



Gradient bounds for a widely degenerate orthotropic parabolic equation

To Cesca, nothing else to say. She knows.

Pasquale Ambrosio

Abstract. In this paper, we consider the following nonlinear parabolic equation

$$\partial_t u = \sum_{i=1}^n \partial_{x_i} \left[(|u_{x_i}| - \delta_i)_+^{p-1} \frac{u_{x_i}}{|u_{x_i}|} \right] \quad \text{in } \Omega \times I,$$

where Ω is a bounded open subset of $\mathbb{R}^n$ for $n \geq 2$, $I \subset \mathbb{R}$ is a bounded open interval, $p \geq 2$, $\delta_1, \dots, \delta_n$ are non-negative numbers and $(\cdot)_+$ denotes the positive part. We prove that the local weak solutions are locally Lipschitz continuous in the spatial variable. The main novelty here is that the above equation combines an orthotropic structure with a strongly degenerate behavior. We emphasize that our result can be considered, on the one hand, as the parabolic counterpart of the elliptic result established in [12], and on the other hand as an extension to a significantly more degenerate framework of the findings contained in [13].

Mathematics Subject Classification. 35B45, 35B65, 35K10, 35K65, 35K92.

Keywords. Degenerate parabolic equations, Anisotropic equations, Lipschitz continuity, Moser iteration.

1. Introduction

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set and $I \subset \mathbb{R}$ a bounded open interval. We are interested in the gradient regularity of the local weak solutions to the following parabolic equation

$$\partial_t u = \sum_{i=1}^n \partial_{x_i} \left[(|u_{x_i}| - \delta_i)_+^{p-1} \frac{u_{x_i}}{|u_{x_i}|} \right] \quad \text{in } \Omega \times I, \quad (1.1)$$

where $p \geq 2$, $\delta_1, \dots, \delta_n$ are non-negative numbers and $(\cdot)_+$ stands for the positive part. Throughout the paper, we denote by $T_0 < T_1$ the endpoints of the time interval I .

Evolutionary equations of the above form have been studied since the 1960s, notably by the Soviet school; see, for instance, the work [33] by Vishik. Equation (1.1) with all δ_i set to zero is also explicitly presented in the monographs [28], [30, Example 4.A, Chapter III] and [35, Example 30.8], among others.

At first glance, (1.1) looks quite similar to the parabolic p -Laplace equation

$$\partial_t u = \sum_{i=1}^n (|Du|^{p-2} u_{x_i})_{x_i} \quad \text{in } \Omega \times I. \quad (1.2)$$

However, the main novelty of Eq. (1.1) lies in the combination of two features, namely an *orthotropic structure* and a *strongly degenerate behavior*. Indeed, unlike the parabolic p -Laplace equation, for which the loss of ellipticity of the operator $\operatorname{div}(|Du|^{p-2} Du)$ is restricted to a single point, Eq. (1.1) becomes degenerate on the larger set

$$\bigcup_{i=1}^n \{|u_{x_i}| \leq \delta_i\}.$$

A more recent work in which Eq. (1.1) appears with all δ_i equal to zero is [13]. There, the authors derive local L^∞ bounds for the spatial gradient Du of local weak solutions to (1.1), but confining their analysis to the case $p \geq 2$ and $\max\{\delta_i\} = 0$. In this special case, as already observed in [13], the basic regularity theory equally applies to both (1.1) and (1.2). A classical reference in the field is DiBenedetto's monograph [21], which provides boundedness results for the solution u (see [21, Chapter V]), Hölder continuity estimates for u (see [21, Chapter III]), as well as Harnack inequalities for non-negative solutions (see [21, Chapter VI]). From a technical point of view, there is no distinction to be made between (1.2) and (1.1) with all δ_i set to zero. Consequently, the results in [13] and [21, Chapter V] imply that, for $p \geq 2$ and $\max\{\delta_i\} = 0$, the local weak solutions of (1.1) are locally Lipschitz continuous in the spatial variable.

Concerning the gradient regularity of weak solutions to Eq. (1.2), we refer again to DiBenedetto's book for a comprehensive account of results on the subject, specifically to [21, Chapter VIII]. Since then, the literature on the regularity for nonlinear, possibly degenerate or singular, parabolic equations (or systems) has been steadily expanding, with the evolutionary p -Laplace Eq. (1.2) serving as a prototypical model. Without any claim of exhaustiveness, we can mention a few classical references [17, 18, 20, 22, 23, 34], up to more recent contributions on the subject, such as [8, 26, 27], among others.

However, none of these results apply to Eq. (1.1), as they all rely on the fact that the loss of ellipticity of the operator in divergence form is restricted to a single point, as in the model case (1.2). As previously noted, such a property dramatically fails for our Eq. (1.1). Therefore, the aforementioned

references do not provide any regularity results for the spatial gradient Du of its solutions.

In [4], we have recently proved that local weak solutions of (1.1) are locally bounded also in the case $p \geq 2$ and $\max \{\delta_i\} > 0$, thus extending DiBenedetto's result [21, Chapter V, Theorem 4.1] to our anisotropic and more degenerate setting.

The primary goal of this paper is to establish a local L^∞ bound on Du for our Eq. (1.1), in order to extend the result of [13] to the markedly more degenerate case $\max \{\delta_i\} > 0$. To this end, we will need to adapt the techniques developed in [9–14, 16] for degenerate equations with orthotropic structure (see also [29], which deals with the higher differentiability of minimizers for non-autonomous orthotropic functionals). Indeed, the main result of this work is the following theorem, which can be considered as the parabolic counterpart of the elliptic result [12, Theorem 1.1]. For notation and definitions we refer to Sect. 2.

Theorem 1.1. *Let $n \geq 2$ and $p \geq 2$. Moreover, assume that $u \in L^p_{loc}(I; W^{1,p}_{loc}(\Omega))$ is a local weak solution of Eq. (1.1). Then*

$$Du \in L^\infty_{loc}(\Omega \times I, \mathbb{R}^n).$$

More precisely, there exists a constant $C > 1$, depending only on n , p and $\max \{\delta_1, \dots, \delta_n\}$, such that for every parabolic cylinder $Q_r(x_0, t_0) \subset Q_R(x_0, t_0) \Subset \Omega \times I$ with $R \in (0, 1]$, we have

$$\|Du\|_{L^\infty(Q_r(x_0, t_0))} \leq \frac{C}{(R-r)^{\vartheta p}} \left[1 + \left(\iint_{Q_R(x_0, t_0)} |Du|^p dx dt \right)^{\frac{1}{2}} \right], \quad (1.3)$$

where

$$\vartheta = \begin{cases} \frac{n+2}{2} & \text{if } n \geq 3, \\ \text{any number} > 2 & \text{if } n = 2. \end{cases}$$

Remark 1.2. We explicitly note that for $p \geq 2$, the local weak solutions of (1.1) are locally Lipschitz continuous in the spatial variable. Indeed, Theorem 1.1 provides a local L^∞ estimate for the spatial gradient Du . In order to interpret this estimate as a genuine Lipschitz continuity property for u , one also needs the local boundedness of the solution itself. This is ensured by [4, Theorem 1.1], which holds under the same assumptions on p and $\delta_1, \dots, \delta_n$ considered here. We also observe that, as recalled in (2.4), any local weak solution $u \in L^p_{loc}(I; W^{1,p}_{loc}(\Omega))$ of (1.1) belongs to $C^0(\bar{J}; L^2(\Omega'))$ for every subinterval $J \Subset I$ and every open set $\Omega' \Subset \Omega$, so that the result of [4] is applicable in our framework. Combining these facts, we obtain the claimed local Lipschitz continuity with respect to the spatial variable.

Remark 1.3. (Comparison with other results) More generally, one may consider the following evolutionary PDE

$$\partial_t u = \sum_{i=1}^n \partial_{x_i} \left[(|u_{x_i}| - \delta_i)_+^{p_i-1} \frac{u_{x_i}}{|u_{x_i}|} \right] \quad \text{in } \Omega \times I, \quad (1.4)$$

which still exhibits an orthotropic structure. Now we have a full range of exponents $1 < p_1 \leq p_2 \leq \dots \leq p_n$, one for each spatial direction. We refer to [32], where some *global* Lipschitz regularity results are established for solutions to the Cauchy-Dirichlet problem associated with (1.4), but only in the case $\max \{\delta_i\} = 0$ and under suitable regularity assumptions on the data. We emphasize that, due to their global character, for $p_1 = \dots = p_n = p \geq 2$, such results are not comparable with the one proved here. We also mention [19] for a refined Harnack inequality for non-negative local weak solutions, as well as for additional references on the problem. Furthermore, in [24] the authors provide an extensive analysis of the Cauchy problem in the case $\max \{p_i\} < 2$ and $\max \{\delta_i\} = 0$, along with related regularity results.

However, as for the analog of our Theorem 1.1 for *local solutions* of Eq. (1.4), this is still an open problem, as far as we know.

It is worth noting that Theorem 1.1 can also be viewed as an extension to the orthotropic framework of our previous result in [3], where we established local L^∞ bounds for the spatial gradient of weak solutions to strongly degenerate parabolic systems, whose model case is given by the equation

$$\partial_t u - \operatorname{div} \left((|Du| - \lambda)_+^{p-1} \frac{Du}{|Du|} \right) = f, \quad \text{with } \lambda > 0. \quad (1.5)$$

The main feature of this PDE is that the diffusion part is uniformly elliptic only outside a ball with radius λ , while it behaves asymptotically, that is, for large values of $|Du|$, like the parabolic p -Laplace operator. Therefore, equations or systems of the form (1.5) fall into the class of *asymptotically regular parabolic problems* (for a comprehensive overview of this topic, see [1, 5, 6] and the references therein). As already pointed out in [2, 6], no more than Lipschitz regularity can be expected for solutions of equations or systems as in (1.5). In fact, when $f = 0$, any time-independent λ -Lipschitz function solves (1.5), and even more, it is a solution of the associated stationary equation or system.

Finally, we mention the very recent contribution [31], where the author studies parabolic equations of the type

$$\partial_t u - \operatorname{div} D_\xi \mathcal{F}(x, t, Du) = g \quad \text{in } \Omega \times I,$$

with $\mathcal{F} : \Omega \times I \times \mathbb{R}^n \rightarrow [0, \infty)$ satisfying the following conditions:

- (i) $\mathcal{F}$ is only elliptic for values of Du outside a bounded and convex set $E \subset \mathbb{R}^n$ with the property that $0 \in \operatorname{Int} E$;
- (ii) the partial map $\xi \mapsto \mathcal{F}(x, t, \xi)$ is regular on $\mathbb{R}^n \setminus \overline{E}$ and vanishes whenever $\xi \in \overline{E}$.

Assuming that $g \in L^{n+2+\sigma}(\Omega \times I)$ for some $\sigma > 0$, in [31] it is shown that $\mathcal{K}(Du) \in C^0(\Omega \times I)$ for any function $\mathcal{K} \in C^0(\mathbb{R}^n)$ that vanishes on E . This

result extends the C^1 -regularity theorem proved in [6] for Eq. (1.5) in the case $f = 0$, which fulfill (i) and (ii) with

$$\mathcal{F}(x, t, \xi) = \frac{1}{p} (|\xi| - \lambda)_+^p \quad \text{and} \quad E = \{\xi \in \mathbb{R}^n : |\xi| < \lambda\}.$$

1.1. Strategy of the proof

The key step in the proof of Theorem 1.1 is to derive an *a priori* Lipschitz estimate for smooth solutions of the regularized parabolic equation

$$\partial_t u_\varepsilon = \operatorname{div} [D_\xi F_\varepsilon(Du_\varepsilon)], \quad (1.6)$$

where F_ε is a smooth, uniformly convex approximation of the orthotropic function defined in (2.2) below. Our goal is to establish a local Lipschitz estimate for u_ε that is independent of the regularization parameter ε . Passing to the limit as $\varepsilon \rightarrow 0$, we then show that the family $\{u_\varepsilon\}$ converges to the original solution u . This allows us to obtain the Lipschitz estimate for u itself.

In what follows, we emphasize the main difficulties to get such a Lipschitz estimate on u_ε . At first sight, the overall strategy may seem rather classical: we rely on a Moser iterative scheme of reverse Hölder-type inequalities, arising from the interplay between Caccioppoli estimates and Sobolev embeddings. However, as precisely explained in [13, Sect. 1.3], this standard approach cannot be applied directly, due to the severe degeneracy of the orthotropic equation under consideration. For this reason, we need to borrow the machinery developed in the elliptic setting in [9] and [12], and later successfully exploited in [13].

In particular, we make use of standard and nonstandard (so-called “weird”) Caccioppoli inequalities, both introduced in [13] to deal with the solutions of the regularized Eq. (1.6) (see Lemmas 3.3 and 3.4 below). Here, the term *weird* refers to a family of parabolic Caccioppoli-type estimates where two different components of the spatial gradient Du_ε are coupled. More precisely, such estimates involve mixed terms depending simultaneously on $\partial_{x_j} u_\varepsilon$ and $\partial_{x_k} u_\varepsilon$, rather than a single spatial derivative. This feature is crucial in the orthotropic setting, where the degeneracy of the operator precludes the direct use of standard elliptic estimates. By coupling different spatial directions, the aforementioned “weird” Caccioppoli inequalities allow one to compensate for the strong degeneracy of the orthotropic operator and recover enough coercivity to carry out a Moser iterative scheme. Thanks to these inequalities, we can proceed up to a certain point as in the proof of [13, Proposition 4.1], where the parameters δ_i are all equal to 0. In that setting, the combination of standard and nonstandard Caccioppoli inequalities is sufficient to implement the iteration scheme. A substantial difference arises in the present framework, where $\max \{\delta_i\} > 0$. In this case, the degeneracy region of the orthotropic operator on the right-hand side of (1.1) is no longer of measure zero, and this prevents a direct application of the argument used in [13]. To overcome this difficulty and handle the more degenerate situation where $\max \{\delta_i\} > 0$, we have to resort to an idea exploited in [12, Sect. 5.2], in the elliptic setting. More specifically, we

need to perform the Moser iteration using a new measure $d\mu$, which is absolutely continuous with respect to the $(n+1)$ -dimensional Lebesgue measure and is supported on the set whose complement coincides with the region¹

$$\left\{ \max_{1 \leq i \leq n} \left| \frac{\partial u_\varepsilon}{\partial x_i} \right| \leq \delta \right\}, \quad \text{where } \delta = 1 + \max \{ \delta_1, \dots, \delta_n \}.$$

From the L^∞ bound for Du_ε obtained via the Moser scheme with respect to the measure $d\mu$, we can easily deduce the desired local L^∞ estimate for Du_ε in terms of the usual Lebesgue measure, by exploiting the support properties of $d\mu$.

From a technical viewpoint, the introduction of the measure $d\mu$ allows the Moser iteration to be localized away from the degeneracy set of the orthotropic operator. This modification of the underlying measure in the iteration scheme makes the Moser iterative procedure fully compatible with the orthotropic structure of the equation and with the presence of multiple degeneracy thresholds $\delta_i \geq 0$. This idea represents the main new ingredient in our approach compared to [13] and allows us to achieve the proof of Theorem 1.1.

