



Optimal regularity for degenerate Kolmogorov equations in non-divergence form with rough-in-time coefficients

STEFANO PAGLIARANI , GIACOMO LUCERTINI AND ANDREA PASCUCCI

Abstract. We consider a class of degenerate equations in non-divergence form satisfying a parabolic Hörmander condition, with coefficients that are measurable in time and Hölder continuous in the space variables. By utilizing a generalized notion of strong solution, we establish the existence of a fundamental solution and its optimal Hölder regularity, as well as Gaussian estimates. These results are key to study the backward Kolmogorov equations associated to a class of Langevin diffusions.

1. Introduction

We study existence and optimal regularity properties of the fundamental solution to a Kolmogorov operator that satisfies a parabolic Hörmander condition. The coefficients of the operator are Hölder continuous in the space variables, but only measurable in time.

Precisely, for fixed $d \leq N$ and $T_0 > 0$, we consider the second order operator in non-divergence form $\mathcal{A} + Y$ with

$$\mathcal{A} = \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t, x) \partial_{x_i} \partial_{x_j} + \sum_{i=1}^d a_i(t, x) \partial_{x_i} + a(t, x), \quad (t, x) \in \mathcal{S}_{T_0} :=]0, T_0[\times \mathbb{R}^N, \quad (1.1)$$

$$Y = \partial_t + \langle Bx, \nabla \rangle = \partial_t + \sum_{i,j=1}^N b_{ij} x_j \partial_{x_i}, \quad x \in \mathbb{R}^N,$$

where B is a constant matrix of dimension $N \times N$. Here, \mathcal{A} is an elliptic operator on \mathbb{R}^d and Y is a first order differential operator on $\mathbb{R} \times \mathbb{R}^N$, also called *transport* or *drift term*. The focus of this paper is mainly on the case $d < N$, that is when $\mathcal{A} + Y$ is fully degenerate, namely no coercivity condition on \mathbb{R}^N is satisfied.

Throughout the paper, $\mathcal{A} + Y$ verifies the following two structural

Assumption 1.1. (*Coercivity on \mathbb{R}^d*) For $1 \leq i, j \leq d$, the coefficients a_{ij}, a_i, a are in $L^\infty([0, T_0]; C_b(\mathbb{R}^N))$, where $C_b(\mathbb{R}^N)$ denotes the space of bounded continuous

Mathematics Subject Classification: 35K65, 35B65, 35A08

Keywords: Kolmogorov equations, Fundamental solution, Hörmander condition, Hölder estimates, Measurable coefficients, Parametrix technique, Anisotropic diffusion.

functions on \mathbb{R}^N . The diffusion matrix $(a_{ij})_{i,j=1,\dots,d}$ is symmetric and there exists a positive constant μ such that

$$\mu^{-1}|\xi|^2 \leq \sum_{i,j=1}^d a_{ij}(t, x)\xi_i\xi_j \leq \mu|\xi|^2, \quad x \in \mathbb{R}^N, \xi \in \mathbb{R}^d,$$

for almost every $t \in [0, T_0]$.

Assumption 1.2. (*Hörmander condition*) The vector fields $\partial_{x_1}, \dots, \partial_{x_d}$ and Y satisfy

$$\text{rank Lie}(\partial_{x_1}, \dots, \partial_{x_d}, Y) = N + 1. \quad (1.2)$$

We refer to (1.2) as a *parabolic Hörmander condition* since the drift term Y plays a key role in the generation of the Lie algebra. Under Assumption 1.2, the prototype Kolmogorov operator

$$\frac{\delta}{2} \sum_{i=1}^d \partial_{x_i x_i} + Y \quad (1.3)$$

is hypoelliptic for any $\delta > 0$. Kolmogorov [19] (see also [15]) constructed the explicit Gaussian fundamental solution for (1.3), which is the transition density of a linear stochastic differential equation.

Motivations for the study of $\mathcal{A} + Y$ come from physics and finance. In its most basic form, with $N = 2$ and $d = 1$,

$$\frac{1}{2} \partial_{x_1 x_1} + x_1 \partial_{x_2} + \partial_t \quad (1.4)$$

is the backward Kolmogorov operator of the system of stochastic equations

$$\begin{cases} dV_t = dW_t \\ dX_t = V_t dt \end{cases} \quad (1.5)$$

where W is a real Brownian motion. In the classical Langevin model, (V, X) describes the velocity and position of a particle in the phase space and it is the prototype of more general kinetic models (cf. [17, 18, 22]). In mathematical finance, (V, X) represents the log-price and average processes used in modeling path-dependent financial derivatives, such as Asian options (cf. [2, 31]).

1.1. Overview of the results: motivations and comparison with the literature

The study of the fundamental solution and its regularity properties is a relevant task in the theory of partial differential equations. In this regard, we point out that the global Schauder estimates recently proved in [25] strongly rely on the results of this paper (see Sect. 4 for a brief anticipation of the results in [25]). The study of the fundamental solution is also crucial in tackling the martingale problem for the

corresponding stochastic differential equations (SDEs) (see [29]). For instance, it is known that uniform bounds on the first and second order derivatives of the solutions to the Cauchy problem can be utilized to prove weak and pathwise well-posedness of the SDE, respectively. We point out that the matter of pathwise uniqueness for the SDE is still partially open in the Hölder setting (see [5]). These issues will be addressed in a future work.

The innovative aspects of our contribution revolve around the regularity assumptions on the coefficients and the regularity properties that we prove for the solutions. For Kolmogorov operators with coefficients that are Hölder continuous (in both space and time) in a classical sense, the study of the existence of a fundamental solution goes back to the early papers [16,28,38,39]. More modern and natural approaches make use of different Hölder spaces, which better adapt to the geometric properties of the operator and of the underlying diffusion processes. Deferring the precise definitions until Sect. 1.2, we may identify two main notions of Hölder space:

- *Anisotropic* these are the Hölder spaces utilized in [26] and later adopted by several authors (see [24,37] among others). They comprise functions $g : \mathbb{R}^N \rightarrow \mathbb{R}$ that are Hölder continuous with respect to an *anisotropic norm* $|\cdot|_B$ related to the scaling properties of the underlying diffusion process (cf. [6,21]). For example, the anisotropic norm for the Langevin operator (1.4) reads as $|(v, x)|_B = |v| + |x|^{\frac{1}{3}}$ for $(v, x) \in \mathbb{R}^2$ and reflects the time-scaling behavior of the stochastic system (1.5), i.e. $(\Delta V)^2 \approx \Delta t$ and $(\Delta X)^2 \approx (\Delta t)^3$.
- *Intrinsic* these are the Hölder spaces introduced in [36] and then studied by several authors (see [30] among others). They are defined for functions $f : \mathbb{R} \times \mathbb{R}^N \rightarrow \mathbb{R}$, by specifying the regularity along the Hörmander vector fields $\partial_{x_1}, \dots, \partial_{x_d}$ and Y . The Hörmander condition then allows to recover joint Hölder regularity in both time and space with respect to an intrinsic distance on $\mathbb{R} \times \mathbb{R}^N$, which reduces to the anisotropic distance when restricted to \mathbb{R}^N , i.e. $|t - s|^{1/2} + |x - e^{(t-s)B}y|_B$ for $(t, x), (s, y) \in \mathbb{R} \times \mathbb{R}^N$ (cf. Remarks 1.1 and 1.2).

In regards to these Hölder spaces, we recognize two main streams of literature that consider anisotropic or intrinsic regularity for both the data and the solutions:

- *Anisotropic-to-anisotropic* in this approach, initiated in [26], the coefficients of the operator are assumed anisotropically Hölder continuous in space and only measurable in time, and the same type of regularity is proved for the solutions. These are typically defined in the distributional sense and do not benefit from the time-smoothing effect that is typical of parabolic equations (see, for instance, Theorem 4.3 in [37]). In particular, no regularity along Y , and thus in the time-variable, is provided.
- *Intrinsic-to-intrinsic* in this approach, initiated by [36] and then developed by several authors (e.g. [1,7,13,32] among others), the data are typically assumed to be intrinsically Hölder continuous in time-space. Solutions are defined in the Lie sense, regarding Y as a directional derivative. In this approach, regularity properties in space and time are strictly intertwined: this allows to fully exploit

the smoothing effect of the equation. However, the regularity assumptions on the data makes the analysis less suitable for applications to stochastic equations.

Only recently (see [3, 10]), the last two approaches have been combined into a third one, which we refer to as *anisotropic-to-intrinsic* and provides intrinsic regularity of the solutions under anisotropic regularity of the coefficients. Our contributions go exactly in this direction. In particular:

- Theorem 1.1 establishes the existence of the fundamental solution, together with Gaussian estimates for the solution and its derivatives. It generalizes the results in [7] by relaxing the regularity hypotheses on the coefficients: as pointed out in [10], intrinsic Hölder continuity is a rather restrictive assumption and simple smooth bounded functions like $f(t, x_1, x_2) = \sin x_2$ fail to satisfy this condition. Theorem 1.1 is also comparable with the results in [6]: however, in the latter reference, it is not proved that the kernel constructed with the parametrix method is a fundamental solution, which would require careful potential estimates; on the other hand, the framework considered in [6] is more general because the drift Y is not necessarily linear in x .
- Theorem 1.2 contains global intrinsic Hölder estimates on the fundamental solution. At the level of the fundamental solution, this result extends some partial results in [8, Theorem 3.2] obtained for Hölder continuous coefficients in the intrinsic sense. Most importantly, the estimates of Theorem 1.2, together with the intermediate estimates their proof rely on, are key to obtain the optimal global Schauder estimates in [25] (cf. Sect. 4 below). The latter are stronger than previous results obtained in the *anisotropic-to-anisotropic* and the *intrinsic-to-intrinsic* settings (e.g. [9, 14, 17, 24, 26, 37] among others) and also refine the estimates of *intrinsic-to-intrinsic* type recently obtained in [3, 10]. We refer to [25, Section 1] for a detailed comparison for the literature on Schauder estimates for degenerate Kolmogorov equations.

Finally, the anisotropic Hölder regularity of solutions to $\mathcal{A} + Y$ unveils the full smoothing effect of the the differential operator in both the space and time variables. A notable implication is the possibility to prove intrinsic Taylor formulas like those proved in [30] (cf. Remark 1.2), on which stochastic tools like the Itô formula rely (see [21]). Furthermore, the setting of coefficients that are rough in the time variable naturally arises in the study of stochastic partial differential equations (see, for example, [20, 35]).

1.2. Anisotropic and intrinsic Hölder spaces

We start by noticing that condition (1.2) is equivalent to the well-known Kalman rank condition for controllability in linear systems theory (cf., for instance, [33]). Also, it was shown in [23] that (1.2) is equivalent to B having the block-form

$$B = \begin{pmatrix} * & * & \cdots & * & * \\ B_1 & * & \cdots & * & * \\ 0 & B_2 & \cdots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_q & * \end{pmatrix} \tag{1.6}$$

where the $*$ -blocks are arbitrary and B_j is a $(d_{j-1} \times d_j)$ -matrix of rank d_j with

$$d \equiv d_0 \geq d_1 \geq \cdots \geq d_q \geq 1, \quad \sum_{i=0}^q d_i = N.$$

This allows to introduce natural definitions of anisotropic norm and Hölder continuity on \mathbb{R}^N .

Definition 1.1. (*Anisotropic norm and anisotropic Hölder spaces*) For any $x \in \mathbb{R}^N$ let

$$|x|_B := \sum_{j=0}^q \sum_{i=\bar{d}_{j-1}+1}^{\bar{d}_j} |x_i|^{\frac{1}{2j+1}}, \quad \bar{d}_j := \sum_{k=0}^j d_k. \tag{1.7}$$

For $\alpha \in]0, 1]$ we denote by $C_B^\alpha(\mathbb{R}^N)$ the set of functions $g \in C_b(\mathbb{R}^N)$ such that

$$\|g\|_{C_B^\alpha(\mathbb{R}^N)} := \sup_{x \in \mathbb{R}^N} |g(x)| + \sup_{x, y \in \mathbb{R}^N} \frac{|g(x) - g(y)|}{|x - y|_B^\alpha} < \infty,$$

and by $L^\infty([0, T_0]; C_B^\alpha(\mathbb{R}^N))$ the set of measurable functions $f : [0, T_0] \rightarrow C_B^\alpha(\mathbb{R}^N)$ such that

$$\|f\|_{L^\infty([0, T_0]; C_B^\alpha(\mathbb{R}^N))} := \sup_{t \in [0, T_0]} \|f(t)\|_{C_B^\alpha(\mathbb{R}^N)} < \infty.$$

To give the precise definition of the Hölder space $C_B^{2,\alpha}$ we first introduce the intrinsic Hölder regularity along the vector fields appearing in the Hörmander condition (1.2). As it is standard in the framework of functional analysis on homogeneous groups (cf. [11]), the idea is to weight the Hölder exponent in terms of the formal degree of the vector fields, which is equal to 1 for $\partial_{x_1}, \dots, \partial_{x_d}$ and equal to 2 for Y .

Definition 1.2. Let $\alpha \in]0, 1]$, $\beta \in]0, 2]$ and $T > 0$. We denote respectively by $C_d^\alpha(\mathcal{S}_T)$ and $C_Y^\beta(\mathcal{S}_T)$ the set of the functions $f : \mathcal{S}_T \rightarrow \mathbb{R}$ such that the following semi-norms are finite

$$\begin{aligned} \|f\|_{C_d^\alpha(\mathcal{S}_T)} &:= \sum_{i=1}^d \sup_{\substack{(t,x) \in \mathcal{S}_T \\ h \in \mathbb{R}}} \frac{|f(t, x + h\mathbf{e}_i) - f(t, x)|}{|h|^\alpha}, \\ \|f\|_{C_Y^\beta(\mathcal{S}_T)} &:= \sup_{\substack{t,s \in [0,T] \\ x \in \mathbb{R}^N}} \frac{|f(s, e^{(s-t)B}x) - f(t, x)|}{|t - s|^{\frac{\beta}{2}}}. \end{aligned}$$

Here e_i denotes the i -th element of the canonical basis of \mathbb{R}^N .

Moreover, we say that f is a.e. Lie differentiable along Y on \mathcal{S}_T if there exists $F \in L^1_{loc}]0, T[; C_b(\mathbb{R}^N)$ such that

$$f(s, e^{(s-t)B}x) = f(t, x) + \int_t^s F(r, e^{(r-t)B}x)dr, \quad (t, x) \in \mathcal{S}_T, s \in]0, T[. \tag{1.8}$$

In that case, we set $Yf = F$ and call it an a.e. Lie derivative of f on \mathcal{S}_T .

