The Cauchy-Dirichlet problem for a general class of parabolic equations

note finali coverpage

(Article begins on next page)

23 April 2024

THE CAUCHY-DIRICHLET PROBLEM
FOR A GENERAL CLASS OF PARABOLIC EQUATIONS

PAOLO BARONI AND CASIMIR LINDFORS

ABSTRACT. We prove regularity results such as interior Lipschitz regularity and boundary continuity for the Cauchy-Dirichlet problem associated to a class of parabolic equations inspired by the evolutionary p-Laplacian, but extending it at a wide scale. We employ a regularization technique of viscosity-type that we find interesting in itself.

1. INTRODUCTION

The aim of this paper is the study of the behaviour of solutions to a wide class of nonlinear parabolic equations modeled after

$$u_t - \operatorname{div}\Big(\frac{g(|Du|)}{|Du|}Du\Big) = 0 \qquad \text{in} \qquad \Omega_T := \Omega \times (0,T) \subset \mathbb{R}^n \times \mathbb{R}, \qquad (1.1)$$

$n \geq 2$, where Ω is a bounded domain with $C^{1,\beta}$ boundary and $g : \mathbb{R}_+ \to \mathbb{R}_+$ is a C^1 function satisfying

$$g_0 - 1 \leq \mathcal{O}_g(s) := \frac{sg'(s)}{g(s)} \leq g_1 - 1 \qquad \text{for every } s > 0 \qquad (1.2)$$

with $1 < g_0 \leq g_1 < \infty$. Notice that we can assume $g_0 < g_1$ without loss of generality. Indeed, if $\mathcal{O}_g(s)$ is constant, say $\mathcal{O}_g(s) = p-1$ for some $p > 1$, a simple integration shows that $g(s) = s^{p-1}$ up to a constant factor, and therefore in this case (1.1) gives back the evolutionary p-Laplacian widely studied in particular by DiBenedetto, see the monograph [14]. This reveals that (1.1) is a natural generalization of the p-Laplacian, and in effect this class of growth conditions was mathematically introduced exactly in these terms by Lieberman in [28], even if this kind of condition appears earlier in the applications, see the forthcoming lines.

We stress that quite a comprehensive study of non-negative solutions to the equation

$$u_t - \operatorname{div}\big[\varphi'(u)Du\big] = 0 \qquad (1.3)$$

where the function $\varphi : [0,\infty) \to [0,\infty)$ satisfies

$$0 < a \leq \mathcal{O}_\varphi(s) := \frac{s\varphi'(s)}{\varphi(s)} \leq \frac{1}{a}, \quad \text{for } s > 0; \qquad 1 + a \leq \mathcal{O}_\varphi(s) \quad \text{for } s > s_0 \qquad (1.4)$$

for some $a \in (0,1)$ and some $s_0 > 0$ has been provided by Dahlberg and Kenig [10, 11]; see also the books [12, 35]. Clearly, while (1.3) is a generalization of the porous medium equation that happens when $\varphi(u) = u^m$, $m > 0$, in the same spirit (1.1) can be seen as a generalization of the p-Laplacian.

As in (1.4), we shall also consider a more stringent growth assumption for g for large values of its argument. In addition to (1.2), we shall assume that there exist constants $c_\ell, \epsilon > 0$ such that

$$g(s) \geq c_\ell s^{\frac{n-2}{n+2}+\epsilon} \qquad \text{for any } s \geq 1. \qquad (1.5)$$

Note that in the p-Laplacian case (1.5) reads precisely as $p > 2n/(n+2)$, a completely natural assumption in the theory of the evolutionary p-Laplacian operator, see [14, 25, 1]. Note moreover that (1.5) is implied by assuming $g_0 > 2n/(n+2)$, see Paragraph 2.2.

The regularity for the elliptic and variational counterpart of (1.1) is quite well understood, see for instance [28, 3, 9, 16] for the first argument and [17, 20, 7, 8] for the second, just to cite some indicative references. In the parabolic setting, however, very few results are available, and some of them only in particular cases: to our knowledge, only [22, 23, 29, 30], all by Lieberman and Hwang, and the recent [5].

The difficulty, in particular in finding zero-order results, stems from several facts, the main one perhaps being that the equation has very different behaviour, already in the p-Laplacian case, in the degenerate ($p \geq 2$) and singular ($p < 2$) cases. In the degenerate case phenomena such as expansion of positivity occur, see [15, 26], and the diffusion dominates [13]. On the other hand, in the singular case the evolutionary character dominates [6] and extinction of positive solutions in finite time could happen, see [14]. In our general setting the degenerate case occurs when $s \mapsto g(s)/s$ is increasing, and when it is decreasing we have the singular case. However, it might also happen that $s \mapsto g(s)/s$ has no monotonicity whatsoever, making the handling of the equation all the more difficult. The comprehension of the interaction of these different phenomena is the key for a better understanding of the behaviour of local solutions to (1.1), and in this paper we hope to start to clarify this difficult point, which will be the object of future investigations.

The class of differential operators we study, besides being quite a general extension of a well-known operator, finds important applications in the applied sciences, also in view of the following observation. Take the convex primitive G of g and consider the general minimization problem

$$u \in u_0 + W_0^{1,1}(\Omega) \mapsto \int_\Omega G(|Du|)\,dx; \qquad (1.6)$$

it is often convenient to have energies with a precise dependence on $|Du|$ of more general type than monomial (that is, the case of the p-Dirichlet energy or appropriate extensions). For instance, in mechanics, fluid dynamics and magnetism, as first approximation it is customary to have dependencies of the energy on the modulus of the gradient of monomial type but with exponent depending on the size of $|Du|$, in order to have mathematical models fitting the experimental data. In this case g is given by the gluing of different monomials (see the example in Paragraph 2.3). At this point, elliptic and parabolic equations having the growth described in (1.1) arise naturally as Euler equations or flows of the functional in (1.6). In [34], for instance, the two-dimensional stationary, irrotational subsonic flow of a compressible fluid is described using an energy defined in the following way:

$$G(s) = -\left(1 - \frac{\gamma-1}{2}s^2\right)^{\frac{\gamma}{\gamma-1}} \quad \text{for small } s, \qquad G(s) = \text{quadratic} \quad \text{otherwise,} \quad (1.7)$$

where $\gamma \in (1,2)$ is the exponent in the law $p \approx \rho^\gamma$ characteristic of polytropic gases.

More in general, see [4, 18, 19], one is lead to consider quasilinear static equations in dimension two and three of the type

$$\operatorname{div}\left[\rho(|Du|^2)Du\right] = 0,$$

with Du representing the *velocity field of the flow* and $q = |Du|$ being the *speed of the flow*. In this context one introduces the *Mach number*

$$M^2 \equiv [M(q)]^2 := -\frac{2q^2}{\rho(q^2)}\rho'(q^2)$$

(note that we must have $\rho' < 0$). In our context, where $g(s) = \rho(s^2)s$, we compute $\mathcal{O}_g(s) = 1 - M(s)^2$. The general theory asserts that a point is *elliptic* if $M < 1$ and in this case the flow is *subsonic*, while if $M > 1$ the point is *hyperbolic* and the flow there is

supersonic. If $M = 1$ the flow is called *sonic*. A solution of the boundary value problem is called a subsonic (supersonic) flow according to whether all points are subsonic (supersonic); note that *mixed, or transonic flows* can exists, with obvious meaning. However, if for some reason we know that the flow maintains a controlled, small speed q, then the problem falls in the class of operators we consider; the approximation in (1.7) is a way to study flows in the subsonic regime.

The object of our study will be the Cauchy-Dirichlet problem

$$\begin{cases} u_t - \operatorname{div} \mathcal{A}(Du) = 0 & \text{in } \Omega_T, \\ u = \psi & \text{on } \partial_p \Omega_T, \end{cases} \tag{1.8}$$

where $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$ is a C^1 vector field modeled after the one appearing in (1.1). In particular, we assume it satisfies the following ellipticity and growth conditions:

$$\begin{cases} \langle D\mathcal{A}(\xi)\lambda, \lambda \rangle \geq \nu \dfrac{g(|\xi|)}{|\xi|}|\lambda|^2 \\ |D\mathcal{A}(\xi)| \leq L \dfrac{g(|\xi|)}{|\xi|} \end{cases}, \tag{1.9}$$

for any $\xi \in \mathbb{R}^n \setminus \{0\}, \lambda \in \mathbb{R}^n$ and with structural constants $0 < \nu \leq 1 \leq L$; we assume without loss of generality that $\mathcal{A}(0) = 0$. The function g is a C^1 function as in (1.1), satisfying only (1.2) and (1.5). For what concerns ψ, we assume it to be continuous in $\partial_p \Omega_T$ with modulus of continuity ω_ψ with respect to the natural distance $\operatorname{dist}_{\mathrm{par},G}$, that is, there exists a continuous, concave function $\omega_\psi : \mathbb{R}_+ \to \mathbb{R}_+$ with $\omega_\psi(0) = 0$ such that

$$|\psi(x,t) - \psi(y,s)| \leq \omega_\psi\big(\max\{|x-y|, [G^{-1}(1/|t-s|)]^{-1}\}\big)$$

for every $(x,t), (y,s) \in \partial_p \Omega_T$. As already mentioned, Ω is a bounded domain of $\mathbb{R}^n$, $n \geq 2$, whose boundary is of class $C^{1,\beta}$ for some $\beta \in (0,1)$; we shall provide some more details at the beginning of Section 2.

In this setting, we state the main result of our paper, which concerns at the same time the existence and regularity of a (unique) solution to (1.8).

Theorem 1.1. *There exists a unique solution u, in the sense of Definition 2, to the Cauchy-Dirichlet problem* (1.8)*, where the vector field $\mathcal{A}$ satisfies the assumptions* (1.9)*, with $g \in C^1(\mathbb{R}_+)$ satisfying* (1.2) *and* (1.5)*. In particular, u is continuous up to the boundary and moreover if the boundary datum ψ is Hölder continuous with respect to the natural metric $\operatorname{dist}_{\mathrm{par},G}$ defined in* (2.1)*, then so is u.*

The following theorem gives some properties together with quantitative estimates for the solution described in the previous statement.

Theorem 1.2. *Let u be the solution to* (1.8) *given by Theorem 1.1. Then u is locally Lipschitz continuous and the following estimate holds:*

$$\|Du\|_{L^\infty(Q_R)} \leq c\bigg(\fint_{Q_{2R}} \big[G(|Du|) + 1\big]\, dx\, dt\bigg)^{\max\left\{\frac{1}{2}, \frac{2}{\epsilon(n+2)}\right\}} \tag{1.10}$$

for every parabolic cylinder $Q_{2R} \Subset \Omega_T$. The constant c depends on $n, g_0, g_1, \nu, L, \epsilon$ and c_ℓ. Moreover, there exists a modulus of continuity $\omega_u : \mathbb{R}_+ \mapsto \mathbb{R}_+$ depending on $n, g_0, g_1, \nu, L, \epsilon, c_\ell, \|\psi\|_{L^\infty}, \omega_\psi, \partial\Omega$ such that

$$|u(x,t) - u(y,s)| \leq \omega_u\big(\max\big\{|x-y|, \big[G^{-1}\big(1/|t-s|\big)\big]^{-1}\big\}\big) \tag{1.11}$$

for every $(x,t), (y,s) \in \overline{\Omega_T}^p$.

We refer the reader to Paragraph 2.1 for the definitions of the standard parabolic cylinders $Q_R(x_0, t_0)$ and of the parabolic closure of Ω_T. We also mention that in the standard case of the evolutionary p-Laplacian our estimate (1.10) gives back exactly the gradient sup-estimate available for degenerate and singular equations, see [14, Chapter VIII, Theorems 5.1 & 5.2].

Remark 1. Theorems 1.1 and 1.2 hold for a wider class of operators generalizing (1.1), which allow the presence of a function g that is not C^1 but merely Lipschitz. Indeed, we may consider Lipschitz functions $g : \mathbb{R}_+ \to \mathbb{R}_+$ satisfying (1.2) almost everywhere and vector fields $\mathcal{A} : \mathbb{R}^n \to \mathbb{R}^n$ in $W^{1,\infty}(\mathbb{R}^n)$ satisfying the monotonicity and Lipschitz assumptions

$$\begin{cases} \langle \mathcal{A}(\xi_1) - \mathcal{A}(\xi_2), \xi_1 - \xi_2 \rangle \geq \nu \dfrac{g(|\xi_1| + |\xi_2|)}{|\xi_1| + |\xi_2|} |\xi_1 - \xi_2|^2 \\[3mm] |\mathcal{A}(\xi_1) - \mathcal{A}(\xi_2)| \leq L \dfrac{g(|\xi_1| + |\xi_2|)}{|\xi_1| + |\xi_2|} |\xi_1 - \xi_2|, \end{cases} \tag{1.12}$$

for every $\xi_1, \xi_2 \in \mathbb{R}^n$ such that $|\xi_1| + |\xi_2| \neq 0$ and for some $0 < \nu \leq 1 \leq L$. For a proof of this fact see the end of Section 6.

1.1. **Novelties and technical tools.** We believe that the main interest of this paper, apart from the results of Theorems 1.1 and 1.2 themselves (that will be used for instance in [31]), is the development of some tools for the treatment of the difficult equation (1.8) (see Paragraph 2.3). We prove the Lipschitz estimate as an *a priori estimate* for problems enjoying further regularity. Instead of using a regularization of the type used in [28, 22, 29], the regularization we employ is of viscosity type, closer to that in [2]: we consider a vector field of the type

$$\mathcal{A}_\varepsilon(\xi) := (\phi_\varepsilon * \mathcal{A})(\xi) + \varepsilon \big(1 + |\xi|\big)^{p-2}\xi, \qquad \xi \in \mathbb{R}^n, \varepsilon \in (0, 1),$$

where $p \gg 1$ is a large exponent and $\{\phi_\varepsilon\}$ a family of mollifiers. This allows us to overcome the difficulties of deriving regularity estimates for the approximant problems, which we were not able to find in the literature. At this point continuity up to the boundary becomes an essential ingredient in the proof of the convergence, as well as the fact that we are solving a Cauchy-Dirichlet problem and therefore have a uniform bound on $\|u_\varepsilon\|_{L^\infty}$ given by the maximum principle.

We use the a priori Lipschitz continuity (and the further regularity) of the approximating solutions in a way inspired by [27]. First, we employ the fact that the function $v = |Du|^2$ is a subsolution to a similar problem, see Lemma 3.1. Then, we define an appropriate intrinsic geometry (see (3.7)) depending on the growth of the approximating vector field $\mathcal{A}_\varepsilon$, which allows us to rebalance estimates, in the sense that the weight appearing in the Caccioppoli estimate for the equation satisfied by v turns out to be essentially constant, see (3.8). Here the fact that we can bound the supremum of Du, and thus of v, from above is essential. Finally, we conclude the proof using an argument based on an alternative in order to get rid of the possible dependence on ε in terms of the aforementioned geometry, depending in turn on the growth of $\mathcal{A}_\varepsilon$.

2. PRELIMINARY MATERIAL: NOTATION, THE FUNCTION g, MISCELLANEA

For what concerns $\partial\Omega$, we assume that there exists a radius $R_\Omega > 0$ such that for every point $x_0 \in \partial\Omega$ there is a unit vector $\hat{e}_{x_0}$ such that the restriction of $\partial\Omega$ is a graph of a $C^{1,\beta}$ function in B_{R_Ω} along the $\hat{e}_{x_0}$ direction, in the following sense: with T being an orthogonal transformation that maps $\hat{e}_{x_0}$ into $(0, 0, \ldots, 0, 1)$, for every $0 < r \leq R_\Omega$ it holds

$$T^{-1}(\partial\Omega - x_0) \cap \big(B'_r \times (-r, r)\big) = \operatorname{graph}\theta$$

(see below for the precise meaning of these symbols) with $\theta \equiv \theta_{x_0} \in C^{1,\beta}(B'_r)$, $\theta(B'_r) \subset (-r, r)$ and the $C^{1,\beta}$ norm of θ uniformly bounded:

$$[\theta]_{C^{1,\beta}} \leq \Theta.$$

Note that without loss of generality, we can take $\hat{e}_{x_0}$ as the inner normal vector in x_0: $\{v : \langle v, \hat{e}_{x_0} \rangle = 0\}$ is the tangent hyperplane to Ω in x_0; therefore $D\theta(0) = 0$. $D\theta$ is the full gradient of θ with respect to its $n-1$ variables. Finally, by saying that a constant depends on $\partial\Omega$, we shall mean it depends on Θ.

2.1. **Notation.** We denote by c a general constant *always larger than or equal to one*, possibly varying from line to line; relevant dependencies on parameters will be emphasized using parentheses, i.e., $c_1 \equiv c_1(n, p, q)$ means that c_1 depends on n, p, q. For the ease of notation, we shall also use the following abbreviation:

$$\texttt{data} := \{n, g_0, g_1, \nu, L\}.$$

We denote by

$$B_R(x_0) := \{x \in \mathbb{R}^n : |x - x_0| < R\}$$

the open ball with center x_0 and radius $R > 0$; when clear from the context or otherwise not important, we shall omit denoting the center as follows: $B_R \equiv B_R(x_0)$. The standard parabolic cylinder is defined as

$$Q_R(x_0, t_0) := B_R(x_0) \times (t_0 - R^2, t_0),$$

while we define the natural cylinder as

$$Q_R^G(x_0, t_0) := B_R(x_0) \times (t_0 - [G(1/R)]^{-1}, t_0).$$

The latter is strictly linked to the scaling of the equation, see Paragraph 2.6. Unless otherwise explicitly stated, different balls and cylinders in the same context will have the same center. We shall denote, for a factor $\alpha > 0$, by αB_R the ball $B_{\alpha R}$ and by $\alpha Q_R(x_0, t_0)$ the cylinder $B_{\alpha R}(x_0) \times (t_0 - (\alpha R)^2, t_0)$; similarly for $\alpha Q_R^G(x_0, t_0)$. The parabolic boundary of a cylindrical domain $\mathcal{K} = \mathcal{D} \times \Gamma$, where $\mathcal{D}$ is an open domain and Γ an open interval of the real line, is defined as

$$\partial_p \mathcal{K} := \left(\overline{\mathcal{D}} \times \inf \Gamma\right) \cup \left(\partial \mathcal{D} \times \Gamma\right).$$

Naturally, the parabolic closure of $\mathcal{K}$ is then $\overline{\mathcal{K}}^p := \mathcal{K} \cup \partial_p \mathcal{K}$. Accordingly with the customary use in the parabolic setting, when considering a sub-cylinder $\mathcal{K}$ (as above) compactly contained in Ω_T, we shall mean that $\mathcal{D} \Subset \Omega$ and $0 < \inf \Gamma < \sup \Gamma \leq T$; we will write in this case $\mathcal{K} \Subset \Omega_T$. By $\partial\Omega - x_0$ we mean the set $\{x \in \mathbb{R}^n : x + x_0 \in \partial\Omega\}$. The standard parabolic distance is

$$\text{dist}_{\text{par}}\big((x, t), (y, s)\big) := \max\left\{|x - y|, \sqrt{|t - s|}\right\}$$

for any $(x, t), (y, s) \in \mathbb{R}^{n+1}$, while a distance strictly related to the scaling properties of the differential operator is

$$\text{dist}_{\text{par}, G}\big((x, t), (y, s)\big) := \max\left\{|x - y|, \left[G^{-1}\Big(\frac{1}{|t - s|}\Big)\right]^{-1}\right\}. \tag{2.1}$$

Note that $Q_R^G(x_0, t_0) = \{(x, t) \in \mathbb{R}^{n+1} : \text{dist}_{\text{par}, G}((x, t), (x_0, t_0)) < R, t < t_0\}$ and similarly for $Q_R(x_0, t_0)$. Accordingly we define the parabolic distance between sets as

$$\text{dist}_{\text{par}}(A, B) := \inf_{\substack{(x,t) \in A \\ (y,s) \in B}} \text{dist}_{\text{par}}\big((x, t), (y, s)\big)$$

for $A, B \subset \mathbb{R}^{n+1}$; similarly for $\text{dist}_{\text{par}, G}(A, B)$.