1.2. Plan of the paper

The paper is organized as follows. Section 2 collects the preliminary material, including classical notations, basic properties of local weak solutions to (1.1) and a few auxiliary lemmas. In Sect. 3, we define the exact structure of the regularized Eq. (1.6), whose local weak solution u_ε satisfies the Caccioppoli-type inequalities established in [13] and recalled in Subsect. 3.1. Section 4 is devoted to the Moser iterative scheme, which yields a uniform L^∞ bound for Du_ε through the uniform energy estimates derived in Sect. 5. Finally, in Sect. 6 we complete the proof of Theorem 1.1, by transferring the *a priori* estimates obtained for the approximating solutions u_ε to the original solution u . In treating the limiting case $p = 2$, we also rely on a lemma proved in the Appendix.

2. Notation and preliminaries

In this paper we shall denote by C or c a general positive constant that may vary on different occasions, even within the same line of estimates. Relevant dependencies on parameters and special constants will be suitably emphasized using parentheses or subscripts. The norm we use on $\mathbb{R}^k$, $k \in \mathbb{N}$, will be the standard Euclidean one and it will be denoted by $|\cdot|$. In particular, for the vectors $\xi, \eta \in \mathbb{R}^k$, we write $\langle \xi, \eta \rangle$ for the usual inner product and $|\xi| := \langle \xi, \xi \rangle^{\frac{1}{2}}$ for the corresponding Euclidean norm.

In what follows, we use the notation

$$Q_R(x_0) := x_0 + (-R, R)^n, \quad x_0 \in \mathbb{R}^n, R > 0,$$

¹More precisely, in the limiting case $p = 2$, the parameter δ is replaced by $\delta + 1$.

for the n -dimensional open cube centered at x_0 with side length $2R$. In the paper we will also work with *anisotropic* parabolic cylinders of the type

$$Q_R(x_0, t_0) := Q_R(x_0) \times (t_0 - R^p, t_0), \quad t_0 \in \mathbb{R}. \quad (2.1)$$

If $E \subseteq \mathbb{R}^k$ is a Lebesgue-measurable set, then we will denote by $|E|$ its k -dimensional Lebesgue measure.

In the sequel, we will also adopt the following notation for convenience: given $v \in L^1(\Omega \times I)$,

$$\int_{\Omega \times \{\tau\}} v \, dx := \int_{\Omega} v(x, \tau) \, dx \quad \text{for a.e. } \tau \in I.$$

In this work, we define a local weak solution to (1.1) as follows.

Definition 2.1. Let $F_0 : \mathbb{R}^n \rightarrow \mathbb{R}$ be the function defined by

$$F_0(\xi) := \sum_{i=1}^n \frac{1}{p} (|\xi_i| - \delta_i)_+^p. \quad (2.2)$$

We say that $u \in L_{loc}^p(I; W_{loc}^{1,p}(\Omega))$ is a *local weak solution* of Eq. (1.1) if and only if, for any test function $\varphi \in C_0^\infty(\Omega \times I)$, we have

$$\iint_{\Omega \times I} (u \partial_t \varphi - \langle D_\xi F_0(Du), D\varphi \rangle) \, dx \, dt = 0. \quad (2.3)$$

Let $u \in L_{loc}^p(I; W_{loc}^{1,p}(\Omega))$ be a local weak solution of (1.1). We now recall some additional properties of u : for every subinterval $J \Subset I$ and every open set $\Omega' \Subset \Omega$, we have

$$\partial_t u \in L^{p'}(J; W^{-1,p'}(\Omega')) \quad \text{and} \quad u \in C^0(\bar{J}; L^2(\Omega')). \quad (2.4)$$

Here, $p' := p/(p-1)$ is the conjugate exponent of p and $W^{-1,p'}(\Omega')$ is the topological dual space of $W_0^{1,p}(\Omega')$, which in turn is the completion of $C_0^\infty(\Omega')$ with respect to the L^p norm of the gradient.

In what follows, we briefly recall the argument used to obtain (2.4). Fix J and Ω' as above. By using Eq. (2.3) and the fact that $D_\xi F_0(Du) \in L_{loc}^{p'}(\Omega \times I, \mathbb{R}^n)$, one easily gets

$$\begin{aligned} \left| \iint_{\Omega' \times J} u \partial_t \psi \, dx \, dt \right| &= \left| \iint_{\Omega' \times J} \langle D_\xi F_0(Du), D\psi \rangle \, dx \, dt \right| \\ &\leq n \|Du\|_{L^p(\Omega' \times J)}^{p-1} \|D\psi\|_{L^p(\Omega' \times J)} \\ &\leq n \|Du\|_{L^p(\Omega' \times J)}^{p-1} \|\psi\|_{L^p(J; W_0^{1,p}(\Omega'))}, \quad \text{for every } \psi \in C_0^\infty(\Omega' \times J). \end{aligned}$$

By density, we can extend the linear functional

$$\Lambda : \psi \mapsto \iint_{\Omega' \times J} u \partial_t \psi \, dx \, dt$$

to the whole space $L^p(J; W_0^{1,p}(\Omega'))$. This implies that (see, for example, [30, Theorem 1.5, Chapter III])

$$\Lambda \in \left(L^p(J; W_0^{1,p}(\Omega')) \right)^* = L^{p'}(J; W^{-1,p'}(\Omega')).$$

From the definition of Λ and of weak derivative, we obtain the first property in (2.4).

The second property in (2.4) follows by recalling that for every open set $\mathcal{O} \subset \mathbb{R}^n$, we have (see [30, Proposition 1.2, Chapter III])

$$\mathcal{W}_p(\mathcal{O} \times J) := \left\{ v \in L^p(J; W_0^{1,p}(\mathcal{O})) : \partial_t v \in L^{p'}(J; W^{-1,p'}(\mathcal{O})) \right\} \subset C^0(\bar{J}; L^2(\mathcal{O})).$$

Indeed, it is sufficient to take $\Omega' \Subset \Omega'' \Subset \Omega$ and use the previous inclusion for the function ηu , where $\eta \in C_0^\infty(\Omega'')$ is such that $\eta \equiv 1$ on Ω' . By the first property in (2.4) (used with Ω'' in place of Ω') and the properties of η , we have that

$$\eta u \in \mathcal{W}_p(\Omega'' \times J) \subset C^0(\bar{J}; L^2(\Omega'')).$$

Since $\eta \equiv 1$ on Ω' , the above fact implies that $u \in C^0(\bar{J}; L^2(\Omega'))$, as claimed.

We now gather some lemmas that will be useful to prove our results. Let $J_\lambda : \mathbb{R} \rightarrow \mathbb{R}$ and $H_\lambda : \mathbb{R} \rightarrow \mathbb{R}$ be the auxiliary functions defined respectively by

$$J_\lambda(s) := \begin{cases} (|s| - \lambda)_+^{p-1} \frac{s}{|s|} & \text{if } s \neq 0, \\ 0 & \text{if } s = 0, \end{cases} \quad (2.5)$$

and

$$H_\lambda(s) := \begin{cases} (|s| - \lambda)_+^{\frac{p}{2}} \frac{s}{|s|} & \text{if } s \neq 0, \\ 0 & \text{if } s = 0, \end{cases} \quad (2.6)$$

where $\lambda \geq 0$ is a parameter. We record the following estimate, which can be obtained by suitably modifying the proof of [15, Lemma 4.1] in the case $N = 1$.

Lemma 2.2. *Let $p \geq 2$ and $\lambda \geq 0$. Then, for every $s, t \in \mathbb{R}$ we get*

$$(J_\lambda(s) - J_\lambda(t))(s - t) \geq \frac{4}{p^2} |H_\lambda(s) - H_\lambda(t)|^2.$$

The next result follows from an elementary computation (see, e.g., [16, Lemma A.1]).

Lemma 2.3. *Let $v : \mathbb{R} \rightarrow [0, \infty)$ be a $C^{1,1}$ convex function. Let us set*

$$\mathcal{G}(s) = \int_0^s \sqrt{v''(\tau)} \, d\tau. \quad (2.7)$$

Then, for every $a, b \in \mathbb{R}$ we have

$$(v'(a) - v'(b))(a - b) \geq |\mathcal{G}(a) - \mathcal{G}(b)|^2.$$

Remark 2.4. Since the function v in Lemma 2.3 belongs to $C^{1,1}(\mathbb{R})$, its derivative v' is Lipschitz continuous. Hence, v'' exists almost everywhere in $\mathbb{R}$ (by Rademacher's theorem) and belongs to $L^\infty(\mathbb{R})$. In particular, the integral in (2.7) is well defined as a Lebesgue integral.

We conclude this section with the following classical result; see [25, Lemma 6.1] for a proof.

Lemma 2.5. *Let $0 \leq \rho_0 < \rho_1 < \infty$ and assume that $Z : [\rho_0, \rho_1] \rightarrow [0, \infty)$ is a bounded function satisfying*

$$Z(\rho) \leq \theta Z(r) + \frac{A}{(r-\rho)^\alpha} + \frac{B}{(r-\rho)^\beta} + C$$

for all $\rho_0 \leq \rho < r \leq \rho_1$, for some $\theta \in (0, 1)$ and fixed non-negative constants $A, B, C, \alpha > \beta > 0$. Then, there exists a constant $\kappa = \kappa(\alpha, \theta) > 0$ such that

$$Z(\rho_0) \leq \kappa \left(\frac{A}{(\rho_1 - \rho_0)^\alpha} + \frac{B}{(\rho_1 - \rho_0)^\beta} + C \right).$$

3. Estimates for a regularized equation

We define

$$G(\xi) := \frac{1}{p} (1 + |\xi|^2)^{\frac{p}{2}}, \quad \xi \in \mathbb{R}^n, \quad (3.1)$$

and, for $i \in \{1, \dots, n\}$, we set

$$g_i(s) := \frac{1}{p} (|s| - \delta_i)_+^p, \quad s \in \mathbb{R}.$$

For $p > 2$ and for every $\varepsilon \in (0, 1)$, we consider the convex function

$$F_\varepsilon(\xi) := \sum_{i=1}^n g_i(\xi_i) + \varepsilon G(\xi), \quad \xi \in \mathbb{R}^n. \quad (3.2)$$

Remark 3.1. For $p = 2$ and $\delta_i > 0$, we have $g_i \in C^1(\mathbb{R}) \cap C^\infty(\mathbb{R} \setminus \{-\delta_i, \delta_i\})$, but g_i is not in $C^2(\mathbb{R})$. In this case, one would need to replace g_i by a regularized version $g_{i,\varepsilon}$, in particular for the derivation of the ellipticity bounds in Lemma 3.2 below. For $0 < \varepsilon < \min\{1, \delta_i\}$, a regularized version of g_i is given by

$$g_{i,\varepsilon}(s) := \begin{cases} 0 & \text{if } |s| \leq \delta_i - \varepsilon, \\ \frac{1}{12\varepsilon} (|s| - \delta_i + \varepsilon)^3 & \text{if } \delta_i - \varepsilon \leq |s| \leq \delta_i + \varepsilon, \\ \frac{\varepsilon^2}{6} + \frac{1}{2} (|s| - \delta_i)^2 & \text{if } |s| \geq \delta_i + \varepsilon, \end{cases} \quad (3.3)$$

which converges in C^1 to $(|s| - \delta_i)_+^2/2$ as ε goes to 0 (see Lemma A.1 in the Appendix and [14, Sect. 2]). Therefore, for $p = 2$, to ensure that $F_\varepsilon \in C^2(\mathbb{R}^n)$, we need to define it as follows:

- first, we denote $\Delta^+ := \{\delta_i : i \in \{1, \dots, n\}, \delta_i > 0\}$;
- then, for $\varepsilon \in (0, \min\{1, \inf \Delta^+\})$, we define

$$F_\varepsilon(\xi) = \sum_{i=1}^n \tilde{g}_{i,\varepsilon}(\xi_i) + \frac{\varepsilon}{2} (1 + |\xi|^2), \quad (3.4)$$

with the convention that $\inf \Delta^+ = +\infty$ if $\Delta^+ = \emptyset$ and

$$\tilde{g}_{i,\varepsilon} := \begin{cases} g_i & \text{if } \delta_i = 0, \\ g_{i,\varepsilon} & \text{if } \delta_i > 0, \end{cases} \quad \forall i \in \{1, \dots, n\}. \quad (3.5)$$

Lemma 3.2. *Let $p \geq 2$, $\varepsilon \in (0, \min\{1, \inf \Delta^+\})$ and $\xi \in \mathbb{R}^n$. Then, for every $\zeta \in \mathbb{R}^n$ we have*

$$\varepsilon (1 + |\xi|^2)^{\frac{p-2}{2}} |\zeta|^2 \leq \langle D^2 F_\varepsilon(\xi) \zeta, \zeta \rangle \leq (1 + \varepsilon) (p-1) (1 + |\xi|^2)^{\frac{p-2}{2}} |\zeta|^2. \quad (3.6)$$

Proof. For $p > 2$, a straightforward computation reveals that

$$D^2 F_\varepsilon(\xi) = \text{diag}(g_1''(\xi_1), \dots, g_n''(\xi_n)) + \varepsilon D^2 G(\xi),$$

where, for every $i \in \{1, \dots, n\}$,

$$g_i''(s) = (p-1) (|s| - \delta_i)_+^{p-2}, \quad s \in \mathbb{R},$$

and

$$D^2 G(\xi) = (1 + |\xi|^2)^{\frac{p-4}{2}} [(1 + |\xi|^2) \mathbb{I} + (p-2) \xi \otimes \xi].$$

Thus, for every $\zeta \in \mathbb{R}^n$ we get

$$\begin{aligned} \langle D^2 F_\varepsilon(\xi) \zeta, \zeta \rangle &= \sum_{i=1}^n g_i''(\xi_i) \zeta_i^2 \\ &\quad + \varepsilon (1 + |\xi|^2)^{\frac{p-2}{2}} |\zeta|^2 + \varepsilon (p-2) (1 + |\xi|^2)^{\frac{p-4}{2}} \sum_{i,j=1}^n \xi_i \xi_j \zeta_i \zeta_j \\ &= (p-1) \sum_{i=1}^n (|\xi_i| - \delta_i)_+^{p-2} \zeta_i^2 \\ &\quad + \varepsilon (1 + |\xi|^2)^{\frac{p-2}{2}} |\zeta|^2 + \varepsilon (p-2) (1 + |\xi|^2)^{\frac{p-4}{2}} \langle \xi, \zeta \rangle^2. \end{aligned} \quad (3.7)$$

Using the Cauchy-Schwarz inequality, from (3.7) we obtain

$$\begin{aligned} \langle D^2 F_\varepsilon(\xi) \zeta, \zeta \rangle &\leq (p-1) (1 + |\xi|^2)^{\frac{p-2}{2}} |\zeta|^2 + \varepsilon (1 + |\xi|^2)^{\frac{p-2}{2}} |\zeta|^2 \\ &\quad + \varepsilon (p-2) (1 + |\xi|^2)^{\frac{p-4}{2}} |\xi|^2 |\zeta|^2 \\ &\leq (1 + \varepsilon) (p-1) (1 + |\xi|^2)^{\frac{p-2}{2}} |\zeta|^2. \end{aligned} \quad (3.8)$$

For the derivation of the lower bound, it is sufficient to observe that the first and third terms in the right-hand side of (3.7) are non-negative.

