ON THE HÖLDER REGULARITY OF SIGNED SOLUTIONS TO A DOUBLY NONLINEAR EQUATION

VERENA BÖGELEIN, FRANK DUZAAR, AND NAIAN LIAO

ABSTRACT. We establish the interior and boundary Hölder continuity of possibly sign-changing solutions to a class of doubly nonlinear parabolic equations whose prototype is

\[ \partial_t (|u|^{p-2}u) - \Delta_p u = 0, \quad p > 1. \]

The proof relies on the property of expansion of positivity and the method of intrinsic scaling, all of which are realized by De Giorgi’s iteration. Our approach, while emphasizing the distinct roles of sub(super)-solutions, is flexible enough to obtain the Hölder regularity of solutions to initial-boundary value problems of Dirichlet type or Neumann type in a cylindrical domain, up to the parabolic boundary. In addition, based on the expansion of positivity, we are able to give an alternative proof of Harnack’s inequality for non-negative solutions. Moreover, as a consequence of the interior estimates, we also obtain a Liouville-type result.

CONTENTS

1. Introduction and Main Results 2
   1.1. Interior Regularity 2
   1.2. Boundary Regularity 3
   1.3. Novelty and Significance 4
   1.4. Notations and Definitions 5
2. Some Technical Tools 7
3. Energy Estimates 8
4. Expansion of Positivity 12
   4.1. Propagation of Positivity in Measure 14
   4.2. A Shrinking Lemma 15
   4.3. A DeGiorgi-type Lemma 17
   4.4. Proof of the Expansion of Positivity 19
5. Proof of Theorem 1.1 When $1 < p < 2$ 19
   5.1. The Proof Begins 19
   5.2. Reduction of Oscillation Near Zero 20
6. Proof of Theorem 1.1 When $p > 2$ 22
   6.1. The Proof Begins 22
   6.2. Reduction of Oscillation Near Zero–Part I 22
   6.3. Reduction of Oscillation Near Zero–Part II 24
   6.4. Reduction of Oscillation Near Zero Concluded 29
   6.5. Reduction of Oscillation Away From Zero 29
7. Proof of Boundary Regularity 30
   7.1. Proof of Theorem 1.2 31
   7.2. Proof of Theorem 1.3 33
   7.3. Proof of Theorem 1.4 34
Appendix A. More on the Notion of Parabolicity 36
Appendix B. Harnack’s Inequality 38
References 39

2010 Mathematics Subject Classification. 35K65, 35K67, 35B65.
Key words and phrases. Doubly nonlinear parabolic equations, signed solutions, intrinsic scaling, expansion of positivity, Hölder continuity.
1. **Introduction and Main Results**

Let $E$ be an open set in $\mathbb{R}^N$. For $T > 0$ let $E_T$ denote the cylindrical domain $E \times (0, T]$. We shall consider quasi-linear, parabolic partial differential equations of the form

\begin{equation}
\partial_t (|u|^{p-2}u) - \text{div} A(x, t, u, Du) = 0 \quad \text{weakly in } E_T
\end{equation}

where the function $A(x, t, u, \zeta): E_T \times \mathbb{R}^{N+1} \rightarrow \mathbb{R}^N$ is only assumed to be measurable with respect to $(x, t) \in E_T$ for all $(u, \zeta) \in \mathbb{R} \times \mathbb{R}^N$, continuous with respect to $(u, \zeta)$ for a.e. $(x, t) \in E_T$, and subject to the structure conditions

\begin{equation}
\begin{cases}
A(x, t, u, \zeta) \cdot \zeta \geq C_o |\zeta|^p & \text{for a.e. } (x, t) \in E_T, \forall u \in \mathbb{R}, \forall \zeta \in \mathbb{R}^N, \\
|A(x, t, u, \zeta)| \leq C_1 |\zeta|^{p-1}
\end{cases}
\end{equation}

where $C_o$ and $C_1$ are given positive constants, and $p > 1$. The prototype equation is

\begin{equation}
\partial_t (|u|^{p-2}u) - \Delta_p u = 0 \quad \text{weakly in } E_T.
\end{equation}

Here $\Delta_p := \text{div}(|Du|^{p-2}Du)$ is the $p$-Laplace operator. When $p = 2$ it becomes the heat equation.

The motivations to study such an equation will be explored in Section 1.3. We however proceed to present our main results on the interior regularity in Section 1.1 and the boundary regularity in Section 1.2.

When we speak of the structural data, we refer to the set of parameters $\{p, N, C_o, C_1\}$. We also write $\gamma$ as a generic positive constant that can be quantitatively determined a priori only in terms of the data and that can change from line to line.

1.1. **Interior Regularity.** Let $\Gamma := \partial E_T - \overline{E} \times \{T\}$ be the parabolic boundary of $E_T$, and for a compact set $K \subset E_T$ introduce the parabolic $p$-distance from $K$ to $\Gamma$ by

$$\text{dist}_\rho(K; \Gamma) := \inf_{(x, t) \in K} \left\{ |x - y| + |t - s|^{\frac{1}{p}} \right\}.$$

For $\rho > 0$ let $K_\rho(x_o)$ be the cube with center at $x_o \in \mathbb{R}^N$ and edge $\rho$. When $x_o = 0$ we simply write $K_\rho$. We define backward cylinders scaled by a positive parameter $\theta$ by

$$(x_o, t_o) + Q_\theta(\theta) = (x_o, t_o) + K_\rho(0) \times (-\theta^p, 0] = K_\rho(x_o) \times (t_o - \theta^p, t_o].$$

If $\theta = 1$, we simply write $Q_\rho$.

We postpone the formal definition of local weak solution to Section 1.4. It is however noteworthy to mention here that local boundedness of local weak solutions is inherent in our notion of local solution (cf. Section 1.4.2). Thus we may always work with locally bounded solutions.

Now we state our main result concerning the interior Hölder continuity of weak solutions to (1.1), subject to the structure conditions (1.2).

**Theorem 1.1.** Let $u$ be a bounded, local, weak solution to (1.1)–(1.2) in $E_T$. Then $u$ is locally Hölder continuous in $E_T$. More precisely, there exist constants $\gamma > 1$ and $\beta \in (0, 1)$ that can be determined a priori only in terms of the data, such that for every compact set $K \subset E_T$,

$$|u(x_1, t_1) - u(x_2, t_2)| \leq \gamma \|u\|_{E_T} \left( \frac{|x_1 - x_2| + |t_1 - t_2|^{\frac{1}{p}}}{\text{dist}_\rho(K; \Gamma)} \right)^{\beta},$$

for every pair of points $(x_1, t_1), (x_2, t_2) \in K$.

**Remark 1.1.** We have stated Theorem 1.1 for globally bounded weak solutions. However the proof has a local thrust. As a matter of fact, we will show the following oscillation decay:

$$\text{ess osc}_{(x_o, t_o) + Q_\rho} u \leq \gamma \left( \text{ess osc}_{(x_o, t_o) + Q_\rho} u \left( \frac{r}{\theta} \right) \right)^{\beta},$$

for any pair of cylinders \((x_0, t_0) + Q_r \subset (x_0, t_0) + Q_{\rho} \subset E_T\). The conclusion of Theorem 1.1 can be derived from this oscillation estimate via a standard covering argument.

The oscillation decay in Remark 1.1, while local in nature, has a global implication. Indeed, let \(u\) be a bounded, local weak solution to (1.1) – (1.2) in the semi-infinite strip \(S_T := \mathbb{R}^N \times (-\infty, T)\) for some \(T \in \mathbb{R}\). Then we have

\[
\text{ess osc}_{(x_0, t_0) + Q_r} u \leq \gamma \|u\|_{\infty, S_T} \left( \frac{r}{\rho} \right)^\beta,
\]

for any pair of cylinders \((x_0, t_0) + Q_r \subset (x_0, t_0) + Q_{\rho} \subset S_T\). Now fixing \(r\) and letting \(\rho \to \infty\), we immediately arrive at a Liouville-type result.

**Corollary 1.1.** A bounded, local weak solution to (1.1) – (1.2) in \(S_T\) must be a constant.

**Remark 1.2.** One-sided boundedness of solutions in \(S_T\) is generally not sufficient to imply they are constants. This is evident from the non-negative solution \(u(x, t) = e^{x+t}\) to the one-dimensional heat equation.

**Remark 1.3.** The global boundedness condition in Corollary 1.1 can be easily relaxed to allow \(u\) to grow slower than \((|x| + |t|^{\frac{1}{2}})^\beta\) as \(|x| \to \infty\) and \(t \to -\infty\). Other variants of Liouville-type results may be obtained, for which we refer to [7, 8].

### 1.2. Boundary Regularity.

We will establish regularity of weak solutions to (1.1) – (1.2) up to the lateral boundary \(S_T := \partial E \times (0, T]\), provided the solution satisfies proper Dirichlet or Neumann boundary data and \(\partial E\) possesses certain geometry or smoothness. Likewise, regularity of weak solutions up to the initial level \(t = 0\) can also be obtained, provided the given initial value is regular enough.

The arguments employed will be local in nature. As a result, it suffices to require the boundary data to be taken just on a portion of the parabolic boundary. Nevertheless, we choose to present the results globally for simplicity, in terms of initial-boundary value problems.

To this end, let us first consider formally the following initial-boundary value problem of Dirichlet type:

(1.4)

\[
\begin{aligned}
\partial_t \left( |u|^{p-2} u \right) - \text{div} \ A(x, t, u, Du) & = 0 \quad \text{weakly in } E_T, \\
\frac{\partial u(\cdot, t)}{\partial \partial E} & = g(\cdot, t) \quad \text{for a.e. } t \in (0, T], \\
u(\cdot, 0) & = u_0(\cdot),
\end{aligned}
\]

where the structure conditions (1.2) are retained. Regarding the Dirichlet datum \(g\) and the initial datum \(u_0\), we assume

(\textbf{D}) \quad \text{\(g \in L^p(0, T; W^{1,p}(E))\), and \(g\) is continuous on \(S_T\) with modulus of continuity \(\omega_g(\cdot)\);}  

(\textbf{I}) \quad \text{\(u_0\) is continuous in \(\mathbb{T} E\) with modulus of continuity \(\omega_o(\cdot)\).}

In order to establish Hölder regularity of \(u\) up to \(S_T\), we need to impose some geometric conditions on \(\partial E\). For this purpose, we introduce the property of *positive geometric density of \(\partial E\), i.e.,

\[
\begin{aligned}
\text{there exists } \alpha_+ \in (0, 1) \text{ and } \varrho_0 > 0, \text{ such that for all } x_0 \in \partial E, \text{ for } \\
every \text{cube } K_{\varrho}(x_0) \text{ and } 0 < \varrho \leq \varrho_0, \text{ there holds } \\
|E \cap K_{\varrho}(x_0)| \leq (1 - \alpha_+)|K_{\varrho}|.
\end{aligned}
\]

Intuitively, this means one can place an exterior cone whose vertex is attached to \(x_0\) (uniformly with respect to \(x_0\)).

Next, we consider the Neumann problem. In order to deal with possible variational data on \(S_T\), we assume \(\partial E\) is of class \(C^1\), such that the outward unit normal, which we denote by \(n\), is
defined on \( \partial E \). Let us consider the initial-boundary value problem of Neumann type:

\[
\begin{aligned}
& \partial_t (|u|^{p-2} u) - \text{div} \mathbf{A}(x, t, u, Du) = 0 \quad \text{weakly in } E_T, \\
& \mathbf{A}(x, t, u, Du) \cdot \mathbf{n} = \psi(x, t, u) \quad \text{on } S_T, \\
& u(\cdot, 0) = u_0(\cdot),
\end{aligned}
\]  

(1.6)

where the structure conditions (1.2) and the initial condition (1.1) are retained. On the Neumann datum \( \psi \) we assume for simplicity that, for some absolute constant \( C_2 \), there holds

\[
|\psi(x, t, u)| \leq C_2 \quad \text{for a.e. } (x, t, u) \in S_T \times \mathbb{R}.
\]

More general conditions should also work (cf. Sec. 2, Chap. II, [5]). The formal definitions of weak solutions to (1.4) and (1.6) will be given in Section 1.4. Now we are ready to present the results concerning regularity of solutions to (1.4) or (1.6) up to the parabolic boundary \( \Gamma' \).

1.2.1. Near the Initial Time.

**Theorem 1.2.** Let \( u \) be a bounded weak solution to the Dirichlet problem (1.4) under the assumption (1.2). Assume (D) holds. Then \( u \) is continuous in \( K \times [0, T] \) for any compact set \( K \subset E \). More precisely, there is a modulus of continuity \( \omega(\cdot) \), determined by the data, dist\((K, \partial E)\), \( \|u\|_{\infty, E_T} \) and \( \omega_0(\cdot) \), such that

\[
|u(x_1, t_1) - u(x_2, t_2)| \leq \omega\left(|x_1 - x_2| + |t_1 - t_2|^{\frac{1}{\beta}}\right),
\]

for every pair of points \((x_1, t_1), (x_2, t_2) \in K \times \mathbb{R}\). In particular, if \( u_0 \) is Hölder continuous with exponent \( \beta_0 \), then \( \omega(r) = \gamma r^\beta \|u\|_{\infty, E_T} \) with some \( \gamma > 0 \) and \( \beta \in (0, \beta_0) \) depending on the data, dist\((K, \partial E)\) and \( \beta_0 \).

1.2.2. Near \( S_T \)–Dirichlet Type Data.

**Theorem 1.3.** Let \( u \) be a bounded weak solution to the Dirichlet problem (1.4) under the assumption (1.2). Assume (D) and (1.5) hold. Then \( u \) is continuous in any compact set \( K \subset \overline{E_T} \). More precisely, there is a modulus of continuity \( \omega(\cdot) \), determined by the data, \( \alpha_s, g_o, d_0 \), dist\((K; \{t = 0\})\), \( \|u\|_{\infty, E_T} \) and \( \omega_0(\cdot) \), such that

\[
|u(x_1, t_1) - u(x_2, t_2)| \leq \omega\left(|x_1 - x_2| + |t_1 - t_2|^{\frac{1}{\beta}}\right),
\]

for every pair of points \((x_1, t_1), (x_2, t_2) \in K \). In particular, if \( g \) is Hölder continuous with exponent \( \beta_g \), then \( \omega(r) = \gamma r^\beta \|u\|_{\infty, E_T} \) with some \( \gamma > 0 \) and \( \beta \in (0, \beta_g) \) depending on the data, \( \alpha_s, g_o, \) dist\((K; \{t = 0\})\) and \( \beta_g \).

1.2.3. Near \( S_T \)–Neumann Type Data.

**Theorem 1.4.** Let \( u \) be a bounded weak solution to the Neumann problem (1.6). Assume \( \partial E \) is of class \( C^1 \) and (N) holds. Then \( u \) is Hölder continuous in any compact set \( K \subset \overline{E_T} \). More precisely, there exist constants \( \gamma > 1 \) and \( \beta \in (0, 1) \) determined by the data, \( C_2, \) dist\((K; \{t = 0\})\) and the structure of \( \partial E \), such that

\[
|u(x_1, t_1) - u(x_2, t_2)| \leq \gamma ||u||_{\infty, E_T} \left(|x_1 - x_2| + |t_1 - t_2|^{\frac{1}{\beta}}\right)^\beta,
\]

for every pair of points \((x_1, t_1), (x_2, t_2) \in K \).

1.3. Novelty and Significance. The equation (1.1) – (1.2) has been referred to as a doubly nonlinear parabolic equation in the literature, due to the nonlinearity of both the solution and its spatial gradient. It is a particular form of a more general equation whose prototype is

\[
\partial_t (|u|^{m-1} u) - \Delta_p u = 0, \quad p > 1, \ m > 0.
\]

(1.7)
The interest in such an equation stems from its mathematical structure, in understanding doubly nonlinear phenomena that generate mixed types of degeneracy and/or singularity in partial differential equations, and its connection to physical models, including dynamics of glaciers ([26]), shallow water flows ([2, 9, 14]) and friction dominated flow in a gas network ([23]).

In particular, the prototype equation (1.3) is naturally connected to the nonlinear eigenvalue problem \( -\Delta_p u = \lambda |u|^{p-2} u \) (cf. [25]), which plays an important role in the nonlinear potential theory.

The equation (1.1) – (1.2) has been observed by Trudinger ([27]), via Moser’s iteration, to possess a Harnack inequality for non-negative solutions, analogous to the one for the heat equation. See also [12, 18]. Such a Harnack inequality has been used to establish the interior Hölder regularity for non-negative solutions in [20, 21].

Our main contribution is to remove the sign restriction on solutions for the Hölder regularity to hold. The Harnack inequality seems not applicable in this setting due to changing signs of solutions and the power-like nonlinearity with respect to the solution itself. Instead, we employ a more basic tool – expansion of positivity – to handle the current situation. Our approach emphasizes the different roles played by sub-solutions and super-solutions. As a by-product, the expansion of positivity also leads to an alternative proof of the Harnack inequality. See Appendix B. The interior estimates also give us a Liouville type result for global solutions, which seems new in the literature. Moreover, our approach is flexible enough to obtain the Hölder regularity of solutions to the initial-boundary value problems of both Dirichlet type and Neumann type, up to the parabolic boundary. As far as we know, the boundary regularity has not ever been dealt with in the literature even in the case of non-negative solutions.

Our proofs of Hölder regularity – interior or boundary – all unfold along two main cases, i.e., when the solution is close to zero or when it is away from zero, through comparisons between the oscillation and the supremum/infimum of the solution. In the first case, we will take advantage of the scaling invariant property of the equation and obtain the expansion of positivity – Proposition 4.1 – without intrinsic scaling techniques. This treatment parallels the classical parabolic theory \( (p = 2) \) in [22], the new input being that we need to trace the competition between the oscillation and the extrema of the solution (see Remark 4.4). Whereas in the second case, the solution behaves like the one to the parabolic \( p \)-Laplacian equation, i.e. \( u_t = \Delta_p u \). Thus this latter case hinges upon the possibility to treat such a degenerate \( (p > 2) \) or singular \((1 < p < 2)\) equation, for which we exploit the existing theory in [5, 8].

The Hölder regularity for doubly nonlinear equations has also been considered in [15, 16, 17, 28, 29], under various conditions on the structure of the equation. The local regularity theory for the doubly nonlinear equation (1.7) seems fragmented and it deserves future investigations.

1.4. Notations and Definitions.

1.4.1. Notion of Local Solution. A function

\[
(1.8) \quad u \in C((0, T; L^p_{\text{loc}}(E)) \cap L^p(0, T; W^{1,p}_{\text{loc}}(E))
\]

is a local, weak sub(super)-solution to (1.1) with the structure conditions (1.2), if for every compact set \( K \subset E \) and every sub-interval \([t_1, t_2] \subset (0, T)\]

\[
(1.9) \quad \int_K |u|^{p-2} u \zeta \, dx \bigg|_{t_1}^{t_2} + \int_{K \times (t_1, t_2)} \left[ -|u|^{p-2} u \zeta_t + A(x, t, u, Du) \cdot D\zeta \right] \, dxdt \leq (\geq) 0
\]

for all non-negative test functions

\[
\zeta \in W^{1,p}_{\text{loc}}(0, T; L^p(K)) \cap L^p(0, T; W^{1,p}_0(K)).
\]

This guarantees that all the integrals in (1.9) are convergent.

A function \( u \) that is both a local weak sub-solution and a local weak super-solution to (1.1) – (1.2) is a local weak solution.
1.4.2. Notion of Parabolicity and Local Boundedness of Solutions. For any $k \in \mathbb{R}$, let

$$(u - k)_- = \max\{-(u - k), 0\}, \quad (u - k)_+ = \max\{u - k, 0\}.$$ 

Accordingly, we notice that

$$k - (u - k)_- = \min\{u, k\}, \quad k + (u - k)_+ = \max\{u, k\}.$$ 

Using (1.2) and employing a similar method as in (A.2) of [14, Chapter II] or Lemma 1.1 of [8, Chapter 3], we can show that the equation (1.1) with (1.2) is parabolic, in the sense that

- whenever $u$ is a local weak sub(super)-solution,
- the function $k \pm (u - k)_\pm$ is a local weak sub(super)-solution, for all $k \in \mathbb{R}$.

We will give a proof of this claim in Appendix A. In particular, when $u$ is a local weak solution, $u_+$ and $u_-$ are non-negative, local weak sub-solutions to (1.1) – (1.2). Since it has been shown that non-negative, local sub-solutions are locally bounded, we may always work with locally bounded solutions. See [12, 18] in this regard.

