Unified approach to $C^{1,\alpha}$ regularity for quasilinear parabolic equations

Karthik Adimurthi$^a$, Agnid Banerjee$^{a,\dagger}$

$^a$Tata Institute of Fundamental Research, Centre for Applicable Mathematics, Bangalore, Karnataka, 560065, India

Abstract

In this paper, we are interested in obtaining a unified approach for $C^{1,\alpha}$ estimates for weak solutions of quasilinear parabolic equations, the prototype example being

$$u_t - \text{div}(|\nabla u|^{p-2}\nabla u) = 0,$$

without having to consider the singular and degenerate cases separately. This is achieved via a new scaling and a delicate adaptation of the covering argument developed by E. DiBenedetto and A. Friedman.

Keywords: quasilinear parabolic equations, Lipschitz estimates, $C^{1,\alpha}$ regularity, unified approach

2020 MSC: 35K59, 35K92, 35B65

Contents

1 Introduction ................................................................. 2
  1.1 Discussion about Lipschitz regularity ................................. 2
  1.2 Discussion about $C^{1,\alpha}$ regularity ............................... 3

2 Preliminaries ................................................................. 3
  2.1 Notation ........................................................................ 5

3 Main Theorems .................................................................. 5

4 Proof of rough Lipschitz bound - Theorem 3.1 ......................... 7
  4.1 Energy type estimate ....................................................... 8
  4.2 Proof of rough Lipschitz bound ........................................ 10

5 Proof of Uniform Lipschitz estimate - Theorem 3.3 .................... 11
  5.1 First iteration .............................................................. 11
  5.2 Second iteration ........................................................ 13

6 Proof of Corollary 3.7 ....................................................... 13

7 Proof of Corollary 3.8 ....................................................... 14

8 Proof of Uniform $C^{1,\alpha}$ estimate - Theorem 3.10 .................. 14
  8.1 Covering argument ....................................................... 16

$^\dagger$Supported by the Department of Atomic Energy, Government of India, under project no. 12-R&D-TFR-5.01-0520

$^{\ddagger}$Supported in part by SERB Matrix grant MTR/2018/000267 and by Department of Atomic Energy, Government of India, under project no. 12-R & D-TFR-5.01-0520

$^*$Corresponding author

Email addresses: karthikaditi@gmail.com and kadimurthi@tifrbng.res.in (Karthik Adimurthi$^a$), agnidban@gmail.com and agnid@tifrbng.res.in (Agnid Banerjee$^{a,\dagger}$)
1. Introduction

In this paper, we study gradient regularity of weak solutions of equations of the form

\[ u_t - \text{div} A(\nabla u) = 0, \quad (1.1) \]

where the nonlinearity is modelled on the p-Laplace operator. Moreover, we assume \( A : \mathbb{R}^N \rightarrow \mathbb{R}^N \) satisfying the following structure conditions for some \( s \geq 0 \):

\[
\begin{align*}
|A(\zeta)| + |A'(\zeta)||\zeta|^2 + s^2 &= C_1(|\zeta|^2 + s^2)^{\frac{p-1}{2}}, \\
\langle A'(\zeta)\eta, \eta \rangle &= C_0(|\zeta|^2 + s^2)^{\frac{p-2}{2}}|\eta|^2,
\end{align*}
\]

(1.2)

where we have denoted \( A'(\zeta) := \frac{dA(\zeta)}{d\zeta} \).

**Remark 1.1.** We can also consider more general structures of the form \( A(x, t, \zeta) \) with appropriate assumptions made regarding the behaviour with respect to \( x, t \). These assumptions are given in [7, Section 1-(ii) of Chapter VIII]. However for the sake of clarity of exposition, we will avoid dealing with such general structures noting that the techniques of this paper directly carry over to those situations with obvious modifications.

Since we prove two main results, one being Lipschitz regularity and the other being \( C^{1,\alpha} \) regularity, we shall discuss them separately in the following subsections. Concerning the \( C^{1,\alpha} \) results in the elliptic case, we refer the reader to the well known papers [6], [13] and [14].

1.1. Discussion about Lipschitz regularity

There has been two approaches to study Lipschitz regularity for quasilinear parabolic equations, one developed in [8] that makes use of Moser's iteration to first show that the solution \( u \in W^{1,q}_{\text{loc}} \) for any \( q \in [1, \infty) \) and then adapting the DeGiorgi iteration techniques to obtain \( u \in W^{1,\infty}_{\text{loc}} \). The other approach was developed in [4], where only the Moser iteration was used to obtain Lipschitz regularity starting from \( u \in W^{1,p}_{\text{loc}} \) solutions. It is important to note that all the Lipschitz estimates have been obtained for the singular and degenerate cases using different techniques.

