**Global Lipschitz regularizing effects for linear and nonlinear parabolic equations**

# GLOBAL LIPSCHITZ REGULARIZING EFFECTS FOR LINEAR AND NONLINEAR PARABOLIC EQUATIONS

A. PORRETTA AND E. PRIOLA

**Abstract**

In this paper we prove global bounds on the spatial gradient of viscosity solutions to second order linear and nonlinear parabolic Cauchy problems in $(0,T)\times\mathbb{R}^N$. Our assumptions include the case that the coefficients be both unbounded and with very mild local regularity (possibly weaker than the Dini continuity), the estimates only depending on the parabolicity constant and on the modulus of continuity of coefficients (but not on their $L^\infty$-norm). Our proof provides the analytic counterpart to the probabilistic proof used in Priola and Wang [PW06] (J. Funct. Anal. 2006) to get this type of gradient estimates in the linear case. We actually extend such estimates to the case of possibly unbounded data and solutions as well as to the case of nonlinear operators including Bellman-Isaacs equations. We investigate both the classical regularizing effect (at time $t>0$) and the possible conservation of Lipschitz regularity from $t=0$, and similarly we prove global Hölder estimates under weaker assumptions on the coefficients. The estimates we prove for unbounded data and solutions seem to be new even in the classical case of linear equations with bounded and Hölder continuous coefficients. Finally, we compare in an appendix the analytic and the probabilistic approach discussing the analogy between the doubling variables method of viscosity solutions and the probabilistic coupling method.

Contents



This work has been supported by the Italian GNAMPA project 2008 “Problemi di Diffusione degeneri” and GNAMPA project 2010 “Proprietà di regolarità in Equazioni alle Derivate Parziali nonlineari legate a problemi di controllo”.

## 1. Introduction

It is well known that the bounded solution $u(t,x)$ of the heat equation in $\mathbb{R}^N$, even if it is only bounded at time $t=0$, becomes Lipschitz as soon as $t>0$ with a global Lipschitz bound of the type $\|Du(t,\cdot)\|_\infty \le \frac{\|u_0\|_\infty}{\sqrt{t}}$.

In this paper we prove a similar global gradient bound for solutions of the Cauchy problem concerning both linear and nonlinear parabolic operators. In particular, our main goal is to obtain this global gradient estimate independently of the $L^\infty$-bounds of the coefficients of the operator but only depending on their modulus of continuity. In case of solutions of linear diffusions, such bounds are well-known whenever either the coefficients are uniformly continuous and bounded (see [St74]) or the coefficients are unbounded but at least $C^1$ with a suitable control on the derivatives (see [EL94], [Ce96], [L98], [Ce01], [BF04], [KLL10] and the references therein). Note that linear parabolic equations with unbounded coefficients also arise in some models of financial mathematics (see, for instance, [HR98] and the references therein). Recently, in [PW06] E. Priola and F.Y. Wang obtained uniform gradient estimates when the coefficients are unbounded and singular at the same time. More precisely, they proved that the uniform estimate

$$\|Du(t,\cdot)\|_\infty \le \frac{K}{\sqrt{t\wedge 1}}\,\|u_0\|_\infty, \quad t>0, \tag{1.1}$$

holds for the diffusion semigroup solution $u = P_t u_0$ of the Cauchy problem

$$\begin{cases} \partial_t u - A(u) = 0 & \text{in } (0,T)\times\mathbb{R}^N \\ u(0,\cdot) = u_0 \end{cases} \quad \text{where} \quad A = \sum_{i,j=1}^N q_{ij}(x)\partial^2_{x_i x_j} + \sum_{i=1}^N b_i(x)\partial_{x_i}, \tag{1.2}$$

only assuming that $q_{ij}$ and $b_i$ are continuous and that $q(x)$ is elliptic, more precisely if

$$q(x) = \lambda\, I + \sigma(x)^2, \quad x\in\mathbb{R}^N, \tag{1.3}$$

for some $\lambda>0$ and some symmetric $N\times N$ matrix $\sigma(x)$ satisfying, jointly with the drift $b(x)$, the following condition:

$$\|\sigma(x)-\sigma(y)\|^2 + (b(x)-b(y))\cdot(x-y) \le g(|x-y|)\,|x-y|, \quad x,y\in\mathbb{R}^N,\ 0<|x-y|\le 1, \tag{1.4}$$

for some nonnegative function $g$ such that $\int_0^1 g(s)\,ds < \infty$. In this case, the constant $K$ in (1.1) only depends on $\lambda$ and $g$. Observe that, when $b\equiv 0$, this is a slightly weaker assumption than asking $\sigma(x)$ to be Dini-continuous. Moreover, both the diffusion and the drift could be unbounded, in particular $b$ could be the gradient of whatever concave $C^1$-function (as in standard Ornstein-Uhlenbeck operators).

The proof of the result in [PW06] stands on probabilistic arguments and relies in particular on the extension to the case of non locally Lipschitz coefficients of the coupling method for diffusions (see e.g. [LR86], [CL89], [Cr91], [Cr92]). Possible extensions of this regularity result to infinite dimensions in the setting of [DZ02] are given in [WZ10].

The aim of our work is twofold. On one hand we give a new proof, using entirely analytic arguments, of the result by Priola and Wang; the only difference here is that, in place of the assumption that the underlying stochastic process is nonexplosive, we assume directly the existence of a smooth Lyapunov function, which is a stronger condition though more adapted to our analytical approach. On the other hand, we extend the same result in two main features, including in our results both the case of *unbounded data and solutions* (as well as possibly larger classes of unbounded coefficients) and the case of *nonlinear operators*. Further minor improvements with respect to [PW06] concern the fact that we consider time dependent coefficients or the case of Hölder estimates or unbounded potentials (see e.g. Theorem 3.12).

In order to allow for unbounded data and solutions, we review the gradient bound (1.1) by distinguishing two main steps in the a priori estimate. In a first step, one obtains a global gradient bound depending only on the oscillation of the solution $u$ (see Theorem 3.3). In a second step, one looks for conditions which allow to estimate the oscillation of $u$. In case of bounded data, the second step is easy since solutions turn out to be globally bounded. However, it is also possible to consider unbounded data, though yet with bounded oscillation, and prove that solutions have bounded oscillation as well, and then, by the first step, be globally Lipschitz for $t > 0$. Let us stress, however, that this latter estimate on the oscillation of $u$ requires (1.4) to hold *for every* $x, y \in \mathbb{R}^N$ and a linear growth assumption on the function $g$ at infinity is required.

Moreover, recall that our estimates rely on the assumption that a Lyapunov function $\varphi$ exists for $A$, which may implicitly be linked to a compatibility condition in the growth at infinity of the drift and diffusion coefficients. For instance, the requirement that $\varphi(x) = 1 + |x|^2$, $x \in \mathbb{R}^N$, is a Lyapunov function for $A$ corresponds to the hypothesis

$$\mathrm{Tr}(q(x)) + b(x) \cdot x \leq C(1 + |x|^2), \quad x \in \mathbb{R}^N, \tag{1.5}$$

for some $C \geq 0$.

In order to fix the ideas, our global result in the linear case (at least when the equation has no source terms and coefficients are independent on time as in (1.2)) reads as follows: *assume (1.3) and that (1.4) holds true for every* $x, y \in \mathbb{R}^N$*, where* $g \in L^1(0,1) \cap C(0,+\infty)$*,* $g(r)r \to 0$ *as* $r \to 0^+$ *and* $g(r) = O(r)$ *as* $r \to +\infty$*, that a Lyapunov function* $\varphi$ *exists satisfying (3.3) and that* $u_0 \in C(\mathbb{R}^N)$ *satisfies*

$$|u_0(x) - u_0(y)| \leq k_0 + k_1|x - y| + k_\alpha|x - y|^\alpha, \qquad x, y \in \mathbb{R}^N,$$

*for some* $\alpha \in (0,1)$ *and* $k_0, k_\alpha, k_1 \geq 0$*. Then any viscosity solution* $u \in C([0,T] \times \mathbb{R}^N)$ *of the Cauchy problem (1.2) such that* $u = o(\varphi)$ *as* $|x| \to \infty$ *(uniformly in* $[0,T]$*) satisfies that* $u(t) = u(t,\cdot)$ *is Lipschitz continuous for* $t \in (0,T)$ *and*

$$\|Du(t)\|_\infty \leq c_T \Big\{ \frac{k_0}{\sqrt{t \wedge 1}} + \frac{k_\alpha}{(t \wedge 1)^{\frac{1-\alpha}{2}}} + k_1 \Big\}, \tag{1.6}$$

*for some* $c_T$*, possibly depending on* $\alpha, \lambda$ *and* $g$*.*

We refer the reader to Sections 3.2 and 3.3 for more general results of this type (see Theorem 3.12, Lemma 3.19 and Theorem 3.24). Due to the fact that $u_0$ can be unbounded the previous result seems to be new even in the classical case of linear equations with bounded and Hölder continuous coefficients (note that, in particular, if $k_0 = k_\alpha = 0$ the above estimate (1.6) implies that unbounded Lipschitz data, as $u_0 = |x|$, yield globally Lipschitz solutions). We also mention that the same approach is used to get Hölder estimates, in which case we relax the assumptions on the coefficients (the function $g$ needs only satisfy $g(r)r \to 0$ as $r \to 0^+$) including, for instance, merely uniformly continuous coefficients in the diffusion part.

The extensions of the previous results to nonlinear operators include both the case of Bellman-Isaacs' type operators, namely sup or inf (sup) of linear operators, which arise naturally from stochastic control problems as well as from differential game theory (see e.g. [Kr80], [FS06], [DL06], [Ko09]), and the case of operators with general nonlinear first order terms, including possibly superlinear growth in the gradient. Similar extensions are possible since we prove the Lipschitz bound in the framework of viscosity solutions (this notion also generalizes, in the linear case, the notion of probabilistic solution, see [FS06]), combining the method introduced by I. Ishii and P.L. Lions ([IL90]) with the coupling idea used in [PW06]. The Ishii and Lions method is by now classical to obtain gradient bounds for viscosity solutions though mainly used in bounded domains or for local estimates (see e.g. [Ba91], [Ba08], [Ch93] and references therein).

Assumptions on the coefficients like (1.4) (see also the nonlinear generalization in Hypothesis 4.1) are close to typical hypotheses imposed to get local Lipschitz regularity in the viscosity setting (e.g. in the above quoted papers). The main novelty here consists in getting *global* Lipschitz estimates on $\mathbb{R}^N$ for possibly unbounded coefficients. To this purpose our assumption (see Hypothesis 4.1) takes into account both the possible dissipation at infinity of the drift term $b$ and the minimal (local) regularity of the coefficients. In this respect, we mention that (1.4) gives a slight improvement, at least in the model linear case, of the assumptions used in [IL90, Section III.1] to get local Lipschitzianity (see also Remark 4.3 and [Ba91, Section III.1] for similar improvements). Let us point out that some dissipation condition on the drift term is somehow necessary when this is unbounded. A counterexample in this sense is given in [BF04], even in the linear case and for $C^1$-coefficients.

We give, for nonlinear equations, similar results as those mentioned before for the linear case. Let us notice that, in the viscosity solutions approach, it is quite clear that a general gradient estimate should depend firstly on the oscillation of $u$, while this latter one is related to the growth of the solution and the initial data (see [GGIS91] for an approach of this type). To get rid of the infinity, again we require the existence of a Lyapunov function satisfying a (possibly strong) condition as supersolution. In this way we obtain estimates and regularity results for viscosity solutions which are possibly unbounded, with the only condition that their growth be dominated by the Lyapunov function's growth. Unfortunately, in the nonlinear case the existence of such a Lyapunov function remains as an implicit condition (see Hypothesis 4.4) to be possibly checked on the particular examples of equations. At the end of Section 4, we give simplified, more readable versions of this condition in at least two cases; both for the case of Bellman-Isaacs operators and for the case of operators with linear second order part and nonlinear first order terms. In particular, our result for Bellman-Isaacs equations of the type

$$\partial_t u + \inf_{\beta\in\mathcal{B}} \sup_{\alpha\in\mathcal{A}} \left\{ -\mathrm{tr}\left(q_{\alpha,\beta}(t,x)D^2u\right) - b_{\alpha,\beta}(t,x)\cdot Du - f_{\alpha,\beta}(t,x)\right\} = 0 \tag{1.7}$$

reads very similar as in the linear case (see in particular Corollary 4.19). We obtain indeed an estimate like (1.6) by requiring that the coefficients $q_{\alpha,\beta}(t,x)$ and $b_{\alpha,\beta}(t,x)$ satisfy hypotheses (1.3) and (1.4) uniformly with respect to $t\in(0,T)$, $\alpha$, $\beta$ and that a Lyapunov function exists, again uniformly in $\alpha,\beta$. It should be noticed that in some examples, if the coefficients have a controlled growth at infinity, Lyapunov functions with polynomial growth at infinity can be constructed at hand (see e.g. [BBBL03]). For example, in case of Bellman-Isaacs operators, $\varphi = 1+|x|^2$ is a Lyapunov function provided (as in (1.5))

$$\mathrm{tr}\left(q_{\alpha,\beta}(t,x)\right) + b_{\alpha,\beta}(t,x)\cdot x \le C(1+|x|^2) \quad (t,x)\in(0,T)\times\mathbb{R}^N\,, \qquad \alpha\in\mathcal{A},\beta\in\mathcal{B}\,.$$

Of course, the Lyapunov function condition is no more needed if one is only interested in *local* Lipschitz estimates (see Theorem 4.20). Moreover, we point out that the continuity assumptions made on the coefficients, see e.g. (1.4), are far beyond the assumptions currently used for the comparison principle of viscosity solutions to hold. Therefore, the Lipschitz, or Hölder estimates proved, allowing for local compactness in the uniform topology, can also be used to obtain the existence of continuous viscosity solutions (possibly globally Lipschitz for $t>0$), whereas the Perron's method is not usable since the comparison principle is missing.

Last but not least, revisiting the method of [IL90] we also show the common denominator which exists between the probabilistic coupling method and the analytic approach to such gradient estimates, which we closely compare in Section A in the model linear case. Indeed, when the doubling variables method of viscosity solutions is specialized to linear diffusions, it offers a truly analytic counterpart of the probabilistic approach; this latter one, in turn, gives an interesting interpretation of the viscosity techniques in terms of coupling of probability measures. Although the links between the viscosity solutions theory with stochastic processes are well known, especially in the theory of controlled Markov processes (see e.g. [Kr80] and [FS06]), it seems that a direct comparison between the coupling method and the doubling variable technique was not addressed before, and may be interesting in itself.

The plan of the paper is the following. We start with some notations and preliminary notions about viscosity solutions in Section 2. In Section 3 we deal with the linear case, offering, in particular, an analytical proof of the main result of [PW06]. Section 4 contains the extension to nonlinear operators and further comments. Let us stress here that *our approach can be extended to deal with more general quasilinear operators*, but we have chosen to postpone this extension eventually to future work, in order to avoid an increasing number of technicalities. Finally, Appendix A is devoted to a possible comparison between the probabilistic and the analytical approach, which we hope be of some interest to experts of different techniques. In particular, restricting to the linear case, we rephrase the viscosity solutions approach in a form which is closer to the probabilistic coupling method.

## 2. Preliminaries and notations

In the following, we set $Q_T = (0,T)\times\mathbb{R}^N$ and indicate by $\bar{Q}_T$ its closure. We denote by $C(Q_T)$ the set of real continuous functions defined on $Q_T$ and by $C^{1,2}(Q_T)$ the subset of functions which are $C^1$ in $t$ and $C^2$ in $x$. If $u \in C(Q_T)$ we often set $u(t) = u(t,\cdot)$, $t \in (0,T)$. By $W^{1,\infty}(\mathbb{R}^N)$ we denote, as usual, the set of real Lipschitz functions on $\mathbb{R}^N$; for $f \in W^{1,\infty}(\mathbb{R}^N)$, $\|f\|_\infty$ is the supremum norm of $f$ on $\mathbb{R}^N$ and $\mathrm{Lip}(f)$ the smallest Lipschitz constant of $f$.

Given a function $v : I \times \mathbb{R}^N \to \mathbb{R}$, where $I \subset \mathbb{R}$ is an interval, we say that $v$ has *bounded oscillation* in $I$ if

$$\mathrm{osc}_I(v) = \sup_{x,y\in\mathbb{R}^N,\,|x-y|\le 1,\,t\in I} |v(t,x)-v(t,y)| < +\infty. \tag{2.1}$$

If the function $v$ only depends on $x$, we will omit the subscribe $I$. It is easy to see that $v$ has bounded oscillation if and only if there exists $\delta_0 \in (0,1]$ such that

$$\mathrm{osc}_{I,\,\delta_0}(v) = \sup_{x,y\in\mathbb{R}^N,\,|x-y|\le\delta_0,\,t\in I} |v(t,x)-v(t,y)| < +\infty \tag{2.2}$$

(note that $osc_{I,\,1}(v) \le \frac{1}{\delta_0} osc_{I,\,\delta_0}(v)$). The sum of a bounded function and a uniformly continuous one is an example of function having bounded oscillation. Functions with bounded oscillation have always at most linear growth in $x$ (uniformly in $t$). In particular, $v$ has bounded oscillation if and only if

$$|v(t,x)-v(t,y)| \le k_0 + k_1|x-y|, \qquad x,y,\in\mathbb{R}^N\,,\ t\in I, \tag{2.3}$$

for some constants $k_0$, $k_1 \ge 0$. Note that this inequality is not restricted to $x$, $y$ such that $|x-y|\le 1$. However, one can easily show that (2.1) implies (2.3) with $k_0 = k_1 = \mathrm{osc}_I(v)$; and conversely, (2.3) obviously implies (2.1).

In the sequel, given two functions $\varphi$, $v : [0,T]\times\mathbb{R}^N \to \mathbb{R}$, we will write that $v = o(\varphi)$ in $\bar{Q}_T$ if, for any $\epsilon > 0$, there exists $C_\epsilon > 0$ such that $|v(t,x)| \le \epsilon\varphi(t,x)$, for $|x| > C_\epsilon$, $t \in [0,T]$.

Given two non-negative symmetric $N\times N$ (real) matrices $A, B$, we write $A \ge B$ if $A\xi\cdot\xi \ge B\xi\cdot\xi$, for any $\xi\in\mathbb{R}^N$ ($\cdot$ denotes the inner product and $|\cdot|$ the Eucliden norm in $\mathbb{R}^N$). Moreover, we always indicate with $\|A\|$ the Hilbert-Schmidt norm of a matrix $A$, defined as $\|A\| = \sqrt{\mathrm{Tr}\,(AA^*)}$ (here $A^*$ denote the adjoint or transpose matrix of $A$).

Let us briefly recall the notion of viscosity solution (in the standard continuous setting) for a general parabolic fully nonlinear equation

$$\partial_t u + F(t, x, u, Du, D^2u) = 0 \tag{2.4}$$

where $F(t, x, u, p, X)$ is a continuous function of $(t, x) \in Q_T$, $u \in \mathbb{R}$, $p \in \mathbb{R}^N$, $X \in \mathcal{S}_N$, being $\mathcal{S}_N$ the set of symmetric $N \times N$ matrices.

By $USC(Q_T)$ and $LSC(Q_T)$ we denote the sets of upper semicontinuous, respectively lower semicontinuous, functions in $Q_T$. Given a point $(t_0, x_0) \in Q_T$, we denote by $P^{2,+}_{Q_T} u(t_0, x_0)$ the set of "generalized derivatives" $(a, p, X) \in \mathbb{R} \times \mathbb{R}^N \times \mathcal{S}_N$ such that

$$u(t, x) \le u(t_0, x_0) + a(t - t_0) + p \cdot (x - x_0) + \frac{1}{2} X(x - x_0) \cdot (x - x_0) + o(t - t_0) + o(|x - x_0|^2)$$

$$\text{as } (t, x) \in Q_T \text{ and } t \to t_0,\ x \to x_0.$$

The set $P^{2,-}_{Q_T} u(t_0, x_0)$ is the set of $(a, p, X)$ satisfying the opposite inequality. Note that

$$P^{2,-}_{Q_T} u(t_0, x_0) = -P^{2,+}_{Q_T}(-u)(t_0, x_0). \tag{2.5}$$

Of course, $P^{2,+}_{Q_T} u(t_0, x_0)$ as well as $P^{2,-}_{Q_T} u(t_0, x_0)$ are reduced to the unique triplet $(\partial_t u(t_0, x_0)$, $Du(t_0, x_0)$, $D^2u(t_0, x_0))$ whenever $u \in C^{1,2}(Q_T)$, by simply applying Taylor's expansion. Moreover, for any $\psi \in C^{1,2}(Q_T)$ and any $(t_0, x_0) \in Q_T$ such that $u - \psi$ has a local maximum (respectively, a local minimum) at $(t_0, x_0)$ we have that $(\partial_t \psi(t_0, x_0), D\psi(t_0, x_0), D^2\psi(t_0, x_0))$ belongs to $P^{2,+}_{Q_T} u(t_0, x_0)$ (respectively, to $P^{2,-}_{Q_T} u(t_0, x_0)$). Finally, the set $\overline{P}^{2,+}_{Q_T} u(t_0, x_0)$ is defined as containing limits of elements of $P^{2,+}_{Q_T} u(t_n, x_n)$ whenever $(t_n, x_n) \in Q_T$, $t_n \to t_0$, $x_n \to x_0$ and $u(t_n, x_n) \to u(t_0, x_0)$ (and similarly for $\overline{P}^{2,-}_{Q_T} u(t_0, x_0)$). We refer to the key-reference [CIL92] for more details and comments.

**Definition 2.1.** *A function $u \in USC(Q_T)$ is a viscosity subsolution of (2.4) if, for every $(t_0, x_0) \in Q_T$, we have*

$$a + F(t_0, x_0, u(t_0, x_0), p, X) \le 0, \quad (a, p, X) \in P^{2,+}_{Q_T} u(t_0, x_0).$$

*A function $u \in LSC(Q_T)$ is a supersolution if, for every $(t_0, x_0) \in Q_T$, we have*

$$a + F(t_0, x_0, u(t_0, x_0), p, X) \ge 0, \quad (a, p, X) \in P^{2,-}_{Q_T} u(t_0, x_0).$$

*A continuous function $u \in C(Q_T)$ is a viscosity solution of (2.4) if it is both a subsolution and a supersolution.*

By the above remarks, one can alternatively define a viscosity subsolution (respectively, supersolution) requiring that for any $\psi \in C^{1,2}(Q_T)$ and any $(t_0, x_0) \in Q_T$ such that $u - \psi$ has a local maximum (respectively, a local minimum) at $(t_0, x_0)$ we have

$$\partial_t \psi(t_0, x_0) + F(t_0, x_0, u(t_0, x_0), D\psi(t_0, x_0), D^2\psi(t_0, x_0)) \le 0$$

(respectively, $\partial_t \psi(t_0, x_0) + F(t_0, x_0, u(t_0, x_0), D\psi(t_0, x_0), D^2\psi(t_0, x_0)) \ge 0$). Moreover, thanks to the continuity of $F$, the inequalities required in the above definition remain true for elements of $\overline{P}^{2,+}_{Q_T} u(t_0, x_0)$ (respectively, $\overline{P}^{2,-}_{Q_T} u(t_0, x_0)$) in case of subsolutions (respectively, supersolutions).

We remind that such a formulation is consistent (i.e., classical solutions $u \in C^{1,2}(Q_T)$ are viscosity solutions) provided $F(t, x, u, p, X)$ is nonincreasing with respect to the matrix $X \in \mathcal{S}_N$, namely $F(t, x, u, p, X) \ge F(t, x, u, p, Y)$ whenever $X \le Y$ (using the order relation of symmetric matrices). Such an assumption corresponds to a (possibly degenerate) ellipticity condition.

**Remark 2.2.** We work for simplicity of notation only on $Q_T$. However all the results can be easily extended to the case in which $Q_T$ is replaced by the more general domain $Q_{s,T} = [s,T] \times \mathbb{R}^N$, with $s < T$. Indeed if $v \in C(Q_{s,T})$ is a viscosity solution to (2.4) on $Q_{s,T}$ then introducing $u(t,x) = v(t+s,x)$ we get a viscosity solution on $Q_{T-s}$ for $\partial_t u + F(\cdot + s, x, u, Du, D^2u) = 0$. In addition, for example, gradient estimates for $u$ like $\|Du(t,\cdot)\|_\infty \le \frac{C_T}{\sqrt{t}} \|u(0,\cdot)\|_\infty$, $t \in (0, T-s)$ (cf. Corollary 4.11 and Theorem 4.5) become $\|Dv(t,\cdot)\|_\infty \le \frac{C_T}{\sqrt{t-s}} \|v(s,\cdot)\|_\infty$, $t \in (s,T)$. All the regularity results in the paper concerning viscosity solutions $u$ on $Q_T$ can be transferred in an obvious way to viscosity solutions $v$ on $Q_{s,T}$ for the corresponding "translated in time" parabolic equations.

Let us explicitly recall the following fundamental result, which is a special case of [CIL92, Theorem 8.3].

**Theorem 2.3.** *Let $u_i$, $i = 1,\dots,k$, be real continuous functions on $Q_T$. Assume that, for any $i = 1,\dots,k$, either $u_i$ is a subsolution or $-u_i$ is a supersolution to the parabolic equation (2.4). Let $z(t,x_1,\dots,x_k)$ be a $C^{1,2}$-function on $(0,T) \times (\mathbb{R}^N)^k$ such that the mapping*

$$u_1(t,x_1) + \ldots + u_k(t,x_k) - z(t,x_1,\dots,x_k)$$

*has a local maximum in $(\hat{t},\hat{x}_1,\dots,\hat{x}_k) \in (0,T) \times (\mathbb{R}^N)^k$.*

*Then, for every $n > 0$ there exist matrices $X_i \in \mathcal{S}_N$ and $a_i \in \mathbb{R}$, such that*

$$a_1 + \ldots + a_k = \partial_t z(\hat{t},\hat{x}_1,\dots,\hat{x}_k), \quad (a_i, D_{x_i} z(\hat{t},\hat{x}_1,\dots,\hat{x}_k), X_i) \in \overline{P}^{2,+}_{Q_T} u_i(\hat{t},\hat{x}_i),$$

$$-(n + c_N\|A\|)I \le \begin{pmatrix} X_1 & \dots & 0 \\ \dots & \dots & \dots \\ 0 & \dots & X_k \end{pmatrix} \le A + \frac{1}{n} A^2 \tag{2.6}$$

*where $A = D_x^2 z(\hat{t},\hat{x}_1,\dots,\hat{x}_k)$ (and the constant $c_N$ appears since we are using the Hilbert-Schmidt norm of $A$).*

**Remark 2.4.** If $u_i$ were regular, then $a_i = \partial_t u_i(\hat{t},\hat{x}_i)$, $D_{x_i} u(\hat{t},\hat{x}_i) = D_{x_i} z(\hat{t},\hat{x}_1,\dots,\hat{x}_k)$, $X_i = D^2 u_i(\hat{t},\hat{x}_i)$, and the above relations are consequence of usual conditions on local maximum points. The remarkable fact (see [CIL92]) is that the lemma holds for semicontinuous functions, and gives the possibility to use the viscosity formulation of $u_i$ as a subsolution.

## 3. Linear equations

### 3.1. **The main gradient estimate in terms of the oscillation of the solution.**

In this section we deal with the linear equation

$$\partial_t u - \operatorname{tr}\left(q(t,x) D^2 u\right) - b(t,x) \cdot Du = h(t,x) \qquad \text{in } Q_T. \tag{3.1}$$

Throughout the whole Section 3, we will always assume the following basic continuity hypothesis.

**Hypothesis 3.1.** We suppose that $q(t,x) = (q_{ij}(t,x))$ is a symmetric matrix depending continuously on $(t,x) \in Q_T$, $b(t,x)$ is a continuous vector field defined on $Q_T$ with values in $\mathbb{R}^N$, and $h : Q_T \to \mathbb{R}$ is a continuous function. ■

We also suppose that the matrix $q(t,x)$ satisfies the parabolicity (or ellipticity) condition

$$q(t,x)\xi \cdot \xi \ge \lambda\,|\xi|^2, \quad (t,x) \in Q_T, \;\; \xi \in \mathbb{R}^N, \tag{3.2}$$

for some $\lambda > 0$ and that the following condition (existence of a Lyapunov function) holds:

$$\exists\, \varphi \in C^{1,2}(\bar{Q}_T),\; M \ge 0 \;:\; \begin{cases} A_t(\varphi) \le M\varphi + \partial_t \varphi, & (t,x) \in Q_T, \\ \varphi(t,x) \to +\infty & \text{as } |x| \to \infty, \text{ uniformly in } t \in (0,T), \end{cases} \tag{3.3}$$

where we have set

$$A_t(u) = \mathrm{tr}\,\big(q(t,x)D^2u\big) + b(t,x)\cdot Du. \tag{3.4}$$

**Remark 3.2.** There is no loss of generality if one requires that (3.3) is satisfied with $M=0$, since we can always replace $\varphi$ with $e^{Mt}\varphi$. Thus, (3.3) is actually equivalent to

$$\exists\,\varphi\in C^{1,2}(\bar Q_T)\ :\ \begin{cases} A_t(\varphi)\le \partial_t\varphi, & (t,x)\in Q_T,\\ \varphi(t,x)\to+\infty & \text{as } |x|\to\infty, \text{ uniformly in } t\in[0,T]. \end{cases} \tag{3.5}$$

Moreover, possibly adding a constant to $\varphi$, we may suppose that $\varphi>0$.

Following [PW06], for any $\lambda$ such that (3.2) is satisfied we denote by $\sigma(t,x)$ the symmetric $N\times N$ nonnegative matrix such that

$$\sigma^2(t,x) = q(t,x) - \lambda\, I, \quad (t,x)\in Q_T. \tag{3.6}$$

Note that $\sigma(t,x)$ actually depends on $\lambda$ but for simplicity we drop in our notation this dependence.

**Theorem 3.3.** *Assume that (3.2), (3.3) hold true, and, in addition, that there exists a nonnegative $g\in C(0,1)$ such that $\int_0^1 g(s)ds<\infty$ and*

$$\frac{1}{|x-y|}\Big(\|\sigma(t,x)-\sigma(t,y)\|^2 + (b(t,x)-b(t,y))\cdot(x-y)\Big) \le g(|x-y|), \qquad x,y\in\mathbb{R}^N,\ 0<|x-y|\le 1, \quad t\in(0,T). \tag{3.7}$$

*Let $u\in C(Q_T)$ be a viscosity solution of (3.1) and assume that $u$ and $h$ have bounded oscillation in $Q_T$ (see (2.1)). Then $u(t)$ is Lipschitz continuous and, setting $t\wedge 1=\min(t,1)$, we have*

$$\|Du(t)\|_\infty \le \frac{\hat c_1}{\sqrt{t\wedge 1}}\,\omega(t,u) + \ \hat c_2\sqrt{t\wedge 1}\;\omega(t,h), \quad t\in(0,T), \tag{3.8}$$

*where $\hat c_1 = \frac{1+2\lambda}{\lambda}\, e^{\frac{1}{4\lambda}\int_0^1 g(s)ds}$, $\hat c_2=\frac{1}{4\lambda}\, e^{\frac{1}{4\lambda}\int_0^1 g(s)ds}$ and (according to (2.1))*

$$\omega(t,u) = osc_{(\frac t2, T\wedge\frac32 t)}(u), \quad \omega(t,h) = osc_{(\frac t2, T\wedge \frac32 t)}(h).$$

Before the proof we make some comments on hypothesis (3.7) and on related conditions.

**Remark 3.4.** If $b=0$ and if we assume that $sg(s)\to 0$ as $s\to 0^+$ (which is reasonable, being $g\in L^1(0,1)$), then (3.7) implies that

$$\sigma(t,\cdot) \ \text{ is uniformly continuous on } \mathbb{R}^N, \text{ uniformly with respect to } t. \tag{3.9}$$

Note that this does not require $q(t,x)$ to be uniformly continuous (e.g. take $N=1$ and $q(x)=x^2+1$).

In the general case that $b\neq 0$, assumption (3.7) does not imply any growth restriction on $\sigma(t,x)$, since the growth of the drift term and of the diffusion matrix may compensate adequately. As an example, consider

$$A_t = A = (1+|x|^4)\triangle u - 4N|x|^2 x\cdot Du. \tag{3.10}$$

Here taking $\lambda=1$ and $\sigma=|x|^2 I$ we have that (3.7) holds trivially with $g(r)=r$, and for *every* $x,y\in\mathbb{R}^N$ (not only for $x,y$ such that $|x-y|\le 1$). Notice also that in this case $\varphi(x)=1+|x|^2$ is a Lyapunov function satisfying (3.3).

**Remark 3.5.** If the matrix $q(t,x) = \Sigma(t,x)\Sigma(t,x)^*$ with $\Sigma(t,\cdot) \in \mathbb{R}^N \times \mathbb{R}^k$ (as it happens if the operator $A_t$ is associated to the stochastic differential equation $dX_t = b(t,X_t)dt + \Sigma(t,X_t)dB_t$, being $B_t$ a $k$-dimensional Brownian motion), then, assuming that

$$\begin{cases} \lambda(t,x)I \le q(t,x) \le \Lambda(t,x)I \\ \lambda(t,x) \ge \lambda > 0\,, \qquad \frac{\Lambda(t,x)}{\lambda(t,y)} \le M \quad \text{for any } x,y \in \mathbb{R}^N : |x-y| \le 1,\ t \in (0,T), \end{cases} \tag{3.11}$$

one can check assumption (3.7) directly on $\Sigma(t,x)$.

Indeed, if we take $\sigma(t,x) = \sqrt{q(t,x) - \frac{\lambda}{2}I}$, we have $\sigma(t,x) \ge \sqrt{\frac{\lambda(t,x)}{2}}I$ and the same for $\sigma(t,y)$. Recall that if $A$ and $B$ are positive symmetric $N \times N$ matrices such that $A \ge \lambda' I$ and $B \ge \lambda' I$ for some $\lambda' > 0$ then $\|\sqrt{A} - \sqrt{B}\| \le \frac{1}{2\sqrt{\lambda'}}\|A - B\|$ (see e.g. [PW06, Lemma 3.3]), hence

$$\|\sigma(t,x) - \sigma(t,y)\| \le \frac{1}{\sqrt{2}\sqrt{\lambda(t,x) \wedge \lambda(t,y)}}\,\|q(t,x) - q(t,y)\|, \quad t \in (0,T),\ x,y \in \mathbb{R}^N. \tag{3.12}$$

Since we also have

$$\|q(t,x) - q(t,y)\| \le \sqrt{2N}\sqrt{\max\{\Lambda(t,x),\Lambda(t,y)\}}\|\Sigma(t,x) - \Sigma(t,y)\| \tag{3.13}$$

we deduce that, under condition (3.11),

$$\|\sigma(t,x) - \sigma(t,y)\| \le \frac{\sqrt{N}\sqrt{\max\{\Lambda(t,x),\Lambda(t,y)\}}}{\sqrt{\min\{\lambda(t,x),\lambda(t,y)\}}}\,\|\Sigma(t,x) - \Sigma(t,y)\| \le \sqrt{M\,N}\,\|\Sigma(t,x) - \Sigma(t,y)\|$$

if $|x-y| \le 1$. Therefore, in this case the regularity of $\Sigma(t,x)$ can be used to verify (3.7). Note that (3.11) allows $\lambda(t,x)$ and $\Lambda(t,x)$ to have (the same) polynomial growth in $x$ of any order (uniformly in $t$).

