Estimates of the derivatives for a class of parabolic degenerate operators with unbounded coefficients in $\mathbb{R}^N$

LUCA LORENZI

Abstract. We consider a class of perturbations of the degenerate Ornstein-Uhlenbeck operator in $\mathbb{R}^N$. Using a revised version of Bernstein's method we provide several uniform estimates for the semigroup $\{T(t)\}_{t \geq 0}$ associated with the realization of the operator $\mathcal{A}$ in the space of all the bounded and continuous functions in $\mathbb{R}^N$.

Mathematics Subject Classification (2000): 35K65 (primary); 35B65, 47D06 (secondary).

1. Introduction

In the last decades, the interest towards elliptic (and parabolic) operators with unbounded coefficients grew considerably, also in view of their wide applications to stochastic partial differential equations. In the uniformly elliptic case, it is well-known that, under quite minimal regularity assumptions on the coefficients of the operator

$$\mathcal{A} = \sum_{i,j=1}^{N} q_{ij} D_{ij} + \sum_{i=1}^{N} b_j D_j + cu$$

and assuming that c is bounded from above (but without any growth assumptions on the diffusion and the drift coefficients), the Cauchy problem

$$\text{(HCP)} \quad \begin{cases} D_t u(t,x) = \mathcal{A}u(t,x), \ (t,x) \in \mathbb{R}_+ \times \mathbb{R}^N, \\ u(0,x) = f(x), \qquad x \in \mathbb{R}^N, \end{cases}$$

admits a classical solution u, which, in general, is not the unique classical solution to problem (HCP) but, when $f \geq 0$, it is the minimal positive solution. This allows us to associate a semigroup $\{T(t)\}_{t \geq 0}$ of bounded operators in $C_b(\mathbb{R}^N)$ with the operator $\mathcal{A}$: for any $t > 0$ and any $f \geq 0$, $T(t)f$ is the value at t of the minimal

classical solution to (HCP). Such a semigroup is not strongly continuous and, in general, it is not analytic in $C_b(\mathbb{R}^N)$ but it enjoys some properties which are typical of analytic semigroups. For instance, under rather general assumptions on the growth rate of the coefficients at infinity, it has been proved (both by analytic and by probabilistic methods) that the behaviour of the space derivatives of the semigroup with respect to t is very similar to what we can expect when dealing with analytic semigroups. See [3, 4, 14]. The determination of such estimates was the key point also to prove Schauder estimates for the solutions to the elliptic equation

$$(\text{NHE}) \quad \lambda u - \mathcal{A}u = f$$

and to the nonhomogeneous Cauchy problem

$$(\text{NHCP}) \quad \begin{cases} D_t u(t, x) = \mathcal{A}u(t, x) + g(t, x), & (t, x) \in \mathbb{R}_+ \times \mathbb{R}^N, \\ u(0, x) = f(x), & x \in \mathbb{R}^N. \end{cases}$$

This was done in [14] and, recently, the method has been applied in [3] to a wider class of elliptic operators. The degenerate elliptic case is much more difficult to handle and, to the author's knowledge, there are only a few results in the literature. The best known example of a degenerate elliptic operator with unbounded coefficients is the so-called degenerate Ornstein-Uhlenbeck operator which is defined by

$$\mathcal{A}u = \frac{1}{2} \sum_{i,j=1}^{N} q_{ij} D_{ij} u + \sum_{i,j=1}^{N} b_{ij} x_i D_j u, \tag{1.1}$$

where Q is any symmetric non-negative definite matrix, and B is a suitable matrix such that the hypoellipticity condition $\det(Q_t) > 0$ is satisfied at any positive t, where

$$Q_t = \int_0^t e^{sB} Q e^{sB^*} ds, \qquad t > 0.$$

This operator has been deeply studied by A. Lunardi in [14], where she proved that the Cauchy problem (HCP), associated with the operator (1.1), admits a unique classical solution u for any $f \in C_b(\mathbb{R}^N)$. This allowed her to associate a semigroup of linear operators with $\mathcal{A}$, as mentioned above. Further, she gave a precise description of the behaviour of the space derivatives of u near $t = 0$ because an explicit representation formula for the solution to problem (HCP) is available in this particular case. The author obtained the estimates for the space derivatives of u by means of direct computations on this formula.

   As in the non-degenerate case, such estimates are the key point to prove Schauder estimates for the solutions to (NHE) and (NHCP). Since the behaviour near $t = 0$ of the space derivatives of $T(t)f$ is worse than in the non-degenerate case, it was only possible to prove Schauder estimates for the distributional solution in anisotropic Hölder spaces. To prove the existence of a classical solution to

problem (NHCP) one has to assume conditions on g which are much more restrictive than in the non-degenerate case. Using a perturbation argument, the author of [14] was also able to prove similar results in the case when the diffusion matrix $Q = (q_{ij})$ depends on the space variables, its entries vanish when $\max(i, j) > r$, and it converges to some positive definite matrix Q_0 as $|x|$ tends to $+\infty$.

More recently, Da Prato in [5] dealt with the operator $\mathcal{A}$ obtained perturbing the drift term B by a suitable smooth and bounded function $F : \mathbb{R}^N \to \mathbb{R}$. He still assumed the hypoellipticity condition (1.2). The techniques of [14] of course could not be extended to the case when $F \neq 0$, since no explicit representation formulas were available. To overcome such a difficulty, Da Prato took advantage of a revised version of Bernstein's method (see [1]) to obtain the *a priori* estimates for the solution to problem (HCP). In fact, he estimates the behaviour of the first order derivatives of the solution u to (HCP) in terms of the sup norm of f and of the norm of the matrix $\Lambda(t) = Q_t^{-1/2} e^{tB}$, recovering the same estimates proved in [14] in the case where $F = 0$.

Bernstein's method, which seems to best fit for *uniformly* elliptic/parabolic operators, has been successfully carried out very recently also in the degenerate case, in [17], to find *a priori* local in time gradient estimates for solutions to a class of quasilinear degenerate parabolic equations with bounded coefficients in bounded domains.

Here, we consider a class of degenerate elliptic operators of the type

$$\mathcal{A}u(x) = \sum_{i,j=1}^{r} q_{ij}(x)D_{ij}u(x) + \sum_{i,j=1}^{N} b_{ij}x_j D_i u(x), \qquad (1.2)$$

which covers all the cases when the diffusion coefficients are bounded as well as some cases in which they are unbounded. So the main topics to be discussed are:

(i) existence (and uniqueness) of the classical solution to problem (HCP) with $\mathcal{A}$ defined in (1.2);

(ii) uniform estimates (with respect to the x variable) for the space derivatives, up to the third order, of the function $T(t)f$ when f belongs to suitable functional spaces;

(iii) continuity properties of the semigroup $\{T(t)\}_{t \geq 0}$ in $C_b(\mathbb{R}^N)$ and characterization of the domain of its weak generator;

(iv) Schauder-type estimates for the solutions to (NHE) and (NHCP).

In this paper, we deal with point (i) and (ii) whereas in [12] we deal with the remaining points. The main assumptions that we make here on the coefficients are the following:

H1) $N/2 \leq r < N$ and

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

for some function $\nu : \mathbb{R}^N \to \mathbb{R}_+$ such that $\nu_0 := \inf_{x \in \mathbb{R}^N} \nu(x) > 0$;

H2) $q_{ij} \in C_{loc}^{3+\delta}(\mathbb{R}^N)$ $(i, j = 1, \ldots, r)$ for some $\delta \in (0, 1)$ and there exists a positive constant C such that

$$|D^\alpha q_{ij}(x)| \le C|x|^{(1-|\alpha|)^+}\sqrt{v(x)}, \quad x \in \mathbb{R}^N, \quad i, j = 1, \ldots, r, \quad |\alpha| \le 3;$$

H3) the matrix B can be split into blocks as follows

$$B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix},$$

with $B_1 \in L(\mathbb{R}^r)$, $B_2, B_3^* \in L(\mathbb{R}^{N-r}, \mathbb{R}^r)$, $B_4 \in L(\mathbb{R}^{N-r})$ and $\mathrm{rank}(B_3) = N - r$.

We observe that, due to our assumptions, the hypoellipticity condition $\det(Q_t) > 0$ is satisfied at any $x \in \mathbb{R}^N$.

A class of degenerate elliptic operators similar to ours has been considered in [20], where the author deals with the case when the diffusion coefficients are bounded (and they may also depend on t). Under assumptions on the rank of the matrix Q less restrictive than ours, but under stronger assumptions on the matrix B, the author of [20] proves the existence of a fundamental solution to problem $D_t u - \mathcal{A}u = 0$. Then, in [16] the author deals with the Dirichlet-Cauchy problem associated with the operator $\mathcal{A}$ considered in [20], in bounded open sets $\Omega \subset \mathbb{R}^{N+1}$. Under suitable assumptions on the geometry of Ω, Manfredini proves existence and Hölder estimates for the solution to the Dirichlet-Cauchy problem.

We also quote [19], where the author proves interior Schauder estimates for the solution to the parabolic equation $D_t u - \mathcal{A}u = f$ in $(0, +\infty) \times \mathbb{R}^N$, when f is a smooth function and $\mathcal{A}u = \sum_{j=1}^{N-1} D_j^2 u + b D_N u$, b being a smooth function, not necessarily bounded at infinity, satisfying suitable conditions. Such conditions are satisfied, for instance, in the particular case when we take $b(x) = Bx$ and the blocks B_1, B_2 and B_4 of the matrix B identically vanish in $\mathbb{R}$.

Here, under the above set of assumptions, we prove existence and uniqueness of the solution to problem (HCP) associated with (1.2), and, therefore, we define a semigroup of bounded operators $\{T(t)\}_{t \ge 0}$ in $C_b(\mathbb{R}^N)$ as described above. Moreover, we show that, for any $\omega > 0$, there exists a constant $C = C(\omega)$ such that, if $f \in C_b(\mathbb{R}^N)$, then

$$\|D_i T(t)f\|_\infty \le Ce^{\omega t} t^{-(1/2+H(i-r))}\|f\|_\infty, \tag{1.3}$$
$$t > 0, \quad i = 1, \ldots, N,$$

$$\|D_{ij} T(t)f\|_\infty \le Ce^{\omega t} t^{-(1+H(i-r)+H(j-r))}\|f\|_\infty, \tag{1.4}$$
$$t > 0, \quad i, j = 1, \ldots, N,$$

$$\|D_{ijh} T(t)f\|_\infty \le Ce^{\omega t} t^{-(3/2+H(i-r)+H(j-r)+H(h-r))}\|f\|_\infty, \tag{1.5}$$
$$t > 0, \quad i, j, h = 1, \ldots, N,$$

where $H(s) = 0$ if $s \leq 0$ and $H(s) = 1$ if $s > 0$. Furthermore, we prove that the more f is regular, the more the estimates (1.4)-(1.3) can be improved. More precisely, we show that

$$\|T(t)f\|_{C_b^k(\mathbb{R}^N)} \leq Ce^{\omega t}\|f\|_{C_b^k(\mathbb{R}^N)}, \tag{1.6}$$
$$t > 0, \quad k = 1, 2,$$

$$\|D_{ij}T(t)f\|_\infty \leq Ce^{\omega t}t^{-1/2-H(i-r)}\|f\|_{C_b^1(\mathbb{R}^N)}, \tag{1.7}$$
$$t > 0, \quad i \leq j,$$

$$\|D_{ijh}T(t)f\|_\infty \leq Ce^{\omega t}t^{c_{ijk}^k}\|f\|_{C_b^k(\mathbb{R}^N)}, \tag{1.8}$$
$$t > 0, \quad i \leq j \leq h, \quad k = 1, 2, 3,$$

where $c_{ijh}^k = (3-k)/2 + H(i-r) + ((2-k) - (1-k)^+)H(j-r) + (1-k)^+H(h-r)$, and C, ω are as above. As in all the cases considered above, also in this situation the uniform estimates (1.4)-(1.6) are the main ingredients to prove Schauder estimates both for the solutions to problem (NHE) and (NHCP) associated with the operator (1.2). But we stress that they also provide the basic tools to investigate point (iii).

The assumption $r \geq N/2$ is essential to obtain the estimates (1.3)-(1.8). Indeed, when $r < N/2$ and $\mathcal{A}$ is the degenerate Ornstein-Uhlenbeck operator, it is well-known that the behaviour of the derivatives of $T(t)f$ near $t = 0$ is worse than the one in the quoted estimates. To let the reader understand the differences, we consider only the first derivative case, when $r < N/2$. In [14] the author shows that there exists a suitable spitting of the indexes $r + 1, \ldots, N$ into blocks $A_l = \{r + j_l + 1, \ldots, r + j_{l+1}\}$ for $l = 0, \ldots, n$ and some $n \in \mathbb{N}$, such that

$$\|D_iT(t)f\|_\infty \leq Ce^{\omega t}t^{-(3/2+l)}\|f\|_\infty, \qquad t > 0,$$

for any $i \in A_l$, some positive constants C, ω and any $f \in C_b(\mathbb{R}^N)$.

Although we believe that our method can be adapted also to such a situation, as well as to more general degenerate elliptic operators, we prefer to show it in the simplest case, which is however rather technical. We stress that the case treated in this paper covers several interesting situations. For instance, for even N's, we can consider non trivial perturbations of the well-known Kolmogorov operator, which can be obtained taking $r = N/2$, $q_{ij} = 1$, for any $i = j = 1, \ldots, r$, $q_{ij} = 0$ otherwise, and $B_1 = 0$, $B_2 = 0$, $B_4 = 0$ and $B_3 = I$. It is worth stressing also that, in some situations, the operator $D_t - \mathcal{A}$ occurs as a linearization prototype of the Fokker-Plank operator, arising in the study of the Brownian motion of a particle in a fluid. To prove that problem (HCP) is uniquely solvable, we replace the operator $\mathcal{A}$ with the uniformly elliptic operator $\mathcal{A}_\varepsilon$ defined by

$$\mathcal{A}_\varepsilon = \mathcal{A} + \varepsilon \sum_{j=r+1}^N D_{jj}, \tag{1.9}$$

and deal with the Cauchy problem

$$
(\text{HCP}_\varepsilon) \quad
\begin{cases}
D_t u(t, x) = \mathcal{A}_\varepsilon u(t, x), & (t, x) \in \mathbb{R}_+ \times \mathbb{R}^N, \\
u(0, x) = f(x), & x \in \mathbb{R}^N.
\end{cases}
$$

It is well-known, see [18], that under our assumptions on the coefficients q_{ij}, for any $f \in C_b(\mathbb{R}^N)$, problem (HCP_ε) admits a unique classical solution $u = T_\varepsilon(\cdot)f$, where $\{T_\varepsilon(t)\}_{t\geq 0}$ is the semigroup associated with $\mathcal{A}_\varepsilon$. We show that for any $\omega > 0$ there exists a positive constant $C = C(\omega)$, *independent* of ε, such that $T_\varepsilon(t)$ satisfies the estimates (1.3)-(1.8) for any $t > 0$ and any $1 \leq i, j, h \leq N$. This allows us to prove, by an approximation argument and a maximum principle, the following fact: first that problem (HCP) corresponding to (1.2) is uniquely solvable for any $f \in C_b(\mathbb{R}^N)$, and, then, that the family of linear operators $\{T(t)\}_{t\geq 0}$, defined by $T(t)f = \text{“}\lim_{\varepsilon\to 0} T_\varepsilon(t)f\text{”}$ for any $t > 0$ and any $f \in C_b(\mathbb{R}^N)$, gives rise to a semigroup of linear operators in $C_b(\mathbb{R}^N)$ satisfying (1.3)-(1.8). The estimates (1.3)-(1.8) are obtained by adapting the Bernstein method. We stress that our method is not a straightforward generalization of the method used by Da Prato in [5] since his method seems to be not applicable to get uniform estimates for the second and third order space derivatives. The paper is structured as follows. First in Section 2 we introduce the function spaces that we need throughout the paper, and collect some preliminaries. In particular, in Subsection 2.2, we recall some results from [18] on uniformly elliptic operators with unbounded coefficients and we prove some preliminary results which will be used in the following section. In Section 3, the main body of the paper, we construct the semigroup $\{T(t)\}_{t\geq 0}$ and we prove the estimates (1.3)-(1.8). First in Subsection 3.1, we prove that, for any $\varepsilon > 0$, $\{T_\varepsilon(t)\}_{t\geq 0}$ satisfies the estimates (1.3)-(1.8) with constants independent of ε, and then, in Subsection 3.2, we use such estimates to show that problem (HCP) admits a unique classical solution for any $f \in C_b(\mathbb{R}^N)$, and that the semigroup $\{T(t)\}_{t\geq 0}$ satisfies (1.3)-(1.8).

Notation. Throughout the paper, for any $u : \mathbb{R}_+ \times \mathbb{R}^N \to \mathbb{R}$ we indifferently write $u(t, \cdot)$ and $u(t)$ when we want to stress the dependence of u on the time variable t. Moreover, for any smooth real valued function v defined on a domain of $\mathbb{R}^N$, we denote by Dv the gradient of v and by $|Dv(x)|$ its Euclidean norm at x. Similarly, by $D^k v$ ($k \in \mathbb{N}$) we denote the vector consisting of all the k-th order derivatives of v, and by $|D^k v(x)|$ its Euclidean norm at x. By $\mathbb{1}$ we denote the function which is identically equal to 1.

By I_k we denote the identity $k \times k$ matrix. If A is a matrix, we denote by A^* its transpose matrix. When a is a vector we denote by a^T its transpose. For any matrix A we denote by $\|A\|$ its Euclidean norm. For any symmetric matrix A we denote by $\lambda_{\max}(A)$ and by $\lambda_{\min}(A)$, respectively, its maximum and minimum eigenvalues. For any square matrix A we denote by $\text{Tr}(A)$ its trace, i.e. the sum of the elements on the main diagonal. By $L(\mathbb{R}^m, \mathbb{R}^n)$ we denote the set of all the linear operators from $\mathbb{R}^m$ to $\mathbb{R}^n$ (or, equivalently, the set of all the $n \times m$ matrices). When $m = n$ we simply write $L(\mathbb{R}^m)$.

Finally, by $a \vee b$ (resp. $a \wedge b$) we denote the maximum (resp. the minimum) between a and b, and we set $a^+ = a \vee 0$.

2. Function spaces and preliminaries

In this section we both introduce the function spaces we deal with throughout this paper and collect all the preliminary results that we need in what follows. We begin with the following definitions.

