J. Math. Anal. Appl. 533 (2024) 127996

Contents lists available at ScienceDirect

Journal of Mathematical Analysis and Applications

journal homepage: www.elsevier.com/locate/jmaa

Regular Articles

Schauder estimates for Kolmogorov-Fokker-Planck operators with coefficients measurable in time and Hölder continuous in space

Stefano Biagi, Marco Bramanti *

Dipartimento di Matematica, Politecnico di Milano, Via Bonardi 9, 20133 Milano, Italy

A R T I C L E   I N F O

Article history:
Received 3 March 2023
Available online 29 November 2023
Submitted by A.B. Cruzeiro

Keywords:
Kolmogorov-Fokker-Planck operators
Schauder estimates
Measurable coefficients

A B S T R A C T

We consider degenerate Kolmogorov-Fokker-Planck operators

$$\mathcal{L}u = \sum_{i,j=1}^{q} a_{ij}(x,t)\partial^2_{x_i x_j} u + \sum_{k,j=1}^{N} b_{jk} x_k \partial_{x_j} u - \partial_t u, \quad (x,t) \in \mathbb{R}^{N+1}, N \geq q \geq 1$$

such that the corresponding model operator having constant a_{ij} is hypoelliptic, translation invariant w.r.t. a Lie group operation in $\mathbb{R}^{N+1}$ and 2-homogeneous w.r.t. a family of nonisotropic dilations. The coefficients a_{ij} are bounded and Hölder continuous in space (w.r.t. some distance induced by $\mathcal{L}$ in $\mathbb{R}^N$) and only bounded measurable in time; the matrix $\{a_{ij}\}_{i,j=1}^{q}$ is symmetric and uniformly positive on $\mathbb{R}^q$. We prove "partial Schauder a priori estimates" of the kind

$$\sum_{i,j=1}^{q} \|\partial^2_{x_i x_j} u\|_{C^\alpha_x(S_T)} + \|Yu\|_{C^\alpha_x(S_T)} \leq c \left\{ \|\mathcal{L}u\|_{C^\alpha_x(S_T)} + \|u\|_{C^0(S_T)} \right\}$$

for suitable functions u, where $Yu = \sum_{k,j=1}^{N} b_{jk} x_k \partial_{x_j} u - \partial_t u$ and

$$\|f\|_{C^\alpha_x(S_T)} = \sup_{t \leq T} \sup_{x_1, x_2 \in \mathbb{R}^N, x_1 \neq x_2} \frac{|f(x_1,t) - f(x_2,t)|}{\|x_1 - x_2\|^\alpha} + \|f\|_{L^\infty(S_T)}.$$

We also prove that the derivatives $\partial^2_{x_i x_j} u$ are locally Hölder continuous in space and time while $\partial_{x_i} u$ and u are globally Hölder continuous in space and time.

Contents

* Corresponding author.
E-mail addresses: stefano.biagi@polimi.it (S. Biagi), marco.bramanti@polimi.it (M. Bramanti).

https://doi.org/10.1016/j.jmaa.2023.127996

1. Introduction and main results

1.1. The problem and its context

Let $N \geq q \geq 1$ be fixed. We consider a Kolmogorov-Fokker-Planck (KFP, in short) operator of the form

$$\mathcal{L}u = \sum_{i,j=1}^{q} a_{ij}(x,t)\partial^2_{x_i x_j} u + \sum_{k,j=1}^{N} b_{jk} x_k \partial_{x_j} u - \partial_t u, \qquad (x,t) \in \mathbb{R}^{N+1}. \tag{1.1.1}$$

The first-order part of the operator, also called *the drift term*, will be briefly denoted by

$$Yu = \sum_{k,j=1}^{N} b_{jk} x_k \partial_{x_j} u - \partial_t u. \tag{1.1.2}$$

Throughout the paper, points of $\mathbb{R}^{N+1}$ will be sometimes denoted by the compact notation

$$\xi = (x,t), \quad \eta = (y,s).$$

We will make the following assumptions:

(H1) $A_0(x,t) = (a_{ij}(x,t))_{i,j=1}^{q}$ is a symmetric uniformly positive matrix on $\mathbb{R}^q$ of bounded coefficients defined in $\mathbb{R}^{N+1}$, so that

$$\nu|\xi|^2 \leq \sum_{i,j=1}^{q} a_{ij}(x,t)\xi_i\xi_j \leq \nu^{-1}|\xi|^2 \tag{1.1.3}$$

for some constant $\nu > 0$, every $\xi \in \mathbb{R}^q$, every $x \in \mathbb{R}^N$ and a.e. $t \in \mathbb{R}$. The coefficients will be assumed measurable w.r.t. t and Hölder continuous w.r.t. x, in a sense that will be made precise later. (See Assumption (H3).)

(H2) The matrix $B = (b_{ij})_{i,j=1}^{N}$ satisfies the following condition: for $m_0 = q$ and suitable positive integers $m_1, \ldots, m_k$ such that

$$m_0 \geq m_1 \geq \ldots \geq m_k \geq 1 \quad \text{and} \quad m_0 + m_1 + \ldots + m_k = N, \tag{1.1.4}$$

we have

$$B = \begin{pmatrix} \mathbb{O} & \mathbb{O} & \dots & \mathbb{O} & \mathbb{O} \\ B_1 & \mathbb{O} & \dots & \dots & \dots \\ \mathbb{O} & B_2 & \dots & \mathbb{O} & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \dots & B_k & \mathbb{O} \end{pmatrix} \tag{1.1.5}$$

where every block B_j is an $m_j \times m_{j-1}$ matrix of rank m_j (for $j = 1, 2, \dots, k$).

We explicitly note that, when $q < N$, the operator $\mathcal{L}$ is *ultraparabolic*; in this context, the model operator is the so-called *Kolmogorov operator* $\mathcal{K}$, which arose in the seminal paper by Kolmogorov [16] on Brownian motion and the theory of gases. Assuming that the ambient space $\mathbb{R}^N$ has even dimension, say $N = 2n$, this model operator $\mathcal{K}$ has the following explicit expression

$$\mathcal{K} = \Delta_u + \langle u, \nabla_v \rangle - \partial_t, \quad \text{with } u, v \in \mathbb{R}^n \text{ and } t \in \mathbb{R}.$$

Clearly, $\mathcal{K}$ can be obtained from (1.1.1) by choosing

$$q = n < N, \quad A_0 = \mathrm{Id}_n, \quad m_0 = m_1 = n, \quad B = \begin{pmatrix} \mathbb{O}_n & \mathbb{O}_n \\ \mathrm{Id}_n & \mathbb{O}_n \end{pmatrix}.$$

Even if it fails to be parabolic, one can easily check that $\mathcal{K}$ satisfies *Hörmander's rank condition*, and thus $\mathcal{K}$ is C^∞-hypoelliptic by Hörmander's hypoellipticity theorem [13]; however, this fact was implicitly proved by Kolmogorov himself several years prior to [13] by exhibiting the explicit fundamental solution for $\mathcal{K}$. It is worth mentioning that, in the introduction of his paper [13], Hörmander presents the operator $\mathcal{K}$ as the main 'inspiration' for his study: in fact, $\mathcal{K}$ is a hypoelliptic operator *not satisfying* the sufficient conditions for the hypoellipticity established by Hörmander himself in his previous work [12]. That condition, much stronger than the famous Hörmander's rank condition given in [13], applies to special operators which can be seen as "small local perturbations" of hypoelliptic constant coefficient operators.

Starting with the results by Hörmander, at the beginning of the '90s the class of KFP operators *with constant coefficients* a_{ij} (of which the degenerate Kolmogorov operator $\mathcal{K}$ is a particular example) has been deeply studied by Lanconelli and Polidoro [18] under a *geometric viewpoint*. More precisely, they proved that the operator

$$\mathcal{L}u = \sum_{i,j=1}^{q} a_{ij} \partial^2_{x_i x_j} u + Yu$$

possesses the following rich underlying geometric structure:

(a) $\mathcal{L}$ is left-invariant on the *non-commutative Lie group* $\mathbb{G} = (\mathbb{R}^{N+1}, \circ)$, where the composition law $\circ$ is defined as follows

$$(y, s) \circ (x, t) = (x + E(t)y, t + s)$$
$$(y, s)^{-1} = (-E(-s)y, -s),$$

and $E(t) = \exp(-tB)$ (which is defined for every $t \in \mathbb{R}$ since the matrix B is nilpotent). For a future reference, we explicitly notice that

$$(y, s)^{-1} \circ (x, t) = (x - E(t - s)y, t - s), \tag{1.1.6}$$

and that the Lebesgue measure is the Haar measure, which is also invariant with respect to the inversion.

(b) $\mathcal{L}$ is homogeneous of degree 2 with respect to a nonisotropic family of *dilations* in $\mathbb{R}^{N+1}$, which are automorphisms of $\mathbb{G}$ and are defined by

$$D(\lambda)(x,t) \equiv (D_0(\lambda)(x), \lambda^2 t) = (\lambda^{q_1}x_1, \dots, \lambda^{q_N}x_N, \lambda^2 t), \tag{1.1.7}$$

where the N-tuple $(q_1, \dots, q_N)$ is given by

$$(q_1, \dots, q_N) = (\underbrace{1, \dots, 1}_{m_0}, \underbrace{3, \dots, 3}_{m_1}, \dots, \underbrace{2k+1, \dots, 2k+1}_{m_k}).$$

The integer

$$Q = \textstyle\sum_{i=1}^N q_i > N \tag{1.1.8}$$

is called the *homogeneous dimension* of $\mathbb{R}^N$, while $Q+2$ is the homogeneous dimension of $\mathbb{R}^{N+1}$. We explicitly point out that the exponential matrix $E(t)$ satisfies the following homogeneity property

$$E(\lambda^2 t) = D_0(\lambda)E(t)D_0\Big(\frac{1}{\lambda}\Big), \tag{1.1.9}$$

for every $\lambda > 0$ and every $t \in \mathbb{R}$ (see [18, Rem. 2.1.]).

Actually, in [18] the Authors study constant-coefficients KFP operators corresponding to a wider class of matrices B, which are not nilpotent; these more general operators are hypoelliptic, left-invariant with respect to the above operation $\circ$, but they are not necessarily homogeneous. For these operators, an explicit fundamental solution is exhibited in [18]. We refer to the introduction of the paper [4] for more details and references about the quest of a fundamental solution for KFP operators *before* the paper [18].

After the seminal paper [18], more general families of degenerate KFP operators of the kind (1.1.1), satisfying the same structural conditions on the matrices A_0 and B but with variable coefficients $a_{ij}(x,t)$, have been studied by several authors. In particular, Schauder estimates have been investigated first by Manfredini [23] and later, under more general assumptions on B, by Di Francesco-Polidoro in [8], on bounded domains, assuming the coefficients a_{ij} Hölder continuous with respect to the intrinsic distance induced in $\mathbb{R}^{N+1}$ by the vector fields $\partial_{x_1}, \dots \partial_{x_q}, Y$. We point out also the papers by Lunardi [22], Priola [28], Imbert-Mouhot [14], Wang-Zhang [30], Henderson-Snelson [10], Anceschi-Zhu [1], and the references therein, on related issues about Schauder estimates for KFP operators.

We recall that motivations to study KFP operators come both from the theory of stochastic differential equations, as well as from collisional kinetic theory. For a brief discussion of the first motivation, the reader is referred to [2, §2.1] and references therein. As to kinetic theory, we quote the essay [29] for a thorough introduction, or [25].

Recent research, especially in the field of stochastic differential equations (see e.g. [27]), suggest the importance of developing a theory allowing the coefficients a_{ij} to be rough in t (say, L^∞), and Hölder continuous (in a suitable sense) only w.r.t. the space variables. The Schauder estimates that one can reasonably expect under this mild assumption consist in controlling the Hölder seminorms w.r.t. x of the derivatives involved in the equations, uniformly in time (we will be more precise in a moment).

For *uniformly parabolic* operators, *partial Schauder estimates*, i.e. the control of the supremum in t of the Hölder quotient in space of $\partial^2_{x_i x_j} u$, under the analogous assumption on the coefficients and the right-hand side of the equation, have been proved already in 1969 by Brandt [5]. In 1980 Knerr [15] proved that, under the same assumptions, $\partial^2_{x_i x_j} u$ are actually Hölder continuous also in time, on bounded cylinders. See also the paper [19] by Lieberman, 1992, containing a unified presentation of these and related results. More recently, Krylov and Priola [17], 2010, have extended partial Schauder estimates (on the whole space) to parabolic operators with lower order unbounded terms while Lorenzi [20], 2011, has proved similar global estimates for operators with possibly unbounded coefficients. We also point out the more recent paper [9] by Dong-Kim, 2019, containing further generalizations to operators with coefficients merely measurable w.r.t. *several* variables.

In the present paper we establish global partial Schauder estimates for degenerate KFP operators (1.1.1) satisfying assumptions (H1)-(H2), with coefficients a_{ij} Hölder continuous in space, bounded measurable in time (see Assumption (H3) here below and Theorem 1.7 for the precise statement). We also show that the second derivatives $\partial^2_{x_i x_j} u$ (for $i, j = 1, 2, ..., q$) are actually locally Hölder continuous also w.r.t. time, so extending to this degenerate context the result proved for uniformly parabolic operators by Knerr [15].

Our technique to establish partial Schauder estimates is deeply rooted in the study of the model operator

$$\mathcal{L}u = \sum_{i,j=1}^{q} a_{ij}(t)\partial^2_{x_i x_j} u + \sum_{k,j=1}^{N} b_{jk} x_k \partial_{x_j} u - \partial_t u, \tag{1.1.10}$$

with coefficients *only depending on time* (in a merely L^∞ way), which has been started in [4]. In that paper, an explicit fundamental solution is built for operators (1.1.10). This fundamental solution will be the key tool used in the present paper.

Partial Schauder estimates for degenerate KFP operators have been proved also in the recent paper [6] by Chaudru de Raynal, Honoré, Menozzi, with different techniques and without getting the Hölder control in time of second order derivatives.

We also quote two other papers on KFP operators with coefficients Hölder continuous in space and L^∞ in time: [21], by Lucertini, Pagliarani, Pascucci, dealing with the construction of a fundamental solution for these operators, with consequent results about the Cauchy problem; and the preprint [11], by Henderson and Wang, containing partial Schauder estimates for a special class of KFP operators, with applications to the Landau equation. The results in [21], [11] are independent from and do not contain our results.

Finally, we point out the paper [24] by Menozzi, containing L^p estimates for the second order derivatives for KFP operators with coefficients a_{ij} continuous in space and L^∞ in time.

1.2. Assumptions and main results

We can now start giving some precise definitions which will allow to state our main result.

Let us introduce the metric structure related to the operator $\mathcal{L}$ that will be used throughout the following. The vector fields

$$X_1 = \partial_{x_1}, \ldots, X_q = \partial_q, X_0 = Y$$

form a system of Hörmander vector fields in $\mathbb{R}^{N+1}$, left-invariant w.r.t. the composition law $\circ$. The vector fields $X_i = \partial_{x_i}$ (with $i = 1, ..., q$) are homogeneous of degree 1, while $X_0 = Y$ is homogeneous of degree 2 w.r.t. the dilations $D(\lambda)$. As every set of Hörmander vector fields with drift, the system

$$\mathbf{X} = \{X_0, X_1, \ldots, X_q\}$$

induces a (weighted) control distance $d_{\mathbf{X}}$ in $\mathbb{R}^{N+1}$; we now review this definition in our special case. First of all, given $\xi = (x, t)$, $\eta = (y, s) \in \mathbb{R}^{N+1}$ and $\delta > 0$, we denote by $C_{\xi,\eta}(\delta)$ the class of *absolutely continuous* curves

$$\varphi =: [0, 1] \longrightarrow \mathbb{R}^{N+1}$$

which satisfy the following properties:

(i) $\varphi(0) = \xi$ and $\varphi(1) = \eta$;
(ii) for almost every $t \in [0,1]$ one has

$$\varphi'(t) = \sum_{i=1}^{q} a_i(t)\varphi_i(t) + a_0(t)Y_{\varphi(t)},$$

where $a_0, \dots, a_q : [0,1] \to \mathbb{R}$ are measurable functions such that

$$|a_i(t)| \leq \delta \text{ (for } i = 1, \dots, q) \quad \text{and} \quad |a_0(t)| \leq \delta^2 \qquad \text{a.e. on } [0,1].$$

Here φ_i are the components of the vector function φ and $Y_{\varphi(t)}$ stands for the vector field Y evaluated at the point $\varphi(t)$. We then define

$$d_{\mathbf{X}}(\xi, \eta) = \inf\big\{\delta > 0 : \exists\, \varphi \in C_{\xi,\eta}(\delta)\big\}.$$

Since $X_0, X_1, \dots, X_q$ satisfy Hörmander's rank condition, it is well-known that the function $d_{\mathbf{X}}$ is a distance in $\mathbb{R}^{N+1}$ (see, e.g., [26, Prop. 1.1]); in particular, for every fixed $\xi, \eta \in \mathbb{R}^{N+1}$ there always exists $\delta > 0$ such that $C_{\xi,\eta}(\delta) \neq \varnothing$. In addition, by the invariance/homogeneity properties of the X_i's, we see that

(a) $d_{\mathbf{X}}$ is left-invariant with respect to $\circ$, that is,

$$d_{\mathbf{X}}(\xi, \eta) = d_{\mathbf{X}}(\eta^{-1} \circ \xi, 0) \tag{1.2.1}$$

(b) $d_{\mathbf{X}}$ is jointly 1-homogeneous with respect to $D(\lambda)$, that is

$$d_{\mathbf{X}}(D(\lambda)\xi, D(\lambda)\eta) = \lambda d_{\mathbf{X}}(\xi, \eta) \qquad \text{for every } \lambda > 0. \tag{1.2.2}$$

As a consequence of (1.2.1), the function $\rho_{\mathbf{X}}(\xi) := d_{\mathbf{X}}(\xi, 0)$ satisfies

(1) $\rho_{\mathbf{X}}(\xi^{-1}) = \rho_{\mathbf{X}}(\xi)$;
(2) $\rho_{\mathbf{X}}(\xi \circ \eta) \leq \rho_{\mathbf{X}}(\xi) + \rho_{\mathbf{X}}(\eta)$;

moreover, by (1.2.2) we also have

(1)' $\rho_X(\xi) \geq 0$ and $\rho_{\mathbf{X}}(\xi) = 0 \Leftrightarrow \xi = 0$;
(2)' $\rho_{\mathbf{X}}(D(\lambda)\xi) = \lambda\rho_{\mathbf{X}}(\xi)$,

and this means that $\rho_{\mathbf{X}}$ is a *homogeneous norm* in $\mathbb{R}^{N+1}$.

We now observe that also the function

$$\rho(\xi) = \rho(x,t) := \|x\| + \sqrt{|t|} = \sum_{i=1}^{N} |x_i|^{1/q_i} + \sqrt{|t|} \tag{1.2.3}$$

is a homogeneous norm in $\mathbb{R}^{N+1}$ (i.e., it satisfies properties (1)'-(2)' above), and therefore it is *globally equivalent* to ρ_X: there exist $c_1, c_2 > 0$ such that

$$c_1\rho_{\mathbf{X}}(\xi) \leq \rho(\xi) \leq c_2\rho_{\mathbf{X}}(\xi) \qquad \forall\, \xi \in \mathbb{R}^{N+1}.$$

As a consequence of this fact, the map

$$d(\xi,\eta) := \rho(\eta^{-1}\circ\xi) \tag{1.2.4}$$

is a left-invariant, 1-homogeneous *quasi-distance* on $\mathbb{R}^{N+1}$. This means, precisely, that there exists a 'structural constant' $\boldsymbol{\kappa} > 0$ such that

$$d(\xi,\eta) \leq \boldsymbol{\kappa}\big(d(\xi,\zeta) + d(\eta,\zeta)\big) \qquad \forall\, \xi,\eta,\zeta \in \mathbb{R}^{N+1}; \tag{1.2.5}$$

$$d(\xi,\eta) \leq \boldsymbol{\kappa}\, d(\eta,\xi) \qquad \forall\, \xi,\eta \in \mathbb{R}^{N+1}. \tag{1.2.6}$$

The quasi-distance d is *globally equivalent* to the control distance $d_{\mathbf{X}}$; hence, we will systematically use this quasi-distance d and the associated balls

$$B_r(\xi) := \big\{\eta \in \mathbb{R}^{N+1} : \, d(\eta,\xi) < r\big\} \qquad (\text{for } \xi \in \mathbb{R}^{N+1} \text{ and } r > 0).$$

Remark 1.1. For a future reference, we list below some properties d.

(1) Owing to (1.1.6), we see that d has the following explicit expression

$$d(\xi,\eta) = \|x - E(t-s)y\| + \sqrt{|t-s|}, \tag{1.2.7}$$

for every $\xi = (x,t)$, $\eta = (y,s) \in \mathbb{R}^{N+1}$.

(2) Since $E(0) = \mathrm{Id}$, from (1.2.7) we get

$$d((x,t),(y,t)) = \|x-y\| \qquad \text{for every } x,y \in \mathbb{R}^N \text{ and } t \in \mathbb{R}, \tag{1.2.8}$$

from which we derive that the quasi-distance d *is symmetric when applied to points with the same t-coordinate*. We explicitly emphasize that an analogous property for points with the same x-coordinate *does not hold*: in fact, for every fixed $x \in \mathbb{R}^N$ and $t,s \in \mathbb{R}$ we have

$$d((x,t),(x,s)) = \|x - E(t-s)x\| + \sqrt{|t-s|} \neq \sqrt{|t-s|}.$$

(3) Let $\xi \in \mathbb{R}^{N+1}$ be fixed, and let $r > 0$. Since d satisfies the quasi-triangular inequality (1.2.5), if $\eta_1,\eta_2 \in B_r(\xi)$ we have

$$d(\eta_1,\eta_2) < 2\boldsymbol{\kappa} r.$$

(4) Taking into account the very definition of d, and bearing in mind that ρ is a *homogeneous norm* in $\mathbb{R}^{N+1}$, it is readily seen that

$$B_r(\xi) = \xi \circ B_r(0) = \xi \circ D_r\big(B_1(0)\big) \quad \forall\, \xi \in \mathbb{R}^{N+1},\ r > 0. \tag{1.2.9}$$

From this, since the Lebesgue measure is a Haar measure on $\mathbb{G} = (\mathbb{R}^{N+1},\circ)$, we immediately obtain the following identity

$$|B_r(\xi)| = |B_r(0)| = \omega\, r^{Q+2} \tag{1.2.10}$$

where $\omega := |B_1(0)| > 0$. Identity (1.2.10) illustrates the role of $Q+2$ as the homogeneous dimension of $\mathbb{R}^{N+1}$ (w.r.t. the dilations $D(\lambda)$).

The quasi-distance d allows us to define the Hölder spaces which will be used in the paper. We are interested both in Hölder norms which measure the *joint continuity in* (x,t) and in Hölder norms which measure the *continuity in* x *alone, for fixed* t.

Definition 1.2. For $-\infty \le \tau < T \le +\infty$, let $\Omega = \mathbb{R}^N \times (\tau, T)$, and let $f : \Omega \to \mathbb{R}$. Given any number $\alpha \in (0,1)$, we introduce the notation:

$$|f|_{C^\alpha(\Omega)} = \sup\left\{\frac{|f(\xi) - f(\eta)|}{d(\xi,\eta)^\alpha} : \xi,\eta \in \Omega \text{ and } \xi \neq \eta\right\}$$

$$|f|_{C^\alpha_x(\Omega)} = \operatorname*{ess\,sup}_{t\in(\tau,T)} \sup\left\{\frac{|f(x,t) - f(y,t)|}{d((x,t),(y,t))^\alpha} : x,y \in \mathbb{R}^N, x \neq y\right\}$$

$$= \operatorname*{ess\,sup}_{t\in(\tau,T)} \sup\left\{\frac{|f(x,t) - f(y,t)|}{\|x-y\|^\alpha} : x,y \in \mathbb{R}^N, x \neq y\right\}$$

(where the last equality holds by (1.2.8)). Accordingly, we define the spaces $C^\alpha(\Omega)$ and $C^\alpha_x(\Omega)$ as follows:

$$C^\alpha(\Omega) := \left\{f \in C(\Omega) \cap L^\infty(\Omega) : |f|_{C^\alpha(\Omega)} < \infty\right\} \tag{1.2.11}$$

$$C^\alpha_x(\Omega) := \left\{f \in L^\infty(\Omega) : |f|_{C^\alpha_x(\Omega)} < \infty\right\} \tag{1.2.12}$$

Remark 1.3. The space $C^\alpha(\Omega)$ endowed with the following norm

$$\|f\|_{C^\alpha(\Omega)} := \|f\|_{L^\infty(\Omega)} + |f|_{C^\alpha(\Omega)} \qquad (f \in C^\alpha(\Omega))$$

is a Banach space. Analogously, the space $C^\alpha_x(\Omega)$ endowed with the norm

$$\|f\|_{C^\alpha_x(\Omega)} := \|f\|_{L^\infty(\Omega)} + |f|_{C^\alpha_x(\Omega)} \qquad (f \in C^\alpha_x(\Omega))$$

is a Banach space.

We can now make precise our regularity assumption on the coefficients a_{ij}.

(H3) There exists $\alpha \in (0,1)$ such that

$$a_{ij} \in C^\alpha_x\left(\mathbb{R}^{N+1}\right) \qquad \text{for every } 1 \le i,j \le q.$$

In all the estimates appearing in next sections, the number

$$\Lambda = \max_{i,j=1,\dots,q} \|a_{ij}\|_{C^\alpha_x(\mathbb{R}^{N+1})}, \tag{1.2.13}$$

together with the ellipticity constant ν in (1.1.3), will quantify the dependence of the constants on the coefficients a_{ij}.

We now turn to define the functions spaces to which our *solution* u will belong.

Definition 1.4. Throughout the following, given $T \in \mathbb{R}$ we set

$$S_T := \mathbb{R}^N \times (-\infty, T).$$

We then define $\mathcal{S}^0(S_T)$ as the space of all functions $u : \overline{S}_T \to \mathbb{R}$ such that

(i) $u \in C(\overline{S_T}) \cap L^\infty(S_T)$;
(ii) for every $1 \le i, j \le q$, the *distributional derivatives* $\partial_{x_i} u, \partial^2_{x_i x_j} u \in L^\infty(S_T)$;
(iii) the *distributional derivative* $Yu \in L^\infty(S_T)$.

Moreover, given any number $\alpha \in (0, 1)$, we define

$$\mathcal{S}^\alpha(S_T) := \{u \in \mathcal{S}^0(S_T) : \partial_{x_i} u, \partial_{x_i x_j} u, Yu \in C^\alpha_x(S_T) \text{ for } 1 \le i, j \le q\}.$$

Finally, given any $\tau \in \mathbb{R}$ with $\tau < T$, we define

$$\begin{aligned} &\mathcal{S}^0(\tau; T) = \{u \in \mathcal{S}^0(S_T) : u(x, t) = 0 \text{ for every } t \le \tau\}, \\ &\mathcal{S}^\alpha(\tau; T) := \mathcal{S}^\alpha(S_T) \cap \mathcal{S}^0(\tau; T) \qquad (\text{for } \alpha \in (0, 1)). \end{aligned}$$

Remark 1.5. On account of assumption (H3), we immediately obtain the following facts which shall be repeatedly used throughout the rest of the paper.

(1) If $u \in \mathcal{S}^0(S_T)$, then $\mathcal{L}u \in L^\infty(S_T)$.
(2) If $u \in \mathcal{S}^\alpha(S_T)$, then $\mathcal{L}u \in C^\alpha_x(S_T)$.
(3) If $u \in \mathcal{S}^0(S_T)$ and $\partial^2_{x_i x_j} u$, $\mathcal{L}u \in C^\alpha_x(S_T)$ $(1 \le i, j \le q)$, then $Yu \in C^\alpha_x(S_T)$.

Some of the results in the next sections are proved under the assumption that $u \in \mathcal{S}^0(S_T)$ and $\mathcal{L}u \in C^\alpha_x(S_T)$; this is slightly weaker than assuming $u \in \mathcal{S}^\alpha(S_T)$.

Remark 1.6 *(Regularity of functions in $S^\alpha(S_T)$).* We will prove in the subsequent sections the following 'higher-regularity' results:

(1) if $u \in \mathcal{S}^0(S_T)$, then u and $\partial_{x_1} u, \dots, \partial_{x_q} u$ are locally Hölder-continuous in the joint variables (see, precisely, Proposition 4.3);
(2) if $u \in \mathcal{S}^\alpha(S_T)$ for some $\alpha \in (0, 1)$, then the distributional derivatives $\partial^2_{x_i x_j} u$ (for $1 \le i, j \le q$) are actually continuous (and locally Hölder continuous in a weaker sense) on S_T (see Theorem 4.9).

As a consequence, every function $u \in \mathcal{S}^\alpha(S_T)$ actually has *classical continuous derivatives* $\partial_{x_i} u, \partial^2_{x_i x_j} u$ (for $1 \le i, j \le q$); instead, the distributional derivative Yu is continuous in space for every fixed t, but it may be only L^∞ w.r.t. time.

We are finally in position to state our main result.

Theorem 1.7 *(Schauder estimates). Let $\mathcal{L}$ be an operator as in* (1.1.1)*, and assume that* (H1), (H2), (H3) *hold, for some $\alpha \in (0, 1)$.*

Then, the following Schauder-type estimates hold true.

(1) *For every $T > 0$ there exists a constant $c > 0$, depending on T, α, the matrix B in* (1.1.5) *and the numbers ν and Λ in* (1.1.3)-(1.2.13)*, respectively, such that*

$$\sum_{i,j=1}^q \|\partial^2_{x_i x_j} u\|_{C^\alpha_x(S_T)} + \|Yu\|_{C^\alpha_x(S_T)} + \sum_{i=1}^q \|\partial_{x_i} u\|_{C^\alpha(S_T)} + \|u\|_{C^\alpha(S_T)} \le c\big(\|\mathcal{L}u\|_{C^\alpha_x(S_T)} + \|u\|_{C^0(S_T)}\big),$$

for every function $u \in \mathcal{S}^\alpha(S_T)$.

(2) *For every* $T > \tau > -\infty$ *and every compact set* $K \subset \mathbb{R}^N$ *there exists a constant* $c > 0$*, depending on* $K, \tau, T, \alpha, B, \nu, \Lambda$*, such that*

$$|\partial^2_{x_i x_j} u(\xi) - \partial^2_{x_i x_j} u(\eta)| \leq c\big(\|\mathcal{L}u\|_{C^\alpha_x(S_T)} + \|u\|_{C^\alpha(S_T)}\big)\big(d(\xi,\eta)^\alpha + |t-s|^{\alpha/q_N}\big)$$

for every $\xi = (x,t)$*,* $\eta = (y,s) \in K \times [\tau, T]$ *and every* $u \in \mathcal{S}^\alpha(S_T)$*. Here, the number* q_N *is the largest exponent in the dilations* $D(\lambda)$*, see* (1.1.7)*.*

Remark 1.8. (i). As observed in Remark 1.6, the finiteness of the quantities

$$\textstyle\sum_{i=1}^q \|\partial_{x_i} u\|_{C^\alpha(S_T)}, \quad \|u\|_{C^\alpha(S_T)},$$

as well as the finiteness of the space-time Hölder quotient in point (2) of the above theorem, are not obvious *a priori* for a function in $\mathcal{S}^\alpha(S_T)$, but they will be actually proved.

