arXiv:2405.09358v3 [math.AP] 24 Nov 2025

GLOBAL SOBOLEV THEORY FOR KOLMOGOROV-FOKKER-PLANCK OPERATORS WITH COEFFICIENTS MEASURABLE IN TIME AND VMO IN SPACE

STEFANO BIAGI, MARCO BRAMANTI

Abstract. We consider Kolmogorov-Fokker-Planck operators of the form

$$\mathcal{L}u = \sum_{i,j=1}^{q} a_{ij}(x,t)u_{x_i x_j} + \sum_{k,j=1}^{N} b_{jk}x_k u_{x_j} - \partial_t u,$$

with $(x,t) \in \mathbb{R}^{N+1}, N \geq q \geq 1$. We assume that $a_{ij} \in L^\infty\left(\mathbb{R}^{N+1}\right)$, the matrix $\{a_{ij}\}$ is symmetric and uniformly positive on $\mathbb{R}^q$, and the *drift*

$$Y = \sum_{k,j=1}^{N} b_{jk}x_k \partial_{x_j} - \partial_t$$

has a structure which makes the model operator with constant a_{ij} hypoelliptic, translation invariant w.r.t. a suitable Lie group operation, and 2-homogeneus w.r.t. a suitable family of dilations. We also assume that the coefficients a_{ij} are VMO w.r.t. the space variable, and only bounded measurable in t. We prove, for every $p \in (1,\infty)$, global Sobolev estimates of the kind:

$$\|u\|_{W_X^{2,p}(S_T)} \equiv \sum_{i,j=1}^{q} \left\|u_{x_i x_j}\right\|_{L^p(S_T)} + \|Yu\|_{L^p(S_T)} + \sum_{i=1}^{q} \|u_{x_i}\|_{L^p(S_T)}$$
$$+ \|u\|_{L^p(S_T)} \leq c\left\{\|\mathcal{L}u\|_{L^p(S_T)} + \|u\|_{L^p(S_T)}\right\}$$

with $S_T = \mathbb{R}^N \times (-\infty, T)$ for any $T \in (-\infty, +\infty]$. Also, the well-posedness in $W_X^{2,p}(\Omega_T)$, with $\Omega_T = \mathbb{R}^N \times (0,T)$ and $T \in \mathbb{R}$, of the Cauchy problem

$$\begin{cases} \mathcal{L}u = f & \text{in } \Omega_T \\ u(\cdot,0) = g & \text{in } \mathbb{R}^N \end{cases}$$

is proved, for $f \in L^p(\Omega_T), g \in W_X^{2,p}\left(\mathbb{R}^N\right)$.

1. Introduction and main results

1.1. **The problem and its context.** In this paper we deal with Kolmogorov-Fokker-Planck (KFP, in short) operators of the form

$$\mathcal{L}u = \sum_{i,j=1}^{q} a_{ij}(x,t)\partial^2_{x_i x_j}u + \sum_{k,j=1}^{N} b_{jk}x_k\partial_{x_j}u - \partial_t u, \qquad (x,t) \in \mathbb{R}^{N+1} \tag{1.1}$$

where $N \geq q \geq 1$. The first-order part of the operator, also called *the drift term*, will be briefly denoted by

$$Yu = \sum_{k,j=1}^{N} b_{jk}x_k\partial_{x_j}u - \partial_t u. \tag{1.2}$$

2010 *Mathematics Subject Classification.* 35K65, 35K70, 35B45, 35A08, 35K15, 42B20, 42B25.
Key words and phrases. Kolmogorov-Fokker-Planck operators; Sobolev estimates; measurable coefficients; VMO coefficients; Cauchy problem.

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 coefficients defined in $\mathbb{R}^{N+1}$ and belonging to $L^\infty(\mathbb{R}^{N+1})$, so that

$$\nu|\xi|^2 \le \sum_{i,j=1}^{q} a_{ij}(x,t)\xi_i\xi_j \le \nu^{-1}|\xi|^2 \tag{1.3}$$

for some constant $\nu > 0$, every $\xi \in \mathbb{R}^q$, a.e. $(x,t) \in \mathbb{R}^{N+1}$.

The coefficients will be also assumed VMO w.r.t. x (and merely measurable w.r.t. t) in a sense that will be made precise later (see Assumption (H3) in Section 1.3).

(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

$$m_0 \ge m_1 \ge \ldots \ge m_k \ge 1 \quad \text{and} \quad m_0 + m_1 + \ldots + m_k = N, \tag{1.4}$$

we have

$$B = \begin{pmatrix} \mathbb{O} & \mathbb{O} & \ldots & \mathbb{O} & \mathbb{O} \\ B_1 & \mathbb{O} & \ldots & \ldots & \ldots \\ \mathbb{O} & B_2 & \ldots & \mathbb{O} & \mathbb{O} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \mathbb{O} & \mathbb{O} & \ldots & B_k & \mathbb{O} \end{pmatrix} \tag{1.5}$$

where $\mathbb{O}$ is the null matrix, and every block B_j is an $m_j \times m_{j-1}$ matrix of rank m_j (for $j = 1, 2, \ldots, k$).

We explicitly note that, when $q < N$, the operator $\mathcal{L}$ is *ultraparabolic*; nevertheless, under assumptions **(H1)-(H2)**, as proved by Lanconelli and Polidoro [24], the *model operator* $\mathcal{L}_0$ corresponding to the case of *constant* a_{ij} (as well as its formal transpose $\mathcal{L}_0^*$) is hypoelliptic and satisfies the following properties:

(a) $\mathcal{L}_0$ 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

$$(y,s) \circ (x,t) = (x + E(t)y, t+s) \tag{1.6}$$
$$(y,s)^{-1} = (-E(-s)y, -s),$$

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

$$(y,s)^{-1} \circ (x,t) = (x - E(t-s)y, t-s), \tag{1.7}$$

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

(b) $\mathcal{L}_0$ 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

$$D(\lambda)(x,t) \equiv (D_0(\lambda)(x), \lambda^2 t) = (\lambda^{q_1}x_1, \ldots, \lambda^{q_N}x_N, \lambda^2 t), \tag{1.8}$$

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

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

The integer

$$Q = \sum_{i=1}^{N} q_i > N \tag{1.9}$$

is called the *homogeneous dimension* of $\mathbb{R}^N$, while $Q+2$ is the homogeneous dimension of $\mathbb{R}^{N+1}$.

Actually, Lanconelli-Polidoro in [24] have studied 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. After the seminal paper [24], more general families of degenerate KFP operators of the kind (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 [25] and later, under more general assumptions on B, by Di Francesco-Polidoro in [18], 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 $\partial_{x_1}, ..., \partial_{x_q}, Y$. In the framework of Sobolev spaces, local L^p estimates for the derivatives $\partial^2_{x_i x_j} u$ $(i,j = 1,2,...,q)$ and Yu in terms of $\mathcal{L}u$ have been established, for operators (1.1) with VMO coefficients $a_{ij}(x,t)$, by Bramanti, Cerutti, Manfredini [11]. We recall that the VMO assumption allows for some kind of discontinuity.

In recent years, there has been growing interest, especially motivated from the research in the field of stochastic differential equations (see e.g. [28]), in the study of KFP operators with coefficients a_{ij} rough in t (say, L^∞), and with some mild regularity (for instance, Hölder continuity) 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. Such estimates are sometimes called "partial Schauder estimates". In the paper [1] we have established global partial Schauder estimates for degenerate KFP operators (1.1) satisfying assumptions **(H1)**-**(H2)**, with coefficients a_{ij} Hölder continuous in space, bounded measurable in time. We have also shown that the second order derivatives $\partial^2_{x_i x_j} u$ (for $i,j = 1,2,...,q$) are actually locally Hölder continuous also w.r.t. time. In [3] we have proved analogous results on bounded cylinders, while more general global results in the context of partially Dini continuous functions have been established by Biagi, Bramanti, Stroffolini in [4]. Partial Schauder estimates for degenerate KFP operators have been proved also in [14] by Chaudru de Raynal, Honoré, Menozzi, with different techniques and without getting the Hölder control in time of second order derivatives. We refer to the introduction of [1] for further references on both standard and partial Schauder estimates for these operators.

On the other hand, Sobolev estimates have been proved by Menozzi [26] for a family of KFP operators with very general drift (containing as a special case the class (1.1)), assuming the coefficients $a_{ij}(x,t)$ uniformly continuous in x and bounded measurable in t. Later, Dong and Yastrzhembskiy in [19], have proved Sobolev estimates of this kind for a class of operators (1.1) with the drift of special type ("kinetic KFP operator"), assuming the coefficients $a_{ij}(x,t)$ to be VMO w.r.t. x and bounded measurable w.r.t. time. We also quote the paper [20], by the same Authors, where similar results are proved in the case of KFP operators in divergence form.

In the present paper we prove global Sobolev estimates, on $\mathbb{R}^{N+1}$ and on infinite strips $S_T = \mathbb{R}^N \times (-\infty, T)$, for KFP operators (1.1) satisfying assumptions **(H1)-(H2)**, with coefficients $a_{ij}(x,t)$ VMO w.r.t. x and L^∞ w.r.t. time (see assumption (**H3**) in Section 1.3 and Theorems 1.4 and 4.1 for the precise results). Moreover, we prove the unique solvability in $W_X^{2,p}(S_T)$ of the equation

$$\mathcal{L}u - \lambda u = f \text{ in } S_T$$

for $\lambda > 0$ large enough and $f \in L^p(S_T), T \in (-\infty, +\infty]$ (see Theorem 4.7). From this fact we also deduce well-posedness results for the Cauchy problem for $\mathcal{L}$ on strips $\Omega_T = \mathbb{R}^N \times (0,T), T \in (0, +\infty)$,

$$\begin{cases} \mathcal{L}u = f \text{ in } \Omega_T \\ u(\cdot, 0) = g \end{cases}$$

with $f \in L^p(\Omega_T), g \in W_X^{2,p}(\mathbb{R}^N)$ (see Theorem 1.7).

Our class of operators contains that one studied in [19], but our technique is completely different. Actually, here we apply the technique that we have recently introduced in [2], which is based on a combination of different ingredients. To describe this technique and the strategy of this paper, let us briefly recall two different approaches which have been followed so far to handle uniformly elliptic or parabolic nonvariational operators with VMO coefficients. The first one dates back to the early 1990s with the works of Chiarenza-Frasca-Longo [15], [16] for uniformly elliptic operators. That technique heavily relies on representation formulas of $\partial^2_{x_i x_j} u$ in terms of $\mathcal{L}u$, via a singular kernel "with variable coefficients", obtained from the fundamental solution of the model operator with constant coefficients. This singular kernel needs to be expanded in series of spherical harmonics, getting singular kernels "with constant coefficients" (i.e., of convolution type). One then needs to apply results of L^p continuity of singular integral operators and of the *commutator* of a singular integral operator with a BMO function. This technique, and its furher extensions to parabolic operators (see [9]), KFP operators (see [11]), and nonvariational operators structured on Hörmander vector fields (see [6], [7]), exploits both translation invariance and homogeneity (in suitable senses) of the model operator with constant a_{ij} or, in some cases, the possibility of approximating locally the operator under study with another one possessing these properties. Moreover, it relies on the knowledge of fine properties of the fundamental solution of the constant coefficient operator, with bounds on the derivatives of every order of this fundamental solution, which have to be uniform as the constant matrix (a_{ij}) ranges in a fixed ellipticity class.

A different technique to prove $W^{2,p}$ a-priori estimates for nonvariational elliptic or parabolic operators with VMO coefficients has been devised in 2007 by Krylov [21], and exploited in a series of subsequent papers. A key step in Krylov' technique constists in establishing a pointwise estimate on the sharp maximal function of the second derivatives of u in terms of $\overline{\mathcal{L}}u$, where $\overline{\mathcal{L}}$ is the model operator with constant a_{ij} (if the final goal is to study elliptic operators with VMO coefficients) or coefficients $a_{ij}(t) \in L^\infty(\mathbb{R})$ (if the final goal is to study parabolic operators with coefficients $a_{ij}(x,t)$ VMO w.r.t. x and L^∞ w.r.t. t). Once this estimate is proved, a clever procedure allows to exploit the VMO assumption on a_{ij} for replacing the model operator $\overline{\mathcal{L}}$ with $\mathcal{L}$. In turn, in order to establish this estimate on the sharp maximal function of $\partial^2_{x_i x_j} u$, Krylov makes use of many results on elliptic or parabolic operators with constant coefficients (or, in the parabolic case, coefficients only depending on t), including several classical solvability results, together with pointwise estimates on the derivatives of the solutions, and exploits translations, dilations, and the Poincaré's inequality in the space variables. The extension of *this part* of Krylov' technique to degenerate operators of Hörmander type seems very difficult. So far, we can quote only the paper [13] by Bramanti-Toschi where this technique has been successfully implemented, for novariational operators modeled on Hörmander vector fields in Carnot groups.

In [2] we have introduced a new approach which combines some ideas of both the strategies described above (by Chiarenza-Frasca-Longo and by Krylov), which we are now going to describe. In the present situation, our *model operator* is the KFP operator with coefficients *only depending on time* (in a merely L^∞ way):

$$\overline{\mathcal{L}}u = \sum_{i,j=1}^{q} a_{ij}(t)\partial^2_{x_i x_j} u + \sum_{k,j=1}^{N} b_{jk} x_k \partial_{x_j} u - \partial_t u. \tag{1.10}$$

We follow Krylov in the idea of exploiting an estimate on the sharp maximal function of $\partial^2_{x_i x_j} u$ (in terms of $\overline{\mathcal{L}}u$) to prove L^p estimates for the operator $\mathcal{L}$ with coefficients VMO w.r.t. x and L^∞ w.r.t. t (see our Theorem 2.1). In order to prove this bound on the sharp maximal function, however, we do not follow Krylov' technique but use representation formulas for second order derivatives in terms of the model operator $\overline{\mathcal{L}}$. Since an explicit fundamental solution for $\overline{\mathcal{L}}$ has been built in [12], the singular kernel appearing in this representation formula, although not of convolution type, can be directly studied thanks to the theory of singular integrals in spaces of homogeneous type. The link between the two ingredients of our proof (singular integral operators *and* sharp maximal function) is contained in an abstract real analysis result in spaces of homogeneous type which has been proved in [2]. By the way, we note that the lack of homogeneity, and the lack of regularity w.r.t. the t variable, of the fundamental solution of $\overline{\mathcal{L}}$, prevents us from applying Chiarenza-Frasca-Longo's technique in this situation.

As usual for operators with variable coefficients, we first prove a-priori estimates for smooth functions with *small* support. Passing from this result to global a-priori estimates for any $W^{2,p}_X$-function requires the use of suitable cutoff functions, a covering theorem, and interpolation inequalities for first order derivatives. In the present setting, cutoff functions are easily built thanks to the presence of translations and dilations. A general covering theorem in spaces of homogeneous type, proved in [2], applies also to our situation. Finally, in the present situation the required interpolation inequality is just the Euclidean one on ∂_{x_i}.

1.2. **Structure of the paper.** In Section 1.3 we introduce the geometric structure related to our operator (translations, dilations, distance), and point out some known facts about these structures. We also introduce some real variable notions related to maximal functions, and recall the related known inequalities. We then introduce the function spaces that will be used throughout the paper, make our assumptions and state our main results. In Section 2 we study the model operator $\overline{\mathcal{L}}$ (1.10) with coefficients depending only on t. We recall the known facts about its fundamental solution Γ, prove suitable representation formulas, prove the L^p continuity of the singular integral operator corresponding to the singular kernel $\partial^2_{x_i x_j}\Gamma$, and with these tools we prove Theorem 2.1, which gives a bound on the oscillation of $\partial^2_{x_i x_j}u$ in terms of $\overline{\mathcal{L}}u$. In Section 3 we consider operators with coefficients $a_{ij}(x,t)$ satisfying our VMO assumption in the space variables. Applying Theorem 2.1, we prove an analogous bound on the oscillation of $\partial^2_{x_i x_j}u$ in terms of $\mathcal{L}u$ (Section 3.1). This bound implies, via the maximal inequalities and exploiting our VMO assumption, an L^p estimate on $\partial^2_{x_i x_j}u$ in terms of $\mathcal{L}u$, for smooth functions with compact support (Section 3.2). Finally, we extend the L^p bound to a global Sobolev estimate, with the techiques already sketched at the end of Section 1.1. In Section 4 we come to the existence results. First (Section 4.1) we show that the global estimates proved in Section 3.2 still hold on strips $S_T = \mathbb{R}^N \times (-\infty, T)$ for every $T \in \mathbb{R}$, and prove the well-posedness of the equation

$$\mathcal{L}u - \lambda u = f \text{ in } S_T$$

for $\lambda > 0$ large enough and $f \in L^p(S_T)$ with $T \in (-\infty, +\infty]$. Then, in Section 4.3, we deduce a well-posedness result for the Cauchy problem for $\mathcal{L}$.

1.3. **Assumptions and main results.** We can now start giving some precise definitions which will allow to state our main result.

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

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

Let us introduce the metric structure related to the operator $\mathcal{L}$ that will be used throughout the following. The vector fields $\partial_{x_1}, \ldots, \partial_{x_q}, Y$ form a system of Hörmander vector fields in $\mathbb{R}^{N+1}$, left-invariant w.r.t. the composition law $\circ$ defined in (1.6). The vector fields ∂_{x_i} $(i = 1, ..., q)$ are homogeneous of degree 1, while Y is homogeneous of degree 2 w.r.t. the dilations $D(\lambda)$ defined in (1.8). As every set of Hörmander vector fields with drift, this system

$$X = \{\partial_{x_1}, \ldots, \partial_{x_q}, Y\}$$

induces a (weighted) control distance d_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 $a_0, \ldots, a_q : [0,1] \to \mathbb{R}$ are measurable functions such that

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

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

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

Since $\partial_{x_1}, \ldots, \partial_{x_q}, Y$ satisfy Hörmander's rank condition, it is well-known that the function d_X is a distance in $\mathbb{R}^{N+1}$ (see, e.g., [27, Prop. 1.1] or [8, Chap.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 vector fields, we see that

(a) d_X is left-invariant with respect to $\circ$, that is,

$$d_X(\xi, \eta) = d_X(\eta^{-1} \circ \xi, 0) \tag{1.11}$$

(b) d_X is is jointly 1-homogeneous with respect to $D(\lambda)$, that is

$$d_X(D(\lambda)\xi, D(\lambda)\eta) = \lambda d_X(\xi, \eta) \qquad \text{for every } \lambda > 0. \tag{1.12}$$

As a consequence of (1.11), the function $\rho_X(\xi) := d_X(\xi, 0)$ satisfies

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

moreover, by (1.12) we also have

(1)' $\rho_X(\xi) \geq 0$ and $\rho_X(\xi) = 0 \Leftrightarrow \xi = 0$;
(2)' $\rho_X(D(\lambda)\xi) = \lambda \rho_X(\xi)$,

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

We now observe that also the function

$$\rho(\xi) = \rho(x, t) := \|x\| + \sqrt{|t|} = \sum_{i=1}^{N} |x_i|^{1/q_i} + \sqrt{|t|} \tag{1.13}$$

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 (see e.g. [8, Thm.3.12]): there exist $c_1, c_2 > 0$ such that

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

As a consequence of this fact, the map

$$d(\xi, \eta) := \rho(\eta^{-1} \circ \xi) \tag{1.14}$$

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

$$d(\xi, \eta) \leq \boldsymbol{\kappa} \{d(\xi, \zeta) + d(\eta, \zeta)\} \qquad \forall\, \xi, \eta, \zeta \in \mathbb{R}^{N+1}; \tag{1.15}$$

$$d(\xi, \eta) \leq \boldsymbol{\kappa}\, d(\eta, \xi) \qquad \forall\, \xi, \eta \in \mathbb{R}^{N+1}. \tag{1.16}$$

The quasidistance d is *globally equivalent* to the control distance d_X; hence, we will systematically use this quasidistance d and the associated balls

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

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

(1) By (1.7), d has the following explicit expression

$$d(\xi, \eta) = \|x - E(t-s)y\| + \sqrt{|t-s|}, \tag{1.17}$$

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

(2) Since $E(0) = \mathrm{Id}$, from (1.17) we get

$$d((x,t),(y,t)) = \|x-y\| \qquad \text{for every } x,y \in \mathbb{R}^N \text{ and } t \in \mathbb{R}, \tag{1.18}$$

from which we derive that the quasidistance 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 (1.15), 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

$$B_r(\xi) = \xi \circ B_r(0) = \xi \circ D(r)\,(B_1(0)) \quad \forall\, \xi \in \mathbb{R}^{N+1},\ r > 0. \tag{1.19}$$

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

$$|B_r(\xi)| = |B_r(0)| = \omega\, r^{Q+2} \tag{1.20}$$

where $\omega := |B_1(0)| > 0$. Identity (1.20) illustrates the role of $Q+2$ as the *homogeneous dimension* of $\mathbb{R}^{N+1}$ (w.r.t. the dilations $D(\lambda)$).