*Proof.* According to Remark 3.2, we assume that (3.5) holds true. Let us fix $t_0 \in (0,T)$ and consider $\delta \in (0,1]$. We define the open set

$$\Delta = \Delta(t_0,\delta) = \Big\{(t,x,y) \in (0,T) \times \mathbb{R}^N \times \mathbb{R}^N \ :\ |x-y| < \delta\,,\ \frac{t_0}{2} < t < (T \wedge \frac{3}{2}t_0)\Big\}$$

and the function

$$\Phi_\epsilon(t,x,y) = u(t,x) - u(t,y) - K\psi(x-y) - \varepsilon\left(\varphi(t,x) + \varphi(t,y)\right) - C_0(t-t_0)^2 - \frac{\varepsilon}{T-t},$$

where $K$, $C_0$ will be chosen later (depending also on $t_0$), $\varphi$ satisfies (3.5) and $\psi(x-y) = f(|x-y|)$, where $f \in C^2((0,\delta])$ is a positive, increasing, Lipschitz function (to be fixed later) such that $f(0) = 0$.

The goal is to prove that, if $K$ and $C_0$ are large enough (independently on $\varepsilon$), then

$$\Phi_\epsilon(t,x,y) \le 0, \quad (t,x,y) \in \Delta\,. \tag{3.14}$$

Once (3.14) is proved, then choosing $t = t_0$ and letting $\varepsilon \to 0$ one gets

$$u(t_0,x) - u(t_0,y) \le K\psi(x-y) = Kf(|x-y|), \quad x\,,y\ :|x-y| < \delta\,. \tag{3.15}$$

Reversing the roles of $x$, $y$ and using the Lipschitz continuity of $f$ will allow us to conclude.

In order to prove (3.14), we argue by contradiction, assuming that

$$\sup_{\Delta} \Phi_\epsilon(t,x,y) > 0\,. \tag{3.16}$$

In the following, let us set (cf. (2.2))

$$\omega_{0,\delta}(u) = osc_{(\frac{t_0}{2}\,, T\wedge\frac{3}{2}t_0),\delta}(u) = \sup\Big\{|u(t,x) - u(t,y)|\,,\ |x-y| \le \delta\,,\, t \in \big(\frac{t_0}{2}\,,\, T \wedge \frac{3}{2}t_0\big)\Big\}\,.$$

Since we have, for $(t,x,y)\in\Delta$,

$$\Phi_\epsilon(t,x,y)\le \omega_{0,\delta}(u)-Kf(|x-y|)-\varepsilon\left(\varphi(t,x)+\varphi(t,y)\right)-C_0(t-t_0)^2-\frac{\epsilon}{T-t},$$

we deduce the following assertions.

1. Thanks to the fact that $\varphi$ blows-up at infinity, we have, for any $\epsilon>0$,

$$\Phi_\epsilon\to-\infty \qquad \text{as } |x|\to\infty \text{ or } |y|\to\infty$$

hence $\Phi_\epsilon$ has a global maximum on $\overline{\Delta}$.

2. Choosing $C_0\ge \frac{4\omega_{0,\delta}(u)}{t_0^2}$ we have $\Phi_\epsilon\le 0$ if $t=\frac{t_0}{2}$ or $t=\frac{3}{2}t_0$, and, since $\lim_{t\to T^-}\frac{\varepsilon}{T-t}=\infty$, we conclude that the positive maximum cannot be attained on $\{t=\frac{t_0}{2}\}\cup\{t=T\wedge\frac{3}{2}t_0\}$.

3. Choosing $K\ge\frac{\omega_{0,\delta}(u)}{f(\delta)}$ we have $\Phi_\epsilon\le 0$ if $|x-y|=\delta$.

4. Since the maximum value is positive, it cannot be reached when $x=y$.

Therefore, (3.16) implies that the global maximum of $\Phi_\epsilon$ cannot be attained when $x=y$ nor on $\partial\Delta$ if $C_0$ and $K$ satisfy the above conditions; in particular, the value of $C_0$ is fixed by now as

$$C_0=\frac{4\omega_{0,\delta}(u)}{t_0^2}.$$

We deduce that $u(t,x)-u(t,y)-z(t,x,y)$ has a local maximum (possibly depending on $\epsilon$) at some $(\hat t,\hat x,\hat y)\in\triangle$, where

$$z(t,x,y)=K\,\psi(x-y)+\varepsilon\left(\varphi(t,x)+\varphi(t,y)\right)+C_0(t-t_0)^2+\frac{\varepsilon}{T-t}. \tag{3.17}$$

We can apply Theorem 2.3 with $u_1(t,x)=u(t,x)$, $u_2(t,y)=-u(t,y)$ and get that, for every $n>0$, there exist matrices $X$, $Y\in\mathcal{S}_N$ and $a$, $b\in\mathbb{R}$ (we omit here and later the dependence on $n$ of $a,b,X,Y$) such that

$$a-b=\partial_t z(\hat t,\hat x,\hat y)\,,\quad (a,D_xz(\hat t,\hat x,\hat y),X)\in\overline{P}^{2,+}_{Q_T}u(\hat t,\hat x)\,,$$
$$(b,-D_yz(\hat t,\hat x,\hat y),Y)\in\overline{P}^{2,-}_{Q_T}u(\hat t,\hat y)$$

and

$$-\left(n+c_N\|D^2z(\hat t,\hat x,\hat y)\|\right)I\le\begin{pmatrix}X&0\\0&-Y\end{pmatrix}\le D^2z(\hat t,\hat x,\hat y)+\frac{1}{n}\left(D^2z(\hat t,\hat x,\hat y)\right)^2. \tag{3.18}$$

Using that $u(t,x)$ is a subsolution and $u(t,y)$ a supersolution, we obtain the two inequalities

$$a-\operatorname{tr}\left(q(\hat t,\hat x)X\right)-b(\hat t,\hat x)\cdot D_xz(\hat t,\hat x,\hat y)\le h(\hat t,\hat x)$$
$$b-\operatorname{tr}\left(q(\hat t,\hat y)Y\right)+b(\hat t,\hat y)\cdot D_yz(\hat t,\hat x,\hat y)\ge h(\hat t,\hat y).$$

Subtracting and using $a-b=\partial_tz(\hat t,\hat x,\hat y)$ we get

$$\partial_tz(\hat t,\hat x,\hat y)-\operatorname{tr}\left(q(\hat t,\hat x)X-q(\hat t,\hat y)Y\right)\le b(\hat t,\hat x)\cdot D_xz(\hat t,\hat x,\hat y)+b(\hat t,\hat y)\cdot D_yz(\hat t,\hat x,\hat y)+\omega_{0,\delta}(h) \tag{3.19}$$

where

$$\omega_{0,\delta}(h)=osc_{(\frac{t_0}{2},T\wedge\frac{3}{2}t_0),\delta}(h)=\sup\{|h(t,x)-h(t,y)|\,,\ |x-y|\le\delta\,,\,t\in(\frac{t_0}{2}\,,T\wedge\frac{3}{2}t_0)\}\,.$$

Recalling the definition of $z$ we compute and find

$$\begin{aligned}&\tfrac{\epsilon}{(T-\hat t)^2}+2C_0(\hat t-t_0)+\varepsilon\left(\partial_t\varphi(\hat t,\hat x)+\partial_t\varphi(\hat t,\hat y)\right)\\&-\operatorname{tr}\left(q(\hat t,\hat x)X-q(\hat t,\hat y)Y\right)\le\left(b(\hat t,\hat x)-b(\hat t,\hat y)\right)\cdot D\psi(\hat x-\hat y)\\&\quad+\varepsilon\left(b(\hat t,\hat x)\cdot D\varphi(\hat t,\hat x)+b(\hat t,\hat y)\cdot D\varphi(\hat t,\hat y)\right)+\omega_{0,\delta}(h).\end{aligned} \tag{3.20}$$

Next step consists in finding estimates on the matrices $X$, $Y$ in terms of $D^2z$, using (3.18). In particular, we have, for $(t,x,y) \in \triangle$,

$$\operatorname{tr}\left[\begin{pmatrix} q(t,x) & c(t,x,y) \\ c^*(t,x,y) & q(t,y)\end{pmatrix}\begin{pmatrix} X & 0 \\ 0 & -Y\end{pmatrix}\right] \tag{3.21}$$
$$\leq \operatorname{tr}\left[\begin{pmatrix} q(t,x) & c(t,x,y) \\ c^*(t,x,y) & q(t,y)\end{pmatrix}\left[D^2z(\hat t,\hat x,\hat y) + \frac{1}{n}\left(D^2z(\hat t,\hat x,\hat y)\right)^2\right]\right]$$

for every $N \times N$ matrix $c(t,x,y)$ such that $\begin{pmatrix} q(t,x) & c(t,x,y) \\ c^*(t,x,y) & q(t,y)\end{pmatrix}$ is symmetric and nonnegative. Following the choice of the coupling matrix in [PW06], we choose

$$c(t,x,y) = \sigma(t,x)\sigma(t,y) + \lambda\left(I - 2\left(\frac{x-y}{|x-y|}\otimes\frac{x-y}{|x-y|}\right)\right) \tag{3.22}$$

where $\sigma$ is given by (3.6). It follows that $c^*(t,x,y) = \sigma(t,y)\sigma(t,x) + \lambda\left(I - 2\left(\frac{x-y}{|x-y|}\otimes\frac{x-y}{|x-y|}\right)\right)$. Recalling the definition of $z$ in (3.17) we have

$$D^2z(\hat t,\hat x,\hat y) = K\begin{pmatrix} D^2\psi(\hat x-\hat y) & -D^2\psi(\hat x-\hat y) \\ -D^2\psi(\hat x-\hat y) & D^2\psi(\hat x-\hat y)\end{pmatrix} + \varepsilon\begin{pmatrix} D^2\varphi(\hat t,\hat x) & 0 \\ 0 & D^2\varphi(\hat t,\hat y)\end{pmatrix} \tag{3.23}$$

hence computing (3.21) we get

$$\begin{aligned}\operatorname{tr}\left(q(\hat t,\hat x)X - q(\hat t,\hat y)Y\right) &\leq K\operatorname{tr}\left((q(\hat t,\hat x) + q(\hat t,\hat y) - 2c(\hat t,\hat x,\hat y))D^2\psi(\hat x-\hat y)\right)\\ &+\varepsilon\operatorname{tr}\left(q(\hat t,\hat x)D^2\varphi(\hat t,\hat x) + q(\hat t,\hat y)D^2\varphi(\hat t,\hat y)\right)\\ &+\tfrac{1}{n}\operatorname{tr}\left[\begin{pmatrix} q(\hat t,\hat x) & c(\hat t,\hat x,\hat y) \\ c^*(\hat t,\hat x,\hat y) & q(\hat t,\hat y)\end{pmatrix}\left(D^2z(\hat t,\hat x,\hat y)\right)^2\right].\end{aligned}$$

On account of (3.20), we deduce that

$$\begin{aligned}2C_0(\hat t - t_0) &\leq K\operatorname{tr}\left((q(\hat t,\hat x) + q(\hat t,\hat y) - 2c(\hat t,\hat x,\hat y))D^2\psi(\hat x-\hat y)\right)\\ &+\left(b(\hat t,\hat x) - b(\hat t,\hat y)\right)\cdot D\psi(\hat x-\hat y)\\ &+\varepsilon\left(A_{\hat t}(\varphi)(\hat t,\hat x) - \partial_t\varphi(\hat t,\hat x) + A_{\hat t}(\varphi)(\hat t,\hat y) - \partial_t\varphi(\hat t,\hat y)\right)\\ &+\tfrac{1}{n}\operatorname{tr}\left[\begin{pmatrix} q(\hat t,\hat x) & c(\hat t,\hat x,\hat y) \\ c^*(\hat t,\hat x,\hat y) & q(\hat t,\hat y)\end{pmatrix}\left(D^2z(\hat t,\hat x,\hat y)\right)^2\right] + \omega_{0,\delta}(h)\end{aligned}$$

where $A_t$ is defined in (3.4). Letting now $n\to\infty$, and using (3.5), we obtain

$$\begin{aligned}2C_0(\hat t - t_0) &\leq K\operatorname{tr}\left((q(\hat t,\hat x) + q(\hat t,\hat y) - 2c(\hat t,\hat x,\hat y))D^2\psi(\hat x-\hat y)\right)\\ &+\left(b(\hat t,\hat x) - b(\hat t,\hat y)\right)\cdot D\psi(\hat x-\hat y) + \omega_{0,\delta}(h)\end{aligned} \tag{3.24}$$

Since $\psi(x-y) = f(|x-y|)$, we have

$$D^2\psi(x-y) = \frac{f'(|x-y|)}{|x-y|}\left(I - \frac{x-y}{|x-y|}\otimes\frac{x-y}{|x-y|}\right) + f''(|x-y|)\frac{x-y}{|x-y|}\otimes\frac{x-y}{|x-y|} \tag{3.25}$$

hence

$$\begin{aligned}&\operatorname{tr}\left((q(\hat t,\hat x) + q(\hat t,\hat y) - 2c(\hat t,\hat x,\hat y))D^2\psi(\hat x-\hat y)\right)\\ &= \tfrac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}\operatorname{tr}(A(\hat t,\hat x,\hat y)) - \tfrac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}A(\hat t,\hat x,\hat y)\hat p\cdot\hat p + f''(|\hat x-\hat y|)A(\hat t,\hat x,\hat y)\hat p\cdot\hat p\end{aligned} \tag{3.26}$$

where $A(\hat t,\hat x,\hat y)=q(\hat t,\hat x)+q(\hat t,\hat y)-2c(\hat t,\hat x,\hat y)$ and $\hat p=\frac{\hat x-\hat y}{|\hat x-\hat y|}$. Let us give at this point a brief motivation for the choice of the coupling matrix $c(t,x,y)$. Since the previous matrix $A$ has to be nonnegative, we have $\mathrm{tr}(A)-A\hat p\cdot\hat p\geq 0$ so that, in order to handle the singular term $\frac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}$, we need that $\mathrm{tr}(A)$ be compensated by $A\hat p\cdot\hat p$, which is the case thanks to the choice (3.22). [1] Indeed, we have $\mathrm{tr}(A(\hat t,\hat x,\hat y))=4\lambda+\|\sigma(\hat t,\hat x)-\sigma(\hat t,\hat y)\|^2$ and $A(\hat t,\hat x,\hat y)\hat p\cdot\hat p\geq 4\lambda$ so that (recall that we assume $f'>0$)

$$\frac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}\,\mathrm{tr}(A(\hat t,\hat x,\hat y))-\frac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}A(\hat t,\hat x,\hat y)\hat p\cdot\hat p\leq\frac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}\,\|\sigma(\hat t,\hat x)-\sigma(\hat t,\hat y)\|^2 .$$

On the other hand, we are also interested in taking a function $f$ which is concave. Suppose now that this is the case, namely that $f''<0$. Then we obtain from (3.26)

$$\begin{aligned}&\mathrm{tr}\left((q(\hat t,\hat x)+q(\hat t,\hat y)-2c(\hat t,\hat x,\hat y))D^2\psi(\hat x-\hat y)\right)\\&\leq\tfrac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}\,\|\sigma(\hat t,\hat x)-\sigma(\hat t,\hat y)\|^2+4\lambda\,f''(|\hat x-\hat y|)\,.\end{aligned}\tag{3.27}$$

Combining this estimate with (3.24) we get

$$\begin{aligned}&2C_0(\hat t-t_0)\\&\leq K\left(4\lambda\,f''(|\hat x-\hat y|)+\tfrac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}\left(\|\sigma(\hat t,\hat x)-\sigma(\hat t,\hat y)\|^2+(b(\hat t,\hat x)-b(\hat t,\hat y))\cdot(\hat x-\hat y)\right)\right)+\omega_{0,\delta}(h)\end{aligned}$$

and using (3.7)

$$2C_0(\hat t-t_0)\leq K\left(4\lambda\,f''(|\hat x-\hat y|)+f'(|\hat x-\hat y|)\,g(|\hat x-\hat y|)\right)+\omega_{0,\delta}(h).\tag{3.28}$$

Choose now $f$ as the solution of

$$\begin{cases}4\lambda\,f''(r)+g(r)f'(r)=-1 & r\in(0,\delta]\,,\\ f(0)=0\,,\ f'(\delta)=0\end{cases}\tag{3.29}$$

which is possible since $\int_0^1 g(s)ds<\infty$. Namely we have

$$f(r)=\frac{1}{4\lambda}\int_0^r e^{-\frac{G(\xi)}{4\lambda}}\int_\xi^\delta e^{\frac{G(\tau)}{4\lambda}}d\tau d\xi\,,\qquad\text{where } G(\xi)=\textstyle\int_0^\xi g(\tau)d\tau\,.\tag{3.30}$$

Observe that $f'>0$ and $f''<0$ as we used above, moreover $f\in C^2((0,\delta])$. Using also that $\hat t>t_0/2$, we get

$$K\leq\omega_{0,\delta}(h)+2C_0(t_0-\hat t)<\omega_{0,\delta}(h)+C_0t_0$$

and since we have chosen $C_0=\frac{4\omega_{0,\delta}(u)}{t_0^2}$, we deduce

$$K<\frac{4\omega_{0,\delta}(u)}{t_0}+\omega_{0,\delta}(h).$$

Therefore, if $K$ is bigger than this bound we reach a contradiction and (3.16) cannot hold true. We conclude that (3.14) is verified provided that

$$K\geq\max\left\{\frac{4\omega_{0,\delta}(u)}{t_0}+\omega_{0,\delta}(h),\ \frac{\omega_{0,\delta}(u)}{f(\delta)}\right\}\,.\tag{3.31}$$

[1] We could also consider

$$c(t,x,y)=\sigma(t,x)\sigma(t,y)+\lambda\left(I-t\left(\frac{x-y}{|x-y|}\otimes\frac{x-y}{|x-y|}\right)\right),\quad\text{for any } t\in(0,2].$$

However the present choice $t=2$ as in [PW06] leads to a better constant in the gradient estimates.

From (3.14), choosing $t = t_0$ and letting $\varepsilon \to 0$, we obtain (3.15) and reversing the roles of $x$ and $y$, we find

$$|u(t_0, x) - u(t_0, y)| \le K\, f(|x-y|), \qquad |x-y| < \delta\,.$$

Using that $f$ is concave we then obtain

$$|u(t_0, x) - u(t_0, y)| \le K\, f'(0)|x-y|, \qquad |x-y| < \delta. \tag{3.32}$$

We fix now $K = \frac{4\omega_{0,\delta}(u)}{t_0} + \omega_{0,\delta}(h) + \frac{\omega_{0,\delta}(u)}{f(\delta)}$ so that (3.31) is satisfied. Since $f(\delta) \ge \frac{1}{8\lambda}\,\delta^2$ and $f'(0) \le \hat{c}_2\delta$, with $\hat{c}_2 = \frac{1}{4\lambda}\, e^{\frac{1}{4\lambda}\int_0^1 g(s)ds}$, we have

$$|u(t_0, x) - u(t_0, y)| \le \hat{c}_2\delta \left(\frac{4\omega_{0,\delta}(u)}{t_0 \wedge 1} + \omega_{0,\delta}(h) + \frac{8\lambda\omega_{0,\delta}(u)}{\delta^2}\right)|x-y|,$$

for every $x$, $y$: $|x-y| \le \delta$. Choosing $\delta = \sqrt{t_0 \wedge 1}$, we deduce

$$|u(t_0, x) - u(t_0, y)| \le \hat{c}_2\sqrt{t_0 \wedge 1}\left(\omega(t_0, u)\frac{4+8\lambda}{t_0 \wedge 1} + \omega(t_0, h)\right)|x-y|,$$

for every $x$, $y$: $|x-y| \le \sqrt{t_0 \wedge 1}$, and then for every $x$, $y$ such that $|x-y| \le 1$. The inequality then extends, in a standard way, to any $x$, $y \in \mathbb{R}^N$. This shows (3.8) and finishes the proof.

■

**Remark 3.6.** We point out from the above proof the estimate on the Lipschitz constant of $u(t_0)$ which is

$$Lip(u(t_0)) \le C_\lambda\, \delta \left(\frac{\omega_{0,\delta}(u)}{t_0} + \omega_{0,\delta}(h) + \frac{\omega_{0,\delta}(u)}{\delta^2}\right),$$

where $C_\lambda$ only depends on $\lambda$ and $g$. In particular, *any additional estimate on the oscillation of $u$ and $h$ can provide further specifications of such estimate* (see Theorem 3.24 below).

**Remark 3.7.** A global version of the previous result up to $t = 0$ is possible if $u_0$ is assumed to be Lipschitz continuous. In this case one can take $C_0 = 0$ and modify suitably the previous proof. One obtains

$$|u(t, x) - u(t, y)| \le K|x-y|, \quad t \in [0, T],$$

where $K = C(\lambda, g) \cdot \max\{Lip(u_0), T\text{osc}_{[0,T]}(h), \text{osc}_{[0,T]}(u)\}$. Note that here the global Lipschitz continuity of $u$ is established in terms of the (bounded) oscillation of $u$. Later on (cf. Section 3.3), assuming a further condition (at infinity) on the function $g$ appearing in (3.7), we will able to prove estimates on the oscillation of $u$ in order to get a complete Lipschitz bound of $u(t)$ only depending on the data $u_0$ and $h$.

**Remark 3.8.** Theorem 3.3 also provides an analytical proof of other regularity results existing in the literature which were only proved by probabilistic techniques. *For instance, our result applies to the case that $q_{ij}(t, \cdot) \in C^1(\mathbb{R}^N)$, $t \in (0, T)$, that (3.2) holds and if, setting $\sigma_0(t, x) = \sqrt{q(t, x)}$, we have*

$$2\|D_h\sigma_0(t, x)\|^2 + (b(t, x+h) - b(t, x)) \cdot h \le c|h|^2, \quad x, y \in \mathbb{R}^N,\ t \in (0, T),\ h \in \mathbb{R}^N. \tag{3.33}$$

(for a smooth $N \times N$ matrix $a(t, x)$, we write $D_h a(t, x)$ to denote the matrix $(D_h a_{ij}(t, x))$, where $D_h a_{ij}$ stands for the directional derivative along the direction $h$).

A time-independent version of (3.35), with some additional assumptions of polynomial growth on $b$ and $\sigma_0$, is used in [Ce96] and [Ce01, Chapter 1] to prove gradient estimates for the diffusion semigroup, with a probabilistic approach.

To see that (3.33) allows to apply our Theorem 3.3, we first note that we can take $\varphi(x) = 1 + |x|^2$ as Lyapunov function. Moreover, we set $\sigma_\mu(t,x) = \sqrt{q(t,x) - \mu I}$, $\mu = \lambda/2$, and we check that (3.7) holds for $\sigma_\mu$ with $g(r) = Kr$. To this purpose, since for every $h \in \mathbb{R}^N$, we have

$$\|\sigma_\mu(t,x+h) - \sigma_\mu(t,x)\|^2 \le \sum_{i,j=1}^N \int_0^1 \langle D\sigma_\mu^{ij}(t,x+rh), h\rangle^2 dr\,,$$

it is enough to show that

$$\sum_{i,j=1}^N \langle D\sigma_\mu^{ij}(t,x), h\rangle^2 = \|D_h\sigma_\mu(t,x)\|^2 \le 2\|D_h\sigma_0(t,x)\|^2, \tag{3.34}$$

so that (3.33) will imply the validity of (3.7).

To prove (3.34), we denote $D_h\sigma_\mu(t,x)$ by $\sigma'_\mu(t,x)$ and, similarly, $D_h\sigma_0(t,x)$ by $\sigma'_0(t,x)$ and $D_hq(t,x)$ by $q'(t,x)$. The next proof is inspired by the one of [Fr85, Theorem 2.1 in Section 3.2]. Note that the following computations hold even if $\mu$ is equal to 0. Since $\sigma_\mu(t,x)^2 = q(t,x) - \mu I$, differentiating with respect to $h$, we find

$$q'(t,x) = \sigma'_\mu(t,x)\sigma_\mu(t,x) + \sigma_\mu(t,x)\sigma'_\mu(t,x). \tag{3.35}$$

Take an orthogonal matrix $\theta(t,x)$ (independent of $\mu$) such that

$$\theta(t,x)\sigma_\mu(t,x)\theta^*(t,x) = \Lambda_\mu(t,x),$$

where $\Lambda_\mu(t,x)$ is diagonal, having $\sqrt{\lambda_i(t,x) - \mu}$, $i = 1,\dots,N$, on the diagonal. Now multiplying both sides of (3.35) on the left by $\theta(t,x)$ and on the right by $\theta(t,x)^*$ we get, setting $a(t,x) = \theta(t,x)q'(t,x)\theta(t,x)^*$,

$$a(t,x) = Y_\mu(t,x)\Lambda_\mu(t,x) + \Lambda_\mu(t,x)Y_\mu(t,x), \quad \text{where } Y_\mu(t,x) = \theta(t,x)\sigma'_\mu(t,x)\theta(t,x)^*.$$

It follows that $Y_\mu(t,x)$ has components

$$Y_\mu^{ij}(t,x) = \frac{a_{ij}(t,x)}{\sqrt{\lambda_i(t,x) - \mu} + \sqrt{\lambda_j(t,x) - \mu}}$$

so that, using that $\lambda_i(t,x) - \mu \ge \lambda_i(t,x)/2$, $i = 1,\dots,N$,

$$Y_\mu^{ij}(t,x)^2 \le 2\frac{a_{ij}^2(t,x)}{(\sqrt{\lambda_i(t,x)} + \sqrt{\lambda_j(t,x)})^2} = 2(Y_0^{ij}(t,x))^2. \tag{3.36}$$

Now note that $\sigma'_\mu(t,x) = \theta(t,x)^*Y_\mu(t,x)\theta(t,x)$.

Moreover, since $\theta(t,x)$ is orthogonal, we have $\|\sigma'_\mu(t,x)\|^2 = \|Y_\mu(t,x)\|^2$ even for $\mu = 0$. It follows that

$$\|\sigma'_\mu(t,x)\|^2 = \sum_{i,j=1}^N Y_\mu^{ij}(t,x)^2 \le 2\sum_{i,j=1}^N Y_0^{ij}(t,x)^2 = \|Y_0(t,x)\|^2 = \|\sigma'_0(t,x)\|^2$$

which proves (3.34).

Finally, let us provide a similar result concerning the case of Hölder regularity in which we relax the assumption $g \in C(0,1) \cap L^1(0,1)$ used in Theorem 3.3. We adopt the same notations as in the previous statement.

**Proposition 3.9.** *Assume the same conditions of Theorem 3.3, only replacing the hypothesis $g \in C(0,1) \cap L^1(0,1)$ with the following condition*

$$g \in C(0,1) \quad \text{and} \quad \lim_{s\to 0^+} sg(s) = 0. \tag{3.37}$$

*Let $u \in C(Q_T)$ be a viscosity solution of (3.1) as in Theorem 3.3. Then $u(t)$ is $\alpha$–Hölder continuous for any $\alpha \in (0,1)$ and verifies*

$$|u(t,x) - u(t,y)| \le \frac{\tilde c}{\lambda \alpha (1-\alpha)} \Big( \frac{1}{(t \wedge 1)^{\alpha/2}}\, \omega(t,u) + (t\wedge 1)^{1-\alpha/2} \omega(t,h) \Big) |x-y|^\alpha, \tag{3.38}$$

*$x, y \in \mathbb{R}^N$, $|x-y| \le 1$, $t \in (0,T)$, where $\tilde c$ only depends on $\alpha$, $\lambda$ and the modulus of continuity of $sg(s)$.*

*Moreover, if we replace (3.37) with the condition $g \in C(0,1)$ and $\limsup_{s\to 0^+} sg(s) < 4\lambda$, then there exists $\alpha = \alpha(g,\lambda) \in (0,1)$ such that (3.38) holds.*

*Proof.* We proceed as in the proof of Theorem 3.3 with few changes. First we fix $\delta_0 \in (0,1)$ such that $sg(s) < 2\lambda(1-\alpha)$ if $s < \delta_0$ and then we consider the set $\triangle(t_0, \delta)$ with $\delta < \delta_0$.

We also choose from the beginning $f(s) = s^\alpha$, and we take $K \ge \frac{\omega_{0,\delta}(u)}{\delta^\alpha}$. Being $f$ increasing and concave, we still arrive at the inequality (see (3.28))

$$2C_0(\hat t - t_0) \le K\big(4\lambda\, f''(|\hat x - \hat y|) + f'(|\hat x - \hat y|)\, g(|\hat x - \hat y|)\big) + \omega_{0,\delta}(h)$$

which now becomes

$$2C_0(\hat t - t_0) \le \alpha\, K\, |\hat x - \hat y|^{\alpha-2}\big(4\lambda\,(\alpha-1) + |\hat x - \hat y| g(|\hat x - \hat y|)\big) + \omega_{0,\delta}(h).$$

Since $|\hat x - \hat y| < \delta_0$, we get

$$2C_0(\hat t - t_0) \le 2\lambda\, \alpha(\alpha-1)\, K\, |\hat x - \hat y|^{\alpha-2} + \omega_{0,\delta}(h)$$

and so $2\lambda\, \alpha(1-\alpha)\, K\delta^{\alpha-2} \le \omega_{0,\delta}(h) + 2C_0(t_0 - \hat t) < \omega_{0,\delta}(h) + C_0 t_0$. We get a contradiction if

$$K \ge \frac{1}{2\lambda\, \alpha(1-\alpha)\delta^{\alpha-2}} \Big(\omega_{0,\delta}(h) + \frac{4\omega_{0,\delta}(u)}{t_0}\Big).$$

Choosing again $\delta = \sqrt{t_0 \wedge \delta_0^2}$ we get the desired estimate.

To prove the last assertion, we set $\gamma = \limsup_{s\to 0^+} sg(s)$ and we proceed similarly, but first we choose $\alpha \in (0,1)$ such that $4\lambda\,(\alpha-1) + \gamma < 0$, which is possible since $\gamma < 4\lambda$. ∎

We point out that (3.38) holds only for $x, y$ such that $|x-y| \le 1$. However, replacing $|x-y|^\alpha$ with $(|x-y|^\alpha + |x-y|)$, the estimate holds for every $x, y \in \mathbb{R}^N$, $t \in (0,T)$.

**Remark 3.10.** We discuss here the previous result when $b = 0 = h$. In such case assertion (3.37) is equivalent to say that $\sigma$ is uniformly continuous on $\mathbb{R}^N$ (uniformly in $t$). In general this is weaker than requiring that $q$ is uniformly continuous on $\mathbb{R}^N$ (cf. Remark 3.4). On the other hand, we recall that if $q$ is elliptic, *bounded, uniformly continuous and independent on time*, then gradient estimates hold for $u(t,\cdot)$ (see [St74, Theorem 6] whose proof uses the theory of analytic semigroups). However, in our case we do not know if gradient estimates hold assuming only (3.37) rather than $g \in L^1(0,1)$.

Note that condition (3.37) seems to be almost sharp to get Hölder continuity for any $\alpha \in (0,1)$. Indeed, in cases when $sg(s)$ is only bounded, it is known (see [KS80, Section 4] for the more general situation with $q_{ij}$ bounded and measurable) that any bounded regular solution $u(t,\cdot)$ of (3.1) is $\alpha$-Holder continuous on $\mathbb{R}^N$ only for *some* $\alpha = \alpha(\lambda, N) \in (0,1)$, with a universal Hölder constant depending on $\|u\|_\infty, \lambda$ and $N$.

Finally note that the assertion $\limsup_{s\to 0^+} sg(s) < 4\lambda$ holds in particular (with $b = h = 0$) if there exists $0 < \lambda < \Lambda$ such that

$$\lambda |h|^2 \le q(t,x)h\cdot h \le \Lambda |h|^2, \quad x, h \in \mathbb{R}^N, \quad t \in (0,T),$$

and the following Cordes type condition holds: $\frac{\Lambda}{\lambda} < \frac{N+4}{N}$. To this purpose take $\mu$ such that $0 < \mu < \lambda$ and $\frac{\Lambda}{\mu} < \frac{N+4}{N}$, and define a symmetric positive definite matrix $\sigma(t,x)$, such that $\sigma(t,x)^2 +$

$\mu I = q(t,x)$. We have $0 \le \big(\sigma(t,x)-\sigma(t,y)\big)^2 \le (\Lambda-\mu)I$ and so $\|\sigma(t,x)-\sigma(t,y)\|^2 = \mathrm{Tr}\big[(\sigma(t,x)-\sigma(t,y))^2\big] \le N(\Lambda-\mu)$. Note that we can take $g(r) = \frac{N(\Lambda-\mu)}{r}$ so that $\limsup_{s\to 0^+} sg(s) = N(\Lambda-\mu) < 4\lambda$ and the assertion follows.

3.2. **The Cauchy problem with bounded initial datum and potential term.** We use Theorem 3.3 and generalize [PW06, Theorem 3.4] concerning Lipschitz continuity of bounded solutions to the Cauchy problem on $Q_T$. Due to the fact that we are considering bounded solutions, we can treat more general equations having also a *possibly unbounded potential term* $V$*,* i.e.,

$$\partial_t u - \mathrm{tr}\,\big(q(t,x)D^2u\big) - b(t,x)\cdot Du + V(t,x)u = h(t,x) \qquad \text{in } Q_T. \tag{3.39}$$

We always assume Hypothesis 3.1 and that $V : Q_T \to \mathbb{R}$ *is continuous and non-negative.*

We start with a quite standard lemma.