(ii). While $\partial^2_{x_i x_j} u$ $(i,j = 1,2,...,q)$ are locally Hölder continuous in space and time, note that a similar property cannot be assured, in general, for Yu. To see this, it is enough to consider an equation of the kind

$$\mathcal{L}u(x,t) = f(t)$$

with f bounded discontinuous function, and u independent of x.

(iii). Since, in the degenerate case, $q_N \geq 3$, the term $|t_1 - t_2|^{\alpha/q_N}$ in the right-hand side of (3.5.1) is larger than the 'expected'

$$|t_1 - t_2|^{\alpha/2},$$

(at least when $|t_1 - t_2| \leq 1$). Also, the constant $c > 0$ depends on the fixed compact set $K \times [\tau, T] \subseteq S_T$. On the other hand, we observe that this mild t-continuity of $\partial^2_{x_i x_j} u$ is obtained *without any* t*-continuity assumption on* $\mathcal{L}u$. Moreover, from the proof of Theorem 3.18 it will be apparent that in the uniformly parabolic case ($B = 0$ and $q = N$) our argument would give exactly

$$|\partial^2_{x_i x_j} u(\xi) - \partial^2_{x_i x_j} u(\eta)| \leq c\big\{\|\mathcal{L}u\|_{C^\alpha_x(S_T)} + \|u\|_{C^\alpha(S_T)}\big\}\big(|x-y| + |t-s|^{\alpha/2}\big).$$

Our result is therefore consistent with the classical result by Knerr [15] which holds for uniformly parabolic operators on bounded cylinders.

1.3. Structure of the paper

After a short section of preliminaries (§2), the paper will proceed in two main steps: the study of the model operator (1.1.10) with coefficients only depending on t (§3) and the study of operators (1.1.1) of general type (§ 4). In section 3 we deepen the study of the fundamental solution for model operators (1.1.10) computed in the previous paper [4]. Thanks to the stronger assumption that we make in this paper on the matrix B with respect to those assumed in [4] (the corresponding operators with *constant* a_{ij} in this paper are both left invariant and homogeneous, while in [4] they are only left invariant) it is possible to sharpen the estimates on the fundamental solutions. Actually, in § 3.2 we establish sharp upper bounds on the fundamental solution and its space derivatives *of every order*, and other relevant properties of this kernel. These upper bounds and properties allow us to establish, in § 3.3, suitable representation formulas for a function u and its derivatives $\partial^2_{x_i x_j} u$ in terms of $\mathcal{L}u$. In turn, thanks to these representation formulas we will establish Hölder estimates in space for $\partial^2_{x_i x_j} u$ in § 3.4, and local Hölder estimates in space and time for $\partial^2_{x_i x_j} u$ in § 3.5. These results are established with techniques of singular integrals, and refer to operators with coefficients only depending on t. Starting with these results, in §4 analogous results are established for operators with coefficients $a_{ij}(x,t)$, exploiting the classical perturbative method used for Schauder estimates. First, in §4.1, Hölder estimates for $\partial^2_{x_ix_j}u$ are proved for functions with small support. Then, in § 4.2, some interpolation inequalities on first order derivatives are proved, which allow to get, in § 4.3, global Schauder estimates in space, extended in § 4.4 to Hölder estimates in space and time on $\partial^2_{x_ix_j}u$.

Acknowledgments. We wish to thank Sergio Polidoro for several useful discussions on the subject of this paper. We also wish to thank the anonymous Referee who careful read the paper, gave useful suggestions, and pointed out to us several relevant bibliography items.

2. Preliminaries and known results

The following mean value theorem, which is well known for systems of left-invariant homogeneous Hörmander vector fields, will be useful.

Theorem 2.1. *There exist an absolute constant $c > 0$ and a number $\delta \in (0,1)$, depending on $\boldsymbol{\kappa}$ in (1.2.5)-(1.2.6), such that, for every fixed $\xi_0 \in \mathbb{R}^{N+1}$, every $r > 0$ and every f Lipschitz-continuous in $\overline{B}_r(\xi_0)$, one has*

$$|f(\xi) - f(\xi_0)| \le c\Big(d(\xi,\xi_0)\cdot \sup_{B_R(\xi_0)} \sqrt{\sum_{i=1}^q |\partial_{x_i} f|^2} + d(\xi,\xi_0)^2 \cdot \sup_{B_R(\xi_0)} |Yf|\Big),$$

for every $\xi \in B_r(\xi_0)$. Moreover, one also has

$$|f(\xi) - f(\eta)| \le c\Big(d(\xi,\eta)\cdot \sup_{B_R(\xi_0)} \sqrt{\sum_{i=1}^q |\partial_{x_i} f|^2} + d(\xi,\eta)^2 \cdot \sup_{B_R(\xi_0)} |Yf|\Big)$$

for every $\xi,\eta \in B_{\delta r}(\xi_0)$.

The next geometric lemma follows by standard computations in doubling metric measure spaces, recalling (1.2.10).

Lemma 2.2. *Let $\alpha > 0$ be fixed, and let Q be as in (1.1.8). Then, there exists a constant $c_\alpha > 0$ such that, for every $\xi \in \mathbb{R}^{N+1}$ and every $r > 0$, one has*

$$\int_{\{\eta:\, d(\xi,\eta)<r\}} \frac{1}{d(\xi,\eta)^{Q+2-\alpha}}\, \eta \le c_\alpha r^\alpha \tag{2.0.1}$$

$$\int_{\{\eta:\, d(\xi,\eta)>r\}} \frac{1}{d(\xi,\eta)^{Q+2+\alpha}}\, d\eta \le \frac{c_\alpha}{r^\alpha}. \tag{2.0.2}$$

We also state the following simple fact which shall be repeatedly used throughout the rest of the paper.

Lemma 2.3. *There exists an absolute constant $\boldsymbol{\vartheta} > 0$ such that, if ξ_1, ξ_2 and η are points in $\mathbb{R}^{N+1}$ which satisfy $d(\xi_1,\eta) \ge 2\boldsymbol{\kappa}\, d(\xi_1,\xi_2)$, one has*

$$\boldsymbol{\vartheta}^{-1} d(\xi_2,\eta) \le d(\xi_1,\eta) \le \boldsymbol{\vartheta} d(\xi_2,\eta). \tag{2.0.3}$$

Here, $\boldsymbol{\kappa} > 0$ is the constant appearing in (1.2.5)-(1.2.6).

Thanks to Lemmas 2.2-2.3 we can establish the following C^α continuity result about "fractional integrals" which will be useful in our estimates.

Proposition 2.4 *(Fractional integrals). Let Q be as in (1.1.8) and $\beta \in [1, Q+2)$. Moreover, let $k = k(\xi,\eta)$ be a kernel satisfying the following properties:*

(1) *there exists a constant $c_1 > 0$ such that*

$$|k(\xi,\eta)| \leq \frac{c_1}{d(\xi,\eta)^{Q+2-\beta}} \quad \forall\, \xi \neq \eta \in \mathbb{R}^{N+1}; \tag{2.0.4}$$

(2) *there exist constants σ, $c_2 > 0$ such that*

$$|k(\xi_1,\eta) - k(\xi_2,\eta)| \leq c_2 \frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+3-\beta}} \quad \forall\, d(\xi_1,\eta) \geq \sigma\, d(\xi_1,\xi_2). \tag{2.0.5}$$

For every fixed $\overline{\xi} \in \mathbb{R}^{N+1}$ and $r > 0$, we introduce the function space

$$\mathbb{X}_\infty(B_r(\overline{\xi})) := \{f \in L^\infty(\mathbb{R}^{N+1}) : \ f \equiv 0 \ a.e. \ in \ \mathbb{R}^{N+1} \setminus B_r(\overline{\xi})\},$$

and we define the linear operator

$$\mathbb{X}_\infty(B_r(\overline{\xi})) \ni f \mapsto Tf(\xi) = \int_{\mathbb{R}^{N+1}} k(\xi,\eta) f(\eta)\, d\eta.$$

Then, for every $\alpha \in (0,1)$ there exists an 'absolute' constant $c > 0$, depending on α, β but independent of $f, \overline{\xi}, r$ and of the kernel k, *such that*

$$\|Tf\|_{L^\infty(B_r(\overline{\xi}))} \leq c c_1 r^\beta \|f\|_{L^\infty(B_R(\overline{\xi}))} \tag{2.0.6}$$

$$|Tf|_{C^\alpha(B_r(\overline{\xi}))} \leq c(c_1 + c_2) r^{\beta-\alpha} \|f\|_{L^\infty(B_R(\overline{\xi}))}. \tag{2.0.7}$$

Proof. Let $f \in \mathbb{X}_\infty(B_r(\overline{\xi}))$ be arbitrarily fixed. Using (2.0.1) and (2.0.4), and taking into account Remark 1.1-(3), for every $\xi \in B_r(\overline{\xi})$ we have

$$\begin{aligned} |Tf(\xi)| &\leq \int_{\{d(\eta,\overline{\xi})<r\}} \frac{c_1}{d(\xi,\eta)^{Q+2-\beta}} |f(\eta)|\, d\eta \\ &\leq c_1 \|f\|_{L^\infty(B_r(\overline{\xi}))} \int_{\{d(\eta,\overline{\xi})<r\}} \frac{1}{d(\xi,\eta)^{Q+2-\beta}}\, d\eta \\ &\leq c_1 \|f\|_{L^\infty(B_r(\overline{\xi}))} \int_{\{d(\xi,\eta)<2\boldsymbol{\kappa} r\}} \frac{1}{d(\xi,\eta)^{Q+2-\beta}}\, d\eta \\ &\leq c_1 c_\beta (2\boldsymbol{\kappa})^\beta r^\beta \|f\|_{L^\infty(B_R(\overline{\xi}))}, \end{aligned}$$

hence

$$\|Tf\|_{L^\infty(B_r(\overline{\xi}))} \leq c' c_1 r^\beta \|f\|_{L^\infty(B_R(\overline{\xi}))} \qquad (\text{with } c' := c_\beta (2\boldsymbol{\kappa})^\beta),$$

which is (2.0.6). Moreover, for every $\xi_1, \xi_2 \in B_r(\overline{\xi})$ one has

$$
\begin{aligned}
|Tf(\xi_1) - Tf(\xi_2)| &\leq \|f\|_{L^\infty(B_r(\overline{\xi}))} \int\limits_{B_r(\overline{\xi})} |k(\xi_1,\eta) - k(\xi_2,\eta)|\, d\eta \\
&= \|f\|_{L^\infty(B_r(\overline{\xi}))} \Bigg(\int\limits_{\{\eta:\, d(\xi_1,\eta)\geq \sigma d(\xi_1,\xi_2)\}} + \int\limits_{\{\eta\in B_r(\overline{\xi}):\, d(\xi_1,\eta)<\sigma d(\xi_1,\xi_2)\}} \Bigg) \{\cdots\}\, d\eta \\
&\equiv \|f\|_{L^\infty(B_r(\overline{\xi}))} \cdot \big(\mathrm{A}+\mathrm{B}\big).
\end{aligned} \tag{2.0.8}
$$

Next, by (2.0.1), (2.0.5) and Remark 1.1-(3), we get

$$
\begin{aligned}
\mathrm{A} &\leq c_2 \int\limits_{B_r'(\overline{\xi})} \frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+3-\beta}}\, d\eta \leq \frac{c_2}{\sigma^\alpha} \cdot d(\xi_1,\xi_2)^\alpha \int\limits_{B_r(\overline{\xi})} \frac{d(\xi_1,\eta)^{1-\alpha}}{d(\xi_1,\eta)^{Q+3-\beta}}\, d\eta \\
&\leq \frac{c_2}{\sigma^\alpha} \cdot d(\xi_1,\xi_2)^\alpha \int\limits_{\{d(\xi_1,\eta)<2\boldsymbol{\kappa} r\}} \frac{1}{d\,(\xi_1,\eta)^{Q+2-(\beta-\alpha)}}\, d\eta \\
&\quad \text{(by (2.0.1), since } 0<\alpha<1\leq\beta) \\
&\leq c\, d(\xi_1,\xi_2)^\alpha r^{\beta-\alpha}.
\end{aligned} \tag{2.0.9}
$$

As to B, again by (2.0.1) and (2.0.4) we get

$$
\begin{aligned}
\mathrm{B} &\leq \int\limits_{B_r''(\overline{\xi})} (|k(\xi_1,\eta)| + |k(\xi_2,\eta)|\, d\eta \\
&\leq c_1 \int\limits_{B_r''(\overline{\xi})} \Big(\frac{1}{d(\xi_1,\eta)^{Q+2-\beta}} + \frac{1}{d(\xi_2,\eta)^{Q+2-\beta}} \Big) d\eta \\
&\leq c_1 \Bigg(\int\limits_{\{d(\xi_1,\eta)<\sigma d(\xi_1,\xi_2)\}} \frac{1}{d(\xi_1,\eta)^{Q+2-\beta}}\, d\eta \\
&\qquad + \int\limits_{\{d(\xi_2,\eta)<\boldsymbol{\kappa}^2(\sigma+1) d(\xi_1,\xi_2)\}} \frac{1}{d(\xi_1,\eta)^{Q+2-\beta}}\, d\eta \Bigg) \\
&\leq c\, d(\xi_1,\xi_2)^\beta \\
&\quad \text{(by Remark 1.1-(3), since } \xi_1,\xi_2\in B_r(\overline{\xi})) \\
&\leq c\, d(\xi_1,\xi_2)^\alpha r^{\beta-\alpha}.
\end{aligned} \tag{2.0.10}
$$

Due to the arbitrariness of $\xi_1,\xi_2 \in B_r(\overline{\xi})$, by (2.0.8)-to-(2.0.10) we get

$$
|Tf|_{C^\alpha(B_r(\overline{\xi}))} \leq c\, r^{\beta-\alpha} \|f\|_{L^\infty(B_r(\overline{\xi}))},
$$

so the proof is complete. □

We end this section with another useful technical lemma.

Lemma 2.5. *There exists an absolute constant $c>0$ such that*

$$
\|E(t)x\| \leq c\rho(x,t) = c\big(\|x\| + \sqrt{|t|}\big) \qquad \forall\ x\in\mathbb{R}^N,\ t\in\mathbb{R}. \tag{2.0.11}
$$

Proof. First of all, since the function $(x,t) \mapsto \|E(t)x\|$ is continuous on $\mathbb{R}^{N+1}$, it is possible to find a constant $M > 0$ such that

$$\|E(\tau)\xi\| \le M \quad \text{for every } \|\xi\| \le 1 \text{ and } |\tau| \le 1.$$

We then fix $(x,t) \in \mathbb{R}^{N+1} \setminus \{(0,0)\}$ and define

$$\lambda = \|x\| + \sqrt{|t|} \quad \text{and} \quad (\xi,\tau) := \left(D_0\left(\frac{1}{\lambda}\right)x, \frac{t}{\lambda^2}\right).$$

Since $\|\cdot\|$ is D_0-homogeneous of degree 1, it is immediate to recognize that $\|\xi\|, |\tau| \le 1$; thus, by (1.1.9) we get

$$M \ge \|E(\tau)\xi\| = \left\|E\left(\frac{t}{\lambda^2}\right)D_0\left(\frac{1}{\lambda}\right)x\right\| = \left\|D_0\left(\frac{1}{\lambda}\right)E(t)x\right\| = \frac{1}{\lambda}\|E(t)x\|,$$

so that

$$\|E(t)x\| \le M\lambda = c\left(\|x\| + \sqrt{|t|}\right),$$

and this gives the desired (2.0.11) for $(x,t) \neq (0,0)$. Since this estimate is clearly satisfied when $x = t = 0$, the proof is complete. □

3. Operators with measurable coefficients $a_{ij}(t)$

3.1. Known results on the fundamental solution

Throughout this section, we consider an operator $\mathcal{L}$ of the form (1.1.1) and satisfying (H1)-(H2), with bounded measurable coefficients a_{ij} only depending on t, that is,

$$\mathcal{L}u = \sum_{i,j=1}^{q} a_{ij}(t)\,\partial^2_{x_ix_j}u + \sum_{k,j=1}^{N} b_{jk}x_k\partial_{x_j}u - \partial_t u, \qquad (x,t) \in \mathbb{R}^{N+1}. \tag{3.1.1}$$

In [4], an explicit fundamental solution for $\mathcal{L}$ is computed, and its properties are studied. The next theorem summarizes some results in [4] that we will need.

We point out that, since our assumption (H1) on the matrix B is stronger than the one made in [4] (here the model operator with constant a_{ij} is both left invariant and homogeneous, while in [4] it is only left invariant), here we specialize the formulas and results to our simpler situation.

Theorem 3.1 *(Fundamental solution for operators with t-variable coefficients). Under assumptions* (H1)-(H2) *above, let $C(t,s)$ be the $N \times N$ matrix defined as*

$$C(t,s) = \int_s^t E(t-\sigma)\cdot\begin{pmatrix} A_0(\sigma) & 0 \\ 0 & 0\end{pmatrix}\cdot E(t-\sigma)^T\,d\sigma \quad (\textit{with } t > s) \tag{3.1.2}$$

(we recall that $E(\sigma) = \exp(-\sigma B)$, see (1.1.5)). Then, the matrix $C(t,s)$ is symmetric and positive definite *for every $t > s$. Moreover, if we define*

$$\Gamma(x,t;y,s) = \frac{1}{(4\pi)^{N/2}\sqrt{\det C(t,s)}}e^{-\frac{1}{4}\langle C(t,s)^{-1}(x-E(t-s)y),\,x-E(t-s)y\rangle}\cdot \mathbf{1}_{\{t>s\}} \tag{3.1.3}$$

(*where* $\mathbf{1}_A$ *denotes the indicator function of a set* A), *then* Γ *enjoys the following properties, so that* Γ *is the* fundamental solution *for* $\mathcal{L}$ *with pole at* (y, s).

(1) *In the open set* $\mathcal{O} := \{(x, t; y, s) \in \mathbb{R}^{2N+2} : (x, t) \neq (y, s)\}$, *the function* Γ *is* jointly continuous *in* $(x, t; y, s)$ *and* C^∞ *with respect to* x, y. *Moreover, for every multi-indexes* α, β *the functions*

$$\partial_x^\alpha \partial_y^\beta \Gamma = \frac{\partial^{\alpha+\beta}\Gamma}{\partial x^\alpha \partial y^\beta}$$

are jointly continuous *in* $(x, t; y, s) \in \mathcal{O}$. *Finally,* Γ *and* $\partial_x^\alpha \partial_y^\beta \Gamma$ *are* Lipschitz continuous *with respect to* t, s *in any region* $\mathcal{R}$ *of the form*

$$\mathcal{R} = \{(x, t; y, s) \in \mathbb{R}^{2N+2} : H \leq s + \delta \leq t \leq K\},$$

where $H, K \in \mathbb{R}$ *and* $\delta > 0$ *are arbitrarily fixed.*

(2) *For every fixed* $y \in \mathbb{R}^N$ *and* $t > s$, *we have*

$$\lim_{|x| \to +\infty} \Gamma(x, t; y, s) = 0.$$

(3) *For every fixed* $(y, s) \in \mathbb{R}^{N+1}$, *we have*

$$\big(\mathcal{L}\Gamma(\cdot; y, s)\big)(x, t) = 0 \qquad \textit{for every } x \in \mathbb{R}^N \textit{ and a.e. } t.$$

(4) *For every fixed* $x \in \mathbb{R}^N$ *and every* $t > s$, *we have*

$$\int_{\mathbb{R}^N} \Gamma(x, t; y, s)\, dy = 1. \tag{3.1.4}$$

(5) *For every* $f \in C(\mathbb{R}^N) \cap L^\infty(\mathbb{R}^N)$ *and every* $s \in \mathbb{R}$, *the function*

$$u(x, t) = \int_{\mathbb{R}^N} \Gamma(x, t; y, s) f(y)\, dy$$

is the unique solution to the Cauchy problem

$$\begin{cases} \mathcal{L}u = 0 & \textit{in } \mathbb{R}^N \times (s, \infty) \\ u(\cdot, s) = f \end{cases} \tag{3.1.5}$$

In particular, $u(\cdot, s) \to f$ *uniformly in* $\mathbb{R}^N$ *as* $t \to s^+$.

Finally, the function $\Gamma^*(x, t; y, s) := \Gamma(y, s; x, t)$ *satisfies dual properties of* (2)-(4) *with respect to the* formal adjoint *of* $\mathcal{L}$, *that is,*

$$\mathcal{L}^* = \textstyle\sum_{i,j=1}^q a_{ij}(s)\partial_{y_i y_j} - \sum_{k,j=1}^N b_{jk} y_k \partial_{y_i} + \partial_s,$$

and thus Γ^* *is the fundamental solution of* $\mathcal{L}^*$.

The precise definition of *solution to the Cauchy problem* (3.1.5) requires some care, see [4, Definitions 1.2 and 1.3] for the details. Let us now further specialize our class of operators to the model operators with *constant coefficients* a_{ij}. Keeping our assumption (H2) on the matrix B, let

$$\mathcal{L}_\alpha u = \alpha \sum_{i=1}^{q} \partial^2_{x_i x_i} u + \sum_{k,j=1}^{N} b_{jk} x_k \partial_{x_j} u - \partial_t u \tag{3.1.6}$$

for some $\alpha > 0$. Then the results of the above theorem apply in a simpler form. Actually, the following facts are proved already in [18].

Theorem 3.2 *(Fundamental solution for operators with constant coefficients). Let $\alpha > 0$ be fixed, and let Γ_α be the fundamental solution of the operator $\mathcal{L}_\alpha$ in (3.1.6), whose existence is guaranteed by Theorem 3.1. Then:*

(1) Γ_α *is a* kernel of convolution type*, that is,*

$$\begin{aligned} \Gamma_\alpha(x,t;y,s) &= \Gamma_\alpha\big(x - E(t-s)y, t-s; 0,0\big) \\ &= \Gamma_\alpha\big((y,s)^{-1} \circ (x,t); 0,0\big); \end{aligned} \tag{3.1.7}$$

(2) *The matrix $C(t,s)$ in (3.1.2) takes the simpler form*

$$C(t,s) = C_0(t-s), \tag{3.1.8}$$

where $C_0(\tau)$ is the $N \times N$ matrix defined as

$$C_0(\tau) = \alpha \int_0^\tau E(t-\sigma) \cdot \begin{pmatrix} I_q & 0 \\ 0 & 0 \end{pmatrix} \cdot E(t-\sigma)^T d\sigma \qquad (\tau > 0).$$

Furthermore, one has the 'homogeneity property'

$$C_0(\tau) = D_0(\sqrt{\tau}) C_0(1) D_0(\sqrt{\tau}) \qquad \forall\, \tau > 0. \tag{3.1.9}$$

In particular, by combining (3.1.3) with (3.1.8)-(3.1.9), we can write

$$\begin{aligned} \Gamma_\alpha(x,t;0,0) &= \frac{1}{(4\pi\alpha)^{N/2}\sqrt{\det C_0(t)}} e^{-\frac{1}{4\alpha} x^T C_0(t)^{-1} x} \\ &= \frac{1}{(4\pi\alpha)^{N/2} t^{Q/2} \sqrt{\det C_0(1)}} e^{-\frac{1}{4\alpha}\langle C_0(1)^{-1}(D_0(\frac{1}{\sqrt{t}})x),\, D_0(\frac{1}{\sqrt{t}})x\rangle}. \end{aligned} \tag{3.1.10}$$

In [4, Thm.1.7], the next useful comparison result is proved.

Theorem 3.3. *Let Γ be as in Theorem 3.1, and let $\nu > 0$ be as in (1.1.3). Then, for every $s, t \in \mathbb{R}$ with $s < t$, one has the following estimate*

$$\nu\, C_0(t-s)^{-1} \le C(t,s)^{-1} \le \nu^{-1} C_0(t-s)^{-1}, \tag{3.1.11}$$

in the sense of quadratic forms in $\mathbb{R}^N$*. As a consequence, we obtain*

$$\nu^N \Gamma_\nu(x,t;y,s) \le \Gamma(x,t;y,s) \le \frac{1}{\nu^N} \Gamma_{\nu^{-1}}(x,t;y,s), \tag{3.1.12}$$

where Γ_ν is the fundamental solution of the operator $\mathcal{L}_\nu$ in (3.1.6).

3.2. Sharp estimates on the fundamental solution

Taking into account all the results recalled so far, we now aim at proving *sharp Gaussian estimates* for the space derivatives of the fundamental solution Γ of the operator $\mathcal{L}$. As we shall see, these estimates will play a key rôle in our argument.

In order to clearly state our results, we first introduce an *ad-hoc* multi-index notation which shall be useful to deal with differential operators acting *on the $2N$ variables $x, y \in \mathbb{R}^N$*. For a multi-index

$$\boldsymbol{\ell} = (\ell_1, \ldots, \ell_{2N}) \in \mathbb{N}^{2N},$$

let

$$D^{\boldsymbol{\ell}}_{(x,y)} f(x,y) := (\partial_{x_1})^{\ell_1} \cdots (\partial_{x_N})^{\ell_N} (\partial_{y_1})^{\ell_{N+1}} \cdots (\partial_{y_N})^{\ell_{2N}} f(x,y).$$

Moreover, setting $\upsilon = (q_1, \ldots, q_N, q_1, \ldots, q_N) \in \mathbb{R}^{2N}$ (where the q_i's are the exponents appearing in the dilation $D_0(\lambda)$, see (1.1.7)), we define

$$|\boldsymbol{\ell}| := \textstyle\sum_{i=1}^{2N} \ell_i \qquad \text{and} \qquad \omega(\boldsymbol{\ell}) := \sum_{i=1}^{2N} \upsilon_i \ell_i.$$

We will refer to $|\boldsymbol{\ell}|$ and $\omega(\boldsymbol{\ell})$ as, respectively, the *length* and the *order* of $\boldsymbol{\ell}$.

Remark 3.4. Throughout the rest of the paper, we will sometimes need to give a meaning to $\omega(\boldsymbol{\alpha})$ when $\boldsymbol{\alpha}$ is a multi-index in $\mathbb{N}^N$, that is, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_N)$. By analogy, if this is the case we agree to define

$$\omega(\boldsymbol{\alpha}) := \omega(\boldsymbol{\alpha}' = (\boldsymbol{\alpha}, \mathbf{0})) = \textstyle\sum_{i=1}^N \alpha_i q_i.$$

Using the notion of *length*, we can introduce an *order relation* between multi-indexes: if $\boldsymbol{\ell} = (\ell_1, \ldots, \ell_{2N})$, $\boldsymbol{\kappa} = (\kappa_1, \ldots, \kappa_{2N}) \in \mathbb{N}^{2N}$, we say that

$$\boldsymbol{\ell} \prec \boldsymbol{\kappa}$$

if one of the following conditions is satisfied:

(i) $|\boldsymbol{\ell}| < |\boldsymbol{\kappa}|$;
(ii) $|\boldsymbol{\ell}| = |\boldsymbol{\kappa}|$ and $\ell_1 < \kappa_1$;
(iii) $|\boldsymbol{\ell}| = |\boldsymbol{\kappa}|$ and there exists $1 \leq i \leq 2N-1$ such that

$$\ell_1 = \kappa_1, \ldots, \ell_i = \kappa_i \quad \text{and} \quad \ell_{i+1} < \kappa_{i+1}.$$

After all these preliminaries, we can state our first main result.

Theorem 3.5. *Let Γ be as in Theorem 3.1, and let $\nu > 0$ be as in (1.1.3). Moreover, let $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \boldsymbol{\alpha}_2) \in \mathbb{N}^{2N}$ be a fixed multi-index. Then, there exist $c = c(\nu, \boldsymbol{\alpha}) > 0$ and a constant $c_1 > 0$, independent of ν and $\boldsymbol{\alpha}$, such that*

$$\begin{aligned}\left|D^{\boldsymbol{\alpha}}_{(x,y)}\,\Gamma(\xi;\eta)\right| &= \left|D^{\boldsymbol{\alpha}_1}_x D^{\boldsymbol{\alpha}_2}_y \Gamma(\xi;\eta)\right| \\ &\leq \frac{c}{(t-s)^{\omega(\boldsymbol{\alpha})/2}}\,\Gamma_{c_1\nu^{-1}}(\xi;\eta) \\ &\leq \frac{c}{d(\xi,\eta)^{Q+\omega(\boldsymbol{\alpha})}}\end{aligned} \tag{3.2.1}$$

for every $\xi,\eta \in \mathbb{R}^N$ with $t \neq s$. The resulting inequality

$$\left|D^{\boldsymbol{\alpha}}_{(x,y)}\Gamma(\xi;\eta)\right| \leq \frac{c}{d(\xi,\eta)^{Q+\omega(\boldsymbol{\alpha})}}$$

actually holds for every $\xi,\eta \in \mathbb{R}^N$ with $\xi \neq \eta$.

Remark 3.6. Let $(y,s) \in \mathbb{R}^{N+1}$ be fixed. Since we know from Theorem 3.1-(3) that $(\mathcal{L}\Gamma(\cdot;y,s))(x,t) = 0$ for every $x \in \mathbb{R}^N$ and a.e. t, we can express

$$\partial_t(D^{\boldsymbol{\alpha}}_{(x,y)}\,\Gamma) \qquad (\text{for every } \boldsymbol{\alpha} \in \mathbb{N}^{2N})$$

as a combination of quantities that, by (3.2.1) and the exponential decay of the right-hand of this inequality as $t \to s^+$ (with $x \neq y$), are locally essentially bounded in $\mathbb{R}^{N+1} \setminus \{(y,s)\}$. In particular, the same is true of $Y(D^{\boldsymbol{\alpha}}_{(x,y)}\,\Gamma)$.

Before proving Theorem 3.5, we establish the following technical lemma.

