A lower spatially Lipschitz bound for solutions to fully nonlinear parabolic equations and its optimality

Nao Hamamuki [*] Suguru Kikkawa [†]

August 30, 2020

Abstract

We derive a lower spatially Lipschitz bound for viscosity solutions to fully nonlinear parabolic partial differential equations when the initial datum belongs to the Hölder space. The resulting estimate depends on the initial Hölder exponent and the growth rates of the equation with respect to the first and second order derivative terms. Our estimate is applicable to equations which are possibly singular at the initial time. Moreover, it gives the optimal rate of the regularizing effect for solutions, which occurs for some uniformly parabolic equations and first order Hamilton–Jacobi equations. In the proof of our lower estimate, we construct a subsolution and a supersolution by optimally rescaling the solution of the heat equation and then compare them with the solution. For linear equations, the lower spatially Lipschitz bound for solutions can be obtained in a different way if the fundamental solution satisfies the Aronson estimate. Examples include the heat convection equation whose convection term has singularities.

Key words: Lower spatially Lipschitz bound, Viscosity solutions, Fully nonlinear equations, Regularizing effect, Aronson estimates
Mathematics Subject Classification 2020: 35B65; 35D40; 35K15

1 Introduction

Purpose. In this paper, we consider fully nonlinear parabolic partial differential equations of the form

$$u_t(x,t) + F(x,t,u(x,t),\nabla u(x,t),\nabla^2 u(x,t)) = 0 \quad \text{in } \mathbf{R}^n \times (0,T) \tag{1.1}$$

[*]Department of Mathematics, Faculty of Science, Hokkaido University, Kita 10, Nishi 8, Kita-Ku, Sapporo, Hokkaido, 060-0810, Japan. E-mail: `hnao@math.sci.hokudai.ac.jp`
[†]Department of Mathematics, Faculty of Science, Hokkaido University, Kita 10, Nishi 8, Kita-Ku, Sapporo, Hokkaido, 060-0810, Japan. E-mail: `kikkawa.suguru@gmail.com`

under the initial condition

$$u(x,0)=\phi(x) \quad \text{in } \mathbf{R}^n. \tag{1.2}$$

Here $u=u(x,t):\mathbf{R}^n\times[0,T)\to\mathbf{R}$ is an unknown function and $u_t=\partial_t u$, $\nabla u=(\partial_{x_i}u)_{i=1}^n$, $\nabla^2 u=(\partial_{x_ix_j}u)_{i,j=1}^n$ denote its derivatives. The goal of this paper is to derive a lower spatially Lipschitz bound for viscosity solutions u to (1.1)–(1.2). To be more precise, we give a lower bound of

$$\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho]:=\sup\left\{\frac{|u(x,t)-u(0,t)|}{|x|}\;\middle|\; x\in\overline{B}_\rho\setminus\{0\}\right\}$$

when the initial datum ϕ belongs to the Hölder space $C^\theta(\mathbf{R}^n)$ for some $\theta\in(0,1)$ and is not differentiable at the origin; a typical one is $\phi(x)=A|x|^\theta$ $(A>0)$. Here $\overline{B}_\rho$ denotes the closed ball in $\mathbf{R}^n$ with radius ρ centered at the origin. The resulting estimate will be of the form

$$\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho]\ge Ct^{-\gamma_0}, \tag{1.3}$$

where γ_0 depends on the Hölder exponent θ of the initial datum and the growth rates of F with respect to ∇u and $\nabla^2 u$. For example, when F is subquadratic with respect to ∇u and sublinear with respect to $\nabla^2 u$, it turns out that (1.3) holds for $\gamma_0=\frac{1-\theta}{2}$. First order Hamilton–Jacobi equations are also covered by our result.

The structure condition on F we assume is rather mild and weaker than the one needed for a comparison principle between viscosity solutions. Moreover, our lower estimate applies to equations with some singular terms. In fact, our assumptions include the heat convection equation

$$u_t(x,t)-\Delta u(x,t)+\langle B(x,t),\nabla u(x,t)\rangle=0 \quad \text{in } \mathbf{R}^n\times(0,T) \tag{1.4}$$

with a singular velocity field B at the initial time. Here $\Delta u=\sum_{i=1}^n\partial_{x_i}^2 u$ and $\langle\cdot,\cdot\rangle$ denotes the standard inner product in $\mathbf{R}^n$. Such equations appear in problems in fluid mechanics. Especially, (1.4) is the two dimensional vorticity equation for the incompressible fluid when B stands for the velocity field of the fluid; see, e.g., [16, Section 2]. Typical fully nonlinear operators are also included in our formulation. For instance, our results apply to Bellman–Isaacs equations:

$$\begin{aligned} u_t(x,t)+\inf_{b\in\mathbf{B}}\sup_{a\in\mathbf{A}}\big\{ &-\mathrm{trace}(A_{a,b}(x,t)\nabla^2u(x,t))+\langle B_{a,b}(x,t),\nabla u(x,t)\rangle \\ &-f_{a,b}(x,t)\big\}=0 \quad \text{in } \mathbf{R}^n\times(0,T). \end{aligned} \tag{1.5}$$

These equations appear in differential games and stochastic control problems, where the unknown function u represents the value function ([4, 10]).

It is known that the solution u to (1.1)–(1.2) can have spatially Lipschitz regularity even if the initial datum ϕ is singular. Such a property is often called a regularizing effect. For example, the regularizing occurs for uniformly parabolic equations and Hamilton–Jacobi equations with superlinear growth with respect to ∇u. In the literature, when the initial Hölder exponent is θ, an upper bound of $\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho]$ is known to be $Ct^{-\frac{1-\theta}{2}}$ for some classes of parabolic equations as stated below. Then, combining this upper estimate with our lower estimate (1.3), we conclude that the rate $t^{-\frac{1-\theta}{2}}$ is optimal.

Known results. There is vast literature on an upper bound estimate for the gradient (Lipschitz constant) of viscosity solutions to elliptic or parabolic problems. Among them, regularizing effects for parabolic problems are investigated in [18, 19, 7, 23, 8]. Especially, Porretta and Priola [23] study regularizing effects for (1.1) under mild regularity assumptions on the equation and, for both the linear and nonlinear problems, they derive upper spatially gradient bounds depending on the initial Hölder exponent. A simplified version of their theorem [23, Theorem 1.4] is stated as follows. See also [23, Theorem 1.1, Theorem 3.21, Corollary 4.18]. Let $\mathbf{S}^n$ stand for the space of real $n \times n$ symmetric matrices with the usual ordering.

Theorem 1.1 ([23, Theorem 1.4 (ii)]). *Assume that $F = F(x,t,p,X)$ satisfies the following:*

(P1) *There exist $\lambda > 0$, $M_0 \geq 0$ and non-negative functions $\eta(x,y,t)$ and $g \in C((0,1)) \cap L^1((0,1))$ satisfying $g(s)s \to 0$ $(s \to 0)$ and $g(s) = O(s)$ $(s \to \infty)$ such that*

$$\begin{aligned}&F(x,t,\mu(x-y),X) - F(y,t,\mu(x-y),Y)\\&\geq -\lambda \mathrm{trace}(X-Y) - \mu|x-y|g(|x-y|) - M_0 - \nu\eta(t,x,y)\end{aligned}$$

for any $\mu > 0$, $\nu \geq 0$, $x, y \in \mathbf{R}^n$, $t \in (0,T)$ and $X, Y \in \mathbf{S}^n$ with

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

(P2) *There exists $w \in C^{2,1}(\mathbf{R}^n \times [0,T])$ and $\varepsilon_0 > 0$ such that $w(x,t) \to \infty$ $(|x| \to \infty)$ uniformly in $t \in [0,T]$ and*

$$\varepsilon w_t(x,t) + F(x,t,p+\varepsilon\nabla w(x,t), X + \varepsilon\nabla^2 w(x,t)) - F(x,t,p,X) \geq 0$$

for any $(x,t,p,X) \in \mathbf{R}^n \times (0,T) \times \mathbf{R}^n \times \mathbf{S}^n$ and $\varepsilon \in (0,\varepsilon_0]$.

Let $\phi \in C^\theta(\mathbf{R}^n)$ for some $\theta \in (0,1)$. Let u be a viscosity solution of (1.1)–(1.2) such that $u(x,t) = o(w(x,t))$ $(|x| \to \infty)$ uniformly in $t \in [0,T]$. Then, for some $C' > 0$,

$$\|\nabla u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)} \leq C' t^{-\frac{1-\theta}{2}} \quad (t \in (0,T),\ t < 1). \tag{1.6}$$

Actually, the estimates obtained in [23] have more general forms. See also [24, 11] for related results.

On the other hand, it seems that the optimality of the regularizing rate in (1.6) as $t \to 0$ is not generally known, except [14]. To show the optimality, one needs to derive a lower bound of $\|\nabla u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)}$ for some special initial data. In [14] Fujita derives such lower bounds for two types of equations. One is a linear equation of divergence form:

$$u_t(x,t) - \sum_{i,j=1}^{n} \partial_{x_i}\left(a_{ij}(x,t)u_{x_j}(x,t)\right) = 0 \quad \text{in } \mathbf{R}^n \times (0,T), \tag{1.7}$$

where $(a_{ij}(x,t))_{i,j=1}^n$ is uniformly elliptic and satisfies some regularity conditions.

Theorem 1.2 ([14, Theorem 3.3 (iii)]). *Assume that $\phi \in C(\mathbf{R}^n)$ satisfies $A_1|x|^{\theta_1}+B_1 \le \phi(x) \le A_2|x|^{\theta_2} + B_2$ $(x \in \mathbf{R}^n)$ for some $A_i > 0$, $B_i \in \mathbf{R}$ and $\theta_i \in (0,1)$ $(i = 1,2)$ with $\theta_1 \le \theta_2$. Let u be a classical solution of* (1.7)–(1.2)*. Then, for some $C > 0$,*

$$\|\nabla u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)} \ge C\left(t^{-\frac{\theta_2}{2}} + B_2 - B_1\right)^{\frac{1-\theta_1}{\theta_1}} \quad (t \in (0,T)). \tag{1.8}$$

In particular, if $\theta_1 = \theta_2 =: \theta$ and $B_1 = B_2$, then the right-hand side becomes $Ct^{-\frac{1-\theta}{2}}$. Thus this theorem shows that the upper estimate (1.6) is optimal for the equation (1.7).

The other equation treated in [14] is a Hamilton–Jacobi equation:

$$u_t(x,t) + \frac{1}{p}|\nabla u(x,t)|^p = 0 \quad \text{in } \mathbf{R}^n \times (0,T), \tag{1.9}$$

where $p > 1$ is a constant. Let $q > 1$ be the conjugate of p, that is, the constant satisfying $\frac{1}{p} + \frac{1}{q} = 1$, and let $UC(\mathbf{R}^n)$ be the set of uniformly continuous functions $\phi : \mathbf{R}^n \to \mathbf{R}$.

Theorem 1.3 ([14, Theorem 4.3 (iii)]). *Assume that $\phi \in UC(\mathbf{R}^n)$ satisfies $-A_1|x|^{\theta_1} + B_1 \le \phi(x) \le -A_2|x|^{\theta_2}+B_2$ $(x \in \mathbf{R}^n)$ for some $A_i > 0$, $B_i \in \mathbf{R}$ and $\theta_i \in (0,1)$ $(i = 1,2)$ with $\theta_1 \ge \theta_2$. Let u be a viscosity solution of* (1.9)–(1.2)*. Then, for some $C > 0$,*

$$\|\nabla u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)} \ge C\left(t^{-\frac{q-1}{q-\theta_1}\theta_1} + B_2 - B_1\right)^{\frac{1-\theta_2}{\theta_2}} \quad (t \in (0,T)). \tag{1.10}$$

The upper bound estimate is also derived in [14, Proposition 4.1], which asserts that, if $\phi \in C^\theta(\mathbf{R}^n)$ for some $\theta \in (0,1)$, the viscosity solution u of (1.9)–(1.2) satisfies

$$\|\nabla u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)} \le Ct^{-\frac{q-1}{q-\theta}(1-\theta)} \quad (t \in (0,T)). \tag{1.11}$$

When $\theta_1 = \theta_2 =: \theta$ and $B_1 = B_2$ in (1.10), the right-hand side is $t^{-\frac{q-1}{q-\theta}(1-\theta)}$. Therefore, this is the optimal rate.