In this paper, we give the details for obtaining uniform Lipschitz estimates for the prototype equation

\[ u_t - \text{div}(|\nabla u|^{p-2}\nabla u) = 0, \]
since the energy estimates needed for this are available in [7, Chapter VIII] and the calculations are more illustrative, noting that extending it to more general operators can be done in a standard way. Moreover, in this paper, we start with the a priori assumption that \( u \in W^{1,q}_{\text{loc}} \) for any \( q \in [1, \infty) \) and then prove \( u \in W^{1,\infty}_{\text{loc}} \) by adapting the DeGiorgi iteration as in [8]. This is because, we are interested in obtaining a unified quantitative Lipschitz estimate which is useful in many applications, an important example being to study Calderon-Zygmund type estimates.

The novelty here is that we obtain Lipschitz estimates without having to differentiate between singular and degenerate cases. Our approach to proving this result is based on De Giorgi’s approach by suitably adapting the ideas developed in [3] and combining it with [8]. The main idea is to note that when applying Sobolev embedding, there is still some available flexibility than what is used in [8], and it is this extra information that we make use of in order to obtain uniform Lipschitz estimates. We would like to refer the reader to the recent interesting paper [5] where a unified approach to Lipschitz estimate has been developed based on Moser iteration. See also [9] where uniform Lipschitz estimates are obtained for viscosity solutions to the prototypical parabolic \( p \)-Laplace equation by an adaptation of Ishii-Lions approach. We also highlight the fact that \( C^{1,\alpha} \) regularity is obtained in [9] using non-divergence techniques for the prototypical case.

1.2. Discussion about \( C^{1,\alpha} \) regularity

There has been a long history regarding the developed of \( C^{1,\alpha} \) theory for quasilinear parabolic equations and we refer the reader to the detailed chronological development given in [7]. The highly influential method using intrinsic scaling was developed in [8] to prove \( C^{1,\alpha} \) regularity for quasilinear parabolic systems. This required studying the solution on cylinders whose size intrinsically depended on the solution itself. Though it must be noted that their method (see [7, Chapter IX]) actually does not need any intrinsic geometry, even though it is not explicitly written that way. To be consistent with existing notation, we shall also use the term 'intrinsic scaling' to denote the geometry considered in [8]. Subsequently, in [12], building upon all the previous ideas, the authors obtained a clean and elegant way of proving \( C^{1,\alpha} \) Hölder regularity for quasilinear parabolic equations.

Remark 1.2. In all the discussion about \( C^{1,\alpha} \) regularity, we start with the assumption that the solution is a priori Lipschitz continuous and restrict our discussion to this situation. Moreover, by \( C^{1,\alpha} \) regularity, we mean that the spatial derivative \( \nabla_x u \) is Hölder continuous.

It is important to note that the standard technique of proving \( C^{1,\alpha} \) regularity for quasilinear parabolic equations required using intrinsic geometry and studying singular and degenerate cases separately. In this paper, we develop new scaling and suitably adapt the covering argument from [7, Chapter IX] using which we prove \( C^{1,\alpha} \) regularity without having to differentiate the singular and degenerate cases separately.

Acknowledgments

The first author thanks Adi Adimurthi and Ugo Gianazza for many helpful discussions and suggestions.

2. Preliminaries

In this section, we shall collect all the preliminary material needed in subsequent sections. Before we recall some useful results, let us define the notion of solutions considered in this paper. In order to do this, let us first define
Steklov average as follows: let $h \in (0, 2T)$ be any positive number, then we define

$$u_h(\cdot, t) := \begin{cases} \int_t^{t+h} u(\cdot, \tau) \, d\tau & t \in (-T, T-h), \\ 0 & \text{else.} \end{cases}$$

We shall now define the notion of weak solutions to (1.1).