Next we recall the intrinsic Hölder spaces of order 0 and 1 introduced in [30,31].

Definition 1.3. (Intrinsic Hölder spaces: orders 0 and 1) For $\alpha \in]0, 1]$, $C_B^{0,\alpha}(\mathcal{S}_T)$ and $C_B^{1,\alpha}(\mathcal{S}_T)$ denote, respectively, the set of the functions $f : \mathcal{S}_T \rightarrow \mathbb{R}$ such that the following semi-norms are finite

$$\begin{aligned} \|f\|_{C_B^{0,\alpha}(\mathcal{S}_T)} &:= \|f\|_{C_Y^\alpha(\mathcal{S}_T)} + \|f\|_{C_d^\alpha(\mathcal{S}_T)}, \\ \|f\|_{C_B^{1,\alpha}(\mathcal{S}_T)} &:= \|f\|_{C_Y^{1+\alpha}(\mathcal{S}_T)} + \sum_{i=1}^d \|\partial_{x_i} f\|_{C_B^{0,\alpha}(\mathcal{S}_T)} \\ &= \|f\|_{C_Y^{1+\alpha}(\mathcal{S}_T)} + \sum_{i=1}^d (\|\partial_{x_i} f\|_{C_Y^\alpha(\mathcal{S}_T)} + \|\partial_{x_i} f\|_{C_d^\alpha(\mathcal{S}_T)}). \end{aligned}$$

Remark 1.1. It is standard to show (see [30]) that, if $f \in C_B^{0,\alpha}(\mathcal{S}_T)$, then f is α -Hölder continuous w.r.t. the intrinsic distance, namely

$$|f(t, x) - f(t, y)| \leq C(|t - s|^{\frac{\alpha}{2}} + |x - e^{(t-s)B}y|_B^\alpha), \quad (t, x)(s, t) \in \mathcal{S}_T,$$

where C is a positive constant that depends only on the matrix B and T . In particular, f enjoys anisotropic Hölder regularity, and thus also regularity in the degenerate variables x_i for $i > d$, namely

$$\sup_{x, y \in \mathbb{R}^N} \frac{|f(t, x) - f(t, y)|}{|x - y|_B^\alpha} \leq C \|f\|_{C_B^{0,\alpha}(\mathcal{S}_T)}, \quad t \in [0, T].$$

Remark 1.2. In [30] it was also shown that if $f \in C_B^{1,\alpha}(\mathcal{S}_T)$ then the following intrinsic Taylor formula holds:

$$\begin{aligned} &\left| f(s, y) - f(t, x) - \sum_{i=1}^d \partial_{x_i} f(t, x)(y - e^{(s-t)B}x)_i \right| \\ &\leq C(|s - t|^{\frac{1}{2}} + |y - e^{(s-t)B}x|_B)^{1+\alpha}, \quad (t, x), (s, y) \in \mathcal{S}_T. \end{aligned}$$

where C is a positive constant that depends only on the matrix B . We stress that the Taylor ‘‘polynomial’’ above only contains the first derivatives of f w.r.t. the first d components of x .

Table 1. Functional spaces

Notation	Functional space	Reference
C_B^α	Anisotropic Hölder spaces on \mathbb{R}^N	Def. 1.1
C_d^α, C_Y^α	Lie Hölder spaces on \mathcal{S}_T	Def. 1.2
$C_B^{k,\alpha}, k = 0, 1, 2$	Intrinsic Hölder spaces on \mathcal{S}_T	Def. 1.3, 1.4

To cope with the lack of regularity of the coefficients in the time-direction, the definition of $C_B^{2,\alpha}(\mathcal{S}_T)$ differs from the one given in [30], specifically with regards to the regularity along Y .

Definition 1.4. (*Intrinsic Hölder spaces: order 2*) For $\alpha \in]0, 1]$, $C_B^{2,\alpha}(\mathcal{S}_T)$ is the set of functions $f : \mathcal{S}_T \rightarrow \mathbb{R}$ such that following semi-norm is finite:

$$\begin{aligned} \|f\|_{C_B^{2,\alpha}(\mathcal{S}_T)} &= \sum_{i=1}^d \|\partial_{x_i} f\|_{C_B^{1,\alpha}(\mathcal{S}_T)} + \|Yf\|_{L^\infty([0,T];C_B^\alpha(\mathbb{R}^N))} \\ &= \sum_{i=1}^d \|\partial_{x_i} f\|_{C_Y^{1+\alpha}(\mathcal{S}_T)} + \sum_{i,j=1}^d (\|\partial_{x_i x_j} f\|_{C_Y^\alpha(\mathcal{S}_T)} + \|\partial_{x_i x_j} f\|_{C_d^\alpha(\mathcal{S}_T)}) \\ &\quad + \|Yf\|_{L^\infty([0,T];C_B^\alpha(\mathbb{R}^N))}. \end{aligned} \tag{1.9}$$

Here Yf is the a.e. Lie derivative as in Definition 1.2.

Remark 1.3. Definition 1.4 is similar in spirit to that proposed in [9,27,30] for the study of Kolmogorov operators with Hölder coefficients: according to their definition, if f belongs to $C^{2,\alpha}$ then Yf belongs to $C_B^{0,\alpha}$.

This is the regularity the fundamental solution enjoys in case the coefficients of \mathcal{A} are Hölder continuous in both space and time, namely if they belong to $C_B^{0,\alpha}$. By contrast, if $f \in C_B^{2,\alpha}$ in the sense of Definition 1.4 then f is generally at most Lipschitz continuous along Y : this is the optimal result one can prove without assuming further regularity of the coefficients in the time variable other than measurability, as it was shown in [4] in the case of coefficients only dependent on time.

For reader’s convenience, we recall that we shall always denote by \mathcal{S}_T the strip $]0, T[\times \mathbb{R}^N$; also, in the following table we collect the notations used for the main functional spaces (Table 1):

1.3. Main results

As previously explained, major questions in the study of Kolmogorov equations are the very definition of solution and its optimal regularity properties. It is well-known that, in general, the fundamental solution is not regular enough to support the derivatives ∂_{x_i} , for $d < i \leq N$, appearing in the transport term Y . Indeed, under

the Hörmander condition (1.2), these derivatives are of order three and higher in the intrinsic sense. For this reason, even for equations with Hölder coefficients, weak notions of solution have been introduced. We employ a notion of solution, recently proposed in [35] with the aim of studying Langevin *stochastic PDEs* with rough coefficients. In accordance with the intrinsic Hölder space $C_B^{2,\alpha}(\mathcal{S}_T)$ in Definition 1.4, this solution supports the existence of an a.e. Lie derivative and is the strongest possible under our regularity assumptions on the coefficients.

Definition 1.5. (*Strong Lie solution*) Let $0 < T \leq T_0$ and $f \in L^1_{\text{loc}}([0, T[; C_b(\mathbb{R}^N)))$. A solution to equation

$$\mathcal{A}u + Yu = f \text{ on } \mathcal{S}_T \tag{1.10}$$

is a continuous function u such that there exist $\partial_{x_i}u, \partial_{x_i x_j}u \in L^1_{\text{loc}}([0, T[; C_b(\mathbb{R}^N)))$, for $i, j = 1, \dots, d$, and $f - \mathcal{A}u$ is an a.e. Lie derivative of u in the sense of Definition 1.2, i.e.

$$u(s, e^{(s-t)B}x) = u(t, x) - \int_t^s \left(\mathcal{A}u(\tau, e^{(\tau-t)B}x) - f(\tau, e^{(\tau-t)B}x) \right) d\tau, \tag{1.11}$$

$$(t, x) \in \mathcal{S}_T, s < T.$$

Remark 1.4. Notice that $s \mapsto (s, e^{(s-t)B}x)$ is the integral curve of Y starting from (t, x) : for any suitably regular function u the limit

$$Yu(t, x) := \lim_{s \rightarrow t} \frac{u(s, e^{(s-t)B}x) - u(t, x)}{s - t} \tag{1.12}$$

is the directional (or Lie) derivative along Y of u at (t, x) . Thus, if the integrand in (1.11) is continuous then u is a classical (pointwise) solution of (1.10). However, as noticed in Remark 2.6, in general a solution u in the sense of Definition 1.5 is only a.e. differentiable along Y and Eq. (1.10) is satisfied for almost every $(t, x) \in \mathcal{S}_T$.

In order to state our first main result, we give the following

Definition 1.6. (*Fundamental solution*) A fundamental solution of $\mathcal{A} + Y$ is a function $p = p(t, x; T, y)$ defined for $t < T$ and $x, y \in \mathbb{R}^N$ such that, for any fixed $(T, y) \in \mathcal{S}_{T_0}$, we have:

- (i) $p(\cdot, \cdot; T, y)$ is a solution of equation $\mathcal{A}u + Yu = 0$ on \mathcal{S}_T in the sense of Definition 1.5;
- (ii) For any $g \in C_b(\mathbb{R}^N)$ we have

$$\lim_{\substack{(t,x) \rightarrow (T,y) \\ t < T}} \int_{\mathbb{R}^N} p(t, x; T, \eta) g(\eta) d\eta = g(y).$$

Before stating our main results, we introduce our last standing assumption, which is concerned with the regularity of the coefficients of \mathcal{A} .

Assumption 1.3. The coefficients a_{ij}, a_i, a of \mathcal{A} belong to $L^\infty([0, T_0]; C_B^\alpha(\mathbb{R}^N))$ for some $\alpha \in]0, 1]$.

Remark 1.5. The boundedness requirement in Assumption 1.3 is a crucial ingredient of the parametrix method: it is used to get the preliminary Gaussian estimates (cf. Proposition A.1) on which the entire technique relies. In particular, the key property is boundedness in space for any fixed time. On the other hand, a detailed examination of the proof reveals the possibility for an extension that includes coefficients that are only locally integrable, as opposed to bounded, in the time variable. We aim to investigate this extension in a future work.

The following result states the existence of the fundamental solution p of $\mathcal{A} + Y$, as well as uniform Gaussian bounds for p and its derivatives with respect to the non-degenerate variables x_1, \dots, x_d .

Theorem 1.1. (Existence and Gaussian bounds) *Under Assumptions 1.1, 1.2 and 1.3, $\mathcal{A} + Y$ has a fundamental solution $p = p(t, x; T, y)$ in the sense of Definition 1.6. For every $\varepsilon > 0$ there exists a positive constant C , only dependent on $T_0, \mu, B, \varepsilon, \alpha$ and the α -Hölder norms of the coefficients, such that*

$$p(t, x; T, y) \leq C \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{1.13}$$

$$|\partial_{x_i} p(t, x; T, y)| \leq \frac{C}{\sqrt{T-t}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{1.14}$$

$$|\partial_{x_i x_j} p(t, x; T, y)| \leq \frac{C}{T-t} \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{1.15}$$

for any $(T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T$ and $i, j = 1, \dots, d$, where Γ^δ is the Gaussian fundamental solution of (1.3), whose explicitly expression is given in (2.4). Moreover, there exist two positive constants $\bar{\mu}, \bar{c}$ such that

$$\bar{c} \Gamma^{\bar{\mu}}(t, x; T, y) \leq p(t, x; T, y), \tag{1.16}$$

for any $(T, y) \in \mathcal{S}_{T_0}$ and $(t, x) \in \mathcal{S}_T$.

In Sect. 4 we present several results for the Cauchy problem that are straightforward consequences of Theorem 1.1. The proof of Theorem 1.1 is based on a modification of Levi’s parametrix technique, which allows to deal with the lack of regularity of the coefficients along the drift term Y . The main tool is the fundamental solution of a Kolmogorov operator with time-dependent measurable coefficients, also recently studied in [4]. This approach allows for a careful analysis of the optimal regularity properties of the fundamental solution p : Theorem 1.2 below states that p belongs to the intrinsic Hölder space $C_B^{2,\alpha}$ as given by Definition 1.4. As the notation could be misleading, we explicitly remark that for $u \in C_B^{2,\alpha}$ not even the first order derivatives $\partial_{x_i} u$, for $i > d$, necessarily exist. However, in general we cannot expect higher regularity properties for solutions to (1.10) and $C_B^{2,\alpha}$ -regularity is indeed optimal. The following theorem refines the known results about the smoothness of the fundamental solution in [26], [27] and [9] and exhibits its maximal regularity properties.

Theorem 1.2. (Regularity of the fundamental solution) *Under the assumptions of Theorem 1.1, $p(\cdot, \cdot; T, y) \in C_B^{2,\beta}(\mathcal{S}_\tau)$ for every $(T, y) \in \mathcal{S}_{T_0}$, $0 < \tau < T$ and $\beta < \alpha$. Precisely, there exists a positive constant C only dependent on $T_0, \mu, B, \beta, \alpha$ and the α -Hölder norms of the coefficients, such that*

$$\|p(\cdot, \cdot; T, y)\|_{C_B^{2,\beta}(\mathcal{S}_\tau)} \leq \frac{C}{(T - \tau)^{\frac{Q+2+\beta}{2}}},$$

where Q is the so-called homogeneous dimension of \mathbb{R}^N with respect to the quasi-norm $|\cdot|_B$ defined by

$$Q = \sum_{i=0}^q (2i + 1)d_i. \quad (1.17)$$

Remark 1.6. As previously explained, the Hölder space $C^{2+\alpha}$ introduced in [26] (and adopted in [24, 37] to prove Schauder estimates), consists of functions f that, together with their second order derivatives $\partial_{x_i x_j} f$ in the non-degenerate directions $i, j = 1, \dots, d$, are Hölder continuous w.r.t. the anisotropic norm (1.7). This notion is weaker than Definition 1.4 both in terms of the regularity of $\partial_{x_i} f$ and, more importantly, in terms of the Lipschitz continuity of f along Y (cf. (1.8)) which reveals the regularizing effect of the associated evolution semigroup.

The rest of the paper is structured as follows. Section 2 contains the construction of the fundamental solution by means of the parametrix method: in particular, Sect. 2.2 includes the proof of Theorem 1.1. In Sect. 3 we prove the regularity estimates of the fundamental solution, in particular Theorem 1.2. In Sect. 4, we anticipate some results for the Cauchy problem for $\mathcal{A} + Y$, namely well-posedness and Schauder estimates, which can be derived by utilizing our results on the fundamental solution. The appendices contain the Gaussian and potential estimates that are employed in the proofs.