Lemma 3.7. *Let $A = (a_{ij})_{i,j=1}^N$ and $B = (b_{ij})_{i,j=1}^N$ be two $N \times N$* symmetric and positive definite matrices *such that, in the sense of quadratic forms, one has*

$$A \leq cB \qquad \textit{for some } c > 0. \tag{3.2.2}$$

Then, denoting by $\|\cdot\|$ the maximum norm of a matrix, we have

$$\|A\| \leq 2c\|B\|. \tag{3.2.3}$$

In particular, if G is any $N \times N$ matrix with real coefficients, then

$$\|G\,A\,G^T\| \leq 2c\|G\,B\,G^T\|. \tag{3.2.4}$$

Proof. First of all, since (3.2.2) holds in the sense of quadratic forms, we have

$$\langle B\xi,\xi\rangle \leq c\,\langle A\xi,\xi\rangle \qquad \forall\,\xi \in \mathbb{R}^N. \tag{3.2.5}$$

As a consequence, choosing $\xi = e_i$ (for $i = 1,\dots,N$) and reminding that both A and B are *positive definite*, we readily have

$$0 < a_{ii} \leq cb_{ii} \leq c\max_{h,k}|b_{hk}| = c\|B\|. \tag{3.2.6}$$

On the other hand, choosing $\xi = e_i \pm e_j$ in (3.2.5) (with $i \neq j$), we get

$$a_{ii} + a_{jj} \pm 2a_{ij} \leq c(b_{ii} + b_{jj} \pm 2b_{ij}) \leq 4c\|B\|;$$

from this, since $a_{ii},\, a_{jj} > 0$, we derive

$$|a_{ij}| \leq 2c\|B\|. \tag{3.2.7}$$

Gathering (3.2.6)-(3.2.7), we immediately obtain (3.2.3). To prove (3.2.4) we observe that, if G is any $N \times N$ matrix, from (3.2.2) it easily follows that

$$GAG^T \leq c\, GBG^T;$$

hence, the desired (3.2.4) is an immediate consequence of (3.2.3). □

Using Lemma 3.7, we can now give the proof of Theorem 3.5.

Proof (of Theorem 3.5). We first observe that, if $\boldsymbol{\alpha} = \mathbf{0}$, estimate (3.2.1) is already contained in Theorem 3.3; hence, we can assume in what follows that

$$\boldsymbol{\alpha} \neq \mathbf{0}.$$

We now fix *once and for all* $s, t \in \mathbb{R}$ satisfying $s < t$ and we notice that, by using the explicit expression of Γ given in (3.1.3), we can write

$$\Gamma(x, t; y, s) = (f_{t,s} \circ p_{t,s})(x, y) \qquad \forall\, x, y \in \mathbb{R}^N, \tag{3.2.8}$$

where the functions $f_{t,s}$ and $p_{t,s}$ are given, respectively, by

$$\begin{aligned} f_{t,s}(z) &= \frac{1}{(4\pi)^{N/2}\sqrt{\det C(t,s)}} e^z \qquad \text{and} \\ p_{t,s}(x,y) &= -\frac{1}{4}\langle C(t,s)^{-1}(x - E(t-s)y), x - E(t-s)y\rangle. \end{aligned}$$

Starting from (3.2.8), and exploiting the multivariate version of the Faà di Bruno formula established in [7, formula (2.1)], we obtain

$$\begin{aligned} D^{\boldsymbol{\alpha}}_{(x,y)}\,\Gamma(x,t;y,s) &= D^{\boldsymbol{\alpha}}_{(x,y)}(f_{t,s} \circ p_{t,s})(x,y) \\ &= \Gamma(x,t;y,s) \cdot \sum_{\lambda=1}^{r} \sum_{m=1}^{r} \sum_{p_m(\lambda,\boldsymbol{\alpha})} \prod_{i=1}^{m} \frac{\boldsymbol{\alpha}!}{k_i!\,(\boldsymbol{\ell}_i!)^{k_i}} \left[D^{\boldsymbol{\ell}_i}_{(x,y)}\, p_{t,s}(x,y)\right]^{k_i}, \end{aligned} \tag{3.2.9}$$

where $r := |\boldsymbol{\alpha}| \geq 1$ and

$$\begin{aligned} p_m(\lambda, \boldsymbol{\alpha}) = \big\{&(k_1, \dots, k_m; \boldsymbol{\ell}_1, \dots, \boldsymbol{\ell}_m) \in \mathbb{N}^m \times (\mathbb{N}^{2N})^m : \, k_i > 0, \\ &\mathbf{0} \prec \boldsymbol{\ell}_1 \cdots \prec \boldsymbol{\ell}_m \text{ and } \textstyle\sum_{i=1}^m k_i = \lambda, \ \sum_{i=1}^m k_i \boldsymbol{\ell}_i = \boldsymbol{\alpha}\big\}. \end{aligned}$$

We now observe that, since the function $p_{t,s}$ is a homogeneous polynomial of degree 2 in the variables x, y, one obviously has

$$D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s} \equiv 0 \qquad \forall\, \boldsymbol{\ell} \in \mathbb{N}^{2N} \text{ with } |\boldsymbol{\ell}| \geq 3;$$

hence, formula (3.2.9) can be rewritten as follows

$$D^{\boldsymbol{\alpha}}_{(x,y)}\,\Gamma(x,t;y,s) = \Gamma(x,t;y,s) \cdot \sum_{(\lambda,m) \in \mathcal{S}} \sum_{p_m(\lambda,\boldsymbol{\alpha})} \prod_{i=1}^{m} \frac{\boldsymbol{\alpha}!}{k_i!\,(\boldsymbol{\ell}_i!)^{k_i}} \left[D^{\boldsymbol{\ell}_i}_{(x,y)}\, p_{t,s}(x,y)\right]^{k_i}, \tag{3.2.10}$$

where $\mathcal{S}$ is the subset of $\{1,\dots,r\}\times\{1,\dots,r\}$ defined as

$$\mathcal{S} := \big\{(\lambda,m) :\, |\boldsymbol{\ell}_i| \leq 2 \text{ for all } (k_1,\dots,k_m;\boldsymbol{\ell}_1,\dots,\boldsymbol{\ell}_m)\in p_m(\lambda,\alpha)\big\}.$$

Then, by combining formula (3.2.10) with the *global pointwise estimates* for Γ contained in Theorem 3.3, for every $x,y\in\mathbb{R}^n$ we obtain

$$|D^{\boldsymbol{\alpha}}_{(x,y)}\,\Gamma(x,t;y,s)| \leq c\,\Gamma_{\nu^{-1}}(x,t;y,s)\cdot \sum_{(\lambda,m)\in\mathcal{S}}\ \sum_{p_m(\lambda,\boldsymbol{\alpha})}\prod_{i=1}^m \big|D^{\boldsymbol{\ell}_i}_{(x,y)}\,p_{t,s}(x,y)\big|^{k_i}, \tag{3.2.11}$$

where $c>0$ is a constant only depending on $\boldsymbol{\alpha}$ and ν. On account of (3.2.11), in order to prove (3.2.1) we need to provide precise estimates for

$$|D^{\boldsymbol{\ell}}_{(x,y)}\,p_{t,s}(x,y)| \qquad (\text{when } 0<|\boldsymbol{\ell}|\leq 2).$$

To this end, we distinguish some different cases. In what follows, we denote by the same c any positive constant which depends only on ν and $\boldsymbol{\alpha}$.

Case I: $\boldsymbol{\ell} = (\boldsymbol{e}_i,\mathbf{0})$. In this case, a direct computation gives

$$D^{\boldsymbol{\ell}}_{(x,y)}\,p_{t,s}(x,y) = \partial_{x_i}p_{t,s}(x,y) = -\frac{1}{2}\big[C(t,s)^{-1}(x-E(t-s)y)\big]_i;$$

hence, setting $v := x-E(t-s)y$ and reminding that

$$\big[D_0(\lambda)v\big]_i = \lambda^{q_i}v_i,$$

we obtain the following chain of inequalities:

$$\begin{aligned}
|D^{\boldsymbol{\ell}}_{(x,y)}\,p_{t,s}(x,y)| &= \frac{1}{2}\big|\big[C(t,s)^{-1}v\big]_i\big| = \frac{c}{(t-s)^{q_i/2}}\big|\big[D_0(\sqrt{t-s})C(t,s)^{-1}v\big]_i\big| \\
&\quad (\text{setting } M(t,s) := D_0(\sqrt{t-s})C(t,s)^{-1}D_0(\sqrt{t-s})) \\
&= \frac{c}{(t-s)^{q_i/2}}\Big|\Big[M(t,s)\cdot D_0\Big(\frac{1}{\sqrt{t-s}}\Big)v\Big]_i\Big| \\
&\leq \frac{c}{(t-s)^{q_i/2}}\|M(t,s)\|\cdot\Big|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)v\Big| =: (\bigstar).
\end{aligned}$$

Now, by combining (3.1.11) with Lemma 3.7, we readily infer that

$$\begin{aligned}
\|M(t,s)\| &\leq 2\nu^{-1}\|D_0(\sqrt{t-s})C_0(t-s)^{-1}D_0(\sqrt{t-s})\| \\
&\quad (\text{see identity (3.1.9)}) \\
&= 2\nu^{-1}\|C_0(1)^{-1}\|;
\end{aligned}$$

as a consequence, we obtain

$$(\bigstar) \leq \frac{c}{(t-s)^{q_i/2}}\,\Big|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(x-E(t-s)y)\Big|.$$

In particular, since $q_i = \omega(\boldsymbol{\ell})$, we conclude that

$$|D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s}(x,y)| \le \frac{c}{(t-s)^{\omega(\boldsymbol{\ell})/2}}\left|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(x-E(t-s)y)\right|. \tag{3.2.12}$$

Case II: $\boldsymbol{\ell} = (\mathbf{0}, \boldsymbol{e}_i)$. In this case, we first rewrite $p_{t,s}$ as follows:

$$\begin{aligned} p_{t,s}(x,y) &= -\frac{1}{4}\langle C(t,s)^{-1}E(t-s)(y-E(s-t)x), E(t-s)(y-E(s-t)x)\rangle \\ &\quad (\text{setting } \widehat{C}(t,s) = E(t-s)^T C(t,s)^{-1}E(t-s)) \\ &= -\frac{1}{4}\langle \widehat{C}(t,s)(y-E(s-t)x), y-E(s-t)x\rangle; \end{aligned} \tag{3.2.13}$$

hence, by proceeding exactly as in Case I, we get

$$|D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s}(x,y)| = |\partial_{y_i} p_{t,s}(x,y)| \le \frac{c}{(t-s)^{q_i/2}}\|\widehat{M}(t,s)\|\cdot\left|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)w\right| =: (\bigstar),$$

where $w := y - E(s-t)x$ and

$$\widehat{M}(t,s) := D_0(\sqrt{t-s})\widehat{C}(t,s)D_0(\sqrt{t-s}).$$

Now, using again (3.1.11) and Lemma 3.7, we get

$$\begin{aligned} \|\widehat{M}(t,s)\| &= \big\|\big[(D_0(\sqrt{t-s})E(t-s)^T\big]\, C(t,s)^{-1}\big[E(t-s)D_0(\sqrt{t-s})\big]\big\| \\ &\le 2\nu^{-1}\big\|\big[(D_0(\sqrt{t-s})E(t-s)^T\big]\, C_0(t-s)^{-1}\big[E(t-s)D_0(\sqrt{t-s})\big]\big\| \\ &\quad (\text{see identities (1.1.9) and (3.1.9)}) \\ &= 2\nu^{-1}\|E(1)^T C_0(1)^{-1}E(1)\|; \end{aligned}$$

as a consequence, we obtain

$$\begin{aligned} (\bigstar) &\le \frac{c}{(t-s)^{q_i/2}}\left|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(y-E(s-t)x)\right| \\ &= \frac{c}{(t-s)^{q_i/2}}\left|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)E(s-t)\cdot(x-E(t-s)y)\right| \\ &\quad (\text{again by (1.1.9)}) \\ &= \frac{c}{(t-s)^{q_i/2}}\left|E(-1)D_0\Big(\frac{1}{\sqrt{t-s}}\Big)\cdot(x-E(t-s)y)\right| \\ &\le \frac{c}{(t-s)^{q_i/2}}\left|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)\cdot(x-E(t-s)y)\right| \end{aligned}$$

In particular, since $q_i = \omega(\boldsymbol{\ell})$, we conclude that

$$|D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s}(x,y)| \le \frac{c}{(t-s)^{\omega(\boldsymbol{\ell})/2}}\left|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(x-E(t-s)y)\right|. \tag{3.2.14}$$

Case III: $\boldsymbol{\ell} = (\boldsymbol{e}_i + \boldsymbol{e}_j, \mathbf{0})$. In this case, a direct computation gives

$$D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s}(x,y) = \partial^2_{x_i x_j} p_{t,s}(x,y) = -\frac{1}{2}C(t,s)^{-1}_{ij};$$

hence, setting $C(t,s)^{-1} := \big(\gamma_{hk}(t,s)\big)_{h,k=1}^N$, we get

$$|D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s}(x,y)| \le c\,|\gamma_{ij}(t,s)|. \tag{3.2.15}$$

Now, taking into account (3.1.11), for every $\varepsilon > 0$ we have

$$\begin{aligned}\gamma_{ii}(t,s) + \varepsilon^2\gamma_{jj}(t,s) \pm 2\varepsilon\gamma_{ij}(t,s) &= \langle C(t,s)^{-1}(e_i \pm \varepsilon e_j), e_i \pm \varepsilon e_j\rangle \\ &\le \nu^{-1}\big(\theta_{ii}(t-s) + \varepsilon^2\theta_{jj}(t-s) \pm 2\varepsilon\theta_{ij}(t-s)\big),\end{aligned}$$

where we have used the notation

$$C_0(\tau)^{-1} = \big(\theta_{hk}(\tau)\big)_{h,k=1}^N.$$

From this, since $C(t,s)^{-1}$ and $C_0(t-s)^{-1}$ are positive definite, we obtain

$$|\gamma_{ij}(t,s)| \le \frac{1}{2\nu}\Big(\frac{1}{\varepsilon}\theta_{ii}(t-s) + \varepsilon\theta_{jj}(t-s) + 2|\theta_{ij}(t-s)|\Big). \tag{3.2.16}$$

To estimate the rhs of (3.2.16) we remind that, by (3.1.9), one has

$$C_0(t-s)^{-1} = D_0\Big(\frac{1}{\sqrt{t-s}}\Big)C_0(1)^{-1}D_0\Big(\frac{1}{\sqrt{t-s}}\Big);$$

as a consequence, we obtain

$$|\theta_{hk}(t-s)| = \frac{\theta_{hk}(1)}{(t-s)^{(q_h+q_k)/2}} \qquad \forall\ 1 \le h,k \le N. \tag{3.2.17}$$

Gathering (3.2.16)-(3.2.17), and choosing $\varepsilon := (t-s)^{(q_j-q_i)/2}$, we then derive

$$|\gamma_{ij}(t,s)| \le \frac{c}{(t-s)^{(q_i+q_j)/2}}. \tag{3.2.18}$$

Finally, since $q_i + q_j = \omega(\boldsymbol{\ell})$, from (3.2.15) and (3.2.18) we conclude that

$$|D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s}(x,y)| \le \frac{c}{(t-s)^{\omega(\boldsymbol{\ell})/2}}. \tag{3.2.19}$$

Case IV: $\boldsymbol{\ell} = (\mathbf{0}, \boldsymbol{e}_i + \boldsymbol{e}_j)$. In this case, using the expression of $p_{t,s}$ given in (3.2.13) (where $\widehat{C}(t,s) = E(t-s)^T C(t,s)^{-1}E(t-s)$), we readily infer that

$$D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s}(x,y) = \partial^2_{y_iy_j} p_{t,s}(x,y) = -\frac{1}{2}\widehat{C}(t,s)_{ij};$$

hence, setting $\widehat{C}(t,s) := \big(\widehat{\gamma}_{hk}(t,s)\big)_{h,k=1}^N$, we get

$$|D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s}(x,y)| \le c\,|\widehat{\gamma}_{ij}(t,s)|. \tag{3.2.20}$$

Now, taking into account (3.1.11), it is easy to see that

$$\begin{aligned}\widehat{C}(t,s) &= E(t-s)^T C(t,s)^{-1}E(t-s) \\ &\le \nu^{-1}E(t-s)^T C_0(t-s)^{-1}E(t-s) \equiv \nu^{-1}\widehat{C}_0(t-s);\end{aligned}$$

from this, by arguing *exactly* as in Case III, for every $\varepsilon > 0$ we obtain

$$|\widehat{\gamma}_{ij}(t,s)| \leq \frac{1}{2\nu}\Big(\frac{1}{\varepsilon}\widehat{\theta}_{ii}(t-s) + \varepsilon\widehat{\theta}_{jj}(t-s) + 2|\widehat{\theta}_{ij}(t-s)|\Big), \tag{3.2.21}$$

where we have used the notation

$$\widehat{C}_0(\tau) = E(\tau)^T C_0(\tau)^{-1} E(\tau) = \big(\widehat{\theta}_{hk}(\tau)\big)_{h,k=1}^N.$$

In order to estimate the rhs of (3.2.21), we observe that

$$\begin{aligned}
\widehat{C}_0(t-s) &= E(t-s)^T C_0(t-s)^{-1} E(t-s) \\
&\quad \text{(see (3.1.9))} \\
&= E(t-s)^T \Big[D_0\Big(\frac{1}{\sqrt{t-s}}\Big) C_0(1)^{-1} D_0\Big(\frac{1}{\sqrt{t-s}}\Big)\Big] E(t-s) \\
&\quad \text{(see (1.1.9))} \\
&= D_0\Big(\frac{1}{\sqrt{t-s}}\Big)\widehat{C}_0(1) D_0\Big(\frac{1}{\sqrt{t-s}}\Big);
\end{aligned}$$

as a consequence, we obtain

$$|\widehat{\theta}_{hk}(t-s)| = \frac{\widehat{\theta}_{hk}(1)}{(t-s)^{(q_h+q_k)/2}} \qquad \forall\, 1 \leq h,k \leq N. \tag{3.2.22}$$

Gathering (3.2.21)-(3.2.22), and choosing $\varepsilon := (t-s)^{(q_j-q_i)/2}$, we then derive

$$|\widehat{\gamma}_{ij}(t,s)| \leq \frac{c}{(t-s)^{(q_i+q_j)/2}}. \tag{3.2.23}$$

Finally, since $q_i + q_j = \omega(\boldsymbol{\ell})$, from (3.2.20) and (3.2.23) we conclude that

$$|D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s}(x,y)| \leq \frac{c}{(t-s)^{\omega(\boldsymbol{\ell})/2}}. \tag{3.2.24}$$

Case V: $\boldsymbol{\ell} = (\boldsymbol{e}_i, \boldsymbol{e}_j)$. In this last case, a direct computation gives

$$\begin{aligned}
D^{\boldsymbol{\ell}}_{(x,y)}\, p_{t,s}(x,y) &= \partial^2_{x_i y_j} p_{t,s}(x,y) = \partial_{y_j}\big(\partial_{x_i} p_{t,s}\big)(x,y) \\
&= \partial_{y_j}\Big(\frac{1}{2}\big[C(t,s)^{-1}(x - E(t-s)y)\big]_i\Big) \\
&= \frac{1}{2}\big[C(t,s)^{-1}E(t-s)\big]_{ij} \\
&= \frac{1}{2}\sum_{k=1}^n \gamma_{ik}(t,s)\, e_{kj}(t-s),
\end{aligned} \tag{3.2.25}$$

where we have used the notation

$$C(t,s) = \big(\gamma_{hk}(t,s)\big)_{h,k=1}^N \qquad \text{and} \qquad E(\tau) = \big(e_{hk}(\tau)\big)_{h,k=1}^N.$$

We now observe that, on account of (1.1.9), we have

$$\begin{aligned}
e_{hk}(t-s) &= \big[E(t-s)\big]_{hk} = \Big[D_0(\sqrt{t-s})E(1)D_0\Big(\frac{1}{\sqrt{t-s}}\Big)\Big]_{hk} \\
&= (t-s)^{(q_h-q_k)/2} e_{hk}(1) \qquad \forall\, 1 \leq h,k \leq N;
\end{aligned} \tag{3.2.26}$$

thus, by combining (3.2.26) with (3.2.18), we get

$$\begin{aligned}\Big|\sum_{k=1}^{n}\gamma_{ik}(t,s)\,e_{kj}(t-s)\Big| &\le \sum_{k=1}^{n}|\gamma_{ik}(t,s)|\,|e_{kj}(t-s)|\\ &\le c\sum_{k=1}^{n}\frac{1}{(t-s)^{(q_i+q_k)/2}}\cdot(t-s)^{(q_k-q_j)/2}e_{kj}(1)\\ &\le \frac{c}{(t-s)^{(q_i+q_j)/2}}.\end{aligned}$$

From this, since $q_i+q_j=\omega(\boldsymbol{\ell})$, we immediately conclude that

$$|D^{\boldsymbol{\ell}}_{(x,y)}\,p_{t,s}(x,y)|\le\frac{1}{2}\Big|\sum_{k=1}^{n}\gamma_{ik}(t,s)\,e_{kj}(t-s)\Big|\le\frac{c}{(t-s)^{\omega(\boldsymbol{\ell})/2}}. \tag{3.2.27}$$

Now we have estimated all the non-vanishing derivatives of $p_{t,s}$ (with respect to both x and y), we are ready to complete the proof. Namely, by combining estimate (3.2.11) with (3.2.12), (3.2.14), (3.2.19), (3.2.24) and (3.2.27), we get

$$\begin{aligned}|D^{\boldsymbol{\alpha}}_{(x,y)}\,\Gamma(x,t;y,s)| &\le c\,\Gamma_{\nu^{-1}}(x,t;y,s)\sum_{(\lambda,m)\in\mathcal{S}}\ \sum_{p_m(\lambda,\boldsymbol{\alpha})}\prod_{i=1}^{m}\frac{1}{(t-s)^{k_i\omega(\boldsymbol{\ell}_i)/2}}\,|v|^{k_i(2-|\boldsymbol{\ell}_i|)}\\ &\le c\,\Gamma(x,t;y,s)\sum_{(\lambda,m)\in\mathcal{S}}\ \sum_{p_m(\lambda,\boldsymbol{\alpha})}\frac{1}{(t-s)^{\sum_{i=1}^{m}k_i\omega(\boldsymbol{\ell}_i)/2}}\,|v|^{\sum_{i=1}^{m}(2k_i-k_i|\boldsymbol{\ell}_i|)},\end{aligned}$$

where we have used the simplified notation

$$v:=D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(x-E(t-s)y).$$

On the other hand, owing to the very definition of $p_m(\lambda,\boldsymbol{\alpha})$, we have

$$\begin{aligned}&\text{(a)}\ \sum_{i=1}^{m}k_i\omega(\boldsymbol{\ell}_i)/2=\frac{1}{2}\,\omega\Big(\sum_{i=1}^{m}k_i\boldsymbol{\ell}_i\Big)=\omega(\boldsymbol{\alpha})/2;\\ &\text{(b)}\ \sum_{i=1}^{m}(2k_i-k_i|\boldsymbol{\ell}_i|)=2\sum_{i=1}^{m}k_i-\Big|\sum_{i=1}^{m}k_i\boldsymbol{\ell}_i\Big|=2\lambda-|\boldsymbol{\alpha}|.\end{aligned}$$

As a consequence, we obtain

$$\begin{aligned}|D^{\boldsymbol{\alpha}}_{(x,y)}\,\Gamma(x,t;y,s)| &\le \frac{c}{(t-s)^{\omega(\boldsymbol{\alpha})/2}}\times\\ &\times\Gamma_{\nu^{-1}}(x,t;y,s)\sum_{(\lambda,m)\in\mathcal{S}}\Big|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(x-E(t-s)y)\Big|^{2\lambda-|\boldsymbol{\alpha}|}.\end{aligned} \tag{3.2.28}$$

We explicitly stress that, if $(\lambda,m)\in\mathcal{S}$, one has $2\lambda-|\boldsymbol{\alpha}|\ge 0$. In fact, taking into account the very definition of $\mathcal{S}$, we know that

$$k_i>0\ \text{and}\ 0<|\boldsymbol{\ell}_i|\le 2\quad\forall\ (k_1,\dots,k_m;\boldsymbol{\ell}_1,\dots,\boldsymbol{\ell}_m)\in p_m(\lambda,\boldsymbol{\alpha});$$

this, together with identity (b), immediately implies that $2\lambda-|\boldsymbol{\alpha}|\ge 0$.

Now, using the explicit expression of Γ_ρ given in Theorem 3.2, together with the fact that the matrix $C_0(1)^{-1}$ is positive definite, we easily see that

$$\Gamma_{\nu^{-1}}(x,t;y,s) \cdot \Big|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(x-E(t-s)y)\Big|^{2\lambda-|\boldsymbol{\alpha}|} \leq c\,\Gamma_{c_1\nu^{-1}}(x,t;y,s), \tag{3.2.29}$$

where $c_1 > 0$ is an absolute constant independent of ν and $\boldsymbol{\alpha}$. Then, by gathering (3.2.28) and (3.2.29), we obtain the first inequality in (3.2.1).

To prove the second inequality in (3.2.1) we will show that for every $\alpha > 0$ and $\omega \geq 0$ there exists a constant $c > 0$ such that, for every $(x,t),(y,s)$ with $t \neq s$ one has

$$\frac{1}{(t-s)^{\omega/2}}\Gamma_\alpha(x,t;y,s) \leq \frac{c}{d((x,t),(y,s))^{\omega+Q}},$$

where $Q > 0$ is the homogeneous dimension of $\mathbb{R}^N$, see (1.1.8).

To this aim, we first observe that, since the matrix $C_0(1)^{-1}$ is (symmetric and) positive definite, by combining (3.1.7) with (3.1.10) we get

$$\begin{aligned}\Gamma_\alpha(x,t;y,s) &= \Gamma_\alpha\big(x-E(t-s)y, t-s; 0,0\big)\\ &\leq \frac{c_0}{(t-s)^{Q/2}}\exp\Big(-c_0\Big|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(x-E(t-s)y)\Big|^2\Big),\end{aligned}$$

where $c_0 > 0$ is a suitable constant depending on α; as a consequence, taking into account the explicit expression of d provided in (1.2.7) (and since $\|\cdot\|$ is D_0-homogeneous of degree 1), we obtain the following estimate

$$\begin{aligned}&\frac{d((x,t),(y,s))^{\omega+Q}}{(t-s)^{\omega/2}}\cdot\Gamma_\alpha(x,t;y,s)\\ &\quad= \frac{\big(\|x-E(t-s)y\|+\sqrt{|t-s|}\big)^{\omega+Q}}{(t-s)^{\omega/2}}\cdot\Gamma_\alpha(x,t;y,s)\\ &\quad= (t-s)^{Q/2}\Big(\Big\|D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(x-E(t-s)y)\Big\|+1\Big)^{\omega+Q}\Gamma_\alpha(x,t;y,s)\\ &\quad\leq c_0\,\mathcal{U}\Big(D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(x-E(t-s)y)\Big),\end{aligned}$$

where we have introduced the notation

$$\mathcal{U}(z) := (\|z\|+1)^{\omega+Q}e^{-c_0|z|^2} \qquad (z\in\mathbb{R}^N).$$

To complete the proof it suffices to show that the function $\mathcal{U}$ is *globally bounded* in $\mathbb{R}^N$. To this end, bearing in mind the explicit definition of $\|\cdot\|$, we notice that

$$\begin{aligned}0 \leq \mathcal{U}(z) &\leq \Big(\sum_{i=1}^N |z|^{1/q_i}+1\Big)^{\omega+Q}e^{-c_0|z|^2}\\ &= \Big[\Big(\sum_{i=1}^N |z|^{1/q_i}+1\Big)e^{-\frac{c_0}{\omega+Q}|z|^2}\Big)\Big]^{\omega+Q};\end{aligned}$$

from this, since the map $\tau \mapsto \tau^\alpha e^{-\beta\tau^2}$ is globally bounded on $[0,+\infty)$ for every choice of $\alpha \geq 0$ and $\beta > 0$, we conclude that $\mathcal{U} \in L^\infty(\mathbb{R}^N)$, as desired.

Finally, combining the two inequalities in (3.2.1) we get

$$\left|D^{\boldsymbol{\alpha}}_{(x,y)}\Gamma(x,t;y,s)\right| \leq \frac{c}{d(x,t;y,s)^{Q+\omega(\boldsymbol{\alpha})}}$$

for every $(x,t),(y,s)$ with $t \neq s$. However, for $x \neq y$ and $s \to t^-$, the first bound in (3.2.1) shows that $D^{\boldsymbol{\alpha}}_{(x,y)}\Gamma(x,t;y,t) = 0$, hence the above inequality actually holds for every $(x,t) \neq (y,s)$, and we are done. □

We highlight a simple consequence of Theorem 3.5 and of (3.1.4) which will be repeatedly exploited in the sequel.

Lemma 3.8. *Let Γ be as in Theorem 3.5, and let $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N) \in \mathbb{N}^N$ be a fixed non-zero multi-index. Then, we have*

$$\int_{\mathbb{R}^N} D^{\boldsymbol{\alpha}}_x\Gamma(x,t;y,s)\,dy = 0 \qquad \text{for every } x \in \mathbb{R}^N \text{ and every } s < t. \tag{3.2.30}$$

Proof. Let x,s,t be as in the statement. Using the global estimates for $D^{\boldsymbol{\alpha}}_x\Gamma$ given in Theorem 3.5, and taking into account identity (3.1.4), we can perform a standard dominated-convergence argument, yielding

$$\int_{\mathbb{R}^N} D^{\boldsymbol{\alpha}}_x\Gamma(x,t;y,s)\,dy = D^{\boldsymbol{\alpha}}_x\Big(x \mapsto \int_{\mathbb{R}^N}\Gamma(x,t;y,s)\,dy\Big) = 0.$$

This ends the proof. □

The next theorem will also be a key tool in our a-priori estimates.