(5) Property (1.20) in particular implies that *a global doubling condition holds*:

$$|B_{2r}(\xi)| \le c\,|B_r(\xi)| \tag{1.21}$$

for some constant $c > 0$, every $\xi \in \mathbb{R}^{N+1}$ and $r > 0$. In turn, this fact implies that $\mathbb{R}^{N+1}$, endowed with the quasidistance d (or the equivalent control distance d_X) and the Lebesgue measure is a *space of homogeneous type*, in the sense of Coifman-Weiss [17]. This will be a key point, in order to apply several deep known results from real analysis.

(6) Another consequence of (1.20) that we will apply in the following is the inequality

$$|B_{kr}(\xi)| \le c k^{Q+2}\,|B_r(\xi)| \tag{1.22}$$

for every $\xi \in \mathbb{R}^{N+1}, r > 0, k \ge 1$.

The quasidistance d allows us to define the function spaces which will be used throughout the paper.

Definition 1.2. For any $f \in L^1_{loc}\left(\mathbb{R}^{N+1}\right)$ we define the (partial) VMO_x modulus of f as the function

$$\eta_f(r) = \sup_{\xi_0 \in \mathbb{R}^{N+1}, \rho \le r} \frac{1}{|B_\rho(\xi_0)|} \int_{B_\rho(\xi_0)} \left| f(x,t) - f(\cdot,t)_{B_\rho(x_0,t_0)} \right| dxdt,$$

for any $r > 0$, where, throughout the following, we let

$$f(\cdot,t)_B = \frac{1}{|B|} \int_B f(y,t)\,dyds. \tag{1.23}$$

We say that $f \in BMO_x\left(\mathbb{R}^{N+1}\right)$ if η_f is bounded; we say that $f \in VMO_x\left(\mathbb{R}^{N+1}\right)$ if, moreover, $\eta_f(r) \to 0$ as $r \to 0^+$.

Note that in (1.23) we compute the integral average, over a ball of $\mathbb{R}^{N+1}$, of the function $f(\cdot,t)$ which only depends on the space variable in $\mathbb{R}^N$ (since t is fixed). Nevertheless, due to the nontrivial structure of metric balls, this quantity cannot be rewritten as an integral average over some fixed subset of $\mathbb{R}^N$. This phenomenon is different from what happens in the parabolic case dealt in [21].

Note also that if $f \in L^\infty\left(\mathbb{R}^{N+1}\right)$ then obviously $f \in BMO_x\left(\mathbb{R}^{N+1}\right)$ with $\eta_f(r) \leq 2\|f\|_{L^\infty(\mathbb{R}^n)}$. Our last assumption on the variable coefficients a_{ij} will be the following:

(H3) We ask that the coefficients a_{ij} in (1.1) belong to $VMO_x\left(\mathbb{R}^{N+1}\right)$.

Letting, for any $R>0$,

$$a^\sharp(R) = \max_{i,j=1,\dots,q} \eta_{a_{ij}}(R), \tag{1.24}$$

our bounds will depend quantitatively on the coefficients through the function $a^\sharp$ and the number ν in (1.3).

Let us now introduce the Sobolev spaces related to our system of Hörmander vector fields, fixing the related notation (see [8, Chap.2] for details).

Definition 1.3 (Sobolev spaces)**.** Under the above assumptions, for any $p \in [1,\infty]$, $k=1,2$, and domain $\Omega \subseteq \mathbb{R}^{N+1}$, we define the Sobolev space

$$W_X^{k,p}(\Omega) = \left\{ f : \Omega \to \mathbb{R} : \|f\|_{W_X^{k,p}(\Omega)} < \infty \right\}$$

where

$$\|f\|_{W_X^{1,p}(\Omega)} = \|f\|_{L^p(\Omega)} + \sum_{i=1}^{q} \|\partial_{x_i} f\|_{L^p(\Omega)}$$

$$\|f\|_{W_X^{2,p}(\Omega)} = \|f\|_{W_X^{1,p}(\Omega)} + \sum_{i,j=1}^{q} \left\|\partial^2_{x_i x_j} f\right\|_{L^p(\Omega)} + \|Yf\|_{L^p(\Omega)}$$

and all the derivatives are meant in weak sense.

We are finally in position to state our first main result.

Theorem 1.4 (Global Sobolev estimates)**.** *Let $\mathcal{L}$ be an operator as in* (1.1)*, and assume that* **(H1)**, **(H2)**, **(H3)** *hold. Then, for every $p \in (1,\infty)$ there exists a constant $c>0$, depending on p, the matrix B in (1.5), the number ν in (1.3), and the function $a^\#$ in (1.24), such that*

$$\|u\|_{W_X^{2,p}(\mathbb{R}^{N+1})} \leq c\left\{ \|\mathcal{L}u\|_{L^p(\mathbb{R}^{N+1})} + \|u\|_{L^p(\mathbb{R}^{N+1})} \right\} \tag{1.25}$$

for every function $u \in W_X^{2,p}\left(\mathbb{R}^{N+1}\right)$.

Our second main result is the well-posedness of the Cauchy problem for $\mathcal{L}$, that is,

$$\begin{cases} \mathcal{L}u = f & \text{in } \Omega_T \equiv \mathbb{R}^N \times (0,T) \\ u(\cdot,0) = g & \text{in } \mathbb{R}^N, \end{cases} \tag{1.26}$$

where $T \in (0,+\infty)$ is fixed, $f \in L^p(\Omega_T)$ and $g \in W_X^{2,p}\left(\mathbb{R}^N\right)$. By the way, saying that a function g depending on x alone belongs to $W_X^{2,p}\left(\mathbb{R}^N\right)$, obviously means that it belongs to $W_X^{2,p}(\Omega_T)$ if we regard g as a function of (x,t). We the obvious meaning of norms, we have

$$\|g\|_{W_X^{2,p}(\Omega_T)} = T^{\frac{1}{p}} \|g\|_{W_X^{2,p}(\mathbb{R}^N)}. \tag{1.27}$$

In order to give sense to the initial condition in (1.26) avoiding a delicate notion of trace, following [22], let us introduce in an indirect way the Sobolev spaces of $W_X^{2,p}(\Omega_T)$ functions "vanishing for $t = 0$":

Definition 1.5. We say that $u \in \mathring{W}_X^{2,p}(\Omega_T)$ if $u \in W_X^{2,p}(\Omega_T)$ and the function $\widetilde{u}$, obtained from u extending it as zero for $t < 0$, belongs to $W_X^{2,p}(S_T)$. Functions in $\mathring{W}_X^{2,p}(\Omega_T)$ will be thought as defined either on Ω_T or on S_T (vanishing for $t < 0$).

We can now give the precise definition of *solution* to the Cauchy problem.

Definition 1.6. Let $f \in L^p(\Omega_T)$ and $g \in W_X^{2,p}(\mathbb{R}^N)$ for some $p \in (1,\infty)$ and $T \in (0,+\infty)$. We say that a function $u \in W_X^{2,p}(\Omega_T)$ is a solution to problem (1.26) if

(1) $\mathcal{L}u = f$ a.e. in Ω_T;
(2) $u - g \in \mathring{W}_X^{2,p}(\Omega_T)$.

Then, our second main result is the following:

Theorem 1.7 (Well-posedness of the Cauchy problem)**.** *Under the same assumptions on $\mathcal{L}$ of Theorem 1.4, let $f \in L^p(\Omega_T)$ and $g \in W_X^{2,p}(\mathbb{R}^N)$ for some $p \in (1,\infty)$ and $T \in (0,+\infty)$. Then, there exists a unique solution $u \in W_X^{2,p}(\Omega_T)$ of the Cauchy problem (1.26). Moreover, the following estimate holds*

$$\|u\|_{W_X^{2,p}(\Omega_T)} \le c \left\{ \|f\|_{L^p(\Omega_T)} + \|g\|_{W_X^{2,p}(\mathbb{R}^N)} \right\}, \tag{1.28}$$

where c depends on p, T, the matrix B in (1.5)*, the number ν in* (1.3)*, and the function $a^{\#}$ in* (1.24)*.*

We end this Section introducing some more known facts from real analysis in spaces of homogeneous type. We will use in the following two different kinds of *maximal functions*.

Definition 1.8. For $f \in L^1_{loc}(\mathbb{R}^{N+1})$, $\xi \in \mathbb{R}^{N+1}$, we define the *Hardy-Littlewood* (uncentered) *maximal function* as:

$$\mathcal{M}f(\xi) = \sup_{\substack{B_r(\overline{\xi}) \ni \xi \\ \overline{\xi} \in \mathbb{R}^{N+1}, r>0}} \frac{1}{|B_r(\overline{\xi})|} \int_{B_r(\overline{\xi})} |f(\eta)|\, d\eta. \tag{1.29}$$

Since $(\mathbb{R}^{N+1}, d, |\cdot|)$ is a space of homogeneous type, by [17, Thm.2.1], we have:

Theorem 1.9. *For every $p \in (1,\infty]$ there exists $c > 0$ such that, for every $f \in L^p(\mathbb{R}^{N+1})$,*

$$\|\mathcal{M}f\|_{L^p(\mathbb{R}^{N+1})} \le c_p \|f\|_{L^p(\mathbb{R}^{N+1})}. \tag{1.30}$$

Another kind of maximal function, which can also be introduced in any space of homogeneous type, is the following:

Definition 1.10. For $f \in L^1_{loc}\left(\mathbb{R}^{N+1}\right)$, $\xi \in \mathbb{R}^{N+1}$, we define the *sharp maximal function* of f as:

$$f^{\#}(\xi) = \sup_{\substack{B_r(\overline{\xi}) \ni \xi \\ \overline{\xi} \in \mathbb{R}^{N+1}, r>0}} \frac{1}{\left|B_r\left(\overline{\xi}\right)\right|} \int_{B_r(\overline{\xi})} \left| f(\eta) - f_{B_r(\overline{\xi})} \right| d\eta, \tag{1.31}$$

where

$$f_{B_r(\overline{\xi})} = \frac{1}{\left|B_r\left(\overline{\xi}\right)\right|} \int_{B_r(\overline{\xi})} f(\eta)\, d\eta.$$

Applying to our context the result proved in [29, Prop.3.4] in the general setting of spaces of homogeneous type of infinite measure, we have the following result, generalizing the well-known Fefferman-Stein inequality which holds in Euclidean spaces:

Theorem 1.11. *For every $p \in (1, \infty)$ there exists C_p (depending on p and the doubling constant in (1.21)) such that for every $f \in L^\infty\left(\mathbb{R}^{N+1}\right)$, f with bounded support, we have*

$$\|f\|_{L^p(\mathbb{R}^{N+1})} \leq C_p \left\| f^{\#} \right\|_{L^p(\mathbb{R}^{N+1})}. \tag{1.32}$$

Let us recall also the following abstract result proved in [2, Thm.3.10], which will be crucial for us:

Theorem 1.12. *Let (X, d, μ) be a space of homogeneous type and let T be a singular integral operator which we already know to be bounded on $L^p(X)$ for some $p \in (1, \infty)$. The operator T has kernel $K(x, y)$, which means that*

$$Tf(x) = \int_X K(x, y) f(y)\, d\mu(y)$$

when f is compactly supported and x does not belong to sprt f. *We assume that the kernel K satisfies the mean value inequality*

$$|K(x_0, y) - K(x, y)| \leq C \frac{d(x_0, x)}{d(x_0, y)\, B(x; y)} \text{ if } d(x_0, y) \geq M d(x_0, x) \tag{1.33}$$

where

$$B(x; y) = \mu(B(x, d(x, y)))$$

and $M > 1$ is such that the condition $d(x_0, y) \geq M d(x_0, x)$ implies the equivalence of $d(x_0, y)$ and $d(x, y)$ (if d is a distance, like in our case, $M = 2$ is a good choice). Let $\beta > 1$ be an exponent such that

$$|B(x, kr)| \leq c k^\beta |B(x, r)| \tag{1.34}$$

for every $k \geq 1, r > 0, x \in X$.

Then, there exists $c > 0$ such that for every $f \in L^p(X)$, $x_0 \in X$, ball $B_r = B(\overline{x}, r) \ni x_0$ (for some $r > 0$, $\overline{x} \in X$), $k \geq M$, we have:

$$\begin{aligned} &\frac{1}{|B_r(\overline{x})|} \int_{B_r} \left| Tf(x) - (Tf)_{B_r(\overline{x})} \right| d\mu(x) \\ &\leq c \left\{ \frac{1}{k} \mathcal{M} f(x_0) + k^{\frac{\beta}{p}} \left(\frac{1}{|B_{kr}(\overline{x})|} \int_{B_{kr}(\overline{x})} |f(x)|^p\, d\mu(x) \right)^{1/p} \right\}. \end{aligned}$$

Finally, we have to define some function spaces of Hölder or partially Hölder continuous functions which will be useful in the following.

Definition 1.13. Let $f:\mathbb{R}^{N+1}\to\mathbb{R}$. Given any number $\alpha\in(0,1)$, we introduce the notation:

$$|f|_{C^\alpha(\mathbb{R}^{N+1})}=\sup\left\{\frac{|f(\xi)-f(\eta)|}{d(\xi,\eta)^\alpha}:\xi,\eta\in\mathbb{R}^{N+1}\text{ and }\xi\neq\eta\right\}$$

$$\begin{aligned}|f|_{C^\alpha_x(\mathbb{R}^{N+1})}&=\operatorname*{ess\,sup}_{t\in\mathbb{R}}\left\{\frac{|f(x,t)-f(y,t)|}{d((x,t),(y,t))^\alpha}:x,y\in\mathbb{R}^N,x\neq y\right\}\\&=\operatorname*{ess\,sup}_{t\in\mathbb{R}}\left\{\frac{|f(x,t)-f(y,t)|}{\|x-y\|^\alpha}:x,y\in\mathbb{R}^N,x\neq y\right\}\end{aligned}$$

(where the last equality holds by (1.18)). Accordingly, we define the spaces $C^\alpha(\mathbb{R}^{N+1})$ and $C^\alpha_x(\mathbb{R}^{N+1})$ as follows:

$$C^\alpha(\mathbb{R}^{N+1}):=\left\{f\in C(\mathbb{R}^{N+1})\cap L^\infty(\mathbb{R}^{N+1}):|f|_{C^\alpha(\mathbb{R}^{N+1})}<\infty\right\}\tag{1.35}$$

$$C^\alpha_x(\mathbb{R}^{N+1}):=\left\{f\in L^\infty(\mathbb{R}^{N+1}):|f|_{C^\alpha_x(\mathbb{R}^{N+1})}<\infty\right\}.\tag{1.36}$$

Finally, we define the spaces

$$C^\alpha_0(\mathbb{R}^{N+1}):=\left\{f\in C^\alpha(\mathbb{R}^{N+1}):f\text{ is compactly supported in }\mathbb{R}^{N+1}\right\}\tag{1.37}$$

$$\mathcal{D}^\alpha_x(\mathbb{R}^{N+1}):=\left\{f\in C^\alpha_x(\mathbb{R}^{N+1}):f\text{ is compactly supported in }\mathbb{R}^{N+1}\right\}.\tag{1.38}$$

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

In this Section we consider a KFP operator $\overline{\mathcal{L}}$ of the form (1.1) but with coefficients a_{ij} *only depending on* t, that is,

$$\overline{\mathcal{L}}u=\sum_{i,j=1}^q a_{ij}(t)\partial^2_{x_ix_j}u+Yu,\qquad(x,t)\in\mathbb{R}^{N+1},\tag{2.1}$$

satisfying assumptions **(H1) - (H2)** in Section 1.1. For this operator, the goal of this Section is to prove an estimate on the mean oscillation of $\partial^2_{x_ix_j}u$ over balls, in terms of $\overline{\mathcal{L}}u$. This will be the key result that we will exploit, in Section 3, to get the desired Sobolev estimates for operators with partially VMO coefficients $a_{ij}(x,t)$. The result is the following, and will be proved at the end of Section 2.2:

Theorem 2.1. *Let assumptions* **(H1)-(H2)** *be in force. Then, for every $p\in(1,\infty)$ there exists $c>0$ such that, for every $u\in C^\infty_0(\mathbb{R}^{N+1})$, $r>0$, $\xi_0=(x_0,t_0)$ and $\bar\xi=(\overline{x},\overline{t})\in\mathbb{R}^{N+1}$ such that $\xi_0\in B_r(\bar\xi)$, every $1\le i,j\le q$ and $k\ge4\boldsymbol{\kappa}$ (where $\boldsymbol{\kappa}$ is the constant in (1.15)), we have the following estimate:*

$$\begin{aligned}&\frac{1}{|B_r(\bar\xi)|}\int_{B_r(\bar\xi)}|\partial^2_{x_ix_j}u(x,t)-(\partial^2_{x_ix_j}u)_{B_r(\bar\xi)}|\,dxdt\\&\quad\le c\Big\{\frac{1}{k}\mathcal{M}(\overline{\mathcal{L}}u)(\xi_0)\\&\qquad+k^{\frac{Q+2}{p}}\left(\frac{1}{\left|B_{kr}\left(\bar\xi\right)\right|}\int_{B_{kr}(\bar\xi)}\left|\overline{\mathcal{L}}u(x,t)\right|^p\,dx\,dt\right)^{1/p}\Big\}.\end{aligned}\tag{2.2}$$

The constant c depends on the matrix A only through the number ν in (1.3)*; moreover, $\mathcal{M}$ is the Hardy-Littlewood maximal operator defined in* (1.29)*.*

2.1. **Fundamental solution and representation formulas for $\overline{\mathcal{L}}$.** We begin by reviewing some known results concerning the fundamental solution of $\overline{\mathcal{L}}$ which we will need in the sequel; such results are proved in [1], [12], to which we refer for the proofs and for further details.

To begin with, we state the following *existence* result.