The proofs of both Theorems 1.2 and 1.3 are based on the following inequality ([14, Theorem 1.1]):

$$\frac{C}{\alpha}e^{\frac{\alpha}{n}\sup\{f(x)|x\in\mathbf{R}^n\}}\left(\int_{\mathbf{R}^n} e^{\alpha f(x)}\,dx\right)^{-\frac{1}{n}} \le \|\nabla f\|_{L^\infty(\mathbf{R}^n)}. \tag{1.12}$$

Here $\alpha > 0$ and $f : \mathbf{R}^n \to \mathbf{R}$ is a locally Lipschitz continuous function such that $e^f \in L^\alpha(\mathbf{R}^n)$. This inequality is derived from the logarithmic Sobolev inequality established in [13]. In the proofs of Theorems 1.2 and 1.3, this estimate is applied for $f(x) = -u(x,t)$ or $f(x) = u(x,t)$, and, to get the desired lower estimate, $\sup\{f(x) \mid x \in \mathbf{R}^n\}$ and $\int_{\mathbf{R}^n} e^{\alpha f(x)}\, dx$ are carefully estimated. In this process, representation formulas of solutions play crucial roles. In fact, the linear equation (1.7) studied in [14] admits a fundamental solution $\Gamma(x,t;y,s)$, and the solution u is given by

$$u(x,t) = \int_{\mathbf{R}^n} \Gamma(x,t;y,0)\phi(y)\, dy.$$

For the Hamilton–Jacobi equation (1.9), the unique viscosity solution is given by the Hopf–Lax formula.

A lower bound for the gradient is also studied in [17] from different points of view. In [17, Theorem 4.2], it is shown that a solution $u(x,t)$ of a Hamilton–Jacobi equation satisfies $|\nabla u| \geq C(t)$ for some $C(t) > 0$ in the sense of viscosity solutions when the Hamiltonian $H(x,t,p)$ is convex in p.

Methods and results in this paper. Similar to (1.8) and (1.10), we will give a lower Lipschitz bound including two different Hölder exponents of an upper function and a lower function for the initial datum. Then, letting the exponents be the same, we will deduce the optimal regularizing rate. Both the results in Theorems 1.2 and 1.3 are covered as special cases of our result. Moreover, we mention that our proof does not rely on representation formulas of solutions, which enables us to handle fully nonlinear equations (1.1).

We first prove key gradient estimates in Section 2 for functions u satisfying some upper and lower estimate. This estimate is derived in a rather direct way, but it turns out that the estimate can apply to a wide class of parabolic equations and generate optimal estimates for the spatially Lipschitz constant. A typical situation for which this key estimate is applicable is that u lies between two modified solutions $w^\pm$ of the heat equation. Namely,

$$w^-(x,t) \leq u(x,t) \leq w^+(x,t). \tag{1.13}$$

The gradient estimate under (1.13) is discussed in Section 3.

As an application to partial differential equations, we first consider linear equations in Section 4. We derive the lower bound estimate (1.3) when the fundamental solution Γ of the equation satisfies the Aronson estimate, which is an upper and lower pointwise estimate for Γ by Gaussian-like functions. This estimate was first proved by Aronson [1] for equations (1.7) of divergence form, and in [2] it was extended to more general equations with lower order terms. See also [3, 9, 22].

The Aronson estimate can be available even for singular equations. For instance, it is known that the heat convection equation (1.4) admits a fundamental solution satisfying the Aronson estimate if the vector field B in the equation fulfills

$$\sup_{t\in(0,T)} \sqrt{t}\|B(\cdot,t)\|_{L^\infty(\mathbf{R}^n)} < \infty \tag{1.14}$$

([6, 21, 20]). This covers the case of the two dimensional vorticity equation.

Fully nonlinear equations are studied in Section 5. In this case, we let $w^{\pm}(x,t) = \pm D_0 t^{\alpha_0} + h_\phi(x, t^{-\beta_0})$ ($D_0 \gg 1$) in (1.13), where $h_\phi(x,t)$ denotes the solution of the heat equation with the initial datum ϕ, and prove that w^- and w^+ are respectively a subsolution and a supersolution to (1.1). We then deduce (1.13) from a comparison principle. The point here is that we change the scale of h_ϕ by $t \mapsto t^{-\beta_0}$. In fact, the estimate (1.3) is derived by optimizing the constants α_0 and β_0 so that (1.3) has the best rate.

Our assumptions allow the nonlinear operator F in (1.1) to be singular at $t = 0$. In particular, in the case of the heat convection equation, the vector field B satisfying (1.14) is covered even in this nonlinear setting.

We would like to mention that our assumption on the initial datum ϕ is localized. In [14], in order to apply (1.12), the datum ϕ is required to grow at infinity. We do not need such a growth condition on ϕ, and thanks to this localization, our lower estimate applies to equations whose solvability is known only for decaying initial functions such as functions in the Lebesgue space. See Example 4.5. Moreover, with a bit of modification of the above argument, we derive a lower spatially Lipschitz bound for solutions to Dirichlet boundary value problems; see Section 6.

Organization. This paper is organized as follows: In Section 2, we derive a key lower bound estimate for the spatially Lipschitz constant. Section 3 is devoted to studying some properties of solutions to the heat equation with a Hölder continuous initial datum. In Sections 4 and 5, we derive a lower spatially Lipschitz bound for solutions to linear and nonlinear equations, respectively. Finally, we consider Dirichlet boundary value problems in Section 6.

2 Key estimates

For an interval $I = [0,a]$ or $I = [0,\infty)$ we define

$$\mathcal{M}(I) = \{m \in C(I) \mid m(0) = 0,\ m \text{ is non-decreasing in } I\}.$$

Let $\rho, T > 0$. We consider a function $u = u(x,t) : \overline{B}_\rho \times (0,T) \to \mathbf{R}$ satisfying the following:

(A1) (i) There exist $\psi \in C(\overline{B}_\rho)$, $m \in \mathcal{M}([0,\rho])$ and $I_1, I_2 \in \mathcal{M}([0,T])$ such that

$$\begin{aligned} \psi(x) - I_1(t) &\le u(x,t) \quad ((x,t) \in \overline{B}_\rho \times (0,T)),\\ u(0,t) &\le I_2(t) \quad (t \in (0,T)). \end{aligned}$$

(ii) $\psi(0) = 0$ and $m(r) \le \max_{|x|=r} \psi(x)$ ($r \in [0,\rho]$).

(iii) $m \in C^2((0,\rho])$ and $m'' < 0$ in $(0,\rho)$.

(iv) $I(t) := I_1(t) + I_2(t) > 0$ $(t \in (0,T))$.

By (A1)-(iii), m is concave in $[0,\rho]$.

We will deduce a lower bound for $\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho]$ (Theorem 2.3) under (A1) with a general $m \in \mathcal{M}([0,\rho])$ as above, though we are mainly interested in the case $m(r) = Ar^\gamma$ $(A>0,\ \gamma \in (0,1))$ for solutions u to (1.1). For such a Hölder type m, see Example 2.4.

Remark 2.1. When (A1)-(iv) does not hold, a lower bound of $\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho]$ is immediately obtained. Assume that u satisfies (A1)-(i)–(iii) and that there is $T_0 \in (0,T)$ such that $I_1(t) = I_2(t) = 0$ for any $t \in [0,T_0]$. For such t, we have $u(0,t)=0$. Now, for $r \in (0,\rho]$, we choose $x_r \in \overline{B}_\rho$ such that $|x_r| = r$ and $\max_{|x|=r}\psi(x) = \psi(x_r)$. Then

$$\frac{u(x_r,t)-u(0,t)}{|x_r|} \geq \frac{\psi(x_r)}{r} \geq \frac{m(r)}{r}.$$

This yields the estimate

$$\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho] \geq \sup_{r\in(0,\rho]} \frac{m(r)}{r} = \lim_{r\to+0}\frac{m(r)}{r}.$$

Lemma 2.2. *Assume that $m \in \mathcal{M}([0,\rho])$ satisfies* (A1)-(iii). *Define $M : [0,\rho] \to \mathbf{R}$ by*

$$M(0) = 0, \quad M(r) = m(r) - m'(r)r \quad (0 < r \leq \rho). \tag{2.1}$$

Then $M \in C([0,\rho]) \cap C^1((0,\rho])$ and M is increasing in $[0,\rho]$.

Proof. By the definition of M, it is clear that $M \in C^1((0,\rho])$. Let us prove the continuity of M at 0. By the concavity of m in $[0,\rho]$, we see that $m'(r) \leq \frac{m(r)}{r}$ $(0 < r \leq \rho)$. This shows that $M \geq 0$ in $[0,\rho]$. We also have $m' \geq 0$ in $(0,\rho]$ since m is non-decreasing. Accordingly,

$$0 \leq M(r) = m(r) - m'(r)r \leq m(r) \to 0 \quad (r \to +0).$$

Thus M is continuous at 0.

Next, observe

$$M'(r) = m'(r) - m''(r)r - m'(r) = -m''(r)r > 0 \quad (0 < r < \rho).$$

This implies that M is increasing in $[0,\rho]$. □

Lemma 2.2 implies that the function $M : [0,\rho] \to [0,M(\rho)]$ is one-to-one. Thus there is the inverse function $M^{-1} : [0,M(\rho)] \to [0,\rho]$.

We now give a lower bound of $\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho]$.

Theorem 2.3 (Lower spatially Lipschitz bound under (A1))**.** *Assume that u satisfies* (A1). *Then, for all $t \in (0,T)$ such that $I(t) \leq M(\rho)$,*

$$\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho] \geq m'(r_t) \quad (r_t := M^{-1}(I(t)) \in (0,\rho]). \tag{2.2}$$

Here M^{-1} is the inverse function of M which is defined by (2.1).

Proof. Let $r \in (0, \rho]$ and select $x_r \in \overline{B}_\rho$ such that $|x_r| = r$ and $\max_{|x|=r} \psi(x) = \psi(x_r)$. For $t \in (0, T)$, we have

$$\frac{u(x_r, t) - u(0, t)}{|x_r|} \geq \frac{\{\psi(x_r) - I_1(t)\} - I_2(t)}{r} \geq \frac{m(r) - I(t)}{r}.$$

Let us maximize the right-hand side $f(r) := \frac{m(r)-I(t)}{r}$ over $(0, \rho]$. The derivative of f is

$$f'(r) = \frac{m'(r)r - m(r) + I(t)}{r^2} = \frac{-M(r) + I(t)}{r^2}.$$

Thus f attains a maximum at $r_t = M^{-1}(I(t))$ when $I(t) \leq M(\rho)$, and the maximum value is

$$f(r_t) = \frac{m(r_t) - I(t)}{r_t} = \frac{m'(r_t)r_t}{r_t} = m'(r_t).$$

Consequently, $\mathrm{Lip}[u(\cdot, t); \overline{B}_\rho] \geq f(r_t) = m'(r_t)$. $\square$

Example 2.4. Let $\rho > 0$ and $m(r) = Ar^\gamma$ $(0 \leq r \leq \rho)$ with $A > 0$ and $\gamma \in (0, 1)$. Then,

$$M(r) = m(r) - m'(r)r = Ar^\gamma - A\gamma r^{\gamma-1} \cdot r = A(1 - \gamma)r^\gamma.$$

Thus, for any $t \in (0, T)$ such that $I(t) \leq M(\rho)$, we have

$$r_t = M^{-1}(I(t)) = \left(\frac{I(t)}{A(1-\gamma)}\right)^{\frac{1}{\gamma}}.$$

Therefore,

$$\mathrm{Lip}[u(\cdot, t); \overline{B}_\rho] \geq m'(r_t) = A\gamma r_t^{\gamma-1} = A^{\frac{1}{\gamma}}\gamma(1-\gamma)^{\frac{1-\gamma}{\gamma}} I(t)^{-\frac{1-\gamma}{\gamma}}.$$

We shall give two more examples of m which are not of Hölder type.