2. Parametrix construction

The parametrix method, introduced by E. Levi in 1907, is a constructive technique to prove the existence of the fundamental solution to an elliptic operator, later extended by several authors to numerous parabolic settings. The monograph [12] is a classical reference for the application of the method to uniformly parabolic operators; the case of degenerate operators of the form (1.1) with Hölder continuous coefficients was studied in [7, 36]. We refer to [4] for a derivation of the fundamental solution via classical Fourier methods, in the case of time-dependent measurable coefficients.

Let Assumptions 1.1, 1.2 and 1.3 be satisfied. The first step of the parametrix method is to set a kernel $\mathbf{P} = \mathbf{P}(t, x; T, y)$ that serves as proxy for the fundamental solution,

called *parametrix*. We denote by $\mathcal{A}^{(s,v)}$ the operator obtained by freezing the second-order coefficients of \mathcal{A} along the integral curve of the vector field Y passing through $(s, v) \in \mathcal{S}_{T_0}$ and neglecting the lower order terms. Namely we consider the operator

$$\mathcal{A}^{(s,v)} := \frac{1}{2} \sum_{i,j=1}^d a_{ij}(t, e^{(t-s)B}v) \partial_{x_i x_j}, \quad (t, x) \in \mathcal{S}_{T_0}. \tag{2.1}$$

One can directly prove that the fundamental solution of

$$\mathcal{A}^{(s,v)} + Y,$$

in the sense of Definition 1.6, is given by

$$\Gamma^{(s,v)}(t, x; T, y) = \mathbf{G}(\mathcal{C}^{(s,v)}(t, T), y - e^{(T-t)B}x), \quad (T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T,$$

where

$$\mathbf{G}(\mathcal{C}, z) := \frac{1}{\sqrt{(2\pi)^N \det \mathcal{C}}} e^{-\frac{1}{2} \langle \mathcal{C}^{-1}z, z \rangle}$$

is the Gaussian kernel on \mathbb{R}^N and

$$\mathcal{C}^{(s,v)}(t, T) := \int_t^T e^{(T-\tau)B} A^{(s,v)}(\tau) e^{(T-\tau)B^*} d\tau, \tag{2.2}$$

$$A^{(s,v)}(\tau) := \begin{pmatrix} A_0(\tau, e^{(\tau-s)B}v) & 0 \\ 0 & 0 \end{pmatrix}, \quad A_0 = (a_{ij})_{i,j=1,\dots,d}. \tag{2.3}$$

Remark 2.1. Clearly $\Gamma^{(s,v)}(t, x; T, y)$ is of class C^∞ as a function of x and only absolutely continuous along the integral curves of Y as a function of (t, x) .

Remark 2.2. In the particular case of $A_0 \equiv \delta I_d$ for some $\delta > 0$, where I_d is the $(d \times d)$ -identity matrix, the Kolmogorov operator $\mathcal{A}^{(s,v)} + Y$ reads as in (1.3) and its fundamental solution reduces to

$$\Gamma^\delta(t, x; T, y) := \mathbf{G}(\delta \mathcal{C}(T-t), y - e^{(T-t)B}x), \tag{2.4}$$

with

$$\mathcal{C}(t) = \int_0^t e^{(t-\tau)B} \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} e^{(t-\tau)B^*} d\tau. \tag{2.5}$$

Proceeding as in [29, 35], we define the parametrix function $\mathbf{P}(t, x; T, y)$ as

$$\mathbf{P}(t, x; T, y) := \Gamma^{(T,y)}(t, x; T, y), \quad (T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T, \tag{2.6}$$

and we refer to it as to the *time-dependent parametrix* in order to emphasize the fact that it is obtained by freezing only the space variable of the coefficients of \mathcal{A} .

Remark 2.3. Since $\Gamma^{(s,v)}$ is the fundamental solution of $\mathcal{A}^{(s,v)} + Y$, we have

$$(\mathcal{A}^{(T,y)} + Y)\mathbf{P}(\cdot, \cdot; T, y) = 0 \quad \text{on } \mathcal{S}_T, \tag{2.7}$$

in the sense of Definition 1.5, for any $(T, y) \in \mathcal{S}_{T_0}$.

Remark 2.4. In [7], where the variable coefficients of \mathcal{A} are assumed intrinsically Hölder continuous in space and time, the parametrix is defined as the fundamental solution of the operator obtained by freezing the second order coefficients of \mathcal{A} in both time and space variables, i.e.

$$\frac{1}{2} \sum_{i,j=1}^d a_{ij}(s, v) \partial_{x_i x_j} + Y.$$

As we shall see below, the choice of freezing the coefficients only in the space variable, along the integral curve of Y as in (2.1), is necessary in order to deal with the lack of regularity in the time variable.

Once the parametrix function is defined, the parametrix construction prescribes that a fundamental solution of $\mathcal{A} + Y$ is sought in the form

$$p(t, x; T, y) = \mathbf{P}(t, x; T, y) + \int_t^T \int_{\mathbb{R}^N} \mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau, \tag{2.8}$$

where ϕ is an unknown function. We now perform some heuristic computations that will lead to a fixed-point equation for ϕ . Assuming that $p(t, x; T, y)$ in (2.8) is a fundamental solution of $\mathcal{A} + Y$, we obtain

$$\begin{aligned} 0 &= (\mathcal{A} + Y)p(t, x; T, y) = (\mathcal{A} + Y)\mathbf{P}(t, x; T, y) + (\mathcal{A} + Y) \\ &\quad \times \int_t^T \int_{\mathbb{R}^N} \mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau. \end{aligned}$$

Furthermore, by formally differentiating and employing $p(t, x; t, \cdot) = \delta_x$ we also have

$$\begin{aligned} &(\mathcal{A} + Y) \int_t^T \int_{\mathbb{R}^N} \mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau \\ &= \int_t^T \int_{\mathbb{R}^N} (\mathcal{A} + Y)\mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau - \phi(t, x; T, y). \end{aligned}$$

Therefore, $\phi(t, x; T, y)$ must solve the Volterra integral equation

$$\phi(t, x; T, y) = (\mathcal{A} + Y)\mathbf{P}(t, x; T, y) + \int_t^T \int_{\mathbb{R}^N} (\mathcal{A} + Y)\mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau. \tag{2.9}$$

Now, owing to Remark 2.3, Eq. (2.9) can be written as

$$\begin{aligned} \phi(t, x; T, y) &= (\mathcal{A} - \mathcal{A}^{(T,y)})\mathbf{P}(t, x; T, y) + \int_t^T \int_{\mathbb{R}^N} (\mathcal{A} - \mathcal{A}^{(\tau,\eta)}) \\ &\quad \mathbf{P}(t, x; \tau, \eta)\phi(\tau, \eta; T, y)d\eta d\tau, \end{aligned} \tag{2.10}$$

whose solution can be determined by an iterative procedure, which leads to the series representation

$$\phi(t, x; T, y) = \sum_{k \geq 1} \phi_k(t, x; T, y) \tag{2.11}$$

where

$$\begin{cases} \phi_1(t, x; T, y) := (\mathcal{A} - \mathcal{A}^{(T,y)})\mathbf{P}(t, x; T, y), \\ \phi_{k+1}(t, x; T, y) := \int_t^T \int_{\mathbb{R}^N} (\mathcal{A} - \mathcal{A}^{(\tau,\eta)})\mathbf{P}(t, x; \tau, \eta)\phi_k(\tau, \eta; T, y)d\eta d\tau, \quad k \in \mathbb{N}. \end{cases} \tag{2.12}$$

In order to make the previous arguments rigorous one has to prove that:

- The series defined by (2.11)–(2.12) is uniformly convergent on \mathcal{S}_T . At this stage one also obtains a uniform upper bound and a Hölder estimate for ϕ ;
- p defined by (2.8) is actually a fundamental solution of $\mathcal{A} + Y$. In this step one also establishes the Gaussian estimates on p and its derivatives that appear in Theorem 1.1.

2.1. Convergence of the series and estimates on ϕ

Proposition 2.1. *For every $(T, y) \in \mathcal{S}_{T_0}$ the series in (2.11) converges uniformly in $(t, x) \in \mathcal{S}_T$ and the function $\phi = \phi(t, x; T, y)$ solves the integral Eq. (2.10) on \mathcal{S}_T . Furthermore, for every $\varepsilon > 0$ and $0 < \delta < \alpha$, there exists a positive constant C , only dependent on $T_0, \mu, B, \delta, \alpha, \varepsilon$ and the α -Hölder norms of the coefficients, such that*

$$|\phi(t, x; T, y)| \leq \frac{C}{(T-t)^{1-\frac{\alpha}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{2.13}$$

$$|\phi(t, x; T, y) - \phi(t, v; T, y)| \leq \frac{C|x-v|_B^{\alpha-\delta}}{(T-t)^{1-\frac{\delta}{2}}} (\Gamma^{\mu+\varepsilon}(t, x; T, y) + \Gamma^{\mu+\varepsilon}(t, v; T, y)), \tag{2.14}$$

for any $(T, y) \in \mathcal{S}_{T_0}$ and $(t, x), (t, v) \in \mathcal{S}_T$.

To avoid repeating the arguments already used in [7], we limit ourselves to highlighting the parts of the proof that differ significantly from the classical case.

Proof. We first prove that there exists a positive κ such that

$$\begin{aligned} |(\mathcal{A} - \mathcal{A}^{(T,y)})\mathbf{P}(t, x; T, y)| &\leq \frac{\kappa}{(T-t)^{1-\alpha/2}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \\ &(T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T. \end{aligned} \tag{2.15}$$

Assume for simplicity that $a_i, a \equiv 0$, the general case being a straightforward extension. By definition (2.1) we have

$$\begin{aligned}
 |(\mathcal{A} - \mathcal{A}^{(T,y)})\mathbf{P}(t, x; T, y)| &\leq \frac{1}{2} \sum_{i,j=1}^d |a_{ij}(t, x) - a_{ij}(t, e^{-(T-t)B} y)| \\
 &\quad \times |\partial_{x_i x_j} \mathbf{P}(t, x; T, y)| \leq
 \end{aligned}
 \tag{2.16}$$

(by the Hölder regularity of a_{ij} and the Gaussian estimate (A.11))

$$\leq \kappa \frac{|x - e^{-(T-t)B} y|_B^\alpha}{T - t} \Gamma^{\mu+\varepsilon/2}(t, x; T, y).
 \tag{2.17}$$

The estimate (A.9) then yields (2.15).

For any $(T, y) \in \mathcal{S}_{T_0}$ and $(t, x) \in \mathcal{S}_T$, (2.12) and (2.15) imply

$$|\phi_1(t, x; T, y)| \leq \frac{\kappa}{(T - t)^{1-\alpha/2}} \Gamma^{\mu+\varepsilon}(t, x; T, y)$$

and

$$\begin{aligned}
 |\phi_2(t, x; T, y)| &\leq \int_t^T \int_{\mathbb{R}^N} \left| (\mathcal{A} - \mathcal{A}^{(\tau,\eta)}) \mathbf{P}(t, x; \tau, \eta) \right| \times |\phi_1(\tau, \eta; T, y)| d\eta d\tau \\
 &\leq \kappa^2 \int_t^T \frac{1}{(\tau - t)^{1-\alpha/2}} \frac{1}{(T - \tau)^{1-\alpha/2}} \\
 &\quad \times \int_{\mathbb{R}^N} \Gamma^{\mu+\varepsilon}(t, x; \tau, \eta) \Gamma^{\mu+\varepsilon}(\tau, \eta; T, y) d\eta d\tau =
 \end{aligned}$$

(by the Chapman–Kolmogorov identity and solving the integral in $d\tau$)

$$= \kappa^2 \frac{\Gamma_{\text{Euler}}^2\left(\frac{\alpha}{2}\right)}{(T - t)^{1-\alpha} \Gamma_{\text{Euler}}(\alpha)} \Gamma^{\mu+\varepsilon}(t, x; T, y).$$

Proceeding by induction, it is straightforward to verify that

$$|\phi_n(t, x; T, y)| \leq \kappa^n \frac{\Gamma_{\text{Euler}}^n\left(\frac{\alpha}{2}\right)}{(T - t)^{1-\frac{\alpha}{2}} \Gamma_{\text{Euler}}\left(\frac{\alpha}{2}\right)} \Gamma^{\mu+\varepsilon}(t, x; T, y), \quad n \in \mathbb{N}.$$

This proves the uniform convergence of the series on \mathcal{S}_T , which in turn implies that ϕ satisfies (2.10), as well as the estimate (2.13).

The proof of (2.14) is a technical modification of the arguments in [7, Lemma 6.1], which is necessary to account for the different parametrix function. For sake of brevity, we leave the details to the reader. □

Remark 2.5. The proof above is particularly informative to understand the choice of the parametrix function in relation to the lack of regularity of the coefficients with respect to the time variable. In particular, in passing from (2.17) to (2.15), we take advantage of the increment $|x - e^{-(T-t)B} y|_B^\alpha$ in order to recover the integrability of

the singularity in time. In the classical case, namely when the coefficient a_{ij} is also Hölder continuous in time, the parametrix function is obtained by freezing the variable coefficients in both space and time (see Remark 2.4). In (2.16), this choice leads to increments of type

$$|a_{ij}(t, x) - a_{ij}(T, y)|,$$

which is clearly not helpful if a_{ij} does not exhibit any regularity in time.

Furthermore, note that the coefficients have to be frozen in the space variable along the integral curve of Y : freezing the coefficients at a fixed point y would yield an increment of type $|x - y|_B^\alpha$ in (2.17), which does not allow to employ the Gaussian estimates in (A.9) to control the singularity.

2.2. Proof that p is a fundamental solution and Gaussian bounds

We now prove the first part of Theorem 1.1, concerning the existence of the fundamental solution of $\mathcal{A} + Y$. This is achieved by proving that the candidate solution $p = p(t, x; T, y)$ defined through (2.8) satisfies points i) and ii) of Definition 1.6. The innovative part of the proof consists in showing point i), which is $p(\cdot, \cdot; T, y)$ solves the equation

$$\mathcal{A}u + Yu = 0 \quad \text{on } \mathcal{S}_T \tag{2.18}$$

in the integral sense of Definition 1.5. Once more, we provide the details of the parts that significantly depart from the classical case.

