Big\}$.
Having $\Big\|\delta_{d(x,\xi)^{-1}}\big(\xi^{-1} \circ x\big)\Big\|_{\mathbb{G}} = 1$ and $\beta + \gamma < Q$, we deduce the existence of a positive constant c_1 such that

$$I = c^2\, d(x,\xi)^{\gamma+\beta-Q} \int_{\Omega_3} \frac{1}{d(w)^{Q-\gamma} d\left(\delta_{d(x,\xi)^{-1}}\left(\xi^{-1} \circ x\right), w\right)^{Q-\beta}} dw \le c_1 d(x,\xi)^{\gamma+\beta-Q},$$

since Proposition 3.2.9 implies that the last integral is finite. This concludes the proof of (3.3.11), and hence of the proposition. □

If the functions g and h are more regular, the function f defined in (3.3.10) is more regular as well, as the propositions below show.

Proposition 3.3.13. *Suppose, in addition to the hypothesis of Proposition 3.3.12, that* $\gamma > 1$, $X_i g(\cdot,\xi) \in C\left(\Omega_1 \setminus \{\xi\}\right)$ *and* $|X_i g(x,\xi)| \le c\, d(x,\xi)^{\sigma-Q}$ *for some* $\sigma, c > 0$. *Then* $X_i f(\cdot,\xi) \in C\left(\Omega_1 \setminus \{\xi\}\right)$ *and*

$$X_i f(x,\xi) = \int_{\Omega_1} X_i g(x,y) h(y,\xi) dy,$$

for every $\xi \in \Omega_1$ *and* $x \in \Omega_1 \setminus \{\xi\}$.

Proof. By Proposition (3.3.12) the function

$$v(x) := \int_{\Omega_1} X_i g(x,y) h(y,\xi) dy, \qquad x \in \Omega_1 \setminus \{\xi\},$$

belongs to $C\big(\Omega_1 \setminus \{\xi\}\big)$ for every $\xi \in \Omega_1$. We claim that $v = X_i f$.
Let $\eta \in C^1(\mathbb{R})$ such that $0 \le \eta \le 1$, $0 \le \eta' \le 2$, $\eta(t) = 0$ if $t \le 1$ and $\eta(t) = 1$ if $t \ge 2$. For every $\varepsilon > 0$ we define

$$w_\varepsilon(x) := \int_{\Omega_1} g(x,y) \eta\left(\frac{d(x,y)}{\varepsilon}\right) h(y,\xi) dy, \qquad x \in \Omega_1 \setminus \{\xi\}.$$

Since $\eta\left(\frac{d(x,y)}{\varepsilon}\right) = 0$ if $d(x,y) \le \varepsilon$, we have $w_\varepsilon \in C^1_{\mathbb{G}}(\Omega_1)$. We pick some $\varepsilon > 0$ such that $d(y,\xi)^{\beta-Q}$ is bounded for every y such that $d(x,y) < 2\varepsilon$. Then, by the assumption on g and

$X_i g$, we get

$$
\begin{aligned}
&|v(x) - X_i w_\varepsilon(x)| \\
&\qquad \le \int_{\Omega_1} \left| X_i \left[\left((1 - \eta\left(\frac{d(x,y)}{\varepsilon}\right)\right) g(x,y) \right] h(y,\xi) \right| dy \\
&\qquad = \int_{d(x,y)\le 2\varepsilon} \left| X_i \left[\left((1 - \eta\left(\frac{d(x,y)}{\varepsilon}\right)\right) g(x,y) \right] h(y,\xi) \right| dy \\
&\qquad \le c_1 \int_{d(x,y)\le 2\varepsilon} \left[\frac{1}{\varepsilon} |g(x,y)| + |X_i g(x,y)| \right] |h(y,\xi)| dy \\
&\qquad \le c_2 \int_{d(x,y)\le 2\varepsilon} \left(\frac{d(x,y)^{\gamma - Q}}{\varepsilon} + d(x,y)^{\sigma - Q} \right) d(y,\xi)^{\beta - Q} dy \\
&\qquad\qquad = c_3 \Big(\varepsilon^\sigma + \varepsilon^{\gamma - 1} \Big),
\end{aligned}
$$

for some $c_1, c_2, c_3 > 0$. Thus we have uniform convergence on every compact set, and since by the same argument it is easy to see that w_ε converges uniformly to f, the thesis follows. □

Proposition 3.3.14. *Suppose, in addition to the hypothesis of Proposition 3.3.12, that there exists $\eta \in (0,1]$, $c > 0$ such that*

$$
|g(x_1,\xi) - g(x_2,\xi)| \le c\, d(x_1,x_2)^\eta \Big(d(x_1,\xi)^{-Q} + d(x_2,\xi)^{-Q} \Big), \tag{3.3.15}
$$

for every $x_1, x_2, \xi \in \Omega_1$ with $d(\xi, x_1) \ge 2\, d(x_1,x_2) > 0$. Then, set $\mu := \min\{\beta,\gamma\}$, for some $C > 0$ we have

$$
|f(x_1,\xi) - f(x_2,\xi)| \le C\; d(x_1,x_2)^\sigma \Big(d(x_1,\xi)^{\mu - Q} + d(x_2,\xi)^{\mu - Q} \Big),
$$

for every $x_1, x_2 \in \Omega_1$ and $\xi \in \Omega_1 \setminus \{x_1, x_2\}$, with $\sigma \in (0,\eta)$ and $\sigma \le \mu$.

Proof. Let us set $a := d(x_1,x_2) > 0$, and fix $\xi \in \Omega_1 \setminus \{x_1,x_2\}$. We have

$$
\begin{aligned}
f(x_1,\xi) - f(x_2,\xi) &= \int_{\Omega_1} [g(x_1,y) - g(x_2,y)] h(y,\xi) dy \\
&= \int_{\Omega_1 \cap B_{2a}(x_1)} [g(x_1,y) - g(x_2,y)] h(y,\xi) dy \\
&\quad + \int_{\Omega_1 \setminus \overline{B_{2a}(x_1)}} [g(x_1,y) - g(x_2,y)] h(y,\xi) dy.
\end{aligned}
$$

By (3.3.9) and (3.3.15) we get

$$
|f(x_1,\xi) - f(x_2,\xi)| \le c_1 \Big(d(x_1,x_2)^\eta I_1 + I_2 \Big), \tag{3.3.16}
$$

for some $c_1 > 0$, where

$$I_1 := \int_{\Omega_1 \setminus \overline{B_{2a}(x_1)}} \left[d(x_1, y)^{-Q} + d(x_2, y)^{-Q} \right] |h(y, \xi)| dy,$$
$$I_2 := \int_{\Omega_1 \cap B_{2a}(x_1)} \left[d(x_1, y)^{\beta - Q} + d(x_2, y)^{\beta - Q} \right] |h(y, \xi)| dy.$$

We first focus on I_1. Preliminarily, if we set $k := k_1 k_2$, we observe that

$$d(\xi, y) < \frac{1}{2k} d(x_1, \xi) \Longrightarrow d(x_1, y) > \frac{1}{2k_1} d(x_1, \xi). \tag{3.3.17}$$

Indeed by (3.2.4) and (3.2.5) we have

$$d(x_1, \xi) \le k_1 d(x_1, y) + k_1 k_2 d(\xi, y) \le k_1 d(x_1, y) + \frac{1}{2} d(x_1, \xi).$$

We employ this relation to bound I_1. If we set $R := \operatorname{diam}(\Omega_1)$, by (3.3.9) and (3.3.17) we get

$$\begin{aligned}
&\int_{\Omega_1 \setminus \overline{B_{2a}(x_1)}} d(x_1, y)^{-Q} |h(y, \xi)| dy \\
&\qquad = \int_{\Omega_1 \setminus \overline{B_{2a}(x_1)},\ d(\xi, y) \ge \frac{1}{2k} d(x_1, \xi)} d(x_1, y)^{-Q} |h(y, \xi)| dy \\
&\qquad + \int_{\Omega_1 \setminus \overline{B_{2a}(x_1)},\ d(\xi, y) < \frac{1}{2k} d(x_1, \xi)} d(x_1, y)^{-Q} |h(y, \xi)| dy \\
&\qquad \le \ c_1 \, d(x_1, \xi)^{\gamma - Q} \int_{2a < d(y, x_1) < R} d(y, x_1)^{-Q} dy \\
&\qquad + c_1 \, d(x_1, \xi)^{-Q} \int_{d(\xi, y) < \frac{1}{2k} d(x_1, \xi)} d(y, \xi)^{\gamma - Q} dy \\
&\qquad = \ c_2 \, d(x_1, \xi)^{\gamma - Q} \left[\int_{2a}^{R} \frac{ds}{s} + c_3 \right] \\
&\qquad\qquad = c_2 \, d(x_1, \xi)^{\gamma - Q} \left(\ln \left(\frac{R}{2a} \right) + c_3 \right),
\end{aligned}$$

for some $c_1, c_2, c_3 > 0$. Similarly, by using the same argument as above, we prove that

$$\int_{\Omega_1 \setminus \overline{B_{2a}(x_1)}} d(x_2, y)^{-Q} |h(y, \xi)| dy \le c_4 \, d(x_2, \xi)^{\gamma - Q} \left(\ln \left(\frac{R}{2a} \right) + c_5 \right),$$

for some $c_4, c_5 > 0$. Therefore, recalled that $a = d(x_1, x_2)$ and $\mu \le \gamma$, we infer that for every $\varepsilon > 0$ there exists $C_1 > 0$ such that

$$I_1 \le C_1 \, d(x_1, x_2)^{-\varepsilon} \left(d(x_1, \xi)^{\mu - Q} + d(x_2, \xi)^{\mu - Q} \right). \tag{3.3.18}$$

The integral I_2 it is similarly treated. Indeed as follows by (3.3.17), $d(y, x_1) < 2a$ and

$d(\xi, y) < \frac{1}{2k_1} d(x_1, \xi)$ imply $\frac{1}{2k_1} d(x_1, \xi) < k_1 d(y, x_1) < 2\, a k_1$. Hence

$$\begin{aligned}
&\int_{\Omega_1 \cap B_{2a}(x_1)} d(x_1, y)^{\beta - Q} |h(y, \xi)| dy \\
&\qquad = \int_{\Omega_1 \cap B_{2a}(x_1),\; d(\xi, y) \geq \frac{1}{2k_1} d(x_1, \xi)} d(x_1, y)^{\beta - Q} |h(y, \xi)| dy \\
&\qquad + \int_{\Omega_1 \cap B_{2a}(x_1),\; d(\xi, y) < \frac{1}{2k_1} d(x_1, \xi)} d(x_1, y)^{\beta - Q} |h(y, \xi)| dy \\
&\qquad\qquad \leq\; c_1\, d(x_1, \xi)^{\gamma - Q} \int_{B_{2a}(x_1)} d(x_1, y)^{\beta - Q} dy + c_1\, d(x_1, \xi)^{\beta - Q} \int_{d(\xi, y) < 2ak_1} d(y, \xi)^{\gamma - Q} dy \\
&\qquad\qquad\qquad = c_2 \left[d(x_1, \xi)^{\gamma - Q} d(x_1, x_2)^{\beta} + d(x_1, \xi)^{\beta - Q} d(x_1, x_2)^{\gamma} \right],
\end{aligned}$$

for some $c_1, c_2 > 0$. Similarly, we have

$$\begin{aligned}
&\int_{\Omega_1 \cap B_{2a}(x_1)} d(x_2, y)^{\beta - Q} |h(y, \xi)| dy \\
&\qquad\qquad \leq c_3 \left(d(x_2, \xi)^{\gamma - Q} d(x_1, x_2)^{\beta} + d(x_2, \xi)^{\beta - Q} d(x_1, x_2)^{\gamma} \right),
\end{aligned}$$

for some $c_3 > 0$. Hence, there exists $C_2 > 0$ such that

$$I_2 \leq C_2\, d(x_1, x_2)^{\mu} \left(d(x_1, \xi)^{\mu - Q} + d(x_2, \xi)^{\mu - Q} \right). \tag{3.3.19}$$

In conclusion, by (3.3.16), (3.3.18) and (3.3.19) we obtain the desired bound for (3.3.16). $\square$

Proposition 3.3.15. *Let $\Omega_0 \supseteq \Omega_1$ be such that $\partial\Omega_0$ is a smooth $N-1$ manifold. Suppose that $g \in C^2_{\mathbb{G}}\Big(\big(\overline{\Omega}_0 \times \overline{\Omega}_0\big) \setminus D_{\overline{\Omega}_0}\Big)$ is such that*

$$\begin{aligned}
|g(x, \xi)| &\leq c\, d(x, \xi)^{\beta - Q}, \\
|X_i g(x, \xi)| &\leq c\, d(x, \xi)^{\beta - 1 - Q}, \\
|X_i X_j g(x, \xi)| &\leq c\, d(x, \xi)^{\beta - 2 - Q},
\end{aligned} \tag{3.3.20}$$

for some $\beta \geq 2$, $c > 0$ and for every $x, \xi \in \Omega_0, x \neq \xi$. Let $h \in C\left(\left(\overline{\Omega}_1 \times \overline{\Omega}_1\right) \setminus D_{\overline{\Omega}_1}\right)$, and assume that for every $\xi \in \Omega_1$, $y \mapsto h(y, \xi)$ belongs to $C^{\alpha}_{\mathbb{G}}(\Omega)$ for some $\alpha \in (0, 1]$. Then the function

$$f(x, \xi) = \int_{\Omega_1} g(x, y) h(y, \xi) dy, \qquad (x, \xi) \in (\Omega_1 \times \Omega_1) \setminus D_{\Omega_1},$$

belongs to $C^2_{\mathbb{G}}\left(\Omega_1 \times \Omega_1\right) \setminus D_{\Omega_1}$. Moreover, if we set $h(y, \xi) = 0$ for all $(y, \xi) \in \left(\Omega_0 \setminus \overline{\Omega}_1\right) \times \Omega_1$, we can represent the second order Lie derivatives of f as

$$\begin{aligned}
X_i X_j f(x, \xi) = &\int_{\Omega_0} X_i X_j g(x, y) \left[h(y, \xi) - h(x, \xi) \right] dy \\
&\qquad - h(x, \xi) \int_{\partial\Omega_0} X_j g(x, y) \widetilde{\nu}_i(y) d\mathcal{H}^{N-1}(y),
\end{aligned} \tag{3.3.21}$$

for every $(x,\xi) \in (\Omega_1 \times \Omega_1) \setminus D_{\Omega_1}$. *Here, we have set* $\widetilde{\nu}_i := \sum_{j=1}^N \varphi_j^i \nu_j$, *having recalled that* $X_i = \sum_{j=1}^N \varphi_j^i \partial_{x_j}$.

Proof. We fix $\xi \in \Omega_1$, and we first notice that by (3.3.20) and Proposition 3.3.13 the function

$$w(x,\xi) := X_j f(x,\xi) = \int_{\Omega_0} X_j g(x,y) h(y,\xi) dy,$$

is continuous for every $x \neq \xi$. Let us set

$$v(x,\xi) := \int_{\Omega_0} X_i X_j g(x,y) \left[h(y,\xi) - h(x,\xi)\right] dy \\ + h(x,\xi) \int_{\partial\Omega_0} X_j g(x,y) \widetilde{\nu}_i d\mathcal{H}^{N-1}(y),$$

that is well defined, thanks to (3.3.20) and the Hölder continuity of h. Take η as in the proof of Proposition 3.3.13, and for every $\varepsilon > 0$ set

$$w_\varepsilon(x,\xi) := \int_{\Omega_0} X_j(x,y) \eta\left(\frac{d(x,y)}{\varepsilon}\right) h(y,\xi) dy.$$

By the same argument employed in the proof of Proposition 3.3.13, we have that $w_\varepsilon \to w$ uniformly on every compact set. Moreover, $X_i w_\varepsilon \in C\Big(\Omega_1 \setminus \{\xi\}\Big)$, and by differentiating under the integral sign we get

$$\begin{aligned} X_i w_\varepsilon(x,\xi) &= \int_{\Omega_0} X_i \left(X_j g(x,y) \eta\left(\frac{d(x,y)}{\varepsilon}\right)\right) h(y,\xi) dy \\ &= \int_{\Omega_0} X_i \left(X_j g(x,y) \eta\left(\frac{d(x,y)}{\varepsilon}\right)\right) \left[h(y,\xi) - h(x,\xi)\right] dy \\ &\quad + h(x,\xi) \int_{\Omega_0} X_i \left(X_j g(x,y) \eta\left(\frac{d(x,y)}{\varepsilon}\right)\right) dy \\ &=: I_1 + I_2. \end{aligned}$$

In order to prove that $f \in C^2_{\mathbb{G}}(\Omega_1 \times \Omega_1) \setminus D_{\Omega_1}$ and that the representation formula (3.3.21) holds true, it only remains to prove that $X_i w_\varepsilon \to v$ uniformly.
Let us examine I_2: by the divergence theorem we get

$$I_2 = -h(x,\xi) \int_{\partial\Omega_0} X_j g(x,y) \eta\left(\frac{d(x,y)}{\varepsilon}\right) \widetilde{\nu}_i(y) d\mathcal{H}^{N-1}(y).$$

By the definition of η, we see that there exists $\overline{\varepsilon}$ such that $\eta\left(\frac{d(x,y)}{\overline{\varepsilon}}\right) = 1$ whenever $x \in \Omega_1$ and $y \in \partial\Omega_0$, therefore for every $\varepsilon < \overline{\varepsilon}$ we have

$$I_2 = -h(x,\xi) \int_{\partial\Omega_0} X_j g(x,y) \widetilde{\nu}_i(y) d\mathcal{H}^{N-1}(y).$$

Exploiting the Hölder continuity of h, for every $\varepsilon < \bar{\varepsilon}$ we conclude that

$$
\begin{aligned}
&|v(x,\xi) - X_i w_\varepsilon(x,\xi)| \\
&\quad \leq \int_{d(x,y)\leq 2\varepsilon} \left| X_i \left[\left((1 - \eta\left(\frac{d(x,y)}{\varepsilon}\right)\right) X_j g(x,y) \right] [h(y,\xi) - h(x,\xi)] \right| dy \\
&\qquad \leq c_1 \int_{d(x,y)\leq 2\varepsilon} \left[\frac{1}{\varepsilon} |X_j g(x,y)| + |X_i X_j g(x,y)| \right] |h(y,\xi) - h(x,\xi)|\, dy \\
&\qquad\quad \leq \frac{c_2}{\varepsilon} \int_{d(x,y)\leq 2\varepsilon} d(x,y)^{\alpha+1-Q} dy + c_2 \int_{d(x,y)\leq 2\varepsilon} d(x,y)^{\alpha - Q} dy \leq c_3\, \varepsilon^\alpha,
\end{aligned}
$$

for some $c_1, c_2, c_3 > 0$, and then the thesis follows. □

3.4 Proof of Theorem 3.1.4

We recall that by assumption **(H4)** we have

$$
\begin{aligned}
&|a_{ij}(x)|, |b_i(x)|, |c(x)| \leq M_1, && \forall x \in \Omega, \\
&|a_{ij}(x) - a_{ij}(y)|, |b_i(x) - b_i(y)|, |c(x) - c(y)| \leq M_2\, d(x,y)^\alpha, && \forall x, y \in \Omega,
\end{aligned}
\tag{3.4.1}
$$

for some $M_1, M_2 > 0$ and $\alpha \in (0, 1]$. In order to simplify the exposition, we suppose that the number $\frac{Q}{\alpha}$ is not an integer. Since Ω is bounded, this assumption is not restrictive.

We first show a preliminary result for the parametrix.

Lemma 3.4.16. *There exists* $c_1 = c_1(M_1, M_2, \lambda, \Omega) > 0$ *such that*

$$
|\mathscr{L}_{\mathbb{G}} Z(x,\xi)| \leq c_1 d(x,\xi)^{\alpha - Q}, \qquad \forall x, \xi \in \Omega, \tag{3.4.2}
$$

moreover $|\Omega_1| \leq |\Omega_2|$ *implies* $c_1(M_1, M_2, \lambda, \Omega_1) \leq c_1(M_1, M_2, \lambda, \Omega_2)$.
If in addition $d(x,y) < \frac{1}{2k_1 b}\, d(x,\xi)$*, then there exists* $c_2 = c_2(M_1, M_2, \lambda, \Omega) > 0$ *such that*

$$
|\mathscr{L}_{\mathbb{G}} Z(x,\xi) - \mathscr{L}_{\mathbb{G}} Z(y,\xi)| \leq c_2 d(x,y)^\alpha \left[d(x,\xi)^{-Q} + d(y,\xi)^{-Q} \right], \tag{3.4.3}
$$

moreover $|\Omega_1| \leq |\Omega_2|$ *implies* $c_2(M_1, M_2, \lambda, \Omega_1) \leq c_2(M_1, M_2, \lambda, \Omega_2)$. *Here,* k_1 *is the constant of the pseudo-triangular inequality* (3.2.4)*, while* b *is the constant appearing in Theorem 3.2.10.*

Proof. The first inequality follows from (3.3.6), (3.3.7), (3.4.1) and the boundedness of Ω, since

$$
\mathscr{L}_{\mathbb{G}} Z(x,\xi) = \sum_{j=1}^{m_1} (a_{ij}(x) - a_{ij}(\xi))\, X_i X_j Z(x,\xi) + \sum_{i=1}^{m_1} b_i(x) X_i Z(x,\xi) + c(x) Z(x,\xi).
$$

The proof of the second inequality relies on Theorem 3.2.10. We have

$$\begin{aligned}|\mathscr{L}_{\mathbb{G}}Z(x,\xi)-\mathscr{L}_{\mathbb{G}}Z(y,\xi)| \leq & \underbrace{\boxed{\sum_{i,j=1}^{m}|a_{ij}(x)-a_{ij}(y)||X_iX_jZ(y,\xi)|}}_{=:I_1} \\ & + \underbrace{\boxed{\sum_{i,j=1}^{m_1}|a_{ij}(x)-a_{ij}(\xi)||X_iX_jZ(x,\xi)-X_iX_jZ(y,\xi)|}}_{=:I_2} \\ & + \underbrace{\boxed{\sum_{i=1}^{m_1}|b_i(x)-b_i(y)||X_iZ(y,\xi)|}}_{=:I_3} \\ & + \underbrace{\boxed{\sum_{i=1}^{m_1}|b_i(x)||X_iZ(x,\xi)-X_iZ(y,\xi)|}}_{=:I_4} \\ & + \underbrace{\boxed{|c(x)-c(y)||Z(y,\xi)|}}_{=:I_5} \\ & + \underbrace{\boxed{|c(x)||Z(x,\xi)-Z(y,\xi)|}}_{=:I_6}.\end{aligned}$$

As before, the terms I_1, I_3 and I_5 are bounded by $c_2 d(x,y)^\alpha d(y,\xi)^{-Q}$ for some positive c_2 that depends on M_1, M_2, λ and Ω.
We claim that there exists some $c > 0$ such that

$$\begin{aligned}|Z(x,\xi)-Z(y,\xi)| &\leq c\, d(x,\xi)^{1-Q}d(x,y), \\ |X_iZ(x,\xi)-X_iZ(y,\xi)| &\leq c\, d(x,\xi)^{-Q}d(x,y), \\ |X_iX_jZ(x,\xi)-X_iX_jZ(y,\xi)| &\leq c\, d(x,\xi)^{-1-Q}d(x,y),\end{aligned} \tag{3.4.4}$$

whenever $d(x,y) < \frac{1}{2k_1 b}\, d(x,\xi)$. As a consequence of the above inequalities we conclude that also I_2, I_4 and I_6 are bounded by $c_2 d(x,y)^\alpha d(y,\xi)^{-Q}$. Indeed, by (3.4.4) the estimate is straightforward for I_4 and I_6, by the boundedness of Ω and (3.4.1). Moreover, it plainly follows for the second term as well, since by (3.4.1) we get

$$|a_{ij}(x)-a_{ij}(\xi)|\cdot|X_iX_jZ(x,\xi)-X_iX_jZ(y,\xi)| \leq M_2\, c\, d(x,\xi)^{-1-Q}d(x,y)^{1+\alpha},$$

and we conclude reminding that by assumption $d(x,y) < \frac{1}{2k_1 b}\, d(x,\xi)$.
It only remains to prove (3.4.4). Let $\eta \in C^1(\mathbb{R})$ such that $0 \leq \eta \leq 1$, $\eta(t) = 0$ if $t \leq \frac{2b-1}{4k_1 b}$ and $\eta(t) = 1$ if $t \geq \frac{2b-1}{2k_1 b}$. We observe that, since $k_1 \geq 1$, we have $\eta(1) = 1$. With the purpose of proving the first inequality in (3.4.4), we fix $x \neq \xi$ and we apply Theorem 3.2.10 to the function $w \mapsto Z(w,\xi)\eta\left(\frac{d(w,\xi)}{d(x,\xi)}\right)$. We obtain

$$\left|Z(y,\xi)\eta\left(\frac{d(y,\xi)}{d(x,\xi)}\right)-Z(x,\xi)\eta(1)\right| \leq c\, d(x,y) \sup_{\substack{z:d(x,z)\leq b\, d(x,y)\\ i=1,\dots,m_1}} \left|\left\{X_iZ(z,\xi)\eta\left(\frac{d(z,\xi)}{d(x,\xi)}\right)\right\}\right|,$$

for some absolute $c > 0$. We observe that, since $d(x,y) < \frac{1}{2k_1 b}\, d(x,\xi)$, by the pseudo-triangular inequality (3.2.4) we have

$$\frac{d(y,\xi)}{d(x,\xi)}, \frac{d(z,\xi)}{d(x,\xi)} \geq \frac{2b-1}{2k_1 b}.$$

Then, by the definition of η and (3.3.7) we get

$$|Z(x,\xi) - Z(y,\xi)| \leq c_1\, d(x,\xi)^{1-Q} d(x,y),$$

for some c_1 that depends only on λ. We conclude the proof observing that the second and third inequalities follow by the same argument, by using $p = 2$ and $p = 3$, respectively. □

Proof of Theorem 3.1.4. Our main aim is to prove that a fundamental solution for (3.1.1) is given by

$$\Gamma(x,\xi) = Z(x,\xi) + \int_\Omega Z(x,y)G(y,\xi)dy, \tag{3.4.5}$$

where Z is the parametrix and G is defined as

$$G(x,\xi) = \sum_{k=1}^{\infty} (\mathscr{L}_{\mathbb{G}} Z)_k(x,\xi), \tag{3.4.6}$$

where $(\mathscr{L}_{\mathbb{G}} Z)_1(x,\xi) = \mathscr{L}_{\mathbb{G}} Z(x,\xi)$ and

$$(\mathscr{L}_{\mathbb{G}} Z)_{k+1}(x,\xi) = \int_\Omega (\mathscr{L}_{\mathbb{G}} Z)_1(x,y)(\mathscr{L}_{\mathbb{G}} Z)_k(y,\xi)dy. \tag{3.4.7}$$

We divide the proof in several steps.

Step 1. We show that, if $|\Omega| < \delta_0$ for some $\delta_0 > 0$, (3.4.6) is the unique solution of

$$G(x,\xi) = \mathscr{L}_{\mathbb{G}} Z(x,\xi) + \int_\Omega \mathscr{L}_{\mathbb{G}} Z(x,y)G(y,\xi)dy. \tag{3.4.8}$$

To this end, we introduce the operator $T : C\left(\overline{\Omega} \times \overline{\Omega}\right) \to C\left(\overline{\Omega} \times \overline{\Omega}\right)$ defined as follows

$$Tf(x,\xi) := \int_\Omega \mathscr{L}_{\mathbb{G}} Z(x,y)f(y,\xi)dy, \qquad x,\xi \in \overline{\Omega},$$

which is well posed due to (3.4.2) and Proposition 3.3.10. Since Ω is bounded, there exists a positive $R = R(\Omega)$, such that $\Omega \subset B_R(x)$ for every $x \in \Omega$. Then, by (3.4.2), (3.2.5) and (3.2.6) it holds

$$\int_\Omega |\mathscr{L}_{\mathbb{G}} Z(x,\xi)|d\xi \leq c_1 \int_{B_R(x)} d(\xi,x)^{\alpha-Q} d\xi = \frac{c_1 Q \omega_Q}{\alpha} R^\alpha,$$

for some positive $c_1 = c_1(M_1, M_2, \lambda, \Omega)$, that does not increase if $|\Omega|$ decreases. This implies that there exists $\delta_0 > 0$ such that

$$\sup_{x\in\Omega} \int_\Omega |\mathscr{L}_{\mathbb{G}} Z(x,\xi)| d\xi < 1, \tag{3.4.9}$$

whenever $|\Omega| < \delta_0$. If (3.4.9) holds, the Neumann series $\sum_{k=0}^{\infty} T^k f$ converges uniformly. Then, once we show that there exists $m \in \mathbb{N}$ such that $(\mathscr{L}_{\mathbb{G}} Z)_m \in C\left(\overline{\Omega} \times \overline{\Omega}\right)$, the series

$$\sum_{k=0}^{\infty} T^k (\mathscr{L}_{\mathbb{G}} Z)_m(x,\xi) \overset{(3.4.7)}{=} \sum_{k=m}^{\infty} (\mathscr{L}_{\mathbb{G}} Z)_k(x,\xi),$$

converges uniformly, and thus (3.4.6) solves (3.4.8). Indeed we have

$$\begin{aligned}\int_\Omega \mathscr{L}_{\mathbb{G}} Z(x,y) G(y,\xi) dy &= \sum_{k=1}^{\infty} \int_\Omega \mathscr{L}_{\mathbb{G}} Z(x,y) (\mathscr{L}_{\mathbb{G}} Z)_k(y,\xi) dy \\ &= \sum_{k=1}^{\infty} (\mathscr{L}_{\mathbb{G}} Z)_{k+1}(x,\xi) = G(x,\xi) - \mathscr{L}_{\mathbb{G}} Z(x,\xi).\end{aligned}$$

Moreover, (3.4.9) guarantees that the integral equation (3.4.8) has at most one solution.
We prove that there exists $m \in \mathbb{N}$ such that $(\mathscr{L}_{\mathbb{G}} Z)_m \in C\left(\overline{\Omega} \times \overline{\Omega}\right)$ by induction: we show that for every $k \in \mathbb{N}$ there exists $c_k > 0$ such that

$$(\mathscr{L}_{\mathbb{G}} Z)_k(x,\xi) \le c_k\, d(x,\xi)^{k\alpha - Q}, \ \text{ if } \ k < \frac{Q}{\alpha}, \qquad (\mathscr{L}_{\mathbb{G}} Z)_k(x,\xi) \le c_k, \ \text{ if } \ k > \frac{Q}{\alpha}, \tag{3.4.10}$$

and

$$(\mathscr{L}_{\mathbb{G}} Z)_k \in C\left(\overline{\Omega} \times \overline{\Omega} \setminus D_{\overline{\Omega}}\right), \ \text{ if } \ k < \frac{Q}{\alpha}, \qquad (\mathscr{L}_{\mathbb{G}} Z)_k \in C\left(\overline{\Omega} \times \overline{\Omega}\right), \ \text{ if } \ k > \frac{Q}{\alpha}. \tag{3.4.11}$$

Since the parametrix $Z(x,\xi)$ is smooth when $x \neq \xi$ and the coefficients are continuous, we infer $(\mathscr{L}_{\mathbb{G}} Z)_1 \in C\left(\overline{\Omega} \times \overline{\Omega} \setminus D_{\overline{\Omega}}\right)$. Moreover, by (3.4.2) we conclude that (3.4.10) and (3.4.11) hold for $k = 1$. Suppose now true the claim for k: since

$$(\mathscr{L}_{\mathbb{G}} Z)_{k+1}(x,\xi) = \int_\Omega (\mathscr{L}_{\mathbb{G}} Z)_1(x,y) (\mathscr{L}_{\mathbb{G}} Z)_k(y,\xi) dy,$$

by Proposition 3.3.12 it follows that (3.4.10) and (3.4.11) hold for $k+1$. We end by choosing $m > \frac{Q}{\alpha}$.

Step 2. We prove that for every $\xi \in \Omega$ we have that $x \mapsto \Gamma(x,\xi) \in C^2_{\mathbb{G}}(\Omega \setminus \{\xi\}) \cap L^1(\Omega)$. By (3.4.10) there exists $C_1 = C_1(M_1, M_2, \lambda, \Omega) > 0$ such that

$$|G(x,\xi)| \le C_1 d(x,\xi)^{\alpha - Q}. \tag{3.4.12}$$

From (3.3.6), (3.4.12) and Proposition 3.3.10 it follows that

$$|\Gamma(x,\xi)| = \left| Z(x,\xi) + \int_\Omega Z(x,y)G(y,\xi)dy \right| \leq c\, d(x,\xi)^{2-Q}, \tag{3.4.13}$$

for some $c > 0$. This implies that $\Gamma(\cdot,\xi) \in L^1(\Omega)$ for every $\xi \in \Omega$, according to Remark 3.2.9.

Next, we prove that Γ belongs to $C^2_{\mathbb{G}}\Big((\Omega\times\Omega) \setminus D_\Omega\Big)$. Since Z is smooth in $(\Omega\times\Omega)\setminus D_\Omega$, we only need to prove that

$$J(x,\xi) = \int_\Omega Z(x,y)G(y,\xi)dy, \tag{3.4.14}$$

belongs to $C^2_{\mathbb{G}}\Big((\Omega\times\Omega) \setminus D_\Omega\Big)$. According to (3.3.7), (3.4.12) and Proposition 3.3.13, we infer that $J \in C^1_{\mathbb{G}}\Big((\Omega\times\Omega) \setminus D_\Omega\Big)$ and

$$X_j J(x,\xi) = X_j \int_\Omega Z(x,y)G(y,\xi)dy = \int_\Omega X_j Z(x,y)G(y,\xi)dy, \tag{3.4.15}$$

therefore $\Gamma \in C^1_{\mathbb{G}}\Big((\Omega\times\Omega) \setminus D_\Omega\Big)$. To prove the $C^2_{\mathbb{G}}$ regularity, we introduce the function $(x,y) \mapsto \Gamma_{A_x}(x,y)$, where Γ_{A_x} is the fundamental solution of the constant coefficient operator

$$\mathscr{L}_{\mathbb{G},A_x} = \sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j.$$

We fix $\eta > 0$ such that $x \notin B_\eta(\xi)$, and we write J as follows

$$\begin{aligned} J(x,\xi) = \int_\Omega Z(x,y)G(y,\xi)dy &= \int_{B_\eta(\xi)} Z(x,y)G(y,\xi)dy \\ &\quad + \int_{\Omega\setminus B_\eta(\xi)} \Big(Z(x,y) - \Gamma_{A(x)}(x,y)\Big)\, G(y,\xi)dy \\ &\quad + \int_{\Omega\setminus B_\eta(\xi)} \Gamma_{A(x)}(x,y)G(y,\xi)dy \\ &\qquad =: J_1(x,\xi) + J_2(x,\xi) + J_3(x,\xi). \end{aligned} \tag{3.4.16}$$

Since $Z(\cdot,\xi) \in C^\infty(\Omega\setminus\{\xi\})$, from estimate (3.4.12) and Proposition 3.3.13 we infer that $J_1 \in C^2_{\mathbb{G}}\Big((\Omega\times\Omega) \setminus D_\Omega\Big)$ and

$$X_iX_jJ_1(x,\xi) = X_iX_j \int_{B_\eta(\xi)} Z(x,y)G(y,\xi)dy = \int_{B_\eta(\xi)} X_iX_jZ(x,y)G(y,\xi)dy. \tag{3.4.17}$$

Moreover, since $a_{ij} \in C^\alpha_{\mathbb{G}}\left(\overline{\Omega}\right)$, by (3.3.8) we get

$$|X_iX_jZ(x,\xi) - X_iX_j\Gamma_{A_x}(x,\xi)| \leq c_1\, d(x,\xi)^{\frac{\alpha}{\kappa}-Q},$$

for some $c_1 > 0$. Thus by (3.4.12) and Proposition 3.3.13 we have $J_2 \in C^2_{\mathbb{G}}\Big((\Omega \times \Omega) \setminus D_\Omega \Big)$ and

$$X_i X_j J_2(x,\xi) = X_i X_j \int_{\Omega \setminus B_\eta(\xi)} \left(Z(x,y) - \Gamma_{A_x}(x,y)\right) G(y,\xi) dy$$
$$= \int_{\Omega \setminus B_\eta(\xi)} \left(X_i X_j Z(x,y) - X_i X_j \Gamma_{A_x}(x,y)\right) G(y,\xi) dy. \tag{3.4.18}$$

It only remains to prove that $J_3 \in C^2_{\mathbb{G}}\Big((\Omega \times \Omega) \setminus D_\Omega \Big)$. To this end, we prove that $y \mapsto G(y,\xi) \in C^\sigma_{\mathbb{G}}\big(\Omega \setminus B_\eta(\xi)\big)$ for some $0 < \sigma \le 1$. Then we conclude by Proposition 3.3.15, reminding that $\Gamma_{A_x} \in C^\infty\left(\left(\mathbb{R}^N \times \mathbb{R}^N\right) \setminus D_{\mathbb{R}^N}\right)$ and taking into account estimate (3.3.7). To show this Hölder regularity, we recall that G solves

$$G(x,\xi) = \mathscr{L}_{\mathbb{G}} Z(x,\xi) + W(x,\xi), \tag{3.4.19}$$

where

$$W(x,\xi) := \int_\Omega \mathscr{L}_{\mathbb{G}} Z(x,y) G(y,\xi) dy.$$

Being $Z(\cdot,\xi)$ smooth in $\mathbb{R}^N \setminus \{\xi\}$, by assumption **(H4)** we have $\mathscr{L}_{\mathbb{G}} Z(\cdot,\xi) \in C^\sigma_{\mathbb{G}}\big(\Omega \setminus B_\eta(\xi)\big)$ for every $\sigma \le \alpha$. We conclude observing that, by (3.4.3) and Proposition 3.3.14, there exists a positive constant c_2 and $\sigma \in (0,\alpha)$ such that

$$|W(x_1,\xi) - W(x_2,\xi)| \le c_2 \, d(x_1,x_2)^\sigma \left(d(x_1,\xi)^{\alpha - Q} + d(x_2,\xi)^{\alpha - Q} \right). \tag{3.4.20}$$

Step 3. We prove that for every $\xi \in \Omega$ the function Γ defined in (3.4.5) solves $\mathscr{L}_{\mathbb{G}} \Gamma(\cdot,\xi) = 0$ in $\Omega \setminus \{\xi\}$. To this end, we examine the action of the differential operator $\sum a_{ij}(x) X_i X_j$ on the function J defined in (3.4.14). We consider again the splitting (3.4.16): (3.4.17) and (3.4.18) yield

$$\sum_{i,j=1}^{m_1} X_i X_j a_{ij}(x) \int_{B_\eta(\xi)} Z(x,y) G(y,\xi) dy = \int_{B_\eta(\xi)} \sum_{i,j=1}^{m_1} X_i X_j a_{ij}(x) Z(x,y) G(y,\xi) dy,$$
$$\sum_{i,j=1}^{m_1} X_i X_j a_{ij}(x) \int_{\Omega \setminus B_\eta(\xi)} \left(Z(x,y) - \Gamma_{A_x}(x,y)\right) G(y,\xi) dy$$
$$= \int_{\Omega \setminus B_\eta(\xi)} \left(\sum_{i,j=1}^{m_1} X_i X_j a_{ij}(x) Z(x,y) - \sum_{i,j=1}^{m_1} X_i X_j a_{ij}(x) \Gamma_{A_x}(x,y) \right) G(y,\xi) dy. \tag{3.4.21}$$

We focus on the second equation in (3.4.21). As $\Gamma_{A_x}(x,y)$ solves

$$\sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j \Gamma_{A_x}(x,y) = 0, \tag{3.4.22}$$

if $x \neq y$, we get

$$\begin{aligned}
&\int_{\Omega\setminus B_\eta(\xi)} \left(\sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j Z(x,y) - \sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j \Gamma_{A_x}(x,y) \right) G(y,\xi) dy \\
&\quad = \lim_{\varepsilon\to 0} \int_{(\Omega\setminus B_\eta(\xi))\setminus B_\varepsilon(x)} \left(\sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j Z(x,y) - \sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j \Gamma_{A_x}(x,y) \right) G(y,\xi) dy \\
&\quad = \lim_{\varepsilon\to 0} \int_{(\Omega\setminus B_\eta(\xi))\setminus B_\varepsilon(x)} \sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j Z(x,y) G(y,\xi) dy \\
&\quad = P.V. \int_{\Omega\setminus B_\eta(\xi)} \sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j Z(x,y) G(y,\xi) dy.
\end{aligned} \tag{3.4.23}$$

It only remains to deal with the term J_3 in (3.4.16). More precisely, we analyse

$$\sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j \int_{\Omega\setminus B_\eta(\xi)} \Gamma_{A_x}(x,y) G(y,\xi) dy.$$

For every $r > 0$ we set

$$\psi_r(x) := \left\{ y \in \mathbb{R}^N : \Gamma_{A_x}(x,y) > \frac{1}{r} \right\}.$$

We first observe that, since $\Gamma_{A_x}(x,y)$ is smooth for $x \neq y$, by Sard's lemma there exists $R > 0$ such that $\partial\psi_R(x)$ is a smooth $N-1$ manifold and $\Omega \subset \psi_R(x)$. Therefore, we can apply Proposition 3.3.15 with $\Omega_1 = \Omega \setminus B_\eta(\xi)$ and $\Omega_0 = \psi_R(x)$, concluding that

$$\begin{aligned}
&\sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j \int_{\Omega\setminus B_\eta(\xi)} \Gamma_{A_x}(x,y) G(y,\xi) dy \\
&\quad = \int_{\psi_R(x)} \sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j \Gamma_{A_x}(x,y) \big[G(y,\xi) - G(x,\xi) \big] dy \\
&\qquad - G(x,\xi) \int_{\partial\psi_R(x)} \sum_{i,j=1}^{m_1} a_{ij}(x) X_j \Gamma_{A_x}(x,y) \widetilde{\nu}_i d\mathcal{H}^{N-1}(y).
\end{aligned}$$

From (3.4.22) we deduce that the first integral in the right-hand side vanishes, moreover from Theorem 3.5 in [23] it follows that the last integral equals 1. Thus we infer that

$$\sum_{i,j=1}^{m_1} a_{ij}(x) X_i X_j \int_{\Omega\setminus B_\eta(\xi)} \Gamma_{A(x)}(x,y) G(y,\xi) dy = -G(x,\xi). \tag{3.4.24}$$

By (3.4.15),(3.4.16),(3.4.21), (3.4.23) and (3.4.24) we conclude that

$$\mathscr{L}_{\mathbb{G}} \int_\Omega Z(x,y) G(y,\xi) dy = \int_\Omega \mathscr{L}_{\mathbb{G}} Z(x,y) G(y,\xi) dy - G(x,\xi). \tag{3.4.25}$$

Then, the function Γ defined in (3.4.5) solves $\mathscr{L}_{\mathbb{G}}\Gamma(x,\xi)=0$ for $x\neq\xi$, as G solves (3.4.8).

Step 4. We prove that for every $\varphi\in C_c^\infty(\Omega)$ the function

$$w(x)=\int_\Omega \Gamma(x,\xi)\varphi(\xi)d\xi, \tag{3.4.26}$$

is a classical solution to $\mathscr{L}_{\mathbb{G}}w=-\varphi$. The $C^2_{\mathbb{G}}$ regularity plainly follows from the same arguments we have used above for the $C^2_{\mathbb{G}}$ regularity of Γ. Then we observe that, by the same reasoning that gave us (3.4.25), if $f\in C^\beta_{\mathbb{G}}(\Omega)$ for some $\beta\in(0,1]$ we have

$$\mathscr{L}_{\mathbb{G}}\int_\Omega Z(x,\xi)f(\xi)d\xi=\int_\Omega \mathscr{L}_{\mathbb{G}}Z(x,\xi)f(\xi)d\xi-f(x). \tag{3.4.27}$$

Moreover, by (3.4.19),(3.4.20) and Remark 3.2.9 the function

$$\int_\Omega G(y,\xi)\varphi(\xi)d\xi,$$

belongs to $C^\delta_{\mathbb{G}}(\Omega)$ for every $\delta<\alpha$. Then

$$\begin{aligned}&\mathscr{L}_{\mathbb{G}}\int_\Omega Z(x,y)\int_\Omega G(y,\xi)\varphi(\xi)d\xi dy\\&\qquad=\int_\Omega \mathscr{L}_{\mathbb{G}}Z(x,y)\int_\Omega G(y,\xi)\varphi(\xi)d\xi dy-\int_\Omega G(x,\xi)\varphi(\xi)d\xi.\end{aligned} \tag{3.4.28}$$

By (3.4.27) and (3.4.28) we conclude that

$$\begin{aligned}\mathscr{L}_{\mathbb{G}}\int_\Omega \Gamma(x,\xi)\varphi(\xi)d\xi&=\mathscr{L}_{\mathbb{G}}\int_\Omega Z(x,\xi)\varphi(\xi)d\xi+\int_\Omega\int_\Omega Z(x,y)G(y,\xi)\varphi(\xi)dyd\xi\\&=-\varphi(x)+\int_\Omega\left(\mathscr{L}_{\mathbb{G}}Z(x,\xi)-G(x,\xi)\right)\varphi(\xi)d\xi+\int_\Omega \mathscr{L}_{\mathbb{G}}Z(x,y)\int_\Omega G(y,\xi)\varphi(\xi)d\xi dy\\&=-\varphi(x)+\int_\Omega\Big(\mathscr{L}_{\mathbb{G}}Z(x,\xi)-G(x,\xi)+\int_\Omega \mathscr{L}_{\mathbb{G}}Z(x,y)G(y,\xi)dy\Big)\varphi(\xi)d\xi=-\varphi(x),\end{aligned} \tag{3.4.29}$$

as G solves (3.4.8).

Step 5. It only remains to prove the $C^{2+\alpha}_{\mathbb{G}}$ regularity stated in the last part of Theorem 3.1.4. We show that for every $\xi\in\Omega$ the function $x\mapsto\Gamma(x,\xi)$ belongs to $C^{2+\alpha}_{\mathbb{G},\mathrm{loc}}(\Omega\setminus\{\xi\})$, and that for every $f\in C^\alpha(\Omega)\cap L^1(\Omega)$ the function

$$u(x):=\int_\Omega \Gamma(x,\xi)f(\xi)d\xi,\qquad x\in\Omega,$$

solves $\mathscr{L}_{\mathbb{G}}u=-f$ and belongs to the space $C^{2+\alpha}_{\mathbb{G}}(\Omega)$.

The first claim follows from Theorem 1.4 in [43]. Regarding the second, it is straightforward to prove that u verifies the equation, following the steps made to prove that (3.4.26) solves

$\mathscr{L}_{\mathbb{G}} w = -\varphi$. Indeed, if f is Hölder continuous we remind (3.4.27). Moreover by (3.4.19), (3.4.20) and Remark 3.2.9 the function

$$\int_\Omega G(x,y)f(y)dy,$$

belongs to $C^\delta_{\mathbb{G}}(\Omega)$ for every $\delta < \alpha$, inferring the validity of (3.4.28) and hence concluding as in (3.4.29). Finally, the $C^{2+\alpha}_{\mathbb{G}}$ regularity follows again from Theorem 1.4 in [43]. □

3.5 Maximum principle and Harnack inequality

We now employ the previous result to prove a maximum principle and a Harnack inequality for (3.1.4). With this aim we apply the mean value formula developed in [91], that holds for divergence form operators. We recall the notation (3.1.3), (3.1.4) and (3.1.5) provided in the introduction. Hereafter in this section we denote as z a generic point in Ω, and we choose $R > 0$ such that $B_R(z) \subset \Omega$ and $|B_R(z)| < \delta_0$, with δ_0 given by Theorem 3.1.4. In $B_R(z)$, as proved in Section 3.4, there exists a fundamental solution Γ to (3.1.4), and a fundamental solution Γ^* to (3.1.5). Given their existence, we can use the mean value formula in [91].

For every $x, \xi \in B_R(z)$ and $\varrho > 0$ we set

$$\begin{aligned} M_{\mathbb{G}}(x,\xi) &:= \frac{Q}{Q-2} \frac{\langle A(\xi)\nabla_{\mathbb{G}}\Gamma^*(x,\xi), \nabla_{\mathbb{G}}\Gamma^*(x,\xi)\rangle}{\Gamma^*(x,\xi)^{\frac{2(Q-1)}{Q-2}}}, \\ \Omega_\varrho(\xi) &:= \Big\{ y \in B_R(z) : \Gamma^*(y,\xi) > \varrho^{2-Q} \Big\}. \end{aligned} \tag{3.5.1}$$

The following representation formula holds true.

Theorem 3.5.17 (Theorem 3.3 in [91]). *Let $f \in C\big(B_R(z)\big)$, and let u be a classical solution to $\mathscr{L}_{\mathbb{G}} u = f$ in $B_R(z)$. Then, for every $\xi \in B_R(z)$ and for every $\varrho > 0$ such that $\overline{\Omega_\varrho(\xi)} \subset B_R(z)$ we have*

$$\begin{aligned} u(\xi) =& \frac{1}{\varrho^Q} \int_{\Omega_\varrho(\xi)} M_{\mathbb{G}}(x,\xi)u(x)dx + \frac{Q}{\varrho^Q} \int_0^\varrho \left(s \int_{\Omega_s(\xi)} (\mathrm{div}_{\mathbb{G}} b(x) - c(x))\, u(x) dx \right) ds \\ &+ \frac{Q}{\varrho^Q} \int_0^\varrho \left(s^{Q-1} \int_{\Omega_s(\xi)} f(x) \left(\frac{1}{s^{Q-2}} - \Gamma^*(x,\xi) \right) dx \right) ds. \end{aligned} \tag{3.5.2}$$

Remark 3.5.18. *The second term in the right-hand side of* (3.5.2) *does not appear in Theorem 3.3 in [91], since the authors assume* $\mathrm{div}_{\mathbb{G}} b(x) - c(x) = 0$.

We now give some bounds for Γ^* and for its Lie derivatives $X_i\Gamma^*$, in order to prove some useful estimates of the kernel $M_{\mathbb{G}}$ and to characterize the set Ω_ϱ. To this end, we recall that the fundamental solution Γ^* given by Theorem 3.1.4 is

$$\Gamma^*(x,\xi) = Z^*(x,\xi) + \int_{B_R(z)} Z^*(x,y)G^*(y,\xi)dy, \tag{3.5.3}$$

where Z^* is the parametrix of (3.1.5) (see Section 3.3). We recall that by Theorem 3.3.11 there exist $c_\lambda, c_{\lambda,1} > 0$ such that

$$\begin{aligned} \frac{1}{c_\lambda} d(x,\xi)^{2-Q} \leq Z^*(x,\xi) \leq c_\lambda\, d(x,\xi)^{2-Q}, \\ |X_i Z^*(x,\xi)| \leq c_{\lambda,1}\, d(x,\xi)^{1-Q}. \end{aligned} \tag{3.5.4}$$

Notably, one can identify a homogeneous distance d_ξ by setting $d_\xi(x,\xi) = Z^*(x,\xi)^{\frac{1}{Q-2}}$, then

$$Z^*(x,\xi) = d_\xi(x,\xi)^{2-Q}, \qquad X_i Z^*(x,\xi) = (2-Q)d_\xi(x,\xi)^{1-Q} X_i d_\xi(x,\xi). \tag{3.5.5}$$

From (3.5.4), (3.5.5), (3.4.12) and Proposition 3.3.10, we deduce that there exist positive constants c_1, c_2 (depending on M_1, M_2, λ and Ω) such that

$$\begin{aligned} \left| \int_{B_R(z)} Z^*(x,y)G^*(y,\xi)dy \right| \leq c_1 d_\xi(x,\xi)^{2-Q+\alpha}, \\ \left| \int_{B_R(z)} X_i Z^*(x,y)G^*(y,\xi)dy \right| \leq c_2 d_\xi(x,\xi)^{1-Q+\alpha}. \end{aligned} \tag{3.5.6}$$

Then, by (3.5.3), (3.5.5) and (3.5.6) we infer that there exist $c_1, c_2, c_3 > 0$ (depending on M_1, M_2, λ and Ω) such that

$$1 - c_1 d_\xi(x,\xi)^\alpha \leq \frac{\Gamma^*(x,\xi)}{d_\xi(x,\xi)^{2-Q}} \leq 1 + c_1 d_\xi(x,\xi)^\alpha, \tag{3.5.7}$$

$$X_i d_\xi(x,\xi) - c_2 d_\xi(x,\xi)^\alpha \leq \frac{X_i \Gamma^*(x,\xi)}{(2-Q)d_\xi(x,\xi)^{1-Q}} \leq X_i d_\xi(x,\xi) + c_2 d_\xi(x,\xi)^\alpha \leq c_3, \tag{3.5.8}$$

for every $x, \xi \in B_R(z)$. We first examine Ω_ϱ by using (3.5.7).

Lemma 3.5.19. *There exist $r_1, C_1, C_2 > 0$ (depending on M_1, M_2, λ and Ω) such that*

$$\Omega_{C_1 r}(\xi) \subseteq B_r(\xi) \subseteq \Omega_{C_2 r}(\xi), \tag{3.5.9}$$

for every $\xi \in B_R(z)$ and $r \leq r_1$.

Proof. The proof follows from the definition of Ω_ϱ in (3.5.1), along with estimate (3.5.7). □

We now focus on the kernel $M_{\mathbb{G}}$ defined in (3.5.1). By the ellipticity assumption in **(H4)**, (3.5.7), (3.5.8), (3.5.4) and (3.5.5) imply that there exist $C_3, r_2 > 0$ (depending on M_1, M_2, λ and Ω) such that

$$0 \leq \frac{1}{\lambda} \frac{|\nabla_{\mathbb{G}} \Gamma^*(x,\xi)|^2}{\Gamma^*(x,\xi)^{\frac{2(Q-1)}{Q-2}}} \leq M_{\mathbb{G}}(x,\xi) \leq \lambda \frac{|\nabla_{\mathbb{G}} \Gamma^*(x,\xi)|^2}{\Gamma^*(x,\xi)^{\frac{2(Q-1)}{Q-2}}} \leq C_3, \qquad \forall x, \xi \in B_R(z) : d_\xi(x,\xi) < r_2. \tag{3.5.10}$$

As said in the introduction, the proof of the maximum principle and the Harnack inequality would be straightforward if the kernel $M_{\mathbb{G}}$ were also strictly positive, but this is not the case when considering subelliptic operators. The following lemma states that a strictly positive lower bound for $M_{\mathbb{G}}$ holds along the integral paths of the generating vector fields $\{X_i\}_{i=1}^{m_1}$ (we refer to the definition of integral paths provided in Definition 3.1.1).

Lemma 3.5.20. *Let $\xi \in B_R(z)$ and let $\gamma_{X_i,\xi}$ be the integral path of X_i from ξ. Then, there exist $C_4, \bar{s} > 0$ (depending on M_1, M_2, λ and Ω) such that*

$$C_4 \leq M_{\mathbb{G}}\Big(\gamma_{X_i,\xi}(s), \xi\Big), \qquad \textit{for every } s \neq 0, |s| \leq \bar{s}.$$

Proof. By Corollary 1.26 in [119] (and the remark there below), it follows that

$$d_\xi\Big(\gamma_{X_i,\xi}(s), \xi\Big) = C_i |s|, \qquad \text{for every } \; s \in \mathbb{R}, \tag{3.5.11}$$

for some $C_i = C_i(X_i) > 0$. This implies (see Definition 3.1.1) that $|X_i d_\xi\Big(\gamma_{X_i,\xi}(s), \xi\Big)| \geq c$ for some $c > 0$, and then by (3.5.10) (3.5.7), and (3.5.8) the thesis follows. □

This lemma suggests us to introduce the following set: let $\xi \in B_R(z)$ and let $\gamma_{X_i,\xi}$ be the integral path of X_i starting from ξ; for every $\delta, s^* > 0$ we set

$$\mathcal{A}_{X_i,\xi,\delta,s^*} := \Big\{ x \in B_R(z) : d\Big(x, \gamma_{X_i,\xi}(s)\Big) < \delta |s|, \text{ for all } s \in \mathbb{R} : |s| \leq s^* \Big\}.$$

Crucially, in this set $M_{\mathbb{G}}$ preserves a strictly positive lower bound.

Lemma 3.5.21. *There exist $\nu, \delta, s^* > 0$ (depending on M_1, M_2, λ and Ω) such that for every $\xi \in B_R(z)$ it holds*

$$\nu \leq M_{\mathbb{G}}(x,\xi), \qquad \textit{for every } x \in \mathcal{A}_{X_i,\xi,\delta,s^*}. \tag{3.5.12}$$

Proof. According to Lemma 3.5.20, it suffices to prove that for every $\varepsilon > 0$ there exist $\delta > 0$ and $0 < s^* < \bar{s}$ (depending on M_1, M_2, λ and Ω) such that

$$\Big| M_{\mathbb{G}}\Big(\gamma_{X_i,\xi}(s), \xi\Big) - M_{\mathbb{G}}(x,\xi) \Big| < \varepsilon, \qquad \text{for every } x \in \mathcal{A}_{X_i,\xi,\delta,s^*}. \tag{3.5.13}$$

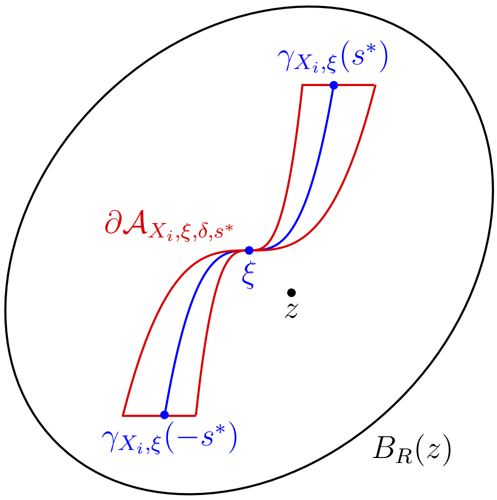

Figure 3.1: An example of $\mathcal{A}_{X_i,\xi,\delta,s^*}$.

If we employ the ellipticity assumption in **(H4)** in (3.5.1), by (3.2.4), (3.5.11) (3.5.7), (3.5.8) and (3.5.10) we see that there exist $\bar{\delta}, \bar{s} > 0$ (depending on M_1, M_2, λ and Ω) such that

$$\left|M_{\mathbb{G}}\Big(\gamma_{X_i,\xi}(s),\xi\Big) - M_{\mathbb{G}}(x,\xi)\right| < \frac{Q}{Q-2} 2\sqrt{\lambda C_3} \left| \frac{\nabla_{\mathbb{G}}\Gamma^*\Big(\gamma_{X_i,\xi}(s),\xi\Big)}{\Gamma^*\Big(\gamma_{X_i,\xi}(s),\xi\Big)^{\frac{(Q-1)}{Q-2}}} - \frac{\nabla_{\mathbb{G}}\Gamma^*(x,\xi)}{\Gamma^*(x,\xi)^{\frac{(Q-1)}{Q-2}}} \right|, \tag{3.5.14}$$

if $s < \bar{s}$ and $x \in \mathcal{A}_{X_i,\xi,\bar{\delta},\bar{s}}$. We use (3.5.7) and (3.5.8) to bound (3.5.14). By (3.2.4) and (3.5.11), for every $\widetilde{\varepsilon} > 0$ there exist $\widetilde{\delta}, \widetilde{s} > 0$ (again depending only on M_1, M_2, λ and Ω) such that

$$1 - \widetilde{\varepsilon} \le \frac{\Gamma^*(x,\xi)}{d_\xi(x,\xi)^{2-Q}} \le 1 + \widetilde{\varepsilon}, \qquad X_i d_\xi(x,\xi) - \widetilde{\varepsilon} \le \frac{X_i\Gamma^*(x,\xi)}{(2-Q)d_\xi(x,\xi)^{1-Q}} \le X_i d_\xi(x,\xi) + \widetilde{\varepsilon}, \tag{3.5.15}$$

if $x \in \mathcal{A}_{X_i,\xi,\widetilde{\delta},\widetilde{s}}$. Then, by (3.5.8) there exist c_1, c_2 (depending on M_1, M_2, λ and Ω) such that

$$\begin{aligned}
&\left| \frac{\nabla_{\mathbb{G}}\Gamma^*\Big(\gamma_{X_i,\xi}(s),\xi\Big)}{\Gamma^*\Big(\gamma_{X_i,\xi}(s),\xi\Big)^{\frac{(Q-1)}{Q-2}}} - \frac{\nabla_{\mathbb{G}}\Gamma^*(x,\xi)}{\Gamma^*(x,\xi)^{\frac{(Q-1)}{Q-2}}} \right| \\
&\qquad \le \left| \left(\frac{2-Q}{1-\widetilde{\varepsilon}}\right) \frac{\nabla_{\mathbb{G}}\Gamma^*\Big(\gamma_{X_i,\xi}(s),\xi\Big)}{(2-Q)d_\xi\Big(\gamma_{X_i,\xi}(s),\xi\Big)^{1-Q}} - \left(\frac{2-Q}{1+\widetilde{\varepsilon}}\right) \frac{\nabla_{\mathbb{G}}\Gamma^*(x,\xi)}{(2-Q)d_\xi(x,\xi)^{1-Q}} \right| \\
&\qquad\qquad \le \frac{c_1}{1-\widetilde{\varepsilon}^2} \left| \frac{\nabla_{\mathbb{G}}\Gamma^*\Big(\gamma_{X_i,\xi}(s),\xi\Big)}{(2-Q)d_\xi\Big(\gamma_{X_i,\xi}(s),\xi\Big)^{1-Q}} - \frac{\nabla_{\mathbb{G}}\Gamma^*(x,\xi)}{(2-Q)d_\xi(x,\xi)^{1-Q}} \right| + \widetilde{\varepsilon}\, c_2,
\end{aligned} \tag{3.5.16}$$

whenever $s < \widetilde{s}$ and $x \in \mathcal{A}_{X_i,\xi,\widetilde{\delta},\widetilde{s}}$. To conclude with the proof of (3.5.13), from (3.5.14), (3.5.15) and (3.5.16) we see that it remains to prove that for every $\varepsilon_i > 0$ there exists $\delta_i > 0$

such that

$$\Big|X_i d_\xi\Big(\gamma_{X_i,\xi}(s),\xi\Big) - X_i d_\xi(x,\xi)\Big| \leq \varepsilon_i, \quad \forall x \in B_R(z) : d_\xi\Big(x,\gamma_{X_i,\xi}(s)\Big) < \delta_i|s|. \tag{3.5.17}$$

Let C_i as in Lemma 3.5.20, b as in Theorem 3.2.10 and k_1, k_2 as in (3.2.4) and (3.2.5). For every $\tau_i < \frac{C_i}{k_1}$ there exists $0 < \delta_i < \tau_i$, with $\delta_i \leq \frac{\tau_i}{2k_1 b}$, such that $d(x,\xi) \geq \tau_i|s|$ if $d\Big(x,\gamma_{X_i,\xi}(s)\Big) < \delta_i|s|$. Then, arguing as in the proof of (3.4.4), by (3.5.5), (3.5.4) and (3.2.5) we deduce that there exists $c_\lambda > 0$ such that

$$\Big|X_i d_\xi\Big(\gamma_{X_i,\xi}(s),\xi\Big) - X_i d_\xi(x,\xi)\Big| \leq c_\lambda k_2 \, d(x,\xi)^{-1} d\Big(x,\gamma_{X_i,\xi}(s)\Big) \leq c_\lambda \frac{\delta_i}{\tau_i},$$

whenever $d\Big(x,\gamma_{X_i,\xi}(s)\Big) < \delta_i|s|$. Hence, (3.5.17) follows by the arbitrariness of δ_i. $\square$

We are now in position to prove the maximum principle.

Proof of Theorem 3.1.5. Let $\xi \in B_R(z) \subseteq \Omega$ such that $u(\xi) = \max_\Omega u$, and take ν, δ, s^* as in Lemma 3.5.21. We first show that the following relation holds

$$u(\xi) = u\Big(\gamma_{X_i,\xi}(s)\Big), \qquad \text{for every } s \in \mathbb{R}, |s| \leq s^*. \tag{3.5.18}$$

Since $\mathscr{L}_{\mathbb{G}} 1 = c$, by the mean value formula (3.5.2) it holds

$$\begin{aligned} 0 = & \underbrace{\boxed{\frac{1}{\varrho^Q}\int_{\mathcal{A}_{X_i,\xi,\delta\sigma^*}\cap\Omega_\varrho(\xi)} M_{\mathbb{G}}(x,\xi)(u(x)-u(\xi))dx}}_{=:I_1} \\ & + \underbrace{\boxed{\frac{1}{\varrho^Q}\int_{\Omega_\varrho(\xi)\setminus\mathcal{A}_{X_i,\xi,\delta,s^*}} M_{\mathbb{G}}(x,\xi)(u(x)-u(\xi))dx}}_{=:I_2} \\ & + \underbrace{\boxed{\frac{Q}{\varrho^Q}\int_0^\varrho\left(s\int_{\Omega_s(\xi)}(\mathrm{div}_{\mathbb{G}} b(x) - c(x))\,(u(x)-u(\xi))dx\right)ds}}_{=:I_3} \\ & + \underbrace{\boxed{\frac{Q}{\varrho^Q}\int_0^\varrho\left(s^{Q-1}\int_{\Omega_s(\xi)}(f(x)-u(\xi)c(x))\left(\frac{1}{s^{Q-2}} - \Gamma^*(x,\xi)\right)dx\right)ds}}_{=:I_4}. \end{aligned} \tag{3.5.19}$$

By definition $s^{2-Q} - \Gamma^*(x,\xi) < 0$ for every $x \in \Omega_s(\xi)$. Since by assumption we have $\mathrm{div}_{\mathbb{G}} b - c \geq 0$, $c \leq 0$ and $u(\xi), f \geq 0$, we deduce that $I_1, I_2, I_3, I_4 \leq 0$, and by (3.5.19) we infer that these inequalities are in fact equalities. In particular, $I_1 = 0$ and (3.5.12) imply $u(\xi) = u(x)$ for almost every $x \in \mathcal{A}_{X_i,\xi,\delta,s^*}$, and then by the continuity of u (3.5.18) holds. We conclude the proof employing the Carathéodory–Chow–Rashevsky connectivity theorem (see Theorem 6.22. in [19]): by Hörmander's rank condition **(H1)**, we can join every $x, y \in B_R(z)$ by a finite number of integral paths of the vector fields $\{X_i\}_{i=1}^{m_1}$; then, $u(x) = u(\xi)$ for every $x \in B_R(z)$, and by a covering argument we conclude that this holds for every $x \in \Omega$. Finally, we notice that if $c = 0$ the integral I_3 in (3.5.19) is strictly positive even if $u(\xi) \leq 0$, and therefore the assumption on the sign of $u(\xi)$ can be dropped. □

In order to prove the Harnack inequality, we need to investigate more thoroughly the second integral in (3.5.2).

Lemma 3.5.22. *For every $\varrho > 0$ such that $\overline{\Omega_\varrho(\xi)} \subset B_R(z)$ it holds*

$$\begin{aligned}\int_0^\varrho s\Big(\int_{\Omega_s(\xi)}(\mathrm{div}_{\mathbb{G}} b(x) - c(x))u(x)dx\Big)ds \\ = \int_{\Omega_\varrho(\xi)}(\mathrm{div}_{\mathbb{G}} b(x) - c(x))u(x)\left(\frac{\varrho^2 - \Gamma^*(x,\xi)^{\frac{2}{2-Q}}}{2}\right)dx.\end{aligned} \tag{3.5.20}$$

Proof. Since $\Omega_s(\xi) \subseteq \Omega_\varrho(\xi)$ for every $s < \varrho$, it holds

$$\begin{aligned}\int_0^\varrho s\Big(\int_{\Omega_s(\xi)}(\mathrm{div}_{\mathbb{G}} b(x) - c(x))u(x)dx\Big)ds \\ = \int_0^\varrho s\Big(\int_{\Omega_\varrho(\xi)}\chi_{\Omega_s(\xi)}(x)(\mathrm{div}_{\mathbb{G}} b(x) - c(x))u(x)dx\Big)ds.\end{aligned} \tag{3.5.21}$$

In order to swap the integration order in the integral on the right-hand side, we establish that

$$(x,s) \mapsto s\,\chi_{\Omega_s(\xi)}(x)(\mathrm{div}_{\mathbb{G}} b(x) - c(x))u(x) \in L^1\Big(\Omega_\varrho(\xi) \times [0,\varrho)\Big). \tag{3.5.22}$$

Since $\overline{\Omega_\varrho(\xi)} \subset B_R(z) \subset \Omega$, by **(H4)** and **(H5)** we infer the existence of $C > 0$ such that $\sup_{\Omega_\varrho(\xi)} |u|, \sup_{\Omega_\varrho(\xi)} |\mathrm{div}_{\mathbb{G}} b - c| \leq C$. Hence, by the definition of Ω_ϱ given in (3.5.1) we see that

$$\begin{aligned}\int_0^\varrho \int_{\Omega_\varrho(\xi)} s\,\chi_{\Omega_s(\xi)}(x)|(\mathrm{div}_{\mathbb{G}} b(x) - c(x))u(x)|dxds \\ \leq C^2\int_0^\varrho s^{Q-1}\int_{\Omega_s(\xi)}\Gamma^*(x,\xi)dxds \leq C^2\varrho^{Q-1}\int_{\Omega_\varrho(\xi)}\Gamma^*(x,\xi)dxds,\end{aligned}$$

since $\Omega_s(\xi) \subseteq \Omega_\varrho(\xi)$ for every $s < \varrho$ and $\Gamma^*(\cdot,\xi) > 0$ in $\Omega_\varrho(\xi)$. By (3.4.13) the last integral is finite, consequently (3.5.22) holds and we are able to interchange the integration order in (3.5.21), concluding that

$$\begin{aligned}\int_0^\varrho s\Big(\int_{\Omega_s(\xi)}(\mathrm{div}_{\mathbb{G}} b(x) - c(x))u(x)dx\Big)ds \\ = \int_{\Omega_\varrho(\xi)}(\mathrm{div}_{\mathbb{G}} b(x) - c(x))u(x)\int_{\Gamma^*(x,\xi)^{\frac{1}{2-Q}}}^\varrho s\,dsdx \\ = \int_{\Omega_\varrho(\xi)}(\mathrm{div}_{\mathbb{G}} b(x) - c(x))u(x)\left(\frac{\varrho^2 - \Gamma^*(x,\xi)^{\frac{2}{2-Q}}}{2}\right)dx.\end{aligned}$$

□

We are finally in position to prove the Harnack inequality.