**Lemma 3.11.** *Let $h \in C(Q_T)\cap L^\infty(Q_T)$. Let $u \in C\left(\bar Q_T\right)$ be a viscosity solution of (3.39) such that $u_0 := u(0,\cdot)$ is bounded on $\mathbb{R}^N$ and moreover $u$ is $o(\varphi)$ in $\bar Q_T$, where $\varphi$ satisfies (3.3). Then $u$ is bounded on $[0,T]\times\mathbb{R}^N$ and*

$$\|u(t)\|_\infty \le \|u_0\|_\infty + t\|h\|_{T,\infty}, \quad t\in[0,T], \tag{3.40}$$

*where $\|h\|_{T,\infty} = \sup_{t\in(0,T)}\|h(t)\|_\infty$.*

*Proof.* Let us consider, for $\epsilon > 0$,

$$\phi_\epsilon(t,x) = u(t,x) - \varepsilon\,\varphi(t,x) - \frac{\varepsilon}{T-t} - \|u_0\|_\infty - t\|h\|_{T,\infty},$$

where $\varphi$ is the Lyapunov function which we may assume to satisfy (3.5). If for any $\epsilon > 0$, we have $\phi_\epsilon(t,x) \le 0$, then, letting $\epsilon \to 0^+$, we deduce $u(t,x) \le \|u_0\|_\infty + t\|h\|_{T,\infty}$. Arguing by contradiction, suppose that, for some $\epsilon > 0$, $\max\limits_{[0,T)\times\mathbb{R}^N} \phi_\epsilon(t,x) > 0$. Let $(t_\epsilon, x_\epsilon)$ be the point where this maximum is attained. Note that $t_\epsilon \in (0,T)$ and so $(t_\epsilon,x_\epsilon)$ is a local maximum. By definition of subsolution we have

$$\|h\|_{T,\infty} + \frac{\varepsilon}{(T-t_\epsilon)^2} + (\partial_t - A_{t_\epsilon})\big(\varepsilon\,\varphi\big)(t_\epsilon,x_\epsilon) \le h(t_\epsilon,x_\epsilon) - V(t_\epsilon,x_\epsilon)u(t_\epsilon,x_\epsilon),$$

which yields a contradiction. Indeed $(\partial_t - A_{t_\epsilon})\varphi(t_\epsilon,x_\epsilon) \ge 0$ and $u(t_\epsilon,x_\epsilon) > 0$ imply

$$\|h\|_{T,\infty} + \frac{\varepsilon}{(T-t_\epsilon)^2} \le h(t_\epsilon,x_\epsilon).$$

Applying the same reasoning to $-u$, which is a subsolution to (3.39) with $h$ replaced by $-h$, we conclude. ∎

In the next result we will assume that the potential $V : Q_T \to \mathbb{R}_+$ satisfies, for every $x,y\in\mathbb{R}^N$, $|x-y|\le 1$, $t\in(0,T)$,

$$|V(t,x)-V(t,y)| \le k_0 + k_1|x-y|\cdot\max\big\{V(t,x),V(t,y)\big\}, \tag{3.41}$$

for some non-negative constants $k_0$ and $k_1$. This condition allows us to consider unbounded potentials and operators like $\triangle u - (|x|^2+|x|^\alpha)u$, $\alpha\in(0,1)$.

**Theorem 3.12.** *Assume that (3.2), (3.3) hold true, and that there exists a non-negative $g \in C(0,1)$ such that $\int_0^1 g(s)ds < \infty$ and (3.7) holds. Assume that $h\in C(Q_T)\cap L^\infty(Q_T)$ and that $V \in C(Q_T)$ is non-negative and satisfies (3.41). Let $u \in C(\bar Q_T)$ be a viscosity solution of (3.39) such that $u_0 := u(0,\cdot)$ is bounded on $\mathbb{R}^N$ and moreover $u$ is $o(\varphi)$ in $\bar Q_T$, where $\varphi$ satisfies (3.3).*

*Then $u(t) \in W^{1,\infty}(\mathbb{R}^N)$, for every $t \in (0,T)$, and there exists $C_0 = C_0(\lambda, g)$, such that*

(3.42)
$$\|Du(t)\|_\infty \le C_0 \left\{ \left[ \frac{1}{\sqrt{t\wedge 1}} + \sqrt{t\wedge 1}\, k_0 + k_1 \right] \left( \|u_0\|_\infty + (T \wedge \frac{3}{2}t)\|h\|_{T,\infty} \right) + \sqrt{t\wedge 1}\, \|h\|_{T,\infty} \right\}.$$

**Remark 3.13.** We compare now this result with [PW06, Theorem 3.4], where the assumption (3.7) was introduced and a similar gradient estimate was proved but using a probabilistic approach. In particular, the concept of solution of (3.1) used there is the probabilistic one given through the expectation of $u_0(X_t)$ where $X_t$ is the diffusion process associated to the operator $A_t$ in (3.4) and starting from $x \in \mathbb{R}^N$ (see Appendix A for more details). Similarly, assumption (3.3) is replaced there by the condition that $X_t$ is not explosive (observe that the existence of a Lyapunov function implies non-explosion of the diffusion process, see e.g. [SV79, Chapter 10]). Up to such changes due to the different settings, Theorem 3.12 provides an analytical proof of [PW06, Theorem 3.4] with a few more generality, since in the latter paper the coefficients were not supposed to be time-dependent, there was no source term $h$ and the potential $V$ was supposed to be either bounded or Lipschitz continuous, in which case a stronger hypothesis on $g$ was required.

We point out that (3.42) implies
$$\|Du(t)\|_\infty \lesssim C\, \frac{\|u_0\|_\infty}{\sqrt{t}} \qquad \text{as } t \to 0^+.$$
On the other hand, when $u_0 = 0$, (3.42) implies that, for small $t$,
$$\|Du(t)\|_\infty \le C\sqrt{t}\, \|h\|_{T,\infty}$$
which generalizes the classical gradient estimate for solutions to the Cauchy problem involving the inhomogeneous heat equation with zero initial condition.

*Proof of Theorem 3.12.* Using Lemma 3.11, we know that $u$ is bounded and moreover

(3.43)
$$\omega(t,u) := osc_{(\frac{t}{2}, T\wedge\frac{3}{2}t)}(u) \le 2 \sup_{s\in(\frac{t}{2}, T\wedge\frac{3}{2}t)} \|u(s)\|_\infty \le 2\|u_0\|_\infty \,+\, 2(T \wedge \frac{3}{2}t)\|h\|_{T,\infty}.$$

In particular, if $V = 0$ the conclusion is a direct consequence of Theorem 3.3.

In the case $V \ne 0$, we modify a bit the proof of Theorem 3.3, choosing $f$ as the solution of

(3.44)
$$\begin{cases} 4\lambda\, f''(r) + g(r) f'(r) = -1 & r \in (0, 2\delta]\,, \\ f(0) = 0\,,\ \ f'(2\delta) = 0, \end{cases}$$

where now $\delta \le \frac{1}{2}$. Therefore we have
$$f(r) = \frac{1}{4\lambda} \int_0^r e^{-\frac{G(\xi)}{4\lambda}} \int_\xi^{2\delta} e^{\frac{G(\tau)}{4\lambda}} d\tau d\xi\,, \qquad \text{where } G(\xi) = \textstyle\int_0^\xi g(\tau)d\tau\,.$$
In particular, note that, for every $r \in (0,\delta)$ we have

(3.45)
$$f(r) \ge f'(\delta) r \ge \frac{\delta}{4\lambda} r\,.$$

Following the same proof of Theorem 3.3 (with the same notations) we obtain (see (3.28)) that

(3.46)
$$\begin{aligned} 2C_0(\hat t - t_0) \le K\,(4\lambda\, f''(|\hat x - \hat y|) + f'(|\hat x - \hat y|)\, g(|\hat x - \hat y|)) + \omega_{0,\delta}(h) \\ + \big(V(\hat t, \hat y) u(\hat t, \hat y) - V(\hat t, \hat x) u(\hat t, \hat x)\big). \end{aligned}$$

We concentrate on the last term setting

$$J(\hat{t},\hat{x},\hat{y}) = V(\hat{t},\hat{y})u(\hat{t},\hat{y}) - V(\hat{t},\hat{x})u(\hat{t},\hat{x}).$$

Assume that $V(\hat{t},\hat{y}) \geq V(\hat{t},\hat{x})$; then we write[2]

$$J(\hat{t},\hat{x},\hat{y}) = V(\hat{t},\hat{y})\big(u(\hat{t},\hat{y}) - u(\hat{t},\hat{x})\big) + u(\hat{t},\hat{x})\big(V(\hat{t},\hat{y}) - V(\hat{t},\hat{x})\big)$$

and since $u(\hat{t},\hat{y}) - u(\hat{t},\hat{x}) < -Kf(|\hat{x}-\hat{y}|)$, which is consequence of (3.16), it follows, using also (3.41), that

$$J(\hat{t},\hat{x},\hat{y}) \leq -KV(\hat{t},\hat{y})f(|\hat{x}-\hat{y}|) + \|u(\hat{t})\|_\infty \left(k_0 + k_1|\hat{x}-\hat{y}|\, V(\hat{t},\hat{y})\right).$$

Since $|\hat{x}-\hat{y}| < \delta$, using (3.45) we deduce

$$J(\hat{t},\hat{x},\hat{y}) \leq -K\,\frac{\delta}{4\lambda} V(\hat{t},\hat{y})|\hat{x}-\hat{y}| + \|u(\hat{t})\|_\infty \left(k_0 + k_1|\hat{x}-\hat{y}|\, V(\hat{t},\hat{y})\right).$$

Therefore, if $K \geq k_1\|u(\hat{t})\|_\infty \frac{4\lambda}{\delta}$ we get

$$J(\hat{t},\hat{x},\hat{y}) \leq k_0\, \|u(\hat{t})\|_\infty$$

so that (3.46) implies

$$2C_0(\hat{t}-t_0) + K \leq 2\|h\|_{T,\infty} + k_0\, \|u(\hat{t})\|_\infty.$$

Finally, we obtain here a contradiction if we choose

$$K \geq \max\left\{ \frac{4\omega_{0,\delta}(u)}{t_0} + 2\|h\|_{T,\infty} + k_0\, \|u(\hat{t})\|_\infty\, , k_1\|u(\hat{t})\|_\infty \frac{4\lambda}{\delta}\, ,\, \frac{\omega_{0,\delta}(u)}{f(\delta)} \right\}.$$

Using that $f(\delta) \geq \frac{\delta^2}{4\lambda}$ and setting $\delta = \sqrt{t_0} \wedge \frac{1}{2}$ we find

$$|u(t_0,x) - u(t_0,y)| \leq Kf(|x-y|) \leq Kf'(0)|x-y| \leq c_2K\delta|x-y|$$

$$\leq C_0\Big(\frac{4\omega_{0,\delta}(u)}{\sqrt{t_0\wedge 1}} + \sqrt{t_0 \wedge 1}\left(\|h\|_{T,\infty} + k_0\|u(\hat{t})\|_\infty\right) + 4\lambda k_1\|u(\hat{t})\|_\infty + 4\lambda\frac{\omega_{0,\delta}(u)}{\sqrt{t_0\wedge 1}}\Big)\, |x-y|.$$

Recall that (3.43) implies $\omega_{0,\delta}(u) \leq 2\|u_0\|_\infty \,+\, 2(T\wedge \frac{3}{2}t_0)\|h\|_{T,\infty}$, hence we conclude

$$\begin{aligned}|u(t_0,x) - u(t_0,y)| \;\; &\leq \Big\{C(\lambda,g)\left(\tfrac{1}{\sqrt{t_0\wedge 1}} + k_0\sqrt{t_0\wedge 1} + k_1\right)\left[\|u_0\|_\infty \,+ (T\wedge \tfrac{3}{2}t_0)\|h\|_{T,\infty}\right] \\ &\quad + C(\lambda,g)\,\sqrt{t_0\wedge 1}\,\|h\|_{T,\infty}\Big\}\, |x-y|\,.\end{aligned}$$

■

**Remark 3.14.** It is not difficult to modify the previous proof and obtain gradient estimates, assuming instead of (3.7), the following more general assumption

$$\tfrac{1}{|x-y|}\left(\|\sigma(t,x)-\sigma(t,y)\|^2 + (b(t,x)-b(t,y))\cdot(x-y)\right)$$

$$\leq g(|x-y|) + k_3|x-y|\max(V(t,x),V(t,y)),$$

for $x,y \in \mathbb{R}^N$ such that $|x-y| \leq \delta_1$, $t\in(0,T)$, for some $\delta_1, k_3 > 0$. This condition is comparable with the assumption $Db(x)h\cdot h \leq (s_1 + s_2V(x))|h|^2$, $x,h \in \mathbb{R}^N$, for some $s_1, s_2 > 0$ used in [BF04]. According to [BF04] the last hypothesis together with $|DV(x)| \leq c_0 + c_1V(x)$, $x \in \mathbb{R}^N$, and $q_{ij}$ Lipschitz continuous imply gradient estimates (3.42) for the autonomous non-degenerate Cauchy problem with $h = 0$.

[2]In the case that $V(\hat{t},\hat{x}) \geq V(\hat{t},\hat{y})$, we use the identity $J(\hat{t},\hat{x},\hat{y}) = (V(\hat{t},\hat{y}) - V(\hat{t},\hat{x}))u(\hat{t},\hat{y}) + (u(\hat{t},\hat{y}) - u(\hat{t},\hat{x}))V(\hat{t},\hat{x})$ and then we proceed in a similar way.

**Remark 3.15.** Theorem 3.12 can be also interpreted as an a priori estimate on classical bounded solutions to the Cauchy problem involving $A_t$. On this respect (see [KLL10, Section 2] and [MPW02, Section 4] for the autonomous case), if we assume Hypothesis 3.1, existence of a Lyapunov function $\varphi$ and the fact that coefficients $q_{ij}$, $b_i$ belong to $C^{\alpha/2,\alpha}_{loc}(\bar{Q}_T)$, for some $\alpha \in (0,1)$, then it is well-known that there exists a unique *bounded classical solution* to the Cauchy problem (3.39) with $h = V = 0$ and $u_0$ which is continuous and bounded on $\mathbb{R}^N$.

We finish the section with a variant of Theorem 3.12 establishing Hölder continuity of the solution. In this case we may assume that the potential $V$ satisfies, for every $x, y \in \mathbb{R}^N$, $|x-y| \le 1$, $t \in (0,T)$,

$$|V(t,x) - V(t,y)| \le k_0 + k_1|x-y|^\alpha \cdot \max\big\{V(t,x), V(t,y)\big\}, \tag{3.47}$$

for some $\alpha \in (0,1)$.

**Proposition 3.16.** *Assume that (3.2), (3.3) hold true, and that (3.7) holds for some $g$ satisfying (3.37). Assume that $h \in C(Q_T) \cap L^\infty(Q_T)$ and that $V \in C(Q_T)$ is nonnegative and satisfies (3.47). If $u \in C(\bar{Q}_T)$ is a viscosity solution of (3.1) as in Theorem 3.12, then there exists $C_0 = C_0(\lambda, \alpha, g)$ such that*

$$\begin{aligned}|u(t,x) - u(t,y)| \le C_0 \bigg\{ \Big[\frac{1}{(t\wedge 1)^{\alpha/2}} + (t\wedge 1)^{1-\alpha/2}\, k_0 + k_1\Big] \big(\|u_0\|_\infty + (T \wedge \tfrac{3}{2}t)\|h\|_{T,\infty}\big) \\ +(t\wedge 1)^{1-\alpha/2}\|h\|_{T,\infty}\bigg\}\ |x-y|^\alpha,\end{aligned} \tag{3.48}$$

*$x, y \in \mathbb{R}^N$, $|x-y| \le 1$, $t \in (0,T)$, where $C_0$ only depends on $\alpha$, $\lambda$ and the modulus of continuity of $sg(s)$.*

*Proof.* Now we follow the proof of Proposition 3.9. We fix $\delta_0 \in (0,1)$ such that $sg(s) < 2\lambda(1-\alpha)$ if $s < \delta_0$ and then we consider the set $\triangle(t_0, \delta)$ with $\delta < \delta_0$. We obtain

$$2C_0(\hat{t} - t_0) \le \alpha\, K\, |\hat{x} - \hat{y}|^{\alpha-2}\big(4\lambda\,(\alpha-1) + |\hat{x}-\hat{y}| g(|\hat{x}-\hat{y}|)\big) + \omega_{0,\delta}(h) + J(\hat{t}, \hat{x}, \hat{y}),$$

which implies, since $|\hat{x} - \hat{y}| < \delta_0$,

$$2C_0(\hat{t} - t_0) \le 2\lambda\, \alpha(\alpha-1)\, K\, |\hat{x}-\hat{y}|^{\alpha-2} + 2\|h\|_{T,\infty} + J(\hat{t}, \hat{x}, \hat{y})\,.$$

We estimate last term as in Theorem 3.12 and using (3.47) we get

$$J(\hat{t}, \hat{x}, \hat{y}) \le -K\, |\hat{x}-\hat{y}|^\alpha V(\hat{t}, \hat{y}) + \|u(\hat{t})\|_\infty \left(k_0 + k_1 |\hat{x}-\hat{y}|^\alpha\, V(\hat{t}, \hat{y})\right).$$

In particular, if $K \ge k_1\, \|u(\hat{t})\|_\infty$ we conclude

$$2C_0(\hat{t} - t_0) \le 2\lambda\, \alpha(\alpha-1)\, K\, |\hat{x}-\hat{y}|^{\alpha-2} + 2\|h\|_{T,\infty} + k_0\, \|u(\hat{t})\|_\infty\,,$$

and we obtain a contradiction if

$$K \ge \frac{1}{2\lambda\, \alpha(1-\alpha)\delta^{\alpha-2}}\Big(2\|h\|_{T,\infty} + k_0\, \|u(\hat{t})\|_\infty + \frac{4\omega_{0,\delta}(u)}{t_0}\Big).$$

Putting together all the conditions required on $K$, we need to choose $K$ such that

$$K \ge \max\left\{\frac{1}{2\lambda\, \alpha(1-\alpha)\delta^{\alpha-2}}\Big(2\|h\|_{T,\infty} + k_0\, \|u(\hat{t})\|_\infty + \frac{4\omega_{0,\delta}(u)}{t_0}\Big),\ k_1\, \|u(\hat{t})\|_\infty\,,\ \frac{\omega_{0,\delta}(u)}{\delta^\alpha}\right\}.$$

Choosing $\delta = \sqrt{t_0} \wedge \delta_0$, and using the bounds on $u$ (see (3.40)), we can choose

$$\begin{aligned}K =\ & C_0 \bigg\{\Big[\tfrac{1}{(\sqrt{t_0}\wedge 1)^{\frac{\alpha}{2}}} + (\sqrt{t_0 \wedge 1})^{1-\frac{\alpha}{2}}\, k_0 + k_1\Big]\ \big(\|u_0\|_\infty\ +\ (T\wedge \tfrac{3}{2}t_0)\|h\|_{T,\infty}\big) \\ &+(\sqrt{t_0 \wedge 1})^{1-\frac{\alpha}{2}}\ \|h\|_{T,\infty}\bigg\},\end{aligned}$$

for some $C_0 = C_0(\lambda, g, \alpha)$. In this way we have proved (3.48). ∎

We finish the section by noting that all the previous results can be further improved in dimension $N = 1$. Indeed in this case we only need that $q$ is uniformly positive and we do not need any "control" for $\|\sigma(t,x) - \sigma(t,y)\|$. When $N = 1$ Theorems 3.3, 3.12, Propositions 3.9 and 3.16 hold replacing (3.7) with

$$(b(t,x) - b(t,y)) \cdot (x - y) \leq g(|x-y|)|x-y|, \quad x, y \, : \, 0 < |x-y| \leq 1, \tag{3.49}$$

$t \in (0,T)$. We only formulate an analogous of Theorem 3.3.

**Proposition 3.17.** *Let $N = 1$. Assume that (3.2), (3.3) hold true, and, in addition, that there exists $g \in C(0,1;\mathbb{R}_+)$ such that $\int_0^1 g(s)ds < \infty$ and (3.49) holds. Let $u \in C(Q_T)$ be a viscosity solution of (3.1). Moreover, suppose that $u$ and $h$ have bounded oscillation (see (2.1)).*

*Then the function $u(t)$ is Lipschitz continuous and estimate (3.8) holds.*

*Proof.* We follow the argument of the proof of Theorem 3.3 and arrive at the identity (3.26). Now note that $\mathrm{tr}\big(A(\hat{t},\hat{x},\hat{y})\big) - A(\hat{t},\hat{x},\hat{y})\hat{p}\cdot\hat{p} = 0$. Using that $A(\hat{t},\hat{x},\hat{y})\hat{p}\cdot\hat{p} \geq 4\lambda$, we obtain

$$(q(\hat{t},\hat{x}) + q(\hat{t},\hat{y}) - 2c(\hat{t},\hat{x},\hat{y}))D^2\psi(\hat{x}-\hat{y}) \leq 4\lambda\, f''(|\hat{x}-\hat{y}|)$$

and so (cf. (3.24)) $2C_0(\hat{t} - t_0) \leq 4\lambda K\, f''(|\hat{x}-\hat{y}|) + \big(b(\hat{t},\hat{x}) - b(\hat{t},\hat{y})\big)\cdot D\psi(\hat{x}-\hat{y}) + \omega_{0,\delta}(h)$. Using the assumption (3.49) we can continue as in the previous proof and get the assertion. ∎

**Remark 3.18.** In [BF04] it is considered a one-dimensional operator $A_t = A$, $Au = u'' + b_0(x)u'$, where $b_0$ does not satisfy (3.49) and for which uniform gradient estimates for the associated parabolic Cauchy problem do not hold.

3.3. **Unbounded data: estimates on the oscillation of the solutions.** In this section we generalize the gradient estimate (3.42) for the operator (3.1) to the case in which the data $u_0$, $h$ are not necessarily bounded but, more generally, have bounded oscillation (see (2.1) and (2.3)). In particular, we cover the case of (possibly unbounded) Lipschitz and Hölder continuous $u_0$.

With respect to the assumptions in Theorem 3.12, we need to assume an additional control at infinity on the function $g$ appearing in (3.7), namely that $g \in C(0,+\infty)$ and $g(r)$ is $O(r)$ as $r \to +\infty$. We stress that this condition is always satisfied whenever $\sigma(t,x)$ has bounded oscillation and $b(t,x)$ verifies $(b(t,x) - b(t,y))\cdot(x-y) \leq C|x-y|^2$, for any $x, y \in \mathbb{R}^N$, $t \in [0,T]$ (notice that in this case $\varphi = 1 + |x|^2$ is a Lyapunov function). However, the operator in (3.10) also satisfies our condition at infinity for $g$. Roughly speaking, $g = O(r)$ as $r \to +\infty$ should be regarded as a control on the superlinear growth of the coefficients which is not compensated by the interaction drift–diffusion.

In order to bound the oscillation of the solutions, we will need the following growth type estimates which may be of independent interest. This is why we give them in a rather sharp form (distinguishing between the Hölder and the Lipschitz case, as in the previous subsections, according to the behaviour of $g$ as $r \to 0^+$).

The result explains that if a viscosity solution $u$ of (3.1) is $o(\varphi)$ on $\bar{Q}_T$ and, moreover, $u_0$, $h$ have bounded oscillation, then $u(t)$ also has bounded oscillation with a precise control which may imply, in particular, the conservation of the Lipschitz and Hölder continuity of $u_0$. We mention that a related result for equations like $\partial_t u + F(\nabla u, \nabla^2 u)$ is given in [GGIS91, Proposition 2.3].

**Lemma 3.19.** *Assume that (3.2) and (3.3) hold true and that (3.7) holds in $Q_T$ with some non-negative $g \in C(0,+\infty)$ such that $g(r)$ is $O(r)$ as $r \to +\infty$. In addition, assume that $u_0$, $h$ satisfy*

$$|u_0(x) - u_0(y)| \leq k_0 + k_\alpha|x-y|^\alpha + k_1|x-y|, \;\; x, y \in \mathbb{R}^N, \tag{3.50}$$

(3.51) $$|h(t,x)-h(t,y)| \le h_0 + h_\alpha|x-y|^\alpha + h_1|x-y|, \quad x,y \in \mathbb{R}^N,\, t\in(0,T).$$

*for some $\alpha \in (0,1)$, with $k_0, k_\alpha, k_1, h_0, h_\alpha, h_1 \ge 0$. Let $u \in C\left(\bar{Q}_T\right)$ be a viscosity solution of (3.1) such that $u$ is $o(\varphi)$ in $\bar{Q}_T$. Then we have:*

**(i)** *If $rg(r) \to 0$ as $r \to 0^+$, we have*

$$|u(t,x)-u(t,y)| \le k_0 + (K+LMt)|x-y|^\alpha + (k_1 + Mt(h_1+k_1))|x-y|\,, \quad x,y\in\mathbb{R}^N,\, t\in[0,T],$$

*where $K = \max(\frac{h_0}{2\alpha\lambda(1-\alpha)}, k_\alpha, k_1)$, $L = \max(h_0, h_\alpha, h_1, k_\alpha, k_1)$ and $M = M(g,\lambda,\alpha,T)$.*

**(ii)** *If $g \in L^1(0,1)$ (and, in case $k_\alpha \neq 0$, if also $rg(r) \to 0$ as $r \to 0^+$), we have*

$$|u(t,x)-u(t,y)| \le k_0 + (k_\alpha + L\,k_\alpha\,t)|x-y|^\alpha + c_0(\max\{h_0,h_\alpha,k_1\} + MLt)|x-y|, \quad x,y\in\mathbb{R}^N,\, t\in[0,T],$$

*where $L = L(g,\lambda,\alpha,T)$, $M = \max(h_0,h_\alpha,h_1,k_1)$, $c_0 = c_0(g,\lambda,\alpha)$.*

**Remark 3.20.** Let us comment on the technical form of estimates (i) and (ii) above. Such form is meant to show that, if $k_0 = 0$, then (i) or (ii) imply the conservation of the Hölder, respectively Lipschitz, continuity from $u_0$ to $u(t)$. To our knowledge, no similar results are available in the literature unless the coefficients $q_{ij}$ and $b_i$ are assumed to be Lipschitz continuous.

In particular, notice that if $k_0 = k_1 = 0$ and if $h_1 = 0$, estimate (i) implies a global Hölder estimate for $u(t)$, and in this case $u(t)$ grows sublinearly at infinity. Similarly, (ii) shows that if $k_0 = k_\alpha = 0$, then $u(t)$ is globally Lipschitz continuous. The reader should keep in mind that here we are not dealing with the regularizing effect, but with merely conservation of the growth and continuity properties from $u_0$ to $u$. On the other hand, the regularizing effect will follow coupling together the above oscillation estimates with Theorem 3.3 and will be stated later (see Theorem 3.24).

*Proof.* According to Remark 3.2, there is no loss of generality in taking $M = 0$ in (3.3). Let us consider the set

(3.52) $$\Delta = \{(t,x,y) \in (0,T')\times\mathbb{R}^N\times\mathbb{R}^N\},$$

where $T' \in [0,T]$ will be chosen later, and the function

(3.53) $$\Phi_\epsilon(t,x,y) = u(t,x)-u(t,y)-f(t,|x-y|)-\varepsilon\left(\varphi(t,x)+\varphi(t,y)\right)-\frac{\varepsilon}{T'-t}\,, \quad \epsilon>0,$$

where

$$f(t,r) = (k_0+at)+(\beta+bt)r^\alpha+(\gamma+ct)(r+\tilde{f}(r)),$$

with $a,b,c,\beta,\gamma \ge 0$ and $\tilde{f}(r) \in C^2(0,\infty)$ is a nondecreasing concave function with $\tilde{f}(0) = 0$, to be fixed later. As usual, we wish to prove that, independently on $\varepsilon$,

(3.54) $$\Phi_\epsilon(t,x,y) \le 0, \quad (t,x,y)\in\Delta\,,$$

and we argue by contradiction, assuming that $\sup_\Delta \Phi_\epsilon(t,x,y) > 0$. Since $\varphi$ blows-up at infinity and $u$ is $o(\varphi)$, we have

$$\Phi_\epsilon \to -\infty \qquad \text{as } |x|\to\infty \text{ or } |y|\to\infty$$

hence $\Phi_\epsilon$ has a global maximum at some point $(\hat{t},\hat{x},\hat{y}) \in \overline{\Delta}$ and clearly $\hat{x} \neq \hat{y}$ since the maximum is positive, and $\hat{t} < T'$. Assuming that

(3.55) $$\beta \ge k_\alpha\,, \quad \gamma \ge k_1,$$

we deduce that the maximum cannot be reached at $t = 0$ because, thanks to (3.50) and (3.55), we have

$$u_0(x)-u_0(y) \le f(0,|x-y|), \quad x,y\in\mathbb{R}^N.$$

It follows that $(\hat t,\hat x,\hat y)$ is also a local maximum in which $\Phi_\epsilon$ is positive. Henceforth, proceeding like in Theorem 3.7 (using (3.5), and that $f'\geq 0$ and $f''\leq 0$ (we have set $\partial_x f = f'$ and $\partial^2_{xx} f = f''$), we obtain

$$\begin{aligned}&\tfrac{\varepsilon}{(T'-\hat t)^2} + (a + b|\hat x-\hat y|^\alpha + c(|\hat x-\hat y| + \tilde f(|\hat x-\hat y|))\\ &\leq \left(4\lambda\, f''(\hat t,|\hat x-\hat y|) + f'(\hat t,|\hat x-\hat y|)\, g(|\hat x-\hat y|)\right) + |h(\hat t,\hat x) - h(\hat t,\hat y)|.\end{aligned}$$

Setting $\hat r = |\hat x-\hat y|$, using the expression of $f(t,r)$ and (3.51), we get

(3.56)

$$\begin{aligned}&\tfrac{\varepsilon}{(T'-\hat t)^2} + (a + b\hat r^\alpha + c(\hat r + \tilde f(\hat r)))\\ &\leq (\beta + b\hat t)\,\alpha\left[4\lambda(\alpha-1)\hat r^{\alpha-2} + \hat r^{\alpha-1} g(\hat r)\right] + (\gamma + c\hat t) g(\hat r) + (\gamma + c\hat t)\left[4\lambda\tilde f''(\hat r) + g(\hat r)\tilde f'(\hat r)\right]\\ &\quad + h_0 + h_\alpha \hat r^\alpha + h_1\hat r\end{aligned}$$

We distinguish now according to the assumption on $g(s)$ near $s=0$.

**(i)** Assume that $g(s)s\to 0$ as $s\to 0$. We take here $\tilde f = 0$ and we assume that $b\geq c$, $\beta\geq\gamma$; in particular we obtain

$$\begin{aligned}&(\beta + b\hat t)\,\alpha\left[4\lambda(\alpha-1)\hat r^{\alpha-2} + \hat r^{\alpha-1} g(\hat r)\right] + (\gamma + c\hat t) g(\hat r)\\ &\leq (\beta + b\hat t)\left\{\alpha\left[4\lambda(\alpha-1)\hat r^{\alpha-2} + \hat r^{\alpha-1} g(\hat r)\right] + g(\hat r)\right\}.\end{aligned}$$

Since $sg(s)\to 0$ as $s\to 0$, there exists $r_0<1$ (only depending on $g$, $\lambda$, $\alpha$) such that

$$\alpha\left[4\lambda(\alpha-1)s^{\alpha-2} + s^{\alpha-1} g(s)\right] + g(s) \leq 2\alpha\lambda(\alpha-1)s^{\alpha-2}, \qquad s\in(0,r_0)\,. \tag{3.57}$$

Since $g(s) = O(s)$ as $s\to+\infty$, for $s\geq r_0$ there exists a constant $L_0$ such that $g(s)\leq L_0 s$, hence

$$(\beta + b\hat t)\,\alpha\left[4\lambda(\alpha-1)s^{\alpha-2} + s^{\alpha-1} g(s)\right] + (\gamma + c\hat t) g(s) \leq (\beta + b\hat t)\,\alpha L_0\, s^\alpha + (\gamma + c\hat t) L_0 s, \quad s\geq r_0.$$

Therefore, we conclude from (3.56) (where $\tilde f = 0$) that

$$\frac{\varepsilon}{(T'-\hat t)^2} + (a + b\hat r^\alpha + c\hat r) \leq -2\alpha\lambda(1-\alpha)\beta\,\hat r^{\alpha-2}\chi_{r<r_0} + (\beta + b\hat t)\,\alpha L_0\,\hat r^\alpha + (\gamma + c\hat t) L_0\hat r + h_0 + h_\alpha\hat r^\alpha + h_1\hat r\,.$$

We choose $T' = 1/2L_0$, so that $L_0 t\leq\frac12$ and we deduce

$$\frac{\varepsilon}{(T'-\hat t)^2} + a + (\frac12 b - L_0\beta)\hat r^\alpha + (\frac12 c - L_0\gamma)\hat r \leq -2\alpha\lambda(1-\alpha)\beta\,\hat r^{\alpha-2}\chi_{r<r_0} + h_0 + h_\alpha\hat r^\alpha + h_1\hat r\,. \tag{3.58}$$

Since $h_0\leq h_0\hat r^{\alpha-2}\chi_{(0,r_0)}(\hat r) + \frac{1}{r_0^\alpha}\,h_0\hat r^\alpha\chi_{r\geq r_0}(\hat r)$, we obtain

$$\begin{aligned}&\tfrac{\varepsilon}{(T'-\hat t)^2} + a + \left(\tfrac12 b - L_0\beta\right)\hat r^\alpha + (\tfrac12 c - L_0\gamma)\hat r\\ &\leq (h_0 - 2\alpha\lambda(1-\alpha)\beta)\hat r^{\alpha-2}\chi_{(0,r_0)}(\hat r) + (\tfrac{1}{r_0^\alpha} h_0 + h_\alpha)\hat r^\alpha + h_1\hat r\,.\end{aligned}$$

Here we choose $a=0$, $\beta\geq\frac{h_0}{2\alpha\lambda(1-\alpha)}$, $b\geq 2(L_0\beta + 2\frac{\max\{h_0,h_\alpha\}}{r_0^\alpha})$, $\gamma = k_1$, $c = 2(L_0\gamma + h_1)$ and we get a contradiction.

Recalling that $\beta\geq\gamma$, $b\geq c$ and (3.55) were used before, we have just proved that

$$u(t,x) - u(t,y)\leq k_0 + (\beta + bt)|x-y|^\alpha + (k_1 + ct)|x-y|, \qquad t\in[0,T'] \text{ and } x,y\in\mathbb{R}^N, \tag{3.59}$$

where $\beta = \max(\frac{h_0}{2\alpha\lambda(1-\alpha)}, k_\alpha, k_1)$, $c = 2(L_0k_1 + h_1)$ and

$$b = \max\left\{2\left(L_0\beta + 2\frac{\max\{h_0,h_\alpha\}}{r_0^\alpha}\right),\ 2(L_0k_1 + h_1)\right\}.$$

In particular we have $u(T',x) - u(T',y)\leq k_0 + (\beta + bT')|x-y|^\alpha + 2(k_1 + h_1T')|x-y|$, $x,y\in\mathbb{R}^N$. Since $T'$ only depends on $\lambda,\alpha$, $g$, and not on the data $h$, $u_0$, we can restart the same method on the interval $I = [T', (2T' \wedge T)]$ (here it is enough to use $f(t,r) = k_0 + b_1 t\, r^\alpha + c_1 tr$, for suitable constants $b_1$ and $c_1$).