**Definition 2.1** (Weak solution). We say that $u \in C^0(-T, T; L^2_{\text{loc}}(\Omega)) \cap L^p(-T, T; W^{1,p}_{\text{loc}}(\Omega))$ is a weak solution of (1.1) if, for any $\phi \in C_c^\infty(\Omega)$ and any $t \in (-T, T)$, the following holds:

$$\int_{\Omega} \left\{ \frac{d[u]}{dt} \cdot \phi + \langle [A(x,t,\nabla u)]_h, \nabla \phi \rangle \right\} \, dx = 0 \quad \text{for any} \quad 0 < t < T - h.$$

**Definition 2.2** (Function Space). For any $1 < \bar{p} < \infty$ and any $m > 1$, we define the following Banach spaces:

$$V^{m,\bar{p}}(\Omega_T) := L^\infty(-T, T; L^m(\Omega)) \cap L^{\bar{p}}(-T, T; W^{1,\bar{p}}(\Omega)), \quad V_0^{m,\bar{p}}(\Omega_T) := L^\infty(-T, T; L^m(\Omega)) \cap L^{\bar{p}}(-T, T; W^{1,\bar{p}}_{\text{loc}}(\Omega)).$$

These function spaces have the norm

$$\|f\|_{V^{m,\bar{p}}(\Omega_T)} := \sup_{-T < t < T} \|f(\cdot, t)\|_{L^\infty(\Omega)} + \|\nabla f\|_{L^{\bar{p}}(\Omega_T)}.$$

We have the following parabolic Sobolev embedding theorem from [7, Proposition 3.1 from Chapter I].

**Lemma 2.3.** For any $v \in V_0^{2,\bar{p}}(\Omega_T)$, there exists a constant $C = C(N, \bar{p})$ such that

$$\int_{\Omega_T} |v(x,t)|^{\bar{q}} \, dz \leq C \left( \sup_{0 < t < T} \int_{\Omega} |v(x,t)|^2 \, dx \right)^{\frac{\bar{p}}{\bar{q}}} \left( \int_{\Omega_T} |\nabla v(x,t)|^{\bar{p}} \, dz \right),$$

where $\bar{q} := \frac{\bar{p}N + 2}{N}$.

The next result that we need is a parabolic Sobolev embedding, see [7, Corollary 3.1 of Chapter I] for the details.

**Lemma 2.4.** Let $1 < s < \infty$ and $v \in V_0^s(Q)$ in some cylinder $Q = B \times I$, then

$$\|v\|_{L^s(Q)}^s \leq C \|v\|_{L^\infty(\Omega)} \|v\|_{L^s(Q)}^s.$$

Let us recall a form of Poincaré inequality, see [7, Lemma 2.2 of Chapter I] for the details.

**Lemma 2.5.** Let $v \in W^{1,1}(B_\rho) \cap C^0(B_\rho)$ for some $\rho > 0$ and let $k, l \in \mathbb{R}$ with $k < l$. Then there exists a constant $\gamma = \gamma(N, \rho)$, independent of $k, l, v, \rho$, such that

$$(l-k) |B_\rho \cap \{v > l\}| \leq \gamma \rho^{N+1} \int_{B_\rho \cap \{v \leq k\}} |\nabla v| \, dx.$$

Next we recall a well known iteration lemma, see [7, Lemma 4.1 of Chapter I] for the details.

**Lemma 2.6.** Let $\{X_n\}$ for $n = 0, 1, 2, \ldots$ be a sequence of positive numbers, satisfying the recursive inequalities

$$X_{n+1} \leq CB^n X_n^{1+\alpha},$$

where $C, b > 1$ and $\alpha > 0$ are given numbers. If

$$X_0 \leq C^{-\frac{1}{\alpha}}b^{-\frac{\alpha}{\alpha}},$$

then $\{X_n\}$ converges to zero as $n \to \infty$.
Lemma 2.7. Let \( \{Y_n\}_{n=1}^{\infty} \) be a sequence of equibounded positive numbers satisfying the recursive inequalities

\[
Y_n \leq Cb^n Y_{n+1}^{1-\alpha},
\]

where \( C, b > 1 \) and \( \alpha \in (0, 1) \) are given constants, then the following holds:

\[
Y_0 \leq \left( \frac{2C}{b^{1-\alpha}} \right)^{\frac{1}{\alpha}}.
\]

2.1. Notation

We shall clarify the notation that will be used throughout the paper:

(i) We shall denote a point in \( \mathbb{R}^{N+1} \) by \( z = (x, t) \in \mathbb{R}^{N} \times \mathbb{R} \).

(ii) We shall use the notation \( Q_{a,b}(x_0, t_0) \) to denote a parabolic cylinder of the form \( B_a(x_0) \times (t_0 - b, t_0 + b) \).