For any $(T, y) \in \mathcal{S}_{T_0}$, let us rewrite $p(t, x; T, y)$ as

$$p(t, x; T, y) = \mathbf{P}(t, x; T, y) + \Phi(t, x; T, y), \quad (t, x) \in \mathcal{S}_T,$$

where we set

$$\Phi(t, x; T, y) := \int_t^T \int_{\mathbb{R}^N} \mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau. \tag{2.19}$$

The strategy of the proof is to first show that p possesses the regularity required in order to qualify as a fundamental solution, and then to check that it actually solves Eq. (2.18). As pointed out in Remark 2.7, the parametrix $\mathbf{P} = \mathbf{P}(t, x; T, y)$ is an integral solution to (2.7). In particular, it is smooth in the variable x and absolutely continuous along Y . As for $\Phi = \Phi(t, x; T, y)$, the next result shows that it is twice differentiable w.r.t. x_1, \dots, x_d and states some Gaussian bounds on the derivatives.

Proposition 2.2. *For any $(T, y) \in \mathcal{S}_{T_0}$, $(t, x) \in \mathcal{S}_T$ and $i, j = 1, \dots, d$, there exist*

$$\begin{aligned} \partial_{x_i} \Phi(t, x; T, y) &= \int_t^T \int_{\mathbb{R}^N} \partial_{x_i} \mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau, \\ \partial_{x_i x_j} \Phi(t, x; T, y) &= \int_t^T \int_{\mathbb{R}^N} \partial_{x_i x_j} \mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau, \end{aligned}$$

and, for any $\varepsilon > 0$ we have

$$\begin{aligned} |\Phi(t, x; T, y)| &\leq C(T - t)^{\frac{\alpha}{2}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \\ |\partial_{x_i} \Phi(t, x; T, y)| &\leq \frac{C}{(T - t)^{\frac{1-\alpha}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \\ |\partial_{x_i x_j} \Phi(t, x; T, y)| &\leq \frac{C}{(T - t)^{\frac{2-\alpha}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \end{aligned}$$

where C denotes a positive constant, only dependent on $T_0, \mu, B, \alpha, \varepsilon$ and the α -Hölder norms of the coefficients.

Proof. By the definition of Φ in (2.19) we have

$$\Phi(t, x; T, y) = \int_t^T J(t, x; \tau; T, y) d\tau,$$

with J defined as in (B.1). The potential estimates of Proposition B.1 upon integrating in τ , yield the result. \square

The following result shows that $\Phi(\cdot, \cdot; T, y)$ is also Lipschitz continuous along the integral curves of Y .

Lemma 2.1. *For every $(T, y) \in \mathcal{S}_{T_0}$ and $(t, x) \in \mathcal{S}_T$, we have*

$$\Phi(s, e^{(s-t)B}x; T, y) - \Phi(t, x; T, y) = - \int_t^s F(\tau, x; T, y) d\tau, \quad s \in [t, T[,$$

where

$$\begin{aligned} F(\tau, x; T, y) &:= \int_{\mathbb{R}^N} \int_{\mathbb{R}^N} \mathcal{A}^{(r, \eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta dr \\ &\quad + \phi(\tau, e^{(\tau-t)B}x; T, y). \end{aligned} \tag{2.20}$$

Proof. For any $s \in [t, T[$ one can write

$$\begin{aligned} &\Phi(s, e^{(s-t)B}x; T, y) - \Phi(t, x; T, y) \\ &= \underbrace{\int_s^T \int_{\mathbb{R}^N} \left(\mathbf{P}(s, e^{(s-t)B}x; r, \eta) - \mathbf{P}(t, x; r, \eta) \right) \phi(r, \eta; T, y) d\eta dr}_{=: G(t, x)} \\ &\quad - \underbrace{\int_t^s \int_{\mathbb{R}^N} \mathbf{P}(t, x; r, \eta) \phi(r, \eta; T, y) d\eta dr}_{=: H(t, x)}. \end{aligned}$$

First, we study the term $G(t, x)$. Remark 2.3 yields

$$G(t, x) = \int_s^T \int_{\mathbb{R}^N} \left(\int_t^s -\mathcal{A}^{(r, \eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) d\tau \right) \phi(r, \eta; T, y) d\eta dr.$$

By (A.11) and Assumption 1.3, for every $\varepsilon > 0$ we have

$$\left| \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \right| \leq \frac{C}{r-\tau} \Gamma^{\mu+\varepsilon}(\tau, e^{(\tau-t)B}x; r, \eta), \quad t < \tau < s < r < T.$$

Therefore, considering also (2.13), for any $r \in]s, T]$, the function

$$(\tau, \eta) \mapsto |\mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \phi(r, \eta; T, y)|$$

is integrable on $[t, s] \times \mathbb{R}^N$. Thus Fubini's theorem yields

$$\begin{aligned} & \int_{\mathbb{R}^N} \left(\int_t^s \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) d\tau \right) \phi(r, \eta; T, y) d\eta \\ &= \int_t^s \int_{\mathbb{R}^N} \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta d\tau. \end{aligned}$$

Moreover, by the potential estimate (B.3) with $\delta = \frac{\alpha}{2}$, for any $\varepsilon > 0$ we have

$$\begin{aligned} & \left| \int_{\mathbb{R}^N} \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta \right| \\ & \leq \frac{C}{(T-r)^{1-\frac{\alpha}{4}} (r-\tau)^{1-\frac{\alpha}{4}}} \Gamma^{\mu+\varepsilon}(\tau, e^{(\tau-t)B}x; T, y). \end{aligned} \tag{2.21}$$

As the right-hand side term is integrable over $[t, s] \times [s, T]$ as a function of (τ, r) , we can apply once more Fubini's theorem to conclude that

$$G(t, x) = - \int_t^s \int_s^T \int_{\mathbb{R}^N} \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta dr d\tau. \tag{2.22}$$

Let us consider $H(t, x)$. For every $n \in \mathbb{N}$, we define $\varepsilon_n(r) := \frac{1}{n}(r-t)$. Note that, for every $r \in]t, s[$ we have $r - \varepsilon_n(r) \geq t$. Hence

$$\begin{aligned} H(t, x) &= \underbrace{\int_t^s \int_{\mathbb{R}^N} \mathbf{P}(r - \varepsilon_n(r), e^{(r-\varepsilon_n(r)-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta dr}_{=: \tilde{H}_n(t, x)} \\ &\quad - \underbrace{\int_t^s \int_{\mathbb{R}^N} \left(\mathbf{P}(r - \varepsilon_n(r), e^{(r-\varepsilon_n(r)-t)B}x; r, \eta) - \mathbf{P}(t, x; r, \eta) \right) \phi(r, \eta; T, y) d\eta dr}_{=: H_n(t, x)}. \end{aligned}$$

Once more, Remark 2.3 yields

$$H_n(t, x) = \int_t^s \int_{\mathbb{R}^N} \left(\int_t^{r-\varepsilon_n(r)} \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) d\tau \right) \phi(r, \eta; T, y) d\eta dr$$

(applying Fubini's theorem as above)

$$= \int_t^s \int_t^{r-\varepsilon_n(r)} \int_{\mathbb{R}^N} \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta d\tau dr$$

(setting $\delta_n(\tau) = \frac{\tau-t}{n-1}$ and applying Fubini's theorem again)

$$= \int_t^{s-\varepsilon_n(s)} \int_{\tau+\delta_n(\tau)}^s \int_{\mathbb{R}^N} \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta dr d\tau$$

(by (2.21) and applying Lebesgue's dominated convergence theorem)

$$\xrightarrow{n \rightarrow \infty} \int_t^s \int_{\tau}^s \int_{\mathbb{R}^N} \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta dr d\tau.$$

On the other hand, by the potential estimate (B.2), for any $n \in \mathbb{N}$ we have

$$\left| \int_{\mathbb{R}^N} \mathbf{P}(r - \varepsilon_n(r), e^{(r-\varepsilon_n(r)-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta \right| \leq C \frac{\Gamma^{\mu+\varepsilon}(\tau, e^{(\tau-t)B}x; T, y)}{(T-r)^{1-\frac{\mu}{2}}(T-r)^{\frac{\mu}{2}}},$$

$r \in [t, s].$

Thus Lebesgue's dominated convergence theorem yields

$$\lim_{n \rightarrow \infty} \tilde{H}_n(t, x) = \int_t^s \lim_{n \rightarrow \infty} \int_{\mathbb{R}^N} \mathbf{P}(r - \varepsilon_n(r), e^{(r-\varepsilon_n(r)-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta dr$$

(by (2.24), since $\eta \mapsto \phi(r, \eta; T, y)$ is a bounded and continuous function for every $r \in [t, s]$)

$$= \int_t^s \phi(r, e^{(r-t)B}x; T, y) dr.$$

We have proved that

$$H(t, x) = \int_t^s \int_{\tau}^s \int_{\mathbb{R}^N} \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta dr d\tau + \int_t^s \phi(\tau, e^{(\tau-t)B}x; T, y) d\tau.$$

This and (2.22) prove the statement. □

We are now in the position to prove Theorem 1.1, namely that $p = p(t, x; T, y)$ defined by (2.8) is a fundamental solution of $\mathcal{A} + Y$ in the sense of definition Definition 1.6, and that the Gaussian bounds from (1.13) to (1.16) are satisfied.

Proof of Theorem 1.1. Let $p = p(t, x; T, y)$ be defined by (2.8).

Step 1. We show that $p = p(t, x; T, y)$ satisfies point i) of Definition 1.6, namely that $p(\cdot, \cdot; T, y)$ is an integral solution to (2.18) on \mathcal{S}_T in the sense of Definition 1.5. By Lemma 2.1, we have

$$p(s, e^{(s-t)B}x; T, y) - p(t, x; T, y) = \mathbf{P}(s, e^{(s-t)B}x; T, y) - \mathbf{P}(t, x; T, y) + \Phi(s, e^{(s-t)B}x; T, y)$$

$$\begin{aligned}
 & - \Phi(t, x; T, y) \\
 & = - \int_t^S \left(\mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) + F(\tau, x; T, y) \right) d\tau. \tag{2.23}
 \end{aligned}$$

Furthermore, by (2.20) and since $\phi(t, x; T, y)$ solves the integral Eq. (2.9), we obtain

$$\begin{aligned}
 & \mathcal{A}^{(r,\eta)} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) + F(\tau, x; T, y) = \mathcal{A} \mathbf{P}(\tau, e^{(\tau-t)B}x; T, y) \\
 & + \int_\tau^T \int_{\mathbb{R}^N} \mathcal{A} \mathbf{P}(\tau, e^{(\tau-t)B}x; r, \eta) \phi(r, \eta; T, y) d\eta dr
 \end{aligned}$$

(by Proposition 2.2)

$$= \mathcal{A} \mathbf{P}(\tau, e^{(\tau-t)B}x; T, y) + \mathcal{A} \Phi(\tau, e^{(\tau-t)B}x; T, y) = \mathcal{A} p(\tau, e^{(\tau-t)B}x; T, y),$$

which, together with (2.23), concludes the proof.

Step 2. We show that $p = p(t, x; T, y)$ satisfies point ii) of Definition 1.6. In light of the estimate (2.13), it is straightforward to see that

$$|\Phi(t, x; T, y)| \leq C(T - t)^{\frac{\alpha}{2}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \quad (T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T.$$

Therefore, it is enough to prove that, for any fixed $(T, y) \in]0, T_0[\times \mathbb{R}^N$, we have

$$\lim_{\substack{(t,x) \rightarrow (T,y) \\ \tau < T}} \int_{\mathbb{R}^N} \mathbf{P}(t, x; T, \eta) f(\eta) d\eta = f(y), \quad f \in C_b(\mathbb{R}^N). \tag{2.24}$$

Recalling the definition of the parametrix \mathbf{P} , we add and subtract to obtain

$$\begin{aligned}
 \int_{\mathbb{R}^N} \mathbf{P}(t, x; T, \eta) f(\eta) d\eta & = \int_{\mathbb{R}^N} \Gamma^{(T,\eta)}(t, x; T, \eta) f(\eta) d\eta \\
 & = \int_{\mathbb{R}^N} \Gamma^{(T,y)}(t, x; T, \eta) f(\eta) d\eta \\
 & + \underbrace{\int_{\mathbb{R}^N} \left(\Gamma^{(T,\eta)}(t, x; T, \eta) - \Gamma^{(T,y)}(t, x; T, \eta) \right) f(\eta) d\eta}_{=: J(t,x)}.
 \end{aligned}$$

Furthermore, by estimate (A.12), for every $\varepsilon > 0$ one has

$$|J(t, x)| \leq C \int_{\mathbb{R}^N} |y - \eta|_B^\alpha \Gamma^{\mu+\varepsilon}(t, x; T, \eta) d\eta.$$

Eventually, (2.24) follows from classical arguments.

Step 3. We show the upper Gaussian bounds (1.13)–(1.14)–(1.15) for p and its derivatives. The proof of the lower Gaussian bound (1.16) is similar to that of Theorem 4.7 in [34] and Section 5.1.4. in [35], thus we omit it for sake of brevity.

The Gaussian bounds of Proposition A.1 and the definition of parametrix (2.6) yield the estimates (1.13)–(1.14)–(1.15) for $\mathbf{P} = \mathbf{P}(t, x; T, y)$. The estimates of Proposition 2.2 and the fact that $p = \mathbf{P} + \Phi$ conclude the proof. \square

Remark 2.6. Any integral solution u to Eq. (1.11) on \mathcal{S}_T in the sense of Definition 1.5, is Lie differentiable along Y almost everywhere on \mathcal{S}_T . Indeed, the set H_T of $(t, x) \in \mathcal{S}_T$ such that $Yu(t, x)$ in (1.12) exists finite, is measurable as the limit

$$\limsup_{\tau \rightarrow t^+} \frac{u(\tau, e^{(\tau-t)B}x) - u(t, x)}{\tau - t}$$

is a measurable function of (t, x) and the same holds for \liminf . This is a straightforward consequence of the continuity of u along the integral curves of Y . The fact that H_T has null Lebesgue measure stems from Fubini’s theorem, as u is absolutely continuous along the integral curves of Y and the map

$$(\tau, y) \mapsto (\tau, e^{\tau B}y)$$

is a diffeomorphism on \mathcal{S}_T .

3. Regularity of the fundamental solution

In this section we prove Theorem 1.2. Since $p(\cdot, \cdot; T, y)$ can be represented as in (2.8), we need to study the regularity of $\mathbf{P}(\cdot, \cdot; T, y)$ and $\Phi(\cdot, \cdot; T, y)$. While the former term can be easily dealt with by means of the Gaussian estimates of Appendix A, the latter has to be treated more carefully. We start with the proof of Theorem 1.2, which is based on the regularity estimates for $\Phi(\cdot, \cdot; T, y)$ and $\mathbf{P}(\cdot, \cdot; T, y)$ proved in Sects. 3.1 and 3.2, respectively.