Example 2.5. Let $0 < \rho \leq e^{-1}$ and $m(r) = -r \log r$ $(0 < r \leq \rho)$, $m(0) = 0$. Then $m \in \mathcal{M}([0, \rho])$ and it satisfies (A1)-(iii). Compute

$$M(t) = m(r) - m'(r)r = -r \log r - (-\log r - 1)r = r.$$

This implies that $r_t = M^{-1}(I(t)) = I(t)$ when $t \in (0, T)$ satisfies $I(t) \leq M(\rho) = \rho$. Then,

$$\mathrm{Lip}[u(\cdot, t); \overline{B}_\rho] \geq m'(r_t) = -\log r_t - 1 = -\log I(t) - 1.$$

Example 2.6. Let $0 < \rho \leq e^{-2}$ and $m(r) = -\frac{1}{\log r}$ $(0 < r \leq \rho)$, $m(0) = 0$. One can check that $m \in \mathcal{M}([0, \rho])$ and that (A1)-(iii) holds. We have

$$M(r) = m(r) - m'(r)r = -\frac{1}{\log r} - \frac{1}{r(\log r)^2} \cdot r = -\frac{1}{\log r} - \frac{1}{(\log r)^2}.$$

Let $t \in (0,T)$ satisfy $I(t) \le M(\rho)$. Solving $M(r_t) = I(t)$, we obtain

$$-\frac{1}{\log r_t} = \frac{1-\sqrt{1-4I(t)}}{2} = \frac{2I(t)}{1+\sqrt{1-4I(t)}} \iff r_t = e^{-\frac{1+\sqrt{1-4I(t)}}{2I(t)}}.$$

Hence

$$\begin{aligned}\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho] \ge m'(r_t) &= \frac{1}{r_t(\log r_t)^2} \\ &= e^{\frac{1+\sqrt{1-4I(t)}}{2I(t)}}\left(\frac{2I(t)}{1+\sqrt{1-4I(t)}}\right)^2 \ge e^{\frac{1}{2I(t)}}\{I(t)\}^2.\end{aligned}$$

3 Preliminaries for application to PDEs

We give some results on solutions to the heat equation, especially when the initial datum is Hölder continuous. In addition, we derive a sufficient condition which guarantees that u satisfies (A1). Roughly speaking, we show that (A1) holds if u lies between two solutions of the heat equation. In Sections 4 and 5, we apply this result to parabolic partial differential equations to obtain a lower spatially Lipschitz bound for solutions.

For $\phi : \mathbf{R}^n \to \mathbf{R}$, its modulus of continuity $\omega_\phi : [0,\infty) \to [0,\infty)$ is defined by

$$\omega_\phi(r) := \sup\{|\phi(x)-\phi(y)| \mid x, y \in \mathbf{R}^n,\ |x-y| \le r\} \quad (r \ge 0).$$

By definition $\omega_\phi(0) = 0$ and ω_ϕ is non-decreasing in $[0,\infty)$. Moreover, if $\phi \in UC(\mathbf{R}^n)$, then we see that ω_ϕ is continuous in $[0,\infty)$ and has at most linear growth at infinity.

Let $\phi \in C^\theta(\mathbf{R}^n)$ for some $\theta \in (0,1)$. In what follows, $C_\phi > 0$ denotes a constant such that

$$\omega_\phi(r) \le C_\phi r^\theta \quad (r \ge 0).$$

Later, we derive a lower Lipschitz bound for solutions when the initial datum $\phi \in C^\theta(\mathbf{R}^n)$ satisfies

$$\begin{cases}\phi(0) = 0, \\ c_\phi r^{\theta_0} \le \max\limits_{|x|=r} \phi(x)\ (r \in [0,\rho]) \text{ for some } c_\phi > 0,\ \theta_0 \in [\theta,1) \text{ and } \rho > 0.\end{cases} \tag{3.1}$$

Let

$$g(x) := \frac{1}{(4\pi)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4}} \quad (x \in \mathbf{R}^n)$$

and define the *Gaussian kernel* as

$$G_t(x) := \frac{1}{t^{\frac{n}{2}}} g\left(\frac{x}{\sqrt{t}}\right) = \frac{1}{(4\pi t)^{\frac{n}{2}}} e^{-\frac{|x|^2}{4t}} \quad (x \in \mathbf{R}^n,\ t > 0).$$

It is well-known that $\int_{\mathbf{R}^n} G_t(x)\,dx = 1$ $(t > 0)$. For $\phi \in UC(\mathbf{R}^n)$, we define

$$\begin{cases} h_\phi(x,t) := \displaystyle\int_{\mathbf{R}^n} G_t(x-y)\phi(y)\,dy \quad (x \in \mathbf{R}^n,\ t > 0), \\ h_\phi(x,0) := \phi(x) \quad (x \in \mathbf{R}^n). \end{cases} \tag{3.2}$$

Then, $h_\phi \in C^\infty(\mathbf{R}^n \times (0,\infty)) \cap C(\mathbf{R}^n \times [0,\infty))$, and $h_\phi(x,kt)$ $(k > 0)$ solves

$$\partial_t h_\phi(x,t) - k\Delta h_\phi(x,t) = 0 \quad \text{in } \mathbf{R}^n \times (0,\infty).$$

For $\phi \in UC(\mathbf{R}^n)$ we set

$$J_\phi(t) := \int_{\mathbf{R}^n} G_t(x)\omega_\phi(|x|)\,dx \quad (t > 0), \tag{3.3}$$

where ω_ϕ is the modulus of continuity of ϕ.

Lemma 3.1. *Let $\phi \in C^\theta(\mathbf{R}^n)$ for some $\theta \in (0,1)$. Then, there exists $C > 0$ such that*

$$J_\phi(t) \le C t^{\frac{\theta}{2}} \quad (t > 0).$$

The constant C depends on n, θ and C_ϕ.

Proof. We have

$$J_\phi(t) = \int_{\mathbf{R}^n} G_t(x)\omega_\phi(|x|)\,dx \le \int_{\mathbf{R}^n} \frac{1}{t^{\frac{n}{2}}} g\left(\frac{x}{\sqrt{t}}\right) C_\phi |x|^\theta\,dx,$$

and then, changing the variables by $x = \sqrt{t}z$, we get

$$J_\phi(t) \le C_\phi t^{\frac{\theta}{2}} \int_{\mathbf{R}^n} g(z)|z|^\theta\,dz = C t^{\frac{\theta}{2}}.$$

$\square$

Lemma 3.2. *For $\phi \in UC(\mathbf{R}^n)$,*

$$|h_\phi(x,t) - \phi(x)| \le J_\phi(t) \quad (x \in \mathbf{R}^n,\ t > 0).$$

Proof. Let $(x,t) \in \mathbf{R}^n \times (0,\infty)$. Since $h_\phi(x,t) - \phi(x) = \int_{\mathbf{R}^n} G_t(x-y)(\phi(y) - \phi(x))\,dy$, we have

$$\begin{aligned} |h_\phi(x,t) - \phi(x)| &\le \int_{\mathbf{R}^n} G_t(x-y)|\phi(y) - \phi(x)|\,dy \\ &\le \int_{\mathbf{R}^n} G_t(x-y)\omega_\phi(|x-y|)\,dy = J_\phi(t). \end{aligned}$$ $\square$

Let us consider $u(x,t)$ which lies between two solutions of the heat equation. More precisely, we assume

(A2) There exist $\alpha_i \in \mathcal{M}([0,T])$, $\phi_i \in UC(\mathbf{R}^n)$ and $k_i, \gamma_i > 0$ $(i = 1, 2)$ such that

$$-\alpha_1(t) + h_{\phi_1}(x, k_1 t^{\gamma_1}) \le u(x,t) \le \alpha_2(t) + h_{\phi_2}(x, k_2 t^{\gamma_2}) \qquad ((x,t) \in \overline{B}_\rho \times (0,T)).$$

Theorem 3.3 (Lower spatially Lipschitz bound under (A2))**.** *Assume that u satisfies* (A2)*. Then*

$$-\alpha_1(t) + \phi_1(x) - J_{\phi_1}(k_1 t^{\gamma_1}) \le u(x,t) \le \alpha_2(t) + \phi_2(x) + J_{\phi_2}(k_2 t^{\gamma_2}) \qquad ((x,t) \in \overline{B}_\rho \times (0,T)). \tag{3.4}$$

In addition, if $\phi_1, \phi_2 \in C^\theta(\mathbf{R}^n)$ for some $\theta \in (0,1)$, ϕ_1 satisfies (3.1) *and $\phi_2(0) = 0$, then there exists $C > 0$ such that*

$$\mathrm{Lip}[u(\cdot,t); \overline{B}_\rho] \ge C\left(\alpha_1(t) + \alpha_2(t) + t^{\frac{\theta\gamma_1}{2}} + t^{\frac{\theta\gamma_2}{2}}\right)^{-\frac{1-\theta_0}{\theta_0}} \qquad (t \in (0,T) \text{ sufficiently small}). \tag{3.5}$$

The constant C depends on n, θ, θ_0, c_{ϕ_1}, C_{ϕ_i} and k_i $(i = 1, 2)$.

Proof. (3.4) is an immediate consequence of Lemme 3.2. Let us prove (3.5). By Lemma 3.1 we have $J_{\phi_i}(t) \le C t^{\frac{\theta}{2}}$ $(i = 1, 2)$, and so (3.4) gives

$$\begin{aligned} \phi_1(x) - \alpha_1(t) - C t^{\frac{\theta\gamma_1}{2}} &\le u(x,t) \quad ((x,t) \in \overline{B}_\rho \times (0,T)), \\ u(0,t) &\le \alpha_2(t) + C t^{\frac{\theta\gamma_2}{2}} \quad (t \in (0,T)). \end{aligned}$$

Thus Theorem 2.3 and Example 2.4 yield (3.5). □

We further state a couple of properties of h_ϕ.

Lemma 3.4. *Let $\phi \in C^\theta(\mathbf{R}^n)$ for some $\theta \in (0,1)$. Then $h_\phi(\cdot,t) \in C^\theta(\mathbf{R}^n)$ uniformly in $t \in [0,\infty)$.*

Proof. Let $x_1, x_2 \in \mathbf{R}^n$ and $t > 0$. Then

$$\begin{aligned} |h_\phi(x_1,t) - h_\phi(x_2,t)| &= \left| \int_{\mathbf{R}^n} G_t(y)\phi(x_1 - y)\, dy - \int_{\mathbf{R}^n} G_t(y)\phi(x_2 - y)\, dy \right| \\ &\le \int_{\mathbf{R}^n} G_t(y) |\phi(x_1 - y) - \phi(x_2 - y)|\, dy \\ &\le \int_{\mathbf{R}^n} G_t(y) C_\phi |x_1 - x_2|^\theta\, dy = C_\phi |x_1 - x_2|^\theta. \end{aligned}$$

The constant C_ϕ does not depend on t. □

For $\beta > 0$ let us define

$$h_\phi^\beta(x,t) := h_\phi\left(x, t^\beta\right) \quad (x \in \mathbf{R}^n,\, t \geq 0), \tag{3.6}$$

where h_ϕ is given by (3.2). We prepare the derivative estimates of h_ϕ^β.

Lemma 3.5. *Let $\phi \in C^\theta(\mathbf{R}^n)$ for some $\theta \in (0,1)$. Then, there exists $C > 0$ such that*

$$\begin{aligned} &|\partial_t h_\phi^\beta(x,t)| \leq C\beta t^{-\left(1-\frac{\theta}{2}\beta\right)}, \quad |\partial_{x_i} h_\phi^\beta(x,t)| \leq C t^{-\left(\frac{1}{2}-\frac{\theta}{2}\right)\beta}, \\ &|\partial_{x_i x_j} h_\phi^\beta(x,t)| \leq C t^{-\left(1-\frac{\theta}{2}\right)\beta} \quad (x \in \mathbf{R}^n,\, t > 0). \end{aligned} \tag{3.7}$$

Here all of C depend on n, θ and C_ϕ.