Proof of Theorem 3.1.6. Take k_1 as in (3.2.4), δ_0 as in Theorem 3.1.4, r_1 as in Lemma 3.5.19, C_3, r_2 as in (3.5.10), ν, δ, s^* as in Lemma 3.5.21, $C_i > 0$ as in (3.5.11), and set $C := \max_i C_i$, $a := k_1 + \frac{k_1 C}{\delta}$ and $k := k_1\left(\frac{C_2}{C_1}a\delta + 1\right)$. Let $x_0 \in H$ and pick $0 < r_0 < s^*$ such that $B_{kr_0}(x_0) \subseteq H$, $|B_{kr_0}(x_0)| < \delta_0$, and $d(x_1, x_2) \leq r_2$ for every $x_1, x_2 \in B_{r_0}(x_0)$. By the stratification of the algebra (3.2.1) and the Carathéodory–Chow–Rashevsky connectivity theorem, there exists $\bar{n} \in \mathbb{N}$ such that we can connect every $\xi_s, \xi_e \in B_{r_0}(x_0)$ with at most $\bar{n}$ integral path of the generating vector fields $\{X_i\}_{i=1}^{m_1}$. We set $\xi_0 := \xi_s,\ \xi_{\bar{n}+1} := \xi_e$, and $\xi_{k+1} := \gamma_{X_{l(k)},\xi_k}(s_k)$ for every $k = 1, \dots, \bar{n}$, with $l(k) \in \{1, \dots, m_1\}$. By choosing r_0 small enough, it is not restrictive to assume that $|s_k| \leq s^*$ and $|s_k| \leq \frac{C_1 r_1}{C_2 a\delta}$. Then, from (3.2.4) and (3.5.9) it follows that

$$\Omega_{C_1\delta|s_k|}(\xi_{k-1}) \subset B_{\delta|s_k|}(\xi_{k-1}) \subset \Omega_{C_2 a\delta|s_k|}(\xi_k) \subset B_{\frac{C_2}{C_1}a\delta|s_k|}(\xi_k) \subset B_{kr_0}(x_0). \tag{3.5.23}$$

Since $u, \mathrm{div}_{\mathbb{G}} b - c \geq 0$, by (3.5.10), (3.5.12), (3.5.20) and (3.5.23) it holds

$$\int_{\Omega_{C_1\delta|s_k|}(\xi_{k-1})} M_{\mathbb{G}}(x, \xi_{k-1}) u(x) dx \leq \frac{C_3}{\nu} \int_{\Omega_{C_2 a\delta|s_k|}(\xi_k)} M_{\mathbb{G}}(x, \xi_k) u(x) dx,$$

$$\int_0^{C_1\delta|s_k|}\left(s\int_{\Omega_s(\xi_{k-1})} (\mathrm{div}_{\mathbb{G}} b(x) - c(x))\, u(x) dx\right) ds \leq \int_0^{C_2 a\delta|s_k|}\left(s\int_{\Omega_s(\xi_k)} (\mathrm{div}_{\mathbb{G}} b(x) - c(x)) u(x) dx\right) ds.$$

We can assume without loss of generality that $\frac{C_3}{\nu} > 1$, then by (3.5.2) these inequalities imply $u(\xi_{k-1}) \leq c_{x_0} u(\xi_k)$ for every $k = 1, \dots, \bar{n}$, with $c_{x_0} = \frac{C_3}{\nu}\left(\frac{C_2 a}{C_1}\right)^Q$. Hence, we conclude that

$$u(\xi_s) \leq c_H u(\xi_e), \qquad \text{for every } \xi_s, \xi_e \in B_{r_0}(x_0),$$

with $c_H = (c_{x_0})^{\bar{n}}$, implying

$$\sup_{B_{r_0}(x_0)} u \leq c_H \inf_{B_{r_0}(x_0)} u.$$

Finally, by a covering argument (see for instance Theorem 5.7.4 in [27]) it holds

$$\sup_H u \leq C_H \inf_H u,$$

for some $C_H > 0$ that depends on $M_1, M_2, \lambda, \Omega$ and H. $\square$

Chapter 4

Fundamental solution, maximum principle and Harnack inequality for subelliptic evolution operators

4.1 Introduction

In this chapter we analyse the following class of subelliptic evolution operators

$$\mathscr{H}_{\mathbb{G}}u(z) = \sum_{i,j=1}^{m_1} a_{ij}(z)X_iX_ju(z) + \sum_{i=1}^{m_1} b_i(z)X_iu(z) + c(z)u(z) - \partial_t u(z) = f(z), \quad z \in \Omega, \tag{4.1.1}$$

focusing on the regularity theory of the corresponding divergence form operators

$$\mathscr{H}_{\mathbb{G}}^{\text{div}}u(z) = \sum_{i,j=1}^{m_1} X_i\Big(a_{ij}(z)X_j\Big)u(z)+\sum_{i=1}^{m_1} b_i(z)X_iu(z)+c(z)u(z)-\partial_t u(z) = f(z), \quad z \in \Omega, \tag{4.1.2}$$

in some open set $\Omega \subset \mathbb{R}^{N+1}$. As in Chapter 3, we establish the existence of a fundamental solution to (4.1.1), which we then employ to prove a strong maximum principle and an invariant Harnack inequality for (4.1.2). Since our analysis is primarily concerned with the divergence form operator, we restrict our attention to this case. Hence, with a slight abuse of notation, we denote by $\mathscr{H}_{\mathbb{G}}$ the divergence form operator in (4.1.2), and by $\mathscr{H}_{\mathbb{G}}^*$ its adjoint, in order to keep the notation as light as possible.

Throughout the rest of the chapter, $z = (x, t)$ and $\zeta = (\xi, \tau)$ denote points in $\mathbb{R}^N \times \mathbb{R}$, and

X_i's are smooth vector fields on $\mathbb{R}^N$, namely

$$X_i(x) = \sum_{j=1}^{N} \varphi_j^i(x)\partial_{x_j}, \qquad i = 1, \ldots, m_1,$$

with $\varphi_j^i \in C^\infty(\mathbb{R}^N)$. Moreover, $A(z) = \Big(a_{i,j}(z)\Big)_{i,j=1}^{m_1}$ is a $m_1 \times m_1$ symmetric matrix with continuous and bounded real entries that satisfies the usual *uniform ellipticity condition*. We also assume that the low order coefficients $b(z)$ and $c(z)$ are continuous and bounded. Furthermore, we suppose that $m_1 < N$, and therefore $\mathscr{H}_{\mathbb{G}}$ is strongly degenerate. However, we recover some regularity properties for $\mathscr{H}_{\mathbb{G}}$ by assuming that the vector fields $\{X_i\}_{i=1}^{m_1}$ satisfy Hörmander's rank condition.
The vector fields are assumed to satisfy the following hypotheses

(H1) *Hörmander's rank condition holds*

$$\text{rank}\, \mathfrak{g}(x) = N, \qquad \forall x \in \mathbb{R}^N,$$

and for every $i = 1, \ldots, m_1$ *we have* $X_i(0) = \partial_{x_i}$.

(H2) $X_1, \ldots, X_{m_1}$ *are linearly independent smooth vector fields, and*

$$\text{rank}\, \mathfrak{g} = N.$$

(H3) *There exists a family of (non-isotropic) dilations* $\{\delta_r\}_{r>0}$ *on* $\mathbb{R}^N$ *of the kind*

$$\delta_r(x) = \Big(r^{\sigma_1}x_1, r^{\sigma_2}x_2, \ldots, r^{\sigma_N}x_N\Big),$$

where $1 = \sigma_1 \leq \sigma_2 \leq \ldots \leq \sigma_N$ *are integers such that the vector fields* $X_1, \ldots, X_{m_1}$ *are* δ_r*-homogeneous of degree* 1, i.e.

$$X_i\left(f\Big(\delta_r(x)\Big)\right) = r\Big(X_i f\Big)\Big(\delta_r(x)\Big), \qquad \forall r > 0,\ x \in \mathbb{R}^N,\ f \in C^\infty(\mathbb{R}^N),\ i = 1, \ldots, m_1.$$

We recall now Theorem 2.0.1, that asserts that the assumptions **(H1)**, **(H2)** and **(H3)** yield the existence of a Carnot group $\mathbb{G} = \Big(\mathbb{R}^N, \circ, \delta_r\Big)$ (where the dilations $\{\delta_r\}_{r>0}$ are the ones given by **(H3)**) such that $\{X_i\}_{i=1}^{m_1}$ are left-invariant w.r.t. $\circ$, and the operator

$$\Delta_{\mathbb{G}} = \sum_{i=1}^{m_1} X_i^2,$$

is the canonical sublaplacian on $\mathbb{G}$. Despite being degenerate, it is an hypoelliptic operator, namely every distributional solution u to $\Delta_{\mathbb{G}} u = f$ is smooth whenever f is smooth.

We associate to the smooth vector fields some function spaces, that are standard to study classical solutions to equation $\mathscr{H}_{\mathbb{G}} u = f$.

Definition 4.1.1 (Lie derivative)**.** *Let $X = \sum_{j=1}^{N} \varphi_j \partial_{x_j}$ be a smooth vector field on $\mathbb{R}^N$. For every $x \in \mathbb{R}^N$ we define as* integral path of X starting at x *the solution to the following Cauchy problem*

$$\begin{cases} \dot{\gamma}_{X,x}(s) = X\Big(\gamma_{X,x}(s)\Big), \\ \gamma_{X,x}(0) = x. \end{cases}$$

We say that a function u is Lie differentiable at $(x,t) \in \mathbb{R}^{N+1}$ w.r.t. X if the following limit

$$Xu(x,t) = \lim_{s \to 0} \frac{u\Big(\gamma_{X,x}(s), t\Big) - u(x,t)}{s}$$

exists and is finite.

We observe that if Ω is an open subset of $\mathbb{R}^{N+1}$ and $\partial_{x_j} u \in C(\Omega)$ for every $j = 1, \dots, N$, then u is Lie differentiable w.r.t. X and its Lie derivative is given by $Xu = \sum_{j=1}^{N} \varphi_j \partial_{x_j} u$. On the other hand, the converse is not necessarily true.
We introduce the function space

$$C^2_{\mathbb{G}}(\Omega) := \Big\{ u \in C(\Omega) : X_i u, X_i X_j u, \partial_t u \in C(\Omega), \quad \text{for every } \; i,j = 1, \dots, m_1 \Big\},$$

where the derivatives $X_i u$ and $X_i X_j u$ are meant as first and second order Lie derivatives.
We are now ready to give the definition of classical solutions to (4.1.2).

Definition 4.1.2 (Classical solutions for $\mathscr{H}_{\mathbb{G}}$)**.** *Let Ω be an open subset of $\mathbb{R}^{N+1}$ and let $f \in C(\Omega)$. We say that a function u is a classical solution to*

$$\mathscr{H}_{\mathbb{G}} u(z) = \sum_{i,j=1}^{m_1} X_i\Big(a_{ij}(z) X_j\Big) u(z) + \sum_{i=1}^{m_1} b_i(z) X_i u(z) + c(z) u(z) - \partial_t u(z) = f(z), \quad z \in \Omega,$$

if $u \in C^2_{\mathbb{G}}(\Omega)$ and the above equation is satisfied at every point $z \in \Omega$.

As previously said, we construct the fundamental solution to $\mathscr{H}_{\mathbb{G}}$, and then we prove a strong maximum principle and an invariant Harnack inequality. To this end, we improve upon a mean value formula recently derived by Pallara and Polidoro in [92], which relies on the fundamental solution Γ^* of the adjoint operator $\mathscr{H}^*_{\mathbb{G}}$. Therefore, our first step is to prove the existence of Γ^* (together with regularity estimates) by adapting the classical Levi's parametrix method (see Section 4.3 below). To apply this method, we need to assume that the coefficients belong to a Hölder space suited to $\mathscr{H}_{\mathbb{G}}$, that is defined w.r.t. a parabolic distance d related to the Carnot group distance d_X (see (4.2.7) and Definition 4.2.9 in Section

4.2 below). We therefore rely on the following assumptions on the coefficients

(H4) *For every* $i, j = 1, \dots, m_1$ *we have* $a_{ij}, b_i, c, X_j a_{ij}, X_i b_i \in C(\mathbb{R}^{N+1})$*, and there exist* $M_1, M_2 > 0$ *and* $\alpha \in (0,1]$ *such that*

$$
\begin{aligned}
&|a_{ij}(z)|, |b_i(z)|, |c(z)| \le M_1, && \forall z \in \mathbb{R}^{N+1},\\
&|a_{ij}(z) - a_{ij}(\zeta)|, |b_i(z) - b_i(\zeta)|, |c(z) - c(\zeta)| \le M_2\, d(z,\zeta)^\alpha, && \forall z, \zeta \in \mathbb{R}^{N+1},\\
&|X_j a_{ij}(z)|, |X_i b_i(z)| \le M_1, && \forall z \in \mathbb{R}^{N+1},\\
&|X_j a_{ij}(z) - X_j a_{ij}(\zeta)|, |X_i b_i(z) - X_i b_i(\zeta)| \le M_2\, d(z,\zeta)^\alpha, && \forall z, \zeta \in \mathbb{R}^{N+1},
\end{aligned}
\tag{4.1.3}
$$

where $d(z,\zeta)$ *denotes the* $\mathbb{G}$*-parabolic distance between* z *and* ζ *as defined in* (4.2.7)*. Moreover, there exists* $\lambda > 0$ *such that* A *belongs to the uniform ellipticity class*

$$
\boldsymbol{M}_\lambda := \left\{ M \in \mathit{Sym}_{m_1}(\mathbb{R}) : \quad \frac{1}{\lambda}|\xi|^2 \le \langle M(z)\xi, \xi\rangle \le \lambda|\xi|^2, \quad \forall z \in \mathbb{R}^{N+1}, \forall \xi \in \mathbb{R}^{m_1} \right\}. \tag{4.1.4}
$$

Throughout the rest of the chapter, we assume **(H1)**, **(H2)**, **(H3)** and **(H4)**.

One of the main contributions of this chapter is the existence of the global fundamental solution to (4.1.2), the construction of which proceeds by adapting the classical Levi's parametrix method. Along with the existence, we show Gaussian upper and lower bounds on the fundamental solution and its Lie derivatives, as well as many interesting properties regarding the relative Cauchy problems (see Theorem 4.3.11 below). Moreover, we establish the corresponding dual properties for the fundamental solution to the adjoint operator. These results on the fundamental solution are then employed to derive new regularity results for these divergence form operators.
The first regularity result is a maximum principle for (4.1.2). In order to state this result, we need to introduce the definition of *attainable set*.

We say that a curve $\gamma(\tau) = \Big(\gamma_1(\tau), \dots, \gamma_{N+1}(\tau)\Big) : [0,1] \to \mathbb{R}^{N+1}$ is $\mathscr{H}_{\mathbb{G}}$*-admissible* if it is absolutely continuous and solves the following differential equation

$$
\dot\gamma(\tau) = \sum_{i=1}^{m_1} \omega_i(\tau) X_i\Big(\gamma(s)\Big) + \Big(0, \dots, 0, -1\Big),
$$

for a.e. $\tau \in [0,1]$, where $\omega_1, \dots, \omega_{m_1} \in L^\infty\Big([0,1]\Big)$.

Definition 4.1.3 ($\mathscr{H}_{\mathbb{G}}$-attainable set)**.** *Let* Ω *be an open subset of* $\mathbb{R}^{N+1}$*, and let* $\zeta \in \Omega$*. The* $\mathscr{H}_{\mathbb{G}}$*-attainable set of* ζ *in* Ω *is*

$$
\mathscr{A}_\zeta(\Omega) := \Big\{ z \in \Omega : \text{ there exists an } \mathscr{H}_{\mathbb{G}}\text{-admissible curve } \gamma \text{ s.t. } \gamma(0) = \zeta \text{ and } \gamma(1) = z \Big\}.
$$

Theorem 4.1.4. *Assume that **(H1)**,**(H2)**,**(H3)** and **(H4)** hold, and let u be a classical solution to $\mathscr{H}_{\mathbb{G}} u = f$ in some open set $\Omega \subset \mathbb{R}^{N+1}$. Suppose that $c \le 0$, $\mathrm{div}_{\mathbb{G}} b - c \ge 0$ and $f \ge 0$. If there exists $\zeta \in \Omega$ such that $u(\zeta) = \max_\Omega u \ge 0$, then*

$$u(z) = u(\zeta) \quad \textit{and} \quad f(z) = u(\zeta) c(z), \qquad \textit{for every } z \in \overline{\mathscr{A}_\zeta(\Omega)}.$$

An analogous result holds if $u(\zeta) = \min_\Omega u \le 0$ and $f \le 0$. Finally, the assumption on the sign of $u(\zeta)$ can be dropped if $c = 0$.

Remark 4.1.5. *We also point out that, due to the generality of* (4.1.2)*, it is not straightforward to get rid of assumption* $\mathrm{div}_{\mathbb{G}} b - c \ge 0$. *However, this additional hypothesis seems to be natural in this setting, see [92]. In particular, Theorem 4.1.4 improves upon Theorem 2.5 in [92] since it includes the endpoint case* $\mathrm{div}_{\mathbb{G}} b - c = 0$.

The second regularity result is an invariant Harnack inequality for nonnegative solutions to $\mathscr{H}_{\mathbb{G}} u = 0$, where $\mathscr{H}_{\mathbb{G}}$ is the operator in (4.1.2). In order to state the invariant Harnack inequality, we first need to introduce the future and past cylinder where we carry out our analysis. Here and in the following, we denote by $B_r(x_0)$ the $\mathbb{G}$*-ball of radius r and center* x_0 (defined w.r.t. the CC-distance d_X introduced in (4.2.4) below). Correspondingly, we define the $\mathbb{G}$*-cylinder of radius r and center* $z_0 = (x_0, t_0)$ as the set

$$Q_r(z_0) := B_r(x_0) \times \Big(t_0 - r^2, t_0\Big).$$

Moreover, for given constants $\nu, \eta, \mu, \vartheta$ with $0 < \nu < \eta < \mu < 1$ and $0 < \vartheta < 1$, we set

$$Q_r^+(z_0) := B_{\vartheta r}(x_0) \times \Big(t_0 - \nu r^2, t_0\Big), \qquad Q_r^-(z_0) := B_{\vartheta r}(x_0) \times \Big(t_0 - \mu r^2, t_0 - \eta r^2\Big), \quad (4.1.5)$$

see Figure 4.1. We use the notation Q^+ and Q^- whenever $z_0 = 0$ and $r = 1$. Given the definition above, we are in a position to state the following result.

Theorem 4.1.6. *Let Ω be an open subset of $\mathbb{R}^{N+1}$ and let u be a nonnegative solution to $\mathscr{H}_{\mathbb{G}} u = 0$ in Ω under assumptions **(H1)**,**(H2)**,**(H3)** and **(H4)**. Moreover, consider three constants ν, η, μ such that $0 < \nu < \eta < \mu < 1$. Then, there exist $\vartheta_0 \in (0,1)$, only depending on operator $\mathscr{H}_{\mathbb{G}}$, and two positive constants r_0, C_H, depending also on ν, η, μ, such that*

$$\sup_{Q_r^-(z_0)} u \le C_H \inf_{Q_r^+(z_0)} u, \tag{4.1.6}$$

for every $z_0 \in \Omega$, for every $\vartheta \le \vartheta_0$, and for every $r > 0$ such that $r \le r_0$ and $Q_r(z_0) \subset \Omega$.

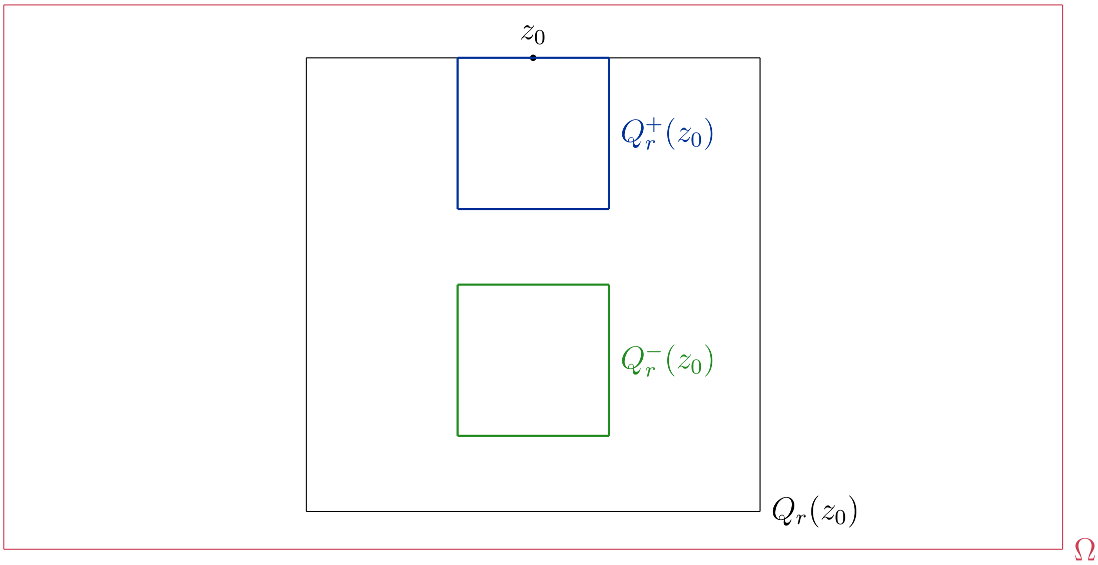

Figure 4.1: Geometric setting of the Harnack inequality (the time variable is represented vertically, and upwardly increasing).

Remark 4.1.7. *Theorem 4.1.6 also applies to classical solutions of the stationary operator*

$$\mathscr{L}_{\mathbb{G}} = \sum_{i,j=1}^{m_1} X_i\Big(a_{ij}X_j\Big) + \sum_{i=1}^{m_1} b_i X_i + c.$$

Hence, the above theorem improves Theorem 3.1.6 in Chapter 3 in that the technical assumption $\mathrm{div}_{\mathbb{G}} b - c \geq 0$ *is removed. Such an assumption is often included when applying the mean value formulas, but can be removed by a standard change of variables in the time variable.*

Before outlining the structure of the chapter, we provide an overview of known results on the maximum principle and the Harnack inequality and we briefly discuss their theoretical interest. We furthermore describe the strategy of the proof of our main results.

A maximum principle for (4.1.2) with C^∞ smooth coefficients was first obtained by Bony in [30] with a barrier argument, and then extended by Amano in [5] for C^1 coefficients. We also mention [29] by Bonfiglioli and Uguzzoni, where the authors proved a maximum principle for (4.1.2) with continuous coefficients under the additional assumption that a Picone function exists. We point out that Theorem 4.1.4 is stronger than [5], and does not require the existence of a Picone function as in [29].
On the other hand, since Bony's work [30] several results for the Harnack inequality are known for operators with smooth coefficients. Concerning operators with Hölder continuous coefficients, we mention Bramanti, Brandolini, Lanconelli and Uguzzoni [34]: here, the authors prove a *parabolic* Harnack inequality for positive solutions to (4.1.2) with $c = 0$,

adapting the method introduced by Krylov and Safonov in [72]. In this chapter we also prove a parabolic Harnack inequality (see Proposition 4.6.15 below) by extending the result of [34] to operators with $c \neq 0$. The main contribution of this chapter is the derivation of an *invariant* Harnack inequality expressed in terms of future and past cylinders, in the same spirit of the one obtained in [85] for uniformly parabolic operators. To the best of the author knowledge, this result is new when dealing with operator (4.1.2). This is obtained by adapting the optimal control strategy introduced in [31], and strongly relies on the sub-Riemannian geometric properties induced by the vector fields $\{X_i\}_{i=1}^{m_1}$.

The proofs of the maximum principle and the parabolic Harnack inequality rely on a technique involving mean value formulas originally developed in the context of harmonic functions (see Chapter 1), though it is not straightforward to adapt this method to our case. We recall that mean value formulas for (4.1.2) have been recently established by Pallara and Polidoro in [92], by employing the fundamental solution Γ^* of the adjoint operator of (4.1.2). In this work the authors managed to extend to the case of Hölder coefficients the results previously known only for smooth coefficients. The main difficulty one encounters when dealing with non-smooth coefficients is due to the loss of regularity of the fundamental solution: while Hörmander's condition theorem ensures smoothness when the coefficients are smooth, the fundamental solution fails to be even C^1 when the coefficients are merely Hölder continuous (as we will see in Definition 4.3.10). To overcome this issue, in [92] the authors used some of the tools developed in the last years in the field of Geometric Measure Theory on Carnot groups (see [110] for an overview of this topic). In order to employ the mean value formula in [92], we first construct the fundamental solution of (4.1.2) and of its adjoint by means of Levi's parametrix method [82, 83] (see Theorem 4.3.11 below). We also derive several key properties of the fundamental solution (most notably, Gaussian estimates), which are then employed to prove the mean value formula in our setting.
The main difficulty encountered in extending this "mean value formula technique" to our subelliptic evolutive setting (in particular, in the proof of the parabolic Harnack inequality) lies in the behavior of the kernel appearing in the formula. More precisely, the kernel is not bounded from above and below by some strictly positive constants (an issue that already arises in the case of the heat operator on $\mathbb{R}^2$). This prevents us from taking advantage of the classical elementary arguments developed for harmonic functions to establish the maximum principle and the parabolic Harnack inequality. To overcome this problem, we employ the technique introduced by Kuptsov in [73], based on Hadamard's descent method, to construct an improved bounded kernel. The resulting kernel is strictly positive, thereby allowing us to make use of the classical theory developed for harmonic functions to prove the maximum principle. Nevertheless, the proof of the parabolic Harnack inequality is more involving, since it is not straightforward to prove the boundedness of the kernel. A parabolic Har-

nack inequality obtained via Hadamard's descent method was first established by Garofalo and Lanconelli in [59] and later by Malagoli, Pallara and Polidoro in [85] for uniformly parabolic operators, and by Polidoro [102] and Rebucci [105] for Kolmogorov operators. In order to prove the boundedness of the kernel we follow the approach introduced by Polidoro in [102], although several technical difficulties need to be addressed in order to obtain the required estimates. This is, in particular, due to the fact that there is no explicit expression for the fundamental solutions Γ associated with the constant coefficient operators within the class under study. In order to tackle the problem, we employ the equivalence of sublaplacians on free Carnot groups derived by Bonfiglioli and Uguzzoni in [28]. To make use of this equivalence on a general Carnot group $\mathbb{G}$ (*i.e.* whenever $\mathbb{G}$ is not free), we apply the Rothschild and Stein lifting technique [106], according to the construction given by Bonfiglioli, Lanconelli and Uguzzoni in [25]. Thanks to this lifting procedure, we can derive the desired estimates and the parabolic Harnack inequality for the lifted operator, which in turn implies the parabolic Harnack inequality for the class of operators under study. The proof of the invariant Harnack inequality presented in Theorem 4.1.6 requires repeated application of the parabolic Harnack inequality to construct a *Harnack chain*. We finally remark that the Carnot–Carathéodory distance associated to the vector fields $\{X_i\}_{i=1}^{m_1}$ plays a key role in this last step.

This chapter is organized as follows: in Section 4.2 we recall some known facts on Carnot groups and the Carnot-Carathéodory distance; in Section 4.3 we prove Theorem 4.3.11 concerning the existence of the fundamental solution of operator $\mathscr{H}_{\mathbb{G}}$, together with Gaussian upper and lower bounds and other properties; in Section 4.4 we introduce the mean value formula and a suitable improvement, that will allow us to prove Theorem 4.1.4 in Section 4.5; in Section 4.6 we make use of the lifting technique to prove a "parabolic-type" Harnack inequality, that will be employed in Section 4.7 to prove the invariant one stated in Theorem 4.1.6; finally, in Appendix A we prove a comparison principle for solutions to (4.1.2) which is used in Section 4.3.

We give some notational remarks. In the following we consistently refrain from restating the phrase 'for $i, j = 1, \dots, m_1$', as it is always implied. Moreover, $\mathcal{H}^d$ will denote the d-dimensional Hausdorff measure, ν the Euclidean inward normal unit vector and $\langle \cdot, \cdot \rangle$ the usual inner product.

4.2 Preliminaries on Carnot groups

In this section we recall some properties concerning the Carnot group that defines the sub-Riemannian setting of (4.1.2). For a more detailed exposition of this topic we refer to the monograph [27]. We subsequently introduce the Carnot-Carathéodory distance associated to the vector fields $X = \{X_1, \ldots, X_{m_1}\}$ in (4.1.2). We eventually describe Rothschild and Stein lifting technique [106] to lift $\mathbb{G}$ to a free group which preserves the homogeneous structure of $\mathbb{G}$.

Let $\mathbb{G} = \left(\mathbb{R}^N, \circ, \delta_r\right)$ be the Carnot group induced by the smooth vector fields $\{X_i\}_{i=1}^{m_1}$ (whose existence is given by Theorem 1.7 in [22], thanks to **(H1)**,**(H2)** and **(H3)**). We first recall that a Carnot group is in particular homogeneous Lie group, namely a group with a smooth composition law $\circ$ and a dilation law $\{\delta_r\}_{r>0}$ that is an automorphism of the group, namely

$$\delta_r(x \circ y) = \left(\delta_r(x)\right) \circ \left(\delta_r(y)\right), \qquad \text{for every } x, y \in \mathbb{R}^N,\ r > 0.$$

By **(H1)**, **(H2)** and **(H3)**, $\mathfrak{g}$ is the Lie algebra of the group $\mathbb{G}$. Moreover, **(H3)** induces a stratification on $\mathfrak{g}$

$$\mathfrak{g} = V_1 \oplus \ldots \oplus V_\kappa, \tag{4.2.1}$$

where $V_1 = \text{span}\{X_i\}_{i=1}^{m_1}$, $V_{k+1} = \text{span}\Big\{[X, Y], X \in V_1, Y \in V_k\Big\}$, for $k = 1, \ldots, \kappa - 1$, and $\Big\{[X, Y], X \in V_1, Y \in V_\kappa\Big\} = \emptyset$, with $[X, Y] = XY - YX$. We observe that the stratification (4.2.1) implies that the dilations $\{\delta_r\}_{r>0}$ take the following form (see Remark 1.4.2 in [27])

$$\delta_r(x) = \left(r x^{(1)}, r^2 x^{(2)}, \ldots, r^\kappa x^{(\kappa)}\right), \qquad x^{(k)} \in \mathbb{R}^{m_k},\ k = 1, \ldots, \kappa, \tag{4.2.2}$$

where m_k denotes the dimension of V_k, for every $k = 1, \ldots, \kappa$. We will refer to the number

$$Q := \sum_{k=1}^{\kappa} k\, m_k,$$

as homogeneous dimension of $\mathbb{G}$. Hörmander's rank condition implies that $\sum_{k=1}^{\kappa} m_k = N$, inferring that $Q \geq N$. In particular, the equality holds if and only if $\mathbb{G}$ is the Euclidean group. A group $\mathbb{G}$ with the properties above is often referred to as Carnot group of step κ and m_1 generators.

As it will be useful in Section 4.3, we recall that Proposition 1.3.5 in [27] provides us with the following characterization of smooth δ_r-homogeneous vector fields: let δ_r be as in (4.2.2) and let $X = \sum_{j=1}^{N} \varphi_j \partial_{x_j}$ be a smooth vector field on $\mathbb{R}^N$. Then X is δ_r-homogeneous of degree $n \in \mathbb{R}$ if and only if φ_j is a polynomial function δ_r-homogeneous of degree $\sigma_j - n$ (unless $\varphi_j \equiv 0$). In particular, by **(H3)** we deduce that the coefficients φ_j^i of the smooth vector fields $\{X_i\}_{i=1}^{m_1}$ are polynomial functions.

We now introduce a homogeneous distance on $\mathbb{G}$, *i.e.* the X-*control distance* (also known as the *Carnot–Carathéodory distance*) associated with the family of vector fields $X =$

$\{X_1, \dots, X_{m_1}\}$. This distance will play a crucial role in the proof of the maximum propagation and the Harnack inequality stated in Theorem 4.1.4 and in Theorem 4.1.6, respectively.

Definition 4.2.8 (Carnot-Carathéodory distance)**.** *The Carnot-Carathéodory distance (CC, in short) d_X associated to the set of Hörmander vector fields $X = \{X_1, \dots, X_{m_1}\}$ is defined as follows. Let $\mathcal{C}$ denote the set of the absolutely continuous maps $\gamma : [0,1] \to \mathbb{R}^N$ satisfying*

$$\dot{\gamma}(\tau) = \sum_{i=1}^{m_1} \alpha_i(\tau) X_i\big(\gamma(\tau)\big),$$

for a.e. $\tau \in [0,1]$, with $\alpha = (\alpha_1, \dots, \alpha_{m_1})$ real-valued function on $[0,1]$ such that

$$\|\alpha\|_\infty := \sup_{\tau \in [0,1]} \left(\sum_{j=1}^{m_1} |\alpha_j(\tau)|^2 \right)^{\frac{1}{2}} < +\infty. \tag{4.2.3}$$

The CC-distance is then defined as

$$d_X(x,y) := \inf \left\{ \int_0^1 |\alpha(\tau)| d\tau \,:\, \gamma \in \mathcal{C},\, \gamma(0) = x,\ \gamma(1) = y \right\}. \tag{4.2.4}$$

In the following we will refer to $\mathcal{C}$ as the set of X-trajectories on $\mathbb{R}^N$*, and the function $\tau \mapsto \alpha(\tau)$ will be referred to as* control *of the X-trajectory γ.*

We now list some properties of the CC-distance that will be useful in the sequel.

- The CC-distance can equivalently be defined in terms of *X-subunit curves*, that is, by requiring $\|\alpha\|_\infty \le 1$ (see, *e.g.*, Proposition 7.12 in [19]).
- Another equivalent definition of the CC-distance is the following

$$d_X^{(p)}(x,y) = \begin{cases} \inf\limits_{\gamma \in \mathcal{C}} \left(\int_0^1 |\alpha(\tau)|^p \, d\tau \right)^{1/p} & \text{if } p < +\infty, \\ \inf\limits_{\gamma \in \mathcal{C}} \sup\limits_{\tau \in [0,1]} |\alpha(\tau)|, & \text{if } p = +\infty. \end{cases}$$

  It is clear that $d_X \equiv d_X^{(1)}$. Moreover, the following property holds

$$d_X^{(p)} = d_X, \qquad \forall p \in [1, +\infty]. \tag{4.2.5}$$

  see for instance Proposition 7.13 in [19].

- The CC-distance associated to a system of Hörmander vector fields is a metric (see Theorem 6.22 in [19]).

- There exists $k_1 \geq 1$ such that

$$d_X(x,\xi) \leq k_1\Big(d_X(x,y) + d_X(y,\xi)\Big), \qquad \text{for every } x, \xi, y \in \mathbb{R}^N. \tag{4.2.6}$$

  If $k_1 = 1$, this is the classical triangular inequality; for such reason this property is usual referred to as *pseudo-triangular inequality*.

- The CC-distance is δ_r-homogeneous of degree 1, hence the Euclidean metric is continuos w.r.t. the metric induced by the CC-distance (see Proposition 5.15.1 in [26]).

In order to study the evolution operator (4.1.2) on the Carnot group $\mathbb{G}$ (that is relevant to the space variable $x \in \mathbb{R}^N$), we introduce the $\mathbb{G}$-parabolic distance on $\mathbb{R}^{N+1}$

$$d(z,\zeta) = d\Big((x,t),(\xi,\tau)\Big) := d_X(x,\xi) + |t-\tau|^{\frac{1}{2}}, \qquad \forall z = (x,t), \zeta = (\xi,\tau) \in \mathbb{R}^{N+1}. \tag{4.2.7}$$

We are now in a position to define the Hölder spaces naturally induced by the differential operator $\mathscr{H}_{\mathbb{G}}$.

Definition 4.2.9 (Hölder continuous function). *Let $\alpha \in (0,1]$. We say that $u : \mathbb{R}^{N+1} \to \mathbb{R}$ is an α-Hölder continuous function with exponent α in $\mathbb{R}^{N+1}$ if there exists a positive constant $c > 0$ such that*

$$|u(z) - u(\zeta)| \leq c\, d(z,\zeta)^\alpha, \qquad \forall z, \zeta \in \mathbb{R}^{N+1}.$$

This definition of Hölder space plays a central role in our analysis: in Section 4.3 we prove that, under the assumption that the coefficients belong to this Hölder space (namely **(H4)**), we are able to apply the parametrix method. Hence, we are able to prove the existence of the fundamental solution Γ to $\mathscr{H}_{\mathbb{G}}$ as well as the fundamental solution Γ^* to $\mathscr{H}^*_{\mathbb{G}}$, along with upper and lower bounds. This, in turn, enables us to use the mean value formula developed in [92], which we then refine in Section 4.4 and exploit in the proof of Theorem 4.1.4 and Theorem 4.1.6. Mean value formulas hold for divergence form operators: accordingly, we recall the definition of the differential operators $\mathrm{div}_{\mathbb{G}}$ and $\nabla_{\mathbb{G}}$, that are the natural extension of the divergence and gradient operator in the non-Euclidean setting of $\mathbb{G}$. In order to define these operators, we agree to identify a horizontal section $F = \sum_{i=1}^{m_1} F_i X_i$ with its canonical coordinates $F = (F_1, \dots, F_{m_1})$; moreover, we define as $C^1_{\mathbb{G}}$ the space of functions with continuous first order Lie derivatives. Then, for every $F \in C^1_{\mathbb{G}}\Big(\mathbb{R}^{N+1}, \mathbb{R}^{m_1}\Big)$ and for every real valued $f \in C^1_{\mathbb{G}}(\mathbb{R}^{N+1})$ we set

$$\mathrm{div}_{\mathbb{G}} F := \sum_{i=1}^{m_1} X_i F_i, \qquad \nabla_{\mathbb{G}} f := \sum_{i=1}^{m_1} \Big(X_i f\Big) X_i. \tag{4.2.8}$$

By employing the operators defined above, equation (4.1.2) can be recast in $\mathbb{G}$-divergence form as follows

$$\mathscr{H}_{\mathbb{G}} u = \operatorname{div}_{\mathbb{G}}\Big(A\nabla_{\mathbb{G}} u\Big) + \langle b, \nabla_{\mathbb{G}} u\rangle + c\,u - \partial_t u = f.$$

We conclude this section by introducing the lifting technique that will be employed in Section 4.6 to establish a parabolic-type Harnack inequality. We first recall that $\mathbb{G}$ is a free Carnot group if the generating vector fields $\{X_i\}_{i=1}^{m_1}$ and their commutators do not satisfy any linear relation other then those which hold automatically as a consequence of the antisymmetry of the Lie bracket and the Jacobi identity (see Definition 2.0.2). Following Theorem 2.0.3, a Carnot group $\mathbb{G}$ of homogeneous dimension Q with generators $\{X_i\}_{i=1}^{m_1}$ can be lifted to a free Carnot group $\widetilde{\mathbb{G}}$ in the sense of Rothschild and Stein [106]. More precisely, there exists $\widetilde{N} \in \mathbb{N}$ and a free Carnot group $\widetilde{\mathbb{G}} = \Big(\mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N, \widetilde{\circ}, \widetilde{\delta}_r\Big)$ with homogeneous dimension $\widetilde{Q}$ and Jacobian basis $\Big\{\widetilde{X}_i\Big\}_{i=1}^{m_1}$, such that the following lifting property holds: let $\pi : \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \to \mathbb{R}^N$ denote the canonical projection onto the last N coordinates, then

$$\widetilde{X}_i u\Big(\pi(\widetilde{x}, x)\Big) = X_i u(x), \qquad (\widetilde{x}, x) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N, \tag{4.2.9}$$

for every $u \in C^\infty(\mathbb{R}^N)$.
We now choose a suitable homogeneous distance on $\widetilde{\mathbb{G}}$. Since the Lie algebra of $\widetilde{\mathbb{G}}$ is stratified as in (4.2.1), we have the following decomposition

$$\mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N = \mathbb{R}^{\widetilde{m}_1} \times \mathbb{R}^{\widetilde{m}_2} \times \ldots \times \mathbb{R}^{\widetilde{m}_\kappa} \times \mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \ldots \times \mathbb{R}^{m_\kappa},$$

for some $\widetilde{m}_1, \ldots, \widetilde{m}_\kappa \in \mathbb{N}$, where $m_1, \ldots, m_\kappa$ are as in (4.2.2). Accordingly, we decompose any $(\widetilde{x}, x) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N$ as

$$\Big(\widetilde{x}^{(1)}, \widetilde{x}^{(2)}, \ldots, \widetilde{x}^{(\kappa)}, x^{(1)}, x^{(2)}, \ldots, x^{(\kappa)}\Big) \qquad \widetilde{x}^{(k)} \in \mathbb{R}^{\widetilde{m}_k}, x^{(k)} \in \mathbb{R}^{m_k},\ k = 1, \ldots, \kappa,$$

and we set

$$d_{\widetilde{\mathbb{G}}}(\widetilde{x}, x) := \left(\sum_{i=1}^{\kappa}\sum_{j=1}^{\widetilde{m}_i} |\widetilde{x}_j^{(i)}|^{\frac{2}{i}}\right)^{\frac{1}{2}} + d_X(x, 0), \tag{4.2.10}$$

where d_X is the CC-distance on $\mathbb{G}$ introduced in Definition 4.2.8.
If $\mathbb{G}$ is a free Carnot group, the following remarkable property holds: let $A = (a_{ij})_{i,j=1}^{m_1}$ be a symmetric, positive definite matrix with constant entries; then, there exists a diffeomorphism $T_A : \mathbb{R}^N \to \mathbb{R}^N$ such that

$$\Gamma_A\Big((x,t), (\xi, \tau)\Big) = J_A\, \Gamma_0\Big(T_A\Big(\xi^{-1} \circ x\Big), t - \tau\Big),$$

where Γ_0 denote the fundamental solution of the canonical heat operator $\Delta_{\mathbb{G}} - \partial_t$, while Γ_A

is the fundamental solution of of the constant coefficients operator

$$\mathscr{H}_A = \sum_{i,j=1}^{m_1} a_{ij} X_i X_j - \partial_t,$$

and J_A is the Jacobian determinant of T_A. It is clear that T_A acts as a classical change of basis in our subelliptic setting. We remark that it is possible to derive ad hoc uniform estimates for the diffeomorphism T_A, see for instance [28]. Hence, lifting $\mathbb{G}$ to $\widetilde{\mathbb{G}}$ allows us to take advantage of the previous correspondence for the lifted fundamental solution $\widetilde{\Gamma}_A$ (see equation (4.6.10)) to derive precise estimates for $\widetilde{\Gamma}_A$.

4.3 Fundamental solution

In this section we prove the existence and some useful properties of the fundamental solution Γ associated to operator $\mathscr{H}_{\mathbb{G}}$. As previously mentioned, the main tool that allows us to employ the mean value formula in [92] and to prove our main results is the fundamental solution Γ^* of the adjoint operator

$$\mathscr{H}_{\mathbb{G}}^* u = \sum_{i,j=1}^{m_1} X_i \Big(a_{ij} X_j \Big) u - \sum_{i=1}^{m_1} b_i X_i u + \left(c - \sum_{i=1}^{m_1} X_i b_i \right) u + \partial_t u. \tag{4.3.1}$$

We give now the definition of fundamental solution for (4.1.2) and for (4.3.1).

Definition 4.3.10 (Fundamental solution)**.** *We say that the real valued function Γ defined on $\Big(\mathbb{R}^{N+1} \times \mathbb{R}^{N+1}\Big) \setminus \Big\{(z,z) \in \mathbb{R}^{N+1} \times \mathbb{R}^{N+1}\Big\}$ is a fundamental solution for* (4.1.2) *if:*

1. *for every $\zeta \in \mathbb{R}^{N+1}$ the function $\Gamma(\cdot,\zeta)$ is a classical solution to $\mathscr{H}_{\mathbb{G}}\Gamma(\cdot,\zeta) = 0$ in $\mathbb{R}^{N+1} \setminus \{\zeta\}$, and belongs to $L^1_{\text{loc}}\Big(\mathbb{R}^{N+1} \setminus \{\zeta\}\Big)$;*
2. *for every $\varphi \in C_c(\mathbb{R}^N)$ the function*

$$u(x,t) := \int_{\mathbb{R}^N} \Gamma(x,t,\xi,\tau)\varphi(\xi) d\xi, \qquad (x,t) \in \mathbb{R}^N,$$

is a classical solution to $\mathscr{H}_{\mathbb{G}} u = 0$ in $\mathbb{R}^N \times (\tau, +\infty)$, and satisfies

$$\lim_{(x,t)\to(\xi,\tau)} u(x,t) = \varphi(\xi), \qquad \textit{for every } \ \xi \in \mathbb{R}^N.$$

We say that Γ^* is a fundamental solution for (4.3.1) *if it satisfies the dual statement.*

We now focus our attention on proving the existence of Γ and Γ^* under the assumption that the coefficients are Hölder continuous. In addition, we establish several fundamental properties that will play a crucial role in our analysis.

Theorem 4.3.11 (Fundamental solution)**.** *Assume that **(H1)**,**(H2)**,**(H3)** and **(H4)** hold. Then*

1. *There exists a fundamental solution* Γ *to* (4.1.2).

2. *The following Gaussian upper bounds hold for every* $T > 0$ *and* $0 < t - \tau \leq T$

$$\Gamma(x,t,\xi,\tau) \leq \frac{c_T}{\left(t-\tau\right)^{\frac{Q}{2}}} \, \exp\left(-c_u \frac{d_X(x,\xi)^2}{(t-\tau)}\right), \tag{4.3.2}$$

$$|X_i\Gamma(x,t,\xi,\tau)| \leq \frac{c_T}{\left(t-\tau\right)^{\frac{Q+1}{2}}} \exp\left(-c_u \frac{d_X(x,\xi)^2}{(t-\tau)}\right), \tag{4.3.3}$$

$$|X_iX_j\Gamma(x,t,\xi,\tau)|, |\partial_t\Gamma(x,t,\xi,\tau)| \leq \frac{c_T}{\left(t-\tau\right)^{\frac{Q+2}{2}}} \exp\left(-c_u \frac{d_X(x,\xi)^2}{(t-\tau)}\right), \tag{4.3.4}$$

for some $c_u > 0$ *(depending only on structural assumptions) and* $c_T > 0$ *(depending also on* T*).*

3. *For any* $(x,t),(\xi,\tau) \in \mathbb{R}^{N+1}$ *with* $t \leq \tau$*, we have that* $\Gamma(x,t,\xi,\tau) = 0$.

4. *Let* f *be a continuous function on* $\mathbb{R}^N \times (T_1,T_2)$ *such that for every* x,ξ *belonging to a compact set* $\mathcal{K} \subset \mathbb{R}^N$*, and for every* $t \in (T_1,T_2)$*, there exists* $c_{\mathcal{K}} = c_{\mathcal{K}}(\mathcal{K}) > 0$ *and* $0 < \eta \leq 1$ *satisfying*

$$|f(x,t) - f(\xi,t)| \leq c_{\mathcal{K}}\, d_X(x,\xi)^\eta.$$

Moreover, suppose that f *has an exponential growth uniform in* t*, namely that there exist* $h_1,h_2 > 0$ *such that*

$$|f(x,t)| \leq h_2 \, \exp\left(h_1 d_X(x,0)^2\right). \tag{4.3.5}$$

Let g *be a continuous function on* $\mathbb{R}^N$ *with the same exponential growth*

$$|g(x)| \leq h_2 \, \exp\left(h_1 d_X(x,0)^2\right). \tag{4.3.6}$$

Then, there exists $T_1 < T < T_2$ *(only depending on the growth constant* h_1*) such that the function*

$$u(x,t) := \int_{\mathbb{R}^N} \Gamma(x,t,\xi,T_1)g(\xi)d\xi - \int_{T_1}^{t}\int_{\mathbb{R}^N} \Gamma(x,t,\xi,\tau)f(\xi,\tau)d\xi d\tau, \tag{4.3.7}$$

is a classical solution to

$$\begin{cases} \mathscr{H}_{\mathbb{G}} u = f, & \text{in } \mathbb{R}^N \times (T_1, T), \\ u(x, T_1) = g(x), & \text{in } \mathbb{R}^N, \end{cases} \tag{4.3.8}$$

as it solves $\mathscr{H}_{\mathbb{G}} u = f$ according to Definition 4.1.2, and it takes the initial condition in the following sense

$$\lim_{t \to T_1} \left(\int_{\mathbb{R}^N} \Gamma(x, t, \xi, T_1) g(\xi) d\xi - \int_{T_1}^{t} \int_{\mathbb{R}^N} \Gamma(x, t, \xi, \tau) f(\xi, \tau) d\xi d\tau \right) = g(x). \tag{4.3.9}$$

5. *Γ is nonnegative, and moreover there exists $S > 0$ (depending only on structural assumptions) such that $\sup_{t \geq \tau+1} \Gamma(x, t, \xi, \tau) = S$.*

6. *If $c(z) = c$ is constant, then*

$$\int_{\mathbb{R}^N} \Gamma(x, t, \xi, \tau) d\xi = \exp\Big(c\,(t - \tau)\Big), \qquad \forall x \in \mathbb{R}^N, \tau < t. \tag{4.3.10}$$

7. *The reproduction property holds: for every $x, \xi \in \mathbb{R}^N$ and $t, \tau \in \mathbb{R}$ with $\tau < s < t$ we have*

$$\Gamma(x, t, \xi, \tau) = \int_{\mathbb{R}^N} \Gamma(x, t, y, s) \Gamma(y, s, \xi, \tau) dy. \tag{4.3.11}$$

8. *The following Gaussian lower bound holds for every $T > 0$ and $0 < t - \tau \leq T$*

$$\Gamma(x, t, \xi, \tau) \geq \frac{\widetilde{c}_T}{\Big(t - \tau\Big)^{\frac{Q}{2}}} \exp\left(-c_l \frac{d_X(x, \xi)^2}{(t - \tau)} \right), \tag{4.3.12}$$

for some $c_l, \widetilde{c}_T > 0$ (depending on structural assumptions and T).

9. *If the exponential growth of f and g in* (4.3.5) *and* (4.3.6) *satisfies*

$$\min\left(\frac{c_u}{2h_1(2k_1^2 - 1)}, \frac{3c_u}{h_1} \right) > T. \tag{4.3.13}$$

where T is the constant given by (4.)*, c_u is the same as in* (2.)*, and k_1 is given by* (4.2.6)*, then there exist $C, M > 0$ such that for every $(x, t) \in \mathbb{R}^N \times (T_1, T)$ the function u defined in* (4.3.7) *verifies*

$$|u(x, t)| \leq C \exp\Big(M d_X(x, 0)^2 \Big). \tag{4.3.14}$$

Furthermore, there exists at most one classical solution u *to* (4.3.8) *such that*

$$\int_{T_1}^{T}\int_{\mathbb{R}^N} |u(x,t)| \exp\Big(-k\, d_X(x,0)^2\Big) dxdt < +\infty, \tag{4.3.15}$$

for some $k > 0$.

10. *There exists a fundamental solution* Γ^* *to* (4.3.1) *satisfying the dual properties of the previous statements. Moreover, it holds*

$$\Gamma^*(x,t,\xi,\tau) = \Gamma(\xi,\tau,x,t), \qquad \textit{for every} \quad (x,t),(\xi,\tau) \in \mathbb{R}^{N+1}, (x,t) \neq (\xi,\tau). \tag{4.3.16}$$

As mentioned above, the existence of the fundamental solution (and more precisely properties *(1.)-(4.)*) follows from the parametrix method introduced by Levi in [82] and [83], and here adapted to our subelliptic evolutive setting. Let us briefly outline the parametrix technique for evolution operators: for every $\zeta \in \mathbb{R}^{N+1}$, we evaluate the coefficients of the principal part matrix A at the point ζ, and we denote by $\Gamma_{A_\zeta} = \Gamma_{A_\zeta}(\cdot,\zeta)$ the fundamental solution of the following frozen operator

$$\mathscr{H}_{A_\zeta} = \sum_{i,j=1}^{m_1} a_{ij}(\zeta) X_i X_j - \partial_t. \tag{4.3.17}$$

We call parametrix the function $Z(z,\zeta) := \Gamma_{A_\zeta}(z,\zeta)$. We look for a fundamental solution Γ as a solution to the following equation

$$\Gamma(z,\zeta) = Z(z,\zeta) + J(z,\zeta), \tag{4.3.18}$$

where J is a correction term that we assume of the following form

$$J(z,\zeta) = \int_\tau^t \int_{\mathbb{R}^N} Z(x,t,y,s) G(y,s,\xi,\tau) dyds, \tag{4.3.19}$$

for some unknown function G. In particular, we construct a function G such that Γ is a classical solution to

$$\mathscr{H}_{\mathbb{G}}\Gamma(z,\zeta) = \mathscr{H}_{\mathbb{G}} Z(z,\zeta) + \mathscr{H}_{\mathbb{G}} J(z,\zeta) = 0, \qquad \text{for every } z \neq \zeta. \tag{4.3.20}$$

Moreover, assuming that we can differentiate J under the integral sign, we get

$$\mathscr{H}_{\mathbb{G}} J(z,\zeta) = \int_\tau^t \int_{\mathbb{R}^N} \mathscr{H}_{\mathbb{G}} Z(x,t,y,s) G(y,s,\xi,\tau) dyds - G(z,\zeta). \tag{4.3.21}$$

Combining (4.3.20) and (4.3.21) we get the following equation for G

$$G(z,\zeta) = \mathscr{H}_{\mathbb{G}} Z(z,\zeta) + \int_\tau^t \int_{\mathbb{R}^N} \mathscr{H}_{\mathbb{G}} Z(x,t,y,s) G(y,s,\xi,\tau) dyds, \tag{4.3.22}$$

which is a Fredholm integral equation of the second kind. We solve (4.3.22) by means of successive approximation method, obtaining

$$G(z,\zeta) = \sum_{k=1}^{\infty} (\mathscr{H}_{\mathbb{G}} Z)_k(z,\zeta), \tag{4.3.23}$$

where $(\mathscr{H}_{\mathbb{G}} Z)_1 := \mathscr{H}_{\mathbb{G}} Z$, and

$$(\mathscr{H}_{\mathbb{G}} Z)_{k+1}(z,\zeta) := \int_{\tau}^{t} \int_{\mathbb{R}^N} (\mathscr{H}_{\mathbb{G}} Z)_1(x,t,y,s)(\mathscr{H}_{\mathbb{G}} Z)_k(y,s,\xi,\tau)dyds. \tag{4.3.24}$$

To prove the well-posedness of the parametrix method, we must verify that the integral (4.3.24) is properly defined, that the series (4.3.23) converges, and that the identity (4.3.21) is satisfied, thereby ensuring the validity of (4.3.20). As it can be seen from (4.3.4) and (4.3.24), the Hölder continuity assumption is crucial to address the existence of some singular integrals.

Once the existence of the fundamental solution is established, the proof of properties *(5.)*-*(10.)* follows from classical PDEs techniques. We emphasize that the proof of the lower bound (4.3.12) strongly depends on the parametrix construction of Γ. Indeed, this method enables us to establish a *local* lower bound (see (4.3.33) below), that, by means of the technique developed in [66], we are able to turn into a *global* one.

We remark again that the proof of properties *(1.)*-*(8.)* still goes through if we consider the non-divergence form operator

$$\mathscr{H} u = \sum_{i,j=1}^{m_1} \bar{a}_{ij} X_i X_j u + \sum_{i=1}^{m_1} \bar{b}_i X_i u + \bar{c}\, u,$$

whenever $\bar{a}$, $\bar{b}$ and $\bar{c}$ are continuous and satisfy (4.1.3).

Proof of Theorem 4.3.11. We recall for reader's convenience that we aim to prove the existence of the fundamental solution to operators (4.1.2) and (4.3.1), *i.e.*

$$\mathscr{H}_{\mathbb{G}} u = \sum_{i,j=1}^{m_1} a_{ij} X_i X_j u + \sum_{i=1}^{m_1} \left(\sum_{j=1}^{m_1} X_j a_{ij} + b_i \right) X_i u + cu - \partial_t u,$$
$$\mathscr{H}_{\mathbb{G}}^* u = \sum_{i,j=1}^{m_1} a_{ij} X_i X_j u + \sum_{i=1}^{m_1} \left(\sum_{j=1}^{m_1} X_j a_{ij} - b_i \right) X_i u + \left(c - \sum_{i=1}^{m_1} X_i b_i \right) u + \partial_t u.$$

(1.)-(5.)

Since the proof of *(1.)-(5.)* follows very closely the one presented in [26], here we only give a sketch of it and we refer the interested reader to [26] for details. In the following we focus our attention on to the fundamental solution Γ of $\mathscr{H}_{\mathbb{G}}$. Clearly, the fundamental solution Γ^*

of the adjoint equation $\mathscr{H}_{\mathbb{G}}^*$ satisfies the dual statement.

The starting point is to show the existence of a smooth fundamental solution to (4.3.17), hence the existence of the parametrix Z. This is achieved thanks to Theorem 2.5 in [25]. To be more precise, Theorem 2.5 in [25] states that the fundamental solution of constant coefficient operators (hence the parametrix) exists and satisfies properties *(2.)-(10.)* with upper and lower Gaussian bounds

$$\frac{1}{c_\lambda\Big(t-\tau\Big)^{\frac{Q}{2}}}\exp\left(-c_\lambda\frac{d_X(x,\xi)^2}{(t-\tau)}\right)\leq \Gamma_{A_\zeta}(z,\zeta)\leq \frac{c_\lambda}{\Big(t-\tau\Big)^{\frac{Q}{2}}}\exp\left(-\frac{d_X(x,\xi)^2}{c_\lambda(t-\tau)}\right), \quad (4.3.25)$$

that hold for every $z,\zeta\in\mathbb{R}^{N+1}$. In addition, similar Gaussian bounds hold also for its derivatives. Crucially, c_λ depends only on the ellipticity constant λ given by **(H4)**. In particular, the Gaussian estimates for the parametrix are *uniform*, as they do not depend on the point $\zeta\in\mathbb{R}^{N+1}$ where we compute (4.3.17).
The core principle underlying Levi's method is the following observation: since $Z(z,\zeta)$ is the fundamental solution associated to (4.3.17), it holds

$$\begin{aligned}\mathscr{H}_{\mathbb{G}}Z(z,\zeta) &= \sum_{i,j=1}^{m_1}\Big(a_{ij}(z)-a_{ij}(\zeta)\Big)X_iX_jZ(z,\zeta)\\ &\quad+\sum_{i=1}^{m_1}\left(\sum_{j=1}^{m_1}X_ja_{ij}(z)+b_i(z)\right)X_iZ(z,\zeta)+c(z)Z(z,\zeta).\end{aligned}$$

Thanks to the Hölder continuity assumption, the Gaussian upper bounds ensure the integrability required to make the integral in (4.3.24) well-defined, and to guarantee that the series in (4.3.23) converges and provides a solution to (4.3.22) (without requiring an analogous of (3.4.9), see Proposition 2.1 and Corollary 2.2 in [26]). Moreover, the following estimate

$$|G(z,\zeta)|\leq \frac{c_T(t-\tau)^{\frac{\alpha}{2}-1}}{(t-\tau)^{\frac{Q}{2}}}\exp\left(-\frac{d_X(x,\xi)^2}{c(t-\tau)}\right), \quad (4.3.26)$$

holds true for some $c>0$ depending only on structural assumptions, while c_T depends also on T. By (4.3.25) and (4.3.26), together with the Lagrange mean value theorem on integral paths, we can prove that Γ is 'almost' a fundamental solution (see Theorem 2.7 in [26] for further details). More precisely, for every $\zeta\in\mathbb{R}^{N+1}$ the function $\Gamma(\cdot,\zeta)$ belongs to the weaker regularity space $C^2_{x,t}$, which is the space of continuous functions w such that $w(\cdot,t)$ has continuous Lie derivatives X_iX_j for every fixed t, and $w(x,\cdot)$ has continuous time derivative ∂_t for every fixed x. Moreover, $\mathscr{H}_{\mathbb{G}}\Gamma(z,\zeta)=0$ in $\mathbb{R}^{N+1}\setminus\{\zeta\}$, and if f and g satisfy the properties of statement *(4.)*, it follows that the function u defined in (4.3.7) belongs to $C^2_{x,t}$ and solves the Cauchy problem (4.3.8) (see Theorem 3.1 in [26]). By (4.3.26)

and the Gaussian estimates for the parametrix, we also infer that Γ satisfies the Gaussian upper bounds (4.3.2), (4.3.3) and (4.3.4).
Hence, to complete the proof of *(1.)-(4.)* we need to recover the $C^2_{\mathbb{G}}$ regularity for Γ (since it implies the same regularity for the function u defined in (4.3.7)). This regularity result follows from a combination of the comparison principle stated in Proposition A.1 (see Appendix (A) below), together with a regularization argument and the Schauder estimates proved in [32] (see Corollary 4.8, Corollary 4.9 and the proof of Theorem 1.2 in [26] for more details). In addition, the comparison principle implies also *(5.)* (see for instance Proposition 4.4 in [26]).

(6.)
We prove (4.3.10) when $c = 0$, the general case follows by a standard change of variable.
Firstly, we show that for every $\varphi \in C_c^\infty(\mathbb{R}^{N+1})$ such that $\sup_{(y,s),(\xi,\tau)\in \mathrm{supp}(\varphi)} |s-\tau| < \varepsilon$, and $0 \le \varphi \le 1$, it holds

$$u(x,t) := \int_{\mathbb{R}^{N+1}} \Gamma(x,t,\xi,\tau)\varphi(\xi,\tau)d\xi d\tau \le \varepsilon, \qquad \forall (x,t) \in \mathbb{R}^{N+1}. \tag{4.3.27}$$

Suppose, without loss of generality, that

$$0 = \min\Big\{ s \in \mathbb{R} : (y,s) \in \mathrm{supp}(\varphi) \text{ for some } y \in \mathbb{R}^N \Big\}. \tag{4.3.28}$$

Since φ is smooth with compact support, by *(4.)* we infer that $\mathscr{H}_{\mathbb{G}} u(x,t) = -\varphi(x,t)$, by *(3.)* and (4.3.28) we deduce that $u(x,t) = 0$ for every $t \le 0$, and by (4.3.2) we see that u vanishes as $|x| \to +\infty$, thanks to the continuity of the Euclidean metric w.r.t. the CC-metric. Thus, for every $(x,t) \in \mathbb{R}^{N+1}$ we set $v(x,t) := t - u(x,t)$, and we deduce that it solves

$$\begin{cases} \mathscr{H}_{\mathbb{G}} v(x,t) = -1 + \varphi(x,t) \le 0, & \text{in } \mathbb{R}^N \times (0,\varepsilon), \\ v(x,0) = 0, & \text{in } \mathbb{R}^N, \\ \lim\limits_{|x|\to+\infty} v(x,t) = t \ge 0, & \text{in } (0,\varepsilon). \end{cases}$$

By comparison principle Proposition A.1, we infer that $v(x,t) \ge 0$, thus implying $u(x,t) \le t$ and therefore the validity of (4.3.27), since $t < \varepsilon$. If $t \ge \varepsilon$, by (4.3.28) and the assumption on the support of φ we infer $u(x,t) = 0$, and then (4.3.27) follows.
Next, we want to show that (4.3.27) implies

$$\int_{\mathbb{R}^N} \Gamma(x,t,\xi,\tau)d\xi \le 1, \tag{4.3.29}$$

for every $x \in \mathbb{R}^N$ and $t > \tau$. Indeed, for every $\varepsilon > 0$ let $(\varphi_n)_{n\in\mathbb{N}} \in C_c^\infty\Big(\mathbb{R}^N \times (\tau, \tau+\varepsilon)\Big)$ be a sequence of function with $0 \le \varphi_n \le 1$, pointwise converging to the characteristic function of $\mathbb{R}^N \times (\tau, \tau+\varepsilon)$. By (4.3.27), as n goes to infinity we get

$$\frac{1}{\varepsilon}\int_\tau^{\tau+\varepsilon}\int_{\mathbb{R}^N}\Gamma(x,t,\xi,\tau)d\xi \le 1,$$

implying (4.3.29) by Lebesgue differentiation theorem.
Finally, we conclude the proof of *(6.)* by showing that (4.3.29) is indeed an equality. For every $T > 0$ we set

$$w(x,t) := t - \int_0^t\int_{\mathbb{R}^N}\Gamma(x,t,\xi,s)d\xi ds, \qquad (x,t) \in \mathbb{R}^N \times (0,T).$$

By the same arguments previously employed, we see that w is a nonnegative function satisfying $\mathscr{H}_{\mathbb{G}} w = 0$ in $\mathbb{R}^N \times (0,T)$, and $w(x,0) = 0$ for every $x \in \mathbb{R}^N$. From the comparison principle we deduce that $w = 0$, that implies

$$\int_0^t \Big(1 - \int_{\mathbb{R}^N}\Gamma(x,t,\xi,s)d\xi ds\Big) = 0.$$

Finally, this relation, combined with (4.3.29), yields the desired conclusion.

(7.)
Let us fix $\xi \in \mathbb{R}^N$ and $s,\tau \in \mathbb{R}$ with $s > \tau$. For every $x \in \mathbb{R}^N$ and $t > s$ we set

$$w(x,t) := \int_{\mathbb{R}^N}\Gamma(x,t,y,s)\Gamma(y,s,\xi,\tau)dy, \qquad v(x,t) := \Gamma(x,t,\xi,\tau).$$

The reproduction property follows once we prove that $w(x,t) = v(x,t)$ for every $x \in \mathbb{R}^N$ and $t > s$. Since v is a fundamental solution, for every $T > s$ it solves $\mathscr{H}_{\mathbb{G}} v = 0$ in $\mathbb{R}^N \times (s,T)$, moreover $v(x,t) \to \Gamma(x,s,\xi,\tau)$ as $t \to s$. Let us now turn our attention to w. Since $\Gamma(\cdot,\xi,\tau) \in C^2_{\mathbb{G}}\big(\mathbb{R}^{N+1} \setminus \{\xi,\tau\}\big)$, it is bounded in $\mathbb{R}^N \times (s,T)$ (we recall that $\tau < s$). Hence, by *(4.)* we infer that for every $T > s$ the function w solves $\mathscr{H}_{\mathbb{G}} w = 0$ in $\mathbb{R}^N \times (s,T)$, and that $w(x,t) \to \Gamma(x,s,\xi,\tau)$ as $t \to s$. The thesis follows then by comparison principle, once we prove that

$$\sup_{t\in(s,T),\,|x|<R}|v(x,t) - w(x,t)| \underset{R\to\infty}{\longrightarrow} 0. \tag{4.3.30}$$

From the Gaussian estimates (4.3.2) we easily infer that

$$|v(x,t) - w(x,t)| \\ \le C_T^2 \left|\frac{\exp\Big(-c_u\frac{d_X(x,\xi)^2}{(t-\tau)}\Big)}{\Big(t-\tau\Big)^{\frac{Q}{2}}} + \int_{\mathbb{R}^N}\frac{\exp\Big(-c_u\frac{d_X(x,y)^2}{(t-s)}\Big)}{\Big(t-s\Big)^{\frac{Q}{2}}}\frac{\exp\Big(-c_u\frac{d_X(y,\xi)^2}{(s-\tau)}\Big)}{\Big(s-\tau\Big)^{\frac{Q}{2}}}dy\right|. \tag{4.3.31}$$

Let Γ_0 denote the fundamental solution of $\Delta_{\mathbb{G}} - \partial_t$. From the Gaussian estimates (4.3.25) and the reproduction property of Γ_0, we infer that there exist four constants $c_1, c_2, c_3, c_4 > 0$ such that

$$\begin{aligned}\int_{\mathbb{R}^N} &\frac{\exp\Big(-c_u \frac{d_X(x,y)^2}{(t-s)}\Big)}{\Big(t-s\Big)^{\frac{Q}{2}}} \frac{\exp\Big(-c_u \frac{d_X(y,\xi)^2}{(s-\tau)}\Big)}{(s-\tau)^{\frac{Q}{2}}} dy \\ &\le c_2 \int_{\mathbb{R}^N} \Gamma_0(x, c_1 t, y, c_1 s)\Gamma_0(y, c_1 s, \xi, c_1 \tau) dy \\ &\le \frac{c_4}{(t-\tau)^{\frac{Q}{2}}} \exp\left(-\frac{d_X(x,\xi)^2}{c_3(t-\tau)}\right),\end{aligned} \tag{4.3.32}$$

for every $(x,t), (\xi,\tau) \in \mathbb{R}^{N+1}$. Thus, we can bound (4.3.31) by $C \exp\Big(- c\, d_X(x,\xi)^2\Big)$, for some $c, C > 0$ that depend on t and τ. Finally, (4.3.30) follows from the continuity of the Euclidean metric w.r.t. the CC-metric, concluding the proof.

(8.)

We begin by proving a local Gaussian lower bound, which follows from the construction (4.3.18) provided by the parametrix method. We claim that there exist $C_1, C_2 > 0$ and $0 < \widetilde{T} < 1$ (depending only on structural assumptions) such that

$$\Gamma(x,t,\xi,\tau) \ge \frac{1}{2c_\lambda \Big(t-\tau\Big)^{\frac{Q}{2}}} \exp\left(-c_\lambda \frac{d_X(x,\xi)^2}{(t-\tau)}\right), \tag{4.3.33}$$

for every $(x,t), (\xi,\tau) \in \mathbb{R}^{N+1}$ such that $0 < t-\tau \le \widetilde{T}$ and $\frac{d_X(x,\xi)^2}{t-\tau} \le C_1 \left|\ln\Big(C_2(t-\tau)\Big)\right|$ (where c_λ is the constant appearing in (4.3.25)).

In order to prove claim (4.3.33), we first estimate the function J introduced in (4.3.19). By (4.3.25), (4.3.26), and (4.3.32) we infer that for every $z = (x,t), \zeta = (\xi,\tau) \in \mathbb{R}^{N+1}$ such that $0 < t-\tau \le T$, the following relation holds

$$\begin{aligned}|J(z,\zeta)| &\le \int_\tau^t \int_{\mathbb{R}^N} |Z(x,t,y,s)||G(y,s,\xi,\tau)| dy ds \\ &\le \bar{C}_T \int_\tau^t (s-\tau)^{\frac{\alpha}{2}-1} \int_{\mathbb{R}^N} \frac{\exp\Big(-\frac{d_X(x,y)^2}{c_\lambda(t-s)}\Big)}{\Big(t-s\Big)^{\frac{Q}{2}}} \frac{\exp\Big(-\frac{d_X(y,\xi)^2}{c(s-\tau)}\Big)}{(s-\tau)^{\frac{Q}{2}}} dy ds \\ &\le \frac{C_T}{(t-\tau)^{\frac{Q}{2}}} \exp\left(-\frac{d_X(x,\xi)^2}{\widetilde{c}(t-\tau)}\right) \int_\tau^t (s-\tau)^{\frac{\alpha}{2}-1} ds \\ &= \frac{2C_T}{\alpha} \frac{(t-\tau)^{\frac{\alpha}{2}}}{(t-\tau)^{\frac{Q}{2}}} \exp\left(-\frac{d_X(x,\xi)^2}{\widetilde{c}(t-\tau)}\right),\end{aligned} \tag{4.3.34}$$

for some $\bar{C}_T, C_T > 0$ and $\widetilde{c} \ge c_\lambda$, that depend only on structural assumptions and T. Hence, the local lower bound (4.3.33) follows from (4.3.18), (4.3.25) and (4.3.34), with C_1 only depending on the constant $\widetilde{c}$ and C_2 depending on C_T and c_λ.

