Lecture Notes of
Seminario Interdisciplinare di Matematica
Vol. 6(2007), pp. 171–193.

On semigroups associated with degenerate elliptic operators with unbounded coefficients in $\mathbb{R}^N$

Luca Lorenzi

Abstract[1]. In this note, partly based on the communication held in the workshop *"Meeting on subelliptic pde's and applications to geometry and finance"*, we present a method developed in [16, 17] to determine a classical solution u to the Cauchy problem associated with a class of degenerate elliptic operators $\mathcal{A}$ in $\mathbb{R}^N$, as well as uniform and pointwise estimates for the spatial derivatives of the function u. As a byproduct, this allows to associate a semigroup of bounded operators $\{T(t)\}$ with $\mathcal{A}$. In particular, the uniform estimates allow to prove Schauder estimates for the (distributional) solution to the elliptic equation $\lambda u - \mathcal{A}u = f$ and the nonhomogeneous Cauchy problem associated with $\mathcal{A}$. Finally, under more restrictive assumptions on the coefficients of the operator $\mathcal{A}$, we show that one can associate an invariant measure with $\{T(t)\}$ and we prove some basic properties of the extension of $\{T(t)\}$ to the L^p-spaces related to this measure.

1. Introduction

Starting from the pioneering papers by Aronson and Besala ([1, 2]), Krzyżański ([11, 12]), Krzyżański and Szybiak ([9, 10]) the interest in elliptic operators with unbounded coefficients in $\mathbb{R}^N$ and its unbounded domains has grown considerably, due to their applications to various fields of sciences such as mathematical finance.

Nondegenerate elliptic operators with unbounded coefficients have been studied using several different approaches, with ideas and methods from partial differential equations, Dirichlet forms, stochastic processes, stochastic differential equations, martingale theory. In fact, nowadays the theory of these operators is rather complete and satisfactory. For most of the results in the literature, we refer the reader to the monograph [4].

It is well known that, under very general assumptions on the coefficients of the nondegenerate elliptic operator

$$\mathcal{A} = \sum_{i,j=1}^{N} q_{ij}(x)D_{ij} + \sum_{j=1}^{N} b_j(x)D_j + c(x) \quad , \quad x \in \mathbb{R}^N , \tag{1.1}$$

[1]Author's address: L. Lorenzi, Università degli Studi di Parma, Dipartimento di Matematica, Viale G.P. Usberti 53/A, I-43100 Parma, Italy; e-mail: `luca.lorenzi@unipr.it` .
URL: `www.unipr.it/∼lorluc99/index.html` .
Keywords. Degenerate elliptic operators with unbounded coefficients in $\mathbb{R}^N$, uniform and pointwise estimates, distributional solutions to elliptic and parabolic problems, Schauder estimates, invariant measures.
AMS Subject Classification. 35K65, 35J70, 35B65, 35K15.

the initial value problem $u(0,\cdot) = f$ for the parabolic equation $D_t u = \mathcal{A}u$ admits at least a classical solution, in the sense that

(i) u is continuously differentiable in $]0,+\infty[\times\mathbb{R}^N$, once with respect to time and twice with respect to the spatial variables,
(ii) it is bounded and continuous in $[0,T]\times\mathbb{R}^N$ for any $T>0$;
(iii) it satisfies the above differential equation pointwise in $]0,+\infty[\times\mathbb{R}^N$ and $u(0,\cdot) = f$.

It is worth stressing that such a problem needs not to admit a unique classical solution. This is due to the lack of a maximum principle for elliptic operators with unbounded coefficients when no additional (algebraic) conditions are prescribed.

In any case one can associate a semigroup $\{T(t)\}$ of bounded linear operators in $C_b(\mathbb{R}^N)$ with the operator $\mathcal{A}$. For any positive f, $T(\cdot)f$ is the minimal positive solution to the above Cauchy problem, in the sense that, if $f \geq 0$, then any other solution v is greater than $T(\cdot)f$.

In particular, uniform estimates for the spatial derivatives of the function $T(t)f$, when f belongs to several spaces of Hölder continuous functions defined in $\mathbb{R}^N$, are available under more restrictive assumptions on the smoothness and the growth at infinity of the coefficients of the operator $\mathcal{A}$.

Schauder estimates for the elliptic equation

$$\lambda v(x) - \mathcal{A}v(x) = h(x) \quad , \quad x \in \mathbb{R}^N,\ \lambda > 0\ , \tag{1.2}$$

as well as for the Cauchy problem

$$\begin{cases} D_t u(t,x) = \mathcal{A}u(t,x) + g(t,x) & , \quad (t,x) \in]0,T[\times\mathbb{R}^N\ , \\ u(0,x) = f(x) & , \quad x \in \mathbb{R}^N\ , \end{cases} \tag{1.3}$$

when the data f,g,h belong to suitable classes of Hölder continuous functions, are then obtained as a byproduct of the uniform estimates.

As in the classical case of bounded coefficients, the solutions to (1.2) and (1.3) are given, respectively by the (pointwise) Laplace transform of the semigroup, i.e.,

$$v(x) = \int_0^{+\infty} e^{-\lambda t}(T(t)f)(x)\,dt \quad , \quad x \in \mathbb{R}^N\ , \tag{1.4}$$

and the (pointwise) variation of constants formula

$$u(t,x) = (T(t)f)(x) + \int_0^t (T(t-s)g(s,\cdot))(x)\,ds \quad , \quad t \in [0,T],\ x \in \mathbb{R}^N\ . \tag{1.5}$$

In the case when $\mathcal{A}$ is a degenerate elliptic operator with unbounded coefficients, global uniform estimates for the spatial derivatives of the associated semigroup as well as Schauder estimates for problems (1.2) and (1.3), seem, to the best of our knowledge, to be available only in a few situations. The simplest one occurs when $\mathcal{A}$ is the degenerate Ornstein-Uhlenbeck operator, i.e., when the operator $\mathcal{A}$ is given by

$$\mathcal{A} = \sum_{i,j=1}^N q_{ij}D_{ij} + \sum_{i,j=1}^N b_{ij}x_jD_i \quad , \quad x \in \mathbb{R}^N\ , \tag{1.6}$$

where $Q = (q_{ij})$ and $B = (b_{ij})$ are suitable constant square matrices, Q being semi-positive definite, such that the hypoellitícity condition

$$\det\left(\int_0^t e^{sB} Q e^{sB^*}\, ds\right) > 0\ ,$$

is satisfied at any $t > 0$. As it is shown in [13], the previous condition is equivalent to both the Kalman rank condition and the hypoellipticity of $\mathcal{A}$ in the sense of Hörmander.

Up to an orthogonal change of variables, the operator $\mathcal{A}$ in (1.6) may be reduced to the following one:

$$\mathcal{A} = \sum_{i,j=1}^{p_0} q_{ij} D_{ij} + \sum_{i,j=1}^{N} b_{ij} x_j D_i \quad , \quad x \in \mathbb{R}^N\ , \tag{1.7}$$

for some $p_0 < N$ and some new matrices $Q = (q_{ij}) \in L(\mathbb{R}^{p_0})$ and $B \in L(\mathbb{R}^N)$, with Q strictly positive definite and

$$B = \begin{pmatrix} \star & \star & \cdots & \cdots & \star \\ B_1 & \star & \cdots & \cdots & \star \\ 0 & B_2 & \star & \cdots & \star \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & B_r & \star \end{pmatrix}\ , \tag{1.8}$$

where "$\star$" denotes suitable submatrices and B_j are $p_{j-1} \times p_j$ matrices with full rank p_j, $j = 1, \cdots, r$, p_j $(j = 1, \cdots, r)$ being positive integers such that $p_0 \geq p_1 \geq \cdots \geq p_r$ and $p_0 + \cdots + p_r = N$, see e.g., [13, 9]. One of the main features of the Ornstein-Uhlenbeck operator is that an explicit representation formula for the associated semigroup is available. Of course, this simplifies the analysis of the properties of the semigroup. Direct computations on the representation formula for $T(t)f$, allow Lunardi ([20]) to estimate the behaviour of the spatial derivatives of the functions $T(t)f$ when t approaches 0 and f belongs to several spaces of Hölder continuous functions. In fact, such behaviour heavily depends on the variable with respect to one differentiates.

The anisotropic behaviour of the semigroup with respect to the directions where differentiation is performed, suggests that suitable spaces where to look for Schauder estimates for both the problems (1.2) and (1.3) are anisotropic spaces of Hölder continuous functions, modelled on the degeneracy of the operator $\mathcal{A}$.

Denoting, roughly speaking, by $\mathcal{C}^\theta$ ($\theta \in]0,1[$), such anisotropic spaces, the results in [20] show that

$$\|T(t)f\|_{\mathcal{C}^\theta} \leq C t^{-(\theta-\alpha)/2} \|f\|_{\mathcal{C}^\alpha} \quad , \quad t \in]0, T]\ , \tag{1.9}$$

for any $\alpha, \theta \in [0, +\infty[$, with $\alpha \leq \theta$, any $T > 0$ and some $C = C(T)$. Estimate (1.9) represents what one can expect when $\mathcal{A}$ is a nondegenerate elliptic operator with bounded coefficients, and $\mathcal{C}^\theta$ and $\mathcal{C}^\alpha$ are replaced with the usual Hölder spaces. The estimates in (1.9) are the keystone to prove, via formulas (1.4) and (1.5), the following results:

1.1. **Elliptic equation** (1.2)**.** Fix $\lambda > 0$ and $\theta \in]0,1[$. Then, for any $h \in \mathcal{C}^\theta$, the function v defined by formula (1.4), is the unique *distributional* solution to problem

(1.2). Moreover, there exists a positive constant C, *independent* of h, such that

$$\|v\|_{\mathcal{C}^{2+\theta}} \le C\|h\|_{\mathcal{C}^{\theta}} .$$

1.2. **Cauchy problem** (1.3)**.** Fix $T > 0$ and $\theta \in]0, 1[$. Then, for any $f \in \mathcal{C}^{2+\theta}$ and any continuous function $g : [0, T] \times \mathbb{R}^N \to \mathbb{R}$ such that

$$[[g]]_\theta := \sup_{t\in[0,T]} \|g(t, \cdot)\|_{\mathcal{C}^{\theta}} < +\infty ,$$

the function u in (1.5) is the unique *distributional* solution to problem (1.3). Moreover, there exists a positive constant C, *independent* of f and g, such that

$$\sup_{t\in[0,T]} \|u(t, \cdot)\|_{\mathcal{C}^{2+\theta}} \le C \left(\|f\|_{\mathcal{C}^{2+\theta}} + [[g]]_\theta\right) .$$

The anisotropic behavior of the spatial derivatives of the semigroup reflects in the fact that only the spatial derivatives up to the second-order with respect to the first p_0 variables exist in the classical sense, whereas all the other spatial derivatives, as well as the time derivative, are meant in the distributional sense.

Recently, in [6, 16, 17, 24, 26], more general classes of degenerate elliptic operators, modelled on the Ornstein-Uhlenbeck operator, have been studied with both probabilistic ([24]) and analytical methods ([6, 16, 17]) with the aim of proving estimates similar to (1.9) and, consequently, to prove Schauder estimates for both the solutions of (1.2) and (1.3).

In [16, 17] operators of the type (1.7) have been considered assuming that $r = 1$, $p_0 \ge N/2$, and allowing the diffusion coefficients of $\mathcal{A}$ to depend on the spatial variables and (possibly) to blow up at infinity with a certain rate. The forthcoming paper [7] is devoting to generalizing the results in [16, 17], dropping out the rather restrictive condition $r = 1$.

On the other hand, [6, 24, 26] deal with operators of type (1.7) with a perturbation in the drift part.

For other results concerned with existence and uniqueness of solutions to Cauchy problems associated with elliptic operators and Hölder estimates for such solutions, we refer the reader e.g., to [14, 15, 22, 23].

In the first part of this paper, we briefly present the technique developed in [16, 17] and, then, used in [7, 26], with the aim of pointing out the main techniques and skipping, as much as possible, the technicalities, since we deem that the arguments in the quoted papers may be applied to situations more general than those in [7, 26]. Then, in Sections 4 and 5 we prove some pointwise estimates in the spirit of [4, Chapter 6] and we study some basic properties of the extension of $\{T(t)\}$ to L^p-spaces related to the invariant measure μ associated with the semigroup, giving also a sufficient condition for its existence.

2. Function spaces, Assumptions and a Maximum principle

2.1. **Function spaces.** Here, we introduce the function spaces we deal with in this note.