Proof. It suffices to prove (3.7) when $\beta = 1$; the case of a general $\beta > 0$ is derived by changing the variables $t \mapsto t^\beta$. We want to prove

$$\begin{aligned} &|\partial_t h_\phi(x,t)| \leq C t^{-\left(1-\frac{\theta}{2}\right)}, \quad |\partial_{x_i} h_\phi(x,t)| \leq C t^{-\left(\frac{1}{2}-\frac{\theta}{2}\right)}, \\ &|\partial_{x_i x_j} h_\phi(x,t)| \leq C t^{-\left(1-\frac{\theta}{2}\right)} \quad (x \in \mathbf{R}^n,\, t > 0). \end{aligned} \tag{3.8}$$

First, observe

$$\begin{aligned} \partial_{x_i} h_\phi(x,t) &= \int_{\mathbf{R}^n} \partial_{x_i} G_t(x-y)\phi(y)\, dy \\ &= \int_{\mathbf{R}^n} \partial_{x_i} G_t(x-y)(\phi(y)-\phi(x))\, dy + \phi(x)\int_{\mathbf{R}^n} \partial_{x_i} G_t(x-y)\, dy. \end{aligned}$$

Here we have

$$\partial_{x_i} G_t(x-y) = \partial_{x_i}\left(\frac{1}{t^{\frac{n}{2}}} g\left(\frac{x-y}{\sqrt{t}}\right)\right) = \frac{1}{t^{\frac{n}{2}+\frac{1}{2}}} g_{x_i}\left(\frac{x-y}{\sqrt{t}}\right),$$

$$\int_{\mathbf{R}^n} \partial_{x_i} G_t(x-y)\, dy = \partial_{x_i}\int_{\mathbf{R}^n} G_t(x-y)\, dy = \partial_{x_i}(1) = 0.$$

These facts imply

$$\partial_{x_i} h_\phi(x,t) = \frac{1}{t^{\frac{1}{2}}}\int_{\mathbf{R}^n} \frac{1}{t^{\frac{n}{2}}} g_{x_i}\left(\frac{x-y}{\sqrt{t}}\right)(\phi(y)-\phi(x))\, dy, \tag{3.9}$$

and therefore

$$\begin{aligned} |\partial_{x_i} h_\phi(x,t)| &\leq \frac{1}{t^{\frac{1}{2}}}\int_{\mathbf{R}^n} \frac{1}{t^{\frac{n}{2}}}\left|g_{x_i}\left(\frac{x-y}{\sqrt{t}}\right)\right| |\phi(y)-\phi(x)|\, dy \\ &\leq \frac{1}{t^{\frac{1}{2}}}\int_{\mathbf{R}^n} \frac{1}{t^{\frac{n}{2}}}\left|g_{x_i}\left(\frac{x-y}{\sqrt{t}}\right)\right| C_\phi |x-y|^\theta\, dy \\ &= \frac{C_\phi}{t^{\frac{1}{2}}}\int_{\mathbf{R}^n} |g_{x_i}(z)||z|^\theta t^{\frac{\theta}{2}}\, dz = C t^{-\left(\frac{1}{2}-\frac{\theta}{2}\right)}. \end{aligned}$$

The second estimate in (3.8) has been proved. Next, differentiating (3.9), we get

$$\partial_{x_i x_j} h_\phi(x,t) = \frac{1}{t}\int_{\mathbf{R}^n} \frac{1}{t^{\frac{n}{2}}} g_{x_i x_j}\left(\frac{x-y}{\sqrt{t}}\right)(\phi(y)-\phi(x))\,dy.$$

Changing the variables in a similar way, we obtain the third estimate in (3.8). We also deduce the first estimate in (3.8) since $\partial_t h_\phi(x,t) = \Delta h_\phi(x,t)$. □

4 Linear equations

When the equation is linear and its fundamental solution satisfies some conditions, the gradient estimate is deduced from the results in the previous section. Before we discuss general nonlinear equations in Section 5, let us study such a linear case.

Consider a linear equation

$$u_t(x,t) - \operatorname{trace}(A(x,t)\nabla^2 u(x,t)) + \langle B(x,t), \nabla u(x,t)\rangle + c(x,t)u(x,t) = 0 \quad \text{in } \mathbf{R}^n \times (0,T) \tag{4.1}$$

with the initial condition (1.2). Here $A(x,t) = (a_{ij}(x,t))_{i,j=1}^n$ and $B(x,t) = (b_i(x,t))_{i=1}^n$.

Set

$$\mathcal{O} := \{(x,t,y,s) \mid x,y \in \mathbf{R}^n,\ 0 \le s < t < T\}.$$

It is known that (see, e.g, [12, Chapter 1.6]), when A, B and c satisfy some suitable regularity conditions, there exists a Fundamental solution $\Gamma \in C(\mathcal{O})$ of (4.1), and the (weak) solution of the initial value problem (4.1)–(1.2) is given by

$$u(x,t) = \int_{\mathbf{R}^n} \Gamma(x,t;y,0)\phi(y)\,dy \quad ((x,t) \in \mathbf{R}^n \times (0,T)). \tag{4.2}$$

In the following, we study the gradient estimate for the above u.

We say that the *Aronson estimate* holds for Γ when there exist $d_i, e_i > 0$ $(i = 1,2)$ such that

$$e_1 G_{d_1(t-s)}(x-y) \le \Gamma(x,t;y,s) \le e_2 G_{d_2(t-s)}(x-y) \quad ((x,t,y,s) \in \mathcal{O}). \tag{4.3}$$

Theorem 4.1 (Lower spatially Lipschitz bound for linear equations)**.** *Assume that the Aronson estimate holds for $\Gamma \in C(\mathcal{O})$. Let $\phi \in C^\theta(\mathbf{R}^n)$ for some $\theta \in (0,1)$, and assume that ϕ satisfies* (3.1)*. Let u be defined by* (4.2)*. Then there exists $C > 0$ such that*

$$\operatorname{Lip}[u(\cdot,t); \overline{B}_\rho] \ge C t^{-\frac{\theta}{2}\frac{1-\theta_0}{\theta_0}} \quad (t \in (0,T) \text{ sufficiently small}). \tag{4.4}$$

The constant C depends on n, θ, C_ϕ, θ_0, c_ϕ, d_i and e_i $(i = 1,2)$. In particular, if $\theta_0 = \theta$, then

$$\operatorname{Lip}[u(\cdot,t); \overline{B}_\rho] \ge C t^{-\frac{1-\theta}{2}} \quad (t \in (0,T) \text{ sufficiently small}). \tag{4.5}$$

Proof. If $\phi \geq 0$ in $\mathbf{R}^n$, the proof is immediate. In fact, by the Aronson estimate we have

$$\int_{\mathbf{R}^n} e_1 G_{d_1 t}(x-y)\phi(y)\,dy \leq \int_{\mathbf{R}^n} \Gamma(x,t;y,0)\phi(y)\,dy \leq \int_{\mathbf{R}^n} e_2 G_{d_2 t}(x-y)\phi(y)\,dy \tag{4.6}$$

and so

$$h_{e_1\phi}(x, d_1 t) \leq u(x,t) \leq h_{e_2\phi}(x, d_2 t) \quad ((x,t) \in \mathbf{R}^n \times (0,T)).$$

Thus u satisfies (A2) with $\alpha_i \equiv 0$, $\phi_i = e_i\phi$, $k_i = d_i$ and $\gamma_i = 1$ ($i = 1,2$). Theorem 3.3 gives the conclusion.

In the case of a general ϕ, let us express it as $\phi = \phi_+ - \phi_-$, where $\phi_\pm(x) := \max\{\pm\phi(x), 0\}$ denote the positive and negative part of ϕ. Then $\phi_\pm \in C^\theta(\mathbf{R}^n)$. Also, $\phi_\pm$ satisfy (4.6), and subtracting them gives

$$\begin{aligned} h_{e_1\phi_+}(x, d_1 t) - h_{e_2\phi_-}(x, d_2 t) \leq u(x,t) \leq h_{e_2\phi_+}(x, d_2 t) - h_{e_1\phi_-}(x, d_1 t) \\ ((x,t) \in \mathbf{R}^n \times (0,T)). \end{aligned} \tag{4.7}$$

Recall that, by Lemmas 3.1 and 3.2, we have

$$|h_{e_i\phi_\pm}(x, d_i t) - e_i\phi_\pm(x)| \leq J_{e_i\phi_\pm}(d_i t) \leq C t^{\frac{\theta}{2}} \quad ((x,t) \in \mathbf{R}^n \times (0,T),\ i = 1,2).$$

Applying this to (4.7), we find

$$-Ct^{\frac{\theta}{2}} + \psi_1(x) \leq u(x,t) \leq Ct^{\frac{\theta}{2}} + \psi_2(x) \quad ((x,t) \in \mathbf{R}^n \times (0,T)), \tag{4.8}$$

where $\psi_1(x) := e_1\phi_+(x) - e_2\phi_-(x)$ and $\psi_2(x) := e_2\phi_+(x) - e_1\phi_-(x)$. Then ψ_1 satisfies (3.1) with $e_1 c_\phi r^{\theta_0}$ instead of $c_\phi r^{\theta_0}$. Moreover, $\psi_2(0) = 0$. Hence Theorem 2.3 is applicable and we get (4.4) by Example 2.4. $\square$

Remark 4.2. As the proof indicates, it suffices to assume the Aronson estimate (4.3) only for $s = 0$.

Remark 4.3. The numbers θ and θ_0 in Theorem 4.1 respectively correspond to θ_2 and θ_1 in Theorem 1.2, and (4.4) gives the same exponent of t as that in (1.8). Indeed,

$$\frac{\theta}{2} = \frac{\theta_2}{2}, \qquad \frac{1-\theta_0}{\theta_0} = \frac{1-\theta_1}{\theta_1}$$

for $\theta = \theta_2$ and $\theta_0 = \theta_1$. In particular, (4.5) shows that the upper estimate (1.6) is optimal when it is available for (4.1).

Remark 4.4. When Γ further satisfies the gradient estimate, an upper bound for $|\nabla u|$ is easily obtained. To be more precise, we assume that there are $d_3, e_3 > 0$ such that

$$|\partial_{x_i}\Gamma(x,t;y,s)| \leq \frac{e_3}{(t-s)^{\frac{1}{2}}} G_{d_3(t-s)}(x-y) \quad ((x,t,y,s) \in \mathcal{O}). \tag{4.9}$$

See, e.g, [12, Chapter 1.6, (6.13)] for this type of estimates. We also assume that

$$\int_{\mathbf{R}^n} \Gamma(x,t;y,s)\,dy = 1 \quad (x \in \mathbf{R}^n,\ 0 \le s < t < T). \tag{4.10}$$

Let $\phi \in C^\theta(\mathbf{R}^n)$ for some $\theta \in (0,1)$. To derive an upper bound for the gradient of u given in (4.2), we can follow the same calculation as in the proof of Lemma 3.5. In fact, thanks to (4.10) we have

$$u_{x_i}(x,t) = \int_{\mathbf{R}^n} \partial_{x_i}\Gamma(x,t;y,0)(\phi(y)-\phi(x))\,dy.$$

Thus, from (4.9) it follows that

$$\begin{aligned}|u_{x_i}(x,t)| &\le \int_{\mathbf{R}^n} |\partial_{x_i}\Gamma(x,t;y,0)|\,|\phi(y)-\phi(x)|\,dy\\ &\le \frac{C_\phi e_3}{t^{\frac{1}{2}}}\int_{\mathbf{R}^n} G_{d_3 t}(x-y)|x-y|^\theta\,dy.\end{aligned}$$

Changing the variables, we obtain

$$\begin{aligned}|u_{x_i}(x,t)| &\le \frac{C_\phi e_3}{t^{\frac{1}{2}}}\int_{\mathbf{R}^n} \frac{1}{(d_3 t)^{\frac{n}{2}}} g\left(\frac{x-y}{\sqrt{d_3 t}}\right)|x-y|^\theta\,dy\\ &= \frac{C_\phi e_3}{t^{\frac{1}{2}}}\int_{\mathbf{R}^n} g(z)|z|^\theta (d_3 t)^{\frac{\theta}{2}}\,dz = Ct^{-\frac{1-\theta}{2}}.\end{aligned}$$

Example 4.5. The Arronson estimate can be available for the fundamental solution of the heat convection equation (1.4) with a singular velocity field B. In [20] Maekawa proves the lower Gaussian estimate of (4.3) ([20, Theorem 1 (see also (28))]) as well as existence of the fundamental solution for (1.4) ([20, Theorem 2]). There B is assumed to be continuous in $\mathbf{R}^n \times (0,T)$ and satisfy

$$\operatorname{div} B(\cdot,t) = 0 \quad (t \in (0,T)) \quad \text{in the sense of distributions}$$

and (1.14). Under these conditions, it is shown that there exists a unique fundamental solution Γ of (1.4) and that, for any initial datum $\phi \in L^1(\mathbf{R}^n)$, the function u given by (4.2) is the unique mild solution.

