Alma Mater Studiorum
Università di Bologna

DOTTORATO DI RICERCA IN MATEMATICA
Ciclo XXX

Settore Concorsuale: 01/A3
Settore Scientifico Disciplinare: MAT 06

Averaged stochastic processes and Kolmogorov operators

Presentata da: Michele Pignotti

Coordinatore Dottorato:
Chiar.ma Prof.ssa
Giovanna Citti

Supervisore:
Chiar.mo Prof.
Andrea Pascucci

Esame finale anno 2018

Acknowledgements

I deeply thank professors Carlos Vázquez Cendón, Stefano Pagliarani and Andrea Pascucci for their guidance through the three years of my Ph.D., countless useful discussions and teachings.
Without them this thesis could not have been possible.

Contents

Introduction

In this thesis we will study different boundary and final value problems for Kolmogorov operators arising from two dimensional stochastic differential equations (SDEs in the sequel) in which the second component represents the time integral of the first. More precisely:

$$\begin{cases} dX_t^1 = \sigma(t, X_t^1)dW_t, \\ dX_t^2 = X_t^1 dt, \end{cases} \tag{0.1}$$

where W is a real Brownian motion. Such setting is wide enough to accommodate various applications from different fields. For example, a particular case of (0.1) is the well-known Langevin equation from kinetic theory, which in simplified form reads

$$\begin{cases} \mathrm{d}X_t^1 = \mathrm{d}W_t \\ \mathrm{d}X_t^2 = X_t^1 \mathrm{d}t. \end{cases}$$

In financial applications, our main motivation for this study, the conditional expectation

$$u(t, x_1, x_2) = \mathbb{E}_{t,x_1,x_2}\Big[\varphi(X_{\tau\wedge T}^1, X_{\tau\wedge T}^2) - \int_t^{\tau\wedge T} f(X_s^1, X_s^2)ds\Big], \tag{0.2}$$

where τ is a stopping time, may represent the prices of a number of things suitably specifying φ and f. Among them, any Asian options (both European and American style) in any local volatility model. In this class, we will focus on European Asian options with fixed strike but we will also present new models for the value of a mine whose expression is again in the form above.

We choose to study the dynamic in (0.1) by the analytic point of view by means of the associated Kolmogorov backward operator $\mathcal{K}$. This is feasible as the problem to find the conditional expectations we are interested in is equivalent, via Feynman-Kac type theorems (see, e.g. Karatzas and Shreve (1991)), to a Cauchy-Dirichlet problems for $\mathcal{K}$, in the case the stopping time τ is the first exit time from a domain D.

For the dynamic in (0.1) we have

$$\mathcal{K} = \frac{1}{2}\sigma^2(t, x_1)\partial_{x_1,x_1} + x_1\partial_{x_2} + \partial_t, \qquad (t, x_1, x_2) \in \mathbb{R}^3,$$

which can be generalized to

$$\mathcal{K} = \frac{1}{2}\sum_{i,j=1}^{p_0} a_{i,j}(t, x)\partial_{x_i,x_j} + \langle Bx, \nabla\rangle + \partial_t, \qquad (t, x) \in \mathbb{R} \times \mathbb{R}^d, \tag{0.3}$$

where $p_0 \leq d$, B is a constant $(d \times d)$-matrix and $A_0 = (a_{i,j})_{i,j=1,\dots,p_0}$ is a symmetric and semi positive definite matrix. From now on, such operators will be referred to simply as *Komogorov operators* or *KO*.

Note that, if the matrix A_0 is constant and σ is a $(d \times p_0)$ matrix such that

$$\sigma\sigma^* = \begin{pmatrix} A_0 & 0 \\ 0 & 0 \end{pmatrix},$$

then the operator in (0.3) is the Kolmogorov backward operator associated to a d-dimensional process X satisfying

$$dX_t = BX_t\,dt + \sigma\,dW_t,$$

where W is a p_0-dimensional Brownian motion. The equation above is a *linear* SDE with constant coefficients. Under suitable conditions we will specify in later chapters, X admits an explicit and smooth transition density which is the fundamental solution of the operator in (0.3) as well.

This is a remarkable fact as many standard techniques used to study variable coefficients operators employ "frozen" version of them. Consider a standard Cauchy Problem (CP) for the a general Kolmogorov operator $\mathcal{K}$: replacing it with the constant coefficient one $\mathcal{K}^{(\bar{z})}$ obtained by freezing the second order part at a point $\bar{z} = (\bar{t}, \bar{x}) \in \mathbb{R} \times \mathbb{R}^d$, we are able to find a explicit solution to CP that, theoretically, should be close to the real one near $\bar{z}$.

It is then natural to ask if we can do better, that is: are we able to explicitly compute higher order approximation? Are they possible, under which conditions and how do they look like? Moreover, is it possible to rigorously prove bounds on the error? The first three chapters of the thesis are dedicated to the answers.

The idea is to mimic a procedure developed in Lorig et al. (2015) in the case of uniformly parabolic operators, a particular case of our setting as we shall see.

The approximation carried out in Lorig et al. (2015) was based on the Taylor expansion of the coefficients which were supposed to be Hölder continuous in the classical sense. However standard Euclidean regularity is not the best choice in our setting.

This particular feature is essentially due to a symmetry property of the fundamental solution of Kolmogorov Operators. Such symmetries, in the form of invariance with respect to particular translations, were known since the pioneering works Kolmogorov (1934) and Hörmander (1967) but the paper Lanconelli and Polidoro (1994) was the first proposing to use them in order to study the operators in the homogeneous Lie group framework.

In such groups the translations and intrinsic distance in use are quite different to the Euclidean ones and have to be accounted when we want to define what regular function means in this context. In fact, whereas the Euclidean geometry behave the same along different directions, more general Lie groups show anisotropic behaviours, usually according to the structure of the corresponding Lie algebra. The main example here is a sub Riemannian manifold: at any point the tangent space, that can be thought as the space of "directions" in which a curve passing trough the point can go, has a special subspace, the so called horizontal tangent space, in which the metric is strictly non degenerate and is thus a preferred choice. Thinking to the tangent space as the space of derivations, we can endow it with the structure of a stratified Lie algebra in which the firs layer V_1 is exactly the horizontal tangent space.

Roughly speaking, we will think to the Lie algebra of a Kolmogorov Group as the space of possible directions and assign a *formal* order to vector fields in the first layer. Any other vector field formal order would then be automatically determined as, by Hörmander condition, the i-th layer can be recovered commuting the first layer i times. Note that in this way some vector fields will have order greater than one. Then, a regular function of order n is expected to be Lie-differentiable along any vector field Z of formal order less or equal to n. The key idea here is not to *assume* smoothness along Z but to give enough regularity along the *first* layer and then prove Lie differentiability (regularity) along Z. The precise statement is given in Definition 1.14.

Of course, Euclidean regularity can be employed as it was in the seminal work Folland and Stein (1982) and its later improvement Bonfiglioli (2009). Essentially, the hypothesis $u \in C^n$ allows to write the n-th order intrinsic Taylor polynomial while the hypothesis $u \in C^{n+1}$ to prove an estimate for it but, as we shall see, such requirements can be significantly weakened on the case of Kolmogorov operators, using the novel spaces.

We prove that for functions in the intrinsic Hölder spaces of order n the corresponding intrinsic Taylor polynomial exists and investigate its form. This is carried out in three main parts: first we prove that any two points can be connected via integral curves of vector fields in $V_1 \subset \mathfrak{g}$ or concatenation of them. Then, we prove Theorem 2.20 in some special cases, namely when we move along such curves.Finally, we reduce the full statement to the particular cases treated above. It turns out the n-th order polynomial is expressible, in a rather compact way, purely in terms of composition of vector fields in the Lie algebra of formal degree less or equal to n and the group law and therefore it truly deserves the adjective intrinsic. Moreover, a Taylor type estimate of the same order as the space holds. For the precise statement see Theorem 2.20.

With both the right definition of regularity and the intrinsic Taylor formula in our hands, we turn our attention to applications. The first one we give is an analytic approximation expansion for the function in (0.2) with null f and $\tau \equiv T$. This corresponds to study European style Asian options. By Feynman-Kac theorem, the function u is the solution to the following Cauchy Problem

$$\begin{cases} \mathcal{K}u = 0, & \text{on } [0, T[\times D, \\ u(T, \cdot) = \varphi, & \text{on } D. \end{cases}$$

The idea is to replace operator $\mathcal{K}$ by operators $\mathcal{K}_n$ in which the coefficients of the second order part are replaced by their intrinsic Taylor polynomial of order n. Ultimately, we are able to prove the following short time estimate:

$$u(t,x) = u_0(t,x) + \sum_{n=1}^{N} \mathcal{L}_n(t,T,x)\, u_0(t,x) + \mathrm{O}\Big((T-t)^{\frac{N+1+k}{2}}\Big) \qquad \text{as } t \to T^-,$$

uniformly with respect to $x \in D$, where:

- the leading term u_0 is the solution of the Cauchy problem for $\mathcal{K}_0$ with final datum φ;

- $(\mathcal{L}_n)_{1 \le n \le N}$ is a family of differential operators, acting on x, that can be explicitly computed in terms of the intrinsic Taylor polynomials of the $a_{i,j}$ (see Theorem 3.55);

- the positive exponent k, contributing to the asymptotic rate of convergence, is the intrinsic Hölder exponent of φ.

We refer to Chapter 3 for the precise statements.

However, we will also study the case $\varphi = 0$, and τ the exit time from a domain D. In this case, we will prove that to find the function u in (0.2) is equivalent to solve

$$\begin{cases} \mathcal{K}u = f, & \text{on } [0,T[\times D, \\ u(T,\cdot) = 0, & \text{on } D, \\ u = 0 & \text{on }]0,T[\times \partial D. \end{cases}$$

This case corresponds to the mine valuation problem in a new model we propose. As we are not interested in approximating the function u analytically but numerically, we tackle the well poseness of the problem i.e. we provide existence and uniqueness results for the Cauchy Dirichlet problem above.

The plan of the thesis is the following: Chapter 1 is mainly introductory; in it we precisely introduce the class of operators we will study as well as the notion of Kolmogorov Lie group, furnishing motivating examples and comparisons. Later, we define novel intrinsic Hölder spaces of any order and compare them to the ones used in the literatures.

In Chapter 2 we extensively investigate the intrinsic Taylor polynomial. The core of the chapter is dedicated to the proof of the Taylor formula.At the end of the chapter we present an extension of the Taylor formula in the more general setting of non-homogeneous Kolmogorov groups.

In Chapter 3 we propose an analytical expansion for solutions to Cauchy problems for Kolmogorov Operators and provide short-time estimate for the error even in the case the operator degenerates outside of a compact domain.

Finally, in Chapter 4 we provide existence and uniqueness results for the value function of a mine as discussed above.

Many of the results presented here are taken from our articles Pagliarani et al. (2016), Pagliarani et al. (2017) (together with S. Pagliarani and A. Pascucci) and Pagliarani and Pignotti (2017) (together with S. Pagliarani). We deeply thank them all.

Chapter 1

Regularity in Kolmogorov groups

In this first chapter we introduce the notion of Kolmogorov Lie group using Kolmogorov operators as a motivating example. Later, we study its peculiar geometry introducing a suitable distance and proving some of its properties. Finally, we define intrinsic Hölder spaces of any order and compare them with the existing ones in the literature.

1.1 Constant coefficients Kolmogorov Operators

The constant coefficients differential operators of the form

$$\mathcal{K} := \frac{1}{2}\sum_{i,j=1}^{p_0} a_{i,j}\partial_{x_i,x_j} + \langle Bx, \nabla_x\rangle + \partial_t, \quad (t,x) \in \mathbb{R}\times\mathbb{R}^d, \quad p_0 \le d, \tag{1.1}$$

where $\langle\cdot,\cdot\rangle$ and $\nabla = (\partial_{x_1}, \dots, \partial_{x_d})$ denote the inner product and the gradient in $\mathbb{R}^d$ respectively, $A_0 = (a_{i,j})_{i,j=1,\dots,p_0}$ is a $p_0 \times p_0$ symmetric and semi positive definite matrix and B is a $d \times d$ matrix are known as Kolmogorov operators since the pioneering work Kolmogorov (1934).

Operators of the form (1.1) appear in several applications in physics, biology and mathematical finance. We recall that $\mathcal{K}$ is the linearized prototype of the Fokker-Planck operator arising in fluidodynamics (cf. Chandresekhar (1943)). Moreover $\mathcal{K}$ was extensively studied by Kolmogorov (1991) as the infinitesimal generator of the linear stochastic equation in $\mathbb{R}^d$

$$\mathrm{d}X_t = BX_t\mathrm{d}t + \sigma\mathrm{d}W_t, \tag{1.2}$$

where W is a p_0-dimensional standard Brownian motion and σ is a $d \times p_0$ matrix such that

$$\sigma\sigma^T = \begin{pmatrix} A_0 & 0 \\ 0 & 0 \end{pmatrix},$$

We also refer to Bossy et al. (2011) for a recent study of Navier-Stokes equations involving more general Kolmogorov-type operators.

In mathematical finance, Kolmogorov equations arise in models incorporating some sort of dependence on the past: typical examples are Asian options (see, for instance, Ingersoll (1987), Barucci et al. (2001), Pascucci (2008), Frentz et al. (2010)) and some volatility models (see, for instance, Hobson and Rogers (1998) and Foschi and Pascucci (2008)).

To shorten notation, let us consider the case $A_0 = I_{p_0}$ with I_{p_0} being the $p_0 \times p_0$ identity matrix. It is natural to place operator $\mathcal{K}$ in the framework of Hörmander's theory; indeed, let us set

$$X_j = \partial_{x_j}, \quad j = 1, \dots, p_0, \qquad \text{and} \qquad Y = \langle Bx, \nabla \rangle + \partial_t, \tag{1.3}$$

Then $\mathcal{K}$ can be written as a sum of vector fields:

$$\mathcal{K} = \frac{1}{2}\sum_{j=1}^{p_0} X_j^2 + Y.$$

Under the Hörmander's condition

$$\operatorname{rank}\left(\operatorname{Lie}(X_1, \dots, X_{p_0}, Y)\right) = d + 1, \tag{1.4}$$

operator $\mathcal{L}$ is hypoelliptic and Kolmogorov (1934) and Hörmander (1967) constructed an explicit fundamental solution of $\mathcal{K}u = 0$, which is the transition density of X in (1.2). We remark that X is a Gaussian process and condition (1.4) turns out to be equivalent to the non-degeneracy of the covariance matrix of X_t for any positive t (see, for instance, Karatzas and Shreve (1991) and ?? (Pas)).

Operator $\mathcal{K}$ in (1.1) is the prototype of the more general class of *Kolmogorov operators with variable coefficients.* The study of general Kolmogorov operators has been successfully carried out by several authors in the framework of the theory of homogeneous groups: Folland (1975), Folland and Stein (1982), Varopoulos et al. (1992) and Bonfiglioli et al. (2007) serve as a reference for the analysis of homogeneous groups. We recall that ?? (Pol) and Di Francesco and Pascucci (2005) proved the existence of a fundamental solution under optimal regularity assumptions on the coefficients; in particular, ?? (Pol) generalized and greatly improved the classical results by Weber (1951), Il′in (1964), Sonin (1967) and Gencev (1963) where unnecessary Euclidean-type regularity was required. Variable coefficients operators will be later addressed in Chapters 3, 4.

The intrinsic Lie group structure modeled on the vector fields $X_1, \dots, X_{p_0}, Y$ and the related non-Euclidean functional analysis (Hölder and Sobolev spaces) were studied by several authors, among others Polidoro and Ragusa (1998), Di Francesco and Polidoro (2006), Bramanti et al. (1996), Manfredini (1997), Lunardi (1997), Kunze et al. (2010), Nyström et al. (2010), Priola (2009) and Menozzi (2011).

1.2 Kolmogorov Lie groups

As first observed by Lanconelli and Polidoro (1994), operator $\mathcal{K}$ in (1.1) has the remarkable property of being invariant with respect to left translations in the group $\left(\mathbb{R} \times \mathbb{R}^d, \circ\right)$, where the non-commutative group law "$\circ$" is defined by

$$z \circ \zeta \equiv (t, x) \circ (s, \xi) = \left(t + s, e^{sB}x + \xi\right), \qquad z, \zeta \in \mathbb{R} \times \mathbb{R}^d. \tag{1.5}$$

Precisely, we have

$$(\mathcal{K}u^{(\zeta)})(z) = (\mathcal{K}u)(\zeta \circ z), \qquad z, \zeta \in \mathbb{R} \times \mathbb{R}^d,$$

where

$$u^{(\zeta)}(z) := u(\zeta \circ z).$$

Notice that in $(\mathbb{R} \times \mathbb{R}^d, \circ)$ the identity element is $\mathrm{Id} = (0, 0)$ while the inverse is given by $(t, x)^{-1} = \left(-t, -e^{-tB}x\right)$.

The translation above were suggested by the form of the fundamental solution of $\mathcal{K}$. Let us define

$$\Gamma(t, x) = \frac{1}{\sqrt{(2\pi)^d \det C(-t)}} e^{-\frac{1}{2}\langle C^{-1}(-t)x, x\rangle},$$

where $C(t)$ is a $d \times d$ matrix that, under Hörmander condition, is positive definite for every $t > 0$ (see equation (3.12) for its precise definition). Then, the fundamental solution with pole in $\zeta = (s, \xi)$ reads as

$$\Gamma(t, x; s, \xi) := \Gamma(s - t, x - e^{(s-t)B}\xi) = \Gamma(\zeta^{-1} \circ z), \qquad t < s.$$

In Lanconelli and Polidoro (1994), the authors proved that Hörmander's condition (1.4) is equivalent to the following assumption:

Assumption 1.1. There exists a basis in $\mathbb{R}^d$ in which the the matrix B has the block structure

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

where each B_j is a $p_j \times p_{j-1}$ matrix of rank p_j,

$$p_0 \geq p_1 \geq \cdots \geq p_r \geq 1, \qquad \sum_{j=0}^{r} p_j = d,$$

and the $*$-blocks are arbitrary.

Moreover, if (and only if) the $*$-blocks in (1.6) are null then $\mathcal{K}$ *is homogeneous of degree two* with respect the dilations $(D(\lambda))_{\lambda>0}$ on $\mathbb{R} \times \mathbb{R}^d$ given by

$$D(\lambda) = \mathrm{diag}\big(\lambda^2, \lambda I_{p_0}, \lambda^3 I_{p_1}, \cdots, \lambda^{2r+1} I_{p_r}\big), \tag{1.7}$$

where I_{p_j} are $p_j \times p_j$ identity matrices: specifically, we have

$$\big(\mathcal{K}u^{(\lambda)}\big)(t, x) = \lambda^2 (\mathcal{K}u)\big(D(\lambda)(t, x)\big), \qquad (t, x) \in \mathbb{R} \times \mathbb{R}^d,\ \lambda > 0,$$

where

$$u^{(\lambda)}(t, x) = u(D(\lambda)(t, x)).$$

For convenience, we also denote by

$$D_0(\lambda) = \mathrm{diag}\big(\lambda I_{p_0}, \lambda^3 I_{p_1}, \ldots, \lambda^{2r+1} I_{p_r}\big),$$

the spacial part of the dilations.

Historically, the homogeneous operators were the first to be studied being simpler than their non-homogeneous counterpart. Moreover, in this case it holds

$$D(\lambda)(z \circ \zeta) = (D(\lambda)z) \circ (D(\lambda)\zeta), \qquad \lambda > 0,\ z, \zeta \in \mathbb{R} \times \mathbb{R}^d.$$

Remark 1.2. The above formula implies that the dilations form a one parameter family of continuous automorphism or, in other words, the group

$$\mathcal{G}_B := \left(\mathbb{R} \times \mathbb{R}^d, \circ, D(\lambda)\right),$$

is homogeneous in the sense of Folland and Stein (1982). We stress that the group only depends on the matrix B.

From this point on, unless explicitly specified, we will work under the additional stronger assumption:

Assumption 1.3. The matrix B in (1.6) is supposed to have the $*$-block null, i.e. B takes the form

$$B = \begin{pmatrix} 0_{p_0 \times p_0} & 0_{p_0 \times p_1} & \cdots & 0_{p_0 \times p_{r-1}} & 0_{p_0 \times p_r} \\ B_1 & 0_{p_1 \times p_1} & \cdots & 0_{p_1 \times p_{r-1}} & 0_{p_1 \times p_r} \\ 0_{p_2 \times p_0} & B_2 & \cdots & 0_{p_2 \times p_{r-1}} & 0_{p_2 \times p_r} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0_{p_r \times p_0} & 0_{p_r \times p_1} & \cdots & B_r & 0_{p_r \times p_r} \end{pmatrix}, \tag{1.8}$$

where $0_{p_i \times p_j}$ is a $p_i \times p_j$ null block. In other words, the corresponding group $\mathcal{G}_B$ is supposed to be homogeneous.

To better understand the role the dilations play in studying the operator $\mathcal{K}$ we need to look at the Lie algebra generated by the vector fields $X_1, \ldots, X_{p_0}, Y$ which we know has dimension $d+1$ at every point of $\mathbb{R} \times \mathbb{R}^d$. As the vector fields $X_1, \ldots, X_{p_0}$ commute between themselves it is sufficient to characterize the iterated commutators $[\cdots[X_i, Y], \cdots Y]$. This is done in the general lemma below.

Lemma 1.4. *Let $v \in \mathbb{R}^d$ be a vector and u any smooth function on R^d. Then*

$$[\cdots[\nabla \cdot v, \underbrace{Y], \cdots, Y]}_{n \text{ times}}\, u = \langle B^n v, \nabla u \rangle, \qquad n \in \mathbb{N}_0.$$

Proof. We prove (2.8) by induction on n. For $n = 0$ the formula is trivially true. Now, supposing it holds for $n \in \mathbb{N}$ and recalling the definition of Y in (1.3), for $n+1$ we have

$$\begin{aligned}
[\cdots[\nabla \cdot v, \underbrace{Y], \cdots, Y]}_{n+1 \text{ times}} u =& [\langle B^n v, \nabla \rangle, \partial_t + \sum_{k,l=1}^{d} b_{k,l} x_l \partial_{x_k}] u \\
=& \sum_{i,j=1}^{d} B^n_{i,j} v_j \Big(\sum_{k=1}^{d} b_{k,i} \Big) \partial_{x_k} u \\
=& \langle B^{n+1} v, \nabla u \rangle.
\end{aligned}$$

□

It is therefore crucial to understand the behaviour of B's powers especially when the vector field $\nabla \cdot v$ is a linear combination of $X_1, \ldots, X_{p_0}$ that is $v_i = 0$ for $i > p_0$.

As a direct consequence of (1.8), we have that for any $n \leq r$

$$B^n = \begin{pmatrix}
0_{\bar{p}_{n-1} \times p_0} & 0_{\bar{p}_{n-1} \times p_1} & \cdots & 0_{\bar{p}_{n-1} \times p_{r-n}} & 0_{\bar{p}_{n-1} \times (\bar{p}_r - \bar{p}_{r-n})} \\
\prod_{j=1}^{n} B_j & 0_{p_n \times p_1} & \cdots & 0_{p_n \times p_{r-n}} & 0_{p_n \times (\bar{p}_r - \bar{p}_{r-n})} \\
0_{p_{n+1} \times p_0} & \prod_{j=2}^{n+1} B_j & \cdots & 0_{p_{n+1} \times p_{r-n}} & 0_{p_{n+1} \times (\bar{p}_r - \bar{p}_{r-n})} \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0_{p_r \times p_0} & 0_{p_r \times p_1} & \cdots & \prod_{j=r-n+1}^{r} B_j & 0_{p_r \times (\bar{p}_r - \bar{p}_{r-n})}
\end{pmatrix}, \qquad (1.9)$$

with

$$\bar{p}_i = p_0 + p_1 + \cdots + p_i, \qquad 0 \leq i \leq r,$$

$\bar{p}_{-1} \equiv 0$ and where

$$\prod_{j=1}^{n} B_j = B_n B_{n-1} \cdots B_1.$$

Moreover $B^n = 0$ for $n > r$, so that

$$e^{\delta B} = \sum_{h=0}^{r} \frac{B^h}{h!} \delta^h. \qquad (1.10)$$

Due to the block structure of the matrix B in (1.8) it is very convenient to split $\mathbb{R}^d$ accordingly. Precisely, let π_n be the projection

$$\pi_n : \mathbb{R}^d \longrightarrow \{0\}^{\bar{p}_{n-1}} \times \mathbb{R}^{p_n} \times \{0\}^{d - \bar{p}_n} \qquad n = 0, \ldots, r,$$

and denote by V_n its image. Then we have

$$\mathbb{R}^d = \bigoplus_{n=0}^{r} V_n, \qquad \dim V_n = p_n, \quad n = 0, \dots, r. \tag{1.11}$$

Definition 1.5. Due to this splitting, which is going to be used extensively, it will be convenient to have a short notation to denote the image of π_n. Let us set

$$x^{[n]} := (x_{\bar{p}_{n-1}}, \dots, x_{\bar{p}_n}) \in \mathbb{R}^{p_n}, \qquad x \in \mathbb{R}^d,$$

or equivalently,

$$\pi_n(x) = (\underbrace{0, \dots, 0}_{\bar{p}_{n-1}}, x^{[n]}, \underbrace{0, \dots, 0}_{d-\bar{p}_n}).$$

Moreover, we will refer to a variable x_i such that $\bar{p}_n < i \le \bar{p}_{n+1}$ as a variable of *level* n, $n = 0, \dots, r$.

By (1.9) it is clear that

$$B^n v \in \bigoplus_{k=n}^{r} V_k, \qquad v \in \mathbb{R}^d, \tag{1.12}$$

and if $v \in V_0$ then

$$B^n v \in V_n, \qquad n = 0, \dots, r. \tag{1.13}$$

More precisely, let us set

$$\bar{B}_n = \begin{pmatrix} 0_{\bar{p}_{n-1} \times p_0} & 0_{\bar{p}_{n-1} \times (r-p_0)} \\ \prod_{j=1}^{n} B_j & 0_{p_n \times (r-p_0)} \\ 0_{(\bar{p}_r - \bar{p}_n) \times p_0} & 0_{(\bar{p}_r - \bar{p}_n) \times (r-p_0)} \end{pmatrix},$$

where the $p_n \times p_0$ matrix $\prod_{j=1}^{n} B_j$ has full rank. Then we have

$$B^n v = \bar{B}_n v, \qquad v \in V_0,$$

and the linear application $\bar{B}_n : V_0 \to V_n$ is surjective but, in general, not injective. For this reason, for any $n = 1, \cdots, r$, we define the subspaces $V_{0,n} \subseteq V_0$ as

$$V_{0,n} = \{x \in V_0 | x_j = 0 \ \forall j \notin \Pi_{B,n}\},$$

with $\Pi_{B,n}$ being the set of the indexes corresponding to the first p_n linear independent columns of $\prod_{j=1}^{n} B_j$. It is now trivial that the linear map

$$\bar{B}_n : V_{0,n} \to V_n$$

is also injective. Notice that

$$V_{0,r} \subseteq V_{0,r-1} \subseteq \cdots \subseteq V_{0,1} \subseteq V_{0,0} := V_0. \tag{1.14}$$

Remark 1.6. Equation (1.13) together with the surjectivity of the linear maps $\bar{B}_n$ imply that *any* spatial derivative ∂_{x_i} can be expressed as iterated commutators of Y and a linear combination of $X_1, \dots, X_{p_0}$. In fact, given an index $1 \leq i \leq d$, say such that $\bar{p}_n < i \leq \bar{p}_{n+1}$, there exist an unique vector $v \in V_{0,n} \subset V_0$ for which it holds $B^n v = e_i$, e_i being the i-th vector of the canonic basis. Plugging v in (2.8) we obtain

$$[\cdots[\nabla \cdot v, \underbrace{Y], \cdots, Y]}_{n \text{ times}} = \langle B^n v, \nabla \rangle = \langle e_i, \nabla \rangle = \partial_{x_i}. \tag{1.15}$$

By the above remark, under Assumption 1.3 the decomposition in (1.11) can be translated in the Lie algebra $\mathfrak{g}$ as follow:

$$\mathfrak{g} = \underbrace{\mathrm{span}\{Y\}}_{=:U_1} \oplus \underbrace{\mathrm{span}\{X_1, \dots, X_{p_0}\}}_{=:U_2} \oplus \underbrace{[U_1, U_2]}_{=:U_3} \oplus \cdots \oplus \underbrace{[U_1, U_r]}_{=:U_{r+1}}. \tag{1.16}$$

where each of the U_j, $2 \leq j \leq r+1$, is isomorphic to V_{j-1}. Moreover, formula (1.16) defines a *gradation* i.e. it holds $[U_i, U_j] \subset U_{i+j}$ for every $i, j \in \mathbb{N}$ (setting $U_i = 0$ for $i > r+1$).

The resemblance of (1.16) with (1.7) which we repeat here below is strikingly

$$D(\lambda) = \mathrm{diag}\big(\lambda^2, \lambda I_{p_0}, \lambda^3 I_{p_1}, \cdots, \lambda^{2r+1} I_{p_r}\big),$$

but expected as both are a consequence of the block structure of B.

As the vector fields $X_1, \dots, X_{p_0}$ and Y are $D(\lambda)$-homogeneous of degree one and two respectively, it follows that any partial derivative ∂_{x_i} obtained as in (1.15) commuting $X_1, \dots, X_{p_0}$ with Y n times should be $D(\lambda)$-homogeneous of degree $1 + 2 + \cdots + 2 = 2n + 1$ and the block structure in (1.7) follows.

As it is customary in the heat operator framework, we regard to the time derivative, here generalized by Y, as a formally second order operator. Moreover, given a variable x_i of level n, it is natural by equations (1.15) and (1.16) to assign to ∂_{x_i} an intrinsic degree of $2n + 1$.

Naturally, also the distance should reflect the anisotropic behaviour of the dilations $D(\lambda)$ and so we look for a homogeneous norm ρ, that is a continuous function $\rho : \mathbb{R} \times \mathbb{R}^d \to [0, +\infty)$ such that $\rho(z) = 0$ iff $z = 0$ and

$$\rho(D(\lambda)z) = \lambda\rho(z), \qquad z \in \mathbb{R} \times \mathbb{R}^d, \quad \lambda > 0.$$

In Folland and Stein (1982) the norm is also required to be smooth out of the origin. As we shall not need this property, it will be convenient the choice

$$\|(t,x)\|_B := |t|^{1/2} + [x]_B, \qquad [x]_B := \sum_{j=1}^{d} |x_j|^{1/\sigma_j}, \qquad (t,x) \in \mathbb{R} \times \mathbb{R}^d, \tag{1.17}$$

where $(\sigma_j)_{1\le j\le d}$ are the integers such that

$$D_0(\lambda) = \mathrm{diag}\big(\lambda^{\sigma_1}, \dots, \lambda^{\sigma_d}\big), \tag{1.18}$$

that is $\sigma_1 = \cdots = \sigma_{p_0} = 1$, $\sigma_{p_0+1} = \cdots = \sigma_{p_0+p_1} = 3$ and so forth.

Definition 1.7. Let $\beta = (\beta_1, \cdots, \beta_d) \in \mathbb{N}_0^d$ denote any multi-index. As usual

$$|\beta| := \sum_{j=1}^{d} \beta_j \quad \text{and} \quad \beta! := \prod_{j=1}^{d} (\beta_j!)$$

are called the length and the factorial of β respectively. Moreover, for any $x \in \mathbb{R}^d$, we set

$$x^\beta = x_1^{\beta_1} \cdots x_d^{\beta_d} \quad \text{and} \quad \partial^\beta = \partial_x^\beta = \partial_{x_1}^{\beta_1} \cdots \partial_{x_d}^{\beta_d}.$$

Accordingly to the dilations $D_0(\lambda)$ we also define the intrinsic length of β as

$$|\beta|_B := \sum_{i=0}^{r} (2i+1)\big|\beta^{[i]}\big|$$

where $\beta^{[i]} \in \mathbb{N}_0^d$ is the multi-index

$$\beta_k^{[i]} := \begin{cases} \beta_k & \text{for } \bar{p}_{i-1} < k \le \bar{p}_i, \\ 0 & \text{otherwise,} \end{cases} \tag{1.19}$$

coherently with Definition 1.5.

We conclude this section with two lemmas regarding the homogeneous norm $\|\cdot\|_B$.

Lemma 1.8. *There exist two constants $C_1 \geq 1$ and $C_2 > 0$, both depending only on B, such that*

$$\begin{aligned} &\|\zeta \circ z\|_B \leq C_1 \left(\|\zeta \circ \eta\|_B + \left\|\eta^{-1} \circ z\right\|_B\right), && z, \zeta, \eta \in \mathbb{R} \times \mathbb{R}^d, \\ &\frac{1}{C_2}|z-\zeta| \leq \left\|\zeta^{-1} \circ z\right\|_B \leq C_2 |z-\zeta|^{\frac{1}{2r+1}}, && \textit{for } |z-\zeta|, \left\|\zeta^{-1} \circ z\right\|_B \leq 1. \end{aligned}$$

The first inequality implies that $\|\cdot\|_B$ is a quasi-norm, while the second formula shows that the intrinsic distance is locally equivalent to the Euclidean one. For a proof we refer to Manfredini (1997), Proposition 2.1.

In the case the matrix B assumes the more general form (1.6) one is lead to still use the same norm but, unfortunately, the relations above are no longer true globally. Instead, the following lemma holds true

Lemma 1.9. *Suppose that the matrix B as in* (1.6)*. Then, for every positive constant T and compact set $H \subset \mathbb{R}^d$ there exist a constant $C_{T,H} \geq 1$, depending also on B, such that*

$$\begin{aligned} &\left\|z^{-1}\right\|_B \leq C_{T,H} \|z\|_B, && z \in [-T, T] \times H, \\ &\|\zeta \circ z\|_B \leq C_{T,H}\big(\|\zeta\|_B + \|z\|_B \big), && z \in [-T, T] \times \mathbb{R}^d, \;\; \zeta \in \mathbb{R} \times H. \end{aligned}$$

A proof can be found in Di Francesco and Polidoro (2006), Lemma 2.1.

1.3 Intrinsic Hölder spaces

Next we introduce the notions of intrinsic regularity and Hölder space. Let X be a Lipschitz vector field on $\mathbb{R} \times \mathbb{R}^d$. For any $z \in \mathbb{R} \times \mathbb{R}^d$, we denote by $\delta \mapsto e^{\delta X}(z)$ the integral curve of X defined as the unique solution of

$$\begin{cases} \frac{d}{d\delta} e^{\delta X}(z) = X\left(e^{\delta X}(z)\right), & \delta \in \mathbb{R}, \\ e^{\delta X}(z)|_{\delta=0} = z. \end{cases}$$

Explicitly, if $X \in \{X_1, \cdots, X_{p_0}, Y\}$ is one of the vector fields in (1.3), we have

$$e^{\delta X_i}(t,x) = (t, x + \delta e_i), \quad i = 1, \cdots, p_0, \qquad e^{\delta Y}(t,x) = (t+\delta, e^{\delta B} x), \quad (1.20)$$

for any $(t, x) \in \mathbb{R} \times \mathbb{R}^d$.

Next we recall the general notion of Lie differentiability and Hölder regularity.

Definition 1.10. Let X be a Lipschitz vector field and u be a real-valued function defined in a neighborhood of $z \in \mathbb{R} \times \mathbb{R}^d$. We say that u is *X-differentiable* in z if the function $\delta \mapsto u\left(e^{\delta X}(z)\right)$ is differentiable in 0. We will refer to the function $z \mapsto \frac{d}{d\delta} u\left(e^{\delta X}(z)\right)\big|_{\delta=0}$ as *X-Lie derivative of* u, or simply *Lie derivative of* u when the dependence on the field X is clear from the context.

As explained in the previous section, it is natural to connect the structure of the dilations $D(\lambda)$ to the vector fields in $\mathfrak{g}$ associating to them their order of $D(\lambda)$-homogeneity.

Assumption 1.11. To any vector field $X \in \{X_1, \cdots, X_{p_0}, Y\}$, we associate a formal degree $m_X > 0$ in the following canonical way: $m_{X_j} = 1$ for $1 \le j \le p_0$ and $m_Y = 2$.

Definition 1.12. Let X be a Lipschitz vector field on $\mathbb{R} \times \mathbb{R}^d$ with formal degree $m_X > 0$. For $\alpha \in]0, m_X]$, we say that $u \in C_X^\alpha$ if the semi-norm

$$\|u\|_{C_X^\alpha} := \sup_{\substack{z \in \mathbb{R} \times \mathbb{R}^d \\ \delta \in \mathbb{R} \setminus \{0\}}} \frac{\left|u\left(e^{\delta X}(z)\right) - u(z)\right|}{|\delta|^{\frac{\alpha}{m_X}}}$$

is finite.

Now, let Ω be a domain in $\mathbb{R} \times \mathbb{R}^d$. For any $z \in \Omega$ we set

$$\delta_z = \sup\left\{\bar{\delta} \in]0,1] \mid e^{\delta X}(z) \in \Omega \text{ for any } \delta \in [-\bar{\delta}, \bar{\delta}]\right\}.$$

If Ω_0 is a bounded domain with $\overline{\Omega}_0 \subseteq \Omega$, we set

$$\delta_{\Omega_0} = \min_{z \in \overline{\Omega}_0} \delta_z.$$

Note that $\delta_{\Omega_0} \in]0,1]$.