In the case $p = 2$, we replace each g_i with the function $\tilde{g}_{i,\varepsilon} \in C^2(\mathbb{R})$ defined by (3.3) and (3.5). Noting that $0 \leq \tilde{g}_{i,\varepsilon}'' \leq 1$ for any $i \in \{1, \dots, n\}$, when $p = 2$ we immediately have

$$\varepsilon |\zeta|^2 \leq \langle D^2 F_\varepsilon(\xi) \zeta, \zeta \rangle = \sum_{i=1}^n \tilde{g}_{i,\varepsilon}''(\xi_i) \zeta_i^2 + \varepsilon |\zeta|^2 \leq (1 + \varepsilon) |\zeta|^2. \quad (3.9)$$

This completes the proof. $\square$

Now, for every $\varepsilon \in (0, \min\{1, \inf \Delta^+\})$, we consider a local weak solution $u_\varepsilon \in L_{loc}^p(I; W_{loc}^{1,p}(\Omega))$ of the equation

$$\partial_t v = \text{div}[D_\xi F_\varepsilon(Dv)] \quad \text{in } \Omega \times I.$$

This means that u_ε verifies

$$\iint_{\Omega \times I} (u_\varepsilon \partial_t \varphi - \langle D_\xi F_\varepsilon(Du_\varepsilon), D\varphi \rangle) dx dt = 0, \quad \text{for every } \varphi \in C_0^\infty(\Omega \times I). \quad (3.10)$$

Since F_ε belongs to $C^2(\mathbb{R}^n)$ and satisfies (3.6), we can rely on the classical regularity theory for quasilinear parabolic equations, see e.g. [21, Theorem 5.1, Chapter VIII] and [7, Lemma 3.1], to get:

$$Du_\varepsilon \in L_{loc}^\infty(\Omega \times I, \mathbb{R}^n) \quad \text{and} \quad u_\varepsilon \in L_{loc}^2(I; W_{loc}^{2,2}(\Omega)).$$

For convenience of notation, from now on we drop the index $\varepsilon \in (0, \min\{1, \inf \Delta^+\})$ and simply write u and F in place of u_ε and F_ε , unless otherwise specified.

3.1. Caccioppoli-type inequalities

The first technical tools in the proof of Theorem 1.1 are the following standard and nonstandard Caccioppoli inequalities, established in [13, Lemmas 3.1 and 3.2].

Lemma 3.3. (Standard Caccioppoli inequality) *Let $\eta \in C_0^\infty(\Omega)$ and $\chi \in C_0^\infty((T_0, T_1])$ be two non-negative functions, with χ non-decreasing. Let $h : \mathbb{R} \rightarrow \mathbb{R}$ be a C^1 convex non-negative function. Then, for almost every $\tau \in I$ and every $j \in \{1, \dots, n\}$, we have*

$$\begin{aligned} \chi(\tau) \int_{\Omega \times \{\tau\}} h^2(u_{x_j}) \eta^2 dx &+ \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Dh(u_{x_j}), Dh(u_{x_j}) \rangle \chi \eta^2 dx dt \\ &\leq \iint_{\Omega \times (T_0, \tau)} (\partial_t \chi) \eta^2 h^2(u_{x_j}) dx dt \\ &+ 4 \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) D\eta, D\eta \rangle h^2(u_{x_j}) \chi dx dt. \end{aligned}$$

While the standard Caccioppoli inequality in Lemma 3.3 provides an estimate for suitable convex functions of a single spatial derivative u_{x_j} , it is not sufficient in the orthotropic setting due to the strong degeneracy of the operator. To overcome this issue, we need a more refined, nonstandard estimate involving two different components of Du . This estimate is the content of the next lemma, which is referred to as a *weird* Caccioppoli inequality in [13]. The key idea is to mix together the components of the spatial gradient Du with respect to two orthogonal directions, in such a way as to compensate for the severe degeneracy of the orthotropic operator, as already explained in Subsect. 1.1 and in [13, Sect. 1.3].

Lemma 3.4. (Weird Caccioppoli inequality) *Let $\eta \in C_0^\infty(\Omega)$ and $\chi \in C_0^\infty((T_0, T_1])$ be two non-negative functions, with χ non-decreasing. Let $\Phi : [0, \infty) \rightarrow \mathbb{R}$ and $\Psi : [0, \infty) \rightarrow \mathbb{R}$ be two C^1 non-decreasing and non-negative convex functions. Then, for almost every $\tau \in I$, every $j, k \in \{1, \dots, n\}$ and every $\alpha \in [0, 1]$, we have*

$$\chi(\tau) \int_{\Omega \times \{\tau\}} \Phi(u_{x_j}^2) \Psi(u_{x_k}^2) \eta^2 dx$$

$$\begin{aligned}
& + \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle \Phi'(u_{x_j}^2) \Psi(u_{x_k}^2) \chi \eta^2 dx dt \\
& \leq \iint_{\Omega \times (T_0, \tau)} \Phi(u_{x_j}^2) \Psi(u_{x_k}^2) (\partial_t \chi) \eta^2 dx dt \\
& + 4 \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) D\eta, D\eta \rangle [u_{x_j}^2 \Phi'(u_{x_j}^2) \Psi(u_{x_k}^2) + u_{x_k}^2 \Phi(u_{x_j}^2) \Psi'(u_{x_k}^2)] \chi dx dt \\
& + 8 \left(\iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle u_{x_j}^2 [\Phi'(u_{x_j}^2)]^2 [\Psi'(u_{x_k}^2)]^\alpha \chi \eta^2 dx dt \right)^{\frac{1}{2}} \\
& \times \left(\iint_{\Omega \times (T_0, \tau)} \left[\frac{1}{4} (\partial_t \chi) \eta^2 + \langle D^2 F(Du) D\eta, D\eta \rangle \chi \right] |u_{x_k}|^{2\alpha} [\Psi(u_{x_k}^2)]^{2-\alpha} dx dt \right)^{\frac{1}{2}}.
\end{aligned}$$

4. A quantitative L^∞ bound for Du_ε

In this section, we establish a uniform L^∞ estimate for Du_ε , working with the anisotropic parabolic cylinders defined in (2.1). As in the previous section, we will drop the index $\varepsilon \in (0, \min \{1, \inf \Delta^+\})$ and simply write u and F in place of u_ε and F_ε , respectively. Moreover, we set

$$\delta := 1 + \max \{\delta_i : i = 1, \dots, n\}. \quad (4.1)$$

Proposition 4.1. *Let $n \geq 2$ and $p \geq 2$. Moreover, let δ be defined according to (4.1). Then, there exist constants $\vartheta = \vartheta(n) > 2$ and $C = C(n, p, \delta) > 1$ such that, for every $\varepsilon \in (0, \min \{1, \inf \Delta^+\})$ and for every $Q_r(x_0, t_0) \subset Q_R(x_0, t_0) \subseteq \Omega \times I$ with $R \leq 1$, we have*

$$\|Du_\varepsilon\|_{L^\infty(Q_r(x_0, t_0))} \leq \frac{C}{(R-r)^{\vartheta p}} \left[1 + \left(\iint_{Q_R(x_0, t_0)} |Du_\varepsilon|^p dx dt \right)^{\frac{1}{2}} \right]. \quad (4.2)$$

Proof. For simplicity, we limit ourselves to the case $n \geq 3$. This allows us to apply the Sobolev inequality valid for every $v \in W_0^{1,2}(\Omega)$

$$\|v\|_{L^{2^*}(\Omega)} \leq C_n \|Dv\|_{L^2(\Omega)} \quad \text{with } 2^* = \frac{2n}{n-2}.$$

Here C_n denotes a constant depending only on n . The case $n = 2$ requires only minor modifications, whose details are left to the reader.

As the proof is rather intricate, we divide it into several steps for clarity. The first two steps follow the proof of the analogous Proposition 4.1 in [13], where, however, $\delta_i = 0$ for every $i \in \{1, \dots, n\}$. For the sake of completeness, we include them here as well.

Step 1: the choices of Φ and Ψ . We apply Lemma 3.4 with the following choices:

$$\Phi(t) = t^s \quad \text{and} \quad \Psi(t) = t^m, \quad \text{for } t \geq 0,$$

with $1 \leq s \leq m$. We also take

$$\alpha = \begin{cases} \frac{m-s}{m-1} \in [0, 1] & \text{if } m > 1, \\ 1 & \text{if } m = 1. \end{cases}$$

This yields

$$\begin{aligned}
& \chi(\tau) \int_{\Omega \times \{\tau\}} |u_{x_j}|^{2s} |u_{x_k}|^{2m} \eta^2 dx \\
& + s \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_j}|^{2s-2} |u_{x_k}|^{2m} \chi \eta^2 dx dt \\
& \leq \iint_{\Omega \times (T_0, \tau)} |u_{x_j}|^{2s} |u_{x_k}|^{2m} (\partial_t \chi) \eta^2 dx dt \\
& + 4 \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) D\eta, D\eta \rangle [s |u_{x_j}|^{2s} |u_{x_k}|^{2m} + m |u_{x_k}|^{2m} |u_{x_j}|^{2s}] \chi dx dt \\
& + 8 s m^{\frac{\alpha}{2}} \left(\iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_j}|^{4s-2} |u_{x_k}|^{2m-2s} \chi \eta^2 dx dt \right)^{\frac{1}{2}} \\
& \times \left(\iint_{\Omega \times (T_0, \tau)} \left[\frac{1}{4} (\partial_t \chi) \eta^2 + \langle D^2 F(Du) D\eta, D\eta \rangle \chi \right] |u_{x_k}|^{2(s+m)} dx dt \right)^{\frac{1}{2}}.
\end{aligned}$$

On the product of the last two integrals, we apply Young's inequality in the form

$$ab \leq a^2 + \frac{b^2}{4}.$$

Thus we obtain

$$\begin{aligned}
& \chi(\tau) \int_{\Omega \times \{\tau\}} |u_{x_j}|^{2s} |u_{x_k}|^{2m} \eta^2 dx \\
& + \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_j}|^{2s-2} |u_{x_k}|^{2m} \chi \eta^2 dx dt \\
& \leq \iint_{\Omega \times (T_0, \tau)} |u_{x_j}|^{2s} |u_{x_k}|^{2m} (\partial_t \chi) \eta^2 dx dt \\
& + 4(s+m) \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) D\eta, D\eta \rangle |u_{x_j}|^{2s} |u_{x_k}|^{2m} \chi dx dt \\
& + 16 s^2 m \iint_{\Omega \times (T_0, \tau)} \left[\frac{1}{4} (\partial_t \chi) \eta^2 + \langle D^2 F(Du) D\eta, D\eta \rangle \chi \right] |u_{x_k}|^{2(s+m)} dx dt \\
& + \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_j}|^{4s-2} |u_{x_k}|^{2m-2s} \chi \eta^2 dx dt,
\end{aligned}$$

where we have also used that $s \geq 1$ in the left-hand side and $m^\alpha \leq m$ in the right-hand side. By Young's inequality again, we can estimate

$$|u_{x_j}|^{2s} |u_{x_k}|^{2m} \leq |u_{x_j}|^{2(s+m)} + |u_{x_k}|^{2(s+m)}.$$

This finally gives

$$\begin{aligned}
& \chi(\tau) \int_{\Omega \times \{\tau\}} |u_{x_j}|^{2s} |u_{x_k}|^{2m} \eta^2 dx \\
& + \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_j}|^{2s-2} |u_{x_k}|^{2m} \chi \eta^2 dx dt \\
& \leq 16(s+m+s^2 m) \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \langle D^2 F(Du) D\eta, D\eta \rangle \chi] \\
& \times [|u_{x_j}|^{2(s+m)} + |u_{x_k}|^{2(s+m)}] dx dt
\end{aligned}$$

$$+ \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_j}|^{4s-2} |u_{x_k}|^{2m-2s} \chi \eta^2 dx dt. \quad (4.3)$$

Step 2: the staircase. Let $\ell_0 \in \mathbb{N} \setminus \{0\}$ and set $M = 2^{\ell_0} - 1$. We define the two families of indices

$$s_\ell = 2^\ell \quad \text{and} \quad m_\ell = M + 1 - 2^\ell, \quad \text{for } \ell \in \{0, \dots, \ell_0\}.$$

By construction, for every $\ell \in \{0, \dots, \ell_0 - 1\}$ we have

$$s_\ell + m_\ell = M + 1, \quad 2s_\ell - 1 = s_{\ell+1} - 1 \quad \text{and} \quad m_\ell - s_\ell = m_{\ell+1}.$$

We also use that $s_\ell + m_\ell + s_\ell^2 m_\ell \leq 2(M+1)^3$. Then, inequality (4.3) written for $s = s_\ell$ and $m = m_\ell$, with $0 \leq \ell \leq \ell_0 - 1$, gives

$$\begin{aligned} & \chi(\tau) \int_{\Omega \times \{\tau\}} |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\ & + \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_j}|^{2s_\ell-2} |u_{x_k}|^{2m_\ell} \chi \eta^2 dx dt \\ & \leq 32(M+1)^3 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \langle D^2 F(Du) D\eta, D\eta \rangle \chi] \\ & \quad \times \left[|u_{x_j}|^{2(M+1)} + |u_{x_k}|^{2(M+1)} \right] dx dt \\ & + \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_j}|^{2s_{\ell+1}-2} |u_{x_k}|^{2m_{\ell+1}} \chi \eta^2 dx dt. \end{aligned}$$

By summing with respect to ℓ from 0 to $\ell_0 - 1$, and erasing the common terms on both sides, we get

$$\begin{aligned} & \chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\ & + \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_k}|^{2M} \chi \eta^2 dx dt \\ & \leq 260 M^3 \ell_0 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \langle D^2 F(Du) D\eta, D\eta \rangle \chi] \\ & \quad \times \left[|u_{x_j}|^{2(M+1)} + |u_{x_k}|^{2(M+1)} \right] dx dt \\ & + \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_j}|^{2M} \chi \eta^2 dx dt. \end{aligned}$$

Now, in order to estimate the last term on the right-hand side of the previous inequality, we use Lemma 3.3 with the choice

$$h(y) = \frac{|y|^{M+1}}{M+1}, \quad y \in \mathbb{R}.$$

We thus obtain

$$\iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_j}|^{2M} \chi \eta^2 dx dt$$

$$\begin{aligned}
&= \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Dh(u_{x_j}), Dh(u_{x_j}) \rangle \chi \eta^2 dx dt \\
&\leq \frac{4}{(M+1)^2} \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \langle D^2 F(Du) D\eta, D\eta \rangle \chi] |u_{x_j}|^{2(M+1)} dx dt.
\end{aligned}$$

Combining the two previous estimates and using the fact that

$$M = 2^{\ell_0} - 1 \geq \ell_0 \geq 1 \quad \text{for every } \ell_0 \in \mathbb{N} \setminus \{0\},$$

we find

$$\begin{aligned}
&\chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\
&\quad + \iint_{\Omega \times (T_0, \tau)} \langle D^2 F(Du) Du_{x_j}, Du_{x_j} \rangle |u_{x_k}|^{2M} \chi \eta^2 dx dt \\
&\leq 261 M^4 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \langle D^2 F(Du) D\eta, D\eta \rangle \chi] \\
&\quad \times \left[|u_{x_j}|^{2(M+1)} + |u_{x_k}|^{2(M+1)} \right] dx dt. \tag{4.4}
\end{aligned}$$

Step 3: weak ellipticity and boundedness of $D^2 F$ for $p > 2$. From now on, we shall assume that $p > 2$, unless otherwise specified.