Proof of Theorem 1.2. Let $\beta < \alpha$. For fixed $(T, y) \in \mathcal{S}_{T_0}$, we set

$$f(t, x) := p(t, x; T, y), \quad (t, x) \in \mathcal{S}_T.$$

We first note that, by definition of fundamental solution, (1.8) is satisfied with $Yf = -\mathcal{A}f$. Furthermore, for any $t \in]0, T[$, by (1.9) and the representation (2.8) we have

$$\|f\|_{C_B^{2,\beta}(\mathcal{S}_t)} = N_{\mathbf{P},1} + N_{\mathbf{P},2} + N_{\Phi,1} + N_{\Phi,2},$$

where

$$\begin{aligned} N_{\mathbf{P},1} &:= \sum_{i=1}^d \|\partial_{x_i} \mathbf{P}(\cdot, \cdot; T, y)\|_{C_Y^{1+\beta}(\mathcal{S}_t)} + \sum_{i,j=1}^d (\|\partial_{x_i x_j} \mathbf{P}(\cdot, \cdot; T, y)\|_{C_Y^\beta(\mathcal{S}_t)} \\ &\quad + \|\partial_{x_i x_j} \mathbf{P}(\cdot, \cdot; T, y)\|_{C_d^\beta(\mathcal{S}_t)}), \\ N_{\mathbf{P},2} &:= \|\mathcal{A} \mathbf{P}(\cdot, \cdot; T, y)\|_{L^\infty([0,t]; C_B^\beta(\mathbb{R}^N))}, \\ N_{\Phi,1} &:= \sum_{i=1}^d \|\partial_{x_i} \Phi(\cdot, \cdot; T, y)\|_{C_Y^{1+\beta}(\mathcal{S}_t)} + \sum_{i,j=1}^d (\|\partial_{x_i x_j} \Phi(\cdot, \cdot; T, y)\|_{C_Y^\beta(\mathcal{S}_t)} \end{aligned}$$

$$\begin{aligned}
 & + \|\partial_{x_i x_j} \Phi(\cdot, \cdot; T, y)\|_{C_d^\beta(\mathcal{S}_t)}, \\
 N_{\Phi,2} & := \|\mathcal{A}\Phi(\cdot, \cdot; T, y)\|_{L^\infty([0,t]; C_B^\beta(\mathbb{R}^N))}.
 \end{aligned}$$

Now, the estimates of Lemma 3.2 yield

$$N_{\mathbf{P},1} \leq \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}. \tag{3.1}$$

To bound $N_{\mathbf{P},2}$, first fix $i, j = 1, \dots, d$ and note that, by estimate (A.11), we obtain

$$\sup_{x \in \mathbb{R}^N} |\partial_{x_i x_j} \mathbf{P}(s, x; T, y)| \leq \frac{C}{(T-t)^{\frac{Q+2}{2}}}, \quad s < t. \tag{3.2}$$

Furthermore, (3.1) combined with Remark 1.1 yield

$$\sup_{x, v \in \mathbb{R}^N} \frac{|\partial_{x_i x_j} \mathbf{P}(s, x; T, y) - \partial_{x_i x_j} \mathbf{P}(s, v; T, y)|}{|x - v|_B^\beta} \leq \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}, \quad s < t. \tag{3.3}$$

Thus, by (3.2)–(3.3) we obtain

$$\|\partial_{x_i x_j} \mathbf{P}(s, \cdot; T, y)\|_{C_B^\beta(\mathbb{R}^N)} \leq \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}, \quad s < t,$$

which in turn implies

$$\|\partial_{x_i x_j} \mathbf{P}(\cdot, \cdot; T, y)\|_{L^\infty([0,t]; C_B^\beta(\mathbb{R}^N))} \leq \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}.$$

This, together with Assumption 1.3, prove

$$N_{\mathbf{P},2} \leq \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}.$$

The bound for $N_{\Phi,1}$ stems from the estimates of Proposition 3.1, which yield

$$N_{\Phi,1} \leq \frac{C}{(T-t)^{\frac{Q+2-(\alpha-\beta)}{2}}} \leq \frac{C}{(T-t)^{\frac{Q+2+\beta}{2}}}.$$

Eventually, the bound for $N_{\Phi,2}$ follows from the same arguments used to bound $N_{\mathbf{P},2}$. □

The rest of this section is devoted to the results utilized in the proof of Theorem 1.2. It is useful to introduce the following

Notation 3.1. Let $f = f(t, x; T, y)$ be a function defined for $(T, y) \in \mathcal{S}_{T_0}$ and $(t, x) \in \mathcal{S}_T$, suitably differentiable w.r.t. x . For any $i = 1, \dots, N$, we set

$$\partial_i f(t, x; T, y) := \partial_{x_i} f(t, x; T, y),$$

and we adopt analogous notations for the higher-order derivatives.

This notation is useful in order to compose partial derivatives with other functions. For instance, if $g = g(t, x)$ is a given function, then

$$\partial_i f(t, g(t, x); T, y) = \partial_{z_i} f(t, z; T, y)|_{z=g(t,x)}.$$

3.1. Regularity estimates of Φ

Now prove the Hölder estimates for $\Phi(\cdot, \cdot; T, y)$. We recall that Q denotes the homogeneous dimension of \mathbb{R}^N as in (1.17).

Proposition 3.1. *For every $\varepsilon > 0$ and $0 < \beta < \alpha$ there exists a positive constant C , only dependent on $T_0, \mu, B, \varepsilon, \alpha, \beta$ and the α -Hölder norms of the coefficients, such that, for any $i, j, k = 1, \dots, d$, we have*

$$\begin{aligned} & \left| \partial_i \Phi(s, e^{(s-t)B}x; T, y) - \partial_i \Phi(t, x; T, y) \right| \\ & \leq C(s-t)^{\frac{1+\beta}{2}} \frac{(T-t)^{Q/2}}{(T-s)^{\frac{Q+2-(\alpha-\beta)}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \end{aligned} \tag{3.4}$$

$$\begin{aligned} & \left| \partial_{ij} \Phi(s, e^{(s-t)B}x; T, y) - \partial_{ij} \Phi(t, x; T, y) \right| \\ & \leq C(s-t)^{\frac{\beta}{2}} \frac{(T-t)^{Q/2}}{(T-s)^{\frac{Q+2-(\alpha-\beta)}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \end{aligned} \tag{3.5}$$

$$\begin{aligned} & \left| \partial_{ij} \Phi(t, x + h\mathbf{e}_k; T, y) - \partial_{ij} \Phi(t, x; T, y) \right| \\ & \leq C|h|^\beta \frac{\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y)}{(T-t)^{\frac{2-(\alpha-\beta)}{2}}}, \end{aligned} \tag{3.6}$$

for every $(T, y) \in \mathcal{S}_{T_0}$, $(t, x) \in \mathcal{S}_T$, $t < s < T$ and $h \in \mathbb{R}$.

The proof of estimates (3.4)–(3.5) relies on the following

Lemma 3.1. *Let $(T, y) \in \mathcal{S}_{T_0}$. Then, for any $i = 1, \dots, d$, the function $u := \partial_i \mathbf{P}(\cdot, \cdot; T, y)$ is a strong Lie solution to the equation*

$$\mathcal{A}u + Yu = - \sum_{j=1}^{d+d_1} b_{ji} \partial_j \mathbf{P}(\cdot, \cdot; T, y) \text{ on } \mathcal{S}_T,$$

in the sense of Definition 1.5.

Proof. We note that

$$[\partial_i, Y]\mathbf{P}(t, x; T, y) = [\partial_i, \langle Bx, \nabla \rangle + \partial_t]\mathbf{P}(t, x; T, y) = \sum_{j=1}^{d+d_1} b_{ji} \partial_j \mathbf{P}(t, x; T, y),$$

for every $x \in \mathbb{R}^N$ and for almost every $t \in [0, T]$, where, in the last equality, we used that $b_{ji} = 0$ if $j > d+d_1$. While it is obvious that the previous identity holds for smooth functions of (t, x) , one can directly check that $\partial_i \partial_t \mathbf{P}(t, x; T, y) = \partial_t \partial_i \mathbf{P}(t, x; T, y)$ and thus the identity holds for the parametrix too. Therefore, we obtain

$$\begin{aligned} & \partial_i \mathbf{P}(s, e^{(s-t)B}x; \tau, \eta) - \partial_i \mathbf{P}(t, x; \tau, \eta) \\ & = \int_t^s (Y \partial_i \mathbf{P})(r, e^{(r-t)B}x; \tau, \eta) dr \end{aligned}$$

$$= \int_t^s \left((\partial_i Y \mathbf{P})(r, e^{(r-t)B}x; \tau, \eta) - [\partial_i, Y] \mathbf{P}(r, e^{(r-t)B}x; \tau, \eta) \right) dr$$

(by Remark 2.3)

$$= - \int_t^s \left((\partial_i \mathcal{A}^{(\tau, \eta)} \mathbf{P})(r, e^{(r-t)B}x; \tau, \eta) + \sum_{j=1}^{d+d_1} b_{ji} \partial_j \mathbf{P}(r, e^{(r-t)B}x; \tau, \eta) \right) dr$$

(since $\partial_i \mathcal{A}^{(\tau, \eta)} = \mathcal{A}^{(\tau, \eta)} \partial_i$)

$$= - \int_t^s \left((\mathcal{A}^{(\tau, \eta)} \partial_i \mathbf{P})(r, e^{(r-t)B}x; \tau, \eta) + \sum_{j=1}^{d+d_1} b_{ji} \partial_j \mathbf{P}(r, e^{(r-t)B}x; \tau, \eta) \right) dr.$$

□

We are now in the position to prove Proposition 3.1.

Proof of Proposition 3.1. Let $(T, y) \in \mathcal{S}_{T_0}$, $(t, x) \in \mathcal{S}_T$, $t < s < T$ and $h \in \mathbb{R}$ be fixed. Also fix $i, j, k \in \{1, \dots, d\}$. First we prove (3.4). By adding and subtracting, we have

$$\begin{aligned} & \partial_i \Phi(s, e^{(s-t)B}x; T, y) - \partial_i \Phi(t, x; T, y) \\ &= \int_s^T \int_{\mathbb{R}^N} \underbrace{\left(\partial_i \mathbf{P}(s, e^{(s-t)B}x; \tau, \eta) - \partial_i \mathbf{P}(t, x; \tau, \eta) \right)}_{=: I(\tau, \eta)} \phi(\tau, \eta; T, y) d\eta d\tau \\ & \quad - \underbrace{\int_t^s \int_{\mathbb{R}^N} \partial_i \mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau}_{=: L}. \end{aligned}$$

We consider the first term. By Lemma 3.1 and swapping the integrals as in the proof of Proposition 2.1, we have

$$\begin{aligned} & \int_s^T \int_{\mathbb{R}^N} I(\tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau \\ &= - \int_s^T \int_t^s \int_{\mathbb{R}^N} \left((\mathcal{A}^{(\tau, \eta)} \partial_i \mathbf{P})(r, e^{(r-t)B}x; \tau, \eta) \right. \\ & \quad \left. + \sum_{j=1}^{d+d_1} b_{ji} \partial_j \mathbf{P}(r, e^{(r-t)B}x; \tau, \eta) \right) \phi(\tau, \eta; T, y) d\eta dr d\tau. \end{aligned}$$

Therefore, the estimates of Proposition B.1 with $\delta = (\alpha - \beta)/2$ yield

$$\begin{aligned} & \left| \int_s^T \int_{\mathbb{R}^N} I(\tau, \eta) \phi(\tau, \eta; T, y) d\eta d\tau \right| \\ & \leq \int_s^T \int_t^s \frac{C}{(T - \tau)^{1 - \frac{\alpha - \beta}{4}} (\tau - r)^{\frac{3}{2} - \frac{\alpha + \beta}{4}}} \Gamma^{\mu + \varepsilon}(r, e^{(r-t)B}x; T, y) dr d\tau \end{aligned}$$

(by a standard estimate on $\Gamma^{\mu+\varepsilon}(r, e^{(r-t)B}x; T, y)$)

$$\leq C \underbrace{\int_s^T \int_t^s \frac{1}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-r)^{\frac{3}{2}-\frac{\alpha+\beta}{4}}} dr d\tau}_{=:K} \left(\frac{T-t}{T-s}\right)^{Q/2} \Gamma^{\mu+\varepsilon}(t, x; T, y). \tag{3.7}$$

We now bound K :

$$\begin{aligned} K &= \int_t^s \int_s^T \frac{1}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-r)^{\frac{3}{2}-\frac{\alpha+\beta}{4}}} d\tau dr \\ &\leq \int_t^s \int_s^T \frac{1}{(T-\tau)^{1-\frac{\alpha-\beta}{4}}(\tau-r)^{1-\frac{\alpha-\beta}{4}}} d\tau \frac{1}{(s-r)^{\frac{1}{2}-\frac{\beta}{2}}} dr \end{aligned}$$

(solving the integral in $d\tau$)

$$\begin{aligned} &\leq C \int_t^s \frac{1}{(T-r)^{1-\frac{\alpha-\beta}{2}}} \frac{1}{(s-r)^{\frac{1}{2}-\frac{\beta}{2}}} dr \leq \frac{C}{(T-s)^{1-\frac{\alpha-\beta}{2}}} \int_t^s \frac{1}{(s-r)^{\frac{1-\beta}{2}}} dr \\ &\leq \frac{C}{(T-s)^{1-\frac{\alpha-\beta}{2}}} (s-t)^{\frac{1+\beta}{2}}. \end{aligned} \tag{3.8}$$

On the other hand, estimate (B.3) with $\delta = \alpha - \beta$ yields

$$\begin{aligned} |L| &\leq \int_t^s \frac{C}{(T-\tau)^{1-\frac{\alpha-\beta}{2}}(\tau-t)^{\frac{1}{2}-\frac{\beta}{2}}} d\tau \Gamma^{\mu+\varepsilon}(t, x; T, y) \\ &\leq \frac{C}{(T-s)^{1-\frac{\alpha-\beta}{2}}} \int_t^s \frac{1}{(\tau-t)^{\frac{1-\beta}{2}}} d\tau \Gamma^{\mu+\varepsilon}(t, x; T, y) \\ &\leq \frac{C}{(T-s)^{1-\frac{\alpha-\beta}{2}}} (s-t)^{\frac{1+\beta}{2}} \Gamma^{\mu+\varepsilon}(t, x; T, y). \end{aligned}$$

This, together with (3.7)–(3.8), proves (3.4). Estimate (3.5) can be obtained following the same arguments.