2.1.1. *Isotropic spaces.* For any $k \in [0, +\infty]$, $C_b^k(\mathbb{R}^N)$ denotes the usual set of functions differentiable up to the $[k]$-th order and with all the derivatives bounded in $\mathbb{R}^N$, those of maximum order being Hölder continuous with exponent $k - [k]$. Here, $[k]$ denotes the integer part of k. All these spaces are endowed with the

Euclidean norm. When the index "b" is replaced by "c" it means that we are dealing with compactly supported functions.

For any set $F \subset [0,+\infty[\times\mathbb{R}^N$, $C^{1,2}(F)$ denotes the space of the u's which are once continuously differentiable with respect to time and twice continuously differentiable with respect to the space variables in F; $C^{1+\alpha/2,2+\alpha}(F)$ ($\alpha \in]0,1[$) is its subset of all functions u with Hölder continuous derivatives of order α in F, with respect to the parabolic distance $d((t,x),(s,y)) = (|t-s| + |x-y|^2)^{1/2}$.

2.1.2. *Anisotropic spaces.* For any $\theta \in]0,3]$, $\mathcal{C}^\theta$ is the subset of $C_b(\mathbb{R}^N)$ of functions f such that

$$\|f\|_{\mathcal{C}^\theta} = \sup_{y\in\mathbb{R}^{N-p_0}} \|f(\cdot,y)\|_{C_b^\theta(\mathbb{R}^{p_0})} + \sup_{x\in\mathbb{R}^{p_0}} [f(x,\cdot)]_{C_b^{\theta/3}(\mathbb{R}^{N-p_0})} < +\infty\ .$$

2.2. **Hypotheses.** Here, we state the assumptions on the coefficients of the operator $\mathcal{A}$ defined by

$$\mathcal{A} = \sum_{i,j=1}^{p_0} q_{ij}(x)D_{ij} + \sum_{i,j=1}^{N} b_{ij}x_jD_i \quad , \quad x\in\mathbb{R}^N\ . \tag{2.1}$$

Hypotheses 2.1.

(i) $q_{ij} = q_{ji} \in C^{3+\delta}(\mathbb{R}^N)$ $(i,j = 1,\cdots,p_0)$ *for some* $\delta \in]0,1[$*, and there exist a positive constant* C *and a function* ν*, with* $\inf_{x\in\mathbb{R}^N}\nu(x) > 0$*, such that*

$$\sum_{i,j=1}^{p_0} q_{ij}(x)\xi_i\xi_j \geq \nu(x)|\xi|^2 \quad , \quad x,\xi\in\mathbb{R}^N\ ,$$

and

$$|D^\alpha q_{ij}(x)| \leq C|x|^{(1-|\alpha|)^+}\sqrt{\nu(x)} \quad , \quad x\in\mathbb{R}^N,\ i,j=1,\cdots,p_0,\ |\alpha|\leq 3\ ,$$

s^+ *denoting the positive part of* $s\in\mathbb{R}$*;*

(ii) *the matrix* B_1 (*see* (1.8)) *has full rank* $N-p_0$.

Remark 2.2. Note that Hypotheses 2.1 imply, in particular, that the diffusion coefficients q_{ij} $(i,j = 1,\cdots,p_0)$ may grow *at most* quadratically as $|x|$ tends to $+\infty$.

2.3. **Maximum principles.** As it is well known, in the classical case of bounded coefficients, the uniqueness of the classical solutions to problem (1.2) and (1.3) are implied by the maximum principle. In the case when the coefficients are unbounded, the maximum principle fails, in general, to hold without any additional assumption on the coefficients. However, it may be proved assuming the existence of a Lyapunov function, i.e. a smooth function $\varphi \in C^2(\mathbb{R}^N)$ blowing up at infinity such that $\lambda_0\varphi - \mathcal{A}\varphi$ is bounded from above in $\mathbb{R}^N$ for some $\lambda_0 \in [0,+\infty[$. More precisely:

Proposition 2.3 (Generalized maximum principle)**.** *Let* $\mathcal{A}$ *be a* (*possibly degenerate*) *elliptic operator with unbounded and continuous coefficients in* $\mathbb{R}^N$*, which admits a Lyapunov function* φ*. Further, let* $u \in C^{1,2}(]0,T[\times\mathbb{R}^N)\cap C([0,T]\times\mathbb{R}^N)$ *be bounded* (*with respect to the sup-norm*) *and satisfy* $D_tu \leq \mathcal{A}u$ *in* $]0,T[\times\mathbb{R}^N$*, for some* $T>0$*. Then,* $\sup_{[0,T]\times\mathbb{R}^N} u \leq \sup_{\mathbb{R}^N} u(0,\cdot)$*. In particular, the Cauchy problem* (1.3) (*with* $g\equiv 0$) *admits, at most, a unique classical solution* u*. Moreover,* $\sup_{t\in[0,T]}\|u(t,\cdot)\|_\infty \leq \|f\|_\infty$.

Proof. See e.g., [16, Proposition 2.7].

□

Remark 2.4. Notice that the proof of the previous proposition may be adapted to show that problem (1.2) resp. (1.3)), with $\mathcal{A}$ being given by (2.1), admits at most a unique bounded and continuous distributional solution u such that the derivatives $D_j u$ and $D_{ij} u$ $(i, j = 1, \cdots, p_0)$ exist in the classical sense in $\mathbb{R}^N$ (resp. in $]0, T[\times\mathbb{R}^N$). For more details, see [7].

Remark 2.5. In view of Remark 2.2 it is an easy task to check that the function $\varphi(x) = 1 + |x|^2$, for any $x \in \mathbb{R}^N$, is a Lyapunov function for the operator $\mathcal{A}$ in (2.1).

3. Construction of the semigroup associated with the operator $\mathcal{A}$ in (2.1) and uniform estimates for the function $T(t)f$

In this section we provide a method to construct a (unique) classical solution of the homogeneous Cauchy problem associated with the operator $\mathcal{A}$. As it has been already remarked in the Introduction, this will allow us to associate a semigroup $\{T(t)\}$ of bounded operators with $\mathcal{A}$. At the same time our technique provides us with uniform estimates for the function $T(t)f$ when f belongs to several spaces of (Hölder-) continuous functions.

Here, we list briefly the main steps in this direction:

(i) we approximate the operator $\mathcal{A}$ with a family of uniformly elliptic operators $\mathcal{A}_\varepsilon$ $(\varepsilon \in]0, 1[)$ defined by

$$\mathcal{A}_\varepsilon = \mathcal{A} + \varepsilon \sum_{i=p_0+1}^{N} D_{ii} := \mathrm{Tr}(Q^{(\varepsilon)}(x)D^2) + \langle Bx, D\rangle \quad , \quad x \in \mathbb{R}^N , \tag{3.1}$$

and we prove uniform estimates for the spatial derivatives of the function $T_\varepsilon(t)f$, with constants independent of ε, $\{T_\varepsilon(t)\}$ being the semigroup associated with the operator $\mathcal{A}_\varepsilon$;

(ii) we use such estimates in place of the usual interior Schauder estimates holding in the nondegenerate case to prove the existence of a solution to problem (1.3) (with $g \equiv 0$), via a compactness argument (the uniqueness of such a solution is a consequence of Proposition 2.3 and Remark 2.5);

(iii) since "$T(t)f = \lim_{\varepsilon\to 0^+} T_\varepsilon(t)f$" for any $t > 0$, taking the limit in the estimates obtained in Step (ii), we get uniform estimates for the derivatives of the function $T(t)f$.

3.1. Step (i).

It is well known since the sixties that the Cauchy problem (1.3) (with $g \equiv 0$) associated with uniformly elliptic operators $\mathcal{A}$ of the type (1.1), admits at least a classical solution u for any $f \in C_b(\mathbb{R}^N)$, when the coefficients of $\mathcal{A}$ are locally Hölder continuous in $\mathbb{R}^N$ with exponent $\alpha \in]0, 1[$ and the potential c is bounded from above. (One of) its solution may be obtained as the "limit" (as R tends to $+\infty$) of the classical solutions u_R to the Dirichlet Cauchy problems in the ball $B(0, R)$, with initial condition $u_R(0, \cdot) = f\eta_R$, where η_R is any $C_c^\infty(\mathbb{R}^N)$ smooth function such that $\eta_R \equiv 1$ in $B(0, R/2)$ and $\eta_R \equiv 0$ outside $B(0, R)$. If $\mathcal{A}$ is given by (3.1), Proposition 2.3 applies, taking the function $\varphi(x) = 1 + |x|^2$, for any $x \in \mathbb{R}^N$, as a Lyapunov function. In particular, Proposition 2.3 immediately implies that $\{T_\varepsilon(t)\}$ is a positive semigroup of contractions. We refer the reader to [4, 21] for further details.

The uniform estimates for the function $T_\varepsilon(t)f$. This is the main step to prove the existence of a classical solution to the Cauchy problem (1.3) as well as the Schauder estimates in 1.1 and 1.2. In fact, one can show the following.

3.1.1. **Estimates in isotropic spaces.**

$$\begin{cases} (a) & \|D_i T_\varepsilon(t)f\|_\infty \le Ct^{-(1/2)-\kappa(i)}\,\|f\|_\infty\,, \\ (b) & \|D_{ij} T_\varepsilon(t)f\|_\infty \le Ct^{-1-\kappa(i)-\kappa(j)}\,\|f\|_\infty\,, \\ (c) & \|D_{ijh} T_\varepsilon(t)f\|_\infty \le Ct^{-(3/2)-\kappa(i)-\kappa(j)-\kappa(h)}\,\|f\|_\infty\,, \end{cases} \tag{3.2}$$

for any $t \in]0,T]$, some $T > 0$ and $C = C(T)$, independent of ε. Here, $\kappa(s) = \chi_{]p_0,+\infty[}(s)$. As in the nondegenerate case, the smoothness of f reflects in the behavior of the derivatives of $T_\varepsilon(t)f$ near 0. More precisely,

$$\begin{cases} (a) & \|D_{ij} T_\varepsilon(t)f\|_\infty \le Ct^{-(1/2)-\kappa(i)}\,\|f\|_{C^1_b(\mathbb{R}^N)}\,, \\ (b) & \|D_{ijh} T_\varepsilon(t)f\|_\infty \le Ct^{-1-\kappa(i)-\kappa(j)}\,\|f\|_{C^1_b(\mathbb{R}^N)}\,, \\ (c) & \|D_{ijh} T_\varepsilon(t)f\|_\infty \le Ct^{-(1/2)-\kappa(i)}\,\|f\|_{C^2_b(\mathbb{R}^N)}\,, \\ (d) & \|D^k T_\varepsilon(t)f\|_\infty \le C\|f\|_{C^k_b(\mathbb{R}^N)}\,, \quad k = 1,2,3\,, \end{cases} \tag{3.3}$$

for any $t \in]0,T]$, where T and C are as above, and $D^k T_\varepsilon(t)f$ denotes the vector of all k-th order spatial derivatives of the function $T_\varepsilon(t)f$.

3.1.2. **Estimates in the anisotropic spaces $\mathcal{C}^\theta$.** Combining the estimates in (3.2), it is immediate to check that there exists $C = C(T)$ such that

$$\|T_\varepsilon(t)f\|_{\mathcal{C}^3} \le Ct^{-(3/2)}\,\|f\|_\infty\,, \tag{3.4}$$

for any $t \in]0,T]$ and some positive T. So, the main step in order to obtain estimates of the type (1.9) for $T_\varepsilon(t)f$ in the anisotropic Hölder spaces $\mathcal{C}^\theta$ consists in showing that

$$\|T_\varepsilon(t)f\|_{\mathcal{C}^3} \le C\|f\|_{\mathcal{C}^3}\,, \tag{3.5}$$

for any $t \in]0,T]$, where C and T are as above. Indeed, since

$$(C_b(\mathbb{R}^N), C^{3,1}_b(\mathbb{R}^N))_{\theta,\infty} = \mathcal{C}^{3\theta} \quad , \quad (\mathcal{C}^\alpha, \mathcal{C}^\beta)_{\theta,\infty} = \mathcal{C}^{\alpha+\theta(\beta-\alpha)}\,,$$

for any $\theta \in]0,1[\setminus\{1/3, 2/3\}$ and any $\alpha, \beta \in]0,3[\setminus\{1,2\}$ with $\alpha \le \beta$ and $\alpha+\theta(\beta-\alpha) \notin \mathbb{N}$, the estimate (3.4) and (3.6) allow to show, via a classical interpolation argument, that

$$\|T_\varepsilon(t)f\|_{\mathcal{C}^{\gamma_2}} \le Ct^{-(\gamma_2-\gamma_1)/2}\,\|f\|_{\mathcal{C}^{\gamma_1}}\,, \tag{3.6}$$

for any $t \in]0,T]$, any $T > 0$, any $\gamma_1, \gamma_2 \in]0,3[\setminus\{1,2\}$ with $\gamma_1 \le \gamma_2$ and some positive constant $C = C(T, \gamma_1, \gamma_2)$.