From (3.7) and (3.8), it follows that

$$(p-1) \sum_{i=1}^n (|\xi_i| - \delta_i)_+^{p-2} \zeta_i^2 \leq \langle D^2 F(\xi) \zeta, \zeta \rangle \leq c(p) (1 + |\xi|^{p-2}) |\zeta|^2 \text{ for every } \xi, \zeta \in \mathbb{R}^n.$$

Inserting these estimates into (4.4), one gets

$$\begin{aligned}
&\chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\
&\quad + (p-1) \sum_{i=1}^n \iint_{\Omega \times (T_0, \tau)} (|u_{x_i}| - \delta_i)_+^{p-2} u_{x_i x_j}^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt \\
&\leq c(p) M^4 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi (1 + |Du|^{p-2}) |D\eta|^2] \\
&\quad \times [|u_{x_j}|^{2(M+1)} + |u_{x_k}|^{2(M+1)}] dx dt.
\end{aligned}$$

We now consider the second term on the left-hand side. By keeping in the sum only the term with $i = k$ and dropping the others, we obtain

$$\begin{aligned}
& \sum_{i=1}^n \iint_{\Omega \times (T_0, \tau)} (|u_{x_i}| - \delta_i)_+^{p-2} u_{x_i x_j}^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt \\
& \geq \iint_{\Omega \times (T_0, \tau)} (|u_{x_k}| - \delta_k)_+^{p-2} u_{x_k x_j}^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt.
\end{aligned}$$

Note that, by Young's inequality, one has

$$\begin{aligned}
& \iint_{\Omega \times (T_0, \tau)} \left| \left[(|u_{x_k}| - \delta_k)_+^{\frac{p}{2}} |u_{x_k}|^M \right]_{x_j} \right|^2 \chi \eta^2 dx dt \\
& \leq 2 \iint_{\Omega \times (T_0, \tau)} \left| \left[(|u_{x_k}| - \delta_k)_+^{\frac{p}{2}} \right]_{x_j} \right|^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt \\
& \quad + 2 \iint_{\Omega \times (T_0, \tau)} (|u_{x_k}| - \delta_k)_+^p \left| [u_{x_k}]_{x_j}^M \right|^2 \chi \eta^2 dx dt \\
& \leq \left(\frac{p^2}{2} + 2M^2 \right) \iint_{\Omega \times (T_0, \tau)} (|u_{x_k}| - \delta_k)_+^{p-2} u_{x_k x_j}^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt \\
& \leq p^2 M^2 \iint_{\Omega \times (T_0, \tau)} (|u_{x_k}| - \delta_k)_+^{p-2} u_{x_k x_j}^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt,
\end{aligned}$$

where, in the last line, we have used that $M \geq 1$ and $\frac{p^2}{2} > 2$. Then, combining the three previous estimates and summing over $j \in \{1, \dots, n\}$ the resulting inequality, we get

$$\begin{aligned}
& \chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} \sum_{j=1}^n |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\
& \quad + \frac{p-1}{p^2 M^2} \iint_{\Omega \times (T_0, \tau)} \left| D \left[(|u_{x_k}| - \delta_k)_+^{\frac{p}{2}} |u_{x_k}|^M \right] \right|^2 \chi \eta^2 dx dt \\
& \leq c(p) M^4 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi (1 + |Du|^{p-2}) |D\eta|^2] \\
& \quad \times \left[\sum_{j=1}^n |u_{x_j}|^{2M+2} + n |u_{x_k}|^{2M+2} \right] dx dt.
\end{aligned}$$

We now use that $p-1 < p^2 M^2$ and add the term

$$\iint_{\Omega \times (T_0, \tau)} (|u_{x_k}| - \delta_k)_+^p |u_{x_k}|^{2M} \chi |D\eta|^2 dx dt$$

to both sides of the preceding inequality. With some algebraic manipulations, this gives

$$\begin{aligned}
& \chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} \sum_{j=1}^n |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\
& \quad + \iint_{\Omega \times (T_0, \tau)} \left| D \left[(|u_{x_k}| - \delta_k)_+^{\frac{p}{2}} |u_{x_k}|^M \eta \right] \right|^2 \chi dx dt
\end{aligned}$$

$$\begin{aligned}
&\leq c(p) M^6 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi (1 + |Du|^{p-2}) |D\eta|^2] \\
&\quad \times \left[\sum_{j=1}^n |u_{x_j}|^{2M+2} + n |u_{x_k}|^{2M+2} \right] dx dt \\
&\quad + 2 \iint_{\Omega \times (T_0, \tau)} (|u_{x_k}| - \delta_k)_+^p |u_{x_k}|^{2M} \chi |D\eta|^2 dx dt \\
&\leq c(p) M^6 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi (1 + |Du|^{p-2}) |D\eta|^2] \\
&\quad \times \left[\sum_{j=1}^n |u_{x_j}|^{2M+2} + n |u_{x_k}|^{2M+2} \right] dx dt \\
&\quad + c(p) M^6 \iint_{\Omega \times (T_0, \tau)} |u_{x_k}|^{p+2M} \chi |D\eta|^2 dx dt.
\end{aligned}$$

By using the Sobolev inequality in the spatial variable for the second term on the left-hand side, we obtain

$$\begin{aligned}
&\chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} \sum_{j=1}^n |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\
&\quad + \int_{T_0}^\tau \chi \left(\int_{\Omega} (|u_{x_k}| - \delta_k)_+^{p \frac{2^*}{2}} |u_{x_k}|^{2^* M} \eta^{2^*} dx \right)^{\frac{2}{2^*}} dt \\
&\leq c M^6 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi (1 + |Du|^{p-2}) |D\eta|^2] \\
&\quad \times \left[\sum_{j=1}^n |u_{x_j}|^{2M+2} + n |u_{x_k}|^{2M+2} \right] dx dt \\
&\quad + c M^6 \iint_{\Omega \times (T_0, \tau)} |u_{x_k}|^{p+2M} \chi |D\eta|^2 dx dt,
\end{aligned}$$

where c is now a positive constant depending only on n and p . Finally, we sum over $k \in \{1, \dots, n\}$ and apply Minkowski's inequality to the second term on the left-hand side. This yields

$$\begin{aligned}
&\chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} \sum_{j=1}^n |u_{x_j}|^{2s_\ell} \sum_{k=1}^n |u_{x_k}|^{2m_\ell} \eta^2 dx \\
&\quad + \int_{T_0}^\tau \chi \left(\int_{\Omega} \left| \sum_{k=1}^n (|u_{x_k}| - \delta_k)_+^p |u_{x_k}|^{2M} \right|^{\frac{2^*}{2}} \eta^{2^*} dx \right)^{\frac{2}{2^*}} dt \\
&\leq c M^6 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi (1 + |Du|^{p-2}) |D\eta|^2]
\end{aligned}$$

$$\begin{aligned}
& \times \left[\sum_{j=1}^n |u_{x_j}|^{2M+2} + \sum_{k=1}^n |u_{x_k}|^{2M+2} \right] dx dt \\
& + c M^6 \iint_{\Omega \times (T_0, \tau)} \sum_{k=1}^n |u_{x_k}|^{p+2M} \chi |D\eta|^2 dx dt. \quad (4.5)
\end{aligned}$$

We now introduce the auxiliary function

$$\mathcal{U}(x, t) := \frac{1}{2\delta} \max_{1 \leq i \leq n} |u_{x_i}(x, t)|,$$

where the parameter δ is defined in (4.1). A few elementary calculations show that

$$\sum_{k=1}^n (|u_{x_k}| - \delta_k)_+^p |u_{x_k}|^{2M} \geq (2\delta)^{p+2M} \left(u - \frac{1}{2}\right)_+^p \mathcal{U}^{2M}$$

and

$$(2\delta\mathcal{U})^q \leq \sum_{k=1}^n |u_{x_k}|^q \leq n (2\delta\mathcal{U})^q \quad \text{for every } q \geq 0.$$

In particular, for $q = 2$ we obtain

$$2\delta\mathcal{U} \leq |Du| \leq \sqrt{n} 2\delta\mathcal{U}. \quad (4.6)$$

Inserting these estimates into (4.5) yields

$$\begin{aligned}
& \chi(\tau) \ell_0 (2\delta)^{2M+2} \int_{\Omega \times \{\tau\}} \mathcal{U}^{2M+2} \eta^2 dx \\
& + (2\delta)^{p+2M} \int_{T_0}^{\tau} \chi \left(\int_{\Omega} \left(u - \frac{1}{2}\right)_+^{p \frac{2^*}{2}} \mathcal{U}^{2^*M} \eta^{2^*} dx \right)^{\frac{2}{2^*}} dt \\
& \leq c M^6 (2\delta)^{p+2M} \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi (1 + \mathcal{U}^{p-2}) |D\eta|^2] \mathcal{U}^{2M+2} dx dt \\
& \quad + c M^6 (2\delta)^{p+2M} \iint_{\Omega \times (T_0, \tau)} \mathcal{U}^{p+2M} \chi |D\eta|^2 dx dt \\
& \leq c M^6 (2\delta)^{p+2M} \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi |D\eta|^2] (1 + \mathcal{U}^{p+2M}) dx dt, \quad (4.7)
\end{aligned}$$

where, in the first line, we have also used that $s_\ell + m_\ell = M + 1$ for every $\ell \in \{0, \dots, \ell_0 - 1\}$. Dividing both sides of (4.7) by $(2\delta)^{2M+2}$ and using that $\ell_0 \geq 1$ together with $(2\delta)^{p-2} \geq 2^{p-2} > 1$, we get

$$\begin{aligned}
& \chi(\tau) \int_{\Omega \times \{\tau\}} \mathcal{U}^{2M+2} \eta^2 dx + \int_{T_0}^{\tau} \chi \left(\int_{\Omega} \left(u - \frac{1}{2}\right)_+^{p \frac{2^*}{2}} \mathcal{U}^{2^*M} \eta^{2^*} dx \right)^{\frac{2}{2^*}} dt \\
& \leq C M^6 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi |D\eta|^2] (1 + \mathcal{U}^{p+2M}) dx dt, \quad (4.8)
\end{aligned}$$

where C is a positive constant depending only on n , p and δ .

Step 4: choice of the cut-off functions. Let $(x_0, t_0) \in \Omega \times I$ and $0 < r < R \leq 1$ be such that the cube $Q_R(x_0) := x_0 + (-R, R)^n$ is compactly contained in Ω . In addition, we require that

$$(t_0 - R^p, t_0) \Subset I,$$

so that we must have $T_0 < t_0 < T_1$ and $R^p < t_0 - T_0$. Let $\chi : [T_0, T_1] \rightarrow \mathbb{R}$ be a non-decreasing Lipschitz continuous function such that

$$\chi \equiv 0 \quad \text{on } [T_0, t_0 - R^p], \quad \chi \equiv 1 \quad \text{on } [t_0 - r^p, t_0] \quad \text{and} \quad \partial_t \chi \leq \frac{\tilde{c}}{(R - r)^p}.$$

Let $\eta \in C_0^\infty(Q_R(x_0))$ be such that

$$0 \leq \eta \leq 1, \quad \eta \equiv 1 \quad \text{on } Q_r(x_0) \quad \text{and} \quad |D\eta| \leq \frac{\tilde{c}}{R - r}.$$

We recall the notation for the anisotropic parabolic cylinder

$$Q_\rho(x_0, t_0) := Q_\rho(x_0) \times (t_0 - \rho^p, t_0), \quad \rho > 0.$$

With this choice of χ and η , we apply estimate (4.8) twice: firstly, by discarding the second term on the left-hand side and taking the supremum over τ in the interval $(t_0 - r^p, t_0)$; secondly, by dropping the first term on the left-hand side and taking $\tau = t_0$. Summing the two resulting inequalities yields

$$\begin{aligned} & \sup_{\tau \in (t_0 - r^p, t_0)} \int_{Q_r(x_0) \times \{\tau\}} \mathcal{U}^{2M+2} dx \\ & + \int_{t_0 - r^p}^{t_0} \left(\int_{Q_r(x_0)} \left(\mathcal{U} - \frac{1}{2} \right)_+^{p \frac{2^*}{2}} \mathcal{U}^{2^*M} dx \right)^{\frac{2}{2^*}} dt \\ & \leq C \frac{M^6}{(R - r)^p} \iint_{Q_R(x_0, t_0)} (1 + \mathcal{U}^{p+2M}) dx dt, \end{aligned} \quad (4.9)$$

where we have also used that $(R - r)^p \leq (R - r)^2$, since $R \leq 1$ and $p > 2$. By Hölder's inequality, we have

$$\begin{aligned} & \iint_{Q_r(x_0, t_0)} \left(\mathcal{U} - \frac{1}{2} \right)_+^p \mathcal{U}^{2M + \frac{4(M+1)}{n}} dx dt \\ & \leq \int_{t_0 - r^p}^{t_0} \left(\int_{Q_r(x_0)} \left(\mathcal{U} - \frac{1}{2} \right)_+^{p \frac{2^*}{2}} \mathcal{U}^{2^*M} dx \right)^{\frac{2}{2^*}} \left(\int_{Q_r(x_0)} \mathcal{U}^{2M+2} dx \right)^{\frac{2}{n}} dt \\ & \leq \left(\sup_{\tau \in (t_0 - r^p, t_0)} \int_{Q_r(x_0) \times \{\tau\}} \mathcal{U}^{2M+2} dx \right)^{\frac{2}{n}} \\ & \quad \times \int_{t_0 - r^p}^{t_0} \left(\int_{Q_r(x_0)} \left(\mathcal{U} - \frac{1}{2} \right)_+^{p \frac{2^*}{2}} \mathcal{U}^{2^*M} dx \right)^{\frac{2}{2^*}} dt. \end{aligned}$$

Combining the previous estimate with (4.9), we obtain

$$\iint_{Q_r(x_0, t_0)} \left(\mathcal{U} - \frac{1}{2} \right)_+^p \mathcal{U}^{2M + \frac{4(M+1)}{n}} dx dt$$

$$\leq \left[C \frac{M^6}{(R-r)^p} \iint_{Q_R(x_0, t_0)} (1 + \mathcal{U}^{p+2M}) dx dt \right]^{\frac{2}{n}+1}. \quad (4.10)$$

We now estimate

$$\begin{aligned} \iint_{Q_R(x_0, t_0)} (1 + \mathcal{U}^{p+2M}) dx dt &\leq 2|Q_R(x_0, t_0)| + \iint_{Q_R(x_0, t_0) \cap \{\mathcal{U} \geq 1\}} \mathcal{U}^{p+2M} dx dt \\ &\leq 2^{n+1} + \iint_{Q_R(x_0, t_0) \cap \{\mathcal{U} \geq 1\}} \mathcal{U}^{p+2M} dx dt, \end{aligned}$$

where, in the last line, we have used that $R \leq 1$. Observe that on the set $\{\mathcal{U} \geq 1\}$, we have $\mathcal{U} \leq 2(\mathcal{U} - \frac{1}{2})_+$. Hence,

$$\begin{aligned} \iint_{Q_R(x_0, t_0)} (1 + \mathcal{U}^{p+2M}) dx dt &\leq 2^{n+1} + 2^p \iint_{Q_R(x_0, t_0) \cap \{\mathcal{U} \geq 1\}} \left(\mathcal{U} - \frac{1}{2}\right)_+^p \mathcal{U}^{2M} dx dt \\ &\leq 2^{n+1} + 2^p \iint_{Q_R(x_0, t_0)} \left(\mathcal{U} - \frac{1}{2}\right)_+^p \mathcal{U}^{2M} dx dt. \end{aligned} \quad (4.11)$$

Joining (4.10) and (4.11), we then find

$$\begin{aligned} &\iint_{Q_r(x_0, t_0)} \mathcal{U}^{2M + \frac{4(M+1)}{n}} \left(\mathcal{U} - \frac{1}{2}\right)_+^p dx dt \\ &\leq \left[C \frac{M^6}{(R-r)^p} \left(1 + \iint_{Q_R(x_0, t_0)} \mathcal{U}^{2M} \left(\mathcal{U} - \frac{1}{2}\right)_+^p dx dt \right) \right]^{\frac{2}{n}+1}. \end{aligned} \quad (4.12)$$