At a certain point it will be useful to split $\mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R}$. We agree here that we shall write a point $x \in \mathbb{R}^n$ as $(x', x_n) \in \mathbb{R}^{n-1} \times \mathbb{R}$; moreover, with $B'_r(x'_0)$ we shall denote the ball of $\mathbb{R}^{n-1}$ with radius r and center $x'_0 \in \mathbb{R}^{n-1}$.

With $\mathcal{B} \subset \mathbb{R}^\ell$ being a measurable set, $\chi_\mathcal{B}$ denotes its characteristic function. If furthermore $\mathcal{B}$ has positive and finite measure and $f : \mathcal{B} \to \mathbb{R}^k$ is a measurable map, we shall denote by

$$(f)_\mathcal{B} \equiv \fint_\mathcal{B} f(y)\,dy := \frac{1}{|\mathcal{B}|} \int_\mathcal{B} f(y)\,dy$$

the integral average of f over $\mathcal{B}$. If $\mathcal{B}$ is a cylinder, $\mathcal{B} := K \times \Gamma \subset \mathbb{R}^{n+1}$, then we shall denote the slicewise average by

$$(f)_K(\tau) := \fint_K f(y,\tau)\,dy.$$

for almost every $\tau \in \Gamma$. By sup we shall mean possibly the essential supremum, and similarly for inf. We shall also as usual denote

$$\operatorname*{osc}_\mathcal{B} f := \sup_\mathcal{B} f - \inf_\mathcal{B} f, \qquad [f]_{C^{0,\gamma}(\mathcal{B})} := \sup_{\substack{x,y \in \mathcal{B} \\ x \neq y}} \frac{|f(x) - f(y)|}{|x - y|^\gamma}.$$

$D_i f := \partial f/\partial x_i$, for $i \in \{1,\dots,n\}$, will stand for the partial derivative of f in the $\hat{e}_i$ direction, and $D_{i,j}^2 f$ will denote $\partial^2 f/\partial x_i \partial x_j$. Here $\hat{e}_i$ is the i-th element of the standard orthonormal basis of $\mathbb{R}^n$. By 2^* we shall denote the Sobolev conjugate exponent of 2, with the agreement that in the case $n = 2$ we fix the value of 2^* as 4, i.e.,

$$2^* := \begin{cases} \dfrac{2n}{n-2} & n > 2, \\[2mm] 4 & n = 2. \end{cases} \tag{2.2}$$

With s being a real number, we shall denote $s_+ := \max\{s,0\}$ and $s_- := \max\{-s,0\}$. For a vector $\xi = (\xi_1,\dots,\xi_n) \in \mathbb{R}^n$, $\operatorname{diag}\xi$ denotes the diagonal matrix $(\xi_i \delta_{i,j})_{i,j=1}^n$. Finally, $\mathbb{R}_+ := [0,\infty)$, $\mathbb{N}$ is the set $\{1,2,\dots\}$ and $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

By "equation structurally similar to $(1.8)_1$" we mean an equation of the type $\partial_t u - \operatorname{div}\widetilde{\mathcal{A}}(Du) = 0$ with $\widetilde{\mathcal{A}}$ satisfying assumptions (1.9) with ν, L and g replaced by $\tilde{\nu}$, $\widetilde{L}$ and $\tilde{g}$. Both $\tilde{\nu}, \widetilde{L}$ will depend on data, while $\tilde{g}$ will satisfy (1.2) and (1.5) with $\widetilde{g}_0$, $\widetilde{g}_0$, $\widetilde{c}_\ell$ depending on data and c_ℓ.

2.2. Properties of g.

Without loss of generality we assume that

$$\int_0^1 g(\rho)\,d\rho = 1. \tag{2.3}$$

Since (1.2) implies that the map $r \mapsto g(r) r^{-(g_0-1)}$ is increasing, while $r \mapsto g(r) r^{-(g_1-1)}$ turns out to be decreasing, we have

$$\min\{\alpha^{g_0-1}, \alpha^{g_1-1}\} g(r) \leq g(\alpha r) \leq \max\{\alpha^{g_0-1}, \alpha^{g_1-1}\} g(r)$$

for every $r, \alpha > 0$; clearly $g(0) = 0$ and $\lim_{r \to \infty} g(r) = \infty$. Since moreover g is strictly increasing, it has a strictly increasing inverse function $g^{-1} \in C^1(\mathbb{R}_+)$ with

$$\left(g^{-1}\right)'(r) = \frac{1}{g'(g^{-1}(r))} \qquad \text{for every } r > 0.$$

Using (1.2) we then see that also g^{-1} satisfies an Orlicz-type condition

$$\frac{1}{g_1 - 1} \leq \frac{(r g^{-1})'(r)}{g^{-1}(r)} \leq \frac{1}{g_0 - 1} \qquad \text{for every } r > 0. \tag{2.4}$$

Therefore, anything derived from (1.2) for g holds for g^{-1} with $g_0 - 1$ and $g_1 - 1$ replaced by $1/(g_1 - 1)$ and $1/(g_0 - 1)$, respectively.

Define the function $G : \mathbb{R}_+ \to \mathbb{R}_+$ as

$$G(r) := \int_0^r g(\rho)\,d\rho. \tag{2.5}$$

Clearly $G'(r) = g(r) > 0$ and $G''(r) = g'(r) > 0$ implying that G is both strictly increasing and strictly convex in $(0, \infty)$. Moreover, $G(0) = 0$ and $G(1) = 1$ due to (2.3). We also define $1/G(1/s) = 1/G^{-1}(1/s) = 0$ for $s = 0$. It is simple to check by integrating the function $r \mapsto rg(r)$ by parts and using (1.2) that also

$$g_0 \leq \frac{G'(r)r}{G(r)} \leq g_1 \tag{2.6}$$

holds true for $r > 0$.

Define the Young complement of G as

$$\widetilde{G}(r) = \sup_{s>0} (rs - G(s)) \qquad \text{or} \qquad \widetilde{G}(r) := \int_0^r g^{-1}(\rho) \, d\rho; \tag{2.7}$$

in our setting these definitions are equivalent, see [33]. Note that the Young's inequality

$$sr \leq G(s) + \widetilde{G}(r) \tag{2.8}$$

holds true for every $r, s > 0$ and by (2.4) and the second definition in (2.7) also $\widetilde{G}$ satisfies an Orlicz-type condition

$$\frac{g_1}{g_1 - 1} \leq \frac{\widetilde{G}'(r)r}{\widetilde{G}(r)} \leq \frac{g_0}{g_0 - 1}. \tag{2.9}$$

Now starting from (2.6) and (2.9), we deduce precisely as for g the inequalities

$$\min\{\alpha^{g_0}, \alpha^{g_1}\}G(r) \leq G(\alpha r) \leq \max\{\alpha^{g_0}, \alpha^{g_1}\}G(r), \tag{2.10}$$

and

$$\min\left\{\alpha^{\frac{g_1}{g_1-1}}, \alpha^{\frac{g_0}{g_0-1}}\right\}\widetilde{G}(r) \leq \widetilde{G}(\alpha r) \leq \max\left\{\alpha^{\frac{g_1}{g_1-1}}, \alpha^{\frac{g_0}{g_0-1}}\right\}\widetilde{G}(r)$$

for every $\alpha, r \geq 0$. These, together with Young's inequality (2.8), imply for $0 < \varepsilon < 1$

$$sr \leq G(\varepsilon^{\frac{1}{g_0}} s) + \widetilde{G}(\varepsilon^{-\frac{1}{g_0}} r) \leq \varepsilon G(s) + c(g_0, \varepsilon)\widetilde{G}(r).$$

Another useful property is

$$\widetilde{G}\left(\frac{G(r)}{r}\right) \leq G(r) \qquad \text{for every } r > 0,$$

see again [33] for the easy proof.

From the second assumption of (1.9) we easily derive an upper bound for $\mathcal{A}$. Indeed, when $\xi \in \mathbb{R}^n \setminus \{0\}$ we have

$$
\begin{aligned}
|\mathcal{A}(\xi)| \leq |\xi| \int_0^1 |D\mathcal{A}(s\xi)| \, ds &\leq L|\xi| \int_0^1 \frac{g(s|\xi|)}{s|\xi|} \, ds \\
&\leq c(L, g_0) \int_0^{|\xi|} g'(r) \, dr \leq c(g_0, g_1, L)\frac{G(|\xi|)}{|\xi|}; \quad (2.11)
\end{aligned}
$$

this holds also for $\xi = 0$ by our conventions, since $\mathcal{A}(0) = 0$. Similarly, the first assumption of (1.9) yields

$$\langle \mathcal{A}(\xi), \xi \rangle = \int_0^1 \langle D\mathcal{A}(s\xi)\xi, \xi \rangle \, ds \geq c(g_1, \nu)|\xi| \int_0^{|\xi|} g'(r) \, dr \geq c(g_0, g_1, \nu)G(|\xi|). \tag{2.12}$$

We define the quantity $V_g : \mathbb{R}^n \to \mathbb{R}^n$ by

$$V_g(\xi) = \left(\frac{g(|\xi|)}{|\xi|}\right)^{\frac{1}{2}} \xi$$

when $\xi \neq 0$ and set $V_g(0) = 0$. Clearly V_g is a continuous bijection of $\mathbb{R}^n$ and, moreover, has a continuous inverse by the inverse function theorem. Furthermore, the following monotonicity formula holds true:

$$\langle \mathcal{A}(\xi_1) - \mathcal{A}(\xi_2), \xi_1 - \xi_2 \rangle \geq c \frac{g(|\xi_1| + |\xi_2|)}{|\xi_1| + |\xi_2|} |\xi_1 - \xi_2|^2 \geq c |V_g(\xi_1) - V_g(\xi_2)|^2 \quad (2.13)$$

for a constant $c \equiv c(g_0, g_1, \nu)$ and for every $\xi_1, \xi_2 \in \mathbb{R}^n$, see [16, 17].

2.3. **A concrete example.** We give here a nontrivial example of a Lipschitz function g satisfying our assumptions - see Remark 1. This example is inspired by [28]. In particular we want to demonstrate the possibility that g oscillates between degenerate and singular behaviour. Suppose $2n/(n+2) < g_0 < g_1$ and set $\delta = (g_1 - g_0)/3 > 0$. Define the sequence $s_k = 2^{2^k}$ for $k \in \mathbb{N}_0$ and the function

$$g(s) = \begin{cases} s^{g_0 - 1 + \delta}, & 0 < s < 2 \\ s_{2k+1}^{-\delta} s^{g_1 - 1}, & s_{2k} \leq s < s_{2k+1} \\ s_{2k+2}^{\delta} s^{g_0 - 1}, & s_{2k+1} \leq s < s_{2k+2} \end{cases}.$$

Clearly g is Lipschitz and it satisfies (1.2). Moreover, (2.3) holds after scaling by a suitable normalization constant. We observe that

$$\limsup_{s \to \infty} \frac{g(s)}{s} = \begin{cases} \infty, & g_1 > 2 + \delta & (\text{iff} \quad g_0 + 2g_1 > 6) \\ 1, & g_1 = 2 + \delta & (\text{iff} \quad g_0 + 2g_1 = 6), \\ 0, & g_1 < 2 + \delta & (\text{iff} \quad g_0 + 2g_1 < 6) \end{cases}$$

$$\liminf_{s \to \infty} \frac{g(s)}{s} = \begin{cases} \infty, & g_0 > 2 - \delta & (\text{iff} \quad 2g_0 + g_1 > 6) \\ 1, & g_0 = 2 - \delta & (\text{iff} \quad 2g_0 + g_1 = 6). \\ 0, & g_0 < 2 - \delta & (\text{iff} \quad 2g_0 + g_1 < 6) \end{cases}$$

By taking $g_0 = 2 - \frac{3}{2n}, g_1 = 2 + \frac{3}{2n}$ we obtain a particularly interesting case, that is, we have $\liminf_{s \to \infty} g(s)/s = 0$ but $\limsup_{s \to \infty} g(s)/s = \infty$. Furthermore, if we consider the function

$$\widetilde{g}(s) = \frac{1}{g(1/s)},$$

we find similar behaviour as $s \to 0$. This is to say, we can build a structure function g (and accordingly a vector field $\mathcal{A}$ as in (1.1)) that, for $\ell \in \mathbb{N}$, along the sequence $\{\ell^{-k}\}_{k \in \mathbb{N}_0}$ the function $g(s)/s$ is at the same time as large and as close to zero as we wish, and therefore it does not enjoy any monotonicity properties. This gives a clue about the difficulty of the application of De Giorgi-type methods, in particular when they have to be matched with intrinsic geometries: note that the expressions of the type $G(s)/s^2 \approx g(s)/s$ appear already in the energy estimate for (1.1), see Lemma 2.3. On the other hand, when the quantity $g(|Du|)/|Du|$ is known to be under control, then the equation becomes treatable, see for instance Proposition 3.4 and in particular (3.8).

2.4. **Orlicz spaces.** For G as in (2.5), a measurable function $u : A \to \mathbb{R}$, $A \subset \mathbb{R}^k$, $k \in \mathbb{N}$ belongs to the Orlicz space $L^G(A)$ if it satisfies

$$\int_A G(|u|) \, dx < \infty.$$

The space $L^G(A)$ is a vector space, since G satisfies the Δ_2-condition (2.10), and it can be shown to be a Banach space if endowed with the Luxemburg norm

$$\|u\|_{L^G(A)} := \inf \left\{ \lambda > 0 : \int_A G\left(\frac{|u|}{\lambda}\right) dx \leq 1 \right\}.$$

A function u belongs to $L_{\mathrm{loc}}^G(A)$, if $u \in L^G(A')$ for every $A' \Subset A$. If also the weak gradient of u belongs to $L^G(A)$, we say that $u \in W^{1,G}(A)$. The corresponding space with zero boundary values, denoted $W_0^{1,G}(A)$, is the completion of $C_c^\infty(A)$ under the norm

$$\|u\|_{W^{1,G}(A)} := \|u\|_{L^G(A)} + \|Du\|_{L^G(A)}.$$

We denote by $V^G(\Omega_T)$ the space of functions $u \in L^G(\Omega_T) \cap L^1(0, T; W^{1,1}(\Omega))$ for which also the weak spatial gradient Du belongs to $L^G(\Omega_T)$. The space $V^G(\Omega_T)$ is also a Banach space with the norm

$$\|u\|_{V^G(\Omega_T)} := \|u\|_{L^G(\Omega_T)} + \|Du\|_{L^G(\Omega_T)}.$$

Moreover, we denote by $V_0^G(\Omega_T)$ the space of functions $u \in V^G(\Omega_T)$ that belong to $W_0^{1,G}(\Omega)$ for almost every $t \in (0, T)$, while the localized version $V_{\mathrm{loc}}^G(\Omega_T)$ is defined, as above, in the customary way. We also shorten

$$V^{2,G}(\Omega_T) := L^\infty\big(0, T; L^2(\Omega)\big) \cap V^G(\Omega_T)$$

and similarly for the localized and the zero trace versions. We shall moreover denote $V^{2,p}(\Omega_T)$, for $p > 1$, the space $V^{2,G}(\Omega_T)$ for the choice $G(s) = s^p$.

2.5. The concept of solution and consequences.
We fix here the notions of solution employed in this paper.

Definition 1. A function u is a *weak solution* to $(1.8)_1$ in a cylindrical domain $\mathcal{K} \subset \mathbb{R}^{n+1}$, with the vector field $\mathcal{A}$ satisfying the assumptions (1.9), if $u \in V_{\mathrm{loc}}^{2,G}(\mathcal{K})$ and it satisfies the weak formulation

$$\int_{\mathcal{K}} \big[-u \partial_t \eta + \langle \mathcal{A}(Du), D\eta \rangle \big] \, dx \, dt = 0 \tag{2.14}$$

for every test function $\eta \in C_c^\infty(\mathcal{K})$. If instead of equality we have the $\leq$ $(\geq)$ sign for every nonnegative $\eta \in C_c^\infty(\mathcal{K})$, we say that u is a *weak subsolution (supersolution)* in $\mathcal{K}$.

Definition 2. A function u is a solution to the Cauchy-Dirichlet problem (1.8) if $u \in C^0(\overline{\Omega}_T)$ is a weak solution to $(1.8)_1$ in Ω_T and moreover $u = \psi$ pointwise on $\partial_p \Omega_T$.