Iterating this argument we obtain the assertion on the whole $[0,T]$. In particular, we have proved, for any $x, y \in \mathbb{R}^N$,

$$u(t,x) - u(t,y) \le k_0 + (K + LMt)|x-y|^\alpha + (k_1 + M(k_1+h_1)t)|x-y|, \tag{3.60}$$

where $K = \beta = \max(\frac{h_0}{2\alpha\lambda(1-\alpha)}, k_\alpha, k_1)$, $L = \max(h_0, h_\alpha, h_1, k_\alpha, k_1)$ and $M = M(g,\lambda,\alpha,T)$.

Note that an alternative estimate can be obtained after (3.58) if we choose $a = h_0$, $b \ge 2(L_0\beta + h_\alpha)$ and $c = 2(L_0\gamma + h_1)$. In this way we get the estimate (recall that $\beta \ge \gamma = k_1$ and $b \ge c$)

$$|u(t,x) - u(t,y)| \le (k_0 + h_0 t) + (\max\{k_\alpha, k_1\} + bt)|x-y|^\alpha + (k_1 + ct)|x-y|, \tag{3.61}$$

for every $t \in [0,T']$ and $x, y \in \mathbb{R}^N$, where $b = 2(L_0 \max\{k_\alpha, k_1\} + \max\{h_\alpha, h_1\})$ and $c = 2(L_0 k_1 + h_1)$. Such a choice better points out the continuity as $t \to 0$ but, in case $k_0 = 0$, (3.60) is preferable to show the continuity properties of $u$.

**(ii)** Assume now in addition that $g \in L^1(0,1)$. Since $g(r) = O(r)$ as $r \to +\infty$, we may deduce that there exist $m > 0$ and a continuous function $\tilde g(r)$ such that

$$g(r) \le mr + \tilde g(r) \qquad \forall r > 0\,, \quad \tilde g \in L^1(0,\infty)\,.$$

Without loss of generality we can also assume that $1 \le \tilde g(r)$ if $r \le 1$, hence taking $\gamma \ge \max\{h_0, h_\alpha\}$ implies

$$h_0 + h_\alpha r^\alpha + h_1 r \le (h_0 + h_\alpha)\tilde g(r) + (h_0 + h_\alpha + h_1) r \le 2\gamma \tilde g(r) + (h_0 + h_\alpha + h_1) r, \ \ r > 0.$$

Therefore, we deduce from (3.56)

$$\begin{aligned} &\tfrac{\varepsilon}{(T'-\hat t)^2} + (a + b\hat r^\alpha + c(\hat r + \tilde f(\hat r))) \le (\beta + b\hat t)\,\alpha \left[4\lambda(\alpha-1)\hat r^{\alpha-2} + \hat r^{\alpha-1} g(\hat r)\right] \\ &+ (\gamma + c\hat t)(1 + \tilde f'(\hat r)) m\hat r + (\gamma + c\hat t)\left[4\lambda \tilde f''(\hat r) + \tilde g(\hat r)\tilde f'(\hat r) + 3\tilde g(\hat r)\right] + (h_0 + h_\alpha + h_1)\hat r\,. \end{aligned} \tag{3.62}$$

We define now $\tilde f(r)$ as the solution of the ODE

$$\begin{cases} 4\lambda\, \tilde f''(r) + \tilde g(r)\tilde f'(r) + 3\tilde g(r) = 0 & r \in (0,\infty)\,, \\ \tilde f(0) = 0\,, \ \ \tilde f'(\infty) = 0\,, \end{cases}$$

which is nothing but

$$\tilde f(r) = 3\int_0^r \left(e^{\frac{\tilde G(\xi)}{4\lambda}} - 1\right) d\xi\,, \qquad \tilde G(\xi) = \int_\xi^\infty \tilde g(\tau) d\tau$$

Moreover, since $rg(r) = o(1)$ as $r \to 0$, we still still use [3] (3.57) and we deduce, for some positive constant $L_0$,

$$\frac{\varepsilon}{(T'-\hat t)^2} + (a + b\hat r^\alpha + c(\hat r + \tilde f(\hat r))) \le (\beta + b\hat t) L_0\, \hat r^\alpha + (\gamma + c\hat t) L_0 \hat r + (h_0 + h_\alpha + h_1)\hat r\,.$$

As before choosing $T' = \frac{1}{2L_0}$ we get

$$\frac{\varepsilon}{(T'-\hat t)^2} + (a + \frac12 b\hat r^\alpha + \frac12 c\hat r) \le \beta L_0\, \hat r^\alpha + \gamma L_0 \hat r + (h_0 + h_\alpha + h_1)\hat r$$

and we conclude choosing $a = 0$, $b = 2L_0\beta$, $c = 2(L_0\gamma + h_0 + h_\alpha + h_1)$.

[3] Note that if $k_\alpha = 0$ we can take $\beta = b = 0$, in which case we do not need anymore the assumption $rg(r) = o(1)$ as $r \to 0^+$.

Recalling the previous conditions on $\beta$ and $\gamma$ ($\beta = k_\alpha$, $\gamma \ge k_1$), we have then obtained that for every $t \in [0,T']$ and $x, y \in \mathbb{R}^N$

$$|u(t,x) - u(t,y)| \le k_0 + k_\alpha(1 + 2L_0 t)|x-y|^\alpha + c_0(\max\{h_0, h_\alpha, k_1\} + ct)|x-y|),$$

where $c = 2(L_0 \max\{h_0, h_\alpha, k_1\} + h_0 + h_\alpha + h_1)$ and $c_0 = 1 + \tilde{f}'(0)$ only depends on $g$, $\lambda$.

Next we iterate the estimate as in the previous case concluding that

$$|u(t,x) - u(t,y)| \le k_0 + (k_\alpha + Lk_\alpha t)|x-y|^\alpha + c_0(\max\{h_0, h_\alpha, k_1\} + MLt)|x-y|$$

where $L = L(g, \lambda, \alpha, T)$, $M = \max(k_1, h_0, h_\alpha, h_1)$ and $c_0 = c_0(g, \lambda, \alpha)$.

Note that also in this case an alternative estimate can be obtained if we choose $a = h_0$, $b = 2(L_0 k_\alpha + h_\alpha)$ (as before) and $c = 2(L_0 k_1 + h_1)$. In this way we obtain the different estimate

$$|u(t,x) - u(t,y)| \le (k_0 + h_0 t) + (k_\alpha + bt)|x-y|^\alpha + c_0(k_1 + ct)|x-y|, \tag{3.63}$$

for every $t \in [0,T']$ and $x, y \in \mathbb{R}^N$. This estimate may be interesting for small time $t$. ■

**Remark 3.21.** We explicitly point out that we also proved the alternative estimates (3.61) and (3.63), namely that

$$|u(t,x) - u(t,y)| \le (k_0 + h_0 t) + (\beta + bt)|x-y|^\alpha + c_0(k_1 + 2(L_0 k_1 + h_1)t)|x-y| \tag{3.64}$$

for every $t \in [0,T']$ and $x, y \in \mathbb{R}^N$, where $L_0$ only depends on $g$, $\lambda$, $\alpha$, $T$ and the constants $\beta$, $b$ are different according to case (i) or (ii); precisely, we have $\beta = \max\{k_\alpha, k_1\}$ and $b = 2(L_0 \max\{k_\alpha, k_1\} + \max\{h_\alpha, h_1\})$ in case (i), while $\beta = k_\alpha$ and $b = 2(L_0 k_\alpha + h_\alpha)$ in case (ii).

**Remark 3.22.** *Under the stronger condition that there exist $M$, $L \ge 0$ such that $g(r) \le M + Lr$ for every $r > 0$, the above proof works even if $\lambda = 0$, i.e., for a degenerate problem, providing a bound for the oscillation of $u$.*

More precisely, if (3.50) and (3.51) hold with $k_\alpha = 0$ and $h_\alpha = 0$, respectively, then $u$ has bounded oscillation and

$$|u(t,x) - u(t,y)| \le k_0 + C_0 t(h_0 + M(h_1 + k_1)) + (k_1 + C_0 t(h_1 + k_1))|x-y|\big), \tag{3.65}$$

for any $x, y \in \mathbb{R}^N$, $t \in [0,T]$, where $C_0 = C_0(T, L) > 0$.

The proof runs as above; taking $b = \beta = 0$ and $\tilde{f} = 0$, and $\gamma = k_1$, (3.56) reads as

$$\frac{\varepsilon}{(T' - \hat{t})^2} + (a + c\hat{r}) \le (k_1 + c\hat{t})g(\hat{r}) + h_0 + h_1 \hat{r}$$

which implies, since $g(r) \le M + Lr$, $r > 0$,

$$\frac{\varepsilon}{(T' - \hat{t})^2} + (a + c\hat{r}) \le (k_1 + c\hat{t})(M + L\hat{r}) + h_0 + h_1 \hat{r}\,.$$

Choosing $T' = \frac{1}{2L}$ we get

$$\frac{\varepsilon}{(T' - \hat{t})^2} + (a + c\hat{r}) \le M(k_1 + \frac{c}{2L}) + (k_1 L + \frac{1}{2}c)\hat{r} + h_0 + h_1 \hat{r}$$

and we conclude choosing $c = 2(h_1 + k_1 L)$ and $a = M(k_1 + \frac{c}{2L}) + h_0$. We obtain then

$$|u(t,x) - u(t,y)| \le k_0 + at + (k_1 + (2h_1 + 2k_1 L)t)|x-y|\big), \quad x, y \in \mathbb{R}^N, \; t \in [0,T'],$$

where $a = M(2k_1 + \frac{h_1}{L}) + h_0$, and next we extend the estimate in $[0,T]$.

Notice that (3.65) allows one to estimate the oscillation of $u$ in terms of the oscillation of $h$ and $u_0$. If in addition we have $M = 0$ (i.e., $g(r) \le Lr$, $r > 0$), we also deduce that the Lipschitz continuity can be preserved if $h_0 = k_0 = 0$; *in that case, $u_0$ and $h$ globally Lipschitz imply that $u(t)$ is globally Lipschitz.*

Let us observe the following corollary of the above result.

**Corollary 3.23.** *Assume that (3.2) and (3.3) hold true and that (3.7) holds in $Q_T$ with some non-negative $g \in C(0,+\infty)$ such that $g(r)$ is $O(r)$ as $r \to +\infty$ and one of the two conditions (i) or (ii) of Lemma 3.19 holds. Let $u \in C\left(\bar{Q}_T\right)$ be a viscosity solution of (3.1) such that $u$ is $o(\varphi)$ in $\bar{Q}_T$. If data $u_0$, $h$ have bounded oscillation, then $u$ has bounded oscillation and*

$$osc(u(t)) \le K\left\{osc(u_0) + t\left[osc(u_0) + osc_{(0,T)}(h)\right]\right\},$$

*for every $t \in [0,T]$, where $K = K(\lambda, g, \alpha, T)$.*

Combining Theorems 3.3 and Lemma 3.19, assuming the additional control of $g$ at infinity, we obtain a global estimate for the case that the initial datum and the source term are possibly unbounded.

**Theorem 3.24.** *Assume that (3.2), (3.3) hold true, and, in addition, that there exists a non-negative $g \in C(0,+\infty)$ such that $\int_0^1 g(s)ds < \infty$, $g(r)r \to 0$ as $r \to 0^+$, $g$ is $O(r)$ as $r \to +\infty$ and (3.7) holds for every $x, y \in \mathbb{R}^N$.*

*Assume that $u_0$, $h$ satisy (3.50) and (3.51). Let $u \in C(\bar{Q}_T)$ be a viscosity solution of (3.1) which is $o(\varphi)$ in $\bar{Q}_T$, where $\varphi$ satisfies (3.3). Then $u(t)$ is Lipschitz continuous, for every $t \in (0,T)$, and, moreover, there exists $c = c(T, \lambda, g, \alpha) > 0$ such that, for $t \in (0,T)$,*

(3.66)

$$\|Du(t)\|_\infty \le c\left\{\frac{k_0}{\sqrt{t\wedge 1}} + \frac{k_\alpha}{(t\wedge 1)^{1/2-\alpha/2}} + k_1 + (\sqrt{t\wedge 1})\left(h_0 + h_\alpha (t\wedge 1)^{\alpha/2} + h_1\sqrt{t\wedge 1}\right)\right\}.$$

The above estimate shows the sharp dependence in $t$ of the Lipschitz constant of $u(t)$ in terms of Lipschitz and Hölder constants of $u_0$ and $h$. For example, if $h = 0$ and $u_0$ is (possibly unbounded) $\alpha$-Hölder continuous on $\mathbb{R}^N$ we can set $k_\alpha = [u_0]_\alpha$ ($[u_0]_\alpha$ denotes the $\alpha$-Hölder constant or the $C^\alpha$-seminorm of $u_0$) and obtain with $k_0 = k_1 = 0$

$$\|Du(t)\|_\infty \le \frac{c_T}{t^{1/2-\alpha/2}}[u_0]_\alpha, \quad t \in (0,T).$$

A similar estimate under stronger assumptions on the coefficients of the operator $A_t$ (i.e., $h = 0$, $q \ge \lambda I$, $q_{ij}$ bounded, regular and Lipschitz continuous, $b$ possibly unbounded and Hölder continuous) was recently proved in [FGP10, Lemma 4].

*Proof.* Let $t_0 \in (0,T)$. By assumption, we estimate the oscillation of $h$ as

$$\omega_{0,\delta}(h) := osc_{(\frac{t_0}{2}, T\wedge\frac{3}{2}t_0),\delta}(h) \le h_0 + h_\alpha\delta^\alpha + h_1\delta.$$

By Remark 3.21, thanks to the assumptions on $u_0$ and $g$, we estimate the oscillation of $u$ as

$$\omega_{0,\delta}(u) := osc_{(\frac{t_0}{2}, T\wedge\frac{3}{2}t_0),\delta}(u) \le C_T\left\{(k_0 + h_0t_0) + (k_\alpha + h_\alpha t_0)\delta^\alpha + (k_1 + h_1t_0)\delta\right\}$$

where $C_T$ depends on $T, g, \lambda, \alpha$. Using Remark 3.6 with $\delta = \sqrt{t_0\wedge 1}$, we deduce

$$Lip(u(t_0)) \le C_\lambda\left(\frac{C_T\left\{(k_0 + h_0t_0) + (k_\alpha + h_\alpha t_0)\delta^\alpha + (k_1 + h_1t_0)\delta\right\}}{\delta} + \delta(h_0 + h_\alpha\delta^\alpha + h_1\delta)\right)$$

and we conclude (3.66). ∎

3.4. **Possible extensions to coefficients with more general growth.** We give here an extension of Theorem 3.3 to operators with more general growth, like for example

$$\tilde{A}_t u = (1+|x|^2)\,\mathrm{tr}(q(t,x)D^2u) + b(t,x)\cdot Du, \tag{3.67}$$

where $q(t,x) \geq \lambda_0 I$ for some $\lambda_0 > 0$, $q_{i,j}(t,\cdot) \in C_b^{\alpha}(\mathbb{R}^N)$ (uniformly in $t \in [0,T]$) for some $\alpha \in (0,1)$ and, say, $b(t,\cdot)$ is uniformly continuous (uniformly in time).

Generalizing (3.6), we first consider positive numbers $\lambda_{t,x}$ satisfying

$$q(t,x)\xi\cdot\xi \geq \lambda_{t,x}\,|\xi|^2, \quad (t,x)\in Q_T,\ \ \xi\in\mathbb{R}^N, \tag{3.68}$$

and such that, for some $\lambda \in (0,1)$, we have $\lambda_{t,x} \geq \lambda$, $(t,x) \in Q_T$. We write $\sigma(t,x,y) = \sqrt{q(t,x) - (\lambda_{t,x}\wedge\lambda_{t,y})I}$ to denote the symmetric $N\times N$ non-negative matrix such that

$$\sigma^2(t,x,y) = q(t,x) - (\lambda_{t,x}\wedge\lambda_{t,y})\,I, \quad t\in(0,T),\ \ x,y\in\mathbb{R}^N. \tag{3.69}$$

Similarly, $\sigma(t,y,x) = \sqrt{q(t,y) - (\lambda_{t,x}\wedge\lambda_{t,y})I}$.

**Theorem 3.25.** *Let $A_t$ be given in (3.4). Assume the same hypotheses of Theorem 3.3 only replacing (3.7) with the following weaker condition*

$$\begin{aligned} &\frac{1}{|x-y|}\Big(\Big\|\sqrt{q(t,x)-(\lambda_{t,x}\wedge\lambda_{t,y})I} - \sqrt{q(t,y)-(\lambda_{t,x}\wedge\lambda_{t,y})I}\Big\|^2 \\ &\quad + (b(t,x)-b(t,y))\cdot(x-y)\Big) \leq 4(\lambda_{t,x}\wedge\lambda_{t,y})\,g_0(|x-y|), \\ &\qquad x,y\in\mathbb{R}^N,\ 0<|x-y|\leq 1, \quad t\in(0,T), \end{aligned} \tag{3.70}$$

*for some $g_0 \in C(0,1;\mathbb{R}_+)\cap L^1(0,1)$ and some $\lambda_{t,x}$ satisfying (3.68).*

*Let $u \in C(Q_T)$ be as in Theorem 3.3. Then $u(t)$ is Lipschitz continuous and, setting $G = e^{\int_0^1 g_0(s)ds}$,*

$$\|Du(t)\|_\infty \leq \frac{G(1+2\lambda)}{\lambda\sqrt{t\wedge 1}}\ osc_{(\frac{t}{2},T\wedge\frac{3}{2}t)}(u)\ +\ \frac{G}{4\lambda}\sqrt{t\wedge 1}\ osc_{(\frac{t}{2},T\wedge\frac{3}{2}t)}(h)\ ,\quad t\in(0,T). \tag{3.71}$$

Note that, by taking $\lambda_{t,x} = \lambda$ condition (3.70) becomes (3.7) with $g = 4\lambda g_0$.

**Remark 3.26.** A sufficient condition for (3.70) in terms of $q_{ij}$ is

$$\begin{aligned} &\frac{1}{|x-y|}\left(\frac{1}{2(\lambda_{t,x}\wedge\lambda_{t,y})}\,\|q(t,x)-q(t,y)\|^2 + (b(t,x)-b(t,y))\cdot(x-y)\right) \\ &\qquad \leq 2(\lambda_{t,x}\wedge\lambda_{t,y})\,\cdot\,g_0(|x-y|), \end{aligned} \tag{3.72}$$

$x,y\in\mathbb{R}^N$, $0<|x-y|\leq 1$, $t\in(0,T)$. To see that (3.72) implies (3.70), we notice that $\sqrt{q(t,x) - \frac{\lambda_{t,x}\wedge\lambda_{t,y}}{2}I} \geq \sqrt{\frac{\lambda_{t,x}}{2}}I$, and $\sqrt{q(t,y) - \frac{\lambda_{t,x}\wedge\lambda_{t,y}}{2}I} \geq \sqrt{\frac{\lambda_{t,y}}{2}}I$. Then, using (3.12), we have

$$\begin{aligned} &\Big\|\sqrt{q(t,x)-\frac{\lambda_{t,x}\wedge\lambda_{t,y}}{2}I} - \sqrt{q(t,y)-\frac{\lambda_{t,x}\wedge\lambda_{t,y}}{2}I}\Big\| \\ &\qquad \leq \frac{1}{\sqrt{2}\sqrt{\lambda_{t,x}\wedge\lambda_{t,y}}}\,\|q(t,x)-q(t,y)\|, \end{aligned}$$

and we obtain (3.70) with $\frac{\lambda_{t,x}}{2}$ and $\frac{\lambda_{t,y}}{2}$. In conclusion, the statement (3.71) holds if we replace the condition (3.70) with (3.72). Note that the operator in (3.67) satisfies (3.72).

*Proof.* According to Remark 3.2 there is no loss of generality in taking $M = 0$ in (3.3).

As in the proof of Theorem 3.3, we fix $t_0 \in (0,T)$ and consider $\delta \in (0,1]$ and the open set $\Delta = \Delta(t_0,\delta)$ with the same notation. Also the function $\Phi_\epsilon(t,x,y)$ is defined as in the previous proof. Arguing by contradiction, we assume that $\sup_{\Delta} \Phi_\epsilon(t,x,y) > 0$ and find that $u(t,x) - u(t,y) - z(t,x,y)$ has a local maximum at some $(\hat t,\hat x,\hat y) \in \triangle$. Applying Theorem 2.3 we arrive again at formula (3.20), with matrices $X$, $Y \in \mathcal{S}_N$ which satisfy (3.21) for any $N \times N$ matrix $c(t,x,y)$ such that $Q(t,x,y) = \begin{pmatrix} q(t,x) & c(t,x,y) \\ c^*(t,x,y) & q(t,y) \end{pmatrix}$ is symmetric and non-negative.

Generalizing the choice of the coupling matrix in [PW06], we choose (according to (3.69))

$$c(t,x,y) = \left(\sqrt{q(t,x) - (\lambda_{t,x} \wedge \lambda_{t,y})I}\right)\left(\sqrt{q(t,y) - (\lambda_{t,x} \wedge \lambda_{t,y})I}\right) + (\lambda_{t,x} \wedge \lambda_{t,y})\left(I - 2\left(\frac{x-y}{|x-y|} \otimes \frac{x-y}{|x-y|}\right)\right). \tag{3.73}$$

Let us check that $Q(t,x,y)$ is symmetric and non-negative. According to (3.69), we set

$$J(t,x,y) = \sqrt{q(t,x) - (\lambda_{t,x} \wedge \lambda_{t,y})I}\ \sqrt{q(t,y) - (\lambda_{t,x} \wedge \lambda_{t,y})I} = \sigma(t,x,y)\sigma(t,y,x),$$

and we have $Q(t,x,y) = Q_1(t,x,y) + Q_2(t,x,y)$, where

$$Q_1(t,x,y) = \begin{pmatrix} \sigma(t,x,y)^2 & J(t,x,y) \\ J(t,x,y)^* & \sigma(t,y,x)^2 \end{pmatrix},$$

and (we write $p = \frac{x-y}{|x-y|}$)

$$Q_2(t,x,y) = (\lambda_{t,x} \wedge \lambda_{t,y}) \begin{pmatrix} I & (I - 2p \otimes p) \\ (I - 2p \otimes p) & I \end{pmatrix}.$$

It is clear that $Q_1(t,x,y)$ and $Q_2(t,x,y)$ are non-negative and so the same holds for $Q(t,x,y)$.

Recalling the definition of $z$ we arrive at

$$\begin{aligned} 2C_0(\hat t - t_0) \le{}& K\operatorname{tr}\left((q(\hat t,\hat x) + q(\hat t,\hat y) - 2c(\hat t,\hat x,\hat y))D^2\psi(\hat t,\hat x - \hat y)\right) \\ &+ \left(b(\hat t,\hat x) - b(\hat t,\hat y)\right) \cdot D\psi(\hat x - \hat y) + \omega_{0,\delta}(h). \end{aligned} \tag{3.74}$$

We have

$$\begin{aligned} &\operatorname{tr}\left((q(\hat t,\hat x) + q(\hat t,\hat y) - 2c(\hat t,\hat x,\hat y))D^2\psi(\hat x - \hat y)\right) \\ &= \tfrac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}\operatorname{tr}(A(\hat t,\hat x,\hat y)) - \tfrac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}A(\hat t,\hat x,\hat y)\hat p\cdot\hat p + f''(|\hat x-\hat y|)A(\hat t,\hat x,\hat y)\hat p\cdot\hat p, \end{aligned}$$

where

$$\begin{aligned} A(\hat t,\hat x,\hat y) &= q(\hat t,\hat x) + q(\hat t,\hat y) - 2c(\hat t,\hat x,\hat y) \\ &= (\sigma(\hat t,\hat x,\hat y) - \sigma(\hat t,\hat y,\hat x))^2 + 4(\lambda_{\hat t,\hat x} \wedge \lambda_{\hat t,\hat y})\,\hat p \otimes \hat p, \end{aligned}$$

with $\hat p = \frac{\hat x-\hat y}{|\hat x-\hat y|}$. We get

$$\operatorname{tr}(A(\hat t,\hat x,\hat y)) = \|\sigma(\hat t,\hat x,\hat y) - \sigma(\hat t,\hat y,\hat x)\|^2 + 4(\lambda_{\hat t,\hat x} \wedge \lambda_{\hat t,\hat y})$$

and $A(\hat t,\hat x,\hat y)\hat p\cdot\hat p \ge 4(\lambda_{\hat t,\hat x} \wedge \lambda_{\hat t,\hat y})$ so that (recall that $f' > 0$ and $f'' < 0$)

$$\begin{aligned} &\operatorname{tr}\left((q(\hat t,\hat x) + q(\hat t,\hat y) - 2c(\hat t,\hat x,\hat y))D^2\psi(\hat x - \hat y)\right) \\ &\le 4(\lambda_{\hat t,\hat x} \wedge \lambda_{\hat t,\hat y})\,f''(|\hat x-\hat y|) + \tfrac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}\,\|\sigma(\hat t,\hat x,\hat y) - \sigma(\hat t,\hat y,\hat x)\|^2. \end{aligned}$$

From (3.74), using (3.70) we obtain then

$$2C_0(\hat{t}-t_0) \le K4(\lambda_{\hat{t},\hat{x}} \wedge \lambda_{\hat{t},\hat{y}})\,(f''(|\hat{x}-\hat{y}|)+f'(|\hat{x}-\hat{y}|)\,g_0(|\hat{x}-\hat{y}|)) + \omega_{0,\delta}(h).$$

Choose now $f$ as the solution of (3.30) with $\lambda$ replaced by 1/4, namely,

$$f(r) = \int_0^r e^{-G(\xi)} \int_\xi^\delta e^{G(\tau)} d\tau d\xi\,, \qquad \text{where } G(\xi) = \textstyle\int_0^\xi g_0(\tau)d\tau\,.$$

We get

$$4(\lambda_{\hat{t},\hat{x}} \wedge \lambda_{\hat{t},\hat{y}})\,K \le \omega_0(h) + 2C_0(t_0-\hat{t}) < \omega_{0,\delta}(h) + C_0 t_0,$$

and since we have chosen $C_0 = \frac{4\omega_{0,\delta}(u)}{t_0^2}$, we deduce

$$K < \frac{\omega_{0,\delta}(u)}{(\lambda_{\hat{t},\hat{x}} \wedge \lambda_{\hat{t},\hat{y}})\,t_0} + \frac{\omega_{0,\delta}(h)}{4(\lambda_{\hat{t},\hat{x}} \wedge \lambda_{\hat{t},\hat{y}})} \le \frac{\omega_{0,\delta}(u)}{\lambda\,t_0} + \frac{\omega_{0,\delta}(h)}{4\lambda}.$$

Therefore, if $K$ is bigger than this bound we reach the desired contradiction, and we conclude as in the proof of Theorem 3.3 choosing

$$K = \frac{1}{\lambda}\left(\frac{\omega_{0,\delta}(u)}{t_0} + \frac{\omega_{0,\delta}(h)}{4} + \frac{\lambda\omega_{0,\delta}(u)}{f(\delta)}\right)\,. \tag{3.75}$$

■

Proceeding as in the proof of Theorem 3.25, all the results in Sections 3.1 and 3.2 concerning Lipschitz and Hölder regularity can be extended replacing condition (3.7) with the more general hypothesis (3.70) (clearly, to generalize Proposition 3.9 and obtain (3.38) one has to require that $sg_0(s) \to 0$ as $s \to 0^+$). Below we only state a generalization of Theorem 3.12 concerning regularity of bounded solutions to Cauchy problems.

On the other hand, we are not able to generalize the results in Section 3.3 replacing (3.7) with (3.70) (in particular, we can not extend Lemma 3.19).

**Corollary 3.27.** *Assume the same assumptions of Theorem 3.12, only replacing (3.7) with the more general (3.70). Let $u \in C(\bar{Q}_T)$ be a viscosity solution of the Cauchy problem (3.39) such that $u_0 := u(0,\cdot)$ is bounded on $\mathbb{R}^N$ and moreover $u$ is $o(\varphi)$ in $\bar{Q}_T$, where $\varphi$ satisfies (3.3).*

*Then $u(t) \in W^{1,\infty}(\mathbb{R}^N)$, for every $t \in (0,T)$, and there exists $C_0 = C_0(\lambda, g)$, such that*

$$\|Du(t)\|_\infty \le C_0\left\{\left[\frac{1}{\sqrt{t\wedge 1}} + \sqrt{t\wedge 1}\,k_0 + k_1\right]\left(\|u_0\|_\infty + (T \wedge \frac{3}{2}t)\|h\|_{T,\infty}\right) + \sqrt{t\wedge 1}\,\|h\|_{T,\infty}\right\}. \tag{3.76}$$

## 4. Nonlinear equations

Here we consider the fully nonlinear equation

$$\partial_t u + F(t,x,Du,D^2u) = h(t,x) \ \text{ in } Q_T, \tag{4.1}$$

where $F$ *is a real continuous function on* $Q_T \times \mathbb{R}^N \times \mathcal{S}_N$ *and* $h$ *is a continuous function on* $Q_T$.

### 4.1. Regularizing effect in terms of the oscillation of the solution.

We start with the global regularizing effect which is proved in terms of the oscillation of a solution. To this purpose, we introduce a set of assumptions which generalize those of the linear case and cover interesting examples as the sup/inf of linear operators, as well as the case of nonlinear Hamiltonians.

**Hypothesis 4.1.** *There exist $\lambda > 0$, $M \ge 0$, $q > 1$, $c_0$, $c_1 \ge 0$, non-negative functions $\eta(t,x,y)$, $\omega : [0,1] \to \mathbb{R}_+$ such that $\lim_{s\to 0^+}\omega(s) = 0$ and $g \in C((0,1);\mathbb{R}_+) \cap L^1(0,1)$ such that*

$$\begin{aligned} F(t,x,\mu(x-y),X) - F(t,y,\mu(x-y),Y) &\ge -\lambda \mathrm{tr}\,(X-Y) - \mu|x-y|\,g(|x-y|) \\ &\quad -(\mu|x-y|)^2\left(c_0 + c_1(\mu|x-y|^2)^{q-1}\right)\omega(|x-y|) - M - \nu\,\eta(t,x,y)\,, \end{aligned} \tag{4.2}$$

*for any* $\mu > 0$, $\nu \geq 0$, $x, y \in \mathbb{R}^N$, $0 < |x-y| \leq 1$, $t \in (0,T)$, $X, Y \in \mathcal{S}_N$*:*

$$\begin{pmatrix} X & 0 \\ 0 & -Y \end{pmatrix} \leq \mu \begin{pmatrix} I & -I \\ -I & I \end{pmatrix} + \nu \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix}. \tag{4.3}$$

A few comments on Hypothesis 4.1 are in order.

**Remark 4.2.** (i) To check that, in the linear case, (3.7) implies (4.2), we first multiply the matrix inequality (4.3) by

$$\begin{pmatrix} \sigma(t,x)^2 & \sigma(t,x)\sigma(t,y) \\ \sigma(t,y)\sigma(t,x) & \sigma(t,y)^2 \end{pmatrix}.$$

Then, taking traces, we deduce

$$\begin{aligned} -\mathrm{Tr}\,((q(t,x)X - q(t,y)Y) \\ \geq -\lambda \mathrm{Tr}\,(X-Y) - \mu \mathrm{Tr}\left((\sigma(t,x)-\sigma(t,y))^2\right) - \nu \mathrm{Tr}\,(q(t,x)+q(t,y)-2\lambda I) \\ \geq -\lambda \mathrm{Tr}\,(X-Y) - \mu \|(\sigma(t,x)-\sigma(t,y)\|^2 - \nu\eta(t,x,y)\,, \end{aligned}$$

where $\eta(t,x,y) = \mathrm{Tr}(q(t,x)) + \mathrm{Tr}(q(t,y)) - 2\lambda N$. By (3.7) we obtain

$$\begin{aligned} -\mathrm{Tr}\,((q(t,x)X - q(t,y)Y) - \mu[b(t,x)-b(t,y)]\cdot(x-y) \\ \geq -\lambda \mathrm{Tr}\,(X-Y) - \mu|x-y|g(|x-y|) - \nu\eta(t,x,y), \end{aligned}$$

hence (4.2).

(ii) Hypothesis 4.1 implies that $F(t,x,0,X)$ is elliptic, at least assuming that $g(s)s \to 0$ as $s \to 0^+$ (which is consistent with requiring that $g$ is integrable). In that case, taking $x = y$ we get

$$F(t,x,0,X) - F(t,x,0,Y) \geq -\lambda \mathrm{tr}\,(X-Y) - M - \nu\eta(t,x,x), \tag{4.4}$$

for every $X$, $Y \in \mathcal{S}_N$ satisfying the matrix inequality. One can prove, as in [CIL92, Remark 3.4], that if $X \leq Y$, then $X$, $Y + \varepsilon I$ satisfy such matrix inequality with $\mu = \left(1 + \frac{\|Y\|}{\varepsilon}\right)\|Y\|$ and $\nu = 0$. Hence (4.4) is satisfied by $X$, $Y + \varepsilon I$, and letting $\varepsilon \to 0$ shows that (4.4) holds for any $X \leq Y$ with $\nu = 0$.

(iii) The terms with $c_0$, $c_1$ account for possibly nonlinear terms in the function $F$ which may depend superlinearly on the gradient. In particular, the term with $c_0$ includes typical terms with (at most) quadratic growth in the gradient, while the term with $c_1$ includes further terms with possibly larger growth. Similar kind of structure conditions are standard in the fully nonlinear framework (see, for instance, [IL90], [CIL92], [Ba08]). *Let us stress that, if the function* $g(s)$ *satisfies* $sg(s) \to 0$ *as* $s \to 0^+$ *(as it is in most cases when* $g \in L^1(0,1)$*), we can assume that the function* $\omega(\cdot)$ *which appears in those terms is only bounded, and therefore it could be dropped being absorbed by the constants* $c_0, c_1$*. This modification however requires a refinement of the proof below, since one needs first to prove an Hölder estimate (of some order* $\alpha$*, possibly small) and then obtain the Lipschitz bound in a second step.*

(iv) The above Hypothesis 4.1 is not meant to cover general situations of quasilinear operators. It is however possible to extend our approach to such situations (see e.g. [Ba91], [Ch93] for similar general frameworks), by suitably modifying Hypothesis 4.1. In order to avoid too many additional tehcnicalities, we decided to defer the analysis of the quasilinear case to the next future.

