Available online at www.sciencedirect.com
ScienceDirect
J. Differential Equations 234 (2007) 177–245
Journal of Differential Equations
www.elsevier.com/locate/jde

Schauder estimates for parabolic nondivergence operators of Hörmander type

Marco Bramanti [a,*], Luca Brandolini [b]

[a] Dipartimento di Matematica, Politecnico di Milano, Via Bonardi 9, 20133 Milano, Italy
[b] Dipartimento di Ingegneria Gestionale e dell'Informazione, Università degli Studi di Bergamo, Viale Marconi 5, 24044 Dalmine BG, Italy

Received 27 April 2006; revised 10 July 2006

Available online 13 December 2006

Abstract

Let $X_1, X_2, \dots, X_q$ be a system of real smooth vector fields satisfying Hörmander's rank condition in a bounded domain Ω of $\mathbb{R}^n$. Let $A = \{a_{ij}(t,x)\}_{i,j=1}^q$ be a symmetric, uniformly positive definite matrix of real functions defined in a domain $U \subset \mathbb{R} \times \Omega$. For operators of kind

$$H = \partial_t - \sum_{i,j=1}^q a_{ij}(t,x) X_i X_j - \sum_{i=1}^q b_i(t,x) X_i - c(t,x)$$

we prove local a-priori estimates of Schauder-type, in the natural (parabolic) $C^{k,\alpha}(U)$ spaces defined by the vector fields X_i and the distance induced by them. Namely, for $a_{ij}, b_i, c \in C^{k,\alpha}(U)$ and $U' \Subset U$, we prove

$$\|u\|_{C^{k+2,\alpha}(U')} \leqslant c\big\{\|Hu\|_{C^{k,\alpha}(U)} + \|u\|_{L^\infty(U)}\big\}.$$

Keywords: Hörmander's operators; A-priori estimates; Hölder spaces; Singular integrals

* Corresponding author.
E-mail addresses: marbra@mate.polimi.it (M. Bramanti), luca.brandolini@unibg.it (L. Brandolini).

0022-0396/$ – see front matter
doi:10.1016/j.jde.2006.07.015

1. Introduction

Let Ω be a bounded domain of $\mathbb{R}^n$, and let $X_1, X_2, \dots, X_q$ be a system of smooth real vector fields satisfying Hörmander's rank condition in Ω. In this setting, "sum of squares" operators

$$\sum_{i=1}^{q} X_i^2$$

or their "parabolic" analog

$$\partial_t - \sum_{i=1}^{q} X_i^2 \tag{1.1}$$

have been widely studied since Hörmander's famous paper [20]: these operators are hypoelliptic, and share with elliptic and parabolic operators several deep analogies. In recent years, nondivergence operators modeled on the above classes, namely,

$$L = \sum_{i,j=1}^{q} a_{ij}(x) X_i X_j \quad \text{or} \tag{1.2}$$

$$H = \partial_t - \sum_{i,j=1}^{q} a_{ij}(t,x) X_i X_j \tag{1.3}$$

have also been studied, assuming that $A = \{a_{ij}\}_{i,j=1}^{q}$ is a symmetric, uniformly positive definite matrix of real functions defined in Ω (in case (1.2)) or in a bounded domain $U \subset \mathbb{R} \times \Omega$ (in case (1.3)), and $\lambda > 0$ is a constant such that

$$\lambda^{-1}|\xi|^2 \leqslant \sum_{i,j=1}^{q} a_{ij}\xi_i\xi_j \leqslant \lambda|\xi|^2 \quad \text{for every } \xi \in \mathbb{R}^q, \tag{1.4}$$

uniformly in Ω or U (see [2] for (1.3) and [6] and references therein for (1.2)). These classes of operators naturally arise in some problems related to geometry in several complex variables (see [27] and references therein) as well as in some models of human vision (see [14] and references therein); moreover, these operators realize a framework where a suitable theory of nonlinear equations modeled on Hörmander's vector fields can be settled.

A system of Hörmander vector fields can be thought as the natural substitute of the "Cartesian" derivatives ∂_{x_i}, in the study of degenerate equations like (1.2) or (1.3). Moreover, it induces a "Carnot–Carathéodory distance," which is (locally) doubling with respect to the Lebesgue measure. These facts allow to define several function spaces shaped on the vector fields, such as Hölder spaces, Sobolev spaces, *BMO*, *VMO*, etc. It is then natural to use these spaces to express the required regularity of the coefficients a_{ij}. Clearly, as soon as the coefficients a_{ij} are not C^∞, the corresponding operator (1.2) or (1.3) is no longer hypoelliptic, and no result can be drawn on it from the classical theory of Hörmander's sums of squares. Nevertheless, many classical results about elliptic and parabolic operators, which do not require, in principle, high regularity of the coefficients, when properly reformulated in the language of vector fields, look like desirable properties of these operators, and reasonable—although nontrivial conjectures. Two typical instances of this situation are (local) L^p estimates and C^α estimates on the "second order" derivatives $X_i X_j u$. In [5,6] we have proved L^p estimates of this kind for operators of type (1.2) or some more general classes, assuming the coefficients a_{ij} in the space *VMO*, extending the classical results of Rothschild and Stein [29] for Hörmander's sum of squares. In this paper, we prove local C^α estimates of Schauder type for an operator (1.3). Our main result is the following (all symbols will be defined in the following sections).

Theorem 1.1. *Let Ω be a bounded domain of $\mathbb{R}^n$, and let $X_1, X_2, \ldots, X_q$ be a system of smooth real vector fields defined in a neighborhood Ω_0 of Ω and satisfying Hörmander's rank condition in Ω_0. Let U be a bounded domain of $\mathbb{R}^{n+1}$, $U \subset \mathbb{R} \times \Omega$; let $A = \{a_{ij}(t,x)\}_{i,j=1}^q$ be a symmetric, uniformly positive definite matrix of real functions defined in U, and $\lambda > 0$ a constant such that* (1.4) *holds in U. Assume $a_{ij}, b_i, c \in C^{k,\alpha}(U)$ for some integer $k \geqslant 0$ and some $\alpha \in (0,1)$. Let*

$$H = \partial_t - \sum_{i,j=1}^q a_{ij}(t,x) X_i X_j - \sum_{i=1}^q b_i(t,x) X_i - c(t,x). \tag{1.5}$$

Then, for every domain $U' \Subset U$ there exists a constant $c > 0$ depending on U, U', $\{X_i\}$, α, k, λ and the $C^{k,\alpha}$ norms of the coefficients such that for every $u \in C^{k+2,\alpha}_{\text{loc}}(U)$ with $Hu \in C^{k,\alpha}(U)$ one has

$$\|u\|_{C^{k+2,\alpha}(U')} \leqslant c\left\{\|Hu\|_{C^{k,\alpha}(U)} + \|u\|_{L^\infty(U)}\right\}.$$

Analogous Schauder estimates for stationary operators (1.2) obviously follow from the above theorem, as a particular case.

Let us briefly compare our result with the existing literature. In [33], Xu states local estimates of Schauder type for operators of type (1.2), under an additional assumption on the structure of the Lie algebra generated by the X_i's. In [12], Capogna and Han prove "pointwise Schauder estimates" (in the spirit of Caffarelli's work [9] on fully nonlinear equations) for equations of type (1.2) in Carnot groups. In [26], Montanari proves local Schauder estimates for a particular class of operators of type (1.3), namely, tangential operators on CR manifolds, where the vector fields are allowed to be nonsmooth (namely, $C^{1,\alpha}$).

The main feature of the present paper, besides the "evolutionary" case it covers, is that our theory applies to *any* system of Hörmander vector fields.

The general strategy we use (described in detail in Section 5) is similar to that we have followed in [6,7]. A basic role is played by C^α continuity of singular and fractional integrals on spaces of homogeneous type (in the sense of Coifman and Weiss [15]), coupled with the machinery introduced in [29] and adapted to nondivergence form of operators in [6]. These results about C^α continuity of singular and fractional integrals are proved in Theorems 2.7 and 2.11 and can be of independent interest. Again, the main feature of these results, compared with the existing literature, is their generality, which makes them suitable for application to the context of general Hörmander's vector fields.

Once we have proved Theorem 1.1, a more subtle question poses, namely the possibility of using the above a-priori estimates to show that, whenever a function $u \in C^{2,\alpha}_{\text{loc}}(U)$ solves $Hu = f$ in U with $C^{k,\alpha}(U)$ coefficients and data, then actually $u \in C^{k+2,\alpha}_{\text{loc}}(U)$. This natural regularization result follows from the a-priori estimates as soon as one can solve the classical Dirichlet problem, for operators of kind (1.5) but with smooth coefficients, provided a good mollification technique, suited to this context, is available. Solvability of the Dirichlet problem is a classical result, due to Bony [3], while in Section 11 we will construct a family of mollifiers adapted to our context. This construction, which can be of independent interest, makes use of the existence and properties of the "heat kernel" for the model operator (1.1), and also of the abstract theory of singular integrals developed in Section 2. The desired regularization result is proved in Theorem 11.5. For technical difficulties, our technique allows to prove this result only for even k.

A first application of the theory contained in this paper is the following. In [8], Bramanti et al. prove that operators of type (1.3) possess a fundamental solution, which satisfies sharp Gaussian estimates. The "Schauder theory" developed in this paper allows to show that this fundamental solution has a finite $C^{2,\alpha}$ norm, in any bounded domain excluding the pole, depending only on the vector fields, the C^{α} norms of the coefficients, and the ellipticity constant λ. This fact will be proved in [8].

Plan of the paper. In Section 2 we prove some abstract results about the action of singular and fractional integrals on spaces of homogeneous type. The next two sections are of preliminary nature: in Section 3 we prove some properties of the "parabolic Carnot–Carathéodory distance" induced by the vector fields, which will allow to apply the abstract theory of Section 2 to our setting, while in Section 4 we collect some properties of parabolic Hölder spaces C^{α} and $C^{k,\alpha}$ induced by Hörmander's vector fields. In Section 5 we state precisely our main results and illustrate the general strategy of the proof: our basic result, that is the $C^{2,\alpha}$ estimate for an operator without lower order terms, will be proved in three steps, which are briefly explained in Section 5. These three steps constitute Sections 6, 7, 8, respectively. The basic result is then extended to higher order derivatives in Section 9, and to operators with lower order terms in Section 10. The construction of a family of mollifiers which allow to control $C^{k,\alpha}$-norms, and the proof of regularity results, are performed in Section 11. Finally, Appendix A collects some notation and known results which are employed throughout the paper, and should be known to any reader who is familiar with the two classical papers [16,29].

2. Singular integrals on spaces of homogeneous type and continuity on Hölder spaces

Let X be a set. A function $d : X \times X \to \mathbb{R}$ is called a quasidistance on X if there exists a constant $c_d \geqslant 1$ such that for any $x, y, z \in X$:

$$d(x,y) \geqslant 0 \quad \text{and} \quad d(x,y) = 0 \quad \Longleftrightarrow \quad x = y;$$
$$d(x,y) = d(y,x);$$
$$d(x,y) \leqslant c_d\big(d(x,z) + d(z,y)\big). \tag{2.1}$$

We will say that two quasidistances d, d' on X are equivalent, and we will write $d \simeq d'$, if there exist two positive constants c_1, c_2 such that $c_1 d'(x,y) \leqslant d(x,y) \leqslant c_2 d'(x,y)$ for any $x, y \in X$.

For $r > 0$, let $B_r(x) = \{y \in X\colon d(x,y) < r\}$. These "balls" satisfy the axioms of a complete system of neighborhoods in X, and therefore induce a (separated) topology. With respect to this topology, the balls $B_r(x)$ need not be open. We will explicitly exclude the above kind of pathology.

Definition 2.1. Let (X,d) be a set endowed with a quasidistance d such that the d-balls are open with respect to the topology induced by d, and let μ be a positive Borel measure on X satisfying the doubling condition: there exists a positive constant c_μ such that

$$\mu\big(B_{2r}(x)\big)\leqslant c_\mu\cdot\mu\big(B_r(x)\big)\quad\text{for any } x\in X,\ r>0. \tag{2.2}$$

Then (X,d,μ) is called a space of homogeneous type.

To simplify notation, the measure $d\mu(x)$ will be denoted simply by dx, and $\mu(A)$ will be written $|A|$. We will also set

$$B(x;y)=B_{d(x,y)}(x).$$

Definition 2.2 *(Hölder spaces).* For any $\alpha>0$, $u:X\to\mathbb{R}$, let:

$$|u|_{C^\alpha(X)}=\sup\left\{\frac{|u(x)-u(y)|}{d(x,y)^\alpha}\colon\ x,y\in X,\ x\neq y\right\},$$
$$\|u\|_{C^\alpha(X)}=|u|_{C^\alpha(X)}+\|u\|_{L^\infty(X)},$$
$$C^\alpha(X)=\big\{u:X\to\mathbb{R}\colon\ \|u\|_{C^\alpha(X)}<\infty\big\}.$$

Also, we denote by $C_0^\alpha(X)$ the subspace of boundedly supported $C^\alpha(X)$ functions.

A basic result proved by Macías and Segovia (see [24, Theorem 2]) states.

Proposition 2.3. *Let d be any quasidistance on a set X. Then there exists another quasidistance d' on X, equivalent to d, a constant $c>0$ and an exponent $\alpha_0\in(0,1]$ such that for every $r>0$, $x,y,z\in X$ with $d'(x,z)<r$, $d'(y,z)<r$,*

$$\big|d'(x,z)-d'(y,z)\big|\leqslant cd'(x,y)^{\alpha_0}r^{1-\alpha_0}. \tag{2.3}$$

Remark 2.4. This proposition says that the function $x\mapsto d'(x,z)$ (for z fixed) is locally Hölder continuous (with respect to d' and therefore also to d). This allows to prove, under reasonable assumptions on the measure μ (for instance, if μ is a Radon measure) that on the space of homogeneous type (X,d,μ), $C_0^\alpha(X)$ is dense in $L^p(X)$ for any $p\in[1,\infty)$ and any $\alpha\leqslant\alpha_0$ (with α_0 as in (2.3)). In particular, if d is (equivalent to) a distance, then $\alpha_0=1$ in (2.3). So, in a general space of homogeneous type, Hölder spaces are always interesting spaces for α small enough. On the opposite side, we cannot say, in general, that for α large enough the space $C^\alpha(X)$ is reduced to constant functions; this will be the case in our application to Carnot–Carathéodory distance, due to the presence of a suitable "gradient" related to the distance.

Definition 2.5. Let (X,d,dx) be a space of homogeneous type.

We will say that a measurable function $k(x,y):X\times X\to\mathbb{R}$ is a standard kernel on X if k satisfies the following properties:

$$\big|k(x,y)\big|\leqslant\frac{c}{|B(x;y)|}\quad\text{for any } x,y\in X \tag{2.4}$$

(“growth estimate”);

$$\left|k(x,y)-k(x_0,y)\right|\leqslant\frac{c}{|B(x_0;y)|}\left(\frac{d(x_0,x)}{d(x_0,y)}\right)^{\beta}\tag{2.5}$$

for any $x_0, x, y \in X$, with $d(x_0, y) \geqslant Md(x_0, x)$, $M > 1$, $c, \beta > 0$ (“mean value inequality”).

Remark 2.6. Condition (2.4) and the doubling condition immediately imply that for any fixed $c_1, c_2 > 0$,

$$\int\limits_{c_1r<d(x,y)<c_2r}\left|k(x,y)\right|dy\leqslant c\tag{2.6}$$

for any $r > 0$, with c independent of r.

Note also that, if condition (2.5) holds for some $M_0 > 1$, then it holds for any $M \geqslant M_0$. We can assume M large enough, so that the condition $d(x_0, y) \geqslant Md(x_0, x)$ implies that $d(x_0, y) \simeq d(x, y)$. We will use systematically this equivalence. Moreover, just not to use one more constant, we will assume that this “large” value of M is 2. This means to assume that the constant c_d in (2.1) is < 2. The reader will excuse this little abuse of notation.

Theorem 2.7. *Let (X, d, dx) be a bounded space of homogeneous type, and let $k(x, y)$ be a standard kernel. Let*

$$K_\varepsilon f(x)=\int\limits_{d'(x,y)>\varepsilon}k(x,y)f(y)\,dy,\tag{2.7}$$

where d' is any quasidistance on X, equivalent to d, and fixed once and for all. Assume that for every $f \in C^\alpha(X)$ and $x \in X$ the following limit exists:

$$Kf(x)=PV\int\limits_X k(x,y)f(y)\,dy=\lim_{\varepsilon\to0}K_\varepsilon f(x).$$

Also, assume that:

$$\left|\int\limits_{d'(x,y)>r}k(x,y)\,dy\right|\leqslant c_K\tag{2.8}$$

for any $r > 0$ (with c_K independent of r) and

$$\lim_{\varepsilon\to0}\left|\int\limits_{d'(x,y)>\varepsilon}k(x,y)\,dy-\int\limits_{d'(x_0,y)>\varepsilon}k(x_0,y)\,dy\right|\leqslant c_Kd(x,x_0)^{\gamma}\tag{2.9}$$

for some $\gamma \in (0, 1]$, where d' is the same quasidistance appearing in (2.7). *Then the operator K is continuous on $C^\alpha(X)$; more precisely*:

$$|Kf|_{C^\alpha(X)}\leqslant c_K\|f\|_{C^\alpha(X)}\quad\text{for every }\alpha\leqslant\gamma,\ \alpha<\beta,\tag{2.10}$$

where γ is the number in (2.9) *and β is the number in* (2.5). *Moreover,*

$$\|Kf\|_\infty\leqslant c_{K,R,\alpha}\|f\|_\alpha,\quad\text{where }R=\operatorname{diam}X.\tag{2.11}$$

Remark 2.8. The fact that classical singular integrals "with variable kernels" (those arising in the study of linear elliptic equations) preserve Hölder spaces was already proved by Calderón and Zygmund in [10, Theorem 2, p. 909]. In the context of "homogeneous spaces with gauge," continuity of singular integrals on Hölder spaces was proved by Korányi and Vági [21]; this result has also been applied by Folland [16], in the context of homogeneous groups. These results are particular cases of the previous proposition, while the lack of any kind of homogeneity in the space is the main feature of our result. It is worthwhile to mention that the boundedness of the space X is necessary only for (2.11).

Remark 2.9 *(On the role of different quasidistances).* Here we want to clarify the role of the two, possibly different, quasidistances d, d'. In some applications of the abstract theory of singular integrals on spaces of homogeneous type (included the present application to the proof of Schauder estimates), it is useful to switch from one quasidistance to another one, having different good properties. In particular, in the definition of principal value of a singular integral, the small region around the pole which is removed and shrinked needs not to be a ball with respect to the original quasidistance. Also, it is worthwhile to note that properties (2.4), (2.5) are preserved replacing the quasidistance with an equivalent one; the same is true for (2.8), provided also (2.4) is assumed (therefore, in (2.8) the presence of d' instead of d is not relevant, but only written for consistence with (2.7)); on the other hand, property (2.9) is not obviously preserved replacing the quasidistance with an equivalent one. Therefore, the possibility of choosing in (2.9) and (2.7) a suitable quasidistance d', possibly different from d, will be crucial to check these assumptions in our context of Hörmander vector fields.

In the proof of the above proposition we need the following lemma, that can be proved by a standard computation (see [7, Lemma 2.8]).

Lemma 2.10. *Let X be any space of homogeneous type. Then*

$$\text{(a)} \quad \int_{d(x,y)<r} \frac{d(x,y)^{\beta}}{|B(x;y)|}\,dy \leqslant cr^{\beta} \quad \textit{for any } \beta > 0;$$

$$\text{(b)} \quad \int_{d(x,y)>r} \frac{d(x,y)^{-\beta}}{|B(x;y)|}\,dy \leqslant cr^{-\beta} \quad \textit{for any } \beta > 0.$$

Proof of Theorem 2.7. To prove (2.10), let us write

$$\begin{aligned} Kf(x) - Kf(x_0) = {} & \left\{ \int_X k(x,y)\big[f(y) - f(x)\big]\,dy - \int_X k(x_0,y)\big[f(y) - f(x_0)\big]\,dy \right\} \\ & + \lim_{\varepsilon\to 0}\left\{ f(x) \int_{d'(x,y)>\varepsilon} k(x,y)\,dy - f(x_0) \int_{d'(x_0,y)>\varepsilon} k(x_0,y)\,dy \right\} \\ \equiv {} & A + B, \end{aligned}$$

$$
\begin{aligned}
A &= \left\{ \int_{d(x_0,y)\geqslant 2d(x_0,x)} \left\{k(x,y)\left[f(y)-f(x)\right]-k(x_0,y)\left[f(y)-f(x_0)\right]\right\} dy \right\} \\
&\quad + \left\{ \int_{d(x_0,y)<2d(x_0,x)} \left\{k(x,y)\left[f(y)-f(x)\right]-k(x_0,y)\left[f(y)-f(x_0)\right]\right\} dy \right\} \\
&\equiv A_1 + A_2, \\
A_1 &= \int_{d(x_0,y)\geqslant 2d(x_0,x)} \left\{\left[k(x,y)-k(x_0,y)\right]\left[f(y)-f(x_0)\right]\right\} dy \\
&\quad + \left[f(x_0)-f(x)\right] \int_{d(x_0,y)\geqslant 2d(x_0,x)} k(x,y)\, dy \\
&\equiv A_{11} + A_{12}, \\
|A_{11}| &\leqslant \int_{d(x_0,y)\geqslant 2d(x_0,x)} \frac{c}{|B(x_0;y)|}\left(\frac{d(x_0,x)}{d(x_0,y)}\right)^{\beta} |f|_\alpha d(x_0,y)^\alpha\, dy \\
&= c|f|_\alpha d(x_0,x)^\beta \int_{d(x_0,y)\geqslant 2d(x_0,x)} \frac{1}{|B(x_0;y)|d(x_0,y)^{\beta-\alpha}}\, dy \\
&\leqslant c|f|_\alpha d(x_0,x)^\beta d(x_0,x)^{\alpha-\beta} = c|f|_\alpha d(x_0,x)^\alpha \quad \text{if } \alpha<\beta \text{ (by Lemma 2.10(b))}.
\end{aligned}
$$

As to the second term,

$$
|A_{12}| \leqslant |f|_\alpha d(x_0,x)^\alpha \left| \int_{d(x_0,y)\geqslant 2d(x_0,x)} k(x,y)\, dy \right|.
$$

By Remark 2.6, $d(x_0,y)\geqslant 2d(x_0,x) \Rightarrow d(x,y)\geqslant cd(x_0,x)$ for some $c>0$. Then

$$
\int_{d(x_0,y)\geqslant 2d(x_0,x)} k(x,y)\, dy = \int_{d(x,y)\geqslant cd(x_0,x)} k(x,y)\, dy - \int_{\substack{d(x_0,y)<2d(x_0,x)\\ d(x,y)\geqslant cd(x_0,x)}} k(x,y)\, dy
$$

and, by (2.8) and (2.6),

$$
\begin{aligned}
|A_{12}| &\leqslant |f|_\alpha d(x_0,x)^\alpha \left\{ \left| \int_{d(x,y)\geqslant cd(x_0,x)} k(x,y)\, dy \right| + \int_{d(x_0,y)<2d(x_0,x),\ d(x,y)\geqslant cd(x_0,x)} \left|k(x,y)\right| dy \right\} \\
&\leqslant |f|_\alpha d(x_0,x)^\alpha \left\{ c_K + \int_{cd(x_0,x)\leqslant d(x,y)\leqslant c_1 d(x_0,x)} \left|k(x,y)\right| dy \right\} \\
&\leqslant c_K |f|_\alpha d(x_0,x)^\alpha,
\end{aligned}
$$

$$|A_2| \leqslant \int\limits_{d(x_0,y)<2d(x_0,x)} |k(x,y)||f(y)-f(x)|\,dy + \int\limits_{d(x_0,y)<2d(x_0,x)} |k(x_0,y)||f(y)-f(x_0)|\,dy$$

since $d(x_0, y) < 2d(x_0, x) \Rightarrow d(x, y) < cd(x_0, x)$

$$\leqslant \int\limits_{d(x,y)<cd(x_0,x)} |k(x,y)||f(y)-f(x)|\,dy + \int\limits_{d(x_0,y)<2d(x_0,x)} |k(x_0,y)||f(y)-f(x_0)|\,dy$$
$$\equiv A_{21} + A_{22},$$

$$|A_{21}| \leqslant c_K |f|_\alpha \int\limits_{d(x,y)<cd(x_0,x)} \frac{d(x,y)^\alpha}{|B(x;y)|}\,dy$$
$$\leqslant c_K |f|_\alpha d(x,x_0)^\alpha \quad \text{(by Lemma 2.10(a))}.$$

Analogously,

$$|A_{22}| \leqslant c_K |f|_\alpha d(x,x_0)^\alpha.$$

We have therefore proved that

$$|A| \leqslant c_K |f|_\alpha d(x,x_0)^\alpha.$$

Let us come to B:

$$B = \lim_{\varepsilon\to 0}\Biggl\{ f(x) \int\limits_{d'(x,y)>\varepsilon} k(x,y)\,dy - f(x_0) \int\limits_{d'(x_0,y)>\varepsilon} k(x_0,y)\,dy \Biggr\}$$
$$= \bigl[f(x) - f(x_0)\bigr] \lim_{\varepsilon\to 0} \int\limits_{d'(x,y)>\varepsilon} k(x,y)\,dy$$
$$+ f(x_0) \lim_{\varepsilon\to 0}\Biggl\{ \int\limits_{d'(x,y)>\varepsilon} k(x,y)\,dy - \int\limits_{d'(x_0,y)>\varepsilon} k(x_0,y)\,dy \Biggr\} \equiv B_1 + B_2,$$

$$|B_1| \leqslant |f|_\alpha d(x,x_0)^\alpha \sup_{\varepsilon>0} \Biggl| \int\limits_{d'(x,y)>\varepsilon} k(x,y)\,dy \Biggr| \leqslant c_K |f|_\alpha d(x,x_0)^\alpha \quad \text{(by (2.8))}.$$

Moreover, by (2.9), we can conclude

$$|B| \leqslant c_K |f|_\alpha d(x,x_0)^\alpha + c_K \|f\|_\infty d(x,x_0)^\gamma.$$

This ends the proof of (2.10).

To prove (2.11), let us write

$$Kf(x) = \int\limits_X k(x,y)\bigl[f(y) - f(x)\bigr]\,dy + f(x) \lim_{\varepsilon\to 0} \int\limits_{d(x,y)>\varepsilon} k(x,y)\,dy = A + B,$$

$$
\begin{aligned}
|A| &\leqslant c_K |f|_\alpha \int_X \frac{d(x,y)^\alpha}{|B(x;y)|}\,dy \\
&\leqslant c_K |f|_\alpha \int_{d(x,y)\leqslant R} \frac{d(x,y)^\alpha}{|B(x;y)|}\,dy \quad \text{for some fixed } R>0, \text{ since the space is bounded} \\
&\leqslant c_K |f|_\alpha R^\alpha \quad \text{(by Lemma 2.10(a))}, \\
|B| &\leqslant \|f\|_\infty \sup_{\varepsilon>0} \left| \int_{d'(x,y)>\varepsilon} k(x,y)\,dy \right| \leqslant c_K \|f\|_\infty
\end{aligned}
$$

and this concludes the proof. □

The next theorem provides a result of C^α continuity for fractional integrals.

Theorem 2.11. *Let (X, d, dx) be a bounded space of homogeneous type, and assume that X does not contain atoms (that is, points of positive measure). Let $k_\delta(x, y)$ be a "fractional integral kernel," that is,*

$$
0 \leqslant k_\delta(x,y) \leqslant \frac{c d(x,y)^\delta}{|B(x;y)|} \tag{2.12}
$$

for any $x, y \in X$, some $c, \delta > 0$;

$$
\left| k_\delta(x,y) - k_\delta(x_0,y) \right| \leqslant \frac{c d(x_0,y)^\delta}{|B(x_0;y)|} \left(\frac{d(x_0,x)}{d(x_0,y)} \right)^\beta \tag{2.13}
$$

for any $x_0, x, y \in X$, with $d(x_0, y) \geqslant 2d(x_0, x)$, some $c, \beta > 0$ ("mean value inequality"). Then the operator

$$
I_\delta f(x) = \int_X k_\delta(x,y) f(y)\,dy
$$

is continuous on $C^\alpha(X)$ for any $\alpha < \min(\beta, \delta)$.

Remark 2.12. If the space X contains atoms, the definition of I_δ has to be modified as

$$
I_\delta f(x) = \int_{X\setminus\{x\}} k_\delta(x,y) f(y)\,dy,
$$

in order to assure the convergence of the integral; we want to avoid these technicalities. Fractional integrals on spaces of homogeneous type have been extensively studied by Gatto and Vági, see [17,18]; see also [19] and references therein. However, our result is not comparable with theirs because, on one side, they make the extra assumption of normality of the space, while, on the other side, they do not require boundedness of X. Moreover, our result is not sharp: one should expect I_δ to map C^α in $C^{\alpha+\delta}$; here we have limited ourselves to prove, in the shortest way, the result which we need for subsequent applications to Schauder estimates.