A very useful formulation, equivalent to (2.14), is the one involving Steklov averages. Indeed, the mild regularity of a solution does not allow us to use it as a test function. Furthermore, it is sometimes useful to have a weak formulation allowing for test functions independent of time, or test functions possibly vanishing only on the parabolic boundary of a cylinder. Apart from mollification, the possible way to have such properties involve the so-called Steklov averaging regularization of a function: for $f : \mathcal{K} = \mathcal{D} \times (t_1, t_2) \to \mathbb{R}$ measurable and $0 < |h| \ll 1$ appropriate, it is defined as

$$f_h(x, t) := \frac{1}{h} \int_{t-h}^t f(x, s) \, ds \qquad \text{for } (x, t) \in \mathcal{D} \times (t_1 + h, t_2);$$

note that we employ the *backward regularization*. If $f \in L^q(\mathcal{K})$ for some $q \geq 1$, then $f_h \to f$ in $L^q(\mathcal{D} \times (t_1 + \varepsilon, t_2))$ for every $\varepsilon > 0$; the same holds in the L^G spaces. Moreover, if $f \in C^0(t_1, t_2; L^q(\mathcal{D}))$ then $f_h(\cdot, \tau) \to f(\cdot, \tau)$ in $L^q(\mathcal{D})$ for a.e. $\tau \in (t_1 + \varepsilon, t_2)$ and for every $\varepsilon > 0$.

At this point it is quite easy to infer the following slicewise formulation for weak solutions (see [14]) using density arguments with respect to the spatial variable:

$$\int_{\mathcal{D}} \big[\partial_t u_h(\cdot, \tau) \eta + \langle [\mathcal{A}(Du)]_h(\cdot, \tau), D\eta \rangle \big] \, dx = 0 \tag{2.15}$$

for every $\eta \in W_0^{1,G}(\mathcal{D})$, almost every $\tau \in (t_1 + h, t_2)$, and $h > 0$ such that the functions are well defined. Similar results hold also for weak super- and subsolutions.

Proposition 2.1. (Comparison principle) *Let* $\mathcal{K} := \mathcal{D} \times (t_1, t_2) \subset \Omega_T$ *and let* $u \in C^0(\overline{\mathcal{K}}^p)$ *be a weak subsolution to* (1.8)$_1$ *and* $v \in C^0(\overline{\mathcal{K}}^p)$ *a weak supersolution to* (1.8)$_1$ *in* $\mathcal{K}$. *If* $u \leq v$ *on* $\partial_p \mathcal{K}$, *then* $u \leq v$ *in* $\overline{\mathcal{K}}^p$.

Proof. For $\varepsilon > 0$ fixed define $\varphi_\varepsilon(t) := (t_2 - \varepsilon - t)_+$ and test (2.14) formally with

$$\eta = (u_h - v_h - \varepsilon)_+ \varphi_\varepsilon.$$

Note that η is compactly supported in $\mathcal{K}$ due to the continuity of u and v and the fact that $u \leq v$ on $\partial_p Q$. Subtracting the Steklov version of the variational inequality of v from that of u and integrating over (t_1, t_2) yields

$$\int_{\mathcal{K}} \partial_t(u_h - v_h)\eta \, dx \, dt + \int_{\mathcal{K}} \langle [\mathcal{A}(Du)]_h - [\mathcal{A}(Dv)]_h, D\eta \rangle \, dx \, dt \leq 0.$$

By the monotonicity of $\mathcal{A}$, Lemma 2.13, we have

$$\int_{\mathcal{K}} \langle [\mathcal{A}(Du)]_h - [\mathcal{A}(Dv)]_h, D\eta \rangle \, dx \, dt$$

$$\rightarrow \int_{\mathcal{K} \cap \{u > v + \varepsilon\}} \langle \mathcal{A}(Du) - \mathcal{A}(Dv), Du - Dv \rangle \varphi_\varepsilon \, dx \, dt \geq 0,$$

and for the parabolic term we obtain using integration by parts

$$\int_{\mathcal{K}} \partial_t(u_h - v_h)(u_h - v_h - \varepsilon)_+ \varphi_\varepsilon \, dx \, dt = -\frac{1}{2} \int_{\mathcal{K}} (u_h - v_h - \varepsilon)_+^2 \partial_t \varphi_\varepsilon \, dx \, dt$$

$$\rightarrow \frac{1}{2} \int_{t_1}^{t_2 - \varepsilon} \int_{\mathcal{D}} (u - v - \varepsilon)_+^2 \, dx \, dt$$

as $h \rightarrow 0$. Combining these gives

$$\int_{t_1}^{t_2 - \varepsilon} \int_{\mathcal{D}} (u - v - \varepsilon)_+^2 \, dx \, dt \leq 0,$$

which implies $u \leq v + \varepsilon$ almost everywhere in $\mathcal{D} \times (t_1, t_2 - \varepsilon)$. Since this holds for every $\varepsilon > 0$ and $u, v \in C(\overline{Q})$, the result follows. $\square$

Observe that the uniqueness of a solution to the Cauchy-Dirichlet problem (1.8) follows immediately from the previous result. Moreover, we have the following corollary.

Corollary 2.2. (Maximum principle) *Let* $\mathcal{K} \subset \Omega_T$ *and let* $u \in C(\overline{\mathcal{K}}^p)$ *be a weak solution to* (1.8)$_1$ *in* $\mathcal{K}$. *Then*

$$\inf_{\partial_p \mathcal{K}} u \leq u \leq \sup_{\partial_p \mathcal{K}} u$$

in $\overline{\mathcal{K}}^p$ *and, moreover,*

$$\sup_{\overline{\mathcal{K}}} |u| = \sup_{\partial_p \mathcal{K}} |u|.$$

We recall the following standard energy inequality for local weak solutions. We give it in a more general form for future reference.

Lemma 2.3 (Caccioppoli's inequality). *Let* $\mathcal{K} := \mathcal{D} \times (t_1, t_2) \Subset \Omega_T$ *and let* u *be a weak solution to* (1.8) *in* $\mathcal{K}$. *Then there exists a constant* $c \equiv c(g_0, g_1, \nu, L)$ *such that*

$$\sup_{\tau \in (t_1, t_2)} \int_{\mathcal{D}} \left[(u - k)_\pm^2 \varphi^{g_1} \right](\cdot, \tau) \, dx + \int_{\mathcal{K}} G\big(|D(u - k)_\pm|\big) \varphi^{g_1} \, dx \, dt$$

$$\leq \int_{\mathcal{D}} \left[(u - k)_\pm^2 \varphi^{g_1} \right](\cdot, t_1) \, dx + c \int_{\mathcal{K}} \left[G\big(|D\varphi|(u - k)_\pm\big) + (u - k)_\pm^2 |\partial_t \varphi| \right] \, dx \, dt$$

for any $k \in \mathbb{R}$ *and for every* $\varphi \in W^{1,\infty}(\mathcal{K})$ *vanishing in a neighborhood of* $\partial \mathcal{D} \times (t_1, t_2)$ *and with* $0 \leq \varphi \leq 1$. *The same inequality but only with the "+" sign holds for weak subsolutions.*

Proof. Fix $\varphi \in W^{1,\infty}(\mathcal{K})$ as in the statement of the Lemma, call $w := \pm(u - k)_{\pm}$ and choose $\eta = w_h \varphi^{g_1}$ as the test function in (2.15). Then we integrate over (t_1, τ) for $\tau \in (t_1, t_2)$ to obtain

$$\int_{\mathcal{K}} \partial_t u_h \, w_h \varphi^{g_1} \chi_{(t_1,\tau)} \, dx \, dt + \int_{\mathcal{K}} \langle [\mathcal{A}(Du)]_h, D(w_h \varphi^{g_1}) \rangle \chi_{(t_1,\tau)} \, dx \, dt = 0. \quad (2.16)$$

Integration by parts gives

$$\int_{\mathcal{K}} \partial_t u_h \, w_h \varphi^{g_1} \chi_{(t_1,\tau)} \, dx \, dt = \frac{1}{2} \int_{t_1}^{\tau} \int_{\mathcal{D}} \partial_t(w_h^2) \varphi^{g_1} \, dx \, dt \quad (2.17)$$

$$= \frac{1}{2} \int_{\mathcal{D}} w_h^2 \varphi^{g_1} \, dx \Big|_{t=t_1}^{\tau} - \frac{1}{2} \int_{t_1}^{\tau} \int_{\mathcal{D}} w_h^2 \partial_t(\varphi^{g_1}) \, dx \, dt$$

$$\to \frac{1}{2} \int_{\mathcal{D}} w^2 \varphi^{g_1} \, dx \Big|_{t=t_1}^{\tau} - \frac{1}{2} \int_{t_1}^{\tau} \int_{\mathcal{D}} w^2 \partial_t(\varphi^{g_1}) \, dx \, dt$$

as $h \to 0$. For the elliptic part we have by (2.12)

$$\int_{\mathcal{K}} \langle [\mathcal{A}(Du)]_h, D(w_h \varphi^{g_1}) \rangle \chi_{(t_1,\tau)} \, dx \, dt$$

$$\to \int_{t_1}^{\tau} \int_{\mathcal{D}} \langle \mathcal{A}(Du), Dw \rangle \varphi^{g_1} \, dx \, dt + g_1 \int_{t_1}^{\tau} \int_{\mathcal{D}} \langle \mathcal{A}(Du), D\varphi \rangle w \, \varphi^{g_1-1} \, dx \, dt$$

$$\geq c_1 \int_{t_1}^{\tau} \int_{\mathcal{D}} G(|Dw|) \varphi^{g_1} \, dx \, dt - \left| g_1 \int_{\mathcal{K}} \langle \mathcal{A}(Du), D\varphi \rangle w \, \varphi^{g_1-1} \, dx \, dt \right|,$$

where c_1 depends on g_0, g_1, ν. Furthermore, by (2.11), Young's inequality with $\varepsilon \in (0,1)$ to be chosen and the properties of g we obtain

$$\left| g_1 \int_{\mathcal{K}} \langle \mathcal{A}(Du), D\varphi \rangle w \, \varphi^{g_1-1} \, dx \, dt \right| \leq g_1 \int_{\mathcal{K}} |\mathcal{A}(Dw)||D\varphi||w|\varphi^{g_1-1} \, dx \, dt$$

$$\leq \varepsilon c_2 \int_{\mathcal{K}} \widetilde{G}\left(\frac{G(|Dw|)}{|Dw|} \varphi^{g_1-1} \right) dx \, dt + c(\varepsilon) \int_{\mathcal{K}} G(|D\varphi||w|) \, dx \, dt$$

$$\leq \varepsilon c_2 \int_{\mathcal{K}} G(|Dw|)\varphi^{g_1} \, dx \, dt + c(\varepsilon) \int_{\mathcal{K}} G(|D\varphi||w|) \, dx \, dt, \quad (2.18)$$

where c_2 depends on g_0, g_1, L and $c(\varepsilon)$ depends on g_0, g_1, L as well as on ε. Now, combining (2.17)-(2.18) with (2.16) yields

$$\frac{1}{2} \int_{\mathcal{D}} w^2 \varphi^{g_1} \, dx \Big|_{t=t_1}^{\tau} - \frac{1}{2} \int_{\mathcal{K}} w^2 \partial_t(\varphi^{g_1}) \, dx \, dt + c_1 \int_{t_1}^{\tau} \int_{\mathcal{D}} G(|Dw|)\varphi^{g_1} \, dx \, dt$$

$$\leq \varepsilon c_2 \int_{\mathcal{K}} G(|Dw|)\varphi^{g_1} \, dx \, dt + c(\varepsilon) \int_{\mathcal{K}} G(|D\varphi||w|) \, dx \, dt.$$

We conclude by taking the essential supremum with respect to $\tau \in (t_1, t_2)$, choosing $\varepsilon \in (0,1)$ such that $\varepsilon c_2 \leq c_1/2$, reabsorbing the term on the right-hand side and recalling the definition of w.

The proof for subsolutions is very similar, taking into account that the test function η must be nonnegative. $\qquad \square$

2.6. **The geometry of the problem.** In order to understand the equation, the first thing we want to stress is its scaling. Suppose u solves the model equation (1.1) in $Q_1 = B_1 \times (-1, 0)$ and let $\kappa > 0$. Then the function

$$\bar{u}(x,t) := \kappa u\left(\frac{x - x_0}{r}, \frac{1}{\kappa^2} G\left(\frac{\kappa}{r} \right)(t - t_0) \right)$$

solves in

$$Q_r^\kappa(x_0, t_0) := B_r(x_0) \times \left(t_0 - \kappa^2 \left[G\left(\frac{\kappa}{r} \right) \right]^{-1}, t_0 \right)$$

the equation

$$\bar{u}_t - \operatorname{div}\left(\frac{\bar{g}(|D\bar{u}|)}{|D\bar{u}|}D\bar{u}\right) = 0, \tag{2.19}$$

where

$$\bar{g}(s) := \frac{\kappa}{r}\left[G\left(\frac{\kappa}{r}\right)\right]^{-1}g\left(\frac{\kappa}{r}s\right). \tag{2.20}$$

The function $\bar{g}$ has the same structure as g, in the sense that it satisfies (1.2) exactly with parameters g_0 and g_1 and moreover, we have $\overline{G}(1) = 1$, where

$$\overline{G}(s) := \int_0^s \bar{g}(\sigma)\,d\sigma = \left[G\left(\frac{\kappa}{r}\right)\right]^{-1}G\left(\frac{\kappa}{r}s\right).$$

Conversely, if we have a solution w to (1.8) in Q_r^κ, then

$$\bar{w}(x,t) := \frac{1}{\kappa}w\left(x_0 + rx, t_0 + \kappa^2\left[G\left(\frac{\kappa}{r}\right)\right]^{-1}t\right)$$

solves (2.19) in Q_1 with $\bar{g}$ as in (2.20). In case we consider the general equation (1.8), the same scaling argument holds if we consider the vector field

$$\overline{\mathcal{A}}(\xi) := \frac{\kappa}{r}\left[G\left(\frac{\kappa}{r}\right)\right]^{-1}\mathcal{A}\left(\frac{\kappa}{r}\xi\right)$$

which satisfies the structural conditions (1.9) with g replaced by the function $\bar{g}$.

2.7. **Other auxiliary results.** The following Lemma encodes the self-improving property of reverse Hölder inequalities. We take the form proposed in [27, Lemma 5.1] with slight changes in order to meet our purposes.

Lemma 2.4. *Let μ be a nonnegative Borel measure with finite total mass. Moreover, let $\gamma > 1$ and $\{\sigma\mathcal{Q}\}_{0<\sigma\leq 1}$ be a family of open sets with the property*

$$\sigma'\mathcal{Q} \subset \sigma\mathcal{Q} \subset 1\mathcal{Q} = \mathcal{Q}$$

whenever $0 < \sigma' < \sigma \leq 1$. If $w \in L^2(\mathcal{Q})$ is a nonnegative function satisfying

$$\left(\int_{\sigma'\mathcal{Q}} w^{2\gamma}\,d\mu\right)^{1/(2\gamma)} \leq \frac{c_0}{\sigma - \sigma'}\left(\int_{\sigma\mathcal{Q}} w^2\,d\mu\right)^{1/2}$$

for all $1/2 \leq \sigma' < \sigma \leq 1$, then for any $0 < q < 2$ there is a positive constant $c \equiv c(c_0, \gamma, q)$ such that

$$\left(\int_{\sigma\mathcal{Q}} w^{2\gamma}\,d\mu\right)^{1/(2\gamma)} \leq \frac{c}{(1-\sigma)^\xi}\left(\int_{\mathcal{Q}} w^q\,d\mu\right)^{1/q},$$

for all $0 < \sigma < 1$, where $\xi := \frac{2\gamma - q}{q(\gamma - 1)}$.

The next one is a classic iteration Lemma.

Lemma 2.5. *Let $\phi : [R, 2R] \to [0, \infty)$ be a function such that*

$$\phi(r) \leq \frac{1}{2}\phi(s) + \frac{A}{(s-r)^\beta} + B \qquad \text{for every } R \leq r < s \leq 2R,$$

where $A, B \geq 1$ and $\beta > 0$. Then

$$\phi(R) \leq c(\beta)\left[\frac{A}{R^\beta} + B\right].$$

3. A priori Lipschitz estimates

In this section we impose on u an additional regularity assumption and prove intrinsic estimates for the gradient of u. To be precise, we shall suppose

$$u, Du \in C^0_{\text{loc}}(\Omega_T), \qquad u \in L^2_{\text{loc}}(0, T; W^{2,2}_{\text{loc}}(\Omega)). \tag{3.1}$$

This is to say, we shall prove the estimates of this section as *a priori estimates*, leaving to Section 4 the approximation procedure which will explain how to deduce the desired estimates without the additional assumption (3.1). Notice that the continuity of u and Du allows us to treat their pointwise values. Due to the assumed extra regularity it will be possible to differentiate the equation; this will be done by showing that the function

$$v := |Du|^2 \tag{3.2}$$

is a subsolution to a similar equation.

Lemma 3.1. *Let u be a weak solution to $(1.8)_1$ in Ω_T and, moreover, assume that the regularity assumptions (3.1) hold. Then v is a weak subsolution to*

$$\partial_t v - \operatorname{div}(D\mathcal{A}(Du)Dv) = 0 \qquad \text{in } \Omega_T. \tag{3.3}$$

Proof. Formally, the idea is to differentiate equation $(1.8)_1$ with respect to x_j for $j = 1, \ldots, n$, then multiply by $D_j u$, and finally sum over j. To this end, let $0 \le \varphi \in C_c^\infty(\Omega_T)$, and test (2.14) with

$$\eta = -D_j(D_j u \varphi).$$

This choice can be justified by using Steklov averages, as done previously in the paper; we shall proceed formally. Integration by parts yields

$$0 = -\int_{\Omega_T} u\, \partial_t(-D_j(D_j u\varphi))\, dx\, dt + \int_{\Omega_T} \langle \mathcal{A}(Du), D(-D_j(D_j u\varphi)) \rangle\, dx\, dt$$

$$= \int_{\Omega_T} \partial_t(D_j u) D_j u \varphi\, dx\, dt + \int_{\Omega_T} \langle D_j \mathcal{A}(Du), D(D_j u\varphi) \rangle\, dx\, dt$$

$$= -\frac{1}{2} \int_{\Omega_T} |D_j u|^2 \partial_t \varphi\, dx\, dt + \frac{1}{2} \int_{\Omega_T} \langle D\mathcal{A}(Du) D(|D_j u|^2), D\varphi \rangle\, dx\, dt$$

$$+ \int_{\Omega_T} \langle D\mathcal{A}(Du) DD_j u, DD_j u \rangle \varphi\, dx\, dt.$$

Now, since

$$\int_{\Omega_T} \langle D\mathcal{A}(Du) DD_j u, DD_j u \rangle \varphi\, dx\, dt \ge \nu \int_{\Omega_T} \frac{g(|Du|)}{|Du|} |DD_j u|^2 \varphi\, dx\, dt \ge 0$$

by $(1.9)_1$, summing up over $j = 1, \ldots, n$ leads to

$$-\int_{\Omega_T} |Du|^2 \partial_t \varphi\, dx\, dt + \int_{\Omega_T} \langle D\mathcal{A}(Du) D|Du|^2, D\varphi \rangle\, dx\, dt \le 0.$$

This proves the claim. $\qquad \square$

Next we prove a Caccioppoli inequality of porous medium type for the function v.