Definition 2.1. For any $k \geq 0$ and any open set $\Omega \subset \mathbb{R}^N$ (not necessarily bounded), we denote by $C_b^k(\overline{\Omega})$ the space of all the continuously differentiable up to the $[k]$-order functions $f : \Omega \to \mathbb{R}$ such that $D^\alpha f$ is bounded a continuous in $\overline{\Omega}$ for any $|\alpha| \leq [k]$ ($[k]$ denoting the integer part of k) and $D^\alpha f$ is Hölder continuous of order $k - [k]$ for any $|\alpha| = [k]$. We endow $C_b^k(\overline{\Omega})$ with the Euclidean norm, i.e.

$$\|f\|_{C_b^k(\overline{\Omega})} = \sum_{|\alpha| \leq [k]} \|D^\alpha f\|_\infty + \sum_{|\alpha| = [k]} [D^\alpha f]_{C_b^{k-[k]}(\overline{\Omega})},$$

where $\|D^\alpha f\|_\infty$ denotes the sup-norm of $D^\alpha f$ and $[D^\alpha f]_{C_b^\alpha(\overline{\Omega})} = \sup_{x,y \in \overline{\Omega}, \, x \neq y} |x - y|^{-\alpha}|D^\alpha f(x) - D^\alpha f(y)|$. We say that $u \in C_b^\infty(\overline{\Omega})$ if it belongs to $C_b^k(\overline{\Omega})$ for any $k \geq 0$.

We drop the index "b" when we do not require boundedness of the functions and their derivatives and when Ω is bounded.

Finally, by $C_{loc}^k(\mathbb{R}^N)$ $k \in \mathbb{R}_+ \setminus \mathbb{N}$, we denote the set of all the functions $u : \mathbb{R}^N \to \mathbb{R}$ which belong to $C^k(K)$ for any compact set $K \subset \mathbb{R}^N$.

Definition 2.2. By $C^{1,2}((0, +\infty) \times \mathbb{R}^N)$ we denote 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 $(0, +\infty) \times \mathbb{R}^N$.

For any $\alpha \in (0, 1)$, $C_{loc}^{1+\alpha/2, 2+\alpha}((0, +\infty) \times \mathbb{R}^N)$ is the subset of $C^{1,2}((0, +\infty) \times \mathbb{R}^N)$ of all the functions u such that for any compact set $F \subset (0, +\infty) \times \mathbb{R}^N$, $D_t u, D_x^\beta u$ ($|\beta| \leq 2$) are Hölder continuous of order α in F with respect to the parabolic distance $d((t, x), (s, y)) = (|t - s| + |x - y|^2)^{1/2}$.

Definition 2.3. A function $u : [0, +\infty) \times \mathbb{R}^N \to \mathbb{R}$ is a classical solution to problem (HCP), associated with the operator $\mathcal{A}$ in (1.2), if u is continuous in $[0, +\infty) \times \mathbb{R}^N$, it is continuously differentiable once with respect to time and twice with respect to the space variables in $(0, +\infty) \times \mathbb{R}^N$, and it satisfies the Cauchy problem (HCP).

2.1. General preliminary results

Lemma 2.4. Let A be a $m \times n$ matrix. Then, there exists a $n \times m$ matrix C such that

$$AC + C^* A^*$$

is strictly positive definite if and only if $n \geq m$ and $\mathrm{rank}(A) = m$. Let B be a $m \times n$

matrix. Then, there exists a $n \times m$ matrix C such that

$$CB + B^*C^*$$

is strictly positive definite if and only if $m \geq n$ and $\mathrm{rank}(B) = n$. In such a case, we can take $C = B^*$.

Proof. Suppose that the matrix $AC + C^*A^*$ is strictly positive definite. Then, for any $\xi \in \mathbb{R}^m \setminus \{0\}$ we have

$$0 < \langle (AC + C^*A^*)\xi, \xi \rangle = 2\langle AC\xi, \xi \rangle.$$

Hence, the matrix AC is not singular i.e. its rank equals m. Since $\mathrm{rank}(AC) \leq \min(\mathrm{rank}(A), \mathrm{rank}(C))$ it follows that $n \geq m = \mathrm{rank}(A)$.

Vice versa, let us assume that $n \geq m = \mathrm{rank}(A)$. Moreover, let $D \in L(\mathbb{R}^n)$ be an invertible matrix such that

$$AD = \begin{pmatrix} A_1 & A_2 \end{pmatrix},$$

where $A_1 \in L(\mathbb{R}^m)$ is invertible, $A_2 \in L(\mathbb{R}^{n-m}, \mathbb{R}^m)$. Let $\widetilde{C} \in L(\mathbb{R}^m, \mathbb{R}^n)$ be the matrix defined by

$$\widetilde{C} = \begin{pmatrix} A_1^{-1}K \\ 0 \end{pmatrix},$$

$K \in L(\mathbb{R}^m)$ being any strictly positive definite matrix. We set $C = D\widetilde{C}$ and observe that $AC = AD\widetilde{C} = K$. Hence, $AC + C^*A^* = 2K$ is a positive definite matrix. □

Lemma 2.5. *Let $k, m, n \in \mathbb{N}$ and let $A(t)$ be the $m \times m$ square matrix defined by*

$$A(t) = \begin{pmatrix} A_{11}t^k & A_{12}t^{k+1} & \cdots & A_{1n}t^{k+n-1} \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ A_{1n}^*t^{k+n-1} & A_{2n}^*t^{k+n} & \cdots & A_{nn}t^{k+2n-2} \end{pmatrix}, \qquad t > 0,$$

where $A_{ij} \in L(\mathbb{R}^{m_j}, \mathbb{R}^{m_i})$ ($m_1 + \ldots + m_n = m$) and $A_{ii} = A_{ii}^*$ for any $i = 1, \ldots, n$. Then, $A(t)$ is positive definite for any $t > 0$ if and only if it is positive definite at $t = 1$. In such a case, if for any $\xi \in \mathbb{R}^m$ we split $\xi^T = (\xi_1^T, \cdots, \xi_n^T)$ with $\xi_i \in \mathbb{R}^{m_i}$, we have

$$\langle A(t)\xi, \xi \rangle \geq \lambda_{\min}(A(1)) \sum_{j=1}^{n-1} t^{k+2j} |\xi_j|^2, \qquad t > 0. \tag{2.1}$$

Proof. Of course, if $A(t)$ is strictly positive definite for any $t > 0$, then, in particular, it is strictly positive definite at $t = 1$. Vice versa assume that $A(1)$ is strictly positive definite. For any $t > 0$ and any $\xi = (\xi_1, \dots, \xi_m) \in \mathbb{R}^m$, split as in the statement of the lemma, let

$$\overline{\xi}^T = (t^{k/2}\xi_1^T, t^{k/2+1}\xi_2^T, \dots, t^{k/2+n-1}\xi_n^T).$$

Since $\langle A(t)\xi, \xi \rangle = \langle A(1)\overline{\xi}, \overline{\xi} \rangle$, we deduce that $A(t)$ is strictly positive definite and we get (2.1). $\square$

Lemma 2.6. *Suppose that $Q = (q_{ij})_{i,j=1}^N$ and A are non-negative definite $N \times N$ square matrices. Further, assume that the submatrix $Q_0 = (q_{ij})_{i,j=1}^r$ is strictly positive definite and $q_{ij} = 0$ if $i \vee j > r$. Then*

$$\mathrm{Tr}\,(QA) \geq \lambda_{\min}(Q_0)\mathrm{Tr}\,(A_1),$$

where A_1 is the submatrix obtained from A by erasing the last $N - r$ rows and lines.

Proof. Since Q_0 is strictly positive definite, then there exist an orthogonal $r \times r$ square matrix $B = (b_{ij})$ and a diagonal matrix $\Lambda = \mathrm{diag}(\lambda_1, \dots, \lambda_r)$ such that $Q_0 = B^*\Lambda B$. This implies that the matrix $\widetilde{B} \in L(\mathbb{R}^N)$, defined by $\widetilde{b}_{ij} = b_{ij}$ if $1 \leq i, j \leq r$, $\widetilde{b}_{ij} = \delta_{ij}$ if $i \vee j > r$, is orthogonal and $\widetilde{B}^*Q\widetilde{B} = \widetilde{\Lambda} = \mathrm{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$. Hence,

$$\mathrm{Tr}\,(QA) = \mathrm{Tr}\,(\widetilde{B}^*\widetilde{\Lambda}\widetilde{B}A) = \mathrm{Tr}\,(\widetilde{B}^{-1}\widetilde{\Lambda}\widetilde{B}A) = \mathrm{Tr}\,(\widetilde{\Lambda}\widetilde{B}A\widetilde{B}^{-1}) = \mathrm{Tr}\,(\widetilde{\Lambda}\widetilde{B}A\widetilde{B}^*).$$

We now observe that, since A is positive definite, then $\widetilde{B}A\widetilde{B}^* =: (c_{ij})$ is. This, in particular, implies that $c_{jj} \geq 0$ for any $j = 1, \dots, N$. Therefore,

$$\mathrm{Tr}\,(\Lambda\widetilde{B}A\widetilde{B}^*) = \sum_{j=1}^N \lambda_j c_{jj} = \sum_{j=1}^r \lambda_j c_{jj} \geq \lambda \sum_{j=1}^r c_{jj} = \lambda\mathrm{Tr}\,(BA_1B^*)$$

$$= \lambda\mathrm{Tr}\,(BA_1B^{-1}) = \lambda\mathrm{Tr}\,(A_1),$$

where $\lambda = \lambda_{\min}(Q_0)$, and the assertion follows. $\square$

2.2. Preliminaries on uniformly elliptic operators with unbounded coefficients in $\mathbb{R}^N$

We now recall some basic results on the Cauchy problem

$$(HCP) \quad \begin{cases} D_t u(t,x) = \mathcal{A}u(t,x), & t > 0, \ x \in \mathbb{R}^N, \\ u(0,x) = f(x), & x \in \mathbb{R}^N, \end{cases} \tag{2.2}$$

where $f \in C_b(\mathbb{R}^N)$ and $\mathcal{A}$ is the uniformly elliptic operator defined on the smooth functions φ by

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

which satisfies

$$\sum_{i,j=1}^{N} q_{ij}(x)\xi_i\xi_j \geq \nu(x)|\xi|^2, \qquad x, \xi \in \mathbb{R}^N, \tag{2.3}$$

for some function $\nu : \mathbb{R}^N \to \mathbb{R}$ with $\inf_{x \in \mathbb{R}^N} \nu(x) = \nu_0 > 0$. If we assume that

H1) $q_{ij}, b_j \in C^{\delta}_{\text{loc}}(\mathbb{R}^N)$ for some $\delta \in (0, 1)$ and $q_{ij}(x) = q_{ji}(x)$ for any $i, j = 1, \dots, N$ and any $x \in \mathbb{R}^N$,

it can be shown that the problem (2.2) admits a classical solution

$$u \in C^{1+\delta/2, 2+\delta}_{\text{loc}}((0, +\infty) \times \mathbb{R}^N),$$

which is bounded and continuous in $\overline{\mathbb{R}}_+ \times \mathbb{R}^N$ (see [18, Theorems 4.2 & 4.5]). Without any additional assumption on the coefficients, in general the function u is not the unique classical bounded solution to problem (2.2) (see [18], [2, Chapters 2 and 3] [9, Section 5.2] for examples of nonuniqueness). If we also assume

H2) there exist $\lambda > 0$ and a function $\varphi \in C^2(\mathbb{R}^N)$ such that $\lim_{|x|\to+\infty} \varphi(x) = +\infty$ and

$$\sup_{x \in \mathbb{R}^N} (\mathcal{A}\varphi(x) - \lambda\varphi(x)) < +\infty,$$

then the classical bounded solution to (2.2) is unique, as the following maximum principle shows.

Proposition 2.7. *Suppose that (2.3) and assumptions H1-H2 hold true and let u : $[0, T] \times \mathbb{R}^N \to \mathbb{R}$ ($T > 0$) be a bounded classical solution of the Cauchy problem*

$$\begin{cases} D_t u(t, x) = \mathcal{A}u(t, x) + g(t, x), & (t, x) \in (0, T) \times \mathbb{R}^N, \\ u(0, x) = f(x), & x \in \mathbb{R}^N, \end{cases} \tag{2.4}$$

where $f \in C_b(\mathbb{R}^N)$ and $g \in C((0, T) \times \mathbb{R}^N)$. If $g(t, x) \leq 0$ for any $(t, x) \in (0, T) \times \mathbb{R}^N$, then

$$\sup_{x \in \mathbb{R}^N} u(t, x) \leq \sup_{x \in \mathbb{R}^N} f(x), \qquad t \in [0, T]. \tag{2.5}$$

Similarly, if $g(t, x) \geq 0$ for any $(t, x) \in (0, T) \times \mathbb{R}^N$, then

$$\inf_{x \in \mathbb{R}^N} u(t, x) \geq \inf_{x \in \mathbb{R}^N} f(x), \qquad t \in [0, T]. \tag{2.6}$$

In particular, if $g \equiv 0$, then

$$\|u(t, \cdot)\|_\infty \leq \|f\|_\infty, \qquad t \in [0, T]. \tag{2.7}$$

Proof. The proof is similar to that of [15, Proposition 2.1]. Nevertheless for the reader convenience, and since in the sequel we need to adapt it to the degenerate case, we are forced to give a detailed proof. We restrict ourselves to proving (2.5), since (2.6) can be obtained applying (2.5) to the function $-u$, and (2.7) is a straightforward consequence of (2.5) and (2.6). We first prove that if $\sup_{\mathbb{R}^N} f \leq 0$, then $\sup_{[0,T] \times \mathbb{R}^N} u \leq 0$. Without loss of generality, we can assume that $\sup_{\mathbb{R}^N} (\mathcal{A}\varphi - \lambda\varphi) < 0$. Indeed, if this is not the case, we replace φ with $\varphi + C$ for a suitable constant $C > 0$.

Let $v(t, \cdot) = e^{-\lambda t} u(t, \cdot)$ for any $t \in [0, T]$. A straightforward computation shows that v is a classical bounded solution to the differential equation $D_t v = (\mathcal{A} - \lambda)v + e^{-\lambda t} g$, satisfying $v(0, \cdot) = f$. For any $k \in \mathbb{N}$, let $v_k : [0, T] \times \mathbb{R}^N \to \mathbb{R}$ be defined by

$$v_k(t, x) = v(t, x) - \frac{1}{k}\varphi(x), \qquad (t, x) \in [0, T] \times \mathbb{R}^N.$$

Let us observe that

$$\lim_{k \to +\infty} \sup_{[0,T] \times \mathbb{R}^N} v_k = \sup_{[0,T] \times \mathbb{R}^N} v.$$

Moreover, since v is bounded and $\varphi(x)$ tends to $+\infty$ as $|x|$ tends to $+\infty$, there exists a sequence $\{(t_k, x_k)\}_{k \in \mathbb{N}} \subset [0, T] \times \mathbb{R}^N$ such that $v_k(t_k, x_k) = \sup_{(0,T) \times \mathbb{R}^N} v_k$ for any $k \in \mathbb{N}$. Hence

$$\sup_{[0,T] \times \mathbb{R}^N} v = \lim_{k \to +\infty} v_k(t_k, x_k).$$

If $t_k = 0$ for any k sufficiently large, we are done. Indeed, in such a situation

$$\sup_{[0,T] \times \mathbb{R}^N} v = \lim_{k \to +\infty} v_k(0, x_k) = \lim_{k \to +\infty} \left(f(x_k) - \frac{1}{k}\varphi(x_k) \right) \leq 0,$$

since $f \leq 0$ and φ is bounded from below. It follows that $\sup_{[0,T] \times \mathbb{R}^N} u \leq 0$ as well. So, let us assume that $t_k > 0$ for infinitely many k. Then, we can find out a subsequence $\{(t_{n_k}, x_{n_k})\}_{k \in \mathbb{N}}$ such that $D_t v_{n_k}(t_{n_k}, x_{n_k}) \geq 0$ and $\mathcal{A}v_{n_k}(t_{n_k}, x_{n_k}) \leq 0$. Since $D_t v_k = D_t v$ and $\mathcal{A}v_k = \mathcal{A}v - \frac{1}{k}\mathcal{A}\varphi$, we deduce that

$$\lambda v_{n_k}(t_{n_k}, x_{n_k}) \leq (\lambda + D_t - \mathcal{A}) \, v_{n_k}(t_{n_k}, x_{n_k})$$

$$= (\lambda + D_t - \mathcal{A}) \, v(t_{n_k}, x_{n_k}) + \tfrac{1}{n_k}(\mathcal{A} - \lambda)\varphi(x_{n_k})$$

$$= e^{-\lambda t_{n_k}} g(t_{n_k}, x_{n_k}) + \frac{1}{n_k}(\mathcal{A} - \lambda)\varphi(x_{n_k}),$$

and the last side of the previous chain of inequalities is negative since both g and $(\mathcal{A} - \lambda)\varphi$ are. Hence, $\sup_{[0,T] \times \mathbb{R}^N} v_{n_k} \leq 0$ and as above, letting k go to $+\infty$, we deduce that u is nonpositive in $[0, T] \times \mathbb{R}^N$.

To prove (2.5) in its generality, it suffices to apply the previous result to the function $\widetilde{u} : [0, T] \times \mathbb{R}^N \to \mathbb{R}$ defined by $\widetilde{u}(t, x) = u(t, x) - \sup_{x \in \mathbb{R}^N} f$ for any $(t, x) \in [0, T] \times \mathbb{R}^N$. This concludes the proof. $\qquad\square$

The family of bounded operators $\{T(t)\}_{t \geq 0}$ defined by $T(t)f = u(t, \cdot)$ for any $f \in C_b(\mathbb{R}^N)$ and any $t > 0$, where u is the classical solution to (HCP), gives rise to contractive semigroup of linear operators in $C_b(\mathbb{R}^N)$. A straightforward consequence of (2.6) yields that $\{T(t)\}_{t \geq 0}$ is order preserving, namely

$$f_1, f_2 \in C_b(\mathbb{R}^N), \quad f_1 \leq f_2 \quad \Longrightarrow \quad T(t)f_1 \leq T(t)f_2, \quad t \geq 0. \qquad (2.8)$$

$\{T(t)\}_{t \geq 0}$ is not strongly continuous in $C_b(\mathbb{R}^N)$, and in general, is not analytic neither in $C_b(\mathbb{R}^N)$ nor in $BUC(\mathbb{R}^N)$. Moreover, the following property holds:

If $\{f_n\}_{n \in \mathbb{N}} \subset C_b(\mathbb{R}^N)$ is a bounded sequence such that $\lim_{n \to \infty} f_n = f \in$
$C_b(\mathbb{R}^N)$ uniformly in $B(0, k)$ for any $k > 0$, then $\lim_{n \to \infty} T(\cdot)f_n = T(\cdot)f$ $\quad (2.9)$
uniformly in $[0, T] \times B(0, k)$ for any $T, k > 0$.

We refer the reader to [2, 8, 18] for the proofs of the previous results.

One of the possible methods to construct the classical solution u to (HCP) (and, hence the semigroup $\{T(t)\}_{t \geq 0}$) consists in seeing it as the "limit" (as R tends to $+\infty$) of the solutions u_R to the Dirichlet Cauchy problems

$$\begin{cases} D_t u_R(t, x) = \mathcal{A}u_R(t, x), & t > 0, \ x \in B(0, R), \\ u_R(t, x) = 0, & t > 0, \ x \in \partial B(0, R), \\ u_R(0, x) = \eta_R(x)f(x), & x \in B(0, R), \end{cases}$$

where η_R is any $C_0^\infty(\mathbb{R}^N)$ smooth function such that $\eta_R \equiv 1$ in $B(0, R/2)$ and $\eta_R \equiv 0$ outside $B(0, R)$. The function u_R is defined by $u_R(t, \cdot) = T_R(t)(\eta_R f)$, where $\{T_R(t)\}_{t \geq 0}$ is the (analytic) semigroup associated with the realization in $C(\overline{B(0, R)})$ of the operator $\mathcal{A}$ with homogeneous Dirichlet boundary conditions. This is the approach followed in [3].