Remark 3.1. Note that it suffices to prove the estimates in (3.2), (3.3) and (3.6) in a small neighborhood of $t = 0$. Indeed, the semigroup property allows then to extend them to all the positive t's.

As it has been already claimed in the Introduction, the estimate (3.6), written for the semigroup $\{T(t)\}$, will be the keystone to prove the Schauder estimates in (1.1) and (1.2) via an interpolation technique from [19].

3.1.3. **A method to obtain the uniform estimates** (3.2)**,** (3.3)**,** (3.5)**.** Here, we explain the main ideas and techniques that can be used to prove the quoted uniform estimates.

If $\mathcal{A}$ is a nondegenerate elliptic operator (admitting a Lyapunov function) and $\{T(t)\}$ is the associated semigroup, one can expect that $\|D^kT(t)f\|_\infty$ behaves like $t^{-k/2}$ when t approaches 0 (and $k = 1, 2, 3$). The classical Bernstein method ([3]) suggests to introduce the function $v_k : [0, T] \times \mathbb{R}^N \to \mathbb{R}$ defined by

$$v_k(t, x) = \alpha_0|(T(t)f)(x)|^2 + \cdots + \alpha_k t^k|(D^kT(t)f)(x)|^2 \quad , \quad t \in]0, T],\ x \in \mathbb{R}^N ,$$

α_j beng positive parameters, and look for the problem fulfilled by v_k. After a few computations and technicalities, one obtains that v_k is a bounded classical solution to the Cauchy problem

$$\begin{cases} D_t v_k(t, x) = \mathcal{A}v_k(t, x) + \psi_k(t, x) & , \quad (t, x) \in]0, T[\times\mathbb{R}^N , \\ v_k(0, x) = |f(x)|^2 & , \quad x \in \mathbb{R}^N , \end{cases}$$

where ψ_k is a suitable continuous function. A proper choice of the parameters α_j $(j = 0, \cdots, k)$, allows to show that $\psi_k \le 0$ in $]0, T[\times\mathbb{R}^N$. Hence, Proposition 2.3 applies and gives $v_k \le f^2$ in $[0, T] \times \mathbb{R}^N$, which provides us with the desired estimates for the derivatives of $T(t)f$ up to the k-th order.

Having this situation in mind, one might think to apply *verbatim* the same technique to the degenerate case. In the simplest case when

$$(\mathcal{A}\varphi)(x, y) = D_{xx}\varphi(x, y) + xD_y\varphi(x, y) \quad , \quad (x, y) \in \mathbb{R}^2 ,$$

one should show that the function

$$v(t, \cdot) = |T_\varepsilon(t)f|^2 + \alpha t|D_xT_\varepsilon(t)f|^2 + \beta t^3|D_yT_\varepsilon(t)f|^2 \quad , \quad t \in]0, T] , \tag{3.7}$$

satisfies the first estimate in (3.2) for a suitable choice of the parameters α and β. A straightforward computation shows that, in this situation,

$$\psi(t, \cdot) = -(2 - \alpha)|D_xT_\varepsilon(t)f|^2 - (2\varepsilon - 3\beta\, t^2)|D_yT_\varepsilon(t)f|^2+$$
$$+2\alpha\, t\, D_xT_\varepsilon(t)f \cdot D_yT_\varepsilon(t)f - 2\alpha\, t\, |D_{xx}T_\varepsilon(t)f|^2 - 2\beta\, t^3|D_{xy}T_\varepsilon(t)f|^2-$$
$$-2\varepsilon\, \alpha\, t\, |D_{xy}T_\varepsilon(t)f|^2 - 2\varepsilon\, \beta\, t^3|D_{yy}T_\varepsilon(t)f|^2 .$$

Consequently, there is no hope to fix the parameters, independently of ε, such that the quadratic form in the gradient of $T_\varepsilon(t)f$ (and, consequently, ψ) is nonpositive in $]0, T] \times \mathbb{R}^N$ for some $T > 0$.

This suggests that in the construction of the function v one has to take into account the structure of the operator $\mathcal{A}$. In fact, the function v in (3.7) should be replaced with the function

$$v_1(t, \cdot) = \alpha^3|T_\varepsilon(t)f|^2 + \langle\mathcal{F}_1(t)DT_\varepsilon(t)f, DT_\varepsilon(t)f\rangle \quad , \quad t \in]0, T] ,$$

where

$$\mathcal{F}_1(t) = \begin{pmatrix} \alpha\, t\, I_{p_0} & t^2F_1^* \\ t^2F_1 & \beta\, t^3 I_{N-p_0} \end{pmatrix} , \tag{3.8}$$

F_1 being a matrix to be chosen later on as well as the parameters α and β. Finally, I_r denotes the $r \times r$ identity matrix. A straightforward computation shows that

$\psi = \psi_1 + \psi_2 + \psi_3$, where

$$\psi_1(t,\cdot) = -2\alpha^3\langle Q^{(\varepsilon)} DT_\varepsilon(t)f, DT_\varepsilon(t)f\rangle - 2\sum_{i,j=1}^{N} q_{ij}^{(\varepsilon)}\langle \mathcal{F}_1 DD_iT_\varepsilon(t)f, DD_jT_\varepsilon(t)f\rangle\ ,$$

$$\psi_2(t,\cdot) = \langle (B\mathcal{F}_1(t) + \mathcal{F}_1(t)B^*)DT_\varepsilon(t)f, DT_\varepsilon(t)f\rangle\ ,$$

$$\psi_3(t,\cdot) = \langle \mathcal{F}_1'(t)DT_\varepsilon(t)f, DT_\varepsilon(t)f\rangle + 2\sum_{i,j=1}^{p_0} D_{ij}T_\varepsilon(t)f\langle \mathcal{F}_1(t)Dq_{ij}, DT_\varepsilon(t)f\rangle\ ,$$

for any $t \in]0,T]$. A long but straightforward computation shows that

$$\psi_1(t,\cdot) \le -2\alpha^3\nu|D_{\star,1}T_\varepsilon(t)f|^2 - 2\alpha\, t\,\nu\,\langle K_1D^2_{\star,1}T_\varepsilon(t)f, D^2_{\star,1}T_\varepsilon(t)f\rangle -$$
$$-2\nu\,\beta\, t^3|D^2_{\star,2}T_\varepsilon(t)f|^2 - 2\nu\, t^2\langle K_2D^2_{\star,1}T_\varepsilon(t)f, D^2_{\star,2}T_\varepsilon(t)f\rangle\ ,$$

for any $t \in]0,T]$, where K_1 is a suitable diagonal matrices whose minimum eigenvalue is 1, whereas the entries of the matrix K_2 depend linearly only on the entries of F_1. In particular, the two previous matrices are *independent* of α. Here, $D_{\star,1}T_\varepsilon(t)f$ contains all the first-order derivatives $D_jT_\varepsilon(t)f$ such that $j \le p_0$. Similarly, $D^2_{\star,1}T_\varepsilon(t)f$, $D^2_{\star,2}T_\varepsilon(t)f$, $D^2_{\star,3}T_\varepsilon(t)f$ denote the vectors of all the second-order derivatives $D_{ij}T_\varepsilon(t)f$ (with no repetitions) such that $i,j \le p_0$, respectively, $i \le p_0 < j$, respectively $i,j \ge p_0$, ordered lexicographically, i.e., if $D_{ij}T_\varepsilon(t)f$ and $D_{hk}T_\varepsilon(t)f$ belong to the same sub-block $D^2_{\star,l}T_\varepsilon(t)f$, we say that $D_{ij}T_\varepsilon(t)f$ precedes $D_{hk}T_\varepsilon(t)f$ if either $i < h$ or $i = j$ and $j < k$. Hence, using properly Young inequality, it follows that

$$\psi_1(t,\cdot) \le -2\alpha^3\nu\,|D_{\star,1}T_\varepsilon(t)f|^2 + (-2\alpha + \alpha^{1/2}\,\|K_2\|)t\,\nu\,|D^2_{\star,1}T_\varepsilon(t)f|^2 +$$
$$+(-2\beta + \alpha^{-1/2}\,\|K_2\|)\nu\, t^3|D^2_{\star,2}T_\varepsilon(t)f|^2\ , \tag{3.9}$$

for any $t \in]0,T]$. As it is immediately seen, if α and β are properly chosen, all the terms in the right-hand side of (3.9) are negative. Hence, they allow us to control part of the terms contained in ψ_2 and ψ_3. Note that in the right-hand side of (3.9) the term $D_{\star,2}T_\varepsilon(t)f$, consisting of all the derivatives $D_jT_\varepsilon(t)f$ with $j > p_0$, is missing, whereas it appears in the definition of the function ψ_3. We can recover the term $|D_{\star,2}T_\varepsilon(t)f|^2$, with a negative coefficient in front, from a part of ψ_2. Indeed,

$$\psi_2(t,\cdot) = 4\langle (B_1F_1 + F_1^*B_1^*)D_{\star,2}T_\varepsilon(t)f, D_{\star,2}T_\varepsilon(t)f\rangle + \cdots\ .$$

Since the matrix B_1 has full rank, we can take $F_1 = -B_1^*$ obtaining that

$$4\langle (B_1F_1 + F_1^*B_1)D_{\star,2}T_\varepsilon(t)f, D_{\star,2}T_\varepsilon(t)f\rangle \le -8\iota|D_{\star,2}T_\varepsilon(t)f|^2\ , \tag{3.10}$$

for any $t \in]0,T]$, where ι is the least positive eigenvalue of the matrix $B_1B_1^*$. The other terms in the definition of ψ_2 and in ψ_3 can be estimated, uniformly with respect to ε. In particular, choosing properly α, β and taking T sufficiently small, they can be controlled by the right-hand sides of (3.9) and (3.10), so that $\psi \le 0$ in $]0,T] \times \mathbb{R}^N$. This means that $v \le \|f\|^2_\infty$ in $]0,T] \times \mathbb{R}^N$. Taking a larger value of α, if necessary, we can assume that the matrix $\mathcal{F}_1(t)$ is not singular. With this choice of α, the condition $\langle \mathcal{F}_1(t)DT_\varepsilon(t)f, DT_\varepsilon(t)f\rangle \le \|f\|^2_\infty$ is equivalent to the first two estimates in (3.2).

