UNIVERSITÀ DEGLI STUDI DI BERGAMO
DIPARTIMENTO DI INGEGNERIA DELL'INFORMAZIONE
E METODI MATEMATICI°

QUADERNI DEL DIPARTIMENTO

Department of Information Technology and Mathematical Methods

Working Paper

Series "*Mathematics and Statistics*"

n. 3/MS – 2007

Non-divergence equations structured on Hörmander vector fields: heat kernels and Harnack inequalities

by

M. Bramanti, L. Brandolini, E. Lanconelli, and F. Uguzzoni

° Viale Marconi. 5, I – 24044 Dalmine (BG), ITALY, Tel. +39-035-2052339; Fax. +39-035-562779

Non-divergence equations structured on Hörmander vector fields: heat kernels and Harnack inequalities*

Marco Bramanti
Dipartimento di Matematica, Politecnico di Milano
Via Bonardi 9, 20133 Milano, Italy
marbra@mate.polimi.it

Luca Brandolini
Dipartimento di Ingegneria Gestionale e dell'Informazione
Università di Bergamo
Viale Marconi 5, 24044 Dalmine, Italy
luca.brandolini@unibg.it

Ermanno Lanconelli
Dipartimento di Matematica, Università di Bologna
Piazza di Porta S. Donato 5, 40126 Bologna, Italy
lanconel@dm.unibo.it

Francesco Uguzzoni
Dipartimento di Matematica, Università di Bologna
Piazza di Porta S. Donato 5, 40126 Bologna, Italy
uguzzoni@dm.unibo.it

Abstract

In this work we deal with linear second order partial differential operators of the following type:

$$H = \partial_t - L = \partial_t - \sum_{i,j=1}^{q} a_{ij}(t,x) X_i X_j - \sum_{k=1}^{q} a_k(t,x) X_k - a_0(t,x)$$

where $X_1, X_2, \ldots, X_q$ is a system of real Hörmander's vector fields in some bounded domain $\Omega \subseteq \mathbb{R}^n$, $A = \{a_{ij}(t,x)\}_{i,j=1}^{q}$ is a real symmetric

*2000 AMS Classification. **Primary:** 35H20, 35A08, 35K65 **Secondary:** 35H10, 35A17 **Keywords**: Hörmander's vector fields, heat kernels, Gaussian bounds, Harnack inequalities

uniformly positive definite matrix such that:

$$\lambda^{-1}|\xi|^2 \leq \sum_{i,j=1}^{q} a_{ij}(t,x)\,\xi_i\xi_j \leq \lambda|\xi|^2 \;\; \forall \xi \in \mathbb{R}^q, x \in \Omega, t \in (T_1, T_2)$$

for a suitable constant $\lambda > 0$ a for some real numbers $T_1 < T_2$. The coefficients a_{ij}, a_k, a_0 are Hölder continuous on $(T_1, T_2) \times \Omega$ with respect to the parabolic CC-metric

$$d_P((t,x),(s,y)) = \sqrt{d(x,y)^2 + |t-s|}$$

(where d is the Carnot-Carathéodory distance induced by the vector fields X_i's). We prove the existence of a fundamental solution $h(t,x;s,y)$ for H, satisfying natural properties and sharp Gaussian bounds of the kind:

$$\frac{e^{-cd(x,y)^2/(t-s)}}{c\left|B\left(x,\sqrt{t-s}\right)\right|} \leq h(t,x;s,y) \leq c\frac{e^{-d(x,y)^2/c(t-s)}}{\left|B\left(x,\sqrt{t-s}\right)\right|}$$

$$|X_i h(t,x;s,y)| \leq \frac{c}{\sqrt{t-s}}\frac{e^{-d(x,y)^2/c(t-s)}}{\left|B\left(x,\sqrt{t-s}\right)\right|}$$

$$|X_i X_j h(t,x;s,y)| + |\partial_t h(t,x;s,y)| \leq \frac{c}{t-s}\frac{e^{-d(x,y)^2/c(t-s)}}{\left|B\left(x,\sqrt{t-s}\right)\right|}$$

where $|B(x,r)|$ denotes the Lebesgue measure of the d-ball $B(x,r)$. We then use these properties of h as a starting point to prove a *scaling invariant* Harnack inequality for positive solutions to $Hu = 0$, when $a_0 \equiv 0$. All the constants in our estimates and inequalities will depend on the coefficients a_{ij}, a_k, a_0 only through their Hölder norms and the number λ.

Contents

Introduction

Object and main results of the paper

Let us consider the heat-type operator in $\mathbb{R}^{n+1}$

$$H=\partial_t-L=\partial_t-\sum_{i,j=1}^{q}a_{ij}\left(t,x\right)X_iX_j-\sum_{k=1}^{q}a_k\left(t,x\right)X_k-a_0\left(t,x\right) \tag{0.1}$$

where:

(H1) $X_1,X_2,\ldots,X_q$ is a system of real smooth vector fields which are defined in some bounded domain $\Omega\subseteq\mathbb{R}^n$ and satisfy Hörmander's condition in Ω: rank Lie$\{X_i,i=1,2,...,q\}=n$ at any point of Ω (more precise definitions will be given later);

(H2) $A=\{a_{ij}\left(t,x\right)\}_{i,j=1}^{q}$ is a real symmetric uniformly positive definite matrix satisfying, for some positive constant λ,

$$\lambda^{-1}\left|\xi\right|^2\leq\sum_{i,j=1}^{q}a_{ij}\left(t,x\right)\xi_i\xi_j\leq\lambda\left|\xi\right|^2$$

for every $\xi\in\mathbb{R}^q$, $x\in\Omega,t\in(T_1,T_2)$ for some $T_1<T_2$.

If $d\left(x,y\right)$ denotes the Carnot-Carathéodory metric generated in Ω by the X_i's and

$$d_P\left(\left(t,x\right),\left(s,y\right)\right)=\sqrt{d\left(x,y\right)^2+\left|t-s\right|}$$

is its "parabolic" counterpart in $\mathbb{R}\times\Omega$, we will assume that:

(H3) a_{ij},a_k,a_0 are Hölder continuos on $\mathcal{C}=(T_1,T_2)\times\Omega$ with respect to the distance d_P.

Under assumptions (H1),(H2),(H3), we shall prove the existence and basic properties of a fundamental solution h for the operator H, including a representation formula for solutions to the Cauchy problem, a "reproduction property" for h, and regularity results: namely, we will show that h is locally Hölder continuous, far off the pole, together with its derivatives X_jh,X_iX_jh, ∂_th. To be more precise, an explanation is in order here. The operator H is defined only on the cylinder $\mathcal{C}$. On the other hand, dealing with fundamental solutions, it is convenient to work with an operator defined on the whole space. For this reason we will extend the operator H to the whole $\mathbb{R}^{n+1}$, in such a way that, outside a compact set in the space variables, it coincides with the classical heat operator, and henceforth we will study the fundamental solution for this extended operator.

Strictly related to the proof of the existence of h, and of independent interest, are several sharp Gaussian bounds for h that we will establish:

$$\begin{gathered}\frac{e^{-cd(x,y)^2/(t-s)}}{c\left|B\left(x,\sqrt{t-s}\right)\right|}\leq h\left(t,x;s,y\right)\leq c\frac{e^{-d(x,y)^2/c(t-s)}}{\left|B\left(x,\sqrt{t-s}\right)\right|}\\ \left|X_i h\left(t,x;s,y\right)\right|\leq\frac{c}{\sqrt{t-s}}\frac{e^{-d(x,y)^2/c(t-s)}}{\left|B\left(x,\sqrt{t-s}\right)\right|}\\ \left|X_iX_j h\left(t,x;s,y\right)\right|+\left|\partial_t h\left(t,x;s,y\right)\right|\leq\frac{c}{t-s}\frac{e^{-d(x,y)^2/c(t-s)}}{\left|B\left(x,\sqrt{t-s}\right)\right|}\end{gathered}\tag{0.2}$$

where $x,y\in\mathbb{R}^n, 0<t-s<T$ and $|B(x,r)|$ denotes Lebesgue measure of the d-ball $B(x,r)$. The constant c in these estimates depends on the coefficients a_{ij}, a_k, a_0 only through their Hölder moduli of continuity and the ellipticity constant λ.

A precise list of the results we prove about h is contained in Theorem 10.7, stated at the beginning of Part II (see also Remark 10.9).

A remarkable consequence of these bounds is a scaling invariant Harnack inequality for H, and for its stationary counterpart L in (0.1), which will be proved throughout Part III. In that part we will assume a_0, the zero order term of H, to be identically zero. Precise results are stated in Theorems 15.1 and 15.3 at the beginning of Part III.

As we mentioned before, all the results we have described so far are proved for an operator defined on the whole space, which extends H, initially defined only locally. At the end of this work (see Section 19) we will also show how to come back to the original operator, deducing local results from the above global theorems (see Theorems 19.1 and 19.2). We could also say that the final goal of all our theory is to prove local properties of our operators, so that the theory itself is local, in spirit, although it exploits, for technical convenience, objects that are defined globally.

An announcement of the results contained in this paper has appeared in [13].

Previous results and bibliographic remarks

Gaussian estimates for the fundamental solution of second order partial differential operators of parabolic type, or, somehow more generally, for the density function of heat diffusion semigroups, have a long history, starting with Aronson's work [1]. The relevance of two-sided Gaussian estimates to get scaling invariant Harnack inequalities for positive solutions was firstly pointed out by Nash in the Appendix of his celebrated paper [46]. However, a complete implementation of the method outlined by Nash was given much later by Fabes and Stroock in [22], also inspired by some ideas of Krylov and Safonov (see [32], [33], [50]). Since then, the full strength of Gaussian estimates has been enlightened by several authors, showing their deep relationship not only with the scaling invariant Harnack inequality, but also with the ultracontractivity property of heat diffusion semigroups, with inequalities of Nash, Sobolev or Poincaré type, and with the doubling property of the measure of "intrinsic" balls. We directly refer to the recent monograph by Saloff-Coste [51] for a beautiful exposition of this circle of ideas, and for an exhaustive list of references on these subjects. Here we explicitly recall just the results in literature strictly close to the core of our work.

For heat operators of the kind

$$H=\partial_t-\sum_{i=1}^{q}X_i^2 \tag{0.3}$$

with X_i left invariant homogeneous vector fields on a Carnot group in $\mathbb{R}^n$, Gaussian bounds have been proved by Varopoulos ([57], [58], see also [59]):

$$\frac{1}{ct^{Q/2}}e^{-c\left\|y^{-1}\circ x\right\|^2/t}\leq h\left(t,x,y\right)\leq\frac{c}{t^{Q/2}}e^{-\left\|y^{-1}\circ x\right\|^2/ct} \tag{0.4}$$

for any $x,y\in\mathbb{R}^n,t>0$, where Q is the homogeneous dimension of the group, and $\|\cdot\|$ any homogeneous norm of the group. Two-sided Gaussian estimates and a scaling invariant Harnack inequality for the operator

$$H=\partial_t-\sum_{i,j=1}^{q}X_i\left(a_{ij}X_j\right)$$

have been proved by Saloff-Coste and Stroock in [52], where $\{a_{ij}\}$ is a uniformly positive matrix with measurable entries, and the vector fields X_i are left invariant with respect to a connected unimodular Lie group with polynomial growth.

In absence of a group structure, Gaussian bounds for operators (0.3) have been proved, on a compact manifold and for finite time, by Jerison-Sanchez-Calle [30], with an analytic approach (see also the previous partial result in [53]), and, on the whole $\mathbb{R}^{n+1}$, by Kusuoka-Stroock, [35], [36], using the Malliavin stochastic calculus.

Unlike the study of "sum of squares" Hörmander's operators, the investigation of non-divergence operators of Hörmander type has a relatively recent history. Stationary operators of kind

$$L=\sum_{i,j=1}^{q}a_{ij}\left(x\right)X_iX_j \tag{0.5}$$

with $X_1,...,X_q$ system of Hörmander's vector fields have been studied by Xu [60], Bramanti, Brandolini [10], [11], Capogna, Han [15]. A first attempt to study Cordes and/or Alexandrov-Bakelman-Pucci estimates for operators (0.5) with measurable coefficients a_{ij} and particular classes of vector fields X_i are contained in [19], [20], [21].

Evolution operators of kind (0.1) have been considered by Bonfiglioli, Lanconelli, Uguzzoni [3], [4], [6], Bramanti, Brandolini [12]. In [9] also more general operators of kind

$$L=\sum_{i,j=1}^{q}a_{ij}\left(x\right)X_iX_j+a_0\left(x\right)X_0 \tag{0.6}$$

with $X_0, X_1, ..., X_q$ system of Hörmander's vector fields have been studied.

In these papers, the matrix $\{a_{ij}\}$ is assumed symmetric and uniformly elliptic, and the entries a_{ij} typically belong to some function space defined in terms of the vector fields X_i and the metric they induce. In particular, these operators do not have smooth coefficients, so they are no longer hypoelliptic. Therefore the mere existence of a fundamental solution is troublesome. For the operators (0.1) (without lower order terms) with X_i left invariant homogeneous Hörmander's vector fields on a stratified Lie group and under assumptions (H1), (H2), (H3), it has been proved by Bonfiglioli, Lanconelli, Uguzzoni in [3], [4], [5] that the fundamental solution h exists and satisfies Gaussian bounds of the kind (0.4). As a consequence of these estimates, in [6] it is proved a scaling invariant Harnack inequality for the operator H.

A particular class of operators (0.6), namely ultraparabolic operators of Kolmogorov-Fokker-Planck type, has been studied by Pascucci and Polidoro in relation both with Harnack inequality and Gaussian bounds for the fundamental solution; see [47].

Previous results about Harnack inequality for general Hörmander's operators date back to Bony's seminal paper [8], where a first qualitative version of this result is proved. A first scaling invariant Harnack inequality for heat-type Hörmander's operators was proved later by Kusuoka-Stroock [35].

Strategy and structure of the paper

Following the general strategy used in the case of homogeneous groups in [3], [4], [6], our study proceeds in three steps, corresponding to the Parts of this paper. In Part I we consider operators of kind (0.1) with *constant* coefficients a_{ij}, and no lower order terms. For these operators, existence and basic properties of the fundamental solution h_A are guaranteed by known results (see Section 3). Here the point is to prove sharp Gaussian bounds on h_A, which have to be *uniform* in the ellipticity class of the matrix $A = \{a_{ij}\}$.

In Part II we study operators with variable Hölder continuous coefficients a_{ij}, a_k, a_0, and apply the results of Part I to establish existence and Gaussian bounds for the fundamental solution of these operators. This is accomplished by a suitable adaptation to our subelliptic context of the classical Levi's parametrix method.

Finally, thanks to the results of Part II, the proof of a Harnack inequality for H can follow the lines drawn in [6], and inspired by Fabes-Stroock's paper [22]: this is accomplished in Part III.

For the reader's convenience, we have included at the beginning of each Part of the paper more details about the strategy, the techniques, and the main new difficulties we had to overcome to reach our results.

A motivation

Many problems in geometric theory of several complex variables lead to fully nonlinear second order equations, whose linearizations are *nonvariational* operators of Hörmander type (0.5). Here we would like to present one of these problems whose source goes back to some papers by Bedford, Gaveau, Slodkowsky and Tomassini, see [2], [56], [54].

Let M be a real hypersurface, embedded in the Euclidean complex space $\mathbb{C}^{n+1}$. The Levi form of M at a point $p \in M$ is a Hermitian form on the complex tangent space whose eigenvalues $\lambda_1(p), ..., \lambda_n(p)$ determine in the directions of each corresponding eigenvector a kind of "principal curvature". Then, given a generalized symmetric function s, in the sense of Caffarelli-Nirenberg-Spruck [14], one can define the s-Levi curvature of M at p, as follows:

$$S_p(M) = s(\lambda_1(p), ..., \lambda_n(p)).$$

When M is the graph of a function u and one imposes that its s-Levi curvature is equal to a given function, one obtains a second order fully nonlinear partial differential equation, which can be seen as the pseudoconvex counterpart of the usual fully nonlinear elliptic equations of Hessian type, as studied e.g. in [14]. In linearized form, the equations of this new class can be written as (see [43, equation (34) p.324])

$$\mathcal{L}u \equiv \sum_{i,j=1}^{2n} a_{ij}\left(Du, D^2u\right) X_i X_j u = K(x, u, Du) \text{ in } \mathbb{R}^{2n+1} \tag{0.7}$$

where:

the X_j's are first order differential operators, with coefficients depending on the gradient of u, which form a real basis for the complex tangent space to the graph of u;

the matrix $\{a_{ij}\}$ depends on the function s;

K is a prescribed function.

It has to be noticed that $\mathcal{L}$ only involves $2n$ derivatives, while it lives in a space of dimension $2n+1$. Then, $\mathcal{L}$ is never elliptic, on any reasonable class of functions. However, the operator $\mathcal{L}$, when restricted to the set of strictly s-pseudoconvex functions, becomes "elliptic" along the $2n$ linearly independent directions given by the X_i's, while the missing one can be recovered by a commutation. Precisely,

$$\dim\left(\text{span}\left\{X_j, [X_i, X_j], i, j = 1, ..., 2n\right\}\right) = 2n+1$$

at any point (see [43, equation (36) p. 324]). This is a Hörmander-type rank condition of step 2.

The parabolic counterpart of (0.7), i.e. equation

$$\partial_t u(t,x) = \mathcal{L}u(t,x) \text{ for } t \in \mathbb{R},\ x \in \mathbb{R}^{2n+1} \tag{0.8}$$

arises studying the evolution by s-Levi curvature of a real hypersurface of $\mathbb{C}^{n+1}$ (see [29], [42]).

Satisfactory existence results of viscosity solution for equation (0.7) are already known (see [18], and references therein). In [44] a sort of hypoellipticity theorem is proved for the s-Levi equation: every strictly s-pseudoconvex

$C^{2,\alpha}_{loc}$ solution to the equation (0.7) is of class C^∞ whenever K is of class C^∞. Existence of classical solutions and optimal regularity results for the viscosity solutions are still widely open problems. One of the motivations of the present work is to provide the linear framework for s-Levi equations (0.7) and (0.8), by performing a deep analysis of the general class of Hörmander heat-type operators (0.1). For instance, an application of our stationary Harnack inequality is the following: let u be a positive smooth strictly s-pseudoconvex solution to the s-Levi equation (0.7), with K of class C^∞. Then u satisfies a scaling invariant Harnack inequality of type:

$$\sup_{B_r} u \le C \inf_{B_r} u$$

where B_r is the Carnot-Carathéodory ball of radius r, related to the vector fields $X_1, X_2, ..., X_{2n}$ in (0.7). The unpleasant fact is that the constant C depends on the solution u in an unspecified way. Understanding how C depends on u is an interesting and seemingly difficult open problem.

Part I

Operators with constant coefficients

1 Overview of Part I

Let us consider the heat-type operator in $\mathbb{R}^{n+1}$

$$H_A = \partial_t - L_A = \partial_t - \sum_{i,j=1}^{q} a_{ij} X_i X_j \tag{1.1}$$

where:

(H1) $X_1, X_2, \ldots, X_q$ is a system of real smooth vector fields which are defined in some bounded domain $\Omega \subseteq \mathbb{R}^n$ and satisfy Hörmander's condition of some step s in Ω. Explicitly, this means that:

$$X_i = \sum_{k=1}^{n} b_{ik}(x)\, \partial_{x_k}$$

with $b_{ik} \in C^\infty(\Omega)$, and the vector space spanned at every point of Ω by: the fields X_i; their commutators $[X_i, X_j] = X_i X_j - X_j X_i$; the commutators of the X_k's with the commutators $[X_i, X_j]$;...and so on, up to some step s, is the whole $\mathbb{R}^n$.

(H2) $A = \{a_{ij}\}_{i,j=1}^{q}$ is a real symmetric positive definite matrix with constant entries, and $\lambda > 0$ a constant such that:

$$\lambda^{-1} |\xi|^2 \le \sum_{i,j=1}^{q} a_{ij} \xi_i \xi_j \le \lambda |\xi|^2$$

for every $\xi \in \mathbb{R}^q$. When condition (H2) is fulfilled, we will say briefly that

$$A \in \mathcal{E}_\lambda.$$

As we have already noted, for these operators (or, more precisely, for a suitable extension of these operators to the whole $\mathbb{R}^{n+1}$, see Section 3), existence and basic properties of the fundamental solution h_A are guaranteed by known results. The goal of Part I is to prove the following:

Theorem 1.1 (uniform Gaussian bounds on h_A) *For any $T > 0$ there exists $c > 0$ such that, for any $t \in (0, T)$, $x, y \in \mathbb{R}^n$ the following bounds hold:*

1. Upper and lower bounds on h_A:

$$\frac{1}{c\left|B\left(x, \sqrt{t}\right)\right|} e^{-cd(x,y)^2/t} \le h_A(t, x, y) \le \frac{c}{\left|B\left(x, \sqrt{t}\right)\right|} e^{-d(x,y)^2/ct} \tag{1.2}$$

2. Upper bounds on the derivatives of h_A of arbitrary order:

$$\left|X_x^I X_y^J \partial_t^i h_A(t,x,y)\right| \leq \frac{c}{t^{i+\frac{|I|+|J|}{2}}\left|B\left(x,\sqrt{t}\right)\right|} e^{-d(x,y)^2/ct} \tag{1.3}$$

3. Estimate on the difference of the fundamental solutions of two operators (and their derivatives):

$$\left|X_x^I X_y^J \partial_t^i h_A(t,x,y) - X_x^I X_y^J \partial_t^i h_B(t,x,y)\right| \leq \frac{c\,\|A-B\|}{t^{i+\frac{|I|+|J|}{2}}\left|B\left(x,\sqrt{t}\right)\right|} e^{-d(x,y)^2/ct} \tag{1.4}$$

(here I, J are arbitrary multiindices, $A, B \in \mathcal{E}_\lambda$). The constants depend on the matrix A only through the number λ; in (1.3), (1.4), the constant also depends on the multiindices. The same estimates hold for $h_A(t,y,x)$.

The above theorem will be proved in Corollary 6.18 and Theorem 7.1 (bounds (1.2)), Theorem 8.1 (bounds (1.3)), and Theorem 9.1 (bounds (1.4)). We also refer to these theorems for more precise statements.

To prove Theorem 1.1, the techniques used in [3] for homogeneous left invariant vector fields are not suitable. Instead, here we follow an approach that just uses the results in [3] and which is basically inspired to the work of Jerison and Sanchez-Calle [30], integrated with several other devices to overcome the new difficulties. The main of them are the following: first, we have to take into account the dependence on the matrix A, getting estimates depending on A only through the number λ; second, our estimates have to be global in space, while in [30] the Authors work on a compact manifold; third, our estimates on the difference of the fundamental solutions of two operators have no analog in [30].

The strategy and plan of Part I is as follows. In Section 2 we will show how to extend to the whole space the vector fields X_i, so that the distance induced in $\mathbb{R}^n$ by these extended vector fields could enjoy suitable global properties, which will be used throughout the paper; in particular, the Lebesgue measure will satisfy globally the doubling condition w.r.t. metric balls. Consequently, in Section 3 we will extend to the whole space $\mathbb{R}^{n+1}$ the operator H_A, in order to assure the existence of a global fundamental solution h_A, satisfying natural properties. The rest of Part I is devoted to the proof of the uniform Gaussian bounds (1.2), (1.3), (1.4) for this fundamental solution h_A. The hardest step is the proof of the upper bound in (1.2), which will go through Sections 4 to 6. The strategy is the following. First, one proves the upper bound for $t \in (0,1)$ and $\varepsilon < d(x,y) < R$. In this range, the bound is equivalent to:

$$h_A(t,x,y) \leq ce^{-1/ct} \tag{1.5}$$

and is proved by means of estimates of Gevray type. This means that the exponential decay of h_A for vanishing t is deduced by a control on the supremum of the time derivative of any order of a solution to $H_A u = 0$. Establishing these bounds is the object of Section 4. This technique makes the constant c in (1.5) depend on:

$$\sup_{y\in\mathbb{R}^n}\int_0^T d\tau\int_{\varepsilon<d(x,y)<R} h_A(\tau,x,y)\,dx.$$

So the next problem is to prove a uniform upper bound on this quantity (i.e., depending on A only through λ). This is accomplished in Section 5, exploiting suitable estimates on fractional and singular integrals on spaces of homogeneous type, and uniform subelliptic estimates. Next, one has to prove the upper bound in (1.2) for $t\in(0,1)$ and $d(x,y)<\varepsilon$. This is performed in Section 6, applying Rothschild-Stein's technique of "lifting and approximation". This allows, by a rather involved procedure, to deduce the desired uniform bound from the analogous result proved, in the context of homogeneous groups, by Bonfiglioli-Lanconelli-Uguzzoni [3], and therefore completes the proof of the upper bound in (1.2) for $t\in(0,1)$ and $d(x,y)<R$. To prove the same upper bound for any $x,y\in\mathbb{R}^n$ and $t\in(0,T)$, we will use a comparison argument, exploiting the *ad hoc* extension of the operator H_A performed in Section 3.

In Section 7 we prove the lower bound in (1.2), exploiting the same construction already used in Section 6. Again, uniformity of the lower bound relies on the analogous uniform lower bound which holds in the case of homogeneous groups.

In Section 8 we prove the Gaussian bound (1.3) on the derivatives of h_A. Like in [30], this bound is deduced by the upper bound on h_A (proved in Sections 4 to 6), applying a powerful result proved by Fefferman and Sanchez-Calle [24], which assures the existence of a local change of coordinates which is a good substitute of dilations (which in our context do not generally exist).

Finally, in Section 9 we prove our estimate (1.4). The basic estimate, on the difference of two fundamental solutions h_A-h_B, relies on a suitable use of basic properties of the fundamental solution and on the uniform bound (1.3) on the derivatives of h_A. The estimate on the difference of derivatives of two fundamental solution is then derived by the basic estimate, with the same techniques used in Section 8.

2 Global extension of Hörmander's vector fields and geometric properties of the CC-distance

The aim of this section is to show how a system of Hörmander's vector fields initially defined in a bounded domain of $\mathbb{R}^n$ can be extended to the whole space to a new system of Hörmander's vector fields enjoying some good properties, among which a global doubling condition for the induced CC-distance. These facts can be also of independent interest. Moreover, in the next section we will apply this procedure to extend the differential operator H_A to the whole space, and assure the existence of a fundamental solution defined in the whole $\mathbb{R}^{n+1}$. We will also prove several estimates related to the CC-distance, which will be used throughout the paper.

We start recalling the standard definition of Carnot-Carathéodory distance induced by a system of vector fields $X_1, ..., X_q$.

Definition 2.1 *Let Ω be a domain in $\mathbb{R}^n$. We say that an absolutely continuous curve $\gamma : [0,T] \to \Omega$ is a sub-unit curve with respect to the system of vector fields $X_1, X_2, \ldots, X_q$ if*

$$\gamma'(t) = \sum_{j=1}^{q} \lambda_j(t) X_j(\gamma(t))$$

for a.e. $t \in [0,T]$, with $\sum_{j=1}^{q} \lambda_j(t)^2 \leq 1$ a.e. In the following, this number T will be denoted by $l(\gamma)$.

For any $x, y \in \Omega$, we define

$$d_\Omega(x,y) = \inf\{l(\gamma) \,|\, \gamma \text{ is } X\text{-subunit},\ \gamma(0) = x,\ \gamma(l(\gamma)) = y\}.$$

It is well known (Chow's theorem) that, if the vector fields satisfy Hörmander's condition, the above set in nonempty, so that $d_\Omega(x,y)$ is finite for every pair of points. Moreover, d_Ω is a distance in Ω, called the Carnot-Carathéodory distance (CC-distance) induced by the vector fields X_i's.

A known result by Fefferman-Phong [23] states that

$$c^{-1} |x-y| \leq d_\Omega(x,y) \leq c\,|x-y|^{1/s} \tag{2.1}$$

for every $x, y \in K \Subset \Omega$, where s is the step of Hörmander condition. In particular, this means that d_Ω induces the usual topology of $\mathbb{R}^n$. Moreover, Sanchez-Calle [53] and Nagel-Stein-Weinger [45] prove that the CC-distance is *locally* doubling with respect to the Lebesgue measure, i.e., denoting by $B(x,r)$ the d_Ω-ball of center x and radius r:

$$|B(x,2r)| \leq c\,|B(x,r)| \tag{2.2}$$

at least for x ranging in a compact set and r bounded by some r_0.

We will now proceed as follows. First, we will assume to have a system of Hörmander's vector fields defined in the whole $\mathbb{R}^n$ and such that, outside a compact set, it coincides with $(0,0,...,0,\partial_{x_1},\partial_{x_2},...,\partial_{x_n})$. Under these assumptions, we will prove several properties of the induced CC-distance and metric balls. Next, we will show how any Hörmander's system in a bounded domain can be extended to $\mathbb{R}^n$ in order to satisfy these assumptions.

2.1 Some global geometric properties of CC-distances

Throughout this subsection, $X = (X_1, X_2, ..., X_m)$ $(m \geq n)$ will denote a fixed system of Hörmander's vector fields defined in the whole $\mathbb{R}^n$, and such that

$$X = (0,0,...,0,\partial_{x_1},\partial_{x_2},...,\partial_{x_n}) \text{ in } \mathbb{R}^n \setminus \Omega_0$$

where Ω_0 is a fixed bounded domain.

Notation 2.2 *We shall denote by d the Carnot-Carathéodory distance induced by X, and by $B(x,r)$ the balls in the metric d. Euclidean balls in $\mathbb{R}^n$ will be denoted by $B_E(x,r)$.*

Also, throughout Section 2 we will denote by $\mathbf{c}$ any positive constant only depending on $X_1,\ldots,X_m$; we will write $\mathbf{c}\,(f_1,f_2,...,f_k)$ for any positive constant also depending on the arguments $f_1,f_2,...,f_k$. Note that a different convention on the constants will be made in Section 3, for the remaining sections of Part I (see Notation 3.3).

We are interested in establishing some *global* properties of d. We first prove the following

Lemma 2.3 *If either $B_E(x,r)\subseteq\mathbb{R}^n\setminus\Omega_0$ or $B(x,r)\subseteq\mathbb{R}^n\setminus\Omega_0$, then we have*

$$d(x,y)=|x-y|\qquad\forall y\in\overline{B_E(x,r)}$$

and

$$B(x,r)=B_E(x,r).$$

Proof. Clearly, if γ is an absolutely continuous path contained in $\mathbb{R}^n\setminus\Omega_0$, then γ is X-subunit iff $|\gamma'|\leq 1$ a.e.
First suppose $B_E(x,r)\subseteq\mathbb{R}^n\setminus\Omega_0$ and let $y\in\overline{B_E(x,r)}$. Obviously $d(x,y)\leq|x-y|$, since the segment $\overrightarrow{xy}$ is X-subunit. Let now γ be a X-subunit path connecting x and y. If γ is contained in $B_E(x,r)$, then $|\gamma'|\leq 1$ a.e. and thus $l(\gamma)\geq|x-y|$. Otherwise, there exists $t_0\leq l(\gamma)$ such that $\gamma(t_0)\in\partial B_E(x,r)$ and $\gamma|_{[0,t_0]}$ is contained in $\mathbb{R}^n\setminus\Omega_0$. Hence $|\gamma'|\leq 1$ a.e. in $[0,t_0]$ and we get $|x-y|\leq r\leq t_0\leq l(\gamma)$. Therefore $d(x,y)=|x-y|$. This immediately yields $B(x,r)\supseteq B_E(x,r)$. On the other hand, if $\xi\in\mathbb{R}^n\setminus B_E(x,r)$, then for every X-subunit path γ connecting x and ξ we have (arguing as above) $l(\gamma)\geq r$ and then $d(x,\xi)\geq r$. This proves that also $B(x,r)\subseteq B_E(x,r)$.
Now suppose $B(x,r)\subseteq\mathbb{R}^n\setminus\Omega_0$. It is sufficient to see that $B(x,r)\supseteq B_E(x,r)$ and then use what we have already proved. Let us argue by contradiction and suppose that there exists $y\in B_E(x,r)\setminus B(x,r)$. Letting $\gamma=\overrightarrow{xy}$, there exists $t_0\leq l(\gamma)=|x-y|<r$ such that $y'=\gamma(t_0)\in\partial B(x,r)$ and $\sigma=\gamma|_{[0,t_0]}$ is contained in $\overline{B(x,r)}\subseteq\mathbb{R}^n\setminus\Omega_0$. Then σ is X-subunit and $d(x,y')\leq l(\sigma)=t_0<r$, which gives a contradiction. ■

Lemma 2.4 *We have*

$$|x-y|\leq\mathbf{c}\,d(x,y)\qquad\forall x,y\in\mathbb{R}^n,\tag{2.3}$$

$$d(x,y)\leq\mathbf{c}(\sigma)\,|x-y|\quad\forall x,y\in\mathbb{R}^n:\,\max\{|x-y|,d(x,y)\}\geq\sigma>0.\tag{2.4}$$

As a consequence, for every domain $\Omega\Supset\Omega_0$, d is equivalent to the Euclidean distance in $\mathbb{R}^n\setminus\Omega$. Moreover,

$$B(x,r)\subseteq B_E(x,\mathbf{c}\,r)\qquad\forall x\in\mathbb{R}^n,\,r>0,\tag{2.5}$$

$$B(x,r)\supseteq B_E(x,\mathbf{c}(\sigma)^{-1}\,r)\qquad\forall x\in\mathbb{R}^n,\,r\geq\sigma>0.\tag{2.6}$$

In particular

$$\mathbf{c}(\sigma)^{-1}r^n \leq |B(x,r)| \leq \mathbf{c}\, r^n \qquad \forall x \in \mathbb{R}^n,\ r \geq \sigma > 0. \tag{2.7}$$

Proof. The proof of (2.3) is a standard consequence of the boundedness of the vector fields $X = (X_1, \ldots, X_m)$. Indeed, for every X-subunit path γ connecting x and y, we have

$$\begin{aligned}|x-y| &= \left|\int_0^{l(\gamma)} \gamma'(t)\,dt\right| \leq \left|\int_0^{l(\gamma)} \sum_{i=1}^m \lambda_i(t)\, X_i(\gamma(t))\,dt\right| \leq \\ &\leqslant \int_0^{l(\gamma)} \sqrt{\sum_{i=1}^m \lambda_i(t)^2}\sqrt{\sum_{i=1}^m [X_i(\gamma(t))]^2}dt \leq l(\gamma)\max_{\mathbb{R}^n}|X| \leq \mathbf{c}d(x,y)\,.\end{aligned}$$

Let us now prove (2.4). Let Ω_0 be as above and denote

$$d(x,\overline{\Omega_0}) = \min_{\xi\in\overline{\Omega_0}} d(x,\xi), \qquad d_E(x,\overline{\Omega_0}) = \min_{\xi\in\overline{\Omega_0}} |x-\xi|$$

$$\operatorname{diam}_d \overline{\Omega_0} = \max_{\xi,\xi'\in\overline{\Omega_0}} d(\xi,\xi') \qquad \operatorname{diam}_E \overline{\Omega_0} = \max_{\xi,\xi'\in\overline{\Omega_0}} |\xi-\xi'|.$$

With no loss of generality we can assume $d_E(x,\overline{\Omega_0}) \geq d_E(y,\overline{\Omega_0})$. Fix $R > 2\operatorname{diam}_d \overline{\Omega_0}$ such that $\overline{\Omega_0} \subseteq B_E(0,R/2)$.

If $|x-y| < d_E(x,\overline{\Omega_0})$, we have $y \in B_E(x, d_E(x,\overline{\Omega_0})) \subseteq \mathbb{R}^n \setminus \Omega_0$ and then $d(x,y) = |x-y|$ by Lemma 2.3.

Assume now $|x-y| \geq d_E(x,\overline{\Omega_0})$ and $x \notin B_E(0,R)$. There exist $\overline{x},\overline{y} \in \overline{\Omega_0}$ such that $d_E(x,\overline{\Omega_0}) = |x-\overline{x}|$, $d_E(y,\overline{\Omega_0}) = |y-\overline{y}|$. From Lemma 2.3, it follows that $d(x,\overline{x}) = |x-\overline{x}|$, $d(y,\overline{y}) = |y-\overline{y}|$. Thus

$$\begin{aligned}d(x,y) &\leq d(x,\overline{x}) + \operatorname{diam}_d \overline{\Omega_0} + d(y,\overline{y}) \\ &\leq d_E(x,\overline{\Omega_0}) + R/2 + d_E(y,\overline{\Omega_0}) \leq 3\, d_E(x,\overline{\Omega_0}) \leq 3|x-y|.\end{aligned}$$

Finally, if $|x-y| \geq d_E(x,\overline{\Omega_0})$ and $x \in B_E(0,R)$, then we have $|y| \leq d_E(y,\overline{\Omega_0}) + R/2 \leq d_E(x,\overline{\Omega_0}) + R/2 \leq 2R$. Therefore, by the continuity of d,

$$\frac{d(x,y)}{|x-y|} \leq \max_{\xi,\vartheta\in\overline{B_E(0,2R)},\,|\xi-\vartheta|\geq\sigma} \frac{d(\xi,\vartheta)}{|\xi-\vartheta|} = \mathbf{c}(\sigma), \qquad \text{if } |x-y| \geq \sigma,$$

$$\frac{|x-y|}{d(x,y)} \geq \min_{\xi,\vartheta\in\overline{B_E(0,2R)},\,d(\xi,\vartheta)\geq\sigma} \frac{|\xi-\vartheta|}{d(\xi,\vartheta)} = \mathbf{c}(\sigma)^{-1}, \qquad \text{if } d(x,y) \geq \sigma.$$

This completes the proof of (2.4).

Now, if $\Omega \Supset \Omega_0$, to prove the equivalence of d with the Euclidean distance, let

$$\delta = d(\Omega_0, \Omega^c)$$

and let $x,y\in\Omega$. If $d(x,y)<\delta$, then $y\in B(x,\delta)\subset(\Omega_0)^c$ and by Lemma 2.3 $d(x,y)=|x-y|$. If $d(x,y)\geq\delta$, then by (2.3)-(2.4) we have

$$|x-y|\leq\mathbf{c}\,d(x,y)\leq\mathbf{c}(\delta)\,|x-y|.$$

Finally, (2.5) follows by (2.3), (2.6) follows by (2.4), and (2.7) follows by (2.5)-(2.6). ■

Next, we prove that for the CC-distance d, the doubling condition holds *globally*. This will have several useful consequences.

Proposition 2.5 *We have*

$$|B(x,2r)|\leq\mathbf{c}|B(x,r)|\qquad\forall\,x\in\mathbb{R}^n,\,r>0.\tag{2.8}$$

As a consequence, there exists a positive constant Q $(\geq n)$ *such that*

$$|B(x,Mr)|\leq\mathbf{c}M^Q|B(x,r)|\quad\forall\,M\geq1,\,x\in\mathbb{R}^n,\,r>0.\tag{2.9}$$

We also have

$$|B(x,r)|\leq\mathbf{c}(\sigma)r^Q\qquad\forall\,r\geq\sigma>0,\,x\in\mathbb{R}^n,\tag{2.10}$$
$$|B(x,r)|\geq\mathbf{c}(R)^{-1}r^Q\qquad\forall\,0<r\leq R,\,x\in\mathbb{R}^n.\tag{2.11}$$

Proof. As we have already noted (see (2.2)), we know that (2.8) holds locally, i.e.,

$$\begin{aligned}&\forall K\Subset\mathbb{R}^n\;\exists\,r_0(K)>0,A(K)\geq1:\\&|B(x,2r)|\leq A(K)\,|B(x,r)|\qquad\forall x\in K,\;0<r\leq r_0(K)\end{aligned}$$

Let $K_0=\{y\in\mathbb{R}^n|d(y,\overline{\Omega_0})\leq2\}$. This is a compact set, by (2.3). Set $\sigma=\min\{r_0(K_0),1\}$. If $x\in K_0$, $0<r\leq\sigma$, then we have

$$|B(x,2r)|\leq A(K_0)\,|B(x,r)|=\mathbf{c}\,|B(x,r)|\,.$$

If $x\notin K_0$, $0<r\leq\sigma$, then we have $B(x,2r)\subseteq B(x,2)\subseteq\mathbb{R}^n\setminus\Omega_0$ and, by Lemma 2.3, $B(x,r)=B_E(x,r)$, $B(x,2r)=B_E(x,2r)$, which immediately gives $|B(x,2r)|=2^n|B(x,r)|$.

Finally, if $r\geq\sigma$, then (2.7) yields $|B(x,2r)|\leq\mathbf{c}(\sigma)\,|B(x,r)|=\mathbf{c}|B(x,r)|$. This proves (2.8).

The proof of (2.9) is now standard: let $\mathbf{c}_0\geq2^n$ be as in (2.8) and set $Q=\log_2\mathbf{c}_0$. We have $2^p\leq M<2^{p+1}$ for some nonnegative integer p. Applying (2.8) $p+1$ times, we get

$$|B(x,Mr)|\leq\left|B(x,2^{p+1}r)\right|\leq\mathbf{c}_0^{p+1}\,|B(x,r)|=\mathbf{c}_0 2^{pQ}\,|B(x,r)|\leq\mathbf{c}_0M^Q|B(x,r)|.$$

Finally, (2.10) and (2.11) are easy consequences of (2.7) and (2.9). ■

A first easy consequence of the global doubling condition is the following (see also Proposition 5.11 in [36]):

Lemma 2.6 *For every $\beta' < \beta$ there exists $\mathbf{c}(\beta,\beta') > 0$, such that for every $x, y \in \mathbb{R}^n, t > 0$,*

$$\frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|} \leq \mathbf{c}(\beta,\beta') \frac{e^{-\beta' d(x,y)^2/t}}{\left|B\left(y,\sqrt{t}\right)\right|} \tag{2.12}$$

Proof. If $d(x,y) \leq \sqrt{t}$, then the doubling condition implies $\left|B\left(x,\sqrt{t}\right)\right| \simeq \left|B\left(y,\sqrt{t}\right)\right|$, and (2.12) holds for any $\beta' \leq \beta$ with $\mathbf{c}$ independent of β, β'. If $d(x,y) \geq \sqrt{t}$, by (2.9)

$$\begin{aligned}\left|B\left(y,\sqrt{t}\right)\right| &\leq \left|B\left(x, d(x,y)+\sqrt{t}\right)\right| \leq \mathbf{c}\left(\frac{d(x,y)+\sqrt{t}}{\sqrt{t}}\right)^Q \left|B\left(x,\sqrt{t}\right)\right| \\ &\leq \mathbf{c}\left(\frac{d(x,y)}{\sqrt{t}}\right)^Q \left|B\left(x,\sqrt{t}\right)\right|.\end{aligned}$$

Then

$$\frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|} \leq \mathbf{c}\left(\frac{d(x,y)}{\sqrt{t}}\right)^Q \frac{e^{-\beta \frac{d(x,y)^2}{t}}}{\left|B\left(y,\sqrt{t}\right)\right|} \leq \mathbf{c}(\beta,\beta') \frac{e^{-\beta' \frac{d(x,y)^2}{t}}}{\left|B\left(y,\sqrt{t}\right)\right|}$$

because for any $\beta > 0, \beta' < \beta$, there exists $c > 0$ such that

$$u^Q e^{-\beta u^2} \leq c e^{-\beta' u^2} \text{ for any } u \geq 0.$$

■

2.2 Global extension of Hörmander's vector fields

We start with the following general fact on CC-distances:

Lemma 2.7 *Let X^1, X^2 be two systems of Hörmander vector fields defined, respectively, in two domains A_1, A_2 of $\mathbb{R}^n$, and assume that in some domain $A' \Subset A_1 \cap A_2$ the two systems coincide. Then, for every subdomain $A'' \Subset A'$, the CC-distances d_1, d_2 induced by X^1, X^2 are equivalent in A''.*

Proof. Let $\delta_1 = d_1\left(A'', (A')^c\right)$, $M = \max_{\xi,\eta\in\overline{\Omega''}} d_2(\xi,\eta)$, and let $x, y \in A''$.

If $d_1(x,y) \geq \delta_1$, then $d_2(x,y) \leq M \leq \frac{M}{\delta_1} d_1(x,y)$.

On the other hand, if $d_1(x,y) < \delta_1$, then $y \in B_{d_1}(x,\delta_1) \subseteq A'$, and there exists a sequence γ_n of X^1-subunit curves joining x to y, contained in $B_{d_1}(x,\delta_1)$ and realizing $d_1(x,y)$. Since $B_{d_1}(x,\delta_1) \subseteq A'$ and $X^1 = X^2$ in A', γ_n will be also X^2-subunit curves. Hence $d_2(x,y) \leq d_1(x,y)$.

We have therefore proved that $d_2(x,y) \leq c d_1(x,y)$ for every $x, y \in A'$. Exchanging the roles of d_1, d_2 we get the assertion. ■

Next, we need a technical lemma:

Lemma 2.8 *For any couple of bounded open subsets of $\mathbb{R}^n, A_1 \Subset A_2$, there exists another open set A, with $A_1 \Subset A \Subset A_2$ such that:*

1. *A has smooth boundary;*

2. *there exists* $\phi \in C_0^\infty(A_2), 0 \le \phi \le 1$, *such that* $\phi(x) = 1$ *if and only if* $x \in \overline{A}$.

Proof. Let ψ_ε be a standard family of mollifiers, that is $\psi_\varepsilon(x) = \varepsilon^{-n}\psi(x/\varepsilon)$ for any $\varepsilon > 0$, where

$$\psi \in C_0^\infty(\mathbb{R}^n) \text{ with } \operatorname{supp}\psi = B_E(0,1); \psi \ge 0; \int_{\mathbb{R}^n} \psi(x)\,dx = 1. \tag{2.13}$$

Let us prove 1. It is known that there exists a function $f \in C_0(A_2), 0 \le f \le 1$, such that $f(x) = 1$ if $x \in A_1$. Let $f_\varepsilon = f * \psi_\varepsilon$. Since $f \in C_0(A_2)$, we know that $f_\varepsilon \to f$ uniformly. Then, pick ε_0 such that $\sup|f - f_{\varepsilon_0}| < 1/8$. Hence

$$f_{\varepsilon_0}(x) \begin{cases} \ge 1 - \frac{1}{8} \text{ in } \overline{A_1} \\ \le \frac{1}{8} \text{ in } (A_2)^c \end{cases} \tag{2.14}$$

Let

$$\Gamma_t = \{x : f_{\varepsilon_0}(x) = t\}.$$

Since f_{ε_0} is smooth, by Sard's theorem, for a.e. $t \in \left[\frac{1}{4}, \frac{3}{4}\right]$ the level set Γ_t does not contain critical values of f_{ε_0}, and therefore is the smooth boundary of the open set $\{x : f_{\varepsilon_0}(x) > t\}$. For a fixed t_0 for which this happens, define $A = \{x : f_{\varepsilon_0}(x) > t_0\}$. Then by (2.14), we have $A_1 \Subset A \Subset A_2$, and point 1 is proved.

To prove 2, we will prove the following more general fact: if $A_1 \Subset A \Subset A_2$ where A is a bounded open subset of $\mathbb{R}^n$ satisfying a uniform exterior ball condition, then property 2 holds.

Let $\varepsilon_0 > 0$ be such that:

$$\forall z \in \partial A, \exists B_E(x,\varepsilon_0) \text{ s.t. } B_E(x,\varepsilon_0) \cap A = \varnothing \text{ and } \overline{B_E(x,\varepsilon_0)} \cap \overline{A} = \{z\}.$$

(This is exactly the uniform exterior ball condition). Now, for some $\varepsilon < \varepsilon_0/2$, let

$$A_\varepsilon = \{x : d_E(x,A) < \varepsilon\},$$

χ_{A_ε} be the characteristic function of A_ε, and

$$\phi_\varepsilon(x) = (\chi_{A_\varepsilon} * \psi_\varepsilon) = \int_{A_\varepsilon} \psi_\varepsilon(x-y)\,dy.$$

Then $0 \le \phi_\varepsilon \le 1, \operatorname{supp}\phi_\varepsilon \subset \overline{A_{2\varepsilon}}$, and by (2.13), $\phi_\varepsilon(x) = 1$ if and only if $B_E(x,\varepsilon) \subseteq A_\varepsilon$. Clearly, if $x \in \overline{A}$, then $B_E(x,\varepsilon) \subseteq A_\varepsilon$, and so $\phi_\varepsilon(x) = 1$. The point is to prove that, conversely,

$$B_E(x,\varepsilon) \subseteq A_\varepsilon \Rightarrow x \in \overline{A}. \tag{2.15}$$

To do this, pick an $x \notin \overline{A}$, and let us show that $B_E(x,\varepsilon) \nsubseteq A_\varepsilon$. If $d_E(x,A) \ge \varepsilon$, this is obvious because $x \notin A_\varepsilon$. If $x \notin \overline{A}$ and $d_E(x,A) < \varepsilon$, let $z \in \partial A$ such that $d_E(x,z) = d_E(x,A) = \delta < \varepsilon$. Let x_0 be such that $\overline{B_E(x_0,\varepsilon_0/2)} \cap \overline{A} = \{z\}$, and let r be the diameter of $B_E(x_0,\varepsilon_0/2)$ passing through z. Then $x \in r$, and for any point $x' \in r$, one can say that $d_E(x',z) = d_E(x',A)$, because the segment r is also the radius of the ball $B_E(x_0,\varepsilon_0) \subset A^c$. Therefore we can find a point $x' \in r$, such that $d_E(x',x) = \varepsilon - \frac{\delta}{2}$ and $d_E(x',z) = d_E(x,z) + d_E(x',x) = \varepsilon + \frac{\delta}{2}$. Hence $x' \in B_E(x,\varepsilon) \setminus A_\varepsilon$ and (2.15) is proved. We have therefore constructed, for any $\varepsilon < \varepsilon_0/2$, a function $\phi_\varepsilon \in C_0^\infty(\mathbb{R}^n)$ such that

$$0 \leq \phi_\varepsilon \leq 1, \operatorname{supp} \phi_\varepsilon \subset \overline{A_{2\varepsilon}}, \phi_\varepsilon(x) = 1 \text{ if and only if } x \in \overline{A}.$$

Choosing $2\varepsilon < d_E(A,A_2)$,we have $\operatorname{supp}\phi_\varepsilon \subset A_2$ and the assertion is proved. ■

Theorem 2.9 *Let $Z = (Z_1,\ldots,Z_q)$ be a system of vector fields defined in a bounded domain $\Omega \subset \mathbb{R}^n$ and satisfying Hörmander's condition of step s in Ω. Then, for any domains $\Omega_1 \Subset \Omega_0 \Subset \Omega$, there exists a new system $X = (X_1, X_2, ..., X_m)$ $(m = q+n)$ of vector fields, such that the vector fields X_i's are defined on the whole space $\mathbb{R}^n$ and satisfy Hörmander's condition of step s in $\mathbb{R}^n$; moreover:*

$$X = (Z_1, Z_2, ..., Z_q, 0, 0, ..., 0) \text{ in } \Omega_1; \tag{2.16}$$

$$X = (0, 0, ..., 0, \partial_{x_1}, \partial_{x_2}, ..., \partial_{x_n}) \text{ in } \mathbb{R}^n \setminus \Omega_0. \tag{2.17}$$

Furthermore, denoting by d_Z,d_X, respectively, the CC-distances induced by Z in Ω and X in $\mathbb{R}^n$, we have:

1. *for any domain $\Omega_2 \Subset \Omega_1$, d_X is equivalent to d_Z in Ω_2;*
2. *d_X is equivalent to the Euclidean distance in $\mathbb{R}^n \setminus \Omega$;*
3. *d_X satisfies the global doubling condition:*

$$|B(x,2r)| \leq c\,|B(x,r)| \;\; \forall x \in \mathbb{R}^n, r > 0.$$

Proof. Applying Lemma 2.8 to $\Omega_1 \Subset \Omega_0$, let A be a smooth open set and $\varphi \in C_0^\infty(\Omega_0)$ a cut-off function such $\Omega_1 \Subset A \Subset \Omega_0$ and $\varphi(x) = 1$ if and only if $x \in \overline{A}$. Let us define:

$$X_i = \varphi Z_i, \quad i = 1,\ldots,q, \qquad X_{q+k} = (1-\varphi)\partial_{x_k}, \quad k = 1,\ldots,n.$$

Relations (2.16), (2.17) are obvious, so we only need to check Hörmander's condition. Fix a point $x \in \mathbb{R}^n$; if $\varphi(x) \neq 1$, then in a neighborhood of x the system $X_1,\ldots,X_m$ contains nonvanishing multiples of the n fields ∂_{x_k}, which span; if $\varphi(x) = 1$, then $x \in \overline{A}$ and the fields $X_i = \varphi Z_i$, $i = 1,...,q$, satisfy Hörmander's condition at x because at that point

$$[X_i, X_j] = [\varphi Z_i, \varphi Z_j] = \varphi^2 [Z_i, Z_j] + \varphi(Z_i\varphi) Z_j - \varphi(Z_j\varphi) Z_i = [Z_i, Z_j]$$

since, in $\overline{A}$, $\varphi = 1$ and $\varphi_{x_k} = 0$ for every k. Iterating the above relation, we see that at the point x the system X_i $(i = 1,...,q)$ and the system Z_i $(i = 1,...,q)$ generate the same Lie algebra, that is the whole $\mathbb{R}^n$.

Property 1 follows from Lemma 2.7; property 2 follows from Lemma 2.4; property 3 follows from Proposition 2.5 ■

3 Global extension of the operator H_A and existence of a fundamental solution

As we have already mentioned in the Introduction, it is convenient to deal with an operator H_A satisfying (H1)-(H2) on the whole space $\mathbb{R}^{n+1}$. By Theorem 2.9, we immediately have the following:

Theorem 3.1 *Given an operator of type (1.1),*

$$H_A = \partial_t - L_A = \partial_t - \sum_{i,j=1}^{q} a_{ij} Z_i Z_j,$$

where $Z_1, \dots, Z_q$ satisfy the assumptions (H1) in some bounded domain $\Omega \subseteq \mathbb{R}^n$, and the matrix $A = \{a_{ij}\}$ satisfies assumptions (H2), and given $\Omega_1 \Subset \Omega$, there exists a new operator of type (1.1)

$$H'_A = \partial_t - L'_A = \partial_t - \sum_{i,j=1}^{m} a'_{ij} X_i X_j, \qquad \{a'_{ij}\} = \begin{bmatrix} \{a_{ij}\}_{i,j=1}^{q} & 0 \\ 0 & I_n \end{bmatrix}, \tag{3.1}$$

such that:
i) the vector fields X_i's are defined on the whole space $\mathbb{R}^{n+1}$;
ii) H'_A coincides with H_A for $x \in \Omega_1$;
iii) H'_A coincides with the classical heat operator for x outside Ω;
iv) H'_A satisfies (H1) and (H2), with the same constant λ.