The upper Gaussian estimate of (4.3) is obtained in [6, 21]. Therefore, Theorem 4.1 is applicable to (1.4) by choosing $\phi \in C^\theta(\mathbf{R}^n) \cap L^1(\mathbf{R}^n)$.

Examples of (1.4) include the two dimensional vorticity equation, in which $n = 2$ and

$$\begin{aligned}B &= \int_{\mathbf{R}^2} K(x-y)u(y,t)\,dy,\\ K(x) &:= \frac{1}{2\pi|x|}(-x_2,x_1) \quad (x=(x_1,x_2) \in \mathbf{R}^2 \setminus \{0\}).\end{aligned}$$

Indeed, B satisfies (1.14) for a solution u to (1.4); see [21, Page 531], [16, Section 2.4.1].

Remark 4.6. The inhomogeneous problem associated with (4.1) is

$$u_t(x,t) - \operatorname{trace}(A(x,t)\nabla^2 u(x,t)) + \langle B(x,t), \nabla u(x,t)\rangle + c(x,t)u(x,t) = f(x,t) \quad \text{in } \mathbf{R}^n \times (0,T). \tag{4.11}$$

By Duhamel's principle, the solution to (4.11)–(1.2) can be represented as

$$u(x,t) = \int_0^t \int_{\mathbf{R}^n} \Gamma(x,t;y,s) f(y,s)\, dy\, ds + \int_{\mathbf{R}^n} \Gamma(x,t;y,0)\phi(y)\, dy \quad ((x,t) \in \mathbf{R}^n \times (0,T)) \tag{4.12}$$

by using the fundamental solution Γ for (4.1). For such u, the estimates in Theorem 4.1 are still valid if

$$C_f := \sup_{t \in (0,T)} \sqrt{t} \|f(\cdot,t)\|_{L^\infty(\mathbf{R}^n)} < \infty,$$

$\Gamma \geq 0$ and (4.10) holds. Indeed, by these assumptions

$$\begin{aligned} &\left| \int_0^t \int_{\mathbf{R}^n} \Gamma(x,t;y,s) f(y,s)\, dy\, ds \right| \\ &\leq C_f \int_0^t \int_{\mathbf{R}^n} \Gamma(x,t;y,s) \frac{1}{\sqrt{s}}\, dy\, ds = 2C_f\sqrt{t} \quad ((x,t) \in \mathbf{R}^n \times (0,T)). \end{aligned}$$

We thus have, instead of (4.8),

$$-2C_f\sqrt{t} - Ct^{\frac{\theta}{2}} + \psi_1(x) \leq u(x,t) \leq 2C_f\sqrt{t} + Ct^{\frac{\theta}{2}} + \psi_2(x) \ ((x,t) \in \mathbf{R}^n \times (0,T)).$$

Adapting Theorem 3.3, we are again led to (4.4) by applying $t^{\frac{1}{2}} \leq t^{\frac{\theta}{2}}$ for t small.

5 Fully nonlinear equations

We now discuss a lower spatially Lipschitz bound for viscosity solutions to (1.1)–(1.2). We always assume that $F = F(x,t,r,p,X) : \mathbf{R}^n \times (0,T) \times \mathbf{R} \times \mathbf{R}^n \times \mathbf{S}^n \to \mathbf{R}$ is a continuous function. In addition, we impose the following conditions on F. For $X = (x_{ij})_{i,j=1}^n \in \mathbf{S}^n$ let us set $\|X\| := \sqrt{\sum_{i,j=1}^n (x_{ij})^2}$.

(F1) There exist $C_F > 0$, $\mu \geq 0$ and $\sigma \geq 0$ such that

$$|F(x,t,r,p,X)| \leq C_F(t^{-\frac{\sigma}{1+\sigma}} + t^{-\zeta}|p|^\mu + \|X\|^\sigma) \tag{5.1}$$

for all $(x,t,r,p,X) \in \mathbf{R}^n \times (0,T) \times \mathbf{R} \times \mathbf{R}^n \times \mathbf{S}^n$. Here $\zeta \in [0,1)$ is a constant given by

$$\zeta = \zeta(\mu,\sigma) := \begin{cases} 0 & \text{if } \mu \geq 2\sigma, \\ 1 - \dfrac{\mu}{2\sigma} & \text{if } \dfrac{2\sigma}{1+\sigma} \leq \mu < 2\sigma, \\ 1 - \dfrac{1}{1+\sigma} & \text{if } \mu < \dfrac{2\sigma}{1+\sigma}. \end{cases} \tag{5.2}$$

(F2) For all $(x, t, p, X) \in \mathbf{R}^n \times (0, T) \times \mathbf{R}^n \times \mathbf{S}^n$,

$$r \mapsto F(x, t, r, p, X)$$

is non-decreasing in $\mathbf{R}$.

(F3) For any $R > 0$ there exists $\omega_{R,F} \in \mathcal{M}([0, \infty))$ such that

$$|F(x, t, r, p_1, X_1) - F(x, t, r, p_2, X_2)| \leq \omega_{R,F}(|p_1 - p_2| + \|X_1 - X_2\|)$$

for all $(x, t, r, p_i, X_i) \in \mathbf{R}^n \times (0, T) \times \mathbf{R} \times \mathbf{R}^n \times \mathbf{S}^n$ with $|p_i|, \|X_i\| \leq R$ $(i = 1, 2)$.

The constants μ and σ in (F1) will appear in our lower Lipschitz bound for solutions. Due to appearance of $t^{-\frac{\sigma}{1+\sigma}}$ and $t^{-\zeta}$ in (5.1), the zeroth and first order derivative terms of F can have singularities at $t = 0$.

We point out that, if we allow (F1) to depend on the Hölder exponent θ of a given initial datum, then we can improve the rate of t, and furthermore we can allow singularity even for the second order derivative terms. However, we adapt the current assumption (F1) so that the rates of singularities become uniform in θ.

The conditions (F2) and (F3) are used just for a comparison principle between a viscosity solution and a classical solution. They are weaker than the usual structure condition which is required to compare two viscosity solutions.

Example 5.1. When $\mu = \sigma = 1$ in (F1), the condition (5.1) becomes

$$|F(x, t, r, p, X)| \leq C_F(t^{-\frac{1}{2}} + t^{-\frac{1}{2}}|p| + \|X\|). \tag{5.3}$$

If the linear equation (4.1) has bounded coefficients, then (5.3) is satisfied. Moreover, (5.3) holds for the heat convection equation (1.4) with a singular B satisfying (1.14).

Example 5.2. Consider a viscous Hamilton–Jacobi equation

$$u_t(x, t) - \Delta u(x, t) + H(x, t, \nabla u(x, t)) = 0 \quad \text{in } \mathbf{R}^n \times (0, T).$$

If there exists some $\mu > 0$ such that

$$|H(x, t, p)| \leq C_H(t^{-\frac{1}{2}} + t^{-\zeta(\mu,1)}|p|^\mu)$$

for all $(x, t, p) \in \mathbf{R}^n \times (0, T) \times \mathbf{R}^n$, then (F1) holds for this μ and $\sigma = 1$.

One may generalize the second order term $-\Delta u$. For example, typical nonlinear second order operators with $\sigma = 1$ are *Pucci extremal operators* defined by

$$\mathcal{P}^+_{\lambda,\Lambda}(X) := \max\{-\mathrm{trace}(AX) \mid A \in \mathbf{S}^n,\ \lambda I \leq A \leq \Lambda I\},$$
$$\mathcal{P}^-_{\lambda,\Lambda}(X) := \min\{-\mathrm{trace}(AX) \mid A \in \mathbf{S}^n,\ \lambda I \leq A \leq \Lambda I\},$$

where $0 < \lambda \leq \Lambda$.

Remark 5.3. The monotonicity assumption (F2) can be relaxed when we study bounded solutions. Suppose that F satisfies (F1), (F3) and

(F2)$'$ There exists $\kappa \in \mathbf{R}$ such that, for all $(x,t,p,X) \in \mathbf{R}^n \times (0,T) \times \mathbf{R}^n \times \mathbf{S}^n$,

$$r \mapsto \kappa r + F(x,t,r,p,X)$$

is non-decreasing in $\mathbf{R}$.

Set $\tilde{u}(x,t) = e^{-\kappa t} u(x,t)$. Then the equation for $\tilde{u}$ is (1.1) with

$$\tilde{F}(x,t,r,p,X) := \kappa r + e^{-\kappa t} F(x,t,e^{\kappa t} r, e^{\kappa t} p, e^{\kappa t} X).$$

One can check that $\tilde{F}$ satisfies (F2) and (F3), but (F1) may not hold for $\tilde{F}$ globally in $r \in \mathbf{R}$ due to the term κr. However, we easily see that, if solutions are bounded, it suffices to assume (5.1) for r in a bounded interval to have the results below.

Definition 5.4 (Viscosity solution)**.** Let $u \in C(\mathbf{R}^n \times (0,T))$. We say that u is a *viscosity subsolution* (resp. *viscosity supersolution*) of (1.1) if

$$\phi_t(\hat{x},\hat{t}) + F(\hat{x},\hat{t},u(\hat{x},\hat{t}),\nabla\phi(\hat{x},\hat{t}),\nabla^2\phi(\hat{x},\hat{t})) \le 0 \quad (\text{resp. } \ge 0)$$

whenever $u - \phi$ attains a maximum (resp. minimum) at $(\hat{x},\hat{t})$ over $\mathbf{R}^n \times (0,T)$ for $\phi \in C^2(\mathbf{R}^n \times (0,T))$. A *viscosity solution* is a function which is both a viscosity subsolution and a viscosity supersolution.

For the initial(-boundary) value problem, we say that u is a viscosity solution if u takes the prescribed values at the initial time (and on the boundary).

A notion of viscosity solutions is defined for possibly discontinuous functions by taking the semicontinuous envelopes ([5]), and our results in this section holds even for such discontinuous solutions. Here, to simplify presentation and notation, we stick to continuous solutions. Handling discontinuities is not essential work for our gradient estimate.

Definition 5.5. (1) For a Hölder constant $\theta \in (0,1)$, we define

$$\theta^* := 1 + \frac{1}{1-\theta} \in (2,\infty).$$

(2) When we are further given the constants $\mu, \sigma \ge 0$ in (F1), let us define $\alpha_0, \beta_0, \gamma_0 > 0$ by

$$\gamma_0 = \gamma_0(\theta,\mu,\sigma) := \begin{cases} \dfrac{1-\theta}{(1-\theta)\mu + \theta} & \text{if } \mu \ge \theta^*\sigma, \\ \dfrac{1-\theta}{(2-\theta)\sigma + \theta} & \text{if } \mu \le \theta^*\sigma, \end{cases} \tag{5.4}$$

and

$$\alpha_0 = \alpha_0(\theta,\mu,\sigma) := \frac{\theta}{1-\theta}\gamma_0, \quad \beta_0 = \beta_0(\theta,\mu,\sigma) := \frac{2}{1-\theta}\gamma_0. \tag{5.5}$$

Note that we have

$$\alpha_0 = \frac{\theta}{2}\beta_0. \tag{5.6}$$

Example 5.6. Suppose that $\sigma = 1$ in (F1). Then

$$\gamma_0 = \begin{cases} \dfrac{1-\theta}{(1-\theta)\mu + \theta} & \text{if } \mu \geq \theta^*, \\ \dfrac{1-\theta}{2} & \text{if } \mu \leq \theta^*. \end{cases}$$

If furthermore $\mu \leq 2$, we have

$$\gamma_0 = \frac{1-\theta}{2} \quad \text{for any } \theta \in (0,1) \tag{5.7}$$

since $\theta^* > 2$.

Example 5.7. Assume next that $\sigma = 0$ in (F1). This is the case where (1.1) is a first order equation. Then

$$\gamma_0 = \frac{1-\theta}{(1-\theta)\mu + \theta} \quad \text{for any } \mu \geq 0 \text{ and } \theta \in (0,1). \tag{5.8}$$

We prepare some properties of the constants in Definition 5.5.