Theorem 2.2 (See [1, Thm. 3.11] and [12, Thm. 1.4])**.** *Under the above assumptions* **(H1)**-**(H2)***, let $C(t,s)$ be the $N\times N$ matrix defined as*

$$C(t,s)=\int_s^t E(t-\sigma)\cdot\begin{pmatrix}A_0(\sigma) & 0\\ 0 & 0\end{pmatrix}\cdot E(t-\sigma)^T\,d\sigma\quad (\textit{with } t>s) \tag{2.3}$$

(*we recall that $E(\sigma)=\exp(-\sigma B)$, see* (1.5)). *Then, the matrix $C(t,s)$ is* symmetric and positive definite *for every $t>s$. Moreover, if we define*

$$\begin{aligned}&\Gamma(x,t;y,s)\\ &=\frac{1}{(4\pi)^{N/2}\sqrt{\det C(t,s)}}e^{-\frac14\langle C(t,s)^{-1}(x-E(t-s)y),\,x-E(t-s)y\rangle}\cdot\mathbf{1}_{\{t>s\}}\end{aligned} \tag{2.4}$$

(*where $\mathbf{1}_A$ denotes the indicator function of a set A*)*, then for every $u\in C_0^\infty(\mathbb{R}^{N+1})$ we have the following representation formula*

$$u(x,t)=-\int_{\mathbb{R}^{N+1}}\Gamma(x,t;y,s)\overline{\mathcal{L}}u(y,s)\,dy\,ds\quad \textit{for every } (x,t)\in\mathbb{R}^{N+1}, \tag{2.5}$$

so that Γ is the fundamental solution *for $\overline{\mathcal{L}}$ with pole at (y,s).*

Moreover, Γ satisfies the following properties:

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

$$(\overline{\mathcal{L}}\Gamma(\cdot;\eta))(x,t)=0\quad \textit{for every } x\in\mathbb{R}^N \textit{ and a.e. } t\in\mathbb{R}. \tag{2.6}$$

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

$$\int_{\mathbb{R}^N}\Gamma(x,t;y,s)\,dy=1. \tag{2.7}$$

Remark 2.3. It is worth mentioning that the existence and the explicit expression of a fundamental solution Γ for the operator $\overline{\mathcal{L}}$ is proved in [12, Thm. 1.4] under a *weaker version* of assumption **(H2)**; moreover, several other properties of Γ (which we will not exploit in this paper) are established.

In the particular case when the coefficients a_{ij} of $\overline{\mathcal{L}}$ are *constant*, the associated fundamental solution Γ constructed in Theorem 2.2 takes a simpler form; due to its relevance in our arguments (see Theorem 2.5 and the proof of Lemma 2.14), we explicitly state this expression in the next theorem.

Theorem 2.4 (See [12])**.** *Let $\alpha>0$ be fixed, and let*

$$\mathcal{L}_\alpha u=\alpha\sum_{i=1}^q\partial^2_{x_ix_i}u+Yu. \tag{2.8}$$

Moreover, let Γ_α be the fundamental solution of $\mathcal{L}_\alpha$, whose existence is guaranteed by Theorem 2.2. Then, the following facts hold true:

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

$$\begin{aligned}\Gamma_\alpha(x,t;y,s) &= \Gamma_\alpha\big(x - E(t-s)y, t-s; 0,0\big) \\ &= \Gamma_\alpha\big((y,s)^{-1}\circ(x,t);0,0\big);\end{aligned} \tag{2.9}$$

(2) *the matrix* $C(t,s)$ *in* (2.3) *takes the simpler form*

$$C(t,s) = C_0(t-s), \tag{2.10}$$

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

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

Furthermore, one has the 'homogeneity property'

$$C_0(\tau) = D_0(\sqrt{\tau})C_0(1)D_0(\sqrt{\tau}) \qquad \forall\,\tau>0. \tag{2.11}$$

In particular, by combining (2.4) *with* (2.10)*-*(2.11)*, we can write*

$$\begin{aligned}\Gamma_\alpha(x,t;0,0) &= \frac{1}{(4\pi\alpha)^{N/2}\sqrt{\det C_0(t)}}e^{-\frac{1}{4\alpha}\langle C_0(t)^{-1}x,x\rangle} \\ &= \frac{1}{(4\pi\alpha)^{N/2}t^{Q/2}\sqrt{\det C_0(1)}}e^{-\frac{1}{4\alpha}\langle C_0(1)^{-1}\left(D_0\left(\frac{1}{\sqrt{t}}\right)x\right), D_0\left(\frac{1}{\sqrt{t}}\right)x\rangle}.\end{aligned} \tag{2.12}$$

With Theorems 2.2-2.4 at hand, we then proceed by recalling some *fine properties* of Γ and of its (spatial) derivatives, which will play a key rôle in our argument. To clearly state such properties (see Theorem 2.5 below), we first introduce a notation: if $\boldsymbol{\ell} = (\ell_1,\ldots,\ell_N)\in(\mathbb{N}\cup\{0\})^N$ is a given multi-index, we set

$$D_x^{\boldsymbol{\ell}} = \partial_{x_1}^{\ell_1}\cdots\partial_{x_N}^{\ell_N}, \qquad \omega(\boldsymbol{\ell}) = \textstyle\sum_{i=1}^N \ell_i q_i,$$

where $q_1,\ldots,q_N$ are the exponents appearing in the dilation $D_0(\lambda)$, see (1.8).

Theorem 2.5 (Fine properties of Γ)**.** *(See* [1, Thm. 3.5, Thm. 3.9, Prop. 3.13 and Thm. 3.16]*) Let* Γ *be as in Theorem 2.2, and let* $\nu>0$ *be as in* (1.3)*. Then, the following assertions hold.*

(1) *There exists* $c_1>0$ *and, for every pair of multi-indices* $\boldsymbol{\ell}_1,\boldsymbol{\ell}_2\in(\mathbb{N}\cup\{0\})^N$*, there exists* $c = c(\nu,\boldsymbol{\ell}_1,\boldsymbol{\ell}_2)>0$*, such that, for every* $(x,t),(y,s)\in\mathbb{R}^{N+1}$ *with* $t\neq s$*, we have*

$$\begin{aligned}|D_x^{\boldsymbol{\ell}_1}D_y^{\boldsymbol{\ell}_2}\Gamma(x,t;y,s)| &\le \frac{c}{(t-s)^{\omega(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2)/2}}\,\Gamma_{c_1\nu^{-1}}(x,t;y,s) \\ &\le \frac{c}{d((x,t),(y,s))^{Q+\omega(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2)}},\end{aligned} \tag{2.13}$$

In particular, we have

$$|D_x^{\boldsymbol{\ell}_1}D_y^{\boldsymbol{\ell}_2}\Gamma(\xi,\eta)| \le \frac{c}{d(\xi,\eta)^{Q+\omega(\boldsymbol{\ell}_1+\boldsymbol{\ell}_2)}} \quad \forall\,\xi\neq\eta. \tag{2.14}$$

(2) *For every multi-index* $\boldsymbol{\ell}\in(\mathbb{N}\cup\{0\})^N$ *there exists* $c = c(\boldsymbol{\ell},\nu)>0$ *such that*

$$|D_x^{\boldsymbol{\ell}}\Gamma(\xi_1,\eta) - D_x^{\boldsymbol{\ell}}\Gamma(\xi_2,\eta)| \le c\frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+\omega(\boldsymbol{\alpha})+1}} \tag{2.15}$$

$$|D_x^{\boldsymbol{\ell}}\Gamma(\eta,\xi_1) - D_x^{\boldsymbol{\ell}}\Gamma(\eta,\xi_2)| \le c\frac{d(\xi_1,\xi_2)}{d(\xi_1,\eta)^{Q+\omega(\boldsymbol{\alpha})+1}} \tag{2.16}$$

for every $\xi_1,\xi_2,\eta\in\mathbb{R}^{N+1}$ *such that*

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

(3) *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*

$$\int_{\mathbb{R}^N\times(\tau,t)}|\partial^2_{x_ix_j}\Gamma(x,t;y,s)|\cdot\|E(s-t)x-y\|^\alpha\,dy\,ds\leq c(t-\tau)^{\alpha/2}. \tag{2.17}$$

(4) *There exists a constant* $c>0$ *such that, for every fixed* $1\leq i,j\leq q$, *one has the estimate*

$$I_{r,\tau}(x,t):=\int_\tau^t\left|\int_{\{y\in\mathbb{R}^N:\,d((x,t),(y,s))\geq r\}}\partial^2_{x_ix_j}\Gamma(x,t;y,s)\,dy\right|ds\leq c, \tag{2.18}$$

$$J_{r,\tau}(y,s):=\int_s^\tau\left|\int_{\{x\in\mathbb{R}^N:\,d((y,s),(x,t))\geq r\}}\partial^2_{x_ix_j}\Gamma(x,t;y,s)\,dx\right|dt\leq c, \tag{2.19}$$

for every $(x,t),(y,s)\in\mathbb{R}^{N+1}$, $s<\tau<t,r>0$.

Remark 2.6. As a matter of fact, the mean-value inequality (2.16) is not explicitly proved in [1]; however, the proof of this inequality is *totally analogous* to that of (2.15) (see, precisely, [1, Thm. 3.9]), and it relies on the 'subelliptic' mean value theorem for the system of Hörmander vector fields

$$\{\partial_{x_1},\ldots,\partial_{x_q},Y\}$$

(see [1, Thm. 2.1]), together with the estimates (2.14) (which apply to *every spatial derivative* of Γ, both with respect to x and y). The same comment applies to the cancellation property (2.19) (see the proof of [1, Thm. 3.16]).

Before proceeding we highlight, for a future reference, an easy consequence of Theorem 2.5 and of identity (2.7), which will be repeatedly used in the sequel.

Lemma 2.7. *Let* Γ *be as in Theorem 2.2, and let* $\boldsymbol{\ell}$ *be a fixed non-zero multi-index. Then, for every* $x\in\mathbb{R}^N$ *and every* $s<t$ *we have*

$$\int_{\mathbb{R}^N}D^{\boldsymbol{\ell}}_x\Gamma(x,t;y,s)\,dy=0. \tag{2.20}$$

Starting from the representation formula (2.5), and using the fine properties of Γ collected in Theorem 2.5, one can prove the following representation formula for the second-order derivatives of a function $u\in C^\infty_0(\mathbb{R}^{N+1})$. Throughout what follows, we tacitly understand that α is a fixed number in $(0,1)$.

Theorem 2.8 (See [1, Cor. 3.12 and Thm. 3.14])**.** *For every* $u\in C^\infty_0(\mathbb{R}^{N+1})$ *and for every* $1\leq i,j\leq q$, *we have the following representation formulas*

$$\partial_{x_i}u(x,t)=-\int_{\mathbb{R}^{N+1}}\partial_{x_i}\Gamma(x,t;\cdot)\,\overline{\mathcal{L}}u\,dy\,ds, \tag{2.21}$$

$$\partial^2_{x_ix_j}u(x,t)=\int_{\mathbb{R}^{N+1}}\partial^2_{x_ix_j}\Gamma(x,t;y,s)\left[\overline{\mathcal{L}}u(E(s-t)x,s)-\overline{\mathcal{L}}u(y,s)\right]dy\,ds, \tag{2.22}$$

holding true for every $(x,t)\in\mathbb{R}^{N+1}$. *In particular, the two integrals appearing in the above* (2.21) - (2.22) *are absolutely convergent, and the operator*

$$T_{ij}f(x,t)=\int_{\mathbb{R}^{N+1}}\partial^2_{x_ix_j}\Gamma(x,t;y,s)\big[f(E(s-t)x,s)-f(y,s)\big]\,dy\,ds \tag{2.23}$$

is well defined for every $f \in \mathcal{D}_x^\alpha(\mathbb{R}^{N+1})$ *(this space has been defined in (1.38)).*

2.2. **Global $W^{2,p}$ and mean-oscillation estimates for $\overline{\mathcal{L}}$.** Taking into account the results recalled so far and following the strategy described at the beginning of the Section, we now turn to establish *global $W^{2,p}$ and mean-oscillation estimates for the solutions of $\overline{\mathcal{L}}u = f$.* Our first result in this direction is the following.

Theorem 2.9. *Let assumptions* **(H1)** - **(H2)** *be in force, and let $p \in (1, \infty)$ be fixed. Moreover, let T_{ij} be the operator defined in* (2.23)*, with $1 \le i, j \le q$.*

Then, T_{ij} can be extended to a linear and continuous operator from $L^p(\mathbb{R}^{N+1})$ into itself. In particular, there exists $c = c(p) > 0$ such that

$$\sum_{i,j=1}^{q} \|\partial^2_{x_i x_j} u\|_{L^p(\mathbb{R}^{N+1})} \le c\|\overline{\mathcal{L}}u\|_{L^p(\mathbb{R}^{N+1})}, \tag{2.24}$$

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

Remark 2.10. We wish to stress that, in view of our general strategy, the relevant point of the previous theorem is the L^p continutiy of the operator T_{ij} and not the estimate (2.24) *per se.* Also, we recall that, in the paper [26], L^p estimates for $\partial^2_{x_i x_j} u$ have been proved for KFP operators with coefficients $a_{ij}(x, t)$ uniformly continuous in x and L^∞ in t. Since our operator $\overline{\mathcal{L}}$ with coefficients only depending on t can be seen as a special case of the operators studied in [26], the estimate (2.24) should be contained in the results proved in [26]. However, we could not find in that paper neither a statement nor a proof of an explicit representation formula of the kind $\partial^2_{x_i x_j} u = T_{ij}\left(\overline{\mathcal{L}}u\right)$, linking the derivatives $\partial^2_{x_i x_j} u$ to a specific singular integral operator. The arguments in this Section make our proof logically independent from the results in [26], which we cannot apply directly to our context.

In order to prove Theorem 2.9, we first observe that, if $u \in C_0^\infty(\mathbb{R}^{N+1})$, then $\overline{\mathcal{L}}u \in \mathcal{D}_x^\alpha(\mathbb{R}^{N+1})$. Hence, if $1 \le i, j \le q$ are fixed, by the representation formula (2.22) we can write

$$\begin{aligned}\partial^2_{x_i x_j} u(x,t) &= \int_{\mathbb{R}^{N+1}} \partial^2_{x_i x_j}\Gamma(x,t;y,s)\left[\overline{\mathcal{L}}u(E(s-t)x,s) - \overline{\mathcal{L}}u(y,s)\right] dyds \\ &= T_{ij}(\overline{\mathcal{L}}u),\end{aligned}$$

where T_{ij} is the operator defined in (2.23). On the other hand, since by Theorem 2.8 the above integral *converges absolutely*, we have

$$\begin{aligned}T_{ij}(\overline{\mathcal{L}}u) &= \lim_{\varepsilon\to 0^+}\int_{-\infty}^{t-\varepsilon}\int_{\mathbb{R}^N} \partial^2_{x_i x_j}\Gamma(x,t;y,s)\left[\overline{\mathcal{L}}u(E(s-t)x,s) - \overline{\mathcal{L}}u(y,s)\right] dyds \\ &= \lim_{\varepsilon\to 0^+}\Big(-\int_{-\infty}^{t-\varepsilon}\int_{\mathbb{R}^N} \partial^2_{x_i x_j}\Gamma(\xi;\eta)\overline{\mathcal{L}}u(\eta)\, d\eta\Big),\end{aligned} \tag{2.25}$$

where we have used the fact that

$$\begin{aligned}&\int_{-\infty}^{t-\varepsilon}\int_{\mathbb{R}^N} \Gamma(x,t;y,s)\overline{\mathcal{L}}u(E(s-t)x,s)\, dyds \\ &\qquad = \int_{-\infty}^{t-\varepsilon} \overline{\mathcal{L}}u(E(s-t)x,s)\Big(\int_{\mathbb{R}^N} \partial^2_{x_i x_j}\Gamma(x,t;y,s)\, dy\Big) ds = 0,\end{aligned}$$

see Lemma 2.7 (and take into account the regularity of Γ out of the pole).

In order to apply L^p continuity results of singular integrals, we need to rewrite the limit (2.25) in a different way. To this aim, let us introduce a cut-off function $\phi_\varepsilon \in C_0^\infty(\mathbb{R})$ such that

i) $0 \le \phi_\varepsilon \le 1$ on $\mathbb{R}$;
ii) $\phi_\varepsilon = 0$ on $(-\infty, \varepsilon]$ and $\phi_\varepsilon = 1$ on $[2\varepsilon, +\infty)$;
iii) $|\phi'_\varepsilon| \le c/\varepsilon$ (for some $c > 0$),

and define the *truncated operator*

$$T_{ij}^\varepsilon(f)(x,t) = -\int_{\mathbb{R}^{N+1}} \phi_\varepsilon(t-s)\partial^2_{x_i x_j}\Gamma(x,t;y,s) f(y,s)\, dy\, ds, \tag{2.26}$$

which is well defined for every compactly supported $f \in L^\infty(\mathbb{R}^{N+1})$ since for every (x,t), the function

$$(y,s) \mapsto \phi_\varepsilon(t-s)\partial^2_{x_i x_j}\Gamma(x,t;y,s)$$

is locally integrable (see Lemma 2.14 below). We then claim that, for every fixed $f \in \mathcal{D}^\alpha_x(\mathbb{R}^{N+1})$ and for every $\xi \in \mathbb{R}^{N+1}$, we have

$$T_{ij}(f)(\xi) = \lim_{\varepsilon \to 0^+} T^\varepsilon_{ij}(f)(\xi). \tag{2.27}$$

Indeed, by the above computation we can write, for $f \in \mathcal{D}^\alpha_x(\mathbb{R}^{N+1})$,

a) $T_{ij}\,(f)\,(x,t) = \int_{\mathbb{R}^{N+1}} \partial^2_{x_i x_j}\Gamma(x,t;y,s)\,[f(E(s-t)x,s) - f(y,s)]\, dyds;$

b) $$\begin{aligned} T^\varepsilon_{ij}\,(f)\,(x,t) &= -\int_{\mathbb{R}^{N+1}} \phi_\varepsilon(t-s)\partial^2_{x_i x_j}\Gamma(x,t;y,s) f(y,s)\, dy\, ds \\ &= \int_{\mathbb{R}^{N+1}} \phi_\varepsilon(t-s)\partial^2_{x_i x_j}\Gamma(x,t;y,s)\,[f(E(s-t)x,s) - f(y,s)]\, dy\, ds; \end{aligned}$$

hence, by using (2.17) in Theorem 2.5 we have

$$\begin{aligned} &\left|\left(T_{ij}\,(f) - T^\varepsilon_{ij}\,(f)\right)(x,t)\right| \\ &= \left|\int_{t-2\varepsilon}^{t} [1-\phi_\varepsilon(t-s)] \left(\int_{\mathbb{R}^N} \partial^2_{x_i x_j}\Gamma(x,t;y,s)\,[f(E(s-t)x,s) - f(y,s)]\, dy\right) ds\right| \\ &\le |f|_{C^\alpha_x(\mathbb{R}^{N+1})} \int_{t-2\varepsilon}^{t} \left(\int_{\mathbb{R}^N} \left|\partial^2_{x_i x_j}\Gamma(x,t;y,s)\,\|E(s-t)x - y\|^\alpha\right| dy\right) ds \\ &\le c\,|f|_{C^\alpha_x(\mathbb{R}^{N+1})}\,\varepsilon^{\alpha/2}. \end{aligned}$$

This proves (2.27), and allows us to rewrite (2.22) as follows:

$$\begin{aligned} \partial^2_{x_i x_j} u(x,t) &= \lim_{\varepsilon \to 0^+} \left(-\int_{\mathbb{R}^{N+1}} \phi_\varepsilon(t-s)\partial^2_{x_i x_j}\Gamma(x,t;y,s)\overline{\mathcal{L}}u(y,s)\, dy\, ds\right) \\ &\equiv \lim_{\varepsilon \to 0^+} T^\varepsilon_{ij}(\overline{\mathcal{L}}u), \end{aligned} \tag{2.28}$$

In view of the above facts, in order to prove Theorem 2.9 it then suffices to show the following result.