The global lower bound (4.3.12) follows from the positivity of Γ given by *(5.)*, combined with the reproduction property (4.3.11). Indeed, let C_1, C_2 and $\widetilde{T}$ be the constants given by the local lower bound, and set

$$A := C_1 \left| \ln \left(C_2 \widetilde{T} \right) \right|. \tag{4.3.35}$$

We suppose without loss of generality that $\widetilde{T}$ is such that

$$\sqrt{A} > \left(\frac{c_\lambda}{2\omega_Q} \right)^{\frac{1}{Q}} \left(k_1 + k_1^2 \right),$$

where ω_Q is the measure of the d_X-unit ball, Q is the homogeneous dimension of $\mathbb{G}$, k_1 is defined in (4.2.6), and c_λ is the constant appearing in (4.3.33). Let us fix $x, \xi \in \mathbb{R}^N$, $0 < t - \tau < T$, and let k be the smallest integer greater then

$$\max \left(\frac{T}{\widetilde{T}}, c_1 \frac{d_X(x,\xi)^2}{T(t-\tau)} \right), \tag{4.3.36}$$

with $c_1 > 0$ to be fixed later. By the segment property and the equivalence of homogeneous distance (see for instance Corollary 5.15.6 and Proposition 5.1.4 in [27]), there exist $x = x_0, x_1, \ldots, x_k, x_{k+1} = \xi$ and $k_3 > 0$ (depending only on the group $\mathbb{G}$) such that

$$d_X(x_j, x_{j+1}) \le k_3 \frac{d_X(x,\xi)}{k+1}, \qquad \text{for every } j = 0, \ldots, k. \tag{4.3.37}$$

We fix

$$r := \left(\frac{2c_\lambda}{\omega_Q} \right)^{\frac{1}{Q}} \sqrt{\frac{(t-\tau)}{k+1}}, \tag{4.3.38}$$

and pick $t = s_0, s_1, \ldots, s_k, s_{k+1} = \tau$ such that

$$s_j - s_{j+1} = \frac{t-\tau}{k+1}, \qquad \text{for every } j = 0, \ldots, k. \tag{4.3.39}$$

Iterating the reproduction property (4.3.11) k times, by the positivity of Γ we get

$$\begin{aligned}
\Gamma(x,t,\xi,\tau) &= \int_{\mathbb{R}^N} \Gamma(x,t,y_1,s_1)\Gamma(y_1,s_1,\xi,\tau)dy_1 \\
&= \int_{(\mathbb{R}^N)^k} \prod_{j=0}^{k} \Gamma(y_j,s_j,y_{j+1}s_{j+1})dy_1 \ldots dy_k \\
&\ge \int_{\prod_{j=1}^k B_r(x_j)} \prod_{j=0}^{k} \Gamma(y_j,s_j,y_{j+1},s_{j+1})dy_1 \ldots dy_k =: I,
\end{aligned} \tag{4.3.40}$$

where we have set in the last two integrals $y_0 = x$ and $y_{j+1} = \xi$, (we recall that $B_r(x_j)$

denote the d_X-ball of radius r and center x_j). We set

$$c_1 := \left(\frac{k_1^2 k_3 \sqrt{T}}{\sqrt{A} - \left(\frac{2c_\lambda}{\omega_Q}\right)^{\frac{1}{Q}} \left(k_1 + k_1^2\right)} \right)^2,$$

and we claim that

$$\frac{d_X(y_j, y_{j+1})^2}{s_j - s_{j+1}} \leq C_1 \left| \ln \left(C_2(s_j - s_{j+1}) \right) \right|,$$

for every $y_j \in B_r(x_j)$ and for all $j = 1, \ldots, k$. This, in particular, allows us to employ the lower bound (4.3.12). By (4.2.6),(4.3.38), (4.3.37), (4.3.36), and (4.3.39) we get

$$\begin{aligned} d_X(y_j, y_{j+1}) &\leq r(k_1 + k_1^2) + k_1^2 k_3 \frac{d_X(x, \xi)}{k+1} \\ &\leq r(k_1 + k_1^2) + \frac{k_1^2 k_3}{k+1} \sqrt{\frac{T\, k}{c_1}(t - \tau)} \leq \sqrt{A}\sqrt{s_j - s_{j+1}}. \end{aligned}$$

The claim is therefore a consequence of (4.3.35). Hence, we are in position to apply (4.3.33), from which it follows that

$$\begin{aligned} I &\geq \left(\frac{1}{2c_\lambda}\right)^k \int_{\prod_{j=1}^k B_r(x_j)} \prod_{j=0}^k \frac{\exp\left(-c_\lambda \frac{d_X(y_j, y_{j+1})^2}{(s_j - s_{j+1})}\right)}{\left(s_j - s_{j+1}\right)^{\frac{Q}{2}}} dy_1 \ldots dy_k \\ &\geq \left(\frac{1}{2c_\lambda}\right)^k \left(\prod_{j=0}^k \frac{\exp\left(-c_\lambda A\right)}{\left(s_j - s_{j+1}\right)^{\frac{Q}{2}}} \right) \left(\prod_{j=1}^k \omega_Q \left(\frac{2c_\lambda}{\omega_Q}\right) (s_j - s_{j+1})^{\frac{Q}{2}} \right) \\ &\geq \frac{\exp\left(-k c_\lambda A\right)}{(t - \tau)^{\frac{Q}{2}}}. \end{aligned} \tag{4.3.41}$$

If $c_1 d_X(x, \xi)^2 \geq (t - \tau)T$, then $k = c_1 d_X(x, \xi)^2/(t - \tau) + \eta_1$ for some $\eta_1 \in [0, 1)$. This, combined with (4.3.40) and (4.3.41), implies

$$\Gamma(x, t, \xi, \tau) \geq \frac{\exp\left(-c_\lambda A\right)}{(t - \tau)^{\frac{Q}{2}}} \exp\left(-c_\lambda A c_1 \frac{d_X(x, \xi)^2}{(t - \tau)}\right).$$

Otherwise, if $c_1 d_X(x, \xi)^2 < (t - \tau)T$, then $k = \frac{T}{\tilde{T}} + \eta_2$ for some $\eta_2 \in [0, 1)$. Hence, again by (4.3.40) and (4.3.41), we deduce that

$$\begin{aligned} \Gamma(x, t, \xi, \tau) &\geq \exp\left(-c_\lambda A\right) \frac{\exp\left(-c_\lambda A \frac{T}{\tilde{T}}\right)}{(t - \tau)^{\frac{Q}{2}}} \\ &\geq \frac{\exp\left(-c_\lambda A \left(\frac{T}{\tilde{T}} + 1\right)\right)}{(t - \tau)^{\frac{Q}{2}}} \exp\left(-c_\lambda A c_1 \frac{d_X(x, \xi)^2}{(t - \tau)}\right), \end{aligned}$$

and this concludes the proof.

(10.) The proof of the existence and the properties of Γ^* is analogous to that of Γ and therefore we are left with proving (4.3.16).
We first observe that by *(1.)-(4.)* the function

$$u(x,t) := \int_{\mathbb{R}^N} \Gamma^*(x,t,\xi,T_2)g(\xi)d\xi - \int_t^{T_2}\int_{\mathbb{R}^N} \Gamma^*(x,t,\xi,\tau)f(\xi,\tau)d\xi d\tau,$$

is a classical solution to

$$\begin{cases} \mathscr{H}^*_{\mathbb{G}} u = f, & \text{in } \mathbb{R}^N \times (T, T_2), \\ u(x,T_2) = g(x), & \text{in } \mathbb{R}^N, \end{cases}$$

as it solves the equation, and the final condition is taken in the following limit sense

$$\lim_{t\to T_2}\left(\int_{\mathbb{R}^N} \Gamma^*(x,t,\xi,T_2)g(\xi)d\xi - \int_t^{T_2}\int_{\mathbb{R}^N} \Gamma^*(x,t,\xi,\tau)f(\xi,\tau)d\xi d\tau\right) = g(x). \quad (4.3.42)$$

We now notice that the *Green's identity* holds

$$v\mathscr{H}_{\mathbb{G}}u - u\mathscr{H}^*_{\mathbb{G}}v = \mathrm{div}_{\mathbb{G}}\left[\sum_{j=1}^{m_1}\Big(va_{ij}X_j u - ua_{ij}X_j v - uvX_j a_{ij}\Big) + b_i(uv)\right] - \partial_t(uv), \quad (4.3.43)$$

for every $u, v \in C^2_{\mathbb{G}}$. Let $d_{\mathbb{G}}$ denote a homogeneous distance on $\mathbb{G}$ smooth out of the origin, and let $B_r(x)$ be the corresponding $d_{\mathbb{G}}$-ball of radius $r > 0$ centered at $x \in \mathbb{R}^N$. We fix $R > 0$, $(x,t), (\xi,\tau) \in \mathbb{R}^{N+1}$, $\varepsilon < \tau - t$, and we integrate on $\Omega_T := B_R(0) \times (\tau-\varepsilon, t+\varepsilon)$ the Green's identity with $u(\cdot) = \Gamma(\cdot, x, t)$ and $v(\cdot) = \Gamma^*(\cdot,\xi,\tau)$. Since $\mathscr{H}_{\mathbb{G}}u = \mathscr{H}^*_{\mathbb{G}}v = 0$ in Ω_T, by applying the divergence theorem on $\mathbb{G}$ (see [91], pag.13) we get

$$\begin{aligned}
&\int_{B_R(0)} \Gamma(y,t+\varepsilon,x,t)\Gamma^*(y,t+\varepsilon,\xi,\tau)dy \\
&\qquad - \int_{B_R(0)} \Gamma(y,\tau-\varepsilon,x,t)\Gamma^*(y,\tau-\varepsilon,\xi,\tau)dy \\
&\qquad\quad = \int_{\tau-\varepsilon}^{t+\varepsilon}\int_{\partial B_R(0)} \sum_{j=1}^{m_1}\Big\langle\Big(\Gamma^* a_{ij}X_j\Gamma - \Gamma a_{ij}X_j\Gamma^* - \Gamma\Gamma^* X_j a_{ij}\Big) + b_i\Gamma\Gamma^*, \widetilde{\nu}\Big\rangle d\mathcal{H}^{n-1},
\end{aligned}$$

where $\widetilde{\nu}_i = \sum_{j=1}^N \nu_j\varphi^i_j$(here we recall that ν denotes the Euclidean normal outward vector of $\partial B_R(0)$). Since the components φ^i_j of the vector fields X_i are polynomial, by (4.3.2), and (4.3.3) and the equivalence of $d_{\mathbb{G}}$ and d_X, the integral in the right-hand side vanishes as $R \to +\infty$. Hence, we conclude that

$$\int_{\mathbb{R}^N} \Gamma(y,t+\varepsilon,x,t)\Gamma^*(y,t+\varepsilon,\xi,\tau)dy = \int_{\mathbb{R}^N} \Gamma(y,\tau-\varepsilon,x,t)\Gamma^*(y,\tau-\varepsilon,\xi,\tau)dy. \quad (4.3.44)$$

The proof follows by letting $\varepsilon \to 0$. Indeed, if we set $g_\varepsilon(y) := \Gamma^*(y,t+\varepsilon,\xi,\tau)$, then we rewrite the left-hand side of the previous equality as follows

$$\begin{aligned}
&\int_{\mathbb{R}^N} \Gamma(y, t+\varepsilon, x, t) g_\varepsilon(y) dy \\
&\quad = \int_{\mathbb{R}^N} \Gamma(y, t+\varepsilon, x, t)\Gamma^*(y, t, \xi, \tau) dy + \int_{\mathbb{R}^N} \Gamma(y, t+\varepsilon, x, t)\Big(g_\varepsilon(y) - \Gamma^*(y, t, \xi, \tau)\Big) dy \\
&\quad =: I_1 + I_2.
\end{aligned} \tag{4.3.45}$$

By the regularity of Γ outside the pole and (4.3.2), we infer that I_2 vanishes as $\varepsilon \to 0$. Moreover, by (4.3.9) we deduce that I_1 converges to $\Gamma^*(x, t, \xi, \tau)$ as $\varepsilon \to 0$.
The same arguments apply to the right-hand side of (4.3.44) (by using (4.3.42) instead of (4.3.9)). Hence, the proof is complete.

(9.)
We begin with estimate (4.3.14). Let k_1 as in (4.2.6): for every $H_1, H_2 > 0$ such that

$$\sqrt{\frac{H_1 + H_2}{H_1}} > k_1, \tag{4.3.46}$$

we claim that there exists $M > 0$ such that

$$H_1\, d_X(x, y)^2 - (H_1 + H_2)\, d_X(y, 0)^2 \leq M\, d_X(x, 0)^2, \qquad \text{for every } x, y \in \mathbb{R}^N. \tag{4.3.47}$$

The claim is clearly true for $y = 0$, taking $M \geq H_1$. Then, if $y \neq 0$, we set $r := d_X(y, 0)$ and

$$\xi := \delta_{\frac{1}{r}}(x), \qquad w := \delta_{\frac{1}{r}}(y).$$

Since d_X is δ_r-homogeneous of degree one, (4.3.47) becomes

$$H_1\, d_X(\xi, w)^2 - (H_1 + H_2) \leq M\, d_X(\xi, 0)^2. \tag{4.3.48}$$

If $d_X(\xi, w)^2 \leq \frac{(H_1+H_2)}{H_1}$, (4.3.48) holds for every $M > 0$. Otherwise, by (4.2.6) we get

$$d_X(\xi, 0) > \frac{1}{k_1}\sqrt{\frac{H_1 + H_2}{H_1}} - 1 := \vartheta.$$

We remark that ϑ is strictly positive, thanks to (4.3.46). Then, by (4.2.6) inequality (4.3.47) follows if

$$d_X(\xi, 0)^2\Big(H_1 k_1^2 - M\Big) + 2H_1 k_1^2 d_X(\xi, 0) + \Big(H_1 k_1^2 - H_1 - H_2\Big) < 0.$$

This inequality holds true if we pick $M > 2H_1 k_1^2/\vartheta$.

Let us now take $H_1 = h_1$ and $H_2 = \frac{c_u}{4T} - \frac{H_1}{2}$ (where c_u is given by the Gaussian upper bound (4.3.2)), and recall that h_1 satisfies (4.3.13). By (4.3.2) and (4.3.13) we get

$$\begin{aligned} |\Gamma(x,t,\xi,\tau)| &\le \frac{c_T(t-\tau)^{\frac{\alpha}{2}}}{(t-\tau)^{\frac{Q}{2}}} \exp\left(-c_u \frac{d_X(x,\xi)^2}{(t-\tau)}\right) \\ &\le c_T \frac{\exp\left(-c\frac{d_X(x,\xi)^2}{(t-\tau)}\right)}{(t-\tau)^{\frac{Q}{2}}} \exp\Big(-(H_1+H_2)d_X(x,\xi)^2\Big), \end{aligned} \tag{4.3.49}$$

with $c = \Big(c_u - T(H_1+H_2)\Big)/2$. From the representation formula (4.3.7), together with (4.3.5), (4.3.6), and (4.3.49), we infer the following inequalities

$$\begin{aligned} |u(x,t)| &\le \int_{\mathbb{R}^N} |\Gamma(x,t,\xi,T_1)g(\xi)|d\xi + \int_{T_1}^t \int_{\mathbb{R}^N} |\Gamma((x,t,\xi,\tau)f(\xi,\tau)|d\xi d\tau \\ &\le h_2 c_T \int_{\mathbb{R}^N} \frac{\exp\left(-c\frac{d_X(x,\xi)^2}{(t-T_1)}\right)}{(t-T_1)^{\frac{Q}{2}}} \exp\Big(-(H_1+H_2)d_X(x,\xi)^2 + H_1 d_X(\xi,0)^2\Big)d\xi \\ &+ h_2 c_T \int_{T_1}^t \int_{\mathbb{R}^N} \frac{\exp\left(-c\frac{d_X(x,\xi)^2}{(t-\tau)}\right)}{(t-\tau)^{\frac{Q}{2}}} \exp\Big(-(H_1+H_2)d_X(x,\xi)^2 + H_1 d_X(\xi,0)^2\Big)d\xi d\tau \\ &:= I_1. \end{aligned}$$

By setting $\xi^{-1} \circ x = y$, from (4.3.47) we get

$$\begin{aligned} I_1 &= h_2 c_T \int_{\mathbb{R}^N} \frac{\exp\left(-c\frac{d_X(y,0)^2}{(t-T_1)}\right)}{(t-T_1)^{\frac{Q}{2}}} \exp\Big(-(H_1+H_2)d_X(y,0)^2 + H_1 d_X(x,y)^2\Big)dy \\ &+ h_2 c_T \int_{T_1}^t \int_{\mathbb{R}^N} \frac{\exp\left(-c\frac{d_X(y,0)^2}{(t-\tau)}\right)}{(t-\tau)^{\frac{Q}{2}}} \exp\Big(-(H_1+H_2)d_X(y,0)^2 + H_1 d_X(x,y)^2\Big)dy d\tau \\ &\le h_2 c_T \exp\Big(M d_X(x,0)^2\Big) \left(\int_{\mathbb{R}^N} \frac{\exp\left(-c\frac{d_X(y,0)^2}{(t-T_1)}\right)}{(t-T_1)^{\frac{Q}{2}}} dy + \int_{T_1}^t \int_{\mathbb{R}^N} \frac{\exp\left(-c\frac{d_X(y,0)^2}{(t-\tau)}\right)}{(t-\tau)^{\frac{Q}{2}}} dy d\tau \right) \\ &:= I_2. \end{aligned}$$

Finally, employing once again the fundamental solution Γ_0 of $\Delta_{\mathbb{G}} - \partial_t$, together with its normalization and the Gaussian upper bound (4.3.25), we conclude

$$\begin{aligned} I_2 \le h_2 C_T \exp\Big(M d_X(x,0)^2\Big) \Bigg(&\int_{\mathbb{R}^N} \Gamma_0(y, c_1 t, 0, c_1 T_1) dy \\ &+ \int_{T_1}^t \int_{\mathbb{R}^N} \Gamma_0(y, c_1 t, 0, c_1 \tau) dy d\tau \Bigg) \le C \exp\Big(M d_X(x,0)^2\Big), \end{aligned}$$

for some $C_T, c_1 > 0$ depending only on structural assumptions.

We conclude by addressing the second part of *(9.)*
Let us suppose that u, v are classical solutions to (4.3.8) satisfying (4.3.15). The thesis follows once we show that there exists $\bar{\sigma} \in \mathbb{R}$ such that, for every $T_1 < \sigma < \bar{\sigma}$, it holds $u(x,t) = v(x,t)$ in $\mathbb{R}^N \times (T_1, \sigma)$.
To this end we consider the function $w := u - v$, which, for every $T_1 < \sigma < T_2$, is a solution to the equation

$$\begin{cases} \mathscr{H}_{\mathbb{G}} w = 0, & \text{in } \mathbb{R}^N \times (T_1, \sigma), \\ w(x, T_1) = 0, & \text{in } \mathbb{R}^N, \end{cases}$$

and by (4.3.15) satisfies

$$\int_{T_1}^{\sigma} \int_{\mathbb{R}^N} |w(x,t)| \exp\left(-k d_X(x,0)^2\right) dx dt < +\infty, \tag{4.3.50}$$

for some $k > 0$. We fix $(x,t) \in \mathbb{R}^N \times (T_1, \sigma)$, we let again $d_{\mathbb{G}}$ denote a homogeneous distance smooth out of the origin, and we denote as B_R the $d_{\mathbb{G}}$-ball of radius $R \geq 1$ centered at x. Let $0 \leq h \leq 1$ be a smooth cut-off function such that $h(\xi) = 0$ if $\xi \notin B_{R+1}$, and $h(\xi) = 1$ if $\xi \in B_R$. If we pick $\varepsilon > 0$ and we integrate over $B_{R+1} \times (T_1, t - \varepsilon)$ the Green's identity (4.3.43) with $u := w$ and $v := h(\cdot)\Gamma(x,t,\cdot)$, we get

$$\begin{aligned} &\int_{T_1}^{t-\varepsilon} \int_{B_{R+1}} w(\xi,\tau) \mathscr{H}^*_{\mathbb{G}} \Big(h(\xi)\Gamma(x,t,\xi,\tau)\Big) d\xi d\tau \\ &\quad = \int_{T_1}^{t-\varepsilon} \int_{B_{R+1}} \partial_\tau \Big(h(\xi)\Gamma(x,t,\xi,\tau) w(\xi,\tau)\Big) d\xi d\tau \\ &\qquad - \int_{T_1}^{t-\varepsilon} \int_{B_{R+1}} \mathrm{div}_{\mathbb{G}} \left[\sum_{j=1}^{m} \Big(h\Gamma a_{ij} X_j w - w a_{ij} X_j (h\Gamma) - w h \Gamma X_j a_{ij} \Big) + b_i w h \Gamma \right] d\xi d\tau. \end{aligned}$$

Since $h = 0$ on ∂B_{R+1}, the divergence theorem on $\mathbb{G}$ implies that the last integral vanishes. Then $w(x,0) = 0$ yields

$$\begin{aligned} \int_{T_1}^{t-\varepsilon} \int_{B_{R+1}} w(\xi,\tau) \mathscr{H}^*_{\mathbb{G}} \Big(h(\xi)\Gamma(x,t,\xi,\tau)\Big) d\xi d\tau \\ = \int_{\mathbb{R}^N} h(\xi)\Gamma(x,t,\xi,t-\varepsilon) w(\xi, t-\varepsilon) d\xi. \end{aligned}$$

As shown in (4.3.45), the integral on the right-hand side converges to $h(x)w(x,t) = w(x,t)$ as $\varepsilon \to 0$. Since $h = 0$ in B_R and $\mathscr{H}^*_{\mathbb{G}}\Gamma(x,t,\cdot) = 0$ in $B_{R+1} \setminus B_R$ (by (4.3.16)), we get

$$w(x,t) = \int_{T_1}^{t} \int_{B_{R+1} \setminus B_R} w \left[2 \sum_{i,j=1}^{m_1} a_{ij} X_i h X_j \Gamma + \Gamma \sum_{i,j=1}^{m_1} a_{ij} X_i X_j h + \Gamma \sum_{i=1}^{m_1} X_i h \right] d\xi d\tau.$$

Given that h and its derivatives are bounded, and exploiting once again the fact that the vector fields have polynomial components, by (4.3.2) and (4.3.3) and the equivalence of $d_{\mathbb{G}}$ and d_X, we infer that there exits $\bar{\sigma} \in \mathbb{R}$ such that, if $T_1 < \sigma < \bar{\sigma}$, the following relation holds for some $C, \bar{C}_T > 0$

$$\begin{aligned}
|w(x,t)| &\le C_T \int_{T_1}^{t} \int_{B_{R+1} \setminus B_R} |w(\xi,\tau)| \frac{\exp\left(-C\frac{d_X(x,\xi)^2}{(t-\tau)}\right)}{(t-\tau)^{\frac{Q+1}{2}}} d\xi d\tau \\
&\le \bar{C}_T \int_{T_1}^{\sigma} \int_{B_{R+1} \setminus B_R} |w(\xi,\tau)| \exp\left(-kk_1^2 d_X(x,\xi)^2\right) d\xi d\tau \\
&\le \bar{C}_T \exp\left((1-kk_1) d_X(x,0)^2\right) \int_{T_1}^{\sigma} \int_{B_{R+1} \setminus B_R} |w(\xi,\tau)| \exp\left(-k d_X(\xi,0)^2\right) d\xi d\tau.
\end{aligned}$$

By (4.3.50), we infer that the right-hand side vanishes as $R \to +\infty$, implying $w(x,t) = 0$ for every $(x,t) \in \mathbb{R}^N \times (T_1, \sigma)$. Repeating this argument finitely many times yields the desired result. $\square$

4.4 Mean value formulas

In this section we prove a representation formula that will be exploited in Sections 4.5 and 4.6 to prove the maximum principle and a parabolic Harnack inequality. This formula is based on the solid mean value formula recently developed in [92], that we here improve employing Hadamard's descent method, in order to overcome the unboundedness of the kernel. The solid mean value formula from Theorem 2.2 in [92] holds true once the following conditions are satisfied

(M1) There exists a homogeneous Lie group $\widehat{\mathbb{G}} = \left(\mathbb{R}^{N+1}, \widehat{\circ}, \widehat{\delta}_r\right)$ such that $X_1, \dots, X_{m_1}$ and $X_0 = \partial_t$ are left-invariant on $\widehat{\mathbb{G}}$ and $\widehat{\delta}_r$-homogeneous of degree one.

(M2) There exists a fundamental solution Γ^* to the adjoint operator (4.3.1) such that:

- there exists $r_0 > 0$ such that for all $\zeta = (\xi, \tau) \in \mathbb{R}^{N+1}$ the set

  $$\Omega_r(\zeta) := \left\{ z = (x,t) \in \mathbb{R}^N \times (-\infty, \tau) : \Gamma^*(z,\zeta) > \frac{1}{r} \right\},$$

  is bounded for every $r \le r_0$;

- for all $\zeta = (\xi, \tau) \in \mathbb{R}^{N+1}$ and $r, \varepsilon > 0$, the set

  $$\mathcal{I}_{r,\varepsilon}(\zeta) := \left\{ x \in \mathbb{R}^N : \Gamma^*(x, \tau - \varepsilon, \zeta) > \frac{1}{r} \right\},$$

satisfies the *pointwise vanishing integral condition*

$$\lim_{\varepsilon\to 0}\mathscr{H}^N\Big(\mathcal{I}_{r,\varepsilon}(\zeta)\Big)=0,\qquad \lim_{\varepsilon\to 0}\int_{\mathbb{R}^N\setminus\mathcal{I}_{r,\varepsilon}(\zeta)}\Gamma^*(x,\tau-\varepsilon,\zeta)d\xi=0. \tag{4.4.1}$$

Before proving our improved mean value formula, we briefly discuss the previous conditions. We first focus our attention on **(M1)**. In regularity theory, $\widehat{\mathbb{G}}$ is not the usual group where one investigates the properties of $\mathscr{H}_{\mathbb{G}}$. This is because operator $\mathscr{H}_{\mathbb{G}}$ is not homogeneous of degree 2 w.r.t. $\widehat{\delta}_r$, since the vector field $X_0=\partial_t$ is $\widehat{\delta}_r$-homogeneous of degree one (namely, X_0 belong to the first layer of the Lie algebra of $\widehat{\mathbb{G}}$). The natural geometry when studying operator $\mathscr{H}_{\mathbb{G}}$ is determined by the parabolic group relevant to $\mathbb{G}$. More precisely, the natural framework for the regularity theory of operator $\mathscr{H}_{\mathbb{G}}$ is the group $\mathbb{G}_p=\Big(\mathbb{R}^{N+1},\bar{\circ},\bar{\delta}_r\Big)$, where

$$z\,\bar{\circ}\,\zeta:=(x\circ\xi,t+\tau)\,,\qquad \bar{\delta}_r(x,t)=\Big(\delta_r(x),r^2t\Big),\qquad \forall z,\zeta\in\mathbb{R}^{N+1},r>0,$$

with $\circ$ and $\{\delta_r\}_{r>0}$ given by $\mathbb{G}$. It is straightforward to verify that on $\mathbb{G}_p$ the operator $\mathscr{H}_{\mathbb{G}}$ is left-invariant and homogeneous of degree 2. However, the divergence theorem on $\mathbb{G}_p$ would not be enough to prove the mean value formula, since in this setting the surfaces $\{t=\text{constant}\}$ would have codimension 2, hence being negligible. For further comments on this topic, we refer the reader to the introduction of [92], or to the monograph [110] about Geometric Measure Theory on Carnot groups.

We now comment on **(M2)**. The pointwise vanishing integral condition (4.4.1), as emphasized in [44], is a fundamental requirement for the proof of mean value formulas, and is closely related to the local properties of Γ^*. In the stationary case, this property is straightforward to establish. In the evolutive setting, however, its proof is more delicate and relies on precise estimates, that in our case will be the Gaussian lower and upper bounds, together with the long time bound given by *(5.)* in Theorem 4.3.11.

We conclude with a final remark. In Theorem 4.3.11 we established the relation (4.3.16) between Γ and Γ^*, if **(H1)**, **(H2)**, **(H3)** and **(H4)** hold. As a result, we can express the mean value formula in terms of the fundamental solution Γ, instead of Γ^*. In particular, this means that

$$\Omega_r(\zeta)=\left\{z\in\mathbb{R}^N\times(-\infty,\tau):\Gamma(\zeta,z)>\frac{1}{r}\right\},\quad \mathcal{I}_{r,\varepsilon}(\zeta)=\left\{x\in\mathbb{R}^N:\Gamma(\zeta,x,\tau-\varepsilon)>\frac{1}{r}\right\}.$$

Finally, before presenting the representation formula, we introduce the relevant kernel. For every $z,\zeta\in\mathbb{R}^{N+1}$ we set

$$M_{\mathbb{G}}(\zeta,z):=\frac{\langle A(z)\nabla_{\mathbb{G}}\Gamma(\zeta,z),\nabla_{\mathbb{G}}\Gamma(\zeta,z)\rangle}{\Gamma(\zeta,z)^2}. \tag{4.4.2}$$

Proposition 4.4.12 (Mean value formula)**.** *Let $\mathscr{H}_{\mathbb{G}}$ be a differential operator as in* (4.1.2) *satisfying assumptions **(H1)**, **(H2)**, **(H3)** and **(H4)**. Let Ω be an open subset of $\mathbb{R}^{N+1}$, $f \in C(\Omega)$ and let u be a classical solution to $\mathscr{H}_{\mathbb{G}} u = f$ in Ω. Then, for every $\zeta \in \Omega$ and for every $r > 0$ such that $\Omega_r(\zeta) \subset \Omega$, we have*

$$
\begin{aligned}
u(\zeta) =& \frac{1}{r}\int_{\Omega_r(\zeta)} M_{\mathbb{G}}(\zeta,z)u(z)\,dz + \frac{1}{r}\int_0^r \left(\int_{\Omega_\varrho(\zeta)} f(z)\Big(\tfrac{1}{\varrho} - \Gamma(\zeta,z)\Big)dz\right)d\varrho \\
&+ \frac{1}{r}\int_0^r \frac{1}{\varrho}\left(\int_{\Omega_\varrho(\zeta)} \Big(\mathrm{div}_{\mathbb{G}} b(z) - c(z)\Big)u(z)\,dz\right)d\varrho.
\end{aligned}
$$

Proof. This representation formula corresponds to the solid mean value formula stated in Theorem 2.2 in [92]. We need to prove that we are in position to apply the aforementioned theorem, namely that **(M1)** and **(M2)** hold under assumptions **(H1)**, **(H2)**, **(H3)** and **(H4)**.

The existence of the Carnot group $\mathbb{G} = \left(\mathbb{R}^N, \circ, \delta_r\right)$ introduced in Section 4.2, directly implies **(M1)**. Indeed, we define on $\mathbb{R}^{N+1}$ the following composition and dilations law

$$
z \,\widehat{\circ}\, \zeta := (x \circ \xi, t + \tau)\,, \qquad \widehat{\delta}_r(x,t) = \Big(\delta_r(x), rt\Big), \qquad \forall z, \zeta \in \mathbb{R}^{N+1}, r > 0.
$$

Given that $\mathbb{G}$ is a Carnot group on $\mathbb{R}^N$, as an obvious consequence it follows that $\widehat{\mathbb{G}} = \left(\mathbb{R}^{N+1}, \widehat{\circ}, \widehat{\delta}_r\right)$ is a homogeneous Lie group on $\mathbb{R}^{N+1}$. Moreover, the vector fields $X_0 = \partial_t$ and $X_1, \dots, X_{m_1}$ are left-invariant w.r.t. $\widehat{\circ}$ and $\widehat{\delta}_r$-homogeneous of degree 1, namely

$$
X_i\Big(f\big(z \,\widehat{\circ}\, \zeta\big)\Big) = \Big(X_i f\Big)\Big(z \,\widehat{\circ}\, \zeta\Big), \quad X_i\left(f\left(\widehat{\delta}_r(z)\right)\right) = r\Big(X_i f\Big)\left(\widehat{\delta}_r(z)\right), \quad i = 0, \dots, m_1,
$$

for every $f \in C^\infty(\mathbb{R}^{N+1})$, thus implying **(M1)**.

Let us turn our attention to condition **(M2)**. By *(5.)* and *(8.)* in Theorem 4.3.11, we infer that there exists $\bar{r} > 0$ such that

$$
\Omega_r(\zeta) = \left\{ z \in \mathbb{R}^N \times (\tau - 1, \tau) : \Gamma(\zeta, z) > \frac{1}{r} \right\}, \tag{4.4.3}
$$

for every $r \le \bar{r}$. Let us now consider the Gaussian upper bound (4.3.2) with $T = 1$. If we set $r_0 \le \min\left(\bar{r}, c_1^{-1}\right)$, it follows

$$
\begin{aligned}
\Omega_r(\zeta) &\subset \left\{ z \in \mathbb{R}^N \times (\tau - 1, \tau) : d_X(\xi, x)^2 \le c_u(\tau - t)\ln\left(\frac{c_1\, r}{(\tau - t)^{\frac{Q}{2}}}\right)\right\}, \\
\mathcal{I}_{r,\varepsilon}(\zeta) &\subset \left\{ x \in \mathbb{R}^N : d_X(\xi, x)^2 \le c_u \varepsilon \ln\left(\frac{c_1\, r}{\varepsilon^{\frac{Q}{2}}}\right)\right\},
\end{aligned}
$$

provided $\varepsilon < 1$. The local boundedness of $\Omega_r(\zeta)$ and the first point of (4.4.1) follows. We conclude by addressing the second point of (4.4.1). With this aim, we employ the Gaussian lower bound (4.3.12), from which it follows that

$$\left\{ x \in \mathbb{R}^N : d_X(\xi, x)^2 \leq c_l \varepsilon \ln \left(\frac{\widetilde{c}_\varepsilon r}{\varepsilon^{\frac{Q}{2}}} \right) \right\} \subset \mathcal{I}_{r,\varepsilon}(\zeta),$$

for every $\varepsilon > 0$. Hence, since Γ is positive, by the Gaussian upper bound (4.3.2) we infer that for every $\varepsilon < 1$ there exists $c_1, c_u > 0$ such that

$$\begin{aligned}
0 &\leq \int_{\mathbb{R}^N \setminus \mathcal{I}_{r,\varepsilon}(\zeta)} \Gamma(\zeta, x, \tau - \varepsilon) dx \leq \int_{\left\{ d_X(\xi,x)^2 \geq c_l \varepsilon \ln \left(\frac{\widetilde{c}_\varepsilon r}{\varepsilon^{\frac{Q}{2}}} \right) \right\}} \Gamma(\zeta, x, \tau - \varepsilon) dx \\
&\leq \frac{c_1}{\varepsilon^{\frac{Q}{2}}} \int_{\left\{ d_X(\xi,x)^2 \geq c_l \varepsilon \ln \left(\frac{\widetilde{c}_\varepsilon r}{\varepsilon^{\frac{Q}{2}}} \right) \right\}} \exp \left(-c_u \frac{d_X(\xi, x)^2}{\varepsilon} \right) dx \\
&\left(\text{by setting } y = \delta_{\frac{1}{\sqrt{\varepsilon}}} \left(x^{-1} \circ \xi \right) \right) \\
&= c_1 \int_{\left\{ d_X(y,0)^2 \geq c_l \ln \left(\frac{\widetilde{c}_\varepsilon r}{\varepsilon^{\frac{Q}{2}}} \right) \right\}} \exp \Big(- c_u d_X(y, 0)^2 \Big) dy,
\end{aligned}$$

and the proof is complete, since the last integral vanishes as $\varepsilon \to 0$. □

Remark 4.4.13. *To be more precise, the kernel appearing in Theorem 2.2 in [92] is actually of the following form*

$$M_{\mathbb{G}}(\zeta, z) := \frac{\langle \bar{A}(z) \nabla_{\widehat{\mathbb{G}}} \Gamma(\zeta, z), \nabla_{\widehat{\mathbb{G}}} \Gamma(\zeta, z) \rangle}{\Gamma(\zeta, z)^2},$$

where $\bar{A}$ denotes the matrix with entries a_{ij}, and $a_{(m_1+1)j} = a_{j(m_1+1)} = 0$, while $\nabla_{\widehat{\mathbb{G}}}$ denotes the horizontal gradient on $\widehat{\mathbb{G}}$, namely

$$\nabla_{\mathbb{G}} f := \sum_{i=1}^{m_1} \Big(X_i f \Big) X_i + (\partial_t f)(\partial_t).$$

However, the previous kernel is formally identical to the one introduced in (4.4.2).

As mentioned before, a questionable feature displayed in the previous proposition is the unboundedness of the kernels. This can be easily checked in the case of the classical heat operator, where $M_{\mathbb{G}}(\zeta, z) = \frac{|\xi - x|^2}{4(\tau - t)^2}$, the so-called Pini-Watson kernel. Moreover, another limitation is that the kernel cannot be bounded from below by a strictly positive constant. These two issues prevent us from taking advantage of the classical elementary argument developed for harmonic functions to establish both the maximum propagation and the Harnack inequality. Hence, starting from Proposition 4.4.12, we aim to prove a representation formula with a kernel that is not only bounded but, also possesses a degree of regularity arbitrarily high. This boundedness will be crucial in particular to prove the parabolic Harnack inequality given in Section 4.6. Our technique to prove the mean value formula with a more regular kernel is based on Hadamard's descent method, already used by Kuptsov in [73], and later developed in [44, 59, 60, 85, 105]. In order to outline the procedure we follow, we first introduce some further notation. For every $m \in \mathbb{N}$, we introduce the operator

$$\widetilde{\mathscr{H}_{\mathbb{G}}}^{(m)} := \Delta_y^{(m)} + \mathscr{H}_{\mathbb{G}}, \tag{4.4.4}$$

where $\Delta_y^{(m)}$ is the classical Laplace operator in $\mathbb{R}^m$ in the variables $y = (y_1, \dots, y_m)$. Moreover, we define the function

$$\begin{aligned} \widetilde{\Gamma}^{(m)}(\eta, \xi, \tau, y, x, t) &= H^{(m)}(\eta, \tau, y, t)\, \Gamma(\xi, \tau, x, t) \\ &= \frac{1}{\big(4\pi(\tau - t)\big)^{m/2}} \exp\left(-\frac{|\eta - y|^2}{4(\tau - t)}\right) \Gamma(\xi, \tau, x, t), \end{aligned} \tag{4.4.5}$$

where Γ and $H^{(m)}$ denote the fundamental solutions of $\mathscr{H}_{\mathbb{G}}$ and of the heat equation in $\mathbb{R}^m$, respectively. By standard arguments it follows that $\widetilde{\Gamma}^{(m)}$ is a fundamental solution to $\widetilde{\mathscr{H}_{\mathbb{G}}}^{(m)}$ and that, if $u = u(x,t)$ is a solution to $\mathscr{H}_{\mathbb{G}} u = f$, then the function

$$\widetilde{u}^{(m)}(y, x, t) := u(x, t), \quad y \in \mathbb{R}^m, \tag{4.4.6}$$

is a solution to $\widetilde{\mathscr{H}_{\mathbb{G}}}^{(m)} \widetilde{u}^{(m)} = \widetilde{f}^{(m)}$, where

$$\widetilde{f}^{(m)}(y, x, t) = f(x, t), \quad y \in \mathbb{R}^m.$$

By integrating the mean value formula of Proposition 4.4.12 with respect to the variable y, we then obtain new kernels which are bounded for any $m > 2$. In order to state and prove our modified mean value formula, we first introduce some further notation:

$$\begin{aligned} \Gamma^{(m)}(\zeta, z) &:= \frac{\Gamma(\zeta, z)}{\big(4\pi(\tau - t)\big)^{m/2}}, \\ \Omega_r^{(m)}(\zeta) &:= \left\{ z \in \mathbb{R}^{N+1} : \Gamma^{(m)}(\zeta, z) > \frac{1}{r} \right\}, \\ N_r(\zeta, z) &:= 2\sqrt{\tau - t}\sqrt{\ln\big(r\, \Gamma^{(m)}(\zeta, z)\big)}, \\ M_r^{(m)}(\zeta, z) &:= \omega_m\, N_r^m(\zeta, z) \left(M_{\mathbb{G}}(\zeta, z) + \frac{m}{m+2} \frac{N_r^2(\zeta, z)}{4(\tau - t)^2} \right), \\ W_r^{(m)}(\zeta, z) &:= m\, \omega_m \left(\frac{N_r^m(\zeta, z)}{rm} - \Gamma^{(m)}(\zeta, z)\, \frac{4(\tau - t)^{m/2}}{2}\, \gamma\left(\frac{m}{2}; \frac{N_r^2(\zeta, z)}{4(\tau - t)} \right) \right), \end{aligned} \tag{4.4.7}$$

where $M_{\mathbb{G}}(\zeta, z)$ is the kernel introduced in (4.4.2), ω_m denotes the Lebesgue measure of the unit ball in $\mathbb{R}^m$ w.r.t. the Euclidean metric, and γ is the lower incomplete gamma function, namely

$$\gamma(s; x) := \int_0^x t^{s-1} \mathrm{e}^{-t} dt. \tag{4.4.8}$$

We are now in a position to state and prove the following proposition.

Proposition 4.4.14 (Improved mean value formula)**.** *Let $\mathscr{H}_{\mathbb{G}}$ be a differential operator of the form* (4.1.2) *satisfying assumptions **(H1)**, **(H2)**, **(H3)** and **(H4)**. Let Ω be an open subset of $\mathbb{R}^{N+1}$, $f \in C(\Omega)$ and let u be a classical solution to $\mathscr{H}_{\mathbb{G}} u = f$ in Ω. Then, for every $\zeta \in \Omega$ and for every $r > 0$ such that $\Omega_r^{(m)}(\zeta) \subset \Omega$, we have*

$$\begin{aligned} u(\zeta) =& \frac{1}{r} \int_{\Omega_r^{(m)}(\zeta)} M_r^{(m)}(\zeta, z) u(z)\, dz + \frac{1}{r} \int_0^r \left(\int_{\Omega_\varrho^{(m)}(\zeta)} f(z)\, W_\varrho^{(m)}(\zeta, z)\, dz \right) d\varrho \\ &+ \frac{1}{r} \int_0^r \frac{\omega_m}{\varrho} \left(\int_{\Omega_\varrho^{(m)}(\zeta)} N_\varrho^m(\zeta, z) \left(\mathrm{div}_{\mathbb{G}} b(z) - c(z) \right) u(z)\, dz \right) d\varrho. \end{aligned} \tag{4.4.9}$$

Proof. As a first step, we observe that, for every $m \in \mathbb{N}$, the operator $\widetilde{\mathscr{H}_{\mathbb{G}}}^{(m)}$ in (4.4.4) belongs to the class of operators (4.1.2) with the matrix A replaced by the $(m+m_1) \times (m+m_1)$ real matrix

$$\widetilde{A}(y, z) = \begin{pmatrix} \mathbb{I}_m & \mathbb{O} \\ \mathbb{O} & A(z) \end{pmatrix}, \qquad y \in \mathbb{R}^m, z \in \mathbb{R}^{N+1},$$

where A denotes the matrix related to operator $\mathscr{H}_{\mathbb{G}}$ in (4.1.2). Hence, operator $\widetilde{\mathscr{H}_{\mathbb{G}}}^{(m)}$ clearly satisfies the assumptions of Proposition 4.4.12, and we infer

$$\begin{aligned} \widetilde{u}^{(m)}(\eta, \zeta) =& \frac{1}{r} \int_{\widetilde{\Omega}_r^{(m)}(\eta, \zeta)} \widetilde{M}_{\mathbb{G}}^{(m)}(\eta, \zeta, y, z) \widetilde{u}^{(m)}(y, z)\, dy\, dz \\ &+ \frac{1}{r} \int_0^r \left(\int_{\widetilde{\Omega}_\varrho^{(m)}(\eta, \zeta)} \widetilde{f}^{(m)}(y, z) \left(\frac{1}{\varrho} - \widetilde{\Gamma}^{(m)}(\eta, \zeta, y, z) \right) dy\, dz \right) d\varrho \\ &+ \frac{1}{r} \int_0^r \frac{1}{\varrho} \left(\int_{\widetilde{\Omega}_\varrho^{(m)}(\eta, \zeta)} \left(\mathrm{div}_{\mathbb{G}} b(z) - c(z) \right) \widetilde{u}^{(m)}(y, z)\, dy\, dz \right) d\varrho, \end{aligned} \tag{4.4.10}$$

where $\widetilde{\Omega}_r^{(m)}(\eta, \zeta)$ is the super-level set relevant to $\widetilde{\Gamma}^{(m)}$ in (4.4.5), *i.e.*

$$\widetilde{\Omega}_r^{(m)}(\eta, \zeta) := \left\{ (y, z) \in \mathbb{R}^{m+N} \times (-\infty, \tau) : \widetilde{\Gamma}^{(m)}(\eta, \zeta, y, z) > \frac{1}{r} \right\}. \tag{4.4.11}$$

We first point out that, in virtue of (4.2.8), the kernel $\widetilde{M}^{(m)}(\eta, \zeta, y, z)$ in (4.4.10) corresponding to operator $\widetilde{\mathscr{H}_{\mathbb{G}}}^{(m)}$ is

$$\widetilde{M}^{(m)}(\eta,\zeta,y,z) = \frac{\Gamma(z,\zeta)\sum_{j=1}^{m}\partial^2_{y_j}H^{(m)}(\eta,\tau,y,t) + H^{(m)}(\eta,\tau,y,t)^2\sum_{i,j=1}^{m_1}a_{ij}(z)X_i\Gamma(\zeta,z)X_j\Gamma(\zeta,z)}{\widetilde{\Gamma}^{(m)}(\eta,\zeta,y,z)^2}.$$

Then, owing to (4.2.8) and (4.4.5), the kernel $\widetilde{M}^{(m)}(\eta,\zeta,y,z)$ becomes

$$\widetilde{M}^{(m)}(\eta,\zeta,y,z) = M_{\mathbb{G}}(\zeta,z) + \frac{|\eta-y|^2}{4(\tau-t)^2}, \tag{4.4.12}$$

where $M_{\mathbb{G}}(\zeta,z)$ is the kernel defined in (4.4.2), and $\frac{|\eta-y|^2}{4(\tau-t)^2}$ is the classical Pini-Watson kernel in $\mathbb{R}^{m+1}$. We observe that, as a consequence of (4.4.5), we can rewrite the super-level set $\widetilde{\Omega}_r^{(m)}(\eta,\zeta)$ in (4.4.11) as follows

$$\widetilde{\Omega}_r^{(m)}(\eta,\zeta) = \left\{(y,z)\in\mathbb{R}^{m+N+1} : r\,\frac{\Gamma(\zeta,z)}{\Big(4\pi(\tau-t)\Big)^{m/2}} > \exp\left(\frac{|\eta-y|^2}{4(\tau-t)}\right)\right\}. \tag{4.4.13}$$

In virtue of (4.4.6), we now rewrite equation (4.4.10) as follows

$$\begin{aligned} u(\zeta) =& \frac{1}{r}\int_{\widetilde{\Omega}_r^{(m)}(\eta,\zeta)} \widetilde{M}^{(m)}(\eta,\zeta,y,z)u(z)\,dy\,dz \\ &+\frac{1}{r}\int_0^r\left(\int_{\widetilde{\Omega}_\varrho^{(m)}(\eta,\zeta)} f(z)\left(\frac{1}{\varrho}-\widetilde{\Gamma}^{(m)}(\eta,\zeta,y,z)\right)dy\,dz\right)d\varrho \\ &+\frac{1}{r}\int_0^r\frac{1}{\varrho}\left(\int_{\widetilde{\Omega}_\varrho^{(m)}(\eta,\zeta)}\Big(\mathrm{div}_{\mathbb{G}}b(z)-c(z)\Big)u(z)\,dy\,dz\right)d\varrho \\ =&: I_1+I_2+I_3. \end{aligned} \tag{4.4.14}$$

We first focus on the term I_1. Exploiting (4.4.12) and (4.4.13), we get rid of the variable y by integrating as follows

$$\begin{aligned} I_1 &= \frac{1}{r}\int_{\widetilde{\Omega}_r^{(m)}(\eta,\zeta)}\left(M_{\mathbb{G}}(\zeta,z)+\frac{|\eta-y|^2}{4(\tau-t)^2}\right)u(z)\,dy\,dz \\ &= \frac{1}{r}\int_{\Omega_r^{(m)}(\zeta)}\int_{\{|\eta-y|<N_r(\zeta,z)\}}\left(M_{\mathbb{G}}(\zeta,z)+\frac{|\eta-y|^2}{4(\tau-t)^2}\right)u(z)\,dy\,dz \\ &= \frac{1}{r}\int_{\Omega_r^{(m)}(\zeta)} M_r^{(m)}(\zeta,z)u(z)\,dz, \end{aligned} \tag{4.4.15}$$

where the quantities $\Omega_r^{(m)}(\zeta)$, $N_r(\zeta,z)$ and $M_r^{(m)}(\zeta,z)$ are defined in (4.4.7). In order to take care of I_2, we first recall the following equality

$$\int_0^R \mathrm{e}^{-\frac{r^2}{c}}r^{m-1}dr = \frac{c^{\frac{m}{2}}}{2}\int_0^{\frac{R^2}{c}}\mathrm{e}^{-r}r^{\frac{m}{2}-1}dr = \frac{c^{\frac{m}{2}}}{2}\gamma\left(\frac{m}{2};\frac{R^2}{c}\right),$$

where γ is the lower incomplete gamma function (we recall the definition given in (4.4.8)). As a consequence of the previous equality and (4.4.13), we obtain

$$\begin{aligned} I_2 &= \frac{1}{r}\int_0^r \int_{\Omega_\varrho^{(m)}(\eta,\zeta)} \left(\int_{\{|\eta-y|<N_\varrho(\zeta,z)\}} \left(\frac{1}{\varrho} - \Gamma(\zeta,z) \frac{\exp\left(-\frac{|\eta-y|^2}{\left(4(\tau-t)\right)^{\frac{m}{2}}}\right)}{\left(4\pi(\tau-t)\right)^{\frac{m}{2}}} \right) dy \right) f(z)dzd\varrho \\ &= \frac{1}{r}\int_0^r \left(\int_{\Omega_\varrho^{(m)}(\zeta)} W_\varrho^{(m)}(\zeta,z) f(z)\, dz \right) d\varrho, \end{aligned} \tag{4.4.16}$$

where $W_\varrho^{(m)}(\zeta,z)$ was introduced in (4.4.7). Finally, exploiting once again (4.4.13), we infer

$$\begin{aligned} I_3 &= \frac{1}{r}\int_0^r \frac{1}{\varrho} \left(\int_{\Omega_\varrho^{(m)}(\zeta)} \left(\mathrm{div}_{\mathbb{G}} b(z) - c(z)\right) u(z) \left(\int_{\{|\eta-y|<N_\varrho(\zeta,z)\}} dy \right) dz \right) d\varrho \\ &= \frac{1}{r}\int_0^r \frac{1}{\varrho} \left(\int_{\Omega_\varrho^{(m)}(\zeta)} \left(\mathrm{div}_{\mathbb{G}} b(z) - c(z)\right) u(z)\, N_\varrho^m(\zeta,z)\, \omega_m\, dz \right) d\varrho. \end{aligned} \tag{4.4.17}$$

The thesis follows combining (4.4.14), (4.4.15), (4.4.16) and (4.4.17). $\square$

4.5 Proof of Theorem 4.1.4

Let $\zeta \in \Omega$ such that $u(\zeta) = \max_\Omega u \geq 0$. Since $\mathscr{H}_{\mathbb{G}} 1 = c$, from (4.4.9) we get

$$\begin{aligned} 0 = &\underbrace{\boxed{\frac{1}{r}\int_{\Omega_r^{(m)}(\zeta)} M_r^{(m)}(\zeta,z)\left(u(z)-u(\zeta)\right)dz}}_{:=I_1} \\ &+ \underbrace{\boxed{\frac{1}{r}\int_0^r \frac{\omega_m}{\varrho} \left(\int_{\Omega_\varrho^{(m)}(\zeta)} N_\varrho^m(\zeta,z)\left(\mathrm{div}_{\mathbb{G}} b(z) - c(z)\right)\left(u(z)-u(\zeta)\right)dz \right) d\varrho}}_{:=I_2} \\ &+ \underbrace{\boxed{\frac{1}{r}\int_0^r \left(\int_{\Omega_\varrho^{(m)}(\zeta)} \left(f(z) - \frac{\omega_m}{\varrho} u(\zeta) c(z)\right) W_\varrho^{(m)}(\zeta,z) dz \right) d\varrho}}_{:=I_3}, \end{aligned} \tag{4.5.1}$$

for every $r > 0$ such that $\Omega_r^{(m)}(\zeta) \subset \Omega$. We now claim that for every $\zeta \in \Omega$ such that $u(\zeta) = \max_\Omega u$, we have

$$u(z) = u(\zeta), \qquad \text{for every } z \in \Omega_r^{(m)}(\zeta). \tag{4.5.2}$$

In order to prove the claim, we take a closer look to the kernels appearing in (4.5.1) (that we recall to be defined in (4.4.7)). By definition, we notice that $N_\varrho^m(\zeta,z), W_\varrho^{(m)}(\zeta,z) > 0$ for every $z \in \Omega_r^{(m)}(\zeta)$. Moreover, thanks to the ellipticity assumption, the kernel $M_{\mathbb{G}}$ defined in (4.4.2) is nonnegative, thus implying $M_r^{(m)}(\zeta,z) > 0$ for every $z \in \Omega_r^{(m)}(\zeta)$. Since $f \le 0$, $c \ge 0$, $\mathrm{div}_{\mathbb{G}} b - c \ge 0$, and $u(\zeta)$ is a nonnegative maximum point, we deduce that $I_1, I_2, I_3 \le 0$. Hence, the three integrals in (4.5.1) vanish and consequently, by the strictly positivity of the kernel, $u(z) - u(\zeta) = 0$ for $\mathcal{H}^{N+1}$ almost every $z \in \Omega_r^{(m)}(\zeta)$. Claim (4.5.2) then follows from the continuity of u.

We now take advantage of claim (4.5.2) to conclude the proof of Theorem 4.1.4. Let $z \in \mathcal{A}_\zeta(\Omega)$, and let $\gamma : [0,1] \to \Omega$ be an $\mathscr{H}_{\mathbb{G}}$-admissible path such that $\gamma(0) = \zeta$ and $\gamma(1) = z$. We want to prove that $u\Big(\gamma(s)\Big) = u(\zeta)$ for every $s \in [0,1]$. To this end we set

$$I := \Big\{ t \in [0,1] \, : \, u\Big(\gamma(s)\Big) = u(\zeta) \text{ for every } s \in [0,t] \Big\},$$

and define $\bar{t} := \sup I$. Clearly, $I \neq \emptyset$ since $0 \in I$. Moreover, by the continuity of u and γ, I is closed, hence $\bar{t} \in I$. Suppose by contradiction that $\bar{t} < 1$, then we set $\bar{z} = (\bar{x}, \bar{t}) := \gamma\Big(\bar{t}\Big)$, and we recall that $\zeta \in \Omega$ and $u(\zeta) = \max_\Omega u$. We claim that there exists $r_1, s_1 > 0$ such that $\Omega_{r_1}^{(m)}(\bar{z}) \subset \Omega$, and

$$\gamma\Big(\bar{t} + s\Big) \in \Omega_{r_1}^{(m)}(\bar{z}), \qquad \text{for every } s \in [0, s_1). \tag{4.5.3}$$

Then, as a consequence of (4.5.2), we get $u\Big(\gamma(\bar{t}+s)\Big) = u(\bar{z}) = u(\zeta)$ for every $s \in (0, s_1)$, which contradicts $\bar{t} = \sup I$ and concludes the proof of our statement.

Thus, we are left with proving claim (4.5.3). Since $\bar{z} \in \Omega$ by the definition of $\mathscr{H}_{\mathbb{G}}$-admissible path, by Gaussian estimates we infer that there exists $\bar{r}$ small enough such that $\Omega_{\bar{r}}^{(m)}(\bar{z}) \subset \Omega$. Hence, to prove the claim we only need to show that there exist $r_1 \le \bar{r}$ and $s_1 > 0$ such that (4.4.3) holds. From *(5.)* and *(8.)* in Theorem 4.3.11, we infer the existence of $r_1 \le \bar{r}, c, \bar{c} > 0$ (depending only on structural assumptions and m) such that

$$\Omega_{r_1}^{(m)}(\bar{z}) \supseteq \Omega_{r_1}^{*}(\bar{z}) := \Big\{ (x,t) \in \mathbb{R}^N \times \Big(\bar{t}-1, \bar{t}\Big) : d_X(x,\bar{x})^2 \le c\Big(\bar{t}-t\Big)\Big(\ln\Big(r_1\bar{c}\Big) - \ln\Big(\bar{t}-t\Big)\Big)\Big\}.$$

Therefore, to prove (4.5.3) it is enough to show that there exists s_1 such that $\gamma\Big(\bar{t}+s\Big) \in \Omega_{r_1}^{*}(\bar{z})$ for every $s < s_1$. By definition of $\mathscr{H}_{\mathbb{G}}$-admissible path we have

$$\gamma\Big(\bar{t}+s\Big) = \Big(\gamma_*(\bar{t}+s), \bar{t}-s\Big)$$

for some X-trajectory γ_* with control function $\omega = (\omega_1, \ldots, \omega_{m_1})$, namely $\gamma_* \in \mathcal{C}$ such that

$$\dot{\gamma_*}(\tau) = \sum_{i=1}^{m_1} \omega_i(\tau) X_i\Big(\gamma^*(\tau)\Big), \qquad \tau \in [0,1].$$

We set

$$\gamma^*(\tau) := \gamma_*\Big(\bar{t} + \tau s\Big), \qquad \tau \in [0,1].$$

Clearly, γ^* is an X-trajectory with control function

$$\omega^*(\tau) = s\,\omega\Big(\bar{t} + \tau s\Big), \qquad \tau \in [0,1],$$

satisfying $\gamma^*(0) = \gamma_*(\bar{t}) = \bar{x}$ and $\gamma^*(1) = \gamma_*(\bar{t}+s)$. Hence, by (4.2.5) we obtain

$$d_X\Big(\gamma_*(\bar{t}+s), \bar{x}\Big) = d_X^{(2)}\Big(\gamma_*(\bar{t}+s), \bar{x}\Big) \leq \sqrt{\int_{\bar{t}}^{\bar{t}+s} s|\omega(\tau)|^2 d\tau} \leq \sqrt{s}\,\|\omega\|_{L^2(0,1)}.$$

Consequently, $\gamma\Big(\bar{t}+s\Big) = \Big(\gamma_*(\bar{t}+s), \bar{t}-s\Big)$ belongs to $\Omega^*_{r_1}(\bar{z})$ provided we pick $s_1 \leq r_1 \bar{c}\,\exp\Big(-\|\omega\|^2_{L^2(0,1)}/c\Big)$, and this concludes the proof of claim (4.5.3).
To summarize the above, we proved that $u(z) = u(\zeta)$ for every $z \in \mathcal{A}_\zeta(\Omega)$. By the continuity of u we conclude that this holds also for every $z \in \overline{\mathcal{A}_\zeta(\Omega)}$. Finally, since u is constant in $\overline{\mathcal{A}_\zeta(\Omega)}$, by $\mathscr{H}_{\mathbb{G}} u = f$ we conclude that $u(\zeta)c(z) = f(z)$ for every $z \in \overline{\mathcal{A}_\zeta(\Omega)}$.

We finally point out that the condition on the sign of $u(\zeta)$ was exploited only to ensure that $u(\zeta)c(z)$ has the desired sign. If $c \equiv 0$, the needed condition is always satisfied, hence we conclude that $f = 0$. This concludes the proof of Theorem 4.1.4.

4.6 Parabolic Harnack inequality

In this section we establish a parabolic-type Harnack inequality, which will be crucial in deriving the invariant Harnack inequality stated in Theorem 4.1.6. To this end, we first recall the notion of $\mathbb{G}$-ball of radius r and center x_0 and of $\mathbb{G}$-parabolic cylinder of radius r and center $z_0 = (x_0, t_0)$

$$B_r(x_0) = \Big\{x \in \mathbb{R}^N \,:\, d_X(x, x_0) \leq r\Big\}, \qquad Q_r(z_0) = B_r(x_0) \times \Big(t_0 - r^2, t_0\Big),$$

where d_X denotes the Carnot-Carathéodory distance defined in Definition 4.2.8.

Proposition 4.6.15. *Let $\mathscr{H}_{\mathbb{G}}$ be a differential operator of the form* (4.1.2) *satisfying assumptions **(H1)**, **(H2)**, **(H3)** and **(H4)**. Then, there exist four positive constants $C_P, r_1, \varepsilon_1, \vartheta_1$ (only depending on structural assumptions), with ε_1, $\vartheta_1 \in (0,1)$, such that for every $z_0 \in \Omega$ and every $r > 0$ with $r \leq r_1$ and $Q_r(z_0) \subset \Omega$, the following estimate holds*

$$\sup_{D_r(z_0)} u \leq C_P\, u(z_0), \tag{4.6.1}$$

for every $u \geq 0$ classical solution to $\mathscr{H}_{\mathbb{G}} u = 0$ in Ω, where

$$D_r(z_0) := B_{\vartheta_1 r}(x_0) \times \Big\{t = t_0 - \varepsilon_1 r^2\Big\}. \tag{4.6.2}$$

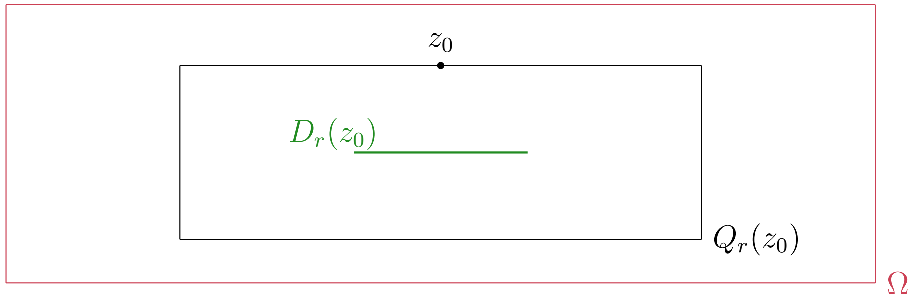

Figure 4.2: The sets appearing in the statement of Proposition 4.6.15 (the time variable is represented vertically, and upwardly increasing).

We first describe the strategy of the proof, which relies on the improved mean value formula derived in Proposition 4.4.14. We begin by discussing a similar strategy employed for uniformly parabolic operators with Hölder continuous coefficients. Then we highlight the main difficulties arising in our framework and outline how we overcome them.

In the uniformly parabolic setting, a key properties of (4.4.9) is the boundedness of the kernel $M_r^{(m)}$: more precisely, there exist $m^-, m^+ > 0$ and $r_0, \varepsilon_0 \in (0,1)$ such that

$$\begin{aligned} &M_r^{(m)}\Big(\xi,\tau,x,t\Big) \le m^+ r^{\frac{m-2}{m+N}}, && \text{for every } (x,t) \in \Omega_r^{(m)}\Big(\xi,\tau\Big), \\ &M_{5r}^{(m)}\Big(x_0,t_0,\xi,\tau\Big) \ge m^- r^m, && \text{for every } \Big(\xi,\tau\Big) \in \Omega_{4r}^{(m)}(x_0,t_0) \cap \Big\{\tau \le t_0 - \varepsilon\, r^2\Big\}, \end{aligned} \tag{4.6.3}$$

whenever $r \in (0, r_0)$ and $\varepsilon \in (0, \varepsilon_0)$. In the constant coefficient case, (4.6.3) follows from the explicit expression of the corresponding fundamental solution. For instance, in the case of the heat equation in $\mathbb{R} \times (0,T)$, one has, for every $r > 0$ and for all $(\xi,\tau), (x,t) \in \mathbb{R} \times (0,T)$,

$$\begin{aligned} M_r^{(m)}(\xi,\tau,x,t) = \frac{2^m m \omega_m}{(m+2)(\tau-t)} \left((\tau - t) \ln \left(\frac{r}{(4\pi(\tau-t))^{\frac{m+1}{2}}} \right) - \frac{(\xi-x)^2}{4} \right)^{\frac{m}{2}} \cdot \\ \cdot \left(\ln \left(\frac{r}{(4\pi(\tau-t))^{\frac{m+1}{2}}} \right) - \frac{(\xi-x)^2}{2m(\tau-t)} \right), \end{aligned}$$

and

$$\Omega_r^{(m)}(\xi,\tau) = \left\{ (\xi - x)^2 < 4(\tau-t) \ln \left(\frac{r}{(4\pi(\tau-t))^{\frac{m+1}{2}}} \right) \right\}.$$

It is straightforward to verify the validity of (4.6.3) when $m > 2$. We now define the set

$$K_{r,\varepsilon}(x_0,t_0) := \overline{\Omega_r^{(m)}(x_0,t_0)} \cap \Big\{ t \le t_0 - \varepsilon\, r^2 \Big\},$$

for some $r \in (0, r_0)$ and $\varepsilon \in (0, \varepsilon_0)$. Then, thanks again to the explicit expression of the fundamental solution, it is easy to show that there exist constants $\vartheta, r_0, \varepsilon_0 \in (0,1)$ such that the following assertion holds

$$\Omega^{(m)}_{\vartheta r}(z) \subset \Omega^{(m)}_{4r}(z_0) \cap \left\{ t \leq t_0 - \varepsilon\, r^2 \right\} \tag{4.6.4}$$

for every $r \in (0, r_0)$, $\varepsilon \in (0, \varepsilon_0)$ and $z \in K_{r,\varepsilon}(z_0)$. Then, the parabolic Harnack inequality follows from an elementary argument similar to the one employed for harmonic functions: let (ξ, τ) be a point in the set $K_{r,\varepsilon}(x_0, t_0)$, for some $r \in (0, r_0)$ and $\varepsilon \in (0, \varepsilon_0)$, then there exists $\vartheta \in (0, 1)$ such that

$$\begin{aligned}
u(\xi,\tau) &\overset{(4.4.9)}{=} \frac{1}{\vartheta r} \int_{\Omega^{(m)}_{\vartheta r}(\xi,\tau)} \widetilde{M}^{(m)}_{\vartheta r}(\xi,\tau,x,t) u(x,t)\, dxdt \\
&\overset{(4.6.3)}{\leq} \frac{m^+ (\vartheta r)^{\frac{m-2}{m+N}}}{\vartheta r} \int_{\Omega^{(m)}_{\vartheta r}(\xi,\tau)} u(x,t)\, dxdt \\
&\overset{(4.6.4)}{\leq} \frac{m^+ (\vartheta r)^{\frac{m-2}{m+N}}}{\vartheta r} \int_{\Omega^{(m)}_{4r}(x_0,t_0)\cap\{\tau \leq t_0 - \varepsilon\, r^2\}} u(x,t)\, dxdt \\
&\overset{(4.6.3)}{\leq} \frac{5 m^+ (\vartheta r)^{\frac{m-2}{m+N}}}{\vartheta m^- r^m} \frac{1}{5r} \int_{\Omega^{(m)}_{5r}(x_0,t_0)} M^{(m)}_{5r}(x_0,t_0,x,t) u(x,t)\, dxdt \\
&\overset{(4.4.9)}{\leq} \frac{5 m^+ r_0^{\frac{m-2}{m+N} - m}}{m^- \vartheta^{\frac{2+N}{m+N}}} u(x_0,t_0).
\end{aligned}$$

The Harnack inequality (4.6.1) follows from the inclusion $D_r(z_0) \subseteq K^{(m)}_{5r,\varepsilon}(x_0,t_0) \subseteq Q_r(z_0)$. In the Hölder continuous coefficients case the explicit expression of the fundamental solution Γ is no longer available. As explained above, the existence of Γ is proved employing the parametrix method outlined in Section 4.3. A key feature of this method is that Γ can be written as the sum of the parametrix Z and a perturbative term (see (4.3.18)). Since the parametrix Z is the fundamental solution of an operator with constant coefficients, we have its explicit expression. Consequently, if one can show that Γ is locally equivalent to Z (in analogy with the stationary case, see (3.5.7)), then the desired bounds for the kernel, equation (4.6.3), follow as in the constant coefficients case. The proof of this local equivalence relies on a perturbative argument originally introduced by Polidoro in [102] and recently employed in [105]. This technique relies on a quantitative dependence of the fundamental solution on the principal part matrix A, which can be obtained thanks to the explicit form of the parametrix Z. More precisely, let A^-, A^+ be two symmetric positive definite matrices satisfying

$$\frac{1}{\Lambda}\, \mathbb{I}_{m_1} \leq A^-, A^+ \leq \Lambda\, \mathbb{I}_{m_1}, \qquad \text{and} \qquad \mathbb{O} < A^+ - A^- < \mu\, \mathbb{I}_{m_1},$$

for some $\Lambda, \mu > 0$. Here, $\mathbb{I}_{m_1}$ denotes the m_1-dimensional unit matrix and $\mathbb{O}$ the zero matrix. Denote by Γ_{A^-} and Γ_{A^+} the fundamental solutions of the constant coefficients operators

$$\mathscr{H}^- = \mathrm{div}(A^-\nabla) - \partial_t, \qquad \text{and} \qquad \mathscr{H}^+ = \mathrm{div}(A^+\nabla) - \partial_t.$$

Then, for every $K > 0$ the following estimates hold

$$\begin{cases} \Gamma_{A^-}(z,\zeta) \leq c_1 \Gamma_{A^+}(z,\zeta), & \forall z,\zeta \in \mathbb{R}^{N+1}, \\ \Gamma_{A^+}(z,\zeta) \leq c_2 \Gamma_{A^-}(z,\zeta)(t-\tau)^{-\frac{Q}{2}c_3\mu}, & \forall z,\zeta : \Gamma_{A^+}(z,\zeta) \geq K, \end{cases} \tag{4.6.5}$$

where c_1 and c_3 are structural constants, while c_2 depends also on K.
In order to apply Polidoro's perturbative method in our subelliptic framework, we need an estimate of the form (4.6.5) for the fundamental solution Γ_A of the constant coefficients operator

$$\mathscr{H}_A = \sum_{i,j=1}^{m_1} a_{ij} X_i X_j - \partial_t, \tag{4.6.6}$$

for $A \in \mathbf{M}_\lambda$. We prove a version of the estimates in (4.6.5) by employing the results in [25, 28], provided that $\mathbb{G}$ is free. The key difficulty in the proof of (4.6.5) lies in the lack of an explicit expression for the fundamental solution of (4.6.6), which prevents a direct derivation of the kernel bounds. However, when $\mathbb{G}$ is a free Carnot group, this obstacle is partially overcome thanks to a result in [28], which ensures the existence of a Lie group automorphism T_A that turns $\mathscr{H}_A$ into the canonical heat operator $\Delta_{\mathbb{G}} - \partial_t$. More precisely, the fundamental solution Γ_A satisfies

$$\Gamma_A(z,\zeta) = \Gamma_A\Big(\xi^{-1} \circ x, t-\tau\Big) = J_A\, \Gamma_0\Big(T_A\Big(\xi^{-1} \circ x\Big), t-\tau\Big), \tag{4.6.7}$$

where Γ_0 is the fundamental solution of the canonical heat operator, and $J_A = \Big|\mathrm{det}\mathcal{J}_{T_A}(x)\Big|$, with $\mathcal{J}_{T_A}$ denoting the Jacobian matrix of T_A. By exploiting the properties of the automorphism T_A, we are able to derive the second estimate in (4.6.5) when dealing with a free Carnot group. Regarding the first estimate in (4.6.5), we are able to prove it in a weaker setting, namely when A^- and A^+ are scalar multiples of the identity matrix (see Lemma 4.6.16 below). However, combining Lemma 4.6.16 with the aforementioned automorphism, we are able to apply Polidoro's method and establish the local comparison between Γ and its parametrix, Proposition 4.6.20, in the context of a free Carnot group.

To summarize, we have described the strategy to prove the parabolic Harnack inequality whenever $\mathbb{G}$ is free. To address the general case, *i.e.*, when $\mathbb{G}$ is not assumed to be free, we adopt the following approach: we lift $\mathbb{G}$ to a free Carnot group $\widetilde{\mathbb{G}}$ (as explained in Section 4.2); then, the Harnack inequality is obtained for the corresponding lifted operator $\mathscr{H}_{\widetilde{\mathbb{G}}}$; finally, since every classical solution to $\mathscr{H}_{\mathbb{G}} u = 0$ solves also $\mathscr{H}_{\widetilde{\mathbb{G}}} u = 0$, the Harnack inequality obtained in the lifted group $\widetilde{\mathbb{G}}$ directly implies the validity of Proposition 4.6.15 in the original group $\mathbb{G}$.