Lemma 5.8. *Let $\mu, \sigma \geq 0$. Then*

$$\inf_{\theta \in (0,1)} (1 - \alpha_0) = \frac{\sigma}{1+\sigma}. \tag{5.9}$$

Moreover, if $\mu < 2\sigma$, then

$$\inf_{\theta \in (0,1)} \left\{ \left(1 - \frac{\theta}{2}\right) \sigma\beta_0 - \left(\frac{1}{2} - \frac{\theta}{2}\right) \mu\beta_0 \right\} = \zeta. \tag{5.10}$$

Here ζ is the constant given by (5.2).

Proof. (5.9): We have

$$1 - \alpha_0 = \begin{cases} \dfrac{(1-\theta)\mu}{(1-\theta)\mu + \theta} & \text{if } \mu \geq \theta^*\sigma, \\ \dfrac{(2-\theta)\sigma}{(2-\theta)\sigma + \theta} & \text{if } \mu \leq \theta^*\sigma. \end{cases}$$

By direct computation we see that both the functions on the right-hand side are non-increasing with respect to $\theta \in (0,1)$. If $\sigma > 0$, we have $\mu \leq \theta^*\sigma$ when θ is sufficiently close to 1. Therefore

$$\inf_{\theta \in (0,1)} (1 - \alpha_0) = \frac{(2-1)\sigma}{(2-1)\sigma + 1} = \frac{\sigma}{\sigma + 1}.$$

If $\sigma = 0$, then $\mu \geq \theta^*\sigma$ for any $\theta \in (0,1)$. Thus

$$\inf_{\theta\in(0,1)}(1-\alpha_0) = \frac{(1-1)\mu}{(1-1)\mu+1} = 0.$$

(5.10): Since $\mu < 2\sigma$, we have $\mu < \theta^*\sigma$ ($\theta \in (0,1)$), and so $\beta_0 = \frac{2}{(2-\theta)\sigma+\theta}$ ($\theta \in (0,1)$). Now, taking the derivative of

$$A(\theta) := \left(1-\frac{\theta}{2}\right)\sigma\beta_0 - \left(\frac{1}{2}-\frac{\theta}{2}\right)\mu\beta_0 = \frac{(2-\theta)\sigma-(1-\theta)\mu}{(2-\theta)\sigma+\theta},$$

we see that $A(\cdot)$ is non-decreasing if $\mu \geq \frac{2\sigma}{1+\sigma}$ while $A(\cdot)$ is non-increasing if $\mu < \frac{2\sigma}{1+\sigma}$. This implies that

$$\inf_{\theta\in(0,1)} A(\theta) = \begin{cases} \dfrac{(2-0)\sigma-(1-0)\mu}{(2-0)\sigma+0} = 1-\dfrac{\mu}{2\sigma} & \text{if } \mu \geq \dfrac{2\sigma}{1+\sigma}, \\ \dfrac{(2-1)\sigma-(1-1)\mu}{(2-1)\sigma+1} = 1-\dfrac{1}{\sigma+1} & \text{if } \mu < \dfrac{2\sigma}{1+\sigma}. \end{cases}$$

□

Lemma 5.9. *For $\theta \in (0,1)$ and $\mu, \sigma \geq 0$ we define*

$$\Delta(\beta) := \max\left\{1-\frac{\theta}{2}\beta,\ \zeta+\left(\frac{1}{2}-\frac{\theta}{2}\right)\mu\beta,\ \left(1-\frac{\theta}{2}\right)\sigma\beta\right\} \quad (\beta > 0).$$

Here ζ is the constant given by (5.2). *Then*

$$\inf_{\beta>0}\Delta(\beta) = \Delta(\beta_0) = 1-\alpha_0. \tag{5.11}$$

Proof. We first discuss the case $\mu \geq 2\sigma$, so that $\zeta = 0$. Note that

$$\left(\frac{1}{2}-\frac{\theta}{2}\right)\mu\beta \geq \left(1-\frac{\theta}{2}\right)\sigma\beta \iff \mu \geq \theta^*\sigma,$$
$$\left(\frac{1}{2}-\frac{\theta}{2}\right)\mu\beta \leq \left(1-\frac{\theta}{2}\right)\sigma\beta \iff \mu \leq \theta^*\sigma.$$

Let us assume that $\mu \geq \theta^*\sigma$. Then $\Delta(\beta) = \max\left\{1-\frac{\theta}{2}\beta,\ \left(\frac{1}{2}-\frac{\theta}{2}\right)\mu\beta\right\}$, and the minimum of $\Delta(\cdot)$ over $(0,\infty)$ is attained at β satisfying

$$1-\frac{\theta}{2}\beta = \left(\frac{1}{2}-\frac{\theta}{2}\right)\mu\beta \iff \beta = \frac{2}{(1-\theta)\mu+\theta} = \beta_0.$$

The minimum value is

$$\Delta(\beta_0) = 1-\frac{\theta}{2}\beta_0 = 1-\alpha_0.$$

Thus (5.11) holds. In a similar way, one can prove (5.11) when $\mu \leq \theta^*\sigma$.

Next, let us consider the case $\mu < 2\sigma$. Then $\mu < \theta^*\sigma$, and so $\left(\frac{1}{2}-\frac{\theta}{2}\right)\mu\beta < \left(1-\frac{\theta}{2}\right)\sigma\beta$ ($\beta > 0$). This implies that, to prove (5.11), it suffices to show that $\zeta + \left(\frac{1}{2}-\frac{\theta}{2}\right)\mu\beta_0 \leq \left(1-\frac{\theta}{2}\right)\sigma\beta_0$. We have already proved this inequality in (5.10). □

The following sub- and supersolution play an important role to derive a lower Lipschitz bound for solutions.

Theorem 5.10 (Sub-/supersolution to (1.1))**.** *Assume* (F1)*. Let $\phi \in C^{\theta}(\mathbf{R}^n)$ for some $\theta \in (0,1)$. For $D_0 > 0$ we define*

$$w^{\pm}(x,t) := \pm D_0 t^{\alpha_0} + h_{\phi}^{\beta_0}(x,t) \quad ((x,t) \in \mathbf{R}^n \times [0,T)), \tag{5.12}$$

where $h_{\phi}^{\beta_0}$ is given by (3.6)*. If D_0 is sufficiently large, then w^- is a classical subsolution of* (1.1) *and w^+ is a classical supersolution of* (1.1)*. The choice of D_0 depends on n, θ, C_{ϕ}, T, C_F, μ and σ.*

Proof. We only give the proof for w^- since that for w^+ is parallel. To simplify notation we write $h = h_{\phi}^{\beta_0}$. Let $(x,t) \in \mathbf{R}^n \times (0,T)$. Using (F1), (3.7), (5.9) and (5.11), we see that

$$\begin{aligned}
&h_t(x,t) + F(x,t,h(x,t),\nabla h(x,t),\nabla^2 h(x,t)) \\
&\le h_t(x,t) + C_F\left(t^{-\frac{\sigma}{1+\sigma}} + t^{-\zeta}|\nabla h(x,t)|^{\mu} + \|\nabla^2 h(x,t)\|^{\sigma}\right) \\
&\le Ct^{-\left(1-\frac{\theta}{2}\beta_0\right)} + C_F\left(t^{-\frac{\sigma}{1+\sigma}} + Ct^{-\zeta-\left(\frac{1}{2}-\frac{\theta}{2}\right)\beta_0\mu} + Ct^{-\left(1-\frac{\theta}{2}\right)\beta_0\sigma}\right) \\
&\le Ct^{-\Delta(\beta_0)} = Ct^{-(1-\alpha_0)}.
\end{aligned}$$

Therefore

$$\begin{aligned}
&w^-_t(x,t) + F(x,t,w^-(x,t),\nabla w^-(x,t),\nabla^2 w^-(x,t)) \\
&\le -D_0\alpha_0 t^{-(1-\alpha_0)} + Ct^{-(1-\alpha_0)}.
\end{aligned}$$

This shows that w^- is a classical subsolution of (1.1) when D_0 is large enough. □

Let $\phi \in C^{\theta}(\mathbf{R}^n)$ for some $\theta \in (0,1)$. In view of Lemma 3.4 we notice that the functions $w^{\pm}$ in (5.12) satisfy

$$|w^{\pm}(x,t)| \le C(1+|x|^{\theta}) \quad (x \in \mathbf{R}^n) \quad \text{for some } C > 0. \tag{5.13}$$

Moreover, by Lemmas 3.1 and 3.2 we have

$$\begin{aligned}
|w^{\pm}(x,t) - \phi(x)| &\le D_0 t^{\alpha_0} + |h_{\phi}(x,t^{\beta_0}) - \phi(x)| \le D_0 t^{\alpha_0} + J_{\phi}(t^{\beta_0}) \\
&\le D_0 t^{\alpha_0} + Ct^{\frac{\theta}{2}\beta_0} = Ct^{\frac{\theta}{2}\beta_0} \quad (x \in \mathbf{R}^n,\ t > 0).
\end{aligned} \tag{5.14}$$

For the last equality we used (5.6).

We next compare $w^{\pm}$ with a viscosity solution u of (1.1)–(1.2). To do this, we assume that the solution u satisfies

$$\lim_{|x|\to\infty} \sup_{t\in[0,T)} \frac{|u(x,t)|}{1+|x|} = 0 \tag{5.15}$$

and

$$|u(x,t)-\phi(x)| \le q(t) \quad ((x,t) \in \mathbf{R}^n \times [0,T)) \text{ for some } q \in \mathcal{M}([0,T]). \tag{5.16}$$

These are not restrictive conditions in the sense that there always exists a viscosity solution u of (1.1)–(1.2) satisfying (5.15) and (5.16). In fact, Perron's method ([5, Section 4]) guarantees that

$$u_P(x,t) := \sup \left\{ v(x,t) \;\middle|\; \begin{array}{c} v : \mathbf{R}^n \times [0,T) \to \mathbf{R} \text{ is a viscosity subsolution} \\ \text{of (1.1) such that} \\ w^- \le v \le w^+ \text{ in } \mathbf{R}^n \times [0,T) \end{array} \right\} \tag{5.17}$$

is a (possibly discontinuous) viscosity solution of (1.1)–(1.2). By definition u_P lies between w^- and w^+, and therefore it satisfies (5.15) and (5.16) by (5.13) and (5.14).

The comparison result is stated as follows. Since $w^\pm$ are smooth functions, the proof does not need the so-called doubling variables technique in the viscosity solutions theory.

Proposition 5.11 (Comparison principle)**.** *Assume* (F2) *and* (F3)*. Let* $\phi \in C^\theta(\mathbf{R}^n)$ *for some* $\theta \in (0,1)$*. Let* u *be a viscosity solution of* (1.1)–(1.2) *satisfying* (5.15) *and* (5.16)*. Let* $w^\pm$ *be the functions defined by* (5.12)*. Then*

$$w^-(x,t) \le u(x,t) \le w^+(x,t) \quad ((x,t) \in \mathbf{R}^n \times [0,T)). \tag{5.18}$$

Proof. Let us prove that $u \le w^+$ in $\mathbf{R}^n \times [0,T)$. Suppose by contradiction that $\sup_{\mathbf{R}^n \times [0,T)}(u - w^+) \ge M$ for some $M > 0$. For $\varepsilon, \eta \in (0,1]$ let us define

$$\tilde{w}^+(x,t) := w^+(x,t) + \varepsilon \langle x \rangle + \frac{\eta}{T-t} \quad ((x,t) \in \mathbf{R}^n \times [0,T)), \tag{5.19}$$

where $\langle x \rangle = \sqrt{1+|x|^2}$. Since both u and $-w^+$ grow slower than a linear function by (5.15) and (5.13), we see that $u - \tilde{w}^+$ attains a maximum over $\mathbf{R}^n \times [0,T)$ at some $(\hat{x},\hat{t}) = (\hat{x}_{\varepsilon,\eta}, \hat{t}_{\varepsilon,\eta})$. When ε and η are small enough, we have

$$(u - \tilde{w}^+)(\hat{x},\hat{t}) \ge \frac{M}{2}; \tag{5.20}$$

in particular,

$$u(\hat{x},\hat{t}) > \tilde{w}^+(\hat{x},\hat{t}) > w^+(\hat{x},\hat{t}). \tag{5.21}$$

We also remark that $\hat{t}$ are away from 0 uniformly in ε and η, that is, there is $t_0 \in (0,T)$ independent of ε and η such that $\hat{t} \ge t_0$. Indeed, by (5.16) and (5.14),

$$u(x,t) - \tilde{w}^+(x,t) < u(x,t) - w^+(x,t) \le q(t) + Ct^{\frac{\theta}{2}\beta_0} \quad ((x,t) \in \mathbf{R}^n \times [0,T)).$$

This and (5.20) show the existence of such t_0.

