

**This is a self-archived version of an original article. This version may differ from the original in pagination and typographic details.**

**Author(s):** Feng, Yawen; Parviainen, Mikko; Sarsa, Saara

**Title:** A systematic approach on the second order regularity of solutions to the general parabolic p-Laplace equation

**Year:** 2023

**Version:** Published version

**Copyright:** © The Author(s) 2023

**Rights:** CC BY 4.0

**Rights url:** <https://creativecommons.org/licenses/by/4.0/>

**Please cite the original version:**

Feng, Y., Parviainen, M., & Sarsa, S. (2023). A systematic approach on the second order regularity of solutions to the general parabolic p-Laplace equation. *Calculus of Variations and Partial Differential Equations*, 62, Article 204. <https://doi.org/10.1007/s00526-023-02537-z>



# A systematic approach on the second order regularity of solutions to the general parabolic $p$ -Laplace equation

Yawen Feng<sup>1,2</sup> · Mikko Parviainen<sup>1</sup> · Saara Sarsa<sup>1</sup>

Received: 5 April 2023 / Accepted: 10 June 2023  
© The Author(s) 2023

## Abstract

We study a general form of a degenerate or singular parabolic equation

$$u_t - |Du|^\gamma (\Delta u + (p - 2)\Delta_\infty^N u) = 0$$

that generalizes both the standard parabolic  $p$ -Laplace equation and the normalized version that arises from stochastic game theory. We develop a systematic approach to study second order Sobolev regularity and show that  $D^2u$  exists as a function and belongs to  $L^2_{loc}$  for a certain range of parameters. In this approach proving the estimate boils down to verifying that a certain coefficient matrix is positive definite. As a corollary we obtain, under suitable assumptions, that a viscosity solution has a Sobolev time derivative belonging to  $L^2_{loc}$ .

**Mathematics Subject Classification** 35B65 · 35K55 · 35K65 · 35K67

## Contents

1	Introduction	.....
2	Preliminaries	.....
3	Plan of proof	.....
3.1	Derivation of a basic estimate	.....
3.2	Formal calculation for smooth solutions with a nonvanishing gradient	.....
3.3	Solutions without smoothness assumptions and regularized equation	.....
4	Hidden divergence structures, the key estimate and auxiliary lemmas	.....

---

Communicated by Jan Kristensen.

---

✉ Mikko Parviainen  
mikko.j.parviainen@jyu.fi

Yawen Feng  
yawen.y.feng@jyu.fi

Saara Sarsa  
saara.m.sarsa@jyu.fi

<sup>1</sup> Department of Mathematics and Statistics, University of Jyväskylä, PO Box 35, 40014 Jyväskylä, Finland

<sup>2</sup> School of Mathematical Science, Beihang University, Changping District Shahe Higher Education Park South Third Street No. 9, Beijing 102206, People's Republic of China

- 4.1 Hidden divergence structures . . . . .
- 4.2 The key estimate . . . . .
- 4.3 Auxiliary lemmas . . . . .
- 5 Smooth case with non-zero gradient . . . . .
- 6 Removing the smoothness assumption . . . . .
  - 6.1 Regularization . . . . .
  - 6.2 Passing to the original equation . . . . .
- Appendix A: Proof of Lemma 4.5 . . . . .
- Appendix B: Positive definiteness condition for the coefficient matrix . . . . .
- References . . . . .

## 1 Introduction

Recently, the second order regularity for parabolic  $p$ -Laplace type equations has been studied by Høeg and Lindqvist [13], Dong et al. [10], and the authors [11]. In this article, we consider a rather general class of parabolic equations

$$u_t - |Du|^\gamma (\Delta u + (p - 2)\Delta_\infty^N u) = 0 \tag{1.1}$$

with  $1 < p < \infty$  and  $-1 < \gamma < \infty$ , where

$$\Delta_\infty^N u := |Du|^{-2} \sum_{i,j=1}^n u_{x_i} u_{x_j} u_{x_i x_j} = |Du|^{-2} \langle Du, D^2 u Du \rangle = |Du|^{-2} \Delta_\infty u$$

denotes the normalized infinity Laplacian. The equation contains the game theoretic or normalized  $p$ -parabolic equation and the divergence form standard  $p$ -parabolic equation as special cases. The equation is not uniformly parabolic or in divergence form except in special cases, and it can be highly degenerate or singular in the gradient variable. Regularity for such equations has been recently studied for example by Imbert, Jin and Silvestre as well as Parviainen and Vázquez as discussed below. The objective of this article is to develop a systematic approach to study the second order spatial regularity of viscosity solutions to (1.1). In this approach proving the estimate reduces down to verifying that a certain coefficient matrix is positive definite. For the further notation and the definition of viscosity solutions to (1.1), we refer to Sect. 2.

In [11] we considered second order Sobolev regularity of the parabolic  $p$ -Laplace equation

$$u_t - \Delta_p u = 0 \tag{1.2}$$

where  $\Delta_p u := \operatorname{div}(|Du|^{p-2} Du)$  is the  $p$ -Laplace operator. Notice that, in the special case  $\gamma = p - 2$ , Eq. (1.1) can be formally, and also rigorously by [16], rewritten as (1.2). One of the key tools is the fundamental inequality (the name stems from Dong, Peng, Zhang and Zhou [10] for a related inequality)

$$|Du|^4 |D^2 u|^2 \geq 2|Du|^2 |D^2 u Du|^2 + \frac{(|Du|^2 \Delta u - \Delta_\infty u)^2}{n - 1} - (\Delta_\infty u)^2 \tag{1.3}$$

which holds for any smooth function  $u$  as shown by Sarsa [25]. Curiously, in [11] it was sufficient to use the above inequality in a simpler form just estimating  $(|Du|^2 \Delta u - \Delta_\infty u)^2 \geq 0$  on the right hand side. With the general equation in this paper, we use the inequality in the full generality. A natural approach to obtain second order Sobolev estimates is to differentiate (1.1), multiply the equation with suitable quantities containing gradients, and manipulate in a suitable way. Thus, among other terms, one can obtain terms in divergence form, which can be controlled. In the case of (1.2), one then uses (1.3) in a simple form as explained above and

thus gets an upper bound for a quantity containing second derivatives. Part of the difficulty in dealing with the general equation instead of the  $p$ -parabolic equation stems from the fact that this approach gives rise to the mixed terms of the type

$$|Du|^{-\gamma} u_t \Delta_\infty^N u$$

which are difficult to handle.

Another difficulty arises from the fact that of course  $u$  is not known to be smooth a priori when differentiating the equation, and negative powers of the gradient are problematic as the gradient might vanish. A natural approach to these problems is regularizing the equation by adding a small regularization parameter, which removes the singularity. Unfortunately, when differentiating the regularized equation, one gets another set of problematic terms that no longer match the terms in the fundamental inequality. Treating these terms is a subtle issue, and we need to guarantee that a sum of certain terms remain nonnegative by carefully analyzing explicit coefficients of the terms.

In order to analyze the nonnegativity of the problematic terms and their coefficients systematically, we develop several techniques. We interpret the terms and their coefficients, as a quadratic form and derive a range condition for the parameters from the positive definiteness condition of this quadratic form. In order to improve the range obtained in this way, we use a hidden divergence structure. Indeed, suitable mixed terms can actually be written in a divergence form, and thus by adding such terms, we can manipulate the coefficients at the cost of adding divergence form terms that can be estimated.

Some steps, in particular checking that the quadratic form is positive definite, of the above plan when written down explicitly are quite complicated, and thus for the convenience of the reader we first provide a formal calculation in Sect. 5, where we assume that the solution is smooth and the gradient nonvanishing. In this case, the above plan gives an optimal (optimality is discussed in Example 5.1) a priori estimate (Proposition 5.1),

$$\int_{Q_r} \left| D(|Du|^{\frac{p-2+s}{2}} Du) \right|^2 dxdt \leq \frac{C}{r^2} \left( \int_{Q_{2r}} |Du|^{p+s} dxdt + \int_{Q_{2r}} |Du|^{p+s-\gamma} dxdt \right)$$

in the range

$$1 < p < \infty, \quad -1 < \gamma < \infty \quad \text{and} \quad n \geq 2.$$

with the range condition

$$s > \max \left\{ -1 - \frac{p-1}{n-1}, \gamma + 1 - p \right\}. \tag{1.4}$$

The left hand side in the above estimate is of the same form as the estimate in [11]. In particular, we may set  $s = 2 - p$ ,  $s = 0$  and  $s = p - 2$  giving

$$D^2u, \quad D(|Du|^{\frac{p-2}{2}} Du) \quad \text{and} \quad D(|Du|^{p-2} Du)$$

as special cases.

Perhaps surprisingly, removing the smoothness assumption and the assumption on the nonvanishing gradient by using the regularized equation turns out to be a problem. In particular, the additional terms resulting from the regularization add to the technical complication of showing that the quadratic form is positive definite. To reduce technical complication partly for expository reasons, we have decided to restrict ourselves to the case  $n = 2$  in the regularized case. In this context we obtain the following result.

**Theorem 1.1** *Let  $n = 2$ . Let  $u : \Omega_T \rightarrow \mathbb{R}$  be a viscosity solution to the general  $p$ -parabolic equation (1.1). If  $p$  and  $\gamma$  satisfy one of the following conditions:*

- (i)  $1 < p \leq 5$  and  $-1 < \gamma < 1$ ; or
- (ii)  $1 < p < \infty$  and  $-1 < \gamma < \sqrt{2} - \frac{1}{2}$ ,

*then  $D^2u$  exists and belongs to  $L^2_{loc}(\Omega_T)$ . Moreover, we have the estimate*

$$\int_{Q_r} |D^2u|^2 dxdt \leq \frac{C}{r^2} \left( \int_{Q_{2r}} |Du|^2 dxdt + \int_{Q_{2r}} |Du|^{2-\gamma} dxdt \right),$$

*where  $C = C(p, \gamma) > 0$  and  $Q_r \subset Q_{2r} \Subset \Omega_T$  are concentric parabolic cylinders.*

This also implies that time derivative exists as an  $L^2$ -function, which is not evident directly by the definition.

**Corollary 1.2** (Time derivative) *Let  $n = 2$ . Let  $u : \Omega_T \rightarrow \mathbb{R}$  be a viscosity solution to the general  $p$ -parabolic Eq. (1.1). If  $p$  and  $\gamma$  satisfy one of the following conditions:*

- (i)  $1 < p \leq 5$  and  $0 \leq \gamma < 1$ ; or
- (ii)  $1 < p < \infty$  and  $0 \leq \gamma < \sqrt{2} - \frac{1}{2}$ ,

*then the time derivative  $u_t$  exists as a function and  $u_t \in L^2_{loc}(\Omega_T)$ .*

At least to some extent the range condition in Theorem 1.1 is an artifact as we explain later. It would be interesting to know whether the theorem is valid in the whole range of parameters.

Next we review the known regularity results of Eq. (1.1) and explain how our results fit into the existing literature. If  $\gamma = p - 2$ , then Eq. (1.1) is the parabolic  $p$ -Laplace Eq. (1.2). For the regularity theory of weak solutions to (1.2) we refer to the monograph of DiBenedetto [8]. In particular, if  $u$  is a continuous weak solution to (1.2), then  $u \in C^\alpha_{loc}$  and  $Du \in C^\beta_{loc}$  for some  $0 < \alpha, \beta < 1$ .

Moreover, Lindqvist [18] showed in the degenerate case  $2 < p < \infty$  that

$$D(|Du|^{\frac{p-2}{2}} Du) \in L^2_{loc},$$

and further that

$$D(|Du|^{p-2} Du) \in L^{\frac{p}{p-1}}_{loc}.$$

The singular case is treated in [20]. The results then imply the existence of time derivative  $u_t$  as a function in suitable spaces similar to Corollary 1.2. In the case of the obstacle problem the existence of the time derivative was established in [19]. Dong, Peng, Zhang and Zhou [10] gave a proof that  $D^2u \in L^2_{loc}$  with a sharp range  $1 < p < 3$ . This range of  $p$  can be recovered from assumption (i) of Theorem 1.1. In the global case, estimates for  $D(|Du|^{p-2} Du)$  have been derived by Cianchi and Maz'ya in [7].

If  $\gamma = 0$ , Eq. (1.1) is the normalized parabolic  $p$ -Laplace equation

$$u_t - \Delta_p^N u = 0$$

where  $\Delta_p^N u := \Delta u + (p - 2)\Delta_\infty^N u$  is the normalized or game theoretic  $p$ -Laplace operator. This equation arises from a two-player stochastic game with a fixed running time, see Manfredi et al. [21], or from image processing, see Does [9]. Banerjee and Garofalo [5, 6] studied the potential theoretic aspects and boundary regularity of the normalized  $p$ -Laplacian

evolution. These papers also contain Lipschitz regularity results for solutions to the normalized  $p$ -parabolic equation. The regularity method in [21] is global whereas in [23] a local game theoretic method is applied in this context. Later Jin and Silvestre [15] established  $C_{loc}^{1,\alpha}$ -regularity in space and  $C_{loc}^{0, \frac{1+\alpha}{2}}$ -regularity in time. In [13], Høeg and Lindqvist studied the second order Sobolev regularity for the normalized  $p$ -parabolic equation and showed that when  $\frac{6}{5} < p < \frac{14}{5}$ , the second order spatial derivatives  $D^2u$  and the time derivative  $u_t$  belong to  $L_{loc}^2$ . Moreover, they also proved that when  $1 < p < 2$ ,  $u_t$  also belongs to  $L_{loc}^2$ . In [3],  $C_{loc}^{1,\alpha}$ -regularity was established to the normalized  $p$ -parabolic equation with a source term. The work of Dong et al. [10] also applies to the normalized  $p$ -parabolic equation; in this case they obtained  $D^2u \in L_{loc}^{2+\delta}$  and  $u_t \in L_{loc}^{2+\delta}$  for some  $\delta > 0$  if  $1 < p < 3 + \frac{2}{n-2}$ . The key result of [10] with  $\delta = 0$  can be recovered from assumption (ii) of Theorem 1.1. Recently Andrade and Santos [1] established improved Sobolev regularity estimates when  $p$  is close to 2.

As stated, (1.1) is in non-divergence form and can be highly degenerate or singular. Thus even defining viscosity solutions in such a way that existence and uniqueness can be obtained becomes a nontrivial issue. This was done by Ohnuma and Sato [22], see also Giga’s monograph [12]. For viscosity solutions to the general Eq. (1.1), where  $1 < p < \infty$  and  $-1 < \gamma < \infty$  are allowed to be independent of each other, Imbert, Jin and Silvestre [14] proved in particular that  $Du \in C_{loc}^\alpha$  for suitable  $0 < \alpha < 1$ . In [24], Parviainen and Vázquez established Harnack’s inequality and asymptotic behaviour by using the fact that for radial solutions Eq. (1.1) is equivalent to a divergence form equation but in fictitious dimension. Attouchi [2] in the degenerate case and Attouchi-Ruosteenoja [4] in the singular case established spatial  $C_{loc}^{1,\alpha}$ -regularity for an equation of type (1.1) but with a source term. The elliptic Harnack’s inequality in the singular range was obtained in [17].

This article is organized as follows. In Sect. 2 we provide the necessary preliminaries. In Sect. 3 we explain the ideas of the proof of Theorem 1.1. In Sect. 4 we state several auxiliary lemmas needed in the proofs, including the fundamental inequality (1.3). Sections 5 and 6 are parallel to each other. In the former, we provide the formal calculation. In the latter, we provide a similar calculation in a regularized setting, which eventually yields Theorem 1.1. In Sect. 6.2 we prove Theorem 1.1 and Corollary 1.2. Some of the proofs for the technical lemmas are postponed to the appendix.

## 2 Preliminaries

We use the following notation. Let  $\Omega \subset \mathbb{R}^n$ ,  $n \geq 2$ , be a domain and define the cylinder

$$\Omega_T := \Omega \times (0, T).$$

If  $U$  is compactly contained in  $\Omega$ , i.e.  $U \subset \Omega$  and the closure of  $U$  is a compact subset of  $\Omega$ , we write  $U \Subset \Omega$ . For  $0 < t_1 < t_2 < \infty$ , we set

$$U_{t_1,t_2} := U \times (t_1, t_2).$$

Moreover, we will use parabolic cylinders of the form

$$Q_r(x_0, t_0) := B_r(x_0) \times (t_0 - r^2, t_0],$$

where  $B_r(x_0)$  denotes the open ball with radius  $r > 0$  and center point  $x_0 \in \Omega$ . When no confusion arises, we may drop the reference point  $(x_0, t_0)$  and write  $Q_r$ .