Remark 3.2 *Conditions (H1)-(H2) imply that H'_A can be rewritten as a standard "sum of squares" Hörmander's operator, and therefore it is hypoelliptic in $\mathbb{R}^{n+1}$. Let us recall that a linear differential operator P with C^∞ coefficients is said to be hypoelliptic if, whenever the equation $Pu = f$ is satisfied, in distributional sense, on some open set Ω, then $f \in C^\infty(\Omega)$ implies $u \in C^\infty(\Omega)$.*

>From now on we shall always work with the new vector fields $X_1, \dots, X_m$ defined in Theorem 2.9 and we shall denote H'_A simply by H_A and a'_{ij} by a_{ij}. More generally, throughout the rest of Part I it will be enough to work in the following setting:

Hypotheses. *Let $X = (X_1, X_2, ..., X_m)$ $(m = n + q)$ be a fixed system of Hörmander's vector fields defined in the whole $\mathbb{R}^n$, and such that*

$$X = (0, 0, ..., 0, \partial_{x_1}, \partial_{x_2}, ..., \partial_{x_n}) \text{ in } \mathbb{R}^n \setminus \Omega_0$$

where Ω_0 is a fixed bounded domain. For a fixed constant $\lambda \geq 1$, let $\mathcal{B}_\lambda$ be the set of matrices of the kind

$$A = \{a_{ij}\}_{i,j=1}^{m} = \begin{bmatrix} \{a_{ij}\}_{i,j=1}^{q} & 0 \\ 0 & I_n \end{bmatrix}$$

with $\{a_{ij}\}_{i,j=1}^q \in \mathcal{E}_\lambda$. *For every* $A \in \mathcal{B}_\lambda$, *we set*

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

Notation 3.3 *From now on, in the rest of Part I, we will denote by* $\mathbf{c}$ *any positive constant only depending on* $X_1, \ldots, X_m$ *and the number* λ. *We will write* $\mathbf{c}\,(f_1, f_2, ..., f_k)$ *for any positive constant also depending on the arguments* $f_1, f_2, ..., f_k$. *Note that a different convention on constants will be made in Section 10 (see Notation 10.6) for Parts II and III.*

The remarkable fact is that the operator H_A possesses a global fundamental solution, as stated in the following result, which is essentially contained in the paper [37] by Lanconelli, Pascucci.

Theorem 3.4 *There exists a global fundamental solution* $h_A(t, x; s, y)$ *for* H_A *in* $\mathbb{R}^{n+1}$, *with the properties listed below.*

(i) h_A *is smooth away from the diagonal of* $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$.

(ii) $h_A \geq 0$ *and* h_A *vanishes for* $t \leq s$.

(iii) For every $(s, y) \in \mathbb{R}^{n+1}$, $h_A(\cdot; s, y)$ *is locally integrable and*

$$H_A h_A(\cdot; s, y) = \delta_{(s,y)}$$

(the Dirac measure supported at $\{(s, y)\}$*).*

(iv) For every test function $\varphi \in C_0^\infty(\mathbb{R}^{n+1})$, *we have*

$$H_A \left(\int_{\mathbb{R}^{n+1}} h_A(\cdot; s, y)\, \varphi(s, y)\, ds dy \right) = \int_{\mathbb{R}^{n+1}} h_A(\cdot; s, y)\, H_A \varphi(s, y)\, ds dy = \varphi.$$

(v) $h_A^*(t, x; s, y) = h_A(s, y; t, x)$ *is a fundamental solution for the formal adjoint operator* $H_A^* = -\partial_t - \sum_{i,j=1}^m a_{i,j} X_j^* X_i^*$ *and it satisfies the dual statements of (iii) and (iv).*

(vi) For every $t > s$, *we have*

$$\int_{\mathbb{R}^n} h_A(t, x; s, y)\, dy = 1. \tag{3.2}$$

(vii) For every compact set $K \subset \mathbb{R}^n$ *and for every* $T > 0$ *there exist positive constants* M, R, δ *such that, for* $|x| > R$

$$\sup_{y \in K,\, s<t<T} h_A(t, x; s, y) + \sup_{y \in K,\, s<t<T} h_A(t, y; s, x) \;\leq M \exp(-\delta |x|^2). \tag{3.3}$$

(viii) $h_A(t,x;s,y)$ *depends on* t,s *only through* $t-s$. *Hence, from now on we will always write*

$$h_A(t,x;s,y)=h_A(t-s,x,y)$$

Corollary 3.5 *Let* $f:\mathbb{R}^n\to\mathbb{R}$ *be a bounded continuous function. Then, the function* $u(t,x)=\int_{\mathbb{R}^n}h_A(t,x,y)\,f(y)\,dy$ *is a classical solution to the Cauchy problem*

$$H_Au=0 \ \text{ in } \ (0,\infty)\times\mathbb{R}^n, \quad u(0,\cdot)=f.$$

Proof. Let $\phi_k\in C_0^\infty(\mathbb{R}^n)$, $k\in\mathbb{N}$, be cutoff functions such that $0\le\phi_k\le 1$, $\phi_k(x)=1$ for $|x|\le k$. Let us set $u_k(t,x)=\int_{\mathbb{R}^n}h_A(t,x,y)\,f(y)\,\phi_k(y)\,dy$. From Theorem 3.4-(i),(iii), it follows that we can differentiate under the integral sign and obtain $H_Au_k(t,x)=0$ for $t>0$. Moreover $u_k\to u$ pointwise. Indeed

$$|u_k(t,x)-u(t,x)|\le\sup|f|\int_{|y|>k}h_A(t,x,y)\,dy\to 0 \qquad \text{as } k\to\infty.$$

Thus, for every test function $\psi\in C_0^\infty((0,\infty)\times\mathbb{R}^n)$, we have

$$0=\int_{\mathbb{R}^{n+1}}\psi H_Au_k=\int_{\mathbb{R}^{n+1}}u_kH_A^*\psi\to\int_{\mathbb{R}^{n+1}}uH_A^*\psi \qquad \text{as } k\to\infty,$$

by dominated convergence, observing that

$$|u_k(t,x)|\le\sup|f|\int_{\mathbb{R}^n}h_A(t,x,y)\,dy=\sup|f|$$

by (3.2). Therefore $H_Au=0$ in $(0,\infty)\times\mathbb{R}^n$ (recall H_A is hypoelliptic). Let now $x_0\in\mathbb{R}^n$ be fixed and let us prove that $u(t,x)\to f(x_0)$, as $(t,x)\to(0,x_0)$. Recalling again (3.2), we get

$$\begin{aligned}|u(t,x)-f(x_0)|\le&\int_{|y-x_0|\le\delta}h_A(t,x,y)\,|f(y)-f(x_0)|\,dy\\&+2\sup|f|\int_{|y-x_0|>\delta}h_A(t,x,y)\,dy.\end{aligned}$$

The first integral can be made small choosing δ small enough, since f is continuous and (3.2) holds. Once δ is fixed, the second integral goes to zero as $(t,x)\to(0,x_0)$ by dominated convergence, using Theorem 3.4-(i),(ii),(vii). ■

We will use several times the following weak maximum principle:

Proposition 3.6 *Let* Ω *be an open subset of* $\mathbb{R}^n$ *and let* $T_0<T_1$. *For any* $u\in C^2((T_0,T_1)\times\Omega)$, *if*

$$\begin{aligned}Hu\le 0 \ &\text{ in } \ (T_0,T_1)\times\Omega\\ \limsup u\le 0 \ &\text{ in } \ ([T_0,T_1)\times\partial\Omega)\cup(\{T_0\}\times\Omega) \ \text{ and at infinity,}\end{aligned}$$

then $u\le 0$ *in* $(T_0,T_1)\times\Omega$.

The above result is essentially due to Picone; a proof can be found in [37, Proposition 2.2].

Corollary 3.7 *For every $x, y \in \mathbb{R}^n$, $t > 0$ and $s > 0$, the following* reproduction property *holds*

$$h_A(t+s, x, y) = \int_{\mathbb{R}^n} h_A(t, x, z)\, h_A(s, z, y)\, dz. \tag{3.4}$$

Proof. Let us denote by $v\,(t, x)$ and $u\,(t, x)$ respectively the functions in the left and in the right hand side of (3.4). From Theorem 3.4 and Corollary 3.5 it follows that $H_A v = 0 = H_A u$ in $(0, \infty) \times \mathbb{R}^n$, $u(0, \cdot) = v(0, \cdot)$. Moreover v vanishes as $|x| \to \infty$, uniformly in any strip $0 < t < T$, by means of (3.3). Thus we only need to see that the same holds for u, in order to get (3.4) from the weak maximum principle for H_A (Proposition 3.6). Let us write

$$\begin{aligned} u\,(t, x) &= \int_{|z-y|>R} h_A(t, x, z)\, h_A(s, z, y)\, dz + \int_{|z-y|\leq R} h_A(t, x, z)\, h_A(s, z, y)\, dz \\ &= I_1 + I_2. \end{aligned}$$

>From (3.3) it follows that, for any fixed $\varepsilon > 0$, we can choose $R = R_\varepsilon > 0$ such that $h_A(s, z, y) \leq \varepsilon$ if $|z - y| > R$. Hence

$$I_1 \leq \varepsilon \int_{\mathbb{R}^n} h_A(t, x, z)\, dz = \varepsilon$$

by (3.2). On the other hand, once R is fixed, (3.3) gives

$$\sup_{0<t<T} I_2 \leq M \exp(-\delta\, |x|^2) \int_{|z-y|\leq R} h_A(s, z, y)\, dz \to 0, \qquad \text{as } |x| \to \infty.$$

Therefore $\sup_{0<t<T} u\,(t, x) \to 0$, as $|x| \to \infty$, and the proof is completed. ■

4 Uniform Gevray estimates and upper bounds of fundamental solutions for large $d\,(x, y)$

Following the plan we have explained in Section 1, from now on in Part I we will study the fundamental solution h_A, whose existence and basic properties are granted by Theorem 3.4. We will adopt Hypotheses and Notation stated in Section 3.

In this Section we follow, adapt and complete the arguments of [30], §2-3, the important difference being that we are considering a class of operators (depending on the matrix $A = \{a_{ij}\}$) instead of a single one.

The strategy is to deduce the exponential decay, for vanishing t, of solutions to the equation $H_A u = 0$, from sharp estimates on time derivatives of any order of u. This will allow to establish the exponential decay of h_A, for vanishing t and (x, y) away from the diagonal, uniformly for $A \in \mathcal{B}_\lambda$. This argument is based on the following fact:

Lemma 4.1 *Let* $f \in C^{\infty}([0,1])$ *be such that the* k*-th derivative* $f^{(k)}(0) = 0, k = 0,1,2,...$ *and such that*

$$\sup_{0\leq t\leq 1}\left|f^{(k)}(t)\right| \leq R^{k+1}(k!)^{2} \quad \text{for some } R>0.$$

Then $|f(t)| \leq R\,e^{-1/(2eRt)}$ *for* $t < 1/(eR)$.

Proof. See Lemma 2 in [30]. ■

Remark 4.2 *The bound on* $f^{(k)}$ *which appears in the assumptions of this Lemma is the condition appearing in the definition of the Gevray class* G^{2}, *whence the term "Gevray estimates".*

Notation 4.3 *Only in this section, we will denote by* $B_{E}(r)$ *the Euclidean balls in* $\mathbb{R}^{n+1}$, *centered at the origin and we will use the notation:*

$$\langle f,g\rangle = \int_{\mathbb{R}^{n+1}} f(t,x)\,g(t,x)\,dtdx; \|f\| = \|f\|_{L^{2}(\mathbb{R}^{n+1})}$$

We will use several times the following instances of Cauchy-Schwarz inequality: for $\mathbf{f} = (f_{1},f_{2},...,f_{m})$, $\mathbf{g} = (g_{1},g_{2},...,g_{m})$, and

$$b(\mathbf{f},\mathbf{g}) = \sum_{i,j=1}^{m} a_{ij}\langle f_{i},g_{j}\rangle,$$

$$|b(\mathbf{f},\mathbf{g})| \leq \sqrt{b(\mathbf{f},\mathbf{f})\cdot b(\mathbf{g},\mathbf{g})}. \tag{4.1}$$

Also, we will use

$$\lambda^{-1}\sum_{i=1}^{m}\|f_{i}\|^{2} \leq b(\mathbf{f},\mathbf{f}) \leq \lambda\sum_{i=1}^{m}\|f_{i}\|^{2}$$

We will also use the more elementary form of (4.1):

$$\left|\sum_{i,j=1}^{m} a_{ij}\xi_{i}\eta_{j}\right| \leq \sqrt{\sum_{i,j=1}^{m} a_{ij}\xi_{i}\xi_{j}}\sqrt{\sum_{i,j=1}^{m} a_{ij}\eta_{i}\eta_{j}} \tag{4.2}$$

Remark 4.4 *on cutoff functions. For a couple of concentric Euclidean balls* $B_{E}(R), B_{E}(r)$ *in* $\mathbb{R}^{n+1}$, *with* $0<r<R<1$, *we will write*

$$B_{E}(r) \preceq \phi \preceq B_{E}(R)$$

to say that ϕ *is a cutoff function with the following standard properties:*

$\phi \in C_{0}^{\infty}\left(\mathbb{R}^{n+1}\right), 0 \leq \phi(t,x) \leq 1$ *for every* $(t,x) \in \mathbb{R}^{n+1}$;

$\phi = 1$ *in* $B_{E}(r)$, $\phi = 0$ *outside* $B_{E}(R)$, $\left|\partial_{t}^{k}\partial_{x}^{\alpha}\phi(t,x)\right| \leq \mathbf{c}(\alpha,k)(R-r)^{-|\alpha|-k}$.

Due to the global boundedness of the coefficients of the vector fields X_{i} *(and their derivatives), if* ϕ *is as above, then it is also true that*

$$\left|\partial_{t}^{k}X_{I}\phi(t,x)\right| \leq \mathbf{c}(k,I)(R-r)^{-|I|-k}$$

Lemma 4.5 *There exists* $\mathbf{c} > 0$ *such that*

$$\|\partial_t u\|^2 \leq \mathbf{c} \left\{ \|(\partial_t - L_A) u\|^2 + \|u\|^2 \right\} \tag{4.3}$$

for every $u \in C_0^\infty\left(\mathbb{R}^{n+1}\right)$.

Proof. We start noting that, if $X_1, X_2, ..., X_m$ are general Hörmander vector fields, $X_j = \sum_{i=1}^n b_{ji}(x)\, \partial_{x_i}$, and the matrix A is symmetric, then

$$X_j^* = -X_j + a_j \text{ with } a_j = -\sum_{i=1}^n \partial_{x_i}(b_{ji})$$

$$L_A^* = \left(\sum_{i,j=1}^m a_{ij} X_i X_j \right)^* = \sum_{i,j=1}^m a_{ij} X_i^* X_j^* = \sum_{i,j=1}^m a_{ij} \left(-X_i + a_i\right)\left(-X_j + a_j\right) =$$
$$= L_A + \sum_{i,j=1}^m a_{ij} \left(-X_i\left(a_j \cdot\right) - a_i X_j + a_i a_j\right).$$

Suppose that $u \in C_0^\infty\left(\mathbb{R}^{n+1}\right)$; then

$$\langle \partial_t u, u \rangle = \langle u, -\partial_t u \rangle = -\langle \partial_t u, u \rangle, \text{ hence } \langle \partial_t u, u \rangle = 0$$

$$\langle (\partial_t - L_A) u, u \rangle = \langle -L_A u, u \rangle = \left\langle -\sum_{i,j=1}^m a_{ij} X_i X_j u, u \right\rangle =$$
$$= \sum_{i,j=1}^m a_{ij} \langle X_j u, X_i u - a_i u \rangle \geq$$
$$\geq \frac{1}{\lambda} \sum_{i=1}^m \|X_i u\|^2 - \lambda \sqrt{\sum_{i=1}^m \|X_i u\|^2} \sqrt{\sum_{i=1}^m \|a_i u\|^2} \geq$$
$$\geq \frac{1}{2\lambda} \sum_{i=1}^m \|X_i u\|^2 - \mathbf{c} \|u\|^2$$

so that

$$\sum_{i=1}^m \|X_i u\|^2 \leq \mathbf{c} \left\{ \|(\partial_t - L_A) u\|^2 + \|u\|^2 \right\}. \tag{4.4}$$

Next, we note that, since ∂_t commutes with L_A and L_A^*,

$$\langle \partial_t u, (L_A + L_A^*) u \rangle = \langle (L_A + L_A^*) u, -\partial_t u \rangle \text{ implies } \langle \partial_t u, (L_A + L_A^*) u \rangle = 0.$$

Therefore (subtracting $\frac{1}{2}\langle\partial_t u, (L_A + L_A^*) u\rangle$), we have

$$\begin{aligned}\|\partial_t u\|^2 &= \langle\partial_t u, \partial_t u\rangle = \langle(\partial_t - L_A) u, \partial_t u\rangle + \langle\partial_t u, L_A u\rangle = \\ &= \langle(\partial_t - L_A) u, \partial_t u\rangle + \frac{1}{2}\langle\partial_t u, (L_A - L_A^*) u\rangle = \\ &= \langle(\partial_t - L_A) u, \partial_t u\rangle + \frac{1}{2}\left\langle\partial_t u, \left(\sum_{i,j=1}^m a_{ij}(X_i(a_j\cdot) + a_i X_j - a_i a_j)\right) u\right\rangle \leq \\ &\leq \|(\partial_t - L_A) u\| \|\partial_t u\| + \mathbf{c}\|\partial_t u\| \left\{\sqrt{\sum_{i=1}^m \|X_i u\|^2} + \|u\|\right\} \leq \\ &\leq \frac{1}{2}\|\partial_t u\|^2 + \mathbf{c}\left\{\|(\partial_t - L_A) u\|^2 + \sum_{i=1}^m \|X_i u\|^2 + \|u\|^2\right\} \leq\end{aligned}$$

by (4.4)

$$\leq \frac{1}{2}\|\partial_t u\|^2 + \mathbf{c}\left\{\|(\partial_t - L_A) u\|^2 + \|u\|^2\right\}$$

and finally

$$\|\partial_t u\|^2 \leq \mathbf{c}\left\{\|(\partial_t - L_A) u\|^2 + \|u\|^2\right\}$$

■

Lemma 4.6 *There is a constant R (depending on A only through λ) such that if $H_A u(t,x) = 0$ in $B_E(1)$, then*

$$\left\|\partial_t^k u\right\|_{L^2(B_E(1/2))} \leq R^{k+1}(k!)^2 \|u\|_{L^2(B_E(1))} \text{ for } k = 0, 1, 2...$$

Proof. We will prove by induction that there exists R_0 such that for any ε, $0 < \varepsilon < 1$ and any nonnegative integer k,

$$\left\|\partial_t^k u\right\|_{L^2(B_E(1-2k\varepsilon))} \leq R_0^{k+1}\varepsilon^{-2k}\|u\|_{L^2(B_E(1))}. \tag{4.5}$$

This implies the Lemma because if we put $\varepsilon = \frac{1}{4k}$ we obtain

$$\left\|\partial_t^k u\right\|_{L^2(B_E(1/2))} \leq R_0^{k+1} 16^k k^{2k}\|u\|_{L^2(B_E(1))} \leq R^{k+1}(k!)^2\|u\|_{L^2(B_E(1))}$$

for $R = 16e^2 R_0$ (where we have used $k! \geq k^k e^{-k}\sqrt{2\pi k}$).
The case $k = 0$ is trivially true provided $R_0 \geq 1$. For $k \geq 1$, choose two cutoff functions $\varphi_{k,\varepsilon}, \psi_{k,\varepsilon}$ such that

$$\begin{gathered}B_E(1 - 2k\varepsilon) \preceq \varphi_{k,\varepsilon} \preceq B_E(1 - (2k-1)\varepsilon) \\ B_E(1 - (2k-1)\varepsilon) \preceq \psi_{k,\varepsilon} \preceq B_E(1 - (2k-2)\varepsilon)\end{gathered}$$

Then we have

$$\begin{aligned}\left\|\partial_t^k u\right\|_{L^2(B_E(1-2k\varepsilon))} &\leq \left\|\partial_t\left(\varphi_{k,\varepsilon}\partial_t^{k-1}u\right)\right\| \\ \text{by (4.3)} &\leq \mathbf{c}\left(\left\|(\partial_t - L_A)\left(\varphi_{k,\varepsilon}\partial_t^{k-1}u\right)\right\| + \left\|\varphi_{k,\varepsilon}\partial_t^{k-1}u\right\|\right).\end{aligned}$$

Let us compute

$$\begin{aligned}&\left(\partial_t - L_A\right)\left(\varphi_{k,\varepsilon}\partial_t^{k-1}u\right) = \left(\left(\partial_t - L_A\right)\varphi_{k,\varepsilon}\right)\partial_t^{k-1}u + \\ &-2\sum_{i,j=1}^{m} a_{ij}\left(X_j\varphi_{k,\varepsilon}\right)\left(X_i\partial_t^{k-1}u\right) + \varphi_{k,\varepsilon}\partial_t^{k-1}\left(\left(\partial_t - L_A\right)u\right)\end{aligned}$$

Because $\left(\partial_t - L_A\right)u = 0$ and using the hypotheses on $\varphi_{k,\varepsilon}$, we have

$$\begin{aligned}&\left\|\partial_t^k u\right\|_{L^2(B_E(1-2k\varepsilon))} \\ &\leq \mathbf{c}\left(\varepsilon^{-2}\left\|\partial_t^{k-1}u\right\|_{L^2(B_E(1-(2k-1)\varepsilon))} + \left\|\sum_{i,j=1}^{m} a_{ij}\left(X_j\varphi_{k,\varepsilon}\right)\left(X_i\partial_t^{k-1}u\right)\right\|\right).\end{aligned} \tag{4.6}$$

Now,

$$\begin{aligned}&\left|\sum_{i,j=1}^{m} a_{ij}\left(X_j\varphi_{k,\varepsilon}\right)\left(X_i\partial_t^{k-1}u\right)\right| \\ &\leq \sqrt{\sum_{i,j=1}^{m} a_{ij}\left(X_i\varphi_{k,\varepsilon}\right)\left(X_j\varphi_{k,\varepsilon}\right)} \cdot \sqrt{\sum_{i,j=1}^{m} a_{ij}\left(X_i\partial_t^{k-1}u\right)\left(X_j\partial_t^{k-1}u\right)} \\ &\leq \mathbf{c}\varepsilon^{-1}\psi_{k,\varepsilon}\sqrt{\sum_{i,j=1}^{m} a_{ij}\left(X_i\partial_t^{k-1}u\right)\left(X_j\partial_t^{k-1}u\right)}\end{aligned}$$

so that

$$\begin{aligned}&\left\|\sum_{i,j=1}^{m} a_{ij}\left(X_j\varphi_{k,\varepsilon}\right)\left(X_i\partial_t^{k-1}u\right)\right\|^2 \\ &\leq \mathbf{c}\varepsilon^{-2}\sum_{i,j=1}^{m} a_{ij}\int_{\mathbb{R}^{n+1}}\psi_{k,\varepsilon}^2\left(X_i\partial_t^{k-1}u\right)\left(X_j\partial_t^{k-1}u\right)dxdt \\ &= \mathbf{c}\varepsilon^{-2}\sum_{i,j=1}^{m} a_{ij}\left\langle\psi_{k,\varepsilon}^2\left(X_i\partial_t^{k-1}u\right),\left(X_j\partial_t^{k-1}u\right)\right\rangle.\end{aligned} \tag{4.7}$$

On the other hand,

$$\begin{aligned}
&\sum_{i,j=1}^{m} a_{ij}\left\langle \psi_{k,\varepsilon}^{2}\left(X_{i}\partial_{t}^{k-1}u\right),\left(X_{j}\partial_{t}^{k-1}u\right)\right\rangle = \\
&=-\sum_{i,j=1}^{m} a_{ij}\left\langle X_{j}\left(\psi_{k,\varepsilon}^{2}\right)\left(X_{i}\partial_{t}^{k-1}u\right),\partial_{t}^{k-1}u\right\rangle + \\
&+\sum_{i,j=1}^{m} a_{ij}\left\langle a_{j}\psi_{k,\varepsilon}^{2}\left(X_{i}\partial_{t}^{k-1}u\right),\partial_{t}^{k-1}u\right\rangle + \\
&-\left\langle \psi_{k,\varepsilon}^{2}\partial_{t}^{k-1}\left(\sum_{i,j=1}^{m} a_{ij}X_{i}X_{j}u\right),\partial_{t}^{k-1}u\right\rangle = \\
&=A+B+C \qquad (4.8)
\end{aligned}$$

Using the equation $\partial_t u = L_A u$,we have

$$C=-\left\langle \psi_{k,\varepsilon}^{2}\partial_{t}^{k}u,\partial_{t}^{k-1}u\right\rangle$$

We note that

$$\begin{aligned}
\left\langle \psi_{k,\varepsilon}^{2}\partial_{t}^{k}u,\partial_{t}^{k-1}u\right\rangle &= \left\langle \partial_{t}\left(\psi_{k,\varepsilon}^{2}\partial_{t}^{k-1}u\right),\partial_{t}^{k-1}u\right\rangle-\left\langle \left(\partial_{t}\psi_{k,\varepsilon}^{2}\right)\partial_{t}^{k-1}u,\partial_{t}^{k-1}u\right\rangle \\
&=-\left\langle \psi_{k,\varepsilon}^{2}\partial_{t}^{k-1}u,\partial_{t}^{k}u\right\rangle-\left\langle \left(\partial_{t}\psi_{k,\varepsilon}^{2}\right)\partial_{t}^{k-1}u,\partial_{t}^{k-1}u\right\rangle
\end{aligned}$$

hence

$$\begin{aligned}
|C| &= \left|\left\langle \psi_{k,\varepsilon}^{2}\partial_{t}^{k}u,\partial_{t}^{k-1}u\right\rangle\right| = \frac{1}{2}\left|\left\langle \left(\partial_{t}\psi_{k,\varepsilon}^{2}\right)\partial_{t}^{k-1}u,\partial_{t}^{k-1}u\right\rangle\right| \\
&\leq \mathbf{c}\varepsilon^{-1}\left\|\partial_{t}^{k-1}u\right\|_{L^{2}(B_{E}(1-(2k-2)\varepsilon))}^{2} \qquad (4.9)
\end{aligned}$$

To bound A, B in (4.8), we make the following remark. For every $\psi \in C_0^\infty\left(\mathbb{R}^{n+1}\right)$,

$$\begin{aligned}
\left\langle \psi X_{j}v,v\right\rangle &= \left\langle X_{j}\left(\psi v\right),v\right\rangle-\left\langle \left(X_{j}\psi\right)v,v\right\rangle = \left\langle \psi v,X_{j}^{*}v\right\rangle-\left\langle \left(X_{j}\psi\right)v,v\right\rangle = \\
&=-\left\langle \psi v,X_{j}v\right\rangle+\left\langle \psi v,a_{j}v\right\rangle-\left\langle \left(X_{j}\psi\right)v,v\right\rangle
\end{aligned}$$

hence

$$\left\langle \psi X_{j}v,v\right\rangle=\frac{1}{2}\left(\left\langle \psi v,a_{j}v\right\rangle-\left\langle \left(X_{j}\psi\right)v,v\right\rangle\right)$$

and

$$\left|\left\langle \psi X_{j}v,v\right\rangle\right| \leq \mathbf{c}\left(\sup|\psi|\left\|v\right\|_{L^{2}(\operatorname{supp}\psi)}^{2}+\sup\left|X_{j}\psi\right|\left\|v\right\|_{L^{2}(\operatorname{supp}X_{j}\psi)}^{2}\right). \qquad (4.10)$$

For $v = \partial_t^{k-1}u$, (4.10) implies

$$|B| = \left|\sum_{i,j=1}^{m} a_{ij}\left\langle a_{j}\psi_{k,\varepsilon}^{2}\left(X_{i}\partial_{t}^{k-1}u\right),\partial_{t}^{k-1}u\right\rangle\right| \leq \mathbf{c}\varepsilon^{-1}\left\|\partial_{t}^{k-1}u\right\|_{L^{2}(B_{E}(1-(2k-2)\varepsilon))}^{2} \qquad (4.11)$$

Similarly,

$$
\begin{aligned}
|A| &= \left|\sum_{i,j=1}^{m} a_{ij}\left\langle X_j\left(\psi_{k,\varepsilon}^2\right)\left(X_i\partial_t^{k-1}u\right),\partial_t^{k-1}u\right\rangle\right| = \\
&= \left|\sum_{i,j=1}^{m} a_{ij}\left\langle 2\psi_{k,\varepsilon}\left(X_j\psi_{k,\varepsilon}\right)\left(X_i\partial_t^{k-1}u\right),\partial_t^{k-1}u\right\rangle\right| \le \\
&\le \mathbf{c}\varepsilon^{-2}\left\|\partial_t^{k-1}u\right\|_{L^2(B_E(1-(2k-2)\varepsilon))}^2 \qquad (4.12)
\end{aligned}
$$

Inequalities (4.6), (4.7), (4.8), (4.9), (4.11), (4.12) imply

$$
\begin{aligned}
&\left\|\partial_t^k u\right\|_{L^2(B_E(1-2k\varepsilon))} \\
&\le \mathbf{c}\varepsilon^{-2}\left\|\partial_t^{k-1}u\right\|_{L^2(B_E(1-(2k-1)\varepsilon))} + \\
&+ \mathbf{c}\varepsilon^{-1}\sqrt{\sum_{i,j=1}^{m} a_{ij}\left\langle \psi_{k,\varepsilon}^2\left(X_i\partial_t^{k-1}u\right),\left(X_j\partial_t^{k-1}u\right)\right\rangle} \\
&\le \mathbf{c}\varepsilon^{-2}\left\|\partial_t^{k-1}u\right\|_{L^2(B_E(1-(2k-1)\varepsilon))} + \\
&+ \mathbf{c}\varepsilon^{-1}\left\{\varepsilon^{-1/2}\left\|\partial_t^{k-1}u\right\|_{L^2(B_E(1-(2k-2)\varepsilon))} + \varepsilon^{-1}\left\|\partial_t^{k-1}u\right\|_{L^2(B_E(1-(2k-2)\varepsilon))}\right\} \qquad (4.13)
\end{aligned}
$$

Therefore,

$$
\left\|\partial_t^k u\right\|_{L^2(B_E(1-2k\varepsilon))} \le \mathbf{c}\varepsilon^{-2}\left\|\partial_t^{k-1}u\right\|_{L^2(B_E(1-(2k-2)\varepsilon))}
$$

by the inductive assumption

$$
\le \mathbf{c}\varepsilon^{-2}R_0^k\varepsilon^{-2k+2}\left\|u\right\|_{L^2(B_E(1))} = R_0^{k+1}\varepsilon^{-2k}\left\|u\right\|_{L^2(B_E(1))}
$$

choosing $R_0 = \mathbf{c}$. This proves (4.5), and so the Lemma. ■

Remark 4.7 *We will need Lemma 4.6 in the following form, adapted to a ball of radius r: for any $r > 0$, there is a constant R such that if $H_A u(t,x) = 0$ in $B_E(r)$, then*

$$
\left\|\partial_t^k u\right\|_{L^2(B_E(r/2))} \le R^{k+1}(k!)^2\left\|u\right\|_{L^2(B_E(r))}. \qquad (4.14)
$$

To get this inequality, we can argue as follows. With the same reasoning of the proof of Lemma 4.6, using cutoff functions adapted to the balls $B_E(r(1-2k\varepsilon))$ etc., one proves, by induction, that there exists R_0 such that for any ε, $0 < \varepsilon < 1$ and any nonnegative integer k,

$$
\left\|\partial_t^k u\right\|_{L^2(B_E(r(1-2k\varepsilon)))} \le R_0^{k+1}(\varepsilon r)^{-2k}\left\|u\right\|_{L^2(B_E(r))}. \qquad (4.15)
$$

This implies (4.14) because if we put $\varepsilon = \frac{1}{4k}$ we obtain

$$
\left\|\partial_t^k u\right\|_{L^2(B_E(r/2))} \le R_0^{k+1}16^k k^{2k} r^{-2k}\left\|u\right\|_{L^2(B_E(r))}
$$

Now,

$$R_0^{k+1}16^k k^{2k} r^{-2k} \leq R^{k+1} k^{2k} e^{-2k} 2\pi k \leq R^{k+1} (k!)^2$$

where the first inequality holds provided we choose

$$R = R_0 \text{ if } \frac{16e^2}{r^2} \leq 1; \; R = R_0 \frac{16e^2}{r^2} \text{ if } \frac{16e^2}{r^2} > 1.$$

To transform the L^2 bound of the previous Lemma in a supremum bound on $\partial_t^k u$, we will make use of the following subelliptic estimate, uniform in A:

Lemma 4.8 *For every couple of cutoff functions φ_1, φ_2, with $\varphi_2 = 1$ on $\operatorname{supp}\varphi_1$, and every $s > 0$, there exists a constant $\mathbf{c}(\varphi_1, \varphi_2, s)$ (depending on A only through λ), such that*

$$\|\varphi_1 w\|_{H^s} \leq \mathbf{c}(\varphi_1, \varphi_2, s) \{\|\varphi_2 (\partial_t - L_A) w\|_{H^s} + \|\varphi_2 w\|\} \tag{4.16}$$

for every $w \in C^\infty(\operatorname{supp}\varphi_2)$. Here H^s denotes the ordinary Sobolev space of fractional order on $\mathbb{R}^{n+1}$.

This classical bound (see for instance Kohn [31]) holds uniformly with respect to A in $\mathcal{B}_\lambda$, as pointed out in [9]; moreover, the same reasoning of [9] applies when $X_1, ..., X_m$ are general Hörmander vector fields (not necessarily left invariant with respect to a Lie group structure). Recall also that, by Sobolev' Lemma, for any $s > (n+1)/2$ we have

$$\sup |\varphi_1 w| \leq \mathbf{c}(s, \varphi_1) \|\varphi_1 w\|_{H^s} \tag{4.17}$$

Proposition 4.9 *Let $r > 0$; there exists a constant $R_1 = \mathbf{c}(r)$ such that, for any ball $B_E(r) \subset \mathbb{R}^{n+1}$, if $(\partial_t - L_A) v = 0$ in $B_E(r)$, then for any nonnegative integer k,*

$$\sup_{B_E(r/4)} \left|\partial_t^k v\right| \leq R_1^{k+1} (k!)^2 \|v\|_{L^2(B_E(r))}. \tag{4.18}$$

Proof. Applying (4.16) to $w = \partial_t^k v$ (which satisfies $(\partial_t - L_A) w = 0$ in $B_E(r)$) and (4.17), and choosing two cutoff functions

$$B_E(r/4) \preceq \varphi_1 \preceq B_E(3r/8); B_E(3r/8) \preceq \varphi_2 \preceq B_E(r/2)$$

we get

$$\begin{aligned}
\sup_{B_E(r/4)} \left|\partial_t^k v\right| &\leq \sup_{B_E(r/2)} \left|\varphi_1 \partial_t^k v\right| \leq \mathbf{c} \left\|\varphi_1 \partial_t^k v\right\|_{H^s} \\
&\leq \mathbf{c} \left\{\left\|\varphi_2 (\partial_t - L_A) \partial_t^k v\right\|_{H^s} + \left\|\varphi_2 \partial_t^k v\right\|\right\} = \\
&= \mathbf{c} \left\|\varphi_2 \partial_t^k v\right\| \leq \mathbf{c} \left\|\partial_t^k v\right\|_{L^2(B_E(r/2))} \leq \mathbf{c} R^{k+1} (k!)^2 \|v\|_{L^2(B_E(r))}
\end{aligned}$$

which gives (4.18) with $R_1 = \mathbf{c}R$. In the last inequality we have applied (4.14). ■

We now apply to the fundamental solution of H_A the previous results, to get an upper Gaussian bound for h_A, which holds when $d(x, y)$ is bounded away from zero.

Theorem 4.10 *There exists a constant τ such that for any $R > \varepsilon > 0$ we have*

$$h_A(t,x,y) \le \mathbf{c}_1(\varepsilon,R)\, e^{-\mathbf{c}/t}$$

for every $t \in (0,\tau)$, $x,y \in \mathbb{R}^n$ with $\varepsilon \le d(x,y) \le R$.

Proof. For any fixed $y \in \mathbb{R}^n$, let

$$C = C_{\varepsilon,R} = \{(t,x) : 0 \le t \le 1, \varepsilon < d(x,y) < R\}.$$

Let us cover C with a finite number of balls B_j such that the balls B_j^* concentric with B_j and with radius four times the radius of B_j are at positive distance from the origin, let $C' = \bigcup B_j^*$. Applying Proposition 4.9 to h_A in each ball, and then passing to the sup with respect to the family B_j we have

$$\sup_C \left|\partial_t^k h_A(\cdot,\cdot,y)\right| \le R^k (k!)^2 \|h_A(\cdot,\cdot,y)\|_{L^2(C')} \le \mathbf{c}R^k (k!)^2 \tag{4.19}$$

for the uniform L^2 bound on h_A contained in the next theorem. Note also that the number of balls B_j can be chosen independently of y. Then Lemma 4.1 implies the theorem. ■

Theorem 4.11 *For any $R > \varepsilon > 0, T > 0$, there exists a constant $\mathbf{c}(\varepsilon,R,T)$, such that*

$$\sup_{y\in\mathbb{R}^n} \int_0^T d\tau \int_{\varepsilon<d(x,y)<R} h_A(\tau,x,y)^2\, dx \le \mathbf{c}(\varepsilon,R,T)$$

This fact will be proved in the next section.

5 Fractional integrals and uniform L^2 bounds of fundamental solutions for large $d(x,y)$

The goal of this section is to prove Theorem 4.11, and therefore complete the proof of Theorem 4.10. To prove this uniform bound on h_A, we will follow the approach used by Bramanti, Brandolini in [9], §4, which in turn is based on ideas contained in the papers by Rothschild-Stein [49] and Kohn [31]. The basic tool is an L^2 estimate for the fractional integral operator with kernel h_A, uniform for A ranging in the class $\mathcal{B}_\lambda$ (see Section 3). The suitable framework for this fractional integral estimate is that of spaces of homogeneous type, in the sense of Coifman-Weiss [16].

Recall that the symbol $B(x,r) = B_r(x)$ denotes balls with respect to the CC-distance d, that we have defined and studied in Section 2. We will also use the "parabolic CC-distance" in $\mathbb{R}^{n+1}$,

$$d_P((t,x),(s,y)) = \sqrt{d(x,y)^2 + |t-s|}$$

and the corresponding balls

$$B_P((t,x),r) = \left\{(s,y) \in \mathbb{R}^{n+1} : d_P((t,x),(s,y)) < r\right\}.$$

Note that

$$\left|B_P\left((s,y), \sqrt{t-s}\right)\right|_{\mathbb{R}^{n+1}} \simeq (t-s)\left|B\left(y, \sqrt{t-s}\right)\right|_{\mathbb{R}^n}$$

where $|\cdot|_{\mathbb{R}^n}$ denotes the Lebesgue measure in $\mathbb{R}^n$. It is easy to prove that d_P is a distance, which satisfies a global doubling condition (since the same is true for d, by Proposition 2.5). Furthermore:

Lemma 5.1 *For any* $(t_0, x_0) \in \mathbb{R}^{n+1}, R_0 > 0$

$$\left(B_P\left((t_0,x_0), R_0\right), d_P, dtdx\right)$$

is a space of homogeneous type, that is the doubling condition holds within this space; explicitly, this means that

$$\left|B_P\left((t,x), 2r\right) \cap B_P\left((t_0,x_0), R_0\right)\right| \leq \mathbf{c}\left|B_P\left((t,x), r\right) \cap B_P\left((t_0,x_0), R_0\right)\right|$$

for any $(t,x) \in B_P\left((t_0,x_0), R_0\right), r > 0$. *Moreover, the constant appearing in the doubling condition can be chosen independently of* (t_0, x_0).

Proof. This fact immediately follows from the global doubling condition which holds for d_P, plus the following property, proved in Proposition 3.8 of [12]: there exists $\mathbf{c} > 0$ such that

$$\left|B_P\left((t,x), r\right) \cap B_P\left((t_0,x_0), R_0\right)\right| \geqslant \mathbf{c}\left|B_P\left((t,x), r\right)\right|$$

for every $(t,x) \in B_P\left((t_0,x_0), R_0\right), 0 < r < 2R_0$. ■

Now, let

$$B_0 = B_P\left((t_0,x_0), R_0\right)$$

and let $h_A(t,x,y)$ be the fundamental solution of H_A in $\mathbb{R}^{n+1}$; also, set

$$G_A\left((t,x),(s,y)\right) = h_A(t-s,x,y).$$

Notation 5.2 *Keeping our convention about constants, in this section we will write* $\mathbf{c}(A)$ *for a constant depending on the coefficients of the matrix* A *in any unspecified way, and* $\mathbf{c}$ *for a constant depending on* A *only through the ellipticity constant* λ.

By the (nonuniform) Gaussian bounds proved by Jerison-Sanchez-Calle [30] or Kusuoka-Stroock [35], [36], we know that, for $(t,x),(s,y) \in B_0$,

$$\begin{aligned} G_A\left((t,x),(s,y)\right) &\leq \mathbf{c}(A,R_0,T)\frac{e^{-\mathbf{c}\frac{d(x,y)^2}{t-s}}}{\left|B\left(y,\sqrt{t-s}\right)\right|} \\ &\leq \mathbf{c}(A,R_0,T)(t-s)\frac{e^{-\mathbf{c}\frac{d(x,y)^2}{t-s}}}{\left|B_P\left((s,y),\sqrt{t-s}\right)\right|}. \end{aligned} \tag{5.1}$$

Similarly, setting

$$G_A'\left((t,x),(s,y)\right) = X_j^x h_A(t-s,x,y) \text{ for any } j = 1,2,..,m$$

we have

$$G'_A((t,x),(s,y)) \leq \mathbf{c}(A,R_0,T)\sqrt{t-s}\frac{e^{-\mathbf{c}\frac{d(x,y)^2}{t-s}}}{\left|B_P\left((s,y),\sqrt{t-s}\right)\right|} \tag{5.2}$$

With these bounds, we can prove the following:

Lemma 5.3 *For* $(t,x),(s,y)\in B_0$,

$$G_A((t,x),(s,y)) \leq \mathbf{c}(A,R_0,T)\frac{d_P^2((t,x),(s,y))}{|B_P((s,y),d_P((t,x),(s,y)))|}$$

$$G'_A((t,x),(s,y)) \leq \mathbf{c}(A,R_0,T)\frac{d_P((t,x),(s,y))}{|B_P((s,y),d_P((t,x),(s,y)))|}.$$

Proof. Let us assume $s<t$, otherwise the kernels $G_A((t,x),(s,y)), G'_A((t,x),(s,y))$ vanish. The doubling condition for d_P implies the existence of some $\alpha>0$ such that

$$\frac{|B_P(\xi,r)|}{|B_P(\xi,\rho)|} \leq \mathbf{c}\left(\frac{r}{\rho}\right)^{\alpha} \text{ for any } \rho<r \tag{5.3}$$

It is not restrictive to assume $\alpha>2$, say $\alpha=2+\varepsilon$. Then

$$\begin{aligned} &G_A((t,x),(s,y)) \leq \\ &\leq \mathbf{c}(A,R_0,T)\frac{d_P((t,x),(s,y))^2 e^{-\mathbf{c}\frac{d(x,y)^2}{t-s}}}{|B_P((s,y),d_P((t,x),(s,y)))|}\cdot\left(\frac{d(x,y)^2+|t-s|}{|t-s|}\right)^{\varepsilon/2} \\ &\leq \mathbf{c}(A,R_0,T)\frac{d_P((t,x),(s,y))^2}{|B_P((s,y),d_P((t,x),(s,y)))|} \end{aligned}$$

because

$$A \equiv e^{-\mathbf{c}\frac{d(x,y)^2}{t-s}}\left(\frac{d(x,y)^2+|t-s|}{|t-s|}\right)^{\varepsilon/2}$$

is bounded: if $d(x,y)^2 \leq |t-s|$, then $A \leq \mathbf{c}e^{-\mathbf{c}\frac{d(x,y)^2}{t-s}} \leq \mathbf{c}$; if $d(x,y)^2 > |t-s|$, then

$$A \leq \mathbf{c}e^{-\mathbf{c}\frac{d(x,y)^2}{t-s}}\left(\frac{d(x,y)^2}{|t-s|}\right)^{\varepsilon/2} \leq \mathbf{c}$$

because $e^{-u}u^{\varepsilon/2}$ is bounded for $u\geq 0$. Analogously, by (5.2), one proves the bound on G'. ∎

Lemma 5.4 *In a bounded space of homogeneous type* (X,d,dx), *let*

$$Tf(x)=\int_X G(x,y)f(y)\,dy$$

with

$$|G(x,y)| \leq c\frac{d(x,y)^{\alpha}}{|B(x,d(x,y))|} \text{ for some } \alpha > 0.$$

Then T is continuous on $L^p(X)$ (for $1 \leq p \leq \infty$), with norm dependent on c, α, the doubling constant and the diameter of X.

Proof. Let us show that

$$\sup_x \int |G(x,y)|\,dy + \sup_y \int |G(x,y)|\,dx \leq \mathbf{c} < \infty.$$

This implies the continuity of T on L^p for every $p \in [1,\infty]$ (see e.g. Theorem 6.18 p.193 in [25]). Since X is bounded, for some $R > 0$ we have

$$\begin{aligned}\int |G(x,y)|\,dy &\leq c\sum_k \int_{d(x,y)\simeq R2^{-k}} \frac{d(x,y)^{\alpha}}{|B(x,d(x,y))|}dy \leq \\ &\leq c\sum_k \frac{\left(R2^{-k}\right)^{\alpha}}{|B(x,R2^{-k-1})|}\left|B\left(x,R2^{-k}\right)\right| \leq \\ &\leq cR^{\alpha}\sum_k 2^{-k\alpha} = \mathbf{c}\end{aligned}$$

by the doubling condition. The other estimate is similar, since

$$|B(x,d(x,y))| \simeq |B(y,d(x,y))|,$$

again by the doubling condition. ∎

Proposition 5.5 *Let*

$$T_A g(t,x) = \int_{B_0} h_A(t-\tau,x,y)\,g(\tau,y)\,d\tau dy.$$

Then, for $1 < p < \infty$:
i)

$$\|T_A g\|_{L^p(B_0)} + \|X_j T_A g\|_{L^p(B_0)} \leq \mathbf{c}(A,p,R_0)\,\|g\|_{L^p(B_0)} \text{ for } j = 1,...m$$

ii)

$$\|X_i X_j T_A g\|_{L^p(B_0)} \leq \mathbf{c}(A,p,R_0)\,\|g\|_{L^p(B_0)} \text{ for } i,j = 1,...m$$

iii)

$$\|T_A g\|_{L^2(B_0)} \leq \mathbf{c}(R_0)\,\|g\|_{L^2(B_0)}.$$

Remark 5.6 *The key feature of the above proposition is that the constant in estimate iii) depends on the matrix A only through the number λ.*

Proof. i) follows by Lemma 5.3 and Lemma 5.4. ii) is a consequence of the following estimate, which is essentially contained in Rothschild-Stein [49]: for every $u \in C^{\infty}\left(\mathbb{R}^{n+1}\right)$,

$$\begin{aligned} & \left\|X_i X_j u\right\|_{L^p(B_0)} \leq \\ & \leq \mathbf{c}\left(A, p, R_0\right)\left\{\|u\|_{L^p(2B_0)}+\sum_k\left\|X_k u\right\|_{L^p(2B_0)}+\left\|H_A u\right\|_{L^p(2B_0)}\right\} \end{aligned} \tag{5.4}$$

(with $2B_0 = B_P\left(\left(t_0, x_0\right), 2R_0\right)$). As to the dependence of $\mathbf{c}$ on the domain B_0, we note that: for R_0 fixed, when B_0 is contained in a fixed compact set, then $\mathbf{c}$ can be chosen independently of (t_0, x_0) by a simple covering argument; on the other side, if (t_0, x_0) is far away from the origin, then H_A is the heat operator, and estimates (5.4) are the standard translation invariant parabolic estimates. Now, let $g \in C_0^{\infty}\left(B_0\right)$, and $u = T_A g$. Then the above inequality and point i) imply

$$\begin{aligned} & \left\|X_i X_j T_A g\right\|_{L^p(B_0)} \leq \\ & \leq \mathbf{c}\left(A, p, R_0\right)\left\{\left\|T_A g\right\|_{L^p(2B_0)}+\sum_k\left\|X_k T_A g\right\|_{L^p(2B_0)}+\|g\|_{L^p(2B_0)}\right\} \\ & \leq \mathbf{c}\left(A, p, R_0\right)\|g\|_{L^p(2B_0)}=\mathbf{c}\left(A, p, R_0\right)\|g\|_{L^p(B_0)}. \end{aligned}$$

iii) follows from i) and ii) with the same reasoning of [9], Lemma 3 p.414-5. ■

Recall that our task is to prove Theorem 4.11, that is a uniform bound of the kind

$$\sup_{y \in \mathbb{R}^n} \int_0^T d\tau \int_{\varepsilon<d(x,y)<R} h_A(\tau, x, y)^2 \, dx \leq \mathbf{c}(\varepsilon, R, T).$$

Remark 5.7 *It is enough to prove the above bound for $T = R^2$, since*

$$\int_0^T d\tau \int_{\varepsilon<d(x,y)<R} h_A(\tau, x, y)^2 \, dx \leq \int_0^{r^2} d\tau \int_{\varepsilon<d(x,y)<r} h_A(\tau, x, y)^2 \, dx$$

with $r = \max\left(R, \sqrt{T}\right)$.

Hence, for any fixed $x \in \mathbb{R}^n$, let

$$\mathcal{C} = \mathcal{C}_{x,\varepsilon,R} = \left\{(t, y) : 0 \leq t \leq R^2, \varepsilon < d(x, y) < R\right\};$$

$$\mathcal{C}' = \mathcal{C}'_{x,\varepsilon,R} = \left\{(t, y) : -2\varepsilon \leq t \leq R^2 + 2\varepsilon, \varepsilon < d(x, y) < R\right\};$$

$$Q = \left\{(t, y) : -2\varepsilon \leq t \leq R^2 + 2\varepsilon, d(x, y) < R\right\},$$

and choose R_0, comparable to R, such that

$$Q \subset B_0.$$

For any $g \in C_0^\infty(\mathcal{C}')$, define

$$T_A g(t,x) = \int_Q h_A(t-\tau,x,y)\, g(\tau,y)\, d\tau dy$$

The following proposition states a result similar to what we have to prove, with an interchanged role of the variables x, y.

Proposition 5.8

$$\sup_{x\in\mathbb{R}^n} \int_0^{R^2} d\tau \int_{\varepsilon<d(x,y)<R} h_A(\tau,x,y)^2\, dy \leq \mathbf{c}(\varepsilon,R)$$

Proof. For a fixed $x \in \mathbb{R}^n$, pick two cutoff functions φ, φ_1 such that

$$\left(0,R^2\right)\times B_{\varepsilon/4}(x) \prec \varphi \prec \left(-\varepsilon, R^2+\varepsilon\right)\times B_{\varepsilon/2}(x) \prec \varphi_1 \prec (-2\varepsilon, R+2\varepsilon)\times B_\varepsilon(x).$$

Let $g \in C_0^\infty(\mathcal{C})$ with $\|g\|_2 \leq 1$, and let $f = T_A g$. Since $H_A f = g$, by the hypoellipticity of H_A, f is smooth. Since $g=0$ in $\left(0,R^2\right)\times B_\varepsilon$, then $\varphi_1 g \equiv 0$, and by the uniform subelliptic estimate (4.16) and Proposition 5.5 iii), we can write

$$\begin{aligned}
\|\varphi f\|_{H^s} &\leq \mathbf{c}(s,\varphi_1,\varphi)\|\varphi_1 f\|_{L^2} \leq \\
&\leq \mathbf{c}(s,\varphi_1,\varphi)\|f\|_{L^2((-2\varepsilon,R^2+2\varepsilon)\times B_R(x))} = \mathbf{c}(s,\varphi_1,\varphi)\|T_A g\|_{L^2(Q)} \leq \\
&\leq \mathbf{c}(s,\varphi_1,\varphi)\|T_A g\|_{L^2(B_0)} \leq \mathbf{c}(s,\varphi_1,\varphi)\|g\|_{L^2(B_0)} \leq \mathbf{c}(s,\varphi_1,\varphi).
\end{aligned}$$

This bound, with a suitable choice of s, implies, by the standard Sobolev embedding theorem, that

$$\sup_{[0,R^2]\times B_{\varepsilon/4}(x)} |f(t,z)| \leq \mathbf{c}(\varepsilon,R).$$

For $t = R^2$ and $z = x$ we get

$$\left|\int_{\mathcal{C}} h_A\left(R^2-\tau,x,y\right) g(\tau,y)\, d\tau dy\right| \leq \mathbf{c}(\varepsilon,R)$$

for any $g \in C_0^\infty(\mathcal{C})$ such that $\|g\|_{L^2(\mathcal{C})} \leq 1$. This implies

$$\int_{\mathcal{C}} h_A(\tau,x,y)^2\, d\tau dy \leq \mathbf{c}(\varepsilon,R).$$

■

Conclusion of the proof of Theorem 4.11. To get Theorem 4.11 from the previous proposition, we have to interchange the variables x, y. To do this, we now use the fact that $h_A^*(t,x,y) = h_A(t,y,x)$ is the fundamental solution of

$$H_A^* = \partial_t - \sum_{i,j=1}^m a_{ij} X_i^* X_j^*$$

where we recall that

$$X_i^* = -X_i + a_i$$

for some smooth function a_i. Then, if we set

$$T_A^* g(t,x) = \int_{\mathcal{C}'} h_A^*(t-\tau,x,y)\, g(\tau,y)\, d\tau dy$$

we can see that Proposition 5.5 holds with X_i, T_A replaced by X_i^*, T_A^*. Namely:

i) holds because it relies on the estimates of Lemma 5.3, which are symmetric in the variables x, y.

ii) holds because $-X_i$ and X_i^*, H_A and H_A^* differ by lower order terms; therefore (5.4) holds with X_i, T_A replaced by X_i^*, T_A^*:

$$\begin{aligned}
&\left\|X_i^* X_j^* u\right\|_{L^p(B_0)} \leq \\
&\leq \left\|X_i X_j u\right\|_{L^p(B_0)} + \mathbf{c}\left[\|u\|_{L^p(B_0)} + \sum_k \|X_k u\|_{L^p(B_0)}\right] \leq \\
&\leq \mathbf{c}\left\{\|u\|_{L^p(2B_0)} + \sum_k \|X_k u\|_{L^p(2B_0)} + \|H_A u\|_{L^p(2B_0)}\right\} \leq \\
&\leq \mathbf{c}\left\{\|u\|_{L^p(2B_0)} + \sum_k \|X_k^* u\|_{L^p(2B_0)} + \|H_A^* u\|_{L^p(2B_0)} + \|(H_A - H_A^*)\, u\|_{L^p(2B_0)}\right\} \\
&\leq \mathbf{c}\left\{\|u\|_{L^p(2B_0)} + \sum_k \|X_k^* u\|_{L^p(2B_0)} + \|H_A^* u\|_{L^p(2B_0)}\right\}.
\end{aligned}$$

iii) follows from i) and ii) as said before. With this remark, we can repeat the proof of Proposition 5.8 and get

$$\sup_{x\in\mathbb{R}^n} \int_0^{R^2} d\tau \int_{\varepsilon < d(x,y) < R} h_A^*(\tau,x,y)^2\, dy \leq \mathbf{c}(\varepsilon, R)$$

which, by Remark 5.7, is Theorem 4.11. ■

6 Uniform global upper bounds for fundamental solutions

In this section and the following one, we will make extensive use of results and formalism mainly developed by Folland [26] and Rothschild-Stein [49]. We have collected in the following subsection a number of known definition and results, to fix notation and make more self-contained the exposition.