(iii) Henceforth, we shall fix a cylinder \( Q_0 = B_{R_0} \times (-R_0^2, R_0^2) \) centered at \( (0,0) \) and its scaled version \( 4Q_0 \).

(iv) We shall denote the boundary of \( 4Q_0 \) by

\[
\Gamma = [B_{4R_0} \times \{ t = -(4R_0)^2 \}] \cup [B_{4R_0} \times \{ t = (4R_0)^2 \}] \cup [\partial B_{4R_0} \times (-(4R_0)^2, (4R_0)^2)].
\]

(v) Let \( \rho > 0, \lambda > 1 \) and \( R_0 > 0 \) be fixed numbers, then for a given point \( z_0 = (x_0, t_0) \in \mathbb{R}^{N+1} \), we define the following cylinders:

\[
Q_\rho(x_0, t_0) := B_\rho(x_0) \times (t_0 - \rho^2, t_0 + \rho^2) \quad \text{and} \quad Q_\lambda(x_0, t_0) := B_{\lambda^{-1}\rho}(x_0) \times (t_0 - \lambda^{-p}\rho^2, t_0 + \lambda^{-p}\rho^2).
\]

(vi) Let \( \lambda > 1 \) be given, then for given two points \( z_1 = (x_1, t_1) \in \mathbb{R}^{N+1} \) and \( z_2 = (x_2, t_2) \in \mathbb{R}^{N+1} \), we need the following metrics:

\[
d(z_1, z_2) := \max\{|x_1 - x_2|, |t_1 - t_2|^{1/2}\}, \quad d_\lambda(z_1, z_2) := \max\{\lambda|x_1 - x_2|, \lambda^{p/2}|t_1 - t_2|^{1/2}\},
\]

\[
d(z_1, K) := \inf_{z_2 \in K} d(z_1, z_2), \quad d_\lambda(z_1, K) := \inf_{z_2 \in K} d_\lambda(z_1, z_2).
\]

(vii) Given any exponent \( q \in (1, \infty) \), we shall denote \( q' = \frac{q}{q-1} \) to be it’s conjugate exponent.

(viii) For a given space-time cylinder \( Q = B_R \times (a, b) \), we denote the parabolic boundary of \( Q \) to be the union of the bottom and the lateral boundaries, i.e., \( \partial_p Q = B_r \times \{ t = a \} \cup \partial B_R \times (a, b) \).

3. Main Theorems

The first theorem we prove gives a rough Lipschitz bound.

**Theorem 3.1.** Let \( \frac{2N}{N+2} < p < \infty \) and \( u \) be a local weak solution of (4.1) satisfying \( |\nabla u| \in L^q_{loc} \) for all \( q \in [1, \infty) \), then \( u \) is Lipschitz continuous.

**Remark 3.2.** The proof of Theorem 3.3 (and Corollary 3.7, Corollary 3.8) makes use of the fact that \( |\nabla u| \in L^\infty \) to obtain (5.2) and (5.3). Thus we first prove a rough estimate in Theorem 3.1 which gives that the gradient is bounded and then we use this fact to obtain improved and optimal quantitative estimates.
Theorem 3.3. Let \( \frac{2N}{N+2} < p < \infty \) and \( u \) be a local weak solution of (4.1) satisfying \( |\nabla u| \in L_q^{\infty} \) for all \( q \in (0, \infty) \), then the following quantitative bound holds: For any \( \sigma \in (0,1) \), \( \varepsilon \in (0,1) \) and parabolic cylinder \( Q_{\rho,\theta} = B_{\rho} \times (-\theta, \theta) \), there holds

\[
\sup_{Q_{(\rho,\sigma,\theta)}} |\nabla u| \leq \left( 2\left( \int_{Q_{(\rho,\theta)}} |\nabla u|^{p+\varepsilon} \, dz \right)^{\frac{2}{p+\varepsilon}} \right)^{\frac{1}{2} + \varepsilon \delta} \frac{C_A}{(1-\sigma)^2} \left( 4^\varepsilon \right)^{\frac{2\varepsilon}{p-2\varepsilon}} \wedge 1,
\]

where \( B, \Sigma, A, X \) are the constants defined in Definition 5.1 and \( \alpha, \beta, \gamma \) to be the constants as defined in Definition 4.2.