Let $\widetilde{\mathbb{G}}$ be the free Carnot group introduced in Section 4.2, which lifts $\mathbb{G}$, and let $\left\{\widetilde{X}_i\right\}_{i=1}^{m_1}$ be the corresponding set of vector fields. The lifted operator associated to (4.1.2) is

$$\mathscr{H}_{\widetilde{\mathbb{G}}}\widetilde{u}(\widetilde{z}) := \sum_{i,j=1}^{m_1} \widetilde{X}_i\Big(a_{ij}(z)\widetilde{X}_j\Big)\widetilde{u}(\widetilde{z}) + \sum_{i=1}^{m_1} b_i(z)\widetilde{X}_i\widetilde{u}(\widetilde{z}) + c(z)\widetilde{u}(\widetilde{z}) - \partial_t\widetilde{u}(\widetilde{z}), \quad \widetilde{z}\in\widetilde{\Omega}, \tag{4.6.8}$$

where a_{ij}, b_i and c are the coefficients appearing in (4.1.2), assumed to satisfy **(H4)**. Here, $\widetilde{z}$ denote the point $(\widetilde{x}, x, t) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times \mathbb{R}$, and we have set $\widetilde{\Omega} := \mathbb{R}^{\widetilde{N}} \times \Omega$. For the rest of this section, we will use the letters $\widetilde{z}_0$ and $\widetilde{\zeta}$ to denote respectively the points $(\widetilde{x}_0, x_0, t_0)$ and $(\widetilde{\xi}, \xi, \tau)$ in $\widetilde{\Omega}$, and z, z_0 and ζ to denote respectively the points $(x,t), (x_0,t_0)$ and (ξ,τ) in Ω. Notice that the coefficients a_{ij}, b_i and c are bounded and α-Hölder continuous with respect to the $\widetilde{\mathbb{G}}$-parabolic metric induced by the distance $d_{\widetilde{\mathbb{G}}}$ defined in (4.2.10). Hence, Theorem 4.3.11 applies to $\mathscr{H}_{\widetilde{\mathbb{G}}}$, ensuing the existence of a fundamental solution $\widetilde{\Gamma}$ to (4.6.8), and Proposition 4.4.14 yields the following representation formula for any $\widetilde{u}$ classical solution to $\mathscr{H}_{\widetilde{\mathbb{G}}}\widetilde{u} = 0$

$$\begin{aligned}\widetilde{u}(\widetilde{\zeta}) &= \frac{1}{r}\int_{\widetilde{\Omega}_r^{(m)}(\widetilde{\zeta})} \widetilde{M}_r^{(m)}(\widetilde{\zeta},\widetilde{z})\widetilde{u}(\widetilde{z})\,d\widetilde{z} \\ &+ \frac{1}{r}\int_0^r \frac{\omega_m}{\varrho}\left(\int_{\widetilde{\Omega}_\varrho^{(m)}(\widetilde{\zeta})} \widetilde{N}_\varrho^m(\widetilde{\zeta},\widetilde{z})\left(\operatorname{div}_{\mathbb{G}} b(z) - c(z)\right)\widetilde{u}(\widetilde{z})\,d\widetilde{z}\right)d\varrho.\end{aligned} \tag{4.6.9}$$

Here, $\widetilde{\Omega}_r^{(m)}$, $\widetilde{M}_r^{(m)}$, and $\widetilde{N}_r^{(m)}$ are the lifted analogues of the quantities introduced in (4.4.7), in the context of $\widetilde{\mathbb{G}}$.

The following observation will be crucial at the end of the section: if u is a classical solution to $\mathscr{H}_{\mathbb{G}} u = 0$ in Ω, then by the lifting property (4.2.9) we infer that the function $(\widetilde{x}, x, t) \to \widetilde{u}(\widetilde{x}, x, t) := u(x,t)$ is a classical solution to $\mathscr{H}_{\widetilde{\mathbb{G}}}\widetilde{u} = 0$ in $\widetilde{\Omega}$.

As previously mentioned, our goal is to establish bounds of the kind (4.6.3) for the kernel $\widetilde{M}_r^{(m)}$ in terms of r. To this end, we begin by deriving some estimates for the fundamental solution of the constant coefficients lifted operator

$$\widetilde{\mathscr{H}_A} = \sum_{i,j=1}^{m_1} a_{ij}\widetilde{X}_i\widetilde{X}_j - \partial_t, \tag{4.6.10}$$

focusing on its dependence on the coefficients matrix $A \in \mathbf{M}_\lambda$.

Lemma 4.6.16. *Assume that **(H1)**,**(H2)** and **(H3)** hold, and pick $0 < \mu^- < \mu^+$. Denote by $\widetilde{\Gamma}_{A^-}$ and $\widetilde{\Gamma}_{A^+}$ the fundamental solutions of* (4.6.10) *corresponding to the matrices $A^- := \mu^-\mathbb{I}_{m_1}$ and $A^+ := \mu^+\mathbb{I}_{m_1}$, respectively. Then, there exists $c_1 > 0$ such that the following* *inequality holds*

$$\widetilde{\Gamma}_{A^-}(\widetilde{x},x,t) \le c_1\,\widetilde{\Gamma}_{A^+}(\widetilde{x},x,t), \qquad \forall (\widetilde{x},x,t)\in\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N\times(0,+\infty).$$

Proof. Since $\widetilde{\Gamma}_{A^-}$ and $\widetilde{\Gamma}_{A^+}$ are the fundamental solutions of

$$\widetilde{\mathscr{H}_{A^-}} = \mu^-\Delta_{\widetilde{\mathbb{G}}} - \partial_t, \qquad \widetilde{\mathscr{H}_{A^+}} = \mu^+\Delta_{\widetilde{\mathbb{G}}} - \partial_t,$$

the thesis immediately follows from a homogeneity argument. □

Lemma 4.6.17. *Assume that **(H1)**,**(H2)** and **(H3)** hold, and let A^-, A^+ be two positive definite, symmetric matrices satisfying*

$$\frac{1}{\Lambda}\,\mathbb{I}_{m_1} \le A^-, A^+ \le \Lambda\,\mathbb{I}_{m_1}, \qquad \textit{and} \qquad \mathbb{O} < A^+ - A^- < \mu\,\mathbb{I}_{m_1}, \tag{4.6.11}$$

for some $\Lambda,\mu>0$. Denote by $\widetilde{\Gamma}_{A^-}$ and $\widetilde{\Gamma}_{A^+}$ the fundamental solution of (4.6.10) *corresponding to the matrices A^- and A^+, respectively. Then, there exists $c_2, c_3 > 0$ such that for every $K>0$ it holds*

$$\widetilde{\Gamma}_{A^+}(\widetilde{x},x,t) \le c_2 K^{-c_3\mu} t^{-\frac{\widetilde{Q}}{2}c_3\mu}\,\widetilde{\Gamma}_{A^-}(\widetilde{x},x,t), \tag{4.6.12}$$

for every $(\widetilde{x},x,t)\in\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N\times(0,+\infty)$ such that $\widetilde{\Gamma}_{A^+}(\widetilde{x},x,t)\ge K$. Here, we recall that $\widetilde{Q}$ denote the homogeneous dimension of the lifted group $\widetilde{\mathbb{G}}$.

Proof. Since $\widetilde{\mathbb{G}}$ is free, Theorem 1.1 in [28] yields the existence of two Lie group automorphisms T_{A^-} and T_{A^+} such that

$$\begin{aligned}
\widetilde{\Gamma}_{A^-}(\widetilde{x},x,t) &= J_{A^-}\,\widetilde{\Gamma}_0\Big(T_{A^-}(\widetilde{x},x),t\Big) = J_{A^-}\,\frac{\exp\left(-\frac{d_0\left(T_{A^-}(\widetilde{x},x)\right)^2}{t}\right)}{t^{\frac{\widetilde{Q}}{2}}},\\
\widetilde{\Gamma}_{A^+}(\widetilde{x},x,t) &= J_{A^+}\,\widetilde{\Gamma}_0\Big(T_{A^+}(\widetilde{x},x),t\Big) = J_{A^+}\,\frac{\exp\left(-\frac{d_0\left(T_{A^+}(\widetilde{x},x)\right)^2}{t}\right)}{t^{\frac{\widetilde{Q}}{2}}},
\end{aligned} \tag{4.6.13}$$

where $\widetilde{\Gamma}_0$ denotes the fundamental solution of the canonical heat operator $\sum_i \widetilde{X}_i^2 - \partial_t$, d_0 is the corresponding homogeneous distance on $\widetilde{\mathbb{G}}$, and $J_A = \det\mathcal{J}_{T_A}(\widetilde{x},x)$, with $\mathcal{J}_{T_A}$ denoting the Jacobian matrix of T_A. By Theorem 2.7 in [28] we infer that J_A is constant, and that there exists $c_\Lambda > 0$ (only depending on the ellipticity constant Λ) such that

$$\begin{aligned} c_\Lambda^{-1} &\le J_{A^-}, J_{A^+} \le c_\Lambda, \\ c_\Lambda^{-1} d_0(\widetilde{x}, x) &\le d_0\Big(T_{A^-}(\widetilde{x}, x)\Big), d_0\Big(T_{A^+}(\widetilde{x}, x)\Big) \le c_\Lambda d_0(\widetilde{x}, x), \\ d_0\Big((T_{A^-}(\widetilde{x}, x))^{-1} \,\widetilde{\circ}\, T_{A^+}(\widetilde{x}, x)\Big) &\le c_\Lambda \|A^+ - A^-\|^{\frac{1}{\kappa}} d_0(\widetilde{x}, x), \end{aligned} \tag{4.6.14}$$

for any $(\widetilde{x}, x) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N$, where we recall that $\widetilde{\circ}$ is the composition law of $\widetilde{\mathbb{G}}$, and κ its step. From the second inequality in (4.6.14), together with the assumption $\widetilde{\Gamma}_{A^+}(\widetilde{x}, x, t) \ge K$, we infer

$$\frac{d_0(\widetilde{x}, x)^2}{t} \le c_\Lambda \ln\left(\frac{1}{t^{\frac{\widetilde{Q}}{2}} K}\right). \tag{4.6.15}$$

The desired estimate (4.6.12) then follows from (4.6.13), (4.6.14), and (4.6.15), concluding the proof of the lemma. $\square$

Remark 4.6.18. *When dealing with uniformly parabolic operators,* i.e. *in the Euclidean setting, the estimate given by Lemma 4.6.16 is known to hold even if A^- and A^+ are more general matrices verifying* (4.6.11)*. The same estimate has also been established for a broad class of degenerate Kolmogorov operators (see for instance [102, 105]). Moreover, some precise estimates obtained for the heat kernel in the Heisenberg group in [84, 104] seem to suggest the validity of such estimates for Carnot groups in general. A precise characterization of this property and its general validity remains an interesting open problem.*

Lemma 4.6.19. *Assume that **(H1)**,**(H2)** and **(H3)** hold, let $\frac{1}{\Lambda}\,\mathbb{I}_{m_1} < A < \Lambda\,\mathbb{I}_{m_1}$ be a symmetric positive matrix, and let $\widetilde{\Gamma}_A$ be the fundamental solution of* (4.6.10) *with matrix A. Then, there exists $c > 0$ (depending only on structural assumptions and Λ) such that*

$$|X_i \widetilde{\Gamma}_A(\widetilde{x}, x, t)| \le c\left(\frac{1}{\sqrt{t}} \frac{d_{\widetilde{\mathbb{G}}}(\widetilde{x}, x)}{\sqrt{t}}\right) \widetilde{\Gamma}_A(\widetilde{x}, x, t), \qquad \forall (\widetilde{x}, x, t) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times (0, +\infty).$$

Proof. As shown in the proof of Lemma 4.6.17, there exists a Lie group automorphism T_A and a homogeneous distance d_0 (smooth out of the origin and homogeneous of degree one w.r.t. the dilations $\widetilde{\delta}_r$) such that

$$\widetilde{\Gamma}_A(\widetilde{x}, x, t) = J_A \frac{\exp\left(-\frac{d_0\left(T_A\left(\widetilde{x}, x\right)\right)^2}{t}\right)}{t^{\frac{\widetilde{Q}}{2}}}.$$

By the chain rule we get

$$\widetilde{X}_i \widetilde{\Gamma}_A(\widetilde{x}, x, t) = \widetilde{\Gamma}_A(\widetilde{x}, x, t) \left(-\frac{2}{\sqrt{t}} \frac{d_0\Big(T_A(\widetilde{x}, x)\Big)}{\sqrt{t}} \widetilde{X}_i \Big(d_0 \Big(T_A(\widetilde{x}, x) \Big) \Big) \right). \tag{4.6.16}$$

From the second estimate in (4.6.14), together with the equivalence of the homogeneous distances d_0 and $d_{\widetilde{\mathbb{G}}}$, we infer

$$d_0\Big(T_A(\widetilde{x}, x)\Big) \leq c\, d_{\widetilde{\mathbb{G}}}(\widetilde{x}, x),$$

for some $c > 0$ depending on Λ. Moreover, from Theorem 2.2 in [28] it follows that

$$X_i \Big(d_0 \Big(T_A(\widetilde{x}, x) \Big) \Big) = \sum_{j=1}^{m_1} \Big(A^{-\frac{1}{2}} \Big)_{ij} X_j d_0 \Big(T_A(\widetilde{x}, x) \Big).$$

Since d_0 is homogeneous of degree one, its Lie derivatives are homogeneous of degree zero and hence bounded away from the origin. Therefore, $X_i\Big(d_0\Big(T_A(\widetilde{x}, x)\Big)\Big)$ is uniformly bounded for all A satisfying the ellipticity bounds.
The lemma is an immediate consequence of (4.6.16) and the above observations. $\square$

We are now in position to establish a local comparison between the fundamental solution $\widetilde{\Gamma}$ of the lifted operator (4.6.8) and its parametrix $\widetilde{Z}$.

Proposition 4.6.20. *Assume that **(H1)**, **(H2)**, **(H3)** and **(H4)** hold. Then, for every $\eta > 0$ there exists $K > 0$ (depending only on η and structural assumptions) such that*

$$(1 - \eta)\widetilde{Z}(\widetilde{z}, \widetilde{\zeta}) \leq \widetilde{\Gamma}(\widetilde{z}, \widetilde{\zeta}) \leq (1 + \eta)\widetilde{Z}(\widetilde{z}, \widetilde{\zeta}), \tag{4.6.17}$$

for every $\widetilde{z}, \widetilde{\zeta} \in O_K := \Big\{ \widetilde{Z}(\widetilde{z}, \widetilde{\zeta}) \geq K \Big\}$.

Proof. Without loss of generality we may assume $\widetilde{\zeta} = 0$ and adopt the simpler notation $\widetilde{\Gamma}(\widetilde{z}, 0) = \widetilde{\Gamma}(\widetilde{z})$ and $\widetilde{Z}(\widetilde{z}, 0) = \widetilde{Z}(\widetilde{z})$. We also recall that we here employ the notation $\widetilde{z} = (\widetilde{x}, x, t), \widetilde{\zeta} = (\widetilde{\xi}, \xi, \tau) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times \mathbb{R}$.

We first assume that $A(0) = \mathbb{I}_{m_1}$, and later we explain how to remove such assumption.
Let $\psi \in C^2(\mathbb{R})$ be a cut-off function such that $\psi(w) = 0$ if $w \geq 1$, and $\psi(w) = 1$ if $w \leq 1/2$. For a fixed $r > 0$ to be chosen later, we set $A^r(\widetilde{z}) := A(0) + \psi\Big(\widetilde{\delta}_{\frac{1}{r}} d_{\widetilde{\mathbb{G}}}(\widetilde{z})\Big)\Big(A(\widetilde{z}) - A(0)\Big)$,

and consider the perturbed operator

$$\mathscr{H}^r_{\widetilde{\mathbb{G}}}\widetilde{u}(\widetilde{z}) := \sum_{i,j=1}^{m_1} \widetilde{X}_i\Big(a^r_{ij}(\widetilde{z})\widetilde{X}_j\Big)\widetilde{u}(\widetilde{z}) + \sum_{i=1}^{m_1} b_i(z)\widetilde{X}_i\widetilde{u}(\widetilde{z}) + c(z)\widetilde{u}(\widetilde{z}) - \partial_t\widetilde{u}(\widetilde{z}), \qquad \widetilde{z} \in \widetilde{\Omega}. \tag{4.6.18}$$

By Theorem 4.3.11, $\mathscr{H}^r_{\widetilde{\mathbb{G}}}$ has a fundamental solution $\widetilde{\Gamma}^r$ given by

$$\widetilde{\Gamma}^r(\widetilde{z}) = \widetilde{Z}^r(\widetilde{z}) + \int_0^t \int_{\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N} \widetilde{Z}^r(\widetilde{z},\widetilde{y},y,s)\widetilde{G}^r(\widetilde{y},y,s,0)d\widetilde{y}\,dy\,ds,$$

where $\widetilde{Z}^r$ is the parametrix, and $\widetilde{G}^r$ is given by

$$\widetilde{G}^r(\widetilde{z},0) = \sum_{k=1}^{\infty}\Big(\mathscr{H}^r_{\widetilde{\mathbb{G}}}\widetilde{Z}^r\Big)_k(\widetilde{z},0),$$

where $\Big(\mathscr{H}^r_{\widetilde{\mathbb{G}}}\widetilde{Z}^r\Big)_1 := \mathscr{H}^r_{\widetilde{\mathbb{G}}}\widetilde{Z}^r$ and

$$\Big(\mathscr{H}^r_{\widetilde{\mathbb{G}}}\widetilde{Z}^r\Big)_{k+1}(\widetilde{z},0) := \int_0^t \int_{\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N}\Big(\mathscr{H}^r_{\widetilde{\mathbb{G}}}\widetilde{Z}^r\Big)_1(\widetilde{z},\widetilde{y},y,s)\Big(\mathscr{H}^r_{\widetilde{\mathbb{G}}}\widetilde{Z}^r\Big)_k(\widetilde{y},y,s,0)d\widetilde{y}\,dy\,ds.$$

Observe that, by construction, $A^r(0) = A(0)$, hence from the definition of parametrix we deduce $\widetilde{Z}^r(\widetilde{z}) = \widetilde{Z}(\widetilde{z})$ for every $\widetilde{z} \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times \mathbb{R}$. For the same reason, by the properties of ψ we get

$$\widetilde{Z}^r(\widetilde{z},\widetilde{\zeta}) = \widetilde{Z}(\widetilde{z},\widetilde{\zeta}), \qquad \text{if } d_{\widetilde{\mathbb{G}}}(\widetilde{\zeta}) \leq \frac{r}{2}. \tag{4.6.19}$$

We now prove the comparison (4.6.17) by expanding (4.3.18) as follows

$$\begin{aligned}
\widetilde{\Gamma}(\widetilde{z}) &= \widetilde{Z}(\widetilde{z}) + \int_0^t \int_{\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N} \widetilde{Z}^r(\widetilde{z},\widetilde{y},y,s)\widetilde{G}^r(\widetilde{y},y,s,0)d\widetilde{y}dyds \\
&\quad + \int_0^t \int_{\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N} \Big(\widetilde{Z}^r(\widetilde{z},\widetilde{y},y,s) - \widetilde{Z}(\widetilde{z},\widetilde{y},y,s)\Big)\widetilde{G}^r(\widetilde{y},y,s,0)d\widetilde{y}dyds \\
&\quad + \int_0^t \int_{\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N} \widetilde{Z}^r(\widetilde{z},\widetilde{y},y,s)\Big(\widetilde{G}(\widetilde{y},y,s,0) - \widetilde{G}^r(\widetilde{y},y,s,0)\Big)d\widetilde{y}dyds \\
&=: \widetilde{Z}(\widetilde{z}) + I + II + III.
\end{aligned} \tag{4.6.20}$$

The proof of the proposition is completed once we show that $|I|, |II|, |III| \leq \frac{\eta}{3}\widetilde{Z}(\widetilde{z})$ on the set $\Big\{\widetilde{z} \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times \mathbb{R} : Z(\widetilde{z}) \geq K\Big\}$.
Let us begin by estimating the term I. We start by fixing $T > 0$ and noticing that, thanks to (4.3.25), we can choose $K = K(T) > 0$ such that

$$O_K \subset \Big\{(\widetilde{x},x,t) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times \mathbb{R} : 0 \leq t \leq T\Big\}. \tag{4.6.21}$$

We now focus our attention on $\widetilde{Z}^r$ and $\widetilde{G}^r$. To this end we apply Lemma 4.6.16 and Lemma

4.6.17 with matrices $A^- := A(0) = A^r(0)$ and $A^+ := A(0) + \mu\,\mathbb{I}_{m_1}$, for some $0 < \mu < \lambda$ to be chosen later, where λ is the ellipticity constant appearing in (4.1.4). By definition $\widetilde{Z}(\widetilde{z}) = \widetilde{\Gamma}_{A^-}(\widetilde{z})$, and the aforementioned lemmas yield

$$\widetilde{Z}(\widetilde{x}, x, t) \le c_1 \widetilde{\Gamma}_{A^+}(\widetilde{x}, x, t), \qquad \forall (\widetilde{x}, x, t) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times (0, +\infty), \tag{4.6.22}$$

and

$$\widetilde{\Gamma}_{A^+}(\widetilde{x}, x, t) \le c_2 K^{-c_3\mu} t^{-\frac{\widetilde{Q}}{2} c_3 \mu}\, \widetilde{Z}(\widetilde{x}, x, t), \tag{4.6.23}$$

in the set $O_K^+ := \left\{ \widetilde{\Gamma}_{A^+}(\widetilde{x}, x, t) \ge K \right\}$, for every fixed $K > 0$. We now choose a positive $r = r(M_1, \mu)$ such that $r^\alpha \le \frac{\mu}{4M_1}$, where α and M_1 are given by **(H4)**. Then by the definition of ψ we obtain

$$\left(1 - \frac{\mu}{2}\right) \mathbb{I}_{m_1} = A(0) - \frac{\mu}{2}\,\mathbb{I}_{m_1} \le A^r(\widetilde{z}) \le A(0) + \frac{\mu}{2}\,\mathbb{I}_{m_1} = \left(1 + \frac{\mu}{2}\right) \mathbb{I}_{m_1} < A^+,$$

for every $\widetilde{z} \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times \mathbb{R}$. From this estimate we deduce that A^r satisfies a uniform ellipticity condition with constant $1+\mu/2$. This fact is crucial, as Theorem B in [25] provides the following Gaussian upper bounds for the parametrix.

$$\widetilde{Z}^r(\widetilde{z}, \widetilde{\zeta}) \le c_\mu \widetilde{\Gamma}_\mu(\widetilde{z}, \widetilde{\zeta}),$$

for some $c_\mu > 0$ depending solely on μ, with $\widetilde{\Gamma}_\mu$ denoting the fundamental solution of $\widetilde{\mathscr{H}_\mu} = (1 + \mu/2)\Delta_{\widetilde{\mathbb{G}}} - \partial_t$. Hence, by Lemma 4.6.16 we deduce the following bound for the parametrix

$$\widetilde{Z}^r(\widetilde{z}, \widetilde{\zeta}) \le c_1 \widetilde{\Gamma}_{A^+}(\widetilde{z}, \widetilde{\zeta}), \tag{4.6.24}$$

in place of the general Gaussian upper bound (4.3.25). Similarly, the derivatives $X_i \widetilde{Z}^r(\widetilde{z}, \widetilde{\zeta})$ and $X_i X_j \widetilde{Z}^r(\widetilde{z}, \widetilde{\zeta})$ can also be bounded by $\widetilde{\Gamma}_{A^+}(\widetilde{z}, \widetilde{\zeta})$ (as in (4.3.3) and (4.3.4)). Hence, by a careful inspection of the computation made in the proof of Theorem 4.3.11, we deduce the following estimate

$$|\widetilde{G}^r(\widetilde{z}, \widetilde{\zeta})| \le \frac{c_T (t-\tau)^{\frac{\alpha}{2}-1}}{(t-\tau)^{\frac{Q}{2}}} \widetilde{\Gamma}_{A^+}(\widetilde{z}, \widetilde{\zeta}), \tag{4.6.25}$$

for some c_T depending only on T. Thus, by (4.6.24), (4.6.25), and the reproduction property we infer that

$$|I| \le \frac{2c_T}{\alpha} T^{\frac{\alpha}{2}} \widetilde{\Gamma}_A^+(\widetilde{x}, x, t), \qquad \forall (\widetilde{x}, x, t) \in O_K.$$

Finally, using (4.6.22), we observe that $O_K \subset O_{K/c_4}^+$. This inclusion, together with (4.6.23), yields

$$|I| \le \frac{2c_T}{\alpha} T^{\frac{\alpha}{2}} \widetilde{\Gamma}_A^+(\widetilde{x}, x, t) \le \frac{C_T}{\alpha} T^{\frac{\alpha}{2}} c_2 K^{-c_3\mu} t^{-\frac{\widetilde{Q}}{2} c_3 \mu}\, \widetilde{Z}(\widetilde{x}, x, t), \qquad \forall (\widetilde{x}, x, t) \in O_K,$$

implying the estimate

$$|I| \le \frac{\eta}{3}\widetilde{Z}(\widetilde{x}, x, t), \qquad \forall(\widetilde{x}, x, t) \in O_K, \tag{4.6.26}$$

provided we pick $\mu = \frac{\alpha}{\widetilde{Q}c_3}$ and $K = K(\mathscr{H}_{\mathbb{G}}, \lambda, \eta)$ sufficiently large.
We now turn to the estimate of term $|II|$ in (4.6.20). By (4.3.25) and (4.6.19), we infer that

$$\sup_{(\widetilde{y},y,s)\in\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N\times\mathbb{R}} \left|\widetilde{Z}^r(\widetilde{z}, \widetilde{y}, y, s) - \widetilde{Z}(\widetilde{z}, \widetilde{y}, y, s)\right| \le c,$$

for some $c > 0$. Hence, from (4.6.25) and the normalization (4.3.10) of $\widetilde{\Gamma}_{A^+}$ we deduce $|II| \le C_{II}T^{\frac{\alpha}{2}}$, for some $C_{II} > 0$ only depending on $\mathscr{H}_{\mathbb{G}}$, λ, r and T. In particular, this implies

$$|II| \le C_{II}T^{\frac{\alpha}{2}}K^{-1}\,\widetilde{Z}(\widetilde{x}, x, t), \qquad \forall(\widetilde{x}, x, t) \in O_K. \tag{4.6.27}$$

Finally, we estimate $|III|$ in an analogous manner. By using the uniform bound

$$\left|\widetilde{G}(\widetilde{y}, y, s, 0) - \widetilde{G}^r(\widetilde{y}, y, s, 0)\right| \le C,$$

whose proof follows the same argument as in Lemma 4.7 of [102], we obtain

$$|III| \le C_{III}T, \qquad \forall(\widetilde{x}, x, t) \in O_K, \tag{4.6.28}$$

for some constant $C_{III} > 0$ only depending on $\mathscr{H}_{\mathbb{G}}$, λ, r and T. Eventually combining (4.6.20), (4.6.26), (4.6.27) and (4.6.28), we infer

$$\widetilde{\Gamma}(\widetilde{z}) \le \left(1 + \frac{\eta}{3} + \frac{C_{II}T^{\frac{\alpha}{2}}}{K} + \frac{C_{III}T}{K}\right)\widetilde{Z}(\widetilde{z}) \le (1+\eta)\widetilde{Z}(\widetilde{z}),$$

provided that we choose $K = K(\mathscr{H}_{\mathbb{G}}, \lambda, \eta, T, \alpha)$ big enough. A symmetric argument yields the corresponding lower estimate $(1-\eta)\widetilde{Z}(\widetilde{z}) \le \widetilde{\Gamma}(\widetilde{z})$.

We now remove the assumption $A(0) = \mathbb{I}_{m_1}$. We consider the constant coefficients operator

$$\widetilde{\mathscr{H}_{A_0}} := \sum_{i,j=1}^{m_1} a_{ij}(0)\widetilde{X}_i\widetilde{X}_j - \partial_t,$$

and let $T_{A_0}(\widetilde{x}, x)$ denote the automorphism turning $\widetilde{\mathscr{H}_{A_0}}$ into the canonical heat operator (as follows from [28]). The change of variable $(x, t) \mapsto \Big(T_{A_0}(\widetilde{x}, x), t\Big) = (\widetilde{y}, y, t)$ turns $\mathscr{H}_{\mathbb{G}}$ into the operator

$$\overline{\mathscr{H}_{\widetilde{\mathbb{G}}}} = \sum_{i,j=1}^{m_1} \widetilde{X}_i\Big(\bar{a}_{ij}(\widetilde{y}, y, t)\widetilde{X}_j\Big) + \sum_{i=1}^{m_1} \bar{b}_i(\widetilde{y}, y, t)\widetilde{X}_i + c(\widetilde{y}, y, t) - \partial_t,$$

where

$$\bar{A}(\widetilde{y},y,t) = A^{-\frac{1}{2}}(0,0,0)A(\widetilde{y},y,t)A^{-\frac{1}{2}}(0,0,0), \qquad \bar{b}(\widetilde{y},y,t) = A^{-\frac{1}{2}}(0,0,0)b(\widetilde{y},y,t).$$

Clearly, $\overline{\mathscr{H}_{\widetilde{\mathbb{G}}}}$ satisfies **(H1)**, **(H2)**, **(H3)** and **(H4)**, and since $\bar{A}(0,0,0) = \mathbb{I}_{m_1}$, the previous result ensures that

$$(1-\eta)\bar{Z}(\widetilde{y},y,t) \leq \bar{\Gamma}(\widetilde{y},y,t) \leq (1+\eta)\bar{Z}(\widetilde{y},y,t),$$

for every $(\widetilde{y},y,t) \in \Big\{\bar{Z}(\widetilde{y},y,t) \geq K\Big\}$. The conclusion then follows from the identities

$$\bar{Z}(\widetilde{y},y,t) = \bar{Z}\Big(T_{A_0}(\widetilde{x},x,t)\Big) = Z(\widetilde{x},x,t), \qquad \bar{\Gamma}(\widetilde{y},y,t) = \bar{\Gamma}\Big(T_{A_0}(\widetilde{x},x,t)\Big) = \Gamma(\widetilde{x},x,t).$$

This completes the proof of the proposition. □

By a similar argument one can show that, locally, the Lie derivatives of $\widetilde{\Gamma}$ can be estimated in terms of $\widetilde{\Gamma}$ itself, rather than relying on the Gaussian upper bound (4.3.3).

Proposition 4.6.21. *Assume that **(H1)**,**(H2)**,**(H3)** and **(H4)** hold. Then, there exist $c, K > 0$ (depending only on structural assumptions) such that*

$$\begin{aligned} |\widetilde{X}_i\widetilde{\Gamma}(\widetilde{z},\widetilde{\zeta})| &\leq c\left(\frac{1}{\sqrt{t-\tau}}d_{\widetilde{\mathbb{G}}}(\widetilde{\xi},\xi,\widetilde{x},x)+1\right)\widetilde{\Gamma}(\widetilde{z},\widetilde{\zeta}), \\ |\widetilde{X}_i\widetilde{\Gamma}(\widetilde{\zeta},\widetilde{z})| &\leq c\left(\frac{1}{\sqrt{t-\tau}}d_{\widetilde{\mathbb{G}}}(\widetilde{x},x,\widetilde{\xi},\xi)+1\right)\widetilde{\Gamma}(\widetilde{z},\widetilde{\zeta}), \end{aligned} \tag{4.6.29}$$

for every $\widetilde{z},\widetilde{\zeta} \in \widetilde{O}_K := \Big\{\widetilde{\Gamma}(\widetilde{z},\widetilde{\zeta}) \geq K\Big\}$.

Proof. The proof is a consequence of Lemma 4.6.19, and follows the same lines as the proof of Proposition 4.6.20. We suppose again, without loss of generality, that $\widetilde{\zeta} = 0$, $A(0) = \mathbb{I}_{m_1}$, and we define the perturbed operator $\mathscr{H}^r_{\widetilde{\mathbb{G}}}$ as in (4.6.18). Thanks to (4.3.3) we can take the Lie derivative $\widetilde{X}_i$ under the integral sign in (4.6.20), obtaining

$$\begin{aligned} \widetilde{X}_i\widetilde{\Gamma}(\widetilde{z}) = \widetilde{X}_i\widetilde{Z}(\widetilde{z}) &+ \int_0^t\int_{\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N}\widetilde{X}_i\widetilde{Z}^r(\widetilde{z},\widetilde{y},y,s)\widetilde{G}^r(\widetilde{y},y,s,0)d\widetilde{y}dyds \\ &+ \int_0^t\int_{\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N}\widetilde{X}_i\Big(\widetilde{Z}^r(\widetilde{z},\widetilde{y},y,s) - \widetilde{Z}(\widetilde{z},\widetilde{y},y,s)\Big)\widetilde{G}^r(\widetilde{y},y,s,0)d\widetilde{y}dyds \\ &+ \int_0^t\int_{\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N}\widetilde{X}_i\widetilde{Z}^r(\widetilde{z},\widetilde{y},y,s)\Big(\widetilde{G}(\widetilde{y},y,s,0) - \widetilde{G}^r(\widetilde{y},y,s,0)\Big)d\widetilde{y}dyds \\ =:&\ \widetilde{X}_i\widetilde{Z}(\widetilde{z}) + I' + II' + III'. \end{aligned}$$

We fix $\eta = 1/2$ and choose $T > 0$. By Proposition 4.6.20 we infer that there exists $K > 0$

such that $\widetilde{O}_K \subset O_{\frac{2}{3}K} \subset \Big\{(\widetilde{x}, x, t) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times \mathbb{R} : 0 \le t \le T\Big\}$. Arguing as in the proof of Proposition 4.6.20 and applying Lemma 4.6.19, we infer that there exists $C > 0$ such that

$$|I'| \le \frac{C}{K^{c_3\mu}} \frac{\widetilde{Z}(\widetilde{x}, x, t)}{\sqrt{t}}, \qquad |II'| \le CT^{\frac{\alpha-1}{2}}, \qquad |III'| \le C\sqrt{T}, \qquad \forall (\widetilde{x}, x, t) \in \widetilde{O}_K,$$

concluding the validity of the first inequality in (4.6.29). In order to prove the second inequality, we first point out that the differentiation is carried out with respect to the adjoint variable $\widetilde{z}$. This is then overcome by invoking the property $\widetilde{\Gamma}(\widetilde{\xi}, \xi, \tau, \widetilde{x}, x, t) = \widetilde{\Gamma}^*(\widetilde{x}, x, t, \widetilde{\xi}, \xi, \tau)$ established in Theorem 4.3.11, and applying the first inequality in (4.6.29). Finally, assumption $A(0) = \mathbb{I}_{m_1}$ is removed as in the proof of Proposition (4.6.20). This concludes the proof of the proposition. □

We are now in position to state and prove a parabolic Harnack inequality for the lifted operator (4.6.8). For every $\widetilde{z}_0 = (\widetilde{x}_0, x_0, t_0) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times \mathbb{R}$, $r, \varepsilon > 0$ and $m \in \mathbb{N}$, we define the truncated kernel set

$$\widetilde{K}^{(m)}_{r,\varepsilon}(\widetilde{x}_0, x_0, t_0) := \overline{\widetilde{\Omega}^{(m)}_r(\widetilde{x}_0, x_0, t_0)} \cap \Big\{t \le t_0 - \varepsilon\, r^2\Big\}. \tag{4.6.30}$$

Proposition 4.6.22. *Let $\mathscr{H}_{\widetilde{\mathbb{G}}}$ be a differential operator of the form* (4.6.8) *satisfying assumptions **(H1)**, **(H2)**, **(H3)** and **(H4)**. For every $m \in \mathbb{N}$ with $m > 2$, there exists three positive constants r_0, ε_0 and C_P (only depending on structural assumptions and m) such that the following holds: for every $(\widetilde{x}_0, x_0, t_0) \in \widetilde{\Omega}$, $0 < r \le r_0$, $0 < \varepsilon \le \varepsilon_0$ and $\widetilde{\Omega}^{(m)}_{5r}(\widetilde{x}_0, x_0, t_0) \subset \widetilde{\Omega}$ we have*

$$\sup_{\widetilde{K}^{(m)}_{r,\varepsilon}(\widetilde{x}_0, x_0, t_0)} \widetilde{u} \le C_P\, \widetilde{u}(\widetilde{x}_0, x_0, t_0), \tag{4.6.31}$$

for every $\widetilde{u} \ge 0$ classical solution to $\mathscr{H}_{\widetilde{\mathbb{G}}}\widetilde{u} = 0$ in $\widetilde{\Omega}$.

Proof. Let $m > 2$, $(\widetilde{x}_0, x_0, t_0) \in \widetilde{\Omega}$ and $r > 0$ such that $\widetilde{\Omega}^{(m)}_{4r}(\widetilde{x}_0, x_0, t_0) \subset \widetilde{\Omega}$. We show that there exist $m^-, m^+ > 0$ and $r_0, \varepsilon_0, \vartheta \in (0, 1)$ such that, for every $r \le r_0$ and $\varepsilon \le \varepsilon_0$, the following properties hold

1. $\widetilde{K}^{(m)}_{r,\varepsilon}(\widetilde{x}_0, x_0, t_0) \ne \emptyset$;
2. $\widetilde{M}^{(m)}_{\vartheta r}\Big(\widetilde{\xi}, \xi, \tau, \widetilde{x}, x, t\Big) \le m^+ r^{\frac{m-2}{m+\widetilde{Q}}}$ for every $(\widetilde{x}, x, t) \in \widetilde{\Omega}^{(m)}_{\vartheta r}\Big(\widetilde{\xi}, \xi, \tau\Big)$;
3. $\widetilde{\Omega}^{(m)}_{\vartheta r}\Big(\widetilde{\xi}, \xi, \tau\Big) \subset \widetilde{\Omega}^{(m)}_{4r}(\widetilde{x}_0, x_0, t_0) \cap \Big\{\tau \le t_0 - \varepsilon\, r^2\Big\}$ for every $\Big(\widetilde{\xi}, \xi, \tau\Big) \in K^{(m)}_{r,\varepsilon}(\widetilde{x}_0, x_0, t_0)$;
4. $\widetilde{M}^{(m)}_{5r}\Big(\widetilde{x}_0, x_0, t_0, \widetilde{x}, x, t\Big) \ge m^- r^m$ for every $(\widetilde{x}, x, t) \in \widetilde{\Omega}^{(m)}_{4r}(\widetilde{x}_0, x_0, t_0) \cap \Big\{t \le t_0 - \varepsilon\, r^2\Big\}$.

Once these properties have been established, the proof of (4.6.31) is straightforward. Indeed, if $\mathrm{div}_{\mathbb{G}} b - c = 0$, for every $\left(\widetilde{\xi}, \xi, \tau\right) \in \widetilde{K}_r^{(m)}(\widetilde{x}_0, x_0, t_0)$ we get

$$\begin{aligned}
\widetilde{u}\left(\widetilde{\xi}, \widetilde{\xi}, \tau\right) &\overset{(4.6.9)}{=} \frac{1}{\vartheta r} \int_{\widetilde{\Omega}_{\vartheta r}^{(m)}(\widetilde{\xi},\xi,\tau)} \widetilde{M}_{\vartheta r}^{(m)}\left(\widetilde{\xi}, \xi, \tau, \widetilde{x}, x, t\right) \widetilde{u}(\widetilde{x}, x, t)\, d\widetilde{x}dxdt \\
&\overset{(2.)}{\leq} \frac{m^+ r^{\frac{m-2}{m+\widetilde{Q}}}}{\vartheta r} \int_{\widetilde{\Omega}_{\vartheta r}^{(m)}(\widetilde{\xi},\xi,\tau)} \widetilde{u}(\widetilde{x}, x, t)\, d\widetilde{x}dxdt \\
&\overset{(3.)}{\leq} \frac{m^+ r^{\frac{m-2}{m+\widetilde{Q}}}}{\vartheta r} \int_{\widetilde{\Omega}_{4r}^{(m)}(\widetilde{x}_0,x_0,t_0)\cap\{\tau\leq t_0-\varepsilon r^2\}} \widetilde{u}(\widetilde{x}, x, t)\, d\widetilde{x}dxdt \\
&\overset{(4.)}{\leq} \frac{5m^+ r^{\frac{m-2}{m+\widetilde{Q}}}}{\vartheta m^- r^m} \frac{1}{5r} \int_{\widetilde{\Omega}_{5r}^{(m)}(\widetilde{x}_0,x_0,t_0)} \widetilde{M}_{5r}^{(m)}\left(\widetilde{x}_0, x_0, t_0, \widetilde{x}, x, t\right) \widetilde{u}(\widetilde{x}, x, t)\, d\widetilde{x}dxdt \\
&\overset{(4.6.9)}{\leq} \frac{5m^+ r_0^{\frac{m-2}{m+\widetilde{Q}}-m}}{\vartheta m^-} \widetilde{u}(\widetilde{x}_0, x_0, t_0),
\end{aligned} \tag{4.6.32}$$

and (4.6.31) follows with $C_P = \frac{5m^+ r_0^{\frac{m-2}{m+\widetilde{Q}}-m}}{\vartheta m^-}$.

We now remove the assumption $\mathrm{div}_{\mathbb{G}} b - c = 0$. From **(H4)** we infer that there exists $k > 0$ such that $|\mathrm{div}_{\mathbb{G}} b - c| \leq k := (m_1 + 1)M_1$. We introduce the auxiliary functions

$$\overline{u}(\widetilde{x}, x, t) := \mathrm{e}^{k(t-t_0)} \widetilde{u}(\widetilde{x}, x, t), \qquad \underline{u}(\widetilde{x}, x, t) := \mathrm{e}^{-k(t-t_0)} \widetilde{u}(\widetilde{x}, x, t).$$

Since $\mathscr{H}_{\widetilde{\mathbb{G}}} \widetilde{u} = 0$, we have that

$$\overline{\mathscr{H}_{\widetilde{\mathbb{G}}}} \overline{u} := \mathscr{H}_{\widetilde{\mathbb{G}}} \overline{u} + k\overline{u} = 0, \qquad \underline{\mathscr{H}_{\widetilde{\mathbb{G}}}} \underline{u} := \mathscr{H}_{\widetilde{\mathbb{G}}} \underline{u} - k\underline{u} = 0.$$

We observe that $\overline{\mathscr{H}_{\widetilde{\mathbb{G}}}}$ and $\underline{\mathscr{H}_{\widetilde{\mathbb{G}}}}$ are operators of the form (4.6.8) with zero-order term $\overline{c} := c + k$ and $\underline{c} := c - k$, respectively. It is clear that *1.-4.* hold with the same constants $r_0, \varepsilon_0, \vartheta, m^-, m^+$. Moreover, by Gaussian estimates we infer that there exists $T, r^* > 0$ such that $0 < \tau - t_0 < T$ for every $\left(\xi, \widetilde{\xi}, \tau\right) \in \widetilde{\Omega}_{5r}^{(m)}(x_0, \widetilde{x}_0, t_0)$, provided $r < r^*$. Assuming without loss of generality that $r_0 \leq r^*$, we infer that there exists $\overline{c}, \underline{c} > 0$ such that

$$\overline{c}\widetilde{u}\left(\widetilde{\xi}, \xi, \tau\right) \leq \overline{u}\left(\widetilde{\xi}, \xi, \tau\right) \leq \widetilde{u}\left(\widetilde{\xi}, \xi, \tau\right), \qquad \widetilde{u}\left(\widetilde{\xi}, \xi, \tau\right) \leq \underline{u}\left(\widetilde{\xi}, \xi, \tau\right) \leq \underline{c}\widetilde{u}\left(\widetilde{\xi}, \xi, \tau\right), \tag{4.6.33}$$

for every $\left(\xi, \widetilde{\xi}, \tau\right) \in \widetilde{\Omega}_{5r}^{(m)}(x_0, \widetilde{x}_0, t_0)$. Let $\overline{M}_r^{(m)}$ and $\underline{M}_r^{(m)}$ denote the mean value formula kernels relevant to $\overline{\mathscr{H}_{\widetilde{\mathbb{G}}}}$ and $\underline{\mathscr{H}_{\widetilde{\mathbb{G}}}}$, respectively, and by $\underline{\Omega}_r^{(m)}$ and $\overline{\Omega}_r^{(m)}$ the corresponding super-level sets. Since $\mathrm{div}_{\mathbb{G}} b - \overline{c} = \mathrm{div}_{\mathbb{G}} b - c - k \leq 0$, the first two lines in (4.6.32) become

$$\begin{aligned}\widetilde{u}\Big(\widetilde{\xi},\xi,\tau\Big) &\overset{(4.6.33)}{\leq} \frac{1}{\overline{c}}\overline{u}\Big(\widetilde{\xi},\xi,\tau\Big)\\ &\overset{(4.6.9)}{\leq} \frac{1}{\vartheta r\overline{c}}\int_{\overline{\Omega}^{(m)}_{\vartheta r}(\widetilde{\xi},\xi,\tau)} \overline{M}^{(m)}_{\vartheta r}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big)\overline{u}(\widetilde{x},x,t)\,d\widetilde{x}dxdt\\ &\overset{(2.)}{\leq} \frac{m^+ r^{\frac{m-2}{m+\widetilde{Q}}}}{\vartheta r\overline{c}}\int_{\overline{\Omega}^{(m)}_{\vartheta r}(\widetilde{\xi},\xi,\tau)} \overline{u}(\widetilde{x},x,t)\,d\widetilde{x}dxdt =: I.\end{aligned}$$

Since *3.* holds also in the form $\overline{\Omega}^{(m)}_{\vartheta r}(\widetilde{\xi},\xi,\tau) \subset \underline{\Omega}^{(m)}_{4r}(\widetilde{x}_0,x_0,t_0) \cap \Big\{\tau \leq t_0 - \varepsilon\, r^2\Big\}$, by using $\mathrm{div}_{\mathbb{G}} b - \underline{c} = \mathrm{div}_{\mathbb{G}} b - c + k \geq 0$ we conclude

$$\begin{aligned}I \leq \frac{5m^+ r^{\frac{m-2}{m+\widetilde{Q}}}}{\vartheta m^- r^m}\frac{\underline{c}}{\overline{c}}\frac{1}{5r}\int_{\underline{\Omega}^{(m)}_{5r}(\widetilde{x}_0,x_0,t_0)} \underline{M}^{(m)}_{5r}\Big(\widetilde{x}_0,x_0,t_0,\widetilde{x},x,t\Big)\underline{u}(\widetilde{x},x,t)\,d\widetilde{x}dxdt&\\ \leq \frac{5m^+}{\vartheta m^-}\frac{\underline{c}}{\overline{c}}\, r_0^{\frac{m-2}{m+\widetilde{Q}}-m}\,\widetilde{u}(\widetilde{x}_0,x_0,t_0),&\end{aligned}$$

where the last inequality follows again from (4.6.9) and (4.6.33). Hence, inequality (4.6.31) follows with $C_P = \frac{5m^+}{\vartheta m^-}\frac{\underline{c}}{\overline{c}}\, r_0^{\frac{m-2}{m+\widetilde{Q}}-m}$.

To conclude the proof of Proposition 4.6.22, we are now left with proving claims *1.-4.*

1. We prove that the set

$$\widetilde{K}^{(m)}_{r,\varepsilon}(\widetilde{x}_0,x_0,t_0) = \overline{\left\{\frac{\widetilde{\Gamma}(\widetilde{x}_0,x_0,t_0,\widetilde{x},x,t)}{(4\pi(t_0-t))^{m/2}} > \frac{1}{r}\right\}} \cap \Big\{t \leq t_0 - \varepsilon\, r^2\Big\},$$

is non-empty for every $r,\varepsilon \leq 1$. From the Gaussian lower bound (4.3.12) we infer that there exists $c, c_l > 0$ (depending on structural assumptions) such that $\widetilde{K}^{(m)}_{r,\varepsilon}(\widetilde{x}_0,x_0,t_0) \subset \mathcal{K}$, where

$$\mathcal{K} := \overline{\left\{\frac{c}{\Big(t_0-t\Big)^{\frac{\widetilde{Q}}{2}}}\ \exp\left(-c_l\frac{d_{\widetilde{\mathbb{G}}}(\widetilde{x}_0,x_0,\widetilde{x},x)^2}{(t_0-t)}\right)\frac{1}{\Big(4\pi(t_0-t)\Big)^{m/2}} > \frac{1}{r}\right\}} \cap \Big\{t \leq t_0 - \varepsilon\, r^2\Big\}.$$

We now show that $\mathcal{K}$ is non-empty. Let $t := t_0 - \varepsilon\, r^2$, and notice that any $(\widetilde{x},x) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N$ satisfying

$$d_{\widetilde{\mathbb{G}}}(\widetilde{x}_0,x_0,\widetilde{x},x) \leq \frac{1}{\varepsilon\, r^2 c_\lambda}\ln\left(\frac{(4\pi)^{\frac{m}{2}}}{c}\Big(\varepsilon^{\,2} r\Big)^{m+\widetilde{Q}-1}\right),$$

belongs to $\mathcal{K}$. Therefore $\mathcal{K}$ is non-empty, and so is $\widetilde{K}^{(m)}_{r,\varepsilon}(\widetilde{x}_0,x_0,t_0)$.

2. We show that there exist $r_0, m^+ > 0$ such that, for every $r \leq r_0$, it holds

$$\widetilde{M}^{(m)}_r\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big) \leq m^+ r^{\frac{m-2}{m+\widetilde{Q}}}, \qquad \text{for every } (\widetilde{x},x,t) \in \widetilde{\Omega}^{(m)}_r\Big(\widetilde{\xi},\xi,\tau\Big).$$

Recall that the explicit formula for the kernel $\widetilde{M}_r^{(m)}$ is given by

$$\begin{aligned}\widetilde{M}_r^{(m)}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big) &= \frac{2^m m\omega_m}{m+2}(\tau-t)^{\frac{m-2}{2}}\left(\ln\left(r\,\frac{\widetilde{\Gamma}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big)}{\Big(4\pi(\tau-t)\Big)^{m/2}}\right)\right)^{\frac{m+2}{2}}\\ &\quad + 2^m\omega_m\left((\tau-t)\ln\left(r\,\frac{\widetilde{\Gamma}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big)}{\Big(4\pi(\tau-t)\Big)^{m/2}}\right)\right)^{\frac{m}{2}}\cdot\\ &\qquad\cdot\frac{\langle A(x,t)\nabla_{\widetilde{\mathbb{G}}}\widetilde{\Gamma}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big),\nabla_{\widetilde{\mathbb{G}}}\widetilde{\Gamma}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big)\rangle}{\widetilde{\Gamma}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big)^2}\\ &=: K_1+K_2.\end{aligned} \tag{4.6.34}$$

We first focus on the term K_2. To this end, we show that there exist constants $r_0>0$ and $c>0$ such that, for every $r\le r_0$, the following inequality holds

$$\frac{\langle A(x,t)\nabla_{\widetilde{\mathbb{G}}}\widetilde{\Gamma}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big),\nabla_{\widetilde{\mathbb{G}}}\widetilde{\Gamma}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big)\rangle}{\widetilde{\Gamma}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big)^2}\le\frac{c}{(\tau-t)},\quad (\widetilde{x},x,t)\in\widetilde{\Omega}_r^{(m)}\Big(\widetilde{\xi},\xi,\tau\Big). \tag{4.6.35}$$

To prove (4.6.35) we apply Proposition 4.6.21. We claim that this proposition remains valid when replacing $\Big\{\widetilde{\Gamma}\ge K\Big\}$ with the level set $\widetilde{\Omega}_{1/K}^{(m)}$. Indeed, Proposition 4.6.21 applied to $\widetilde{\Gamma}^{(m)}$ on the set $\Big\{\widetilde{\Gamma}^{(m)}(0,\widetilde{\xi},\xi,\tau,\widetilde{x},x,y,t\Big)\ge K\Big\}\supseteq\widetilde{\Omega}_{1/K}^{(m)}\Big(\widetilde{\xi},\xi,\tau\Big)$ implies that

$$\begin{aligned}H^{(m)}(0,\tau,0,t)\widetilde{X}_i\widetilde{\Gamma}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big) &= X_i\widetilde{\Gamma}^{(m)}(0,\widetilde{\xi},\xi,\tau,0,\widetilde{x},x,y,t)\\ &\le c\left(\frac{1}{\sqrt{t-\tau}}d_{\widetilde{\mathbb{G}}}\Big(\widetilde{\xi},\xi,\widetilde{x},x\Big)+1\right)\widetilde{\Gamma}^{(m)}\Big(0,\widetilde{\xi},\xi,\tau,0,\widetilde{x},x,t\Big)\\ &= c\left(\frac{1}{\sqrt{t-\tau}}d_{\widetilde{\mathbb{G}}}\Big(\widetilde{\xi},\xi,\widetilde{x},x\Big)+1\right)H^{(m)}(0,\tau,0,t)\widetilde{\Gamma}\Big(\widetilde{\xi},\xi,\tau,\widetilde{x},x,t\Big),\end{aligned}$$

proving the claim. Let K be the constant given by the aforementioned proposition, and set $r_0:=1/K$. By the ellipticity assumption given by **(H4)** and Proposition 4.6.21, we finally conclude the validity of estimate (4.6.35).

We focus now on K_1. By Proposition 4.6.20 and (4.6.21), we deduce that, if we pick r_0 small enough, we can choose $T>0$ such that

$$\widetilde{\Omega}_r^{(m)}(\widetilde{\xi},\xi,\tau)\subseteq\Big\{(\widetilde{x},x,t)\in\mathbb{R}^{\widetilde{N}}\times\mathbb{R}^N\times\mathbb{R}:0\le\tau-t\le T\Big\},\qquad\text{for every } r\le r_0.$$

Then, by (4.3.2) we infer that there exists $c = c(r_0) > 0$ such that

$$K_1 \leq \frac{2^m m\omega_m}{m+2}(\tau - t)^{\frac{m}{2}} \left(\ln \left(\frac{c\, r}{(\tau - t)^{\frac{m+\widetilde{Q}}{2}}} \right) \right)^{\frac{m+2}{2}} . \tag{4.6.36}$$

Combining (4.6.34), (4.6.35), and (4.6.36), the validity of *2.* follows.

3. The key ingredient to prove this inclusion is the parabolic dilation relevant to $\widetilde{\mathbb{G}}$. Indeed, by Proposition 4.6.20 we infer that there exist $r_0, \varepsilon_0 > 0$ such that for every $r \leq r_0$ and $\varepsilon \leq \varepsilon_0$ it holds

$$\widetilde{K}^{(m)}_{r,\varepsilon}(\tilde{x}_0, x_0, t_0) \subset \overline{\widetilde{\Omega}^{(m)^*}_{3r}(\tilde{x}_0, x_0, t_0)} \cap \left\{ t \leq t_0 - \varepsilon\, r^2 \right\}, \tag{4.6.37}$$

where

$$\widetilde{\Omega}^{(m)^*}_{r}(\tilde{x}_0, x_0, t_0) := \left\{ (\tilde{x}, x, t) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times \mathbb{R} : \frac{\widetilde{Z}(\tilde{x}_0, x_0, t_0, \tilde{x}, x, t)}{\Big(4\pi(t_0 - t)\Big)^{m/2}} > \frac{2}{r} \right\},$$

with $\widetilde{Z}$ being the parametrix of $\widetilde{\Gamma}$. We claim that there exists $\vartheta \in (0, 1)$ such that

$$\widetilde{\Omega}^{(m)^*}_{3\vartheta r}(\tilde{x}, x, t) \subset \widetilde{\Omega}^{(m)^*}_{4r}(\tilde{x}_0, x_0, t_0) \cap \left\{ \tau \leq t_0 - \varepsilon\, r^2 \right\}, \tag{4.6.38}$$

for every $(\tilde{x}, x, t) \in \overline{\widetilde{\Omega}^{(m)^*}_{3r}(\tilde{x}_0, x_0, t_0)} \cap \{t \leq t_0 - \varepsilon\, r^2\}$. The validity of *3.* follows then by Proposition 4.6.20, that implies $\widetilde{\Omega}^{(m)^*}_{4r}(\tilde{x}_0, x_0, t_0) \subset \widetilde{\Omega}^{(m)}_{4r}(\tilde{x}_0, x_0, t_0)$. To prove claim (4.6.38), we employ the parabolic dilation relevant to $\widetilde{\mathbb{G}} = \left(\mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N, \tilde{\circ}, \tilde{\delta}_r \right)$: by the homogeneity of the parametrix we infer that for every positive k it holds

$$\begin{gathered} \Big((\tilde{x}_0, x_0) \mathbin{\tilde{\circ}} \tilde{\delta}_r(\tilde{x}, x), t_0 + r^2 t \Big) \in \widetilde{\Omega}^{(m)^*}_{kr}(\tilde{x}_0, x_0, t_0) \\ \Updownarrow \\ \Big((\tilde{x}_0, x_0) \mathbin{\tilde{\circ}} (\tilde{x}, \tilde{x}), t_0 + t \Big) \in \widetilde{\Omega}^{(m)^*}_{k}(\tilde{x}_0, x_0, t_0) \end{gathered} \tag{4.6.39}$$

Hence, it is sufficient to show (4.6.38) whenever $r = 1$. Let $d(\tilde{x}_0, x_0, t_0)$ denote the distance between the set $\overline{\widetilde{\Omega}^{(m)^*}_{3}(\tilde{x}_0, x_0, t_0)} \cap \{t \leq t_0 - \varepsilon\}$ and the boundary of $\widetilde{\Omega}^{(m)^*}_{4}(\tilde{x}_0, x_0, t_0)$. Clearly, this defines a positive function continuously depending on $(\tilde{x}_0, x_0, t_0)$. In particular, it depends on the automorphism $T_{A(\tilde{x}_0, x_0, t_0)}$, associated with the matrix A evaluated at the point $(\tilde{x}_0, x_0, t_0)$ (see (4.6.7)). By (4.6.14), we infer that there exists $d_0, \varepsilon_0 > 0$ (depending only on structural assumption) such that $d(\tilde{x}_0, x_0, t_0) \geq d_0$ for every $(\tilde{x}_0, x_0, t_0) \in \widetilde{\Omega}$ and $\varepsilon \leq \varepsilon_0$. Since the diameter of $\widetilde{\Omega}^{(m)^*}_{1}(\tilde{x}, x, t)$ is uniformly bounded, by (4.6.39) we deduce that it is possible to find $\vartheta \in (0, 1)$ such that the diameter of $\widetilde{\Omega}^{(m)^*}_{\vartheta}(\tilde{x}, x, t)$ is not greater than

d_0, and then we conclude the proof of the claim.

4. We recall again that the explicit expression of the kernel is given by (4.6.34). By the ellipticity assumption **(H4)** we infer $K_2 \geq 0$. To prove the estimate is therefore sufficient to show that $K_1 \geq m^- r^m$. This follows once we set

$$m^- := \frac{2^m m\omega_m}{4(m+2)} \left(\ln(5/4)\right)^{\frac{m+2}{2}} .$$

□

Building on the previous proposition, we are eventually able to prove Proposition 4.6.15.

Proof of Proposition 4.6.15. Let $u \geq 0$ be a classical solution to $\mathscr{H}_{\mathbb{G}} u = 0$ in $\Omega \subset \mathbb{R}^N \times \mathbb{R}$, and for every $(x,t) \in \Omega$ and $\widetilde{x} \in \mathbb{R}^{\widetilde{N}}$ set $\widetilde{u}(\widetilde{x},x,t) := u(x,t)$. We notice that, by (4.2.9), $\widetilde{u}$ solves $\widetilde{\mathscr{H}_{\mathbb{G}}}\widetilde{u} = 0$ in $\widetilde{\Omega} := \mathbb{R}^{\widetilde{N}} \times \Omega$. Choose $m = 4$, and let r_0, ε_0 and C_P be the corresponding constant given by Proposition 4.6.22. We claim that there exists $0 < r_1 \leq r_0 < 1$, $0 < \varepsilon_1 \leq \varepsilon_0 < 1$, $0 < \vartheta_1 < 1$, and $C > 0$ such that

$$\{0\} \times D_r(z_0) = \{0\} \times B_{\vartheta_1 r}(x_0) \times \left\{t - t_0 = -\varepsilon_1 r^2\right\} \subseteq \widetilde{K}^{(4)}_{r,\vartheta_1 \varepsilon}(0,x_0,t_0), \tag{4.6.40}$$

$$\widetilde{K}^{(4)}_{5r,\vartheta_1 \varepsilon}(0,x_0,t_0) \subseteq \widetilde{Q}_r(0,x_0,t_0) = \widetilde{B}_r(0,x_0) \times \left\{0 \leq t_0 - t \leq r^2\right\}, \tag{4.6.41}$$

$$\widetilde{Q}_r(0,x_0,t_0) \subseteq \mathbb{R}^{\widetilde{N}} \times Q_r(x_0,t_0) \subseteq \mathbb{R}^{\widetilde{N}} \times \Omega = \widetilde{\Omega}, \tag{4.6.42}$$

for every $(x_0,t_0) \in \Omega$, $\varepsilon \leq \varepsilon_1$ and $r \leq r_1$ such that $Q_r(x_0,t_0) \subset \Omega$. Once the claim is proved, we are in position to apply Proposition 4.6.22 to the lifted function $\widetilde{u}$, obtaining

$$\sup_{D_r(x_0,t_0)} u = \sup_{\{0\}\times D_r(x_0,t_0)} \widetilde{u} \leq \sup_{\widetilde{K}^{(4)}_{r,\vartheta_1\varepsilon}(0,x_0,t_0)} \widetilde{u} \leq C_P\, \widetilde{u}(0,x_0,t_0) = C_P\, u(x_0,t_0),$$

for every $(x_0,t_0) \in \Omega$. This proves the desired parabolic Harnack inequality.

We prove the inclusions in the claim by employing Gaussian estimates and the boundedness of the set $\widetilde{\Omega}^{(m)}$. By Theorem 4.3.11 we infer that there exists two positive constants c_λ and $\widetilde{c}$ such that

$$\widetilde{\Gamma}(0,x_0,t_0,\widetilde{x},x,t) \geq \frac{\widetilde{c}}{\left(t_0 - t\right)^{\frac{\widetilde{Q}}{2}}} \exp\left(-c_l \frac{d_{\widetilde{\mathbb{G}}}(0,x_0,\widetilde{x},x)^2}{(t_0 - t)}\right),$$

for every $(0,x_0,t_0), (\widetilde{x},x,t) \in \mathbb{R}^{\widetilde{N}} \times \mathbb{R}^N \times \mathbb{R}$ such that $0 \leq t_0 - t \leq 1$. By the definition

(4.2.10) of the distance $d_{\widetilde{\mathbb{G}}}$ we deduce

$$\widetilde{\Gamma}(0,x_0,t_0,\widetilde{x},x,t) \geq \frac{\widetilde{c}}{\Big(t_0-t\Big)^{\frac{\widetilde{Q}}{2}}} \exp\left(-c_\lambda \frac{d_X(x_0,x)^2}{(t_0-t)}\right) \exp\left(-c_\lambda \frac{\sum_{i=1}^{\kappa}\sum_{j=1}^{\widetilde{m}_i} |\widetilde{x}_j^{(i)}|^{\frac{2}{i}}}{(t_0-t)}\right) =: I.$$

If $(\widetilde{x},x,t) \in \{0\} \times D_r(z_0)$, then $t_0 - t = \varepsilon_1 r^2 \leq 1$, $d_X(x_0,x)^2 \leq \vartheta_1 r$ and $\widetilde{x} = 0$, this implying

$$I \geq \frac{\widetilde{c}}{\varepsilon_1^{\widetilde{Q}/2} r^{\widetilde{Q}}} \exp\Big(- c_\lambda \vartheta_1^2/\varepsilon_1\Big),$$

concluding that for every fixed $\vartheta_1 \in (0,1)$ (4.6.40) holds for a suitable choice of ε_1 and for every $r \leq r_*$, for some $r_* \in (0,1)$.

The proof of (4.6.41) follows by a similar argument: combining the Gaussian upper bound (4.3.2) with (4.6.30), (4.6.37), and the scaling property (4.6.39), we infer that there exist $r^*, \vartheta_1 \in (0,1)$ such that (4.6.41) holds for every $r \leq r^*$. Inclusion (4.6.42) is a direct consequence of the definition of the lifted distance (4.2.10), which completes the proof of the claim by picking $r_1 \leq \min\Big(r_0, r^*, r_*\Big)$. □

4.7 Proof of Theorem 4.1.6

We can first observe that it is not restrictive to assume $z_0 = 0$ and $r = 1$. We let r_1, ϑ_1, ε_1 and C_P be the constants given by Proposition 4.6.15, and we set

$$\vartheta_0 := \min\left\{\vartheta_1, \frac{1}{12k_1^3}\right\}, \tag{4.7.1}$$

where k_1 is the constant given by the pseudo-triangular inequality (4.2.6). Moreover, we choose four positive constants $\nu, \eta, \mu, \vartheta$ with $0 < \nu < \eta < \mu < 1$ and $0 < \vartheta \leq \vartheta_0$, and consider the cylinders Q^+ and Q^- introduced in (4.1.5). Finally, we set

$$r_0 := \min\left\{r_1, \frac{1}{2k_1}, \sqrt{1-\gamma}\right\}. \tag{4.7.2}$$

We next consider $z^- := \Big(x^-, t^-\Big) \in Q^- = B_\vartheta(0) \times (-\nu, 0)$, and $z^+ := \Big(x^+, t^+\Big) \in Q^+ = B_\vartheta(0) \times (-\mu, -\eta)$, and exploit Proposition 4.6.15 to construct a Harnack chain, *i.e* a finite set of points $\{z_0, z_1, \dots, z_m\} \in Q_1(0)$ such that

$$z_0 = z^+, \qquad z_m = z^-, \qquad Q_r(z_j) \subset Q_1(0) \subset \Omega, \qquad z_j \in D_r(z_{j-1}), \tag{4.7.3}$$

for every $j = 1, \dots, m$ and $r \leq r_0$. If (4.7.3) holds, by iterating Proposition 4.6.15 we conclude that

$$u\Big(z^-\Big) \leq C_P^m u\Big(z^+\Big).$$

Once we prove that m is uniformly bounded w.r.t. $z^+ \in Q^+$ and $z^- \in Q^-$, we conclude the proof of the theorem.

In order to prove (4.7.3), we need to make a suitable choice of z_j's. To this end, we consider a X-trajectory $\underline{\gamma} : [0,1] \to \Omega$ in $\mathcal{C}$ steering x^+ to x^-, *i.e.* $\underline{\gamma}$ is a solution to the following system

$$\begin{cases} \dot{\underline{\gamma}}(\tau) = \sum_{i=1}^{m_1} \underline{\alpha}_i(\tau) X_i\Big(\underline{\gamma}(\tau)\Big), \\ \underline{\gamma}(0) = x^+, \, \underline{\gamma}(1) = x^-, \end{cases}$$

where $\underline{\alpha} = (\underline{\alpha}_1, ., \underline{\alpha}_{m_1})$ is the control function satisfying condition (4.2.3). We further require $\underline{\gamma}$ to be optimal in the d_X^∞ metric, meaning that

$$d_X^\infty\Big(x^+, x^-\Big) = \sup_{\tau \in [0,1]} |\underline{\alpha}(\tau)|. \tag{4.7.4}$$

Since $\underline{\gamma}$ connects only the space variables x^+ and x^-, we need to suitably incorporate the time variables t^+ and t^- into the construction. To this end, we reparametrize this curve by stretching it over the time interval $[0, t^+ - t^-]$, *i.e.* we define the X-curve $\widetilde{\gamma} : \Big[0, t^+ - t^-\Big] \to \Omega$ as

$$\widetilde{\gamma}(\tau) := \underline{\gamma}\left(\frac{\tau}{t^+ - t-}\right), \qquad \tau \in \Big[0, t^+ - t^-\Big],$$

whose control function $\widetilde{\alpha}$ is given by

$$\widetilde{\alpha}(\tau) = \frac{1}{t^+ - t^-} \underline{\alpha}\left(\frac{\tau}{t^+ - t^-}\right), \qquad \tau \in [0, t^+ - t^-]. \tag{4.7.5}$$

The following observation will be crucial in the sequel. For every $s \in \Big[0, t^+ - t^-\Big]$, let $\widehat{\gamma} : [0,1] \to \Omega$ denote the X-trajectory $\widehat{\gamma}(\tau) = \widetilde{\gamma}(\tau s)$. We denote its control function as $\widehat{\alpha}(\tau) = s\, \widetilde{\alpha}(\tau s)$, and we notice that $\widehat{\gamma}(0) = x^+$ and $\widehat{\gamma}(1) = \widetilde{\gamma}(s)$. Then, the following inequality holds true

$$d_X\Big(\widetilde{\gamma}(s), x^+\Big) \leq \int_0^1 |\widehat{\alpha}(\tau)| d\tau = \int_0^s |\widetilde{\alpha}(\tau)| d\tau. \tag{4.7.6}$$

We employ $\widetilde{\gamma}$ to choose the z_j's: for an arbitrary $m \in \mathbb{N}$, we choose a positive r such that

$$\varepsilon_1 r^2 = \frac{t^+ - t^-}{m}, \tag{4.7.7}$$

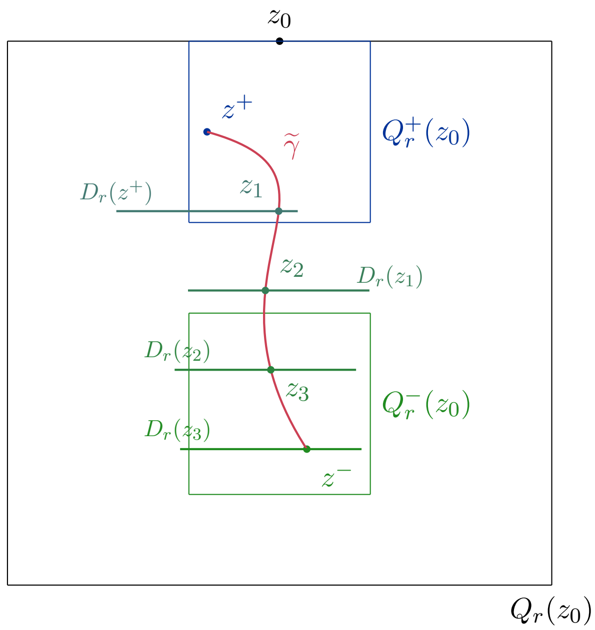

Figure 4.3: Harnack chain (the time variable is represented vertically, and upwardly increasing). The author is very proud of this figure.

and for every $j = 0, \ldots, m$ we define

$$z_j = \Big(x_j, t_j\Big) := \Big(\widetilde{\gamma}\Big(j\varepsilon_1 r^2\Big), t^+ - j\varepsilon_1 r^2\Big),$$

see Figure 4.3. We first observe that from the previous definition it follows

$$z_j = \Big(\widetilde{\gamma}_{x_{j-1}}\Big(\varepsilon_1 r^2\Big), t_{j-1} - \varepsilon_1 r^2\Big), \tag{4.7.8}$$

where $\widetilde{\gamma}_{x_{j-1}}$ is the section of the curve $\widetilde{\gamma}$ steering x_{j-1} to x_j, *i.e.* $\widetilde{\gamma}_{x_{j-1}} : \Big[0, \varepsilon_1 r^2\Big] \to \Omega$ with $\widetilde{\gamma}_{x_{j-1}}(0) = x_{j-1}$, $\widetilde{\gamma}_{x_{j-1}}\Big(\varepsilon_1 r^2\Big) = x_j$ for every $j = 1, \ldots, m$.
It is clear that $z_0 = z^+$, $z_m = z^-$. Hence, to prove (4.7.3) we begin by showing that for every $s \in \Big[0, t^+ - t^-\Big]$, $\vartheta \leq \vartheta_0$ and $r \leq r_0$ it holds $B_r\Big(\widetilde{\gamma}(s)\Big) \subset B_1(0)$. This implies $Q_r(z_j) \subset Q_1(0)$ if $r \leq r_0$, in virtue of the choice of γ and t_j's. Let $y \in B_r\Big(\widetilde{\gamma}(s)\Big)$, then for every $r \leq r_0$ it holds

$$\begin{aligned}
d_X(y, 0) &\overset{(4.2.6)}{\leq} k_1 d_X\Big(y, \widetilde{\gamma}(s)\Big) + k_1^2 d_X\Big(x^+, 0\Big) + k_1^2 d_X\Big(\widetilde{\gamma}(s), x^+\Big) \\
&\overset{(4.7.6)}{\leq} k_1 r + k_1^2 \vartheta + k_1^2 \int_0^s |\widetilde{\alpha}(\tau)| d\tau \\
&\leq k_1 r + k_1^2 \vartheta + k_1^2 \Big(t^+ - t^-\Big) \sup_{\tau \in [0, t^+ - t^-]} |\widetilde{\alpha}(\tau)|,
\end{aligned}$$

where in the last inequality we simply used the fact that $s \in \left[0, t^+ - t^-\right]$. We then compute

$$\begin{aligned}
d_X(y,0) &\overset{(4.7.5)}{\leq} k_1 r + k_1^2 \vartheta + k_1^2 \sup_{\tau\in[0,1]} |\underline{\alpha}(\tau)| \\
&\overset{(4.7.4)}{=} k_1 r + k_1^2 \vartheta + k_1^2 d_X^{(\infty)}(x^+, x^-) \\
&\overset{(4.2.5)}{=} k_1 r + k_1^2 \vartheta + k_1^2 d_X(x^+, x^-) \\
&\overset{(4.7.2)}{\leq} \frac{1}{2} + 3k_1^3 \vartheta \overset{(4.7.1)}{<} 1,
\end{aligned}$$

and therefore the trajectory $\widetilde{\gamma}$ we constructed does not exit $B_1(0)$. We conclude by showing the last inclusion in (4.7.3): in virtue of (4.6.2) and (4.7.8), $z_j \in D_r(z_{j-1})$ if $d_X(x_j, x_{j-1}) \leq \vartheta_1 r$. Arguing as in (4.7.6) we get

$$\begin{aligned}
d_X(x_j, x_{j-1}) &\leq \int_{(j-1)\varepsilon_1 r^2}^{j\varepsilon_1 r^2} |\widetilde{\alpha}(\tau)| d\tau \\
&\overset{(4.7.7)}{\leq} \frac{t^+ - t^-}{m} \sup_{\tau\in[0,t^+-t^-]} |\widetilde{\alpha}(\tau)| \\
&\overset{(4.7.5)}{=} \frac{1}{m} \sup_{\tau\in[0,1]} |\underline{\alpha}(\tau)| \\
&\overset{(4.7.4)}{=} \frac{1}{m} d_X^{(\infty)}(x^+, x^-) \\
&\overset{(4.2.5)}{=} \frac{1}{m} d_X(x^+, x^-).
\end{aligned}$$

Hence, $z_j \in D_r(z_{j-1})$ if we choose $m \in \mathbb{N}$ such that

$$\frac{1}{m} d_X(x^+, x^-) \leq \vartheta_1 r \overset{(4.7.7)}{=} \frac{\vartheta_1 \sqrt{t^+ - t^-}}{\sqrt{m}\sqrt{\varepsilon_1}} \iff m \geq \frac{d_X(x^+, x^-)^2 \varepsilon_1}{\vartheta_1^2 (t^+ - t^-)}. \tag{4.7.9}$$

The proof of the theorem follows once we obtain a bound for m that is uniform w.r.t. $z^+ \in Q^+$ and $z^- \in Q^-$. In order for the previous construction to hold, we need to satisfy (4.7.9) and $r \leq r_0$. Since $t^+ - t^- \geq \eta - \nu$, we need to require $m \geq \frac{4k_1^2 \varepsilon_1}{(\eta-\nu)}$ so that (4.7.9) holds. Moreover, by (4.7.7) condition $r \leq r_0$ holds if $m \geq \frac{t^+ - t^-}{\varepsilon_1 r_0^2}$, and therefore it is sufficient to require that $m \geq \frac{\mu}{\varepsilon_1 r_0^2}$ since $t^+ - t^- \leq \mu$. To summarize the above, we have proved that inequality (4.1.6) holds with $C_H = C_P^{\bar{m}}$, provided that we choose $\bar{m} \in \mathbb{N}$ such that $\bar{m} \geq \max\left\{\frac{4^2 k_1 \varepsilon_1}{(\eta-\nu)}, \frac{\mu}{\varepsilon_1 r_0^2}\right\}$.