We now provide some global estimates for $T(\cdot)f$ and its derivatives up to the third order, when the coefficients of the drift are linear (i.e. we assume that $b_i(x) = \sum_{i,j=1}^N b_{ij}x_j$, for some matrix B) and the diffusion coefficients satisfy suitable regularity and growth assumptions at infinity. Such results will be used in the next section to prove our estimates. The proof that we provide is similar to the one in [3, Theorem 3.3]. Hence, we just sketch it.

Theorem 2.8. *Suppose that the coefficients $q_{ij} = q_{ji} \in C_{\mathrm{loc}}^{k+\delta}(\mathbb{R}^N)$ $(i, j = 1, ..., N)$ for some $k = 0, 1, 2, 3$ are such that $q_{ij} = 0$ if $i \leq r$, $j > r$, $q_{ij} \in C_b^k(\mathbb{R}^N)$*

$(i, j = r+1, \dots, N)$ *and* q_{ij} $(i, j = 1, \dots, r)$ *satisfy assumptions* (2.3) *and*

$$\sum_{i,j=1}^{N-r} q_{i+r\, j+r}(x)\xi_i\xi_j \geq \nu_1|\xi|^2, \qquad \xi \in \mathbb{R}^{N-r}, \tag{2.10}$$

for some positive constant ν_1. *Further assume that*

$$|D^\alpha q_{ij}(x)| \leq C_0|x|^{(1-|\alpha|)^+}\sqrt{\nu(x)}, \quad x \in \mathbb{R}^N, \quad i, j = 1, \dots, r, \quad |\alpha| \leq k, \tag{2.11}$$

for some positive constant C_0, *and that* $b_i(x) = \sum_{j=1}^N b_{ij}x_j$ $(i = 1, \dots, N)$ *for some* $N \times N$ *square matrix* B.

Let $T(\cdot)f$ *be the solution to problem* (2.2) *corresponding to* $f \in C_b^j(\mathbb{R}^N)$ $(j \in \mathbb{N}, j \leq k)$. *Then, for any* $T > 0$ *there exists a positive constant* $\widetilde{C} = \widetilde{C}_T$ *such that*

$$\sum_{j=0}^3 t^{(j-k)^+/2}\|D^jT(t)f\|_\infty \leq \widetilde{C}\|f\|_{C^k(\mathbb{R}^N)}, \qquad t \in (0, T]. \tag{2.12}$$

Remark 2.9. Since q_{ij} $(i, j = 1, \dots, r)$ need to satisfy both (2.10) and (2.11), the q_{ij}'s and ν $(i, j = 1, \dots, r)$ may grow *at most* as $|x|^2$ as $|x|$ tends to $+\infty$.

Remark 2.10. Due to the previous remark, it is immediate to check that if $\mathcal{A}$ is as in the statement of the theorem, then there exist $\lambda > 0$ and a (Lyapunov) function $\varphi \in C^2(\mathbb{R}^N)$ satisfying the assumption H2 of this section. It suffices to take $\varphi(x) = 1 + |x|^2$ for any $x \in \mathbb{R}^N$ and $\lambda = 2Cr^2 + 2\|B\|$. Therefore, for any $f \in C_b(\mathbb{R}^N)$, $T(\cdot)f$ is the unique classical solution to problem (2.4) (with $g = 0$).

Remark 2.11. Since we are assuming different growth condition on the coefficients q_{ij} $(i, j = 1, \dots, N)$, neither the estimates of [15, Theorem 2.4] nor those of [3, Theorem 3.3] may be applied to our situation. Indeed, the quoted estimates to be applied need that the modulus of the derivatives of the diffusion coefficients of $\mathcal{A}$ should be estimated by $C\overline{\nu}(x)$ where, at any $x \in \mathbb{R}^N$, $\overline{\nu}(x)$ is the minimum eigenvalue of the matrix $Q = (q_{ij})$. In our situation $\overline{\nu} = \nu_0 \wedge \nu_1$. Hence, to apply the quoted result the coefficients q_{ij} $(i, j = 1, \dots, N)$ should grow at infinity at most linearly, while our result can be applied also to coefficients which grow faster at infinity. For instance, Theorem 2.8 covers the case where $Q = (q_{ij}) \in L(\mathbb{R}^4)$, $q_{11}(x) = q_{22}(x) = (1/2 + |x|^2)^{3/5} + (1 + |x|^2)^{4/5}$, $q_{12}(x) = q_{21}(x) = -(1/2 + |x|^2)^{3/5} + (1 + |x|^2)^{4/5}$, $q_{ij} = \delta_{ij}, i, j = 3, 4$.

Remark 2.12. In the following section we apply Theorem 2.8 in the particular case where $q_{ij}(x) = \varepsilon\delta_{ij}$ for any $i, j \geq r+1$.

Proof of Theorem 2.8. We restrict ourselves to showing (2.12) in the case where $(k, l) = (0, 3)$, the other cases being similar and even easier. Without loss of generality, we can assume that $1 = \nu_0 := \inf_{x \in \mathbb{R}^N} \nu(x)$.

Let $\varphi \in C_b^\infty(\mathbb{R})$ be a nonincreasing function such that $\varphi(t) = 1$ for any $t \in (-1/2, 1/2)$, $\varphi(t) = 0$ for any $t \in \mathbb{R} \setminus (-1, 1)$. With any $R > 0$ we associate the function $\eta_R : \mathbb{R}^N \to \mathbb{R}$ defined by $\eta_R(x) = \varphi(|x|/R)$. Of course $\eta \in C_b^\infty(\mathbb{R}^N)$, $\eta_R = 1$ in $B(0, R/2)$ and $\eta = 0$ outside the ball $B(0, R)$.

Denote by v the function defined by

$$v_R(t, x) = |u_R(t, x)|^2 + at\eta_R^2(x)|Du_R(t, x)|^2 + a^2 t^2 \eta_R^4(x)|D^2 u_R(t, x)|^2 + a^3 t^3 \eta_R^6(x)|D^3 u_R(t, x)|^2,$$

for any $t \in [0, T]$ and any $x \in B(0, R)$ where $u_R(t, x) = T_R(t)f(x)$. To simplify the notation, in the sequel we simply write (u, v, η) instead of (u_R, v_R, η_R).

Classical results on parabolic equations in bounded domains show that v is continuous in $[0, T] \times \overline{B(0, R)}$ (the continuity of v can also be checked by adapting the proof of the forthcoming Theorem 2.13) and it solves the Cauchy problem

$$\begin{cases} D_t v(t, x) = \mathcal{A}v(t, x) + g(t, x), & t \in [0, T], \ x \in B(0, R), \\ v(t, x) = 0, & t \in [0, T], \ x \in \partial B(0, R), \\ v(0, x) = (f(x))^2, & x \in B(0, R), \end{cases}$$

where $g(t, x) = \sum_{j=1}^5 g_j(t, x)$, for any $t > 0$ and any $x \in B(0, R)$, with

$$g_1 = -2 \sum_{i,j=1}^N q_{ij} D_i u D_j u - 2at\eta^2 \sum_{i,j,h=1}^N q_{ij} D_{ih} u D_{jh} u$$
$$-2a^2 t^2 \eta^4 \sum_{i,j,h,k=1}^N q_{ij} D_{ihk} u D_{jhk} u - 2a^3 t^3 \eta^6 \sum_{i,j,h,k,l=1}^N q_{ij} D_{ihkl} u D_{jhkl} u,$$

$$g_2 = a\eta^2 |Du|^2 + 2a^2 t \eta^4 |D^2 u|^2 + 3a^3 t^2 \eta^6 |D^3 u|^2,$$

$$g_3 = -2at \left(|Du|^2 + 6at\eta^2 |D^2 u|^2 + 15a^2 t^2 \eta^4 |D^3 u|^2 \right) \sum_{i,j=1}^N q_{ij} D_i \eta D_j \eta,$$

$$g_4 = -2at\eta \left(|Du|^2 + 2at\eta^2 |D^2 u|^2 + 3a^2 t^2 \eta^4 |D^3 u|^2 \right) \mathcal{A}\eta$$
$$-8at\eta \sum_{i,j,h=1}^N q_{ij} D_j \eta D_h u D_{ih} u - 16a^2 t^2 \eta^3 \sum_{i,j,h,k=1}^N q_{ij} D_j \eta D_{hk} u D_{ihk} u$$
$$-24a^3 t^3 \eta^5 \sum_{i,j,h,k,l=1}^N q_{ij} D_j \eta D_{hkl} u D_{ihkl} u$$
$$+2at\eta^2 \sum_{j,h=1}^N b_{jh} D_j u D_h u + 4a^2 t^2 \eta^4 \sum_{j,h,k=1}^N b_{jh} D_{jk} u D_{hk} u$$
$$+6a^3 t^3 \eta^6 \sum_{j,h,k,l=1}^N b_{jh} D_{jkl} u D_{hkl} u,$$

$$g_5 = 2at\eta^2 \sum_{i,j,h=1}^{N} D_h q_{ij} D_h u D_{ij} u + 4a^2 t^2 \eta^4 \sum_{i,j,h,k=1}^{N} D_h q_{ij} D_{hk} u D_{ijk} u$$

$$+6a^3 t^3 \eta^6 \sum_{i,j,h,k,l=1}^{N} D_h q_{ij} D_{hkl} u D_{ijkl} u + 2a^2 t^2 \eta^4 \sum_{i,j,h,k=1}^{N} D_{hk} q_{ij} D_{ij} u D_{hk} u$$

$$+6a^3 t^3 \eta^6 \sum_{i,j,h,k,l=1}^{N} D_{hk} q_{ij} D_{ijl} u D_{hkl} u$$

$$+2a^3 t^3 \eta^6 \sum_{i,j,h,k,l=1}^{N} D_{hkl} q_{ij} D_{ij} u D_{hkl} u.$$

Using the ellipticity assumption on q_{ij} we get

$$g_1 \le -2\widetilde{\nu}|Du|^2 - 2at\eta^2 \nu |D_{\#,1}^2 u|^2 - 2at\eta^2 \nu_1 |D_{\#,2}^2 u|^2 - 2a^2 t^2 \eta^4 \nu |D_{\#,1}^3 u|^2$$

$$-2a^2 t^2 \eta^4 \nu_1 |D_{\#,2}^3 u|^2 - 2a^3 t^3 \eta^6 \nu |D_{\#,1}^4 u|^2 - 2a^3 t^3 \eta^6 \nu_1 |D_{\#,2}^4 u|^2, \quad (2.13)$$

where $\widetilde{\nu} = 1 \wedge \nu_1$, and $g_3 \le 0$. Here, to simplify the notation, we have denoted by $D_{\#,1}^k u$ (resp. $D_{\#,2}^k u$) ($k = 2, 3$) the vectors whose entries are the k-order derivatives $\dfrac{\partial^k u}{\partial x_{i_1} \dots \partial x_{i_k}}$ with $i_j \le r$ for some $j = 1, \dots, k$ (resp. $i_j > r$ for any $j = 1, \dots, k$).

To estimate the function g_4 we observe that, by virtue of our assumptions on the coefficients q_{ij}, it can be easily shown that $\nu(x) \le C_0^2 |x|^2$ for any $x \in \mathbb{R}^N$. Therefore, by $H2$, $|q_{ij}(x)| \le C_0^3 |x|^2$ for any $x \in \mathbb{R}^N$ and any $i, j = 1, \dots, N$. Now, a straightforward computation shows that

$$\left| \mathcal{A}\eta(x) \right| \le C_1, \qquad \left| (Q(x) D\eta(x))_i \right| \le C_1 \begin{cases} \sqrt{\nu(x)}, & \text{if } i \le r, \\ 1, & \text{if } i > r, \end{cases} \quad (2.14)$$

for any $x \in \mathbb{R}^N$ and some positive constant C_1. Taking (2.14) into account and recalling that for any $a, b, \varepsilon > 0$ it holds that $ab \le (4\varepsilon)^{-1} a^2 + \varepsilon^2 b^2$, we can now show that

$$g_4 \le 2at \left(C_1 + C_1 \tfrac{N}{\varepsilon} + \|B\| \right) |Du|^2$$

$$+4at\eta^2 \left(2C_1 \varepsilon + at(C_1 + \|B\|) + atC_1 \tfrac{N}{\varepsilon} \right) \left(\nu |D_{\#,1}^2 u|^2 + |D_{\#,2}^2 u|^2 \right)$$

$$+2a^2 t^2 \eta^4 \left(8C_1 \varepsilon + 3at(C_1 + \|B\|) + 3atC_1 \tfrac{N}{\varepsilon} \right) \left(\nu |D_{\#,1}^3 u|^2 + |D_{\#,2}^3 u|^2 \right)$$

$$+ 24a^3 t^3 C_1 \varepsilon \eta^6 \left(\nu |D_{\#,1}^4 u|^2 + |D_{\#,2}^4 u|^2 \right).$$

$$(2.15)$$

The term g_5 can be estimated similarly, taking (2.11) into account, and it gives

$$g_5 \le atC_2\frac{N^2}{2\varepsilon}|Du|^2$$

$$+atC_2N\left\{2\varepsilon + 2atN + at\frac{N}{\varepsilon} + a^2t^2\frac{N^2}{2\varepsilon}\right\}\eta^2\left(\nu|D^2_{\#,1}u|^2 + |D^2_{\#,2}u|^2\right)$$

$$+a^2t^2C_2N\left(4\varepsilon + 3at\frac{N}{2\varepsilon} + 6atN + 2at\varepsilon N\right)\eta^4\left(\nu|D^3_{\#,1}u|^2 + |D^3_{\#,2}u|^2\right)$$

$$+6a^3t^3\eta^6C_2\varepsilon N\left(\nu|D^4_{\#,1}u|^2 + |D^4_{\#,2}u|^2\right),$$

(2.16)

where $C_2 = \max\{C, \|Dq_{ij}\|_\infty, \; i, j = r+1, \ldots, N\}$. From (2.13), (2.15) and (2.16) we get, for any $t \in [0, T]$,

$$g \le \left\{-2\tilde{\nu} + a\left[1 + 2T\|B\| + 2C_3T\left(1 + \frac{N}{\varepsilon} + \frac{N^2}{4\varepsilon}\right)\right]\right\}|Du|^2$$

$$+2a\left\{-\tilde{\nu} + \varepsilon C_3(4 + N) + a\left[1 + 2T\|B\|\right.\right.$$

$$\left.\left.+C_3T\left(2 + \frac{N}{2\varepsilon}(4 + N) + N^2 + aT\frac{N^3}{4\varepsilon}\right)\right]\right\}t\eta^2\left(\nu|D^2_{\#,1}u|^2 + |D^2_{\#,2}u|^2\right)$$

$$+a^2\left\{-2\tilde{\nu} + 4C_3\varepsilon(4 + N) + a\left[3 + 6T\|B\|\right.\right.$$

$$\left.\left.+C_3T\left(6 + 3\frac{N}{2\varepsilon}(4 + N) + 2N^2(3 + \varepsilon)\right)\right]\right\}t^2\eta^4\left(\nu|D^3_{\#,1}u|^2 + |D^3_{\#,2}u|^2\right)$$

$$+2a^3\left[-\tilde{\nu} + 3C_3\varepsilon(2 + N)\right]t^3\eta^6\left(\nu|D^4_{\#,1}u|^2 + |D^4_{\#,2}u|^2\right),$$

where $C_3 := C_1 \vee C_2$. A straightforward computation shows that we can choose (a, ε), independent of R, and such that $g(t, x) \le 0$ for any $t \in [0, T]$ and any $x \in B(0, R)$. The classical maximum principle yields

$$|v(t, x)| \le \|f\|_\infty, \qquad (t, x) \in [0, T] \times \overline{B(0, R)}.$$

Now, (2.12) follows letting R go to $+\infty$. $\qquad\square$

The following theorem guarantees the continuity of the functions $(t, x) \mapsto t^{j/2}(D^jT(t)f)(x)$ at $t = 0$, for any $j = 1, 2, 3$ and any $f \in C_b(\mathbb{R}^N)$. Its proof is based upon property (2.9) and the interior estimates of [7] and [11].

Theorem 2.13. *Under the same assumptions as in Theorem* 2.8*, if $f \in C^k_b(\mathbb{R}^N)$ ($k = 0, 1, 2, 3$), then the function $(t, x) \mapsto t^{(j-k)^+/2}(D^jT(t)f)(x)$ is continuous in $[0, +\infty) \times \mathbb{R}^N$ for any $j = 0, \ldots, 3$. In particular,*

$$\lim_{t\to 0^+} t^{(j-k)^+/2}(D^jT(t)f)(x) = 0, \quad x \in \mathbb{R}^N, \quad j = 0, \ldots, 3. \qquad (2.17)$$

Proof. We begin the proof by checking (2.17) in the case where $k > 0$. We restrict ourselves to dealing with the case where $k = 1$, since the other cases are similar and even simpler.