Proposition 2.11. *Let $p \in (1,\infty)$ be fixed, and let T^ε_{ij} be the operator defined in* (2.26). *Then, T^ε_{ij} can be extended to a linear and continuous operator from $L^p(\mathbb{R}^{N+1})$ into itself: there exists $c = c(\nu, p) > 0$ such that*

$$\left\|T^\varepsilon_{ij} f\right\|_{L^p(\mathbb{R}^{N+1})} \le c\,\|f\|_{L^p(\mathbb{R}^{N+1})} \quad \textit{for every } f \in L^p\left(\mathbb{R}^{N+1}\right). \tag{2.29}$$

We stress that the constant c appearing in (2.29) *is independent of* ε.

Before embarking on the (quite technical) proof of Proposition 2.11, let us show how this proposition allows us to easily prove Theorem 2.9.

Proof (of Theorem 2.9). We first observe that, given any $f \in \mathcal{D}_x^\alpha(\mathbb{R}^{N+1})$, by combining (2.29) with Fatou's lemma, we get

$$\begin{aligned}\int_{\mathbb{R}^{N+1}} |T_{ij}(f)|^p\,d\xi &= \int_{\mathbb{R}^{N+1}} \lim_{\varepsilon\to 0^+} |T_{ij}^\varepsilon(f)|^p\,d\xi \\ &\leq \liminf_{\varepsilon\to 0^+} \int_{\mathbb{R}^{N+1}} |T_{ij}^\varepsilon(f)|^p\,d\xi \leq c^p \|f\|^p_{L^p(\mathbb{R}^{N+1})}.\end{aligned}$$

From this, since $\mathcal{D}_x^\alpha(\mathbb{R}^{N+1})$ is dense in $L^p(\mathbb{R}^{N+1})$, we conclude that T_{ij} can be extended to a linear and continuous operator from $L^p(\mathbb{R}^{N+1})$ into itself.

In particular, given any $u \in C_0^\infty(\mathbb{R}^{N+1})$, since $f = \overline{\mathcal{L}}u \in \mathcal{D}_x^\alpha(\mathbb{R}^{N+1})$, from the representation formula (2.22) we get

$$\|\partial^2_{x_i x_j} u\|_{L^p(\mathbb{R}^{N+1})} = \|T_{ij}(\overline{\mathcal{L}}u)\|_{L^p(\mathbb{R}^{N+1})} \leq c\|\overline{\mathcal{L}}u\|_{L^p(\mathbb{R}^{N+1})}.$$

This ends the proof. □

We now turn to prove Proposition 2.11. We rely on the following abstract result from the theory of singular integrals.

Theorem 2.12 (See [10, Thm. 4.1])**.** *Let $K : \mathbb{R}^{N+1} \times \mathbb{R}^{N+1} \setminus \{\xi \neq \eta\} \to \mathbb{R}$ be a kernel satisfying the following conditions: there exist $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C} > 0$ such that:*

(i) *for every $\xi \neq \eta \in \mathbb{R}^{N+1}$ one has*

$$|K(\xi,\eta)| + |K(\eta,\xi)| \leq \frac{\boldsymbol{A}}{\left|B_{d(\xi,\eta)}(\xi)\right|};$$

(ii) *there exists a constant $\beta > 0$ such that for every $\xi, \xi_0, \eta \in \mathbb{R}^{N+1}$ satisfying the condition $d(\xi_0,\eta) \geq 4\boldsymbol{\kappa}\, d(\xi_0,\xi) > 0$, we have*

$$\begin{aligned}&|K(\xi,\eta) - K(\xi_0,\eta)| + |K(\eta,\xi) - K(\eta,\xi_0)| \\ &\qquad \leq \boldsymbol{B} \left(\frac{d(\xi_0,\xi)}{d(\xi_0,\eta)}\right)^\beta \cdot \frac{1}{\left|B_{d(\xi_0,\eta)}(\xi_0)\right|};\end{aligned}$$

(iii) *for every $\zeta \in \mathbb{R}^{N+1}$ and every $0 < r < R < \infty$, one has*

$$\left|\int_{\{r<d(\zeta,\eta)<R\}} K(\zeta,\eta)\,d\eta\right| + \left|\int_{\{r<d(\zeta,\xi)<R\}} K(\xi,\zeta)\,d\xi\right| \leq \boldsymbol{C}.$$

Suppose, in addition, that K_ε (for $\varepsilon > 0$) is a 'regularized kernel' defined in such a way that the following properties holds:

(a) *$K_\varepsilon(\xi,\cdot)$ and $K_\varepsilon(\cdot,\xi)$ are locally integrable for every $\xi \in \mathbb{R}^{N+1}$;*

(b) *K_ε satisfies the standard estimates (i), (ii), with constant bounded by $c'(\boldsymbol{A}+\boldsymbol{B})$ and c' absolute constant (independent of ε);*

(c) *there exists an absolute constant c' such that for every $\zeta \in \mathbb{R}^{N+1}$*

$$\left|\int_{\{r<d(\zeta,\eta)<R\}} K_\varepsilon(\zeta,\eta)\,d\eta\right| + \left|\int_{\{r<d(\zeta,\xi)<R\}} K_\varepsilon(\xi,\zeta)\,d\xi\right| \leq c'\boldsymbol{C}.$$

For every $f \in C_0^\alpha(\mathbb{R}^{N+1})$ *(see (1.37)) set*

$$T_\varepsilon f(\xi) = \int_{\mathbb{R}^{N+1}} K_\varepsilon(\xi,\eta) f(\eta)\, d\eta.$$

Then, for every $p \in (1,\infty)$*, the operator* T_ε *can be extended to a linear continuous operator on* $L^p(\mathbb{R}^{N+1})$*, and*

$$\|T_\varepsilon f\|_{L^p(\mathbb{R}^{N+1})} \le c\,\|f\|_{L^p(\mathbb{R}^{N+1})}$$

for every $f \in L^p(\mathbb{R}^{N+1})$*, with* $c > 0$ *independent of* ε*. Moreover, the constant* c *in the last estimate depends on the quantities involved in the assumptions as follows:*

$$c \le c'\,(\boldsymbol{A} + \boldsymbol{B} + \boldsymbol{C}),$$

with c' *'absolute' constant.*

Remark 2.13. It is worth mentioning that Theorem 2.12 can be applied in the present context since $(\mathbb{R}^{N+1}, d, |\cdot|)$ is a *space of homogeneous type*, in the sense of Coifman-Weiss (see Remark 1.1). In fact, such a result actually holds if we replace the space $(\mathbb{R}^{N+1}, d, |\cdot|)$ with any *space of homogeneous type* (X, d, μ). More precisely, the standard definition of space of homogeneous type requires $d\,(x,y)$ to be symmetric, while our d is actually only quasisymmetric (see (1.16)). However, this problem can be overcome in a standard way: given a quasisymmetric quasidistance d, the function

$$d'\,(x,y) = d\,(x,y) + d\,(y,x)$$

is a (symmetric) quasidistance, equivalent to d. Now, properties (i)-(ii) in Theorem 2.12 are clearly stable under replacement of d with an equivalent d'. This is also true for the cancellation property (iii), for a less obvious reason, which is discussed for instance in [10, Remark 4.6].

In order to deduce Proposition 2.11 from Theorem 2.12 we first note that, if T_{ij}^ε is the operator defined in (2.26), we can clearly write

$$T_{ij}^\varepsilon(f) = \int_{\mathbb{R}^{N+1}} K_{ij}^\varepsilon(\xi;\eta) f(\eta)\, d\eta,$$

where the kernel K_{ij}^ε is given by

$$K_{ij}^\varepsilon(\xi;\eta) = -\phi_\varepsilon(t-s)\partial^2_{x_i x_j}\Gamma(\xi;\eta). \tag{2.30}$$

Hence, if we set $K_{ij}(\xi;\eta) = -\partial^2_{x_i x_j}\Gamma(\xi;\eta)$, we can directly derive Proposition 2.11 from Theorem 2.12 as soon as we are able to prove the following facts:

1) K_{ij} satisfies properties (i)- (iii) in the statement of Theorem 2.12;
2) K_{ij}^ε satisfies properties (a)-(c) in the statement of Theorem 2.12.

Since the validity of assertion 1) follows directly from Theorem 2.5 (jointly with the fact that $|B_r(\xi)| = c\,r^{Q+2}$, see (1.20)), we turn to prove assertion 2).

Lemma 2.14. *The 'regularized kernel'* K_{ij}^ε *defined in* (2.30) *satisfies the properties* (a) - (c) *in the statement of Theorem 2.12.*

Proof. We prove the validity of the three properties separately.

- *Proof of property* (a). On account of Theorem 2.5 - 1), and recalling all the properties satisfied by ϕ_ε, for every fixed $\xi = (x,t) \in \mathbb{R}^{N+1}$ we have

$$\begin{aligned}|K^\varepsilon_{ij}(\xi,\cdot)| = |\phi_\varepsilon(t-s)\partial^2_{x_ix_j}\Gamma(\xi;\eta)| &\le c\,\mathbf{1}_{\{t-s\ge\varepsilon\}}(s)\,\frac{1}{d(\xi,\eta)^Q}\\ &\text{(by the explicit expression of } d\text{, see (1.14))}\\ &\le c\,\mathbf{1}_{\{t-s\ge\varepsilon\}}(s)\,\frac{1}{|t-s|^{Q/2}} \le \frac{c}{\varepsilon^{Q/2}},\end{aligned}$$

and this proves that $K^\varepsilon_{ij}(\xi,\cdot) \in L^1_{\mathrm{loc}}(\mathbb{R}^{N+1})$. Analogously one can show that $K^\varepsilon_{ij}(\cdot,\xi)$ is bounded, and thus locally integrable.

- *Proof of property* (b). We begin by proving that K^ε_{ij} satisfies the standard estimate (i) (with a constant $\le c'(\mathbf{A}+\mathbf{B})$ and c' independent of ε). To this end it suffices to observe that, since $0 \le \phi_\varepsilon \le 1$, for every $\xi,\eta \in \mathbb{R}^{N+1}$ we have

$$\begin{aligned}|K^\varepsilon_{ij}(\xi,\eta)| + |K^\varepsilon_{ij}(\eta,\xi)| &\le |\partial^2_{x_ix_j}\Gamma(\xi,\eta)| + |\partial^2_{x_ix_j}\Gamma(\eta,\xi)|\\ &\le \frac{\boldsymbol{A}}{d(\xi,\eta)^{Q+2}} = \frac{\boldsymbol{A}}{|B_{d(\xi,\eta)}(\xi)|},\end{aligned}$$

where we have used the fact that $K_{ij} = -\partial^2_{x_ix_j}\Gamma$ satisfies (i), see (2.14).

We then turn to prove the validity of the standard estimate (b), namely

$$\begin{aligned}&|K^\varepsilon_{ij}(\xi,\eta) - K^\varepsilon_{ij}(\xi_0,\eta)| + |K^\varepsilon_{ij}(\eta,\xi) - K^\varepsilon_{ij}(\eta,\xi_0)|\\ &\qquad \le \boldsymbol{B}\,\frac{d(\xi_0,\xi)}{d(\xi_0,\eta)}\cdot\frac{1}{|B_{d(\xi_0,\eta)}(\xi_0)|}\end{aligned} \tag{2.31}$$

for every $\xi_0,\xi,\eta \in \mathbb{R}^{N+1}$, provided that $d(\xi_0,\eta) > 4\boldsymbol{\kappa}d(\xi_0,\xi)$ (the choice $\beta = 1$ follows from the fact that K_{ij} satisfies (ii) with $\beta = 1$, see Theorem 2.5). We limit ourselves to prove the above estimate for the term

$$|K^\varepsilon_{ij}(\xi,\eta) - K^\varepsilon_{ij}(\xi_0,\eta)|$$

since the other one is totally analogous.

To begin with, we observe that, by definition of K^ε_{ij}, we have

$$\begin{aligned}&|K^\varepsilon_{ij}(\xi,\eta) - K^\varepsilon_{ij}(\xi_0,\eta)|\\ &\quad = |\phi_\varepsilon(t-s)\partial^2_{x_ix_j}\Gamma(x,t;y,s) - \phi_\varepsilon(t_0-s)\partial^2_{x_ix_j}\Gamma(x_0,t_0;y,s)|\\ &\quad \le |\partial^2_{x_ix_j}\Gamma(\xi;\eta) - \partial^2_{x_ix_j}\Gamma(\xi_0,\eta)|\,\phi_\varepsilon(t_0-s)\\ &\qquad + \partial^2_{x_ix_j}\Gamma(\xi;\eta)|\phi_\varepsilon(t_0-s) - \phi_\varepsilon(t-s)|\\ &\quad \equiv A_1 + A_2;\end{aligned}$$

moreover, since $0 \le \phi_\varepsilon \le 1$, from (2.15) and (1.20) we get

$$A_1 \le \boldsymbol{B}\frac{d(\xi_0,\xi)}{d(\xi_0,\eta)^{Q+3}} = \boldsymbol{B}\frac{d(\xi_0,\xi)}{d(\xi_0,\eta)}\frac{1}{|B_{d(\xi_0,\eta)}(\xi_0)|}. \tag{2.32}$$

Hence, we only need to estimate the term A_2. To this end we first notice that, since $\phi_\varepsilon \in C^\infty_0(\mathbb{R})$, by the mean value theorem we can write

$$|\phi_\varepsilon(t_0-s) - \phi_\varepsilon(t-s)| = |\phi'_\varepsilon(\tau)|\,|t_0-t|;$$

from this, since $\phi' \not\equiv 0$ only on $(\varepsilon, 2\varepsilon)$ and since $|\phi'_\varepsilon| \leq c/\varepsilon$ (for some absolute constant $c > 0$), we obtain the following estimates

$$\begin{aligned} &1)\ |\phi_\varepsilon(t_0 - s) - \phi_\varepsilon(t - s)| \leq c\,\frac{|t_0 - t|}{\varepsilon} \leq c\,\frac{|t_0 - t|}{|t - s|} \quad (\text{if } \varepsilon < t - s < 2\varepsilon) \\ &2)\ |\phi_\varepsilon(t_0 - s) - \phi_\varepsilon(t - s)| \leq c\,\frac{|t_0 - t|}{\varepsilon} \leq c\,\frac{|t_0 - t|}{|t_0 - s|} \quad (\text{if } \varepsilon < t_0 - s < 2\varepsilon). \end{aligned} \tag{2.33}$$

We then consider the two cases 1) - 2) separately. (Note that if none of these cases occurs, then $A_2 = 0$ and there is nothing to prove).

- **Case 1):** $\varepsilon < t - s < 2\varepsilon$. In this first case, owing to (2.13) we get

$$A_2 \leq c\frac{|t_0 - t|}{|t - s|}|\partial^2_{x_i x_j}\Gamma(\xi, \eta)| \leq c\frac{|t_0 - t|}{|t - s|^2}\Gamma_{c_1\nu^{-1}}(x, t; y, s) \leq \frac{c|t_0 - t|}{d(\xi, \eta)^{Q+4}};$$

from this, using (1.17) (and since the condition $d(\xi_0, \eta) \geq 4\boldsymbol{\kappa}\, d(\xi_0, \xi) > 0$ ensures that $d(\xi_0, \eta)$ and $d(\xi, \eta)$ are equivalent), we obtain

$$A_2 \leq c\frac{d\,(\xi_0, \xi)^2}{d\,(\xi, \eta)^{Q+4}} \leq c\frac{d\,(\xi_0, \xi)}{d\,(\xi_0, \eta)^{Q+3}}.$$

By combining this last estimate with (2.32), we obtain (2.31).

- **Case 2):** $\varepsilon < t_0 - s < 2\varepsilon$. In this second case, we first write

$$\begin{aligned} &|K^\varepsilon_{ij}(\xi, \eta) - K^\varepsilon_{ij}(\xi_0, \eta)| \\ &\quad = |\phi_\varepsilon(t - s)\partial^2_{x_i x_j}\Gamma(x, t; y, s) - \phi_\varepsilon(t_0 - s)\partial^2_{x_i x_j}\Gamma(x_0, t_0; y, s)| \\ &\quad \leq |\partial^2_{x_i x_j}\Gamma(\xi; \eta) - \partial^2_{x_i x_j}\Gamma(\xi_0, \eta)|\, \phi_\varepsilon(t - s) \\ &\qquad + \partial^2_{x_i x_j}\Gamma(\xi_0; \eta)|\phi_\varepsilon(t_0 - s) - \phi_\varepsilon(t - s)| \\ &\quad \equiv B_1 + B_2. \end{aligned}$$

Now, since $0 \leq \phi_\varepsilon \leq 1$, from (2.15) and (1.20) we get

$$B_1 \leq \boldsymbol{B}\frac{d(\xi_0, \xi)}{d(\xi_0, \eta)^{Q+3}} = \boldsymbol{B}\frac{d(\xi_0, \xi)}{d(\xi_0, \eta)}\frac{1}{|B_{d(\xi_0, \eta)}(\xi_0)|}. \tag{2.34}$$

Moreover, by arguing exactly as in the previous case, we have

$$\begin{aligned} B_2 &\leq c\frac{|t_0 - t|}{|t_0 - s|}|\partial^2_{x_i x_j}\Gamma(\xi_0; \eta)| \leq c\frac{|t_0 - t|}{|t_0 - s|^2}\Gamma_{c_1\nu^{-1}}(x_0, t_0; y, s) \\ &\leq \frac{c|t_0 - t|}{d(\xi_0, \eta)^{Q+4}} \leq c\frac{d(\xi_0, \xi)^2}{d(\xi_0, \eta)^{Q+4}} \leq c\frac{d(\xi_0, \xi)}{d(\xi_0, \eta)^{Q+3}}. \end{aligned}$$

By combining this last estimate with (2.34), we obtain (2.31) also in this case.

- *Proof of property* (c). Owing to the *cancellation property* of $K_{ij} = \partial^2_{x_i x_j}\Gamma$ contained in Theorem 2.5 - 3), for every $\zeta = (z,u) \in \mathbb{R}^{N+1}$ we have

$$\begin{aligned}
&\Big|\int_{\{r<d(\zeta,\xi)<R\}} K^{\varepsilon}_{ij}(\xi,\zeta)d\xi\Big| + \Big|\int_{\{r<d(\zeta,\eta)<R\}} K^{\varepsilon}_{ij}(\zeta,\eta)d\eta\Big| \\
&\qquad \text{(setting } \xi=(x,t),\ \eta=(y,s)\text{)} \\
&\qquad \le \int_u^{u+R^2} \Big|\phi_\varepsilon(t-u)\Big(\int_{\{x\in\mathbb{R}^N: r<d((z,u),(x,t))<R\}} \partial^2_{x_i x_j}\Gamma(x,t;z,u)dx\Big)\Big|\, dt \\
&\qquad\quad + \int_{u-R^2}^{u} \Big|\phi_\varepsilon(u-s)\Big(\int_{\{y\in\mathbb{R}^N: r<d((z,u),(y,s))<R\}} \partial^2_{x_i x_j}\Gamma(z,u;y,s)dy\Big)\Big|\, ds \\
&\qquad \le \int_u^{u+R^2} \Big|\int_{\{x\in\mathbb{R}^N: r<d((z,u),(x,t))<R\}} \partial^2_{x_i x_j}\Gamma(x,t;z,u)dx\Big|\, dt \\
&\qquad\quad + \int_{u-R^2}^{u} \Big|\int_{\{y\in\mathbb{R}^N: r<d((z,u),(y,s))<R\}} \partial^2_{x_i x_j}\Gamma(z,u;y,s)dy\Big|\, ds \\
&\qquad \le c,
\end{aligned}$$

for some constant $c > 0$ independent of ζ.