Theorem 3.9 *(Mean value inequality for fractional and singular kernels). Let Γ be as in Theorem 3.1, and let $\eta = (y,s) \in \mathbb{R}^{N+1}$ be fixed. Moreover, let*

$$\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)$$

be a fixed multi-index. Then, there exists a constant $c = c(\boldsymbol{\alpha}) > 0$ such that

$$|D^{\boldsymbol{\alpha}}_x\Gamma(\xi_1,\eta) - D^{\boldsymbol{\alpha}}_x\Gamma(\xi_2,\eta)| \leq c\frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+\omega(\boldsymbol{\alpha})+1}}$$

for every $\xi_1 = (x_1,t_1), \xi_2 = (x_2,t_2) \in \mathbb{R}^{N+1}$ such that

$$d(\xi_1,\eta) \geq 4\boldsymbol{\kappa}d(\xi_1,\xi_2) > 0.$$

Proof. Let $\xi_1, \xi_2 \in \mathbb{R}^{N+1}$ be as in the statement, and let $r := 2d(\xi_1,\xi_2) > 0$. Owing to (1.2.5), one can easily recognize that $\eta \notin \overline{B}_r(\xi_2)$; thus, taking into account the *regularity of* Γ stated in Theorem 3.1-(1) and Remark 3.6, we are entitled to apply Theorem 2.1 to the function $f := D^{\boldsymbol{\alpha}}_x\Gamma(\cdot;\eta)$ on the ball $\overline{B}_r(\xi_2) \ni \xi_1$, obtaining

$$\begin{aligned} |D^{\boldsymbol{\alpha}}_x\Gamma(\xi_1,\eta) - D^{\boldsymbol{\alpha}}_x\Gamma(\xi_2,\eta)| &= |f(\xi_1) - f(\xi_2)| \\ &\leq c\Big(d(\xi_1,\xi_2)\cdot\sup_{B_r(\xi_2)}\sqrt{\sum_{k=1}^q|\partial_{x_k}D^{\boldsymbol{\alpha}}_x\Gamma(\cdot;\eta)|^2} + d(\xi_1,\xi_2)^2\cdot\sup_{B_r(\xi_2)}|YD^{\boldsymbol{\alpha}}_x\Gamma(\cdot;\eta)|\Big). \end{aligned} \tag{3.2.31}$$

Now, since $\xi_1 \in B_r(\xi_2)$ and $q_k = 1$ for $1 \le k \le q$, by Theorem 3.5 we have

$$\sup_{B_r(\xi_2)} \sqrt{\sum_{k=1}^q |\partial_{x_k} D_x^{\boldsymbol{\alpha}} \Gamma(\cdot;\eta)|^2} = \sup_{B_r(\xi_2)} \sqrt{\sum_{k=1}^q |D_x^{\boldsymbol{\alpha}+\boldsymbol{e}_k} \Gamma(\cdot;\eta)|^2} \\ \le c \sup_{\zeta \in B_r(\xi_2)} \frac{1}{d(\zeta,\eta)^{Q+\omega(\boldsymbol{\alpha})+1}} \le \frac{c}{d(\xi_1,\eta)^{Q+\omega(\boldsymbol{\alpha})+1}}. \tag{3.2.32}$$

We then claim that we also have

$$\sup_{B_r(\xi_2)} |Y D_x^{\boldsymbol{\alpha}} \Gamma(\cdot;\eta)| \le \frac{c}{d(\xi_1,\eta)^{Q+\omega(\boldsymbol{\alpha})+2}}. \tag{3.2.33}$$

Taking this claim for granted for a moment, we can conclude the proof of theorem: indeed, by combining (3.2.31), (3.2.32) and (3.2.33) we immediately obtain

$$|D_x^{\boldsymbol{\alpha}} \Gamma(\xi_1,\eta) - D_x^{\boldsymbol{\alpha}} \Gamma(\xi_2,\eta)| \le c d(\xi_1,\xi_2) \Big(\frac{1}{d(\xi_1,\eta)^{Q+\omega(\boldsymbol{\alpha})+1}} + \frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+\omega(\boldsymbol{\alpha})+2}} \Big) \\ (\text{since } d(\xi_1,\eta) \ge 4\boldsymbol{\kappa} d(\xi_1,\xi_2)) \\ \le c \frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+\omega(\boldsymbol{\alpha})+1}},$$

which is exactly what we wanted to prove.

Hence, we are left to prove the claimed (3.2.33). To this end we first notice that, since $\mathcal{L}\Gamma(\cdot;\eta) = 0$ a.e. in $\mathbb{R}^{N+1} \setminus \{\eta\}$, we can write

$$Y D_x^{\boldsymbol{\alpha}} \Gamma(\cdot;\eta) = D_x^{\boldsymbol{\alpha}} \big(Y \Gamma(\cdot;\eta) \big) + [Y, D_x^{\boldsymbol{\alpha}}] \Gamma(\cdot;\eta) \\ = - \sum_{i,j=1}^q a_{ij}(t) D_x^{\boldsymbol{\alpha}+\boldsymbol{e}_i+\boldsymbol{e}_j} \Gamma(\cdot;\eta) + [Y, D_x^{\boldsymbol{\alpha}}] \Gamma(\cdot;\eta), \tag{3.2.34}$$

where $[Y, D_x^{\boldsymbol{\alpha}}] = Y D_x^{\boldsymbol{\alpha}} - D_x^{\boldsymbol{\alpha}} Y$. Moreover, since the coefficients a_{ij} are globally bounded (and $q_k = 1$ for every $1 \le k \le q$), again by Theorem 3.5 we get

$$\left| - \sum_{i,j=1}^q a_{ij}(t) D_x^{\boldsymbol{\alpha}+\boldsymbol{e}_i+\boldsymbol{e}_j} \Gamma(\zeta;\eta) \right| \le \frac{c}{d(\zeta,\eta)^{Q+\omega(\boldsymbol{\alpha})+2}} \qquad \forall\, \zeta \in B_r(\xi_2). \tag{3.2.35}$$

We now turn to estimate the term $[Y, D_x^{\boldsymbol{\alpha}}]\Gamma(\cdot;\eta)$. First of all, using the explicit expression of the vector field Y in (1.1.2), it is easy to see that

$$[Y, D_x^{\boldsymbol{\alpha}}] = Y D_x^{\boldsymbol{\alpha}} - D_x^{\boldsymbol{\alpha}} Y = \sum_{j,k=1}^N b_{jk} \alpha_k D_x^{\boldsymbol{\alpha}+\boldsymbol{e}_j-\boldsymbol{e}_k},$$

where $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)$ and the b_{jk}'s are the entries of the matrix B. On the other hand, taking into account the specific block form of B in assumption (H2), it is not difficult to recognize that

$$q_j - q_k = 2 \qquad \text{for every } 1 \le j,k \le N \text{ such that } b_{jk} \ne 0.$$

As a consequence, using once again Theorem 3.5, we get

$$\begin{aligned} \big|[Y, D_x^{\boldsymbol{\alpha}}]\Gamma(\zeta;\eta)\big| &\leq c \sum_{j,k=1}^{N} |b_{jk}| \cdot \frac{1}{d(\zeta,\eta)^{Q+\omega(\boldsymbol{\alpha})+q_j-q_k}} \\ &\leq \frac{c}{d(\zeta,\eta)^{Q+\omega(\boldsymbol{\alpha})+2}} \qquad \forall\, \zeta \in B_r(\xi_2). \end{aligned} \tag{3.2.36}$$

Finally, by combining (3.2.34), (3.2.35) and (3.2.36) we obtain

$$\sup_{\zeta\in B_r(\xi_2)} |Y D_x^{\boldsymbol{\alpha}}\Gamma(\zeta;\eta)| \leq c \sup_{\zeta\in B_r(\xi_2)} \frac{1}{d(\zeta,\eta)^{Q+\omega(\boldsymbol{\alpha})+2}} \leq \frac{c}{d(\xi_1,\eta)^{Q+\omega(\boldsymbol{\alpha})+2}},$$

which is precisely the claimed (3.2.33). This ends the proof. □

3.3. Representation formulas for u and $\partial^2_{x_ix_j}u$ in terms of $\mathcal{L}u$

We continue to consider an operator $\mathcal{L}$ with coefficients $a_{ij}(t)$ satisfying (H1)-(H2), and its fundamental solution Γ (see Theorem 3.1). Here, we are going to establish some representation formulas for u and for its derivatives in terms of $\mathcal{L}u$.

We start with the following proposition.

Proposition 3.10. *Let $T \in \mathbb{R}$ be fixed, and let $g : S_T \to \mathbb{R}$ be* continuous and bounded*. For every $\varepsilon > 0$, we consider the function*

$$v_\varepsilon : S_T \to \mathbb{R}, \qquad v_\varepsilon(x,t) := \int_{\mathbb{R}^N} \Gamma(x,t;y,t-\varepsilon)\, g(y,t-\varepsilon)\, dy.$$

Then, $v_\varepsilon \to g$ pointwise in S_T *as $\varepsilon \to 0^+$.*

Proof. Let $(x,t) \in S_T$. By combining (3.1.4) with (3.1.12), we can write

$$\begin{aligned} |v_\varepsilon(x,t) - g(x,t)| &= \Big|\int_{\mathbb{R}^N} \Gamma(x,t;y,t-\varepsilon)\big(g(y,t-\varepsilon) - g(x,t)\big)\, dy\Big| \\ &\leq \frac{1}{\nu^N} \int_{\mathbb{R}^N} \Gamma_{\nu^{-1}}(x,t;y,t-\varepsilon) \cdot |g(y,t-\varepsilon) - g(x,t)|\, dy \\ &\leq \frac{c_0}{\varepsilon^{Q/2}} \int_{\mathbb{R}^N} e^{-c_0\, |D_0(\frac{1}{\sqrt{\varepsilon}})(x-E(\varepsilon)y)|^2} \cdot |g(y,t-\varepsilon) - g(x,t)|\, dy = (\bigstar), \end{aligned}$$

where $c_0 > 0$ is a suitable constant only depending on $\nu > 0$. On the other hand, taking into account (1.1.9) and performing the change of variables

$$y = E(-\varepsilon)x - D_0(\sqrt{\varepsilon})E(-1)z,$$

we derive

$$(\bigstar) = c_0 \int_{\mathbb{R}^N} e^{-c_0\, |z|^2} \big|g(E(-\varepsilon)x - D_0(\sqrt{\varepsilon})E(-1)z, t-\varepsilon) - g(x,t)\big|\, dz,$$

since $\det(E(-1)) = 1$. Summing up, we obtain the estimate

$$|v_\varepsilon(x,t) - g(x,t)| \leq c_0 \int_{\mathbb{R}^N} e^{-c_0\,|z|^2} h_\varepsilon(z)\,dz, \tag{3.3.1}$$

where we have introduced the simplified notation

$$h_\varepsilon(z) := \big|g(E(-\varepsilon)x - D_0(\sqrt{\varepsilon})E(-1)z, t-\varepsilon) - g(x,t)\big|.$$

Now, since $E(-\varepsilon) \to E(0) = \mathrm{Id}_N$ and $D_0(\sqrt{\varepsilon}) \to \mathbb{O}_N$ as $\varepsilon \to 0^+$ (see (1.1.7)), from the continuity of g we immediately derive that

$$\lim_{\varepsilon\to 0^+} h_\varepsilon(z) \to 0 \qquad \text{for every fixed } z \in \mathbb{R}^N.$$

Moreover, since g is globally bounded on S_T, we have

$$0 \leq |h_\varepsilon(z)| \leq 2\,\|g\|_{L^\infty(S_T)}.$$

Gathering these two facts, we can apply the dominated convergence theorem in the right-hand side of (3.3.1), yielding

$$|v_\varepsilon(x,t) - g(x,t)| \to 0 \text{ as } \varepsilon \to 0^+.$$

By the arbitrariness of $(x,t) \in S_T$, this completes the proof. □

Thanks to Proposition 3.10, we can now prove the next key result. Throughout the sequel, when dealing with integral over strips we tacitly understand that

$$\int_{\mathbb{R}^N\times(a,b)} \cdots = -\int_{\mathbb{R}^N\times(b,a)} \{\cdots\} \quad \text{when } b < a.$$

Theorem 3.11. *Let $T \in \mathbb{R}$ be fixed, and let $\tau < T$. Moreover, let $u \in \mathcal{S}^0(\tau;T)$. Then, we have the following* representation formula

$$u(x,t) = -\int_{\mathbb{R}^N\times(\tau,t)} \Gamma(x,t;y,s)\mathcal{L}u(y,s)\,dy\,ds, \tag{3.3.2}$$

for every point $(x,t) \in S_T$.

Proof. Since $u \in \mathcal{S}^0(\tau;T)$, then $\mathcal{L}u \in L^\infty(S_T)$. Thus, taking into account (3.1.4) in Theorem 3.1, for every $(x,t) \in S_T$ we get

$$\Bigg|\int_{\mathbb{R}^N\times(\tau,t)} \big|\Gamma(x,t;y,s)\mathcal{L}u(y,s)\big|\,dy\,ds\Bigg| \leq \|\mathcal{L}u\|_{L^\infty(S_T)}\Bigg|\int_\tau^t\bigg(\int_{\mathbb{R}^N}\Gamma(x,t;y,s)\,dy\bigg)ds\Bigg| = |t-\tau| < \infty, \tag{3.3.3}$$

and this proves that the right-hand side of (3.3.2) is *finite*. Now, in order to establish the representation formula (3.3.2), we proceed by steps.

Step I. Let us first prove (3.3.2) by assuming that $u \in \mathcal{S}^0(\tau;T)$ satisfies the following *additional properties*:

(i) $u \in C^\infty(S_T)$;
(ii) there exists $r > 0$ such that

$$u(x,t) = 0 \text{ for every } (x,t) \in S_T \text{ with } |x| > r.$$

Then, owing to (3.3.3) we have

$$\int_{\mathbb{R}^N \times (\tau,t)} \Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds = \lim_{\varepsilon\to 0^+} \int_{\mathbb{R}^N \times (\tau,t-\varepsilon)} \Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds; \tag{3.3.4}$$

moreover, since we are assuming that $u \in C^\infty(S_T)$, we can write

$$\int_{\mathbb{R}^N \times (\tau,t-\varepsilon)} \Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds = \int_\tau^{t-\varepsilon} \Big(\int_{\mathbb{R}^N} \Gamma(x,t;\cdot)\,\mathcal{L}_0 u\,dy \Big) ds - \int_{\mathbb{R}^N} \Big(\int_\tau^{t-\varepsilon} \Gamma(x,t;\cdot)\,\partial_s u\,ds \Big) dy, \tag{3.3.5}$$

where we have written $\mathcal{L} = \mathcal{L}_0 - \partial_s$, that is,

$$\mathcal{L}_0 = \textstyle\sum_{i,j=1}^q a_{ij}(s)\partial_{y_i y_j} + \sum_{j,k=1}^N b_{jk} y_k \partial_{y_j}.$$

Now, owing to Theorem 3.1-(1), we readily see that $y \mapsto \Gamma(x,t;y,s) \in C^\infty(\mathbb{R}^N)$ for every fixed point $(x,t) \in S_T$ and every $s < t - \varepsilon$; as a consequence, taking into account the additional assumptions (i)-(ii), we have

$$\int_{\mathbb{R}^N} \Gamma(x,t;\cdot)\,\mathcal{L}_0 u\,dy = \int_{\mathbb{R}^N} (\mathcal{L}_0)^*\Gamma(x,t;\cdot)\,u\,dy, \tag{3.3.6}$$

where $\mathcal{L}_0^*$ denotes the formal adjoint of $\mathcal{L}_0$, that is,

$$\mathcal{L}_0^* = \textstyle\sum_{i,j=1}^q a_{ij}(s)\partial_{y_i y_j} - \sum_{j,k=1}^N b_{jk} y_k \partial_{y_j}.$$

On the other hand, since from Theorem 3.1-(1) we also derive that $s \mapsto \Gamma(x,t;y,s)$ is Lipschitz-continuous on $(\tau, t-\varepsilon)$, again by (i)-(ii) we have

$$\int_\tau^{t-\varepsilon} \Gamma(x,t;\cdot)\,\partial_s u\,ds = \Gamma(x,t;y,t-\varepsilon)u(y,t-\varepsilon) - \int_\tau^{t-\varepsilon} \partial_s\Gamma(x,t;\cdot)\,u\,ds, \tag{3.3.7}$$

where we have also used the fact that $u \in \mathcal{S}^0(\tau;T)$. Gathering (3.3.6)-(3.3.7), from the above (3.3.5) we then obtain the following identity

$$\begin{aligned}\int_{\mathbb{R}^N \times (\tau,t-\varepsilon)} \Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds &= -\int_{\mathbb{R}^N} \Gamma(x,t;y,t-\varepsilon)u(y,t-\varepsilon)\,dy \\ &\quad + \int_{\mathbb{R}^N \times (\tau,t-\varepsilon)} (\mathcal{L}_0^* + \partial_s)\Gamma(x,t;\cdot)\,u\,dy\,ds.\end{aligned}$$

Then, taking into account (3.3.4), in order to establish formula (3.3.2) it is enough to prove the following fact:

$$\left(\int\limits_{\mathbb{R}^N\times(\tau,t-\varepsilon)} (\mathcal{L}_0^*+\partial_s)\Gamma(x,t;\cdot)\, u\, dy\, ds - \int\limits_{\mathbb{R}^N} \Gamma(x,t;y,t-\varepsilon)u(y,t-\varepsilon)\, dy \right) \\ \to \ -u(x,t) \qquad \text{for every } (x,t)\in S_T \text{ as } \varepsilon\to 0^+. \tag{3.3.8}$$

To this end we first notice that, owing to Theorem 3.1, we have

$$(\mathcal{L}_0^*+\partial_s)\Gamma(x,t;\cdot) = \mathcal{L}^*\Gamma(x,t;\cdot) = 0 \quad \text{a.e. on } \mathbb{R}^N\times(\tau,t-\varepsilon); \tag{3.3.9}$$

moreover, since $u\in\mathcal{S}^0(\tau;T)$ (hence, in particular, u is continuous and bounded on the strip S_T), from Proposition 3.10 we infer that

$$\lim_{\varepsilon\to 0^+} \int\limits_{\mathbb{R}^N} \Gamma(x,t;y,t-\varepsilon)u(y,t-\varepsilon)\, dy = u(x,t) \quad \text{pointwise on } S_T. \tag{3.3.10}$$

By combining (3.3.9)-(3.3.10), we immediately obtain (3.3.8).

Step II. Let us now prove the representation formula (3.3.2) by dropping the additional assumption (ii) on u, that is, we only suppose that

$$u\in\mathcal{S}^0(\tau;T)\cap C^\infty(S_T).$$

To begin with, we fix a cut-off function $\phi_0\in C_0^\infty(\mathbb{R}^N)$ such that

(a) $0\le\phi_0\le 1$ in $\mathbb{R}^N$;
(b) $\phi_0\equiv 1$ on $\{|x|<1\}$ and $\phi_0\equiv 0$ on $\{|x|>2\}$.

Moreover, for every $n\ge 1$ we set $\phi_n(x):=\phi_0(x/n)$, and we define

$$u_n := u\cdot\phi_n.$$

Owing to (a)-(b), it is readily seen that $u_n\in\mathcal{S}^0(\tau;T)\cap C^\infty(S_T)$ and $u_n(x,t)=0$ for every $(x,t)\in S_T$ with $|x|>n$; hence, by Step I we can write

$$u_n(x,t) = -\int\limits_{\mathbb{R}^N\times(\tau,t)} \Gamma(x,t;\cdot)\,\mathcal{L}u_n\, dy\, ds \quad \text{for every } (x,t)\in S_T. \tag{3.3.11}$$

We now aim to pass to the limit as $n\to\infty$ in the above (3.3.11). By definition of ϕ_n, we have

$$\lim_{n\to\infty} u_n(x,t) = u(x,t) \quad \text{for every fixed } (x,t)\in S_T. \tag{3.3.12}$$

As to the right-hand side, instead, we rely on the dominated convergence theorem. First of all, since $u\in C^\infty(S_T)$ and $\phi_n\in C_0^\infty(\mathbb{R}^N)$, we have

$$\mathcal{L}u_n = \mathcal{L}(u\phi_n) = (\mathcal{L}u)\cdot\phi_n + u\cdot(\mathcal{L}\phi_n) + 2\sum_{i,j=1}^q a_{ij}(t)\partial_{x_i}u\,\partial_{x_j}\phi_n;$$

moreover, since $u\in\mathcal{S}^0(\tau;T)$ and $\phi_n=\phi_0(\cdot/n)$, there exists a constant $\mathbf{c}>0$, depending on u and ϕ_0 but independent of n, such that

$$\Big| u \cdot (\mathcal{L}\phi_n) + 2\sum_{i,j=1}^{q} a_{ij}(t)\partial_{x_i} u\,\partial_{x_j}\phi_n \Big| \leq \frac{\mathbf{c}}{n} \qquad \text{pointwise on } S_T.$$

This, together with the fact that $\phi_n \equiv 1$ on $\{|x| < n\}$, implies

$$\lim_{n\to\infty} \mathcal{L}u_n = \mathcal{L}u \qquad \text{pointwise on } S_T.$$

On the other hand, since $\mathcal{L}u \in L^\infty(S_T)$ and $0 \leq \phi_n \leq 1$, we also have

$$|\mathcal{L}u_n| \leq \|\mathcal{L}u\|_{L^\infty(S_T)} + \frac{\mathbf{c}}{n} \leq \|\mathcal{L}u\|_{L^\infty(S_T)} + \mathbf{c} =: \mathbf{c}' \qquad \text{for every } n \geq 1;$$

gathering these facts, and taking into account (3.1.4), we can then apply the dominated convergence theorem in the right-hand side of (3.3.11), yielding

$$\lim_{n\to\infty} \int\limits_{\mathbb{R}^N\times(\tau,t)} \Gamma(x,t;\cdot)\,\mathcal{L}u_n\,dy\,ds = \int\limits_{\mathbb{R}^N\times(\tau,t)} \Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds. \tag{3.3.13}$$

Finally, by combining (3.3.12) and (3.3.13) we can let $n \to \infty$ in (3.3.11), thus obtaining the desired representation formula (3.3.2) for u.

Step III: Let us finally prove the representation formula (3.3.2) for every $u \in \mathcal{S}^0(\tau;T)$.

To begin with, we fix a point $\xi_0 = (x_0,t_0) \in S_T$ and we choose $0 < \varepsilon_0 < 1$ in such a way that $\xi_0 \in S_{T-\varepsilon_0}$. Moreover, we choose a function $J \in C_0^\infty(\mathbb{R}^{N+1})$ such that $J \geq 0$ pointwise in $\mathbb{R}^{N+1}$, $\mathrm{supp}(J) \subseteq B_1(0)$ and

$$\int\limits_{\mathbb{R}^{N+1}} J(\eta)\,d\eta = \int\limits_{B_1} J(\eta)\,d\eta = 1, \tag{3.3.14}$$

where $B_1(0) = \{\eta : d(\eta,0) < 1\}$ is the d-ball with centre 0 and radius 1. We then define, for every fixed $0 < \varepsilon < \varepsilon_0$, the $(\varepsilon,\mathbb{G})$-*convolution kernel*

$$J_\varepsilon(\eta) := \varepsilon^{-Q-2} J\big(D(1/\varepsilon)\eta\big)$$

(where $D(\cdot)$ and $Q > 0$ are as in (1.1.7) and (1.1.8), respectively), and we consider the so-called *mollifier of u related to the kernel J_ε*, that is,

$$\begin{aligned}
&u_\varepsilon : S_{T-\varepsilon_0} \to \mathbb{R},\\
&u_\varepsilon(\xi) := \int\limits_{S_T} J_\varepsilon(\xi\circ\eta^{-1})\,u(\eta)\,d\eta = \int\limits_{B_1(0)} J(\zeta)\,u\big((D(\varepsilon)\zeta^{-1})\circ\xi\big)\,d\zeta.
\end{aligned}$$

We explicitly point out, for the sake of completeness, that the definition of u_ε is *meaningful*: in fact, using (1.1.6), (1.1.7) and (1.2.4) we easily see that

(a) for every fixed $\xi = (x,t) \in S_{T-\varepsilon_0}$, one has

$$\mathrm{supp}\big(\eta \mapsto J_\varepsilon(\xi\circ\eta^{-1})\big) \subseteq \{\eta = (y,s) : |t-s| < \varepsilon\} \subseteq S_T; \tag{3.3.15}$$

(b) for every $\zeta \in B_1(0)$ and $\xi \in S_{T-\varepsilon_0}$, one has $(D(\varepsilon)\zeta^{-1})\circ\xi \in S_T$.

We now claim that:

$$u_\varepsilon \in \mathcal{S}^0(\tau-\varepsilon_0; T-\varepsilon_0) \cap C^\infty(S_{T-\varepsilon_0}). \tag{3.3.16}$$

Indeed, since $J \in C_0^\infty(\mathbb{R}^{N+1})$, by a standard dominated-convergence argument we easily infer that $u_\varepsilon \in C^\infty(S_{T-\varepsilon_0})$; moreover, taking into account that $u(x,t) \equiv 0$ for every $(x,t) \in S_T$ with $t \le \tau$, by (3.3.15) we derive that

$$u_\varepsilon(\xi) = \int_{\{|t-s|<\varepsilon\}} J_\varepsilon((x,t)\circ(y,s)^{-1})\, u(y,s)\, dy\, ds = 0$$

for every $\xi = (x,t) \in S_{T-\varepsilon_0}$ with $t \le \tau - \varepsilon_0$. Hence, to prove the claimed (3.3.16) we are left to show that the derivatives $\partial^2_{x_ix_j} u_\varepsilon, Yu_\varepsilon$, which exist pointwise and in the classical sense on $S_{T-\varepsilon_0}$, are globally bounded in $S_{T-\varepsilon_0}$ (for $1 \le i,j \le q$).

To this end it suffices to observe that, since $u \in \mathcal{S}^0(\tau;T)$ and since the vector fields $\partial_{x_1},\ldots,\partial_{x_q}, Y$ are left-invariant with respect to $\circ$, we can write

$$\begin{aligned} \partial^2_{x_ix_j} u_\varepsilon\,(\xi) &= \int_{B_1(0)} J(\zeta)\,(\partial^2_{x_ix_j}u)\big(D(\varepsilon)\zeta^{-1})\circ\xi\big)\, d\zeta \qquad (\text{for } i=1,\ldots,q), \\ Yu_\varepsilon\,(\xi) &= \int_{B_1(0)} J(\zeta)\,(Yu)\big(D(\varepsilon)\zeta^{-1})\circ\xi\big)\, d\zeta, \end{aligned} \tag{3.3.17}$$

thus, since $\partial^2_{x_ix_j}u, Yu \in L^\infty(S_T)$, from (3.3.14) we obtain

$$\begin{aligned} \|\partial^2_{x_ix_j}u_\varepsilon\|_{L^\infty(S_{T-\varepsilon_0})} &\le \|\partial^2_{x_ix_j}u\|_{L^\infty(S_T)} \quad (\text{for } 1\le i,j\le q); \\ \|Yu_\varepsilon\|_{L^\infty(S_{T-\varepsilon_0})} &\le \|Yu\|_{L^\infty(S_T)}, \end{aligned} \tag{3.3.18}$$

and this completes the proof of (3.3.16).

Now we have established (3.3.16), thanks to Step II we know that the representation formula (3.3.2) holds for the function u_ε on the strip $S_{T-\varepsilon_0}$: in particular, since we have that $\xi_0 = (x_0,t_0) \in S_{T-\varepsilon_0}$, we can write

$$u_\varepsilon(x_0,t_0) = -\int_{\mathbb{R}^N\times(\tau-\varepsilon_0,t_0)} \Gamma(x_0,t_0;\cdot)\,\mathcal{L}u_\varepsilon\, dy\, ds. \tag{3.3.19}$$

We then pass to the limit as $\varepsilon \to 0^+$ in (3.3.19). As to the left-hand, since u is continuous and bounded on S_T, it is easily seen that

$$\lim_{\varepsilon\to 0^+} u_\varepsilon(x_0,t_0) = u(x_0,t_0). \tag{3.3.20}$$

As to the right-hand side, taking into account (3.3.17) and the fact that

$$\partial^2_{x_ix_j}u, Yu \in L^\infty(S_T),$$

we can use a classical approximation argument to prove that $\partial^2_{x_ix_j}u_\varepsilon \to \partial^2_{x_ix_j}u$ (for every $1\le i,j\le q$) and $Yu_\varepsilon \to Y_u$ in $L^1_{\text{loc}}(S_{T-\varepsilon_0})$ as $\varepsilon \to 0^+$; as a consequence, by possibly choosing a sequence $\varepsilon_n \to 0$ as $n\to\infty$, we get

$$\lim_{\varepsilon\to 0^+}\mathcal{L}u_\varepsilon = \lim_{\varepsilon\to 0^+}\left(\sum_{i,j=1}^q a_{ij}(\cdot)\partial^2_{x_ix_j}u_\varepsilon + Yu_\varepsilon\right) = \mathcal{L}u \quad \text{a.e. in } S_{T-\varepsilon_0}.$$

On the other hand, using (3.3.18) and the fact that the coefficients a_{ij} are globally bounded, we also have the following estimate

$$\begin{aligned}|\mathcal{L}u_\varepsilon| &\le \sum_{i,j=1}^q \|a_{ij}\|_{L^\infty(\mathbb{R})}\cdot\|\partial^2_{x_ix_j}u\|_{L^\infty(S_T)} + \|Yu\|_{L^\infty(S_T)} \\ &=: \mathbf{c}, \qquad \text{for every } 0<\varepsilon<\varepsilon_0.\end{aligned}$$

Gathering these facts, and recalling that $J\in C_0^\infty(\mathbb{R}^{N+1})$, we can then apply the dominated convergence theorem in the right-hand side of (3.3.19), getting

$$\lim_{\varepsilon\to 0^+}\int_{\mathbb{R}^N\times(\tau-\varepsilon_0,t_0)}\Gamma(x_0,t_0;\cdot)\,\mathcal{L}u_\varepsilon\,dy\,ds = \int_{\mathbb{R}^N\times(\tau-\varepsilon_0,t_0)}\Gamma(x_0,t_0;\cdot)\,\mathcal{L}u\,dy\,ds \tag{3.3.21}$$

Finally, by combining (3.3.20)-(3.3.21) and by taking into account that

$$u = \mathcal{L}u \equiv 0 \text{ a.e. on } \mathbb{R}^N\times(-\infty,\tau),$$

we can pass to the limit as $\varepsilon\to 0^+$ in (3.3.19), thus obtaining the desired representation formula (3.3.2) for u. This completes the proof. □

Starting from the representation formula (3.3.2), we easily obtain the following representation formula for the *first-order derivatives* of u.