Note that, in the nondegenerate case, the matrix $\mathcal{F}(t)$ reduces to tI_N, whereas in the degenerate case the terms off the main diagonal turn out to play a crucial role. Hence, what we have shown so far suggests that, to prove the remaining estimates in (3.2), the terms $t^{k/2}|D^kT_\varepsilon(t)f|^2$ $(k=2,3)$ have to be replaced with the quadratic forms

$$\langle \mathcal{F}_2(t)D^2_\star T_\varepsilon(t)f, D^2_\star T_\varepsilon(t)f\rangle \quad \text{and} \quad \langle \mathcal{F}_3(t)D^3_\star T_\varepsilon(t)f, D^3_\star T_\varepsilon(t)f\rangle\ ,$$

where

(a) $D^h_\star T_\varepsilon(t)$ denotes the vector of all h-th derivatives of $T_\varepsilon(t)f$ $(h=2,3)$ with no repetitions, ordered in sub-blocks as follows:

$$D^2_\star T_\varepsilon(t)f = (D^2_{\star,1}T_\varepsilon(t)f, D^2_{\star,2}T_\varepsilon(t)f, D^2_{\star,3}T_\varepsilon(t)f)\ ,$$

$$D^3_\star T_\varepsilon(t)f = (D^3_{\star,1}T_\varepsilon(t)f, D^3_{\star,2}T_\varepsilon(t)f, D^3_{\star,3}T_\varepsilon(t)f, D^3_{\star,4}T_\varepsilon(t)f)\ ,$$

where the vector $D^2_{\star,3}T_\varepsilon(t)f$ contains the second-order derivatives $D^2_{ij}T_\varepsilon(t)f$ with $p_0 \le i \le j$, ordered lexicographically, whereas

$$D^3_{\star,1}T_\varepsilon(t)f = \{D^3_{ijh}T_\varepsilon(t)f : i \le j \le h \le p_0\}\ ,$$

$$D^3_{\star,2}T_\varepsilon(t)f = \{D^3_{ijh}T_\varepsilon(t)f : i \le j \le p_0 < h\}\ ,$$

$$D^3_{\star,3}T_\varepsilon(t)f = \{D^3_{ijh}T_\varepsilon(t)f : i \le p_0 < j \le h\}\ ,$$

$$D^3_{\star,4}T_\varepsilon(t)f = \{D^3_{ijh}T_\varepsilon(t)f : p_0 < i \le j \le h\}\ ,$$

each of these sets of derivatives being ordered lexicographically;

(b) the matrices $\mathcal{F}_2(t)$, $\mathcal{F}_3(t)$, which have to be positive definite at any $t \in]0,T]$, and split into blocks according to the above splitting of the vectors $D^2_\star T_\varepsilon(t)f$ and $D^3_\star T_\varepsilon(t)f$, may be of the following form:

$$\mathcal{F}_2(t) = \begin{pmatrix} t^2 I_{n_1^2} & 0 & 0 \\ 0 & \alpha^{\gamma_1}t^4 I_{n_2^2} & \alpha^{\gamma_2}t^5F_2 \\ 0 & \alpha^{\gamma_2}t^5F_2^* & \alpha^{\gamma_3}t^6 I_{n_3^2} \end{pmatrix}, \tag{3.11}$$

$$\mathcal{F}_3(t) = \begin{pmatrix} \alpha^{\gamma_4}t^3 I_{n_1^3} & 0 & 0 & 0 \\ 0 & \alpha^{\gamma_5}t^5 I_{n_2^3} & 0 & 0 \\ 0 & 0 & \alpha^{\gamma_6}t^7 I_{n_3^3} & \alpha^{\gamma_7}t^8F_3 \\ 0 & 0 & \alpha^{\gamma_7}t^8F_3^* & \alpha^{\gamma_8}t^9 I_{n_4^3} \end{pmatrix}, \tag{3.12}$$

where, n_j^k denotes the length of the vector $D^k_{\star,j}T_\varepsilon(t)f$ and γ_j $(j=1,\cdots,8)$ are to be properly chosen together with the constant α and the matrix F_2 and F_3, whose role is to provide us with the terms $|D^2_{\star,3}T_\varepsilon(t)f|^2$ and $|D^3_{\star,4}T_\varepsilon(t)f|^2$ with negative coefficients in front of them.

Some further details can be found in the proof of the forthcoming Theorem 5.1.

The estimates (3.3) and (3.5) may be obtained in a completely similar way with obvious modifications in the powers of t appearing in the matrices (3.8), (3.11) and (3.12).

Remark 3.2. Of course, the technique described above can be used also to estimate higher order derivatives of the functions $T_\varepsilon(t)f$ provided one requires much more regularity on the diffusion coefficients q_{ij} $(i,j=1,\cdots,p_0)$ as well as some growth assumptions at infinity of the derivatives $D^\alpha q_{ij}$. We refer the reader to [7] for further details.

3.2. **Step (ii).** Here, we show how the previous estimates can be used to prove the existence of a classical solution u of problem (1.3) with $g\equiv 0$. Fix $f\in C_b(\mathbb{R}^N)$. Using the estimates (3.2) and the fact that $D_tT_\varepsilon(t)f=\mathcal{A}_\varepsilon T_\varepsilon(t)f$, one can easily show that

$$\|T_\varepsilon(\cdot)f\|_{C^{1+\alpha/2,2+\alpha}([a,b]\times B(0,R))}\leq C\|f\|_\infty\ ,$$

for any $\alpha\in]0,1[$, any $0<a<b$, any $R>0$ and some positive constant $C=C(a,b,R)$, independent of ε. A compactness argument, based on the Ascoli-Arzelà theorem, shows that there exists an infinitesimal sequence $\{\varepsilon_n\}$ such that $T_{\varepsilon_n}(\cdot)f$ converges in $C^{1,2}([a,b]\times B(0,R))$ to a function $u:]0,+\infty[\times\mathbb{R}^N\to\mathbb{R}$.

The main effort consists in showing that u can be extended up to $t=0$ where it equals f. This can be done in three steps depending on the smoothness of f:

(a) first, one considers the case when $f\in C_c^2(\mathbb{R}^N)$;
(b) next, one deals with the case when $f\in C_c(\mathbb{R}^N)$;
(c) finally, one proves the assertion in the general case when $f\in C_b(\mathbb{R}^N)$.

The two first cases are rather simple. Indeed, one can use the fact that $T_\varepsilon(t)$ and $\mathcal{A}_\varepsilon$ commute on $C_c^2(\mathbb{R}^N)$ to prove that, when $f\in C_c^2(\mathbb{R}^N)$, $T_\varepsilon(\cdot)f$ is Lipschitz continuous in $[0,+\infty[$, with its Lipschitz constant being independent of ε. This, of course ensures the continuity at 0 of the function u as well as the fulfilling of the condition $u(0,\cdot)=f$. In particular, one sees that $T_\varepsilon(\cdot)f$ converges to $T(\cdot)f$ in $[a,b]\times B(0,R)$, as $\varepsilon\to 0^+$, for any a,b,R as above. Then, the density of $C_c^2(\mathbb{R}^N)$ in $C_c(\mathbb{R}^N)$ allows to extend the previous result to any function $f\in C_c(\mathbb{R}^N)$. In particular, one can show that the semigroup $\{T(t)\}$ is well defined in $C_c(\mathbb{R}^N)$ and $T_\varepsilon(\cdot)f$ converges to $T(\cdot)f$ in $C^{1,2}(F)$ for any compact set $F\subset]0,+\infty[\times\mathbb{R}^N$.

The case when $f\in C_b(\mathbb{R}^N)$ follows from (b) via an argument from [8, Proposition 2.2], based on a localization procedure. In fact, using the maximum principle, one can show that

$$T_\varepsilon(t)((1-\varphi)f)\leq\|f\|_\infty(1-T_\varepsilon(t)\varphi)\quad ,\quad t>0\ , \tag{3.13}$$

for any $\varphi\in C_b(\mathbb{R}^N)$ and any $\varepsilon>0$. Let $\{\varepsilon_n\}$ be an infinitesimal sequence such that $T_{\varepsilon_n}(\cdot)$ converges to u. Moreover, fix $R>0$ and a smooth and compactly supported function φ such that $\varphi\equiv 1$ in $B(0,R)$. Splitting $T_{\varepsilon_n}(t)f=T_{\varepsilon_n}(t)(\varphi f)+T_{\varepsilon_n}(t)((1-\varphi)f)$ and letting $n\to+\infty$, from (3.13) it is immediate to check that $u(t,\cdot)-T(t)(\varphi f)$ vanishes uniformly in $B(0,R)$ as t tends to 0. Since φf is compactly supported in $\mathbb{R}^N$, it follows easily that u can be extended by continuity to $\{0\}\times B(0,R)$ (and, hence, to $\{0\}\times\mathbb{R}^N$, by the arbitrariness of R) where it equals f.

3.3. **Step (iii).** Now, letting $\varepsilon\to 0^+$ in the estimates (a), (b) and (d) (with $k\leq 2$) in (3.2), (a), (d) in (3.3) and (3.6), we immediately get the corresponding estimates for the semigroup $\{T(t)\}$.

Showing that also the remaining estimates in (3.2) and (3.3) hold with $T(t)f$ instead of $T_\varepsilon(t)f$, is a bit more tricky. In fact, the argument in Step (ii) does not guarantee that the function $T(t)f$ is thrice continuously differentiable in $\mathbb{R}^N$ for any $t>0$. We briefly sketch how one can prove that the derivative $D^3_{ijh}T(t)f$ exists and that the estimate (3.2)(c) holds with $T(t)f$ instead of $T_\varepsilon(t)f$, in the particular (and simplest) case when $i = 1$ and $j, h \le p_0$. It is immediate to realize that one has only to prove the existence of the third-order derivative $D^3_{1jh}T(t)f$. Indeed, (3.2)(a)-(b) imply that the function $D_{jh}T(t)f(\cdot, x_2, \cdots, x_N)$ is Lipschitz continuous in $\mathbb{R}$ with a Lipschitz constant bounded from above by $Ct^{-5/2}\|f\|_{C_b(\mathbb{R}^N)}$ for any $t \in]0, 1]$ and some $C > 0$, *independent* of $x_2, \cdots, x_N$. Therefore, if the derivative $D^3_{1jh}T(t)f$ exists, then it satisfies (3.2)(c).

To prove the existence of $D^3_{1jh}u$, it suffices to adapt to our situation the classical Nirenberg translations' method. For technical reasons it is useful to assume that $f \in C^3_b(\mathbb{R}^N)$: this is not a loss in the generality at all. Indeed, first a density argument allows to show that $D^3_{ijh}T(t)f$ exists for any $t > 0$ and any bounded and uniformly continuous function f (say $f \in BUC(\mathbb{R}^N)$). Then, splitting $T(t)f = T(t/2)T(t/2)f$, one can obtain the same result in the general case when $f \in C_b(\mathbb{R}^N)$, observing that, by (3.2)(a), $T(t/2)f \in BUC(\mathbb{R}^N)$.

Let τ_k be the operator defined by $\tau_k\psi = [\psi(\cdot + ke_1) - \psi]/k$ for any function $\psi \in C_b(\mathbb{R}^N)$. Moreover, let $v_{k,R} = \tau_k(\eta_R T(t)f)$, where $\eta_R \in C^\infty_c(\mathbb{R}^N)$ satisfies $\chi_{B(0,R/2)} \le \eta_R \le \chi_{B(0,R)}$. By an approximation argument, one can show that the function $v_{k,R}$ solves the Cauchy problem (1.3) with f and g being replaced, respectively, by $\eta_R f$ and a continuous (and compactly supported) function $g_{k,R}$. The function $v_{k,R}$ may be expressed by the *pointwise*[2] variation of constant formula

$$v_{k,R}(t, x) = (T(t)(\tau_k(\eta_R f)))(x) + \int_0^t (T(t-s)g_{k,R}(s, \cdot))(x)\, ds\,, \tag{3.14}$$

$t > 0$, $x \in \mathbb{R}^N$. Letting $k \to 0$ in (3.14)[3] yields us to the *fundamental* formula

$$D_1(\eta_R T(t)f)(x) = (T(t)(D_1(\eta_R f)))(x) + \int_0^t (T(t-s)g_R(s, \cdot))(x)\, ds\,, \tag{3.15}$$

for any $t > 0$, $x \in \mathbb{R}^N$, where g_R is compactly supported in $B(0, R)$ and depends on the derivatives of $T(t)f$, up to the second order and in the directions corresponding to the first p_0 variables, on the drift coefficients and on the diffusion coefficients and their first-order derivatives.

Since

$$\|D^2_{ij}T(t-s)g_R(s, \cdot)\|_{C_b(\mathbb{R}^N)} \le C\|D^2_{ij}T(t-s)\|_{L(C^\theta_b(\mathbb{R}^N);C_b(\mathbb{R}^N))}\|g(s, \cdot)\|_{C^\theta_b(\mathbb{R}^N)}\,,$$

for any $t \in]0, +\infty[$ and any $\theta \in]0, 1[$, choosing properly θ and using (3.2), (3.3) and an interpolation argument, one can easily show that the function

$$s \mapsto \|D^2_{ij}T(t-s)g_R(s, \cdot)\|_{C(K)}$$

is integrable in $[0, t]$ for any compact set $K \subset \mathbb{R}^N$. Therefore, from (3.15) we immediately deduce that we can differentiate under the integral sign to get

$$D_{1jh}(\eta_R T(t)f)(x) = D_{jh}(T(t)(D_1(\eta_R f)))(x) + \int_0^t D_{jh}(T(t-s)g_R(s, \cdot))(x)\, ds\,,$$

for any $t > 0$ and any $x \in \mathbb{R}^N$. Whence we also deduce that $D_{1jh}T(\cdot)f$ is continuous in $]0, +\infty[\times\mathbb{R}^N$.

[2]The function $t \mapsto T(t)\psi$ might not be integrable with respect to the sup-norm, when $\psi \in C_b(\mathbb{R}^N)$, since, in general, the semigroup $\{T(t)\}$ is neither strongly continuous nor analytic in $C_b(\mathbb{R}^N)$ and $BUC(\mathbb{R}^N)$.

[3]Here, the smoothness assumption on f is needed!

The proof of the remaining estimates in (3.2), (3.3) and (3.5) is similar, but it requires a bootstrap argument and some more interpolation arguments.