We use a localization argument. For any $x_0 \in \mathbb{R}^N$, let Ω', Ω be two (sufficiently smooth) bounded open neighborhoods of x_0 such that $\Omega' \subset\subset \Omega$, and let ϑ be any $C_0^\infty(\mathbb{R}^N)$ smooth function compactly supported in Ω and such that $\vartheta \equiv 1$ in Ω'. Since u is a classical solution to problem (2.2), the function $v = \vartheta u$ turns out to be a classical solution to the nonhomogeneous Cauchy problem

$$\begin{cases} D_t v(t, x) = \mathcal{A}v(t, x) + g(t, x), & (t, x) \in \mathbb{R}_+ \times \Omega, \\ v(t, x) = 0, & (t, x) \in \mathbb{R}_+ \times \partial\Omega, \\ v(0, x) = \vartheta(x) f(x), & x \in \Omega, \end{cases}$$

where

$$g(t, x) = -u(t, x)\mathcal{A}\vartheta(x) - 2\sum_{i,j=1}^N q_{ij}(x) D_i u(t, x) D_j \vartheta(x), \quad (t, x) \in \mathbb{R}_+ \times \Omega.$$

Since the realization A_Ω in $C(\overline{\Omega})$ of the operator $\mathcal{A}$, with homogeneous Dirichlet boundary condition, generates an analytic semigroup $\{T_\Omega(t)\}_{t \geq 0}$ (see e.g. [13, Corollary 3.1.21(ii)]), v is given by the usual variation of constants formula

$$v(t, \cdot) = T_\Omega(t)(\vartheta f) + \int_0^t T_\Omega(t - s)g(s, \cdot)ds, \qquad t > 0. \qquad (2.18)$$

Let us estimate separately the two terms in the right-hand side of (2.18). Of course, without loss of generality, we can let t vary in $(0, 1)$. Throughout the rest of the proof, to simplify the notation, we denote by C positive constants, *independent* of $t \in (0, 1)$, which may vary from line to line. To begin with, we consider the function $T_\Omega(\cdot)(\vartheta f)$. From the classical uniform estimates for parabolic problems with regular coefficients in bounded domains, we know that for any $T > 0$ there exists a positive constant $C = C(T)$ such that

$$\|D^j T_\Omega(t)h\|_{C(\overline{\Omega})} \leq \frac{C}{t^{(j-m)^+/2}} \|h\|_{C^m(\overline{\Omega})}, \qquad t \in (0, 1), \quad j = 1, 2, 3, \quad (2.19)$$

for any $h \in C^m(\overline{\Omega})$ ($m = 0, 1, 2$). By [10, Theorem 5.2], if $h \in C_b^\infty(\mathbb{R}^N)$ has compact support in Ω, then the function $T_\Omega(\cdot)h$ and its space derivatives up to the third order are continuous up to $t = 0$. In particular, $(DT_\Omega(t)h)_{|t=0} = Dh$. Since $\vartheta f \in C_b^1(\overline{\Omega})$ is compactly supported in Ω, we can determine a sequence $\{h_n\}_{n \in \mathbb{N}} \subset C_b^\infty(\overline{\Omega})$ of smooth functions, with compact support in Ω, converging to ϑf in $C^1(\overline{\Omega})$. Applying (2.19) with $h = h_n - \vartheta f$, we deduce that $t^{(j-1)/2}D^j T_\Omega(\cdot)h_n$ ($j = 1, 2, 3$) converges uniformly in $[0, T] \times \overline{\Omega}$ to $t^{(j-1)/2}D^j T_\Omega(\cdot)(\vartheta f)$. This implies that $t^{(j-1)/2}D^j T_\Omega(\cdot)(\vartheta f)$ ($j = 1, 2, 3$) is continuous in $[0, +\infty) \times \overline{\Omega}$ and $DT_\Omega(t)(\vartheta f)_{|t=0} = D(\vartheta f)$ while, if $j > 1$, $t^{(j-1)/2}D^j T_\Omega(\vartheta f)_{|t=0} = 0$.

Let us now consider the integral term in (2.18). A straightforward computation shows that

$$\|g(t,\cdot)\|_{C^m(\overline{\Omega})} \le C\|u(t,\cdot)\|_{C^{m+1}(\overline{\Omega})}, \qquad t > 0, \qquad m = 0, 1, 2. \qquad (2.20)$$

Hence, taking (2.12) and (2.20) into account, we deduce that

$$\|g(t,\cdot)\|_{C^m(\overline{\Omega})} \le \frac{C}{t^{m/2}}\|f\|_{C_b^1(\mathbb{R}^N)}, \qquad t \in (0,1), \qquad m = 0, 1, 2. \qquad (2.21)$$

Combining (2.19) and (2.21) (with $m = 1$) gives

$$t^{(j-1)/2}\left\|D^j \int_0^t T_\Omega(t-s)g(s,\cdot)ds\right\|_{C(\overline{\Omega})} \le C\|f\|_{C^1(\overline{\Omega})} t^{(j-1)/2}\int_0^t (t-s)^{-(j-1)/2}s^{-1/2}ds$$
$$\le C\sqrt{t}\|f\|_{C^1(\overline{\Omega})},$$

for any $t \in (0,1)$, $j = 1, 2$.

By interpolation, from (2.19) and (2.21) (with $k = 1, 2$) we get

$$\|D^3 T_\Omega(t)h\|_{C(\overline{\Omega})} \le \frac{C}{t^{1-\alpha/2}}\|h\|_{C^{1+\alpha}(\overline{\Omega})}, \quad \|g(t,\cdot)\|_{C^{1+\alpha}(\overline{\Omega})} \le \frac{C}{t^{(1+\alpha)/2}}\|f\|_{C_b^1(\mathbb{R}^N)}, \qquad (2.22)$$

for any $t \in (0,T)$. Now, (2.22) yields

$$t\left\|D^3 \int_0^t T_\Omega(t-s)g(s,\cdot)ds\right\|_{C(\overline{\Omega})} \le C\|f\|_{C^1(\overline{\Omega})} t\int_0^t (t-s)^{-1+\alpha/2}s^{-(1+\alpha)/2}ds$$
$$\le C\sqrt{t}\|f\|_{C^1(\overline{\Omega})},$$

for any $t \in (0,1)$. Summing up, we have shown that the functions $(t,x) \mapsto t^{(j-1)/2}D^j v(t,x)$ are continuous in $[0,1) \times \overline{\Omega}$. Moreover, $Dv(0,\cdot) = \vartheta f$ and $(\sqrt{t}D^2 v(t,\cdot))_{|t=0} = (tD^3 v(t,\cdot))_{|t=0} = 0$. Since $v \equiv u$ in $\mathbb{R}_+ \times \Omega'$ and $\vartheta \equiv 1$ in Ω', we obtain that (2.17) (with $k = 1$) holds true at $x = x_0$, and, by the arbitrariness of x_0, it holds true for any $x \in \mathbb{R}^N$.

To conclude the proof, we now check (2.17) in the case where $k = 0$. For this purpose, we observe that the classical interior estimates of [7, Theorem 3.5] and [11, Exercise 4.5] imply that if $d(\Omega',\Omega) > 1$, then

$$\sup_{(t,x)\in\Omega_1'} t^{1/2}|(DT(t)h)(x)| + \sup_{(t,x)\in\Omega_1'} t|(D^2 T(t)h)(x)| + \sup_{(t,x)\in\Omega_1'} t^{3/2}|(D^3 T(t)h)(x)|$$
$$\le C \sup_{(t,x)\in\Omega_1} |(T(t)h)(x)|, \qquad (2.23)$$

for any $h \in C_b(\mathbb{R}^N)$ and some positive constant $C = C(\Omega,\Omega')$, *independent* of h. Here, to simplify the notation, we set $\Omega_1 = (0,1) \times \Omega$, $\Omega_1' = (0,1) \times \Omega'$. Now, with any $f \in C_b(\mathbb{R}^N)$, we fix a sequence $\{f_n\}_{n\in\mathbb{N}} \in C_b^1(\mathbb{R}^N)$ of smooth

functions converging in a dominated way to f as n tends to $+\infty$. Applying (2.23) with $g = f - f_n$ and taking (2.9) into account, we deduce that the function $(t, x) \mapsto \psi_{j,n}(t, x) := t^{j/2}(D^j T(t) f_n)(x)$ $(j = 1, 2, 3)$ converges uniformly in Ω'_T to the function $(t, x) \mapsto \psi_j(t, x) := t^{j/2}(D^j T(t) f)(x)$. By the previous results we know that each $\psi_{n,j}$ is continuous in $\overline{\Omega}'_T$ and it vanishes at $t = 0$. It follows that ψ_j $(j = 1, 2, 3)$ is continuous in $\overline{\Omega}'_T$ and it vanishes at $t = 0$, as well. Now, (2.17) follows. $\qquad\square$

3. Existence and uniqueness results, and uniform estimates

In this section, the main body of the paper, we show that for any $f \in C_b(\mathbb{R}^N)$ problem (HCP), where $\mathcal{A}$ is given by (1.2), admits a unique classical solution u_f (see Definition 2.3) and that the family of linear operators $\{T(t)\}_{t\geq 0} \in L(C_b(\mathbb{R}^N))$, defined by $T(t)f = u_f(t, \cdot)$ for any $t \geq 0$, gives rise to a semigroup of linear operators satisfying (1.3)-(1.8). Such a semigroup is not strongly continuous and, in general, is not analytic. Moreover, the behaviour of the space derivatives of $T(t)f$ near $t = 0$ differs from the case where the semigroup is associated to a uniformly elliptic operator in $C_b(\mathbb{R}^N)$ (*cf.* Theorems 2.8 and 3.1).

3.1. Uniform estimates for the approximating semigroup $\{T_\varepsilon(t)\}_{t\geq 0}$

We begin this subsection observing that the coefficients of the operator $\mathcal{A}_\varepsilon$ in (1.9) satisfy the assumptions of Theorem 2.8. Therefore, by Remarks 2.9 and 2.10, for any $\varepsilon > 0$, the Cauchy problem (HCP$_\varepsilon$) admits a unique bounded classical solution u_ε. It follows that the semigroup $\{T_\varepsilon(t)\}_{t\geq 0}$ is well-defined and it satisfies (2.12) and (2.17).

Our purpose in this section consists in showing that, for any $\varepsilon > 0$, $\{T_\varepsilon(t)\}_{t\geq 0}$ satisfies the estimates (1.3)-(1.8) with constants *independent* of ε. Such estimates will be fundamental to solve the degenerate Cauchy problem (HCP) and, consequently, to construct the semigroup $\{T(t)\}_{t\geq 0}$. Before going on, let us introduce some notation. At any $x \in \mathbb{R}^N$, we denote by $Q(x) \in L(\mathbb{R}^r)$ the (strictly) positive definite matrix defined by $(Q(x))_{ij} = q_{ij}(x)$ for any $1 \leq i, j \leq r$. Similarly, we denote by $Q_0(x) = (q_{ij}^0(x)) \in L(\mathbb{R}^N)$ the matrix defined as follows: $q_{ij}^0(x) = q_{ij}(x)$ if $i, j \leq r$, $q_{ij}^0(x) = 0$ otherwise. Moreover, by $Q_\varepsilon(x)$ we denote the matrix differing from $Q_0(x)$ only in the diagonal elements $q_{jj}^\varepsilon(x)$ $(j > r)$ where $q_{jj}^\varepsilon(x) = \varepsilon$. Using this notation, we can write

$$\mathcal{A}u(x) \;\; = \operatorname{Tr}(Q(x)D^2 u(x)) + \langle Bx, Du(x)\rangle,$$

$$\mathcal{A}_\varepsilon u(x) = \operatorname{Tr}(Q_\varepsilon(x)D^2 u(x)) + \langle Bx, Du(x)\rangle,$$

for any $x \in \mathbb{R}^N$ (see (1.2) and (1.9)).

Next, for any integer $k \in \{2, 3, 4\}$, any k-tuple $(i_1, \ldots, i_k)$ with $1 \leq i_1 \leq \ldots \leq i_k \leq N$ and any smooth function w, we set

$$D_{i_1,\ldots,i_k} w = \frac{\partial^k w}{\partial x_{i_1} \ldots \partial x_{i_k}}.$$

Moreover, we set $D_{*,1} w = (D_1 w, \ldots, D_r w)$ and $D_{*,2} w = (D_{r+1} w, \ldots, D_N w)$. Then, we introduce, instead of the corresponding tensors, the vectors $D_*^k w$ ($k = 2, 3, 4$) consisting of all the derivatives $D_{i_1,\ldots,i_k} w$ ordered as follows: $D_{i_1,\ldots,i_k} w$ precedes $D_{j_1,\ldots,j_k} w$ if $i_l \leq j_l$, for any $l = 1, \ldots, k$, and $i_{l_0} < j_{l_0}$ for some $l_0 \in \{1, \ldots, k\}$, or $\{j_1, \ldots, j_k\}$ contains more indexes $j_l \geq r + 1$ than the set $\{i_1, \ldots, i_k\}$. Finally, we set $D_*^k w^T = ((D_{*,1}^k w)^T, \ldots, (D_{*,k+1}^k w)^T)$, where the vector $D_{*,j}^k w$ contains all the derivatives $D_{i_1,\ldots,i_k}^k w$ with $i_{k+1-j} \leq r < i_{k+2-j}$ (when such inequalities are meaningful).

For instance, if $N = 4$, $k = 3$ and $r = 2$, then

$$D_{*,1}^3 w = (D_{111} w, D_{112} w, D_{122} w, D_{222} w),$$

$$D_{*,2}^3 w = (D_{113} w, D_{114} w, D_{123} w, D_{124} w, D_{223} w, D_{224} w),$$

$$D_{*,3}^3 w = (D_{133} w, D_{134} w, D_{144} w, D_{233} w, D_{234} w, D_{244} w),$$

$$D_{*,4}^3 w = (D_{333} w, D_{334} w, D_{344} w, D_{444} w).$$

Theorem 3.1. *Let $\varepsilon > 0$ and assume that hypothesis H1-H3 are satisfied. Then, for any $k = 1, 2, 3$ and any $\omega > 0$, there exists a positive constant $C = C(\omega)$, independent of ε, such that, for any $\varepsilon > 0$,*

$$\|D_{*,j}^k T_\varepsilon(t) f\|_\infty \leq C e^{\omega t} t^{-(k+2j-2)/2} \|f\|_\infty, \quad t > 0, \quad j = 1, \ldots, k+1, \quad f \in C_b(\mathbb{R}^N). \tag{3.1}$$

Proof. As it has already been mentioned in the Introduction, we use the classical Bernstein's method to prove (3.1). We restrict ourselves to considering the case where $k = 3$, the other cases being similar and even easier. We introduce the function $\xi_\varepsilon : [0, +\infty) \times \mathbb{R}^N \to \mathbb{R}$ defined by

$$\begin{aligned}
\xi_\varepsilon(t, x) = \;& \frac{1}{2} \alpha^3 (u_\varepsilon(t, x))^2 + \langle F(t) D u_\varepsilon(t, x), D u_\varepsilon(t, x) \rangle \\
& + \langle G(t) D_*^2 u_\varepsilon(t, x), D_*^2 u_\varepsilon(t, x) \rangle \\
& + \langle H(t) D_*^3 u_\varepsilon(t, x), D_*^3 u_\varepsilon(t, x) \rangle,
\end{aligned}$$

for any $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^N$, where

$$F(t) = \begin{pmatrix} \alpha t I_r & 4t^2 F_1 \\ 4t^2 F_1 & t^3 I_{N-r} \end{pmatrix}, \quad G(t) = \begin{pmatrix} t^2 I_{n_r^1} & 0 & 0 \\ 0 & \alpha^{-7/16} t^4 I_{r(N-r)} & \alpha^{-4/5} t^5 G_1 \\ 0 & \alpha^{-4/5} t^5 G_1^* & \alpha^{-7/8} t^6 I_{n_{N-r}^1} \end{pmatrix}, \tag{3.2}$$

$$H(t) = \begin{pmatrix} \alpha^{-7/16}t^3 I_{n_r^2} & 0 & 0 & 0 \\ 0 & \alpha^{-7/8}t^5 I_{(N-r)n_r^1} & 0 & 0 \\ 0 & 0 & \alpha^{-1}t^7 I_{rn_{N-r}^1} & \alpha^{-13/12}t^8 H_1 \\ 0 & 0 & \alpha^{-13/12}t^8 H_1^* & \alpha^{-9/8}t^9 I_{n_{N-r}^2} \end{pmatrix}, \quad (3.3)$$

for any $t > 0$. Here F_1 is any matrix such that $B_3 F_1 + F_1^* B_3^*$ is strictly negative definite (this matrix exists by virtue of Lemma 2.4), $-\iota = \lambda_{\max}(B_3 F_1 + F_1^* B_3^*)$; $G_1 \in L(\mathbb{R}^{(N-r)(N-r+1)/2}, \mathbb{R}^{r(N-r)})$, $H_1 \in L(\mathbb{R}^{(N-r)(N-r-1)(N-r-2)/6}, \mathbb{R}^{r(N-r)(N-r-1)/2})$ are suitable matrices (with entries *independent* of α) to be determined later on, as well as the constant α. Finally, $n_m^1 = m(m+1)/2$ and $n_m^2 = m(m+1)(m+2)/6$ for any $m \in \mathbb{N}$.

We require that the matrices $F(t)$, $G(t)$ and $H(t)$ are strictly positive definite for any $t > 0$. By Lemma 2.5 it suffices to assume that $F(1)$, $G(1)$ and $H(1)$ are strictly positive definite and, as a straightforward computation shows, this is the case if we assume that

$$\begin{cases} \iota\alpha - 4\|F_1\|^2 > 0, \\ \alpha^{-7/16} - \alpha^{-4/5}\|G_1\|^2 > 0, \\ \alpha^{-17/8} - \alpha^{-13/12}\|H_1\|^2 > 0. \end{cases} \quad (3.4)$$

Thanks to Theorems 2.8 and 2.13, it is easy to check that the function ξ_ε is a classical solution of the Cauchy problem

$$\begin{cases} D_t\xi_\varepsilon(t, \cdot) = \mathcal{A}_\varepsilon\xi_\varepsilon(t, \cdot) + g_\varepsilon(t, \cdot), \ t > 0, \\ \xi_\varepsilon(0, \cdot) = \frac{1}{2}\alpha^3 f^2, \end{cases}$$

where $g_\varepsilon = \sum_{j=1}^3 g_{j,\varepsilon}$ and

$$\begin{aligned} g_{1,\varepsilon}(t, x) = &-\alpha^3\langle Q_\varepsilon(x)Du_\varepsilon(t, x), Du_\varepsilon(t, x)\rangle \\ &-2\text{Tr}(Q_\varepsilon(x)D^2 u_\varepsilon(t, x)F(t)D^2 u_\varepsilon(t, x)) \\ &-2\sum_{i,j=1}^N q_{ij}^\varepsilon(x)\langle G(t)D_*^2 D_i u_\varepsilon(t, x), D_*^2 D_j u_\varepsilon(t, x)\rangle \\ &-2\sum_{i,j=1}^N q_{ij}^\varepsilon(x)\langle H(t)D_*^3 D_i u_\varepsilon(t, x), D_*^3 D_j u_\varepsilon(t, x)\rangle; \end{aligned}$$

$$\begin{aligned} g_{2,\varepsilon}(t, x) = &\langle F'(t)Du_\varepsilon(t, x), Du_\varepsilon(t, x)\rangle + \langle G'(t)D_*^2 u_\varepsilon(t, x), D_*^2 u_\varepsilon(t, x)\rangle \\ &+\langle H'(t)D_*^3 u_\varepsilon(t, x), D_*^3 u_\varepsilon(t, x)\rangle \\ &+\langle(BF(t) + F(t)B^*)Du_\varepsilon(t, x), Du_\varepsilon(t, x)\rangle \\ &+2\langle G(t)[D_*^2, \langle Bx, D\rangle]u_\varepsilon(t, x), D_*^2 u_\varepsilon(t, x)\rangle \\ &+2\langle H(t)[D_*^3, \langle Bx, D\rangle]u_\varepsilon(t, x), D_*^3 u_\varepsilon(t, x)\rangle; \end{aligned}$$

$$g_{3,\varepsilon}(t, x) = 2 \sum_{i,j=1}^{r} D_{ij} u_\varepsilon(t, x) \langle F(t) Dq_{ij}(x), Du_\varepsilon(t, x) \rangle$$

$$+2 \sum_{i,j=1}^{r} \langle G(t)[D_*^2, q_{ij}(x) D_{ij}] u_\varepsilon(t, x), D_*^2 u_\varepsilon(t, x) \rangle$$

$$+2 \sum_{i,j=1}^{r} \langle H(t)[D_*^3, q_{ij}(x) D_{ij}] u_\varepsilon(t, x), D_*^3 u_\varepsilon(t, x) \rangle.$$

Here by $[A, B]$ we have denoted the commutator between the operators A and B.