Step 5: the local L^∞ estimate on Du in the case $p > 2$. We now take $M = M_j = 2^{j+1} - 1$ with $j \in \mathbb{N}$. Then, we set

$$\gamma_j := 2M_j = 2^{j+2} - 2, \quad \widehat{\gamma}_j := 2M_j + \frac{4(M_j + 1)}{n} = 2^{j+2} - 2 + \frac{4}{n} 2^{j+1}, \quad \text{for } j \in \mathbb{N},$$

and

$$\tau_j := \frac{\widehat{\gamma}_j - \gamma_j}{\widehat{\gamma}_j - \gamma_{j-1}} \frac{\gamma_{j-1}}{\gamma_j}, \quad \text{for } j \in \mathbb{N} \setminus \{0\}.$$

We note that $\gamma_{j-1} < \gamma_j < \widehat{\gamma}_j$ and $\tau_j \in (0, 1)$ is defined in such a way that

$$\frac{1}{\gamma_j} = \frac{\tau_j}{\gamma_{j-1}} + \frac{1 - \tau_j}{\widehat{\gamma}_j}.$$

In order to simplify our notation, we also introduce the absolutely continuous measure

$$d\mu := \left(\mathcal{U} - \frac{1}{2}\right)_+^p d\mathcal{L}^{n+1}, \quad (4.13)$$

where $\mathcal{L}^{n+1}$ denotes the $(n+1)$ -dimensional Lebesgue measure. Thus, estimate (4.12) can be rewritten as follows:

$$\int_{Q_r(x_0, t_0)} \mathcal{U}^{\widehat{\gamma}_j} d\mu \leq \left[C \frac{M_j^6}{(R-r)^p} \left(1 + \int_{Q_R(x_0, t_0)} \mathcal{U}^{\gamma_j} d\mu \right) \right]^{\frac{2}{n}+1}.$$

By interpolation in Lebesgue spaces, we obtain

$$\int_{Q_r(x_0, t_0)} \mathcal{U}^{\gamma_j} d\mu \leq \left(\int_{Q_r(x_0, t_0)} \mathcal{U}^{\gamma_{j-1}} d\mu \right)^{\tau_j \frac{\gamma_j}{\gamma_j - 1}} \left(\int_{Q_r(x_0, t_0)} \mathcal{U}^{\widehat{\gamma}_j} d\mu \right)^{(1-\tau_j) \frac{\gamma_j}{\widehat{\gamma}_j}}.$$

Now, a few elementary computations reveal that

$$\tau_j \frac{\gamma_j}{\gamma_{j-1}} = \frac{4}{n+4} \quad \text{and} \quad (1-\tau_j) \frac{\gamma_j}{\widehat{\gamma}_j} = \frac{n}{n+4}.$$

Thus, the combination of the two previous inequalities leads to

$$\begin{aligned} & \int_{Q_r(x_0, t_0)} \mathcal{U}^{\gamma_j} d\mu \\ & \leq \left(\int_{Q_R(x_0, t_0)} \mathcal{U}^{\gamma_{j-1}} d\mu \right)^{\frac{4}{n+4}} \left[C \frac{M_j^6}{(R-r)^p} \left(1 + \int_{Q_R(x_0, t_0)} \mathcal{U}^{\gamma_j} d\mu \right) \right]^{\frac{n+2}{n+4}}. \end{aligned}$$

By Young's inequality, we get

$$\begin{aligned} \int_{Q_r(x_0, t_0)} \mathcal{U}^{\gamma_j} d\mu & \leq \frac{n+2}{n+4} \int_{Q_R(x_0, t_0)} \mathcal{U}^{\gamma_j} d\mu \\ & \quad + \frac{2}{n+4} \left(\frac{CM_j^6}{(R-r)^p} \right)^{\frac{n+2}{2}} \left(\int_{Q_R(x_0, t_0)} \mathcal{U}^{\gamma_{j-1}} d\mu \right)^2 + \frac{n+2}{n+4}. \end{aligned} \quad (4.14)$$

We can now invoke Lemma 2.5 to absorb the term on the right-hand side of (4.14) involving $\mathcal{U}^{\gamma_j}$, in a standard way. By using the definition of M_j and the fact that $R \leq 1$, we get

$$\begin{aligned} \int_{Q_r(x_0, t_0)} \mathcal{U}^{\gamma_j} d\mu & \leq C \frac{2^{3(n+2)j}}{(R-r)^p \frac{n+2}{2}} \left(\int_{Q_R(x_0, t_0)} \mathcal{U}^{\gamma_{j-1}} d\mu \right)^2 + C \\ & \leq C \frac{2^{3(n+2)j}}{(R-r)^p \frac{n+2}{2}} \left[1 + \left(\int_{Q_R(x_0, t_0)} \mathcal{U}^{\gamma_{j-1}} d\mu \right)^2 \right] \\ & \leq C \frac{2^{3(n+2)j}}{(R-r)^p \frac{n+2}{2}} \left(1 + \int_{Q_R(x_0, t_0)} \mathcal{U}^{\gamma_{j-1}} d\mu \right)^2, \end{aligned}$$

for some constant $C = C(n, p, \delta) > 1$. By summing 1 on both sides of the previous estimate, and exploiting that

$$C \frac{2^{3(n+2)j}}{(R-r)^p \frac{n+2}{2}} > 1 \quad \text{for every } j \in \mathbb{N} \setminus \{0\},$$

we obtain

$$1 + \int_{Q_r(x_0, t_0)} \mathcal{U}^{\gamma_j} d\mu \leq C \frac{2^{3(n+2)j}}{(R-r)^p \frac{n+2}{2}} \left[1 + \left(1 + \int_{Q_R(x_0, t_0)} \mathcal{U}^{\gamma_{j-1}} d\mu \right)^2 \right]. \quad (4.15)$$

Now we want to iterate the above estimate on a sequence of shrinking parabolic cylinders. To this end, we consider the decreasing sequence

$$R_j := r + \frac{R-r}{2^{j-1}}, \quad j \in \mathbb{N} \setminus \{0\},$$

and apply (4.15) with $R_{j+1} < R_j$ in place of $r < R$. To simplify our notation, we define

$$Y_j := 1 + \int_{Q_{R_j}(x_0, t_0)} \mathcal{U}^{\gamma_{j-1}} d\mu, \quad j \in \mathbb{N} \setminus \{0\}, \quad (4.16)$$

and

$$\vartheta := \frac{n+2}{2}. \quad (4.17)$$

Using the fact that $Y_j \geq 1$ for every $j \in \mathbb{N} \setminus \{0\}$, and up to redefining the constant $C > 1$, from (4.15) we get

$$Y_{j+1} \leq C 2^{3p(n+2)j} (R-r)^{-\vartheta p} Y_j^2,$$

for any $j \in \mathbb{N} \setminus \{0\}$. By iterating the previous inequality starting from $j = 1$, we obtain for every $k \in \mathbb{N} \setminus \{0\}$

$$\begin{aligned} Y_{k+1} &\leq Y_1^{2^k} \prod_{j=1}^k \left[C 8^{p(n+2)j} (R-r)^{-\vartheta p} \right]^{2^{k-j}} \\ &= Y_1^{2^k} \left[C (R-r)^{-\vartheta p} \right]^{\sum_{j=1}^k 2^{k-j}} 8^{p(n+2) \sum_{j=1}^k j 2^{k-j}}. \end{aligned}$$

Now we observe that

$$\sum_{j=1}^k 2^{k-j} = 2^k \sum_{j=1}^k \left(\frac{1}{2} \right)^j < 2^k$$

and

$$\sum_{j=1}^k j 2^{k-j} = 2^k \sum_{j=1}^k j \left(\frac{1}{2} \right)^j \leq 2^k \frac{\left(\frac{1}{2} \right)}{\left(1 - \frac{1}{2} \right)^2} = 2^{k+1}.$$

Therefore, for every $k \in \mathbb{N} \setminus \{0\}$, we have

$$Y_{k+1} \leq \left[C 64^{p(n+2)} (R-r)^{-\vartheta p} Y_1 \right]^{2^k}.$$

Thus, by redefining the constant $C > 1$ again and recalling the definition of Y_j in (4.16), we obtain

$$\int_{Q_{R_{k+1}}(x_0, t_0)} \mathcal{U}^{\gamma_k} d\mu < Y_{k+1} \leq \left[\frac{C}{(R-r)^{\vartheta p}} \left(1 + \int_{Q_R(x_0, t_0)} \mathcal{U}^2 d\mu \right) \right]^{2^k},$$

for any $k \in \mathbb{N} \setminus \{0\}$. Taking both sides of the previous inequality to the power γ_k^{-1} and then letting $k \rightarrow \infty$, we find

$$\|\mathcal{U}\|_{L^\infty(Q_r(x_0, t_0), d\mu)} \leq \frac{C}{(R-r)^{p \frac{n+2}{8}}} \left(1 + \int_{Q_R(x_0, t_0)} \mathcal{U}^2 d\mu \right)^{\frac{1}{4}}.$$

Here, we have also used the definition of ϑ in (4.17) and the fact that $\gamma_k \sim 2^{k+2}$ as k tends to ∞ . Exploiting once again the condition $R \leq 1$ and recalling the definition of $d\mu$ in (4.13), we deduce from the previous estimate that

$$\begin{aligned} \|\mathcal{U}\|_{L^\infty(Q_r(x_0, t_0), d\mu)} &\leq \frac{C}{(R-r)^p \frac{n+2}{4}} \left(1 + \iint_{Q_R(x_0, t_0)} \mathcal{U}^2 \left(\mathcal{U} - \frac{1}{2} \right)_+^p dx dt \right)^{\frac{1}{4}} \\ &\leq \frac{C}{(R-r)^p \frac{n+2}{4}} \left(1 + \iint_{Q_R(x_0, t_0)} \mathcal{U}^{p+2} dx dt \right)^{\frac{1}{4}} \\ &\leq \frac{C}{(R-r)^p \frac{n+2}{4}} \left(1 + \iint_{Q_R(x_0, t_0)} |Du|^{p+2} dx dt \right)^{\frac{1}{4}}, \end{aligned}$$

where, in the last line, we have applied the inequalities $\mathcal{U} \leq \frac{1}{2\delta}|Du| \leq |Du|$. Recalling the definition of $d\mu$ again, using the above estimate and taking into account that $R \leq 1 < C$, we also get

$$\begin{aligned} \|\mathcal{U}\|_{L^\infty(Q_r(x_0, t_0))} &\leq \max \left\{ \|\mathcal{U}\|_{L^\infty(Q_r(x_0, t_0), d\mu)}, \sup_{Q_r(x_0, t_0) \cap \{\mathcal{U} \leq \frac{1}{2}\}} \mathcal{U} \right\} \\ &\leq \max \left\{ \|\mathcal{U}\|_{L^\infty(Q_r(x_0, t_0), d\mu)}, \frac{1}{2} \right\} \\ &\leq \frac{1}{2} + \|\mathcal{U}\|_{L^\infty(Q_r(x_0, t_0), d\mu)} \\ &\leq \frac{C}{(R-r)^p \frac{n+2}{4}} \left(1 + \iint_{Q_R(x_0, t_0)} |Du|^{p+2} dx dt \right)^{\frac{1}{4}}. \end{aligned}$$

We now apply the second inequality in (4.6) together with the previous estimate. This yields

$$\begin{aligned} \|Du\|_{L^\infty(Q_r(x_0, t_0))} &\leq \sqrt{n} 2\delta \|\mathcal{U}\|_{L^\infty(Q_r(x_0, t_0))} \\ &\leq \frac{C}{(R-r)^p \frac{n+2}{4}} \left[1 + \left(\iint_{Q_R(x_0, t_0)} |Du|^{p+2} dx dt \right)^{\frac{1}{4}} \right]. \end{aligned} \quad (4.18)$$

Finally, in order to remove the dependence on the L^{p+2} norm of the gradient, we use a standard interpolation argument. We write

$$\left(\iint_{Q_R(x_0, t_0)} |Du|^{p+2} dx dt \right)^{\frac{1}{4}} \leq \|Du\|_{L^\infty(Q_R(x_0, t_0))}^{\frac{1}{2}} \left(\iint_{Q_R(x_0, t_0)} |Du|^p dx dt \right)^{\frac{1}{4}}.$$

Inserting this estimate into (4.18) and applying Young's inequality, we obtain

$$\begin{aligned}
\|Du\|_{L^\infty(Q_r(x_0, t_0))} &\leq \frac{C}{(R-r)^p \frac{n+2}{4}} \\
&+ \frac{C}{(R-r)^p \frac{n+2}{4}} \|Du\|_{L^\infty(Q_R(x_0, t_0))}^{\frac{1}{2}} \left(\iint_{Q_R(x_0, t_0)} |Du|^p dx dt \right)^{\frac{1}{4}} \\
&\leq \frac{1}{2} \|Du\|_{L^\infty(Q_R(x_0, t_0))} + \frac{C}{(R-r)^p \frac{n+2}{2}} \left[1 + \left(\iint_{Q_R(x_0, t_0)} |Du|^p dx dt \right)^{\frac{1}{2}} \right],
\end{aligned}$$

where, in the last line, we have also used that $R \leq 1$. By Lemma 2.5 again, we get

$$\|Du\|_{L^\infty(Q_r(x_0, t_0))} \leq \frac{C}{(R-r)^p \frac{n+2}{2}} \left[1 + \left(\iint_{Q_R(x_0, t_0)} |Du|^p dx dt \right)^{\frac{1}{2}} \right].$$

This concludes the proof for the case $p > 2$.

Step 6: the local L^∞ estimate on Du in the case $p = 2$. We now detail the modifications of the above proof to obtain the local estimate (4.2) in the case $p = 2$.