Lemma 3.2. *Let u be a weak solution of (1.8) in Ω_T and assume that (3.1) holds. Let $\mathcal{K} := \mathcal{D} \times (t_1, t_2) \Subset \Omega_T$ and $k \in \mathbb{R}$. Then there exists a constant $c \equiv c(\nu, L)$ such that*

$$\sup_{\tau \in (t_1, t_2)} \int_{\mathcal{D}} [(v - k)_+^2 \varphi^2](\cdot, \tau)\, dx + \int_{\mathcal{K}} \frac{g(|Du|)}{|Du|} |D(v - k)_+|^2 \varphi^2\, dx\, dt$$

$$\le c \int_{\mathcal{K}} (v - k)_+^2 \left(\frac{g(|Du|)}{|Du|} |D\varphi|^2 + |\partial_t \varphi| \right) dx\, dt$$

for every $\varphi \in C^\infty(\mathcal{K})$ vanishing in a neighborhood of $\partial_p \mathcal{K}$.

Proof. We can take

$$\eta = (v - k)_+ \varphi^2 \chi_{(t_1, \tau)}$$

for $\tau \in (t_1, t_2)$ as the test function in the weak formulation of (3.3), up to a regularization similar to the previous ones. For the parabolic part we have

$$-\int_{t_1}^{\tau} \int_{\mathcal{D}} v \, \partial_t \big((v - k)_+ \varphi^2\big) \, dx \, dt = \frac{1}{2} \int_{t_1}^{\tau} \int_{\mathcal{D}} \partial_t (v - k)_+^2 \varphi^2 \, dx \, dt$$

$$= \frac{1}{2} \int_{\mathcal{D}} [(v - k)_+^2 \varphi^2](\cdot, \tau) \, dx - \frac{1}{2} \int_{t_1}^{\tau} \int_{\mathcal{D}} (v - k)_+^2 \partial_t \varphi^2 \, dx \, dt.$$

The elliptic term can be estimated from below by using the assumptions (1.9) and Young's inequality with $\varepsilon = \nu/(2L)$. This gives

$$\int_{t_1}^{\tau} \int_{\mathcal{D}} \big\langle D\mathcal{A}(Du)Dv, D\big((v - k)_+ \varphi^2\big) \big\rangle \, dx \, dt$$

$$= \int_{t_1}^{\tau} \int_{\mathcal{D}} \big\langle D\mathcal{A}(Du)D(v - k)_+, D(v - k)_+ \big\rangle \varphi^2 \, dx \, dt$$

$$+ 2 \int_{t_1}^{\tau} \int_{\mathcal{D}} \big\langle D\mathcal{A}(Du)D(v - k)_+, D\varphi \big\rangle (v - k)_+ \varphi \, dx \, dt$$

$$\geq \nu \int_{t_1}^{\tau} \int_{\mathcal{D}} \frac{g(|Du|)}{|Du|} |D(v - k)_+|^2 \varphi^2 \, dx \, dt$$

$$- 2L \int_{t_1}^{\tau} \int_{\mathcal{D}} \frac{g(|Du|)}{|Du|} |D(v - k)_+||D\varphi| \, (v - k)_+ \varphi \, dx \, dt$$

$$\geq \frac{\nu}{2} \int_{t_1}^{\tau} \int_{\mathcal{D}} \frac{g(|Du|)}{|Du|} |D(v - k)_+|^2 \varphi^2 \, dx \, dt$$

$$- c(\nu, L) \int_{t_1}^{\tau} \int_{\mathcal{D}} \frac{g(|Du|)}{|Du|} |D\varphi|^2 \, (v - k)_+^2 \, dx \, dt,$$

and thus, we obtain

$$\int_{\mathcal{D}} [(v - k)_+^2 \varphi^2](\cdot, \tau) \, dx + \nu \int_{t_1}^{\tau} \int_{\mathcal{D}} \frac{g(|Du|)}{|Du|} |D(v - k)_+|^2 \varphi^2 \, dx \, dt$$

$$\leq c \int_{\mathcal{K}} \frac{g(|Du|)}{|Du|} |D\varphi|^2 \, (v - k)_+^2 \, dx \, dt + \int_{\mathcal{K}} (v - k)_+^2 |\partial_t \varphi| \, dx \, dt.$$

Since $\tau \in (t_1, t_2)$ was arbitrary, the result follows. $\qquad\square$

Combining the previous lemma with Sobolev's inequality leads to the following estimate.

Lemma 3.3. *Let the assumptions of Lemma 3.2 be in force. Then there exists a constant* $c \equiv c(n, g_1, \nu, L)$ *such that*

$$\fint_{\mathcal{K}} \frac{g(|Du|)}{|Du|} (v - k)_+^{2\gamma} \varphi^{2\gamma} \, dx \, dt$$

$$\leq c \, |\mathcal{D}|^{2/n} (t_2 - t_1)^{\gamma - 1} \left(\fint_{\mathcal{K}} (v - k)_+^2 \left(\frac{g(|Du|)}{|Du|} |D\varphi|^2 + |\partial_t \varphi| \right) dx \, dt \right)^{\gamma}, \quad (3.4)$$

where - recall (2.2) -

$$\gamma := 2 - \frac{2}{2^*} > 1.$$

Proof. By Hölder's and Sobolev's inequalities we have

$$\fint_{\mathcal{K}} \frac{g(|Du|)}{|Du|} (v - k)_+^{2\gamma} \varphi^{2\gamma} \, dx \, dt$$

$$= \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \fint_{\mathcal{D}} \frac{g(|Du|)}{|Du|} (v-k)_+^2 \varphi^2 \big((v-k)_+^2 \varphi^2 \big)^{1-2/2^*} \, dx \, dt$$

$$\le \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \left(\fint_{\mathcal{D}} \left(\Big(\frac{g(|Du|)}{|Du|} \Big)^{1/2} (v-k)_+ \varphi \right)^{2^*} dx \right)^{2/2^*}$$

$$\times \left(\fint_{\mathcal{D}} (v-k)_+^2 \varphi^2 \, dx \right)^{1-2/2^*} dt$$

$$\le c(n) |\mathcal{D}|^{2/n} \left(\sup_{\tau \in (t_1, t_2)} \fint_{\mathcal{D}} [(v-k)_+^2 \varphi^2](\cdot, \tau) \, dx \right)^{1-2/2^*}$$

$$\times \fint_{\mathcal{K}} \left| D\left(\Big(\frac{g(|Du|)}{|Du|} \Big)^{1/2} (v-k)_+ \varphi \right) \right|^2 dx \, dt. \quad (3.5)$$

A straightforward calculation yields

$$\left| D\left(\Big(\frac{g(|Du|)}{|Du|} \Big)^{1/2} (v-k)_+ \varphi \right) \right|^2$$

$$= \left| \left[\frac{(v-k)_+}{4v} \left(\frac{|Du| g'(|Du|)}{g(|Du|)} - 1 \right) + 1 \right] \Big(\frac{g(|Du|)}{|Du|} \Big)^{1/2} D(v-k)_+ \varphi \right.$$

$$\left. + \Big(\frac{g(|Du|)}{|Du|} \Big)^{1/2} (v-k)_+ D\varphi \right|^2$$

$$\le c(g_1) \frac{g(|Du|)}{|Du|} |D(v-k)_+|^2 \varphi^2 + 2 \frac{g(|Du|)}{|Du|} (v-k)_+^2 |D\varphi|^2,$$

and thus, integrating and estimating the first term using Lemma 3.2 yields

$$\fint_{\mathcal{K}} \left| D\left(\Big(\frac{g(|Du|)}{|Du|} \Big)^{\frac{1}{2}} (v-k)_+ \varphi \right) \right|^2 dx \, dt$$

$$\le c \fint_{\mathcal{K}} (v-k)_+^2 \left(\frac{g(|Du|)}{|Du|} |D\varphi|^2 + |\partial_t \varphi| \right) dx \, dt,$$

where the constant c depends only on g_1, ν, and L. From Lemma 3.2 it also follows that

$$\sup_{\tau \in (t_1, t_2)} \fint_{\mathcal{D}} [(v-k)_+^2 \varphi^2](\cdot, \tau) \, dx$$

$$\le c \, (t_2 - t_1) \fint_{\mathcal{K}} (v-k)_+^2 \left(\frac{g(|Du|)}{|Du|} |D\varphi|^2 + |\partial_t \varphi| \right) dx \, dt;$$

therefore, by inserting the previous two inequalities into (3.5) we obtain (3.4). $\qquad \square$

Next the aim is to prove an intrinsic reverse Hölder's inequality. To this end, let $Q_\rho(x_0, t_0) \subset \Omega_T$, let $\lambda \ge 1$ be such that

$$\lambda \ge \frac{1}{4} \sup_{Q_\rho(x_0, t_0)} |Du|, \tag{3.6}$$

and set

$$\theta_\lambda := \frac{g(\lambda)}{\lambda}.$$

We introduce the intrinsic cylinder

$$Q_\rho^\lambda \equiv Q_\rho^\lambda(x_0, t_0) := \min\{1, \theta_\lambda\}^{1/2} B_\rho(x_0) \times \big(t_0 - \min\{1, \theta_\lambda^{-1}\} \rho^2, t_0 \big). \tag{3.7}$$

Note that we have the alternative expression

$$Q_\rho^\lambda = \begin{cases} B_\rho(x_0) \times \left(t_0 - \theta_\lambda^{-1}\rho^2, t_0\right), & \theta_\lambda \geq 1 \\ \\ \theta_\lambda^{1/2} B_\rho(x_0) \times \left(t_0 - \rho^2, t_0\right), & 0 < \theta_\lambda < 1, \end{cases}$$

from which we easily see the analogy with the intrinsic geometry used to handle the parabolic p-Laplacian, recalling that in this case $g(s)/s = s^{p-2}$ and λ is "dimensionally comparable" to $|Du|$. Observe that we clearly have $Q_\rho^\lambda(x_0, t_0) \subset Q_\rho(x_0, t_0)$ in any case.

Lemma 3.4. *Let u be a weak solution to* $(1.8)_1$ *in Ω_T, assume that* (3.1) *and* (3.6) *hold and let $q > 0$. Then there exists a constant $c \equiv c(n, g_1, \nu, L, q)$ such that*

$$\left(\fint_{Q_{\rho/2}^\lambda} (v-k)_+^{2\gamma}\, dx\, dt\right)^{1/(2\gamma)} \leq c\left(\fint_{Q_\rho^\lambda} (v-k)_+^q\, dx\, dt\right)^{1/q}$$

for every $k \geq \lambda^2$.

Proof. Let $1/2 \leq \sigma' < \sigma \leq 1$ and choose a cut-off function $\varphi \in C^\infty(\sigma Q_\rho^\lambda)$ vanishing in the neighborhood of $\partial_p(\sigma Q_\rho^\lambda)$ such that $0 \leq \varphi \leq 1$, $\varphi = 1$ in $\sigma' Q_\rho^\lambda$, and

$$|D\varphi| \leq \frac{c}{\rho(\sigma - \sigma')} \min\left\{1, \theta_\lambda\right\}^{-1/2}, \qquad |\partial_t \varphi| \leq \frac{c}{\rho^2(\sigma - \sigma')^2} \min\left\{1, \theta_\lambda^{-1}\right\}^{-1}.$$

Observe that by the inclusion $Q_\rho^\lambda(x_0, t_0) \subset Q_\rho(x_0, t_0)$ and (3.6) we have

$$|Du| \leq 4\lambda \qquad \text{in } Q_\rho^\lambda.$$

Moreover, we have $|Du| \geq \lambda$ in the support of $(v-k)_+$, since $k \geq \lambda^2$ and $v = |Du|^2$. Thus, by using the properties of g we obtain

$$\frac{1}{4}\theta_\lambda \leq \frac{g(|Du|)}{|Du|} \leq c(g_1)\theta_\lambda \tag{3.8}$$

in $Q_\rho^\lambda \cap \{v \geq k\}$. Now Lemma 3.3 yields

$$\fint_{\sigma' Q_\rho^\lambda} (v-k)_+^{2\gamma}\, dx\, dt \leq c(n)\,\theta_\lambda^{-1} \fint_{\sigma Q_\rho^\lambda} \frac{g(|Du|)}{|Du|}(v-k)_+^{2\gamma}\varphi^{2\gamma}\, dx\, dt$$

$$\leq c\,\theta_\lambda^{-1}\Big| \min\{1, \theta_\lambda\}^{1/2} B_{\sigma\rho}\Big|^{2/n} \Big(\min\{1, \theta_\lambda^{-1}\}(\sigma\rho)^2 \Big)^{\gamma-1}$$

$$\times \left(\fint_{\sigma Q_\rho^\lambda} (v-k)_+^2 \left(\frac{g(|Du|)}{|Du|}|D\varphi|^2 + |\partial_t\varphi|\right) dx\, dt\right)^\gamma$$

$$\leq c\,\theta_\lambda^{-1} \min\{1, \theta_\lambda\} \min\{1, \theta_\lambda^{-1}\}^{\gamma-1}\rho^{2\gamma}$$

$$\times \left(\frac{\theta_\lambda \min\{1, \theta_\lambda\}^{-1} + \min\{1, \theta_\lambda^{-1}\}^{-1}}{\rho^2(\sigma - \sigma')^2} \fint_{\sigma Q_\rho^\lambda} (v-k)_+^2\, dx\, dt\right)^\gamma$$

$$= \frac{c}{(\sigma - \sigma')^{2\gamma}} \left(\fint_{\sigma Q_\rho^\lambda} (v-k)_+^2\, dx\, dt\right)^\gamma.$$

This is to say

$$\left(\fint_{\sigma' Q_\rho^\lambda} (v-k)_+^{2\gamma}\, dx\, dt\right)^{1/(2\gamma)} \leq \frac{c}{\sigma - \sigma'}\left(\fint_{\sigma Q_\rho^\lambda} (v-k)_+^2\, dx\, dt\right)^{1/2},$$

where the constant c depends only on n, g_1, ν, L.

Next we use Lemma 2.4 with $w = (v - k)_+$ and $d\mu = \frac{1}{|Q_\rho^\lambda|} \, dx \, dt$. This gives for every $0 < q < 2$ a constant $c \equiv c(n, g_1, \nu, L, q)$ such that

$$\left(\fint_{Q_{\rho/2}^\lambda} (v - k)_+^{2\gamma} \, dx \, dt \right)^{1/(2\gamma)} \le c \left(\fint_{Q_\rho^\lambda} (v - k)_+^q \, dx \, dt \right)^{1/q};$$

the case $q \ge 2$ now follows from Hölder's inequality. $\qquad\square$

Iterating the previous result yields the following pointwise estimate.

Proposition 3.5. *Let u be a weak solution to (1.8) in Ω_T and assume that (3.1) holds. Then for every $q > 0$ there exists a constant $c \equiv c(n, g_1, \nu, L, q)$ such that*

$$|Du(x_0, t_0)| \le \lambda + c \left(\fint_{Q_\rho^\lambda(x_0, t_0)} \left(|Du|^2 - \lambda^2 \right)_+^q \, dx \, dt \right)^{1/(2q)}$$

holds for every λ satisfying (3.6).

Proof. The idea is to apply De Giorgi's iteration method with the aid of Lemma 3.4. Let us first consider the case $0 < q < 2$. To this end, choose for $j \in \mathbb{N}_0$

$$\rho_j = 2^{-j} \rho, \qquad k_j = \lambda^2 + (1 - 2^{-j}) d,$$

where $d > 0$ is to be determined later. Observe that $\rho_0 = \rho$, $k_0 = \lambda^2$, and ρ_j decreases to zero and k_j increases to $\lambda^2 + d$ as j tends to infinity; clearly $k_j \ge \lambda^2$. Denote $Q_j := Q_{\rho_j}^\lambda(x_0, t_0)$ and

$$Y_j := \left(\fint_{Q_j} (v - k_j)_+^q \, dx \, dt \right)^{1/q} \qquad \text{for } j \in \mathbb{N}_0.$$

By Lemma 3.4 we have

$$\left(\fint_{Q_{j+1}} (v - k_j)_+^{2\gamma} \, dx \, dt \right)^{1/(2\gamma)} \le c \left(\fint_{Q_j} (v - k_j)_+^q \, dx \, dt \right)^{1/q},$$

and since $k_{j+1} > k_j$ implies

$$(v - k_j)_+^{2\gamma} \ge (k_{j+1} - k_j)^{2\gamma - q} (v - k_{j+1})_+^q \chi_{\{v \ge k_{j+1}\}},$$

we obtain

$$Y_{j+1} \le \frac{c}{(k_{j+1} - k_j)^\beta} \left(\fint_{Q_{j+1}} (v - k_j)_+^{2\gamma} \, dx \, dt \right)^{2\gamma/q} \le c^* d^{-\beta} 2^{\beta j} Y_j^{1+\beta},$$

for every $j \in \mathbb{N}_0$, where $\beta := 2\gamma/q - 1 > 0$ and $c^* \equiv c^*(n, g_1, \nu, L, q)$. Then a standard hyper-geometric iteration lemma implies $Y_j \to 0$ as $j \to \infty$, provided that

$$Y_0 \le (2c^*)^{-\frac{1}{\beta}} d$$

and this can be guaranteed by choosing

$$d = (2c^*)^{\frac{1}{\beta}} \left(\fint_{Q_\rho^\lambda(x_0, t_0)} (v - \lambda^2)_+^q \, dx \, dt \right)^{1/q}.$$

Now Lebesgue's differentiation theorem yields

$$\left(v(x_0, t_0) - (\lambda^2 + d) \right)_+ = \lim_{j \to \infty} \left(\fint_{Q_j} \left(v - (\lambda^2 + d) \right)_+^q \, dx \, dt \right)^{1/q} \le \lim_{j \to \infty} Y_j = 0,$$

which implies, recalling the choice of d,

$$v(x_0, t_0) \le \lambda^2 + c \left(\fint_{Q_\rho^\lambda(x_0, t_0)} (v - \lambda^2)_+^q \, dx \, dt \right)^{1/q}.$$

The case $q \ge 2$ follows again by Hölder's inequality. $\qquad\square$

4. Approximation

In this section we regularize the equation in order to apply the results of the previous section and show that the gradient of the solution to the regularized equation is uniformly bounded. Then all we have left to prove is that the approximating solutions converge to a function that solves the original equation.