This completes the proof of Lemma 2.14 and therefore of Theorem 2.9. □

We are now ready to prove the main result of this Section:

Proof of Theorem 2.1. Given any $p \in (1,\infty)$, we know from Theorem 2.9 that the operator T_{ij} defined in (2.23) can be extended to a linear and continuous operator form $L^p(\mathbb{R}^{N+1})$ into itself; moreover, by (2.22) we have

$$\partial^2_{x_i x_j} u = T_{ij}(\overline{\mathcal{L}}u)$$

for every $u \in C_0^\infty(\mathbb{R}^{N+1})$ and $1 \le i,j \le q$.

Given any $f \in C_0^\infty(\mathbb{R}^{N+1})$, reasoning as in the proof of Theorem 2.9 we can write

$$T_{ij}(f)\,(\xi) = \lim_{\varepsilon\to 0^+}\left(-\int_{-\infty}^{t-\varepsilon}\int_{\mathbb{R}^N} \partial^2_{x_i x_j}\Gamma(\xi;\eta)f(\eta)\,d\eta\right).$$

On the other hand, if $\xi \notin \mathrm{supp}(f)$, the last limit equals the integral

$$-\int_{-\infty}^{t}\int_{\mathbb{R}^N} \partial^2_{x_i x_j}\Gamma(\xi;\eta)f(\eta)\,d\eta,$$

which is absolutely convergent. Therefore in this case

$$T_{ij}(f)\,(\xi) = \int_{\mathbb{R}^{N+1}} K_{ij}(\xi;\eta)f(\eta)\,d\eta.$$

Moreover, the kernel K_{ij} satisfies the mean value inequality (2.15).

Finally, owing to (1.22), we have $|B_{kr}(\xi)| \le c\,k^{Q+2}|B_r(\xi)|$, which is (1.34) with $\beta = Q+2$.

Hence we can apply Theorem 1.12 and conclude (2.2). So the theorem is proved. □

3. Operators with coefficients depending on (x,t)

With Theorem 2.1 at hand, we can now prove Theorem 1.4. Henceforth we assume that $\mathcal{L}$ is a KFP operator (1.1), with coefficients $a_{ij}(x,t)$ *depending on both space and time* and fulfilling assumptions **(H1)**, **(H2)**, **(H3)** stated in Section 1.

3.1. **Estimates on the mean oscillation of $\partial^2_{x_ix_j}u$ in terms of $\mathcal{L}u$.** According to [2], the first step for the proof of Theorem 1.4 consists in establishing a control on the mean oscillation of $\partial^2_{x_ix_j}u$ for functions $u\in C_0^\infty(\mathbb{R}^{N+1})$ *with small support*, in terms of $\mathcal{L}u$. To prove this result we combine Theorem 2.1 with the VMO_x assumption on the a_{ij}'s, following as far as possible Krylov' technique [21].

Theorem 3.1. *Let $p,\alpha,\beta\in(1,\infty)$ with $\alpha^{-1}+\beta^{-1}=1$. Then, there exists a constant $c>0$, depending on p,B,ν, such that for every $R,r>0$, $\xi^*,\xi_0,\overline{\xi}\in\mathbb{R}^{N+1}$ with $\xi_0\in B_r(\overline{\xi})$ and $u\in C_0^\infty(B_R(\xi^*))$, and every $k\geq 4\boldsymbol{\kappa}$, we have*

$$\begin{aligned}
&\frac{1}{|B_r(\overline{\xi})|}\int_{B_r(\overline{\xi})}\left|\partial^2_{x_ix_j}u(\xi)-(\partial^2_{x_ix_j}u)_{B_r(\overline{\xi})}\right|d\xi\\
&\leq\frac{c}{k}\sum_{h,l=1}^{q}\mathcal{M}(\partial^2_{x_ix_j}u)(\xi_0)+ck^{\frac{Q+2}{p}}(\mathcal{M}(|\mathcal{L}u|^p)(\xi_0))^{1/p}\\
&\quad+ck^{\frac{Q+2}{p}}a^\sharp(R)^{1/p\beta}\sum_{h,l=1}^{q}(\mathcal{M}(|\partial^2_{x_hx_l}u|^{p\alpha})(\xi_0))^{1/p\alpha}
\end{aligned}\tag{3.1}$$

for $i,j=1,2,...,q$. Recall that $a^\sharp(R)$ has been defined in (1.24).

Proof. We can assume that $B_r(\overline{\xi})\cap B_R(\xi^*)\neq\varnothing$, because otherwise

$$\int_{B_r(\overline{\xi})}\left|\partial^2_{x_ix_j}u(\xi)-(\partial^2_{x_ix_j}u)_{B_r(\overline{\xi})}\right|d\xi=0,$$

and thus there is nothing to prove.

We then observe that, if $\overline{A}=(\gamma_{hl}(t))_{hl}$ is any fixed matrix satisfying assumption **(H1)** (that is, $\gamma_{hl}\in L^\infty(\mathbb{R}^{N+1})$ and the ellipticity condition (1.3) holds), by Theorem 2.1 we can write (provided that k is large enough)

$$\begin{aligned}
&\frac{1}{|B_r(\overline{\xi})|}\int_{B_r(\overline{\xi})}\left|\partial^2_{x_ix_j}u(\xi)-(\partial^2_{x_ix_j}u)_{B_r(\overline{\xi})}\right|d\xi\\
&\leq c\Big\{\frac{1}{k}\mathcal{M}(\overline{\mathcal{L}}u)(\xi_0)\\
&\quad+k^{\frac{Q+2}{p}}\Big(\frac{1}{|B_{kr}(\overline{\xi})|}\int_{B_{kr}(\overline{\xi})}|\overline{\mathcal{L}}u(x,t)|^p\,dxdt\Big)^{1/p}\Big\},
\end{aligned}\tag{3.2}$$

where $\overline{\mathcal{L}}$ is the KFP operator with coefficient matrix $\overline{A}$, that is,

$$\overline{\mathcal{L}}=\textstyle\sum_{i,j=1}^q\gamma_{ij}(t)\partial^2_{x_ix_j}+Y.\tag{3.3}$$

We now turn to bound the right hand-side of (3.2). First, to handle the term $\mathcal{M}\left(\overline{\mathcal{L}}u\right)$, let us write, by (3.3),

$$\overline{\mathcal{L}}u=\sum_{i,j=1}^{q}\gamma_{ij}(t)\partial^2_{x_ix_j}u+\mathcal{L}u-\sum_{i,j=1}^{q}a_{ij}(x,t)\partial^2_{x_ix_j}u$$

so that

$$\mathcal{M}\left(\overline{\mathcal{L}}u\right) \le \mathcal{M}\left(\mathcal{L}u\right) + c\sum_{i,j=1}^{q}\mathcal{M}\left(\partial^2_{x_ix_j}u\right) \tag{3.4}$$

with c only depending on ν, q.

To handle the second term at the right hand side of (3.2), we write

$$\|\overline{\mathcal{L}}u\|_{L^p(B_{kr}(\overline{\xi}))} \le \|\mathcal{L}u\|_{L^p(B_{kr}(\overline{\xi}))} + \|\overline{\mathcal{L}}u - \mathcal{L}u\|_{L^p(B_{kr}(\overline{\xi}))}$$

and we exploit the fact that, since $\xi_0 \in B_r(\overline{\xi})$, we have

$$\Big(\frac{1}{|B_{kr}(\overline{\xi})|}\int_{B_{kr}(\overline{\xi})}|\mathcal{L}u(\xi)|^p\,d\xi\Big)^{1/p} \le \big(\mathcal{M}(|\mathcal{L}u|^p)(\xi_0)\big)^{1/p};$$

as a consequence of these facts, we get

$$\begin{aligned}
&\frac{1}{|B_r(\overline{\xi})|}\int_{B_r(\overline{\xi})}\big|\partial^2_{x_ix_j}u(\xi) - (\partial^2_{x_ix_j}u)_{B_r(\overline{\xi})}\big|\,d\xi \\
&\quad\le \frac{c}{k}\left\{\sum_{h,l=1}^{q}\mathcal{M}(\partial^2_{x_hx_l}u)(\xi_0) + \mathcal{M}\left(\mathcal{L}u\right)(\xi_0)\right\} + ck^{\frac{Q+2}{p}}\big(\mathcal{M}(|\mathcal{L}u|^p)(\xi_0)\big)^{1/p} \\
&\qquad + ck^{\frac{Q+2}{p}}\frac{1}{|B_{kr}(\overline{\xi})|^{1/p}}\|\overline{\mathcal{L}}u - \mathcal{L}u\|_{L^p(B_{kr}(\overline{\xi}))}.
\end{aligned} \tag{3.5}$$

Furthermore, we have (setting, as usual, $\xi = (x,t)$)

$$\begin{aligned}
&\int_{B_{kr}(\overline{\xi})}|\overline{\mathcal{L}}u(\xi) - \mathcal{L}u(\xi)|^p\,d\xi \\
&\quad\le c\sum_{h,l=1}^{q}\Big(\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)}|\gamma_{hl}(t) - a_{hl}(x,t)|^{p\beta}\,dxdt\Big)^{1/\beta}\times \\
&\qquad\times\Big(\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)}|\partial^2_{x_hx_l}u(\xi)|^{p\alpha}\,d\xi\Big)^{1/\alpha} \\
&\quad\text{(since the coefficients } \gamma_{hl},\ a_{hl} \text{ are bounded by } \nu^{-1}) \\
&\quad\le c(2\nu^{-1})^{p-1/\beta}\sum_{h,l=1}^{q}\Big(\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)}|\gamma_{hl}(t) - a_{hl}(x,t)|\,dxdt\Big)^{1/\beta}\times \\
&\qquad\times\Big(\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)}|\partial^2_{x_hx_l}u(\xi)|^{p\alpha}\,d\xi\Big)^{1/\alpha}.
\end{aligned} \tag{3.6}$$

Now, since the above estimates hold *for every fixed matrix* $\overline{A} = (\gamma_{hl}(t))_{hl}$ satisfying assumption **(H1)**, we choose the particular matrix $(\gamma_{hl}(t))_{hl}$, depending on the fixed quantities $r, k, R, \xi^*, \overline{\xi}$, defined as follows

$$\gamma_{hl}(t) = \begin{cases} (a_{hl}(\cdot,t))_{B_R(\xi^*)} & \text{if } kr \ge R \\ (a_{hl}(\cdot,t))_{B_{kr}(\overline{\xi})} & \text{if } kr \le R. \end{cases}$$

We recall that, according to Definition 1.2, we have

$$(a_{hl}(\cdot,t))_B = \frac{1}{|B|}\int_B a_{hl}(x,t)\,dx\,ds \quad \text{for every } d\text{-ball } B \subseteq \mathbb{R}^{N+1}.$$

We then observe that, since the functions $a_{hl} \in L^\infty\left(\mathbb{R}^{N+1}\right)$, by Fubini's theorem we see that γ_{hl} is measurable for every $1 \leq h, l \leq q$; on account of this fact, and since $A_0 = (a_{hl})_{hl}$ fulfills assumption **(H1)**, we easily conclude that also the chosen matrix $\overline{A} = (\gamma_{hl})_{hl}$ satisfies **(H1)**, and we can exploit the above estimates (3.5) - (3.6) with this choice of A.

As to estimate (3.6), we notice that by Definition 1.2 we have

$$\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)} \left|a_{hl}(x,t) - \gamma_{hl}(t)\right| dxdt \leq \begin{cases} |B_R(\xi^*)| a^\sharp(R) & \text{if } kr \geq R \\ |B_{kr}(\overline{\xi})| a^\sharp(R) & \text{if } kr \leqslant R. \end{cases}$$

On the other hand, since we are assuming that $B_r(\overline{\xi})$ and $B_R(\xi^*)$ intersect, in the case $kr \geq R$ we can write, by the doubling condition,

$$|B_R(\xi^*)| \leq |B_{kr}(\xi^*)| \leq c\left|B_{kr}\left(\overline{\xi}\right)\right|;$$

as a consequence, we obtain

$$\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)} \left|a_{hl}(x,t) - \gamma_{hl}(t)\right| dxdt \leq c|B_{kr}(\overline{\xi})| a^\sharp(R).$$

Summing up, estimate (3.6) with this choice of $\overline{A}$ boils down to

$$\begin{aligned} &\int_{B_{kr}(\overline{\xi})} |\overline{\mathcal{L}}u(\xi) - \mathcal{L}u(\xi)|^p \, d\xi \\ &\quad \leq c\,|B_{kr}(\overline{\xi})|^{1/\beta} a^\#(R)^{1/\beta} \sum_{h,l=1}^q \left(\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)} |\partial^2_{x_h x_l} u(\xi)|^{p\alpha} \, d\xi\right)^{1/\alpha}. \end{aligned} \tag{3.7}$$

With estimate (3.7) at hand, we are now ready to conclude the proof of the theorem: indeed, by combining (3.4), (3.5), (3.7), we obtain

$$\begin{aligned} &\frac{1}{|B_r(\overline{\xi})|}\int_{B_r(\overline{\xi})} \left|\partial^2_{x_i x_j} u(\xi) - (\partial^2_{x_i x_j} u)_{B_r(\overline{\xi})}\right| d\xi \\ &\quad \leq \frac{c}{k}\left\{\sum_{i,j=1}^q \mathcal{M}(\partial^2_{x_i x_j} u)(\xi_0) + \mathcal{M}\left(\mathcal{L}u\right)(\xi_0)\right\} + ck^{\frac{Q+2}{p}}\left(\mathcal{M}(|\mathcal{L}u|^p)(\xi_0)\right)^{1/p} \\ &\qquad + ck^{\frac{Q+2}{p}} |B_{kr}(\overline{\xi})|^{\frac{1}{p\beta} - \frac{1}{p}} a^\sharp(R)^{1/p\beta} \times \\ &\qquad\quad \times \sum_{h,l=1}^q \Big(\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)} |\partial^2_{x_h x_l} u(\xi)|^{p\alpha} \, d\xi\Big)^{1/p\alpha} \end{aligned} \tag{3.8}$$

furthermore, recalling that $\alpha^{-1}+\beta^{-1}=1$ and using (1.29) (note that $\xi_0\in B_{kr}(\overline{\xi})$), we can bound the the last line of (3.8) as

$$
\begin{aligned}
&|B_{kr}(\overline{\xi})|^{\frac{1}{p\beta}-\frac{1}{p}} a^{\sharp}\,(R)^{1/p\beta}\sum_{h,l=1}^{q}\Big(\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)}|\partial^2_{x_hx_l}u(\xi)|^{p\alpha}\,d\xi\Big)^{1/p\alpha}\\
&= a^{\sharp}\,(R)^{1/p\beta}\sum_{h,l=1}^{q}\Big(\frac{1}{|B_{kr}(\overline{\xi})|}\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)}|\partial^2_{x_hx_l}u(\xi)|^{p\alpha}\,d\xi\Big)^{1/p\alpha}\\
&\le a^{\sharp}\,(R)^{1/p\beta}\sum_{h,l=1}^{q}\big(\mathcal{M}(|\partial^2_{x_ix_j}u|^{p\alpha})(\xi_0)\big)^{1/p\alpha},
\end{aligned}
\tag{3.9}
$$

for some c also depending on ν. Note also that in the right-hand side of (3.8), the term

$$
\frac{c}{k}\mathcal{M}\left(\mathcal{L}u\right)(\xi_0)\ \text{ can be absorbed in } ck^{\frac{Q+2}{p}}\left(\left(\mathcal{M}\left(|\mathcal{L}u|^p\right)\right)(\xi_0)\right)^{1/p}
$$

since k is large and for every ball $B_\rho\ni\xi_0$ we can write

$$
\frac{1}{|B_\rho|}\int_{B_\rho}|\mathcal{L}u\,(x)|\,dx\le\left(\frac{1}{|B_\rho|}\int_{B_\rho}|\mathcal{L}u\,(x)|^p\,dx\right)^{1/p}\le\left(\left(\mathcal{M}\left(|\mathcal{L}u|^p\right)\right)(\xi_0)\right)^{1/p},
$$

hence

$$
\mathcal{M}\left(\mathcal{L}u\right)(\xi_0)\le\left(\left(\mathcal{M}\left(|\mathcal{L}u|^p\right)\right)(\xi_0)\right)^{1/p}. \tag{3.10}
$$

Gathering (3.8), (3.9) and (3.10), we finally obtain the desired (3.1), and the proof is complete. □

3.2. **L^p estimates.** With Theorem 3.1 at hand, we can easily prove *local L^p estimates* for functions $u\in C_0^\infty(\mathbb{R}^{N+1})$ with *small support*:

Theorem 3.2. *For every $p\in(1,\infty)$ there exist constants $R,c>0$ such that, for every ball $B_R(\xi^*)$ in $\mathbb{R}^{N+1}$ and every $u\in C_0^\infty(B_R(\xi^*))$, we have*

$$
\sum_{h,l=1}^{q}\|u_{x_hx_l}\|_{L^p(B_R(\xi^*))}\le c\,\|\mathcal{L}u\|_{L^p(B_R(\xi^*))}. \tag{3.11}
$$

The constants c,R in (3.11) depends on the numbers p,ν,B and on the function $a^\sharp$ in (1.24), but do not depend on ξ^*.

Proof. The present theorem can be established by arguing *exactly* as in the proof of [2, Thm. 4.2]; the idea is to combine the mean-oscillation estimates in Theorem 3.1 (which is the analog of [2, Thm. 4.1]) with the Hardy-Littlewood maximal inequality and the Fefferman-Stein-like maximal inequality for spaces of homogeneous type (see Theorem 1.9 and Theorem 1.10, respectively), and to exploit assumption **(H3)** to take to the left hand side the term involving $a^\sharp\,(R)$ in (3.1). □

Starting from Theorem 3.2, we can then establish the following global a priori estimates, which are a 'less refined' version of Theorem 1.4.

Theorem 3.3. *For every $p\in(1,\infty)$ there exists $c>0$, depending on p,ν, B and on the function $a^\sharp$, such that, for every $u\in C_0^\infty(\mathbb{R}^{N+1})$, one has*

$$
\|u\|_{W^{2,p}_X(\mathbb{R}^{N+1})}\le c\big\{\,\|\mathcal{L}u\|_{L^p(\mathbb{R}^{N+1})}+\|u\|_{W^{1,p}_X(\mathbb{R}^{N+1})}\,\big\}.
$$

Moreover, for every $R > 0$ and every $u \in C^\infty(\mathbb{R}^{N+1})$ we have

$$\|u\|_{W^{2,p}_X(B_R(0))} \leq c_1 \big\{ \|\mathcal{L}u\|_{L^p(B_{2R}(0))} + \|u\|_{W^{1,p}_X(B_{2R}(0))} \big\},$$

where $c_1 > 0$ is a constant independent of R.

According to [2], in order to prove the above result we need to show the existence of suitable cutoff functions and of a suitable covering of $\mathbb{R}^{N+1}$.

We begin with the existence of an ad-hoc family of cutoff functions.

Lemma 3.4. *For every fixed $R > 0$, there exist a constant $c > 0$ and a family $\{\phi^\xi(\cdot)\}_{\xi\in\mathbb{R}^{N+1}}$ of cutoff functions in $\mathbb{R}^{N+1}$ such that*

i) $\phi^\xi \in C_0^\infty(B_{2R}(\xi))$

ii) $\phi^\xi = 1$ *on* $B_R(\xi)$ *and* $0 \leq \phi^\xi \leq 1$ *on* $\mathbb{R}^{N+1}$;

iii) $\displaystyle \sup_\eta |\phi^\xi(\eta)| + \sum_{i=1}^q \sup_\eta |\partial_{y_i}\phi^\xi(\eta)| + \sum_{i,j=1}^q \sup_\eta |\partial^2_{y_iy_j}\phi^\xi(\eta)| + \sup_\eta |Y_y\phi^\xi(\eta)| \leq c$

(where $\eta = (y, s)$). Here, the relevant fact is that the constant c is independent of ξ (while it may depend on R).