Definition 1.13. For $\alpha \in]0, m_X]$, we say that $u \in C^\alpha_{X,\text{loc}}(\Omega)$ if for any bounded domain Ω_0 with $\overline{\Omega}_0 \subseteq \Omega$, the semi-norm

$$\|u\|_{C_X^\alpha(\Omega_0)} := \sup_{\substack{z \in \Omega_0 \\ 0 < |\delta| < \delta_{\Omega_0}}} \frac{\left|u\left(e^{\delta X}(z)\right) - u(z)\right|}{|\delta|^{\frac{\alpha}{m_X}}}$$

is finite.

Now we define the intrinsic Hölder spaces on the homogeneous group $\mathcal{G}_B$.

Definition 1.14. Let $\alpha \in]0,1]$, then:

i) $u \in C_B^{0,\alpha}$ if $u \in C_Y^\alpha$ and $u \in C_{\partial_{x_i}}^\alpha$ for any $i = 1, \dots, p_0$. For any $u \in C_B^{0,\alpha}$ we define the semi-norm

$$\|u\|_{C_B^{0,\alpha}} := \|u\|_{C_Y^\alpha} + \sum_{i=1}^{p_0} \|u\|_{C_{\partial_{x_i}}^\alpha} .$$

ii) $u \in C_B^{1,\alpha}$ if $u \in C_Y^{1+\alpha}$ and $\partial_{x_i} u \in C_B^{0,\alpha}$ for any $i = 1, \dots, p_0$. For any $u \in C_B^{1,\alpha}$ we define the semi-norm

$$\|u\|_{C_B^{1,\alpha}} := \|u\|_{C_Y^{\alpha+1}} + \sum_{i=1}^{p_0} \|\partial_{x_i} u\|_{C_B^{0,\alpha}} .$$

iii) For $k \in \mathbb{N}$ with $k \geq 2$, $u \in C_B^{k,\alpha}$ if $Yu \in C_B^{k-2,\alpha}$ and $\partial_{x_i} u \in C_B^{k-1,\alpha}$ for any $i = 1, \dots, p_0$. For any $u \in C_B^{k,\alpha}$ we define the semi-norm

$$\|u\|_{C_B^{k,\alpha}} := \|Yu\|_{C_B^{k-2,\alpha}} + \sum_{i=1}^{p_0} \|\partial_{x_i} u\|_{C_B^{k-1,\alpha}} .$$

Similarly, according to Definition 1.13, we define the spaces $C_{B,\mathrm{loc}}^{k,\alpha}(\Omega)$ of *locally* Hölder continuous functions on a domain Ω of $\mathbb{R} \times \mathbb{R}^d$, and the related semi-norms $\|\cdot\|_{C_B^{k,\alpha}(\Omega_0)}$ on bounded domains Ω_0 with $\overline{\Omega}_0 \subseteq \Omega$.

Remark 1.15. The following inclusion holds: $C_{B,\mathrm{loc}}^{k,\alpha} \subseteq C_{B,\mathrm{loc}}^{k',\alpha'}$ for $0 \leq k' \leq k$ and $0 < \alpha' \leq \alpha \leq 1$. Moreover we have $C_B^{k,\alpha} \subseteq C_{B,\mathrm{loc}}^{k,\alpha}$ for $k \geq 0$.

1.3.1 Intrinsic Hölder spaces in the literature

Intrinsic Hölder spaces play a central role in the study of the existence and the regularity properties of solutions to Kolmogorov operators with variables coefficients. Slightly different notions of Hölder spaces have been proposed by several authors (see, for instance, Manfredini (1997), Lunardi (1997), Pascucci (2003), Di Francesco and Polidoro (2006) and Frentz et al. (2010)): note that some authors introduce only the definition of $C_B^{0,\alpha}$ and $C_B^{2,\alpha}$. Indeed, the definition of $C_B^{1,\alpha}$ is technically more elaborate because it involves derivatives of fractional (in the intrinsic sense) order and therefore is sometimes omitted.

In Manfredini (1997) and Di Francesco and Polidoro (2006), $C_B^{0,\alpha}$ is defined as the space of functions that are bounded and Hölder continuous with respect to the homogeneous group structure: precisely, a function $u \in C_B^{0,\alpha}$ on a domain Ω of $\mathbb{R} \times \mathbb{R}^d$ if

$$|u|_{\alpha,\Omega} := \sup_{z\in\Omega} |u(z)| + \sup_{\substack{z,\zeta\in\Omega \\ z\neq\zeta}} \frac{|u(z)-u(\zeta)|}{\|\zeta^{-1}\circ z\|_B^{\alpha}} < \infty. \tag{1.21}$$

It can be easily seen that this definition implies ours except for the L^∞ control. Corollary 2.22 in Chapter 2 shows that definition (1.21) is equivalent to Definition 1.14-i). Similarly, Frentz et al. (2010) define the following norm in the space $C_B^{1,\alpha}$:

$$\begin{aligned} |u|_{1+\alpha,\Omega} :=& |u|_{\alpha,\Omega} + \sum_{i=1}^{p_0} |\partial_{x_i} u|_{\alpha,\Omega} + \\ & \sup_{\substack{z,\zeta\in\Omega \\ z\neq\zeta}} \frac{|u(z)-u(\zeta)-\sum_{i=1}^{p_0}\partial_{x_i}u(\zeta)(x-\xi)_i|}{\|\zeta^{-1}\circ z\|_B^{1+\alpha}}, \end{aligned}$$

where $z=(t,x)$, $\zeta=(s,\xi)\in\mathbb{R}\times\mathbb{R}^d$. As we shall see, with this definition, the Taylor theorem 2.20 trivially follows.

Various definitions of the space $C_B^{2,\alpha}(\Omega)$ are used in the literature. Manfredini (1997) requires bounded and Hölder continuous second order derivatives, while Di Francesco and Polidoro (2006) and Frentz et al. (2010) also require the function u and its first p_0 spatial derivatives to be Hölder continuous. Precisely, Manfredini (1997) introduces the norm

$$|u|_{2+\alpha,\Omega}^{(M)} := \sup_{\Omega} |u| + \sum_{i=1}^{p_0} \sup_{\Omega} |\partial_{x_i} u| + \sum_{i,j=1}^{p_0} |\partial_{x_i,x_j} u|_{\alpha,\Omega} + |Yu|_{\alpha,\Omega}, \tag{1.22}$$

while Di Francesco and Polidoro (2006) and Frentz et al. (2010) define

$$|u|_{2+\alpha,\Omega} := |u|_{\alpha,\Omega} + \sum_{i=1}^{p_0} |\partial_{x_i} u|_{\alpha,\Omega} + \sum_{i,j=1}^{p_0} |\partial_{x_i,x_j} u|_{\alpha,\Omega} + |Yu|_{\alpha,\Omega}.$$

In light of the main result of this paper, the Taylor formula in Theorem 2.20, the notion of Hölder spaces in Definition 1.14 turns out to be optimal in the sense that it is given under more explicit and less restrictive assumptions compared to the literature.

In obtaining Shauder type estimates for Kolmogorov operators on bounded domains, it is very common to use weighted version of the spaces above (see e.g. Manfredini (1997), Di Francesco and Polidoro (2006)). More precisely, for any $z,\zeta\in\Omega$ they set

$$d_{z,\zeta} = \min\{d_z, d_\zeta\}, \qquad d_z = \inf_{w\in\partial\Omega} \left\|w^{-1}\circ z\right\|_B.$$

Then, for $m \in \{0, 2\}$ the following norm is used

$$|u|_{m+\alpha,d,\Omega} := \sup_{z\in\Omega} d_z^m |u(z)| + \sup_{\substack{z,\zeta\in\Omega \\ z\neq\zeta}} d_{z,\zeta}^{m+\alpha} \frac{|u(z) - u(\zeta)|}{\|\zeta^{-1}\circ z\|_B^\alpha}.$$

Note the similarity with (1.21). Those norms, although the index m may be misleading, are zero-th order norms. The space $C^{0,\alpha}_{B,d}(\Omega)$ is defined as the space of function u such that $|u|_{\alpha,d,\Omega} < \infty$. Note that under this requirement also the norm $|u|_{2+\alpha,d,\Omega}$ is finite. A second order norm should allow to control also first and second derivatives to some extend. In fact, the weighted space $C^{2,\alpha}_{B,d}$ is then defined by finiteness of the norm

$$\begin{aligned}|u|_{2,\alpha,d,\Omega} :=& \sup_{z\in\Omega}|u(z)| + \sup_{\substack{z,\zeta\in\Omega \\ z\neq\zeta}} d_{z,\zeta}^{2+\alpha}\frac{|u(z)-u(\zeta)|}{\|\zeta^{-1}\circ z\|_B^\alpha} + \sum_{i=1}^{p_0}\sup_{z\in\Omega} d_z|\partial_{x_i}u| + \\ & \sum_{i=1}^{p_0} d_{z,\zeta}^{2+\alpha}\frac{|\partial_{x_i}u(z) - \partial_{x_i}u(\zeta)|}{\|\zeta^{-1}\circ z\|_B^\alpha} + \sum_{i,j=1}^{p_0}|\partial_{x_i,x_j}u|_{2+\alpha,d,\Omega} + \\ & |Yu|_{2+\alpha,d,\Omega}.\end{aligned}$$

We will make use of the norm in the last paragraph in Chapter 4.

1.3.2 Examples of intrinsically regular functions

For comparison, we give some examples of functions with different intrinsic and Euclidean regularity in the simplest case. We set $d = 2$ and consider the prototype Kolmogorov operator

$$\mathcal{K} = \frac{1}{2}\partial_{x,x} + x\partial_y + \partial_t, \qquad (t, x, y) \in \mathbb{R}\times\mathbb{R}^2. \tag{1.23}$$

Corresponding to

$$B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}.$$

Example 1.16. Consider the function $u : \mathbb{R}\times\mathbb{R}^2 \longrightarrow \mathbb{R}$ given by $u(t, x_1, x_2) = |x_2 - c|$, with $c \in \mathbb{R}$. As we shall see in Chapter 3, this function is particularly relevant for financial applications since it is often related to the payoff of so-called Asian-style derivatives. Clearly u is Lipschitz continuous in the Euclidean sense, but intrinsically we have $u \in C^{1,1}_{B,\text{loc}}(\mathbb{R}\times\mathbb{R}^2)$ because $\partial_{x_1}u \in C^{0,1}_{B,\text{loc}}(\mathbb{R}\times\mathbb{R}^2)$ and $u \in C^2_{Y,\text{loc}}(\mathbb{R}\times\mathbb{R}^2)$. Note that $u \notin C^{2,\alpha}_{B,\text{loc}}(\mathbb{R}\times\mathbb{R}^2)$ because u is not Y-differentiable in $x_2 = c$: nevertheless a (2.3)-like estimate for $n = 2$ and $\alpha = 1$

holds for two points $z, \zeta \in \mathbb{R} \times \mathbb{R}^2$ sharing the same time-component, i.e.

$$|u(z) - u(\zeta)| \leq |x_2 - \xi_2| \leq \|\zeta^{-1} \circ z\|_B^3, \qquad z = (t, x),\ \zeta = (t, \xi) \in \mathbb{R} \times \mathbb{R}^2.$$

This is an instance of a more general phenomenon that we shall study in Remark 3.48

Example 1.17. As a variant of the previous example let us consider the function $u : \mathbb{R} \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ given by $u(t, x_1, x_2) = |x_2 - c|^{\frac{3}{2}}$, with $c \in \mathbb{R}$. This time $u \in C^{1,1/2}$, that is differentiable with Hölder continuous derivatives in the Euclidean sense, but intrinsically we have $u \in C^{2,1}_{B,\mathrm{loc}}(\mathbb{R} \times \mathbb{R}^2)$ because $\partial_{x_1} u \equiv 0$ and

$$Yu(t, x_1, x_2) = \frac{3}{2} x_1 \, |x_2 - c|^{\frac{1}{2}} \,\mathrm{sgn}(x_2 - c) \in C^{0,1}_{B,\mathrm{loc}}.$$

Also in the present example the function shows higher intrinsic regularity than the Euclidean one.

Example 1.18. It is easy to check that any function of the form $u(t, x_1, x_2) = f(x_2 - t x_1)$ is constant along the integral curves $e^{\delta Y}(z) = (t + \delta, x_1, x_2 + \delta x_1)$ for any $z \in \Omega$. Therefore, we have $Y^n u \equiv 0$ for any $n \in \mathbb{N}$. In this particular case, we have that $u \in C^{n,\alpha}_{B,\mathrm{loc}}$ if and only if $u \in C^{n,\alpha}_{\mathrm{loc}}$ in the Euclidean sense.

Example 1.19. The following function belongs to $C^{2,\alpha}_{B,\mathrm{loc}}$ but only to $C^{0,\alpha}_{\mathrm{loc}}$:

$$u(t, x_1, x_2) = \begin{cases} \frac{1}{\sqrt{2\pi x_1^4}} \int_{\mathbb{R}} \exp\left(- \frac{(y - x_2)^2}{2 x_1^4} \right) |y| \,\mathrm{d}y & \text{if } x_1 \neq 0, \\ |x_2| & \text{if } x_1 = 0. \end{cases}$$

Indeed u is continuous and smooth on $\{x_1 \neq 0\}$; in particular, $u \in C^{2,1}_{\mathrm{loc}}(\{x_1 \neq 0\})$ and $u \in C^{2,1}_{B,\mathrm{loc}}(\{x_1 \neq 0\})$. On the plane $\{x_1 = 0\}$ the Euclidean derivative $\partial_{x_2} u$ does not exist in $x_2 = 0$ for any t and thus $u \notin C^{2,\alpha}_{\mathrm{loc}}$ for any $\alpha \in (0, 1]$. On the other hand, $\partial_{x_1} u$, $\partial_{x_1 x_1} u$ and Yu exist on $\{x_1 = 0\}$ and they are all equal to 0. In particular, we have $\partial_{x_1 x_1} u, Yu \in C^1_{Y,\mathrm{loc}}$ and $\partial_{x_1} u \in C^2_{Y,\mathrm{loc}}$. Moreover, one can directly prove that $\partial_{x_1 x_1} u, Yu \in C^1_{\partial_{x_1},\mathrm{loc}}$ and thus, $u \in C^{2,1}_{B,\mathrm{loc}}$.

Chapter 2

Intrinsic Taylor formula

In this chapter we prove an intrinsic Taylor formula for Hölder regular functions on the homogeneous group $\mathcal{G}_B$ previously defined in 1.14. After some notations we state the main theorem followed by some corollaries. The proof is divided in a preliminary part and the proper proof. In the former we collect results on how to approximate the integrals curves of the higher order vector fields in the gradation (1.16) and how they interact with intrinsic Hölder functions. In the latter we prove the main theorem by induction. Finally, we extend some of the results to the non homogeneous case.

When dealing with intrinsic Hölder spaces, Taylor-type formulas (and the related estimates for the remainder) form one of the cornerstones for the development of the theory. Classical results about intrinsic Taylor polynomials on homogeneous groups were proved in great generality by Folland and Stein (1982). Recently, Bonfiglioli (2009) (see also Arena et al. (2010)) derived explicit formulas for Taylor polynomials on homogeneous groups and the corresponding remainders by adapting the classical Taylor formula with integral remainder.

Here we give a new and more explicit representation of the intrinsic Taylor polynomials for Kolmogorov-type groups. The distinguished features of our formulas are as follows:

i) in Folland and Stein (1982) and Bonfiglioli (2009), Taylor polynomials of order n are defined for functions that are differentiable up to order n in the Euclidean sense; the constants in the error estimates for the remainders (that is, the differences between the function and its Taylor polynomials) depend on the norms of the function in the Euclidean Hölder spaces. Conversely, in this paper we define n-th order Taylor polynomials for functions that are regular *in the intrinsic sense* and the constants appearing in the error estimates *depend only on the norms of the intrinsic derivatives up to order* n. At the best of our knowledge, a similar result under such intrinsic regularity assumptions only appeared in Arena et al. (2010), but limited to the particular case of the Heisenberg group. Moreover, the fact that we assume intrinsic regularity on the function, as opposed to Euclidean one, allows us to yield some global error bounds for the remainders when the function belongs to the *intrinsic global* Hölder spaces. This represents another key difference with respect to the existing literature, where such bounds are only local;

ii) since the vector fields $X_1, \dots, X_{p_0}$ do not commute with Y, there are different representations for the Taylor polynomials depending on the order of the derivatives: specifically, the representation in Folland and Stein (1982) and Bonfiglioli (2009) is given as a sum over all possible permutations of the derivatives. Thus, computing explicitly the n-th order Taylor polynomials can be very lengthy since the number of terms involved grows proportionally to d^n. On the contrary, even though our Taylor polynomials are algebraically equivalent to those given by Folland and Stein (1982) and Bonfiglioli (2009), as the Taylor polynomial is unique, in Theorem 2.20 we determine a privileged way to order the vector fields so that we are able to get compact Taylor polynomials with a number of terms increasing linearly with respect to the order of the polynomial itself (see (2.4) below); this is quite relevant for practical computations, as we will show through a simple example in Section 2.1 and in Chapter 3.

Recall that $|\beta|_B$ denotes the intrinsic length of a multi index $\beta \in \mathbb{N}_0^d$ as defined in 1.7. We next state the Taylor formula in its global version.

Theorem 2.20. *Let $\alpha \in]0,1]$ and $n \in \mathbb{N}_0$. If $u \in C_B^{n,\alpha}$ then we have:*

1) there exist the Lie derivatives

$$Y^k \partial_x^\beta u, \qquad 0 \le 2k + |\beta|_B \le n; \tag{2.1}$$

2) they lie in the spaces

$$Y^k \partial_x^\beta u \in C_B^{n-2k-|\beta|_B,\alpha} \qquad \text{for } \ 0 \le 2k + |\beta|_B \le n, \tag{2.2}$$

and

$$\left|u(z) - \mathcal{T}_n(u,\zeta)(z)\right| \leq c_B \|u\|_{C^{n,\alpha}_B} \|\zeta^{-1} \circ z\|_B^{n+\alpha}, \qquad z, \zeta \in \mathbb{R} \times \mathbb{R}^d, \quad (2.3)$$

where c_B *is a constant that depends on* B*, while* $\mathcal{T}_n(u,\zeta)(z)$ *is the* n-th order intrinsic Taylor polynomial of u around $\zeta = (s,\xi)$*, calculated in* $z = (t,x) \in \mathbb{R} \times \mathbb{R}^d$*, defined as*

$$\mathcal{T}_n(u,\zeta)(z) := \sum_{0 \leq 2k + |\beta|_B \leq n} \frac{1}{k!\,\beta!} \left(Y^k \partial_\xi^\beta u(s,\xi)\right)(t-s)^k \left(x - e^{(t-s)B}\xi\right)^\beta. \tag{2.4}$$

Remark 2.21. The local version of this theorem also holds true. Let Ω be a domain in $\mathbb{R} \times \mathbb{R}^d$ and $u \in C^{n,\alpha}_{B,\mathrm{loc}}(\Omega)$. Then for any $\zeta \in \Omega$ there exist a neighbourhood V such that $\overline{V} \subseteq \Omega$ and

$$|u(z) - \mathcal{T}_n(u,\zeta)(z)| \leq c_{B,\zeta} \|u\|_{C^{n,\alpha}_{B,\mathrm{loc}}(V)} \|\zeta^{-1} \circ z\|_B^{n+\alpha}, \qquad z \in V.$$

A proof can be found in our work Pagliarani et al. (2016) but here we shall not follow it as we will obtain the same result under a much weaker assumption. Namely, we will discard the Assumption 1.3 but maintain the hypoellepticity condition (1.6). This is done in section 2.4.

A direct consequence of estimate (2.3) in the particular case $n = 0$ is the following

Corollary 2.22. *A function* $u \in C^{0,\alpha}_B$ *if and only if there exists a positive constant* c *such that*

$$|u(z) - u(\zeta)| \leq c \left\|\zeta^{-1} \circ z\right\|_B^\alpha, \qquad z, \zeta \in \mathbb{R} \times \mathbb{R}^d,$$

i.e. u *is* B*-Hölder continuous in the sense of Definition 1.2 in ?? (Pol).*

For a comparison between intrinsic and Euclidean Hölder continuity we refer to Proposition 2.1 in ?? (Pol).

We stress that the derivatives Y^k in (2.1) are meant in the Lie sense. However, if the function is regular enough they are equivalent to classical Euclidean derivatives. In fact, the next result can also be seen as the embedding $C^{2r+1,\alpha}_B \subset C^1$.

Corollary 2.23. *If* $u \in C^{2r+1,\alpha}_B$*, then there exists* $\partial_t u \in C^{0,\alpha}_{B,\mathrm{loc}}$*. Moreover, we have*

$$\partial_t u(t,x) = Y u(t,x) - \langle Bx, \nabla u(t,x) \rangle. \tag{2.5}$$

Proof. In Theorem 2.20 we take $\zeta = (t, x)$, $z = (t + \delta, x)$ and note that, in this case, the spatial increments become

$$x - e^{\delta B}x = -\delta Bx + O(\delta^2) \qquad \text{as } \delta \to 0.$$

Now, by Theorem 2.20 all the spatial first-order derivatives exist and

$$\begin{aligned} u(z) - &\mathcal{T}_{2r+1}(u, \zeta)(z) \\ &= u(t + \delta, x) - u(t, x) - \delta Y u(t, x) + \delta \sum_{i=1}^{d} \partial_{x_i} u(t, x)(Bx)_i + O(\delta^2), \end{aligned}$$

as $\delta \to 0$. Since

$$\|\zeta^{-1} \circ z\|_B^{2r+1+\alpha} = \|(\delta, x - e^{\delta B}x)\|_B^{2r+1+\alpha} = O(|\delta|^{1+\frac{\alpha}{2r+1}}), \qquad \text{as } \delta \to 0,$$

we get

$$\frac{u(t + \delta, x) - u(t, x)}{\delta} - Y u(t, x) + \sum_{i=1}^{d} (Bx)_i \partial_{x_i} u(t, x) = O(|\delta|^{\frac{\alpha}{2r+1}}) \qquad \text{as } \delta \to 0.$$

This implies that the time-derivative exists and formula (2.5) holds. Now, it also easily follows that $\partial_t u \in C^{0,\alpha}_{B,\mathrm{loc}}$ since , by the inclusions in Remark 1.15, all the derivatives appearing in the right-hand side of (2.5) are in $C^{0,\alpha}_{B,\mathrm{loc}}$. □

Note that, as it apparent from the proof, the result holds also under the weaker assumption $u \in C^{2r+1,\alpha}_{B,\mathrm{loc}}$.

2.1 Comparison with known results and examples

In the more general theory of homogeneous Lie groups developed in Folland and Stein (1982) the Taylor polynomials are expressed in terms of left invariant vector fields which form a basis for the Lie algebra. As such groups are automatically nilpotent the exponential map Exp between the Lie algebra $\mathfrak{g}$ and the corresponding Lie group $\mathcal{G}$ is a global diffeomorphism whose inverse is denoted by Log. We can therefore identify $\mathcal{G}$ with a Lie group $(\mathbb{R}^N, *)$ as we will from now on.

We suppose then that the abstract dilations $D(\lambda)$ take the form

$$D(\lambda)(x_1, \ldots, x_N) = (\lambda^{\sigma_1} x_1, \ldots, \lambda^{\sigma_N} x_N), \qquad x \in \mathbb{R}^N,$$

where $1 = \sigma_1 \leq \cdots \leq \sigma_N$ are positive real numbers. On such a group there is a privileged basis for the Lie algebra $\mathfrak{g}$, the Jacobian one, whose elements are the left-invariant vector fields Z_i uniquely defined by

$$Z_i|_{x=0} \equiv \partial_{x_i} \qquad i = 1, \dots, N.$$

As can be proved, In this framework it is natural to define the intrinsic degree of Z_i as σ_i and the D_λ-homogeneous norm

$$|x|_{\mathbb{G}} = \sum_{i=1}^{N} |x_i|^{\frac{1}{\sigma_i}}.$$

Following Bonfiglioli (2009), the n-th order intrinsic Taylor polynomial $P_n f(x_0, \cdot)$ of a function f around the point x_0, can be defined as the unique polynomial function such that

$$f(x) - P_n(f, x_0)(x) = O(|x_0^{-1} * x|_{\mathbb{G}}^{n+\varepsilon}) \qquad \text{as } |x_0^{-1} * x|_{\mathbb{G}} \to 0,$$

for some $\varepsilon > 0$. For $f \in C^{n+1}$ existence and uniqueness of $P_n f$ was proved in Folland and Stein (1982); under the same hypothesis, a more explicit expression and a better estimate of the remainder was given in Bonfiglioli (2009). Precisely, in the latter the author proved that

$$P_n(f, x_0)(x) = f(x_0) + \sum_{k=1}^{n} \sum_{\substack{1 \leq i_1, \dots, i_k \leq N \\ I = (i_1, \dots, i_k),\ \sigma(I) \leq n}} \frac{Z_I f(x_0)}{k!} \mathrm{Log}_{i_1}(x_0^{-1} * x) \cdots \mathrm{Log}_{i_k}(x_0^{-1} * x). \tag{2.6}$$

Here $\sigma(I) := i_1 \sigma_{i_1} + \cdots + i_k \sigma_{i_k}$ denotes the intrinsic order of the operator $Z_I := Z_{i_1} \cdots Z_{i_k}$ and Log_i is the i-th component of the Log map in the basis $\{Z_1, \dots, Z_N\}$.

Note that, in general, operators Z_i do not commute. Therefore, formula (2.6) typically involves a large number of terms. In the special case of a Kolmogorov-type group, the Taylor polynomial (2.4) is much more compact that (2.6) because we can exploit the fact that all but one of the Z_i coincide with Euclidean derivatives and thus commute with each other; moreover, our increments along the integral curves of the vector fields are different from those in (2.6). We illustrate this fact in the following example.

Let us consider the simplest Kolmogorov group, namely the one induced by the operator defined in (1.23). This case corresponds to the matrix

$$B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

and the dilations $D(\lambda)$ take the following explicit form:

$$D(\lambda)(t, x_1, x_2) = (\lambda^2 t, \lambda x_1, \lambda^3 x_2), \qquad (t, x_1, x_2) \in \mathbb{R} \times \mathbb{R}^2.$$

Moreover, if $z = (t, x_1, x_2), \zeta = (s, \xi_1, \xi_2)$, then we also have

$$\zeta \circ z = (s+t, x_1+\xi_1, x_2+\xi_2+t\xi_1), \qquad \zeta^{-1} \circ z = (t-s, x_1-\xi_1, x_2-\xi_2-(t-s)\xi_1).$$

The components of left-hand side vector in the previous formula are exactly the increments appearing in (2.4). With regard to formula (2.6), we have

$$Z_0 = Y, \qquad Z_1 = \partial_{x_1}, \qquad Z_2 = \partial_{x_2},$$

while the corresponding components of the Log map are

$$\begin{aligned} &\mathrm{Log}_0(\zeta^{-1} \circ z) = t - s, \qquad \mathrm{Log}_1(\zeta^{-1} \circ z) = x_1 - \xi_1, \\ &\mathrm{Log}_2(\zeta^{-1} \circ z) = x_2 - \xi_2 - (t-s)\xi_1 - \frac{(t-s)(x_1-\xi_1)}{2}. \end{aligned}$$

Note that the first two components coincide with the increments mentioned above while the third one is different. It follows that, up to order two, the two versions of the Taylor polynomial coincide. They are

$$\begin{aligned} &\mathcal{T}_0(u, \zeta)(z) = u(\zeta), \\ &\mathcal{T}_1(u, \zeta)(z) = u(\zeta) + \partial_x u(\zeta)(x_1 - \xi_1). \end{aligned}$$

On the other hand, according to our definition, the third and fourth polynomials are given by

$$\begin{aligned} \mathcal{T}_3(u, \zeta)(z) &= \mathcal{T}_2(u, \zeta)(z) + \frac{1}{3!}\partial_{x_1}^3 u(\zeta)(x_1 - \xi_1)^3 + Y\partial_{x_1} u(\zeta)(x_1 - \xi_1)(t-s) \\ &\quad + \partial_{x_2} u(\zeta)(x_2 - \xi_2 - (t-s)\xi_1), \\ \mathcal{T}_4(u, \zeta)(z) &= \mathcal{T}_3(u, \zeta)(z) + \frac{1}{4!}\partial_{x_1}^4 u(\zeta)(x_1 - \xi_1)^4 + \frac{1}{2!}Y\partial_{x_1}^2 u(\zeta)(x_1 - \xi_1)^2(t-s) \\ &\quad + \frac{1}{2!}Y^2 u(\zeta)(t-s)^2 + \partial_{x_2}\partial_{x_1} u(\zeta)(x_1 - \xi_1)(x_2 - \xi_2 - (t-s)\xi_1), \end{aligned}$$

while, according to formula (2.6), we have

$$\begin{aligned}
\mathcal{T}_3(u,\zeta)(z) &= \mathcal{T}_2(u,\zeta)(z) + \frac{1}{2!}(Y\partial_{x_1} + \partial_{x_1}Y)u(\zeta)(x_1-\xi_1)(t-s)\\
&\quad + \frac{1}{3!}\partial_{x_1}^3 u(\zeta)(x_1-\xi_1)^3\\
&\quad + \partial_{x_2}u(\zeta)\big(x_2-\xi_2-(t-s)\xi_1 - \frac{(t-s)(x_1-\xi_1)}{2}\big),\\
\mathcal{T}_4(u,\zeta)(z) &= \mathcal{T}_3(u,\zeta)(z) + \frac{1}{2!}Y^2u(\zeta)(t-s)^2 + \frac{1}{4!}\partial_{x_1}^4 u(\zeta)(x_1-\xi_1)^4\\
&\quad + \frac{1}{3!}(Y\partial_{x_1}^2 + \partial_{x_1}Y\partial_{x_1} + \partial_{x_1}^2 Y)u(\zeta)(x_1-\xi_1)^2(t-s)\\
&\quad + \partial_{x_2}\partial_{x_1}u(\zeta)(x_1-\xi_1)\big(x_2-\xi_2-(t-s)\xi_1 - \frac{(t-s)(x_1-\xi_1)}{2}\big).
\end{aligned}$$

Notice that the above expressions of the Taylor polynomials can be proved to be algebraically equivalent by using the identity $\partial_{x_1}Y = Y\partial_{x_1} + \partial_{x_2}$.

Regarding the type of estimates the use of intrinsic regularity leads to two other remarkable properties that are not present in the work of Bonfiglioli (2009). The first one is that the estimate of order n only depends on the matrix B and on the norm of the function in $C^{n,\alpha}_{B,\mathrm{loc}}$ which, we recall, just depends on the derivatives up to *intrinsic order* n. This in contrast with the bound given in Bonfiglioli (2009) which depends on all the (Euclidean) derivatives up to order $n+1$. The second feature is the possibility to give global estimates of the remainder. This is not possible in Bonfiglioli (2009) due to the presence in the bound of different powers of the intrinsic distance which are not asymptotically equivalent. Instead, our approach produces homogeneous estimates in terms of the distance of the same degree of the approximation.

2.2 Commutators and integral paths

In this section we construct approximations of the integral paths of the commutators of the vector fields $X_1,\dots,X_{p_0}$ and Y in (1.3). In the sequel we shall use the following notations: for any $v\in\mathbb{R}^d$ we set

$$Y_v^{(0)} = \sum_{i=1}^d v_i\partial_{x_i}.$$

Hereafter we will always consider $v\in V_0$. In such way $Y_v^{(0)}$ will be actually a linear combination of $X_1,\dots,X_{p_0}$. Moreover we define recursively

$$Y_v^{(n)} = [Y_v^{(n-1)}, Y] = Y_v^{(n-1)}Y - YY_v^{(n-1)}, \qquad n\in\mathbb{N}. \tag{2.7}$$

Remark 2.24. By Lemma 2.25 it holds for any $u \in C^\infty(\mathbb{R} \times \mathbb{R}^d)$

$$Y_v^{(n)} u = \langle B^n v, \nabla u \rangle, \qquad n \in \mathbb{N}, \tag{2.8}$$

with $B^n v \in V_n$ by (1.13).

When applied to functions in $C^{n,\alpha}_{B,\mathrm{loc}}$, operator $Y_v^{(n)}$ can be interpreted as a composition of Lie derivatives. Indeed we have the following.

Lemma 2.25. *Let $n \in \mathbb{N}$ and $u \in C^{n,\alpha}_{B,loc}$. Then, for any $v \in V_0$ and $k \in \mathbb{N} \cup \{0\}$ with $2k+1 \le n$, we have $Y_v^{(k)} u \in C^{n-2k-1,\alpha}_{B,loc}$.*

Proof. If $k = 0$, the thesis is obvious since, by assumption, $\partial_{x_i} u \in C^{n-1,\alpha}_{B,\mathrm{loc}}$ for $i = 1, \dots, p_0$. To prove the general case we proceed by induction on n. If $n \le 2$ there is nothing to prove because we only have to consider the case $k = 0$. Fix now $n \ge 2$. We assume the thesis to hold for any $m \le n$ and prove it true for $n+1$. We proceed by induction on k. We have already shown the case $k = 0$. Thus, we assume the statement to hold for $k \in \mathbb{N} \cup \{0\}$ with $2(k+1)+1 \le n+1$ and we prove it true for $k+1$. Note that, by definition (2.7) we clearly have

$$Y_v^{(k+1)} u = Y_v^{(k)} Y u - Y\, Y_v^{(k)} u,$$

with $v \in V_0$. Then the thesis follows by inductive hypothesis and since, by definition of intrinsic Hölder space, $Yu \in C^{n-1,\alpha}_{B,\mathrm{loc}}$. □

Next we show how to approximate the integral curves of the commutators $Y_v^{(k)}$ by using a rather classical technique from control theory. For any $n \in \{0, \dots, r\}$, $z = (t,x) \in \mathbb{R} \times \mathbb{R}^d$, $\delta \in \mathbb{R}$ and $v \in V_0$, we define iteratively the family of trajectories $\left(\gamma^{(n,k)}_{v,\delta}(z)\right)_{k=n,\dots,r}$ as

$$\gamma^{(n,n)}_{v,\delta}(z) = e^{\delta^{2n+1} Y_v^{(n)}}(z) = \big(t, x + \delta^{2n+1} B^n v\big), \tag{2.9}$$

$$\gamma^{(n,k+1)}_{v,\delta}(z) = e^{-\delta^2 Y}\left(\gamma^{(n,k)}_{v,-\delta}\left(e^{\delta^2 Y}\left(\gamma^{(n,k)}_{v,\delta}(z)\right)\right)\right), \tag{2.10}$$

for $n \le k \le r-1$. We also set

$$\gamma^{(-1,k)}_{v,\delta}(z) = \gamma^{(0,k)}_{v,\delta}(z), \qquad 0 \le k \le r. \tag{2.11}$$

Lemma 2.26. *For any $n \in \{0, \cdots, r\}$, $(t,x) \in \mathbb{R} \times \mathbb{R}^d$, $\delta \in \mathbb{R}$ and $v \in V_0$ we have*

$$\gamma^{(n,k)}_{v,\delta}(t,x) = (t, x + S_{n,k}(\delta) v), \qquad k = n, \dots, r, \tag{2.12}$$

where

$$S_{n,n}(\delta) = \delta^{2n+1}B^n \qquad \textit{and} \qquad S_{n,k}(\delta) = (-1)^{k-n}\delta^{2n+1}B^n \sum_{\substack{h\in\mathbb{N}^{k-n}\\ |h|\le r}} \frac{(-B)^{|h|}}{h!}\delta^{2|h|}, \tag{2.13}$$

for $k = n+1, \dots, r$, *and with* $|h| = h_1 + \dots + h_k$.