Our aim consists in showing that there exists $T_0 > 0$, *independent* of ε, such that $g_\varepsilon \leq 0$ in $(0, T_0] \times \mathbb{R}^N$. The maximum principle in Proposition 2.7 (see (2.5)), will then imply that

$$\xi_\varepsilon(t, x) \leq \tfrac{1}{2} \alpha^3 (f(t, x))^2, \qquad (t, x) \in (0, T_0] \times \mathbb{R}^N. \tag{3.5}$$

Since we are assuming that the matrices $F(t)$, $G(t)$ and $H(t)$ are strictly positive definite for any $t > 0$, then, by (3.5), all the terms in the definition of ξ_ε will turn out to be bounded by $\tfrac{1}{2} \alpha^3 f^2$ in $(0, T_0] \times \mathbb{R}^N$ and, consequently, Lemma 2.5 will imply that the estimate (3.1) is satisfied in $(0, T_0]$ with $\omega = 0$ and C replaced with a new constant C_1. The semigroup rule then will allow us to extend the previous estimate to all the positive times. Indeed, consider, to fix the ideas, the case when $k = 1$. For any $\omega > 0$, let $C_\omega > C_1$ be such that

$$C_\omega t^{-(1/2 + H(i-r))} e^{\omega t} \geq C_1, \qquad t \geq T_0, \quad i = 1, \ldots, N.$$

Then, for any $t > T_0$, split $u_\varepsilon(t, \cdot) = T_\varepsilon(t) f = T_\varepsilon(T_0) T_\varepsilon(t - T_0) f$. Since, as it has been pointed out in Subsection 2.2, $\{T_\varepsilon(t)\}_{t \geq 0}$ is a semigroup of contractions, then

$$\|D_i T_\varepsilon(t) f\|_\infty = \|D_i T_\varepsilon(T_0) T_\varepsilon(t - T_0) f\|_\infty$$
$$\leq C_1 T_0^{-1/2 - H(i-r)} \|T_\varepsilon(t - T_0) f\|_\infty \leq C_\omega e^{\omega t} \|f\|_\infty,$$

for any $\varepsilon > 0$, and (3.1) (with $k = 1$) follows with $C = C_\omega$.

So, let us prove that the parameter α and the matrices G_1 and H_1 can be fixed such that $g_\varepsilon \leq 0$ in $(0, T_0] \times \mathbb{R}^N$ for some $T_0 > 0$, *independent* of ε. For this purpose, we begin by observing that, since the matrices $F(t)$, $(\langle G(t) D_*^2 D_i u_\varepsilon(t), D_*^2 D_j u_\varepsilon(t) \rangle)_{ij}$ and $(\langle H(t) D_*^3 D_i u_\varepsilon(t), D_*^3 D_j u_\varepsilon(t) \rangle)_{ij}$ are strictly positive definite, then

$$g_{1,\varepsilon}(t, x) \leq g_{1,0}(t, x), \qquad (t, x) \in \mathbb{R}_+ \times \mathbb{R}^N,$$

where $g_{1,0}$ is obtained from $g_{1,\varepsilon}$ replacing, at any $x \in \mathbb{R}^N$, the matrix $Q_\varepsilon(x)$ with

$Q_0(x)$. Now, Lemma 2.6 implies that

$$g_{1,0}(t) \leq -\alpha^3 \nu |D_{*,1}u_\varepsilon(t)|^2 - 2\nu \sum_{j=1}^{r}(D^2 u_\varepsilon(t) F(t) D^2 u_\varepsilon(t))_{jj}$$

$$-2\nu \sum_{j=1}^{r}\langle G(t) D_*^2 D_j u_\varepsilon(t), D_*^2 D_j u_\varepsilon(t)\rangle$$

$$-2\nu \sum_{j=1}^{r}\langle H(t) D_*^3 D_j u_\varepsilon(t), D_*^3 D_j u_\varepsilon(t)\rangle.$$

Hence, using properly the inequality

$$2\alpha^\gamma t^\beta ab \leq \alpha^{\gamma_1} t^{\beta_1} a^2 + \alpha^{\gamma_2} t^{\beta_2} b^2 \quad (2\gamma = \gamma_1 + \gamma_2, \ 2\beta = \beta_1 + \beta_2), \qquad (3.6)$$

holding for any $\alpha, a, b, t > 0$, it follows that

$$g_{1,0}(t) \leq -\alpha^3 \nu |D_{*,1}u_\varepsilon(t)|^2 - 2\alpha t \nu \langle K_1 D_{*,1}^2 u_\varepsilon(t), D_{*,1}^2 u_\varepsilon(t)\rangle - 2\iota \nu t^3 |D_{*,2}^2 u_\varepsilon(t)|^2$$

$$+\|K_2\|\nu(\alpha^{1/2} t |D_{*,1}^2 u_\varepsilon(t)|^2 + \alpha^{-1/2} t^3 |D_{*,2}^2 u_\varepsilon(t)|^2)$$

$$-2t^2 \nu \langle K_3 D_{*,1}^3 u_\varepsilon(t), D_{*,1}^3 u_\varepsilon(t)\rangle - 2\alpha^{-7/16} t^4 \nu \langle K_4 D_{*,2}^3 u_\varepsilon(t), D_{*,2}^3 u_\varepsilon(t)\rangle$$

$$-2\alpha^{-7/8} t^6 \nu |D_{*,3}^3 u_\varepsilon(t)|^2$$

$$+\|K_5\|\nu(\alpha^{-3/5} t^4 |D_{*,2}^3 u_\varepsilon(t)|^2 + \alpha^{-1} t^6 |D_{*,3}^3 u_\varepsilon(t)|^2)$$

$$-2\alpha^{-7/16} t^3 \nu \langle K_6 D_{*,1}^4 u_\varepsilon(t), D_{*,1}^4 u_\varepsilon(t)\rangle$$

$$-2\alpha^{-7/8} t^5 \nu \langle K_7 D_{*,2}^4 u_\varepsilon(t), D_{*,2}^4 u_\varepsilon(t)\rangle$$

$$-2\alpha^{-1} t^7 \nu \langle K_8 D_{*,3}^4 u_\varepsilon(t), D_{*,3}^4 u_\varepsilon(t)\rangle - 2\alpha^{-9/8} t^9 \nu |D_{*,4}^4 u_\varepsilon(t)|^2$$

$$+\|K_9\|\nu(\alpha^{-49/48} t^7 |D_{*,3}^4 u_\varepsilon(t)|^2 + \alpha^{-55/48} t^9 |D_{*,4}^4 u_\varepsilon(t)|^2)$$

$$\leq -\alpha^3 \nu |D_{*,1}u_\varepsilon(t)|^2 + (-2\alpha + \alpha^{1/2}\|K_2\|)t\nu|D_{*,1}^2 u_\varepsilon(t)|^2$$

$$+(-2\iota + \alpha^{-1/2}\|K_2\|)\nu t^3 |D_{*,2}^2 u_\varepsilon(t)|^2 - 2t^2 \nu |D_{*,1}^3 u_\varepsilon(t)|^2$$

$$+(-2\alpha^{-7/16} + \alpha^{-3/5}\|K_5\|)\nu t^4 |D_{*,2}^3 u_\varepsilon(t)|^2$$

$$+(-2\alpha^{-7/8} + \alpha^{-1}\|K_5\|)\nu t^6 |D_{*,3}^3 u_\varepsilon(t)|^2$$

$$-2\alpha^{-7/16} t^3 \nu |D_{*,1}^4 u_\varepsilon(t)|^2 - 2\alpha^{-7/8} t^5 \nu |D_{*,2}^4 u_\varepsilon(t)|^2$$

$$+(-2\alpha^{-1} + \alpha^{-49/48}\|K_9\|)\nu t^7 |D_{*,3}^4 u_\varepsilon(t)|^2$$

$$+(-2\alpha^{-9/8} + \alpha^{-55/48}\|K_9\|)\nu t^9 |D_{*,4}^4 u_\varepsilon(t)|^2,$$

$$(3.7)$$

where $K_1, K_3, K_4, K_6, K_7, K_8$ are suitable diagonal matrices whose minimum eigenvalue is 1, whereas the entries of the matrices K_2, K_5 and K_9 depend linearly only on the entries of F_1, G_1 and H_1, respectively. In particular, all the previous matrices are *independent* of α.

Throughout the rest of the proof, to simplify the notation, we denote by $o_t(t^k)$ ($k \in \mathbb{R}_+ \cup \{0\}$) any function h, depending on t, and possibly on α, $\|G_1\|$, $\|H_1\|$, but *independent* of x, such that $\lim_{t \to 0} t^{-k} h(t) = 0$. Moreover, in the estimates for $g_{2,\varepsilon}$ and $g_{3,\varepsilon}$ we write explicitly only the terms which are not negligible, as t tends to 0, with respect to the terms in the right-hand side of (3.7), and we use the notation now introduced to denote all the other ones. For instance, if a term is negligible with respect to $-2\iota \nu t^3 |D_{*,2}^2 u_\varepsilon(t)|^2$, we simply denote it by $o_t(t^3)\nu|D_{*,2}^2 u_\varepsilon(t)|^2$, or by $o_t(t^3)|D_{*,2}^2 u_\varepsilon(t)|^2$, if it is independent of ν.

In order to estimate the function $g_{2,\varepsilon}$ we observe that

$$[D_*^2, \langle Bx, D\rangle]u_\varepsilon(t,x) = \mathcal{L}D_*^2 u_\varepsilon(t,x), \quad [D_*^3, \langle Bx, D\rangle]u_\varepsilon(t,x) = \mathcal{M}D_*^3 u_\varepsilon(t,x),$$
$$x \in \mathbb{R}^N,$$

where

$$\mathcal{L} = \begin{pmatrix} L_1 & L_2 & 0 \\ L_3 & L_4 & L_5 \\ 0 & L_6 & L_7 \end{pmatrix}, \qquad \mathcal{M} = \begin{pmatrix} M_1 & M_2 & 0 & 0 \\ M_3 & M_4 & M_5 & 0 \\ 0 & M_6 & M_7 & M_8 \\ 0 & 0 & M_9 & M_{10} \end{pmatrix}, \tag{3.8}$$

L_j ($j = 1, \dots, 7$) and M_j ($j = 1, \dots, 10$) being suitable matrices whose entries linearly depend on the entries of B, but are *independent* of α, G_1, H_1. Using properly inequality (3.6), it follows that

$$2\langle G(t)[D_*^2, \langle Bx, D\rangle]u_\varepsilon(t), D_*^2 u_\varepsilon(t)\rangle$$
$$\leq 2t^2 \|L_1\| |D_{*,1}^2 u_\varepsilon(t)|^2 + \|L_2\|(\alpha^{1/2}t|D_{*,1}^2 u_\varepsilon(t)|^2 + \alpha^{-1/2}t^3|D_{*,2}^2 u_\varepsilon(t)|^2)$$
$$+\alpha^{-4/5}\|G_1^*\|\|L_3\| \left(t^2|D_{*,1}^2 u_\varepsilon(t)|^2 + t^8|D_{*,3}^2 u_\varepsilon(t)|^2\right)$$
$$+\alpha^{-7/16}\|L_3\| \left(t^2|D_{*,1}^2 u_\varepsilon(t)|^2 + t^6|D_{*,2}^2 u_\varepsilon(t)|^2\right)$$
$$+2\alpha^{-7/16}t^4\|L_4\||D_{*,2}^2 u_\varepsilon(t)|^2 + 2\alpha^{-4/5}t^5\|G_1\|\|L_6\||D_{*,2}^2 u_\varepsilon(t)|^2$$
$$+\|L_5\|(\alpha^{-1/24}t^3|D_{*,2}^2 u_\varepsilon(t)|^2 + \alpha^{-5/6}t^5|D_{*,3}^2 u_\varepsilon(t)|^2)$$
$$+\alpha^{-4/5}\|G_1\| (\|L_4\| + \|L_7\|) \left(t^4|D_{*,2}^2 u_\varepsilon(t)|^2 + t^6|D_{*,3}^2 u_\varepsilon(t)|^2\right)$$
$$+\alpha^{-7/8}t^6\|L_6\| \left(|D_{*,2}^2 u_\varepsilon(t)|^2 + |D_{*,3}^2 u_\varepsilon(t)|^2\right)$$
$$+\alpha^{-4/5}\lambda_{\max}(G_1^*L_5 + L_5^*G_1)t^5|D_{*,3}^2 u_\varepsilon(t)|^2 + 2\alpha^{-7/8}t^6\|L_7\||D_{*,3}^2 u_\varepsilon(t)|^2$$
$$= \{\alpha^{1/2}\|L_2\| + o_t(1)\}t|D_{*,1}^2 u_\varepsilon(t)|^2$$
$$+\{\alpha^{-1/2}\|L_2\| + \alpha^{-1/24}\|L_5\| + o_t(1)\}t^3|D_{*,2}^2 u_\varepsilon(t)|^2$$
$$+\{\alpha^{-4/5}\lambda_{\max}(G_1^*L_5 + L_5^*G_1) + \alpha^{-5/6}\|L_5\| + o_t(1)\}t^5|D_{*,3}^2 u_\varepsilon(t)|^2,$$
$$\tag{3.9}$$

and, similarly,

$$2\langle H(t)[D_*^3, \langle Bx, D\rangle]u_\varepsilon(t), D_*^3 u_\varepsilon(t)\rangle$$

$$\leq \|M_2\|(\alpha^{-1/8}t^2|D_{*,1}^3 u_\varepsilon(t)|^2 + \alpha^{-3/4}t^4|D_{*,2}^3 u_\varepsilon(t)|^2)$$

$$+\|M_5\|(\alpha^{-1/2}t^4|D_{*,2}^3 u_\varepsilon(t)|^2 + \alpha^{-5/4}t^6|D_{*,3}^3 u_\varepsilon(t)|^2)$$

$$+\|M_8\|(\alpha^{-43/48}t^6|D_{*,3}^3 u_\varepsilon(t)|^2 + \alpha^{-53/48}t^8|D_{*,4}^3 u_\varepsilon(t)|^2)$$

$$+\alpha^{-13/12}t^8\lambda_{\max}(H_1^* M_8 + M_8^* H_1)|D_{*,4}^3 u_\varepsilon(t)|^2 + o_t(t^2)|D_{*,1}^3 u_\varepsilon(t)|^2$$

$$+o_t(t^4)|D_{*,2}^3 u_\varepsilon(t)|^2 + o_t(t^6)|D_{*,3}^3 u_\varepsilon(t)|^2 + o(t^8)|D_{*,4}^3 u_\varepsilon(t)|^2$$

$$= \{\alpha^{-1/8}\|M_2\| + o_t(1)\}|D_{*,1}^3 u_\varepsilon(t)|^2$$

$$+\{\|M_2\|\alpha^{-3/4} + \|M_5\|\alpha^{-1/2} + o_t(1)\}t^2|D_{*,2}^3 u_\varepsilon(t)|^2$$

$$+\{\|M_5\|\alpha^{-5/4} + \|M_8\|\alpha^{-53/48} + o_t(1)\}t^6|D_{*,3}^3 u_\varepsilon(t)|^2$$

$$+\{\|M_8\|\alpha^{-53/48} + \alpha^{-13/12}\lambda_{\max}(H_1^* M_8 + M_8^* H_1) + o_t(1)\}t^8|D_{*,4}^3 u_\varepsilon(t)|^2. \tag{3.10}$$

Therefore, from (3.9) and (3.10) and observing that

$$\langle (4(B_3 F_1 + F_1 B_3^*)D_{*,2}u_\varepsilon(t), D_{*,2}u_\varepsilon(t)\rangle \leq -4\iota|D_{*,2}u_\varepsilon(t)|^2, \qquad t > 0,$$

we get

$$g_{2,\varepsilon}(t) \leq a_1(t)|D_{*,1}u_\varepsilon(t)|^2 + a_2(t)|D_{*,2}u_\varepsilon(t)|^2 + a_3(t)|D_{*,1}^2 u_\varepsilon(t)|^2$$

$$+ a_4(t)|D_{*,2}^2 u_\varepsilon(t)|^2 + a_5(t)|D_{*,3}^2 u_\varepsilon(t)|^2 + a_6(t)|D_{*,1}^3 u_\varepsilon(t)|^2 \tag{3.11}$$

$$+ a_7(t)|D_{*,2}^3 u_\varepsilon(t)|^2 + a_8(t)|D_{*,3}^3 u_\varepsilon(t)|^2 + a_9(t)|D_{*,4}^3 u_\varepsilon(t)|^2,$$

where

$$a_1(t) = \alpha + 8\|\tilde{C}\|\alpha^{1/2} + \|B_3\|\alpha^{5/2} + o_t(1),$$

$$a_2(t) = \{-\iota + (8\|F_1\| + \|B_3\|)\alpha^{-1/2} + o_t(1)\}t^2,$$

$$a_3(t) = \{4 + \alpha^{1/2}\|L_2\| + o_t(1)\}t,$$

$$a_4(t) = \{4\alpha^{-7/16} + \alpha^{-1/2}\|L_2\| + \alpha^{-1/24}\|L_5\| + o_t(1)\}t^3,$$

$$a_5(t) = \{6\alpha^{-7/8} + \alpha^{-4/5}\lambda_{\max}(G_1^* L_5 + L_5^* G_1) + \alpha^{-5/6}\|L_5\| + o_t(1)\}t^5,$$

$$a_6(t) = \{15\alpha^{-7/16} + \alpha^{-1/8}\|M_2\| + o_t(1)\}t^2,$$

$$a_7(t) = \{10\alpha^{-7/8} + \|M_2\|\alpha^{-3/4} + \|M_5\|\alpha^{-1/2} + o_t(1)\}t^4,$$

$$a_8(t) = \{7\alpha^{-1} + \|M_5\|\alpha^{-5/4} + \|M_8\|\alpha^{-53/48} + o_t(1)\}t^6,$$

$$a_9(t) = \{9\alpha^{-9/8} + \|M_8\|\alpha^{-53/48} + \alpha^{-13/12}\lambda_{\max}(H_1^* M_8 + M_8^* H_1) + o_t(1)\}t^8.$$