A Comparison principle

In this appendix we provide the proof of the following comparison principle.

Proposition A.1. *Let $\Omega_T = \Omega \times (T_1, T)$ be an open bounded subset of $\mathbb{R}^{N+1}$, and let $w \in C^2_{x,t}(\Omega_T)$ be a solution to*

$$\begin{cases} \mathscr{H}_{\mathbb{G}} w \geq 0, & \text{in } \Omega_T, \\ \limsup w \leq 0, & \text{in } \partial\Omega \times (T_1, T) \cup \Omega \times \{T_1\}, \end{cases}$$

then $w \leq 0$ in Ω_T. The same result holds in $\Omega = \mathbb{R}^N$ if $\limsup w \leq 0$ in $\mathbb{R}^N \times \{T_1\}$ and at infinity.

Proof. The scheme of the proof is classical, once we follow the $C^2_{x,t}$ regularity of w. For every $k \in \mathbb{N}$ we set $\Omega_{T_k} := \Omega_T \cap \left\{t \leq \left(T - \frac{1}{k}\right)\right\}$, and we prove the result by showing that $w \leq 0$ in Ω_{T_k} for every $k \in \mathbb{N}$. The result on $\mathbb{R}^N$ straightforwardly follows.
Let us pick some $M > M_1$ (where M_1 is given by **(H4)**), and set $v(x,t) := w(x,t)\exp(-Mt)$. By a direct computation we see that v solves $\mathscr{H}_{\mathbb{G}} v - Mv \geq 0$ in Ω_{T_k}, and that $\limsup v \leq 0$ in $\partial\Omega \times (T_1, T) \cup \Omega \times \{T_1\}$. Since the function $v(\cdot, t)$ has continuous first and second order Lie derivatives w.r.t. the generating vector fields X_i, then it has continuous Lie derivative w.r.t. every $X \in \text{span}\{X_i\}_{i=1}^{m_1}$(see Lemma 4.2 in [26]). Therefore, if $v \in C^2_{x,t}(\Omega_T)$ has a positive maximum point in $(x_0, t_0) \in \Omega \times \{t = T_k\}$, we claim that

$$X_i v(x_0, t_0) = 0, \qquad \partial_t v(x_0, t_0) \leq 0, \qquad \sum_{i,j=1}^{m_1} a_{ij}(x_0, t_0) X_i X_j v(x_0, t_0) \leq 0, \tag{A.1}$$

whereas when $(x_0, t_0) \in \Omega_{T_k}$ the second inequality becomes an equality. From this claim it follows $\Big(M - c(x_0, t_0)\Big) v(x_0, t_0) \leq 0$, hence $v \leq 0$ in Ω_{T_k}, whereby the thesis follows.
We prove the claim by proving the last inequality in (A.1), since the others are straightforward: by contradiction, assume there exists $d \in \mathbb{R}^{m_1}$ such that $\sum_{i,j=1}^{m_1} d_i d_j X_i X_j v(x_0, t_0) > 0$. Then, let $\gamma_{Z,0}(s)$ be the integral curve of $Z = \sum_{i=1}^{m_1} d_i X_i$ starting from the origin, and for every $h \in \mathbb{R}$ set $g(h) := v\Big(x_0 \circ \gamma_{Z,0}(h), t_0\Big)$. Since $v(\cdot, t_0)$ has continuous second order Lie derivatives, it follows $g \in C^2(-\sigma, \sigma)$ for a sufficiently small $\sigma > 0$. Therefore, recalling that (x_0, t_0) is a maximum point, by a Taylor expansion we get

$$\begin{aligned} v(x_0, t_0) \geq g(h) =& g(0) + \frac{h^2}{2} g''(0) + o_{h \to 0}\Big(h^2\Big) \\ =& v(x_0, t_0) + \frac{h^2}{2} \sum_{i,j=1}^{m_1} d_i d_j X_i X_j v(x_0, t_0) + o_{h \to 0}\Big(h^2\Big) > v(x_0, t_0), \end{aligned}$$

hence the contradiction. $\square$

Part II

Applications to degenerate McKean-Vlasov processes

Chapter 5

Preliminaries: Kolmogorov equations

In this preliminary chapter we present the analytical framework of Kolmogorov operators. This setting will be employed to investigate the Cauchy problems associated with some McKean-Vlasov stochastic processes. For our purpose, we focus on the appropriate functional spaces and the fundamental solution, along with its main properties. These tools will serve as the main instruments to address the following initial value problem

$$\begin{cases} \mathscr{K} f(v,x,t) = \partial^2_{vv} f(v,x,t) \\ \qquad -\partial_v\Big(\Big(b(v,x) + \int_{\mathbb{R}^2} K(\eta - v, \xi - x, t) f(\eta,\xi,t)d\eta d\xi\Big) f(v,x,t)\Big) \\ \qquad\qquad -v\partial_x f(v,x,t) - \partial_t f(v,x,t) = 0, \\ \qquad\qquad\qquad (v,x,t) \in \mathbb{R}^2 \times (0,T), \\ f(v,x,0) = f_0(v,x), \qquad\qquad (v,x) \in \mathbb{R}^2, \end{cases} \tag{5.0.1}$$

where b is a drift term, while K is an interaction kernel. This analytical framework proves to be optimal to establish existence results for (5.0.1), to obtain quantitative estimates, and to develop a stable and consistent numerical scheme to approximate its solution.

We begin by introducing the setting of the classical theory for Kolmogorov operators on $\mathbb{R}^{N+1}$, which will be employed in Chapter 6. For a broader overview of the topic, we refer the interested reader to the survey [9]. Consider first the following class of differential operators

$$\mathcal{K} = \sum_{i,j=1}^{m_1} a_{ij}\partial^2_{x_i x_j} + \langle Bx, \nabla\rangle - \partial_t.$$

Here, $A = (a_{i,j})_{i,j=1}^{m_1}$ denotes a symmetric, positive definite $m_1 \times m_1$ matrix (with $m_1 \leq N$), while B is an $N \times N$ matrix. The operator $\mathcal{K}$ (which has a more structured principal part than the one appearing in (5.0.1)) allows us to understand the Lie group structure of Kolmogorov operators, which is mainly determined by the the drift matrix B. We introduce some further notation: let A_0 denote the $N \times N$ matrix

$$A_0 := \begin{pmatrix} A & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix},$$

and for every $t > 0$ set

$$E(t) := \exp(-tB), \qquad C(t) := \int_0^t E(s) A_0 E^T(s) ds.$$

The matrix $C(t)$ is symmetric and nonnegative for every $t > 0$, but generally may fail to be strictly positive (as this is the case when $m_1 < N$ and $B = \mathbb{O}$). In his famous work [63], Hörmander gave a sufficient condition on the matrices A_0 and B ensuring the hypoellipticity of $\mathcal{K}$. This condition implies $C(t) > 0$ for every $t > 0$, and yields the following expression for its fundamental solution

$$\Gamma(x,t,\xi,\tau) = \Gamma\Big(x - E(t-\tau)\xi, t-\tau\Big), \qquad (x,t),(\xi,\tau) \in \mathbb{R}^{N+1}, \tag{5.0.2}$$

with $\Gamma(x,t) = 0$ whenever $t \leq 0$, and

$$\Gamma(x,t) = \frac{1}{(4\pi)^{\frac{N}{2}} \sqrt{\det C(t)}} \exp\Big(-\frac{1}{4}\langle C^{-1}(t)x, x\rangle - t\,\mathrm{Tr}(B)\Big), \qquad x \in \mathbb{R}, t > 0.$$

As suggested by (5.0.2), $\mathcal{K}$ is left-invariant w.r.t. the following non abelian group law

$$(x,t) \circ (\xi,\tau) := \Big(\xi + E(\tau)x, t+\tau\Big), \qquad (x,t),(\xi,\tau) \in \mathbb{R}^{N+1}. \tag{5.0.3}$$

Starting from this observation, Hörmander's results were later investigated by Lanconelli and Polidoro [76] in the setting of Lie groups. In their analysis, they proved that the operator is hypoelliptic if and only if the following condition holds

(B) *There exist κ natural integers with $1 \leq m_{\kappa+1} \leq \ldots \leq m_2 \leq m_1$ and $m_2 + \ldots + m_{\kappa+1} = N - m_1$, such that the matrix B takes the following block form*

$$B = \begin{pmatrix} \star & \star & \ldots & \star & \star \\ B_1 & \star & \ldots & \star & \star \\ \mathbb{O} & B_2 & \ldots & \star & \star \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \ldots & B_\kappa & \star \end{pmatrix}, \tag{5.0.4}$$

where B_i is an $m_{i+1} \times m_i$ matrix of rank m_{i+1}, while the $\star$ blocks are arbitrary.

If the $\star$ blocks are zero matrices, $\mathcal{K}$ is homogeneous of degree 2 w.r.t. the dilation law

$$\delta_r(x,t) := \mathrm{diag}\Big(r\mathbb{I}_{m_1}, r^3\mathbb{I}_{m_2}, \ldots, r^{2\kappa+1}\mathbb{I}_{m_{\kappa+1}}, r^2\Big)(x,t), \qquad r>0. \tag{5.0.5}$$

The homogeneous Lie group $\mathbb{K} = \Big(\mathbb{R}^N, \circ, \delta_r\Big)$ provides a natural framework for the classical theory of Kolmogorov operators. In particular, it enable us to introduce a suitable Hölder space, in analogy with the ones introduced in Chapters 3 and 4.

Definition 5.0.1 ($\mathbb{K}$-metric and $\mathbb{K}$-Hölder space)**.** *Suppose that **(B)** holds, and accordingly to* (5.0.4) *split $\mathbb{R}^N$ as $\mathbb{R}^{m_1} \times \mathbb{R}^{m_2} \times \ldots \times \mathbb{R}^{m_{\kappa+1}}$. We introduce on $\mathbb{R}^{N+1}$ the norm*

$$\|(x,t)\|_{\mathbb{K}} := \sum_{i=1}^{\kappa+1} |x^{(i)}|^{\frac{1}{2i-1}} + |t|^{\frac{1}{2}}, \qquad x^{(i)} \in \mathbb{R}^{m_i},$$

that is homogeneous of degree 1 *w.r.t. the dilation law δ_r given by* (5.0.5)*. Accordingly, we define on $\mathbb{R}^{N+1}$ the quasi-distance*

$$d_{\mathbb{K}}\Big((x,t),(\xi,\tau)\Big) := \|(\xi,\tau)^{-1} \circ (x,t)\|_{\mathbb{K}},$$

where the composition law $\circ$ is defined in (5.0.3)*.*
Let $\alpha \in (0,1]$ and $\Omega \subseteq \mathbb{R}^{N+1}$. We say that $u:\Omega \to \mathbb{R}$ belongs to the $\mathbb{K}$-Hölder space $C^{\alpha}_{\mathbb{K}}(\Omega)$ if there exists $C>0$ such that

$$|u(x,t) - u(\xi,\tau)| \le C\, d_{\mathbb{K}}\Big((x,t),(\xi,\tau)\Big)^{\alpha}, \qquad \forall (x,t),(\xi,\tau) \in \Omega.$$

This intrinsic structure was exploited by Polidoro in [102], who analysed Kolmogorov operators with $\mathbb{K}$-Hölder continuous coefficients

$$\mathcal{K}_H = \sum_{i,j=1}^{m_1} a_{ij}(x,t)\partial^2_{x_i x_j} + \langle Bx, \nabla\rangle - \partial_t. \tag{5.0.6}$$

As in Part I, Polidoro assumed the following *uniform ellipticity condition*

(A) *A is a $m_1 \times m_1$ symmetric matrix, and there exists $\lambda > 0$ such that*

$$\frac{1}{\lambda}|\xi|^2 \le \langle A(x,t)\xi,\xi\rangle \le \lambda|\xi|^2, \quad \forall (x,t) \in \mathbb{R}^{N+1}, \forall \xi \in \mathbb{R}^{m_1}. \tag{5.0.7}$$

In line with the method employed in Chapter 4, he proved the existence of the fundamental solution (together with some Gaussian estimates) by means of Levi's parametrix method. In this Hölder coefficients setting, classical solutions needs to be addressed in a suitable functional space: let $\Omega \subset \mathbb{R}^{N+1}$, classical solutions are formulated in the following kinetic space

$$C^2_{\mathbb{K}}(\Omega) := \Big\{u \in C(\Omega) : \partial_{x_i}u, \partial^2_{x_ix_j}u, Yu \in C(\Omega), \quad \text{for every} \ \ i,j = 1,\dots,m_1\Big\}.$$

Here, $Yu(x,t) = \langle Bx, \nabla u(x,t)\rangle - \partial_t u(x,t)$ needs to be addressed as *Lie derivative*

$$Yu(x,t) := \lim_{s\to 0} \frac{u\Big(E(-s)x, t-s\Big) - u(x,t)}{s}, \qquad (x,t) \in \Omega.$$

Definition 5.0.2 (Classical Lie solution)**.** *Let $g \in C(\Omega)$. We say that a function u is a classical solution to*

$$\mathcal{K}_H u(x,t) = \sum_{i,j=1}^{m_1} a_{ij}(x,t)\partial^2_{x_ix_j}u(x,t) + \langle Bx, \nabla u(x,t)\rangle - \partial_t u(x,t) = g(x,t),$$
$$(x,t) \in \Omega,$$

if $u \in C^2_{\mathbb{K}}(\Omega)$ and the above equation is satisfied at every point $(x,t) \in \Omega$.

Polidoro results were later extended by Di Francesco and Pascucci in [47,48]. In these works, the authors considered an extension of (5.0.6) with low order terms

$$\mathcal{K}_L = \sum_{i,j=1}^{m_1} a_{ij}(x,t)\partial^2_{x_ix_j} + \sum_{i=1}^{m_1} b_i(x,t)\partial_{x_i} + c(x,t) + \langle Bx, \nabla\rangle - \partial_t. \tag{5.0.8}$$

Beside extending Polidoro's results on the fundamental solution, Di Francesco and Pascucci investigated the relevant Cauchy problems in the strip $S_T := \mathbb{R}^N \times (0,T)$. We summarize some of the results we will employ in Chapter 6.

Theorem 5.0.3. *Let $\mathcal{K}_L$ as in* (5.0.8) *verifying **(A)** and **(B)**. Suppose that $a_{ij} \in C^\alpha_{\mathbb{K}}(S_T)$ for some $\alpha \in (0,1]$, and assume that the for every $i = 1,\dots,m_1$ the low order coefficients b_i and c are bounded, continuous, and satisfy*

$$|b_i(x,t) - b_i(\xi,t)|, |c(x,t) - c(\xi,t)| \le C\, d_{\mathbb{K}}\Big((x,t),(\xi,t)\Big)^\alpha, \qquad \forall (x,t),(\xi,t) \in S_T, \tag{5.0.9}$$

for some $C > 0$. Then, there exists a fundamental solution Γ to (5.0.8) *such that:*

1. *for every $\zeta = (\xi,\tau) \in \mathbb{R}^N \times (0,T)$, $\Gamma(\cdot,\zeta)$ is a classical solution to $\mathcal{K}_L\Gamma(\cdot,\zeta) = 0$ in $\mathbb{R}^N \times (\tau,T)$, moreover $\Gamma(\cdot,\zeta) \in L^1_{\text{loc}}\Big(\mathbb{R}^N \times (\tau,T)\Big)$;*

2. *the following Gaussian estimates holds: for every $\varepsilon, T > 0$ there exists $C > 0$ (de-*

pending only on ε, T, and the Hölder norm of the coefficients) such that

$$\begin{aligned} \Gamma(x,t,\xi,\tau) &\le C\,\Gamma^{\varepsilon}(x,t,\xi,\tau), & \left|\partial_{x_i}\Gamma(x,t,\xi,\tau)\right| &\le \frac{C}{\sqrt{t-\tau}}\,\Gamma^{\varepsilon}(x,t,\xi,\tau), \\ \left|\partial^2_{x_ix_j}\Gamma(x,t,\xi,\tau)\right| &\le \frac{C}{t-\tau}\,\Gamma^{\varepsilon}(x,t,\xi,\tau), & \left|Y\Gamma(x,t,\xi,\tau)\right| &\le \frac{C}{t-\tau}\,\Gamma^{\varepsilon}(x,t,\xi,\tau), \end{aligned} \tag{5.0.10}$$

for every $x, \xi \in \mathbb{R}^N$ and $0 < t-\tau \le T$. Here, Γ^{ε} denotes the fundamental solution of

$$\mathcal{K}^{\varepsilon} := \sum_{i,j=1}^{m_1} (\lambda+\varepsilon)\partial^2_{x_ix_j} + \langle Bx, \nabla\rangle - \partial_t,$$

where λ is the uniform ellipticity constant given by (5.0.7)*;*

3. *the reproduction property holds*

$$\Gamma(x,t,\xi,\tau) = \int_{\mathbb{R}^N} \Gamma(x,t,y,s)\Gamma(y,s,\xi,\tau)dy,$$

for every $x, \xi \in \mathbb{R}^N$ and $T > t > s > \tau > 0$;

4. *the normalization property holds: if $c(x,t) \equiv c$, then*

$$\int_{\mathbb{R}^N} \Gamma(x,t,\xi,\tau)d\xi = \mathrm{e}^{c(t-\tau)},$$

for every $x \in \mathbb{R}^N$ and $T > t > \tau > 0$;

5. *let $g \in C(S_T)$ such that for any compact $K \subset \mathbb{R}^N$ the following relation holds*

$$|g(x,t) - g(y,t)| \le c\,\frac{d_{\mathbb{K}}\big((x,t),(y,t)\big)^{\beta}}{t^{1-\delta}}, \qquad x, y \in K, t \in (0,T), \tag{5.0.11}$$

for some $c = c(K)$ and $\beta, \delta \in (0,1)$, and assume it satisfies the exponential growth

$$|g(x,t)| \le C\,\frac{\mathrm{e}^{C|x|^2}}{\sqrt{t}}, \qquad (x,t) \in S_T, \tag{5.0.12}$$

for some $C > 0$; let $g_0 \in C(\mathbb{R}^N)$ satisfying the same exponential growth

$$|g_0(x)| \le C\,\mathrm{e}^{C|x|^2}, \qquad x \in \mathbb{R}^N,$$

then there exists $0 < T_0 \le T$ (only depending on C) such that the function

$$\begin{aligned} u(x,t) = \int_{\mathbb{R}^N} \Gamma(x,t,\xi,0)g_0(\xi)d\xi - \int_0^t \int_{\mathbb{R}^N} \Gamma(x,t,\xi,\tau)g(\xi,\tau)d\xi d\tau, \\ (x,t) \in \mathbb{R}^N \times (0,T_0), \end{aligned} \tag{5.0.13}$$

belongs to $C^2_{\mathbb{K}}\Big(\mathbb{R}^N \times (0,T_0)\Big)$ *and solves in classical sense*

$$\begin{cases} \mathcal{K}_L u(x,t) = g(x,t), & (x,t) \in \mathbb{R}^N \times (0,T_0), \\ \lim\limits_{(x,t)\to(x_0,0)} u(x,t) = g_0(x_0), & x_0 \in \mathbb{R}^N; \end{cases} \tag{5.0.14}$$

6. *if* u *is a solution to the Cauchy problem* (5.0.14) *with null* g *and* g_0*, and verifies estimate* (5.0.12)*, then* $u \equiv 0$*; in particular, the function in* (5.0.13) *is the unique classical solution to* (5.0.14) *verifying estimate* (5.0.12).

Remark 5.0.4. *Assumption* (5.0.9) *is weaker than intrinsic Hölder continuity imposed on the low order terms in [47, 48]. Nevertheless, a careful inspection of the parametrix method reveals that* (5.0.9) *suffices to ensure the validity of the theorem. Moreover, it aligns with the regularity requirements in* (5.0.11).

In Chapter 6 we employ Theorem 5.0.3 to study the Cauchy problem relevant to the Kuramoto model (1.3.5). To address existence, estimates, and uniqueness, we rely on the fundamental solution and its Gaussian bounds (5.0.10), which allow us to handle the nonlinearity. We will show that the solution belongs to the kinetic space $C^2_{\mathbb{K}}\Big(\mathbb{R}^2 \times (0,T)\Big)$, and therefore possesses a continuous Lie derivative. Building on this regularity, we propose a finite difference method specifically tailored to the structure of the equation, based on the discretization (1.3.6) of the Lie derivative. This subelliptic approximation was introduced by Polidoro and Mogavero in [103], and later analysed by Di Francesco, Foschi, and Pascucci in [46] in the context of a financial model. We emphasize that this scheme is well suited to our setting, as it allows us to establish stability conditions analogous to those for the heat equation.

In Chapter 7, we study a mean field optimal control problem by employing the theory of Kolmogorov operators in a less regular setting. First, we address the Cauchy problem (5.0.1) with an interaction kernel K that is merely measurable w.r.t. the time variable t. In this setting, we prove the existence within an appropriate strong solutions framework. Subsequentially, we derive some first order optimality conditions for the optimal controls by means of the Lagrangian formalism, which requires a suitable weak solutions setting.

We start by introducing the framework of strong solutions. This setting has been investigated in recent years, mainly because it provides a robust analytical framework for dealing with stochastic processes whose coefficients lack regularity in time (see for instance [71, 94]). Regarding the fundamental solution for (5.0.8), Bramanti and Polidoro [36] proved its existence whenever the coefficients a_{ij}, b_i, c are measurable and only time depending. Their results were later generalized by Lucertini, Pagliarani, and Pascucci [90] to include coefficients that may also depend on the space variable. In this work, the authors employed the intrinsic Hölder continuity assumption introduced in [94], that is tailored to the low regularity in time.

Definition 5.0.5. *Suppose that **(B)** holds, and let $\alpha \in (0,1]$. We denote by $C_b^\alpha(\mathbb{R}^N)$ the set of continuous and bounded function h such that*

$$\|h\|_{C_b^\alpha(\mathbb{R}^N)} := \sup_{x\in\mathbb{R}^N} |h(x)| + \sup_{x,y,\in\mathbb{R}^N} \frac{|h(x)-h(y)|}{d_{\mathbb{K}}\Big((x,0),(y,0)\Big)^\alpha} < +\infty.$$

Accordingly, we define as $L^\infty\Big((0,T);C_b^\alpha(\mathbb{R}^N)\Big)$ the set of measurable functions $u:(0,T)\to C_b^\alpha(\mathbb{R}^N)$ such that

$$\|u\|_{L^\infty((0,T);C_b^\alpha(\mathbb{R}^N))} := \sup_{t\in(0,T)} \|u(t)\|_{C_b^\alpha(\mathbb{R}^N)} < +\infty.$$

Following the intrinsic framework developed in [94], Lucertini, Pagliarani and Pascucci in [90] introduced a suitable notion of strong solutions associated with (5.0.8). In analogy with the case of classical solutions, the crucial step consists in formulating a rigorous definition of strong Lie derivative: we say that u is a.e. Lie differentiable along Y on the strip $S_T = \mathbb{R}^N \times (0,T)$ if there exists $F \in L^1_{\text{loc}}\Big((0,T);C_b(\mathbb{R}^N)\Big)$ such that

$$u\Big(E(s-t)x,s\Big) = u(x,t) + \int_s^t F\Big(E(\tau-t)x,\tau\Big)d\tau,$$

for every $(x,t)\in S_T$ and $s\in(0,T)$.

Definition 5.0.6 (Strong Lie solution)**.** *Let $g \in L^1_{\text{loc}}\Big((0,T);C_b(\mathbb{R}^N)\Big)$. We say that a function u is a strong solution to*

$$\begin{aligned}\mathcal{K}_L u(x,t) = &\sum_{i,j=1}^{m_1} a_{ij}(x,t)\partial^2_{x_i x_j}u(x,t) + \sum_{i=1}^{m_1} b_i(x,t)\partial_{x_i}u(x,t) + c(x,t)u(x,t)\\ &+ \langle Bx, \nabla u(x,t)\rangle - \partial_t u(x,t) = g(x,t), \quad (x,t)\in S_T,\end{aligned} \tag{5.0.15}$$

if u is a continuous function such that $\partial_{x_i}u, \partial^2_{x_i x_j}u \in L^1_{\text{loc}}\Big((0,T);C_b(\mathbb{R}^N)\Big)$ for every $i,j=1,\dots,m_1$, and such that $Yu = \langle Bx,\nabla u\rangle - \partial_t u$ is an a.e. strong Lie derivative, i.e.

$$u\Big(E(s-t)x,s\Big) = u(x,t) - \int_s^t \Big(\sum_{i,j=1}^{m_1} a_{ij}\partial^2_{x_i x_j}u + \sum_{i=1}^{m_1} b_i\partial_{x_i}u + cu - g\Big)\Big(E(\tau-t)x,\tau\Big)d\tau,$$

for every $(x,t)\in S_T$ and $s\in(0,T)$.

Combining the results of [36] and Levi's parametrix method, the authors of [90] proved the existence of a strong fundamental solution, assuming the minimal $L^\infty\Big((0,T);C_b^\alpha(\mathbb{R}^N)\Big)$ regularity on the coefficients. We recall some results from [36,90] on the fundamental solution, the corresponding Gaussian bounds, and the relevant Cauchy problems, which will play a crucial role in Chapter 7.

Theorem 5.0.7. *Let $\mathcal{K}_L$ as in* (5.0.8) *verifying **(A)** and **(B)**, and assume that the coefficients a_{ij}, b_i, c belongs to $L^\infty\Big((0,T); C_b^\alpha(\mathbb{R}^N)\Big)$ for some $\alpha \in (0,1]$. Then, there exists a fundamental solution Γ to* (5.0.15) *such that:*

1. *for every $\zeta = (\xi,\tau) \in \mathbb{R}^N \times (0,T)$, $\Gamma(\cdot,\zeta)$ is a strong solution to $\mathcal{K}_L\Gamma(\cdot,\zeta) = 0$ in $\mathbb{R}^N \times (\tau,T)$, in particular $\Gamma(\cdot,\zeta) \in L^1_{\mathrm{loc}}\Big(\mathbb{R}^N \times (\tau,T)\Big)$;*

2. *the following Gaussian estimates hold: for every $\varepsilon, T > 0$ there exists $C > 0$ (depending only on ε, T, and the coefficients) such that*

$$\begin{aligned}
\Gamma(x,t,\xi,\tau) &\le C\,\Gamma^\varepsilon(x,t,\xi,\tau),\\
\Big|\partial_{x_i}\Gamma(x,t,\xi,\tau)\Big| &\le \frac{C}{\sqrt{t-\tau}}\,\Gamma^\varepsilon(x,t,\xi,\tau),\\
\Big|\partial^2_{x_ix_j}\Gamma(x,t,\xi,\tau)\Big| &\le \frac{C}{t-\tau}\,\Gamma^\varepsilon(x,t,\xi,\tau),
\end{aligned}$$

   for every $x,\xi \in \mathbb{R}^N$ and $0 < t-\tau \le T$, where Γ^ε denotes the fundamental solution of

$$\mathcal{K}^\varepsilon = \sum_{i,j=1}^{m_1} (\lambda+\varepsilon)\partial^2_{x_ix_j} + \langle Bx, \nabla\rangle - \partial_t,$$

   with λ given by (5.0.7)*;*

3. *let $g_0 \in L^q\Big(\mathbb{R}^N\Big)$ for some $q \in [1,+\infty)$, then the function*

$$u(x,t) = \int_{\mathbb{R}^N} \Gamma(x,t,\xi,0)g_0(\xi)d\xi, \qquad (x,t) \in S_T,$$

   solves in strong sense

$$\begin{cases}\mathcal{K}_L u(x,t) = 0, & (x,t) \in S_T,\\ \lim\limits_{t\to 0} \|u(\cdot,t) - g_0\|_{L^q(\mathbb{R}^N)} = 0.\end{cases}$$

Remark 5.0.8. *As mentioned in Remark 5.0.4, requiring the $L^\infty\Big((0,T); C_b^\alpha(\mathbb{R}^N)\Big)$ regularity for the coefficients appears to be a more natural and flexible condition, even when the coefficients are assumed to be continuous in time. Indeed, as noted by Pascucci and Pesce in [94], intrinsic Hölder continuity $C^\alpha_{\mathbb{K}}$ (namely, intrinsic Hölder regularity coupling space and time variables) may be a rather restrictive assumption, even in the simple case of smooth* *bounded functions. For instance, consider on $\mathbb{R}^3$ the function $f(x, y, t) = \sin(y)$: as follows from Example 1.3 in [94], f does not belongs to $C^{\alpha}_{\mathbb{K}}(S_T)$, but clearly $f(\cdot, \cdot, t) \in C^{\alpha}_b(\mathbb{R}^N)$ for every $t \in (0, T)$.*

We conclude this preliminary chapter by introducing the functional setting of weak solutions. For a comprehensive overview of the weak theory, we refer the interested reader to the survey [8] by Anceschi, Piccinini and Rebucci. In the following, we consider divergence form Kolmogorov operators on the strip $S_T = \mathbb{R}^N \times (0, T)$

$$\mathcal{K}_D = \sum_{i,j=1}^{m_1} \partial_{x_i}\Big(a_{ij}(x,t)\partial_{x_j}\Big) + \sum_{i=1}^{m_1} b_i(x,t)\partial_{x_i} + c(x,t) + \langle Bx, \nabla\rangle - \partial_t, \qquad (x,t) \in S_T, \tag{5.0.16}$$

where $A = (a_{i,j})_{i,j=1}^{m_1}$ satisfies **(A)**, while B satisfies **(B)** with the $\star$ blocks being zero matrices. As we have seen for classical and strong solutions, a central issue in defining an appropriate variational formulation lies in the suitable treatment of the Lie derivative $Y = \langle Bx, \nabla\rangle - \partial_t$. Since we have diffusion in the first m_1 variables, it is natural to introduce the Sobolev space

$$H^1_{m_1}(\mathbb{R}^{m_1}) := \Big\{u \in L^2(\mathbb{R}^{m_1}) : \partial_{x_i} u \in L^2(\mathbb{R}^{m_1}), \text{ for every } i = 1, \dots, m_1\Big\},$$

equipped with the norm

$$\|u\|^2_{H^1_{m_1}(\mathbb{R}^{m_1})} := \|u\|^2_{L^2(\mathbb{R}^{m_1})} + \sum_{i=1}^{m_1} \|\partial_{x_i} u\|^2_{L^2(\mathbb{R}^{m_1})}.$$

We denote its dual space as $H^{-1}_{m_1}(\mathbb{R}^{m_1})$. For every $x \in \mathbb{R}^N$ we write $\bar{x}$ to denote its first m_1 components, and $\underline{x}$ to denote its remaining $\mathbb{R}^{N-m_1}$ components. The functional space suited to the weak formulation of (5.0.16) is the space W_T defined as the closure of the space $C^{\infty}_c\Big(\mathbb{R}^N \times [0, T]\Big)$ w.r.t. the norm

$$\|u\|_{W_T} := \int_0^T \int_{\mathbb{R}^{N-m_1}} \|u(\cdot, \underline{x}, t)\|^2_{H^1_{m_1}(\mathbb{R}^{m_1})} d\underline{x} dt + \int_0^T \int_{\mathbb{R}^{N-m_1}} \|Yu(\cdot, \underline{x}, t)\|^2_{H^{-1}_{m_1}(\mathbb{R}^{m_1})} d\underline{x} dt.$$

We are now in position to give a suitable definition of weak solutions.

Definition 5.0.9 (Weak Lie solution)**.** *Let $g \in L^2\Big(\mathbb{R}^{N-m_1} \times (0,T); H^{-1}_{m_1}(\mathbb{R}^{m_1})\Big)$ and $g_0 \in L^2(\mathbb{R}^N)$. We say that a function u is a weak solution to*

$$\begin{cases} \mathcal{K}_D u(x,t) = g(x,t), & (x,t) \in S_T, \\ u(x,0) = g_0(x), & x \in \mathbb{R}^N, \end{cases}$$

if $u \in W_T$, $u(\cdot, 0) = g_0$ in L^2 sense, and for every $\varphi \in C^{\infty}_c\Big(\mathbb{R}^N \times [0,T]\Big)$ it holds

$$\int_0^T \int_{\mathbb{R}^N} \Big(- \sum_{i,j=1}^{m_1} a_{ij} \partial_{x_j} u \partial_{x_i} \varphi + \sum_{i=1}^{m_1} b_i \partial_{x_i} u \varphi + cu\varphi \Big)(x,t) dx dt$$
$$= - \int_0^T \int_{\mathbb{R}^{N-m_1}} \langle Yu(\cdot, \underline{x}, t), \varphi(\cdot, \underline{x}, t) \rangle d\underline{x} dt + \int_0^T \int_{\mathbb{R}^{N-m_1}} \langle g(\cdot, \underline{x}, t), \varphi(\cdot, \underline{x}, t) \rangle d\underline{x} dt,$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H^{-1}_{m_1}(\mathbb{R}^{m_1})$ and $H^1_{m_1}(\mathbb{R}^{m_1})$.

In Chapter 7, we shall make use the above weak formulation to employ the Lagrangian formalism and derive optimality conditions for the McKean–Vlasov control problem introduced in Chapter 1. Our analysis is based on the recent results by Auscher, Imbert and Niebel [16, 17], who extended Lions' form method to the following class of kinetic Kolmogorov operators

$$\mathcal{K}_K = \mathcal{A}_v - v\partial_x - \partial_t, \qquad (v,x,t) \in \mathbb{R}^N \times \mathbb{R}^N \times (0,T),$$

with $\mathcal{A}_v$ being a general linear second order diffusion operator acting only on the velocity variable v (for instance $\mathcal{A} = \Delta_v$). The main ingredient of these results are suitable kinetic embeddings, which play a crucial role in establishing compactness and continuity properties. These embeddings deeply rely on the structure of the functional space W_T, which is specifically tailored to capture the regularity imposed by the Lie derivative $Y = -v\partial_x - \partial_t$.

Chapter 6

On the study of a Kuramoto model based on degenerate Kolmogorov-Fokker-Planck equations

6.1 Introduction

In this chapter we analyse a particular stochastic process, that is the Kuramoto model with inertia and noise. Our analysis focuses on the Cauchy problems associated with the corresponding mean field PDE. When all oscillators are identical, the PDE describing the evolution of the density is

$$\frac{D}{m^2}\partial^2_{\omega\omega}\varrho(\omega,\vartheta,t) - \frac{1}{m}\partial_\omega\Big(\Big(-\omega + K\int_{\mathbb{R}}\int_0^{2\pi}\sin(\vartheta'-\vartheta)\varrho(\omega,\vartheta,t)d\omega d\vartheta\Big)\varrho_t(\omega,\vartheta)\Big)$$
$$-\,\omega\partial_\vartheta\varrho(\omega,\vartheta,t) - \partial_t\varrho(\omega,\vartheta,t) = 0.$$

where ϱ denotes the density of oscillators with frequency $\omega \in \mathbb{R}$ and phase $\vartheta \in [0, 2\pi]$ at time $t \in [0, T]$, while $D > 0$ represents the noise amplitude and $m > 0$ is a mass-type parameter. If we momentarily disregard the nonlinear term, we see that this PDE falls within the class of Kolmogorov equations introduced in the previous chapter.

Kuramoto models are among the most widely studied models for synchronization phenomena: despite their comparatively simple mathematical structure, they are able to capture with remarkable accuracy a wide range of collective behaviours observed in different frameworks. Synchronization is one of the most fascinating and pervasive phenomena in nature. It pervades nature at every scale from the nucleus to the cosmos, and even our heart exhibits this phenomenon. A cluster of about $10,000$ cells, called the sinoatrial node, generates the electrical rhythm that commands the rest of the heart to beat, and it must do so reliably, minute after minute, for about three billion beats in a lifetime. These cells are a collection of oscillators, *i.e.* entities that cycle automatically in more or less regular time intervals, and need to coordinate their rhythm in order to achieve some kind of synchronization. Huygens was one of the pioneers in the study of synchronization. In February of 1665 the Dutch physicist was confined to his bedroom for several days, and observed a curious phenomenon: the two pendulum clocks in the room, separated by two feet, kept oscillating together without any variation; moreover, mixing up the swings of the pendulums the synchronization returned after some time. Intrigued by this phenomenon, he carried out several experiments, finding out that the synchronization could take place only if they could communicate in some way. In the following centuries many mathematicians studied the problem, coming up with different models: some of these describe specific problems (see [101] for the description of sinoatrial node synchronization), while others have been created with the aim of being general enough to describe the dynamics of different systems, even relevant to very different scientific fields such as biology, physics, and social sciences as well. One of the most famous model that belongs to the latter group is the one proposed by Kuramoto [74], from which many variations have been devised. We refer to the article [111] by Spigler for an overview of this model and some of its derivatives. We next describe the Kuramoto model and some of its modifications we are interested in, then we discuss the contributions of this thesis on this subject.

The Kuramoto model is a system of ordinary differential equations, that describes the collective behaviour of a population that can be seen as a group of oscillators. Let's consider a population of N oscillators, we denote as $\vartheta_k(t)$ the phase related to the k-th oscillator, and as Ω_k its natural frequency, which is an intrinsic parameter of the k-th oscillator. We suppose that these natural frequencies are distributed according to an unimodal distribution $g = g(\Omega)$ that is normalized and symmetric w.r.t. its mean frequency $\overline{\Omega}$. The governing equation is given by the following system of ordinary differential equations

$$\dot{\vartheta}_k = \Omega_k + \frac{K}{N} \sum_{j=1}^{N} \sin\Big(\vartheta_j - \vartheta_k\Big), \qquad k = 1, \dots, N, \tag{6.1.1}$$

where K is a real positive constant sizing the coupling strength between the oscillators. Note that the coupling in the above system is nonlinear, thus the ensuing phenomena is expected to be rather complicated. On the other hand, a mean field coupling, which is present in (6.1.1), significantly simplify the evolution of the system. In order to point out this coupling effect, we introduce the following complex number, called *order parameter*

$$r_N(t)\mathrm{e}^{i\psi_N(t)} = \frac{1}{N} \sum_{j=1}^{N} \mathrm{e}^{i\vartheta_j(t)}. \tag{6.1.2}$$

The phase $\psi_N(t)$ is the mean phase of the system. The modulus $|r_N(t)|$ provides us with a measure of the synchronization of the system. When $|r_N(t)| = 0$ the system is fully incoherent, when $0 < |r_N(t)| < 1$ the system is partly synchronized, and $|r_N(t)| = 1$ matches the full coherence state. Equation (6.1.1) can be rewritten using the order parameter

$$\dot{\vartheta}_k = \Omega_k + K r_N \sin\Big(\psi_N - \vartheta_k\Big), \qquad k = 1, \dots, N, \tag{6.1.3}$$

which is the equation of an overdamped pendulum with torque Ω_k and restoring force proportional to $K r_N$. This formulation simplifies the analytical treatment.
Despite the simplicity of this model, Wiesenfeld, Colet and Strogatz show in [120, 121] that (6.1.3) describes some important physical phenomena such as the interaction of quasi-optical oscillators with a cavity and some Josephson arrays, and it gives us a starting point to study synchronization in generic oscillators. Moreover, Kuramoto model shows how synchronization in a population of many coupled oscillators may occur, and the type of synchronization: indeed it may occur frequency synchronization, where each oscillator completes its cycle at the same time, phase synchronization, where each oscillator is at the same point in the cycle, or even a more complex scenario.
The Kuramoto model can be extended to a population of infinitely many oscillators, as it was described by Strogatz in [112]. Let's consider a continuum of oscillators, and let $\varrho = \varrho(\vartheta, \Omega, t)$ be a density function describing the fraction of oscillators at phase ϑ with natural frequency Ω at time t. Then, ϱ is a nonnegative function 2π periodic in ϑ, and satisfies

$$\int_0^{2\pi} \varrho(\vartheta, \Omega, t) d\vartheta = 1, \qquad \forall\ (\Omega, t) \in \mathbb{R} \times [0, +\infty).$$

The evolution of ϱ is described by the continuity equation

$$\frac{\partial \varrho}{\partial t} = -\frac{\partial}{\partial \vartheta}(\varrho v), \tag{6.1.4}$$

which expresses conservation of oscillators of frequency Ω. Here, $v = v(\vartheta, \Omega, t)$ is the instantaneous velocity of an oscillator at position ϑ, given that it has natural frequency Ω, and is interpreted in the Eulerian sense. From (6.1.3) we see that that velocity is

$$v(\vartheta, \Omega, t) = \Omega + K r(t) \sin(\psi(t) - \vartheta),$$

where $r(t)$ and $\psi(t)$ follow from (6.1.2) by letting $N \to +\infty$

$$r(t) \mathrm{e}^{\mathrm{i}\psi(t)} = \int_0^{2\pi} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\vartheta'} \varrho(\vartheta', \Omega', t) g(\Omega') d\Omega' d\vartheta'. \tag{6.1.5}$$

Combining (6.1.4) and(6.1.5) we obtain the following nonlinear partial integro-differential equation for the density ϱ

$$\partial_t \varrho(\vartheta,\Omega,t) = -\partial_\vartheta \left[\varrho(\vartheta,\Omega,t)\left(\Omega + K\int_0^{2\pi}\int_{\mathbb{R}} \sin(\vartheta'-\vartheta)\varrho(\vartheta',\omega',t)g(\Omega')d\Omega' d\vartheta'\right)\right]. \tag{6.1.6}$$

Sakaguchi in [108] modified the deterministic equations (6.1.3) and (6.1.6) by adding noise to describe stochastic phenomena, in order to consider rapid stochastic fluctuations in the natural frequencies. The governing equation (6.1.3) for N oscillators takes now the form of a system of Langevin equations

$$\dot{\vartheta}_k = \Omega_k + \xi_k + \frac{K}{N}\sum_{j=1}^{N} \sin\Big(\vartheta_j - \vartheta_k\Big), \qquad k = 1,\dots,N, \tag{6.1.7}$$

where $\xi(t)dt = D\,dW_t$ being $(W_t)_{t\geq 0}$ a N-dimensional Wiener process, with $D > 0$.
Then, Sakaguchi argued intuitively that since (6.1.7) is a system of Langevin equations with mean field coupling, as $N \to \infty$ the density ϱ should satisfy the following Fokker–Planck equation

$$\partial_t \varrho = D\,\partial^2_{\vartheta\vartheta}\varrho - \partial_\vartheta \left[\varrho\left(\Omega + K\int_0^{2\pi}\int_{\mathbb{R}} \sin(\vartheta'-\vartheta)\varrho(\vartheta',\Omega',t)g(\Omega')d\Omega' d\vartheta'\right)\right]. \tag{6.1.8}$$

Note that (6.1.8) is a nonlinear parabolic partial integro-differential equation, that reduces to (6.1.6) when $D = 0$. We finally recall that linear stability of the incoherent state, concerning both the deterministic and the stochastic equations (6.1.6) and (6.1.8), respectively, have been studied by Strogatz and Mirollo in [113]. The Cauchy problem relevant to (6.1.8) has been studied in recent years by Lavrentiev, Spigler and Tani in [79], [80] and [81].
Based on some ideas developed by Ermentrout in his studies about fireflies synchronization [52], Tanaka, Lichtenberg and Oishi in [115] elaborated the following second order variation of the Kuramoto model

$$m\ddot{\vartheta}_k + \dot{\vartheta}_k = \Omega_k + \frac{K}{N}\sum_{j=1}^{N} \sin\Big(\vartheta_j - \vartheta_k\Big), \qquad k = 1,\dots,N, \tag{6.1.9}$$

with $m > 0$. We point out that the term $m\ddot{\vartheta}_k$ introduces an inertial effect in the model, in order to better explain some synchronization phenomena. Unlike the original Kuramoto model, Ermentrout's second order model has the remarkable property to allow near-perfect phase synchronization, where the phase shift between synchronized oscillators is inversely proportional to the mass. The equation (6.1.9) can be rewritten in terms of the order parameter (6.1.2) as follows

$$m\ddot{\vartheta}_k + \dot{\vartheta}_k = \Omega_k + Kr_N \sin\Big(\psi_N - \vartheta_k\Big), \qquad i = 1,\dots,N,$$

which is the governing equation for a single damped driven pendulum with torque Ω_k. Starting from the contributions due to Tanaka, Lichtenberg and Oishi [115], Acebroan and Spigler introduced a noise term $\xi_k = \xi_k(t)$ in (6.1.9), so that they considered in [1] a system of second order Langevin equations that consider an inertial term

$$m\ddot{\vartheta}_k + \dot{\vartheta}_k = \Omega_k + Kr_N \sin\Big(\psi_N - \vartheta_k\Big) + \xi_k, \qquad k = 1, \dots, N. \tag{6.1.10}$$

Once we set $\dot{\vartheta}_k = \omega_k$, equation (6.1.10) takes the form of the following system of Stochastic Differential Equations

$$\begin{cases} d\omega_{k,t} = \dfrac{1}{m}\Bigg(-\omega_{k,t} + \Omega_k + \dfrac{K}{N}\displaystyle\sum_{j=1}^{N} \sin\Big(\vartheta_{j,t} - \vartheta_{k,t}\Big)\Bigg)dt + \dfrac{D}{m}\sqrt{2}\, dW_{k,t}, & k = 1, \dots, N, \\ d\vartheta_{k,t} = \omega_{k,t} dt, & k = 1, \dots, N, \end{cases}$$

that, as N goes to infinity, converges in the sense of propagation of chaos to the corresponding McKean-Vlasov process. Its density solves the following nonlinear, degenerate, parabolic integro-differential equation

$$\frac{D}{m^2}\, \partial^2_{\omega\omega}\varrho + \frac{1}{m}\partial_\omega\Big[\Big(\omega - \Omega - K_\varrho(\vartheta, t)\Big)\varrho\Big] - \omega\partial_\vartheta\varrho - \partial_t\varrho = 0, \tag{6.1.11}$$

where

$$K_\varrho(\vartheta, t) = K\int_{\mathbb{R}}\int_{\mathbb{R}}\int_0^{2\pi} g(\Omega')\sin(\vartheta' - \vartheta)\varrho(\omega', \vartheta', \Omega', t)d\vartheta' d\omega' d\Omega'. \tag{6.1.12}$$

In order to simplify the notation, in the sequel we set $D = 1$ and $m = 1$. For a given $\varrho_0 = \varrho_0(\omega, \vartheta, \Omega)$ we consider the following Cauchy problem

$$\begin{cases} \partial^2_{\omega\omega}\varrho + \partial_\omega\Big[\Big(\omega - \Omega - K_\varrho(\vartheta, t)\Big)\varrho\Big] - \omega\partial_\vartheta\varrho - \partial_t\varrho = 0, & (\omega, \vartheta, \Omega, t) \in \mathbb{R}^3 \times (0, T), \\ \varrho(\omega, \vartheta, \Omega, 0) = \varrho_0(\omega, \vartheta, \Omega), & (\omega, \vartheta, \Omega) \in \mathbb{R}^3. \end{cases} \tag{6.1.13}$$

The Cauchy problem (6.1.13) has been addressed by Akhmetov, Lavrentiev and Spigler in [77], [78], [2] and [3] by using a parabolic regularization of the governing equation.
The main result proved in [2] is the existence and the uniqueness of a classical solution ϱ to (6.1.13). Since $(\omega, \vartheta) \mapsto \varrho(\omega, \vartheta, \Omega, t)$ denotes the density of a probability measure, the authors of the aforementioned references focus on positive solutions with the property

$$\int_{(0,2\pi)\times\mathbb{R}} \varrho(\omega, \vartheta, \Omega, t)d\vartheta\, d\omega = 1, \qquad t \geq 0,\ \Omega \in \mathbb{R}.$$

Moreover, the solutions considered there have the following *exponential decay property*.

Definition 6.1.1. *We say that a function $f : \mathbb{R}^4 \to \mathbb{R}$ has the* exponential decay property *if there exist two positive constants M, C such that:*

$$(\mathbf{E}) \qquad |f(\omega, \vartheta, \Omega, t)| \leq C\mathrm{e}^{-M\omega^2}, \qquad \forall (\omega, \vartheta, \Omega, t) \in \mathbb{R}^4.$$

The main result proved in [2] is the existence of a unique positive classical solution ϱ to the Cauchy problem (6.1.13), with the properties listed above, provided that ϱ_0 and some of its derivatives have the exponential decay property.

In this chapter we address the Cauchy problem (6.1.13) by using the theory of subelliptic Kolmogorov equations in homogeneous Lie groups, as described in Chapter 5. In particular, the first order differential operator $Y = \omega\partial_\omega - \omega\partial_\vartheta - \partial_t$ is considered as a Lie derivative.

Definition 6.1.2 (Lie derivative)**.** *For every for $z = (\omega, \vartheta, \Omega, t) \in \mathbb{R}^4$ we define as* integral path of Y starting at z *the solution to the following Cauchy problem*

$$\begin{cases} \dot\gamma_{Y,z}(s) = Y\Big(\gamma_{Y,z}(s)\Big), \\ \gamma_{Y,z}(0) = z. \end{cases}$$

We say that a function ϱ is Lie differentiable at $z = (\omega, \vartheta, \Omega, t) \in \mathbb{R}^4$ w.r.t. Y if the following limit

$$Y\varrho(z) = \lim_{s \to 0} \frac{\varrho\Big(\gamma_{Y,z}(s)\Big) - \varrho(z)}{s},$$

exists and is finite.

This definition clarifies the meaning of *classical solution* to (6.1.11). A function ϱ is a classical solution if it is continuous, its derivatives $\partial_\omega\varrho$ and $\partial^2_{\omega\omega}\varrho$ are defined as continuous functions, the Lie derivative $Y\varrho$ is continuous and the equation (6.1.11) is satisfied at every point. The main advantage in our approach, with respect to [2], is that we don't require unnecessary regularity on the data ϱ_0. Moreover, our hypothesis (6.1.14) below does not require the compactness of the support of g assumed in [2]. Finally, a numerical approximation of the Lie derivative $Y\varrho$ allows us to define a stable numerical method for the Cauchy problem (6.1.13).
The main result of this chapter is the following theorem.

Theorem 6.1.3. *Let g be a nonnegative function such that*

$$\int_{\mathbb{R}} g(\Omega) d\Omega = 1,$$

and assume that exists $\beta > 2$ *such that*

$$\int_{\mathbb{R}} g(\Omega) e^{|\Omega|^\beta} d\Omega < +\infty. \tag{6.1.14}$$

Let $\varrho_0 \in C(\mathbb{R}^3)$ *be a strictly positive function,* 2π *periodic in* ϑ *such that:*

1. ϱ_0 *verifies the exponential decay property* (**E**) *as stated in Definition 6.1.1;*

2. *for every* $\Omega \in \mathbb{R}$ *we have*

$$\int_{(0,2\pi)\times\mathbb{R}} \varrho_0(\omega, \vartheta, \Omega) d\vartheta d\omega = 1.$$

Then, there exists a strictly positive classical solution ϱ *to the Cauchy problem* (6.1.13)*, defined for* $(\omega, \vartheta, \Omega, t) \in \mathbb{R}^3 \times [0, +\infty)$ *such that*

$$\int_{(0,2\pi)\times\mathbb{R}} \varrho(\omega, \vartheta, \Omega, t) d\vartheta d\omega = 1, \qquad \textit{for every} \quad t \geq 0, \ \Omega \in \mathbb{R}.$$

Moreover, ϱ *is* 2π *periodic in* ϑ*, continuously depends on* Ω *and for every* $T > 0$ *and* $\Omega \in \mathbb{R}$ *there exist positive constants* $C_{\Omega,T}$*,* $\overline{M}$ *such that:*

1. *the function* $(\omega, \vartheta, \Omega, t) \mapsto \varrho(\omega, \vartheta, \Omega, t)$ *has the exponential decay property* (**E**)*, with constants* $\overline{M}$ *and* $C_{\Omega,T}$*;*

2. *the function* $(\omega, \vartheta, \Omega, t) \mapsto \partial_\omega \varrho(\omega, \vartheta, \Omega, t)$ *has the exponential decay property* (**E**)*, with constants* $\overline{M}$ *and* $C_{\Omega,T}/\sqrt{t}$*;*

3. *the function* $(\omega, \vartheta, \Omega, t) \mapsto \partial^2_{\omega\omega} \varrho(\omega, \vartheta, \Omega, t)$ *has the exponential decay property* (**E**)*, with constants* $\overline{M}$ *and* $C_{\Omega,T}/t$*. The same assertion holds true for the Lie derivative* $Y\varrho = (\omega\partial_\omega - \omega\partial_\vartheta - \partial_t)\varrho$*.*

Furthermore, ϱ *is the unique positive solution satisfying the properties listed above whenever* g *has compact support.*

Following [2], we look for a solution to the nonlinear Cauchy problem (6.1.13) as the fixed point of a map that associates to a given function ϱ the solution to a suitable linearized problem, which in our case is relevant to a degenerate Kolmogorov equation. With this aim, we introduce the following Cauchy problem

$$\begin{cases} \partial^2_{\omega\omega} u + \omega\partial_\omega u - \omega\partial_\vartheta u + \Phi_{\widetilde{\varrho}}(\vartheta, \Omega, t)\partial_\omega u + u - \partial_t u = f, & (\omega, \vartheta, \Omega, t) \in \mathbb{R}^3 \times (0, T_1), \\ u(\omega, \vartheta, \Omega, 0) = u_0(\omega, \vartheta, \Omega), & (\omega, \vartheta, \Omega) \in \mathbb{R}^3, \end{cases} \tag{6.1.15}$$

where

$$\Phi_{\widetilde{\varrho}}(\vartheta, \Omega, t) = -\Omega - K \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^{2\pi} g(\Omega') \sin(\vartheta' - \vartheta) \widetilde{\varrho}(\omega', \vartheta', \Omega', t) d\vartheta' d\omega' d\Omega'. \qquad (6.1.16)$$

If $\widetilde{\varrho}$ is a given continuous function satisfying the exponential decay property (see Definition 6.1.1), then there exists a fundamental solution $\Gamma_{\Omega,\widetilde{\varrho}}$ useful to represent the solution u to (6.1.15) as follows

$$\begin{aligned} u(\omega, \vartheta, \Omega, t) = &\int_{\mathbb{R}^2} \Gamma_{\Omega,\widetilde{\varrho}}(\omega, \vartheta, t; \xi, \eta, 0) u_0(\xi, \eta, \Omega) d\eta d\xi - \\ &- \int_0^t \int_{\mathbb{R}^2} \Gamma_{\Omega,\widetilde{\varrho}}(\omega, \vartheta, t; \xi, \eta, \tau) f(\xi, \eta, \Omega, \tau) d\eta d\xi d\tau, \end{aligned}$$

when f and u_0 verify appropriate assumptions. In Section 6.2 we study the existence of $\Gamma_{\Omega,\widetilde{\varrho}}$ and some bounds of $\Gamma_{\Omega,\widetilde{\varrho}}$ and its derivatives, that provide us with the proof of the existence of the unique solution to (6.1.13).

This chapter is organized as follows: in Section 6.2 we recall some results about the fundamental solution and we give some bounds useful in the proof of Theorem 6.1.3, Section 6.3 is devoted to the proof of Theorem 6.1.3. Finally, in Section 6.4 we define a stable and consistent finite difference scheme that follows the structure of (6.1.11).

6.2 Fundamental Solution

In this section we recall some known facts about the partial differential operator

$$\mathscr{L}_{\Omega,\widetilde{\varrho}} = \partial^2_{\omega\omega} + \omega\partial_\omega - \omega\partial_\vartheta + \Phi_{\widetilde{\varrho}}(\vartheta, \Omega, t)\partial_\omega + 1 - \partial_t, \qquad (6.2.1)$$

appearing in the Cauchy problem (6.1.15), where the function $\Phi_{\widetilde{\varrho}}$ has been introduced in (6.1.16)

$$\Phi_{\widetilde{\varrho}}(\vartheta, \Omega, t) = -\Omega - K \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^{2\pi} g(\Omega') \sin(\vartheta' - \vartheta) \widetilde{\varrho}(\omega', \vartheta', \Omega', t) d\vartheta' d\omega' d\Omega'.$$

Note that no derivatives w.r.t. Ω appear in $\mathscr{L}_{\Omega,\widetilde{\varrho}}$, then from now on we fix $\Omega \in \mathbb{R}$ and we consider it as a parameter.
Let us now recall some known facts about the degenerate differential operator $\mathscr{L}_{\Omega,\widetilde{\varrho}}$ that will be used in the proof of Theorem 6.1.3. For a reason that will be clear in the following, we also introduce a parameter $\varepsilon > 0$, and we set

$$\mathscr{L}^\varepsilon := (1 + \varepsilon)\partial^2_{\omega\omega} + \omega\partial_\omega - \omega\partial_\vartheta - \partial_t. \qquad (6.2.2)$$

As follows from Chapter 5, this strongly degenerate operator has a fundamental solution $\Gamma^\varepsilon = \Gamma^\varepsilon(\omega, \vartheta, t, \omega_0, \vartheta_0, t_0)$ which is smooth w.r.t. the variable (ω, ϑ, t) belonging to the set $\mathbb{R}^3 \setminus \{(\omega_0, \vartheta_0, t_0)\}$. In order to align the exposition with that of Chapter 5 we introduce some further notation, that will allow us to write the explicit expression of Γ^ε. We first write $\mathscr{L}^\varepsilon$ in the form

$$\mathscr{L}^\varepsilon = \sum_{i,j=1}^{2} a_{ij}\partial^2_{x_i x_j} + \sum_{i,j=1}^{2} b_{ij} x_j \partial_{x_i} - \partial_t,$$

where $x = (x_1, x_2) = (\omega, \vartheta)$, and

$$A_\varepsilon := (a_{ij})_{i,j=1,2} = \begin{pmatrix} 1+\varepsilon & 0 \\ 0 & 0 \end{pmatrix}, \quad B := (b_{ij})_{i,j=1,2} = \begin{pmatrix} 1 & 0 \\ -1 & 0 \end{pmatrix}.$$

Following Chapter 5, for every $t > 0$ we set

$$E(t) := \exp\Big(-tB\Big), \qquad C_\varepsilon(t) := \int_0^t E(s) A_\varepsilon E^T(s)\, ds.$$

A plain computation shows that

$$E(t) = \begin{pmatrix} \mathrm{e}^{-\mathrm{t}} & 0 \\ 1-\mathrm{e}^{-\mathrm{t}} & 1 \end{pmatrix}, \qquad C_\varepsilon(s) = \frac{1+\varepsilon}{2} \begin{pmatrix} 1-\mathrm{e}^{-2t} & 1-2\mathrm{e}^{-t}+\mathrm{e}^{-2t} \\ 1-2\mathrm{e}^{-t}+\mathrm{e}^{-2t} & 2t-3+4\mathrm{e}^{-t}-\mathrm{e}^{-2t} \end{pmatrix}.$$

Note that the matrix $C_\varepsilon(t)$ is symmetric and strictly positive for every $t > 0$, then it is invertible and the fundamental solution Γ^ε of $\mathscr{L}^\varepsilon u = 0$ is defined for every $(x,t) = (\omega, \vartheta, t) \in \mathbb{R}^2 \times (0, +\infty)$ as follows

$$\Gamma^\varepsilon(x,t) = \frac{1}{4\pi\sqrt{\det C_\varepsilon(t)}} \exp\left(-\tfrac{1}{4}\langle C_\varepsilon^{-1}(t)x, x\rangle - t\right).$$

As customary in the setting of parabolic equation, $\Gamma^\varepsilon(x,t) = 0$ whenever $t \le 0$. Finally, the fundamental solution with singularity at any point $(x_0, t_0) \in \mathbb{R}^3$ is defined as

$$\Gamma^\varepsilon(x, t, x_0, t_0) := \Gamma^\varepsilon\Big(x - E(t-t_0)x_0, t - t_0\Big). \tag{6.2.3}$$

The following property of Γ^ε will be useful in the sequel

$$\int_{\mathbb{R}^2} \Gamma^\varepsilon(\omega, \vartheta, t, \xi, \eta, \tau) d\eta d\xi = 1, \qquad \forall (\omega, \vartheta) \in \mathbb{R}^2, \tau < t. \tag{6.2.4}$$

Moreover, a direct computation shows that the following identity holds true

$$\int_{\mathbb{R}} \Gamma^\varepsilon(\omega, \vartheta, t, \xi, \eta, \tau) d\eta = \widetilde{\Gamma}^\varepsilon(\omega, t, \xi, \tau), \qquad \forall (\omega, \vartheta, t, \xi, \tau) \in \mathbb{R}^5, \quad \text{with } \tau < t.$$

Here

$$\widetilde{\Gamma}^{\varepsilon}(\omega,t,\xi,\tau)=\frac{1}{\sqrt{(1+\varepsilon)\pi(1-e^{-2(t-\tau)})}}\exp\left(-\tfrac{|\omega-e^{-(t-\tau)}\xi|^2}{(1+\varepsilon)(1-e^{-2(t-\tau)})}-(t-\tau)\right), \tag{6.2.5}$$

is the fundamental solution to the operator $(1+\varepsilon)\partial^2_{\omega\omega}+\omega\partial_\omega-\partial_t$.
Let's turn our attention to the Cauchy problem (6.1.15)

$$\begin{cases}\mathscr{L}_{\Omega,\widetilde{\varrho}}\,u=f, & (\omega,\vartheta,t)\in\mathbb{R}^2\times(0,T_1),\\ u(\omega,\vartheta,0)=u_0(\omega,\vartheta), & (\omega,\vartheta)\in\mathbb{R}^2.\end{cases}$$

The existence of a fundamental solution $\Gamma_{\Omega,\widetilde{\varrho}}$ to the equation $\mathscr{L}_{\Omega,\widetilde{\varrho}}u=0$ has been proved by Di Francesco and Pascucci in [47] via Levi's parametrix method (see Theorem 5.0.3 for the full statement). To present Theorem 5.0.3 in our setting (which include some bounds for $\Gamma_{\Omega,\widetilde{\varrho}}$ that will be useful in the proof of Theorem 6.1.3), we first recall some additional notation and known results concerning the degenerate Kolmogorov operator (6.2.2).
The expression $\Big(x-E(t-t_0)x_0,t-t_0\Big)$ appearing in (6.2.3) is related to an invariance property of the differential operator $\mathscr{L}^{\varepsilon}$, which is needed to state the Hölder continuity assumption on the coefficient $\Phi_{\widetilde{\varrho}}$ appearing in the operator $\mathscr{L}_{\Omega,\widetilde{\varrho}}$. We first recall that the operator $\mathscr{L}^{\varepsilon}$ is invariant w.r.t. the change of variable $(\omega,\vartheta,t)\mapsto(\omega,\vartheta,t)\circ(\xi,\eta,\tau)$, where

$$\begin{aligned}(\omega,\vartheta,t)\circ(\xi,\eta,\tau):=&\Big((\xi,\eta)+E(\tau)(\omega,\vartheta),t+\tau\Big)\\ =&\Big(\xi+\omega\mathrm{e}^{-\tau},\eta+\vartheta+\omega(1-\mathrm{e}^{-\tau}),t+\tau\Big),\qquad(\omega,\vartheta,t),(\xi,\eta,\tau)\in\mathbb{R}^3.\end{aligned}$$

The set $\Big(\mathbb{R}^3,\circ\Big)$ is a Lie group with identity $(0,0,0)$, and the inverse of a point $(\omega,\vartheta,t)\in\mathbb{R}^3$ is

$$(\omega,\vartheta,t)^{-1}=\Big(-E(-t)(\omega,\vartheta),-t\Big)=\Big(-\omega\mathrm{e}^{t},-\vartheta-\omega(1-\mathrm{e}^{t}),-t\Big).$$

As usual in the regularity theory for subelliptic Kolmogorov operators on Lie groups, we employ the anisotropic norm and the quasi-metric introduced in Definition 5.0.1. In this case we consider the norm

$$\|(\omega,\vartheta,t)\|=|\omega|+|\vartheta|^{\frac{1}{3}}+|t|^{\frac{1}{2}},\qquad\forall(\omega,\vartheta,t)\in\mathbb{R}^3,$$

and accordingly we define the quasi-distance

$$\begin{aligned}d\Big((\omega,\vartheta,t),(\xi,\eta,\tau)\Big):=&\;\|(\xi,\eta,\tau)^{-1}\circ(\omega,\vartheta,t)\|\\ =&\;|\omega-\xi\mathrm{e}^{\tau-t}|+|\vartheta-\eta+\xi(\mathrm{e}^{\tau-t}-1)|^{\frac{1}{3}}+|t-\tau|^{\frac{1}{2}},\\ &\qquad(\omega,\vartheta,t),(\xi,\eta,\tau)\in\mathbb{R}^3.\end{aligned}$$

We are now in position to give the following definition.

Definition 6.2.4 (Hölder continuous function)**.** *Let $\alpha \in (0,1]$, and let U be an open subset of $\mathbb{R}^3$. We say that a function $f : U \longrightarrow \mathbb{R}$ is Hölder continuous with exponent α in U (in short $f \in C^\alpha(U)$) if there exists a positive constant L such that*

$$|f(\omega,\vartheta,t) - f(\xi,\eta,\tau)| \leq L\, d\Big((\omega,\vartheta,t),(\xi,\eta,\tau)\Big)^\alpha, \qquad \forall (\omega,\vartheta,t),(\xi,\eta,\tau) \in U.$$

We are now in position to state Theorem 5.0.3 in the setting of (6.2.1), which provides us with an existence result for a fundamental solution $\Gamma_{\Omega,\widetilde{\varrho}}$ to $\mathscr{L}_{\Omega,\widetilde{\varrho}} u = 0$, as well as many useful properties. In the following, we fix $T_1 > 0$ and consider $\Phi_{\widetilde{\varrho}}$ as a function of the variable (ω,ϑ,t) by letting $\Phi_{\widetilde{\varrho}}(\omega,\vartheta,\Omega,t) := \Phi_{\widetilde{\varrho}}(\vartheta,\Omega,t)$.

Theorem 6.2.5. *Suppose that there exists $c > 0$ such that*

$$\|\Phi_{\widetilde{\varrho}}\|_{L^\infty(\mathbb{R}^2\times(0,T_1))} \leq c,$$

and assume that there exist $\alpha \in (0,1]$ and a positive constant L such that

$$|\Phi_{\widetilde{\varrho}}(\omega_1,\vartheta_1,\Omega,t) - \Phi_{\widetilde{\varrho}}(\omega_2,\vartheta_2,\Omega,t)| \leq L\ \mathrm{d}\Big((\omega_1,\vartheta_1,t),(\omega_2,\vartheta_2,t)\Big)^\alpha,$$

for every $(\omega_1,\vartheta_1),(\omega_2,\vartheta_2) \in \mathbb{R}^2$ and $t \in (0,T_1)$. Then, there exists a fundamental solution $\Gamma_{\Omega,\widetilde{\varrho}}$ to $\mathscr{L}_{\Omega,\widetilde{\varrho}}$ with the following properties:

1. *$\Gamma_{\Omega,\widetilde{\varrho}}(\cdot,\xi,\eta,\tau) \in \mathrm{L}^1_{\mathrm{loc}}\Big(\mathbb{R}^2\times(\tau,T_1)\Big) \cap C\Big(\mathbb{R}^2\times(\tau,T_1)\Big)$ for every $(\xi,\eta,\tau) \in \mathbb{R}^2\times(0,T_1)$;*

2. *$\Gamma_{\Omega,\widetilde{\varrho}}(\cdot,\xi,\eta,\tau)$ is a classical solution to $\mathscr{L}_{\Omega,\widetilde{\varrho}} u = 0$ in $\mathbb{R}^2\times(\tau,T_1)$ for every $(\xi,\eta,\tau) \in \mathbb{R}^2\times(0,T_1)$;*

3. *the following identity holds*

$$\int_{\mathbb{R}^2} \Gamma_{\Omega,\widetilde{\varrho}}(\omega,\vartheta,t,\xi,\eta,\tau)\, d\eta\, d\xi = \mathrm{e}^{(t-\tau)}, \qquad \forall (\omega,\vartheta) \in \mathbb{R}^2, \tau < t;$$

4. *for every $\varepsilon > 0$ and $T \in (0,T_1]$ there exists a constant $\overline{C} = \overline{C}\Big(\varepsilon, T, \|\Phi_{\widetilde{\varrho}}\|_{L^\infty}\Big)$, such that*

$$\begin{aligned} 0 <\ & \Gamma_{\Omega,\widetilde{\varrho}}(\omega,\vartheta,t,\xi,\eta,\tau) \leq \overline{C}\,\Gamma^\varepsilon(\omega,\vartheta,t,\xi,\eta,\tau), \\ & |\partial\Gamma_{\Omega,\widetilde{\varrho}}(\omega,\vartheta,t,\xi,\eta,\tau)| \leq \frac{\overline{C}}{\sqrt{(t-\tau)^i}}\Gamma^\varepsilon(\omega,\vartheta,t,\xi,\eta,\tau), \end{aligned} \tag{6.2.6}$$

where Γ^ε is the function defined in (6.2.3), and ∂ stands for ∂_ω ($i = 1$), $\partial^2_{\omega\omega}$ or Lie derivative Y ($i = 2$), for every $(\omega,\vartheta,t),(\xi,\eta,\tau) \in \mathbb{R}^3$ with $0 < t-\tau < T$.