6.1 Preliminaries on homogeneous groups and the "lifting and approximation" technique

Homogeneous groups

Following a terminology introduced by Stein (see [55], p. 618-622) we say that a *homogeneous group* $\mathbb{G}$ is $\mathbb{R}^N$ endowed with a Lie group structure (such that the group operation is written $u \circ v$ and called *translation*; the inverse is denoted by u^{-1}, and the identity is the origin, 0), and a one parameter family of automorphisms (called *dilations* and denoted by $D(\lambda)$), which act as follows:

$$D(\lambda) : (u_1, \ldots, u_N) \mapsto (\lambda^{a_1} u_1, \ldots, \lambda^{a_N} u_N) \ \ \forall \lambda > 0$$

for suitable fixed integers $0 < a_1 \leq a_2 \leq \ldots \leq a_N$. The number

$$Q = \sum_{i=1}^{n} a_i$$

is called the *homogeneous dimension* of $\mathbb{G}$. If $\varphi : \left(\mathbb{R}^N, \circ\right) \to \left(\mathbb{R}^N, *\right)$ is any group isomorphism, we can also say that

$$v = \varphi(u)$$

is another choice of a system of coordinates in $\mathbb{G}$.

The following structures can be defined in a standard way in $\mathbb{G}$.

• *Homogeneous norm* $\|\cdot\|$: for any $u \in \mathbb{G}$, $u \neq 0$, set

$$\|u\| = \rho \quad \Leftrightarrow \quad \left| D(\frac{1}{\rho}) u \right| = 1,$$

where $|\cdot|$ denotes the Euclidean norm; also, let $\|0\| = 0$. Then:

$\|D(\lambda)u\| = \lambda \|u\|$ for every $u \in G$, $\lambda > 0$;

the set $\{u \in \mathbb{G}\colon \|u\| = 1\}$ coincides with the Euclidean unit sphere $\sum_N$;

the function $u \mapsto \|u\|$ is smooth outside the origin;

there exists $c(\mathbb{G}) \geq 1$ such that for every u, $v \in \mathbb{G}$

$$\|u \circ v\| \leq c\left(\|u\| + \|v\|\right) \qquad \text{and} \qquad \left\|u^{-1}\right\| \leq c\,\|u\|;$$

$$\frac{1}{c}\,|v| \leq \|v\| \leq c\,|v|^{1/s} \qquad \text{if} \quad \|v\| \leq 1.$$

• *Quasidistance* d:

$$d(u, v) = \left\|v^{-1} \circ u\right\|$$

for which the following hold:

$$d(u, v) \geq 0 \quad \text{and} \quad d(u, v) = 0 \text{ if and only if } u = v;$$

$$\frac{1}{c}\,d(v,u) \leq d(u, v) \leq c\,d(v,u);$$

$$d(u,v) \leq c\ \{d(u,z) + d(z,v)\}$$

for every $u,\ v,\ z \in \mathbb{R}^N$ and some positive constant $c(\mathbb{G}) \geq 1$.

If we denote by $B(u,r) \equiv B_r(u) \equiv \left\{v \in \mathbb{R}^N\colon\ d(u,v) < r\right\}$ the metric balls, then we see that $B(0,r) = D(r)B(0,1)$. Moreover, it can be proved that the Lebesgue measure in $\mathbb{R}^N$ is the *Haar measure* of $\mathbb{G}$. Therefore

$$|B(u,r)| = |B(0,1)|\ r^Q,$$

for every $u \in \mathbb{G}$ and $r > 0$, where Q is the homogeneous dimension of $\mathbb{G}$.

• The *convolution* of two functions in $\mathbb{G}$ is defined as

$$(f * g)(x) = \int_{\mathbb{R}^N} f(x \circ y^{-1})\, g(y)\, dy = \int_{\mathbb{R}^N} g(y^{-1} \circ x)\, f(y)\, dy,$$

for every couple of functions for which the above integrals make sense.

Let τ_u be the left translation operator acting on functions: $(\tau_u f)(v) = f(u \circ v)$. We say that a *differential operator* P *on* $\mathbb{G}$ *is left invariant* if $P(\tau_u f) = \tau_u(Pf)$ for every smooth function f. From the above definition of convolution we read that if P is any left invariant differential operator,

$$P(f * g) = f * Pg$$

(provided the integrals converge).

We say that *a differential operator* P *on* $\mathbb{G}$ *is homogeneous of degree* $\delta > 0$ if

$$P\ (f\ (D(\lambda)u)) = \lambda^\delta\ (Pf)(D(\lambda)u)$$

for every test function f, $\lambda > 0$, $u \in \mathbb{R}^N$. Also, we say that *a function* f *is homogeneous of degree* $\delta \in \mathbb{R}$ if

$$f\,(D(\lambda)u) = \lambda^\delta\, f\,(u) \ \text{ for every } \lambda > 0,\ u \in \mathbb{R}^N.$$

Clearly, if P is a differential operator homogeneous of degree δ_1 and f is a homogeneous function of degree δ_2, then Pf is homogeneous of degree $\delta_2 - \delta_1$. For example, $u_i \frac{\partial}{\partial u_j}$ is homogeneous of degree $|\alpha_j| - |\alpha_i|$.

A differential operator P on $\mathbb{G}$ is said to have *local degree less than or equal to* ℓ if, after taking the Taylor expansion at 0 of its coefficients, each term obtained is homogeneous of degree $\leq \ell$. For instance, if P is a vector field (that is a differential operator of degree one), saying that P has local degree $\leq \ell$ explicitly means that for any positive integer K, we can write

$$P = \sum_{\alpha \in A} \left(\sum_{k=0}^{K} c_{\alpha kj} p_{kj}\,(u)\, \frac{\partial}{\partial u_\alpha} + g\,(u)\, \frac{\partial}{\partial u_\alpha} \right)$$

where: $c_{\alpha kj}$ are suitable constants; $p_{kj}\,(u)$ are all the homogeneous monomials of degree k, such that each differential operator $p_{kj}\,(u)\,\frac{\partial}{\partial u_\alpha}$ is homogeneous of degree $|\alpha| - k \leq \ell$; $g\,(u) = o\left(\|u\|^K\right)$ for $u \to 0$.

Homogeneous Lie algebras

Let $\mathcal{G}(s,m)$ be the *free Lie algebra of step s on m generators*, that is the quotient of the free Lie algebra with m generators by the ideal generated by the commutators of length at least $s+1$, and let $N = \dim\mathcal{G}(s,m)$, as a vector space. One always has $N \geq n$.

If $e_1, \ldots, e_m$ are generators of $\mathcal{G}(m,s)$, for every multiindex $\alpha = (\alpha_1, \ldots, \alpha_d)$ with $1 \leq \alpha_i \leq m$, we define

$$e_\alpha = \left[e_{\alpha_d}, \left[e_{\alpha_{d-1}}, \ldots \left[e_{\alpha_2}, e_{\alpha_1}\right] \ldots\right]\right],$$

and $|\alpha| = d$. We call e_α a commutator of the e_i's of length d. Then there exists a set A of multiindices α so that $\{e_\alpha\}_{\alpha\in A}$ is a basis of $\mathcal{G}(m,s)$ as a vector space. This allows us to identify $\mathcal{G}(m,s)$ with $\mathbb{R}^N$. Note that $\operatorname{Card} A = N$ while $\max_{\alpha\in A} |\alpha| = s$, the step of the Lie algebra. The Campbell-Hausdorff series defines a multiplication in $\mathbb{R}^N$ (see e.g. [53]) that makes $\mathbb{R}^N$ the group $N(m,s)$, that is the simply connected Lie group associated to $\mathcal{G}(m,s)$. We can naturally define dilations in $N(m,s)$ by

$$D(\lambda)\left((u_\alpha)_{\alpha\in A}\right) = \left(\lambda^{|\alpha|} u_\alpha\right)_{\alpha\in A}.$$

These are automorphisms of $N(m,s)$, which is therefore a homogeneous group. We will call it $\mathbb{G}$, leaving the numbers m, s implicitly understood.

Once we have introduced this structure of homogeneous group in $\mathbb{R}^N$, we can give a concrete visualization to the abstract elements e_i of the Lie algebra, as follows. Denote by Y_j $(j = 1, \ldots, m)$ the *left-invariant vector fields* on $\mathbb{G}$ which agree with $\frac{\partial}{\partial u_j}$ at 0. Then Y_j is homogeneous of degree 1 and, for every multiindex α, Y_α is homogeneous of degree $|\alpha|$. The system of vector fields $\{Y_j\}_{j=1}^m$ satisfies Hörmander's condition of step s in $\mathbb{R}^N$, and their Lie algebra coincides with $\mathcal{G}(m,s)$. These Y_j's are uniquely determined by the choice of a system of coordinates u in $\mathbb{G}$.

It is sometimes useful to consider also the *right-invariant vector fields Z_j*, which agree with $\frac{\partial}{\partial u_j}$ (and therefore with Y_j) at 0; also these Z_i are homogeneous of degree one.

As to the *structure of the left (or right) invariant vector fields*, it can be proved that the systems $\{Y_i\}$ and $\{Z_i\}$ have the following "triangular form" with respect to Cartesian derivatives:

$$Y_i = \frac{\partial}{\partial u_i} + \sum_{k=i+1}^{N} q_i^k(u) \frac{\partial}{\partial u_k}$$

$$Z_i = \frac{\partial}{\partial u_i} + \sum_{k=i+1}^{N} \overline{q}_i^k(u) \frac{\partial}{\partial u_k}$$

where $q_i^k, \overline{q}_i^k$ are polynomials, homogeneous of degree $a_k - a_i$ (the a_i's are the dilation exponents). This implies that any Cartesian derivative ∂_{u_k} can be written as a linear combination of the Y_i's (and, analogously, of the Z_j's). In particular, any homogeneous differential operator can be rewritten as a linear combination of left invariant (or, similarly, right invariant) homogeneous vector fields, with polynomial coefficients. The above structure of the Y_i's also implies that the formal transpose Y_i^* of Y_i is just $-Y_i$ (as in a standard integration by parts).

Sublaplacians and homogeneous fundamental solutions

With the above notation, we see that the *sublaplacian*:

$$L = \sum_{i=1}^{m} Y_i^2$$

is a left invariant, homogeneous of degree 2, hypoelliptic operator on $\mathbb{G}$; also, note that $L^* = L$. At this point we recall a fundamental result by Folland [26]:

Theorem 6.1 (Existence of a homogeneous fundamental solution) *Let $\mathcal{L}$ be a left invariant differential operator homogeneous of degree two on a homogeneous group $\mathbb{G}$, such that $\mathcal{L}$ and $\mathcal{L}^*$ are both hypoelliptic. Moreover, assume $Q \geq 3$. Then there is a unique fundamental solution Γ such that:*
(*a*) $\Gamma \in C^\infty\left(\mathbb{R}^N \setminus \{0\}\right)$;
(*b*) Γ *is homogeneous of degree* $(2 - Q)$;
(*c*) *for every distribution* τ,

$$\mathcal{L}(\tau * \Gamma) = (\mathcal{L}\tau) * \Gamma = \tau.$$

The previous representation formula also implies that, for any test function u:

$$Y_i Y_j u = PV\left((\mathcal{L}u) * Y_i Y_j \Gamma\right) + c_{ij}\mathcal{L}u$$

where c_{ij} are constants.

General Hörmander's vector fields: "lifting and approximation"

Let us consider now a generic system of Hörmander's vector fields $X_1, X_2, ..., X_m$. Rothschild and Stein [49] have found a way to exploit Folland's theory for homogeneous groups to study the more general operators

$$L = \sum_{i=1}^{q} X_i^2 \text{ or } H = \partial_t - L.$$

This is accomplished by the famous "lifting and approximation" result contained in [49]:

Theorem 6.2 *Let $X_1, \ldots, X_m$ be C^∞ real vector fields on a domain $\Omega \subset \mathbb{R}^n$ satisfying Hörmander's condition of step s at some point $x_0 \in \Omega$,*

$$X_i = \sum_{j=1}^{n} b_{ij}(x)\,\partial_{x_j}.$$

Let $\mathcal{G}(s,m)$ be the free Lie algebra of step s on q generators, $\mathbb{G}$ the corresponding homogeneous group on $\mathbb{R}^N$ ($N = \dim \mathcal{G}(s,m) > n$). Then in terms of new variables, $h_1,\ldots,h_{N-n}$, there exist smooth functions $c_{ij}(x,h)$ ($1 \leq i \leq m$, $1 \leq j \leq N-n$) defined in a neighborhood $\widetilde{U}$ of $\xi_0 = (x_0, 0) \in \Omega \times \mathbb{R}^{N-n} = \widetilde{\Omega}$ such that the vector fields $\widetilde{X}_i$ given by

$$\widetilde{X}_i = X_i + \sum_{j=1}^{N-n} c_{ij}(x, h_1, h_2, ..., h_{j-1})\, \partial_{h_j} \quad i = 1,\ldots,m$$

satisfy Hörmander's condition of step s. Moreover, if we choose a system $\{\widetilde{X}_\alpha(\xi)\}_{\alpha\in A}$ such that $\{\widetilde{X}_\alpha(\xi_0)\}_{\alpha\in A}$ be a basis for $\mathbb{R}^N$, then there exists a choice of coordinates u_α in $\mathbb{G}$ such that, for $\xi,\eta \in \widetilde{U}$, the map

$$\Theta_\eta(\xi) = (u_\alpha)_{\alpha\in A}$$

with

$$\xi = \exp\left(\sum_{\alpha\in A} u_\alpha\, \widetilde{X}_\alpha\right)\eta$$

is well-defined, and satisfies the following properties. There exist open neighborhoods U of 0 and V, W of ξ_0 in $\mathbb{R}^N$, with $W \Subset V$ such that:
a) $\Theta_\eta \mid V$ is a diffeomorphism onto the image, for every $\xi \in V$;
b) $\Theta_\eta(V) \supseteq U$ for every $\eta \in W$;
c) $\Theta\colon V \times V \to \mathbb{R}^N$, defined by $\Theta(\xi,\eta) = \Theta_\eta(\xi)$ is $C^\infty(V \times V)$;
d) In the coordinates given by Θ_η, we can write $\widetilde{X}_i = Y_i + R_i^\eta$ on U, where Y_i are the homogeneous left invariant vector fields on $\mathbb{G}$, coinciding with ∂_{u_i} at the origin, and R_i^η are vector fields of local degree ≤ 0 depending smoothly on $\eta \in W$ (the superscript η does not denote the variable of differentiation but the dependence on the point η). Explicitly, this means that for every $f \in C_0^\infty(\mathbb{G})$:

$$\widetilde{X}_i\,(f(\Theta_\eta(\cdot)))(\xi) = (Y_i f + R_i^\eta f)(\Theta_\eta(\xi)).$$

e) More generally, for every $\alpha \in A$ we can write

$$\widetilde{X}_\alpha = Y_\alpha + R_\alpha^\eta$$

where R_α^η is a vector field of local degree $\leq |\alpha| - 1$ depending smoothly on η.

Roughly speaking, the above theorem says that the original system of vector fields $\{X_i\}_{i=1}^m$ defined in $\mathbb{R}^n$ can be locally lifted to another system $\left\{\widetilde{X}_i\right\}_{i=1}^m$ defined in $\widetilde{U} \subset \mathbb{R}^N$ ($N > n$), such that the $\widetilde{X}_i$ can be locally approximated by the homogeneous left invariant vector fields Y_i. The remainder in this approximation process is expressed by the vector fields R_i^η which have the following good property: when they act on a homogeneous function, typically of negative degree (that is, with some singularity), the singularity does not become worse. The vector fields Y_i, R_i^η must be thought as acting on the group $\mathbb{G}$; the vector fields $\widetilde{X}_i$ as acting on the "manifold" $\mathbb{R}^N$, the change of variables between the two environments being realized by the map Θ_η. Here below we add some other miscellaneous facts, related to the above concepts, which will be needed in the following.

• Under the change of variables $\xi \mapsto u$ given by $u = \Theta_\eta(\xi)$, the measure element becomes:

$$d\xi = c(\eta) \cdot (1 + O(\|u\|))\, du,$$

where $c(\eta)$ is a smooth function, bounded and bounded away from zero in V. For the change of coordinates $\eta \mapsto u$ given by $u = \Theta_\eta(\xi)$, an analogous relation holds:

$$d\eta = c(\xi) \cdot (1 + O(\|u\|))\, du.$$

• If, for $\xi, \eta \in V$, we define

$$\rho(\xi, \eta) = \|\Theta(\xi, \eta)\|$$

where $\|\cdot\|$ is the homogeneous norm defined above, then ρ is a quasidistance, locally equivalent to the CC-distance $\widetilde{d}$ induced by the vector fields $\left\{\widetilde{X}_i\right\}$.

• Although there is no easy relation between the CC-distance d induced in $\mathbb{R}^n$ by the X_i's and the CC-distance $\widetilde{d}$ induced in $\mathbb{R}^N$ by the $\widetilde{X}_i$'s, a more transparent relation holds between the volumes of corresponding balls. This fact is described by the following result by Sanchez-Calle:

Lemma 6.3 *(See [53], Lemma 5). Let $B, \widetilde{B}$ denote metric balls with respect to d (in $\mathbb{R}^n$) and $\widetilde{d}$ (in $\mathbb{R}^N$), respectively. There exist $r_0 > 0$ and $\delta \in (0,1)$ such that for any $(x,h) \in \widetilde{\Omega}, r < r_0, y \in B(x, \delta r)$ one has*

$$r^Q \simeq \left|\widetilde{B}((x,h), r)\right| \simeq |B(x,r)| \cdot \left|\left\{h' \in \mathbb{R}^{N-n} : (y,h') \in \widetilde{B}((x,h), r)\right\}\right|$$

where $|\cdot|$ denotes Lebesgue measure in the appropriate $\mathbb{R}^m$, and the equivalence $a \simeq b$ means $c_1 a \leq b \leq c_2 a$ for positive constants c_1, c_2 independent of r, x, y, h.

• Sometimes we will integrate a function of N variables with respect to the "lifted variables" only; the resulting function is defined in $\mathbb{R}^n$, and its derivatives with respect to the original vector fields X_i can be expressed by the following useful identity:

Lemma 6.4 *Let $f \in C_0^\infty(U)$, then*

$$X_i^x \int_{\mathbb{R}^{N-n}} f(\xi)\, dh = \int_{\mathbb{R}^{N-n}} \widetilde{X}_i^\xi f(\xi)\, dh$$

(writing $\xi = (x,h)$).

Proof. Recall that:

$$\widetilde{X}_i = X_i + \sum_{j=1}^{N-n} c_{ij}(x, h_1, h_2, ..., h_{j-1})\, \partial_{h_j}$$

Then

$$\int_{\mathbb{R}^{N-n}} \widetilde{X}_i^{\xi} f(\xi)\, dh$$

$$= \int_{\mathbb{R}^{N-n}} X_i^x f(\xi) + \sum_{j=1}^{N-n} \partial_{h_j} \left(c_{ij} \left(x, h_1, h_2, ..., h_{j-1} \right) f(\xi) \right) dh =$$

$$= X_i^x \int_{\mathbb{R}^{N-n}} f(\xi)\, dh + \sum_{j=1}^{N-n} \int_{\mathbb{R}^{N-n-1}} dh_1 ... dh_{j-1} dh_{j+1} ... dh_{N-n} \int_{\mathbb{R}} \partial_{h_j} \left(g(\xi) \right) dh_j =$$

$$= X_i^x \int_{\mathbb{R}^{N-n}} f(\xi)\, dh$$

because $\int_{\mathbb{R}} \partial_{h_j} \left(g(\xi) \right) dh_j = 0$. ■

Parabolic context

In the following, we will apply the procedures described above to the space variables, while time will play the role of a parameter. It is worthwhile to note explicitly the following fact:

Remark 6.5 *If $\mathbb{G} = \left(\mathbb{R}^N, \circ, D(\lambda) \right)$ is a homogeneous group, we can naturally define its "parabolic version", setting, in $\mathbb{R}^{N+1}$:*

$$(t,u) \circ_P (s,v) = (t+s, u \circ v)\, ; \;\; D_P(\lambda)(t,u) = \left(\lambda^2 t, D(\lambda) u \right).$$

The "parabolic homogeneous group" $\mathbb{G}_P = \left(\mathbb{R}^{N+1}, \circ_P, D_P(\lambda) \right)$ has homogeneous dimension $Q+2$.

With respect to this structure, the hypoelliptic differential operator in $\mathbb{R}^{N+1}$

$$\mathcal{H}_A = \partial_t - \sum_{i,j=1}^{m} a_{ij} Y_i Y_j$$

is translation invariant and homogeneous of degree 2. Then, by the general result of Folland recalled above, $\mathcal{H}_A$ has a fundamental solution $g_A(t,u)$, homogeneous of degree $-Q$. Moreover, for this operator, uniform Gaussian bounds have been proved in [3, Theorem 2.5 p.1160]; we recall here these estimates, which will be crucial in the following (actually, we will use only (6.1), (6.2)):

Theorem 6.6 *For every nonnegative integers p, q, for every $u \in \mathbb{R}^n$, $t > 0$ and $A, B \in \mathcal{B}_\lambda$, the following uniform Gaussian bounds hold*

$$\mathbf{c}^{-1} t^{-Q/2} e^{-\|u\|^2/\mathbf{c}^{-1}t} \le g_A(t,u) \le \mathbf{c} t^{-Q/2} e^{-\|u\|^2/\mathbf{c}t} \tag{6.1}$$

$$\left| Y_{i_1} \cdots Y_{i_p} (\partial_t)^q \, g_A(t,u) \right| \le \mathbf{c}(p,q) t^{-(Q+p+2q)/2} e^{-\|u\|^2/\mathbf{c}t} \tag{6.2}$$

$$\left| Y_{i_1} \cdots Y_{i_p} (\partial_t)^q \left(g_A - g_B \right)(t,u) \right| \le \mathbf{c}(p,q) \; \|A - B\|^{1/s} \; t^{-(Q+p+2q)/2} e^{-\|u\|^2/\mathbf{c}t} \tag{6.3}$$

where s denotes the step of the Lie algebra of $\mathbb{G}$.

6.2 Upper bounds on fundamental solutions

As we mentioned in the Overview of Part I, the procedure to deduce the desired uniform upper bound on h_A from the analogous result proved in [3] for homogeneous groups, is rather involved. So, to orient the reader, we want to sketch here the main steps of this proof, adapted from [30]. Applying Rothschild-Stein technique, we "lift", locally, our operator H_A to another operator $\widetilde{H}_A$, living in a higher dimensional space, which is locally approximated by a third operator $\mathcal{H}_A$, left invariant and 2-homogeneous on a homogeneous group, whose fundamental solution $g_A(t,u)$ satisfies the uniform Gaussian bounds stated in Theorem 6.6. Starting with this g_A, we form the kernel (see (6.9))

$$\widetilde{K}_0(t,\xi,\eta) = \chi(t)\, a(\xi)\, b(\eta)\, g_A(t,\Theta(\xi,\eta)),$$

(where $\chi(t), a(\xi), b(\eta)$ are suitable cutoff functions), which, morally speaking, should be a first approximation of the fundamental solution of $\widetilde{H}_A$. Namely (see (6.10)),

$$\widetilde{H}^{\xi}\widetilde{K}_0(t,\xi,\eta) = a(\xi)\,\delta_{(0,\eta)}(t,\xi) + \widetilde{E}_0(t,\xi,\eta)$$

where the "error term" $\widetilde{E}_0$ satisfies the Gaussian-type estimate that we would expect from the *first order* derivative of $\widetilde{K}_0$. A suitable inductive procedure allows to improve this gain, building a sequence of kernels $\widetilde{K}_i$ (see (6.12)) such that:

1) $\widetilde{K}_i$ satisfies the Gaussian bound we expect for the fundamental solution of $\widetilde{H}_A$;

2)

$$\widetilde{H}^{\xi}\widetilde{K}_i(t,\xi,\eta) = a(\xi)\,\delta_{(0,\eta)}(t,\xi) + \widetilde{E}_i(t,\xi,\eta);$$

3) the "error term" $\widetilde{E}_i$ is bounded by a function of the kind $t^{(i-1)/2}g_A$ (we can think that "it is small in sense of Gaussian bounds").

To keep track of the uniformity (with respect to A) of these bounds on $\widetilde{K}_i, \widetilde{E}_i$, the preliminary study carried out from Definition 6.7 through Lemma 6.11 is crucial; this study makes extensive use of (6.1), (6.2).

Then, one could hope to prove that this sequence $\widetilde{K}_i$ approximates the fundamental solution $\widetilde{h}_A$ of $\widetilde{H}_A$ (thus giving the Gaussian bounds for $\widetilde{h}_A$) and this, in turn, could imply analogous Gaussian bounds for h_A, the fundamental solution of H_A. However, this idea does not work (one of the reasons being that the fundamental solution of H_A *is not unique*). Instead, a more indirect link is established from $\widetilde{K}_i$ to h_A: first (see (6.13)), one defines a new sequence of kernels K_i, living in the "original" space $\mathbb{R}^n$ where H_A is defined:

$$K_j(t,x,y) = \int_{\mathbb{R}^{N-n}} \int_{\mathbb{R}^{N-n}} \widetilde{K}_j(t,\xi,\eta)\, dh\, dh'.$$

Then, one proves that $H_A(K_j - h_A)$ is smaller and smaller, for increasing j, in the sense of Gaussian bounds (see Lemma 6.13). Finally, the construction of suitable barriers (see Lemma 6.15) enables to show (see the proof of Theorem 6.16) that in the region $d(x,y) < \varepsilon, x \in B(0,R)$ (for ε small enough), the desired Gaussian bound holds for $(K_j - h_A)$, if j is large enough, and therefore holds for h_A. Since, on the other side, in the region $d(x,y) \geq \varepsilon$ we have already proved the upper bound in Sections 4 and 5, we are done. We now come to the precise construction.

Here we keep all notation introduced in the previous subsection; in particular, we will denote by $B, \widetilde{B}$ the metric balls with respect to d (in $\mathbb{R}^n$) and $\widetilde{d}$ (in $\mathbb{R}^N$), respectively.

Fix a large ball $B(0,R) \subset \mathbb{R}^n$, with R to be chosen later, and cover it with a finite collection of balls $B(x_\ell, r)$, with r small enough so that in each ball we can perform the procedure of lifting and approximation. Let $\{\alpha_\ell(x)\}$ be a partition of unity of $B(0,R)$ induced by the family $\{B(x_\ell, r)\}$. Set $\xi_\ell = (x_\ell, 0)$, and define

$$a_\ell(\xi) = \alpha_\ell(x)\,\varphi(h) \ \text{ for } \xi = (x,h) \tag{6.4}$$

where φ is a cutoff function in the h variables, fixed once and for all, with

$$\int_{\mathbb{R}^{N-n}} \varphi(h)\,dh = 1. \tag{6.5}$$

We now fix one of the points ξ_ℓ, and in a suitable neighborhood $\widetilde{U}_\ell$ of the kind $B(x_\ell, r) \times (-\delta,\delta)^{N-n} \subset \mathbb{R}^N$ we perform the following construction. Let $\widetilde{X}_1^\ell, \widetilde{X}_2^\ell, ..., \widetilde{X}_m^\ell$ be the lifted vector fields defined in $\widetilde{U}^\ell$; with the notation of Theorem 6.2,

$$\widetilde{X}_i^\ell\left(f\left(\Theta^\ell(\xi,\eta)\right)\right) = \left(Y_i f + R_i^{\ell,\eta} f\right)\left(\Theta^\ell(\xi,\eta)\right)$$

where the map $\Theta_\eta^\ell(\xi) = \Theta^\ell(\xi,\eta)$ is defined for ξ and η in $\widetilde{U}^\ell$. In the above formula, the derivative $\widetilde{X}_i^\ell$ is taken with respect to ξ; the derivatives $Y_i, R_i^{\ell,\eta}$ are taken with respect to the variable u in the group $\mathbb{G} = \left(\mathbb{R}^N, \circ\right)$, and the coefficients of the vector fields $R_i^{\ell,\eta}$ depend on the fixed point η (and on u). Let

$$\widetilde{H}_A^\ell = \partial_t - \sum_{i,j=1}^m a_{ij} \widetilde{X}_i^\ell \widetilde{X}_j^\ell, \tag{6.6}$$

and

$$\mathcal{H}_A = \partial_t - \sum_{i,j=1}^m a_{ij} Y_i Y_j, \tag{6.7}$$

then the above approximation relation becomes:

$$\widetilde{H}_A^\ell f\left(t, \left(\Theta^\ell(\cdot,\eta)\right)\right)(\xi) = \left(\mathcal{H}_A f\right)\left(t, \Theta^\ell(\xi,\eta)\right) + \left(E_A^{\ell,\eta} f\right)\left(t, \Theta^\ell(\xi,\eta)\right) \tag{6.8}$$

with

$$\left(E_A^{\ell,\eta} f\right) = \sum_{i,j=1}^{m} a_{ij} \left(E_{ij}^{\ell,\eta} f\right) =$$
$$= \sum_{i,j=1}^{m} a_{ij} \left(Y_i R_j^{\ell,\eta} f + R_i^{\ell,\eta} Y_j f + R_i^{\ell,\eta} R_j^{\ell,\eta} f\right)$$

(observe that the differential operator $E_A^{\ell,\eta}$ does not depend on t).

The reason why this approximation procedure is helpful is that the operator $\mathcal{H}_A$ in (6.7) is homogeneous of degree 2 and translation invariant on $\mathbb{G}_P$ (see Remark 6.5), and for its fundamental solution Gaussian bounds are known (see Theorem 6.6).

Definition 6.7 *Denote with $g_A(t,u)$ the homogeneous fundamental solution of $\mathcal{H}_A$. For any integer μ, we define the class F_μ^A as follows: $k \in F_\mu^A$ if:*

i) $k \in C^\infty(\mathbb{R} \times \mathbb{G} \setminus \{0,0\})$*;*

ii) $k(t,u) = 0$ *for* $t < 0$*;*

iii) for every integer p and every multi-index α

$$|\partial_t^p Y^\alpha k(t,u)| \leqslant \mathbf{c}(p,\alpha,\mu)\, t^{(\mu-2-2p-|\alpha|)/2} g_A(\mathbf{C}t,u) \text{ for any } t > 0, u \in \mathbb{G}$$

where the constants $\mathbf{c}(p,\alpha,\mu)$ depend on the matrix $\{a_{ij}\}$ only through the ellipticity constant λ and $\mathbf{C}$ depends on p,α and the kernel k, but not on A.

Note that, by (6.1), (6.2), $g_A \in F_2^A$. Also, note that

$$\mu_1 < \mu_2 \Rightarrow F_{\mu_2}^A \subset F_{\mu_1}^A.$$

In the next lemma we summarize some of the properties of the classes F_μ^A.

Lemma 6.8

1) If $k \in F_\mu^A$ with $\mu > Q+2$, and we define $k(0,0) = 0$, then $k \in C^0(\mathbb{R} \times \mathbb{G})$.

2) If $k \in F_\mu^A$ and P is a homogeneous differential operator of degree d in the u variable, then $\partial_t^p P(k) \in F_{\mu-d-2p}^A$. In particular one can take in the definition of F_μ^A the right invariant vector fields $Z_1, Z_2, \ldots, Z_m$ instead of $Y_1, \ldots Y_m$. Also if $k \in F_\mu^A$ and P is a differential operator of local degree $\leqslant d$, then $\partial_t^p P(k) \in F_{\mu-d-2p}^A$.

3) Given $\gamma > 0$ we can find μ so that $F_\mu^A \subseteq C^\gamma(\mathbb{R} \times \mathbb{G})$.

4) If $k \in F_{\mu_1}^A$, $h \in F_{\mu_2}^A$, $\mu_1, \mu_2 > 0$ and

$$(k * h)(t,u) = \int\int k\left(t-s, u \circ v^{-1}\right) h(s,v)\, dvds$$

then $(k * h) \in F_{\mu_1+\mu_2}$.

Proof. 1) is a simple consequence of Definition 6.7 and (6.1).

To show 2) observe that any homogeneous differential operator of degree d can be written as $\sum \eta_\alpha(u) Y^\alpha$ where η_α is a homogeneous function of degree $|\alpha| - d$. We have

$$\begin{aligned}
\left|\partial_t^p P(k)\right| &= \left|\sum_\alpha \eta_\alpha(u)\, \partial_t^p Y^\alpha k\right| \leqslant \\
&\leqslant \sum_\alpha \|u\|^{|\alpha|-d} \mathbf{c}(p,\alpha,\mu)\ t^{(\mu-2-2p-|\alpha|)/2} g_A(\mathbf{C}t,u) \\
&\leqslant \sum_\alpha \left(\frac{\|u\|}{\sqrt{t}}\right)^{|\alpha|-d} \mathbf{c}(p,\alpha,\mu)\ t^{(\mu-2-2p-d)/2} g_A(\mathbf{c}t,u). \\
&\leqslant \mathbf{c}(p,\alpha,\mu)\ t^{(\mu-2-2p-d)/2} g_A(\mathbf{C}'t,u)
\end{aligned}$$

where the last inequality is a consequence of (6.1). Indeed,

$$\begin{aligned}
\left(\frac{\|u\|}{\sqrt{t}}\right)^{|\alpha|-d} g_A(\mathbf{C}t,u) &\leqslant \mathbf{c}\left(\frac{\|u\|}{\sqrt{t}}\right)^{|\alpha|-d} t^{-\frac{Q}{2}} \exp\left(-\|u\|^2/\mathbf{c}t\right) \\
&\leqslant \mathbf{c} g_A(\mathbf{C}'t,u).
\end{aligned}$$

This proves the first assertion in (2). Moreover, since the right invariant vector fields $Z_1,\dots,Z_m$ are homogeneous of degree 1, and since it is also true that any homogeneous differential operator of degree d can be written as $\sum \eta'_\alpha(u) Z^\alpha$, where η'_α is a homogeneous function of degree $|\alpha|-d$, the above reasoning also implies that in the definition of the class F_μ^A one can use these Z_i's instead of the Y_i's. The case when P is a differential operator of local degree $\leqslant d$ can be proved in the same way observing that if $P = \sum f_\alpha(u) D^\alpha$ then the Taylor expansion of $f_\alpha(u) D^\alpha \sim \sum c_{\alpha\beta} u^\beta D^\alpha$ is the formal sum of homogeneous differential operator of degree $\leqslant d$.

3) The fact that $F_\mu^A \subseteq C^\gamma(\mathbb{R}\times\mathbb{G})$ when μ is large enough is a simple consequence of (6.1).

4) Let $\phi \in C_0^\infty(\mathbb{R}\times\mathbb{G})$ such that $\phi(t,u) = 1$ if $\|(t,u)\| \leqslant 1$ and $\phi(t,u) = 0$ if $\|(t,u)\| > 2$. Define $\phi_\delta(t,u) = \phi(D_P(\delta)(t,u))$ and write

$$\begin{aligned}
(k*h)(t,u) &= \int\int k\left(t-s, u\circ v^{-1}\right) h(s,v)\, dvds \\
&= \int\int \phi_\delta(s,v)\, k(s,v)\, h\left(t-s, v^{-1}\circ u\right) dvds \\
&+ \int\int \phi_\delta(s,v)\, h(s,v)\, k\left(t-s, u\circ v^{-1}\right) dvds \\
&+ \int\int \left[1-\phi_\delta\left(t-s,u\circ v^{-1}\right) - \phi_\delta(s,v)\right] k\left(t-s,u\circ v^{-1}\right) h(s,v)\, dvds \\
&= I_1 + I_2 + I_3.
\end{aligned}$$

Let C such that $\|(t+s, u\circ v)\| \leqslant C(\|(t,u)\| + \|(s,v)\|)$ and observe that for

$(s,v)\in \operatorname{supp}\phi_\delta$ one has

$$\left\|\left(t-s,v^{-1}\circ u\right)\right\| \geqslant \frac{\|(t,u)\|}{C}-\|(s,v)\| \geqslant \frac{\|(t,u)\|}{C}-\frac{2}{\delta}$$
$$\geqslant \frac{\|(t,u)\|}{2C}$$

whenever $\delta \geqslant \frac{4C}{\|(t,u)\|}$. This shows that I_1, I_2 and I_3 are in $C^\infty\left(\mathbb{R}\times\mathbb{G}\setminus\{0,0\}\right)$. The fact that $(k*h)(t,u)=0$ when $t<0$ follows from the analogous property of k and h. Let us write

$$\begin{aligned}
&(k*h)(t,u)=\\
&=\int_0^t\int_{\mathbb{G}} k\left(t-s,u\circ v^{-1}\right)h(s,v)\,dvds\\
&=\int_0^{t/2}\int_{\mathbb{G}} k\left(t-s,u\circ v^{-1}\right)h(s,v)\,dvds+\int_{t/2}^{t}\int_{\mathbb{G}} k\left(t-s,u\circ v^{-1}\right)h(s,v)\,dvds\\
&=\int_0^{t/2}\int_{\mathbb{G}} k\left(t-s,u\circ v^{-1}\right)h(s,v)\,dvds+\int_0^{t/2}\int_{\mathbb{G}} k(s,w)\,h\left(t-s,w^{-1}\circ u\right)dvds\\
&I+II.
\end{aligned}$$

Let Z^a denote any right invariant differential operator homogeneous of degree α. Then

$$\begin{aligned}
\partial_t^p Z^\alpha(I) &= \partial_t^p\int_0^{t/2}\int_{\mathbb{G}} Z^\alpha k\left(t-s,u\circ v^{-1}\right)h(s,v)\,dvds\\
&=\partial_t^p t\int_0^{\frac{1}{2}}\int_{\mathbb{G}} Z^\alpha k\left(t-ts,u\circ v^{-1}\right)h(ts,v)\,dvds\\
&=p\int_0^{\frac{1}{2}}\int_{\mathbb{G}}(1-s)^{p-1}\,\partial_t^{p-1}Z^\alpha k\left(t(1-s),u\circ v^{-1}\right)h(ts,v)\,dvds\\
&+t\int_0^{\frac{1}{2}}\int_{\mathbb{G}}(1-s)^{p}\,\partial_t^{p}Z^\alpha k\left(t(1-s),u\circ v^{-1}\right)h(ts,v)\,dvds
\end{aligned}$$

where we used the fact that $\frac{d^n}{dt^n}(tg(t))=ng^{(n-1)}(t)+tg^{(n)}(t)$.

For the first term we have

$$\left|\int_0^{\frac12}\int_{\mathbb{G}}(1-s)^{p-1}\partial_t^{p-1}Z^\alpha k\left(t(1-s),u\circ v^{-1}\right)h(ts,v)\,dvds\right|$$
$$\leqslant \mathbf{c}\int_0^{\frac12}(1-s)^{p-1}(t(1-s))^{(\mu_1-2p-|\alpha|)/2}(ts)^{(\mu_2-2)/2}\cdot$$
$$\cdot\int_{\mathbb{G}}g_A\left(\mathbf{c}t(1-s),u\circ v^{-1}\right)g_A(\mathbf{c}ts,v)\,dvds$$
$$\leqslant \mathbf{c}g_A(\mathbf{c}t,u)\int_0^{\frac12}(1-s)^{p-1}(t(1-s))^{(\mu_1-2p-|\alpha|)/2}(ts)^{(\mu_2-2)/2}ds$$
$$\leqslant \mathbf{c}g_A(\mathbf{c}t,u)\,t^{(\mu_1-2p-|\alpha|)/2}t^{(\mu_2-2)/2}\int_0^{\frac12}s^{(\mu_2-2)/2}ds$$
$$\leqslant \mathbf{c}g_A(\mathbf{c}t,u)\,t^{(\mu_1+\mu_2-2-2p-|\alpha|)/2}.$$

The second term can be controlled in the same way. In a similar way one can estimate $\partial_t Y^\alpha(II)$ where Y^α is a homogeneous left invariant differential operator of degree α. ∎

Definition 6.9 *We say that K is a kernel of type μ in $\widetilde{U}^\ell$ if*

$$K(t,\xi,\eta)=\sum_{j=1}^r\chi_j(t)\,a_j(\xi)\,b_j(\eta)\,k_j\left(t,\Theta^\ell(\xi,\eta)\right)$$

where $\chi_j\in C_0^\infty(\mathbb{R})$, $a_j,b_j\in C_0^\infty\left(\mathbb{R}^N\right)$ (with supports small enough so that $\Theta^\ell(\xi,\eta)$ is well defined for $\xi\in\operatorname{supp}a_j$ and $\eta\in\operatorname{supp}b_j$), $\chi_j(t)=1$ for small t, $b_j=1$ on $\operatorname{supp}a_j$ and $k_j\in F_\mu^A$. We also assume that

$$\sup|D^\alpha a_j|+\sup|D^\alpha b_j|\leqslant\mathbf{c}(\alpha)$$
$$\sup\left|\frac{d^k\chi}{dt^k}(t)\right|\leqslant\mathbf{c}(k)$$

for every integer k and multiindex α.

In particular, note that if K is of type μ in $\widetilde{U}^\ell$ then

$$|K(t,\xi,\eta)|\leq t^{(\mu-2)/2}g_A\left(\mathbf{c}t,\Theta^\ell(\xi,\eta)\right)\sum_{j=1}^r\chi_j(t)\,a_j(\xi)\,b_j(\eta)$$

Remark 6.10 *We are now going to make some computations in a single fixed neighborhood $\widetilde{U}^\ell$. In order to make more readable our expressions, the index ℓ will be suppressed. We will use it again later, to explain how global objects can be built gluing together local pieces.*

Lemma 6.11 *Let $K(t,\xi,\eta) = \chi(t)\, a(\xi)\, b(\eta)\, k(t,\Theta(\xi,\eta))$ be a kernel of type μ. Then*

$$\widetilde{H}^{\xi} K(t,\xi,\eta) = \chi(t)\, a(\xi)\, b(\eta)\, \mathcal{H}k(t,\Theta(\xi,\eta)) + E(t,\xi,\eta)$$

where E is a kernel of type $\mu-1$. Moreover, $\widetilde{H}^{\xi}K$ is a kernel of type $\mu-2$.

Proof. By (6.8) we have

$$\begin{aligned}
&\widetilde{H}^{\xi} K(t,\xi,\eta) \\
&= \chi(t)\, a(\xi)\, b(\eta)\, \widetilde{H}^{\xi}[k(t,\Theta(\xi,\eta))] + \chi'(t)\, a(\xi)\, b(\eta)\, k(t,\Theta(\xi,\eta)) \\
&- \chi(t)\, b(\eta) \sum_{i,j=1}^{m} a_{ij} \left\{ \left(\widetilde{X}_i \widetilde{X}_j a(\xi)\right) k(t,\Theta(\xi,\eta)) + 2\widetilde{X}_i a(\xi)\, \widetilde{X}_j [k(t,\Theta(\xi,\eta))] \right\} \\
&= \chi(t)\, a(\xi)\, b(\eta) [\mathcal{H}k(t,\Theta(\xi,\eta)) + (E_\eta k)(\Theta(\xi,\eta),t)] + \\
&+ \chi'(t)\, a(\xi)\, b(\eta)\, k(t,\Theta(\xi,\eta)) - \chi(t)\, b(\eta) \sum_{i,j=1}^{m} a_{ij} \left\{ \left(\widetilde{X}_i \widetilde{X}_j a(\xi)\right) k(t,\Theta(\xi,\eta)) + \right. \\
&\left. +2\widetilde{X}_i a(\xi) \left[(Y_j k)(t,\Theta(\xi,\eta)) + \left(R_j^{\eta} k\right)(t,\Theta(\xi,\eta))\right] \right\} \\
&= \chi(t)\, a(\xi)\, b(\eta)\, \mathcal{H}k(t,\Theta(\xi,\eta)) + E(t,\xi,\eta)
\end{aligned}$$

where $E(t,\xi,\eta)$ is the sum of the following terms:

$\chi'(t)\, a(\xi)\, b(\eta)\, k(t,\Theta(\xi,\eta))$, kernel of type μ;

$- \chi(t)\, b(\eta) \sum_{i,j=1}^{m} a_{ij} \left(\widetilde{X}_i \widetilde{X}_j a(\xi)\right) k(t,\Theta(\xi,\eta))$, kernel of type μ;

$- \chi(t)\, b(\eta) \sum_{i,j=1}^{m} a_{ij} 2\widetilde{X}_i a(\xi)\, (Y_j k)(t,\Theta(\xi,\eta))$, kernel of type $\mu-1$;

$\chi(t)\, a(\xi)\, b(\eta)\, (E_\eta k)(t,\Theta(\xi,\eta))$, kernel of type $\mu-1$;

$- \chi(t)\, b(\eta) \sum_{i,j=1}^{m} a_{ij} 2\widetilde{X}_i a(\xi) \left(R_j^{\eta} k\right)(t,\Theta(\xi,\eta))$, kernel of type $\mu-1$.

This follows from Lemma 6.8, point 2, since Y_j is a homogeneous differential operator of degree 1, while E_η and R_j^η are differential operators of local degree ≤ 1. Therefore E is a kernel of type $\mu-1$. Finally, since $\mathcal{H}$ is a homogeneous differential operator of degree 2,

$$\chi(t)\, a(\xi)\, b(\eta)\, \mathcal{H}k(t,\Theta(\xi,\eta)) \text{ is a kernel of type } \mu-2$$

and therefore $\widetilde{H}^{\xi} K(t,\xi,\eta)$ is a kernel of type $\mu-2$. ■

We now set:

$$\widetilde{K}_0(t,\xi,\eta) = \chi(t)\, a(\xi)\, b(\eta)\, g_A(t,\Theta(\xi,\eta)), \tag{6.9}$$

where $\chi(t), a(\xi), b(\eta)$ are cutoff functions as in the Definition 6.9, and $a(\xi)$ is the function $a_\ell(\xi)$ defined in (6.4)-(6.5). Note that $\widetilde{K}_0$ is a kernel of type 2 in $\widetilde{U}^\ell$. By Lemma 6.11,

$$\begin{aligned}\widetilde{H}^\xi \widetilde{K}_0(t,\xi,\eta) &= \chi(t)\, a(\xi)\, b(\eta)\, \mathcal{H}_A g_A(t,\Theta(\xi,\eta)) + \widetilde{E}_0(t,\xi,\eta) = \\ &= a(\xi)\, \delta_{(0,\eta)}(t,\xi) + \widetilde{E}_0(t,\xi,\eta)\end{aligned} \tag{6.10}$$

where $\widetilde{E}_0$ is a kernel of type 1, which can be written as:

$$\widetilde{E}_0(t,\xi,\eta) = \sum_{j=1}^{r} \chi_j(t)\, a_j(\xi)\, b_j(\eta)\, k_j(t,\Theta(\xi,\eta)) \tag{6.11}$$

with $k_j \in F_1^A$. Next, we define:

$$\widetilde{K}_1(t,\xi,\eta) = \widetilde{K}_0(t,\xi,\eta) - \sum_{j=1}^{r} \chi_j(t)\, a_j(\xi)\, b_j(\eta)\, (k_j * g_A)(t,\Theta(\xi,\eta))$$

Claim 6.12
i) *The convolution* $(k_j * g_A)$ *makes sense*
ii)

$$\widetilde{H}^\xi \widetilde{K}_1(t,\xi,\eta) = a(\xi)\, \delta_{(0,\eta)}(t,\xi) + \widetilde{E}_1(t,\xi,\eta)$$

where $\widetilde{E}_1$ *is a kernel of type* 2.

Proof. Since $g_A \in F_2^A$ and $k_j \in F_1^A$, $(k_j * g_A)$ exists and belongs to F_3^A. By (6.10), (6.11) and Lemma 6.11, we can compute

$$\begin{aligned}&\widetilde{H}^\xi \widetilde{K}_1(t,\xi,\eta) = \\ &= \widetilde{H}^\xi \widetilde{K}_0(t,\xi,\eta) - \widetilde{H}^\xi \left(\sum_{j=1}^{r} \chi_j(t)\, a_j(\xi)\, b_j(\eta)\, (k_j * g_A)(t,\Theta(\xi,\eta)) \right) = \\ &= a(\xi)\, \delta_{(0,\eta)}(t,\xi) + \sum_{j=1}^{r} \chi_j(t)\, a_j(\xi)\, b_j(\eta)\, k_j(t,\Theta(\xi,\eta)) - \\ &- \sum_{j=1}^{r} \chi_j(t)\, a_j(\xi)\, b_j(\eta)\, [\mathcal{H}(k_j * g_A)](t,\Theta(\xi,\eta)) + \widetilde{E}_1(t,\xi,\eta)\end{aligned}$$

where $\widetilde{E}_1$ is a kernel of type 2. Finally, since $[\mathcal{H}(k_j * g_A)] = k_j$ we get

$$\widetilde{H}^\xi \widetilde{K}_1(t,\xi,\eta) = a(\xi)\, \delta_{(0,\eta)}(t,\xi) + \widetilde{E}_1(t,\xi,\eta)$$

■

Proceeding inductively, using Lemma 6.11 and the argument of the previous Claim, we can build two sequences $\widetilde{K}_i, \widetilde{E}_i$ such that:

$$\widetilde{H}^\xi \widetilde{K}_i(t,\xi,\eta) = a(\xi)\, \delta_{(0,\eta)}(t,\xi) + \widetilde{E}_i(t,\xi,\eta) \tag{6.12}$$

where $\widetilde{E}_i$ is a kernel of type $i+1$ which can be written as

$$\widetilde{E}_i(t,\xi,\eta) = \sum_{j=1}^{r} \chi_{i,j}(t)\, a_{i,j}(\xi)\, b_{i,j}(\eta)\, k_{i,j}(t,\Theta(\xi,\eta))$$

with $k_{i,j} \in F^A_{i+1}$, and

$$\widetilde{K}_{i+1}(t,\xi,\eta) = \widetilde{K}_i(t,\xi,\eta) - \sum_{j=1}^{r} \chi_{i,j}(t)\, a_{i,j}(\xi)\, b_{i,j}(\eta)\, (k_{i,j} * g_A)(t,\Theta(\xi,\eta)).$$

Moreover, $\widetilde{K}_i$ is a kernel of type 2 for every i.

We now recall that this construction can be performed in each of the neighborhoods $\widetilde{U}^\ell$. Therefore we can define

$$\widetilde{K}_i(t,\xi,\eta) = \sum_\ell \widetilde{K}^\ell_i(t,\xi,\eta)$$

and

$$\widetilde{E}_i(t,\xi,\eta) = \sum_\ell \widetilde{E}^\ell_i(t,\xi,\eta).$$

We now want to come back to the original (unlifted) variables. Let:

$$\xi = (x,h), \eta = (y,h')$$

and define:

$$K_j(t,x,y) = \int_{\mathbb{R}^{N-n}} \int_{\mathbb{R}^{N-n}} \widetilde{K}_j(t,\xi,\eta)\, dh\ dh'; \tag{6.13}$$
$$E_j(t,x,y) = \int_{\mathbb{R}^{N-n}} \int_{\mathbb{R}^{N-n}} \widetilde{E}_j(t,\xi,\eta)\, dh\ dh'.$$

Then:

Lemma 6.13 *There exists a constant $\beta > 0$ and, for every positive integer j there exists a constant $\mathbf{c}(j) > 0$ such that, for any $x, y \in B(0,R), t \in (0,1)$,*

$$H^{(t,x)}_A (K_j - h_A)(t,x,y) = E_j(t,x,y)$$

with

$$|E_j(t,x,y)| \le \mathbf{c}(j)\, t^{(j-1)/2} \frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|}.$$

Also,

$$|K_j(t,x,y)| \le \mathbf{c}(j) \frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|}.$$

Proof. By Lemma 6.4 we have

$$\begin{aligned}
H_A \int_{\mathbb{R}^{N-n}} \widetilde{K}_j(t,\xi,\eta)\,dh &= H_A \int_{\mathbb{R}^{N-n}} \sum_\ell \widetilde{K}_i^\ell(t,\xi,\eta)\,dh \\
&= \sum_\ell \int_{\mathbb{R}^{N-n}} \widetilde{H}_A^\ell \widetilde{K}_i^\ell(t,\xi,\eta)\,dh \\
&= \sum_\ell \int_{\mathbb{R}^{N-n}} \Big(a_\ell(\xi)\,\delta_{(0,\eta)}(t,\xi) + \widetilde{E}_i^\ell(t,\xi,\eta)\Big)\,dh \\
&= \int_{\mathbb{R}^{N-n}} \Big(a(\xi)\,\delta_{(0,\eta)}(t,\xi) + \widetilde{E}_i(t,\xi,\eta)\Big)\,dh.
\end{aligned}$$

Therefore

$$\begin{aligned}
H_A^{(t,x)} K_j(t,x,y) &= H_A^{(t,x)} \int_{\mathbb{R}^{N-n}} \int_{\mathbb{R}^{N-n}} \widetilde{K}_j(t,\xi,\eta)\,dh\ dh' \\
&= \int_{\mathbb{R}^{N-n}} H_A^{(t,x)} \int_{\mathbb{R}^{N-n}} \widetilde{K}_j(t,\xi,\eta)\,dh\ dh' \\
&= \int_{\mathbb{R}^{N-n}} \int_{\mathbb{R}^{N-n}} \Big(a(\xi)\,\delta_{(0,\eta)}(t,\xi) + \widetilde{E}_i(t,\xi,\eta)\Big)\,dhdh' \\
&\int_{\mathbb{R}^{N-n}} \int_{\mathbb{R}^{N-n}} a(\xi)\,\delta_{(0,\eta)}(t,\xi)\,dhdh' + E_i(t,x,y)\,.
\end{aligned}$$

By (6.5) we have

$$\int_{\mathbb{R}^{N-n}} \int_{\mathbb{R}^{N-n}} \varphi(h)\,\delta_{(0,\eta)}(t,\xi)\,dh\ dh' = \delta_{(0,y)}(t,x)$$

and since $H_A h_A(t,x,y) = \delta_{(0,y)}(t,x)$ we obtain

$$H_A^{(t,x)} K_j(t,x,y) = H_A h_A(t,x,y) + E_i(t,x,y)\,.$$

Now recall that

$$\left|\widetilde{E}_j(t,\xi,\eta)\right| \leq \mathbf{c}(j)\,t^{(j-1)/2} \sum_\ell g_A\left(\mathbf{c}t, \Theta^\ell(\xi,\eta)\right) \sum_{i=1}^r \chi_i^\ell(t)\,a_i^\ell(\xi)\,b_i^\ell(\eta)$$

hence

$$\begin{aligned}
&|E_j(t,x,y)| \leq \\
&\leq \mathbf{c}(j)\,t^{(j-1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} \left[\sum_\ell g_A\left(\mathbf{c}t, \Theta^\ell(\xi,\eta)\right) \sum_{i=1}^r \chi_i^\ell(t)\,a_i^\ell(\xi)\,b_i^\ell(\eta)\right] dh' \qquad (6.14) \\
&\leq \mathbf{c}(j)\,t^{(j-1)/2} \frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|}
\end{aligned}$$

where the last inequality follows from the next Lemma; the same argument proves the bound on K_j, remembering that $\widetilde{K}_j$ is a kernel of type 2. This finishes the proof. ■

Lemma 6.14 *For each term appearing in the integral in (6.14) the following bound holds,*

$$\int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} g_A\left(\mathbf{c}t, \Theta\left(\xi,\eta\right)\right) a\left(\xi\right) b\left(\eta\right) dh' \leq \mathbf{c}' \frac{e^{-\mathbf{c}'' d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|}.$$

This Lemma can be proved as the Remark on p.848 of [30], making also use of the uniform Gaussian bounds of [3]; in [30] the Authors write that "the proof is very similar to that of Lemmas 8, 9 in [53], where the 'elliptic' case is dealt". For seek of completeness we present a detailed proof of this fact.