Proof of Theorem 2.11. Let R be the diameter of X. We will check that k_δ satisfies assumptions (2.4), (2.5), (2.8), and (2.9); then the result will follow by Theorem 2.7. Namely, property (2.12) implies (2.4) with the constant c replaced by cR^δ; analogously, property (2.13) implies (2.5), with the same exponent β. By Lemma 2.10(a)

$$\left|\int_{d'(x,y)>\delta} k_\delta(x,y)\,dy\right| \leqslant c \int_{d(x,y)<R} \frac{d(x,y)^\delta}{|B(x;y)|}\,dy \leqslant cR^\delta$$

hence (2.8) holds. Finally, to prove (2.9), we start by noting that in this case

$$\lim_{\varepsilon\to 0}\left|\int_{d'(x,y)>\varepsilon} k_\delta(x,y)\,dy - \int_{d'(x_0,y)>\varepsilon} k_\delta(x_0,y)\,dy\right| = \left|\int_X k_\delta(x,y)\,dy - \int_X k_\delta(x_0,y)\,dy\right|$$

because, by (2.12), the integral of $k_\delta(x,\cdot)$ is convergent, hence

$$\int_{d'(x,y)\leqslant\varepsilon} k_\delta(x,y)\,dy \to 0 \quad \text{for } \varepsilon\to 0$$

since X has no atoms. By (2.12), (2.13) and Lemma 2.10,

$$\begin{aligned}
&\left|\int_X k_\delta(x,y)\,dy - \int_X k_\delta(x_0,y)\,dy\right| \\
&\quad\leqslant \int_{d(x_0,y)>2d(x_0,x)} \left|k_\delta(x,y)-k_\delta(x_0,y)\right|dy + \int_{d(x_0,y)\leqslant 2d(x_0,x)} \left|k_\delta(x,y)-k_\delta(x_0,y)\right|dy \\
&\quad\leqslant cd(x_0,x)^\beta \int_{d(x_0,y)>2d(x_0,x)} \frac{dy}{|B(x_0;y)|d(x_0,y)^{\beta-\delta}} \\
&\qquad + c\int_{d(x,y)\leqslant cd(x_0,x)} \frac{d(x,y)^\delta}{|B(x;y)|}\,dy + c\int_{d(x_0,y)\leqslant 2d(x_0,x)} \frac{d(x_0,y)^\delta}{|B(x_0;y)|} \equiv I.
\end{aligned}$$

Now,

$$\begin{aligned}
I &\leqslant cd(x_0,x)^\beta\cdot d(x_0,x)^{\delta-\beta} + cd(x_0,x)^\delta \leqslant cd(x_0,x)^\delta \quad \text{if } \beta>\delta, \\
I &\leqslant cd(x_0,x)^\beta \int_{d(x_0,y)<R} \frac{d(x_0,y)^{\delta-\beta}}{|B(x_0;y)|}\,dy + cd(x_0,x)^\delta \\
&\leqslant cd(x_0,x)^\beta R^{\delta-\beta} + cd(x_0,x)^\delta \leqslant cd(x_0,x)^\beta R^{\delta-\beta} \quad \text{if } \beta<\delta,
\end{aligned}$$

since we can assume $d(x_0,x)<R$; finally, if $\beta=\delta$,

$$\begin{aligned} I &\leqslant cd(x_0,x)^\beta \int\limits_{d(x_0,y)>2d(x_0,x)} \frac{1}{|B(x_0;y)|}\,dy + cd(x_0,x)^\beta \\ &\leqslant cd(x_0,x)^\beta \int\limits_{d(x_0,y)<R} \left(\frac{d(x_0,y)}{d(x_0,x)}\right)^\varepsilon \frac{1}{|B(x_0;y)|}\,dy + cd(x_0,x)^\beta \\ &\leqslant c_\varepsilon d(x_0,x)^{\beta-\varepsilon}R^\varepsilon + cd(x_0,x)^\beta \leqslant c_\varepsilon d(x_0,x)^{\beta-\varepsilon}R^\varepsilon. \end{aligned}$$

Hence (2.9) holds for any $\gamma < \min(\beta,\delta)$; by Theorem 2.7, I_δ is continuous on $C^\alpha(X)$ for any $\alpha \leqslant \gamma, \alpha < \beta$, that is for any $\alpha < \min(\beta,\delta)$. □

3. Parabolic Carnot–Carathéodory distance

Let Ω be a bounded domain of $\mathbb{R}^n$, and let $X_1, X_2, \ldots, X_q$ be a system of smooth real vector fields defined in a neighborhood Ω_0 of Ω and satisfying Hörmander's condition of step s in Ω_0. Explicitly, this means that

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

with $b_{ik} \in C^\infty(\Omega_0)$, and the vector space spanned at every point of Ω_0 by: the fields X_i; their commutators $[X_i, X_j] = X_iX_j - X_jX_i$; 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 $\mathbb{R}^n$.

Let us recall the following definition.

Definition 3.1 *(Carnot–Carathéodory distance).* For $x, y \in \Omega_0$, let

$$d(x,y) = \inf\bigl\{T(\gamma) \mid \gamma : \bigl[0, T(\gamma)\bigr] \to \mathbb{R}^n\ X\text{-subunit},\ \gamma(0) = x,\ \gamma\bigl(T(\gamma)\bigr) = y\bigr\},$$

where we call X-subunit any absolutely continuous path γ such that

$$\gamma'(t) = \sum_{j=1}^m \lambda_j(t)X_j\bigl(\gamma(t)\bigr) \quad \text{a.e. with } \sum_{j=1}^m \lambda_j(t)^2 \leqslant 1 \quad \text{a.e.}$$

For $x \in \Omega$, we set

$$B_r(x) = \bigl\{y \in \Omega_0\colon d(x,y) < r\bigr\}.$$

It is well known (see [28]) that d is a distance (called Carnot–Carathéodory distance, or briefly CC-distance, induced by the system of Hörmander's vector fields X_i) and that there exist positive constants c, r_0, c_1, c_2 depending on Ω such that:

$$\begin{gathered} \bigl|B_{2r}(x)\bigr| \leqslant c\bigl|B_r(x)\bigr| \quad \text{for any } x \in \Omega, r \leqslant r_0, \\ c_1|x-y| \leqslant d(x,y) \leqslant c_2|x-y|^{1/s} \quad \text{for any } x, y \in \Omega, \end{gathered} \tag{3.1}$$

where s is the step appearing in Hörmander's condition.

In order to apply to a domain $A \subseteq \Omega$ the abstract theory of spaces of homogeneous type developed in Section 2, we need to know that in (A, d, dx) the doubling condition holds. Explicitly, this means that

$$\left|B_{2r}(x) \cap A\right| \leqslant c\left|B_r(x) \cap A\right| \quad \text{for any } x \in A,\ r > 0.$$

This requires some regularity property of ∂A.

Definition 3.2. Under the above assumptions, we say that a domain $A \subseteq \Omega$ is d-regular if

$$\left|B_r(x) \cap A\right| \geqslant c\left|B_r(x)\right|$$

for every $x \in A$, $0 < r < \operatorname{diam}(A)$.

In [7] we have proved the following criteria of regularity.

Lemma 3.3.

(i) *Let $A = B_R(x_0) \subset \Omega_0$ be a metric ball. Then, $B_R(x_0)$ is d-regular.*
(ii) *The union of a finite number of d-regular domains in Ω_0 is d-regular.*
(iii) *If A is a bounded d-regular domain in Ω_0, then (A, d, dx) is a space of homogeneous type.*

Let us now consider the parabolic Carnot–Carathéodory distance d_{P} corresponding to d, namely,

$$d_{\mathrm{P}}\big((t, x), (s, y)\big) = \sqrt{d(x, y)^2 + |t - s|},$$

defined in the cylinder $\mathbb{R} \times \Omega$.

One can easily prove:

Lemma 3.4. *Whenever $d(x, y)$ is a distance defined on some set Ω, $d_{\mathrm{P}}((t, x), (s, y))$ defined as above is a distance on $\mathbb{R} \times \Omega$.*

Notation 3.5. We will write $B_r(x)$ for the d-ball in Ω with radius r and center x, and $B_r(t, x)$ for the d_{P}-ball in $\mathbb{R} \times \Omega$ with radius r and center (t, x). In other words, with this notation the center of the ball reveals the dimension of the space.

To apply the theory developed in Section 2 to the space

$$\big(B_R(t_0, x_0), d_{\mathrm{P}}, dt\, dx\big)$$

we need to know that a d_{P}-ball $B_R(t_0, x_0)$ is d_{P}-regular. This fact will be actually proved in this section, and will require some labour.

First, we need to introduce some standard subsets related to parabolic geometry, namely:

- "parabolic cones" of the kind

$$C_r(t,x)=\left\{(\tau,z)\colon |t-\tau|<r^2,\ d(x,z)<r-\frac{|t-\tau|}{r}\right\},\quad \text{and}$$

- "parabolic cylinders"

$$Q_r(\tau,x)=\left\{(t,z)\colon |t-\tau|<r^2,\ d(x,z)<r\right\}.$$

Then we have:

Lemma 3.6. *The volume of the sets*

$$B_r(t,x),\quad C_r(t,x),\quad Q_r(t,x)$$

is equivalent to

$$r^2\left|B_r(x)\right|.$$

Moreover, if $d(x,y)\leqslant cr$, then $|B_r(t,x)|$ is equivalent to $|B_r(t,y)|$.

Proof. Obviously,

$$\left|Q_r(t,x)\right|=2r^2\left|B_r(x)\right|.$$

Moreover,

$$B_r(t,x)\subset Q_r(t,x)\quad\text{and}\quad C_r(t,x)\subset Q_r(t,x),$$

hence

$$\left|B_r(t,x)\right|\leqslant 2r^2\left|B_r(x)\right|\quad\text{and}\quad \left|C_r(t,x)\right|\leqslant 2r^2\left|B_r(x)\right|.$$

As to the estimates from below, we can write

$$\begin{aligned}\left|B_r(t,x)\right|&=\int_{t-r^2}^{t+r^2}d\tau\int_{d(x,y)<\sqrt{r^2-|t-\tau|}}dy=2\int_0^{r^2}\left|B_{\sqrt{r^2-\tau}}(x)\right|d\tau\\&\geqslant 2\int_0^{\frac{3}{4}r^2}\left|B_{r/2}(x)\right|d\tau=\frac{3}{2}r^2\left|B_{r/2}(x)\right|;\\ \left|C_r(t,x)\right|&=\int_{t-r^2}^{t+r^2}d\tau\int_{d(x,y)<\frac{r^2-|t-\tau|}{r}}dy=2\int_0^{r^2}\left|B_{\frac{r^2-\tau}{r}}(x)\right|d\tau\geqslant 2\int_0^{\frac{1}{2}r^2}\left|B_{r/2}(x)\right|d\tau=r^2\left|B_{r/2}(x)\right|.\end{aligned}\tag{3.2}$$

By the doubling property of $|B_r(x)|$, the result follows.

Finally, the last assertion holds because, since $d(x,y)\leqslant cr$, by the doubling condition on d,

$$\left|B_r(t,x)\right|\leqslant 2r^2\left|B_r(x)\right|\leqslant c_1r^2\left|B_r(y)\right|\leqslant c_2\left|B_r(t,y)\right|.\qquad\square$$

We also recall the following lemma.

Lemma 3.7. *Let $B_R(x_0)$ be a metric ball, $x \in B_R(x_0)$, $d(x, x_0) = \rho < R$. If $r < 3\rho$, then there exists x_1 such that*:

(i) $B_{r/3}(x_1) \subset B_R(x_0) \cap B_r(x)$;
(ii) $d(x_0, x_1) < \rho - \frac{r}{3}$;
(iii) $d(x_1, x) < \frac{2}{3}r$.

If $r \geqslant 3\rho$, then taking $x_1 = x_0$, properties (i)–(iii) *hold.*

The above lemma is contained in the proof of [7, Lemma 4.2]. We can now prove the following proposition.

Proposition 3.8. *Let $B_R(t_0, x_0)$ be a d_{P}-ball. Then $B_R(t_0, x_0)$ is d_{P}-regular, that is there exists $c > 0$ such that*

$$\left|B_R(t_0, x_0) \cap B_r(t, x)\right| \geqslant c\left|B_r(t, x)\right|$$

for every $(t, x) \in B_R(t_0, x_0)$, $0 < r < 2R$.

Proof. For $(t, x) \in B_R(t_0, x_0)$, let us consider the ball $B_r(t, x)$, for some $r \leqslant 2R$; let $\rho = d(x, x_0)$.

Case 1. We assume $r < 3\rho$. Let x_1 be as in Lemma 3.7. Then, we claim that

$$C_{r/3}(t, x_1) \subset B_R(t_0, x_0) \cap B_r(t, x). \tag{3.3}$$

Namely, let

$$(\tau, z) \in C_{r/3}(t, x_1) \equiv \left\{(\tau, z)\colon\ |t - \tau| < \frac{r^2}{9},\ d(x_1, z) < \frac{r}{3} - \frac{3}{r}|t - \tau|\right\}.$$

To prove that $(\tau, z) \in B_r(t, x)$, we write

$$d_{\mathrm{P}}\big((\tau, z), (t, x)\big) = \sqrt{d(x, z)^2 + |\tau - t|} \leqslant \sqrt{\big(d(x, x_1) + d(x_1, z)\big)^2 + |\tau - t|}$$

by Lemma 3.7(iii) and definition of $C_{r/3}(t, x_1)$

$$\leqslant \sqrt{\left(\frac{2}{3}r + \frac{r}{3} - \frac{3}{r}|t - \tau|\right)^2 + |\tau - t|} = \sqrt{\frac{9}{r^2}|t - \tau|^2 - 5|\tau - t| + r^2} \leqslant r$$

because the function $f(s) = \frac{9}{r^2}s^2 - 5s + r^2$, is decreasing in $[0, \frac{r^2}{9}]$, hence has its maximum at $s = 0$, and $f(0) = r^2$.

To prove that $(\tau, z) \in B_R(t_0, x_0)$, we write

$$\begin{aligned} d_{\mathrm{P}}\big((\tau, z), (t_0, x_0)\big) &= \sqrt{d(x_0, z)^2 + |\tau - t_0|} \\ &\leqslant \sqrt{\big(d(x_0, x_1) + d(x_1, z)\big)^2 + |\tau - t| + |t - t_0|} \end{aligned}$$

$$\leqslant \sqrt{\left(\rho - \frac{r}{3} + \frac{r}{3} - \frac{3}{r}|\tau - t|\right)^2 + |\tau - t| + |t - t_0|}$$

$$\leqslant \sqrt{\left(\rho^2 + |t - t_0|\right) + \left(\frac{9}{r^2}|\tau - t|^2 - \frac{6\rho}{r}|\tau - t| + |\tau - t|\right)} \leqslant R$$

because $\rho^2 + |t - t_0| \leqslant R^2$ and $\frac{9}{r^2}|\tau - t|^2 - \frac{6\rho}{r}|\tau - t| + |\tau - t| \leqslant 0$ for $r < 3\rho$.

Inclusion (3.3) and Lemma 3.6 imply that, in case 1,

$$\left|B_R(t_0, x_0) \cap B_r(t, x)\right| \geqslant \left|C_{r/3}(t, x_1)\right| \geqslant c_1\left|B_r(t, x_1)\right| \geqslant c_2\left|B_r(t, x)\right|.$$

The last inequality follows, again by Lemma 3.6, because $d(x, x_1) < \frac{2}{3}r$.

Case 2. We assume $r \geqslant 3\rho$ and $|t - t_0| \leqslant \frac{5}{9}R^2$. Under the assumption $r \geqslant 3\rho$, Lemma 3.7 states that

(i) $B_{r/3}(x_0) \subset B_R(x_0) \cap B_r(x)$;
(ii) $d(x_0, x) < \frac{2}{3}r$.

Let us show that (3.3) still holds, with $x_1 = x_0$, that is,

$$C_{r/3}(t, x_0) \subset B_R(t_0, x_0) \cap B_r(t, x). \tag{3.4}$$

Inclusion $C_{r/3}(t, x_0) \subset B_r(t, x)$ follows by the same proof as above. To show that $C_{r/3}(t, x_0) \subset B_R(t_0, x_0)$, let $(\tau, z) \in C_{r/3}(t, x_0)$; then

$$d_{\mathrm{P}}\big((\tau, z), (t_0, x_0)\big) = \sqrt{d(x_0, z)^2 + |\tau - t_0|} \leqslant \sqrt{\left(\frac{r}{3} - \frac{3}{r}|t - \tau|\right)^2 + |\tau - t| + |t - t_0|}$$

$$= \sqrt{\frac{9}{r^2}|t - \tau|^2 - |\tau - t| + \frac{r^2}{9} + \frac{5}{9}R^2} \leqslant \sqrt{\frac{r^2}{9} + \frac{5}{9}R^2} \leqslant R,$$

where we used the fact that $\frac{9}{r^2}|t - \tau|^2 - |\tau - t| \leqslant 0$ for $|\tau - t| < \frac{r^2}{9}$, and that $r < 2R$. This shows that also in case 2,

$$\left|B_R(t_0, x_0) \cap B_r(t, x)\right| \geqslant \left|C_{r/3}(t, x_0)\right| \geqslant c_1\left|B_r(t, x_0)\right| \geqslant c_2\left|B_r(t, x)\right|,$$

where the last inequality follows, by Lemma 3.6, because $d(x, x_0) = \rho < \frac{r}{3}$.

Case 3. We assume $r \geqslant 3\rho$ and $\frac{5}{9}R^2 < |t - t_0| < R^2$. Since $r \geqslant 3\rho$, as in case 2 we know that

(i) $B_{r/3}(x_0) \subset B_R(x_0) \cap B_r(x)$;
(ii) $d(x_0, x) < \frac{2}{3}r$.

To fix ideas, assume $t \geqslant t_0$ (the other case is identical), that is $t > t_0 + \frac{5}{9}R^2$. Let us define:

$$C^-_{r/3}(t, x_0) = \left\{(t, z)\colon\ t - \frac{r^2}{9} < \tau < t - \frac{r^2}{18},\ d(x_0, z) < \frac{r}{3} - \frac{3}{r}|t - \tau|\right\}.$$

We claim that

$$C^-_{r/3}(t, x_0) \subset B_R(t_0, x_0) \cap B_r(t, x). \tag{3.5}$$

Inclusion $C^-_{r/3}(t, x_0) \subset B_r(t, x)$ can be proved as in case 1. To show that $C^-_{r/3}(t, x_0) \subset B_R(t_0, x_0)$, let $(\tau, z) \in C^-_{r/3}(t, x_0)$; then $\tau > t_0$ and

$$\begin{aligned} d_P\big((\tau, z), (t_0, x_0)\big) = \sqrt{d(x_0, z)^2 + \tau - t_0} &\leqslant \sqrt{\left(\frac{r}{3} - \frac{3}{r}(t-\tau)\right)^2 + (t - t_0) - (t - \tau)} \\ &= \sqrt{\frac{9}{r^2}(t-\tau)^2 - 3(\tau - t) + \frac{r^2}{9} + R^2} \leqslant R \end{aligned}$$

because the function $f(s) = \frac{9}{r^2}s^2 - 3s + \frac{r^2}{9}$ for $s \in [\frac{r^2}{18}, \frac{r^2}{9}]$ is decreasing and attains its maximum at $s = \frac{r^2}{18}$, $f(\frac{r^2}{18}) = -\frac{r^2}{36} < 0$. Therefore (3.5) holds, and we conclude that in case 3,

$$\left|B_R(t_0, x_0) \cap B_r(t, x)\right| \geqslant \left|C^-_{r/3}(t, x_0)\right| \geqslant c_1\left|B_r(t, x_0)\right| \geqslant c_2\left|B_r(t, x)\right|,$$

where the second inequality follows by a similar computation to that of the proof of Lemma 3.6, and the last inequality follows by Lemma 3.6. This ends the proof of the proposition. □

4. Parabolic Hörmander Hölder spaces

We now define parabolic Hölder spaces adapted to this context. Let Ω be as in previous section. For any bounded domain $U \subset \mathbb{R} \times \Omega$ and any $\alpha > 0$, let

$$|u|_{C^\alpha(U)} = \sup\left\{\frac{|u(t,x) - u(s,y)|}{d_P((t,x),(s,y))^\alpha}\colon (t,x),\ (s,y) \in U,\ (t,x) \neq (s,y)\right\},$$

$$\|u\|_{C^\alpha(U)} = |u|_{C^\alpha(U)} + \|u\|_{L^\infty(U)},$$

$$C^\alpha(U) = \big\{u : U \to \mathbb{R}\colon \|u\|_{C^\alpha(U)} < \infty\big\}.$$

Note that, by (3.1), a function $u \in C^\alpha(U)$ is also continuous on U in Euclidean sense. By Lemma 3.4, d_P is a distance; if U is a d_P-regular domain (for instance, a d_P-ball), then $(U, d_P, dt\,dx)$ is a space of homogeneous type and, by Remark 2.4, the space $C^\alpha_0(U)$ is dense in $L^p(U)$ for any $\alpha \in (0, 1]$ and $p \in [1, \infty)$. We are going to show that for $\alpha > 1$, C^α spaces become trivial.

Proposition 4.1. *Let d be the Carnot–Carathéodory distance induced in a domain Ω by a system of Hörmander's vector fields $X_1, \dots, X_q$, and d_P the corresponding parabolic distance. Then*:

(i) *if $f(x) \in C^\alpha(\Omega)$ for some $\alpha > 1$, then f is constant in Ω*;
(ii) *if $f(t, x) \in C^\alpha(U)$ for some $\alpha > 2$, then f is constant in U; if $1 < \alpha \leqslant 2$, then f does not depend on x.*

Proof. (i) Let us show that $X_i f \equiv 0$ in Ω for $i = 1, 2, \dots, q$; then Hörmander's condition implies that the Euclidean gradient of f vanishes in Ω, hence f is constant. For any $x \in \Omega$, let $\gamma(t)$ be the integral curve of X_i such that

$$\begin{cases} \gamma'(t) = X_i(\gamma(t)), \\ \gamma(0) = x. \end{cases}$$

Then

$$X_i f(x) = \left[\frac{d}{dt} f\bigl(\gamma(t)\bigr)\right](0) = \lim_{t\to 0} \frac{f(\gamma(t)) - f(\gamma(0))}{t}. \tag{4.1}$$

Since γ is subunit (see Section 3), we can write

$$\bigl|f\bigl(\gamma(t)\bigr) - f\bigl(\gamma(0)\bigr)\bigr| \leqslant |f|_a d\bigl(\gamma(t), \gamma(0)\bigr)^\alpha \leqslant |f|_a t^\alpha$$

and, if $\alpha > 1$, this implies $X_i f(x) = 0$, by (4.1).

(ii) Applying (i) to the function $x \mapsto f(t, x)$ for fixed t, we deduce that if $\alpha > 1$ then f does not depend on x. Now, saying that $f(t)$ belongs to the parabolic C^α space means that

$$\bigl|f(t) - f(s)\bigr| \leqslant c|t - s|^{\alpha/2}$$

and this implies that f is constant if $\alpha > 2$. □

By the previous discussion, henceforth we will consider parabolic Hölder spaces C^α for $\alpha \in (0, 1)$.

For any positive integer k, let

$$C^{k,\alpha}(U) = \bigl\{u : U \to \mathbb{R}\colon \|u\|_{C^{k,\alpha}(U)} < \infty\bigr\} \quad \text{with}$$
$$\|u\|_{C^{k,\alpha}(U)} = \sum_{|I|+2h\leqslant k} \bigl\|\partial_t^h X^I u\bigr\|_{C^\alpha(U)},$$

where, for any multiindex $I = (i_1, i_2, \dots, i_s)$, with $1 \leqslant i_j \leqslant q$, we say that $|I| = s$ and

$$X^I u = X_{i_1} X_{i_2} \cdots X_{i_s} u.$$

Note that all the derivatives $\partial_t^h X^I u$ involved in the definition of $C^{k,\alpha}$ are continuous in Euclidean sense, because they belong to C^α.

We will also set $C_0^{k,\alpha}(U)$ for the space of $C^{k,\alpha}(U)$ functions compactly supported in U.

Occasionally, we will also use the space $C^{1,0}(U)$ of continuous functions u with continuous derivatives $X_i u$ (for $i = 1, 2, \dots, q$), and the corresponding space $C_0^{1,0}(U)$ of compactly supported functions.

The following proposition collects some simple facts about parabolic Hölder spaces, which will be used later.

Proposition 4.2. *Let U be as above.*

(i) *For any couple of functions $f, g \in C^{\alpha}(U)$, one has*

$$|fg|_{C^{\alpha}} \leqslant |f|_{C^{\alpha}} \|g\|_{L^{\infty}} + \|f\|_{L^{\infty}} |g|_{C^{\alpha}} \quad \textit{and} \tag{4.2}$$

$$\|fg\|_{C^{\alpha}} \leqslant 2\|f\|_{C^{\alpha}} \|g\|_{C^{\alpha}}. \tag{4.3}$$

Moreover, if both f and g vanish at least at a point of U, then

$$|fg|_{C^{\alpha}} \leqslant 2(\operatorname{diam} U)^{\alpha} |f|_{C^{\alpha}} |g|_{C^{\alpha}}. \tag{4.4}$$

Also, for any couple of functions $f, g \in C^{k,\alpha}(U)$

$$\|fg\|_{C^{k,\alpha}} \leqslant c_k \|f\|_{C^{k,\alpha}} \|g\|_{C^{k,\alpha}} \tag{4.5}$$

for some absolute constant c_k depending only on k.

(ii) *If $B_R(x_0)$ is a d-ball in $\mathbb{R}^n$, $f \in C^{1,0}(B_{5R}(x_0))$ one has*

$$\bigl|f(x) - f(y)\bigr| \leqslant \sup_{B_{5R}(x_0)} |Xf| \cdot d(x,y) \quad \textit{for any } x, y \in B_R(x_0), \tag{4.6}$$

$$\textit{where} \quad |Xf| = \sqrt{\sum_{i=1}^{q} (X_i f)^2}.$$

If $B_R(t_0, x_0)$ is a d_{P}-ball in $\mathbb{R}^{n+1}$, for any $f \in C_0^{1,0}(B_R(t_0, x_0))$ one has

$$\bigl|f(t,x) - f(s,y)\bigr| \leqslant \bigl(\sup|Xf| + R \sup|\partial_t f|\bigr) \cdot d_{\mathrm{P}}\bigl((t,x),(s,y)\bigr). \tag{4.7}$$

In particular,

$$|f|_{C^{\alpha}} \leqslant R^{1-\alpha} \cdot \bigl(\sup|Xf| + R \sup|\partial_t f|\bigr). \tag{4.8}$$

(iii) *If $U' \subset U$, then*

$$|f|_{C^{\alpha}(U')} \leqslant |f|_{C^{\alpha}(U)}. \tag{4.9}$$

(iv) *For any ball $B_R(t_0, x_0) \subset U$, for any $f \in C^{\alpha}(U)$, with* $\operatorname{sprt} f \subset B_R$ *we have*

$$|f|_{C^{\alpha}(U)} = |f|_{C^{\alpha}(B_R)}.$$

(v) *Let B_r^i $(i = 1, 2, \ldots, k)$ be a finite family of balls (in $\mathbb{R}^{n+1}$) of the same radius, such that $\bigcup_{i=1}^{k} B_{2r}^i \subset U$. Then for any $f \in C^{\alpha}(U)$,*

$$\|f\|_{C^{\alpha}(\bigcup_{i=1}^{k} B_r^i)} \leqslant c \sum_{i=1}^{k} \|f\|_{C^{\alpha}(B_{2r}^i)} \tag{4.10}$$

with c depending on the family of balls, but independent of f.

(vi) *The following interpolation inequality holds for the time derivative of any function* $f \in C_0^{2,\alpha}(U)$

$$\|f_t\|_{L^\infty} \leqslant r^{\alpha/2}|f_t|_{C^\alpha} + \frac{2}{r}\|f\|_{L^\infty} \quad \textit{for any } r > 0,\ \alpha \in (0,1). \tag{4.11}$$

Proof. The first two inequalities in (i) are obvious. The third follows from the second by the following remark: if $f(t_0, x_0) = 0$ for some $(t_0, x_0) \in U$, then for any $(t, x) \in U$,

$$\left|f(t,x)\right| = \left|f(t,x) - f(t_0,x_0)\right| \leqslant |f|_\alpha d_{\mathrm{P}}\big((t,x),(t_0,x_0)\big)^\alpha,$$

hence

$$\|f\|_\infty \leqslant |f|_\alpha (\operatorname{diam} U)^\alpha$$

and the same holds for g. Inequality (4.5) obviously follows from (4.3).

To prove (ii), for any fixed $\varepsilon > 0$, let γ be a subunit curve joining x, y such that

$$\gamma'(t) = \sum_{i=1}^q \lambda_i(t) X_i\big(\gamma(t)\big);$$
$$\gamma(0) = x; \quad \gamma(T) = y; \quad T \leqslant (1+\varepsilon)d(x,y).$$

Observe that $\gamma \subset B_{5R}(x_0)$: namely, for any $z \in \gamma$, let γ_z be the portion of γ which joins x to z, $\gamma(T_z) = z$, then

$$d(x,z) \leqslant T_z \leqslant T \leqslant (1+\varepsilon)d(x,y) \leqslant (1+\varepsilon)2R$$

for $x, y \in B_R(x_0)$, hence $d(z, x_0) \leqslant d(x,z) + d(x,x_0) < 5R$.