We finally prove (3.6). By Proposition 2.2 we have

$$\begin{aligned} &\partial_{ij}(t, x + h\mathbf{e}_k; T, y) - \partial_{ij}\Phi(t, x; T, y) \\ &= \int_t^T \underbrace{\int_{\mathbb{R}^N} (\partial_{ij}\mathbf{P}(t, x + h\mathbf{e}_k; \tau, \eta) - \partial_{ij}\mathbf{P}(t, x; \tau, \eta)) \phi(\tau, \eta; T, y) d\eta}_{=:I(\tau)} d\tau. \end{aligned}$$

We first prove that

$$|I(\tau)| \leq C \frac{|h|^\beta}{(T - \tau)^{1-\frac{\alpha-\beta}{4}} (\tau - t)^{1-\frac{\alpha-\beta}{4}}} (\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y)),$$

$\tau \in]t, T[.$

We consider the case $\tau - t \geq h^2$. By the mean-value theorem, there exists a real \bar{h} with $|\bar{h}| \leq |h|$ such that

$$|\partial_{ij}\mathbf{P}(t, x + h\mathbf{e}_k; \tau, \eta) - \partial_{ij}\mathbf{P}(t, x; \tau, \eta)| = |h| |\partial_{ijk}\mathbf{P}(t, x + \bar{h}\mathbf{e}_k; \tau, \eta)|.$$

Therefore, by the estimate (B.3) with $\delta = (\alpha - \beta)/2$, we have

$$|I(\tau)| \leq C \frac{|h|}{(T - \tau)^{1-\frac{\alpha-\beta}{4}} (\tau - t)^{\frac{3}{2}-\frac{\alpha+\beta}{4}}} \Gamma^{\mu+\varepsilon}(t, x + \bar{h}\mathbf{e}_k; T, y)$$

(since $\tau - t \geq h^2$)

$$\leq C \frac{|h|^\beta}{(T - \tau)^{1-\frac{\alpha-\beta}{4}} (\tau - t)^{1-\frac{\alpha-\beta}{4}}} \Gamma^{\mu+\varepsilon}(t, x + \bar{h}\mathbf{e}_k; T, y)$$

(by standard estimates on $\Gamma^{\mu+\varepsilon}(t, x + \bar{h}\mathbf{e}_k; T, y)$ with $\tau - t \geq h^2$)

$$\leq C \frac{|h|^\beta}{(T - \tau)^{1-\frac{\alpha-\beta}{4}} (\tau - t)^{1-\frac{\alpha-\beta}{4}}} (\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y)).$$

We now consider the case $\tau - t < h^2$. Employing triangular inequality and estimate (B.3) with $\delta = (\alpha - \beta)/2$, we get

$$|I(\tau)| \leq \frac{C}{(T - \tau)^{1-\frac{\alpha-\beta}{4}} (\tau - t)^{1-\frac{\alpha+\beta}{4}}} (\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y))$$

(since $\tau - t < h^2$)

$$\leq C \frac{|h|^\beta}{(T - \tau)^{1-\frac{\alpha-\beta}{4}} (\tau - t)^{1-\frac{\alpha-\beta}{4}}} (\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y)).$$

Therefore, combining the previous estimates, we obtain

$$\begin{aligned} \left| \int_t^T I(\tau) d\tau \right| &\leq C |h|^\beta \int_t^T \frac{1}{(T - \tau)^{1-\frac{\alpha-\beta}{4}} (\tau - t)^{1-\frac{\alpha-\beta}{4}}} d\tau (\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) \\ &\quad + \Gamma^{\mu+\varepsilon}(t, x; T, y)) \\ &\leq C |h|^\beta \frac{1}{(T - t)^{1-\frac{\alpha-\beta}{2}}} (\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y)), \end{aligned}$$

which proves (3.6). □

3.2. Regularity estimates for the parametrix

We have the following Hölder estimates for \mathbf{P} . Notice that Notation 3.1 remains in force.

Lemma 3.2. *Let $0 \leq \beta \leq \alpha$. Then for every $\varepsilon > 0$ there exists a positive constant C , only dependent on $T_0, \mu, B, \varepsilon, \alpha, \beta$ and the α -Hölder norms of the coefficients, such that for any $i, j, k = 1, \dots, d$ we have*

$$|\partial_i \mathbf{P}(s, e^{(s-t)B}x; T, y) - \partial_i \mathbf{P}(t, x; T, y)| \leq C(s-t)^{\frac{1+\beta}{2}} \frac{(T-t)^{Q/2}}{(T-s)^{\frac{Q+2+\beta}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{3.9}$$

$$|\partial_{ij} \mathbf{P}(s, e^{(s-t)B}x; T, y) - \partial_{ij} \mathbf{P}(t, x; T, y)| \leq C(s-t)^{\frac{\beta}{2}} \frac{(T-t)^{Q/2}}{(T-s)^{\frac{Q+2+\beta}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{3.10}$$

$$|\partial_{ij} \mathbf{P}(t, x + h\mathbf{e}_k; T, y) - \partial_{ij} \mathbf{P}(t, x; T, y)| \leq C|h|^\beta \frac{1}{(T-t)^{\frac{2+\beta}{2}}} (\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y)), \tag{3.11}$$

for any $(T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T, t < s < T$ and $h \in \mathbb{R}$.

Proof. We first consider (3.9). By Lemma 3.1 we have

$$\begin{aligned} \partial_i \mathbf{P}(s, e^{(s-t)B}x; T, y) - \partial_i \mathbf{P}(t, x; T, y) &= - \int_t^s \left(\mathcal{A}^{(T,y)} \partial_i \mathbf{P}(r, e^{(r-t)B}x; T, y) \right. \\ &\quad \left. + \sum_{j=1}^{d+d_1} b_{ji} \partial_j \mathbf{P}(r, e^{(r-t)B}x; T, y) \right) dr. \end{aligned}$$

Therefore, by boundedness of the coefficients of $\mathcal{A}^{(T,y)}$ and the estimates of Proposition A.1, we obtain

$$\begin{aligned} &|\partial_i \mathbf{P}(s, e^{(s-t)B}x; T, y) - \partial_i \mathbf{P}(t, x; T, y)| \\ &\leq \int_t^s \frac{C}{(T-r)^{\frac{3}{2}}} \Gamma^{\mu+\varepsilon}(r, e^{(r-t)B}x; T, y) dr \\ &\leq \int_t^s \frac{C}{(T-r)^{\frac{3}{2}}} dr \left(\frac{T-t}{T-s} \right)^{Q/2} \Gamma^{\mu+\varepsilon}(t, x; T, y) \end{aligned}$$

(for any $\beta \leq 1$)

$$\leq C \frac{(s-t)^{\frac{1+\beta}{2}}}{(T-s)^{1+\frac{\beta}{2}}} \left(\frac{T-t}{T-s} \right)^{Q/2} \Gamma^{\mu+\varepsilon}(t, x; T, y).$$

The proof of (3.10) is based on analogous arguments.

We finally prove (3.11). As for (3.6), we first consider the case $T - t \geq h^2$. By the mean-value theorem, there exists a real \bar{h} with $|\bar{h}| \leq |h|$ such that

$$|\partial_{ij}\mathbf{P}(t, x + h\mathbf{e}_k; T, y) - \partial_{ij}\mathbf{P}(t, x; T, y)| = |h| |\partial_{ijk}\mathbf{P}(t, x + \bar{h}\mathbf{e}_k; T, y)|$$

(by estimate (A.11))

$$\leq C \frac{|h|}{(T - t)^{\frac{3}{2}}} \Gamma^{\mu+\varepsilon}(t, x + \bar{h}\mathbf{e}_k; T, y)$$

(since $T - t \geq h^2$ and by standard estimates on $\Gamma^{\mu+\varepsilon}(t, x + \bar{h}\mathbf{e}_k; T, y)$)

$$\leq C \frac{|h|^\beta}{(T - t)^{1+\frac{\beta}{2}}} (\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y)).$$

We now consider $T - t < h^2$. Employing triangular inequality and estimate (A.11) yields

$$\begin{aligned} & |\partial_{ij}\mathbf{P}(t, x + h\mathbf{e}_k; T, y) - \partial_{ij}\mathbf{P}(t, x; T, y)| \\ & \leq \frac{C}{T - t} (\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y)) \end{aligned}$$

(since $T - t < h^2$)

$$\leq C \frac{|h|^\beta}{(T - t)^{1+\frac{\beta}{2}}} (\Gamma^{\mu+\varepsilon}(t, x + h\mathbf{e}_k; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y)).$$

This concludes the proof of (3.11). □

4. Cauchy problem

In this section we anticipate some of the results on the Cauchy problem, namely well-posedness and Schauder estimates, which can be derived by utilizing the existence and the regularity estimates on the fundamental solution. We refer the reader to [25] for the details.

Consider the Cauchy problem

$$\begin{cases} \mathcal{A}u + Yu = f & \text{on } \mathcal{S}_T, \\ u(T, \cdot) = g & \text{on } \mathbb{R}^N, \end{cases} \tag{4.1}$$

with $T > 0$ being fixed hereafter.

Assumption 4.1. For some $\alpha \in]0, 1]$ and $\nu \in [0, 2 + \alpha]$, we have $f \in L^\infty([0, T]; C_B^\alpha(\mathbb{R}^N))$ and $g \in C_B^\nu(\mathbb{R}^N)$.

When $\nu > 1$, the definition of the anisotropic space $C_B^\nu(\mathbb{R}^N)$ is a standard extension of that in Definition 1.1. Once more, we refer to [25] for the details.

Theorem 4.1. (Well-posedness and Schauder estimate) *Under Assumptions 1.1, 1.2, 1.3 and 4.1, the function*

$$u(t, x) := \int_{\mathbb{R}^N} p(t, x; T, y)g(y)dy - \int_t^T \int_{\mathbb{R}^N} p(t, x; s, y)f(s, y)dyds$$

is the unique bounded solution of the Cauchy problem (4.1), in the sense of Definition 1.5. Furthermore, for any $\beta \in]0, \alpha[$, u satisfies

$$\|u\|_{C_B^{2,\beta}(\mathcal{S}_t)} \leq C((T-t)^{-\frac{2+\beta-\nu}{2}} \|g\|_{C_B^\nu(\mathbb{R}^N)} + \|f\|_{L^\infty([0,t]; C_B^\beta(\mathbb{R}^N))), \quad t \in]0, T[.$$

Remark 4.1. In [25] it is allowed for f to be unbounded with respect to the time variable, namely $\|f\|_{L^\infty([0,t]; C_B^\alpha(\mathbb{R}^N))} \sim (T-t)^{-\gamma}$ as $t \rightarrow T^-$, for some $\gamma \in [0, 1[$. Furthermore, the proofs in [25] can be extended to include the case of f and g with exponential growth in the space variable, namely:

$$|f(t, x)| + |g(x)| \leq Ce^{C|x|^2}, \quad x \in \mathbb{R}^N.$$

As a consequence of Proposition 4.1, we have the following result, which contains further useful properties that allow to view the fundamental solution as the transition probability density of a Markovian process.

Corollary 4.1. *Under the assumptions of Theorem 1.1 we have:*

(i) *The Chapman–Kolmogorov identity*

$$p(t, x; T, y) = \int_{\mathbb{R}^N} p(t, x; s, \eta)p(s, \eta; T, y)d\eta,$$

$$0 < t < s < T < T_0, \quad x, y \in \mathbb{R}^N;$$

(ii) *If the zero-th order coefficient a of \mathcal{A} is constant, i.e. $a(t, x) = \bar{a}$, then*

$$\int_{\mathbb{R}^N} p(t, x; T, y)dy = e^{\bar{a}(T-t)}, \quad 0 < t < T < T_0, \quad x \in \mathbb{R}^N.$$

Funding Open access funding provided by Alma Mater Studiorum - Università di Bologna within the CRUI-CARE Agreement.

Data availability Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A Gaussian estimates

We prove Gaussian estimates that are crucial in the analysis of Sects. 2 and 3. Here we follow the ideas in [7, Section 3], but with some technical difference. Namely, in the aforementioned paper the Kolmogorov operator acts on the forward variables of $\Gamma^{(s,v)}(t, x; T, y)$, whereas here we consider $\mathcal{A} + Y$ acting on the backward variables (t, x) . This has an impact on the spatial derivatives, which contain additional factors that require a careful analysis.

Throughout the appendix we suppose that Assumptions 1.1, 1.2 and 1.3 are satisfied and fix $(s, v) \in \mathcal{S}_{T_0}$. Denoting by B_0 the matrix B with null $*$ -blocks, we define the $N \times N$ matrices

$$\begin{aligned} \mathcal{C}_0(t) &:= \int_0^t e^{(t-\tau)B_0} \begin{pmatrix} I_d & 0 \\ 0 & 0 \end{pmatrix} e^{(t-\tau)B_0^*} d\tau, \\ \mathcal{C}_0^{(s,v)}(t, T) &:= \int_t^T e^{(T-\tau)B_0} A^{(s,v)}(\tau) e^{(T-\tau)B_0^*} d\tau, \end{aligned}$$

with $A^{(s,v)}$ as defined in (2.3). As an immediate consequence of Assumption 1.1 we can compare the quadratic forms associated to $\mathcal{C}^{(s,v)}$ (as in (2.2)), $\mathcal{C}_0^{(s,v)}$ with $\mathcal{C}(T-t)$ (as in (2.5)), $\mathcal{C}_0(T-t)$, respectively:

$$\begin{aligned} \frac{1}{\mu} \mathcal{C}(T-t) &\leq \mathcal{C}^{(s,v)}(t, T) \leq \mu \mathcal{C}(T-t), \\ \frac{1}{\mu} \mathcal{C}_0(T-t) &\leq \mathcal{C}_0^{(s,v)}(t, T) \leq \mu \mathcal{C}_0(T-t), \end{aligned} \tag{A.1}$$

for any $t \leq T$. Moreover, an asymptotic comparison near 0 of $\mathcal{C}^{(s,v)}$ and $\mathcal{C}_0^{(s,v)}$ holds:

Lemma A.1. *There exist two positive constants C and δ , only dependent on μ and B , such that*

$$\begin{aligned} \frac{1}{2\mu} \mathcal{C}_0(T-t) &\leq \mathcal{C}^{(s,v)}(t, T) \leq 2\mu \mathcal{C}_0(T-t), \\ \frac{1}{(2\mu)^N} \det \mathcal{C}_0(T-t) &\leq \det \mathcal{C}^{(s,v)}(t, T) \leq (2\mu)^N \det \mathcal{C}_0(T-t), \end{aligned}$$

for any $0 < T-t < \delta$. Analogous estimates hold for $(\mathcal{C}^{(s,v)}(t, T))^{-1}$.