Given a function  $u = u(x, t)$  of point  $x \in \mathbb{R}^n$  and time  $t > 0$ , the spatial gradient of  $u$  is denoted by  $Du = (u_{x_1}, \dots, u_{x_n})$ , and the time derivative by  $u_t$ . The Hessian matrix of  $u$  is denoted by  $D^2u = (u_{x_i x_j})_{i,j=1}^n$ . The Laplacian of  $u$  is given by

$$\Delta u := \sum_{i=1}^n u_{x_i x_i}$$

and the infinity Laplacian by

$$\Delta_\infty u := \sum_{i,j=1}^n u_{x_i} u_{x_j} u_{x_i x_j} = \langle Du, D^2u Du \rangle$$

where  $\langle \cdot, \cdot \rangle$  stands for the inner product in  $\mathbb{R}^n$ . The normalized infinity Laplacian is denoted by

$$\Delta_\infty^N u := \frac{\Delta_\infty u}{|Du|^2}.$$

We study viscosity solutions to the general  $p$ -parabolic equation

$$u_t - |Du|^\gamma (\Delta u + (p - 2)\Delta_\infty^N u) = 0 \quad \text{in } \Omega_T, \tag{2.1}$$

where  $1 < p < \infty$  and  $-1 < \gamma < \infty$ . The definition of suitable viscosity solutions to (2.1) requires some care because the operator may be singular. Nonetheless, a definition that fits our needs can be found in [22]. First set

$$F(Du, D^2u) := |Du|^\gamma (\Delta u + (p - 2)\Delta_\infty^N u)$$

whenever  $Du \neq 0$ . We define  $\mathcal{F}$  to be a set of functions  $f \in C^2([0, \infty))$  such that

$$f(0) = f'(0) = f''(0) = 0, \quad f''(r) > 0 \text{ for all } r > 0,$$

and moreover we require for  $g(x) := f(|x|)$  that

$$\lim_{x \rightarrow 0, x \neq 0} F(Dg(x), D^2g(x)) = 0.$$

Further, let

$$\Sigma = \{\sigma \in C^1(\mathbb{R}) : \sigma \text{ is even, } \sigma(0) = \sigma'(0) = 0, \text{ and } \sigma(r) > 0 \text{ for all } r \neq 0\}.$$

**Definition 2.1** A function  $\varphi \in C^2(\Omega_T)$  is admissible if for any  $(x_0, t_0) \in \Omega_T$  with  $D\varphi(x_0, t_0) = 0$ , there are  $\delta > 0$ ,  $f \in \mathcal{F}$  and  $\sigma \in \Sigma$  such that

$$|\varphi(x, t) - \varphi(x_0, t_0) - \varphi_t(x_0, t_0)(t - t_0)| \leq f(|x - x_0|) + \sigma(t - t_0)$$

for all  $(x, t) \in B_\delta(x_0) \times (t_0 - \delta, t_0 + \delta)$ .

If  $D\varphi \neq 0$ , a  $C^2$ -function is automatically admissible.

**Definition 2.2** We say that  $\varphi$  touches  $u$  at  $(x_0, t_0) \in \Omega_T$  (strictly) from below if

- (1)  $u(x_0, t_0) = \varphi(x_0, t_0)$ , and
- (2)  $u(x, t) > \varphi(x, t)$  for all  $(x, t) \in \Omega_T$  such that  $(x, t) \neq (x_0, t_0)$ .

The definition for touching (strictly) from above is analogous.

**Definition 2.3** A function  $u : \Omega_T \rightarrow \mathbb{R} \cup \{\infty\}$  is a viscosity supersolution to (2.1) if

- (i)  $u$  is lower semicontinuous,
- (ii)  $u$  is finite in a dense subset of  $\Omega_T$ ,
- (iii) for all admissible  $\varphi \in C^2(\Omega_T)$  touching  $u$  at  $(x_0, t_0) \in \Omega_T$  from below

$$\begin{cases} \varphi_t(x_0, t_0) - F(D\varphi(x_0, t_0), D^2\varphi(x_0, t_0)) \geq 0 & \text{if } D\varphi(x_0, t_0) \neq 0, \\ \varphi_t(x_0, t_0) \geq 0 & \text{if } D\varphi(x_0, t_0) = 0. \end{cases}$$

The definition of a subsolution  $u : \Omega_T \rightarrow \mathbb{R} \cup \{-\infty\}$  is analogous except that we require upper semicontinuity, touching from above, and we reverse the inequalities above: in other words if  $-u$  is a viscosity supersolution. If a continuous function is both a viscosity super- and subsolution, it is a *viscosity solution*.

It is shown in [16] that if  $\gamma = p - 2 > -1$ , then the above notion coincides with the notion of  $p$ -super/subparabolic functions, having a direct connection to the distributional weak super/subsolutions as well. Moreover, if  $\gamma \geq 0$ , then viscosity solutions can be defined in a standard way by using semicontinuous envelopes, see Proposition 2.2.8 in [12].

### 3 Plan of proof

In this section we explain the idea of the proof of Theorem 1.1 and our plan of the proof.

#### 3.1 Derivation of a basic estimate

In order to prove second order estimates, we first derive a key basic estimate (3.4) (or actually equality at this point). To this end, we regularize the original Eq. (1.1) and consider

$$u_t^\varepsilon - (|Du^\varepsilon|^2 + \varepsilon)^{\gamma/2} \left( \Delta u^\varepsilon + (p - 2) \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \right) = 0 \tag{3.1}$$

for small  $\varepsilon > 0$ . Solutions to this equation are smooth according to the standard theory. We differentiate Eq. (3.1) with respect to  $x_k, k = 1, \dots, n$ , and find that the spatial partial derivatives  $u_{x_k}^\varepsilon, k = 1, \dots, n$ , solve the equation

$$\begin{aligned} & (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2-\gamma}{2}} (u_{x_k}^\varepsilon)_t - \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2}{2}} A Du_{x_k}^\varepsilon \right) \\ & + (p - 2 - \gamma) (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-4-\gamma}{2}} u_t^\varepsilon \langle Du^\varepsilon, Du_{x_k}^\varepsilon \rangle = 0 \end{aligned} \tag{3.2}$$

where

$$A = I + (p - 2) \frac{Du^\varepsilon \otimes Du^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon}$$

is a uniformly positive definite  $n \times n$ -matrix. Here  $I$  denotes the identity matrix.

We continue with the intention to study the derivatives of  $|Du|^{p-2+s} Du$ ; in particular the choice  $s = 2 - p$  corresponds to  $D^2u$ . We multiply the differentiated Eq. (3.2) by  $(|Du^\varepsilon|^2 + \varepsilon)^{s/2} u_{x_k}^\varepsilon$  and obtain

$$\begin{aligned} & (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2-\gamma+s}{2}} u_{x_k}^\varepsilon (u_{x_k}^\varepsilon)_t - (|Du^\varepsilon|^2 + \varepsilon)^{s/2} u_{x_k}^\varepsilon \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2}{2}} A Du_{x_k}^\varepsilon \right) \\ & + (p - 2 - \gamma) (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-4-\gamma+s}{2}} u_t^\varepsilon \langle Du^\varepsilon, Du_{x_k}^\varepsilon \rangle u_{x_k}^\varepsilon = 0. \end{aligned} \tag{3.3}$$



Using the chain rule

$$u_{x_k}^\varepsilon (u_{x_k}^\varepsilon)_t = \frac{1}{2} \left( (u_{x_k}^\varepsilon)^2 + \frac{\varepsilon}{n} \right)_t,$$

and summing (3.3) over  $k = 1, \dots, n$  gives that

$$\begin{aligned} & \frac{(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p+s-\gamma}{2}}}{p+s-\gamma} \Big|_t - (|Du^\varepsilon|^2 + \varepsilon)^{s/2} \sum_{k=1}^n u_{x_k}^\varepsilon \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2}{2}} ADu_{x_k}^\varepsilon \right) \\ & + (p-2-\gamma)(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2-\gamma+s}{2}} u_t^\varepsilon \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} = 0. \end{aligned}$$

Observing that

$$\begin{aligned} & \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} AD^2 u^\varepsilon Du^\varepsilon \right) \\ & = \sum_{k=1}^n \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{s/2} u_{x_k}^\varepsilon (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2}{2}} ADu_{x_k}^\varepsilon \right) \\ & = (|Du^\varepsilon|^2 + \varepsilon)^{s/2} \sum_{k=1}^n u_{x_k}^\varepsilon \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2}{2}} ADu_{x_k}^\varepsilon \right) \\ & \quad + (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ |D^2 u^\varepsilon|^2 + (p-2+s) \frac{|D^2 u^\varepsilon Du^\varepsilon|^2}{|Du^\varepsilon|^2 + \varepsilon} + s(p-2) \frac{(\Delta_\infty u^\varepsilon)^2}{(|Du^\varepsilon|^2 + \varepsilon)^2} \right\}, \end{aligned}$$

we obtain the identity

$$\begin{aligned} & (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ |D^2 u^\varepsilon|^2 + (p-2+s) \frac{|D^2 u^\varepsilon Du^\varepsilon|^2}{|Du^\varepsilon|^2 + \varepsilon} + s(p-2) \frac{(\Delta_\infty u^\varepsilon)^2}{(|Du^\varepsilon|^2 + \varepsilon)^2} \right. \\ & \quad \left. + (p-2-\gamma)(|Du^\varepsilon|^2 + \varepsilon)^{-\gamma/2} u_t^\varepsilon \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \right\} \tag{3.4} \\ & = \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} AD^2 u^\varepsilon Du^\varepsilon \right) - \frac{(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p+s-\gamma}{2}}}{p+s-\gamma} \Big|_t. \end{aligned}$$

Here we assume that  $s \neq \gamma - p$ . This is not restrictive, because eventually such value of  $s$  violates the resulting range condition (1.4) in any case. It is important that the terms on the right hand side are in divergence form and can thus be well estimated. An important step towards the desired result would be a pointwise inequality

$$(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2 u^\varepsilon|^2 \lesssim \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} AD^2 u^\varepsilon Du^\varepsilon \right) - \frac{(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p+s-\gamma}{2}}}{p+s-\gamma} \Big|_t,$$

which then could be integrated to obtain the final result and for this we need to estimate the excess terms on the left hand side of (3.4).

### 3.2 Formal calculation for smooth solutions with a nonvanishing gradient

Compared to our earlier work [11] where we treated the case  $\gamma = p - 2$ , we now have two extra difficulties for the general case  $-1 < \gamma < \infty$ . The first difficulty arises from the fourth term on the left hand side of (3.4), that is,

$$(p-2-\gamma)(|Du^\varepsilon|^2 + \varepsilon)^{-\gamma/2} u_t^\varepsilon \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon}.$$

Note that this mixed term vanishes if  $\gamma = p - 2$ . In general we regard the term mixed in the sense that we cannot determine its sign by the sign of the coefficient  $p - 2 - \gamma$ .

We first discuss the difficulty of mixed terms in the formal case with  $\varepsilon = 0$ , and denote a solution by  $u$ . In this case, we assume in addition that  $Du \neq 0$ . As indicated above, we would like to estimate the excess term in (3.4) and obtain an estimate for  $|Du|^{p-2+s}|D^2u|^2$  with the range (1.4). To this end, we write the fundamental inequality (1.3) in the form

$$2|D_T|Du|^2 + \frac{(\Delta_T u)^2}{n-1} + (\Delta_\infty^N u)^2 \leq |D^2u|^2$$

and employ it in identity (3.4) on the term  $|Du|^{p-2+s}|D^2u|^2$  to obtain that

$$\begin{aligned} &|Du|^{p-2+s} \left\{ \frac{1}{n-1}(\Delta_T u)^2 + (p+s)|D_T|Du|^2 + (p-1)(s+1)(\Delta_\infty^N u)^2 \right. \\ &\quad \left. + (p-2-\gamma)|Du|^{-\gamma} u_t \Delta_\infty^N u \right\} \\ &\leq \operatorname{div} (|Du|^{p-2+s} A D^2u Du) - \frac{(|Du|^{p+s-\gamma})_t}{p+s-\gamma}, \end{aligned} \tag{3.5}$$

where

$$|D_T|Du|^2 := \frac{|D^2u Du|^2}{|Du|^2} - (\Delta_\infty^N u)^2 \quad \text{and} \quad \Delta_T u := \Delta u - \Delta_\infty^N u.$$

Note that  $|D_T|Du|^2 \geq 0$ . Sometimes  $\Delta_T u$  is called the normalized 1-Laplacian for the obvious reason.

Except the mixed term that is the last term on the left hand side in (3.5), the nonnegativity of other terms in the left hand side of (3.5) can be easily obtained by the restriction  $s > -1$ . In order to develop a systematic way of checking nonnegativity of the mixed term utilizing other terms, we use Eq. (1.1) to rewrite

$$|Du|^{-\gamma} u_t \Delta_\infty^N u = \Delta_T u \Delta_\infty^N u + (p-1)(\Delta_\infty^N u)^2,$$

and view the mixed term  $\Delta_T u \Delta_\infty^N u$  as a part of a quadratic form of  $\Delta_T u$  and  $\Delta_\infty^N u$ . That is, we consider

$$\begin{aligned} Q &:= \frac{1}{n-1}(\Delta_T u)^2 + (p-1)(p-1+s-\gamma)(\Delta_\infty^N u)^2 + (p-2-\gamma)\Delta_T u \Delta_\infty^N u \\ &=: \langle \bar{x}, M \bar{x} \rangle, \end{aligned}$$

where  $\bar{x} := (\Delta_T u, \Delta_\infty^N u)^T \in \mathbb{R}^2$  and

$$M = \begin{bmatrix} \frac{1}{n-1} & \frac{1}{2}(p-2-\gamma) \\ \frac{1}{2}(p-2-\gamma) & (p-1)(p-1+s-\gamma) \end{bmatrix}$$

is a symmetric  $2 \times 2$ -matrix.

It turns out that in order to derive the desired estimate, it suffices to ensure along with few other conditions that the quadratic form  $Q$  is strictly positive in  $\mathbb{R}^2 \setminus \{0\}$ , that is,  $M$  is positive definite. However, the range condition in (1.4) does not suffice to guarantee that the positive definiteness of  $Q$ , hence we need to improve the estimate. We employ the following observation: If  $q > 1$ , then

$$u_t |Du|^{q-2} \Delta_q^N u = u_t \operatorname{div} (|Du|^{q-2} Du) = \operatorname{div} (u_t |Du|^{q-2} Du) - \frac{(|Du|^q)_t}{q} \tag{3.6}$$

holds for any smooth function  $u$  with nonvanishing gradient. In other words, the quantity on the left hand side is a ‘good term’ with a hidden divergence structure.

It is easier to utilize this observation with inequality (3.5), if we rewrite the right hand side of that inequality using Eq. (1.1). To be more precise,

$$\begin{aligned} & \operatorname{div} (|Du|^{p-2+s} AD^2uDu) - \frac{(|Du|^{p+s-\gamma})_t}{p+s-\gamma} \\ &= \operatorname{div} (|Du|^{p-2+s} (D^2uDu - \Delta uDu)) + u_t \operatorname{div} (|Du|^{p-2+s-\gamma} Du), \end{aligned} \tag{3.7}$$

where the last term now matches with (3.6) setting  $q := p + s - \gamma$ . On the other hand, for a solution  $u$ , by Eq. (1.1), and by the definition of normalized  $q$ -Laplacian  $\Delta_q^N u$ , one has

$$u_t = |Du|^\gamma (\Delta_T u + (p - 1)\Delta_\infty^N u), \quad \text{and} \quad \Delta_{p+s-\gamma}^N u = \Delta_T u + (p - 1 + s - \gamma)\Delta_\infty^N u$$

and thus

$$\begin{aligned} & u_t \operatorname{div} (|Du|^{p-2+s-\gamma} Du) \\ &= |Du|^\gamma (\Delta_T u + (p - 1)\Delta_\infty^N u) \cdot |Du|^{p-2+s-\gamma} \cdot (\Delta_T u + (p - 1 + s - \gamma)\Delta_\infty^N u) \\ &= |Du|^{p-2+s} \left\{ (\Delta_T u)^2 + (2p - 2 + s - \gamma)\Delta_T u \Delta_\infty^N u + (p - 1)(p - 1 + s - \gamma)(\Delta_\infty^N u)^2 \right\}. \end{aligned} \tag{3.8}$$

The idea is to add  $u_t \operatorname{div} (|Du|^{p-2+s-\gamma} Du)$  with a suitable weight on both sides of (3.5): then by the above equation, it produces new coefficients on the left hand side that can be utilized later to get better range, and controllable terms on the right hand side by (3.7). We also add another positive weight by using

$$\begin{aligned} & |Du|^{p-2+s} \left\{ |D^2u|^2 - (\Delta u)^2 + (p - 2 + s) \frac{|D^2uDu|^2}{|Du|^2} - (p - 2 + s)\Delta u \Delta_\infty^N u \right\} \\ &= \operatorname{div} (|Du|^{p-2+s} (D^2uDu - \Delta uDu)) \end{aligned} \tag{3.9}$$

from Lemma 4.2 below which holds for any smooth function with nonvanishing gradient. This allows us to obtain simplified coefficients in intermediate steps. Thus we obtain

$$\begin{aligned} & |Du|^{p-2+s} \left\{ \left( w_2 - \frac{n-2}{n-1} w_1 \right) (\Delta_T u)^2 + w_1(p+s) |D_T |Du||^2 \right. \\ & \quad \left. + w_2(p-1)(p-1+s-\gamma)(\Delta_\infty^N u)^2 + (w_2(2p-2+s-\gamma) - w_1(p+s)) \Delta_T u \Delta_\infty^N u \right\} \\ & \leq w_1 \operatorname{div} (|Du|^{p-2+s} (D^2uDu - \Delta uDu)) + w_2 u_t \operatorname{div} (|Du|^{p-2+s-\gamma} Du), \end{aligned} \tag{3.10}$$

which reduces to (3.5) if  $w_1 = 1$  and  $w_2 = 1$ . Calculations reveal that if the range condition (1.4) holds, then the weights  $w_1$  and  $w_2$  can be adjusted so that the weighted quadratic form

$$\begin{aligned} & \left( w_2 - \frac{n-2}{n-1} w_1 \right) (\Delta_T u)^2 + w_2(p-1)(p-1+s-\gamma)(\Delta_\infty^N u)^2 \\ & \quad + (w_2(2p-2+s-\gamma) - w_1(p+s)) \Delta_T u \Delta_\infty^N u \end{aligned}$$

is positive in  $\mathbb{R}^2 \setminus \{0\}$ . This positivity in the formal case  $\varepsilon = 0$  is shown in Lemma 5.2. By Proposition 5.1, this then implies the desired estimate

$$\int_{Q_r} \left| D(|Du|^{\frac{p-2+s}{2}} Du) \right|^2 dxdt \leq \frac{C}{r^2} \left( \int_{Q_{2r}} |Du|^{p+s} dxdt + \int_{Q_{2r}} |Du|^{p+s-\gamma} dxdt \right).$$

Heuristically, in order to prove the above estimate, and setting  $s = 2 - p$  for simplicity, we could have left a small piece of  $|D^2u|^2$  when applying the fundamental inequality for (3.5). Then the rest of the terms can be dropped by the above positivity result: in detail this is implemented in Lemma 4.4 also for other values of  $s$ . The obtained pointwise estimate can then be integrated by parts along with a cutoff function to get Proposition 5.1.