We have

$$f(y) - f(x) = f\big(\gamma(T)\big) - f\big(\gamma(0)\big) = \int_0^{\mathrm{T}} \frac{d}{dt}\big(f\big(\gamma(t)\big)\big)\,dt = \int_0^{\mathrm{T}} \sum_{i=1}^q \lambda_i(t) f'\big(\gamma(t)\big) X_i\big(\gamma(t)\big)\,dt.$$

Then

$$\begin{aligned}\left|f(y) - f(x)\right| &\leqslant \int_0^{\mathrm{T}} \sqrt{\sum_{i=1}^q \lambda_i(t)^2}\sqrt{\sum_{i=1}^q X_i f\big(\gamma(t)\big)^2}\,dt \\ &\leqslant \sup_{z \in B_{5R}(x_0)} \left|Xf(z)\right| \cdot T \leqslant (1+\varepsilon)d(x,y) \sup_{z \in B_{5R}(x_0)} \left|Xf(z)\right|.\end{aligned}$$

For vanishing ε we have (4.6). For functions depending also on t, the same reasoning gives

$$\begin{aligned}\left|f(t,x) - f(s,y)\right| &\leqslant \sup|Xf| \cdot d(x,y) + \sup|\partial_t f| \cdot |t-s| \\ &\leqslant \sup|Xf| \cdot d(x,y) + \sup|\partial_t f| \cdot |t-s|^{1/2} R \\ &\leqslant \big(\sup|Xf| + R\sup|\partial_t f|\big) \cdot d_{\mathrm{P}}\big((t,x),(s,y)\big)\end{aligned} \tag{4.12}$$

which is (4.7); this also implies (4.8).

Part (iii) is obvious. To prove (iv): $|f|_{C^\alpha(U)} \geqslant |f|_{C^\alpha(B_R)}$ is obvious; if $(t,x) \in B_R, (s,y) \notin B_R$, pick a subunit curve γ joining x to y, with

$$T(\gamma) \leqslant (1+\varepsilon) d(x,y)$$

and let $y^* \in \gamma$ such that $(t, y^*) \in \partial B_R$; then

$$d_P\big((t,x),(t,y^*)\big) = d(x,y^*) \leqslant (1+\varepsilon) d(x,y) \leqslant (1+\varepsilon) d_P\big((t,x),(s,y)\big).$$

Since $f(s,y) = f(t,y^*) = 0$ we have

$$\begin{aligned}\frac{|f(t,x) - f(s,y)|}{d_P((t,x),(s,y))^\alpha} &= \frac{|f(t,x)|}{d_P((t,x),(s,y))^\alpha} \leqslant (1+\varepsilon)^\alpha \frac{|f(t,x)|}{d_P((t,x),(t,y^*))^\alpha} \\ &= (1+\varepsilon)^\alpha \frac{|f(t,x) - f(t,y^*)|}{d_P((t,x),(t,y^*))^\alpha}\end{aligned}$$

therefore $|f|_{C^\alpha(U)} \leqslant (1+\varepsilon)^\alpha |f|_{C^\alpha(B_R)}$ for any $\varepsilon > 0$, and we are done.

Part (v) can be proved with a similar reasoning to that used in [7, Lemma 4.4]: let ζ_i $(i = 1, 2, \dots, k)$ be smooth cutoff functions such that $\operatorname{sprt} \zeta_i \subset B^i_{2r}$, $\sum_{i=1}^k \zeta_i = 1$ in $\bigcup_{i=1}^k B^i_r$. Then, subadditivity of the seminorm and (4.9) give

$$\begin{aligned}|f|_{C^\alpha(\bigcup_{j=1}^k B^j_r)} &\leqslant \sum_{i=1}^k |f\zeta_i|_{C^\alpha(\bigcup_{j=1}^k B^j_r)} \leqslant \sum_{i=1}^k |f\zeta_i|_{C^\alpha(\bigcup_{j=1}^k B^j_{2r})} \\ &= \sum_{i=1}^k |f\zeta_i|_{C^\alpha(B^i_{2r})} \leqslant 2\sum_{i=1}^k \|f\|_{C^\alpha(B^i_{2r})} \|\zeta_i\|_{C^\alpha(B^i_{2r})} \leqslant c\sum_{i=1}^k \|f\|_{C^\alpha(B^i_{2r})}\end{aligned}$$

by (iv), since $\operatorname{sprt} \zeta_i \subset B^i_{2r}$.

Part (vi) can be proved as in the Euclidean case (see [22, p. 124]):

$$\begin{aligned}f_t(t,x) &= f_t(t,x) - \big[f(t+1,x) - f(t,x)\big] + \big[f(t+1,x) - f(t,x)\big] \\ &= f_t(t,x) - f_t(t+\theta,x) + \big[f(t+1,x) - f(t,x)\big]\end{aligned}$$

for some $\theta \in (0,1)$. Then

$$\begin{aligned}\big|f_t(t,x)\big| &\leqslant \big|f_t(t,x) - f_t(t+\theta,x)\big| + 2\|f\|_{L^\infty} \\ &\leqslant \theta^{\alpha/2} |f_t|_{C^\alpha} + 2\|f\|_{L^\infty}.\end{aligned}$$

The same reasoning applied to the function $f(t,x) = g(rt,x)$ (for any $r > 0$) gives

$$r\big|f_t(t,x)\big| \leqslant (r\theta)^{\alpha/2} r |f_t|_{C^\alpha} + 2\|f\|_{L^\infty}$$

and, finally,

$$\|f_t\|_{L^\infty} \leqslant r^{\alpha/2} |f_t|_{C^\alpha} + \frac{2}{r}\|f\|_{L^\infty}. \qquad \square$$

5. Local Schauder estimates: statement of results and strategy of the proof

We are now in position to summarize our assumptions and main results.

(H1) Let Ω be a bounded domain of $\mathbb{R}^n$, and let $X_1, X_2, \ldots, X_q$ be a system of smooth real vector fields defined in a neighborhood Ω_0 of Ω and satisfying Hörmander's condition of step s in Ω_0.

(H2) Let U be a bounded domain of $\mathbb{R}^{n+1}$, $U \subset \mathbb{R} \times \Omega$; let $A = \{a_{ij}(t,x)\}_{i,j=1}^{q}$ be a symmetric, uniformly positive definite matrix of real functions defined in U, and let $\lambda > 0$ be a constant such that

$$\lambda^{-1}|\xi|^2 \leqslant \sum_{i,j=1}^{q} a_{ij}(t,x)\xi_i\xi_j \leqslant \lambda|\xi|^2 \quad \text{for every } \xi \in \mathbb{R}^q,\ (t,x) \in U.$$

(H3) Assume $a_{ij} \in C^{\alpha}(U)$ for some $\alpha \in (0,1)$.

We consider the differential operator:

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

Our basic result for the operator H is the following theorem.

Theorem 5.1. *Under the assumptions* (H1)–(H3)*, for every domain $U' \Subset U$ and $\alpha \in (0,1)$ there exists a constant $c > 0$ depending on $U, U', \{X_i\}, \alpha, \lambda$ and $\|a_{ij}\|_{C^{\alpha}(U)}$ such that for every $u \in C^{2,\alpha}_{\mathrm{loc}}(U)$ with $Hu \in C^{\alpha}(U)$ one has*

$$\|u\|_{C^{2,\alpha}(U')} \leqslant c\left\{\|Hu\|_{C^{\alpha}(U)} + \|u\|_{L^{\infty}(U)}\right\}.$$

We now outline the strategy of the proof.

To study H, we will use extensively results and techniques from [29] (in particular, Rothschild–Stein's technique of "lifting and approximation"), as well as from our previous papers [6,7]. We will briefly recall the basic definitions and results in Appendix A, which we refer to for our notation. For more details, the reader is referred to the papers quoted in Appendix A.

First of all, by Rothschild–Stein "lifting theorem," we lift the vector fields $X_i(x)$, defined in $\mathbb{R}^n$, to new vector fields $\widetilde{X}_i(\xi)$ defined on $\mathbb{R}^N$, with $\xi = (x,h)$, $h \in \mathbb{R}^{N-n}$. We also set $\tilde{a}_{ij}(t,\xi) = \tilde{a}_{ij}(t,x,h) = a_{ij}(t,x)$, $\widetilde{\Omega} = \Omega \times I$, where I is a neighborhood of the origin in $\mathbb{R}^{N-n}$, $\widetilde{U} = U \times I$ and

$$\widetilde{H} = \partial_t - \sum_{i,j=1}^{q} \tilde{a}_{ij}(t,\xi)\widetilde{X}_i\widetilde{X}_j.$$

All the notations and results introduced in Sections 3, 4 can now be applied to the system of Hörmander vector fields $\{\widetilde{X}_i\}$. To make explicit the context where we are now working, we will denote by $\tilde{d}$ the CC-distance induced in $\widetilde{\Omega}$ by the system $\{\widetilde{X}_i\}$, and by $\tilde{d}_{\mathrm{P}}$ its parabolic counterpart in $N+1$ variables. Accordingly, the symbol $\widetilde{B}_r(t_0,\xi_0)$ will denote the $\tilde{d}_{\mathrm{P}}$-ball of center (t_0,ξ_0) and radius r.

Following a general strategy already employed in [6,7], the proof of Theorem 5.1 will then proceed in three steps.

Step 1. C^α-estimates for $\widetilde{H}$, when u is a test function with small support in $\mathbb{R}^N$.

Theorem 5.2. *There exist $r,c>0$ such that for any $u\in C^{2,\alpha}(\widetilde{U})$, u compactly supported in some ball $\widetilde{B}_r(t_0,\xi_0)\subset\widetilde{U}$,*

$$\|u\|_{C^{2,\alpha}(\widetilde{B}_r)}\leqslant c\left\{\left\|\widetilde{H}u\right\|_{C^\alpha(\widetilde{B}_r)}+\|u\|_{L^\infty(\widetilde{B}_r)}\right\},$$

where c,r depend on $\{X_i\}$, α, λ and $\|a_{ij}\|_{C^\alpha(U)}$.

Step 2. C^α-estimates for $\widetilde{H}$ on a ball, for functions not necessarily vanishing at the boundary.

Theorem 5.3. *There exist positive constants r,c,β such that for any $u\in C^{2,\alpha}(\widetilde{B}_r(t_0,\xi_0))$, $0<t<s<r$,*

$$\|u\|_{C^{2,\alpha}(\widetilde{B}_t)}\leqslant\frac{c}{(s-t)^\beta}\left\{\left\|\widetilde{H}u\right\|_{C^\alpha(\widetilde{B}_s)}+\|u\|_{L^\infty(\widetilde{B}_s)}\right\},$$

where c,r depend on $\{X_i\}$, α, λ and $\|a_{ij}\|_{C^\alpha(U)}$, β depends on $\{X_i\}$, α.

Step 3. C^α-estimates for H on a ball, for any $u\in C^{2,\alpha}_{\mathrm{loc}}(U)$.

Theorem 5.4. *There exist positive constants r,c,β such that for any $u\in C^{2,\alpha}(B_r(t_0,x_0))$, $0<t<s<r$,*

$$\|u\|_{C^{2,\alpha}(B_t)}\leqslant\frac{c}{(s-t)^\beta}\left\{\|Hu\|_{C^\alpha(B_s)}+\|u\|_{L^\infty(B_s)}\right\},$$

where c,r depend on $\{X_i\}$, α, λ and $\|a_{ij}\|_{C^\alpha(U)}$, β depends on $\{X_i\}$, α.

Step 1 will be achieved in Section 6, exploiting the results of Sections 2–4, and adapting ideas and techniques already applied in [6,7,29]. Step 2 will be achieved in Section 7, and will follow from step 1 by standard properties of cutoff functions and suitable interpolation inequalities for Hölder norms, which will be proved there. These, in turn, rely both on results and techniques of Section 6, and on the abstract results proved in Section 2. Step 3 will be achieved in Section 8, and will follows from step 2 by known properties of the metrics induced by the vector fields $\{X_i\}$ and $\{\widetilde{X}_i\}$, provided we use an integral characterization of Hölder spaces, which is also proved in Section 8. Finally, by a covering argument, Theorem 5.1 immediately follows from step 3.

6. Operators of type l, parametrix and local estimates for functions of small support

In this section we will prove Theorem 5.2, that is the first step in the proof of our basic result, Theorem 5.1. We will use systematically notation and results borrowed from [16,29]; the reader is referred to Appendix A for the details.

Let us start again from the lifted operator

$$\widetilde{H}=\partial_t-\sum_{i,j=1}^{q}\tilde{a}_{ij}(t,\xi)\widetilde{X}_i\widetilde{X}_j.$$

By Rothschild–Stein "approximation theorem" (see [6, Theorem 1.6]), we can locally approximate the vector fields $\widetilde{X}_i$ with left invariant vector fields Y_i defined on a homogeneous group $\mathbb{G}$ (which is actually $\mathbb{R}^N$ endowed with a suitable Lie group structure). This approximation is expressed by the following identity which holds for every $f\in C_0^\infty(\mathbb{G})$:

$$\widetilde{X}_i\big(f\big(\Theta_\xi(\cdot)\big)\big)(\eta)=\big(Y_i f+R_i^\xi f\big)\big(\Theta_\xi(\eta)\big), \tag{6.1}$$

where $\Theta_\xi(\eta)=\Theta(\xi,\eta)$ is a local diffeomorphism in $\mathbb{R}^N$, and the vector fields R_i^ξ are remainders in a suitable sense (see Appendix A or [6]). The superscript ξ in R_i^ξ recalls that these vector fields depend on the point ξ, while they act as derivatives with respect to η.

We now freeze $\widetilde{H}$ at some point $(t_0,\xi_0)\in\widetilde{U}$, and consider the frozen lifted operator

$$\widetilde{H}_0=\partial_t-\sum_{i,j=1}^{q}\tilde{a}_{ij}(t_0,\xi_0)\widetilde{X}_i\widetilde{X}_j.$$

To study $\widetilde{H}_0$, we will consider its approximating operator, defined on $\mathbb{G}'=\mathbb{R}\times\mathbb{G}$:

$$\mathcal{H}_0=\partial_t-\sum_{i,j=1}^{q}\tilde{a}_{ij}(t_0,\xi_0)Y_iY_j.$$

Here we regard $\mathbb{G}'$ as a homogeneous group, with translations

$$(t,\xi)\circ(s,\eta)=(t+s,\xi\circ\eta),$$

dilations

$$D(\lambda)(t,\xi)=\big(\lambda^2t,D(\lambda)\xi\big)$$

and homogeneous dimension $Q'=Q+2$, where Q is the homogeneous dimension of $\mathbb{G}$. Since $\mathcal{H}_0$ is left invariant and homogeneous of degree 2 in $\mathbb{G}'$, by known results by Folland (see [16, Sections 2, 3]), it has a fundamental solution, denoted by

$$h(t_0,\xi_0,s,u)$$

which is homogeneous of degree $2-Q'=-Q$. Also, $h(t_0,\xi_0,s,u)$ is nonnegative and vanishes for $s<0$.

Throughout this section, $\tilde{d}$ will denote the CC-distance induced in $\widetilde{\Omega}$ by the system $\{\widetilde{X}_i\}$, and $\tilde{d}_P$ its parabolic counterpart in $N+1$ variables. Moreover, we will use the quasidistance, introduced by Rothschild and Stein in [29]:

$$\tilde{d}'(\xi,\eta)=\big\|\Theta(\xi,\eta)\big\|,$$

where $\|\cdot\|$ is the homogeneous norm in $\mathbb{G}$; note that $\tilde{d}'$ is defined only locally and it is a quasidistance, equivalent to $\tilde{d}$; we will also set

$$\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\eta)\big)=\sqrt{\tilde{d}'(\xi,\eta)^2+|t-s|}.$$

Obviously, $\tilde{d}'_{\mathrm{P}}$ is a quasidistance, equivalent to $\tilde{d}_{\mathrm{P}}$. Note also that, denoting by

$$\widetilde{B}\big((t,\xi);(s,\eta)\big)=\widetilde{B}_{\tilde{d}_{\mathrm{P}}((t,\xi),(s,\eta))}(t,\xi),$$

we have

$$\big|\widetilde{B}\big((t,\xi);(s,\eta)\big)\big|\simeq \tilde{d}_{\mathrm{P}}\big((t,\xi),(s,\eta)\big)^{Q+2}.$$

Notation 6.1. Henceforth we will use the symbol D^k to understand the sum of all space derivatives of order k. For instance, in the statement of the lemma here below, the symbol

$$\big\|f\,\partial_t^h D^k\varphi\big\|_{C^\alpha}$$

stands for

$$\sum_{1\leqslant i_j\leqslant q}\big\|f\,\partial_t^h \widetilde{X}_{i_1}\cdots\widetilde{X}_{i_k}\varphi\big\|_{C^\alpha}.$$

Lemma 6.2 *(Cutoff functions). For any $0<\rho<r$, $(t,\xi)\in\mathbb{R}^{N+1}$ there exists $\varphi\in C_0^\infty(\mathbb{R}^{N+1})$ with the following properties*:

(i) $0\leqslant\varphi\leqslant 1$, $\varphi\equiv 1$ *on* $\widetilde{B}_\rho(t,\xi)$ *and* $\operatorname{sprt}\varphi\subseteq\widetilde{B}_r(t,\xi)$;

(ii)
$$\big|\partial_t^h D^k\varphi\big|\leqslant\frac{c_{k,h}}{(r-\rho)^{k+2h}}\quad\textit{for } k,h\in\mathbb{N};\tag{6.2}$$

(iii) *for any $f\in C^\alpha$,*

$$\big\|f\,\partial_t^h D^k\varphi\big\|_{C^\alpha}\leqslant\frac{c_{k,h}}{(r-\rho)^{k+2h+1}}\|f\|_{C^\alpha}\quad\textit{for } k,h\in\mathbb{N}\tag{6.3}$$

and $r-\rho$ small enough.

We will write

$$\widetilde{B}_\rho(t,\xi)\prec\varphi\prec\widetilde{B}_r(t,\xi)$$

to indicate that φ satisfies all the previous properties.

Proof of Lemma 6.2. Since $\widetilde{B}_\rho\prec\varphi\prec\widetilde{B}_r$ implies $\widetilde{B}_{\rho'}\prec\varphi\prec\widetilde{B}_r$ for any $\rho'<\rho$, we can assume without loss of generality that $\rho\geqslant r/2$.

The proof of (i), (ii) is very similar to the proof of [6, Lemma 3.3]; we repeat it for convenience of the reader. Pick a function $f:[0,\infty)\to[0,1]$ satisfying:

$$f\equiv 1\quad\text{in }[0,\rho],\qquad f\equiv 0\quad\text{in }[r,\infty),\qquad f\in C^\infty(0,\infty),$$
$$\big|f^{(k)}\big|\leqslant\frac{c_k}{(r-\rho)^k}\quad\text{for } k=1,2,\ldots.\tag{6.4}$$

Setting $\varphi(s,\eta) = f(\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta)))$, we can compute

$$\widetilde{X}_i\varphi(s,\eta) = f'\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\eta)\big)\big)\widetilde{X}_i\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\cdot)\big)\big)(\eta), \tag{6.5}$$

$$\begin{aligned}\widetilde{X}_i\widetilde{X}_j\varphi(s,\eta) &= f''\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\eta)\big)\big)\,\widetilde{X}_i\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\cdot)\big)\big)(\eta)\widetilde{X}_j\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\cdot)\big)\big)(\eta)\\ &\quad + f'\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\eta)\big)\big)\,\widetilde{X}_i\widetilde{X}_j\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\cdot)\big)\big)(\eta).\end{aligned} \tag{6.6}$$

Next, we use the approximation theorem:

$$\begin{aligned}\widetilde{X}_i\Big(\sqrt{\big\|\Theta(\xi,\cdot)\big\|^2+|t-s|}\Big)(\eta) &= \big(\big(Y_i+R_i^\xi\big)\big(\sqrt{\|\cdot\|^2+|t-s|}\big)\big)\big(\Theta(\xi,\eta)\big)\\ &= \frac{\|\Theta(\xi,\eta)\|}{\sqrt{\|\Theta(\xi,\eta)\|^2+|t-s|}}\big(\big(Y_i+R_i^\xi\big)\big(\|\cdot\|\big)\big)\big(\Theta(\xi,\eta)\big).\end{aligned} \tag{6.7}$$

By homogeneity of the norm, $Y_i(\|u\|)$ is bounded and, since R_i^ξ has local degree $\leqslant 0$, $R_i^\xi(\|u\|)$ is also uniformly bounded; hence

$$\big|\widetilde{X}_i\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\cdot)\big)\big)(\eta)\big| \leqslant c. \tag{6.8}$$

Analogously,

$$\big|\widetilde{X}_i\widetilde{X}_j\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\cdot)\big)\big)(\eta)\big| \leqslant \frac{c}{\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta))} \tag{6.9}$$

for $\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta))$ small enough. Then (6.4), (6.5), (6.8) imply

$$\big|\widetilde{X}_i\varphi(s,\eta)\big| \leqslant \frac{c}{r-\rho}.$$

Since $f'(\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta))) \neq 0$ for $\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta)) > \rho$, (6.4), (6.6), (6.8), (6.9) imply:

$$\big|\widetilde{X}_i\widetilde{X}_j\varphi(s,\eta)\big| \leqslant \frac{c}{(r-\rho)^2} + \frac{c}{\rho(r-\rho)} \leqslant c\frac{r}{\rho(r-\rho)^2} \leqslant \frac{c}{(r-\rho)^2};$$

proceeding analogously we get

$$\big|D^k\varphi(s,\eta)\big| \leqslant c\sum_{i=0}^{k-1}\frac{1}{\rho^i(r-\rho)^{k-i}} = \frac{c}{(r-\rho)^k}\sum_{i=0}^{k-1}\frac{(r-\rho)^i}{\rho^i} \leqslant \frac{c}{(r-\rho)^k}.$$

Moreover,

$$\partial_s\varphi(s,\eta) = f'\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\eta)\big)\big)\partial_s\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(\cdot,\eta)\big)\big)(s) \quad \text{and}$$

$$\big|\partial_s\big(\tilde{d}'_{\mathrm{P}}\big((t,\xi),(\cdot,\eta)\big)\big)(s)\big| \leqslant \frac{c}{\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta))},$$

hence

$$\left|\partial_s \varphi(s,\eta)\right| \leqslant \frac{c}{\rho(r-\rho)} \leqslant c\frac{r}{\rho(r-\rho)^2} \leqslant \frac{c}{(r-\rho)^2}.$$

Analogously,

$$\left|\partial_s^h \varphi(s,\eta)\right| \leqslant \frac{c}{(r-\rho)^{2h}}.$$

Combining these computations we can complete the proof of (i) and (ii).

To prove (iii), we apply (4.12) to $\partial_s^h D^k \varphi$. By (6.2) we get

$$\begin{aligned}
&\left|\partial_s^h D^k \varphi(u,\zeta) - \partial_s^h D^k \varphi(s,\eta)\right| \\
&\quad \leqslant \sup\left|\partial_s^h D^{k+1}\varphi\right|\tilde{d}(\zeta,\eta) + \sup\left|\partial_s^{h+1} D^k \varphi\right||u-s| \\
&\quad \leqslant c_{k,h}\left[\frac{1}{(r-\rho)^{k+2h+1}}\tilde{d}'_{\mathrm{P}}\big((u,\zeta),(s,\eta)\big) + \frac{1}{(r-\rho)^{k+2h+2}}\tilde{d}'_{\mathrm{P}}\big((u,\zeta),(s,\eta)\big)^2\right].
\end{aligned}$$

Now, if $\tilde{d}'_{\mathrm{P}}((u,\zeta),(s,\eta)) \leqslant r-\rho$, then

$$\left|\partial_s^h D^k \varphi(u,\zeta) - \partial_s^h D^k \varphi(s,\eta)\right| \leqslant \frac{c_{k,h}}{(r-\rho)^{k+2h+1}}\tilde{d}'_{\mathrm{P}}\big((u,\zeta),(s,\eta)\big);$$

if $\tilde{d}'_{\mathrm{P}}((u,\zeta),(s,\eta)) > r-\rho$, then

$$\begin{aligned}
\left|\partial_s^h D^k \varphi(u,\zeta) - \partial_s^h D^k \varphi(s,\eta)\right| &\leqslant \left|\partial_s^h D^k \varphi(u,\zeta)\right| + \left|\partial_s^h D^k \varphi(s,\eta)\right| \\
&\leqslant \frac{c}{(r-\rho)^{k+2h}} \leqslant \frac{c}{(r-\rho)^{k+2h}} \cdot \frac{\tilde{d}'_{\mathrm{P}}((u,\xi),(s,\eta))}{r-\rho} \\
&\leqslant \frac{c_{k,h}}{(r-\rho)^{k+2h+1}}\tilde{d}'_{\mathrm{P}}\big((u,\zeta),(s,\eta)\big).
\end{aligned}$$

This, together with (6.2), means that

$$\left\|\partial_s^h D^k \zeta\right\|_{C^\alpha} \leqslant \frac{c_{k,h}}{(r-\rho)^{k+2h+1}},$$

which by (4.3) implies (6.3). □

Let us recall the key definition which describes the singular and fractional integral operators which appear in this context.

Definition 6.3. As above, let $h(t_0,\xi_0,s,u)$ be the fundamental solution of $\mathcal{H}_0$, homogeneous of degree $2-Q' = -Q$. We say that $k(t_0,\xi_0;t,\xi,\eta)$ is a frozen kernel of type ℓ, for some nonnegative integer ℓ, if for every positive integer m there exists a positive integer H_m such that

$$\begin{aligned}
k(t_0,\xi_0;t,\xi,\eta) = \sum_{i=1}^{H_m} a_i(\xi)b_i(\eta)\big[D_i h(t_0,\xi_0;\cdot)\big]\big(t,\Theta(\eta,\xi)\big) \\
+ a_0(\xi)b_0(\eta)\big[D_0 h(t_0,\xi_0;\cdot)\big]\big(t,\Theta(\eta,\xi)\big),
\end{aligned}$$

where a_i, b_i $(i = 0, 1, \dots, H_m)$ are test functions, D_i are differential operators such that: for $i = 1, \dots, H_m$, D_i is homogeneous of degree $\leqslant 2 - \ell$ (so that $D_i h(t_0, \xi_0; \cdot)$ is a homogeneous function of degree $\geqslant \ell - Q'$), and D_0 is a differential operator such that $D_0 h(t_0, \xi_0; \cdot)$ has m derivatives with respect to the vector fields Y_i $(i = 1, \dots, q)$.

We say that $T(t_0, \xi_0)$ is a frozen operator of type $\ell \geqslant 1$ if $k(t_0, \xi_0; t, \xi, \eta)$ is a frozen kernel of type ℓ and

$$T(t_0, \xi_0) f(t, \xi) = \int_{-\infty}^{t} \int_{\mathbb{R}^N} k(t_0, \xi_0; t - s, \xi, \eta)\, f(s, \eta)\, ds\, d\eta;$$

we say that $T(t_0, \xi_0)$ is a frozen operator of type 0 if $k(t_0, \xi_0; t, \xi, \eta)$ is a frozen kernel of type 0 (or "frozen singular integral") and

$$\begin{aligned} T(t_0, \xi_0) f(t, \xi) = PV \int_{-\infty}^{t} \int_{\mathbb{R}^N} k(t_0, \xi_0; t - s, \xi, \eta)\, f(s, \eta)\, ds\, d\eta \\ + \alpha(t_0, \xi_0) \beta(t, \xi) f(t, \xi), \end{aligned} \tag{6.10}$$

where α is bounded and β is smooth. Explicitly, the principal value of the integral is defined as

$$PV \int_{-\infty}^{t} \int_{\mathbb{R}^N} \dots\, ds\, d\eta = \lim_{\varepsilon \to 0} \int_{\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta)) > \varepsilon} \dots\, ds\, d\eta.$$

The link between this definition and the abstract theory of Section 2 is contained in the following proposition.