Proof. By (6.1),

$$\begin{aligned}&\int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} g_A\left(\mathbf{c}t, \Theta\left(\xi,\eta\right)\right) a\left(\xi\right) b\left(\eta\right) dh' \\ &\leq \mathbf{c}\alpha\left(x\right) \int_{\mathbb{R}^{N-n}} \varphi\left(h\right) dh \int_{\mathbb{R}^{N-n}} t^{-Q/2} e^{-\beta \widetilde{d}(\xi,\eta)^2/t} b\left(\eta\right) dh'.\end{aligned}$$

Then we distinguish two cases.

1st case: $d^2\left(x,y\right) \geq t$. Then

$$\begin{aligned}&\int_{\mathbb{R}^{N-n}} t^{-Q/2} e^{-\beta \widetilde{d}(\xi,\eta)^2/t} b\left(\eta\right) dh' \\ &= \sum_{k=0}^{\infty} \int_{\left\{h' : 2^k d(x,y) \leq \widetilde{d}(\xi,\eta) \leq 2^{k+1} d(x,y)\right\}} t^{-Q/2} e^{-\beta \widetilde{d}(\xi,\eta)^2/t} b\left(\eta\right) dh'.\end{aligned}$$

By Lemma 6.3,

$$\begin{aligned}\left(2^{k+1} d\left(x,y\right)\right)^Q &\simeq \left|\widetilde{B}\left(\left(x,h\right), 2^{k+1} d\left(x,y\right)\right)\right| \simeq \\ &\simeq \left|B\left(x, 2^{k+1} d\left(x,y\right)\right)\right| \cdot \left|\left\{h' : \widetilde{d}\left(\left(x,h\right),\left(y,h'\right)\right) \leq 2^{k+1} d\left(x,y\right)\right\}\right|\end{aligned}$$

(strictly speaking, to apply Lemma 6.3 we should replace 2^k with δ^{-k}, where δ is the small number appearing in the Lemma); then

$$\begin{aligned}&\int_{\mathbb{R}^{N-n}} t^{-Q/2} e^{-\beta \widetilde{d}(\xi,\eta)^2/t} b\left(\eta\right) dh' \\ &\leq \sum_{k=0}^{\infty} t^{-Q/2} e^{-\beta 2^{2k} d(x,y)^2/t} \cdot \frac{\left(2^{k+1} d\left(x,y\right)\right)^Q}{\left|B\left(x, 2^{k+1} d\left(x,y\right)\right)\right|} \leq \\ &\leq \frac{1}{\left|B\left(x, d\left(x,y\right)\right)\right|} \left(\frac{d\left(x,y\right)}{\sqrt{t}}\right)^Q \sum_{k=0}^{\infty} e^{-\beta 2^{2k} d(x,y)^2/t} 2^{(k+1)Q} \leq\end{aligned}$$

since $s=\frac{d(x,y)}{\sqrt{t}}\geq 1$, and $s^{Q}e^{-\beta 2^{2k}s^{2}}\leq \mathbf{c}e^{-\beta 2^{2k-1}s^{2}}$

$$\leq \frac{1}{\left|B\left(x,\sqrt{t}\right)\right|}\sum_{k=0}^{\infty}e^{-\beta 2^{2k-1}d(x,y)^{2}/t}2^{(k+1)Q}=$$
$$=\frac{e^{-\beta d(x,y)^{2}/t}}{\left|B\left(x,\sqrt{t}\right)\right|}\sum_{k=0}^{\infty}e^{-\beta\left(2^{2k-1}-1\right)}2^{(k+1)Q}\leq \mathbf{c}\frac{e^{-\beta d(x,y)^{2}/t}}{\left|B\left(x,\sqrt{t}\right)\right|}$$

2nd case: $d^{2}(x,y)\leq t$. Then

$$\int_{\mathbb{R}^{N-n}}t^{-Q/2}e^{-\beta\widetilde{d}(\xi,\eta)^{2}/t}b(\eta)\,dh'$$
$$=\int_{\left\{h':\widetilde{d}(\xi,\eta)\leq 2\sqrt{t}\right\}}...dh'+\sum_{k=1}^{\infty}\int_{\left\{h':2^{k}\sqrt{t}\leq\widetilde{d}(\xi,\eta)\leq 2^{k+1}\sqrt{t}\right\}}...dh'.$$

Again by Lemma 6.3,

$$\left(2^{k+1}\sqrt{t}\right)^{Q}\simeq\left|\widetilde{B}\left((x,h),2^{k+1}\sqrt{t}\right)\right|\simeq$$
$$\simeq\left|B\left(x,2^{k+1}\sqrt{t}\right)\right|\cdot\left|\left\{h':\widetilde{d}\left((x,h),(y,h')\right)\leq 2^{k+1}\sqrt{t}\right\}\right|$$

hence

$$\sum_{k=1}^{\infty}\int_{\left\{h':2^{k}\sqrt{t}\leq\widetilde{d}(\xi,\eta)\leq 2^{k+1}\sqrt{t}\right\}}...dh'\leq\sum_{k=1}^{\infty}t^{-Q/2}e^{-\beta 2^{2k}}\frac{\left(2^{k+1}\sqrt{t}\right)^{Q}}{\left|B\left(x,2^{k+1}\sqrt{t}\right)\right|}\leq$$
$$\leq\frac{1}{\left|B\left(x,\sqrt{t}\right)\right|}\sum_{k=1}^{\infty}e^{-\beta 2^{2k}}\left(2^{k+1}\right)^{Q}\leq$$
$$\leq\frac{\mathbf{c}}{\left|B\left(x,\sqrt{t}\right)\right|}\leq\mathbf{c}\frac{e^{-\beta d(x,y)^{2}/t}}{\left|B\left(x,\sqrt{t}\right)\right|}$$

because $e^{-\beta d(x,y)^{2}/t}\geq\mathbf{c}>0$, being $d^{2}(x,y)\leq t$. Analogously one bounds the single term

$$\int_{\left\{h':\widetilde{d}(\xi,\eta)\leq 2\sqrt{t}\right\}}...dh'.$$

This completes the proof. ■

Our next task is to build a suitable barrier function f_j, to be compared with (K_j-h_A):

Lemma 6.15 *For any fixed $R,\varepsilon>0$ and every positive integer j large enough, there exist a function $f_j(t,x,y)$ and positive constants $c_j=\mathbf{c}(j,\varepsilon,R)$, $\gamma_1(\varepsilon,R)$, $\gamma_2(\varepsilon,R)$ such that for any $t\in(0,1)$, $x,y\in B(0,R)$:*
i)

$$e^{-\gamma_1 d(x,y)^{2}/t}\leq f_j(t,x,y)\leq\frac{c_j}{\left|B\left(x,\sqrt{t}\right)\right|}e^{-\gamma_2 d(x,y)^{2}/t}\quad \text{if } d(x,y)\simeq\varepsilon \qquad (6.15)$$

ii)

$$H_A^{(t,x)} f_j(t,x,y) \geq c_j \left| E_j(t,x,y) \right|. \tag{6.16}$$

Proof. By (6.14), we know that, for $x, y \in B(0,R)$,

$$\begin{aligned} &\left| E_j(t,x,y) \right| \leq \mathbf{c}(j)\, t^{(j-1)/2} \cdot \\ &\cdot \sum_{\ell} \sum_{i=1}^{r} \chi_i^{\ell}(t) \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} \left[g_A\left(\mathbf{c}t, \Theta^{\ell}(\xi,\eta)\right) a_i^{\ell}(\xi)\, b_i^{\ell}(\eta) \right] dh'. \end{aligned}$$

With this notation, let

$$\begin{aligned} &f_j(t,x,y) = B_j t^{(j+1)/2} \cdot \\ &\cdot \sum_{\ell} \sum_{i=1}^{r} \chi_i^{\ell}(t) \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} \left[g_A\left(\mathbf{c}t, \Theta^{\ell}(\xi,\eta)\right) a_i^{\ell}(\xi)\, b_i^{\ell}(\eta) \right] dh' \end{aligned}$$

with B_j constant to be chosen later. Just to simplify notation, we will write the proof assuming that

$$\left| E_j(t,x,y) \right| \leq \mathbf{c}(j)\, t^{(j-1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} g_A(\mathbf{c}t, \Theta(\xi,\eta))\, a(\xi)\, b(\eta)\, dh'$$

and

$$f_j(t,x,y) = B_j t^{(j+1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} g_A(\mathbf{c}t, \Theta(\xi,\eta))\, a(\xi)\, b(\eta)\, dh'.$$

By Lemma 6.14,

$$f_j(t,x,y) \leq \mathbf{c} B_j t^{(j+1)/2} \frac{e^{-\beta d(x,y)^2/t}}{\left| B\left(x,\sqrt{t}\right) \right|} \leq \mathbf{c} B_j \frac{e^{-\beta d(x,y)^2/t}}{\left| B\left(x,\sqrt{t}\right) \right|}$$

since $t \leq 1$.

To prove the first inequality in (6.15), we start noting that

$$\widetilde{d}(\xi,\eta) \leq \mathbf{c}(\varepsilon,R)\, d(x,y)$$

since $\widetilde{d}(\xi,\eta)$ is bounded and $d(x,y) \simeq \varepsilon$. By the lower bound in (6.1), taking into account the supports of a, b, and since $d(x,y) \simeq \varepsilon$

$$\begin{aligned} f_j(t,x,y) &\geq \mathbf{c} B_j t^{(j+1-Q)/2} \int_{|h| \leq c} dh \int_{|h'| \leq c} e^{-\widetilde{d}(\xi,\eta)^2/\mathbf{c}(\varepsilon,R)t} dh' \\ &\geq \mathbf{c} B_j e^{-d(x,y)^2/\mathbf{c}t} t^{(j+1-Q)/2} \\ &\geq \mathbf{c} B_j e^{-\varepsilon^2/\mathbf{c}t} t^{(j+1-Q)/2} \geq \mathbf{c} B_j e^{-\varepsilon^2/\mathbf{c}' t} \\ &\geq \mathbf{c} B_j e^{-d(x,y)^2/\mathbf{c}' t} \geq e^{-d(x,y)^2/\mathbf{c}' t} \end{aligned}$$

by a suitable choice of B_j.

To prove (6.16), we use Lemma 6.4 and compute:

$$H_A^{(t,x)} f_j(t,x,y) =$$

$$= B_j \frac{j+1}{2} t^{(j-1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} g_A(\mathbf{c}t, \Theta(\xi,\eta))\, a(\xi)\, b(\eta)\, dh' +$$
$$+ B_j t^{(j+1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} \widetilde{H}_A^{(t,x)} \left[g_A(\mathbf{c}t, \Theta(\xi,\eta))\, a(\xi)\right] b(\eta)\, dh'$$

by Lemma 6.11,

$$= B_j \frac{j+1}{2} t^{(j-1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} g_A(\mathbf{c}t, \Theta(\xi,\eta))\, a(\xi)\, b(\eta)\, dh' +$$
$$+ B_j t^{(j+1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} \left[(\mathcal{H}_A g_A)(\mathbf{c}t, \Theta(\xi,\eta))\, a(\xi)\, b(\eta) + E(t,\xi,\eta)\right] dh'$$

for some kernel $E(t,\xi,\eta)$ of type 1

$$= B_j \frac{j+1}{2} t^{(j-1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} g_A(\mathbf{c}t, \Theta(\xi,\eta))\, a(\xi)\, b(\eta)\, dh' +$$
$$+ B_j t^{(j+1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} E(t,\xi,\eta)\, dh'$$

where we used the fact that $t^{(j+1)/2} \delta_{(0,x)}(t,y) = 0$. Now,

$$\left| t^{(j+1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} E(t,\xi,\eta)\, dh' \right|$$
$$\leq \mathbf{c} t^{j/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} g_A(\mathbf{c}t, \Theta(\xi,\eta))\, a(\xi)\, b(\eta)\, dh'$$
$$\leq \mathbf{c} t^{(j-1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} g_A(\mathbf{c}t, \Theta(\xi,\eta))\, a(\xi)\, b(\eta)\, dh'$$

since $t < 1$; therefore, for j large enough,

$$H_A^{(t,x)} f_j(t,x,y) \geq$$
$$\geq \mathbf{c} B_j t^{(j-1)/2} \int_{\mathbb{R}^{N-n}} dh \int_{\mathbb{R}^{N-n}} g_A(\mathbf{c}t, \Theta(\xi,\eta))\, a(\xi)\, b(\eta)\, dh' \geq c_j \left|E_j(t,x,y)\right|.$$

■

Theorem 6.16 *There exist $\tau = \mathbf{c}$ and $\beta = \mathbf{c}$ such that, for every $R > 0$ we have*

$$h_A(t,x,y) \leq \mathbf{c}(R) \frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|}$$

for every $t \in (0,\tau)$, $x, y \in \mathbb{R}^n$, with $d(x,y) < R$.

Proof. By Theorem 4.10, for any fixed $\varepsilon > 0$, if $\varepsilon \leq d(x,y) \leq R$ and $t \in (0,\tau)$ we have:

$$h_A(t,x,y) \leq \mathbf{c}e^{-\mathbf{c}/t} \leq \mathbf{c}e^{-\mathbf{c}d(x,y)^2/R^2t} \leq$$

$$\leq \frac{\mathbf{c}}{|B(x,1)|}e^{-\mathbf{c}d(x,y)^2/R^2t} \leq \mathbf{c}'\frac{e^{-\mathbf{c}d(x,y)^2/R^2t}}{\left|B\left(x,\sqrt{t}\right)\right|}.$$

Hence we have to worry only for $d(x,y) < \varepsilon$. We will proceed in two steps: first we will prove the upper bound assuming $d(x,y) < \varepsilon$ and $x \in B(0,R)$; then we will consider the case $d(x,y) < \varepsilon$ and $x \notin B(0,R)$.

Step 1. It is enough to prove that, for $x \in B(0,R), d(x,y) < \varepsilon, t \in (0,1)$ and j large enough,

$$|K_j(t,x,y) - h_A(t,x,y)| \leq \mathbf{c}(j)\frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|}. \tag{6.17}$$

To show this, we want to apply the maximum principle on the cylinder

$$\{(t,x) : t \in (0,1), d(x,y) < \varepsilon\}, \text{ for } y \text{ fixed,}$$

to the functions

$$\pm(K_j - h_A) - A_j f_j$$

where f_j is as in Lemma 6.15, and A_j will be chosen later. We know that, for $d(x,y) \geq \varepsilon$, both $h_A(t,x,y)$ and $K_j(t,x,y)$ are bounded by $\mathbf{c}(j)\frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|}$, therefore (6.17) holds for $d(x,y) \simeq \varepsilon, t \in (0,1)$. Note also that, for $d(x,y) \geq \varepsilon$,

$$\frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|} \leq \mathbf{c}e^{-\beta'/t} \text{ for any } \beta' < \beta$$

(because $\left|B\left(x,\sqrt{t}\right)\right| \geq ct^{n/2m}$). Then for $d(x,y) \simeq \varepsilon, t \in (0,1)$, we can say that

$$|K_j - h_A| \leq \mathbf{c}e^{-\gamma_1 d(x,y)^2/t} \leq A_j f_j. \tag{6.18}$$

Moreover, by the subelliptic estimates on H_A, since $H_A(K_j - h_A) = E_j$ is as smooth as we need, for large j, even at $t = 0$, then $(K_j - h_A)$ is continuous at $t = 0$, and therefore vanishes (because for $t < 0$ both K_j and h_A vanish), hence (6.18) holds also for $t = 0, d(x,y) \leq \varepsilon$.

On the other hand, by Lemma 6.13,

$$H_A(K_j - h_A) = E_j$$

hence, by (6.15),

$$H_A(\pm(K_j - h_A) - A_j f_j) \leq |E_j(t,x,y)| - A_j c_j |E_j(t,x,y)| \leq 0$$

for suitable A_j. Therefore, by the maximum principle (Proposition 3.6),

$$|K_j - h_A| \leq A_j f_j \text{ for } d(x,y) \leq \varepsilon, t \in (0,1). \tag{6.19}$$

This ends the first step.

Step 2. We can assume R large enough such that H_A is the standard heat operator in the complement of $B(0, R/2)$. Let $x \notin B(0, R)$ and $d(x, y) < \varepsilon$; recall that in this region d equals the Euclidean distance, by Lemma 2.3. Let $h_0(t, x, y)$ be the fundamental solution of the heat operator in $\mathbb{R}^{n+1}$. Adapting the technique of Step 1, we now simply choose

$$f_j(t, x, y) = \mathbf{c} t^{\frac{j+1}{2}} h_0(t, x, y).$$

Then, one easily checks that, in this region,

$$e^{-\gamma_1 d(x,y)^2/t} \leq f_j(t, x, y)$$

for a positive constant γ_1, $t \in (0, 1)$, and $d(x, y) \simeq \varepsilon$, and that

$$H_A^{(t,x)} f_j(t, x, y) \geq 0.$$

One can now repeat the argument of step 1: choosing $K_j = h_0$ and exploiting the fact that

$$H_A^{(t,x)}(h_0 - h_A) = 0$$

one proves (6.17). This completes the proof. ∎

We eventually have to prove the upper bound on h_A removing the assumption $d(x, y) < R$. Again, this is accomplished exploiting the fact that outside a compact set H_A is the heat operator:

Theorem 6.17 *There exist $\tau' = \mathbf{c}$ and $\beta' = \mathbf{c}$ such that,*

$$h_A(t, x, y) \leq \mathbf{c} \frac{e^{-\beta' d(x,y)^2/t}}{\left|B\left(x, \sqrt{t}\right)\right|}$$

for every $t \in (0, \tau')$ and $x, y \in \mathbb{R}^n$.

Proof. Fix R large enough so that $\Omega_0 \subseteq B(0, R)$ and outside $B(0, R)$ the operator H_A reduces to the heat operator. For any fixed $y \in \mathbb{R}^n$ we define

$$\rho(y) = \begin{cases} R & \text{if } d(y, 0) > 3R \\ \\ 5R & \text{otherwise.} \end{cases}.$$

It is easy to check that

$$x \in \partial B(y, \rho(y)) \Longrightarrow B(x, R) \cap B(0, R) = \varnothing. \tag{6.20}$$

Also observe that by Theorem 6.16

$$h_A(t, x, y) \leq \mathbf{c}(R) \frac{e^{-\frac{\beta d(x,y)^2}{t}}}{\left|B\left(x, \sqrt{t}\right)\right|} \tag{6.21}$$

for every $x \in \overline{B(y, \rho(y))}$, $t \in (0, \tau)$.

Let now

$$\gamma(t, x) = h_A(t, x, y)$$

$$w_\delta(t, x) = K(\delta t, x - y) \text{ where } K(t, x) = t^{-n/2} e^{-\frac{|x|^2}{t}}$$

$$D_\tau = (0, \tau') \times \left(\mathbb{R}^n \setminus \overline{B(y, \rho(y))}\right)$$

with τ' to be chosen later, $\tau' < \tau$. We aim to apply the maximum principle (Proposition 3.6) on D_τ to show that $\gamma \leq \mathbf{c}(\delta, R) w_\delta$. The parabolic boundary of D_τ is

$$\partial_p D_\tau = \partial_0 \cup \partial_1$$

where

$$\partial_0 = \{0\} \times (\mathbb{R}^n \setminus B(y, \rho(y)))$$

and

$$\partial_1 = [0, \tau'] \times \partial B(y, \rho(y)).$$

Let $(t, x) \in \partial_1$; by (6.20), $B\left(x, \sqrt{\tau'}\right) \cap B(0, R) = \varnothing$, since we can assume $\sqrt{\tau'} < R$. This implies that $\left|B\left(x, \sqrt{t}\right)\right| = c_n t^{n/2}$, by Lemma 2.3. Then by (6.21), and since $d(x, y) \geq \mathbf{c}|x - y|$ by Lemma 2.4, we have

$$\gamma(t, x) \leq \mathbf{c}(R) t^{-n/2} e^{-\frac{\beta d(x,y)^2}{t}} \leq \mathbf{c}(R) t^{-n/2} e^{-\frac{\mathbf{c}|x-y|^2}{t}} \leq \mathbf{c}(R) \delta^{n/2} w_\delta(t, x)$$

for a suitable δ (large enough). Note that on ∂_0 on has $\gamma(t, x) = 0 = w_\delta(t, x)$. Moreover for $(t, x) \in D_\tau$, using the global, non-uniform estimates in [35] we have

$$\lim_{x \to \infty} \sup_{t \in [0, \tau']} \gamma(t, x) \leq \lim_{x \to \infty} \sup_{t \in [0, \tau']} \mathbf{c}(A) \frac{e^{-\frac{d(x,y)^2}{\mathbf{c}(A)t}}}{\left|B\left(x, \sqrt{t}\right)\right|}$$

$$\leq \lim_{x \to \infty} \sup_{t \in [0, \tau']} \mathbf{c}(A) t^{-n/2} e^{-\frac{|x-y|^2}{\mathbf{c}(A)t}} = 0.$$

The above arguments show that

$$\gamma(t, x) \leq \mathbf{c}(R) \delta^{n/2} w_\delta(t, x)$$

on the parabolic boundary of D_τ and at the infinity, for δ large enough. We now want to prove that in D_τ, again for δ large enough,

$$H_A(\gamma) \leq H_A\left(\mathbf{c}(R) \delta^{n/2} w_\delta\right).$$

Since $H_A(\gamma) = 0$ in D_τ, it is enough to show that $H_A(w_\delta) \geq 0$.

We first compute:

$$\frac{\partial K}{\partial t}(t,x) = K(t,x)\left(-\frac{n}{2}t^{-1} + |x|^2 t^{-2}\right),$$
$$\frac{\partial K}{\partial x_j}(t,x) = K(t,x)\left(-2t^{-1}x_j\right),$$
$$\frac{\partial^2 K}{\partial x_i \partial x_j}(t,x) = K(t,x)\left(-2t^{-1}\delta_{ij} + 4t^{-2}x_i x_j\right).$$

Writing explicitly the differential operator L_A in Cartesian coordinates we have that

$$\begin{aligned}|L_A w_\delta(t,x)| &\le \mathbf{c}\sum_{i,j=1}^{n}\left(\left|\frac{\partial w_\delta}{\partial x_i}(t,x)\right| + \left|\frac{\partial^2 w_\delta}{\partial x_i \partial x_j}(t,x)\right|\right)\\ &\le \mathbf{c}K(\delta t, x-y)\left(\frac{2|x-y|}{\delta t} + \frac{2}{\delta t} + \frac{4|x-y|^2}{\delta^2 t^2}\right).\end{aligned}$$

Moreover

$$\frac{\partial w_\delta}{\partial t}(t,x) = \delta\frac{\partial K}{\partial t}(\delta t, x-y) = K(\delta t, x-y)\left(-\frac{n}{2t} + \frac{|x-y|^2}{\delta t^2}\right).$$

By Lemma 2.3 and the definition of D_τ, for $(t,x) \in D_\tau$ we have $|x-y| \ge R$. Therefore

$$\begin{aligned}&H_A(w_\delta) = \frac{\partial w_\delta}{\partial t}(t,x) - L_A w_\delta(t,x)\\ &\ge \frac{K(\delta t, x-y)}{\delta t^2}\left\{\left(-\frac{n\delta t}{2} + |x-y|^2\right) - \mathbf{c}\left(2t|x-y| + 2t + \frac{4|x-y|^2}{\delta}\right)\right\}\\ &\ge \frac{K(\delta t, x-y)}{\delta t^2}|x-y|^2\left\{\left(-\frac{n}{2}\frac{\delta\tau'}{R^2} + 1\right) - \mathbf{c}\left(\frac{2\tau'}{R} + \frac{2\tau'}{R^2} + \frac{4}{\delta}\right)\right\}\end{aligned}$$

Now, for $\tau' = 1/\delta^2$, and δ large enough depending on R, we get $H_A(w_\delta) \ge 0$ in D_τ. Hence the maximum principle implies that

$$\gamma(t,x) \le \mathbf{c}(R)\,\delta^{n/2} w_\delta(t,x) \text{ in } D_\tau,$$

that is

$$h_A(t,x,y) \le \mathbf{c}(R)\,t^{-n/2}e^{-|x-y|^2/\mathbf{c}t} \text{ for } (t,x) \in D_\tau.$$

We now use the fact that, for $(t,x) \in D_\tau$, $|x-y| \ge R > 1$, so that by (2.3), (2.4) in Lemma 2.4

$$\mathbf{c}^{-1}d(x,y) \le |x-y| \le \mathbf{c}d(x,y).$$

Then:

$$\begin{aligned}h_A(t,x,y) &\le \mathbf{c}(R)\, t^{-n/2} e^{-|x-y|^2/2\mathbf{c}t} e^{-|x-y|^2/2\mathbf{c}t} \le \\ &\le \mathbf{c}(R)\left(t^{-n/2}e^{-1/2\mathbf{c}t}\right) e^{-|x-y|^2/2\mathbf{c}t} \le \\ &\le \mathbf{c}(R)\, e^{-d(x,y)^2/2\mathbf{c}t} \le \\ &\le \mathbf{c}(R)\, \frac{e^{-d(x,y)^2/2\mathbf{c}t}}{\left|B\left(x,\sqrt{t}\right)\right|}\end{aligned}$$

because, by (2.7),

$$\left|B\left(x,\sqrt{t}\right)\right| \le |B(x,1)| \le \mathbf{c}.$$

■

Corollary 6.18 *There exists* $\beta = \mathbf{c}$ *such that, for any* $T > 0$ *we have*

$$h_A(t,x,y) \le \mathbf{c}(T)\, \frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|}$$

for every $0 < t < T$ *and* $x, y \in \mathbb{R}^n$.

Proof. For $T > 0$ fixed, let k be an integer large enough so that $\frac{T}{k} < \tau'$, where τ' is the number appearing in the previous Theorem. Then, by the reproduction property (3.4) and the previous theorem,

$$\begin{aligned}&h_A(t,x,y) = \\ &= \int_{\mathbb{R}^n} \dots \int_{\mathbb{R}^n} h_A\left(\frac{t}{k},x,y_1\right) h_A\left(\frac{t}{k},y_1,y_2\right) \dots h_A\left(\frac{t}{k},y_{k-1},y\right) dy_1 dy_2 \dots dy_{k-1} \le \\ &\le \int_{\mathbb{R}^n} \dots \int_{\mathbb{R}^n} \mathbf{c}\frac{e^{-\frac{\beta d^2(x,y_1)}{t/k}}}{\left|B\left(x,\sqrt{t/k}\right)\right|}\mathbf{c}\frac{e^{-\frac{\beta d^2(y_1,y_2)}{t/k}}}{\left|B\left(y_1,\sqrt{t/k}\right)\right|}\dots\mathbf{c}\frac{e^{-\frac{\beta d^2(y_{k-1},y)}{t/k}}}{\left|B\left(y_{k-1},\sqrt{t/k}\right)\right|} dy_1 dy_2 \dots dy_{k-1}.\end{aligned}$$

We now use the global, long time estimates from below and from above, proved in [36, Corollary 3.25 p.182] for the fundamental solution h_0 of the operator, fixed once and for all,

$$H_0 = \partial_t - \sum_{i=1}^m X_i^* X_i,$$

namely:

$$\mathbf{c}_1 \frac{e^{-\beta_1 d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|} \le h_0(t,x,y) \le \mathbf{c}_2 \frac{e^{-\beta_2 d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|}$$

for $x, y \in \mathbb{R}^n, t \in (0,\infty)$, with constants $\mathbf{c}_1, \mathbf{c}_2, \beta_1, \beta_2$ only depending on n and the vector fields $X_1, \dots, X_m$. Hence

$$h_A(t,x,y) \le$$

$$\leq \mathbf{c}^k \int_{\mathbb{R}^n} ... \int_{\mathbb{R}^n} h_0 \left(c' \frac{t}{k}, x, y_1 \right) h_0 \left(c' \frac{t}{k}, y_1, y_2 \right) ...h_0 \left(c' \frac{t}{k}, y_{k-1}, y \right) dy_1 dy_2 ... dy_{k-1} =$$

$$= \mathbf{c}^k h_0 \left(c't, x, y \right) \leq \mathbf{c}(T) \frac{e^{-\beta d(x,y)^2/t}}{\left| B\left(x, \sqrt{t}\right) \right|}.$$

■

7 Uniform lower bounds for fundamental solutions

In this section we want to prove the following

Theorem 7.1 *There exists $\beta = \mathbf{c}$ such that, for any $T > 0$ we have*

$$h_A(t, x, y) \geq \frac{\mathbf{c}(T)}{\left| B\left(x, \sqrt{t}\right) \right|} e^{-\beta d(x,y)^2/t}$$

for all $t \in (0, T), x, y \in \mathbb{R}^n$.

This result will be deduced by the following

Lemma 7.2 *There exists $\varepsilon = \mathbf{c}$ and $\mathbf{c}_1 \in (0, 1)$ such that*

$$h_A(t, x, y) \geq \frac{\mathbf{c}_0}{\left| B\left(x, \sqrt{t}\right) \right|} \tag{7.1}$$

whenever $d(x, y)^2 \leq \mathbf{c}_1 t$ and $t \leq \varepsilon, x, y \in \mathbb{R}^n$.

Proof of the Lemma. The proof relies on the same construction of the previous section. From the proof of Theorem 6.16, Step 1, we read (see (6.19))

$$\left| \left(K_j - h_A \right)(t, x, y) \right| \leq A_j f_j(t, x, y) \tag{7.2}$$

for a suitable j, $x \in B(0, R), d(x, y)$ small enough and any $t \in (0, 1)$, where

$$f_j(t, x, y) \leq \mathbf{c}(j) \ \frac{t^{(j+1)/2}}{\left| B\left(x, \sqrt{t}\right) \right|}. \tag{7.3}$$

Next we will show that for every j there exists ε such that for $x \in B(0, R), t < \varepsilon$ and $d(x, y)^2 \leq \mathbf{c}_1 t$ one has

$$K_j(t, x, y) \geq \frac{\mathbf{c}(j)}{\left| B\left(x, \sqrt{t}\right) \right|}. \tag{7.4}$$

To prove this, we start recalling the following reformulation of Lemma 6.3: there exists $\mathbf{c}_1 \in (0, 1)$ such that for $d(x, y)^2 \leq \mathbf{c}_1 t$, we have

$$\int_{\left\{ h:\ \widetilde{d}^2((x,h),(y,h')) \leq t \right\}} t^{-Q/2} dh \geq \frac{\mathbf{c}_0}{\left| B\left(x, \sqrt{t}\right) \right|} \tag{7.5}$$

Moreover, we need the following:

Claim 7.3 *For every i there exists ε such that for $t < \varepsilon$ and $\widetilde{d}(\xi,\eta)^2 \leq t$ one has*

$$\widetilde{K}_i(t,\xi,\eta) \geq \mathbf{c}(i)\, t^{-Q/2}. \tag{7.6}$$

Proof of the Claim. We will prove (7.6) by induction. By the lower bound in (6.1), we have

$$\begin{aligned}\widetilde{K}_0(t,\xi,\eta) &= \sum_{\ell} \chi^{\ell}(t)\, a^{\ell}(\xi)\, b^{\ell}(\eta)\, g_A\left(t, \Theta^{\ell}(\xi,\eta)\right) \\ &\geq \mathbf{c} t^{-Q/2} e^{-\widetilde{d}(\xi,\eta)^2/\mathbf{c}t} \geq \mathbf{c} t^{-Q/2}\end{aligned}$$

for $\widetilde{d}(\xi,\eta)^2 \leq t$. We now assume that (7.6) holds for $\widetilde{K}_i$. Recall that

$$\widetilde{K}_{i+1}(t,\xi,\eta) = \widetilde{K}_i(t,\xi,\eta) - \sum_{\ell} \sum_{j=1}^{r} \chi^{\ell}_{i,j}(t)\, a^{\ell}_{i,j}(\xi)\, b^{\ell}_{i,j}(\eta)\, (k_{i,j} * g_A)\left(t, \Theta^{\ell}(\xi,\eta)\right)$$

where $k_{i,j} \in F^A_{i+1}$, so that by Lemma 6.8 we have $k_{i,j} * g_A \in F^A_{i+3}$ and

$$\begin{aligned}&\left| \sum_{\ell} \sum_{j=1}^{r} \chi^{\ell}_{i,j}(t)\, a^{\ell}_{i,j}(\xi)\, b^{\ell}_{i,j}(\eta)\, (k_{i,j} * g_A)\left(t, \Theta^{\ell}(\xi,\eta)\right) \right| \leq \\ &\leq \mathbf{c} t^{\frac{i+1}{2}} \sum_{\ell} \chi^{\ell}_{i,j}(t)\, a^{\ell}_{i,j}(\xi)\, b^{\ell}_{i,j}(\eta)\, g_A\left(t, \Theta^{\ell}(\xi,\eta)\right) \leq \mathbf{c} t^{\frac{i+1-Q}{2}}.\end{aligned}$$

Therefore

$$\widetilde{K}_{i+1}(t,\xi,\eta) \geq \mathbf{c}_1 t^{-Q/2} - \mathbf{c}_2 t^{\frac{i+1-Q}{2}} \geq \mathbf{c}_3 t^{-Q/2}$$

for sufficiently small t. This proves the Claim. ■

We can now prove (7.4) as follows: by the Claim, (7.5), and the definition of K_j,

$$\begin{aligned}K_j(t,x,y) &= \int_{\mathbb{R}^{N-n}} \int_{\mathbb{R}^{N-n}} \widetilde{K}_j(t,\xi,\eta)\, dh\, dh' \geq \\ &\geq \int_{\left\{h:\ \widetilde{d}^2((x,h),(y,h')) \leq t\right\}} t^{-Q/2} dh \\ &\geq \frac{\mathbf{c}_0}{\left|B\left(x,\sqrt{t}\right)\right|}.\end{aligned}$$

Lemma 7.2, under the further assumption $x \in B(0,R)$, then follows from (7.2), (7.3) and (7.4). To end the proof of Lemma 7.2, assume now that $x \notin B(0,R)$, with R large enough such that H_A is the standard heat operator in the complement of $B(0,R/2)$. Reasoning like in Step 2 of the proof of Theorem 6.16, we then prove that

$$\left|(h_0 - h_A)(t,x,y)\right| \leq A_j f_j(t,x,y)$$

where h_0 is the heat kernel,

$$f_j(t,x,y) = \mathbf{c} t^{\frac{j+1}{2}} h_0(t,x,y)$$

and for $t < \varepsilon$ and $d(x,y)^2 \leq \mathbf{c}_1 t$ (recall that here $d(x,y) = |x-y|$) one has

$$h_0(t,x,y) \geq \frac{\mathbf{c}}{\left|B\left(x,\sqrt{t}\right)\right|} = \frac{\mathbf{c}}{t^{n/2}}.$$

Then we conclude as in the case $x \in B(0,R)$. ■

We now come to the

Proof of Theorem 7.1. (This proof is similar to that of Theorem 4 in [30], p.853, with minor corrections). Let $d(x,y) = \rho > 0$ and take a subunit path γ connecting x and y, with $l(\gamma) \leq 2\rho$. Take $k+1$ points $x_0, x_1, ..., x_k$ on γ with $x_0 = x$, $x_k = y$, $d(x_i, x_{i+1}) \leq 2\rho/k, i = 0, ..., k-1$, and let

$$B_i = B(x_i, \sigma), i = 0, ..., k-1$$

with σ to be chosen later. Note that if $y_i \in B_i$, then $d(y_i, y_{i+1}) \leq 2\rho/k + 2\sigma$. By the reproduction property (Corollary 3.7),

$$\begin{aligned} &h_A(t,x,y) = \\ &\int_{\mathbb{R}^n} ... \int_{\mathbb{R}^n} h_A\left(\frac{t}{k}, x, y_1\right) h_A\left(\frac{t}{k}, y_1, y_2\right) ... h_A\left(\frac{t}{k}, y_{k-1}, y\right) dy_1 dy_2 ... dy_{k-1} \geq \\ &\geq \int_{B_1 \times ... \times B_{k-1}} ... \int h_A\left(\frac{t}{k}, x, y_1\right) h_A\left(\frac{t}{k}, y_1, y_2\right) ... h_A\left(\frac{t}{k}, y_{k-1}, y\right) dy_1 dy_2 ... dy_{k-1} \end{aligned} \tag{7.7}$$

To apply Lemma 7.2 to each factor in the integral (7.7), we need to know that

$$d^2(y_i, y_{i+1}) \leq \frac{\mathbf{c}_1 t}{k} \text{ for any } y_i \in B_i, \text{and } \frac{t}{k} < \varepsilon.$$

This follows from

$$\left(\frac{2\rho}{k} + 2\sigma\right)^2 \leq \frac{\mathbf{c}_1 t}{k} \text{ and } \frac{T}{k} < \varepsilon$$

which, in turn, hold provided

$$\sigma = \frac{1}{4}\sqrt{\frac{\mathbf{c}_1 t}{k}}, k \geq \frac{T}{\varepsilon} \text{ and } k \geq \frac{16\rho^2}{\mathbf{c}_1 t}.$$

Then we choose k such that

$$k \leq \frac{16\rho^2}{\mathbf{c}_1 t} + \frac{T}{\varepsilon} + 1 \leq k+1$$

and get, by (7.7) and Lemma 7.2,

$$h_A(t,x,y) \geq \frac{\mathbf{c}_0}{\left|B\left(x,\sqrt{t/k}\right)\right|} \frac{\mathbf{c}_0\left|B\left(x_1, \frac{1}{4}\sqrt{\frac{\mathbf{c}_1 t}{k}}\right)\right|}{\left|B\left(y_1,\sqrt{t/k}\right)\right|} \cdot ... \cdot \frac{\mathbf{c}_0 B\left(x_{k-1}, \frac{1}{4}\sqrt{\frac{\mathbf{c}_1 t}{k}}\right)}{\left|B\left(y_{k-1},\sqrt{t/k}\right)\right|} \geq$$

by the doubling property

$$\geq \frac{\mathbf{c}_0^k \mathbf{c}_2^{k-1}}{\left|B\left(x, \sqrt{t/k}\right)\right|} \geq \mathbf{c} \frac{e^{-\alpha k}}{\left|B\left(x, \sqrt{t}\right)\right|}.$$

Now: if $\frac{16\rho^2}{\mathbf{c}_1 t} \geq \frac{T}{\varepsilon}$, then $k \leq \frac{32\rho^2}{\mathbf{c}_1 t}$ and

$$h_A(t,x,y) \geq \mathbf{c} \frac{e^{-\alpha k}}{\left|B\left(x, \sqrt{t}\right)\right|} \geq \mathbf{c} \frac{e^{-\beta d^2(x,y)/t}}{\left|B\left(x, \sqrt{t}\right)\right|};$$

if $\frac{16\rho^2}{\mathbf{c}_1 t} \leq \frac{T}{\varepsilon}$, then $k \leq \frac{2T}{\varepsilon}$ and

$$h_A(t,x,y) \geq \mathbf{c} \frac{e^{-\alpha k}}{\left|B\left(x, \sqrt{t}\right)\right|} \geq \mathbf{c} \frac{e^{-\beta T/\varepsilon}}{\left|B\left(x, \sqrt{t}\right)\right|} \geq \mathbf{c} e^{-\beta T/\varepsilon} \frac{e^{-\beta d^2(x,y)/t}}{\left|B\left(x, \sqrt{t}\right)\right|}.$$

So in any case we get the conclusion. ■

Remark 7.4 *From Theorem 7.1 and the doubling condition, we also read that, for any $\beta' > 0$, any $A \in \mathcal{B}_\lambda$*

$$\frac{e^{-\beta' d(x,y)^2/t}}{\left|B\left(x, \sqrt{t}\right)\right|} \leq \mathbf{c}(\beta', T)\, h_A(\mathbf{c}_1(\beta')\, t, x, y)$$

for any $t \in (0,T), x, y \in \mathbb{R}^n$. We will use several times this property, in the following sections.

8 Uniform upper bounds for the derivatives of fundamental solutions

Following a technique already exploited in [30], in this section we will show how the uniform upper bound proved in Theorem 6.16 implies an analogous upper bound on the derivatives of h_A:

Theorem 8.1 *There exists $\beta = \mathbf{c}$ such that, for every nonnegative integer i, multiindices I, J, and for any $T > 0$, we have*

$$\left|\partial_t^i X_I^x X_J^y h_A(t,x,y)\right| \leq \mathbf{c}(i,I,J,T)\ \ t^{-i-(|I|+|J|)/2} \frac{e^{-\beta d(x,y)^2/t}}{\left|B\left(x, \sqrt{t}\right)\right|}$$

for every $x, y \in \mathbb{R}^n, t \in (0,T)$.

The key tool to prove the above Theorem is the use of suitable "dilations" which exist even in this non-homogeneous context, due to the following result of Fefferman and Sanchez-Calle (see [24], Lemma 3 p.253; see also [30], Lemma 3 p.851), which we state here in a fashion adapted to our context:

Lemma 8.2 *There exist constants* $\mathbf{c}, \mathbf{c}' \in (0,1)$ *such that given any metric ball* $B = B(x,r) \subset \mathbb{R}^n$ *there is a smooth invertible map* $\Phi_B : [-1,1]^n \rightarrow B$, *satisfying:*
i) $\Phi_B([-1,1]^n) \supseteq B(x, \mathbf{c}r) \supseteq \Phi_B\left([-\mathbf{c}', \mathbf{c}']^n\right), \Phi_B(0) = x;$
ii) the push-forward $\left(\Phi_B^{-1}\right)_* (X_j) \equiv \left(X_j\left[f \circ \Phi_B^{-1}\right]\right) \circ \Phi_B = r^{-1} Z_{j,B}$, *the coefficients of* $Z_{j,B}$ *having all their derivatives bounded independently of* B*;*
iii) If L_A^B *is defined by*

$$L_A^B(f) = \left[r^2 L_A\left(f \circ \Phi_B^{-1}\right)\right] \circ \Phi_B, \quad \left(L_A^B = \sum_{i,j=1}^{m} a_{ij} Z_{i,B} Z_{j,B}\right)$$

then the family $\left\{L_A^B\right\}$ *is uniformly subelliptic, which means in particular that: for every couple of cutoff functions* φ_1, φ_2*,with* $\varphi_2 = 1$ *on* $\operatorname{supp}\varphi_1$*, and every* $s > 0$*, there exists a constant* $\mathbf{c}(s, \varphi_1, \varphi_2)$ *such that*

$$\|\varphi_1 w\|_{H^s} \leq \mathbf{c}(s, \varphi_1, \varphi_2)\left\{\left\|\varphi_2\left(\partial_t - L_A^B\right) w\right\|_{H^s} + \|\varphi_2 w\|_{L^2}\right\} \qquad (8.1)$$

for every $w \in C^\infty(\operatorname{supp}\varphi_2)$*. In particular,* $\mathbf{c}$ *is independent of* B*, and depends on* A *only through* λ*. (Actually, we will prove (8.1) only for* s *even integer, but this will be enough to complete the proof of Theorem 8.1).*
iv) Finally, if we replace L_A *with* L_A^* *and* L_A^B *with* $\left(L_A^*\right)^B$*, (8.1) still holds for* $\left(L_A^*\right)^B$*.*

Proof of the Lemma. The map Φ_B is the one constructed by Fefferman-Sanchez-Calle in [24, §1], depending on the vector fields $X_1, X_2, ..., X_m$ and the ball B (but not the coefficients a_{ij}). They prove i), ii); they also prove iii) for a fixed matrix A. In order to obtain uniform subelliptic estimate adapted to our context (that is, uniform with respect to both the matrix A and the ball B) we can revise the proof of these estimates as given by Kohn in [31]. This inspection shows that:

1. Replacing the operator $\sum Z_i^2$ with $\sum a_{ij} Z_i Z_j$ is harmless, since the few inequalities which explicitly involve the coefficients a_{ij} (and not only the vector fields) can be carried out by our ellipticity assumption (H2), getting constants depending on the a_{ij}'s only through the number λ (this fact has already been pointed out in [9, Theorem 20]).

2. Replacing a "fixed" system of vector fields with one depending on a parameter (as the system $Z_{j,B}$ depending on the ball B) keeps uniform bounds, as soon as we check that:

2.a. The coefficients of the vector fields $Z_{j,B}$ and their derivatives are bounded independently of B (we already know this fact by point (ii)).

2.b. Any direction w_j can be obtained by means of the vector fields $Z_{j,B}$ and their commutators using coefficients that are uniformly bounded. More precisely one can write $\partial_{w_j} = \sum_I \alpha_{I,B} Z_{I,B}$ with coefficients $\alpha_{I,B}$ bounded independently of B. (This is needed in the "Proof of Theorem H", see [31, p. 65]).

2.c. If we define the operator Λ^s setting:

$$\widehat{(\Lambda^s f)}(\xi) = \left(1 + |\xi|^2\right)^{s/2} \widehat{f}(\xi);$$

then the commutator

$$\left[L_A^B, \Lambda^s\right]$$

is of type $s+1$, uniformly with respect to A, B, that is:

$$\left\| \left[L_A^B, \Lambda^s\right] u \right\|_{L^2} \leq c \left\| u \right\|_{H^{s+1}} \text{ for any } s > 0,$$

where H^s is the standard Sobolev space of fractional order. (This is needed in the proof of (11) in [31, p. 64]). So, let us prove 2.b and 2.c.

Proof of 2.b. Let

$$Y_B g\left(x\right) = r^{-1} \partial_{w_j} \left[g\left(\Phi_B\left(\cdot\right)\right)\right] \left(\Phi_B^{-1}\left(x\right)\right).$$

By [24, Lemma 1] $\left|\partial_{w_i} \Phi_B\left(w\right)\right| \leqslant cr$ and therefore Y_B has uniformly bounded coefficients. Assume that $Y_B g\left(x\right) = \sum a_i \partial_{x_i}$. Applying Hörmander's to ∂_{x_i} shows that we can write

$$Y_B g\left(x\right) = \sum_I \alpha_{I,B} X_I g\left(x\right)$$

for suitable uniformly bounded coefficients $\alpha_{I,B}$.

By Lemma 8.2 ii) we have $X_I g = r^{-1} Z_{I,B} \left[g\left(\Phi_B\left(\cdot\right)\right)\right]$ and consequently

$$r^{-1} \partial_{w_j} \left[g\left(\Phi_B\left(\cdot\right)\right)\right] = r^{-1} \sum_I \alpha_{I,B} Z_{I,B} \left[g\left(\Phi_B\left(\cdot\right)\right)\right]$$

that is

$$\partial_{w_j} = \sum_I \alpha_{I,B} Z_{I,B}.$$

Proof of 2.c. The bound

$$\left\| \left[L_A^B, \Lambda^s\right] u \right\|_{L^2} \leq c \left\| u \right\|_{H^{s+1}}$$

will be proved as soon as we show that the commutator $[a, \Lambda^s]$, where a denotes multiplication by a smooth function a, is of type $s-1$, that is

$$\left\| \left[a, \Lambda^s\right] u \right\|_{L^2} \leq c \left\| u \right\|_{H^{s-1}}, \tag{8.2}$$

with c depending on a only through upper bounds on a and its derivatives. Let us prove (8.2) for s even integer. Let Δ be the standard Laplace operator. For any positive integer h, we have

$$\widehat{\left(\Delta^h f\right)}\left(\xi\right) = \left(-1\right)^h \left|\xi\right|^{2h} \widehat{f}\left(\xi\right);$$

hence

$$\widehat{\left(\Lambda^{2k} f\right)}\left(\xi\right) = \left(1 + \left|\xi\right|^2\right)^k \widehat{f}\left(\xi\right) =$$

$$= \sum_{h=0}^{k} \binom{k}{h} \left|\xi\right|^{2h} \widehat{f}\left(\xi\right) = \sum_{h=0}^{k} \binom{k}{h} \left(-1\right)^h \widehat{\left(\Delta^h f\right)}\left(\xi\right)$$

and

$$\Lambda^{2k} f = \sum_{h=0}^{k} \binom{k}{h} (-1)^h \Delta^h f$$

is a constant coefficient differential operator of order $2k$. Therefore $\left[\Lambda^{2k}, a\right]$ is a differential operator of order $2k-1$, with coefficients given by a and derivatives of a of order up to $2k$. This proves (8.2), and therefore (4.16), for s even integer.

As to point iv), the same reasoning of point iii) applies. ■

Proof of the Theorem. Here we follow the argument in [30], pp.851-2.

Fix $t_0 \in (0,T), x_0, y_0 \in \mathbb{R}^n$. While in [30] the Authors only have to care about the case $t_0 \ll d(x_0,y_0)^2$, we have to handle separately two cases.

i) $4t_0 \le d(x_0,y_0)^2$. Making the change of variables $t = t_0 s, x = \Phi_B(z)$, with $B = B\left(x_0, \sqrt{t_0}\right)$, we see that the function $u(s,z) = h(t,x,y_0)$ satisfies, by Corollary 6.18,

$$\begin{aligned}
&\left(\partial_s - L_A^B\right) u(s,z) = 0 \text{ if } |s| \le 2, |z| \le 1 \\
&|u(s,z)| \le \mathbf{c}(T) \frac{e^{-\beta d(x,y_0)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|} \le \mathbf{c}(T) \frac{e^{-\beta d(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|} \text{ if } \frac{1}{2} \le |s| \le 2, |z| \le 1
\end{aligned}$$

because: $\left|B\left(x,\sqrt{t}\right)\right| \simeq \left|B\left(x_0,\sqrt{t_0}\right)\right|$ since $t \simeq t_0$ and $x \in B\left(x_0,\sqrt{t_0}\right)$; $d(x,y_0) \simeq d(x_0,y_0)$ since $d(x_0,y_0) \ge 2\sqrt{t_0} > 2d(x,x_0)$. Now, apply (8.1) to u choosing $\varphi_1 \equiv 1$ on $\{3/4 \le |s| \le 3/2, |z| \le \mathbf{c}'\}, \varphi_2 \equiv 1$ on sprtφ_1, and sprt$\varphi_2 \subset \{1/2 \le |s| \le 2, |z| \le 1\}$:

$$\|\varphi_1 u\|_{H^\sigma} \le \mathbf{c}(\sigma, \varphi_1, \varphi_2) \|\varphi_2 u\|_{L^2}$$

hence, for any integer k (applying the previous inequality with $\sigma = \sigma(k)$ large enough)

$$\|u\|_{C^k(\{3/4 \le |s| \le 3/2, |z| \le \mathbf{c}'\})} \le \mathbf{c}(T,k) \frac{e^{-\beta d(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|}.$$

In particular, by Lemma 8.2 ii),

$$\begin{aligned}
\left|\partial_t^i X_I^x h_A(t_0,x_0,y_0)\right| &= t_0^{-i-|I|/2} \left|\partial_s^i Z_{I,B} u(1,0)\right| \\
&\le \mathbf{c} t_0^{-i-|I|/2} \|u\|_{C^{i+|I|}(\{3/4 \le |s| \le 3/2, |z| \le \mathbf{c}'\})} \\
&\le \mathbf{c}(T,i,|I|)\ t_0^{-i-|I|/2} \frac{e^{-\beta d(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|}.
\end{aligned}$$

ii) $4t_0 \ge d(x_0,y_0)^2$. With the same change of variables, we see that

$$\left(\partial_s - L_A^B\right) u(s,z) = 0 \text{ if } \frac{1}{2} \le s \le 2, |z| \le 1.$$

Note that in this case, conditions $d(x,x_0) < \sqrt{t_0}$ and $4t_0 \ge d(x_0,y_0)^2$ allow x to reach y_0; nevertheless, condition $\frac{1}{2} \le s \le 2$ keeps t far off 0, so $h_A(t,x,y_0)$ is a solution in the corresponding region. By Corollary 6.18, if $\frac{1}{2} \leq s \leq 2, |z| \leq 1$

$$\begin{aligned}
|u(s,z)| &\leq \mathbf{c}(T) \frac{e^{-\beta d(x,y_0)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|} \\
&\leq \mathbf{c}(T) \frac{e^{-\beta d(x,y_0)^2/t}}{\left|B\left(x,\sqrt{t_0}\right)\right|} \leq \mathbf{c}(T) \frac{e^{-\beta d(x,y_0)^2/t}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|}
\end{aligned}$$

because $d(x,x_0) < \sqrt{t_0}$ implies $\left|B\left(x,\sqrt{t_0}\right)\right| \simeq \left|B\left(x_0,\sqrt{t_0}\right)\right|$

$$\leq \mathbf{c}(T) \frac{1}{\left|B\left(x_0,\sqrt{t_0}\right)\right|} \leq \mathbf{c}(T) \frac{e^{-\beta' d(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|}$$

because $4t_0 \geq d(x_0,y_0)^2$. Hence the same argument of case i) can be repeated, to conclude the proof. The estimates of the derivatives in y follow as in the proof of Theorem 3, p.852, [30]: we will present this technique in detail in the proof of next Theorem 9.1. ■

9 Uniform upper bounds on the difference of the fundamental solutions of two operators

In this section we want to prove the following

Theorem 9.1 *For $A, B \in \mathcal{B}_\lambda$, let h_A, h_B be fundamental solutions of H_A, H_B, respectively. Then, for any nonnegative integer i, couple of multiindices I, J, and $T > 0$, we have*

$$\begin{aligned}
&\left|\left(\partial_t^i X_I^x X_J^y h_A - \partial_t^i X_I^x X_J^y h_B\right)(t,x,y)\right| \leq \\
&\leq \mathbf{c}(i,I,J,T)\,\|A-B\|\, t^{-i-(|I|+|J|)/2} \frac{e^{-\mathbf{c}' d(x,y)^2/t}}{\left|B\left(x,\sqrt{t}\right)\right|}
\end{aligned} \tag{9.1}$$

for every $x, y \in \mathbb{R}^n, t \in (0,T)$.

Proof. It is enough to prove (9.1) when A, B differ for a single coefficient, and then iterate; so, assume:

$$H_A - H_B = (\beta - \alpha)(XY + YX)$$

where X, Y are any two of the vector fields $X_1, ..., X_m$. Due to our assumptions on the X_i's and the structure of the matrix $A \in \mathcal{B}_\lambda$ (see the Hypotheses in Section 3), the vector fields X, Y will vanish outside a compact set. Also, we will write $h_\alpha, h_\beta, H_\alpha, H_\beta$ for h_A, h_B, H_A, H_B respectively.

The proof proceeds in three steps.