Proof. Fix $n = 0$ and proceed by induction on k. The case $k = n$ is trivial. Now, assuming (2.12)-(2.13) as inductive hypothesis and noting that $S_k(-\delta) = -S_k(\delta)$, we have

$$\begin{aligned}
\gamma_{v,\delta}^{(k+1)}(t,x) &= e^{-\delta^2 Y}\big(\gamma_{v,-\delta}^{(n,k)}\big(e^{\delta^2 Y}\big(\gamma_{v,\delta}^{(n,k)}(t,x)\big)\big)\big)\\
&= e^{-\delta^2 Y}\big(\gamma_{v,-\delta}^{(n,k)}\big(e^{\delta^2 Y}\big(t, x + S_{n,k}(\delta)v\big)\big)\big)\\
&= e^{-\delta^2 Y}\big(\gamma_{v,-\delta}^{(n,k)}\big(t+\delta^2, e^{\delta^2 B}\big(x + S_{n,k}(\delta)v\big)\big)\big)\\
&= e^{-\delta^2 Y}\big(t+\delta^2, e^{\delta^2 B}\big(x + S_{n,k}(\delta)v\big) - S_{n,k}(\delta)v\big)\\
&= \big(t, e^{-\delta^2 B}\big(e^{\delta^2 B}\big(x + S_{n,k}(\delta)v\big) - S_{n,k}(\delta)v\big)\big)\\
&= \big(t, x + S_{n,k}(\delta)v - e^{-\delta^2 B}S_{n,k}(\delta)v\big).
\end{aligned}$$

On the other hand, by (1.10) we have

$$\begin{aligned}
x + S_{n,k}(\delta)v - e^{-\delta^2 B}S_{n,k}(\delta)v &= x + S_{n,k}(\delta)v - \left(\sum_{j=0}^{r}\frac{(-B)^j}{j!}\delta^{2j}\right)S_{n,k}(\delta)v\\
&= x - \left(\sum_{j=1}^{r}\frac{(-B)^j}{j!}\delta^{2j}\right)S_{n,k}(\delta)v\\
&= x + S_{n,k+1}(\delta)v,
\end{aligned}$$

and this concludes the proof. $\square$

Remark 2.27. Note that

$$S_{n,k}(\delta) = \delta^{2k+1}B^k + \widetilde{S}_{n,k}(\delta), \qquad n \le k \le r,$$

with

$$\widetilde{S}_{n,n}(\delta) := 0 \quad \text{and} \quad \widetilde{S}_{n,k}(\delta) := (-1)^{k-n}\delta^{2n+1}B^n \sum_{\substack{h\in\mathbb{N}^{k-n}\\ k-n<|h|\le r}} \frac{(-B)^{|h|}}{h!}\delta^{2|h|},$$

if $n+1 \le k \le r$. Then we deduce from (2.12) that

$$\gamma_{v,\delta}^{(n,k)}(z) = \big(t, x + \delta^{2k+1}B^k v\big) + \big(0, \widetilde{S}_{n,k}(\delta)v\big), \qquad n \le k. \tag{2.14}$$

It is important to remark that $\widetilde{S}_{n,r}(\delta) = 0$ and, by (1.12), we have

$$\widetilde{S}_{n,k}(\delta)v \in \bigoplus_{j=k+1}^{r} V_j, \qquad k = n, \dots, r; \tag{2.15}$$

since $v \in V_0$, then by (1.13) we have

$$\gamma_{v,\delta}^{(n,n)}(z) = (t,x) + (0, \delta^{2n+1} B^n v), \quad \text{with } B^n v \in V_n.$$

Thus, by using notation (1.19), for any $k = n, \dots, r$ we have

$$\left| \left(\widetilde{S}_{n,k}(\delta)v \right)^{[j]} \right| \le c_B |\delta|^{2j+1} |v|, \qquad j = k+1, \dots, r, \quad \delta \in \mathbb{R}, \tag{2.16}$$

where the constant c_B *depends only on the matrix* B. If $|v| = 1$, (2.16) also implies

$$\begin{aligned}
\left\| \left(\gamma_{v,\delta}^{(n,k)}(z) \right)^{-1} \circ z \right\|_B &= \left\| z^{-1} \circ \gamma_{v,\delta}^{(n,k)}(z) \right\|_B \\
&= \left\| \left((t, x + \delta^{2k+1} B^k v) + (0, \widetilde{S}_{n,k}(\delta)v) \right)^{-1} \circ (t,x) \right\|_B \\
&= \left\| \left(0, -\delta^{2k+1} B^k v - \widetilde{S}_{n,k}(\delta)v \right) \right\|_B \\
&= [-\delta^{2k+1} B^k v - \widetilde{S}_{n,k}(\delta)v]_B \le c_B |\delta|.
\end{aligned} \tag{2.17}$$

Next we show how to connect two points in $\mathbb{R} \times \mathbb{R}^d$ that only differ w.r.t. the spatial components by only moving along the the integral curves $\gamma^{(n,k)}$ previously defined.

Lemma 2.28. *Let* $n \in \{0, \cdots, r\}$, $\zeta = (t, \xi) \in \mathbb{R} \times \mathbb{R}^d$, $y \in \bigoplus_{k=n}^{r} V_k$ *and the points* $\zeta_k = (t, \xi_k)$, *for* $k = n-1, \cdots, r$, *defined as*

$$\zeta_{n-1} := \zeta, \qquad \zeta_k := \gamma_{v_k,\delta_k}^{(n-1,k)}(\zeta_{k-1}), \quad v_k = \frac{w_k}{|w_k|}, \quad \delta_k = |w_k|, \quad k \ge n,$$

where w_k *is the only vector in* $V_{0,k} \subseteq V_0$ *such that* $B^k w_k = y^{[k]} + \xi^{[k]} - \xi_{k-1}^{[k]}$. *Then:*

i) for any $k \in \{n, \cdots, r\}$ *we have:*

$$\delta_k \le c_B [y]_B, \qquad \xi_k^{[j]} = \xi^{[j]} + y^{[j]}, \quad j = 0, \dots, k. \tag{2.18}$$

Note that, in particular, $\zeta_r = \zeta + (0, y)$;

ii) there exists a positive constant c_B, only dependent on the matrix B, *such that*

$$\left\|\zeta_k^{-1}\circ\zeta\right\|_B \le c_B[y]_B,$$

for any $k = n, \cdots, r$ *and* $0 \le \delta \le \delta_k$.

Proof. We first prove i). The second identity in (2.18) easily stems from (2.15) and by definition of v_k and δ_k. We then focus on the first one. By Remark (1.13) along with the expression of $\gamma_{v,\delta}^{(n,k)}$ in (2.14), it is easy to prove that

$$\delta_k \le c_B\left|\xi^{[k]} + y^{[k]} - \xi_{k-1}^{[k]}\right|^{\frac{1}{2k+1}}. \tag{2.19}$$

Moreover, by (2.16) we get

$$\left|\xi_k^{[j]} - \xi_{k-1}^{[j]}\right| \le c_B\left|\delta_k\right|^{2j+1}, \qquad j = k+1, \dots, r. \tag{2.20}$$

We proceed by induction on k. For $k = n$ the thesis immediately follows by (2.19). We now fix $n \le k \le r-1$ and assume the estimate to hold for any $n \le h \le k$. By (2.19) we have

$$\begin{aligned}
\delta_{k+1} &\le c_B\left|\xi^{[k+1]} + y^{[k+1]} - \xi_k^{[k+1]}\right|^{\frac{1}{2(k+1)+1}} \\
&\le c_B\left|y^{[k+1]}\right|^{\frac{1}{2(k+1)+1}} + c_B\sum_{h=n}^{k}\left|\xi_h^{[k+1]} - \xi_{h-1}^{[k+1]}\right|^{\frac{1}{2(k+1)+1}} \\
&\le c_B\left|y^{[k+1]}\right|^{\frac{1}{2(k+1)+1}} + c_B\sum_{h=n}^{k}\delta_h,
\end{aligned}$$

where we used (2.20). The thesis for $k+1$ now follows by inductive hypothesis.

We now prove ii). As first step we prove that

$$\left\|\left(\gamma_{v_k,\delta}^{(n-1,k)}(\zeta_{k-1})\right)^{-1}\circ\zeta_{k-1}\right\|_B \le c_B[y]_B.$$

By equations (2.12), (2.17) and (2.18) we get

$$\begin{aligned}
\left\|\left(\gamma_{v_k,\delta}^{(n-1,k)}(\zeta_{k-1})\right)^{-1}\circ\zeta_{k-1}\right\|_B &= \left\|(t, \xi_{k-1} + S_{n-1,k}(\delta)v_k)^{-1}\circ(t,\xi_{k-1})\right\|_B \\
&= \left\|(0, \xi_{k-1} - (\xi_{k-1} + S_{n-1,k}(\delta)v_k))\right\|_B \\
&= \left\|(0, -S_{n-1,k}(\delta)v_k)\right\|_B \le c_B\delta_k \le c_B[y]_B.
\end{aligned}$$

This estimate along with equations (2.17) and (2.18) allow us to conclude. Precisely, applying the quasi-triangular inequality we get

$$\left\|\zeta_k^{-1}\circ\zeta\right\|_B \le c_B\sum_{i=n}^{k}\left\|\zeta_i^{-1}\circ\zeta_{i-1}\right\|_B \le c_B[y]_B.$$

□

Remark 2.29. Let $n \in \mathbb{N}_0$, $m \in \{0,1\}$ and $u \in C_B^{2n+m,\alpha}$. Then, by Definition 1.14, we have $Y^n u \in C_Y^{m+\alpha}$. Therefore, by the Euclidean mean-value theorem along the vector field Y, for any z and $\delta \in \mathbb{R}$, there exists $\bar{\delta}$ with $|\bar{\delta}| \leq |\delta|$ such that

$$u\big(e^{\delta Y}(z)\big) - u(z) - \sum_{i=1}^{n} \frac{\delta^i}{i!} Y^i u(z) = \delta^n \Big(Y^n u\big(e^{\bar{\delta} Y}(z)\big) - Y^n u(z) \Big),$$

and thus, by Definition 1.12 along with Assumption 1.11,

$$\Big| u\big(e^{\delta Y}(z)\big) - u(z) - \sum_{i=1}^{n} \frac{\delta^i}{i!} Y^i u(z) \Big| \leq \|u\|_{C_B^{2n+m,\alpha}} |\delta|^{n+\frac{m+\alpha}{2}}, \qquad \delta \in \mathbb{R}.$$

We also have a control on the homogeneous distance between points connected by integral curves of Y:

Remark 2.30. By (1.20) and (1.5) we have

$$\big\| z^{-1} \circ e^{\delta Y}(z) \big\|_B = \big\| (e^{\delta Y}(z))^{-1} \circ z \big\|_B = |\delta|^{\frac{1}{2}}, \qquad z \in \mathbb{R} \times \mathbb{R}^d, \quad \delta \in \mathbb{R}. \tag{2.21}$$

We conclude the section with the table below which clarifies the strategy needed to connect two points $z = (t,x)$ and $\zeta = (s,\xi)$, progressively correcting the levels, using the integral curves defined so far

z	$e^{\delta Y}$	$\gamma^{(0,0)}$	$\gamma^{(0,1)}$	$\cdots$	$\gamma^{(0,r)}$	ζ
t	1			$\cdots$		s
$x^{[0]}$		2		$\cdots$		$\xi^{[0]}$
$x^{[1]}$			3	$\cdots$		$\xi^{[1]}$
$\vdots$				$\ddots$		$\vdots$
$x^{[r]}$				$\cdots$	$r+2$	$\xi^{[r]}$

As we can see $r+2$ steps are needed in general (less if temporal and low order spatial variables already coincide). We first correct the time variable using the integral curve $e^{\delta Y}$ with $\delta = s-t$. Thanks to (2.21) the distance is controlled by $|s-t|^{\frac{1}{2}} \leq \big\|\zeta^{-1} \circ z\big\|_B$. This changes also the space components but using the curves $\gamma^{(n,k)}$ as in Lemma 2.28 they can be, one level each time, progressively corrected to ξ in a controlled way.

2.3 Proof of the intrinsic Taylor formula

Theorem 2.20 will be proved by induction on n, through the following steps:

- **Step 1**: Proof for $n = 0$;
- **Step 2**: Induction from $2n$ to $2n + 1$ for any $0 \leq n \leq r$;
- **Step 3**: Induction from $2n + 1$ to $2(n + 1)$ for any $0 \leq n \leq r - 1$;
- **Step 4**: Induction from n to $n + 1$ for any $n \geq 2r + 1$.

A brief explanation is needed: the proof of Theorem 2.20 cannot be carried out by a simple induction on n, due to the qualitative differences in the Taylor polynomials of different orders. For instance, one could suppose the theorem to hold for $n = 2$ and consider a function $u \in C^{3,\alpha}_{B,\text{loc}}$. By the inclusion property

$$C^{3,\alpha}_{B,\text{loc}} \subseteq C^{2,\alpha}_{B,\text{loc}},$$

all the derivatives of second B-order do exist, i.e.

$$Y^k \partial_x^\beta u \in C^{2-2k-|\beta|_B,\alpha}_{B,\text{loc}}, \qquad 2k + |\beta|_B \leq 2.$$

However, $\mathcal{T}_3 u$ also contains the derivatives of intrinsic order equal to 3. These are exactly

$$\partial_{x_i,x_j,x_k} u, \quad Y\partial_{x_i} u \qquad 1 \leq i,j,k \leq p_0,$$

whose existence is granted by definition of $C^{3,\alpha}_{B,\text{loc}}$, and the Euclidean derivatives

$$\partial_{x_l} u, \qquad p_0 < l \leq \bar{p}_1,$$

whose existence must be proved, as it is not trivially implied by definition of $C^{3,\alpha}_{B,\text{loc}}$. In general, such problem arises every time when defining the Taylor expansion of order $2n + 1$, $n = 1, \ldots, r$, i.e. when the Euclidean derivatives w.r.t. the variables of level n appear for the first time in the Taylor polynomial. This motivates the need to treat the inductive step from $2n$ to $2n + 1$ in a separate way and therefore the necessity for Step 2 and Step 3 in the proof. Eventually, Step 4 is justified by the fact that, when $n \geq 2r + 1$, the existence of the Euclidean partial derivatives w.r.t. any variable has already been proved and thus the proof goes smoothly without any further complication.

We now try to summarize the main arguments on which the proof is based. Roughly speaking, in order to prove the estimate (2.3) we shall be able to connect any pair of points $z, \zeta \in \mathbb{R} \times \mathbb{R}^d$ and to have a control of the increment of u along the connecting path. The definition of $C^{n,\alpha}_{B,\text{loc}}$ (and $C^{n,\alpha}_B$) does only specify the regularity along the fields Y and $(\partial_{x_i})_{1 \le i \le p_0}$, but does not give any a priori information about the regularity along all the other Euclidean fields $(\partial_{x_i})_{p_0 < i \le d}$. It seems then clear that, when trying to connect z and ζ, we cannot simply *move* along the canonical directions $(e_i)_{1 \le i \le d}$. We shall indeed take advantage of Lemma 2.28 in order to go from ζ to z by using the integral curves $\gamma^{(n,k)}$ and then control the increment of u along the connecting paths by exploiting the estimates contained in Remark 2.27.

To easy notations, through this chapter, z and ζ will respectively denote the generic points $(t,x),\ (s,\xi) \in \mathbb{R} \times \mathbb{R}^d$ unless differently specified.

In order to prove the main theorem we will need to state three auxiliary results, which will be proved step by step along with Theorem 2.20.

Proposition 2.31. *Let $u \in C^{2n+1,\alpha}_B$ with $\alpha \in]0,1]$ and $n \in \mathbb{N}_0$ with $n \le r$. Then, there exist the Euclidean partial derivatives $\partial_{x_i} u \in C^{0,\alpha}_B$ for any $\bar{p}_{n-1} < i \le \bar{p}_n$ and*

$$Y^{(n)}_{v_i^{(n)}} u(z) = \partial_{x_i} u(z), \tag{2.22}$$

with $\big(v_i^{(n)}\big)_{\bar{p}_{n-1} < i \le \bar{p}_n}$ being the family of vectors such that $v_i^{(n)} \in V_{0,n}$ with $B^n v_i^{(n)} = e_i$. Note that such family of vectors is univocally defined (see Remark 1.6).

Proposition 2.32. *Let $\alpha \in]0,1]$, $n \in \mathbb{N}_0$ with $n \le r$, $m \in \{0,1\}$ and $u \in C^{2n+m,\alpha}_B$. Then, for any $\max\{n-1,0\} \le k \le r$ and $v \in V_{0,k}$ with $|v| = 1$, we have:*

$$\left| u\big(\gamma^{(n-1,k)}_{v,\delta}(z)\big) - \mathcal{T}_{2n+m}(u,z)\big(\gamma^{(n-1,k)}_{v,\delta}(z)\big) \right| \le c_B \|u\|_{C^{2n+m,\alpha}_B} |\delta|^{2n+m+\alpha}, \quad \delta \in \mathbb{R}, \tag{2.23}$$

where c_B is a positive constant that only depends on B.

Proposition 2.33. *Let $\alpha \in]0,1]$, $n \in \mathbb{N}_0$ with $n \le r$, $m \in \{0,1\}$ and $u \in C^{2n+m,\alpha}_B$. Then, we have:*

$$|u(t,x) - \mathcal{T}_{2n+m}(u,(t,x))(t,x+\xi)| \le c_B \|u\|_{C^{2n+m,\alpha}_B} [\xi]_B^{2n+m+\alpha}, \quad \xi \in \bigoplus_{j=0}^{n-1} V_j,$$

where c_B is a positive constant that only depends on B.

Propositions 2.32 and 2.33 are particular cases of the main theorem and are preparatory to its proof.

2.3.1 Step 1

Here we give the proofs for

- Proposition 2.32 for $n = 0$, $m = 0$;
- Theorem 2.20 (Part 2) for $n = 0$.

We start by recalling that:

$$\mathcal{T}_0(u, z)(\zeta) = u(z).$$

Proof of Propostion 2.32 for $n = 0$, $m = 0$.
We prove the thesis by induction on k. For $k = 0$ the estimate (2.23) trivially follows by combining definitions (2.11) and (2.9) with the assumptions $v \in V_0$, $|v| = 1$ and $u \in C^{\alpha}_{\partial_{x_i}}$ for any $i = 1, \dots, p_0$.

We now assume the thesis to hold for $k \geq 0$ and we prove it true for $k + 1$. We recall (2.10) and set

$$z_0 = z, \quad z_1 = \gamma^{(0,k)}_{v,\delta}(z_0), \quad z_2 = e^{\delta^2 Y}(z_1), \quad z_3 = \gamma^{(0,k)}_{v,-\delta}(z_2),$$
$$z_4 = e^{-\delta^2 Y}(z_3) = \gamma^{(0,k+1)}_{v,\delta}(z) = \gamma^{(-1,k+1)}_{v,\delta}(z).$$

Now, by triangular inequality we get

$$\left|u\big(\gamma^{(-1,k+1)}_{v,\delta}(z)\big) - u(z)\right| \leq \sum_{i=1}^{4} |u(z_i) - u(z_{i-1})|,$$

and thus, (2.23) for $k + 1$ follows from the inductive hypothesis and from the assumption $u \in C^{\alpha}_{Y}$. □

We are now ready to prove Part 2 of Theorem 2.20 for $n = 0$.

Proof of Theorem 2.20 (Part 2) for $n = 0$.
We first consider the particular case $z = (t, x)$, $\zeta = (t, \xi)$, with $x, \xi \in \mathbb{R}^d$. Precisely, we show that, if $u \in C^{0,\alpha}_B$ we have

$$|u(t, x) - u(t, \xi)| \leq c_B \|u\|_{C^{0,\alpha}_B} [x - \xi]^{\alpha}_B, \qquad t \in \mathbb{R}, \quad x, \xi \in \mathbb{R}^d. \tag{2.24}$$

By the triangular inequality, we obtain

$$|u(t,x) - u(t,\xi)| \leq \sum_{i=0}^{r} |u(\zeta_i) - u(\zeta_{i-1})|,$$

where the points $\zeta_k = (t, \xi_k)$, for $k = -1, 0, \cdots, r$, are defined as in Lemma 2.28 by setting $n = 0$ and $v = x - \xi$. The estimate (2.24) then stems from (2.23) with $n = 0$, combined with (2.18).

We now prove the general case. For any $z, \zeta \in \mathbb{R} \times \mathbb{R}^d$, by triangular inequality we get

$$\begin{aligned} |u(z) - u(\zeta)| &\leq |u(z) - u(e^{(t-s)Y}(\zeta))| + |u(e^{(t-s)Y}(\zeta)) - u(\zeta)| \\ &= |u(t,x) - u(t, e^{(t-s)B}\xi)| + |u(e^{(t-s)Y}(\zeta)) - u(\zeta)|. \quad (2.25) \end{aligned}$$

As $u \in C_Y^\alpha$ we can estimate the second term with $|s-t|^{\frac{1}{2}} \leq \|\zeta^{-1} \circ z\|$ Now, to prove (2.3), we use (2.24) to bound the first term in (2.25), $u \in C_Y^\alpha$ to bound the second one, and we obtain

$$|u(z) - u(\zeta)| \leq c_B \|u\|_{C_B^{0,\alpha}} \|\zeta^{-1} \circ z\|_B^\alpha,$$

which concludes the proof. □

2.3.2 Step 2

Throughout this section we fix $\bar{n} \in \{0, \cdots, r\}$ and assume to be holding true:

- Proposition 2.31 for any $0 \leq n \leq \bar{n} - 1$, if $\bar{n} \geq 1$;
- Theorem 2.20 for any $0 \leq n \leq 2\bar{n}$.

Then we prove:

- Propositions 2.32 and 2.33 for $n = \bar{n}, m = 1$;
- Proposition 2.31 for $n = \bar{n}$;
- Theorem 2.20 (Part 2) for $n = 2\bar{n} + 1$.

This induction step has to be treated separately because we cannot assume a priori the existence of the first order Euclidean partial derivatives w.r.t. the $\bar{n}$-th level variables. Therefore, we introduce the following alternative definition of $(2\bar{n}+1)$-th order intrinsic Taylor polynomial of u that does not make explicit use of the derivatives $\left(\partial_{\bar{p}_{\bar{n}-1}+i}u\right)_{1\le i\le p_{\bar{n}}}$:

$$\bar{\mathfrak{T}}_{2\bar{n}+1}(u,\zeta)(z) := \sum_{\substack{0\le 2k+|\beta|_B\le 2\bar{n}+1\\ \beta^{[\bar{n}]}=0}} \frac{1}{k!\,\beta!}\left(Y^k\partial_\xi^\beta u(\zeta)\right)(t-s)^k\left(x-e^{(t-s)B}\xi\right)^\beta$$
$$+\sum_{i=\bar{p}_{\bar{n}-1}+1}^{\bar{p}_{\bar{n}}}\left(Y^{(\bar{n})}_{v_i^{(\bar{n})}}u(\zeta)\right)\left(x-e^{B(t-s)}\xi\right)_i, \tag{2.26}$$

with $\left(v_i^{(\bar{n})}\right)_{\bar{p}_{\bar{n}-1}<i\le\bar{p}_{\bar{n}}}$ being the family of vectors such that $v_i^{(\bar{n})}\in V_{0,\bar{n}}$ with $B^{\bar{n}}v_i^{(\bar{n})}=e_i$.

Remark 2.34. The Taylor polynomial $\bar{\mathfrak{T}}_{2\bar{n}+1}$ is well-defined for any $u\in C^{2\bar{n}+1,\alpha}_{B,\mathrm{loc}}$. In fact, by Lemma 2.25 we have

$$Y^{(\bar{n})}_{v_i^{(\bar{n})}}u\in C^{0,\alpha}_{B,\mathrm{loc}},\qquad \bar{p}_{\bar{n}-1}<i\le\bar{p}_{\bar{n}}.$$

On the other hand, by using the inclusion of the spaces $C^{n,\alpha}_{B,\mathrm{loc}}$ and the inductive hypothesis (Theorem 2.20, Part 1), the Euclidean derivatives

$$\partial_\xi^\beta u(\zeta),\qquad 0\le|\beta|_B\le 2\bar{n}+1,\quad \beta^{[\bar{n}]}=0,$$

are well defined. Therefore, by combining the inductive hypothesis on Proposition 2.31 and Lemma 2.25, we have

$$Y^k\partial_\xi^\beta u(\zeta)\in C^{2\bar{n}+1-2k-|\beta|_B,\alpha}_{B,\mathrm{loc}},\qquad 0\le 2k+|\beta|_B\le 2\bar{n}+1,\quad \beta^{[\bar{n}]}=0.$$

In particular, by analogous arguments, if $u\in C^{2\bar{n}+1,\alpha}_B$ and $0\le 2k+|\beta|_B\le 2\bar{n}+1$, $\beta^{[\bar{n}]}=0$ we have that

$$Y^k\partial_\xi^\beta u(\zeta)\in C^{2\bar{n}+1-2k-|\beta|_B,\alpha}_B, \tag{2.27}$$
$$Y^{(\bar{n})}_{v_i^{(\bar{n})}}u\in C^{0,\alpha}_B,\qquad \bar{p}_{\bar{n}-1}<i\le\bar{p}_{\bar{n}}. \tag{2.28}$$

Remark 2.35. By simple linear algebra arguments, it is also easy to show that for a given $\alpha\in[0,1]$, $n\in\{0,\cdots,r\}$ and $u\in C^{2n+1,\alpha}_B$, we have

$$\sum_{i=\bar{p}_{n-1}+1}^{\bar{p}_n}\left(Y^{(n)}_{v_i^{(n)}}u(\zeta)\right)(B^n v)_i = Y^n_v u(\zeta),\qquad \zeta\in\mathbb{R}\times\mathbb{R}^d,\quad v\in V_{0,n}.$$

Proof of Propositions 2.32 and 2.33, for $\mathbf{n = \bar{n}}$ and $\mathbf{m = 1}$

We prove Propositions 2.32 and 2.33 on $\bar{\mathcal{T}}_{2\bar{n}+1}u$, for $n = \bar{n}$ and $m = 1$. Note that, after proving Proposition 2.31 for $n = \bar{n}$, the two versions of the Taylor polynomials $\bar{\mathcal{T}}_{2\bar{n}+1}u$ and $\mathcal{T}_{2\bar{n}+1}u$ will turn out to be equivalent.

Proof of Proposition 2.32 for $n = \bar{n}, m = 1$.
We assume $u \in C_B^{2\bar{n}+1,\alpha}$ and we have to prove that for any $\max\{\bar{n}-1, 0\} \le k \le r$, $v \in V_{0,k}$ with $|v| = 1$, and z, we have

$$u\big(\gamma_{v,\delta}^{(\bar{n}-1,k)}(z)\big) = \bar{\mathcal{T}}_{2\bar{n}+1}(u,z)\big(\gamma_{v,\delta}^{(\bar{n}-1,k)}(z)\big) + R_{v,\delta}^{(\bar{n}-1,k)}(z), \tag{2.29}$$

with

$$|R_{v,\delta}^{(\bar{n}-1,k)}(z)| \le c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} |\delta|^{2\bar{n}+1+\alpha}, \qquad \delta \in \mathbb{R}. \tag{2.30}$$

We prove (2.30) by induction on k.
Proof for $k = \max\{\bar{n} - 1, 0\}$: because of the particular definition of $\gamma_{v,\delta}^{(n,k)}$ we have to treat separately the cases $\bar{n} = 0$, $\bar{n} = 1$ and $\bar{n} > 1$.
<u>Case $\bar{n} = 0$</u>: by (2.11) and (2.9) we have

$$\gamma_{v,\delta}^{(-1,0)}(z) = u(t, x + \delta v),$$

and thus, by (2.26), (2.29) for $k = 0$ reads as

$$u(t, x + \delta v) = u(t, x) + \delta \sum_{i=1}^{p_0} \partial_{x_i} u(t, x) v_i + R_{v,\delta}^{(-1,0)}(z).$$

Now, by the standard mean-value theorem, there exist $(\bar{v}_i)_{i=1,\dots,p_0}$ with $\bar{v}_i \in V_0$ and $|\bar{v}_i| \le |v| \le 1$, such that

$$u(t, x + \delta v) - u(t, x) = \delta \sum_{i=1}^{p_0} \partial_{x_i} u(t, x + \delta \bar{v}_i) v_i,$$

and thus

$$R_{v,\delta}^{(-1,0)}(z) = \delta \sum_{i=0}^{p_0} (\partial_{x_i} u(t, x + \delta \bar{v}_i) - \partial_{x_i} u(t, x)) v_i.$$

Note that $\partial_{x_i} u \in C_B^{0,\alpha}$ for any $1 \le i \le p_0$ because $u \in C_B^{1,\alpha}$ by assumption. Therefore estimate (2.30) stems from Part 2 of Theorem 2.20 for $n = 0$.
<u>Case $\bar{n} = 1$</u>: by (2.9) we have

$$\gamma_{v,\delta}^{(0,0)}(z) = u(t, x + \delta v),$$

and thus, by (2.26), (2.29) for $k = 0$ reads as

$$u(t,x+\delta v) = u(t,x) + \delta \sum_{i=1}^{p_0} \partial_{x_i} u(t,x) v_i + \frac{\delta^2}{2!} \sum_{i,j=1}^{p_0} \partial_{x_i x_j} u(t,x) v_i v_j + \frac{\delta^3}{3!} \sum_{i,j,l=1}^{p_0} \partial_{x_i x_j x_l} u(t,x) v_i v_j v_l + R_{v,\delta}^{(0,0)}(z).$$

Now, by the mean-value theorem, there exist $(\bar{v}_{i,j,k})_{1\le i,j,k \le p_0}$, with $\bar{v}_{i,j,k} \in V_0$ and $|\bar{v}_{i,j,k}| \le |v| \le 1$, such that

$$u(t,x+\delta v) - u(t,x) - \delta \sum_{i=1}^{p_0} \partial_{x_i} u(t,x) v_i - \frac{\delta^2}{2!} \sum_{i,j=1}^{p_0} \partial_{x_i x_j} u(t,x) v_i v_j = \frac{\delta^3}{3!} \sum_{i,j,l=1}^{p_0} \partial_{x_i x_j x_l} u(t,x+\delta \bar{v}_{i,j,k}) v_i v_j v_l,$$

and thus

$$R_{v,\delta}^{(0,0)}(z) = \frac{\delta^3}{3!} \sum_{i,j,l=1}^{p_0} \big(\partial_{x_i,x_j,x_l} u(t,x+\delta \bar{v}_{i,j,l}) - \partial_{x_i,x_j,x_l} u(t,x)\big) v_i v_j v_l.$$

Note that $\partial_{x_i,x_j,x_l} u \in C_B^{0,\alpha}$ for any $1 \le i,j,l \le p_0$ since, by assumption, $u \in C_B^{3,\alpha}$. Estimate (2.30) then stems from Part 2 of Theorem 2.20 for $n = 0$.

Case $\bar{n} > 1$: by (2.9) we have

$$\gamma_{v,\delta}^{(\bar{n}-1,\bar{n}-1)}(z) = u(t, x + \delta^{2\bar{n}-1} B^{\bar{n}-1} v),$$

and thus, by (2.26), (2.29) for $k = \bar{n} - 1$ reads as

$$u\big(t, x + \delta^{2\bar{n}-1} B^{\bar{n}-1} v\big) = u(z) + \delta^{2\bar{n}-1} \sum_{i=\bar{p}_{\bar{n}-2}+1}^{\bar{p}_{\bar{n}-1}} \partial_{x_i} u(z) (B^{\bar{n}-1} v)_i + R_{v,\delta}^{(\bar{n}-1,\bar{n}-1)}(z).$$

Now, by the mean-value theorem, there exists a family of vectors $(\bar{v}_i)_{\bar{p}_{\bar{n}-2} < i \le \bar{p}_{\bar{n}-1}}$, with $\bar{v}_i \in V_{\bar{n}-1}$ and $|\bar{v}_i| \le |B^{\bar{n}-1} v| \le c_B$, such that

$$u(t, x + \delta^{2\bar{n}-1} B^{\bar{n}-1} v) - u(t,x) = \delta^{2\bar{n}-1} \sum_{i=\bar{p}_{\bar{n}-2}+1}^{\bar{p}_{\bar{n}-1}} \partial_{x_i} u(t, x + \delta^{2\bar{n}-1} \bar{v}_i)(B^{\bar{n}-1} v)_i,$$

and thus,

$$R_{v,\delta}^{(\bar{n}-1,\bar{n}-1)}(z) = \delta^{2\bar{n}-1} \sum_{i=\bar{p}_{\bar{n}-2}+1}^{\bar{p}_{\bar{n}-1}} \big(\partial_{x_i} u(t, x + \delta^{2\bar{n}-1} \bar{v}_i) - \partial_{x_i} u(t,x)\big) (B^{\bar{n}-1} v)_i$$

$$= \delta^{2\bar{n}-1} \sum_{i=\bar{p}_{\bar{n}-2}+1}^{\bar{p}_{\bar{n}-1}} \big(\partial_{x_i} u(t, x+\delta^{2\bar{n}-1}\bar{v}_i) - \mathfrak{T}_2(\partial_{x_i} u, (t,x))\big(t, x+\delta^{2\bar{n}-1}\bar{v}_i\big)\big)(B^{\bar{n}-1}v)_i.$$

Now, by (2.27) in Remark 2.34, we have $\partial_{x_i} u \in C_B^{2,\alpha}$ for any $\bar{p}_{\bar{n}-2} < i \leq \bar{p}_{\bar{n}-1}$. Therefore estimate (2.30) stems from Part 2 of Theorem 2.20 for $n = 2$.

Inductive step on k: we assume the thesis to hold true for a fixed $\max\{\bar{n} - 1, 0\} \leq k < r$ and prove it true for $k+1$. Consider thus $v \in V_{0,k+1}$ with $|v| = 1$. Set

$$\widetilde{\mathfrak{T}}_{2\bar{n}+1}(u,\zeta)(z) = \bar{\mathfrak{T}}_{2\bar{n}+1}(u,\zeta)(z) - u(\zeta), \qquad z,\zeta \in \mathbb{R}\times\mathbb{R}^d,$$

and

$$\begin{gathered} z_0 = z, \quad z_1 = \gamma_{v,\delta}^{(\bar{n}-1,k)}(z_0), \quad z_2 = e^{\delta^2 Y}(z_1), \quad z_3 = \gamma_{v,-\delta}^{(\bar{n}-1,k)} \\ (z_2), \quad z_4 = e^{-\delta^2 Y}(z_3) = \gamma_{v,\delta}^{(\bar{n}-1,k+1)}(z). \end{gathered} \tag{2.31}$$

According to this notation we have

$$\begin{aligned} R_{v,\delta}^{(\bar{n}-1,k+1)}(z) &= u\big(\gamma_{v,\delta}^{(\bar{n}-1,k+1)}(z)\big) - \bar{\mathfrak{T}}_{2\bar{n}+1}(u,z)\big(\gamma_{v,\delta}^{(\bar{n}-1,k+1)}(z)\big) \\ &= u(z_4) - \bar{\mathfrak{T}}_{2\bar{n}+1}(u,z_0)(z_4) = \sum_{i=1}^{6} G_i, \end{aligned}$$

with

$$\begin{aligned} G_1 &= u(z_4) - u(z_3) - \sum_{i=1}^{\bar{n}} \frac{(-\delta^2)^i}{i!} Y^i u(z_3), \\ G_2 &= u(z_3) - u(z_2) - \widetilde{\mathfrak{T}}_{2\bar{n}+1}(u,z_2)(z_3), \\ G_3 &= \sum_{i=1}^{\bar{n}} \frac{(-\delta^2)^i}{i!} Y^i u(z_2) + u(z_2) - u(z_1), \\ G_4 &= \widetilde{\mathfrak{T}}_{2\bar{n}+1}(u,z_1)(z_0) + u(z_1) - u(z_0), \\ G_5 &= \sum_{i=1}^{\bar{n}} \frac{(-\delta^2)^i}{i!} \Big(Y^i u(z_3) - Y^i u(z_2) - \widetilde{\mathfrak{T}}_{2(\bar{n}-i)+1}(Y^i u, z_2)(z_3)\Big), \\ G_6 &= \widetilde{\mathfrak{T}}_{2\bar{n}+1}(u,z_2)(z_3) - \widetilde{\mathfrak{T}}_{2\bar{n}+1}(u,z_1)(z_0) - \widetilde{\mathfrak{T}}_{2\bar{n}+1}(u,z_0)(z_4) \\ &\quad + \sum_{i=1}^{\bar{n}} \frac{(-\delta^2)^i}{i!} \widetilde{\mathfrak{T}}_{2(\bar{n}-i)+1}(Y^i u, z_2)(z_3). \end{aligned}$$

Now, by applying Remark 2.29 with $n = \bar{n}$, $m = 1$, on G_1 and G_3, and by using the inductive hypothesis on G_2 and G_4 (note that by (1.14) $V_{0,k+1} \subseteq V_{0,k}$), we have

$$|G_1+G_2+G_3+G_4| \le c_B \|u\|_{C_B^{2\bar n+1,\alpha}} |\delta|^{2\bar n+1+\alpha}, \qquad z=(t,x)\in\mathbb{R}\times\mathbb{R}^d, \quad \delta\in\mathbb{R}.$$

To bound G_5, it is enough to observe that, by Definition 1.14, $u \in C_B^{2\bar n+1,\alpha}$ implies $Y^i u \in C_B^{2(\bar n-i)+1,\alpha}$, for any $i=1,\cdots,\bar n$. Therefore, the bound follows by applying Part 2 of Theorem 2.20 for $n=2(\bar n-i)+1$, combined with (2.17).

In order to estimate G_6 and conclude the proof, we need to distinguish on whether $k=\max\{\bar n-1,0\}$, $k=\bar n$ or $k>\bar n$.

Case $k>\bar n$: there is nothing to prove because, by definitions (2.26) and (2.31), we have $G_6\equiv 0$.