Proposition 6.4. *Let*

$$k_j(t, \xi, s, \eta) = a(\xi) b(\eta) \big[D_j h(t_0, \xi_0; \cdot)\big]\big(t - s, \Theta(\eta, \xi)\big)$$

be a kernel like those appearing in Definition 6.3*, with D_j differential operator homogeneous of degree j* (*we now leave the dependence on the frozen point* (t_0, ξ_0) *implicitly understood*)*. Then*:

(i) (*growth condition*) k_j *satisfies* (2.4) *in the form*:

$$\big|k_j(t, \xi, s, \eta)\big| \leqslant \frac{c}{\tilde{d}_{\mathrm{P}}((t, \xi), (s, \eta))^{Q+j}} \leqslant c \frac{\tilde{d}_{\mathrm{P}}((t, \xi), (s, \eta))^{2-j}}{|\widetilde{B}((t, \xi); (s, \eta))|};$$

(ii) (*mean value inequality*) k_j *satisfies* (2.5) *in the form*:

$$\begin{aligned} \big|k_j(t, \xi, s, \eta) - k_j(t_1, \xi_1, s, \eta)\big| &\leqslant c \frac{\tilde{d}_{\mathrm{P}}((t_1, \xi_1), (t, \xi))}{\tilde{d}_{\mathrm{P}}((t_1, \xi_1), (s, \eta))^{Q+j+1}} \\ &\leqslant c \frac{\tilde{d}_{\mathrm{P}}((t_1, \xi_1), (t, \xi))^{2-j}}{|\widetilde{B}((t, \xi); (s, \eta))|} \cdot \left(\frac{\tilde{d}_{\mathrm{P}}((t_1, \xi_1), (t, \xi))}{\tilde{d}_{\mathrm{P}}((t_1, \xi_1), (s, \eta))}\right) \end{aligned}$$

when $\tilde{d}_{\mathrm{P}}((t_1, \xi_1), (s, \eta)) > 2\tilde{d}_{\mathrm{P}}((t_1, \xi_1), (t, \xi))$;

(iii) (*cancellation properties*) *if* $j=2$, *then* k_j *satisfies property* (2.8), *in the form*

$$\left|\int_{r<\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta))<R} k_j(t,\xi,s,\eta)\,ds\,d\eta\right|\leqslant c$$

with c *independent of* r, R, *and satisfies* (2.9), *in the form*

$$\lim_{\varepsilon\to 0}\left|\int_{\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta))>\varepsilon} k_j(t,\xi,s,\eta)\,ds\,d\eta-\int_{\tilde{d}'_{\mathrm{P}}((t_1,\xi_1),(s,\eta))>\varepsilon} k_j(t_1,\xi_1,s,\eta)\,ds\,d\eta\right|$$
$$\leqslant c\tilde{d}_{\mathrm{P}}(t_1,\xi_1,t,\xi)^{\gamma}$$

for every $\gamma\in(0,1)$.

Remark 6.5. Point (ii) of this proposition is similar to, but sharper than, Proposition 2.17 of [6]. The point is that, to get Schauder estimates for any $\alpha\in(0,1)$, here we need (2.5) with exponent $\beta=1$ at the numerator, while in [6], following Rothschild and Stein, we only get $\beta=1/s$ with $s=$ step of Hörmander's condition. Also, point (iii) of this proposition is similar to, but stronger than, Lemma 4.11 of [7].

Proof of Proposition 6.4. By the uniform Gaussian estimates proved in [1] for the fundamental solution of $\mathcal{H}_0$, we know that

$$\left|\partial_s^k Y_{i_1}Y_{i_2}\cdots Y_{i_r}h(t_0,\xi_0,s,u)\right|\leqslant c_1\frac{e^{-c_2\|u\|^2/s}}{s^{Q/2+k+r/2}} \tag{6.11}$$

with c_1,c_2 independent of (t_0,ξ_0). More generally, if D_k is a differential operator homogeneous of degree k, we can write:

$$\begin{aligned}\left|D_k h(t_0,\xi_0,s,u)\right|&\leqslant c_1\frac{e^{-c_2\|u\|^2/s}}{s^{Q/2+k/2}}\leqslant c_1\left(\frac{\|u\|^2+s}{s}\right)^{\frac{Q+k}{2}}\frac{e^{-c_2\|u\|^2/s}}{(\|u\|^2+s)^{(Q+k)/2}}\\&\leqslant\frac{c_3}{(\|u\|^2+s)^{(Q+k)/2}}\end{aligned} \tag{6.12}$$

because the function $\alpha\mapsto(1+\alpha)^{\frac{Q+k}{2}}e^{-c_2\alpha}$ is bounded on $[0,\infty)$. Again, the constant c_3 is independent of (t_0,ξ_0). This implies, for the kernel k_j in the statement of this proposition,

$$\left|k_j(t,\xi,s,\eta)\right|\leqslant\frac{c_3}{(\|\Theta(\eta,\xi)\|^2+|t-s|)^{(Q+j)/2}}\leqslant\frac{c}{\tilde{d}_{\mathrm{P}}((t,\xi),(s,\eta))^{Q+j}}$$

which is (i).

To prove (ii), fix (t_1,ξ_1), (s,η), and let $2r=\tilde{d}_{\mathrm{P}}(t_1,\xi_1,s,\eta)$; then condition

$$\tilde{d}_{\mathrm{P}}\big((t_1,\xi_1),(s,\eta)\big)>2\tilde{d}_{\mathrm{P}}\big((t_1,\xi_1),(t,\xi)\big)$$

means that (t,ξ) is a point ranging in $\widetilde{B}_r(t_1,\xi_1)$. Let $\varphi(t,\xi)$ be a cutoff function such that

$$\widetilde{B}_r(t_1,\xi_1) \prec \varphi \prec \widetilde{B}_{\frac{3}{2}r}(t_1,\xi_1)$$

(see Lemma 6.2) and let

$$u(t,\xi) = k_j(t,\xi,s,\eta)\varphi(t,\xi).$$

Then $u \in C_0^{1,0}(\widetilde{B}_{\frac{3}{2}r}(t_1,\xi_1))$, and, for $\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(t,\xi)) < r$, we can apply property (ii) of Proposition 4.2:

$$\begin{aligned}
&\bigl|k_j(t,\xi,s,\eta) - k_j(t_1,\xi_1,s,\eta)\bigr| \\
&\quad = \bigl|u(t,\xi) - u(t_1,\xi_1)\bigr| \\
&\quad \leqslant \tilde{d}_{\mathrm{P}}\bigl((t_1,\xi_1),(t,\xi)\bigr)\Bigl\{\sup_{(\tau,\zeta)\in\widetilde{B}_{\frac{3}{2}r}(t_1,\xi_1)}\bigl|\widetilde{X}u(\tau,\zeta)\bigr| + \frac{3}{2}r\sup_{(\tau,\zeta)\in\widetilde{B}_{\frac{3}{2}r}(t_1,\xi_1)}\bigl|u_t(\tau,\zeta)\bigr|\Bigr\}. \qquad (6.13)
\end{aligned}$$

Now,

$$\begin{aligned}
\widetilde{X}_h u(\tau,\zeta) &= \widetilde{X}_h k_j(\tau,\zeta,s,\eta)\varphi(\tau,\zeta) + k_j(\tau,\zeta,s,\eta)\widetilde{X}_h\varphi(\tau,\zeta) = I + II \quad \text{and} \\
\widetilde{X}_h k_j(\tau,\zeta,s,\eta) &= a(\zeta)b(\eta)(Y_h D_j h)\bigl(t_0,\xi_0,\tau-s,\Theta(\zeta,\eta)\bigr) \\
&\quad + a(\zeta)b(\eta)\bigl(R_h^{\zeta} D_j h\bigr)\bigl(t_0,\xi_0,\tau-s,\Theta(\zeta,\eta)\bigr) \\
&\quad + a(\zeta)\widetilde{X}_h b(\eta)(D_j h)\bigl(t_0,\xi_0,\tau-s,\Theta(\zeta,\eta)\bigr),
\end{aligned}$$

so that

$$|I| \leqslant \frac{c\varphi(\tau,\zeta)}{\tilde{d}_{\mathrm{P}}((\tau,\zeta),(s,\eta))^{Q+j+1}} \leqslant \frac{c}{r^{Q+j+1}} \leqslant \frac{c}{\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(s,\eta))^{Q+j+1}},$$

where we have used the fact that, for $(\tau,\zeta) \in \widetilde{B}_{(3/2)r}(t_1,\xi_1)$ and $2r = \tilde{d}_{\mathrm{P}}((t_1,\xi_1),(s,\eta))$, we have $\tilde{d}_{\mathrm{P}}((\tau,\zeta),(s,\eta)) \geqslant cr$. On the other hand,

$$|II| \leqslant \frac{c}{r}\cdot\frac{c}{\tilde{d}_{\mathrm{P}}((\tau,\zeta),(s,\eta))^{Q+j}} \leqslant \frac{c}{r^{Q+j+1}} \leqslant \frac{c}{\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(s,\eta))^{Q+j+1}}.$$

Similarly,

$$u_t(\tau,\zeta) = \partial_t k_j(\tau,\zeta,s,\eta)\varphi(\tau,\zeta) + k_j(\tau,\zeta,s,\eta)\varphi_t(\tau,\zeta) = I_t + II_t$$

with

$$\begin{aligned}
|I_t| &\leqslant c\bigl|\partial_t D_j h\bigl(t_0,\xi_0,\tau-s,\Theta(\zeta,\eta)\bigr)\varphi(\tau,\zeta)\bigr| \leqslant \frac{c\varphi(\tau,\zeta)}{\tilde{d}_{\mathrm{P}}((\tau,\zeta),(s,\eta))^{Q+j+2}} \leqslant \frac{c}{r^{Q+j+2}} \quad \text{and} \\
|II_t| &\leqslant \frac{c}{r^2}\cdot\frac{c}{\tilde{d}_{\mathrm{P}}((\tau,\zeta),(s,\eta))^{Q+j}} \leqslant \frac{c}{r^{Q+j+2}}.
\end{aligned}$$

Therefore,

$$r \sup_{(\tau,\zeta)\in \widetilde{B}_{\frac{3}{2}r}(t_1,\xi_1)} \left|u_t(\tau,\zeta)\right| \leqslant \frac{c}{r^{Q+j+1}} \leqslant \frac{c}{\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(s,\eta))^{Q+j+1}}$$

and, finally, by (6.13), we get

$$\left|k(t,\xi,s,\eta)-k(t_1,\xi_1,s,\eta)\right| \leqslant c\frac{\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(t,\xi))}{\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(s,\eta))^{Q+j+1}}$$

when $\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(s,\eta)) \geqslant 2\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(t,\xi))$.

To prove (iii) when $j=2$, we proceed similarly to the proof of [7, Lemma 4.11]: if

$$k_2(t,\xi,s,\eta)=a(\xi)b(\eta)\left[D_2h(t_0,\xi_0;\cdot)\right]\left(t-s,\Theta(\eta,\xi)\right),$$

where $[D_2h(t_0,\xi_0;\cdot)]$ is homogeneous of degree $-Q'$, we split k_j as follows:

$$\begin{aligned} k_2(t,\xi,s,\eta) &= \frac{a(\xi)b(\xi)c(\xi)[D_2h(t_0,\xi_0;\cdot)](t-s,\Theta(\eta,\xi))}{g(\xi,\Theta(\xi,\eta))} \\ &\quad + \frac{a(\xi)b(\xi)[D_2h(t_0,\xi_0;\cdot)](t-s,\Theta(\eta,\xi))}{g(\xi,\Theta(\xi,\eta))}\left[g\left(\xi,\Theta(\xi,\eta)\right)-c(\xi)\right] \\ &\quad + a(\xi)\left[b(\eta)-b(\xi)\right]\left[D_2h(t_0,\xi_0;\cdot)\right]\left(t-s,\Theta(\eta,\xi)\right) \\ &\equiv k_a(t,\xi,s,\eta)+k_b(t,\xi,s,\eta)+k_c(t,\xi,s,\eta), \end{aligned}$$

where g and $c(\xi)$ are the functions appearing in the following formula of change of variables (see [6, Theorem 1.7]):

$$u=\Theta(\xi,\eta); \qquad d\eta=g(\xi,u)\,du; \quad g(\xi,u)=c(\xi)\left(1+O\left(\|u\|\right)\right).$$

We will prove that k_a, k_b, k_c satisfy (2.8) and (2.9). First, let us note that k_a is singular, but satisfies the strong vanishing property, with respect to the quasidistance $\tilde{d}'_{\mathrm{P}}$:

$$\int_{r<\tilde{d}'_{\mathrm{P}}((\xi,t),(\eta,s))<R} k_a(t,\xi,s,\eta)\,d\eta\,ds = a(\xi)b(\xi)c(\xi)\int_{r<\sqrt{\|u\|^2+|s|}<R} D_2h(t_0,\xi_0;s,u)\,du\,ds=0,$$

where the last integral vanishes by a known property of homogeneous distributions of degree $-Q'$ in homogeneous groups (see [16, Proposition 1.8]). Hence k_a obviously satisfies (2.8) and (2.9), for any γ.

On the other side, let us check that both k_b and k_c are fractional integral kernels that satisfy properties (2.12) and (2.13) with $\beta=\delta=1$. As we have seen in the proof of Theorem 2.11, this implies that k_b and k_c also satisfy (2.8) and (2.9) with any $\gamma<1$. Namely,

$$\begin{aligned} k_b(t,\xi,s,\eta) &= \frac{a(\xi)b(\xi)[D_2h(t_0,\xi_0;\cdot)](t-s,\Theta(\eta,\xi))}{g(\xi,\Theta(\xi,\eta))}\big[g\big(\xi,\Theta(\xi,\eta)\big)-c(\xi)\big] \\ &= \frac{a(\xi)b(\xi)[D_2h(t_0,\xi_0;\cdot)](t-s,\Theta(\eta,\xi))}{1+O(\|\Theta(\xi,\eta)\|)}O\big(\big\|\Theta(\xi,\eta)\big\|\big), \end{aligned}$$

so that

$$\big|k_b(t,\xi,s,\eta)\big| \leqslant \frac{c|a(\xi)b(\xi)|}{\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta))^{Q'-1}} \leqslant \frac{c|a(\xi)b(\xi)|}{\tilde{d}_{\mathrm{P}}((t,\xi),(s,\eta))^{Q'-1}}.$$

Finally, since b is smooth, and

$$|\xi-\eta| \leqslant c\tilde{d}(\xi,\eta) \leqslant c\tilde{d}'(\xi,\eta) \leqslant c\tilde{d}'_{\mathrm{P}}\big((t,\xi),(s,\eta)\big),$$

$$\begin{aligned} \big|k_c(t,\xi,s,\eta)\big| &= \big|a(\xi)\big[b(\eta)-b(\xi)\big]\big[Dh(t_0,\xi_0;\cdot)\big]\big(t-s,\Theta(\eta,\xi)\big)\big| \\ &\leqslant c\big|a(\xi)\big||\eta-\xi|\big|Dh(t_0,\xi_0;\cdot)\big(t-s,\Theta(\eta,\xi)\big)\big| \\ &\leqslant \frac{c|a(\xi)|}{\tilde{d}'_{\mathrm{P}}((t,\xi),(s,\eta))^{Q'-1}} \leqslant \frac{c|a(\xi)|}{\tilde{d}_{\mathrm{P}}((t,\xi),(s,\eta))^{Q'-1}}. \end{aligned}$$

This means that k_b and k_c satisfy (2.12). A similar and more tedious computation shows that (2.13) also holds with $\beta=\delta=1$. □

Theorem 6.6. *Let $T(t_0,\xi_0)$ is a frozen operator of type $\ell \geqslant 0$ and $\widetilde{B}_r$ a $\tilde{d}_{\mathrm{P}}$-ball in $\mathbb{R}^{N+1}$, then $T(t_0,\xi_0)$ is continuous on $C^\alpha(\widetilde{B}_r)$:*

$$\big\|T(t_0,\xi_0)f\big\|_{C^\alpha(\widetilde{B}_r)} \leqslant c\|f\|_{C^\alpha(\widetilde{B}_r)}.$$

Proof. We prove the theorem for frozen operators of type 0, being the other cases implicitly contained in this, by Definition 6.3. So, let $T(t_0,\xi_0)$ be as in (6.10). Throughout this proof, we will apply the results of Section 2 to the homogeneous space

$$\big(\widetilde{B}_r,\tilde{d}_{\mathrm{P}},dt\,d\xi\big).$$

This is possible in view of Proposition 3.8.

The multiplication operator

$$f \mapsto \alpha(t_0,\xi_0)\beta f$$

is obviously continuous on $C^\alpha(\widetilde{B}_r)$, by (4.3), because β is a smooth function. On the other hand, by Definition 6.3, the kernel of $T(t_0,\xi_0)$ is a finite sum of kernels of the kind

$$k_j(t,\xi,s,\eta) = a(\xi)b(\eta)\big[D_jh(t_0,\xi_0;\cdot)\big]\big(t-s,\Theta(\eta,\xi)\big)$$

with $[D_jh(t_0,\xi_0;\cdot)]$ homogeneous of some degree $j \geqslant -Q'$ (that is, D_j is homogeneous of degree $j \leqslant 2$), plus a regular kernel.

This regular part obviously satisfies (2.12) and (2.13) with $\beta = \delta = 1$ on any bounded domain, by Proposition 4.2(ii), hence defines a continuous operator on $C^{\alpha}(\widetilde{B}_r)$, by Theorem 2.11.

By Proposition 6.4, we get that:

If $j < 2$, then k_j satisfies (2.12) and (2.13) with $\beta = 1$ and $\delta = 2 - j$; the operator with kernel k_j is a fractional integral operator, continuous on $C^{\alpha}(\widetilde{B}_r)$ for any $\alpha \in (0, 1)$, by Theorem 2.11.

If $j = 2$, the kernel k_j satisfies (2.4), (2.5) with $\beta = 1$, (2.8), and (2.9), with any $\gamma < 1$; the operator with kernel k_j is a singular integral operator, continuous on $C^{\alpha}(\widetilde{B}_r)$ for any $\alpha < \gamma$, and therefore for any $\alpha \in (0, 1)$, by Theorem 2.7. □

With Theorem 6.6 at hand, we can complete the proof of Theorem 5.2 with a fairly straightforward adaptation of techniques contained in [6,7,29]. For convenience of the reader, we present a reasonably detailed proof.

As in [29, Theorem 8] (for a detailed proof see [6, Lemma 2.9]), we have:

Proposition 6.7. *If $T(t_0, \xi_0)$ is a frozen operator of type $\ell \geqslant 1$, then $\widetilde{X}_i T(t_0, \xi_0)$ is a frozen operator of type $\ell - 1$.*

Next, we recall the basic "representation formula" which holds in this context (compare with [6, Theorem 3.1]).

Theorem 6.8 *(Parametrix for $\widetilde{H}_0$). For every test function a, every t_0, ξ_0, there exist a frozen operator of type two, $P^*(t_0, \xi_0)$, and q^2 frozen operators of type one, $S_{ij}(t_0, \xi_0)$ $(i, j = 1, \dots, q)$, such that for every compactly supported function $f \in C^{2,\alpha}$,*

$$P^*(t_0, \xi_0)\widetilde{H}_0 f(t, \xi) = a(\xi) f(t, \xi) + \sum_{i,j=1}^{q} \tilde{a}_{ij}(t_0, \xi_0)\, S_{ij}(t_0, \xi_0) f(t, \xi). \tag{6.14}$$

In particular,

$$P^*(t_0, \xi_0) f(\xi) = \int_{-\infty}^{t} \int_{\mathbb{R}^N} a(\eta) b(\xi) h\big(t_0, \xi_0; t - s, \Theta(\xi, \eta)\big) f(s, \eta)\, d\eta\, ds.$$

Sketch of the proof. (See [6] for details.) 1. One considers the formally transposed operator

$$\widetilde{H}_0^* = \partial_t - \sum_{i,j=1}^{q} \tilde{a}_{ij}(t_0, \xi_0) \widetilde{X}_i^* \widetilde{X}_j^*$$

and the corresponding approximating operator

$$\mathcal{H}_0 = \partial_t - \sum_{i,j=1}^{q} \tilde{a}_{ij}(t_0, \xi_0) Y_i Y_j$$

(recall that, on the group $\mathbb{G}$, Y_i^* simply coincides with $-Y_i$, see, e.g., [29, p. 252]).

2. One defines

$$P_0(t_0,\xi_0)f(t,\xi)=\int_{-\infty}^{t}\int_{\mathbb{R}^N} a(\xi)b(\eta)h\big(t_0,\xi_0;t-s,\Theta(\eta,\xi)\big)f(s,\eta)\,d\eta\,ds,$$

where $h(t_0,\xi_0;\cdot)$ is the fundamental solution of $\mathcal{H}_0$, and a,b are suitable cutoff functions.

3. One computes $\widetilde{H}_0^* P_0(t_0,\xi_0)f(t,\xi)$ by means of relation (6.1), and finds the identity

$$\widetilde{H}_0^* P_0(t_0,\xi_0)f(t,\xi)=a(\xi)f(t,\xi)+\sum_{i,j=1}^{q}\tilde{a}_{ij}(t_0,\xi_0)\,S_{ij}^*(t_0,\xi_0)f(t,\xi),$$

where $S_{ij}^*(t_0,\xi_0)$ are frozen kernels of type 1, by Proposition 6.7.

4. One transposes the last identity and finds

$$P^*(t_0,\xi_0)\widetilde{H}_0 f(t,\xi)=a(\xi)f(t,\xi)+\sum_{i,j=1}^{q}\tilde{a}_{ij}(t_0,\xi_0)\,S_{ij}(t_0,\xi_0)f(t,\xi),$$

where the $S_{ij}(t_0,\xi_0)$'s are frozen kernels of type 1, and $P^*(t_0,\xi_0)$ is a frozen kernel of type 2; namely,

$$P^*(t_0,\xi_0)f(\xi)=\int_{0}^{t}\int_{\mathbb{R}^N} a(\eta)b(\xi)h\big(t_0,\xi_0;t-s,\Theta(\xi,\eta)\big)f(s,\eta)\,d\eta\,ds.$$

This is exactly the identity (6.14). □

Next, we want to take the second derivative $\widetilde{X}_h\widetilde{X}_k$ of both sides of (6.14), to get a representation formula for the second derivatives of a test function. To perform this computation, the following property is crucial.

Proposition 6.9. *If $S(t_0,\xi_0)$ is any frozen operator of type* 1*, there exist $q+1$ frozen operator of type* 2*, $P(t_0,\xi_0)$, $P^k(t_0,\xi_0)$ $(k=1,2,\dots,q)$ such that*

$$S(t_0,\xi_0)f(t,\xi)=\sum_{k=1}^{q}P^k(t_0,\xi_0)X_k f(t,\xi)+P(t_0,\xi_0)f(t,\xi).$$

In the stationary case, this property is contained in [29, Theorem 9, p. 292]; in this computation the presence of the time variable is irrelevant, hence the proposition holds.

Conclusion of the proof of Theorem 5.2 *(Sketch).* By Proposition 6.9, one can rewrite the parametrix formula (6.14) in the form:

$$P^*(t_0,\xi_0)\widetilde{H}_0 f(t,\xi) = a(\xi) f(t,\xi) + \sum_{i,j=1}^{q} \tilde{a}_{ij}(t_0,\xi_0)\left\{\sum_{k=1}^{q} P_{ij}^k(t_0,\xi_0)\widetilde{X}_k f(t,\xi) + P_{ij}(t_0,\xi_0) f(t,\xi)\right\},$$

where P^*, P_{ij}, P_{ij}^k are frozen operators of type two. Taking two derivatives of both sides of the previous identity, applying Proposition 6.7 and writing $\widetilde{H}_0 = \widetilde{H} + (\widetilde{H}_0 - \widetilde{H})$ we get

$$\widetilde{X}_r\widetilde{X}_s(af)(t,\xi) = T(t_0,\xi_0)\widetilde{H} f(t,\xi) + T(t_0,\xi_0)\sum_{i,j=1}^{q}\big[\tilde{a}_{ij}(t_0,\xi_0) - \tilde{a}_{ij}(t,\xi)\big]\widetilde{X}_i\widetilde{X}_j f(t,\xi) + \sum_{i,j=1}^{q} \tilde{a}_{ij}(t_0,\xi_0)\left\{\sum_{k=1}^{q} T_{ij}^k(t_0,\xi_0)\widetilde{X}_k f(t,\xi) + T_{ij}(t_0,\xi_0) f(t,\xi)\right\}, \quad (6.15)$$

where T, T_{ij}, T_{ij}^k are frozen singular integrals.

Next, we take $C^\alpha(\widetilde{B}_r)$ norm of both sides of (6.15) and apply Theorem 6.6, writing

$$\big\|\widetilde{X}_k\widetilde{X}_h f\big\|_{C^\alpha(\widetilde{B}_r)} \leqslant c\left\{\big\|\widetilde{H} f\big\|_{C^\alpha(\widetilde{B}_r)} + \sum_{i,j=1}^{q}\big\|\big[\tilde{a}_{ij}(t_0,\xi_0) - \tilde{a}_{ij}(\cdot)\big]\widetilde{X}_i\widetilde{X}_j f\big\|_{C^\alpha(\widetilde{B}_r)} + \sum_{l=1}^{q}\big\|\widetilde{X}_l f\big\|_{C^\alpha(\widetilde{B}_r)} + \|f\|_{C^\alpha(\widetilde{B}_r)}\right\}.$$

To handle the term involving $\widetilde{X}_i\widetilde{X}_j f$ in the right-hand side of the last inequality, we now exploit the fact that, for $f \in C_0^{2,\alpha}(\widetilde{B}_r)$, both $\widetilde{X}_i\widetilde{X}_j f$ and $[\tilde{a}_{ij}(t_0,\xi_0) - \tilde{a}_{ij}(\cdot)]$ vanish at a point of $\widetilde{B}_r$; then (4.4) implies

$$\sum_{i,j=1}^{q}\big|\big[\tilde{a}_{ij}(t_0,\xi_0) - \tilde{a}_{ij}(\cdot)\big]\widetilde{X}_i\widetilde{X}_j f\big|_{C^\alpha(\widetilde{B}_r)} \leqslant cr^\alpha |\tilde{a}_{ij}|_{C^\alpha(\widetilde{B}_r)}\big|\widetilde{X}_i\widetilde{X}_j f\big|_{C^\alpha(\widetilde{B}_r)},$$

while obviously

$$\sum_{i,j=1}^{q}\big\|\big[\tilde{a}_{ij}(t_0,\xi_0) - \tilde{a}_{ij}(\cdot)\big]\widetilde{X}_i\widetilde{X}_j f\big\|_{L^\infty(\widetilde{B}_r)} \leqslant cr^\alpha |\tilde{a}_{ij}|_{C^\alpha(\widetilde{B}_r)}\big\|\widetilde{X}_i\widetilde{X}_j f\big\|_{L^\infty(\widetilde{B}_r)}.$$

This allows, for r small enough, to get

$$\big\|\widetilde{X}_k\widetilde{X}_h f\big\|_{C^\alpha(\widetilde{B}_r)} \leqslant c\left\{\big\|\widetilde{H} f\big\|_{C^\alpha(\widetilde{B}_r)} + \sum_{l=1}^{q}\big\|\widetilde{X}_l f\big\|_{C^\alpha(\widetilde{B}_r)} + \|f\|_{C^\alpha(\widetilde{B}_r)}\right\} \quad (6.16)$$

(this is the classical "Korn's trick"). Since from the equation we also read

$$\|\partial_t f\|_{C^\alpha(\widetilde{B}_r)} \leqslant \|\widetilde{H} f\|_{C^\alpha(\widetilde{B}_r)} + c \sum_{h,k=1}^{q} \|\widetilde{X}_k \widetilde{X}_h f\|_{C^\alpha(\widetilde{B}_r)},$$

from (6.16) we have

$$\|f\|_{C^{2,\alpha}(\widetilde{B}_r)} \leqslant c \left\{ \|\widetilde{H} f\|_{C^\alpha(\widetilde{B}_r)} + \sum_{l=1}^{q} \|\widetilde{X}_l f\|_{C^\alpha(\widetilde{B}_r)} + \|f\|_{C^\alpha(\widetilde{B}_r)} \right\}. \tag{6.17}$$

Next, we want to get rid of the term $\|\widetilde{X}_l f\|_{C^\alpha(B_r)}$ in the last inequality. To do this, we start again with (6.14), take only one derivative $\widetilde{X}_i$ and reason like above, getting

$$\begin{aligned} \widetilde{X}_k\big[a(\xi) f(t,\xi)\big] = {} & S(t_0,\xi_0)\widetilde{H} f(t,\xi) + S(t_0,\xi_0)\left(\sum_{i,j=1}^{q} \big[\tilde{a}_{ij}(t_0,\xi_0) - \tilde{a}_{ij}(\cdot)\big]\widetilde{X}_i \widetilde{X}_j f \right)(t,\xi) \\ & + \sum_{i,j=1}^{q} \tilde{a}_{ij}(t_0,\xi_0)\, T^{ij}(t_0,\xi_0) f(t,\xi). \end{aligned}$$

In the last formula, $S(t_0,\xi_0)$, $T^{ij}(t_0,\xi_0)$ are, respectively, frozen operators of type $1, 0$. Taking C^α norms in the last equation and substituting in (6.16) we get

$$\|f\|_{C^{2,\alpha}(\widetilde{B}_r)} \leqslant c\big\{ \|\widetilde{H} f\|_{C^\alpha(\widetilde{B}_r)} + \|f\|_{C^\alpha(\widetilde{B}_r)} \big\} + \varepsilon \sum_{h,k=1}^{q} \|\widetilde{X}_k \widetilde{X}_h f\|_{C^\alpha(\widetilde{B}_r)}$$

with ε small for small r. Hence we conclude

$$\|f\|_{C^{2,\alpha}(\widetilde{B}_r)} \leqslant c\big\{ \|\widetilde{H} f\|_{C^\alpha(\widetilde{B}_r)} + \|f\|_{C^\alpha(\widetilde{B}_r)} \big\}. \tag{6.18}$$

Finally, we want to replace the term $\|f\|_{C^\alpha(\widetilde{B}_r)}$ with $\|f\|_{L^\infty(\widetilde{B}_r)}$ in the last inequality. To this aim, we apply (4.8) and write

$$\|f\|_{C^\alpha(\widetilde{B}_r)} \leqslant \|f\|_{L^\infty(\widetilde{B}_r)} + r^{1-\alpha} \left(\sum_{l=1}^{q} \|\widetilde{X}_l f\|_{L^\infty} + r\|\partial_t f\|_{L^\infty} \right).$$

Substituting this in (6.18), for r small enough the term $(\sum_{l=1}^{q} \|\widetilde{X}_l f\|_{L^\infty} + r\|\partial_t f\|_{L^\infty})$ can be taken to the left-hand side, to get

$$\|f\|_{C^{2,\alpha}(\widetilde{B}_r)} \leqslant c\big\{ \|\widetilde{H} f\|_{C^\alpha(\widetilde{B}_r)} + \|f\|_{L^\infty(\widetilde{B}_r)} \big\}$$

that is Theorem 5.2. □

7. Interpolation inequalities for Hölder norms and local Schauder estimates in the lifted variables

In order to get from Theorem 5.2 a local estimate for $C^{2,\alpha}$ functions (not necessarily with compact support), we need to establish suitable interpolation inequalities. This will require some labour; we start with the following proposition.