Corollary 3.12. *Let $T\in\mathbb{R}$ be fixed, and let $\tau<T$. Moreover, let $u\in\mathcal{S}^0(\tau;T)$ and let $1\le i\le q$ be fixed. Then, we have the representation formula*

$$\partial_{x_i}u(x,t) = -\int_{\mathbb{R}^N\times(\tau,t)}\partial_{x_i}\Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds, \tag{3.3.22}$$

for every $(x,t)\in S_T$. Moreover,

$$\|\partial_{x_i}u\|_{L^\infty(S_T)} \le c\,\|\mathcal{L}u\|_{L^\infty(S_T)}\cdot\sqrt{T-\tau}. \tag{3.3.23}$$

Proof. We start noting that, combining the global estimates for $\partial_{x_i}\Gamma$ contained in Theorem 3.5, see (3.2.1), with identity (3.1.4), we have, for every $x\in\mathbb{R}^N$ and every $\tau<t$,

$$\begin{aligned}\int_{\mathbb{R}^N\times(\tau,t)}|\partial_{x_i}\Gamma(x,t;\cdot)|\,dy\,ds &\le c\int_\tau^t\frac{1}{\sqrt{t-s}}\left(\int_{\mathbb{R}^N}\Gamma_{c_1\nu^{-1}}(x,t;\cdot)\,dy\right)ds \\ &= c\int_\tau^t\frac{1}{\sqrt{t-s}}\,ds = 2c\sqrt{t-\tau}.\end{aligned} \tag{3.3.24}$$

Let us now prove formula (3.3.22). To begin with, since $u\in\mathcal{S}^0(\tau;T)$, we have $\mathcal{L}u\in L^\infty(S_T)$; thus, from (3.3.24) we get

$$\left| \int\limits_{\mathbb{R}^N\times(\tau,t)} |\partial_{x_i}\Gamma(x,t;\cdot)|\,|\mathcal{L}u| dy\,ds \right| \leq c\,\|\mathcal{L}u\|_{L^\infty(S_T)}\cdot\sqrt{|t-\tau|} \qquad \forall\ (x,t)\in S_T \tag{3.3.25}$$

(where $c > 0$ only depends on ν), and this shows that the function

$$g(x,t) := -\int\limits_{\mathbb{R}^N\times(\tau,t)} \partial_{x_i}\Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds,$$

is well-defined on S_T. We then turn to prove that $\partial_{x_i}u \equiv g$ pointwise in S_T by an approximation argument. To this end, we fix $0 < \varepsilon \ll 1$ and we define

$$u_\varepsilon(x,t) := -\int\limits_{\mathbb{R}^N\times(\tau,t-\varepsilon)} \Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds.$$

Owing to the representation formula (3.3.2), it is readily seen that $u_\varepsilon \to u$ pointwise on S_T as $\varepsilon \to 0^+$; moreover, since $t-s \geq \varepsilon > 0$ when $s < t-\varepsilon$, by simple dominated-convergence arguments based on (3.2.1) (and on the regularity of Γ, see Theorem 3.1-(1)) we easily infer that

(i) $u_\varepsilon \in C(S_T)$;
(ii) u_ε is continuously differentiable w.r.t. x_i on S_T, and

$$\partial_{x_i}u_\varepsilon(x,t) = -\int\limits_{\mathbb{R}^N\times(\tau,t-\varepsilon)} \partial_{x_i}\Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds \quad \forall\ (x,t)\in S_T.$$

Finally, by (3.3.24) we also have

$$\begin{aligned} |\partial_{x_i}u_\varepsilon(x,t) - g(x,t)| &= \int\limits_{\mathbb{R}^N\times(t-\varepsilon,t)} |\partial_{x_i}\Gamma(x,t;\cdot)|\,\|\mathcal{L}u\|_{L^\infty(S_T)}\,dy\,ds \\ &\leq c\,\|\mathcal{L}u\|_{L^\infty(S_T)}\sqrt{\varepsilon} \qquad \text{uniformly for } (x,t)\in S_T, \end{aligned}$$

from which we derive that $\partial_{x_i}u_\varepsilon \to g$ uniformly on S_T as $\varepsilon \to 0^+$. As is well-know, all the above facts are enough to conclude that

$$\partial_{x_i}u \equiv g \text{ on } S_T,$$

and this is precisely (3.3.22). By (3.3.25), this also implies (3.3.23). □

With the representation formula (3.3.22) at hand, we now aim to prove a representation formula for the *derivatives* $\partial_{x_ix_j}u$ of a function $u \in \mathcal{S}^0(\tau;T)$.

To this end, we first establish the following proposition.

Proposition 3.13. *Let $\alpha \in (0,1)$ be fixed, and let $1 \leq i,j \leq q$. Then, there exists a constant $c = c(\alpha) > 0$ such that, for every $x \in \mathbb{R}^N$ and every $\tau < t$, one has*

$$\int\limits_{\mathbb{R}^N\times(\tau,t)} |\partial^2_{x_ix_j}\Gamma(x,t;y,s)|\cdot\|E(s-t)x-y\|^\alpha\,dy\,ds \leq c(t-\tau)^{\alpha/2}. \tag{3.3.26}$$

As a consequence, we have

$$\int_{\mathbb{R}^N\times(t-\varepsilon,t)} |\partial^2_{x_ix_j}\Gamma(x,t;y,s)|\cdot\|E(s-t)x-y\|^\alpha\,dy\,ds\to 0$$
$$\text{uniformly w.r.t. } (x,t)\in\mathbb{R}^{N+1}\ \textit{as}\ \varepsilon\to 0^+. \tag{3.3.27}$$

Proof. Let x,τ,t be as in the statement. Owing to the global estimates for $\partial^2_{x_ix_j}\Gamma$ in Theorem 3.5, see (3.2.1), and taking into account (3.1.7)-(3.1.10), we have

$$\begin{aligned}|\partial^2_{x_ix_j}\Gamma(x,t;y,s)| &\le \frac{c}{t-s}\Gamma_{c_1\nu^{-1}}(x,t;y,s)\\ &\le \frac{c_0}{(t-s)^{Q/2+1}}e^{-c_0|D_0(\frac{1}{\sqrt{t-s}})(x-E(t-s)y)|^2},\end{aligned} \tag{3.3.28}$$

where $c_0>0$ is a suitable constant only depending on the number $\nu>0$. On the other hand, taking into account (1.1.9), for every $s<t$ we can write

$$\begin{aligned}D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(x-E(t-s)y) &= \Big[D_0\Big(\frac{1}{\sqrt{t-s}}\Big)E(t-s)\Big](E(s-t)x-y)\\ &= E(1)\Big[D_0\Big(\frac{1}{\sqrt{t-s}}\Big)(E(s-t)x-y)\Big].\end{aligned}$$

As a consequence, since $E(1)$ is non-singular, we get

$$e^{-c_0|D_0(\frac{1}{\sqrt{t-s}})(x-E(t-s)y)|^2}\le e^{-c_0'|D_0(\frac{1}{\sqrt{t-s}})(E(s-t)x-y)|^2}, \tag{3.3.29}$$

where $c_0'>0$ is another constant only depending on ν. Then, by combining (3.3.28) with (3.3.29), we obtain the following estimate:

$$\begin{aligned}&\int_{\mathbb{R}^N\times(\tau,t)} |\partial^2_{x_ix_j}\Gamma(x,t;y,s)|\cdot\|E(s-t)x-y\|^\alpha\,dy\,ds\\ &\le c_0\int_{\mathbb{R}^N\times(\tau,t)}\frac{1}{(t-s)^{Q/2+1}}e^{-c_0'|D_0(\frac{1}{\sqrt{t-s}})(E(s-t)x-y)|^2}\|E(s-t)x-y\|^\alpha\,dy\,ds\\ &= c_0\int_\tau^t\frac{1}{(t-s)^{Q/2+1}}\Big(\int_{\mathbb{R}^N}e^{-c_0'|D_0(\frac{1}{\sqrt{t-s}})(E(s-t)x-y)|^2}\|E(s-t)x-y\|^\alpha\,dy\Big)ds\\ &=: (\bigstar).\end{aligned}$$

To proceed further, we perform in the dy-integral the change of variables

$$y=E(s-t)x-D_0(\sqrt{t-s})z.$$

Reminding that $\det(D_0(\lambda))=\lambda^Q$ for every $\lambda>0$ (see (1.1.7)-(1.1.8)), and since the norm $\|\cdot\|$ is D_0-homogeneous of degree 1, we get

$$\begin{aligned}(\bigstar) &= c_0\int_\tau^t\frac{1}{(t-s)^{1-\frac{\alpha}{2}}}\Big(\int_{\mathbb{R}^N}e^{-c_0'|z|^2}\|z\|^\alpha\,dz\Big)ds\\ &= \frac{c_0}{\alpha}(t-\tau)^{\alpha/2}\Big(\int_{\mathbb{R}^N}e^{-c_0'|z|^2}\|z\|^\alpha\,dz\Big).\end{aligned}$$

To complete the proof of (3.3.26) we only need to show that the dz-integral is finite. To this end we observe that, by definition of $\|\cdot\|$, we have

$$I := \int_{\mathbb{R}^N} e^{-c_0'|z|^2}\|z\|^\alpha\,dz = \int_{\mathbb{R}^N} e^{-c_0'|z|^2}\Big(\sum_{i=1}^N |z_i|^{1/q_i}\Big)^\alpha dz$$
$$\leq c(\alpha)\sum_{i=1}^N \int_{\mathbb{R}^N} e^{-c_0'|z|^2}|z_i|^{\alpha/q_i}\,dz \leq c(\alpha)\sum_{i=1}^N \int_{\mathbb{R}^N} e^{-c_0'|z|^2}|z|^{\alpha/q_i}\,dz;$$

from this, we immediately see that $I < \infty$, and the proof is complete. $\square$

With Proposition 3.13 at hand, we can now prove the following theorem.

Theorem 3.14. *For $T > \tau > -\infty$ and $\alpha \in (0,1)$, let $u \in \mathcal{S}^0(\tau;T)$ be such that $\mathcal{L}u \in C_x^\alpha(S_T)$. Then, we have*

$$\partial^2_{x_ix_j}u(x,t) = \int_{\mathbb{R}^N\times(\tau,t)} \partial^2_{x_ix_j}\Gamma(x,t;y,s)\big[\mathcal{L}u(E(s-t)x,s) - \mathcal{L}u(y,s)\big]\,dy\,ds, \tag{3.3.30}$$

for every $(x,t) \in S_T$ and every $1 \leq i,j \leq q$.

Proof. We first observe that, since $\mathcal{L}u \in C_x^\alpha(S_T)$, by definition we have

$$|\mathcal{L}u(E(s-t)x,s) - \mathcal{L}u(y,s)| \leq |\mathcal{L}u|_{C_x^\alpha(S_T)}\cdot\|E(s-t)x - y\|^\alpha, \tag{3.3.31}$$

for every $x, y \in \mathbb{R}^N$ and every $s,t < T$. Thus, by Proposition 3.13 we get

$$\Big|\int_{\mathbb{R}^N\times(\tau,t)} |\partial^2_{x_ix_j}\Gamma(x,t;y,s)|\cdot|\mathcal{L}u(E(s-t)x,s) - \mathcal{L}u(y,s)|\,dy\,ds\Big|$$
$$\leq |\mathcal{L}u|_{C_x^\alpha(S_T)}\Big|\int_{\mathbb{R}^N\times(\tau,t)} |\partial^2_{x_ix_j}\Gamma(x,t;y,s)|\cdot\|E(s-t)x-y\|^\alpha\,dy\,ds\Big|$$
$$\leq \mathbf{c}\,|\mathcal{L}u|_{C_x^\alpha(S_T)}\cdot|t-\tau|^{\alpha/2}\quad\forall\,(x,t)\in S_T$$

(where $\mathbf{c} > 0$ only depends on α), and this shows that the function

$$g(x,t) := \int_{\mathbb{R}^N\times(\tau,t)} \partial^2_{x_ix_j}\Gamma(x,t;y,s)\big[\mathcal{L}u(E(s-t)x,s) - \mathcal{L}u(y,s)\big]\,dy\,ds$$

is well-defined on S_T. We then turn to prove that $\partial^2_{x_ix_j}u = g$ pointwise in S_T by an approximation argument. To this end, we fix $0 < \varepsilon \ll 1$ and we define

$$v_\varepsilon(x,t) := -\int_{\mathbb{R}^N\times(\tau,t-\varepsilon)} \partial_{x_j}\Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds.$$

Now, arguing as in the proof of Corollary 3.12 and taking into account (3.3.22), we see that

(i) $v_\varepsilon \in C(S_T)$ and $v_\varepsilon \to \partial_{x_j}u$ pointwise in S_T as $\varepsilon \to 0^+$;

(ii) v_ε is continuously differentiable w.r.t. x_i on S_T, and

$$\partial_{x_i} v_\varepsilon(x,t) = -\int_{\mathbb{R}^N\times(\tau,t-\varepsilon)} \partial^2_{x_i x_j}\Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds \quad \forall\ (x,t)\in S_T.$$

On the other hand, owing to Lemma 3.8, we have

$$\begin{aligned}\partial_{x_i} v_\varepsilon(x,t) &= -\int_{\mathbb{R}^N\times(\tau,t-\varepsilon)} \partial^2_{x_i x_j}\Gamma(x,t;\cdot)\,\mathcal{L}u\,dy\,ds \\ &= \int_{\mathbb{R}^N\times(\tau,t-\varepsilon)} \partial^2_{x_i x_j}\Gamma(x,t;y,s)\big[\mathcal{L}u(E(s-t)x,s)-\mathcal{L}u(y,s)\big]\,dy\,ds.\end{aligned}$$

As a consequence, by combining (3.3.31) with Proposition 3.13 we obtain

$$\begin{aligned}|\partial_{x_i} v_\varepsilon(x,t) - g(x,t)| &= \int_{\mathbb{R}^N\times(t-\varepsilon,t)} |\partial^2_{x_i x_j}\Gamma(x,t;\cdot)|\,|\mathcal{L}u(E(s-t)x,s)-\mathcal{L}u|\,dy\,ds \\ &\le |\mathcal{L}u|_{C^\alpha_x(S_T)} \int_{\mathbb{R}^N\times(t-\varepsilon,t)} |\partial^2_{x_i x_j}\Gamma(x,t;\cdot)|\ \cdot\|E(s-t)x-y\|^\alpha\,dy\,ds \\ &\le c\,|\mathcal{L}u|_{C^\alpha_x(S_T)}\varepsilon^{\alpha/2} \qquad \text{uniformly for } (x,t)\in S_T,\end{aligned}$$

from which we derive that $\partial_{x_i} v_\varepsilon \to g$ uniformly on S_T as $\varepsilon\to 0^+$. As in the proof of Corollary 3.12 we then conclude that

$$\partial^2_{x_i x_j} u = \partial_{x_i}(\partial_{x_j} u) = g \quad \text{pointwise in } S_T,$$

and this gives (3.3.30). □

3.4. Schauder estimates in space

We now want to prove the following result:

Theorem 3.15 *(Global Schauder estimates in space). Let $T>\tau>-\infty$ and $\alpha\in(0,1)$. Then, there exists $c>0$, only depending on $(T-\tau),\alpha,\nu,B$, such that*

$$\sum_{i,j=1}^q \|\partial^2_{x_i x_j} u\|_{C^\alpha_x(S_T)} \le c\,|\mathcal{L}u|_{C^\alpha_x(S_T)} \tag{3.4.1}$$

$$\|Yu\|_{C^\alpha_x(S_T)} \le c\,\|\mathcal{L}u\|_{C^\alpha_x(S_T)}, \tag{3.4.2}$$

for every $u\in\mathcal{S}^0(\tau;T)$ with $\mathcal{L}u\in C^\alpha_x(S_T)$.

The estimates in the above theorem will be generalized, in Section 4, in the context of operators with coefficients $a_{ij}(x,t)$; hence, the core of this section consists more in the development of the *tools* necessary to prove the above theorem, then in the result itself. Actually, these tools will be useful also in the following parts of the paper. Also, it is worth noting that the proof of global Schauder estimates in the situation considered in this section is much more straightforward than for coefficients also depending on x. However, note that for the moment we do not prove global estimates on the lower order derivatives $\partial_{x_k} u$ and on u itself.

To prove Theorem 3.15, we need the following auxiliary results.

Theorem 3.16 *(Cancellation property of the singular kernel). There exists a constant $c > 0$ such that, for every $1 \le i, j \le q$, one has the estimate*

$$I_{r,\tau}(x,t) := \int_\tau^t \Big| \int_{\{y \in \mathbb{R}^N :\, d((x,t),(y,s)) \ge r\}} \partial^2_{x_i x_j} \Gamma(x,t;y,s)\, dy \Big|\, ds \le c, \tag{3.4.3}$$

for every $x \in \mathbb{R}^N$, $\tau < t$ and $r > 0$.

Proof. Let x, τ, t and r be as in the statement. We then distinguish two cases.

Case I: $t - \tau > r^2$. In this case we first observe that, taking into account the explicit expression of the quasi-distance d given in (1.2.7), we have

$$d((x,t),(y,s)) = \|x - E(t-s)y\| + \sqrt{t-s} \ge \sqrt{t-s} \ge r,$$

for every $\tau < s < t - r^2$; thus, by Lemma 3.8 we can write

$$\begin{aligned} I_{r,\tau}(x,t) &= \int_\tau^{t-r^2} \Big| \int_{\mathbb{R}^N} \partial^2_{x_i x_j} \Gamma(x,t;y,s)\, dy \Big|\, ds \\ &\quad + \int_{t-r^2}^t \Big| \int_{\{y \in \mathbb{R}^N :\, d((x,t),(y,s)) \ge r\}} \partial^2_{x_i x_j} \Gamma(x,t;y,s)\, dy \Big|\, ds \\ &= \int_{t-r^2}^t \Big| \int_{\{y \in \mathbb{R}^N :\, d((x,t),(y,s)) \ge r\}} \partial^2_{x_i x_j} \Gamma(x,t;y,s)\, dy \Big|\, ds =: J_{r,\tau}(x,t). \end{aligned}$$

In order to prove (3.4.3), we then turn to bound the integral $J_r(x,t)$.

First of all, by combining the global upper estimates for $\partial^2_{x_i x_j}\Gamma$ in Theorem 3.5 with (3.1.7)-(3.1.10) (see also (3.3.28) in the proof of Proposition 3.13), we get

$$\begin{aligned} J_{r,\tau}(x,t) &\le \int_{t-r^2}^t \Big(\int_{\{y \in \mathbb{R}^N :\, d((x,t),(y,s)) \ge r\}} |\partial^2_{x_i x_j} \Gamma(x,t;y,s)|\, dy \Big) ds \\ &\le c_0 \int_{t-r^2}^t \frac{ds}{(t-s)^{Q/2+1}} \times \\ &\quad \times \Big(\int_{\{y \in \mathbb{R}^N :\, d((x,t),(y,s)) \ge r\}} e^{-c_0 \left| D_0\left(\frac{1}{\sqrt{t-s}}\right)(x - E(t-s)y) \right|^2} dy \Big) =: (\bigstar), \end{aligned}$$

where $c_0 > 0$ is a constant only depending on ν. From this, recalling (1.2.7) and using the change of variables

$$y = E(s-t)x - E(s-t)z \tag{3.4.4}$$

in the dy-integral, we obtain

$$(\bigstar) = c_0 \int_{t-r^2}^{t} \frac{|\det(E(s-t))|}{(t-s)^{Q/2+1}} \Bigg(\int_{\{z\in\mathbb{R}^N:\, \|z\|+\sqrt{t-s}\geq r\}} e^{-c_0 |D_0(\frac{1}{\sqrt{t-s}})z|^2}\, dz \Bigg) ds$$

$$\text{(since } \det(E(s-t)) = e^{(t-s)\det B} = 1, \text{ see (1.1.5))}$$

$$= c_0 \int_{\{t-r^2\leq s\leq t,\, \|z\|+\sqrt{t-s}\geq r\}} \frac{1}{(t-s)^{Q/2+1}} e^{-c_0 |D_0(\frac{1}{\sqrt{t-s}})z|^2}\, dz\, ds =: (2\bigstar).$$

To proceed further, we now perform another change of variables, this time involving both z and s: taking into account the D_0-homogeneity of $\|\cdot\|$, we set

$$(z, s) = \big(D_0(r)w, t - r^2\sigma\big). \tag{3.4.5}$$

Recalling that $\det(D_0(r)) = r^Q$, we then get

$$(2\bigstar) = c_0 \int_0^1 \frac{1}{\sigma^{Q/2+1}} \Bigg(\int_{\{w\in\mathbb{R}^N:\, \|w\|+\sqrt{\sigma}\geq 1\}} e^{-c_0 |D_0(\frac{1}{\sqrt{\sigma}})w|^2}\, dw \Bigg) d\sigma \equiv c_0\, \mathbf{J}.$$

Since the integral $\mathbf{J}$ is a constant, to complete the proof of (3.4.3) in this case it suffices to show that $\mathbf{J} < \infty$. To this end, we perform yet another change of variables in the dw-integral: setting

$$w = D_0(\sqrt{\sigma})u,$$

and taking into account that $\det(D_0(\sqrt{\sigma})) = \sigma^{Q/2}$, we obtain

$$\mathbf{J} = \int_0^1 \frac{1}{\sigma} \Bigg(\int_{\{u\in\mathbb{R}^N:\, \|u\|\geq \frac{1}{\sqrt{\sigma}}-1\}} e^{-c_0 |u|^2}\, du \Bigg) d\sigma$$

$$= \int_0^{1/4} \frac{h(\sigma)}{\sigma}\, d\sigma + \int_{1/4}^{1} \frac{h(\sigma)}{\sigma}\, d\sigma =: \mathbf{J}_1 + \mathbf{J}_2,$$

where we have introduced the shorthand notation

$$h(\sigma) := \int_{\{u\in\mathbb{R}^N:\, \|u\|\geq \frac{1}{\sqrt{\sigma}}-1\}} e^{-c_0 |u|^2}\, du.$$

We then turn to show that both the integrals $\mathbf{J}_1$, $\mathbf{J}_2$ are finite. As to $\mathbf{J}_1$ we first notice that, since $\theta_N\|u\| \leq |u|$ when $\|u\| \geq 1$ (here, $\theta_N > 0$ is a constant only depending on the dimension N), and since

$$\frac{1}{\sqrt{\sigma}} - 1 \geq 1 \quad \text{when } 0 < \sigma \leq \frac{1}{4},$$

we have the following estimate on the function h:

$$h(\sigma) \le \int_{\{u\in\mathbb{R}^N:\,|u|\ge \theta_N(\frac{1}{\sqrt{\sigma}}-1)\}} e^{-c_0\,|u|^2}\,du = \omega_N \int_{\theta_N(\frac{1}{\sqrt{\sigma}}-1)}^{+\infty} e^{-c_0\rho^2}\rho^{N-1}\,d\rho$$

$$\text{(since } e^{-c_0\rho^2}\rho^{N-1} \le \gamma\rho e^{-\frac{c_0}{2}\rho^2} \text{ when } N\ge 2)$$

$$= \gamma\,\omega_N \int_{\theta_N(\frac{1}{\sqrt{\sigma}}-1)}^{+\infty} \rho e^{-\frac{c_0}{2}\rho^2}\,d\rho = c_N e^{-\frac{c_0\theta_N^2}{2}(\frac{1}{\sqrt{\sigma}}-1)^2},$$

where $c_N := \gamma\,\omega_N/c_0$. As a consequence, we easily obtain

$$\mathbf{J}_1 \le c_N \int_0^{1/4} \frac{1}{\sigma} e^{-\frac{c_0\theta_N^2}{2}(\frac{1}{\sqrt{\sigma}}-1)^2}\,d\sigma < \infty.$$

As to $\mathbf{J}_2$, instead, taking into account that the map $u\mapsto e^{-c_0|u|^2}$ is integrable on $\mathbb{R}^N$, we immediately get

$$\mathbf{J}_2 \le \int_{1/4}^{1} \frac{1}{\tau}\Big(\int_{\mathbb{R}^N} e^{-c_0|u|^2}\,du\Big)d\sigma \le 4\int_{\mathbb{R}^N} e^{-c_0|u|^2}\,du < \infty.$$

Gathering these facts, we then conclude that $\mathbf{J} < \infty$, as desired.

Case II: $t-\tau \le r^2$. In this case, using once again the global upper estimates for $\partial^2_{x_ix_j}\Gamma$ in Theorem 3.5, and taking into account (3.1.7)-(3.1.10), we get

$$\begin{aligned} I_{r,\tau}(x,t) &\le \int_\tau^t \Big(\int_{\{y\in\mathbb{R}^N:\,d((x,t),(y,s))\ge r\}} |\partial^2_{x_ix_j}\Gamma(x,t;y,s)|\,dy\Big)ds \\ &\le c_0 \int_\tau^t \frac{ds}{(t-s)^{Q/2+1}} \times \\ &\qquad \times \Big(\int_{\{y\in\mathbb{R}^N:\,d((x,t),(y,s))\ge r\}} e^{-c_0|D_0(\frac{1}{\sqrt{t-s}})(x-E(t-s)y)|^2}\,dy\Big) =: (\star). \end{aligned}$$

Starting from this estimate, and performing the change of variables (3.4.4)-(3.4.5), we then obtain

$$\begin{aligned} (\star) &= c_0\int_\tau^t \frac{1}{(t-s)^{Q/2+1}}\Big(\int_{\{z\in\mathbb{R}^N:\,\|z\|+\sqrt{t-s}\ge r\}} e^{-c_0|D_0(\frac{1}{\sqrt{t-s}})z|^2}\,dz\Big)ds \\ &= c_0\int_0^{\frac{t-\tau}{r^2}} \frac{1}{\sigma^{Q/2+1}}\Big(\int_{\{w\in\mathbb{R}^N:\,\|w\|+\sqrt{\sigma}\ge 1\}} e^{-c_0|D_0(\frac{1}{\sqrt{\sigma}})w|^2}\,dw\Big)d\sigma =: (2\star). \end{aligned}$$

Now, since are assuming that $t-\tau\le r^2$, we have

$$(2\star) \le c_0\int_0^1 \frac{1}{\sigma^{Q/2+1}}\Big(\int_{\{w\in\mathbb{R}^N:\,\|w\|+\sqrt{\sigma}\ge 1\}} e^{-c_0|D_0(\frac{1}{\sqrt{\sigma}})w|^2}\,dw\Big)d\sigma = c_0\,\mathbf{J}, \tag{3.4.6}$$

where $\mathbf{J}$ is the same integral considered in the previous case; as a consequence, since we have already recognized that $\mathbf{J} < \infty$, from (3.4.6) we immediately derive (3.4.3) also in this case, and the proof is complete. $\square$

Theorem 3.17 *(Hölder continuity of singular integrals). For $T > \tau > -\infty$ and $\alpha \in (0,1)$, let us introduce the function space*

$$C_x^\alpha(\tau;T) := \{f \in C_x^\alpha(S_T) : \, f(x,t) = 0 \text{ for every } t \le \tau\},$$

and define, on this space $C_x^\alpha(\tau;T)$, the linear operator

$$f \mapsto T_{ij}f(x,t) := \int_{\mathbb{R}^N \times (\tau,t)} \partial^2_{x_i x_j}\Gamma(x,t;y,s)\left[f(E(s-t)x,s) - f(y,s)\right] dy\, ds.$$

Then, there exists a constant $c > 0$, depending on $(T-\tau)$ and α, such that

$$\|T_{ij}f\|_{C_x^\alpha(S_T)} \le c\,|f|_{C_x^\alpha(S_T)} \quad \text{for every } f \in C_x^\alpha(\tau;T). \tag{3.4.7}$$

Proof. Let $f \in C_x^\alpha(\tau;T)$ be arbitrarily fixed. Since $f(\cdot,t) \equiv 0$ for every $t \le \tau$, we have $T_{ij}f(x,t) = 0$ for every $x \in \mathbb{R}^N$ and $t \le \tau$. Thus, we derive that

$$\|T_{ij}f\|_{C_x^\alpha(S_T)} = \|T_{ij}f\|_{C_x^\alpha(\Omega)}, \quad \text{where } \Omega := \mathbb{R}^N \times (\tau,T).$$

Hence, to prove (3.4.7) it suffices to study $T_{ij}f\,(x,t)$ for $(x,t) \in \Omega$.

First of all, owing to Proposition 3.13, for every $(x,t) \in \Omega$ we have

$$\begin{aligned}
|T_{ij}f(x,t)| &\le \int_{\mathbb{R}^N \times (\tau,t)} |\partial^2_{x_i x_j}\Gamma(x,t;y,s)| \cdot |f(E(s-t)x,s) - f(y,s)|\, dy\, ds \\
&\le |f|_{C_x^\alpha(S_T)} \int_{\mathbb{R}^N \times (\tau,t)} |\partial^2_{x_i x_j}\Gamma(x,t;y,s)| \cdot \|E(s-t)x - y\|^\alpha\, dy\, ds \\
&\le c\,|f|_{C_x^\alpha(S_T)} \cdot (t-\tau)^{\alpha/2} \le c\,|f|_{C_x^\alpha(S_T)} \cdot (T-\tau)^{\alpha/2},
\end{aligned}$$

where $c > 0$ is a constant only depending on α. From this, we derive

$$\|T_{ij}f\|_{L^\infty(S_T)} \le c(T-\tau,\alpha)\,|f|_{C_x^\alpha(S_T)}. \tag{3.4.8}$$

On the other hand, if (x_1,t), $(x_2,t) \in \Omega$ are such that $\|x_1 - x_2\| \ge 1$, thanks to estimate (3.4.8) we also obtain the following bound

$$|T_{ij}f(x_1,t) - T_{ij}f(x_2,t)| \le 2\|Tf\|_{L^\infty(S_T)} \le c(T-\tau,\alpha)|f|_{C_x^\alpha(S_T)}\|x_1 - x_2\|^\alpha.$$

Thus, to prove (3.4.7) we are left to show that

$$\begin{gathered}
|T_{ij}f(x_1,t) - T_{ij}f(x_2,t)| \le c(T-\tau,\alpha)\|x_1 - x_2\|^\alpha \\
\text{for every } (x_1,t),\ (x_2,t) \in \Omega \text{ with } \|x_1 - x_2\| < 1.
\end{gathered} \tag{3.4.9}$$

To this end, taking into account the definition of Tf, we write

$$
\begin{aligned}
T_{ij}f(x_1,t) - T_{ij}f(x_2,t) &= \int\limits_{\mathbb{R}^N\times(\tau,t)} \Big\{\partial^2_{x_ix_j}\Gamma(x_1,t;y,s)\big[f(E(s-t)x_1,s)-f(y,s)\big] \\
&\qquad - \partial^2_{x_ix_j}\Gamma(x_2,t;y,s)\big[f(E(s-t)x_2,s)-f(y,s)\big]\Big\}\,dy\,ds \\
&= \int\limits_{\{(y,s):\,d((x_2,t),(y,s))\geq 4\boldsymbol{\kappa}\rho\}} \{\cdots\}\,dy\,ds \\
&\quad + \int\limits_{\{(y,s):\,d((x_2,t),(y,s))<4\boldsymbol{\kappa}\rho\}} \{\cdots\}\,dy\,ds \\
&=: \mathrm{A}_1 + \mathrm{A}_2,
\end{aligned}
\tag{3.4.10}
$$

where $\boldsymbol{\kappa} > 0$ is as in (1.2.5)-(1.2.6) and

$$
\rho := d((x_2,t),(x_1,t)) = \|x_1 - x_2\|.
$$

We then turn to estimate A_1 and A_2.