Let's consider the following Cauchy problem

$$\begin{cases} \mathscr{L}_{\Omega,\widetilde{\varrho}}u(\omega,\vartheta,t) = f(\omega,\vartheta,t), & (\omega,\vartheta,t)\in\mathbb{R}^2\times(0,T_1),\\ u(\omega,\vartheta,T_0) = u_0(\omega,\vartheta), & (\omega,\vartheta)\in\mathbb{R}^2, \end{cases}$$

where $u_0\in C(\mathbb{R}^2)$ *is such that*

$$|u_0(\omega,\vartheta)|\le C_1 \mathrm{e}^{C_1(\omega^2+\vartheta^2)}, \qquad \forall(\omega,\vartheta)\in\mathbb{R}^2,$$

and $f\in C\Big(\mathbb{R}^2\times(0,T_1)\Big)$ *is such that*

$$|f(\omega,\vartheta,t)|\le \frac{C_1\mathrm{e}^{C_1(\omega^2+\vartheta^2)}}{\sqrt{t}}, \qquad \forall(\omega,\vartheta,t)\in\mathbb{R}^2\times(0,T_1),$$

for some positive constant C_1*, and assume that for any compact subset* M *of* $\mathbb{R}^2$ *there exists a positive constant* C*,* $\beta\in(0,1)$ *and* $\delta>0$ *such that*

$$|f(\omega,\vartheta,t)-f(\xi,\eta,t)|\le C\frac{d\Big((\omega,\vartheta,t),(\xi,\eta,t)\Big)^{\beta}}{t^{1-\delta}}, \qquad \forall(\omega,\vartheta),(\xi,\eta)\in M, t\in(0,T_1).$$

Then, there exists $T\in(0,T_1]$*, only depending on growth constant* C_1*, such that the function*

$$u(\omega,\vartheta,t)=$$
$$\int_{\mathbb{R}^2}\Gamma_{\Omega,\widetilde{\varrho}}(\omega,\vartheta,t,\xi,\eta,0)u_0(\xi,\eta)d\eta d\xi-\int_0^t\int_{\mathbb{R}^2}\Gamma_{\Omega,\widetilde{\varrho}}(\omega,\vartheta,t,\xi,\eta,\tau)f(\xi,\eta,\tau)d\eta d\xi d\tau,$$

is solution to the previous Cauchy problem in $(0,T)$*.*
Moreover if u *is a solution to the Cauchy problem with null* f *and* g*, and verifies the following estimate*

$$|u(\omega,\vartheta,t)|\le C\mathrm{e}^{C(\omega^2+\vartheta^2)}, \qquad \forall(\omega,\vartheta,t)\in\mathbb{R}^2\times(0,T),$$

for some $C>0$*, then* $u\equiv 0$*.*

6.3 Proof of Theorem 6.1.3.

In this section we prove existence, uniqueness and regularity of classical solutions to the Cauchy problem (6.1.13) in $S_T=\mathbb{R}^3\times(0,T)$, which we recall here for reader's convenience

$$\begin{cases} \mathscr{L}_{\Omega,\varrho}\varrho = \partial^2_{\omega\omega}\varrho+\partial_\omega\Big[(\omega+\Phi_\varrho(\vartheta,\Omega,t))\varrho\Big]-\omega\partial_\vartheta\varrho-\partial_t\varrho=0, & \text{in } S_T,\\ \varrho(\omega,\vartheta,\Omega,0)=\varrho_0(\omega,\vartheta,\Omega), & \text{in } \mathbb{R}^3, \end{cases}$$

where

$$\Phi_{\varrho}(\vartheta, \Omega, t) = -\Omega - K \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^{2\pi} g(\Omega') \sin(\vartheta' - \vartheta) \varrho(\omega', \vartheta', \Omega', t) d\vartheta' d\omega' d\Omega'.$$

As said in the introduction, we find a solution to the nonlinear problem (6.1.13) as the fixed point of a map that associates to a given function ϱ the solution to the linearized problem (6.1.15), which is obtained by using the fundamental solution $\Gamma_{\Omega,\widetilde{\varrho}}$. We list the assumptions about the initial datum ϱ_0 we will adopt in this iterative argument:

1. $\varrho_0 \in C(\mathbb{R}^3)$;

2. ϱ_0 is strictly positive and verifies (**E**) in Definition 6.1.1;

3. ϱ_0 is 2π periodic with respect to ϑ;

4. for every $\Omega \in \mathbb{R}$ we have

$$\int_{(0,2\pi)\times\mathbb{R}} \varrho_0(\omega, \vartheta, \Omega) d\vartheta d\omega = 1.$$

Moreover we recall that the natural frequency distributions of oscillators $g(\Omega)$ is a probability density, satisfying (6.1.14). We first prove a preliminary result. Here and in the sequel, we set $S_T^* := \mathbb{R}^2 \times (0, T)$, and we fix $\Omega \in \mathbb{R}$. Moreover, we consider the function Γ^ε, defined in (6.2.3), for a fixed $\varepsilon > 0$, as we will rely on (6.2.6).

Lemma 6.3.6. *Consider for $T > 0$ the Cauchy problem* (6.1.15) *in S_T^* with $f = 0$. If $\widetilde{\varrho}$ is continuous and verifies property* (**E**)*, then the function u defined as*

$$u(\omega, \vartheta, \Omega, t) = \int_{\mathbb{R}^2} \Gamma_{\Omega,\widetilde{\varrho}}(\omega, \vartheta, t, \xi, \eta, 0) \varrho_0(\xi, \eta, \Omega) d\eta d\xi, \tag{6.3.1}$$

is a classical solution u to (6.1.15) *with $T_1 = T$ and $u_0 = \varrho_0$. Moreover:*

1. *u verifies property* (**E**) *with constants $\overline{M}, C_{\Omega,T}$; in addition we have*

$$|\partial u(\omega, \vartheta, \Omega, t)| \leq \frac{C_{\Omega,T}}{\sqrt{t^i}} \mathrm{e}^{-\overline{M}\omega^2}, \quad \forall (\omega, \vartheta, t) \in S_T^*, \tag{6.3.2}$$

   where ∂ stands for ∂_ω (and (6.3.2) *holds with $i = 1$), $\partial^2_{\omega\omega}$ or Lie derivative Y (and* (6.3.2) *holds with $i = 2$);*

2. *u is 2π periodic with respect to ϑ;*

3. *for every* $\Omega \in \mathbb{R}$ *and for every* $t \in [0, T)$ *we have*

$$\int_{(0,2\pi)\times\mathbb{R}} u(\omega, \vartheta, \Omega, t) d\vartheta d\omega = 1;$$

4. $u(\omega, \vartheta, \Omega, t) > 0$ *for every* $t \geq 0$.

Proof. As a direct consequence of the property (**E**), we can easily see that the function $\Phi_{\widetilde{\varrho}}$ defined in (6.1.16), considered as a function of the variables (ω, ϑ, t) via $\Phi_{\widetilde{\varrho}}(\omega, \vartheta, \Omega, t) := \Phi_{\widetilde{\varrho}}(\vartheta, \Omega, t)$, is bounded. Moreover, we infer that there exists a positive constant L such that

$$|\Phi_{\widetilde{\varrho}}(\omega_1, \vartheta_1, \Omega, t) - \Phi_{\widetilde{\varrho}}(\omega_2, \vartheta_2, \Omega, t)| \leq L \,\mathrm{d}\Big((\omega_1, \vartheta_1, t), (\omega_2, \vartheta_2, t)\Big), \tag{6.3.3}$$

for every $(\omega_1, \vartheta_1, t), (\omega_2, \vartheta_2, t) \in S_T^*$. Therefore the function u defined in (6.3.1) is a classical solution u to (6.1.15), by Theorem 6.2.5. We next prove that u has the properties listed in the statement of the lemma.

In order to prove that u and all the derivatives ∂u listed in the point *(1.)* verify the property (**E**), we first observe that the identity (6.3.1) can be differentiated under the integral sign, *i.e.*

$$\partial u(\omega, \vartheta, \Omega, t) = \int_{\mathbb{R}^2} \partial \Gamma_{\Omega,\widetilde{\varrho}}(\omega, \vartheta, t, \xi, \eta, 0) \varrho_0(\xi, \eta, \Omega) d\eta d\xi.$$

Then by point *(4.)* of Theorem 6.2.5 we have

$$|\partial u(\omega, \vartheta, \Omega, t)| \leq \frac{\overline{C}_{\Omega,T}}{\sqrt{t^i}} \int_{\mathbb{R}^2} \Gamma^\varepsilon(\omega, \vartheta, t, \xi, \eta, 0) \varrho_0(\xi, \eta, \Omega) d\eta d\xi,$$

for some $\overline{C}_{\Omega,T} > 0$ depending on $\varepsilon, T, \|\Phi_{\widetilde{\varrho}}\|_{L^\infty}$, and hence on Ω. Property (**E**) follows from (6.2.6) and from the exponential decay of ϱ_0. Consider a given $\omega \in \mathbb{R}$, by (6.2.4) we get

$$|\partial u(\omega, \vartheta, \Omega, t)| \leq \frac{\overline{C}_{\Omega,T} C}{\sqrt{t^i}} \int_{\mathbb{R}^2} \Gamma^\varepsilon(\omega, \vartheta, t, \xi, \eta, 0) \mathrm{e}^{-M\xi^2} d\xi d\eta. \tag{6.3.4}$$

If $|\omega| > 1$, we first recall the equation (6.2.5) where $\widetilde{\Gamma}^\varepsilon$ has been defined, then we split the above integral in two parts

$$\begin{aligned}\int_{\mathbb{R}^2} \Gamma^\varepsilon(\omega, \vartheta, t, \xi, \eta, 0) \mathrm{e}^{-M\xi^2} d\xi d\eta &= \int_{-\infty}^{+\infty} \widetilde{\Gamma}^\varepsilon(\omega, t, \xi, 0) \mathrm{e}^{-M\xi^2} d\xi \\ &= \int_{\left\{|\xi| \geq \frac{|\omega|}{2}\right\}} \widetilde{\Gamma}^\varepsilon(\omega, t, \xi, 0) \mathrm{e}^{-M\xi^2} d\xi + \int_{\left\{|\xi| < \frac{|\omega|}{2}\right\}} \widetilde{\Gamma}^\varepsilon(\omega, t, \xi, 0) \mathrm{e}^{-M\xi^2} d\xi.\end{aligned}$$

Concerning the first integral, we immediately have

$$\int_{\left\{|\xi| \geq \frac{|\omega|}{2}\right\}} \widetilde{\Gamma}^\varepsilon(\omega, t, \xi, 0) \mathrm{e}^{-M\xi^2} d\xi \leq \mathrm{e}^{-M\frac{\omega^2}{4}} \int_{-\infty}^{+\infty} \widetilde{\Gamma}^\varepsilon(\omega, t, \xi, 0) d\xi = \mathrm{e}^{-M\frac{\omega^2}{4}}. \tag{6.3.5}$$

We use the change of variable $x = \frac{e^{-t}\xi - \omega}{\sqrt{(1+\varepsilon)(1-e^{-2t})}}$ in the second integral. Note that, if $|\xi| < \frac{|\omega|}{2}$, then $|x| > \frac{|\omega|}{2\sqrt{(1+\varepsilon)(1-e^{-2t})}}$, so that

$$\begin{aligned}
&\int_{\left\{|\xi|<\frac{|\omega|}{2}\right\}} \widetilde{\Gamma}^{\varepsilon}(\omega, t, \xi, 0) \mathrm{e}^{-M\xi^2} d\xi \\
&\qquad \le \int_{\left\{|x|>\frac{|\omega|}{2\sqrt{(1+\varepsilon)(1-e^{-2t})}}\right\}} \frac{\mathrm{e}^{-x^2}}{\sqrt{\pi}} \mathrm{e}^{-M\mathrm{e}^{2t}\left(x\sqrt{M(1+\varepsilon)(1-e^{-2t})}+\omega\right)^2} dx \\
&\qquad\qquad \le \frac{\mathrm{e}^{-\frac{\omega^2}{4(1+\varepsilon)(1-e^{-2t})}}}{M\sqrt{(1+\varepsilon)(1-e^{-2t})}} \\
&\qquad\qquad\qquad < \frac{\mathrm{e}^{-\frac{1}{8(1+\varepsilon)(1-e^{-2t})}}}{M\sqrt{(1+\varepsilon)(1-e^{-2t})}} \cdot \mathrm{e}^{-\frac{\omega^2}{8(1+\varepsilon)}} \\
&\qquad\qquad\qquad\qquad \le \widetilde{C}_{\varepsilon,T} \cdot \mathrm{e}^{-\frac{\omega^2}{8(1+\varepsilon)}},
\end{aligned}$$

where we have used the assumption that $|\omega| > 1$ and the fact that the function appearing in the last line of the above display is bounded in its domain $\mathbb{R}^+$. Then (6.3.2) in the case $|\omega| > 1$ follows from the above inequality and (6.3.5), if we set

$$\overline{M} = \min\left\{\frac{M}{4}, \frac{1}{8(1+\varepsilon)}\right\}.$$

If $|\omega| \le 1$ by (6.3.4) we have

$$|\partial u(\omega, \vartheta, \Omega, t)| \le \frac{\overline{C}_{\Omega,T}\, C\, \mathrm{e}^{\overline{M}}}{\sqrt{t^i}} \mathrm{e}^{-\overline{M}\omega^2},$$

then the thesis follows once we set

$$C_{\Omega,T} := \overline{C}_{\Omega,T}\, C \max\left\{2, 2\,\widetilde{C}_{\varepsilon,T}, \mathrm{e}^{\overline{M}}\right\}. \tag{6.3.6}$$

The proof of the periodicity stated in *(2.)* follows from a standard uniqueness argument. We omit the details.

In order to prove the assertion *(3.)* of the lemma, we first notice that the continuity of the functions $\partial_\omega u$ and $Yu = \omega\partial_\omega u - \omega\partial_\vartheta u - \partial_t u$ implies the continuity of the Lie derivative $\widetilde{Y}u := Yu - \omega\partial_\omega u = -\omega\partial_\vartheta u - \partial_t u$. Let us define for every $R > 0$ and $0 < t_0 < t_1$ the sets

$$\begin{aligned}
C_R &= \left\{(\omega, \vartheta, t) \in \mathbb{R}^3 : -R \le \omega \le R,\ 0 \le \vartheta \le 2\pi,\ t_0 \le t \le t_1\right\}, \\
B_R &= \left\{(\omega, \vartheta) \in \mathbb{R}^2 : -R \le \omega \le R,\ 0 \le \vartheta \le 2\pi\right\}.
\end{aligned}$$

If we integrate in C_R the equation in (6.1.15) we obtain

$$\int_{C_R} -\widetilde{Y} u d\omega d\vartheta dt = \int_{C_R} \partial_\omega \Big(\partial_\omega u + \big(\omega + \Phi_{\widetilde{\varrho}}(\vartheta, \Omega, t)\big) u \Big) d\omega d\vartheta dt.$$

Then by the divergence theorem we get

$$\begin{aligned}\int_{B_R} u(\omega, \vartheta, \Omega, t_1) d\omega d\vartheta - \int_{B_R} u(\omega, \vartheta, \Omega, t_0) d\omega d\vartheta \\ + \int_{t_0}^{t_1} \int_{-R}^{R} \omega u(\omega, 0, \Omega, t) d\omega dt - \int_{t_0}^{t_1} \int_{-R}^{R} \omega u(\omega, 2\pi, \Omega, t) d\omega dt \\ = \int_{C_R} \partial_\omega \Big(\partial_\omega u + \big(\omega + \Phi_{\widetilde{\varrho}}(\vartheta, \Omega, t)\big) u \Big) d\omega d\vartheta dt.\end{aligned} \tag{6.3.7}$$

The last two integrals in the left-hand side sum to zero by periodicity, moreover by the exponential decay in ω we have

$$\lim_{R \to \infty} \int_{C_R} \partial_\omega \Big(\partial_\omega u + \big(\omega + \Phi_{\widetilde{\varrho}}(\vartheta, \Omega, t)\big) u \Big) d\omega d\vartheta dt = 0.$$

Therefore, for every $0 \le t_0 < t_1$ we have

$$\int_{(0,2\pi)\times\mathbb{R}} u(\omega, \vartheta, \Omega, t_0) d\omega d\vartheta = \int_{(0,2\pi)\times\mathbb{R}} u(\omega, \vartheta, \Omega, t_1) d\omega d\vartheta.$$

The thesis follows then from the property of ϱ_0 by choosing $t_0 = 0$.
The proof of the point *(4.)* is a consequence of the representation formula (6.3.1), since both $\Gamma_{\Omega,\widetilde{\varrho}}$ and ϱ_0 are positive functions. $\square$

We are now in position to establish an existence result for the nonlinear Cauchy problem (6.1.13).

Proposition 6.3.7. *For every $\Omega, T \in \mathbb{R}$ with $T > 0$, and ϱ_0 satisfying the property* (**E**)*, there exists a classical solution $\varrho = \varrho(\omega, \vartheta, \Omega, t)$ to* (6.1.13) *that verifies property* (**E**) *in the set $S_T^* = \mathbb{R}^2 \times (0, T)$.*
Moreover ϱ can be represented by the fundamental solution $\Gamma_{\Omega,\varrho}$ of $\mathscr{L}_{\Omega,\varrho}$ as follows

$$\varrho(\omega, \vartheta, \Omega, t) = \int_{\mathbb{R}^2} \Gamma_{\Omega,\varrho}(\omega, \vartheta, t, \xi, \eta, 0) \varrho_0(\xi, \eta, \Omega) d\eta d\xi. \tag{6.3.8}$$

Proof. We define a sequence of function $\big\{\varrho_n\big\}_{n\in\mathbb{N}\cup\{0\}}$ as follows. We set $\varrho_0(\omega, \vartheta, \Omega, t) := \varrho_0(\omega, \vartheta, \Omega)$ for every $(\omega, \vartheta, t) \in S_T^*$. Then, for every $n \in \mathbb{N}$, we let

$$\Phi_{\varrho_{n-1}}(\vartheta, \Omega, t) = -\Omega - K \int_{\mathbb{R}} \int_{\mathbb{R}} \int_0^{2\pi} g(\Omega') \sin(\vartheta' - \vartheta) \varrho_{n-1}(\omega', \vartheta', \Omega', t) d\vartheta' d\omega' d\Omega',$$

and we define ϱ_n as the solution to the linear Cauchy problem

$$\begin{cases} \mathscr{L}_{\Omega,\varrho_{n-1}}\varrho_n = \partial^2_{\omega\omega}\varrho_n + \partial_\omega\Big[\Big(\omega + \Phi_{\varrho_{n-1}}(\vartheta,\Omega,t)\Big)\varrho_n\Big] - \omega\partial_\vartheta\varrho_n - \partial_t\varrho_n = 0, \\ \varrho_n(\omega,\vartheta,\Omega,0) = \varrho_0(\omega,\vartheta,\Omega). \end{cases} \tag{6.3.9}$$

The solution to the nonlinear Cauchy problem will be defined as the limit of the sequence $\left\{\varrho_n\right\}_{n\in\mathbb{N}\cup\{0\}}$. We rely on a compactness argument to prove the convergence of a subsequence of $\left\{\varrho_n\right\}_{n\in\mathbb{N}\cup\{0\}}$. However, due the presence of the term ϱ_{n-1} in the coefficient $\Phi_{\varrho_{n-1}}$ in (6.3.9), we need to show that the whole sequence converges.

We now prove that the sequence $\left\{\varrho_n\right\}_{n\in\mathbb{N}\cup\{0\}}$ has the property (**E**), with constants that does not depend on n.
As $\varrho_0 \in C(\mathbb{R}^3)$ and verifies property (**E**), we can apply Lemma 6.3.6, inferring that for every $n \in \mathbb{N}$ there exists a classical solution ϱ_n in $\mathbb{R}^2 \times [0,T)$ such that:

- ϱ_n is 2π periodic in ϑ;
- ϱ_n is strictly positive for every $t \geq 0$;
- ϱ_n is normalized

$$\int_{(0,2\pi)\times\mathbb{R}} \varrho_n(\omega,\vartheta,\Omega,t)d\vartheta d\omega = 1, \qquad \forall\, t \in [0,T). \tag{6.3.10}$$

We recall that $\Gamma_{\Omega,\varrho_{n-1}}$ denotes the fundamental solution of the operator $\mathscr{L}_{\Omega,\varrho_{n-1}}$ appearing in (6.3.9). The assertion *(4.)* of Theorem 6.2.5 yields

$$\begin{aligned} \Gamma_{\Omega,\varrho_{n-1}}(\omega,\vartheta,t,\xi,\eta,0) &\leq \overline{C}_{\Omega,T}\Gamma^\varepsilon(\omega,\vartheta,t,\xi,\eta,0), \\ |\partial\Gamma_{\Omega,\varrho_{n-1}}(\omega,\vartheta,t,\xi,\eta,0)| &\leq \frac{\overline{C}_{\Omega,T}}{\sqrt{t^i}}\Gamma^\varepsilon(\omega,\vartheta,t,\xi,\eta,0), \end{aligned} \tag{6.3.11}$$

where the constant $\overline{C}_{\Omega,T} > 0$ only depends on ε, T, and on $\|\Phi_{\varrho_{n-1}}\|_{L^\infty}$. Moreover, also using (6.3.3) and (6.3.10), we see that

$$\begin{aligned} \|\Phi_{\varrho_{n-1}}\|_{L^\infty} &\leq |\Omega| + K, \\ |\Phi_{\varrho_{n-1}}(\omega_1,\vartheta_1,\Omega,t) - \Phi_{\varrho_{n-1}}(\omega_2,\vartheta_2,\Omega,t)| &\leq L\,\mathrm{d}\Big((\omega_1,\vartheta_1,t),(\omega_2,\vartheta_2,t)\Big), \end{aligned} \tag{6.3.12}$$

for every $(\omega_1,\vartheta_1,t),(\omega_2,\vartheta_2,t) \in S^*_T$, and for every $n \in \mathbb{N}$. Note that, in particular, the constant $\overline{C}_{\Omega,T}$ in (6.3.11) does not depend on n. In the sequel, we will use the fact that, from the boundedness and the Lipschitz continuity of the function $\Phi_{\varrho_{n-1}}$, the β-Hölder continuity of $\Phi_{\varrho_{n-1}}$ directly follows for every $\beta \in (0,1)$.

As a consequence of Lemma 6.3.6, the sequence $\left\{\varrho_n\right\}_{n\in\mathbb{N}\cup\{0\}}$ has an uniform exponential decay in ω, meaning that

$$0 < \varrho_n(\omega,\vartheta,t) \le C_{\Omega,T}\,\mathrm{e}^{-\overline{M}\omega^2}, \qquad \forall(\omega,\vartheta,t)\in\mathbb{R}^2\times(0,T), \quad n\in\mathbb{N}. \tag{6.3.13}$$

Moreover, $\partial^2_\omega\varrho_n$, and $Y\varrho_n$ have an uniform exponential decay in ω as well, with a constant $C_{\Omega,T}/t$. Thus, the sequence $\left\{\varrho_n\right\}_{n\in\mathbb{N}\cup\{0\}}$ is equi-bounded and equi-continuous on every compact subset of $\mathbb{R}^2\times(0,T]$. Hence, there exists a subsequence $\left\{\varrho_{n_k}\right\}_{k\in\mathbb{N}\cup\{0\}}$ and a function $\varrho\in C(S^*_T)$ such that

$$\varrho_{n_k}\to\varrho, \quad \text{as} \quad k\to+\infty,$$

uniformly on every compact subset of $\mathbb{R}^2\times(0,T]$.
As said above, we need to show that the whole sequence $\left\{\varrho_n\right\}_{n\in\mathbb{N}\cup\{0\}}$ converges to ϱ. We will prove a stronger result, namely

$$\|\varrho_n-\varrho\|_{L^\infty(S^*_T)}\to 0, \quad \text{as} \quad n\to+\infty, \tag{6.3.14}$$

where we recall that $S^*_T=\mathbb{R}^2\times(0,T)$. With the aim of proving (6.3.14), we preliminarily note that the constant $\overline{C}_{\Omega,T}$ in (6.3.11) depends on the quantities Ω and T as follows

$$\overline{C}_{\Omega,T}=C_\varepsilon+\frac{\sqrt{T}}{2}\sum_{j=1}^{+\infty}(|\Omega|+K)^j(1+\sqrt{T})^j\frac{\sqrt{\pi}}{\Gamma_E(\frac{j}{2})}, \tag{6.3.15}$$

where $C_\varepsilon>0$ depends only on ε, which is fixed, and Γ_E denotes the Euler gamma function. We obtain the above expression by repeating the computations made in [47] in the particular case of the operator $\mathscr{L}$. Then, by recalling the definition (6.3.6) of $C_{\Omega,T}$, the hypothesis (6.1.14) yields

$$G:=\int_{\mathbb{R}}g(\Omega)C_{\Omega,T}\,d\Omega<+\infty. \tag{6.3.16}$$

Let us now fix a positive t_0, which needs to be chosen small enough to have

$$q:=2K\pi^{\frac{3}{2}}\overline{C}_{\Omega,T}G\sqrt{\frac{t_0}{A}}<1,$$

where we have set

$$A:=\frac{\overline{M}}{1+2(1+\varepsilon)\overline{M}(\mathrm{e}^{2T}-1)}.$$

Here $\overline{M}$ is the constants in the exponential decay. Then, we define $\widehat{k}=\lfloor T/t_0\rfloor+1$ and we set $T_k=kt_0$ for $k=0,\dots,\widehat{k}-1$, and $T_{\widehat{k}}=T$. For every $n\in\mathbb{N}$ we consider the function

$$v_n(\omega,\vartheta,\Omega,t):=\varrho_{n+1}(\omega,\vartheta,\Omega,t)-\varrho_n(\omega,\vartheta,\Omega,t).$$

This is the unique bounded solution to the following Cauchy problem

$$\begin{cases} \mathscr{L}_{\Omega,\varrho_n} v_n = (\Phi_{\varrho_n} - \Phi_{\varrho_{n-1}})\partial_\omega \varrho_n, & \text{in } \mathbb{R}^2 \times (T_k, T_{k+1}), \\ v_n(\omega,\vartheta,\Omega,T_k) = \varrho_{n+1}(\omega,\vartheta,\Omega,T_k) - \varrho_n(\omega,\vartheta,\Omega,T_k), & \text{in } \mathbb{R}^2. \end{cases}$$

Now we represent the function v_n by using the fundamental solution $\Gamma_{\Omega,\varrho_n}$ of $\mathscr{L}_{\Omega,\varrho_n}$. The above PDE can be written as $\mathscr{L}_{\Omega,\varrho_n} v_n = f_n$, with

$$f_n(\omega,\vartheta,\Omega,t) := \Big(\Phi_{\varrho_n}(\vartheta,\Omega,t) - \Phi_{\varrho_{n-1}}(\vartheta,\Omega,t)\Big)\partial_\omega \varrho_n(\omega,\vartheta,\Omega,t).$$

In order to apply Theorem 6.2.5, we claim that there exist three positive constants C_1,γ and β, with $\beta,\gamma<1$, such that

$$\begin{aligned} |f_n(\omega,\vartheta,\Omega,t)| &\le \frac{C_1}{\sqrt{t}}, \\ |f_n(\omega_1,\vartheta_1,\Omega,t) - f_n(\omega_2,\vartheta_2,\Omega,t)| &\le \frac{C_1}{t^{1-\gamma}}\, \mathrm{d}\Big((\omega_1,\vartheta_1,t),(\omega_2,\vartheta_2,t)\Big)^\beta, \end{aligned} \tag{6.3.17}$$

for every $(\omega,\vartheta,t),(\omega_1,\vartheta_1,t),(\omega_2,\vartheta_2,t) \in S_T^*$, and for every $n\in\mathbb{N}$. The first inequality in (6.3.17) directly follows from the first inequality in (6.3.12) and from the uniform exponential decay of $\partial_\omega\varrho_n$. In order to prove the second inequality, we claim that there exist two positive constants C and β, with $\beta<1$, such that

$$|\partial_\omega\varrho_n(\omega_1,\vartheta_1,\Omega,t) - \partial_\omega\varrho_n(\omega_2,\vartheta_2,\Omega,t)| \le \frac{C}{t^{\frac{1+\beta}{2}}}\, \mathrm{d}\Big((\omega_1,\vartheta_1,t),(\omega_2,\vartheta_2,t)\Big)^\beta, \tag{6.3.18}$$

for every $(\omega_1,\vartheta_1,t),(\omega_2,\vartheta_2,t) \in S_T^*$, and for every $n\in\mathbb{N}$. The conclusion of the proof of (6.3.12) plainly follows from the exponential decay of $\partial_\omega\varrho_n$ combined with (6.3.12) and (6.3.18).
The estimate (6.3.18) is a direct consequence of the identity

$$\partial\varrho_n(\omega,\vartheta,\Omega,t) = \int_{\mathbb{R}^2} \partial_\omega\Gamma_{\Omega,\varrho_n}(\omega,\vartheta,t,\xi,\eta,0)\varrho_0(\xi,\eta,\Omega)d\eta d\xi.$$

Then we use the following property the fundamental solution, proved in Theorem 3.2 of [48]: for every $\beta\in(0,1)$ there exists a positive constant c_β such that

$$\Big|\partial_\omega\Gamma_{\Omega,\varrho_n}(\omega_1,\vartheta_1,t,\xi,\eta,0) - \partial_\omega\Gamma_{\Omega,\varrho_n}(\omega_2,\vartheta_2,t,\xi,\eta,0)\Big| \le \\ c_\beta \frac{\mathrm{d}\Big((\omega_1,\vartheta_1,t),(\omega_2,\vartheta_2,t)\Big)^\beta}{t^{\frac{1+\beta}{2}}}\Gamma^\varepsilon(\omega_1,\vartheta_1,t,\xi,\eta,0),$$

for every $t\in(0,T),(\omega_1,\vartheta_1),(\omega_2,\vartheta_2),(\xi,\eta)\in\mathbb{R}^2$, being Γ^ε the function defined in (6.2.2).

Finally, by using the exponential decay of the initial datum and (6.2.4) we find

$$\begin{aligned}\Big|\partial_\omega \varrho_n(\omega_1,\vartheta_1,\Omega,t) - \partial_\omega \varrho_n(\omega_2,\vartheta_2,\Omega,t)\Big| \leq \\ c_\beta \frac{\mathrm{d}\Big((\omega_1,\vartheta_1,t),(\omega_2,\vartheta_2,t)\Big)^\beta}{t^{\frac{1+\beta}{2}}} \int_{\mathbb{R}^2} \Gamma^\varepsilon(\omega_1,\vartheta_1,t,\xi,\eta,0)\varrho_0(\xi,\eta)d\xi d\eta \leq \\ \overline{c}_\beta \frac{\mathrm{d}\Big((\omega_1,\vartheta_1,t),(\omega_2,\vartheta_2,t)\Big)^\beta}{t^{\frac{1+\beta}{2}}},\end{aligned}$$

for some $\overline{c}_\beta > 0$. This concludes the proof of (6.3.17) with β arbitrarily chosen in $(0,1)$ and $\gamma = \frac{1-\beta}{2}$. Once the bounds (6.3.17) have been proved, we represent v_n as follows

$$\begin{aligned}v_n(\omega,\vartheta,\Omega,t) = \int_{\mathbb{R}^2} \Gamma_{\Omega,\varrho_n}(\omega,\vartheta,t,\xi,\eta,T_k) v_n(\xi,\eta,\Omega,T_k)\, d\eta d\xi \\ - \int_{T_k}^t \int_{\mathbb{R}^2} \Gamma_{\Omega,\varrho_n}(\omega,\vartheta,t,\xi,\eta,\tau) f_n(\xi,\eta,\Omega,\tau)\, d\eta d\xi d\tau.\end{aligned} \tag{6.3.19}$$

In order to simplify the remaining part of the proof, we define recursively the constants

$$M_1 := \frac{\overline{M}}{1+\overline{M}(1+\varepsilon)(\mathrm{e}^{2T_1}-1)}, \qquad M_{k+1} := \frac{M_k}{1+M_k(1+\varepsilon)(\mathrm{e}^{2(T_{k+1}-T_k)}-1)}, \quad k = 1,\dots,\widehat{k}-1,$$

and we note that, by a plain induction argument, we find

$$\begin{aligned}M_k &= \frac{\overline{M}}{1+\overline{M}(1+\varepsilon)k(\mathrm{e}^{2t_0}-1)}, \\ M_{\hat{k}} &= \frac{\overline{M}}{1+\overline{M}(1+\varepsilon)((\hat{k}-1)(\mathrm{e}^{2t_0}-1)+(\mathrm{e}^{2(T-T_{\hat{k}-1})}-1))}.\end{aligned}$$

In particular, the inequalities below hold

$$\overline{M} > M_k > M_{k+1} \geq A, \qquad k = 1,\dots,\widehat{k}-1, \tag{6.3.20}$$

since

$$\begin{aligned}(\hat{k}-1)(\mathrm{e}^{2t_0}-1) + (\mathrm{e}^{2(T-T_{\hat{k}-1})}-1) \\ \leq \hat{k}(\mathrm{e}^{2t_0}-1) \leq \left(\frac{T}{t_0}+1\right)(\mathrm{e}^{2t_0}-1) \leq 2(\mathrm{e}^{2T}-1).\end{aligned}$$

We now introduce the weighted norms

$$\|\varrho_{n+1}-\varrho_n\|_k := \sup_{S_T^*\cap(T_{k-1},T_k]} |\varrho_{n+1}(\omega,\vartheta,\Omega,t) - \varrho_n(\omega,\vartheta,\Omega,t)| \frac{\mathrm{e}^{M_k\omega^2}}{C_{\Omega,T}}, \qquad k = 1,\dots,\widehat{k}. \tag{6.3.21}$$

From (6.3.20) and (6.3.13) it follows that

$$\|\varrho_{n+1} - \varrho_n\|_k \leq 2, \qquad k = 1, \dots, \widehat{k}.$$

The term $\frac{\mathrm{e}^{M_k \omega^2}}{C_{\Omega,T}}$ appearing in the norm introduced in (6.3.21) is useful in order to deal with the term f_n in (6.3.19). Indeed, later on we will rely on the following bound

$$\begin{aligned}
&|\Phi_{\varrho_n}(\vartheta,\Omega,t) - \Phi_{\varrho_{n-1}}(\vartheta,\Omega,t)| \\
&\leq K \int_{\mathbb{R}} g(\Omega) \int_{\mathbb{R}} \int_0^{2\pi} |\sin(\vartheta-\vartheta')(\varrho_n(\omega',\vartheta',\Omega,t) - \varrho_{n-1}(\omega',\vartheta',\Omega,t))| d\vartheta' d\omega' d\Omega \\
&\qquad \leq K \|\varrho_n - \varrho_{n-1}\|_{k+1} \int_{\mathbb{R}} \int_0^{2\pi} \mathrm{e}^{-M_{k+1}\omega'^2} d\vartheta' d\omega' \int_{\mathbb{R}} g(\Omega) C_{\Omega,T} d\Omega \\
&\qquad = \frac{2\pi^{\frac{3}{2}} K G}{\sqrt{M_{k+1}}} \|\varrho_n - \varrho_{n-1}\|_{k+1},
\end{aligned} \tag{6.3.22}$$

for every $\vartheta \in \mathbb{R}$ and $t \in (T_k, T_{k+1})$, which will be used to estimate $|f_n|$.
In order to prove (6.3.14) we claim that the following estimates

$$\begin{aligned}
\|\varrho_{n+1} - \varrho_n\|_1 &\leq q\, \|\varrho_n - \varrho_{n-1}\|_1, \\
\|\varrho_{n+1} - \varrho_n\|_{k+1} &\leq q\, \|\varrho_n - \varrho_{n-1}\|_{k+1} + \overline{C}_{\Omega,T} \|\varrho_{n+1} - \varrho_n\|_k,
\end{aligned} \tag{6.3.23}$$

hold for every $n = 2, 3, ..$ and $k = 1, 2, \dots, \widehat{k} - 1$. To prove the claim we denote by I_1 and I_2 the two integral in the right-hand side of (6.3.19). By (6.2.5) and (6.2.6) a direct computation give us

$$\begin{aligned}
|I_1| &\leq \overline{C}_{\Omega,T} \|\varrho_{n+1} - \varrho_n\|_k \, C_{\Omega,T} \int_{\mathbb{R}} \widetilde{\Gamma}^{\varepsilon}(\omega,t,\xi,T_k) \mathrm{e}^{-M_k \xi^2} d\xi \\
&\leq \overline{C}_{\Omega,T} \|\varrho_{n+1} - \varrho_n\|_k \, C_{\Omega,T} \exp\left(-\omega^2 \frac{M_k}{1 + M_k(1+\varepsilon)(\mathrm{e}^{2(T_{k+1}-T_k)} - 1)} \right) \\
&= \overline{C}_{\Omega,T} \|\varrho_{n+1} - \varrho_n\|_k \, C_{\Omega,T} \mathrm{e}^{-M_{k+1}\omega^2}.
\end{aligned}$$

In order to estimate I_2, we use again (6.2.5) and (6.2.6), together with (6.3.22) and the exponential decay of $\partial_\omega \varrho_n$

$$\begin{aligned}
|I_2| &\leq \frac{2\pi^{\frac{3}{2}} K G \overline{C}_{\Omega,T}}{\sqrt{M_{k+1}}} \|\varrho_n - \varrho_{n-1}\|_{k+1} C_{\Omega,T} \int_{T_k}^t \int_{\mathbb{R}} \frac{\widetilde{\Gamma}^{\varepsilon}(\omega,t,\xi,\tau) \mathrm{e}^{-\overline{M}\xi^2}}{\sqrt{\tau}} d\xi d\tau \\
&\leq 2\pi^{\frac{3}{2}} K G \overline{C}_{\Omega,T} \sqrt{\frac{t_0}{A}} \, \|\varrho_n - \varrho_{n-1}\|_{k+1} C_{\Omega,T} \mathrm{e}^{-M_{k+1}\omega^2} \\
&= q \|\varrho_n - \varrho_{n-1}\|_{k+1} C_{\Omega,T} \mathrm{e}^{-M_{k+1}\omega^2},
\end{aligned}$$

where we have used the inequalities in (6.3.20). This concludes the proof of (6.3.23).

We next recall that Lemma 3.4 in [77] states that (6.3.23) imply the inequalities

$$\|\varrho_{n+1}-\varrho_n\|_k \le q^{n-2}\sum_{i=0}^{k-1}\Big(\overline{C}_{\Omega,T}(n-2)\Big)^i\|\varrho_3-\varrho_2\|_{k-i},$$

for all $n=3,4,\dots$ and $k=1,\dots,\widehat{k}$. We then find

$$\|\varrho_{n+1}-\varrho_n\|_{C(\overline{S_T^*})} \le \sum_{k=1}^{\widehat{k}}\|\varrho_{n+1}-\varrho_n\|_k \le \sum_{k=1}^{\widehat{k}} q^{n-2}\sum_{i=0}^{k-1}\Big(\overline{C}_{\Omega,T}(n-2)\Big)^i\|\varrho_3-\varrho_2\|_{k-i}.$$

By using the elementary inequality $\Big(\overline{C}_{\Omega,T}(n-2)\Big)^i \le (n-2)^{\widehat{k}}\Big(1+\overline{C}_{\Omega,T}\Big)^{\widehat{k}}$ and

$$\sum_{k=1}^{\widehat{k}}\sum_{i=0}^{k-1}\|\varrho_3-\varrho_2\|_{k-i} = \sum_{k=1}^{\widehat{k}}\sum_{j=1}^{k}\|\varrho_3-\varrho_2\|_j \le \sum_{k=1}^{\widehat{k}}\sum_{j=1}^{\widehat{k}}\|\varrho_3-\varrho_2\|_j = \widehat{k}\sum_{j=1}^{\widehat{k}}\|\varrho_3-\varrho_2\|_j,$$

we finally have

$$\|\varrho_{n+1}-\varrho_n\|_{C(\overline{S_T^*})} \le q^{n-2}(n-2)^{\widehat{k}}\widehat{k}\Big(1+\overline{C}_{\Omega,T}\Big)^{\widehat{k}}\sum_{j=1}^{\widehat{k}}\|\varrho_3-\varrho_2\|_j.$$

Then, we set

$$V := \widehat{k}\Big(1+\overline{C}_{\Omega,T}\Big)^{\widehat{k}}\sum_{j=1}^{\widehat{k}}\|\varrho_3-\varrho_2\|_j,$$

and we notice that V does not depend on n. Hence, we eventually obtain

$$\sum_{n=3}^{+\infty}\|\varrho_{n+1}-\varrho_n\|_{C(\overline{S_T^*})} \le V\sum_{n=3}^{+\infty}q^{n-2}(n-2)^{\widehat{k}} = V\sum_{n=1}^{+\infty}q^n n^{\widehat{k}} < \infty.$$

This accomplishes the proof of (6.3.14).
From (6.3.14) and (6.3.13) it follows that ϱ satisfies the identity (6.3.8), it has property (**E**) in the set S_T^*, and

$$\varrho(\omega,\vartheta,\Omega,0)=\varrho_0(\omega,\vartheta,\Omega), \qquad \text{for every } (\omega,\vartheta)\in\mathbb{R}^2.$$

Moreover, ϱ is a classical solution to $\mathscr{L}_{\Omega,\varrho}\varrho=0$ in S_T^*, thanks to the bounds (6.3.2) in Lemma 6.3.6, that hold for ϱ_n, uniformly with respect to n. This completes the proof of Proposition 6.3.7. □

We now prove that, under the additional assumption that g has compact support, there is only one solution satisfying the property (**E**).

Proposition 6.3.8. *Suppose that the function g has support in the set $[-\Omega_1, \Omega_1]$, and let ϱ_1 and ϱ_2 be two classical solutions to the Cauchy problem* (6.1.13). *If ϱ_1 and ϱ_2 both satisfy property* (**E**)*, then $\varrho_1 = \varrho_2$.*

Proof. Suppose that ϱ_1 and ϱ_2 are two solutions to the Cauchy problem (6.1.13), each satisfying property (**E**) with the associated constants $\overline{M}$ and $C_{\Omega,T}$. We introduce the function $\overline{\varrho} := \varrho_1 - \varrho_2$, and we notice that, under these assumptions, the following quantities are finite for every $\tau, t \in [0, T)$

$$\widetilde{\varphi}(\tau) := \int_{\mathbb{R}} \max_{\vartheta\in\mathbb{R},\Omega\in[-\Omega_1,\Omega_1]} \left|\frac{\overline{\varrho}(\omega,\vartheta,\Omega,\tau)}{C_{\Omega,T}}\right| d\omega, \qquad \varphi(t) := \sup_{0\le\tau\le t} \widetilde{\varphi}(\tau). \tag{6.3.24}$$

Moreover, $\overline{\varrho}$ is a classical solution to

$$\begin{cases} \partial^2_{\omega\omega}\overline{\varrho} + \omega\partial_\omega\overline{\varrho} - \omega\partial_\vartheta\overline{\varrho} + \overline{\varrho} - \partial_t\overline{\varrho} = -\Big((\Phi_{\varrho_2} - \Phi_{\varrho_1})\partial_\omega\varrho_2 - \Phi_{\varrho_1}\partial_\omega(\varrho_2 - \varrho_1)\Big), \\ \overline{\varrho}(\omega,\vartheta,\Omega,0) = 0. \end{cases}$$

Arguing as in the proof of Theorem 6.1.3, we see that $\overline{\varrho}$ can be represented as

$$\begin{aligned} &\overline{\varrho}(\omega,\vartheta,\Omega,t) \\ &\quad = \int_0^t \int_{\mathbb{R}^2} \overline{\Gamma}(\omega,\vartheta,t,\xi,\eta,\tau)\Big((\Phi_{\varrho_2} - \Phi_{\varrho_1})\partial_\xi\varrho_2 - \Phi_{\varrho_1}\partial_\xi(\varrho_2 - \varrho_1)\Big)(\xi,\eta,\Omega,\tau)d\xi d\eta d\tau, \end{aligned} \tag{6.3.25}$$

where $\overline{\Gamma}$ is the fundamental solution of

$$\overline{\mathscr{L}} = \partial^2_{\omega\omega} + \omega\partial_\omega - \omega\partial_\vartheta + 1 - \partial_t.$$

Our aim is to obtain an estimate for $\varphi(t)$, by using the representation formula (6.3.25), which will implies $\varphi(t_0) = 0$ for a sufficiently small t_0. We will show that t_0 only depends on the operator $\overline{\mathscr{L}}$ and on the quantities appearing in property (**E**). This implies that this argument can be repeated in the time interval $[t_0, 2t_0]$, then in $[2t_0, 3t_0]$, and so on, hence we conclude that $\overline{\varrho}(\omega,\vartheta,\Omega,t) = 0$ for every $(\omega,\vartheta,\Omega,t) \in \mathbb{R}^3 \times [0,T)$.
For this purpose, we first recall the constant G introduced in (6.3.16), and we note that the following estimate

$$\begin{aligned} &|\Phi_{\varrho_2}(\eta,\Omega,\tau) - \Phi_{\varrho_1}(\eta,\Omega,\tau)| \\ &\qquad K\int_{\mathbb{R}} g(\Omega)\int_{\mathbb{R}}\int_0^{2\pi} |\sin(\eta-\vartheta')(\varrho_2(\omega',\vartheta',\Omega,\tau) - \varrho_1(\omega',\vartheta',\Omega,\tau))|d\vartheta' d\omega' d\Omega \\ &\qquad\qquad \le 2\pi K G\varphi(t), \qquad \eta\in\mathbb{R}, \tau\in[0,t], \end{aligned}$$

plainly follows from the definition of Φ_{ϱ_1} and Φ_{ϱ_2}. Then, from this bound, the exponential decay property of the derivative $\partial_\omega \varrho_2$ and the bound (6.3.12) applied to Φ_{ϱ_1} we find that

$$\begin{aligned}
\int_{\mathbb{R}} |\overline{\varrho}(\omega,\vartheta,\Omega,t)| d\omega \leq \\
& 2\pi K G \varphi(t) \int_0^t \int_{\mathbb{R}^3} \overline{\Gamma}(\omega,\vartheta,t,\xi,\eta,\tau) \frac{C_{\Omega,T}}{\sqrt{\tau}} \mathrm{e}^{-\overline{M}\xi^2} d\eta d\omega d\xi d\tau \\
& \quad + (|\Omega| + K) \int_0^t \int_{\mathbb{R}^3} |\partial_\xi \overline{\Gamma}(\omega,\vartheta,t,\xi,\eta,\tau)| (\varrho_2 - \varrho_1)(\xi,\eta,\Omega,\tau) d\omega d\eta d\xi d\tau \\
& =: I_1 + I_2.
\end{aligned}$$

By a direct computation we find that

$$I_1 \leq C_{\Omega,T} \, \frac{\pi^{\frac{3}{2}} K G}{\sqrt{\overline{M}}} \sqrt{t} \, \varphi(t).$$

Moreover, by (6.2.6), together with definition (6.3.24) and bounds (6.3.6), (6.3.15), we obtain

$$I_2 \leq C_{\Omega,T} \overline{C} \frac{\sqrt{t}}{2} \varphi(t),$$

for some positive constant $\overline{C}$ that depends only on Ω_1. Thus, since the previous estimates gives us

$$\varphi(t) \leq \varphi(t) \left(\frac{\pi^{\frac{3}{2}} K G}{\sqrt{\overline{M}}} \sqrt{t} + \overline{C} \frac{\sqrt{t}}{2} \right),$$

we can state that there exists t_0 such that, if $T \leq t_0$, we have $\varphi(t) \leq c\,\varphi(t)$ for some $c < 1$, and therefore $\varphi(t) = 0$.
Finally, by iterating this arguments in time intervals $[kt_0, (k+1)t_0]$ the thesis follows. □

Proof of Theorem 6.1.3. The proof of the theorem follows from the results we have established in Lemma 6.3.6, Proposition 6.3.7 and Proposition 6.3.8. The continuity dependence of the solution w.r.t. Ω is proved in [48]. □

6.4 Finite difference method

The purpose of this section is to develop a numerical scheme in the spirit of the analytical framework introduced in Section 6.2. We propose therefore a finite difference scheme where we discretize the Lie derivative instead of the classical derivatives w.r.t. the variables ϑ and t. In particular, we are following the pattern described in [46] and [103], where an accurate analysis concerning the discretization of degenerate Kolmogorov operators has been made. The discretization we would like to carry out would make us approximate the Lie derivative

$Y\varrho = \omega\partial_\omega\varrho - \omega\partial_\vartheta\varrho - \partial_t\varrho$ with the following quotient

$$\frac{\varrho(\omega e^\delta, \omega - \omega e^\delta - \vartheta, t-\delta) - \varrho(\omega,\vartheta,\Omega,t)}{\delta}.$$

The main difficulty when setting an appropriate grid is represented by the factor e^δ, which arises from the term $\omega\partial_\omega$ in the Lie derivative. This problem can be overcome by observing that, thanks to the continuity w.r.t. the derivative ∂_ω, the solution is also continuous w.r.t. the Lie derivative $Y_r = -\omega\partial_\vartheta - \partial_t$, whose discretization is easier.
In our analysis we assume that the natural frequency distribution g is compactly supported in $[-\Omega_1, \Omega_1]$, so that the solution we approximate is unique, according to Proposition 6.3.8.
We list here the various approximations that we use for the numerical scheme

$$\begin{aligned}
\partial^2_{\omega\omega}\varrho(\omega,\vartheta,\Omega,t) &= \frac{\varrho(\omega+\Delta\omega,\vartheta,\Omega,t) - 2\varrho(\omega,\vartheta,\Omega,t) + \varrho(\omega-\Delta\omega,\vartheta,\Omega,t)}{(\Delta\omega)^2} + o(\Delta\omega)^2,\\
\partial_\omega\varrho(\omega,\vartheta,\Omega,t) &= \frac{\varrho(\omega+\Delta\omega,\vartheta,\Omega,t) - \varrho(\omega-\Delta\omega,\vartheta,\Omega,t)}{2(\Delta\omega)} + o(\Delta\omega)^2,\\
Y_r\varrho(\omega,\vartheta,\Omega,t) &= \frac{\varrho(\omega,\vartheta-\omega\Delta t,\Omega,t-\Delta t) - \varrho(\omega,\vartheta,\Omega,t)}{\Delta t} + o(\Delta t).
\end{aligned}$$

Finally, we approximate the nonlinear term by

$$-\Omega - K\int_{-\Omega_1}^{\Omega_1}\int_0^{2\pi}\int_{-G_\omega}^{G_\omega} \varrho(\omega',\vartheta',\Omega,t-\Delta t)\sin(\vartheta-\vartheta')g(\Omega')d\omega' d\vartheta' d\Omega' + o(\Delta t).$$

This approximation introduces the constant G_ω to overcome the problem of an unlimited interval in ω. We observe that, thanks to the exponential decay, we can make this last approximation truly negligible. Finally, we approximate this integral by the method of rectangles, and we denote as $\widetilde{\Phi}_\varrho$ this last approximation. As in the classical case, the various remainders depend on the L^∞ norm of the solution and of some of its derivatives.

In the following, in order to provide a comprehensive analysis of the model, we will include the cases where the inertial term m and the noise term D are not necessarily equal to one, in order to investigate the changes in the solution as these parameters vary.
We define the approximate operator $\mathscr{L}_a$ as

$$\begin{aligned}
&\mathscr{L}_a\varrho(\omega,\vartheta,\Omega,t) :=\\
&=\frac{D}{m^2}\frac{\varrho(\omega+\Delta\omega,\vartheta-\omega\Delta t,\Omega,t) - 2\varrho(\omega,\vartheta-\omega\Delta t,\Omega,t) + \varrho(\omega-\Delta\omega,\vartheta-\omega\Delta t,\Omega,t)}{(\Delta\omega)^2}\\
&+\frac{1}{m}\left(\omega + \widetilde{\Phi}_\varrho(\vartheta-\omega\Delta t,\Omega,t)\right)\frac{\varrho(\omega+\Delta\omega,\vartheta-\omega\Delta t,,\Omega,t) - \varrho(\omega-\Delta\omega,\vartheta-\omega\Delta t,\Omega,t)}{2(\Delta\omega)}\\
&+\frac{\varrho(\omega,\vartheta-\omega\Delta t,\Omega,t) - \varrho(\omega,\vartheta,\Omega,t+\Delta t)}{\Delta t} + \frac{1}{m}\varrho(\omega,\vartheta-\omega\Delta t,\Omega,t).
\end{aligned}$$

The previous approximations make $\mathscr{L}_a$ consistent with $\mathscr{L}_\varrho$. Indeed, if $G_\omega = +\infty$ we have

$$\mathscr{L}_\varrho \varrho(\omega, \vartheta, \Omega, t) - \mathscr{L}_a \varrho(\omega, \vartheta, \Omega, t) = o\Big((\Delta\omega)^2 + \Delta t\Big).$$

In order to define the corresponding grid, let us examine the domain in which we solve the problem: as anticipated above, we suppose that ω belongs to the bounded interval $[-G_\omega, G_\omega]$; moreover, by the periodicity in ϑ we can suppose that this variable belongs to $[0, 2\pi]$ and then we work with periodic conditions. Thus, the natural grid where we discretize our problem is

$$\begin{aligned} G_r =& \{i\Delta\omega, j\Delta\omega\Delta t, k\Delta\Omega, n\Delta t, \\ & -\frac{G_\omega}{\Delta\omega} \le i \le \frac{G_\omega}{\Delta\omega},\ 0 \le j \le \frac{2\pi}{\Delta\omega\Delta t},\ 0 \le n \le \frac{T}{\Delta t},\ -\frac{\Omega_1}{\Delta\Omega} \le k \le \frac{\Omega_1}{\Delta\Omega}\Bigg\}, \\ G_0 =& G_r \cap \{n = 0\}, \qquad G_b = G_r \cap \left\{|i| = \frac{G_\omega}{\Delta\omega}\right\}, \qquad G_{in} = G_r \setminus (G_0 \cup G_b). \end{aligned}$$

We conclude with one last remark. As we have shown in Lemma 6.3.6, if the initial datum ϱ_0 is normalized for every Ω, then the same condition holds true for the solution ϱ; in order to achieve this, we introduce the following normalization condition

$$(\Delta\omega)^2(\Delta t)\sum_{\underline{i},\underline{j}} \varrho^n_{\underline{i},\underline{j},k} = 1, \qquad \forall\ 0 \le n \le \frac{T}{\Delta t},\ -\frac{\Omega_1}{\Delta\Omega} \le k \le \frac{\Omega_1}{\Delta\Omega}, \tag{6.4.1}$$

The approximate Cauchy problem is

$$\begin{cases} \mathscr{L}_a \varrho = 0, & \text{in } G_{in}, \\ \varrho = \varrho_0, & \text{in } G_0, \\ \varrho = 0, & \text{in } G_b, \end{cases}$$

with ϱ that verifies (6.4.1) and such that $\varrho(\omega, \vartheta, \Omega, t) = \varrho(\omega, \vartheta + 2\pi, \Omega, t)$. While the first two equations are the classical approximation of the Cauchy problem (6.1.13), the last is intended to approximate the exponential decay condition.
As in the heat equation case, we need to introduce some stability conditions.

Theorem 6.4.9. *Suppose that the grid G_r is such that*

$$\Delta\omega \le \frac{\sqrt{2D}}{m}, \qquad \Delta t \le \frac{m^2(\Delta\omega)^2}{2D - m(\Delta\omega)^2}, \qquad G_\omega \le \frac{2D}{m\Delta\omega} - \Omega_1 - K, \tag{6.4.2}$$

then every solution to

$$\begin{cases} \mathscr{L}_a \varrho \le 0, & \text{in } G_{in}, \\ \varrho \ge 0, & \text{in } G_0, \\ \varrho = 0, & \text{in } G_b, \end{cases}$$

that verifies (6.4.1) *and such that* $\varrho(\omega, \vartheta, \Omega, t) = \varrho(\omega, \vartheta + 2\pi, \Omega, t)$*, is nonnegative in* G_r*.*

Proof. The proof runs by induction. Let us set

$$\varrho^n_{i,j,k} = \varrho(i\Delta\omega, j\Delta\omega\Delta t, k\Delta\Omega, n\Delta t).$$

As $\varrho \geq 0$ in G_0, the condition holds for $n = 0$. Let now suppose true the claim for n. We notice that $\mathscr{L}_a\varrho \geq 0$ implies

$$\begin{aligned}\varrho^{n+1}_{i,j,k} \geq & \left(1 - \frac{2D\Delta t}{m^2(\Delta\omega)^2} + \frac{\Delta t}{m}\right)\varrho^n_{i,j-i,k}\\ &+ \left(\frac{D\Delta t}{m^2(\Delta\omega)^2} - \frac{\Delta t}{2\Delta\omega m}(i\Delta\omega + \widetilde{\Phi}_\varrho((j-i)\Delta\omega, k\Delta\Omega, n\Delta t))\right)\varrho^n_{i-1,j-i,k}\\ &+ \left(\frac{D\Delta t}{m^2(\Delta\omega)^2} + \frac{\Delta t}{2\Delta\omega m}(i\Delta\omega + \widetilde{\Phi}_\varrho((j-i)\Delta\omega, k\Delta\Omega, n\Delta t)\right)\varrho^n_{i+1,j-i,k},\end{aligned}$$

where

$$\begin{aligned}\widetilde{\Phi}_\varrho((j-i)\Delta\omega, &k\Delta\Omega, n\Delta t)\\ &= -k\Delta\Omega - K\Delta\Omega(\Delta\omega)^2(\Delta t)\sum_{\underline{i},\underline{j},\underline{k}} \varrho^n_{\underline{i},\underline{j},\underline{k}} g(\underline{k})\sin(j - i + \underline{i}).\end{aligned}$$

We employ the normalization property (6.4.1) and the normalization of g to notice that $|\widetilde{\Phi}_\varrho| \leq K + \Omega_1$. Then, if (6.4.2) holds we infer that $\varrho^{n+1}_{i,j,k} \geq 0$. This concludes the proof of the theorem. $\square$

We illustrate now some numerical results. The following quantities have been analysed

$$\begin{aligned}|r(t)| &= \left|\int_0^{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\vartheta}\varrho(\omega, \vartheta, \Omega, t)g(\Omega)d\Omega d\omega d\vartheta\right|,\\ |s(t)| &= \left|\int_0^{2\pi}\int_{\mathbb{R}}\int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}\omega}\varrho(\omega, \vartheta, \Omega, t)g(\Omega)d\Omega d\omega d\vartheta\right|.\end{aligned}$$

The first one represents the phase coherence, and gives us information about the phase synchrony, while the latter is the analogous for frequencies. We have considered identical oscillators with natural frequency $\Omega = 0$, and with the following initial conditions

$$\varrho_0(\omega, \vartheta) = \frac{\mathrm{e}^{-\omega^2}(\sin(\vartheta) + 1)}{N(\omega + 1)}, \qquad N = \int_0^{2\pi}\int_{\mathbb{R}} \frac{\mathrm{e}^{-\omega^2}(\sin(\vartheta) + 1)}{(\omega + 1)} d\omega d\vartheta. \tag{6.4.3}$$

As expected, we see in Figure 6.1 that an increase in the coupling strength K leads to a greater phase synchrony. Conversely, in Figure 6.2 we see that the frequency synchronization seems to be asymptotically independent of K. This behavior could be heuristically explained as follows: suppose that, in the large time limit, the solution converges to a stationary density of the separated form $\varrho(\omega,\vartheta)=\bar{\varrho}(\omega)\hat{\varrho}(\vartheta)$ solving the following equation

$$\frac{D}{m^2}\,\partial^2_{\omega\omega}\varrho+\frac{1}{m}\partial_\omega\Big[\Big(\omega-K_\varrho(\vartheta,t)\Big)\varrho\Big]-\omega\partial_\vartheta\varrho=0, \tag{6.4.4}$$

where K_ϱ is defined in (6.1.12). By integrating (6.4.4) w.r.t ϑ and applying the divergence theorem as in (6.3.7), we infer that the marginal $\bar{\varrho}$ solves

$$\frac{D}{m}\partial^2_{\omega\omega}\bar{\varrho}(\omega)+\omega\partial_\omega\bar{\varrho}(\omega)+\bar{\varrho}(\omega)=K\partial_\omega\bar{\varrho}(\omega)\int_{\mathbb{R}}\bar{\varrho}(\omega')\int_0^{2\pi}\int_0^{2\pi}\sin(\vartheta-\vartheta')\hat{\varrho}(\vartheta)\hat{\varrho}(\vartheta')d\vartheta d\vartheta' d\omega',$$

thanks to the periodicity w.r.t. ϑ. By symmetry, the integral in the right-hand side vanishes, hence the marginal density $\bar{\varrho}$ does not depend on K. This observation aligns with the behavior of the classical Kuramoto model (6.1.4) in the case of identical oscillators, where the coupling strength does not affect the asymptotic phase synchrony.
We stress that the above argument is purely heuristic and not rigorously justified; in particular, neither the convergence to a stationary solution nor the fact that such a solution satisfies (6.4.4) is currently established, and will be addressed in future works.

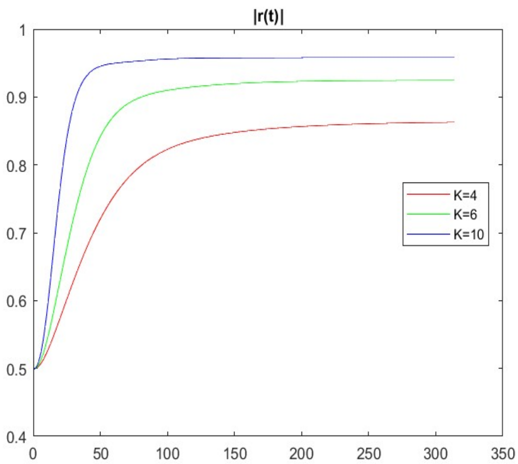

Figure 6.1: Time evolution of the phase coherence, $m=1, D=1, T=10, \Delta t=0.0317$.

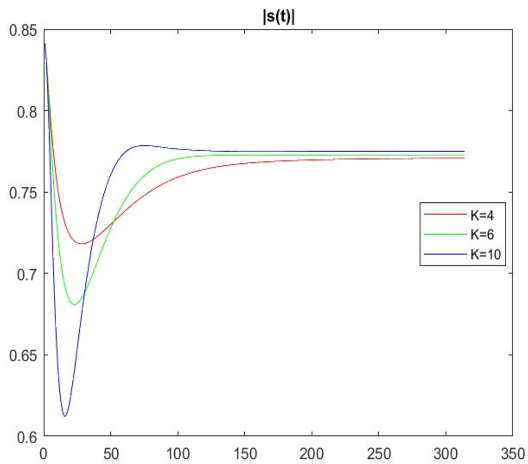

Figure 6.2: Time evolution of the frequency coherence, $m=1, D=1, T=10, \Delta t=0.0317$.

Changes in the inertial term m seem to have the opposite nature: indeed, in Figure 6.3 we see that the asymptotic behaviour of phase synchronization does not change; on the other side, in Figure 6.4 we notice that the behaviour of frequency synchronization significantly depends on this parameter, as one can observe by comparing this case with the one in Figure 6.2. In particular, we see that the asymptotic behaviour of the frequency coherence increases as m increases.

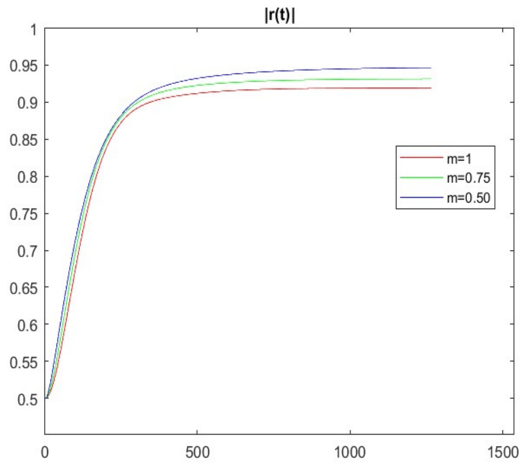

Figure 6.3: Time evolution of the phase coherence, $D = 1, K = 6, T = 10, \Delta t = 0.0079$.

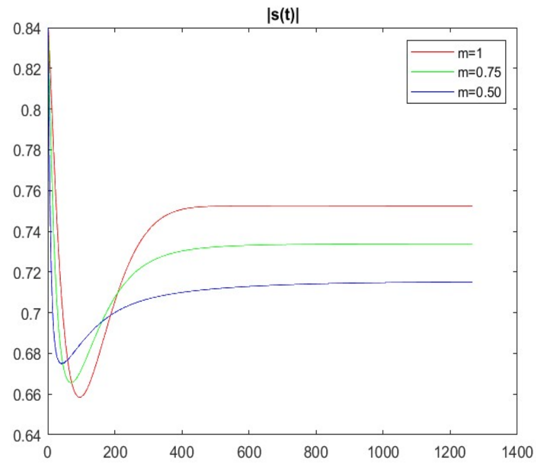

Figure 6.4: Time evolution of the frequency coherence, $D = 1, K = 6, T = 10, \Delta t = 0.0079$.

We now turn our attention to the analysis of the effects induced by variations of the noise parameter D. Consider first the original model (6.1.4) without noise and inertia: the initial condition $\varrho_0(\vartheta) = \frac{(\sin(\vartheta)+1)}{2\pi}$ yields a total phase synchrony (*i.e.* $|r| = 1$) whenever we have identical oscillators (as follows from Proposition 3.1 and Theorem 3.4 in [18]); moreover, these oscillators synchronize at the same frequency (therefore $|s| = 1$). Our analysis with initial condition (6.4.3) and identical oscillators suggests a similar behavior for decreasing D and fixed m, as illustrated in Figures 6.5 and 6.6.

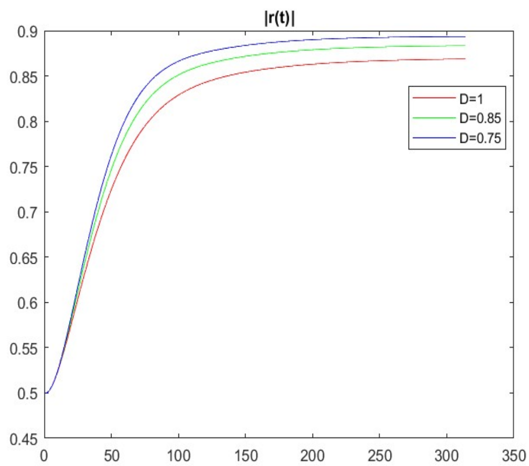

Figure 6.5: Time evolution of the phase coherence, $m = 1, K = 6, T = 10, \Delta t = 0.0317$.

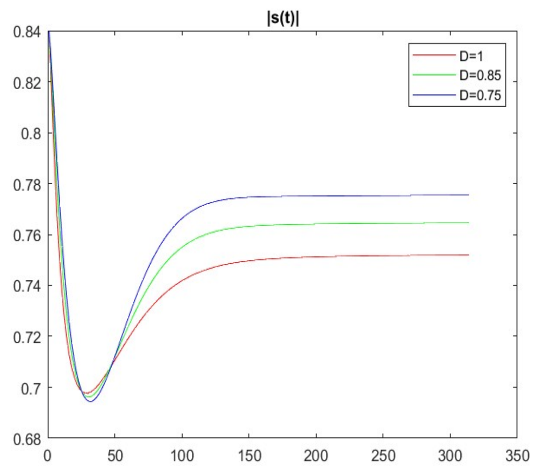

Figure 6.6: Time evolution of the frequency coherence, $m = 1, K = 6, T = 10, \Delta t = 0.0317$.

Finally, in Figure 6.7 and Figure 6.8 we compare two different initial conditions

$$\varrho_{0,1}(\omega,\vartheta) = \frac{e^{-\omega^2}(\sin(\frac{\vartheta}{2})+1)}{N_1(\omega+1)}, \qquad N_1 = \int_{\mathbb{R}}\int_0^{2\pi} \frac{e^{-\omega^2}(\sin(\frac{\vartheta}{2})+1)}{(\omega+1)} d\omega d\vartheta,$$
$$\varrho_{0,2}(\omega,\vartheta) = \frac{e^{-\omega^2}(\sin(\vartheta)+1)}{N_2(\omega+1)}, \qquad N_2 = \int_{\mathbb{R}}\int_0^{2\pi} \frac{e^{-\omega^2}(\sin(\vartheta)+1)}{(\omega+1)} d\omega d\vartheta,$$

and we observe two distinct behaviours, evidencing sensitivity to initial conditions.

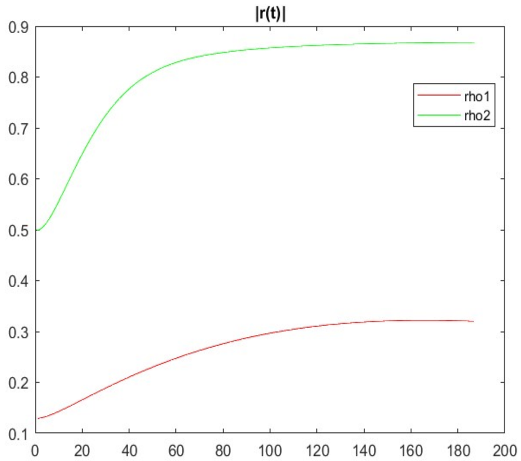

Figure 6.7: Time evolution of the phase coherence, $m = 1, D = 1, K = 4, T = 10$, $\Delta t = 0.053$.

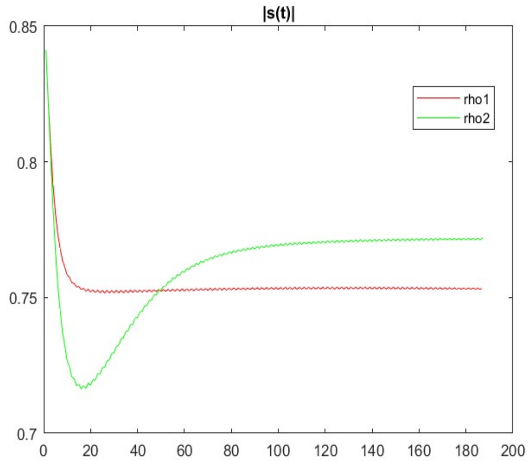

Figure 6.8: Time evolution of the frequency coherence, $m = 1, D = 1, K = 4, T = 10$, $\Delta t = 0.053$.

We conclude with one last remark. These numerical results are in agreement with the ones in [1], where the finite size system (6.1.10) was analysed. The point we are making is that we have used a different method here: indeed, in [1] the authors used a Monte Carlo method (thus a stochastic method) in the finite size case (with $N = 30000$ oscillators), while we have employed a finite difference method to analyse (6.1.11), that is the limit of (6.1.10) as $N \to \infty$.