Proposition 7.1 *(Interpolation inequality for test functions). Let* $\mathbf{H} = \partial_t - \sum \widetilde{X}_i^2$ *and let* $\widetilde{B}_R \subset \mathbb{R}^{N+1}$ *a ball of radius* R*. Then for every* $\alpha \in (0,1)$ *there exist positive constants* $\gamma \geqslant 1$ *and* c*, depending on* α *and* $\{X_i\}$*, such that for every* $\varepsilon > 0$ *and every* $f \in C_0^\infty(\widetilde{B}_R)$

$$\left\| \widetilde{X}_i f \right\|_\alpha \leqslant \varepsilon \|\mathbf{H} f\|_\alpha + \frac{c}{\varepsilon^\gamma} \|f\|_{L^\infty}.$$

This result, in turn, relies on a similar interpolation inequality for operators of type $\ell \geqslant 1$.

Lemma 7.2. *Let* P *be an operator of type* $\ell \geqslant 1$ *and* $\alpha \in (0,1)$*. Then there exist positive constants* $\gamma > 1$ *and* c*, depending on* α *and* $\{X_i\}$*, such that for every* $\varepsilon > 0$ *and every* $f \in C_0^\infty(\widetilde{B}_R)$

$$\|P\mathbf{H} f\|_\alpha \leqslant \varepsilon \|\mathbf{H} f\|_\alpha + \frac{c}{\varepsilon^\gamma} \|f\|_{L^\infty}.$$

Proof. Let

$$P\mathbf{H} f = \int_{\widetilde{B}_R} k(t-s;\xi,\eta)\mathbf{H} f(s,\eta)\,d\eta\,ds,$$

where k satisfies the properties of a frozen kernel of type ℓ, and let ζ_ε be a cutoff function such that $\widetilde{B}_{\varepsilon/2}(t,\xi) \prec \zeta_\varepsilon \prec \widetilde{B}_\varepsilon(t,\xi)$. We split $P\mathbf{H}$ as follows

$$\begin{aligned}
P\mathbf{H} f(t,\xi) &= \int_{\widetilde{B}_R} k(t-s;\xi,\eta)\mathbf{H} f(s,\eta)\,d\eta\,ds \\
&= \int_{\tilde{d}_{\mathrm{P}}((s,\eta),(t,\xi))>\varepsilon/2} k(t-s;\xi,\eta)\bigl[1-\zeta_\varepsilon(s,\eta)\bigr]\mathbf{H} f(s,\eta)\,d\eta\,ds \\
&\quad + \int_{\tilde{d}_{\mathrm{P}}((s,\eta),(t,\xi))\leqslant\varepsilon} k(t-s;\xi,\eta)\zeta_\varepsilon(s,\eta)\mathbf{H} f(s,\eta)\,d\eta\,ds \\
&= \int_{\tilde{d}_{\mathrm{P}}((s,\eta),(t,\xi))>\varepsilon/2} \mathbf{H}^{\mathrm{T}}\bigl[k(t-\cdot;\xi,\cdot)\bigl(1-\zeta_\varepsilon(\cdot,\cdot)\bigr)\bigr](s,\eta) f(s,\eta)\,d\eta\,ds \\
&\quad + \int_{\tilde{d}_{\mathrm{P}}((s,\eta),(t,\xi))\leqslant\varepsilon} k(t-s;\xi,\eta)\zeta_\varepsilon(s,\eta)\mathbf{H} f(s,\eta)\,d\eta\,ds \\
&= I(t,\xi) + II(t,\xi),
\end{aligned}$$

where $\mathbf{H}^{\mathrm{T}}$ denote the transpose of $\mathbf{H}$.

Let $h^\varepsilon(t,\xi;s,\eta)=\mathbf{H}^{\mathrm{T}}[k(t-\cdot;\xi,\cdot)(1-\zeta_\varepsilon(\cdot,\cdot))]$ and observe that for a suitable $\gamma>1$

$$\left|h^\varepsilon(t,\xi;s,\eta)\right|+\left|\partial_t h^\varepsilon(t,\xi;s,\eta)\right|+\sum\left|\widetilde{X}_j^\xi h^\varepsilon(t,\xi;s,\eta)\right|\leqslant c\varepsilon^{-\gamma}.$$

This follows from (6.12), by the definition of h^ε.

By Proposition 4.2(ii), it follows that

$$\left|h^\varepsilon(t_1,\xi_1;s,\eta)-h^\varepsilon(t_2,\xi_2;s,\eta)\right|\leqslant c_R\varepsilon^{-\gamma}\tilde{d}_{\mathrm{P}}\big((t_1,\xi_1),(t_2,\xi_2)\big)$$

and therefore

$$\begin{aligned}\left|I(t_1,\xi_1)-I(t_2,\xi_2)\right|&\leqslant\int\left|h^\varepsilon(t_1,\xi_1;s,\eta)-h^\varepsilon(t_2,\xi_2;s,\eta)\right|\left|f(s,\eta)\right|d\eta\,ds\\&\leqslant c\varepsilon^{-\gamma}\left|\widetilde{B}_R\right|\|f\|_{L^\infty(\widetilde{B}_R)}\tilde{d}_{\mathrm{P}}\big((t_1,\xi_1),(t_2,\xi_2)\big).\end{aligned}$$

Also, since

$$\left|I(t,\xi)\right|\leqslant\int_{\tilde{d}_{\mathrm{P}}((s,\eta),(t,\xi))>\varepsilon/2}c\varepsilon^{-\gamma}\left|f(s,\eta)\right|d\eta\,ds\leqslant c\varepsilon^{-\gamma}\left|\widetilde{B}_R\right|\|f\|_{L^\infty(\widetilde{B}_R)}$$

we obtain

$$\|I\|_\alpha\leqslant c\varepsilon^{-\gamma}\|f\|_{L^\infty}\quad\text{for any }\alpha\in(0,1).$$

Let us consider $II(t,\xi)$, and let

$$k_\varepsilon(t,\xi,s,\eta)=k(t-s;\xi,\eta)\zeta_\varepsilon(s,\eta).$$

By Proposition 6.4, and keeping into account the support of k_ε, for any fixed $\delta\in(0,1)$, the kernel satisfies properties (2.12), (2.13) in the form:

$$\begin{aligned}\left|k_\varepsilon(t,\xi,s,\eta)\right|&\leqslant c\frac{\tilde{d}_{\mathrm{P}}((t,\xi),(s,\eta))}{|\widetilde{B}((t,\xi);(s,\eta))|}\leqslant c\varepsilon^\delta\frac{\tilde{d}_{\mathrm{P}}((t,\xi),(s,\eta))^{1-\delta}}{|\widetilde{B}((t,\xi);(s,\eta))|};\\\left|k_\varepsilon(t,\xi,s,\eta)-k_\varepsilon(t_1,\xi_1,s,\eta)\right|&\leqslant c\frac{\tilde{d}_{\mathrm{P}}((t,\xi),(t_1,\xi_1))}{|\widetilde{B}((t_1,\xi_1);(s,\eta))|}\left(\frac{\tilde{d}_{\mathrm{P}}((t,\xi),(t_1,\xi_1))}{\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(s,\eta))}\right)\\&\leqslant c\varepsilon^\delta\frac{\tilde{d}_{\mathrm{P}}((t,\xi),(t_1,\xi_1))^{1-\delta}}{|\widetilde{B}((t_1,\xi_1);(s,\eta))|}\left(\frac{\tilde{d}_{\mathrm{P}}((t,\xi),(t_1,\xi_1))}{\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(s,\eta))}\right)\end{aligned}$$

for $\tilde{d}_{\mathrm{P}}((t_1,\xi_1),(s,\eta))\geqslant 2\tilde{d}_{\mathrm{P}}((t,\xi),(t_1,\xi_1))$. By Theorem 2.11, this implies

$$\|II\|_\alpha\leqslant c\varepsilon^\delta\|\mathbf{H}f\|_\alpha$$

for every $\alpha<1-\delta$. Therefore, for every $\alpha\in(0,1)$ there exist $\delta,\gamma>0$ such that

$$\|P\mathbf{H}f\|_\alpha\leqslant c\varepsilon^\delta\|\mathbf{H}f\|_\alpha+c\varepsilon^{-\gamma}\|f\|_{L^\infty(\widetilde{B}_R)},$$

which implies the lemma. □

Proof of Proposition 7.1. Let $\{a_{ij}\}$ be the identity matrix. By Theorem 6.8, we can write

$$P\mathbf{H}f(t,\xi)=a(\xi)f(t,\xi)+Sf(t,\xi),$$

where P is an operator of type 2 and S is an operator of type 1. If we assume $a\equiv 1$ on $\widetilde{B}_R$, for $f\in C_0^\infty(\widetilde{B}_R)$ we obtain

$$f=P\mathbf{H}f-Sf \tag{7.1}$$

and therefore, by Proposition 6.7

$$\widetilde{X}_i f=S_1\mathbf{H}f+Tf, \tag{7.2}$$

where S_1 is an operator of type 1 and T is an operator of type 0. Substituting (7.1) in (7.2) yields

$$\widetilde{X}_i f=S_1\mathbf{H}f+TP\mathbf{H}f-TSf$$

and therefore

$$\begin{aligned}\|\widetilde{X}_i f\|_\alpha &\leqslant \|S_1\mathbf{H}f\|_\alpha+\|TP\mathbf{H}f\|_\alpha+\|TSf\|_\alpha \\ &\leqslant c\big\{\|S_1\mathbf{H}f\|_\alpha+\|P\mathbf{H}f\|_\alpha+\|Sf\|_\alpha\big\} \quad \text{(by Theorem 6.6)}, \\ &\leqslant c\big\{\varepsilon\|\mathbf{H}f\|_\alpha+\varepsilon^{-\gamma}\|f\|_{L^\infty}+\|Sf\|_\alpha\big\} \quad \text{(applying Lemma 7.2 to } S_1 \text{ and } P).\end{aligned} \tag{7.3}$$

We end the proof by showing that

$$\|Sf\|_\alpha\leqslant c\|f\|_{L^\infty}.$$

Indeed, if

$$Sf(t,\xi)=\int_{\widetilde{B}_R}k(t,\xi;s,\eta)f(s,\eta)\,d\eta\,ds$$

we have

$$\begin{aligned}\big|Sf(t_1,\xi_1)-Sf(t_2,\xi_2)\big| &= \left|\int\big[k(t_1,\xi_1;s,\eta)-k(t_2,\xi_2;s,\eta)\big]f(s,\eta)\,ds\,d\eta\right| \\ &\leqslant \int\big|k(t_1,\xi_1;s,\eta)-k(t_2,\xi_2;s,\eta)\big|\big|f(y)\big|\,ds\,d\eta \\ &\leqslant \|f\|_{L^\infty(\widetilde{B}_R)}\int\big|k(t_1,\xi_1;s,\eta)-k(t_2,\xi_2;s,\eta)\big|\,ds\,d\eta.\end{aligned}$$

Arguing as in last part of the proof of Theorem 2.11 (with $\beta=\delta=1$), we obtain that for every $\alpha\in(0,1)$

$$\int\big|k(t_1,\xi_1;s,\eta)-k(t_2,\xi_2;s,\eta)\big|\,ds\,d\eta\leqslant c_\alpha\tilde{d}_{\mathrm{P}}\big((t_1,\xi_1),(t_2,\xi_2)\big)^\alpha R^{1-\alpha}.$$

This shows that

$$|Sf|_\alpha \leqslant c\|f\|_{L^\infty}.$$

Moreover,

$$\begin{aligned}\left|Sf(t,\xi)\right| &\leqslant \int_{\widetilde{B}_R} \left|k(t,\xi;s,\eta)f(s,\eta)\right| d\eta\, ds \\ &\leqslant \|f\|_{L^\infty} \int_{\tilde{d}_P((t,\xi),(s,\eta))\leqslant cR} c\frac{\tilde{d}_P((t,\xi),(s,\eta))}{|\widetilde{B}((t,\xi),(s,\eta))|}\, ds\, d\eta \leqslant cR\|f\|_{L^\infty}\end{aligned}$$

by Lemma 2.10. Hence

$$\|Sf\|_\alpha \leqslant c\|f\|_{L^\infty}. \qquad \square$$

We can now follow the technique used in [7], to prove a version of the previous theorem for functions which do not vanish at the boundary of the domain. Some complication will arise to handle extra terms involving the time derivative.

The following technical lemma is adapted from [13, Lemma 4.1, p. 27], and is proved in this form in [7].

Lemma 7.3. *Let $\psi(t)$ be a bounded nonnegative function defined on the interval $[T_0, T_1]$, where $T_1 > T_0 \geqslant 0$. Suppose that for any $T_0 \leqslant t < s \leqslant T_1$, ψ satisfies*

$$\psi(t) \leqslant \vartheta\psi(s) + \frac{A}{(s-t)^\beta} + B,$$

where ϑ, A, B, β are nonnegative constants, and $\vartheta < \frac{1}{3}$. Then

$$\psi(\rho) \leqslant c_\beta\left[\frac{A}{(R-\rho)^\beta} + B\right], \quad \forall\rho,\ T_0 \leqslant \rho < R \leqslant T_1,$$

where c_β only depends on β.

Theorem 7.4 *(Interpolation inequality). There exist positive constants c, R, γ such that for any $f \in C^{2,\alpha}(\widetilde{B}_R)$, $0 < \rho < R$, $0 < \delta < 1/3$,*

$$\|Df\|_{C^\alpha(\widetilde{B}_\rho)} \leqslant \delta\left[\left\|D^2 f\right\|_{C^\alpha(\widetilde{B}_R)} + \|\partial_t f\|_{C^\alpha(\widetilde{B}_R)}\right] + \frac{c}{\delta^\gamma(R-\rho)^{2\gamma}}\|f\|_{L^\infty(\widetilde{B}_R)}.$$

The constants c, R, γ depend on α, $\{X_i\}$; γ is as in Proposition 7.1. (Recall Notation 6.1 for the use of symbols D, D^2.)

Proof. If $f \in C^{2,\alpha}(\widetilde{B}_R)$, $0 < t < s \leqslant R$ and ζ is a cutoff function with $\widetilde{B}_t \prec \zeta \prec \widetilde{B}_s$, applying Proposition 7.1 to $f\zeta$ and using Lemma 6.2, we get

$$\|Df\|_{C^\alpha(\widetilde{B}_t)} \leqslant \|D(\zeta f)\|_{C^\alpha(\widetilde{B}_s)} \leqslant \varepsilon \|\mathbf{H}(\zeta f)\|_{C^\alpha(\widetilde{B}_s)} + \frac{c}{\varepsilon^\gamma}\|f\|_{L^\infty(\widetilde{B}_s)}, \tag{7.4}$$

where

$$\begin{aligned} \|\mathbf{H}(\zeta f)\|_{C^\alpha(\widetilde{B}_s)} &\leqslant \|\zeta \mathbf{H} f\|_{C^\alpha(\widetilde{B}_s)} + c\|D\zeta\, Df\|_{C^\alpha(\widetilde{B}_s)} + \|f\mathbf{H}\zeta\|_{C^\alpha(\widetilde{B}_s)} \\ &\leqslant \frac{c}{s-t}\|\mathbf{H}f\|_{C^\alpha(\widetilde{B}_s)} + \frac{c}{(s-t)^2}\|Df\|_{C^\alpha(\widetilde{B}_s)} \\ &\quad + \frac{c}{(s-t)^2}\|f\|_{L^\infty(\widetilde{B}_s)} + |f\mathbf{H}\zeta|_{C^\alpha(\widetilde{B}_s)}. \end{aligned} \tag{7.5}$$

To bound the last term in the last inequality, we apply (4.8), and write

$$\begin{aligned} |f\mathbf{H}\zeta|_{C^\alpha(\widetilde{B}_s)} &\leqslant R^{1-\alpha}\left(\|D(f\mathbf{H}\zeta)\|_{L^\infty(\widetilde{B}_s)} + R\|\partial_t(f\mathbf{H}\zeta)\|_{L^\infty(\widetilde{B}_s)}\right) \\ &\leqslant R^{1-\alpha}\left\{\frac{c}{(s-t)^2}\|Df\|_{L^\infty(\widetilde{B}_s)} + \frac{c}{(s-t)^3}\|f\|_{L^\infty(\widetilde{B}_s)}\right\} \\ &\quad + R^{2-\alpha}\left\{\frac{c}{(s-t)^2}\|\partial_t f\|_{L^\infty(\widetilde{B}_s)} + \frac{c}{(s-t)^4}\|f\|_{L^\infty(\widetilde{B}_s)}\right\}. \end{aligned}$$

This bound inserted in (7.5) gives

$$\begin{aligned} \|\mathbf{H}(\zeta f)\|_{C^\alpha(\widetilde{B}_s)} &\leqslant \frac{c}{(s-t)^2}\left\{\|D^2 f\|_{C^\alpha(\widetilde{B}_s)} + \|\partial_t f\|_{C^\alpha(\widetilde{B}_s)}\right\} \\ &\quad + \frac{c_1}{(s-t)^2}\|Df\|_{C^\alpha(\widetilde{B}_s)} + \frac{c}{(s-t)^4}\|f\|_{L^\infty(\widetilde{B}_s)}, \end{aligned}$$

where now all the constants c depend also on R. Next, we insert the last inequality in (7.4), choosing $\varepsilon = \delta(s-t)^2/c_1$ and get

$$\begin{aligned} \|Df\|_{C^\alpha(\widetilde{B}_t)} &\leqslant \delta\|Df\|_{C^\alpha(\widetilde{B}_s)} + c\delta\left\{\|D^2 f\|_{C^\alpha(\widetilde{B}_s)} + \|\partial_t f\|_{C^\alpha(\widetilde{B}_s)}\right\} \\ &\quad + \left(\frac{c\delta}{(s-t)^2} + \frac{c}{\delta^\gamma (s-t)^{2\gamma}}\right)\|f\|_{L^\infty(\widetilde{B}_s)}. \end{aligned}$$

Let $\psi(t) = \|Df\|_{L^\infty(\widetilde{B}_t)}$. Then, for R fixed once and for all (small enough), $\vartheta < 1/3$ fixed, and any $\delta < \vartheta$ we get, since $\gamma > 1$,

$$\psi(t) \leqslant \vartheta\psi(s) + \frac{c}{\delta^\gamma (s-t)^{2\gamma}}\|f\|_{L^\infty(\widetilde{B}_R)} + c\delta\left[\|D^2 f\|_{C^\alpha(\widetilde{B}_R)} + \|\partial_t f\|_{C^\alpha(\widetilde{B}_R)}\right]$$

for any $0 < t < s < R$, and by Lemma 7.3 we get

$$\psi(\rho) \leqslant \frac{c}{\delta^\gamma (R-\rho)^{2\gamma}}\|f\|_{L^\infty(\widetilde{B}_R)} + c\delta\left[\|D^2 f\|_{C^\alpha(\widetilde{B}_R)} + \|\partial_t f\|_{C^\alpha(\widetilde{B}_R)}\right]$$

for any $0 < \rho < R$. $\square$

We now come to the goal of this section.

Proof of Theorem 5.3. If $f \in C^{2,\alpha}(\widetilde{B}_R)$ (R small enough to apply Theorem 5.2), $t < R$, $s = (t+R)/2$, and ζ is a cutoff function, $\widetilde{B}_t \prec \zeta \prec \widetilde{B}_s$, we can apply Theorem 5.2 to $f\zeta$, getting

$$\begin{aligned}\|f\|_{C^{2,\alpha}(\widetilde{B}_t)} &\leqslant c\left\{\left\|\widetilde{H}(f\zeta)\right\|_{C^\alpha(\widetilde{B}_s)} + \|f\zeta\|_{L^\infty(\widetilde{B}_s)}\right\} \\ &\leqslant c\left\{\frac{1}{s-t}\left\|\widetilde{H}f\right\|_{C^\alpha(\widetilde{B}_s)} + \frac{1}{(s-t)^2}\|Df\|_{C^\alpha(\widetilde{B}_s)}\right. \\ &\quad \left. + \frac{1}{(s-t)^2}\|f\|_{L^\infty(\widetilde{B}_s)} + \left|f\widetilde{H}\zeta\right|_{C^\alpha(\widetilde{B}_s)}\right\} + \|f\zeta\|_{L^\infty(\widetilde{B}_s)} \end{aligned} \tag{7.6}$$

by computations similar to those already done in the proof of Theorem 7.4.

Now, however, we have to handle the term $|f\widetilde{H}\zeta|_{C^\alpha(\widetilde{B}_s)}$ in a different way. Applying (4.8) and (4.11) we can write, for some small η to be chosen later,

$$\begin{aligned}\left|f\widetilde{H}\zeta\right|_{C^\alpha(\widetilde{B}_s)} &\leqslant R^{1-\alpha}\left(\left\|D\left(f\widetilde{H}\zeta\right)\right\|_{L^\infty(\widetilde{B}_s)} + R\left\|\partial_t\left(f\widetilde{H}\zeta\right)\right\|_{L^\infty(\widetilde{B}_s)}\right) \\ &\leqslant R^{1-\alpha}\left\{\frac{c}{(s-t)^2}\|Df\|_{L^\infty(\widetilde{B}_s)} + \frac{c}{(s-t)^3}\|f\|_{L^\infty(\widetilde{B}_s)}\right\} \\ &\quad + R^{2-\alpha}\left\{\eta^{\alpha/2}\left|\partial_t\left(f\widetilde{H}\zeta\right)\right|_{C^\alpha(\widetilde{B}_s)} + \frac{2}{\eta}\left\|f\widetilde{H}\zeta\right\|_{L^\infty(\widetilde{B}_s)}\right\}. \end{aligned} \tag{7.7}$$

Now,

$$R^{2-\alpha}\frac{2}{\eta}\left\|f\widetilde{H}\zeta\right\|_{L^\infty(\widetilde{B}_s)} \leqslant \frac{cR^{2-\alpha}}{\eta(s-t)^2}\|f\|_{L^\infty(\widetilde{B}_s)}; \tag{7.8}$$

$$\left|\partial_t\left(f\widetilde{H}\zeta\right)\right|_{C^\alpha(\widetilde{B}_s)} \leqslant \left|(\partial_t f)\widetilde{H}\zeta\right|_{C^\alpha(\widetilde{B}_s)} + \left|f\partial_t\widetilde{H}\zeta\right|_{C^\alpha(\widetilde{B}_s)}; \tag{7.9}$$

$$\left|(\partial_t f)\widetilde{H}\zeta\right|_{C^\alpha(\widetilde{B}_s)} \leqslant \frac{c}{(s-t)^3}\|\partial_t f\|_{C^\alpha(\widetilde{B}_s)}; \tag{7.10}$$

$$\begin{aligned}\left|f\partial_t\widetilde{H}\zeta\right|_{C^\alpha(\widetilde{B}_s)} &\leqslant R^{1-\alpha}\left(\left\|D\left(f\partial_t\widetilde{H}\zeta\right)\right\|_{L^\infty(\widetilde{B}_s)} + R\left\|\partial_t\left(f\partial_t\widetilde{H}\zeta\right)\right\|_{L^\infty(\widetilde{B}_s)}\right) \\ &\leqslant R^{1-\alpha}\left\{\frac{c}{(s-t)^4}\|Df\|_{L^\infty(\widetilde{B}_s)} + \frac{c}{(s-t)^5}\|f\|_{L^\infty(\widetilde{B}_s)}\right\} \\ &\quad + R^{2-\alpha}\left\{\frac{c}{(s-t)^4}\|\partial_t f\|_{L^\infty(\widetilde{B}_s)} + \frac{c}{(s-t)^6}\|f\|_{L^\infty(\widetilde{B}_s)}\right\}. \end{aligned} \tag{7.11}$$

Inserting (7.8)–(7.11) in (7.7) and then in (7.6) we get

$$\begin{aligned}&\|f\|_{C^{2,\alpha}(\widetilde{B}_t)} \\ &\quad \leqslant c_1\left\{\frac{1}{s-t}\left\|\widetilde{H}f\right\|_{C^\alpha(\widetilde{B}_s)} + \frac{1}{(s-t)^2}\|Df\|_{C^\alpha(\widetilde{B}_s)} + \left(\frac{1}{(s-t)^3} + \frac{1}{\eta(s-t)^2}\right)\|f\|_{L^\infty(\widetilde{B}_s)}\right\} \\ &\qquad + c_2\eta^{\alpha/2}\left\{\frac{1}{(s-t)^4}\|\partial_t f\|_{C^\alpha(\widetilde{B}_s)} + \frac{1}{(s-t)^4}\|Df\|_{C^\alpha(\widetilde{B}_s)} + \frac{1}{(s-t)^6}\|f\|_{L^\infty(\widetilde{B}_s)}\right\}, \end{aligned} \tag{7.12}$$

where now the constants c_i depend also on R.

For a small ε to be chosen later, we not pick η such that $c_2\eta^{\alpha/2}/(s-t)^4=\varepsilon$, and write

$$\leqslant \|f\|_{C^{2,\alpha}(\widetilde{B}_t)} \leqslant c_1\left\{\frac{1}{s-t}\big\|\widetilde{H}f\big\|_{C^\alpha(\widetilde{B}_R)}+\frac{1}{(s-t)^2}\|Df\|_{C^\alpha(\widetilde{B}_s)}\right\} + \varepsilon\|f\|_{C^{2,\alpha}(\widetilde{B}_R)}+\frac{c}{\varepsilon^{\alpha/2}(s-t)^{\beta'}}\|f\|_{L^\infty(\widetilde{B}_R)}. \tag{7.13}$$

Next, we apply Theorem 7.4 to write

$$\|Df\|_{C^\alpha(\widetilde{B}_s)} \leqslant \delta\|f\|_{C^{2,\alpha}(\widetilde{B}_R)}+\frac{c}{\delta^\gamma(s-t)^\beta}\|f\|_{L^\infty(\widetilde{B}_R)} \tag{7.14}$$

(recall that, by our choice of t,s,R, $(s-t)=(R-s)$). We insert (7.14) in (7.13) with δ such that $c_1\delta/(s-t)^2=\varepsilon$, and get

$$\|f\|_{C^{2,\alpha}(\widetilde{B}_t)} \leqslant \frac{c_1}{R-t}\big\|\widetilde{H}f\big\|_{C^\alpha(\widetilde{B}_R)}+2\varepsilon\|f\|_{C^{2,\alpha}(\widetilde{B}_R)}+\frac{c}{\varepsilon^{\alpha'}(R-t)^{\beta''}}\|f\|_{L^\infty(\widetilde{B}_R)}.$$

Letting $\psi(t)=\|f\|_{C^{2,\alpha}(\widetilde{B}_t)}$ and choosing ε such that $2\varepsilon=\vartheta<1/3$, we can rewrite the last inequality as

$$\psi(t)\leqslant\vartheta\psi(R)+\frac{c}{(R-t)^{\beta''}}\left(\big\|\widetilde{H}f\big\|_{C^\alpha(\widetilde{B}_R)}+\|f\|_{L^\infty(\widetilde{B}_R)}\right)$$

and by Lemma 7.3 we get

$$\|f\|_{C^{2,\alpha}(\widetilde{B}_t)} \leqslant \frac{c}{(R-t)^{\beta''}}\left(\big\|\widetilde{H}f\big\|_{C^\alpha(\widetilde{B}_R)}+\|f\|_{L^\infty(\widetilde{B}_R)}\right) \tag{7.15}$$

for R small enough, with c depending also on an upper bound for R. □

8. Hölder spaces and lifting

To show how Theorem 5.3 implies Theorem 5.4 and then Theorem 5.1, we need some facts about the metric induced by vector fields. Let $d(x,y)$ the CC-distance induced by a system $X_1,X_2,\dots,X_q$ of Hörmander's vector fields in $\mathbb{R}^n$, and let $\tilde{d}(\xi,\eta)$ be the CC-distance induced by the lifted vector fields $\widetilde{X}_1,\widetilde{X}_2,\dots,\widetilde{X}_q$ in $\mathbb{R}^N$; also, let d_P, $\tilde{d}_P$ be the corresponding parabolic distances. For a bounded domain $U\subset\mathbb{R}^{n+1}$, let $\widetilde{U}=U\times I\subset\mathbb{R}^{N+1}$ be the lifted counterpart of U, where I is some neighborhood of the origin in $\mathbb{R}^{N-n}$. Denote by $C_X^\alpha(U)$, $C_{\widetilde{X}}^\alpha(\widetilde{U})$ the Hölder spaces induced by d_P and $\tilde{d}_P$, respectively. We are interested in the following question: if, for any $f:U\to\mathbb{R}$ we set $\tilde{f}:\widetilde{U}\to\mathbb{R}$ with $\tilde{f}(t,x,h)=f(t,x)$, can we say that $f\in C_X^\alpha(U)$ if and only if $\tilde{f}\in C_{\widetilde{X}}^\alpha(\widetilde{U})$? By [30, Lemma 7, p. 153], we know that

$$\tilde{d}\big((x,h),(y,k)\big)\geqslant d(x,y).$$

This obviously implies

$$\tilde{d}_P\big((t,x,h),(s,y,k)\big)\geqslant d_P\big((t,x),(s,y)\big)$$

and therefore

$$\left|\tilde f\right|_{C^{\alpha}_{\widetilde X}(\widetilde U)} \leqslant |f|_{C^{\alpha}(U)}. \tag{8.1}$$

However, no simple inequality of the kind

$$\tilde d\big((x,h),(y,h)\big) \leqslant c d(x,y)$$

seems to be known, so an inequality of the kind

$$|f|_{C^{\alpha}_X(U)} \leqslant c\left|\tilde f\right|_{C^{\alpha}_{\widetilde X}(\widetilde U)} \tag{8.2}$$

is not trivial and, as far we know, has never been proved. (Note that in [33] the lifting technique was avoided, making a stronger assumption on the algebra of the vector fields.)