Proof. It follows from the same arguments of [23, Lemma 3.1]: the proof is only based on the properties of the matrices A and B , and it is not relevant whether A has constant or time-dependent entries. □

Remark A.1. We note that $|\cdot|_B$ is homogeneous with respect to the family of dilations defined by the matrices

$$D(\lambda) := \text{diag}(\lambda I_d, \lambda^3 I_{d_1}, \dots, \lambda^{2q+1} I_{d_q}), \quad \lambda \geq 0.$$

In [23, Proposition 2.3] it is proved that

$$\mathcal{C}_0(t) = D(\sqrt{t})\mathcal{C}_0(1)D(\sqrt{t}), \quad t \geq 0. \tag{A.2}$$

Therefore, for $0 < T - t < \delta$ with δ as in Lemma A.1,

$$\frac{(T - t)^Q}{(2\mu)^N} \det \mathcal{C}_0(1) \leq \det \mathcal{C}^{(s,v)}(t, T) \leq (2\mu)^N (T - t)^Q \det \mathcal{C}_0(1).$$

To compute the spatial derivatives of $\Gamma^{(s,v)}(t, x; T, y)$ it is useful noticing that

$$\Gamma^{(s,v)}(t, x; T, y) = \mathbf{G}(H^{(s,v)}(t, T), e^{-(T-t)B}y - x), \quad (T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T,$$

where

$$H^{(s,v)}(t, T) := e^{-(T-t)B}\mathcal{C}^{(s,v)}(t, T)e^{-(T-t)B*}.$$

Since $\mathcal{C}^{(s,v)}(t, T)$ is symmetric positive definite and $e^{-(T-t)B}$ is non-singular, then $H^{(s,v)}(t, T)$ is symmetric and positive definite for every $0 \leq t < T$.

In order to give estimates on the matrix $H^{(s,v)}$ we need to study the elements of e^{tB} . We recall the block partition (1.6) of the matrix B : for $h, k = 0, \dots, q$, we denote the $d_h \times d_k$ block of B by

$$\mathcal{Q}_{hk} := (b_{ij})_{\substack{i=\bar{d}_{h-1}+1, \dots, \bar{d}_h \\ j=\bar{d}_{k-1}+1, \dots, \bar{d}_k}},$$

with \bar{d}_h as in (1.7). Note that by (1.6) we have

$$\begin{cases} \mathcal{Q}_{hk} = 0_{d_h \times d_k} & \text{if } h > k + 1, \\ \mathcal{Q}_{hk} = B_h & \text{if } h = k + 1, \\ \mathcal{Q}_{hk} = * & \text{if } h < k + 1. \end{cases} \tag{A.3}$$

Analogously, for $n \in \mathbb{N}$, we can consider the same block decomposition for B^n . We denote by $\mathcal{Q}_{hk}^{(n)}$ the $d_h \times d_k$ block of B^n .

Lemma A.2. *Let $h, k = 0, \dots, q$ and $n \in \mathbb{N}$. Then*

$$\mathcal{Q}_{hk}^{(n)} = 0_{d_h \times d_k}, \quad h > k + n, \tag{A.4}$$

which is $(B^n)_{ij} = 0$ if $i \in \{\bar{d}_{h-1} + 1, \dots, \bar{d}_h\}$ and $j \in \{\bar{d}_{k-1} + 1, \dots, \bar{d}_k\}$.

Proof. We proceed by induction on n . The case of $n = 1$ is obvious (see (A.3)). Now we assume that (A.4) holds for a certain $n \in \mathbb{N}$. For $h > k + n + 1$ we have

$$\mathcal{Q}_{hk}^{(n+1)} = \sum_{m=0}^q \mathcal{Q}_{hm}^{(n)} \mathcal{Q}_{mk}.$$

If $m < h - n$, then $\mathcal{Q}_{hm}^{(n)} = 0_{d_h \times d_m}$ by inductive hypothesis; if $m \geq h - n$, then $m > k + 1$ and $\mathcal{Q}_{mk} = 0_{d_m \times d_k}$. Therefore $\mathcal{Q}_{hk}^{(n+1)} = 0_{d_h \times d_k}$. \square

Lemma A.3. *Let $h, k = 1, \dots, q$ such that $h - k =: n \in \mathbb{N}$. For any $i \in \{\bar{d}_{h-1} + 1, \dots, \bar{d}_h\}$ and $j \in \{\bar{d}_{k-1} + 1, \dots, \bar{d}_k\}$ we have*

$$\left(e^{tB} \right)_{ij} = O(t^n), \quad \text{as } t \rightarrow 0.$$

Proof. From Lemma A.2 we have that $(B^m)_{ij} = 0$ for every $m = 0, \dots, n - 1$, since $\mathcal{Q}_{hk}^{(m)} = 0_{d_h \times d_k}$ for $h - k = n > m$. Therefore

$$(e^{tB})_{ij} = \frac{t^n (B^n)_{ij}}{n!} + O(t^{n+1}), \quad \text{as } t \rightarrow 0.$$

□

Lemma A.4. *There exists a positive constant C that only depends on μ, B and T_0 such that, for every $i, j = 1, \dots, d$ and $k = d + 1, \dots, d + d_1$,*

$$\left| (H^{(s,v)}(t, T)^{-1}x)_i \right| \leq \frac{C}{\sqrt{T-t}} \left| D(\sqrt{T-t})^{-1} e^{(T-t)B}x \right|, \tag{A.5}$$

$$\left| (H^{(s,v)}(t, T)^{-1})_{ij} \right| \leq \frac{C}{T-t}, \tag{A.6}$$

$$\left| (H^{(s,v)}(t, T)^{-1}x)_k \right| \leq \frac{C}{(T-t)^{\frac{3}{2}}} \left| D(\sqrt{T-t})^{-1} e^{(T-t)B}x \right|, \tag{A.7}$$

$$\left| (H^{(s,v)}(t, T)^{-1})_{ik} \right| \leq \frac{C}{(T-t)^2}, \tag{A.8}$$

for any $0 < T < T_0$ and $(t, x) \in \mathcal{S}_T$.

Proof. We prove the first inequality. Setting $\tau = T - t$, we have

$$\begin{aligned} \left| (H^{(s,v)}(t, T)^{-1}x)_i \right| &= \frac{1}{\sqrt{\tau}} \left| \left(D(\sqrt{\tau}) e^{\tau B^*} \mathcal{C}^{(s,v)}(t, T)^{-1} e^{\tau B}x \right)_i \right| \\ &\leq \frac{1}{\sqrt{\tau}} \sum_{n=1}^N \left| \left(D(\sqrt{\tau}) e^{\tau B^*} D(\sqrt{\tau})^{-1} \right)_{in} \right| \\ &\quad \left\| D(\sqrt{\tau}) \mathcal{C}^{(s,v)}(t, T)^{-1} D(\sqrt{\tau}) \right\| \\ &\quad \left| D(\sqrt{\tau})^{-1} e^{\tau B}x \right|. \end{aligned}$$

By Lemma A.1 there exists a positive constant δ such that, if $0 < \tau < \delta$, we have

$$\begin{aligned} \left\| D(\sqrt{\tau}) \mathcal{C}^{(s,v)}(t, T)^{-1} D(\sqrt{\tau}) \right\| &\leq \sup_{|y|=1} \langle D(\sqrt{\tau}) \mathcal{C}^{(s,v)}(t, T)^{-1} D(\sqrt{\tau})y, y \rangle \\ &\leq 2\mu \sup_{|y|=1} \langle \mathcal{C}_0(\tau)^{-1} D(\sqrt{\tau})y, D(\sqrt{\tau})y \rangle \\ &= 2\mu \|\mathcal{C}_0(1)^{-1}\|, \end{aligned}$$

where the last equality follows from (A.2). If $\delta \leq \tau < T_0$, by Eq. (A.1) we have

$$\left\| D(\sqrt{\tau})\mathcal{C}^{(s,v)}(t, T)^{-1}D(\sqrt{\tau}) \right\| \leq \mu \left\| D(\sqrt{\tau})\mathcal{C}(\tau)^{-1}D(\sqrt{\tau}) \right\|,$$

which is bounded by a constant that depends only on μ, T_0 and B .

In order to conclude the proof of (A.5), we let h_n be the only $h \in \{0, \dots, q\}$ such that $\bar{d}_{h-1} + 1 \leq n \leq \bar{d}_h$. Then, by Lemma A.3, since $i \in \{1, \dots, d\}$, we obtain

$$\begin{aligned} (D(\sqrt{\tau})e^{\tau B^*}D(\sqrt{\tau})^{-1})_{in} &= D(\sqrt{\tau})_{ii}(e^{\tau B^*})_{in}D(\sqrt{\tau})_{nn}^{-1} \\ &= \tau^{\frac{1}{2}} \left(e^{\tau B} \right)_{ni} \tau^{-\frac{2h_n+1}{2}} = O(1) \quad \text{as } \tau \rightarrow 0. \end{aligned}$$

Estimate (A.6) follows from (A.5) choosing $x = \mathbf{e}_j$. Estimates (A.7) and (A.8) can be proved following the same arguments, noticing that for $k = d + 1, \dots, d + d_1$ we have $D(\tau)_{kk} = \tau^3$. □

Finally, we provide Gaussian estimates for $\Gamma^{(s,v)}(t, x; T, y)$ and its derivatives up to the fourth order that will be used to study the Hölder regularity of the second order derivatives of the fundamental solution via the representation (2.8)–(2.12). The following result can be proved as [7, Proposition 3.5].

Lemma A.5. *For every $\beta \geq 0$ and $\varepsilon > 0$ there exists a positive constant C , only dependent on T_0, μ, B, ε and β , such that*

$$\begin{aligned} |w_j|^\beta \Gamma^{(s,v)}(t, x; T, y) &\leq C\Gamma^{\mu+\varepsilon}(t, x; T, y), \\ (T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T, i = 1, \dots, N, \end{aligned} \tag{A.9}$$

where

$$w = D(\sqrt{T-t})^{-1} \left(y - e^{(T-t)B}x \right).$$

Notation A.1. Let $\nu = (\nu_1, \dots, \nu_N) \in \mathbb{N}_0^N$ be a multi-index. We define the B -length of ν as

$$[\nu]_B := \sum_{j=0}^q (2j + 1) \sum_{i=\bar{d}_{j-1}+1}^{\bar{d}_j} \nu_i.$$

Moreover, as usual $\partial_x^\nu = \partial_{x_1}^{\nu_1} \dots \partial_{x_N}^{\nu_N}$.

Combining Lemmas A.4 and A.5 with [7, Proposition 3.1, 3.6 and Lemma 5.2], some lengthy but straightforward computations show the following

Proposition A.1. *We have*

$$\frac{1}{\mu^N} \Gamma^{\frac{1}{\mu}}(t, x; T, y) \leq \Gamma^{(s,v)}(t, x; T, y) \leq \mu^N \Gamma^\mu(t, x; T, y). \tag{A.10}$$

for any $(T, y) \in \mathcal{S}_{T_0}$ and $(t, x) \in \mathcal{S}_T$. Moreover, for every $\varepsilon > 0$ and $v \in \mathbb{N}_0^N$ with $[v]_B \leq 4$, there exists a positive constant C , only dependent on T_0, μ, B and ε , such that

$$|\partial_x^v \Gamma^{(s,v)}(t, x; T, y)| \leq \frac{C}{(T-t)^{\frac{[v]_B}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{A.11}$$

$$\left| \partial_x^v \Gamma^{(s,v)}(t, x; T, y) - \partial_x^v \Gamma^{(s,w)}(t, x; T, y) \right| \leq C \frac{|v-w|_B^\alpha}{(T-t)^{\frac{[v]_B}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{A.12}$$

for any $(T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T$ and $w \in \mathbb{R}^N$.

B potential estimates

We study $\Phi = \Phi(t, x; T, y)$ in (2.19) and its derivatives w.r.t. to the variables x_1, \dots, x_d . To do so, we have to deal with some singular integrals. We follow the steps in [7, Section 5], but we remark that the estimates of Proposition B.1 extend the ones in the aforementioned paper to higher order derivatives. This is needed to prove the optimal regularity of $\Phi(t, x; T, y)$ and thereafter of $p(t, x; T, y)$.

We set

$$J(t, x; \tau; T, y) := \int_{\mathbb{R}^N} \mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta, \tag{B.1}$$

$(T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T, \tau \in]t, T[.$

Proposition B.1. *For every $\varepsilon > 0, v \in \mathbb{N}_0^N$ with $[v]_B \leq 4$ and $0 < \delta < \alpha$, there exists a positive constant C , only dependent on $N, T_0, \mu, B, \delta, \alpha$ and ε , such that,*

$$|J(t, x; \tau; T, y)| \leq \frac{C}{(T-\tau)^{1-\frac{\alpha}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y) \tag{B.2}$$

$$|\partial_x^v J(t, x; \tau; T, y)| \leq \frac{C}{(T-\tau)^{1-\frac{\delta}{2}} (\tau-t)^{\frac{[v]_B-(\alpha-\delta)}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y), \tag{B.3}$$

for every $(T, y) \in \mathcal{S}_{T_0}, (t, x) \in \mathcal{S}_T$ and $\tau \in]t, T[.$

Proof. The proof relies on Proposition 2.1: (B.2) can be easily obtained by applying estimate (A.10) to $\mathbf{P}(t, x; \tau, \eta)$, estimate (2.13) to $\phi(\tau, \eta; T, y)$ and the Chapman–Kolmogorov identity.