Chapter 7

Kinetic mean field optimal control

7.1 Introduction and main results

In this chapter we investigate a McKean–Vlasov optimal control problem. This problem can be regarded as the mean field approximation of an optimal control problem relevant to a system of N interacting particles, where each particle has position $X_{i,t} \in \mathbb{R} \times [0,T]$ and velocity $V_{i,t} \in \mathbb{R} \times [0,T]$. This system evolves according to the following set of N coupled Stochastic Differential Equations (hereafter referred to as SDEs)

$$(N\textbf{-SDEs}) \begin{cases} dV_{i,t} = \left(\dfrac{u_t}{N} \displaystyle\sum_{k=1}^{N} K\Big(V_{k,t} - V_{i,t}, X_{k,t} - X_{i,t}\Big) \right) dt + \sqrt{2}\, dW_{i,t}, & i = 1, \dots, N, \\ dX_{i,t} = V_{i,t}\, dt, & i = 1, \dots, N, \\ \Big(V_{i,0}, X_{i,0}\Big)_{i \le N} \overset{\text{i.i.d.}}{\sim} F_0. \end{cases}$$

Here, K represents the coupling function, $(W_{i,t})_{t\in[0,T]}$ are N independent Wiener processes, and u is a control function belonging to the admissible set $\mathcal{U} := L^\infty\Big([0,T]; [u_{\min}, u_{\max}]\Big)$. The initial conditions $\Big(V_{i,0}, X_{i,0}\Big)_{i\le N}$ are independently and identically distributed (i.i.d.) copies of a random variable with law F_0.

The corresponding optimal control problem is

$$(\mathbf{P}_N) \quad \begin{aligned} &\min_{u\in\mathcal{U}} J_N\Big(\Big(V_{i,t}, X_{i,t}\Big), u\Big), \\ &J_N\Big(\Big(V_{i,t}, X_{i,t}\Big), u\Big) := \sum_{i=1}^{N} \frac{1}{N} \mathbb{E}\left[\int_0^T h_t\Big(V_{i,t}, X_{i,t}\Big) dt + g\Big(V_{i,T}, X_{i,T}\Big)\right] + \frac{\nu}{2} \int_0^T |u_t|^2 dt, \end{aligned}$$

where h represents the running cost, g the final cost and $\nu \ge 0$ is a penalization parameter. The study of coupled stochastic systems of this type is motivated by their broad applicability in different disciplines. In particular, these systems model collective organization phenomena, and find applications in various areas such as physics, engineering, biology and social sciences (see for instance the surveys [37, 111, 117]). Among the most prominent models in this area are the Kuramoto models introduced in the previous chapter. We recall that these models are systems of differential equations, describing the group behaviour of a population that can be seen as a cluster of interacting oscillators.

For completeness, we briefly revisit the Kuramoto model with inertia and noise studied in the previous chapter. Let us consider a system of N coupled identical oscillators with phase $\vartheta_{i,t}$ and frequency $\omega_{i,t}$. Let K be a constant (modelling the intensity of the coupling between oscillators), m a positive constant (modelling inertial effects), and let $D > 0$ be the noise amplitude. The evolution of the system is described by the following system of SDEs

$$\begin{cases} d\omega_{i,t} = \dfrac{1}{m}\left(- \omega_{i,t} + \dfrac{K}{N}\displaystyle\sum_{k=1}^{N} \sin\left(\vartheta_{k,t} - \vartheta_{i,t}\right)\right)dt + \dfrac{D}{m}\,dW_{i,t}, & i = 1, \dots, N, \\ d\vartheta_{i,t} = \omega_{i,t}dt, & i = 1, \dots, N. \end{cases} \tag{7.1.1}$$

We observe that, except for the term $-\omega_{i,t}$, the system (7.1.1) falls into the class of **(N-SDEs)**, with a sinusoidal coupling kernel acting only on the phase variable $\vartheta_{i,t}$, which in the model plays the role of the position variable $X_{i,t}$.

The study of optimal control problems in collective organization phenomena is motivated by the need to influence how agents coordinate their behaviour. For instance in Kuramoto models these problems typically involve a time-dependent coupling function K_t rather than a constant coupling K, with the objective of steering the system toward a desired synchronization state. Depending on the context, one may seek to enhance synchronization, as in power grids [21] and robotic coordination [39], or suppress it, as in certain neural applications [88].

We refer to the optimal control problem **(N-SDEs)-(P$_N$)** as *all-to-all* or *global* control problem, since the control function acts simultaneously on the entire population by changing the intensity of the interaction between the particles. We briefly present a similar N particles control problem, to further support the motivation behind our work. This problem is referred to as *sparse* optimal control problem (see for instance [14, 56, 89]). In such problems, only a fraction of particles is controlled, while the remaining ones are influenced indirectly through their interaction with the controlled group. This type of control has been extensively studied in the context of Cucker-Smale systems, which, like Kuramoto models, can converge toward certain attractors, though they are highly sensitive to initial conditions. A prominent example, that we illustrate in the following, is the shepherding problem, where a small number of M shepherd dogs are tasked with steering a herd of N sheep to a target location. Let $(x_{i,t}, v_{i,t})$ denote the position and the velocity of the i-th sheep at time t, and let $y_{j,t}$ denote the position of the j-th dog at time t. They evolve according to the following system of Ordinary Differential Equations

$$\begin{cases} \dot{v}_{i,t} = \dfrac{1}{N-1} \displaystyle\sum_{k=1, k\neq i}^{N} \Big(a(v_{k,t} - v_{i,t}) + g(x_{k,t} - x_{i,t}) \Big) \Big(x_{k,t} - x_{i,t} \Big) \\ \qquad + \dfrac{1}{M} \displaystyle\sum_{j=1}^{M} f(y_{j,t} - x_{i,t})(y_{j,t} - x_{i,t}), & i = 1, \dots, N, \\ \dot{x}_{i,t} = v_{i,t}, & i = 1, \dots, N, \\ \dot{y}_{j,t} = u_{j,t}, & j = 1, \dots, M, \\ x_{i|_{t=0}} = x_{i,0}, \quad v_{i|_{t=0}} = 0, \quad y_{j|_{t=0}} = y_{j,0}, \qquad i = 1, \dots, N, & j = 1, \dots, M. \end{cases} \tag{7.1.2}$$

Here, a and g represent the functions modelling the interaction between the sheep, while f describes the influence of the dogs on the sheep. Notably, the control u is applied directly to the dogs, and affects the sheep only indirectly through the interaction term f. The relevant optimal control problem aims to direct the sheep towards a desired position, and is expressed by minimizing the cost functional

$$J_N\Big((v_{i,t}, x_{i,t}), u\Big) = \frac{1}{N} \sum_{i=1}^{N} \int_0^T |x_{i,t} - x_d|^2 dt + \sum_{j=1}^{M} \int_0^T |u_j(t)|^2 dt, \tag{7.1.3}$$

where x_d is the target location.

The main issue inherent the control of interacting particles (as **(N-SDEs)-(P$_N$)** or (7.1.2) - (7.1.3)) is the *curse of dimensionality*, which leads to the impracticality of the optimal control computation as the numbers of particles increases. Here we give the example described in [69] for the shepherd problem (7.1.2) - (7.1.3): computing the optimal control with $M = 4$ dogs and $N = 16$ sheep (*e.g.* computing the strategy that the dogs need to employ to lead the sheep into the grassland) requires nearly two hours of computation on a standard laptop (CPU: i5-4258U 2.4GHz, RAM: 8GB DDR3L 1600MHz) running MATLAB.

This example highlights the need for alternative strategies to overcome the curse of dimensionality. For this purpose, we consider a common approach in the field of interactive particle systems: as stated before, instead of investigating the optimal control problem **(P$_N$)**, we investigate the corresponding mean field problem. By this approach, the system of N particles is replaced by a continuum corresponding to the limit of infinite particles, and the dynamic is described by the McKean-Vlasov equation

$$\textbf{(SDE)} \qquad \begin{cases} dV_t = \Big(u_t \displaystyle\int_{\mathbb{R}^2} K\Big(v - V_t, x - X_t\Big) dF_t(v,x) \Big)\, dt + \sqrt{2}\, dW_t, \\ dX_t = V_t\, dt, \\ \Big(V_0, X_0\Big) \sim F_0, \end{cases}$$

where F_t denotes the law of $\Big(V_t, X_t\Big)$. The optimal control problem $(\mathbf{P}_N)$ becomes

$$
\boxed{\begin{array}{ll}
\textbf{(P)} & \begin{array}{l}\displaystyle\min_{u\in\mathcal{U}} J\Big(\Big(V_t, X_t\Big), u\Big),\\[1ex]
\displaystyle J\Big(\Big(V_t, X_t\Big), u\Big) := \mathbb{E}\left[\int_0^T h_t\Big(V_t, X_t\Big)dt + g\Big(V_T, X_T\Big)\right] + \frac{\nu}{2}\int_0^T |u_t|^2 dt.\end{array}
\end{array}}
$$

In this chapter we address two questions. The first concerns the theoretical validity of the mean field approach: does the optimal control for the mean field limit approximate the optimal control for the finite N particles system? The second question is of computational interest: is there a feasible method to compute the mean field optimal control?

Regarding the first question, we address it by proving the Γ-convergence of the functional J_N associated to $(\mathbf{P}_N)$ to the functional J associated to **(P)**. In particular, this implies the convergence, up to a subsequence, of optimal controls for the N particles system to optimal controls for the mean field system.
To prove this result, we rely on the following assumptions

(K1) The coupling kernel K is Lipschitz continuous w.r.t. the Euclidean metric.

(D1) The law F_0 relevant to (N**-SDEs)** and to **(SDE)** has finite p-th order moment for some $p > 2$.

(J1) There exists $L > 0$ such that

$$
\Big|h_t(v,x) - h_t(\eta,\xi)\Big|, \Big|g(v,x) - g(\eta,\xi)\Big| \leq L\Big(|v-\eta| + |x-\xi|\Big),
$$

for every $v, \eta, x, \xi \in \mathbb{R}, t \in [0,T]$, and there exists $c > 0$ such that

$$
\Big|h_t(v,x)\Big|, \Big|g(v,x)\Big| \leq c\,\Big(1 + |v|^p + |x|^p\Big), \tag{7.1.4}
$$

for every $v, x \in \mathbb{R}, t \in [0,T]$, where p is the same as in **(D1)**.

In the sequel, if **(K1)**, **(D1)** and **(J1)** hold we say that **(A)** holds.

Theorem 7.1.1. *Assume **(A)** holds. Then, the functional J_N Γ-converges to the functional J in the weak * topology of $L^\infty(0,T)$. Moreover, for every $N \in \mathbb{N}$ there exists a solution u_N to $(\boldsymbol{P}_N)$, and there exists a subsequence $(u_{N_k})_{k\in\mathbb{N}}$ such that $u_{N_k} \overset{*}{\rightharpoonup} u^*$, with u^* solving **(P)**. Finally, there exists $C > 0$ (only depending on $T, p, u_{\min}$ and $u_{\max}$) such that for every $N \in \mathbb{N}$ it holds*

$$
\left|J_N\Big(\Big(V^*_{i,t}, X^*_{i,t}\Big), u^*\Big) - J\Big(\Big(V^*_t, X^*_t\Big), u^*\Big)\right| \leq C\,\frac{\ln(1+N)}{\sqrt{N}}, \tag{7.1.5}
$$

where $\left(V_{i,t}^*, X_{i,t}^*\right)_{i\leq N}$ *and* $\left(V_t^*, X_t^*\right)$ *are the stochastic processes that solves respectively* **(***N***-SDEs)** *and* **(SDE)** *with control* u^*.

Remark 7.1.2. *Let* u_N *be an optimal control for* **(***P*$_N$**)**, *and let* $\left(V_{i,t}^N, X_{i,t}^N\right)_{i\leq N}$ *denote the corresponding stochastic processes solving* **(***N***-SDEs)**. *Then, by* (7.1.5) *we infer the following estimate for the cost functionals*

$$J_N\Big(\Big(V_{i,t}^N, X_{i,t}^N\Big), u_N\Big) \leq J_N\Big(\Big(V_{i,t}^*, X_{i,t}^*\Big), u^*\Big) \leq J\Big(\Big(V_t^*, X_t^*\Big), u^*\Big) + C\,\frac{\ln(1+N)}{\sqrt{N}}.$$

Remark 7.1.3. *Theorem 7.1.1 holds also for more general stochastic processes. For instance, consider the* dN*-dimensional, real value stochastic process satisfying*

$$dY_{i,t} = b_t(Y_{i,t})dt + \left(\frac{u_t}{N}\sum_{k=1}^{N} K\Big(Y_{k,t} - Y_{i,t}\Big)\right) dt + \sigma_t\, dW_{i,t}, \qquad i = 1, \dots, N,$$

where b *is the* drift coefficient *and* σ *is the* diffusion coefficient. *This general formulation encompasses several models of interest, including the stochastic Kuramoto model* (7.1.1). *If we assume that* b *and* σ *are measurable and bounded in time, and that* b *is uniformly Lipschitz continuous in space, then the conclusion of Theorem 7.1.1 still holds. We restrict our attention to the* **(***N***-SDEs)** *model to keep the notation as simple and transparent as possible.*

The problem to establish the connection between the optimal control of the N particles system and the corresponding mean field problem has attracted growing attention in recent years. In the sparse optimal control setting, a link between the two problems is established in [55–57] for the deterministic case, and in [6, 15, 89] for the stochastic case, where the authors proved a similar Γ-convergence results. We recall that, in the sparse control problem, the number of controllers (the dogs in the previous example) is kept fixed, while the number of uncontrolled particles (the sheep) tends to infinity.
For the global stochastic control problem **(***N***-SDEs)-(P**$_N$**)** we are investigating in this chapter, some results have been achieved in recent years in [50, 75]. Here the authors, in a fully probabilistic approach, proved the convergence of *approximate* optimal controls for the N particles problem to the optimal controls of the mean field problem (*i.e.* convergence of optimal controls up to an ε_N error, that vanishes as the number of particles N goes to infinity). In Theorem 7.1.1, although in a totally different setting, we recover these results if there is the uniqueness of minimizers (in particular, Theorem 2.12 and Corollary 2.13 in [75]). Indeed, in this case not just a subsequence, but the whole sequence of optimal controls for **(P**$_N$**)** converges to the optimal control for **(P)**.

The second question we address is strictly tied to the previous result. Having established that the optimal controls for the N particles system can be suitably approximated by the mean field optimal controls, is there a way to practically compute it? This question falls within the broader framework of stochastic optimal control problem, and we decide to address it by following a PDEs approach that has been gaining popularity in recent years (see for instance [10, 11, 107], or the recent survey [12]). Assume that the law F of the stochastic process solving **(SDE)** admits density f w.r.t. the Lebesgue measure. By Itô formula, it satisfies the following degenerate Kolmogorov-Fokker-Planck Cauchy problem

$$\textbf{(PDE)}\begin{cases}\partial_t f_t(v,x) + v\partial_x f_t(v,x) \\ \qquad\qquad = \partial^2_{vv} f_t(v,x) - \partial_v\left(u_t \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi\, f_t(v,x)\right), \\ f_{|t=0} = f_0.\end{cases}$$

In this case, the functional of the optimal control problem **(P)** can be expressed in terms of the density as follows

$$J(f,u) := \int_{\mathbb{R}^2}\int_0^T h_t(v,x) f_t(v,x) dt dv dx + \int_{\mathbb{R}^2} g(v,x) f_T(v,x) dv dx + \frac{\nu}{2}\int_0^T |u_t|^2 dt. \tag{7.1.6}$$

The key advantage of this approach is that it recasts the problem from a stochastic framework into a deterministic one. This reformulation allows for a more structured analysis, making it possible to apply powerful techniques developed in the realm of PDE optimal control problems (see for instance the monograph [116]). In particular, it provides us both theoretical and practical methods for computing optimal controls. A key tool in this setting is the Lagrangian formalism (as in [4, 13]), which enables us to derive first order necessary conditions for the minimizers. This formalism relies on the adjoint variable p, defined as the solution to the adjoint Cauchy problem

$$\textbf{(ADJ)}\begin{cases}\partial_t p_t(v,x) + v\partial_x p_t(v,x) = -\partial^2_{vv} p_t(v,x) \\ \qquad -u_t\Bigg(\partial_v p_t(v,x)\int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \\ \qquad\qquad + \int_{\mathbb{R}^2} K(v-\eta, x-\xi)\partial_v p(\xi,\eta) f_t(\eta,\xi) d\eta d\xi\Bigg) - h_t(v,x), \\ p_{|t=T}(v,x) = g(v,x).\end{cases}$$

By using f and p, we are able to compute the gradient of (7.1.6). Consequently, the optimality conditions for the minimizers of **(P)** follow straightforwardly.
To formally carry out this computation, we need the following assumptions

(K2) $K \in L^2(\mathbb{R}^2) \cap L^\infty(\mathbb{R}^2)$, and its derivative $\partial_v K$ is bounded and Lipschitz continuous w.r.t. the Euclidean metric.

(D2) There exists $f_0 \in L^1(\mathbb{R}^2) \cap L^2(\mathbb{R}^2)$ such that the law F_0 admits density f_0 w.r.t. the Lebesgue measure.

(J2) $h \in L^2\Big(\mathbb{R}^2 \times [0,T]\Big)$ and $g \in L^2(\mathbb{R}^2)$.

In the sequel, if **(A)**, **(K2)**, **(D2)** and **(J2)** hold we say that **(B)** holds.

Theorem 7.1.4. *Assume **(B)** holds. If $u \in \mathcal{U}$ is a solution to **(P)**, then*

$$\int_0^T \int_{\mathbb{R}^2} \Big(f_t(v,x) \partial_v p_t(v,x) \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \Big)(w_t - u_t) dv dx dt \\ + \nu \int_0^T u_t(w_t - u_t) dt \geq 0, \qquad \forall w \in \mathcal{U}, \tag{7.1.7}$$

*where f and p are the solutions to **(PDE)** and **(ADJ)** with control function u. Hence, if $\nu > 0$ the following conditions hold for a.e. $t \in [0,T]$:*

1. *if*

$$u_t > -\frac{1}{\nu} \int_{\mathbb{R}^2} \Big(f_t(v,x) \partial_v p_t(v,x) \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \Big) dv dx,$$

   then $u_t = u_{\min}$;

2. *if*

$$u_t < -\frac{1}{\nu} \int_{\mathbb{R}^2} \Big(f_t(v,x) \partial_v p_t(v,x) \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \Big) dv dx,$$

   then $u_t = u_{\max}$;

3. *if $u_{\min} < u_t < u_{\max}$, then the following equality holds*

$$u_t = -\frac{1}{\nu} \int_{\mathbb{R}^2} \Big(f_t(v,x) \partial_v p_t(v,x) \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \Big) dv dx.$$

Theorem 7.1.4 holds significant importance due to its practical applications. More precisely, it enables us to formalize a gradient descent method, which we have implemented in some numerical simulations. We will illustrate these results in details in Section 7.7.

We conclude this introduction with two remarks. First, although assumption **(J1)** is not needed for the proof of Theorem 7.1.4 (which is obtained purely within a PDE framework) it is essential for establishing Theorem 7.1.1, that is, the connection between the N particles optimal control problem and its mean field counterpart. Second, Theorem 7.1.4 does not require the cost functional to be linear in the density variable. We note that the specific form of the functional in (7.1.6) naturally arises from its interpretation as an expectation associated with the underlying stochastic process. However, if we consider only the control problem relevant to the density f, we can focus on cost functionals more general than (7.1.6). For instance, consider the case where we aim to steer the density toward a specific target form. Let $\bar{f}$ denote the target running density, and f_d denote the target final density. We can consider the L^2 cost functional

$$J(f,u) = \int_0^T \int_{\mathbb{R}^2} \Big|f_t(v,x) - \bar{f}_t(v,x)\Big|^2 dvdxdt + \int_{\mathbb{R}^2} \Big|f_T(v,x) - f_d(v,x)\Big|^2 dvdx + \frac{\nu}{2}\int_0^T |u_t|^2 dt.$$

If we set $h_t(v,x) := \Big|f_t(v,x) - \bar{f}_t(v,x)\Big|^2$ and $g(v,x) := \Big|f_T(v,x) - f_d(v,x)\Big|^2$, the optimality conditions stated in Theorem 7.1.4 still holds. In particular, we can apply a gradient descent method also for this class of functionals.

This chapter is organized as follows: in Section 7.2, Section 7.3, and Section 7.4 we give all the theoretical tools we need to prove Theorem 7.1.1 and Theorem 7.1.4; in Section 7.5 we give the proof of Theorem 7.1.1, while Section 7.6 provides the proof of Theorem 7.1.4; finally, we conclude in Section 7.7 with some numerical simulations.

We end the introduction with some notational remarks.
We set $\bar{u} = \max\Big(|u_{\min}|, |u_{\max}|\Big)$, and, without loss of generality, we set all the Lipschitz constants equal to one. For every stochastic process $(Y_t)_{t\in[0,T]}$ we will denote as Y_t its realization at time t, while (Y_t) represents the complete realization over the time interval $[0,T]$. Finally, we point out two minor notational discrepancies with respect to the previous chapters. The first concerns the notation $f_t(v,x)$ instead of the more classical $f(v,x,t)$; this choice reflects the fact that, for each $t \in [0,T]$, the function f_t represents the density of the measure F_t appearing in **(SDE)** (see Section 7.3 for further details). The second concerns the definition of the Lie derivative: in this chapter we adopt the convention $Y = v\partial_x + \partial_t$ rather than the usual $Y = -v\partial_x - \partial_t$ introduced in Chapter 5. We employ this choice in order to remain consistent with the notation used in [16].

7.2 Stochastic processes and Stochastic Differential Equations

In this section we briefly revisit some key definitions and results required to analyse **(N-SDEs)** and **(SDE)**. For a comprehensive discussion on the theory of stochastic processes and SDEs, we direct the interested reader to the monograph [93].
For the rest of the section n will denote a generic natural number.

We start by describing the SDE framework that fits **(N-SDEs)** and **(SDE)**. Let us fix a probability space $(\Omega, \mathscr{F}, \mathbb{P})$. A filtration $(\mathscr{F}_t)_{t\in[0,T]}$ on $(\Omega, \mathscr{F}, \mathbb{P})$ (denoted as $(\mathscr{F}_t)$ in the follow-

ing) is an increasing collection of σ-algebras such that $\mathscr{F}_T \subseteq \mathscr{F}$.
We say that a filtration $(\mathscr{F}_t)$ satisfies the usual condition if it is right continuous and contains the set $\Big\{A \in \mathbb{R}^n \text{ s.t. } A \subseteq B \in \mathscr{F} : \mathbb{P}(B) = 0\Big\}$.
An n-dimensional Wiener process (W_t) on $(\mathscr{F}_t)$ is a stochastic process adapted to $(\mathscr{F}_t)$ such that:

- $W_0 = 0$ a.s.;
- (W_t) is a.s. continuous;
- $W_t - W_s$ is independent of $\mathscr{F}_s$ for every $0 \le s \le t$;
- $W_t - W_s \sim \mathcal{N}(0, t-s)$ for every $0 \le s \le t$.

We are ready to give the definition of strong solutions to the linear SDE

$$\begin{cases} dZ_t = b_t\Big(Z_t, Y_t\Big)\, dt + \sqrt{2}\, dW_t, \\ dY_t = Z_t\, dt, \\ \Big(Z_{|t=0}, Y_{|t=0}\Big) = \Big(Z_0, Y_0\Big). \end{cases} \tag{7.2.1}$$

We notice that the linear system **(N-SDEs)** is a particular instance of (7.2.1) with $n = N$, $(Z_t, Y_t) = \Big(V_{i,t}, X_{i,t}\Big)_{i \le N}$, and $b_t(Z_t, Y_t) = \dfrac{u_t}{N} \sum_{k=1}^{N} K\Big(V_{k,t} - V_{i,t}, X_{k,t} - X_{i,t}\Big)$.
The **(SDE)** that arise in the mean field limit is a nonlinear version of (7.2.1), where we set $n = 1$, $(Z_t, Y_t) = (V_t, X_t)$, and $b_t(Z_t, Y_t) = u_t \int_{\mathbb{R}^2} K\Big(v - V_t, x - X_t\Big) dF_t(v, x)$, with the nonlinearity given by the law F of (V_t, X_t).

Definition 7.2.5 (Strong solutions to SDE)**.** *Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space with filtration $(\mathscr{F}_t)$ and n-dimensional Wiener process (W_t). Let $b : \mathbb{R}^n \times \mathbb{R}^n \times [0, T] \mapsto \mathbb{R}^n$ be a measurable function, and (Z_0, Y_0) be a $2n$-dimensional random variable $\mathscr{F}_0$-measurable. We say that a continuous and adapted $2n$-dimensional stochastic process (Z_t, Y_t) is a solution to* (7.2.1) *if a.s. it holds*

$$\int_0^t \Big| b_s\Big(Z_s, Y_s\Big) \Big| ds < +\infty, \qquad \forall\, 0 \le t \le T,$$

and the relation below holds a.s. for every $t \in [0, T]$

$$\begin{cases} Z_t = Z_0 + \int_0^t b_s\Big(Z_s, Y_s\Big) ds + \sqrt{2}\, W_t, \\ Y_t = Y_0 + \int_0^t Z_s ds. \end{cases}$$

We say that strong uniqueness holds if all strong solutions have the same law.

Existence and uniqueness of strong solutions to (7.2.1) is ensured by the classical Lipschitz continuity and linear growth conditions on b

$$\begin{aligned}
&\Big|b_t(z_1,y_1)-b_t(z_2,y_2)\Big| \le C\Big(|z_1-z_2|+|y_1-y_2|\Big), && \forall z_1,z_2,y_1,y_2\in\mathbb{R}, t\in[0,T],\\
&\Big|b_t(z,y)\Big| \le C(1+|z|+|y|), && \forall z,y\in\mathbb{R}, t\in[0,T],
\end{aligned} \tag{7.2.2}$$

for some constant $C>0$.

Theorem 7.2.6. *Assume that* (7.2.2) *holds. Then, there exists a unique strong solution* (Z_t,Y_t) *to* (7.2.1). *Moreover, if* $\mathbb{E}\Big[\Big|\Big(Z_0,Y_0\Big)\Big|^p\Big]<\infty$ *for some* $p\ge 2$*, then there exists a constant* C_T *such that*

$$\mathbb{E}\left[\sup_{t\le T}\Big|\Big(Z_t,Y_t\Big)\Big|^p\right]\le C_T\left(1+\mathbb{E}\Big[\Big|\Big(Z_0,Y_0\Big)\Big|^p\Big]\right).$$

Proof. The existence is established using a classical fixed-point theorem, while the uniqueness follows from a stopping time argument (see for instance Theorem 17.1.1 and Theorem 17.2.1 in [93]). The moment estimate follows from the properties of the Wiener process and the linear growth (7.2.2) (see Theorem 14.5.2 in [93] for the details). □

7.3 Propagation of chaos and Wasserstein distance

In this section we present the mathematical tools needed to rigorously define and study the link between the N particles system and its mean field limit.

The convergence of (**N-SDEs**) to its mean field limit **(SDE)** is formalized through the concept of propagation of chaos. This notion, introduced by McKean [86] and later popularized after the Saint-Flour lecture notes given by Sznitman [114], describes the asymptotic behaviour of a system of N particles as their number approaches infinity. To be more precise, it states that the sequence of stochastic processes solving (**N-SDEs**) asymptotically behaves like the stochastic process that solves the mean field limit **(SDE)**.

To provide the definition of propagation of chaos, we introduce the notion of empirical measure F^N associated to the stochastic process $\Big(V_{i,t},X_{i,t}\Big)_{i\le N}$

$$F^N:=\frac{1}{N}\sum_{i=1}^N\delta_{(V_{i,t},X_{i,t})},$$

where $\delta_{(V_{i,t},X_{i,t})}$ denotes the Dirac measure centred at $(V_{i,t},X_{i,t})$.

Definition 7.3.7 (Propagation of chaos)**.** *Let F be the law of the stochastic process (V_t,X_t). A sequence of stochastic processes $\Big(V_{i,t},X_{i,t}\Big)_{i\le N}$ converges in the sense of the propagation* *of chaos to* (V_t, X_t) *if its empirical measure converges in probability to* F*, namely*

$$F^N = \frac{1}{N}\sum_{i=1}^{N}\delta_{(V_{i,t},X_{i,t})} \xrightarrow[N\to\infty]{\mathcal{P}} F. \tag{7.3.1}$$

In this case we say that $\Big(V_{i,t}, X_{i,t}\Big)_{i\le N}$ *is* F*-chaotic.*

A key tool for proving this property is the Wasserstein distance, commonly used in the context of probability measures. For a detailed overview, we refer the reader to the monograph [118]. In view of **(SDE)**, **(D1)**, and Theorem 7.2.6, it is clear that we work within the framework of stochastic processes whose law belongs to the space of probability measure on $C = C\Big([0,T];\mathbb{R}^2\Big)$ with first moment bounded (denoted as $\mathcal{P}(C)$ in the following). The Wasserstein distance on this space is defined as

$$\mathcal{W}_{[0,T]}\Big(F_1, F_2\Big) := \inf_{\gamma}\int_{\mathbb{R}^2\times\mathbb{R}^2}\sup_{t\in[0,T]}|X_t(\omega_1) - X_t(\omega_2)|d\gamma(\omega_1,\omega_2), \tag{7.3.2}$$

where X_t is the canonical process on C, and the infimum is taken over all $\gamma \in \mathcal{P}(C\times C)$ with marginal F_1 and F_2 (usually referred to as *coupling* between F_1 and F_2). This distance defines a complete metric space on $\mathcal{P}(C)$, equipping it with the topology of the weak convergence (see Theorem 3.2 in [51]). This implies that the propagation of chaos (7.3.1) can be expressed by means of the Wasserstein distance as follows

$$\lim_{N\to+\infty}\mathbb{E}\Big[\mathcal{W}_{[0,T]}\Big(F^N, F\Big)\Big] = 0. \tag{7.3.3}$$

As can be deduced from **(SDE)**, in our context it is more natural to work with the evaluation of the law F at time $t \in [0,T]$ (denoted to as F_t), rather then considering its entire trajectory. In order to make this precise, we introduce the evaluation map ev_t, which assigns to $g \in C$ the value $\mathrm{ev}_t g := g_t$. For each $F \in \mathcal{P}(C)$ we define $F_t := \mathrm{ev}_t \# F$, where $\#$ denotes the pushforward.
The Wasserstein distance between $F_{1,t}$ and $F_{2,t}$ is straightforwardly defined as

$$\mathcal{W}_t\Big(F_1, F_2\Big) := \inf_{\gamma}\int_{\mathbb{R}^2\times\mathbb{R}^2}|x-y|d\gamma(x,y), \tag{7.3.4}$$

where the infimum is taken over all $\gamma \in \mathcal{P}\Big(\mathbb{R}^2\times\mathbb{R}^2\Big)$ with marginal $F_{1,t}$ and $F_{2,t}$. Clearly, $\mathcal{W}_t(F_1,F_2) \le \mathcal{W}_{[0,T]}(F_1,F_2)$ for every $t \in [0,T]$.

We conclude this subsection by presenting a useful tool to work with the Wasserstein distance (7.3.4).

Lemma 7.3.8 (Kantorovich duality)**.** *The Wasserstein distance* (7.3.4) *can equivalently be* *characterized as*

$$\mathcal{W}_t\Big(F_1, F_2\Big) := \sup_{\varphi} \int_{\mathbb{R}^2} \varphi(x)\Big(dF_{1,t} - dF_{2,t}\Big)(x),$$

where the supremum is taken over all Lipschitz function φ with Lipschitz constant equal to one. In particular, the following inequality holds

$$\int_{\mathbb{R}^2} |x|\Big(dF_{1,t} - dF_{2,t}\Big)(x) \leq \mathcal{W}_t\Big(F_1, F_2\Big), \qquad \forall F_{1,t}, F_{2,t} \in \mathcal{P}(\mathbb{R}^2). \tag{7.3.5}$$

Proof. Let $c : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R} \cup \{+\infty\}$ be a lower semicontinuous function such that

$$c(x, y) \geq a(x) + b(y), \qquad \forall x, y \in \mathbb{R}^2,$$

for some real valued, upper semicontinuous function $a \in L^1_{F_{1,t}}(\mathbb{R}^2)$ and $b \in L^1_{F_{2,t}}(\mathbb{R}^2)$ (this notation refers to the Lebesgue space w.r.t. the measures $F_{1,t}$ and $F_{2,t}$). By Theorem 5.10 in [118] it holds the duality

$$\inf_{\gamma} \int_{\mathbb{R}^2\times\mathbb{R}^2} c(x, y) d\gamma(x, y) = \sup_{\varphi} \left(\int_{\mathbb{R}^2} \varphi^c(x) dF_{1,t}(x) - \int_{\mathbb{R}^2} \varphi(x) dF_{2,t}(x) \right), \tag{7.3.6}$$

where the supremum is taken over the set of c-convex function $\varphi \in L^1_{F_{1,t}}(\mathbb{R}^2)$, and

$$\varphi^c(x) := \inf_{y\in\mathbb{R}^2} \Big[\varphi(y) + c(x, y)\Big].$$

We apply (7.3.6) with $c(x, y) = |x-y| \geq |x|-|y| = a(x)+b(y)$. Since $F_{1,t}$ and $F_{2,t}$ have first moment bounded, we observe that a and b satisfy the integrability assumption. Moreover, since c is a distance, φ^c is a Lipschitz function with Lipschitz constant equal to 1 and it holds $\varphi^c = \varphi$ (see Particular Case 5.4 in [118]). We finally observe that, if φ is Lipschitz continuous, then $\varphi \in L^1_{F_{1,t}}(\mathbb{R}^2) \cap L^1_{F_{2,t}}(\mathbb{R}^2)$, since $|\varphi(x)| \leq |\varphi(x) - \varphi(0)| + |\varphi(0)| \leq |x| + |\varphi(0)|$. Thus, the dual representation holds, concluding the proof. □

7.4 Degenerate Kolmogorov-Fokker-Planck PDEs

In this section we recall some key results from the theory of degenerate Kolmogorov-Fokker-Planck equations, as presented in Chapter 5, which will be used to establish the optimality conditions in Theorem 7.1.4.
The prototype of this class of differential operators is

$$\mathscr{L} f = (\partial_t + v\partial_x) f - \partial^2_{vv} f, \qquad v, x \in \mathbb{R},\ t \in [0, T]. \tag{7.4.1}$$

This operator is degenerate, since it lacks diffusion in the x variable. Nevertheless, Kolmogorov [70] proved that it enjoys the same regularizing property as the heat operator, as discussed in Chapter 1. Kolmogorov's observations where later extended by Hörmander [63], who proved this regularizing property for the class of degenerate operators (1.1.2), known as sum of squares plus drift operators. His key insight was to interpret such operators through their Lie derivative structure, and to recover regularity by iterating suitable Lie brackets. The Lie derivative appearing in (7.4.1) is

$$Yf := (\partial_t + v\partial_x)f.$$

We first encounter a Kolmogorov equation in the nonlinear PDE that describes the evolution of the mean field density. For every $G \in \mathcal{P}(C)$ with density g, we define the differential operator

$$\begin{aligned}\mathscr{L}_g f &= Yf_t(v,x) - \mathcal{A}_g f_t(v,x) \\ &= \partial_t f_t(v,x) + v\partial_x f_t(v,x) - \partial^2_{vv} f_t(v,x) \\ &\quad + \partial_v \left(u_t \int_{\mathbb{R}^2} K(\eta - v, \xi - x) g_t(\eta,\xi) d\eta d\xi \, f_t(v,x) \right),\end{aligned} \tag{7.4.2}$$

that is of Kolmogorov type (7.4.1) with lower order terms. As shown in Proposition 7.6.16 below, if **(B)** holds the law F of **(SDE)** admits density f which solves the nonlinear PDE $\mathscr{L}_f f = Yf - \mathcal{A}_f f = 0$. Since $u \in \mathcal{U} = L^\infty\Big([0,T];[u_{\min},u_{\max}]\Big)$, and in view of assumptions **(K1)** and **(K2)** (which ensure that K and $\partial_v K$ are both bounded and Lipschitz continuous), the lower order coefficients in (7.4.2) are Lipschitz continuous and bounded in v and x (w.r.t. the Euclidean metric), and measurable and bounded in t. It is straightforward to show that this implies that the coefficients belong to the space $L^\infty\Big((0,T);C^\alpha_b(\mathbb{R}^2)\Big)$ introduced in Chapter 5, so that we can work within the framework of strong solutions presented there.

Definition 7.4.9 (Strong Lie solution)**.** *We say that f is a strong solution to $Yf - \mathcal{A}_g f = 0$ in $\mathbb{R}^2 \times [0,T]$ if f is a continuous function with $\partial_v f, \partial^2_{vv} f \in L^1_{\text{loc}}\Big([0,T];C_b(\mathbb{R}^2)\Big)$, and such that $Yf = \mathcal{A}_g f$ is an a.e. Lie derivative, namely*

$$\begin{aligned}&f_t(v, x + v(s-t)) - f_s(v,x) \\ &\qquad = \int_s^t \mathcal{A}_g f_\tau(v, x + v(s-\tau)) d\tau, \quad (v,x) \in \mathbb{R}^2, 0 < s < t < T.\end{aligned} \tag{7.4.3}$$

By employing the tools developed in [20, 36, 90] (in particular, the existence of the fundamental solution and some Gaussian estimates), we show that the measure F of the solution to **(SDE)** admits density f, which is the unique strong solution to the Cauchy problem

$$
\begin{cases}
\mathscr{L}_f f_t(v,x) \\
\quad = Y f_t(v,x) - \partial^2_{vv} f_t(v,x) + \partial_v \Big(u_t \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \, f_t(v,x) \Big) = 0, \\
f_{|t=0} = f_0,
\end{cases}
\tag{7.4.4}
$$

where $f_{|t=0} = f_0$ is understood in L^2 sense.

Once we prove the existence of a strong solution to (7.4.4), we rely on the Lagrangian formalism and the adjoint equation **(ADJ)** to prove Theorem 7.1.4. In this framework, it is natural to consider weak solutions. Accordingly, we employ the weak theory recently developed [16], which is based on Lion's form method.
As we have seen in Chapter 5, we need to address the Lie derivative Y also for weak solutions. To this end, we introduce the relevant functional space W_T.
Since we have diffusion in the v variable, we introduce the Sobolev space

$$
H^1_v(\mathbb{R}) := \Big\{ g \in L^2(\mathbb{R}) : \partial_v g \in L^2(\mathbb{R}) \Big\},
$$

equipped with the norm

$$
\|g\|^2_{H^1_v(\mathbb{R})} := \|g\|^2_{L^2(\mathbb{R})} + \|\partial_v g\|^2_{L^2(\mathbb{R})}.
$$

We denote its dual space as $H^{-1}_v(\mathbb{R})$. The space W_T is then defined as the closure of the space $C^\infty_c\Big(\mathbb{R}^2 \times [0,T]\Big)$ w.r.t. the norm

$$
\|g\|_{W_T} := \int_0^T \int_{\mathbb{R}} \|g_t(\cdot,x)\|^2_{H^1_v(\mathbb{R})} dx dt + \int_0^T \int_{\mathbb{R}} \|Y g_t(\cdot,x)\|^2_{H^{-1}_v(\mathbb{R})} dx dt.
$$

By density (see Theorem 5.5 (iii) in [16] and Theorem 11.3 in [38]) we have

$$
W_T = \Big\{ g \in L^2\Big(\mathbb{R} \times [0,T]; H^1_v\Big) \text{ s.t. } Y g \in L^2\Big(\mathbb{R} \times [0,T]; H^{-1}_v\Big)\Big\}.
$$

In the following, we denote as $L^2_{x,t}H^1_v$ the space $L^2\Big(\mathbb{R} \times [0,T]; H^1_v\Big)$, and as $L^2_{x,t}H^{-1}_v$ the space $L^2\Big(\mathbb{R} \times [0,T]; H^{-1}_v\Big)$.

Working on this functional space, we can consider a wide family of kinetic Cauchy problems: let $\mathscr{A} : L^2_{x,t}H^1_v \to L^2_{x,t}H^{-1}_v$ be a general differential operator (*e.g.* $\mathscr{A} = -\partial^2_{vv}$, or even a diffusion operator with measurable coefficients and low order terms), for every $\bar{f} \in L^2_{x,t}H^{-1}_v$ we consider the Cauchy problem

$$
\begin{cases}
Y f_t(v,x) + \mathscr{A} f_t(v,x) = \bar{f}_t(v,x), & (v,x) \in \mathbb{R}^2, t \in (0,T), \\
f_{|t=0}(v,x) = f_0(v,x), & (v,x) \in \mathbb{R}^2.
\end{cases}
\tag{7.4.5}
$$

Definition 7.4.10 (Weak Lie solution)**.** *We say that a function f is a weak solution to* (7.4.5) *if $f \in W_T$, $f_{|t=0} = f_0$ in L^2 sense, and for every $\varphi \in C_c^\infty\Big(\mathbb{R}^2 \times [0,T]\Big)$ it holds*

$$\int_0^T \int_{\mathbb{R}} \langle Y f_t(\cdot, x), \varphi_t(\cdot, x)\rangle dx dt + \int_0^T \int_{\mathbb{R}} \langle \mathscr{A} f_t(\cdot, x), \varphi_t(\cdot, x)\rangle = \int_0^T \int_{\mathbb{R}} \langle \bar{f}_t(\cdot, x), \varphi_t(\cdot, x)\rangle, \tag{7.4.6}$$

where $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H_v^1(\mathbb{R})$ and $H_v^{-1}(\mathbb{R})$.

We conclude this section with two crucial remarks.

Remark 7.4.11. *We point out that $f_{|t=0} = f_0$ is well posed, thanks to the embedding*

$$W_T \hookrightarrow C\Big([0,T], L^2\Big(\mathbb{R}^2\Big)\Big), \tag{7.4.7}$$

given by Theorem 7.1 and (7.3) in [16].

Remark 7.4.12. *By the density of $C_c^\infty\Big(\mathbb{R}^2 \times [0,T]\Big)$ in W_T, we can test* (7.4.6) *with $\varphi \in W_T$. This observation is useful in combination with the integration by parts formula*

$$\begin{aligned}\int_0^T \int_{\mathbb{R}} \langle Y z_t(\cdot, x), p_t(\cdot, x)\rangle dx dt &+ \int_0^T \int_{\mathbb{R}} \langle Y p_t(\cdot, x), z_t(\cdot, x)\rangle dx dt \\ &= \int_{\mathbb{R}^2} z_T(v,x) p_T(v,x) dv dx - \int_{\mathbb{R}^2} z_0(v,x) p_0(v,x) dv dx, \qquad \forall z, p \in W_T.\end{aligned}$$

7.5 Proof of Theorem 7.1.1

Throughout this section we assume that **(A)** holds.

The strategy for the proof of Theorem 7.1.1 is the following. We first prove that for every $N \in \mathbb{N}$ there exists an optimizer for **($\mathbf{P}_N$)**. Then, since the sequence of optimizers $(u_N)_{N \in \mathbb{N}}$ belongs to $\mathcal{U} = L^\infty\Big([0,T]; [u_{\min}, u_{\max}]\Big)$, we infer the existence of $u^* \in \mathcal{U}$ such that $u_N \overset{*}{\rightharpoonup} u^*$ (up to a subsequence). Next, we establish the propagation of chaos of **(N-SDE)** with control u_N to **(SDE)** with control u^*, *i.e.* propagation of chaos under the convergence of the controls. Hence, we conclude with the proof of the Γ-convergence, showing the $\liminf$ inequality and the existence of a recovery sequence. In particular, from the Γ-convergence it follows that u^* is a solution to **(P)**. Finally, the bound (7.1.5) follows from the propagation of chaos result previously proved.

We start by proving the existence of an optimal control for the N particle optimization problem **($\mathbf{P}_N$)**.

Proposition 7.5.13. *For every $N \in \mathbb{N}$ there exists an optimal control u_N^* solution to **($\boldsymbol{P}_N$)**.*

Proof. Fix $N \in \mathbb{N}$, and let $(u_N^j)_{j\in\mathbb{N}} \in \mathcal{U}$ be a minimizing sequence for $(\mathbf{P}_N)$. By boundedness, we infer the existence of a subsequence (still denoted as $(u_N^j)_{j\in\mathbb{N}}$), and $u_N^* \in \mathcal{U}$ such that $u_N^j \overset{*}{\rightharpoonup} u_N^*$. We claim that u_N^* solves $(\mathbf{P}_N)$. Let us consider the N particles SDEs (N**-SDEs**) with control u_N^j and u_N^*. By the Lipschitz condition **(K1)**, it is not difficult to see that we are in position to apply Theorem 7.2.6, that ensures the existence of strong solutions $\left(V_{i,t}^j, X_{i,t}^j\right)_{i\le N}$ and $\left(V_{i,t}^*, X_{i,t}^*\right)_{i\le N}$, respectively. Moreover, thanks to **(D1)**, they have finite p-th moment.
We claim that a.s. the following relation holds

$$\max_{i\le N}\sup_{t\in[0,T]}\left[\left|V_{i,t}^j - V_{i,t}^*\right| + \left|X_{i,t}^j - X_{i,t}^*\right|\right] \xrightarrow[j\to\infty]{} 0. \tag{7.5.1}$$

Indeed, since $\left(V_{i,t}^j, X_{i,t}^j\right)_{i\le N}$ and $\left(V_{i,t}^*, X_{i,t}^*\right)_{i\le N}$ solve (N**-SDEs**), the following holds a.s.

$$\begin{aligned}
\left|V_{i,t}^j - V_{i,t}^*\right| &\le \bar{u}\int_0^t \frac{1}{N}\sum_{k=1}^N \left|K\left(V_{k,\tau}^j - V_{i,\tau}^j, X_{k,\tau}^j - X_{i,\tau}^j\right) - K\left(V_{k,\tau}^* - V_{i,\tau}^*, X_{k,\tau}^* - X_{i,\tau}^*\right)\right| d\tau \\
&\quad + \int_0^t \left|u_{N,\tau}^j - u_{N,\tau}^*\right| \left|\frac{1}{N}\sum_{k=1}^N K\left(V_{k,\tau}^* - V_{i,\tau}^*, X_{k,\tau}^* - X_{i,\tau}^*\right)\right| d\tau := I_1 + I_2, \\
\left|X_{i,t}^j - X_{i,t}^*\right| &\le \int_0^t \left|V_{i,\tau}^j - V_{i,\tau}^*\right| d\tau.
\end{aligned} \tag{7.5.2}$$

By the Lipschitz assumption on K we infer

$$I_1 \le 2\bar{u}\int_0^t \max_{i\le N}\sup_{\tau\in[0,t]}\left[\left|V_{i,\tau}^j - V_{i,\tau}^*\right| + \left|X_{i,\tau}^j - X_{i,\tau}^*\right|\right] d\tau. \tag{7.5.3}$$

We employ again the Lipschitz continuity to bound I_2, observing that

$$\left|K(v,x)\right| \le \left|K(v,x) - K(0,0)\right| + \left|K(0,0)\right| \le |v| + |x| + \left|K(0,0)\right|.$$

This implies

$$\begin{aligned}
I_2 &\le \int_0^t \left|u_{N,\tau}^j - u_{N,\tau}^*\right| \sup_{\tau\in[0,t]}\left[\left|K(0,0)\right| + \frac{1}{N}\sum_{k=1}^N \left|V_{k,\tau}^* - V_{i,\tau}^*\right| + \left|X_{k,\tau}^* - X_{i,\tau}^*\right|\right] d\tau \\
&\le \int_0^t \left|u_{N,\tau}^j - u_{N,\tau}^*\right| \max_{i\le N}\sup_{\tau\in[0,t]}\left[2\left|V_{i,\tau}^*\right| + 2\left|X_{i,\tau}^*\right| + \left|K(0,0)\right|\right] d\tau.
\end{aligned} \tag{7.5.4}$$

Next, for every $k \in \mathbb{N}$ we set

$$A_k := \left\{\omega \in \Omega : \max_{i\le N}\sup_{t\in[0,T]}\left[\left|V_{i,t}^*\right| + \left|X_{i,t}^*\right|\right] \le k\right\}.$$

Given that $(V_{i,t}^*, X_{i,t}^*)_{i\le N}$ has finite p-th moment, we deduce $\mathbb{P}\left(\bigcup_k A_k\right) = 1$.

Let $\omega \in A_k$ such that (7.5.2) holds, by (7.5.3) and (7.5.4) we infer

$$\begin{aligned}\int_{A_k} \max_{i\leq N} \sup_{t\in[0,T]} \left[\left|V_{i,t}^j - V_{i,t}^*\right| + \left|X_{i,t}^j - X_{i,t}^*\right|\right] d\mathbb{P} \\ \leq 2(1+T)\bar{u}\int_0^t \int_{A_k} \max_{i\leq N} \sup_{\tau\in[0,t]} \left[\left|V_{i,\tau}^j - V_{i,\tau}^*\right| + \left|X_{i,\tau}^j - X_{i,\tau}^*\right|\right] d\mathbb{P} d\tau \\ + (1+T)\Big(2\,k + \left|K(0,0)\right|\Big) \int_0^t \left|u_{N,\tau}^j - u_{N,\tau}^*\right| d\tau,\end{aligned}$$

as the moment bound allows us to apply Fubini and switch the integration order. Hence, the proof of the claim (7.5.1) follows by Gronwall lemma, since $u_N^j \overset{*}{\rightharpoonup} u_N^*$ as $j \to \infty$.
We conclude with the proof of the proposition. Since the stochastic process $\left(V_{i,t}^j, X_{i,t}^j\right)_{i\leq N}$ has finite p-th order moment, by (7.1.4) in **(J1)** and the a.s. convergence (7.5.1) we get

$$\mathbb{E}\left[\int_0^T h_t\Big(V_{i,t}^j, X_{i,t}^j\Big)dt + g\Big(V_{i,T}^j, X_{i,T}^j\Big)\right] \xrightarrow[j\to\infty]{} \mathbb{E}\left[\int_0^T h_t\Big(V_{i,t}^*, X_{i,t}^*\Big)dt + g\Big(V_{i,T}^*, X_{i,T}^*\Big)\right], \tag{7.5.5}$$

thanks to the dominated convergence theorem. Moreover, by the lower semicontinuity of the L^2 norm it holds

$$\int_0^T \left|u_{N,t}^*\right|^2 dt \leq \liminf_{j\to\infty} \int_0^T \left|u_{N,t}^j\right|^2 dt. \tag{7.5.6}$$

Therefore, by (7.5.5), (7.5.6) and the subadditivity of the $\liminf$ we conclude that

$$J_N\Big(\Big(V_{i,t}^*, X_{i,t}^*\Big), u_N^*\Big) \leq \liminf_{j\to\infty} J_N\Big(\Big(V_{i,t}^j, X_{i,t}^j\Big), u_N^j\Big),$$

inferring that u_N^* is a minimizer for J_N. □

Our next aim is to prove the propagation of chaos: to this end we start by proving the existence of a unique strong solution to **(SDE)**.

Proposition 7.5.14. *For every $u \in \mathcal{U}$ there exists an unique strong solution $\Big(V_t, X_t\Big)$ to **(SDE)**. Moreover, there exists $C_T > 0$ such that*

$$\mathbb{E}\left[\sup_{t\leq T}\left|\Big(V_t, X_t\Big)\right|^p\right] \leq C_T\left(1 + \mathbb{E}\left[\left|\Big(V_0, X_0\Big)\right|^p\right]\right). \tag{7.5.7}$$

Proof. The scheme of the proof is classical in the theory of McKean-Vlasov SDE (see for instance [114]), we give it here for the sake of completeness.

Let $\Phi : \mathcal{P}(C) \to \mathcal{P}(C)$ be the map that to $m \in \mathcal{P}(C)$ assigns the law F of the solution to

$$\begin{cases} dV_t = \left(u_t \displaystyle\int_{\mathbb{R}^2} K\Big(v - V_t, x - X_t\Big) dm_t(v,x)\right) dt + \sqrt{2}\, dW_t, \\ dX_t = V_t\, dt, \\ \Big(V_0, X_0\Big) \sim F_0. \end{cases} \tag{7.5.8}$$

By **(K1)**, (7.5.8) is a linear SDE of the type (7.2.1) satisfying (7.2.2). Hence, by Theorem 7.2.6 we infer that there exists an unique strong solution to (7.5.8), verifying (7.5.7) thanks to **(D1)** . To prove the proposition, is sufficient to show that Φ admits a unique fixed point F. Let $m_1, m_2 \in \mathcal{P}(C)$, and denote as $\left(V_t^1, X_t^1\right)$ and $\left(V_t^2, X_t^2\right)$ the respective solutions to (7.5.8) with laws F_1 and F_2. Employing the Lipschitz assumption as in (7.5.3) and (7.5.4), and the Kantorovich duality (7.3.5), the following inequalities hold

$$
\begin{aligned}
&\left|V_t^1 - V_t^2\right| \\
&\leq \bar{u} \int_0^t \left| \int_{\mathbb{R}^2} K\Big(v - V_\tau^1, x - X_\tau^1\Big) dm_{1,\tau}(v,x) - \int_{\mathbb{R}^2} K\Big(v - V_\tau^2, x - X_\tau^2\Big) dm_{2,\tau}(v,x) \right| d\tau \\
&\leq \bar{u} \int_0^t \Big(\left|V_\tau^2 - V_\tau^1\right| + \left|X_\tau^2 - X_\tau^1\right| \Big) d\tau + \bar{u} \int_0^t \int_{\mathbb{R}^2} |(v,x)| (dm_{1,\tau} - dm_{2,\tau})(v,x) d\tau \\
&\leq \bar{u} \int_0^t \Big[\left|V_\tau^1 - V_\tau^2\right| + \left|X_\tau^1 - X_\tau^2\right| \Big] d\tau + \bar{u} \int_0^t \mathcal{W}_\tau(m_1, m_2) d\tau, \\
&\left|X_t^1 - X_t^2\right| \leq \int_0^t \left|V_\tau^1 - V_\tau^2\right| d\tau.
\end{aligned}
$$

Summing up these inequalities, by Gronwall lemma it follows

$$
\Big[\left|V_t^1 - V_t^2\right| + \left|X_t^1 - X_t^2\right| \Big] \leq \bar{u}\, \mathrm{e}^{2\,\bar{u}\,t} \int_0^t \mathcal{W}_\tau(m_1, m_2) d\tau.
$$

From the definition of the Wasserstein distance (7.3.4), we infer

$$
\mathcal{W}_t\Big(\Phi(m_1), \Phi(m_2)\Big) \leq \left|V_t^1 - V_t^2\right| + \left|X_t^1 - X_t^2\right| \leq \bar{u}\, \mathrm{e}^{2\,\bar{u}\,t} \int_0^t \mathcal{W}_\tau(m_1, m_2) d\tau. \tag{7.5.9}
$$

By iterating this inequality, for every $k \in \mathbb{N}$ we get

$$
\sup_{t \in [0,T]} \Big[\left|V_t^1 - V_t^2\right| + \left|X_t^1 - X_t^2\right| \Big] \leq \frac{\Big(\bar{u}\, \mathrm{e}^{2\,\bar{u}\,T} T^k\Big)^k}{k!} \sup_{t \in [0,T]} \mathcal{W}_t(\Phi(m), m), \qquad \forall m \in \mathcal{P}(C).
$$

From (7.3.2) and (7.3.4) we conclude that

$$
\mathcal{W}_{[0,T]}\Big(\Phi^{k+1}(m), \Phi^k(m)\Big) \leq \frac{\Big(\bar{u}\, \mathrm{e}^{2\,\bar{u}\,T} T^k\Big)^k}{k!} \mathcal{W}_{[0,T]}\Big(\Phi(m), m\Big),
$$

inferring that Φ is a Cauchy sequence in $\mathcal{P}(C)$, that equipped with the Wasserstein distance is a Banach space (see Theorem 3.2 in [51]). Therefore, we deduce the existence of an unique fixed point $F = \Phi(F)$, such that the relevant stochastic process satisfies (7.5.7). $\square$

The next step for the proof of Theorem 7.1.1 is the propagation of chaos.

Proposition 7.5.15. *Let $(u_N)_{N\in\mathbb{N}} \in \mathcal{U}$ such that $u_N \stackrel{*}{\rightharpoonup} u$. Then, the unique strong solution*

*$\Big(V_{i,t}, X_{i,t}\Big)_{i\le N}$ to (N-**SDEs**) with control u_N converges in the sense of propagation of chaos to the unique strong solution $\Big(V_t, X_t\Big)$ to **(SDE)** with control u.*

Proof. Let F^{N,u_N} denote the empirical measure of the unique strong solution $\Big(V_{i,t}, X_{i,t}\Big)_{i\le N}$ to (N-**SDEs**) with control u_N, and let F^u be the law of the unique strong solution $\Big(V_t, X_t\Big)$ to **(SDE)** with control u. According to (7.3.3), we need to prove that

$$\mathbb{E}\Big[\mathcal{W}_{[0,T]}\Big(F^{N,u_N}, F^u\Big)\Big] \xrightarrow[N\to\infty]{} 0.$$

Let $\Big(\overline{V}_{i,t}, \overline{X}_{i,t}\Big)_{i\le N}$ be a set of N i.i.d. copies of $\Big(V_t, X_t\Big)$, and denote as $F^{N,u}$ its empirical measure. By Theorem 1 in [58] we know that

$$\mathbb{E}\Big[\mathcal{W}_{[0,T]}\Big(F^{N,u}, F^u\Big)\Big] \le C\left(\mathbb{E}\Big[\Big|(V_t, X_t)\Big|^p\Big]\right)^{\frac{1}{p}}\frac{\ln(1+N)}{\sqrt{N}}, \tag{7.5.10}$$

for some $C > 0$ depending on p. From (7.5.7) and **(D1)**, the expectation in the right-hand side is bounded. Hence, it is sufficient to show that

$$\lim_{N\to\infty}\mathbb{E}\Big[\mathcal{W}_{[0,T]}\Big(F^{N,u_N}, F^{N,u}\Big)\Big] = 0.$$

Reasoning as in (7.5.2),(7.5.3) and (7.5.4) we get

$$\begin{aligned}
\Big|V_{i,t} - \overline{V}_{i,t}\Big| &\le \bar{u}\int_0^t \frac{1}{N}\sum_{k=1}^N \Big|K\Big(V_{k,\tau} - V_{i,\tau}, X_{k,\tau} - X_{i,\tau}\Big) - K\Big(\overline{V}_{k,\tau} - \overline{V}_{i,\tau}, \overline{X}_{k,\tau} - \overline{X}_{i,\tau}\Big)\Big| d\tau\\
&\quad + \bar{u}\int_0^t \Big|\int_{\mathbb{R}^2} K\Big(v - \overline{V}_\tau, x - \overline{X}_\tau\Big)\Big(dF^{N,u}_\tau - dF^u_\tau\Big)(v,x)\Big|\, dt\\
&\quad + \int_0^t \Big|u_{N,\tau} - u_\tau\Big|\,\Big|\int_{\mathbb{R}^2} K\Big(v - \overline{V}_{i,\tau}, x - \overline{X}_{i,\tau}\Big) dF^u_\tau(v,x)\Big|\, d\tau\\
&\le 2\bar{u}\int_0^t \max_{i\le N}\sup_{\tau\in[0,t]}\Big[\Big|V_{i,\tau} - \overline{V}_{i,\tau}\Big| + \Big|X_{i,\tau} - \overline{X}_{i,\tau}\Big|\Big] d\tau\\
&\quad + \bar{u}\int_0^t \Big|\int_{\mathbb{R}^2} |(v,x)|(dF^{N,u}_\tau - dF^u_\tau)(v,x)\Big|\, dt\\
&\quad + \int_0^t \Big|u_{N,\tau} - u_\tau\Big|\Big(\Big|\overline{V}_{i,\tau}\Big| + \Big|\overline{X}_{i,\tau}\Big| + \mathbb{E}\Big[\Big|\overline{V}_\tau\Big| + \Big|\overline{X}_\tau\Big|\Big]\Big) dt\\
&\le 2\bar{u}\int_0^t \max_{i\le N}\sup_{\tau\in[0,t]}\Big[\Big|V_{i,\tau} - \overline{V}_{i,\tau}\Big| + \Big|X_{i,\tau} - \overline{X}_{i,\tau}\Big|\Big] d\tau + \bar{u}\, T\, \mathcal{W}_{[0,T]}\Big(F^{N,u}, F^u\Big)\\
&\quad + \left(\sup_{t\in[0,T]}\Big[\Big|\overline{V}_{i,t}\Big| + \Big|\overline{X}_{i,t}\Big| + \mathbb{E}\Big[\sup_{t\in[0,T]}\Big|\overline{V}_t\Big| + \Big|\overline{X}_t\Big|\Big]\Big]\right)\int_0^t \Big|u_{N,\tau} - u_\tau\Big| d\tau,\\
\Big|X_{i,t} - \overline{X}_{i,t}\Big| &\le \int_0^t \Big|V_{i,\tau} - \overline{V}_{i,\tau}\Big| d\tau.
\end{aligned} \tag{7.5.11}$$

Since $\Big(V_{i,t},X_{i,t}\Big)_{i\le N}$ and $\Big(\overline{V}_{i,t},\overline{X}_{i,t}\Big)_{i\le N}$ have finite p-th moment, by (7.5.11) there exists $C_T>0$ such that

$$\begin{aligned}\mathbb{E}\left[\max_{i\le N}\sup_{t\in[0,T]}\Big[|V_{i,t}-\overline{V}_{i,t}|+|X_{i,t}-\overline{X}_{i,t}|\Big]\right]&\\ \le 2\bar u\int_0^t\mathbb{E}\left[\max_{i\le N}\sup_{\tau\in[0,t]}\Big|V_{i,\tau}-\overline{V}_{i,\tau}\Big|+\Big|X_{i,\tau}-\overline{X}_{i,\tau}\Big|\right]d\tau&\\ +\bar u\,T\,\mathbb{E}\Big[\mathcal{W}_{[0,T]}\Big(F^{N,u},F^u\Big)\Big]+C_T\int_0^t\Big|u_{N,\tau}-u_\tau\Big|d\tau.&\end{aligned}$$

Arguing as in (7.5.9) and relying on Gronwall lemma we get

$$\mathbb{E}\Big[\mathcal{W}_{[0,T]}\Big(F^{N,u_N},F^{N,u}\Big)\Big]\le \mathrm{e}^{2\,\bar u\,T}\left[\bar u\,T\,\mathbb{E}\Big[\mathcal{W}_{[0,T]}(F^{N,u},F^u)\Big]+C_T\int_0^t\Big|u_{N,\tau}-u_\tau\Big|d\tau\right]. \tag{7.5.12}$$

Since the right-hand side vanishes as N goes to infinity (given $u_N\overset{*}{\rightharpoonup}u$ as $N\to\infty$ and (7.5.10)), the thesis follows. $\square$

At this stage, we are in position to proceed with the proof of Theorem 7.1.1.

Proof of Theorem 7.1.1. We divide the proof in several steps. In the first two we prove the Γ-convergence, in the third we show the convergence of a subsequence of minimizers of **(P$_N$)** to a solution of **(P)**, and in the fourth we prove (7.1.5).

Step 1 (Liminf inequality). We start by proving the asymptotic lower bound

$$J\Big(\Big(V_t,X_t\Big),u\Big)\le\liminf_{N\to\infty}J_N\Big(\Big(V_{i,t},X_{i,t}\Big),u_N\Big),\qquad\text{for every }\ u_N\overset{*}{\rightharpoonup}u, \tag{7.5.13}$$

where $\Big(V_{i,t},X_{i,t}\Big)_{i\le N}$ solves **(N-SDEs)** with control u_N and $\Big(V_t,X_t\Big)$ solves **(SDE)** with control u.

Since the L^2 norm is weakly sequentially lower semi continuous, the following inequality holds

$$\int_0^T|u_t|^2dt\le\liminf_{N\to\infty}\int_0^T|u_{N,t}|^2dt.$$

By the subadditivity of the $\liminf$, it remains to show a similar result for the term involving the running cost h and the final cost g. To this end we prove a stronger result, namely

$$\lim_{N\to\infty}\sum_{i=1}^N\frac{1}{N}\mathbb{E}\left[\int_0^T h_t\Big(V_{i,t},X_{i,t}\Big)dt+g\Big(V_{i,T},X_{i,T}\Big)\right]=\mathbb{E}\left[\int_0^T h_t\Big(V_t,X_t\Big)dt+g\Big(V_T,X_T\Big)\right]. \tag{7.5.14}$$

This follows from an argument based on empirical measures, together with the propagation of chaos result stated in Proposition 7.5.15: let F^N denote the empirical measure of

$(V_{i,t}, X_{i,t})_{i\le N}$, the N particles expectation can be written using the measure F^N as follows

$$\begin{aligned}\sum_{i=1}^{N}\frac{1}{N}\mathbb{E}\left[\int_0^T h_t\Big(V_{i,t},X_{i,t}\Big)dt+g\Big(V_{i,T},X_{i,T}\Big)\right]\\ =\mathbb{E}\left[\int_0^T\int_{\mathbb{R}^2}h_t(v,x)dF_t^N(v,x)dt+g(v,x)dF_T^N(v,x)\right].\end{aligned}$$

Analogously, the mean field expectation can be expressed as follows

$$\mathbb{E}\left[\int_0^T h_t\Big(V_t,X_t\Big)dt+g\Big(V_T,X_T\Big)\right]=\int_{\mathbb{R}^2}\int_0^T h_t(v,x)dF_t(v,x)dt+\int_{\mathbb{R}^2}g(v,x)dF_T(v,x),$$

where F is the law of (V_t, X_t). Bearing these observations in mind, from (7.3.5) we infer the following inequality

$$\begin{aligned}&\left|\sum_{i=1}^{N}\frac{1}{N}\mathbb{E}\left[\int_0^T h_t\Big(V_{i,t},X_{i,t}\Big)dt+g\Big(V_{i,T},X_{i,T}\Big)\right]-\mathbb{E}\left[\int_0^T h_t\Big(V_t,X_t\Big)dt+g\Big(V_T,X_T\Big)\right]\right|\\ &\qquad\le\mathbb{E}\left[\int_0^T\int_{\mathbb{R}^2}|h_t(v,x)|\Big(dF_t^N-dF_t\Big)(v,x)dt+\int_{\mathbb{R}^2}|g(v,x)|\Big(dF_T^N-dF_T\Big)(v,x)\right]\\ &\qquad\le(1+T)\mathbb{E}\Big[\mathcal{W}_{[0,T]}\left(F^N,F\right)\Big].\end{aligned}\tag{7.5.15}$$

By Proposition 7.5.15 we know that the last expectation vanishes as N goes to infinity, and therefore we conclude the validity of (7.5.14).

Step 2 (Existence of a recovery sequence and Γ*-convergence).* Next, given any $u\in\mathcal{U}$ we prove the existence of a sequence $(u_N)_{N\in\mathbb{N}}\in\mathcal{U}$ (a recovery sequence) such that $u_N\overset{*}{\rightharpoonup}u$, and

$$\lim_{N\to\infty}J_N\Big(\Big(V_{i,t},X_{i,t}\Big),u_N\Big)=J\Big(\Big(V_t,X_t\Big),u\Big).\tag{7.5.16}$$

This holds by choosing the trivial sequence $u_N = u$ for all $N\in\mathbb{N}$. Indeed, let $(V_{i,t}, X_{i,t})_{i\le N}$ and (V_t, X_t) be the solutions respectively of **(N-SDEs)** and **(SDE)** with control u: by the same computation of *Step 1.* we infer

$$\begin{aligned}J\Big(\Big(V_t,X_t\Big),u\Big)&=\int_0^T|u_t|^2dt+\mathbb{E}\left[\int_0^T h_t\Big(V_t,X_t\Big)dt+g\Big(V_T,X_T\Big)\right]\\ &=\int_0^T|u_{N,t}|^2dt+\lim_{N\to\infty}\sum_{i=1}^{N}\frac{1}{N}\mathbb{E}\left[\int_0^T h_t\Big(V_{i,t},X_{i,t}\Big)dt+g\Big(V_{i,T},X_{i,T}\Big)\right]\\ &=\lim_{N\to\infty}J_N\Big(\Big(V_{i,t},X_{i,t}\Big),u_N\Big).\end{aligned}$$

Finally, combining the asymptotic lower bound (7.5.13) proved in *Step 1.* and the equality (7.5.16) of *Step 2.* the proof of the Γ-convergence follows (we refer the interested reader to the monograph [45] for more details about the Γ-convergence).

Step 3 (Convergence of minimizers). In this step we prove the existence of a minimizer to **(P)** by employing the Γ-convergence result previously established.
Let $(u_N^*)_{N\in\mathbb{N}} \in \mathcal{U}$ be a sequence of minimizers to **(P$_N$)**. By compactness, there exists a subsequence of minimizers (still denoted as $(u_N)_{N\in\mathbb{N}}$) such that u_N^* weak* converges to some $u^* \in \mathcal{U}$. We claim that u^* is a solution to **(P)**.
Let us denote as $\left(V_{i,t}^*, X_{i,t}^*\right)_{i\leq N}$ the solution of **(N-SDEs)** with control u_N^* and $\left(V_t^*, X_t^*\right)$ the solution of **(SDE)** with control u^*. From the previous steps it follows

$$\begin{aligned}
&J\Big(\Big(V_t^*, X_t^*\Big), u^*\Big) \overset{\text{Step 1}}{\leq} \liminf_{N\to\infty} J_N\Big(\Big(V_{i,t}^*, X_{i,t}^*\Big), u_N^*\Big) \\
&\leq \limsup_{N\to\infty} J_N\Big(\Big(V_{i,t}^*, X_{i,t}^*\Big), u_N^*\Big) \leq \limsup_{N\to\infty} J_N\Big(\Big(V_{i,t}, X_{i,t}\Big), u\Big) \overset{\text{Step 2}}{=} J\Big(\Big(V_t, X_t\Big), u\Big),
\end{aligned}$$

for all $u \in U$ and $\left(V_{i,t}, X_{i,t}\right)_{i\leq N}$ solutions of **(N-SDEs)** with control u, where in the last equality we employed again the recovery sequence $(u_N)_{N\in\mathbb{N}} = u$. Thus, u^* is a minimizer for **(P)**.

Step 4 (Performance bound). Finally, we establish the performance bound (7.1.5). Let us denote as $\left(V_{i,t}^*, X_{i,t}^*\right)_{i\leq N}$ the solution to **(N-SDEs)** with control u^* and let F^N be its empirical measure. Let $\left(V_t^*, X_t^*\right)$ be the solution to **(SDE)** with control u^* and law F^*, and let $(\overline{V}_{i,t}^*, \overline{X}_{i,t}^*)_{i\leq N}$ be a set of i.i.d. copies of $\left(V_t^*, X_t^*\right)$ with empirical measure $\bar{F}^N$. Reasoning as in (7.5.12) the following inequality holds

$$\mathbb{E}\Big[\mathcal{W}_{[0,T]}\Big(F^N, F^*\Big)\Big] \leq C\,\mathbb{E}\Big[\mathcal{W}_{[0,T]}\Big(\bar{F}^N, F^*\Big)\Big],$$

for some $C > 0$ that depends only on p, T and $\bar{u}$. Finally, arguing as in (7.5.15) we conclude

$$\begin{aligned}
\Big|J_N(u^*) - J(u^*)\Big| &\leq (1+T)\mathbb{E}\Big[\mathcal{W}_{[0,T]}(F^N, F^*\Big)\Big] \\
&\leq \widetilde{C}\,\mathbb{E}\Big[\mathcal{W}_{[0,T]}\Big(\bar{F}^N, F^*\Big)\Big] \leq \bar{C}\,\frac{\ln(1+N)}{\sqrt{N}},
\end{aligned}$$

where the last inequality follows again from Theorem 1 in [58] and Proposition 7.5.14. □

7.6 Proof of Theorem 7.1.4

Throughout this section we assume that **(B)** holds.