Step 1. We will prove first (9.1) when $i + |I| + |J| = 0$. We start noting that:

$$H_\alpha(h_\alpha - h_\beta) = (H_\beta - H_\alpha)h_\beta = (\alpha - \beta)(XYh_\beta + YXh_\beta) \tag{9.2}$$

where all the fundamental solutions are written as $h(t,x,w)$, with pole at the point $(0,w)$. Let T_1 be the distribution $(XYh_\beta)(0,\cdot,w)$, i.e.

$$\begin{aligned}
\langle T_1,\varphi\rangle &= \int_{\mathbb{R}^n}\int_0^{+\infty} h_\beta(t,x,w)\,Y^*X^*\varphi(t,x)\,dtdx \\
&= \int_{\mathbb{R}^n}\int_0^{+\infty} Y_xh_\beta(t,x,w)\,X^*\varphi(t,x)\,dtdx \\
&= \lim_{\varepsilon\to0}\int_{\mathbb{R}^n}\int_\varepsilon^{+\infty} Y_xh_\beta(t,x,w)\,X^*\varphi(t,x)\,dtdx \\
&= \lim_{\varepsilon\to0}\int_{\mathbb{R}^n}\int_\varepsilon^{+\infty} X_xY_xh_\beta(t,x,w)\,\varphi(t,x)\,dtdx.
\end{aligned}$$

Let

$$\begin{aligned}
u_1(\tau,y) &= \lim_{\varepsilon\to0}\int_{\mathbb{R}^n}\int_\varepsilon^{+\infty} X_xY_xh_\beta(t,x,w)\,h_\alpha(\tau-t,y,x)\,dtdx \\
&= \int_{\mathbb{R}^n}\int_0^{+\infty} Y_xh_\beta(t,x,w)\,X_x^*h_\alpha(\tau-t,y,x)\,dtdx
\end{aligned}$$

We want to show that

$$H_\alpha u_1 = T_1,$$

i.e.

$$\langle u_1, H_\alpha^*\varphi\rangle = \langle T_1,\varphi\rangle.$$

We have

$$\begin{aligned}
&\int_{\mathbb{R}^n}\int_{-\infty}^{+\infty} u_1(\tau,y)\,H_\alpha^*\varphi(\tau,y)\,dyd\tau \\
&= \int_{\mathbb{R}^n}\int_{-\infty}^{+\infty}\left(\int_{\mathbb{R}^n}\int_0^{+\infty} Y_xh_\beta(t,x,w)\,X_x^*h_\alpha(\tau-t,y,x)\,dtdx\right)H_\alpha^*\varphi(\tau,y)\,dyd\tau \\
&= \lim_{\varepsilon\to0}\int_{\mathbb{R}^n}\int_{-\infty}^{+\infty}\left(\int_{\mathbb{R}^n}\int_\varepsilon^{+\infty} Y_xh_\beta(t,x,w)\,X_x^*h_\alpha(\tau-t,y,x)\,dtdx\right)H_\alpha^*\varphi(\tau,y)\,dyd\tau \\
&= \lim_{\varepsilon\to0}\int_{\mathbb{R}^n}\int_{-\infty}^{+\infty}\left(\int_{\mathbb{R}^n}\int_\varepsilon^{+\infty} X_xY_xh_\beta(t,x,w)\,h_\alpha(\tau-t,y,x)\,dtdx\right)H_\alpha^*\varphi(\tau,y)\,dyd\tau \\
&= \lim_{\varepsilon\to0}\int_{\mathbb{R}^n}\int_\varepsilon^{+\infty} X_xY_xh_\beta(t,x,w)\left(\int_{\mathbb{R}^n}\int_{-\infty}^{+\infty} h_\alpha(\tau-t,y,x)\,H_\alpha^*\varphi(\tau,y)\,dyd\tau\right)dtdx \\
&= \lim_{\varepsilon\to0}\int_{\mathbb{R}^n}\int_\varepsilon^{+\infty} X_xY_xh_\beta(t,x,w)\,\varphi(t,x)\,dtdx \\
&= \langle T_1,\varphi\rangle
\end{aligned}$$

where we used the fact that

$$\int_{\mathbb{R}^n}\int_{-\infty}^{+\infty} h_\alpha(\tau-t,y,x)\,H_\alpha^*\varphi(\tau,y)\,dyd\tau = \varphi(t,x)$$

since $H_\alpha^{(\tau,y)} h_\alpha(\tau - t, y, x) = \delta_{(t,x)}(\tau, y)$. Analogously, if we define

$$u(\tau, y) = \int_{\mathbb{R}^n} \int_0^{+\infty} Y_x h_\beta(t, x, w) X_x^* h_\alpha(\tau - t, y, x)\, dt dx + \\ + \int_{\mathbb{R}^n} \int_0^{+\infty} X_x h_\beta(t, x, w) Y_x^* h_\alpha(\tau - t, y, x)\, dt dx$$

we find that $H_\alpha u = (XYh_\beta + YXh_\beta) = H_\alpha \left(\frac{h_\alpha - h_\beta}{\alpha - \beta}\right)$, so that $v = u - \left(\frac{h_\alpha - h_\beta}{\alpha - \beta}\right)$ satisfies the equation $H_\alpha v = 0$ in the whole space, in distributional sense. Next, we recall that, for any fixed α, β,

$$\left|\frac{h_\alpha - h_\beta}{\alpha - \beta}(t, x, y)\right| \leq \mathbf{c}(\alpha, \beta, T)\, h(\mathbf{c}_1 t, x, y)$$

for any $x, y \in \mathbb{R}^n$ and $t \in (0, T)$ by Corollary 6.18 and Remark 7.4. Here h stands for the fundamental solution of the operator H_A where A =identity matrix. We are going to show that u satisfies a similar bound, *uniformly in* α, β, for any $x, y \in \mathbb{R}^n$ and $t \in (0, T)$. This implies that:

i) v vanishes at infinity and therefore is identically zero, by the maximum principle (Proposition 3.6), so $u = (h_\alpha - h_\beta) / (\beta - \alpha)$;

ii)

$$\left|\frac{h_\alpha - h_\beta}{\beta - \alpha}(t, x, y)\right| = |u(t, x, y)| \leq \mathbf{c}(T)\, h(\mathbf{c}_1 t, x, y)$$

and finally

$$\begin{aligned} |(h_\alpha - h_\beta)(t, x, y)| &\leq \mathbf{c}(T)\, |\beta - \alpha|\, h(\mathbf{c} t, x, y) \\ &\leq \mathbf{c}(T)\, |\beta - \alpha| \frac{e^{-\mathbf{c}' d(x,y)^2/t}}{\left|B\left(x, \sqrt{t}\right)\right|} \end{aligned} \tag{9.3}$$

for any $x, y \in \mathbb{R}^n$ and $t \in (0, T)$, again by Corollary 6.18. So, let us prove the bound for u_1; analogously one can prove the bound for the whole u. Recall that

$$u_1(\tau, y) = \int_{\mathbb{R}^n} \int_0^{\tau} Y_x h_\beta(t, x, w) X_x^* h_\alpha(\tau - t, y, x)\, dt dx.$$

By Theorem 8.1 and Remark 7.4

$$|Y_x h_\beta(t, x, w)| \leq \mathbf{c}(T)\, t^{-1/2} h(\mathbf{c}_1 t, x, w) \text{ for } x, w \in \mathbb{R}^n, 0 < t < T.$$

Also,

$$X_x^* h_\alpha(t, y, x) = -X_x h_\alpha(t, y, x) + c(x)\, h_\alpha(t, y, x)$$

hence

$$|X_x^* h_\alpha(t, y, x)| \leq \mathbf{c}(T)\, t^{-1/2} \frac{e^{-\mathbf{c}' d(x,y)^2/t}}{\left|B\left(y, \sqrt{t}\right)\right|} \leq \mathbf{c}(T)\, t^{-1/2} h(\mathbf{c}_1 t, y, x)$$

for $x, y \in \mathbb{R}^n, 0 < t < T$. This implies that

$$\begin{aligned}
|u_1(\tau, y)| &\leq \int_{\mathbb{R}^n} \int_0^{+\infty} |Y_x h_\beta(t, x, w) X_x^* h_\alpha(\tau - t, y, x)| \, dt dx \leq \\
&\leq \mathbf{c}(T) \int_{\mathbb{R}^N} \int_0^{+\infty} t^{-1/2} h(\mathbf{c}_1 t, x, w) (\tau - t)^{-1/2} h(\mathbf{c}_1 (\tau - t), y, x) \, dt dx = \\
&= \mathbf{c}(T) \int_0^\tau t^{-1/2} (\tau - t)^{-1/2} dt \cdot h(\mathbf{c}_1 \tau, y, w) = \mathbf{c}(T) h(\mathbf{c}_1 \tau, y, w)
\end{aligned}$$

where we used the reproduction property (3.4). This ends the proof of (9.3).

Step 2. Our next goal is to extend this result to:

$$\left| \left(\partial_t^i X_I^x h_\alpha - \partial_t^i X_I^x h_\beta \right)(t, x, y) \right| \leq \mathbf{c}(T, i, |I|) |\beta - \alpha| \, t^{-i-|I|/2} h(\mathbf{c}_1 t, x, y). \quad (9.4)$$

Here we follow the same technique of the proof of Theorem 8.1.

So, let $4t_0 \leq d(x_0, y_0)^2$, and make the change of variables $t = t_0 s, x = \Phi_B(z)$ with $B = B\left(x_0; \sqrt{t_0}\right)$. Let

$$u(s, z) = (h_\alpha - h_\beta)(t, x, y_0).$$

By (9.2), we know that, for $1/2 < s < 2, |z| \leq 1$,

$$\left(\partial_s - L_A^B\right) u(s, z) = t_0 (\alpha - \beta) (XY h_\beta + YX h_\beta)(t, x, y_0).$$

Also, by (9.3), in the same region we have

$$\begin{aligned}
|u(s, z)| &\leq \mathbf{c}(T) |\beta - \alpha| \frac{e^{-\mathbf{c} d(x, y_0)^2/t}}{\left|B\left(x, \sqrt{t}\right)\right|} \\
&\leq \mathbf{c}(T) |\beta - \alpha| \frac{e^{-\mathbf{c} d(x_0, y_0)^2/t_0}}{\left|B\left(x_0, \sqrt{t_0}\right)\right|}
\end{aligned}$$

where the last inequality is proved as in the proof of Theorem 8.1, case i). On the other hand,

$$\begin{aligned}
\left| \partial_t^i X_I^x (h_\alpha - h_\beta)(t_0, x_0, y_0) \right| &\leq t_0^{-i-|I|/2} \left| \partial_s^i Z_{I,B} u(1, 0) \right| \leq \\
&\leq t_0^{-i-|I|/2} \|u\|_{C^{i+|I|}\left(\frac{3}{4} < |s| < \frac{3}{2}, |z| < 1\right)}
\end{aligned}$$

for σ large enough

$$\leq \mathbf{c} t_0^{-i-|I|/2} \|\varphi_1 u\|_{H^\sigma}$$

by the uniform subelliptic estimate (4.8), choosing the same cutoff functions as above

$$\leq \mathbf{c} t_0^{-i-|I|/2} \left\{ \mathbf{c}(T) |\beta - \alpha| \frac{e^{-\mathbf{c} d(x_0, y_0)^2/t_0}}{\left|B\left(x_0, \sqrt{t_0}\right)\right|} + \left\| \varphi_2 \left(\partial_s - L_A^B\right) u \right\|_{H^\sigma} \right\}.$$

But:

$$\varphi_2\left(\partial_s - L_A^B\right)u(s,z) = \varphi_2 t_0(\alpha-\beta)(XYh_\beta + YXh_\beta)(t_0 s, \Phi_B(z), y_0).$$

Recalling that, by Lemma 8.2, ii),

$$\left(\Phi_B^{-1}\right)_*(X_j) = t_0^{-1/2} Z_{j,B}$$

we have:

$$\begin{aligned}
&\left|Z_{j,B}\left[t_0(XYh_\beta + YXh_\beta)(t_0 s, \Phi_B(z), y_0)\right]\right| \\
&= \left|t_0^{3/2}(X_j XYh_\beta + X_j YXh_\beta)(t_0 s, \Phi_B(z), y_0)\right| \\
&\leq \mathbf{c}(T)\frac{e^{-\mathbf{c}d(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|}
\end{aligned}$$

by Theorem 8.1. Iterating this argument, we find that

$$\begin{aligned}
&\left|\partial_t^p Z_{j_1,B} Z_{j_2,B} \ldots Z_{j_k,B}\left[t_0(XYh_\beta + YXh_\beta)(t_0 s, \Phi_B(z), y_0)\right]\right| \\
&\leq \mathbf{c}(T,p,k)\frac{e^{-\mathbf{c}d(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|}.
\end{aligned}$$

Expressing standard derivatives in terms of vector fields, one has

$$\left\|\varphi_2\left(\partial_s - L_A^B\right)u\right\|_{H^\sigma} \leq \mathbf{c}(T,\sigma)\frac{e^{-\mathbf{c}d(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|}$$

and finally (recalling that σ only depends on $i, |I|$)

$$\left|\partial_t^i X_I^x(h_\alpha - h_\beta)(t_0, x_0, y_0)\right| \leq \mathbf{c}(T, i, |I|)\, t_0^{-i-|I|/2}\left|\alpha-\beta\right|\frac{e^{-\mathbf{c}d(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|}.$$

The case $4t_0 \geq d(x_0,y_0)^2$ can be handled similarly, adapting the proof of Theorem 8.1, second case.

Step 3. Finally, we come to the estimates on y-derivatives. We start noting that, if $X_1, X_2, \ldots, X_m$ are general Hörmander vector fields, $X_j = \sum_{i=1}^n b_{ji}(x)\,\partial_{x_i}$, then

$$X_j^* = -X_j + a_j \text{ with } a_j = -\sum_{i=1}^n \partial_{x_i}(b_{ji})$$

With a slight abuse of notation, here we will write:

$$H_\alpha^* = \partial_t - L_\alpha^*$$

Since the matrix A is symmetric, the fundamental solution of H_α^* is $h^*(t,x,y) = h(t,y,x)$, and

$$\begin{aligned}
&H_\alpha^*\left(h_\alpha^* - h_\beta^*\right)(t,x,y) = \left(H_\beta^* - H_\alpha^*\right)h_\beta^* = \left(L_\alpha^* - L_\beta^*\right)h_\beta^* = \\
&= (\alpha-\beta)\left\{\left(XYh_\beta^* + YXh_\beta^*\right) + aXh_\beta^* + bYh_\beta^* + ch_\beta^*\right\}(t,x,y)
\end{aligned}$$

for suitable smooth functions a, b, c depending only on the vector fields X_i; the fields X, Y and the operator H_α^* act on the x variable. Differentiating the above identity with respect to $\partial_t^i X_J^y$ (which commutes with X, Y, H_α^*), we write:

$$\begin{aligned}
&H_\alpha^*\left(\partial_t^i X_J^y h_\alpha^* - \partial_t^i X_J^y h_\beta^*\right)(t,x,y) \\
&= (\alpha-\beta)\left\{\left(\partial_t^i XYX_J^y h_\beta^* + \partial_t^i YXX_J^y h_\beta^*\right) + \right. \\
&\left. + a\partial_t^i XX_J^y h_\beta^* + b\partial_t^i YX_J^y h_\beta^* + c\partial_t^i X_J^y h_\beta^*(t,x,y)\right\}.
\end{aligned}$$

We only treat the case $4t_0 \leq d(x_0,y_0)^2$, the other case being similar. With the same change of variables as above, let

$$u(s,z) = \left(\partial_t^i X_J^y h_\alpha^* - \partial_t^i X_J^y h_\beta^*\right)(t,x,y_0).$$

Step 2 of the proof shows that

$$|u(s,z)| \leq \mathbf{c}(T,i,|J|)\, t_0^{-i-|J|/2}\, |\beta-\alpha| \frac{e^{-\mathbf{c}d(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|}.$$

On the other hand

$$\begin{aligned}
\left|\left(\partial_t^i X_I^x X_J^y h_\alpha^* - \partial_t^i X_I^x X_J^y h_\beta^*\right)(t_0,x_0,y_0)\right| &\leq t_0^{-|I|/2}\left|Z_{I,B}u(1,0)\right| \leq \\
&\leq \mathbf{c}t_0^{-|I|/2}\|u\|_{C^{|I|}\left(\frac{3}{4}<|s|<\frac{3}{2},|z|<1\right)} \\
&\leq \mathbf{c}t_0^{-|I|/2}\|\varphi_1 u\|_{H^\sigma}
\end{aligned}$$

$$\leq \mathbf{c}t_0^{-|I|/2}\left\{\mathbf{c}(T,i,|J|)\, t_0^{-i-|J|/2}\,|\beta-\alpha|\frac{e^{-cd(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|} + \left\|\varphi_2\left(\partial_s - (L_\alpha^*)^B\right)u\right\|_{H^\sigma}\right\} \tag{9.5}$$

where we have applied uniform subelliptic estimates (for some σ only depending on $|I|$) to the operator $(L_A^*)^B$, that is point (iv) of Lemma 8.2. Again Lemma 8.2, and the structure of the operator L_α^* imply:

$$\begin{aligned}
&\varphi_2\left(\partial_s - (L_\alpha^*)^B\right)u(s,z) = \\
&= (\alpha-\beta)\left\{\varphi_2\left(t_0\partial_t^i XYX_J^y h_\beta^* + t_0\partial_t^i YXX_J^y h_\beta^* + at_0^{1/2}\partial_t^i XX_J^y h_\beta^*\right.\right. \\
&\left.\left.+bt_0^{1/2}\partial_t^i YX_J^y h_\beta^* + c\partial_t^i X_J^y h_\beta^*\right)\right\}(t_0 s, \Phi_B(z), y_0)
\end{aligned}$$

This, by the same argument of the first part of this proof, implies that

$$\left\|\varphi_2\left(\partial_s - (L_\alpha^*)^B\right)u\right\|_{H^\sigma} \leq \mathbf{c}(|I|)\, t_0^{-i-|J|/2}\,|\beta-\alpha|\frac{e^{-\mathbf{c}d(x_0,y_0)^2/t_0}}{\left|B\left(x_0,\sqrt{t_0}\right)\right|}$$

which inserted in (9.5) gives the desired final result. ■

Part II

Fundamental solution for operators with Hölder continuous coefficients

10 Assumptions, main results and overview of Part II

In this part, we will deal with variable-coefficient complete operators, of the kind:

$$H = \partial_t - \sum_{i,j=1}^{m} a_{i,j}(t,x) X_i X_j - \sum_{k=1}^{m} a_k(t,x) X_k - a_0(t,x). \tag{10.1}$$

To exploit the results proved in Part I, we will make the same assumptions stated in Section 3 on the vector fields and the structure of the matrix of the coefficients in the principal part. Moreover, the coefficients a_{ij}, a_k, a_0 will be assumed globally defined and Hölder continuous with respect to the parabolic CC-distance d_P; the matrix $\{a_{ij}\}_{i,j=1}^{m}$ will be assumed symmetric and uniformly positive definite. Under these assumptions (which we will state more precisely in a while) we will prove the existence of a (global) fundamental solution for H, satisfying natural basic properties and sharp Gaussian bounds. A precise list of our results is contained in Theorem 10.7 here below. Before stating it, we need to introduce some precise definitions, notation and assumptions we will use throughout Parts II and III.

Assumptions on the vector fields

We will assume that:

$X = (X_1, X_2, ..., X_m)$ $(m = n + q)$ is a fixed system of Hörmander's vector fields defined in the whole $\mathbb{R}^n$, and such that

$$X = (0, 0, ..., 0, \partial_{x_1}, \partial_{x_2}, ..., \partial_{x_n}) \text{ in } \mathbb{R}^n \setminus \Omega_0 \tag{10.2}$$

where Ω_0 is a fixed bounded domain.

Function spaces

We start with the following

Definition 10.1 *The intrinsic-derivative along the vector field X_j of a function $v(x)$ at a point $x_0 \in \mathbb{R}^n$, is defined to be*

$$X_j v(x_0) = \left.\frac{d}{d\sigma}\right|_{\sigma=0} v(\gamma(\sigma))$$

(if such derivative exists), where γ is the solution to

$$\dot{\gamma}(\sigma) = X_j(\gamma(\sigma)), \qquad \gamma(0) = x_0.$$

Remark 10.2 *The reader could ask why, only at this point of the paper, we need to give a precise definition of the derivative $X_j v$. This fact is related to a deep difference which exists between Part I and Part II: in Part I we have studied the differential operator H_A with smooth coefficients (being the a_{ij}'s constant), which is hypoelliptic, hence the functions involved in our estimates were always C^∞, and the meaning of the derivative $X_j v$ was obvious. In contrast with this, in this part we will study a differential operator H with Hölder continuous coefficients, and will build a fundamental solution for it, which we cannot expect to be smooth: instead, its degree of regularity will be the object of a careful study.*

We can now introduce some function spaces which will be useful in the following.

Definition 10.3 *Let $U \subseteq \mathbb{R}^{n+1}$ be an open set. We denote by $\mathfrak{C}^2(U)$ the class of functions $u(t,x)$ defined on U which are continuous in U w.r.t. the pair (t,x) and such that $u(t,\cdot)$ has continuous intrinsic-derivatives up to second order along the vector fields $X_1, \dots, X_m$ (w.r.t. x, for every fixed t) and $u(\cdot,x)$ has continuous derivative (w.r.t. t, for every fixed x), in their respective domains of definition.*

We will denote by d the Carnot-Carathéodory distance induced by the system $\{X_i\}_{i=1}^m$ in the whole $\mathbb{R}^n$ and by $B(x,r)$ the balls in the metric d. Moreover d_P will be the corresponding "parabolic-CC-distance". These distances have been introduced and studied in Sections 2 and 5. We can introduce "parabolic CC-Hölder spaces" related to d_P.

Definition 10.4 *For any $\alpha \in (0,1]$ and domain $U \subseteq \mathbb{R}^{n+1}$, let:*

$$|u|_{C^\alpha(U)} = \sup\left\{\frac{|u(t,x) - u(s,y)|}{d_P((t,x),(s,y))^\alpha} : (t,x),(s,y) \in U, (t,x) \neq (s,y)\right\}$$

$$\|u\|_{C^\alpha(U)} = |u|_{C^\alpha(U)} + \|u\|_{L^\infty(U)}$$

$$C^\alpha(U) = \left\{u : U \to \mathbb{R} : \|u\|_{C^\alpha(U)} < \infty\right\}.$$

Also, for any positive integer k, and domain $U \subseteq \mathbb{R}^{n+1}$, let

$$C^{k,\alpha}(U) = \left\{u : U \to \mathbb{R} : \|u\|_{C^{k,\alpha}(U)} < \infty\right\}$$

with

$$\|u\|_{C^{k,\alpha}(U)} = \sum_{|I|+2h\leqslant k} \left\|\partial_t^h X^I u\right\|_{C^\alpha(U)}$$

where, for any multiindex $I = (i_1, i_2, ..., i_s)$, *with* $1 \leqslant i_j \leqslant q$, *we say that* $|I| = s$ *and*

$$X^I u = X_{i_1} X_{i_2} ... X_{i_s} u.$$

We explicitly remark that in the definition of $C^{k,\alpha}$ *we are assuming that the derivatives of* u *involved exist as intrinsic derivatives.*

Remark 10.5 (Continuity and compactness properties of Hölder functions) *Note that, by (2.1), a function* $u \in C^\alpha(U)$ *is also continuous in Euclidean sense. For the same reason, all the derivatives* $\partial_t^h X^I u$ *involved in the definition of* $C^{k,\alpha}$ *are continuous in Euclidean sense. This fact has also the following consequence: if* U *is any bounded domain of* $\mathbb{R}^{n+1}$ *and* $f \in C^\alpha(U)$, *then* f *can be continuously extended up to the boundary of* U, *preserving* $C^\alpha(U)$ *norm. Therefore,* f *can be thought as belonging to* $C^\alpha\left(\overline{U}\right)$. *This fact will be implicitly used is some compactness arguments. For instance, it allows to apply Ascoli-Arzelà's theorem to a bounded sequence of functions in* $C^\alpha(U)$.

Assumptions on the coefficients

Throughout Parts II and III we will assume that:

the matrix $\{a_{ij}\}_{i,j=1}^m$ has the following structure:

$$A = \{a_{ij}\}_{i,j=1}^m = \begin{bmatrix} \{a_{ij}\}_{i,j=1}^q & 0 \\ 0 & I_n \end{bmatrix}; \tag{10.3}$$

the functions $a_{ij} = a_{ji},\ a_k,\ a_0,$ are defined on $\mathbb{R}^{n+1}$ and satisfy, for some $\alpha \in (0,1]$ and for some positive constants λ, K,

$$\begin{gathered} \lambda^{-1}|w|^2 \leq \sum_{i,j=1}^q a_{ij}(t,x)\, w_i w_j \leq \lambda |w|^2 \ \ \forall w \in \mathbb{R}^q, (t,x) \in \mathbb{R}^{n+1} \\ \|a_{ij}\|_{C^\alpha(\mathbb{R}^{n+1})} + \|a_k\|_{C^\alpha(\mathbb{R}^{n+1})} + \|a_0\|_{C^\alpha(\mathbb{R}^{n+1})} \leq K. \end{gathered} \tag{10.4}$$

Notation 10.6 *Here we collect some notation which will be used extensively throughout Part II and III.*

We shall denote by $\mathbf{c}$ *any positive constant only depending on* $X_1, \ldots, X_m$ *and the parameters* λ, K, α *appearing in (10.4). Moreover, we will write* $\mathbf{c}(f_1, \ldots, f_p)$ *if* $\mathbf{c}$ *also depends on* $f_1, \ldots, f_p$.

The points of $\mathbb{R}^{n+1} = \mathbb{R} \times \mathbb{R}^n$ *will be denoted by*

$$z = (t,x)\,;\ \zeta = (\tau, \xi);\ \eta = (s,y).$$

For the sake of brevity, we shall use the notation

$$\mathbf{E}(x,\xi,t) = |B(x,\sqrt{t})|^{-1} \exp\left(-\frac{d(x,\xi)^2}{t}\right), \quad x, \xi \in \mathbb{R}^n,\ t > 0.$$

Main results

We can now state the main results we will prove in Part II.

Theorem 10.7 (Fundamental solution for H) *Let H be as in (10.1). Under the above assumptions, there exists a function*

$$h : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}$$

such that:

i) h is continuous away from the diagonal of $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$;

ii) $h(z, \zeta)$ is nonnegative, and vanishes for $t \le \tau$;

iii) for every fixed $\zeta \in \mathbb{R}^{n+1}$, we have

$$h(\cdot;\zeta) \in C^{2,\alpha}_{loc}(\mathbb{R}^{n+1} \setminus \{\zeta\}), \quad H(h(\cdot;\zeta)) = 0 \text{ in } \mathbb{R}^{n+1} \setminus \{\zeta\};$$

iv) the following estimates hold for every $T > 0, z = (t,x), \zeta = (\tau, \xi) \in \mathbb{R}^{n+1}$, $0 < t - \tau \le T$:

$$\mathbf{c}(T)^{-1}\mathbf{E}(x,\xi,\mathbf{c}^{-1}(t-\tau)) \le h(z;\zeta) \le \mathbf{c}(T)\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)),$$
$$|X_j(h(\cdot;\zeta))(z)| \le \mathbf{c}(T)\,(t-\tau)^{-1/2}\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau));$$
$$|X_iX_j(h(\cdot;\zeta))(z)| + |\partial_t(h(\cdot;\zeta))(z)| \le \mathbf{c}(T)\,(t-\tau)^{-1}\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau));$$

v) for any $f \in C^{\alpha}(\mathbb{R}^{n+1})$, $g \in C(\mathbb{R}^n)$, both satisfying suitable growth condition at infinity (see Theorem 12.1 for an exact statement), $T \in \mathbb{R}$, the function

$$u(t,x) = \int_{\mathbb{R}^n} h(t,x;T,\xi)\,g(\xi)\,d\xi + \int_{[T,t]\times\mathbb{R}^n} h(t,x;\tau,\xi)\,f(\tau,\xi)d\tau\,d\xi$$

is a $C^{2,\alpha}_{loc}$ solution to the following Cauchy problem

$$\begin{cases} Hu = f & \text{in } (T,\infty)\times\mathbb{R}^n, \\ u(T,\cdot) = g & \text{in } \mathbb{R}^n \end{cases}$$

vi) the following reproduction formula holds

$$h(t,x;\tau,\xi) = \int_{\mathbb{R}^n} h(t,x;s,y)\,h(s,y;\tau,\xi)dy,$$

for $t > s > \tau$ and $x, \xi \in \mathbb{R}^n$.

Remark 10.8 *By Lemma 2.6, we know that*

$$\mathbf{E}(x,\xi,t) \le \mathbf{c}\mathbf{E}(\xi,x,\mathbf{c}t)$$

for any $x,\xi \in \mathbb{R}^n, t > 0$. Hence the Gaussian bounds in point (iv) of the above Theorem are not so asymmetric in x,ξ as it could seem.

The above theorem collects several facts which will be proved throughout the following sections of Part II, namely: Proposition 11.4 (Gaussian bound from above for h), Proposition 11.7 (continuity of h outside the diagonal), Theorem 11.8 (upper bounds on the derivatives of h; h is a solution to $Hu = 0$ outside the pole); Theorem 12.1 (solution to the Cauchy problem); Proposition 13.3 (positivity of h); Proposition 13.4 (reproduction property for the fundamental solution); Theorem 13.6 (Gaussian bound from below for h); Theorem 14.4 ($C^{2,\alpha}_{loc}$-regularity of h outside the pole).

The statement of point (v) in this Theorem has been simplified with respect to the sharper result proved in Theorem 12.1.

Remark 10.9 *Our assumptions (10.2), (10.3), as well as the fact that both the vectors fields and the coefficients are defined on the whole space, are just made to have a convenient setting to prove the existence of a global fundamental solution, but are not really restrictive. Namely, assume we have an operator*

$$H_{loc} = \partial_t - \sum_{i,j=1}^{q} a_{i,j}(t,x) X_i X_j - \sum_{k=1}^{q} a_k(t,x) X_k - a_0(t,x)$$

where $X_1, X_2, ..., X_q$ is a system of Hörmander's vector fields defined in a bounded domain Ω of $\mathbb{R}^n$, the coefficients a_{ij}, a_k, a_0 are defined and Hölder continuous in some domain $U \subset \mathbb{R} \times \Omega$, and the matrix $\{a_{ij}\}_{i,j=1}^{q}$ satisfies

$$\lambda^{-1} |w|^2 \le \sum_{i,j=1}^{q} a_{ij}(t,x) w_i w_j \le \lambda |w|^2 \quad \forall w \in \mathbb{R}^q, (t,x) \in U.$$

Then, applying some results proved in Section 2, we will show in Section 19 how to define a new operator H, satisfying all the assumptions we have made above and such that, in some compact subdomain of U, H coincides with H_{loc}, and the CC-distances of the two operators are equivalent. This fact will allow to deduce local results for the operator H_{loc} from the results we have proved for our globally defined operator H (see Theorem 19.1).

Overview of Part II and relations with Part I

The proofs of all the results of Part II, that we have collected in Theorem 10.7, is strictly based on the achievements of Part I. Namely, let

$$H_{\zeta_0} \equiv \partial_t - L_{\zeta_0} \equiv \partial_t - \sum_{i,j=1}^{m} a_{i,j}(\zeta_0) X_i X_j \tag{10.5}$$

be the operator obtained from the "principal part" of H by freezing the coefficients a_{ij} (but not the vector fields X_j) at any point $\zeta_0 \in \mathbb{R}^{n+1}$. By the assumptions we have made above on H, the operator H_{ζ_0} fits the assumptions of the theory developed in Part I; let us denote by h_{ζ_0} its fundamental solution (with the notation of Part I, $h_{\zeta_0}(z,\zeta) = h_A(z,\zeta)$ where $A = (a_{ij}(\zeta_0))_{i,j=1}^{m}$).

For the reader's convenience, we now recall the properties of h_{ζ_0} that we have proved in Part I, and will play a crucial role in the following.

Theorem 10.10 (Properties of h_{ζ_0} proved in Part I) *The function h_{ζ_0} is smooth away from the diagonal of $\mathbb{R}^{n+1}\times\mathbb{R}^{n+1}$ and*

$$H_{\zeta_0}(h_{\zeta_0}(\cdot,\zeta))=0 \qquad \text{in } \mathbb{R}^{n+1}\setminus\{\zeta\}. \tag{10.6}$$

Moreover, for every $\zeta_0\in\mathbb{R}^{n+1}$ and $T>0$ we have

$$\int_{\mathbb{R}^n} h_{\zeta_0}(t,x,\tau,\xi)d\xi=1,\ \text{ if } t>\tau,\ x\in\mathbb{R}^n, \tag{10.7}$$

$$\int_{\mathbb{R}^n} h_{\zeta_0}(t,x,\tau,\xi)\,dx\le \mathbf{c}(T),\ \text{ if } 0<t-\tau<T,\ \xi\in\mathbb{R}^n, \tag{10.8}$$

$$h_{\zeta_0}(t,x,\tau,\xi)=\int_{\mathbb{R}^n} h_{\zeta_0}(t,x,s,y)h_{\zeta_0}(s,y,\tau,\xi)dy \text{ if } t>s>\tau,\ x,\xi\in\mathbb{R}^n. \tag{10.9}$$

The following uniform Gaussian bounds hold: for every $T>0$ and for every nonnegative integers p,q, we have

$$\begin{gathered}\mathbf{c}(T)^{-1}\mathbf{E}(x,\xi,\mathbf{c}^{-1}(t-\tau))\le h_{\zeta_0}(t,x,\tau,\xi)\le \mathbf{c}(T)\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)),\\ \left|X_{i_1}\cdots X_{i_p}(\partial_t)^q\,h_{\zeta_0}(\cdot,\tau,\xi)\,(t,x)\right|\le \mathbf{c}(T,p,q)(t-\tau)^{-(p+2q)/2}\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)),\end{gathered} \tag{10.10}$$

$$\begin{aligned}&\left|X_{i_1}\cdots X_{i_p}(\partial_t)^q\,(\,h_{\zeta_0}(\cdot,\tau,\xi)-h_{\zeta_1}(\cdot,\tau,\xi))\,(t,x)\right|\\ &\le \mathbf{c}(T,p,q)\,d_P(\zeta_0,\zeta_1)^{\alpha}\,(t-\tau)^{-(p+2q)/2}\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)),\end{aligned} \tag{10.11}$$

if $0<t-\tau<T$, $x,\xi\in\mathbb{R}^n$ and $\zeta_0=(\tau_0,\xi_0)$, $\zeta_1=(\tau_1,\xi_1)\in\mathbb{R}^{n+1}$.

We refer to Theorem 3.4, Corollary 3.7, Corollary 6.18, Theorem 7.1, Theorem 8.1 and Theorem 9.1, for the proof of the above statements (recall also (10.4)). We remark that (10.8) follows from (10.7) and (10.10), using (2.12).

The above results will be the key ingredient to prove the existence of a fundamental solution for the operator (10.1), using the Levi parametrix method, as will be explained in Section 11. Other tools will be the geometric properties of d proved in Section 2, plus some other miscellaneous properties which we collect here below.

Some auxiliary estimates

Proposition 10.11 **i)** *There exists* $\mathbf{c}$ *such that*

$$\mathbf{E}(x,\xi,t)\le \mathbf{c}\ \beta^{Q/2}\ \mathbf{E}(x,\xi,\beta t) \tag{10.12}$$

for every $\beta\ge 1$, $x,\xi\in\mathbb{R}^n$, $t>0$.

ii) *For any $\mu\ge 0$, there exists $\mathbf{c}\,(\mu)$ such that*

$$(d(x,\xi)^2/t)^{\mu}\ \mathbf{E}(x,\xi,\lambda t)\le \mathbf{c}(\mu)\ \lambda^{\mu}\ \mathbf{E}(x,\xi,2\lambda t) \tag{10.13}$$

for every $\lambda>0$, $x,\xi\in\mathbb{R}^n$, $t>0$.

iii) *For any $\varepsilon > 0$ and $\mu \geq 0$, there exists $\mathbf{c}\,(\mu, \varepsilon)$ such that*

$$t^{-\mu}\ \mathbf{E}(x, \xi, t) \leq \mathbf{c}(\varepsilon, \mu) \tag{10.14}$$

for every $x, \xi \in \mathbb{R}^n$, $t > 0$ such that $d(x, \xi)^2 + t \geq \varepsilon$.

iv) *There exists a positive constant $\delta = \mathbf{c}^{-1}$ such that, if $0 \leq \mu \leq \delta/T$, then*

$$\mathbf{E}(x, \xi, t)\ \exp(\mu|\xi|^2) \leq \mathbf{c}\ \mathbf{E}(x, \xi, 2t)\ \exp(2\mu|x|^2) \tag{10.15}$$

for every $x, \xi \in \mathbb{R}^n$ and $0 < t \leq T$.

Proof. In order to prove (10.12), it is sufficient to observe that

$$\left|B(x, \sqrt{t})\right| \geq \mathbf{c}^{-1}\beta^{-Q/2}\left|B(x, \sqrt{\beta t})\right|,$$

by (2.9). Let us prove (10.13). Since $\max_{s\in[0,\infty)} s^\mu \exp(-s/2) = \mathbf{c}(\mu) < \infty$, taking $s = d(x, \xi)^2/(\lambda t)$, we get

$$\begin{aligned}(d(x, \xi)^2/t)^\mu\ \mathbf{E}(x, \xi, \lambda t) &\leq \mathbf{c}(\mu)\ \lambda^\mu\ \exp(-s/2)|B(x, \sqrt{\lambda t})|^{-1} \\ &\leq \mathbf{c}(\mu)\ \lambda^\mu\ \mathbf{E}(x, \xi, 2\lambda t),\end{aligned}$$

by the doubling condition (2.8). We now turn to the proof of (10.14). If $t > \varepsilon/2$, we use (2.7) and get

$$t^{-\mu}\ \mathbf{E}(x, \xi, t) \leq \mathbf{c}(\varepsilon)t^{-\mu-n/2}\exp(-\varepsilon^2/t) \leq \mathbf{c}(\varepsilon, \mu).$$

If $0 < t \leq \varepsilon/2$, we use (2.11) and get

$$\begin{aligned}t^{-\mu}\ \mathbf{E}(x, \xi, t) &\leq \mathbf{c}(\varepsilon)t^{-\mu-Q/2}\exp(-d(x, \xi)^2/t) \\ &\leq \mathbf{c}(\varepsilon)t^{-\mu-Q/2}\exp(-\varepsilon/(2t)) \leq \mathbf{c}(\varepsilon, \mu).\end{aligned}$$

We finally prove (10.15). From (2.3), it follows that

$$\mu|\xi|^2 \leq 2\mu(|x|^2 + |x - \xi|^2) \leq 2\mu|x|^2 + \mathbf{c}_0\mu d(x, \xi)^2.$$

Hence, choosing $\delta = (2\mathbf{c}_0)^{-1}$, we have $\mathbf{c}_0\mu \leq (2t)^{-1}$ and, using (2.9), we obtain

$$\begin{aligned}\mathbf{E}(x, \xi, t)\ \exp(\mu|\xi|^2) &\leq \exp(2\mu|x|^2)|B(x, \sqrt{t})|^{-1}\exp((\mathbf{c}_0\mu - 1/t)d(x, \xi)^2) \\ &\leq \mathbf{c}\ \mathbf{E}(x, \xi, 2t)\ \exp(2\mu|x|^2).\end{aligned}$$

■

Corollary 10.12 *If $0 < t - \tau \leq T$ and $t > s > \tau$, then we have*

$$\int_{\mathbb{R}^n} \mathbf{E}(x, y, \mathbf{c}_1(t - s))\mathbf{E}(y, \xi, \mathbf{c}_2(s - \tau))dy \leq \mathbf{c}(T)\mathbf{E}(x, \xi, \mathbf{c}\,(t - \tau)), \tag{10.16}$$

for every $x, \xi \in \mathbb{R}^n$. Moreover

$$\int_{\mathbb{R}^n} \mathbf{E}(x, \xi, \mathbf{c}\,t)d\xi \leq \mathbf{c}(T), \tag{10.17}$$

if $0 < t \leq T$, $x \in \mathbb{R}^n$.

Proof. Fixed $\zeta_0 \in \mathbb{R}^{n+1}$ and setting $\mathbf{c}_3 = \max\{\mathbf{c}_1, \mathbf{c}_2\}$, from (10.12) and (10.10) we obtain

$$\mathbf{E}(x, y, \mathbf{c}_1(t-s)) \leq \mathbf{c}\,\mathbf{E}(x, y, \mathbf{c}_3(t-s)) \leq \mathbf{c}(T) h_{\zeta_0}(x, y, \mathbf{c}_4(t-s)),$$
$$\mathbf{E}(y, \xi, \mathbf{c}_2(s-\tau)) \leq \mathbf{c}\,\mathbf{E}(y, \xi, \mathbf{c}_3(s-\tau)) \leq \mathbf{c}(T) h_{\zeta_0}(y, \xi, \mathbf{c}_4(s-\tau)).$$

We can now use the reproduction property (10.9) for h_{ζ_0} and then the estimate from above in (10.10) to get (10.16). Finally (10.17) is an immediate consequence of (10.10) and (10.8). ■

We close this section pointing out the following mean value theorem on X-subunit paths, which will be useful in the sequel:

Lemma 10.13 *For every d-ball $B(x_0, r)$, $x \in B(x_0, r)$, any continuous function u such that $X_i u$ exists and is continuous in $\overline{B(x_0, r)}$ for $i = 1, 2, ..., m$, we have*

$$|u(x) - u(x_0)| \leq d(x, x_0) \max_{\overline{B(x_0,r)}} |Xu| \tag{10.18}$$

where $|Xu| = \left(\sum_{i=1}^m |X_i u|^2\right)^{1/2}$.

Proof. Fix $x \in B(x_0, r)$, and let $\varepsilon > 0$ be such that $(1+\varepsilon)\, d(x_0, x) < r$; let γ be a subunit curve joining x_0, x such that:

$$\gamma'(t) = \sum_{i=1}^m \lambda_i(t)\, X_i(\gamma(t));$$
$$\gamma(0) = x_0;\ \gamma(T) = x;\ T \leqslant (1+\varepsilon)\, d(x_0, x).$$

Observe that $\gamma \subset B(x_0, r)$: namely, for any $z \in \gamma$, let γ_z be the portion of γ which joins x_0 to z, $\gamma(T_z) = z$, then

$$d(x_0, z) \leqslant T_z \leqslant T \leqslant (1+\varepsilon)\, d(x_0, x) < r.$$

Now, assume for a moment that u is smooth. Then we have

$$u(x) - u(x_0) = u(\gamma(T)) - u(\gamma(0)) =$$
$$= \int_0^T \frac{d}{dt}(u(\gamma(t)))\, dt = \int_0^T \sum_{i=1}^m \lambda_i(t)\, (X_i u)(\gamma(t))\, dt.$$

hence

$$|u(y) - u(x)| \leqslant \sup_{z \in \overline{B(x_0,r)}} |Xu(z)| \cdot T \leqslant (1+\varepsilon)\, d(x_0, x) \cdot \sup_{z \in \overline{B(x_0,r)}} |Xu(z)|$$

and, for vanishing ε, we have the desired result. If u is not smooth, let u_{ε_n} be a sequence of smooth functions obtained from u with a standard (Euclidean) mollification procedure. In has been proved in [7, Proposition 2.2.] that

$$X_j u_{\varepsilon_n} \to X_j u$$

uniformly on compact subsets of U, and this is enough to get our assertion in the general case. ■

Plan of Part II

In Section 11 we build a fundamental solution h for H and prove Gaussian estimates from above for it; in Section 12 we show how the solution of a Cauchy problem can be represented by this fundamental solution; in Section 13 we prove a Gaussian estimate from below on h and a reproduction property for it. The fundamental solution constructed, so far, possesses only a partial regularity: space derivatives of h are continuous for fixed t, and *viceversa.* In Section 14 we prove some regularity results, which allow to show that h belongs to the space $C^{2,\alpha}$ in any region excluding the pole.

11 Fundamental solution for H: the Levi method

The "Levi method" is a classical technique that allows to construct the fundamental solution of a variable coefficient differential operator, starting from the fundamental solution of the corresponding operator with constant coefficients. This method was originally developed by E. E. Levi at the beginning of 20th century to study uniformly elliptic equations of order $2n$ (see [40], [41]) and later extended to uniformly parabolic operators (see e.g. [27]).

In the context of hypoelliptic ultraparabolic operators of Kolmogorov-Fokker-Planck type, Polidoro in [48] managed to adapt this method, thanks to the knowledge of an explicit expression for the fundamental solution of the "frozen" operator, which had been constructed in [38]. For operators of type (10.1), structured on homogeneous and invariant vector fields on Carnot groups, no explicit fundamental solution is available in general. Nevertheless, Bonfiglioli, Lanconelli, Uguzzoni in [4] showed how to adapt the same method, exploiting suitable sharp uniform Gaussian bounds on the fundamental solutions of the frozen operators.

Here we will follow the same line, thanks to the results of Part I. We start with a brief outline of the scheme of Levi method.

Let us consider the fundamental solution $h_{\zeta_0}(z,\zeta)$ of the "frozen" operator H_{ζ_0}; the function $z \mapsto h_\zeta(z,\zeta)$ is called *parametrix.* The idea of the Levi method is to look for a fundamental solution $h\,(z,\zeta)$ for H, which could be written in the form:

$$h\,(z,\zeta) = h_\zeta(z,\zeta) + \int_\tau^t \int_{\mathbb{R}^n} h_\eta(z,\eta)\Phi\,(\eta,\zeta)\,d\eta \tag{11.1}$$

for a suitable, unknown kernel $\Phi\,(z,\zeta)$. This seems reasonable because we expect h to be a small perturbation of the parametrix, as the integral equation (11.1) expresses. The following formal computation suggests how to guess the right form of $\Phi\,(z,\zeta)$. If we set

$$Z_1(z;\zeta) = -H\,(z \mapsto h_\zeta(z,\zeta))\,(z), \qquad z \neq \zeta \in \mathbb{R}^{n+1}$$

and apply the operator H to both sides of (11.1) for $z \neq \zeta$, we find:

$$0 = -Z_1(z;\zeta) + \Phi\,(z,\zeta) - \int_\tau^t \int_{\mathbb{R}^n} Z_1(z,\eta)\Phi\,(\eta,\zeta)\,d\eta.$$

This means that Φ solves the integral equation

$$Z_1(z,\zeta) = \Phi(z;\zeta) - \int_\tau^t \int_{\mathbb{R}^n} Z_1(z,\eta)\Phi(\eta,\zeta)\,d\eta$$

which, defining the integral operator T with kernel Z_1, can be rewritten as

$$Z_1 = (I-T)\,\Phi$$

whence, formally,

$$\Phi = \sum_{k=0}^{\infty} T^k Z_1 \equiv \sum_{k=0}^{\infty} Z_k.$$

To make the above idea rigorous, we will have to reverse the order of the previous steps: we will start studying the properties of the function Z_1, then Z_k, then $\Phi = \sum_{k=0}^{\infty} Z_k$, then

$$J(z,\zeta) = \int_\tau^t \int_{\mathbb{R}^n} h_\eta(z,\eta)\Phi(\eta,\zeta)\,d\eta$$

and finally

$$h(z,\zeta) = h_\zeta(z,\zeta) + J(z,\zeta).$$

Let us now start this hard job.

Definition 11.1 *We set*

$$Z_1(z;\zeta) = -H\left(z \mapsto h_\zeta(z,\zeta)\right)(z), \qquad z \neq \zeta \in \mathbb{R}^{n+1} \tag{11.2}$$

and, for every $j \in \mathbb{N}$,

$$Z_{j+1}(z;\zeta) = \int_{\mathbb{R}^n \times [\tau,t]} Z_1(z;\eta)\,Z_j(\eta;\zeta)\,d\eta, \quad z=(t,x),\ \zeta=(\tau,\xi) \in \mathbb{R}^{n+1},\ t>\tau.$$

Proposition 11.2 *For every $j \in \mathbb{N}$, $Z_j(z;\zeta)$ is well defined and satisfies the following estimate:*

$$|Z_j(z;\zeta)| \leq \mathbf{c}_1(T)^j\, b_j(\alpha)\,(t-\tau)^{-1+j\alpha/2}\, h_{\zeta_0}(x,\xi,\mathbf{c}_2\,(t-\tau)), \quad 0<t-\tau\leq T, \tag{11.3}$$

for any $\zeta_0 \in \mathbb{R}^{n+1}$, where

$$b_j(\alpha) = \Gamma^j\left(\frac{\alpha}{2}\right)/\Gamma\left(\frac{j\alpha}{2}\right)$$

(here Γ denotes the Euler Gamma function, and α is the Hölder exponent appearing in our assumption (10.4)). As a consequence, the series

$$\Phi(z;\zeta) = \sum_{j=1}^{\infty} Z_j(z;\zeta)$$

totally converges on the set

$$\{0 < t-\tau \le T,\, d_P(z,\zeta) \ge 1/T\}$$

(for every $T > 0$), and satisfies the estimate

$$|\Phi(z;\zeta)| \le \mathbf{c}(T)\,(t-\tau)^{\frac{\alpha}{2}-1}\,\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau)), \quad 0 < t-\tau \le T. \tag{11.4}$$

Also, we have

$$\Phi(z;\zeta) = Z_1(z;\zeta) + \int_{\mathbb{R}^n\times[\tau,t]} Z_1(z;\eta)\,\Phi(\eta;\zeta)d\eta, \tag{11.5}$$

for every $z=(t,x)$, $\zeta=(\tau,\xi)\in\mathbb{R}^{n+1}$, $t>\tau$.

Proof. Let us prove (11.3) by induction. For $j=1$, by (10.6), we have

$$\begin{aligned} Z_1(z;\zeta) &= \sum_{i,j=1}^{m} (a_{ij}(z)-a_{ij}(\zeta))\,X_iX_jh_\zeta(\cdot,\zeta)(z) \\ &\quad + \sum_{k=1}^{m} a_k(z)X_kh_\zeta(\cdot,\zeta)(z) + a_0(z)h_\zeta(z,\zeta). \end{aligned} \tag{11.6}$$

Moreover, by (10.4), (10.10) and (10.13), if $0 < t-\tau \le T$, we have

$$\begin{aligned} |Z_1(z;\zeta)| &\le \mathbf{c}(T)\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau))\left(d_P\,(z,\zeta)^\alpha\,(t-\tau)^{-1} + (t-\tau)^{-1/2}+1\right) \\ &\le \mathbf{c}(T)\,(t-\tau)^{\frac{\alpha}{2}-1}\,\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau)) \\ &\le \mathbf{c}(T)\,(t-\tau)^{\frac{\alpha}{2}-1}\,h_{\zeta_0}(x,\xi,\mathbf{c}\,(t-\tau)), \end{aligned} \tag{11.7}$$

for every $\zeta_0\in\mathbb{R}^{n+1}$. Assuming now that (11.3) holds for a given $j\in\mathbb{N}$, we get

$$\begin{aligned} &\int_{\mathbb{R}^n\times[\tau,t]} |Z_1(z;\eta)\,Z_j(\eta;\zeta)|\,d\eta \le \mathbf{c}_1(T)^{j+1}\,b_j(\alpha)\cdot \\ &\cdot\int_\tau^t (t-s)^{-1+\frac{\alpha}{2}}(s-\tau)^{-1+\frac{j\alpha}{2}}\int_{\mathbb{R}^n} h_{\zeta_0}(x,y,\mathbf{c}_2(t-s))\,h_{\zeta_0}(y,\xi,\mathbf{c}_2(s-\tau))dyds \\ &= \mathbf{c}_1(T)^{j+1}\,b_j(\alpha)\,h_{\zeta_0}(x,\xi,\mathbf{c}_2\,(t-\tau))(t-\tau)^{-1+\frac{(j+1)\alpha}{2}}\int_0^1 (1-r)^{-1+\frac{\alpha}{2}}\,r^{-1+\frac{j\alpha}{2}}\,dr. \end{aligned}$$

In the last equality, we have used the reproduction property (10.9) of h_{ζ_0}. Recalling the definition of $b_j(\alpha)$, we obtain (11.3) for $j+1$. Observing now that the power series

$$\sum_{j=1}^{\infty} b_j(\alpha)\,w^j$$

has infinite radius of convergence we get, using again (10.10),

$$\begin{aligned}|\Phi(z;\zeta)| &\leq \sum_{j=1}^{\infty} |Z_j(z;\zeta)| \leq \\ &\leq (t-\tau)^{\frac{\alpha}{2}-1}\, h_{\zeta_0}(x,\xi,\mathbf{c}\,(t-\tau)) \sum_{j=1}^{\infty} b_j(\alpha)\mathbf{c}_1(T)^j T^{(j-1)\alpha/2} \leq \\ &\leq \mathbf{c}(T)(t-\tau)^{\frac{\alpha}{2}-1}\,\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau)).\end{aligned}$$

Finally, to prove (11.5), it is enough to observe that the above computations also show that the series

$$\sum_{j=1}^{\infty} \int_{\mathbb{R}^n\times[\tau,t]} |Z_1(z;\eta)\, Z_j(\eta;\zeta)| d\eta$$

is convergent. Hence

$$\begin{aligned}\int_{\mathbb{R}^n\times[\tau,t]} Z_1(z;\eta)\,\Phi(\eta;\zeta)d\eta &= \sum_{j=1}^{\infty} \int_{\mathbb{R}^n\times[\tau,t]} Z_1(z;\eta)\, Z_j(\eta;\zeta)d\eta \\ &= \sum_{j=1}^{\infty} Z_{j+1}(z;\zeta) = \Phi(z;\zeta) - Z_1(z;\zeta).\end{aligned}$$

This ends the proof. ■

Definition 11.3 *For every* $z=(t,x)\,,\ \zeta=(\tau,\xi)\in\mathbb{R}^{n+1}$, $t>\tau$, *we set*

$$J(z;\zeta) = \int_{\mathbb{R}^n\times[\tau,t]} h_\eta(z;\eta)\,\Phi(\eta;\zeta)d\eta, \tag{11.8}$$

and

$$h(z;\zeta) = J(z;\zeta) + h_\zeta(z;\zeta). \tag{11.9}$$

We also agree to extend $h(z;\zeta)$ *to be zero for* $t\leq\tau$.

Proposition 11.4 *The integral in (11.8) is convergent and the following estimate holds, for* $0<t-\tau\leq T$*:*

$$|J(z;\zeta)| \leq \mathbf{c}(T)\,(t-\tau)^{\frac{\alpha}{2}}\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)). \tag{11.10}$$

Moreover, the function h *satisfies*

$$|h(z;\zeta)| \leq \mathbf{c}(T)\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)), \tag{11.11}$$

for every $z=(t,x)\,,\ \zeta=(\tau,\xi)\in\mathbb{R}^{n+1}$, $0<t-\tau\leq T$.

Proof. By means of (10.10), (11.4) and (10.16), we have (for $0 < t-\tau \le T$)

$$\begin{aligned}
&\int_{\mathbb{R}^n \times [\tau,t]} |h_\eta(z;\eta)\,\Phi(\eta;\zeta)|\,\eta \le \\
&\quad \le \mathbf{c}(T)\int_\tau^t (s-\tau)^{-1+\frac{\alpha}{2}} \int_{\mathbb{R}^n} \mathbf{E}(x,y,\mathbf{c}_1(t-s))\mathbf{E}(y,\xi,\mathbf{c}_2(s-\tau))dyds \quad (11.12)\\
&\le \mathbf{c}(T)\,(t-\tau)^{\frac{\alpha}{2}}\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)).
\end{aligned}$$

This proves the first part of the Proposition. As to h, (11.11) immediately follows from (10.10) and (11.10). ■

To show that the function h actually satisfies the desired properties of the fundamental solution, we have now to investigate some regularity properties of the functions Z_j, Φ and J that we have defined.