### 3.3 Solutions without smoothness assumptions and regularized equation

The second difficulty, which is related to the regularization, is that the left hand side of (3.4) consists of regularized versions of second order derivative quantities,

$$\frac{|D^2u^\varepsilon Du^\varepsilon|^2}{|Du^\varepsilon|^2 + \varepsilon} \quad \text{and} \quad \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon},$$

whereas employing the fundamental inequality (1.3) results in quantities like

$$\frac{|D^2u^\varepsilon Du^\varepsilon|^2}{|Du^\varepsilon|^2} \quad \text{and} \quad \Delta_\infty^N u^\varepsilon.$$

This mismatch causes that some of the formal calculations do not work as such but have further complications: in particular positive definiteness of the quadratic form becomes an issue.

For a certain range of parameters, the main result is obtained by a straightforward generalization of the formal calculation ( $\varepsilon = 0$ ) in the previous section. However, in the process of extending the range, we consider

$$\begin{aligned} S := & w_1 \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} (D^2u^\varepsilon Du^\varepsilon - \Delta u^\varepsilon Du^\varepsilon) \right) \\ & + w_2 u_t^\varepsilon \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s-\gamma}{2}} Du^\varepsilon \right) \\ & + w_3 \varepsilon \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}-1} (D^2u^\varepsilon Du^\varepsilon - \Delta u^\varepsilon Du^\varepsilon) \right) \\ & + w_4 \varepsilon u_t^\varepsilon \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s-\gamma}{2}-1} Du^\varepsilon \right), \end{aligned} \tag{3.11}$$

where  $w_1, w_2, w_3, w_4 \in \mathbb{R}$ . Compared to the right hand side of (3.10), or (6.3), this sum has two additional terms with weights  $w_3$  and  $w_4$ . The latter additional term has a hidden divergence structure, similarly to (3.6). These divergence structures can be used to adjust the coefficients on the left hand side of the estimate (3.10), and thus to improve the range of parameters. To be more precise, we denote

$$\theta := \frac{|Du^\varepsilon|^2}{|Du^\varepsilon|^2 + \varepsilon} \quad \text{and} \quad \kappa := 1 - \theta = \frac{\varepsilon}{|Du^\varepsilon|^2 + \varepsilon}, \tag{3.12}$$

and obtain

$$\frac{|D^2u^\varepsilon Du^\varepsilon|^2}{|Du^\varepsilon|^2 + \varepsilon} = \theta |D |Du^\varepsilon||^2 \quad \text{and} \quad \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} = \theta \Delta_\infty^N u^\varepsilon.$$

The second mixed term of (3.11) can also be written as a part of the quadratic form as follows

$$\begin{aligned} &\varepsilon u_t^\varepsilon \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s-\gamma}{2}-1} Du^\varepsilon \right) \\ &= \theta (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left( (\Delta_T u^\varepsilon)^2 + ((2p - 6 + s - \gamma)\theta + 2) \Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon \right. \\ &\quad \left. + ((p - 2)\theta + 1)((p - 4 + s - \gamma)\theta + 1) (\Delta_\infty^N u^\varepsilon)^2 \right) \end{aligned} \tag{3.13}$$

with weight  $w_4$ , where

$$u_t^\varepsilon = (|Du^\varepsilon|^2 + \varepsilon)^{\gamma/2} \left( \Delta_T u^\varepsilon + ((p - 2)\theta + 1) \Delta_\infty^N u^\varepsilon \right)$$

by using the regularized equation and recalling the shorthand notation  $\Delta_T u^\varepsilon := \Delta u^\varepsilon - \Delta_\infty^N u^\varepsilon$ . This will give rise to new coefficients and thus to a better range condition.

In order to produce new coefficients on the left hand side of (3.4), especially for the second order term  $(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2 u^\varepsilon|^2$ , and also to improve the range of the parameters, we add another divergence structure

$$\begin{aligned} &\varepsilon \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}-1} (D^2 u^\varepsilon Du^\varepsilon - \Delta u^\varepsilon Du^\varepsilon) \right) \\ &= \theta (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ |D^2 u^\varepsilon|^2 - (\Delta u^\varepsilon)^2 + (p - 4 + s)\theta |D| |Du^\varepsilon|^2 \right. \\ &\quad \left. - (p - 4 + s)\theta \Delta u^\varepsilon \Delta_\infty^N u^\varepsilon \right\}. \end{aligned} \tag{3.14}$$

Also observe that the above choice of the power  $(p - 2 + s)/2 - 1$  will be useful in the proof of Lemma 4.5 when deriving an upper bound for the left hand side of the estimate, after integration by parts where we estimate  $\varepsilon/(|Du^\varepsilon|^2 + \varepsilon) \leq 1$  and thus the additional  $-1$  in the power gets canceled out. Besides, the error terms obtained in Lemmas 4.4 and 4.7 in [10] can be seen as special cases of the error terms above.

Then combining (3.8), (3.9), (3.13) and (3.14) together with definition (3.11) of  $S$ , we get

$$\begin{aligned} &( |Du^\varepsilon|^2 + \varepsilon )^{\frac{p-2+s}{2}} \left\{ c_1 |D^2 u^\varepsilon|^2 + c_2 |D_T |Du^\varepsilon|^2 + (c_3 - c_1) (\Delta_T u^\varepsilon)^2 \right. \\ &\quad \left. + \left( (c_3 + c_4) ((p - 2)\theta + 1) - c_1 \right) (\Delta_\infty^N u^\varepsilon)^2 \right. \\ &\quad \left. + \left( c_3 ((p - 2)\theta + 1) + (c_3 + c_4) - (2c_1 + c_2) \right) \Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon \right\} = S, \end{aligned} \tag{3.15}$$

where  $c_1, c_2, c_3$  and  $c_4$  depend on  $w_1, w_2, w_3, w_4$  and  $\theta$  as computed in detail in Sect. 4.2. Then we again use the fundamental inequality on part of  $c_1 |D^2 u^\varepsilon|^2$  and find such weights  $w_1, w_2, w_3$  and  $w_4$  that the last three terms on the left hand side can be interpreted as a positive definite quadratic form and thus removed. Finally,  $S$  on the right hand side can be multiplied by a cutoff function and integrated by parts to get the final estimate. However, the nonnegativity can only be checked in certain ranges, since it needs to hold uniformly for all  $\theta \in [0, 1)$ .

### 4 Hidden divergence structures, the key estimate and auxiliary lemmas

In this section we prove several auxiliary tools. The lemmas in this section will be used to prove estimates for both  $u^\varepsilon$ , that solves (3.1) with  $\varepsilon > 0$ , and  $u$ , that solves (3.1) with  $\varepsilon = 0$

and  $Du \neq 0$ . Therefore we state the lemmas in such a generality that applies to both of these cases.

### 4.1 Hidden divergence structures

In this subsection we gather some useful facts about generic smooth functions. First, if  $u: \Omega_T \rightarrow \mathbb{R}$ ,  $\Omega_T \subset \mathbb{R}^{n+1}$  is a smooth function, then  $|Du|$  is locally Lipschitz continuous and thus, by Rademacher’s theorem, differentiable almost everywhere on each time slice. Here and in similar occurrences in what follows, we write that  $D|Du|$  exists *almost everywhere in space*.

Note that if  $(x_0, t_0) \in \Omega_T$  is a space-time point where  $|Du|$  is differentiable and  $Du(x_0, t_0) = 0$ , then  $D|Du|(x_0, t_0) = 0$ . Indeed, if we had  $D|Du|(x_0, t_0) \neq 0$ , then we could find a point  $\xi \in \Omega \times \{t_0\}$  (close to  $(x_0, t_0)$ ) such that  $|Du|(\xi) < 0$ , which is obviously impossible. On the other hand, if  $Du(x_0, t_0) \neq 0$  for some  $(x_0, t_0) \in \Omega_T$ , then  $|Du|$  is differentiable at  $(x_0, t_0)$  and

$$D|Du|(x_0, t_0) = \frac{D^2u(x_0, t_0)Du(x_0, t_0)}{|Du(x_0, t_0)|}.$$

For each point in  $\Omega_T$  where  $Du \neq 0$ , we fix an orthonormal basis of  $\mathbb{R}^n$ ,  $\{e_1, \dots, e_n\}$ , such that  $e_n = \frac{Du}{|Du|}$ . Hence we have, for those points where  $Du \neq 0$ ,

$$\frac{D^2uDu}{|Du|} = \langle e_1, D|Du| \rangle e_1 + \dots + \langle e_{n-1}, D|Du| \rangle e_{n-1} + \left\langle \frac{Du}{|Du|}, D|Du| \right\rangle \frac{Du}{|Du|}.$$

For those points where  $|Du|$  is differentiable, let us define the part of  $D|Du|$  which is tangential to the spatial level sets of  $u$  as

$$D_T|Du| := \begin{cases} \langle e_1, D|Du| \rangle e_1 + \dots + \langle e_{n-1}, D|Du| \rangle e_{n-1} & \text{if } Du \neq 0, \\ 0 & \text{if } Du = 0, \end{cases}$$

and its orthogonal counterpart, the normalized infinity Laplacian, as

$$\Delta_\infty^N u := \begin{cases} \left\langle \frac{Du}{|Du|}, D|Du| \right\rangle = \frac{\Delta_\infty u}{|Du|^2} & \text{if } Du \neq 0, \\ 0 & \text{if } Du = 0. \end{cases}$$

We employ these notation to write

$$|D|Du||^2 = |D_T|Du||^2 + (\Delta_\infty^N u)^2 \quad \text{a.e. in space in } \Omega_T, \tag{4.1}$$

and

$$\Delta_T u = \Delta u - \Delta_\infty^N u \quad \text{a.e. in } \Omega_T. \tag{4.2}$$

**Lemma 4.1** (Fundamental inequality) *Let  $u: \Omega_T \rightarrow \mathbb{R}$  be a smooth function. Then*

$$|D^2u|^2 \geq 2|D_T|Du||^2 + \frac{(\Delta_T u)^2}{n-1} + (\Delta_\infty^N u)^2 \quad \text{a.e. in space in } \Omega_T. \tag{4.3}$$

If  $n = 2$ , we have equality in the place of inequality.

For the proof of Lemma 4.1, we refer to [11, 25].

The following lemmas show that certain terms that first appear to be in non-divergence form, can actually be expressed in a divergence form. On the other hand, these structures can

be utilized in tuning the coefficients in the quadratic form as explained in Sect. 3.3, and thus they improve the range we obtain. The first Lemma 4.2 will mainly adjust the coefficient of the term  $(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2u^\varepsilon|^2$ . The second divergence structure, Lemma 4.3, will produce certain new coefficients on the quadratic form as  $Q$ . The proofs of both of these lemmas are direct calculations.

**Lemma 4.2** (Hidden divergence structure 1) *Let  $u : \Omega_T \rightarrow \mathbb{R}$  be a smooth function. Then for any  $\alpha \in \mathbb{R}$  and  $\varepsilon > 0$ ,*

$$\begin{aligned} & (|Du|^2 + \varepsilon)^{\alpha/2} \left\{ |D^2u|^2 - (\Delta u)^2 + \alpha \frac{|D^2uDu|^2}{|Du|^2 + \varepsilon} - \alpha \Delta u \frac{\Delta_\infty u}{|Du|^2 + \varepsilon} \right\} \\ &= \operatorname{div} \left( (|Du|^2 + \varepsilon)^{\alpha/2} (D^2uDu - \Delta uDu) \right). \end{aligned}$$

Furthermore, if  $Du \neq 0$ , then the above equality holds also for  $\varepsilon = 0$ .

**Proof** By the derivative rule of composite function, the right hand side

$$\begin{aligned} & \operatorname{div} \left( (|Du|^2 + \varepsilon)^{\alpha/2} (D^2uDu - \Delta uDu) \right) \\ &= \langle D^2uDu - \Delta uDu, D((|Du|^2 + \varepsilon)^{\alpha/2}) \rangle + (|Du|^2 + \varepsilon)^{\alpha/2} \operatorname{div} (D^2uDu - \Delta uDu) \\ &= \langle D^2uDu - \Delta uDu, D((|Du|^2 + \varepsilon)^{\alpha/2}) \rangle + (|Du|^2 + \varepsilon)^{\alpha/2} (|D^2u|^2 - (\Delta u)^2) \\ &= (|Du|^2 + \varepsilon)^{\alpha/2} \left\{ |D^2u|^2 - (\Delta u)^2 + \alpha \frac{|D^2uDu|^2}{|Du|^2 + \varepsilon} - \alpha \Delta u \frac{\Delta_\infty u}{|Du|^2 + \varepsilon} \right\}, \end{aligned}$$

where

$$D((|Du|^2 + \varepsilon)^{\alpha/2}) = \alpha(|Du|^2 + \varepsilon)^{\frac{\alpha-2}{2}} D^2uDu.$$

□

The next lemma demonstrates that a mixed term can be written in a divergence form. On the other hand by using Eq. (3.1), as explained in (3.13), the mixed term adds up in the quadratic form, and thus adding such mixed terms can be used to improve the range.

**Lemma 4.3** (Hidden divergence structure 2) *Let  $u : \Omega_T \rightarrow \mathbb{R}$  be a smooth function. Then for any  $\beta \in \mathbb{R}$  and  $\varepsilon > 0$ ,*

$$\begin{aligned} & u_t (|Du|^2 + \varepsilon)^{\beta/2} \left( \Delta u + \beta \frac{\Delta_\infty u}{|Du|^2 + \varepsilon} \right) \\ &= u_t \operatorname{div} \left( (|Du|^2 + \varepsilon)^{\beta/2} Du \right) \\ &= \begin{cases} \operatorname{div} \left( u_t (|Du|^2 + \varepsilon)^{\beta/2} Du \right) - \left( \frac{(|Du|^2 + \varepsilon)^{\frac{\beta+2}{2}}}{\beta + 2} \right)_t & \text{if } \beta \neq -2, \\ \operatorname{div} \left( u_t (|Du|^2 + \varepsilon)^{-1} Du \right) - \left( \frac{\ln(|Du|^2 + \varepsilon)}{2} \right)_t & \text{if } \beta = -2. \end{cases} \end{aligned}$$

Furthermore, if  $Du \neq 0$ , then the above equality holds also for  $\varepsilon = 0$ .

**Proof** We give the proof when  $\beta \neq -2$ , the second case is similar. By the derivative rule of composite function again, one has

$$\begin{aligned}
 & \operatorname{div} \left( u_t (|Du|^2 + \varepsilon)^{\beta/2} Du \right) - \left( \frac{(|Du|^2 + \varepsilon)^{\frac{\beta+2}{2}}}{\beta + 2} \right)_t \\
 &= u_t \operatorname{div} \left( (|Du|^2 + \varepsilon)^{\beta/2} Du \right) + (|Du|^2 + \varepsilon)^{\beta/2} Du Du_t - \left( \frac{(|Du|^2 + \varepsilon)^{\frac{\beta+2}{2}}}{\beta + 2} \right)_t \\
 &= u_t \operatorname{div} \left( (|Du|^2 + \varepsilon)^{\beta/2} Du \right) \\
 &= u_t \langle D(|Du|^2 + \varepsilon)^{\beta/2}, Du \rangle + u_t (|Du|^2 + \varepsilon)^{\beta/2} \operatorname{div} (Du) \\
 &= u_t (|Du|^2 + \varepsilon)^{\beta/2} \left( \Delta u + \beta \frac{\Delta_\infty u}{|Du|^2 + \varepsilon} \right).
 \end{aligned}$$

□

For  $\alpha \in \mathbb{R}$ , we denote the ‘first good divergence structure’ as

$$GD_1(\alpha) := \operatorname{div} \left( (|Du|^2 + \varepsilon)^{\alpha/2} (D^2u Du - \Delta u Du) \right)$$

and the ‘second good divergence structure’

$$GD_2(\alpha) := \begin{cases} \operatorname{div} \left( u_t (|Du|^2 + \varepsilon)^{\frac{\alpha-\gamma}{2}} Du \right) - \left( \frac{(|Du|^2 + \varepsilon)^{\frac{\alpha-\gamma+2}{2}}}{\alpha - \gamma + 2} \right)_t & \text{if } \alpha \neq \gamma - 2, \\ \operatorname{div} \left( u_t (|Du|^2 + \varepsilon)^{-1} Du \right) - \left( \frac{\ln(|Du|^2 + \varepsilon)}{2} \right)_t & \text{if } \alpha = \gamma - 2. \end{cases}$$

Then as explained in (3.11), we consider the following weighted sum of these ‘good structures’,

$$\begin{aligned}
 S := & w_1 GD_1(p - 2 + s) + w_2 GD_2(p - 2 + s) \\
 & + \varepsilon w_3 GD_1(p - 4 + s) + \varepsilon w_4 GD_2(p - 4 + s)
 \end{aligned} \tag{4.4}$$

for some parameter  $s \in \mathbb{R}$  and some weights  $w_1, w_2, w_3, w_4 \in \mathbb{R}$ . Observe that taking into account Lemmas 4.2 and 4.3, then  $S$  introduced above coincides with  $S$  in (3.11), i.e. the notation is consistent. The reason for using the mixed term form in  $S$  there was to emphasize the idea that we can improve the range by adding the mixed terms. To derive the final estimate, we need terms in the divergence form, and therefore this form was used in the above definition of  $S$ , but as stated they are equivalent.