Finally, we consider the function $g_{3,\varepsilon}$ and we observe that

$$[D_*^2, q_{ij}(x)D_{ij}]u_\varepsilon(t,x) = D_{ij}u_\varepsilon(t,x)D_*^2 q_{ij}(x) + \mathcal{N}D_*^3 u_\varepsilon(t,x),$$

$$[D_*^3, q_{ij}(x)D_{ij}]u_\varepsilon(t,x) = D_{ij}u_\varepsilon(t,x)D_*^3 q_{ij}(x) + \mathcal{R}(x)D_*^3 u_\varepsilon(t,x)$$

$$+\mathcal{S}(x)D_*^4 u_\varepsilon(t,x),$$

where, for any $x \in \mathbb{R}^N$, the matrices $\mathcal{N}(x) \in L(\mathbb{R}^{n_N^2}, \mathbb{R}^{n_N^1})$, $\mathcal{P}(x) \in L(\mathbb{R}^{n_N^3})$ and $\mathcal{R}(x) \in L(\mathbb{R}^{n_N^3}, \mathbb{R}^{n_N^2})$ ($n_N^3 := N(N+1)(N+2)(N+3)/24$), split according to the splitting of the vectors $D_*^k u_\varepsilon$ ($k = 2, 3, 4$), are given by

$$\mathcal{N}(x) = \begin{pmatrix} N_1(x) & 0 & 0 & 0 \\ N_2(x) & N_3(x) & 0 & 0 \\ 0 & N_4(x) & 0 & 0 \end{pmatrix},$$

$$\mathcal{P}(x) = \begin{pmatrix} P_1(x) & 0 & 0 & 0 \\ P_2(x) & P_3(x) & 0 & 0 \\ P_4(x) & P_5(x) & 0 & 0 \\ 0 & P_6(x) & 0 & 0 \end{pmatrix}, \qquad \mathcal{R}(x) = \begin{pmatrix} R_1(x) & 0 & 0 & 0 & 0 \\ R_2(x) & R_3(x) & 0 & 0 & 0 \\ 0 & R_4(x) & R_5(x) & 0 & 0 \\ 0 & 0 & R_6(x) & 0 & 0 \end{pmatrix},$$

for any $x \in \mathbb{R}^N$, the entries of the matrices $N_j(x)$ ($j = 1, \dots, 4$), $P_j(x)$, $R_j(x)$ ($j = 1, \dots, 6$) being linear combinations of the entries of the derivatives of the diffusion coefficients $q_{ij}(x)$. In particular, they are independent of α, G_1, H_1 and there exists a positive constant C_2, *independent* of x, such that

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

Hence, using properly the inequality (3.6) (where now a and b are given, respectively, by $\sqrt{v}|D_{*,i}^k u_\varepsilon|$ and $|D_{*,j}^m u_\varepsilon|$ for suitable i, j, k, m), and taking assumption *H2* into account, it is easy to check that

$$\begin{aligned} g_{3,\varepsilon}(t) \le\ & o_t(1)|D_{*,1}u_\varepsilon(t)|^2 + o_t(t^2)|D_{*,2}u_\varepsilon(t)|^2 \\ & + o_t(t)v|D_{*,1}^2 u_\varepsilon(t)|^2 + o_t(t^3)v|D_{*,2}^2 u_\varepsilon(t)|^2 \\ & + o_t(t^5)|D_{*,3}^2 u_\varepsilon(t)|^2 + o_t(t^2)v|D_{*,1}^3 u_\varepsilon(t)|^2 \\ & + o_t(t^4)v|D_{*,2}^3 u_\varepsilon(t)|^2 + o_t(t^6)v|D_{*,3}^3 u_\varepsilon(t)|^2 \\ & + o_t(t^8)|D_{*,4}^3 u_\varepsilon(t)|^2 + o_t(t^3)v|D_{*,1}^4 u_\varepsilon(t)|^2 \\ & + o_t(t^5)v|D_{*,2}^4 u_\varepsilon(t)|^2 + o_t(t^7)v|D_{*,3}^4 u_\varepsilon(t)|^2, \end{aligned} \qquad (3.12)$$

for any $t > 0$. Summing up, from (3.7), (3.11) and (3.12) we deduce that

$$
\begin{aligned}
g_\varepsilon(t) \leq\ & \{-\alpha^3 \nu + o_\alpha(\alpha^3) + o_t(1)\}|D_{*,1}u_\varepsilon(t)|^2 \\
& + \{-\iota + o_\alpha(1) + o_t(1)\}t^2|D_{*,2}u_\varepsilon(t)|^2 \\
& + \{-2\alpha + o_\alpha(\alpha) + o_t(1)\}\nu t|D_{*,1}^2 u_\varepsilon(t)|^2 \\
& + \{-2\iota\nu + o_\alpha(1)\nu + o_t(1)\nu\}t^3|D_{*,2}^2 u_\varepsilon(t)|^2 \\
& + \{\alpha^{-4/5}\lambda_{\max}(G_1^* L_5 + L_5^* G_1) + o_\alpha(\alpha^{-4/5}) + o_t(1)\}t^5|D_{*,3}^2 u_\varepsilon(t)|^2 \\
& + \{-2\nu + o_\alpha(1) + o_t(1)\nu\}t^2|D_{*,1}^3 u_\varepsilon(t)|^2 \\
& + \{-2\alpha^{-7/16} + o_\alpha(\alpha^{-7/16}) + o_t(1)\}\nu t^4|D_{*,2}^3 u_\varepsilon(t)|^2 \\
& + \{-2\alpha^{-7/8} + o_\alpha(\alpha^{-7/8}) + o_t(1)\}\nu t^6|D_{*,3}^3 u_\varepsilon(t)|^2 \\
& + \{\alpha^{-13/12}\lambda_{\max}(H_1^* M_8 + M_8^* H_1) + o_\alpha(\alpha^{-13/12}) + o_t(1)\}t^8|D_{*,4}^3 u_\varepsilon(t)|^2 \\
& + \{-2\alpha^{-7/16}\nu + o_t(1)\nu\}t^3|D_{*,1}^4 u_\varepsilon(t)|^2 \\
& + \{-2\alpha^{-7/8} + o_t(1)\}\nu t^5|D_{*,2}^4 u_\varepsilon(t)|^2 \\
& + \{-2\alpha^{-1} + o_\alpha(\alpha^{-1}) + o_t(1)\}\nu t^7|D_{*,3}^4 u_\varepsilon(t)|^2 \\
& + \{-2\alpha^{-9/8} + o_\alpha(\alpha^{-9/8})\}\nu t^9|D_{*,4}^4 u_\varepsilon(t)|^2, \tag{3.13}
\end{aligned}
$$

where by $o_\alpha(\alpha^{-k})$ $(k \geq 0)$ we have denoted any function $h : \mathbb{R}_+ \to \mathbb{R}$, depending on α, and possibly on G_1 and H_1, but *independent* of t, such that $\lim_{\alpha \to +\infty} \alpha^k h(\alpha) = 0$. To prove that $g_\varepsilon(t, x) \leq 0$ for any t in a right neighborhood of 0 (*independent* of ε) and any $x \in \mathbb{R}^N$, we show that we can fix α, $T_0 > 0$ and the matrices G_1 and H_1 so that all the terms in the right-hand side of (3.13) are negative in $(0, T_0] \times \mathbb{R}^N$. We denote by $\hat{a}_j(x)$ the coefficients which we obtain from the terms in curly brackets in the right-hand side of (3.13), disregarding the terms of type $o_t(1)$. As a first step we prove that we can fix $\alpha > 0$ and the matrices G_1 and H_1 so that

$$
\sup_{x \in \mathbb{R}^N} \hat{a}_j(x) < 0, \qquad j = 1, \dots, 9. \tag{3.14}
$$

Once (3.14) is proved, it will be an easy task to check that, we can fix $T_0 > 0$ such that the right-hand side of (3.13) is negative in $(0, T_0] \times \mathbb{R}^N$. An easy asymptotic analysis shows that all the coefficients $\hat{a}_j$ $(j = 1, \dots, 9)$ satisfy (3.14), for $\alpha > 0$ sufficiently large, provided that G_1 and H_1 can be chosen so that the matrices $G_1^* L_5 + L_5^* G_1$ and $H_1^* M_8 + M_8^* H_1$ are strictly negative definite. By virtue of Lemma 2.4 and assumption $H1$, this is the case if the ranks of the matrices $L_5 \in L(\mathbb{R}^{n_{N-r}^1}, \mathbb{R}^{r(N-r)})$ and $M_8 \in L(\mathbb{R}^{n_{N-r}^2}, \mathbb{R}^{rn_{N-r}^1})$ are, respectively, n_{N-r}^1 and n_{N-r}^2.

Straightforward computations show that, up to rearranging the rows, we can split L_5 and M_8 into blocks (according to the splitting of the vectors $D_*^3 u_\varepsilon$, $k = 3, 4$) as follows:

$$L_5 = S_0, \qquad M_8 = \begin{pmatrix} S_0 & 0 & \cdots & \cdots & & 0 \\ \star & S_1 & 0 & \cdots & & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & 0 \\ \star & \cdots & \cdots & & \star & S_{N-r-1} \end{pmatrix},$$

where $S_j \in L(\mathbb{R}^{n_{N-r-j-1}^1}, \mathbb{R}^{r(N-r-j)})$ is given by

$$D_j = \begin{pmatrix} B_3^j & 0 & \cdots & \cdots & & 0 \\ \star & B_3^{j+1} & 0 & \cdots & & 0 \\ \vdots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & & 0 \\ \star & \cdots & \cdots & & \star & B_3^{N-r-1} \end{pmatrix}, \qquad j = 0, \dots, N - r - 1,$$

B_3^j being obtained from B_3^* by removing the first j columns. Here and above by "$\star$" we denote suitable matrices whose entries depend linearly only on the entries of B_3. Since, by the assumption H3, $\operatorname{rank}(B_3) = N - r$, it follows, due to the particular structure of the matrices L_5 and M_8, that $\operatorname{rank}(L_5) = n_{N-r}^1$ and $\operatorname{rank}(M_8) = n_{N-r}^2$. Hence, from the second part of Lemma 2.4, we deduce that, if we set $G_1 = -L_5$, and $H_1 = -M_8$, then the matrices $G_1^* L_5 + L_5^* G_1$ and $H_1^* M_8 + M_8^* H_1$ are strictly negative definite. Therefore, we can fix $\alpha > 0$ such that (3.14) holds. Moreover, up to choosing a larger α, we can assume that conditions (3.4) are satisfied. Then, fixing T_0 sufficiently small, we obtain that $g_\varepsilon \le 0$ in $(0, T_0] \times \mathbb{R}^N$ and $F(t)$, $G(t)$ and $H(t)$ are strictly positive definite for any $t > 0$. By the above remarks this concludes the proof. $\qquad \square$

As in the non-degenerate case, the more the initial datum is regular, the more we can improve the estimates of the derivatives of $T_\varepsilon(t)$ near $t = 0$. We state this fact in the following theorem.

Theorem 3.2. *Under the same assumptions of Theorem 3.1, for any $k = 2, 3$ and any $\omega > 0$, there exists a positive constant $C = C(\omega)$, independent of ε, such that for any $\varepsilon > 0$*

$$\|D_{*,j}^k T_\varepsilon(t) f\|_\infty \le C e^{\omega t} t^{-(2(j-h-1)^+ + k - h)/2} \|f\|_{C_b^h(\mathbb{R}^N)}, \tag{3.15}$$

$$t > 0, \quad j \le k + 1, \quad h \le k, \quad f \in C_b^h(\mathbb{R}^N).$$

Proof. Since the proof is close to that of Theorem 3.1, we just sketch it, pointing out the main differences. We confine ourselves to proving (3.15) when $k = h = 3$, the other cases being similar and even simpler. Let us introduce the function $\xi_{3,\varepsilon}$: $[0, +\infty) \times \mathbb{R}^N \to \mathbb{R}$ defined by

$$
\begin{aligned}
\xi_{3,\varepsilon}(t, x) = {} & \frac{1}{2}\alpha^3 (u_\varepsilon(t, x))^2 + \langle F\, Du_\varepsilon(t, x), Du_\varepsilon(t, x)\rangle \\
& + \langle G\, D_*^2 u_\varepsilon(t, x), D_*^2 u_\varepsilon(t, x)\rangle + \langle H\, D_*^3 u_\varepsilon(t, x), D_*^3 u_\varepsilon(t, x)\rangle,
\end{aligned}
\tag{3.16}
$$

for any $(t, x) \in \mathbb{R}_+ \times \mathbb{R}^N$, where $F = F(1)$ (with F_1 being replaced with $(\|B_4\| + 1)F_1$), $G = G(1)$ and $H = H(1)$ (see (3.2), (3.3)) with $G_1 = -L_5$, $H_1 = -M_8$ (where L_5 and M_8 are given by (3.8)) and α is to be determined later on. Using Theorem 2.13 we can show that $\xi_{3,\varepsilon}$ is continuous up to $t = 0$. Furthermore, straightforward computations show that $\xi_{3,\varepsilon}$ satisfies

$$
\begin{cases}
D_t \xi_{3,\varepsilon}(t, \cdot) = \mathcal{A}_\varepsilon \xi_{3,\varepsilon}(t, \cdot) + \hat{g}_\varepsilon(t, \cdot), & t > 0, \\
\xi_{3,\varepsilon}(0, \cdot) = \frac{1}{2}\alpha^3 f^2 + \langle F\, Df, Df\rangle + \langle G\, D_*^2 f, D_*^2 f\rangle + \langle H D_*^3 f, D_*^3 f\rangle,
\end{cases}
$$

where $\hat{g}_\varepsilon = \sum_{j=1}^3 \hat{g}_{j,\varepsilon}$. Here, the functions $\hat{g}_{j,\varepsilon}$ ($j = 1,2,3$) are obtained from the functions $g_{j,\varepsilon}$ in the proof of Theorem 3.1 by replacing everywhere $(F(t),G(t),H(t))$ by (F, G, H) and disregarding the terms containing the matrices F', G' and H'.

The proof now follows the same ideas of the proof of Theorem 3.1. The function $\hat{g}_{1,\varepsilon}$ can be estimated by the right-hand side of (3.7) where we set $t = 1$ and $G_1 = -L_5$, $H_1 = -M_8$. As far as the functions $\hat{g}_{2,\varepsilon}$ and $\hat{g}_{3,\varepsilon}$ are concerned, we write explicitly only the terms which are not negligible as α tends to $+\infty$ with respect to the terms in (3.7) and simply write $o_\alpha(\alpha^k)$ ($k \ge 0$) to denote the remaining ones. Hence, using properly inequality (3.6) (with $\beta = 0$) and arguing as in the proof of (3.11) and (3.12), we get

$$
\begin{aligned}
\hat{g}_{2,\varepsilon}(t) + \hat{g}_{3,\varepsilon}(t) \le {} & -\iota|D_{*,2}u_\varepsilon(t)|^2 - 2\alpha^{-4/5}\lambda_{\min}(L_5^* L_5)|D_{*,3}^2 u_\varepsilon(t)|^2 \\
& -2\alpha^{-13/12}\lambda_{\min}(M_8^* M_8)|D_{*,4}^3 u_\varepsilon(t)|^2 + o_\alpha(\alpha^3)|D_{*,1}u_\varepsilon(t)|^2 \\
& + o_\alpha(1)|D_{*,2}u_\varepsilon(t)|^2 + o_\alpha(\alpha)|D_{*,1}^2 u_\varepsilon(t)|^2 \\
& + o_\alpha(1)|D_{*,2}^2 u_\varepsilon(t)|^2 + o_\alpha(1)|D_{*,3}^2 u_\varepsilon(t)|^2 \\
& + o_\alpha(1)|D_{*,1}^3 u_\varepsilon(t)|^2 + o_\alpha(\alpha^{-7/16})|D_{*,2}^3 u_\varepsilon(t)|^2 \\
& + o_\alpha(\alpha^{-7/8})|D_{*,3}^3 u_\varepsilon(t)|^2 + o_\alpha(\alpha^{-13/12})|D_{*,4}^3 u_\varepsilon(t)|^2 \\
& + o_\alpha(\alpha^{-7/16})\nu|D_{*,1}^4 u_\varepsilon(t)|^2 + o_\alpha(\alpha^{-7/8})\nu|D_{*,2}^4 u_\varepsilon(t)|^2 \\
& + o_\alpha(\alpha^{-1})\nu|D_{*,3}^4 u_\varepsilon(t)|^2,
\end{aligned}
\tag{3.17}
$$

for any $t > 0$. Therefore, from (3.7) (with $t = 1$, $G_1 = -L_5$ and $H_1 = -M_8$) and (3.17), we obtain

$$
\begin{aligned}
\hat{g}_\varepsilon(t) \leq & -\{\alpha^3 v + o_\alpha(\alpha^3)\}|D_{*,1}u_\varepsilon(t)|^2 - \{\iota + o_\alpha(1)\}|D_{*,2}u_\varepsilon(t)|^2 \\
& -\{2\alpha + o_\alpha(\alpha)\}v|D^2_{*,1}u_\varepsilon(t)|^2 - \{2\iota + o_\alpha(1)\}v|D^2_{*,2}u_\varepsilon(t)|^2 \\
& -\{2\alpha^{-4/5}\lambda_{\min}(L_5^*L_5) + o_\alpha(\alpha^{-4/5})\}|D^2_{*,3}u_\varepsilon(t)|^2 \\
& -\{2v + o_\alpha(1)\}|D^3_{*,1}u_\varepsilon(t)|^2 \\
& -\{2\alpha^{-7/16} + o_\alpha(\alpha^{-7/16})\}v|D^3_{*,2}u_\varepsilon(t)|^2 \\
& -\{2\alpha^{-7/8} + o_\alpha(\alpha^{-7/8})\}v|D^3_{*,3}u_\varepsilon(t)|^2 \\
& -\{2\alpha^{-13/12}\lambda_{\min}(M_8^*M_8) + o_\alpha(\alpha^{-13/12})\}|D^3_{*,4}u_\varepsilon(t)|^2 \\
& -2\alpha^{-7/16}v|D^4_{*,1}u_\varepsilon(t)|^2 - 2\alpha^{-7/8}vt^5|D^4_{*,2}u_\varepsilon(t)|^2 \\
& -\{2\alpha^{-1} + o_\alpha(\alpha^{-1})\}v|D^4_{*,3}u_\varepsilon(t)|^2 \\
& -\{2\alpha^{-9/8} + o_\alpha(\alpha^{-9/8})\}v|D^4_{*,4}u_\varepsilon(t)|^2. \tag{3.18}
\end{aligned}
$$

Now, from (3.18) and condition (3.4) (where we replace F_1 by $(\|B_4\| + 1)F_1$), it follows that we can fix $\alpha > 0$ such that F, G, H are strictly positive definite and $\hat{g}_\varepsilon \leq 0$ in $\mathbb{R}_+ \times \mathbb{R}^N$, implying that

$$
\|T_\varepsilon(t)f\|_{C^3_b(\mathbb{R}^N)} \leq C_0\|f\|_{C^3_b(\mathbb{R}^N)}, \quad t \in (0, T_0], \quad f \in C^3_b(\mathbb{R}^N), \tag{3.19}
$$

for some positive T_0 and some constant $C = C(T_0) > 0$. The semigroup rule, then allows us to extend (3.19) to all the positive times obtaining (3.15). Indeed, for any $t > T_0$, it suffices to split $T(t)f = T(T_0/2)T(t - T_0/2)f$ observing that (3.1) and (3.19) give

$$
\|T_\varepsilon(t)f\|_{C^3_b(\mathbb{R}^N)} \leq C_0\|T(t - T_0/2)f\|_{C^3_b(\mathbb{R}^N)} \leq C_1 e^{\omega t}\|f\|_{C^3_b(\mathbb{R}^N)},
$$

for any $\omega > 0$ and some positive constant $C_1 = C_1(\omega)$. To prove (3.15), with $k = 1, 2$, we apply the previous arguments to the functions $\xi_{k,\varepsilon}$ defined by

$$
\begin{aligned}
\xi_{k,\varepsilon}(t, x) = & \frac{1}{2}\alpha^3 u_\varepsilon^2(t, x) + \langle FDu_\varepsilon(t, x), Du_\varepsilon(t, x)\rangle \\
& + (k - 1)\langle GD^2_*u_\varepsilon(t, x), D^2_*u_\varepsilon(t, x)\rangle, \quad k = 1, 2,
\end{aligned}
$$

for any $t > 0$, $x \in \mathbb{R}^N$, where F and G are as in (3.16).