Case $k=\bar n$: first note that, in this case, the term G_6 reduces to

$$\begin{aligned} G_6 &= \widetilde{\mathfrak{T}}_{2\bar n+1}(u,z_2)(z_3) - \widetilde{\mathfrak{T}}_{2\bar n+1}(u,z_1)(z_0) \\ &= \widetilde{\mathfrak{T}}_{2\bar n+1}(u,z_2)\big(\gamma_{v,-\delta}^{(\bar n-1,\bar n)}(z_2)\big) - \widetilde{\mathfrak{T}}_{2\bar n+1}(u,z_1)\big(\gamma_{v,-\delta}^{(\bar n-1,\bar n)}(z_1)\big), \end{aligned}$$

and by definition (2.26), along with (2.14)-(2.15), we get

$$|G_6| = \Big|\delta^{2\bar n+1} \sum_{i=\bar p_{\bar n-1}+1}^{\bar p_{\bar n}} \big(Y_{v_i^{(\bar n)}}^{(\bar n)} u(z_1) - Y_{v_i^{(\bar n)}}^{(\bar n)} u(z_2)\big)\left(B^{\bar n} v\right)_i\Big| =$$

(by Remark 2.35 with $n=\bar n$ and since $v\in V_{0,\bar n+1}\subseteq V_{0,\bar n}$)

$$= \Big|\delta^{2\bar n+1}\big(Y_v^{(\bar n)}u(z_1) - Y_v^{(\bar n)}u(z_2)\big)\Big| \le$$

(by hypothesis $u\in C_B^{2\bar n+1,\alpha}$ and thus, by Lemma 2.25, $Y_v^{(\bar n)}u \in C_B^{0,\alpha}\subseteq C_Y^\alpha$)

$$\le c_B \|u\|_{C_B^{2\bar n+1,\alpha}} |\delta|^{2\bar n+1+\alpha}.$$

Case $k=\max\{\bar n-1,0\}$: we only need to prove the case $\bar n>0$. We first consider $\bar n\ge 2$. We have

$$\begin{aligned} G_6 = {} & \widetilde{\mathfrak{T}}_{2\bar n+1}(u,z_2)\big(\gamma_{v,-\delta}^{(\bar n-1,\bar n-1)}(z_2)\big) - \widetilde{\mathfrak{T}}_{2\bar n+1}(u,z_1)\big(\gamma_{v,-\delta}^{(\bar n-1,\bar n-1)}(z_1)\big) \\ & - \widetilde{\mathfrak{T}}_{2\bar n+1}(u,z_0)\big(\gamma_{v,\delta}^{(\bar n-1,\bar n)}(z_0)\big) + \sum_{i=1}^{\bar n} \frac{(-\delta^2)^i}{i!} \widetilde{\mathfrak{T}}_{2(\bar n-i)+1}(Y^i u, z_2)\big(\gamma_{v,-\delta}^{(\bar n-1,\bar n-1)}(z_2)\big). \end{aligned}$$

Now recall that, by (2.14)-(2.15),

$$\gamma_{v,-\delta}^{(\bar n-1,\bar n-1)}(z) = \big(t, x-\delta^{2(\bar n-1)+1}B^{\bar n-1}v\big),$$

$$\gamma_{v,\delta}^{(\bar n-1,\bar n)}(z) = \big(t, x+\delta^{2\bar n+1}B^{\bar n}v + \widetilde S_{\bar n-1,\bar n}(\delta)v\big), \qquad \widetilde S_{\bar n-1,\bar n}(\delta)v \in \bigoplus_{j=\bar n+1}^{r} V_j,$$

and thus, by definition (2.26), we obtain

$$G_6 = \delta^{2(\bar n-1)+1}\sum_{i=\bar p_{\bar n-2}+1}^{\bar p_{\bar n-1}} \big(\partial_{x_i}u(z_1) - \partial_{x_i}u(z_2) + \delta^2\partial_{x_i}Yu(z_2)\big)\,(B^{\bar n-1}v)_i$$
$$- \delta^{2\bar n+1}\sum_{i=\bar p_{\bar n-1}+1}^{\bar p_{\bar n}} Y^{(\bar n)}_{v_i^{(\bar n)}}u(z_0)\,(B^{\bar n}v)_i$$

(by Proposition 2.31 for $n = \bar n - 1$)

$$= \delta^{2(\bar n-1)+1}\sum_{i=\bar p_{\bar n-2}+1}^{\bar p_{\bar n-1}} \big(Y^{(\bar n-1)}_{v_i^{(\bar n-1)}}u(z_1) - Y^{(\bar n-1)}_{v_i^{(\bar n-1)}}u(z_2) + \delta^2 Y^{(\bar n-1)}_{v_i^{(\bar n-1)}}Yu(z_2)\big)\,(B^{\bar n-1}v)_i$$

$$- \delta^{2\bar n+1}\sum_{i=\bar p_{\bar n-1}+1}^{\bar p_{\bar n}} Y^{(\bar n)}_{v_i^{(\bar n)}}u(z_0)(B^{\bar n}v)_i$$

(by applying Remark 2.35 with $n = \bar n - 1$ and $n = \bar n$, and since $v \in V_{0,\bar n} \subseteq V_{0,\bar n-1}$)

$$= \delta^{2(\bar n-1)+1}\big(Y^{(\bar n-1)}_v u(z_1) - Y^{(\bar n-1)}_v u(z_2) + \delta^2 Y^{(\bar n-1)}_v Yu(z_2)\big) - \delta^{2\bar n+1}Y^{\bar n}_v u(z_0)$$

(since, by definition (2.7), $Y^{(\bar n-1)}_v Y = Y^{(\bar n)}_v + YY^{\bar n-1}_v$)

$$= \delta^{2(\bar n-1)+1}\big(Y^{(\bar n-1)}_v u(z_1) - Y^{(\bar n-1)}_v u(z_2)$$
$$+ \delta^2 YY^{(\bar n-1)}_v u(z_2)\big) + \delta^{2\bar n+1}\big(Y^{(\bar n)}_v u(z_2) - Y^{(\bar n)}_v u(z_0)\big) = \sum_{i=1}^{3} F_i.$$

with

$$F_1 = \delta^{2(\bar n-1)+1}\left(Y^{(\bar n-1)}_v u(z_1) - Y^{(\bar n-1)}_v u(z_2) + \delta^2 YY^{(\bar n-1)}_v u(z_2)\right),$$
$$F_2 = \delta^{2\bar n+1}\left(Y^{(\bar n)}_v u(z_2) - Y^{(\bar n)}_v u(z_1)\right), \quad F_3 = \delta^{2\bar n+1}\left(Y^{(\bar n)}_v u(z_1) - Y^{(\bar n)}_v u(z_0)\right).$$

Now, to bound F_1 it is sufficient to note that, by Lemma 2.25, $Y^{(\bar n-1)}_v u \in C^{2,\alpha}_B$ and thus the bounds directly follow by applying Remark 2.29 with $n = 1$ and

$m = 0$. To bound the terms F_2 and F_3 we use that, by Lemma 2.25, $Y_v^{\bar{n}} u \in C_B^{0,\alpha}$. The estimate for F_3 then follows by Part 2 of Theorem 2.20 for $n = 0$ along with equation (2.17), whereas the one for F_2 is a consequence of the inclusion $C_B^{0,\alpha} \subseteq C_Y^{\alpha}$ and of Remark 2.29.

Finally, the case $\bar{n} = 1$ is analogous, but G_6 contains two more terms:

$$F_4 = \frac{\delta^2}{2!} \sum_{i,j=1}^{p_0} \big(\partial_{x_i,x_j} u(z_2) - \partial_{x_i,x_j} u(z_1)\big) v_i v_j,$$

$$F_5 = -\frac{\delta^3}{3!} \sum_{i,j,l=1}^{p_0} \big(\partial_{x_i,x_j,x_l} u(z_2) - \partial_{x_i,x_j,x_l} u(z_1)\big) v_i v_j v_l,$$

which can be estimated by using that $\partial_{x_i,x_j,x_l} u \in C_B^{0,\alpha} \subseteq C_Y^{\alpha}$ and $\partial_{x_i,x_j} u \in C_B^{1,\alpha} \subseteq C_Y^{\alpha+1}$ for any $1 \le i,j,l \le p_0$. $\square$

Proof of Proposition 2.33 for $n = \bar{n}$ and $m = 1$. We assume $u \in C_B^{2\bar{n}+1,\alpha}$ and we prove that, for any $0 \le k \le \bar{n}$,

$$u(t, x+\xi) = \mathcal{T}_{2\bar{n}+1}(u,(t,x))(t,x+\xi) + R_{\bar{n}}\big(t,x,\xi\big),$$

with

$$|R_{\bar{n}}\big(t,x,\xi\big)| \le c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} [\xi]_B^{2n+1+\alpha}, \qquad (t,x) \in \mathbb{R}\times\mathbb{R}^d, \quad \xi \in \bigoplus_{j=0}^{k-1} V_j.$$

We prove the thesis by induction on k. For $k = 0$ there is nothing to prove since $R_{\bar{n}}\big(t,x,0\big) \equiv 0$. Now, assume $0 \le k < \bar{n}$, $\xi \in \bigoplus_{j=0}^{k-1} V_j$ and $v \in V_k$. Then

$$R_{\bar{n}}(t,x,\xi+v) = F_1 + F_2,$$

with

$$F_1 = u(t,x+\xi+v) - \mathcal{T}_{2\bar{n}+1}\big(u,(t,x+v)\big)(t,x+\xi+v)$$
$$F_2 = \mathcal{T}_{2\bar{n}+1}\big(u,(t,x+v)\big)(t,x+\xi+v) - \mathcal{T}_{2\bar{n}+1}\big(u,(t,x)\big)(t,x+\xi+v).$$

We can apply the inductive hypothesis on F_1 and obtain the estimate

$$|F_1| \le c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} [\xi]_B^{2\bar{n}+1+\alpha} \le c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} [\xi+v]_B^{2\bar{n}+1+\alpha}.$$

Recalling (2.4), F_2 can be written as

$$F_2 = \sum_{\substack{0 \le |\beta|_B \le 2\bar{n}+1 \\ \beta^{[i]}=0 \text{ if } i \ge k}} \frac{1}{\beta!} \partial_x^\beta u(t,x+v)\, \xi^\beta$$

$$
\begin{aligned}
&- \sum_{\substack{0\le|\beta|_B\le 2\bar n+1\\ \beta^{[i]}=0 \text{ if } i\ge k}} \sum_{\substack{0\le|\gamma|_B\le 2\bar n+1-|\beta|_B\\ \gamma=\gamma^{[k]}}} \frac{1}{\beta!\gamma!}\,\partial_x^\gamma\partial_x^\beta u(t,x)\,\xi^\beta v^\gamma\\
&= \sum_{\substack{0\le|\beta|_B\le 2\bar n+1\\ \beta^{[i]}=0 \text{ if } i\ge k}} \frac{1}{\beta!}\Bigg(\partial_x^\beta u(t,x+v) - \sum_{\substack{0\le|\gamma|_B\le 2\bar n+1-|\beta|_B\\ \gamma=\gamma^{[k]}}} \frac{1}{\gamma!}\,\partial_x^\gamma\partial_x^\beta u(t,x)\,v^\gamma\Bigg)\xi^\beta\\
&= \sum_{\substack{0\le|\beta|_B\le 2\bar n+1\\ \beta^{[i]}=0 \text{ if } i\ge k}} \frac{1}{\beta!}\Bigg(\partial_x^\beta u(t,x+v) - \mathcal{T}_{2\bar n+1-|\beta|_B}\big(\partial_x^\beta u,(t,x)\big)(t,x+v)\Bigg)\xi^\beta.
\end{aligned}
$$

By Remark 2.34, we get $\partial_x^\beta u \in C_B^{2\bar n+1-|\beta|_B,\alpha}$. Now, if $|\beta|_B\ge 1$, we can apply Part 2 of Theorem 2.20 for $n=2\bar n+1-|\beta|_B$ on $\partial_x^\beta u$ and get

$$
\begin{aligned}
&\big|\partial_x^\beta u(t,x+v) - \mathcal{T}_{2\bar n+1-|\beta|_B}\big(\partial_x^\beta u,(t,x)\big)(t,x+v)\big|\,\big|\xi^\beta\big|\\
&\quad\le c_B\|u\|_{C_B^{2\bar n+1,\alpha}}[v]_B^{2\bar n+1-|\beta|_B+\alpha}[\xi]_B^{|\beta|_B}\\
&\quad\le c_B\|u\|_{C_B^{2\bar n+1,\alpha}}[\xi+v]_B^{2\bar n+1+\alpha}.
\end{aligned}
$$

On the other hand, if $|\beta|_B=0$ then we have to estimate

$$
u(t,x+v) - \sum_{\substack{0\le|\gamma|_B\le 2\bar n+1\\ \gamma=\gamma^{[k]}}} \frac{1}{\gamma!}\,\partial_x^\gamma u(t,x)\,v^\gamma.
$$

Recall that, by definition, we have $|\gamma|_B=(2k+1)|\gamma|$ if $\gamma=\gamma^{[k]}$. Now, set

$$
j := \max\{i\ge 0 \mid (2k+1)i \le 2\bar n+1\}, \tag{2.32}
$$

and note that $j\ge 1$ because $k<\bar n$. By Remark 2.34 and the mean-value theorem, there exists a family of vectors $(\bar v_\eta)_{\eta\in\mathcal{I}_k^j}$ where

$$
\mathcal{I}_k^j = \{\eta\in\mathbb{N}_0^d \mid \eta=\eta^{[k]} \text{ and } |\eta|_B=(2k+1)j\}, \tag{2.33}
$$

such that $\bar v_\eta\in V_k$, $|\bar v_\eta|\le|v|$ and

$$
u(t,x+v) - \sum_{\substack{0\le|\gamma|_B\le (2k+1)(j-1)\\ \gamma=\gamma^{[i]}}} \frac{v^\gamma}{\gamma!}\partial_x^\gamma u(t,x) = \sum_{\eta\in\mathcal{I}_k^j}\frac{v^\eta}{\eta!}\partial_x^\eta u(t,x+\bar v_\eta).
$$

Therefore, we obtain

$$
\begin{aligned}
&\Big|u(t,x+v) - \sum_{\substack{0\le|\gamma|_B\le 2\bar n+1\\ \gamma=\gamma^{[i]}}} \frac{v^\gamma}{\gamma!}\partial_x^\gamma u(t,x)\Big|\\
&\quad= \Big|\sum_{\eta\in\mathcal{I}_k^j}\frac{v^\eta}{\eta!}\left(\partial_x^\eta u(t,x+\bar v_\eta) - \partial_x^\eta u(t,x)\right)\Big| =
\end{aligned}
$$

(by (2.32))

$$= \Big| \sum_{\eta\in\mathcal{I}_k^j} \frac{1}{\eta!}\big(\partial_x^\eta u(t,x+\bar v_\eta) - \mathcal{T}_{2\bar n+1-(2k+1)j}\big(\partial_x^\eta u,(t,x)\big)(t,x+\bar v_\eta)\big)v^\eta\Big| \le$$

(by Remark 2.34, $\partial_x^\eta u \in C_B^{2\bar n+1-(2k+1)j,\alpha}$ and thus by Part 2 of Theorem 2.20 with $n = 2\bar n+1-(2k+1)j$)

$$\le c_B\|u\|_{C_B^{2\bar n+1,\alpha}} \sum_{\eta\in\mathcal{I}_k^j} \frac{1}{\eta!}\,[\bar v_\eta]_B^{2\bar n+1-(2k+1)j+\alpha}\,[v]_B^{|\eta|_B} \le$$

(since $|\bar v_\eta| \le |v|$ and by (2.33))

$$\le c_B\|u\|_{C_B^{2\bar n+1,\alpha}}[v]_B^{2\bar n+1+\alpha} \le c_B\|u\|_{C_B^{2\bar n+1,\alpha}}[\xi+v]_B^{2\bar n+1+\alpha}$$

which concludes the proof. □

Proof of Proposition 2.31 for $\mathbf{n}=\bar{\mathbf{n}}$

To start we show that if $u \in C_B^{2\bar n+1,\alpha}$ then for any $z=(t,x)$, $\zeta=(t,\xi)\in\mathbb{R}\times\mathbb{R}^d$ we have

$$\big|u(t,x)-\bar{\mathcal{T}}_{2\bar n+1}(u,(t,\xi))(t,x)\big| \le c_B\|u\|_{C_B^{2\bar n+1,\alpha}}\,[x-\xi]_B^{2\bar n+1+\alpha}. \tag{2.34}$$

Define the point $\bar z=(t,\bar x)$ with

$$\bar x^{[i]} = \begin{cases} x^{[i]}, & \text{if } i\ge \bar n,\\ \xi^{[i]}, & \text{if } i<\bar n.\end{cases}$$

It follows that

$$(x-\bar x)^\beta = \begin{cases} (x-\xi)^\beta & \text{if } |\beta|_B\le 2\bar n+1,\ \beta^{[\bar n]}=0,\\ 0, & \text{if } |\beta|_B\le 2\bar n+1,\ \beta^{[\bar n]}\ne 0,\end{cases} \tag{2.35}$$

and

$$[x-\bar x]_B \le [x-\xi]_B, \qquad [\bar x-\xi]_B\le[x-\xi]_B.$$

Then we write

$$u(t,x)-\bar{\mathcal{T}}_{2\bar n+1}(u,(t,\xi))(t,x) = F_1+F_2,$$

with

$$F_1 = u(t,x)-\bar{\mathcal{T}}_{2\bar n+1}u\left((t,\bar x),(t,x)\right),$$

$$F_2 = \bar{\mathfrak{T}}_{2\bar{n}+1}(u,(t,\bar{x}))(t,x) - \bar{\mathfrak{T}}_{2\bar{n}+1}(u,(t,\xi))(t,x).$$

Applying Proposition 2.33 with $n = \bar{n}$ and $m = 1$, we obtain

$$|F_1| \le c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} [x-\bar{x}]_B^{2\bar{n}+1+\alpha} \le c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} [x-\xi]_B^{2\bar{n}+1+\alpha}.$$

Now, by (2.35) we have

$$F_2 = \sum_{\substack{|\beta|_B \le 2\bar{n}+1 \\ \beta^{[\bar{n}]}=0}} \frac{1}{\beta!} \big(\partial_x^\beta u(t,\bar{x}) - \partial_x^\beta u(t,\xi)\big)(x-\xi)^\beta - \sum_{i=\bar{p}_{\bar{n}-1}+1}^{\bar{p}_{\bar{n}}} Y_{v_i^{(\bar{n})}}^{(\bar{n})} u(t,\xi)(x-\xi)_i.$$

Moreover, by Remark 2.34 we have $\partial_x^\beta u \in C_B^{2\bar{n}+1-|\beta|_B,\alpha}$ and therefore, if $|\beta|_B > 0$, by Part 2 of Theorem 2.20 for $n = 2\bar{n}+1-|\beta|_B$, we get

$$\begin{aligned}\big|\big(\partial_x^\beta u(t,\bar{x}) - \partial_x^\beta u(t,\xi)\big)(x-\xi)^\beta\big| &\le c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} [\bar{x}-\xi]_B^{2\bar{n}+1+\alpha-|\beta|_B} |x-\xi|^{|\beta|} \\ &\le c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} [x-\xi]_B^{2\bar{n}+1+\alpha}.\end{aligned}$$

In order to conclude the proof of (2.34), we only have to prove

$$\Big|u(t,\bar{x}) - u(t,\xi) - \sum_{j=\bar{p}_{\bar{n}-1}+1}^{\bar{p}_{\bar{n}}} Y_{v_j^{(\bar{n})}}^{(\bar{n})} u(t,\xi)(x-\xi)_j\Big| \le c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} [x-\xi]_B^{2\bar{n}+1+\alpha}. \tag{2.36}$$

We set the points $\zeta_i = (t,\xi_i)$, for $i = \bar{n}-1,\cdots,r$, as defined in Lemma 2.28 for $n = \bar{n}$ and $v = \bar{x}-\xi$. By (2.18) we have

$$\bar{\mathfrak{T}}_{2\bar{n}+1}(u,\zeta_{i-1})(\zeta_i) = u(\zeta_{i-1}), \qquad i = \bar{n},\ldots,r,$$

and

$$|\delta_i| \le c_B[\bar{x}-\xi]_B \le c_B[x-\xi]_B, \qquad i = \bar{n},\ldots,r. \tag{2.37}$$

It is now clear that

$$\begin{aligned} u(t,\bar{x}) - u(t,\xi) &- \sum_{j=\bar{p}_{\bar{n}-1}+1}^{\bar{p}_{\bar{n}}} Y_{v_j^{(\bar{n})}}^{(\bar{n})} u(t,\xi)(x-\xi)_j \\ &= u(\zeta_r) - \bar{\mathfrak{T}}_{2\bar{n}+1}(u,\zeta_{\bar{n}-1})(\zeta_{\bar{n}}) \\ &= \sum_{i=\bar{n}}^{r} \big(u(\zeta_i) - \bar{\mathfrak{T}}_{2\bar{n}+1}(u,\zeta_{i-1})(\zeta_i)\big), \end{aligned}$$

and formula (2.36) follows from Proposition 2.32 along with (2.37).

We are now ready to prove (2.22) for $n = \bar{n}$. For any $i \in \{\bar{p}_{\bar{n}-1}+1, \dots, \bar{p}_{\bar{n}}\}$ and $\delta \in \mathbb{R}$, set $x = \xi + \delta e_i$ in (2.34), where e_i is the i-th vector of the canonical basis of $\mathbb{R}^d$: we obtain

$$u(t, \xi + \delta e_i) - u(t, \xi) - \delta Y^{(\bar{n})}_{v_i^{(\bar{n})}} u(t, \xi) = O\big(|\delta|^{1+\frac{\alpha}{2\bar{n}+1}}\big), \qquad \text{as } \delta \to 0.$$

This implies that $\partial_{x_i} u(t, \xi)$ exists and

$$\partial_{x_i} u(t, \xi) = Y^{(\bar{n})}_{v_i^{(\bar{n})}} u(t, \xi) \quad t \in \mathbb{R},\ \xi \in \mathbb{R}^d,\ i = \bar{p}_{\bar{n}-1}+1, \dots, \bar{p}_{\bar{n}}.$$

Finally, by Remark 2.34 we have $Y^{(\bar{n})}_{v_i^{(\bar{n})}} u \in C^{0,\alpha}_B$ and thus $\partial_{x_i} u \in C^{0,\alpha}_B$.

Remark 2.36. Incidentally we have just proved a special case of Part 2 of Theorem 2.20 for $n = 2\bar{n}+1$, namely the case when there is no increment in the time variable. Precisely we have shown that, for any function $u \in C^{2\bar{n}+1}_B$, we have

$$\big|u(t,x) - \mathcal{T}_{2\bar{n}+1}(u, (t,\xi))(t,x)\big| \le c_B \|u\|_{C^{2\bar{n}+1,\alpha}_B} [x - \xi]_B^{2\bar{n}+1+\alpha}. \tag{2.38}$$

Proof of Part 2 of Theorem 2.20 for $n = 2\bar{n}+1$

Relation (2.2) is a trivial consequence of Remark 2.35 (see (2.28)-(2.27)) along with Proposition 2.31 for $n = \bar{n}$. We next prove estimate (2.3): by (2.2), for any $z = (t,x)$ and $\zeta = (s,\xi)$, the B-Taylor polynomial $\mathcal{T}_{2\bar{n}+1}(u,\zeta)(z)$ is well defined. Define the point $\zeta_1 := e^{(t-s)Y}(\zeta) = (t, e^{(t-s)B}\xi)$ and note that ζ_1 and z only differ in the spatial variables. Moreover, we have

$$\zeta_1^{-1} \circ z = \Big(0, x - e^{(t-s)B}\xi\Big), \qquad \zeta^{-1} \circ z = \Big(t-s, x - e^{(t-s)B}\xi\Big),$$

and therefore

$$\big\|\zeta_1^{-1} \circ z\big\|_B = \big[x - e^{(t-s)B}\xi\big]_B \le \big\|\zeta^{-1} \circ z\big\|_B. \tag{2.39}$$

Now write

$$u(z) - \mathcal{T}_{2\bar{n}+1}(u,\zeta)(z) = F_1 + F_2,$$

with

$$F_1 = u(z) - \mathcal{T}_{2\bar{n}+1}(u,\zeta_1)(z), \qquad F_2 = \mathcal{T}_{2\bar{n}+1}(u,\zeta_1)(z) - \mathcal{T}_{2\bar{n}+1}(u,\zeta)(z).$$

By (2.38) in Remark 2.36 along with (2.39), we obtain the estimate

$$|F_1| \leq c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} \|\zeta^{-1} \circ z\|_B^{2\bar{n}+1+\alpha}.$$

A convenient rearrangement of the terms in the Taylor polynomials allows us to estimate F_2. Precisely, we have

$$\begin{aligned}
F_2 = & \sum_{|\beta|_B \leq 2\bar{n}+1} \frac{1}{\beta!} \big(\partial_\xi^\beta u(e^{(t-s)Y}(\zeta))\big)(x - e^{(t-s)B}\xi)^\beta \\
& - \sum_{2k+|\beta|_B \leq 2\bar{n}+1} \frac{Y^k \partial_\xi^\beta u(\zeta)}{\beta! k!} (x - e^{(t-s)B}\xi)^\beta (t-s)^k \\
= & \sum_{|\beta|_B \leq 2\bar{n}+1} \Bigg(\Big(\partial_\xi^\beta u(e^{(t-s)Y}(\zeta)) - \sum_{2k \leq 2\bar{n}+1-|\beta|_B} \frac{(t-s)^k}{k!} Y^k \partial_\xi^\beta u(\zeta) \Big) \\
& \times \frac{(x - e^{(t-s)B}\xi)^\beta}{\beta!} \Bigg).
\end{aligned}$$

Now, by (2.2) we have $\partial_x^\beta u \in C_B^{2\bar{n}+1-|\beta|_B,\alpha}$ and thus, by Remark 2.29 we obtain

$$\begin{aligned}
|F_2| \leq & \|u\|_{C_B^{2\bar{n}+1,\alpha}} \sum_{|\beta|_B \leq 2\bar{n}+1} \frac{1}{\beta!} |t-s|^{\frac{2\bar{n}+1-|\beta|_B+\alpha}{2}} \left[x - e^{(t-s)B}\xi\right]_B^{|\beta|_B} \\
\leq & c_B \|u\|_{C_B^{2\bar{n}+1,\alpha}} \|\zeta^{-1} \circ z\|_B^{2\bar{n}+1+\alpha},
\end{aligned}$$

and this concludes the proof.

2.3.3 Step 3

Fix $\bar{n} \in \{0, \cdots, r-1\}$. Assume to be holding true:

- Proposition 2.31 for any $0 \leq n \leq \bar{n}$;
- Theorem 2.20 for any $0 \leq n \leq 2\bar{n}+1$;

we have to prove:

- Propositions 2.32 and 2.33 for $n = \bar{n}+1$, $m = 0$;
- Part 2 of Theorem 2.20 for $n = 2\bar{n}+2$.

In this case, the proof is relatively simpler if compared to the one of Step 2. This is because we do not need to prove the existence of the Euclidean derivatives of the higher level. Hence the proofs are simpler versions of those in Step 2. We skip the details for the sake of brevity.

2.3.4 Step 4

Here we fix a certain $\bar{n} \geq 2r+1$, suppose Theorem 2.20 true for any $0 \leq n \leq \bar{n}$ and prove Part 2 of Theorem 2.20 for $n = \bar{n}+1$. To prove the claim, we will first consider the case with no increment w.r.t. the time variable, as we have done in Step 2. In that case, we used the curves $\gamma^{n,k}_{v,\delta}(z)$ in order to increment those variables w.r.t. which we had no regularity in the Euclidean sense: then we applied Proposition 2.32 to estimate the increment along such curves. This time, this will not be necessary because, since $\bar{n}+1 > 2r+1$, the existence of the Euclidean derivatives is ensured along any direction by the inductive hypothesis.

Proof of Part 2 of Theorem 2.20 for $n = \bar{n}+1$. Recall that, by hypothesis, $u \in C^{\bar{n}+1,\alpha}_B$ with $\bar{n} \geq 2r+1$. It is easy to prove that, for any $z = (t,x)$, $\zeta = (s,\xi) \in \mathbb{R}^d$, we have

$$|u(t,x) - \mathcal{T}_{\bar{n}+1}(u,(t,\xi))(t,x))| \leq c_B \, \|u\|_{C^{\bar{n}+1,\alpha}_B} \, [x-\xi]^{\bar{n}+1+\alpha}_B. \tag{2.40}$$

The proof of the latter identity is identical to that of Proposition 2.33. Precisely, under the assumption $\bar{n} \geq 2r+1$, the technical restriction made on the spatial increments in Proposition 2.33 can be dropped and the proof proceeds exactly in the same way, by making sure that the constant c_B in (2.40) is actually independent of $\bar{n}$.

The proof of Part 2 of Theorem 2.20 then follows exactly as in Step 2, by using the estimate (2.40) instead of (2.38). □

2.4 The non homogeneous case

As previously said, the Taylor formula (2.3) locally holds also for functions defined on a domain $\Omega \subset \mathbb{R} \times \mathbb{R}^d$. To prove this, once the base point ζ is fixed, one has to choose a z so close to ζ to ensure that any integral curves used to connect the two points still completely lies in Ω. As every curve defined so long in this chapter is a continuous function of time (in fact, analytic) and the proof only use a finite amount of them, such procedure is viable and, actually, was the one used in Pagliarani et al. (2016).

More generally, we can allow z, ζ to move in a small domain U to which will correspond a domain V such that $\bar{U} \subset V \subset \bar{V} \subset \Omega$. The result would then read as

$$|u(z) - \mathcal{T}_n(u,\zeta)(z)| \leq c_B \|u\|_{C_B^{n,\alpha}(V)} \|\zeta^{-1} \circ z\|_B^{n+\alpha}, \qquad z, \zeta \in U. \tag{2.41}$$

Note the presence of the norm $\|u\|_{C_B^{n,\alpha}(V)}$.

We shall prove a slightly weaker result but under much weaker assumptions. From this point to the end of the chapter, the matrix B is allowed to take the more general form (1.6). For greater convenience we relabel the blocks as follow:

$$B = \begin{pmatrix} B_{0,0} & B_{0,1} & \cdots & B_{0,r-1} & B_{0,r} \\ B_{1,0} & B_{1,1} & \cdots & B_{1,r-1} & B_{1,r} \\ 0 & B_{2,1} & \cdots & B_{2,r-1} & B_{2,r} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_{r,r-1} & B_{r,r} \end{pmatrix},$$

- $B_{i,j} \in \mathcal{M}^{p_i \times p_j}$
- $B_{j,j-1}$ have rank p_j
- $p_0 \geq p_1 \geq \cdots \geq p_r \geq 1$
- $p_0 + p_1 + \cdots + p_r = d.$

Remark 2.37. The Hölder spaces $C_{B,loc}^{n,\alpha}$ are defined exactly as in 1.14.

In this setting, the following theorem holds true:

Theorem 2.38. *Let Ω be a domain of $\mathbb{R} \times \mathbb{R}^d$, $\alpha \in]0,1]$ and $n \in \mathbb{N}_0$. If $u \in C_{B,loc}^{n,\alpha}(\Omega)$ then it holds:*

1) there exist

$$Y^k \partial_x^\beta u \in C_{B,loc}^{n-2k-|\beta|_B,\alpha}(\Omega), \qquad 0 \leq 2k + |\beta|_B \leq n;$$

2) for any $\zeta_0 \in \Omega$, there exist two bounded domains U, V, such that $\zeta_0 \in U \subset \bar{V} \subset \Omega$ and

$$|u(z) - \mathcal{T}_n(u,\zeta)(z)| \leq c_{B,U} \|u\|_{C_B^{n,\alpha}(V)} \|\zeta^{-1} \circ z\|_B^{n+\alpha}, \qquad z, \zeta \in U,$$

*where $c_{B,U}$ is a positive constant and $\mathcal{T}_n(u,\zeta)$ is the n-*th order intrinsic Taylor polynomial of u centered in ζ *as defined in* (2.4).

Remark 2.39. The estimate above is the same as in (2.41) except for the constant c_B which is replaced by $c_{B,U}$. To explain this discrepancy between the two results we have to go back to Lemma 1.9: in the non homogeneous case in fact, the triangular inequality for $\left\|\zeta^{-1} \circ z\right\|_B \leq c(\|\zeta\|_B + \|z\|_B)$ is not available globally but the points z, ζ must lie in a fixed domain. Moreover, the constant used there depends on such domain.

Remark 2.40. It is worth to note that the non homogeneous Taylor polynomial share the same formal expression with the homogeneous one even though, strictly speaking, it is not a polynomial function of time. This is coherent with the group law expression in this case: in fact, being the matrix B not nilpotent, the exponential matrix $e^{(s-t)B}$ has analytic rather than polynomial entries.

Example 2.41. Consider the case

$$B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix},$$

then, if $z = (t, x, y)$ and $\zeta = (s, \xi, \eta)$, we have

$$\zeta^{-1} \circ z = (s - t, x - e^{t-s}\xi, y - \eta - (e^{t-s} - 1)\xi).$$

Proof of Theorem 2.38

For sake of brevity, we only prove the statement for $r = 1$, which is $B = (B_{i,j})_{i,j\in\{0,1\}}$ with $B_{i,j} \in \mathcal{M}^{p_i\times p_j}$ and $B_{1,0}$ has full rank. This case is complex enough to see the conceptual difficulties that arise from dropping the homogeneity assumption On the other hand, the proof for a general $r \geq 1$ is only a lengthy and technical extension.

Notation 2.42. Throughout this section we will use the notation $z = (t, x, y)$ or $\zeta = (s, \xi, \eta)$ to indicate a general element of $\mathbb{R} \times \mathbb{R}^{p_0} \times \mathbb{R}^{p_1}$. Moreover, we will denote by c any positive constant that depends on B and on the domain U in Theorem 2.38, at most.

As in the homogeneous case, the first task is connecting two points in $\mathbb{R} \times \mathbb{R}^{p_0} \times \mathbb{R}^{p_1}$ using integral curves. To obtain an increment in the x-variables it is enough to move along the integral curves of the fields $X_1, \cdots, X_{p_0}$, i.e.

$$\gamma^{(0)}_{v,\delta}(t, x, y) := (t, x + \delta v, y), \qquad v \in \mathbb{R}^{p_0},\ \delta \in \mathbb{R}.$$

To understand how to obtain an increment in the y-variables, it is useful to observe that

$$[v_1X_1 + \cdots + v_{p_0}X_{p_0}, Y] - \langle\nabla_x, B_{0,0}v\rangle = \langle\nabla_y, B_{1,0}v\rangle, \qquad v \in \mathbb{R}^{p_0}. \tag{2.42}$$

Compare with (2.8) where choosing a vector $v \in \mathbb{R}^{p_0}$ was sufficient to obtain a derivative of the desired intrinsic order. Here instead we need to correct the presence of a *lower* order derivative. It is thus reasonable to approximate the integral curves of the vector field on the right-hand side as

$$\begin{aligned}\gamma_{v,\delta}(t,x,y) &:= \gamma^{(0)}_{B_{0,0}v,-\delta^3}\left(e^{-\delta^2 Y}\left(\gamma^{(0)}_{v,-\delta}\left(e^{\delta^2 Y}\left(\gamma^{(0)}_{v,\delta}(t,x,y)\right)\right)\right)\right) \qquad (2.43)\\ &= \left(t,x,y+\delta^3 B_{1,0}v\right) - \delta^5\left(0,\sum_{n=0}^{\infty}\frac{(-1)^n\delta^{2n}}{(n+2)!}B^{n+2}(v,0)^*\right),\end{aligned}$$

where $v\in\mathbb{R}^{p_0}$, and $\delta\in\mathbb{R}$.The leading order increment is proportional to δ^3, along the y variable only. However, due to the non-homogeneous structure of B (the block $B_{0,0}$ is not null), the higher order increment affects both the components x and y. To correct this, we employ again the curve $\gamma^{(0)}$. Set

$$g_{v,\delta}(t,x,y) := \gamma^{(0)}_{v',\delta'}\left(\gamma_{v,\delta}(t,x,y)\right), \qquad v\in\mathbb{R}^{p_0},\ \delta\in\mathbb{R}, \qquad (2.44)$$

where

$$v' = v'(\delta,v) = \sum_{n=0}^{\infty}\frac{(-1)^n\delta^{2n}}{(n+2)!}B^{n+2}_{0,0}v, \quad \delta' = \delta'(\delta) = \delta^5, \qquad (2.45)$$

and $B^{n+2}_{0,0}$ is the top-left $(p_0\times p_0)$-submatrix of B^{n+2}.

The following lemma is crucial, as it allow us to obtain a (small) increment in the y direction using integral curves.

Lemma 2.43. *There exists $\varepsilon>0$, only dependent on B, such that: for any $\eta\in\mathbb{R}^{p_1}$ with $|\eta|\le\varepsilon$, there exist $v\in\mathbb{R}^{p_0}$ with $|v|=1$ and $\delta\ge 0$ such that*

$$g_{v,\delta}(t,x,y) = (t,x,y+\eta), \qquad \text{and} \qquad |\delta|\le c|\eta|^{\frac{1}{3}}. \qquad (2.46)$$

Proof. By (2.44) and (2.43) we obtain

$$g_{v,\delta}(t,x,y)-(t,x,y) = \left(0,0,\delta^3 R(\delta,v)\right), \quad R(\delta,v) := \sum_{n=0}^{\infty}\frac{(-1)^n\delta^{2n}}{(n+1)!}B^{n+1}_{1,0}v,$$

where $B^{n+1}_{1,0}$ is the bottom-left $(p_1\times p_0)$-submatrix of B^{n+1}. Therefore, denoting by $\mathbb{S}^{p_0-1}$ the unitary sphere in $\mathbb{R}^{p_0}$, we have to find some $(\delta,v)\in[0,\infty[\times\mathbb{S}^{p_0-1}$ that solves the equation

$$\delta^3 R(\delta,v) = \eta. \qquad (2.47)$$

Since $B_{1,0}$ has full rank it is not restrictive to assume $p_0=p_1$, and thus $B_{1,0}$ invertible. In particular, $R(0,v)=B_{1,0}v$, which implies that $R(0,\cdot)$ is a bijective and linear function. Moreover, since R is globally C^1, there exists $\bar\delta>0$ such that $R(\delta,\cdot)$ is still a bijective linear function for any $\delta\le\bar\delta$. In particular, when restricted to $\mathbb{S}^{p_0-1}$, $(\delta^3 R(\delta,\cdot))_{0\le\delta\le\bar\delta}$ is a continuous family of embeddings that collapses to zero at $\delta = 0$. By Theorem A.1 in Appendix A equation (2.47) admits a solution $\big(\delta(\eta), v(\eta)\big) \in [0,\bar\delta] \times \mathbb{S}^{p_0-1}$ for any $|\eta| \le \varepsilon$, where $\varepsilon > 0$ only depends on B. We now prove the second part of (2.46). Choosing ε small enough, it holds $\big|R\big(\delta(\eta), v(\eta)\big)\big| \in [\|B_{1,0}\| - \varepsilon, \|B_{1,0}\| + \varepsilon]$, and by (2.47),

$$\big|\delta(\eta)\big|^3 = \frac{|\eta|}{\big|R\big(\delta(\eta), v(\eta)\big)\big|} \le \frac{|\eta|}{\max(0, \|B_{1,0}\| - \varepsilon)}.$$

Again, taking ε suitably small yields the result. □

We are now in the position to prove Theorem 2.38.