**Remark 4.3.** Hypothesis 4.1 should be compared with the conditions under which, *in a bounded domain* $\Omega$, the Lipschitz regularity is proved in [IL90, Theorem VII.1]. Even if we essentially adopt the method introduced there, assumption (4.2) is a slightly more general condition. For simplicity, assume that $F$ only depends on $x$ and $X$. In that case, it is required in [IL90] that

$$|F(x,X) - F(y,X)| \le \mu(|x-y|)\|X\|, \quad x,y \in \bar\Omega, \quad X \in \mathcal{S}_N,$$

for some non-negative function $\mu(s)$ such that $\frac{\mu}{s} \in L^1(0,1)$. When specialized to the linear case, this assumption is satisfied if

$$\|q(x) - q(y)\| \le \mu(|x-y|)\,.$$

On the other hand, since $q(x) \ge \lambda I$, the previous inequality implies

$$\|\sqrt{q(x)-\lambda/2} - \sqrt{q(y)-\lambda/2}\| \le C\mu(|x-y|),$$

which corresponds to our assumption (3.7) with $\sigma(x) = \sqrt{q(x)-\lambda/2}$, $b = 0$ and $g(s) = \frac{C^2\mu^2}{s}$. Since our result only requires that $\frac{\mu^2}{s} \in L^1$ rather than $\frac{\mu}{s} \in L^1$, there is a small improvement (ex. we can afford $\mu(s) = \frac{1}{(\log s)}$). In fact, it is known that assumptions involving the whole matrix inequality satisfied by $X, Y$ (like Hypothesis 4.1) may yield finer results rather than assumptions made at fixed $X$ (see also [Ba91, Section III.1]). This is the typical case in uniqueness results, and so is the spirit of Hypothesis 4.1 for the Lipschitz estimates as well. We refer to Subsection 4.4 for a result concerning the local Lipschitz regularity under similar assumptions. Of course, in bounded domains we would need an extra assumption concerning the boundary behaviour in order to obtain global estimates up to the boundary.

Let us proceed towards the formulation of a general result. In order to take care of the behaviour at infinity, we need to make extra assumptions; as in Section 2, we require the existence of Lyapunov type functions, suitably related to the growth of $F$. More precisely, we assume

**Hypothesis 4.4.** *For any $L > 0$, $\exists\ \varphi = \varphi_L \in C^{1,2}(\bar Q_T)\,,\ \varepsilon_0 = \epsilon_0(L) > 0$:*

$$\begin{cases} \varepsilon\partial_t\varphi + F(t,x,p+\varepsilon D\varphi, X+\varepsilon D^2\varphi) - F(t,x,p,X) \ge 0 \\ \qquad \text{for every } (t,x) \in Q_T,\ p \in \mathbb{R}^N\text{: } |p| \le L + \varepsilon|D\varphi(t,x)|,\ X \in \mathcal{S}_N, \text{ and every } \varepsilon \le \varepsilon_0 \\ \varphi(t,x) \to +\infty \quad \text{as } |x| \to \infty, \text{ uniformly for } t \in [0,T]. \end{cases}$$

We refer to Subsection 4.5 for examples of operators satisfying Hypotheses 4.1 and 4.4, including the case of Bellman-Isaacs operators, and a discussion concerning the case of nonlinear Hamiltonians. In particular, we will see that in the linear case such assumptions reduce to (3.7) and (3.3) which were made in the previous section. We have then the following nonlinear version of Theorem 3.3.

**Theorem 4.5.** *Assume that $F(t,x,p,X)$ satisfies Hypotheses 4.1 and 4.4. Let $u \in C(Q_T)$ be a viscosity solution of (4.1). Moreover, assume that $u$ and $h$ have bounded oscillation (see (2.1)). Then $u(t) \in W^{1,\infty}(\mathbb{R}^N)$, $t \in (0,T)$, and*

$$\|Du(t)\|_\infty \le \frac{C}{\sqrt{t\wedge 1}}, \tag{4.5}$$

*where $C$ depends on $osc_{(\frac{t}{2}, T\wedge\frac{3}{2}t)}(u)$, $osc_{(\frac{t}{2}, T\wedge\frac{3}{2}t)}(h)$, $\lambda, g, M, q$, $c_0$, $c_1$ and $\omega$ (cf. Hypothesis 4.1).*

*Proof.* As in the proof of Theorem 3.3, with the same notation, we define

$$\begin{aligned}\Phi_\epsilon(t,x,y) &= u(t,x)-u(t,y)-Kf(|x-y|)-\varepsilon\left(\varphi(t,x)+\varphi(t,y)\right)-C_0(t-t_0)^2-\frac{\epsilon}{T-t}\\ &= u(t,x)-u(t,y)-z(t,x,y),\end{aligned}$$

where $\varphi(t,x)$ is the Lyapunov function given in Hypothesis 4.4 corrisponding to some constant $L>0$ to be fixed later. Arguing by contradiction, we deduce that $\Phi_\varepsilon$ has a positive local maximum at $(\hat{t},\hat{x},\hat{y})\in\triangle=\triangle(t_0,\delta)$ provided

$$C_0=\frac{4\omega_{0,\delta}(u)}{t_0^2}\,,\qquad K\geq\frac{\omega_{0,\delta}(u)}{f(\delta)}\,, \tag{4.6}$$

where $\omega_{0,\delta}(u)=osc_{(\frac{t_0}{2},T\wedge\frac{3}{2}t_0),\delta}(u)$. Since $u$ is a viscosity solution of (4.1), we end up with

$$\begin{aligned}\tfrac{\epsilon}{(T-\hat{t})^2}+2C_0(\hat{t}-t_0)+\varepsilon\left(\partial_t\varphi(\hat{t},\hat{x})+\partial_t\varphi(\hat{t},\hat{y})\right)+F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y})+\varepsilon\,D\varphi(\hat{t},\hat{x}),X)&\\ -F(\hat{t},\hat{y},KD\psi(\hat{x}-\hat{y})-\varepsilon\,D\varphi(\hat{t},\hat{y}),Y)\leq h(\hat{t},\hat{x})-h(\hat{t},\hat{y}),&\end{aligned}$$

where $\psi(\cdot)=f(|\cdot|)$. Therefore we obtain

$$\begin{aligned}&\tfrac{\epsilon}{(T-\hat{t})^2}+2C_0(\hat{t}-t_0)+F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y}),X-\varepsilon\,D^2\varphi(\hat{t},\hat{x}))\\ &-F(\hat{t},\hat{y},KD\psi(\hat{x}-\hat{y}),Y+\varepsilon\,D\varphi(\hat{t},\hat{y}))\leq\omega_{0,\delta}(h)+\mathcal{I}_\varepsilon(\varphi(\hat{t},\hat{x}))+\mathcal{I}_\varepsilon(\varphi(\hat{t},\hat{y})),\end{aligned} \tag{4.7}$$

where $\omega_{0,\delta}(h)=osc_{(\frac{t_0}{2},T\wedge\frac{3}{2}t_0),\delta}(h)$ and where

$$\begin{aligned}\mathcal{I}_\varepsilon(\varphi(\hat{t},\hat{x}))=-\varepsilon\,\partial_t\varphi(\hat{t},\hat{x})+F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y}),X-\varepsilon\,D^2\varphi(\hat{t},\hat{x}))&\\ -F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y})+\varepsilon\,D\varphi(\hat{t},\hat{x}),X)&\end{aligned}$$

and similarly

$$\begin{aligned}\mathcal{I}_\varepsilon(\varphi(\hat{t},\hat{y}))=-\varepsilon\,\partial_t\varphi(\hat{t},\hat{y})+F(\hat{t},\hat{y},KD\psi(\hat{x}-\hat{y})-\varepsilon\,D\varphi(\hat{t},\hat{y}),Y)&\\ -F(\hat{t},\hat{y},KD\psi(\hat{x}-\hat{y}),Y+\varepsilon\,D^2\varphi(\hat{t},\hat{y})).&\end{aligned}$$

Observe that the matrix inequality (3.18) implies (see also (3.23))

$$\begin{pmatrix}X-\varepsilon\,D^2\varphi(\hat{t},\hat{x}) & 0\\ 0 & -(Y+\varepsilon\,D^2\varphi(\hat{t},\hat{y}))\end{pmatrix}\leq K\begin{pmatrix}D^2\psi(\hat{x}-\hat{y}) & -D^2\psi(\hat{x}-\hat{y})\\ -D^2\psi(\hat{x}-\hat{y}) & D^2\psi(\hat{x}-\hat{y})\end{pmatrix}+\frac{1}{n}(D^2z(\hat{t},\hat{x},\hat{y}))^2. \tag{4.8}$$

Note that

$$\frac{1}{n}(D^2z(\hat{t},\hat{x},\hat{y}))^2\leq\frac{\theta_\epsilon(\hat{t},\hat{x},\hat{y})}{n}\begin{pmatrix}I & 0\\ 0 & I\end{pmatrix}, \tag{4.9}$$

for some positive function $\theta_\epsilon$ (independent on $n$), and, moreover, using (3.25) and the concavity of $f$, we have

$$KD^2\psi(\hat{x}-\hat{y})\leq K\frac{f'(|\hat{x}-\hat{y}|)}{|\hat{x}-\hat{y}|}\,I.$$

Since for matrices $B, C \in \mathcal{S}_N$, the inequality $B \le C$ implies $\begin{pmatrix} B & -B \\ -B & B \end{pmatrix} \le \begin{pmatrix} C & -C \\ -C & C \end{pmatrix}$, we can use Hypothesis 4.1 with $\mu = K\frac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}$ and $\nu = \frac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}$ to estimate

$$\begin{aligned}
&F(\hat t,\hat x, KD\psi(\hat x-\hat y), X-\varepsilon D^2\varphi(\hat t,\hat x)) - F(\hat t,\hat y, KD\psi(\hat x-\hat y), Y+\varepsilon D^2\varphi(\hat t,\hat y))\\
&\quad \ge -\lambda \mathrm{tr}\left(X-\varepsilon D^2\varphi(\hat t,\hat x) - (Y+\varepsilon D^2\varphi(\hat t,\hat y))\right) - Kf'(|\hat x-\hat y|)\, g(|\hat x-\hat y|)\\
&- (Kf'(|\hat x-\hat y|))^2 \left(c_0 + c_1 \left(Kf'(|\hat x-\hat y|)\,|\hat x-\hat y|\right)^{q-1}\right)\omega(|\hat x-\hat y|) - M - \tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\,\eta(\hat t,\hat x,\hat y)
\end{aligned}$$

(recall that $D\psi(\hat x-\hat y) = K\frac{f'(|\hat x-\hat y|)}{|\hat x-\hat y|}(\hat x-\hat y)$). On the other hand, since $\max\Phi_\varepsilon > 0$, we deduce that

$$Kf(|\hat x-\hat y|) \le u(\hat t,\hat x) - u(\hat t,\hat y) \le \omega_{0,\delta}(u),$$

hence we get, since $f$ is concave,

$$Kf'(|\hat x-\hat y|)\,|\hat x-\hat y| \le \omega_{0,\delta}(u)\,.$$

Therefore we obtain

$$\begin{aligned}
&F(\hat t,\hat x, KD\psi(\hat x-\hat y), X-\varepsilon D^2\varphi(\hat t,\hat x)) - F(\hat t,\hat y, KD\psi(\hat x-\hat y), Y+\varepsilon D^2\varphi(\hat t,\hat y))\\
&\quad \ge -\lambda \mathrm{tr}\left(X-\varepsilon D^2\varphi(\hat t,\hat x) - (Y+\varepsilon D^2\varphi(\hat t,\hat y))\right) - Kf'(|\hat x-\hat y|)\, g(|\hat x-\hat y|)\\
&\quad -K^2\,(f'(|\hat x-\hat y|))^2\,(c_0 + c_1\,(\omega_{0,\delta}(u))^{q-1})\omega(|\hat x-\hat y|) - M - \tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\,\eta(\hat t,\hat x,\hat y).
\end{aligned}$$

Now we multiply (4.8) by the coupling matrix

$$\begin{pmatrix} I & I - 2\left(\frac{\hat x-\hat y}{|\hat x-\hat y|}\otimes\frac{\hat x-\hat y}{|\hat x-\hat y|}\right) \\ I - 2\left(\frac{\hat x-\hat y}{|\hat x-\hat y|}\otimes\frac{\hat x-\hat y}{|\hat x-\hat y|}\right) & I \end{pmatrix}$$

which is non-negative and we take traces. We obtain (using also (4.9))

$$\mathrm{tr}\left(X-\varepsilon D^2\varphi(\hat t,\hat x) - (Y+\varepsilon D^2\varphi(\hat t,\hat y))\right) \le 4\,Kf''(|\hat x-\hat y|) + \tfrac{2N\theta_\epsilon(\hat t,\hat x,\hat y)}{n}$$

hence

$$\begin{aligned}
&F(\hat t,\hat x, KD\psi(\hat x-\hat y), X-\varepsilon D^2\varphi(\hat t,\hat x)) - F(\hat t,\hat y, KD\psi(\hat x-\hat y), Y+\varepsilon D^2\varphi(\hat t,\hat y))\\
&\quad \ge -4\lambda\,Kf''(|\hat x-\hat y|) - Kf'(|\hat x-\hat y|)\,g(|\hat x-\hat y|) - \tfrac{2N\lambda\,\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\\
&\quad -K^2\,(f'(|\hat x-\hat y|))^2\,(c_0 + c_1\,(\omega_{0,\delta}(u))^{q-1})\omega(|\hat x-\hat y|) - M - \tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\,\eta(\hat t,\hat x,\hat y).
\end{aligned} \tag{4.10}$$

Thus (4.7) implies

$$\begin{aligned}
&\tfrac{\epsilon}{(T-\hat t)^2} + 2C_0(\hat t - t_0) \le 4\lambda Kf''(|\hat x-\hat y|) + Kf'(|\hat x-\hat y|)\,g(|\hat x-\hat y|)\\
&\quad +K^2\,(f'(|\hat x-\hat y|))^2\,(c_0 + c_1\,(\omega_{0,\delta}(u))^{q-1})\omega(|\hat x-\hat y|) + M + \tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\,\eta(\hat t,\hat x,\hat y)\\
&\quad +\tfrac{2N\lambda\,\theta_\epsilon(\hat t,\hat x,\hat y)}{n} + \omega_{0,\delta}(h) + \mathcal{I}_\varepsilon(\varphi(\hat t,\hat x)) + \mathcal{I}_\varepsilon(\varphi(\hat t,\hat y)).
\end{aligned} \tag{4.11}$$

We choose $f$ as the solution of (3.29) and we obtain

$$\begin{aligned}
&\tfrac{\epsilon}{(T-\hat t)^2} + 2C_0(\hat t - t_0) \le -K + K^2\,(f'(|\hat x-\hat y|))^2\,(c_0 + c_1\,(\omega_{0,\delta}(u))^{q-1})\omega(|\hat x-\hat y|) + M\\
&\quad +\tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\eta(\hat t,\hat x,\hat y) + \tfrac{2N\lambda\,\theta_\epsilon(\hat t,\hat x,\hat y)}{n} + \omega_{0,\delta}(h) + \mathcal{I}_\varepsilon(\varphi(\hat t,\hat x)) + \mathcal{I}_\varepsilon(\varphi(\hat t,\hat y)).
\end{aligned}$$

Recall that $f$ satisfies $f(\delta) \geq \frac{\delta^2}{8\lambda}$ and $f'(0) \leq \hat{c}_2\delta$, where $\hat{c}_2 = \frac{1}{4\lambda} e^{\frac{1}{4\lambda}\int_0^1 g(s)ds}$. Now, due to the presence of $K^2 (f'(|\hat{x}-\hat{y}|))^2 (c_0 + c_1 (\omega_{0,\delta}(u))^{q-1})\omega(|\hat{x}-\hat{y}|)$ we do not continue exactly as in the proof of Theorem 3.3. We first fix

$$K = \frac{8\lambda\, \omega_{0,\delta}(u)}{\delta^2},$$

so that $K$ satisfies (4.6); we will choose $\delta$ later. Then we estimate

$$\begin{aligned} K^2\ (f'(|\hat{x}-\hat{y}|))^2 (c_0 + c_1 (\omega_{0,\delta}(u))^{q-1})\omega(|\hat{x}-\hat{y}|) &\leq K^2\, \hat{c}_2^2\, \delta^2 (c_0 + c_1 (\omega_{0,\delta}(u))^{q-1})\omega(|\hat{x}-\hat{y}|) \\ &\leq K\, C\, \omega_{0,\delta}(u)(1 + (\omega_{0,\delta}(u))^{q-1})\omega(|\hat{x}-\hat{y}|). \end{aligned}$$

Since $\lim_{r\to 0^+}\omega(r) = 0$, there exists $\delta_0$ (depending on $\omega$ and $\omega_{0,1}(u)$) such that $\delta \leq \delta_0$ implies

$$K^2\ \left(f'(|\hat{x}-\hat{y}|)\right)^2 (c_0 + c_1 (\omega_{0,\delta}(u))^{q-1})\omega(|\hat{x}-\hat{y}|) \leq \frac{K}{2}.$$

Therefore we get

$$\frac{K}{2} + \frac{\epsilon}{(T-\hat{t})^2} \leq C_0 t_0 + M + \frac{\theta_\epsilon(\hat{t},\hat{x},\hat{y})}{n}\eta(\hat{t},\hat{x},\hat{y}) + \frac{2N\lambda\,\theta_\epsilon(\hat{t},\hat{x},\hat{y})}{n} + \omega_{0,\delta}(h) + \mathcal{I}_\varepsilon(\varphi(\hat{t},\hat{x})) + \mathcal{I}_\varepsilon(\varphi(\hat{t},\hat{y})).$$

Choosing again $\delta$ small enough we obtain $\frac{K}{2} \geq C_0 t_0 + M + \omega_{0,\delta}(h)$. Indeed, recalling the value of $C_0$ (see (4.6)) and that $K = \frac{8\lambda\omega_{0,\delta}(u)}{\delta^2}$, it is enough to take

$$\delta \leq \Big(\sqrt{\frac{\lambda\, t_0\, \omega_{0,\delta}(u)}{4\omega_{0,\delta}(u) + M t_0 + \omega_{0,\delta}(h) t_0}}\ \Big) \wedge \delta_0.$$

This means that we can choose $\delta = C_3\sqrt{t_0 \wedge 1}$ for some constant $C_3 = C_3(\delta_0, \omega_{0,\delta}(u), \lambda, g, M, q, c_0, c_1)$; recall that $\omega_{0,\delta}(u) \leq osc_{(\frac{t}{2}, T\wedge\frac{3}{2}t)}(u)$. Finally, since $K$ is now fixed, we use Hypothesis 4.4 with $L = Kf'(0)$, which allows us to deduce that $\mathcal{I}_\varepsilon(\varphi(\hat{t},\hat{x}))$, $\mathcal{I}_\varepsilon(\varphi(\hat{t},\hat{y})) \leq 0$. In this way, we get for, $0 < \epsilon \leq \epsilon_0(L)$,

$$\frac{\epsilon}{(T-\hat{t})^2} \leq \frac{\theta_\epsilon(\hat{t},\hat{x},\hat{y})}{n}\eta(\hat{t},\hat{x},\hat{y}) + \frac{2N\lambda\,\theta_\epsilon(\hat{t},\hat{x},\hat{y})}{n}.$$

Letting $n \to \infty$, we get a contradiction. Continuing as in the proof of Theorem 3.3 we obtain

$$|u(t_0,x) - u(t_0,y)| \leq Kf'(0)|x-y| \leq \hat{c}_2 K\delta|x-y|, \qquad |x-y| \leq \delta\,.$$

The proof is complete. ■

Now we consider a similar regularizing effect but concerning the Hölder continuity of solutions (i.e., the extension of Proposition 3.9). In this case we still assume the structure condition of Hypothesis 4.1 but replacing the condition $g \in L^1(0,1)$ with $\lim_{s\to 0^+} sg(s) = 0$; moreover, we do not need any more the modulus of continuity $\omega(|x-y|)$ in the superlinear terms (see also Remark 4.2 (iii)).

**Hypothesis 4.6.** *There exist $\lambda > 0$, $M \geq 0$, $q > 1$, $c_0 \geq 0$, $c_1 \geq 0$, non-negative functions $\eta(t,x,y)$ and $g \in C((0,1);\mathbb{R}_+)$ satisfying $\lim_{s\to 0^+} sg(s) = 0$, such that*

$$\begin{aligned} F(t,x,\mu(x-y),X) - F(t,y,\mu(x-y),Y) &\geq -\lambda \mathrm{tr}\,(X-Y) - \mu|x-y|\, g(|x-y|) \\ &\quad -(\mu|x-y|)^2 \left(c_0 + c_1(\mu|x-y|^2)^{q-1}\right) - M - \nu\,\eta(t,x,y)\,, \end{aligned} \tag{4.12}$$

*for any $\mu > 0$, $\nu \geq 0$, $x, y \in \mathbb{R}^N$, $0 < |x-y| \leq 1$, $t \in (0,T)$, $X, Y \in \mathcal{S}_N$ satisfying (4.3).*

We also need to replace Hypothesis 4.4 with the following stronger one, which accounts for the fact that solutions will possibly have unbounded gradient.

**Hypothesis 4.7.** $\exists\, \varphi \in C^{1,2}(\bar Q_T)\,,\ \epsilon_0>0\,$*:*

$$\begin{cases} \varepsilon\partial_t\varphi + F(t,x,p+\varepsilon D\varphi, X+\varepsilon D^2\varphi) - F(t,x,p,X)\geq 0 \\ \qquad \text{for every } (t,x)\in Q_T,\ p\in\mathbb{R}^N,\ X\in\mathcal{S}_N, \text{ and every } \varepsilon\leq\varepsilon_0 \\ \varphi(t,x)\to+\infty \quad \text{as } |x|\to\infty, \text{ uniformly for } t\in[0,T]. \end{cases}$$

**Proposition 4.8.** *Assume that $F$ satisfies Hypotheses 4.6 and 4.7. Let $u\in C(Q_T)$ be a viscosity solution of (4.1). Moreover, suppose that $u$ and $h$ have bounded oscillation. Then $u(t)$ is $\alpha$-Holder continuous on $\mathbb{R}^N$, for any $\alpha\in(0,1)$, and moreover*

$$|u(t,x)-u(t,y)|\leq \frac{C}{(t\wedge 1)^{\alpha/2}}|x-y|^\alpha, \tag{4.13}$$

*$x,y\in\mathbb{R}^N$, $|x-y|\leq 1$, $t\in(0,T)$, where $C$ depends on $\alpha$, $osc_{(\frac t2,T\wedge\frac32 t)}(u)$, $osc_{(\frac t2,T\wedge\frac32 t)}(h)$, $\lambda, g, M, q$, $c_0$ and $c_1$.*

*Moreover, if we replace $\lim_{s\to0^+} sg(s)=0$ with the condition $\limsup_{s\to0^+} sg(s)<4\lambda$, then there exists some $\alpha=\alpha(\ osc_{(\frac t2,T\wedge\frac32 t)}(u),\ osc_{(\frac t2,T\wedge\frac32 t)}(h),\ \lambda,g,M,q,c_0,c_1)\in(0,1)$ such that (4.13) holds.*

*Proof.* We proceed as in the proof of Theorem 4.5 (see also the proof of Proposition 3.9). The main difference is that we consider $f(s)=s^\alpha$ and take $\delta\leq\delta_1<1$ such that $sg(s)<2\lambda(1-\alpha)$ if $s<\delta_1$. We also fix $K=\frac{\omega_{0,\delta}(u)}{\delta^\alpha}$. Being $f$ increasing and concave, we arrive again at the inequality (4.11), which implies

$$\begin{aligned}\frac{\epsilon}{(T-\hat t)^2}+2C_0(\hat t-t_0)&\leq \alpha\,K\,|\hat x-\hat y|^{\alpha-2}\big(4\lambda\,(\alpha-1)+|\hat x-\hat y|g(|\hat x-\hat y|)\big)+\omega_{0,\delta}(h)\\ &\quad +K^2\left(f'(|\hat x-\hat y|)\right)^2(c_0+c_1\,(\omega_{0,\delta}(u))^{q-1})+M\\ &\quad +\frac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\eta(\hat t,\hat x,\hat y)+\frac{2N\lambda\,\theta_\epsilon(\hat t,\hat x,\hat y)}{n}+\mathcal{I}_\varepsilon(\varphi(\hat t,\hat x))+\mathcal{I}_\varepsilon(\varphi(\hat t,\hat y)).\end{aligned}$$

Hypothesis 4.7 implies that $\mathcal{I}_\varepsilon(\varphi(\hat t,\hat x)),\mathcal{I}_\varepsilon(\varphi(\hat t,\hat y))\leq 0$. Dropping these two terms and letting $n$ go to infinity, we deduce

$$\begin{aligned}\frac{\epsilon}{(T-\hat t)^2}+2C_0(\hat t-t_0)&\leq \alpha\,K\,|\hat x-\hat y|^{\alpha-2}\big(4\lambda\,(\alpha-1)+|\hat x-\hat y|g(|\hat x-\hat y|)\big)+\omega_{0,\delta}(h)\\ &\quad +K^2\left(f'(|\hat x-\hat y|)\right)^2(c_0+c_1\,(\omega_{0,\delta}(u))^{q-1})+M\,.\end{aligned}$$

Since $|\hat x-\hat y|<\delta\leq\delta_1$, we get

$$\alpha\,K\,|\hat x-\hat y|^{\alpha-2}\big(4\lambda\,(\alpha-1)+|\hat x-\hat y|g(|\hat x-\hat y|)\big)\leq 2\lambda\,\alpha(\alpha-1)\,K\,|\hat x-\hat y|^{\alpha-2}.$$

Using the precise choice of $K$ and that $|\hat x-\hat y|<\delta$, we have

$$K^2\left(f'(|\hat x-\hat y|)\right)^2\leq\ K\omega_{0,\delta}(u)\frac{1}{\delta^\alpha}\alpha^2|\hat x-\hat y|^{2\alpha-2}\leq\ K\omega_{0,\delta}(u)\alpha^2|\hat x-\hat y|^{\alpha-2}. \tag{4.14}$$

Therefore, if $\alpha$ is small enough (eventually depending on the oscillation of $u$), we have

$$K^2\left(f'(|\hat x-\hat y|)\right)^2(c_0+c_1\,(\omega_{0,\delta}(u))^{q-1})\leq\lambda\alpha(1-\alpha)K\,|\hat x-\hat y|^{\alpha-2}. \tag{4.15}$$

It follows that

$$\frac{\epsilon}{(T-\hat t)^2}+\lambda\,\alpha(1-\alpha)K\,|\hat x-\hat y|^{\alpha-2}\leq 2C_0(t_0-\hat t)+\omega_{0,\delta}(h)+M\,.$$

Recalling the values of $C_0$ and $K$, and since $|\hat x-\hat y|<\delta$, we obtain then

$$\frac{\epsilon}{(T-\hat t)^2}+\lambda\,\alpha(1-\alpha)\,\delta^{\alpha-2}\frac{\omega_{0,\delta}(u)}{\delta^\alpha}\leq\frac{4\omega_{0,\delta}(u)}{t_0}+\omega_{0,\delta}(h)+M\,.$$

We can choose now $\delta$ small enough so that $\lambda\,\alpha(1-\alpha)\,\delta^{-2}\omega_{0,\delta}(u) \geq \omega_{0,\delta}(h) + \frac{\omega_{0,\delta}(u)}{t_0} + M$ and we get a contradiction. To this purpose, it is enough to consider

$$\delta \leq \Big(\sqrt{\frac{\lambda\alpha(1-\alpha)\,t_0\,\omega_{0,\delta}(u)}{4\omega_{0,\delta}(u) + Mt_0 + \omega_0(h)t_0}}\Big) \wedge \delta_1. \tag{4.16}$$

Continuing as in the proof of Theorem 4.5 we obtain

$$|u(t_0,x) - u(t_0,y)| \leq \frac{\omega_{0,\delta}(u)}{\delta^\alpha}|x-y|^\alpha, \qquad |x-y| \leq \delta\,.$$

Taking into account (4.16) we find the assertion, which is now proved for $\alpha$ sufficiently small. We repeat now the same scheme with any $\alpha < 1$; knowing that (4.13) holds at least for $\alpha$ small, we estimate the oscillation of $u$ as

$$\omega_{0,\delta}(u) \leq \frac{C}{(t_0 \wedge 1)^{\alpha_0/2}}\delta^{\alpha_0}$$

for some (small) $\alpha_0 > 0$, possibly depending on the oscillation $\omega_{0,1}(u)$. Note that $\delta \leq C_3\sqrt{t_0 \wedge 1}$ for some constant $C_3 = C_3(\omega_{0,\delta}(u),\, \lambda, g, M, q,\, c_0,\, c_1,\, \alpha)$.

Now (4.14) implies

$$K^2\,\left(f'(|\hat{x}-\hat{y}|)\right)^2 \leq\, K\omega_{0,\delta}(u)\alpha^2|\hat{x}-\hat{y}|^{\alpha-2} \leq\, K\frac{\delta^{\alpha_0}}{(t_0 \wedge 1)^{\alpha_0/2}}\,\alpha^2|\hat{x}-\hat{y}|^{\alpha-2}\,,$$

and we deduce again that (4.15) holds true provided $\frac{\delta}{\sqrt{t_0\wedge 1}}$ is sufficiently small. We conclude then as before and obtain the estimate (4.13) for every $\alpha < 1$.

To prove the last statement, we set $\gamma = \limsup_{s\to 0^+} sg(s)$ and proceed similarly. First we choose $\alpha \in (0,1)$ such that $\sigma = 4\lambda\,(1-\alpha) - \gamma > 0$. Then we obtain, for $|\hat{x}-\hat{y}| < \delta \leq \delta_1$,

$$\alpha\,K\,|\hat{x}-\hat{y}|^{\alpha-2}\big(4\lambda\,(\alpha-1) + |\hat{x}-\hat{y}|g(|\hat{x}-\hat{y}|)\big) \leq -\sigma\,K\,|\hat{x}-\hat{y}|^{\alpha-2}.$$

Next, we proceed as before and obtain the assertion. ∎

**Remark 4.9.** It follows from the proof of Theorem 4.5 that if $c_0 = c_1 = 0$ in Hypothesis 4.1 then estimate (4.5) becomes:

$$\|Du(t)\|_\infty \leq \frac{\hat{c}_1}{\sqrt{t\wedge 1}}\,\omega(t,u) + \,\hat{c}_2\sqrt{t\wedge 1}\,\left(\omega(t,h) + M\right),\;\; t \in (0,T), \tag{4.17}$$

(as in (3.8)). Note also that if $c_0 = c_1 = 0$ we have (cf. Remark 3.6):

$$Lip(u(t_0)) \leq C_\lambda\,\delta\,\Big(\frac{\omega_{0,\delta}(u)}{t_0} + \omega_{0,\delta}(h) + M + \frac{\omega_{0,\delta}(u)}{\delta^2}\Big), \tag{4.18}$$

where $C_\lambda$ only depends on $\lambda$ and $g$.

Similarly, if $c_0 = c_1 = 0$ in Hypothesis 4.6 then estimate (4.13) becomes as (3.38) (with $\omega(t,h)$ replaced by $\omega(t,h) + M$).

4.2. **The case of bounded data and solutions.** As a first application of Theorem 4.5 and Proposition 4.8 to the Cauchy problem

$$\begin{cases} \partial_t u + F(t,x,Du,D^2u) = h(t,x) \\ u(0,\cdot) = u_0, \end{cases} \tag{4.19}$$

we consider the easier case in which $u_0$ is bounded on $\mathbb{R}^N$, and $h$ is bounded on $Q_T$. Recall that a viscosity solution to (4.19) is a function $u \in C(\bar{Q}_T)$ which is a a viscosity solution to (4.1) and satisfies $u(0,\cdot) = u_0$. We also assume, which is no loss of generality (up to adding a term to the right hand side $h(t,x)$) that

$$F(t,x,0,0) = 0\,, \qquad \text{for any } (t,x) \in Q_T. \tag{4.20}$$

Note that if $F$ satisfies (4.1) and $F(t,x,0,0)$ has bounded oscillation, then $\tilde F = F(t,x,p,X) - F(t,x,0,0)$ satisfies (4.1) and (4.20). The following lemma, based on the Lyapunov function $\varphi$ (cf. Lemma 3.11), ensures the boundedness of solutions.

**Lemma 4.10.** *Assume (4.20) and Hypothesis 4.4. Let $h \in C(Q_T) \cap L^\infty(Q_T)$. Let $u \in C\left(\bar Q_T\right)$ be a viscosity solution of (4.19) such that $u_0 := u(0,\cdot)$ is bounded on $\mathbb{R}^N$ and moreover $u$ is $o(\varphi)$ in $\bar Q_T$, where $\varphi$ satisfies Hypothesis 4.4. Then $u$ is bounded on $[0,T] \times \mathbb{R}^N$ and*

$$\|u(t)\|_\infty \le \|u_0\|_\infty + t\|h\|_{T,\infty}, \quad t \in (0,T), \tag{4.21}$$

*where $\|h\|_{T,\infty} = \sup_{t\in[0,T]} \|h(t)\|_\infty$.*

*Proof.* Let us consider, for $\epsilon > 0$,

$$f_\epsilon(t,x) = u(t,x) - \varepsilon\,\varphi(t,x) - \frac{\varepsilon}{T-t} - \|u_0\|_\infty - t\|h\|_{T,\infty}.$$

If for any $\epsilon > 0$, we have $f_\epsilon(t,x) \le 0$, then, letting $\epsilon \to 0^+$, we deduce $u(t,x) \le \|u_0\|_\infty + t\|h\|_{T,\infty}$. Arguing by contradiction, suppose that, for some $\epsilon > 0$, $\sup\limits_{[0,T)\times\mathbb{R}^N} f_\epsilon(t,x) > 0$. Since $u = o(\varphi)$ this sup is a maximum attained at some point $(t_\epsilon, x_\epsilon)$. Note that $t_\epsilon \in (0,T)$ and so $(t_\epsilon, x_\epsilon)$ is a local maximum. By definition of subsolution we have

$$\|h\|_{T,\infty} + \frac{\varepsilon}{(T-t_\epsilon)^2} + \varepsilon\partial_t\varphi(t_\epsilon,x_\epsilon) + F(t_\epsilon, x_\epsilon, \varepsilon D\varphi(t_\epsilon,x_\epsilon), \varepsilon D^2\varphi(t_\epsilon,x_\epsilon)) \le h(t_\epsilon,x_\epsilon). \tag{4.22}$$

Using Hypothesis 4.4 with $p = 0$ and $X = 0$, and thanks to (4.20), we deduce

$$\|h\|_{T,\infty} + \frac{\varepsilon}{(T-t_\epsilon)^2} \le h(t_\epsilon,x_\epsilon)$$

which yields a contradiction. To obtain the opposite inequality, i.e., $u(t,x) \ge -\|u_0\|_\infty - t\|h\|_{T,\infty}$, we introduce

$$g_\epsilon(t,x) = u(t,x) + \varepsilon\,\varphi(t,x) + \frac{\varepsilon}{T-t} + \|u_0\|_\infty + t\|h\|_{T,\infty},$$

and we suppose that, for some $\epsilon > 0$, $\min\limits_{[0,T)\times\mathbb{R}^N} g_\epsilon(t,x) < 0$. Let $(t_\epsilon, x_\epsilon)$ be the point where this minimum is attained. Considering now $u$ as a supersolution, we get

$$-\|h\|_{T,\infty} - \frac{\varepsilon}{(T-t_\epsilon)^2} - \varepsilon\partial_t\varphi(t_\epsilon,x_\epsilon) + F(t_\epsilon, x_\epsilon, -\varepsilon D\varphi(t_\epsilon,x_\epsilon), -\varepsilon D^2\varphi(t_\epsilon,x_\epsilon))\Big) \ge h(t_\epsilon,x_\epsilon).$$

Using Hypothesis 4.4 with $p = -\varepsilon D\varphi(t_\epsilon,x_\epsilon)$ and $X = -\varepsilon D^2\varphi(t_\epsilon,x_\epsilon)$, and using (4.20), we obtain again a contradiction, and then we conclude. ■

Using Lemma 4.10, we immediately obtain from Theorem 4.5 and Proposition 4.8:

**Corollary 4.11.** *Assume (4.20) and suppose that $h \in C(Q_T) \cap L^\infty(Q_T)$. Assume also that $u_0$ is bounded on $\mathbb{R}^N$. Let $u \in C(\bar Q_T)$ be a viscosity solution of (4.19) which is also $o(\varphi)$ in $\bar Q_T$. We have the following statements.*

*i) If Hypotheses 4.1 and 4.4 hold, then $u(t) \in W^{1,\infty}(\mathbb{R}^N)$, $t \in (0,T)$, and (4.5) holds with a constant $C$ depending on $\|u_0\|_\infty$, $\|h\|_{T,\infty}$, $\lambda, g, M, q$, $c_0$, $c_1$ and $\omega$ (cf. Hypothesis 4.1).*

*ii) If Hypotheses 4.6 and 4.7 hold, then $u(t)$ is $\alpha$-Holder continuous on $\mathbb{R}^N$, for any $\alpha \in (0,1)$, and (4.13) holds with a constant $C$ depending on $\alpha$, $\|u_0\|_\infty$, $\|h\|_{T,\infty}$, $\lambda, g, M, q$, $c_0$ and $c_1$.*

*Moreover, if in Hypothesis 4.6 the function $g$ only satisfies $\limsup_{s\to 0^+} sg(s) < 4\lambda$, then there exists some $\alpha = \alpha(\|u_0\|_\infty, \|h\|_{T,\infty}, \lambda, g, M, q, c_0, c_1) \in (0,1)$ such that (4.13) holds.*

**Remark 4.12.** Similarly to Remark 4.9 if $c_0 = c_1 = 0$ in Hypotheses 4.1 and 4.6 then the previous global estimates of i) and ii) become like (3.42) and (3.48) respectively.