### 4.2 The key estimate

As explained in (3.15),  $S$  represents the right hand side in our key estimate, and on the left we should have the second derivatives and a positive definite quadratic form. In this section, we derive the key estimate corresponding to (3.15) in detail.

We use Lemmas 4.2 and 4.3 to rewrite  $S$  as a linear combination of time derivatives and second order spatial derivative quantities, similarly to the left hand side of (3.4). First recall shorthand notation  $\theta$  and  $\kappa$  from (3.12)

$$\theta = \frac{|Du^\varepsilon|^2}{|Du^\varepsilon|^2 + \varepsilon} \quad \text{and} \quad \kappa = \frac{\varepsilon}{|Du^\varepsilon|^2 + \varepsilon},$$

thus  $0 \leq \theta, \kappa \leq 1, \theta + \kappa = 1$  and

$$\frac{|D^2u^\varepsilon Du^\varepsilon|^2}{|Du^\varepsilon|^2 + \varepsilon} = \theta |D |Du^\varepsilon||^2 \quad \text{and} \quad \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} = \theta \Delta_\infty^N u^\varepsilon.$$



In particular, if  $\varepsilon = 0$  and the gradient does not vanish, then  $\theta \equiv 1$  and  $\kappa \equiv 0$ . Next we recall the definition of  $S$  from the above, and use the good divergence structures i.e. Lemma 4.2 (with  $\alpha = p - 2 + s$  and  $p - 4 + s$ ) and Lemma 4.3 (with  $\beta = p - 2 + s - \gamma$  and  $p - 4 + s - \gamma$ ). For a smooth solution and indeed for any smooth function, we have

$$\begin{aligned}
 S = & w_1 (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ |D^2u^\varepsilon|^2 - (\Delta u^\varepsilon)^2 + (p - 2 + s) \left( \frac{|D^2u^\varepsilon Du^\varepsilon|^2}{|Du^\varepsilon|^2 + \varepsilon} \right. \right. \\
 & \left. \left. - \Delta u^\varepsilon \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \right) \right\} \\
 & + w_2 u_t^\varepsilon (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s-\gamma}{2}} \left\{ \Delta u^\varepsilon + (p - 2 + s - \gamma) \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \right\} \\
 & + \varepsilon w_3 (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-4+s}{2}} \left\{ |D^2u^\varepsilon|^2 - (\Delta u^\varepsilon)^2 + p - 4 + s \right\} \left( \frac{|D^2u^\varepsilon Du^\varepsilon|^2}{|Du^\varepsilon|^2 + \varepsilon} \right. \\
 & \left. - \Delta u^\varepsilon \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \right) \left\{ \right\} \\
 & + \varepsilon w_4 u_t^\varepsilon (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-4+s-\gamma}{2}} \left\{ \Delta u^\varepsilon + (p - 4 + s - \gamma) \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \right\}.
 \end{aligned}$$

Then by simplifying, we get

$$\begin{aligned}
 S = & (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ (w_1 + w_3\kappa) (|D^2u^\varepsilon|^2 - (\Delta u^\varepsilon)^2) \right. \\
 & + (w_1(p - 2 + s) + w_3(p - 4 + s)\kappa) \theta (|D|Du^\varepsilon|^2 - \Delta u^\varepsilon \Delta_\infty^N u^\varepsilon) \\
 & + (w_2 + w_4\kappa) (|Du^\varepsilon|^2 + \varepsilon)^{-\gamma/2} u_t^\varepsilon \Delta u^\varepsilon \\
 & \left. + (w_2(p - 2 + s - \gamma) + w_4(p - 4 + s - \gamma)\kappa) \theta (|Du^\varepsilon|^2 + \varepsilon)^{-\gamma/2} u_t^\varepsilon \Delta_\infty^N u^\varepsilon \right\} \\
 = & (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ c_1 (|D^2u^\varepsilon|^2 - (\Delta u^\varepsilon)^2) + c_2 (|D|Du^\varepsilon|^2 - \Delta u^\varepsilon \Delta_\infty^N u^\varepsilon) \right. \\
 & \left. + c_3 (|Du^\varepsilon|^2 + \varepsilon)^{-\gamma/2} u_t^\varepsilon \Delta u^\varepsilon + c_4 (|Du^\varepsilon|^2 + \varepsilon)^{-\gamma/2} u_t^\varepsilon \Delta_\infty^N u^\varepsilon \right\}
 \end{aligned}$$

almost everywhere in  $\Omega_T$ , where

$$\begin{cases} c_1 = w_1 + w_3\kappa, & c_2 = (w_1(p - 2 + s) + w_3(p - 4 + s)\kappa)\theta, \\ c_3 = w_2 + w_4\kappa, & c_4 = (w_2(p - 2 + s - \gamma) + w_4(p - 4 + s - \gamma)\kappa)\theta. \end{cases} \tag{4.5}$$

Observe that given  $p, \gamma$  and  $s$ , if  $\varepsilon = 0$ , then  $c_1, \dots, c_4$  reduce to constants that only depend on  $w_1$  and  $w_2$ , which shows that in smooth case by adjusting  $w_1$  and  $w_2$ , we can get the desired estimate as explained in (3.10).

By employing expressions (4.1) and (4.2), we can write

$$\begin{aligned}
 s = & (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ c_1 |D^2u^\varepsilon|^2 + c_2 |D_T|Du^\varepsilon|^2 - c_1 (\Delta_T u^\varepsilon)^2 - c_1 (\Delta_\infty^N u^\varepsilon)^2 \right. \\
 & - (2c_1 + c_2) \Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon + c_3 (|Du^\varepsilon|^2 + \varepsilon)^{-\gamma/2} u_t^\varepsilon \Delta_T u^\varepsilon \\
 & \left. + (c_3 + c_4) (|Du^\varepsilon|^2 + \varepsilon)^{-\gamma/2} u_t^\varepsilon \Delta_\infty^N u^\varepsilon \right\}
 \end{aligned} \tag{4.6}$$

almost everywhere in  $\Omega_T$ . Next we use regularized Eq. (3.1) to replace time derivatives  $u_t$  in (4.6) with spatial derivatives. Thus we arrive to the key estimate for a smooth solution to

the regularized equation (which is actually equality at this point)

$$\begin{aligned}
 & (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ c_1 |D^2 u^\varepsilon|^2 + c_2 |D_T |Du^\varepsilon||^2 + (c_3 - c_1) (\Delta_T u^\varepsilon)^2 \right. \\
 & \quad + ((c_3 + c_4) P_\theta - c_1) (\Delta_\infty^N u^\varepsilon)^2 \\
 & \quad \left. + (c_3 P_\theta + (c_3 + c_4) - (2c_1 + c_2)) \Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon \right\} = S,
 \end{aligned}
 \tag{4.7}$$

where

$$P_\theta := (p - 2)\theta + 1 \in (0, \infty)$$

for the sake of brevity. We rewrite this as

$$(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ c_1 |D^2 u^\varepsilon|^2 + c_2 |D_T |Du^\varepsilon||^2 + R \right\} = S \tag{4.8}$$

where

$$\begin{aligned}
 R := & (c_3 - c_1) (\Delta_T u^\varepsilon)^2 + ((c_3 + c_4) P_\theta - c_1) (\Delta_\infty^N u^\varepsilon)^2 \\
 & + (c_3 P_\theta + (c_3 + c_4) - (2c_1 + c_2)) \Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon
 \end{aligned}$$

is a quadratic form in variables  $\Delta_T u$  and  $\Delta_\infty^N u$ . We rewrite  $R$  as

$$R = \langle \bar{x}, N\bar{x} \rangle,$$

where  $\bar{x} = (\Delta_T u^\varepsilon, \Delta_\infty^N u^\varepsilon)^T \in \mathbb{R}^2$  and  $N \in \mathbb{R}^{2 \times 2}$  is a symmetric matrix whose entries  $N_{ij}$ ,  $i, j = 1, 2$ , are given by

$$\begin{cases} N_{11} & = c_2 - c_1 \\ N_{12} = N_{21} & = \frac{1}{2} (c_3 P_\theta + (c_3 + c_4) - (2c_1 + c_2)) \\ N_{22} & = (c_3 + c_4) P_\theta - c_1. \end{cases}$$

Note that

$$\|N\|_{L^\infty(\Omega_T)} := \sup\{|N(x, t)| : (x, t) \in \Omega_T\}$$

where

$$|N(x, t)| = \sqrt{(N_{11}(x, t))^2 + (N_{12}(x, t))^2 + (N_{21}(x, t))^2 + (N_{22}(x, t))^2},$$

has an upper bound that only depends on  $p, \gamma$  and  $s$  by fixing  $w_1, w_2, w_3$  and  $w_4$ .

### 4.3 Auxiliary lemmas

In this subsection we state two technical lemmas that can be used to conclude our main integral estimate.

We want to apply the fundamental inequality, Lemma 4.1, to estimate  $|D^2 u^\varepsilon|^2$  in (4.8) from below to improve the range condition by using terms we obtain in this application. However, the direct application will eliminate the full Hessian  $|D^2 u^\varepsilon|^2$  that we want to estimate. We could leave a small fraction of  $|D^2 u^\varepsilon|^2$  (like the method was first described at the end of Sect. 3.3 for simplicity) and apply the fundamental inequality only to a remaining part, but actually this will not be necessary: The next lemma shows that already a seemingly weaker lower bound is sufficient. This will simplify the exposition.

**Lemma 4.4** *Let  $u^\varepsilon: \Omega_T \rightarrow \mathbb{R}$  be a smooth solution to (3.1),  $S$  as in (4.4),  $c_1$  as in (4.5), and  $\varepsilon \geq 0$ . If  $\varepsilon = 0$ , we assume in addition that  $Du^\varepsilon \neq 0$ . Suppose that we can select  $w_1, w_2, w_3, w_4 \in \mathbb{R}$  such that  $c_1 = c_1(n, p, \gamma, s, w_1, w_2, w_3, w_4) > 0$ ,  $c = c(n, p, \gamma, s, w_1, w_2, w_3, w_4) > 0$  and*

$$(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ c |D_T |Du^\varepsilon||^2 + Q \right\} \leq S \quad \text{a.e. in space in } \Omega_T, \tag{4.9}$$

where

$$Q = \langle \bar{x}, M\bar{x} \rangle$$

with  $\bar{x} = (\Delta_T u^\varepsilon, \Delta_\infty^N u^\varepsilon)^T \in \mathbb{R}^2$  and a uniformly bounded positive definite (with a uniform constant) symmetric matrix  $M = M(n, p, \gamma, s, w_1, w_2, w_3, w_4) \in \mathbb{R}^{2 \times 2}$ . Then there is  $\lambda = \lambda(n, p, \gamma, s, w_1, w_2, w_3, w_4) > 0$  such that

$$\lambda(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2 u^\varepsilon|^2 \leq S \quad \text{a.e. in space in } \Omega_T.$$

**Proof** Recall that

$$\begin{aligned} S &= w_1 G D_1(p-2+s) + w_2 G D_2(p-2+s) \\ &\quad + \varepsilon w_3 G D_1(p-4+s) + \varepsilon w_4 G D_2(p-4+s) \end{aligned}$$

which, as pointed out in (4.8), can be written as

$$S = (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ c_1 |D^2 u^\varepsilon|^2 + c_2 |D_T |Du^\varepsilon||^2 + \langle \bar{x}, N\bar{x} \rangle \right\}$$

almost everywhere in space  $\Omega_T$ , where  $\bar{x}$  and  $N$  are as in (4.8). Observe that we utilized Eq. (3.1) at this step to get rid of the time derivatives.

For any  $\lambda \in (0, 1)$ , we write

$$S = \lambda S + (1 - \lambda)S$$

and use the assumption (4.9) to estimate  $(1 - \lambda)S$  from below. We end up with

$$\begin{aligned} S &\geq (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ \lambda c_1 |D^2 u^\varepsilon|^2 + (c + \lambda(c_2 - c)) |D_T |Du^\varepsilon||^2 \right. \\ &\quad \left. + \langle \bar{x}, (M + \lambda(N - M))\bar{x} \rangle \right\}. \end{aligned}$$

We claim that we can select  $\lambda > 0$  such that  $c + \lambda(c_2 - c) \geq 0$  and  $M + \lambda(N - M)$  is a positive definite matrix. Indeed, since  $c > 0$ , then

$$c + \lambda(c_2 - c) \geq c - \lambda \|c_2 - c\|_{L^\infty(\Omega_T)} > 0,$$

uniformly if  $\lambda = \lambda(n, p, \gamma, s, w_1, w_2, w_3, w_4) > 0$  is small enough. Next we recall that the boundedness and positive definiteness of  $M$  implies

$$\|M\|_{L^\infty(\Omega_T)} \leq C, \quad \text{and } M_{11} \geq c \quad \text{and } \det(M) \geq c \quad \text{in } \Omega_T$$

by Sylvester’s criterion and choosing small enough  $c > 0$ . For the positive definiteness of the matrix  $M + \lambda(N - M)$  we can use Sylvester’s criterion again and check that the leading principal minors are positive if  $\lambda > 0$  is small enough. The first principal minor is the upper-left corner entry, i.e.

$$(M + \lambda(N - M))_{11} = M_{11} + \lambda(N_{11} - M_{11}) \geq c - \lambda(\|N\|_{L^\infty(\Omega_T)} + \|M\|_{L^\infty(\Omega_T)}),$$

and the second principal minor is the determinant, i.e.

$$\begin{aligned} \det(M + \lambda(N - M)) &= \det(M) + \lambda(M_{11}N_{22} + M_{22}N_{11} - 2M_{12}N_{12} - 2\det(M)) \\ &\quad + \lambda^2 \det(N - M) \\ &\geq c - 2\lambda \left( 2\|M\|_{L^\infty(\Omega_T)}\|N\|_{L^\infty(\Omega_T)} + \|M\|_{L^\infty(\Omega_T)}^2 \right) \\ &\quad - \lambda^2\|N - M\|_{L^\infty(\Omega_T)}^2 \geq c - 4\lambda \left( \|M\|_{L^\infty(\Omega_T)} + \|N\|_{L^\infty(\Omega_T)} \right)^2. \end{aligned}$$

Hence we choose  $\lambda$  such that

$$0 < \lambda < \min \left\{ 1, \frac{c}{\|c_2 - c\|_{L^\infty(\Omega_T)}}, \frac{c}{\|N\|_{L^\infty(\Omega_T)} + \|M\|_{L^\infty(\Omega_T)}}, \frac{c}{4 \left( \|M\|_{L^\infty(\Omega_T)} + \|N\|_{L^\infty(\Omega_T)} \right)^2} \right\}.$$

Since we have now proven the nonnegativity of the excess terms, the result follows. □

The following lemma shows that we can derive the desired integral estimate from the pointwise lower bound. The proof uses rather standard techniques and is based on localization with a suitable cutoff function and then integration by parts. For the convenience of the reader, we give the details in the appendix.

**Lemma 4.5** *Let  $u^\varepsilon : \Omega_T \rightarrow \mathbb{R}$  be a smooth solution to (3.1), and  $S$  as in (4.4). If  $\varepsilon = 0$ , we assume in addition that  $Du^\varepsilon \neq 0$ . Suppose that we can find weights  $w_1, w_2, w_3, w_4 \in \mathbb{R}$  such that*

$$\lambda(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2u^\varepsilon|^2 \leq S \quad \text{a.e. in space in } \Omega_T, \tag{4.10}$$

for some constant  $\lambda = \lambda(n, p, \gamma, s, w_1, w_2, w_3, w_4) > 0$ . If  $s \neq \gamma - p$ , then for any concentric parabolic cylinders  $Q_r \subset Q_{2r} \Subset \Omega_T$  with center point  $(x_0, t_0) \in \Omega_T$ , we have the estimate

$$\begin{aligned} &\int_{Q_r} \left| D \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{4}} Du^\varepsilon \right) \right|^2 dxdt \\ &\leq \frac{C}{r^2} \left( \int_{Q_{2r}} (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} |Du^\varepsilon|^2 dxdt + \int_{Q_{2r}} (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p+s-\gamma}{2}} dxdt \right) \\ &\quad + \varepsilon \left( \frac{C}{r^2} \int_{Q_{2r}} |\ln(|Du^\varepsilon|^2 + \varepsilon)| dxdt + C \int_{B_{2r}} |\ln(|Du^\varepsilon(x, t_0)|^2 + \varepsilon)| dx \right) \end{aligned} \tag{4.11}$$

where  $C = C(n, p, \gamma, s, \lambda, w_1, w_2, w_3, w_4) > 0$ .