1.4.3. Notion of Solution to the Dirichlet Problem. A function

$$u \in C(0, T; L^p(E)) \cap L^p(0, T; W^{1,p}(E))$$

is a weak sub(super)-solution to (1.4), if for every sub-interval $[t_1, t_2] \subset (0, T],

$$\int_E |u|^{p-2} u\zeta \, dx \bigg|_{t_1}^{t_2} + \int_{E \times (t_1, t_2)} \left[ - |u|^{p-2} u\zeta + A(x, t, u, Du) \cdot D\zeta \right] \, dx \, dt \leq (\geq) 0$$

for all non-negative test functions

$$\zeta \in W^{1,p}_{loc}(0, T; L^p(E)) \cap L^p_{loc}(0, T; W^{1,p}_{o}(E)).$$

Moreover, setting $\hat{p} := \min\{2, p\}$, the initial datum is taken in the sense that for any compact set $K \subset \mathbb{E}$,

$$\int_{K \times \{t\}} (u - u_0)^{\hat{p}} \, dx \to 0 \quad \text{as} \quad t \downarrow 0.$$ 

The Dirichlet datum $g$ is attained under $u \leq (\geq) g$ on $\partial E$ in the sense that the traces of $(u - g)_\pm$ vanish as functions in $W^{1,p}(E)$ for a.e. $t \in (0, T]$, i.e. $(u - g)_\pm \in L^p(0, T; W^{1,p}(E))$. Notice that no a priori information is assumed on the smoothness of $\partial E$.

A function $u$ that is both a weak sub-solution and a weak super-solution to (1.4) is a weak solution.

1.4.4. Notion of Solution to the Neumann Problem. A function

$$u \in C(0, T; L^p(E)) \cap L^p(0, T; W^{1,p}(E))$$

is a weak sub(super)-solution to (1.6), if for every compact set $K \subset \mathbb{R}^N$ and every sub-interval $[t_1, t_2] \subset (0, T],

$$\int_{K \cap E} |u|^{p-2} u\zeta \, dx \bigg|_{t_1}^{t_2} + \int_{(K \cap E) \times (t_1, t_2)} \left[ - |u|^{p-2} u\zeta + A(x, t, u, Du) \cdot D\zeta \right] \, dx \, dt$$

$$\leq (\geq) \int_{(K \cap \partial E) \times (t_1, t_2)} \psi(x, t, u)\zeta \, d\sigma \, dt$$

for all non-negative test functions

$$\zeta \in W^{1,p}_{loc}(0, T; L^p(K)) \cap L^p_{loc}(0, T; W^{1,p}_{o}(K)).$$

Here $d\sigma$ denotes the surface measure on $\partial E$. The Neumann datum $\psi$ is reflected in the boundary integral on the right-hand side. Moreover, the initial datum is taken as in the Dirichlet problem.

A function $u$ that is both a weak sub-solution and a weak super-solution to (1.6) is a weak solution.
Acknowledgement. V. Bögelein and N. Liao have been supported by the FWF-Project P31956-N32 “Doubly nonlinear evolution equations”.

2. Some Technical Tools

For $k, w \in \mathbb{R}$ we define two quantities

$$g_{\pm}(w, k) = \pm(p - 1) \int_{k}^{w} |s|^{p-2}(s - k)_{\pm} \, ds.$$  

Note that $g_{\pm}(w, k) \geq 0$. For $b \in \mathbb{R}$ and $\alpha > 0$, we will embolden $b^\alpha$ to denote the signed $\alpha$-power of $b$ as

$$b^\alpha = \begin{cases} 
|b|^{\alpha-1}b, & b \neq 0, \\
0, & b = 0.
\end{cases}$$  

(2.1)

The following lemma can be found in the literature; cf. [1, Lemma 2.2] for $\alpha \in (0, 1)$ and [13, inequality (2.4)] for $\alpha > 1$.

Lemma 2.1. For any $\alpha > 0$, there exists a constant $\gamma = \gamma(\alpha)$ such that, for all $a, b \in \mathbb{R}$, the following inequality holds true:

$$\frac{1}{\gamma} |b^\alpha - a^\alpha| \leq (|a| + |b|)^{\alpha-1}|b - a| \leq \gamma |b^\alpha - a^\alpha|.$$

Based on Lemma 2.1, we prove the following.

Lemma 2.2. There exists a constant $\gamma = \gamma(p)$ such that, for all $w, k \in \mathbb{R}$, the following inequality holds true:

$$\frac{1}{\gamma(|w| + |k|)^{p-2}} (w - k)_{\pm}^2 \leq g_{\pm}(w, k) \leq \gamma(|w| + |k|)^{p-2} (w - k)_{\pm}^2$$

Proof. We only consider $g_-$, since the estimate for $g_+$ is analogous. If $k \leq w$, then $g_-(w, k) = 0 = (w - k)_-$. Therefore it is enough to consider $w, k \in \mathbb{R}$ with $w < k$. Here, we have

$$g_-(w, k) = (p - 1) \int_{w}^{k} |s|^{p-2}(k - s) \, ds \geq (p - 1) \int_{w}^{\frac{1}{2}(k+w)} |s|^{p-2}(k - s) \, ds \geq \frac{p-1}{2}(k - w) \int_{w}^{\frac{1}{2}(k+w)} |s|^{p-2} \, ds.$$

Note that $p - 2 > -1$ and therefore the integral on the right-hand side exists. With Lemma 2.1 we thus obtain

$$g_-(w, k) \geq \frac{1}{2}(k - w)|s|^{p-2}s \int_{w}^{\frac{1}{2}(k+w)} \geq \frac{1}{\gamma(p)} (k - w) \left(|w| + \frac{1}{2}|k + w|\right)^{p-2} \left(\frac{1}{2}(k + w) - w\right) = \frac{1}{2\gamma(p)} (k - w)^2 \left(|w| + \frac{1}{2}|k + w|\right)^{p-2} \geq \frac{1}{\gamma(p)} (k - w)^2 \left(|w| + |k|\right)^{p-2}.$$
In the last line we have used \( \frac{1}{2}(|w| + |k|) \leq |w| + \frac{1}{2}|k + w| \leq 2(|w| + |k|) \). This proves the lower bound for \( g_-(w, k) \). For the upper bound we again apply Lemma 2.1 and obtain
\[
g_-(w, k) = (p - 1) \int_{w}^{k} |s|^{p-2}(k - s) \, ds
\leq (p - 1)(k - w) \int_{w}^{k} |s|^{p-2} \, ds
= (k - w)|s|^{p-2} |_{w}^{k}
\leq \gamma(p)(k - w)^2(|w| + |k|)^{p-2}.
\]
This finishes the proof of the lemma. \( \square \)

The time derivative of a weak solution exists in the sense of distribution only. However we often need to use \( u \) in the test function and thus the term \( u_t \) appears in the integral weak formulation of solution, which is not granted by the preset notion of solution. In order to overcome the lack of regularity in the time variable, we define the following mollification in time:
\[
[v]_{h}(x, t) \overset{\text{def}}{=} \frac{1}{\pi} \int_{0}^{t} e^{-\frac{|x - s|}{\tau}} v(x, s) \, ds \quad \text{for any } v \in L^1(E_T).
\]
Properties of this mollification can for instance be found in [19].

3. Energy Estimates

In this section we exploit the property of weak sub(super)-solutions in order to deduce certain energy estimates. We emphasize the different roles played by sub-solutions and super-solutions. When we state "\( u \) is a sub(super)-solution..." and use "\( \pm \)" or "\( \mp \)" in what follows, we mean the sub-solution corresponds to the upper sign and the super-solution corresponds to the lower sign in the statement.

First of all, we present energy estimates for local weak sub(super)-solutions defined in Section 1.4.1.

Proposition 3.1. Let \( u \) be a local weak sub(super)-solution to (1.1) – (1.2) in \( E_T \). There exists a constant \( \gamma(C_0, C_1, p) > 0 \), such that for all cylinders \( Q_{R,S} = K_R(x_o) \times (t_o - S, t_o) \subseteq E_T \), every \( k \in \mathbb{R} \), and every non-negative, piecewise smooth cutoff function \( \zeta \) vanishing on \( \partial K_R(x_o) \times (t_o - S, t_o) \), there holds
\[
es_\sup_{t_o - S < t < t_o} \int_{K_R(x_o) \times \{ t \}} \zeta^p g_{\pm}(u, k) \, dx + \iint_{Q_{R,S}} \zeta^p |D(u - k)_{\pm}|^p \, dx \, dt
\leq \gamma \iint_{Q_{R,S}} \left[ |u - k|^p |D\zeta|^p + g_{\pm}(u, k)|\partial_t \zeta|^p \right] \, dx \, dt + \int_{K_R(x_o) \times (t_o - S)} \zeta^p g_{\pm}(u, k) \, dx.
\]

Proof. We only consider the case of a local weak sub-solution. Recall the definition (2.1) for the \( \alpha \)-power of a possibly negative number and define \( w_h \) via \( w_h^{\alpha - 1} := \|u^{\alpha - 1}\|_h \), where \( \|\cdot\|_h \) denotes the time mollification from (2.2). From the weak form of the differential inequality for sub-solutions we deduce the mollified version (cf. [19])
\[
\iint_{E_T} \left[ \partial_t w_h^{\alpha - 1} \varphi + [A(x, t, u, Du)]_h \cdot Du \right] \, dx \, dt
\leq \int_{E_T} w_h^{\alpha - 1}(x, 0) \cdot \frac{1}{\pi} \int_{0}^{T} e^{-\frac{\pi}{2}} \varphi(x, s) \, ds \, dx,
\]
for any non-negative test function \( \varphi \in L^p(0, T; W_0^1(E)) \). Let \( Q_{R,S} = K_R(x_o) \times (t_o - S, t_o) \subseteq E_T \) as in the statement of the energy estimate. Let \( \zeta \in C^1(Q_{R,S}, [0, 1]) \) be a cutoff function
vanishing on \( \partial K_R(x_o) \times (t_o - S, t_o) \). Furthermore, for fixed \( t_o - S < t_1 < t_2 < t_o \) and \( \epsilon > 0 \) small enough we define the cutoff function in time \( \psi_{\epsilon} \in W^{1,\infty}((t_o - S, t_o), [0, 1]) \) by

\[
\psi_{\epsilon}(t) := \begin{cases} 
0, & \text{for } t_o - S \leq t \leq t_1 - \epsilon, \\
1 + \frac{t - t_1}{\epsilon}, & \text{for } t_1 - \epsilon < t \leq t_1, \\
1, & \text{for } t_1 < t \leq t_2, \\
1 - \frac{t - t_2}{\epsilon}, & \text{for } t_2 < t \leq t_2 + \epsilon, \\
0, & \text{for } t_2 + \epsilon < t \leq t_o.
\end{cases}
\]

Now, we choose in (3.1) the testing function

\[
Q_{R,S} \ni (x, t) \mapsto \varphi(x, t) = \zeta^p(x, t)\psi_{\epsilon}(t)(u(x, t) - k)_+.
\]

In the following we omit in the notation the reference to the center \( z_o = (x_o, t_o) \). Moreover, we observe that

\[
g_+(w, k) = (p - 1) \int_k^w |s|^{p-2}(s - k)_+ \, ds = \int_k^{\frac{w}{p-1}} \left( s^{\frac{1}{p-1}} - k \right)_+ \, ds,
\]

which can be seen by substituting \( \sigma := s^{p-1} \). Note that the mapping \( \mathbb{R} \ni s \mapsto \phi(s) = s^{p-1} \) is increasing with derivative \( \phi'(s) = (p - 1)|s|^{p-2} (s \neq 0 \text{ in the case } p < 2) \). For the integral in (3.1) containing the time derivative we compute

\[
\begin{align*}
\int_{E_T} \partial_t w_h^{p-1}\varphi \, dxdt &= \int_{Q_{R,S}} \zeta^p \psi_{\epsilon} \partial_t w_h^{p-1}(w_h - k)_+ \, dxdt \\
&\quad + \int_{Q_{R,S}} \zeta^p \psi_{\epsilon} \partial_t w_h^{p-1}((u - k)_+ - (w_h - k)_+) \, dxdt \\
&\geq \int_{Q_{R,S}} \zeta^p \psi_{\epsilon} \partial_t g_+(w_h, k) \, dxdt \\
&= - \int_{Q_{R,S}} \left( \zeta^p \psi_{\epsilon}' + \psi_{\epsilon} \partial_t \zeta^p \right) g_+(w_h, k) \, dxdt.
\end{align*}
\]

Here we used in the second line the identity

\[
\partial_t w_h^{p-1} = \frac{1}{p}(u^{p-1} - w_h^{p-1}),
\]

and the fact that the map \( \tau \mapsto \left( \frac{\tau}{p-1} - k \right)_+ \) is a monotone increasing function, implying that the term in the second last line of the above inequality is non-negative. Since \( w_h^{p-1} = [u^{p-1}]_h \ni u^{p-1} \to w^{p-1} \) in \( L^{\frac{p}{p-1}}(\Omega_T) \) we can pass to the limit \( h \downarrow 0 \) in the integral on the right-hand side. We therefore get

\[
\begin{align*}
\lim_{h \downarrow 0} \int_{Q_{R,S}} \partial_t w_h^{p-1}\varphi \, dxdt &\geq \int_{Q_{R,S}} \left( \zeta^p \psi_{\epsilon}' + \psi_{\epsilon} \partial_t \zeta^p \right) g_+(u, k) \, dxdt \\
&= - [I_\varphi + II_\varphi],
\end{align*}
\]

with the obvious meaning of \( I_\varphi \) and \( II_\varphi \). We now pass to the limit \( \epsilon \downarrow 0 \). For the term \( I_\varphi \) we obtain for any \( t_o - S < t_1 < t_2 < t_o \) that

\[
\lim_{\epsilon \downarrow 0} I_\varphi = \int_{K_R} \zeta^p(x, t_1)g_+(u(x, t_1), k) \, dx - \int_{K_R} \zeta^p(x, t_2)g_+(u(x, t_2), k) \, dx,
\]

while for \( II_\varphi \) we have

\[
\lim_{\epsilon \downarrow 0} II_\varphi = \int_{K_R \times (t_1, t_2)} \partial_t \zeta^p g_+(u, k) \, dxdt.
\]
Next, we observe that the boundary term in (3.1) disappears as $h \downarrow 0$, since by construction \( \varphi(\cdot,0) \equiv 0 \) on \( E \), i.e. we have
\[
\lim_{h \downarrow 0} \int_E u^{p-1}(x,0) \cdot \frac{1}{h} \int_0^T e^{-\frac{t}{h}} \varphi(x,s) \, ds \, dx = \int_E u^{p-1}(x,0) \varphi(x,0) \, dx = 0.
\]

It remains to consider the diffusion term. After passing to the limit \( h \downarrow 0 \), we use the ellipticity and growth assumption (1.2) for the vector-field \( A \), and subsequently Young’s inequality to the integral containing \( (u - k)_+ \) and \( D(u - k)_+ \). In this way we obtain one term that we can absorb in the term arising from the ellipticity condition, the other one is shifted later on to the right hand side.

\[
\lim_{h \downarrow 0} \int_{E_T} [A(x,t,u,Du)]_h \cdot D\varphi \, dx \, dt
\]
\[
= \int_{Q_{R,S}} \psi \mathbf{A}(x,t,u,Du) \cdot [\xi^p D(u - k)_+ + \rho \xi^{p-1}(u - k)_+ D\xi] \, dx \, dt
\]
\[
\geq C_o \int_{Q_{R,S}} \xi^p \psi |D(u - k)_+|^p \, dx \, dt
\]
\[
- C_1 p \int_{Q_{R,S}} \xi^{p-1} \psi_x |D\xi|(u - k)_+ |D(u - k)_+|^{p-1} \, dx \, dt
\]
\[
\geq C_o \int_{Q_{R,S}} \xi^p \psi_x |D(u - k)_+|^p \, dx \, dt - \gamma \int_{Q_{R,S}} \xi^p |D\xi|^p (u - k)_+^p \, dx \, dt.
\]

Combining the preceding estimates and letting \( \varepsilon \downarrow 0 \) we arrive at
\[
\int_{K_R \times \{t_2\}} \xi^p \mathbf{g}_+(u,k) \, dx + \frac{C_o}{p} \int_{K_R \times \{t_1, t_2\}} \xi^p |D(u - k)_+|^p \, dx \, dt
\]
\[
\leq \int_{K_R \times \{t_1, t_2\}} \left[ |D\xi|^p (u - k)_+^p + \partial_t \xi^p \mathbf{g}_+(u,k) \right] \, dx \, dt + \int_{K_R \times \{t_1\}} \xi^p \mathbf{g}_+(u,k) \, dx,
\]
whenever \( t_o - S < t_1 < t_2 < t_o \). The constant \( \gamma \) in the first integral on the right hand side depends only on \( p, C_o \) and \( C_1 \). At this point, a standard argument finishes the proof. We first pass in the last inequality to the limit \( t_1 \downarrow t_o - S \). This poses no problem since \( u \in C(0,T; L^p_{loc}(E)) \). We obtain
\[
\int_{K_R \times \{t_2\}} \xi^p \mathbf{g}_+(u,k) \, dx + \frac{C_o}{p} \int_{K_R \times \{t_o - S, t_2\}} \xi^p |D(u - k)_+|^p \, dx \, dt
\]
\[
\leq \int_{Q_{R,S}} \left[ |D\xi|^p (u - k)_+^p + \partial_t \xi^p \mathbf{g}_+(u,k) \right] \, dx \, dt + \int_{K_R \times \{t_o - S\}} \xi^p \mathbf{g}_+(u,k) \, dx,
\]

Here, we discard the second integral on the left-hand side and take then the essential supremum with respect to \( t_2 \in (t_o - S, t_o) \). This leads to an estimate of the essential supremum of the first integral. On the other hand, discarding the first integral and passing to the limit \( t_2 \uparrow t_o \) we deduce a similar estimate for the second integral of the left-hand side. Together, this gives
\[
\text{ess sup}_{t_o - S < t < t_o} \int_{K_R \times \{t\}} \xi^p \mathbf{g}_+(u,k) \, dx + \frac{C_o}{p} \int_{K_R \times \{t_o - S, t\}} \xi^p |D(u - k)_+|^p \, dx \, dt
\]
\[
\leq \int_{K_R \times \{t_o - S\}} \left[ |D\xi|^p (u - k)_+^p + \partial_t \xi^p \mathbf{g}_+(u,k) \right] \, dx + \int_{K_R \times \{t_o - S\}} \xi^p \mathbf{g}_+(u,k) \, dx.
\]
This finishes the proof of the energy estimate. \( \square \)

Next, we consider the situation near the initial level \( t = 0 \) when a continuous datum \( u_o \) is prescribed. We work in a cylinder \( K_R(x_o) \times (0,S) \subset E_T \), which lies on the bottom of \( E_T \). Let \( \zeta \) be a non-negative, piecewise smooth cutoff function that is independent of \( t \) and vanishes
on \(\partial K_R(x_o)\). The conclusion of Proposition 3.1 holds in any cylinder satisfying \(K_R(x_o) \times (t_1, S) \subseteq E_T\). Suppose the level \(k\) satisfies

\[
\begin{cases}
    k \geq \sup_{K_R(x_o)} u_o & \text{for sub-solutions}, \\
    k \leq \inf_{K_R(x_o)} u_o & \text{for super-solutions}.
\end{cases}
\]

(3.4)

Then in view of the initial datum \(u_o\) being taken in the topology of \(L^p_{loc}(E)\) and letting \(t_1 \downarrow 0\), it is not hard to verify that the space integral on the right-hand side at the time level \(t_1\) will tend to zero. Consequently, we arrive at

**Proposition 3.2.** Let \(u\) be a local weak sub(super)-solution to (1.4) with (1.2) in \(E_T\). There exists a constant \(\gamma(C_o, C_1, p) > 0\), such that for all cylinders \(K_R(x_o) \times (0, S) \subseteq E_T\), every \(k \in \mathbb{R}\) satisfying (3.4) and every non-negative, piecewise smooth cutoff function \(\zeta\) independent of \(t\) and vanishing on \(\partial K_R(x_o)\), there holds

\[
\begin{align*}
\text{ess sup}_{0 < t < S} \int_{K_R(x_o) \times (t)} \zeta^p g_{\pm}(u, k) \, dx + \int_{K_R(x_o) \times (0, S)} \zeta^p |D(u - k)_{\pm}|^p \, dx dt \\
\leq \gamma \int_{K_R(x_o) \times (0, S)} (u - k)_{\pm} |D\zeta|^p \, dx dt
\end{align*}
\]