Remark 3.3. It is worth stressing that the argument in Steps (ii) and (iii) are *independent* of the structure of the operator $\mathcal{A}$, since they rely *only* on the estimates (3.2), (3.3) and (3.6). Hence, they can be repeated *verbatim* whenever similar estimates are available.

4. Proof of the Schauder estimates in 1.1 and 1.2

As it has been claimed in the Introduction, the candidate solutions to the problems (1.2) and (1.3) are the functions v and u given by (1.4) and (1.5). For both the elliptic equation and the Cauchy problem, the main effort consists in proving that the functions v and $u(t,\cdot)$ ($t \in]0,T]$) are in $\mathcal{C}^{2+\theta}$, under the smoothness assumptions on f, g, h in 1.1 and 1.2.

Since the needed techniques for this aim are essentially the same for both v and $u(t,\cdot)$, let us concentrate for simplicity on the elliptic equation. Note that the estimate (3.6), written with the semigroup $\{T_\varepsilon(t)\}$ replaced with $\{T(t)\}$, and with $\gamma_1 = \theta \in]0,1[$ and $\gamma_2 = 2+\theta$, yields

$$\|T(t)f\|_{\mathcal{C}^{2+\theta}} \le Ct^{-1}\|f\|_{\mathcal{C}^\theta} \ ,$$

near $t=0$. Hence, we cannot estimate the $\mathcal{C}^{2+\theta}$-norm of the function v simply by computing the integral of the $\mathcal{C}^{2+\theta}$-norm of $T(t)f$.

This problem can be avoided applying an argument from [19], that we sketch here. Since $\mathcal{C}^{2+\theta} = (\mathcal{C}^\alpha, \mathcal{C}^{2+\alpha})_{1-(\alpha-\theta)/2,\infty}$, with equivalence of the corresponding norms, for any $\alpha > \theta$, it suffices to show that v belongs to this latter space, for some suitable choice of α. Recalling that $(\mathcal{C}^\alpha, \mathcal{C}^{2+\alpha})_{1-(\alpha-\theta)/2,\infty}$ is the set of φ's such that $|||\varphi||| := \sup_{\xi \in]0,1[} \xi^{1-(\alpha-\theta)/2} K(\xi,\varphi) < +\infty$, endowed with the norm $|||\cdot|||$, where

$$K(\xi,\varphi) = \inf\left\{\|a\|_{\mathcal{C}^\alpha} + \xi\|b\|_{\mathcal{C}^{2+\alpha}} \ : \ a \in \mathcal{C}^\alpha \ , \ b \in \mathcal{C}^{2+\alpha} \ , \ a+b=\varphi\right\} \ ,$$

(see e.g., [18, Chapter 1]), it suffices to find, for any $\xi \in]0,1[$, a suitable splitting of the integral defining v into two functions a_ξ and b_ξ such that

$$\xi^{1-(\alpha-\theta)/2}\left(\|a_\xi\|_{\mathcal{C}^\alpha} + \|b_\xi\|_{\mathcal{C}^{2+\alpha}}\right) \le C\|f\|_{\mathcal{C}^\theta} \ ,$$

for some positive constant C, independent of $\xi \in]0,1]$. Taking the estimates (3.6) into account, one can easily realize that a suitable splitting of v may be the following:

$$a_\xi(x) = \int_0^\xi e^{-\lambda t}(T(t)f)(x)\,dt \quad , \quad b_\xi(x) = \int_\xi^{+\infty} e^{-\lambda t}(T(t)f)(x)\,dt \ .$$

5. Pointwise estimates and invariant measures

This section is devoted to prove pointwise estimates for the first-, second- and third-order spatial derivatives of $T(t)f$ and some nice consequences. This is the original part of the paper.

5.1. **Pointwise estimate.**

Theorem 5.1. *For any* $f \in C_b^l(\mathbb{R}^N)$*, any* $l = 1, 2, 3$ *and any* $p \in]1, +\infty[$*, we have*

$$|(D^l T(t) f)(x)|^p \leq M_p \left\{ T(t) \left(|f|^2 + \cdots + |D^l f|^2 \right)^{p/2} \right\} (x) \, , \tag{5.1}$$

for any $t \in]0, +\infty[$*,* $x \in \mathbb{R}^N$ *and some positive constant* M_p*.*

Proof. We begin with the case when $l = 3$ and $p \in]1, 2]$. To simplify the notation, throughout the proof we set $u_\varepsilon = T_\varepsilon(\cdot) f$, where $\{T_\varepsilon(t)\}$ is the approximating semigroup defined in Step (i).

For any $\delta > 0$ we introduce the function $w_{\delta,\varepsilon} : [0, +\infty[\times \mathbb{R}^N \to \mathbb{R}$ defined by

$$w_{\delta,\varepsilon} = \left(\alpha^3 |u_\varepsilon|^2 + \sum_{k=1}^{3} \langle \mathcal{F}_k \, D_\star^k u_\varepsilon, D_\star^k u_\varepsilon \rangle + \delta \right)^{p/2} ,$$

where $D_\star^1 = D$, $D_\star^2$ and $D_\star^3$ are as in Subsection 3.1.2 and

$$\mathcal{F}_1 = \begin{pmatrix} \alpha I_{p_0} & F_1 \\ F_1^* & I_{p_0} \end{pmatrix} \quad , \quad \mathcal{F}_2 = \begin{pmatrix} I_{n_1^2} & 0 & 0 \\ 0 & \alpha^{-7/16} \, I_{n_2^2} & \alpha^{-4/5} \, F_2 \\ 0 & \alpha^{-4/5} \, F_2^* & \alpha^{-7/8} \, I_{n_3^2} \end{pmatrix} ,$$

$$\mathcal{F}_3 = \begin{pmatrix} \alpha^{-7/16} \, I_{n_1^3} & 0 & 0 & 0 \\ 0 & \alpha^{-7/8} \, I_{n_2^3} & 0 & 0 \\ 0 & 0 & \alpha^{-1} \, I_{n_3^3} & \alpha^{-13/12} \, F_3 \\ 0 & 0 & \alpha^{-13/12} \, F_3^* & \alpha^{-9/8} \, I_{n_4^3} \end{pmatrix} ,$$

with the matrices F_1, F_2 and F_3, as well as the parameter α to be determined later on. For simplicity, in what follows we simply write u for u_ε.

By the results in the previous sections, we know that $w_\delta \in C_b([0, T] \times \mathbb{R}^N) \cap C^{1,2}(]0, +\infty[\times \mathbb{R}^N)$ for any $T > 0$. Moreover, $w_{\delta,\varepsilon}$ solves the Cauchy problem (1.3) with

$$\begin{aligned} g = p \Bigg(& -\alpha^3 \sum_{i,j=1}^{N} q_{ij}^{(\varepsilon)} D_i u D_j u - \sum_{k=1}^{3} \sum_{i,j=1}^{N} q_{ij}^{(\varepsilon)} \langle \mathcal{F}_k D_\star^k D_i u, D_\star^k D_j u \rangle + \\ & + \sum_{k=1}^{3} \langle \mathcal{F}_k \, [D_\star^k, \mathcal{A}_\varepsilon] u, D_\star^k u \rangle \Bigg) w^{1-(2/p)} + \\ & + \Bigg\{ \sum_{i,j=1}^{N} q_{ij}^{(\varepsilon)} \left(\alpha^3 u D_i u + \sum_{h=1}^{3} \langle \mathcal{F}_h D_\star^h D_i u, D_\star^h u \rangle \right) \times \\ & \times \left(\alpha^3 u D_j u + \sum_{k=1}^{3} \langle \mathcal{F}_k D_\star^k D_j u, D_\star^k u \rangle \right) \Bigg\} p(2-p) w^{1-(4/p)} \, , \end{aligned}$$

and f being replaced by $(\alpha^3|f|^2+\sum_{k=1}^3\langle\mathcal{F}_k\,D_\star^k f, D_\star^k f\rangle)^{p/2}$. Here, $[\cdot,\cdot]$ denotes the commutator of the quantities in brackets.

Let M and P be, respectively, a $m\times m$ and a $p\times p$ positive and symmetric square matrices. Further let $\xi_j,\xi\in\mathbb{R}^m$ and $\eta_j,\eta\in\mathbb{R}^p$ $(j=1,\cdots,p_0)$. Then, applying twice the Cauchy-Schwarz inequality (firstly to the inner product induced by the matrix $Q=(q_{ij})$ and then to the Euclidean one) we deduce that

$$\sum_{i,j=1}^N q_{ij}^{(\varepsilon)}\langle M\xi_i,\xi\rangle\langle P\eta_j,\eta\rangle \le \sum_{h=1}^m\sum_{k=1}^p (M^{1/2}\xi)_h(P^{1/2}\eta)_k\ \times$$

$$\times\left(\sum_{i,j=1}^N q_{ij}^{(\varepsilon)}(M^{1/2}\xi_i)_h(M^{1/2}\xi_j)_h\right)^{1/2}\left(\sum_{i,j=1}^N q_{ij}^{(\varepsilon)}(P^{1/2}\eta_i)_k(P^{1/2}\eta_j)_k\right)^{1/2}\le$$

$$\le(\langle M\xi,\xi\rangle)^{1/2}(\langle P\eta,\eta\rangle)^{1/2}\left(\sum_{i,j=1}^N q_{ij}^{(\varepsilon)}\langle M\xi_i,\xi_j\rangle\right)^{1/2}\left(\sum_{i,j=1}^N q_{ij}^{(\varepsilon)}\langle P\eta_i,\eta_j\rangle\right)^{1/2}.$$

This estimate can be used in order to get

$$\sum_{i,j=1}^N q_{ij}^{(\varepsilon)}\left(\alpha^3 uD_iu+\sum_{h=1}^3\langle\mathcal{F}_hD_\star^hD_iu,D_\star^hu\rangle\right)\times \tag{5.2}$$

$$\times\left(\alpha^3uD_ju+\sum_{k=1}^3\langle\mathcal{F}_kD_\star^kD_ju,D_\star^ku\rangle\right)\le\left(\alpha^3|u|^2+\sum_{k=1}^3\langle\mathcal{F}_k\,D_\star^ku,D_\star^ku\rangle\right)\times$$

$$\times\left(\alpha^3\sum_{i,j=1}^N q_{ij}^{(\varepsilon)}D_iuD_ju+\sum_{k=1}^3\sum_{i,j=1}^N q_{ij}^{(\varepsilon)}\langle\mathcal{F}_kD_\star^kD_iu,D_\star^kD_ju\rangle\right).$$

Taking (5.2) into account and observing that $[D_\star^k,\mathcal{A}_\varepsilon]=[D_\star^k,\mathcal{A}]$ $(k=1,2,3)$, it is immediate to check that

$$g\le p\left\{(1-p)\left[\alpha^3\sum_{i,j=1}^N q_{ij}^{(\varepsilon)}D_iuD_ju+\sum_{k=1}^3\sum_{i,j=1}^N q_{ij}^{(\varepsilon)}\langle\mathcal{F}_k\,D_\star^kD_iu,D_\star^kD_ju\rangle\right]+\right.$$

$$\left.+\sum_{k=1}^3\langle\mathcal{F}_k\,[D_\star^k,\mathcal{A}_\varepsilon]u,D_\star^ku\rangle\right\}w^{1-(2/p)}\le$$

$$\le p\left\{(1-p)\left[\alpha^3\sum_{i,j=1}^{p_0}q_{ij}D_iuD_ju+\sum_{k=1}^3\sum_{i,j=1}^{p_0}q_{ij}\langle\mathcal{F}_k\,D_\star^kD_iu,D_\star^kD_ju\rangle\right]+\right.$$

$$\left.+\sum_{k=1}^3\langle\mathcal{F}_k\,[D_\star^k,\mathcal{A}]u,D_\star^ku\rangle\right\}w^{1-(2/p)}=p\left\{(1-p)g_1+g_2\right\}w^{1-(2/p)}.$$