We provide a full proof of (B.3) in the case of $\partial_x^v = \partial_{x_i x_j}$, with $i, j \leq d$, the proof for higher order derivatives being analogous. The idea is to combine (2.14) with the techniques in [7, Proposition 5.3] and [36, Proposition 3.2]. Let $(t, x) \in \mathcal{S}_T$ and $\tau \in]t, T[$ be fixed. By estimates (A.11) and (2.13), we have

$$\partial_{x_i x_j} J(t, x; \tau; T, y) = \int_{\mathbb{R}^N} \partial_{x_i x_j} \mathbf{P}(t, x; \tau, \eta) \phi(\tau, \eta; T, y) d\eta.$$

We set $= \frac{t+T}{2}$ and consider two separate cases:

Case $t < \tau < T$. By (A.11) and (2.13), we have that for every $\varepsilon > 0$ and $0 < \delta < \alpha$ there exists a positive constant C such that

$$|\partial_{x_i x_j} J(t, x; \tau; T, y)| \leq \int_{\mathbb{R}^N} \frac{C}{(T - \tau)^{1-\frac{\alpha}{2}}(\tau - t)} \Gamma^{\mu+\varepsilon}(t, x; \tau, \eta) \Gamma^{\mu+\varepsilon}(\tau, \eta; T, y) d\eta$$

(by the Chapman–Kolmogorov equation)

$$\leq \frac{C}{(T - \tau)^{1-\frac{\alpha}{2}}(\tau - t)} \Gamma^{\mu+\varepsilon}(t, x; T, y)$$

(since $T - \tau < \tau - t$)

$$\leq \frac{C}{(T - \tau)^{1-\frac{\delta}{2}}(\tau - t)^{1-\frac{\alpha-\delta}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y).$$

Case $t < \tau \leq T$. Here we need to handle with care the singularity of $\partial_{x_i x_j} \mathbf{P}(t, x; \tau, \eta)$ for small $\tau - t$. Note that in this case the following inequalities hold true:

$$\tau - t \leq \frac{T - t}{2} \leq T - \tau < T - t. \tag{B.4}$$

We have

$$\partial_{x_i x_j} J(t, x; \tau; T, y) = K_1 + K_2 + K_3,$$

where, setting $\xi = e^{(\tau-t)B}x$,

$$K_1 := \int_{\mathbb{R}^N} \partial_{x_i x_j} \Gamma^{(\tau, \eta)}(t, x; \tau, \eta) (\phi(\tau, \eta; T, y) - \phi(\tau, \xi; T, y)) d\eta,$$

$$K_2 := \phi(\tau, \xi; T, y) \int_{\mathbb{R}^N} \left(\partial_{x_i x_j} \Gamma^{(\tau, \eta)}(t, x; \tau, \eta) - \partial_{x_i x_j} \Gamma^{(\tau, v)}(t, x; \tau, \eta) \Big|_{v=\xi} \right) d\eta,$$

$$K_3 := \phi(\tau, x; T, y) \int_{\mathbb{R}^N} \partial_{x_i x_j} \Gamma^{(\tau, v)}(t, x; \tau, \eta) \Big|_{v=\xi} d\eta.$$

We first consider K_1 . By (2.14) and (A.11), for every $\varepsilon > 0$ and $0 < \delta < \alpha$ there exists a positive constant C such that

$$\begin{aligned} |K_1| &\leq \frac{C}{(T - \tau)^{1-\frac{\delta}{2}}} \int_{\mathbb{R}^N} \frac{|\eta - \xi|_B^{\alpha-\delta}}{(\tau - t)} \Gamma^{\mu+\frac{\varepsilon}{2}}(t, x; \tau, \eta) \\ &\quad \times (\Gamma^{\mu+\varepsilon}(\tau, \xi; T, y) + \Gamma^{\mu+\varepsilon}(\tau, \eta; T, y)) d\eta \end{aligned}$$

(by (A.9))

$$\begin{aligned} &\leq \frac{C}{(T - \tau)^{1-\frac{\delta}{2}}} \int_{\mathbb{R}^N} \frac{1}{(\tau - t)^{1-\frac{\alpha-\delta}{2}}} \Gamma^{\mu+\varepsilon}(t, x; \tau, \eta) \\ &\quad \times (\Gamma^{\mu+\varepsilon}(\tau, \xi; T, y) + \Gamma^{\mu+\varepsilon}(\tau, \eta; T, y)) d\eta \end{aligned}$$

(integrating in η and by the Chapman–Kolmogorov identity)

$$\leq \frac{C}{(T - \tau)^{1-\frac{\delta}{2}}(\tau - t)^{1-\frac{\alpha-\delta}{2}}} (\Gamma^{\mu+\varepsilon}(\tau, \xi; T, y) + \Gamma^{\mu+\varepsilon}(t, x; T, y))$$

(by (B.4))

$$\leq \frac{C}{(T - \tau)^{1-\frac{\delta}{2}}(\tau - t)^{1-\frac{\alpha-\delta}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y).$$

Consider now K_2 . By (2.13) and (A.12), we obtain

$$|K_2| \leq C \frac{\Gamma^{\mu+\varepsilon}(\tau, \xi; T, y)}{(T - \tau)^{1-\frac{\alpha}{2}}} \int_{\mathbb{R}^N} \frac{|\eta - \xi|_B^\alpha}{\tau - t} \Gamma^{\mu+\varepsilon}(t, x; \tau, \eta) d\eta$$

(by (A.9) and integrating in η)

$$\leq \frac{C}{(T - \tau)^{1-\frac{\alpha}{2}}(\tau - t)^{1-\frac{\alpha}{2}}} \Gamma^{\mu+\varepsilon}(\tau, \xi; T, y)$$

(again by (B.4))

$$\leq \frac{C}{(T - \tau)^{1-\frac{\delta}{2}}(\tau - t)^{1-\frac{\alpha-\delta}{2}}} \Gamma^{\mu+\varepsilon}(t, x; T, y).$$

Finally, $K_3 = 0$ since

$$\int_{\mathbb{R}^N} \partial_{x_i x_j} \Gamma^{(\tau, v)}(t, x; \tau, \eta) d\eta = \partial_{x_i x_j} \int_{\mathbb{R}^N} \Gamma^{(\tau, v)}(t, x; \tau, \eta) d\eta = 0$$

for any $v \in \mathbb{R}^N$. □

REFERENCES

- [1] Aneschi, F., Rebusci, A.: On the fundamental solution for degenerate kolmogorov equations with rough coefficients. *Journal of Elliptic and Parabolic Equations* (2022). <https://doi.org/10.1007/s41808-022-00191-8>.
- [2] Barucci, E., Polidoro, S., Vespri, V.: Some results on partial differential equations and Asian options. *Math. Models Methods Appl. Sci.* **11**(3), 475–497 (2001). <https://doi.org/10.1142/S0218202501000945>.
- [3] Biagi, S., Bramanti, M.: Schauder estimates for Kolmogorov–Fokker–Planck operators with coefficients measurable in time and Hölder continuous in space. Preprint, [arXiv:2205.10270](https://arxiv.org/abs/2205.10270) (2022)
- [4] Bramanti, M., Polidoro, S.: Fundamental solutions for Kolmogorov-Fokker-Planck operators with time-depending measurable coefficients. *Math. Eng.* **2**(4), 734–771 (2020). <https://doi.org/10.3934/mine.2020035>.
- [5] De Raynal, P.C.: Strong existence and uniqueness for degenerate sde with hölder drift. *Annales de l’Institut Henri Poincaré, Probabilités et Statistiques*, vol. 53, pp. 259–286 (2017)
- [6] Delarue, F., Menozzi, S.: Density estimates for a random noise propagating through a chain of differential equations. *J. Funct. Anal.* **259**(6), 1577–1630 (2010). <https://doi.org/10.1016/j.jfa.2010.05.002>.

- [7] Di Francesco, M., Pascucci, A.: On a class of degenerate parabolic equations of Kolmogorov type. *AMRX Appl. Math. Res. Express* (3), 77–116 (2005)
- [8] Di Francesco, M., Pascucci, A.: A continuous dependence result for ultraparabolic equations in option pricing. *J. Math. Anal. Appl.* **336**(2), 1026–1041 (2007). <https://doi.org/10.1016/j.jmaa.2007.03.031>.
- [9] Di Francesco, M., Polidoro, S.: Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form. *Adv. Differential Equations* **11**(11), 1261–1320 (2006)
- [10] Dong, H., Yastrzhembskiy, T.: Global Schauder estimates for kinetic Kolmogorov–Fokker–Planck equations. Preprint, [arXiv:2209.00769](https://arxiv.org/abs/2209.00769) (2022)
- [11] Folland, G.B., Stein, E.M.: Hardy spaces on homogeneous groups, *Mathematical Notes*, vol. 28. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo (1982)
- [12] Friedman, A.: Partial differential equations of parabolic type. Prentice-Hall Inc., Englewood Cliffs, N.J. (1964)
- [13] Henderson, C., Snelson, S.: C^∞ smoothing for weak solutions of the inhomogeneous Landau equation. *Arch. Ration. Mech. Anal.* **236**(1), 113–143 (2020). <https://doi.org/10.1007/s00205-019-01465-7>.
- [14] Henderson, C., Wang, W.: Kinetic Schauder estimates with time-irregular coefficients and uniqueness for the Landau equation. Preprint, [arXiv:2205.12930](https://arxiv.org/abs/2205.12930) (2022)
- [15] Hörmander, L.: Hypoelliptic second order differential equations. *Acta Math.* **119**, 147–171 (1967). <https://doi.org/10.1007/BF02392081>.
- [16] Il'in, A.M.: On a class of ultraparabolic equations. *Dokl. Akad. Nauk SSSR* **159**, 1214–1217 (1964)
- [17] Imbert, C., Mouhot, C.: The Schauder estimate in kinetic theory with application to a toy nonlinear model. *Ann. H. Lebesgue* **4**, 369–405 (2021). <https://doi.org/10.5802/ahl.75>.
- [18] Imbert, C., Silvestre, L.: The Schauder estimate for kinetic integral equations. *Anal. PDE* **14**(1), 171–204 (2021). <https://doi.org/10.2140/apde.2021.14.171>.
- [19] Kolmogorov, A.: Zufällige Bewegungen (zur Theorie der Brownschen Bewegung). *Ann. of Math.* (2) **35**(1), 116–117 (1934). <https://doi.org/10.2307/1968123>.
- [20] Krylov, N.V.: Hörmander's theorem for stochastic partial differential equations. *Algebra i Analiz* **27**(3), 157–182 (2015). <https://doi.org/10.1090/spmj/1398>.
- [21] Lanconelli, A., Pagliarani, S., Pascucci, A.: Local densities for a class of degenerate diffusions. *Ann. Inst. Henri Poincaré Probab. Stat.* **56**(2), 1440–1464 (2020). <https://doi.org/10.1214/19-AIHP1009>.
- [22] Lanconelli, E., Pascucci, A., Polidoro, S.: Linear and nonlinear ultraparabolic equations of Kolmogorov type arising in diffusion theory and in finance. In: *Nonlinear problems in mathematical physics and related topics, II*, Int. Math. Ser. (N. Y.), vol. 2, pp. 243–265. Kluwer/Plenum, New York (2002). https://doi.org/10.1007/978-1-4615-0701-7_14.
- [23] Lanconelli, E., Polidoro, S.: On a class of hypoelliptic evolution operators. pp. 29–63 (1994). *Partial differential equations, II* (Turin, 1993)
- [24] Lorenzi, L.: Schauder estimates for degenerate elliptic and parabolic problems with unbounded coefficients in \mathbb{R}^N . *Differential Integral Equations* **18**(5), 531–566 (2005)
- [25] Lucertini, G., Pagliarani, S., Pascucci, A.: Optimal Schauder estimates for kinetic Kolmogorov equations with time measurable coefficients. Preprint, [arXiv:2304.13392](https://arxiv.org/abs/2304.13392) (2023)
- [26] Lunardi, A.: Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in \mathbb{R}^n . *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* (4) **24**(1), 133–164 (1997). http://www.numdam.org/item?id=ASNSP_1997_4_24_1_133_0
- [27] Manfredini, M.: The Dirichlet problem for a class of ultraparabolic equations. *Adv. Differential Equations* **2**(5), 831–866 (1997)
- [28] McKean Jr., H.P., Singer, I.M.: Curvature and the eigenvalues of the Laplacian. *J. Differential Geometry* **1**(1), 43–69 (1967). <http://projecteuclid.org/euclid.jdg/1214427880>
- [29] Menozzi, S.: Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electron. Commun. Probab.* **16**, 234–250 (2011). <https://doi.org/10.1214/ECP.v16-1619>.
- [30] Pagliarani, S., Pascucci, A., Pignotti, M.: Intrinsic Taylor formula for Kolmogorov-type homogeneous groups. *J. Math. Anal. Appl.* **435**(2), 1054–1087 (2016). <https://doi.org/10.1016/j.jmaa.2015.10.080>.

- [31] Pagliarani, S., Pascucci, A., Pignotti, M.: Intrinsic expansions for averaged diffusion processes. *Stochastic Process. Appl.* **127**(8), 2560–2585 (2017). <https://doi.org/10.1016/j.spa.2016.12.002>.
- [32] Pagliarani, S., Polidoro, S.: A Yosida’s parametrix approach to Varadhan’s estimates for a degenerate diffusion under the weak Hörmander condition. *J. Math. Anal. Appl.* **517**(1), Paper No. 126538, 42 (2023). <https://doi.org/10.1016/j.jmaa.2022.126538>.
- [33] Pascucci, A.: PDE and martingale methods in option pricing, *Bocconi & Springer Series*, vol. 2. Springer, Milan; Bocconi University Press, Milan (2011). <https://doi.org/10.1007/978-88-470-1781-8>.
- [34] Pascucci, A., Pesce, A.: The parametrix method for parabolic SPDEs. *Stochastic Process. Appl.* **130**(10), 6226–6245 (2020). <https://doi.org/10.1016/j.spa.2020.05.008>.
- [35] Pascucci, A., Pesce, A.: On stochastic Langevin and Fokker-Planck equations: the two-dimensional case. *J. Differential Equations* **310**, 443–483 (2022). <https://doi.org/10.1016/j.jde.2021.11.004>.
- [36] Polidoro, S.: On a class of ultraparabolic operators of Kolmogorov-Fokker-Planck type. *Matematiche (Catania)* **49**(1), 53–105 (1995) (1994)
- [37] Priola, E.: Global Schauder estimates for a class of degenerate Kolmogorov equations. *Studia Math.* **194**(2), 117–153 (2009). <https://doi.org/10.4064/sm194-2-2>.
- [38] Sonin, I.M.: A class of degenerate diffusion processes. *Teor. Veroyatnost. i Primenen* **12**, 540–547 (1967)
- [39] Weber, M.: The fundamental solution of a degenerate partial differential equation of parabolic type. *Trans. Amer. Math. Soc.* **71**, 24–37 (1951). <https://doi.org/10.2307/1990857>.

*Stefano Pagliarani, Giacomo Lucertini and
Andrea Pascucci
Dipartimento di Matematica
Università di Bologna
Bologna
Italy
E-mail: stefano.pagliarani9@unibo.it*

*Giacomo Lucertini
E-mail: giacomo.lucertini3@unibo.it*

*Andrea Pascucci
E-mail: andrea.pascucci@unibo.it*

Accepted: 10 September 2023