Lemma 11.5 *For every $x, x', \xi \in \mathbb{R}^n$ and $0 < t-\tau \le T$, we have*

$$\begin{aligned}
&|Z_1(t,x;\tau,\xi) - Z_1(t,x';\tau,\xi)| \\
&\le \mathbf{c}(T)\,d(x,x')^{\frac{\alpha}{2}}\,(t-\tau)^{\frac{\alpha}{4}-1}\,\left(\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau)) + \mathbf{E}(x',\xi,\mathbf{c}\,(t-\tau))\right).
\end{aligned}$$

Proof. If $d(x,x') \ge \sqrt{t-\tau}$, it is sufficient to use (11.7). We then suppose that $d(x,x') < \sqrt{t-\tau}$. >From (10.4), (10.10) and (11.6), it follows that ($\zeta = (\tau,\xi), z = (t,x)\,, z' = (t,x')$)

$$\begin{aligned}
Z_1(z;\zeta) - Z_1(z';\zeta) &= \sum_{i,j=1}^m \left(a_{i,j}\,(z) - a_{i,j}(z')\right) X_iX_jh_\zeta(z',\zeta) \\
&+ \sum_{i,j=1}^m \left(a_{i,j}\,(z) - a_{i,j}(\zeta)\right)\,\left(X_iX_jh_\zeta(z,\zeta) - X_iX_jh_\zeta(z',\zeta)\right) \\
&+ \sum_{k=1}^m (a_k\,(z) - a_k(z'))\,X_kh_\zeta(z',\zeta) + \sum_{k=1}^m a_k\,(z)\,\left(X_kh_\zeta(z,\zeta) - X_kh_\zeta(z',\zeta)\right) \\
&+ (a_0\,(z) - a_0(z'))\,h_\zeta(z',\zeta) + a_0\,(z)\,\left(h_\zeta(z,\zeta) - h_\zeta(z',\zeta)\right).
\end{aligned}$$

Hence,

$$\begin{aligned}
|Z_1(z;\zeta) - Z_1(z';\zeta)| &\le \mathbf{c}(T)\,d(x,x')^\alpha\,(t-\tau)^{-1}\,\mathbf{E}(x',\xi,\mathbf{c}\,(t-\tau)) \\
&+ \mathbf{c}\,d_P(z,\zeta)^\alpha \sum_{i,j=1}^m |X_iX_jh_\zeta(z,\zeta) - X_iX_jh_\zeta(z',\zeta)| \\
&+ \mathbf{c}(T)\,d(x,x')^\alpha\,(t-\tau)^{-1/2}\,\mathbf{E}(x',\xi,\mathbf{c}\,(t-\tau)) \\
&+ \mathbf{c}\sum_{k=1}^m |X_kh_\zeta(z,\zeta) - X_kh_\zeta(z',\zeta)| \\
&+ \mathbf{c}(T)\,d(x,x')^\alpha\,\mathbf{E}(x',\xi,\mathbf{c}\,(t-\tau)) + \mathbf{c}\,|h_\zeta(z,\zeta) - h_\zeta(z',\zeta)|\,.
\end{aligned}$$

We now use the mean value inequality proved in Lemma 10.13. Recalling also (10.10), we have

$$\sum_{i,j=1}^{m} |X_iX_jh_\zeta(z,\zeta) - X_iX_jh_\zeta(z',\zeta)|$$
$$\leq d(x,x') \sum_{i,j=1}^{m} \max_{B(x,2d(x,x'))} |X\, X_iX_jh_\zeta(t,\cdot,\zeta)|$$
$$\leq \mathbf{c}(T)d(x,x')(t-\tau)^{-3/2} \sup_{B(x,2d(x,x'))} \mathbf{E}(\cdot,\xi,\mathbf{c}\,(t-\tau)).$$

Moreover, recalling (2.9) and using the fact that $d(x,x') < \sqrt{t-\tau}$, it is not difficult to see that the above supremum is lower than $\mathbf{c}\,\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau))$. Analogous estimates can be made for $\sum_{k=1}^{m} |X_kh_\zeta(z,\zeta) - X_kh_\zeta(t,x',\tau,\xi)|$ and $|h_\zeta(z,\zeta) - h_\zeta(t,x',\tau,\xi)|$. Using again that $d(x,x') < \sqrt{t-\tau}$, we finally get

$$|Z_1(t,x;\zeta) - Z_1(t,x';\zeta)| \leq \mathbf{c}(T)\, d(x,x')^\alpha\, (t-\tau)^{-1}\, \mathbf{E}(x',\xi,\mathbf{c}\,(t-\tau))$$
$$+\, \mathbf{c}\, d_P(z,\zeta)^\alpha d(x,x')(t-\tau)^{-3/2}\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau))$$
$$+\, \mathbf{c}(T)\, d(x,x')^\alpha\, (t-\tau)^{-1/2}\, \mathbf{E}(x',\xi,\mathbf{c}\,(t-\tau))$$
$$+\, \mathbf{c}(T)d(x,x')(t-\tau)^{-1/2}\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau))$$
$$+\, \mathbf{c}(T)\, d(x,x')^\alpha\, \mathbf{E}(x',\xi,\mathbf{c}\,(t-\tau)$$
$$+\, \mathbf{c}(T)d(x,x')\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau))$$
$$\leq \mathbf{c}(T)\, d(x,x')^{\frac{\alpha}{2}}\, (t-\tau)^{\frac{\alpha}{4}-1}\, \left[\mathbf{E}(x',\xi,\mathbf{c}\,(t-\tau)) + \mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau))\right].$$

■

Proposition 11.6 *Let $T > 0$. We have*

$$|\Phi(t,x;\tau,\xi) - \Phi(t,x';\tau,\xi)| \tag{11.13}$$
$$\leq \mathbf{c}(T)\, d(x,x')^{\frac{\alpha}{2}}\, (t-\tau)^{\frac{\alpha}{4}-1}\, \left(\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau)) + \mathbf{E}(x',\xi,\mathbf{c}\,(t-\tau))\right),$$

for every $x,x',\xi \in \mathbb{R}^n$, $0 < t-\tau \leq T$. Moreover, $\Phi(\cdot;\zeta)$ and $\Phi(z;\cdot)$ are continuous functions in their domains of definition.

Proof. >From (11.5), (11.4), Lemma 11.5 and (10.16), it follows that

$$|\Phi(t,x;\tau,\xi) - \Phi(t,x';\tau,\xi)|$$
$$\leq |Z_1(t,x;\tau,\xi) - Z_1(t,x';\tau,\xi)| +$$
$$+ \int_{\mathbb{R}^n\times[\tau,t]} |\Phi(s,y;\tau,\xi)|\, |Z_1(t,x;s,y) - Z_1(t,x';s,y)|\, dyds$$
$$\leq |Z_1(t,x;\tau,\xi) - Z_1(t,x';\tau,\xi)| + \mathbf{c}(T)d(x,x')^{\frac{\alpha}{2}} \int_\tau^t (t-s)^{-1+\frac{\alpha}{4}} (s-\tau)^{-1+\frac{\alpha}{2}} \cdot$$
$$\cdot \int_{\mathbb{R}^n} \mathbf{E}(y,\xi,\mathbf{c}(s-\tau))\, [\mathbf{E}(x,y,\mathbf{c}'\,(t-s)) + \mathbf{E}(x',y,\mathbf{c}''\,(t-s))]\, dyds$$
$$\leq \mathbf{c}(T)d(x,x')^{\frac{\alpha}{2}}\, (t-\tau)^{\frac{\alpha}{4}-1}\, \left[\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau)) + \mathbf{E}(x',\xi,\mathbf{c}\,(t-\tau))\right],$$

which proves (11.13).

We now turn to the proof of the last statement of Proposition 11.6.

Since (11.13) and (10.14) hold, in order to prove the continuity of $\Phi(\cdot;\zeta)$, we only have to see that $\Phi(\cdot,x;\zeta)$ is continuous. This will follow from the continuity of $Z_j(\cdot,x;\zeta)$, $j\in\mathbb{N}$. For $j=1$, such continuity follows immediately from (10.4) and the definition of Z_1. We now fix $j\in\mathbb{N}$ and prove that $Z_{j+1}(\cdot,x;\zeta)$ is continuous at $t_0>\tau$. We have

$$\begin{aligned}
&|Z_{j+1}(t,x;\zeta)-Z_{j+1}(t_0,x;\zeta)|\\
&\le\int_\tau^{t_0-\delta}\int_{\mathbb{R}^n}|Z_j(s,y;\zeta)|\,|Z_1(t_0,x;s,y)-Z_1(t,x;s,y)|dyds\\
&+\int_{t_0-\delta}^{t}\int_{\mathbb{R}^n}|Z_j(s,y;\zeta)\,Z_1(t,x;s,y)|dyds\\
&+\int_{t_0-\delta}^{t_0}\int_{\mathbb{R}^n}|Z_j(s,y;\zeta)\,Z_1(t_0,x;s,y)|dyds.
\end{aligned}$$

By using (11.3), (10.9), (10.10) and (10.14), it is easy to see that the last two integrals in the right-hand side are small if $|t-t_0|<\delta$ and δ is small enough. On the other hand, from (10.4), (10.10), (10.14) and (11.6) it follows that, for $|t-t_0|<\delta/2$ and for every $s\in(\tau,t_0-\delta)$, we have

$$\begin{aligned}
&|Z_1(t_0,x;s,y)-Z_1(t,x;s,y)|\\
&\le\sum_{i,j=1}^m\Big[|a_{ij}(t_0,x)-a_{ij}(t,x)|\left|X_iX_jh_{(s,y)}(t_0,x,s,y)\right|+\\
&+|a_{ij}(t,x)-a_{ij}(s,y)|\left|X_iX_jh_{(s,y)}(t_0,x,s,y)-X_iX_jh_{(s,y)}(t,x,s,y)\right|\Big]\\
&+\sum_{k=1}^m\Big[|a_k(t_0,x)-a_k(t,x)|\left|X_kh_{(s,y)}(t_0,x,s,y)\right|\\
&+|a_k(t,x)|\left|X_kh_{(s,y)}(t_0,x,s,y)-X_kh_{(s,y)}(t,x,s,y)\right|\Big]\\
&+|a_0(t_0,x)-a_0(t,x)|\left|h_{(s,y)}(t_0,x,s,y)\right|\\
&+|a_0(t,x)|\left|h_{(s,y)}(t_0,x,s,y)-h_{(s,y)}(t,x,s,y)\right|\\
&\le\mathbf{c}(t_0,\tau,\delta)\left(|t-t_0|^{\frac{\alpha}{2}}+|t-t_0|+\sum_{k=1}^m|a_k(t_0,x)-a_k(t,x)|+|a_0(t_0,x)-a_0(t,x)|\right)\\
&=\mathbf{c}(t_0,\tau,\delta)\varepsilon_{t_0,x}(t),
\end{aligned}$$

where $\varepsilon_{t_0,x}(t)$ vanishes as $t\to t_0$, recalling that a_k,a_0 are continuous. In the last inequality we have used (10.10) and (10.14) in order to get

$$\begin{aligned}
&|X_iX_jh_{(s,y)}(t_0,x,s,y)-X_iX_jh_{(s,y)}(t,x,s,y)|\\
&\le|t-t_0|\sup_{t^*\in(t_0,t)}\left|\partial_tX_iX_jh_{(s,y)}(t^*,x,s,y)\right|\\
&\le\mathbf{c}(t_0,\tau)\,|t-t_0|(t^*-s)^{-2}\mathbf{E}(x,y,\mathbf{c}(t^*-s))\le\mathbf{c}(t_0,\tau,\delta)|t-t_0|
\end{aligned}$$

and to find analogous estimates for $|X_k h_{(s,y)}(t_0,x,s,y) - X_k h_{(s,y)}(t,x,s,y)|$ and $|h_{(s,y)}(t_0,x,s,y) - h_{(s,y)}(t,x,s,y)|$. As a consequence, using also (10.8) and (11.3), for $|t-t_0| < \delta/2$ we obtain

$$\begin{aligned}
&\int_\tau^{t_0-\delta} \int_{\mathbb{R}^n} |Z_j(s,y;\zeta)|\, |Z_1(t_0,x;s,y) - Z_1(t,x;s,y)| dyds \\
&\leq \mathbf{c}(t_0,\tau,\delta,j)\, h_{t_0,x}(t) \int_\tau^{t_0-\delta} (s-\tau)^{-1+\frac{j\alpha}{2}} \left(\int_{\mathbb{R}^n} h_0(y,\xi,\mathbf{c}(s-\tau)) dy \right) ds \\
&\leq \mathbf{c}(t_0,\tau,\delta,j)\, h_{t_0,x}(t) \int_\tau^{t_0-\delta} (s-\tau)^{-1+\frac{j\alpha}{2}}\, s \leq \mathbf{c}(t_0,\tau,\delta,j)\, h_{t_0,x}(t).
\end{aligned}$$

In this way the continuity at t_0 of $Z_{j+1}(\cdot,x;\zeta)$ is proved.

We are only left to prove that $\Phi(z;\cdot)$ is continuous. To this end, it is sufficient to show the continuity of $Z_j(z;\cdot)$ for every $j \in \mathbb{N}$. For $j=1$, it is easy to see that $Z_1(z;\cdot)$ is continuous by using (11.6), (10.4), (10.11), and (10.14). One can then prove the continuity of $Z_j(z;\cdot)$ by induction, showing that $Z_{j+1}(z;\cdot)$ is a uniform limit, as $\sigma \to 0^+$, of the continuous functions

$$\zeta = (\tau,\xi) \mapsto \int_{\mathbb{R}^n \times [\tau+\sigma,t]} Z_1(z;\eta)\, Z_j(\eta;\zeta) d\eta,$$

on the compact subsets of $\mathbb{R}^n \times (-\infty,t)$ (by using the estimates (11.3), (10.10), (10.14) and (10.7), see also the proof of Proposition 11.2). ∎

Proposition 11.7 *The function h is continuous away from the diagonal of $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$.*

Proof. The function $h(z;\zeta)$ is obviously continuous in the set $\{t=\tau,\, x \neq \xi\}$, since, when $x_0 \neq \xi_0$, (11.11) and (10.14) give

$$\mathbf{E}(x,\xi,\mathbf{c}\,(t-\tau)) \leq \mathbf{c}(x_0,\xi_0)(t-\tau) \to 0$$

as $(t,x,\tau,\xi) \to (x_0,t_0,\xi_0,t_0)$, $t>\tau$. Therefore, we only have to prove that h is continuous in the set $\{t>\tau\}$. Let us show that the function $(z;\zeta) \mapsto h_\zeta(z;\zeta)$ is continuous away from the diagonal $\{z=\zeta\}$. Namely,

$$h_\zeta(z;\zeta) - h_{\zeta_1}(z_1;\zeta_1) = [h_\zeta(z;\zeta) - h_{\zeta_1}(z;\zeta)] + [h_{\zeta_1}(z;\zeta) - h_{\zeta_1}(z_1;\zeta_1)]\,.$$

The first term vanishes as $z \to z_1$ by (10.11) and (10.14). The second term vanishes as $(z;\zeta) \to (z_1;\zeta_1)$ since $h_{\zeta_1}(\cdot;\cdot)$ is smooth outside the diagonal (recall that $t>\tau$).

To complete the proof, it is enough to show that J is continuous in $\{t>\tau\}$. This can be done by showing that the functions

$$J_\sigma(z;\zeta) = \int_{\mathbb{R}^n \times [\tau+\sigma,t-\sigma]} h_\eta(z;\eta)\, \Phi(\eta;\zeta) d\eta$$

are continuous and converge uniformly to J, as $\sigma \to 0^+$, on the compact subsets K of $\{t>\tau\}$. The continuity of J_σ follows from the continuity of h_η and of

$\Phi(\eta;\cdot)$ (see Proposition 11.6), by dominated convergence, observing that, as $(z,\zeta) \to (z_0,\zeta_0) \in K$,

$$\chi_{(\tau+\sigma,t-\sigma)}(s)|h_\eta(z;\eta)\,\Phi(\eta;\zeta)| \leq \mathbf{c}(K,\sigma)\exp(-\mathbf{c}(K)^{-1}d(x,y)^2)\chi_{(\tau_0,t_0)}(s)$$
$$\leq \mathbf{c}(K,\sigma)\exp(-\mathbf{c}(K)^{-1}|y|^2)\chi_{(\tau_0,t_0)}(s) \in L^1_y(\mathbb{R}^n)$$

(χ_I denotes the characteristic function of I). Here we have used (10.10), (11.4), (10.14), (2.11) and (2.3). On the other hand, for every $K \Subset \{t > \tau\}$, arguing as in (11.12) and using (10.14), we obtain

$$\begin{aligned}\sup_K |J_\sigma - J| &\leq \sup_{(z,\zeta)\in K} \left(\int_\tau^{\tau+\sigma} + \int_{t-\sigma}^t\right)\int_{\mathbb{R}^n} |h_\eta(z;\eta)\,\Phi(\eta;\zeta)|\,d\eta \\ &\leq \mathbf{c}(K) \sup_{(z,\zeta)\in K} \mathbf{E}(x,\xi,\mathbf{c}(t-\tau))\left(\int_\tau^{\tau+\sigma} + \int_{t-\sigma}^t\right)(s-\tau)^{-1+\frac{\alpha}{2}}\,ds \\ &\leq \mathbf{c}(K)\,\sigma^{\alpha/2} \longrightarrow 0,\end{aligned}$$

as $\sigma \to 0^+$. ■

To prove that $Hh(\cdot,\zeta) = 0$ in $\mathbb{R}^{n+1}\setminus\{\zeta\}$, we now turn to the study of differentiability properties of the function h. Although our final goal is to show that $h(\cdot,\zeta) \in C^{2,\alpha}_{loc}(\mathbb{R}^{n+1}\setminus\{\zeta\})$, as a first step we shall show that $h(\cdot,\zeta)$ belongs to the larger function space $\mathfrak{C}^2(\mathbb{R}^{n+1}\setminus\{\zeta\})$, that we have introduced in Definition 10.3.

Theorem 11.8 *For every fixed $\zeta \in \mathbb{R}^{n+1}$, we have*

$$h(\cdot;\zeta) \in \mathfrak{C}^2(\mathbb{R}^{n+1}\setminus\{\zeta\}), \quad H(h(\cdot;\zeta)) = 0 \text{ in } \mathbb{R}^{n+1}\setminus\{\zeta\}. \tag{11.14}$$

Moreover, the following estimates hold for $0 < t-\tau \leq T$:

$$\begin{aligned}&|X_j(h(\cdot;\zeta))(z)| \leq \mathbf{c}(T)(t-\tau)^{-1/2}\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)); \qquad (11.15)\\ &|X_iX_j(h(\cdot;\zeta))(z)| + |\partial_t(h(\cdot;\zeta))(z)| \leq \mathbf{c}(T)(t-\tau)^{-1}\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)).\end{aligned}$$

In the proof of Theorem 11.8 we shall use the following simple fact:

Lemma 11.9 *Let $(u_i)_{i\in\mathbb{N}}$ be a sequence of continuous functions, defined on an open set $A \subseteq \mathbb{R}^n$, with continuous intrinsic-derivative along X_j. Suppose that u_i converges pointwise in A to some function u and that X_ju_i converges to some function w uniformly on the compact subsets of A, as $i \to \infty$. Then there exists the intrinsic-derivative of u along X_j, $X_ju(x) = w(x)$, for every $x \in A$.*

Proof. For any fixed $x \in A$, let $\gamma : [-T,T] \to A$, $\gamma(0) = x$ be an integral curve of X_j passing through x. Define the following functions of $t \in [-T,T]$:

$$f_i(t) = u_i(\gamma(t))\,;\, f(t) = u(\gamma(t))\,;\, g(t) = w(\gamma(t))\,.$$

By definition of intrinsic-derivative we also have:

$$f_i'(t) = (X_ju_i)(\gamma(t))\,;\, f'(t) = (X_ju)(\gamma(t))$$

and, by assumption, $f_i \to f$ pointwise and $f_i' \to g$ uniformly in $[-T, T]$. Then there exists $f'(0) = g(0)$, that is there exists

$$X_j u(x) = w(x).$$

■

The main step in the proof of Theorem 11.8 is the following lemma.

Lemma 11.10 *For every fixed* $\zeta = (\tau, \xi) \in \mathbb{R}^{n+1}$,

$$J(\cdot;\zeta) \in \mathfrak{C}^2(\{z = (t,x) \in \mathbb{R}^{n+1} : t > \tau\})$$

and we have

$$X_j (J(\cdot;\zeta))(z) = \int_{\mathbb{R}^n \times [\tau,t]} X_j h_\eta(z,\eta)\, \Phi(\eta;\zeta) d\eta, \tag{11.16}$$

$$X_i X_j (J(\cdot;\zeta))(z) = \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n \times [\tau, t-\varepsilon]} X_i X_j h_\eta(z,\eta)\, \Phi(\eta;\zeta) d\eta, \tag{11.17}$$

$$\partial_t (J(\cdot;\zeta))(z) = \Phi(z;\zeta) + \lim_{\varepsilon \to 0^+} \int_{\mathbb{R}^n \times [\tau, t-\varepsilon]} \partial_t h_\eta(z,\eta)\, \Phi(\eta;\zeta) d\eta. \tag{11.18}$$

Moreover, the following estimates hold for $0 < t - \tau \le T$*:*

$$|X_j (J(\cdot;\zeta))(z)| \le \mathbf{c}(T)\,(t-\tau)^{(\alpha-1)/2}\, \mathbf{E}(x,\xi,\mathbf{c}(t-\tau)), \tag{11.19}$$

$$|X_i X_j (J(\cdot;\zeta))(z)| + |\partial_t (J(\cdot;\zeta))(z)| \tag{11.20}$$

$$\le \mathbf{c}(T)\,(t-\tau)^{(\alpha-2)/2}\, \mathbf{E}(x,\xi,\mathbf{c}(t-\tau)).$$

Proof. The continuity of $J(\cdot;\zeta)$ has been proved in Proposition 11.4. In order to prove that there exist the intrinsic-derivatives in (11.16) and (11.17), we shall use Lemma 11.9. Let us set

$$J_\varepsilon(z;\zeta) = \int_{\mathbb{R}^n \times [\tau, t-\varepsilon]} h_\eta(z;\eta)\, \Phi(\eta;\zeta) d\eta, \tag{11.21}$$

so that J_ε pointwise converges to J, as $\varepsilon \to 0^+$. Making use of (10.10), (11.4), (10.14) and (10.8), it is not difficult to see that $J_\varepsilon(t,\cdot;\zeta)$ has continuous intrinsic-derivatives up to second order along the vector fields $X_1, \ldots, X_m$, obtained deriving (11.21) under the integral sign. In order to prove (11.16) it is then sufficient to show that

$$\sup_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n \times [t-\varepsilon, t]} |X_j h_\eta(z,\eta)\, \Phi(\eta;\zeta)| d\eta \longrightarrow 0, \quad \text{as } \varepsilon \to 0^+.$$

This is an easy consequence of the estimates (10.10), (11.4), (10.16) and (2.11). Indeed the above supremum is bounded by

$$\mathbf{c} \int_{t-\varepsilon}^{t} (t-s)^{-1/2}\,(s-\tau)^{-1+\alpha/2} \sup_{x \in \mathbb{R}^n} \int_{\mathbb{R}^n} \mathbf{E}(x,y,\mathbf{c}_1(t-s))\mathbf{E}(y,\xi,\mathbf{c}_2(s-\tau)) dy ds$$

$$\le \mathbf{c}(t,\tau)\varepsilon^{1/2} \sup_{x \in \mathbb{R}^n} \mathbf{E}(x,\xi,\mathbf{c}(t-\tau))$$

$$\le \mathbf{c}(t,\tau)\varepsilon^{1/2} \sup_{x \in \mathbb{R}^n} |B(x, \mathbf{c}\sqrt{t-\tau})|^{-1} \le \mathbf{c}(t,\tau)\varepsilon^{1/2}. \tag{11.22}$$

By Lemma 11.9, in order to prove (11.17), we now only have to show that there exists the limit in (11.17), uniformly in $x \in \mathbb{R}^n$. To this end, let us consider the integral

$$I = \int_{\mathbb{R}^n} X_i X_j h_\eta(t,x,s,y)\, \Phi(s,y;\tau,\xi) dy, \quad \tau < s < t.$$

Using (10.10), (11.4) and (10.16), it is easy to see that

$$|I| \le \mathbf{c}(T)\,(t-s)^{-1}\,(s-\tau)^{-1+\frac{\alpha}{2}}\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)). \tag{11.23}$$

Moreover, for every fixed $y_0 \in \mathbb{R}^n$, we have $I = I_1 + I_2 + I_3$, where

$$\begin{aligned}
I_1 &= \int_{\mathbb{R}^n} X_i X_j h_\eta(t,x,s,y)\,(\Phi(s,y;\zeta) - \Phi(s,y_0;\zeta)) dy,\\
I_2 &= \Phi(s,y_0;\zeta) \int_{\mathbb{R}^n} X_i X_j \left(h_{(s,y)} - h_{(s,y_0)}\right)(t,x,s,y) dy,\\
I_3 &= \Phi(s,y_0;\zeta) \int_{\mathbb{R}^n} X_i X_j h_{(s,y_0)}(t,x,s,y) dy.
\end{aligned}$$

Since

$$\int_{\mathbb{R}^n} h_{(s,y_0)}(t,x,s,y) dy = 1$$

by (10.7), differentiating the integral we obtain $I_3 \equiv 0$. This is possible in view of (10.10), (2.11) and (2.3), which ensure that

$$\begin{aligned}
&|X_j h_{(s,y_0)}(t,x,s,y)|, |X_i X_j h_{(s,y_0)}(t,x,s,y)|\\
&\quad \le \mathbf{c}(s,\tau)\mathbf{E}(x,y,\mathbf{c}(t-s)) \le \mathbf{c}(s,\tau)\exp(-\mathbf{c}(t,s)^{-1}|x-y|^2).
\end{aligned}$$

We now choose $y_0 = x$ and we estimate I_1, I_2.

Making use of (10.10), Proposition 11.6, (10.13), (10.16), and (10.17), for $0 < t - \tau \le T$ we obtain

$$\begin{aligned}
|I_1| &\le \mathbf{c}(T)\,(t-s)^{-1}(s-\tau)^{-1+\frac{\alpha}{4}} \int_{\mathbb{R}^n} d(x,y)^{\frac{\alpha}{2}}\,\mathbf{E}(x,y,\mathbf{c}(t-s))\cdot\\
&\quad \cdot (\mathbf{E}(y,\xi,\mathbf{c}(s-\tau)) + \mathbf{E}(x,\xi,\mathbf{c}(s-\tau)))\, dy\\
&\le \mathbf{c}(T)\,((t-s)(s-\tau))^{-1+\frac{\alpha}{4}} \Bigg(\mathbf{E}(x,\xi,\mathbf{c}(s-\tau)) \int_{\mathbb{R}^n} \mathbf{E}(x,y,\mathbf{c}(t-s)) dy\\
&\quad + \int_{\mathbb{R}^n} \mathbf{E}(y,\xi,\mathbf{c}(s-\tau))\,\mathbf{E}(x,y,\mathbf{c}(t-s)) dy\Bigg)\\
&\le \mathbf{c}(T)\,((t-s)(s-\tau))^{-1+\frac{\alpha}{4}}\,(\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)) + \mathbf{E}(x,\xi,\mathbf{c}(s-\tau)))\,.
\end{aligned} \tag{11.24}$$

Using (11.4), (10.10), (10.11), (10.13) and (10.17), we obtain

$$\begin{aligned}
|I_2| &\le \mathbf{c}(T)\,(s-\tau)^{-1+\frac{\alpha}{2}}\,\mathbf{E}(x,\xi,\mathbf{c}(s-\tau))\cdot\\
&\quad \cdot \int_{\mathbb{R}^n} d(x,y)^\alpha\,(t-s)^{-1}\,\mathbf{E}(x,y,\mathbf{c}(t-s))\, dy\\
&\quad \le \mathbf{c}(T)\,(t-s)^{-1+\frac{\alpha}{2}}\,(s-\tau)^{-1+\frac{\alpha}{2}}\,\mathbf{E}(x,\xi,\mathbf{c}(s-\tau)).
\end{aligned} \tag{11.25}$$

Collecting the above estimates and recalling (10.12) we finally obtain

$$|I| \leq \mathbf{c}(T) \left((t-\tau)(t-s)\right)^{-1+\frac{\alpha}{4}} \mathbf{E}(x, \xi, \mathbf{c}(t-\tau)), \tag{11.26}$$

if $0 < t - \tau \leq T$, $\frac{\tau+t}{2} < s < t$. Recalling also that

$$\sup_{x \in \mathbb{R}^n} \mathbf{E}(x, \xi, \mathbf{c}(t-\tau)) \leq \mathbf{c}(t, \tau)$$

(see (11.22)), it is now easy to recognize that the limit in (11.17) exists and it is uniform in $x \in \mathbb{R}^n$.
In order to conclude the proof of the first statement of Lemma 11.10, we are only left to prove that $J(\cdot, x; \zeta)$ has continuous derivative given by (11.18). To this end, it is sufficient to show that $J_\varepsilon(\cdot, x; \zeta)$ has continuous derivative, given by

$$\begin{aligned} \partial_t J_\varepsilon(t, x; \zeta) = & \int_{\mathbb{R}^n} h_{(t-\varepsilon, y)}(t, x, t-\varepsilon, y)\, \Phi(t-\varepsilon, y; \zeta) dy \\ & + \int_{\mathbb{R}^n \times [\tau, t-\varepsilon]} \partial_t h_{(s,y)}(t, x, s, y)\, \Phi(s, y; \zeta) dy ds \end{aligned} \tag{11.27}$$

and that

$$\sup_{t \in K} \left| \Phi(z; \zeta) - \int_{\mathbb{R}^n} h_{(t-\varepsilon, y)}(t, x, t-\varepsilon, y)\, \Phi(t-\varepsilon, y; \zeta) dy \right| \longrightarrow 0, \tag{11.28}$$

$$\sup_{t \in K} \int_{t-\varepsilon}^{t} \left| \int_{\mathbb{R}^n} \partial_t h_{(s,y)}(t, x, s, y)\, \Phi(s, y; \zeta) dy \right| ds \longrightarrow 0, \tag{11.29}$$

as $\varepsilon \to 0^+$, for every $K \subset\subset (\tau, \infty)$.

We have

$$\begin{aligned} & \tfrac{1}{h} \left(J_\varepsilon(t+h, x; \zeta) - J_\varepsilon(t, x; \zeta)\right) \\ & \quad = \int_{t-\varepsilon}^{t-\varepsilon+h} \tfrac{1}{h} \int_{\mathbb{R}^n} h_{(s,y)}(t+h, x, s, y)\, \Phi(s, y; \zeta) dy ds \\ & \quad + \int_{\mathbb{R}^n \times [\tau, t-\varepsilon]} \tfrac{1}{h} \left(h_\eta(t+h, x, \eta) - h_\eta(t, x, \eta)\right) \Phi(\eta; \zeta)\, d\eta. \end{aligned} \tag{11.30}$$

The second integral in (11.30) converges (as $h \to 0$) to the second integral in (11.27), by dominated convergence (use the mean value theorem and recall (10.10), (10.14), (11.4), (10.8)).

The first integral in (11.30) is equal to

$$\int_0^1 \int_{\mathbb{R}^n} h_{(t-\varepsilon+rh, y)}(t+h, x, t-\varepsilon+rh, y)\, \Phi(t-\varepsilon+rh, y; \zeta) dy dr,$$

which converges to the first integral in (11.27) (as $h \to 0$) by dominated convergence, as one can easily recognize, using (10.10), (10.12), (10.17), (11.4), (10.14), (10.11) and Proposition 11.6. This proves (11.27). Using the properties just recalled, it is also easy to see that $\partial_t J_\varepsilon(x,\cdot;\zeta)$ is continuous, again by dominated convergence. We now prove (11.28) and (11.29). Recalling (10.7), the supremum in (11.28) is bounded by $S_1+S_2+S_3$, where

$$S_1 = \sup_{t\in K}\int_{\mathbb{R}^n} \left|\left(h_{(t-\varepsilon,y)}(t,x,t-\varepsilon,y) - h_{(t-\varepsilon,x)}(t,x,t-\varepsilon,y)\right)\Phi(t-\varepsilon,y;\zeta)\right| dy,$$

$$S_2 = \sup_{t\in K}\int_{\mathbb{R}^n} h_{(t-\varepsilon,x)}(t,x,t-\varepsilon,y)\left|\Phi(t-\varepsilon,y;\zeta)-\Phi(t-\varepsilon,x;\zeta)\right| dy,$$

$$S_3 = \sup_{t\in K}\left|\Phi(t-\varepsilon,x;\zeta)-\Phi(t,x;\zeta)\right|.$$

>From the continuity of $\Phi(x,\cdot;\zeta)$ (see Proposition 11.6), we infer that $S_3\to 0$, as $\varepsilon\to 0^+$. Using (10.11), (11.4), (10.14), (10.13) and (10.17), we obtain

$$\begin{aligned} S_1 &\le \mathbf{c}\int_{\mathbb{R}^n} d(x,y)^\alpha\, \mathbf{E}(x,y,\mathbf{c}\,\varepsilon)\,(t-\varepsilon)^{-1+\alpha/2}\,\mathbf{E}(y,\xi,\mathbf{c}\,(t-\varepsilon-\tau))dy \\ &\le \mathbf{c}(K,\tau)\,\varepsilon^{\frac{\alpha}{2}}\int_{\mathbb{R}^n}\mathbf{E}(x,y,\mathbf{c}\,\varepsilon)dy \le \mathbf{c}(K,\tau)\,\varepsilon^{\frac{\alpha}{2}}. \end{aligned}$$

Hence $S_1\to 0$, as $\varepsilon\to 0^+$. In a similar way (exploiting (10.10) and Proposition 11.6), one can see that also S_2 vanishes as ε goes to zero. This proves (11.28). The proof of (11.29) closely follows the lines of the proof of (11.17) and therefore is omitted.

We finally turn to the proof of the second statement of Lemma 11.10. The estimate (11.19) can be obtained arguing as in (11.12). Moreover, from (11.23) and (11.26), it follows that

$$\begin{aligned} &\left|X_iX_j\left(J(\cdot;\zeta)\right)(z)\right| \\ &\le \mathbf{c}(T)\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau))\cdot \\ &\cdot\left(\int_\tau^{(t+\tau)/2}(t-s)^{-1}\,(s-\tau)^{-1+\frac{\alpha}{2}}ds + \int_{(t+\tau)/2}^t\left((t-s)(t-\tau)\right)^{-1+\frac{\alpha}{4}}ds\right) \\ &\le \mathbf{c}(T)\,(t-\tau)^{-1+\frac{\alpha}{2}}\,\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)). \end{aligned}$$

The estimate of $\partial_t J(x,\cdot;\zeta)$ is analogous (also recalling (11.4)). ■

Proof of Theorem 11.8. By means of Lemma 11.10 (and recalling (10.10) and Proposition 11.4), we only have to prove that $H(h(\cdot;\zeta))(z)=0$ for $t>\tau$. We explicitly remark that, by (10.14),

$$(t-\tau)^{-1}\mathbf{E}(x,\xi,\mathbf{c}(t-\tau)) \le \mathbf{c}(\delta)(t-\tau)$$

if $d(x,\xi)\ge\delta>0$, $t>\tau$: this allows to recognize that $h(\cdot,x,\tau,\xi)\in C^1(\mathbb{R})$ if

$x \neq \xi$. Making use of (11.8), (11.16), (11.17) and (11.18), we obtain

$$\begin{aligned} H\left(h(\cdot;\zeta)\right)(z) &= H\left(h_\zeta(\cdot;\zeta)\right)(z) \\ &\quad + \lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^n\times[\tau,t-\varepsilon]} H\left(h_\eta(\cdot;\eta)\right)(z)\,\Phi(\eta;\zeta)d\eta + \Phi(z;\zeta) \\ &= -Z_1(z;\zeta) - \int_{\mathbb{R}^n\times[\tau,t]} Z_1(z;\eta)\,\Phi(\eta;\zeta)d\eta + \Phi(z;\zeta) = 0, \end{aligned}$$

by means of (11.2) and (11.5). This completes the proof. ■

12 The Cauchy problem

The aim of this section is to prove the following theorem.

Theorem 12.1 *There exists a positive constant $\delta = \mathbf{c}^{-1}$ such that the following statement holds for every $\mu \geq 0$ and $T_2 > T_1$ satisfying*

$$(T_2 - T_1)\mu < \delta.$$

Given a continuous function $f(t,x)$ on $[T_1,T_2]\times\mathbb{R}^n$, locally d-Hölder continuous in x, uniformly w.r.t. t, and given a continuous function $g(x)$ on $\mathbb{R}^n$, satisfying the growth condition

$$|f(t,x)|\,,\ |g(x)| \leq M\ \exp(\mu\,|x|^2)$$

for some constant $M > 0$, then the function

$$u(t,x) = \int_{\mathbb{R}^n} h(t,x;\xi,T_1)\,g(\xi)\,d\xi + \int_{[T_1,t]\times\mathbb{R}^n} h(t,x;\tau,\xi)\,f(\tau,\xi)d\xi\,d\tau$$

belongs to the class

$$\mathfrak{C}^2((T_1,T_2)\times\mathbb{R}^n)\cap C([T_1,T_2]\times\mathbb{R}^n)$$

and is a solution to the following Cauchy problem

$$\begin{cases} Hu = f & \text{in } (T_1,T_2)\times\mathbb{R}^n, \\ u(T_1,\cdot) = g & \text{in } \mathbb{R}^n. \end{cases}$$

Remark 12.2 *As we will see in Theorem 14.4, if $f \in C^{\alpha}_{loc}$ (in the sense of Definition 10.4), then $u \in C^{2,\alpha}_{loc}$.*

We split the proof of Theorem 12.1 in some lemmas.

Lemma 12.3 *The statement of Theorem 12.1 holds if $f \equiv 0$.*

Proof. We only prove that $u\,(t,x) \to g(x_0)$, as $(t,x) \to (T_1, x_0)$, for every fixed $x_0 \in \mathbb{R}^n$. The other properties of u easily follow from Proposition 11.4 and Theorem 11.8 observing that, if $T_1 + \sigma \le t \le T_2$, $\sigma > 0$, then for any operator $\mathrm{D} \in \{X_j, X_iX_j, \partial_t, \mathrm{Id}\}$ we have

$$\begin{aligned}|\mathrm{D}h(\cdot; T_1, \xi)\,(t,x)\,g(\xi)| &\le \mathbf{c}(\sigma, g, \mu)\mathbf{E}(x, \xi, \mathbf{c}(t - T_1))\exp(\mu|\xi|^2)\\ &\le \mathbf{c}(\sigma, g, \mu)\exp(\mu(|\xi|^2 - (\mathbf{c}\,\delta)^{-1}d(x,\xi)^2))\\ &\le \mathbf{c}(\sigma, g, \mu)\exp(\mu(|\xi|^2 - 2|x - \xi|^2)),\end{aligned}$$

by (10.12), (2.11) and (2.3), if δ is chosen small enough.

Making use of (11.9) and (10.7), we can write

$$|u\,(t,x) - g(x_0)| \le I_1 + I_2 + I_3 + I_4 + I_5,$$

where

$$\begin{aligned}I_1 &= \int_{d(x_0,\xi)\ge\rho} |h(z; T_1, \xi)\,g(\xi)|d\xi,\\ I_2 &= \int_{d(x_0,\xi)<\rho} |J(z; T_1, \xi)\,g(\xi)|d\xi,\\ I_3 &= \int_{d(x_0,\xi)<\rho} \left|\left(h_{(T_1,\xi)}(z; T_1, \xi) - h_{(T_1,x_0,)}(z; T_1, \xi)\right) g(\xi)\right| d\xi,\\ I_4 &= \int_{d(x_0,\xi)<\rho} \left|h_{(T_1,x_0)}(z; T_1, \xi)\,(g(\xi) - g(x_0))\right| d\xi,\\ I_5 &= |g(x_0)| \int_{d(x_0,\xi)\ge\rho} \left|h_{(T_1,x_0)}(z; T_1, \xi)\right| d\xi\end{aligned}$$

and $0 < \rho < 1$, $d(x, x_0) < \rho/2$ and $0 < t - T_1 < 1$. From (11.10), (2.3) and (10.17) we get

$$I_2 \le \mathbf{c}(g, x_0)(t - T_1)^{\alpha/2} \int_{d(x_0,\xi)<\rho} \mathbf{E}(x, \xi, \mathbf{c}(t - T_1))\,d\xi \le \mathbf{c}(g, x_0)(t - T_1)^{\alpha/2},$$

which vanishes as $(t, x) \to (T_1, x_0)$. Moreover, if $\rho = \rho\,(\varepsilon, x_0, g)$ is small enough, we have

$$I_3 \le \mathbf{c}(g, x_0) \int_{d(x_0,\xi)<\rho} d(x_0, \xi)^\alpha \mathbf{E}(x, \xi, \mathbf{c}(t - T_1))\,d\xi \le \mathbf{c}(g, x_0)\rho^\alpha \le \varepsilon,$$

by means of (10.11), (2.3) and (10.17), and

$$I_4 \le \mathbf{c} \int_{d(x_0,\xi)<\rho} \mathbf{E}(x, \xi, \mathbf{c}(t - T_1))|g(\xi) - g(x_0)|d\xi \le \mathbf{c}\varepsilon,$$

by means of (10.10), and (10.17). Finally, using (11.11) and (10.10),

$$\begin{aligned}I_1 + I_5 &\le \mathbf{c}(g, \mu, x_0) \int_{d(x_0,\xi)\ge\rho} \mathbf{E}(x, \xi, \mathbf{c}(t - T_1))\exp(\mu|\xi|^2)d\xi\\ &\le \mathbf{c}(g, \mu, x_0, \rho) \int_{\mathbb{R}^n} \exp\left(\mu|\xi|^2 - (\mathbf{c}(t - T_1))^{-1}\,|x - \xi|^2\right) d\xi\end{aligned}$$

which vanishes as $(t,x) \to (T_1, x_0)$. In the last inequality we have used (10.13), (2.11), (2.3) and the fact that $d(x,x_0) \le \rho/2$, in order to estimate

$$\begin{aligned}
\mathbf{E}(x,\xi,\mathbf{c}(t-T_1)) &\le \mathbf{c}\left(\frac{t-T_1}{d(x,\xi)^2}\right)^{Q/2} \mathbf{E}(x,\xi,\mathbf{c}(t-T_1)) \\
&\le \mathbf{c}(\rho)(t-T_1)^{Q/2}\mathbf{E}(x,\xi,\mathbf{c}(t-T_1)) \\
&\le \mathbf{c}(\rho)\exp(-(\mathbf{c}(t-T_1))^{-1}d(x,\xi)^2) \\
&\le \mathbf{c}(\rho)\exp(-(\mathbf{c}(t-T_1))^{-1}|x-\xi|^2).
\end{aligned}$$

This completes the proof. ∎

Lemma 12.4 *Let $\mu \ge 0$ and $T_2 > T_1$ be such that $(T_2 - T_1)\mu$ is small enough. Let $f(t,x)$ be a continuous function on $[T_1,T_2]\times\mathbb{R}^n$, locally d-Hölder continuous in x, uniformly w.r.t. t, satisfying the growth condition*

$$|f(t,x)| \le M \exp(\mu|x|^2)$$

for some constant $M > 0$. Then the function

$$V_f(z) = \int_{[T_1,t]\times\mathbb{R}^n} h_\eta(z;\eta)\, f(\eta)d\eta, \quad z = (t,x) \in [T_1,T_2]\times\mathbb{R}^n \tag{12.1}$$

belongs to the class $\mathfrak{C}^2((T_1,T_2)\times\mathbb{R}^n) \cap C([T_1,T_2]\times\mathbb{R}^n)$ and we have

$$HV_f(z) = f(z) - \int_{[T_1,t]\times\mathbb{R}^n} Z_1(z;\eta)\, f(\eta)d\eta, \quad z = (t,x) \in (T_1,T_2)\times\mathbb{R}^n. \tag{12.2}$$

Proof. V_f is well-posed since, by means of (10.10), (10.15) and (10.17), we have (we use the notation $z = (t,x)$, $\eta = (s,y)$)

$$\begin{aligned}
|h_\eta(z,\eta)f(\eta)| &\le \mathbf{c}(T_1,T_2,f,\mu)\mathbf{E}(x,y,\mathbf{c}_0(t-s))\exp(\mu|y|^2) \\
&\le \mathbf{c}(T_1,T_2,f,\mu)\mathbf{E}(x,y,2\mathbf{c}_0(t-s))\exp(2\mu|x|^2) \in L^1_\eta((T_1,t)\times\mathbb{R}^n).
\end{aligned}$$

Using the estimate (10.15) in a similar manner and following arguments similar to those in the proof of Lemma 11.10, one can see that V has the required regularity and

$$\begin{aligned}
X_jV_f(z) &= \int_{[T_1,t]\times\mathbb{R}^n} X_j h_\eta(z,\eta)\, f(\eta)d\eta, \\
X_iX_jV_f(z) &= \lim_{\varepsilon\to 0^+}\int_{[T_1,t-\varepsilon]\times\mathbb{R}^n} X_iX_j h_\eta(z,\eta)\, f(\eta)d\eta, \\
\partial_t V_f(z) &= f(z) + \lim_{\varepsilon\to 0^+}\int_{[T_1,t-\varepsilon]\times\mathbb{R}^n} \partial_t h_\eta(z,\eta)\, f(\eta)d\eta.
\end{aligned}$$

Recalling the definition (11.2) of Z_1, we obtain (12.2). ∎

Lemma 12.5 *Let $\mu \geq 0$ and $T_2 > T_1$ be such that $(T_2 - T_1)\mu$ is small enough. Let f be a continuous function on $[T_1, T_2] \times \mathbb{R}^n$, satisfying the growth condition*

$$|f(t,x)| \leq M \exp(\mu |x|^2)$$

for some constant $M > 0$. Then the function

$$\widetilde{f}(z) = \int_{[T_1,t]\times\mathbb{R}^n} \Phi(z;\eta)\, f(\eta) d\eta, \quad z = (t,x) \in [T_1, T_2] \times \mathbb{R}^n \tag{12.3}$$

is continuous on $[T_1, T_2] \times \mathbb{R}^n$. Moreover, $\widetilde{f}$ is locally d-Hölder continuous in x, uniformly w.r.t. t, and satisfies the growth condition

$$|\widetilde{f}(t,x)| \leq \widetilde{M} \exp(2\mu |x|^2)$$

for some constant $\widetilde{M} > 0$.

Proof. The growth estimate of $\widetilde{f}$ immediately follows from (11.4), (10.15) and (10.17). The Hölder continuity easily follows from Proposition 11.6, (10.15) and (10.17). We only need to prove that $\widetilde{f}(x, \cdot) \in C([T_1, T_2])$ for every fixed $x \in \mathbb{R}^n$. We have

$$\left|\widetilde{f}(t+h,x) - \widetilde{f}(t,x)\right| \leq \int_{[T_1,t-\sigma]\times\mathbb{R}^n} |(\Phi(t+h,x;\eta) - \Phi(t,x;\eta))\, f(\eta)|\, d\eta$$
$$+ \int_{[t-\sigma,t+h]\times\mathbb{R}^n} |\Phi(t+h,x;\eta)\, f(\eta)|\, d\eta + \int_{[t-\sigma,t]\times\mathbb{R}^n} |\Phi(t,x;\eta)\, f(\eta)|\, d\eta.$$

The first integral in the right-hand side vanishes as $h \to 0$, by dominated convergence, by making use of (11.4), (10.12), (10.15), (10.17) and Proposition 11.6. On the other hand, the other two integrals are smaller than $\mathbf{c}(f,\mu,x)(\sigma+h)^{\alpha/2}$ and $\mathbf{c}(f,\mu,x)\sigma^{\alpha/2}$ respectively, again by means of (11.4), (10.15) and (10.17). ■

Lemma 12.6 *The statement of Theorem 12.1 holds if $g \equiv 0$.*

Proof. We set

$$w = V_f + V_{\widetilde{f}},$$

where $\widetilde{f}$ is defined by (12.3) and V_f is defined by (12.1). From Lemma 12.4 and Lemma 12.5, it follows that

$$w \in \mathfrak{C}^2((T_1, T_2) \times \mathbb{R}^n) \cap C([T_1, T_2] \times \mathbb{R}^n)$$

and we have (using the estimates (11.7), (11.4), (10.15) and (10.17))

$$
\begin{aligned}
Hw(z) &= f(z) - \int_{[T_1,t]\times\mathbb{R}^n} \left(Z_1(z;\zeta) - \Phi(z;\zeta)\right) f(\zeta) d\zeta \\
&\quad - \int_{[T_1,t]\times\mathbb{R}^n} Z_1(z;\eta) \left(\int_{[T_1,s]\times\mathbb{R}^n} \Phi(\eta;\zeta)\, f(\zeta)\, d\zeta \right) d\eta \\
&= f(z) - \int_{[T_1,t]\times\mathbb{R}^n} \left[Z_1(z;\zeta) - \Phi(z;\zeta) + \int_{[\tau,t]\times\mathbb{R}^n} Z_1(z;\eta)\, \Phi(\eta;\zeta) d\eta \right] f(\zeta) d\zeta \\
&= f(z),
\end{aligned}
$$

since the expression [...] in the last integral vanishes by (11.5).

It is now sufficient to recognize that $w = u$:

$$
\begin{aligned}
w(z) &= V_f(z) + \int_{[T_1,t]\times\mathbb{R}^n} f(\zeta) \left(\int_{[\tau,t]\times\mathbb{R}^n} h_\eta(z;\eta)\, \Phi(\eta;\zeta)\, \eta \right) d\zeta \\
&= \int_{[T_1,t]\times\mathbb{R}^n} f(\zeta) \left(h_\zeta(z;\zeta) + J(z;\zeta) \right) d\zeta = u(z).
\end{aligned}
$$

This ends the proof. ■

Proof of Theorem 12.1. It directly follows by Lemma 12.3 and Lemma 12.6. ■

13 Lower bounds for fundamental solutions

We begin with the following weak maximum principle for H in the class $\mathfrak{C}^2$, which is a consequence of the results in [7].

Theorem 13.1 *Let U be a bounded open subset of $\mathbb{R}^{n+1}$ and let $t_0 \in \mathbb{R}$. If*

$$u \in \mathfrak{C}^2(U),\ Hu \le 0 \text{ in } U \cap \{t < t_0\},\ \limsup u \le 0 \text{ in } \partial U \cap \{t < t_0\},$$

then $u \le 0$ in $U \cap \{t < t_0\}$.

Proof. The scheme of the proof is classical. The new difficulty is due to the "weak regularity" of u, namely $u \in \mathfrak{C}^2(U)$. Let us show that $u \le 0$ in

$$U_T := U \cap \{t < T\},$$

for every $T < t_0$, by proving that

$$v_\varepsilon(t,x) = (u(t,x) - \varepsilon g(t)) \exp(-2Kt) \le 0$$

in U_T for every $\varepsilon > 0$, where

$$g(t) = \exp\left(\frac{2KR^2}{T-t}\right), \qquad R = \max_{(t,x),(t',x')\in\overline{U}} |t - t'|$$

(K has been introduced in (10.4)). Let $\overline{z} \in \overline{U}_T$ be such that $\sup_{V \cap U_T} v_\varepsilon = \sup_{U_T} v_\varepsilon$ for every neighborhood V of $\overline{z}$. If $\overline{z} \in \partial U_T$, then we immediately get $v_\varepsilon \leq 0$ in U_T, since $\limsup v_\varepsilon \leq 0$ in ∂U_T, by the definition of v_ε, the continuity of u and the hypothesis on $\limsup u$ (we stress that u is continuous in the pair (t, x), by the definition of $\mathfrak{C}^2$). Suppose now that $\overline{z} \in U_T$. Then $\overline{z}$ is a maximum point of v_ε in U_T. We can then use Proposition 2.4 in [7] and obtain

$$\begin{cases} X_j v_\varepsilon(\overline{z}) = 0, \quad j = 1, \ldots, m, \qquad \partial_t v_\varepsilon(\overline{z}) = 0, \\ \sum_{i,j=1}^m X_i X_j v_\varepsilon(\overline{z})\, \omega_i \omega_j \leq 0 \quad \forall\, \omega \in \mathbb{R}^m. \end{cases}$$

Since $(a_{i,j}(\overline{z}))_{i,j}$ is positive definite (see (10.4)), we get

$$\sum_{i,j=1}^m a_{i,j}(\overline{z}) X_i X_j v_\varepsilon(\overline{z}) \leq 0.$$

Therefore (recalling the definition (10.1) of H),

$$H v_\varepsilon(\overline{z}) \geq -a_0(\overline{z}) v_\varepsilon(\overline{z}). \tag{13.1}$$

On the other hand, recalling that $|a_0| \leq K$ (see (10.4)) and using the hypothesis $Hu \leq 0$, we also have

$$\begin{aligned} H v_\varepsilon(\overline{z}) &= \exp(-2K\overline{t}) \left(Hu(\overline{z}) - \varepsilon \left(g'(\overline{t}) - a_0(\overline{z}) g(\overline{t}) \right) \right) - 2K v_\varepsilon(\overline{z}) \\ &\leq \exp(-2K\overline{t}) \left(-\varepsilon \left(g'(\overline{t}) - K g(\overline{t}) \right) \right) - 2K v_\varepsilon(\overline{z}). \end{aligned} \tag{13.2}$$

A direct computation gives:

$$g' - Kg = g \cdot \frac{2KR^2}{(T-t)^2} - Kg = Kg \left(\frac{2R^2}{(T-t)^2} - 1 \right) \geq Kg$$

since $(T-t)^2 \leq R^2$. Inserting the last inequality in (13.2), by (13.1) we have

$$-a_0(\overline{z}) v_\varepsilon(\overline{z}) \leq -\varepsilon K g(\overline{t}) \exp(-2K\overline{t}) - 2K v_\varepsilon(\overline{z}).$$

As a consequence,

$$v_\varepsilon(\overline{z})(a_0(\overline{z}) - 2K) \geq \exp(-2K\overline{t}) \varepsilon K g(\overline{t}) > 0$$

and $v_\varepsilon \leq v_\varepsilon(\overline{z}) < 0$ in U_T. This completes the proof. ∎

The following version of the weak maximum principle in infinite strips easily follows from Theorem 13.1.