Using the definition of F in (3.4), the fact that $\tilde{g}_{i,\varepsilon}'' \geq \mathbb{1}_{\{|\tau| \geq \delta_i + \varepsilon\}}$ and (3.9), we obtain that

$$\sum_{i \in E^+} \mathbb{1}_{\{|\tau| \geq \delta_i + \varepsilon\}} (\xi_i) \zeta_i^2 + \sum_{i \in E^-} \zeta_i^2 \leq \langle D^2 F(\xi) \zeta, \zeta \rangle \leq 2|\zeta|^2 \text{ for every } \xi, \zeta \in \mathbb{R}^n,$$

where

$$E^+ := \{i \in \{1, \dots, n\} : \delta_i > 0\} \quad \text{and} \quad E^- := \{i \in \{1, \dots, n\} : \delta_i = 0\}. \quad (4.19)$$

Inserting the previous estimates into (4.4), one gets

$$\begin{aligned}
&\chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\
&+ \sum_{i \in E^+} \iint_{\Omega \times (T_0, \tau)} \mathbb{1}_{\{|u_{x_i}| \geq \delta_i + \varepsilon\}} u_{x_i x_j}^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt \\
&+ \sum_{i \in E^-} \iint_{\Omega \times (T_0, \tau)} u_{x_i x_j}^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt \\
&\leq 522 M^4 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi |D\eta|^2] [|u_{x_j}|^{2(M+1)} + |u_{x_k}|^{2(M+1)}] dx dt.
\end{aligned}$$

Now we consider the last two terms on the left-hand side. By keeping in the sums only the term with $i = k$ and dropping the others, we obtain

$$\begin{aligned}
&\chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\
&+ \iint_{\Omega \times (T_0, \tau)} \mathbb{1}_{\{|u_{x_k}| \geq \delta_k + \varepsilon\}} u_{x_k x_j}^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt
\end{aligned}$$

$$\leq 522 M^4 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi |D\eta|^2] [|u_{x_j}|^{2(M+1)} + |u_{x_k}|^{2(M+1)}] dx dt.$$

Note that, by Young's inequality, one has

$$\begin{aligned} & \iint_{\Omega \times (T_0, \tau)} \left| [(|u_{x_k}| - \delta_k - \varepsilon)_+ |u_{x_k}|^M]_{x_j} \right|^2 \chi \eta^2 dx dt \\ & \leq 2 \iint_{\Omega \times (T_0, \tau)} \left| [(|u_{x_k}| - \delta_k - \varepsilon)_+]_{x_j} \right|^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt \\ & \quad + 2 \iint_{\Omega \times (T_0, \tau)} (|u_{x_k}| - \delta_k - \varepsilon)_+^2 \left| [|u_{x_k}|^M]_{x_j} \right|^2 \chi \eta^2 dx dt \\ & \leq 4M^2 \iint_{\Omega \times (T_0, \tau)} \mathbb{1}_{\{|u_{x_k}| \geq \delta_k + \varepsilon\}} u_{x_k}^2 |u_{x_k}|^{2M} \chi \eta^2 dx dt, \end{aligned}$$

where, in the last line, we have used that $M \geq 1$. Then, combining the two previous estimates and summing over $j \in \{1, \dots, n\}$ the resulting inequality, we get

$$\begin{aligned} & \chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} \sum_{j=1}^n |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\ & \quad + \frac{1}{4M^2} \iint_{\Omega \times (T_0, \tau)} |D[(|u_{x_k}| - \delta_k - \varepsilon)_+ |u_{x_k}|^M]|^2 \chi \eta^2 dx dt \\ & \leq 522 M^4 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi |D\eta|^2] \left[\sum_{j=1}^n |u_{x_j}|^{2M+2} + n |u_{x_k}|^{2M+2} \right] dx dt. \end{aligned}$$

We now use that $4M^2 \geq 4$ and add the term

$$\iint_{\Omega \times (T_0, \tau)} (|u_{x_k}| - \delta_k - \varepsilon)_+^2 |u_{x_k}|^{2M} \chi |D\eta|^2 dx dt$$

to both sides of the preceding inequality. With some algebraic manipulations, this gives

$$\begin{aligned} & \chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} \sum_{j=1}^n |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\ & \quad + \iint_{\Omega \times (T_0, \tau)} |D[(|u_{x_k}| - \delta_k - \varepsilon)_+ |u_{x_k}|^M \eta]|^2 \chi dx dt \\ & \leq c_0 M^6 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi |D\eta|^2] \left[\sum_{j=1}^n |u_{x_j}|^{2M+2} + |u_{x_k}|^{2M+2} \right] dx dt \\ & \quad + 2 \iint_{\Omega \times (T_0, \tau)} (|u_{x_k}| - \delta_k - \varepsilon)_+^2 |u_{x_k}|^{2M} \chi |D\eta|^2 dx dt \\ & \leq c M^6 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi |D\eta|^2] \left[\sum_{j=1}^n |u_{x_j}|^{2M+2} + |u_{x_k}|^{2M+2} \right] dx dt, \end{aligned}$$

where c_0 and c are positive constants depending only on n . By using the Sobolev inequality in the spatial variable for the second term on the left-hand

side, we obtain

$$\begin{aligned} & \chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} \sum_{j=1}^n |u_{x_j}|^{2s_\ell} |u_{x_k}|^{2m_\ell} \eta^2 dx \\ & + \int_{T_0}^{\tau} \chi \left(\int_{\Omega} (|u_{x_k}| - \delta_k - \varepsilon)_+^{2^*} |u_{x_k}|^{2^* M} \eta^{2^*} dx \right)^{\frac{2}{2^*}} dt \\ & \leq c M^6 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi |D\eta|^2] \left[\sum_{j=1}^n |u_{x_j}|^{2M+2} + |u_{x_k}|^{2M+2} \right] dx dt. \end{aligned}$$

At this stage, we sum over $k \in \{1, \dots, n\}$ and apply Minkowski's inequality to the second term on the left-hand side. This yields

$$\begin{aligned} & \chi(\tau) \sum_{\ell=0}^{\ell_0-1} \int_{\Omega \times \{\tau\}} \sum_{j=1}^n |u_{x_j}|^{2s_\ell} \sum_{k=1}^n |u_{x_k}|^{2m_\ell} \eta^2 dx \\ & + \int_{T_0}^{\tau} \chi \left(\int_{\Omega} \left| \sum_{k=1}^n (|u_{x_k}| - \delta_k - \varepsilon)_+^2 |u_{x_k}|^{2M} \right|^{\frac{2^*}{2}} \eta^{2^*} dx \right)^{\frac{2}{2^*}} dt \\ & \leq c M^6 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi |D\eta|^2] \\ & \quad \times \left[\sum_{j=1}^n |u_{x_j}|^{2M+2} + \sum_{k=1}^n |u_{x_k}|^{2M+2} \right] dx dt. \end{aligned} \quad (4.20)$$

Similarly to what has been done in steps 3 and 5, we now introduce the auxiliary function

$$\mathcal{V}(x, t) := \frac{1}{2(\delta + 1)} \max_{1 \leq i \leq n} |u_{x_i}(x, t)|$$

and the absolutely continuous measure

$$d\sigma := \left(\mathcal{V} - \frac{1}{2} \right)_+^2 d\mathcal{L}^{n+1}.$$

A few elementary computations reveal that

$$\sum_{k=1}^n (|u_{x_k}| - \delta_k - \varepsilon)_+^2 |u_{x_k}|^{2M} \geq (2\delta + 2)^{2M+2} \left(\mathcal{V} - \frac{1}{2} \right)_+^2 \mathcal{V}^{2M}$$

and

$$[2(\delta + 1) \mathcal{V}]^q \leq \sum_{k=1}^n |u_{x_k}|^q \leq n [2(\delta + 1) \mathcal{V}]^q \quad \text{for every } q \geq 0.$$

In particular, for $q = 2$ we obtain

$$2(\delta + 1) \mathcal{V} \leq |Du| \leq \sqrt{n} 2(\delta + 1) \mathcal{V}.$$

Inserting these estimates into (4.20), using that $\ell_0 \geq 1$ and recalling that $s_\ell + m_\ell = M + 1$ for every $\ell \in \{0, \dots, \ell_0 - 1\}$, we get

$$\begin{aligned} & \chi(\tau) \int_{\Omega \times \{\tau\}} \mathcal{V}^{2M+2} \eta^2 dx + \int_{T_0}^\tau \chi \left(\int_{\Omega} \left(\mathcal{V} - \frac{1}{2} \right)_+^{2^*} \mathcal{V}^{2^*M} \eta^{2^*} dx \right)^{\frac{2}{2^*}} dt \\ & \leq c M^6 \iint_{\Omega \times (T_0, \tau)} [(\partial_t \chi) \eta^2 + \chi |D\eta|^2] \mathcal{V}^{2M+2} dx dt. \end{aligned}$$

Starting from this estimate and proceeding exactly as in steps 4 and 5, but using $\mathcal{V}$, $\delta + 1$ and $d\sigma$ in place of $\mathcal{U}$, δ and $d\mu$, respectively, we reach the desired conclusion for $p = 2$. $\square$

Remark 4.2. A careful inspection of the previous proof reveals that the exponent ϑ in (4.2) can be taken to be

$$\vartheta = \begin{cases} \frac{n+2}{2} & \text{if } n \geq 3, \\ \text{any number } > 2 & \text{if } n = 2. \end{cases} \quad (4.21)$$

In the case $n = 2$, the constant C in (4.2) blows up as $\vartheta \rightarrow 2$.

5. Uniform energy estimates for a regularized problem

Let us fix an open set $\Omega' \Subset \Omega$ and a subinterval $J := (\tau_0, \tau_1) \Subset I$. Assume that $u \in L^p_{loc}(I; W^{1,p}_0(\Omega))$ is a local weak solution of Eq. (1.1). In light of (2.4), we have

$$\partial_t u \in L^{p'}(J; W^{-1,p'}(\Omega')) \quad \text{and} \quad u \in C^0(\bar{J}; L^2(\Omega')).$$

Now, for any fixed $\varepsilon \in (0, \min\{1, \inf \Delta^+\})$, we consider the approximating Cauchy-Dirichlet problem parametrized by ε

$$\begin{cases} \partial_t v = \operatorname{div} [D_\varepsilon F_\varepsilon(Dv)] & \text{in } \Omega' \times J, \\ v = u & \text{on } \partial\Omega' \times J, \\ v(\cdot, \tau_0) = u(\cdot, \tau_0) & \text{in } \Omega'. \end{cases} \quad (5.1)$$

By [30, Proposition 4.1, Chapter III], this problem admits a unique weak solution $u_\varepsilon \in L^p(J; W^{1,p}(\Omega'))$ such that

$$\partial_t u_\varepsilon \in L^{p'}(J; W^{-1,p'}(\Omega')), \quad \text{and thus } u_\varepsilon \in C^0(\bar{J}; L^2(\Omega')).$$

The condition $u_\varepsilon = u$ on the lateral boundary $\partial\Omega' \times J$ is understood in the sense that

$$u_\varepsilon - u \in L^p(J; W^{1,p}_0(\Omega')),$$

while the initial condition $u_\varepsilon(\cdot, \tau_0) = u(\cdot, \tau_0)$ in Ω' is taken in the usual L^2 -sense, which is feasible due to the continuity properties of both u_ε and u .

The penultimate step in the proof of Theorem 1.1 consists in establishing the uniform energy estimates, as well as the strong convergence results, stated in Propositions 5.1 and 5.2 below. The need to distinguish the cases $p > 2$

and $p = 2$ arises from the fact that, for $p = 2$, the regularizing function F_ε is defined by (3.4) rather than (3.2).

Proposition 5.1. (Uniform energy estimate for $p > 2$) *Let $n \geq 2$, $p > 2$ and $\varepsilon \in (0, \min \{1, \inf \Delta^+\})$. Moreover, let $u \in L^p_{loc}(I; W^{1,p}_{loc}(\Omega))$ be a local weak solution of (1.1) and assume that $u_\varepsilon \in L^p(J; W^{1,p}(\Omega'))$ is the unique weak solution of problem (5.1). Then, the estimate*

$$\begin{aligned} & \int_{\Omega' \times \{\tau_1\}} |u_\varepsilon - u|^2 dx + \sum_{i=1}^n \iint_{\Omega' \times J} |H_{\delta_i}(\partial_{x_i} u_\varepsilon) - H_{\delta_i}(\partial_{x_i} u)|^2 dx dt \\ & + \varepsilon \iint_{\Omega' \times J} |Du_\varepsilon|^p dx dt \\ & \leq c\varepsilon \left(|\Omega' \times J| + \iint_{\Omega' \times J} |Du|^p dx dt \right) \end{aligned} \quad (5.2)$$

holds for some positive constant c depending only on p . In particular, this estimate implies that

$$\iint_{\Omega' \times J} |Du_\varepsilon|^p dx dt \leq c \left(|\Omega' \times J| + \iint_{\Omega' \times J} |Du|^p dx dt \right) \quad (5.3)$$

and

$$H_{\delta_j}(\partial_{x_j} u_\varepsilon) \rightarrow H_{\delta_j}(\partial_{x_j} u) \quad \text{strongly in } L^2(\Omega' \times J) \quad \text{as } \varepsilon \rightarrow 0, \quad (5.4)$$

for each $j \in \{1, \dots, n\}$.

Proof. The function u_ε verifies

$$\iint_{\Omega' \times J} u_\varepsilon \partial_t \varphi dx dt - \iint_{\Omega' \times J} \langle D_\xi F_\varepsilon(Du_\varepsilon), D\varphi \rangle dx dt = 0, \quad (5.5)$$

for every $\varphi \in C_0^\infty(\Omega' \times J)$. Integrating by parts in (5.5) yields

$$\int_J (\partial_t u_\varepsilon, \varphi)_{(W^{-1,p'}, W_0^{1,p})} dt + \iint_{\Omega' \times J} \langle D_\xi F_\varepsilon(Du_\varepsilon), D\varphi \rangle dx dt = 0.$$

By density, the above identity also holds for every $\varphi \in L^p(J; W_0^{1,p}(\Omega'))$. We then choose $\varphi = u_\varepsilon - u$, which gives

$$\int_J (\partial_t u_\varepsilon, u_\varepsilon - u)_{(W^{-1,p'}, W_0^{1,p})} dt + \iint_{\Omega' \times J} \langle D_\xi F_\varepsilon(Du_\varepsilon), Du_\varepsilon - Du \rangle dx dt = 0.$$

By recalling the definition of F_ε in (3.2), the previous integral identity can be rewritten as follows:

$$\begin{aligned} & \int_J (\partial_t u_\varepsilon, u_\varepsilon - u)_{(W^{-1,p'}, W_0^{1,p})} dt + \iint_{\Omega' \times J} \langle D_\xi F_0(Du_\varepsilon), Du_\varepsilon - Du \rangle dx dt \\ & + \varepsilon \iint_{\Omega' \times J} \langle D_\xi G(Du_\varepsilon), Du_\varepsilon - Du \rangle dx dt = 0. \end{aligned}$$

Starting from (2.3), we similarly have

$$\int_J (\partial_t u, u_\varepsilon - u)_{(W^{-1,p'}, W_0^{1,p})} dt + \iint_{\Omega' \times J} \langle D_\xi F_0(Du), Du_\varepsilon - Du \rangle dx dt = 0.$$

By subtracting the two identities above, we get

$$\begin{aligned} & \int_J (\partial_t u_\varepsilon - \partial_t u, u_\varepsilon - u)_{(W^{-1,p'}, W_0^{1,p})} dt \\ & + \iint_{\Omega' \times J} \langle D_\xi F_0(Du_\varepsilon) - D_\xi F_0(Du), Du_\varepsilon - Du \rangle dx dt \\ & + \varepsilon \iint_{\Omega' \times J} \langle D_\xi G(Du_\varepsilon), Du_\varepsilon - Du \rangle dx dt = 0. \end{aligned} \quad (5.6)$$

The term involving the time derivatives can be rewritten as

$$\int_J (\partial_t u_\varepsilon - \partial_t u, u_\varepsilon - u)_{(W^{-1,p'}, W_0^{1,p})} dt = \frac{1}{2} \int_{\Omega' \times \{\tau_1\}} |u_\varepsilon - u|^2 dx.$$

This follows from the fact that the map

$$t \mapsto \frac{1}{2} \int_{\Omega'} |u_\varepsilon(x, t) - u(x, t)|^2 dx$$

is absolutely continuous on J , with derivative given exactly by

$$(\partial_t u_\varepsilon - \partial_t u, u_\varepsilon - u)_{(W^{-1,p'}, W_0^{1,p})} \quad \text{for a.e. } t \in J,$$

see [30, Proposition 1.2, Chapter III].