Now we turn our attention to the energy estimates near \(S_T\). We first deal with Dirichlet data. When we run the calculation in the proof of Proposition 3.1 within \(Q_{R,S} = K_R(x_o) \times (t_o - S, t_o)\) for some \((x_o, t_o) \in S_T\), we need to assume certain restrictions on the level \(k\), i.e.,

\[
\begin{cases}
    k \geq \sup_{Q_{R,S} \cap S_T} g & \text{for sub-solutions}, \\
    k \leq \inf_{Q_{R,S} \cap S_T} g & \text{for super-solutions}.
\end{cases}
\]

(3.5)

In such a way, the test functions in (3.2)

\[Q_{R,S} \cap E_T \ni (x, t) \mapsto \varphi(x, t) = \zeta^p (x, t) \psi_{\varepsilon}(t)(u(x, t) - k)_{\pm}\]

become admissible as the functions \(x \mapsto (u(x, t) - k)_{\pm}\) vanish on \(Q_{R,S} \cap S_T\) in the sense of traces for a.e. \(t \in (t_o - S, t_o)\). This fact does not require any smoothness of \(\partial E\) (cf. [11, Lemma 2.1]). As a result, we have

**Proposition 3.3.** Let \(u\) be a local weak sub(super)-solution to (1.4) with (1.2) in \(E_T\). There exists a constant \(\gamma(C_o, C_1, p) > 0\), such that for all cylinders \(Q_{R,S} = K_R(x_o) \times (t_o - S, t_o)\) with the vertex \((x_o, t_o) \in S_T\), every \(k \in \mathbb{R}\) satisfying (3.5), and every non-negative, piecewise smooth cutoff function \(\zeta\) vanishing on \(\partial K_R(x_o) \times (t_o - S, t_o)\), there holds

\[
\begin{align*}
\text{ess sup}_{t_o - S < t < t_o} \int_{K_R(x_o) \cap E_T} \zeta^p g_{\pm}(u, k) \, dx + \int_{Q_{R,S} \cap E_T} \zeta^p |D(u - k)_{\pm}|^p \, dx dt \\
\leq \gamma \int_{Q_{R,S} \cap E_T} [(u - k)_{\pm} |D\zeta|^p + g_{\pm}(u, k) |\partial_t \zeta|^p] \, dx dt \\
+ \int_{(K_R(x_o) \cap E_T) \times \{t_o - S\}} \zeta^p g_{\pm}(u, k) \, dx.
\end{align*}
\]

Finally, we deal with the energy estimates for the Neumann problem (1.6). Like before, we consider the problem in \(Q_{R,S} = K_R(x_o) \times (t_o - S, t_o)\) with \((x_o, t_o) \in S_T\) and may assume \((x_o, t_o) = (0, 0)\). For a cutoff function \(\zeta\) as in Proposition 3.3, a similar procedure as in the
proof of Proposition 3.1 will give us that
\[
\begin{align*}
\esssup_{t_o-S\leq t\leq t_o} \int_{(K_R \cap E_t) \times \{t\}} \zeta^p g_{+}(u, k) \, dx + \int_{Q_{R,S} \cap E_T} \zeta^p |D(u - k)_+|^p \, dx dt \\
\leq \gamma \int_{Q_{R,S} \cap E_T} [ (u - k)^p_+ |D\zeta|^p + \partial_t \zeta^p g_+ (u, k) ] \, dx dt \\
+ \gamma \int_{Q_{R,S} \cap S_T} \zeta^p \psi(x, t, u)(u - k)_+ \, d\sigma dt \\
+ \int_{(K_R \cap E_t) \times \{t_o - S\}} \zeta^p g_+ (u, k) \, dx.
\end{align*}
\]
Now we make use of (N), apply the trace inequality (cf. [4, Proposition 18.1]) for each time slice and then integrate in time, and use Young’s inequality to estimate the boundary integral:
\[
\begin{align*}
\int_{Q_{R,S} \cap S_T} \psi(x, t, u)(u - k)_+ \zeta^p \, d\sigma dt \\
\leq C_2 \int_{\partial(K_R \cap E_t) \times \{t_o - S, t_o\}} (u - k)_+ \zeta^p \, d\sigma dt \\
\leq \gamma C_2 \int_{Q_{R,S} \cap E_T} [D(u - k)_+ |\zeta|^p + (u - k)_+(\zeta^p + |D\zeta|^p)] \, dx dt \\
\leq \frac{C_2}{2p} \int_{Q_{R,S} \cap E_T} \zeta^p |D(u - k)_+|^p \, dx dt + \gamma \int_{Q_{R,S} \cap E_T} (u - k)^p_+ |D\zeta|^p \, dx dt \\
+ \gamma C_2^{\frac{1}{p}} \int_{Q_{R,S} \cap E_T} \zeta^p \chi_{\{u > k\}} \, dx dt.
\end{align*}
\]
Hence, collecting the above estimates we arrive at

**Proposition 3.4.** Let \( u \) be a local weak sub(super)-solution to (1.6) with (1.2) in \( E_T \). Assume \( \partial E \) is of class \( C^1 \) and (N) holds. There exists a constant \( \gamma(C_o, C_1, p) > 0 \), such that for all cylinders \( Q_{R,S} = K_R(x_o) \times (t_o - S, t_o) \) with the vertex \((x_o, t_o) \in S_T\), every \( k \in \mathbb{R} \), and every non-negative, piecewise smooth cutoff function \( \zeta \) vanishing on \( \partial K_R(x_o) \times (t_o - S, t_o) \), there holds
\[
\begin{align*}
\esssup_{t_o - S \leq t \leq t_o} \int_{(K_R(x_o) \cap E_t) \times \{t\}} \zeta^p g_{\pm}(u, k) \, dx + \int_{Q_{R,S} \cap E_T} \zeta^p |D(u - k)_\pm|^p \, dx dt \\
\leq \gamma \int_{Q_{R,S} \cap E_T} [ (u - k)^p_\pm |D\zeta|^p + g_{\pm}(u, k) |\partial_t \zeta|^p ] \, dx dt \\
+ \gamma C_2^{\frac{1}{p}} \int_{Q_{R,S} \cap E_T} \zeta^p \chi_{\{u - k > 0\}} \, dx dt \\
+ \int_{(K_R(x_o) \cap E_t) \times \{t_o - S\}} \zeta^p g_{\pm}(u, k) \, dx.
\end{align*}
\]

4. **Expansion of Positivity**

We first introduce the notation that is used throughout this section. For a compact set \( K \subset \mathbb{R}^N \) and a cylinder \( Q = K \times (T_1, T_2] \subset E_T \) we introduce numbers \( \mu_\pm \) and \( \omega \) satisfying
\[
\mu^+ \geq \esssup_Q u, \quad \mu^- \leq \essinf_Q u, \quad \omega \geq \mu^+ - \mu^-.
\]
We also assume \((x_o, t_o) \in Q\), such that the forward cylinder
\[
(4.1) \quad K_{\omega}(x_o) \times (t_o, t_o + (8\omega)^p] \subset Q.
\]
Next we state our main proposition of this section, which will be the main ingredient in the proof of Theorem 1.1.
Proposition 4.1. Let \( u \) be a locally bounded, local, weak sub(super)-solution to (1.1) – (1.2) in \( \mathbb{R}^T \). Suppose for some \((x_0, t_0) \in T\\mathbb{R}, M > 0, \alpha \in (0, 1)\) and \( \rho > 0 \) we have (4.1) and
\[
\left\{ \{\pm (\mu^\pm - u(t, t_0)) \geq M \right\} \cap K_\rho(x_0) \geq \alpha |K_\rho|.
\]
There exist constants \( \xi, \delta \) and \( \eta \) in \((0, 1)\) depending only on the data and \( \alpha \), such that either
\[
|\mu^\pm| > \xi M
\]
or
\[
\pm (\mu^\pm - u) \geq \eta M \quad \text{a.e. in } K_\rho(x_0) \times (t_0 + \delta \rho^p, t_0 + \delta \rho^p),
\]
where
\[
\xi = \begin{cases} 
2\eta, & \text{if } p > 2, \\
8, & \text{if } 1 < p \leq 2.
\end{cases}
\]

The proof of Proposition 4.1 is a straightforward consequence of Lemmas 4.1 – 4.3 in the following sections. Before presenting proofs, some remarks are in order.

Remark 4.1. By repeated applications of Proposition 4.1, we could conclude that for an arbitrary \( A > 1 \), there exists some \( \tilde{\eta} \in (0, 1) \) depending on \( \alpha \), the data and also on \( A \), such that
\[
\pm (\mu^\pm - u) \geq \tilde{\eta} M \quad \text{a.e. in } K_\rho(x_0) \times (t_0 + \rho^p, t_0 + A \rho^p),
\]
provided this cylinder is included in \( \mathbb{R}^T \) and \( |\mu^\pm| < \xi M \), where \( \xi \) is the parameter from Proposition 4.1.

Remark 4.2. Proposition 4.1 exhibits the spread of pointwise positivity both in time and in space. Nonetheless, in the proof of Hölder regularity we only need the positivity in a smaller cube than the one of the initial measure information assigned. Incidentally, Proposition 4.1 can be exploited to give an alternative proof of Harnack’s inequality for non-negative solutions (cf. [12, 18, 27]). We will present it in Appendix B.

Remark 4.3. The statement of Proposition 4.1 presents an either-or form. This is typical when power-like nonlinearity appears in an equation and as a result, adding a constant to a solution to (1.1) – (1.2) in general does not yield another solution to the same equation. We mention that a proposition of similar type has been used in [24] to deal with the Hölder regularity for the porous medium type equation.

Remark 4.4. Up to a proper adjustment of coefficients, the prototype equation of (1.1) – (1.2) can be written in a formal way as
\[
|u|^{p-2}u_t = \Delta_\rho u.
\]
Seemingly the equation is homogeneous in \( u \), and cylinders of the type \( Q_\rho = K_\rho \times (-\rho^p, 0] \) appear to be the correct ones to examine the equation. However, a careful inspection of proofs of Lemma 4.1 – 4.3 will reveal a subtle yet crucial difference between the role of \( u \) in the absolute value on the left-hand side and that of others. Loosely speaking, \(|u|\) represents the extrema, while other \( u \)'s stand for the oscillation. Notice also that when we apply Proposition 4.1 to prove Theorem 1.1, the typical \( M \) will be \( a\omega \) for some \( a \) in \((0, 1)\). Under this point of view, the above equation can be interpreted in a probably improper but heuristic manner that
\[
[|\mu|]^{p-2}[\omega] [t] = \frac{|\omega|^{p-1}}{|t|^{p}},
\]
assuming the symbols are self-suggestive. This hints the correct cylinder to examine the equation is actually
\[
Q_\rho(\theta) = K_\rho \times (-\theta \rho^p, 0], \quad \text{where } \theta = \left( \frac{|\omega|}{|\mu|} \right)^{2-p}.
\]
The time scaling by \( \theta \) reflects the competition between \(|\omega| \) and \(|\mu| \) via either-or alternatives. A closer inspection of the proofs of Lemma 4.1 and Lemma 4.3 shows it is required that \( \theta \approx 1 \), while the proof of Lemma 4.2 uses \( \theta \lesssim 1 \) only. This explains why the quantities \( \mu^\pm \) enter into
Proposition 4.1 via an either-or form, such that they are comparable with $\xi \omega$ for $\xi = 2\eta$ when $p > 2$, whereas $\xi = 8$ suffices when $1 < p < 2$.

4.1. Propagation of Positivity in Measure.

**Lemma 4.1.** Let $M > 0$ and $\alpha \in (0, 1)$. Then, there exist $\delta$ and $\varepsilon$ in $(0, 1)$, depending only on the data and $\alpha$, such that whenever $u$ is a locally bounded, local, weak sub(super)-solution to (1.1) – (1.2) in $E_T$ satisfying

\[
\left\{ \pm \left( \mu^\pm - u(\cdot, t_0) \right) \geq M \right\} \cap K_\varepsilon(x_0) \geq \alpha |K_\varepsilon|,
\]

then either

\[|\mu^\pm| > 8M\]

or

\[\left\{ \pm \left( \mu^\pm - u(\cdot, t) \right) \geq \varepsilon M \right\} \cap K_\varepsilon(x_0) \geq \frac{\alpha}{2} |K_\varepsilon| \quad \text{for all } t \in (t_0, t_0 + \delta \varrho^p].\]

**Proof.** We only show the case of super-solutions, the other case of sub-solutions being similar. Assume $(x_0, t_0) = (0, 0)$ and $|\mu^-| \leq 8M$. Otherwise there is nothing to prove. Use the energy estimate in Proposition 3.1 in the cylinder $Q = K_\varepsilon \times (0, \delta \varrho^p]$, with $k = \mu^- + M$ and choose a standard non-negative cutoff function $\zeta(x, t) \equiv \zeta(x)$ independent of time that equals 1 on $K_{(1-\sigma)}\varepsilon$ with $\sigma \in (0, 1)$ to be chosen later and vanishes on $\partial K_\varepsilon$ satisfying $|D\zeta| \leq (\sigma \varrho)^{-1}$; in such a case, we have for all $0 < t < \delta \varrho^p$, that

\[
\int_{K_\varepsilon \times \{t\}} \int_{k}^{0} |s|^{p-2}(s-k)_{-} \, ds \zeta^p \, dx
\]

\[
\leq \int_{K_\varepsilon \times \{0\}} \int_{u}^{k} |s|^{p-2}(s-k)_{-} \, ds \zeta^p \, dx + \gamma \int_{Q} (u-k)^p |D\zeta|^p \, dx \, dt.
\]

In order to estimate the first integral on the right-hand side, we take into consideration the measure theoretical information at the initial time $t = 0$ and the fact $u \geq \mu^-$. This leads to

\[
\int_{K_\varepsilon \times \{0\}} \int_{u}^{k} |s|^{p-2}(s-k)_{-} \, ds \zeta^p \, dx \leq (1 - \alpha) |K_\varepsilon| \int_{\mu^-}^{\mu} |s|^{p-2}(s-k)_{-} \, ds.
\]

The second term on the right-hand side of (4.3) is estimated by

\[
\int_{Q} (u-k)^p |D\zeta|^p \, dx \, dt \leq \frac{M^p}{(\sigma \varrho)^p} |Q| = \frac{\delta M^p}{\sigma \varrho^p} |K_\varepsilon|.
\]

The left-hand side of the energy estimate (4.3) can be bounded from below by

\[
\int_{K_\varepsilon \times \{t\}} \int_{k}^{0} |s|^{p-2}(s-k)_{-} \, ds \zeta^p \, dx \geq |A_{k_{\varepsilon, (1-\sigma)}\varepsilon}(t)| \int_{k_{\varepsilon}}^{0} |s|^{p-2}(s-k)_{-} \, ds
\]

where we have defined

\[A_{k_{\varepsilon, (1-\sigma)}\varepsilon}(t) = \left\{ u(\cdot, t) \leq k_{\varepsilon} \right\} \cap K_{(1-\sigma)}\varepsilon, \quad \text{and} \quad k_{\varepsilon} = \mu^- + \varepsilon M,
\]

with $\varepsilon \in (0, \frac{1}{2})$ to be chosen later. Due to Lemma 2.2 and the fact that $\frac{1}{2}M \leq (1 - \varepsilon)M = k - k_{\varepsilon} \leq |k_{\varepsilon}| + |k| \leq 2(|\mu^-| + M) \leq 18M$ we can further estimate from below

\[
\int_{k_{\varepsilon}}^{0} |s|^{p-2}(s-k)_{-} \, ds = \frac{1}{p-1} g_{-(k_{\varepsilon}, k)} \geq \frac{1}{\gamma(p)} (|k_{\varepsilon}| + |k|)^{p-2} (k - k_{\varepsilon})^2 \geq \frac{1}{\gamma(p)} M^p.
\]

Notice that

\[
|A_{k_{\varepsilon, \varrho}}(t)| = |A_{k_{\varepsilon, (1-\sigma)}\varepsilon}(t) \cup (A_{k_{\varepsilon, \varrho}}(t) \setminus A_{k_{\varepsilon, (1-\sigma)}\varepsilon}(t))| \\
\leq |A_{k_{\varepsilon, (1-\sigma)}\varepsilon}(t)| + |K_{\varrho} \setminus K_{(1-\sigma)}\varepsilon| \\
\leq |A_{k_{\varepsilon, (1-\sigma)}\varepsilon}(t)| + N\sigma |K_{\varrho}|,
\]
Collecting all the above estimates yields that
\[
|A_{k_{-},\varphi}(t)| \leq \int_{k_{-}}^{k_{+}} |s|^{p-2}(s-k)_- \, ds \left(1 -\alpha\right) |K_{\varphi}| + \frac{\gamma\delta}{\sigma^p} |K_{\varphi}| + N\sigma |K_{\varphi}|, \\
\]
for a constant \(\gamma = \gamma(p, C_\alpha, C_1)\). The fractional number in the preceding inequality can be rewritten in the form
\[
1 + I_{\varepsilon} \quad \text{where} \quad I_{\varepsilon} = \frac{\int_{k_{-}}^{k_{+}} |s|^{p-2}(s-k)_- \, ds}{\int_{k_{-}}^{k_{+}} |s|^{p-2}(s-k)_- \, ds}.
\]
At this stage, we need to bound \(I_{\varepsilon}\) from above. Keeping in mind \(|\mu^-| \leq 8M\) and \(|k_{-}| \leq 9M\) and applying Lemma 2.1, we have
\[
\int_{k_{-}}^{k_{+}} |\tau|^{p-2}(\tau-k)_- \, d\tau \leq M \int_{k_{-}}^{k_{+}} |\tau|^{p-2} \, d\tau = M |s|^{p-2} |s|_\mu^- \leq \gamma(p)M^{p}\varepsilon.
\]
Together with inequality (4.4) we obtain
\[
I_{\varepsilon} \leq \gamma(p)\varepsilon.
\]
This allows us to choose the various parameters quantitatively. Indeed, we may choose \(\varepsilon \in (0, 1)\) small enough such that
\[
(1 -\alpha)(1 + \gamma\varepsilon) \leq 1 - \frac{\alpha}{4}.
\]
This fixes \(\varepsilon\) as a constant depending only on \(p\) and \(\alpha\). Next, we define \(\sigma := \frac{\alpha}{8N}\). Finally, we choose \(\delta \in (0, 1)\) small enough so that
\[
\frac{\gamma\delta}{\sigma^p} \leq \frac{\alpha}{8}.
\]
Note that this specifies \(\delta\) as a constant depending on the data and \(\alpha\). With these choices we have
\[
|A_{k_{-},\varphi}(t)| \leq (1 - \frac{\theta}{\delta}) |K_{\varphi}|. \quad \text{This proves the asserted propagation of positivity (4.2), as long as 0 < t \leq \delta g^p.} \quad \square
\]

4.2. A Shrinking Lemma.

**Lemma 4.2.** Suppose that in Lemma 4.1 the second alternative (4.2) holds, let \(Q = K_{\varphi}(x_\alpha) \times (t_\alpha, t_\alpha + \delta g^p)\) be the corresponding cylinder and let \(\hat{Q} = K_{\dot{\varphi}}(x_\alpha) \times (t_\alpha, t_\alpha + \delta g^p) \subseteq E_T\). There exists \(\gamma > 0\) depending only on the data and \(\alpha\), such that for any positive integer \(j^*\), if \(1 < p < 2\), we have
\[
\left|\left\{\pm (\mu^+ - u) \leq \frac{\varepsilon M}{2j^*}\right\} \cap \hat{Q}\right| \leq \frac{\gamma}{j^*} |\hat{Q}|,
\]
while in the case \(p > 2\), the same conclusion holds provided \(|\mu^+| < \varepsilon M^{2-j^*}\).