Proof of Theorem 2.38. Analogously to the homogeneous setting, the cases $n = 0, 1, 2, 3$ have to be proved separately, while for $n > 3$ the proof is by induction on n. For sake of brevity, here we only provide a proof for $n = 0$ and $n = 3$, these being the most interesting and difficult steps. On the one hand, the proof for $n = 0$ allows to appreciate how the connection Lemma 2.43 along with the regularity along the fields can be used, in a rather simple way, in order to obtain the most basic result, namely the Hölderianity with respect to the B-intrinsic norm. On the other hand, the proof for $n = 3$ enlightens the main difficulty of the whole proof, namely proving the existence of the first order partial derivative w.r.t. the y-variable. Note that the existence of the latter is not trivially ensured by the definition of $C^{3,\alpha}_B(\Omega)$, as the existence of $X_i Y u$ and $Y X_i u$, and thus the commutators $[X_i, Y]u$, are only meant in the sense of Lie derivatives. As for the steps $n = 1$, $n = 2$, these are just simplifications of the case $n = 3$, whereas the inductive step for $n > 3$ is totally analogous to the homogeneous case.

<u>Case $n = 0$</u>: We only need to prove Part 2). Let $U \subset \Omega$ be a bounded domain suitably small so as to ensure that all the integral curves that are employed below to connect z and ζ are entirely contained in the bounded domain $V \subset \Omega$. The first step is to bound the increment w.r.t. the time variable by employing the integral curve of Y in (1.20):

$$\begin{aligned} |u(t,x,y) - u(s,\xi,\eta)| &\le \big|u(t,x,y) - u\big(e^{(t-s)Y}(s,\xi,\eta)\big)\big| \\ &\quad + \big|u\big(e^{(t-s)Y}(s,\xi,\eta)\big) - u(s,\xi,\eta)\big| \\ &\le \big|u(t,x,y) - u\big(t, e^{(t-s)B}(\xi,\eta)^*\big)\big| + c\,\|u\|_{C^\alpha_Y(V)}\,|s-t|^{\frac{\alpha}{2}}, \end{aligned}$$

where we used triangular inequality in the first line and $u \in C^{\alpha}_{Y}(\Omega)$ in the second. Note that $\left|(x,y)^{*} - e^{(t-s)B}(\xi,\eta)^{*}\right|^{\alpha}_{B} \le \|\zeta^{-1} \circ z\|^{\alpha}_{B}$ and thus we only need to prove

$$|u(t,\xi,\eta) - u(t,x,y)| \le c\, \|u\|_{C^{0,\alpha}_{B}(V)} \,(|\xi - x| + |\eta - y|^{\frac{1}{3}})^{\alpha}. \tag{2.48}$$

We can use again triangular inequality and write

$$\begin{aligned} u(t,\xi,\eta) - u(t,x,y) &= \big(u(t,\xi,\eta) - u(t,x,\eta)\big) + \big(u(t,x,\eta) - u(t,x,y)\big) \\ &= \big(u(t,\xi,\eta) - u(t,x,\eta)\big) + \Big(u\big(g_{\delta,v}(t,x,y)\big) - u(t,x,y)\Big) \end{aligned}$$

with $|v| = 1$ and $|\delta| \le c|\eta - y|^{\frac{1}{3}}$. By using $u \in C^{\alpha}_{X_i}(\Omega)$, $i = 1,\dots,p_0$, in order to bound the first term, together with $u \in C^{\alpha}_{Y}(\Omega)$ to bound the second, we obtain (2.48), which concludes the proof for $n = 0$.

<u>*Case $n = 3$*</u>: To shorten notation we only prove the case $p_0 = p_1 = 1$. The difficulty in considering multi-dimensional blocks is purely notational. We first prove Part 1). Fix an arbitrary bounded domain $\Omega_0 \subset \Omega$. Proceeding as in the homogeneous case, one obtains

$$\left|u\big(\gamma^{(i)}_{\delta,v}(z)\big) - \bar{T}_3 u\big(z,\gamma^{(i)}_{\delta,v}(z)\big)\right| \le c\, \|u\|_{C^{3,\alpha}_{B}(\Omega_0)} |\delta|^{3+\alpha}, \qquad i = 0,1, \tag{2.49}$$

for any $z \in \Omega_0$, and $v, \delta \in \mathbb{R}$ with $|v| = 1$ and $|\delta|$ suitably small, where we set

$$\bar{\mathfrak{T}}_3(u,z)(\zeta) = \sum_{i=0}^{3} \frac{(\xi - x)^i}{i!} \partial^i_x u(z) + \frac{\eta - y}{B_{1,0}} \big([\partial_x, Y] - B_{0,0}\partial_x\big) u(z),$$

and $z = (t,x,y)$, $\zeta = (t,\xi,\eta)$. The last term in the right-hand side is inspired by (2.42) to mimic a partial derivative w.r.t. y and is well defined when applied to $u \in C^{3,\alpha}_{B}(\Omega)$. We now prove

$$\left|u\big(g_{\delta,v}(z)\big) - \bar{\mathfrak{T}}_3(u,z)\big(g_{\delta,v}(z)\big)\right| \le c\, \|u\|_{C^{3,\alpha}_{B}(\Omega_0)} |\delta|^{3+\alpha}, \tag{2.50}$$

where $g_{v,\delta}$ is as defined in (2.44)-(2.45). Setting $z' := \gamma_{\delta,v}(z)$ and $z'' = g_{v,\delta}(z)$ we have

$$\begin{aligned} u\big(z''\big) - \bar{\mathfrak{T}}_3(u,z)\big(z''\big) &= F_1 + F_2, \\ F_1 &= \big(u(z'') - \bar{\mathfrak{T}}_3(u,z')\big(z''\big)\big) + \big(u(z') - \bar{\mathfrak{T}}_3(u,z)\big(z'\big)\big)\,, \\ F_2 &= \bar{\mathfrak{T}}_3(u,z')\big(z''\big) + \bar{\mathfrak{T}}_3(u,z)\big(z'\big) - u(z') - \bar{\mathfrak{T}}_3(u,z)\big(z''\big). \end{aligned}$$

Now, (2.49) and (2.45) yield $|F_1| \le c\,\|u\|_{C^{3,\alpha}_B(\Omega_0)}|\delta|^{3+\alpha}$; as for F_2 it holds:

$$|F_2| = \left|\sum_{i=1}^{3} \frac{1}{i!}\left(\partial_x^i u(z') - \partial_x^i u(z)\right)(\delta' v')^i\right| \le c\,\|u\|_{C^{3,\alpha}_B(\Omega_0)}|\delta' v'| \le c\,\|u\|_{C^{3,\alpha}_B(\Omega_0)}|\delta|^{3+\alpha},$$

where we used $\partial_x^i u \in C^{3-i,\alpha}_B(\Omega)$ and Theorem 2.38 for $n = 0, 1, 2$, to prove the first inequality, and (2.45) to prove the second one. This proves (2.50). We are now able to prove differentiability along the y direction. For any $z = (t, x, y) \in \Omega_0$ and $\eta \in \mathbb{R}$ with $|\eta|$ small enough, choosing v and δ as given by Lemma 2.43 yields

$$\begin{aligned}
\big|u(t,x,y+\eta) - &\bar{\mathfrak{T}}_3\big(u,(t,x,y)\big)(t,x,y+\eta)\big| \\
&= \big|u\big(g_{v,\delta}(t,x,y)\big) - \bar{\mathfrak{T}}_3\big(u,(t,x,y)\big)\big(g_{\delta,v}(t,x,y)\big)\big| \\
&\le c\|u\|_{C^{3,\alpha}_B(\Omega_0)}|\delta|^{3+\alpha} \le c\|u\|_{C^{3,\alpha}_B(\Omega_0)}|\eta|^{1+\frac{\alpha}{3}},
\end{aligned}$$

where we used (2.50) in to obtain the first inequality, and (2.46) to obtain the second. Thus $\partial_y u(z)$ exists and

$$\partial_y u(t,x,y) = \frac{1}{B_{1,0}}\big([\partial_x, Y] - B_{0,0}\partial_x\big)u(t,x,y).$$

Furthermore, $u \in C^{3,\alpha}_B(\Omega)$ implies $\partial_y u \in C^{0,\alpha}_B(\Omega)$, which is Part 1) of Theorem 2.38 for $n = 3$.

The proof of Part 2) is analogous to the homogeneous case.

Chapter 3

Analytical Expansions and Error Estimates

In this chapter we apply the results obtained in the previous one to prove a error estimate on an asymptotic expansion of the conditional expectation

$$u(t,x) := \mathbb{E}_{t,x}[\varphi(X_T)], \tag{3.1}$$

where $X = (X_t)_{t\in[0,T]}$ is a continuous $\mathbb{R}^d$-valued Feller process and a degenerate diffusion in the sense that the operator $\partial_t + \mathcal{A}_X$, where $\mathcal{A}_X$ is the infinitesimal generator of X, is a Kolmogorov Operator.

The prototype process we have in mind is $X = (S, A)$ solution to the SDE

$$\begin{cases} dS_t = \sigma S_t dW_t, \\ dA_t = S_t dt, \end{cases} \tag{3.2}$$

where W is a real Brownian motion. In financial applications, S and A represent the price and average processes respectively, in the Black&Scholes model for arithmetic Asian options. The infinitesimal generator of (S, A)

$$\mathcal{A}_X := \frac{\sigma^2 s^2}{2}\partial_{ss} + s\partial_a, \qquad (s,a) \in \mathbb{R}_{>0} \times \mathbb{R}_{>0},$$

is degenerate in two ways: on the one hand, the quadratic form of the second order part is singular (it has rank one) and, on the other hand, it degenerates completely on the half-line $\{s = 0, a > 0\}$. However, for any $0 < a < b$, $\mathcal{A}_X$ is a hypoelliptic operator on the strip $D :=]a, b[\times \mathbb{R}_{>0}$ and coincides on D with an operator that satisfies the Hörmander condition globally, the latter obtained by smoothly perturbing the second order coefficient $\sigma^2 s^2$ outside D. By performing a local analysis, we aim at exploiting this fact to prove error estimates, uniform w.r.t. $x = (s, a) \in D$, for the intrinsic asymptotic expansions of the conditional expectation in (3.1).

In general, we assume that the infinitesimal generator of X *coincides, on a domain* D of $\mathbb{R}^d$, with a differential operator of the form

$$\mathcal{A} = \frac{1}{2}\sum_{i,j=1}^{p_0} a_{ij}(t,x)\partial_{x_i x_j} + \sum_{i=1}^{p_0} a_i(t,x)\partial_{x_i} + \langle Bx, \nabla_x\rangle \tag{3.3}$$

where $(t, x) \in \mathbb{R} \times \mathbb{R}^d$, $p_0 \le d$ and $\mathcal{A}$ verifies the following

Assumption 3.44. $A_0 := \big(a_{ij}(t, x)\big)_{i,j=1,\cdots,p_0}$ satisfies the non-degeneracy condition

$$\mu M|\xi|^2 < \sum_{i,j=1}^{p_0} a_{ij}(t,x)\xi_i\xi_j < M|\xi|^2, \quad (t,x) \in \mathbb{R}\times\mathbb{R}^d,\ \xi \in \mathbb{R}^{p_0}, \tag{3.4}$$

for some positive constants M and μ.

We also recall the homogeneity assumption 1.3 made in Chapter 1:

Assumption 3.45. B is a $(d \times d)$-matrix with constant entries of the form

$$B = \begin{pmatrix} 0 & 0 & \cdots & 0 & 0 \\ B_1 & 0 & \cdots & 0 & 0 \\ 0 & B_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_r & 0 \end{pmatrix}$$

where each B_j is a $(p_j \times p_{j-1})$-matrix of rank p_j and

$$p_0 \ge p_1 \ge \cdots \ge p_r \ge 1, \qquad \sum_{j=0}^{r} p_j = d.$$

In this way, all the results obtained in Chapter 1 are available.

Under suitable regularity conditions that will be specified later, the ultra-parabolic operator

$$\mathcal{K} := \mathcal{A} + \partial_t \tag{3.5}$$

admits a fundamental solution (see Polidoro (1994) and Di Francesco and Pascucci (2005)). In the case $p_0 < d$, which is the focus of this work, this is a remarkable fact as the second order part of $\mathcal{A}$ is fully degenerate at any point.

Our analysis takes advantage of the intrinsic geometry and the related regularity structures induced by the Kolmogorov operator $\mathcal{K}$ and studied in the previous chapters. These features bring a number of benefits that are explained here below, and distinguish our approach from others in the literature. It is worth to emphasize further that our results are carried out under strictly *local assumptions* on the generator of X, which coincides with a Kolmogorov operator on a domain D, not necessarily equal to $\mathbb{R}^d$. This allows to include degenerate models with relevant financial applications, such as the well-known CEV model (that is when σ in (3.2) is not a constant but a function of S of the form $\sigma(S) = S^\gamma$ for some $\gamma \in \mathbb{R}$) and the Heston stochastic volatility model as very particular cases. The proof of our main result, Theorem 3.59, will be split in two separate steps: first, in Theorem 3.62, we consider the case $D = \mathbb{R}^d$ for which we employ some Gaussian upper bounds for the transition density of X; second, we adapt a localization procedure, originally introduced in Safonov (1998) and lately extended in Cinti and Polidoro (2009), which is based on the Gaussian bounds for a dummy diffusion $\widetilde{X}$ that is generated by $\mathcal{A}$ in (3.3). The latter localization procedure is coherent with what is known in the theory of diffusions as the *principle of not feeling the boundary* (cf. Hsu (1995), Gatheral et al. (2012)).

Taylor expansion forms the cornerstone of the perturbation technique that we study in this chapter. Here below we summarize the intuitive idea behind it and its primary features.

We recall that, under mild assumptions that will be specified in Section 3.2, the function u in (3.1) satisfies

$$\begin{cases} \mathcal{K}u = 0, & \text{on } [0, T[\times D, \\ u(T, \cdot) = \varphi, & \text{on } D. \end{cases} \tag{3.6}$$

Notice that (3.6) is not a standard Cauchy-Dirichlet problem since no lateral boundary conditions are imposed. In Lorig et al. (2015), Pagliarani and Pascucci (2014), the authors propose a perturbative method to carry out a closed-from approximation of solutions to (3.6) under the assumption that $\mathcal{K}$ in (3.3)-(3.5) is locally parabolic, i.e. $p_0 = d$ and $B = 0$ in (1.6) The basic idea is to approximate the generator by Taylor expanding its coefficients, and take advantage of some symmetry properties of Gaussian kernels. Sharp short-time/small-noise asymptotic estimates for the remainder of the expansion are then proved. In order to generalize the aforementioned technique to the case $p_0 < d$, we perform an expansion that is compatible with the sub-elliptic geometry induced by Kolmogorov operators. Assuming $a_{ij}, a_i \in C_B^{N,1}$, we expand the operator $\mathcal{K}$ through the sequence $\left(\mathcal{K}_n^{(\bar{z})}\right)_{0\leq n\leq N}$ defined as

$$\mathcal{K}_n^{(\bar{z})} = \frac{1}{2}\sum_{i,j=1}^{p_0} \mathcal{T}_n\left(a_{ij},\bar{z}\right)(z)\partial_{x_i x_j} + \sum_{i=1}^{p_0} \mathcal{T}_{n-1}\left(a_i,\bar{z}\right)(z)\partial_{x_i} + Y, \tag{3.7}$$

for $0 \leq n \leq N$. As in the previous chapter, $\mathcal{T}_n\left(a_{ij},\bar{z}\right)(z)$ is the Taylor polynomial of a_{ij}, defined as in (2.4), centered at a fixed point $\bar{z} \in \mathbb{R}\times\mathbb{R}^d$, and calculated in z, Y is the vector field $\langle Bx, \nabla\rangle + \partial_t$ and, by convection, $\mathcal{T}_{-1} \equiv 0$.

We explicitly remark that leading term of the expansion, the operator

$$\mathcal{K}_0^{(\bar{z})} = \frac{1}{2}\sum_{i,j=1}^{p_0} a_{ij}(\bar{z})\partial_{x_i x_j} + Y, \tag{3.8}$$

is a Kolmogorov operator with constant coefficients in the form (1.1)defined on $\mathbb{R}\times\mathbb{R}^d$. It is well-known that $\mathcal{K}_0^{(\bar{z})}$ admits a Gaussian fundamental solution (see equation (3.11) for the precise expression) that satisfies some remarkable symmetry properties written in terms of the increments appearing in the intrinsic Taylor polynomials in (2.4). The main result of this chapter, Theorem 3.59, provides an explicit approximating expansion for $u(t,x)$ in (3.1), equipped with sharp short-time error bounds, and can be roughly summarized as:

$$u(t,x) = u_0(t,x) + \sum_{n=1}^{N} \mathcal{L}_n(t,T,x)\, u_0(t,x) + \mathrm{O}\left((T-t)^{\frac{N+1+k}{2}}\right) \qquad \text{as } t \to T^-, \tag{3.9}$$

uniformly with respect to $x \in D$, where:

- the leading term u_0 is the solution of the Cauchy problem for $\mathcal{K}_0^{(\bar{z})}$ with final datum φ;

- $(\mathcal{L}_n)_{1\leq n\leq N}$ is a family of differential operators, acting on x, that can be explicitly computed in terms of the intrinsic Taylor polynomials $\mathcal{T}_n\left(a_{ij},\bar{z}\right)$ and $\mathcal{T}_n\left(a_i,\bar{z}\right)$ (see Theorem 3.55);

- the positive exponent k, contributing to the asymptotic rate of convergence, is the intrinsic Hölder exponent of φ. Precisely, $\varphi \in C_B^k$ according to Definition 3.47 below.

Such approximation turns out to be optimal to several extents. In particular, the benefit in exploiting the intrinsic regularity is threefold: first, since the intrinsic Taylor polynomial has a shorter expression than the Euclidean one (see Section 2.1) we avoid taking up terms in the expansion that do not improve the quality of the approximation; secondly, the fact that the increments of the intrinsic Taylor polynomial appear in the symmetries of the fundamental solution of $\mathcal{K}_0^{(\bar{z})}$ allows to get compact approximation formulas; finally, the asymptotic rate of convergence of the expansion also depends on the *intrinsic regularity* of the datum φ, which is typically higher than the Euclidean regularity. See Remark 3.46 below).

3.1 Applications to finance and comparison with the existing literature

The application of Kolmogorov operators in mathematical finance is particularly relevant in the pricing of Asian-style derivatives. These are financial claims whose payoff is a function not only of the terminal value of an underlying asset, but also of its average over a certain time-period. In most cases of interest, the problem of computing the conditional expectation (3.1), which defines the no-arbitrage price of such financial claims, is not known to have an explicit solution, and thus a considerably large amount of literature has been developed in the last decades in order to find accurate and quickly computable approximate solutions. Some of these approaches make use of asymptotic techniques that lead to semi-closed approximation formulas. In this section we aim at firming our results within the existing literature on analytical approximations of Asian-style derivatives. Before to proceed we recall that other financial applications, where averaged-diffusion processes are employed, include volatility models with path-dependent coefficients, e.g. the Hobson-Rogers model Hobson and Rogers (1998).

Let us resume our first example (3.2) and now assume that S follows the more general dynamics

$$dS_t = \sigma\left(t, S_t, A_t\right) dW_t.$$

In this case, $a_{11}(t, x_1, x_2) = \sigma^2(t, x_1, x_2)$ and we recall that increments along time in the intrinsic Taylor polynomials only shows from the 2nd order on, whereas the increment along the average variable appears from the 3rd order on. As it was mentioned above, the operators $\mathcal{L}_n^{(\zeta)}$ appearing in the asymptotic expansion in (3.9) can be explicitly computed by applying (3.21)-(3.22)-(3.19)-(3.12). In this case $\mathcal{L}_n^{(\zeta)}(t, T, x)$ reads as

$$\frac{1}{2}\int_t^T \big(\mathcal{T}_n(a_{11}, \zeta) - \mathcal{T}_{n-1}(a_{11}, \zeta)\big)\big(s, \mathcal{M}^{(\zeta)}(s-t, x_1, x_2)\big)\big(\partial_{x_1} - (s-t)\partial_{x_2}\big)^2 ds,$$

while $\mathcal{M}^{(\zeta)}(t, x_1, x_2)$ is the one by two vector

$$\Big(x_1 + a_{11}(\zeta)t\partial_{x_1} - a_{11}(\zeta)\frac{t^2}{2}\partial_{x_2}\ ,\ tx_1 + x_2 - a_{11}(\zeta)\frac{t^2}{2}\partial_{x_1} + a_{11}(\zeta)\frac{t^3}{6}\partial_{x_2}\Big).$$

In order to show an even more explicit sample, at order 1 we have:

$$\mathcal{L}_1^{(\zeta)}(t, T, x) = \frac{\partial_{\xi_1} a_{11}(\zeta)}{2}$$
$$\times \int_t^T \Big((x_1 - \xi_1) + a_{11}(\zeta)(s-t)\partial_{x_1} - \frac{a_{11}(\zeta)}{2}(s-t)^2\partial_{x_2}\Big)\big(\partial_{x_1} - (s-t)\partial_{x_2}\big)^2 ds.$$

Two typical arithmetic Asian options are the so-called *floating strike* and *fixed strike* Call options, whose payoffs are given respectively by

$$\varphi_{\text{float}}(x_1, x_2) = \big(x_1 - x_2/T\big)^+, \qquad \varphi_{\text{fixed}}(x_1, x_2) = \big(x_2/T - K\big)^+,$$

where T is the maturity and K is the strike price.

Remark 3.46. The payoff φ_{fixed} is Lipschitz continuous in the standard Euclidean sense but has higher intrinsic regularity (namely, C_B^3 according to Definition 3.47, see also Example 3.49): this property reflects a higher rate of convergence of the asymptotic expansion (3.9) compared with other expansions based on standard Taylor polynomials. On the other hand, because of its explicit dependence on x_1, the payoff φ_{float} is only $C^1_{B,\text{loc}}$.

Even in the simplest case of constant volatility, i.e. in the Black&Scholes model, both the marginal distribution of A_t and the joint distribution of (S_t, A_t)

are difficult to characterize analytically. The distribution of A_t was given an integral representation in the pioneering work Yor (1992), though that result is of limited practical use in the valuation of Asian options. The approximation formulas that we propose in this chapter were applied heuristically in Foschi et al. (2013), where intensive numerical tests were performed to confirm their accuracy. However, the general hypoelliptic framework that we consider here clearly allows for several generalization, including more general dynamics and more sophisticated Asian style-derivatives including stochastic local volatility models such as the CEV and the Heston models Heston (1993). An interesting example is also given by a generalized type of Asian option, where the average is weighted w.r.t. the volume of traded assets: these options are written on the *Volume Weighted Average Price* (VWAP), a trading benchmark used especially in pension plans (see, for instance, Novikov et al. (2014)). The dynamics of the traded volume V are lead by an additional stochastic factor that has to be chosen as to reflect the corresponding volume statistics, and the average process A is then given by

$$A_t = \frac{\int_0^t S_\tau V_\tau d\tau}{\int_0^t V_\tau d\tau}.$$

As it was previously argued, our technique makes use of the intrinsic Taylor polynomials in (2.4) in order to be consistent with the subelliptic geometry induced by Kolmogorov operators. This differentiates our approach from others appearing in the literature that are based on classical Euclidean expansions. In the relevant paper Gobet and Miri (2014), Malliavin calculus techniques were employed to derive analytical approximations for the law of a general averaged diffusion. When applied to the pricing of arithmetic Asian options, the approach in Gobet and Miri (2014) returns an expansion whose leading term is the price of a geometric Asian option. Correcting terms are computed by Taylor expanding the coefficients of the diffusion and error estimates depend on standard Euclidean regularity assumptions on the coefficients and on the payoff function. In Tsao et al. (2003) and Chung et al. (2003), the authors followed a different approach and carried out a Taylor based-expansion of the joint distribution (S_t, A_t) to analytically approximate the price of an Asian option (possibly, forward-starting); this technique seems to be limited to the Black&Scholes dynamics. Other approximations, based on Taylor expansions and on Watanabe's theory, can be found in Kunitomo and Takahashi (1992), though no rigorous error bounds are provided.

For sake of completeness, we also give a brief, and by no means exhaustive, overview of the existing literature concerning other approaches to the pricing of Asian options. Within the Black&Scholes framework, Geman and Yor (1992) derived an analytical expression for the Laplace transform of A_t. However, several authors pointed out some stability issues related to the numerical inversion of the Laplace transform, which lacks accuracy and efficiency in regimes of small volatility or short time-to-maturity. This is also a disadvantage of the Laguerre expansion proposed in Dufresne (2000). Shaw (2003) used a contour integral approach based on Mellin transforms to improve the accuracy of the results in the case of low volatilities, albeit at a higher computational cost. As opposed to numerical inversion, Linetsky (2004) derived an eigenfunction expansion of the transition density of A_t (see also Donati-Martin et al. (2001)) by employing spectral theory of singular Sturm-Liouville operators. Although it returns in general very accurate results, Linetsky's series formula may converge slowly in the case of low volatility and become computationally expensive. Note that, by opposite, the analytical pricing formulas we propose here do not suffer any lack of accuracy or efficiency in these limiting cases. In actual fact, Theorem 3.59 and Remark 3.61 show that the accuracy improves as volatility and/or time to maturity get smaller. Again in the particular case of the Black&Scholes model, and for special homogeneous payoff functions, it is possible to reduce the pricing PDE in (3.6) to a one state variable PDE. PDE reduction techniques were initiated in Ingersoll (1987) and applied to the problem of pricing Asian options by several authors, including Rogers and Shi (1995); Vecer (2001) and Dewynne and Shaw (2008). Eventually, other approaches include the parametrix expansion in Corielli et al. (2010) and the moment-matching techniques in Dufresne (2001b); Deelstra et al. (2010); Fusai and Tagliani (2002) and Forde and Jacquier (2010) among others.

We consider the prototype Kolmogorov operator obtained by (3.3)-(3.5) with A_0 equal to a scalar $(p_0 \times p_0)$-matrix and $a_i \equiv 0$, $i = 1, \dots, p_0$, i.e.

$$\mathcal{K}^{\Lambda} := \frac{\Lambda}{2} \sum_{i=1}^{p_0} \partial_{x_i}^2 + \langle Bx, \nabla_x \rangle + \partial_t, \quad (t,x) \in \mathbb{R} \times \mathbb{R}^d, \quad \Lambda > 0. \tag{3.10}$$

In this case we say that $\mathcal{K}^{\Lambda}$ is a *constant coefficients Kolmogorov operator*. By Assumption 1.3, the vector fields $\partial_{x_1}, \dots, \partial_{x_{p_0}}$ and Y in (1.3) satisfy the Hörmander's condition and therefore $\mathcal{K}^{\Lambda}$ is hypoelliptic. Definition 1.14 and

Theorem 2.20 will be used in the next section, respectively, to specify suitable regularity conditions on the coefficients of $\mathcal{K}$ in (3.3)-(3.5), and to expand them as in (3.7). However the intrinsic regularity of the terminal datum φ plays as well a key role in the error analysis of the expansion (3.9). This motivates the following

Definition 3.47. Let $k \in]0, 2r+1]$. We denote by $C_B^k(\mathbb{R}^d)$ the space of functions φ on $\mathbb{R}^d$ such that

$$|\varphi(x) - \varphi(y)| \leq C[x-y]_B^k, \qquad x, y \in \mathbb{R}^d,$$

for some positive constant C, where $[\cdot]_B$ is the norm on $\mathbb{R}^d$ defined in (1.17). We also set

$$\|\varphi\|_{C_B^k(\mathbb{R}^d)} = \sup_{x \neq y} \frac{|\varphi(x) - \varphi(y)|}{[x-y]_B^k}.$$

Moreover, by convention, $C_B^0(\mathbb{R}^d)$ is the set of bounded and continuous functions on $\mathbb{R}^d$ and $\|\varphi\|_{C_B^0(\mathbb{R}^d)} = \|\varphi\|_{L^\infty(\mathbb{R}^d)}$.

Remark 3.48. It should be noted that, if in the previous definition $k > 2n+1$ for a $n = 0, 1, \dots, r-1$ then φ is in fact a function of just the variables $x_{\bar{p}_n+1}, \dots, x_d$. This is an effect of the different weights in the definition of the homogeneous norm $[\cdot]_B$. However, it is essentially the analogous of the fact that the only Euclidean Hölder functions of order greater than one are constant.

Conversely, any $\varphi \in C^\alpha(\mathbb{R}^{p_n})$ can be automatically extended to a function $\widetilde{\varphi} \in C_B^{(2n+1)\alpha}(\mathbb{R}^d)$ via

$$\widetilde{\varphi}(x) := \varphi(x^{[n]}), \qquad x \in \mathbb{R}^d$$

where we used definition 1.5. The following example is a particularly important instance of this procedure.

Example 3.49. Consider the case of arithmetic Asian options with fixed strike discussed in Section 3.1, i.e.

$$B = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \qquad \varphi_{\text{fixed}}(x_1, x_2) = (x_2/T - K)^+ .$$

According to Definition 3.47, $\varphi_{\text{fixed}} \in C_B^3(\mathbb{R}^2)$ even if it is only Lipschitz continuous in the Euclidean sense.

3.2 Approximate solutions and error bounds

Let X be a Feller process as defined in the introduction of the chapter: in particular, we assume that the infinitesimal generator of X coincides with operator $\mathcal{A}$ in (3.3) on a fixed domain D of $\mathbb{R}^d$. Moreover, $\mathcal{A}$ satisfies Assumptions 3.44 and 1.3. Throughout this section $N \in \mathbb{N}_0$ and $T > 0$ are fixed and we also require the following assumptions to be in force:

Assumption 3.50. The coefficients a_{ij}, a_i of $\mathcal{A}$ belong to $\in C_B^{N,1}$ and

$$\|a_{ij}\|_{C_B^{N,1}}, \|a_i\|_{C_B^{N,1}} \leq M,$$

with M as in (3.4).

Assumption 3.51. The final datum φ is a continuous function with sub-exponential growth such that $u = u(t,x)$ in (3.1) is well defined and belongs to $L^\infty([0,T] \times D)$. Moreover, there exists $\psi \in C_B^k(\mathbb{R}^d)$, with $k \in [0, 2r+1]$, such that $\varphi = \psi$ on D.

The following preliminary result can be proved as in Janson and Tysk (2006) or Pagliarani and Pascucci (2017), using the Schauder estimates and the results on Green functions proved in Di Francesco and Polidoro (2006).

Proposition 3.52. *Let Assumptions 3.44, 1.3, 3.50 and 3.51 be in force. Then,* $u \in C([0,T] \times D) \cap C_{B,\mathrm{loc}}^{N+2,1}$ *and satisfies* (3.6).

As was mentioned in the introduction, the idea behind our approximation of $u = u(t,x)$ in (3.1) is to expand the generator of X by approximating the coefficients a_{ij} and a_j in (3.3) by means of their *intrinsic* Taylor polynomials. Thus we fix $\bar{z} = (\bar{t}, \bar{x}) \in \mathbb{R} \times \mathbb{R}^d$ and consider the sequence $\left(\mathcal{K}_n^{(\bar{z})}\right)_{0 \leq n \leq N}$ in (3.7). We recall that, by Assumptions 3.44 and 1.3, $\mathcal{K}_0^{(\bar{z})}$ in (3.8) has a fundamental solution $\Gamma_0^{(\bar{z})}$ that is the d-dimensional Gaussian density

$$\Gamma_0^{(\bar{z})}(t,x;T,y) = \frac{1}{\sqrt{(2\pi)^d |\mathbf{C}_{\bar{z}}(T-t)|}} \times \exp\left(-\frac{1}{2}\langle \mathbf{C}_{\bar{z}}^{-1}(T-t)(y - e^{(T-t)B}x)), (y - e^{(T-t)B}x)\rangle\right) \tag{3.11}$$

with covariance matrix $\mathbf{C}_{\bar{z}}(t)$ given by

$$\mathbf{C}_{\bar{z}}(t) := \int_0^t e^{sB} A(\bar{z}) e^{sB^*} ds, \tag{3.12}$$

$$A(\bar z) := \begin{pmatrix} A_0(\bar z) & 0_{p_0\times(d-p_0)} \\ 0_{(d-p_0)\times p_0} & 0_{(d-p_0)\times(d-p_0)}. \end{pmatrix}.$$

Next we formally expand the expected value u in (3.1) as

$$u \;\approx\; U_N^{(\bar z)} := \sum_{n=0}^{N} u_n^{(\bar z)}. \tag{3.13}$$

Inserting (3.7), (3.13) into (3.6) and formally collecting terms of the same order, we find that the functions $u_n^{(\bar z)}$ satisfy the following sequence of nested Cauchy problems

$$\begin{cases} \mathcal{K}_0^{(\bar z)} u_0^{(\bar z)} = 0, & \text{on } [0,T[\times\mathbb{R}^d, \\ u_0^{(\bar z)}(T,\cdot) = \varphi, & \text{on } \mathbb{R}^d, \end{cases} \tag{3.14}$$

and

$$\begin{cases} \mathcal{K}_0^{(\bar z)} u_n^{(\bar z)} = -\sum\limits_{h=1}^{n} \big(\mathcal{K}_h^{(\bar z)} - \mathcal{K}_{h-1}^{(\bar z)}\big) u_{n-h}^{(\bar z)}, & \text{on } [0,T[\times\mathbb{R}^d, \\ u_n^{(\bar z)}(T,\cdot) = 0, & \text{on } \mathbb{R}^d. \end{cases} \tag{3.15}$$

Remark 3.53. In the above construction, the approximation in (3.13) is defined in terms of a sequence of Cauchy problems that admit a unique non-rapidly increasing solution. Conversely, equations (3.6) do not have a unique solution unless additional lateral boundary conditions are posed. Nevertheless, Theorem 3.59 below states that the above expansion is asymptotically convergent in the limit of short-time, uniformly on compact subsets of D. This is in line with the so-called *principle of not feeling the boundary* (cf. Hsu (1995), Gatheral et al. (2012)). Basically, the same asymptotic result would hold for any bounded solution of equations (3.6), with error bounds depending on the L^∞-norm of the solution. Of course, knowing the boundary conditions would allow to construct an approximate sequence that is also accurate near the boundary.

3.3 The approximation closed expression

We now show that the functions $u_n^{(\bar z)}$ in (3.14)-(3.15) can be explicitly computed at any order. It is clear that the leading term $u_0^{(\bar z)}$ is given by

$$u_0^{(\bar z)}(t,x) = \int_{\mathbb{R}^d} \Gamma_0^{(\bar z)}(t,x;T,y)\varphi(y)dy, \qquad (t,x)\in[0,T[\times\mathbb{R}^d, \tag{3.16}$$

where $\Gamma_0^{(\bar z)}$ is the Gaussian density in (3.11). For $n \in \mathbb{N}$ with $n \le N$, the explicit representation for the correcting terms $u_n^{(\bar z)}$ can be derived using the following notable symmetry properties of $\Gamma_0^{(\bar z)}$.

Lemma 3.54. *For any $x, y \in \mathbb{R}^d$, $t < s$ and $\bar z = (\bar t, \bar x) \in \mathbb{R} \times \mathbb{R}^d$, we have*

$$\nabla_x \Gamma_0^{(\bar z)}(t,x;s,y) = -e^{(s-t)B^*} \nabla_y \Gamma_0^{(\bar z)}(t,x;s,y), \tag{3.17}$$

$$y\, \Gamma_0^{(\bar z)}(t,x;s,y) = \mathcal{M}^{(\bar z)}(s-t,x) \Gamma_0^{(\bar z)}(t,x;s,y), \tag{3.18}$$

where $\mathcal{M}^{(\bar z)}(t,x)$ is the operator defined as

$$\mathcal{M}^{(\bar z)}(t,x) = e^{tB}\left(x + \mathbf{M}_{\bar z}(t)\nabla_x\right), \qquad \mathbf{M}_{\bar z}(t) = e^{-tB}\mathbf{C}_{\bar z}(t)e^{-tB^*}. \tag{3.19}$$

Proof. Using the explicit expression of $\Gamma_0^{(\bar z)}$, the proof is a direct computation. □

Finally, the exact expression of $u_n^{(\bar z)}$ in (3.13): remarkably, it can be written as a finite sum of spatial derivatives acting on $u_0^{(\bar z)}$.