We are going to prove (8.2) here. The point is to make use of an integral formulation of Hölder continuity.

Let

$$M_{\alpha,B_R(t_0,x_0)}(f) = \sup_{(t,x)\in B_R,\ r>0}\ \inf_{c\in\mathbb{R}} \frac{1}{r^{\alpha}|B_r(t,x)|} \int\limits_{B_r(t,x)\cap B_R(t_0,x_0)} \left|f(s,y)-c\right| ds\,dy.$$

If $f \in C^{\alpha}_X(B_R(t_0,x_0))$, then $M_{\alpha,B_R(t_0,x_0)}(f) \leqslant c|f|_{C^{\alpha}_X(B_R(t_0,x_0))}$. But the converse is also true.

Lemma 8.1. *If $f \in L^1_{\mathrm{loc}}(B_R(t_0,x_0))$ is a function such that $M_{\alpha,B_R(t_0,x_0)}(f) < \infty$, then there exists a function f^*, a.e. equal to f, such that $f^* \in C^{\alpha}_X(B_R(t_0,x_0))$ and*

$$|f^*|_{C^{\alpha}_X(B_R)} \leqslant c M_{\alpha,B_R(t_0,x_0)}(f)$$

for some c independent of f.

Proof. This result, in the Euclidean case, is due to Campanato (see [11, Theorem I.2, p. 183], see also [25,31] for related results). Reading [11, pp. 177–184], one can see that exactly the same proof holds in a much more general context, namely:

If (X,d,μ) is a space of homogeneous type, and $\Omega \subset X$ is a bounded domain in X, d-regular in the sense of Definition 3.2, then there exists $r>0$ such that

$$\left|f^*(x)-f^*(y)\right| \leqslant c M_{\alpha,\Omega}(f) d(x,y)^{\alpha}$$

for a suitable function $f^* = f$ a.e., and any couple of points $x, y \in \Omega$ with $d(x,y) < r$ (with the obvious meaning of symbols).

We can apply the above statement to $\Omega = B_R(t_0,x_0)$ and $d_{\mathrm P}$ the parabolic CC-distance induced by the vector fields X_i, in view of Proposition 3.8, getting

$$\left|f^*(t,x)-f^*(s,y)\right| \leqslant c_{\alpha} M_{\alpha,B_R(t_0,x_0)}(f) d_{\mathrm P}\big((t,x),(s,y)\big)^{\alpha}$$

for $d_{\mathrm P}((t,x),(s,y)) < r$. It is then easy to extend this bound to any couple of points (t,x), $(s,y) \in B_R(t_0,x_0)$ such that $d_{\mathrm P}((t,x),(s,y)) \geqslant r$. To this aim, it is enough to show that we can choose k points (t_i, x_i), $i = 1, 2, \dots, k$ such that:

(i) $(t_i, x_i) \in B_R(t_0, x_0)$ for $i = 1, 2, \dots, k$, $(t_1, x_1) = (t, x)$ and $(t_k, x_k) = (s, y)$;
(ii) $d_P((t_{i-1}, x_{i-1}), (t_i, x_i)) \leqslant r$ for $i = 2, 3, \dots, k$;
(iii) the integer k and the constant c can be chosen dependently only on r, R.

Once this is done, we can write

$$\begin{aligned}
\left|f^*(t,x) - f^*(s,y)\right| &\leqslant \sum_{i=2}^{k} \left|f^*(t_i, x_i) - f^*(t_{i-1}, x_{i-1})\right| \\
&\leqslant \sum_{i=2}^{k} c M_{\alpha, B_R(t_0,x_0)}(f) d_P\big((t_{i-1}, x_{i-1}), (t_i, x_i)\big)^\alpha \\
&\leqslant kc M_{\alpha, B_R(t_0,x_0)}(f) d_P\big((t,x),(s,y)\big)^\alpha \\
&= c(r, R, \alpha) d_P\big((t,x),(s,y)\big)^\alpha.
\end{aligned}$$

So, let us show how to choose these points. To fix ideas, assume $|t - t_0| \geqslant |s - t_0|$. Recalling that

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

we can join (t, x) with (s, y) along the following line: first we move from (t, x) to (s, x) along a segment; this segment is contained in $B_R(t_0, x_0)$ and can be divided in k_1 equal parts, each of (Euclidean) length $\leqslant r^2$, with k_1 only depending on r, R; then we consider the two points $x, y \in B_{\sqrt{R^2 - |s-t_0|}}(x_0)$; by definition of CC-distance, we can join x to x_0 and y to x_0 with two subunit curves γ_1, γ_2 contained in $B_{\sqrt{R^2 - |s-t_0|}}(x_0)$, with

$$T(\gamma_i) \leqslant R \leqslant \frac{R}{r} d_P\big((t,x),(s,y)\big).$$

Therefore we can also join (s, x) to (t, y) with a line γ contained in $B_R(t_0, x_0)$, with

$$T(\gamma) \leqslant c d_P\big((t,x),(s,y)\big);$$

on this line we can choose k_2 points such that the distance of two subsequent points is $\leqslant r$; moreover, the number k_2 only depends on r, R. This ends the proof. □

The second fact we use is the following property.

Lemma 8.2. *There exist $c > 0$ and $\delta \in (0, 1)$ such that for any positive function g defined in $\widetilde{U} \subset \mathbb{R}^{N+1}$, $(t, x, h) \in \widetilde{U}$, $r > 0$, r small enough,*

$$\frac{1}{|B_{\delta r}(t,x) \cap B_R(t_0,x_0)|} \int_{B_{\delta r}(t,x) \cap B_R(t_0,x_0)} g(s,y)\,ds\,dy \leqslant \frac{c}{|\widetilde{B}_r(t,x,h)|} \int_{\widetilde{B}_r(t,x,h)} g(s,y)\,ds\,dy\,dh'.$$

Proof. By [30, Theorem 4, p. 151] (quoted in Appendix A) we know that, given a point $(x,h)\in\mathbb{R}^N$,

$$\big|\widetilde{B}_r(x,h)\big| \simeq \big|B_r(x)\big|\cdot\big|\big\{h'\in\mathbb{R}^{N-n}\colon (z,h')\in\widetilde{B}_r(x,h)\big\}\big|$$

provided $z\in B_{\delta r}(x)$ for some fixed $\delta<1$. The equivalence holds with respect to $r>0$, and the symbol $|\cdot|$ denotes the volume of a set in the suitable dimension. Multiplying both sides by r^2 we see that this property extends to the parabolic version: given a point $(t,x,h)\in\mathbb{R}^{N+1}$,

$$\big|\widetilde{B}_r(t,x,h)\big| \simeq \big|B_r(t,x)\big|\cdot\big|\big\{h'\in\mathbb{R}^{N-n}\colon (\tau,z,h')\in\widetilde{B}_r(t,x,h)\big\}\big|$$

provided $(\tau,z)\in B_{\delta r}(t,x)$ for some fixed $\delta<1$. Exploiting this fact, for any positive function $g(s,y)$ defined in $\mathbb{R}^{n+1}$ we can write

$$\begin{aligned}
&\frac{1}{|\widetilde{B}_r(t,x,h)|}\int\limits_{\widetilde{B}_r(t,x,h)} g(s,y)\,ds\,dy\,dh' \\
&\quad=\frac{1}{|\widetilde{B}_r(t,x,h)|}\int\limits_{B_r(t,x)} g(s,y)\,ds\,dy \int\limits_{\{h'\in\mathbb{R}^{N-n}\colon (s,y,h')\in\widetilde{B}_r(t,x,h)\}} dh' \\
&\quad\geqslant\frac{c}{|\widetilde{B}_r(t,x,h)|}\int\limits_{B_{\delta r}(t,x)} \frac{|\widetilde{B}_r(t,x,h)|}{|B_r(t,x)|} g(s,y)\,ds\,dy \\
&\quad=\frac{c}{|B_r(t,x)|}\int\limits_{B_{\delta r}(t,x)} g(s,y)\,ds\,dy \geqslant \frac{c}{|B_{\delta r}(t,x)|}\int\limits_{B_{\delta r}(t,x)} g(s,y)\,ds\,dy \\
&\quad\geqslant\frac{c}{|B_{\delta r}(t,x)\cap B_R(t_0,x_0)|}\int\limits_{B_{\delta r}(t,x)\cap B_R(t_0,x_0)} g(s,y)\,ds\,dy,
\end{aligned}$$

where the last inequality holds by Proposition 3.8. □

The above lemma enables us to state the following proposition.

Proposition 8.3. *If f, $\tilde f$ are as above, then*

$$\big|\tilde f\big|_{C^\alpha_{\widetilde X}(\widetilde B_R)} \leqslant |f|_{C^\alpha_X(B_R)} \leqslant c\big|\tilde f\big|_{C^\alpha_{\widetilde X}(\widetilde B_R)}. \tag{8.3}$$

Moreover,

$$\big|\widetilde X_{i_1}\widetilde X_{i_2}\cdots\widetilde X_{i_k}\tilde f\big|_{C^\alpha_{\widetilde X}(\widetilde B_R)} \leqslant |X_{i_1}X_{i_2}\cdots X_{i_k}f|_{C^\alpha_X(B_R)} \leqslant c\big|\widetilde X_{i_1}\widetilde X_{i_2}\cdots\widetilde X_{i_k}\tilde f\big|_{C^\alpha_{\widetilde X}(\widetilde B_R)} \tag{8.4}$$

for $i_j=1,2,\dots,q$.

Proof. For any $k \in \mathbb{R}$, we have the following inequalities:

$$\begin{aligned}
&\frac{1}{r^\alpha}\frac{1}{|B_r(t,x)\cap B_R(t_0,x_0)|}\int\limits_{B_r(t,x)\cap B_R(t_0,x_0)}\big|f(s,y)-k\big|\,ds\,dy\\
&\quad\leqslant \frac{c}{r^\alpha}\frac{1}{|\widetilde{B}_{r/\delta}(t,x,h')|}\int\limits_{\widetilde{B}_{r/\delta}(t,x,h')}\big|\tilde{f}(s,y,h)-k\big|\,ds\,dy\,dh \quad \text{(by Lemma 8.2)}\\
&\quad\leqslant \frac{c}{r^\alpha}\big|\tilde{f}\big|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_R)}(r/\delta)^\alpha = c\big|\tilde{f}\big|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_R)} \quad \text{(choosing } k = f(t,x) = \tilde{f}(t,x,h')\text{)}.
\end{aligned}$$

Taking the sup on $r > 0$ and $(t,x) \in B_R(t_0,x_0)$, we get

$$M_{\alpha,B_R(t_0,x_0)}(f) \leqslant c\big|\tilde{f}\big|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_R)}$$

and, by Lemma 8.1, the second inequality in (8.3) follows (while the first is trivial).

Now, inequality (8.4) is also a consequence of what we have proved, just because $\widetilde{X}_i\tilde{f} = \widetilde{(X_i f)}$. To justify this assertion, it is enough to recall the structure of the lifted vector fields $\widetilde{X}_i$:

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

Since $\tilde{f}$ does not depend on the added variables h_j, $\widetilde{X}_i\tilde{f} = X_i\tilde{f} = \widetilde{(X_i f)}$. The same reasoning can be iterated to higher order derivatives. □

Combining Theorem 5.3 with Proposition 8.3, we immediately get Theorem 5.4.

Proof of Theorem 5.4.

$$\begin{aligned}
\|u\|_{C^{2,\alpha}_X(B_t)} &\leqslant c\|\tilde{u}\|_{C^{2,\alpha}_{\widetilde{X}}(\widetilde{B}_t)} \leqslant \frac{c}{(s-t)^\beta}\big\{\|\widetilde{H}\tilde{u}\|_{C^\alpha_{\widetilde{X}}(\widetilde{B}_s)} + \|\tilde{u}\|_{L^\infty(\widetilde{B}_s)}\big\}\\
&\leqslant \frac{c}{(s-t)^\beta}\big\{\|Hu\|_{C^\alpha_X(B_s)} + \|u\|_{L^\infty(B_s)}\big\}. \qquad \square
\end{aligned}$$

Finally, by a covering argument, Theorem 5.1 follows:

Proof of Theorem 5.1. Let U', U as in the statement of the theorem, and chose a family of balls B_R^i such that

$$U' \subset \bigcup_{i=1}^k B_R^i \subset \bigcup_{i=1}^k B_{2R}^i \subset U.$$

Then by (4.9), (4.10) and Theorem 5.4,

$$\|u\|_{C^{2,\alpha}(U')} \leqslant \|u\|_{C^{2,\alpha}(\bigcup B_R^i)} \leqslant c\sum_{i=1}^k \|u\|_{C^{2,\alpha}(B_R^i)}$$

$$\leqslant c\sum_{i=1}^k \left\{\|Hu\|_{C^\alpha(B_{2R}^i)} + \|u\|_{L^\infty(B_{2R}^i)}\right\} \leqslant c\left\{\|Hu\|_{C^\alpha(U)} + \|u\|_{L^\infty(U)}\right\}. \qquad \square$$

9. Schauder estimates of higher order

In this section we want to extend Theorem 5.1 to higher order derivatives, proving Theorem 1.1 for the operator H (without lower order terms). Explicitly, we are going to prove the following theorem.

Theorem 9.1. *Let k be a positive integer. Under assumptions* (H1), (H2) (*see Section* 5)*, if $a_{ij} \in C^{k,\alpha}(U)$ for some positive integer k and $\alpha \in (0,1)$, then for every domain $U' \Subset U$ there exists a constant $c > 0$ depending on U, U', $\{X_i\}$, α, k, λ and $\|a_{ij}\|_{C^{k,\alpha}(U)}$, such that for every $u \in C^{2+k,\alpha}_{\mathrm{loc}}(U)$ with $Hu \in C^{k,\alpha}(U)$ one has*

$$\|u\|_{C^{2+k,\alpha}(U')} \leqslant c\left\{\|Hu\|_{C^{k,\alpha}(U)} + \|u\|_{L^\infty(U)}\right\}.$$

Let us recall the definition of $C^{k,\alpha}$-norm

$$\|u\|_{C^{k,\alpha}(U)} = \sum_{|I|+2h\leqslant k} \left\|\partial_t^h X^I u\right\|_{C^\alpha(U)},$$

where, for any multiindex $I = (i_1, i_2, \ldots, i_s)$, with $1 \leqslant i_j \leqslant q$, we say that $|I| = s$ and

$$X^I u = X_{i_1} X_{i_2} \cdots X_{i_s} u.$$

The first step to prove the above theorem is to get the analog of Theorem 5.2, for $\|f\|_{C^{k,\alpha}}$.

Theorem 9.2. *Under the assumptions of Theorem* 9.1*, there exists $r > 0$ such that for any $f \in C_0^{k+2,\alpha}(\widetilde{B}_r)$, for some ball $\widetilde{B}_r(t_0, \xi_0) \subset \widetilde{U}$,*

$$\|f\|_{C^{k+2,\alpha}} \leqslant c\left\{\left\|\widetilde{H} f\right\|_{C^{k,\alpha}} + \|f\|_{L^\infty}\right\}.$$

In turn, the proof of this result will be achieved through several lemmas.

Lemma 9.3. *For every $k \geqslant 0$ and every multi-index $J = (j_1, \ldots, j_k)$ there exist operators P_I and S_I that are a linear combination of frozen operators of type* 2 *and type* 1*, respectively, such that*

$$\widetilde{X}^J(af) = \sum_{m=0}^k \sum_{|I|=m} \left(P_I \widetilde{X}_I \widetilde{H}_0 f + S_I \widetilde{X}_I f\right).$$

Proof. When $k=0$ the above formula reduces to

$$af = P\widetilde{H}_0 f + Sf$$

which is (6.14).

Let us prove the formula by induction on k, so assume it holds for $|J|=k$ and let us prove its analog for a derivative of the kind $\widetilde{X}_h\widetilde{X}^J(af)$. By Proposition 6.9, for suitable operators $P_{I,p}$, $P_{I,0}$ of type two, and $S_{I,p}$, $S_{I,0}$ of type one, we have

$$\begin{aligned}
\widetilde{X}_h\widetilde{X}^J(af) &= \sum_{m=0}^{k}\sum_{|I|=m}\big(\widetilde{X}_h P_I\widetilde{X}_I\widetilde{H}_0 f + \widetilde{X}_h S_I\widetilde{X}_I f\big)\\
&= \sum_{m=0}^{k}\sum_{|I|=m}\left(\sum_{p=1}^{q} P_{I,p}\widetilde{X}_p\widetilde{X}_I\widetilde{H}_0 f + P_{I,0}\widetilde{X}_I\widetilde{H}_0 f + \sum_{p=1}^{q} S_{I,p}\widetilde{X}_p\widetilde{X}_I f + S_{I,0}\widetilde{X}_I f\right)\\
&= \sum_{m=0}^{k+1}\sum_{|I'|=m}\big(P_{I'}\widetilde{X}_{I'}\widetilde{H}_0 f + S_{I'}\widetilde{X}_{I'} f\big)
\end{aligned}$$

which is exactly the assertion for $k+1$. □

Lemma 9.4. *For any integer $k \geqslant 0$, there exists $r>0$ such that for any $f \in C_0^{k+2,\alpha}(\widetilde{B}_r)$*

$$\big\|D^{k+2}f\big\|_{C^\alpha} \leqslant c\left(\sup_{i,j}\|a_{ij}\|_{C^{k,\alpha}}\sum_{h\leqslant k+1}\big\|D^h f\big\|_{C^\alpha} + \big\|\widetilde{H}f\big\|_{C^{k,\alpha}}\right) \tag{9.1}$$

and hence, by iteration

$$\big\|D^{k+2}f\big\|_{C^\alpha} \leqslant c_{a,k}\big(\|f\|_{L^\infty} + \big\|\widetilde{H}f\big\|_{C^{k,\alpha}}\big), \tag{9.2}$$

where $c_{a,k}$ depends on $\sup\|a_{ij}\|_{C^{k,\alpha}}$.

Proof. By the previous lemma and Proposition 6.7, for any multiindex J, $|J|=k$ and m, $l=1,\dots,q$, we have

$$\widetilde{X}_m\widetilde{X}_l\widetilde{X}^J(af) = \sum_{i=0}^{k}\sum_{|I|=i}\big(T_I\widetilde{X}_I\widetilde{H}_0 f + \widetilde{X}_m\widetilde{X}_l S_I\widetilde{X}_I f\big),$$

where T_I is of type 0, S_I is of type 1. By Proposition 6.9, this last equals

$$\sum_{i=0}^{k}\sum_{|I|=i}\left(T_I\widetilde{X}_I\widetilde{H}_0 f + \sum_{p=1}^{q} T_{I,p}\widetilde{X}_p\widetilde{X}_I f + T_{I,0}\widetilde{X}_I f\right),$$

for suitable operators $T_{I,p}$, $T_{I,0}$ of type 0. Hence

$$\left\|D^{k+2}f\right\|_{C^\alpha} \leqslant c\sum_{i=0}^{k}\sum_{|I|=i}\left\|\widetilde{X}_I\widetilde{H}_0 f\right\|_{C^\alpha} + c\sum_{j\leqslant k+1}\left\|D^j f\right\|_{C^\alpha}.$$

To estimate $\|\widetilde{X}_I\widetilde{H}_0 f\|_{C^\alpha}$ we write

$$\begin{aligned}
\widetilde{X}_I\widetilde{H}_0 f &= \widetilde{X}_I\big(\widetilde{H}_0-\widetilde{H}\big)f+\widetilde{X}_I\widetilde{H}f\\
&= \sum_{i,j=1}^{q}\widetilde{X}_I\big(\big(a_{ij}(t_0,\xi_0)-a_{ij}(\cdot,\cdot)\big)\widetilde{X}_i\widetilde{X}_j f\big)+\widetilde{X}_I\widetilde{H}f\\
&= \sum_{i,j=1}^{q}\big(a_{ij}(t_0,\xi_0)-a_{ij}(\cdot,\cdot)\big)\widetilde{X}_I\widetilde{X}_i\widetilde{X}_j f - \sum_{\substack{|J'|+|J''|=|I|\\ |J'|>0}}\big(\widetilde{X}_{J'}a_{ij}\big)\widetilde{X}_{J''}\widetilde{X}_i\widetilde{X}_j f+\widetilde{X}_I\widetilde{H}f
\end{aligned}$$

and therefore (by the same "Korn's trick" explained in the proof of Theorem 5.2),

$$\left\|\widetilde{X}_I\widetilde{H}_0 f\right\|_{C^\alpha} \leqslant \varepsilon\left\|D^{k+2}f\right\|_{C^\alpha} + \sup\|a_{ij}\|_{C^{k,\alpha}}\sum_{h\leqslant k+1}\left\|D^h f\right\|_{C^\alpha} + \left\|\widetilde{H}f\right\|_{C^{k,\alpha}}$$

which gives

$$\left\|D^{k+2}f\right\|_{C^\alpha} \leqslant c\left(\varepsilon\left\|D^{k+2}f\right\|_{C^\alpha} + \sup\|a_{ij}\|_{C^{k,\alpha}}\sum_{h\leqslant k+1}\left\|D^h f\right\|_{C^\alpha} + \left\|\widetilde{H}f\right\|_{C^{k,\alpha}}\right),$$

hence

$$\left\|D^{k+2}f\right\|_{C^\alpha} \leqslant c\left(\sup\|a_{ij}\|_{C^{k,\alpha}}\sum_{h\leqslant k+1}\left\|D^h f\right\|_{C^\alpha} + \left\|\widetilde{H}f\right\|_{C^{k,\alpha}}\right).$$

Iteration gives

$$\left\|D^{k+2}f\right\|_{C^\alpha} \leqslant c_{a,k}\big(\|f\|_{C^{2,\alpha}} + \left\|\widetilde{H}f\right\|_{C^{k,\alpha}}\big)$$

and hence, by our "basic estimate" for $k=0$ (that is (5.2))

$$\left\|D^{k+2}f\right\|_{C^\alpha} \leqslant c_{a,k}\big(\|f\|_{L^\infty} + \left\|\widetilde{H}f\right\|_{C^{k,\alpha}}\big). \qquad \square$$

The problem is now to bound also time derivatives of f. Recalling that

$$\|f\|_{C^{k+2,\alpha}} = \sum_{2h+m\leqslant k+2}\left\|\partial_t^h D^m f\right\|_{C^\alpha}$$

let us prove the following lemma.

Lemma 9.5. *For any triple of integers k, h, m such that $k \geqslant 1, h \geqslant 1, m \geqslant 0$, $2h + m \leqslant k + 2$, we have*

$$\left\| \partial_t^h D^m f \right\|_\alpha \leqslant c_{a,k}\left(\left\| \widetilde{H} f \right\|_{C^{k,\alpha}} + \| f\|_{L^\infty}\right). \tag{9.3}$$

Proof. Let us prove (9.3) by induction on h. For $h = 1$ we have to show that

$$\sum_{m=0}^{k} \left\| \partial_t D^m f \right\|_\alpha \leqslant c_{a,k}\left(\left\| \widetilde{H} f \right\|_{C^{k,\alpha}} + \| f\|_{L^\infty}\right).$$

We start from the equation

$$\partial_t f = \widetilde{H} f + \sum_{i,j=1}^{q} a_{ij} \widetilde{X}_i \widetilde{X}_j f, \qquad D^m \partial_t f = D^m \widetilde{H} f + D^m \left(\sum_{i,j=1}^{q} a_{ij} \widetilde{X}_i \widetilde{X}_j f \right)$$

$$\begin{aligned}\left\| \partial_t D^m f \right\|_{C^\alpha} &\leqslant c\left(\left\| \widetilde{H} f \right\|_{C^{k,\alpha}} + \sum_{i,j=1}^{q} \sum_{l=0}^{m} \left\| D^{m-l} a_{ij} \right\|_{C^\alpha} \left\| D^{l+2} f \right\|_{C^\alpha} \right) \\ &\leqslant c_{a,k}\left(\left\| \widetilde{H} f \right\|_{C^{k,\alpha}} + \| f\|_{L^\infty}\right) \quad \text{(by (9.2))}. \end{aligned} \tag{9.4}$$

Assume (9.3) holds up to $h - 1$. Again from (9.4) we get

$$D^m \partial_t^h f = D^m \partial_t^{h-1} \widetilde{H} f + D^m \partial_t^{h-1} \left(\sum_{i,j=1}^{q} a_{ij} \widetilde{X}_i \widetilde{X}_j f \right),$$

$$\begin{aligned}\left\| D^m \partial_t^h f \right\|_{C^\alpha} &\leqslant \left\| \widetilde{H} f \right\|_{C^{m+2h-2,\alpha}} + c_{a,m+2h-2} \sum_{l=0}^{m} \left\| D^{l+2} \partial_t^{h-1} f \right\|_{C^\alpha} \\ &\leqslant c_{a,k}\left(\left\| \widetilde{H} f \right\|_{C^{k,\alpha}} + \| f\|_{L^\infty}\right)\end{aligned}$$

by inductive hypothesis and our assumptions on m, k, h. □

By Lemmas 9.4 and 9.5, Theorem 9.2 is proved. The second step of the proof of higher order Schauder estimates is contained in the following theorem.

Theorem 9.6. *Under the assumptions of Theorem* 9.1*, there exists $R > 0$ such that for any for every $f \in C^{k+2,\alpha}(\widetilde{B}_R(t_0, \xi_0))$, $0 < t < s < R$,*

$$\| f\|_{C^{k+2,\alpha}(\widetilde{B}_t)} \leqslant \frac{c}{(s-t)^{\beta_k}} \left\{ \left\| \widetilde{H} f \right\|_{C^{k,\alpha}(\widetilde{B}_s)} + \| f\|_{L^\infty(\widetilde{B}_s)} \right\}. \tag{9.5}$$

Proof *(Sketch).* The proof is now a tedious but quite straigthforward iteration of the steps of the proofs of Lemmas 9.4 and 9.5, using suitable cutoff function. We state the steps.