**Proof.** We only show the case of super-solutions, the case of sub-solutions being similar. Moreover, we assume \((x_\alpha, t_\alpha) = (0, 0)\). We employ the energy estimate in Proposition 3.1 in \(K_{\varphi} \times (0, \delta g^p)\) with levels
\[
k_j := \mu^- + \frac{\varepsilon M}{2^j}, \quad j = 0, 1, \ldots, j^*,
\]
and introduce a cutoff function $\zeta$ in $K_{8\rho}$ (independent of $t$) that is equal to 1 in $K_{4\rho}$ and vanishes on $\partial K_{8\rho}$, such that $|D\zeta| \leq \frac{1}{\delta\rho}$. Then, we obtain

$$\int_Q |D(u - k_j)|^p \, dx \, dt \leq \int_{K_{8\rho} \times \{0\}} g_-(u, k_j) \, dx + \frac{\gamma}{\delta\rho^p} \int_{K_{8\rho} \times (0, \delta\rho^p)} (u - k_j)^p \, dx \, dt.$$  

Now we treat the individual terms of the right side separately. We begin with the first one. Due to Lemma 2.2 we have

$$g_-(u, k_j) \leq \gamma (|u| + |k_j|)^{p-2} (u - k_j)^2.$$  

When $p \geq 2$, we use $(u - k_j)_- \leq |u| + |k_j|$ as well as $u \geq \mu^-$ and $|\mu^-| \leq \varepsilon M^{2-\delta}$ to estimate

$$g_-(u, k_j) \leq \gamma (|u| + |k_j|)^p \chi_{\{u \leq k_j\}} \leq \gamma \left(\frac{\varepsilon M}{2}\right)^p.$$  

When $1 < p < 2$, we again use $(u - k_j)_- \leq |u| + |k_j|$ and $u \geq \mu^-$ to obtain

$$g_-(u, k_j) \leq \gamma (u - k_j)^p \leq \gamma \left(\frac{\varepsilon M}{2}\right)^p,$$

for a constant $\gamma$ depending only on $p$. This implies in particular that

$$g_-(u, k_j) \leq \frac{\gamma}{\delta\rho^p} \left(\frac{\varepsilon M}{2}\right)^p |Q|$$

holds true in any case. In the second integral appearing on the right-hand side of the energy estimate, we utilize the bound $(u - k_j)_- \leq \varepsilon M^{2-\delta}$. Therefore, in all cases the above estimate yields

$$\int_Q |D(u - k_j)|^p \, dx \, dt \leq \frac{\gamma}{\delta\rho^p} \left(\frac{\varepsilon M}{2}\right)^p |Q|.$$  

Next, we apply [5, Chapter I, Lemma 2.2] slice wise to $u(\cdot, t)$ for $t \in (0, \delta\rho^p)$ over the cube $K_q$, for levels $k_{j+1} < k_j$. Taking into account the measure theoretical information

$$\left|\left\{u(\cdot, t) \big| \mu^+ + \varepsilon M \right\} \cap K_q\right| \geq \frac{\alpha}{2} |K_q|$$

for all $t \in (0, \delta\rho^p)$, this gives

$$(k_j - k_{j+1}) \left|\left\{u(\cdot, t) < k_j \right\} \cap K_{4\rho}\right|$$

$$\leq \frac{\gamma\rho^{N+1}}{\left|\left\{u(\cdot, t) > k_j \right\} \cap K_{4\rho}\right|} \int_{\{k_j+1 < u(\cdot, t) < k_j\} \cap K_{4\rho}} |Du(\cdot, t)| \, dx$$

$$\leq \frac{\gamma\rho}{\alpha} \int_{\{k_j+1 < u(\cdot, t) < k_j\} \cap K_{4\rho}} |Du(\cdot, t)|^p \, dx \left\|\left\{k_j+1 < u(\cdot, t) < k_j\right\} \cap K_{4\rho}\right|^{\frac{1}{p}}$$

$$= \frac{\gamma\rho}{\alpha} \int_{K_{8\rho}} |D(u - k_j)_{-} (\cdot, t)|^p \, dx \left\|\left\{|A_j(t)| - |A_{j+1}(t)|\right\}\right\|^{\frac{1}{p}}.$$  

Here we used in the last line the short hand notation $A_j(t) := \left\{u(\cdot, t) < k_j\right\} \cap K_{4\rho}$. We now integrate the last inequality with respect to $t$ over $(0, \delta\rho^p)$ and apply Hölder’s inequality in time. With the abbreviation $A_j = \left\{u < k_j\right\} \cap \widehat{Q}$ this procedure leads to

$$\frac{\varepsilon M}{2^{1+p} |A_{j+1}|} \leq \frac{\gamma\rho}{\alpha} \left\|\int_{K_{8\rho}} |D(u - k_j)_{-} (\cdot, t)|^p \, dx \big| A_j \big| - |A_{j+1}|^{1+\frac{1}{p}} \right\|^{\frac{1}{p}}$$

$$\leq \frac{\gamma\rho}{\alpha} \frac{|\widehat{Q}|^{\frac{1}{p}}}{|A_{j+1}|^{\frac{1}{p}}} \left\|\left|A_j\right| - |A_{j+1}|\right\|^{1+\frac{1}{p}}.$$  

Recall that $\delta$ depends on the data and $\alpha$. Therefore $\gamma$ depends on only the data and $\alpha$. Now take the power $\frac{p}{p-1}$ on both sides of the above inequality to obtain

$$|A_{j+1}|^{\frac{1}{p}} \leq \gamma|\widehat{Q}|^{\frac{1}{p}} \left\|\left|A_j\right| - |A_{j+1}|\right\|.$$
Add these inequalities from 0 to $j_\ast - 1$ to obtain
\[ j_\ast |A_{j_\ast}|^{\frac{2s}{s-2}} \leq \gamma |\tilde{Q}|^{\frac{2s}{s-2}}. \]
From this we conclude
\[ |A_{j_\ast}| \leq \frac{\gamma}{j_\ast^\frac{s}{2}} |\tilde{Q}|. \]
This completes the proof. \( \Box \)

4.3. A DeGiorgi-type Lemma. Here we prove a DeGiorgi-type Lemma on cylinders of the form $Q_\theta(\theta)$. In the application $\theta$ will be a universal constant depending only on the data, in particular $\theta$ will be independent of the solution.

**Lemma 4.3.** Let $u$ be a locally bounded, local sub(super)-solution to (1.1) – (1.2) in $E_T$ and $(x_0, t_0) + Q_\theta(\theta) = K_\theta(x_0) \times (t_0 - \theta \delta^p, t_0] \subseteq E_T$. There exists a constant $\nu \in (0, 1)$ depending only on the data and $\theta$, such that if
\[ \left| \left\{ \pm (\mu^\pm - u) \leq M \right\} \cap (x_0, t_0) + Q_\theta(\theta) \right| \leq \nu |Q_\theta(\theta)|, \]
then either
\[ |\mu^\pm| > 8M, \]
or
\[ \pm (\mu^\pm - u) \geq \frac{1}{4}M \text{ a.e. in } (x_0, t_0) + Q_\frac{1}{4} \theta(\theta). \]

**Proof.** We prove the case of super-solutions only, the case of sub-solutions being similar. Assume $(x_0, t_0) = (0, 0)$ and $|\mu^-| \leq 8M$. Otherwise there is nothing to prove. In order to employ the energy estimate in Proposition 3.1, we notice first that due to Lemma 2.2 we have
\[ g_-(u, k) \leq \gamma (|u| + |k|)^{p-2}(u - k)_- \leq \gamma (|u| + |k|)^{p-1}(u - k)_- \]
and for $\tilde{k} < k$ there holds $(u - k)_- \geq (u - \tilde{k})_-$. Therefore, the energy estimate yields
\[ \text{ess sup } \int_{-\theta \delta^p < t < 0} \int_{K_\theta} \zeta^p (|u| + |k|)^{p-2}(u - k)_-^2 \, dx + \int_{Q_\theta(\theta)} \zeta^p |D(u - k)_-|^p \, dx \, dt \]
\[ \leq \gamma \int_{Q_\theta(\theta)} (u - k)_-^p |D\zeta|^p \, dx \, dt + \gamma \int_{Q_\theta(\theta)} (|u| + |k|)^{p-1}(u - k)_- |\partial_t \zeta|^p \, dx \, dt, \]
for any non-negative piecewise smooth cutoff function $\zeta$ vanishing on the parabolic boundary of $Q_\theta(\theta)$. In order to use this energy estimate, we set
\[
\begin{align*}
  k_n &= \mu^- + \frac{M}{2} + \frac{M}{2n+1}, \\
  \tilde{k}_n &= \frac{k_n + k_{n+1}}{2}, \\
  \vartheta_n &= \frac{\vartheta}{2} + \frac{\vartheta}{2n+1}, \\
  \tilde{\vartheta}_n &= \frac{\vartheta_n + \vartheta_{n+1}}{2}, \\
  K_n &= K_{\vartheta_n}, \\
  \tilde{K}_n &= \frac{K_n + K_{\tilde{\vartheta}_n}}{2}, \\
  Q_n &= Q_{\vartheta_n}(\theta), \\
  \tilde{Q}_n &= Q_{\tilde{\vartheta}_n}(\theta).
\end{align*}
\]
Recall that $Q_{\vartheta_n}(\theta) = K_n \times (-\vartheta_n \delta^p, 0]$ and $Q_{\tilde{\vartheta}_n}(\theta) = \tilde{K}_n \times (-\tilde{\vartheta}_n \delta^p, 0]$. Introduce the cutoff function $0 \leq \zeta \leq 1$ vanishing on the parabolic boundary of $Q_n$ and equal to identity in $\tilde{Q}_n$, such that
\[ |D\zeta| \leq \gamma \frac{2^n}{\vartheta} \text{ and } |\zeta| \leq \gamma \frac{2^n}{\vartheta \delta^p}. \]
In this setting, the energy estimate may be written as
\[
\text{ess sup}_{-\theta \partial t < t < 0} \int_{K_n} \left( |u| + |k_n| \right)^{p-2} (u - \tilde{k}_n)^2 \, dx + \int_{Q_n} |D(u - \tilde{k}_n)|^p \, dx dt
\leq \gamma \frac{2^{pn}}{\theta^p} \int_{Q_n} (u - k_n)^2 \, dx dt + \gamma \frac{2^{pn}}{\theta^p} \int_{Q_n} \left( |u| + |k_n| \right)^{p-1} \, dx dt
\leq \gamma \frac{2^{pn}}{\theta^p} M^p |A_n|,
\]
where $\gamma$ depends on the data and $\theta$. Here, we used $\mu^- \leq u \leq k_n \leq \mu^- + M$ on $A_n$, where
\[
A_n = \{ u < k_n \} \cap Q_n.
\]
On the other hand, we recall $|\mu^-| \leq 8M$, so that $u \leq \tilde{k}_n$ implies $|u| + |k_n| \leq 18M$ and $|u| + |k_n| \geq k_n - u \geq k_n - \tilde{k}_n = 2^{-(n+3)}M$. Inserting this above, we find that
\[
(4.6) \quad \frac{M^{p-2}}{2^{2n+3}} \text{ess sup}_{-\theta \partial t < t < 0} \int_{K_n} (u - \tilde{k}_n)^2 \, dx + \int_{Q_n} |D(u - \tilde{k}_n)|^p \, dx dt \leq \frac{2^{pn}}{\theta^p} M^p |A_n|.
\]
Now setting $0 \leq \phi \leq 1$ to be a cutoff function which vanishes on the parabolic boundary of $\tilde{Q}_n$ and equals the identity in $Q_{n+1}$, an application of the Hölder inequality and the Sobolev imbedding [5, Chapter I, Proposition 3.1] gives that
\[
\frac{M}{2^{n+1}} |A_{n+1}| \leq \int_{Q_n} (u - \tilde{k}_n)^2 \, dx dt
\leq \left[ \int_{Q_n} \left[ |(u - \tilde{k}_n) - \phi| \right]^{p \frac{N}{N+2}} \, dx dt \right]^\frac{N}{p(N+2)} |A_n|^{1 - \frac{N}{p(N+2)}}
\leq \gamma \left[ \int_{Q_n} |D[(u - \tilde{k}_n) - \phi]|^p \, dx dt \right]^\frac{N}{p(N+2)}
\times \left( \text{ess sup}_{-\theta \partial t < t < 0} \int_{K_n} (u - \tilde{k}_n)^2 \, dx \right)^\frac{1}{N+2} |A_n|^{1 - \frac{N}{N+2}}
\leq \gamma \frac{2^{pn}}{\theta^p} M^p \left( \frac{2^{p(2n+3)}}{\theta^p} M^2 \right)^\frac{1}{N+2} |A_n|^{1 + \frac{N}{N+2}}
= \gamma \frac{2^{2n+3}}{\theta^p} M |A_n|^{1 + \frac{N}{N+2}}.
\]
In the second last line we used the above energy estimate. In terms of $Y_n = |A_n|/|Q_n|$, this can be rewritten as
\[
Y_{n+1} \leq \gamma b^n Y_1^{1 + \frac{N}{N+2}},
\]
for a constant $\gamma$ depending only on the data and with $b \equiv \frac{2^{N+p+1}}{\theta^p}$. Hence, by [5, Chapter I, Lemma 4.1], there exists a positive constant $\nu$ depending only on the data, such that $Y_n \to 0$ if we require that $Y_0 \leq \nu$, which is the same as assuming
\[
|A_0| = \left| \{ u < k_0 \} \cap Q_0 \right| = \left| \{ u < \mu^- + M \} \cap Q_0(\theta) \right| \leq \nu |Q_0(\theta)|.
\]
Since $Y_n \to 0$ in the limit $n \to \infty$ we have
\[
\left| \{ u < \mu^- + \frac{1}{2} M \} \cap Q_{\frac{1}{2} \theta}(\theta) \right| = 0.
\]
This concludes the proof of the lemma. \qed
4.4. Proof of the Expansion of Positivity. We now have all the prerequisites at hand to prove the main result of this section.

Proof of Proposition 4.1. We only show the case of super-solutions, the other case of sub-solutions being similar. Assume $(x_o, t_o) = (0, 0)$. By $\delta, \varepsilon \in (0, 1)$ and $\gamma > 0$ we denote the corresponding constants from Lemma 4.1 and Lemma 4.2 depending on the data and $\alpha$ and by $\nu \in (0, 1)$ we denote the constant from Lemma 4.3 applied with $\theta = \delta$. Then, $\nu$ depends on the data and $\alpha$. Next, we choose an integer $j_\alpha$ in such a way that

$$\frac{\gamma}{j_\alpha} \leq \nu.$$

Then, $j_\alpha$ depends only on the data and $\alpha$. We let $\xi = 8$ if $1 < p \leq 2$ and $\xi = 2^{-j_\alpha}$ if $p > 2$. In the following we may assume that $|\mu^-| \leq \xi M$, since otherwise there is nothing to prove. Applying in turn Lemma 4.1 and Lemma 4.2 we infer that

$$\left| \left\{ u \leq \mu^- + \frac{\varepsilon M}{2j_\alpha} \right\} \cap \widehat{Q} \right| \leq \nu |\widehat{Q}|,$$

where $\widehat{Q} = K_{4\nu} \times (0, \delta \rho^p]$. Applying Lemma 4.3 with $M$ replaced by $\frac{\varepsilon M}{2j_\alpha}$ yields

$$u \geq \mu^- + \frac{\varepsilon M}{2j_\alpha} \quad \text{a.e. in } K_{2\nu} \times (\delta \rho^{p}, \delta \rho^p].$$

This proves the assertion of Proposition 4.1 for $\eta = \frac{\varepsilon M}{2j_\alpha}$ depending only on the data and $\alpha$. Let us point out in fact we have chosen $\xi = 2\eta$ when $p > 2$. \hfill \Box

5. Proof of Theorem 1.1 When $1 < p < 2$

5.1. The Proof Begins. Fix $(x_o, t_o) \in E_T$ and define

$$Q_o \overset{\text{def}}{=} K_{\rho}(x_o) \times (t_o - \rho^p, t_o] \subseteq E_T.$$

We may assume that $(x_o, t_o)$ coincides with the origin. Set

$$\mu^+ = \text{ess sup}_{Q_o} u, \quad \mu^- = \text{ess inf}_{Q_o} u, \quad \omega = \mu^+ - \mu^-.$$

Our proof unfolds along two main cases, namely

$$\begin{cases} 
  \text{when } u \text{ is near zero: } \mu^- \leq \omega \text{ and } \mu^+ \geq -\omega, \\
  \text{when } u \text{ is away from zero: } \mu^- > \omega \text{ or } \mu^+ < -\omega.
\end{cases}\tag{5.1}$$

Note that (5.1)$_1$ is equivalent to the condition that $-2\omega \leq \mu^- \leq \mu^+ \leq 2\omega$ and therefore $|\mu^\pm| \leq 2\omega$. When this case holds, we deal with it in Section 5.2 via a simple application of Proposition 4.1, thanks to the possibility to choose $\xi = 8$. No intrinsic scaling whatsoever is used in Section 5.2, though it seems unavoidable when we deal with the second case of (5.1)$_2$ in Section 5.3

5.2. Reduction of Oscillation Near Zero. In this section assume that the first case in (5.1) holds. Observe that one of the following must be true: either

$$\left| \left\{ u(\cdot, -\frac{1}{2} \rho^p) - \mu^- > \frac{1}{2} \omega \right\} \cap K_{\rho} \right| \geq \frac{1}{2} |K_{\rho}|, \tag{5.2}$$

or

$$\left| \left\{ \mu^+ - u(\cdot, -\frac{1}{2} \rho^p) > \frac{1}{2} \omega \right\} \cap K_{\rho} \right| \geq \frac{1}{2} |K_{\rho}|.$$

Since both cases can be treated similarly, we restrict ourselves to the case (5.2). As mentioned above, $|\mu^\pm| \leq 2\omega$ always holds. An application of Proposition 4.1 (note also Remark 4.1) gives $\eta \in (0, 1)$ depending only on the data, such that

$$u \geq \mu^- + \eta \omega \quad \text{a.e. in } Q_1 \overset{\text{def}}{=} K_{\rho_1} \times (-\rho^p_1, 0], \text{ with } \rho_1 = \frac{1}{2} \rho.$$
This yields a reduction of oscillation, i.e. we have
\[ \text{ess osc } u \leq (1 - \eta)\omega \overset{\text{def}}{=} \omega_1. \]
Now we may proceed by induction. Suppose up to \( i = 1, 2, \cdots j - 1 \), we have built
\[
\begin{cases}
\theta_i = \frac{1}{i}Q_{i-1}, \quad \omega_i = (1 - \eta)\omega_{i-1}, \quad Q_i = K_{\theta_i} \times (-\partial^p, 0], \\
\mu^-_i = \text{ess sup } u, \quad \mu^+_i = \text{ess inf } u, \quad \text{ess osc } u \leq \omega_i.
\end{cases}
\]
For all the indices \( i = 1, 2, \cdots j - 1 \), we always assume the first case in (5.1) holds, i.e.,
\[ \mu^-_i \leq \omega_i \quad \text{and} \quad \mu^+_i \geq -\omega_i. \]
In this way the argument at the beginning can be repeated and we have for all \( i = 1, 2, \cdots j \),
\[ \text{ess osc } u \leq (1 - \eta)\omega_{j-1} = \omega_j. \]
Consequently, iterating the above recursive inequality we obtain for all \( i = 1, 2, \cdots j \),
\[ \text{ess osc } u \leq (1 - \eta)^i\omega = \omega \left( \frac{\theta_i}{\theta} \right)^{\beta_0} \quad \text{where } \beta_0 = \frac{-\ln(1 - \eta)}{\ln 2}. \]

5.3. Reduction of Oscillation Away From Zero. In this section, let us suppose \( j \) is the first index satisfying the second case in (5.1), i.e.,
\[ \text{either } \mu^-_j > \omega_j \text{ or } \mu^+_j < -\omega_j. \]
Let us treat for instance \( \mu^-_j > \omega_j \), for the other case is analogous. We observe that since \( j \) is the first index for this to happen, one should have \( \mu^-_{j-1} \leq \omega_{j-1} \). Moreover, one estimates
\[ \mu^-_j \leq \mu^-_{j-1} + \omega_{j-1} - \omega_j \leq 2\omega_{j-1} - \omega_j = \frac{1 + \eta}{1 - \eta} \omega_j. \]
As a result, we have
\[ \omega_j < \mu^-_j \leq \frac{1 + \eta}{1 - \eta} \omega_j. \]
The condition (5.4) indicates that starting from \( j \) the equation (1.1) resembles the parabolic \( p \)-Laplacian type equation in \( Q_j \). Therefore, the reduction of oscillation hinges upon the possibility to treat the parabolic \( p \)-Laplacian type equation. To render this technically, we drop the suffix \( j \) from our notation temporarily for simplicity, and introduce \( v := u/\mu^- \) in \( Q = K \times (-\partial^p, 0] \).
It is straightforward to verify that \( v \) satisfies
\[ \partial_t v^{p-1} - \text{div } \hat{A}(x, t, v, Dw) = 0 \quad \text{weakly in } Q, \]
where, for \( (x, t) \in Q, v \in \mathbb{R} \) and \( \zeta \in \mathbb{R}^N \), we have defined
\[ \hat{A}(x, t, v, \zeta) = (\mu^-)^{1-p} A(x, t, \mu^- v, \mu^- \zeta), \]
which is subject to the structure conditions
\[
\begin{cases}
A(x, t, v, \zeta) \cdot \zeta \geq C_0 |\zeta|^p \\
|\hat{A}(x, t, v, \zeta)| \leq C_1 |\zeta|^{p-1}
\end{cases}
\quad \text{for a.e. } (x, t) \in Q, \forall v \in \mathbb{R}, \forall \zeta \in \mathbb{R}^N.
\]
Moreover,
\[ 1 \leq v \leq 2 \quad \text{a.e. in } Q. \]
In order to use the known regularity theory for the parabolic \( p \)-Laplacian (see [5, 8] for an account of the theory), it turns out to be more convenient to consider the equation satisfied by \( w := v^{p-1} \), i.e.
\[ \partial_t w - \text{div } \hat{A}(x, t, w, Dw) = 0 \quad \text{weakly in } Q, \]
where for \((x, t) \in Q, y \in \mathbb{R} \) and \(\zeta \in \mathbb{R}^N\), we have defined
\[
\tilde{A}(x, t, y, \zeta) = A\left(x, t, \frac{y}{\sqrt{p^{-1} + 1}}, \frac{1}{p^{-1}} \frac{\tilde{y}}{\tilde{y}^{p^{-1}}} \zeta\right).
\]
In the last line we used the abbreviation
\[
y \overset{\text{def}}{=} \min \{\max \left\{y, \frac{1}{2}\right\}, 2p\}.
\]
It is easy to see that \(w\) belongs to the same kind of functional space \((1.8)\) as \(u\) and \(v\) due to \((5.5)\) which yields \(1 \leq w \leq 2p^{-1}\) in \(Q\). Employing \((5.5)\) again one can verify that there exist absolute positive constants \(\tilde{C}_o = \gamma_o(p)C_o\) and \(\tilde{C}_1 = \gamma_1(p)C_1\), such that
\[
(5.7) \quad \tilde{A}(x, t, y, \zeta) \cdot \zeta \geq \tilde{C}_o|\zeta|^p \quad \text{and} \quad |\tilde{A}(x, t, y, \zeta)| \leq \tilde{C}_1|\zeta|^{p-1},
\]
for a.e. \((x, t) \in Q\), any \(y \in \mathbb{R}\), and any \(\zeta \in \mathbb{R}^N\). In other words, \(w\) is a local weak solution to the parabolic \(p\)-Laplacian type equation. First proved in [6] for \(p > 2\) and then in [3] for \(1 < p < 2\), the power-like oscillation decay for solutions to this kind of degenerate or singular parabolic equation is well known by now. The proofs in [3, 6] exploit the idea of intrinsic scaling. We state the conclusion in the following proposition in a form that favors our application, and refer to the monograph [5] for a comprehensive treatment of this issue.