To this end, define for $\varepsilon \in (0,1)$

$$\mathcal{A}_\varepsilon(\xi) := (\phi_\varepsilon * \mathcal{A})(\xi) + \varepsilon\big(1 + |\xi|\big)^{\widetilde{g}_1 - 2}\xi, \tag{4.1}$$

where $\phi_\varepsilon(\xi) = \phi(\xi/\varepsilon)/\varepsilon^n$; ϕ is a standard mollifier with $\int_{\mathbb{R}^n} \phi\, dx = 1$. That is, we mollify the vector field $\mathcal{A}$ and perturb it with the nondegenerate $\widetilde{g}_1$-Laplacian, where $\widetilde{g}_1 > \max\{g_1, 2\}$; we can take for example $\widetilde{g}_1 := g_1 + 1$. It is straightforward to see that $\mathcal{A}_\varepsilon$ satisfies (1.9) with g replaced by

$$g_\varepsilon(s) := \frac{g(s+\varepsilon)}{s+\varepsilon}s + \varepsilon(1+s)^{\widetilde{g}_1 - 2}s \tag{4.2}$$

and L, ν replaced by $\widetilde{L} = c(n, g_1)L, \widetilde{\nu} = \nu/c(n, g_1)$, see also Paragraph 6.1. Now the key point is that $\mathcal{O}_{g_\varepsilon}$ can be bounded *independently of* ε. Indeed, we have

$$\widetilde{g}_0 - 1 \le \mathcal{O}_{g_\varepsilon}(s) \le \widetilde{g}_1 - 1,$$

where $\widetilde{g}_0 := \min\{g_0, 2\}$. Note that g_ε also satisfies the lower bound in (1.5), since $g_\varepsilon(s) \ge g(s)/2$ for $s \ge 1$.

Let $u_\varepsilon \in V^{2,\widetilde{g}_1}(\Omega_T) \cap C^0(\overline{\Omega}_T)$ be the solution to the Cauchy-Dirichlet problem

$$\begin{cases} \partial_t u_\varepsilon - \operatorname{div}\mathcal{A}_\varepsilon(Du_\varepsilon) = 0 & \text{in } \Omega_T, \\ u_\varepsilon = \psi & \text{on } \partial_p\Omega_T; \end{cases} \tag{4.3}$$

for existence and uniqueness of such solutions see for instance [24]. Since

$$\varepsilon(1+s)^{\widetilde{g}_1 - 2} \le \frac{g_\varepsilon(s)}{s} \le \frac{c(g_1)}{\varepsilon}(1+s)^{\widetilde{g}_1 - 2},$$

in addition to satisfying g_ε-ellipticity and -growth conditions analogous to (1.9), the vector field $\mathcal{A}_\varepsilon$ also enjoys *nondegenerate p-Laplacian growth conditions* with $p = \widetilde{g}_1$. Hence, by standard theory, u_ε satisfies the assumption (3.1), see [14, 27]; therefore the results of the previous section are at our disposal for $u \equiv u_\varepsilon$. Note that all the constants will turn out to be effectively *independent of* ε.

Let us then show how to apply the result of the previous section in order to locally bound the gradient of the approximating solution uniformly in terms of ε. Here we also prove an estimate that, once convergence is established, leads to (1.10). Observe that the assumption (1.5) is crucial in this proof. We shall shorten $\|\psi\|_{L^\infty} \equiv \|\psi\|_{L^\infty(\partial_p\Omega_T)}$.

Proposition 4.1. *Let u_ε be a solution to (4.3) and let $\mathcal{K} \Subset \Omega_T$. Then $\|Du_\varepsilon\|_{L^\infty(\mathcal{K})}$ is bounded by a constant depending on* `data`*, $\epsilon, c_\ell, \|\psi\|_{L^\infty}$, and $\operatorname{dist}_{\mathrm{par}}(\partial_p\Omega_T, \mathcal{K})$, but independent of ε.*

Proof. Let us consider a standard parabolic cylinder $Q_{4R} \equiv Q_{4R}(x^*, t^*) \subset \Omega_T$ and a subcylinder $Q_\rho(x_0, t_0) \subset Q_{2R}$. Moreover, let $\lambda \ge 1$ be such that

$$\lambda \ge \frac{1}{4}\sup_{Q_\rho(x_0,t_0)}|Du_\varepsilon|. \tag{4.4}$$

We divide the proof into two cases depending on which term of g_ε dominates at λ.

Case I. Assume

$$\frac{g(\lambda + \varepsilon)}{\lambda + \varepsilon} \leq \varepsilon(1 + \lambda)^{\widetilde{g}_1 - 2}.$$

Setting

$$\theta_\lambda^\varepsilon := \frac{g_\varepsilon(\lambda)}{\lambda} = \frac{g(\lambda + \varepsilon)}{\lambda + \varepsilon} + \varepsilon(1 + \lambda)^{\widetilde{g}_1 - 2}$$

we clearly have

$$\varepsilon(1 + \lambda)^{\widetilde{g}_1 - 2} \leq \theta_\lambda^\varepsilon \leq 2\varepsilon(1 + \lambda)^{\widetilde{g}_1 - 2}. \tag{4.5}$$

By applying Proposition [3.5](#) to u_ε with $q = \widetilde{g}_1/2$ we obtain

$$|Du_\varepsilon(x_0, t_0)| \leq \lambda + c\left(\fint_{Q_\rho^\lambda(x_0,t_0)} \left(|Du_\varepsilon|^2 - \lambda^2\right)_+^{\widetilde{g}_1/2} dx\,dt\right)^{1/\widetilde{g}_1}$$

$$\leq \lambda + c\left(\frac{\max\{1, \theta_\lambda^\varepsilon\}}{\min\{1, \theta_\lambda^\varepsilon\}^{n/2}} \fint_{Q_\rho(x_0,t_0)} |Du_\varepsilon|^{\widetilde{g}_1} dx\,dt\right)^{1/\widetilde{g}_1},$$

since $Q_\rho^\lambda(x_0, t_0) \subset Q_\rho(x_0, t_0)$.

We further distinguish two cases: in the case when $\theta_\lambda^\varepsilon \geq 1$ we get

$$\frac{\max\{1, \theta_\lambda^\varepsilon\}}{\min\{1, \theta_\lambda^\varepsilon\}^{n/2}} = \theta_\lambda^\varepsilon \leq 2\varepsilon(1 + \lambda)^{\widetilde{g}_1 - 2},$$

while when $0 < \theta_\lambda^\varepsilon < 1$ we have

$$\frac{\max\{1, \theta_\lambda^\varepsilon\}}{\min\{1, \theta_\lambda^\varepsilon\}^{n/2}} = (\theta_\lambda^\varepsilon)^{-n/2} \leq \left(\varepsilon(1 + \lambda)^{\widetilde{g}_1 - 2}\right)^{-n/2} = \varepsilon\left(\varepsilon^{1+2/n}(1 + \lambda)^{\widetilde{g}_1 - 2}\right)^{-n/2}; \tag{4.6}$$

in both cases we have used [(4.5)](#). Since

$$\varepsilon \geq \frac{g(\lambda + \varepsilon)}{(1 + \lambda)^{\widetilde{g}_1 - 2}(\lambda + \varepsilon)}$$

$$\geq c_\ell(\lambda + \varepsilon)^{\frac{n-2}{n+2} + \epsilon - 1}(1 + \lambda)^{2 - \widetilde{g}_1}$$

$$\geq c_\ell(1 + \lambda)^{2 - \widetilde{g}_1 + \min\{\epsilon - 4/(n+2), 0\}} =: c_\ell(1 + \lambda)^{\bar{\eta}}$$

by [(1.5)](#) and the fact that $\lambda \geq 1$, plugging this estimate into [(4.6)](#) yields

$$\left(\varepsilon^{1+2/n}(1 + \lambda)^{\widetilde{g}_1 - 2}\right)^{-n/2} \leq c(n, c_\ell)(1 + \lambda)^{-(\bar{\eta}(1+2/n) + \widetilde{g}_1 - 2)n/2}$$

$$\leq c(n, c_\ell)(1 + \lambda)^{\widetilde{g}_1 - \min\{\epsilon(n+2)/2, 2\}};$$

a direct computation shows indeed the relation between the exponents. Hence we have

$$|Du_\varepsilon(x_0, t_0)| \leq \lambda + c(1 + \lambda)^{1 - \min\{\epsilon(n+2)/(2\widetilde{g}_1), 2/\widetilde{g}_1\}}$$

$$\times \left(\varepsilon \fint_{Q_\rho(x_0,t_0)} |Du_\varepsilon|^{\widetilde{g}_1} dx\,dt\right)^{1/\widetilde{g}_1}$$

$$\leq 2\lambda + c\left(\varepsilon \fint_{Q_\rho(x_0,t_0)} |Du_\varepsilon|^{\widetilde{g}_1} dx\,dt\right)^{\max\left\{\frac{2}{\epsilon(n+2)}, \frac{1}{2}\right\}} + 1$$

by Young's inequality; we also used $\widetilde{g}_1 > 2$.

Case II. Suppose then that

$$\frac{g(\lambda + \varepsilon)}{\lambda + \varepsilon} > \varepsilon(1 + \lambda)^{\tilde{g}_1 - 2}.$$

Here we have

$$\frac{g(\lambda + \varepsilon)}{\lambda + \varepsilon} \leq \theta_\lambda^\varepsilon \leq 2\frac{g(\lambda + \varepsilon)}{\lambda + \varepsilon},$$

and again by Proposition 3.5

$$|Du_\varepsilon(x_0, t_0)| \leq \lambda + c\left(\frac{\max\{1, \theta_\lambda^\varepsilon\}}{\min\{1, \theta_\lambda^\varepsilon\}^{n/2}} \fint_{Q_\rho(x_0, t_0)} \left(|Du_\varepsilon|^2 - \lambda^2\right)_+^q dx\, dt\right)^{\frac{1}{2q}}.$$

When $\theta_\lambda^\varepsilon \geq 1$, choosing $q = 1$ leads to

$$|Du_\varepsilon(x_0, t_0)| \leq \lambda + c\left(\frac{g(\lambda + \varepsilon)}{\lambda + \varepsilon} \fint_{Q_\rho(x_0, t_0)} \left(|Du_\varepsilon|^2 - \lambda^2\right)_+ dx\, dt\right)^{\frac{1}{2}}$$

$$\leq \lambda + c\left(\fint_{Q_\rho(x_0, t_0)} \frac{g(|Du_\varepsilon|)}{|Du_\varepsilon|}\left(|Du_\varepsilon|^2 - \lambda^2\right)_+ dx\, dt\right)^{\frac{1}{2}}$$

$$\leq \lambda + c\left(\fint_{Q_\rho(x_0, t_0)} G(|Du_\varepsilon|)\chi_{\{|Du_\varepsilon| \geq 1\}} dx\, dt\right)^{\frac{1}{2}}.$$

The second inequality stems from the fact that

$$\frac{1}{2}(\lambda + \varepsilon) \leq |Du_\varepsilon| \leq 4(\lambda + \varepsilon) \tag{4.7}$$

in the set $Q_\rho(x_0, t_0) \cap \{|Du_\varepsilon| \geq \lambda\}$ by (4.4), while for the last one we used (2.6) and the fact that $\lambda \geq 1$.

In the case $0 < \theta_\lambda^\varepsilon < 1$ we choose $q = \epsilon(n + 2)/4$ and use (1.5) and again (4.7) to obtain

$$|Du_\varepsilon(x_0, t_0)| \leq \lambda + c\left(\left(\frac{g(\lambda + \varepsilon)}{\lambda + \varepsilon}\right)^{-\frac{n}{2}} \fint_{Q_\rho(x_0, t_0)} \left(|Du_\varepsilon|^2 - \lambda^2\right)_+^q dx\, dt\right)^{\frac{1}{2q}}$$

$$\leq \lambda + c\left(\fint_{Q_\rho(x_0, t_0)} \left(\frac{|Du_\varepsilon|}{g(|Du_\varepsilon|)}\right)^{\frac{n}{2}}\left(|Du_\varepsilon|^2 - \lambda^2\right)_+^q dx\, dt\right)^{\frac{1}{2q}}$$

$$\leq \lambda + c\left(\fint_{Q_\rho(x_0, t_0)} |Du_\varepsilon|^{\left(1 - \frac{n-2}{n+2} - \epsilon\right)\frac{n}{2} + 2q}\chi_{\{|Du_\varepsilon| \geq 1\}} dx\, dt\right)^{\frac{1}{2q}}$$

$$= \lambda + c\left(\fint_{Q_\rho(x_0, t_0)} |Du_\varepsilon|^{1 + \frac{n-2}{n+2} + \epsilon}\chi_{\{|Du_\varepsilon| \geq 1\}} dx\, dt\right)^{\frac{2}{\epsilon(n+2)}}$$

$$\leq \lambda + c\left(\fint_{Q_\rho(x_0, t_0)} G(|Du_\varepsilon|)\chi_{\{|Du_\varepsilon| \geq 1\}} dx\, dt\right)^{\frac{2}{\epsilon(n+2)}};$$

note that

$$\left(1 - \frac{n-2}{n+2} - \epsilon\right)\frac{n}{2} + 2q = \left(\frac{4}{n+2} - \epsilon\right)\frac{n}{2} + \epsilon\frac{n+2}{2} = 1 + \frac{n-2}{n+2} + \epsilon.$$

Therefore in both cases we have

$$|Du_\varepsilon(x_0, t_0)| \leq \lambda + c\left(\fint_{Q_\rho(x_0, t_0)} G(|Du_\varepsilon|)\chi_{\{|Du_\varepsilon| \geq 1\}} dx\, dt\right)^{\max\left\{\frac{1}{2}, \frac{2}{\epsilon(n+2)}\right\}}.$$

Combining Cases I and II and denoting $\tilde{\eta} := \max\left\{\frac{1}{2}, \frac{2}{\epsilon(n+2)}\right\}$ yields

$$|Du_\varepsilon(x_0, t_0)|$$

$$\leq 2\lambda + c\left(\fint_{Q_\rho(x_0,t_0)}\left(G(|Du_\varepsilon|)\chi_{\{|Du_\varepsilon|\geq 1\}} + \varepsilon|Du_\varepsilon|^{\widetilde{g}_1}\right)dx\,dt\right)^{\tilde\eta} + 1$$

$$\leq 2\lambda + c\left(\frac{R}{\rho}\right)^{(n+2)\tilde\eta}\left(\fint_{Q_{2R}} G_\varepsilon(|Du_\varepsilon|)\,dx\,dt\right)^{\tilde\eta} + 1, \tag{4.8}$$

since

$$G(s) \leq \frac{1}{g_0}g(s+\varepsilon)(s+\varepsilon) \leq \frac{4}{g_0}\frac{g(s+\varepsilon)}{s+\varepsilon}s^2 \leq \frac{4}{g_0}sg_\varepsilon(s) \leq \frac{4\widetilde{g}_1}{g_0}G_\varepsilon(s)$$

for $s \geq 1$ and trivially

$$\varepsilon s^{\widetilde{g}_1} \leq \varepsilon(1+s)^{\widetilde{g}_1-2}s^2 \leq \widetilde{g}_1 G_\varepsilon(s).$$

The constant c in (4.8) depends only on `data`, ϵ, c_ℓ.

Let us now choose two intermediate cylinders $Q_R \subset Q_r \Subset Q_s \subset Q_{2R}$ and fix

$$\lambda := 1 + \frac{1}{4}\|Du_\varepsilon\|_{L^\infty(Q_s)} < \infty, \qquad (x_0,t_0) \in Q_r, \qquad \rho := \frac{s-r}{2} > 0.$$

Clearly $Q_\rho(x_0,t_0) \subset Q_s$ so that (4.4) holds. Then (4.8) implies

$$\|Du_\varepsilon\|_{L^\infty(Q_r)} \leq \frac{1}{2}\|Du_\varepsilon\|_{L^\infty(Q_s)}$$

$$+ c\left(\frac{R}{s-r}\right)^{(n+2)\tilde\eta}\left(\fint_{Q_{2R}} G_\varepsilon(|Du_\varepsilon|)\,dx\,dt\right)^{\tilde\eta} + 3.$$

Now, by choosing $\phi(r) = \|Du_\varepsilon\|_{L^\infty(Q_r)}$, iteration Lemma 2.5 gives

$$\|Du_\varepsilon\|_{L^\infty(Q_R)} \leq c\left(\fint_{Q_{2R}}\left[G_\varepsilon(|Du_\varepsilon|)+1\right]dx\,dt\right)^{\tilde\eta}. \tag{4.9}$$

At this point, in order to get rid of the dependence on ε on the right-hand side, the idea is to use the Caccioppoli inequality of Lemma 2.3 to translate the dependence on Du_ε to one on u_ε, and the latter in turn into a dependence on ψ. Indeed, take $\varphi \in C^\infty(Q_{4R})$ vanishing in a neighborhood of $\partial_p Q_{4R}$ such that $0 \leq \varphi \leq 1$, $\varphi = 1$ in Q_{2R}, and $|D\varphi|^2 + |\partial_t\varphi| \leq c/R^2$. Since

$$\sup_{Q_{4R}}|u_\varepsilon| \leq \sup_{\partial_p\Omega_T}|u_\varepsilon| = \sup_{\partial_p\Omega_T}|\psi| \leq \|\psi\|_{L^\infty}$$

by the maximum principle, Corollary 2.2, we can estimate by Lemma 2.3

$$\fint_{Q_{2R}} G_\varepsilon(|Du_\varepsilon|)\,dx\,dt \leq c\fint_{Q_{4R}}\left[G_\varepsilon(|D\varphi||u_\varepsilon|)+u_\varepsilon^2|\partial_t\varphi|\right]dx\,dt$$

$$\leq c\left(1+\frac{\|\psi\|_{L^\infty}}{R}\right)^{\widetilde{g}_1} + c\left(\frac{\|\psi\|_{L^\infty}}{R}\right)^2$$

$$= c\left(\texttt{data},\epsilon,c_\ell,\|\psi\|_{L^\infty},R\right). \tag{4.10}$$

Note that the constant does not depend on ε. Therefore we conclude the proof of the Proposition, modulo a standard covering argument. $\qquad\square$

4.1. A uniform interior modulus of continuity via Lipschitz regularity.

In this section we prove that the approximating solutions u_ε are equicontinuous in the interior of the domain; in particular we shall show their equi-Lipschitz regularity with respect to the parabolic metric.