Corollary 13.2 *Let $u \in \mathfrak{C}^2((T_1, T_2) \times \mathbb{R}^n)$ be such that $Hu \leq 0$ in $(T_1, T_2) \times \mathbb{R}^n$ and $\limsup u \leq 0$ in $\{T_1\} \times \mathbb{R}^n$ and at infinity. Then $u \leq 0$ in $(T_1, T_2) \times \mathbb{R}^n$.*

Proof. Fix $\varepsilon > 0$ and find $R_\varepsilon > 0$ such that

$$\sup_{|z| \geq R_\varepsilon} u(z) \leq \varepsilon \exp(KT_1)$$

(which is possible by our assumption on limsup at infinity). Apply Theorem 13.1 to the function $u - \varepsilon \exp(Kt)$ in the set

$$C_{\varepsilon,p} = \{z \in (T_1, T_2) \times \mathbb{R}^n \,:\, |z| < p + R_\varepsilon\}$$

(for some fixed $p \in \mathbb{R}$). Recalling that $a_0 \le K$ by (10.4), we have

$$H\left(u - \varepsilon \exp(Kt)\right) = Hu + (a_0 - K)\,\varepsilon \exp(Kt) \le Hu \le 0$$

so that

$$u - \varepsilon \exp(Kt) \le 0 \text{ in } C_{\varepsilon,p}.$$

Lettin p go to infinity and ε go to zero, we get the assertion. ■

Using Corollary 13.2 we can prove the following

Proposition 13.3 *h is a nonnegative function.*

Proof. We fix $x_0 \in \mathbb{R}^n$, $t_0 > \tau_0$ and we set

$$v = h(t_0, x_0; \tau_0, \cdot).$$

We will show that $\int_{\mathbb{R}^n} vg \ge 0$, for every non-negative test function $g \in C_0^\infty(\mathbb{R}^n)$. By the continuity of v (see Proposition 11.4), this will imply the assertion. Recall that $v \in L^1(\mathbb{R}^n)$, by (11.11) and (10.17). From Theorem 12.1, it follows that

$$u(z) = \int_{\mathbb{R}^n} h(z; \tau_0, \xi)\, g(\xi) d\xi$$

belongs to the class $\mathfrak{C}^2(\{t > \tau_0\})$, $Hu = 0$ in $(\tau_0, \infty) \times \mathbb{R}^n$ and $u \to g \ge 0$ for $t \to \tau_0$. Moreover, using the estimate (11.11),

$$\begin{aligned}\sup_{\tau_0 < t < t_0+1,\, |x| > R} |u\,(t,x)\,| &\le \mathbf{c}(g, t_0, \tau_0) \int_{\operatorname{supp} g} \sup_{\tau_0 < t < t_0+1,\, |x| > R} \mathbf{E}(x, \xi, \mathbf{c}_0(t - \tau_0)) d\xi \\ &\le \mathbf{c}(g, t_0, \tau_0) \exp(-\mathbf{c}(t_0, \tau_0)^{-1} R^2) \to 0, \qquad \text{as } R \to \infty.\end{aligned}$$

In the last inequality we have used the fact that, if $\xi \in \operatorname{supp} g$ and $|x|$ is large enough, then $d(x, \xi) \ge \mathbf{c}^{-1}$ (see (2.3)) and (2.9), (10.14) give

$$\begin{aligned}\mathbf{E}(x, \xi, \mathbf{c}_0(t - \tau_0)) &\le \mathbf{E}(x, \xi, 2\mathbf{c}_0(t - \tau_0)) \exp(-(2\mathbf{c}_0(t - \tau_0))^{-1} d(x, \xi)^2) \\ &\le \mathbf{c} \exp(-\mathbf{c}(t_0, \tau_0)^{-1} R^2).\end{aligned}$$

Therefore u goes to zero at infinity in the strip $(\tau_0, t_0 + 1) \times \mathbb{R}^n$. We now apply Corollary 13.2 and obtain $u \ge 0$ in $(\tau_0, t_0 + 1) \times \mathbb{R}^n$. In particular, we get $\int_{\mathbb{R}^n} vg = u\,(t_0, x_0) \ge 0$. ■

Proposition 13.4 (Reproduction property for the fundamental solution) *We have*

$$h(t, x; \tau, \xi) = \int_{\mathbb{R}^n} h(t, x; s, y)\, h(s, y; \tau, \xi) dy,$$

for $t > s > \tau$ and x, $\xi \in \mathbb{R}^n$.

Proof. We fix τ, ξ, s as above and we set

$$u = \int_{\mathbb{R}^n} h(\cdot; s, y)\, h(s, y; \tau, \xi) dy, \qquad v = h(\cdot; \tau, \xi).$$

>From Theorem 12.1 it follows that (for $T_2 > T_1 = s$)

$$u \in \mathfrak{C}^2((s, T_2) \times \mathbb{R}^n) \cap C([s, T_2] \times \mathbb{R}^n),$$

$Hu = 0$ in $(s, T_2) \times \mathbb{R}^n$ and $u\,(s, \cdot) = v\,(s, \cdot)$ (note that $v\,(s, \cdot)$ is continuous and bounded, by Proposition 11.4 and (10.14)). On the other hand, from Theorem 11.8, we know that $v \in \mathfrak{C}^2(\mathbb{R}^{n+1} \setminus \{(\tau, \xi)\})$, $Hv = 0$ in $\mathbb{R}^{n+1} \setminus \{(\tau, \xi)\}$. Thus, we only need to show that

$$\sup_{s<t<T_2,\, |x|>R} |u\,(t, x) - v\,(t, x)\,| \longrightarrow 0, \quad \text{as } R \to \infty, \tag{13.3}$$

in order to get $u = v$ in $(s, T_2) \times \mathbb{R}^n$ from the weak maximum principle in Corollary 13.2. Let us prove (13.3). Using the estimates (11.11), (10.16), (2.11) and (2.3), we see that the supremum in (13.3) is lower than

$$\begin{aligned}
&\mathbf{c}(T_2, \tau) \sup_{s<t<T_2,\, |x|>R} \left(\mathbf{E}(x, \xi, \mathbf{c}(t-\tau)) + \int_{\mathbb{R}^n} \mathbf{E}(x, y, \mathbf{c}'(t-s))\mathbf{E}(y, \xi, \mathbf{c}''(s-\tau))dy \right) \\
&\quad \le \mathbf{c}(T_1, T_2, \tau) \sup_{|x|>R} \exp\left(-\frac{d^2(x, \xi)}{\mathbf{c}(T_2, \tau)} \right) \\
&\quad \le \mathbf{c}(T_1, T_2, \tau) \sup_{|x|>R} \exp\left(-\frac{|x-\xi|^2}{\mathbf{c}(T_2, \tau)} \right) \longrightarrow 0, \quad \text{as } R \to \infty.
\end{aligned}$$

■

The following lemma is the main step in the proof of the lower bound of h.

Lemma 13.5 *There exists a positive constant $\delta = \mathbf{c}^{-1}$, such that*

$$h(t, x; \tau, \xi) \ge \mathbf{c}^{-1}\, \mathbf{E}(x, \xi, \mathbf{c}^{-1}(t-\tau)),$$

if $0 < t - \tau < \delta$ *and* $d^2(x, \xi) \le \delta\,(t-\tau)\,|\log(t-\tau)|$.

Proof. Let $0 < t - \tau < \delta < 1$ and $d^2(x, \xi) \le \delta\,(t-\tau)\,|\log(t-\tau)|$. From (11.9), (11.10), (10.10) and (2.9) it follows that ($z = (t, x)$, $\zeta = (\tau, \xi)$)

$$\begin{aligned}
h(z; \zeta) &= h_\zeta(z; \zeta) + J(z; \zeta) \\
&\ge \mathbf{c}_0^{-1}\, \mathbf{E}(x, \xi, \mathbf{c}_1^{-1}(t-\tau)) - \mathbf{c}_2\,(t-\tau)^{\frac{\alpha}{2}}\, \mathbf{E}(x, \xi, \mathbf{c}_3(t-\tau)) \\
&\ge \mathbf{c}_0^{-1}\, \mathbf{E}(x, \xi, \mathbf{c}_1^{-1}(t-\tau)) \left(1 - \mathbf{c}_4\,(t-\tau)^{\frac{\alpha}{2}}\, \exp\left(\mathbf{c}_1\, \frac{d^2(x, \xi)}{t-\tau} \right) \right).
\end{aligned}$$

We now choose $\delta < \min\{1, \frac{\alpha}{4\mathbf{c}_1}, (2\mathbf{c}_4)^{-4/\alpha}\}$ in order to get

$$\mathbf{c}_4\,(t-\tau)^{\frac{\alpha}{2}}\, \exp\left(\mathbf{c}_1\, d^2(x, \xi)/(t-\tau)\right) \le \mathbf{c}_4\,(t-\tau)^{\frac{\alpha}{2} - \delta\,\mathbf{c}_1} \le \mathbf{c}_4\,(t-\tau)^{\frac{\alpha}{4}} \le \tfrac{1}{2}.$$

This completes the proof. ■

We are now in position to prove the main result of this section.

Theorem 13.6 (Lower bound for the fundamental solution) *For every $T > 0$, there exists a positive constant $\mathbf{c}(T)$ such that*

$$h(t, x; \tau, \xi) \geq \mathbf{c}(T)^{-1}\, \mathbf{E}(x, \xi, \mathbf{c}^{-1}(t-\tau)), \tag{13.4}$$

for $0 < t - \tau \leq T$ and x, $\xi \in \mathbb{R}^n$.

Proof. This proof is similar to that of Theorem 7.1 in Part I.

Let $\delta < e^{-1}$ be as in Lemma 13.5 and let us fix $T > \delta$, $x, \xi \in \mathbb{R}^n$ and $0 < t - \tau \leq T$. Let k be the smallest integer greater than

$$\max\{T\,\delta^{-1}, 16\, d^2(x,\xi)\,(\delta(t-\tau))^{-1}\}$$

and let us set

$$\sigma = \frac{1}{4}\sqrt{\delta(t-\tau)/(k+1)}.$$

There exists a chain of points $x = x_0, x_1, \ldots, x_{k+1} = \xi$, laying on a suitable X-subunit path connecting the points x and ξ, such that $d(x_j, x_{j+1}) \leq 2d(x,\xi)/(k+1)$, for $j = 0, \ldots, k$. Moreover, we pick $t = t_0, t_1, \ldots, t_{k+1} = \tau$ such that $t_j - t_{j+1} = (t-\tau)/(k+1)$ for $j = 0, \ldots, k$. Using Proposition 13.4 repeatedly, we obtain

$$\begin{aligned} h(t,x;\tau,\xi) &= \int_{\mathbb{R}^n} h(t,x;t_1,y_1)\, h(t_1,y_1;\tau,\xi) dy_1 \\ &= \int_{(\mathbb{R}^n)^k} h(t,x;t_1,y_1)\, h(t_1,y_1;t_2,y_2)\cdots h(t_k,y_k;\tau,\xi)\, dy_1\cdots dy_k \\ &\geq \int_{B(x_1,\sigma)\times\cdots\times B(x_k,\sigma)} h(t_0,y_0;t_1,y_1,)\cdots h(t_k,y_k;t_{k+1},y_{k+1})\, dy_1\cdots dy_k \end{aligned}$$

(here we have set $y_0 = x$, $y_{k+1} = \xi$). We claim that $0 < t_j - t_{j+1} < \delta$ and

$$d^2(y_j, y_{j+1}) \leq \delta\,(t_j - t_{j+1})\,|\log(t_j - t_{j+1})| \;\text{ if } d(x_j, y_j) < \sigma.$$

Indeed, by the definition of k and of σ, we have

$$t_j - t_{j+1} = (t-\tau)/(k+1) \leq T/(k+1) < \delta,$$

and

$$\begin{aligned} d(y_j, y_{j+1}) &\leq d(y_j, x_j) + d(x_j, x_{j+1}) + d(x_{j+1}, y_{j+1}) \\ &< 2\sigma + \frac{2d(x,\xi)}{k+1} < 2\sigma + \frac{\sqrt{k\delta(t-\tau)}}{2(k+1)} \\ &< \sqrt{\frac{\delta\,(t-\tau)}{(k+1)}} \leq \delta\,(t_j - t_{j+1})\,|\log(t_j - t_{j+1})| \end{aligned}$$

since $t-\tau<\delta<e^{-1}$. Thus, we can apply Lemma 13.5 and obtain that

$$\begin{aligned}
&h(t,x;\tau,\xi)\geq \\
&\geq \mathbf{c}^{-1}\exp(-\mathbf{c}k)\int_{B(x_1,\sigma)\times\cdots\times B(x_k,\sigma)}\prod_{j=0}^{k}|B(y_j,\mathbf{c}_0\sqrt{t_j-t_{j+1}})|^{-1}\,dy_1\cdots dy_k \\
&\geq \mathbf{c}^{-1}\exp(-\mathbf{c}k)|B(x,c_1\sigma)|^{-1}\prod_{j=1}^{k}\frac{|B(x_j,\sigma)|}{|B(x_j,\mathbf{c}_1\sigma)|} \\
&\geq \mathbf{c}^{-1}|B(x,\sqrt{\mathbf{c}(t-\tau)})|^{-1}\exp(-\mathbf{c}\,k),
\end{aligned}$$

by (2.9) and the definition of σ. Now, if $d^2(x,\xi)>T(t-\tau)/16$, from the definition of k it follows that $\exp(-\mathbf{c}\,k)\geq\exp(-\mathbf{c}d^2(x,\xi)/(t-\tau)$. On the other hand, if $d^2(x,\xi)<T(t-\tau)/16$, the definition of k gives $\exp(-\mathbf{c}\,k)\geq\mathbf{c}(T)^{-1}$. Using again (2.9) we finally get (13.4). ■

14 Regularity results

In this section we will prove that $\mathfrak{C}^2$ solutions to the equation $Hu=f$, with $f\in C^\alpha$, actually belong to $C^{2,\alpha}$. In particular, this implies that $h(\cdot,\zeta)\in C^{2,\alpha}_{loc}$. This regularization result will follow using the Schauder theory developed in [12] and the weak maximum principle for $\mathfrak{C}^2$ solutions contained in Theorem 13.1, together with some classical results due to Bony [8]. We start recalling two theorems which are proved, respectively, in §10 and §11 of [12].

Theorem 14.1 (Schauder estimates for H) *For any domain $U'\Subset U$ there exists a constant $\mathbf{c}(U,U')>0$ such that for every $u\in C^{2,\alpha}_{loc}(U)$ with $Hu\in C^\alpha(U)$ one has*

$$\|u\|_{C^{2,\alpha}(U')}\leq\mathbf{c}(U,U')\left\{\|Hu\|_{C^\alpha(U)}+\|u\|_{L^\infty(U)}\right\}. \tag{14.1}$$

Recall that the exponent α is the one appearing in our assumptions on the coefficients of H (see (10.4)).

Theorem 14.2 (mollifiers on C^β) *Let $\mathbf{h}(t,x,y)$ be the fundamental solution of the operator*

$$\mathbf{H}=\partial_t-\mathbf{L}=\partial_t-\sum_{i=1}^m X_i^2,$$

fix a positive test function $\eta\in C_0^\infty(\mathbb{R})$ such that $\int\eta(t)\,dt=1$ and set

$$\phi_\varepsilon(t,x,y)=\varepsilon^{-1}\mathbf{h}(\varepsilon,x,y)\,\eta\left(\frac{t}{\varepsilon}\right).$$

For any $\beta\in(0,1)$, $f\in C^\beta\left(\mathbb{R}^{n+1}\right)$ set

$$f_\varepsilon(t,x)=\int_{\mathbb{R}^{n+1}}\phi_\varepsilon(t-s,x,y)\,f(s,y)\,dsdy.$$

Then:

$$\|f_\varepsilon\|_{C^\beta} \le c\,\|f\|_{C^\beta}$$

with c independent of f, ε;

$$\lim_{\varepsilon\to 0} \|f_\varepsilon - f\|_{L^\infty(\mathbb{R}^{n+1})} = 0;$$

if a_{ij} satisfy ellipticity condition in (10.4), then a_{ij}^ε satisfy the same condition, with ellipticity constant independent of ε.

We will also need the following compactness lemma:

Lemma 14.3 *Let $\{u_j\}$ be a sequence in $C^{k,\beta}(U)$, for some positive integer k, $\beta \in (0,1)$, and U bounded domain in $\mathbb{R}^{n+1}$, such that*

$$\|u_j\|_{C^{k,\beta}(U)} \le c$$

with c independent of j. Then, there exists a subsequence u_{j_h} and a function $u \in C^{k,\beta}(U)$ such that

$$\partial_t^m X^I u_{j_h} \to \partial_t^m X^I u$$

uniformly in U for any m, I with $2m + |I| \le k$.

Proof. For any m, I such that $2m + |I| \le k$, the functions $\partial_t^m X^I u_j$ are equibounded and equicontinuous (in classical sense), hence by Arzelà's theorem there exists a subsequence $\partial_t^m X^I u_{j_h}$ uniformly converging in U to some function $v_{m,I}$. (See Remark 10.5). Moreover, we can extract a single subsequence u_{j_h} such that all these conditions simultaneously hold. Set $u = v_{0,0}$. By Lemma 11.9, this implies that the derivative $\partial_t^m X^I u$ exists and equals $v_{m,I}$. Finally, passing to the limit in the inequality

$$\left|\partial_t^m X^I u_{j_h}(t,x) - \partial_t^m X^I u_{j_h}(s,y)\right| \le c\, d_P\left((t,x),(s,y)\right)^\beta$$

we find that actually $u \in C^{k,\alpha}(U)$. ∎

Theorem 14.4 *Let $u \in \mathfrak{C}^2(U)$ be a solution to the equation $Hu = f$, with $f \in C^\alpha(U)$. Then $u \in C^{2,\alpha}_{loc}(U)$ and satisfies (14.1). In particular, for every $\zeta \in \mathbb{R}^{n+1}$, the fundamental solution $h(\cdot,\zeta)$ belongs to $C^{2,\alpha}_{loc}(\mathbb{R}^{n+1} \setminus \{\zeta\})$.*

Proof. This proof is similar to that of Theorem 11.5 in [12]. Let $u \in \mathfrak{C}^2(U)$, $f = Hu \in C^\alpha(U)$. Assume first that a_0 satisfies the sign condition

$$a_0(t,x) \le -c < 0 \text{ for any } (t,x) \in \mathbb{R}^{n+1}. \tag{14.2}$$

Let now $a_{ij}^\varepsilon, a_j^\varepsilon, a_0^\varepsilon$, f^ε be the mollified versions of a_{ij}, a_j, a_0 and f, and set

$$H^\varepsilon = \partial_t - \sum_{i,j=1}^m a_{ij}^\varepsilon(t,x)\, X_i X_j - \sum_{i=1}^m a_i^\varepsilon(t,x)\, X_i - a_0^\varepsilon(t,x).$$

Recall that the a_{ij}^{ε}'s satisfy the ellipticity condition in (10.4) with constant λ independent of ε. Since H^{ε} has smooth coefficients, it can be written as a Hörmander operator. This, together with condition (14.2) (note that also a_0^{ε} satisfies this condition), allows to apply known results of Bony [8]: for every point of U' we can find a neighborhood D where we can uniquely solve the classical Dirichlet problem:

$$\begin{cases} H^{\varepsilon} u^{\varepsilon} = f^{\varepsilon} & \text{in } D \\ u^{\varepsilon} = u & \text{on } \partial D \end{cases}$$

Moreover, the domain D satisfies the following regularity property which will be useful later (see [8, Corollary 5.2]): for every point $(t_1, x_1) \in \partial D$ there exists an Euclidean ball of center $(t_0, x_0) \notin \overline{D}$ which intersects $\overline{D}$ exactly at (t_1, x_1).

Since H^{ε} is hypoelliptic, the solution u^{ε} belongs to $C^{\infty}(D)$, then we can apply our a-priori estimates (14.1), writing

$$\|u^{\varepsilon}\|_{C^{2,\alpha}(D')} \leq C(\varepsilon, D, D') \left\{ \|f^{\varepsilon}\|_{C^{\alpha}(D)} + \|u^{\varepsilon}\|_{L^{\infty}(D)} \right\}.$$

The constant $C(\varepsilon, D, D')$ depends on the coefficients $a_{ij}^{\varepsilon}, a_j^{\varepsilon}, a_0^{\varepsilon}$ only through their $C^{\alpha}(D)$-norms and the ellipticity constant, hence by Theorem 14.2, C can be bounded independently of ε. For the same reason $\|f^{\varepsilon}\|_{C^{\alpha}(D)} \leq \mathbf{c} \|f\|_{C^{\alpha}(D)}$.

To get a bound independent of ε also on $\|u_{\varepsilon}\|_{L^{\infty}(D)}$, let

$$v^{\varepsilon}(t) = \max_{\partial D} |u^{\varepsilon}| + e^{t-M} \max_{\overline{D}} |f^{\varepsilon}|$$

with $M = \min_{(t,x) \in \overline{D}} t$. Then

$$\begin{aligned} H^{\varepsilon} v^{\varepsilon}(t,x) &= e^{t-M} \max_{\overline{D}} |f^{\varepsilon}| - a_0(t,x) v^{\varepsilon} \geq e^{t-M} \max_{\overline{D}} |f^{\varepsilon}| \\ &\geq \max_{\overline{D}} |f^{\varepsilon}| \geq f(t,x) = H_{\varepsilon} u^{\varepsilon}(t,x) \text{ in } D \end{aligned}$$

while

$$v^{\varepsilon} \geq u^{\varepsilon} \text{ on } \partial D,$$

hence by Theorem 13.1,

$$\begin{aligned} \|u^{\varepsilon}\|_{L^{\infty}(D)} &\leq \|v^{\varepsilon}\|_{L^{\infty}(D)} \leq \|u\|_{L^{\infty}(\partial D)} + \mathbf{c} \max_{\overline{D}} |f^{\varepsilon}| \leq \\ &\leq \|u\|_{L^{\infty}(\partial D)} + \mathbf{c} \|f\|_{C^{\alpha}(D)} \end{aligned}$$

This means that, for any $D' \Subset D$,

$$\|u^{\varepsilon}\|_{C^{2,\alpha}(D')} \leq \mathbf{c}(D', D) \left\{ \|u\|_{L^{\infty}(\partial D)} + \|f\|_{C^{\alpha}(D)} \right\}. \tag{14.3}$$

Hence, by Lemma 14.3, for every $D' \Subset D$ we can find a sequence $\varepsilon_n \to 0$ and a function $v \in C^{2,\alpha}(D')$ such that

$$\partial_t^m X^I u^{\varepsilon_n} \to \partial_t^m X^I v$$

uniformly in D', for $2m+|I|\leq 2$. By a standard "diagonal argument", we can also select a single sequence $\varepsilon_n\to 0$ and a function $v\in C^{2,\alpha}_{loc}(D)$ such that $\partial_t^m X^I u^{\varepsilon_n}\to\partial_t^m X^I v$ locally uniformly and pointwise in D. Moreover,

$$H^{\varepsilon_n}u^{\varepsilon_n}=f_{\varepsilon_n}\to f, \text{ but also } H^{\varepsilon_n}u^{\varepsilon_n}\to Hv$$

hence

$$Hv=f \text{ in } D.$$

Our next task is to show that $v=u$ in D; this will imply $u\in C^{2,\alpha}_{loc}(D)$, that is the desired regularity result. To do this, we will make use of a classical argument of barriers, taken from [8], to show that $u=v$ on ∂D; this will imply that $v=u$ in D, again by the maximum principle (Theorem 13.1).

Fix a point $(t_1,x_1)\in\partial D$; let (t_0,x_0) be the center of the exterior ball that touches ∂D at (t_1,x_1), and set:

$$w(t,x)=e^{-N\left[|x-x_0|^2+(t-t_0)^2\right]}-e^{-N\left[|x_1-x_0|^2+(t_1-t_0)^2\right]}$$

with N a positive constant to be chosen later. By construction, $w(t,x)<0$ in D. A direct computation shows that, by the construction of D made in [8], $Hw(t,x)<0$ in a suitable neighborhood D_1 of (t_1,x_1), choosing N large enough. Next, we compute, for another large constant M:

$$H(Mw\pm(u^\varepsilon-u))=MHw\pm(f^\varepsilon-f)<0 \text{ in } D_1\cap D$$

for M large enough, since $(f^\varepsilon-f)$ is uniformly bounded with respect to ε. Let us show that

$$Mw\pm(u^\varepsilon-u)<0 \text{ on } \partial(D_1\cap D).$$

On $D_1\cap\partial D$, we have $Mw\pm(u^\varepsilon-u)=Mw\leq 0$; on the other hand, on $\partial D_1\cap D$ we have $w\leq c_0$ for some $c_0<0$, while $(u^\varepsilon-u)$ is uniformly bounded with respect to ε; hence for M large enough $Mw\pm(u^\varepsilon-u)\leq 0$. The maximum principle then implies

$$Mw\pm(u^\varepsilon-u)\leq 0 \text{ in } D_1\cap D$$

that is

$$|u^\varepsilon-u|\leq -Mw \text{ in } D_1\cap D, \text{ uniformly in } \varepsilon.$$

For $\varepsilon\to 0$ we get

$$|(v-u)(t,x)|\leq -Mw(t,x) \text{ for } (t,x)\in D_1\cap D$$

and, for $(t,x)\to(t_1,x_1)$ we get $v(t_1,x_1)=u(t_1,x_1)$. This ends the proof of our result, under the additional assumption (14.2). In the general case, if u satisfies $Hu=f$, then

$$v_k(t,x)=e^{kt}u(t,x)$$

satisfies

$$(H-k)v_k(t,x)=e^{kt}f(t,x).$$

For $-k > 0$ large enough, we can apply the previous argument to $H - k$, getting

$$\|u\|_{C^{2,\alpha}(U')} \leq \mathbf{c} \|v_k\|_{C^{2,\alpha}(U')} \leq \mathbf{c} \left\{ \left\| e^{kt} f \right\|_{C^\alpha(U)} + \|v_k\|_{L^\infty(U)} \right\} \leq$$
$$\leq \mathbf{c} \left\{ \|f\|_{C^\alpha(U)} + \|u\|_{L^\infty(U)} \right\}.$$

■

We also point out the following proposition. Although it will never be used in the following, it completes the picture of regularization properties, assuring that, when the coefficients of the operator are smooth, our "weak" $\mathfrak{C}^2$ solutions are also solutions in distributional sense:

Proposition 14.5 *Suppose the coefficients $a_{i,j}, a_k, a_0$ of H are smooth. Then H is hypoelliptic in $\mathbb{R}^{n+1}$ and, for every fixed $\zeta \in \mathbb{R}^{n+1}$,*

$$H(h(\cdot;\zeta)) = \delta_\zeta \quad \text{in } \mathcal{D}'(\mathbb{R}^{n+1}). \tag{14.4}$$

Moreover, given an open set $U \subset \mathbb{R}^{n+1}$, we have

$$(u \in \mathfrak{C}^2(U),\ Hu = 0 \text{ in } U) \Rightarrow (u \in C^\infty(U)). \tag{14.5}$$

Proof. It has already been noted that if the coefficients are smooth, then H can be rewritten as a Hörmander operator, hence H is hypoelliptic. Moreover, the adjoint operator H^* is well defined, hence by standard computation, the representation formulas written in Theorem 12.1 also imply (14.4), once we have proved the following claim: if a continuous function u has continuous intrinsic derivative $X_j u$, then $X_j u$ is also a derivative in the sense of distributions. This fact, together with hypoellipticity, will also imply (14.5). To show this, let u_{ε_n} be a sequence of smooth functions obtained from u with a standard (Euclidean) mollification procedure. It has been proved in [7, Proposition 2.2.] that

$$X_j u_{\varepsilon_n} \to X_j u$$

uniformly on compact subsets of U. Then we have, for any test function φ,

$$\int u X_j^* \varphi = \lim_{n\to\infty} \int u_{\varepsilon_n} X_j^* \varphi = \lim_{n\to\infty} \int (X_j u_{\varepsilon_n}) \varphi = \int (X_j u) \varphi.$$

This proves the claim. ■

Part III

Harnack inequality for operators with Hölder continuous coefficients

15 Overview of Part III

In this part we come to one of the main goals of our theory, namely the proof of invariant Harnack inequalities for evolutionary or stationary operators of the kind

$$H = \partial_t - \sum_{i,j=1}^{m} a_{i,j}\,(t,x)\,X_i X_j - \sum_{k=1}^{m} a_k\,(t,x)\,X_k \tag{15.1}$$

or

$$L = \sum_{i,j=1}^{m} a_{i,j}\,(x)\,X_i X_j + \sum_{k=1}^{m} a_k\,(x)\,X_k, \tag{15.2}$$

respectively. Our assumptions are the same as in Part II (as stated in Section 10; see also Remark 10.9), except for the vanishing of the zero-order term (a_0 in (10.1)), that here will be assumed.

Our main result is the following:

Theorem 15.1 (Parabolic Harnack inequality) *Let H be as above. Let $R_0 > 0$, $0 < h_1 < h_2 < 1$ and $\gamma \in (0,1)$. There exists a positive constant $M = \mathbf{c}(h_1, h_2, \gamma, R_0)$ such that for every $(\tau_0, \xi_0) \in \mathbb{R}^{n+1}$, $R \in (0, R_0]$ and every*

$$u \in \mathfrak{C}^2((\tau_0 - R^2, \tau_0) \times B(\xi_0, R)) \cap C([\tau_0 - R^2, \tau_0] \times \overline{B(\xi_0, R)})$$

satisfying $Hu = 0$, $u \geq 0$ in $(\tau_0 - R^2, \tau_0) \times B(\xi_0, R)$, we have:

$$\max_{[\tau_0 - h_2 R^2, \tau_0 - h_1 R^2] \times \overline{B(\xi_0, \gamma R)}} u \leq M\, u(\tau_0, \xi_0).$$

Remark 15.2 *The space $\mathfrak{C}^2$ has been introduced in Definition 10.3. Recall that, by Theorem 14.4, any solution in $\mathfrak{C}^2$ also belongs to $C^{2,\alpha}$.*

The above theorem immediately implies the stationary version:

Theorem 15.3 (Harnack inequality for stationary operators) *Let L be as above. Let $R_0 > 0$. There exists a positive constant $M = \mathbf{c}(R_0)$ such that for every $\xi_0 \in \mathbb{R}^n$, $R \in (0, R_0]$ and every $u \in \mathfrak{C}^2(B(\xi_0, 3R))$ satisfying*

$$Lu = 0,\ u \geq 0,\ \textit{in } B(\xi_0, 3R)$$

one has

$$\max_{\overline{B(\xi_0, R)}} u \leq M \min_{\overline{B(\xi_0, R)}} u.$$

The strategy we follow to prove "parabolic" Harnack inequality for operators (15.1) is inspired to the paper by Fabes-Stroock [22], who, in turn, exploited the original ideas by Krylov-Safanov about parabolic operators in nondivergence form (see [32], [33], [50]). In Fabes-Stroock's paper, Harnack inequality is derived by a fairly short but clever combination of estimates based only on the Gaussian bounds (from above and below) on the Green function for a cylinder. The radius of the cylinder incorporates the essential geometrical information, giving dilation invariance to the Harnack estimate.

About at the same time of Fabes-Stroock paper, the same deep ideas were applied by Kusuoka-Stroock [35] in the context of Hörmander's operators $\partial_t - \sum_{i=1}^{q} X_i^2$. Much more recently, this general strategy has been adapted by Bonfiglioli, Uguzzoni [6] to study nonvariational operators structured on Hörmander's vector fields in Carnot groups.

Here we will follow the same line. The striking feature of this proof is the "axiomatic" nature of its core: it depends only on the suitable Gaussian estimates for the Green function, a maximum principle for H, the fact that constants are solutions to $Hu = 0$ (absence of the zero order term), and some geometric properties of CC-distance and balls, like the doubling property for the Lebesgue measure of metric balls. Then, also in our subriemannian setting, and for operators in nondivergence form, we recover an axiomatic link between Gaussian bounds, scaling invariant Harnack inequality, and properties of the underlying metric structure, as the one stressed by Saloff-Coste and Grigor'yan for divergence form parabolic operators on Riemannian manifolds (see the book [51] and references therein).

However, in pursuing our aim, a first problem arises: for our operator H with Hölder continuous coefficients, the existence of the Green function is not yet granted. Therefore it is convenient, as a first step, to make the qualitative assumption of smoothness on the coefficients, study the Green function in this setting and derive the Harnack inequality for operators with smooth coefficients. Since the Gaussian bounds on the Green function will be derived by the analogous bounds on the fundamental solution, established in Part II, all the constants will depend on the coefficients a_{ij}, a_k only through their C^{α}-norms and ellipticity (the constants K, λ defined in (10.4)). This will allow to get, by a limiting procedure, Harnack inequality in the non-smooth case. But a second problem arises: even for the operator with smooth coefficients, which is hypoelliptic and fits the assumptions of the classical theory developed by Bony [8], the cylinder based on a metric ball could be a bad domain for the Dirichlet problem, so that the existence of the Green functions for this kind of domain is still troublesome. Nevertheless, Lanconelli, Uguzzoni have recently proved in [39] that given two metric balls $B(\xi, \delta R), B(\xi, R)$, (with $\delta \in (0,1)$), there always exists a domain $A(\xi, R)$, regular for the (stationary) Dirichlet problem, and such that $B(\xi, \delta R) \subseteq A(\xi, R) \subseteq B(\xi, R)$ (see Lemma 16.2). The Green function for H on $\mathbb{R} \times A(\xi, R)$ must be thought as the natural substitute of the Green function for the cylinder $\mathbb{R} \times B(\xi, R)$. In the final limiting procedure, we will also use the fact that the domain $A(\xi, R)$ can be suitably chosen in order for it to be "uniformly regular" for the family of approximating operators H_{ε}

(see Lemma 18.1 for the exact statement of this property). This is another fact proved in [39].

Plan of Part III

In Section 16 we study operators with smooth coefficients, and construct a Green function G for regular domains. Then, we prove the Gaussian bounds from above and below on G. In Section 17 we derive from the Gaussian bounds on G the Harnack inequality for operators with smooth coefficients. This is the axiomatic core of the proof, strictly reflecting the line of [22]. Finally, in Section 18 we settle a suitable limiting procedure to prove parabolic Harnack inequality in the nonsmooth case. Harnack inequality in the stationary case is then an easy consequence.

16 Green function for operators with smooth coefficients on regular domains

Throughout Sections 16 and 17, we shall make the qualitative assumption that the coefficients $a_{i,j}, a_k$ of H are smooth. In Section 18, we shall turn back to Hölder continuous coefficients and complete the proof of the Harnack inequality in Theorem 15.1, by an approximation argument.

We start with the following definitions:

Definition 16.1 *We shall say that a bounded cylinder*

$$D = (T_1, T_2) \times \Omega \subseteq \mathbb{R}^{n+1}$$

is H-regular, if for every continuous function φ on the parabolic boundary

$$\partial_p D = ([T_1, T_2] \times \partial\Omega) \cup (\{T_1\} \times \overline{\Omega}),$$

there exists a (unique, by Theorem 13.1) solution u_φ to

$$u \in C^\infty(D) \cap C(D \cup \partial_p D), \quad Hu = 0 \text{ in } D, \quad u = \varphi \text{ in } \partial_p D. \tag{16.1}$$

We shall also say that a bounded open set $\Omega \subseteq \mathbb{R}^n$ is H-regular if, for any $T_1 < T_2$, the cylinder $(T_1, T_2) \times \Omega$ is H-regular.

By the maximum principle (Theorem 13.1), if D is an H-regular domain, for any fixed $z \in D$, the linear functional

$$T : C(\partial_p D) \to \mathbb{R},$$
$$T : \varphi \longmapsto u_\varphi(z) \text{ with } u_\varphi \text{ as in (16.1)}$$

is continuous. Therefore there exists a measure μ_z^D (supported on $\partial_p D$) so that

$$u_\varphi(z) = \int_{\partial_p D} \varphi(\zeta) d\mu_z^D(\zeta).$$

The measures $\{\mu_z^D\}_{z\in D}$ are called H-caloric measures.

The following lemma, proved in [39], states that it is always possible to approximate any bounded domain $\Omega \subset \mathbb{R}^n$ by H-regular domains, both from the inside and from the outside.

Lemma 16.2 *Let B be a bounded open set of $\mathbb{R}^n$. Then for every $\delta > 0$ there exist H-regular domains A^δ, A_δ such that*

$$\{x \in B | d(x, \partial B) > \delta\} \subseteq A_\delta \subseteq B \subseteq A^\delta \subseteq \{x \in \mathbb{R}^n | d(x, \overline{B}) < \delta\}.$$

The first goal of this section is to prove the existence and basic properties of the Green function for any regular cylinder $\mathbb{R} \times \Omega$.

Theorem 16.3 *Let $\Omega \subseteq \mathbb{R}^n$ be an H-regular domain. Then there exists a Green function $G = G^\Omega$ on the cylinder $\mathbb{R} \times \Omega$, with the properties listed below.*

(i) G is a continuous function defined on the set $\{(z,\zeta) \in (\mathbb{R}\times\overline{\Omega}) \times (\mathbb{R}\times\Omega) : z \neq \zeta\}$. Moreover, for every fixed $\zeta \in \mathbb{R}\times\Omega$, $G(\cdot;\zeta) \in C^\infty((\mathbb{R}\times\Omega)\setminus\{\zeta\})$, and we have

$$H(G(\cdot;\zeta)) = 0 \ \text{ in } (\mathbb{R}\times\Omega)\setminus\{\zeta\}, \qquad G(\cdot;\zeta) = 0 \ \text{ in } \mathbb{R}\times\partial\Omega.$$

(ii) We have $0 \le G \le h$. Moreover $G(t,x;\tau,\xi) = 0$ if $t < \tau$.

(iii) For every $\varphi \in C(\overline{\Omega})$ such that $\varphi = 0$ in $\partial\Omega$ and for every fixed $\tau \in \mathbb{R}$, the function

$$u\,(t,x) = \int_\Omega G(t,x;\tau,\xi)\,\varphi(\xi)\,d\xi, \qquad x \in \overline{\Omega},\ \ t > \tau$$

belongs to the class $C^\infty((\tau,\infty)\times\Omega) \cap C([\tau,\infty)\times\overline{\Omega})$ and solves

$$\begin{cases} Hu = 0 & \text{in } (\tau,\infty)\times\Omega, \\ u = 0 & \text{in } [\tau,\infty)\times\partial\Omega, \\ u(\tau,\cdot) = \varphi & \text{in } \overline{\Omega}. \end{cases}$$

Proof. For a fixed cylinder $D = (T_1, T_2) \times \Omega$ we set

$$\Psi_\zeta^D(z) = \int_{\partial_p D} h(\eta;\zeta)\,d\mu_z^D(\eta), \qquad z,\,\zeta \in D,$$

so that $G^D(z;\zeta) = h(z;\zeta) - \Psi_\zeta^D(z)$ solves

$$G^D(\cdot;\zeta) \in C^\infty(D\setminus\{\zeta\})$$
$$H(G^D(\cdot;\zeta)) = 0 \text{ in } D\setminus\{\zeta\}$$
$$G^D(\cdot;\zeta) = 0 \text{ in } \partial_p D$$

(recall (14.5)). Moreover, applying the weak maximum principle for H (see Theorem 13.1) first in the set $(T_1, \tau) \times \Omega$ and then in the set $(\tau, T_2) \times \Omega$, we obtain

$$G^D(t, x; \tau, \xi) = 0 \text{ if } t \leq \tau, \qquad G^D(t, x; \tau, \xi) \geq 0 \text{ if } t \geq \tau. \tag{16.2}$$

We now consider the sequence $D_n = (-n, n) \times \Omega$ and we observe that

$$G^{D_n} = G^{D_{n+1}} \text{ in } (D_n \cup \partial_p D_n) \times D_n.$$

This easily follows from (16.2) and from the weak maximum principle for H applied to the function $G^{D_n}(\cdot;\zeta) - G^{D_{n+1}}(\cdot;\zeta) = \Psi_\zeta^{D_{n+1}} - \Psi_\zeta^{D_n}$ in the set $(\tau, n) \times \Omega$. As a consequence, the following definition of $G = G^\Omega$ is well-posed:

$$G(z;\zeta) = G^{D_n}(z;\zeta), \qquad \text{for } z \in D_n \cup \partial_p D_n, \;\; \zeta \in D_n.$$

Moreover, G solves $H(G(\cdot;\zeta)) = 0$ in $(\mathbb{R} \times \Omega) \setminus \{\zeta\}$, $G(\cdot, \zeta) = 0$ in $\mathbb{R} \times \partial\Omega$. We also define $\Psi = \Psi^\Omega$ by

$$\Psi(z;\zeta) = h(z;\zeta) - G(z;\zeta)$$

(i.e., $\Psi(z;\zeta) = \Psi_\zeta^D(z)$ for $\zeta \in D = (T_1, T_2) \times \Omega$, $z \in D \cup \partial_p D$). We claim that

$$\Psi \in C\left((\mathbb{R} \times \overline{\Omega}) \times (\mathbb{R} \times \Omega)\right). \tag{16.3}$$

Since $\Psi(\cdot;\zeta)$ is continuous on $\mathbb{R} \times \overline{\Omega}$, it is sufficient to prove that, for any $\zeta_0 \in \mathbb{R} \times \Omega$ and $T > 0$,

$$\sup_{z \in [-T,T] \times \overline{\Omega}} |\Psi(z;\zeta) - \Psi(z;\zeta_0)| \longrightarrow 0, \quad \text{as } \zeta \to \zeta_0.$$

Setting for brevity $D = (-T, T) \times \Omega$ and $w_\zeta = \Psi(\cdot;\zeta) - \Psi(\cdot;\zeta_0)$, w_ζ is a solution to $Hw_\zeta = 0$ in D, $w_\zeta = h(\cdot;\zeta) - h(\cdot;\zeta_0)$ in $[-T, T] \times \partial\Omega$, $w_\zeta(-T, \cdot) = 0$ in $\overline{\Omega}$ (if T is chosen large enough). Hence, by the weak maximum principle for H (Theorem 13.1), since

$$|w_\zeta(z)| \leq \sup_{[-T,T] \times \partial\Omega} |h(\cdot;\zeta) - h(\cdot;\zeta_0)| \text{ for } z \in \partial_p D$$

we obtain

$$\sup_{z \in \overline{D}} |w_\zeta(z)| \leq \sup_{z \in [-T,T] \times \partial\Omega} |h(z;\zeta) - h(z;\zeta_0)|$$

which vanishes as $\zeta \to \zeta_0$, since h is continuous away from the diagonal of $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$ (see Proposition 11.4). Thus (16.3) is proved and the continuity of G away from the diagonal of $(\mathbb{R} \times \overline{\Omega}) \times (\mathbb{R} \times \Omega)$ immediately follows. Therefore, the proof of (i) is completed. On the other hand, (ii) directly follows from (16.2) and by observing that $\Psi \geq 0$ by the definition of Ψ_ζ^D. We now turn to the proof of (iii). We first need the following lemma.

Lemma 16.4 *Let $Y \in \{X_i,\, X_iX_j,\, \partial_t \,|i, j = 1, \dots, m\}$ and let O, U be bounded domains of $\mathbb{R}^{n+1}$ such that $O \Subset U \Subset \mathbb{R} \times \Omega$. We have*

$$\|\Psi(\cdot;\zeta)\|_{C^{2,\alpha}(O)} \le \mathbf{c}(O,U) \sup_{z\in U} |\Psi(z;\zeta)| \qquad \forall\, \zeta \in \mathbb{R}\times\Omega, \tag{16.4}$$

$$(z;\zeta) \mapsto Y(\Psi(\cdot;\zeta))(z) \quad \text{is continuous in } (z;\zeta) \in (\mathbb{R}\times\Omega)\times(\mathbb{R}\times\Omega), \tag{16.5}$$

$$\sup_{z\in\overline{O},\ \zeta\in[-T,T]\times\Omega} |Y(\Psi(\cdot;\zeta))(z)| < \infty. \tag{16.6}$$

Proof. >From Theorem 14.1, recalling that $H(\Psi(\cdot;\zeta)) = 0$ in $\Omega \times \mathbb{R}$, (16.4) straightforwardly follows. Since $Y(\Psi(\cdot;\zeta))$ is continuous on $\mathbb{R}\times\Omega$ for every fixed $\zeta \in \mathbb{R}\times\Omega$, in order to prove (16.5) we only have to show that

$$\sup_{z\in\overline{O}} |Y(\Psi(\cdot;\zeta))(z) - Y(\Psi(\cdot;\zeta_0))(z)| \longrightarrow 0 \quad \text{as } \zeta \to \zeta_0 \tag{16.7}$$

for every fixed $\zeta_0 \in \mathbb{R}\times\Omega$ and every bounded domain $O \Subset \mathbb{R}\times\Omega$. To this end, we apply Theorem 14.1 to the function $\Psi(\cdot;\zeta) - \Psi(\cdot;\zeta_0)$ and obtain that the supremum in (16.7) is lower than

$$\|\Psi(\cdot;\zeta) - \Psi(\cdot;\zeta_0)\|_{C^{2,\alpha}(\overline{O})} \le \mathbf{c}(O,\Omega) \sup_{z\in[-T(O),T(O)]\times\overline{\Omega}} |\Psi(z;\zeta) - \Psi(z;\zeta_0)|,$$

which vanishes as $\zeta \to \zeta_0$, by (16.3). This proves (16.5). Let us now prove (16.6). Let O, U, V be bounded domains of $\mathbb{R}^{n+1}$ such that $O \Subset U \Subset V \Subset \mathbb{R}\times\Omega$. Applying Theorem 14.4 to $\Psi(\cdot;\zeta)$ we get

$$\begin{aligned}
\sup_{z\in\overline{O},\ \zeta\in([-T,T]\times\Omega)\setminus\overline{V}} |Y(\Psi(\cdot;\zeta))(z)| &\le \sup_{\zeta\in([-T,T]\times\Omega)\setminus\overline{V}} \|\Psi(\cdot;\zeta)\|_{C^{2,\alpha}(O)} \\
&\le \mathbf{c}(O,U) \sup_{z\in U,\ \zeta\in([-T,T]\times\Omega)\setminus\overline{V}} |\Psi(z;\zeta)| \\
&\le \mathbf{c}(O,U) \max_{z\in\overline{U},\ \zeta\in([-T,T]\times\overline{\Omega})\setminus V} h(z;\zeta) < \infty,
\end{aligned}$$

since h is continuous away from the diagonal of $\mathbb{R}^{n+1}\times\mathbb{R}^{n+1}$. On the other hand, (16.5) directly yields

$$\sup_{z\in\overline{O},\zeta\in\overline{V}} |Y(\Psi(\cdot;\zeta))(z)| < \infty.$$

Thus (16.6) is completely proved. ■

We are now in position to complete the proof of Theorem 16.3, (iii). Let $\varphi \in C(\overline{\Omega})$, $\varphi = 0$ in $\partial\Omega$ and let $\tau \in \mathbb{R}$ be fixed. We agree to extend φ to be zero outside Ω. Then, by means of Theorem 12.1 and (14.5), the function

$$u_1\,(t,x) = \int_{\mathbb{R}^n} h(t,x;\tau,\xi)\,\varphi(\xi)d\xi, \qquad x \in \mathbb{R}^n,\ \ t > \tau,$$

belongs to the class $C^\infty((\tau,\infty)\times\mathbb{R}^n) \cap C([\tau,\infty)\times\mathbb{R}^n)$ and it is a solution to the Cauchy problem $Hu_1 = 0$ in $(\tau,\infty)\times\mathbb{R}^n$, $u_1(\tau,\cdot) = \varphi$ in $\mathbb{R}^n$. On the other hand, recalling (16.3), the fact that $\Psi \le h$, and the continuity of h away from the diagonal, the function

$$u_2(t,x) = \int_\Omega \Psi(t,x;\tau,\xi)\,\varphi(\xi)d\xi$$

is well-defined and continuous in $(\tau,\infty)\times\overline{\Omega}$. Moreover, by means of Lemma 16.4, u_2 has continuous intrinsic-derivatives up to second order along the vector fields $X_1,\ldots,X_m$, and continuous derivative along ∂_t, on $(\tau,\infty)\times\Omega$, obtained by differentiating under the integral sign. As a consequence, $Hu_2 = 0$ in $(\tau,\infty)\times\Omega$. Furthermore, $u_2 \in C^\infty((\tau,\infty)\times\Omega)$ by (14.5). Recalling that $G = h - \Psi$ and that $u = u_1 - u_2$, in order to complete the proof of Theorem 16.3 we are only left to show that

$$u_2(t,x) \longrightarrow 0, \quad \text{as } (t,x)\to(x_0,\tau), \quad \forall\, x_0 \in \overline{\Omega}. \tag{16.8}$$

If $x_0 \in \Omega$ we choose $\delta > 0$ such that $B = \overline{B(x_0,\delta)} \subset \Omega$ and we write

$$|u_2(t,x)| \le \int_B \Psi(t,x;\tau,\xi)\,|\varphi(\xi)|d\xi + \int_{\Omega\setminus B} \Psi(t,x;\tau,\xi)\,|\varphi(\xi)|d\xi.$$

Recalling (16.3) and the fact that $\Psi(s,y;\tau,\xi) = 0$ for $s \le \tau$, the first integral vanishes as $(t,x)\to(\tau,x_0)$; on the other hand, using the estimate $\Psi \le h$, the second integral goes to zero as well. Thus (16.8) holds for $x_0 \in \Omega$. If $x_0 \in \partial\Omega$, we use Theorem 12.1 and we obtain

$$|u_2(t,x)| \le \int_{\mathbb{R}^n} h(t,x;\tau,\xi)\,|\varphi(\xi)|d\xi \longrightarrow |\varphi(x_0)| = 0, \quad \text{as } (t,x)\to(\tau,x_0).$$

This completes the proof. ■

Corollary 16.5 *Let $\Omega \subseteq \mathbb{R}^n$ be an H-regular domain and let G^Ω denote the related Green function as in Theorem 16.3. The following* reproduction property *of G^Ω holds:*

$$G^\Omega(t,x;\tau,\xi) = \int_\Omega G^\Omega(t,x;s,y)\,G^\Omega(s,y;\tau,\xi)dy,$$

for every $t > s > \tau$ and x, $\xi \in \Omega$.

Proof. We fix τ,ξ,s as above and we set $\varphi = G^\Omega(\cdot,s;\tau,\xi)$. Then $\varphi \in C(\overline{\Omega})$, $\varphi = 0$ in $\partial\Omega$, by Theorem 16.3-(i). Therefore we can apply Theorem 16.3-(iii) and obtain that the function

$$u(t,x) = \int_\Omega G^\Omega(t,x;s,y)\,\varphi(y)dy, \qquad x \in \overline{\Omega}, \ \ t > s,$$

satisfies $u \in C^\infty((s,\infty)\times\Omega)\cap C([s,\infty)\times\overline{\Omega})$, $Hu = 0$ in $(s,\infty)\times\Omega$, $u = 0$ in $[s,\infty)\times\partial\Omega$, $u(s,\cdot) = \varphi$ in $\overline{\Omega}$. It is now sufficient to observe that $G^\Omega(\cdot;\tau,\xi)$ has the same properties and to use the weak maximum principle for H (see Theorem 13.1). ■

We now specialize our study of Green functions to cylinders based on regular domains which approximate metric balls. For these Green functions we shall prove Gaussian bounds from above and below.

Fix $\delta_0 \in (0,1)$. By Lemma 16.2, for every $\xi_0 \in \mathbb{R}^n$ and $R > 0$, there exists an H-regular domain $A(\xi_0, R)$ of $\mathbb{R}^n$ such that

$$B(\xi_0, \delta_0 R) \subseteq A(\xi_0, R) \subseteq B(\xi_0, R). \tag{16.9}$$

Then:

Lemma 16.6 *Let $R_0 > 0$ and $\delta \in (0, \delta_0)$. There exists a constant $\rho = \mathbf{c}(\delta, \delta_0, R_0)^{-1} \in (0,1)$, such that*

$$G^{A(\xi_0,R)}(t,x;\tau,\xi) \geq \mathbf{c}(R_0)^{-1}\, \mathbf{E}(x,\xi,\mathbf{c}^{-1}(t-\tau)),$$

for every $\xi_0 \in \mathbb{R}^n$, $R \in (0, R_0]$, $x \in A(\xi_0, R)$, $\xi \in B(\xi_0, \delta R)$, $t, \tau \in \mathbb{R}$ satisfying $d^2(x,\xi) < t - \tau < \rho\, R^2$.

Proof. Let $\xi_0, R, x, \xi, t, \tau$ be as above. Let us set

$$D = (\tau - 1, t+1) \times A(\xi_0, R).$$

Observing that $\mu^D_{(t,x)}((t, t+1] \times \partial A(\xi_0, R)) = 0$ and using (11.11), (11.9), (2.9) and (2.12), we obtain

$$\begin{aligned}
\Psi^{A(\xi_0,R)}(t,x;\tau,\xi) &= \Psi^D_{(\tau,\xi)}(t,x) = \int_{\partial_p D} h(s,y;\tau,\xi)\, \mu^D_{(t,x)}(s,y) \\
&\leq \mathbf{c}(R_0) \int_{[\tau,t]\times\partial A(\xi_0,R)} \mathbf{E}(\xi, y, \mathbf{c}(s-\tau))\, \mu^D_{(t,x)}(s,y) \\
&\leq \mathbf{c}(R_0) \sup_{0<r<t-\tau} |B(\xi,\sqrt{r})|^{-1} \exp\left(-\frac{(\delta_0-\delta)^2 R^2}{\mathbf{c}\, r}\right).
\end{aligned}$$

In the last inequality we have used the fact that $\xi \in B(\xi_0, \delta R)$ and that

$$\mu^D_{(t,x)}(\partial_p D) \equiv 1$$

(recall the operator H is homogeneous). We now exploit Theorem 13.6, (2.12) and (2.9) and obtain

$$\begin{aligned}
&G^{A(\xi_0,R)}(t,x;\tau,\xi) = h(t,x;\tau,\xi) - \Psi^{A(\xi_0,R)}(t,x;\tau,\xi) \\
&\quad \geq \mathbf{c}_1(R_0)^{-1}\, |B(\xi,\sqrt{t-\tau})|^{-1} \exp\left(-\mathbf{c}\, \frac{d^2(x,\xi)}{t-\tau}\right) \times \\
&\quad \times \left[1 - \mathbf{c}_2(R_0) \sup_{0<r<t-\tau} \frac{|B(\xi, R\sqrt{\rho})|}{|B(\xi,\sqrt{r})|} \exp\left(\mathbf{c}\frac{d^2(x,\xi)}{t-\tau} - \frac{R^2}{\mathbf{c}(\delta,\delta_0)\, r}\right)\right] \\
&\quad \geq \mathbf{c}_3(R_0)^{-1}\, \mathbf{E}(x,\xi,\mathbf{c}^{-1}(t-\tau)) \times \\
&\quad \times \left[1 - \mathbf{c}_4(R_0) \sup_{0<r<t-\tau} \left(\frac{\rho\, R^2}{r}\right)^{Q/2} \exp\left(-\frac{R^2}{\mathbf{c}(\delta,\delta_0)\, r}\right)\right].
\end{aligned}$$

It is now sufficient to observe that the expression between square brackets is greater than 1/2 if $\rho = \rho(\delta, \delta_0, R_0)$ is small enough, as one can easily recognize by showing that the function

$$h \mapsto (\rho\, h)^{Q/2} \exp(-h/\mathbf{c}(\delta, \delta_0))$$

is monotone decreasing on the interval $[\rho^{-1}, \infty)$. ■

Theorem 16.7 *Let $R_0 > 0$, $T > 1$ and $\gamma \in (0, \delta_0)$. We have*

$$G^{A(\xi_0,R)}(t, x; \tau, \xi) \geq \mathbf{c}(T, \gamma, R_0, \delta_0)^{-1}\, |B(\xi, \sqrt{t-\tau})|^{-1}\, \exp\left(-\mathbf{c}(\gamma, R_0, \delta_0)\, \frac{d^2(x,\xi)}{t-\tau}\right),$$

for every $\xi_0 \in \mathbb{R}^n$, $R \in (0, R_0]$, $x, \xi \in B(\xi_0, \gamma R)$, and $0 < t - \tau < T\, R^2$.