- Estimate of A_1. To begin with, we write A_1 as follows:

$$
\begin{aligned}
\mathrm{A}_1 &= \int\limits_{\{(y,s):\,d((x_2,t),(y,s))\geq 4\boldsymbol{\kappa}\rho\}} \Big\{\big[f(E(s-t)x_1,s)-f(y,s)\big]\times \\
&\qquad \times\big[\partial^2_{x_ix_j}\Gamma(x_1,t;y,s) - \partial^2_{x_ix_j}\Gamma(x_2,t;y,s)\big]\Big\}\,dy\,ds \\
&\quad + \int\limits_{\{(y,s):\,d((x_2,t),(y,s))\geq 4\boldsymbol{\kappa}\rho\}} \Big\{\partial^2_{x_ix_j}\Gamma(x_2,t;y,s)\times \\
&\qquad \times\big[f(E(s-t)x_1,s)-f(E(s-t)x_2,s)\big]\Big\}\,dy\,ds \\
&=: \mathrm{A}_{11} + \mathrm{A}_{12}.
\end{aligned}
$$

Estimate of A_{11}. First of all we observe that, owing to the mean value inequalities in Theorem 3.9 (and taking into account the definition of ρ), we have

$$
\begin{aligned}
&\big|\partial^2_{x_ix_j}\Gamma(x_1,t;y,s) - \partial^2_{x_ix_j}\Gamma(x_2,t;y,s)\big| \\
&\qquad \leq c\,\frac{d((x_2,t),(x_1,t))}{d((x_2,t),(y,s))^{Q+3}} = c\,\frac{\|x_1-x_2\|}{d((x_2,t),(y,s))^{Q+3}},
\end{aligned}
$$

for every $(y,s)\in\Omega$ such that $d((x_2,t),(y,s))\geq 4\boldsymbol{\kappa}\rho$. Moreover, using the explicit expression of d in (1.2.7) and the quasi-symmetry property (1.2.6), we get

$$
\begin{aligned}
\big|f(E(s-t)x_1,s)-f(y,s)\big| &\leq |f|_{C^\alpha_x(S_T)}\|E(s-t)x_1 - y\|^\alpha \\
&\leq |f|_{C^\alpha_x(S_T)} d((y,s),(x_1,t))^\alpha \\
&\leq \boldsymbol{\kappa}^\alpha\,|f|_{C^\alpha_x(S_T)} d((x_1,t),(y,s))^\alpha,
\end{aligned}
$$

where we have also used the fact that $f\in C^\alpha_x(\tau;T)$. Hence, by combining these estimates and by using Lemma 2.3, we get

$$
\begin{aligned}
&\left|f(E(s-t)x_1,s)-f(y,s)\right|\cdot\left|\partial^2_{x_ix_j}\Gamma(x_1,t;y,s)-\partial^2_{x_ix_j}\Gamma(x_2,t;y,s)\right| \\
&\quad\le c\,|f|_{C^\alpha_x(S_T)}\|x_1-x_2\|\cdot\frac{d((x_1,t),(y,s))^\alpha}{d((x_2,t),(y,s))^{Q+3}} \\
&\quad\le c\,|f|_{C^\alpha_x(S_T)}\|x_1-x_2\|\cdot\frac{1}{d((x_2,t),(y,s))^{Q+3-\alpha}},
\end{aligned}
\tag{3.4.11}
$$

for every $(y,s)\in\mathbb{R}^N\times(\tau,t)$ satisfying $d((x_2,t),(y,s)\ge 4\boldsymbol{\kappa}\rho>2\boldsymbol{\kappa}\rho$. Owing to (3.4.11), and exploiting (2.0.2) in Lemma 2.2, we finally obtain

$$
\begin{aligned}
|\mathrm{A}_{11}| &\le c\,|f|_{C^\alpha_x(S_T)}\,\|x_1-x_2\|\int_{\{\eta:\,d(\xi,\eta)\ge 4\boldsymbol{\kappa}\rho\}}\frac{1}{d(\xi,\eta)^{Q+3-\alpha}}\,d\eta \\
&\le c\,|f|_{C^\alpha_x(S_T)}\,\|x_1-x_2\|\cdot\rho^{\alpha-1}=c\,|f|_{C^\alpha_x(S_T)}\,\|x_1-x_2\|^\alpha,
\end{aligned}
\tag{3.4.12}
$$

where $c>0$ is a constant only depending on α.

Estimate of A_{12}. First of all, using once again the fact that $f\in C^\alpha_x(S_T)$, jointly with Lemma 2.5, we can bound the integral A_{12} as follows:

$$
\begin{aligned}
|\mathrm{A}_{12}| &\le \int_\tau^t \left|f(E(s-t)x_1,s)-f(E(s-t)x_2,s)\right|\cdot\mathcal{J}(s)\,ds \\
&\le |f|_{C^\alpha_x(S_T)}\int_\tau^t \|E(s-t)(x_1-x_2)\|^\alpha\cdot\mathcal{J}(s)\,ds \\
&\le c\,|f|_{C^\alpha_x(S_T)}\int_\tau^t \left(\|x_1-x_2\|+\sqrt{t-s}\right)^\alpha\cdot\mathcal{J}(s)\,ds,
\end{aligned}
\tag{3.4.13}
$$

where $c>0$ is an absolute constant and

$$
\mathcal{J}(s):=\left|\int_{\{y\in\mathbb{R}^N:\,d((x_2,t),(y,s))\ge 4\boldsymbol{\kappa}\rho\}}\partial^2_{x_ix_j}\Gamma(x_2,t;y,s)\,dy\right|.
$$

We now distinguish two cases, according to the value of $\theta:=t-16\boldsymbol{\kappa}^2\rho^2$.

(i) $\theta>\tau$. In this case, we start from (3.4.13) and we write

$$
\begin{aligned}
|\mathrm{A}_{12}| &\le c\,|f|_{C^\alpha_x(S_T)}\int_\tau^\theta \left(\|x_1-x_2\|+\sqrt{t-s}\right)^\alpha\cdot\mathcal{J}(s)\,ds \\
&\quad + c\,|f|_{C^\alpha_x(S_T)}\int_\theta^t \left(\|x_1-x_2\|+\sqrt{t-s}\right)^\alpha\cdot\mathcal{J}(s)\,ds.
\end{aligned}
\tag{3.4.14}
$$

We now observe that, when $\theta\le s\le t$, we have $0\le t-s\le 16\boldsymbol{\kappa}^2\rho^2$; thus, by using the cancellation property of $\mathcal{J}$ in Theorem 3.16, we get

$$\begin{aligned}
&\int_{\theta}^{t} \left(\|x_1 - x_2\| + \sqrt{t-s}\right)^{\alpha} \cdot \mathcal{J}(s)\,ds \\
&\qquad \leq (1+4\boldsymbol{\kappa})^{\alpha} \|x_1 - x_2\|^{\alpha} \int_{\theta}^{t} \mathcal{J}(s)\,ds \\
&\qquad \leq (1+4\boldsymbol{\kappa})^{\alpha} \|x_1 - x_2\|^{\alpha} \int_{\tau}^{t} \mathcal{J}(s)\,ds \\
&\qquad = (1+4\boldsymbol{\kappa})^{\alpha} \|x_1 - x_2\|^{\alpha} \cdot I_{4\boldsymbol{\kappa}\rho,\tau}(x_2,t) \\
&\qquad \leq c\,\|x_1 - x_2\|^{\alpha},
\end{aligned} \tag{3.4.15}$$

where $c > 0$ is a suitable constant only depending on α.

On the other hand, when $\tau \leq s < \theta$, by (1.2.7) we infer that

$$d((x_2,t),(y,s)) \geq \sqrt{t-s} \geq 4\boldsymbol{\kappa}\rho \quad \forall\, y \in \mathbb{R}^N;$$

as a consequence, from Lemma 3.8 we obtain

$$\begin{aligned}
&\int_{\tau}^{\theta} \left(\|x_1 - x_2\| + \sqrt{t-s}\right)^{\alpha} \cdot \mathcal{J}(s)\,ds \\
&\qquad = \int_{\tau}^{\theta} \left(\|x_1 - x_2\| + \sqrt{t-s}\right)^{\alpha} \cdot \left| \int_{\mathbb{R}^N} \partial^2_{x_i x_j} \Gamma(\xi_1; y, s)\,dy \right| ds = 0.
\end{aligned} \tag{3.4.16}$$

Summing up, by combining (3.4.15)-(3.4.16) with (3.4.14), we conclude that

$$|A_{12}| \leq c\,|f|_{C^{\alpha}_x(S_T)}\,\|x_1 - x_2\|^{\alpha}, \tag{3.4.17}$$

for a suitable constant $c > 0$ only depending on α.

(ii) $\theta \leq \tau$. In this case, starting from (3.4.13) and using once again the cancellation property of $\mathcal{J}$ in Theorem 3.16, we immediately get

$$\begin{aligned}
|\mathrm{A}_{12}| &\leq c\,|f|_{C^{\alpha}_x(S_T)} \int_{\tau}^{t} \left(\|x_1 - x_2\| + \sqrt{t-\tau}\right)^{\alpha} \cdot \mathcal{J}(s)\,ds \\
&\leq c\,|f|_{C^{\alpha}_x(S_T)} \|x_1 - x_2\|^{\alpha} \int_{\tau}^{t} \mathcal{J}(s)\,ds \\
&\leq c\,|f|_{C^{\alpha}_x(S_T)} \|x_1 - x_2\|^{\alpha} \cdot I_{4\boldsymbol{\kappa}\rho,\tau}(x_2,t) \\
&\leq c\,|f|_{C^{\alpha}_x(S_T)} \|x_1 - x_2\|^{\alpha},
\end{aligned} \tag{3.4.18}$$

where $c > 0$ is another constant only depending on α.

All in all, by combining (3.4.12) with (3.4.17)-(3.4.18), we conclude that

$$|\mathrm{A}_1| \leq c\,|f|_{C^{\alpha}_x(S_T)} \|x_1 - x_2\|^{\alpha}, \tag{3.4.19}$$

for a suitable constant $c > 0$ only depending on α.

- Estimate of A_2. We first observe that, since $f \in C_x^\alpha(\tau; T)$, one has

$$|\mathrm{A}_2| \le |f|_{C_x^\alpha(S_T)} \cdot \big(\mathrm{A}_{21} + \mathrm{A}_{22}\big), \tag{3.4.20}$$

where, for $k = 1, 2$, we have introduced the notation

$$\mathrm{A}_{2k} := \int_{\{(y,s)\colon d((x_2,t),(y,s))<4\boldsymbol{\kappa}\rho\}} |\partial^2_{x_i x_j}(x_k, t; y, s)\Gamma| \cdot \|E(s-t)x_k - y\|^\alpha \, dy \, ds.$$

We then proceed by estimating the two integrals A_{21}, A_{22} separately.

Estimate of A_{21}. First of all, by using the estimates for $\partial^2_{x_i x_j}\Gamma$ given in Theorem 3.5, jointly with (1.2.7), we get

$$\begin{aligned} \mathrm{A}_{21} &\le c \int_{\{(y,s)\colon d((x_2,t),(y,s))<4\boldsymbol{\kappa}\rho\}} \frac{\|E(s-t)x_1 - y\|^\alpha}{d((x_1,t),(y,s))^{Q+2}} \, dy \, ds \\ &\le c \int_{\{(y,s)\colon d((x_2,t),(y,s))<4\boldsymbol{\kappa}\rho\}} \frac{d((y,s),(x_1,t))^\alpha}{d((x_1,t),(y,s))^{Q+2}} \, dy \, ds \\ &\le c \int_{\{(y,s)\colon d((x_2,t),(y,s))<4\boldsymbol{\kappa}\rho\}} \frac{1}{d((x_1,t),(y,s))^{Q+2-\alpha}} \, dy \, ds =: (\bigstar). \end{aligned}$$

On the other hand, by the quasi-triangular inequality (1.2.5), we have

$$\begin{aligned} d((x_1,t),(y,s)) &\le \boldsymbol{\kappa}\big(d((x_1,t),(x_2,t)) + d((y,s),(x_2,t))\big) \\ &\le \boldsymbol{\kappa}^2\big(d((x_2,t),(x_1,t)) + d((x_2,t),(y,s))\big) \\ &= \boldsymbol{\kappa}^2(1+4\boldsymbol{\kappa})\rho, \end{aligned} \tag{3.4.21}$$

for every $(y, s) \in \mathbb{R}^{N+1}$ such that $d((x_2,t),(y,s)) < 4\boldsymbol{\kappa}\rho$. On account of (3.4.21), and exploiting (2.0.1) in Lemma 2.2, we finally obtain

$$\begin{aligned} (\bigstar) &\le c \int_{\{(y,s)\colon d((x_1,t),(y,s))<\boldsymbol{\kappa}^2(1+4\boldsymbol{\kappa})\rho\}} \frac{1}{d((x_1,t),(y,s))^{Q+2-\alpha}} \, dy \, ds \\ &= c \int_{\{\eta\colon d(\xi,\eta)<\boldsymbol{\kappa}^2(1+4\boldsymbol{\kappa})\rho\}} \frac{1}{d(\xi,\eta)^{Q+2-\alpha}} \, d\eta \\ &\le c\rho^\alpha = c\,\|x_1 - x_2\|^\alpha. \end{aligned} \tag{3.4.22}$$

Estimate of A_{22}. Using once again the estimates for $\partial^2_{x_i x_j}\Gamma$ in Theorem 3.5, together with (1.2.6)-(1.2.7) and (2.0.1) in Lemma 2.2, we readily obtain

$$\begin{aligned}
\mathrm{A}_{22} &\leq c \int\limits_{\{(y,s)\colon d((x_2,t),(y,s))<4\boldsymbol{\kappa}\rho\}} \frac{\|E(s-t)x_2 - y\|^\alpha}{d((x_2,t),(y,s))^{Q+2}}\,dy\,ds \\
&\leq c \int\limits_{\{(y,s)\colon d((x_2,t),(y,s))<4\boldsymbol{\kappa}\rho\}} \frac{d((y,s),(x_2,t))^\alpha}{d((x_2,t),(y,s))^{Q+2}}\,dy\,ds \\
&= c \int\limits_{\{\eta\colon d(\xi,\eta)<4\boldsymbol{\kappa}\rho\}} \frac{1}{d(\xi,\eta)^{Q+2-\alpha}}\,d\eta \\
&\leq c\rho^\alpha = c\,\|x_1 - x_2\|^\alpha.
\end{aligned} \tag{3.4.23}$$

Summing up, by combining (3.4.22)-(3.4.23) with (3.4.20), we conclude that

$$|\mathrm{A}_2| \leq c\,|f|_{C^\alpha_x(S_T)}\|x_1 - x_2\|^\alpha, \tag{3.4.24}$$

where $c > 0$ is a suitable constant only depending on α.

Now we have estimated A_1 and A_2, we are finally ready to complete the proof: in fact, gathering (3.4.19)-(3.4.24), and recalling (3.4.10), we conclude that

$$|T_{ij}f(x_1,t) - T_{ij}f(x_2,t)| \leq |\mathrm{A}_1| + |\mathrm{A}_2| \leq c\,|f|_{C^\alpha_x(S_T)}\|x_1 - x_2\|^\alpha,$$

which is exactly the desired (3.4.9). □

Thanks to all the results established so far, we can finally give the

Proof of Theorem 3.15. Let T, τ, α be as in the statement, and let $u \in \mathcal{S}^0(\tau;T)$ be such that $\mathcal{L}u \in C^\alpha_x(S_T)$. By the representation formula (3.3.30), we have

$$\begin{aligned}
\partial_{x_ix_j}u(x,t) &= \int\limits_{\mathbb{R}^N\times(\tau,t)} \partial^2_{x_ix_j}\Gamma(x,t;y,s)\cdot\big[\mathcal{L}u(E(s-t)x,s) - \mathcal{L}u(y,s)\big]\,dy\,ds \\
&= T_{ij}(\mathcal{L}u)(x,t) \qquad \text{for every } (x,t)\in S_T \text{ and } 1\leq i,j\leq q,
\end{aligned}$$

where T_{ij} is as in Theorem 3.17. Then, from (3.4.7) we infer that

$$\|\partial_{x_ix_j}u\|_{C^\alpha_x(S_T)} = \|T_{ij}(\mathcal{L}u)\|_{C^\alpha_x(S_T)} \leq c\,|\mathcal{L}u|_{C^\alpha_x(S_T)}, \tag{3.4.25}$$

where $c > 0$ is a constant only depending on $(T-\tau)$ and α, and this is (3.4.1).

On the other hand, using the definition of $\mathcal{L}$, and recalling that the coefficients $a_{ij}(\cdot)$ are globally bounded on $\mathbb{R}$ and *independent of* x, from (3.4.25) we also get

$$\begin{aligned}
\|Yu\|_{C^\alpha_x(S_T)} &= \Big\|\mathcal{L}u - \sum_{i,j=1}^q a_{ij}\,\partial_{x_ix_j}u\Big\|_{C^\alpha_x(S_T)} \\
&\leq \|\mathcal{L}u\|_{C^\alpha_x(S_T)} + \sum_{i,j=1}^q \|a_{ij}\|_{L^\infty(\mathbb{R})}\cdot\|\partial_{x_ix_j}u\|_{C^\alpha_x(S_T)} \\
&\leq c\,\|\mathcal{L}u\|_{C^\alpha_x(S_T)}.
\end{aligned} \tag{3.4.26}$$

This is (3.4.2), and we are done. □

3.5. Schauder estimates in space and time

Theorem 3.15 shows that, for the derivatives $\partial^2_{x_i x_j} u$ (with $1 \le i, j \le q$) we can bound the C^α_x-norm in terms of the quantity $|\mathcal{L}u|_{C^\alpha_x(S_T)}$. We now aim to show how to improve the previous result, giving a control on the Hölder norm of $\partial^2_{x_i x_j} u$ with respect to both space and time, without strengthening the assumptions on $\mathcal{L}u$.

Theorem 3.18 *(Local Schauder estimates in space-time). Let $T > \tau > -\infty$, $\alpha \in (0,1)$, and let $K \subseteq \mathbb{R}^N$ be a* compact set*.*

Then, there exists a constant $c = c(K, \tau, T) > 0$ such that, for every $u \in \mathcal{S}^0(\tau; T)$ such that $\mathcal{L}u \in C^\alpha_x(S_T)$, one has

$$|\partial^2_{x_i x_j} u(x_1, t_1) - \partial^2_{x_i x_j} u(x_2, t_2)| \le c\, |\mathcal{L}u|_{C^\alpha_x(S_T)} \big(d((x_1, t_1), (x_2, t_2))^\alpha + |t_1 - t_2|^{\alpha/q_N} \big) \tag{3.5.1}$$

for every $1 \le i, j \le q$ and every $(x_1, t_1), (x_2, t_2) \in K \times [\tau, T]$. We recall that $q_N \ge 3$ is the largest exponent in the dilations $D_0(\lambda)$, see (1.1.7).

To prove Theorem 3.18, we first establish the following technical lemma.

Lemma 3.19. *Let $K \subseteq \mathbb{R}^N$ be a fixed compact set, and let $T > \tau > -\infty$. There exists a constant $c = c(K, \tau, T) > 0$ such that*

$$\|x - E(t-s)x\| \le c\, |t-s|^{1/q_N} \qquad \text{for every } x \in K \text{ and } t, s \in [\tau, T] \tag{3.5.2}$$

$$\|(E(t) - E(s))x\| \le c\, |t-s|^{1/q_N} \qquad \text{for every } x \in K \text{ and } t, s \in [\tau, T]. \tag{3.5.3}$$

Proof. We begin with the proof (3.5.2). To this end, we fix $x \in K$ and $t, s \in [\tau, T]$, and we choose $\rho = \rho(K) \ge 1$ such that $K \subseteq \{|z| \le \rho\}$. Taking into account the explicit expression of $\|\cdot\|$ given in (3.2.22), we have

$$\begin{aligned} \|x - E(t-s)x\| &\le \sum_{i=1}^N \big| \big(\mathrm{Id}_N - E(t-s)\big)x \big|^{1/q_i} \\ &\le \rho \sum_{i=1}^N \|\mathrm{Id}_N - E(t-s)\|_{\mathrm{Op}}^{1/q_i}, \end{aligned} \tag{3.5.4}$$

where $\|\cdot\|_{\mathrm{Op}}$ denotes the *operator norm* of a matrix. On the other hand, recalling that $E(\sigma) = e^{-\sigma B}$ (and since $\tau \le t, s \le T$), we also have

$$\begin{aligned} \|\mathrm{Id}_N - E(t-s)\|_{\mathrm{Op}} &\le \sum_{k=1}^\infty \frac{|t-s|^k\, \|B\|_{\mathrm{Op}}^k}{k!} \\ &\le |t-s| \sum_{k=1}^\infty \frac{(T-\tau)^{k-1}\, \|B\|_{\mathrm{Op}}^k}{k!} = c(\tau, T) \cdot |t-s|. \end{aligned} \tag{3.5.5}$$

Gathering (3.5.4)-(3.5.5), and recalling that $1 = q_1 \le \ldots \le q_N$, we then get

$$\|x - E(t-s)x\| \le c(\tau, T)\rho \cdot \sum_{i=1}^N |t-s|^{1/q_i} \le c\, |t-s|^{1/q_N},$$

where $c > 0$ only depends on K, τ, T. This completes the proof of (3.5.2).

We now turn to establish (3.5.3). To this end, we fix $x \in K$ and $t, s \in [\tau, T]$. By applying the Mean Value Theorem to the function $\gamma(\sigma) = E(\sigma)x$ (and taking into account that $E(\sigma) = e^{-\sigma B}$), we have the estimate

$$\begin{aligned} |(E(t) - E(s))x| &= |\gamma(t) - \gamma(s)| \leq |\gamma'(\theta)| \cdot |t - s| \\ &= |t - s| \cdot |BE(\theta)x| \\ &\leq \rho\, |t - s| \cdot \|BE(\theta)\|_{\mathrm{Op}}, \end{aligned} \tag{3.5.6}$$

where θ is a suitable point between t and s, and $\rho \geq 1$ is as before. On the other hand, observing that $\tau \leq \theta \leq T$ (as the same is true of t, s), we also get

$$\begin{aligned} \|BE(\theta)\|_{\mathrm{Op}} &\leq \sum_{k=0}^{\infty} \frac{|\theta|^k \, \|B\|_{\mathrm{Op}}^{k+1}}{k!} \\ &\leq \sum_{k=0}^{\infty} \frac{\max\{|\tau|, |T|\}^k \cdot \|B\|_{\mathrm{Op}}^{k+1}}{k!} =: c(\tau, T). \end{aligned} \tag{3.5.7}$$

Gathering (3.5.6)-(3.5.7), we then obtain

$$\begin{aligned} \|((E(t) - E(s))x\| &\leq \sum_{i=1}^{N} |(E(t) - E(s))x|^{1/q_i} \\ &\leq c(\tau, T)\rho \cdot \sum_{i=1}^{N} |t - s|^{1/q_i} \leq c\, |t - s|^{1/q_N}, \end{aligned}$$

where $c > 0$ only depends on K, τ, T. This completes the proof. □

With Lemma 3.19, we can now prove Theorem 3.18.

Proof (of Theorem 3.18). Let $u \in \mathcal{S}^0(\tau; T)$ be such that $\mathcal{L}u \in C_x^\alpha(S_T)$. First of all we observe that, owing to Theorem 3.15 (and taking into account the expression of d given in (1.2.7)), there exists an absolute constant $c > 0$ such that

$$|\partial^2_{x_i x_j} u(x_1, t) - \partial^2_{x_i x_j} u(x_2, t)| \leq |\partial^2_{x_i x_j} u|_{C_x^\alpha(S_T)} \|x_1 - x_2\|^\alpha \leq c\, |\mathcal{L}u|_{C_x^\alpha(S_T)} \|x_1 - x_2\|^\alpha, \tag{3.5.8}$$

for every $(x_1, t), (x_2, t) \in S_T$. As a consequence of (3.5.8), and taking into account Lemma 3.19, to prove (3.5.1) it suffices to show that

$$|\partial^2_{x_i x_j} u(x, t_1) - \partial^2_{x_i x_j} u(x, t_2)| \leq c\, |\mathcal{L}u|_{C_x^\alpha(S_T)} \left\{ d((x, t_1), (x, t_2))^\alpha + |t_1 - t_2|^{\alpha/q_N} \right\}, \tag{3.5.9}$$

for every $(x, t_1), (x, t_2) \in K \times [\tau, T]$, where $c > 0$ is an absolute constant independent of u (but possibly depending on the fixed K, τ, T). In fact, once (3.5.9) has been established, by combining (3.5.8)-(3.5.9) with Lemma 3.19 we get

$$\begin{aligned} &|\partial^2_{x_i x_j} u(x_1, t_1) - \partial^2_{x_i x_j} u(x_2, t_2)| \\ &\quad \leq |\partial^2_{x_i x_j} u(x_1, t_1) - \partial^2_{x_i x_j} u(x_2, t_1)| + |\partial^2_{x_i x_j} u(x_2, t_1) - \partial^2_{x_i x_j} u(x_2, t_2)| \\ &\quad \leq c\, |\mathcal{L}u|_{C_x^\alpha(S_T)} \big(\|x_1 - x_2\|^\alpha + d((x_2, t_1), (x_2, t_2))^\alpha + |t_1 - t_2|^{\alpha/q_N} \big) \\ &\quad \text{(by the explicit expression of } d\text{, see (1.2.7))} \end{aligned}$$

$$\leq c\,|\mathcal{L}u|_{C^\alpha_x(S_T)}\big(\|x_1-x_2\|^\alpha+\|x_2-E(t_1-t_2)x_2\|^\alpha+|t_1-t_2|^{\alpha/2}+|t_1-t_2|^{\alpha/q_N}\big)$$
(recalling that, by assumption, $q_N \geq 3$)
$$\leq c\,|\mathcal{L}u|_{C^\alpha_x(S_T)}\big(\|x_1-x_2\|^\alpha+|t_1-t_2|^{\alpha/q_N}\big) =: (\bigstar);$$

from this, using the quasi-triangle inequality (1.2.5), we obtain

$$(\bigstar) \leq c\,|\mathcal{L}u|_{C^\alpha_x(S_T)}\big(\|x_1-E(t_1-t_2)x_2\|^\alpha+\|x_2-E(t_1-t_2)x_2\|^\alpha+|t_1-t_2|^{\alpha/q_N}\big)$$
$$\leq c\,|\mathcal{L}u|_{C^\alpha_x(S_T)}\big(\|x_1-E(t_1-t_2)x_2\|^\alpha+|t_1-t_2|^{\alpha/q_N}\big)$$
(again by the expression of d in (1.2.7))
$$\leq c\,|\mathcal{L}u|_{C^\alpha_x(S_T)}\big(d((x_1,t_1),(x_2,t_2))^\alpha+|t_1-t_2|^{\alpha/q_N}\big),$$

which is exactly (3.5.1). Hence, we turn to prove (3.5.9).

This can be done adapting several computations exploited in the proof of Theorem 3.17. We will point out just the relevant differences.

Let us fix two points $(x,t_1),(x,t_2) \in K\times[\tau,T]$ and exploit the representation formula (3.3.30) for $\partial^2_{x_ix_j}u$: assuming, to fix the ideas, that $t_2 \geq t_1$, we can write

$$\begin{aligned}
&\partial^2_{x_ix_j}u(x,t_1)-\partial^2_{x_ix_j}u(x,t_2)\\
&= \int_{\mathbb{R}^N\times(\tau,t_1)} \Big\{\partial^2_{x_ix_j}\Gamma(x,t_1;y,s)\big[\mathcal{L}u(E(s-t_1)x,s)-\mathcal{L}u(y,s)\big]\\
&\qquad -\partial^2_{x_ix_j}\Gamma(x,t_2;y,s)\big[\mathcal{L}u(E(s-t_2)x,s)-\mathcal{L}u(y,s)\big]\Big\}\,dy\,ds\\
&\quad -\int_{\mathbb{R}^N\times(t_1,t_2)} \partial^2_{x_ix_j}\Gamma(x,t_2;y,s)\big[\mathcal{L}u(E(s-t_2)x,s)-\mathcal{L}u(y,s)\big]\,dy\,ds\\
&= \int_{\{(y,s):\,d((x,t_2),(y,s))\geq 4\boldsymbol{\kappa}\rho\}} \{\cdots\}\,dy\,ds\\
&\qquad + \int_{\{(y,s):\,d((x,t_2),(y,s))<4\boldsymbol{\kappa}\rho\}} \{\cdots\}\,dy\,ds\\
&\qquad - \int_{\mathbb{R}^N\times(t_1,t_2)} \{\cdots\}\,dy\,ds\\
&=: \mathrm{A}_1+\mathrm{A}_2-\mathrm{A}_3,
\end{aligned} \tag{3.5.10}$$

where $\boldsymbol{\kappa}>0$ is as in (1.2.5)-(1.2.6) and

$$\rho := d((x,t_2),(x,t_1)).$$

We now turn to estimate the integrals A_k (for $k=1,2,3$).

- Estimate of A_1. To begin with, we write A_1 as follows:

$$\begin{aligned}
\mathrm{A}_1 = \int_{\{(y,s):\,d((x,t_2),(y,s))\geq 4\boldsymbol{\kappa}\rho\}} &\Big\{\big[\mathcal{L}u(E(s-t_1)x,s)-\mathcal{L}u(y,s)\big]\times\\
&\times\big[\partial^2_{x_ix_j}\Gamma(x,t_1;y,s)-\partial^2_{x_ix_j}\Gamma(x,t_2;y,s)\big]\Big\}\,dy\,ds
\end{aligned}$$

$$+ \int_{\{(y,s):\, d((x,t_2),(y,s))\geq 4\boldsymbol{\kappa}\rho\}} \Big\{\partial^2_{x_i x_j}\Gamma(x,t_2;y,s)\times$$
$$\times \big[\mathcal{L}u(E(s-t_1)x,s) - \mathcal{L}u(E(s-t_2)x,s)\big]\Big\}\, dy\, ds$$
$$=: \mathrm{A}_{11} + \mathrm{A}_{12}.$$

Estimate of A_{11}. This can be done analogously to what done in the proof of Theorem 3.17 for A_{11}, with $d\left((x,t_1),(x,t_2)\right)$ now replacing $\|x_1 - x_2\|$, getting

$$|\mathrm{A}_{11}| \leq c\, |\mathcal{L}u|_{C^\alpha_x(S_T)}\, d\left((x,t_1),(x,t_2)\right)^\alpha \tag{3.5.11}$$

where $c > 0$ is a constant only depending on α.

Estimate of A_{12}. First of all, using once again the fact that $\mathcal{L}u \in C^\alpha_x(S_T)$, jointly with Lemma 3.19, we can bound the integral A_{12} as follows:

$$\begin{aligned}
|\mathrm{A}_{12}| &\leq \int_\tau^{t_1} \big|\mathcal{L}u(E(s-t_1)x,s) - \mathcal{L}u(E(s-t_2)x,s)\big| \cdot \mathcal{J}(s)\, ds \\
&\leq |\mathcal{L}u|_{C^\alpha_x(S_T)} \int_\tau^{t_1} |(E(s-t_1) - E(s-t_2))x|^\alpha \cdot \mathcal{J}(s)\, ds \\
&\quad (\text{since } |s-t_1|, |s-t_2| \leq T-\tau \text{ for all } \tau \leq s \leq t_1) \\
&\leq c\, |\mathcal{L}u|_{C^\alpha_x(S_T)} \cdot |t_1 - t_2|^{\alpha/q_N} \int_\tau^{t_1} \mathcal{J}(s)\, ds =: (\bigstar)
\end{aligned}$$

where $c > 0$ is an absolute constant and

$$\mathcal{J}(s) := \left| \int_{\{y\in\mathbb{R}^N:\, d((x,t_2),(y,s))\geq 4\boldsymbol{\kappa}\rho\}} \partial^2_{x_i x_j}\Gamma(x,t_2;y,s)\, dy \right|.$$

From this, using the cancellation property of $\mathcal{J}$ in Theorem 3.16, we obtain

$$(\bigstar) \leq c\, |\mathcal{L}u|_{C^\alpha_x(S_T)} |t_1 - t_2|^{\alpha/q_N} \tag{3.5.12}$$

for a suitable constant $c > 0$ only depending on α.