For the second term in (5.6), we apply Lemma 2.2, which, for every $w, z \in \mathbb{R}^n$, yields

$$\begin{aligned} \langle D_\xi F_0(w) - D_\xi F_0(z), w - z \rangle &= \sum_{i=1}^n (J_{\delta_i}(w_i) - J_{\delta_i}(z_i)) (w_i - z_i) \\ &\geq \frac{4}{p^2} \sum_{i=1}^n |H_{\delta_i}(w_i) - H_{\delta_i}(z_i)|^2, \end{aligned}$$

where the functions J_{δ_i} and H_{δ_i} are defined respectively by (2.5) and (2.6) with $\lambda = \delta_i$. Using this pointwise estimate in (5.6) and recalling the definition of G in (3.1), we then get

$$\begin{aligned} & \frac{1}{2} \int_{\Omega' \times \{\tau_1\}} |u_\varepsilon - u|^2 dx + \frac{4}{p^2} \sum_{i=1}^n \iint_{\Omega' \times J} |H_{\delta_i}(\partial_{x_i} u_\varepsilon) - H_{\delta_i}(\partial_{x_i} u)|^2 dx dt \\ & + \varepsilon \iint_{\Omega' \times J} \langle (1 + |Du_\varepsilon|^2)^{\frac{p-2}{2}} Du_\varepsilon, Du_\varepsilon - Du \rangle dx dt \leq 0. \end{aligned} \quad (5.7)$$

By the Cauchy-Schwarz inequality and Young's inequality with $\beta > 0$, from (5.7) we infer

$$\begin{aligned} & \int_{\Omega' \times \{\tau_1\}} |u_\varepsilon - u|^2 dx + \frac{8}{p^2} \sum_{i=1}^n \iint_{\Omega' \times J} |H_{\delta_i}(\partial_{x_i} u_\varepsilon) - H_{\delta_i}(\partial_{x_i} u)|^2 dx dt \\ & + 2\varepsilon \iint_{\Omega' \times J} |Du_\varepsilon|^p dx dt \\ & \leq 2\varepsilon \iint_{\Omega' \times J} (1 + |Du_\varepsilon|^2)^{\frac{p-1}{2}} |Du| dx dt \\ & \leq \frac{2\varepsilon\beta^{p'}}{p'} \iint_{\Omega' \times J} (1 + |Du_\varepsilon|^2)^{\frac{p}{2}} dx dt + \frac{2\varepsilon}{p\beta^p} \iint_{\Omega' \times J} |Du|^p dx dt \end{aligned}$$

$$\leq \frac{2^{\frac{p}{2}} \varepsilon \beta^{p'}}{p'} \iint_{\Omega' \times J} |Du_\varepsilon|^p dx dt + \frac{2^{\frac{p}{2}} \varepsilon \beta^{p'}}{p'} |\Omega' \times J| + \frac{2\varepsilon}{p\beta^p} \iint_{\Omega' \times J} |Du|^p dx dt. \quad (5.8)$$

Upon choosing $\beta = \left(\frac{p'}{2^{p/2}}\right)^{\frac{1}{p'}}$ and absorbing the first integral on the right-hand side of (5.8) into the left-hand side, we arrive at estimate (5.2).

The uniform energy estimate (5.3) follows by discarding the first two terms on the left-hand side of (5.2) and then dividing by ε . Similarly, by dropping the first and third terms on the left-hand side of (5.2) and letting $\varepsilon \rightarrow 0$, we obtain the conclusion (5.4). $\square$

In the case $p = 2$, to obtain a result analogous to Proposition 5.1, we need to introduce, for each $i \in \{1, \dots, n\}$, the auxiliary function

$$K_{i,\varepsilon}(s) := \int_0^s \sqrt{\tilde{g}_{i,\varepsilon}''(\tau)} d\tau, \quad s \in \mathbb{R}, \quad (5.9)$$

where $\tilde{g}_{i,\varepsilon}$ denotes the convex C^2 map defined in (3.3) and (3.5). More precisely, we have the following

Proposition 5.2. (Uniform energy estimate for $p = 2$) *Let $n \geq 2$, $p = 2$ and $\varepsilon \in (0, \min\{1, \inf \Delta^+\})$. Moreover, let $u \in L_{loc}^2(I; W_{loc}^{1,2}(\Omega))$ be a local weak solution of (1.1) and assume that $u_\varepsilon \in L^2(J; W^{1,2}(\Omega'))$ is the unique weak solution of problem (5.1), where F_ε is defined by (3.4). Then, the estimate*

$$\begin{aligned} & \int_{\Omega' \times \{\tau_1\}} |u_\varepsilon - u|^2 dx + \sum_{i=1}^n \iint_{\Omega' \times J} |K_{i,\varepsilon}(\partial_{x_i} u_\varepsilon) - K_{i,\varepsilon}(\partial_{x_i} u)|^2 dx dt \\ & + \varepsilon \iint_{\Omega' \times J} |Du_\varepsilon - Du|^2 dx dt \\ & \leq c\varepsilon \left(|\Omega' \times J| + \iint_{\Omega' \times J} |Du|^2 dx dt \right) \end{aligned} \quad (5.10)$$

holds for some positive constant c depending only on n . In particular, this estimate implies that

$$\iint_{\Omega' \times J} |Du_\varepsilon|^2 dx dt \leq 2(c+1) \left(|\Omega' \times J| + \iint_{\Omega' \times J} |Du|^2 dx dt \right). \quad (5.11)$$

Furthermore, we have

$$K_{j,\varepsilon}(\partial_{x_j} u_\varepsilon) \rightarrow H_{\delta_j}(\partial_{x_j} u) \quad \text{strongly in } L^2(\Omega' \times J) \quad \text{as } \varepsilon \rightarrow 0, \quad (5.12)$$

for each $j \in \{1, \dots, n\}$.

Proof. Arguing exactly as in the first part of the preceding proof, we find that

$$\begin{aligned} & \frac{1}{2} \int_{\Omega' \times \{\tau_1\}} |u_\varepsilon - u|^2 dx + \iint_{\Omega' \times J} \langle D_\xi F_\varepsilon(Du_\varepsilon) - D_\xi F_\varepsilon(Du), Du_\varepsilon - Du \rangle dx dt \\ & = \iint_{\Omega' \times J} \langle D_\xi F_0(Du) - D_\xi F_\varepsilon(Du), Du_\varepsilon - Du \rangle dx dt. \end{aligned}$$

Recalling the definitions of F_ε , $\tilde{g}_{i,\varepsilon}$ and E^+ in (3.4), (3.5) and (4.19), respectively, the above identity can be rewritten as follows:

$$\begin{aligned} & \frac{1}{2} \int_{\Omega' \times \{\tau_1\}} |u_\varepsilon - u|^2 dx + \sum_{i=1}^n \iint_{\Omega' \times J} [\tilde{g}'_{i,\varepsilon}(\partial_{x_i} u_\varepsilon) - \tilde{g}'_{i,\varepsilon}(\partial_{x_i} u)] [\partial_{x_i} u_\varepsilon - \partial_{x_i} u] dx dt \\ & + \varepsilon \iint_{\Omega' \times J} |Du_\varepsilon - Du|^2 dx dt \\ & = \sum_{i \in E^+} \iint_{\Omega' \times J} [g'_i(\partial_{x_i} u) - g'_{i,\varepsilon}(\partial_{x_i} u)] [\partial_{x_i} u_\varepsilon - \partial_{x_i} u] dx dt \\ & - \varepsilon \iint_{\Omega' \times J} \langle Du, Du_\varepsilon - Du \rangle dx dt. \end{aligned}$$

We can estimate the second term on the left-hand side by applying Lemma 2.3 with $v = \tilde{g}_{i,\varepsilon}$ and recalling (5.9). Thus, we obtain

$$\begin{aligned} & \sum_{i=1}^n \iint_{\Omega' \times J} [\tilde{g}'_{i,\varepsilon}(\partial_{x_i} u_\varepsilon) - \tilde{g}'_{i,\varepsilon}(\partial_{x_i} u)] [\partial_{x_i} u_\varepsilon - \partial_{x_i} u] dx dt \\ & \geq \sum_{i=1}^n \iint_{\Omega' \times J} |K_{i,\varepsilon}(\partial_{x_i} u_\varepsilon) - K_{i,\varepsilon}(\partial_{x_i} u)|^2 dx dt. \end{aligned}$$

Combining the two previous estimates and using the Cauchy-Schwarz and Young's inequalities, we get

$$\begin{aligned} & \frac{1}{2} \int_{\Omega' \times \{\tau_1\}} |u_\varepsilon - u|^2 dx + \sum_{i=1}^n \iint_{\Omega' \times J} |K_{i,\varepsilon}(\partial_{x_i} u_\varepsilon) - K_{i,\varepsilon}(\partial_{x_i} u)|^2 dx dt \\ & + \varepsilon \iint_{\Omega' \times J} |Du_\varepsilon - Du|^2 dx dt \\ & \leq \sum_{i \in E^+} \iint_{\Omega' \times J} |g'_i(\partial_{x_i} u) - g'_{i,\varepsilon}(\partial_{x_i} u)| |\partial_{x_i} u_\varepsilon - \partial_{x_i} u| dx dt \\ & + \frac{\varepsilon}{2} \iint_{\Omega' \times J} |Du_\varepsilon - Du|^2 dx dt + \frac{\varepsilon}{2} \iint_{\Omega' \times J} |Du|^2 dx dt. \quad (5.13) \end{aligned}$$

To estimate the first term on the right-hand side, we apply the second inequality of Lemma A.1 from the Appendix, together with Young's inequality in the form

$$ab \leq \frac{a^2}{\varepsilon} + \frac{\varepsilon b^2}{4}.$$

Consequently, we obtain

$$\begin{aligned} & \sum_{i \in E^+} \iint_{\Omega' \times J} |g'_i(\partial_{x_i} u) - g'_{i,\varepsilon}(\partial_{x_i} u)| |\partial_{x_i} u_\varepsilon - \partial_{x_i} u| dx dt \\ & \leq \sum_{i \in E^+} \left[\frac{\varepsilon}{16} |\Omega' \times J| + \frac{\varepsilon}{4} \iint_{\Omega' \times J} |\partial_{x_i} u_\varepsilon - \partial_{x_i} u|^2 dx dt \right] \\ & \leq \frac{n\varepsilon}{16} |\Omega' \times J| + \frac{\varepsilon}{4} \iint_{\Omega' \times J} |Du_\varepsilon - Du|^2 dx dt. \quad (5.14) \end{aligned}$$

Joining (5.13) and (5.14), we deduce the estimate (5.10). As immediate consequences of (5.10), we get

$$\begin{aligned} \iint_{\Omega' \times J} |Du_\varepsilon|^2 dx dt &\leq 2 \iint_{\Omega' \times J} |Du_\varepsilon - Du|^2 dx dt + 2 \iint_{\Omega' \times J} |Du|^2 dx dt \\ &\leq 2(c+1) \left(|\Omega' \times J| + \iint_{\Omega' \times J} |Du|^2 dx dt \right), \end{aligned}$$

and moreover

$$\|K_{j,\varepsilon}(\partial_{x_j} u_\varepsilon) - K_{j,\varepsilon}(\partial_{x_j} u)\|_{L^2(\Omega' \times J)} \longrightarrow 0 \quad \text{as } \varepsilon \rightarrow 0, \quad \text{for each } j \in \{1, \dots, n\}. \quad (5.15)$$

We now proceed to prove (5.12). First observe that, for all $j \in \{1, \dots, n\}$, we have

$$\begin{aligned} H_{\delta_j}(s) &= \int_0^s \mathbf{1}_{\{|w| > \delta_j\}}(\tau) d\tau \quad \text{for every } s \in \mathbb{R}, \\ \sqrt{\tilde{g}_{j,\varepsilon}''(\tau)} &\rightarrow \mathbf{1}_{\{|w| > \delta_j\}}(\tau) \quad \text{as } \varepsilon \rightarrow 0 \quad \text{for a.e. } \tau \in \mathbb{R}, \\ \sqrt{\tilde{g}_{j,\varepsilon}''(\tau)} &\leq 1 \quad \text{for every } \tau \in \mathbb{R}. \end{aligned} \quad (5.16)$$

Hence, by the Dominated Convergence Theorem, we conclude that, for all $j \in \{1, \dots, n\}$,

$$\lim_{\varepsilon \rightarrow 0} K_{j,\varepsilon}(s) = H_{\delta_j}(s) \quad \text{for every } s \in \mathbb{R}.$$

Furthermore, using (2.6), (5.9) and (5.16), we have, for all $j \in \{1, \dots, n\}$,

$$\begin{aligned} |K_{j,\varepsilon}(u_{x_j}) - H_{\delta_j}(u_{x_j})|^2 &\leq 2|K_{j,\varepsilon}(u_{x_j})|^2 + 2|H_{\delta_j}(u_{x_j})|^2 \\ &\leq 4|u_{x_j}|^2 \quad \text{almost everywhere in } \Omega' \times J. \end{aligned}$$

Since $|Du| \in L^2(\Omega' \times J)$, we may apply the Dominated Convergence Theorem again, thus obtaining

$$\|K_{j,\varepsilon}(u_{x_j}) - H_{\delta_j}(u_{x_j})\|_{L^2(\Omega' \times J)} \longrightarrow 0 \quad \text{as } \varepsilon \rightarrow 0, \quad \text{for each } j \in \{1, \dots, n\}. \quad (5.17)$$

Finally, using Minkowski's inequality together with (5.15) and (5.17), we reach the conclusion (5.12). This completes the proof. $\square$

6. Proof of Theorem 1.1

We are now in a position to prove Theorem 1.1. Indeed, using the results from Propositions 5.1 and 5.2, we will now show that estimate (4.2) also holds for any local weak solution u of (1.1), in place of u_ε . Therefore, in the next proof we will adopt the same assumptions and notations as in Propositions 5.1 and 5.2.

Proof of Theorem 1.1. Let $(x_0, t_0) \in \Omega \times I$ and $0 < r < R \leq 1$ be such that the cube $Q_R(x_0)$ is compactly contained in Ω . In addition, we require that $(t_0 - R^p, t_0) \subseteq I$. Now set

$$\rho = \frac{R+r}{2}$$

and let $u_\varepsilon \in L^p(J; W^{1,p}(\Omega'))$ be the unique weak solution to problem (5.1) with $\Omega' = Q_R(x_0)$ and $J = (t_0 - R^p, t_0)$.

Let us first assume that $p > 2$. Then, by Proposition 5.1, we have that

$$H_{\delta_j}(\partial_{x_j} u_\varepsilon) \rightarrow H_{\delta_j}(\partial_{x_j} u) \quad \text{strongly in } L^2(Q_R(x_0, t_0)) \quad \text{as } \varepsilon \rightarrow 0,$$

for each $j \in \{1, \dots, n\}$. Thus, for each fixed $j \in \{1, \dots, n\}$, there exists a sequence $\{\varepsilon_k\}_{k \in \mathbb{N}}$ such that:

- $0 < \varepsilon_k < \min\{1, \inf \Delta^+\}$ for every $k \in \mathbb{N}$ and $\varepsilon_k \searrow 0$ as $k \rightarrow +\infty$;
- $|H_{\delta_j}(\partial_{x_j} u_{\varepsilon_k})| \rightarrow |H_{\delta_j}(\partial_{x_j} u)|$ almost everywhere in $Q_R(x_0, t_0)$ as $k \rightarrow +\infty$.