Proof. Let $\psi \in C_0^\infty(\mathbb{R}^{N+1})$ be a fixed cutoff function such that

a) $0 \leq \psi \leq 1$ on $\mathbb{R}^{N+1}$;

b) $\psi = 1$ on $B_1(0)$ and $\psi = 0$ out of $B_2(0)$

Given any $\xi \in \mathbb{R}^{N+1}$, we define the function

$$\phi^\xi(\eta) = \psi\big(D(1/R)(\xi^{-1} \circ \eta)\big),$$

and we claim that it satisfies the required properties i) - iii).

Clearly, we have $\psi^\xi \in C_0^\infty(\mathbb{R}^{N+1})$ and $0 \leq \psi^\xi \leq 1$; moreover, using (1.14)-(1.19) (and taking into account property b) of ψ), we get

- $\phi^\xi = 1$ on $\xi \circ D(R)(B_1(0)) = B_R(\xi)$;
- $\psi^\xi = 0$ out of $\xi \circ D(R)(B_2(0)) = B_{2R}(\xi)$.

Finally, recalling that ∂_{y_i} (with $1 \leq i \leq q$) and Y are left-invariant and $D(\lambda)$-homogeneous of degree 1 and 2, respectively, for every $\eta \in \mathbb{R}^{N+1}$ we obtain

$$\begin{aligned}
&|\phi^\xi(\eta)| + \sum_{i=1}^q |\partial_{y_i}\phi^\xi(\eta)| + \sum_{i,j=1}^q |\partial^2_{y_iy_j}\phi^\xi(\eta)| + |Y_y\phi^\xi(\eta)| \\
&\qquad (\text{setting } \zeta = D(1/R)(\xi^{-1} \circ \eta)) \\
&\qquad = |\psi(\zeta)| + R^{-1}\sum_{i=1}^q |(\partial_i\psi)(\zeta)| + R^{-2}\Big(\sum_{i,j=1}^q |(\partial^2_{ij}\psi)(\zeta)| + |(Y\psi)(\zeta)|\Big) \\
&\qquad \leq c,
\end{aligned}$$

where $c > 0$ depends on ψ and R. This ends the proof. □

As for the existence of a suitable covering of $\mathbb{R}^{N+1}$, instead, we recall the following *covering theorem in spaces of homogeneous type.*

Proposition 3.5 (See [2, Prop. 2.19])**.** *Let (X, d, μ) be a space of homogeneous type. For every fixed $R > 0$ there exists a family*

$$\mathcal{B} = \{B(x_\alpha, R)\}_{\alpha \in A}$$

of d-balls in $\mathbb{R}^{N+1}$ satisfying the following properties:

i) $\bigcup_{\alpha \in A} B(x_\alpha, R) = X$*;*

ii) *for every $H > 1$, the family of dilated balls*

$$\mathcal{B}^H = \{B(x_\alpha, HR)\}_{\alpha \in A}$$

has the bounded overlapping property*, that is, there exists a constant N (depending on H and the constants of X, but independent of R) such that every point of X belongs to at most N balls $B(x_\alpha, HR)$.*

With Lemma 3.4 and Proposition 3.5 at hand, we can provide the

Proof of Theorem 3.3. The present theorem can be established by arguing exactly as in the proof of [2, Thm. 4.3] and [2, Cor. 4.4]; the idea is to apply the local L^p estimates in Theorem 3.2 to the function $u\phi^\xi$ (where ϕ^ξ is as in Lemma 3.4), and then to exploit Proposition 3.5. □

Thanks to Theorem 3.3, we can finally come to the

Proof of Theorem 1.4. Taking into account all the results established so far, the present theorem can be demonstrated by arguing exactly as in the proof of [2, Thm. 1.4]: first of all, one extends the estimate in Theorem 3.3 to functions $u \in W_X^{2,p}(\mathbb{R}^{N+1})$ by using the *local approximation of $W_X^{2,p}(\mathbb{R}^{N+1})$ by smooth functions* (see [8, Thm. 2.9]); then, one gets rid of the term

$$\|u\|_{W_X^{1,p}(\mathbb{R}^{N+1})}$$

by exploiting the following known (Euclidean!) interpolation inequality,

$$\|\partial_{x_i} u\|_{L^p(\mathbb{R}^{N+1})} \leq \varepsilon \|\partial^2_{x_i x_i} u\|_{L^p(\mathbb{R}^{N+1})} + \frac{c_p}{\varepsilon}\|u\|_{L^p(\mathbb{R}^{N+1})}, \tag{3.12}$$

holding true for every $p \in (1, \infty)$, every $u \in W_X^{2,p}(\mathbb{R}^{N+1})$, every $1 \leq i, j \leq q$ and every $\varepsilon > 0$, with a constant $c_p > 0$ only depending on p. □

4. Existence results

In this last section we exploit the global $W_X^{2,p}$ - estimates contained in Theorem 1.4 to prove some existence results for the equation

$$\mathcal{L}u - \lambda u = f \text{ in } S_T = \mathbb{R}^N \times (-\infty, T),$$

with $f \in L^p(S_T)$, for $\lambda > 0$ large enough, and every $T \in (-\infty, +\infty]$, and for the Cauchy problem

$$\text{(CP)} \qquad \begin{cases} \mathcal{L}u = f & \text{in } \mathbb{R}^N \times (0, T) \\ u(\cdot, 0) = g & \text{in } \mathbb{R}^N \end{cases}$$

with $f \in L^p(\mathbb{R}^N \times (0, T))$, $g \in W_X^{2,p}(\mathbb{R}^N)$ for every $T \in (0, +\infty)$ (see Section 4.3 for the precise definition of solution). In both cases, the solutions we obtain are actually *strong solutions*, that is, they belong to some $W_X^{2,p}$ - space, and the equation $\mathcal{L}u = f$ holds pointwise (a.e.).

In order to prove our results we mainly follow the approach by Krylov [22] for parabolic equations, which essentially consists in the following steps.

(1) First of all, we refine the $W_X^{2,p}$ - estimates in Theorem 1.4 in two different directions: we localize such estimates on any strip $S_T = \mathbb{R}^N \times (-\infty, T)$, and we get rid of the term $\|u\|_{L^p}$ on the right-hand side (by paying the price of considering the 'perturbed operator' $\mathcal{L} - \lambda$ for some $\lambda > 0$ large enough).

(2) Using these refined estimates and the method of continuity, we then prove the existence of a unique $W_X^{2,p}(S_T)$ - solution of the equation
$$\mathcal{L}u - \lambda u = f \quad \text{in } S_T,$$
for an arbitrary $f \in L^p\left(\mathbb{R}^{N+1}\right)$ (and for every $-\infty < T \leq +\infty$).

(3) Finally, using the above solvability result (and the refined estimates in point (1)), we establish the existence of a unique solution of the Cauchy problem (CP).

Throughout this section, we adopt the notation:

- given any $T \in (-\infty, +\infty]$, we set $S_T = \mathbb{R}^N \times (-\infty, T)$;
- given any $T \in (0, +\infty)$, we set $\Omega_T = \mathbb{R}^N \times (0, T)$.

4.1. **Refined $W_X^{2,p}$ - estimates.** We start with the following theorem, localizing the global estimates of Theorem 1.4 on any strip S_T.

Theorem 4.1. *Let $\mathcal{L}$ be an operator as in* (1.1)*, assume that* **(H1)**, **(H2)**, **(H3)** *hold, and let $p \in (1, \infty)$.*

Then, there exists a constant $c > 0$ (depending on p, the matrix B in (1.5)*, the number ν in* (1.3)*, and the function $a^{\#}$ in* (1.24)*) such that, for every $T \in (-\infty, +\infty]$,*
$$\|u\|_{W_X^{2,p}(S_T)} \leq c\left\{\|\mathcal{L}u\|_{L^p(S_T)} + \|u\|_{L^p(S_T)}\right\}$$
for every $u \in W_X^{2,p}(S_T)$. Note that the constant c does not depend on T.

The proof of Theorem 4.1 follows by carefully revising the one of Theorem 1.4, and by exploiting the following proposition.

Proposition 4.2. *There exists $c > 0$ such that if, for every $\xi \in S_T$ and $r > 0$, we let*
$$B_r^T(\xi_0) = B_r(\xi_0) \cap S_T,$$
then;
$$\text{i) } \left|B_{2r}^T(\xi)\right| \leq c\left|B_r^T(\xi)\right|$$
$$\text{ii) } \left|B_{kr}^T(\xi)\right| \leq ck^{Q+2}\left|B_r^T(\xi)\right| \text{ for every } k \geq 1.$$

Proof. We will show that, for every $\xi \in S_T$ and $r > 0$,
$$\left|B_r^T(\xi)\right| \geq \frac{1}{2}\left|B_r(\xi)\right|. \tag{4.1}$$
This implies
$$\left|B_{2r}^T(\xi)\right| \leq \left|B_{2r}(\xi)\right| \leq c\left|B_r(\xi)\right| \leq 2c\left|B_r^T(\xi)\right|$$
and also
$$\left|B_{kr}^T(\xi)\right| \leq \left|B_{kr}(\xi)\right| \leq ck^{Q+2}\left|B_r(\xi)\right| \leq 2ck^{Q+2}\left|B_r^T(\xi)\right|.$$

To prove (4.1), let us write, for $\xi=(x,t)\,,\eta=(y,s)$

$$\left|B_r^T(\xi)\right|=\int_{\rho(\eta^{-1}\circ\xi)<r,s<T}d\xi$$

letting $\eta^{-1}\circ\xi=\zeta=(z,\sigma)\,,$ with $\sigma=t-s$

$$=\int_{\rho(\zeta)<r,\sigma>t-T}d\zeta$$

since $\xi\in S_T$, we have $t-T<0$, hence

$$\begin{aligned}&\geq\int_{\rho(\zeta)<r,\sigma>0}d\zeta=\int_{\|z\|+\sqrt{|\sigma|}<r,\sigma>0}dzd\sigma\\&=\frac{1}{2}\int_{\|z\|+\sqrt{|\sigma|}<r}dzd\sigma=\frac{1}{2}\left|B_r(\xi)\right|.\end{aligned}$$

This ends the proof. □

With Proposition 4.2 at hand, we can provide the

Proof of Theorem 4.1. First of all we observe that, as we read from Proposition 4.2, $(S_T,d,|\cdot|)$ is a space of homogeneous type; in particular, the Hardy-Littlewood maximal operator $\mathcal{M}^T$ (defined on S_T w.r.t. the balls B_r^T) maps the space $L^p(S_T)$ into itself, and the same is true for the analogous sharp maximal operator (since S_T is still unbounded). Moreover, if T_{ij} is the singular - integral operator defined w.r.t. the differential operator $\overline{\mathcal{L}}$ (see (2.23)), we have:

$$T_{ij}(f)=T_{ij}(f\cdot\mathbf{1}_{S_T})\ \text{ in } S_T.$$

As a consequence, we have

$$\|T_{ij}(f)\|_{L^p(S_T)}\leq c\,\|f\|_{L^p(S_T)}\ \text{ for every } f\in L^p(S_T)\ \text{ and } p\in(1,\infty).$$

Gathering all these facts, and revising carefully the proofs of the results we have established on the entire space $\mathbb{R}^{N+1}$, we can then infer the following facts.

a) Theorem 2.1 holds with

$$B_r\mapsto B_r^T;\qquad \mathcal{M}\mapsto\mathcal{M}^T.$$

b) In the proof of Theorem 3.1, we can replace

$$B_r\mapsto B_r^T;\qquad \mathcal{M}\mapsto\mathcal{M}^T.$$

Thus, we have

$$\begin{aligned}&\int_{B_{kr}^T(\overline{\xi})\cap B_R^T(\xi^*)}|\gamma_{hl}(t)-a_{hl}(x,t)|\,dxdt\\&\qquad\leq\int_{B_{kr}(\overline{\xi})\cap B_R(\xi^*)}|\gamma_{hl}(t)-a_{hl}(x,t)|\,dxdt\\&\qquad\text{(reasoning as in the proof of Theorem 3.1)}\\&\qquad\leq c\left|B_{kr}(\overline{\xi})\right|a^{\#}(R)\leq 2c\left|B_{kr}^T(\overline{\xi})\right|a^{\#}(R).\end{aligned}$$

and the proof can be concluded at the same way.

c) On account of a), b), Theorem 3.2 can be proved with $B_R \mapsto B_R^T$, by arguing exactly in the same way and exploiting the continuity of the maximal Hardy Littlewood and sharp maximal operators over S_T.

In view of these facts, and since the interpolation inequalities (3.12) hold for every fixed time (hence, they also hold on the strip S_T), we can conclude the proof of Theorem 4.1 by proceeding as in the proof of Theorem 1.4. □

Now that we have established Theorem 4.1, we proceed by proving the second improvement of our estimates announced at the beginning of the section.

Theorem 4.3. *Let $\mathcal{L}$ be an operator as in* (1.1)*, assume that* **(H1)**, **(H2)**, **(H3)** *hold, and let $p \in (1,\infty)$. Then, there exist positive constants c, λ_0 (depending on p, the matrix B in* (1.5)*, the number ν in* (1.3)*, and the function $a^{\#}$ in* (1.24)*) such that*

$$\|u\|_{W_X^{2,p}(S_T)} \le c\|\mathcal{L}u - \lambda u\|_{L^p(S_T)} \tag{4.2}$$

for every $T \in (-\infty, +\infty]$, $u \in W_X^{2,p}(S_T)$ and $\lambda \ge \lambda_0$. Note that the constants c, λ_0 do not depend on T.

It should be noticed that the appearence of the L^p - norm of $\mathcal{L}_\lambda u$ (instead of that of $\mathcal{L}$) in the right-hand side of (4.2) is somehow unavoidable: indeed, in the special case when $\mathcal{L}$ has *constant coefficients* and $T = +\infty$, a simple homogeneity argument shows that an estimate of the form (4.2) *cannot hold* for $\mathcal{L}$.

Proof. To ease the readability (and to simplify the notation), we only consider the case $T = +\infty$ (the case $T < +\infty$ being completely analogous). To begin with, let $\widetilde{\mathcal{L}}$ be the KFP operator defined on $\mathbb{R}^{N+2} = \mathbb{R}_y \times \mathbb{R}^{N+1}_{(x,t)}$ as

$$\widetilde{\mathcal{L}} = \mathcal{L} + \partial^2_{yy} = \partial^2_{yy} + \sum_{i,j=1}^{q} a_{ij}(x,t)\partial^2_{x_i x_j} + Y.$$

Clearly, this operator $\widetilde{\mathcal{L}}$ satisfies assumptions **(H1)** - **(H3)**, with

$$\widetilde{X} = \{\partial_y, \partial_{x_1}, \dots, \partial_{x_n}, Y\}, \qquad \widetilde{A}_0(x,t) = \begin{pmatrix} 1 & \mathbb{O} \\ \mathbb{O} & A_0(x,t) \end{pmatrix};$$

hence, we can apply Theorem 1.4 to $\widetilde{\mathcal{L}}$, obtaining the estimate

$$\|u\|_{W^{2,p}_{\widetilde{X}}(\mathbb{R}^{N+2})} \le c\big\{\|\widetilde{\mathcal{L}}u\|_{L^p(\mathbb{R}^{N+2})} + \|u\|_{L^p(\mathbb{R}^{N+2})}\big\}, \tag{4.3}$$

which holds for every function $u = u(y,x,t) \in W^{2,p}_{\widetilde{X}}(\mathbb{R}^{N+2})$.

Let now $\phi(y) \in C_0^\infty(-1,1)$, $\phi \not\equiv 0$, and let $u(x,t) \in W_X^{2,p}(\mathbb{R}^{N+1})$. We set

$$\widetilde{u}(y,x,t) = u(x,t)\,\phi(y)\,e^{i\sqrt{\lambda}y} \in W_X^{2,p}(\mathbb{R}^{N+2}).$$

We can apply the bound (4.3) to $\widetilde{u}$ and $\widetilde{\mathcal{L}}$ on $\mathbb{R}^{N+2}$ (since $\widetilde{\mathcal{L}}$ is linear with real coefficients, the bound extends to complex valued functions):

$$\|\widetilde{u}\|_{W^{2,p}_{\widetilde{X}}(\mathbb{R}^{N+2})} \le c\left\{\|\widetilde{\mathcal{L}}\widetilde{u}\|_{L^p(\mathbb{R}^{N+2})} + \|\widetilde{u}\|_{L^p(\mathbb{R}^{N+2})}\right\}. \tag{4.4}$$

Now, by definition of $\widetilde{u}$,

$$\|\widetilde{u}\|_{W^{2,p}_{\widetilde{X}}(\mathbb{R}^{N+2})} \geq c\,\|u\|_{W^{2,p}_{X}(\mathbb{R}^{N+1})} + c\,\|u\|_{L^p(\mathbb{R}^{N+1})} \left\|\phi'' + 2i\sqrt{\lambda}\phi' - \lambda\phi\right\|_{L^p(-1,1)}$$

and

$$\left\|\phi'' + 2i\sqrt{\lambda}\phi' - \lambda\phi\right\|_{L^p(-1,1)} \geq \|\phi'' - \lambda\phi\|_{L^p(-1,1)} \geq c\lambda$$

for λ large enough depending on ϕ, and $c > 0$ depending on ϕ. Hence

$$\|\widetilde{u}\|_{W^{2,p}_{X}(\mathbb{R}^{N+2})} \geq c\left\{\|u\|_{W^{2,p}_{X}(\mathbb{R}^{N+1})} + \lambda\,\|u\|_{L^p(\mathbb{R}^{N+1})}\right\}. \tag{4.5}$$

Also,

$$\|\widetilde{u}\|_{L^p(\mathbb{R}^{N+2})} \leq c\|u\|_{L^p(\mathbb{R}^{N+1})}$$

and

$$\begin{aligned}\widetilde{\mathcal{L}}\widetilde{u}\,(y,x,t) &= \mathcal{L}u\,(x,t)\cdot\phi\,(y)\,e^{i\sqrt{\lambda}y} + u\,(x,t)\,\partial^2_{yy}\left(\phi\,(y)\,e^{i\sqrt{\lambda}y}\right)\\ &= e^{i\sqrt{\lambda}y}\left\{(\mathcal{L}u-\lambda u)\,(x,t)\cdot\phi\,(y) + u\,(x,t)\left(\phi''\,(y) + 2i\sqrt{\lambda}\phi'\,(y)\right)\right\}\end{aligned}$$

so that

$$\|\widetilde{\mathcal{L}}\widetilde{u}\|_{L^p(\mathbb{R}^{N+2})} \leq c\left\{\|\mathcal{L}u-\lambda u\|_{L^p(\mathbb{R}^{N+1})} + \left(1+\sqrt{\lambda}\right)\|u\|_{L^p(\mathbb{R}^{N+1})}\right\}. \tag{4.6}$$

Then (4.4)-(4.5)-(4.6) imply:

$$\begin{aligned}&\|u\|_{W^{2,p}_{X}(\mathbb{R}^{N+1})} + \lambda\,\|u\|_{L^p(\mathbb{R}^{N+1})}\\ &\qquad\leq c\left\{\|\mathcal{L}u-\lambda u\|_{L^p(\mathbb{R}^{N+1})} + \left(1+\sqrt{\lambda}\right)\|u\|_{L^p(\mathbb{R}^{N+1})}\right\}.\end{aligned} \tag{4.7}$$

Now, there exists $\lambda_0 > 0$ such that

$$c\left(1+\sqrt{\lambda}\right) \leq \frac{\lambda}{2} \text{ for every } \lambda \geq \lambda_0.$$

For these λ and λ_0, and the same c as in (4.7), we get

$$\|u\|_{W^{2,p}_{X}(\mathbb{R}^{N+1})} \leq c\|\mathcal{L}u-\lambda u\|_{L^p(\mathbb{R}^{N+1})},$$

so we are done. □

4.2. **Solvability of $\mathcal{L}u - \lambda u = f$.** We now turn to study the solvability of the equation

$$\mathcal{L}u - \lambda u = f \quad \text{in } S_T, \tag{4.8}$$

where $-\infty < T \leq +\infty$, $f \in L^p(S_T)$ and $\lambda \geq \lambda_0$ (with $\lambda_0 > 0$ as in Theorem 4.3). As anticipated, our approach relies on the method of continuity, which allows us to relate the solvability of (4.8) with the solvability of

$$\mathcal{K}u - \lambda u = f \quad \text{in } S_T, \tag{4.9}$$

where $\mathcal{K}$ is the *constant-coefficient* Kolmogorov operator, that is,

$$\mathcal{K} = \Delta_{\mathbb{R}^q} + Y. \tag{4.10}$$

Thus, we begin by investigating the solvability of (4.9). To this end, following the approach by Krylov [22, Chap. 2], we first prove the subsequent results.