**Proposition 5.1.** Let \(p > 1\). Suppose \(w\) is a bounded, local, weak solution to \((5.6) – (5.7)\) in \(Q := Q_\omega\) and define
\[
\tilde{\omega} = \text{ess osc}_{Q} w.
\]
If for some constant \(\sigma\) in \((0, 1)\), there holds
\[
(5.8) \quad \text{ess osc}_{Q_{\sigma}(\theta)} w \leq \tilde{\omega} \quad \text{where} \quad \theta = \tilde{\omega}^{2-p},
\]
then, there exist constants \(\beta_1\) in \((0, 1)\) and \(\gamma > 1\) depending only on the data \(N, p, \tilde{C}_o, \tilde{C}_1\) and \(\sigma\), such that for all \(0 < r < \rho\), we have
\[
\text{ess osc}_{Q_r(\theta)} w \leq \gamma \tilde{\omega} \left(\frac{r}{\theta}\right)^{\beta_1}.
\]

**Remark 5.1.** This proposition has been stated for all \(p > 1\). However the proofs in [6] for \(p > 2\) and in [3] for \(1 < p < 2\) are remarkably different. We mention a recent attempt in [24] to find a unified approach.

To use this proposition properly when \(1 < p < 2\), we first check the condition \((5.8)\) is satisfied. Indeed, recalling \(v = u/\mu^{-}\), \(w = v^{p-1}\) and \(\omega = \text{ess osc}_{Q} u\), we first use \((5.5)\) and the mean value theorem to obtain
\[
(p - 1)2^{p-2} \text{ ess osc}_{Q} v \leq \tilde{\omega} = \text{ess osc}_{Q} w \leq (p - 1) \text{ ess osc}_{Q} v.
\]
Since \(\text{ess osc}_{Q} v = \omega/\mu^{-}\), this amounts to
\[
(p - 1)2^{p-2} \frac{\omega}{\mu} \leq \tilde{\omega} \leq (p - 1) \frac{\omega}{\mu}.
\]
Then by \((5.4)\), we have
\[
c \overset{\text{def}}{=} \frac{1 - \eta}{1 + \eta} (p - 1)2^{p-2} \leq \tilde{\omega} \leq (p - 1) \leq 1.
\]
Now since \(\tilde{\omega} \leq 1\), we have \(Q_{\rho}(\theta_o) \subset Q_{\rho}\). Thus the condition \((5.8)\) in Proposition 5.1 is fulfilled for \(\sigma = 1\). As a result, the conclusion of Proposition 5.1 is obtained. Moreover, the above lower bound of \(\tilde{\omega}\) actually allows us to obtain the set inclusion
\[
Q_{r}(\theta_o) \subset Q_{r}(\theta) \quad \text{where} \quad \theta_o = c^{2-p}.
\]
Using this set inclusion and rephrasing the oscillation decay of Proposition 5.1 in terms of $u$, we have for all $0 < r < \rho$,

$$\text{ess osc } u_{Q_r(\theta_0)} \leq \gamma \omega \left( \frac{r}{\rho} \right)^{\beta_1}.$$ 

Now we revert to using the suffix $j$. The above oscillation estimate reads: for all $0 < r < \rho_j$, we have

$$(5.9) \quad \text{ess osc } u_{Q_r(\theta_0)} \leq \gamma \omega_j \left( \frac{r}{\rho_j} \right)^{\beta_1}.$$ 

Combining (5.3) and (5.9), we arrive at the desired conclusion, i.e., for all $0 < r < \rho$, there holds

$$\text{ess osc } u_{Q_r(\theta_0)} \leq \gamma \omega \left( \frac{r}{\rho} \right)^{\beta} \quad \text{where } \beta = \min \{\beta_0, \beta_1\}.$$ 

A proper rescaling gives the oscillation decay in Remark 1.1 and finishes the proof of Theorem 1.1 in the case $1 < p < 2$.

6. Proof of Theorem 1.1 When $p > 2$

6.1. The Proof Begins. Fix $(x_0, t_0) \in E_T$. Let $A \geq 1$ be determined later in terms of the data and $\rho > 0$ be so small that

$$Q_\rho \overset{\text{def}}{=} K_\rho(x_0) \times (t_0 - Ag^p, t_0] \subset E_T.$$ 

We may assume that $(x_0, t_0)$ coincides with the origin. Set

$$\mu^+ = \text{ess sup}_{Q_\rho} u, \quad \mu^- = \text{ess inf}_{Q_\rho} u, \quad \omega = \mu^+ - \mu^-.$$ 

Like when $1 < p < 2$, our proof unfolds along two main cases, namely

$$\left\{ \begin{array}{ll} \text{when } u \text{ is near zero: } \mu^- \leq \xi \omega \text{ and } \mu^+ \geq -\xi \omega; \\
\text{when } u \text{ is away from zero: } \mu^- > \xi \omega \text{ or } \mu^+ < -\xi \omega. \end{array} \right.$$ 

Strictly speaking, the above $\xi$ should be $\xi/8$ for $\xi$ chosen as in Proposition 4.1 (note also Remark 4.1) depending on the data, $A$, and $\alpha = 1/2 \nu$, whereas $\nu$ is the absolute constant determined in Lemma 4.3 with $\theta = 1$ there. It will be clear shortly from the proof where the various dependences come from. Meanwhile, we will keep using $\xi$ to denote $\xi/8$ for ease of notation bearing in mind the actual meaning of $\xi$, and this substitution will not spoil our reasoning in the following.

When $p > 2$, the number $\xi$ from Proposition 4.1 is in general a very small number. This brings additional technical complication to reducing the oscillation in the first case of (6.1). As we will see in Section 6.2 and Section 6.3, the method of intrinsic scaling is employed. Whereas in Section 6.5 where we deal the second case of (6.1), the treatment more or less parallels Section 5.3 for $1 < p < 2$.

6.2. Reduction of Oscillation Near Zero–Part I. In this section we assume the first case of (6.1) holds. We work with $u$ as a super-solution near its infimum.

Suppose that for some $\ell \in \left( - (A - 1)g^p, 0 \right]$,

$$(6.2) \quad \left| \left\{ u \leq \mu^- + \frac{1}{8} \omega \right\} \cap (0, \ell) + Q_\rho \right| \leq \nu \left| Q_\rho \right|,$$ 

where $\nu$ is the absolute constant appearing in Lemma 4.3. Taking $M = \frac{1}{8} \omega$, then according to Lemma 4.3, we have

$$u \geq \mu^- + \frac{1}{8} \omega \quad \text{a.e. in } (0, \ell) + Q_{\frac{1}{2} \rho}.$$
since the other alternative, i.e., $|\mu^-| \geq 2\omega$, does not hold due to (6.1). An application of Proposition 4.1 (note also Remark 4.1) applied with $2^pA$ instead of $A$ gives $\xi, \eta \in (0, 1)$ depending only on the data and $A$, such that either $|\mu^-| > \xi \omega$ or

$$u \geq \mu^- + \eta \omega \quad \text{a.e. in } \tilde{Q}_1 \equiv K_{\frac{2}{\omega}} \times (-\frac{1}{2}g)^p, 0].$$

If the above line holds, we immediately obtain a reduction of oscillation, i.e. we have

$$\text{ess osc } u \leq (1-\eta)\omega.$$ 

The case $\mu^- > \xi \omega$ does not hold due to (6.1). Therefore it remains to deal with the case $\mu^- < -\xi \omega$. Due to the restriction on $\mu^+$ in (6.1), we must also have $\mu^- > -2\omega$. Thus, we proceed further with the assumptions

$$u \geq \mu^- + \eta \omega \quad \text{a.e. in } K_{\frac{2}{\omega}}.$$ 

In the next lemma we establish that the pointwise information in (6.4) propagates to the top of the cylinder $\tilde{Q}_o$.

**Lemma 6.1.** Suppose the hypothesis (6.4) holds. Then there exists a constant $\eta_1 \in (0, 1)$ depending on $\xi, n$ and the data, such that

$$u \geq \mu^- + \eta_1 \omega \quad \text{a.e. in } K_{\frac{2}{\omega}} \times \left(\tilde{t} - \left(\frac{1}{2}g\right)^p, 0\right].$$

As a result, we have a reduction of oscillation

$$\text{ess osc } u \leq (1-\eta_1)\omega \quad \text{where } \tilde{Q}_1 = K_{\frac{2}{\omega}} \times \left(\tilde{t} - \left(\frac{1}{2}g\right)^p, 0\right].$$

**Proof.** For ease of notation we set $\tilde{t} - \left(\frac{1}{2}g\right)^p = 0$. Define $k_n, \tilde{k}_n, g_n, \tilde{g}_n, K_n$ and $\tilde{K}_n$, according to (4.5) (cf. the proof of Lemma 4.3) with $M$ and $g$ replaced by $2\eta_1\omega$ and $\frac{1}{2}g$ respectively, for some $0 < \eta_1 < \frac{\xi}{2}$ and $\theta > 0$ to be determined later. The only difference is that now the cylinders $Q_n$ and $\tilde{Q}_n$ are of forward type whose vertices are attached to the origin, i.e., $Q_n = K_n \times (0, \tilde{g}_n)$ and $\tilde{Q}_n = \tilde{K}_n \times (0, g_n^p)$). Since we know the “initial datum” at $t = 0$ is as in (6.4), we may choose a cutoff function $\tilde{\zeta}$ in $K_n$ independent of $t$, such that it equals 1 on $\tilde{K}_n$ and vanishes on $\partial K_n$, satisfying $|D\tilde{\zeta}| \leq \gamma 2^n g^{-1}$. Note that the boundary term at $t = 0$ on the right-hand side of the energy inequality vanishes on $K_n$, since

$$u(\cdot, 0) \geq \mu^- + \frac{1}{n} \omega \geq \mu^- + \eta_1 \omega \geq \mu^- + \eta_1 \omega + \eta_1 \omega = \frac{2\eta_1}{2} = k_n$$

a.e. on $K_{\frac{2}{\omega}}$. This requires $2\eta_1 < \frac{1}{n}$. In this way, the terms on the right-hand side of the energy estimate in Proposition 3.1 involving the initial time and $\tilde{\zeta}$ vanish. Thus, using the condition $-2\omega < \mu^- < -\xi \omega$, which leads to a lower bound for the sup-term in the energy estimate, we obtain

$$\omega^{p-2} \sup_{0 < t < \theta g^p} \int_{\tilde{K}_n} (u - \tilde{k}_n)^2 \, dx + \int_{\tilde{Q}_n} |D(u - \tilde{k}_n)|^p \, dx dt \leq \frac{2^{pn}}{g^p} (\eta_1 \omega)^p |A_n|,$$

where

$$A_n = \{ u < k_n \} \cap Q_n.$$ 

Now setting $\tilde{\zeta}$ to be a cutoff function which vanishes on the parabolic boundary of $\tilde{Q}_n$ and equals identity in $Q_{n+1}$, an application of the Sobolev imbedding [5, Chapter I, Proposition 3.1] with
\[ q = p \frac{N+2}{N} \]  and \( m = 2 \) gives that
\[
\left( \frac{\eta \omega}{2^{m+2}} \right)^p |A_{n+1}| \leq \int_{Q_n} \left[(u - \bar{k}_n)^p \zeta \right]^p \, dx \, dt
\]
\[
\leq \left[ \int_{Q_n} \left[(u - \bar{k}_n)^{-\zeta} \right]^p \, dx \, dt \right]^{\frac{N}{N+2}} |A_n|^{\frac{2}{N+2}}
\]
\[
\leq \gamma \left[ \int_{Q_n} \left| D \left[(u - \bar{k}_n)^{-\zeta} \right] \right|^p \, dx \, dt \right]^{\frac{N}{N+2}}
\]
\[
\times \left[ \text{ess sup}_{0 < t < \theta^0} \int_{K} (u - \bar{k}_n)^2 \, dx \right]^{\frac{1}{N+2}} |A_n|^{\frac{2}{N+2}}
\]
\[
\leq \gamma \omega^{\frac{2(p-1)}{N+2}} \left( \frac{2p}{\theta^0} (\eta \omega)^p \right)^{\frac{N}{N+2}} |A_n|^{1+\frac{2}{N+2}}.
\]

Setting \( Y_n = |A_n||Q_n| \), we arrive at
\[
Y_{n+1} \leq \gamma \left( \frac{2^{p+1} N}{m+2} \right)^n \left( \eta \omega \right)^{\frac{2(p-2)\theta}{\theta^0}} Y_n^{1+\frac{2}{N+2}} \equiv \gamma b^n \left( \eta \omega \right)^{\frac{2(p-2)\theta}{\theta^0}} Y_n^{1+\frac{2}{N+2}}.
\]

The meaning of \( b \) is clear in this context. Note that \( b \) and \( \gamma \) depend only on the data. Hence by [5, Chapter I, Lemma 4.1], there exists a constant \( \nu_0 \in (0, 1) \) depending only on the data, such that \( Y_n \to 0 \) in the limit \( n \to \infty \) if we require that
\[
Y_0 \leq \nu_0 \frac{2^{2-p} m}{\theta^0}.
\]

To finish the proof, we fix \( \theta = 2^p A \) and choose \( \eta_1 \) so small that \( \nu_0 \frac{2^{2-p} m}{\theta^0} \geq 1 \). The latter is implied if \( \eta_1^{p-2} < \frac{1}{2b^2} \). Together with the former bound for \( \eta_1 \) determined in the course of the proof, we have to require that
\[
\eta_1 < \min \left\{ \frac{1}{\nu_0} \frac{b}{2b^2}, \left( \frac{\nu_0}{2b^2 A} \right) \right\}.
\]

This proves the asserted claim. \( \square \)

6.3. Reduction of Oscillation Near Zero–Part II. In this section we still assume the first case of (6.1) holds. However, now we work with \( u \) as a sub-solution near its supreme.

Suppose contrary to (6.2) that
\[
(6.6) \quad \left\{ u \leq \mu^- + \frac{1}{4} \omega \right\} \cap (0, \bar{t}) + Q_\theta > \nu |Q_\theta|, \quad \forall \bar{t} \in \left(- (A - 1) \theta^p, 0 \right].
\]

Then for any such \( \bar{t} \), there exists some \( s \in \left[ \bar{t} - \theta^p, \bar{t} - \frac{1}{2} \nu \theta^p \right] \) with
\[
\left\{ u(\cdot, s) \leq \mu^- + \frac{1}{4} \omega \right\} \cap K_\theta > \frac{1}{2} \nu |K_\theta|.
\]

Indeed, if the above inequality does not hold for any \( s \) in the given interval, then
\[
\left\{ u \leq \mu^- + \frac{1}{4} \omega \right\} \cap (0, \bar{t}) + Q_\theta = \int_{\bar{t} - \theta^p}^{\bar{t} - \frac{1}{2} \nu \theta^p} \left\{ u(\cdot, s) \leq \mu^- + \frac{1}{4} \omega \right\} \cap K_\theta \, ds
\]
\[
+ \int_{\bar{t} - \frac{1}{2} \nu \theta^p}^{\bar{t}} \left\{ u(\cdot, s) \leq \mu^- + \frac{1}{4} \omega \right\} \cap K_\theta \, ds
\]
\[
< \frac{1}{2} \nu |K_\theta| \left( \theta^p - \frac{1}{2} \nu \theta^p \right) + \frac{1}{2} \nu \theta^p |K_\theta| < \nu |Q_\theta|,
\]
contradicting (6.6). Since \( \mu^+ - \frac{1}{4} \omega > \mu^- + \frac{1}{4} \omega \) always holds, this implies
\[
\left\{ u(\cdot, s) \leq \mu^+ - \frac{1}{4} \omega \right\} \cap K_\theta \geq \frac{1}{2} \nu |K_\theta|.
\]
By Proposition 4.1, there exist \( \xi, \eta_2 \in (0, 1) \), such that either \( |\mu^+| > \xi \omega \) or
\[
u \leq \mu^+ - \eta_2 \omega \quad \text{a.e. in } \bar{Q}_1,
\]
where \( \bar{Q}_1 \) is defined in (6.3). This implies a reduction of oscillation
\[
es osc u \leq (1 - \eta_2) \omega.
\]
The case \( \mu^+ < -\xi \omega \) does not hold due to (6.1)_1. Next we handle the case when \( \mu^+ > \xi \omega \). Due to the restriction on \( \mu^- \) in (6.1)_1, we must have \( \mu^+ \leq 2\omega \). Thus our assumptions for the following Sections 6.3.1 – 6.3.3 are
\[
\xi \omega \leq \mu^+ \leq 2\omega,
\]
and
\[
\begin{align*}
&\text{for any } t \in \left(-1(A - 1) \rho^p, 0\right) \text{ there exists } s \in \left[\tilde{t} - \rho^p, \tilde{t} - \frac{1}{2} \nu \rho^p\right] \\
&\text{such that } \left\{ u(\cdot, s) \leq \mu^+ - \frac{1}{2} \omega \right\} \cap K_\rho \geq \frac{1}{2} \nu |K_\rho|.
\end{align*}
\]

6.3.1. Propagation of Measure Theoretical Information.

Lemma 6.2. Suppose (6.7) and (6.8) are in force. There exists \( \varepsilon \in (0, 1) \), depending only on \( \nu \), \( \xi \) and the data, such that
\[
\left| \left\{ u(\cdot, t) \leq \mu^+ - \varepsilon \omega \right\} \cap K_\rho \right| \geq \frac{1}{2} \nu |K_\rho| \quad \text{for all } t \in (s, \tilde{t}].
\]