**Remark 4.13.** The previous result can be proved in more generality in order to provide a nonlinear version of Theorem 3.12 and Proposition 3.16. Recall that these results also contain the case of unbounded potential terms $V$ which is not covered in Corollary 4.11.

*Let us only show how to generalize (i) in Corollary 4.11 to the case when $F$ depends also on $u$, namely for the Cauchy problem*

$$\partial_t u + F(t,x,u,Du,D^2u) = 0 \qquad \text{in } Q_T, \quad u(0) = u_0\,. \tag{4.23}$$

*More precisely, we show how to extend Theorem 3.12 to the nonlinear setting.*

As before we may assume that $F(t,x,0,0,0) = 0$. In addition, we assume that

$$F(t,x,u,p,X) \geq F(t,x,v,p,X) \quad \forall u \geq v\,, \quad (t,x) \in Q_T, u,v \in \mathbb{R}, p \in \mathbb{R}^N, X \in \mathcal{S}_N\,. \tag{4.24}$$

Hypotheses 4.1 could be generalized as follows: there exists $\lambda > 0$, $q > 1$, $c_0, c_1, k_0 \geq 0$, $\gamma_R, M_R \geq 0$, non-negative functions $\eta(t,x,y)$, $V(t,x)$, $\omega : [0,1] \to \mathbb{R}_+$ such that $\lim_{s\to 0^+}\omega(s) = 0$ and $g \in C((0,1);\mathbb{R}_+) \cap L^1(0,1)$ such that

$$\begin{aligned} F(t,x,r,\mu(x-y),X) - F(t,y,s,\mu(x-y),Y) &\geq -\lambda \mathrm{tr}\,(X-Y) - \mu|x-y|\,g(|x-y|) \\ &\quad -(\mu|x-y|)^2\left(c_0 + c_1(\mu|x-y|^2)^{q-1}\right)\omega(|x-y|) \\ +\gamma_R(V(t,x)\vee V(t,y))(r-s) - M_R\big(1 + k_0|x-y|(V(t,x)\vee V(t,y))\big) &- \nu\,\eta(t,x,y)\,, \end{aligned} \tag{4.25}$$

for any $\mu > 0$, $\nu \geq 0$, $x,y \in \mathbb{R}^N$: $0 < |x-y| \leq 1$, $t \in (0,T)$, $r,s \in \mathbb{R}$: $-R \leq s \leq r \leq R$, $R > 0$, and $X, Y \in \mathcal{S}_N$ which satisfy (4.3) (we have set $a \vee b = max(a,b)$, $a,b \in \mathbb{R}$).

One can easily check that (4.25) is satisfied in the linear case considered in Theorem 3.12. Finally, Hypothesis 4.4 has to be generalized as follows: for any $L > 0$, $\exists\ \varphi = \varphi_L \in C^2([0,T) \times \mathbb{R}^N)$, $\varepsilon_0 = \epsilon_0(L) > 0$:

$$\varepsilon\partial_t\varphi + F(t,x,r,p+\varepsilon D\varphi, X + \varepsilon D^2\varphi) - F(t,x,r,p,X) \geq 0, \tag{4.26}$$

for every $(t,x) \in Q_T$, $p \in \mathbb{R}^N$: $|p| \leq L + \varepsilon|D\varphi(t,x)|$, $X \in \mathcal{S}_N$, $r \in \mathbb{R}$, and every $\varepsilon \leq \varepsilon_0$, with $\varphi(t,x) \to +\infty$ as $|x| \to \infty$, uniformly for $t \in [0,T]$.

To treat the Cauchy problem (4.19) when $F$ depends also on $u$, we first note that, replacing Hypotheses 4.1 and 4.4 respectively with (4.25) and (4.26), one can still prove Lemma 4.10. To this purpose, the only modification is that one uses (4.26) together with (4.24) and $F(t,x,0,0,0) = 0$ in order to deduce that

$$\varepsilon\partial_t\varphi(t_\epsilon,x_\epsilon) + F(t_\epsilon,x_\epsilon,u(t_\epsilon,x_\epsilon),\varepsilon D\varphi(t_\epsilon,x_\epsilon),\varepsilon D^2\varphi(t_\epsilon,x_\epsilon)) \geq 0,$$

and, of course, a similar argument for the lower bound.

Then, we observe that the proof of Theorem 4.5 still works with small changes under condition (4.25). Recall that here we are assuming that $u_0$ is bounded and so by Lemma 4.10 the solution $u$ is also bounded, which makes possible such a variation. Indeed, we use (4.25) with $R = \|u_0\|_\infty + T\|h\|_{T,\infty}$, $r = u(\hat t,\hat x)$ and $s = u(\hat t,\hat y)$. Then (4.10) now becomes

$$\begin{aligned} &F(\hat t,\hat x,u(\hat t,\hat x),KD\psi(\hat x-\hat y),X - \varepsilon\,D^2\varphi(\hat t,\hat x)) - F(\hat t,\hat y,u(\hat t,\hat y),KD\psi(\hat x-\hat y),Y + \varepsilon\,D^2\varphi(\hat t,\hat y)) \\ &\geq -4\lambda\,Kf''(|\hat x-\hat y|) - Kf'(|\hat x-\hat y|)\,g(|\hat x-\hat y|) - K^2\,(f'(|\hat x-\hat y|))^2\,(c_0 + c_1\,(\omega_{0,\delta}(u))^{q-1})\omega(|\hat x-\hat y|) \\ &\quad +\gamma_R(V(\hat t,\hat x)\vee V(\hat t,\hat y))(u(\hat t,\hat x) - u(\hat t,\hat y)) - M_R(1 + k_0|\hat x-\hat y|(V(\hat t,\hat x)\vee V(\hat t,\hat y))) \\ &\quad -\tfrac{2N\lambda\,\theta_\epsilon(\hat t,\hat x,\hat y)}{n} - \tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\,\eta(\hat t,\hat x,\hat y). \end{aligned}$$

Choosing $f$ as the solution to (3.44), which satisfies $f(r) \geq \frac{\delta}{4\lambda} r$, we use that $(u(\hat t, \hat y) - u(\hat t, \hat x)) < -Kf(|\hat x - \hat y|)$ and it follows

$$\begin{aligned}
&\tfrac{\epsilon}{(T-\hat t)^2} + 2C_0(\hat t - t_0) \leq -K + K^2\left(f'(|\hat x - \hat y|)\right)^2 (c_0 + c_1(2R)^{q-1})\omega(|\hat x - \hat y|)\\
&\quad -\gamma_R K \tfrac{\delta}{4\lambda}(V(\hat t, \hat x) \vee V(\hat t, \hat y)) + M_R(1 + k_0|\hat x - \hat y|(V(\hat t, \hat x) \vee V(\hat t, \hat y)))\\
&\quad + \tfrac{\theta_\epsilon(\hat t, \hat x, \hat y)}{n}\, \eta(\hat t, \hat x, \hat y) + \tfrac{2N\lambda\, \theta_\epsilon(\hat t, \hat x, \hat y)}{n} + \omega_{0,\delta}(h) + \mathcal{I}_\varepsilon(\varphi(\hat t, \hat x)) + \mathcal{I}_\varepsilon(\varphi(\hat t, \hat y)).
\end{aligned}$$

Choosing $K \geq \frac{k_0 M_R}{\gamma_R}\frac{4\lambda}{\delta}$, we drop the terms with the potential $V$ and next we continue as in the proof of Theorem 4.5. Let us just notice that the terms with $\mathcal{I}_\varepsilon(\varphi(\hat t, \hat x))$, $\mathcal{I}_\varepsilon(\varphi(\hat t, \hat y))$ are dealt with now using (4.26) with no additional difficulties.

4.3. **Unbounded data: estimates on the oscillation of the solutions.** We consider now the Cauchy problem (4.19) in the case that the initial datum $u_0$ is not necessarily bounded. The first goal is to obtain estimates on the oscillation of viscosity solutions, so that Theorem 4.5 or Theorem 4.8 can provide a full estimate only depending on the data. To this purpose we will assume that $h$ and $u_0$ have bounded oscillation and we will need to modify Hypotheses 4.1 or 4.6 by requiring some additional condition when $|x-y|$ is large. Hence Hypotheses 4.1 or 4.6 are modified in the following way.

**Hypothesis 4.14.** *There exist $\lambda > 0$, $M_0, M_1 \geq 0$, $q \in (1,2)$, $c_0$, $c_1 \geq 0$, non-negative functions $\eta(t,x,y)$, $\omega \in C(\mathbb{R}_+;\mathbb{R}_+)$ which is bounded and satisfies $\lim_{s\to 0^+}\omega(s) = 0$ and $g \in C((0,+\infty);\mathbb{R}_+) \cap L^1(0,1)$ which is $O(r)$ as $r \to +\infty$, such that*

$$\begin{aligned}
&F(t,x,\mu(x-y),X) - F(t,y,\mu(x-y),Y) \geq -\lambda \mathrm{tr}\,(X-Y) - \mu|x-y|\, g(|x-y|)\\
&\quad -(\mu|x-y|)^2\left(c_0 + c_1(\mu|x-y|^2)^{q-1}\right)\omega(|x-y|) - M_0 - M_1|x-y| - \nu\,\eta(t,x,y)\,,
\end{aligned}$$

*for any $\mu > 0$, $\nu \geq 0$, $x,y \in \mathbb{R}^N$, $t \in (0,T)$, $X, Y \in \mathcal{S}_N$ satisfying (4.3).*

**Hypothesis 4.15.** *There exist $\lambda > 0$, $M_0, M_1 \geq 0$, $q \in (1,2)$, $c_0$, $c_1 \geq 0$, a non-negative functions $\eta(t,x,y)$ and $g \in C((0,+\infty);\mathbb{R}_+)$ which is $O(r)$ as $r \to +\infty$ and satisfies $g(s)s \to 0$ as $s \to 0^+$, such that*

$$\begin{aligned}
&F(t,x,\mu(x-y),X) - F(t,y,\mu(x-y),Y) \geq -\lambda \mathrm{tr}\,(X-Y) - \mu|x-y|\, g(|x-y|)\\
&\quad -(\mu|x-y|)^2\left(c_0 + c_1(\mu|x-y|^2)^{q-1}\right) - M_0 - M_1|x-y| - \nu\,\eta(t,x,y)\,,
\end{aligned}$$

*for any $\mu > 0$, $\nu \geq 0$, $x,y \in \mathbb{R}^N$, $t \in (0,T)$, $X, Y \in \mathcal{S}_N$ satisfying (4.3).*

As a first step, we need the following growth estimates which generalize Lemma 3.19. This result only requires Hypothesis 4.15. On the other hand, Theorem 4.18 will also use Hypothesis 4.14.

**Lemma 4.16.** *Assume that $u_0$, $h$ satisfy (3.50) and (3.51), respectively. Let $u \in C\left(\bar Q_T\right)$ be a viscosity solution of (4.19) such that $u$ is $o(\varphi)$ in $\bar Q_T$. Then we have:*

*(i) If Hypotheses 4.15 and 4.7 hold true, then $u$ has bounded oscillation and there exists $C_T$ such that*

$$|u(t,x) - u(t,y)| \leq C_T\big(1 + |x-y|\big), \quad x,y \in \mathbb{R}^N,\ t \in [0,T], \tag{4.27}$$

*where $C_T$ depends on $T, g, \lambda, \alpha, q$, $c_0$, $c_1$, $M_0$, $M_1$ and $k_i, h_i$, for $i = 0, \alpha, 1$.*

*(ii) If Hypothesis 4.15 holds replacing the condition $\lim_{s\to 0^+} g(s)s = 0$ with $g \in L^1(0,1)$, and if Hypothesis 4.4 holds true, then $u$ has bounded oscillation and satisfies (4.27).*

*(iii) Assume Hypothesis 4.15 with $c_0 = c_1 = 0$ and Hypothesis 4.4. Then $u$ satisfies the same estimate in (i) of Lemma 3.19. If Hypothesis 4.4 holds and we have, in addition, that*

*$g \in L^1(0,1)$, then $u$ satisfies also the estimate (ii) in Lemma 3.19. Moreover, estimate (3.64) holds for $u$.*

**Remark 4.17.** If Hypothesis 4.15 holds with $c_0 = c_1 = 0$, this means, roughly speaking, that the nonlinear terms in the function $F$ have at most linear growth with respect $Du$. As mentioned in (iii), in this case the conclusion of the above Lemma is much stronger since it coincides with the conclusion of Lemma 3.19. In particular, we have conservation of the Lipschitz, or Hölder, continuity from $u_0$ to $u(t)$ if $c_0 = c_1 = 0$.

*Proof.* We essentially follow the proof of Lemma 3.19. For $T' \le T$, let us set

$$\Delta = \{(t, x, y) \in (0, T') \times \mathbb{R}^N \times \mathbb{R}^N\} \tag{4.28}$$

and define the function

$$\Phi_\epsilon(t, x, y) = u(t, x) - u(t, y) - f(t, |x - y|) - \varepsilon\,(\varphi(t, x) + \varphi(t, y)) - \frac{\varepsilon}{T' - t}\,, \quad \epsilon > 0, \tag{4.29}$$

where $C_0$, $T' > 0$ will be chosen later and where

$$f(t, r) = (k_0 + at) + (\beta + bt)r^\alpha + (\gamma + ct)(r + \tilde{f}(r)),$$

where $a, b, c, \beta, \gamma \ge 0$ and $\tilde{f}(r) \in C^2(0, \infty)$ is a nondecreasing concave function with $\tilde{f}(0) = 0$, to be fixed later. We prove that, independently on $\varepsilon$, it holds

$$\Phi_\epsilon(t, x, y) \le 0, \quad (t, x, y) \in \Delta\,.$$

Arguing by contradiction, as in Lemma 3.19 we deduce that $\Phi_\epsilon$ has a local (and global) positive maximum at some point $(\hat{t}, \hat{x}, \hat{y}) \in \overline{\Delta}$ and clearly $\hat{x} \ne \hat{y}$ since the maximum is positive, and $\hat{t} < T'$. Assuming that

$$\beta \ge k_\alpha\,, \quad \gamma \ge k_1, \tag{4.30}$$

we deduce that the maximum cannot be reached at $t = 0$ thanks to (3.50). Since $u$ is a solution, proceeding as usual and setting $\psi(t, x - y) = f(t, |x - y|)$, we end up with

$$\begin{gathered}\tfrac{\varepsilon}{(T'-\hat{t})^2} + (a + b|\hat{x} - \hat{y}|^\alpha + c(|\hat{x} - \hat{y}| + \tilde{f}(|\hat{x} - \hat{y}|)) \\ +\varepsilon\,(\partial_t\varphi(\hat{t}, \hat{x}) + \partial_t\varphi(\hat{t}, \hat{y})) + F(\hat{t}, \hat{x}, D\psi(\hat{t}, \hat{x} - \hat{y}) + \varepsilon\, D\varphi(\hat{t}, \hat{x}), X) \\ -F(\hat{t}, \hat{y}, D\psi(\hat{t}, \hat{x} - \hat{y}) - \varepsilon\, D\varphi(\hat{t}, \hat{y}), Y) \le h(\hat{t}, \hat{x}) - h(\hat{t}, \hat{y}),\end{gathered}$$

where $X$, $Y$ satisfy (3.18) with $z = f(t, |x - y|) + \varepsilon\,(\varphi(t, x) + \varphi(t, y)) + \frac{\varepsilon}{T'-t}$. Therefore we obtain

$$\begin{gathered}\tfrac{\epsilon}{(T'-\hat{t})^2} + (a + b|\hat{x} - \hat{y}|^\alpha + c(|\hat{x} - \hat{y}| + \tilde{f}(|\hat{x} - \hat{y}|)) \\ +F(\hat{t}, \hat{x}, D\psi(t, \hat{x} - \hat{y}), X - \varepsilon\, D^2\varphi(\hat{t}, \hat{x})) \\ -F(\hat{t}, \hat{y}, D\psi(t, \hat{x} - \hat{y}), Y + \varepsilon\, D^2\varphi(\hat{t}, \hat{y})) \le h(\hat{t}, \hat{x}) - h(\hat{t}, \hat{y}) + \mathcal{I}_\varepsilon(\varphi(\hat{t}, \hat{x})) + \mathcal{I}_\varepsilon(\varphi(\hat{t}, \hat{y})),\end{gathered} \tag{4.31}$$

where

$$\begin{aligned}\mathcal{I}_\varepsilon(\varphi(\hat{t}, \hat{x})) = -\varepsilon\, \partial_t\varphi(\hat{t}, \hat{x}) &+ F(\hat{t}, \hat{x}, D\psi(t, \hat{x} - \hat{y}), X - \varepsilon\, D^2\varphi(\hat{t}, \hat{x})) \\ &-F(\hat{t}, \hat{x}, D\psi(t, \hat{x} - \hat{y}) + \varepsilon\, D\varphi(\hat{t}, \hat{x}), X)\end{aligned}$$

and

$$\begin{aligned}\mathcal{I}_\varepsilon(\varphi(\hat{t}, \hat{y})) = -\varepsilon\, \partial_t\varphi(\hat{t}, \hat{y}) &+ F(\hat{t}, \hat{y}, D\psi(t, \hat{x} - \hat{y}) - \varepsilon\, D\varphi(\hat{t}, \hat{y}), Y) \\ &-F(\hat{t}, \hat{y}, D\psi(t, \hat{x} - \hat{y}), Y + \varepsilon\, D^2\varphi(\hat{t}, \hat{y})).\end{aligned}$$

Since $f(t,\cdot)$ is increasing and concave, as in the proof of Theorem 4.5 we can use (3.18) and Hypothesis 4.15 with $\mu = \frac{f'(\hat t,|\hat x-\hat y|)}{|\hat x-\hat y|}$ (we have set $\partial_x f = f'$ and $\partial^2_{xx} f = f''$) and we get

$$\begin{aligned} &F(\hat t,\hat x, D\psi(t,\hat x-\hat y), X-\varepsilon\, D^2\varphi(\hat t,\hat x)) - F(\hat t,\hat y, D\psi(t,\hat x-\hat y), Y+\varepsilon\, D^2\varphi(\hat t,\hat y)) \\ &\geq -\lambda \mathrm{tr}\left(X-\varepsilon\, D^2\varphi(\hat t,\hat x) - (Y+\varepsilon\, D^2\varphi(\hat t,\hat y))\right) - f'(\hat t,|\hat x-\hat y|)\, g(|\hat x-\hat y|) \\ &- \left(f'(\hat t,|\hat x-\hat y|)\right)^2 \left(c_0 + c_1\left(f'(\hat t,|\hat x-\hat y|)|\hat x-\hat y|\right)^{q-1}\right) - M_0 - M_1|x-y| - \tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\,\eta(\hat t,\hat x,\hat y), \end{aligned}$$

for some positive quantity $\theta_\epsilon(\hat t,\hat x,\hat y)$. As in the proof of Theorem 4.5 we estimate

$$\mathrm{tr}\left(X-\varepsilon\, D^2\varphi(\hat t,\hat x) - (Y+\varepsilon\, D^2\varphi(\hat t,\hat y))\right) \leq 4\, f''(\hat t,|\hat x-\hat y|) + \tfrac{2N\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\,.$$

Therefore, setting $\hat r = |\hat x-\hat y|$ and using the growth of $h$, we deduce

$$\begin{aligned} &\tfrac{\epsilon}{(T'-\hat t)^2} + (a + b\hat r^\alpha + c(\hat r + \tilde f(\hat r)) \leq 4\,\lambda f''(\hat t,\hat r) + f'(\hat t,\hat r)\, g(\hat r) \\ &\left(f'(\hat t,\hat r)\right)^2\left(c_0 + c_1\left(f'(\hat t,\hat r)\hat r\right)^{q-1}\right) + M_0 + M_1\hat r + h_0 + h_\alpha \hat r^\alpha + h_1\hat r \\ &+\tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\left(2N\lambda + \eta(\hat t,\hat x,\hat y)\right) + \mathcal{I}_\varepsilon(\varphi(\hat t,\hat x)) + \mathcal{I}_\varepsilon(\varphi(\hat t,\hat y)). \end{aligned} \tag{4.32}$$

We split henceforth the proof in the two cases (i) and (ii).

**(i)** Assume that $g(r)r \to 0$ as $r\to 0^+$ and that Hypothesis 4.7 holds. Thanks to this latter assumption we have $\mathcal{I}_\varepsilon(\varphi(\hat t,\hat x)), \mathcal{I}_\varepsilon(\varphi(\hat t,\hat y)) \leq 0$. We take $\tilde f = 0$; dropping the latter non-positive terms and letting $n\to\infty$, and using the precise form of $f(t,r)$, we find from (4.32)

$$\begin{aligned} &\tfrac{\varepsilon}{(T'-\hat t)^2} + (a + b\hat r^\alpha + c\hat r) \\ &\leq (\beta + b\hat t)\left[4\lambda\alpha(\alpha-1)\hat r^{\alpha-2} + \alpha\hat r^{\alpha-1} g(\hat r)\right] + (\gamma + c\hat t) g(\hat r) + \\ &\left(f'(\hat t,\hat r)\right)^2\left(c_0 + c_1\left(f'(\hat t,\hat r)\hat r\right)^{q-1}\right) + M_0 + h_0 + h_\alpha\hat r^\alpha + (h_1 + M_1)\hat r. \end{aligned}$$

Henceforth, if $c_0 = c_1 = 0$, one can follow the proof of Lemma 3.19 only replacing $h_0$ with $h_0 + M_0$ and $h_1$ with $h_1 + M_1$, and obtaining the same kind of estimate. Assume instead that $c_0$ or $c_1$ are positive. We take $b \geq c$ and $\beta\geq\gamma$; as in the proof of Lemma 3.19, we have that there exist $r_0 < 1$ and $L_0 > 0$ such that

$$\begin{aligned} &(\beta + b\hat t)\left[4\lambda\alpha(\alpha-1)\hat r^{\alpha-2} + \alpha\hat r^{\alpha-1} g(\hat r)\right] + (\gamma + c\hat t) g(\hat r) \\ &\leq -2\alpha(1-\alpha)\lambda\beta\hat r^{\alpha-2}\chi_{\{r<r_0\}} + (\beta + b\hat t)\alpha L_0\hat r^\alpha + (\gamma + c\hat t) L_0\hat r \end{aligned}$$

so that, if we fix $T' = \frac{1}{2L_0}$, we have $L_0\hat t \leq 1/2$ and we deduce

$$\begin{aligned} &\tfrac{\varepsilon}{(T'-\hat t)^2} + a + (\tfrac12 b - L_0\beta)\hat r^\alpha + (\tfrac12 c - L_0\gamma)\hat r \leq -2\beta\alpha(1-\alpha)\lambda\hat r^{\alpha-2}\chi_{\{r<r_0\}} \\ &+ \left(f'(\hat t,\hat r)\right)^2\left(c_0 + c_1\left(f'(\hat t,\hat r)\hat r\right)^{q-1}\right) + M_0 + h_0 + h_\alpha\hat r^\alpha + (h_1 + M_1)\hat r. \end{aligned}$$

We estimate now the superlinear term. If $\hat r \leq r_0$, we have $f'(\hat t,\hat r)\hat r \leq C$, where $C$ depends on $r_0$, $\beta$, $\gamma$, $b, c$, $L_0$ (since $T'$ depends on $L_0$). Then if $\hat r\leq r_0$

$$\left(f'(\hat t,\hat r)\right)^2\left(c_0 + c_1\left(f'(\hat t,\hat r)\hat r\right)^{q-1}\right) \leq C_2\hat r^{2\alpha-2},$$

for some $C_2 = C_2(r_0,\beta,\gamma,b,c,L_0,c_0,c_1)$; so, for some $r_1 < r_0$, we have

$$\left(f'(\hat t,\hat r)\right)^2\left(c_0 + c_1\left(f'(\hat t,\hat r)\hat r\right)^{q-1}\right) \leq \beta\alpha(1-\alpha)\lambda\hat r^{\alpha-2} \quad \text{if } \hat r\leq r_1\,.$$

If instead $\hat r > r_1$, we have

$$\left(f'(\hat t,\hat r)\right)^2\left(c_0+c_1\left(f'(\hat t,\hat r)\hat r\right)^{q-1}\right)\le C\hat r^{q-1}$$

for a possibly different constant $C$, hence, since $q<2$, Young's inequality implies

$$\left(f'(\hat t,\hat r)\right)^2\left(c_0+c_1\left(f'(\hat t,\hat r)\hat r\right)^{q-1}\right)\le \frac{1}{4}c\hat r+K$$

where $K=K(\alpha,r_0,\beta,\gamma,L_0,b,c,c_0,c_1)$. We conclude that

$$\tfrac{\varepsilon}{(T'-\hat t)^2}+a+(\tfrac12 b-L_0\beta)\hat r^\alpha+(\tfrac14 c-L_0\gamma)\hat r\le K+M_0+h_0+h_\alpha\hat r^\alpha+(h_1+M_1)\hat r\,.$$

Here one chooses $c\ge 4(L_0\gamma+h_1+M_1)$, $b\ge 2(L_0\beta+h_\alpha)$ and, lastly, $a\ge K+M_0+h_0$ and the conclusion follows. The estimate is therefore proved in $[0,T']$ for $T'=\frac{1}{2L_0}$. Since $L_0$ only depends on $\lambda,g,\alpha$, the argument can be iterated and yields the global estimate on $[0,T]$.

**(ii)** Assume that $g(r)\in L^1(0,1)$ and that Hypothesis 4.4 holds. Up to using Young's inequality, we can suppose that $k_\alpha=0$; in particular, this allows us to take $b=\beta=0$, so that (4.32) implies

$$\begin{aligned}\tfrac{\epsilon}{(T'-\hat t)^2}+(a+c(\hat r+\tilde f(\hat r))&\le(\gamma+c\hat t)[4\lambda\tilde f''(\hat r)+g(\hat r)(1+\tilde f'(\hat r))]\\ \left((\gamma+c\hat t)(1+\tilde f'(\hat r))\right)^2&\left(c_0+c_1\left((\gamma+c\hat t)(\hat r+\tilde f'(\hat r)\hat r)\right)^{q-1}\right)+M_0+h_0+h_\alpha\hat r^\alpha+(h_1+M_1)\hat r\\ &+\tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\left(2N\lambda+\eta(\hat t,\hat x,\hat y)\right)+\mathcal I_\varepsilon(\varphi(\hat t,\hat x))+\mathcal I_\varepsilon(\varphi(\hat t,\hat y)).\end{aligned}$$

We proceed then as in Lemma 3.19, choosing, with the same notations, $\tilde f(r)=3\int_0^r\left(e^{\frac{\tilde G(\xi)}{4\lambda}}-1\right)d\xi$, where $\tilde G(\xi)=\int_\xi^\infty\tilde g(\tau)d\tau$. In the same way we obtain, for some constant $L_0=L_0(g,\lambda)>0$ (cf. (3.56) and (3.62)):

$$\begin{aligned}\tfrac{\epsilon}{(T'-\hat t)^2}+(a+c(\hat r+\tilde f(\hat r))&\le(\gamma+c\hat t)L_0\hat r+(M_0+h_0+h_\alpha+h_1+M_1)\hat r\\ &\left((\gamma+c\hat t)(1+\tilde f'(\hat r))\right)^2\left(c_0+c_1\left((\gamma+c\hat t)(r+\tilde f'(\hat r)\hat r)\right)^{q-1}\right)\\ &+\tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\left(2N\lambda+\eta(\hat t,\hat x,\hat y)\right)+\mathcal I_\varepsilon(\varphi(\hat t,\hat x))+\mathcal I_\varepsilon(\varphi(\hat t,\hat y)),\end{aligned}$$

and then, choosing $T'=\frac{1}{2L_0}$ and since $t\le T'$ and $\tilde f\ge 0$,

$$\begin{aligned}\tfrac{\epsilon}{(T'-\hat t)^2}+(a+(\tfrac12 c-\gamma L_0)\hat r)&\le(M_0+h_0+h_\alpha+h_1+M_1)\hat r\\ &\left((\gamma+ct)(1+\tilde f'(\hat r))\right)^2\left(c_0+c_1\left((\gamma+ct)(r+\tilde f'(\hat r)\hat r)\right)^{q-1}\right)\\ &+\tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\left(2N\lambda+\eta(\hat t,\hat x,\hat y)\right)+\mathcal I_\varepsilon(\varphi(\hat t,\hat x))+\mathcal I_\varepsilon(\varphi(\hat t,\hat y)).\end{aligned}$$

Since there exists $\tilde l>0$ such that $\tilde f'(r)\le\tilde l$ for every $r>0$, we have $|D\psi(\hat t,\hat x-\hat y)|=(\gamma+c\hat t)(1+\tilde f'(\hat r))\le L$, where $L=(\gamma+\frac{c}{2L_0})(1+\tilde l)$. Therefore, we use Hypothesis 4.4 with such $L$ and we deduce that last two terms are non-positive and can be neglected; letting also $n\to\infty$ we get

$$\begin{aligned}\tfrac{\epsilon}{(T'-\hat t)^2}+(a+(\tfrac12 c-\gamma L_0)\hat r)&\le(M_0+h_0+h_\alpha+h_1+M_1)\hat r\\ &\left((\gamma+c\hat t)(1+\tilde f'(\hat r))\right)^2\left(c_0+c_1\left((\gamma+c\hat t)(r+\tilde f'(\hat r)\hat r)\right)^{q-1}\right).\end{aligned}$$

If $c_0 = c_1 = 0$, we take $a = 0$ and we obtain the same conclusion as in Lemma 3.19, with $k_\alpha = 0$. If $c_0$ or $c_1$ are positive, we observe that, being $q < 2$, we have

$$\left((\gamma + c\hat{t})(1 + \tilde{f}'(\hat{r}))\right)^2 \left(c_0 + c_1 \left((\gamma + c\hat{t})(\hat{r} + \tilde{f}'(\hat{r})\hat{r})\right)^{q-1}\right) \leq K + \frac{1}{4} c\hat{r}$$

for some constant $K = K(\gamma, c, L_0, \tilde{f}, c_0, c_1)$. Then we obtain

$$\tfrac{\epsilon}{(T'-\hat{t})^2} + (a + (\tfrac{1}{4}c - \gamma L_0)\hat{r}) \leq (M_0 + h_0 + h_\alpha + h_1 + M_1)\hat{r} + K,$$

and we conclude as before choosing $a$ and $c$ sufficiently large.

**(iii)** Since $c_0 = c_1 = 0$, it follows exactly as in the proof of Lemma 3.19. ∎

As a consequence of Lemma 4.16, we have now conditions under which solutions have bounded oscillation and Theorem 4.5 can be applied. We deduce a complete regularity result.

**Theorem 4.18.** *Assume that $h$ and $u_0$ have bounded oscillation in $Q_T$ and in $\mathbb{R}^N$, respectively. Let $u \in C(\bar{Q}_T)$ be a viscosity solution of (4.19) which is also $o(\varphi)$ in $\bar{Q}_T$. We have the following statements.*

*i) If Hypotheses 4.14 and 4.4 hold, then $u(t) \in W^{1,\infty}(\mathbb{R}^N)$, $t \in (0,T)$, and (4.5) holds with a constant $C$ depending on $osc(u_0)$, $osc_{(0,T)}(h)$, $\lambda, g, M, q$, $c_0$, $c_1$, $M_0$, $M_1$, $T$ and $\omega$.*

*ii) If Hypotheses 4.15 and 4.7 hold, then $u(t)$ is $\alpha$-Holder continuous on $\mathbb{R}^N$, for any $\alpha \in (0,1)$, and (4.13) holds with a constant $C$ depending on $\alpha$, $osc(u_0)$, $osc_{(0,T)}(h)$, $\lambda, g, M, q$, $c_0$, $c_1$, $M_0$, $M_1$ and $T$.*

*Proof.* Applying Lemma 4.16 we deduce that $u$ has bounded oscillation and its oscillation is estimated in terms of $osc(u_0)$, $osc_{(0,T)}(h)$, $T$ besides the usual constants. Then, we apply either Theorem 4.5 or Proposition 4.8 to conclude the Lipschitz, respectively Hölder, estimate. ∎

The following corollary provides a nonlinear version of Theorem 3.24. It can be applied, for instance, to Bellman-Isaacs type equations.