The last two integrals on the right hand side of (4.11) do not appear if  $s \neq \gamma - p + 2$ . The source of such error terms in the case  $s = \gamma - p + 2$  is the logarithm in Lemma 4.3 when  $\beta = -2$ .

### 5 Smooth case with non-zero gradient

Let  $1 < p < \infty$  and  $-1 < \gamma < \infty$ . In this section we assume that  $u : \Omega_T \rightarrow \mathbb{R}$  is a smooth solution to

$$u_t - |Du|^\gamma (\Delta u + (p - 2)\Delta_\infty^N u) = 0, \tag{5.1}$$

such that  $Du \neq 0$ . That is,  $u$  does not have critical points in space. Our main result in this case is the following a priori estimate. Usually extending a regularity result to a general nonsmooth case is quite straightforward.

**Proposition 5.1** *Let  $n \geq 2, 1 < p < \infty$  and  $-1 < \gamma < \infty$ . Let  $u : \Omega_T \rightarrow \mathbb{R}$  be a smooth solution to (5.1) such that  $Du \neq 0$ . If*

$$s > \max \left\{ -1 - \frac{p-1}{n-1}, \gamma + 1 - p \right\}, \tag{5.2}$$

*then for any concentric parabolic cylinders  $Q_r \subset Q_{2r} \Subset \Omega_T$ , we have the estimate*

$$\int_{Q_r} \left| D \left( |Du|^{\frac{p-2+s}{2}} Du \right) \right|^2 dxdt \leq \frac{C}{r^2} \left( \int_{Q_{2r}} |Du|^{p+s} dxdt + \int_{Q_{2r}} |Du|^{p+s-\gamma} dxdt \right),$$

*where  $C = C(n, p, \gamma, s) > 0$ .*

The following Lemma, Lemma 5.2, is the main ingredient in the proof of Proposition 5.1. Thus we postpone the proof of Proposition 5.1 until after the proof of Lemma 5.2.

In the following lemma we consider the weighted sum

$$S = w_1 G D_1(p - 2 + s) + w_2 G D_2(p - 2 + s)$$

where  $w_1, w_2 \in \mathbb{R}$ , and the notation was defined in (4.4). Note that since  $\varepsilon = 0$  in this section, the terms with weights  $w_3$  and  $w_4$  in (4.4) disappear. The purpose of Lemma 5.2 is to show that under restriction (5.2), we can find positive weights  $w_1 = w_1(n, p, \gamma, s) > 0$  and  $w_2 = w_2(n, p, \gamma, s) > 0$  such that  $S$  has a suitably nonnegative lower bound to make Lemma 4.4 applicable. Moreover, by the proof of Lemma 5.2 and Sylvester’s condition, we can choose the value  $c = c(n, p, \gamma, s) > 0$  small enough such that for  $M$  in the proof it holds

$$M_{11} \geq c \text{ and } \det(M) \geq c.$$

The proof of Proposition 5.1 is then finished by using Lemma 4.5.

**Lemma 5.2** *Let  $n \geq 2, 1 < p < \infty$  and  $-1 < \gamma < \infty$ . Let  $u : \Omega_T \rightarrow \mathbb{R}$  be a smooth solution to (5.1) such that  $Du \neq 0$ . If (5.2) holds, then we can select  $w_1 = w_1(n, p, \gamma, s) > 0$  and  $w_2 = w_2(n, p, \gamma, s) > 0$ , such that*

$$|Du|^{p-2+s} \left\{ c |D_T |Du||^2 + Q \right\} \leq S$$

*where  $c = c(n, p, \gamma, s) > 0$  and*

$$Q = \langle \bar{x}, M \bar{x} \rangle$$

*with  $\bar{x} = (\Delta_T u, \Delta_\infty^N u)^T \in \mathbb{R}^2$  and a uniformly bounded positive definite (with a uniform constant) symmetric matrix  $M = M(n, p, \gamma, s) \in \mathbb{R}^{2 \times 2}$ .*

**Proof** Similarly as in (4.7), recalling that  $\varepsilon = 0$ , by expressions (4.5), we arrive at

$$\begin{aligned} |Du|^{p-2+s} \left\{ w_1 |D^2 u|^2 + w_1 (p - 2 + s) |D_T |Du||^2 + (w_2 - w_1) (\Delta_T u)^2 \right. \\ \left. + (w_2 (p - 1) (p - 1 + s - \gamma) - w_1) (\Delta_\infty^N u)^2 \right. \\ \left. + (w_2 (2p - 2 + s - \gamma) - w_1 (p + s)) \Delta_T u \Delta_\infty^N u \right\} = S. \end{aligned} \tag{5.3}$$

We estimate  $|D^2 u|^2$  on the left hand side of (5.3) from below by the fundamental inequality, Lemma 4.1. This yields the following lower bound for  $S$

$$|Du|^{p-2+s} \left\{ w_1 (p + s) |D_T |Du||^2 + Q \right\} \leq S,$$

where

$$Q := \left( w_2 - \frac{n-2}{n-1} w_1 \right) (\Delta_T u)^2 + w_2(p-1)(p-1+s-\gamma) (\Delta_\infty^N u)^2 + (w_2(2p-2+s-\gamma) - w_1(p+s)) \Delta_T u \Delta_\infty^N u.$$

We write  $Q$  more compactly as

$$Q = \langle \bar{x}, M\bar{x} \rangle,$$

where  $\bar{x} = (\Delta_T u, \Delta_\infty^N u)^T \in \mathbb{R}^2$  is a vector and

$$M := \begin{bmatrix} w_2 - \frac{n-2}{n-1} w_1 & \frac{1}{2} (w_2(2p-2+s-\gamma) - w_1(p+s)) \\ \frac{1}{2} (w_2(2p-2+s-\gamma) - w_1(p+s)) & w_2(p-1)(p-1+s-\gamma) \end{bmatrix}$$

is a symmetric  $2 \times 2$ -matrix. We claim that under assumption (5.2) we can choose  $w_1, w_2 \in \mathbb{R}$  such that  $M$  is uniformly bounded positive definite (with a uniform constant).

If  $n = 2$ , this is easy to see by selecting

$$w_1 = 2p - 2 + s - \gamma \quad \text{and} \quad w_2 = p + s,$$

because then

$$M = \begin{bmatrix} p + s & 0 \\ 0 & (p + s)(p - 1)(p - 1 + s - \gamma) \end{bmatrix}$$

and hence

$$Q = (p + s) \left( (\Delta_T u)^2 + (p - 1)(p - 1 + s - \gamma) (\Delta_\infty^N u)^2 \right).$$

In other words, with such choice of  $w_1$  and  $w_2$ , the mixed term  $\Delta_T u \Delta_\infty^N u$  vanishes. Notice that (5.2) implies that  $w_1 > 0$  and  $w_2 > 0$ .

For the higher dimensional case  $n \geq 3$ , we set  $w_1 = 1$  and find  $w_2 = w_2(n, p, \gamma, s) > 0$  such that  $M$  is uniformly bounded positive definite (with a uniform constant). This is possible precisely when (5.2) holds: Since the proof is quite tedious, we postpone it to Lemma B.1 in the appendix. □

We are ready to give the proof of Proposition 5.1.

**Proof of Proposition 5.1** Let us fix  $w_1 = w_1(n, p, \gamma, s) > 0$  and  $w_2 = w_2(n, p, \gamma, s) > 0$  according to Lemma 5.2. Lemma 4.4 is then applicable because  $w_1 > 0, w_3 = 0$  implies that  $c_1 = w_1 + w_3\kappa > 0$  and the conclusion of Lemma 5.2 implies that (4.9) holds. Therefore, by Lemma 4.4 there exists  $\lambda = \lambda(n, p, \gamma, s, w_1, w_2, w_3, w_4) > 0$  such that

$$\lambda |Du|^{p-2+s} |D^2u|^2 \leq S$$

in  $\Omega_T$ . Now the desired estimate follows from Lemma 4.5. □

Range (5.2) in Proposition 5.1, is optimal in the following sense: In the elliptic case, [10] and [25], the best known range is  $s > -1 - \frac{p-1}{n-1}$ . On the other hand, Example 5.1 below shows that in the parabolic case we cannot hope to reach any better range than  $s > \gamma + 1 - p$ . A counterexample of this type was used in [10, Sect. 1.3] for the standard  $p$ -parabolic equation.

**Example 5.1** (Counterexample) Let  $u : \mathbb{R}^n \times (0, \infty) \rightarrow \mathbb{R}$  be given by

$$u(x, t) := Ct + |x_1|^\alpha$$

for some  $C \in \mathbb{R}$  and  $\alpha > 0$ . Note that

$$|Du|^\gamma \Delta_p^N u = \alpha^{\gamma+1}(\alpha - 1)(p - 1)|x_1|^{(\alpha-1)(\gamma+1)-1}.$$

Hence, if

$$\alpha = 1 + \frac{1}{\gamma + 1} \quad \text{and} \quad C = \alpha^{\gamma+1}(\alpha - 1)(p - 1)$$

then  $u$  solves (2.1) in the classical sense whenever  $x_1 \neq 0$ . Indeed, by a direct computation, we have

$$u_{x_1} = \alpha|x_1|^{\alpha-2}x_1, \quad u_{x_i} = 0 \quad \text{for } i = 2, \dots, n,$$

and

$$u_{x_1x_1} = \alpha(\alpha - 1)|x_1|^{\alpha-2}, \quad u_{x_ix_j} = 0,$$

where  $i, j = 1, \dots, n$  and  $i$  and  $j$  are not both 1.

Next we verify that the function  $u$  is a viscosity solution in the whole  $\mathbb{R}^n$  according to Definition 2.3 also at those points where  $x_1 = 0$ . Whenever  $x_1 \neq 0$ ,  $x_0 = (x_1, \dots, x_n)$ , and the test function  $\varphi$  touches  $u$  at  $(x_0, t_0)$  from below (the argument is analogous from above), we may use the facts that

$$D\varphi(x_0, t_0) = Du(x_0, t_0) \neq 0, \quad \varphi_t(x_0, t_0) = u_t(x_0, t_0) = 0$$

and

$$D^2\varphi(x_0, t_0) \leq D^2u(x_0, t_0).$$

Let us consider the points where  $x_1 = 0$ . We study the degenerate case  $\gamma > 0$  and the singular case  $-1 < \gamma \leq 0$  separately. If  $x_1 = 0$  and  $\gamma > 0$ , then there are no test functions touching  $u$  from above and for a test function  $\varphi$  touching from below, we have  $D\varphi(x_0, t_0) = Du(x_0, t_0) = 0$  and

$$\varphi_t(x_0, t_0) = u_t(x_0, t_0) = C.$$

Since  $C > 0$ , the function  $u$  is a viscosity supersolution.

The given function is also a viscosity solution whenever  $-1 < \gamma \leq 0$ : the proof for the supersolution property is the same as in the degenerate case above. It is also a subsolution because (similarly to the degenerate case) there are no admissible test functions touching  $u$  from above. We provide a detailed proof of this fact. Thriving for a contradiction, suppose that there is an admissible test function  $\varphi$  touching  $u$  at  $(x_0, t_0)$  with  $x_0 = (0, \dots, 0)$  (for simplicity) from above. Then necessarily

$$\varphi_t(x_0, t_0) = C > 0.$$

By the definition of a viscosity solution it holds that

$$\phi(x, t) = u(x_0, t_0) + \varphi_t(x_0, t_0)(t - t_0) + f(|x|) + \sigma(t - t_0)$$

is an admissible test function touching strictly from above. By strict touching and regularity of  $u$ , by translating with respect to  $x_1$  and lifting we may assume that  $\phi$  touches  $u$  at a point  $(x, t_0)$ ,  $x = (\varepsilon, 0, \dots, 0)$ , with small  $\varepsilon > 0$ . Also observe that by an approximation, we

could assume that  $\sigma$  is a  $C^2$  function, but we omit this step as well. Also recall the notation  $g(y) = f(|y|)$  and that

$$\lim_{y \rightarrow 0, y \neq 0} F(Dg(y), D^2g(y)) = 0.$$

Then by this and the counter assumption it holds at a point  $(x, t_0)$  for  $x$  close enough  $x_0$  that

$$\phi_t(x, t_0) - F(D\phi(x, t_0), D^2\phi(x, t_0)) = \varphi_t(x_0, t_0) - F(Dg(x), D^2g(x)) > 0. \tag{5.4}$$

On the other hand, since  $u$  is now  $C^2$ -function with the explicit formula, we have

$$\begin{aligned} \phi_t(x, t_0) - F(D\phi(x, t_0), D^2\phi(x, t_0)) &= \varphi_t(x_0, t_0) - F(Dg(x), D^2g(x)) \\ &= u_t(x, t_0) - F(Dg(x), D^2g(x)) \\ &\leq u_t(x, t_0) - F(Du(x, t_0), D^2u(x, t_0)) = 0, \end{aligned}$$

which contradicts inequality (5.4).

In the above inequality we used the fact that since  $\phi$  touches  $u$  from above at  $(x, t_0)$  we have  $D^2g(x) \geq D^2u(x, t_0)$  and  $Dg(x) = Du(x, t_0) \neq 0$  and thus

$$F(Dg(x), D^2g(x)) \geq F(Du(x, t_0), D^2u(x, t_0)).$$

We study the local  $W^{1,2}$ -regularity of  $|Du|^{\frac{p-2+s}{2}} Du$  for  $s \in \mathbb{R}$  and see what kind of restrictions for  $s$  arise. We have

$$\begin{aligned} \left| D(|Du|^{\frac{p-2+s}{2}} Du) \right| &= \frac{1}{2} \alpha^{\frac{p+s}{2}} (\alpha - 1)(p + s) |x_1|^{\frac{(\alpha-1)(p+s)}{2} - 1} \\ &= C(p, s, \gamma) |x_1|^{\frac{p+s}{2(\gamma+1)} - 1}. \end{aligned}$$

The function  $D(|Du|^{\frac{p-2+s}{2}} Du)$  locally belongs to  $L^2(\mathbb{R}^n \times (0, \infty))$  if and only if

$$2\left(\frac{p + s}{2(\gamma + 1)} - 1\right) > -1,$$

that is,

$$s > \gamma + 1 - p.$$

Observe that range condition (5.2) gives this in the plane, but in higher dimensions we have an additional restriction, which is the same restriction as in the elliptic case.

When  $s = 2 - p$ , then for  $W^{2,2}$ -regularity, the range

$$-1 < \gamma < 1$$

is sharp in the plane.

**Remark 5.1** Also the case  $n = 1$  holds. Recall that the key point is identity (3.4), that is,

$$\begin{aligned} &(|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ |D^2u^\varepsilon|^2 + (p - 2 + s) \frac{|D^2u^\varepsilon Du^\varepsilon|^2}{|Du^\varepsilon|^2 + \varepsilon} + s(p - 2) \frac{(\Delta_\infty u^\varepsilon)^2}{(|Du^\varepsilon|^2 + \varepsilon)^2} \right. \\ &\quad \left. + (p - 2 - \gamma)(|Du^\varepsilon|^2 + \varepsilon)^{-\gamma/2} u_t^\varepsilon \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \right\} \\ &= \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} AD^2u^\varepsilon Du^\varepsilon \right) - \frac{((|Du^\varepsilon|^2 + \varepsilon)^{\frac{p+s-\gamma}{2}})_t}{p + s - \gamma}, \end{aligned}$$



provided that  $s \neq \gamma - p$ . If  $n = 1$  this reduces to

$$\begin{aligned}
 & (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ 1 + (p-2+s)\theta + s(p-2)\theta^2 \right. \\
 & \quad \left. + (p-2-\gamma)((p-2)\theta + 1)\theta \right\} |D^2u^\varepsilon|^2 \\
 & = \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} (D^2u^\varepsilon Du^\varepsilon - \Delta u^\varepsilon Du^\varepsilon) \right) \\
 & \quad + u_t^\varepsilon \operatorname{div} \left( (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s-\gamma}{2}} Du^\varepsilon \right).
 \end{aligned} \tag{5.5}$$

The left hand side of (5.5) is

$$\begin{aligned}
 & (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} \left\{ (p-1)(p-1+s-\gamma)\theta^2 + (2p-2+s-\gamma)\theta\kappa + \kappa^2 \right\} |D^2u^\varepsilon|^2 \\
 & \geq \lambda (|Du^\varepsilon|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2u^\varepsilon|^2,
 \end{aligned}$$

for some constant  $\lambda = \lambda(p, \gamma, s) > 0$ , provided that  $s > \gamma + 1 - p$ . From this it is easy to derive the desired integral estimate. We conclude that Proposition 5.1 holds in case  $n = 1$  without the additional smoothness assumptions for  $u$ , and with the interpretation

$$s > \max \left\{ -1 - \frac{p-1}{n-1}, \gamma + 1 - p \right\} = \max \{-\infty, \gamma + 1 - p\} = \gamma + 1 - p.$$

## 6 Removing the smoothness assumption

Section 5 gives a formal derivation of the regularity estimate under the assumption that the gradient of a solution does not vanish. In this section, we remove the additional assumption in a certain range of parameters by regularizing the equation and then finally pass to a limit to obtain the result for the original equation.