Proof. Assume \( s = 0 \) for ease of notation. Use the energy estimate in Proposition 3.1 in the cylinder \( Q := K_\rho \times \left(0, \delta \varepsilon^{2-p} \rho^p\right) \), with \( k = \mu^+ - \varepsilon \omega \) for some \( \delta > 0 \) and \( 0 < \varepsilon \leq \frac{1}{2} \xi \) to be determined later. Note that \( k \geq \frac{1}{2} \xi \omega \). Choose a standard non-negative time independent cutoff function \( \zeta(x, t) \equiv \zeta(x) \) that equals 1 on \( K_{(1 - \varepsilon) \rho} \) with \( \sigma \in (0, 1) \) and vanishes on \( \partial K_\rho \) satisfying \( |D\zeta| \leq (\sigma \rho)^{-1} \); in such a case, for all \( 0 < t < \delta \varepsilon^{2-p} \rho^p \) we have
\[
\int_{K_\rho \times \{t\}} \int_k^u s^{p-2} (s - k)_+ \, ds \, \zeta^p \, dx \
\leq \int_{K_\rho \times \{0\}} \int_k^u s^{p-2} (s - k)_+ \, ds \, \zeta^p \, dx + \gamma \int_Q (u - k)^p_+ |D\zeta|^p \, dx \, dt.
\]
The first term on the right is bounded from above by taking (6.8) into consideration. Indeed we have
\[
\int_{K_\rho \times \{0\}} \int_k^u s^{p-2} (s - k)_+ \, ds \, \zeta^p \, dx \leq (1 - \frac{1}{2} \nu^p) |K_\rho| \int_k^u s^{p-2} (s - k)_+ \, ds.
\]
The second term on the right is estimated by
\[
\int_Q (u - k)^p_+ |D\zeta|^p \, dx \, dt \leq \gamma \delta^{\frac{2}{\sigma^p} (\varepsilon \omega)^p |K_\rho| \leq \gamma \delta^{\frac{2}{\sigma^p} \varepsilon^{2-p} \rho^p |K_\rho|}.
\]
For the left-hand side, we estimate from below by
\[
\int_{K_\rho \times \{t\}} \int_k^u s^{p-2} (s - k)_+ \, ds \, \zeta^p \, dx \geq \left| \left\{ u(\cdot, t) > k_\varepsilon \right\} \cap K_{(1 - \varepsilon) \rho} \right| \int_k^{k_\varepsilon} s^{p-2} (s - k)_+ \, ds,
\]
where \( k_\varepsilon = \mu^+ - \varepsilon \omega \) for some \( \varepsilon \in (0, \frac{1}{2}) \). Noticing \( \xi \omega \leq \mu^+ \leq 2\omega \), we may estimate
\[
\int_k^{k_\varepsilon} s^{p-2} (s - k)_+ \, ds \geq \gamma \omega^{p-2} (\varepsilon \omega)^2 = \gamma \varepsilon^2 \omega^p.
\]
A similar consideration as in Lemma 4.1 then gives
\[
\left\{ u(\cdot, t) > k_\varepsilon \right\} \cap K_{(1-\sigma)\theta} \leq \frac{\int_{k_\varepsilon}^{\mu_+} s^{p-2}(s-k_\varepsilon) \, ds}{\int_{k_\varepsilon}^{\mu_+} s^{p-2}(s-k_\varepsilon) \, ds} (1 - \frac{1}{2}\nu) |K_{\theta}| + \frac{\gamma \delta}{\sigma^p} |K_{\theta}|.
\]

The fractional number of integral on the right can be rewritten as
\[
1 + I_\varepsilon \quad \text{where} \quad I_\varepsilon = \frac{\int_{k_\varepsilon}^{\mu_+} s^{p-2}(s-k_\varepsilon) \, ds}{\int_{k_\varepsilon}^{\mu_+} s^{p-2}(s-k_\varepsilon) \, ds}.
\]

We estimate by using \( \xi \omega \leq \mu^+ \leq 2\omega \) and \( k \geq \frac{1}{2}\xi \omega \) to obtain the bound \( I_\varepsilon \leq \frac{3}{2}\nu \), where \( \gamma \) depends only on \( p \). Inserting this above leads to the inequality
\[
\left\{ u(\cdot, t) > k_\varepsilon \right\} \cap K_{\theta} \leq (1 - \frac{1}{2}\nu) (1 + \gamma \varepsilon) |K_{\theta}| + \frac{\gamma \delta}{\sigma^p} |K_{\theta}| + N \sigma |K_{\theta}|.
\]

Now we first choose \( \varepsilon \) small enough so that
\[
(1 - \frac{1}{2}\nu) (1 + \gamma \varepsilon) \leq 1 - \frac{3}{4}\nu.
\]

This fixes \( \varepsilon \) in dependence on \( p \) and \( \nu \). Then we fix \( \sigma := \frac{1}{16N} \) and choose \( \delta \) small enough to have \( \frac{2\delta}{\sigma^p} \leq \frac{1}{4}\nu \). Finally, the parameter \( \varepsilon \) is chosen such that \( \delta \varepsilon^{2-p} \geq 1 \). The proof can now be finished by redefining \( \varepsilon \) as \( \varepsilon \).

Since \( \ell \) is arbitrary, we actually obtain the measure theoretical information
\[
\left\{ u(\cdot, t) \leq \mu^+ - \varepsilon \omega \right\} \cap K_{\theta} \geq \frac{1}{2}\nu |K_{\theta}| \quad \text{for all} \quad t \in \left( - (A - 1)\gamma^p, 0 \right].
\]

6.3.2. Shrinking the Measure Near the Supremum. By \( \varepsilon \in (0, 1) \) we denote the constant from Lemma 6.2 depending only on the data. The number \( A \) is still to be determined. We choose \( A \) in the form \( A = 2^{j_*}(p-2) + 1 \) with some \( j_* \) to be fixed later and define \( Q_\theta(\theta) = K_{\theta} \times (-\theta \gamma^p, 0] \) with \( \theta = 2^{j_*}(p-2) \).

Lemma 6.3. Suppose (6.7) and (6.9) hold. There exists \( \gamma > 0 \) depending only on the data, such that for any positive integer \( j_* \), we have
\[
\left| \left\{ u \geq \mu^+ - \frac{\varepsilon \omega}{2j_*} \right\} \cap Q_\theta(\theta) \right| \leq \frac{\gamma}{j_*} |Q_\theta(\theta)|.
\]

Proof. For \( j = 0, \ldots, j_* - 1 \) we employ the energy estimate in Proposition 3.1 in the cylinder \( K_{2\theta} \times (-\theta \gamma^p, 0] \) with levels \( k_j = \mu^+ - 2^{-j} \varepsilon \omega \) and a time independent cutoff function \( \zeta(x, t) \equiv \zeta(x) \), such that \( \zeta \) equals 1 in \( K_{\theta} \), vanishes on \( \partial K_{2\theta} \), and such that \( |D\zeta| \leq 2^{j-1} \). Then, we obtain
\[
\iint_{Q_\theta(\theta)} |D(u-k_j)_+|^p \, dx \, dt \leq \iint_{K_{2\theta} \times (-\theta \gamma^p)} \zeta^p \mathcal{L}_+(u, k_j) \, dx \, dy + \iint_{K_{2\theta} \times (-\theta \gamma^p, 0]} (u-k_j)^p_+ |D\zeta|^p \, dx \, dt.
\]

The first term on the right is estimated by Lemma 2.2 and using \( \xi \omega \leq \mu^+ \leq 2\omega \). We obtain
\[
\iint_{K_{2\theta} \times (-\theta \gamma^p)} \zeta^p \mathcal{L}_+(u, k_j) \, dx \leq \gamma \omega^{p-2} \left( \frac{\varepsilon \omega}{2j_*} \right)^2 |K_{2\theta}| \leq \gamma \omega^{p-2} \left( \frac{\varepsilon \omega}{2j_*} \right)^2 |Q_\theta(\theta)| \leq \frac{\gamma}{\theta^p} \left( \frac{\varepsilon \omega}{2j_*} \right)^p |Q_\theta(\theta)|.
\]
In the second last line we used the fact that the parameter $\varepsilon$ is already fixed in dependence on the data. Hence, the above energy estimate yields
\[
\iint_{Q_\varepsilon(\theta)} |D(u - k_j)|^p \, dx \, dt \leq \frac{\gamma}{\theta^p} \left( \frac{\varepsilon \omega}{2^j} \right)^p |Q_\varepsilon(\theta)|.
\]
Next, we apply [5, Chapter I, Lemma 2.2] slicewise to $u(\cdot, t)$ for $t \in (-\theta \rho^j, 0]$ over the cube $K_\varepsilon$, for levels $k_{j+1} > k_j$ and take into account the measure theoretical information from (6.9), i.e. that
\[
\left\{ u(\cdot, t) \leq \mu^+ - \varepsilon \omega \right\} \cap K_\varepsilon \geq \frac{1}{4} |K_\varepsilon| \quad \text{for all } t \in (-\theta \rho^j, 0].
\]
This leads to
\[
(k_{j+1} - k_j) \left\{ u(\cdot, t) > k_{j+1} \right\} \cap K_\varepsilon
\leq 
\frac{\gamma \rho^{N+1}}{N} \left[ \iint_{Q_\varepsilon(\theta)} |D(u - k_j)| \, dx \, dt \right] \leq 
\frac{\gamma \rho^{N+1}}{N} \nu \left[ \iint_{Q_\varepsilon(\theta)} |D(u - k_j)|^p \, dx \, dt \right] \left\{ k_j < u(\cdot, t) < k_{j+1} \right\} \cap K_\varepsilon \leq 
\frac{\gamma \rho^{N+1}}{N} \nu \left[ \iint_{Q_\varepsilon(\theta)} |D(u - k_j)|^p \, dx \, dt \right] \left| A_j(t) - |A_{j+1}(t)| \right|^{1-\frac{1}{p}}.
\]
In the last line we used the abbreviation $A_j(t) := \left\{ u(\cdot, t) < k_j \right\} \cap K_\varepsilon$. We now integrate the preceding inequality with respect to $t$ over $(-\theta \rho^j, 0]$ and apply Hölder’s inequality slice-wise. Setting $A_j = \{ u < k_j \} \cap Q_\varepsilon(\theta)$ this leads to the measure estimate
\[
\frac{\varepsilon \omega}{2^j+1} |A_{j+1}| \leq \frac{\gamma \rho^{N+1}}{N} \nu \left[ \iint_{Q_\varepsilon(\theta)} |D(u - k_j)|^p \, dx \, dt \right] \left| A_j - |A_{j+1}| \right|^{1-\frac{1}{p}} \leq \left( \frac{\varepsilon \omega}{2^j} \right)^\frac{1}{p} |Q_\varepsilon(\theta)| \left| A_j - |A_{j+1}| \right|^{1-\frac{1}{p}}.
\]
Now take the power $\frac{p}{p-1}$ on both sides. This gives
\[
|A_{j+1}|^{\frac{p}{p-1}} \leq \gamma |Q_\varepsilon(\theta)|^{\frac{p}{p-1}} \left| A_j - |A_{j+1}| \right|.
\]
To finish the proof, we proceed exactly as in the proof of Lemma 4.2. We add the inequalities with respect to $j$ from 0 to $j_*$ - 1 and obtain
\[
j_* |A_{j_*}|^{\frac{p}{p-1}} \leq \gamma |Q_\varepsilon(\theta)|^{\frac{p}{p-1}},
\]
from which we deduce the claim, i.e. that
\[
|A_{j_*}| \leq \frac{\gamma}{j_*} |Q_\varepsilon(\theta)|.
\]
This completes the proof.

6.3.3. A DeGiorgi-type Lemma. As before, $\varepsilon \in (0, 1)$ denotes the constant from Lemma 6.2 depending only on the data.

Lemma 6.4. Suppose that the assumptions (6.7) and (6.8) hold true. Then, there exists a constant $\nu_1 \in (0, 1)$ depending only on the data, such that if for some $j_* > 1$, the measure bound
\[
\left| \{ \mu^+ - u \leq \frac{\varepsilon \omega}{2^j} \right\} \cap Q_\varepsilon(\theta) \right| \leq \nu_1 |Q_\varepsilon(\theta)|,
\]
holds true, where $\theta = 2^{j_*(p-2)}$, then
\[
\mu^+ - u \geq \frac{\varepsilon \omega}{2^{j_*+1}} \quad \text{a.e. in } Q_{\frac{1}{2} \varepsilon}(\theta).
\]
Proof. Let $M := 2^{-j_s} \varepsilon \omega$ and

$$k_n := \mu^+ - \frac{M}{2} \frac{M}{2^n + 1}.$$  

Then, define $\tilde{k}_n, \tilde{q}_n, \tilde{K}_n, \tilde{Q}_n, \nu_n$ and $\tilde{Q}_n$ as in the proof of Lemma 4.3, i.e. as in (4.5). Introduce the cutoff functions $\zeta$ vanishing on the parabolic boundary of $Q_n$ and equal to identity in $\tilde{Q}_n$, such that

$$|D\zeta| \leq \frac{\gamma 2^n}{\nu} \quad \text{and} \quad |\zeta_t| \leq \frac{\gamma 2^n}{\nu \theta^{p/2}}.$$

Thus, using again the condition $\xi \omega \leq \mu^+ \leq 2\omega$ to estimate the terms in the energy inequality (see the proof of Lemma 4.3) we obtain

$$\omega^{p-2} \sup_{-\theta \leq t < t_0} \int_{K_n} (u - \tilde{k}_n)^2 \, dx + \int_{\tilde{Q}_n} |D(u - \tilde{k}_n)|^p \, dx \, dt \leq \gamma \frac{2^n}{\nu} M^p \left( 1 + \frac{\omega^{p-2}}{\theta^{p-2}} \right) |A_n| = \gamma \frac{2^n}{\nu} M^p \left( 1 + \varepsilon^{2-p} \right) |A_n|,$$

where we abbreviated

$$A_n = \{ u > k_n \} \cap Q_n.$$

The constant $\gamma$ depends on the data and $\xi$. The latter dependence enters due to the estimate from below of the sup-term in the energy inequality. Note that $\xi$ is already determined in dependence on the data. Now setting $\zeta$ to be a cutoff function which vanishes on the parabolic boundary of $\tilde{Q}_n$ and equals identity in $Q_{n+1}$, an application of the Sobolev imbedding [5, Chapter I, Proposition 3.1] and the preceding estimate imply that

$$\left( \frac{M}{2^n + 1} \right)^p |A_{n+1}| \leq \int_{\tilde{Q}_n} (u - \tilde{k}_n)^p \, dx \, dt \leq \gamma \left( \frac{2^n}{\nu} M^p \right) \left( 1 + \frac{\omega^{p-2}}{\theta^{p-2}} \right) |A_n|^{1 + \frac{p}{N-2}}.$$

Setting $Y_n = |A_n|/|Q_n|$, we arrive at

$$Y_{n+1} \leq \gamma b^n \left( \frac{\omega^{p-2}}{\theta^{p-2}} \right) \left( 1 + \varepsilon^{2-p} \right) \frac{N+p}{N-2} Y_n^{1 + \frac{p}{N-2}}.$$

where $b = 4$ and $\gamma$ only depends on the data. Hence by [5, Chapter I, Lemma 4.1], there exists a constant $\nu_1 \in (0, 1)$ depending only on the data, such that $Y_n \to 0$ if we require the smallness condition $Y_0 \leq \nu_1$. \hfill \Box

We are now ready to conclude the reduction of oscillation near the supremum in the final case where (6.7) and (6.8) are satisfied. By $\varepsilon \in (0, 1), \gamma > 0$ and $\nu_1 \in (0, 1)$ we denote the corresponding constants from Lemmas 6.2, 6.3 and 6.4. Then, we choose a positive integer $j_s$ in such a way that

$$\frac{\gamma}{2^n} \leq \nu_1.$$

Applying in turn Lemmas 6.2, 6.3 and 6.4 then yields that

$$\mu^+ - u \geq \frac{\varepsilon \omega}{2^{j_s + 1}} \quad \text{a.e. in} \ \tilde{Q}_1.$$
where \( \tilde{Q}_1 \) is defined in (6.3). Here we used the fact that \( Q \subset \tilde{Q}_1 \), which implies a reduction of oscillation, i.e. we have 
\[
\text{ess osc } u \leq \left( 1 - \frac{\varepsilon}{2^{i+1}} \right) \omega.
\]

6.4. Reduction of Oscillation Near Zero Concluded. Let us first define quantities 
\[
\lambda = \min \left\{ \frac{1}{4}, \frac{1}{2}, \frac{1}{2A} \right\}, \quad \eta = \min \left\{ \eta_1, \eta_2, \frac{\varepsilon}{2^{i+1}} \right\}.
\]

Now we may proceed by induction. Suppose up to \( i = 1, 2, \cdots j - 1 \), we have built 
\[
\begin{aligned}
\varrho_i &= \lambda \varrho_{i-1}, \quad \omega_i = (1 - \eta)\omega_{i-1}, \quad Q_i = K_{\varrho_i} \times (-\varrho_i^p, 0], \\
\mu_i^- &= \text{ess sup}_{Q_i} u, \quad \mu_i^+ = \text{ess inf}_{Q_i} u, \quad \text{ess osc } u \leq \omega_i.
\end{aligned}
\]

For all the indices \( i = 1, 2, \cdots j - 1 \), we always assume the first case in (6.1), i.e. 
\[
\mu_i^- \leq \xi \omega_i \quad \text{and} \quad \mu_i^+ \geq \xi \omega_i,
\]
where \( \xi \) is determined in Proposition 4.1. In this way the argument in the previous sections can be repeated, and we have for all \( i = 1, 2, \cdots j \), 
\[
\text{ess osc } u \leq (1 - \eta)\omega_{i-1} = \omega_i.
\]

Consequently, iterating this recursive inequality we obtain for all \( i = 1, 2, \cdots j \), 
\[
\text{ess osc } u \leq (1 - \eta)^{i-1} \omega_1.
\]

6.5. Reduction of Oscillation Away From Zero. In this section, let us suppose \( j \) is the first index satisfying the second case in (6.1), i.e. 
\[
\text{either } \mu_j^- > \xi \omega_j \quad \text{or} \quad \mu_j^+ < -\xi \omega_j.
\]

Let us treat for instance \( \mu_j^- > \xi \omega_j \), for the other case is analogous. We observe that since \( j \) is the first index for this to happen, one should have \( \mu_{j-1}^- \leq \xi \omega_{j-1} \). Moreover, one estimates 
\[
\mu_j^- \leq \mu_{j-1}^- + \omega_{j-1} - \omega_j \leq (1 + \xi)\omega_{j-1} - \omega_j = \frac{\xi + \eta}{1 - \eta} \omega_j.
\]
As a result, we have 
\[
\xi \omega_j \leq \mu_j^- \leq \frac{\xi + \eta}{1 - \eta} \omega_j.
\]

The condition (6.11) indicates that starting from \( j \) the equation (1.1) resembles the parabolic \( p \)-Laplacian type equation in \( Q_j \). Like when \( 1 < p < 2 \) (cf. Section 5.3), we drop the suffix \( j \) from our notation for simplicity, and introduce \( v \) defined in \( Q = K_\varrho \times (-\varrho^p, 0] \). As in Section 5.3, \( v \) satisfies (1.1) – (1.2) with \( A(x, t, u, Du) \) replaced by some properly defined \( \tilde{A}(x, t, v, Dv) \), which is subject to the structural conditions (1.2). Moreover, 
\[
1 \leq v \leq \frac{1 + \xi}{\xi} \quad \text{a.e. in } Q.
\]

As in Section 5.3, it turns out to be more convenient to consider the equation satisfied by \( w := v^{\theta-1} \), i.e. 
\[
\partial_t w - \text{div} \tilde{A}(x, t, w, Dw) = 0 \quad \text{weakly in } Q,
\]
where similarly as in (5.3) we define the vector-field \( \tilde{A} \) by 
\[
\tilde{A}(x, t, y, \zeta) = \tilde{A}(x, t, y^{\varrho_p}, \frac{1}{p-1} y^{\varrho_p - 1} \zeta),
\]
for a.e. \((x, t) \in Q\), any \(y \in \mathbb{R}\) and any \(\zeta \in \mathbb{R}^N\). This time \(\tilde{y}\) is defined by

\[
\tilde{y} \overset{\text{def}}{=} \min \left\{ \max \left\{ y, \frac{1}{2} \right\}, 2 \left( \frac{1 + \xi}{\xi} \right)^{p-1} \right\}.
\]

It is easy to see that \(w\) is in the same kind of functional space (1.8) as \(u\) and \(v\) due to (6.12). Employing (6.12) again, we verify exactly as in Section 5.3 that there exist absolute positive constants \(C_0 = \gamma_0(p, \xi)C_o\) and \(C_1 = \gamma_1(p, \xi)C_1\), such that

\[
\tilde{A}(x, t, y, \zeta) \cdot \zeta \geq C_0 |\zeta|^p \quad \text{and} \quad |\tilde{A}(x, t, y, \zeta)| \leq C_1 |\zeta|^{p-1},
\]

for a.e. \((x, t) \in Q\), any \(y \in \mathbb{R}\), and any \(\zeta \in \mathbb{R}^N\). Note that \(\xi\) is already fixed in dependence of the data. This shows that \(w\) is a local weak solution to the parabolic \(p\)-Laplacian type equation in \(Q\). We tend to use Proposition 5.1. To order for that, we first check the condition (5.8) is satisfied. Indeed, recalling \(v = u/\mu^-\), \(w = v^{p-1}\) and \(\omega = \text{ess osc}_Q u\), we first use (6.12) and the mean value theorem to obtain

\[
(p-1) \text{ess osc}_Q v \leq \tilde{\omega} = \text{ess osc}_Q w \leq (p-1) \left( \frac{1 + \xi}{\xi} \right)^{p-2} \text{ess osc}_Q v.
\]

Since \(\text{ess osc}_Q v = \omega/\mu^-\), this amounts to

\[
(p-1) \frac{\omega}{\mu^-} \leq \tilde{\omega} \leq (p-1) \left( \frac{1 + \xi}{\xi} \right)^{p-2} \omega.
\]

Then by (6.11), we have

\[
c \overset{\text{def}}{=} (p-1) \frac{1-\eta}{\xi + \eta} \leq \tilde{\omega} \leq (p-1) \left( \frac{1 + \xi}{\xi} \right)^{p-2} \overset{\text{def}}{=} C.
\]

Thus we only need to take \(\sigma \leq c^{p-2}\), such that \(Q_{\sigma^\theta}(\theta) \subset Q_{\theta}\) and the condition (5.8) in Proposition 5.1 is fulfilled. As a result, the conclusion of Proposition 5.1 is obtained. Moreover, the above upper bound of \(\tilde{\omega}\) actually allows us to obtain the set inclusion

\[
Q_{\tilde{r}}(\theta_0) \subset Q_{\tilde{r}}(\theta) \quad \text{where} \ \theta_0 = C^{2-p}.
\]

Using this set inclusion and rephrasing the oscillation decay in Proposition 5.1 in terms of \(u\), we have for all \(0 < \tilde{r} < \theta\), that the oscillation decay estimate

\[
\text{ess osc}_{Q_{\tilde{r}}(\theta_0)} u \leq \gamma \omega \left( \frac{r}{\theta_j} \right)^{\beta_1}
\]

holds true. Now we revert to using the suffix \(j\). The above oscillation decay then reads

\[
(6.13) \quad \text{ess osc}_{Q_{\tilde{r}}(\theta_0)} u \leq \gamma \omega_j \left( \frac{r}{\theta_j} \right)^{\beta_1}
\]

whenever \(0 < \tilde{r} < \theta_j\). Combining (6.10) and (6.13), we arrive at the desired conclusion, i.e., for all \(0 < \tilde{r} < \theta\) we have

\[
\text{ess osc}_{Q_{\tilde{r}}(\theta_0)} u \leq \gamma \omega \left( \frac{r}{\theta} \right)^{\beta} \quad \text{where} \ \beta = \min\{\beta_\alpha, \beta_1\}.
\]

A proper rescaling gives the oscillation decay in Remark 1.1 and completes the proof of Theorem 1.1.