By the definition of viscosity subsolutions, we have

$$0 \geq \tilde{w}_t^+(\hat{x}, \hat{t}) + F(\hat{x}, \hat{t}, u(\hat{x}, \hat{t}), \nabla \tilde{w}^+(\hat{x}, \hat{t}), \nabla^2 \tilde{w}^+(\hat{x}, \hat{t})).$$

Next, we apply the monotonicity (F2). By (5.21) we observe

$$\begin{aligned} 0 &\geq \tilde{w}_t^+(\hat{x}, \hat{t}) + F(\hat{x}, \hat{t}, w^+(\hat{x}, \hat{t}), \nabla \tilde{w}^+(\hat{x}, \hat{t}), \nabla^2 \tilde{w}^+(\hat{x}, \hat{t})) \\ &\geq w_t^+(\hat{x}, \hat{t}) + \frac{\eta}{T^2} \\ &\quad + F(\hat{x}, \hat{t}, w^+(\hat{x}, \hat{t}), \nabla w^+(\hat{x}, \hat{t}) + \varepsilon \nabla \langle \hat{x} \rangle, \nabla^2 w^+(\hat{x}, \hat{t}) + \varepsilon \nabla^2 \langle \hat{x} \rangle). \end{aligned}$$

We now choose $R > 0$ independent of ε and η such that $|\nabla w^+|, \|\nabla^2 w^+\| \leq R$ in $\mathbf{R}^n \times [t_0, T)$, which is possible by Lemma 3.5, and $|\nabla \langle \cdot \rangle|, \|\nabla^2 \langle \cdot \rangle\| \leq R$ in $\mathbf{R}^n$. Then, by (F3)

$$\begin{aligned} 0 &\geq w_t^+(\hat{x}, \hat{t}) + \frac{\eta}{T^2} \\ &\quad + F(\hat{x}, \hat{t}, w^+(\hat{x}, \hat{t}), \nabla w^+(\hat{x}, \hat{t}), \nabla^2 w^+(\hat{x}, \hat{t})) - \omega_{2R,F}(\varepsilon |\nabla \langle \hat{x} \rangle| + \varepsilon \|\nabla^2 \langle \hat{x} \rangle\|) \\ &\geq \frac{\eta}{T^2} - \omega_{2R,F}(2\varepsilon R). \end{aligned}$$

Here we have used the fact that w^+ is a supersolution. Sending $\varepsilon \to 0$ in the above inequality gives $0 \geq \frac{\eta}{T^2}$, which is a contradiction.

In a similar way, one can prove that $w^- \leq u$ in $\mathbf{R}^n \times [0, T)$. $\square$

We are now in a position to give a lower bound for $\mathrm{Lip}[u(\cdot, t); \overline{B}_\rho]$.

Theorem 5.12 (Lower spatially Lipschitz bound for nonlinear equations)**.** *Assume* (F1), (F2) *and* (F3). *Let $\phi \in C^\theta(\mathbf{R}^n)$ for some $\theta \in (0,1)$, and assume that ϕ satisfies* (3.1). *Let u be a viscosity solution of* (1.1)–(1.2) *satisfying* (5.15) *and* (5.16). *Then*

$$\mathrm{Lip}[u(\cdot, t); \overline{B}_\rho] \geq C t^{-\frac{1-\theta_0}{\theta_0}\alpha_0} \quad (t \in (0, T) \textit{ sufficiently small}). \tag{5.22}$$

Here C depends on θ_0, c_ϕ, ρ and D_0. In particular, if $\theta_0 = \theta$, then

$$\mathrm{Lip}[u(\cdot, t); \overline{B}_\rho] \geq C t^{-\gamma_0} \quad (t \in (0, T) \textit{ sufficiently small}). \tag{5.23}$$

Proof. By (5.18), we see that u fulfills (A2) with $\alpha_i(t) = D_0 t^{\alpha_0}$, $\phi_i = \phi$, $k_i = 1$ and $\gamma_i = \beta_0$ $(i = 1, 2)$. Hence Theorem 3.3 implies

$$\mathrm{Lip}[u(\cdot, t); \overline{B}_\rho] \geq C \left(2D_0 t^{\alpha_0} + 2t^{\frac{\theta}{2}\beta_0} \right)^{-\frac{1-\theta_0}{\theta_0}} = C t^{-\frac{1-\theta_0}{\theta_0}\alpha_0} \quad (t \in (0, T)),$$

where (5.6) is applied for the last equality. Thus (5.22) is proved. If $\theta_0 = \theta$, we have (5.23) since $\frac{1-\theta}{\theta}\alpha_0 = \gamma_0$. $\square$

Remark 5.13. In Theorem 5.12 one may directly assume (5.18) instead of (F2)–(F3), (5.15) and (5.16). (Actually, in this case u does not have to be a viscosity solution of (1.1)–(1.2).) For example, the Perron solution u_P in (5.17) satisfies (5.18), and therefore the estimates (5.22) and (5.23) hold.

This version of the theorem would be useful if one studies less regular F than (F3) such as a singular F appearing in surface evolution problems ([15]). One of typical surface evolution equations is the level-set mean curvature flow equation with a driving force, for which F is given by

$$F(p, X) = -\mathrm{trace}\left(\left(I - \frac{p \otimes p}{|p|^2}\right) X\right) - \nu|p| \quad ((p, X) \in (\mathbf{R}^n \setminus \{0\}) \times \mathbf{S}^n).$$

Here $\nu \in \mathbf{R}$ is a constant and $p \otimes p = (p_i p_j)_{i,j=1}^n \in \mathbf{S}^n$ for $p = (p_i)_{i=1}^n \in \mathbf{R}^n$. This F satisfies

$$|F(p, X)| \leq C_F(|p| + \|X\|) \quad ((p, X) \in (\mathbf{R}^n \setminus \{0\}) \times \mathbf{S}^n)$$

for some $C_F > 0$, and so one can construct a subsolution w^- and a supersolution w^+ of (1.1)–(1.2) in the same manner. Thus the estimates (5.22) and (5.23) hold for the Perron solution u_P.

Remark 5.14. The resulting estimate (5.23) is understood as follows: Let us fix θ as a Hölder exponent of the initial datum. Recall that, by definition, the constant γ_0 depends on μ and σ, which represent the growth rates of the first order term and the second order term of F, respectively. Moreover, it only depends on the dominant one. Namely, γ_0 is independent of σ if $\mu \geq \theta^*\sigma$, while it is independent of μ if $\mu \leq \theta^*\sigma$. The constant θ^* determines which one is dominant.

Assume now that $\mu \geq \theta^*\sigma$. Then, the larger μ is, the smaller γ_0 is; hence $t^{-\gamma_0}$ gets to be small near the initial time. This means that the gradient can decay fast, that is, the regularizing effect can be strong if it occurs. A similar observation can be made for the case $\mu \leq \theta^*\sigma$.

Example 5.15. Let us check that the exponents of t in Theorem 5.12 agree with those in Theorems 1.2 and 1.3.

(1) We revisit the linear equation (4.1) and compare (5.22) with (1.8). Recall that, if (4.1) has bounded coefficients, (F1) is satisfied for $\mu = 1$ and $\sigma = 1$. Then, by (5.7), we have $\gamma_0 = \frac{1-\theta}{2}$ and $\alpha_0 = \frac{\theta}{1-\theta}\gamma_0 = \frac{\theta}{2}$. Therefore

$$\alpha_0 = \frac{\theta_2}{2}, \quad \frac{1-\theta_0}{\theta_0} = \frac{1-\theta_1}{\theta_1}$$

for $\theta = \theta_2$ and $\theta_0 = \theta_1$. Letting $\theta = \theta_0$ shows that the upper bound in (1.6) is optimal when it is available.

(2) For the Hamilton–Jacobi equation (1.9), (F1) is satisfied when $\mu = p\,(> 1)$ and $\sigma = 0$, and then (5.8) holds. Let us compare (5.22) and (1.10). We check that

$$\alpha_0 = \frac{q-1}{q-\theta_1}\theta_1, \quad \frac{1-\theta_0}{\theta_0} = \frac{1-\theta_2}{\theta_2}$$

for $\mu = p$, $\theta = \theta_1$ and $\theta_0 = \theta_2$. The second one is obvious. The first one is immediate too. Indeed,

$$\alpha_0 = \frac{\theta}{1-\theta}\gamma_0 = \frac{\theta}{(1-\theta)p + \theta} = \frac{\theta}{(1-\theta)\frac{q}{q-1} + \theta} = \frac{q-1}{q-\theta}\theta.$$

The optimality of (1.11) also follows by letting $\theta = \theta_0$.

When the equation is subquadratic in ∇u and sublinear in $\nabla^2 u$, i.e., $\mu \le 2$ and $\sigma = 1$, we have $-\gamma_0 = -\frac{1-\theta}{2}$ ($\theta \in (0,1)$) by (5.7), which is the same exponent as in (1.6). Therefore, combing Theorems 5.12 and 1.1, we obtain the following theorem. Note that (F2) is fulfilled when F is independent of r.

Theorem 5.16 (Optimality of the spatially Lipschitz bound)**.** *Assume that $F = F(x,t,p,X)$ satisfies* (P1), (P2) *in Theorem 1.1. Furthermore, assume* (F1) *with $\mu \le 2$ and $\sigma = 1$, and* (F3). *Let $\phi \in C^{\theta}(\mathbf{R}^n)$ for some $\theta \in (0,1)$, and assume that ϕ satisfies* (3.1) *with $\theta_0 = \theta$. Let u be a viscosity solution of* (1.1)–(1.2) *satisfying* (5.15) *and* (5.16). *Then*

$$Ct^{-\frac{1-\theta}{2}} \le \|\nabla u(\cdot,t)\|_{L^\infty(\mathbf{R}^n)} \le C't^{-\frac{1-\theta}{2}} \quad (t \in (0,T) \text{ sufficiently small}). \tag{5.24}$$

Here C and C' are the constants appearing in Theorem 5.12 and Theorem 1.1, respectively.

Example 5.17. Let us consider the Bellman–Isaacs equation (1.5), where

$$F(x,t,p,X) = \inf_{b\in\mathbf{B}} \sup_{a\in\mathbf{A}} \{-\operatorname{trace}(A_{a,b}(x,t)X) + \langle B_{a,b}(x,t), p\rangle - f_{a,b}(x,t)\}.$$

Here $\mathbf{A}$ and $\mathbf{B}$ are index sets, and $A_{a,b} : \mathbf{R}^n \times (0,T) \to \mathbf{S}^n$, $B_{a,b} : \mathbf{R}^n \times (0,T) \to \mathbf{R}^n$, $f_{a,b} : \mathbf{R}^n \times (0,T) \to \mathbf{R}$ are continuous for any $a \in \mathbf{A}$ and $b \in \mathbf{B}$. To simplify the situation, we now suppose that all of these functions are bounded uniformly in $(a,b) \in \mathbf{A} \times \mathbf{B}$. Namely,

$$\sup_{(a,b)\in\mathbf{A}\times\mathbf{B}} \{\|A_{a,b}\|_{L^\infty(\mathbf{R}^n\times(0,T))} + \|B_{a,b}\|_{L^\infty(\mathbf{R}^n\times(0,T))} + \|f_{a,b}\|_{L^\infty(\mathbf{R}^n\times(0,T))}\} < \infty.$$

Then F is finite in $\mathbf{R}^n \times (0,T) \times \mathbf{R}^n \times \mathbf{S}^n$ and satisfies (F1) with $\mu = 1$ and $\sigma = 1$. Also, (F2) and (F3) hold.