Proof. The way Theorem 16.7 follows from Lemma 16.6 is similar to the way Theorem 7.1 follows from Lemma 7.2. Compare also with the proof of Theorem 13.6.

We set $\delta = (\gamma + \delta_0)/2$ and choose $\rho = \rho(\delta, \delta_0, R_0)$ as in Lemma 16.6. Let us also fix $\xi_0, R, x, \xi, t, \tau$ as above. Let k be the smallest integer greater than

$$M(\gamma, \delta_0)\, \max\left\{T/\rho,\, d^2(x,\xi)/(t-\tau)\right\},$$

where the constant $M(\gamma, \delta_0) > 1$ will be chosen later, and let us set

$$\sigma = \frac{1}{4}\sqrt{(t-\tau)/(k+1)}.$$

We claim that there exists a chain of points of $\mathbb{R}^n$ $x = x_0, x_1, \ldots, x_{k+1} = \xi$ such that

$$d(x_j, x_{j+1}) \leq \mathbf{c}(\gamma, \delta_0)\, \frac{d(x,\xi)}{k+1}, \qquad d(x_j, \xi_0) \leq \tfrac{\gamma+\delta}{2}\, R. \tag{16.10}$$

Indeed, if $d(x,\xi) \leq R(\delta_0 - \gamma)/8$, we can choose $x_1, \ldots, x_k$ laying on a suitable X-subunit path connecting x and ξ, so that $d(x_j, x_{j+1}) \leq 2d(x,\xi)/(k+1)$ and

$$d(x_j, \xi_0) \leq d(x, x_j) + d(x, \xi_0) \leq 2d(x,\xi) + d(x,\xi_0) \leq R(\delta_0 - \gamma)/4 + \gamma R = R(\delta+\gamma)/2.$$

On the other hand, if $d(x,\xi) > R(\delta_0 - \gamma)/8$, then we can choose $x_1, \ldots, x_k$ laying on suitable X-subunit paths connecting x with ξ_0 and ξ_0 with ξ, so that

$$d(x_j, x_{j+1}) \leq 2\gamma R/(k+1) < 16\gamma d(x,\xi)(\delta_0 - \gamma)^{-1}(k+1)^{-1}$$

and $d(x_j, \xi_0) \leq \gamma R$. Observing that, by the definition of k and σ, we have $\sigma \leq M(\gamma, \delta_0)^{-1/2} R$, from (16.10) it follows that we can choose $M(\gamma, \delta_0)$ such that

$$B(x_j, \sigma) \subseteq B(\xi_0, \delta R). \tag{16.11}$$

Moreover, up to a new choice of $M(\gamma, \delta_0)$, we also have

$$d(y_j, y_{j+1}) < \sqrt{\frac{t-\tau}{k+1}} \qquad \text{for every } y_j \in B(x_j, \sigma),\ y_{j+1} \in B(x_{j+1}, \sigma). \tag{16.12}$$

Indeed, from (16.10) and the definition of k it follows that

$$d(y_j, y_{j+1}) \le 2\sigma + d(x_j, x_{j+1}) \le \frac{1}{2}\sqrt{\frac{t-\tau}{k+1}} + \mathbf{c}(\gamma, \delta_0)\,\frac{d(x,\xi)}{k+1}$$
$$\le \sqrt{\frac{t-\tau}{k+1}}\,\left(\tfrac{1}{2} + \mathbf{c}(\gamma, \delta_0)\, M(\gamma, \delta_0)^{-1/2}\right).$$

Let now $t = t_0, t_1, \dots, t_{k+1} = \tau$ be such that $t_j - t_{j+1} = (t-\tau)/(k+1)$ for $j = 0, \dots, k$. Using Corollary 16.5 repeatedly, we obtain (we set $G = G^{A(\xi_0,R)}$, $y_0 = x$, $y_{k+1} = \xi$)

$$G(t, x; \tau, \xi) = \int_{(A(\xi_0,R))^k} G(t, x; y_1, t_1)\, G(y_1, t_1; y_2, t_2) \cdots G(y_k, t_k; \tau, \xi)\, y_1 \cdots y_k$$
$$\ge \int_{\prod_{j=1}^k B(x_j,\sigma)} \prod_{j=0}^{k} G(y_j, t_j; y_{j+1}, t_{j+1})\, y_1 \cdots y_k,$$

by 16.11. Moreover, from (16.11), (16.12) and the definition of k, it follows that $y_{j+1} \in B(\xi_0, \delta R)$,

$$d^2(y_j, y_{j+1}) < (t-\tau)/(k+1) = t_j - t_{j+1} < TR^2/(k+1) < \rho R^2.$$

Therefore, we can apply Lemma 16.6 and obtain (arguing as in the proof of Theorem 13.6)

$$G(t, x; \tau, \xi) \ge \mathbf{c}(R_0)^{-(k+1)} \int_{\prod_{j=1}^k B(x_j,\sigma)} \prod_{j=0}^{k} \mathbf{E}(y_j, y_{j+1}, \mathbf{c}^{-1}(t_j - t_{j+1}))dy_1 \cdots dy_k$$
$$\ge \mathbf{c}(R_0)^{-1}\, \exp(-\mathbf{c}(R_0)\, k) \frac{1}{|B(x, \mathbf{c}\sigma)|} \prod_{j=1}^{k} \frac{|B(x_j, \sigma)|}{|B(x_j, \mathbf{c}\sigma)|}$$
$$\ge \mathbf{c}(R_0)^{-1} |B(x, \mathbf{c}\sigma)|^{-1}\, \exp(-\mathbf{c}(R_0)\, k)$$
$$\ge \mathbf{c}(R_0)^{-1} |B(x, \sqrt{t-\tau})|^{-1}\, \exp(-\mathbf{c}(R_0)\, k),$$

by the definition of σ. Now, if $d^2(x,\xi) \ge T(t-\tau)/\rho$, from the definition of k it follows that $k < \mathbf{c}(\gamma, \delta_0)\, d^2(x,\xi)/(t-\tau)$ and then

$$G(t, x; \tau, \xi) \ge \mathbf{c}(R_0)^{-1} |B(x, \sqrt{t-\tau})|^{-1}\, \exp\left(-\mathbf{c}(\gamma, R_0, \delta_0)\, \frac{d^2(x,\xi)}{t-\tau}\right).$$

On the other hand, if $d^2(x,\xi) < T(t-\tau)/\rho$, the definition of k gives $k < \mathbf{c}(\gamma, \delta_0)T/\rho$ and then

$$G(t, x; \tau, \xi) \ge \mathbf{c}(T, \gamma, R_0, \delta_0)^{-1} |B(x, \sqrt{t-\tau})|^{-1}.$$

This completes the proof of Theorem 16.7. ■

Remark 16.8 *All the results proved in this section actually hold for the* complete *operator H defined in (10.1) (i.e. with a smooth, but not necessarily vanishing, coefficient a_0). Here we have assumed $a_0 \equiv 0$ only for consistence with the rest of Part III. The only points where the presence of a_0 would require slight changes in our arguments are the following:*

1. In the final part of the proof of Theorem 16.3, if a_0 is not zero we should compare w_ζ with the barrier

$$v\,(t,x) = \exp(k(T+t)) \sup_{[-T,T]\times\partial\Omega} |h(\cdot;\zeta) - h(\cdot;\zeta_0)|.$$

2. In the proof of Lemma 16.6, we used the fact that $\mu^D_{(t,x)}(\partial_p D) \equiv 1$; if a_0 is not zero, we would have instead to prove the boundedness of the function $(s,y) \mapsto \mu^D_{(s,y)}(\partial_p D)$ comparing it with $(s,y) \mapsto \exp(k(s-\tau+1))$ by the weak maximum principle (Theorem 13.1).

17 Harnack inequality for operators with smooth coefficients

Also in this section we consider operators with smooth coefficients. We keep the notation introduced in the previous Section; in particular, recall that the regular domain $A(\xi_0, R)$, for which in Section 16 we have estimated the Green function $G^{A(\xi_0,R)}$, depends on the choice of a number $\delta_0 \in (0,1)$ (see (16.9)). Several results in this Section involve this number δ_0, which will be eventually chosen in a suitable way.

Lemma 17.1 *Let $R_0 > 0$ and $\gamma \in (0, \delta_0)$. There exists a constant $\mu \in (0,1)$, $\mu = \mathbf{c}(\gamma, R_0, \delta_0)$, such that for every $(\tau_0, \xi_0) \in \mathbb{R}^{n+1}$, $R \in (0, R_0]$ and every*

$$u \in C^\infty((\tau_0 - R^2, \tau_0) \times B(\xi_0, R)) \cap C([\tau_0 - R^2, \tau_0] \times \overline{B(\xi_0, R)})$$

satisfying $Hu = 0$ in $(\tau_0 - R^2, \tau_0) \times B(\xi_0, R)$, we have

$$\operatorname*{osc}_{[\tau_0-\gamma^2 R^2, \tau_0]\times\overline{B(\xi_0,\gamma R)}} u \le \mu \operatorname*{osc}_{[\tau_0-R^2,\tau_0]\times\overline{B(\xi_0,R)}} u. \tag{17.1}$$

Proof. We set

$$D = (\tau_0 - R^2, \tau_0) \times A(\xi_0, R)$$

and

$$S = \left\{x \in B(\xi_0, \gamma R) \,|\, u(x, \tau_0 - R^2) \ge (M+m)/2\right\},$$

where $M = \max_{\overline{D}} u$, $m = \min_{\overline{D}} u$. We also define $w = u - m$ and (for $x \in \overline{A(\xi_0, R)}$, $t > \tau_0 - R^2$)

$$v\,(t,x) = \int_{A(\xi_0,R)} G^{A(\xi_0,R)}(t,x;\tau_0 - R^2, y)\, w(\tau_0 - R^2, y)\, \varphi(y) d\, y,$$

where $\varphi \in C_0(A(\xi_0, R))$ is a cut-off function such that $0 \le \varphi \le 1$ and $\varphi \equiv 1$ in $B(\xi_0, \gamma R)$. By means of Theorem 16.3-(iii), v is a solution to $Hv = 0$ in D, $v = 0$ in $[\tau_0 - R^2, \tau_0] \times \partial A(\xi_0, R)$, $v(\cdot, \tau_0 - R^2) = w(\tau_0 - R^2, \cdot)\, \varphi$ in $\overline{A(\xi_0, R)}$. Moreover $Hw = 0$ since we are supposing that $a_0 = 0$. Therefore, by the weak maximum principle for H (see Theorem 13.1), $v \le w$ in D. As a consequence, using the estimate in Theorem 16.7, for every

$$(t, x) \in D_\gamma = (\tau_0 - \gamma^2 R^2, \tau_0) \times B(\xi_0, \gamma R)$$

we get

$$\begin{aligned} w\,(t,x) \ge v\,(t,x) &\ge \int_S G^{A(\xi_0,R)}(t, x; \tau_0 - R^2, y)\ \left(\tfrac{M+m}{2} - m\right) dy \\ &\ge \int_S \frac{\mathbf{c}}{\left|B\left(x, \sqrt{t - \tau + R^2}\right)\right|} e^{-d(x,y)^2/\mathbf{c}\left(t-\tau+R^2\right)}\, \tfrac{M-m}{2} dy \\ &\ge \mathbf{c}(\gamma, R_0, \delta_0)^{-1} |B(x, R)|^{-1}\ \tfrac{M-m}{2}\ |S| \\ &\ge \mathbf{c}(\gamma, R_0, \delta_0)^{-1} |B(\xi_0, R)|^{-1}\ \tfrac{M-m}{2}\ |S|. \end{aligned}$$

Now, if $|S| \ge \frac{1}{2}|B(\xi_0, \gamma R)|$, we infer that

$$\min_{\overline{D}_\gamma} u - m \ge \mathbf{c}(\gamma, R_0, \delta_0)^{-1}(M - m)$$

and then

$$\begin{aligned} \operatorname{osc}_{\overline{D}_\gamma} u \le M - \min_{\overline{D}_\gamma} u &\le M - m - \mathbf{c}(\gamma, R_0, \delta_0)^{-1}(M - m) \\ &= \left(1 - \mathbf{c}(\gamma, R_0, \delta_0)^{-1}\right)\ \operatorname*{osc}_{\overline{D}} u. \end{aligned}$$

Recalling that $A(\xi_0, R) \subseteq B(\xi_0, R)$, we have proved (17.1) when $|S| \ge \frac{1}{2}|B(\xi_0, \gamma R)|$. On the other hand, if $|S| < \frac{1}{2}|B(\xi_0, \gamma R)|$, the argument above can be applied to $\widetilde{u} := -u$, since (with the natural notation) $|\widetilde{S}| > \frac{1}{2}|B(\xi_0, \gamma R)|$. As a consequence, we get (17.1) for $\widetilde{u}$ and the proof is completed, since $\operatorname{osc} \widetilde{u} = \operatorname{osc} u$. ∎

Lemma 17.2 *Let $R_0 > 0$, $h \in (0, 1)$ and $\gamma \in (0, \delta_0)$. There exists a positive constant $\beta = \mathbf{c}(h, \gamma, R_0, \delta_0)$ such that*

$$\sup_{\sigma>0,\, s\in[\tau_0-R^2, \tau_0-hR^2]} \sigma\ |\{y \in B(\xi_0, \gamma R) \,|\, u(s, y) \ge \sigma\}| \le \beta\, |B(\xi_0, R)|\, u(\xi_0, \tau_0)$$

for every $(\xi_0, \tau_0) \in \mathbb{R}^{n+1}$, $R \in (0, R_0]$ and every

$$u \in C^\infty(B(\xi_0, R) \times (\tau_0 - R^2, \tau_0)) \cap C(\overline{B(\xi_0, R)} \times [\tau_0 - R^2, \tau_0])$$

satisfying $Hu = 0$, $u \ge 0$ in $B(\xi_0, R) \times (\tau_0 - R^2, \tau_0)$.

Proof. Let us fix $s \in [\tau_0 - R^2, \tau_0 - hR^2]$; set $A = A(\xi_0, R)$ and

$$w(t,x) = \int_A G^A(t,x;s,y)\, u(s,y)\, \varphi(y)\, dy, \qquad x \in \overline{A},\ \ t > s,$$

where $\varphi \in C_0(A)$ is a cut-off function such that $0 \le \varphi \le 1$ and $\varphi \equiv 1$ in $B(\xi_0, \gamma R)$. By means of Theorem 16.3-(iii), w is a solution to $Hw = 0$ in $(s, \infty) \times A$, $w = 0$ in $[s, \infty) \times \partial A$, $w(s, \cdot) = u(s, \cdot)\, \varphi$ in $\overline{A}$. Therefore, by the weak maximum principle for H, $w \le u$ in $[s, \tau_0] \times \overline{A}$. As a consequence, using the estimate in Theorem 16.7, we obtain

$$u(\tau_0, \xi_0) \ge w(\tau_0, \xi_0) \ge \int_A \frac{\mathbf{c} u(s,y)\, \varphi(y)}{|B(\xi_0, \sqrt{\tau_0 - s})|} e^{-d(\xi_0, y)^2 / \mathbf{c}(\tau_0 - s)} dy$$

since $\tau_0 - s \ge hR^2$ and $d(\xi_0, y) \le \mathbf{c}R$

$$\ge \frac{\mathbf{c}(h, \gamma, R_0, \delta_0)}{|B(\xi_0, R)|} \int_{B(\xi_0, \gamma R)} u(s,y)\, dy$$
$$\ge \frac{\mathbf{c}(h, \gamma, R_0, \delta_0)}{|B(\xi_0, R)|} \sigma\ |\{y \in B(\xi_0, \gamma R) \,|\, u(s,y) \ge \sigma\}|$$

for any σ. This ends the proof. ■

Next theorem follows from the previous two Lemmas, with the same technique used in [22]. We will present a detailed proof, anyhow.

Theorem 17.3 *Let $R_0 > 0$, $0 < h_1 < h_2 < 1$ and $\gamma \in (0,1)$. There exists a positive constant $M = \mathbf{c}(h_1, h_2, \gamma, R_0)$ such that for every $(\xi_0, \tau_0) \in \mathbb{R}^{n+1}$, $R \in (0, R_0]$ and every*

$$u \in C^\infty((\tau_0 - R^2, \tau_0) \times B(\xi_0, R)) \cap C([\tau_0 - R^2, \tau_0] \times \overline{B(\xi_0, R)})$$

satisfying $Hu = 0$, $u \ge 0$ in $(\tau_0 - R^2, \tau_0) \times B(\xi_0, R)$, we have

$$\max_{[\tau_0 - h_2 R^2, \tau_0 - h_1 R^2] \times \overline{B(\xi_0, \gamma R)}} u \le M\, u(\xi_0, \tau_0). \tag{17.2}$$

Proof of Theorem 17.3. Recall that our previous estimates depend on a number $\delta_0 \in (0,1)$ which can be arbitrarily chosen (see (16.9)). Then, for any fixed $\gamma \in (0,1)$, pick $\delta_0 = (1+\gamma)/2$. We will apply all our previous results with this particular choice of δ_0.

Let $\mu = \mu(\frac{\delta_0}{2}, R_0, \delta_0) \in (0,1)$, $\beta = \beta(h_1, \frac{\gamma + \delta_0}{2}, R_0, \delta_0) > 0$ be as in Lemma 17.1 and Lemma 17.2 respectively. Let $\mathbf{c}_0$ be as in (2.9). We define $r : (0, \infty) \to (0, \infty)$,

$$r(\sigma) = 2 \left(\frac{4\, \beta \mathbf{c}_0}{\sigma(1 - \mu)} \right)^{1/Q}$$

and we set

$$K = (1 + \mu^{-1})/2$$

and

$$M = r^{-1}\left(\delta_0\,(1-h_2)\,(\delta_0-\gamma)\,(1-K^{-1/Q})/4\right).$$

We now argue by contradiction and suppose that there exist ξ_0, τ_0, R and u satisfying the hypotheses of the theorem, for which (17.2) is not true (with the above choice of M). We first observe that $u(\tau_0, \xi_0) \neq 0$, since otherwise (17.2) would follow from Lemma (17.2). Let now $v = u/u(\tau_0, \xi_0)$. Since v is bounded, in order to get a contradiction and thus prove the theorem, it is sufficient to show that there exists a sequence of points $\{(s_j, y_j)\}_{j\in\mathbb{N}}$ in $[\tau_0 - R^2, \tau_0] \times \overline{B(\xi_0, R)}$ such that

$$v(s_j, y_j) \geq K^j M, \qquad (s_j, y_j) \in [\tau_0 - R^2, \tau_0] \times \overline{B(\xi_0, R)}.$$

Indeed, recalling that $K > 1$, this would give $v(s_j, y_j) \to \infty$. To construct this sequence, we will prove by induction the existence of points $(s_j, y_j) \in [\tau_0 - R^2, \tau_0] \times \overline{B(\xi_0, R)}$ such that:

$$\begin{aligned} &v(s_j, y_j) \geq K^j M \\ &(s_0, y_0) \in [\tau_0 - h_2 R^2, \tau_0 - h_1 R^2] \times \overline{B(\xi_0, \gamma R)} \\ &(s_j, y_j) \in [s_{j-1} - \rho_{j-1}^2, s_{j-1}] \times \overline{B(y_{j-1}, \rho_{j-1})} \text{ if } j \geq 1 \\ &\text{with } \rho_j = 2\delta_0^{-1} r(K^j M) R \end{aligned} \tag{17.3}$$

The existence of $(s_0, y_0) \in [\tau_0 - h_2 R^2, \tau_0 - h_1 R^2] \times \overline{B(\xi_0, \gamma R)}$ such that $v(s_0, y_0) \geq M$ follows from the assumption that u does not satisfy (17.2). We now suppose that, for a fixed $q \in \mathbb{N}$, $(s_0, y_0), \ldots, (s_q, y_q)$ have been defined and satisfy (17.3) for every $j \in \{0, \ldots, q\}$. We have to prove that we can find (s_{q+1}, y_{q+1}) satisfying (17.3) for $j = q + 1$. We claim that

$$\overline{B(y_q, \rho_q)} \subseteq B(\xi_0, (\gamma + \delta_0)R/2). \tag{17.4}$$

Indeed, if $d(y, y_q) \leq \rho_q$, then recalling the definition of M and using (17.3) for $j \in \{0, \ldots, q\}$, we obtain

$$\begin{aligned} d(y, \xi_0) &\leq d(\xi_0, y_0) + \sum_{j=1}^{q} d(y_{j-1}, y_j) + d(y_q, y) \\ &\leq \gamma R + 2\delta_0^{-1} R \sum_{i=0}^{q} r(K^i M) < \gamma R + 2\delta_0^{-1} r(M)\, R \sum_{i=0}^{\infty} K^{-i/Q} \\ &= (\gamma + (1 - h_2)(\delta_0 - \gamma)/2)\, R < (\gamma + \delta_0)R/2. \end{aligned}$$

Moreover, with a similar computation we can prove that

$$[s_q - \rho_q^2, s_q] \subseteq (\tau_0 - R^2, \tau_0 - h_1 R^2]. \tag{17.5}$$

Indeed, $s_q \le s_{q-1} \le \cdots \le s_0 \le \tau_0 - h_1 R^2$ and

$$s_q - \rho_q^2 = s_0 + \sum_{j=1}^{q} (s_j - s_{j-1}) - \rho_q^2$$
$$> \tau_0 - h_2 R^2 - 4\delta_0^{-2} (r(M))^2 R^2 \sum_{i=0}^{\infty} K^{-2i/Q}$$
$$> \tau_0 - \Big(h_2 + (1-h_2)(1-K^{-1/Q})(1+K^{-1/Q})^{-1} \Big) \, R^2 > \tau_0 - R^2.$$

We now apply Lemma 17.2 (with $\sigma = (1-\mu)K^q M/2$) to v and we obtain (recalling (17.5) and the definition of r)

$$|\{y \in B(\xi_0, (\gamma+\delta_0)R/2) \,|\, v(y, s_q) \ge (1-\mu)K^q M/2\}|$$
$$\le \frac{2\beta |B(\xi_0, R)|}{(1-\mu)K^q M} = \frac{1}{2} \mathbf{c}_0^{-1} \left(\frac{r(K^q M)}{2} \right)^Q |B(\xi_0, R)|$$
$$< \mathbf{c}_0^{-1} \left(\frac{r(K^q M)}{2} \right)^Q |B(y_q, 2R)| \le |\{B(y_q, r(K^q M)R)\}| \, .$$

In the last inequality we have used (2.9) and the fact that $r(K^q M) \le 2$, which follows from the definition of M (see the proof of (17.4)). As a consequence, since also (17.4) holds, there exists

$$\overline{y} \in B(y_q, r(K^q M)R)$$

such that

$$v(s_q, \overline{y}) < (1-\mu)K^q M/2.$$

Therefore, recalling that we are supposing that (17.3) holds for $j = q$, we have

$$(1+\mu)K^q M/2 = K^q M - (1-\mu)K^q M/2 < v(s_q, y_q) - v(s_q, \overline{y})$$
$$\le \underset{\{s_q\} \times B(y_q, r(K^q M)R)}{\operatorname{osc}} v \le \mu \underset{[s_q - \rho_q^2, s_q] \times B(y_q, \rho_q)}{\operatorname{osc}} v$$

by means of Lemma 17.1, (17.4) and (17.5) (note that $\rho_q \le R_0$ by the definition of M). Since $v \ge 0$, it follows that there exists

$$(s_{q+1}, y_{q+1}) \in [s_q - \rho_q^2, s_q] \times \overline{B}(y_q, \rho_q)$$

such that

$$v(s_{q+1}, y_{q+1}) > \mu^{-1}(1+\mu)K^q M/2 = K^{q+1} M.$$

This completes the proof of Theorem 17.3. ∎

18 Harnack inequality in the non-smooth case

In this section we complete the proof of the Harnack inequality in Theorem 15.1. Throughout the section the coefficients a_{ij}, a_k of H will not be supposed to be smooth as in Sections 16-17, but only Hölder continuous according to (10.4). We shall approximate H by suitable smooth coefficient operators

$$H_\varepsilon = \partial_t - \sum_{i,j=1}^{m} a_{ij}^{\varepsilon}(t,x)\, X_i X_j - \sum_{k=1}^{m} a_k^{\varepsilon}(t,x)\, X_k,$$

where $a_{i,j}^{\varepsilon}$ and a_k^{ε} are regularized versions of a_{ij} and a_k as in Theorem 14.2. The following result has been proved in [39].

Lemma 18.1 *For every $\xi_0 \in \mathbb{R}^n, R > 0$ and $\delta \in (0,1)$, there exists a domain A, H_ε-regular for any $\varepsilon > 0$, such that*

$$B(\xi_0, \delta R) \subseteq A \subseteq B(\xi_0, R)$$

and with the following property: for any bounded cylinder $D = (T_1, T_2) \times A$ and for every $\varphi \in C(\partial_p D)$, letting u_ε be the solution to $H_\varepsilon u_\varepsilon = 0$ in D, $u_\varepsilon = \varphi$ in $\partial_p D$, we have

$$|u_\varepsilon(z) - \varphi(z_0)| \to 0, \qquad \text{as } z \to z_0, \text{ uniformly in } 0 < \varepsilon < 1, \tag{18.1}$$

for every $z_0 \in \partial_p D$.

We are finally in position to conclude the proof of our invariant Harnack inequality for H.

Proof of Theorem 15.1. Chosen

$$\delta = (\max\{\gamma, h_2^{1/4}\} + 1)/2,$$

let A be as in Lemma 18.1. We set $D = (\tau_0 - R^2, \tau_0) \times A$, $\varphi = u|_{\partial_p D}$ and we denote by u_ε the solution to $H_\varepsilon u_\varepsilon = 0$ in D, $u_\varepsilon = \varphi$ in $\partial_p D$. With this notation, (18.1) holds, thanks to Lemma 18.1. On the other hand, applying to H_ε the *a-priori* estimates in Theorem 14.1 for every ε, and the maximum principle in Theorem 13.1, we obtain, for every bounded domain $O \Subset D$,

$$\|u_\varepsilon\|_{C^{2,\alpha}(O)} \le \mathbf{c}(O,D) \sup_D |u_\varepsilon| \le \mathbf{c}(O,D) \max_{\partial_p D} |u|.$$

By Lemma 14.3, there exists $v \in C^{2,\alpha}_{loc}(D)$ satisfying $Hv = 0$ in D such that for some $\varepsilon_k \to 0$, $u_{\varepsilon_k} \to v$ uniformly on any compact subset of D. Let us show that $v|_{\partial_p D} = \varphi$. For any $z_0 \in \partial D$, we have:

$$|v(z) - \varphi(z_0)| \le |v(z) - u_{\varepsilon_k}(z)| + |u_{\varepsilon_k}(z) - \varphi(z_0)|.$$

By (18.1),

$$|u_{\varepsilon_k}(z) - \varphi(z_0)| < \eta \text{ whenever } d(z, z_0) < \rho \text{ (uniformly in } \varepsilon_k)$$

while, for any such z we have

$$|v(z) - u_{\varepsilon_k}(z)| < \eta$$

for ε_k small enough (possibly depending on z). Hence $v|_{\partial_p D} = \varphi$, and from the maximum principle for H it follows that $v = u$ in D. In particular u_{ε_k} uniformly converges to u on the compact subsets of D. We now want to apply Theorem 17.3 to u_{ε_k} (recall that $u_\varepsilon \geq 0$ since $\varphi \geq 0$). Since $Hu_{\varepsilon_k} = 0$ in $\left(\tau_0 - R^2, \tau_0\right) \times A$, we note that $u_{\varepsilon_k}(\tau_0, \cdot)$ could be discontinuous; hence, to check the assumptions of Theorem 17.3 we have to consider u_{ε_k} on the cylinder

$$(\tau_\sigma - R'^2, \tau_\sigma) \times B(\xi_0, R'),$$

where $R' = \delta R$, $\tau_\sigma = \tau_0 - \sigma R'^2$ and $\sigma > 0$ is small. Let also $h_2' = \sqrt{h_2}$ and $\gamma' = \gamma/\delta$. By our choice of the parameters, we have:

$$\begin{aligned} \gamma' &< 1; \\ \gamma' R' &= \gamma R; \\ h_2' R'^2 &> h_2 R^2; \\ h_1 R'^2 &< h_1 R^2; \end{aligned}$$

therefore

$$\begin{aligned} \max_{[\tau_\sigma - h_2 R^2, \tau_\sigma - h_1 R^2] \times \overline{B(\xi_0, \gamma R)}} u_{\varepsilon_k} &\leq \max_{\left[\tau_\sigma - h_2' R'^2, \tau_\sigma - h_1 R'^2\right] \times \overline{B(\xi_0, \gamma' R')}} u_{\varepsilon_k} \\ &\leq \mathbf{c}(h_1, h_2, \gamma, R_0)\, u_{\varepsilon_k}(\tau_\sigma, \xi_0) \end{aligned}$$

by Theorem 17.3. Letting first k go to infinity and then σ go to zero, from the above inequality we finally get

$$\max_{[\tau_0 - h_2 R^2, \tau_0 - h_1 R^2] \times \overline{B(\xi_0, \gamma R)}} u \leq \mathbf{c}(h_1, h_2, \gamma, R_0)\, u(\tau_0, \xi_0).$$

■

The above theorem easily implies also the analog stationary version.

Proof of Theorem 15.3. Since u is also a nonnegative solution to $Hu = 0$ in $\mathbb{R} \times B(\xi_0, 3R)$, by Theorem 15.1 we know that

$$\max_{\overline{B(\xi_0, R)}} u \leq M u(\xi_0). \tag{18.2}$$

Now, let $x_0 \in \overline{B(\xi_0, R)}$ such that $u(x_0) = \min_{\overline{B(\xi_0, R)}} u$, let $r = d(x_0, \xi_0)$ and $r' = (r + R)/2$. Then $\xi_0 \in \overline{B(x_0, r')}$, $B(x_0, 2r') \subset B(\xi_0, 3R)$, hence again by Theorem 15.1 we have

$$u(\xi_0) \leq \max_{\overline{B(x_0, r')}} u \leq M u(x_0) = M \min_{\overline{B(\xi_0, R)}} u$$

and

$$\max_{\overline{B(\xi_0, R)}} u \leq M^2 \min_{\overline{B(\xi_0, R)}} u.$$

■

Epilogue

19 Applications to operators which are defined only locally

As we have already pointed out, the final goal of our theory is to establish some local properties for an operator H defined in some bounded domain of $\mathbb{R}^{n+1}$, deducing them from the properties of the globally defined operators that we have considered so far. We come at last at this point.

Assume H is an operator of type

$$H = \partial_t - L = \partial_t - \sum_{i,j=1}^{q} a_{ij}(t,x) Z_i Z_j - \sum_{k=1}^{q} a_k(t,x) Z_k - a_0(t,x)$$

where $Z_1, Z_2, ..., Z_q$ are smooth Hörmander vector fields in some bounded domain Ω of $\mathbb{R}^n$; let d_Z be the CC-distance induced by the Z_i's in Ω, and d_{ZP} the corresponding parabolic CC-distance defined in $\mathbb{R} \times \Omega$. Assume that the coefficients a_{ij}, a_k, a_0 are bounded and d_{ZP}-Hölder continuous in a cylinder $\mathcal{C} = (T_1, T_2) \times \Omega$ $(-\infty \leq T_1 < T_2 \leq \infty)$, and let $\{a_{ij}\}$ satisfy the following uniform positive definiteness condition in the same cylinder:

$$\lambda^{-1} |w|^2 \leq \sum_{i,j=1}^{q} a_{ij}(t,x) w_i w_j \leq \lambda |w|^2 \quad \forall w \in \mathbb{R}^q, (t,x) \in (T_1, T_2) \times \Omega. \tag{19.1}$$

Then:

Theorem 19.1 (Local fundamental solution for H) *Under the above assumptions, for any domain $\Omega' \Subset \Omega$, there exists a function*

$$h : \mathbb{R}^{n+1} \times \mathbb{R}^{n+1} \to \mathbb{R}$$

such that, setting $\mathcal{C}' = (T_1, T_2) \times \Omega'$, we have:

i) h is continuous away from the diagonal of $\mathbb{R}^{n+1} \times \mathbb{R}^{n+1}$;

ii) $h(z, \zeta)$ is nonnegative, and vanishes for $t \leq \tau$;

iii) for every fixed $\zeta \in \mathbb{R}^{n+1}$, we have

$$h(\cdot;\zeta) \in C^{2,\alpha}_{loc}(\mathcal{C}' \setminus \{\zeta\}), \quad H(h(\cdot;\zeta)) = 0 \text{ in } \mathcal{C}' \setminus \{\zeta\};$$

iv) there exists $T > 0$, depending on Ω', such that the following estimates hold for every $z = (t,x)$, $\zeta = (\tau, \xi) \in \mathcal{C}'$, $0 < t - \tau \leq T$:

$$c^{-1} E(x, \xi, c^{-1}(t-\tau)) \leq h(z;\zeta) \leq c\, E(x, \xi, c(t-\tau)),$$
$$|Z_j(h(\cdot;\zeta))(z)| \leq c(t-\tau)^{-1/2} E(x, \xi, c(t-\tau));$$
$$|Z_i Z_j(h(\cdot;\zeta))(z)| + |\partial_t(h(\cdot;\zeta))(z)| \leq c(t-\tau)^{-1} E(x, \xi, c(t-\tau));$$

where

$$E(x,\xi,t) = |B_{d_Z}(x,\sqrt{t})|^{-1} \exp\left(-\frac{d_Z(x,\xi)^2}{t}\right)$$

and c is a positive constant that depends only on the vector fields Z_i, the Hölder norms of the coefficients, the number λ and the domains Ω, Ω';

v) for any $f \in C_0^\alpha(\mathcal{C}')$, the function

$$u(t,x) = \int_{\mathbb{R}^{n+1}} h(t,x;\tau,\xi)\, f(\tau,\xi) d\tau d\xi$$

belongs to $C_{loc}^{2,\alpha}(\mathcal{C}')$ and solves the equation

$$Hu = f \text{ in } \mathcal{C}'$$

vi) the following reproduction formula holds

$$h(t,x;\tau,\xi) = \int_{\mathbb{R}^n} h(t,x;s,y)\, h(s,y;\tau,\xi) dy,$$

for $t > s > \tau$ and $x, \xi \in \mathbb{R}^n$.

In the above theorem the Hölder spaces are taken with respect to $Z_1, \ldots, Z_m$ and $(d_Z)_P$ analogously as in Definition 10.4.

Proof. For $\Omega' \Subset \Omega$ fixed, we choose domains Ω_i such that

$$\Omega' \Subset \Omega_2 \Subset \Omega_1 \Subset \Omega_0 \Subset \Omega.$$

By Theorem 2.9, there exists a new system $X = (X_1, X_2, ..., X_m)$ $(m = q+n)$ of vector fields, such that the X_i's are defined on the whole space $\mathbb{R}^n$ and satisfy Hörmander's condition in $\mathbb{R}^n$; moreover:

$$X = (Z_1, Z_2, ..., Z_q, 0, 0, ..., 0) \text{ in } \Omega_1;$$
$$X = (0, 0, ..., 0, \partial_{x_1}, \partial_{x_2}, ..., \partial_{x_n}) \text{ in } \mathbb{R}^n \setminus \Omega_0;$$
$$d_X \text{ is equivalent to } d_Z \text{ in } \Omega_2.$$

Now, we extend the coefficients a_{ij}, a_k, a_0 to the infinite cylinder $\mathbb{R} \times \Omega$, just setting:

$$a_{ij}(t,x) = a_{ij}(T_1,x) \text{ if } t \le T_1;\, a_{ij}(t,x) = a_{ij}(T_2,x) \text{ if } t \ge T_2$$

(and analogously for the a_k's and a_0). Clearly, the C_Z^α norms of the extended coefficients in $\mathbb{R} \times \Omega$ are equal to those of the original coefficients in $(T_1, T_2) \times \Omega$. Next, we take a cutoff function $\phi(x)$ such that

$$\Omega' \prec \phi \prec \Omega_2$$

and define

$$\widetilde{a}_{ij} = \phi a_{ij} + (1-\phi)\,\delta_{ij} \quad \text{for } i,j = 1, ..., q.$$

Since $a_{ij} \in C_Z^\alpha(\mathbb{R}\times\Omega)$, we also have $\widetilde{a}_{ij} \in C_Z^\alpha(\mathbb{R}\times\Omega)$ and therefore $\widetilde{a}_{ij} \in C_X^\alpha(\mathbb{R}\times\Omega_2)$, by the equivalence of d_X and d_Z in Ω_2. On the other hand, in $\mathbb{R}\times\Omega_2^c$ the $\widetilde{a}_{ij}$'s are constant, so $\widetilde{a}_{ij} \in C_X^\alpha\left(\mathbb{R}^{n+1}\right)$, with norms controlled by $\|a_{ij}\|_{C_Z^\alpha(\mathbb{R}\times\Omega)}$. It is also immediate to check that the matrix $\{\widetilde{a}_{ij}\}$ still satisfies (19.1) with the same constant λ. Finally, we set

$$\{\widetilde{a}_{ij}\}_{i,j=1}^m = \begin{bmatrix} \{\widetilde{a}_{ij}\}_{i,j=1}^q & 0 \\ 0 & I_n \end{bmatrix}.$$

As to the coefficients a_k, a_0, we make an analogous (but simpler) extension, just setting

$$\widetilde{a}_k = \phi a_k$$

and repeating the above reasoning.

With this construction, we see that the operator

$$\widetilde{H} = \partial_t - \sum_{i,j=1}^m \widetilde{a}_{ij}(t,x) X_i X_j - \sum_{k=1}^q \widetilde{a}_k(t,x) X_k - \widetilde{a}_0(t,x)$$

fits the assumptions of Part II, as stated in Section 10. Hence we can apply Theorem 10.7 to $\widetilde{H}$. The global fundamental solution h of $\widetilde{H}$ is then a local fundamental solution for H, satisfying all the properties required in points i) to vi). This simply follows from the fact that, in $\mathcal{C}'$, $\widetilde{H}$ coincides with H and d_X is equivalent to d_Z.

We just note that the number T has to be chosen small enough, so that $B_{d_Z}(x, \sqrt{cT}) \subset \Omega_2$ whenever $x \in \Omega'$; this fact assures that $B_{d_Z}(x, \sqrt{t})$ is well defined, and its measure is equivalent to $B_{d_X}(x, \sqrt{t})$, for any $t \in (0, cT)$. ■

Analogously we have:

Theorem 19.2 (Harnack inequality) *Assume H satisfies all the above assumptions; moreover, take $a_0 \equiv 0$. For any domain $\Omega' \Subset \Omega$ there exists a constant $R_0 > 0$ such that, for every $0 < h_1 < h_2 < 1$ and $\gamma \in (0,1)$, there exists $M > 0$, depending only on $h_1, h_2, \gamma, \Omega, \Omega', \lambda$, the vector fields Z_i and the Hölder norms of the coefficients, such that for every $(\tau_0, \xi_0) \in \mathcal{C}' = (T_1, T_2) \times \Omega'$, $R \in (0, R_0]$ and every*

$$u \in \mathfrak{C}^2((\tau_0 - R^2, \tau_0) \times B(\xi_0, R)) \cap C([\tau_0 - R^2, \tau_0] \times \overline{B(\xi_0, R)})$$

satisfying $Hu = 0$, $u \geq 0$ in $(\tau_0 - R^2, \tau_0) \times B(\xi_0, R)$, we have

$$\max_{[\tau_0 - h_2 R^2, \tau_0 - h_1 R^2] \times \overline{B(\xi_0, \gamma R)}} u \leq M\, u(\tau_0, \xi_0).$$

All the balls are taken with respect to d_Z.

Remark 19.3 *Recall that, by Theorem 14.4, if u is a $\mathfrak{C}^2$ solution to $Hu = 0$, then $u \in C^{2,\alpha}_{loc}$. Here the space $\mathfrak{C}^2$ is defined as in Definition 10.3, but with respect to the vector fields Z_i.*

Proof. With the same construction explained in the proof of the previous theorem, we build an operator $\widetilde{H}$ which satisfies the assumptions of Theorem 15.1 and such that in $\mathcal{C}'$, $\widetilde{H}$ coincides with H and d_X is equivalent to d_Z. Then, one can repeat the whole proof of Theorem 15.1 using the distance d_Z instead of d_X, and conclude that our assertion holds. Analogously one can restate the stationary Harnack inequality (Theorem 15.3) in this local setting. ■

20 Further developments and open problems

The results proved in this work leave some problems open, and suggest further developments of our study. Here we shall briefly illustrate some of them, which we would like to address in the future.

Fundamental solution for the stationary operator

Let us consider the stationary operator

$$L = \sum_{i,j=1}^{m} a_{ij}\left(x\right) X_i X_j.$$

A scaling invariant Harnack inequality for L has been deduced by the analogous (parabolic) result for H, see Theorem 15.3.

It would be interesting to prove the existence of a fundamental solution Γ for L, enjoying natural properties, among with bounds of the kind:

$$\begin{aligned} \frac{d\left(x,y\right)^2}{c\left|B_d\left(x,y\right)\right|} \leq \Gamma\left(x,y\right) &\leq \frac{cd\left(x,y\right)^2}{\left|B_d\left(x,y\right)\right|}; \\ \left|X_i^x \Gamma\left(x,y\right)\right| &\leq \frac{cd\left(x,y\right)}{\left|B_d\left(x,y\right)\right|}; \\ \left|X_i^x X_j^x \Gamma\left(x,y\right)\right| &\leq \frac{c}{\left|B_d\left(x,y\right)\right|}. \end{aligned}$$

A possible way to prove these results, already followed in [4] for vector fields on Carnot groups, is to consider:

$$\Gamma\left(x,y\right) = \int_0^{\infty} h\left(t,x,y\right) dt$$

where h is the fundamental solution of the evolution operator $\partial_t - L$. Obviously, to make this definition meaningful, one needs to establish *long time estimates for h*, like

$$h\left(t,x,y\right) \leq \frac{c}{\left|B_d\left(x,\sqrt{t}\right)\right|} \text{ for } t \in \left(0,\infty\right), x,y \in \mathbb{R}^n.$$

The proof of these long time estimates on h is a nontrivial extension of our results, which seems to require some new ideas.

General Hörmander's vector fields defined on the whole $\mathbb{R}^n$

Our theory has been developed to treat the case of a system of Hörmander's vector fields defined on a bounded domain of $\mathbb{R}^n$. The system has then been extended to the whole space in a suitable fashion. A harder challenge would be to consider any system of Hörmander's vector fields defined on the whole $\mathbb{R}^n$, i.e. without assuming that outside a compact set they take a particular form. Clearly, some requirement must be done on the behavior of the vector fields $X_i = \sum_{j=1}^{n} b_{ij}(x)\,\partial_{x_j}$ at infinity, in terms of the coefficients b_{ij} and of their derivatives, and / or in terms of some global property of the Carnot-Carathéodory metric induced by the vector fields.

Time-dependent and rough vector fields

Motion by Levi curvature involves fully nonlinear partial differential equations whose linearizations take the form:

$$H \equiv \partial_t - \sum_{i,j=1}^{2n} a_{ij}(t,x)\, X_i X_j u = f(t,x) \text{ for } (t,x) \in \mathbb{R}^{2n+2}$$

with

$$X_i = \sum_{j=1}^{2n+1} b_{ij}(t,x)\,\partial_{x_j}.$$

Hence, the vector fields have time-dependent coefficients, even though they still act as differential operators in the space variables only. For any fixed t, these X_i's are a system of vector fields in $\mathbb{R}^{2n+1}$, satisfying a Hörmander's rank condition of step two when, for instance, the previous equation describes the motion of strictly pseudoconvex real hypersurfaces. On the other hand, the variable t in the coefficients b_{ij} cannot be seen just as a "parameter", since the operator H also involves the derivative with respect to t. Extending our theory to this setting is likely to pose new interesting problems.

Deepening the above analysis, one should also take into account the necessity to deal with rough coefficients $b_{ij}(t,x)$. Indeed, in the original fully-nonlinear equation these coefficients depend on the solution $u(t,x)$, which *a-priori* could be nonsmooth. A general study of second order "Hörmander's operators" built with nonsmooth vector fields seems at present out of sight, but one could try to handle some more simply structured situation, sufficient to cover the particular application we have described.

References

[1] D. G. Aronson: Bounds for the fundamental solution of a parabolic equation. Bull. Amer. Math. Soc. 73 1967 890-896.

[2] E. Bedford, B. Gaveau: Hypersurfaces with bounded Levi form. Indiana Univ. Math. J. 27 (1978), no. 5, 867–873.

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

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

[5] A. Bonfiglioli, F. Uguzzoni: Families of diffeomorphic sub-Laplacians and free Carnot groups. Forum Math. 16 (2004), no. 3, 403-415.

[6] A. Bonfiglioli, F. Uguzzoni: Harnack inequality for non-divergence form operators on stratified groups, to appear in Trans. Amer. Math. Soc.

[7] A. Bonfiglioli, F. Uguzzoni: Maximum principle and propagation for intrinsicly regular solutions of differential inequalities structured on vector fields, J. Math. Anal. Appl., vol. 322, n. 2 (2006), 886-900.

[8] 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 fasc. 1, 277-304 xii.

[9] M. Bramanti, L. Brandolini: L^p-estimates for uniformly hypoelliptic operators with discontinuous coefficients on homogeneous groups. Rend. Sem. Mat. dell'Univ. e del Politec. di Torino. Vol. 58, 4 (2000), 389-433.

[10] M. Bramanti, L. Brandolini: L^p-estimates for nonvariational hypoelliptic operators with VMO coefficients. Trans. Amer. Math. Soc. 352 (2000), no. 2, 781-822.

[11] M. Bramanti, L. Brandolini: Estimates of BMO type for singular integrals on spaces of homogeneous type and applications to hypoelliptic PDEs, Rev. Mat. Iberoamericana 21 (2005), no. 2, 511–556.

[12] M. Bramanti, L. Brandolini: Schauder estimates for parabolic nondivergence operators of Hörmander type. To appear on Journal of Differential Equations.

[13] M. Bramanti, L. Brandolini, E. Lanconelli, F. Uguzzoni: Heat kernels for non-divergence operators of Hörmander type. To appear on C. R. Math. Acad. Sci. Paris.

[14] L. Caffarelli, L. Nirenberg, J. Spruck: The Dirichlet problem for nonlinear second-order elliptic equations. III. Functions of the eigenvalues of the Hessian. Acta Math. 155 (1985), no. 3-4, 261–301.

[15] L. Capogna, Q. Han: Pointwise Schauder estimates for second order linear equations in Carnot groups. Harmonic analysis at Mount Holyoke (South Hadley, MA, 2001), 45-69, Contemp. Math., 320, Amer. Math. Soc., Providence, RI, 2003.

[16] R. Coifman, G. Weiss: Analyse Harmonique Non-Commutative sur Certains Espaces Homogenes. Lecture Notes in Mathematics, n. 242. Springer-Verlag, Berlin-Heidelberg-New York, 1971.

[17] M.G. Crandall, H. Ishii, P.L. Lions: User's guide to viscosity solutions of second order partial differential equations. Bull. Amer. Math. Soc. (N.S.) 27 (1992), no. 1, 1–67.

[18] F. Da Lio, A. Montanari: Existence and uniqueness of Lipschitz continuous graphs with prescribed Levi curvature. Ann. Inst. H. Poincaré Anal. Non Linéaire 23 (2006), no. 1, 1–28.

[19] D. Danielli, N. Garofalo, D. M. Nhieu: On the best possible character of the norm in some a priori estimates for non-divergence form equations in Carnot groups. Proc. Amer. Math. Soc. 131 (2003), no. 11, 3487-3498.

[20] G. Di Fazio, A. Domokos, M.S. Fanciullo, J. Manfredi: Subelliptic Cordes estimates for Hörmander vector fields and applications to p-sublaplacian. Manuscripta Mathematica, vol. 120, no. 4 (2006), 419-433.

[21] A. Domokos, J. Manfredi: Cordes conditions and subelliptic estimates. Proc. Amer. Math. Soc. 133 (2005), no. 4, 1047-1056.

[22] E. B. Fabes, D. W. Stroock, A new proof of Moser's parabolic Harnack inequality using the old ideas of Nash, Arch. Rational Mech. Anal. **96** (1986), 327–338.

[23] C. Fefferman, D.H. Phong: Subelliptic eigenvalue problems. Conference on harmonic analysis in honor of Antoni Zygmund, Vol. I, II (Chicago, Ill., 1981), 590-606, Wadsworth Math. Ser., Wadsworth, Belmont, CA, 1983.

[24] C. Fefferman, A. Sánchez-Calle: Fundamental solutions for second order subelliptic operators. Ann. of Math. (2) 124 (1986), no. 2, 247-272.

[25] G. B. Folland: Real Analysis. John Wiley & sons, 1999.

[26] G. B. Folland: Subelliptic estimates and function spaces on nilpotent Lie groups, Arkiv for Mat. 13, (1975), 161-207.

[27] A. Friedman: Partial differential equations of parabolic type. Prentice-Hall, Inc., Englewood Cliffs, N.J. 1964.

[28] L. Hörmander: Hypoelliptic second order differential equations. Acta Math. 119 (1967) 147–171.

[29] G. Huisken, W. Klingenberg: Flow of real hypersurfaces by the trace of the Levi form. Math. Res. Lett. 6 (1999), no. 5-6, 645–661.

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

[31] J.J. Kohn: Pseudo-differential operators and hypoellipticity. Partial differential equations (Proc. Sympos. Pure Math., Vol. XXIII, Univ. California, Berkeley, Calif., 1971), pp. 61-69. Amer. Math. Soc., Providence, R.I., 1973.

[32] N. V. Krylov, M. V. Safonov: An estimate for the probability of a diffusion process hitting a set of positive measure. Dokl. Akad. Nauk SSSR 245 (1979), no. 1, 18–20.

[33] N. V. Krylov, M. V. Safonov: A property of the solutions of parabolic equations with measurable coefficients. Izv. Akad. Nauk SSSR Ser. Mat. 44 (1980), no. 1, 161–175, 239.

[34] S. Kusuoka, D. Stroock: Applications of the Malliavin calculus. II. J. Fac. Sci. Univ. Tokyo Sect. IA Math. 32 (1985), no. 1, 1–76.

[35] S. Kusuoka, D. Stroock: Applications of the Malliavin calculus. III. J. Fac. Sci. Univ. Tokyo Sect. IA Math. 34 (1987), no. 2, 391-442.

[36] S. Kusuoka, D. Stroock: Long time estimates for the heat kernel associated with a uniformly subelliptic symmetric second order operator. Ann. of Math. (2) 127 (1988), no. 1, 165-189.

[37] E. Lanconelli, A. Pascucci: On the fundamental solution for hypoelliptic second order partial differential equations with non-negative characteristic form, Ricerche Mat. 48 (1999), no. 1, 81–106.

[38] E. Lanconelli, S. Polidoro: On a class of hypoelliptic evolution operators. Partial differential equations, II (Turin, 1993). Rend. Sem. Mat. Univ. Politec. Torino 52 (1994), no. 1, 29–63.

[39] E. Lanconelli, F. Uguzzoni: Potential theory for a class of diffusion equations: a Gaussian bounds approach. In preparation.

[40] E. E. Levi: Sulle equazioni lineari totalmente ellittiche alle derivate parziali. Rend. Circ. Mat. Palermo 24, 275-317 (1907) (Also with corrections in: Eugenio Elia Levi, Opere, vol. II, 28-84. Roma, Edizioni Cremonese 1960).

[41] E. E. Levi: I problemi dei valori al contorno per le equazioni lineari totalmente ellittiche alle derivate parziali, Memorie Mat. Fis. Soc. Ital. Scienza (detta dei XL) (3) 16, 3-113 (1909). (Also with corrections in: Eugenio Elia Levi, Opere, vol. II, 207-343. Roma, Edizioni Cremonese 1960).

[42] A. Montanari: Real hypersurfaces evolving by Levi curvature: smooth regularity of solutions to the parabolic Levi equation. Comm. Partial Differential Equations 26 (2001), no. 9-10, 1633–1664.

[43] A. Montanari, E. Lanconelli: Pseudoconvex fully nonlinear partial differential operators: strong comparison theorems. J. Differential Equations 202 (2004), no. 2, 306–331.

[44] A. Montanari, F. Lascialfari: The Levi Monge-Ampère equation: smooth regularity of strictly Levi convex solutions. J. Geom. Anal. 14 (2004), no. 2, 331–353.

[45] A. Nagel, E. M. Stein, S. Wainger: Balls and metrics defined by vector fields I: Basic properties. Acta Mathematica, 155 (1985), 130-147.

[46] J. Nash: Continuity of solutions of parabolic and elliptic equations. Amer. J. Math. 80 (1958) 931–954.

[47] A. Pascucci, S. Polidoro: On the Harnack inequality for a class of hypoelliptic evolution equations. Trans. Amer. Math. Soc. 356 (2004), no. 11, 4383–4394.

[48] S. Polidoro: On a class of ultraparabolic operators of Kolmogorov-Fokker-Planck type. Le Matematiche (Catania) 49 (1994), no. 1, 53–105.

[49] L. P. Rothschild, E. M. Stein: Hypoelliptic differential operators and nilpotent groups. Acta Math., 137 (1976), 247-320.

[50] M. V. Safonov: Harnack's inequality for elliptic equations and Hölder property of their solutions. Boundary value problems of mathematical physics and related questions in the theory of functions, 12. Zap. Nauchn. Sem. Leningrad. Otdel. Mat. Inst. Steklov. (LOMI) 96 (1980), 272–287, 312.

[51] L. Saloff-Coste, Aspects of Sobolev-type inequalities. London Mathematical Society Lecture Note Series, 289. Cambridge University Press, Cambridge, 2002.

[52] L. Saloff-Coste, D. W. Stroock: Opérateurs uniformément sous-elliptiques sur les groupes de Lie. J. Funct. Anal. 98 (1991), no. 1, 97–121.

[53] A. Sanchez-Calle: Fundamental solutions and geometry of sum of squares of vector fields. Inv. Math., 78 (1984), 143-160.

[54] Z. Slodkowski, G. Tomassini: Weak solutions for the Levi equation and envelope of holomorphy. J. Funct. Anal. 101 (1991), no. 2, 392–407.

[55] E. M. Stein: Harmonic Analysis: Real-Variable Methods, Orthogonality and Oscillatory Integrals. Princeton Univ. Press. Princeton, New Jersey, 1993.

[56] G. Tomassini: Geometric properties of solutions of the Levi-equation. Ann. Mat. Pura Appl. (4) 152 (1988), 331–344.

[57] N. Th. Varopoulos: Théorie du potentiel sur les groupes nilpotents. C. R. Acad. Sci. Paris Sér. I Math. 301 (1985), no. 5, 143-144.

[58] N. Th. Varopoulos, Analysis on nilpotent groups. J. Funct. Anal. 66 (1986), no. 3, 406-431.

[59] N. Th.Varopoulos, L. Saloff-Coste, T. Coulhon: Analysis and geometry on groups. Cambridge Tracts in Mathematics, 100. Cambridge University Press, Cambridge, 1992.

[60] C.J. Xu: Regularity for quasilinear second-order subelliptic equations. Comm. Pure Appl. Math. 45 (1992), no. 1, 77-96.