### 6.1 Regularization

Let  $u^\varepsilon : \Omega_T \rightarrow \mathbb{R}$  be a smooth solution to the equation

$$u_t^\varepsilon - (|Du^\varepsilon|^2 + \varepsilon)^{\gamma/2} \left( \Delta u^\varepsilon + (p-2) \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \right) = 0 \tag{6.1}$$

where  $1 < p < \infty$ ,  $-1 < \gamma < \infty$ , and  $\varepsilon > 0$  is a regularization parameter. As explained in Sect. 3.3, the mismatch between the second order differential quantities in the fundamental inequality and the regularized equation and consequently in the basic estimate causes that some of the formal calculations do not work as such even if most of the steps work for general  $s$ . In particular positive definiteness of the quadratic form becomes an issue.

In this section, partly for the convenience of the reader, we have decided to limit ourselves to the planar case  $n = 2$  and focus on the square-integrability of the second order derivatives  $D^2u$ , that is, we consider the case  $s = 2 - p$ . In this case the range condition in (1.4) that is

$$s > \max \left\{ -1 - \frac{p-1}{n-1}, \gamma + 1 - p \right\}$$

reduces to

$$1 < p < \infty \quad \text{and} \quad -1 < \gamma < 1.$$

Then range (i)

$$1 < p \leq 5 \quad \text{and} \quad -1 < \gamma < 1, \tag{6.2}$$

in Theorem 1.1 and the Proposition below will be obtained by a straightforward generalization of the formal calculation ( $\varepsilon = 0$ ). That is, we consider the sum

$$S = w_1 \operatorname{div} (D^2 u^\varepsilon Du^\varepsilon - \Delta u^\varepsilon Du^\varepsilon) + w_2 u_t^\varepsilon \operatorname{div} ((|Du^\varepsilon|^2 + \varepsilon)^{-\frac{\gamma}{2}} Du^\varepsilon), \tag{6.3}$$

and show that if (6.2) holds, then we can find  $w_1, w_2 > 0$  such that

$$c |D_T |Du^\varepsilon|^2 + Q \leq S,$$

where  $c > 0$  and  $Q$  is positive definite. For range (ii) in the Proposition below which is the same as in Theorem 1.1, we instead consider the full  $S$  as defined in (3.11) or equivalently in (4.4).

Our main result for  $u^\varepsilon$  is the following.

**Proposition 6.1** *Let  $n = 2$ . Let  $u^\varepsilon : \Omega_T \rightarrow \mathbb{R}$  be a smooth solution to (6.1). If  $p$  and  $\gamma$  satisfy one of the following conditions:*

- (i)  $1 < p \leq 5$  and  $-1 < \gamma < 1$ ; or
- (ii)  $1 < p < \infty$  and  $-1 < \gamma < \sqrt{2} - \frac{1}{2}$ ,

*then for any concentric parabolic cylinders  $Q_r \subset Q_{2r} \Subset \Omega_T$  with center point  $(x_0, t_0) \in \Omega_T$ , we have the estimate*

$$\begin{aligned} \int_{Q_r} |D^2 u^\varepsilon|^2 dx dt &\leq \frac{C}{r^2} \left( \int_{Q_{2r}} |Du^\varepsilon|^2 dx dt + \int_{Q_{2r}} (|Du^\varepsilon|^2 + \varepsilon)^{\frac{2-\gamma}{2}} dx dt \right) \\ &\quad + \varepsilon \left( \frac{C}{r^2} \int_{Q_{2r}} |\ln(|Du^\varepsilon|^2 + \varepsilon)| dx dt + C \int_{B_{2r}} |\ln(|Du^\varepsilon(x, t_0)|^2 + \varepsilon)| dx \right) \end{aligned}$$

where  $C = C(p, \gamma) > 0$ .

The proof of Proposition 6.1 is postponed to the end of the section. The main ingredients of the proof of Proposition 6.1 are the following lemmas, Lemmas 6.2 and 6.3. The first lemma, Lemma 6.2, yields case (i). The second lemma, Lemma 6.3 yields case (ii). In both lemmas we consider the same weighted sum as before now selecting  $s = 2 - p$  i.e.

$$\begin{aligned} S &= w_1 G D_1(p - 2 + s) + w_2 G D_2(p - 2 + s) + \varepsilon w_3 G D_1(p - 4 + s) \\ &\quad + \varepsilon w_4 G D_2(p - 4 + s) \\ &= w_1 G D_1(0) + w_2 G D_2(0) + w_3 \varepsilon G D_1(-2) + w_4 \varepsilon G D_2(-2), \end{aligned} \tag{6.4}$$

where  $w_1, w_2, w_3, w_4 \in \mathbb{R}$  are some weights, and the notation was defined in (4.4).

The purpose of Lemmas 6.2 and 6.3 is to show that under restrictions (i) and (ii), respectively, we can find suitable weights  $w_1, w_2, w_3$  and  $w_4$ , that only depend on  $p$  and  $\gamma$ , such that  $S$  has a suitable lower bound.

**Lemma 6.2** *Let  $n = 2$ ,  $S$  be as in (6.4), and (i) in Proposition 6.1 hold. For  $\eta = \eta(p, \gamma) > 0$  small enough, if*

$$\begin{cases} w_1 = p - \gamma - 2\sqrt{(p-1)(1-\gamma)} + \eta, & w_2 = 2, \\ w_3 = 0, & w_4 = 0, \end{cases}$$

then

$$S \geq c|D_T|Du^\varepsilon|^2 + Q$$

where  $c = c(p, \gamma) > 0$  and

$$Q = \langle \bar{x}, M\bar{x} \rangle$$

with  $\bar{x} = (\Delta_T u^\varepsilon, \Delta_\infty^N u^\varepsilon)^T \in \mathbb{R}^2$  and a uniformly bounded positive definite (with a uniform constant) symmetric matrix  $M = M(p, \gamma) \in \mathbb{R}^{2 \times 2}$ .

**Lemma 6.3** *Let  $n = 2$ ,  $S$  be as in (6.4), and (ii) in Proposition 6.1 hold. If*

$$\begin{cases} w_1 = p - \gamma, & w_2 = 2 \\ w_3 = 4 - p + \gamma, & w_4 = 2, \end{cases}$$

then a statement similar to that in Lemma 6.2 holds.

To begin with, recall from (4.7) that  $S$  can be written as

$$S = c_1|D^2u^\varepsilon|^2 + c_2|D_T|Du^\varepsilon|^2 + (c_3 - c_1)(\Delta_T u^\varepsilon)^2 + ((c_3 + c_4)P_\theta - c_1)(\Delta_\infty^N u^\varepsilon)^2 + (c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2))\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon.$$

where

$$\begin{cases} c_1 = w_1 + w_3\kappa, & c_2 = -2w_3\theta\kappa, \\ c_3 = w_2 + w_4\kappa, & c_4 = -(w_2\gamma + w_4\kappa(2 + \gamma))\theta, \end{cases} \tag{6.5}$$

and

$$P_\theta = (p - 2)\theta + 1 \in (0, \infty), \quad \theta = \frac{|Du|^2}{|Du|^2 + \varepsilon} \in [0, 1),$$

$$\kappa = 1 - \theta = \frac{\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \in (0, 1].$$

Fundamental equality (4.3) in the plane yields that

$$S = c_1(2|D_T|Du^\varepsilon|^2 + (\Delta_T u^\varepsilon)^2 + (\Delta_\infty^N u^\varepsilon)^2) + c_2|D_T|Du^\varepsilon|^2 + (c_3 - c_1)(\Delta_T u^\varepsilon)^2 + ((c_3 + c_4)P_\theta - c_1)(\Delta_\infty^N u^\varepsilon)^2 + (c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2))\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon = (2c_1 + c_2)|D_T|Du^\varepsilon|^2 + Q. \tag{6.6}$$

where

$$Q = c_3(\Delta_T u^\varepsilon)^2 + (c_3 + c_4)P_\theta(\Delta_\infty^N u^\varepsilon)^2 + (c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2))\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon, \tag{6.7}$$

is a quadratic form in  $\Delta_T u^\varepsilon$  and  $\Delta_\infty^N u^\varepsilon$ . We write  $Q$  compactly as

$$Q = \langle \bar{x}, M\bar{x} \rangle,$$

where  $\bar{x} = (\Delta_T u^\varepsilon, \Delta_\infty^N u^\varepsilon)^T \in \mathbb{R}^2$  is a vector and

$$M := \begin{bmatrix} c_3 & \frac{1}{2}(c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2)) \\ \frac{1}{2}(c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2)) & (c_3 + c_4)P_\theta \end{bmatrix}$$

is a symmetric  $2 \times 2$ -matrix.

To prove Lemmas 6.2 and 6.3, it now suffices to check that (6.6) satisfies all the requirements of the lemmas: The coefficient of  $|D_T|Du^\varepsilon|^2$  in (6.6) needs to be bounded from below by a positive constant, that is,

$$2c_1 + c_2 = 2(w_1 + w_3\kappa) - 2w_3\theta\kappa \geq c \tag{6.8}$$

uniformly in  $\Omega_T$ . For the quadratic form  $Q$ , we need to analyse the uniform boundedness and uniform positive definiteness of matrix  $M$ . Uniform boundedness is quite straightforward, so we focus our attention on the uniform positive definiteness. By Sylvester’s condition, it suffices to check that

$$c_3 = w_2 + w_4\kappa \geq c, \tag{6.9}$$

and

$$\det(M) = c_3(c_3 + c_4)P_\theta - \frac{(c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2))^2}{4} \geq c \tag{6.10}$$

uniformly in  $\Omega_T$ . Next we prove Lemma 6.2, which implies nonnegativity of the necessary terms when  $1 < p \leq 5$  and  $-1 < \gamma < 1$ . In this case a simple choice of the weights  $w_3 = w_4 = 0$  will work.

**Proof** (Proof of Lemma 6.2) Similarly to the smooth case, we start with  $w_3 = w_4 = 0$ , plug these values into (6.5), and obtain

$$\begin{cases} c_1 = w_1, & c_2 = 0, \\ c_3 = w_2, & c_4 = -w_2\gamma\theta. \end{cases}$$

This together with (6.6) gives

$$S = 2w_1|D_T|Du^\varepsilon|^2 + w_2(\Delta_T u^\varepsilon)^2 + w_2P_\theta R_\theta(\Delta_\infty^N u^\varepsilon)^2 + (w_2(P_\theta + R_\theta) - 2w_1)\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon,$$

where we denote  $R_\theta := 1 - \gamma\theta \in (0, 2)$ , for the sake of brevity. To simplify the above identity, we select  $w_2 = 2$ . Thus

$$\begin{aligned} S &= 2w_1|D_T|Du^\varepsilon|^2 + 2\left((\Delta_T u^\varepsilon)^2 + P_\theta R_\theta(\Delta_\infty^N u^\varepsilon)^2 + (P_\theta + R_\theta - w_1)\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon\right) \\ &= 2w_1|D_T|Du^\varepsilon|^2 + Q, \end{aligned}$$

where the matrix of the quadratic form  $Q$  is

$$M(\theta) := \begin{bmatrix} 2 & P_\theta + R_\theta - w_1 \\ P_\theta + R_\theta - w_1 & 2P_\theta R_\theta \end{bmatrix}.$$

The determinant of  $M(\theta)$  is uniformly positive if and only if

$$P_\theta R_\theta - \frac{(P_\theta + R_\theta - w_1)^2}{4} \geq c > 0,$$

that is,

$$X_2(\theta) := (\sqrt{P_\theta} + \sqrt{R_\theta})^2 > w_1 > (\sqrt{P_\theta} - \sqrt{R_\theta})^2 := X_1(\theta)$$

uniformly in  $\Omega_T$ . Thus it suffices to verify

$$\inf_\theta X_2(\theta) > \sup_\theta X_1(\theta).$$

Computing the derivative of  $X_1$  with respect to  $\theta$ , one has

$$\begin{aligned} X_1'(\theta) &= p - 2 - \gamma - \frac{(p - 2)R_\theta - \gamma P_\theta}{\sqrt{P_\theta R_\theta}} \\ &= \frac{(\sqrt{P_\theta} - \sqrt{R_\theta})((p - 2)\sqrt{R_\theta} + \gamma\sqrt{P_\theta})}{\sqrt{P_\theta R_\theta}}. \end{aligned}$$

Then in the eligible range of parameters  $X_1'(\theta) = 0$  if and only if  $p - 2 + \gamma = 0$ . Hence, by considering the values at the endpoints, we obtain the supremum of  $X_1$  with respect to  $\theta$ :

$$\begin{aligned} \sup_{\theta} X_1(\theta) &= \max\{X_1(0), X_1(1)\} \\ &= \max\left\{0, (\sqrt{p - 1} - \sqrt{1 - \gamma})^2\right\} \\ &= (\sqrt{p - 1} - \sqrt{1 - \gamma})^2. \end{aligned}$$

Similarly, we obtain the derivative of  $X_2$  with respect of  $\theta$ :

$$X_2'(\theta) = \frac{(\sqrt{P_\theta} + \sqrt{R_\theta})((p - 2)\sqrt{R_\theta} - \gamma\sqrt{P_\theta})}{\sqrt{P_\theta R_\theta}},$$

and thus the eligible stationary point is

$$\theta_2 = \frac{p - 2 - \gamma}{(p - 2)\gamma}$$

if  $(p - 2)\gamma > 0$ . Then

$$\begin{aligned} \inf_{\theta} X_2(\theta) &= \min\left\{X_2(0), X_2\left(\frac{p - 2 - \gamma}{(p - 2)\gamma}\right), X_2(1)\right\} \\ &= \min\left\{4, \frac{p - 2}{\gamma} + \frac{\gamma}{p - 2} + 2, (\sqrt{p - 1} + \sqrt{1 - \gamma})^2\right\} \\ &= \min\left\{4, (\sqrt{p - 1} + \sqrt{1 - \gamma})^2\right\}. \end{aligned}$$

Obviously, we have

$$(\sqrt{p - 1} + \sqrt{1 - \gamma})^2 > (\sqrt{p - 1} - \sqrt{1 - \gamma})^2.$$

Note that

$$4 > (\sqrt{p - 1} - \sqrt{1 - \gamma})^2$$

is equivalent to

$$1 < p \leq 5 \quad \text{and} \quad -1 < \gamma < 1,$$

or  $5 < p < 7 + 4\sqrt{2}$  and  $-1 < \gamma < -2 - p + 4\sqrt{p - 1}$ . Thus if  $p$  and  $\gamma$  satisfy range (i), for small enough  $\eta = \eta(p, \gamma) > 0$ , in addition to the above choice  $w_2 = 2$ , we set

$$w_1 = (\sqrt{p - 1} - \sqrt{1 - \gamma})^2 + \eta.$$

The proof is finished. □

Next we prove Lemma 6.3, which implies nonnegativity of the necessary terms when  $1 < p < \infty$  and  $-1 < \gamma < \sqrt{2} - \frac{1}{2}$ . In this case, we use a choice of the weights which leads to the vanishing coefficient of the mixed term  $\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon$  in (6.7).

To be more precise, at the beginning of this section, we obtained three conditions (6.8), (6.9) and (6.10), i.e. that

$$\begin{cases} 2c_1 + c_2 = 2(w_1 + w_3\kappa) - 2w_3\theta\kappa \geq c, \\ c_3 = w_2 + w_4\kappa \geq c, \\ \det(M) = c_3(c_3 + c_4)P_\theta - \frac{(c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2))^2}{4} \geq c \end{cases} \tag{6.11}$$

need to hold uniformly in  $\Omega_T$ . Here

$$M = \begin{bmatrix} & c_3 & \frac{1}{2}(c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2)) \\ \frac{1}{2}(c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2)) & & (c_3 + c_4)P_\theta \end{bmatrix}$$

is the coefficient matrix of quadratic form

$$Q = c_3(\Delta_T u^\varepsilon)^2 + (c_3 + c_4)P_\theta(\Delta_\infty^N u^\varepsilon)^2 + (c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2))\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon.$$

To simplify the computations in checking the last condition in (6.11), we will consider a special case where the coefficient  $c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2)$  of the mixed term  $\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon$  vanishes.

**Lemma 6.4** (Vanishing mixed term) *The mixed term  $\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon$  in  $Q$  vanishes, i.e.*

$$c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2) = 0$$

uniformly in  $\Omega_T$  if and only if

$$\begin{cases} 2w_1 = (p - \gamma)w_2, \\ 2w_3 = (4 - p + \gamma)w_4, \\ (p - 2 - \gamma)(w_4 - w_2) = 0. \end{cases}$$

**Proof** Recall that  $\theta = 1 - \kappa, \kappa > 0$ , and  $P_\theta = (p - 2)\theta + 1$ . Then recalling the expressions of  $c_1, \dots, c_4$  in (6.5), we can write the coefficient of the mixed term  $\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon$  as a polynomial of  $\kappa$  as

$$\begin{aligned} & c_3P_\theta + (c_3 + c_4) - (2c_1 + c_2) \\ &= ((4 - p + \gamma)w_4 - 2w_3)\kappa^2 + (p - 2 - \gamma)(w_4 - w_2)\kappa + (p - \gamma)w_2 - 2w_1. \end{aligned}$$

Set all the coefficients to be zero, we have the desired condition. □

By the above Lemma, we can easily to obtain the following result.