A straightforward computation shows that

$$g_1(t)\ge\alpha^3\nu|D_{\star,1}u|^2+\alpha\nu|D_{\star,1}^2u|^2+\nu|D_{\star,2}^2u|^2+\nu\langle K_1D_{\star,1}^2u,D_{\star,2}^2u\rangle+\nu|D_{\star,1}^3u|^2+$$

$$+\alpha^{-7/16}\nu|D_{\star,2}^3u|^2+\alpha^{-7/8}\nu|D_{\star,3}^3u|^2+\alpha^{-4/5}\nu\langle K_2D_{\star,2}^3u,D_{\star,3}^3u\rangle+\alpha^{-7/16}\nu|D_{\star,1}^4u|^2+$$

$$+\alpha^{-7/8}\nu|D_{\star,2}^4u|^2+\alpha^{-1}\nu|D_{\star,3}^4u|^2+\alpha^{-9/8}\nu|D_{\star,4}^4u|^2+\nu\alpha^{-13/12}\langle K_3D_{\star,3}^4u,D_{\star,4}^4u\rangle\,,$$

for any $t > 0$, where K_j $(j = 1, 2, 3)$ are suitable square matrix, whose coefficients are independent of α. Hence, using Young inequality, we get

$$
\begin{aligned}
(5.3)\qquad g_1(t) \geq\ & \alpha^3\nu|D_{\star,1}u|^2 + \alpha\nu|D^2_{\star,1}u|^2 + \nu|D^2_{\star,2}u|^2 + \nu|D^3_{\star,1}u|^2 + \\
& +\alpha^{-7/16}\nu|D^3_{\star,2}u|^2 + \alpha^{-7/8}\nu|D^3_{\star,3}u|^2 + \alpha^{-7/16}\nu|D^4_{\star,1}u|^2 + \\
& +\alpha^{-7/8}\nu|D^4_{\star,2}u|^2 + \alpha^{-1}\nu|D^4_{\star,3}u|^2 + \alpha^{-9/8}\nu|D^4_{\star,4}u|^2 - \\
& -\|K_1\|\nu(\alpha^{1/2}|D^2_{\star,1}u|^2 + \alpha^{-1/2}|D^2_{\star,2}u|^2) - \\
& -\|K_2\|\nu(\alpha^{-3/5}|D^3_{\star,2}u|^2 + \alpha^{-1}|D^3_{\star,3}u|^2) - \\
& -\|K_3\|\nu(\alpha^{-49/48}|D^4_{\star,3}u|^2 + \alpha^{-55/48}|D^4_{\star,4}u|^2) \geq \\
& \geq \alpha^3\nu|D_{\star,1}u(t)|^2 + (2+o(1))\alpha\nu|D^2_{\star,1}u(t)|^2 + \\
& +(2+o(1))\nu|D^2_{\star,2}u(t)|^2 + 2\nu|D^3_{\star,1}u(t)|^2 + \\
& +(2+o(1))\alpha^{-7/16}\nu|D^3_{\star,2}u(t)|^2 + (2+o(1))\alpha^{-7/8}\nu|D^3_{\star,3}u(t)|^2 + \\
& +2\alpha^{-7/16}\nu|D^4_{\star,1}u(t)|^2 + 2\alpha^{-7/8}\nu|D^4_{\star,2}u(t)|^2 + \\
& +(2+o(1))\alpha^{-1}\nu|D^4_{\star,3}u(t)|^2 + (2+o(1))\alpha^{-9/8}\nu|D^4_{\star,4}u(t)|^2 ,
\end{aligned}
$$

for any $t > 0$, where by $o(1)$ we have denoted suitable functions of α (depending also on F_2 and F_3) vanishing as $\alpha \to +\infty$.

As far as g_2 is concerned, in [16, Thm 3.1] we have proved that $[D, \langle B\cdot, D\rangle]u = B^*Du$, $[D^k_\star, \langle B\cdot, D\rangle]u = \mathcal{M}_k D^k_\star u$ $(k = 2, 3)$, where

$$
\mathcal{M}_1 = \begin{pmatrix} \star & \star & 0 \\ \star & \star & M_1 \\ 0 & \star & \star \end{pmatrix} \quad , \quad \mathcal{M}_2 = \begin{pmatrix} \star & \star & 0 & 0 \\ \star & \star & \star & 0 \\ 0 & \star & \star & M_2 \\ 0 & 0 & \star & \star \end{pmatrix} ,
$$

$\star$, M_1 and M_2 being suitable matrices whose entries linearly depend on the entries of B, but are *independent* of α, F_2, F_3. In particular, the matrices M_1 and M_2 have full rank. See also [7] for further details. Hence,

$$
\begin{aligned}
(5.4)\qquad & 2\langle \mathcal{F}_1[D, \langle B\cdot, D\rangle]u, Du\rangle = \langle(\mathcal{F}B^* + B\mathcal{F}^*)Du, Du\rangle = \\
& = 2(\alpha\|\star\| + \|F_1\|\cdot\|\star\|)|D_{\star,1}u|^2 + \|B_1\|(\alpha^{5/2}|D_{\star,1}u|^2 + \alpha^{-1/2}|D_{\star,2}u|^2) + \\
& +(\|F_1\|\cdot\|\star\| + \|F_1\|\cdot\|\star\| + \|\star\|)(\alpha|D_{\star,1}u|^2 + \alpha^{-1}|D_{\star,2}u|^2) + \\
& +(\lambda_{\max}(F_1B_1^* + B_1F_1^*) + 2\|\star\|)|D_{\star,2}u|^2 = \\
& = \alpha^{5/2}(\|B_1\| + o(1)\}|D_{\star,1}u|^2 + \{\lambda_{\max}(F_1B_1^* + B_1F_1^*) + 2\|\star\| + o(1)\}|D_{\star,2}u|^2 ,
\end{aligned}
$$

where by $\lambda_{\max}(E)$ we denote the maximum eigenvalue of the matrix E and "$\star$" are as in (1.8). In particular, in the last term of (5.4), "$\star$" denotes the last block of B. Similarly,

$$
\begin{aligned}
(5.5)\qquad & 2\langle \mathcal{F}_2[D^2_\star, \langle B\cdot, D\rangle]u, D^2_\star u\rangle \leq \\
& \leq 2\|\star\|\cdot|D^2_{\star,1}u|^2 + \|\star\|(\alpha^{1/2}|D^2_{\star,1}u|^2 + \alpha^{-1/2}|D^2_{\star,2}u|^2) + \\
& +\|F_2\|\cdot\|\star\|(|D^2_{\star,1}u|^2 + \alpha^{-8/5}|D^2_{\star,3}u|^2) + \alpha^{-7/16}\|\star\|(|D^2_{\star,1}u|^2 + |D^2_{\star,2}u|^2) + \\
& +2\alpha^{-7/16}\|\star\||D^2_{\star,2}u|^2 + 2\alpha^{-4/5}\|F_2\|\cdot\|\star\|\cdot|D^2_{\star,2}u|^2 +
\end{aligned}
$$

$$+\|M_1\|(\alpha^{-1/24}|D^2_{\star,2}u|^2+\alpha^{-5/6}|D^2_{\star,3}u|^2)+$$
$$+\|F_2\|(\|\star\|+\|\star\|)(\alpha^{-3/5}|D^2_{\star,2}u|^2+\alpha^{-1}|D^2_{\star,3}u|^2)+$$
$$+\alpha^{-7/8}\|\star\|\left(|D^2_{\star,2}u|^2+|D^2_{\star,3}u|^2\right)+\alpha^{-4/5}\lambda_{\max}(F_2^*M_1+M_1^*F_2)|D^2_{*,3}u|^2+$$
$$+2\alpha^{-7/8}\|\star\|\cdot|D^2_{\star,3}u|^2=$$
$$=\{\|\star\|+o(1)\}\alpha^{1/2}|D^2_{*,1}u|^2+\alpha^{-1/24}\{\|M_1\|+o(1)\}|D^2_{\star,2}u|^2+$$
$$+\alpha^{-4/5}\{\lambda_{\max}(F_2^*M_1+M_1^*F_2)+o(1)\}|D^2_{\star,3}u|^2\ ,$$

and

$$(5.6)\qquad 2\langle\mathcal{F}_3[D^3_\star,\langle B\cdot,D\rangle]u,D^3_\star u\rangle\leq$$
$$\leq 2\alpha^{-7/16}\|\star\|\cdot|D^3_{\star,1}u|^2+\|\star\|(\alpha^{-1/8}|D^3_{\star,1}u|^2+\alpha^{-3/4}|D^3_{\star,2}u|^2)+$$
$$+\|\star\|\alpha^{-7/8}(|D^3_{\star,1}u|^2+|D^3_{\star,2}u|^2)+2\alpha^{-7/8}\|\star\|\cdot|D^3_{\star,2}u|^2+$$
$$+\|\star\|(\alpha^{-1/2}|D^3_{\star,2}u|^2+\alpha^{-5/4}|D^3_{\star,3}u|^2)+\|\star\|(\alpha^{-3/4}|D^3_{\star,2}u|^2+\alpha^{-5/4}|D^3_{\star,3}u|^2)+$$
$$+2(\alpha^{-1}\|\star\|+\alpha^{-13/12}\|F_3\|\cdot\|\star\|)|D^3_{\star,3}u|^2+\|M_2\|(\alpha^{-43/48}|D^3_{\star,3}u|^2+\alpha^{-53/48}|D^3_{\star,4}u|^2)+$$
$$+\|F_3\|\cdot\|\star\|(\alpha^{-25/24}|D^3_{\star,3}u|^2+\alpha^{-27/24}|D^3_{\star,4}u|^2)+$$
$$+\|F_3\|\cdot\|\star\|(\alpha^{-1}|D^3_{\star,2}u|^2+\alpha^{-7/6}|D^3_{\star,4}u|^2)+$$
$$+\|F_3\|\cdot\|\star\|(\alpha^{-25/24}|D^3_{\star,3}u|^2+\alpha^{-27/24}|D^3_{\star,4}u|^2)+\alpha^{-9/8}\|\star\|(|D^3_{\star,3}u|^2+|D^3_{\star,4}u|^2)+$$
$$+(\alpha^{-13/12}\lambda_{\max}(F_3^*M_2+M_2^*F_3)+2\alpha^{-9/8}\|\star\|)|D^3_{\star,4}u|^2=$$
$$=\alpha^{-1/8}\{\|\star\|+o(1)\}|D^3_{\star,1}u|^2+\alpha^{-1/2}\{\|\star\|+o(1)\}|D^3_{\star,2}u|^2+$$
$$+\alpha^{-43/48}\{\|M_2\|+o(1)\}|D^3_{\star,3}u|^2+\alpha^{-13/12}\{\lambda_{\max}(F_3^*M_2+M_2^*F_3)+o(1)\}|D^3_{\star,4}u|^2\ .$$

Finally, we observe that, denoting by $\mathcal{A}_0$ the principal part of the operator $\mathcal{A}$, from [16, Theorem 3.1], we have

$$(5.7)\qquad \sum_{k=1}^{3}\langle\mathcal{F}_k\,[D^k_\star,\mathcal{A}_0]u,D^k_\star u\rangle=\sum_{k=1}^{3}\sum_{i,j=1}^{p_0}D_{ij}u\langle\mathcal{F}_k\,D^k_\star q_{ij},D^k_\star u\rangle+$$
$$+\langle\mathcal{F}_2\,\mathcal{N}D^3_\star u,D^2_\star u\rangle+\langle\mathcal{F}_3\,\mathcal{P}D^3_\star u,D^3_\star u\rangle+\langle\mathcal{F}_3\,\mathcal{R}D^4_\star u,D^3_\star u\rangle\ ,$$

where, for any $x\in\mathbb{R}^N$, the matrices $\mathcal{N}(x)$, $\mathcal{P}(x)$ and $\mathcal{R}(x)$, split according to the splitting of the vectors $D^k_\star u_\varepsilon$ ($k=2,3,4$), are given by

$$\mathcal{N}(x)=\begin{pmatrix}N_1(x)&0&0&0\\N_2(x)&N_3(x)&0&0\\0&N_4(x)&0&0\end{pmatrix}\quad,\quad\mathcal{P}(x)=\begin{pmatrix}P_1(x)&0&0&0\\P_2(x)&P_3(x)&0&0\\P_4(x)&P_5(x)&0&0\\0&P_6(x)&0&0\end{pmatrix},$$

$$\mathcal{R}(x)=\begin{pmatrix}R_1(x)&0&0&0&0\\R_2(x)&R_3(x)&0&0&0\\0&R_4(x)&R_5(x)&0&0\\0&0&R_6(x)&0&0\end{pmatrix},$$

for any $x \in \mathbb{R}^N$. Moreover, there exists a positive constant C, independent of α, G_1, H_1 and x, such that

$$\|N_i(x)\| + \|P_j(x)\| + \|R_j(x)\| \le C_2\sqrt{\nu(x)} \quad , \quad x \in \mathbb{R}^N,\ i = 1, \cdots, 4,\ j = 1, \cdots, 6 .$$

Therefore, the terms in the right-hand side of (5.7) are all negligible with respect to the terms in (5.3)-(5.6), provided we fix the matrices F_1, F_2 and F_3 such that

$$\lambda_{\max}(F_1B_1^* + B_1F_1^*) + 2\|\star\| < 0 \quad , \quad \lambda_{\max}(F_2^*M_1 + M_1^*F_2) < 0 ,$$
$$\lambda_{\max}(F_3^*M_2 + M_2^*F_3) < 0 .$$

This is possible since the matrices B_1, M_1 and M_2 have full rank.