Proposition 4.4 (Liouville-type property). *Let $\lambda > 0$ be fixed, and let $\mathcal{K}_\lambda = \mathcal{K} - \lambda$ (with $\mathcal{K}$ as in (4.10)). Moreover, let $\mathcal{K}_\lambda^*$ be the formal adjoint of $\mathcal{K}_\lambda$, that is,*

$$\mathcal{K}_\lambda^* = \Delta_{\mathbb{R}^q} - Y - \lambda. \tag{4.11}$$

If $f \in C^\infty(\mathbb{R}^{N+1}) \cap L^\infty(\mathbb{R}^{N+1})$ is such that

$$\mathcal{K}_\lambda^* f = 0 \text{ in } \mathbb{R}^{N+1},$$

then we necessarily have $f \equiv 0$ in $\mathbb{R}^{N+1}$.

The previous proposition is stated for the adjoint operator $\mathcal{K}_\lambda^*$ just because we will need it in this form; the same property holds for $\mathcal{K}_\lambda$ as well.

In order to prove Proposition 4.4 we need the following lemma.

Lemma 4.5. *There exists a function $F : \mathbb{R} \to \mathbb{R}$ such that $F \in C^\infty(\mathbb{R})$, $F(z) = 1$ for $z \in [-1, 1]$ and, for some constant $c > 0$,*

$$\left|\frac{F''(z)}{F(z)}\right| \leq c \quad \text{and} \quad \left|\frac{F'(z)}{F(z)}\right| \leq c\,|z| \qquad \text{for every } z \in \mathbb{R}.$$

Proof. Let $\phi \in C^\infty(\mathbb{R})$ be such that

i) $0 \leq \phi \leq 1$ pointwise on $\mathbb{R}$;
ii) $\phi \equiv 0$ on $[-1, 1]$ and $\phi \equiv 1$ on $\mathbb{R} \setminus [-2, 2]$,

and let

$$F(z) = \cosh(z\phi(z)).$$

Then

$$F'(z) = \sinh(z\phi(z))(\phi(z) + z\phi'(z));$$
$$F''(z) = \cosh(z\phi(z))(\phi(z) + z\phi'^2 + \sinh(z\phi(z))(2\phi'(z) + z\phi''(z)).$$

From this, using the properties i) - ii) of ϕ and observing that $\phi'(z), \phi''(z) \neq 0$ only for $1 < |z| < 2$, we get

$$\left|\frac{F''(z)}{F(z)}\right| \leq (\phi(z) + z\phi'^2 + |\tanh(z\phi(z))| \cdot |2\phi'(z) + z\phi''(z)| \leq c.$$

Moreover, since $|\tanh(z)| \leq |z|$ for every $z \in \mathbb{R}$, we get

$$\left|\frac{F'(z)}{F(z)}\right| \leq |\tanh(z\phi(z))| \cdot |\phi(z) + z\phi'(z)| \leq |z|\big(\phi(z)^2 + \phi(z)|z\phi'(z)|\big) \leq c\,|z|\,.$$

This ends the proof. □

With Lemma 4.5, we can provide the

Proof of Proposition 4.4. We argue by contradiction, assuming that there exists some $\bar{\xi} \in \mathbb{R}^{N+1}$ such that $f(\bar{\xi}) \neq 0$. Following [22], we then fix $\varepsilon \in (0, 1)$ (to be chosen conveniently small later on), and we introduce the auxiliary function

$$\zeta(x, t) = F(\varepsilon\hat{p}(x))\cosh(\varepsilon t),$$

where $F(z)$ is as in Lemma 4.5, and $\hat{p}$ is a homogeneous norm which is globally equivalent to $\|\cdot\|$, and it is also smooth outside the origin, e.g.

$$\hat{p}(x) = \left(\sum_{i=1}^{N} |x_i|^{2q_N!/q_i}\right)^{1/(2q_N!)}.$$

Accordingly, we define $g = f/\zeta$.

We now observe that, since $\hat{p}$ is smooth outside the origin and since $F \equiv 1$ in $[-1, 1]$, we have $\zeta \in C^\infty(\mathbb{R}^{N+1})$; hence, the same is true of g (notice that $\zeta \geq 1$). Moreover, since f is bounded, we have

$$g(\xi) \to 0 \text{ as } \rho(\xi) = \|x\| + \sqrt{|t|} \to +\infty.$$

In particular, there exists $\xi_0 \in \mathbb{R}^{N+1}$ such that

$$g(\xi_0) = \max_{\mathbb{R}^{N+1}} g > 0.$$

Finally, since $\mathcal{K}_\lambda^* f = 0$ on $\mathbb{R}^{N+1}$, we have

$$\begin{aligned} 0 = \mathcal{K}_\lambda^* f = \mathcal{K}_\lambda^*(\zeta g) = \zeta \mathcal{K}^*(g) + 2\sum_{i=1}^q \partial_{x_i}\zeta\partial_{x_i} g + c\,(\xi)\, g, \\ \text{where } c(\xi) = \mathcal{K}_\lambda^*(\zeta) = \Delta_{\mathbb{R}^q}\zeta - \langle Bx, \nabla\zeta\rangle + \partial_t\zeta - \lambda\zeta. \end{aligned} \tag{4.12}$$

To proceed further we claim that, by choosing $\varepsilon > 0$ small enough, we have

$$c(\xi) < 0 \text{ pointwise on } \mathbb{R}^{N+1}. \tag{4.13}$$

Indeed, since the vector fields ∂_{x_i} (for $1 \leq i \leq q$) and $Z = \langle Bx, \nabla\rangle$ are homogeneous of degree 1 and 2, respectively (notice that $Z = Y + \partial_t$, and both Y and ∂_t are homogeneous of degree 2), and since the norm $\hat{p}$ is $D_0(\lambda)$-homogeneous of degree 1, by a direct computation we get:

$$\begin{aligned} \Delta_{\mathbb{R}^q}\zeta &= \cosh(\varepsilon t)\Delta_{\mathbb{R}^q}\big[x \mapsto F\big(\hat{p}(D_0(\varepsilon)x)\big)\big] \\ &= \cosh(\varepsilon t)\Delta_{\mathbb{R}^q}\big((F \circ \hat{p}) \circ D_0(\varepsilon)\big)(x) \\ &= \cosh(\varepsilon t) \cdot \varepsilon^2 \Delta_{\mathbb{R}^q}(F \circ \hat{p})(D_0(\varepsilon)x) \\ &= \cosh(\varepsilon t)\big(\varepsilon^2(F'' \circ \hat{p})|\nabla_{\mathbb{R}^q}\hat{p}|^2 + \varepsilon^2(F' \circ \hat{p})\Delta_{\mathbb{R}^q}\hat{p}\big)(D_0(\varepsilon)x) \\ &= \zeta\Big(\varepsilon^2\frac{F''}{F}(\varepsilon\hat{p}(x))|\nabla_{\mathbb{R}^q}\hat{p}|^2(D_0(\varepsilon)x) + \varepsilon^2\frac{F'}{F}(\varepsilon\hat{p}(x))\Delta_{\mathbb{R}^q}\hat{p}(D_0(\varepsilon)x)\Big). \end{aligned}$$

Since $|\nabla_{\mathbb{R}^q}\hat{p}|$ and $\Delta_{\mathbb{R}^q}\hat{p}$ are $D_0\,(\lambda)$-homogeneous of degrees 0 and -1, respectively, using the estimates in Lemma 4.5 we get

$$|\Delta_{\mathbb{R}^q}\zeta| \leq \kappa\zeta\Big(\varepsilon^2 + \frac{\varepsilon^3\hat{p}(x)}{\hat{p}(D_0(\varepsilon)x)}\Big) = 2\kappa\zeta\varepsilon^2.$$

Next, we compute

$$\begin{aligned} |Z\zeta| &= \cosh(\varepsilon t) \cdot \big|Z\big[x \mapsto F\big(\hat{p}(D_0(\varepsilon)x)\big)\big]\big| \\ &= \cosh(\varepsilon t) \cdot \big|Z\big((F \circ \hat{p}) \circ D_0(\varepsilon)\big)(x)\big| \\ &= \cosh(\varepsilon t) \cdot \varepsilon^2\big|Z(F \circ \hat{p})(D_0(\varepsilon)x)\big| \\ &= \cosh(\varepsilon t) \cdot \varepsilon^2\big|F'(\varepsilon\hat{p}(x)) \cdot (Z\hat{p})(D_0(\varepsilon)x)\big| \\ &= \zeta\varepsilon^2\Big|\frac{F'}{F}(\varepsilon\hat{p}(x)) \cdot (Z\hat{p})(D_0(\varepsilon)x)\Big| \\ &\qquad \text{(again by the estimates in Lemma 4.5)} \\ &\leq \kappa\zeta \cdot \frac{\varepsilon^3\hat{p}(x)}{\hat{p}(D_0(\varepsilon)x)} = \kappa\zeta\varepsilon^2; \end{aligned}$$

$$\partial_t\zeta = \varepsilon\sinh(\varepsilon t)F(\varepsilon\hat{p}(x)) = \zeta\big(\varepsilon\tanh(\varepsilon t)\big);$$

for some 'universal' constant $\kappa > 0$ independent of ε.

From this, since $\tanh(z) \leq 1$ for all $z \geq 0$ (and since $c(\xi)$ is a smooth function on the whole of $\mathbb{R}^{N+1}$), we conclude that

$$\begin{aligned} c(\xi) &= \Delta_{\mathbb{R}^q}\zeta - \langle Bx, \nabla\zeta\rangle + \partial_t\zeta - \lambda\zeta \\ &\leq \Delta_{\mathbb{R}^q}\zeta + |Z\zeta| + \partial_t\zeta - \lambda\zeta \\ &\leq \zeta\big(3\kappa\varepsilon^2 + \varepsilon - \lambda\big) < 0, \end{aligned}$$

provided that $\varepsilon > 0$ is sufficiently small.

With (4.13) at hand, we can finally come to the end of the proof. In fact, since the point ξ_0 is a (global) maximum point for g, we clearly have

$$\mathcal{K}^* g(\xi_0) = \Delta_{\mathbb{R}^q} g(\xi_0) - \langle Bx, \nabla g(\xi_0)\rangle + \partial_t g(\xi_0) = \Delta_{\mathbb{R}^q} g(\xi_0) \leq 0;$$

on the other hand, by combining (4.12) - (4.13), we have, since $\partial_{x_i} g\,(x_0) = 0$,

$$0 = \zeta\mathcal{K}^*(g)(\xi_0) + c(\xi_0)g(\xi_0) \leq c(\xi_0)g(\xi_0) < 0$$

(recall that $g(\xi_0) > 0$), which is clearly a contradiction. □

Proposition 4.6. *Let $\lambda > 0$ be fixed, and let $\mathcal{K}_\lambda = \mathcal{K} - \lambda$. Then,*

$$X = \mathcal{K}_\lambda(C_0^\infty(\mathbb{R}^{N+1})) \text{ is dense in } L^p(\mathbb{R}^{N+1}).$$

Proof. We argue by contradiction, assuming that the vector space X is not dense in $L^p(\mathbb{R}^{N+1})$. Then, by the Hahn-Banach theorem, there exists a nontrivial linear continuous functional on $L^p(\mathbb{R}^{N+1})$ which vanishes on X. Hence, by Riesz' representation theorem, we can find some non-zero function $f \in L^q(\mathbb{R}^{N+1})$ (where q is the conjugate exponent of p, that is, $1/q = 1 - 1/p$) such that

$$\int_{\mathbb{R}^{N+1}} f(\xi)\mathcal{K}_\lambda\varphi(\xi)\,d\xi = 0 \quad \text{for every } \varphi \in C_0^\infty(\mathbb{R}^{N+1}). \tag{4.14}$$

This means that

$$\mathcal{K}_\lambda^* f = 0 \text{ in } \mathcal{D}'(\mathbb{R}^{N+1}); \tag{4.15}$$

thus, since $\mathcal{K}_\lambda^*$ is C^∞-hypoelliptic, we derive that $f \in C^\infty(\mathbb{R}^{N+1})$, and (4.15) holds pointwise on $\mathbb{R}^{N+1}$. Despite this fact, we cannot directly apply Proposition 4.4 to deduce that $f \equiv 0$, since we do not know that *f is bounded*. To overcome this issue, following the argument in [5, Prop. 5.3.9] we introduce the mollified function f_ε defined by

$$f_\varepsilon(\xi) = \int_{\mathbb{R}^{N+1}} f(\eta)J_\varepsilon(\xi\circ\eta^{-1})\,d\eta$$

(with J_ε the usual family of mollifiers 'adapted' to the group $\mathbb{G}$). Then, we easily see that $f_\varepsilon \in C^\infty(\mathbb{R}^{N+1})$ and we can write, for every $\phi \in C_0^\infty\left(\mathbb{R}^{N+1}\right)$,

$$\begin{aligned} \int_{\mathbb{R}^{N+1}} f_\varepsilon(\xi)\mathcal{K}_\lambda\phi\,(\xi)\;d\xi &= \int_{\mathbb{R}^{N+1}} \left(\int_{\mathbb{R}^{N+1}} f(\eta^{-1}\circ\xi)J_\varepsilon(\eta)\,d\eta\right)\mathcal{K}_\lambda\phi\,(\xi)\;d\xi \\ &= \int_{\mathbb{R}^{N+1}} J_\varepsilon(\eta)\left(\int_{\mathbb{R}^{N+1}} f(\eta^{-1}\circ\xi)\mathcal{K}_\lambda\phi\,(\xi)\;d\xi\right)d\eta \\ &= \int_{\mathbb{R}^{N+1}} J_\varepsilon(\eta)\left(\int_{\mathbb{R}^{N+1}} f(\xi)\mathcal{K}_\lambda\phi\,(\eta\circ\xi)\;d\xi\right)d\eta = 0 \end{aligned}$$

because, by (4.14),

$$\int_{\mathbb{R}^{N+1}} f(\xi)\mathcal{K}_\lambda\phi\,(\eta\circ\xi)\;d\xi = \int_{\mathbb{R}^{N+1}} f(\xi)\mathcal{K}_\lambda\phi^\eta\,(\xi)\;d\xi = 0$$

since the function $\phi^\eta\,(\xi) = \phi\,(\eta\circ\xi)$ belongs to $C_0^\infty(\mathbb{R}^{N+1})$. As a consequence, we have $\mathcal{K}_\lambda^*(f_\varepsilon) = 0$ in $\mathcal{D}'(\mathbb{R}^{N+1})$ and thus, by the C^∞-hypoellipticity of $\mathcal{K}_\lambda^*$,

$$f_\varepsilon \in C^\infty(\mathbb{R}^{N+1}) \quad \text{and} \quad \mathcal{K}_\lambda^*(f_\varepsilon) = 0 \text{ pointwise in } \mathbb{R}^{N+1}.$$

On the other hand, since $f \in L^q(\mathbb{R}^{N+1})$, we easily see that the (smooth) function f_ε is bounded (for every fixed $\varepsilon > 0$): in fact, we have

$$\begin{aligned} |f_\varepsilon(\xi)| &\leq \|f\|_{L^q(\mathbb{R}^{N+1})}\left(\int_{\mathbb{R}^{N+1}} J_\varepsilon^p(\xi\circ\eta^{-1})\,d\eta\right)^{1/p} \\ &\qquad (\text{setting } \eta = \zeta^{-1}\circ\xi, \text{ and noting that } d\eta = d\zeta) \\ &= \|f\|_{L^q(\mathbb{R}^{N+1})}\left(\int_{\mathbb{R}^{N+1}} J_\varepsilon^p(\zeta)\,d\zeta\right)^{1/p} = c\,(\varepsilon) < +\infty. \end{aligned}$$

Hence, we can apply Proposition 4.4, obtaining

$$f_\varepsilon \equiv 0 \text{ pointwise in } \mathbb{R}^{N+1}.$$

From this, since $f_\varepsilon \to f$ as $\varepsilon \to 0^+$ in $L^1_{\mathrm{loc}}(\mathbb{R}^{N+1})$ (see, e.g., [5, Rem. 5.3.8]), we conclude that $f \equiv 0$ a.e. in $\mathbb{R}^{N+1}$, but this is a contradiction. □

Thanks to the above results, we can then prove the solvability of (4.9).

Theorem 4.7 (Solvability of $\mathcal{K}_\lambda u = f$)**.** *Let $\lambda \geq \lambda_0$ be fixed (where $\lambda_0 > 0$ is as in Theorem 4.1), and let $-\infty < T \leq +\infty$ and $p \in (1,\infty)$. Then, for every $f \in L^p(S_T)$ there exists a* unique *function $u \in W_X^{2,p}(S_T)$ such that*

$$\mathcal{K}u - \lambda u = f \text{ a.e. in } S_T. \tag{4.16}$$

Moreover, we have the estimate

$$\|u\|_{W_X^{2,p}(S_T)} \leq c\,\|f\|_{L^p(S_T)}, \tag{4.17}$$

with the constant c as in Theorem 4.3.

Proof. First of all we observe that, if a solution $u \in W_X^{2,p}(S_T)$ of equation (4.16) does exist (for some fixed $f \in L^p(\mathbb{R}^{N+1})$), by Theorem 4.3 we have

$$\|u\|_{W_X^{2,p}(S_T)} \leq c\,\|\mathcal{K}u - \lambda u\|_{L^p(S_T)} = c\,\|f\|_{L^p(S_T)};$$

this proves at once the validity of the estimate (4.17), and the uniqueness part of the theorem.

As for the existence part, we split the proof into two steps.