By combining (3.5.11) with (3.5.12), we conclude that

$$|\mathrm{A}_1| \leq c\, |\mathcal{L}u|_{C^\alpha_x(S_T)} \left\{ d((x,t_1),(x,t_2))^\alpha + |t_1 - t_2|^{\alpha/q_N} \right\}, \tag{3.5.13}$$

for a suitable constant $c > 0$ only depending on α.

- Estimate of A_2. This can be done analogously to what done in the proof of Theorem 3.17 for A_2, with $d\left((x,t_1),(x,t_2)\right)$ now replacing $\|x_1 - x_2\|$, getting

$$|\mathrm{A}_2| \leq c\, |\mathcal{L}u|_{C^\alpha_x(S_T)}\, d((x,t_1),(x,t_2))^\alpha, \tag{3.5.14}$$

where $c > 0$ is a suitable constant only depending on α.

- Estimate of A_3. Using once again the fact that $\mathcal{L}u \in C_x^\alpha(S_T)$, together with the estimate (3.3.26) in Proposition 3.13, we immediately obtain

$$\begin{aligned}|A_3| &\leq |\mathcal{L}u|_{C_x^\alpha(S_T)} \int_{\mathbb{R}^N \times (t_1,t_2)} |\partial^2_{x_i x_j}\Gamma(x,t_2;y,s)| \cdot \|E(s-t_2)x-y\|^\alpha \, dy\, ds \\ &\leq c\,(t_2-t_1)^{\alpha/2} \leq c\,|t_1-t_2|^{\alpha/q_N}\end{aligned} \tag{3.5.15}$$

where $c > 0$ only depends on α.

Now we have estimated A_1, A_2 and A_3, we can complete the proof: in fact, gathering (3.5.13), (3.5.14) and (3.5.15), and recalling (3.5.10), we conclude that

$$\begin{aligned}|\partial^2_{x_i x_j}u(x,t_1) - \partial^2_{x_i x_j}u(x,t_2)| &\leq |A_1| + |A_2| + |A_3| \\ &\leq c\,|\mathcal{L}u|_{C_x^\alpha(S_T)}\big(d((x,t_1),(x,t_2))^\alpha + |t_1-t_2|^{\alpha/2}\big),\end{aligned}$$

which is exactly the desired (3.5.9). □

4. Schauder estimates for operators with coefficients depending on (x,t)

Throughout this section we study operators (1.1.1) with coefficients $a_{ij}(x,t)$ depending on both space and time, fulfilling assumptions (H1), (H2), (H3) stated in section 1.

Here we will prove our main result, Theorem 1.7, exploiting all the results proved so far.

4.1. Local Schauder estimates in space

Throughout this section we will consider metric balls $B_r(\xi)$ centred at points $\xi \in \mathbb{R}^N \times (0,T)$ but possibly overlapping the hyperplanes $t = 0$ and $t = T$ (since these balls will eventually build a covering of $\mathbb{R}^N \times (0,T)$). Our functions $u \in \mathcal{S}^\alpha(0;T)$, so that they are actually defined and jointly continuous in the whole ball $B_r(\xi) \cap S_T$; however, the derivative Yu is merely an L^∞ function of the joint variables.

Notation. Throughout this section, we will set

$$B_\rho^T(\xi) := B_\rho(\xi) \cap S_T \qquad \text{for every } \xi \in \mathbb{R}^{N+1} \text{ and } \rho > 0.$$

Theorem 4.1. *Let $\mathcal{L}$ be an operator of type (1.1.1) satisfying assumptions (H1), (H2), (H3) stated in Section 1, for some $\alpha \in (0,1)$.*

Then, there exist constants $c, r_0 > 0$ depending on T, α, the matrix B in (1.1.5) and the numbers ν and Λ in (1.1.3) and (1.2.13), respectively, such that, for every point $\overline{\xi} \in S_T$, $r \leq r_0$ and $u \in \mathcal{S}^\alpha(S_T)$ with $\operatorname{supp}(u) \subseteq B_r(\overline{\xi}) \cap \overline{S_T}$, one has

$$\|\partial^2_{x_k x_h}u\|_{C_x^\alpha(B_r^T(\overline{\xi}))} \leq c\,|\mathcal{L}u|_{C_x^\alpha(B_r^T(\overline{\xi}))}, \tag{4.1.1}$$

for every $1 \leq h, k \leq q$. We stress that the constant c in (4.1.1) is independent of the ball $B_r(\overline{\xi})$.

Proof. Let $r \leq 1$ to be chosen later. For a fixed $\overline{\xi} = (\overline{x}, \overline{t})$, we consider the operator $\mathcal{L}_{\overline{x}}$ with coefficients $a_{ij}(\overline{x}, t)$ (frozen in space, variable in time). Let $\Gamma^{\overline{x}}$ be its fundamental solution, as described in Theorem 3.1. Let $u \in \mathcal{S}^\alpha(S_T)$ with $\operatorname{supp}(u) \subseteq B_r(\overline{\xi}) \cap \overline{S_T}$; then $\mathcal{L}_{\overline{x}}u \in C_x^\alpha(S_T)$ and, by Theorem 3.14, we can write

$$\partial^2_{x_k x_h} u(x,t) = \int_{\overline{t}-1}^{t} \left(\int_{\mathbb{R}^N} \partial^2_{x_i x_j} \Gamma^{\overline{x}}(x,t;y,s) \left[\mathcal{L}_{\overline{x}} u(E(s-t)x,s) - \mathcal{L}_{\overline{x}} u(y,s) \right] dy \right) ds,$$

for every $(x,t) \in B_r^T(\overline{\xi})$ (so that, in particular, $|t-\overline{t}| \le r \le 1$). Writing

$$\mathcal{L}_{\overline{x}} = \mathcal{L} + (\mathcal{L}_{\overline{x}} - \mathcal{L}),$$

we then have

$$\begin{aligned} \partial^2_{x_k x_h} u(x,t) &= \int_{\overline{t}-1}^{t} \left(\int_{\mathbb{R}^N} \partial^2_{x_k x_h} \Gamma^{\overline{x}}(x,t;y,s) \left[\mathcal{L} u(E(s-t)x,s) - \mathcal{L} u(y,s) \right] dy \right) ds \\ &\quad + \sum_{i,j=1}^{q} \int_{\overline{t}-1}^{t} \int_{\mathbb{R}^N} \partial^2_{x_k x_h} \Gamma^{\overline{x}}(x,t;y,s) \cdot \\ &\quad \cdot \Big\{ \left[a_{ij}(\overline{x},s) - a_{ij}(E(s-t)x,s) \right] \partial^2_{x_i x_j} u(E(s-t)x,s) \\ &\quad - \left[a_{ij}(\overline{x},s) - a_{ij}(y,s) \right] \partial^2_{x_i x_j} u(y,s) \Big\} dy ds \\ &\equiv A + \sum_{i,j=1}^{q} B_{ij}. \end{aligned}$$

For the term A we have, by Theorem 3.17,

$$\|A\|_{C^\alpha_x(S_T)} \le c\, |\mathcal{L}u|_{C^\alpha_x(S_T)}. \tag{4.1.2}$$

On the other hand,

$$B_{ij} = \int_{\overline{t}-1}^{t} \int_{\mathbb{R}^N} \partial^2_{x_k x_h} \Gamma^{\overline{x}}(x,t;y,s) \left[f_{ij}(E(s-t)x,s) - f_{ij}(y,s) \right] dy ds \tag{4.1.3}$$

with

$$f_{ij}(y,s) = [a_{ij}(\overline{x},s) - a_{ij}(y,s)] \partial^2_{x_i x_j} u(y,s),$$

hence, again by Theorem 3.17,

$$\|B_{ij}\|_{C^\alpha_x(S_T)} \le c\, |f_{ij}|_{C^\alpha_x(S_T)}.$$

We point out that the constant c in (4.1.2)-(4.1.3) is independent of the ball $B_r(\overline{\xi})$, since $\operatorname{supp}(u) \subseteq B_r(\overline{\xi}) \subseteq \{(x,t) : |t-\overline{t}| \le 1\}$, so that we can apply Theorem 3.17 with $T - \tau \le 2$.

We then turn to bound $|f_{ij}|_{C^\alpha_x(S_T)}$. We now exploit the fact that u has small support in space, namely $u(x,t) \ne 0$ only if $\|x - \overline{x}\| < r$; therefore we can assume that $\|x_k - \overline{x}\| < r$ for $k = 1,2$. Hence, we have

$$\begin{aligned} f_{ij}(x_1,s) - f(x_2,s) &= [a_{ij}(\overline{x},s) - a_{ij}(x_1,s)]\, \partial^2_{x_i x_j} u(x_1,s) - [a_{ij}(\overline{x},s) - a_{ij}(x_2,s)]\, \partial^2_{x_i x_j} u(x_2,s) \\ &= [a_{ij}(x_2,s) - a_{ij}(x_1,s)]\, \partial^2_{x_i x_j} u(x_1,s) \\ &\quad + [a_{ij}(\overline{x},s) - a_{ij}(x_2,s)]\, [\partial^2_{x_i x_j} u(x_1,s) - \partial^2_{x_i x_j} u(x_2,s)]. \end{aligned}$$

Then, writing briefly $|\cdot|_\alpha$ for $|\cdot|_{C_x^\alpha(S_T)}$

$$|f_{ij}(x_1,s)-f(x_2,s)|\leq |a_{ij}|_\alpha \|x_2-x_1\|^\alpha\cdot\sup|\partial^2_{x_ix_j}u|+|a_{ij}|_\alpha r^\alpha|\partial^2_{x_ix_j}u|_\alpha\|x_2-x_1\|^\alpha$$

so that

$$|\partial^2_{x_kx_h}u|_\alpha+\sup|\partial^2_{x_kx_h}u|\leq c\,|\mathcal{L}u|_\alpha+c\big\{|a_{ij}|_\alpha\sup|\partial^2_{x_ix_j}u|+|a_{ij}|_\alpha r^\alpha|\partial^2_{x_ix_j}u|_\alpha\big\}.$$

Exploiting again the fact that u has compact support, we have

$$\sup_{B_r(\overline{\xi})}|\partial^2_{x_ix_j}u|\leq|\partial^2_{x_ix_j}u|_\alpha(cr)^\alpha,$$

so that

$$|\partial^2_{x_kx_h}u|_\alpha+\sup|\partial^2_{x_kx_h}u|\leq c\,|\mathcal{L}u|_\alpha+c\,|a_{ij}|_\alpha r^\alpha|\partial^2_{x_ix_j}u|_\alpha,$$

and for r small enough we get (4.1.1). Note that the small number r and the constant c are independent of the fixed point $\overline{x}$. The independence of the constant on $\overline{x}$ also relies on the uniformity (in $\overline{x}$) of the upper bounds on $\partial^2_{x_ix_j}\Gamma^{\overline{x}}$. Actually these bounds depend on the coefficients $a_{ij}(x,t)$ only through the number ν. □

4.2. Some interpolation inequalities

Interpolation inequalities are a typical tool to deduce global estimates starting with local estimates for compactly supported functions. We will need the following:

Theorem 4.2. *For every $r>0$ there exist $c>0$ and $\gamma>1$ such that for every $\varepsilon\in(0,1)$, $\overline{\xi}\in S_T$ and $u\in\mathcal{S}^0(S_T)$,*

$$\sum_{h=1}^q\|\partial_{x_h}u\|_{C^\alpha(B_r^T(\overline{\xi}))}+\|u\|_{C^\alpha(B_r^T(\overline{\xi}))}$$
$$\leq\varepsilon\left\{\sum_{h,k=1}^q\left\|\partial^2_{x_kx_h}u\right\|_{C^0(B_{4r}^T(\overline{\xi}))}+\|Yu\|_{C^0(B_{4r}^T(\overline{\xi}))}\right\}+\frac{c}{\varepsilon^\gamma}\|u\|_{C^0(B_{4r}^T(\overline{\xi}))}. \tag{4.2.1}$$

The proof of the above inequality will be reached in several steps. The first step is based on the analysis of fractional integral operators carried out in Proposition 2.4 and has an independent interest, since it contains a regularity result for functions in $\mathcal{S}^0(S_T)$.

Proposition 4.3. (i) *Let $\mathcal{L}_0$ be the constant-coefficient operator*

$$\mathcal{L}_0=\textstyle\sum_{i=1}^q\partial^2_{x_ix_i}+Y$$

and let $R>0$ be fixed. For every $\alpha\in(0,1)$ there exists $\gamma>2$ and $c>0$ such that, for every $\overline{\xi}\in S_T, u\in\mathcal{S}^0(S_T)$ with $\operatorname{supp}(u)\subseteq B_R(\overline{\xi})\cap\overline{S_T}$ and every $\varepsilon\in(0,1)$ we have:

$$\|\partial_{x_k}u\|_{C^\alpha(B_R^T(\overline{\xi}))}+\|u\|_{C^\alpha(B_R^T(\overline{\xi}))}\leq\varepsilon\,\|\mathcal{L}_0u\|_{C^0(B_R^T(\overline{\xi}))}+\frac{c}{\varepsilon^\gamma}\|u\|_{C^0(B_R^T(\overline{\xi}))}\qquad \text{for } k=1,2,...,q.$$

(*The constant c depends on r and α but not on $\overline{\xi},u$ and ε.*)

(ii) *Let* $u \in \mathcal{S}^0(S_T)$, $\overline{\xi} \in S_T$ *and* $R > 0$. *Then, we have*

$$u,\, \partial_{x_k} u \in C^\alpha(B_R^T(\overline{\xi})) \text{ for every } 1 \le k \le q.$$

Proof. Point (ii) will simply follow applying point (i) with $\varepsilon = 1$ to the function $u\phi$, where $\phi \in C_0^\infty(B_{2R}(\overline{\xi}))$ and $\phi \equiv 1$ on $B_R(\overline{\xi})$. So, let us prove (i). This proof is inspired to [3, Prop. 7.1].

Let Γ^0 be the fundamental solution of $\mathcal{L}_0$ and let us write

$$u(\xi) = \int \Gamma^0(\xi,\eta)\, \mathcal{L}_0 u(\eta)\, d\eta$$
$$\partial_{x_k} u(\xi) = \int \partial_{x_k}\Gamma^0(\xi,\eta)\, \mathcal{L}_0 u(\eta)\, d\eta.$$

For a fixed $\varepsilon > 0$ (that we can assume $< \min(1, R)$) let $k_\varepsilon(\xi,\eta)$ a cutoff function such that

$$B_{\varepsilon/2}(\xi) \prec k_\varepsilon(\xi,\cdot) \prec B_\varepsilon(\xi).$$

We will prove the desired bound for $|\partial_{x_k} u|_{C^\alpha(B_R(\overline{\xi}))}$. A completely analogous proof, starting from the above representation formula for $u(\xi)$, gives an analogous bound for $|u|_{C^\alpha(B_R(\overline{\xi}))}$, possibly with a different exponent γ in the constant c/ε^γ. Since $\varepsilon \in (0,1)$, the assertion then follows choosing the bigger exponent.

Let us write

$$\begin{aligned}
\partial_{x_k} u(\xi) &= \int \partial_{x_k}\Gamma^0(\xi,\eta)\, k_\varepsilon(\xi,\eta)\, \mathcal{L}_0 u(\eta)\, d\eta \\
&\qquad + \int \partial_{x_k}\Gamma^0(\xi,\eta)\,[1 - k_\varepsilon(\xi,\eta)]\, \mathcal{L}_0 u(\eta)\, d\eta \\
&= \int \partial_{x_k}\Gamma^0(\xi,\eta)\, k_\varepsilon(\xi,\eta)\, \mathcal{L}_0 u(\eta)\, d\eta \\
&\qquad + \int (\mathcal{L}_0^*)^\eta \left(\partial_{x_k}\Gamma^0(\xi,\eta)\,[1 - k_\varepsilon(\xi,\eta)]\right) u(\eta)\, d\eta \\
&= T_1(\mathcal{L}_0 u) + T_2(u)
\end{aligned} \tag{4.2.2}$$

where

$$\mathcal{L}_0^* = \sum_{i=1}^q \partial^2_{x_i x_i} - Y.$$

Now we handle T_1 as a fractional integral. Since the kernel

$$K_1(\xi,\eta) = \partial_{x_k}\Gamma^0(\xi,\eta)\, k_\varepsilon(\xi,\eta)$$

does not vanish only if $d(\xi,\eta) < \varepsilon$, owing to Theorems 3.5-3.9 we see that, for every $\delta \in (0,1)$, the kernel K_1 satisfies the bounds

$$\begin{aligned}
|K_1(\xi,\eta)| &\le \frac{c}{d(\xi,\eta)^{Q+1}} \le \frac{c\varepsilon^\delta}{d(\xi,\eta)^{Q+1+\delta}} \\
|K_1(\xi_1,\eta) - K_1(\xi_2,\eta)| &\le c\frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+2}} \le c\varepsilon^\delta \frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+2+\delta}} \\
&\qquad \text{when } d(\xi_1,\eta) > 4\boldsymbol{\kappa} d(\xi_1,\xi_2).
\end{aligned}$$

For a fixed $\alpha \in (0,1)$, choosing $\delta < 1-\alpha$, by Proposition 2.4 (applied by extending our functions equal to 0 out of S_T) we get

$$\|T_1(\mathcal{L}_0 u)\|_{C^\alpha(B_R^T(\overline{\xi}))} \leq c(R)\,\varepsilon^\delta \|\mathcal{L}_0 u\|_{C^0(B_R^T(\overline{\xi}))}. \tag{4.2.3}$$

As to $T_2(u)$, let us consider the kernel

$$K_2(\xi,\eta) = \mathcal{L}_0^*\left(\partial_{x_k}\Gamma^0(\xi,\cdot)\left[1-k_\varepsilon(\xi,\cdot)\right]\right)(\eta). \tag{4.2.4}$$

We now claim that the kernel $K_2(\xi,\eta)$ satisfies the following fractional integral estimates:

$$|K_2(\xi,\eta)| \leq \frac{c}{\varepsilon^4}\frac{1}{d(\xi,\eta)^{Q-1}} \tag{4.2.5}$$

$$|K_2(\xi_1,\eta)-K_2(\xi_2,\eta)| \leq \frac{c}{\varepsilon^4}\frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^Q} \text{ for } d(\xi_1,\eta) > 4\boldsymbol{\kappa} d(\xi_1,\xi_2). \tag{4.2.6}$$

These bounds will be proved in Lemma 4.4. Taking these bounds for granted, by Proposition 2.4 we get

$$\|T_2(u)\|_{C^\alpha(B_R^T(\overline{\xi}))} \leq \frac{c(R)}{\varepsilon^4}\|u\|_{C^0(B_R^T(\overline{\xi}))}$$

and then, by (4.2.2) and (4.2.3), for some constants c_1, c_2 depending on R but independent of $\varepsilon, \overline{\xi}$ and u,

$$\|\partial_{x_k} u\|_{C^\alpha(B_R^T(\overline{\xi}))} \leq c_1\varepsilon^\delta \|\mathcal{L}_0 u\|_{C^0(B_R^T(\overline{\xi}))} + \frac{c_2}{\varepsilon^4}\|u\|_{C^0(B_R^T(\overline{\xi}))}.$$

Rescaling $c_1\varepsilon^\delta = \varepsilon_1$ we get

$$\|\partial_{x_k} u\|_{C^\alpha(B_R^T(\overline{\xi}))} \leq \varepsilon_1 \|\mathcal{L}_0 u\|_{C^0(B_R^T(\overline{\xi}))} + \frac{c}{\varepsilon_1^{4/\delta}}\|u\|_{C^0(B_R^T(\overline{\xi}))}$$

for some c depending on R but not on ε_1. So the assertion is proved, with $\gamma = 4/\delta$ and some fixed $\delta \in (0,1)$.

The analogous bound on $\|u\|_{C^\alpha(B_R^T(\overline{\xi}))}$ can be proved, with a completely analogous reasoning, starting with the representation formula

$$\begin{aligned} u(\xi) &= \int \Gamma^0(\xi,\eta)\,\mathcal{L}_0 u(\eta)\,d\eta \\ &= \int \Gamma^0(\xi,\eta)\,k_\varepsilon(\xi,\eta)\,\mathcal{L}_0 u(\eta)\,d\eta + \int (\mathcal{L}_0^*)^\eta\left(\Gamma^0(\xi,\eta)\left[1-k_\varepsilon(\xi,\eta)\right]\right)u(\eta)\,d\eta \\ &= T_1'(\mathcal{L}_0 u) + T_2'(u) \end{aligned}$$

where T_1', T_2' are fractional integral operators with kernels K_1', K_2', respectively, satisfying the following bounds:

$$|K_1'(\xi,\eta)| \leq \frac{c}{d(\xi,\eta)^Q} \leq \frac{c\varepsilon^\delta}{d(\xi,\eta)^{Q+\delta}}$$

$$|K_1'(\xi_1,\eta)-K_1'(\xi_2,\eta)| \leq c\frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+1}} \leq c\varepsilon^\delta \frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+1+\delta}}$$

$$\text{when } d(\xi_1,\eta) > 4\boldsymbol{\kappa}(\xi_1,\xi_2)$$

$$|K_2'(\xi,\eta)| \leq \frac{c}{\varepsilon^4}\frac{1}{d(\xi,\eta)^{Q-2}}$$

$$|K_2'(\xi_1,\eta) - K_2'(\xi_1,\eta)| \le \frac{c}{\varepsilon^4}\frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q-1}}$$
$$\text{when } d(\xi_1,\eta) > 4\boldsymbol{\kappa} d(\xi_1,\xi_2)..$$

The bounds on K_1' are immediate, while those on K_2' can be proved with the same reasoning used in the proof of Lemma 4.4 here below, exploiting the corresponding upper bounds on the derivatives of Γ^0. The upper bound on $\|u\|_{C^\alpha(B_R^T(\overline{\xi}))}$ leads to an exponent γ' possibly different from the exponent γ found in the bound on $\|\partial_{x_k}u\|_{C^\alpha(B_R^T(\overline{\xi}))}$, but since $\varepsilon \in (0,1)$ it is enough to choose $\max(\gamma,\gamma')$. □

Lemma 4.4. *For every $\varepsilon \in (0,1)$, the kernel K_2 defined in (4.2.4) satisfies the bounds (4.2.5)-(4.2.6).*

Proof. Recalling that $\mathcal{L}_0^*\left(\Gamma^0(\xi,\cdot)\right) = 0$ and the x and y derivatives of Γ_0 commute, we have:

$$\begin{aligned} K_2(\xi,\eta) &= \left(\sum_{i=1}^q \partial^2_{y_iy_i} - Y^{(y,s)}\right)\left(\partial_{x_k}\Gamma^0((x,t),(y,s))\left[1-k_\varepsilon((x,t),(y,s))\right]\right) \\ &= \partial_{x_k}\Gamma^0((x,t),(y,s))\left(\sum_{i=1}^q \partial^2_{y_iy_i} - Y^{(y,s)}\right)\left[1-k_\varepsilon((x,t),(y,s))\right] \\ &+ 2\sum_{i=1}^q \partial^2_{x_ky_i}\Gamma^0((x,t),(y,s))\left[1-k_\varepsilon((x,t),(y,s))\right]_{y_i}. \end{aligned} \tag{4.2.7}$$

Exploiting the growth estimates of $\Gamma^0_{x_k}, \Gamma^0_{x_ky_i}$ (see Theorem 3.5) we get

$$|K_2(\xi,\eta)| \le \frac{c}{d(\xi,\eta)^{Q+1}}\frac{c}{\varepsilon^2} + \frac{c}{d(\xi,\eta)^{Q+2}}\frac{c}{\varepsilon} \le \frac{c}{\varepsilon^4}\frac{1}{d(\xi,\eta)^{Q-1}}$$

since $K_2(\xi,\eta)$ vanishes for $d(\xi,\eta) < \varepsilon/2$. So we have (4.2.5). In order to prove (4.2.6) we are going to bound $\partial_{x_h}K_2$ for $h = 1,2,...,q$ and $Y^{(x,t)}K_2$ and then apply the mean value theorem with respect to the vector fields. Note that the operator $\mathcal{L}_0$ has *smooth* coefficients, independent of t.

$$\begin{aligned} |\partial_{x_h}K_2(\xi,\eta)| &\le \frac{c}{d(\xi,\eta)^{Q+2}}\frac{c}{\varepsilon^2} + \frac{c}{d(\xi,\eta)^{Q+1}}\frac{c}{\varepsilon^3} + \frac{c}{d(\xi,\eta)^{Q+3}}\frac{c}{\varepsilon} \\ &\le \frac{c}{d(\xi,\eta)^{Q+1}}\frac{c}{\varepsilon^3} \le \frac{c}{\varepsilon^4}\frac{1}{d(\xi,\eta)^{Q}} \end{aligned} \tag{4.2.8}$$

since $K_2(\xi,\eta)$ vanishes for $d(\xi,\eta) < \varepsilon/2$. Moreover, using the bounds for $YD_x^{\boldsymbol{\alpha}}\Gamma$ established in the proof of Theorem 3.9, we have

$$|Y^{(x,t)}\partial_{x_k}\Gamma^0| \le \frac{c}{d(\xi,\eta)^{Q+3}} \quad \text{and} \quad |Y^{(x,t)}\partial^2_{x_ky_i}\Gamma^0| \le \frac{c}{d(\xi,\eta)^{Q+4}}.$$

Therefore, by (4.2.7) we obtain

$$\begin{aligned} \left|Y^{(x,t)}K_2(\xi,\eta)\right| &\le \frac{c}{d(\xi,\eta)^{Q+3}}\frac{c}{\varepsilon^2} + \frac{c}{d(\xi,\eta)^{Q+1}}\frac{c}{\varepsilon^4} \\ &\quad + \frac{c}{d(\xi,\eta)^{Q+4}}\frac{c}{\varepsilon} + \frac{c}{d(\xi,\eta)^{Q+2}}\frac{c}{\varepsilon^3} \\ &\le \frac{c}{d(\xi,\eta)^{Q+1}}\frac{c}{\varepsilon^4} \end{aligned} \tag{4.2.9}$$

where we have used again the vanishing of $K_2(\xi,\eta)$ for $d(\xi,\eta)<\varepsilon/2$.

Hence, by Lagrange' theorem (Theorem 2.1), (4.2.8)-(4.2.9) imply (4.2.6). This completes the proof of Lemma 4.4 and therefore of Proposition 4.3. □

The second ingredient of the proof of Theorem 4.2 is the following inequality, which seems a standard Euclidean result. The only difference is that the norms are based on *metric* balls.

Proposition 4.5. *For every $r>\varepsilon>0$ and $u\in C^0(\overline{B_{2r}(\overline{\xi})\cap S_T})$ possessing continuous derivatives $\partial_{x_h}u$ and $\partial^2_{x_hx_h}u$ in $\overline{B_{2r}(\overline{\xi})\cap S_T}$ for some $1\le h\le q$, we have*

$$\|\partial_{x_h}u\|_{C^0(B_r^T(\overline{\xi}))}\le\varepsilon\|\partial^2_{x_hx_h}u\|_{C^0(B_{2r}^T(\overline{\xi}))}+\frac{2}{\varepsilon}\|u\|_{C^0(B_{2r}^T(\overline{\xi}))}.$$

Proof. For a fixed $\xi\in B_r^T\left(\overline{\xi}\right)$, let

$$f(t)=u(\xi+t\varepsilon e_h)\text{ for }t\in[0,1],$$

with e_h the h-th unit vector. Then the identity

$$f(1)-f(0)=f'(0)+\int_0^1(1-s)f''(s)\,ds$$

gives

$$u(\xi+\varepsilon e_h)-u(\xi)=\varepsilon\partial_{x_h}u(\xi)+\varepsilon^2\int_0^1(1-s)\,\partial^2_{x_hx_h}u(\xi+s\varepsilon e_h)\,ds.$$

Moreover, for $\xi=(x,t)$ ranging in $B_r^T\left(\overline{\xi}\right)$ and $\varepsilon<r$, $s\in(0,1)$, we claim that

$$\xi+s\varepsilon e_h\in B_{2r}^T\left(\overline{\xi}\right).$$

This fact is not trivial because B_r are not Euclidean balls but balls w.r.t. the quasidistance d. Let us compute

$$\begin{aligned}d\left(\xi+s\varepsilon e_h,\overline{\xi}\right)&=\left\|x+s\varepsilon e_h-E\left(t-\overline{t}\right)\overline{x}\right\|+\sqrt{\left|t-\overline{t}\right|}\\&=\sum_{i\neq h}\left|\left(x-E\left(t-\overline{t}\right)\overline{x}\right)_i\right|^{1/q_i}+\left|\left(x-E\left(t-\overline{t}\right)\overline{x}\right)_h+s\varepsilon\right|+\sqrt{\left|t-\overline{t}\right|}\\&\le\left(\sum_i\left|\left(x-E\left(t-\overline{t}\right)\overline{x}\right)_i\right|^{1/q_i}+\sqrt{\left|t-\overline{t}\right|}\right)+\varepsilon<2r,\end{aligned}$$

where we have exploited the fact that the h-th variable (for $h=1,2,...,q$) has homogeneity 1. Hence $\xi+s\varepsilon e_h\in B_{2r}(\overline{\xi})$. Note also that ξ and $\xi+s\varepsilon e_h$ have the same t-component. Therefore

$$\varepsilon\sup_{B_r^T}|\partial_{x_h}u|\le 2\sup_{B_{2r}^T}|u|+\varepsilon^2\sup_{B_{2r}^T}\left|\partial^2_{x_hx_h}u\right|$$

as desired. □

We can now come to the

Proof of Theorem 4.2. Let us first prove the result under the additional assumption that $u \in C^{2,\alpha}(\overline{B_{4r}^T(\overline{\xi})})$. Let $\mathcal{L}_0$ be as in Proposition 4.3, and choose $\phi \in C_0^\infty\left(B_{2r}\left(\overline{\xi}\right)\right)$ with $\phi = 1$ in $B_r\left(\overline{\xi}\right)$. To be more precise, we can fix a "mother function" $\Phi \in C_0^\infty\left(B_R\left(0\right)\right)$ with $\Phi = 1$ in $B_{R/2}\left(0\right)$ and define $\phi\left(\xi\right) = \Phi(\overline{\xi}^{-1} \circ \xi)$ so that, by left invariance of $\mathcal{L}_0$ and ∂_{x_h} $(h = 1, 2, .., q)$ the quantities

$$\|\phi\|_{C^0(B_{2r}(\overline{\xi}))}, \|\partial_{x_h}\phi\|_{C^0(B_{2r}(\overline{\xi}))}, \|\mathcal{L}_0\phi\|_{C^0(B_{2r}(\overline{\xi}))}$$

do not depend on the centre $\overline{\xi}$ of the ball (they will depend on r, which however is fixed).