**Corollary 6.5** *If*

$$\begin{cases} w_1 = p - \gamma, & w_2 = 2, \\ w_3 = 4 - p + \gamma, & w_4 = 2, \end{cases} \tag{6.12}$$

then the mixed term  $\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon$  in  $Q$  vanishes.

The above corollary gives a choice of the coefficients  $w_1, w_2, w_3$  and  $w_4$  to obtain the vanishing coefficient of mixed term  $\Delta_T u^\varepsilon \Delta_\infty^N u^\varepsilon$ . This then helps us in proving Lemma 6.3.

**Proof of Lemma 6.3** If  $w_1, w_2, w_3$  and  $w_4$  satisfy (6.12), then by Corollary 6.5, the last condition in (6.11) reduces to checking that

$$\det(M) = c_3(c_3 + c_4)P_\theta \geq c.$$

Since

$$P_\theta = (p - 2)\theta + 1 \geq \min\{p - 1, 1\} > 0,$$

sufficient conditions to obtain (6.11) can be written as

$$\begin{cases} 2c_1 + c_2 = 2(w_1 + w_3\kappa) - 2w_3\theta\kappa \geq c, \\ c_3 = w_2 + w_4\kappa \geq c, \\ c_3 + c_4 = w_2 + w_4\kappa - (w_2\gamma + w_4\kappa(2 + \gamma))\theta \geq c \end{cases}$$

uniformly in  $\Omega_T$ .

First, using values (6.12) in the first condition and replacing  $\theta$  by  $1 - \kappa$ , we have

$$2c_1 + c_2 = 2(p - \gamma) + 2(4 - p + \gamma)\kappa^2.$$

Since  $\kappa$  is positive, the sign of the derivative with respect to  $\kappa$  that is  $4(4 - p + \gamma)\kappa$  is fixed. Then  $2c_1 + c_2$  with respect to  $\kappa$  is monotone and the minimum point corresponds either  $\kappa = 0$  or  $\kappa = 1$ . Thus

$$2c_1 + c_2 \geq \min\{2(p - \gamma), 8\} > 0.$$

For the second condition, when  $w_2 = w_4 = 2$ , it is obvious that

$$c_3 = 2 + 2\kappa \geq 2 > 0.$$

Finally, for the last condition plugging values (6.12) in and rewriting as

$$c_3 + c_4 = 2(1 - \gamma) - 2\kappa + 2(2 + \gamma)\kappa^2.$$

When the derivative of  $c_3 + c_4$  with respect to  $\kappa$  vanishes, that is,  $-2 + 4(2 + \gamma)\kappa = 0$ , one has

$$\kappa_1 = \frac{1}{2(2 + \gamma)} \in (0, 1].$$

Then the minimum point is one of the boundary points or the extreme point  $\kappa_1$ . Selecting  $\kappa = \kappa_1$ , we have

$$c_3 + c_4 = 2(1 - \gamma) - \frac{1}{2(2 + \gamma)} > 0 \tag{6.13}$$

if and only if

$$-1 < \gamma < \sqrt{2} - \frac{1}{2}.$$

If  $\kappa = 0$ , we have  $c_3 + c_4 = 2(1 - \gamma)$ , and if  $\kappa = 1$ , then  $c_3 + c_4 = 4$ . It follows that the minimum is given by strictly positive expression (6.13), and the proof is finished.  $\square$

The proof of Proposition 6.1 now immediately follows.

**Proof of Proposition 6.1** The result immediately follows from the previous lemmas, since under assumption (i), Lemma 6.2 implies that (4.9) holds and thus Lemma 4.4 is applicable. Similarly under assumption (ii), Lemma 6.3 implies that Lemma 4.4 is applicable. Now the desired estimate follows from Lemma 4.5.  $\square$

### 6.2 Passing to the original equation

In this section we justify the limiting argument to let  $\varepsilon \rightarrow 0$  in Proposition 6.1 and thus derive our main result, Theorem 1.1.

**Proof of Theorem 1.1** Let  $u : \Omega_T \rightarrow \mathbb{R}$  be a viscosity solution to

$$u_t - |Du|^\gamma (\Delta u + (p - 2)\Delta_\infty^N u) = 0.$$

Let us fix concentric parabolic cylinders  $Q_r \subset Q_{2r} \Subset \Omega_T$  with center point  $(x_0, t_0) \in \Omega_T$  and moreover, let us fix a smooth subdomain  $U \Subset \Omega$  and  $0 < t_1 < t_2 < T$  such that  $Q_{2r} \Subset U_{t_1, t_2} \Subset \Omega_T$ . For  $\varepsilon > 0$  small, let us consider the Dirichlet problem

$$\begin{cases} u_t^\varepsilon - (|Du^\varepsilon|^2 + \varepsilon)^{\gamma/2} \left( \Delta u^\varepsilon + (p - 2) \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \right) = 0 & \text{in } U_{t_1, t_2}; \\ u^\varepsilon = u & \text{on } \partial_p U_{t_1, t_2}, \end{cases}$$

where

$$\partial_p U_{t_1, t_2} := (\bar{U} \times \{t_1\}) \cup (\partial U \times (t_1, t_2])$$

is the parabolic boundary of  $U_{t_1, t_2}$ . By the classical theory of uniformly parabolic equations, the above problem has a unique solution  $u^\varepsilon \in C^\infty(U_{t_1, t_2}) \cap C(\bar{U}_{t_1, t_2})$ .

Proposition 6.1 is applicable to  $u^\varepsilon$  and we conclude that

$$\begin{aligned} \int_{Q_r} |D^2 u^\varepsilon|^2 dx dt &\leq \frac{C}{r^2} \left( \int_{Q_{2r}} |Du^\varepsilon|^2 dx dt + \int_{Q_{2r}} (|Du^\varepsilon|^2 + \varepsilon)^{\frac{2-\gamma}{2}} dx dt \right) \\ &+ \varepsilon \left( \frac{C}{r^2} \int_{Q_{2r}} |\ln(|Du^\varepsilon|^2 + \varepsilon)| dx dt \right. \\ &\left. + C \int_{B_{2r}} |\ln(|Du^\varepsilon(x, t_0)|^2 + \varepsilon)| dx \right) \end{aligned} \tag{6.14}$$

where  $C = C(p, \gamma) > 0$ . By [14], for any  $Q_R \Subset U_{t_1, t_2}$  there exist positive constants  $\alpha \in (0, 1)$  and  $C > 0$ , that are allowed to depend on  $p, \gamma, \text{dist}(Q_R, \partial U_{t_1, t_2})$  and  $\|u\|_{L^\infty(U_{t_1, t_2})}$ , such that

$$\|Du^\varepsilon\|_{C^\alpha(Q_R)} \leq C. \tag{6.15}$$

Arzelà-Ascoli theorem gives that  $u^\varepsilon$  and  $Du^\varepsilon$  both converge locally uniformly, up to a subsequence, and

$$u^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} \bar{u} \quad \text{and} \quad Du^\varepsilon \xrightarrow{\varepsilon \rightarrow 0} D\bar{u}$$

for some continuous function  $\bar{u} : U_{t_1, t_2} \rightarrow \mathbb{R}$ , which by a barrier argument is continuous up to the parabolic boundary, and whose spatial gradient  $D\bar{u}$  is locally continuous.

By the well known [12] stability properties of viscosity solutions,  $\bar{u}$  is a viscosity solution to

$$\begin{cases} \bar{u}_t - |D\bar{u}|^\gamma (\Delta \bar{u} + (p - 2)\Delta_\infty^N \bar{u}) = 0 & \text{in } U_{t_1, t_2}; \\ \bar{u} = u & \text{on } \partial_p U_{t_1, t_2}. \end{cases}$$

By the uniqueness theorem for viscosity solutions [22], we conclude that  $\bar{u} = u$ .

By employing bound (6.15), we find that the right hand side of (6.14) is bounded from above by a constant independent of  $\varepsilon$ . Thus  $\{D^2 u^\varepsilon\}_\varepsilon$  is bounded in  $L^2(Q_r)$ , and consequently



we may extract a subsequence that converges weakly in  $L^2(Q_r)$ . Further, using integration by parts, we see that the limit is  $D^2u$ , and thus  $D^2u \in L^2_{\text{loc}}(\Omega_T)$ . Finally, we conclude that

$$\begin{aligned} \int_{Q_r} |D^2u|^2 dxdt &\leq \liminf_{\varepsilon \rightarrow 0} \int_{Q_r} |D^2u^\varepsilon|^2 dxdt \\ &\leq \liminf_{\varepsilon \rightarrow 0} \left( \frac{C}{r^2} \left( \int_{Q_{2r}} |Du^\varepsilon|^2 dxdt + \int_{Q_{2r}} (|Du^\varepsilon|^2 + \varepsilon)^{\frac{2-\gamma}{2}} dxdt \right) \right. \\ &\quad \left. + \varepsilon \left( \frac{C}{r^2} \int_{Q_{2r}} |\ln(|Du^\varepsilon|^2 + \varepsilon)| dxdt + C \int_{B_{2r}} |\ln(|Du^\varepsilon(x, t_0)|^2 + \varepsilon)| dx \right) \right) \\ &= \frac{C}{r^2} \left( \int_{Q_{2r}} |Du|^2 dxdt + \int_{Q_{2r}} |Du|^{2-\gamma} dxdt \right), \end{aligned}$$

which is the desired estimate. □

It is possible to improve the ranges in Theorem 1.1. However, the computations get more technical, even if they follow the same ideas as above, and thus we have chosen to omit them. In any case the question whether the full range obtained in the smooth case in Proposition 5.1 can also be obtained here remains an open problem.

Next we give the proof of Corollary 1.2.

**Proof of Corollary 1.2** Assume that  $u^\varepsilon$  is a smooth solution to (6.1), and observe

$$\begin{aligned} |u_t^\varepsilon| &= \left| (|Du^\varepsilon|^2 + \varepsilon)^{\gamma/2} \left( \Delta u^\varepsilon + (p-2) \frac{\Delta_\infty u^\varepsilon}{|Du^\varepsilon|^2 + \varepsilon} \right) \right| \\ &\leq (|Du^\varepsilon|^2 + \varepsilon)^{\gamma/2} (|\Delta u^\varepsilon| + |p-2| |D^2u^\varepsilon|) \\ &\leq (p+2) (|Du^\varepsilon|^2 + \varepsilon)^{\gamma/2} |D^2u^\varepsilon|. \end{aligned}$$

As above, the spatial gradient is Hölder continuous and since  $\gamma$  is nonnegative, we have in  $Q_{2r}$

$$(|Du^\varepsilon|^2 + \varepsilon)^{\gamma/2} \leq C.$$

For all  $Q_r \subset Q_{2r} \Subset \Omega_T$ , we have

$$\begin{aligned} \int_{Q_r} |u_t^\varepsilon|^2 dxdt &\leq (p+2)^2 \int_{Q_r} (|Du^\varepsilon|^2 + \varepsilon)^\gamma |D^2u^\varepsilon|^2 dxdt \\ &\leq (p+2)^2 \| (|Du^\varepsilon|^2 + \varepsilon)^\gamma \|_{L^\infty(Q_r)} \int_{Q_r} |D^2u^\varepsilon|^2 dxdt. \end{aligned}$$

Then we use (6.14) estimate the right hand side of the above estimate. Similarly to the proof of Theorem 1.1, up to a subsequence,  $\{u_t^\varepsilon\}_\varepsilon$  converges weakly in  $L^2(Q_r)$ . By integration by parts, the weak limit is  $u_t$ . In particular  $u_t$  exists as a function and  $u_t \in L^2_{\text{loc}}(\Omega_T)$ . □

**Acknowledgements** The first author was supported by China Scholarship Council, no. 202006020186. The third author was supported by the Academy of Finland, Center of Excellence in Randomness and Structures and the Academy of Finland, project 308759.

**Funding** Open Access funding provided by University of Jyväskylä (JYU).

**Data availability statement:** Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

**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/>.

### Appendix A: Proof of Lemma 4.5

Next we prove Lemma 4.5. For convenience of the reader, we recall its statement here: Let  $S$  be as in (4.4), and  $u : \Omega_T \rightarrow \mathbb{R}$  be a smooth solution to (3.1). (If  $\varepsilon = 0$ , we assume in addition that  $Du \neq 0$ .) Suppose that we can find weights  $w_1, w_2, w_3, w_4 \in \mathbb{R}$  such that

$$S \geq \lambda(|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2u|^2 \quad \text{a.e. in space in } \Omega_T, \tag{A.1}$$

for some constant  $\lambda = \lambda(n, p, \gamma, s, w_1, w_2, w_3, w_4) > 0$ . If  $s \neq \gamma - p$ , then for any concentric parabolic cylinders  $Q_r \subset Q_{2r} \Subset \Omega_T$  with center point  $(x_0, t_0) \in \Omega_T$  we have the estimate

$$\begin{aligned} & \int_{Q_r} \left| D \left( (|Du|^2 + \varepsilon)^{\frac{p-2+s}{4}} Du \right) \right|^2 dx dt \\ & \leq \frac{C}{r^2} \left( \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} |Du|^2 dx dt + \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p+s-\gamma}{2}} dx dt \right) \\ & \quad + \varepsilon \left( \frac{C}{r^2} \int_{Q_{2r}} |\ln(|Du|^2 + \varepsilon)| dx dt + C \int_{B_{2r}} |\ln(|Du(x, t_0)|^2 + \varepsilon)| dx \right) \end{aligned}$$

where  $C = C(n, p, \gamma, s, \lambda, w_1, w_2, w_3, w_4) > 0$ .

**Proof of Lemma 4.5** Let us assume that  $s \neq \gamma - p$  and  $s \neq \gamma - p + 2$ . As remarked after the lemma when  $s \neq \gamma - p + 2$ , the logarithmic term does not appear in  $S$  and in the estimate of the lemma. Assumption (A.1) can be written as

$$\begin{aligned} & \lambda(|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2u|^2 \\ & \geq w_1 \operatorname{div} \left( (|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} (D^2u Du - \Delta u Du) \right) \\ & \quad + w_2 \operatorname{div} \left( u_t (|Du|^2 + \varepsilon)^{\frac{p-2+s-\gamma}{2}} Du \right) - w_2 \left( \frac{(|Du|^2 + \varepsilon)^{\frac{p+s-\gamma}{2}}}{p + s - \gamma} \right)_t \tag{A.2} \\ & \quad + \varepsilon w_3 \operatorname{div} \left( (|Du|^2 + \varepsilon)^{\frac{p-4+s}{2}} (D^2u Du - \Delta u Du) \right) \\ & \quad + \varepsilon w_4 \operatorname{div} \left( u_t (|Du|^2 + \varepsilon)^{\frac{p-4+s-\gamma}{2}} Du \right) - \varepsilon w_4 \left( \frac{(|Du|^2 + \varepsilon)^{\frac{p-2+s-\gamma}{2}}}{p - 2 + s - \gamma} \right)_t. \end{aligned}$$

Let us fix any concentric parabolic cylinders  $Q_r \subset Q_{2r} \Subset \Omega_T$  and then select a nonnegative cutoff function  $\phi : \mathbb{R}^n \times [0, t_0] \rightarrow [0, 1]$  such that

$$\phi \equiv 1 \text{ in } Q_r, \quad \phi \equiv 0 \text{ outside } Q_{2r}, \quad |D\phi| \leq \frac{C}{r} \quad \text{and} \quad |\phi_t| \leq \frac{C}{r^2} \tag{A.3}$$

for some absolute constant  $C > 0$ . We multiply (A.2) with  $\phi^2$  and integrate over  $Q_{2r}$ , apply integration by parts to each integral on the right hand side to obtain

$$\begin{aligned} & \lambda \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2u|^2 \phi^2 dx dt \\ & \leq -2w_1 \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} \langle D^2u Du - \Delta u Du, D\phi \rangle \phi dx dt \\ & \quad - 2w_2 \int_{Q_{2r}} u_t (|Du|^2 + \varepsilon)^{\frac{p-2+s-\gamma}{2}} \langle Du, D\phi \rangle \phi dx dt \\ & \quad + \frac{2w_2}{p+s-\gamma} \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p+s-\gamma}{2}} \phi_t \phi dx dt \\ & \quad - 2\varepsilon w_3 \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p-4+s}{2}} \langle D^2u Du - \Delta u Du, D\phi \rangle \phi dx dt \\ & \quad - 2\varepsilon w_4 \int_{Q_{2r}} u_t (|Du|^2 + \varepsilon)^{\frac{p-4+s-\gamma}{2}} \langle Du, D\phi \rangle \phi dx dt \\ & \quad + \frac{2\varepsilon w_4}{p-2+s-\gamma} \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p-2+s-\gamma}{2}} \phi_t \phi dx dt. \end{aligned}$$

Above we dropped the nonpositive boundary terms that appear when we integrate by parts with respect to time. Next we take absolute values and estimate  $\varepsilon/(|Du|^2 + \varepsilon) \leq 1$  in the last three integrals of the above display. We arrive at

$$\begin{aligned} & \lambda \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2u|^2 \phi^2 dx dt \\ & \leq C \left( \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2u| |Du| |D\phi| \phi dx dt \right. \\ & \quad + \int_{Q_{2r}} |u_t| (|Du|^2 + \varepsilon)^{\frac{p-2+s-\gamma}{2}} |Du| |D\phi| \phi dx dt \\ & \quad \left. + \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p+s-\gamma}{2}} |\phi_t| \phi dx dt \right), \end{aligned}$$

where  $C = C(n, p, \gamma, s, w_1, w_2, w_3, w_4) > 0$ . By Young’s inequality

$$\begin{aligned} & (\lambda - 2\eta) \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2u|^2 \phi^2 dx dt \\ & \leq \frac{C}{\eta} \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} |Du|^2 |D\phi|^2 dx dt + C \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{p+s-\gamma}{2}} |\phi_t| \phi dx dt, \end{aligned}$$

for any  $\eta > 0$  and some  $C = C(n, p, \gamma, s, w_1, w_2, w_3, w_4) > 0$ . Above we also employed equation (5.1) and estimated

$$(|Du|^2 + \varepsilon)^{-\gamma/2} |u_t| = \left| \Delta u + (p-2) \frac{\Delta_\infty u}{|Du|^2 + \varepsilon} \right| \leq C |D^2u| \tag{A.4}$$

for some  $C = C(n, p) > 0$ . Finally, we select  $\eta > 0$  small enough and employ (A.3) together with the fact that

$$\left| D \left( (|Du|^2 + \varepsilon)^{\frac{p-2+s}{4}} Du \right) \right|^2 \leq C (|Du|^2 + \varepsilon)^{\frac{p-2+s}{2}} |D^2u|^2 \tag{A.5}$$

where  $C = C(p, s) > 0$ , to arrive to the desired estimate.