Therefore, using the definition of H_{δ_j} , Proposition 4.1, the fact that $\rho < R \leq 1$, and (5.3) with $\Omega' \times J = Q_R(x_0, t_0)$, we have for almost every $z \in Q_r(x_0, t_0)$ that

$$\begin{aligned} |u_{x_j}(z)| &\leq |H_{\delta_j}(u_{x_j}(z))|^{\frac{2}{p}} + \delta_j = \lim_{k \rightarrow \infty} |H_{\delta_j}(\partial_{x_j} u_{\varepsilon_k}(z))|^{\frac{2}{p}} + \delta_j \\ &= \lim_{k \rightarrow \infty} \max\{\delta_j, |\partial_{x_j} u_{\varepsilon_k}(z)|\} \\ &\leq \limsup_{k \rightarrow \infty} \operatorname{ess\,sup}_{Q_r(x_0, t_0)} (\max\{\delta_j, |\partial_{x_j} u_{\varepsilon_k}|\}) \\ &\leq \limsup_{k \rightarrow \infty} \max\{\delta_j, \|Du_{\varepsilon_k}\|_{L^\infty(Q_r(x_0, t_0))}\} \\ &\leq \limsup_{k \rightarrow \infty} \left\{ \delta + \|Du_{\varepsilon_k}\|_{L^\infty(Q_r(x_0, t_0))} \right\} \\ &\leq \limsup_{k \rightarrow \infty} \left\{ \delta + \frac{C}{(\rho - r)^{\vartheta p}} \left[1 + \left(\iint_{Q_\rho(x_0, t_0)} |Du_{\varepsilon_k}|^p dx dt \right)^{\frac{1}{2}} \right] \right\} \\ &\leq \limsup_{k \rightarrow \infty} \frac{C_1}{(\rho - r)^{\vartheta p}} \left[1 + \left(\iint_{Q_R(x_0, t_0)} |Du_{\varepsilon_k}|^p dx dt \right)^{\frac{1}{2}} \right] \\ &\leq \frac{C_2}{(R - r)^{\vartheta p}} \left[1 + \left(\iint_{Q_R(x_0, t_0)} |Du|^p dx dt \right)^{\frac{1}{2}} \right], \end{aligned}$$

where ϑ is defined by (4.21), while the constants C , C_1 and C_2 depend only on n , p and δ . Since the above inequality holds for almost every $z \in Q_r(x_0, t_0)$, we immediately get

$$\|u_{x_j}\|_{L^\infty(Q_r(x_0, t_0))} \leq \frac{C_2}{(R - r)^{\vartheta p}} \left[1 + \left(\iint_{Q_R(x_0, t_0)} |Du|^p dx dt \right)^{\frac{1}{2}} \right] < +\infty.$$

This bound holds uniformly for all $j \in \{1, \dots, n\}$. Thus, taking the maximum over j , we obtain estimate (1.3), which proves that $Du \in L_{loc}^\infty(\Omega \times I, \mathbb{R}^n)$.

We now consider the case $p = 2$. By (5.12), we conclude that, for each fixed $j \in \{1, \dots, n\}$, there exists a sequence $\{\tilde{\varepsilon}_k\}_{k \in \mathbb{N}}$ such that:

- $0 < \tilde{\varepsilon}_k < \min\{1, \inf \Delta^+\}$ for every $k \in \mathbb{N}$ and $\tilde{\varepsilon}_k \searrow 0$ as $k \rightarrow +\infty$;
- $|K_{j, \tilde{\varepsilon}_k}(\partial_{x_j} u_{\tilde{\varepsilon}_k})| \rightarrow |H_{\delta_j}(\partial_{x_j} u)|$ almost everywhere in $Q_R(x_0, t_0)$ as $k \rightarrow +\infty$.

Then, arguing as above, but this time also using the definition of $K_{j,\tilde{\varepsilon}_k}$ and the fact that $\sqrt{\tilde{g}_{j,\tilde{\varepsilon}_k}''} \leq 1$, as well as (5.11) in place of (5.3), we obtain for almost every $z \in Q_r(x_0, t_0)$ that

$$\begin{aligned} |u_{x_j}(z)| &\leq |H_{\delta_j}(u_{x_j}(z))| + \delta_j = \lim_{k \rightarrow \infty} |K_{j,\tilde{\varepsilon}_k}(\partial_{x_j} u_{\tilde{\varepsilon}_k}(z))| + \delta_j \\ &\leq \limsup_{k \rightarrow \infty} |\partial_{x_j} u_{\tilde{\varepsilon}_k}(z)| + \delta \\ &\leq \limsup_{k \rightarrow \infty} \frac{C}{(\rho - r)^{2\vartheta}} \left[1 + \left(\iint_{Q_\rho(x_0, t_0)} |Du_{\tilde{\varepsilon}_k}|^2 dx dt \right)^{\frac{1}{2}} \right] + \delta \\ &\leq \limsup_{k \rightarrow \infty} \frac{C_1}{(R - r)^{2\vartheta}} \left[1 + \left(\iint_{Q_R(x_0, t_0)} |Du_{\tilde{\varepsilon}_k}|^2 dx dt \right)^{\frac{1}{2}} \right] \\ &\leq \frac{C_2}{(R - r)^{2\vartheta}} \left[1 + \left(\iint_{Q_R(x_0, t_0)} |Du|^2 dx dt \right)^{\frac{1}{2}} \right]. \end{aligned}$$

This yields the same conclusion as before and thus completes the proof. $\square$

Remark 6.1. A careful inspection of the proofs of Proposition 4.1 and Theorem 1.1 reveals that the constant C in (1.3) depends polynomially on

$$\delta = 1 + \max \{ \delta_i : i = 1, \dots, n \}.$$

In particular, this constant converges to a finite quantity depending only on n and p as $\max \{ \delta_i \} \rightarrow 0$. This is in agreement with the fact that, in [13], the constant in the corresponding local L^∞ estimate for Du depends only on n and p .

Acknowledgements

The author would like to thank the reviewers for their valuable comments, which helped to improve this work.

Author contributions P. A. is the sole author of this work and is responsible for the conception, analysis, and writing of the paper.

Funding Open access funding provided by Uppsala University. This work has been partially supported by the Wallenberg AI, Autonomous Systems and Software Program (WASP) funded by the Knut and Alice Wallenberg Foundation. The author is a member of the GNAMPA group of INdAM, which partially supported his research through the INdAM–GNAMPA 2026 Project “Esistenza e regolarità per soluzioni di equazioni ellittiche e paraboliche anisotrope” (CUP E53C25002010001).

The author also acknowledges financial support from the IADE_CITTL2020 Project “Intersectorial applications of differential equations” (CUP J34I20000 980006).

Data Availability Statement No datasets were generated or analysed during the current study.

Declarations

Competing interests The authors declare no competing interests.

Open Access. This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

A. Appendix

Let us fix $i \in \{1, \dots, n\}$ and let $0 < \varepsilon < \delta_i$. We recall the C^2 function $g_{i,\varepsilon} : \mathbb{R} \rightarrow [0, \infty)$ defined in (3.3). In this appendix, we show that $g_{i,\varepsilon}$ converges in $C^1(\mathbb{R})$ to

$$g_i(s) := \frac{1}{2} (|s| - \delta_i)_+^2 \quad (\text{A.1})$$

as $\varepsilon \rightarrow 0$. More precisely, we establish the following result.

Lemma A.1. *Let $0 < \varepsilon < \delta_i$. Then, for every $s \in \mathbb{R}$ we have*

$$|g_{i,\varepsilon}(s) - g_i(s)| \leq \frac{\varepsilon^2}{6} \quad \text{and} \quad |g'_{i,\varepsilon}(s) - g'_i(s)| \leq \frac{\varepsilon}{4}. \quad (\text{A.2})$$

Proof. Since $g_{i,\varepsilon}$ and g_i are even functions, it suffices to prove the claim for every $s \geq 0$. For convenience of notation, we set

$$r = s - \delta_i.$$

Recalling the definitions in (3.3) and (A.1), we immediately have:

- $g_{i,\varepsilon}(r) = g_i(r) = 0$ if $r \in [-\delta_i, -\varepsilon]$;
- $|g_{i,\varepsilon}(r) - g_i(r)| = \frac{1}{12\varepsilon} (r + \varepsilon)^3 \leq \frac{\varepsilon^2}{12}$ if $r \in [-\varepsilon, 0]$;
- $|g_{i,\varepsilon}(r) - g_i(r)| = \frac{\varepsilon^2}{6}$ if $r \in [\varepsilon, \infty)$.

Furthermore, for every $r \in [0, \varepsilon]$ one has

$$|g_{i,\varepsilon}(r) - g_i(r)| = \left| \frac{1}{12\varepsilon} (r + \varepsilon)^3 - \frac{1}{2} r^2 \right| \leq \frac{\varepsilon^2}{6},$$

since the function

$$\phi(\tau) := \frac{1}{12\varepsilon} (\tau + \varepsilon)^3 - \frac{1}{2} \tau^2$$

is increasing and, moreover, $\phi(0) = \frac{\varepsilon^2}{12}$ and $\phi(\varepsilon) = \frac{\varepsilon^2}{6}$. We have thus obtained the first inequality in (A.2) for every $s \geq 0$.

The second inequality in (A.2) follows by a similar argument. We leave the details to the reader. $\square$

References

- [1] Ambrosio, P.: Fractional Sobolev regularity for solutions to a strongly degenerate parabolic equation. *Forum Math.* **35**(6), 1485–1497 (2023)
- [2] Ambrosio, P.: Sharp Sobolev regularity for widely degenerate parabolic equations. *Calc. Var.* **64**, 32 (2025)
- [3] Ambrosio, P., Bäuerlein, F.: Gradient bounds for strongly singular or degenerate parabolic systems. *J. Differ. Equ.* **401**, 492–549 (2024)
- [4] Ambrosio, P., Ciani, S.: Local boundedness for weak solutions to strongly degenerate orthotropic parabolic equations. *Ricerche mat.* **75**, 561–579 (2026)
- [5] Ambrosio, P., Passarelli di Napoli, A.: Regularity results for a class of widely degenerate parabolic equations. *Adv. Calc. Var.* **17**(3), 805–829 (2023)
- [6] Bögelein, V., Duzaar, F., Giova, R., Passarelli di Napoli, A.: Gradient regularity for a class of widely degenerate parabolic systems. *SIAM J. Math. Anal.* **56**(4), 5017–5078 (2024)
- [7] Bögelein, V., Duzaar, F., Marcellini, P.: Parabolic equations with p, q -growth. *J. Math. Pures Appl.* **100**(4), 535–563 (2013)
- [8] Bögelein, V., Duzaar, F., Mingione, G.: The regularity of general parabolic systems with degenerate diffusion. *Mem. Amer. Math. Soc.* **221**, 1–143 (2013)
- [9] Bousquet, P., Brasco, L.: C^1 regularity of orthotropic p -harmonic functions in the plane. *Anal. PDE* **11**(4), 813–854 (2018)
- [10] Bousquet, P., Brasco, L.: Lipschitz regularity for orthotropic functionals with nonstandard growth conditions. *Rev. Mat. Iberoam.* **36**(7), 1989–2032 (2020)

- [11] Bousquet, P., Brasco, L., Julin, V.: Lipschitz regularity for local minimizers of some widely degenerate problems. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **16**(4), 1235–1274 (2016)
- [12] Bousquet, P., Brasco, L., Leone, C., Verde, A.: On the Lipschitz character of orthotropic p -harmonic functions. *Calc. Var.* **57**, 88 (2018)
- [13] Bousquet, P., Brasco, L., Leone, C., Verde, A.: Gradient estimates for an orthotropic nonlinear diffusion equation. *Adv. Calc. Var.* **16**(3), 705–730 (2023)
- [14] Brasco, L., Carlier, G.: On certain anisotropic elliptic equations arising in congested optimal transport: local gradient bounds. *Adv. Calc. Var.* **7**, 379–407 (2014)
- [15] Brasco, L., Carlier, G., Santambrogio, F.: Congested traffic dynamics, weak flows and very degenerate elliptic equations, [corrected version of mr2584740], *J. Math. Pures Appl.* (9) **93** (6), 652–671 (2010)
- [16] Brasco, L., Leone, C., Pisante, G., Verde, A.: Sobolev and Lipschitz regularity for local minimizers of widely degenerate anisotropic functionals. *Nonlinear Anal.* **153**, 169–199 (2017)
- [17] Chen, Y.Z., DiBenedetto, E.: On the local behavior of solutions of singular parabolic equations. *Arch. Ration. Mech. Anal.* **103**, 319–345 (1988)
- [18] Choe, H.J.: Hölder regularity for the gradient of solutions of certain singular parabolic systems. *Comm. Partial Differential Equations* **16**(11), 1709–1732 (1991)
- [19] Ciani, S., Mosconi, S., Vespri, V.: Parabolic Harnack estimates for anisotropic slow diffusion. *JAMA* **149**, 611–642 (2023)
- [20] DiBenedetto, E.: On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients. *Ann. Sc. Norm. Super. Pisa Cl. Sci.* **13**(3), 487–535 (1986)
- [21] DiBenedetto, E.: *Degenerate Parabolic Equations*. Universitext, Springer, New York (1993)
- [22] DiBenedetto, E., Friedman, A.: Regularity of solutions of nonlinear degenerate parabolic systems. *J. Reine Angew. Math.* **349**, 83–128 (1984)
- [23] DiBenedetto, E., Friedman, A.: Hölder estimates for nonlinear degenerate parabolic systems. *J. Reine Angew. Math.* **357**, 1–22 (1985)
- [24] Feo, F., Vázquez, J.L., Volzone, B.: Anisotropic p -Laplacian evolution of fast diffusion type. *Adv. Nonlinear Stud.* **21**(3), 523–555 (2021)
- [25] Giusti, E.: *Direct Methods in the Calculus of Variations*, World Scientific Publishing Co (2003)
- [26] Kuusi, T., Mingione, G.: Potential estimates and gradient boundedness for nonlinear parabolic systems. *Rev. Mat. Iberoam.* **28**(2), 535–576 (2012)
- [27] Kuusi, T., Mingione, G.: Riesz potentials and nonlinear parabolic equations. *Arch. Ration. Mech. Anal.* **212**(3), 727–780 (2014)

- [28] Lions, J.-L.: Quelques Méthodes de Résolution des Problèmes Aux Limites Non Linéaires. Gauthier-Villars, Paris (1969)
- [29] Russo, S.: Higher differentiability of minimizers for non-autonomous orthotropic functionals. *Nonlinear Anal. RWA* **87**, 104450 (2026)
- [30] Showalter, R.E.: Monotone Operators in Banach Space and Nonlinear Partial Differential Equations, *Math. Surveys Monogr.* **49**, American Mathematical Society, Providence, (1997)
- [31] Strunk, M.: Gradient regularity for widely degenerate parabolic equations, preprint (2025), [arXiv: 2510.07999](https://arxiv.org/abs/2510.07999)
- [32] Tersenov, A.S., Tersenov, A.S.: Existence of Lipschitz continuous solutions to the Cauchy-Dirichlet problem for anisotropic parabolic equations. *J. Funct. Anal.* **272**(10), 3965–3986 (2017)
- [33] Vishik, M.I.: Solubility of boundary-value problems for quasi-linear parabolic equations of higher orders. *Mat. Sb. (N. S.)* **59**(101), 289–325 (1962)
- [34] Wiegner, M.: On C_α -regularity of the gradient of solutions of degenerate parabolic systems. *Ann. Mat. Pura Appl. I.* **4**(145), 385–405 (1986)
- [35] Zeidler, E.: Nonlinear Functional Analysis and its Applications. II/A - Linear Monotone Operators, Springer, New York (1989)

Pasquale Ambrosio
Department of Mathematics
Uppsala University
P.O. Box 480, 751 06 Uppsala
Sweden
e-mail: pasquale.ambrosio@math.uu.se

Received: 3 November 2025.

Revised: 1 June 2026.

Accepted: 7 June 2026.