Proposition 4.2. *Let u_ε be a solution to (4.3). Then $u_\varepsilon \in \mathrm{Lip}(1,1/2)(\Omega_T)$ locally, uniformly in ε; this is to say, for every subcylinder $\mathcal{K} \Subset \Omega_T$ there exists a constant c depending on* `data`, $\epsilon, c_\ell, \|\psi\|_{L^\infty}$, *and* $\mathrm{dist}_{\mathrm{par}}(\partial_p\Omega_T,\mathcal{K})$ *such that*

$$|u_\varepsilon(x,t) - u_\varepsilon(y,s)| \leq c\,\mathrm{dist}_{\mathrm{par}}\big((x,t),(y,s)\big) \tag{4.11}$$

for every $(x,t), (y,s) \in \mathcal{K}$ and for every $\varepsilon \in (0,1)$.

Proof. Fix an intermediate set $\mathcal{K}'$ such that $\mathcal{K} \Subset \mathcal{K}' \Subset \Omega_T$ and

$$\operatorname{dist}_{\mathrm{par}}(\hat{z}, \partial_p \Omega_T) = \operatorname{dist}_{\mathrm{par}}(\mathcal{K}, \partial_p \Omega_T)/2 =: d/2$$

for every $\hat{z} \in \partial_p \mathcal{K}'$. Take also a cylinder $Q_r(x_0, t_0) \subset \mathcal{K}'$ with $(x_0, t_0) \in \mathcal{K}$; this will happen for instance if $r \leq d/2$. Since Du_ε is continuous, by applying the divergence theorem and using the bound for $\mathcal{A}_\varepsilon$ in (2.11) we infer

$$\fint_{B_r(x_0)} u_\varepsilon(\cdot, \tau)\, dx \Big|_{\tau = t_1}^{t_2} = \frac{n}{r} \int_{t_1}^{t_2} \fint_{\partial B_r(x_0)} \left\langle \mathcal{A}_\varepsilon(Du_\varepsilon), \frac{x - x_0}{|x - x_0|} \right\rangle d\mathcal{H}^{n-1}\, dt$$

$$\leq \frac{c}{r} \int_{t_1}^{t_2} \fint_{\partial B_r(x_0)} g_\varepsilon(Du_\varepsilon)\, d\mathcal{H}^{n-1}\, dt$$

for all $t_0 - r^2 < t_1 \leq t_2 < t_0$, where $\mathcal{H}^{n-1}$ stands for the $(n-1)$-dimensional Hausdorff measure. We thus estimate

$$\operatorname*{osc}_{\tau \in (t_0 - r^2, t_0)} (u_\varepsilon)_{B_r(x_0)}(\tau) = \sup_{t_0 - r^2 < t_1 \leq t_2 < t_0} \left| \fint_{B_r(x_0)} u_\varepsilon(\cdot, \tau)\, dx \Big|_{\tau = t_1}^{t_2} \right|$$

$$\leq \frac{c}{r} \int_{t_0 - r^2}^{t_0} \fint_{\partial B_r(x_0)} g_\varepsilon(Du_\varepsilon)\, d\mathcal{H}^{n-1}\, dt$$

$$\leq c\, r \big(1 + \|Du_\varepsilon\|_{L^\infty(Q_r(x_0, t_0))}\big)^{\widetilde{g}_1 - 1}$$

$$\leq c\, r \big(1 + \|Du_\varepsilon\|_{L^\infty(\mathcal{K}')}\big)^{\widetilde{g}_1 - 1}.$$

Now by Proposition 4.1, in particular by (4.9)-(4.10), we have

$$\operatorname*{osc}_{\tau \in (t_0 - r^2, t_0)} (u_\varepsilon)_{B_r(x_0)}(\tau) \leq c\big(\mathtt{data}, c_\ell, \epsilon, \|\psi\|_{L^\infty}, d\big)\, r. \tag{4.12}$$

At this point we simply split for $(x_1, t_1), (x_2, t_2) \in Q_r(x_0, t_0)$

$$|u_\varepsilon(x_1, t_1) - u_\varepsilon(x_2, t_2)| \leq \left| u_\varepsilon(x_1, t_1) - \fint_{B_r(x_0)} u_\varepsilon(\cdot, t_1)\, dx \right|$$

$$+ \left| \fint_{B_r(x_0)} u_\varepsilon(\cdot, t_1)\, dx - \fint_{B_r(x_0)} u_\varepsilon(\cdot, t_2)\, dx \right| + \left| u_\varepsilon(x_2, t_2) - \fint_{B_r(x_0)} u_\varepsilon(\cdot, t_2)\, dx \right|.$$

While in order to bound the second term we shall use (4.12), the first and last terms can be estimated using the mean value theorem as follows:

$$\left| u_\varepsilon(x_i, t_i) - \fint_{B_r(x_0)} u_\varepsilon(\cdot, t_i)\, dx \right| \leq \fint_{B_r(x_0)} |u_\varepsilon(x_i, t_i) - u_\varepsilon(x, t_i)|\, dx$$

$$\leq 2r\, \|Du_\varepsilon\|_{L^\infty(\mathcal{K}')},$$

for $i \in \{1, 2\}$. Therefore, using again Proposition 4.1, we have

$$\operatorname*{osc}_{Q_r(x_0, t_0)} u_\varepsilon \leq c\, r \tag{4.13}$$

with c as in (4.12), in particular *not depending* on ε. To conclude the proof, for (x_1, t_1), $(x_2, t_2) \in \mathcal{K}$, we simply check whether $\operatorname{dist}_{\mathrm{par}}\big((x_1, t_1), (x_2, t_2)\big) \leq d/4$ holds true or not; if so, then there exists a cylinder $Q_r(x_0, t_0)$ with $r = \operatorname{dist}_{\mathrm{par}}\big((x_1, t_1), (x_2, t_2)\big)$ such that $(x_1, t_1), (x_2, t_2) \in Q_r(x_0, t_0)$ and we can apply (4.13) that directly yields (4.11). If on the other hand $\operatorname{dist}_{\mathrm{par}}\big((x_1, t_1), (x_2, t_2)\big) > d/4$, then, again simply using the maximum principle, we have

$$|u_\varepsilon(x, t) - u_\varepsilon(y, s)| \leq 2\|u_\varepsilon\|_{L^\infty(\Omega_T)} \leq 8\, \frac{\operatorname{dist}_{\mathrm{par}}\big((x_1, t_1), (x_2, t_2)\big)}{d} \|\psi\|_{L^\infty}; \tag{4.14}$$

the proof is concluded. $\qquad\square$

Remark 2. Notice that, tracking the dependence on d of the constant in Proposition 4.2 and in turn the dependence on R of estimate (4.10), and also slightly modifying the previous proof, we deduce that estimate (4.11) can be rewritten as

$$|u_\varepsilon(x,t) - u_\varepsilon(y,s)| \leq \frac{c}{d_{z,w}^\gamma} \operatorname{dist}_{\mathrm{par}}\big((x,t),(y,s)\big), \qquad (4.15)$$

for an exponent $\gamma \equiv \gamma(n, g_1, \epsilon) \geq 1$ and a constant c depending only on $\texttt{data}, \epsilon, c_\ell$, $\|\psi\|_{L^\infty}$, with $z = (x,t), w = (y,s)$ and accordingly

$$d_{z,w} := \min\big\{\operatorname{dist}_{\mathrm{par}}(z, \partial_p\Omega_T), \operatorname{dist}_{\mathrm{par}}(w, \partial_p\Omega_T), 1\big\}.$$

Indeed, if $\operatorname{dist}_{\mathrm{par}}(z,w) \leq d_{z,w}/8$, then we can apply the argument in the first part of the proof of Proposition 4.2 with $r = \operatorname{dist}_{\mathrm{par}}(z,w)$ to get (suppose $s \leq t$)

$$|u_\varepsilon(z) - u_\varepsilon(w)| \leq \operatorname*{osc}_{Q_r(z)} u_\varepsilon \leq \frac{c}{d_{z,w}^\gamma} \operatorname{dist}_{\mathrm{par}}(z,w),$$

where $\gamma = \widetilde{g}_1(\widetilde{g}_1 - 1)\widetilde{\eta}$, since we have $Q_r(z) \subset Q_{d_{z,w}/8}(z)$, $Q_{d_{z,w}/2}(z) \subset \Omega_T$ and so

$$\|Du_\varepsilon\|_{L^\infty(Q_r(z))} \leq \|Du_\varepsilon\|_{L^\infty(Q_{d_{z,w}/8}(z))} \leq \frac{c}{d_{z,w}^{\widetilde{g}_1\widetilde{\eta}}}$$

The case where $d_{z,w} < 8\operatorname{dist}_{\mathrm{par}}(z,w)$ can be approached exactly as in (4.14).

5. CONTINUITY AT THE BOUNDARY

In this section we prove that the solution to the approximating problem (4.3) is continuous up to the boundary independently of ε by building an explicit barrier. We do not want to enter the details of the theory and the general relation between existence of barriers and regularity of the boundary points; the interested reader can see the nice paper [24] for the evolutionary p-Laplacian, while [21, 32] summarize the results in the elliptic setting.

We shall begin with the proof of the continuity at the lateral boundary; here we shall give all the details needed. For the continuity at the initial boundary we shall however only sketch the proof, which on the other hand is very similar and easier than the lateral case. Again, we will prove the existence of a uniform (in the sense that it will be independent of ε) modulus of continuity for u_ε; in the last section we shall show that this modulus is easily inherited by the limit of u_ε.

Let us begin with the construction of an explicit barrier at the lateral boundary. Due to a scaling argument that will be clear soon it is enough to consider a very special case.

5.1. **An explicit construction of a supersolution at the boundary.** We define the function

$$v^+(x,t) := |x'|^2 + M\sqrt{x_n} + (2t+1)_-,$$

where $M \geq 1$ is to be chosen depending on $\texttt{data}$. We aim to show that v^+ is a weak supersolution in

$$\mathcal{Q} := \big\{(x,t) \in \mathbb{R}^{n+1} : |x'| \leq 1,\ x_n \in [0,2],\ t \in [-1,0]\big\}.$$

Simple calculations show that

$$Dv^+ = (2x', Mx_n^{-1/2}/2), \qquad \partial_t v^+ = -2\chi_{\{-1 < t < -1/2\}},$$

$$D^2v^+ = \operatorname{diag}\big(2, \dots, 2, -Mx_n^{-3/2}/4\big),$$

and moreover, since $D_{i,j}^2 v^+ = 0$ whenever $i \neq j$, we have

$$\operatorname{div}\mathcal{A}(Dv^+) = \sum_{i=1}^n D_i\mathcal{A}_i(Dv^+) = \sum_{i,j=1}^n D_{\xi_j}\mathcal{A}_i(Dv^+)D_{i,j}^2 v^+$$

$$= 2\sum_{i=1}^{n-1} D_{\xi_i}\mathcal{A}_i(Dv^+) - \frac{M}{4}D_{\xi_n}\mathcal{A}_n(Dv^+)x_n^{-3/2}.$$

The first term we estimate from above using $(1.9)_2$ and for the second term we can apply $(1.9)_1$, since $D_{\xi_n}\mathcal{A}_n(Dv^+) = \langle D\mathcal{A}(Dv^+)\hat{e}_n, \hat{e}_n\rangle$. Furthermore, if we require $M \geq 2^{3/2}16(n-1)L/\nu$, we obtain

$$
\operatorname{div}\mathcal{A}(Dv^+) \leq \Big(2(n-1)L - \frac{\nu}{4}Mx_n^{-3/2}\Big)\frac{g(|Dv^+|)}{|Dv^+|} \tag{5.1}
$$
$$
\leq -\frac{\nu}{8}Mx_n^{-3/2}\frac{g(|Dv^+|)}{|Dv^+|}.
$$

Now, observe that since $M \geq 4$ we also get

$$
|Dv^+| = \sqrt{4|x'|^2 + \big(Mx_n^{-1/2}/2\big)^2} \leq Mx_n^{-1/2}
$$

in $\mathcal{Q}$. On the other hand, we have

$$
|Dv^+| \geq Mx_n^{-1/2}/2 \geq 1.
$$

Using these estimates we obtain

$$
\frac{g(|Dv^+|)}{|Dv^+|} \geq |Dv^+|^{g_0-2} \geq \begin{cases} 1, & g_0 \geq 2 \\ \big(Mx_n^{-1/2}\big)^{g_0-2} & g_0 < 2 \end{cases},
$$

and thus

$$
\operatorname{div}\mathcal{A}(Dv^+) \leq -\frac{\nu}{8}M^{\min\{g_0,2\}-1}x_n^{-(\min\{g_0,2\}+1)/2}.
$$

The exponent of x_n is negative, so that by choosing $M \equiv M(\texttt{data})$ large enough (recall that $g_0 > 1$), we finally obtain

$$
\partial_t v^+ - \operatorname{div}\mathcal{A}(Dv^+) \geq -2 + \frac{\nu}{8}2^{-(\min\{g_0,2\}+1)/2}M^{\min\{g_0,2\}-1} \geq 0.
$$

It is easy to see that $v^+ \in V_{\mathrm{loc}}^{2,G}(\mathcal{Q})$ and thus v^+ is a (weak) supersolution in $\mathcal{Q}$.

5.2. **A reduction of the oscillation in a significant case.** We set ourselves now in what seems to be a very particular, unitary case; it will be clear soon that, up to a simple rescaling procedure, this will be the significant case for the proof.

Let $\bar{\Omega}$ be a bounded $C^{1,\beta}$ domain and $\bar{\Omega}_T := \bar{\Omega} \times (-1,0)$. Suppose that $0 \in \partial\bar{\Omega}$ and the orthonormal system where the boundary is a graph is the standard cartesian one, with the direction where $\partial\bar{\Omega}$ is a graph given by $\hat{e}_n$. We hence have

$$
\partial\bar{\Omega} \cap \{|x'| < 1, |x_n| < 1\} = \operatorname{graph}\bar{\theta}, \quad \text{with} \quad \bar{\theta}: B_1'(0) \to (-1,1) \text{ and } \bar{\theta}(0) = 0
$$

and $\bar{\Omega} \cap \{|x'| < 1, |x_n| < 1\}$ is the epigraph of $\bar{\theta}$. Let $\bar{u}$ be a weak solution to $(1.8)_1$ in

$$
\bar{\Omega}_T \cap \mathcal{Q}_1 \quad \text{with} \quad \mathcal{Q}_1 := B_1' \times (-1,1) \times (-1,0) \subset \mathbb{R}^{n+1},
$$

such that $\bar{u} = \bar{\psi}$ in $\partial_p\bar{\Omega}_T \cap \mathcal{Q}_1$. Moreover, we suppose $\bar{\psi}(0) = \bar{u}(0) = 0$. Take $\delta \in (0,1)$ to be fixed later. We assume that

$$
\text{the graph of } \bar{\theta} \text{ over } B_1' \text{ is contained in the cylinder } B_1' \times (-\delta,\delta) \tag{5.2}
$$

and moreover that

$$
\operatorname*{osc}_{\partial\bar{\Omega}\cap B_1'\times(-1,1)} \bar{\psi} \leq \delta \quad \text{and} \quad \operatorname*{osc}_{\bar{\Omega}_T\cap\mathcal{Q}_1} \bar{u} \leq 1 \tag{5.3}
$$

Let us take the barrier v_+ built in the previous paragraph and shift it in the $\hat{e}_n$ direction as follows:

$$
v_\delta^+(x', x_n, t) := v^+(x', x_n + \delta, t) + \delta.
$$

Now v_δ^+ is defined and continuous, in particular, over the parabolic closure of $\bar{\Omega}_T \cap \mathcal{Q}_\delta$, where $\mathcal{Q}_\delta = B_1' \times (-\delta,1) \times (-1,0)$, and there it is still a supersolution to an equation structurally similar to $(1.8)_1$. The aim is to prove that $\bar{u} \leq v_\delta^+$ on $\partial_p(\bar{\Omega}_T \cap \mathcal{Q}_\delta)$ by considering the different pieces:

- on $[\partial\bar\Omega \times (-1,0)] \cap \mathcal{Q}_\delta$ we estimate

$$\bar u - v_\delta^+ \le \bar\psi - \delta \le 0$$

using $(5.3)_1$ and since $v^+ \ge 0$;
- on $\left[(\partial B_1' \times [-\delta,1]) \cap \overline{\overline\Omega}\right] \times (-1,0)$ we have

$$\bar u - v_\delta^+ \le 1 - 1 - \delta \le 0,$$

by $(5.3)_2$ together with $\bar u(0) = 0$ and the fact that $v^+ \ge 1$, since $|x'| = 1$;
- on $\overline{B_1'} \times \{1\} \times (-1,0)$ we have

$$\bar u - v_\delta^+ \le 1 - M \le 0, \tag{5.4}$$

since $\bar u \le 1$ as above and on $\{x_n = 1\}$ we have $v_\delta^+ \ge M \ge 1$;
- finally, on $\overline{\overline\Omega \cap (B_1' \times (-\delta,1))} \times \{-1\}$ we again have $v_\delta^+ \ge 1$ due to the expression of the time-dependent part, and therefore the conclusion again follows.