Theorem 3.55. *Let Assumptions 3.44, 1.3 and 3.50 be in force. Then, for any $n \in \mathbb{N}$ with $n \le N$, and for any $\bar z \in \mathbb{R} \times \mathbb{R}^d$, we have*

$$u_n^{(\bar z)}(t,x) = \mathcal{L}_n^{(\bar z)}(t,T,x) u_0^{(\bar z)}(t,x), \qquad (t,x) \in [0,T[\times \mathbb{R}^d. \tag{3.20}$$

In (3.20), $\mathcal{L}_n^{(\bar z)}(t,T,x)$ denotes the differential operator

$$\mathcal{L}_n^{(\bar z)}(t,T,x) = \sum_{h=1}^{n} \int_t^T ds_1 \int_{s_1}^T ds_2 \cdots \int_{s_{h-1}}^T ds_h \sum_{i \in I_{n,h}} \mathcal{G}_{i_1}^{(\bar z)}(t,s_1,x) \cdots \mathcal{G}_{i_h}^{(\bar z)}(t,s_h,x), \tag{3.21}$$

where

$$I_{n,h} = \{ i = (i_1, \dots, i_h) \in \mathbb{N}^h \mid i_1 + \dots + i_h = n \}, \qquad 1 \le h \le n,$$

and

$$\begin{aligned} &\mathcal{G}_n^{(\bar z)}(t,s,x) \\ =&\frac{1}{2}\sum_{i,j=1}^{p_0} \big(\mathcal{T}_n - \mathcal{T}_{n-1}\big)(a_{ij},\bar z)\big(s, \mathcal{M}^{(\bar z)}(s-t,x)\big)\big(e^{-(s-t)B^*}\nabla_x\big)_i \big(e^{-(s-t)B^*}\nabla_x\big)_j \\ &+ \sum_{i=1}^{p_0} \big(\mathcal{T}_{n-1}(a_i,\bar z) - \mathcal{T}_{n-2}(a_i,\bar z)\big)\big(s, \mathcal{M}^{(\bar z)}(s-t,x)\big)\big(e^{-(s-t)B^*}\nabla_x\big)_i, \end{aligned} \tag{3.22}$$

with $\mathcal{M}^{(\bar z)}(t,x)$ as in (3.19) and, by convention, $\mathcal{T}_{-1} f \equiv 0$.

It should be noted that the result above is essentially the same in Lorig et al. (2015) but with a more general reach. In fact, in that article only the case $B \equiv 0$ or, in other words, the uniformly parabolic case is studied, a setting which greatly simplify the symmetries in Lemma 3.54 and the expression of the operators $\mathcal{G}_n$ consequently.

The proof of Theorem 3.55 goes exactly as its counterpart in Lorig et al. (2015) one the correct symmetries are used. In fact, the other ingredients are an extensive use of the Duhamel's principle and the Chapman-Kolmogorov equation which are available also in our setting. For the sake of clarity, we report here the most important parts of the proof but we avoid to furnish complete proofs as they are very lengthy and refer to the original article for the interested reader. Since the choice of $\bar{z}$ is unimportant through this section, we drop the explicit dependence on $\bar{z}$ in the following formulas. First, we generalize formula (3.18) to polynomial functions p with time-dependent coefficients, that is $p = p(t, \cdot)$ is a polynomial for every fixed $t \in \mathbb{R}$: this will be used to deal with the operators $\mathcal{K}_n$ in (3.7) that have coefficients of this form.

Proposition 3.56. *For any $t, s, s_1 \in [0, T]$, with $t < s$, $x, y \in \mathbb{R}^d$, we have*

$$p(s_1, y)\Gamma_0(t, x; s, y) = p\left(s_1, \mathcal{M}(s - t, x)\right)\Gamma_0(t, x; s, y). \tag{3.23}$$

Proof. Let us recall that operator $\mathcal{M}(t, x)$ acts only on the variable x. First, we prove that the components $\mathcal{M}_j(t, x)$, $i = 1, \dots, d$, commute when applied to $\Gamma_0 = \Gamma_0(t, x; s, y)$ and to its derivatives (notice however that this is not true in general when they are applied to a generic function). Notice also that formula (3.17) expresses an x-derivative as a linear combination of y-derivatives with coefficients that depend only on t and s. This is obviously true also for higher orders and we express it through the differential operator $S^\beta_y(s - t)$, acting on y, defined by

$$D^\beta_x \Gamma_0(t, x; s, y) = S^\beta_y(s - t)\Gamma_0(t, x; s, y).$$

Now we have

$$\begin{aligned}
&\mathcal{M}_i(s - t, x)\mathcal{M}_j(s - t, x)D^\beta_x \Gamma_0 \\
=&\mathcal{M}_i(s - t, x)\mathcal{M}_j(s - t, x)S^\beta_y(s - t)\Gamma_0 && \text{(by the definition above)} \\
=&S^\beta_y(s - t)(\mathcal{M}_i(s - t, x)\mathcal{M}_j(s - t, x)\Gamma_0) && (S^\beta_y \text{ and } \mathcal{M}_j \text{ commute}) \\
=&S^\beta_y(s - t)(\mathcal{M}_i(s - t, x)y_j\Gamma_0) && \text{(by (3.18))}
\end{aligned}$$

$$
\begin{aligned}
&= S_y^{\beta}(s-t)(y_j \mathcal{M}_i(s-t,x)\Gamma_0) \\
&= S_y^{\beta}(s-t)(y_j y_i \Gamma_0) && \text{(again, by (3.18))} \\
&= \mathcal{M}_j(s-t,x)\mathcal{M}_i(s-t,x) D_x^{\beta}\Gamma_0. && \text{(by reversing the steps above)}
\end{aligned}
$$

Since $p(s_1,\cdot)$ is a polynomial by definition, we therefore have that the operators $p\left(s_1, \mathcal{M}(s-t,x)\right)$ are defined unambiguously when applied to $\Gamma_0(t,x;s,y)$ and to its derivatives. Moreover, clearly (3.23) is now a straightforward consequence of (3.18). □

Remark 3.57. By Proposition 3.56, the operators $\mathcal{G}_n(t,s,x)$ are defined unambiguously when applied to $\Gamma_0 = \Gamma_0(t,x;s,y)$, to its derivatives and, more generally, by the representation formula (3.16), to solutions of the Cauchy problem (3.14).

The next proposition, essentially based on the symmetries of Lemma 3.54, is the key of the proof of Theorem 3.55.

Proposition 3.58. *For any $x, y \in \mathbb{R}^d$, $t < s$ and $n \in \mathbb{N}$ with $n \leq N$, we have*

$$
\int_{\mathbb{R}^d} \Gamma_0(t,x;s,\xi)\big(\left(\mathcal{K}_n - \mathcal{K}_{n-1}\right) f\big)(s,\xi) d\xi = \mathcal{G}_n(t,s,x) \int_{\mathbb{R}^d} \Gamma_0(t,x;s,\xi) f(\xi) d\xi, \tag{3.24}
$$

for any $f \in C_0^2(\mathbb{R}^d)$.

Proof. To keep formulas at a reasonable size we suppose that the functions a_i, $i = 1, \dots, p_0$, in (3.3) are identically zero. By the definition (3.7) we have

$$
\begin{aligned}
&\int_{\mathbb{R}^d} \Gamma_0(t,x;s,\xi)\big((\mathcal{K}_n - \mathcal{K}_{n-1}) f\big)(s,\xi) d\xi \\
&= \frac{1}{2} \sum_{i,j=1}^{p_0} \int_{\mathbb{R}^d} \left(\mathcal{T}_n\left(a_{ij}, \bar{z}\right) - \mathcal{T}_{n-1}(a_{ij}, \bar{z})\right)(s,\xi) \Gamma_0(t,x;s,\xi) \partial_{\xi_i \xi_j} f(\xi) d\xi \\
&= \frac{1}{2} \sum_{i,j=1}^{p_0} \left(\mathcal{T}_n\left(a_{ij}, \bar{z}\right) - \mathcal{T}_{n-1}(a_{ij}, \bar{z})\right)\big(s, \mathcal{M}(s-t,x)\big) \int_{\mathbb{R}^d} \Gamma_0(t,x;s,\xi) \partial_{\xi_i \xi_j} f(\xi) d\xi
\end{aligned}
$$

(by (3.23))

$$
= \frac{1}{2} \sum_{i,j=1}^{p_0} \left(\mathcal{T}_n\left(a_{ij}, \bar{z}\right) - \mathcal{T}_{n-1}(a_{ij}, \bar{z})\right)\big(s, \mathcal{M}(s-t,x)\big) \int_{\mathbb{R}^d} \partial_{\xi_i \xi_j} \Gamma_0(t,x;s,\xi) f(\xi) d\xi
$$

(by parts)

$$= \mathcal{G}_n(t,s,x)\int_{\mathbb{R}^d}\Gamma_0(t,x;s,\xi)f(\xi)d\xi.$$

(by (3.17) and (3.22))

□

The proof of Theorem 3.55 consists of mostly formal and tedious computations that are totally analogous to those given for the parabolic case in Section 5 in Lorig et al. (2015). This may not be surprising since our framework contains the parabolic one as a special case. Therefore, we only give a proof for $n = 1$, which still sheds light on the origin of the operators $\mathcal{L}_n$.

By definition, u_1 is the solution of the Cauchy problem (3.15) with $n = 1$. By Duhamel's principle we have

$$u_1(t,x) = \int_t^T \int_{\mathbb{R}^d} \Gamma_0(t,x;s,\xi)\left((\mathcal{K}_1 - \mathcal{K}_0)u_0\right)(s,\xi)\mathrm{d}\xi ds$$

$$= \int_t^T \mathcal{G}_1(t,s,x)\int_{\mathbb{R}^d}\Gamma_0(t,x;s,\xi)\,u_0(s,\xi)d\xi ds \qquad \text{(by (3.24) with } n=1)$$

$$= \int_t^T \mathcal{G}_1(t,s,x)\int_{\mathbb{R}^d}\Gamma_0(t,x;s,\xi)\int_{\mathbb{R}^d}\Gamma_0(s,\xi;T,y)\varphi(y)dyd\xi ds \qquad \text{(by (3.16))}$$

$$= \int_t^T \mathcal{G}_1(t,s,x)\int_{\mathbb{R}^d}\varphi(y)\int_{\mathbb{R}^d}\Gamma_0(t,x;s,\xi)\,\Gamma_0(s,\xi;T,y)d\xi dy ds \qquad \text{(Fubini's theorem)}$$

$$= \int_t^T \mathcal{G}_1(t,s,x)ds\,u_0(t,x) \qquad \text{(Chapman-Kolmogorov and (3.16))}$$

$$= \mathcal{L}_1(t,T,x)u_0(t,x). \qquad \text{(by (3.21))}$$

□

3.4 Error estimates

The choice of the basis point $\bar{z}$ is somewhat arbitrary, but only some particular choices allow for performing a rigorous error analysis. For instance, here below we consider the case $\bar{z} = z = (t, x)$. However, although we omit to write separate proofs, the same results hold by setting $\bar{z} = (T, x)$. In the following statement, we put

$$U_N(z) := U_N^{(z)}(z), \qquad z \in [0, T] \times D, \tag{3.25}$$

with $U_N^{(z)}$ defined by (3.13)-(3.14)-(3.15).

Theorem 3.59. *Let Assumptions 3.44, 1.3, 3.50 and 3.51 be in force. Then for any compact subset K of D, we have*

$$|u(t, x) - U_N(t, x)| \leq C(T - t)^{\frac{N+k+1}{2}}, \qquad (t, x) \in [0, T] \times K, \tag{3.26}$$

where C is a positive constant that depends only on $M, \mu, B, T, N, K, \|\psi\|_{C_B^k(\mathbb{R}^d)}$ and $\|u\|_{L^\infty([0,T]\times D)}$.

Theorem 3.59 will be proved in Section 3.4.2.

Remark 3.60. As shown in Example 3.49, for a fixed-strike Asian option we have $\varphi \in C_B^3(\mathbb{R}^2)$ and therefore we get $(T - t)^{\frac{N+4}{2}}$ in the error estimate (3.26). This is coherent with the previous results proved in Gobet and Miri (2014) in the scalar case for $N \leq 2$, and sheds some light on why the order of convergence of Asian call options is improved w.r.t. their European counterparts, for which the error is of order $(T - t)^{\frac{N+2}{2}}$. When placed within our framework, this improvement of convergence can be seen as part of a wider phenomenon related to the intrinsic geometry of Kolmogorov operators.

Remark 3.61. If the coefficients a_{ij}, a_i only depend on the first p_0 variables, then it is possible to prove the error bounds in (3.26) to be also asymptotic in the limit of small M. Precisely,

$$|u(t, x) - U_N(t, x)| \leq C\big(M(T - t)\big)^{\frac{N+k+1}{2}}, \qquad (t, x) \in [0, T] \times K,$$

with C independent of M as $M \to 0^+$. This is the case, for instance, of classical volatility models for Asian options where the volatility coefficient depends at most on the underlying asset S_t (local volatility) and on some exogenous factors (stochastic volatility), but not on the average process A_t.

In the global case, when $D = \mathbb{R}^d$, we have some stronger results. Aside from the error bounds in (3.26) becoming global in space, we are also able to obtain analogous asymptotic error bounds for the transition density of X. We start by observing that when $D = \mathbb{R}^d$ our assumptions imply that X has a transition density Γ that coincides with the fundamental solution of $\mathcal{K}$ as in (3.3)-(3.5) (see, for instance, Polidoro (1994)). We denote by Γ_N the N-th order approximation of Γ defined as

$$\Gamma_N(t,x;T,y) = \sum_{n=0}^{N} u_n(t,x;T,y) \qquad 0 \le t < T,\ x,y \in \mathbb{R}^d,$$

where $u_0(t,x;T,y) = \Gamma_0^{(t,x)}(t,x;T,y)$ in (3.11), and the correcting terms $u_n(t,x;T,y)$ are defined recursively by (3.15) with $\bar{z} = (t,x)$. We have the following

Theorem 3.62. *Let Assumptions 3.44, 1.3, 3.50 and 3.51 be in force with $D = \mathbb{R}^d$. Then, we have*

$$|u(t,x) - U_N(t,x)| \le C(T-t)^{\frac{N+k+1}{2}}, \qquad (t,x) \in [0,T] \times \mathbb{R}^d, \quad (3.27)$$

where C depends only on M, μ, B, T, N and $\|\varphi\|_{C_B^k(\mathbb{R}^d)}$. Moreover, for any $c > 1$, we have

$$|\Gamma(t,x;T,y) - \Gamma_N(t,x;T,y)| \le C(T-t)^{\frac{N+1}{2}} \Gamma^{cM}(t,x;T,y), \ \ 0 \le t < T, \quad (3.28)$$

where, for any $\Lambda > 0$, Γ^Λ denotes the fundamental solution of the constant-coefficient Kolmogorov operator $\mathcal{K}^\Lambda$ as defined in (3.10)*, and C is a positive constant that depends only on M, μ, B, T, N and c.*

3.4.1 Proof of the global estimates

The proof of Theorem 3.62 is based on the following two propositions. The first one provides some Gaussian estimates for the fundamental solution $\Gamma = \Gamma(t,x;T,y)$ of the operator $\mathcal{K}$ in (3.5)-(3.3): for the proof see Polidoro (1994) and Di Francesco and Pascucci (2005). Throughout this section we suppose the assumptions of Theorem 3.62 to be in force.

Proposition 3.63. *For any $k \in \mathbb{R}_{\ge 0}$, $c > 1$ and $\gamma \in \mathbb{N}_0^d$, with $|\gamma|_B \le N+2$, we have*

$$\left[y - e^{(T-t)B}x\right]_B^k \left|D_x^\gamma \Gamma(t,x;T,y)\right| \le C(T-t)^{\frac{k-|\gamma|_B}{2}} \Gamma^{cM}(t,x;T,y),$$

for any $0 \le t < T,$, $x,y \in \mathbb{R}^d$, and where Γ^{cM} is the fundamental solution of the operator in (3.10) *and C is a positive constant, only dependent on M, μ, B, T, N, k and c.*

The following result is proved in Appendix B.

Proposition 3.64. *Let $\varphi \in C_B^k(\mathbb{R}^d)$ with $k \in [0, 2r+1]$ and $n \in \mathbb{N}$ with $n \leq N$. Then we have*

$$\left|D_x^\beta u_n^{(\bar{z})}(t,x)\right| \leq C\,(T-t)^{\frac{k-|\beta|_B}{2}} \left((T-t)^{\frac{n}{2}} + \left[x - e^{(t-\bar{t})B}\bar{x}\right]_B^n\right),$$

for $0 \leq t < T$, $x \in \mathbb{R}^d$ and where C is a constant that depends only on $M, \mu, B, T, N, |\beta|_B$ and $\|\varphi\|_{C_B^k(\mathbb{R}^d)}$.

Proof of Theorem 3.62. To keep formulas at a reasonable size we suppose that the functions a_i, $i = 1, \dots, p_0$, in (3.3) are identically zero. We first remark that a straightforward computation (see Lemma 6.3 in Lorig et al. (2015)) shows that

$$u(t,x) - U_N(t,x) = \sum_{n=0}^{N} E_n^{(\bar{z})}(t,x)\Big|_{\bar{z}=(t,x)}. \tag{3.29}$$

where

$$\begin{aligned} E_n^{(\bar{z})}(t,x) &:= \int_t^T \int_{\mathbb{R}^d} \Gamma(t,x;s,\xi)\left(\mathcal{K} - \mathcal{K}_n^{(\bar{z})}\right) u_{N-n}^{(\bar{z})}(s,\xi)d\xi ds \\ &= \frac{1}{2}\sum_{i,j=1}^{p_0} \int_t^T \int_{\mathbb{R}^d} \Gamma(t,x;s,\xi) \\ &\qquad \times \Big(a_{ij}(s,\xi) - \mathcal{T}_n\left(a_{ij},\bar{z}\right)(s,\xi)\Big)\partial_{\xi_i\xi_j} u_{N-n}^{(\bar{z})}(s,\xi)\, d\xi ds. \end{aligned} \tag{3.30}$$

Now, if $k > 0$, by Theorem 2.20 and Proposition 3.64 we have

$$\begin{aligned} \left|E_n^{(t,x)}(t,x)\right| &\leq C \int_t^T \int_{\mathbb{R}^d} \Gamma(t,x;s,\xi)\left\|(t,x)^{-1} \circ (s,\xi)\right\|_B^{n+1} \\ &\qquad \times (T-s)^{\frac{k-2}{2}}\left((T-s)^{\frac{N-n}{2}} + \left[\xi - e^{(s-t)B}x\right]_B^{N-n}\right)d\xi ds \end{aligned}$$

(by Proposition 3.63)

$$\begin{aligned} &\leq C\int_t^T (s-t)^{\frac{n+1}{2}}(T-s)^{\frac{k-2}{2}}\left((T-s)^{\frac{N-n}{2}} + (s-t)^{\frac{N-n}{2}}\right)ds \\ &\leq C\,(T-t)^{\frac{N+k+1}{2}} \end{aligned}$$

where we have used the identity

$$\int_t^T (T-s)^n (s-t)^k\, ds = \frac{\Gamma_E(k+1)\Gamma_E(n+1)}{\Gamma_E(k+n+2)}(T-t)^{k+n+1}, \qquad n, k > -1,$$

with Γ_E denoting the Euler Gamma function. The case $k = 0$ can be handled similarly performing first an integration by parts in (3.30).

Finally, estimate (3.28) can be proved by a straightforward modification of the proof of (3.27), using also the Chapman-Kolmogorov equation. We omit the details for brevity. □

Remark 3.65. Under the assumptions of Theorem 3.62, we have also error bounds for the approximation of the derivatives of u; precisely, we have

$$\left| D_x^\alpha u(t,x) - D_x^\alpha U_N^{(\bar z)}(t,x)|_{\bar z=(t,x)} \right| \leq C(T-t)^{\frac{N+k+1-|\alpha|_B}{2}}, \qquad |\alpha|_B \leq N. \quad (3.31)$$

The proof of this formula is analogous to the proof of Theorem 3.62, once D_x^α is applied to the representation formulas (3.29) and (3.30). When $u(t,x)$ represents the price of an arithmetic Asian option, formula (3.31) provides error bounds on the approximate sensitivities or, as they are usually called in finance, the Greeks. For instance, in the case of a fixed-strike Asian option (see Example 3.49), we have $k = 3$ and thus

$$\begin{aligned} \left| \mathrm{Delta} - \partial_{x_1} U_N^{(\bar z)}|_{\bar z=(t,x_1,x_2)} \right| &\leq C(T-t)^{\frac{N+3}{2}}, \\ \left| \mathrm{Gamma} - \partial_{x_1,x_1} U_N^{(\bar z)}|_{\bar z=(t,x_1,x_2)} \right| &\leq C(T-t)^{\frac{N+2}{2}}, \end{aligned}$$

where $\mathrm{Delta} := \partial_{x_1} u$ and $\mathrm{Gamma} := \partial_{x_1,x_1} u$.

3.4.2 Proof of the local estimates

Throughout this section we suppose the assumptions of Theorem 3.59 to be in force. The proof of Theorem 3.59 is based on some estimates on short cylinders initially introduced in Safonov (1998) for uniformly parabolic operators and later generalized to Kolmogorov operators in Cinti and Polidoro (2009).

First, we introduce the "cylinder" of radius R and height h centered in $\zeta = (s,\xi) \in \mathbb{R} \times \mathbb{R}^d$ and its lateral and parabolic boundaries, respectively:

$$\begin{aligned} H_\zeta(h,R) &:= \{(t,x) \in \mathbb{R} \times \mathbb{R}^d \mid s-h < t < s,\ [x - e^{(t-s)B}\xi]_B < R\}, \\ \Sigma_\zeta(h,R) &:= \{(t,x) \in \mathbb{R} \times \mathbb{R}^d \mid s-h < t < s,\ [x - e^{(t-s)B}\xi]_B = R\}, \\ \partial_P H_\zeta(h,R) &:= \Sigma_\zeta(h,R) \cup \{(s,x) \in \mathbb{R} \times \mathbb{R}^d \mid [x-\xi]_B < R\}. \end{aligned}$$

We explicitly observe that these cylinders are invariant with respect to the left translations in $\mathcal{G}_B$, meaning that $z \circ H_\zeta(h,R) = H_{z\circ\zeta}(h,R)$ for any $z, \zeta \in \mathbb{R}\times\mathbb{R}^d$.

We also recall the following result from Lemma 1.8:

$$\|z \circ \zeta\|_B \leq c_B \left(\|z\|_B + \|\zeta\|_B\right), \qquad z, \zeta \in \mathbb{R} \times \mathbb{R}^d, \tag{3.32}$$

where $c_B \geq 1$ is a constant that depends only on the matrix B. In particular, taking $z = (0, x)$ and $\zeta = (t, 0)$, (3.32) implies that

$$[e^{tB}x]_B \leq \|(t, e^{tB}x)\|_B = \|z \circ \zeta\|_B \leq c_B\big(|t|^{\frac{1}{2}} + [x]_B\big), \qquad t \in \mathbb{R},\ x \in \mathbb{R}^d. \tag{3.33}$$

Lemma 3.66. *There exist $C > 0$, $\varepsilon \in]0, 1[$, only dependent on M, μ, B, and a nonnegative function $v \in C([0, T] \times \mathbb{R}^d) \cap C^{2,1}_{B,\mathrm{loc}}$ such that, for every $R > 0$ we have*

$$\mathcal{K}v(t, x) = 0 \qquad (t, x) \in H_{(T,0)}(\varepsilon R^2, R), \tag{3.34}$$

$$v(t, x) \geq 1, \qquad (t, x) \in \Sigma_{(T,0)}(\varepsilon R^2, R), \tag{3.35}$$

$$v(t, x) \leq C \exp\left(-\frac{R^2}{C(T-t)}\right) \qquad (t, x) \in H_{(T,0)}\big(\varepsilon R^2, \frac{R}{8c_B^2}\big), \tag{3.36}$$

where c_B is the constant in (3.32).

Proof. Let Γ denote the fundamental solution of $\mathcal{K}$ in (3.5): Γ can be thought as the transition density of a dummy process $\widetilde{X}$ whose infinitesimal generator is $\mathcal{A}$ and can be used to approximate the original process X locally on D. The proof of the lemma is based on a Gaussian upper bound for Γ. More precisely, since $\mathcal{K}$ is a global Kolmogorov operator, by Proposition 3.63 we have: there exists a positive constant c^+, only depending on M, μ and B, such that

$$\Gamma(t, x; s, \xi) \leq c^+ \Gamma^\Lambda(t, x; s, \xi), \qquad 0 \leq t < s \leq T,\ x, \xi \in \mathbb{R}^d, \tag{3.37}$$

where Γ^Λ is the fundamental solution of the constant coefficients Kolmogorov operator in (3.10) and Λ is strictly greater than M, say $\Lambda = 2M$.

Next, we set

$$v(t, x) = 2 \int_{\mathbb{R}^d} \Gamma(t, x; T, y) \chi_R(y) dy, \qquad t < T,\ x \in \mathbb{R}^d,$$

where $\chi_R \in C^\infty(\mathbb{R}^d, [0, 1])$ is a smooth function such that $\chi_R(y) = 0$ if $[y]_B < \frac{R}{2}$ and $\chi_R(y) = 1$ if $[y]_B > \frac{3}{4}R$. By definition, it is clear that v satisfies (3.34). Moreover, we have

$$\lim_{t \to T^-} v(t, x) = 2\chi_R(x) = 2, \tag{3.38}$$

uniformly w.r.t. $x \in \mathbb{R}^d$ *such that* $[x]_B = R$*:* this follows by noting that

$$\begin{aligned}
|v(t,x) - 2\chi_R(x)| &\le 2 \int\limits_{\mathbb{R}^d} \Gamma(t,x;T,y)\, |\chi_R(y) - \chi_R(x)|\, dy \\
&\le 2c^+ \int\limits_{\mathbb{R}^d} \Gamma^\Lambda(t,x;T,y)\, |\chi_R(y) - \chi_R(x)|\, dy. \qquad \text{(by (3.37))}
\end{aligned}$$

Now, by (3.38) there exists $\varepsilon > 0$, which we can safely assume to be less than $\frac{1}{16c_B^4}$ and $\frac{1}{64c_B^2}$, such that (3.35) holds.

The proof of (3.36) depends on the reverse triangle inequality for the norm $[\cdot]_B$:

$$[y - e^{tB}x]_B \ge \frac{1}{c_B}[y]_B - c_B\big(|t|^{\frac12} + [x]_B\big), \qquad t \in \mathbb{R},\ x, y \in \mathbb{R}^d,$$

whose proof is an easy consequence of (3.33). In particular, if $[y]_B \ge \frac{R}{2}$ and $(t,x) \in H_{(T,0)}\big(\varepsilon R^2, \frac{R}{8c_B^2}\big)$, then in light of the first bound for ε we get

$$[y - e^{(T-t)B}x]_B \ge \frac{R}{8c_B}. \tag{3.39}$$

Hence, for such (t,x) we get

$$\begin{aligned}
v(t,x) &\le 2c^+ \int\limits_{\mathbb{R}^d} \Gamma^\Lambda(t,x;T,y)\chi_R(y)dy \le 2c^+ \int\limits_{[y]_B \ge \frac{R}{2}} \Gamma^\Lambda(t,x;T,y)dy \\
&= \frac{2c^+(2\pi)^{-\frac{d}{2}}}{\sqrt{|\mathbf{C}(T-t)|}} \times \\
&\quad \int\limits_{[y]_B \ge \frac{R}{2}} \exp\left(-\frac12 \langle \mathbf{C}^{-1}(T-t)(y - e^{(T-t)B}x), (y - e^{(T-t)B}x)\rangle\right) dy
\end{aligned}$$

(by (3.39) and denoting by $\mathbf{C}$ the matrix in (3.12) with $A_0 = \Lambda I_{p_0}$ and I_{p_0} being the $(p_0 \times p_0)$ identity matrix)

$$\begin{aligned}
&\le \frac{2c^+(2\pi)^{-\frac{d}{2}}}{\sqrt{|\mathbf{C}(T-t)|}} \times \\
&\quad \int\limits_{[y - e^{(T-t)B}x]_B \ge \frac{R}{8c_B}} \exp\left(-\frac12 \langle \mathbf{C}^{-1}(T-t)(y - e^{(T-t)B}x), (y - e^{(T-t)B}x)\rangle\right) dy
\end{aligned}$$

(by the change of variables $\eta = D_0(\frac{1}{\sqrt{T-t}})(y - e^{(T-t)B}x)$ and the homogeneity relation (B.1))

$$= \frac{2c^+(2\pi)^{-\frac{d}{2}}}{\sqrt{|\mathbf{C}(1)|}} \int\limits_{[\eta]_B \geq \frac{R}{8c_B\sqrt{T-t}}} \exp\left(-\frac{1}{2}\langle \mathbf{C}^{-1}(1)\eta, \eta\rangle\right) d\eta. \tag{3.40}$$

Since we are assuming $T - t \leq \varepsilon R^2$, thanks to the second bound on ε we have $[\eta]_B \geq \frac{R}{8c_B\sqrt{T-t}} \geq 1$ and thus, there exists $C_0 > 0$ only dependent on μ, M, B, such that

$$\begin{aligned}
\langle \mathbf{C}^{-1}(1)\eta, \eta\rangle \geq C_0|\eta|^2 &= C_0 \sum_{j=1}^{d} \frac{|\eta_j|^2}{[\eta]_B^{2\sigma_j}} [\eta]_B^{2\sigma_j} \\
&= C_0 \sum_{j=1}^{d} \left(\frac{|\eta_j|^{1/\sigma_j}}{[\eta]_B}\right)^{2\sigma_j} [\eta]_B^{2\sigma_j} \\
&\geq C_0[\eta]_B^2 \sum_{j=1}^{d} \left(\frac{|\eta_j|^{1/\sigma_j}}{[\eta]_B}\right)^{2(2r+1)} \\
&\geq \frac{C_0}{d^{4r+1}}[\eta]_B^2 \left(\sum_{j=1}^{d} \frac{|\eta_j|^{1/\sigma_j}}{[\eta]_B}\right)^{2(2r+1)} = \frac{C_0}{d^{4r+1}}[\eta]_B^2.
\end{aligned}$$

Setting $C_1 := \frac{C_0}{d^{4r+1}}$ we get

$$\begin{aligned}
\int\limits_{[\eta]_B \geq \frac{R}{8c_B\sqrt{T-t}}} \exp\Big(-\frac{1}{2}\langle \mathbf{C}^{-1}(1)\eta, \eta\rangle\Big) d\eta &\leq \int\limits_{[\eta]_B \geq \frac{R}{8c_B\sqrt{T-t}}} \exp\Big(-\frac{1}{2}C_1[\eta]_B^2\Big) d\eta \\
&\leq \max_{[y]_B \geq \frac{R}{8c_B\sqrt{T-t}}} \exp\Big(-\frac{1}{4}C_1[y]_B^2\Big) \int\limits_{[\eta]_B \geq \frac{R}{8c_B\sqrt{T-t}}} \exp\Big(-\frac{1}{4}C_1[\eta]_B^2\Big) d\eta \\
&\leq \exp\Big(-\frac{C_1R^2}{2^8c_B^2(T-t)}\Big) \int_{\mathbb{R}^d} \exp\Big(-\frac{1}{4}C_1[\eta]_B^2\Big) d\eta,
\end{aligned}$$

which, combined with (3.40), proves (3.36). $\square$

Proof of Theorem 3.59. Since the statement is a short-time estimate on a compact subset, it is enough to prove (3.26) for $(t, x) \in H_{(T,\xi)}(\varepsilon R^2, R) \subseteq \,]0, T[\times D$ for suitably small $\varepsilon, R > 0$. Secondly, we can suppose $\xi = 0$. In fact, if u is a solution to $\mathcal{K}u = 0$ in $H_{(T,\xi)}(\varepsilon R^2, R)$ then $w(t, x) = u(t, x - e^{-TB}\xi)$ solves on $H_{(T,0)}(\varepsilon R^2, R)$ the operator obtained through $\mathcal{K}$ by translating its coefficients.

Let us denote by u^ψ the unique solution (with polynomial growth) to the Cauchy problem

$$\begin{cases} \mathcal{K}f = 0, & \text{on } [0,T[\times\mathbb{R}^d, \\ f(T,\cdot) = \psi, & \text{on } \mathbb{R}^d, \end{cases}$$

with ψ as in Assumption 3.51, and by U_N^ψ its N-th order approximation as defined in Section 3.2. By triangular inequality we have

$$|u - U_N| \le |u - u^\psi| + |u^\psi - U_N^\psi| + |U_N^\psi - U_N|. \tag{3.41}$$

We now aim at estimating each of the terms in the sum above.

We start with $|u - u^\psi|$. Let v be the function appearing in Lemma 3.66. By Proposition 3.52 and (3.34), $u - u^\psi$ and v solve $\mathcal{K}w = 0$ in $H_{(T,0)}(\varepsilon R^2, R)$ and are continuous on $\overline{H_{(T,0)}(\varepsilon R^2, R)}$. Moreover, $(u - u^\psi)(T,x) = 0$ if $[x]_B < R$, and thus, by setting

$$C_1 := \max_{\Sigma_{\varepsilon R^2, R}(T,0)} |u - u^\psi|,$$

we get $|u - u^\psi| \le C_1 v$ on $\partial_P H_{(T,0)}(\varepsilon R^2, R)$. Therefore, by the Feynman-Kac theorem we have

$$\big|\big(u - u^\psi\big)(t,x)\big| = \big|\mathbb{E}_{t,x}\big[(u - u^\psi)\big(\tau, X_\tau\big)\big]\big| \le C_1 \mathbb{E}_{t,x}\big[v\big(\tau, X_\tau\big)\big] = C_1 v(t,x),$$

where τ denotes the exit time from $H_{(T,0)}(\varepsilon R^2, R)$ of the process (s, X_s) starting from $(t,x) \in H_{(T,0)}(\varepsilon R^2, R)$. By estimate (3.36) of Lemma 3.66 we obtain

$$\big|(u - u^\psi)(t,x)\big| \le C_1 C_2 \exp\left(-\frac{R^2}{C_2(T-t)}\right), \qquad (t,x) \in H_{\varepsilon R^2, \frac{R}{8c_B^2}}(T,0), \tag{3.42}$$

with $C_2 > 0$ depending only on M, μ, B.

We continue by estimating $|u^\psi - U_N^\psi|$. By Theorem 3.62 there exists $C_3 > 0$, only dependent on M, μ, B, T, N and $\|\psi\|_{C_B^k(\mathbb{R}^d)}$, such that

$$\big|u^\psi(t,x) - U_N^\psi(t,x)\big| \le C_3 (T-t)^{\frac{N+k+1}{2}}, \qquad (t,x) \in [0,T] \times \mathbb{R}^d. \tag{3.43}$$

We conclude by estimating $|U_N^\psi - U_N|$. First observe that, by (3.16), for any multi-index $\alpha \in \mathbb{N}_0^d$ we have

$$\begin{aligned} D_x^\alpha\big(u_0^{(\bar z)} - u_0^{(\bar z),\psi}\big)(t,x) &= D_x^\alpha \int_{\mathbb{R}^d} \Gamma_0^{(\bar z)}(t,x;T,y)\big(\varphi(y) - \psi(y)\big)dy \\ &= \int_{\mathbb{R}^d} D_x^\alpha \Gamma_0^{(\bar z)}(t,x;T,y)\big(\varphi(y) - \psi(y)\big)dy, \end{aligned}$$

with $\Gamma_0^{(\bar z)}$ as in (3.11). Now, $\Gamma_0^{(\bar z)}$ is the fundamental solution of the constant-coefficients Kolmogorov operator $\mathcal{K}_0^{(\bar z)}$ in (3.8), for which Assumptions 3.44, 1.3 and 3.50 are trivially satisfied. Therefore, the bounds in Lemma 3.63 also apply to $\Gamma_0^{(\bar z)}$ and yield

$$\left|D_x^\alpha\big(u_0^{(\bar z),\psi}-u_0^{(\bar z)}\big)(t,x)\right| \le C_4(T-t)^{-\frac{|\alpha|_B}{2}}w(t,x), \qquad \bar z\in\mathbb{R}\times\mathbb{R}^d,\ (t,x)\in[0,T[\times\mathbb{R}^d, \tag{3.44}$$

with

$$w(t,x) := \int_{\mathbb{R}^d}\Gamma^{2M}(t,x;T,y)\big|\big(\varphi(y)-\psi(y)\big)\big|dy,$$

where Γ^{2M} is the fundamental solution of the Kolmogorov operator $\mathcal{K}^{2M}$ as in (3.10), and $C_4>0$ only depends on $M,\mu,B,T,|\alpha|_B$. Now note that, by (3.25) and (3.20), we have

$$\big(U_N^\psi-U_N\big)(t,x) = \big(u_0^{(\bar z),\psi}-u_0^{(\bar z)}\big)(t,x) + \sum_{n=1}^N \mathcal{L}_n^{(\bar z)}\big(u_0^{(\bar z),\psi}-u_0^{(\bar z)}\big)(t,x)\Big|_{\bar z=(t,x)}.$$

Thus by Lemma B.7 with (3.44) we get

$$\left|\big(U_N^\psi-U_N\big)(t,x)\right| \le C_5|w(t,x)|, \qquad (t,x)\in[0,T[\times\mathbb{R}^d,$$

where $C_5>0$ only depends on M,μ,B,T and N. By repeating step by step the same proof of (3.42) it is straightforward to obtain an estimate for $|w(t,x)|$ analogous to (3.42), which finally yields

$$\left|\big(U_N^\psi-U_N\big)(t,x)\right| \le C_5C_6C_7\exp\left(-\frac{R^2}{C_7(T-t)}\right), \qquad (t,x)\in H_{(T,0)}\left(\varepsilon R^2,\frac{R}{8c_B^2}\right), \tag{3.45}$$

with $C_7>0$ depending only on M,μ,B,T,N, and

$$C_6 := \max_{\Sigma_{\varepsilon R^2,R}(T,0)}|w|.$$

Plugging (3.42)-(3.43)-(3.45) into (3.41) yields (3.26) for $(t,x)\in H_{(T,0)}\left(\varepsilon R^2,\frac{R}{8c_B^2}\right)$ and concludes the proof.