**Corollary 4.19.** *Assume that $u_0$, $h$ satisfy (3.50) and (3.51), respectively. Assume Hypothesis 4.14 with $c_0 = c_1 = 0$ and requiring that $sg(s) \to 0$ as $s \to 0^+$. Assume also Hypothesis 4.4.*

*Let $u \in C\left(\bar{Q}_T\right)$ be a viscosity solution of (4.19) such that $u$ is $o(\varphi)$ in $\bar{Q}_T$. Then $u(t) \in W^{1,\infty}(\mathbb{R}^N)$, $t \in (0,T)$, and there exist $c = c_T(T, \lambda, g, \alpha) > 0$ such that, for $t \in (0,T)$,*

$$\begin{aligned} \|Du(t)\|_\infty \leq c_T \Big\{ & \tfrac{k_0}{\sqrt{t \wedge 1}} + \tfrac{k_\alpha}{(t\wedge 1)^{1/2-\alpha/2}} + k_1 \\ & + (\sqrt{t \wedge 1})\left(h_0 + M_0 + h_\alpha (t \wedge 1)^{\alpha/2} + (h_1 + M_1)\sqrt{t \wedge 1}\right)\Big\}. \end{aligned} \tag{4.33}$$

*Proof.* We already know by Lemma 4.16 that $u$ has bounded oscillation. We follow the proof of Theorem 4.5 and we arrive at (4.18). Then using (iii) in Lemma 4.16, we can conclude similarly to the proof of Theorem 3.24. ∎

4.4. **Local Lipschitz continuity.** The approach developed so far can also provide estimates and regularity results concerning the local Lipschitz continuity. We give an example in the following theorem.

**Theorem 4.20.** *Assume that Hypothesis 4.1 holds true and in addition that, for every compact set $S \subset \mathbb{R}^N$ we have*

$$\begin{aligned} |F(t,x,p,X) - F(t,x,q,Y)| \leq C_S(1 + (|p|^m + |q|^m))\{|p-q| + \|X - Y\|\}, \\ x \in S\,,\ \ p, q \in \mathbb{R}^N\,,\ t \in (0,T)\,. \end{aligned} \tag{4.34}$$

*for some constants $C_S > 0$ and $m < 2$.*

*Let $u \in C(Q_T)$ be a viscosity solution of equation (4.1) with $h \in C(Q_T)$. Then $u(t) \in W^{1,\infty}_{\rm loc}(\mathbb{R}^N)$ for every $t \in (0,T)$ and, for any ball $B_R(x_0)$, there exists $M_0$ such that*

$$\|Du(t)\|_{L^\infty(B_R(x_0))} \leq \frac{M_0}{\sqrt{t\wedge 1}}\,,$$

*where* $M_0 = (g,\lambda,q,\omega,c_0,c_1,M,x_0,N,R,\|u\|_{L^\infty((\frac{t}{2},T\wedge\frac{3}{2}t)\times B_{2R}(x_0))},\|h\|_{L^\infty((\frac{t}{2},T\wedge\frac{3}{2}t)\times B_{2R}(x_0))})$.

*Proof.* Let us fix $x_0 \in \mathbb{R}^N$, $t_0 > 0$, $\delta > 0$, $r > 0$ and consider the set

$$\Delta = \Big\{(t,x,y) \in (0,T)\times B_r(x_0)\times B_r(x_0)\,:\,|x-y|<\delta\,,\ \frac{t_0}{2}<t<\frac{3}{2}t_0\wedge T\Big\}$$

and the function

$$\Phi_\epsilon(t,x,y) = u(t,x)-u(t,y)-K\psi(x-y)-L|x-x_0|^2-C_0(t-t_0)^2-\frac{\epsilon}{T-t},$$

where $K$, $C_0$, $L$ will be chosen later and $\psi(x-y)=f(|x-y|)$ with $f$ as in (3.30). As usual, we claim that

$$\Phi_\epsilon(t,x,y)\leq 0, \qquad (t,x,y)\in\Delta\,, \tag{4.35}$$

and, arguing by contradiction, we suppose that

$$\sup_\Delta \Phi_\epsilon(t,x,y) > 0\,. \tag{4.36}$$

In the sequel, in order to simplify the notation, we write $\|u\|_\infty := \|u\|_{L^\infty((\frac{t_0}{2},\frac{3}{2}t_0\wedge T)\times B_r(x_0))}$ and $\|h\|_\infty := \|h\|_{L^\infty((\frac{t_0}{2},\frac{3}{2}t_0\wedge T)\times B_r(x_0))}$. Since

$$\Phi_\epsilon(t,x,y) < 2\|u\|_\infty - Kf(|x-y|)-L|x-x_0|^2-C_0(t-t_0)^2\,,$$

then if we choose

$$L\geq\frac{2\|u\|_\infty}{r^2}\,,\qquad C_0\geq\frac{8\|u\|_\infty}{t_0^2}\,,\qquad K\geq\frac{2\|u\|_\infty}{f(\delta)}$$

we can exclude that the maximum of $\Phi_\epsilon$ in $\overline{\Delta}$ be attained when $t=\frac{t_0}{2}$, $t=\frac{3}{2}t_0\wedge T$, or when $|x-y|=\delta$ or when $|x-x_0|=r$. If $|y-x_0|=r$, then $|x-x_0|\geq r-\delta$ and choosing $r\geq 2\delta$ and $L\geq\frac{8\|u\|_\infty}{r^2}$ we also exclude that the maximum be attained when $|y-x_0|=r$.

On account of the above choices, we deduce that the maximum is attained inside $\Delta$. Moreover, this maximum being positive, it cannot be attained when $x=y$. Let $(\hat t,\hat x,\hat y)$ be the point in which the maximum is achieved. We proceed now as in the proof of Theorem 4.1 obtaining

$$\begin{aligned}\tfrac{\epsilon}{(T-\hat t)^2}+2C_0(\hat t-t_0)+F(\hat t,\hat x,KD\psi(\hat x-\hat y)+2L(\hat x-x_0),X)&\\ -F(\hat t,\hat y,KD\psi(\hat x-\hat y),Y)\leq h(\hat t,\hat x)-h(\hat t,\hat y),&\end{aligned} \tag{4.37}$$

where $\psi(\cdot)=f(|\cdot|)$, and where $X$, $Y$ satisfy

$$\begin{pmatrix}X-2LI & 0\\ 0 & -Y\end{pmatrix}\leq K\begin{pmatrix}D^2\psi(\hat x-\hat y) & -D^2\psi(\hat x-\hat y)\\ -D^2\psi(\hat x-\hat y) & D^2\psi(\hat x-\hat y)\end{pmatrix}+\frac{1}{n}(D^2z(\hat t,\hat x,\hat y))^2,$$

where $D^2z = D^2\left(K\psi(x-y)+L|x-x_0|^2\right)$. Using (4.9) and Hypothesis 4.1 we have

$$\begin{aligned}F(\hat t,\hat x,D\psi(\hat x-\hat y),X-2LI)-F(\hat t,\hat y,D\psi(\hat x-\hat y),Y)\geq -\lambda\mathrm{tr}\,(X-2LI-Y)\\ -Kf'(|\hat x-\hat y|)\,g(|\hat x-\hat y|)-(Kf'(|\hat x-\hat y|))^2\,[c_0+c_1\big(Kf'(|\hat x-\hat y|)\,|\hat x-\hat y|\big)^{q-1}]\omega(|\hat x-\hat y|)\\ -M-\tfrac{\theta_\epsilon(\hat t,\hat x,\hat y)}{n}\,\eta(\hat t,\hat x,\hat y).\end{aligned}$$

On the other hand, since $\max\Phi_\varepsilon>0$, we deduce that

$$Kf(|\hat x-\hat y|)\leq u(\hat t,\hat x)-u(\hat t,\hat y)\leq 2\|u\|_\infty,$$

hence we get, since $f$ is concave,

$$Kf'(|\hat{x}-\hat{y}|)\,|\hat{x}-\hat{y}| \le 2\|u\|_\infty\,.$$

Therefore we obtain

$$\begin{aligned} &F(\hat{t},\hat{x},D\psi(\hat{x}-\hat{y}),X-2LI)-F(\hat{t},\hat{y},D\psi(\hat{x}-\hat{y}),Y)\\ &\ge -\lambda \mathrm{tr}\,(X-2LI-Y)-Kf'(|\hat{x}-\hat{y}|)\,g(|x-y|)\\ &-K^2\,(f'(|\hat{x}-\hat{y}|))^2\,[c_0+c_1\,(2\|u\|_\infty)^{q-1}]\omega(|\hat{x}-\hat{y}|)-M-\tfrac{\theta_\epsilon(\hat{t},\hat{x},\hat{y})}{n}\,\eta(\hat{t},\hat{x},\hat{y}). \end{aligned}$$

Using the coupling argument as in the proof of Theorem 4.1 we estimate $\mathrm{tr}\,(X-2LI-Y)$ and we conclude that

$$\begin{aligned} &F(\hat{t},\hat{x},D\psi(\hat{x}-\hat{y}),X-2LI)-F(\hat{t},\hat{y},D\psi(\hat{x}-\hat{y}),Y)\\ &\ge -4\lambda Kf''(|\hat{x}-\hat{y}|)-Kf'(|\hat{x}-\hat{y}|)\,g(|x-y|)\\ &-K^2\,(f'(|\hat{x}-\hat{y}|))^2\,[c_0+c_1\,(2\|u\|_\infty)^{q-1}]\omega(|\hat{x}-\hat{y}|)-M-\tfrac{\theta_\epsilon(\hat{t},\hat{x},\hat{y})}{n}\,\left(2N\lambda+\eta(\hat{t},\hat{x},\hat{y})\right). \end{aligned}$$

Therefore we have from (4.37)

$$\begin{aligned} &\tfrac{\epsilon}{(T-\hat{t})^2}+2C_0(\hat{t}-t_0)\le 4\lambda Kf''(|\hat{x}-\hat{y}|)-Kf'(|\hat{x}-\hat{y}|)\,g(|\hat{x}-\hat{y}|)\\ &+K^2\,(f'(|\hat{x}-\hat{y}|))^2\,[c_0+c_1\,(2\|u\|_\infty)^{q-1}]\omega(|\hat{x}-\hat{y}|)+M+\tfrac{\theta_\epsilon(\hat{t},\hat{x},\hat{y})}{n}\,\left(2N\lambda+\eta(\hat{t},\hat{x},\hat{y})\right)+2\|h\|_\infty\\ &+F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y}),X-2LI)-F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y})+2L(\hat{x}-x_0),X)\,. \end{aligned}$$

Using the equation satisfied by $f$, we get

$$\begin{aligned} &\tfrac{\epsilon}{(T-\hat{t})^2}+2C_0(\hat{t}-t_0)\le -K+c_\lambda K[c_0+c_1\,(2\|u\|_\infty)^{q-1}]\omega(|\hat{x}-\hat{y}|)+M+2\|h\|_\infty\\ &+\tfrac{\theta_\epsilon(\hat{t},\hat{x},\hat{y})}{n}\,\left(2N\lambda+\eta(\hat{t},\hat{x},\hat{y})\right)\\ &+F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y}),X-2LI)-F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y})+2L(\hat{x}-x_0),X)\,. \end{aligned} \tag{4.38}$$

We fix henceforth $L=\frac{8\|u\|_\infty}{r^2}$, $C_0=\frac{8\|u\|_\infty}{t_0^2}$ and $K=\frac{\kappa\|u\|_\infty}{\delta^2}$ for some $\kappa=\kappa(g,\lambda)$ so that $K\ge\frac{2\|u\|_\infty}{f(\delta)}$. We will later choose $\delta$ suitably small (this implies in turn that $K$ is sufficiently large).

The last term in (4.38) is estimated using assumption (4.34) with $S=\overline{B}_r(x_0)$. Since $|KD\psi(\hat{x}-\hat{y})|\le C\frac{\|u\|_\infty}{\delta}$, with $C=C(\lambda,g)$, we deduce that there exists some constant $c=c(N,L,\lambda,g,S)$ such that

$$\begin{aligned} &F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y}),X-2LI)-F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y})+2L(\hat{x}-x_0),X)\le\\ &\le c\left\{(1+\left(\tfrac{\|u\|_\infty}{\delta}+Lr\right)^m L(|\hat{x}-x_0|+1)\right\}\,. \end{aligned}$$

Recalling that $K\simeq\frac{\|u\|_\infty}{\delta^2}$, and $L$ and $|\hat{x}-x_0|$ are only estimated in terms of $r$, we get

$$F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y}),X-2LI)-F(\hat{t},\hat{x},KD\psi(\hat{x}-\hat{y})+2L(\hat{x}-x_0),X)\le C(\|u\|_\infty,r)K^{\frac{m}{2}}.$$

We obtain then from (4.38)

$$\begin{aligned} &\tfrac{\epsilon}{(T-\hat{t})^2}+2C_0(\hat{t}-t_0)\le -K+c_\lambda K\,(2\|u\|_\infty)^{q-1}\,\omega(|\hat{x}-\hat{y}|)+M+2\|h\|_\infty\\ &C(\|u\|_\infty,r)K^{\frac{m}{2}}+\tfrac{\theta_\epsilon(\hat{t},\hat{x},\hat{y})}{n}\,\left(2N\lambda+\eta(\hat{t},\hat{x},\hat{y})\right). \end{aligned}$$

We let $n\to\infty$. Then, since $m<2$ and $\omega(0)=0$, we choose $\delta$ small (depending also on $r$, $\|h\|_\infty$ and $\|u\|_\infty$) so that

$$\frac{\epsilon}{(T-\hat{t})^2}+2C_0(\hat{t}-t_0)\le-\frac{K}{2}$$

and we conclude taking $\delta$ eventually smaller so that $K > 2C_0t_0$ (i.e., $\delta \le C_3\sqrt{t_0 \wedge 1}$). In this way we have proved (4.35), which implies, if we take $x = x_0$ and $t = t_0$ (and since $Kf(|x-y|) \le Kf'(0)|x-y| \le \tilde{c}\frac{\|u\|_\infty}{\delta}|x-y|$)

$$u(t_0, x_0) - u(t_0, y) \le \tilde{c}|x_0 - y|, \quad |x_0 - y| < \delta\,.$$

If we take now $x$, $y$ belonging to some ball $B_R(x_0)$, then we easily extend a similar estimate to $x$, $y$. ■

4.5. **Examples and applications.** In this subsection we consider examples of operators to which the previous results can be applied, in particular checking the Hypotheses 4.1 and 4.4.

4.5a. **Bellman-Isaacs operators.** Let us consider the case of Bellman-Isaacs equations appearing in stochastic control problems or game theory (see [Kr80], [FS06], [DL06], [Ko09] and the references therein) where, for instance,

$$F(t,x,Du,D^2u) = \inf_{\beta\in\mathcal{B}}\sup_{\alpha\in\mathcal{A}}\left\{-\mathrm{tr}\left(q_{\alpha,\beta}(t,x)D^2u\right) - b_{\alpha,\beta}(t,x)\cdot Du - f_{\alpha,\beta}(t,x)\right\}\,.$$

As a preliminary assumption, we suppose that $F(t,x,p,X)$ is finite for every $(t,x) \in Q_T$, $p \in \mathbb{R}^N$, $X \in \mathcal{S}_N$. For example, this is certainly true if

$$\inf_{\beta\in\mathcal{B}}\sup_{\alpha\in\mathcal{A}}\left\{\mathrm{tr}\left(q_{\alpha,\beta}(t,x)\right) + |b_{\alpha,\beta}(t,x)| + |f_{\alpha,\beta}(t,x)|\right\} < \infty\,.$$

We show now that Hypotheses 4.1 and 4.4 are satisfied provided that the following conditions hold.

(i) $q_{\alpha,\beta}(t,x) = \lambda I + \sigma_{\alpha,\beta}(t,x)^2$, for some $\lambda > 0$, where the coefficients $\sigma_{\alpha,\beta}(t,x)$ and $b_{\alpha,\beta}(t,x)$ are continuous on $Q_T$, uniformly in $\alpha \in \mathcal{A}$ and $\beta \in \mathcal{B}$, and satisfy (3.7), uniformly in $\alpha \in \mathcal{A}$, $\beta \in \mathcal{B}$.

(ii) $f_{\alpha,\beta}$ are continuous and have bounded oscillation on $\bar{Q}_T$ (uniformly in $\alpha$ and $\beta$).

(iii) For any $(t,x) \in Q_T$, we have

$$\sup_{\beta\in\mathcal{B}}\sup_{\alpha\in\mathcal{A}}\left\{\mathrm{tr}\left(q_{\alpha,\beta}(t,x)\right)\right\} < \infty,$$

(iv) There exists $\varphi \in C^{1,2}(\bar{Q}_T)$ such that $\varphi \to +\infty$ as $|x| \to \infty$ (uniformly in $[0,T]$) and

$$\partial_t\varphi + \left\{-\mathrm{tr}\left(q_{\alpha,\beta}(t,x)D^2\varphi\right) - b_{\alpha,\beta}(t,x)\cdot D\varphi\right\} \ge 0 \quad \text{in } Q_T, \text{ for every } \alpha\in\mathcal{A}, \beta\in\mathcal{B}.$$

To check Hypothesis 4.1, we multiply the matrix inequality (4.3) by

$$\begin{pmatrix} \sigma_{\alpha,\beta}(t,x)^2 & \sigma_{\alpha,\beta}(t,x)\sigma_{\alpha,\beta}(t,y) \\ \sigma_{\alpha,\beta}(t,y)\sigma_{\alpha,\beta}(t,x) & \sigma_{\alpha,\beta}(t,y)^2 \end{pmatrix},$$

then, taking traces, we deduce

$$\begin{aligned} &-\mathrm{Tr}\left((q_{\alpha,\beta}(t,x)X - q_{\alpha,\beta}(t,y)Y\right) \qquad (4.39)\\ &\ge -\lambda\mathrm{Tr}\,(X-Y) - \mu\mathrm{Tr}\left((\sigma_{\alpha,\beta}(t,x)-\sigma_{\alpha,\beta}(t,y))^2\right) - \nu\mathrm{Tr}\left(\sigma^2_{\alpha,\beta}(t,x)+\sigma^2_{\alpha,\beta}(t,y)\right)\\ &\ge -\lambda\mathrm{Tr}\,(X-Y) - \mu\|(\sigma_{\alpha,\beta}(t,x)-\sigma_{\alpha,\beta}(t,y)\|^2 - \nu\mathrm{Tr}\left(q_{\alpha,\beta}(t,x)+q_{\alpha,\beta}(t,y)-2\lambda I\right), \end{aligned}$$

for every $\alpha \in \mathcal{A}$, $\beta \in \mathcal{B}$. Therefore

$$\begin{aligned} &\left\{-\mathrm{tr}\left(q_{\alpha,\beta}(t,x)X\right) - \mu b_{\alpha,\beta}(t,x)\cdot(x-y) - f_{\alpha,\beta}(t,x)\right\} \ge\\ &\ge \left\{-\mathrm{tr}\left(q_{\alpha,\beta}(t,y)Y\right) - \mu b_{\alpha,\beta}(t,y)\cdot(x-y) - f_{\alpha,\beta}(t,y)\right\}\\ &-\lambda\mathrm{Tr}\,(X-Y) - \mu\left[\|(\sigma_{\alpha,\beta}(t,x)-\sigma_{\alpha,\beta}(t,y)\|^2 + (b_{\alpha,\beta}(t,x)-b_{\alpha,\beta}(t,y))\cdot(x-y)\right]\\ &-(f_{\alpha,\beta}(t,x)-f_{\alpha,\beta}(t,y)) - \nu\eta(t,x,y) \end{aligned}$$

where

$$\eta(t,x,y)=\sup_{\beta\in\mathcal{B}}\sup_{\alpha\in\mathcal{A}}\left\{\mathrm{Tr}(q_{\alpha,\beta}(t,x))+\mathrm{Tr}(q_{\alpha,\beta}(t,y))\right\}-2\lambda N.$$

Note that $\eta(t,x,y)$ is finite thanks to (iii). Thus we get

$$\begin{aligned}&\{-\mathrm{tr}\,(q_{\alpha,\beta}(t,x)X)-\mu b_{\alpha,\beta}(t,x)\cdot(x-y)-f_{\alpha,\beta}(t,x)\}\geq\\ &\geq\{-\mathrm{tr}\,(q_{\alpha,\beta}(t,y)Y)-\mu b_{\alpha,\beta}(t,y)\cdot(x-y)-f_{\alpha,\beta}(t,y)\}\\ &-\lambda\mathrm{Tr}\,(X-Y)-\mu|x-y|g(|x-y|)-M-\nu\eta(t,x,y),\end{aligned}$$

for every $x,y\in\mathbb{R}^N$, $|x-y|\leq 1$, $t\in(0,T)$.

Taking the sup on $\alpha$ and the inf on $\beta$ implies that $F$ satisfies Hypothesis 4.1. In order to see that Hypothesis 4.4 is satisfied, we use (iv) which implies

$$\begin{aligned}&\left\{-\mathrm{tr}\left(q_{\alpha,\beta}(t,x)(X+\varepsilon D^2\varphi\right)-b_{\alpha,\beta}(t,x)\cdot(p+\varepsilon D\varphi)-f_{\alpha,\beta}(t,x)\right\}\\ &=\{-\mathrm{tr}\,(q_{\alpha,\beta}(t,x)X)-b_{\alpha,\beta}(t,x)\cdot p-f_{\alpha,\beta}(t,x)\}\\ &-\varepsilon\left\{\mathrm{tr}\left(q_{\alpha,\beta}(t,x)D^2\varphi\right)+b_{\alpha,\beta}(t,x)\cdot D\varphi\right\}\\ &\geq\{-\mathrm{tr}\,(q_{\alpha,\beta}(t,x)X)-b_{\alpha,\beta}(t,x)\cdot p-f_{\alpha,\beta}(t,x)\}-\varepsilon\partial_t\varphi\end{aligned}$$

and taking $\sup_\alpha$ and $\inf_\beta$ on both sides we deduce that Hypothesis 4.4 is satisfied.

Let us note that a sufficient condition which implies both (iii) and (iv) (hence Hypothesis 4.4) is that there exists $C\geq 0$ such that

$$\mathrm{tr}\,(q_{\alpha,\beta}(t,x))+b_{\alpha,\beta}(t,x)\cdot x\leq C(1+|x|^2), \tag{4.40}$$

$(t,x)\in Q_T$, $\alpha\in\mathcal{A},\beta\in\mathcal{B}$. Indeed, in this case (iv) is satisfied with $\varphi(t,x)=e^{Mt}(1+|x|^2)$ for some suitable $M>0$.

We point out that if $F$ consists only of $\sup_{\alpha\in\mathcal{A}}\{\cdot\}$ (i.e., we have a Bellman operator), then Lipschitz continuity of solutions of (4.1) holds even if $q_\alpha$ are degenerate, assuming Lipschitz continuity of coefficients (see [YZ99] and [BCQ10] for the joint Lipschitz continuity in $(t,x)$).

4.5b. **The case of nonlinear Hamiltonians.** Here we consider the case when the nonlinearity only concerns the first order terms, namely the equation

$$\partial_t u-\mathrm{tr}\left(q(t,x)D^2u\right)=H(t,x,Du)+h(t,x)\qquad\text{in } Q_T, \tag{4.41}$$

where $H$ is continuous on $Q_T\times\mathbb{R}^N$, $q$ is continuous on $Q_T$ and $h$ is continuous on $Q_T$ and has bounded oscillation. We also assume that

$$H(t,x,0)=0\,.$$

In general, one can reduce to this case whenever $H(t,x,0)$ has bounded oscillation. We assume the ellipticity condition (3.2), and we suppose that $q(t,x)$ and $H(t,x,p)$ satisfy:

$$\begin{aligned}&\exists\text{ a non-negative function } g\in C(0,1)\text{ such that }\textstyle\int_0^1 g(s)ds<\infty\text{ and}\\ &\mu\|\sigma(t,x)-\sigma(t,y)\|^2+\big(H(t,x,\mu(x-y))-H(t,y,\mu(x-y)\big)\leq\\ &\leq\mu|x-y|\,g(|x-y|)+(\mu|x-y|)^2(c_0+c_1(\mu|x-y|^2)^{q-1})\omega(|x-y|)+M,\end{aligned} \tag{4.42}$$

for every $\mu>0$, $x,y\in\mathbb{R}^N$ such that $0<|x-y|\leq 1$ and every $t\in(0,T)$, where $\omega$ is some function such that $\lim_{r\to 0^+}\omega(r)=0$, $q>1$, $c_0,c_1,M\geq 0$ and $\sigma(t,x)=\sqrt{q(t,x)-\lambda I}$.

Assumptions (3.2) and (4.42) imply that Hypothesis 4.1 holds true for the operator

$$F(t,x,p,X)=-\mathrm{tr}\,(q(t,x)X)-H(t,x,p)$$

Indeed, whenever $X$, $Y$ satisfy (4.3), we deduce

$$\mathrm{Tr}(q(t,x)X - q(t,y)Y) \le \lambda \mathrm{Tr}(X-Y) + \mu\|\sigma(t,x)-\sigma(t,y)\|^2 + \nu\left(\mathrm{Tr}(q(t,x)) + \mathrm{Tr}(q(t,y)) - 2\lambda N\right).$$

Therefore, using (4.42) we deduce that $F(t,x,p,X)$ satisfies Hypothesis 4.1 with $\eta(t,x,y) = \big(\mathrm{Tr}(q(t,x)) + \mathrm{Tr}(q(t,y))\big) - 2\lambda N$. As far as the Lyapunov condition is concerned, in this example Hypothesis 4.4 reads as follows:

For any $L > 0$, $\exists\ \varphi = \varphi_L \in C^{1,2}(\bar{Q}_T)$, $\varepsilon_0 = \epsilon_0(L) > 0$:

$$\begin{cases} \partial_t \varphi - \mathrm{Tr}(q(t,x)D^2\varphi) - \frac{1}{\varepsilon}\{H(t,x,p+\varepsilon D\varphi) - H(t,x,p)\} \ge 0 \\ \qquad \text{for every } (t,x) \in Q_T,\ p \in \mathbb{R}^N\text{: } |p| \le L + \varepsilon|D\varphi|, \text{ and every } \varepsilon \le \varepsilon_0 \\ \varphi(t,x) \to +\infty \quad \text{as } |x| \to \infty, \text{ uniformly for } t \in [0,T]. \end{cases} \tag{4.43}$$

In particular, under the previous assumptions (3.2), (4.42) and (4.43), the conclusion of Theorem 4.5 applies to equation (4.41).

**Remark 4.21.** Assume that $H(t,x,p)$ satisfies, for every $(t,x) \in Q_T$, every $p$, $\xi \in \mathbb{R}^N$:

$$|H(t,x,p+\xi) - H(t,x,p) - H(t,x,\xi)| \le \gamma(t,x)|\xi| + c(t,x)(|\xi| + |p|)^q|\xi| \tag{4.44}$$

for some non-negative functions $c(t,x)$, $\gamma(t,x)$, with $q \ge 0$. Observe that the function $c(t,x)$ accounts for the possibly superlinear growth of $H(x,p)$ with respect to $p$. If we also assume that $c(t,x)$ is bounded and

$$\begin{gathered} \exists\ \varphi \in C^{1,2}(\bar{Q}_T),\ \hat{\varepsilon}_0 > 0\ :\ \text{ for every } (t,x) \in Q_T \text{ and } \varepsilon \le \hat{\varepsilon}_0 \\ \begin{cases} \mathrm{tr}\left(q(t,x)D^2\varphi\right) + \frac{1}{\varepsilon}H(t,x,\varepsilon D\varphi) + \gamma(t,x)\,|D\varphi| + \hat{\varepsilon}_0|D\varphi|^{q+1} \le \partial_t\varphi \\ \varphi(t,x) \to +\infty \quad \text{as } |x| \to \infty, \text{ uniformly in } t \in (0,T), \end{cases} \end{gathered} \tag{4.45}$$

then condition (4.43) is satisfied. Indeed, in this case one has

$$\begin{aligned} \tfrac{1}{\varepsilon}\{H(t,x,p+\varepsilon D\varphi) - H(t,x,p)\} &\le \tfrac{1}{\varepsilon}H(t,x,\varepsilon D\varphi) + \gamma(t,x)\,|D\varphi| \\ &\quad + c(t,x)(|\varepsilon D\varphi| + |p|)^q|D\varphi|\,. \end{aligned}$$

Since, for every $p$: $|p| \le L + \varepsilon|D\varphi|$, we have

$$c(t,x)(|\varepsilon D\varphi| + |p|)^q|D\varphi| \le \varepsilon^q K(\|c\|_\infty)|D\varphi|^{q+1} + \tfrac{\hat{\varepsilon}_0}{2}\,|D\varphi|^{q+1} + K(L,\|c\|_\infty)$$

if $\varepsilon$ is small we deduce that

$$\frac{1}{\varepsilon}\{H(t,x,p+\varepsilon D\varphi) - H(t,x,p)\} \le \frac{1}{\varepsilon}H(t,x,\varepsilon D\varphi) + \gamma(t,x)\,|D\varphi| + \hat{\varepsilon}_0|D\varphi|^{q+1} + K(L,\|c\|_\infty)$$

hence (4.45) implies

$$\partial_t\varphi - \mathrm{Tr}(q(t,x)D^2\varphi) - \frac{1}{\varepsilon}\{H(t,x,p+\varepsilon D\varphi) - H(t,x,p)\} \ge -K(L,\|c\|_\infty)\,.$$

Then $\tilde{\varphi} := \varphi + K(L,\|c\|_\infty)t$ satisfies (4.43).

As a model case which can be dealt with, consider the following equation

$$\partial_t u - \mathrm{tr}\left(q(t,x)D^2u\right) - b(t,x)\cdot Du + \Psi(t,x,Du) + c(t,x)|Du|^{q+1} = h(t,x) \quad \text{in } Q_T. \tag{4.46}$$

We deduce then the following

**Corollary 4.22.** *Consider the equation (4.46), with $q \le 1$. Assume that $q(t,x)$ satisfies (3.2) and that $\sqrt{q(t,x) - \lambda I}$ and $b(t,x)$ satisfy (3.7). Suppose that $h$ has bounded oscillation and let*

*$\Psi(t,x,\xi)$ be a continuous function on $Q_T\times\mathbb{R}^N$ which satisfies $\Psi(t,x,0)=0$ and, for $\xi,\eta\in\mathbb{R}^N$, $x,y\in\mathbb{R}^N$, $0<|x-y|\le 1$, $t\in(0,T)$,*

$$\begin{aligned} &(i) \qquad |\Psi(t,x,\xi)-\Psi(t,x,\eta)|\le\gamma(t,x)|\xi-\eta| \\ &(ii) \qquad |\Psi(t,x,\xi)-\Psi(t,y,\xi)|\le g(|x-y|)\,|\xi|\,, \end{aligned}$$

*for some continuous functions $\gamma(t,x)$ on $Q_T$ and $g\in L^1(0,1)\cap C(0,1;\mathbb{R}_+)$. Assume that $c(t,x)$ is bounded and continuous on $Q_T$, uniformly continuous with respect to $x$ (uniformly for $t\in(0,T)$). Assume in addition that*

$$\exists\,\varphi\in C^{1,2}(\bar Q_T)\ :\ \begin{cases}\mathrm{tr}\left(q(t,x)D^2\varphi\right)+b(t,x)\cdot D\varphi+\gamma(t,x)|D\varphi|\le\partial_t\varphi \ \textit{ in } Q_T,\\ \varphi(t,x)\to+\infty \quad \textit{as } |x|\to\infty, \textit{ uniformly for } t\in[0,T]\end{cases} \tag{4.47}$$

*Then the conclusion of Theorem 4.5 holds true.*

*Proof.* Here we have $H(t,x,p)=b(x)\cdot p+\Psi(t,x,p)+c(t,x)|p|^q$. Using (3.7) and the assumption (ii) on $\Psi$, we immediately deduce that (4.42) is satisfied with $c_1=0$ and $\omega$ being the modulus of uniform continuity of $c(t,\cdot)$.

Given $\varphi$ satisfying (4.47), consider now $\phi:=\log(\varphi+\eta)+Rt$ (with constant $\eta$ such that, $\varphi+\eta\ge 1$). Since the assumptions on $\Psi$ imply

$$\frac{1}{\varepsilon}\left\{H(t,x,p+\varepsilon D\varphi)-H(t,x,p)\right\}\le b(t,x)\cdot D\varphi+\gamma(t,x)\,|D\varphi|+\frac{1}{\varepsilon}c(t,x)\left\{|p+\varepsilon D\varphi|^{q+1}-|p|^{q+1}\right\}$$

we deduce, for $\varepsilon$ small and for every $p$: $|p|\le L+\varepsilon|D\phi|$, and using that $q\le 1$, that $\phi$ satisfies

$$\begin{aligned} &\partial_t\phi-\mathrm{Tr}(q(t,x)D^2\phi)-\tfrac{1}{\varepsilon}\left\{H(t,x,p+\varepsilon D\phi)-H(t,x,p)\right\}\\ &\ge R+\tfrac{1}{\varphi+\eta}\left(\partial_t\varphi-\mathrm{tr}\left(q(t,x)D^2\varphi(x)\right)-b(t,x)\cdot D\varphi-\gamma(t,x)|D\varphi|\right)\\ &\quad -K(q,\|c\|_\infty)\tfrac{|D\varphi|}{\varphi+\eta}\left(L+2\varepsilon\tfrac{|D\varphi|}{\varphi+\eta}\right)^q+\tfrac{q(t,x)D\varphi\cdot D\varphi}{(\varphi+\eta)^2}\\ &\ge R-\tfrac{\lambda}{2}\tfrac{|D\varphi|^2}{(\varphi+\eta)^2}-\tilde K+\tfrac{q(t,x)D\varphi\cdot D\varphi}{(\varphi+\eta)^2} \end{aligned}$$

for some constant $\tilde K=\tilde K(q,L,\lambda,\|c\|_\infty)$. Since $q(t,x)\ge\lambda I$, choosing $R$ sufficiently large we deduce that (4.43) is satisfied. Therefore Theorem 4.5 applies. ■

**Remark 4.23.** When $H(t,x,\xi)=b(t,x)\cdot\xi$, then (4.47) reduces to (3.5), and we recover the result of the linear case. In fact, whenever $\gamma(t,x)$ is also bounded, with the same argument as above we see that the term $\gamma(t,x)|D\varphi|$ can also be neglected in (4.47). In conclusion, if both the coefficients $\gamma(t,x)$ and $c(t,x)$ are bounded, and $q\le 1$, the Lyapunov condition of the linear case is enough to ensure the same condition for the nonlinear equation (4.46).