Finally, to prove (3.15) with $(h, k) = (1, 2)$ and with $h = 1, 2$, $k = 3$, it suffices to repeat the previous arguments applied, respectively, to the functions $\xi_{1,2,\varepsilon}$

and $\xi_{1,3,\varepsilon}, \xi_{2,3,\varepsilon}$ defined by

$$\xi_{1,k,\varepsilon}(t,x) = \xi_{1,\varepsilon}(t,x) + \langle G(t)D_*^2 u_\varepsilon(t,x), D_*^2 u_\varepsilon(t,x)\rangle$$

$$+(k-1)\langle H_1(t)D_*^3 u_\varepsilon(t,x), D_*^3 u_\varepsilon(t,x)\rangle, \qquad k = 2, 3,$$

for any $t > 0$, $x \in \mathbb{R}^N$, and

$$\xi_{2,3,\varepsilon}(t,x) = \xi_{2,\varepsilon}(t,x) + \langle H_2(t)D_*^3 u_\varepsilon(t,x), D_*^3 u_\varepsilon(t,x)\rangle, \qquad t > 0, \quad x \in \mathbb{R}^N,$$

where

$$G(t) = \begin{pmatrix} tI_{n_r^1} & 0 & 0 \\ 0 & \alpha^{-7/16}tI_{(N-r)n_r^1} & -\alpha^{-4/5}t^2 L_5 \\ 0 & -\alpha^{-4/5}t^2 L_5^* & \alpha^{-7/8}t^3 I_{n_{N-r}^1} \end{pmatrix}, \qquad t > 0;$$

$$H_k(t) = \begin{pmatrix} \alpha^{-7/16}t^{3-k}I_{n_r^2} & 0 & 0 & 0 \\ 0 & \alpha^{-7/8}t^{3-k}I_{(N-r)n_r^1} & 0 & 0 \\ 0 & 0 & \alpha^{-1}t^{7-3k}I_{rn_{N-r}^1} & -\alpha^{-13/12}t^{8-3k}M_8 \\ 0 & 0 & -\alpha^{-13/12}t^{8-3k}M_8^* & \alpha^{-9/8}t^{3(3-k)}I_{n_{N-r}^2} \end{pmatrix},$$

for any $t > 0$ ($k = 1, 2$), with α sufficiently large. Here, as in the proof of Theorem 3.1, we have set $n_m^1 = (m-1)m/2$, $n_m^2 = m(m+1)(m+2)/6$ for any $m \in \mathbb{N}$. $\square$

3.2. Construction of the semigroup

In this section we prove that, for any $f \in C_b(\mathbb{R}^N)$, the Cauchy problem (HCP), associated with the degenerate elliptic operator $\mathcal{A}$ in (1.2), admits a unique classical solution u_f. This will allow us to define a semigroup of bounded operators in $C_b(\mathbb{R}^N)$ by setting $T(t)f = u_f(t,\cdot)$ for any $t > 0$. At the same time we also show that the semigroup $\{T(t)\}_{t\geq 0}$ satisfies the uniform estimates (1.3)-(1.8). The following remark will be fundamental in order to prove our results, since it provides us an useful maximum principles for the classical solution to Cauchy problem (HCP).

Remark 3.3. *(maximum principle)* Let us observe that the maximum principle as stated in Proposition 2.7 holds true also when $\mathcal{A}$ is given by (1.2). Indeed, in the proof of the quoted proposition we never took advantage of the fact that $\mathcal{A}$ was uniformly elliptic. We just took advantage of the existence of a Lyapunov function $\varphi \in C^2(\mathbb{R}^N)$ such that $\sup_{\mathbb{R}^N}(\mathcal{A}\varphi - \lambda\varphi) < +\infty$ for some $\lambda > 0$, and, as it is easily seen, the function $\varphi(x) = 1 + |x|^2$ is a Lyapunov function for the operator $\mathcal{A}$.

Theorem 3.4. *Under assumptions H1-H3, for any $f \in C_b(\mathbb{R}^N)$ there exists a unique classical solution u to problem* (HCP). *Moreover, if we set $u(t,\cdot) = T(t)f$, the family $\{T(t)\}_{t\geq 0}$ is an order preserving semigroup of linear operators in $C_b(\mathbb{R}^N)$ satisfying* (1.3)-(1.8).

Proof. We split the proof into several steps. First in Steps 1, 2 and 3 we prove that, for any $f \in C_b(\mathbb{R}^N)$, problem (HCP) admits a (unique) classical solution u_f; we define the operator $T(t)$ ($t > 0$) and we show that, for any compact set $F \subset (0, +\infty) \times \mathbb{R}^N$, $T_\varepsilon(\cdot)f$ converges to $T(\cdot)f$ in $C^{1,2}(F)$ as ε tends to 0. Moreover, we show that $\{T(t)\}_{t \geq 0}$ is an order preserving semigroup in $C_b(\mathbb{R}^N)$ and it satisfies (1.3), (1.4), (1.6) (with $k = 1, 2$) and (1.7). Then in Steps 4 and 5, we show that u_f is thrice-continuously differentiable with respect to the space variables in $(0, +\infty) \times \mathbb{R}^N$ and $T(t)$ satisfies the estimates (1.6) (with $k = 3$) and (1.8) (with $k = 0, 1, 2$).

Throughout the proof, we consider the spaces $B(I; C_b^k(K))$, $\mathrm{Lip}(I; C_b^k(K))$ and $C^h(I; C_b^k(K))$ where $I \subset \mathbb{R}$ is an interval, $K = \mathbb{R}^N$ or $K = \overline{B(0, m)}$ for some $m > 0$, and $h \in (0, 1)$, $k \geq 0$. These spaces are defined as follows. $B(I; C_b^k(\mathbb{R}^N))$ is the space of all the functions $u : I \times K \to \mathbb{R}$ such that $\|u\|_{B(I;C_b^k(K))} := \sup_{t \in [a,b]} \|u(t, \cdot)\|_{C_b^k(\mathbb{R}^N)}$ is finite. $\mathrm{Lip}(I; C_b^k(K))$ is the subset of $B(I; C_b^k(K))$ of all the functions u such that $\|u(t, \cdot) - u(s, \cdot)\|_{C_b^k(K)} \leq C|t - s|$ for any $s, t \in I$ and some $C \geq 0$. Finally, $C^h(I; C_b^k(K))$ is the set of all the $u : I \times K \to \mathbb{R}$ such that $\|u(t, \cdot) - u(s, \cdot)\|_{C_b^k(K)} \leq C|t - s|^h$, for any $s, t \in I$ and some $C \geq 0$.

Step 1. For any $f \in C_b(\mathbb{R}^N)$ and any $\varepsilon \in (0, 1]$, let $u_\varepsilon = T_\varepsilon(\cdot)f$ be, as usual, the solution to the Cauchy problem (HCP$_\varepsilon$). Using the estimates (1.4) and (1.5), we deduce that $\{u_\varepsilon\}_{\varepsilon > 0} \subset B([T_0, T]; C_b^3(\mathbb{R}^N))$, for any $0 < T_0 < T$, with norm *independent* of ε. Moreover, since $D_t u_\varepsilon = \mathcal{A}_\varepsilon u_\varepsilon$, then $\{D_t u_\varepsilon\}_{\varepsilon \in (0,1)}$ is equibounded and equicontinuous in $[T_0, T] \times \overline{B(0, R)}$ for any $R > 0$. It follows that $u_\varepsilon \in \mathrm{Lip}([T_0, T]; C(\overline{B(0, R)})) \cap B([T_0, T] \times C^3(\overline{B(0, R)}))$. From [13, Proposition 1.1.4(i) & Corollary 1.2.19] we deduce that $u_\varepsilon \in C^{(1-\alpha)/3}([T_0, T]; C^{2+\alpha}(\overline{B(0, R)}))$ and $D_t u_\varepsilon \in C^{(1-\alpha)/3}([T_0, T]; C(\overline{B(0, R)}))$ with norms *independent* of ε. Therefore, the families of functions $\{D_t^\alpha D_x^\beta u_\varepsilon\}_{\varepsilon \in (0,1]}$ ($2\alpha + |\beta| \leq 2$) are equibounded and equicontinuous in $[T_0, T] \times \overline{B(0, R)}$, for any $0 < T_0 < T$ and any $R > 0$. Hence, there exists an infinitesimal sequence $\{\varepsilon_n\}_{n \in \mathbb{N}}$ such that u_{ε_n} converges in $C^{1,2}(F)$, for any compact set $F \subset (0, +\infty) \times \mathbb{R}^N$, to some function u_f which, of course, solves the differential equation $D_t u_f - \mathcal{A} u_f = 0$.

We now assume that $f \in C_0^2(\mathbb{R}^N)$ and prove that u_f is continuous up to $t = 0$. For this purpose, we observe that by [18, Proposition 4.3] we can write

$$(T_{\varepsilon_n}(t)f)(x) - f(x) = \int_0^t (T_{\varepsilon_n}(s)\mathcal{A}_{\varepsilon_n}f)(x)\,ds, \qquad t > 0,\ x \in \mathbb{R}^N.$$

Since $\{T_\varepsilon(t)\}_{t \geq 0}$ is a semigroup of contractions for any $\varepsilon > 0$, we get

$$\|T_{\varepsilon_n}(t)f - f\|_\infty \leq \sup_{x \in \mathbb{R}^N} \int_0^t |(T_{\varepsilon_n}(s)\mathcal{A}_{\varepsilon_n}f)(x)|ds$$

$$\leq \|\mathcal{A}_{\varepsilon_n}f\|_\infty t \leq Ct\|f\|_{C_b^2(\mathbb{R}^N)}, \qquad t > 0,$$

where C is a positive constant, *independent* of n. Letting n go to $+\infty$, we deduce that $u_f(t, \cdot)$ tends to f uniformly as t tends to 0. Hence, the function u_f is a classical solution to problem (HCP). Repeating the same arguments as above, and taking Remark 3.3 into account, we can show that any sequence $\{u_{\varepsilon'_n}\}_{n\in\mathbb{N}}$ with ε'_n vanishing as n go to $+\infty$, admits a subsequence $\{u_{\varepsilon'_{n_k}}\}_{k\in\mathbb{N}}$ converging to u_f in $C^{1,2}(F)$ for any F as above. This implies that $T_\varepsilon(\cdot)f$ converges to u_f in $C^{1,2}(F)$ as ε tends to 0^+.

Step 2. We now assume that $f \in C_b(\mathbb{R}^N)$ has compact support, and let $u_{\varepsilon_n} = T_{\varepsilon_n}(t)f$ be as in Step 1. Moreover, let $\{f_m\}_{m\in\mathbb{N}} \subset C_b^2(\mathbb{R}^N)$ be a sequence of compactly supported functions converging to f in $C_b(\mathbb{R}^N)$. We fix $m \in \mathbb{N}$ and observe that, by Step 1, $T_{\varepsilon_n}(t)f_m$ converges to $T(t)f_m$, locally uniformly in $\mathbb{R}^N$, as n tends to $+\infty$, for any $t > 0$. Writing (2.7) with u being replaced with $T_{\varepsilon_n}(t)f - T_{\varepsilon_n}(t)f_m$ and then, letting n go to $+\infty$, we get

$$\|u_f(t, \cdot) - T(t)f_m\|_\infty \leq \|f - f_m\|_\infty, \qquad t > 0, \qquad m \in \mathbb{N}. \qquad (3.20)$$

Hence, from (3.20) it follows that

$$\|u_f(t, \cdot) - f\|_\infty \leq \|u_f(t, \cdot) - T(t)f_m\|_\infty + \|T(t)f_m - f_m\|_\infty + \|f_m - f\|_\infty$$

$$\leq 2\|f - f_m\|_\infty + \|T(t)f_m - f_m\|_\infty,$$

for any $t > 0$ and any $m \in \mathbb{N}$. Hence, by Step 1, $u_f(t, \cdot)$ tends to f uniformly in $\mathbb{R}^N$ as t tends to 0. This implies that the function u_f is a classical solution to problem (HCP). Then, with the same arguments as in Step 1, we can easily show that $D_t^\alpha D_x^\beta T_\varepsilon(\cdot)f$ converges to $D_t^\alpha D_x^\beta T(\cdot)f$ locally uniformly in $(0, +\infty) \times \mathbb{R}^N$ as ε goes to 0^+, for any $2|\alpha| + |\beta| \leq 2$.

Step 3. We now assume that $f \in C_b(\mathbb{R}^N)$ and adapt to our situation the technique in [6, Proposition 2.2]. We preliminarily observe that, by Proposition 2.7, $T_\varepsilon(t)\mathbb{1} = \mathbb{1}$ for any $t > 0$. Hence, $T(t)\mathbb{1}$ is well-defined for any $t > 0$ and $T(\cdot)\mathbb{1} = \mathbb{1}$.

We now fix a compact set $K \subset \mathbb{R}^N$ and a smooth function η compactly supported in $\mathbb{R}^N$ such that $\eta = \mathbb{1}$ in K and $0 \leq \eta \leq \mathbb{1}$. By linearity and Step 2, we easily see that, for any $t > 0$, the function $T(t)(\mathbb{1} - \eta)$ is well-defined. Moreover, since $T_\varepsilon(t)(\mathbb{1} - \eta) \geq 0$ for any $t > 0$ and any $\varepsilon \in (0, 1)$ (see (2.8)), then

$$0 \leq T(t)(\mathbb{1} - \eta) = \mathbb{1} - T(t)\eta, \qquad t > 0.$$

By Step 2 we know that $T(t)\eta$ tends to η, uniformly in $\mathbb{R}^N$. Therefore, $T(t)(\mathbb{1} - \eta)$ tends to 0 as t tends to 0, uniformly in K.

Let now $u_{\varepsilon_n} = T_{\varepsilon_n}(\cdot)f$ and u_f be as in Step 1. Since $T_{\varepsilon_n}(t)(\mathbb{1} - \eta)f) = T_{\varepsilon_n}(t)f - T_{\varepsilon_n}(\eta f)$ and ηf is compactly supported in $\mathbb{R}^N$, then $T_{\varepsilon_n}(t)(\mathbb{1} - \eta)f)$ converges in $C^{1,2}(F)$ to $u_f(t, \cdot) - T(t)(\eta f)$, for any compact set $F \subset (0, +\infty) \times \mathbb{R}^N$. Since the semigroup $\{T_\varepsilon(t)\}_{t\geq 0}$ is order preserving, then

$$|(T_{\varepsilon_n}(t)((\mathbb{1} - \eta)f))(x)| \leq \|f\|_\infty (T_{\varepsilon_n}(t)(\mathbb{1} - \eta))(x), \qquad t > 0, \quad x \in \mathbb{R}^N, \quad n \in \mathbb{N},$$

so that,

$$|u_f(t, x) - (T(t)(\eta f))(x)| \leq \|f\|_\infty (T(t)(\mathbb{1} - \eta))(x), \qquad t > 0, \quad x \in \mathbb{R}^N.$$

It follows that $u_f(t, \cdot) - T(t)(\eta f)$ vanishes uniformly in K, as t tends to 0. Since $T(t)(\eta f)$ tends to f, uniformly in K, as t tends to 0, then $u_f(t, \cdot)$ converges to f, as t tends to 0, uniformly in K. By the arbitrariness of K, u_f turns out to be continuous up to $t = 0$ and it is a classical solution to problem (HCP). Arguing once more as in Step 1, we can then easily show that $D_t^\alpha D_x^\beta T_\varepsilon(\cdot) f$ converges to $D_t^\alpha D_x^\beta T(\cdot) f$ locally uniformly in $(0, +\infty) \times \mathbb{R}^N$ as ε goes to 0.

Now, the estimates (1.3), (1.4), (1.6) (with $k = 1, 2$) and (1.7) easily follow letting ε go to 0^+ in (3.1), (3.15), recalling that the constants occuring in these last estimates are *independent* of ε.

To conclude, we observe that the family of bounded operators $\{T(t)\}_{t \geq 0}$, $T(0) = I$, is an order preserving semigroup of linear operators in $C_b(\mathbb{R}^N)$. To check the semigroup rule, it suffices to notice that, for any $f \in C_b(\mathbb{R}^N)$ and any $s > 0$, both the functions $u(t, \cdot) = T(t)T(s)f$ and $v(t, \cdot) = T(t+s)f$ are classical solutions to the Cauchy problem (HCP) with f replaced with $T(s)f$. Hence, from Remark 3.3 we deduce that $u \equiv v$, i.e. $T(t+s)f = T(t)T(s)f$. The maximum principle also implies that $\{T(t)\}_{t \geq 0}$ is an order preserving semigroup.

Step 4. We now prove (1.5), (1.6) (with $k = 3$) and (1.8) (all of them in the case where $i \leq r$), using a localization argument. Without loss of generality, we can assume that $f \in C_b^3(\mathbb{R}^N)$. Indeed, once (1.5) and (1.8) are established for any $f \in C_b^3(\mathbb{R}^N)$, then they can be extended, respectively, to the case where $f \in BUC(\mathbb{R}^N)$ and $f \in BUC^k(\mathbb{R}^N)$ ($k = 0, 1, 2$) by a density argument, approximating $f \in BUC^k(\mathbb{R}^N)$ with a sequence of functions in $C_b^3(\mathbb{R}^N)$, converging uniformly to f in $BUC^k(\mathbb{R}^N)$. Here, $BUC^k(\mathbb{R}^N)$ ($k \in \mathbb{N}$) denotes the set of all the functions $f \in C_b^k(\mathbb{R}^N)$ such that $D^\alpha f$ is uniformly continuous in $\mathbb{R}^N$ for any $|\alpha| = k$. Finally, for a general $f \in C_b^k(\mathbb{R}^N)$, it suffices to split, for any $t > 0$, $T(t)f = T(t/2)T(t/2)f$ and apply the above results with f replaced with $T(t/2)f \in BUC^k(\mathbb{R}^N)$. Moreover, to prove the previous estimates, it suffices to show that for any $i \in \{1, \dots, r\}$ and any $j, h \in \{1, \dots, N\}$, the function $D_{jh}T(t)f$ is continuously differentiable with respect to the i-th space variable. Indeed, suppose for instance that $i = 1$, $j \leq r$, $h > r$ and $f \in C_b(\mathbb{R}^N)$. Estimate (3.1) (with $k = 3$) implies that, for any $x_2, \dots, x_N$, the function $x \mapsto (D_{jh}T_\varepsilon(t)f)(x, x_2, \dots, x_N)$ is Lipschitz continuous in $\mathbb{R}$ and

$$[(D_{jh}T_\varepsilon(t)f)(\cdot, x_2, \dots, x_N)]_{\mathrm{Lip}(\mathbb{R})} \leq Ce^{\omega t}t^{-5/2}\|f\|_{C_b(\mathbb{R}^N)}, \qquad t > 0,$$

for suitable $C, \omega > 0$, *independent* of $x_2, \dots, x_N$. Since $D_{jh}T_\varepsilon(t)f$ converges to $D_{jh}T(t)f$ locally uniformly, then the function $(D_{jh}T(t)f)(\cdot, x_2, \dots, x_N)$ is Lipschitz continuous in $\mathbb{R}$ as well, and

$$[(D_{jh}T(t)f)(\cdot, x_2, \dots, x_N)]_{\mathrm{Lip}(\mathbb{R})} \leq Ce^{\omega t}t^{-5/2}\|f\|_{C_b(\mathbb{R}^N)}, \qquad t > 0.$$

Therefore, if $D_{jh}T(t)f$ is continuously differentiable with respect to the direction e_1, then the function $D_{ijh}T(t)f$ satisfies (1.8).