□

Chapter 4

A new mine valuation model

We investigate the valuation problem for a mine, starting from the work Brennan and Schwartz (1985) in which a three dimensional boundary problem is proposed. The problem is to express the price of a mine as a function of the relevant quantities, which, in the simplest case, are supposed to be the mineral price S, the remaining quantity to be extracted Q and, of course, time.

We suppose the total mineable resource quantity Q_{M} to be known. This is a realistic assumption as it is estimated during the preliminary study of the terrain. Obviously, new findings or improved technologies can change such amount but we think that incorporate uncertainty in this direction would not significantly improve the model.

On the other hand, an important parameter is the ore grade ratio G, the quantity of pure mineral (usually in grams) per excavated tonne. The importance of this parameter is clear: as the cost of running the mine essentially depends on the quantity of excavated material, the profitability of the mine can roughly be expressed by a function of the product GS.

As a matter of fact, G is rarely constant through the life of a mine but changes as different lodes are excavated. A more realistic approach to the mine valuation would then consider G to be stochastic. However, implementing this feature would increase the complexity of the model that would then suffer greatly from the curse of dimensionality.In the sequel we present both the models in which G is constant or stochastic and refer to them by the number of "spacial" components, as the two and three dimensional model respectively.

We then proceed to introduce our modified models and compare them with the existent ones, later we prove that our value function satisfies a Dirichlet boundary value problem for a degenerate Kolmogorov operator. The main part of the chapter is dedicated to the proof of existence and uniqueness of the price function. For the uniqueness we employ a probabilistic approach, using a Feynman-Kac type theorem (see, for example, Pascucci (2011), Theorem 9.44) while the existence result is achieved adapting a procedure used in Barucci et al. (2001) for Asian option. It should be noted that we were not able to provide uniqueness in the three dimensional case due to our inability to obtain an explicit bound, see Section 4.3.

Along the way we will point out some peculiar properties of the problem and the operator, both from the geometric and analytic point of view. Such features, while taking us out of the general theory framework, both make the problem worth to study and, in our humble opinion, beautiful.

At the end, we will uniformly bound the value function from below and exactly identify the region in which it is positive.

4.1 Literature review

We begin by describing the model proposed in Brennan and Schwartz (1985) for the price of a mine and its generalization to stochastic ore grade made in Evatt et al. (2010). As we will see, the modification proposed to us, while reasonable by the financial point of view, make the problem analytically tractable.

The mineral price S is supposed to be stochastic and to follow a classic geometric Brownian motion dynamic

$$dS = \mu S\, dt + \sigma_1 S\, dW,$$

with μ, σ_1 two positive constants. Defining $\bar{Q} := Q_{\mathrm{M}} - Q$, the already extracted quantity, the extraction rate $d\bar{Q}$ is supposed to be deterministic with dynamic

$$d\bar{Q} = w(t)\, dt = -dQ \tag{4.1}$$

where w is a positive and deterministic function of time. In particular, in this model the extraction rate is decided a priori and it is not influenced by external factors as price or ore grading. Thus, the mine is supposed to work until either it is exhausted ($Q = 0$) or the option to use it expires ($t = T$). However, in practical implementation, w is supposed to be constant.

The economic value of the mineral is the difference between the extraction cost per unit ε_M and the cash generated by selling the mineral. Usually, an extra stage of processing is required after the extraction to be able to sell the mineral. In Evatt et al. (2010) the authors model the case in which this secondary stage is done only if economically convenient i.e. if $wSG > \varepsilon_P$ with ε_P being the processing cost per unit. Note that wSG is the instantaneous value of the extracted mineral.

By financial arguments, the value function $V(t, s, q)$ has to solve the equation

$$L_1 V := \frac{1}{2} s^2 \sigma_1^2 \frac{\partial^2 V}{\partial s^2} + rs \frac{\partial V}{\partial s} - w \frac{\partial V}{\partial q} + \frac{\partial V}{\partial t} = rV - f(t, s), \tag{4.2}$$

in the unbounded domain $\Omega_1 := \mathbb{R}_{>0} \times (0, Q_M) \times (0, T)$. The function f represents the instantaneous gains:

$$f(s) := (wGs - \varepsilon_P)^+ - \varepsilon_M,$$

while r the constant (non-negative) interest rate.

Regarding the boundary conditions for the above equation, it is natural to impose the homogeneous condition $V = 0$ on the part of $\partial\Omega_1$ described either by $t = T$ or $Q = 0$. The meaning is clear: if the contract to use the mine expires or there is no more mineral to be extracted, the mine has no value. However, the conditions on the other parts of $\partial\Omega_1$ are more puzzling: in the paper above the authors do not provide financial motivations but essentially restrict equation (4.2) to the boundary. As we shall see, in our (different) model, these additional conditions are not required: we will prove existence and uniqueness for the two dimensional model under just the homogeneous conditions.

No attempt were made in Brennan and Schwartz (1985) to prove the existence of solutions. In fact, this problem exceed the standard framework of PDE analysis as the operator $\mathcal{L}_1$ is strongly degenerate: the second order q derivative is missing and is totally degenerate at the boundary points described by $s = 0$. Moreover, the analysis of L_1 also lies outside the realm of Hörmander theory. Introducing the vector fields

$$X_1 = \frac{1}{\sqrt{2}} \sigma_1 s \partial_s, \qquad Y = -w\partial_q + \partial_t, \tag{4.3}$$

we have $[X_1, Y] = 0$ and thus the Lie algebra generated by them has only dimension 2 at every point of $\Omega_1 \subset \mathbb{R}^3$. Therefore, Hörmander's condition is not met.

Three dimensional model

In the paper Evatt et al. (2010) the model is extended to the case of a stochastic ore grade G following a mean reverting process such as

$$dG = k(\alpha - G)\, d\bar{Q} + \sigma_2 \sqrt{G}\, d\widetilde{W}^2, \tag{4.4}$$

where $\widetilde{W}^2$ is a normally distributed as $\mathcal{N}(0, \sqrt{\bar{Q}})$. This corresponds to a classic CIR model under the *deterministic* operational time given by $\bar{Q}$: this is necessary as the ore grade does not change directly as time flows but as the ore is excavated. See Karatzas and Shreve (1991), Sections 3.4 B and 5.5 A for further details on time-changed processes. Note that k and α represent the long-term mean and the velocity of the reversion; together with σ_2 they are positive parameters to be calibrated. The dynamic above take into account the fact that ore grade may vary during the mine lifetime but should not divert too much from the mean.

On the other hand, the extension introduces a new state variable transforming the PDE problem in a 4-dimensional one. Its numerical solution suffers of the so called curse of dimensionality. Also, the operator has the expression

$$L_2 = L_1 + \frac{1}{2} w g \sigma_2^2 \frac{\partial^2}{\partial g^2} + w k \alpha \frac{\partial}{\partial g}$$

and thus still suffers of the same difficulties that characterize the analysis of $\mathcal{L}_1$. Note that if we suppose G to be constant L_2 reduces to L_1. Regarding boundary conditions, the discussion for the two dimensional case can be repeated here almost word by word.

4.2 The new models

Our models build on the aforementioned ones simply changing the dynamic in (4.1) to

$$d\bar{Q} = wS\, dt = -dQ. \tag{4.5}$$

As simple as the change can seem, it has remarkable consequences: first of all, the extraction rate is directly proportional to the mineral price meaning that, as the price rises the owner of the mine is willing to extract more while if it drops the excavations slow down accordingly. To the best of our knowledge this is the first PDE model to take into account price as a factor for extraction rate.

Secondly, the new operators fall in the Hörmander theory framework. Consider for example the two dimensional case: the analogous the vector fields in (4.3) reads as

$$X = \frac{1}{\sqrt{2}}\sigma s_1 \partial_s, \qquad Y = -ws\partial_Q + \partial_t,$$

and now it holds $[X, Y] = \frac{1}{\sqrt{2}} q\sigma_1 s\partial_q$ and therefore the Lie algebra generated by X, Y has full dimension at every point of Ω_1. In fact, the corresponding operator is a (degenerate) non homogeneous Kolmogorov Operator with matrix

$$B = \begin{pmatrix} r & 0 \\ -w & 0 \end{pmatrix}.$$

The degeneracy is due to the coefficient s^2 in the second order part and cannot be avoided. On the other hand, we will find a suitable change of variable able to transform the operator in a homogeneous one. However, we were not able to find such a change for the three dimensional case. We will discuss this topic later on Section 4.4

Finally, in Section 4.3 we will be able to describe exactly in which part of the boundary Dirichlet type conditions can be imposed.

Remark 4.67. An important problem is to compute the probability that the mine is exhausted before the option to use it expires or, in other terms, $Q_t = 0$ for some time $t < T$. To address this problem, we note that the dynamic in (4.5) can be directly integrated to give

$$Q_t = Q_M - \bar{Q}_t, \quad \bar{Q}_t = w\int_0^t S_s\, ds$$

where we promptly see that Q is a strictly decreasing process. Thus, abandonment of the mine occurs before T if and only if $Q_T < 0$. The density function of Q_t, γ_t, is known, albeit in a rather involved form, see for example Dufresne (2001a). Therefore, the probability above can be computed as

$$\mathbb{P}(Q_T < 0) = \int_0^\infty \gamma_T(-u)\, du.$$

We also remark that we can compute the probability to abandon the mine prior to any fixed time t_0 as easily.

Two dimensional model

Changing the dynamic as in (4.5) the operator L_1 in (4.2) changes to the Kolmogorov Operator

$$\mathcal{K}_1 := \frac{1}{2}s^2\sigma_1^2\frac{\partial^2}{\partial s^2} + rs\frac{\partial}{\partial s} - ws\frac{\partial}{\partial q} + \frac{\partial}{\partial t},$$

The Dirichlet problem that the value function V has to satisfy reads as

$$\begin{cases} \mathcal{K}_1 V = rV - f & \text{in } \Omega_1 \\ V = 0 & \text{on } \partial_p\Omega_1 \end{cases} \tag{4.6}$$

A few clarifications are needed. In the above problem, Ω_1 stands for the unbounded domain $(0,T)\times\mathbb{R}_{>0}\times(0,Q_M)$ whose components represent the domain of variation of t, S and Q respectively. The set $\partial_p\Omega_1$ denotes what we call the *parabolic boundary* of Ω_1 in analogy with the uniformly parabolic PDE theory. Precisely we set

$$\partial_p\Omega_1 = \bar{\Omega}_1 \cap \big(\{t = T\} \cup \{Q = 0\}\big). \tag{4.7}$$

The function f still represents the instantaneous gain which in this model is

$$f(s) := \big(wGs^2 - \varepsilon_P\big)^+ - \varepsilon_M.$$

Problem (4.6) shows various features which put it outside of the classical theory for Dirichlet problems. First of all, the operator, as already noted, is not parabolic as one could expect dealing with evolution processes. Furthermore, it degenerates at the boundary and the coefficients grow more than linearly. On the domain side instead, Ω_1 is unbounded in the price direction and the very partial boundary condition does not make clear if uniqueness holds, at least at first sight. We also remark that the vector field Y in this model has the same expression as before.

Three dimensional model

Applying the new the dynamic (4.5) to equation (4.4), the SDE for G in (4.4) is transformed into

$$dG = wk(\alpha - G)S\,dt + \sigma_2\sqrt{wGS}\,dW^2.$$

This can be interpreted as a standard CIR model under the *stochastic* operational time $\bar{Q}$. Note the difference with the previous case. It should be noted that by Remark 4.67 $\bar{Q}$ is a strictly increasing process so that the process G paths looks exactly like a standard CIR process ones, but travelled at a different (stochastic) velocity. In particular, properties like staying positive at every time are preserved while others, in general, are not. An example is the probability to reach a certain threshold *before a fixed* time. Now, all the comments made for the two dimensional case also apply here after both $\mathcal{K}_1$ and Ω_1 are lifted to their four dimensional counterparts

$$\mathcal{K}_2 = \mathcal{K}_1 + \frac{1}{2} wsg\sigma_2^2 \frac{\partial^2}{\partial g^2} + wk\alpha s \frac{\partial}{\partial g}, \qquad \Omega_2 = (0,T) \times \mathbb{R}_{>0} \times \mathbb{R}_{>0} \times (0, Q_M).$$

In particular, the parabolic boundary for Ω_2 is defined similarly to (4.7) as

$$\partial_p \Omega_2 = \overline{\Omega}_2 \cap \big(\{t = T\} \cup \{Q = 0\}\big).$$

That being said, the full problem is

$$\begin{cases} \mathcal{K}_2 V = rV - f & \text{in } \Omega_2 \end{cases} \tag{4.8}$$

Note that the function $f(s,g) = \big(wgs^2 - \varepsilon_P\big)^+ - \varepsilon_M$ now also depends from the variable g.

The analysis of $\mathcal{K}_2$ does not differ too much from the one of $\mathcal{K}_1$. Both are non-homogeneous Kolmogorov Operators structurally and their Lie algebra has step two. In other words, in the decomposition (1.11) we have $V_0 \oplus V_1$. For $\mathcal{K}_1$ we have $\dim V_0 = \dim V_1 = 1$ while for $\mathcal{K}_2$ we have $\dim V_0 = 2$ and $\dim V_1 = 1$.

4.3 Uniqueness

In this section we study the uniqueness of solutions to Problems (4.6),(4.8). We begin by clarifying what we mean by classical solution:

Definition 4.68. A function $u \in C(\Omega)$ is called a classical solution to problem (4.6) ((4.8)) if the derivatives $\partial_s u$, $\partial^2_{s^2} u$, Yu (and $\partial_g u$, $\partial^2_{g^2} u$) belong to $C(\Omega)$ and $\mathcal{K}u - ru = f$ pointwise in Ω. Moreover, u has to attain the boundary condition by continuity.

As usual, we treat the two dimensional problem first and then extend the results to the three dimensional one. As a matter of fact, it is important to understand the behaviour of the stochastic process $X = (S, Q)$. The first component S is a GBM hence, with probability one, it stays positive at every time and therefore this component does not escape from Ω_1. The second one, as already noted in Remark 4.67 is strictly decreasing. Thus, if we let the process X start at any point of Ω_1, the only way it has to escape Ω_1 is to pass through the parabolic boundary $\partial_p\Omega_1$.

In the three dimensional case we consider the process $X = (S, G, Q)$. The extra component G is driven by a CIR process and thus remains positive under the so called Feller condition (see Feller (1951)) $2k\alpha \geq \sigma_2^2$ which we assume to hold. Note that, even if the condition is not met, G does not became negative. The heuristic reason lies in equation (4.4): when $G = 0$ it reduces to

$$dG = k\alpha \, d\bar{Q}$$

and, being k, α both positive (and $\bar{Q}$ increasing) G is immediately reflected to the positive semiaxis. However, if Feller's condition does not hold then the process reaches zero with probability one. We refer to Cox et al. (1985) for a deeper analysis of the CIR model.

The following Faynmac-Kac type representation formula can be proved in quite a standard way (see e.g. ?? (Pas) or Karatzas and Shreve (1991)). Note that a polynomial growth condition has to be imposed due to the unboundedness of the domain. To short notations, we denote by x the couple of variables (s, g).

Proposition 4.69. *If u is a classical solution to problems* (4.6)*such that for some positive constants C and m*

$$|u(t, x)| \leq C(1 + |x|^2)^m, \qquad (t, x) \in \Omega,$$

then

$$u(t, x) = -\mathbb{E}\Big[\int_t^{\tau\wedge T} e^{-(h-t)r} f(h, X_h)\, dh \,\Big|\, X_t = x\Big], \qquad (t, x) \in \Omega, \tag{4.9}$$

where $\tau = \inf\{h > t \,|\, (s, X_s) \notin \Omega\}$ is the first exit time from Ω.

In proving Feynman-Kac type theorems the main difficulty is to prove that the expectation in (4.9) is finite. This usually involves proving that the exit time from the domain has finite expectation and the integral's argument is bounded. The first task usually regard the purely Dirichlet type problems. In fact, as we have a Cauchy-Dirichlet problem, our exit time is capped by T giving the trivial bound $\mathbb{E}[\tau \wedge T] \leq T$. On the other hand, the classical assumption $f \in L^\infty$ is not satisfied in our case. However, we have the bound

$$e^{(t-h)r}|f(S,G)| \leq C_1(1 + GS^2), \qquad t \leq h \leq T.$$

In the two dimensional case G is constant and, being S a geometric Brownian motion, also its square is. Then it follows than expectation (4.9) is bounded by the expectation of the integral of a geometric Brownian motion and it is known that this last one is finete: see, again, Dufresne (2001a). Of course, formula (4.9) gives uniqueness in the set of classical solutions with polynomial growth.

A three dimensional counterpart of the above theorem should be natural to obtain. However in this case G is no longer constant but a time-changed CIR process. This greatly complicates the problem. In fact, it is not clear at all the joint density of the process (S, G) and how to obtain the desired bound. Note that, by the nature of the time change, S and G are not independent. However, there is a heuristic argument supporting the finiteness: G is a mean reverting process and therefore large fluctuations from its long term mean are probable only at short times. On the other hand the relevant time for G is given by $\bar{Q}$, directly proportional to the time integral of the price S. Thus, if the time is small then S cannot have been too large and conversely, if S is large, then G should be near its mean. Moreover, in case of large prices, we should have $\tau > t$ and this case does not contribute to the expectation (4.9).

4.4 Existence: 2D model

In this section we prove that classical solutions to problems (4.6), (4.8) with polynomial growth do exist. As already discussed for the uniqueness, also existence is a delicate matter, essentially for the same reasons. Our proof is inspired by a procedure used in Barucci et al. (2001) to prove that the asian option pricing problem admits a solution. Coincidentally, the same solution to which we sought an approximation in Chapter 3. The main differences is that we have a boundary value problem while they studied a Cauchy problem. We illustrated the method in the two dimensional case in which important simplifications can be made. At the end we will illustrate how to modify the proof for the three dimensional model. To simplify notation, in this section we will drop the index

1 in the operator $\mathcal{K}_1$ and the domain Ω_1.

The strategy is the following:

1. apply a change of variable in order to simplify the operator;
2. find sub and super solution for the new problem;
3. get solutions u_k to suitable problems in bounded subdomains $\Omega_k \subset \Omega$;
4. prove that a subsequence of $(u_k)_k$ converges to the classical solution u.

The first step will allow us to simplify the operator treatment. The crucial step however, is the second: find upper and sub solutions will allow us to control the functions u_k found in the third step and prove convergence (step four). Eventually, in the last step, we prove u has enough regularity and attains the boundary conditions. In the last step a priori estimates will play a crucial role.

Change of Variable

We operate a change of variable or, to be more precise, we transform problem (4.6) in an equivalent one but with a more tractable operator. Let us define

$$u(t,x,y) = e^{pt}x^m V(T - t/a, x, y/b) \tag{4.10}$$

where p, m, a, b are suitable constant that will be determined later. We get

$$\begin{aligned}
s\frac{\partial V}{\partial s} &= (e^{pt}x^m)^{-1}\left(x\frac{\partial u}{\partial x} - mu\right)\\
\frac{\partial V}{\partial q} &= b(e^{pt}x^m)^{-1}\frac{\partial u}{\partial y}\\
s^2\frac{\partial^2 V}{\partial s^2} &= (e^{pt}x^m)^{-1}\left(x^2\frac{\partial^2 u}{\partial x^2} - 2mx\frac{\partial u}{\partial x} + m(m+1)u\right)\\
\frac{\partial V}{\partial t} &= -a(e^{pt}x^m)^{-1}\left(\frac{\partial u}{\partial t} - pu\right)
\end{aligned}$$

Now, choosing the parameters to be

$$a = \frac{\sigma_1^2}{2}, \quad b = -\frac{\sigma_1^2}{2w}, \quad m = \frac{r}{\sigma_1^2}, \quad p = \frac{2}{\sigma_1^2}mr - m(m+1),$$

problem (4.6) is transformed into

$$\begin{cases}
(\mathcal{K}u)(t,x,y) = \widetilde{f}(t,x,y) & \text{if } (t,x,y) \in \widetilde{\Omega}\\
u(t,x,y) = 0 & \text{if } (t,x,y) \in \partial_p\widetilde{\Omega}
\end{cases} \tag{4.11}$$

where $\widetilde{\Omega} := (0, aT) \times \mathbb{R}_{>0} \times (bQ_M, 0)$ and

$$\mathcal{K}u = x^2 \frac{\partial^2 u}{\partial x^2} + x \frac{\partial u}{\partial y} - \frac{\partial u}{\partial t} \qquad \widetilde{f}(t,x,y) = \frac{2}{\sigma_1^2} x^m e^{pt} \big(\varepsilon_M - \big(wGx^2 - \varepsilon_P\big)^+\big) \tag{4.12}$$

The parabolic boundary $\partial_p \Omega$ is transformed into

$$\partial_p \widetilde{\Omega} = \overline{\widetilde{\Omega}} \cap (\{t = 0\} \cup \{y = 0\}).$$

Note that as b is negative the interval $(bQ_M, 0)$ lies in the negative semi axis and that the lateral condition $y = 0$ is formally the same as before ($Q = 0$) but is approached by the other side.

We remark that the new operator is the same as the Kolmogorov operator for the asian options studied in Chapter 3.

Sub- and super-solutions

We call a function u a sub-solution to Problems (4.11) *for a general domain* Ω if it holds

$$\begin{cases} \mathcal{K}u \geq \widetilde{f} & \text{in } \widetilde{\Omega} \\ u \leq 0 & \text{on } \partial\widetilde{\Omega} \end{cases}$$

and a super-solution if the above inequalities are reversed. Note that the second inequality is required to hold on the *topological* boundary of Ω.

The importance of having explicit sub- and super-solutions for our problems is made clearer by the Picone's maximum principle: any *classical* solutions we seek must be bounded below and above by a sub-solution and super-solution respectively.

Let $\underline{u}(t,x,y) := -\frac{2\varepsilon_M}{\sigma_1^2} x^m e^{\alpha t}$, then, for $\alpha > \max\{m(m-1)+1, p\}$ we have

$$\begin{aligned} \mathcal{K}\underline{u} &= (m(m-1) - \alpha)\underline{u} = \frac{2\varepsilon_M}{\sigma_1^2} x^m e^{\alpha t} (\alpha - m(m-1)) x^m e^{\alpha t} \\ &\geq \frac{2\varepsilon_M}{\sigma_1^2} x^m e^{\alpha t} \geq \frac{2\varepsilon_M}{\sigma_1^2} x^m e^{pt} \geq \widetilde{f} \end{aligned}$$

and, being clearly negative, $\underline{u}$ is a sub-solution. Similarly, defining $\bar{u} = \frac{2qG}{\sigma_1^2} x^{m+2} e^{\beta t}$ with $\beta > \max\{(m+1)(m+2)+1, p\}$ we get

$$\begin{aligned} \mathcal{K}\overline{u} &= ((m+1)(m+2) - \beta)\overline{u} = \frac{2wG}{\sigma_1^2} x^{m+2} e^{\beta t} ((m+1)(m+2) - \beta) \\ &\leq -\frac{2wG}{\sigma_1^2} x^{m+2} e^{\beta t} \leq -\frac{2wG}{\sigma_1^2} x^{m+2} e^{pt} \leq \widetilde{f}, \end{aligned}$$

a super-solution. We also explicitly note that $\underline{u} \leq \bar{u}$ in $\widetilde{\Omega}$.

The problem on subdomains

We proceed to study the problem (4.11) in the bounded subdomains

$$\widetilde{\Omega}_n := \widetilde{\Omega} \cap \{(t, x, y) \,|\, n^{-1} < x < n\} \qquad n \in \mathbb{N}$$

Precisely we study

$$\begin{cases} \mathcal{K}u = \widetilde{f} & \text{in } \widetilde{\Omega}_n \\ u = 0 & \text{on } \partial\widetilde{\Omega}_n \end{cases}$$

And note that now the vanishing condition holds on the whole boundary.

We will use the following two results taken from the literature. The first one concerns existence of a so-called *generalized* solution to the problem abovewhile the second studies the regularity of the boundary.

Proposition 4.70. *Let $D \subset \mathbb{R}^3$ be any bounded domain such that $\bar{D} \subset \{x \neq 0\}$, let $h \in C^{\alpha}_{B,M}(D)$ and $g \in C(\partial D)$. Then there exists a function $u \in C^{2+\alpha}_{B,M,loc}(D)$, solution to $\mathcal{K}u = h$ in D and such that*

$$\lim_{z \to \zeta} u(z) = g(\zeta),$$

for every $\mathcal{K}$-regular point of ∂D. The function u is called generalized solution.

A few comments are needed. The requirement $\bar{D} \subset \{x \neq 0\}$ is not explicitly stated in Manfredini (1997) where it is required to second order part of the Kolmogorov operator to do not degenerate. Of course, this assumption is equivalent to ours for the operator $\mathcal{K}$. The spaces $C^{\alpha}_{B,M}(D)$, $C^{2+\alpha}_{B,M,\text{loc}}(D)$ are defined by the fineteness of the norms (1.21),(1.22) respectively. At last, few definitions taken from potential theory.

Definition 4.71. A point $\zeta \in \partial D$ is called regular for the operator $\mathcal{K}$ if there exists a so-called *barrier function* ω_ζ, a function defined in a neighbourhood V of ζ such that

- $\omega_\zeta \in C^{2+\alpha}_{B}(V)$;
- $\omega_\zeta(\zeta) = 0$;
- $\omega_\zeta(z) > 0$ in $\overline{V \cap D} \setminus \{\zeta\}$;

- $\mathcal{K}\omega_\zeta < 0$ in $V \cap D$.

An *outer normal vector to the open set D at the point $z_0 \in \partial D$* is a vector $\nu \in \mathbb{R}^3$ such that the Euclidean ball of center $z_0 + \nu$ and radius $|\nu|$ is contained in $\mathbb{R}^3 \setminus D$.

The next result, a particular case of Theorem 6.1 in Manfredini (1997), gives two simple criteria to verify if a point is regular:

Proposition 4.72. *Under the same hypothesis as Proposition 4.70 above let (t_0, x_0, y_0) be a point of ∂D. If there exists a normal outer vector $\nu = (\nu_t, \nu_x, \nu_y)$ such that one of*

1. *$\nu_x \neq 0$ or*
2. *$\nu_x = 0$ but $x_0\nu_y - \nu_t > 0$ and there exists a positive constant ε such that $x_0^2\varepsilon^2 \leq x_0\nu_y - \nu_t$ and*
$$\{(t,x,y) \in \mathbb{R}^3 | (t-t_0-\varepsilon^2\nu_t)^2 + \varepsilon^2(x-x_0)^2 + (y-y_0-\varepsilon^2\nu_y)^2 \leq \varepsilon^4\} \subset \mathbb{R}^3 \setminus D$$

holds, then (t_0, x_0, y_0) is a regular point.

Now, let apply this proposition to the study of $\partial\widetilde{\Omega}_n$. It is clear that any boundary point described by one of the equations $x = n^{-1}$, $x = n$ satisfies criterion 1. Points described by the equation $t = 0$, $y = 0$ instead satisfy criterion 2 for an ε suitably small and with normal outer vectors $\nu = (-1, 0, 0)$ and $\nu = (0, 0, 1)$ respectively.

Note that the last two class of points are exactly the one in the parabolic boundary of $\widetilde{\Omega}$. If we were to cut $\widetilde{\Omega}$ away from the degenerate plane $x = 0$ then a similar result would apply also for the problem (4.11).

We now wish to apply the results stated so far. To do so we only need to verify that $\widetilde{f} \in C^\alpha_{B,M}(\widetilde{\Omega}_n)$. Consider the set $\widetilde{\Omega}$. As a function of time, $\widetilde{f}$ is smooth in it while it is just locally (Euclidean) Lipschitz in s. Nonetheless, if we confine ourself to the bounded domains $\widetilde{\Omega}_n$ then there is no problem and $\widetilde{f} \in C^\alpha_{B,M}(\widetilde{\Omega}_n)$. Note however that the norm explodes as n goes to infinity.

Collecting all together we obtain a sequence of function $(u_n)_{n\in\mathbb{N}}$, each in $C^\alpha_{B,M}(\widetilde{\Omega}_n)$ and attaining the null boundary condition on $\partial\widetilde{\Omega} \cap (\{t = 0\} \cup \{y = 0\})$ and verifying the uniform estimates

$$\underline{u} \leq u_n \leq \overline{u}, \quad \text{in } \widetilde{\Omega}_n, \quad n \in \mathbb{N}. \tag{4.13}$$

In particular, each function u_n is bounded in its domain.

The candidate solution

We now proceed to extract a convergent subsequence from (u_n). We will make heavy use of the following theorem whose proof can be found in Di Francesco and Polidoro (2006).

Proposition 4.73. *Let $D \subset \mathbb{R}^3$ be any bounded domain such that $\bar{D} \subset \{x \neq 0\}$, $h \in C^{0,\alpha}_{B,d}(D)$ and u a bounded function belonging to $C^{2+\alpha}_{B,loc}(D)$ such that $\mathcal{K}u = h$ in D. Then $u \in C^{2+\alpha}_{B,d}(D)$ and*

$$|u|_{2,\alpha,d,D} \leq C\Big(\|u\|_{L^\infty(D)} + |h|_{2+\alpha,d,D}\Big) \tag{4.14}$$

where the constant C does not depend on u.

For the definitions of the norms in (4.14) and the intrinsic spaces in the statement, see Section 1.3.1.

Now, define the sets

$$A_n = \{(x, y, t) \in \widetilde{\Omega}_k \,\big|\, aT/n < t < aT(1 - 1/n)),\, bQ_M(1 - 1/n) < y < 1/n\}$$

then $\widetilde{\Omega} = \bigcup_n A_n$ and every A_n is compactly contained in A_{n+1}.

As any u_m solves $\mathcal{K}u_m = \widetilde{f}$ in A_n for every $m \geq n$ we have, by (4.14) and (4.13)

$$\begin{aligned}|u_m|_{2,\alpha,d,A_n} &\leq C\Big(\|u_m\|_{L^\infty(A_n)} + |\widetilde{f}|_{2+\alpha,d,A_n}\Big)\\ &\leq C\Big(\|\overline{u}\|_{L^\infty(A_n)} + \|\underline{u}\|_{L^\infty(A_n)} + |\widetilde{f}|_{2+\alpha,d,A_n}\Big),\end{aligned}$$

and the last bound does *not* depend on $m \geq n$.

Consider $(u_n)_{n\geq 2}$ in A_1. By the above estimate, this sequence is uniformly bounded and thus, in particular, uniformly intrinsically Hölder continuous. By Lemma 1.8 this implies the uniform Euclidean Hölder continuity, hence by the classical Ascoli-Arzel theorem, there exists an uniformly convergent subsequence $(u_{1,j})_j$ with limit a continuous function on $\bar{A}_1$, say v_1. It follows that v_1 is a weak solution to $\mathcal{K}u = \widetilde{f}$ and thus, by the hypoellipticity of $\mathcal{K}$, v_1 is actually a classical solution.

The procedure can be repeated with the subsequence $(u_{1,j})_{j\in\mathbb{N}}$ in place of (u_n) on the set $\bar{A}_2$. This gives us a classical solution, v_2, on A_2, *that agrees with v_1 in A_1*. Iterating the reasoning we thus obtain a sequence of functions $(v_k)_k$ that allow us to define a candidate solution to $\mathcal{K}u = \widetilde{f}$ in $\widetilde{\Omega}$ as follow

$$u(t, x, y) = v_j(t, x, y) \qquad \text{if } (t, x, y) \in A_j$$

the definition is well posed as every v_j agrees with v_i in A_j if $i \geq j$ and the A_j cover $\widetilde{\Omega}$.

We are left with the boundary condition. Recall that, as discussed in the previous subsection, every u_j meets the null boundary condition on $\partial_p\widetilde{\Omega} \cap \overline{\widetilde{\Omega}}_j$. However, this is not enough to conclude that this holds for the limit u. Nevertheless, the barrier functions give an uniform estimates on the rate of convergence and we can conclude that u is the (unique) classical solution to Problem (4.11).

Remark 4.74. By (4.13), u exhibits polynomial growth and, by the Feynman-Kaz formula, is the only solution of (4.6) with such property.

Remark 4.75. A sub solution $\underline{V}$ for the original problem (4.6) is simply given by $-\varepsilon_M/r$. It is worth noting that such bound is negative. This should may be surprising but in fact is expected as, in the model, we mine even in the case the price S is not high enough to cover the costs ε_M. In the case $r = 0$, an alternative sub-solution is given by $\underline{V} = -\varepsilon_M e^{T-t}$. Note that also this function is strictly negative on Ω_1 even though it is not constant (nevertheless, it is uniformly bounded).

Building on the previous remark, it is interesting to ask in which condition the mine is profitable i.e. when $V > 0$. Note that we are interested in the true value function V so we analyse the original problem (4.11). It turns out, the answer depends on the sign of the datum $-f = \varepsilon_M - (ws^2g - \varepsilon_P)^+$; its sign divides the domain Ω into the two regions Ω^+, Ω^- separated by the plane

$$s = \sqrt{\frac{\varepsilon_M + \varepsilon_P}{wG}}.$$

It is now clear that the zero function represent both a sub-solution for $\mathcal{K}_1$ in Ω^+, (the set in which $-f \geq 0$) and a super-solution for $\mathcal{K}_1$ in Ω^- and we have thus exactly identified the region of positivity of V. We also note that, by continuity, $V = 0$ on the plane.

4.5 Existence: 3D model

We start recalling the three dimensional problem (4.8)

$$\begin{cases} \mathcal{K}_2 V = rV - f & \text{in } \Omega_2 \\ V = 0 & \text{on } \partial_p \Omega_2 \end{cases}$$

where $\Omega_2 = (0,T) \times \mathbb{R}^2_{>0} \times (0, Q_M)$, $\partial_p \Omega_2 = \overline{\Omega}_2 \cap \big(\{t=T\} \cup \{Q=0\}\big)$ and

$$\mathcal{K}_2 = \frac{1}{2} s^2 \sigma^2 \frac{\partial^2}{\partial s^2} + \frac{1}{2} wsg\sigma_2^2 \frac{\partial^2}{\partial g^2} + rs\frac{\partial}{\partial s} + wk\alpha s \frac{\partial}{\partial g} - ws\frac{\partial}{\partial q} + \frac{\partial}{\partial t}.$$

As we have seen in the two dimensional case, the main ingredients we need are three: the ability to solve Cauchy Problems on bounded subdomains, Shauder estimates and sub- and super-solutions. The main problem in treating the three dimensional model is that the operator $\mathcal{K}_2$ cannot be significantly simplified by a change of variable analogous to the one in (4.10). This is essentially due to the presence of g instead of g^2 in the coefficient of ∂_g^2. However, the extra first order derivatives ∂_s, ∂_g are derivatives along directions on which the second order part of the operator is not degenerate or, in other words, the operator is in the form (3.3). Such operators were studied in Di Francesco and Polidoro (2006) where the exact same results in Propositions 4.70 and 4.73 were proved for this larger class of operators.

Remark 4.76. We explicitly remark that all the result for Kolmogorov operators are available under the assumption that the second order part of the operator must be elliptic on the space spanned by its derivatives uniformly on the domain. For the operator $\mathcal{K}_2$ treated here this condition is fulfilled if the domain is compactly contained in the set $\{(t,s,w,q) \in \mathbb{R}^4 \,|\, s>0, g>0\}$.

Particular attention should be given to the analogous of Proposition 4.72. First of all, we now have an extra dimension; secondly the vector field $Y = -ws\partial_q + \partial_t$ has the opposite sign of the analogous $x\partial_y - \partial_t$ in (4.12). The correct statement is then the following.