Note that the first three pieces exhaust the lateral boundary of $\bar\Omega_T \cap \mathcal{Q}_\delta$, while the fourth one makes up its initial boundary. Therefore, we have $\bar u \le v_\delta^+$ on the parabolic boundary of $\bar\Omega_T \cap \mathcal{Q}_\delta$ and hence, by Proposition 2.1, $\bar u \le v_\delta^+$ in $\bar\Omega_T \cap \mathcal{Q}_\delta$. Now, if $\delta \le 1/2$, we have

$$v_\delta^+ \le \delta^2 + M(2\delta)^{1/2} \qquad \text{in } [(B_\delta' \times (-\delta,\delta)) \cap \bar\Omega] \times (-\delta,0).$$

Therefore, if we choose δ small enough, depending only on M and so ultimately on data, such that $\delta^2 + M(2\delta)^{1/2} \le 1/4$, then we have

$$\sup_{[(B_\delta' \times (-\delta,\delta)) \cap \bar\Omega] \times (-\delta,0)} \bar u \le \frac{1}{4}.$$

Completely analogously we may consider the subsolution $v_-(x', x_n, t) = -v_+(x', x_n, t)$ to obtain a corresponding bound from below. All in all, we conclude with

$$\operatorname*{osc}_{[B_\delta \cap \bar\Omega] \times (-\delta,0)} \bar u \le \frac{1}{2}. \tag{5.5}$$

5.3. Iteration. Let $R_0 \le \min\{R_\Omega, 1\}$ be fixed and let $Q_r^\omega(x_0, t_0)$ be a cylinder not intersecting the initial boundary, with $x_0 \in \partial\Omega$, $\omega > 0$ and $r \le R_0$. Since we are supposing $R_0 \le R_\Omega$, we have that the boundary of Ω can be written as a $C^{1,\beta}$ graph in B_r: there exists a unitary vector $\hat e \in \mathbb{R}^n$ such that if we set $T : \mathbb{R}^n \to \mathbb{R}^n$ for the orthogonal transformation that maps $\hat e_n = (0, \ldots, 0, 1)$ into $\hat e$, we have

$$T^{-1}(\partial\Omega - x_0) \cap \left(B_r' \times (-r,r)\right) = \operatorname{graph} \theta$$

for some $\theta \in C^{1,\beta}(B_r')$ with values in $(-r,r)$. We now start from the assumption

$$\operatorname*{osc}_{\Omega_T \cap Q_r^\omega(x_0,t_0)} u_\varepsilon \le \omega. \tag{5.6}$$

We define, for $j \in \mathbb{N}$, the quantities

$$\omega_j := 2^{-j}\omega, \qquad r_{j+1} = \min\{\sigma r_j, \bar r_{j+1}\}, \qquad r_0 = r$$

where $\sigma \in (0, 1/2)$ is such that $\sigma \le \frac{\delta}{\sqrt 2}$ and $(2\sigma)^{g_0} \le 4(\sqrt 2)^{-g_1}\delta$ (see (5.9)), with $\delta \in (0, 1/2)$ being the constant defined in the previous paragraph, and $\bar r_j$ is such that

$$\operatorname*{osc}_{\partial_{\mathrm{lat}}\Omega_T \cap Q_{\bar r_j}^{\omega_j}(x_0,t_0)} \psi \le \delta\,\omega_j. \tag{5.7}$$

Note that this is possible, since ψ is continuous so that at ω_j fixed the map $\rho \mapsto \operatorname{osc}_{Q_\rho^{\omega_j}(x_0,t_0)} \psi$ vanishes as $\rho \to 0$. We prove by induction

$$\operatorname*{osc}_{\Omega_T \cap Q_{r_j}^{\omega_j}(x_0,t_0)} u_\varepsilon \le \omega_j. \tag{5.8}$$

Now $(5.8)_0$ is simply (5.6), so we suppose that $(5.8)_j$ holds and we prove $(5.8)_{j+1}$, for $j \in \mathbb{N}_0$. Rescale u_ε as follows:

$$\bar{u}(x,t) := \frac{1}{\omega_j}\, u_\varepsilon\Big(x_0 + \frac{r_j}{\sqrt{2}}Tx, t_0 + \omega_j^2\Big[G\Big(\frac{\sqrt{2}\omega_j}{r_j}\Big)\Big]^{-1}t\Big) - \frac{u_\varepsilon(x_0,t_0)}{\omega_j}.$$

This is a solution to an equation structurally similar to $(1.8)_1$, see Paragraph 2.6, in particular in $[(B_1' \times (-1,1)) \cap \Omega] \times (-1,0)$, with boundary datum

$$\bar{\psi}(x,t) := \frac{1}{\omega_j}\, \psi\Big(x_0 + \frac{r_j}{\sqrt{2}}Tx, t_0 + \omega_j^2\Big[G\Big(\frac{\sqrt{2}\omega_j}{r_j}\Big)\Big]^{-1}t\Big) - \frac{u_\varepsilon(x_0,t_0)}{\omega_j}$$

and where the boundary of $\bar{\Omega} := [\sqrt{2}T^{-1}(\Omega - x_0)/r_j] \cap (B_1' \times (-1,1))$ is given by the graph of the function $\bar{\theta}(x') = \theta(r_j x'/\sqrt{2})/r_j$ over B_1'. We have

$$\operatorname*{osc}_{B_1'} |D\bar{\theta}| = \frac{1}{\sqrt{2}} \operatorname*{osc}_{B_{r_j/\sqrt{2}}'} |D\theta| \leq \frac{1}{\sqrt{2}} \operatorname*{osc}_{B_{R_0}'} |D\theta| \leq \frac{R_0^\beta}{\sqrt{2}}\Theta.$$

Now we choose R_0 small enough so that the right hand side of the chain of inequalities in the above display is smaller than δ, where δ is the quantity fixed in the previous paragraph. This ensures that (5.2) is satisfied (since $D\theta(0) = D\bar{\theta}(0) = 0 = \theta(0) = \bar{\theta}(0)$). Since all the other assumptions in Paragraph 5.2 are satisfied (in particular by our choice of $\bar{r}_j$), we have estimate (5.5) at hand; therefore $(5.8)_{j+1}$ follows by our definition of r_{j+1} and ω_{j+1}. Indeed, scaling back we have

$$\operatorname*{osc}_{\Omega_T \cap \hat{\mathcal{Q}}_{r_j}} u_\varepsilon \leq \frac{1}{2}\omega_j \quad \text{with} \quad \hat{\mathcal{Q}}_{r_j} := B_{\delta r_j/\sqrt{2}}(x_0) \times \Big(t_0 - \delta\,\omega_j^2\Big[G\Big(\frac{\sqrt{2}\omega_j}{r_j}\Big)\Big]^{-1}, t_0\Big)$$

and by (2.10) and our definition of σ, we infer

$$\omega_{j+1}^2\Big[G\Big(\frac{\omega_{j+1}}{r_{j+1}}\Big)\Big]^{-1} \leq \frac{(2\sigma)^{g_0}}{4}\omega_j^2\Big[G\Big(\frac{\omega_j}{r_j}\Big)\Big]^{-1}$$

$$\leq \frac{\delta}{(\sqrt{2})^{g_1}}\omega_j^2\Big[G\Big(\frac{\omega_j}{r_j}\Big)\Big]^{-1} \leq \delta\,\omega_j^2\Big[G\Big(\frac{\sqrt{2}\omega_j}{r_j}\Big)\Big]^{-1}. \quad (5.9)$$

Finally, we note that the lengths of the time intervals also go to zero, that is, the cylinders are shrinking. Indeed, the first inequality in the above computation shows that the ratio of two consecutive time scales is bounded by $(2\sigma)^{g_0}/4$, which is clearly strictly smaller than one.

5.4. Some quantitative estimates. Let us set

$$\omega := 2\|\psi\|_{L^\infty} + 1,$$

fix a radius $r < R_0$, and take a point $(x_0,t_0) \in \partial_{\mathrm{lat}}\Omega_T$ such that $Q^G_{\max\{1,\omega^{2/g_0-1}\}r}(x_0,t_0)$ does not intersect the initial boundary. Clearly (5.6) holds by the maximum principle. Now we recall that ψ has the modulus of continuity ω_ψ:

$$|\psi(x,t) - \psi(y,s)| \leq \omega_\psi\big(\mathrm{dist}_{\mathrm{par},G}((x,t),(y,s))\big)$$

for all $(x,t),(y,s) \in \partial_p \Omega_T$. Since

$$Q_\rho^{\omega_j}(x_0,t_0) \subset B_\rho(x_0) \times (t_0 - [G(2^{-j(1-2/g_1)}\omega^{1-2/g_0}/\rho)]^{-1}, t_0),$$

we have

$$Q_\rho^{\omega_j}(x_0,t_0) \subset Q^G_{A_j\rho}(x_0,t_0),$$

with $A_j := \max\{1, 2^{j(1-2/g_1)}\omega^{2/g_0-1}\} \geq 1$. Thus we see that if we want (5.7) satisfied, it is enough to require

$$A_j\bar{r}_j \leq \max\{1, \omega^{2/g_0-1}\}r,$$

so that $Q^G_{A_j \bar{r}_j}(x_0, t_0)$ does not intersect the initial boundary, and

$$\omega_\psi(\bar{r}_j) \leq \delta\, A_j^{-1} \omega_j = \delta \min\{2^{-j}\omega, 2^{-2j(1-1/g_1)}\omega^{2(1-1/g_0)}\}$$

by the concavity of $\omega_\psi(\cdot)$. At this point we have (5.8) at our disposal, and this will be used noting that in particular we have

$$Q^{\omega_j}_{r_j}(x_0, t_0) \supset Q^G_{r_j/B_j}(x_0, t_0)$$

with $B_j := \max\{1, 2^{-j(1-2/g_0)}\omega^{1-2/g_1}\} \geq 1$. Hence, for $(x, t) \in \Omega_T \cap Q^G_{r/B_0}(x_0, t_0)$ fixed we find the largest $j \in \mathbb{N}_0$ such that

$$\frac{r_{j+1}}{B_{j+1}} \leq \operatorname{dist}_{\mathrm{par},G}((x, t), (x_0, t_0)) < \frac{r_j}{B_j}.$$

Note that this is possible, since clearly $r_j/B_j \leq r_j \to 0$ as $j \to \infty$. At this point

$$|u_\varepsilon(x, t) - u_\varepsilon(x_0, t_0)| \leq \operatorname*{osc}_{\Omega_T \cap Q^{\omega_j}_{r_j}(x_0, t_0)} u_\varepsilon \leq 2^{-j}\omega.$$

Let $\{(r_{j+1}/B_{j+1}, 2^{-j}\omega)\}_{j\in\mathbb{N}_0}$ be a sequence of points in $\mathbb{R}^2$ and call ω_u the smallest concave function such that $\omega_u(r_{j+1}/B_{j+1}) \geq 2^{-j}\omega$; note that ω_u is a modulus of continuity. For instance, one can take the piecewise linear interpolation of the sequence $\{(x_j, y_j)\}_{j\in\mathbb{N}}$ given by $x_j = \max_{k\geq j+1} r_k/B_k, y_j = 2^{-j}\omega$, which is component-wise decreasing as j increases. This finally leads to

$$|u_\varepsilon(x, t) - u_\varepsilon(x_0, t_0)| \leq 2^{-j}\omega \leq \omega_u(r_{j+1}/B_{j+1}) \leq \omega_u\big(\operatorname{dist}_{\mathrm{par},G}((x, t), (x_0, t_0))\big),$$
(5.10)

and this holds for $(x, t) \in \Omega_T \cap Q^G_{r/B_0}(x_0, t_0)$. In fact, it also holds for points (x, t) outside $Q^G_{r/B_0}(x_0, t_0)$, since then we have $\operatorname{dist}_{\mathrm{par},G}((x, t), (x_0, t_0)) > r/B_0$ and thus

$$|u_\varepsilon(x, t) - u_\varepsilon(x_0, t_0)| \leq \frac{2B_0}{r}\|\psi\|_{L^\infty}\frac{r}{B_0} \leq c\operatorname{dist}_{\mathrm{par},G}((x, t), (x_0, t_0))$$

by the maximum principle. Note that the modulus of continuity ω_u at this point depends on $\mathtt{data}, \|\psi\|_{L^\infty}, \omega_\psi$ but also on r.

If now ψ is γ-Hölder continuous with respect to the G-parabolic metric, then we see that it is enough to take $\bar{r}_j = c(\mathtt{data}, \omega, \gamma)2^{-\eta j}r$ for some $\eta \equiv \eta(g_1, \gamma)$. This yields that the numbers r_j can be written as $\bar{\eta}^j r$ for some $\bar{\eta} \in (0, 1)$. Now the Hölder continuity follows, for instance, similarly to [14, Chapter III, Lemma 3.1].

5.5. **Continuity at the initial boundary.** We begin by modifying the barrier built in Paragraph 5.1 to meet the different situations at the initial boundary. We start by considering the case where, before rescaling, we have a solution in a cylinder $B_r(x_0) \times (0, \omega^2/G(\omega/r))$, with $B_r(x_0) \subset \Omega$, equal to $\bar{\psi}$ over $B_r(x_0) \times \{0\}$; that is, the true case of initial boundary continuity. Later on we shall face the "corner case", that is the case of cylinders $B_r(x_0) \times (0, \omega^2/G(\omega/r))$ with $x_0 \in \partial\Omega$.

After rescaling, one sees that it is enough to build a supersolution in $\mathcal{Q} := B_1 \times (0, 1)$. In this case the explicit expression is simply $v^+(x, t) := |x|^{1/2}$. We then have $v^+ \in V^{2,G}_{\mathrm{loc}}(\mathcal{Q})$ and v^+ is a supersolution to $(1.8)_1$ in $\mathcal{Q}$. Moreover, if we further suppose that

$$\operatorname*{osc}_{B_1 \times \{0\}} \bar{\psi} \leq \delta, \qquad \operatorname*{osc}_{\mathcal{Q}} \bar{u} \leq 1, \qquad \bar{u}(0, 0) = \bar{\psi}(0, 0) = 0$$

for some $\delta \in (0, 1)$, it is easy to see that $\bar{u} \leq v^+ + \delta$ on $\partial_p\mathcal{Q}$. Indeed on $\partial B_1 \times (0, 1)$ we have $v^+ = 1$ but $\bar{u} \leq 1$, while on $\overline{B_1} \times \{0\}$ we have $\bar{u} = \bar{\psi} \leq \delta$ and $v^+ \geq 0$. Therefore we can deduce by Proposition 2.1 that $\bar{u} \leq v^+ + \delta$ in $\mathcal{Q}$. Now the proof goes on similarly as in Paragraphs 5.2 to 5.4, with possibly new constants δ and R_0.

For the "corner situation", we are lead to consider a solution in a domain of the type $\mathcal{Q} := B'_1 \times (-1, 1) \times (0, 1)$; the supersolution in this case is $v^+ = |x'|^2 + Mx_n^{1/2}$, with

M as in Paragraph 5.1. The fact that the function is a supersolution follows plainly from (5.1). Assuming now that the boundary graph $\bar{\theta}$ over B_1' takes values in $(-\delta, \delta)$ and

$$\operatorname*{osc}_{\partial_p \bar{\Omega}_T \cap \mathcal{Q}} \bar{\psi} \leq \delta, \quad \operatorname*{osc}_{\mathcal{Q}} \bar{u} \leq 1, \quad \bar{u}(0,0) = \bar{\psi}(0,0) = 0,$$

we have $\bar{u} \leq v_\delta^+$ in $\partial_p \bar{\Omega}_T \cap \mathcal{Q}$, since $\bar{u} = \bar{\psi} \leq \delta$ there; on the remaining part of the parabolic boundary of $\mathcal{Q}$ we use the fact that v_δ^+ is larger than one, as in (5.4). Again, now the proof is similar as above.

In both cases, a scaling and iteration procedure like the one used in Paragraph 5.4 allows us to prove the reduction of oscillation in a sequence of nested cylinders of the type $(\Omega \cap B_{r_j}(x_0)) \times (0, \omega_j^2 G(\omega_j/r_j))$, with $x_0 \in \overline{\Omega}$. This leads to

$$|u_\varepsilon(x, t) - u_\varepsilon(x_0, 0)| \leq \omega_u\big(\operatorname{dist}_{\mathrm{par},G}((x,t),(x_0,0))\big). \tag{5.11}$$

for every $(x, t) \in \Omega_T$. Moreover, opportune statements similar to above still hold in the case ψ is Hölder continuous.

At this point we call $\widetilde{R}$ the smallest value of R_0 coming from the three different cases, ultimately a constant depending on `data` and $\partial\Omega$. Choose $r = \widetilde{R}/2$. Now r is a constant depending only on `data` and $\partial\Omega$. By combining the boundary estimates (5.10) and (5.11) with the interior estimate in Remark 2 we obtain

$$|u_\varepsilon(x, t) - u_\varepsilon(y, s)| \leq \omega_u\big(\operatorname{dist}_{\mathrm{par},G}((x,t),(y,s))\big) \tag{5.12}$$

for every $(x, t), (y, s) \in \overline{\Omega_T}^p$, where ω_u depends on `data`, $\epsilon, c_\ell, \omega_\psi, \|\psi\|_{L^\infty}, \partial\Omega$. Indeed, if one of the points is in $\partial_p\Omega_T$, then (5.12) is either (5.10) or (5.11). In the case where both $(x, t), (y, s) \in \Omega_T$ we consider two different cases. Either the mutual distance of (x, t) and (y, s) is small compared to their distance to the boundary, in which case we use the interior estimate, or otherwise we can again use the boundary estimates.

Let us make this rigorous. Denote $z = (x, t), w = (y, s)$. If $\operatorname{dist}_{\mathrm{par},G}(z, w) \geq 1$, we are done by the maximum principle. Note now that if $\operatorname{dist}_{\mathrm{par},G}(z, w) \leq 1$, we have

$$\operatorname{dist}_{\mathrm{par},G}(z, w)^{\max\{1, g_1/2\}} \leq \operatorname{dist}_{\mathrm{par}}(z, w) \leq \operatorname{dist}_{\mathrm{par},G}(z, w)^{\min\{1, g_0/2\}}.$$

Observe that (4.15) can be written in terms of the parabolic G-distance as follows:

$$|u_\varepsilon(z) - u_\varepsilon(w)| \leq c \left[d_{z,w}^G\right]^{-\gamma \max\{1, g_1/2\}} \operatorname{dist}_{\mathrm{par},G}(z, w)^{\min\{1, g_0/2\}},$$

where

$$d_{z,w}^G := \min\big\{\operatorname{dist}_{\mathrm{par},G}(z, \partial_p\Omega_T), \operatorname{dist}_{\mathrm{par},G}(w, \partial_p\Omega_T), 1\big\}.$$

If now $\operatorname{dist}_{\mathrm{par},G}(z, w) \leq \left[d_{z,w}^G\right]^{2\gamma \max\{1, g_1/2\}}$, that is, the mutual distance of z and w is small compared to their distance to the boundary, then we have

$$|u_\varepsilon(z) - u_\varepsilon(w)| \leq c \operatorname{dist}_{\mathrm{par},G}(z, w)^{\min\{1, g_0-1\}/2}.$$

On the other hand, when $\operatorname{dist}_{\mathrm{par},G}(z, w) > \left[d_{z,w}^G\right]^{2\gamma \max\{1, g_1/2\}}$ there exists a cylinder $Q_\rho^G(x_0, t_0) \ni z, w$ with $\rho = 2\operatorname{dist}_{\mathrm{par},G}(z, w)^{\min\{1/2, 1/g_1\}/\gamma}$ such that either $x_0 \in \partial\Omega$ or the bottom of $Q_\rho^G(x_0, t_0)$ touches the initial boundary. Now using triangle inequality and the boundary estimates yields

$$|u_\varepsilon(z) - u_\varepsilon(w)| \leq 2\,\omega_u(\rho) \leq 4\,\omega_u\big(\operatorname{dist}_{\mathrm{par},G}(z, w)^{\min\{1/2, 1/g_1\}/\gamma}\big).$$

Finally, we take the largest modulus of continuity ω_u for which all the conditions proved above are satisfied, and this proves (5.12). The proof in the Hölder case is similar, since in this case we can quantify all the moduli.