Having established the connection between the N particles optimal control problem and its

mean field counterpart, we now employ the PDE relevant to the mean field dynamic to derive the first order optimality conditions stated in Theorem 7.1.4.
For reader's convenience, we recall that, if the law F of the solution to **(SDE)** admits density f, this solves the following PDE

$$\begin{cases} Yf_t(v,x) = \partial^2_{vv} f_t(v,x) - \partial_v \left(u_t \int_{\mathbb{R}^2} \mathcal{K}(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \, f_t(v,x) \right), \\ f_{|t=0} = f_0, \end{cases} \tag{7.6.1}$$

where $Yf = (\partial_t + v\partial_x)f$. The optimal control cost functional is

$$J(f,u) = \int_{\mathbb{R}^2} \int_0^T h_t(v,x) f_t(v,x) dt dv dx + \int_{\mathbb{R}^2} g(v,x) f_T(v,x) dv dx + \frac{\nu}{2} \int_0^T |u_t|^2 dt.$$

The strategy for the proof of Theorem 7.1.4 is the following. We first prove that for every $u \in \mathcal{U}$ there exists an unique strong solution f to (7.6.1) with control function u. We express then the cost functional as $J(f,u) = J(G(u),u) = \widetilde{J}(u)$, for some map G. We show that G is Fréchet differentiable, and by chain rule we compute the Fréchet derivative of $\widetilde{J}$. Finally, we get the first order optimality condition (7.1.7) by expressing the Newton quotient of $\widetilde{J}$ in terms of f and the adjoint variable p, that solves the adjoint PDE

$$\begin{cases} Yp_t(v,x) + \partial^2_{vv} p_t(v,x) + u_t \partial_v p_t(v,x) \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \\ \qquad\qquad + u_t \int_{\mathbb{R}^2} K(v-\eta, x-\xi) \partial_v p_t(\eta,\xi) f_t(\eta,\xi) d\eta d\xi = -h_t(v,x), \\ p_{|t=T} = g. \end{cases} \tag{7.6.2}$$

This approach relies on the weak theory of degenerate Kolmogorov-Fokker-Planck PDEs discussed in Section 7.4.
We start by proving that, for every control $u \in \mathcal{U}$, the law F of the solution to **(SDE)** admits density f, which is a strong solution to (7.6.1) in the sense of the Definition 7.4.9.

Proposition 7.6.16. *For every $u \in \mathcal{U}$ there exists an unique strong solution f to* (7.6.1) *that belongs to W_T and that is the density of the solution to **(SDE)**.*

Proof. Since u is bounded and K is Lipschitz continuous and bounded, Theorem 1.4 in [7] guarantees that the law $F \in \mathcal{P}(C)$ of the solution to **(SDE)** is the unique measure solution to (7.6.1), namely it solves

$$\begin{aligned} \int_{\mathbb{R}^2} \varphi(v,x) dF_T(v,x) - \int_{\mathbb{R}^2} \varphi(v,x) dF_0(v,x) = \int_0^T \int_{\mathbb{R}^2} \partial^2_{vv} \varphi(v,x) dF_t(v,x) dt \\ + \int_0^T \int_{\mathbb{R}^2} \partial_v \varphi(v,x) u_t \int_{\mathbb{R}^2} K(\eta - v, \xi - x) dF_t(\eta,\xi) dF_t(v,x) dt, \qquad \forall \varphi \in C_c^\infty(\mathbb{R}^2). \end{aligned} \tag{7.6.3}$$

We define the following functions

$$b_t(v,x) := u_t \int_{\mathbb{R}^2} K(\eta-v,\xi-x)dF_t(\eta,\xi), \quad c_t(v,x) := u_t \int_{\mathbb{R}^2} \partial_v K(\eta-v,\xi-x)dF_t(\eta,\xi).$$

Since u belongs to $\mathcal{U} = L^\infty\Big([0,T];[u_{\min},u_{\max}]\Big)$, these functions are bounded and Lipschitz continuous (w.r.t. the Euclidean metric) in the v and x variables, and measurable and bounded in the t variable (thanks to **(K1)** and **(K2)**). Hence, it is straightforward to deduce that Theorem 5.0.7 applies, implying the existence of a fundamental solution Γ to

$$Yf_t(v,x) - \partial^2_{vv} f_t(v,x) + b_t(v,x)\partial_v f_t(v,x) - c_t(v,x) f_t(v,x) = 0. \tag{7.6.4}$$

This means that for every $1 \le q < +\infty$, and for every $f_0 \in L^q(\mathbb{R}^2)$ the function

$$f_t(v,x) := \int_{\mathbb{R}^2} \Gamma_t(v-\eta, x-\xi-t\eta) f_0(\eta,\xi) d\eta d\xi, \tag{7.6.5}$$

is well defined, solves in strong sense (7.6.4) (according to Definition 7.4.9), and satisfies $f_{|t=0} = f_0$ in L^q sense. We aim to show that f is the function described in the statement.
By Theorem 5.0.7, the fundamental solution Γ and its derivative $\partial_v\Gamma$ satisfy the following Gaussian estimates: for every $\varepsilon > 0$ there exist $C_1, C_2, C_3 > 0$ (only depending on $\bar{u}, T$ and ε) such that

$$\begin{aligned} \Gamma_t(v,x) &\ge \frac{C_1}{t^2} \exp\left(-\frac{1}{1-\varepsilon} \left(\frac{v^2}{t} - \frac{3vx}{t^2} + \frac{3x^2}{t^3} \right) \right), \\ \Gamma_t(v,x) &\le \frac{C_2}{t^2} \exp\left(-\frac{1}{1+\varepsilon} \left(\frac{v^2}{t} - \frac{3vx}{t^2} + \frac{3x^2}{t^3} \right) \right), \\ \Big|\partial_v \Gamma_t(v,\xi)\Big| &\le \frac{C_3}{t^2\sqrt{t}} \exp\left(-\frac{1}{1+\varepsilon} \left(\frac{v^2}{t} - \frac{3vx}{t^2} + \frac{3x^2}{t^3} \right) \right). \end{aligned} \tag{7.6.6}$$

From the Gaussian upper bounds we conclude that $f \in W_T$, and that for every $t > 0$ both $f_t(v,x)$ and $\partial_v f_t(v,x)$ vanish as v and x tend to $\pm\infty$.
We now prove that f is a probability density function, namely it is positive and normalized for every $t \ge 0$. The positivity follows from the Gaussian lower bound and the representation formula (7.6.5), since f_0 is a probability density function. To prove the normalization property, we recall that, as f solves (7.6.1) in strong sense, it has strong Lie derivative according to (7.4.3). Integrating this relation yields

$$\begin{aligned} \int_{\mathbb{R}^2} f_t(v,x)dvdx &= \int_{\mathbb{R}^2} f_t(v,x+v(s-t))dvdx \\ &= \int_{\mathbb{R}^2} f_s(v,x)dvdx + \int_s^t \int_{\mathbb{R}^2} \partial_v \Big(-\partial_v f_\tau(v,x+v(s-\tau)) \\ &\quad + u_\tau \int_{\mathbb{R}^2} K(\eta-v,\xi-x-v(s-\tau))dF_t(\eta,\xi) f_\tau(v,x+v(s-\tau)) \Big) dvdxd\tau, \end{aligned} \tag{7.6.7}$$

for every $0 < s < t < T$. Since $f_\tau(v, x + v(s - \tau))$ and $\partial_v f_\tau(v, x + v(s - \tau))$ vanish as $|v|$ goes to infinity, the last integral in (7.6.7) vanishes. From (7.6.7) we infer that the L^1 norm of f is constant, then the normalization property follows from $f_{|t=0} = f_0$ in L^1 sense.
To conclude the proof of the proposition we need to show that f is the density of F. To this end, we define a measure $\widetilde{F}$ whose density is f, and we claim that $\widetilde{F} = F$. To prove the claim we show that $\widetilde{F}$ is a distributional solution in the sense of (7.6.3), then by the distributional uniqueness we conclude.
Since f is a strong solution to (7.6.4), we only need to prove that $\widetilde{F}_t$ has first moment bounded for every $t \in [0, T]$. To this end, we begin by noticing that

$$
\begin{aligned}
&\int_{\mathbb{R}^2} \Big(|v| + |x|\Big) d\widetilde{F}_t(v, x) \\
&\qquad = \int_{\mathbb{R}^2} \Big(|v| + |x|\Big) f_t(v, x) dv dx \\
&\qquad\qquad = \int_{\mathbb{R}^2} \Big(|v| + |x|\Big) \int_{\mathbb{R}^2} \Gamma_t(v - \eta, x - \xi - t\eta) f_0(\eta, \xi) d\eta d\xi dv dx.
\end{aligned}
$$

We can bound the last integral with the sum of the following two terms

$$
\begin{aligned}
I_1 &:= \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \Big(|v - \eta| + |x - \xi - t\eta|\Big) \Gamma_t(v - \eta, x - \xi - t\eta) f_0(\eta, \xi) d\eta d\xi dv dx, \\
I_2 &:= (T + 1) \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \Big(|\eta| + |\xi|\Big) \Gamma_t(v - \eta, x - \xi - t\eta) f_0(\eta, \xi) d\eta d\xi dv dx.
\end{aligned}
$$

Thanks to (7.6.6), it is not difficult to establish (see for instance Lemma 2.2 in [102]) that there exists c_1 (depending on $\bar{u}, T$ and ε) such that

$$
\begin{aligned}
&\Big(|v - \eta| + |x - \xi - t\eta|\Big) \Gamma_t(v - \eta, x - \xi - t\eta) \\
&\qquad \le \frac{c_1}{t^2} \exp\left(-\frac{1}{1 + \varepsilon} \left(\frac{(v - \eta)^2}{t} - \frac{3(v - \eta)(x - \xi - t\eta)}{t^2} + \frac{3(x - \xi - t\eta)^2}{t^3} \right) \right).
\end{aligned}
$$

The right-hand side belongs to $L^1(\mathbb{R}^2)$ for every $t > 0$. Therefore, since $f_0 \in L^1(\mathbb{R}^2)$, by Young's convolution inequality we conclude that $I_1 < +\infty$.
By a similar argument we conclude that I_2 is finite, since **(D1)** and **(D2)** implies that the measure F_0 with density f_0 has first moment bounded, and hence the map $(\xi, \eta) \mapsto (|\eta| + |\xi|) f_0(\eta, \xi)$ belongs to $L^1(\mathbb{R}^2)$. This concludes the proof of the claim, and hence of the proposition. $\square$

We now introduce the control map $G : \mathcal{U} \to W_T$, assigning to each $u \in \mathcal{U}$ the probability density function $f \in W_T$ strong solution to (7.6.1) with control u. In Proposition 7.6.16 we established that f is uniquely determined by u. In order to apply the chain rule to J, our goal is to prove that G is Fréchet differentiable. To establish this, we rely on the implicit function theorem on Banach spaces. Consider the map $\mathcal{G} : W_T \times \mathcal{U} \to L^2_{x,t}H^{-1}_v \times L^2(\mathbb{R}^2)$ defined by $\mathcal{G}(f,u) = (\mathcal{G}_1, \mathcal{G}_2)$, where

$$\begin{aligned} \mathcal{G}_{1,t}(v,x) &:= Y f_t(v,x) - \partial^2_{vv} f_t(v,x) + \partial_v\Big(u_t\int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi\, f_t(v,x)\Big) \\ \mathcal{G}_2(v,x) &:= f_0(v,x) - f_{|t=0}(v,x). \end{aligned} \tag{7.6.8}$$

Note that finding a weak solution to the Cauchy problem (7.6.1) is equivalent to solve the equation $\mathcal{G}(f,u) = 0$.
The Fréchet differentiability of G follows once we show that $\mathcal{G}$ is smooth and implicitly defines G. We start by proving its regularity.

Lemma 7.6.17. *The map $\mathcal{G}$ defined in* (7.6.8) *is smooth.*

Proof. The smoothness of $\mathcal{G}_2$ is immediate due to linearity. To prove the smoothness of $\mathcal{G}_1$, we introduce the auxiliary map $\mathcal{K} : W_T \times \mathcal{U} \to L^2\Big(\mathbb{R}^2 \times (0,T)\Big)$ defined by

$$\mathcal{K}_t(f,u)(v,x) := u_t \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi\; f_t(v,x) = u_t\, K \star f_t(v,x)\, f_t(v,x).$$

The definition of $\mathcal{G}_1$ implies

$$\begin{aligned} &\int_0^T \int_{\mathbb{R}} \langle \mathcal{G}_{1,t}(\cdot,x), \varphi_t(\cdot,x)\rangle dx dt \\ &\quad = \int_0^T \int_{\mathbb{R}} \langle Y f_t(\cdot,x), \varphi_t(\cdot,x)\rangle dx dt \\ &\qquad\quad + \int_0^T \int_{\mathbb{R}^2} \partial_v f_t(v,x) \partial_v \varphi(v,x) dv dx dt \\ &\qquad\qquad - \int_0^T \int_{\mathbb{R}^2} \mathcal{K}_t(f,u)(v,x) \partial_v \varphi(v,x) dv dx dt, \qquad \forall \varphi \in L^2_{x,t}H^1_v. \end{aligned}$$

The first two summands are smooth by linearity. The smoothness of the third term involving $\mathcal{K}$ follows from its cubic nature. Indeed, for every variation $g \in W_T$ and $w \in \mathcal{U}$ we have the Taylor expansion

$$\begin{aligned} \mathcal{K}_t(f+g,&u+w)(v,x) \\ &= u_t\, K \star f_t(v,x)\, f_t(v,x) \\ &+ u_t\, K \star f_t(v,x)\, g_t(v,x) + u_t\, K \star g_t(v,x)\, f_t(v,x) + w_t\, K \star f_t(v,x)\, f_t(v,x) \\ &+ u_t\, K \star g_t(v,x)\, g_t(v,x) + w_t\, K \star g_t(v,x)\, f_t(v,x) + w_t\, K \star f_t(v,x)\, g_t(v,x) \\ &+ w_t\, K \star g_t(v,x)\, g_t(v,x), \\ &\qquad\qquad = u_t\, K \star f_t(v,x)\, f_t(v,x) + L_t(v,x) + Q_t(v,x) + C_t(v,x), \end{aligned}$$

which decomposes into the linear term L, the quadratic one Q and the cubic term C, with respect to $g \in W_T$ and $w \in \mathcal{U}$. To prove the smoothness, we only need to prove the continuity of L, Q and C, but this is straightforward. Indeed, L is a linear bounded operator. Analogously, let us consider C: thanks to **(K2)** and (7.4.7), the Hölder and Young inequalities imply

$$\|C\|_{L^2(\mathbb{R}^2\times(0,T))} \leq \|w\|_{\mathcal{U}}\|K \star g_t\|_{L^\infty(\mathbb{R}^2\times(0,T))}\|g_t\|_{L^2(\mathbb{R}^2\times(0,T))} \leq c\|w\|_{\mathcal{U}}\|g\|^2_{W_T},$$

for some $c > 0$ that depends on T and $\|K\|_{L^2(\mathbb{R}^2)}$. The proof for the quadratic term Q is simpler, and therefore is omitted. This completes the proof. □

We now prove that $\mathcal{G}$ implicitly defines G. Hence, the previous lemma proves also its smoothness.

Lemma 7.6.18. *The map G, which assigns to each control function u the unique density function f strong solution to* (7.6.1) *with control u, is implicitly defined by $\mathcal{G}$. Therefore, G is smooth.*

Proof. Let $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2)$ be defined as in (7.6.8). Since $\mathcal{G}$ is smooth, we can compute its Fréchet derivative. For every $z \in W_T$ we obtain

$$\begin{aligned}
\partial_f\mathcal{G}_1(f,u)z_t(v,x) &= Yz_t(v,x) - \partial^2_{vv}z_t(v,x) \\
&\quad + u_t\partial_v\Big(\int_{\mathbb{R}^2} K(\eta-v,\xi-x)z_t(\eta,\xi)d\eta d\xi\, f_t(v,x) \\
&\quad + \int_{\mathbb{R}^2} K(\eta-v,\xi-x)f_t(\eta,\xi)d\eta d\xi\, z_t(v,x)\Big), \\
\partial_f\mathcal{G}_2(f,u)z_t &= z_{|t=0}.
\end{aligned}$$

In order to apply the implicit function theorem, we need to show that for every $\bar{z} \in L^2_{x,t}H^{-1}_v$ and $z_0 \in L^2(\mathbb{R}^2)$ there exists an unique $z \in W_T$ satisfying $\partial_f\mathcal{G}(f,u)z = (\bar{z}, z_0)$. Hence, we need to establish weak existence and uniqueness to the Cauchy problem

$$\begin{cases}
Yz_t(v,x) - \partial^2_{vv}z_t(v,x) + u_t\partial_v\Bigg(\displaystyle\int_{\mathbb{R}^2}K(\eta-v,\xi-x)z_t(\eta,\xi)d\eta d\xi\, f_t(v,x) \\
\qquad\qquad\qquad + \displaystyle\int_{\mathbb{R}^2}K(\eta-v,\xi-x)f_t(\eta,\xi)d\eta d\xi\, z_t(v,x)\Bigg) = \bar{z}_t(v,x), \\
z_{|t=0} = z_0.
\end{cases} \tag{7.6.9}$$

This implies that $\partial_f\mathcal{G}$ is an isomorphism between the Banach spaces W_T and $L^2_{x,t}H^{-1}_v \times L^2$. Once this is established, we are in position to apply the implicit function theorem, from which it follows that equation $\mathcal{G}(f,u) = 0$ implicitly defines the map G. Thanks to Lemma

7.6.17, we then infer its smoothness.
Therefore, to complete the proof of the lemma we show that there exists an unique weak solution to (7.6.9), *i.e.* $z \in W_T$ such that $z_{|t=0} = z_0$ in L^2 sense and

$$\begin{aligned}\int_0^T \int_{\mathbb{R}} \langle Y z_t(\cdot, x), \varphi_t(\cdot, x)\rangle dxdt \\ + \int_0^T \int_{\mathbb{R}^2} \partial_v z_t(v, x) \partial_v \varphi_t(v, x) dvdxdt \\ - \int_0^T \int_{\mathbb{R}^2} u_t \Bigg(\int_{\mathbb{R}^2} K(\eta - v, \xi - x) z_t(\eta, \xi) d\eta d\xi \, f_t(v, x) \\ + \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta, \xi) d\eta d\xi \, z_t(v, x) \Bigg) \partial_v \varphi_t(v, x) \\ = \int_0^T \int_{\mathbb{R}} \langle \bar{z}_t(\cdot, x), \varphi_t(\cdot, x)\rangle dxdt,\end{aligned} \tag{7.6.10}$$

for every $\varphi \in C_c^\infty\Big(\mathbb{R}^2 \times [0, T]\Big)$.
To this end, we rely on Theorem 5.2 in [16] (together with the remark of Section 5.3 in [16]). For every $(x, t) \in \mathbb{R} \times [0, T]$ we define the continuous bilinear form $a_{x,t} : H_v^1 \times H_v^1 \to \mathbb{R}$

$$\begin{aligned}a_{x,t}(g_1, g_2) :=& \int_{\mathbb{R}} \partial_v g_1 \partial_v g_2 dv \\ &- \int_{\mathbb{R}} u_t g_1 \Big(\hat{K} \star f_t \partial_v g_2\Big) dv - \int_{\mathbb{R}} u_t g_1 \partial_v g_2 (K \star f_t) dv,\end{aligned} \tag{7.6.11}$$

where $\hat{K} \star f \partial_v g_2$ and $K \star f$ denote the following convolutions

$$\begin{aligned}\Big(\hat{K} \star f_t \partial_v g_1\Big)(v, x) &:= \int_{\mathbb{R}^2} K(v - \eta, x - \xi) f_t(\eta, \xi) \partial_v g_2(\eta) d\eta d\xi, \\ (K \star f_t)(v, x) &:= \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta, \xi) d\eta d\xi.\end{aligned} \tag{7.6.12}$$

We claim that $a_{x,t}$ is uniformly bounded and weakly coercive, namely that there exist three positive constants Λ, λ, c_0 such that for every $(x, t) \in \mathbb{R} \times [0, T]$ and $g_1, g_2, g \in H_v^1(\mathbb{R})$ it holds

$$|a_{x,t}(g_1, g_2)| \leq \Lambda \|\partial_v g_1\|_{L^2(\mathbb{R})} \|\partial_v g_2\|_{L^2(\mathbb{R})}, \qquad a_{x,t}(g, g) \geq \lambda \|\partial_v g\|^2_{L^2(\mathbb{R})} - c_0 \|g\|^2_{L^2(\mathbb{R})}. \tag{7.6.13}$$

The first term in the bilinear form (7.6.11) is easily bounded as follows

$$\left| \int_{\mathbb{R}} \partial_v g_1 \partial_v g_2 \right| \leq \|\partial_v g_1\|_{L^2(\mathbb{R})} \|\partial_v g_2\|_{L^2(\mathbb{R})}, \qquad \int_{\mathbb{R}} (\partial_v g)^2 = \|\partial_v g\|^2_{L^2(\mathbb{R})}. \tag{7.6.14}$$

For the second and third term in (7.6.11), we rely on the regularity of K given by **(K2)**, and the normalization property of f for every $t \geq 0$ (as follows from Proposition 7.6.16). From the Hölder and Young inequalities we get

$$
\begin{aligned}
\left|\int_{\mathbb{R}} u_t g_1 (K \star f_t \partial_v g_2) dv\right| &\leq \bar{u}\|g_1\|_{L^2(\mathbb{R})}\|f_t\|_{L^2(\mathbb{R}^2)}\|K\|_{L^2(\mathbb{R}^2)}\|\partial_v g_2\|_{L^2(\mathbb{R})} \\
&\leq c_1\|g_1\|_{L^2(\mathbb{R})}\|\partial_v g_2\|_{L^2(\mathbb{R})}, \\
\left|\int_{\mathbb{R}} u_t g_1 \partial_v g_2 (K \star f_t) dv\right| &\leq \bar{u}\|g_1\|_{L^2(\mathbb{R})}\|\partial_v g_2\|_{L^2(\mathbb{R})}\|K\|_{L^\infty(\mathbb{R}^2)}\|f_t\|_{L^1(\mathbb{R}^2)} \\
&\leq c_2\|g_1\|_{L^2(\mathbb{R})}\|\partial_v g_2\|_{L^2(\mathbb{R})}, \\
-\int_{\mathbb{R}} u_t g (K \star f_t \partial_v g) dv &\geq -\frac{c_1}{2\varepsilon}\|g\|^2_{L^2(\mathbb{R})} - \frac{c_1\varepsilon}{2}\|\partial_v g\|^2_{L^2(\mathbb{R})}, \\
-\int_{\mathbb{R}} u_t g \partial_v g (K \star f_t) dv &\geq -\frac{c_2}{2\varepsilon}\|g\|^2_{L^2(\mathbb{R})} - \frac{c_2\varepsilon}{2}\|\partial_v g\|^2_{L^2(\mathbb{R})},
\end{aligned}
\tag{7.6.15}
$$

for every $\varepsilon > 0$. Here c_1 depends on $\bar{u}$, $\|K\|_{L^2(\mathbb{R}^2)}$ and $\sup_t \|f_t\|_{L^2(\mathbb{R}^2)}$ (that is bounded by $\|f\|_{W_T}$ thanks to the embedding (7.4.7)), while c_2 depends on $\bar{u}$ and $\|K\|_{L^\infty(\mathbb{R}^2)}$. Choosing a suitable $\varepsilon > 0$, by (7.6.14) and (7.6.15) we see that (7.6.13) holds. In order to apply Theorem 5.2 in [16], we need a coercive form. To this end, we pick some $c > c_0$ and define the continuous bilinear form $\bar{a}_{x,t} : H^1_v \times H^1_v \to \mathbb{R}$

$$
\bar{a}_{x,t}(g_1, g_2) := a_{x,t}(g_1, g_2) + c\int_{\mathbb{R}} g_1 g_2 dv.
$$

From (7.6.13) we infer that $\bar{a}_{x,t}$ is coercive, thus we apply Theorem 5.2 in [16] to the PDE relevant to $\bar{a}_{x,t}$. Hence, we conclude that there exists an unique $\bar{z} \in W_\infty$ that solves $z_{|t=0} = z_0$ in L^2 sense, and

$$
\begin{aligned}
&\int_0^{+\infty}\int_{\mathbb{R}} \langle Y\bar{z}_t(\cdot, x), \varphi_t(\cdot, x)\rangle dx dt \\
&+ \int_0^{+\infty}\int_{\mathbb{R}^2} \partial_v \bar{z}_t(v, x)\partial_v\varphi_t(v, x) dv dx dt \\
&+ \int_0^{+\infty}\int_{\mathbb{R}^2} \bar{c}\,\bar{z}_t(v, x)\varphi_t(v, x) dv dx dt \\
&- \int_0^{+\infty}\int_{\mathbb{R}^2} u_t\Bigg(\int_{\mathbb{R}^2} K(\eta - v, \xi - x)\bar{z}_t(\eta, \xi) d\eta d\xi\, f_t(v, x) \\
&+ \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta, \xi) d\eta d\xi\, \bar{z}_t(v, x)\Bigg)\partial_v\varphi_t(v, x) \\
&\qquad\qquad = \int_0^{+\infty}\int_{\mathbb{R}} \langle h_t(\cdot, x), \varphi_t(\cdot, x)\rangle dx dt,
\end{aligned}
$$

for every $\varphi \in C^\infty_c\Big(\mathbb{R}^2 \times [0, +\infty)\Big)$. We conclude with the existence and uniqueness of (7.6.10) by observing that the function $z = \bar{z}e^{ct}$ belongs to W_T for every $T < +\infty$ and solves (7.6.10). This completes the proof of the lemma. $\square$

We are now ready to introduce some optimality conditions for **(P)**, which will be made explicit in Theorem 7.1.4 in terms of the density f and the adjoint variable p. To this end, we introduce the reduced cost functional

$$\widetilde{J}(u) = J\Big(G(u), u\Big). \tag{7.6.16}$$

Proposition 7.6.19. *Let u be a solution to **(P)**. Then, for every $w \in \mathcal{U}$ it holds*

$$\widetilde{J}'(u)(w-u) \geq 0, \tag{7.6.17}$$

where $\widetilde{J}'(u)$ denotes the Fréchet derivative of the reduced cost functional (7.6.16).

Proof. Let $u \in \mathcal{U}$ be a minimizer, then there exists $\bar{\varepsilon} > 0$ such that

$$\widetilde{J}\Big(u + \varepsilon(w-u)\Big) - \widetilde{J}(u) \geq 0, \qquad \forall w \in \mathcal{U},$$

for every $\varepsilon < \bar{\varepsilon}$. By Lemma 7.6.18, we know that $\widetilde{J}$ is Fréchet differentiable. Thus we are allowed to divide by ε and passing to the limit, obtaining the thesis. $\square$

Theorem 7.1.4 turns the previous optimality conditions in a computable form, employing the adjoint variable p that solves the Cauchy problem (7.6.2) (which can be derived by means of the Lagrangian formalism, see for instance [116]). First, we need to show that there exists an unique weak solution to (7.6.2), namely $p \in W_T$ such that

$$\begin{aligned}
&\int_0^T \int_{\mathbb{R}} \langle Y p_t(\cdot, x), \varphi_t(\cdot, x)\rangle dx dt \\
&\quad - \int_0^T \int_{\mathbb{R}^2} \partial_v p_t(v,x) \partial_v \varphi_t(v,x) dv dx dt \\
&\qquad + \int_0^T \int_{\mathbb{R}^2} u_t \partial_v p_t(v,x)(K \star f_t)(v,x)\varphi_t(v,x) dv dx dt \\
&\qquad\quad + \int_0^T \int_{\mathbb{R}^2} u_t \Big(\hat{K} \star \partial_v p_t f_t\Big)(v,x)\varphi_t(v,x) dv dx dt \\
&\qquad\qquad = - \int_0^T h_t(v,x)\varphi_t(v,x) dv dx dt,
\end{aligned} \tag{7.6.18}$$

for every $\varphi \in C_c^\infty\Big(\mathbb{R}^2 \times [0,T]\Big)$.
This follows again by employing the weak theory of the degenerate Kolmogorov-Fokker-Planck equations developed in [16].

Proposition 7.6.20. *There exists an unique weak solution p to* (7.6.18).

Proof. The proof follows as in Lemma 7.6.18, keeping in mind Remark 1.10 in [16] regarding the backward equations. □

We are now ready to prove Theorem 7.1.4, where we express the optimality condition (7.6.17) as (7.1.7) by explicitly computing the derivative of the reduced cost functional (7.6.16).

Proof of Theorem 7.1.4. Let G be the map that associates to each control function $u \in \mathcal{U}$ the density $f \in W_T$. For every $w \in \mathcal{U}$ it holds

$$\widetilde{J}'(u)(w-u) = \partial_f J(G(u),u)G'(u)(w-u) + \partial_u J(G(u),u)(w-u), \tag{7.6.19}$$

where we recall that $\widetilde{J}$ is the reduced cost functional introduced in (7.6.16), and

$$J(f,u) = \int_{\mathbb{R}^2}\int_0^T h_t(v,x)f_t(v,x)dtdvdx + \int_{\mathbb{R}^2} g(v,x)f_T(v,x)dvdx + \frac{\nu}{2}\int_0^T |u_t|^2 dt.$$

Letting $G'(u)(w-u) = z$, by standard computation we get

$$\begin{aligned}\partial_f J(G(u),u)z &= \int_0^T \int_{\mathbb{R}^2} h_t(v,x)z_t(v,x)dvdxdt + \int_{\mathbb{R}^2} g(v,x)z_T(v,x)dvdx,\\ \partial_u J(G(u),u)(w-u) &= \nu\int_0^T u_t(w_t-u_t)dt.\end{aligned}$$

Since $u \in \mathcal{U}$ is a minimizer, by (7.6.17) and (7.6.19) we infer that

$$\int_0^T \int_{\mathbb{R}^2} h_t(v,x)z_t(v,x)dvdxdt + \int_{\mathbb{R}^2} g(v,x)z_T(v,x)dvdx + \nu\int_0^T u_t(w_t-u_t)dt \geq 0. \tag{7.6.20}$$

We recall that, by Lemma 7.6.18, G is implicitly defined by the map $\mathcal{G}$ introduced in (7.6.8). Hence, from the implicit function theorem, $G'(u)(w-u) =: z$ satisfies

$$\begin{aligned}&\int_0^T \int_{\mathbb{R}} \langle Y z_t(\cdot,x), \varphi_t(\cdot,x)\rangle dxdt\\ &+ \int_0^T \int_{\mathbb{R}^2} \partial_v z_t(v,x)\partial_v\varphi_t(v,x)dvdxdt\\ &\quad - \int_0^T \int_{\mathbb{R}^2} u_t\Big((K\star z_t)(v,x)f_t(v,x)\\ &\qquad + (K\star f_t)(v,x)z_t(v,x)\Big)\partial_v\varphi_t(v,x)dvdxdt\\ &\qquad\quad = \int_0^T \int_{\mathbb{R}^2} (K\star f_t)(v,x)f_t(v,x)\partial_v\varphi_t(v,x)(w_t-u_t)dvdxdt,\\ &\hspace{20em} \forall \varphi \in C_c^\infty\Big([0,T],\mathbb{R}^2\Big),\end{aligned} \tag{7.6.21}$$

with $z_{|_{t=0}} = 0$ in L^2 sense (here we recall the notation (7.6.12)). We now want to employ the adjoint equation (7.6.18) to express (7.6.20) in terms of the weak solution p of the adjoint equation, the density f, and the difference $(w-u)$. If we test (7.6.21) with p (see Remark 7.4.12), the integration by parts formula yields

$$
\begin{aligned}
&-\int_0^T \int_{\mathbb{R}} \langle Y p_t(\cdot,x), z_t(\cdot,x)\rangle dx dt \\
&\quad + \int_0^T \int_{\mathbb{R}^2} \partial_v z_t(v,x) \partial_v p_t(v,x) dv dx dt \\
&\quad\quad - \int_0^T \int_{\mathbb{R}^2} u_t \Big((K \star z_t)(v,x) f_t(v,x) \\
&\quad\quad\quad + (K \star f_t)(v,x) z_t(v,x) \Big) \partial_v p_t(v,x) dv dx dt \\
&\quad\quad\quad\quad = - \int_{\mathbb{R}^2} g(v,x) z_T(v,x) dv dx \\
&\quad\quad\quad\quad\quad + \int_0^T \int_{\mathbb{R}^2} f_t(v,x) \partial_v p_t(v,x) (K \star f_t)(v,x)(w_t - u_t) dv dx dt,
\end{aligned}
\tag{7.6.22}
$$

where we employed the final condition $p_{|t=T} = g$ and the initial condition $z_{|t=0} = 0$.
We now test the weak formulation of the adjoint equation (7.6.18) with z (see again Remark 7.4.12), obtaining

$$
\begin{aligned}
&\int_0^T \int_{\mathbb{R}} \langle Y p_t(\cdot,x), z_t(\cdot,x)\rangle dx dt \\
&\quad - \int_0^T \int_{\mathbb{R}^2} \partial_v p_t(v,x) \partial_v z_t(v,x) dv dx dt \\
&\quad\quad + \int_0^T \int_{\mathbb{R}^2} u_t \partial_v p_t(v,x) (K \star f_t)(v,x) z_t(v,x) dv dx dt \\
&\quad\quad\quad + \int_0^T \int_{\mathbb{R}^2} u_t \Big(\hat{K} \star \partial_v p_t f_t \Big)(v,x) z_t(v,x) dv dx dt \\
&\quad\quad\quad\quad = - \int_0^T h_t(v,x) z_t(v,x) dv dx dt.
\end{aligned}
\tag{7.6.23}
$$

Summing up (7.6.22) and (7.6.23) we get

$$
\begin{aligned}
&\int_0^T \int_{\mathbb{R}^2} h_t(v,x) z_t(v,x) dv dx dt \\
&\quad + \int_{\mathbb{R}^2} g(v,x) z_T(v,x) dv dx \\
&\quad\quad = \int_0^T \int_{\mathbb{R}^2} f_t(v,x) \partial_v p_t(v,x) (K \star f_t)(v,x)(w_t - u_t) dv dx dt.
\end{aligned}
\tag{7.6.24}
$$

By (7.6.24) and (7.6.20) we conclude that for the minimizer u it holds

$$
\begin{aligned}
&\int_0^T \int_{\mathbb{R}^2} \Big(f_t(v,x) \partial_v p_t(v,x) \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \Big)(w_t - u_t) dv dx dt \\
&\qquad\qquad + \int_0^T \nu\, u_t (w_t - u_t) dt \geq 0, \qquad \forall w \in \mathcal{U}.
\end{aligned}
$$

This completes the proof of the theorem. □

7.7 Numerical simulations

In this section we present some numerical simulations illustrating the evolution of the density of the stochastic process solving **(SDE)**. This evolution occurs under the action of the optimal control, given by the solution to **(P)**. It is worth noticing that controlling the density means influencing every possible realization of the stochastic process described by **(SDE)**.
For reader's convenience, we recall that the density f solves

$$\begin{cases} Yf_t(v,x) = \partial^2_{vv} f_t(v,x) - \partial_v \left(u_t \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \, f_t(v,x) \right), \\ f_{|t=0} = f_0, \end{cases} \tag{7.7.1}$$

and the solution u to **(P)** minimizes over $\mathcal{U}$ the functional

$$\begin{aligned} \widetilde{J}(u) &= J(f,u) \\ &= \int_{\mathbb{R}^2} \int_0^T h_t(v,x) f_t(v,x) dt dv dx + \int_{\mathbb{R}^2} g(v,x) f_T(v,x) dv dx + \frac{\nu}{2} \int_0^T |u_t|^2 dt. \end{aligned} \tag{7.7.2}$$

We numerically minimize (7.7.2) by using a gradient descent method. To this end, we recall that in Theorem 7.1.4 we formally established that

$$\nabla \widetilde{J}(u) = \nu u_t + \int_{\mathbb{R}^2} f_t(v,x) \partial_v p_t(v,x) \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \, dv dx,$$

where f is the density of the solution to **(SDE)** corresponding to u (that is, the solution to (7.7.1) relevant to the control function u), and the adjoint variable p is the solution to

$$\begin{cases} Yp_t(v,x) + \partial^2_{vv} p_t(v,x) + u_t \partial_v p_t(v,x) \int_{\mathbb{R}^2} K(\eta - v, \xi - x) f_t(\eta,\xi) d\eta d\xi \\ \qquad\qquad + u_t \int_{\mathbb{R}^2} K(v - \eta, x - \xi) \partial_v p_t(\eta,\xi) f_t(\eta,\xi) d\eta d\xi = -h_t(v,x), \\ p_{|t=T}(v,x) = g(v,x). \end{cases} \tag{7.7.3}$$

In order to compute the solutions f and p to (7.7.1) and (7.7.3), we employ a discretization method based on the finite difference scheme introduced in [103], and refined in Chapter 6 for the kinetic mean field model under study. We recall that this particular scheme is well suited for handling degenerate Kolmogorov-Fokker-Planck equations, since it effectively discretizes the Lie derivative with the following first order approximation

$$Yf_t(v,x) \approx -\frac{f_{t-\Delta t}(v, x - v\Delta t) - f_t(v,x)}{\Delta t}. \tag{7.7.4}$$

We recall that the crucial advantage of this method is that it allows for the formulation of stability conditions. In fact, following Theorem 6.4.9, the stability follows provided that the discretization steps Δ_v and Δ_t in the v and t variables are chosen as follows

$$\Delta_v \leq \frac{2}{\bar{u}\,\|K\|_\infty}, \qquad \Delta_t \leq \frac{\Delta_v^2}{2 + \bar{u}\,\Delta_v^2\,\|\partial_v K\|_\infty}. \tag{7.7.5}$$

We recall that $\bar{u} = \max\Big(|u_{\min}|, |u_{\max}|\Big)$, and the control u_t satisfies $u_{\min} \leq u_t \leq u_{\max}$ for a.e. $t \in [0,T]$. In order to discretize the quotient (7.7.4), we set the discretization step in the x variable as $\Delta_x = \Delta_v \Delta_t$.

In the first simulation, we focus on showing the typical evolution of degenerate Kolmogorov equations. To this end, we use as initial density

$$f_0(v,x) = C \exp\Big(- 0.5\left(v^2 + x^2\right)\Big), \tag{7.7.6}$$

with $C > 0$ being a normalization constant, see Figure 7.1 below. We start by examining the behaviour of the system in the absence of interaction, *i.e.*, when there is no control ($u_t = 0$ for every $t \in [0,T]$), or when the interaction kernel K is the null function. As can be inferred from the corresponding stochastic process **(SDE)**, the density exhibits diffusion in the velocity variable v and transport in the position variable x. This behaviour is illustrated in Figure 7.2, which shows the evolution of the initial density up to time $T = 2$.

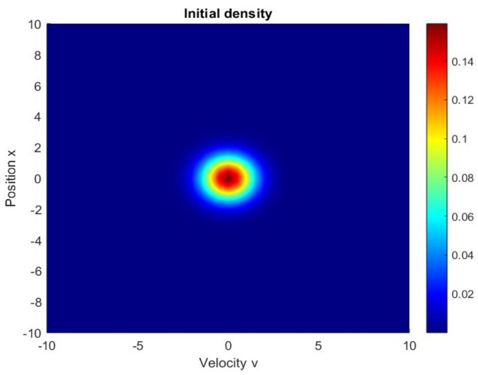

Figure 7.1
$\Delta_v = 0,5,\ \Delta_x = 0,0417,\ \Delta_t = 0,0833.$

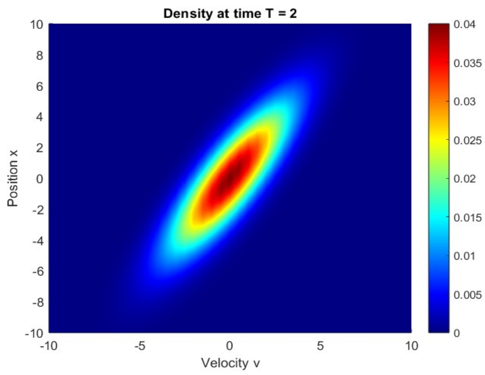

Figure 7.2
Evolution with no interaction.

We study the evolution of the initial density in a model where the interaction is based on the position x, as in Kuramoto models. We recall that in these models (see for instance (7.1.1)) the particles interact according to how they are positioned relative to one another: if the interaction is attractive, they move closer together; if repulsive, they tend to separate and disperse.

We consider the following interaction kernel

$$K(v,x) = x\,\exp\Big(- 0.05\left(v^2 + x^2\right)\Big).$$

We constrain the control u within the range $u_{\min} = -2$ and $u_{\max} = 2$, ensuring that the previously employed parameters ($\Delta_v = 0,5,\ \Delta_x = 0,0417,\ \Delta_t = 0,0833$) satisfy the stability condition (7.7.5). Firstly, we consider again the initial density given by (7.7.6). Our target is to steer the density as close as possible to $x = 0$ at the time $T = 2$. To this end, we consider the cost functional

$$J(f,u) = \int_{\mathbb{R}^2} 10\, x^2 \exp\Big(-0.05\big(v^2 + x^2\big)\Big) f_T(v,x)\, dv dx. \tag{7.7.7}$$

To measure the effectiveness of the control strategy, we introduce the *order parameter*

$$r_t = \int_{\mathbb{R}^2} \frac{1}{x^2+1} f_t(v,x)\, dv dx, \qquad t \in [0,T].$$

Clearly, $0 \le r_t \le 1$. In particular, if the density is concentrated around $x = 0$, then $r_t = 1$. On the other hand, as the support of the density moves away from $x = 0$, r_t tends to vanishes. The first case corresponds to a situation where the particles have the same position $x = 0$, whereas the second describes a scenario where the particles are widely separated.
In Figure 7.3 we illustrate the results of the controlled dynamic, which can be compared to the uncontrolled case shown in Figure 7.2. In Figure 7.4, we compare the evolution of the order parameter in the uncontrolled and the controlled dynamics, respectively. The simulation with initial density (7.7.6) shows that, even if the density is initially concentrated around $x = 0$, the dynamics tends to disperse the distribution over time, both in the spatial variable x and in the velocity variable v. However, the action of the control mitigates this spreading effect, helping to keep the density more concentrated in the position variable near $x = 0$, as can be seen by comparing Figure 7.2 with Figure 7.3, and by inspecting the comparison of the order parameters in Figure 7.4.

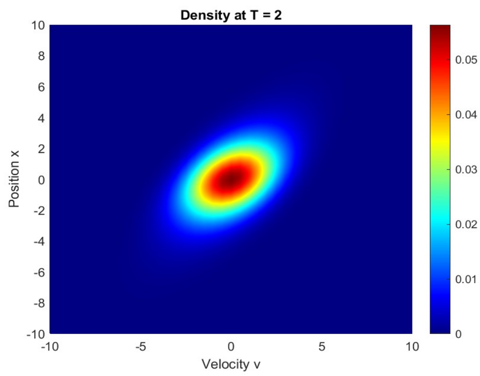

Figure 7.3
Evolution with clustering interaction kernel.

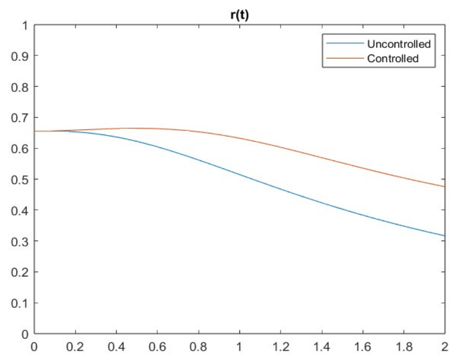

Figure 7.4
Order parameters comparison.

We conclude examining a more interesting case, where the initial density is given by

$$f_0(v,x) = C\exp\Big(-0.5\left(v^2+(x-2)^2\right)\Big) + \exp\Big(-0.5\left(v^2+(x+2)^2\right)\Big),$$

with $C > 0$ being a normalization constant. This scenario corresponds to the case where there are two clusters of particles around $x = \pm 2$, as shown in Figure 6.5. In the absence of interactions ($u_t = 0$ for every $t \in [0,T]$, or $K(v,x) = 0$ for all $(v,x) \in \mathbb{R}^2$), the evolution of the density is depicted in Figure 6.6. In particular, we observe that the density spreads along the position variable x, just as we have seen in the case with initial density (7.7.6).

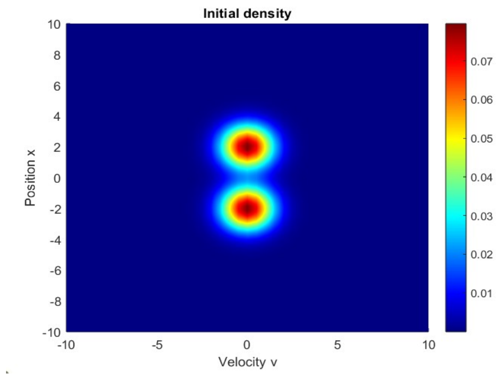

Figure 7.5
$\Delta_v = 0,5,\ \Delta_x = 0,0417,\ \Delta_t = 0,0833.$

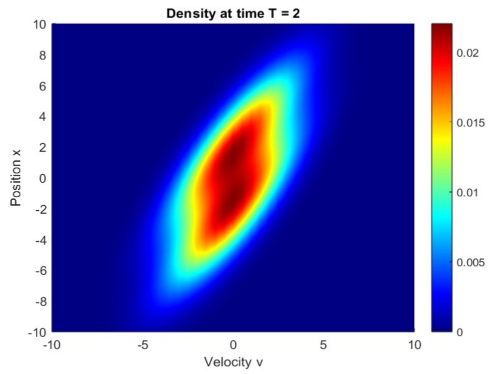

Figure 7.6
Evolution with no interaction.

Our goal, as in the previous examples, is to drive the density (namely the particles) toward $x = 0$. To achieve this, we use the same cost functional as in (7.7.7). In Figure 7.7 we illustrate the effect of the control on the density, while in Figure 7.8 we compare the order parameters of the uncontrolled and controlled cases. In this simulation, we observe a behaviour similar to that of the one with initial density (7.7.6). In this case, however, the system starts away from $x = 0$. In the absence of control, the dynamics tend to preserve this dispersion, while the action of the control aggregates the particles around $x = 0$, as shown in Figure 7.7 and Figure 7.8.

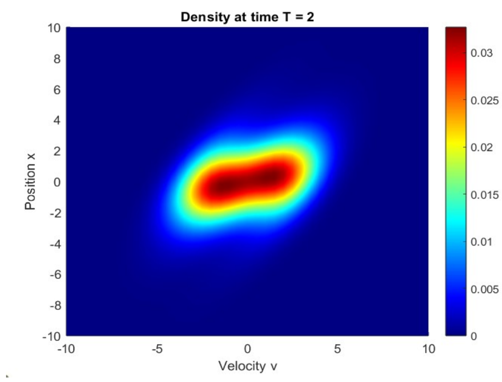

Figure 7.7
Evolution with clustering interaction kernel.

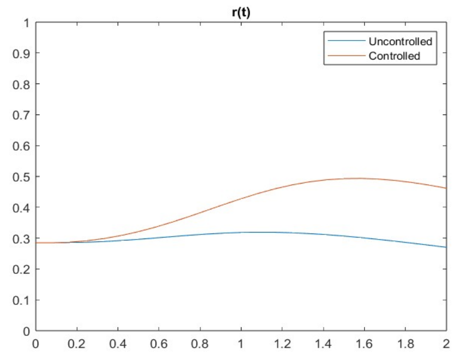

Figure 7.8
Order parameters comparison.

The numerical results obtained via the gradient descent method exhibit the expected behavior, in agreement with the formal derivation of **(SDE)-(P)** as the mean field limit of **(N-SDEs)-(P$_N$)**. Although a direct numerical verification of (7.1.5) lies beyond the scope of this work (mainly because of the curse of dimensionality), these simulations nonetheless support this theoretical framework.

Acknowledgements

First and foremost, I wish to express my sincere gratitude to Prof. Sergio Polidoro: I am deeply grateful for his constant guidance, encouragement, and for sharing with me his ideas, insights, and enthusiasm throughout these years.
Secondly, I would like to express my heartfelt thanks to Prof. Cecilia Vernia, whose kindness and support, as well as her guidance as one of my first mentors, have been invaluable in my mathematical research.
I owe my sincere gratitude to Dr. Annalaura Rebucci, who has been much more than a collaborator and has also become a truly precious friend.
I am also sincerely thankful to Prof. Enrique Zuazua for inviting me to Erlangen, and to Maureen, Ilias, Enrico, and all the wonderful people I met there, for the unforgettable moments we shared together.
In addition, I would like to express my deepest appreciation to all the people in Modena with whom I spent so much time, in particular Filippo, Gianluca, Elena, Danilo, Andrea, Carmelo, Davide, Giorgia and many others, whose company made my experience truly memorable.
Finally, I sincerely thank Prof. Matteo Colangeli and Prof. Stefano Pagliarani for agreeing to referee this thesis.
Outside of the academic context, I would like to thank my mother, Gigliola, and my sister, Chiara, for supporting me patiently since childhood. I would also like to thank my friends for their constant understanding and for filling my life with many joyful moments. A special thanks goes to Alessandro, Francesco, Antonella, and Alberto for their unwavering support and friendship.

Bibliography

[1] J. A. ACEBRÓN AND R. SPIGLER, *Adaptive frequency model for phase-frequency synchronization in large populations of globally coupled nonlinear oscillators*, Phys. Rev. Lett., 81 (1998), pp. 2229–2232.

[2] D. R. AKHMETOV, M. M. LAVRENTIEV, JR., AND R. SPIGLER, *Existence and uniqueness of classical solutions to certain nonlinear integro-differential Fokker-Planck type equations*, Electron. J. Differential Equations, (2002), pp. No. 24, 17.

[3] D. R. AKHMETOV AND R. SPIGLER, *Uniform and optimal estimates for solutions to singularly perturbed parabolic equations*, J. Evol. Equ., 7 (2007), pp. 347–372.

[4] G. ALBI, Y.-P. CHOI, M. FORNASIER, AND D. KALISE, *Mean field control hierarchy*, Appl. Math. Optim., 76 (2017), pp. 93–135.

[5] K. AMANO, *Maximum principles for degenerate elliptic-parabolic operators*, Indiana Univ. Math. J., 28 (1979), pp. 545–557.

[6] F. ANCESCHI, G. ASCIONE, D. CASTORINA, AND F. SOLOMBRINO, *Optimal control problems driven by nonlinear degenerate Fokker-Planck equations*, arXiv preprint arXiv:2410.24000, (2024).

[7] ——, *Well-posedness of Kolmogorov-Fokker-Planck equations with unbounded drift*, Journal of Mathematical Analysis and Applications, 543 (2025), p. 128909.

[8] F. ANCESCHI, M. PICCININI, AND A. REBUCCI, *New perspectives on recent trends for Kolmogorov operators*, in Kolmogorov operators and their applications, vol. 56 of Springer INdAM Ser., Springer, Singapore, [2024] ©2024, pp. 57–92.

[9] F. ANCESCHI AND S. POLIDORO, *A survey on the classical theory for Kolmogorov equation*, Matematiche (Catania), 75 (2020), pp. 221–258.

[10] M. ANNUNZIATO AND A. BORZÍ, *Optimal control of probability density functions of stochastic processes*, Math. Model. Anal., 15 (2010), pp. 393–407.

[11] ——, *Fokker-Planck-based control of a two-level open quantum system*, Math. Models Methods Appl. Sci., 23 (2013), pp. 2039–2064.

[12] M. Annunziato and A. Borzí, *A Fokker-Planck control framework for stochastic systems*, EMS Surv. Math. Sci., 5 (2018), pp. 65–98.

[13] M. S. Aronna and F. Tröltzsch, *First and second order optimality conditions for the control of Fokker-Planck equations*, ESAIM Control Optim. Calc. Var., 27 (2021), pp. Paper No. 15, 26.

[14] G. Ascione, D. Castorina, and F. Solombrino, *Mean-field sparse optimal control of systems with additive white noise*, SIAM J. Math. Anal., 55 (2023), pp. 6965–6990.

[15] ——, *Mean-field sparse optimal control of systems with additive white noise*, SIAM Journal on Mathematical Analysis, 55 (2023), pp. 6965–6990.

[16] P. Auscher, C. Imbert, and L. Niebel, *Weak solutions to Kolmogorov-Fokker-Planck equations: regularity, existence and uniqueness*, arXiv preprint arXiv:2403.17464, (2024).

[17] ——, *Fundamental solutions to Kolmogorov-Fokker-Planck equations with rough coefficients: existence, uniqueness, upper estimates*, SIAM J. Math. Anal., 57 (2025), pp. 2114–2137.

[18] D. Benedetto, E. Caglioti, and U. Montemagno, *On the complete phase synchronization for the Kuramoto model in the mean-field limit*, Commun. Math. Sci., 13 (2015), pp. 1775–1786.

[19] S. Biagi and A. Bonfiglioli, *An introduction to the geometrical analysis of vector fields—with applications to maximum principles and Lie groups*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2019.

[20] S. Biagi and M. Bramanti, *Schauder estimates for Kolmogorov-Fokker-Planck operators with coefficients measurable in time and Hölder continuous in space*, J. Math. Anal. Appl., 533 (2024), pp. Paper No. 127996, 65.

[21] U. Biccari and E. Zuazua, *A stochastic approach to the synchronization of coupled oscillators*, Frontiers in Energy Research, 8 (2020), p. 115.

[22] A. Bonfiglioli, *Hörmander vector fields equipped with dilations: lifting, Lie-group construction, applications*, Matematiche (Catania), 75 (2020), pp. 67–89.

[23] A. Bonfiglioli, G. Citti, G. Cupini, M. Manfredini, A. Montanari, D. Morbidelli, A. Pascucci, S. Polidoro, and F. Uguzzoni, *The role of fundamental solution in potential and regularity theory for subelliptic PDE*, in Geometric methods in PDE's, vol. 13 of Springer INdAM Ser., Springer, Cham, 2015, pp. 341–373.

[24] A. Bonfiglioli and E. Lanconelli, *Liouville-type theorems for real sub-Laplacians*, Manuscripta Math., 105 (2001), pp. 111–124.

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

[26] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Fundamental solutions for non-divergence form operators on stratified groups*, Trans. Amer. Math. Soc., 356 (2004), pp. 2709–2737.

[27] A. Bonfiglioli, E. Lanconelli, and F. Uguzzoni, *Stratified Lie groups and potential theory for their sub-Laplacians*, Springer Monographs in Mathematics, Springer, Berlin, 2007.

[28] A. Bonfiglioli and F. Uguzzoni, *Families of diffeomorphic sub-Laplacians and free Carnot groups*, Forum Math., 16 (2004), pp. 403–415.

[29] ——, *Maximum principle and propagation for intrinsicly regular solutions of differential inequalities structured on vector fields*, J. Math. Anal. Appl., 322 (2006), pp. 886–900.

[30] J.-M. Bony, *Principe du maximum, inégalite de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés*, Ann. Inst. Fourier (Grenoble), 19 (1969), pp. 277–304 xii.

[31] U. Boscain and S. Polidoro, *Gaussian estimates for hypoelliptic operators via optimal control*, Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl., 18 (2007), pp. 333–342.

[32] M. Bramanti and L. Brandolini, *Schauder estimates for parabolic nondivergence operators of Hörmander type*, J. Differential Equations, 234 (2007), pp. 177–245.

[33] ——, *Hörmander operators*, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, [2023] ©2023. With a foreword by Ermanno Lanconelli.

[34] M. Bramanti, L. Brandolini, E. Lanconelli, and F. Uguzzoni, *Non-divergence equations structured on Hörmander vector fields: heat kernels and Harnack inequalities*, Mem. Amer. Math. Soc., 204 (2010), pp. vi+123.

[35] M. Bramanti, L. Brandolini, M. Manfredini, and M. Pedroni, *Fundamental solutions and local solvability for nonsmooth Hörmander's operators*, Mem. Amer. Math. Soc., 249 (2017), pp. v+79.

[36] M. Bramanti and S. Polidoro, *Fundamental solutions for Kolmogorov-Fokker-Planck operators with time-depending measurable coefficients*, Math. Eng., 2 (2020), pp. 734–771.

[37] J. A. Carrillo, Y.-P. Choi, and S. P. Perez, *A review on attractive-repulsive hydrodynamics for consensus in collective behavior*, in Active particles. Vol. 1. Advances in theory, models, and applications, Model. Simul. Sci. Eng. Technol., Birkhäuser/Springer, Cham, 2017, pp. 259–298.

[38] M. Chipot, *Elements of nonlinear analysis*, Birkhäuser Advanced Texts: Basler Lehrbücher. [Birkhäuser Advanced Texts: Basel Textbooks], Birkhäuser Verlag, Basel, 2000.

[39] N. Chopra and M. W. Spong, *Passivity-Based Control of Multi-Agent Systems*, Springer Berlin Heidelberg, Berlin, Heidelberg, 2006, pp. 107–134.

[40] G. Citti, N. Garofalo, and E. Lanconelli, *Harnack's inequality for sum of squares of vector fields plus a potential*, Amer. J. Math., 115 (1993), pp. 699–734.

[41] G. Citti, E. Lanconelli, and A. Montanari, *Smoothness of Lipchitz-continuous graphs with nonvanishing Levi curvature*, Acta Math., 188 (2002), pp. 87–128.

[42] G. Citti and A. Sarti, *A cortical based model of perceptual completion in the roto-translation space*, J. Math. Imaging Vision, 24 (2006), pp. 307–326.

[43] G. Citti and B. Stroffolini, *Internal Schauder estimates for Hörmander type equations with Dini continuous source*, arXiv preprint arXiv:2306.02799, (2023).

[44] G. Cupini and E. Lanconelli, *On mean value formulas for solutions to second order linear PDEs*, Ann. Sc. Norm. Super. Pisa Cl. Sci., 22 (2021), pp. 1–33.

[45] G. Dal Maso, *An introduction to Γ-convergence*, vol. 8 of Progress in Nonlinear Differential Equations and their Applications, Birkhäuser Boston, Inc., Boston, MA, 1993.

[46] M. Di Francesco, P. Foschi, and A. Pascucci, *Analysis of an uncertain volatility model*, J. Appl. Math. Decis. Sci., (2006), pp. Art. ID 15609, 17.

[47] M. Di Francesco and A. Pascucci, *On a class of degenerate parabolic equations of Kolmogorov type*, AMRX Appl. Math. Res. Express, (2005), pp. 77–116.

[48] ——, *A continuous dependence result for ultraparabolic equations in option pricing*, J. Math. Anal. Appl., 336 (2007), pp. 1026–1041.

[49] M. F. Djete, *Extended mean field control problem: a propagation of chaos result*, Electron. J. Probab., 27 (2022), pp. Paper No. 20, 53.

[50] M. F. Djete, D. Possamaï, and X. Tan, *McKean-Vlasov optimal control: limit theory and equivalence between different formulations*, Math. Oper. Res., 47 (2022), pp. 2891–2930.

[51] R. L. Dobrushin, *Prescribing a system of random variables by conditional distributions*, Theory of Probability & Its Applications, 15 (1970), pp. 458–486.

[52] B. Ermentrout, *An adaptive model for synchrony in the firefly pteroptyx malaccae*, J. Math. Biol., 29 (1991), pp. 571–585.

[53] M. Fischer and G. Livieri, *Continuous time mean-variance portfolio optimization through the mean field approach*, ESAIM Probab. Stat., 20 (2016), pp. 30–44.

[54] G. B. Folland, *Subelliptic estimates and function spaces on nilpotent Lie groups*, Ark. Mat., 13 (1975), pp. 161–207.

[55] M. Fornasier, S. Lisini, C. Orrieri, and G. Savarè, *Mean-field optimal control as gamma-limit of finite agent controls*, European Journal of Applied Mathematics, 30 (2019), p. 1153–1186.

[56] M. Fornasier, B. Piccoli, and F. Rossi, *Mean-field sparse optimal control*, Philos. Trans. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci., 372 (2014), pp. 20130400, 21.

[57] M. Fornasier and F. Solombrino, *Mean-field optimal control*, ESAIM Control Optim. Calc. Var., 20 (2014), pp. 1123–1152.

[58] N. Fournier and A. Guillin, *On the rate of convergence in Wasserstein distance of the empirical measure*, Probab. Theory Related Fields, 162 (2015), pp. 707–738.

[59] N. Garofalo and E. Lanconelli, *Asymptotic behavior of fundamental solutions and potential theory of parabolic operators with variable coefficients*, Math. Ann., 283 (1989), pp. 211–239.

[60] ——, *Level sets of the fundamental solution and Harnack inequality for degenerate equations of Kolmogorov type*, Trans. Amer. Math. Soc., 321 (1990), pp. 775–792.

[61] D. GILBARG AND N. S. TRUDINGER, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001. Reprint of the 1998 edition.

[62] Z. HAO, M. RÖCKNER, AND X. ZHANG, *Second order fractional mean-field SDEs with singular kernels and measure initial data*, arXiv preprint arXiv:2302.04392, (2023).

[63] L. HÖRMANDER, *Hypoelliptic second order differential equations*, Acta Math., 119 (1967), pp. 147–171.

[64] E. ISSOGLIO, S. PAGLIARANI, F. RUSSO, AND D. TREVISANI, *Degenerate McKean-Vlasov equations with drift in anisotropic negative Besov spaces*, arXiv preprint arXiv:2401.09165, (2024).

[65] D. JERISON AND J. M. LEE, *The Yamabe problem on CR manifolds*, J. Differential Geom., 25 (1987), pp. 167–197.

[66] D. S. JERISON AND A. SÁNCHEZ-CALLE, *Estimates for the heat kernel for a sum of squares of vector fields*, Indiana Univ. Math. J., 35 (1986), pp. 835–854.

[67] M. KAC, *Foundations of kinetic theory*, in Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, 1954–1955, vol. III, Univ. California Press, Berkeley-Los Angeles, Calif., 1956, pp. 171–197.

[68] H. KALF, *On E. E. Levi's method of constructing a fundamental solution for second-order elliptic equations*, Rend. Circ. Mat. Palermo (2), 41 (1992), pp. 251–294.

[69] D. KO AND E. ZUAZUA, *Model predictive control with random batch methods for a guiding problem*, Math. Models Methods Appl. Sci., 31 (2021), pp. 1569–1592.

[70] A. KOLMOGOROFF, *Zufallige bewegungen (zur theorie der brownschen bewegung)*, The Annals of Mathematics, 35 (1934), p. 116.

[71] N. V. KRYLOV, *Hörmander's theorem for stochastic partial differential equations*, Algebra i Analiz, 27 (2015), pp. 157–182.

[72] N. V. KRYLOV AND M. V. SAFONOV, *A certain property of solutions of parabolic equations with measurable coefficients*, Mathematics of the USSR-Izvestiya, 16 (1981), p. 151.

[73] L. KUPTSOV, *On parabolic means*, Dokl. Akad. Nauk, 252 (1980), pp. 296–301.

[74] Y. KURAMOTO, *Chemical oscillations, waves, and turbulence*, vol. 19 of Springer Series in Synergetics, Springer-Verlag, Berlin, 1984.

[75] D. LACKER, *Limit theory for controlled McKean-Vlasov dynamics*, SIAM J. Control Optim., 55 (2017), pp. 1641–1672.

[76] E. LANCONELLI AND S. POLIDORO, *On a class of hypoelliptic evolution operators*, vol. 52, 1994, pp. 29–63. Partial differential equations, II (Turin, 1993).

[77] M. M. LAVRENT′ EV, JR., R. SPIGLER, AND D. R. AKHMETOV, *Nonlinear integro-parabolic equations in unbounded domains. The existence of classical solutions with special properties*, Sibirsk. Mat. Zh., 42 (2001), pp. 585–609, ii.

[78] ——, *Regularization of the nonlinear integroparabolic Fokker-Planck equation with spatially periodic solutions. The existence of strong solutions*, Sibirsk. Mat. Zh., 42 (2001), pp. 825–848, ii.

[79] M. M. LAVRENTIEV, JR. AND R. SPIGLER, *Existence and uniqueness of solutions to the Kuramoto-Sakaguchi nonlinear parabolic integrodifferential equation*, Differential Integral Equations, 13 (2000), pp. 649–667.

[80] ——, *Time-independent estimates and a comparison theorem for a nonlinear integroparabolic equation of the Fokker-Planck type*, Differential Integral Equations, 17 (2004), pp. 549–570.

[81] M. M. LAVRENTIEV, JR., R. SPIGLER, AND A. TANI, *Existence, uniqueness, and regularity for the Kuramoto-Sakaguchi equation with unboundedly supported frequency distribution*, Differential Integral Equations, 27 (2014), pp. 879–892.

[82] E. E. LEVI, *Sulle equazioni lineari totalmente ellittiche alle derivate parziali*, Rendiconti del Circolo Matematico di Palermo (1884-1940), 24 (1907), pp. 275–317.

[83] E. E. LEVI, *I problemi dei valori al contorno per le equazioni lineari totalmente ellittiche alle derivate parziali*, Memorie di matematica e di fisica della Società italiana delle scienze, 16 (1909), pp. 3–113.

[84] H.-Q. LI, *Estimations optimales du noyau de la chaleur sur les groupes de type Heisenberg*, J. Reine Angew. Math., 646 (2010), pp. 195–233.

[85] E. MALAGOLI, D. PALLARA, AND S. POLIDORO, *Mean value formulas for classical solutions to uniformly parabolic equations in the divergence form with non-smooth coefficients*, Math. Nachr., 296 (2023), pp. 4236–4263.

[86] H. P. McKean, Jr., *Propagation of chaos for a class of non-linear parabolic equations*, in Stochastic Differential Equations (Lecture Series in Differential Equations, Session 7, Catholic Univ., 1967), vol. Session 7 of Lecture Series in Differential Equations, Air Force Office of Scientific Research, Office of Aerospace Research, United States Air Force, Arlington, VA, 1967, pp. 41–57.

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

[88] A. Nabi and J. Moehlis, *Single input optimal control for globally coupled neuron networks*, Journal of neural engineering, 8 (2011), p. 065008.

[89] G. Orlando, *Mean-field optimal control in a multi-agent interaction model for prevention of maritime crime*, Adv. Contin. Discrete Models, (2023), pp. Paper No. 24, 68.

[90] S. Pagliarani, G. Lucertini, and A. Pascucci, *Optimal regularity for degenerate Kolmogorov equations in non-divergence form with rough-in-time coefficients*, J. Evol. Equ., 23 (2023), pp. Paper No. 69, 37.

[91] D. Pallara and S. Polidoro, *Mean value formulas for classical solutions to some degenerate elliptic equations in Carnot groups*, Discrete Contin. Dyn. Syst. Ser. S, 17 (2024), pp. 1733–1749.

[92] ——, *Mean value formulas for classical solutions to subelliptic evolution equations in stratified Lie groups*, Calc. Var. Partial Differential Equations, 64 (2025), p. Paper No. 208.

[93] A. Pascucci, *Probability theory. Vol. II. Stochastic calculus*, vol. 166 of Unitext, Springer, Cham, italian ed., [2024] ©2024. La Matematica per il 3+2.

[94] A. Pascucci and A. Pesce, *On stochastic Langevin and Fokker-Planck equations: the two-dimensional case*, J. Differential Equations, 310 (2022), pp. 443–483.

[95] A. Pascucci and A. Rondelli, *Well-posedness of kinetic McKean-Vlasov equations*, arXiv preprint arXiv:2501.10987, (2025).

[96] A. Pascucci, A. Rondelli, and A. Y. Veretennikov, *Existence and uniqueness results for strongly degenerate McKean-Vlasov equations with rough coefficients*, arXiv preprint arXiv:2409.14451, (2024).

[97] G. Pecorella, *Kinetic mean field optimal control*, (submitted).

[98] ——, *Fundamental solution, maximum principle and Harnack inequality for second order subelliptic operators*, J. Elliptic Parabol. Equ., 11 (2025), pp. 45–81.

[99] G. PECORELLA, S. POLIDORO, AND C. VERNIA, *A study of the Kuramoto model for synchronization phenomena based on degenerate Kolmogorov-Fokker-Planck equations*, Journal of Mathematical Analysis and Applications, 543 (2025), p. 128837.

[100] G. PECORELLA AND A. REBUCCI, *Fundamental solution and Harnack inequality for subelliptic evolution operators on Carnot groups*, arXiv preprint arXiv:2509.15982, (2025).

[101] C. S. PESKIN, *Mathematical aspects of heart physiology*, Courant Institute of Mathematical Sciences, New York University, New York, 1975. Notes based on a course given at New York University during the year 1973/74.

[102] S. POLIDORO, *On a class of ultraparabolic operators of Kolmogorov-Fokker-Planck type*, Matematiche (Catania), 49 (1994), pp. 53–105 (1995).

[103] S. POLIDORO AND C. MOGAVERO, *A finite difference method for a boundary value problem related to the Kolmogorov equation*, Calcolo, 32 (1995), pp. 193–205.

[104] B. QIAN, *Gradient estimates for the heat kernels in higher dimensional Heisenberg groups*, Chinese Ann. Math. Ser. B, 31 (2010), pp. 305–314.

[105] A. REBUCCI, *Harnack inequality and maximum principle for degenerate kolmogorov operators in divergence form*, Journal of Mathematical Analysis and Applications, 538 (2024), p. 128371.

[106] L. P. ROTHSCHILD AND E. M. STEIN, *Hypoelliptic differential operators and nilpotent groups*, Acta Math., 137 (1976), pp. 247–320.

[107] S. ROY, M. ANNUNZIATO, AND A. BORZÍ, *A Fokker-Planck feedback control-constrained approach for modelling crowd motion*, J. Comput. Theor. Transp., 45 (2016), pp. 442–458.

[108] H. SAKAGUCHI, *Cooperative phenomena in coupled oscillator systems under external fields*, Progr. Theoret. Phys., 79 (1988), pp. 39–46.

[109] A. SÁNCHEZ-CALLE, *Fundamental solutions and geometry of the sum of squares of vector fields*, Invent. Math., 78 (1984), pp. 143–160.

[110] F. SERRA CASSANO, *Some topics of geometric measure theory in Carnot groups*, in Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. 1, EMS Ser. Lect. Math., Eur. Math. Soc., Zürich, 2016, pp. 1–121.

[111] R. SPIGLER, *The mathematics of Kuramoto models which describe synchronization phenomena*, Mat. Cult. Soc. Riv. Unione Mat. Ital. (I), 1 (2016), pp. 123–132.

[112] S. H. STROGATZ, *From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators*, Phys. D, 143 (2000), pp. 1–20. Bifurcations, patterns and symmetry.

[113] S. H. STROGATZ AND R. E. MIROLLO, *Stability of incoherence in a population of coupled oscillators*, J. Statist. Phys., 63 (1991), pp. 613–635.

[114] A.-S. SZNITMAN, *Topics in propagation of chaos*, in École d'Été de Probabilités de Saint-Flour XIX—1989, vol. 1464 of Lecture Notes in Math., Springer, Berlin, 1991, pp. 165–251.

[115] H.-A. TANAKA, A. J. LICHTENBERG, AND S. OISHI, *Self-synchronization of coupled oscillators with hysteretic responses*, Physica D: Nonlinear Phenomena, 100 (1997), pp. 279–300.

[116] F. TRÖLTZSCH, *Optimal control of partial differential equations*, vol. 112 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2010. Theory, methods and applications, Translated from the 2005 German original by Jürgen Sprekels.

[117] T. VICSEK AND A. ZAFEIRIS, *Collective motion*, Physics Reports, 517 (2012), pp. 71–140. Collective motion.

[118] C. VILLANI, *Topics in optimal transportation*, vol. 58 of Graduate Studies in Mathematics, American Mathematical Society, Providence, RI, 2003.

[119] D. VITTONE, *Submanifolds in Carnot groups*, vol. 7 of Tesi. Scuola Normale Superiore di Pisa (Nuova Series) [Theses of Scuola Normale Superiore di Pisa (New Series)], Edizioni della Normale, Pisa, 2008. Thesis, Scuola Normale Superiore, Pisa, 2008.

[120] K. WIESENFELD, P. COLET, AND S. H. STROGATZ, *Synchronization transitions in a disordered Josephson series array*, Phys. Rev. Lett., 76 (1996), pp. 404–407.

[121] ——, *Frequency locking in Josephson arrays: connection with the Kuramoto model*, Phys. Rev. E, 57 (1998), pp. 1563–1569.

[122] X. ZHANG, *Second order McKean-Vlasov SDEs and kinetic Fokker-Planck-Kolmogorov equations*, Discrete and Continuous Dynamical Systems - S, 18 (2025), pp. 2973–3026.