1. We start from (9.1), in the following slightly sharper form, which is actually what has been proved in Lemma 9.4 (here the norm of $\widetilde{H} f$ involves only spatial derivatives):

$$\left\| D^{k+2} f \right\|_{C^\alpha} \leqslant c_{k,a}\left(\sum_{j \leqslant k+1} \left\| D^j f \right\|_{C^\alpha} + \sum_{j \leqslant k} \left\| D^j \widetilde{H} f \right\|_{C^\alpha} \right)$$

and apply this to $f\zeta$, with $f \in C^{k+2,\alpha}(\widetilde{B}_s)$ and ζ cutoff function with $\widetilde{B}_t \prec \zeta \prec \widetilde{B}_s$. Then we get, with the usual techniques:

$$\left\| D^{k+2} f \right\|_{C^\alpha(\widetilde{B}_t)} \leqslant c_{k,a} \left(\sum_{j \leqslant k+1} \frac{1}{(s-t)^{j+1}} \left\| D^j f \right\|_{C^\alpha(\widetilde{B}_s)} + \frac{1}{(s-t)^{k+1}} \left\| \widetilde{H} f \right\|_{C^{k,\alpha}(\widetilde{B}_s)} \right). \tag{9.6}$$

2. Next, we refine the above argument as follows. For fixed $0 < t < s$, we set $t_j = t + \frac{j}{k+1}(s-t)$ for $j = 0, 1, 2, \ldots, k+1$, and rewrite (9.6) as

$$\begin{aligned} &\left\| D^{j+2} f \right\|_{C^\alpha(\widetilde{B}_{t_{j-1}})} \\ &\quad \leqslant c_{k,a} \left(\sum_{i \leqslant j+1} \frac{1}{(t_j - t_{j-1})^{i+1}} \left\| D^i f \right\|_{C^\alpha(\widetilde{B}_{t_j})} + \frac{1}{(t_j - t_{j-1})^{i+1}} \left\| \widetilde{H} f \right\|_{C^{j,\alpha}(\widetilde{B}_{t_j})} \right) \end{aligned}$$

for $j = 1, 2, \ldots, k$. Collecting all these inequalities and our basic estimate for $k = 0$ (that is, (7.15)) we get, by iteration

$$\left\| D^{k+2} f \right\|_{C^\alpha(\widetilde{B}_t)} \leqslant \frac{c_{k,a}}{(s-t)^{\beta_k}} \left(\left\| \widetilde{H} f \right\|_{C^{k,\alpha}(\widetilde{B}_s)} + \| f \|_{L^\infty(\widetilde{B}_s)} \right). \tag{9.7}$$

3. We now have to add, at the left-hand side of our inequalities, the terms involving time derivatives. To do this, we apply (9.3) to $f\zeta$, where ζ is a cutoff function with $\widetilde{B}_t \prec \zeta \prec \widetilde{B}_s$. By standard computations this yields, for any triple of integers k, h, m such that $k \geqslant 1$, $h \geqslant 1$, $m \geqslant 0$, $2h + m \leqslant k + 2$,

$$\left\| \partial_t^h D^m f \right\|_{C^\alpha(\widetilde{B}_t)} \leqslant c_{a,k} \left(\frac{1}{(s-t)^{k+1}} \left\| \widetilde{H} f \right\|_{C^{k,\alpha}(\widetilde{B}_s)} + \| f \|_{L^\infty(\widetilde{B}_s)} + \frac{1}{(s-t)^{k+3}} \| f \|_{C^{k+1,\alpha}(\widetilde{B}_s)} \right).$$

Together with (9.7), this allows to write

$$\begin{aligned} &\| f \|_{C^{k+2,\alpha}(\widetilde{B}_t)} \\ &\quad \leqslant c_{a,k} \left(\frac{1}{(s-t)^{\beta_k}} \left\| \widetilde{H} f \right\|_{C^{k,\alpha}(\widetilde{B}_s)} + \frac{1}{(s-t)^{\beta_k}} \| f \|_{L^\infty(\widetilde{B}_s)} + \frac{1}{(s-t)^{k+3}} \| f \|_{C^{k+1,\alpha}(\widetilde{B}_s)} \right). \end{aligned} \tag{9.8}$$

4. Reasoning like in step 2 of this proof, (9.8) iteratively implies

$$\| f \|_{C^{k+2,\alpha}(\widetilde{B}_t)} \leqslant \frac{c_{a,k}}{(s-t)^{\beta_k'}} \left(\left\| \widetilde{H} f \right\|_{C^{k,\alpha}(\widetilde{B}_s)} + \| f \|_{L^\infty(\widetilde{B}_s)} \right)$$

which ends the proof. □

Finally, we note that Theorem 9.6 immediately implies Theorem 9.1, by the same arguments of Section 8.

10. Operators with lower order terms

We now complete the proof of Theorem 1.1, considering an operator with lower order terms. We start with $C^{2,\alpha}$-estimates.

Theorem 10.1. *Let*

$$H_1 = \partial_t - \sum_{i,j=1}^{q} a_{ij}(t,x) X_i X_j + \sum_{j=1}^{q} b_j(t,x) X_j + c(t,x).$$

If (H1), (H2) *hold, then for every domain $U' \Subset U$, $\alpha \in (0,1)$, $a_{ij}, b_j, c \in C^\alpha(U)$, there exists a constant $c > 0$, depending on U, U', $\{X_i\}$, α, λ and the $C^\alpha(U)$-norms of the coefficients a_{ij}, b_j, c, such that for every $u \in C^{2,\alpha}_{\mathrm{loc}}(U)$ with $H_1 u \in C^\alpha(U)$ one has*

$$\|u\|_{C^{2,\alpha}(U')} \leqslant c\bigl\{\|H_1 u\|_{C^\alpha(U)} + \|u\|_{L^\infty(U)}\bigr\}.$$

Proof. The proof will follow the same three steps of the proof of Theorem 5.1. Let

$$\widetilde{H}_1 = \partial_t - \sum_{i,j=1}^{q} a_{ij}(t,x) \widetilde{X}_i \widetilde{X}_j + \sum_{k=1}^{q} b_k(t,x) \widetilde{X}_k + c(t,x) = \widetilde{H} + \sum_{k=1}^{q} b_k(t,x) \widetilde{X}_k + c(t,x).$$

By (6.16) and (4.3) of Proposition 4.2 we can write

$$\begin{aligned}\bigl\|\widetilde{X}_k \widetilde{X}_h f\bigr\|_{C^\alpha(\widetilde{B}_r)} &\leqslant c\left\{\bigl\|\widetilde{H} f\bigr\|_{C^\alpha(\widetilde{B}_r)} + \sum_{l=1}^{q}\bigl\|\widetilde{X}_l f\bigr\|_{C^\alpha(\widetilde{B}_r)} + \|f\|_{C^\alpha(\widetilde{B}_r)}\right\} \\ &\leqslant c\left\{\bigl\|\widetilde{H}_1 f\bigr\|_{C^\alpha(\widetilde{B}_r)} + \sum_{l=1}^{q}\bigl\|\widetilde{X}_l f\bigr\|_{C^\alpha(\widetilde{B}_r)} + \|f\|_{C^\alpha(\widetilde{B}_r)}\right\} \qquad (10.1)\end{aligned}$$

for every $f \in C_0^{2,\alpha}(\widetilde{B}_r)$, with r small enough. To get rid of the term containing $\widetilde{X}_l f$,we now apply the interpolation inequality of Theorem 7.4 which, for functions with compact support, rewrites as

$$\|Df\|_{C^\alpha(\widetilde{B}_r)} \leqslant \delta\bigl[\bigl\|D^2 f\bigr\|_{C^\alpha(\widetilde{B}_r)} + \|\partial_t f\|_{C^\alpha(\widetilde{B}_r)}\bigr] + \frac{c}{\delta^\gamma r^\beta}\|f\|_{L^\infty(\widetilde{B}_r)}. \qquad (10.2)$$

From (10.1) and (10.2) we get

$$\|f\|_{C^{2,\alpha}(\widetilde{B}_r)} \leqslant c\bigl\{\bigl\|\widetilde{H}_1 f\bigr\|_{C^\alpha(\widetilde{B}_r)} + \|f\|_{C^\alpha(\widetilde{B}_r)}\bigr\}$$

and the same reasoning of the last lines of Section 6 then gives

$$\|f\|_{C^{2,\alpha}(\widetilde{B}_r)} \leqslant c\bigl\{\bigl\|\widetilde{H}_1 f\bigr\|_{C^\alpha(\widetilde{B}_r)} + \|f\|_{L^\infty(\widetilde{B}_r)}\bigr\} \qquad (10.3)$$

that is step 1 for the operator H_1.

Now, look at the proof of Theorem 5.3, at the end of Section 7. If $f \in C^{2,\alpha}(\widetilde{U})$, $\widetilde{B}_R \subseteqq \widetilde{U}$ (R small enough to apply Theorem 5.2), $t < R$, $s = (t+R)/2$, and ζ is a cutoff function, $\widetilde{B}_t \prec \zeta \prec \widetilde{B}_s$, we can apply (10.3) to $f\zeta$, getting

$$\|f\|_{C^{2,\alpha}(\widetilde{B}_t)} \leqslant c\left\{\left\|\widetilde{H}_1(f\zeta)\right\|_{C^\alpha(\widetilde{B}_s)} + \|f\zeta\|_{L^\infty(\widetilde{B}_s)}\right\}.$$

Now, expanding the expression $\widetilde{H}_1(f\zeta)$ and bounding the $C^\alpha(\widetilde{B}_s)$-norm of each term, we get essentially the same terms obtained in the proof of Theorem 5.3; so the rest of the proof can be repeated without changes, and we get

$$\|f\|_{C^{2,\alpha}(\widetilde{B}_t)} \leqslant \frac{c}{(t-s)^\beta}\left(\left\|\widetilde{H}_1 f\right\|_{C^\alpha(\widetilde{B}_s)} + \|f\|_{L^\infty(\widetilde{B}_s)}\right) \tag{10.4}$$

that is, step 2 for the operator H_1. Finally, by the same arguments of Section 8, (10.4) implies Theorem 10.1. □

We can now easily extend to the operator with lower order terms also the $C^{k+2,\alpha}$-estimates of Theorem 9.1, completing the proof of Theorem 1.1.

Proof of Theorem 1.1. By Theorem 9.1 we can write

$$\begin{aligned}
\|u\|_{C^{2+k,\alpha}(U')} &\leqslant c\left\{\|Hu\|_{C^{k,\alpha}(U)} + \|u\|_{L^\infty(U)}\right\} \\
&\leqslant c\left\{\|H_1 u\|_{C^{k,\alpha}(U)} + \|u\|_{L^\infty(U)} + \sum_{j=1}^{q}\|b_j X_j u\|_{C^{k,\alpha}(U)} + \|cu\|_{C^{k,\alpha}(U)}\right\} \\
&\leqslant c\left\{\|H_1 u\|_{C^{k,\alpha}(U)} + \|u\|_{L^\infty(U)} + \|u\|_{C^{k+1,\alpha}(U)}\right\} \quad \text{(by (4.5)).}
\end{aligned} \tag{10.5}$$

Next, we choose an increasing family of domains U_j $(j = 0, 1, 2, \ldots, k+1)$ such that

$$U_0 = U' \Subset U_1 \Subset U_2 \Subset \cdots \Subset U_k \Subset U_{k+1} = U,$$

and rewrite (10.5) as

$$\|u\|_{C^{2+j,\alpha}(U_{k-j})} \leqslant c\left\{\|H_1 u\|_{C^{j,\alpha}(U)} + \|u\|_{L^\infty(U)} + \|u\|_{C^{j+1,\alpha}(U_{k-j+1})}\right\}$$

for $j = 1, 2, \ldots, k$. Collecting these inequalities and our basic estimate on $\|u\|_{C^{2,\alpha}(U_k)}$, that is, Theorem 5.1, we get

$$\|u\|_{C^{2+k,\alpha}(U')} \leqslant c\left\{\|H_1 u\|_{C^{k,\alpha}(U)} + \|u\|_{L^\infty(U)}\right\}$$

which is our desired result. □

11. Regularization of solutions

In this section we will prove a regularization result for the complete operator H_1 considered in Section 10. The main tool for this result is a family of mollifiers adapted to the vector fields X_j. We start with a technical lemma borrowed from [8].

Lemma 11.1. *Given an operator of type*

$$H_A = \partial_t - \sum_{i,j=1}^{q} a_{ij}(t,x) Z_i Z_j,$$

where $Z_1, \dots, Z_q$ satisfy the assumptions (H1) *in some bounded domain $\Omega \subseteq \mathbb{R}^n$, the matrix $A = \{a_{ij}\}$ satisfies assumptions* (H2), (H3) *in $U \subset \mathbb{R} \times \Omega$, and given $U' \Subset U$, there exists a new operator of type*

$$H'_A = \partial_t - \sum_{i,j=1}^{m} a'_{ij}(t,x) X_i X_j \tag{11.1}$$

such that:

(i) *the vector fields X_i's and the coefficients a'_{ij} are defined on the whole space $\mathbb{R}^{n+1}$*;
(ii) *H'_A coincides with H_A in U'*;
(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.*

Proof. Let $\varphi \in C_0^\infty(\Omega)$ be a fixed cutoff function such that $\varphi(x) = 1$ iff $x \in \overline{\Omega'}$, being Ω' an open set with $\overline{\Omega'} \Subset \Omega$ (see Lemma 2.8 in [8] for details on the existence of this cutoff function). Let us define the new system of vector fields $X_1, \dots, X_m$ $(m = q + n)$, as follows:

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

Next, let $\psi \in C_0^\infty(U)$, $\psi \equiv 1$ in U', and set

$$\{b_{ij}\}_{i,j=1}^{n+q} = \begin{bmatrix} \{a_{hk}\}_{h,k=1}^{q} & 0 \\ 0 & I_n \end{bmatrix}; \qquad a'_{ij} = \psi b_{ij} + (1-\psi)\delta_{ij}.$$

For the operator H'_A defined as in (11.1) by these vector fields X_i, conditions (i), (ii), (iv) and (H2) are obviously satisfied, 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 $X_1, \dots, X_m$ contains nonvanishing multiples of the n fields ∂_{x_k}, which span; if $\varphi(x) = 1$, then the fields $X_i = \varphi Z_i$, $i = 1, \dots, q$, 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 $\overline{\Omega'}$ $\varphi = 1$ and $\varphi_{x_k} = 0$ for every k. Iterating the above relation, we see that at the point x the system X_i $(i = 1, \dots, q)$ and the system Z_i $(i = 1, \dots, q)$ generate the same Lie algebra, that is the whole $\mathbb{R}^n$. $\square$

Let

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

By the above lemma, the vector fields X_j have globally bounded coefficients and by known results of Kusuoka and Stroock (see [23, Section 4]), there exists a fundamental solution $h(t,x,y)$ such that

$$\frac{\partial h}{\partial t}(t,x,y) = \big[Lh(t,\cdot,y)\big](x) = \big[L^* h(t,x,\cdot)\big](y) \tag{11.2}$$

for $(t,x,y) \in (0,+\infty) \times \mathbb{R}^n \times \mathbb{R}^n$, satisfying the estimates

$$\frac{1}{c|B(x,\sqrt{t})|} e^{-\frac{cd(x,y)^2}{t}} \leqslant h(t,x,y) \leqslant \frac{c}{|B(x,\sqrt{t})|} e^{-\frac{d(x,y)^2}{ct}}, \tag{11.3}$$

$$\big|X_x^I X_y^J h(t,x,y)\big| \leqslant \frac{c}{t^{(|I|+|J|)/2}|B(x,\sqrt{t})|} e^{-\frac{d(x,y)^2}{ct}} \tag{11.4}$$

for $(t,x,y) \in (0,1) \times \mathbb{R}^n \times \mathbb{R}^n$, for every multiindexes I and J. By construction $h(t,x,y)$ is the density of a probability measure and therefore

$$\int_{\mathbb{R}^n} h(t,x,y)\,dy = 1$$

for every $(t,x) \in (0,+\infty) \times \mathbb{R}^n$.

We now use this "Gaussian kernel" to build a family of mollifiers adapted to the vector fields X_i.

Theorem 11.2 *(Mollifiers). Let $\eta \in C_0^\infty(\mathbb{R})$ be a positive test function with $\int \eta(t)\,dt = 1$ and let*

$$\phi_\varepsilon(t,x,y) = \varepsilon^{-1} h(\varepsilon,x,y)\eta\left(\frac{t}{\varepsilon}\right).$$

For any $f \in C^\alpha(\mathbb{R}^{n+1})$, $\varepsilon \in (0,1)$, set

$$f_\varepsilon(t,x) = \int_{\mathbb{R}^{n+1}} \phi_\varepsilon(t-s,x,y) f(s,y)\,ds\,dy.$$

Then, there exists a constant c depending on α, $\{X_i\}$, such that

$$\|f_\varepsilon\|_{C^\alpha} \leqslant c\|f\|_{C^\alpha}. \tag{11.5}$$

Moreover,

$$\lim_{\varepsilon \to 0} \|f_\varepsilon - f\|_{L^\infty(\mathbb{R}^{n+1})} = 0. \tag{11.6}$$

Proof. To prove (11.5), we will show that $\phi_\varepsilon(t-s,x,y)$ satisfies the properties of singular integral kernels, (2.4), (2.5), (2.8), (2.9), with $\beta=\gamma=1$, uniformly in ε. By (11.3) and Lemma 3.6, we have

$$0 \leqslant \phi_\varepsilon(t-s,x,y) \leqslant c\frac{\eta(\frac{t-s}{\varepsilon})e^{-\frac{d(x,y)^2}{c\varepsilon}}}{\varepsilon|B(x,\sqrt{\varepsilon})|} \leqslant c\frac{\eta(\frac{t-s}{\varepsilon})e^{-\frac{d(x,y)^2}{c\varepsilon}}}{|B((t,x),\sqrt{\varepsilon})|} \tag{11.7}$$

and therefore when $d_{\mathrm P}((t,x),(s,y)) \leqslant \sqrt{\varepsilon}$ we obtain

$$\phi_\varepsilon(t-s,x,y) \leqslant \frac{c}{|B((t,x),(s,y))|}.$$

If now $d_{\mathrm P}((t,x),(s,y)) \geqslant \sqrt{\varepsilon}$, by the doubling condition there exists $\sigma>0$ such that

$$\frac{|B((t,x),(s,y))|}{|B((t,x),\sqrt{\varepsilon})|} \leqslant c\left(\frac{d_{\mathrm P}((t,x),(s,y))}{\sqrt{\varepsilon}}\right)^\sigma.$$

Hence

$$\frac{\eta(\frac{t-s}{\varepsilon})e^{-\frac{d(x,y)^2}{c\varepsilon}}}{|B((t,x),\sqrt{\varepsilon})|} \leqslant \frac{c}{|B((t,x),(s,y))|}\left(\frac{d_{\mathrm P}((t,x),(s,y))}{\sqrt{\varepsilon}}\right)^\sigma \eta\left(\frac{t-s}{\varepsilon}\right)e^{-\frac{d(x,y)^2}{c\varepsilon}}.$$

Since $\eta\in C_0^\infty(\mathbb{R})$ we have $\eta(v)\leqslant ce^{-|v|}$ and therefore

$$\eta\left(\frac{t-s}{\varepsilon}\right)e^{-\frac{d(x,y)^2}{c\varepsilon}} \leqslant e^{-|\frac{t-s}{\varepsilon}|}e^{-\frac{d(x,y)^2}{c\varepsilon}} \leqslant e^{-\frac{|t-s|+d(x,y)^2}{c\varepsilon}} = e^{-\frac{d_{\mathrm P}((t,x),(s,y))^2}{c\varepsilon}}. \tag{11.8}$$

Since the function $t\mapsto t^\sigma e^{-t^2}$ is bounded on $(0,\infty)$ we conclude

$$\begin{aligned}\frac{\eta(\frac{t-s}{\varepsilon})e^{-\frac{d(x,y)^2}{c\varepsilon}}}{|B((t,x),\sqrt{\varepsilon})|} &\leqslant \frac{c}{|B((t,x),(s,y))|}\left(\frac{d_{\mathrm P}((t,x),(s,y))}{\sqrt{\varepsilon}}\right)^\sigma e^{-\frac{d_{\mathrm P}((t,x),(s,y))^2}{c\varepsilon}}\\ &\leqslant \frac{c}{|B((t,x),(s,y))|}\end{aligned}$$

that is (2.4).

Let now $R=d_{\mathrm P}((t_0,x_0),(t,x))$. By 4.2(ii) we have

$$\begin{aligned}&\bigl|\phi_\varepsilon(t-s,x,y)-\phi_\varepsilon(t_0-s,x_0,y)\bigr|\\ &\quad\leqslant \Bigl(\sup_{(\tau,z)}\bigl|X_i^x\phi_\varepsilon(\tau-s,z,y)\bigr| + R\sup_{(\tau,z)}\bigl|\partial_t\phi_\varepsilon(\tau-s,z,y)\bigr|\Bigr)d_{\mathrm P}\bigl((t,x),(t_0,x_0)\bigr),\end{aligned}$$

where the sup is taken for $(\tau,z)\in B((t_0,x_0),5R)$.

Assume that

$$d_{\mathrm P}\bigl((t_0,x_0),(s,y)\bigr) \geqslant Md_{\mathrm P}\bigl((t_0,x_0),(t,x)\bigr)$$

with $M>5$; then for a suitable constant c we have $d_{\mathrm P}((t_0,x_0),(s,y)) \leqslant cd_{\mathrm P}((\tau,z),(s,y))$.

Using (11.4) and reasoning as in (11.8) we obtain

$$\left|X_i^x\phi_\varepsilon(\tau-s,z,y)\right| \leqslant \frac{c}{\varepsilon^{3/2}}\frac{e^{-\frac{d(z,y)^2}{c\varepsilon}}}{|B(z,\sqrt{\varepsilon})|}\eta\left(\frac{\tau-s}{\varepsilon}\right) \leqslant \frac{c}{\sqrt{\varepsilon}}\frac{e^{-\frac{d_{\mathrm{P}}((\tau,z),(s,y))^2}{c\varepsilon}}}{|B((\tau,z),\sqrt{\varepsilon})|} \leqslant \frac{c}{\sqrt{\varepsilon}}\frac{e^{-\frac{d_{\mathrm{P}}((t_0,x_0),(s,y))^2}{c\varepsilon}}}{|B((\tau,z),\sqrt{\varepsilon})|}$$

Assume that $d_{\mathrm{P}}((\tau,z),(s,y)) \leqslant \sqrt{\varepsilon}$. Then

$$\begin{aligned}&\left|B\big((t_0,x_0),d_{\mathrm{P}}\big((t_0,x_0),(s,y)\big)\big)\right|\\ &\quad\leqslant c\left|B\big((s,y),d_{\mathrm{P}}\big((t_0,x_0),(s,y)\big)\big)\right| \leqslant c\left|B\big((s,y),d_{\mathrm{P}}\big((\tau,z),(s,y)\big)\big)\right|\\ &\quad\leqslant c\left|B\big((\tau,z),d_{\mathrm{P}}\big((\tau,z),(s,y)\big)\big)\right| \leqslant c\left|B\big((\tau,z),\sqrt{\varepsilon}\big)\right|.\end{aligned}$$

Therefore

$$\left|X_i^x\phi_\varepsilon(\tau-s,z,y)\right| \leqslant \frac{c}{\sqrt{\varepsilon}}\frac{e^{-\frac{d_{\mathrm{P}}((t_0,x_0),(s,y))^2}{c\varepsilon}}}{|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|}.$$

Let now $d_{\mathrm{P}}((\tau,z),(s,y)) \geqslant \sqrt{\varepsilon}$ then

$$\begin{aligned}\frac{|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|}{|B((\tau,z),\sqrt{\varepsilon})|} &\leqslant c\frac{|B((\tau,z),d_{\mathrm{P}}((\tau,z),(s,y)))|}{|B((\tau,z),\sqrt{\varepsilon})|}\\ &\leqslant \left(\frac{d_{\mathrm{P}}((\tau,z),(s,y))}{\sqrt{\varepsilon}}\right)^\sigma\end{aligned}$$

so that

$$\begin{aligned}\left|X_i^x\phi_\varepsilon(\tau-s,z,y)\right| &\leqslant \frac{c}{\sqrt{\varepsilon}}\frac{e^{-\frac{d_{\mathrm{P}}((\tau,z),(s,y))^2}{c\varepsilon}}}{|B((\tau,z),\sqrt{\varepsilon})|}\\ &\leqslant \frac{c}{\sqrt{\varepsilon}}\frac{e^{-\frac{d_{\mathrm{P}}((\tau,z),(s,y))^2}{c\varepsilon}}}{|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|}\left(\frac{d_{\mathrm{P}}((\tau,z),(s,y))}{\sqrt{\varepsilon}}\right)^\sigma\\ &\leqslant \frac{c}{\sqrt{\varepsilon}}\frac{e^{-\frac{d_{\mathrm{P}}((t_0,x_0),(s,y))^2}{c\varepsilon}}}{|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|}.\end{aligned}$$

Similarly

$$\begin{aligned}R\left|\partial_t\phi_\varepsilon(\tau-s,z,y)\right| &\leqslant cR\frac{|\eta'(\frac{\tau-s}{\varepsilon})|e^{-\frac{d(z,y)^2}{c\varepsilon}}}{\varepsilon^2|B(z,\sqrt{\varepsilon})|}\\ &\leqslant \frac{R}{\sqrt{\varepsilon}}\frac{ce^{-\frac{d_{\mathrm{P}}((\tau,z),(s,y))^2}{c\varepsilon}}}{\sqrt{\varepsilon}|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|}\\ &\leqslant \frac{d_{\mathrm{P}}((t_0,x_0),(s,y))}{\sqrt{\varepsilon}}\frac{ce^{-\frac{d_{\mathrm{P}}((t_0,x_0),(s,y))^2}{c\varepsilon}}}{\sqrt{\varepsilon}|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|}\\ &\leqslant \frac{ce^{-\frac{d_{\mathrm{P}}((t_0,x_0),(s,y))^2}{c\varepsilon}}}{\sqrt{\varepsilon}|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|}.\end{aligned}$$

Hence

$$\begin{aligned}
&\left|\phi_\varepsilon(t-s,x,y)-\phi_\varepsilon(t_0-s,x_0,y)\right| \\
&\quad\leqslant \frac{c e^{-\frac{d_{\mathrm{P}}((t_0,x_0),(s,y))^2}{c\varepsilon}}}{\sqrt{\varepsilon}|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|} d_{\mathrm{P}}\big((t,x),(t_0,x_0)\big) \\
&\quad\leqslant \frac{d_{\mathrm{P}}((t_0,x_0),(s,y))}{\sqrt{\varepsilon}} \frac{c d_{\mathrm{P}}((t,x),(t_0,x_0)) e^{-\frac{d_{\mathrm{P}}((t_0,x_0),(s,y))^2}{c\varepsilon}}}{d_{\mathrm{P}}((t_0,x_0),(s,y))|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|} \\
&\quad\leqslant \frac{c d_{\mathrm{P}}((t,x),(t_0,x_0))}{d_{\mathrm{P}}((t_0,x_0),(s,y))|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|} \\
&\quad\leqslant \frac{c d_{\mathrm{P}}((t,x),(t_0,x_0))}{d_{\mathrm{P}}((t,x),(s,y))|B((t_0,x_0),d_{\mathrm{P}}((t_0,x_0),(s,y)))|}.
\end{aligned}$$

This is exactly (2.5) with $\beta = 1$.

Next, we have

$$\left|\int\limits_{d'_{\mathrm{P}}((t,x),(s,y))>r} \phi_\varepsilon(t-s,x,y)\,ds\,dy\right| \leqslant \int\limits_{\mathbb{R}^{n+1}} \varepsilon^{-1} h(\varepsilon,x,y)\eta\left(\frac{t-s}{\varepsilon}\right) ds\,dy = 1$$

which is (2.8). Also,

$$\begin{aligned}
&\lim_{r\to 0}\left|\int\limits_{d'_{\mathrm{P}}((t,x),(s,y))>r} \phi_\varepsilon(t-s,x,y)\,ds\,dy - \int\limits_{d'_{\mathrm{P}}((t_0,x_0),(s,y))>r} \phi_\varepsilon(t_0-s,x_0,y)\,ds\,dy\right| \\
&\quad= \left|\int\limits_{\mathbb{R}^{n+1}} \phi_\varepsilon(t-s,x,y)\,ds\,dy - \int\limits_{\mathbb{R}^{n+1}} \phi_\varepsilon(t-s,x,y)\,ds\,dy\right| = |1-1| = 0
\end{aligned}$$

which trivially implies (2.9) with $\gamma = 1$.

By Theorem 2.7 we get

$$|f_\varepsilon|_{C^\alpha} \leqslant c\|f\|_{C^\alpha} \quad \text{for every } \alpha \in (0,1).$$

Since we also have

$$\left|f_\varepsilon(t,x)\right| \leqslant \int\limits_{\mathbb{R}^{n+1}} \phi_\varepsilon(t-s,x,y)\left|f(s,y)\right| ds\,dy \leqslant \|f\|_\infty \int\limits_{\mathbb{R}^{n+1}} \phi_\varepsilon(t-s,x,y)\,dy = 1\cdot\|f\|_\infty,$$

we conclude

$$\|f_\varepsilon\|_{C^\alpha} \leqslant c\|f\|_{C^\alpha} \quad \text{for every } \alpha \in (0,1)$$

that is (11.5). Note that we have applied only (2.10) in Theorem 2.7, which does not require boundedness of the space.