**Remark 4.24.** It could still be possible to consider the case when the function $c(t,x)$ in (4.46) is unbounded. As in Corollary 4.22, one can use (through a log-transform) the ellipticity of $q(t,x)$ to control the term with $c(t,x)$. In particular, whenever (4.47) is satisfied for some $\varphi$: $|D\varphi|\to\infty$ as $|x|\to\infty$ (uniformly in $t$), if we have $c(t,x)=$o $(l(t,x))$ as $|x|\to+\infty$, uniformly in $t$, where $l(t,x)=\inf_{|\xi|=1}q(t,x)\xi\cdot\xi$, it is still possible to use the above argument to conclude.

4.5c. **A Liouville type theorem.** Here we show a related Liouville type theorem which extends [PW06, Theorem 3.6] to the nonlinear setting.

Let $F:\mathbb{R}^N\times\mathbb{R}^N\times\mathcal{S}_N\to\mathbb{R}$ be a continuous function. We consider *bounded viscosity solutions* $v\in C_b(\mathbb{R}^N)$ to

$$F(x,Dv,D^2v)=0 \ \text{ in } \ \mathbb{R}^N. \tag{4.48}$$

We suppose that $F$ verifies (cf. Hypothesis 4.1): there exist $\lambda > 0$, non-negative functions $\eta(x, y)$ and $g \in C((0, +\infty); \mathbb{R}_+) \cap L^1(0, 1)$ such that
(4.49)
$$F(x, \mu(x-y), X) - F(y, \mu(x-y), Y) \geq -\lambda \mathrm{tr}\,(X - Y) - \mu|x-y|\, g(|x-y|) - \nu\, \eta(x, y),$$
for any $\mu > 0$, $\nu \geq 0$, $x, y \in \mathbb{R}^N$, matrices $X, Y \in \mathcal{S}_N$ which verify (4.3).

**Theorem 4.25.** *Suppose that $F$ satisfies Hypothesis 4.4 and condition (4.49) with $g$ such that*
$$\int_0^{+\infty} e^{-\frac{1}{4\lambda}\int_0^r g(s)ds} dr = +\infty. \tag{4.50}$$
*Then any bounded viscosity solution $v \in C_b(\mathbb{R}^N)$ of (4.48) is constant.*

*Proof.* First, one checks easily that $v(t, x) = v(x)$, $t \in (0, T)$, $x \in \mathbb{R}^N$, is also a viscosity solution to the parabolic equation
$$\partial_t v + F(x, Dv, D^2 v) = 0 \text{ in } Q_T.$$
Then as in the proof of Theorem 4.5, with the same notation, we fix $t_0 > 0$. We also consider $\delta \geq 1$ and define the open set
$$\Delta = \Delta(t_0, \delta) = \Big\{(t, x, y) \in (0, +\infty) \times \mathbb{R}^N \times \mathbb{R}^N \,:\, |x-y| < \delta\,,\ \frac{t_0}{2} < t < \frac{3}{2} t_0\Big\}$$
and the function
$$\Phi_\epsilon(t, x, y) = v(x) - v(y) - Kf(|x-y|) - \varepsilon\,(\varphi(t, x) + \varphi(t, y)) - C_0(t - t_0)^2,$$
where $\varphi(t, x)$ is the usual Lyapunov function given in Hypothesis 4.4 corrisponding to some constant $L > 0$ to be fixed later. Arguing by contradiction, we deduce that $\Phi_\varepsilon$ has a positive local maximum at $(\hat{t}, \hat{x}, \hat{y}) \in \triangle$, provided
$$C_0 = \frac{8\|v\|_\infty}{t_0^2}, \qquad K \geq \frac{2\|v\|_\infty}{f(\delta)},$$
where $f$ is the solution of (3.29), i.e., $f(r) = f_\delta(r)$, $r \in [0, \delta]$,
$$f(r) = \frac{1}{4\lambda}\int_0^r e^{-\frac{G(\xi)}{4\lambda}} \int_\xi^\delta e^{\frac{G(\tau)}{4\lambda}} d\tau d\xi\,, \qquad G(\xi) = \textstyle\int_0^\xi g(\tau)d\tau.$$
Note that $f(\delta) = f_\delta(\delta) \to +\infty$, as $\delta \to +\infty$. We end up with
$$\begin{aligned}2C_0(\hat{t} - t_0) + \varepsilon\,(\partial_t\varphi(\hat{t}, \hat{x}) + \partial_t\varphi(\hat{t}, \hat{y})) + F(\hat{x}, KD\psi(\hat{x} - \hat{y}) + \varepsilon\, D\varphi(\hat{t}, \hat{x}), X)\\ -F(\hat{y}, KD\psi(\hat{x} - \hat{y}) - \varepsilon\, D\varphi(\hat{t}, \hat{y}), Y) \leq 0,\end{aligned}$$
where $\psi(\cdot) = f(|\cdot|)$. Continuing as in the proof of Theorem 4.5, since now $c_0 = c_1 = M = 0$, we get
$$2C_0(\hat{t} - t_0) \leq -K + \tfrac{\theta_\epsilon(\hat{t}, \hat{x}, \hat{y})}{n}\eta(\hat{x}, \hat{y}) + \tfrac{2N\lambda\,\theta_\epsilon(\hat{t}, \hat{x}, \hat{y})}{n} + \mathcal{I}_\varepsilon(\varphi(\hat{t}, \hat{x})) + \mathcal{I}_\varepsilon(\varphi(\hat{t}, \hat{y})).$$
Now set
$$K = \left\{\frac{8\|v\|_\infty}{t_0} + \frac{2\|v\|_\infty}{f(\delta)}\right\}.$$
Since $K$ is fixed, we use Hypothesis 4.4 with $L = Kf'(0)$, which allows us to deduce that $\mathcal{I}_\varepsilon(\varphi(\hat{t}, \hat{x}))$, $\mathcal{I}_\varepsilon(\varphi(\hat{t}, \hat{y})) \leq 0$. In this way, letting $n \to \infty$, we get a contradiction.

As in the proof of Theorem 3.3, we obtain
$$|v(x) - v(y)| \leq Kf(|x-y|) \leq \left\{\frac{8\|v\|_\infty}{t_0} + \frac{2\|v\|_\infty}{f(\delta)}\right\} f(|x-y|), \quad |x-y| \leq \delta.$$

Using the concavity of $f$ we find

$$|v(x)-v(y)| \le \left\{ \frac{8\|v\|_\infty}{t_0} + \frac{2\|v\|_\infty}{f(\delta)} \right\} f'(0)|x-y|.$$

Since $f'(0) = \frac{1}{4\lambda}\int_0^\delta e^{\frac{1}{4\lambda}\int_0^r g(s)ds}dr$ we find, for any $t>0$, any $\delta \ge 1$, and any $x, y \in \mathbb{R}^N$ such that $|x-y| \le \delta$,

$$|v(x)-v(y)| \le \frac{1}{4\lambda} \left\{ \frac{8\|v\|_\infty}{t} + \frac{2\|v\|_\infty}{f(\delta)} \right\} \int_0^\delta e^{\frac{1}{4\lambda}\int_0^r g(s)ds}dr \ \cdot |x-y|.$$

Now letting $t \to +\infty$, we find, for $|x-y| \le \delta$,

$$|v(x)-v(y)| \le \frac{\|v\|_\infty}{2\lambda} \frac{1}{f(\delta)} \int_0^\delta e^{\frac{1}{4\lambda}\int_0^r g(s)ds}dr \ \cdot |x-y|. \tag{4.51}$$

Let us fix any $x$ and $y \in \mathbb{R}^N$. For any $\delta \ge |x-y|$, we have (4.51) (recall that $f = f_\delta$). Letting $\delta \to +\infty$, by using Hôpital's rule thanks to (4.50), we have

$$\lim_{\delta\to+\infty} \frac{1}{f(\delta)} \int_0^\delta e^{\frac{1}{4\lambda}\int_0^r g(s)ds}dr = 0$$

and so $v(x) = v(y)$. It follows that $v$ is constant. ∎

4.5d. **A remark on possible existence results.** Our estimates are particularly useful in cases in which the comparison principle is not known due to the irregularity of coefficients and so one can not perform the Perron's method to get existence of viscosity solutions. We recall (see [BBL02, Lemma 9.1]) that, under fairly general conditions on the function $F$, once a local modulus of continuity is established for the $x$-variable, then it is possible to deduce a local modulus of continuity for the $t$-variable. In particular, if Lipschitz continuity holds in the $x$-variable, then $\frac{1}{2}$-Hölder continuity holds in the $t$-variable. Thanks to such a result, the above Lipschitz, or Hölder estimates which we proved imply a local uniform equi–continuity for the viscosity solutions, and therefore a local compactness in the uniform (space–time) topology. As a consequence, suppose that the function $F$ can be approximated by a sequence $F_n$ of functions satisfying Hypotheses 4.1 and 4.4 or Hypotheses 4.6 and 4.7 uniformly with respect to $n$ and such that viscosity solutions $u_n$ are known to exist for the Cauchy problem

$$\partial_t u_n + F_n(t, x, Du_n, D^2 u_n) = 0,$$

with $u_n(0) = u_{0n}$ converging locally uniformly to $u_0$. Then, if $u_0 \in C_b(\mathbb{R}^N)$ (i.e. $u_0$ is continuous and bounded) we conclude the existence of a bounded viscosity solution $u$ which is Lipschitz, or Hölder, as $t > 0$. Similarly, we deduce the existence of a viscosity solution in case $u_0$ is continuous and has bounded oscillation provided the stronger Hypotheses 4.14 or 4.15 are satisfied. In particular, in the linear case, as well as in the case of Bellman-Isaacs equations, we deduce the existence of a viscosity solution by simply approximating the coefficients through a standard convolution regularization.

## Appendix A. Probabilistic Vs analytic approach

Let us spend a few words in the comparison between the probabilistic proof of [PW06, Theorem 3.4] and the previous analytical proof of Theorem 3.3. For simplicity, *we consider here that coefficients $q$ and $b$ are independent of time as in* [PW06], namely we consider

$$A = \sum_{i,j=1}^N q_{ij}(x)\partial^2_{x_i x_j} + \sum_{i=1}^N b_i(x)\partial_{x_i}.$$

The probabilistic approach used to prove estimate (1.1) relies on a so-called coupling method (see also [LR86], [CL89], [Cr91], [Cr92]). Let us briefly explain such method. Under the assumptions of Theorem 3.3, it is well known that there exist unique probability measures $\mathbb{P}^x$ and $\mathbb{P}^y$ on $\Omega_N := C([0,\infty);\mathbb{R}^N)$ which are martingale solutions for $A$ (starting from $x \in \mathbb{R}^N$ and $y \in \mathbb{R}^N$ respectively; see [SV79]). In particular, for $u_0 : \mathbb{R}^N \to \mathbb{R}$ continuous and bounded, this allows to define the probabilistic solution to the Cauchy problem

$$\begin{cases} \partial_t u - \operatorname{tr}\left(q(x)D^2u\right) - b(x)\cdot Du = 0 \\ u(0,\cdot) = u_0, \end{cases} \tag{A.1}$$

as

$$u(t,x) = \mathbb{E}^x\left[u_0(x_t)\right], \quad t \geq 0,$$

where $(x_t)$ denotes the canonical process over $\Omega_N$ with values in $\mathbb{R}^N$. Note, incidentally, that this is a viscosity solution (see [FS06]). Equivalently, we can also write $u = P_t u_0$ where $P_t$ is the diffusion Markov semigroup associated to (A.1).

A coupling for the measures $\mathbb{P}^x$ and $\mathbb{P}^y$ is a probability measure $\mathbb{P}^{x,y}$ on $\Omega_{2N} = C([0,\infty);\mathbb{R}^{2N})$ such that the two marginal distributions are, respectively, $\mathbb{P}^x$ and $\mathbb{P}^y$. For any coupling $\mathbb{P}^{x,y}$, we have that

$$u(t,x) - u(t,y) = \mathbb{E}^x\left[u_0(x_t)\right] - \mathbb{E}^y\left[u_0(x_t)\right] = \mathbb{E}^{x,y}\left[u_0(x_t) - u_0(y_t)\right], \tag{A.2}$$

where $\mathbb{E}^{x,y}$ denotes the expectation with respect to $\mathbb{P}^{x,y}$ and $(z_t) = \big((x_t,y_t)\big)$ denotes the canonical process over $\Omega_{2N}$ with values in $\mathbb{R}^{2N}$.

A Lipschitz estimate for $u(t)$ is then obtained if we are able to estimate the last term in (A.2). Introducing the coupling time $T_c := \inf\{t \geq 0 : x_t = y_t\}$ (with the usual convention $\inf(\emptyset) = \infty$), this estimate can be obtained by constructing a coupling $\mathbb{P}^{x,y}$ which verifies the property ($\mathbb{P}^{x,y}$-a.s.):

$$x_t = y_t \quad \text{for all } t \geq T_c \text{ on } \{T_c < \infty\}. \tag{A.3}$$

Indeed, if the coupling $\mathbb{P}^{x,y}$ satisfies the property (A.3), then we have clearly

$$\mathbb{E}^{x,y}\left[u_0(x_t) - u_0(y_t)\right] \leq 2\|u_0\|_\infty \mathbb{P}^{x,y}(t < T_c), \tag{A.4}$$

so that, in the end, one is left with the estimate of the first hitting time of the diagonal set $\Delta := \{(x,y) \in \mathbb{R}^{2N} : x = y\}$ for the process $\big((x_t,y_t)\big)$ starting from $(x,y) \notin \Delta$.

It is proved in [PW06] that a coupling measure $\mathbb{P}^{x,y}$ satisfying (A.3) can be constructed starting from any martingale solution for the differential operator in $\mathbb{R}^{2N}$

$$\mathcal{A}_c := \sum_{i,j}\left(q_{ij}(x)\partial^2_{x_i,x_j} + 2c_{ij}(x,y)\partial^2_{x_i,y_j} + q_{ij}(y)\partial^2_{y_i,y_j}\right) + \sum_i (b_i(x)\partial_{x_i} + b_i(y)\partial_{y_i}), \tag{A.5}$$

where the only requirement is that $c_{ij}$ are continuous on $\mathbb{R}^{2N}$ and the matrix $\begin{pmatrix} q(x) & c(x,y) \\ c^*(x,y) & q(y) \end{pmatrix}$ is symmetric and nonnegative. Indeed, if $\tilde{\mathbb{P}}^{x,y}$ is a martingale solution for $\mathcal{A}_c$, then one can define on $(\Omega_{2N}, \tilde{\mathbb{P}}^{x,y})$ the process

$$x'_t = \begin{cases} x_t, & t \leq T_c, \\ y_t, & t > T_c, \end{cases}$$

and, using that the martingale problem for $A$ is well-posed, one can prove that the process $(x_t)$ and $(x'_t)$ have the same law. On $(\Omega_{2N}, \tilde{\mathbb{P}}^{x,y})$ we define the process $\big((x'_t, y_t)\big)$. Its law is a coupling measure $\mathbb{P}^{x,y}$ satisfying also (A.3). Note that the martingale problem for $\mathcal{A}_c$ could be not well-posed due to the fact that $\mathcal{A}_c$ is possibly degenerate and has coefficients which are not locally Lipschitz in general.

Summing up, the probabilistic approach relies on the possibility that the Lipschitz regularity can be deduced in two main steps:

(i) construction of a suitable coupling $\mathbb{P}^{x,y}$ (satisfying (A.3) and associated to $\mathcal{A}_c$);
(ii) estimate of $\mathbb{P}^{x,y}(t < T_c)$ in terms of $|x-y|$.

In particular, if one has

$$\mathbb{P}^{x,y}(t < T_c) \leq \frac{k}{\sqrt{t \wedge 1}}|x-y|, \quad t > 0 \tag{A.6}$$

then (A.2) and (A.4) imply the desired Lipschitz estimate (1.1).

At the second stage (ii), the differential operator $\mathcal{A}_c$ (and so the right choice of $c(x,y)$) play a crucial role, since a quite standard procedure (see e.g. [Fr75]) to obtain estimates on the hitting time of the diagonal $\Delta$ is through the construction of suitable supersolutions $W$ for $\mathcal{A}_c$, i.e.,

$$\partial_t W - \mathcal{A}_c(W) \geq 0 \quad \text{in} \quad (0,T) \times (\mathbb{R}^{2N} \setminus \Delta).$$

Indeed, estimate (A.6) is proved in [PW06] after choosing $c(x,y)$ as in (3.22) and finding suitable supersolutions for the corresponding $\mathcal{A}_c$.

**Remark A.1.** The above approach can be also rephrased in terms of a Wasserstein distance between the transition probabilities $p(t,x,\cdot)$ and $p(t,y,\cdot)$, $t \geq 0$, $x,y \in \mathbb{R}^N$ (i.e., $p(t,x,A) = \mathbb{P}^x(x_t \in A)$, for any Borel set $A \subset \mathbb{R}^N$). More precisely if, given two Borel probability measures $\mu$, $\nu$ in $\mathbb{R}^N$, we define a distance (see e.g. [CL89]) as

$$d_W(\mu,\nu) = \inf_{Q \in \pi(\mu,\nu)} \int\int \chi(z,w) dQ(z,w), \qquad \chi(z,w) = \begin{cases} 1 & \text{if } z \neq w \\ 0 & \text{if } z = w, \end{cases}$$

where $\pi(\mu,\nu)$ is the set of all couplings of $\mu, \nu$ on $\mathbb{R}^{2N}$, then we have, being $\mathbb{P}^{x,y}$ a coupling of $\mathbb{P}^x$ and $\mathbb{P}^y$,

$$\mathbb{E}^{x,y}\left(u_0(x_t) - u_0(y_t)\right) \leq 2\|u_0\|_\infty d_W(p(t,x,\cdot), p(t,y,\cdot)).$$

Therefore from (A.2) we get

$$u(t,x) - u(t,y) \leq 2\|u_0\|_\infty \, d_W(p(t,x,\cdot), p(t,y,\cdot)) \tag{A.7}$$

(recall that a classical result by Dobrushin says that $d_W$ is just half of the total variation distance). In this framework, the Lipschitz estimate on $u(t,x)$ is reduced to an estimate of the Wasserstein distance between $p(t,x,\cdot)$ and $p(t,y,\cdot)$. On the other hand, if one has constructed a coupling process such that the corresponding measure $\mathbb{P}^{x,y}$ over $\Omega_{2N}$ satisfies (A.3), then it follows

$$d_W(p(t,x,\cdot), p(t,y,\cdot)) \leq \mathbb{P}^{x,y}(t < T_c)$$

and we end up with (A.4) again.

Let us now rephrase the analytic proof given in Theorem 3.3 in order to see its close correspondence with the probabilistic approach. The key point in such proof is the "maximum principle for semicontinuous functions" (Theorem 2.3) which is the heart of the viscosity solutions theory. In the linear framework, this fundamental result has the following immediate consequence, which is interesting in its own.

**Lemma A.2.** *Let $u$, $v$ be respectively a viscosity sub and supersolution of (A.1). Then, for every open set $\mathcal{O} \subset \mathbb{R}^{2N}$ and every $N \times N$ matrix $c(x,y)$ (with $c_{ij} \in C(\mathcal{O})$) such that $\begin{pmatrix} q(x) & c(x,y) \\ c^*(x,y) & q(y) \end{pmatrix}$ is nonnegative, we have that $u(t,x) - v(t,y)$ is a viscosity subsolution of*

$$\partial_t z - \mathcal{A}_c(z) = 0 \qquad \text{in } (0,T) \times \mathcal{O}. \tag{A.8}$$

*where $\mathcal{A}_c$ is the operator defined in (A.5).*

*In particular, if $u$ is a viscosity solution of (3.1), then $u(t,x)-u(t,y)$ is a viscosity subsolution of (A.8).*

*Proof.* For $\hat t \in (0,T), (\hat x,\hat y) \in \mathcal{O}$, and $z \in C^{1,2}((0,T)\times\mathcal{O})$, let $(\hat t,\hat x,\hat y)$ be a local maximum point of $u(t,x)-v(t,y)-z(t,x,y)$. Applying Theorem 2.3, for every $n>0$ there exist $a,b\in\mathbb{R}$, $X,Y\in\mathcal{S}_N$ such that $(a, D_x z(\hat t,\hat x,\hat y), X)\in \overline{P}^{2,+}u(\hat t,\hat x)$, $(b,-D_y z(\hat t,\hat x,\hat y),Y)\in\overline{P}^{2,-}v(\hat t,\hat y)$, $a-b=\partial_t z(\hat t,\hat x,\hat y)$ and

$$-\left(n+c_N\|D^2z\|\right)I \le \begin{pmatrix} X & 0\\ 0 & -Y\end{pmatrix} \le D^2z+\frac{1}{n}\left(D^2z\right)^2 \tag{A.9}$$

where $D^2z$ is computed at $(\hat t,\hat x,\hat y)$. We proceed then as in the proof of Theorem 3.3; using the equations of $u$ and $v$, subtracting the two and since $a-b=\partial_t z(\hat t,\hat x,\hat y)$, we find inequality (3.19). Now, if $c(x,y)$ is any $N\times N$ matrix such that $\begin{pmatrix} q(x) & c(x,y)\\ c^*(x,y) & q(y)\end{pmatrix}$ is nonnegative, multiplying (A.9) and taking traces we obtain the same as (3.21), hence

$$\begin{aligned}&\partial_t z(\hat t,\hat x,\hat y)-\operatorname{tr}\left(q(\hat x)D^2_x z+q(\hat y)D^2_y z-2c(\hat x,\hat y)D^2_{xy}z\right)\\ &\le b(\hat x)\cdot D_x z+b(\hat y)\cdot D_y z+\tfrac{1}{n}\operatorname{tr}\left(\begin{pmatrix} q(\hat x) & c(\hat x,\hat y)\\ c^*(\hat x,\hat y) & q(\hat y)\end{pmatrix}\left(D^2z\right)^2\right)\end{aligned}$$

Letting $n\to\infty$ we obtain

$$\partial_t z(\hat t,\hat x,\hat y)-\mathcal{A}_c(z)(\hat t,\hat x,\hat y)\le 0$$

and since this is true for every $z$ and $(\hat t,\hat x,\hat y)$ being a local maximum of $u(t,x)-v(t,y)-z(t,x,y)$, this means that $u(t,x)-v(t,y)$ is a viscosity subsolution of (A.8). ■

The previous lemma offers a nice interpretation, at least in the linear setting, to the use of matrix inequality (A.9) which is usually the essential part of the doubling variable technique for viscosity solutions. Indeed, it suggests that manipulations of the matrix inequality (A.9) amount to an optimization among all possible coupling operators. In this viewpoint, the Lipschitz estimate for $u$ is reduced to the existence of "good"supersolutions for some coupling operator $\mathcal{A}_c$. Since $u(t,x)-u(t,y)$ is a viscosity subsolution of a family of linear problems, thanks to the comparison principle for viscosity sub-super solutions, we expect that $u(t,x)-u(t,y)\le w(t,x,y)$ for every $w(t,x,y)$ being a supersolution of some operator $\mathcal{A}_c$. This fact could be shortly rephrased in the following suggestive form closer to inequality (A.7)

$$u(t,x)-u(t,y)\le\inf_{\mathcal{A}_c}\inf\{\psi(t,x,y),\ :\ \partial_t\psi-\mathcal{A}_c(\psi)\ge 0\ \}.$$

Thus, in the analytic approach the oscillation of $u$ can be estimated by choosing the best among all possible supersolutions of such family of operators. Let us make it more precise as follows.

**Lemma A.3.** *Assume (3.3) and let $u\in C(Q_T)$ be a solution of (3.1) such that $u(t,x)=o(\varphi)$ as $|x|\to\infty$. Assume that there exist some matrix $c(x,y)\in C(\mathbb{R}^{2N}\setminus\Delta;\mathbb{R}^{N^2})$ such that $\begin{pmatrix} q(x) & c(x,y)\\ c^*(x,y) & q(y)\end{pmatrix}$ is nonnegative, and a nonnegative function $\psi\in C^{1,2}((0,T)\times(\mathbb{R}^{2N}\setminus\Delta))$ such that*

$$\partial_t\psi-\mathcal{A}_c(\psi)\ge 0\qquad \textit{in } (0,T)\times(\mathbb{R}^{2N}\setminus\Delta).$$

*If, for some $\tau\in[0,T)$, we have $\psi(\tau,x,y)\ge u(\tau,x)-u(\tau,y)$, then*

$$u(t,x)-u(t,y)\le\psi(t,x,y),\qquad t\in[\tau,T).$$

*Proof.* Replacing $u(t,x)$ with $u^\varepsilon(t,x) = u(t,x) - \varepsilon\varphi(t,x)$ and $u(t,y)$ with $u^\varepsilon(t,y) = u(t,y) + \varepsilon\varphi(t,y)$ we have that $u^\varepsilon(t,x) - u^\varepsilon(t,y)$ is still a subsolution, and goes to $-\infty$ when $|x| \to \infty$ or $|y| \to \infty$, uniformly for $t \in (0,T)$. Consider the function $\Phi_\varepsilon = u^\varepsilon(t,x) - u^\varepsilon(t,y) - (\psi(t,x,y) + \frac{\varepsilon}{T-t})$. Assume by contradiction that $\sup_{[\tau,T)\times\mathbb{R}^{2N}} \Phi_\varepsilon > 0$. Since $\Phi_\varepsilon \to -\infty$ at infinity as $t \to T$ too, $\Phi_\varepsilon$ has a maximum at some point $(\hat{t},\hat{x},\hat{y})$. Since $u^\varepsilon(t,x) - u^\varepsilon(t,y) \le 0 \le \psi$ on $\Delta$, it cannot be $\hat{x} = \hat{y}$, and since $u^\varepsilon(\tau,x) - u^\varepsilon(\tau,y) \le \psi(\tau,x,y)$ by assumption, we have $\hat{t} \in (\tau,T)$. Therefore $(\hat{t},\hat{x},\hat{y})$ is a local maximum with $(\hat{x},\hat{y})$ lying in $\mathbb{R}^{2N} \setminus \Delta$ and by Lemma A.2 we deduce that

$$\frac{\varepsilon}{(T-t)^2} + \partial_t\psi - \mathcal{A}_c(\psi) \le 0$$

which gives a contradiction since $\psi$ is a supersolution. This proves that $\sup \Phi_\varepsilon \le 0$, i.e., $u^\varepsilon(t,x) - u^\varepsilon(t,y) \le \psi(t,x,y) + \frac{\varepsilon}{T-t}$. Letting $\varepsilon \to 0$ we conclude. ∎

The above approach represents actually a sort of analytic translation of the probabilistic coupling method. Indeed, if Lemma A.2 gives an analytic counterpart of the (A.7), Lemma A.3 gives the way for the estimate related to (ii). Indeed, one might expect, roughly speaking, that $z(t,x,y) = 2\|u_0\|_\infty \mathbb{P}^{x,y}(t < T_c)$ be a supersolution to the Cauchy parabolic problem involving $\mathcal{A}_c$ and having $u_0(x) - u_0(y)$ as initial datum at $t = 0$, with the additional property of being "minimal" in the sense that $z(t,x,x) = 0$, in other words that $\mathbb{P}^{x,y}(t < T_c)$ solves a Cauchy-Dirichlet parabolic problem for $\mathcal{A}_c$ with values 1 at $t = 0$ and 0 at $\Delta$. On the other hand, instead of checking this property on $\mathbb{P}^{x,y}(t < T_c)$, it is enough to build a supersolution $\psi$ with the desired features. Observe actually that the analytic proof of Theorem 3.3 basically follows from the previous lemmas up to showing that

$$\psi(t,x,y) = K\, f(|x-y|) + C_0(t-t_0)^2$$

is a supersolution in $(\frac{t_0}{2},T)$ for the coupling $c(x,y)$ given in (3.22), for a suitable choice of $K$, $C_0$ and $f$ satisfying (3.29). This is indeed the computational part given in the proof of Theorem 3.3 in Section 3.

To conclude this discussion, we have tried to underline the close analogy between the probabilistic and analytic approach, though expressed through different concepts and tools. We stress however some advantage of the approach through viscosity solutions: it only relies on maximum principle, which is a pointwise device, and it admits natural extensions to *nonlinear* operators as we have showed in Section 4.

## References


[Ba91] G. Barles, Interior gradient bounds for the mean curvature equation by viscosity solutions methods, Differential Integral Eq. **4** (1991), n.2, 263-275.

[Ba08] G. Barles, $C^{0,\alpha}$-regularity and estimates for solutions of elliptic and parabolic equations by the Ishii & Lions method, Gakuto International Series, Mathematical Sciences and Applications, vol. 30 (2008), 33-47.

[BBBL03] G. Barles, S. Biton, M. Bourgoing, O. Ley, Uniqueness results for quasilinear parabolic equations through viscosity solutions' methods. Calc. Var. Partial Differential Equations 18 (2003), no. 2, 159-179.

[BBL02] G. Barles, S. Biton, O. Ley, A geometrical approach to the study of unbounded solutions of quasilinear parabolic equations, Arch. Ration. Mech. Anal. 162 (2002), no. 4, 287-325.

[BF04] M. Bertoldi, S. Fornaro, Gradient estimates in parabolic problems with unbounded coefficients, Studia Math. 165 (2004), no. 3, 221-254,

[BCQ10] R. Buckdahn, P. Cannarsa, M. Quincampoix, Lipschitz continuity and semiconcavity properties of the value function of a stochastic control problem, NoDEA Nonlinear Differential Equations Appl. 17 (2010), no. 6, 715-728.

[CC95] L. Caffarelli, X. Cabré, Fully nonlinear elliptic equations, Amer. Math. Soc. Colloq. Publ., vol. 43, Amer. Math. Soc., Providence, RI, 1995.

[Ce96] S. Cerrai, Elliptic and parabolic equations in $\mathbb{R}^n$ with coefficients having polynomial growth, Comm. Partial Differential Equations 21 (1996), no. 1-2, 281-317.

[Ce01] S. Cerrai, Second order PDE's in finite and infinite dimensions. A probabilistic approach, Lectures Notes in Math. 1762, Springer Verlag, 2001.

[Ch93] Y. Z. Chen, $C^{1,\alpha}$ regularity of viscosity solutions of fully nonlinear elliptic PDE under natural structure conditions, J. Partial Differential Equations 6 (1993), 193-216.

[CL89] M. F. Chen, S. F. Li, Coupling methods for multi-dimensional diffusion processes, Ann. Probab. 17 (1989), 151-177.

[Cr91] M. Cranston, Gradient estimates on manifolds using coupling, J. Funct. Anal. 99 (1991), no. 1, 110-124.

[Cr92] M. Cranston, A probabilistic approach to gradient estimates, Canad. Math. Bull. 35 (1992), 46-55.

[CIL92] M. G. Crandall, H. Ishii, P. L. Lions, User's guide to viscosity solutions of second order partial differential equations. Bull. Amer. Math. Soc. (N.S.) 27 (1992), no. 1, 1-67.

[DZ02] G. Da Prato, J. Zabczyk, Second Order Partial Differential Equations in Hilbert Spaces, London Math. Soc. Lecture Note Ser., vol. 293, Cambridge, 2002.

[DL06] F. Da Lio, O. Ley, Uniqueness results for second-order Bellman-Isaacs equations under quadratic growth assumptions and applications. SIAM J. Control Optim. 45 (2006), no. 1, 74-106.

[EL94] K. D. Elworthy, X. M. Li, Formulae for the derivatives of heat semigroups. J. Funct. Anal. 125 (1994), no. 1, 252-286.

[FGP10] F. Flandoli, M. Gubinelli, E. Priola, Flow of diffeomorphisms for SDEs with unbounded Hölder continuous drift, Bulletin des Sciences Mathématiques, 134 (2010), 405-422.

[FS06] W. H. Fleming, H. M. Soner, Controlled Markov processes and viscosity solutions. (Second edition). Stochastic Modelling and Applied Probability, 25. Springer, New York, 2006.

[Fr75] A. Friedman, Stochastic differential equations and applications. Vol. 1, Academic Press, New York, San Francisco, London, 1975.

[Fr85] M. I. Freidlin, Functional Integration and Partial Differential Equations, Princeton University Press, 1985.

[GGIS91] Y. Giga, S. Goto, H. Ishii, and M. H. Sato, Comparison principle and convexity preserving properties for singular degenerate parabolic equations on unbounded domains, Indiana Univ. Math. J. 40 (1991), 443-469.

[HR98] D. G. Hobson, L. C. G. Rogers, Complete models with stochastic volatility, Math. Finance 8 (1998), no. 1, 27-48.

[IL90] H. Ishii, P. L. Lions, Viscosity solutions of fully nonlinear second-order elliptic partial differential equations, J. Differential Equations 83 (1990), no. 1, 26-78.

[Ko09] J. Kovats, Differentiability properties of solutions of nondegenerate Isaacs equations, Nonlinear Analysis: Theory, Methods & Applications, 71, Issue 12, 2009, 2418-2426.

[Kr80] N. V. Krylov, Controlled Diffusion Processes, Applications of Mathematics 14, Springer-Verlag 1980.

[KS80] N. V. Krylov, M. V. Safonov, A property of the solutions of parabolic equations with measurable coefficients. Izv. Akad. Nauk SSSR Ser. Mat. 44 (1980), no. 1, 161-175.

[KLL10] M. Kunze, L. Lorenzi, A. Lunardi, Nonautonomous Kolmogorov parabolic equations with unbounded coefficients, Trans. Amer. Math. Soc. 362 (2010), no. 1, 169-198.

[LR86] T. Lindvall, C. Rogers, Coupling of multidimensional diffusions by reflection, Ann. Probab. 14 (1986), 860-872.

[L98] A. Lunardi, Schauder theorems for linear elliptic and parabolic problems with unbounded coefficients in $\mathbb{R}^n$, Studia Math. 128 (1998), no. 2, 171-198.

[MPW02] G. Metafune, D. Pallara, M. Wacker, Feller semigroups on $\mathbb{R}^N$, Semigroup Forum 65 (2002), no. 2, 159-205.

[PW06] E. Priola, F. Y. Wang, Gradient estimates for diffusion semigroups with singular coefficients, J. Funct. Anal. 236 (2006), no. 1, 244-264.

[St74] H. B. Stewart, Generation of Analytic Semigroups by Strongly Elliptic Operators, Trans. Amer. Math. Soc. 199 (1974), 141-162.

[SV79] D. W. Stroock, S. R. S. Varadhan, Multidimensional diffusion processes, Springer-Verlag, 1979.

[YZ99] J. Yong, X. Y. Zhou, Stochastic controls: Hamiltonian systems and HJB equations, Springer, 1999.

[WZ10] F. Y. Wang, T. S. Zhang, Gradient estimates for stochastic evolution equations with non-Lipschitz coefficients, J. Math. Anal. Appl., 365, 2010, 1-11.

Dipartimento di Matematica, Università di Roma "Tor Vergata", Via della Ricerca Scientifica 1, 00133 Roma

*E-mail address*: porretta@mat.uniroma2.it

Dipartimento di Matematica, Università di Torino, Via Carlo Alberto 10, 10123 Torino, Italy

*E-mail address*: enrico.priola@unito.it