Proposition 4.77. *Let D be a bounded domain of $\mathbb{R}^4$ such that $\bar{D} \subset \{(t,s,w,q) \in \mathbb{R}^4 \,|\, s>0, g>0\}$. And let $P = (t_0, s_0, g_0, q_0)$ be a point of ∂D. If there exists a normal outer vector $\nu = (\nu_t, \nu_s, \nu_g, \nu_q)$ such that one of the two condition below holds, then P is a regular point.*

1. $(\nu_s, \nu_g) \neq (0,0)$;
2. $(\nu_s, \nu_g) = (0,0)$ *but* $\nu_t - x_0\nu_q > 0$ *and there exists a positive constant* ε *such that* $\lambda_0 \varepsilon^2 \leq \nu_t - x_0 \nu_q$ *and the set of* $(t,s,g,q) \in \mathbb{R}^4$ *with*
$$(t - t_0 - \varepsilon^2\nu_t)^2 + \varepsilon^2(s-s_0)^2 + \varepsilon^2(g-g_0)^2 + (q - q_0 - \varepsilon^2\nu_q)^2 \leq \varepsilon^4$$
does not intersect D. *Here* λ_0 *is the constant* $\frac{s_0^3 g_0 \sigma^2 \sigma_2^2 w}{4}$.

With this result one can easily check that any point in $\partial_p\Omega_2$ is regular for $\mathcal{K}_2$. In particular we are able to obtain generalized solutions to the problems

$$\begin{cases} \mathcal{K}_2 V = rV - f & \text{in } \Omega_{2,n} \\ V = 0 & \text{on } \partial\Omega_{2,n} \end{cases}$$

where

$$\Omega_{2,n} := \Omega_2 \cap \{(t,s,g,q) \in \mathbb{R}^4 \,|\, n^{-1} < s, g < n\}.$$

Such solutions, as in the 2D case, attain the null boundary condition on $\partial\Omega_{2,n} \cap \{t = T \text{ or } q = 0\}$.

The only ingredient missing now to get the machinery to work are the sub and super-solution. We first re-define what we mean as the operator contains a zero-order term $-r \leq 0$. A sub-solution for $\mathcal{K}$ in the domain D is a regular function u such that

$$\begin{cases} \mathcal{K}_0 u - ru \geq -f & \text{in } D \\ u \leq 0 & \text{on } \partial D \end{cases}$$

A function which satisfies the reversed inequalities is called super-solution.

As already seen in the two dimensional case, the constant function $\underline{V} = -\frac{\varepsilon_M}{r}$ provides a sub-solution.

The super-solution is a bit more trickier: define

$$\overline{V}(s,g,q) = sgq + Q_M(k\alpha)\,sq,$$

which is clearly positive in Ω_2. It is clear that $rs\partial_s\overline{V} = r\overline{V}$ and thus

$$\begin{aligned} \mathcal{K}\overline{V} - r\overline{V} &= w(k\alpha s\partial_g - s\partial_q)\overline{V} \\ &= w\big(-s^2 g + k\alpha\, s^2 q - k\alpha Q_M\, s^2\big) \\ &\leq -ws^2 g \leq -(ws^2 g - \varepsilon_P)^+ \leq -f, \end{aligned}$$

where in the first estimate we used $q \leq Q_M$ in Ω_2.

Remark 4.78. A discussion similar to the one in Remark 4.75 can be made also in this case. We just note that the equation

$$s = \frac{C}{\sqrt{g}}, \qquad C = \sqrt{\frac{\varepsilon_M + \varepsilon_P}{w}},$$

now describes an hypersurface in $\mathbb{R}^4$.

Appendix A

Connectivity results

Here we prove an intermediate value theorem for a scaled family of embeddings of the sphere in the space. More precisely, we suppose to have a family of embeddings of the sphere that degenerates to a constant as time approaches zero (or, equivalently, that ”grows” from a constant to a regular embedding). Then, we prove that any intermediate point between the constant and the final surface is covered by at least an embedding for a positive time. In other words, such a family cannot have holes in its image. This is intuitively true but not completely trivial to prove. The key idea of the proof was provided us by Stefano Pagliarani to whom we are much in debt.

Theorem A.1. *Let $f : [0,1] \times \mathbb{S}^n \to \mathbb{R}^{n+1}$ a continuous map such that:*

1. *$f(t,\cdot)$ is an embedding (homeomorphism on its image) for any $t \in]0,1]$;*
2. *$f(0,\cdot) \equiv 0$.*

Then, every $y \in f(1,\overset{\circ}{\mathbb{S}^n})$ belongs to the image of f.

Here above $f(1,\mathbb{S}^n)$ denotes, with a slight abuse of notation, the image of the function $f(1,\cdot)$, whereas $f(1,\overset{\circ}{\mathbb{S}^n})$ denotes the “internal” part of $f(1,\mathbb{S}^n)$, meaning the bounded domain (open connected set) of $\mathbb{R}^{n+1}$ whose border is $f(1,\mathbb{S}^n)$ (Jordan-Brouwer separation theorem).

The idea of the proof is to use an argument of contraction of the volume. We first have the following preliminary

Lemma A.2. *For any $t_0 \in]0,1]$ we have*

$$\lim_{t\to t_0} \nu\left(f(t, \overset{\circ}{\mathbb{S}^n})\right) = \nu\left(f(t_0, \overset{\circ}{\mathbb{S}^n})\right) > 0, \tag{A.1}$$

where ν denotes the Lebesgue's measure on $\mathbb{R}^{n+1}$.

Proof. Let $B(0,1)$ denote the unitary ball in $\mathbb{R}^{n+1}$ centered at the origin. It is not restrictive to assume that f can be extended to a continuous map $\widetilde{f}$: $[0,1] \times \overline{B(0,1)} \to \mathbb{R}^{n+1}$ such that $\widetilde{f}\big(t, B(0,1)\big) = f(t, \overset{\circ}{\mathbb{S}^n})$ for any $t \in]0,1]$. Now, by the uniform continuity of $\widetilde{f}$ along with dominated convergence theorem, one has

$$\lim_{t\to t_0} \nu\left(f(t, \overset{\circ}{\mathbb{S}^n})\right) = \lim_{t\to t_0} \nu\left(\widetilde{f}\big(t, B(0,1)\big)\right) = \nu\left(\widetilde{f}\big(t_0, B(0,1)\big)\right) = \nu\left(f(t_0, \overset{\circ}{\mathbb{S}^n})\right).$$

$\square$

Proof of Proposition A.1. We prove it by contradiction. Assume there is $y_0 \in f(1, \overset{\circ}{\mathbb{S}^n})$ such that $f(t,v) \neq y_0$ for any $(t,y) \in [0,1] \times \mathbb{S}^n$. Let us define

$$t_0 := \inf\{t \in [0,1] : y_0 \in f(t, \overset{\circ}{\mathbb{S}^n})\}.$$

Note that $y_0 \neq 0$ by hypothesis 2, and thus, by continuity of f it has to be $t_0 > 0$. We can now distinguish two cases: $y_0 \in f(t_0, \overset{\circ}{\mathbb{S}^n})$ (t_0 is a minimum), $y_0 \notin f(t_0, \overset{\circ}{\mathbb{S}^n})$ (t_0 is only an infimum).

<u>1st case: $y_0 \in f(t_0, \overset{\circ}{\mathbb{S}^n})$</u>: Since f is uniformly continuous, for any $\varepsilon > 0$ there exists $t_1 < t_0$ such that

$$f(t, \mathbb{S}^n) \subseteq \operatorname{ann}\left(f(t_0, \mathbb{S}^n), \varepsilon\right) \qquad \forall t \in]t_1, t_0[,$$

where $\operatorname{ann}\left(f(t_0, \mathbb{S}^n), \varepsilon\right)$ is the annulus of the points whose distance from $f(t_0, \mathbb{S}^n)$ is ε at most. Observe now that, for ε suitably small, we also have

$$f(t, \overset{\circ}{\mathbb{S}^n}) \subseteq \operatorname{ann}\left(f(t_0, \mathbb{S}^n), \varepsilon\right) \qquad \forall t \in]t_1, t_0[. \tag{A.2}$$

In fact, for ε small enough we have $\operatorname{dist}\big(y_0, f(t_0, \mathbb{S}^n)\big) > \varepsilon$. Therefore, if there was $t \in]t_1, \delta_0[$ and $y \in f(t, \overset{\circ}{\mathbb{S}^n}) \cap \big(\mathbb{R}^{n+1} \setminus \operatorname{ann}\left(f(t_0, \mathbb{S}^n), \varepsilon\right)\big)$, then by Jordan-Brouwer separation theorem it would also be $y_0 \in f(t, \overset{\circ}{\mathbb{S}^n})$, which is impossible by definition of δ_0. Finally, by (A.2) we have

$$\nu\left(f(t, \overset{\circ}{\mathbb{S}^n})\right) \leq \nu\big(\operatorname{ann}\left(f(t_0, \mathbb{S}^n), \varepsilon\right)\big),$$

and thus

$$\nu\left(f(t,\overset{\circ}{\mathbb{S}^n})\right) \to 0 \qquad \text{as } t \to t_0^-,$$

which contradicts (A.1) and thus concludes the proof.

<u>2st case: $y_0 \notin f(t_0,\overset{\circ}{\mathbb{S}^n})$:</u> It is a simple modification of the proof for the 1st case. $\square$

Appendix B

Estimates on derivatives

In this appendix we prove some preliminary estimates on the spatial derivatives of solutions of constant coefficient-Kolmogorov operators: in particular, we prove estimates for the derivatives of $u_n^{(\bar{z})}$ defined by (3.14)-(3.15). Throughout this section $\bar{z} \in \mathbb{R} \times \mathbb{R}^d$ is fixed.

Proposition B.1. *Let $k \in [0, 2r+1]$, $\beta \in \mathbb{N}_0^d$ with $|\beta|_B > 0$. If $\psi \in C_B^k(\mathbb{R}^d)$ then the solution $u_0^{(\bar{z})}$ of the Cauchy problem* (3.14) *satisfies*

$$\left| D_x^\beta u_0^{(\bar{z})}(t,x) \right| \leq C(T-t)^{\frac{k-|\beta|_B}{2}}, \qquad 0 \leq t < T, \ x \in \mathbb{R}^d,$$

where C is a positive constant that depends only on M, μ, B, T, β and $\|\psi\|_{C_B^k(\mathbb{R}^d)}$.

Proof. We prove the case $k \in]0, 2r+1]$ since the case $k = 0$ is straightforward. We first note that, since $\Gamma_0^{(\bar{z})}$ is a density and $|\beta|_B > 0$, we have

$$D_x^\beta \int_{\mathbb{R}^d} \Gamma_0^{(\bar{z})}(t,x;T,y)dy = 0.$$

and therefore

$$\begin{aligned} D_x^\beta u_0^{(\bar{z})}(t,x) &= \int_{\mathbb{R}^d} \psi(y) D_x^\beta \Gamma_0^{(\bar{z})}(t,x;T,y)dy \\ &= \int_{\mathbb{R}^d} \Big(\psi(y) - \psi\big(e^{(T-t)B}x\big) \Big) D_x^\beta \Gamma_0^{(\bar{z})}(t,x;T,y)dy. \end{aligned}$$

Since $\psi \in C_B^k(\mathbb{R}^d)$, we obtain

$$\left| D_x^\beta u_0^{(\bar{z})}(t,x) \right| \leq \|\psi\|_{C_B^k(\mathbb{R}^d)} \int_{\mathbb{R}^d} \left[y - e^{(T-t)B}x \right]_B^k \left| D_x^\beta \Gamma_0^{(\bar{z})}(t,x;T,y) \right| dy$$

$$\leq C\|\psi\|_{C_B^k(\mathbb{R}^d)}(T-t)^{\frac{k-|\beta|_B}{2}}\int_{\mathbb{R}^d}\Gamma^{2M}(t,x;T,y)dy,$$

where the second inequality follows from a direct estimate on the derivatives of $\Gamma_0^{(\bar{z})}$ (see, for example, Section 2 in Polidoro (1994)) and Γ^{2M} is the fundamental solution of the Kolmogorov operator $\mathcal{K}^{2M}$ as defined in (3.10). □

In the next lemmas we will use the following result proved in Lanconelli and Polidoro (1994).

Lemma B.2. *The following homogeneity relations hold*

$$\mathbf{C}_{\bar{z}}(t) = D_0(\sqrt{t})\mathbf{C}_{\bar{z}}(1)D_0(\sqrt{t}), \tag{B.1}$$

$$\mathbf{M}_{\bar{z}}(t) = D_0(\sqrt{t})\mathbf{M}_{\bar{z}}(1)D_0(\sqrt{t}), \tag{B.2}$$

$$e^{tB} = D_0(\sqrt{t})e^{B}D_0\left(\frac{1}{\sqrt{t}}\right), \tag{B.3}$$

for any $t > 0$.

Notation B.3. From now to the end of this section, we use the Greek letters $\alpha, \beta, \gamma, \delta, \nu$ to denote multi-indexes in $\mathbb{N}_0^d$, and $|\alpha| = \sum_{i=1}^d \alpha_i$ is the standard Euclidean height of α. To simplify notations, if I is any family of indexes, we use the unconventional notation

$$\sideset{}{^\bullet}\sum_{\ell\in I}\pi_\ell = \sum_{\ell\in I}c_\ell\pi_\ell$$

for a sum where the constants c_ℓ depend only on $\bar{z}, B, N, T, a_{ij}, a_i$ and are uniformly bounded by a constant that depends only on M, μ, T, N and B.

Lemma B.4. *Let*

$$\mathcal{W}(t) = e^{-tB^*}\nabla_x, \qquad t\in\mathbb{R},$$

denote the differential operators appearing in (3.22) *and by $\mathcal{W}^\alpha(t)$ the composition*[1]

$$\mathcal{W}^\alpha(t) = \mathcal{W}_1^{\alpha_1}(t)\cdots\mathcal{W}_d^{\alpha_d}(t). \tag{B.4}$$

The following representation holds true:

$$\mathcal{W}^\beta(t) = \sideset{}{^\bullet}\sum_{\substack{|\alpha|=|\beta|\\ |\alpha|_B\geq|\beta|_B}} t^{\frac{|\alpha|_B-|\beta|_B}{2}}D_x^\alpha.$$

[1]Operator $\mathcal{W}^\alpha(t)$ in (B.4) is well defined since the components of $\mathcal{W}(t)$ commute.

Proof. It suffices to prove the statement for a single $\mathcal{W}_i(t)$. Using the relations in Lemma B.2, we have

$$\begin{aligned}\mathcal{W}_i(t) &= \sum_{j=1}^{d} D_0\left(\frac{1}{\sqrt{t}}\right)_{ii} e^{-B^*}_{ij} D_0\big(\sqrt{t}\big)_{jj}\partial_{x_j}\\ &= t^{-\frac{\sigma_i}{2}}\sum_{j=1}^{d} e^{-B^*}_{ij} t^{\frac{\sigma_j}{2}}\partial_{x_j},\end{aligned}$$

with σ_i as in (1.18). The result follows noting that the intrinsic order of ∂_{x_j} is exactly σ_j. Moreover, as the matrix e^{-B^*} is upper triangular the sum actually ranges over $j = i, \dots, d$ and thus $\sigma_j - \sigma_i$ is always a nonnegative integer. $\square$

Next step is the study of the operator $\mathcal{M}^{(\bar z)}(t,x)$: we recall that, by Proposition 3.56, the components of $\mathcal{M}^{(\bar z)}(t,x)$ commute when applied to $\Gamma_0^{(\bar z)}$ and more generally to $u_n^{(\bar z)}$ and its derivatives.

Lemma B.5. *For any $\beta \in \mathbb{N}_0^d$, we have*

$$\begin{aligned}&\big(\mathcal{M}^{(\bar z)}(s-t,x) - e^{(s-\bar t)B}\bar x\big)^{\beta} = \qquad\qquad (\text{B.5})\\ &\qquad \dot{\sum_{\substack{|\delta|+|\alpha|\le|\beta|\\ |\delta|_B-|\alpha|_B\le|\beta|_B}}} (s-t)^{\frac{|\beta|_B+|\alpha|_B-|\delta|_B}{2}}\big(x-e^{(t-\bar t)B}\bar x\big)^{\delta} D_x^{\alpha}.\end{aligned}$$

Proof. First of all, let us note that

$$\mathcal{M}^{(\bar z)}(s-t,x) - e^{(s-\bar t)B}\bar x = e^{(s-t)B}\left(x - e^{(t-\bar t)B}\bar x + \mathbf{M}_{\bar z}(s-t)\nabla_x\right),$$

and it is not restrictive to take $\bar x = 0$ and $t = 0$. We proceed now by induction on $|\beta|$. If $|\beta| = 1$ then $\beta = \mathbf{e}_i$ where $\mathbf{e}_i$ is the i-th element of the canonical basis of $\mathbb{R}^d$. A direct computation shows

$$\begin{aligned}\big(\mathcal{M}^{(\bar z)}(s,x)\big)^{\mathbf{e}_i} &= \dot{\sum_{\substack{|\delta|=1\\ |\delta|_B\le|\mathbf{e}_i|_B}}} s^{\frac{|\mathbf{e}_i|_B-|\delta|_B}{2}}\big(x^{\delta} + (\mathbf{M}_{\bar z}(s)\nabla_x)^{\delta}\big) && (\text{by (B.3)})\\ &= \dot{\sum_{\substack{|\delta|=1\\ |\delta|_B\le|\mathbf{e}_i|_B}}} s^{\frac{|\mathbf{e}_i|_B-|\delta|_B}{2}}\left(x^{\delta} + s^{\frac{|\delta|_B}{2}}\dot{\sum_{|\nu|=1}} s^{\frac{|\nu|_B}{2}} D_x^{\nu}\right), && (\text{by (B.2)})\end{aligned}$$

which proves (B.5) with $\beta = \mathbf{e}_i$. We now assume the statement to hold for $|\beta| \le n$, and prove it true for $\beta + \mathbf{e}_i$. By inductive hypothesis applied to both

β and $\mathbf{e}_i$ we get

$$\begin{aligned}\left(\mathcal{M}^{(\bar{z})}(s,x)\right)^{\beta+\mathbf{e}_i} &= \dot{\sum_{\substack{|\delta^1|+|\alpha^1|\le 1\\ |\delta^1|_B-|\alpha^1|_B\le|\mathbf{e}_i|_B}}} \Bigg(s^{\frac{|\mathbf{e}_i|_B+|\alpha^1|_B-|\delta^1|_B}{2}} \\ &\times \dot{\sum_{\substack{|\delta^2|+|\alpha^2|\le|\beta|\\ |\delta^2|_B-|\alpha^2|_B\le|\beta|_B}}} s^{\frac{|\beta|_B+|\alpha^2|_B-|\delta^2|_B}{2}} x^{\delta^1} D_x^{\alpha^1}\left(x^{\delta^2} D_x^{\alpha^2}\right)\Bigg) \\ &= \dot{\sum_{\substack{|\delta|+|\alpha|\le|\beta+\mathbf{e}_i|\\ |\delta|_B-|\alpha|_B\le|\beta+\mathbf{e}_i|_B}}} s^{\frac{|\beta+\mathbf{e}_i|_B+|\alpha|_B-|\delta|_B}{2}} x^{\delta} D_x^{\alpha},\end{aligned}$$

where we set $\delta = \delta^1 + \delta^2$ and $\alpha = \alpha^1 + \alpha^2$. □

Lemma B.6. *For any $n \in \mathbb{N}$, with $n \le N$, we have the following representation*

$$\mathcal{G}_n^{(\bar{z})}(t,s,x) = \dot{\sum_{(\alpha,\delta)\in I_n}} (s-t)^{\frac{|\alpha|_B-|\delta|_B+n-2}{2}} (x - e^{(t-\bar{t})B}\bar{x})^{\delta} D_x^{\alpha}, \tag{B.6}$$

where

$$I_n = \{(\alpha,\delta) \in \mathbb{N}_0^d \times \mathbb{N}_0^d \mid 1 \le |\alpha| \le n+2,\ |\delta|_B \le n,\ |\alpha|_B - |\delta|_B + n - 2 \ge 0\}.$$

Proof. Using the definition of $\mathcal{G}_n^{(\bar{z})}(t,s,x)$ in (3.22), the proof is a straightforward application of Lemmas B.4 and B.5. □

Lemma B.7. *For any $n \in \mathbb{N}$, with $n \le N$, we have the following representation*

$$\mathcal{L}_n^{(\bar{z})}(t,T,x) = \dot{\sum_{(\alpha,\delta)\in J_n}} (T-t)^{\frac{|\alpha|_B-|\delta|_B+n}{2}} (x - e^{(t-\bar{t})B}\bar{x})^{\delta} D_x^{\alpha}, \tag{B.7}$$

where

$$J_n = \{(\alpha,\delta) \in \mathbb{N}_0^d \times \mathbb{N}_0^d \mid 1 \le |\alpha| \le 3n,\ |\delta|_B \le n,\ |\alpha|_B - |\delta|_B + n \ge 0\}. \tag{B.8}$$

Proof. For greater convenience we recall the expression of $\mathcal{L}_n^{(\bar{z})}(t,T,x)$ as given in (3.21):

$$\mathcal{L}_n^{(\bar{z})}(t,T,x) = \sum_{h=1}^{n} \sum_{i\in I_{n,h}} L_{h,i}(t,T,x),$$

where

$$L_{h,i}(t,T,x) := \int_t^T ds_1 \int_{s_1}^T ds_2 \cdots \int_{s_{h-1}}^T ds_h \mathcal{G}_{i_1}^{(\bar{z})}(t,s_1,x) \cdots \mathcal{G}_{i_h}^{(\bar{z})}(t,s_h,x),$$

and $I_{n,h} = \{i = (i_1, \dots, i_h) \in \mathbb{N}^h \mid i_1 + \cdots + i_h = n\}$, for $1 \le h \le n$. We prove that, for fixed $h \in \{1, \dots, n\}$ and $i \in I_{n,h}$ it holds

$$L_{h,i}(t,T,x) = \sum^{\bullet}_{(\alpha,\delta)\in J_n} (T-t)^{\frac{|\alpha|_B - |\delta|_B + n}{2}} (x - e^{(t-\bar{t})B}\bar{x})^{\delta} D_x^{\alpha},$$

the result will then readily follow. We only consider the case $\bar{x} = 0$. Plugging equation (B.6) into the definition of $L_{h,i}$ we obtain

$$\begin{aligned} L_{h,i}(t,T,x) = &\sum^{\bullet}_{(\alpha^1,\delta^1)\in I_{i_1}} \cdots \sum^{\bullet}_{(\alpha^h,\delta^h)\in I_{i_h}} x^{\delta^1} D_x^{\alpha^1}\left(x^{\delta^2} D_x^{\alpha^2}\left(\cdots\left(x^{\delta^h} D_x^{\alpha^h}\right)\right)\right) \times \\ &\times \int_t^T \cdots \int_{s_{h-1}}^T \prod_{j=1}^h (s_j - t)^{\frac{|\alpha^j|_B - |\delta^j|_B + i_j - 2}{2}} ds_1 \cdots ds_h. \end{aligned}$$

Now, setting $\alpha = \alpha^1 + \cdots + \alpha^h$, $\delta = \delta^1 + \cdots + \delta^h$ and recalling that $i_1 + \cdots + i_h = n$, the integral above can be easily computed to be equal to

$$(T-t)^{\frac{|\alpha|_B - |\delta|_B + n}{2}},$$

times a constant. The statement follows applying Leibniz rule and noticing that $(\alpha, \delta) \in J_n$ if $(\alpha^j, \delta^j) \in I_{i_j}$ for $j = 1, \dots, h$. $\square$

Proof of Proposition 3.64. By (3.20)-(B.7), we get

$$D_x^{\beta} u_n^{(\bar{z})}(t,x) = D_x^{\beta} \sum^{\bullet}_{(\alpha,\delta)\in J_n} (T-t)^{\frac{|\alpha|_B - |\delta|_B + n}{2}} \left(x - e^{(t-\bar{t})B}\bar{x}\right)^{\delta} D_x^{\alpha} u_0^{(\bar{z})}(t,x)$$

(by applying Leibniz rule and reordering the indexes of J_n in (B.8))

$$= \sum^{\bullet}_{\substack{(\alpha,\delta)\in J_n \\ \nu \le \min\{\beta,\delta\}}} (T-t)^{\frac{|\alpha|_B - |\delta|_B + n}{2}} \left(x - e^{(t-\bar{t})B}\bar{x}\right)^{\delta-\nu} D_x^{\alpha+\beta-\nu} u_0^{(\bar{z})}(t,x),$$

where $\nu \le \min\{\beta, \delta\}$ means that $\nu_i \le \min\{\beta_i, \delta_i\}$ for any $i = 1, \dots, d$. Now, by applying Proposition B.1 and the property

$$\left|y^{\delta}\right| = \prod_{i=1}^d |y_i|^{\delta_i} \le \prod_{i=1}^d [y]_B^{\sigma_i \delta_i} = [y]_B^{|\delta|_B}, \qquad y \in \mathbb{R}^d,$$

we obtain

$$\left|D_x^{\beta} u_n^{(\bar{z})}(t,x)\right| \le \sum^{\bullet}_{\substack{(\alpha,\delta)\in J_n \\ \nu \le \min\{\beta,\delta\}}} (T-t)^{\frac{-|\delta|_B + n + k - |\beta|_B + |\nu|_B}{2}} \left[x - e^{(t-\bar{t})B}\bar{x}\right]_B^{|\delta|_B - |\nu|_B}$$

$$= \sum_{0\le m\le n}^{\bullet} (T-t)^{\frac{-m+n+k-|\beta|_B}{2}} \left[x - e^{(t-\bar{t})B}\bar{x}\right]_B^m,$$

and the statement follows by the elementary inequality

$$a^m b^{n-m} \le a^n + b^n, \qquad a, b \in \mathbb{R}_{>0},\ 0 \le m \le n.$$

$\square$

Bibliography

Arena, G., A. O. Caruso, and A. Causa (2010). Taylor formula on step two Carnot groups. *Rev. Mat. Iberoam. 26*(1), 239–259.

Barucci, E., S. Polidoro, and V. Vespri (2001). Some results on partial differential equations and Asian options. *Math. Models Methods Appl. Sci. 11*(3), 475–497.

Bonfiglioli, A. (2009). Taylor formula for homogeneous groups and applications. *Math. Z. 262*(2), 255–279.

Bonfiglioli, A., E. Lanconelli, and F. Uguzzoni (2007). *Stratified Lie groups and potential theory for their sub-Laplacians*. Springer Monographs in Mathematics. Springer, Berlin.

Bossy, M., J.-F. Jabir, and D. Talay (2011). On conditional McKean Lagrangian stochastic models. *Probab. Theory Related Fields 151*(1-2), 319–351.

Bramanti, M., M. C. Cerutti, and M. Manfredini (1996). L^p estimates for some ultraparabolic operators with discontinuous coefficients. *J. Math. Anal. Appl. 200*(2), 332–354.

Brennan, M. J. and E. S. Schwartz (1985). Evaluating natural resource investments. *The Journal of Business 58*(2), 135–157.

Chandresekhar, S. (1943). Stochastic problems in physics and astronomy. *Rev. Modern Phys. 15*, 1–89.

Chung, S. L., M. Shackleton, and R. Wojakowski (2003). Efficient quadratic approximation of floating strike Asian option values. *Finance (AFFI) 24*(1), 49–62.

Cinti, C. and S. Polidoro (2009). Bounds on short cylinders and uniqueness in cauchy problem for degenerate kolmogorov equations. *Journal of Mathematical Analysis and Applications 359*(1), 135 – 145.

Corielli, F., P. Foschi, and A. Pascucci (2010). Parametrix approximation of diffusion transition densities. *SIAM J. Financial Math. 1*, 833–867.

Cox, J. C., J. E. Ingersoll Jr, and S. A. Ross (1985). A theory of the term structure of interest rates. *Econometrica: Journal of the Econometric Society*, 385–407.

Deelstra, G., I. Diallo, and M. Vanmaele (2010). Moment matching approximation of Asian basket option prices. *J. Comput. Appl. Math. 234*(4), 1006–1016.

Dewynne, J. N. and W. T. Shaw (2008). Differential equations and asymptotic solutions for arithmetic Asian options: 'Black-Scholes formulae' for Asian rate calls. *European J. Appl. Math. 19*(4), 353–391.

Di Francesco, M. and A. Pascucci (2005). On a class of degenerate parabolic equations of Kolmogorov type. *AMRX Appl. Math. Res. Express* (3), 77–116.

Di Francesco, M. and S. Polidoro (2006). Schauder estimates, Harnack inequality and Gaussian lower bound for Kolmogorov-type operators in non-divergence form. *Adv. Differential Equations 11*(11), 1261–1320.

Donati-Martin, C., R. Ghomrasni, and M. Yor (2001). On certain Markov processes attached to exponential functionals of Brownian motion; application to Asian options. *Rev. Mat. Iberoamericana 17*(1), 179–193.

Dufresne, D. (2000). Laguerre series for Asian and other options. *Math. Finance 10*(4), 407–428.

Dufresne, D. (2001a). The integral of geometric brownian motion. *Advances in Applied Probability 33*(1), 223–241.

Dufresne, D. (2001b). The integrated square-root process.

Evatt, G., P. Johnson, P. Duck, and S. Howell (2010, 11). The measurement and inclusion of a stochastic ore-grade uncertainty in mine valuations using pdes. *IAENG International Journal of Applied Mathematics 40*(4), 1–7.

Feller, W. (1951). Two singular diffusion problems. *Annals of Mathematics 54*(1), 173–182.

Folland, G. B. (1975). Subelliptic estimates and function spaces on nilpotent Lie groups. *Ark. Mat. 13*(2), 161–207.

Folland, G. B. and E. M. Stein (1982). *Hardy spaces on homogeneous groups*, Volume 28 of *Mathematical Notes*. Princeton University Press, Princeton, N.J.; University of Tokyo Press, Tokyo.

Forde, M. and A. Jacquier (2010). Robust approximations for pricing Asian options and volatility swaps under stochastic volatility. *Appl. Math. Finance 17*(3), 241–259.

Foschi, P., S. Pagliarani, and A. Pascucci (2013). Approximations for Asian options in local volatility models. *J. Comput. Appl. Math. 237*(1), 442–459.

Foschi, P. and A. Pascucci (2008). Path dependent volatility. *Decis. Econ. Finance 31*(1), 13–32.

Frentz, M., K. Nyström, A. Pascucci, and S. Polidoro (2010). Optimal regularity in the obstacle problem for Kolmogorov operators related to American Asian options. *Math. Ann. 347*(4), 805–838.

Fusai, G. and A. Tagliani (2002). An accurate valuation of Asian options using moments. *Int. J. Theor. Appl. Finance 5*(2), 147–169.

Gatheral, J., E. P. Hsu, P. Laurence, C. Ouyang, and T.-H. Wang (2012). Asymptotics of implied volatility in local volatility models. *Math. Finance 22*(4), 591–620.

Geman, H. and M. Yor (1992). Quelques relations entre processus de Bessel, options asiatiques et fonctions confluentes hypergéométriques. *C. R. Acad. Sci. Paris Sér. I Math. 314*(6), 471–474.

Gencev, T. G. (1963). On ultraparabolic equations. *Dokl. Akad. Nauk SSSR 151*, 265–268.

Gobet, E. and M. Miri (2014). Weak approximation of averaged diffusion processes. *Stochastic Process. Appl. 124*(1), 475–504.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Financ. Stud. 6*(2), 327–343.

Hobson, D. G. and L. C. G. Rogers (1998). Complete models with stochastic volatility. *Math. Finance 8*(1), 27–48.

Hörmander, L. (1967). Hypoelliptic second order differential equations. *Acta Math. 119*, 147–171.

Hsu, E. P. (1995). On the principle of not feeling the boundary for diffusion processes. *J. London Math. Soc. (2) 51*(2), 373–382.

Il′in, A. M. (1964). On a class of ultraparabolic equations. *Dokl. Akad. Nauk SSSR 159*, 1214–1217.

Ingersoll, J. E. (1987). *Theory of Financial Decision Making.* Oxford: Blackwell.

Janson, S. and J. Tysk (2006). Feynman-Kac formulas for Black-Scholes-type operators. *Bull. London Math. Soc. 38*(2), 269–282.

Karatzas, I. and S. E. Shreve (1991). *Brownian motion and stochastic calculus* (Second ed.), Volume 113 of *Graduate Texts in Mathematics.* New York: Springer-Verlag.

Kolmogorov, A. (1934). Zufllige Bewegungen. (Zur Theorie der Brownschen Bewegung.). *Ann. of Math., II. Ser. 35*, 116–117.

Kolmogorov, A. N. (1991). *Selected works of A. N. Kolmogorov. Vol. I*, Volume 25 of *Mathematics and its Applications (Soviet Series).* Kluwer Academic Publishers Group, Dordrecht. Mathematics and mechanics.

Kunitomo, N. and A. Takahashi (1992). Pricing average options. *Japan Financial Review 14*, 1–20 (in Japanese).

Kunze, M., L. Lorenzi, and A. Lunardi (2010). Nonautonomous Kolmogorov parabolic equations with unbounded coefficients. *Trans. Amer. Math. Soc. 362*(1), 169–198.

Lanconelli, E. and S. Polidoro (1994). On a class of hypoelliptic evolution operators. *Rend. Sem. Mat. Univ. Politec. Torino 52*(1), 29–63. Partial differential equations, II (Turin, 1993).

Linetsky, V. (2004). Spectral expansions for Asian (average price) options. *Oper. Res. 52*(6), 856–867.

Lorig, M., S. Pagliarani, and A. Pascucci (2015). Analytical expansions for parabolic equations. *SIAM J. Appl. Math. 75*(2), 468–491.

Lunardi, A. (1997). 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.

Manfredini, M. (1997). The Dirichlet problem for a class of ultraparabolic equations. *Adv. Differential Equations 2*(5), 831–866.

Menozzi, S. (2011). Parametrix techniques and martingale problems for some degenerate Kolmogorov equations. *Electron. Commun. Probab. 16*, 234–250.

Novikov, A. A., T. G. Ling, and N. Kordzakhia (2014). Pricing of volume-weighted average options: analytical approximations and numerical results. In *Inspired by finance*, pp. 461–474. Springer, Cham.

Nyström, K., A. Pascucci, and S. Polidoro (2010). Regularity near the initial state in the obstacle problem for a class of hypoelliptic ultraparabolic operators. *J. Differential Equations 249*(8), 2044–2060.

Pagliarani, S. and A. Pascucci (2014). Asymptotic expansions for degenerate parabolic equations. *Comptes Rendus Mathematique 352*(12), 1011–1016.

Pagliarani, S. and A. Pascucci (2017). The exact Taylor formula of the implied volatility. *Finance Stoch. 21*(3), 661–718.

Pagliarani, S., A. Pascucci, and M. Pignotti (2016). Intrinsic Taylor formula for Kolmogorov-type homogeneous groups. *J. Math. Anal. Appl. 435*(2), 1054–1087.

Pagliarani, S., A. Pascucci, and M. Pignotti (2017). Intrinsic expansions for averaged diffusion processes. *Stochastic Processes and their Applications 127*(8), 2560 – 2585.

Pagliarani, S. and M. Pignotti (2017, July). Intrinsic Taylor formula for non-homogeneous Kolmogorov-type Lie groups. *ArXiv e-prints*.

Pascucci, A. (2003). Hölder regularity for a Kolmogorov equation. *Trans. Amer. Math. Soc. 355*(3), 901–924.

Pascucci, A. (2008). Free boundary and optimal stopping problems for American Asian options. *Finance Stoch. 12*(1), 21–41.

Pascucci, A. (2011). *PDE and martingale methods in option pricing.* Bocconi&Springer Series. New York: Springer-Verlag.

Polidoro, S. (1994). On a class of ultraparabolic operators of Kolmogorov-Fokker-Planck type. *Matematiche (Catania) 49*(1), 53–105 (1995).

Polidoro, S. and M. A. Ragusa (1998). Sobolev-Morrey spaces related to an ultraparabolic equation. *Manuscripta Math. 96*(3), 371–392.

Priola, E. (2009). Global Schauder estimates for a class of degenerate Kolmogorov equations. *Studia Math. 194*(2), 117–153.

Rogers, L. and Z. Shi (1995). The value of an Asian option. *J. Appl. Probab. 32*(4), 1077–1088.

Safonov, M. (1998). Estimates near the boundary for solutions of second order parabolic equations. In *Proceedings of the International Congress of Mathematicians, Vol. I (Berlin, 1998)*, Number Extra Vol. I, pp. 637–647 (electronic).

Shaw, W. T. (2003). Pricing Asian options by contour integration, including asymptotic methods for low volatility. *Working paper,* *http://www.mth.kcl.ac.uk/~shaww/web_page/papers/Mathematicafin.htm*.

Sonin, I. M. (1967). A class of degenerate diffusion processes. *Teor. Verojatnost. i Primenen 12*, 540–547.

Tsao, C. Y., C. C. Chang, and C. G. Lin (2003). Analytic approximation formulae for pricing forward-starting Asian options. *Journal of Futures Markets 23*(5), 487–516.

Varopoulos, N. T., L. Saloff-Coste, and T. Coulhon (1992). *Analysis and geometry on groups*, Volume 100 of *Cambridge Tracts in Mathematics*. Cambridge University Press, Cambridge.

Vecer, J. (2001). A new pde approach for pricing arithmetic average Asian options. *J. Comput. Finance*, 105–113.

Weber, M. (1951). The fundamental solution of a degenerate partial differential equation of parabolic type. *Trans. Amer. Math. Soc. 71*, 24–37.

Yor, M. (1992). On some exponential functionals of Brownian motion. *Adv. in Appl. Probab. 24*(3), 509–531.