7. Proof of Boundary Regularity

The proofs of Theorems 1.2 - 1.4 present many similarities with the interior case. However, contrary to the interior case we do not need to distinguish between the cases \(p < 2\) and \(p > 2\). All the technical tools needed near the parabolic boundary have been presented previously. Therefore, we will give sketchy proofs only, while keeping reference to the tools and strategies used in the interior and highlighting the main modifications.
7.1. **Proof of Theorem 1.2.** Consider the cylinder of forward type \( Q_o = K_o(x_o) \times (0, \theta_p^o] \subset E_T \) whose vertex \((x_o, 0)\) is attached to the bottom of \( E_T \). We may assume \( x_o = 0 \) and set
\[
\mu^+ = \text{ess sup}_{Q_o} u, \quad \mu^- = \text{ess inf}_{Q_o} u, \quad \omega = \mu^+ - \mu^-.
\]
Like in the proof of interior regularity, there are two main cases to consider, namely
\[
\begin{align*}
\text{when } u & \text{ is near zero: } \mu^- \leq \omega \text{ and } \mu^+ \geq -\omega; \\
\text{when } u & \text{ is away from zero: } \mu^- > \omega \text{ or } \mu^+ < -\omega.
\end{align*}
\]
Let us suppose the first case holds, which implies \( |\mu^\pm| \leq 2\omega \). The proof continues with a comparison to the initial datum \( u_o \), i.e., we may assume
\[
\text{either } \mu^+ - \frac{4}{3} \omega > \sup_{K_o} u_o \text{ or } \mu^- + \frac{4}{3} \omega < \inf_{K_o} u_o.
\]
For otherwise, we would arrive at
\[
\text{ess osc } u \leq 2 \text{ ess osc } u_o.
\]
Let us assume for instance the second inequality with \( \mu^- \) holds and work with \( u \) as a supersolution. Therefore, we let \( \theta \in (0, 1) \) to be chosen later and define \( \tilde{k}_o, \tilde{Q}_o, K_o \), and \( \tilde{K}_o \), as in the proof of Lemma 4.3 according to (4.5), with \( M \) replaced by \( \frac{1}{4} \omega \). The only difference is that now the cylinders \( Q_o \) and \( \tilde{Q}_o \) are of forward type whose vertices are attached to the origin, i.e., \( Q_o = K_o \times (0, \theta_p^o] \) and \( \tilde{Q}_o = K_o \times (0, \theta_p^o] \). With these choices, we may apply the energy estimates in Proposition 3.2 within \( Q_o \), since the levels \( k_o \) are admissible according to (3.4), i.e.,
\[
k_o \leq \mu^- + \frac{4}{3} \omega < \inf_{K_o} u_o.
\]
Using \( |\mu^-| \leq 2\omega \), a similar analysis as in the proof of Lemma 4.3 leads us to the analogue of (4.6), i.e. to the energy estimate
\[
\frac{\omega^{p-2}}{2p(n-\beta)} \text{ess sup}_{0<\tau<\theta_p^o} \int_{\tilde{K}_o} (u - \tilde{k}_o)^2 \, dx + \int_{Q_o} |D(u - \tilde{k}_o)|^p \, dx dt \leq \frac{\gamma 2^{pn} \theta_p^o}{\theta_p^o} \omega^p |A_o|,
\]
where we have abbreviated
\[
A_n = \{ u > k_o \} \cap Q_o.
\]
Note that in (4.6) we only have to replace \( M \) by \( \frac{1}{4} \omega \). Now we are in a situation similar to Lemma 6.1. More precisely, (6.5) holds true with \( \eta_1 = 1 \) and – on the right-hand side – \( 2^{pn} \) replaced by \( 4^{pn} \). Now, applying the Sobolev imbedding as in the proof of Lemma 6.1 and rewriting the resulting estimate in terms of \( Y_n = |A_n|/|Q_o| \), we arrive at
\[
Y_{n+1} \leq \gamma b^\alpha \theta_p^o \omega^p Y_n^{1 + \frac{\beta}{p} - \frac{\gamma}{2}}.
\]
where \( \gamma \) and \( b \) are positive constants depending only on the data. Hence by [5, Chapter I, Lemma 4.1], there exists a constant \( \nu_o \in (0, 1) \) depending only on the data, such that \( Y_n \to 0 \) as \( n \to \infty \) if we require that
\[
Y_o \leq \frac{\nu_o}{\theta_p^o}.
\]
Upon choosing \( \theta = \nu_o \), the above line is automatically satisfied and as a result we obtain
\[
u \geq \mu^- + \frac{4}{3} \omega \text{ a.e. in } \tilde{Q}_1 \overset{\text{def}}{=} K_{\nu_o}^1 \times (0, \theta(\frac{1}{2} \theta)^p] \).
\]
This in turn yields a reduction of oscillation of the form
\[
\text{ess osc } u \leq \frac{7}{8} \omega.
\]
we use the and introduce $p$ will satisfy the parabolic...

\[ \lambda(Q, t) \]

Consequently, iterating the above recursive inequality, we obtain for all $p > -\mu$ and $\sigma$ concerning the regularity of solutions to the parabolic $p$-Laplacian type equation (7.1). For all the indices $i$ from our notation temporarily, $w_o := (u_o / \mu^-)$, $w := v^{\mu^-}$ in $Q := K_\delta \times (0, \vartheta]$. In this way, the function $w$ will satisfy the parabolic $p$-Laplacian type equation (5.6) – (5.7). Moreover, it attains the initial datum $w_o := (u_o / \mu^-)^{\rho - 1}$ in the sense of $L^2(K_\delta)$. Next, we state in the following proposition concerning the regularity of solutions to the parabolic $p$-Laplacian type equation up to the initial time (cf. [5]).

**Proposition 7.1.** Let $p > 1$ and $\sigma$ in $(0, 1)$. Suppose $w$ is a bounded, local, weak solution to (5.6) in $Q := Q_\delta$ such that the structure conditions (5.7) are in force, and such that $w(\cdot, t) \to w_o$ as $t \downarrow 0$ in the sense of $L^2(K_\delta)$. Assume $w_o$ is continuous in $K_\delta$ with modulus of continuity $\omega_{w_o}(\cdot)$. Let

\[ \tilde{\omega} = \text{ess osc} w_{Q, \delta} \quad \text{and} \quad \theta = \tilde{\omega}^{2 - p}. \]

Then, there exist constants $\beta_1$ in $(0, 1)$ and $\gamma > 1$ depending only on the data $N, p, C_o, \tilde{C}_1$ and $\sigma$ (but independent of $w$), such that there holds: whenever we have

\[ \text{ess osc} w_{Q, \delta}(\theta) \leq \omega_{w_o}(\cdot), \]

then the oscillation decay estimate

\[ \text{ess osc} w_{Q, \delta}(\theta) \leq \gamma \omega_{w_o}(\cdot) \]
holds true for all $0 < r < g$. Here, we use the notation $Q_r(\theta) = B_r \times [0, \theta r^p]$.

As in Sections 5.3 and 6.5 (distinguishing the cases $1 < p < 2$ and $p > 2$), one quickly checks that there exist absolute constants $c, \tilde{c} > 0$, such that $c \leq \tilde{\omega} \leq \tilde{C}$ and the condition in Proposition 7.1 is fulfilled for some proper $\sigma$. Moreover, the lower/upper bounds of $\tilde{\omega}$ actually allow us to obtain the set inclusion

$$Q_r(\theta_o) \subset Q_r(\theta) \quad \text{where} \quad \theta_o = \min \{c^2-p, \tilde{C}^{-2}p\}.$$ 

Using this set inclusion and rephrasing the oscillation decay of Proposition 7.1 in terms of $u$, we have

$$\text{ess osc}_{Q_r(\theta_o)} u \leq \gamma \omega \left( \frac{r}{\theta} \right)^{\beta_1} + \gamma \omega_{\theta_o}(g),$$

whenever $0 < r < g$; here we argue similarly to the proof of (6.13). Now we revert to using the suffix $j$. The above oscillation estimate then reads as

$$\text{ess osc}_{Q_r(\theta_o)} u \leq \gamma \omega_j \left( \frac{r}{\theta_j} \right)^{\beta_1} + \gamma \omega_{\theta_o}(g_j)$$

for all $0 < r < \theta_j$. Combining the above two cases, we arrive at the desired conclusion, i.e., for all $0 < r < g$, there holds

$$\text{ess osc}_{Q_r(\theta_o)} u \leq \gamma \omega_{\theta_o}(g) \quad \text{where} \quad \beta = \min\{\beta_o, \beta_1\}.$$ 

Observe that we may replace $g$ by any $\bar{g} \in (r, g)$. In particular, we may set $\bar{g} = \sqrt{r g}$. In this way, we end up with

$$\text{ess osc}_{Q_r(\theta_o)} u \leq \gamma \omega \left( \frac{r}{\sqrt{r g}} \right)^2 + \gamma \omega_{\theta_o}(\sqrt{r g}).$$

7.2. Proof of Theorem 1.3. For $A \geq 1$ to be determined, consider the cylinder $Q_o = K_o(x_o) \times (t_o - A \rho^g, t_o)$ whose vertex $(x_o, t_o)$ is attached to $S_T$. The number $g$ is so small that $t_o - A \rho^g > 0$ and $g < g_o$, where $g_o$ is the constant from the property of positive geometric density (1.5). We may also assume $(x_o, t_o) = (0, 0)$ and set

$$\mu^+ = \text{ess sup}_{Q_o \cap E_T} u, \quad \mu^- = \text{ess inf}_{Q_o \cap E_T} u, \quad \omega = \mu^+ - \mu^-.$$ 

Like in the proof of interior regularity, there are two main cases to consider, namely

$$\text{(7.3)} \left\{ \begin{array}{ll} \text{when} \ u \text{ is near zero:} & \mu^- \leq \xi \omega \text{ and } \mu^+ \geq -\xi \omega; \\ \text{when} \ u \text{ is away from zero:} & \mu^- > \xi \omega \text{ or } \mu^+ < -\xi \omega. \end{array} \right.$$ 

Here $\xi$ is the positive constant fixed in Proposition 4.1 through dependence on the data and $\alpha = \alpha_o$, whereas $\alpha_o$ comes from the property of positive geometric density of $\partial E$.

Let us suppose the first case holds. The proof continues with a comparison to the boundary datum $g$, i.e., we may assume

either $\mu^+ - \frac{1}{2} \omega > \sup_{Q_o \cap S_T} g$ or $\mu^- + \frac{1}{2} \omega < \inf_{Q_o \cap S_T} g$.

For otherwise, we would arrive at

$$\text{ess osc}_{Q_o} u \leq 2 \text{ ess osc}_{Q_o} g.$$ 

Let us suppose for instance the second inequality holds. To proceed, we turn our attention to the energy estimates in Proposition 3.2 for super-solutions. Since $(u - k)_-$ vanishes on $Q_o \cap S_T$ for all $k \leq \mu^- + \frac{1}{2} \omega$ (i.e. $k$ satisfies (3.5) with $Q_{R, S}$ replaced by $Q_o$), we may extend all integrals in the energy estimates to zero outside of $E_T$. The extended $(u - k)_-$ will be denoted by the same symbol and it is still a member of the functional space in (1.8) within $Q_o$. 


The proofs of Lemma 4.2 and Lemma 4.3 can be carried over to the current situation with properly chosen parameters, bearing in mind that we have assumed \( \partial E \) fulfills the property of positive geometric density (1.5), and therefore for any \( k \leq \mu^- + \frac{4}{3} \omega \), we have
\[
\left\{ u(\cdot, t) - k = 0 \right\} \cap K_\varrho(x_0) \geq \alpha_* \left| K_\varrho \right| \quad \text{for all } t \in (-A\varrho^p, 0].
\]
Thus the conclusion of Proposition 4.1 can be reached. As a result the oscillation is reduced in the case \( 1 < p < 2 \) under the condition (7.3)_1 with \( \xi = 1 \) just like in Section 5, whereas this is true for \( p > 2 \) only when \( |\mu^-| < \xi \omega \) with some very small \( \xi \). As a result, one still needs to handle the situation when \( \mu^- < -\xi \omega \) since this is not excluded in (7.3)_1, for \( p > 2 \).

Like in Section 6, there seems to be some technical complication due to the smallness of the parameter \( \xi \) in the case \( p > 2 \). However, the property (7.4) offers considerable simplification. Indeed, we do not need to split the proof into two parts. Our current hypothesis to continue consist of (7.4) and \( -2 \omega < \mu^- < -\xi \omega \) as we have assumed \( \mu^+ \geq -\xi \omega \) in (7.3)_1. They are analogues of (6.7) and (6.9) formulated near the infimum instead of the supremum, with which one can run the machinery employed in Lemma 6.3 and Lemma 6.4. The only difference is that Lemma 6.3 and Lemma 6.4 have been presented in terms of sub-solutions near the supreme, whereas now one needs to reproduce similar arguments in terms of super-solutions near the infimum. Therefore we can reduce the oscillation under the condition (7.3)_1, for \( p > 2 \) as well.

Next, we can proceed by induction just like the interior case until a certain index \( j \), when the second case of (7.3) happens for the first time. Starting from \( j \), the equation will behave like the parabolic \( p \)-Laplacian type equation within \( Q_j \cap E_T \). We may render this point technically just like in the interior case. Accordingly, we need the following result near the lateral boundary (cf. [5]).

**Proposition 7.2.** Let \( p > 1 \) and \( \sigma \in (0, 1) \). Suppose \( w \) is a bounded, local, weak solution to (5.6) – (5.7) in \( Q_\varrho \cap E_T \) and \( w = g \) on \( Q_\varrho \cap S_T \) in the sense defined in Section 1.4.3. Assume \( g \) is continuous on \( S_T \) with modulus of continuity \( \omega_p(\cdot) \). Define
\[
\tilde{\omega} = \text{ess osc}_{Q_\varrho \cap E_T} w \quad \text{and} \quad \theta = \tilde{\omega}^{2-p}.
\]
Then, there exist constants \( \beta_1 \) in \( (0, 1) \) and \( \gamma > 1 \) depending only on the data \( N, p, \bar{C}_0, \bar{C}_1 \) and \( \sigma \) (but independent of \( w \)), such that there holds: if
\[
\text{ess osc}_{Q_\varrho(\cdot) \cap E_T} w \leq \tilde{\omega},
\]
then, the oscillation decay estimate
\[
\text{ess osc}_{Q_\varrho(\cdot) \cap E_T} w \leq \gamma \tilde{\omega}
\]
holds true for all \( 0 < r < \varrho \).

We refrain from further elaboration due to the similarity of the arguments. The proof may be concluded as in the previous section.

**Remark 7.1.** We have omitted actual computations due to similarities with the proof of interior regularity. Nevertheless the conclusions in the interior, such as Proposition 4.1, can be applied directly to the current situation without repeating their proofs, thanks to their emphasis on the distinct roles of sub-solutions and super-solutions, provided we can extend \( u \) properly to the outside of \( E_T \) and generate sub(super)-solutions across the lateral boundary. In this regard, we refer to Lemma A.2 for such extensions.

7.3. **Proof of Theorem 1.4.** First of all, we observe that the proof of interior regularity (Theorem 1.1) consists of two main components, namely, the expansion of positivity (Proposition 4.1) which is based solely on the energy estimates in Proposition 3.1 and a corresponding Hölder estimate for solutions to the parabolic \( p \)-Laplacian type equation.
This observation is also essentially the gist in the proofs of Theorem 1.2–1.3. Similar calculations have to be reproduced mainly due to the variant energy estimates in Proposition 3.2–3.3 that have incorporated either initial data or Dirichlet data. In particular, a key ingredient – the Sobolev imbedding (cf. [5, Chapter I, Proposition 3.1]) – was used in all these situations, assuming the functions \((u - \hat{k})_{\pm}z^p\) vanish on the lateral boundary of the domain of integration. This assumption in turn is fulfilled either by choosing a proper cutoff function \(\zeta\) or by restricting the value of the level \(k\) according to the Dirichlet data as in (3.5) or the initial data as in (3.4).

The main difference in the current situation lies in that such a Sobolev imbedding cannot be used because in general the functions \((u - \hat{k})_{\pm}z^p\) under conditions of Proposition 3.4 do not vanish on \(S_T\). However, a similar Sobolev imbedding (cf. [5, Chapter I, Proposition 3.2]) that does not require functions to vanish on the boundary still holds for the functional space

\[ u \in C(0, T; L^p(E)) \cap L^p(0, T; W^{1,p}(E)). \]

It is remarkable that the imbedding constant now depends on \(N\), the structure of \(\partial E\) and the ratio \(T/|E|^{\frac{1}{p}}\), which is invariant for cylinders of the type \(Q_\varepsilon = K_\varepsilon \times (-\varepsilon, 0]\) and \(Q_\varepsilon \cap E_T\) as well, provided \(\partial E\) is smooth enough.

As an example, we exhibit in the following how to modify the proof of Lemma 4.3 technically. Based on Proposition 3.4 and under the similar notations in Lemma 4.3, with the interior cylinders replaced by their intersection with \(E_T\), the energy estimate (4.6) becomes, assuming \(|\mu| \leq 8M\),

\[
\frac{M^{p-2}}{2^{p(n+3)}} \text{ess sup}_{-\theta \varepsilon, \varepsilon < 0} \int_{\hat{K}_n \cap E} (u - \hat{k}_n)_z^2 \, dx + \int_{\hat{Q}_n \cap E_T} |D(u - \hat{k}_n)|_{-}^p \, dz \, dt \\
\leq \gamma \frac{2^{pn}}{\varepsilon^p} M^p|A_n| + \gamma_n |A_n|,
\]

where we have abbreviated

\(A_n = \{u < \hat{k}_n\} \cap Q_n \cap E_T\).