Let us come to the proof of (11.6). Since $\int_{\mathbb{R}^{n+1}} \phi_\varepsilon(t,x,y)\,dt\,dy = 1$ we have

$$f_\varepsilon(t,x) - f(t,x) = \int \phi_\varepsilon(t-s,x,y)\big[f(s,y) - f(t,x)\big]\,ds\,dy.$$

Hence, using (11.7) and (11.8) we get

$$\begin{aligned}
&\big|f_\varepsilon(t,x) - f(t,x)\big| \\
&\quad \leqslant \int \phi_\varepsilon(t-s,x,y)\big|f(s,y) - f(t,x)\big|\,ds\,dy \\
&\quad \leqslant \frac{1}{|B((t,x),\sqrt{\varepsilon})|}\int e^{-\frac{d_{\mathrm{P}}((t,x),(s,y))^2}{c\varepsilon}} d_{\mathrm{P}}\big((s,y),(t,x)\big)^\alpha\,ds\,dy \\
&\quad = \frac{1}{|B((t,x),\sqrt{\varepsilon})|}\int\limits_{B((t,x),\sqrt{\varepsilon})} e^{-\frac{d_{\mathrm{P}}((t,x),(s,y))^2}{c\varepsilon}} d_{\mathrm{P}}\big((s,y),(t,x)\big)^\alpha\,ds\,dy \\
&\qquad + \sum_{k=0}^{+\infty}\frac{1}{|B((t,x),\sqrt{\varepsilon})|}\int\limits_{B((t,x),2^{k+1}\sqrt{\varepsilon})\setminus B((t,x),2^k\sqrt{\varepsilon})} e^{-\frac{d_{\mathrm{P}}((t,x),(s,y))^2}{c\varepsilon}} d_{\mathrm{P}}\big((s,y),(t,x)\big)^\alpha\,ds\,dy \\
&\quad \leqslant \varepsilon^{\alpha/2} + \sum_{k=0}^{+\infty}\frac{|B((t,x),2^{k+1}\sqrt{\varepsilon})|}{|B((t,x),\sqrt{\varepsilon})|}e^{-\frac{2^{2k}}{c}}\big(2^{k+1}\sqrt{\varepsilon}\big)^\alpha \\
&\quad \leqslant \varepsilon^{\alpha/2} + \sum_{k=0}^{+\infty}2^{(k+1)\sigma}e^{-\frac{2^{2k}}{c}}\big(2^{k+1}\sqrt{\varepsilon}\big)^\alpha = c\varepsilon^{\alpha/2}. \qquad \square
\end{aligned}$$

Proposition 11.3. *For any $\alpha \in (0,1)$, k even integer, U, U' bounded open sets, with $U' \Subset U$, there exists a constant c such that for any $f \in C^{k,\alpha}(U)$, $\varepsilon \in (0,1)$,*

$$\|f_\varepsilon\|_{C^{k,\alpha}(U')} \leqslant c\|f\|_{C^{k,\alpha}(U)}.$$

Proof. By (11.2), we have

$$\begin{aligned}
Lf_\varepsilon(t,x) &= \int\limits_{\mathbb{R}^{n+1}} \varepsilon^{-1}\eta\left(\frac{t-s}{\varepsilon}\right)\big[Lh(\varepsilon,\cdot,y)\big](x)f(s,y)\,ds\,dy \\
&= \int\limits_{\mathbb{R}^{n+1}} \varepsilon^{-1}\eta\left(\frac{t-s}{\varepsilon}\right)\big[L^{\mathrm{T}}h(\varepsilon,x,\cdot)\big](y)f(s,y)\,ds\,dy \\
&= \int\limits_{\mathbb{R}^{n+1}} \varepsilon^{-1}\eta\left(\frac{t-s}{\varepsilon}\right)h(\varepsilon,x,y)Lf(s,y)\,ds\,dy = (Lf)_\varepsilon(t,x).
\end{aligned}$$

Also

$$\begin{aligned}\frac{\partial}{\partial t} f_\varepsilon(t,x) &= \int\limits_{\mathbb{R}^{n+1}} \varepsilon^{-1} \frac{\partial}{\partial t}\left[\eta\left(\frac{t-s}{\varepsilon}\right)\right] h(\varepsilon,x,y) f(s,y)\,ds\,dy \\ &= -\int\limits_{\mathbb{R}^{n+1}} \varepsilon^{-1} \frac{\partial}{\partial s}\left[\eta\left(\frac{t-s}{\varepsilon}\right)\right] h(\varepsilon,x,y) f(s,y)\,ds\,dy \\ &= \int\limits_{\mathbb{R}^{n+1}} \varepsilon^{-1} \eta\left(\frac{t-s}{\varepsilon}\right) h(\varepsilon,x,y) \frac{\partial f}{\partial s}(s,y)\,ds\,dy = \left(\frac{\partial f}{\partial t}\right)_\varepsilon (t,x).\end{aligned}$$

Therefore we have

$$\mathbf{H} f_\varepsilon = (\mathbf{H} f)_\varepsilon.$$

Iterating, we obtain for any positive integer m,

$$\mathbf{H}^m f_\varepsilon = \left(\mathbf{H}^m f\right)_\varepsilon. \tag{11.9}$$

Also, we need to iterate the inequality in Theorem 9.1, as follows. Let $U' = U_m \Subset U_{m-1} \Subset \cdots \Subset U_1 = U$, then

$$\begin{aligned}\|f\|_{C^{2m,\alpha}(U_m)} &\leqslant c\left\{\|\mathbf{H} f\|_{C^{2m-2,\alpha}(U_{m-1})} + \|f\|_{L^\infty(U_{m-1})}\right\} \\ &\leqslant c\left\{\left\|\mathbf{H}^2 f\right\|_{C^{2m-4,\alpha}(U_{m-2})} + \|\mathbf{H} f\|_{L^\infty(U_{m-2})} + \|f\|_{L^\infty(U_{m-2})}\right\} \\ &\vdots \\ &\leqslant c\left\{\left\|\mathbf{H}^m f\right\|_{C^\alpha(U_1)} + \left\|\mathbf{H}^{m-1} f\right\|_{L^\infty(U_1)} + \cdots + \|f\|_{L^\infty(U_1)}\right\}.\end{aligned} \tag{11.10}$$

By (11.9), (11.10) and Proposition 11.2 we can write

$$\begin{aligned}\|f_\varepsilon\|_{C^{2m,\alpha}(U')} &\leqslant c\left\{\left\|\mathbf{H}^m f_\varepsilon\right\|_{C^\alpha(U_1)} + \left\|\mathbf{H}^{m-1} f_\varepsilon\right\|_{L^\infty(U_1)} + \cdots + \|f_\varepsilon\|_{L^\infty(U_1)}\right\} \\ &= c\left\{\left\|\left(\mathbf{H}^m f\right)_\varepsilon\right\|_{C^\alpha(U_1)} + \left\|\left(\mathbf{H}^{m-1} f\right)_\varepsilon\right\|_{L^\infty(U_1)} + \cdots + \|f_\varepsilon\|_{L^\infty(U_1)}\right\} \\ &\leqslant c\left\{\left\|\mathbf{H}^m f\right\|_{C^\alpha(U_1)} + \left\|\mathbf{H}^{m-1} f\right\|_{L^\infty(U_1)} + \cdots + \|f\|_{L^\infty(U_1)}\right\} \\ &\leqslant c\|f\|_{C^{2m,\alpha}(U)}. \qquad \square\end{aligned}$$

We will also need the following compactness lemma.

Lemma 11.4. *Let* $\{u_n\}$ *be a sequence of* $C^{k,\alpha}(U)$ *functions such that*

$$\|u_n\|_{C^{k,\alpha}(U)} \leqslant c$$

with c independent of n. Then, there exists a subsequence u_{n_h} *and a function* $u \in C^{k,\alpha}(U)$ *such that* $u_{n_h} \to u$ *in* $C^k(U)$*. Explicitly, this means that*

$$\partial_t^m X^I u_{n_h} \to \partial_t^m X^I u$$

uniformly in U for any m, I *such that* $2m + |I| \leqslant k$*.*

Proof. For any m, I such that $2m+|I| \leqslant k$, the functions $\partial_t^m X^I u_n$ are equibounded and equicontinuous (in classical sense), hence by Arzelà's theorem there exists a subsequence $\partial_t^m X^I u_{n_h}$ uniformly converging in U to some function $v_{m,I}$. Moreover, we can extract a single subsequence u_{n_h} such that all these conditions simultaneously hold. Set $u = v_{0,0}$. By [4, Proposition 2.2] (see also [8, Lemma 11.9]), this implies that $u \in C^k(U)$ and $v_{m,I} = \partial_t^m X^I u$, hence $u_{n_h} \to u$ in $C^k(U)$. Finally, passing to the limit in the inequality

$$\left|\partial_t^m X^I u_{n_h}(t,x) - \partial_t^m X^I u_{n_h}(s,y)\right| \leqslant c d_{\mathrm{P}}\big((t,x),(s,y)\big)^\alpha$$

we find that actually $u \in C^{k,\alpha}(U)$. □

Next, we apply the previous mollification machinery to prove that the a-priori estimates of higher order that we have proved in Section 9 also imply a regularization result.

Theorem 11.5. *Under the assumptions of Theorem* 1.1*, for every* $\alpha \in (0,1)$*, if* $u \in C^{2,\alpha}_{\mathrm{loc}}(U)$ *and* $H_1 u \in C^{k,\alpha}(U)$ *for some even integer k, then* $u \in C^{2+k,\alpha}_{\mathrm{loc}}(U)$*. Moreover, for every domain* $U' \Subset U$ *there exists a constant* $c > 0$ *depending on* $U, U', \{X_i\}, \alpha, k, \lambda$ *and* $\|a_{ij}\|_{C^{k,\alpha}(U)}$, $\|b_i\|_{C^{k,\alpha}(U)}$, $\|c\|_{C^{k,\alpha}(U)}$ *such that*

$$\|u\|_{C^{2+k,\alpha}(U')} \leqslant c\big\{\|H_1 u\|_{C^{k,\alpha}(U)} + \|u\|_{L^\infty(U)}\big\}.$$

Proof. Let $u \in C^{2,\alpha}_{\mathrm{loc}}(U)$, $f = H_1 u \in C^{k,\alpha}(U)$, and let a_{ij}, b_j, c be the coefficients of H_1. By Lemma 11.1, we can assume that $a_{ij} \in C^{k,\alpha}(\mathbb{R}^{n+1})$ and satisfy the ellipticity condition (H2) on the whole space. Analogously, we can extend the function f and the coefficients b_j, c to the whole space in such a way that $f, b_j, c \in C^{k,\alpha}(\mathbb{R}^{n+1})$. Assume first that c satisfies the sign condition

$$c(t,x) \geqslant c_0 > 0 \quad \text{for any } (t,x) \in \mathbb{R}^{n+1}. \tag{11.11}$$

Let now $a^\varepsilon_{ij}, b^\varepsilon_j, c^\varepsilon, f^\varepsilon$ be the mollified versions of a_{ij}, b_j, c and f, and set

$$H_1^\varepsilon = \partial_t - \sum_{i,j=1}^m a^\varepsilon_{ij}(t,x) X_i X_j + \sum_{i=1}^m b^\varepsilon_i(t,x) X_i + c^\varepsilon(t,x).$$

Note that the a^ε_{ij}'s satisfy (H2) with constant λ independent of ε. Since H_1^ε has smooth coefficients, it can be written as a Hörmander operator. This, together with condition (11.11), allows to apply known results of Bony [3]: for every point of U' we can find a neighborhood $D \Subset U$, where we can uniquely solve the classical Dirichlet problem:

$$\begin{cases} H_1^\varepsilon u^\varepsilon = f^\varepsilon & \text{in } D, \\ u^\varepsilon = u & \text{on } \partial D. \end{cases}$$

Moreover, the domain D satisfies the following regularity property (see [3, Corollary 5.2]) which will be useful later: 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_1^ε is hypoelliptic, the solution u^ε belongs to $C^\infty(D)$; in particular, $u^\varepsilon \in C^{k+2,\alpha}_{\text{loc}}(D)$, hence we can apply our a-priori estimates (Theorem 9.1), writing

$$\|u_\varepsilon\|_{C^{k+2,\alpha}(D')} \leqslant c_\varepsilon \bigl\{ \|f_\varepsilon\|_{C^{k,\alpha}(D)} + \|u_\varepsilon\|_{L^\infty(D)} \bigr\}.$$

The constant c_ε depends on the coefficients $a_{ij}^\varepsilon, b_j^\varepsilon, c^\varepsilon$ only through their $C^{k,\alpha}(D)$-norms and the ellipticity constant, hence by Proposition 11.3, if k is an even integer c_ε can be bounded independently of ε. For the same reason $\|f^\varepsilon\|_{C^{k,\alpha}(D)} \leqslant c\|f\|_{C^{k,\alpha}(U)}$, while, by the classical maximum principle (for operators with nonnegative characteristic form satisfying (11.11)),

$$\bigl\|u^\varepsilon\bigr\|_{L^\infty(D)} \leqslant \bigl\|u^\varepsilon\bigr\|_{L^\infty(\partial D)} = \|u\|_{L^\infty(\partial D)}.$$

This means that, for any $D' \Subset D$,

$$\|u_\varepsilon\|_{C^{k+2,\alpha}(D')} \leqslant c \tag{11.12}$$

with c depending on D' but not on ε. By Lemma 11.4, for every $D' \Subset D$ we can find a sequence $\varepsilon_n \to 0$ and a function $v \in C^{k+2,\alpha}(D')$ such that

$$u_{\varepsilon_n} \to v \quad \text{in } C^{k+2,0}(D').$$

By a standard "diagonal argument," we can also select a single sequence $\varepsilon_n \to 0$ and a function $v \in C^{k+2,\alpha}_{\text{loc}}(D)$ such that

$$u_{\varepsilon_n} \to v \quad \text{in } C^{k+2,0}_{\text{loc}}(D) \text{ and pointwise in } D.$$

In particular, this means that $H_1 u_{\varepsilon_n} \to H_1 v$. On the other hand, $H_1 u_{\varepsilon_n} = f_{\varepsilon_n} \to f$ by (11.6), hence

$$H_1 v = f \quad \text{in } D.$$

Our next task is to show that $v = u$ in D; this will imply $u \in C^{k+2,\alpha}_{\text{loc}}(D)$, that is the desired regularity result. To do this, we will make use of a classical argument of barriers, taken from [3], to show that $u = v$ on ∂D; this will imply that $v = u$ in D, again by the maximum principle, applied to H_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:

$$w(t,x) = e^{-K[|x-x_0|^2 + (t-t_0)^2]} - e^{-K[|x_1-x_0|^2 + (t_1-t_0)^2]}$$

with K 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 [3], $H_1 w(t,x) < 0$ in a suitable neighborhood D_1 of (t_1, x_1), for K large enough. Next, we compute, for a large constant M:

$$H_1\big(Mw \pm \big(u^\varepsilon - u\big)\big) = MH_1 w \pm \big(f^\varepsilon - f\big) < 0 \quad \text{in } D_1 \cap D$$

for M large enough, since $(f^\varepsilon - f)$ is uniformly bounded with respect to ε. Let us show that

$$Mw \pm \big(u^\varepsilon - u\big) < 0 \quad \text{on } \partial(D_1 \cap D).$$

On $D_1 \cap \partial D$, we have $Mw \pm (u^\varepsilon - u) = Mw \leqslant 0$; on the other hand, on $\partial D_1 \cap D$ we have $w \leqslant c < 0$, while $(u^\varepsilon - u)$ is uniformly bounded with respect to ε; hence for M large enough $Mw \pm (u^\varepsilon - u) \leqslant 0$. The maximum principle then implies

$$Mw \pm \big(u^\varepsilon - u\big) \leqslant 0 \quad \text{in } D_1 \cap D,$$

that is,

$$\big|u^\varepsilon - u\big| \leqslant -Mw \quad \text{in } D_1 \cap D, \text{ uniformly in } \varepsilon.$$

For $\varepsilon \to 0$ we get

$$\big|(v-u)(t,x)\big| \leqslant -Mw(t,x) \quad \text{for } (t,x) \in D_1 \cap D$$

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 (11.11). In the general case, since c is bounded we can rewrite the equation $H_1 u = f$ in the form

$$(H_1 + c_0)u = f + c_0 u,$$

where c_0 is a constant such that $c + c_0$ satisfies condition (11.11). Since $f + c_0 u \in C^{2,\alpha}(U)$, the above reasoning implies $u \in C^{4,\alpha}_{\mathrm{loc}}(U)$. Iterating this argument yields our result in the general case. $\square$

Acknowledgments

We thank Ermanno Lanconelli and Francesco Uguzzoni for some useful discussions on the subject of this paper.

Appendix A. Homogeneous groups, Rothschild–Stein "lifting and approximation" technique and their parabolic version

Let $X_1, \dots, X_q$ be C^∞ real vector fields on a domain $\Omega \subset \mathbb{R}^n$. For every multiindex $\alpha = (\alpha_1, \dots, \alpha_d)$ with $1 \leqslant \alpha_i \leqslant q$, we define

$$X_\alpha = \big[X_{\alpha_d}, \big[X_{\alpha_{d-1}}, \dots [X_{\alpha_2}, X_{\alpha_1}] \dots\big]\big],$$

and $|\alpha| = d$. We call X_α a commutator of the X_i's of length d. Assume that $X_1, \dots, X_q$ satisfy Hörmander's condition of step s at some point $x_0 \in \mathbb{R}^n$; this means that $\{X_\alpha(x_0)\}_{|\alpha| \leqslant s}$ spans $\mathbb{R}^n$. Let $\mathcal{G}(s,q)$ be the free Lie algebra of step s on q generators, that is the quotient of the free Lie algebra with q generators by the ideal generated by the commutators of length at least $s+1$, and let $N = \dim \mathcal{G}(s,q)$, as a vector space. One always has $N \geqslant n$. If $e_1, \dots, e_q$ are generators of the free Lie algebra $\mathcal{G}(q,s)$ and

$$e_\alpha = \big[e_{\alpha_d}, \big[e_{\alpha_{d-1}}, \dots [e_{\alpha_2}, e_{\alpha_1}] \dots\big]\big],$$

then there exists a set A of multiindices α so that $\{e_\alpha\}_{\alpha \in A}$ is a basis of $\mathcal{G}(q,s)$ as a vector space. This allows us to identify $\mathcal{G}(q,s)$ with $\mathbb{R}^N$. Note that $\operatorname{Card} A = N$ while, $\max_{\alpha \in A} |\alpha| = s$. The Campbell–Hausdorff series defines a multiplication in $\mathbb{R}^N$ (see, e.g., [29] or [30]) that makes $\mathbb{R}^N$ the group $N(q,s)$, that is the simply connected Lie group associated to $\mathcal{G}(q,s)$. We can naturally define dilations in $N(q,s)$ by

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

These are automorphisms of $N(q,s)$, which is therefore a homogeneous group, in the sense of Stein (see [32, pp. 618–622]). We will call it $\mathbb{G}$, leaving the numbers q, s implicitly understood. Note that the $\mathbb{G}$ is uniquely determined by the number q of the vector fields X_i and the step s of the Hörmander's condition they satisfy.

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

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

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

$$\|u \circ v\| \leqslant c\big(\|u\| + \|v\|\big) \quad \text{and} \quad \big\|u^{-1}\big\| \leqslant c\|u\|;$$
$$\frac{1}{c}|v| \leqslant \|v\| \leqslant c|v|^{1/s} \quad \text{if } \|v\| \leqslant 1.$$

- Quasidistance d:

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

  for which the following hold:

$$d(u,v) \geqslant 0 \quad \text{and} \quad d(u,v) = 0 \quad \text{if and only if } u = v;$$
$$\frac{1}{c} d(v,u) \leqslant d(u,v) \leqslant c\, d(v,u); \qquad d(u,v) \leqslant c\big\{d(u,z) + d(z,v)\big\}$$

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

If we denote by $B(u,r) \equiv B_r(u) \equiv \{v \in \mathbb{R}^N \colon d(u,v) < r\}$ 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

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

for every $u \in \mathbb{G}$ and $r > 0$, where $Q = \sum_{\alpha \in A} |\alpha|$ is called the homogeneous dimension of $\mathbb{G}$.

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

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

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

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

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

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

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, $u_\alpha \frac{\partial}{\partial u_\beta}$ is homogeneous of degree $|\beta| - |\alpha|$.

Denote by Y_j $(j = 1, \dots, q)$ the left-invariant vector field on $\mathbb{G}$ which agrees 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 $\{Y_j\}_{j=1}^q$ satisfies Hörmander's condition of step s in $\mathbb{R}^N$, and their Lie algebra coincides with $\mathcal{G}(q,s)$. Again, the Y_j's are uniquely determined by the numbers q, s, related to the original vector fields X_i defined in $\mathbb{R}^n$.

A differential operator 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 $\leqslant \ell$.

We are now in position to state the famous "Lifting and approximation" result by Rothschild and Stein [29].

Theorem A.1. *Let $X_1, \dots, X_q$ be C^∞ real vector fields on a domain $\Omega \subset \mathbb{R}^n$ satisfying Hörmander's condition of step s at some point $x_0 \in \Omega$. Then in terms of new variables, $h_1, \dots, h_{N-n}$,* *there exist smooth functions $c_{ij}(x,h)$ $(1 \leqslant i \leqslant q,\ 1 \leqslant j \leqslant N-n)$ 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*

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

satisfy Hörmander's condition of step s. Moreover, denoting by $\{\widetilde{X}_\alpha(\xi)\}_{\alpha \in A}$ a basis for $\mathbb{R}^N$ for every $\xi \in \widetilde{U}$, let us define, for $\xi, \eta \in \widetilde{U}$, the map

$$\Theta_\xi(\eta) = (u_\alpha)_{\alpha \in A} \quad \textit{with} \quad \eta = \exp\biggl(\sum_{\alpha \in A} u_\alpha \widetilde{X}_\alpha\biggr)\xi.$$

Then there exist open neighborhoods U of 0 and V, W of ξ_0 in $\mathbb{R}^N$, with $W \Subset V$ such that:

(a) *$\Theta_\xi | V$ is a diffeomorphism onto the image, for every $\xi \in V$.*
(b) *$\Theta_\xi(V) \supseteq U$ for every $\xi \in W$.*
(c) *$\Theta : V \times V \to \mathbb{R}^N$, defined by $\Theta(\xi, \eta) = \Theta_\xi(\eta)$ is $C^\infty(V \times V)$.*
(d) *In the coordinates given by Θ_ξ, we can write $\widetilde{X}_i = Y_i + R_i^\xi$ on U, where Y_i are the homogeneous left invariant vector fields defined above, and R_i^ξ are vector fields of local degree $\leqslant 0$ depending smoothly on $\xi \in W$ (the superscript ξ does not denote the variable of differentiation but dependence on the point ξ). Explicitly, this means that for every $f \in C_0^\infty(\mathbb{G})$*

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

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

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

with R_α^ξ a vector field of local degree $\leqslant |\alpha| - 1$ depending smoothly on ξ.

Roughly speaking, the above theorem says that the original system of vector fields $\{X_i\}_{i=1}^q$ defined in $\mathbb{R}^n$ can be lifted to another system $\{\widetilde{X}_i\}_{i=1}^q$ defined in $\mathbb{R}^N$ $(N > n)$, 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^ξ 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^ξ 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_ξ. Here below we add some other miscellaneous facts, related to the above concepts, which are used in this paper.

- Under the change of variables $u = \Theta_\xi(\eta)$, the measure element becomes:

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

  where $c(\xi)$ is a smooth function, bounded and bounded away from zero in V. The same is true for the change of coordinates $u = \Theta_\eta(\xi)$.

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

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

where $\|\cdot\|$ is the homogeneous norm defined above, then ρ is a quasidistance, locally equivalent to the CC-distance $\tilde{d}$ induced by the vector fields $\{\widetilde{X}_i\}$. Note, however, that $\tilde{d}$ is globally defined in $\widetilde{\Omega}$, while the map Θ is only defined in each neighborhood of $\widetilde{\Omega}$.
- 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 $\tilde{d}$ 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 A.2. (See [30, Theorem 4].) *Let $B, \widetilde{B}$ denote metric balls with respect to d (in $\mathbb{R}^n$) and $\tilde{d}$ (in $\mathbb{R}^N$), respectively. For any $r > 0$ (small enough), $x, y \in \mathbb{R}^n$, $d(x, y) \leqslant \delta r$ ($\delta < 1$ fixed), $h \in \mathbb{R}^{N-n}$, one has*

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

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

References

[1] 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.

[2] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni, Fundamental solutions for non-divergence form operators on stratified groups, Trans. Amer. Math. Soc. 356 (7) (2004) 2709–2737.

[3] J.-M. Bony, Principe du maximum, inégalité de Harnack et unicité du problème de Cauchy pour les opérateurs elliptiques dégénérés, Ann. Inst. Fourier (Grenoble) 19 (fasc. 1) (1969) 277–304 xii.

[4] A. Bonfiglioli, F. Uguzzoni, Maximum principle and propagation for intrinsically regular solutions of differential inequalities structured on vector fields, J. Math. Anal. Appl. 322 (2) (2006) 886–900.

[5] M. Bramanti, L. Brandolini, L^p-estimates for uniformly hypoelliptic operators with discontinuous coefficients on homogeneous groups, Rend. Sem. Mat. Univ. Politec. Torino 58 (4) (2000) 389–433.

[6] M. Bramanti, L. Brandolini, L^p-estimates for nonvariational hypoelliptic operators with VMO coefficients, Trans. Amer. Math. Soc. 352 (2) (2000) 781–822.

[7] 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 (2) (2005) 511–556.

[8] M. Bramanti, L. Brandolini, E. Lanconelli, F. Uguzzoni, Non-divergence equations structured on Hörmander vector fields: Heat kernels and Harnack inequalities, preprint, 2006.

[9] L.A. Caffarelli, Interior a-priori estimates for solutions of fully nonlinear equations, Ann. of Math. (2) 130 (1) (1989) 189–213.

[10] A.P. Calderón, A. Zygmund, Singular integral operators and differential equations, Amer. J. Math. 79 (1957) 901–921.

[11] S. Campanato, Proprietà di hölderianità di alcune classi di funzioni, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) 17 (1963) 175–188.

[12] L. Capogna, Q. Han, Pointwise Schauder estimates for second order linear equations in Carnot groups, in: Harmonic Analysis at Mount Holyoke, South Hadley, MA, 2001, in: Contemp. Math., vol. 320, Amer. Math. Soc., Providence, RI, 2003, pp. 45–69.

[13] Y.Z. Chen, L.C. Wu, Second Order Elliptic Equations and Elliptic Systems, Transl. Math. Monogr., vol. 174, Amer. Math. Soc., Providence, RI, 1998.

[14] G. Citti, A. Sarti, A cortical based model of perceptual completion in the roto-translation space, J. Math. Imaging Vision 24 (3) (2006) 307–326.

[15] R. Coifman, G. Weiss, Analyse Harmonique Non Commutative sur Certains Espaces Homogènes, Lecture Notes in Math., vol. 242, Springer, New York, 1971.
[16] G.B. Folland, Subelliptic estimates and function spaces on nilpotent Lie groups, Ark. Math. 13 (1975) 161–207.
[17] A.E. Gatto, S. Vági, Fractional integrals on spaces of homogeneous type, in: Analysis and Partial Differential Equations, in: Lecture Notes in Pure and Appl. Math., vol. 122, Dekker, New York, 1990, pp. 171–216.
[18] A.E. Gatto, S. Vági, On molecules and fractional integrals on spaces of homogeneous type with finite measure, Studia Math. 103 (1) (1992) 25–39.
[19] A.E. Gatto, C. Segovia, S. Vági, On fractional differentiation and integration on spaces of homogeneous type, Rev. Mat. Iberoamericana 12 (1) (1996) 111–145.
[20] L. Hörmander, Hypoelliptic second order differential equations, Acta Math. 119 (1967) 147–171.
[21] A. Korányi, S. Vági, Singular integrals on homogeneous spaces and some problems of classical analysis, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) 25 (1971) 575–648 (1972).
[22] N.V. Krylov, Lectures on Elliptic and Parabolic Equations in Hölder Spaces, Grad. Stud. Math., vol. 12, Amer. Math. Soc., Providence, RI, 1996.
[23] S. Kusuoka, D. Stroock, Applications of the Malliavin calculus. III, J. Fac. Sci. Univ. Tokyo Sect. IA Math. 34 (2) (1987) 391–442.
[24] R.A. Macías, C. Segovia, Lipschitz functions on spaces of homogeneous type, Adv. Math. 33 (3) (1979) 257–270.
[25] N.G. Meyers, Mean oscillation over cubes and Hölder continuity, Proc. Amer. Math. Soc. 15 (1964) 717–721.
[26] A. Montanari, Hölder a-priori estimates for second order tangential operators on CR manifolds, Ann. Sc. Norm. Super. Pisa Cl. Sci. (5) 2 (2) (2003) 345–378.
[27] A. Montanari, E. Lanconelli, Pseudoconvex fully nonlinear partial differential operators, strong comparison theorems, J. Differential Equations 202 (2) (2004) 306–331.
[28] A. Nagel, E.M. Stein, S. Wainger, Balls and metrics defined by vector fields I: Basic properties, Acta Math. 155 (1985) 130–147.
[29] L.P. Rothschild, E.M. Stein, Hypoelliptic differential operators and nilpotent groups, Acta Math. 137 (1976) 247–320.
[30] A. Sanchez-Calle, Fundamental solutions and geometry of sum of squares of vector fields, Invent. Math. 78 (1984) 143–160.
[31] S. Spanne, Some function spaces defined using the mean oscillation over cubes, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3) 19 (1965) 593–608.
[32] E.M. Stein, Harmonic Analysis: Real-Variable Methods, Orthogonality and Oscillatory Integrals, Princeton Univ. Press, Princeton, NJ, 1993.
[33] C.J. Xu, Regularity for quasilinear second-order subelliptic equations, Comm. Pure Appl. Math. 45 (1) (1992) 77–96.