Sufficient conditions for (P1) and (P2) are discussed in [23, Section 5.1]. They are satisfied if

(P1)′ There exist $\lambda > 0$, $\sigma_{a,b} : \mathbf{R}^n \times (0,T) \to \mathbf{S}^n$ and a non-negative function $g \in C((0,1)) \cap L^1((0,1))$ satisfying $g(s)s \to 0$ ($s \to 0$) and $g(s) = O(s)$ ($s \to \infty$) such that

- $A_{a,b} = \lambda I + \sigma_{a,b}^2$ in $\mathbf{R}^n \times (0,T)$.
- $\sigma_{a,b}$ and $B_{a,b}$ are continuous in $\mathbf{R}^n \times (0,T)$ uniformly in $a \in \mathbf{A}$ and $b \in \mathbf{B}$.
-
$$\|\sigma_{a,b}(x,t) - \sigma_{a,b}(y,t)\|^2 - \langle B_{a,b}(x,t) - B_{a,b}(y,t), x-y\rangle \leq |x-y|g(|x-y|)$$
for any $x, y \in \mathbf{R}^n$, $t \in (0,T)$, $a \in \mathbf{A}$ and $b \in \mathbf{B}$.

(P2)′ There exists $w \in C^{2,1}(\mathbf{R}^n \times [0,T])$ such that $w(x,t) \to \infty$ ($|x| \to \infty$) uniformly in $t \in [0,T]$ and

$$w_t(x,t) - \mathrm{trace}(A_{a,b}(x,t)\nabla^2 w(x,t)) + \langle B_{a,b}(x,t), \nabla w(x,t)\rangle \geq 0$$

for any $(x,t) \in \mathbf{R}^n \times (0,T)$, $a \in \mathbf{A}$ and $b \in \mathbf{B}$.

A further sufficient condition for (P2)′ is given in [23, (5.2)], which asserts that, if there exists some $C > 0$ such that

$$\mathrm{trace} A_{a,b}(x,t) - \langle B_{a,b}(x,t), \nabla x\rangle \leq C(1+|x|^2)$$

for any $(x,t) \in \mathbf{R}^n \times (0,T)$, $a \in \mathbf{A}$ and $b \in \mathbf{B}$, then (P2)′ holds for $w(x,t) = e^{Mt}(1+|x|^2)$ ($M \gg 1$).

6 Dirichlet boundary value problems

Our method also applies to solutions of Dirichlet boundary value problems. In this section, we briefly present how to modify the argument in the previous section for Dirichlet problems.

Let $\Omega \subset \mathbf{R}^n$ be a nonempty, bounded open subset such that $0 \in \Omega$. We study

$$u_t(x,t) + F(x,t,u(x,t),\nabla u(x,t),\nabla^2 u(x,t)) = 0 \quad \text{in } \Omega \times (0,T) \tag{6.1}$$

with the Dirichlet boundary condition

$$u(x,t) = d(x,t) \quad \text{on } \partial\Omega \times [0,T) \tag{6.2}$$

and the initial condition

$$u(x,0) = \phi(x) \quad \text{in } \overline{\Omega}. \tag{6.3}$$

Here we assume

- $F \in C(\Omega \times (0,T) \times \mathbf{R} \times \mathbf{R}^n \times \mathbf{S}^n)$, and F satisfies (F1) and (F2), where the domains $\mathbf{R}^n$ of x are replaced by Ω.
- $d \in C(\partial\Omega \times [0,T))$ and

$$|d(x,t) - d(x,0)| \leq L_d t \quad ((x,t) \in \partial\Omega \times (0,T)) \quad \text{for some } L_d > 0. \tag{6.4}$$

- $\phi \in C^\theta(\overline{\Omega})$ for some $\theta \in (0,1)$. We choose $C_\phi > 0$ such that $|\phi(x) - \phi(y)| \le C_\phi |x-y|^\theta$ $(x, y \in \overline{\Omega})$.
- The compatibility condition holds:

$$d(x,0) = \phi(x) \quad (x \in \partial\Omega). \tag{6.5}$$

To derive the gradient estimate, we first extend ϕ to $\mathbf{R}^n$. Let us define

$$\Phi(x) := \max_{y \in \overline{\Omega}} \left\{ \phi(y) - C_\phi |x-y|^\theta \right\} \quad (x \in \mathbf{R}^n).$$

Then $\Phi \in C^\theta(\mathbf{R}^n)$, $\Phi = \phi$ in $\overline{\Omega}$ and $\omega_\Phi(r) \le C_\phi r^\theta$ $(r \ge 0)$. We define $w^\pm$ as in (5.12) with Φ instead of ϕ.

We now claim that

$$w^-(x,t) \le d(x,t) \le w^+(x,t) \quad ((x,t) \in \partial\Omega \times [0,T)) \tag{6.6}$$

holds by choosing D_0 large if necessarily. To see this, we note that

$$\begin{aligned} &w^+(x,t) = D_0 t^{\alpha_0} + h_\Phi^{\beta_0}(x,t) \ge D_0 t^{\alpha_0} + \phi(x) - J_\Phi\left(t^{\beta_0}\right), \\ &J_\Phi\left(t^{\beta_0}\right) \le C t^{\frac{\theta}{2}\beta_0} = C t^{\alpha_0}. \end{aligned}$$

Using these inequalities, (6.4) and (6.5), we have, for $(x,t) \in \partial\Omega \times [0,T)$

$$\begin{aligned} d(x,t) - w^+(x,t) &\le d(x,0) + L_d t - (D_0 t^{\alpha_0} + \phi(x) - C t^{\alpha_0}) \\ &\le t^{\alpha_0}\left(L_d T^{1-\alpha_0} - D_0 + C\right). \end{aligned}$$

Thus the last term is nonpositive for large D_0. Similarly, we deduce $w^- \le d$ on $\partial\Omega \times [0,T)$.

Applying the comparison principle, we see that any viscosity solution u of (6.1)–(6.3) satisfies

$$w^-(x,t) \le u(x,t) \le w^+(x,t) \quad ((x,t) \in \overline{\Omega} \times [0,T)). \tag{6.7}$$

The proof of the comparison principle is almost the same as that of Proposition 5.11. For the Dirichlet problem (6.1)–(6.3), we can remove the term $\varepsilon\langle x\rangle$ from the definition (5.19) of $\tilde{w}^+$ since Ω is bounded. By (6.6), a maximizer $(\hat{x}, \hat{t})$ of $u - \tilde{w}^+$ lies in $\Omega \times (0,T)$, and therefore the definition of viscosity solutions can be applied at $(\hat{x}, \hat{t})$. We also note that, since the term $\varepsilon\langle x\rangle$ does not appear, the assumptions (F3) and (5.16) are not needed in this case.

Since (6.7) holds, the argument to deduce a lower bound for $\mathrm{Lip}[u(\cdot,t); \overline{B}_\rho]$ runs as before. Consequently, we have

Theorem 6.1 (Lower spatially Lipschitz bound for Dirichlet problems)**.** *Assume* (F1) *and* (F2)*. Let* $\phi \in C^{\theta}(\overline{\Omega})$ *for some* $\theta \in (0,1)$*, and assume that* ϕ *satisfies* (3.1) *with* ρ *such that* $\overline{B}_\rho \subset \Omega$*. Assume that* $d \in C(\partial\Omega \times [0,T))$ *satisfies* (6.4)*. Let* u *be a viscosity solution of* (6.1)*–*(6.3)*. Then*

$$\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho] \geq Ct^{-\frac{1-\theta_0}{\theta_0}\alpha_0} \quad (t \in (0,T) \textit{ sufficiently small}).$$

Here C *depends on* θ_0*,* c_ϕ*,* ρ*,* L_d *and* D_0*. In particular, if* $\theta_0 = \theta$*, then*

$$\mathrm{Lip}[u(\cdot,t);\overline{B}_\rho] \geq Ct^{-\gamma_0} \quad (t \in (0,T) \textit{ sufficiently small}).$$

Acknowledgements

The authors are grateful to Professors Yasuhiro Fujita and Yoshikazu Giga for fruitful discussions. The work of the first author was partly supported by No. 19K14564 (Early-Career Scientists).

References

[1] D. G. Aronson, *Bounds for fundamental solutions of a parabolic equation*, Bull. Am. Math. Soc. **73** (1967), 890–896.

[2] D. G. Aronson, *Non-negative solutions of linear parabolic equations*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. **22** (1968), 607–694.

[3] D. G. Aronson, J. Serrin, *Local behavior of solutions of quasilinear parabolic equations*, Arch. Ration. Mech. Anal. **25** (1967), 81–122.

[4] M. Bardi, I. Capuzzo-Dolcetta, *Optimal control and viscosity solutions of Hamilton–Jacobi–Bellman equations*, With appendices by Maurizio Falcone and Pierpaolo Soravia, Systems & Control: Foundations & Applications, Birkhäuser Boston, Inc., Boston, MA, 1997, xviii+570 pp.

[5] 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), 1–67.

[6] E. A. Carlen, M. Loss, *Optimal smoothing and decay estimates for viscously damped conservation laws, with applications to the 2-D Navier-Stokes equation*, Duke Math. J. **81** (1995), 135–157.

[7] M. G. Crandall, P.-L. Lions, P. E. Souganidis, *Maximal solutions and universal bounds for some partial differential equations of evolution*, Arch. Rational Mech. Anal. **105** (1989), 163–190.

[8] N. Dirr, V. D. Nguyen, *Some new results on Lipschitz regularization for parabolic equations*, J. Evol. Equ. **19** (2019), 1149–1166.

[9] E. B. Fabes, D. W. Stroock, *A new proof of Moser's parabolic Harnack inequality using the old idea of Nash*, Arch. Ration. Mech. Anal. **96** (1986), 327–338.

[10] W. H. Fleming, H. M. Soner, *Controlled Markov processes and viscosity solutions*, Second edition, Stochastic Modelling and Applied Probability **25**, Springer, New York, 2006, xviii+429 pp.

[11] F. Flandoli, M. Gubinelli, E. Priola, *Well-posedness of the transport equation by stochastic perturbation*, Invent. Math. **180** (2010), 1–53.

[12] A. Friedman, *Partial differential equations of parabolic type*, Prentice-Hall, Inc., Englewood Cliffs, N.J. 1964, xiv+347 pp.

[13] Y. Fujita, *An optimal logarithmic Sobolev inequality with Lipschitz constants*, J. Funct. Anal. **261** (2011), 1133–1144.

[14] Y. Fujita, *Lower estimates of L^∞-norm of gradients for Cauchy problems*, J. Math. Anal. Appl. **458** (2018), 910–924.

[15] Y. Giga, *Surface evolution equations: A level set approach*, Monographs in Mathematics **99**, Birkhäuser Verlag, Basel, 2006, xii+264 pp.

[16] M.-H. Giga, Y. Giga, J. Saal, *Nonlinear partial differential equations: Asymptotic behavior of solutions and self-similar solutions*, Progress in Nonlinear Differential Equations and their Applications **79**, Birkhäuser Boston, Inc., Boston, MA, 2010, xviii+294 pp.

[17] O. Ley, *Lower-bound gradient estimates for first-order Hamilton–Jacobi equations and applications to the regularity of propagating fronts*, Adv. Differential Equations **6** (2001), 547–576.

[18] P.-L. Lions, *Generalized solutions of Hamilton–Jacobi equations*, Pitman (Advanced Publishing Program), Boston, Mass.-London, 1982, iv+317 pp.

[19] P.-L. Lions, *Regularizing effects for first-order Hamilton–Jacobi equations*, Applicable Anal. **20** (1985), 283–307.

[20] Y. Maekawa, *A lower bound for fundamental solutions of the heat convection equations*, Arch. Ration. Mech. Anal. **189** (2008), 45–58.

[21] S. Matsui, S. Tokuno, *Remark on fundamental solutions for vorticity equation of two dimensional Navier-Stokes flows*, Hokkaido Math. J. **26** (1997), 529–539.

[22] H. Osada, *Diffusion processes with generators of generalized divergence form*, J. Math. Kyoto Univ. **27** (1987), 597–619.

[23] A. Porretta, E. Priola, *Global Lipschitz regularizing effects for linear and nonlinear parabolic equations*, J. Math. Pures Appl. **100** (2013), 633–686.

[24] E. Priola, F. Y. Wang, *Gradient estimates for diffusion semigroups with singular coefficients*, J. Funct. Anal. **236** (2006), 244–264.