- *Step I).* Let us first prove the solvability of (4.16) assuming that $T = +\infty$ (that is, when $S_T = \mathbb{R}^{N+1}$). In this case, given any $f \in L^p(\mathbb{R}^{N+1})$, we know from Proposition 4.6 that there exists a sequence $\{\varphi_n\}_n \subseteq C_0^\infty(\mathbb{R}^{N+1})$ such that

$$f_n = (\mathcal{K}-\lambda)\varphi_n \to f \quad \text{as } n \to +\infty \text{ in } L^p(\mathbb{R}^{N+1});$$

on the other hand, from Theorem 4.3 we infer that

$$\begin{aligned} \|\varphi_n - \varphi_m\|_{W_X^{2,p}(\mathbb{R}^{N+1})} &\leq c\,\|(\mathcal{K}-\lambda)(\varphi_n-\varphi_m)\|_{L^p(\mathbb{R}^{N+1})} \\ &= c\,\|f_n - f_m\|_{L^p(\mathbb{R}^{N+1})}, \end{aligned}$$

and this shows that $\{\varphi_n\}_n$ is a *Cauchy sequence* in *the Banach space* $W^{2,p}_X(\mathbb{R}^{N+1})$. Thus, we can find some $u \in W^{2,p}_X(\mathbb{R}^{N+1})$ such that

$$\varphi_n \to u \text{ as } n \to +\infty \text{ in } W^{2,p}_X(\mathbb{R}^{N+1}). \tag{4.18}$$

Now, since $\mathcal{K} = \Delta_{\mathbb{R}^q} + Y$, by (4.18) (and the definition of $W^{2,p}_X$, see Definition 1.3) we infer that $f_n = (\mathcal{K} - \lambda)\varphi_n \to (\mathcal{K} - \lambda)u$ as $n \to +\infty$ in $L^p(\mathbb{R}^{N+1})$; from this, since we also have $f_n \to f$ as $n \to +\infty$ in $L^p(\mathbb{R}^{N+1})$, we conclude that

$$\mathcal{K}u - \lambda u = f \text{ a.e. in } \mathbb{R}^{N+1},$$

and this proves that u is a solution of (4.16).

- *Step II).* Let us now turn to prove the solvability of equation (4.16) in the case $T < +\infty$. Let then $f \in L^p(S_T)$ be fixed, and let

$$g = f \cdot \mathbf{1}_{S_T}.$$

Since, obviously, we have $g \in L^p(\mathbb{R}^{N+1})$, from the previous *Step I)* we know that there exists a (unique) function $v \in W^{2,p}(\mathbb{R}^{N+1})$ such that

$$\mathcal{K}v - \lambda v = g \quad \text{a.e. in } \mathbb{R}^{N+1}.$$

Then, setting $u = v|_{S_T}$, we immediately see that $u \in W^{2,p}(S_T)$, and that u is a solution of equation (4.16) (since $g = f$ a.e. in S_T). This ends the proof. □

By combining Theorem 4.7 with the localized estimates in Theorem 4.1, we can now prove the solvability of (4.8).

Theorem 4.8 (Solvability of $\mathcal{L}u - \lambda u = f$)**.** *Let $\lambda \geq \lambda_0$ be fixed (where $\lambda_0 > 0$ is as in Theorem 4.1), and let $-\infty < T \leq +\infty$. Moreover, let $\mathcal{L}$ be an operator as in* (1.1)*, assume that* **(H1)**, **(H2)**, **(H3)** *hold and let $p \in (1, \infty)$. Then, for every $f \in L^p(S_T)$ there exists a* unique *function $u \in W^{2,p}_X(S_T)$ such that*

$$\mathcal{L}u - \lambda u = f \text{ a.e. in } S_T. \tag{4.19}$$

Moreover, we have the estimate

$$\|u\|_{W^{2,p}_X(S_T)} \leq c\,\|f\|_{L^p(S_T)} \tag{4.20}$$

with the constant c as in Theorem 4.3.

Proof. We are going to prove the theorem by using the method of continuity. To this end we consider, for every fixed $r \in [0, 1]$, the operator

$$\mathcal{P}_r : W^{2,p}_X(S_T) \to L^p(S_T), \qquad \mathcal{P}_r u = \mathcal{L}_r u - \lambda u,$$

where $\mathcal{L}_r = (1 - r)\mathcal{K} + r\mathcal{L}$, that is,

$$\mathcal{L}_r = \sum_{i,j=1}^{q} \big((1 - r)\delta_{ij} + r a_{ij}(x, t)\big)\partial^2_{x_i x_j} + Y \equiv \sum_{i,j=1}^{q} a^r_{ij}(x, t)\partial^2_{x_i x_j} + Y.$$

We then observe that, since the coefficients a_{ij} are globally bounded in $\mathbb{R}^{N+1}$ (see assumption **(H1)**), and since $0 \leq r \leq 1$, for every $u \in W^{2,p}_X(S_T)$ we get

$$\begin{aligned}\|\mathcal{P}_r u\|_{L^p(S_T)} &\leq \sum_{i,j=1}^q \|a_{ij}^r \partial^2_{x_i x_j} u\|_{L^p(S_T)} + \|Yu\|_{L^p(S_T)} + \lambda \|u\|_{L^p(S_T)} \\ &\leq (1 + \max_{i,j} \|a_{ij}\|_{L^\infty(S_T)}) \sum_{i,j=1}^q \|\partial^2_{x_i x_j} u\|_{L^p(S_T)} \\ &\qquad + \|Yu\|_{L^p(S_T)} + \lambda \|u\|_{L^p(S_T)} \\ &\leq c\, \|u\|_{W^{2,p}_X(S_T)},\end{aligned}$$

and this proves that $\mathcal{P}_r$ is a (linear and) continuous operator from $W^{2,p}(S_T)$ into $L^p(S_T)$, with norm bounded independently of r. Moreover, by Theorem 4.3 there exists a constant $c > 0$, independent of r, such that

$$\|u\|_{W^{2,p}_X(S_T)} \leq c\, \|\mathcal{P}_r u\|_{L^p(S_T)}. \tag{4.21}$$

We stress that we are entitled to apply Theorem 4.3 to the operator $\mathcal{P}_r$, since it is a KFP operator of the form (1.1), satisfying assumptions **(H1)**, **(H2)**, **(H3)**, with uniform constants as r ranges in $[0, 1]$. We can then apply the method of continuity: since we know from Theorem 4.7 that the operator $\mathcal{P}_0 = \mathcal{K} - \lambda$ is surjective, we conclude that the same is true of the operator $\mathcal{P}_1 = \mathcal{L} - \lambda$, that is, for every fixed $f \in L^p(\mathbb{R}^{N+1})$ there exists a unique function $u \in W^{2,p}_X(S_T)$ such that

$$\mathcal{L}u - \lambda u = f \text{ a.e. in } S_T.$$

Finally, the validity of estimate (4.20) follows from (4.21) (with $r = 1$). □

4.3. **The Cauchy problem for $\mathcal{L}$.** We can now prove our well-posedness result for the $\mathcal{L}$-Cauchy problem, that is, Theorem 1.7; we refer to Section 1.3 for the definition of *solution of the Cauchy problem*, and for the definition of the involved function spaces $\mathring{W}^{2,p}_X(\Omega_T)$, $W^{2,p}_X(\mathbb{R}^N)$.

Proof of Theorem 1.7. We split the proof into three steps.

- *Step I).* Here we prove the *existence part* of the theorem, together with the validity of estimate (1.28), in the special case $g = 0$.

To this end, we fix $f \in L^p(\Omega_T)$, and we set

$$h = e^{-\lambda_0 t} f \cdot \mathbf{1}_{[0,T]},$$

where $\lambda_0 > 0$ is as in Theorem 4.8. (As usual, the symbol $f \cdot \mathbf{1}_{[0,T]}$ means that the function f has been defined in the whole strip S_T, letting it equal to zero for $t < 0$). Since, obviously, $h \in L^p(S_T)$, by Theorem 4.8 we know that there exists a (unique) $v \in W^{2,p}_X(S_T)$ such that

$$\mathcal{L}v - \lambda_0 v = h \text{ a.e. in } S_T;$$

moreover, by applying the estimate (4.20) on the strip S_0 (to the 'restricted' function $v|_{S_0} \in W^{2,p}_X(S_0)$), we have

$$\|v\|_{W^{2,p}(S_0)} \leq c\, \|\mathcal{L}v - \lambda_0 v\|_{L^p(S_0)} = \|h\|_{L^p(S_0)} = 0,$$

from which we derive that

$$v = 0 \text{ a.e. in } S_0. \tag{4.22}$$

Then, we set $u = e^{\lambda_0 t}v$, and we claim that u is a solution to the Cauchy problem (1.26). Indeed, since $v \in W^{2,p}_X(S_T)$, also $u \in W^{2,p}_X(S_T)$ (here it is important that $T < \infty$), and by (4.22) we conclude that

$$u \in \mathring{W}^{2,p}_X(\Omega_T).$$

Moreover, since $\mathcal{L}v - \lambda_0 v = h$, by a direct computation we have

$$\begin{aligned} \mathcal{L}u &= \mathcal{L}(e^{\lambda_0 t}v) = e^{\lambda_0 t}(\mathcal{L}v - \lambda_0 v) = e^{\lambda_0 t}h \\ &\text{(by definition of } h) \\ &= f \cdot \mathbf{1}_{[0,T]} = f \quad \text{a.e. in } \Omega_T, \end{aligned}$$

and this proves that u is a solution of the Cauchy problem (1.26), as claimed.

As for the validity of estimate (1.28) it suffices to note that, by applying (4.20) on the strip S_T to v (and exploiting again the finiteness of T), we get

$$\begin{aligned} \|u\|_{W^{2,p}_X(\Omega_T)} &= \|e^{\lambda_0 t}v\|_{W^{2,p}_X(\Omega_T)} \le c\,(T)\,\|v\|_{W^{2,p}_X(S_T)} \\ &\le c\,\|\mathcal{L}v - \lambda_0 v\|_{L^p(S_T)} = c\,\|h\|_{L^p(S_T)} \\ &= c\,\|e^{-\lambda_0 t}f\|_{L^p(\Omega_T)} \le c\,\|f\|_{L^p(\Omega_T)}, \end{aligned}$$

for some absolute constant $c > 0$ (possibly different from line to line).

- *Step II).* We now consider general data $f \in L^p(\Omega_T)$ and $g \in W^{2,p}_X(\mathbb{R}^N)$ for the Cauchy problem. The function $\widetilde{f} = f - \mathcal{L}g$ clearly belongs to $L^p(\Omega_T)$; hence, by *Step I)*, we can find $v \in \mathring{W}^{2,p}_X(\Omega_T)$ such that v is a solution to

$$\begin{cases} \mathcal{L}v = f - \mathcal{L}g & \text{in } \Omega_T \\ v(\cdot, 0) = 0 & \text{in } \mathbb{R}^N \end{cases} \tag{4.23}$$

and

$$\|v\|_{W^{2,p}_X(\Omega_T)} \le c\,\left\|\widetilde{f}\right\|_{L^p(\Omega_T)} \le c\,\left\{\|f\|_{L^p(\Omega_T)} + \|g\|_{W^{2,p}_X(\mathbb{R}^N)}\right\}. \tag{4.24}$$

Define $u(x,t) = v(x,t) + g(x)$. Then $u \in W^{2,p}_X(\Omega_T)$ and $\mathcal{L}u = f$ in Ω_T. Moreover, $u - g = v \in \mathring{W}^{2,p}_X(\Omega_T)$, hence u is a solution to the Cauchy problem (1.26).

Finally,

$$\|u\|_{W^{2,p}_X(\Omega_T)} \le \|v\|_{W^{2,p}_X(\Omega_T)} + \|g\|_{W^{2,p}_X(\Omega_T)}. \tag{4.25}$$

By (1.27), (4.24), and (4.25), we get (1.28).

- *Step III).* We now prove the *uniqueness part* of the theorem. This *does not follow* directly from estimate (1.28), since this estimate has been proved *only for the solution* u constructed in the previous *Steps I-II).* However, let $u_1, u_2 \in W^{2,p}_X(\Omega_T)$ be two solutions of the Cauchy problem (1.26). Then the function $\bar{u} = u_1 - u_2$ belongs to $\mathring{W}^{2,p}_X(\Omega_T)$ and solves $\mathcal{L}\bar{u} = 0$ in Ω_T. Hence also the function

$$v = e^{-\lambda_0 t}\bar{u}$$

belongs to $\mathring{W}^{2,p}_X(\Omega_T)$. Moreover, by proceeding as in *Step I)*, we have

$$\mathcal{L}v - \lambda_0 v = e^{-\lambda_0 t}\mathcal{L}\bar{u} = 0 \quad \text{a.e. in } \Omega_T.$$

As a consequence, by applying (4.2) to the strip S_T and to the function v, we obtain the following estimate (recall that $v = 0$ a.e. in S_0)

$$\|v\|_{W^{2,p}_X(S_T)} \leq c\,\|\mathcal{L}v - \lambda_0 v\|_{L^p(S_T)} = c\,\|\mathcal{L}v - \lambda_0 v\|_{L^p(\Omega_T)} = 0.$$

This proves that $v = 0$ a.e. in S_T, and thus

$$u_1 = u_2 \text{ a.e. in } S_T.$$

This ends the proof. □

Remark 4.9. It is clear from the above proof of Theorem 1.7 that, differently from the previous results, in this theorem we have to assume $0 < T < +\infty$.

Indeed, if $T = +\infty$ (so that $\Omega_T = \mathbb{R}^N \times (0, +\infty)$), the change of variable

$$v \longleftrightarrow e^{\lambda_0 t} v = u$$

(used in *Step I)* to link the solvability of the Cauchy problem to that of the equation $\mathcal{L}v - \lambda_0 v = h$) *does not preserve* the property of belonging to $W^{2,p}_X(\Omega_T)$.

Acknowledgements. The Authors are members of the research group "Gruppo Nazionale per l'Analisi Matematica, la Probabilità e le loro Applicazioni" of the Italian "Istituto Nazionale di Alta Matematica". The first Author is partially supported by the PRIN 2022 project 2022R537CS *NO^3 - Nodal Optimization, NOnlinear elliptic equations, NOnlocal geometric problems, with a focus on regularity*, funded by the European Union - Next Generation EU; the second Author is partially supported by the PRIN 2022 project *Partial differential equations and related geometric-functional inequalities*, financially supported by the EU, in the framework of the "Next Generation EU initiative".

Address. S. Biagi, M. Bramanti:
Dipartimento di Matematica. Politecnico di Milano.
Via Bonardi 9. 20133 Milano. Italy.
e-mail: stefano.biagi@polimi.it; marco.bramanti@polimi.it

References

[1] S. Biagi, M. Bramanti: Schauder estimates for Kolmogorov-Fokker-Planck operators with coefficients measurable in time and Hölder continuous in space. J. Math. Anal. Appl. 533 (2024), no. 1, Paper No. 127996, 65 pp.

[2] S. Biagi, M. Bramanti: Global Sobolev regularity for nonvariational operators built with homogeneous Hörmander vector fields. J. Differential Equations 423 (2025), 708–764.

[3] S. Biagi, M. Bramanti: Schauder estimates on bounded domains for KFP operators with coefficients measurable in time and Hölder continuous in space. Anal. Geom. Metr. Spaces 12 (2024), no. 1, Paper No. 20240009, 16 pp.

[4] S. Biagi, M. Bramanti, B. Stroffolini: KFP operators with coefficients measurable in time and Dini continuous in space. J. Evol. Equ. 24 (2024), no. 2, Paper No. 32, 52 pp.

[5] A. Bonfiglioli, E. Lanconelli, F. Uguzzoni: 'Stratified Lie Groups and Potential Theory for their sub-Laplacians', Springer Monographs in Mathematics **26**, Springer, New York, N.Y., 2007.

[6] M. Bramanti, L. Brandolini: L^p estimates for uniformly hypoelliptic operators with discontinuous coefficients on homogeneous groups. Rend. Sem. Mat. Univ. Politec. Torino 58 (2000), no. 4, 389–433 (2003).

[7] M. Bramanti, L. Brandolini: L^p estimates for nonvariational hypoelliptic operators with VMO coefficients. Trans. Amer. Math. Soc. 352 (2000), no. 2, 781–822.

[8] M. Bramanti, L. Brandolini: Hörmander operators. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2023, xxviii+693 pp.

[9] M. Bramanti, M. C. Cerutti: $W^{1,2,p}$ solvability for the Cauchy-Dirichlet problem for parabolic equations with VMO coefficients. Comm. Partial Differential Equations 18 (1993), no. 9-10, 1735–1763.

[10] M. Bramanti, M. C. Cerutti: Commutators of singular integrals on homogeneous spaces. Boll. Un. Mat. Ital. B (7) 10 (1996), no. 4, 843-883.

[11] Bramanti, M.; Cerutti, M. C.; Manfredini, M.: L^p-estimates for some ultraparabolic operators with discontinuous coefficients. J. Math. Anal. Appl. 200 (1996), no. 2, 332-354.

[12] M. Bramanti, S. Polidoro: Fundamental solutions for Kolmogorov-Fokker-Planck operators with time-depending measurable coefficients. Math. Eng. **2** (2020), no. 4, 734–771.

[13] M. Bramanti, M. Toschi: The sharp maximal function approach to estimates for operators structured on Hörmander vector fields. Rev. Mat. Complut. (2016) 29: 531-557.

[14] 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 (2021), no. 3, 989–1089.

[15] F. Chiarenza, M. Frasca, P. Longo: Interior $W^{2,p}$ estimates for nondivergence elliptic equations with discontinuous coefficients. Ricerche Mat. 40 (1991), no. 1, 149–168.

[16] F. Chiarenza, M. Frasca, P. Longo: $W^{2,p}$-solvability of the Dirichlet problem for nondivergence elliptic equations with VMO coefficients. Trans. Amer. Math. Soc. 336 (1993), no. 2, 841–853.

[17] R. R. Coifman, G. Weiss: Analyse harmonique non-commutative sur certains espaces homogènes. Lecture Notes in Math., Vol. 242. Springer-Verlag, Berlin-New York, 1971, v+160 pp.

[18] M. Di Francesco, S. Polidoro: Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form. Adv. Differential Equations 11 (2006), no. 11, 1261–1320.

[19] H. Dong, T. Yastrzhembskiy: Global L^p estimates for kinetic Kolmogorov-Fokker-Planck equations in nondivergence form. Arch. Ration. Mech. Anal. 245 (2022), no. 1, 501–564.

[20] H. Dong, T. Yastrzhembskiy: Global L^p estimates for kinetic Kolmogorov-Fokker-Planck equations in divergence form. SIAM J. Math. Anal.56(2024), no.1, 1223–1263.

[21] N. V. Krylov: Parabolic and elliptic equations with VMO coefficients. Comm. Partial Differential Equations 32 (2007), no. 1-3, 453–475.

[22] N. V. Krylov: Lectures on elliptic and parabolic equations in Sobolev spaces, Amer. Math. Soc., Providence, RI, 2008.

[23] N. V. Krylov, E. Priola: Elliptic and parabolic second-order PDEs with growing coefficients. Comm. Partial Differential Equations 35 (2010), no. 1, 1–22.

[24] E. Lanconelli, S. Polidoro: On a class of hypoelliptic evolution operators. Partial differential equations, II (Turin, 1993). Rend. Sem. Mat. Univ. Politec. Torino 52 (1994), no. 1, 29–63.

[25] M. Manfredini: The Dirichlet problem for a class of ultraparabolic equations. Adv. Differential Equations 2 (1997), no. 5, 831–866.

[26] S. Menozzi: Martingale problems for some degenerate Kolmogorov equations. Stochastic Process. Appl. 128 (2018), no. 3, 756–802.

[27] A. Nagel, E.M. Stein, S. Wainger: Balls and metrics defined by vector fields. I. Basic properties, Acta Math. **155** (1985), 103–147.

[28] A. Pascucci, A. Pesce: On stochastic Langevin and Fokker-Planck equations: the two-dimensional case. J. Differential Equations 310 (2022), 443–483.

[29] G. Pradolini, O. Salinas: Commutators of singular integrals on spaces of homogeneous type. Czechoslovak Math. J. 57(132) (2007), no. 1, 75–93.