Applying Proposition 4.3 to $u\phi$ we get, for every $\varepsilon \in (0, 1), h = 1, 2, ..., q$,

$$\begin{aligned}
\|\partial_{x_h}u\|_{C^\alpha(B_r^T(\overline{\xi}))} + \|u\|_{C^\alpha(B_r^T(\overline{\xi}))} &\le \|\partial_{x_h}(u\phi)\|_{C^\alpha(B_{2r}^T(\overline{\xi}))} + \|u\phi\|_{C^\alpha(B_{2r}^T(\overline{\xi}))} \\
&\le \varepsilon\|\mathcal{L}_0(u\phi)\|_{C^0(B_{2r}^T(\overline{\xi}))} + \frac{c}{\varepsilon^\gamma}\|u\phi\|_{C^0(B_{2r}^T(\overline{\xi}))} \\
&\le c\varepsilon\Bigg\{ \sum_{h,k=1}^q \|\partial^2_{x_k x_h}u\|_{C^0(B_{2r}^T(\overline{\xi}))} + \|Yu\|_{C^0(B_{2r}^T(\overline{\xi}))} \\
&\qquad + \sum_{h=1}^q \|\partial_{x_h}u\|_{C^0(B_{2r}^T(\overline{\xi}))} + \|u\|_{C^0(B_{2r}^T(\overline{\xi}))}\Bigg\} \\
&\quad + \frac{c}{\varepsilon^\gamma}\|u\|_{C^0(B_{2r}^T(\overline{\xi}))}
\end{aligned}$$

(with c independent of $\overline{\xi}$, by the construction of ϕ), by Proposition 4.5 (applied with $\varepsilon = 1$), which can be applied because we are assuming $u \in C^{2,\alpha}(\overline{B_{4r}^T(\overline{\xi})})$,

$$\begin{aligned}
&\le c\varepsilon\left\{ \sum_{h,k=1}^q \left\|\partial^2_{x_k x_h}u\right\|_{C^0(B_{4r}^T(\overline{\xi}))} + \|Yu\|_{C^0(B_{2r}^T(\overline{\xi}))} + \|u\|_{C^0(B_{4r}^T(\overline{\xi}))}\right\} + \frac{c}{\varepsilon^\gamma}\|u\|_{C^0(B_{2r}^T(\overline{\xi}))} \\
&\le c\varepsilon\left\{ \sum_{h,k=1}^q \left\|\partial^2_{x_k x_h}u\right\|_{C^0(B_{4r}^T(\overline{\xi}))} + \|Yu\|_{C^0(B_{4r}^T(\overline{\xi}))}\right\} + \frac{c}{\varepsilon^\gamma}\|u\|_{C^0(B_{4r}^T(\overline{\xi}))}.
\end{aligned}$$

Next, let $u \in \mathcal{S}^0(S_T)$, extend it to 0 out of S_T and define its mollified version u_δ as in the proof of Theorem 3.11. Then u_δ satisfies (4.2.1), however ∂_{x_k} and Y commute with the mollification, so that

$$\begin{aligned}
&\sum_{h=1}^q \left\|(\partial_{x_h}u)_\delta\right\|_{C^\alpha(B_r^T(\overline{\xi}))} + \|u_\delta\|_{C^\alpha(B_r^T(\overline{\xi}))} \\
&\le \varepsilon\left\{ \sum_{h,k=1}^q \left\|\left(\partial^2_{x_k x_h}u\right)_\delta\right\|_{C^0(B_{4r}^T(\overline{\xi}))} + \left\|(Yu)_\delta\right\|_{C^0(B_{4r}^T(\overline{\xi}))}\right\} + \frac{c}{\varepsilon^\gamma}\|u_\delta\|_{C^0(B_{4r}^T(\overline{\xi}))} \\
&\le \varepsilon\left\{ \sum_{h,k=1}^q \left\|\partial^2_{x_k x_h}u\right\|_{C^0(B_{4r}^T(\overline{\xi}))} + \|Yu\|_{C^0(B_{4r}^T(\overline{\xi}))}\right\} + \frac{c}{\varepsilon^\gamma}\|u\|_{C^0(B_{4r}^T(\overline{\xi}))}.
\end{aligned}$$

We already know that u_δ uniformly converges to u, which is a priori continuous, on $S_{T-\varepsilon_0}$ for every $\varepsilon_0 > 0$. The uniform bound on $\|(\partial_{x_h}u)_\delta\|_{C^\alpha(B_r^T(\overline{\xi}))}, \|u_\delta\|_{C^\alpha(B_r^T(\overline{\xi}))}$ implies that the functions u_δ, $(\partial_{x_h}u)_\delta$ are equicontinuous and equibounded, then by Ascoli-Arzelà's theorem we can extract a sequence $(\partial_{x_k}u)_\delta$ uniformly converging to some function v_k which must coincide with $\partial_{x_k}u$.

The uniform convergence allows to get the bound

$$\sum_{h=1}^{q} \|\partial_{x_h} u\|_{C^\alpha(B_r^{T-\varepsilon_0}(\bar{\xi}))} + \|u\|_{C^\alpha(B_r^{T-\varepsilon_0}(\bar{\xi}))}$$
$$\leq \varepsilon \left\{ \sum_{h,k=1}^{q} \left\|\partial^2_{x_k x_h} u\right\|_{C^0(B_{4r}^T(\bar{\xi}))} + \|Yu\|_{C^0(B_{4r}^T(\bar{\xi}))} \right\} + \frac{c}{\varepsilon^\gamma} \|u\|_{C^0(B_{4r}^T(\bar{\xi}))}.$$

Since this holds for every $\varepsilon_0 > 0$ with a constant c independent of ε_0, we obtain (4.2.1), and we are done. □

4.3. Global Schauder estimates in space

Here we want to get global Schauder estimates on the strip S_T, starting with the local Schauder estimates proved in Theorem 4.1 for functions which are compactly supported on small balls. To this aim, we will basically make use of cutoff functions and the interpolation inequalities proved in the previous section.

We start with a brief discussion about how a control of $C_x^\alpha(S_T)$-norm can be obtained starting with the control of $C_x^\alpha(B_r(\xi_i))$-norms for a suitable family of balls $\{B_r(\xi_i)\}_i$.

Let us start defining, for some fixed small $r > 0$, the seminorms

$$|f|_{C^\alpha_{x,r}(S_T)} \equiv \sup_{\substack{(x_1,t),(x_2,t)\in S_T \\ 0<\|x_1-x_2\|\leq r}} \frac{|f(x_1,t) - f(x_2,t)|}{\|x_1 - x_2\|^\alpha}$$
$$|f|_{C^\alpha_{r}(S_T)} \equiv \sup_{\substack{\xi_1,\xi_2\in S_T \\ 0<d(\xi_1,\xi_2)\leq r}} \frac{|f(\xi_1) - f(\xi_2)|}{d(\xi_1,\xi_2)^\alpha}$$

and let

$$\|f\|_{C^\alpha_{x,r}(S_T)} = |f|_{C^\alpha_{x,r}(S_T)} + \|f\|_{C^0(S_T)}.$$

Then the following holds:

Proposition 4.6. *Let $r > 0$ and $\alpha \in (0,1)$ be fixed, then:*

(i) *There exists $c > 0$, depending on α and r, such that*

$$\|f\|_{C^\alpha_x(S_T)} \leq c \|f\|_{C^\alpha_{x,r}(S_T)}. \tag{4.3.1}$$

(ii) *Moreover, let $\left\{B_r\left(\bar{\xi}_i\right)\right\}_{i=1}^\infty$ be a covering of S_T, then*

$$|f|_{C^\alpha_{x,r}(S_T)} \leq \sup_i |f|_{C^\alpha_x(B^T_{\theta r}(\bar{\xi}_i))} \tag{4.3.2}$$
$$|f|_{C^\alpha_{r}(S_T)} \leq \sup_i |f|_{C^\alpha(B^T_{\theta r}(\bar{\xi}_i))} \tag{4.3.3}$$

where $\theta \geq 1$ is an absolute constant.

Proof. (i) Noting that

$$\sup_{\substack{(x_1,t),(x_2,t)\in S_T \\ \|x_1-x_2\|>r}} \frac{|f(x_1,t) - f(x_2,t)|}{\|x_1 - x_2\|^\alpha} \leq \frac{2}{r^\alpha} \|f\|_{C^0(S_T)}$$

we immediately derive

$$|f|_{C_x^\alpha(S_T)} \leq \max\left(|f|_{C_{x,r}^\alpha(S_T)}, \frac{2}{r^\alpha}\|f\|_{C^0(S_T)}\right)$$

which in turn implies (4.3.1).

(ii) Next, for any two points $(x_1,t),(x_2,t)\in S_T$ such that $\|x_1-x_2\|\leq r$, let $(x_1,t)\in B_r(\overline{\xi}_{i_1})$ for some i_1. Then $(x_2,t)\in B_{\theta r}(\overline{\xi}_{i_1})$ for some absolute $\theta\geq 1$, and hence

$$\frac{|f(x_1,t)-f(x_2,t)|}{\|x_1-x_2\|^\alpha}\leq |f|_{C_x^\alpha(B_{\theta r}^T(\overline{\xi}_{i_1}))}.$$

Therefore

$$\sup_{\substack{(x_1,t),(x_2,t)\in S_T\\ 0<\|x_1-x_2\|\leq r}}\frac{|f(x_1,t)-f(x_2,t)|}{\|x_1-x_2\|^\alpha}\leq \sup_i |f|_{C_x^\alpha(B_{\theta r}^T(\overline{\xi}_i))},$$

which is (4.3.2). Analogously one can prove (4.3.3). □

We are now ready for

Theorem 4.7 *(Global Schauder estimates). Let $\mathcal{L}$ be the operator* (1.1.1) *in S_T and assume* (H1), (H2), (H3) *hold, for some $\alpha\in(0,1)$.*

Then, there exists a constant $c>0$, depending on T, α, the matrix B in (1.1.5) *and the numbers ν and Λ in* (1.1.3), (1.2.13), *respectively, such that*

$$\sum_{h,k=1}^q \left\|\partial^2_{x_h x_k}u\right\|_{C_x^\alpha(S_T)}+\|Yu\|_{C_x^\alpha(S_T)}+\sum_{k=1}^q\|\partial_{x_k}u\|_{C^\alpha(S_T)}+\|u\|_{C^\alpha(S_T)}$$
$$\leq c\left\{\|\mathcal{L}u\|_{C_x^\alpha(S_T)}+\|u\|_{C^0(S_T)}\right\}$$

for every $u\in\mathcal{S}^\alpha(S_T)$.

Proof. For a fixed $r>0$, small enough so that the local Schauder estimates of Theorem 4.1 hold on balls of radius $2\theta r$ (with $\theta\geq 1$ as in Proposition 4.6), let $\{B_r(\overline{\xi}_i)\}_{i=1}^\infty$ be a covering of S_T.

Let $\Phi\in C_0^\infty(B_{2\theta r}(0))$ such that $\Phi\equiv 1$ in $B_{\theta r}(0)$, and let $\phi_i(\xi)=\Phi(\overline{\xi}_i^{-1}\circ\xi)$, so that $\phi_i\in C_0^\infty\left(B_{2\theta r}\left(\overline{\xi}_i\right)\right),\phi_i=1$ in $B_{\theta r}\left(\overline{\xi}_i\right)$. Moreover, by construction of ϕ_i and left invariance of Y and ∂_{x_k} for $k=1,2,...,q$, the C^α norms of $\phi_i,\partial_{x_k}\phi_i,\mathcal{L}(\phi_i)$ are bounded independently of i. Throughout this proof the constants involved may depend on r, which however is by now fixed.

To begin with, applying Theorem 4.1 to $u\phi_i$ on $B_{2\theta r}(\overline{\xi}_i)$ we have

$$\left\|\partial^2_{x_k x_h}u\right\|_{C_x^\alpha(B_{\theta r}^T(\overline{\xi}_i))}\leq\left\|\partial^2_{x_k x_h}(u\phi_i)\right\|_{C_x^\alpha(B_{2\theta r}^T(\overline{\xi}_i))}\leq c\,|\mathcal{L}(u\phi_i)|_{C_x^\alpha(B_{2\theta r}^T(\overline{\xi}_i))},\tag{4.3.4}$$

where the constant $c>0$ is independent of the ball. On the other hand,

$$\mathcal{L}(u\phi_i)=(\mathcal{L}u)\,\phi_i+u(\mathcal{L}\phi_i)+2\sum_{h,k=1}^q a_{hk}\partial_{x_h}u\cdot\partial_{x_k}\phi_i$$

hence, for some constant c independent of $\overline{\xi}_i$,

$$|\mathcal{L}(u\phi_i)|_{C_x^\alpha(B_{2\theta r}^T(\overline{\xi}_i))} \leq c\left\{\|\mathcal{L}u\|_{C_x^\alpha(B_{2\theta r}^T(\overline{\xi}_i))} + \sum_{h=1}^q \|\partial_{x_h}u\|_{C_x^\alpha(B_{2\theta r}^T(\overline{\xi}_i))} + \|u\|_{C_x^\alpha(B_{2\theta r}^T(\overline{\xi}_i))}\right\} \tag{4.3.5}$$

Inserting (4.3.5) in (4.3.4) and adding to both sides

$$\sum_{k=1}^q \|\partial_{x_k}u\|_{C^\alpha(B_{\theta r}^T(\overline{\xi}_i))} + \|u\|_{C^\alpha(B_{\theta r}^T(\overline{\xi}_i))}$$

(note that this quantity is finite by Proposition 4.3 (ii) since $u \in \mathcal{S}^0(S_T)$) we get:

$$\sum_{h,k=1}^q \left\|\partial^2_{x_kx_h}u\right\|_{C_x^\alpha(B_{\theta r}^T(\overline{\xi}_i))} + \sum_{k=1}^q \|\partial_{x_k}u\|_{C^\alpha(B_{\theta r}^T(\overline{\xi}_i))} + \|u\|_{C^\alpha(B_{\theta r}^T(\overline{\xi}_i))}$$

$$\leq c\left\{\|\mathcal{L}u\|_{C_x^\alpha(B_{2\theta r}^T(\overline{\xi}_i))} + \sum_{h=1}^q \|\partial_{x_h}u\|_{C^\alpha(B_{2\theta r}^T(\overline{\xi}_i))} + \|u\|_{C^\alpha(B_{2\theta r}^T(\overline{\xi}_i))}\right\}$$

by Theorem 4.2, for any $\varepsilon \in (0,1)$ (to be fixed later)

$$\leq c\left\{\|\mathcal{L}u\|_{C_x^\alpha(B_{2\theta r}^T(\overline{\xi}_i))} + \varepsilon\left[\sum_{h,k=1}^q \left\|\partial^2_{x_kx_h}u\right\|_{C^0(B_{8\theta r}^T(\overline{\xi}_i))} + \|Yu\|_{C^0(B_{8\theta r}^T(\overline{\xi}_i))}\right] + \frac{1}{\varepsilon^\gamma}\|u\|_{C^0(B_{8\theta r}^T(\overline{\xi}_i))}\right\}$$

from the equation $Yu = \mathcal{L}u - \sum_{h,k=1}^q a_{hk}\partial^2_{x_hx_k}u$

$$\leq c\left\{\|\mathcal{L}u\|_{C_x^\alpha(B_{2\theta r}^T(\overline{\xi}_i))}\right.$$

$$+\varepsilon\left[(1+c(\nu))\sum_{h,k=1}^q \left\|\partial^2_{x_kx_h}u\right\|_{C^0(B_{8\theta r}^T(\overline{\xi}_i))} + \|\mathcal{L}u\|_{C^0(B_{8\theta r}^T(\overline{\xi}_i))}\right]$$

$$\left.+\frac{1}{\varepsilon^\gamma}\|u\|_{C^0(B_{8\theta r}^T(\overline{\xi}_i))}\right\}$$

$$\leq c\left\{\|\mathcal{L}u\|_{C_x^\alpha(S_T)} + c_1\varepsilon\sum_{h,k=1}^q \left\|\partial^2_{x_kx_h}u\right\|_{C^0(S_T)} + \frac{1}{\varepsilon^\gamma}\|u\|_{C^0(S_T)}\right\}$$

We now fix $\varepsilon > 0$ small enough so that $cc_1\varepsilon \leq 1/2$, so that for every ball $B_r(\overline{\xi}_i)$ of the fixed covering we have

$$\sum_{h,k=1}^q \left\|\partial^2_{x_kx_h}u\right\|_{C_x^\alpha(B_{\theta r}^T(\overline{\xi}_i))} + \sum_{k=1}^q \|\partial_{x_k}u\|_{C^\alpha(B_{\theta r}^T(\overline{\xi}_i))} + \|u\|_{C^\alpha(B_{\theta r}^T(\overline{\xi}_i))}$$

$$\leq c\left\{\|\mathcal{L}u\|_{C_x^\alpha(S_T)} + \|u\|_{C^0(S_T)}\right\} + \frac{1}{2}\sum_{h,k=1}^q \left\|\partial^2_{x_kx_h}u\right\|_{C^0(S_T)}.$$

Finally, taking the supremum for $i = 1,2,3...$ we get, by (4.3.2)-(4.3.2)

$$\sum_{h,k=1}^q \left\|\partial^2_{x_kx_h}u\right\|_{C_{x,r}^\alpha(S_T)} + \sum_{k=1}^q \|\partial_{x_k}u\|_{C_r^\alpha(S_T)} + \|u\|_{C_r^\alpha(S_T)}$$

$$\leq c\left\{\|\mathcal{L}u\|_{C_x^\alpha(S_T)} + \|u\|_{C^0(S_T)}\right\} + \frac{1}{2}\sum_{h,k=1}^q \left\|\partial^2_{x_kx_h}u\right\|_{C^0(S_T)}$$

so that

$$\sum_{h,k=1}^{q}\left\|\partial^2_{x_k x_h}u\right\|_{C^\alpha_{x,r}(S_T)}+\sum_{k=1}^{q}\|\partial_{x_k}u\|_{C^\alpha_r(S_T)}+\|u\|_{C^\alpha_r(S_T)}\leq c\left\{\|\mathcal{L}u\|_{C^\alpha_x(S_T)}+\|u\|_{C^0(S_T)}\right\}$$

and by (4.3.1) we conclude

$$\left\|\partial^2_{x_k x_h}u\right\|_{C^\alpha_x(S_T)}+\sum_{k=1}^{q}\|\partial_{x_k}u\|_{C^\alpha(S_T)}+\|u\|_{C^\alpha(S_T)}\leq c\left\{\|\mathcal{L}u\|_{C^\alpha_x(S_T)}+\|u\|_{C^0(S_T)}\right\}. \tag{4.3.6}$$

Finally, from the equation $Yu=\mathcal{L}u-\sum_{h,k=1}^{q}a_{hk}\partial^2_{x_h x_k}u$ we also get

$$\|Yu\|_{C^\alpha_x(S_T)}\leq c\left\{\|\mathcal{L}u\|_{C^\alpha_x(S_T)}+\sum_{h,k=1}^{q}\left\|\partial^2_{x_h x_k}u\right\|_{C^\alpha_x(S_T)}\right\}$$

with c also depending on the Hölder norms of the coefficients a_{ij}, and by (4.3.6)

$$\leq c\left\{\|\mathcal{L}u\|_{C^\alpha_x(S_T)}+\|u\|_{C^0(S_T)}\right\}.$$

So we are done. □

4.4. Schauder estimates in space and time

For an arbitrary set $\Omega\subseteq\overline{S_T}$, let us define the seminorms:

$$|f|_{C^\alpha_t(\Omega)}=\sup_{\substack{(x_1,t_1),(x_2,t_2)\in\Omega\\(x_1,t_1)\neq(x_2,t_2)}}\frac{|f(x_1,t_1)-f(x_2,t_2)|}{d((x_1,t_1),(x_2,t_2))^\alpha+|t_1-t_2|^{\alpha/q_N}}$$

$$|f|_{C^\alpha_{t,r}(\Omega)}\equiv\sup_{\substack{(x_1,t_1),(x_2,t_2)\in\Omega\\0<d((x_1,t_1),(x_2,t_2))\leq r}}\frac{|f(x_1,t_1)-f(x_2,t_2)|}{d((x_1,t_1),(x_2,t_2))^\alpha+|t_1-t_2|^{\alpha/q_N}}.$$

Here the number q_N is the largest homogeneity exponent in the dilations, see (1.1.7). Let also:

$$\|f\|_{C^\alpha_t(\Omega)}=|f|_{C^\alpha_t(\Omega)}+\|f\|_{C^0(\Omega)}$$
$$\|f\|_{C^\alpha_{t,r}(\Omega)}=|f|_{C^\alpha_{t,r}(\Omega)}+\|f\|_{C^0(\Omega)}.$$

Then the following holds, with a proof perfectly analogous to that of Proposition 4.6:

Proposition 4.8. *Let $r>0$ and $\alpha\in(0,1)$ be fixed, then:*

(i) *There exists $c>0$, depending on α and r, such that*

$$\|f\|_{C^\alpha_t(\Omega)}\leq c\|f\|_{C^\alpha_{t,r}(\Omega)}. \tag{4.4.1}$$

(ii) *Moreover, let $\{B_r(\overline{\xi}_i)\}_{i=1}^\infty$ be a covering of Ω, then*

$$|f|_{C^\alpha_{t,r}(\Omega)}\leq\sup_i|f|_{C^\alpha_t(B^T_{\theta r}(\overline{\xi}_i))} \tag{4.4.2}$$

where $\theta\geq 1$ is an absolute constant.

We can now state our Hölder estimate in space and time:

Theorem 4.9. *Let $\mathcal{L}$ be the operator (1.1.1) in S_T and assume (H1), (H2), (H3) hold, for some $\alpha \in (0,1)$. For every $T > \tau > -\infty$ and every compact set $K \subset \mathbb{R}^N$ there exists $c > 0$ depending on $K, \tau, T, \alpha, B, \nu, \Lambda$ such that, for every $u \in \mathcal{S}^\alpha(S_T)$ the derivatives $\partial^2_{x_h x_k} u$ satisfy the following local Hölder continuity in space-time:*

$$\left|\partial^2_{x_i x_j} u\right|_{C_t^\alpha(K\times[\tau,T])} + \|\partial^2_{x_i x_j} u\|_{C^0(S_T)} \le c\left\{\|\mathcal{L}u\|_{C_x^\alpha(S_T)} + \|u\|_{C^0(S_T)}\right\}.$$

In particular, even the second derivatives $\partial^2_{x_i x_j} u$ are jointly continuous in S_T.

Proof. Fix a compact set $K \subset \mathbb{R}^N$, let $T > \tau > -\infty$ and let $\psi(t)$ be a smooth function such that $\psi(t) = 1$ for $t \ge \tau$, $\psi(t) = 0$ for $t \le \tau - 1$, $0 \le \psi(t) \le 1$.

For $\overline{\xi} = (\overline{x}, \overline{t})$, let us consider the frozen operator $\mathcal{L}_{\overline{x}}$ with coefficients $a_{ij}(\overline{x}, t)$. Applying Theorem 3.18 to the operator $\mathcal{L}_{\overline{x}}$ we get the existence of a constant, depending on $K, \tau, T, \alpha, B, \nu$ but not on $\overline{\xi}$, such that for every $u \in \mathcal{S}^\alpha(S_T)$, since $u\psi \in \mathcal{S}^0(\tau - 1, T)$,

$$\left|\partial^2_{x_i x_j} u(x_1,t_1) - \partial^2_{x_i x_j} u(x_2,t_2)\right| \le c\left|\mathcal{L}_{\overline{x}}(u\psi)\right|_{C_x^\alpha(B_r^T(\overline{\xi}_i))}\left\{d((x_1,t_1),(x_2,t_2))^\alpha + |t_1 - t_2|^{\alpha/q_N}\right\}$$

for $(x_1,t_1),(x_2,t_2) \in K \times [\tau, T]$. However, since $\mathcal{L}_{\overline{x}}(u\psi) = \psi\mathcal{L}_{\overline{x}}u - \psi_t u$, we have

$$\begin{aligned}
\left|\mathcal{L}_{\overline{x}}(u\psi)\right|_{C_x^\alpha(B_r^T(\overline{\xi}_i))} &\le |\psi\mathcal{L}_{\overline{x}}u|_{C_x^\alpha(S_T)} + |\psi_t u|_{C_x^\alpha(S_T)} \\
&\le |\mathcal{L}_{\overline{x}}u|_{C_x^\alpha(S_T)} + c\,|u|_{C_x^\alpha(S_T)} \\
&\le |\mathcal{L}u|_{C_x^\alpha(S_T)} + c\,|u|_{C_x^\alpha(S_T)} + \sum_{i,j=1}^{q}\left|\left[a_{ij}(\overline{x},t) - a_{ij}(\cdot,t)\right]\partial^2_{x_i x_j} u\right|_{C_x^\alpha(S_T)}.
\end{aligned}$$

On the other hand, since

$$\begin{aligned}
\left|\left[a_{ij}(\overline{x},t) - a_{ij}(\cdot,t)\right]\partial^2_{x_i x_j} u\right|_{C_x^\alpha(S_T)} &\le 2\Lambda\left|\partial^2_{x_i x_j} u\right|_{C_x^\alpha(S_T)} + |a_{ij}(\cdot,t)|_{C_x^\alpha(S_T)}\left\|\partial^2_{x_i x_j} u\right\|_{L^\infty(S_T)} \\
&\le 2\Lambda\left\|\partial^2_{x_i x_j} u\right\|_{C_x^\alpha(S_T)},
\end{aligned}$$

by Theorem 4.7 we conclude

$$\left|\partial^2_{x_i x_j} u(x_1,t_1) - \partial^2_{x_i x_j} u(x_2,t_2)\right| \le c\left\{\|\mathcal{L}u\|_{C_x^\alpha(S_T)} + \|u\|_{C^0(S_T)}\right\}\left\{d((x_1,t_1),(x_2,t_2))^\alpha + |t_1 - t_2|^{\alpha/q_N}\right\}.$$

So we are done. □

References

[1] F. Anceschi, Y. Zhu, On a spatially inhomogeneous nonlinear Fokker-Planck equation: Cauchy problem and diffusion asymptotics, Preprint https://arxiv.org/abs/2102.12795, 2022.
[2] M. Bramanti, An Invitation to Hypoelliptic Operators and Hörmander's Vector Fields, SpringerBriefs Math., Springer, Cham, 2014, xii+150 pp.
[3] M. Bramanti, M. Brandolini, Schauder estimates for parabolic nondivergence operators of Hörmander type, J. Differ. Equ. 234 (2007) 177–245.
[4] M. Bramanti, S. Polidoro, Fundamental solutions for Kolmogorov-Fokker-Planck operators with time-depending measurable coefficients, Math. Eng. 2 (4) (2020) 734–771.
[5] A. Brandt, Interior Schauder estimates for parabolic differential- (or difference-) equations via the maximum principle, Isr. J. Math. 7 (1969) 254–262.

[6] P.-E. Chaudru de Raynal, I. Honoré, S. Menozzi, Sharp Schauder estimates for some degenerate Kolmogorov equations, Ann. Sc. Norm. Super. Pisa, Cl. Sci. (5) 22 (3) (2021) 989–1089.

[7] G.M. Constantine, T.H. Savits, A multivariate Faà di Bruno formula with applications, Trans. Am. Math. Soc. 348 (2) (1996) 503–520.

[8] M. Di Francesco, S. Polidoro, Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form, Adv. Differ. Equ. 11 (11) (2006) 1261–1320.

[9] H. Dong, S. Kim, Partial Schauder estimates for second-order elliptic and parabolic equations: a revisit, Int. Math. Res. Not. (7) (2019) 2085–2136.

[10] C. Henderson, S. Snelson, C^∞ smoothing for weak solutions of the inhomogeneous Landau equation, Arch. Ration. Mech. Anal. 236 (1) (2020) 113–143.

[11] C. Henderson, W. Wang, Kinetic Schauder estimates with time-irregular coefficients and uniqueness for the Landau equation, Preprint https://arxiv.org/abs/2205.12930, 2022.

[12] L. Hörmander, Pseudo-differential operators and hypoelliptic equations, in: Singular Integrals, in: Proc. Sympos. Pure Math., vol. X, Chicago, Ill., 1966, Amer. Math. Soc., Providence, R.I., 1967, pp. 138–183.

[13] L. Hörmander, Hypoelliptic second order differential equations, Acta Math. 119 (1967) 147–171.

[14] C. Imbert, C. Mouhot, The Schauder estimate in kinetic theory with application to a toy nonlinear model, Ann. Henri Lebesgue 4 (2021) 369–405.

[15] B.F. Knerr, Parabolic interior Schauder estimates by the maximum principle, Arch. Ration. Mech. Anal. 75 (1) (1980/1981) 51–58.

[16] A.N. Kolmogorov, Zufällige Bewegungen (Zur Theorie der Brownschen Bewegung), Ann. Math. 35 (2) (1934) 116–117.

[17] N.V. Krylov, E. Priola, Elliptic and parabolic second-order PDEs with growing coefficients, Commun. Partial Differ. Equ. 35 (1) (2010) 1–22.

[18] E. Lanconelli, S. Polidoro, On a class of hypoelliptic evolution operators, in: Partial Differential Equations, II, Turin, 1993, Rend. Semin. Mat. Univ. Politec. Torino 52 (1) (1994) 29–63.

[19] G.M. Lieberman, Intermediate Schauder theory for second order parabolic equations. IV. Time irregularity and regularity, Differ. Integral Equ. 5 (6) (1992) 1219–1236.

[20] L. Lorenzi, Optimal Hölder regularity for nonautonomous Kolmogorov equations, Discrete Contin. Dyn. Syst., Ser. (4) (1) (2011) 169–191.

[21] G. Lucertini, S. Pagliarani, A. Pascucci, Optimal regularity for degenerate Kolmogorov equations with rough coefficients, J. Evol. Equ. 23 (4) (2023) 69.

[22] A. Lunardi, Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in $\mathbb{R}^n$, Ann. Sc. Norm. Super. Pisa, Cl. Sci. (4) 24 (1) (1997) 133–164.

[23] M. Manfredini, The Dirichlet problem for a class of ultraparabolic equations, Adv. Differ. Equ. 2 (5) (1997) 831–866.

[24] S. Menozzi, Martingale problems for some degenerate Kolmogorov equations, Stoch. Process. Appl. 128 (3) (2018) 756–802.

[25] C. Mouhot, De Giorgi–Nash–Moser and Hörmander Theories: New Interplays, World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2018, pp. 2467–2493.

[26] A. Nagel, E.M. Stein, S. Wainger, Balls and metrics defined by vector fields. I. Basic properties, Acta Math. 155 (1985) 103–147.

[27] A. Pascucci, A. Pesce, On stochastic Langevin and Fokker-Planck equations: the two-dimensional case, J. Differ. Equ. 310 (2022) 443–483.

[28] E. Priola, Global Schauder estimates for a class of degenerate Kolmogorov equations, Stud. Math. 194 (2) (2009) 117–153.

[29] C. Villani, A review of mathematical topics in collisional kinetic theory, in: Handbook of Mathematical Fluid Dynamics, vol. I, North-Holland, Amsterdam, 2002, pp. 71–305.

[30] W.D. Wang, L.Q. Zhang, The C^α regularity of a class of non-homogeneous ultraparabolic equations, Sci. China Ser. A 52 (8) (2009) 1589–1606.