Combining the estimates so far proved, we get

$$
\begin{aligned}
g \le p\big\{&\alpha^3\nu[1 - p + o(1)]|D_{\star,1}u|^2 + \\
&+[\lambda_{\max}(F_1B_1^* + B_1F_1^*) + 2\|\star\| + o(1)]\,|D_{\star,2}u|^2 + \\
&+[2 - 2p + o(1)]\alpha\nu|D^2_{\star,1}u|^2 + [2 - 2p + o(1)]\nu|D^2_{\star,2}u|^2 + \\
&+[\lambda_{\max}(F_2^*M_1 + M_1^*F_2) + o(1)]\alpha^{-4/5}|D^2_{\star,3}u|^2 + [2 - 2p + o(1)]\nu|D^3_{\star,1}u|^2 + \\
&+[2 - 2p + o(1)]\alpha^{-7/16}\nu|D^3_{\star,2}u|^2 + [2 - 2p + o(1)]\alpha^{-7/8}\nu|D^3_{\star,3}u|^2 + \\
&+[\lambda_{\max}(F_3^*M_2 + M_2^*F_3) + o(1)]\alpha^{-13/12}|D^3_{\star,4}u|^2 + [2 - 2p + o(1)]\alpha^{-7/16}\nu|D^4_{\star,1}u|^2 + \\
&+[2 - 2p + o(1)]\alpha^{-7/8}\nu|D^4_{\star,2}u|^2 + [2 - 2p + o(1)]\alpha^{-1}\nu|D^4_{\star,3}u|^2 + \\
&+[2 - 2p + o(1)]\alpha^{-9/8}\nu|D^4_{\star,4}u|^2\big\}w^{1-(2/p)} .
\end{aligned} \tag{5.8}
$$

Taking α sufficiently large, we can make all the coefficients of the terms in the right-hand side of (5.8) negative in $[0, +\infty[\times\mathbb{R}^N$. Hence, with this choice of α, w_δ satisfies

$$
\begin{cases}
D_t w_\delta \le \mathcal{A} w_\delta , & \text{in } [0, +\infty[\times\mathbb{R}^N , \\
w_\delta(0, \cdot) = \left(|f|^2 + \sum_{k=1}^3 \langle \mathcal{F}_k\, D^k_\star f, D^k_\star f\rangle + \delta\right)^{p/2} , & \text{in } \mathbb{R}^N .
\end{cases}
$$

Then, by comparison, using the maximum principle in Proposition 2.3, we get

$$w_{\delta,\varepsilon}(t, \cdot) \le T_\varepsilon(t) \left\{ \left(|f|^2 + \sum_{k=1}^3 \langle \mathcal{F}_k\, D^k_\star f, D^k_\star f\rangle + \delta \right)^{p/2} \right\} \quad , \quad t > 0 ,$$

for any $\delta > 0$. Taking the limit, first, as $\varepsilon \to 0^+$ and, then, as $\delta \to 0^+$, the assertion immediately follows when $p \in]1, 2]$, provided we assume that the matrices $\mathcal{F}$, $\mathcal{G}$ and $\mathcal{H}$ are positive definite. This is surely the case if α is sufficiently large.

The case when $p > 2$ now follows from Jensen inequality, observing that, by the forthcoming formula (5.10), $(T(t)\psi)^{p/2} \le T(t)(\psi^{p/2})$ for any $\psi \in C_b(\mathbb{R}^N)$ and any $t > 0$. Indeed,

$$
\begin{aligned}
|D^3T(t)f|^p &= \left(|D^3T(t)f)|^2\right)^{p/2} \le \\
&\le \left(M_2\, T(t)\left(|f|^2 + |Df|^2 + |D^2f|^2 + |D^3f|^2\right)\right)^{p/2} \le \\
&\le M_2^{p/2}\, T(t)\left[\left(|f|^2 + |Df|^2 + |D^2f|^2 + |D^3f|^2\right)^{p/2}\right] ,
\end{aligned}
$$

for any $t > 0$.

To get (5.1) in the case when $p \in]1, 2]$ and $l = 1, 2$, it suffices to apply the previous arguments to the function

$$w_{\delta,\varepsilon}(t, \cdot) = \left(|u_\varepsilon(t, \cdot)|^2 + \langle \mathcal{F}_1\, Du_\varepsilon(t, \cdot), Du_\varepsilon(t, \cdot)\rangle + \delta\right)^{p/2} ,$$

if $l = 1$ and to the function

$$w_{\delta,\varepsilon}(t,\cdot) = \left(|u_\varepsilon(t,\cdot)|^2 + \langle \mathcal{F}_1\, Du_\varepsilon(t,\cdot), Du_\varepsilon(t,\cdot)\rangle + \langle \mathcal{F}_2\, D^2_\star u_\varepsilon(t,\cdot), D^2_\star u_\varepsilon(t,\cdot)\rangle\right)^{p/2} ,$$

if $l = 2$. Estimate (5.1) with $l = 1, 2$ and $p > 2$ can, then, be obtained by means of Jensen inequality as in the case $l = 3$.

□

Remark 5.2. One can easily see that all the techniques used in this and in the previous sections can be applied to the operator

$$\mathcal{A}_F\varphi = \sum_{i,j=1}^{p_0} q_{ij} D_{ij}\varphi + \sum_{i,j=1}^{N} b_{ij} x_j D_i\varphi + \sum_{j=1}^{p_0} F_j D_j\varphi ,$$

obtained perturbing the operator $\mathcal{A}$ by a drift term F which satisfies the following conditions:

(i) $F_j \in C^3(\mathbb{R}^N)$ for any $j = 1, \cdots, p_0$;

(ii) there exists a positive constant K such that

$$|D^\alpha F_j(x)| \le K\sqrt{\nu(x)} \quad , \quad x \in \mathbb{R}^N,\ |\alpha| \le 3\ ,\ \ j = 1, \cdots, p_0\ .$$

5.2. **Invariant measures.** By definition, an invariant measure of the semigroup $\{T(t)\}$ is any (probability) measure μ such that

$$(5.9) \qquad \int_{\mathbb{R}^N} T(t)f\ d\mu = \int_{\mathbb{R}^N} f\ d\mu\ ,$$

for any $t > 0$ and any $f \in C_b(\mathbb{R}^N)$. Note that an invariant measure of $\{T(t)\}$ may not exist. Indeed, in the case when $\mathcal{A}$ is the Ornstein-Uhlenbeck operator, an invariant measure exists if and only if B is of negative type, i.e., if and only if the eigenvalues of B have negative real part.

As in the nondegenerate case, a sufficient condition ensuring the existence of an invariant measure of $\{T(t)\}$, is the following:

Hypothesis 5.3.

There exists a (Lyapunov) function $\varphi \in C^2(\mathbb{R}^N)$ such that $\mathcal{A}\varphi$ tends to $-\infty$ as $|x| \to +\infty$.

In our situation, in view of Remark 2.2, Hypothesis 5.3 is satisfied taking $\varphi(x) = 1 + |x|^2$ provided we assume that B is a matrix of negative type such that

$$\lambda_{\max}(B + B^*) < -2 \sup_{x\in\mathbb{R}^N} \frac{\mathrm{Tr}(Q(x))}{|x|^2}\ .$$

As it has been proved in [17, Corollary 4.2], there exists a family of probability Borel measures $\{p(t, x, dy)\}$ such that

$$(5.10) \qquad (T(t)f)(x) = \int_{\mathbb{R}^N} f(y)p(t, x, dy) \quad , \quad t > 0,\ x \in \mathbb{R}^N\ ,$$

for any $f \in C_b(\mathbb{R}^N)$.

Let us define the family of probability measures $\{r(t, x; dy),\ t > 0,\ x \in \mathbb{R}^N\}$ by setting

$$r(t, x; B) = \frac{1}{t}\int_0^t p(s, x; B)\ ds\ ,$$

for any Borel set B, where $p(s, x; B) = (T(s)\chi_B)(x)$.

Theorem 5.4 (Krylov-Bogoliubov + Khas'minskii). *Suppose that Hypotheses* 2.1 *and* 5.3 *are satisfied. Then, the semigroup* $\{T(t)\}$ *admits an invariant measure* μ.

Proof. To prove the assertion, it suffices to show that the family of measures $\{r(t,0;dy),\ t>1\}$ is tight.[4] Indeed, once this property is proved, we can apply Prokhorov theorem which implies the existence of a sequence $\{t_n\}$ diverging to $+\infty$ as $n\to+\infty$ such that $r(t_n,0;dy)$ converges weakly* to a probability measure μ. The measure μ is an invariant measure of $\{T(t)\}$. Indeed, for any $f\in C_b(\mathbb{R}^N)$, we have

$$\int_{\mathbb{R}^N}(T(t)f)(y)r(t_n,x_0;dy)=\frac{1}{t_n}\int_0^{t_n}(T(s+t)f)(0)\,ds=$$
$$=\int_{\mathbb{R}^N}f(y)r(t_n,x_0;dy)\ +\frac{1}{t_n}\int_{t_n}^{t_n+t}(T(s)f)(x_0)\,ds-$$
$$-\frac{1}{t_n}\int_0^t(T(s)f)(x_0)\,ds\ .$$

Since $\|T(t)\|_{L(C_b(\mathbb{R}^N))}=1$ for any $t\geq 0$, letting n tend to $+\infty$ we immediately deduce that μ is an invariant measure of $\{T(t)\}$.

So, let us check the tightness of the measures $\{r(t,0;dy):t>1\}$. Let $\{\psi_n\}\in C^\infty([0,+\infty[)$ be a sequence of smooth functions with the following properties:

(i) $\psi_n(t)=t$ for any $t\in[0,n]$;
(ii) $\psi_n(t)=n+1$ for any $t\geq n+2$;
(iii) $\psi_n'(t)\in[0,1]$ and $\psi_n''(t)\leq 0$ for any $t\geq 0$.

Note that $D_tT(t)\varphi_n=T(t)\mathcal{A}\varphi_n$ for any $n\in\mathbb{N}$ and any $t>0$. Indeed, for any $n\in\mathbb{N}$, the function φ_n-n-1 belongs to $C_c^2(\mathbb{R}^N)$ and the semigroup commutes with $\mathcal{A}$ in $C_c^2(\mathbb{R}^N)$ as it has been already recalled. Hence,

$$D_sT(s)\varphi_n=D_sT(s)\left(\varphi_n-n-1\right)=T(s)\mathcal{A}\left(\varphi_n-n-1\right)=T(s)\mathcal{A}\varphi_n\ \ ,\ \ s>0\ ,$$

so that

$$(D_sT(s)\varphi_n)(0)=\int_{\mathbb{R}^N}\left(\psi_n'(\varphi)\mathcal{A}\varphi+\psi_n''(\varphi)\sum_{i,j=1}^{p_0}q_{ij}D_i\varphi D_j\varphi\right)p(t,0;dy)\ , \tag{5.11}$$

for any $s>0$. We now fix $t>0$ and integrate (5.11) with respect to $s\in[0,t]$. Recalling that $T(\cdot)\varphi_n\geq 0$ in $[0,+\infty[\times\mathbb{R}^N$ and $\psi_n''\leq 0$ in $[0,+\infty[$, we get

$$-\varphi_n(0)\leq\int_0^t ds\int_E\psi_n'(\varphi)(\mathcal{A}\varphi)p(s,0;dy)+ \tag{5.12}$$
$$+\int_0^t ds\int_{\mathbb{R}^N\setminus E}\psi_n'(\varphi)(\mathcal{A}\varphi)p(s,0;dy)\ ,$$

where $E=\{y\in\mathbb{R}^N:\mathcal{A}\varphi(y)\geq 0\}$. Since E is a bounded set, the first integral term in the last side of (5.12) converges to

$$\int_0^t ds\int_E(\mathcal{A}\varphi)p(s,0;dy)\ ,$$

[4]i.e., we have to show that for any $\varepsilon>0$ there exists $R>0$ such that $r(t,0,B(R))\geq 1-\varepsilon$ for any $t>1$.

as $n \to +\infty$, by dominated convergence. On the other hand, since the sequence $\{\psi_n'(s)\}$ is increasing to 1 for any $s \in [0, +\infty[$, and $\mathcal{A}\varphi \le 0$ in $\mathbb{R}^N \setminus E$, the second integral term tends to

$$\int_0^t ds \int_{\mathbb{R}^N \setminus E} (\mathcal{A}\varphi) p(s, 0; dy) \ ,$$

by monotone convergence. Note that this latter integral is finite due to (5.12). Therefore,

$$-\varphi(0) \le \int_0^t ds \int_{\mathbb{R}^N} (\mathcal{A}\varphi) p(s, 0; dy) \ . \tag{5.13}$$

Now, fix ε and let $R > 0$ be such that $\mathcal{A}\varphi < -\varepsilon^{-1}$ outside the ball $B(0, R)$. From (5.13) we obtain that

$$-\varphi(0) \le Mt - \frac{1}{\varepsilon} \int_0^t p(s, 0; \mathbb{R}^N \setminus B(0, R))\, ds = Mt - \frac{t}{\varepsilon} r(t, 0; \mathbb{R}^N \setminus B(0, R)) \ ,$$

where $M = \sup_{\mathbb{R}^N} \mathcal{A}\varphi$. Consequently,

$$r(t, 0, \mathbb{R}^N \setminus B(0, R)) \le \varepsilon \frac{Mt + \varphi(0)}{t} \le \varepsilon (M + \varphi(0)) \quad , \quad t > 1 \ ,$$

which is, of course, the claim.