Now, let us assume that  $s = \gamma - p + 2$ . Assumption (A.1) is now

$$\begin{aligned} \lambda(|Du|^2 + \varepsilon)^{\gamma/2}|D^2u|^2 &\leq w_1 \operatorname{div} \left( (|Du|^2 + \varepsilon)^{\gamma/2}(D^2uDu - \Delta uDu) \right) \\ &\quad + w_2 \operatorname{div} (u_t Du) - \frac{1}{2}w_2(|Du|^2 + \varepsilon)_t \\ &\quad + \varepsilon w_3 \operatorname{div} \left( (|Du|^2 + \varepsilon)^{\frac{\gamma-2}{2}}(D^2uDu - \Delta uDu) \right) \\ &\quad + \varepsilon w_4 \operatorname{div} \left( u_t(|Du|^2 + \varepsilon)^{-1}Du \right) - \frac{\varepsilon}{2}w_4 \left( \ln(|Du|^2 + \varepsilon) \right)_t. \end{aligned} \tag{A.6}$$

Let us fix any concentric parabolic cylinders  $Q_r \subset Q_{2r} \Subset \Omega_T$  and then select a nonnegative cutoff function  $\phi : \mathbb{R}^n \times [0, t_0] \rightarrow [0, 1]$  such that (A.3) holds. We multiply (A.6) with  $\phi^2$ , integrate over  $Q_{2r}$ , apply integration by parts to each integral on the right hand side to obtain

$$\begin{aligned} &\lambda \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\gamma/2}|D^2u|^2 \phi^2 dx dt \\ &\leq -2w_1 \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\gamma/2} \langle D^2uDu - \Delta uDu, D\phi \rangle \phi dx dt \\ &\quad - 2w_2 \int_{Q_{2r}} u_t \langle Du, D\phi \rangle \phi dx dt + w_2 \int_{Q_{2r}} (|Du|^2 + \varepsilon) \phi_t \phi dx dt \\ &\quad - 2\varepsilon w_3 \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\frac{\gamma-2}{2}} \langle D^2uDu - \Delta uDu, D\phi \rangle \phi dx dt \\ &\quad - 2\varepsilon w_4 \int_{Q_{2r}} u_t (|Du|^2 + \varepsilon)^{-1} \langle Du, D\phi \rangle \phi dx dt + \varepsilon w_4 \int_{Q_{2r}} \ln(|Du|^2 + \varepsilon) \phi_t \phi dx dt \\ &\quad - \frac{\varepsilon w_4}{2} \int_{B_{2r}} \ln(|Du(x, t_0)|^2 + \varepsilon) \phi^2(x, t_0) dx \end{aligned}$$

Above we dropped the nonpositive boundary term that appears when we integrate by parts with respect to time. However, we cannot drop the boundary term that appears from integrating by parts the last term of the right hand side of (A.6), because logarithm may change sign.

Next we take absolute values and employ again the estimate  $\varepsilon/(|Du|^2 + \varepsilon) \leq 1$  to arrive at

$$\begin{aligned} &\lambda \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\gamma/2}|D^2u|^2 \phi^2 dx dt \\ &\leq C \left( \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\gamma/2}|D^2u||Du||D\phi|\phi dx dt + \int_{Q_{2r}} |u_t||Du||D\phi|\phi dx dt \right. \\ &\quad + \int_{Q_{2r}} (|Du|^2 + \varepsilon)|\phi_t|\phi dx dt + \varepsilon \int_{Q_{2r}} |\ln(|Du|^2 + \varepsilon)||\phi_t|\phi dx dt \\ &\quad \left. + \varepsilon \int_{B_{2r}} |\ln(|Du(x, t_0)|^2 + \varepsilon)| \phi^2(x, t_0) dx \right), \end{aligned}$$

where  $C = C(n, p, \gamma, w_1, w_2, w_3, w_4) > 0$ . By Young’s inequality

$$\begin{aligned} & (\lambda - 2\eta) \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\gamma/2} |D^2u|^2 \phi^2 dx dt \\ & \leq \frac{C}{\eta} \int_{Q_{2r}} (|Du|^2 + \varepsilon)^{\gamma/2} |Du|^2 |D\phi|^2 dx dt + C \left( \int_{Q_{2r}} (|Du|^2 + \varepsilon) |\phi_t| \phi dx dt \right. \\ & \quad \left. + \varepsilon \int_{Q_{2r}} |\ln(|Du|^2 + \varepsilon)| |\phi_t| \phi dx dt + \varepsilon \int_{B_{2r}} |\ln(|Du(x, t_0)|^2 + \varepsilon)| \phi^2(x, t_0) dx \right) \end{aligned}$$

for any  $\eta > 0$  and some  $C = C(n, p, \gamma, w_1, w_2, w_3, w_4) > 0$ . Above we also employed estimate (A.4). Finally, we select  $\eta > 0$  small enough and employ (A.3) and (A.5) to arrive to the desired estimate.  $\square$

### Appendix B: Positive definiteness condition for the coefficient matrix

In the proof of Lemma 5.2, we wrote one of the key estimates as

$$|Du|^{p-2+s} \{w_1(p + s) |D_T |Du||^2 + Q\} \leq S$$

where

$$\begin{aligned} Q = & \left( w_2 - \frac{n-2}{n-1} w_1 \right) (\Delta_T u)^2 + w_2(p-1)(p-1+s-\gamma) (\Delta_\infty^N u)^2 \\ & + (w_2(2p-2+s-\gamma) - w_1(p+s)) \Delta_T u \Delta_\infty^N u. \end{aligned}$$

This can also be written as

$$Q = \langle \bar{x}, M\bar{x} \rangle,$$

where  $\bar{x} = (\Delta_T u, \Delta_\infty^N u)^T \in \mathbb{R}^2$  is a vector and

$$M = \begin{bmatrix} w_2 - \frac{n-2}{n-1} w_1 & \frac{1}{2} (w_2(2p-2+s-\gamma) - w_1(p+s)) \\ \frac{1}{2} (w_2(2p-2+s-\gamma) - w_1(p+s)) & w_2(p-1)(p-1+s-\gamma) \end{bmatrix}.$$

Then we stated that if  $w_1 = 1$  and the range condition is satisfied, we can select  $w_2 = w_2(n, p, \gamma, s) > 0$  in such a way that  $Q$  is positive definite, which then allows us to get rid of the excess terms. Next we prove this fact.

**Lemma B.1** *Let  $n \geq 2$ ,  $1 < p < \infty$ ,  $-1 < \gamma < \infty$ ,  $w_1 = 1$  and let  $M$  be as above. Then if*

$$s > \max \left\{ -1 - \frac{p-1}{n-1}, \gamma + 1 - p \right\},$$

*there is  $w_2 = w_2(n, p, \gamma, s) > 0$  such that  $M$  is uniformly bounded positive definite (with a uniform constant).*

**Proof** We will show that  $\det(M) > 0$  and  $w_2 - \frac{n-2}{n-1} > 0$  with uniform lower bound, and thus by Sylvester’s condition  $M$  is uniformly bounded positive definite with a uniform constant. We fix  $w_1 = 1$  and introduce the following shorthand notation,

$$P := p - 1 \quad \text{and} \quad K := \gamma + 1,$$

and

$$G := p - 1 + s - \gamma \quad \text{and} \quad E := s + 1 + \frac{p - 1}{n - 1}.$$

We observe that  $P, K, G, E > 0$  under the assumptions of the lemma. Using this notation, one has

$$M = \begin{bmatrix} w_2 - \frac{n - 2}{n - 1} & \frac{1}{2}(w_2(P + G) - (K + G)) \\ \frac{1}{2}(w_2(P + G) - (K + G)) & w_2 P \cdot G \end{bmatrix}.$$

Then we rewrite the determinant

$$\det(M) = a \left( w_2 - \frac{n - 2}{n - 1} \right)^2 + b \left( w_2 - \frac{n - 2}{n - 1} \right) + c$$

where

$$a = -\frac{1}{4}(G - P)^2, \quad b = P \cdot E + \frac{1}{2}(G - P) \left( \frac{G}{n - 1} + K - \frac{(n - 2)P}{n - 1} \right)$$

and

$$c = -\frac{1}{4} \left( \frac{G}{n - 1} + K - \frac{(n - 2)P}{n - 1} \right)^2.$$

The discriminant of such a polynomial is

$$b^2 - 4ac = G \cdot P \cdot E \left( \frac{G}{n - 1} + K \right).$$

Notice that  $b^2 - 4ac > 0$  and hence our polynomial has two distinct roots, unless  $G = P$ , in which case our polynomial is of the first order and has one root. Moreover  $\det(M) > 0$  if and only if  $w_2 - \frac{n-2}{n-1}$  lies between these roots, that is,

$$Root_+ < w_2 - \frac{n - 2}{n - 1} < Root_-,$$

where

$$\begin{aligned} Root_{\pm} &= \frac{-\left( P \cdot E + \frac{1}{2}(G - P) \left( \frac{G}{n - 1} + K - \frac{(n - 2)P}{n - 1} \right) \right) \pm \sqrt{b^2 - 4ac}}{-\frac{1}{2}(G - P)^2} \\ &= \frac{\left( \sqrt{P \cdot E} \mp \sqrt{G \left( \frac{G}{n - 1} + K \right)} \right)^2}{(G - P)^2} \\ &= \left( \frac{\sqrt{E}}{\sqrt{G} \pm \sqrt{P}} - \frac{\sqrt{G} \left( \sqrt{E} - \sqrt{\frac{G}{n - 1} + K} \right)}{(\sqrt{G} + \sqrt{P})(\sqrt{G} - \sqrt{P})} \right)^2 \geq 0. \end{aligned}$$

These formulas are valid if  $G \neq P$ . Indeed, recall that  $a < 0$  if  $G \neq P$ , then

$$Root_- - Root_+ = -\frac{\sqrt{b^2 - 4ac}}{a}$$

$$= 4 \frac{\sqrt{G \cdot P \cdot E\left(\frac{G}{n-1} + K\right)}}{(G - P)^2} > 0.$$

On the other hand, by

$$\lim_{G \rightarrow P} E = \lim_{s \rightarrow \gamma} \left( s + 1 + \frac{p-1}{n-1} \right) = \frac{P}{n-1} + K,$$

and by l'Hopital's rule, one has

$$\lim_{G \rightarrow P} \frac{\sqrt{E} - \sqrt{\frac{G}{n-1} + K}}{\sqrt{G} - \sqrt{P}} = \frac{(n-2)\sqrt{P}}{(n-1)\sqrt{\frac{P}{n-1} + K}}.$$

We conclude that for the smaller root

$$Root_+ \xrightarrow{G \rightarrow P} \left( \frac{\sqrt{\frac{P}{n-1} + K}}{2\sqrt{P}} - \frac{(n-2)\sqrt{P}}{2(n-1)\sqrt{\frac{P}{n-1} + K}} \right)^2.$$

For the bigger root, it is easy to see that

$$Root_- \xrightarrow{G \rightarrow P} \infty.$$

The proof is finished.  $\square$

## References

1. Andrade, P.D.S., Santos, M.S.: Improved regularity for the parabolic normalized  $p$ -Laplace equation. *Calc. Var. Partial Differ. Equ.* **61**(5), 196 (2022)
2. Attouchi, A.: Local regularity for quasi-linear parabolic equations in non-divergence form. *Nonlinear Anal.* **199**, 112051 (2020)
3. Attouchi, A., Parviainen, M.: Hölder regularity for the gradient of the inhomogeneous parabolic normalized  $p$ -Laplacian. *Commun. Contemp. Math.* **20**(4), 1750035 (2018)
4. Attouchi, A., Ruosteenoja, E.: Gradient regularity for a singular parabolic equation in non-divergence form. *Discrete Contin. Dyn. Syst.* **40**(10), 5955–5972 (2020)
5. Banerjee, A., Garofalo, N.: Gradient bounds and monotonicity of the energy for some nonlinear singular diffusion equations. *Indiana Univ. Math. J.* **62**(2), 699–736 (2013)
6. Banerjee, A., Garofalo, N.: On the Dirichlet boundary value problem for the normalized  $p$ -Laplacian evolution. *Commun. Pure Appl. Anal.* **14**(1), 1–21 (2015)
7. Cianchi, A., Maz'ya, V.G.: Second-order regularity for parabolic  $p$ -Laplace problems. *J. Geom. Anal.* **30**(2), 1565–1583 (2020)
8. DiBenedetto, E.: *Degenerate Parabolic Equations*. Universitext. Springer, New York (1993)
9. Does, K.: An evolution equation involving the normalized  $p$ -Laplacian. *Commun. Pure Appl. Anal.* **10**(1), 361–396 (2011)
10. Dong, H., Peng, F., Zhang, Y., Zhou, Y.: Hessian estimates for equations involving  $p$ -Laplacian via a fundamental inequality. *Adv. Math.* **370**, 107212 (2020)
11. Feng, Y., Parviainen, M., Sarsa, S.: On the second-order regularity of solutions to the parabolic  $p$ -Laplace equation. *J. Evol. Equ.*, **22**(1), Paper No. 6, 17 pp (2022)
12. Giga, Y.: *Surface Evolution Equations: A Level Set Approach*, vol. 99 of *Monographs in Mathematics*. Birkhauser Verlag (2006)
13. Høeg, F.A., Lindqvist, P.: Regularity of solutions of the parabolic normalized  $p$ -Laplace equation. *Adv. Nonlinear Anal.* **9**(1), 7–15 (2020)
14. Imbert, C., Jin, T., Silvestre, L.: Hölder gradient estimates for a class of singular or degenerate parabolic equations. *Adv. Nonlinear Anal.* **8**(1), 845–867 (2019)
15. Jin, T., Silvestre, L.: Hölder gradient estimates for parabolic homogeneous  $p$ -Laplacian equations. *J. Math. Pures Appl.* **108**(9), 63–87 (2017)

16. Juutinen, P., Lindqvist, P., Manfredi, J.J.: On the equivalence of viscosity solutions and weak solutions for a quasi-linear equation. *SIAM J. Math. Anal.* **33**(3), 699–717 (2001)
17. Kurkinen, T., Parviainen, M., Siltakoski, J.: Elliptic harnack's inequality for a singular nonlinear parabolic equation in non-divergence form. arXiv preprint [arXiv:2202.02350](https://arxiv.org/abs/2202.02350) (2022)
18. Lindqvist, P.: On the time derivative in a quasilinear equation. *Skr. K. Nor. Vidensk. Selsk.* **2**, 1–7 (2008)
19. Lindqvist, P.: On the time derivative in an obstacle problem. *Rev. Mat. Iberoam.* **28**(2), 577–590 (2012)
20. Lindqvist, P.: The time derivative in a singular parabolic equation. *Differ. Integral Equ.* **30**(9–10), 795–808 (2017)
21. Manfredi, J.J., Parviainen, M., Rossi, J.D.: An asymptotic mean value characterization for  $p$ -harmonic functions. *Proc. Am. Math. Soc.* **138**, 881–889 (2010)
22. Ohnuma, M., Sato, K.: Singular degenerate parabolic equations with applications to the  $p$ -Laplace diffusion equation. *Commun. Partial Differ. Equ.* **22**(3–4), 381–411 (1997)
23. Parviainen, M., Ruosteenoja, E.: Local regularity for time-dependent tug-of-war games with varying probabilities. *J. Differ. Equ.* **261**(2), 1357–1398 (2016)
24. Parviainen, M., Vázquez, J.L.: Equivalence between radial solutions of different parabolic gradient-diffusion equations and applications. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, **21**, 303–359 (2020)
25. Sarsa, S.: Note on an elementary inequality and its application to the regularity of  $p$ -harmonic functions. *Ann. Fenn. Math.* **47**(1), 139–153 (2022)

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