6. CONCLUSION

Call $u_j := u_\varepsilon$ for $\varepsilon = 1/j$, $j \in \mathbb{N}$, and similarly $\mathcal{A}_j, g_j, \phi_j$. From the results of the preceding section, that is, from the equi-boundedness of the sequence $\{u_j\}_{j\in\mathbb{N}}$ following from the maximum principle Corollary 2.2 and the global equi-continuity coming from the results of Sections 4 and 5, using Ascoli-Arzelà theorem we see that $u_j \to u$ uniformly in $C^0(\overline{\Omega_T}^p)$ for some $u \in C^0(\overline{\Omega_T}^p)$. Now all we have left to prove is that u is a weak solution to $(1.8)_1$, which follows easily from the next proposition.

Proposition 6.1. *Let $u_j \in V_{\mathrm{loc}}^{2,G}(\Omega_T) \cap C^0(\overline{\Omega_T}^p)$ be the solutions to (4.3) defined above. Suppose there exists a function u such that $u_j \to u$ almost everywhere in Ω_T. Then $Du_j \to Du$ almost everywhere.*

Proof. Take $\mathcal{K} \Subset \Omega_T$ and choose a cutoff function $\varphi \in C_c^\infty(\Omega_T)$ such that $0 \leq \varphi \leq 1$, $\varphi = 1$ in $\mathcal{K}$, and $\|\partial_t \varphi\|_{L^\infty(\Omega_T)}, \|D\varphi\|_{L^\infty(\Omega_T)} \leq c$ for some $c \geq 1$ depending on $\mathrm{dist}(\mathcal{K}, \partial_p \Omega_T)$. Let $j, k \in \mathbb{N}$ and test the weak formulations of u_j and u_k with $\eta = w_{j,k}\varphi$, where $w_{j,k} := u_j - u_k$. This choice can be justified by standard methods such as Steklov averages. By subtracting we obtain

$$0 = -\int_{\Omega_T} w_{j,k}\partial_t(w_{j,k}\varphi)\,dx\,dt + \int_{\Omega_T} \langle \mathcal{A}_j(Du_j) - \mathcal{A}_k(Du_k), D(w_{j,k}\varphi)\rangle\,dx\,dt$$

$$= -\frac{1}{2}\int_{\Omega_T} w_{j,k}^2 \partial_t\varphi\,dx\,dt + \int_{\Omega_T} \langle \mathcal{A}_j(Du_j) - \mathcal{A}_k(Du_k), D\varphi\rangle w_{j,k}\,dx\,dt$$

$$+ \int_{\Omega_T} \langle \mathcal{A}_j(Du_j) - \mathcal{A}(Du_j), Dw_{j,k}\rangle\varphi\,dx\,dt$$

$$+ \int_{\Omega_T} \langle \mathcal{A}(Du_j) - \mathcal{A}(Du_k), Dw_{j,k}\rangle\varphi\,dx\,dt$$

$$+ \int_{\Omega_T} \langle \mathcal{A}(Du_k) - \mathcal{A}_k(Du_k), Dw_{j,k}\rangle\varphi\,dx\,dt =: I + II + III + IV + V.$$

Since $\|Du_j\|_{L^\infty(\Omega_T)} \leq c$ uniformly with respect to j by Proposition 4.1, we also have

$$\|\mathcal{A}_j(Du_j)\|_{L^\infty(\Omega_T)} \leq cg_j\big(\|Du_j\|_{L^\infty(\Omega_T)}\big) \leq c.$$

Thus, by the definition of $\mathcal{A}_j$

$$|I + II + III + V| \leq c\,\|u_j - u_k\|_{L^2(\Omega_T)} + c\|(\phi_j * \mathcal{A})(Du_j) - \mathcal{A}(Du_j)\|_{L^2(\Omega_T)}$$

$$+ c\|\mathcal{A}(Du_k) - (\phi_k * \mathcal{A})(Du_k)\|_{L^2(\Omega_T)} + c\,(1/j + 1/k).$$

The first term on the right-hand side tends to zero as $j, k \to \infty$ by Lebesgue's dominated convergence theorem and the second and third by the properties of mollifiers; the last one is obvious. On the other hand, by (2.13)

$$IV \geq c \int_{\mathcal{K}} |V_g(Du_j) - V_g(Du_k)|^2\,dx\,dt.$$

Thus

$$c \int_{\mathcal{K}} |V_g(Du_j) - V_g(Du_k)|^2\,dx\,dt \leq IV \leq |I + II + III + V| \to 0$$

as $j, k \to \infty$. We have shown that the sequence $\{V_g(Du_j)\}_{j\in\mathbb{N}}$ is Cauchy in $L^2(\mathcal{K})$ and therefore there exists a function $w \in L^2(\mathcal{K})$ such that $V_g(Du_j) \to w$ in $L^2(\mathcal{K})$ as $j \to \infty$. This implies that there exists a (nonrelabeled) subsequence $V_g(Du_j)$ converging to w almost everywhere in $\mathcal{K}$. Now the fact that V_g has a continuous inverse yields

$$Du_j = V_g^{-1}(V_g(Du_j)) \to V_g^{-1}(w) =: v$$

almost everywhere in $\mathcal{K}$.

Now, since $u_j \to u$ almost everywhere in $\mathcal{K}$, we have for any $\phi \in C_c^\infty(\mathcal{K})$ that

$$\int_\mathcal{K} u D\phi \, dx \, dt = \lim_{j\to\infty} \int_\mathcal{K} u_j D\phi \, dx \, dt = -\lim_{j\to\infty} \int_\mathcal{K} Du_j \, \phi \, dx \, dt = -\int_\mathcal{K} v\phi \, dx \, dt$$

by Lebesgue's dominated convergence theorem and the definition of weak gradient, showing that $v = Du$. Thus, we have $Du_j \to Du$ almost everywhere in $\mathcal{K}$ for any $\mathcal{K} \Subset \Omega_T$, which implies that $Du_j \to Du$ almost everywhere in Ω_T. $\square$

To conclude, (1.10) follows from (4.9) simply using the local almost everywhere convergence of Du_ε, and (1.11) follows from (5.12) using the global uniform convergence of u_ε.

6.1. **Weakening the assumptions.**
As mentioned in Remark 1, in this paragraph we show how to modify the proofs of the paper in order to obtain Theorems 1.1 and 1.2 for vector fields satisfying the weaker assumptions (1.12).

We observe that assumptions (1.9) are only used in order to have the analogous properties for the regularized vector field $\mathcal{A}_\varepsilon$ defined in (4.1). Moreover, (2.11) and (2.12) trivially hold by taking $\xi_2 = 0$ in (1.12). Thus, it suffices to show that under the assumptions (1.12) we still have (1.9) for $\mathcal{A}_\varepsilon$ with g replaced by g_ε defined in (4.2).

We shall focus only on the convolution part of the vector field $\mathcal{A}_\varepsilon$, since for the part involving the nondegenerate $\widetilde{g}_1$-Laplacian the corresponding estimates are classic and easy to verify. Therefore, we only need to prove (1.9) with $\mathcal{A}$ replaced by $\phi_\varepsilon * \mathcal{A}$ and $g(s)$ replaced by $\frac{g(s+\varepsilon)}{s+\varepsilon}s$. Using (1.12) we have

$$\langle D(\phi_\varepsilon * \mathcal{A})(\xi)\lambda, \lambda \rangle = \lim_{h\to 0} \frac{1}{h} \int_{B_\varepsilon(0)} \langle \mathcal{A}(\xi + h\lambda - \eta) - \mathcal{A}(\xi - \eta), \lambda \rangle \phi_\varepsilon(\eta) \, d\eta$$

$$\geq 2^{g_0 - 2}\nu \int_{B_1(0)} \frac{g(|\xi - \varepsilon\eta|)}{|\xi - \varepsilon\eta|} \phi(\eta) \, d\eta \, |\lambda|^2$$

and

$$|D_j(\phi_\varepsilon * \mathcal{A})(\xi)| = \lim_{h\to 0} \frac{1}{h} \left| \int_{B_\varepsilon(0)} (\mathcal{A}(\xi + h\hat{e}_j - \eta) - \mathcal{A}(\xi - \eta))\phi_\varepsilon(\eta) \, d\eta \right|$$

$$\leq 2^{g_1 - 2}L \int_{B_1(0)} \frac{g(|\xi - \varepsilon\eta|)}{|\xi - \varepsilon\eta|} \phi(\eta) \, d\eta.$$

Hence, if we can show that

$$\int_{B_1(0)} \frac{g(|\xi - \varepsilon\eta|)}{|\xi - \varepsilon\eta|} \phi(\eta) \, d\eta \approx \frac{g(|\xi| + \varepsilon)}{|\xi| + \varepsilon} \tag{6.1}$$

independently of ε, we are done.

Consider first the case $|\xi| \geq 2\varepsilon$. This implies $|\xi - \varepsilon\eta| \geq \frac{1}{3}(|\xi| + \varepsilon)$, and thus $|\xi - \varepsilon\eta| \approx |\xi| + \varepsilon$ so that (6.1) holds. On the other hand, if $|\xi| < 2\varepsilon$, we have

$$\int_{B_1(0)} \frac{g(|\xi - \varepsilon\eta|)}{|\xi - \varepsilon\eta|} \phi(\eta) \, d\eta \leq \sup_{B_1(0)} \phi \int_{B_3(\xi/\varepsilon)} \frac{1}{|\xi/\varepsilon - \eta|} \, d\eta \, \frac{g(|\xi| + \varepsilon)}{\varepsilon}$$

$$\leq c(n) \sup_{B_1(0)} \phi \, \frac{g(|\xi| + \varepsilon)}{|\xi| + \varepsilon}$$

and

$$\int_{B_1(0)} \frac{g(|\xi - \varepsilon\eta|)}{|\xi - \varepsilon\eta|} \phi(\eta) \, d\eta \geq \int_{B_{1/2}(0)\setminus B_{1/4}(\xi/\varepsilon)} \frac{g(|\xi - \varepsilon\eta|)}{|\xi - \varepsilon\eta|} \phi(\eta) \, d\eta$$

$$\geq \inf_{B_{1/2}(0)} \phi \, |B_{1/2}(0) \setminus B_{1/4}(\xi/\varepsilon)| \, \frac{g(\varepsilon/4)}{|\xi| + \varepsilon}$$

$$\geq c(n, g_1) \inf_{B_{1/2}(0)} \phi \, \frac{g(|\xi| + \varepsilon)}{|\xi| + \varepsilon}.$$

Note that we can assume without loss of generality that $\sup_{B_1(0)} \phi \leq c$ and $\inf_{B_{1/2}(0)} \phi \geq 1/c$ for some $c \equiv c(n) > 0$.

References

[1] E. ACERBI, G. MINGIONE: Gradient estimates for a class of parabolic systems, *Duke Math. J.* **136** (2007), no. 2, pp. 285–320.

[2] E. ACERBI, G. MINGIONE, G. A. SEREGIN: Regularity results for parabolic systems related to a class of non-newtonian fluids, *Ann. Inst. H. Poincaré Anal. Non Linéaire* **21** (2004), no. 1, pp. 25–60.

[3] P. BARONI: Riesz potential estimates for a general class of quasilinear equations, *Calc. Var. Partial Differential Equations* **53** (2015), no. 3-4, pp. 803–846.

[4] L. BERS: *Mathematical aspects of subsonic and transonic gas dynamics*, Surveys in Applied Mathematics, 3, John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London 1958.

[5] J. BURCZAK, P. KAPLICKÝ: Interior regularity of space derivatives to an evolutionary, symmetric φ-Laplacian, *preprint*, http://arxiv.org/abs/1507.05843v1

[6] Y.Z. CHEN, E. DIBENEDETTO: Hölder estimates of solutions of singular parabolic equations with measurable coefficients, *Arch. Rational Mech. Anal.* **118** (1992), no. 3, pp. 257–271.

[7] A. CIANCHI: Boundedness of solutions to variational problems under general growth conditions, *Comm. Partial Differential Equations* **22** (1997), no. 9-10, pp. 1629–1646.

[8] A. CIANCHI, N. FUSCO: Gradient regularity for minimizers under general growth conditions, *J. Reine Angew. Math. (Crelle's J.)* **507** (1999), pp. 15–36.

[9] A. CIANCHI, V. MAZ'YA: Gradient regularity via rearrangements for p-Laplacian type elliptic boundary value problems. *J. Eur. Math. Soc. (JEMS)* **16** (2014), no. 3, pp. 571–595.

[10] B.E.J. DAHLBERG, C.E. KENIG: Non-negative solution of generalized porous medium equations, *Rev. Mat. Iberoamericana* **2 (3)** (1986), pp. 267–305.

[11] B.E.J. DAHLBERG, C.E. KENIG: Nonnegative solutions of the initial-Dirichlet problem for generalized porous medium equations in cylinders, *J. Amer. Math. Soc.* **1** (1988), no. 2, pp. 401–412.

[12] P. DASKALOPOULOS, C. KENIG: *Degenerate diffusions. Initial value problems and local regularity theory.* EMS Tracts in Mathematics, 1. Zürich, 2007.

[13] E. DIBENEDETTO: On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients, *Ann. Sc. Norm. Sup. Pisa Cl. Sci.* (IV) **13** (1986), no. 3, pp. 487–535.

[14] E. DIBENEDETTO: *Degenerate parabolic equations*, Universitext, Springer, New York, 1993.

[15] E. DIBENEDETTO, U. GIANAZZA, V. VESPRI: Harnack estimates for quasi-linear degenerate parabolic differential equations, *Acta Math.* **200** (2008), no. 2, 181–209.

[16] L. DIENING, F. ETTWEIN: Fractional estimates for non-differentiable elliptic systems with general growth, *Forum Math.* **20** (2008), pp. 523–556.

[17] L. DIENING, B. STROFFOLINI, A. VERDE: Everywhere regularity of functionals with φ-growth, *Manuscripta Math.* **129** (2009), no. 4, pp. 449–481.

[18] G. C. DONG: *Nonlinear partial differential equations of second order*, Translations of Mathematical Monographs, 95. American Mathematical Society, Providence, RI, 1991.

[19] R. FINN, D. GILBARG: Three-dimensional subsonic flows, and asymptotic estimates for elliptic partial differential equations, *Acta Math.* **98** (1957), pp. 265–296.

[20] N. FUSCO, C. SBORDONE: Higher integrability of the gradient of minimizers of functionals with nonstandard growth conditions, *Comm. Pure Appl. Math.* **43** (1990), no. 5, pp. 673–683.

[21] J. HEINONEN, T. KILPELÄINEN, O. MARTIO: *Nonlinear potential theory of degenerate elliptic equations*, Oxford Mathematical Monographs, New York, 1993.

[22] S. HWANG: Hölder regularity of solutions of generalized p-Laplacian type parabolic equations, PhD thesis (Paper 12667), Iowa State University, 2012.

[23] S. HWANG, G. M. LIEBERMAN: Hölder continuity of a bounded weak solution of generalized parabolic *p*-Laplacian equations, *preprint*, http://arxiv.org/abs/1407.0531

[24] T. KILPELÄINEN, P. LINDQVIST: On the Dirichlet boundary value problem for a degenerate parabolic equation, *SIAM J. Math. Anal.* **27** (1996), no. 3, pp. 661–683.

[25] T. KUUSI, G. MINGIONE: New perturbation methods for nonlinear parabolic problems, *J. Math. Pures Appl.* (9) **98** (2012), no. 4, pp. 390–427.

[26] T. KUUSI: Harnack estimates for weak supersolutions to nonlinear degenerate parabolic equations, *Ann. Sc. Norm. Super. Pisa Cl. Sci.* (5) **7** (2008), no. 4, 673–716.

[27] T. KUUSI, G. MINGIONE: Potential estimates and gradient boundedness for nonlinear parabolic systems, *Rev. Mat. Iberoamericana* **28** (2012), no. 2, pp. 535–576.

[28] G. M. LIEBERMAN: The natural generalization of the natural conditions of Ladyzhenskaya and Ural'tseva for elliptic equations, *Comm. Partial Differential Equations* **16** (1991), pp. 311–361.

[29] G. M. LIEBERMAN: Boundary and initial regularity for solutions of degenerate parabolic equations, *Nonlinear Anal.* **20** (1993), no. 5, pp. 551–569.

[30] G. M. LIEBERMAN: Hölder regularity for the gradients of solutions of degenerate parabolic systems, *Nelineĭn. Granichnye Zadachi* **16** (2006), pp. 69–85.

[31] C. LINDFORS: Obstacle problem for a class of parabolic equations of generalized *p*-Laplacian type, preprint.

[32] P. LINDQVIST: *Notes on the p-Laplace equation*, Report 102, University of Jyväskylä, 2006.

[33] M.M. RAO, Z.D. REN: *Theory of Orlicz spaces*, Marcel Dekker Inc., 1991.

[34] M. SHIFFMAN: On the existence of subsonic flows of a compressible fluid, *Proc. Nat. Acad. Sci. U.S.A.* **38**, (1952), pp. 434–438.

[35] J.J. VÁZQUEZ: *The Porous Medium Equation, Mathematical Theory*, Oxford Mathematical Monographs. The Caledon Press, Oxford University Press, Oxford, 2007.

PAOLO BARONI, DIPARTIMENTO DI MATEMATICA E APPLICAZIONI "R. CACCIOPPOLI", UNIVERSITÀ DEGLI STUDI DI NAPOLI "FEDERICO II", I-80125 NAPOLI, ITALY
    E-mail address: paolo.baroni@unina.it

DEPARTMENT OF MATHEMATICS AND SYSTEMS ANALYSIS, AALTO UNIVERSITY, P.O. BOX 11000, 00076 AALTO, FINLAND
    E-mail address: casimir.lindfors@aalto.fi