The term with \(\gamma_n\), comes from the extra term generated by the Neumann datum \(\psi\) and \(\gamma_n\) depends on \(C_2\) through (N). We may assume that the first term on the right dominates the second, for otherwise we would have \(M \leq \left(\frac{2^n}{\varepsilon^p}\right)^\frac{1}{p}\). When \(p > 2\), the first integral on the left may be estimated from below by

\[
\int_{\hat{K}_n \cap E} (u - \hat{k}_n)_z^2 \, dx \geq M^{2-p} \int_{\hat{K}_n \cap E} (u - \hat{k}_n)_z^p \, dx.
\]

When \(1 < p < 2\), we introduce \(\tilde{k}_n = \frac{1}{4}\bar{k}_{n+1} + \frac{1}{4}\bar{k}_n < \hat{k}_n\) and estimate

\[
\int_{\hat{K}_n \cap E} (u - \hat{k}_n)_z^2 \, dx \geq (\tilde{k}_n - \tilde{k}_n)^{2-p} \int_{\hat{K}_n \cap E} (u - \tilde{k}_n)_z^p \, dx = \left(\frac{M}{2^n}\right)^{2-p} \int_{\hat{K}_n \cap E} (u - \tilde{k}_n)_z^p \, dx.
\]

In all cases, the energy estimate becomes

\[
\frac{1}{2^{p+2n}} \text{ess sup}_{-\theta \varepsilon, \varepsilon < 0} \int_{\hat{K}_n \cap E} (u - \hat{k}_n)_z^p \, dx + \int_{\hat{Q}_n \cap E_T} |D(u - \hat{k}_n)|_{-}^p \, dz \, dt \leq \gamma \frac{2^{pn}}{\varepsilon^p} M^p|A_n|.
\]

Then one may proceed to use the previously mentioned Sobolev imbedding (cf. [5, Chapter I, Proposition 3.2]) to establish a recursive inequality of fast geometric convergence for \(Y_n = |A_n|/|Q_n \cap E_T|\) as in Lemma 4.3.

As shown above, whenever we use the energy estimate in Proposition 3.4, the extra term containing \(C_2\) is always absorbed into other terms via the assumption \(M \geq \left(\frac{2^n}{\varepsilon^p}\right)^\frac{1}{p}\), and as such it contributes in the proof of Hölder regularity only by an extra control on the oscillation via \(\omega \leq \gamma_\omega\) with \(\gamma\) depending also on \(C_2\). This remark also holds for the proof of a result like Lemma 4.1.

Finally, we remark that the use of De Giorgi’s isoperimetric inequality (cf. [5, Chapter I, Lemma 2.2]) is permitted for all convex domains. This is not restrictive in our case upon a
local flattening of $\partial E$. In other words, since $\partial E$ is of class $C^1$, the portion of $\partial E$ within $K_R(x_o)$ can be represented in a local coordinate system as part of the hyperplane $x_N = 0$ and $K_R(x_o) \cap E \subset \{ x_N > 0 \}$. Set $K_R^+ := K_R \cap \{ x_N > 0 \}$ and $Q_{R,S} := Q_{R,S} \cap \{ x_N > 0 \}$. Without loss of generality we may assume that the weak formulation in Section 1.4.4 is written in such a coordinate system. Consequently, the energy estimate in Proposition 3.4 is written with $K_R^+$ and $Q_{R,S}$ respectively. Thus the machinery used in Lemma 4.2 can be reproduced with modifications as indicated above.

As usual, another main component of the induction argument will be a corresponding result from the regularity theory for the parabolic $p$-Laplacian type equation (cf. [5]), which we record in the following.

**Proposition 7.3.** Let $p > 1$ and $\partial E$ be of class $C^1$. Suppose $w$ is a bounded, local, weak solution to (5.6) – (5.7) in $Q_{\varrho} \cap E_T$ with the Neumann datum $\psi$ on $Q_{\varrho} \cap S_T$ taken in the sense defined in Section 1.4.4. Assume $\psi$ satisfies (N) and define

$$\hat{\omega} = \text{ess osc} w,$$

If for some constant $\sigma$ in $(0,1)$, there holds

$$\text{ess osc} w \leq \hat{\omega}$$

where $\theta = \hat{\omega}^{2-p}$.

then, there exist constants $\beta_1$ in $(0,1)$ and $\gamma > 1$ depending only on the data $N, p, \tilde{C}_0, \tilde{C}_1, C_2, \sigma$ and the structure of $\partial E$, such that for all $0 < r < \varrho$, we have

$$\text{ess osc} w \leq \gamma \hat{\omega} \left( \frac{r}{\varrho} \right)^{\beta_1}.$$

**Appendix A. More on the Notion of Parabolicity**

**Lemma A.1.** Let $u$ be a local weak sub(super)-solution to (1.1) – (1.2). Then, for any $k \in \mathbb{R}$, the truncation $k \pm (u - k)_\pm$ is a local weak super(sub)-solution to (1.1) – (1.2).

**Proof.** Without loss of generality, let $u$ be a local weak sub-solution to (1.1) – (1.2). We show that $k - (u - k)_+$ is a local weak sub-solutions to (1.1) – (1.2). Write down the mollified equation (3.1) and follow the introduction of $Q_{R,S} \in E_T$, the function $w_h$ and the piecewise smooth functions $\zeta$ and $\psi_\varepsilon$ in the proof of Proposition 3.1. Instead of (3.2), we choose here, for some $\sigma > 0$, the test function

$$Q_{R,S} \ni (x,t) \mapsto \varphi(x,t) = \zeta^p(x,t)\psi_\varepsilon(t)\frac{(u(x,t) - k)_+}{(u(x,t) - k)_+ + \sigma}.$$

Like in the proof of Proposition 3.1, we treat the various terms in (3.1). First of all, we consider the time part. We have

$$\int_{E_T} \partial_t w_h^{p-1}\varphi \, dx \, dt = \int_{Q_{R,S}} \zeta^p\psi_\varepsilon\partial_t w_h^{p-1} \frac{(w_h - k)_+}{(w_h - k)_+ + \sigma} \, dx \, dt$$

$$+ \int_{Q_{R,S}} \zeta^p\psi_\varepsilon \partial_t w_h^{p-1} \left[ \frac{(u - k)_+}{(u - k)_+ + \sigma} - \frac{(w_h - k)_+}{(w_h - k)_+ + \sigma} \right] \, dx \, dt$$

$$\geq \int_{Q_{R,S}} \zeta^p\psi_\varepsilon \partial_t h_+(w_h, \sigma, k) \, dx \, dt$$

$$= - \int_{Q_{R,S}} (\zeta^p\psi_\varepsilon' + \psi_\varepsilon \partial_t \zeta^p) h_+(w_h, \sigma, k) \, dx \, dt,$$

where we have defined

$$h_+(u, \sigma, k) \overset{\text{def}}{=} (p - 1) \left[ k^{p-1} + \int_k^u |s|^{p-2}(s - k)_+ \, ds \right].$$
Note that $\lim_{x,t \to 0} h_\sigma(u(x,t), \sigma, k) = [k + (u - k)_+]$. As in the proof of Proposition 3.1, we have used the fact that the second line in the above estimates has a non-negative contribution, due to (3.3) and the fact that the map

$$s \mapsto \frac{(s^{1+p} - k)_+}{(s^{1+p} - k)_+ + \sigma}$$

is a monotone increasing function. We now send $h \downarrow 0$ and then $\varepsilon \downarrow 0$, as in the proof of Proposition 3.1, to obtain

$$\lim_{\varepsilon \downarrow 0} \left( \lim_{h \downarrow 0} \int_{E_T} \partial_t \varphi_{\varepsilon} \, dx \right) \geq \int_{K_R} \zeta^p(x, t_1) h_\sigma(u(x, t_1), \sigma, k) \, dx - \int_{K_R} \zeta^p(x, t_2) h_\sigma(u(x, t_2), \sigma, k) \, dx$$

$$\quad - \int_{K_R \times (t_1, t_2)} \partial_t \zeta^p h_\sigma(u, \sigma, k) \, dx dt.$$

Next, we consider the diffusion term. To this end, we again send $h \downarrow 0$ and then $\varepsilon \downarrow 0$, and use (1.2)$_1$ to obtain

$$\lim_{\varepsilon \downarrow 0} \left( \lim_{h \downarrow 0} \int_{E_T} [A(x, t, u, Du)]_R \cdot \partial_t \varphi_{\varepsilon} \, dx \right) = \int_{K_R \times (t_1, t_2)} A(x, t, u, Du) \cdot \left[ D\zeta^p \frac{(u - k)_+}{(u - k)_+ + \sigma} + \zeta^p \frac{\sigma D(u - k)_+}{((u - k)_+ + \sigma)^2} \right] \, dx dt$$

$$\geq \int_{K_R \times (t_1, t_2)} A(x, t, u, Du) \cdot D\zeta^p \frac{(u - k)_+}{(u - k)_+ + \sigma} \, dx dt.$$

Combining all above estimates gives

$$\int_{K_R} \zeta^p(x, t) h_\sigma(u(x, t), \sigma, k) \, dx \bigg|_{t_1}^{t_2} - \int_{K_R \times (t_1, t_2)} \partial_t \zeta^p h_\sigma(u, \sigma, k) \, dx dt$$

$$\quad + \int_{K_R \times (t_1, t_2)} A(x, t, u, Du) \cdot D\zeta^p \frac{(u - k)_+}{(u - k)_+ + \sigma} \, dx dt \leq 0.$$

Finally we send $\sigma \downarrow 0$ to finish the proof. \qed

The above Lemma A.1 has an analog near the lateral boundary $S_T$. Suppose $u$ is a sub(super)-solution to (1.4). The cylinder $Q_{R,S} = K_R(x_\sigma) \times (t_\sigma - S, t_\sigma)$ has its vertex $(x_\sigma, t_\sigma)$ attached to $S_T$ and the level $k$ satisfies (3.5). We define the following truncated extension of $u$ in $Q_{R,S}$:

$$u_{k}^\pm \overset{\text{def}}{=} \begin{cases} k \pm (u - k)_\pm & \text{in } Q_{R,S} \cap E_T, \\ k & \text{in } Q_{R,S} \setminus E_T. \end{cases}$$

An extension of $A$ can be defined as

$$\tilde{A}(x, t, u, \zeta) \overset{\text{def}}{=} \begin{cases} A(x, t, u, \zeta) & \text{in } Q_{R,S} \cap E_T, \\ |\zeta|^{p-2} \zeta & \text{in } Q_{R,S} \setminus E_T. \end{cases}$$

In this way, $\tilde{A}$ is a Carathéodory function satisfying (1.2) with $C_0$ and $C_1$ replaced by $\min\{1, C_0\}$ and $\max\{1, C_1\}$ respectively. Furthermore, we have

**Lemma A.2.** Suppose $u$ is a sub(super)-solution to (1.4) with (1.2) and the level $k$ satisfies (3.5). Let $u_{k}^\pm$ be defined as above. Then $u_{k}^\pm$ is a local weak sub(super)-solution to (1.1) with $\tilde{A}$ in $Q_{R,S}$.\hspace{1cm}
Proof. The calculations are similar to the proof of Lemma A.1. One only has to notice that due to our choice of \( k \) satisfying (3.5), the test function (A.1) is still admissible if we extend it to zero in \( Q_{R,S} \setminus E_T \) (cf. [11, Lemma 2.1]). Notice also this extension does not require any smoothness of the lateral boundary \( ST \) \( a \ priori \). In this way, all the subsequent integrals are carried over into the whole \( Q_{R,S} \).

Remark A.1. We point out that Lemma A.2 exhibits an important character of parabolic equations. So-extended sub(super)solutions across the lateral boundary often play a basic role in investigating the boundary regularity of solutions on rough domains. See for instance [10, 11] in this regard.

Appendix B. Harnack’s Inequality

First of all, let us rephrase Proposition 4.1 for non-negative super-solutions. In such a case, \( \mu^- = 0 \) and local boundedness is not needed. In addition, the either-or alternative does not appear.

Proposition B.1. Let \( u \) be a non-negative, local, weak super-solution to (1.1) – (1.2) in \( E_T \). Suppose for some \( (x_0, t_0) \in E_T, M > 0, \alpha \in (0, 1) \) and \( \varphi > 0 \) we have (4.1) and

\[
\left\{ u(\cdot, t_0) \geq M \right\} \cap K_\varphi(x_0) \geq \alpha |K_\varphi |.
\]

There exist constants \( \delta \) and \( \eta \) in \( (0, 1) \) depending only on the data and \( \alpha \), such that

\[
u \geq \eta M \quad a.e. \text{in } K_{2\varphi}(x_0) \times (t_0 + \delta \varphi, t_0 + \delta \varphi^p].
\]

It is worth mentioning that a similar remark as Remark 4.1 also holds under current circumstance. The following Harnack’s inequality has been shown in [12, 18, 27]. However, we give an alternative proof based on Proposition 4.1 and Theorem 1.1.

Theorem B.1. Let \( u \) be non-negative, continuous, local weak solution to (1.1) – (1.2) in \( E_T \). Assume the set inclusion

\[
K_{2\varphi}(x_0) \times (t_0 - (2\varphi)^p, t_0 + (2\varphi)^p] \subseteq E_T.
\]

There exist \( \theta \in (0, 1) \) and \( \gamma > 1 \) depending only on the data, such that

\[
\gamma^{-1} \sup_{K_\varphi(x_0)} u(\cdot, t_0 - \theta \varphi^p) \leq u(x_0, t_0) \leq \gamma \inf_{K_\varphi(x_0)} u(\cdot, t_0 + \theta \varphi^p).
\]

Proof. We only prove the right-hand inequality, as the left-hand one is a direct consequence (cf. [8, Chapter 5, Section 3]). Introduce a new function

\[
v(x, t) \overset{\text{def}}{=} \frac{u(x_0 + \theta x, t_0 + \theta^p t)}{u(x_0, t_0)},
\]

which satisfies the same type of equation as (1.1) – (1.2) in \( K_2 \times (-2^p, 2^p) \). Thus we only need to show that there exist \( \theta \in (0, 1) \) and \( \gamma > 1 \), such that

\[
\inf_{K_1} v(\cdot, \theta) \geq \gamma^{-1}.
\]

To this end, we introduce, for \( \tau \in (0, 1) \), the family of nested cylinders \( \{Q_\tau\} \) and the families of non-negative numbers \( \{M_\tau\} \) and \( \{N_\tau\} \) as follows:

\[
M_\tau = \sup_{Q_\tau} v, \quad N_\tau = (1 - \tau)^{-\sigma},
\]

where \( \sigma > 1 \) is to be chosen. The two functions \( [0, 1] \ni \tau \to M_\tau, N_\tau \) are increasing, and \( M_0 = N_0 = 1 \) since \( v(0, 0) = 1 \). Moreover, \( N_\tau \to \infty \) as \( \tau \to 1 \) whereas \( M_\tau \) is bounded since \( v \) is locally bounded. Therefore the equation \( M_\tau = N_\tau \) has roots and we denote the largest one as \( \tau_* \). By the continuity of \( v \), there exists \( (y, s) \in Q_{\tau_*} \), such that

\[
v(y, s) = M_{\tau_*} = N_{\tau_*} = (1 - \tau_*)^{-\sigma}.
\]
Moreover,\[ (y,s) + Q_R \subset Q_{1+\tau_*} \subset Q_1, \]where \( R \equiv \frac{1 - \tau_*}{2} \).

Therefore by the definition of \( \tau_* \),
\[
\sup_{(y,s)+Q_R} v \leq \sup_{Q_{1+\tau_*}} v \leq 2^{\sigma}(1 - \tau_*)^{-\sigma} \equiv M_*.\]

Now let \( \varepsilon_* \in (0,1) \) and set \( r = \varepsilon_* R \). By Theorem 1.1 (note also Remark 1.1), for all \( r < R \) and for all \( x \in K_r(y) \), we have
\[
v(x,s) - v(y,s) \geq -\gamma M_* \left( \frac{r}{R} \right)^\beta = -\gamma 2^\sigma (1 - \tau_*)^{-\sigma} \left( \frac{r}{R} \right)^\beta \geq -\frac{1}{2} (1 - \tau_*)^{-\sigma},\]

provided we choose \( \varepsilon_* \) so small that
\[
\gamma 2^\sigma \varepsilon_*^\beta \leq \frac{1}{2}.
\]

This in turn gives
\[
v(x,0) \geq \frac{1}{2} (1 - \tau_*)^{-\sigma} \equiv M \quad \text{for all} \ x \in K_r(y).
\]

From this we may start employing Proposition B.1 with \( \alpha = 1 \) to conclude that there exist positive constants \( \eta \) and \( \delta \) as indicated, such that
\[
v \geq \eta M\]
in the cylinder
\[
Q^{(1)}(y) \equiv K_{2r}(y) \times [s + \delta r^p, s + \delta (2r)^p].
\]

Repeating this process, we conclude that for any positive integer \( n \),
\[
v \geq \eta^n M\]
in the cylinder
\[
Q^{(n)}(y) \equiv K_{2^nr}(y) \times [s + \delta (2^n r)^p, s + \delta (2^n r)^p].
\]

We may assume \( \varepsilon_* (1 - \tau_*) \) is a negative, integral power of \( 2 \). Then choose \( n \) such that \( 2^n r = 2 \).

In this way, we calculate
\[
\eta^n M = \frac{1}{2} (1 - \tau_*)^{-\sigma} \eta^n = \frac{1}{2} (2R)^{-\sigma} \eta^n = \frac{1}{2} \left( \frac{2r}{\varepsilon_*} \right)^{-\sigma} \eta^n = 2^{-2^\sigma - 1} \varepsilon_*^\sigma (2^\sigma \eta)^n.
\]

Finally, we may choose \( \sigma \) such that \( 2^\sigma \eta = 1 \). As a result, setting \( \gamma^{-1} = 2^{-2^\sigma - 1} \varepsilon_*^\sigma \), we have
\[
v \geq \gamma^{-1} \quad \text{in} \ Q^{(n)}(y) = K_2(y) \times [s + \frac{1}{2} \delta, s + \delta].
\]

A further application of Proposition B.1 (note also Remark 4.1) gives us the desired conclusion. \( \square \)

REFERENCES

[1] E. Acerbi and N. Fusco, Regularity for minimizers of nonquadratic functionals: the case \( 1 < p < 2 \), J. Math. Anal. Appl., 140(1), (1989), 115–135.

[2] R. Alonso, M. Santillana and C. Dawson, On the diffusive wave approximation of the shallow water equations, European J. Appl. Math., 19(5), (2008), 575–606.

[3] Y.-Z. Chen and E. DiBenedetto, \( \text{H} \ddot{\text{o}} \text{lder estimates of solutions of singular parabolic equations with measurable coefficients} \), Arch. Rational Mech. Anal., 118(3), (1992), 257–271.

[4] E. DiBenedetto, “Real analysis”, Second edition, Birkhäuser/Springer, New York, 2016.

[5] E. DiBenedetto, “Degenerate Parabolic Equations”, Universitext, Springer-Verlag, New York, 1993.

[6] E. DiBenedetto, On the local behaviour of solutions of degenerate parabolic equations with measurable coefficients, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4), 13(3), (1986), 487–535.

[7] E. DiBenedetto, U. Gianazza and V. Vespri, Liouville-type theorems for certain degenerate and singular parabolic equations, C. R. Math. Acad. Sci. Paris, 348(15-16), (2010), 873–877.

[8] E. DiBenedetto, U. Gianazza and V. Vespri, “Harnack’s Inequality for Degenerate and Singular Parabolic Equations”, Springer Monographs in Mathematics, Springer-Verlag, New York, 2012.
[9] K. Feng and F.J. Molz, A 2-d diffusion based, wetland flow model, J. Hydrol., 196, (1997), 230–250.
[10] U. Gianazza and N. Liao, A boundary estimate for degenerate parabolic diffusion equations, Potential Analysis, in press, 19pp. https://doi.org/10.1007/s11118-019-09794-8
[11] U. Gianazza, N. Liao and T. Lukkari, A boundary estimate for singular parabolic diffusion equations, NoDEA Nonlinear Differential Equations Appl., 25(4), (2018), 24pp.
[12] U. Gianazza and V. Vespri, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 1(3), (2006), 271–284.
[13] M. Giaquinta and G. Modica, A remark on the regularity of the minimizers of certain degenerate functionals, Manuscripta Math., 57(1), (1986), 55–99.
[14] T.V. Hromadka, C.E. Berenbrock, J.R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[15] T. V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[16] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[17] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[18] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[19] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[20] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[21] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[22] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[23] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[24] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[25] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[26] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[27] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[28] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.
[29] T.V. Hromadka, C. E. Berenbrock, J. R. Freckleton and G. L. Guymon, Remarks on the regularity of the minimizers of certain degenerate functionals, J. Funct. Anal., 57(1), (1986), 55–99.