□

Before going on describing some of the properties of μ, let us prove the following lemma.

Lemma 5.5. *The semigroup $\{T(t)\}$ extends to $B_b(\mathbb{R}^N)$ (the set of all bounded Borel functions $f : \mathbb{R}^N \to \mathbb{R}$) with a strong Feller semigroup, i.e., $T(t)$ maps $B_b(\mathbb{R}^N)$ into $C_b(\mathbb{R}^N)$ (and, hence, into $C_b^3(\mathbb{R}^N)$) for any $t \in]0, +\infty[$.*

Proof. Since any $f \in B_b(\mathbb{R}^N)$ is the pointwise limit of a bounded sequence of functions in $C_b(\mathbb{R}^N)$, it is immediate to extend $\{T(\cdot)\}$ to $B_b(\mathbb{R}^N)$ through (5.10).

To prove that the so extended semigroup is strong Feller, we observe that using the uniform estimates (3.2)(a), it is immediate to check that there exists a positive constant $C = C(t)$ such that

$$[T(t)f]_{\mathrm{Lip}(\mathbb{R}^N)} \le C \|f\|_\infty \quad , \quad t \in]0, +\infty[\ ,$$

for any $f \in C_b(\mathbb{R}^N)$. Of course, the previous estimate can be extended, by dominated convergence, to any $f \in B_b(\mathbb{R}^N)$ and this leads to the assertion.

□

We now introduce the set $W^{k,p}(\mathbb{R}^N, \mu)$ ($k \in \mathbb{N}$, $p \in [1, +\infty[$) of all functions $u \in L^p(\mathbb{R}^N, \mu)$ such that the distributional derivatives of u up to the k-th order are in $L^p(\mathbb{R}^N, \mu)$. It is endowed with the norm $\|u\|^p_{W^{k,p}(\mathbb{R}^N,\mu)} = \sum_{j=0}^k \|D^j u\|_{L^p(\mathbb{R}^N,\mu)}$. Moreover, $W_0^{k,p}(\mathbb{R}^N)$ denotes the closure of $C_c^\infty(\mathbb{R}^N)$ in $W^{k,p}(\mathbb{R}^N)$.

As a consequence of the previous theorem, we can prove the following result.

Proposition 5.6. *The semigroup $\{T(t)\}$ extends to $L^p(\mathbb{R}^N, \mu)$ with a strongly continuous semigroup of contractions for any $p \in [1, +\infty[$. Moreover, there exists a positive constant C such that*

$$\|T(t)f\|_{W^{k,p}(\mathbb{R}^N,\mu)} \le C \|f\|_{W^{k,p}(\mathbb{R}^N,\mu)} \quad , \quad t > 0 \ ,$$

for any $f \in W_0^{k,p}(\mathbb{R}^N, \mu)$ and any $k = 1, 2, 3$.

Proof. Using formula (5.10) and Jensen inequality, it follows that $|T(t)\psi|^p \le T(t)(|\psi|^p)$ for any $t > 0$, any $p \in [1, +\infty[$ and any $\psi \in C_b(\mathbb{R}^N)$. Hence, from (5.9) it is immediate to check that

$$\|T(t)f\|_{L^p(\mathbb{R}^N,\mu)} \le \|f\|_{L^p(\mathbb{R}^N,\mu)} \quad , \quad t > 0 \, , \tag{5.14}$$

for any $f \in C_b(\mathbb{R}^N)$, by Lusin theorem. Since $C_b(\mathbb{R}^N)$ is dense in $L^p(\mathbb{R}^N, \mu)$ for any p as above, (5.14) can be extended to any $f \in L^p(\mathbb{R}^N, \mu)$.

To check that $\{T(t)\}$ is strongly continuous, it suffices to observe that $T(t)f$ tends to f as $t \to 0^+$, by dominated convergence, if $f \in C_b(\mathbb{R}^N)$. Still, by virtue of (5.14), such a convergence can be extended to any $f \in L^p(\mathbb{R}^N, \mu)$.

The last part of the proof can be proved similarly, taking Theorem 5.1 into account.

□

Remark 5.7. In the nondegenerate case, the spaces $W^{k,p}(\mathbb{R}^N, \mu)$ and $W_0^{k,p}(\mathbb{R}^N, \mu)$ coincide. A proof of this property can be obtained in the following way. By a truncation argument, one can easily see that any $u \in W^{k,p}(\mathbb{R}^N, \mu)$ is the limit of a sequence $\{u_n\} \subset W^{k,p}(\mathbb{R}^N, \mu)$ of compactly supported functions. Moreover, as a consequence of the regularity results proved in [5], μ and the Lebesgue measure dx are equivalent and there exists a positive function $\varrho \in L^1(\mathbb{R}^N) \cap C_b(\mathbb{R}^N)$ such that $\mu = \varrho dx$. Therefore, $W^{k,p}(B(0, R), \mu)$ and $W^{k,p}(B(0, R), dx)$ coincide for any $R > 0$. Since $C_c^\infty(\mathbb{R}^N)$ is dense in $W^{k,p}(B(0, R), dx)$, the assertion immediately follows. See, e.g., [4, Chapter 8] for further results on invariant measures in the nondegenerate case.

In the degenerate case, there are, to the best of our knowledge, only a few results on the regularity of the invariant measures and none of them can be applied to the case where $\mathcal{A}$ degenerates identically in $\mathbb{R}^N$. A complete understanding of the regularity of the invariant measures in the degenerate case is still an open problem.

Remark 5.8. The results in Proposition 5.6 show that the L^p-spaces of the invariant measure are the right L^p-spaces where to study elliptic operators with unbounded coefficients. Indeed, it is known also in the nondegenerate case, that, in general, the realizations of the elliptic operators with unbounded coefficients in the usual L^p spaces related to the Lebesgue measure dx have no nice properties. For instance, in [25, Introduction] it has been shown that whatever $\varepsilon > 0$ is fixed, the operator

$$(\mathcal{A}\varphi)(x) = \varphi''(x) - \mathrm{sign}(x)|x|^{1+\varepsilon}\varphi'(x) \quad , \quad x \in \mathbb{R} \, ,$$

does not generate a strongly continuous semigroup in $L^p(\mathbb{R}, dx)$ for any $p \in]1, +\infty[$. On the contrary, it generates semigroups of contractions in $L^p(\mathbb{R}, \mu)$ when μ is the invariant measure.

References

[1] D.G. Aronson & P. Besala, *Parabolic equations with unbounded coefficients*, J. Differential Equations, 3(1967), 1–14.

[2] D.G. Aronson & P. Besala, *Uniqueness of the positive solutions of parabolic equations with unbounded coefficients*, Colloq. Math., 18(1967), 126–135.

[3] S. Bernstein, *Sur la généralisation du probléme de Dirichlet, I*, Math. Ann., 62(1906), 253–271.

[4] M. Bertoldi & L. Lorenzi, *Analytical methods for Markov semigroups*, *Pure and applied mathematics*, 283, Chapman Hall/CRC Press, 2006.

[5] V.I. Bogachev, N.V. Krylov & M. Röckner, *On regularity of transition probabilities and invariant measures of singular diffusions under minimal conditions*, Comm. Partial Differential Equations, (11-12)26(2001), 2037–2080.

[6] G. Da Prato, *Regularity results for some degenerate parabolic equations*, Riv. Mat. Univ. Parma (6), 2*(1999), 245–257.

[7] B. Farkas & L. Lorenzi, *On a class of degenerate elliptic and parabolic operators with unbounded coefficients in* $\mathbb{R}^N$, in preparation.

[8] S. Fornaro, G. Metafune & E. Priola, *Gradient estimates for Dirichlet parabolic problems in unbounded domains*, J. Differential Equations, (2)205(2004), 329–353.

[9] M. Krzyżański & A. Szybiak, *Construction et étude de la solution fondamentale de l'équation linéaire du type parabolique dont le dernier coefficient est non borné. I*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat., (8)27(1959), 26–30.

[10] M. Krzyżański & A. Szybiak, *Construction et étude de la solution fondamentale de l'équation linéaire du type parabolique dont le dernier coefficient est non borné. II*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat., (8)27(1959), 113–117.

[11] M. Krzyżański, *Sur la solution fondamentale de l'équation linéaire normale du type parabolique dont le dernier coefficient est non borné. I*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat., (8)32(1962), 326–330.

[12] M. Krzyżański, *Sur la solution fondamentale de l'équation linéaire normale du type parabolique dont le dernier coefficient est non borné. II*, Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Nat., (8)32(1962), 471–476.

[13] E. Lanconelli & S. Polidoro, *On a class of hypoelliptic evolution operators*, Partial differential equations II, Turin, 1993, Rend. Sem. Mat. Univ. Politec. Torino, (1)52(1994), 29–63.

[14] E. Lanconelli, *On a class of Kolmogorov-Fokker-Planck operators*, Progress in elliptic and parabolic partial differential equations, Capri 1994, Pitman Res. Notes Math. Ser., 350, Longman, Harlow, 1996, 173–183.

[15] E. Lanconelli & A. Pascucci, *On the fundamental solution for hypoelliptic second order partial differential equations with non-negative characteristic form*, Ricerche Mat., (1)48(1999), 81–106.

[16] L. Lorenzi, *Estimates of the derivatives for a class of parabolic degenerate operators with unbounded coefficients in* $\mathbb{R}^N$, Ann. Sc. Norm. Sup. Pisa (5), (2)4(2005), 255–293.

[17] L. Lorenzi, *Schauder estimates for a class of degenerate elliptic and parabolic problems with unbounded coefficients*, Differential Integral Equations, (5)18(2005), 531–566.

[18] A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems*, Birkhäuser, Basel, 1995.

[19] A. Lunardi, *An interpolation method to characterize domains of generators of semigroups*, Semigroup Forum, 53(1996), 321–329.

[20] A. Lunardi, *Schauder estimates for a class of degenerate elliptic and parabolic operators with unbounded coefficients in* $\mathbb{R}$, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), (1)24(1997), 133–164.

[21] G. Metafune, D. Pallara & M. Wacker, *Feller semigroups on* $\mathbb{R}^N$, Semigroup Forum, 65(2002), 159–205.

[22] A. Pascucci, *Hölder regularity for a Kolmogorov equation*, Trans. Amer. Math. Soc., (3)355(2003), 901–924.

[23] S. Polidoro & M.A. Ragusa, *Hölder regularity for solutions of ultraparabolic equations in divergence form*, Potential Anal., (4)14(2001), 341–350.

[24] E. Priola, *Formulae for the derivatives of degenerate diffusion semigroups*, J. Evol. Eq., (4)6(2006), 577–600.

[25] J. Prüss, A. Rhandi & R. Schnaubelt, *The domain of elliptic operators on* $L^p(\mathbb{R}^d)$ *with unbounded drift coefficients*, Houston Journal of Mathematics, to appear.

[26] N. Santier, *Schauder estimates for degenerate elliptic and parabolic equations in* $\mathbb{R}^N$ *with Lipschitz drift*, to appear.