So, we fix $i \leq r$ and $j, h \leq N$ and we prove that the function $D_{jh}u = D_{jh}T(\cdot)f$ is continuously differentiable in $\mathbb{R}_+ \times \mathbb{R}^N$ with respect to the i-th variable. For this purpose, let $\eta_R : \mathbb{R}^N \to \mathbb{R}$ be a smooth function compactly supported in $B(0, R)$ $(R > 0)$ such that $\eta_R \equiv 1$ in $B(0, R/2)$. For any $k \in \mathbb{R}$, with $|k| \leq 1$, we introduce the operator τ_h^k defined on $C_b(\mathbb{R}^N)$ by

$$\tau_k^h \psi(x) = \frac{\psi(x + ke_h) - \psi(x)}{k}, \qquad x \in \mathbb{R}^N, \qquad \psi \in C_b(\mathbb{R}^N),$$

where e_h is the h-th vector of the Euclidean basis of $\mathbb{R}^N$. Moreover, we set $v_{\varepsilon,k,R}^h = \tau_k^h v_{\varepsilon,R}$, where $v_{\varepsilon,R} = u_\varepsilon \eta_R$ and $u_\varepsilon = T_\varepsilon(\cdot)f$. As is easily seen, the function $v_{\varepsilon,k,R}^h$ is the classical solution to the Cauchy problem

$$\begin{cases} D_t v_{\varepsilon,k,R}^h(t, x) = \mathcal{A}_\varepsilon v_{\varepsilon,k,R}^h(t, x) + g_{\varepsilon,k,R}^h(t, x), & (t, x) \in \mathbb{R}_+ \times \mathbb{R}^N, \\ v_{\varepsilon,h,R}^k(0, x) = \tau_k^h(\eta_R f)(x), & x \in \mathbb{R}^N, \end{cases}$$

where, for any $t > 0$,

$$g_{\varepsilon,k,R}^h(t, \cdot) = -\tau_k^h(u_\varepsilon(t, \cdot)\mathcal{A}_\varepsilon \eta_R) - 2\sum_{l,m=1}^r q_{lm}D_l u_\varepsilon(t, \cdot + ke_h)\tau_k^h(D_m \eta_R)$$

$$-2\sum_{l,m=1}^r q_{lm}(\tau_k^h D_l u_\varepsilon(t, \cdot))D_m \eta_R$$

$$-2\sum_{l,m=1}^r (\tau_k^h q_{lm})D_l u_\varepsilon(t, \cdot + ke_h)D_m \eta_R(\cdot + ke_h)$$

$$-2\varepsilon \sum_{m=r+1}^N D_m u_\varepsilon(t, \cdot + ke_h)(\tau_k^h D_m \eta_R)$$

$$-2\varepsilon \sum_{m=r+1}^N (\tau_k^h D_m u_\varepsilon(t, \cdot))D_m \eta_R + \sum_{l,m=1}^r (\tau_k^h q_{lm})D_{lm}v_{\varepsilon,R}(t, \cdot)$$

$$+\sum_{l=1}^N b_{lh}D_l v_{\varepsilon,R}(t, \cdot). \tag{3.21}$$

By [21, Theorem 3.5], $v_{\varepsilon,k,R}^h$ can be represented by

$$v_{\varepsilon,k,R}^h(t, x) = (T_\varepsilon(t)(\tau_k^h(\eta_R f)))(x) + \int_0^t (T_\varepsilon(t - s)g_{\varepsilon,k,R}^h(s, \cdot))(x)ds, \tag{3.22}$$

$$t > 0, \ x \in \mathbb{R}^N.$$

We are going to show that we can take the limit as ε tends to 0^+ in (3.23) and write

$$v_{k,R}^h(t,x) := (\tau_k^h u(t,\cdot))(x)\eta_R(x) = (T(t)(\tau_k^h(\eta_R f)))(x)$$

$$+ \int_0^t (T(t-s)g_{k,R}^h(s,\cdot))(x)ds, \qquad t > 0, \ x \in \mathbb{R}^N, \qquad (3.23)$$

where $g_{k,R}^h$ is obtained from $g_{\varepsilon,k,R}^h$ by replacing u_ε with u and letting $\varepsilon = 0$ in (3.21). Of course, thanks to the previous steps, it suffices to deal with the convolution term in (3.22). Since u_ε converges to u in $C^{1,2}(F)$ for any compact set $F \subset (0,+\infty) \times \mathbb{R}^N$ (see Step 1) and η_R is compactly supported in $\mathbb{R}^N$, then the continuous function $g_{\varepsilon,k,R}^h$ converges uniformly in $\mathbb{R}^N$ to the function $g_{k,R}^h$ as ε tends to 0. This implies that, for any $s,t > 0$, $T_\varepsilon(t)g_{\varepsilon,k,R}^h(s,\cdot)$ converges to $T(t)g_{k,R}^h(s,\cdot)$ locally uniformly in $\mathbb{R}^N$ as ε tends to 0. Indeed, for any compact set $K \subset \mathbb{R}^N$, we have

$$\sup_{x\in K} |(T_\varepsilon(t)g_{\varepsilon,k,R}^h(s,\cdot))(x) - (T(t)g_{k,R}^h(s,\cdot))(x)|$$

$$\leq \sup_{x\in K} |(T_\varepsilon(t)g_{\varepsilon,k,R}^h(s,\cdot))(x) - (T_\varepsilon(t)g_{k,R}^h(s,\cdot))(x)|$$

$$+ \sup_{x\in K} |(T_\varepsilon(t)g_{k,R}^h(s,\cdot))(x) - (T(t)g_{k,R}^h(s,\cdot))(x)|$$

$$\leq \|g_{\varepsilon,k,R}^h - g_{k,R}^h\|_\infty + \sup_{x\in K} |(T_\varepsilon(t)g_{k,R}^h(s,\cdot))(x) - (T(t)g_{k,R}^h(s,\cdot))(x)|,$$

and, by virtue of Step 1, the last side of the previous chain of inequalities vanishes as ε tends to 0^+. Moreover, since the semigroups $\{T_\varepsilon(t)\}_{t\geq 0}$ are contractive, then $T_\varepsilon(\cdot)g_{\varepsilon,k,R}^h$ is bounded in $[0,T] \times \mathbb{R}^N$ for any $T > 0$, uniformly with respect to $\varepsilon \in (0,1)$. Therefore, letting ε go to 0^+ in (3.22), by the dominated convergence theorem, we get (3.23).

Next step consists in showing that we can let k go to 0 in (3.23) getting the fundamental representation formula

$$D_h v_R(t,x) = (T(t)(D_h(\eta_R f)))(x) + \int_0^t (T(t-s)g_R^h(s,\cdot))(x)ds, \qquad (3.24)$$

$$t > 0, \ x \in \mathbb{R}^N,$$

where

$$g_R^h = -D_h(u\mathcal{A}\eta_R) - 2\sum_{l,m=1}^r q_{lm}D_lu D_{mh}\eta_R - 2\sum_{l,m=1}^r q_{lm}D_{lh}u D_m\eta_R$$

$$-2\sum_{l,m=1}^r (D_h q_{lm})D_lu D_m\eta_R + \sum_{l,m=1}^r (D_h q_{lm})D_{lm}v_R + \sum_{l=1}^N b_{lh}D_l v_R, \quad (3.25)$$

and $v_R = u\eta_R$.

The convergence of the integral term in (3.23) follows from the dominated convergence theorem, since $g_{k,R}^h$ converges uniformly in $\mathbb{R}^N$ to the function g_R^h, as k tends to 0. To show this it suffices to observe that, for any $\psi \in C(\mathbb{R}^N)$ such that $D_h\psi \in C(\mathbb{R}^N)$, the function $\tau_k^h\psi$ converges to $D_h\psi$ locally uniformly in $\mathbb{R}^N$ as k tends to 0. Similarly $\tau_k^h(\eta_R f)$ and $v_{k,R}^h(t,\cdot)$ $(t > 0)$ converge uniformly, respectively, to $D_h(\eta_R f)$ and $D_h v_R(t,\cdot)$ as k tends to 0.

Now, taking advantage of the representation formula (3.24), we can show that the function $D_h v_R$ is twice continuously differentiable in $(0,T) \times \mathbb{R}^N$ with respect to the i-th and j-th space variable, for any $T > 0$. To simplify the notation, in the rest of the proof, we denote by C_j positive constants which may depend on R, but are *independent* of t, ε and k.

We begin by observing that, by interpolation, from (3.15) (with $h = k = 2, 3$) we deduce that $T_\varepsilon(\cdot)f \in B((0,T); C_b^{2+\alpha}(\mathbb{R}^N))$ with norm *independent* of ε. Therefore, letting ε go to 0, we get $T(\cdot)f \in B((0,T); C_b^{2+\alpha}(\mathbb{R}^N))$ as well and, consequently, $g_R^h \in B((0,T); C_b^\alpha(\mathbb{R}^N))$ for any $\alpha \in (0,1)$.

Next, interpolating the estimates (1.4) and (1.7) we get

$$\|D_{ij}T(t)\psi\|_\infty \le C_1 t^{-3/4}\|\psi\|_{C_b^{5/6}(\mathbb{R}^N)}, \qquad (3.26)$$

for any $\psi \in C_b^\alpha(\mathbb{R}^N)$. The estimate (3.26) implies that the function $s \to \|D_{ij}T(t-s)g_R^h(s,\cdot)\|_\infty$ is in $L^1(0,t)$. Hence, from (3.24) and Step 3, we immediately deduce that the function $D_h v_R$ is twice-continuously differentiable with respect to the i-th and j-th space variables in $(0,T) \times \mathbb{R}^N$. Since $v_R \equiv u$ in $B(0,R/2)$ and R is arbitrarily fixed, we deduce that $D_h u$ is twice-continuously differentiable with respect to the i-th and j-th space variables as well. The estimate (1.5) follows.

Step 5. We now show (1.5), (1.6) (with $k = 3$) and (1.8) in the case where $i, j, h > r$. Repeating the same arguments as in Step 4, we easily see that it is not restrictive to consider the case where $f \in C_b^4(\mathbb{R}^N)$. We are going to prove that the right-hand side of (3.24) defines a function which is twice continuously differentiable in $\mathbb{R}_+ \times \mathbb{R}^N$, with respect to space variables x_i and x_j. As a first step, we show that $g_R^h \in B((0,T); C_b^{3/2}(\mathbb{R}^N))$ for any $T > 0$. Of course, this is the case if $D_{lm}u \in B((0,T); C^{3/2}(\overline{B(0,R)}))$ for $1 \le l \le r$. To prove that $D_{lm}u \in B((0,T); C^{3/2}(\overline{B(0,R)}))$ we use a bootstrap argument, first showing that it belongs to $B((0,T); C_b^{1+\theta/3}(B(0,2R)))$ for any $\theta \in (0,1)$. For this purpose, we replace the function η_R defined in Step 4, with the function η_{4R} which satisfies $\eta_{4R} = \mathbb{1}$ in $B(0,2R)$. Interpolating (1.7) and (1.8) (with $k = 1$) we get

$$\|D_{lm}T(t)\psi\|_{C_b^{\theta/3}(\mathbb{R}^N)} \le \frac{C_2}{t^{(\theta+1)/2}}\|\psi\|_{C_b^1(\mathbb{R}^N)}, \qquad t \in (0,T), \quad 1 \le l \le r,$$

for any $\psi \in C_b^1(\mathbb{R}^N)$. It follows that the function $t \mapsto \|D_{lm}T(t)\psi\|_{C_b^{\theta/3}(\mathbb{R}^N)}$ is integrable in $(0,T)$. Therefore, since by Step 4, $T(\cdot)\psi$ and $D_{lm}T(\cdot)\psi$ belong to $B((0,T); C_b^1(\mathbb{R}^N))$ for any $1 \le l \le r$, it follows that $g_{4R}^h \in B((0,T); C_b^1(\mathbb{R}^N))$ as

well and then, by (3.24), that $D_{lm}u \in B((0,T); C^{1+\theta/3}(\overline{B(0,2R)}))$, since $v_{4R} \equiv u$ in $\mathbb{R}_+ \times B(0,2R)$. Hence, by (3.25) (with R replaced with $2R$) we deduce that $g_{2R}^h \in B((0,T); C_b^{1+\theta/3}(\mathbb{R}^N))$ for any $\theta \in (0,1)$. Now, we interpolate first (1.8), respectively, with $k = 1$ and $k = 2$, and then (1.4) and (1.7), obtaining

$$\|D_{lm}T(t)\psi\|_{C_b^1(\mathbb{R}^N)} \leq \frac{C_3}{t^{2-\theta/2}}\|\psi\|_{C_b^{1+\theta/3}(\mathbb{R}^N)}, \quad t \in (0,T), \quad 1 \leq l \leq r, \quad (3.27)$$

$$\|D_{lm}T(t)\psi\|_\infty \leq \frac{C_4}{t^{(3-\theta)/6}}\|\psi\|_{C_b^{1+\theta/3}(\mathbb{R}^N)}, \quad t \in (0,T) \quad 1 \leq l \leq r. \quad (3.28)$$

Then, interpolating (3.27) and (3.28), yields

$$\|D_{lm}T(t)\psi\|_{C_b^{1/2}(\mathbb{R}^N)} \leq \frac{C_5}{t^{(15-4\theta)/12}}\|\psi\|_{C_b^{1+\theta/3}(\mathbb{R}^N)}, \quad t \in (0,T), \quad 1 \leq l \leq r. \quad (3.29)$$

Taking $\theta = 4/5$ in (3.29), we easily see that the function $s \mapsto \|D_{lm}T(t-s)g_{2R}^h(s)\|_{C_b^{1/2}(\mathbb{R}^N)}$ is integrable in $(0,T)$. Hence, from formula (3.24) we deduce that $D_{lm}v_{2R} \in B((0,T); C_b^{3/2}(\mathbb{R}^N))$, and, eventually, we obtain that $D_{lm}u \in B((0,T); C^{3/2}(\overline{B(0,R)}))$, since $v_{2R} \equiv u$ in $\mathbb{R}_+ \times B(0,R)$.

Now, we are almost done. Indeed, interpolating (1.6) and (1.7), we easily see that

$$\|D_{ij}T(t)\psi\|_\infty \leq \frac{C_6}{t^{3/4}}\|\psi\|_{C_b^{3/2}(\mathbb{R}^N)}, \quad t \in (0,T), \quad i,j > r,$$

which, due to the above results, implies that the map $s \mapsto \|D_{ij}T(t-s)g_R^h(s,\cdot)\|_\infty$ is integrable in $(0,t)$. Therefore, from (3.24), we easily obtain that $D_{ijk}u(t,\cdot)$ exists for any $t > 0$. $\square$

Remark 3.5. In fact, in Step 2 of the proof of Theorem 3.4 we have shown that, for any $f \in C_0(\mathbb{R}^N)$, $T(t)f$ converges to f uniformly in $\mathbb{R}^N$ as t tends to 0.

References

[1] S. BERNSTEIN, *Sur la généralisation du probléme de Dirichlet, I*, Math. Ann. **62** (1906), 253-271.

[2] M. BERTOLDI and L. LORENZI, Analytic methods for Markov semigroups, Preprint 401, Dipartimento di Matematica, Università di Parma, 2005.

[3] M. BERTOLDI and L. LORENZI, *Estimates of the derivatives for parabolic operators with unbounded coefficients*, Trans. Amer. Math. Soc. (to appear).

[4] S. CERRAI, *Some results for second order elliptic operators having unbounded coefficients*, Differential Integral Equations **11** (1998), 561-588.

[5] G. DA PRATO, *Regularity results for some degenerate parabolic equations*, Riv. Mat. Univ. Parma (6) **2*** (1999), 245-257.

[6] S. FORNARO, G. METAFUNE and E. PRIOLA, *Gradient estimates for Dirichlet parabolic problems in unbounded domains*, J. Differential Equations **205** (2004), 329-353.

[7] A. FRIEDMAN, "Partial Differential Equations of Parabolic Type", Prentice Hall, Englewood Cliffs, N.J., 1964.

[8] R.Z. HAS'MINSKII, "Stochastic Stability of Differential Equations", Nauka 1969 (in Russian), English translation: Sijthoff and Noordhoff 1980.

[9] N.V. KRYLOV, "Introduction to the Theory of Diffusion Processes", American Mathematical Society **142**, (1992).

[10] O. A. LADYZHENSKAJA, V. A. SOLONNIKOV and N. N. URAL'CEVA, "Linear and Quasilinear Equations of Parabolic Type", Nauka, English transl.: American Mathematical Society, Providence, 1968.

[11] G. LIEBERMAN, "Second Order Parabolic Differential Equations", World Scientific Publishing Co. Pte. Ltd, Singapore, New Jersey, London Hong Kong, 1996.

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

[13] A. LUNARDI, "Analytic Semigroups and Optimal Regularity in Parabolic Problems", Birkhäuser, Basel, 1995.

[14] A. LUNARDI, *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** (1997), 133-164.

[15] A. LUNARDI, *Schauder theorems for linear elliptic and parabolic problems with unbounded coefficients in $\mathbb{R}^N$*, Studia Math. **128** (1998), 171-198.

[16] M. MANFREDINI, *The Dirichlet problem for a class of ultraparabolic equations*, Adv. Differential Equations **2** (1997), 831-866.

[17] M. MANFREDINI and A. PASCUCCI, *A priori estimates for quasilinear degenerate parabolic equations*, Proc. Amer. Math. Soc. **131** (2002), 1115-1120.

[18] G. METAFUNE, D. PALLARA and M. WACKER, *Feller semigroups on $\mathbb{R}^N$*, Semigroup Forum **65** (2002), 159-205.

[19] A. PASCUCCI, *Hölder regularity for a Kolmogorov equation*, Trans. Amer. Math. Soc. **355** (2002), 901-924.

[20] S. POLIDORO, *On a class of ultraparabolic operators of Kolmogorov-Fokker-Plank type*, Matematiche (Catania) **49** (1994), 53-105 (1995).

[21] E. PRIOLA, *The Cauchy problem for a class of Markov-type semigroups*, Comm. Appl. Anal. **5** (2001), 49-75.

Dipartimento di Matematica
Università di Parma
Parco Area delle Scienze 53/A
43100 Parma, Italy
luca.lorenzi@unipr.it