Remark 3.4. The right hand side of Theorem 3.3 has \( \int_{Q_{(\rho,\theta)}} v^{p+\varepsilon} \, dz \) which holds for any \( \varepsilon \in (0,1) \). It is well known from higher integrability results proved first in [10] (see [2, Theorem 6.1] where a unified proof without having to differentiate the singular and degenerate regimes was given based on the ideas from [1]) that \( v = |\nabla u| \in L^{p+\varepsilon_0} \) for some universal \( \varepsilon_0 \in (0,1) \). Let us choose \( \varepsilon = \varepsilon_0 \), then the higher integrability result gives

\[
\int_{Q_{(\rho,\theta)}} v^{p+\varepsilon_0} \, dz \leq (\Sigma, p, C_1) \left( \int_{Q_{(2\rho,2\theta)}} v^p \, dz \right)^{1+\varepsilon_0 \delta},
\]

where \( \delta = \frac{1}{-\frac{N}{p} + \frac{(N+2)d}{2}} \) and \( \min \left\{ \frac{2}{p}, 1 \right\} > d > \frac{2N}{(N+2)p} \) is some fixed exponent.

Remark 3.5. We note that (3.1) can be further weakened with the use of what is called very weak solutions. This version of higher integrability was established in [11] and a unified approach has been developed more recently in [1]. Thus there exists an \( \varepsilon_0 \) depending only on data such that the following estimate holds

\[
\int_{Q_{(\rho,\theta)}} v^p \, dz \leq (\Sigma, p, C_1) \left( \int_{Q_{(2\rho,2\theta)}} v^{p-\varepsilon_0} \, dz \right)^{1+\varepsilon_0 \delta},
\]

where \( \delta = \frac{1}{-\frac{N}{p} + \frac{(N+2)d}{2} - \varepsilon_0} \) and \( \min \left\{ \frac{2}{p}, 1 \right\} > d > \frac{2N}{(N+2)p} \) is some fixed exponent.

Remark 3.6. In the strictly degenerate regime \( p \geq 2 \) or the singular regime \( p \leq 2 \), such an explicit Lipschitz estimate is given in [7, Chapter VIII], though we were unable to verify the calculations. The main difficulty we encountered was in verifying [7, Lemma 4.2 of Chapter VIII] where specific choices for \( f(v) \) were made, but the choices made did not seem to satisfy the conditions from [7, Corollary 3.1 of Chapter VIII]. In view of this, we instead follow the calculations from [8] (see proof of Theorem 3.1) where this Lipschitz regularity result was first proved. In particular, we obtain the corrected versions of [7, Theorem 5.1 and Theorem 5.2 of Chapter VIII] and their interpolated versions given in [7, Theorem 5.1’ and Theorem 5.2’ of Chapter VIII] which are given below.

The first is the analogue of [7, Theorem 5.1’ from Chapter VIII]:

Corollary 3.7 (Degenerate case). Let \( p \geq 2 \) and \( u \) be a local weak solution of (4.1) satisfying \( v = |\nabla u| \in L_q^{\infty} \) for all \( q \in [1, \infty) \), then \( u \) is Lipschitz continuous. Moreover, for any \( \varepsilon \in (0,2) \), the following estimate is satisfied:

\[
\sup_{Q_{(\rho,\sigma,\theta)}} v \leq \left( 2\left( \int_{Q_{(\rho,\theta)}} v^{p-2+\varepsilon} \, dz \right)^{\frac{2}{p-2+\varepsilon}} \right)^{\frac{1}{2} + \varepsilon \delta} \frac{C_A}{(1-\sigma)^2} \left( 4^\varepsilon \right)^{\frac{2\varepsilon}{p-2\varepsilon}} \wedge 1,
\]

where \( X, \Sigma, A \) and \( B \) are analogously computed constants with the choice \( \alpha = \gamma = 0 \) and \( \beta = p - 1 \).

Now we state the analogue of [7, Theorem 5.2’ from Chapter VIII]:

...
Corollary 3.8 (Singular case). Let \( \frac{2N}{N + 2} < p \leq 2 \) and \( u \) be a weak solution of (4.1) satisfying \( v = |\nabla u| \in L^q_{\text{loc}} \) for all \( q \in [1, \infty) \), then \( u \) is Lipschitz continuous. Moreover, for any \( \epsilon \in (2 - p, 3) \), the following estimate is satisfied:

\[
|\nabla u(z_0) - \nabla u(z_1)| \leq C \mu_0 \left( \frac{d(z_0, z_1)}{R_0} \right)^\alpha,
\]

where \( \mu_0 = \max \{1, \sup_{Q_0} |\nabla u|\} \), \( C = C(N, p, C_0, C_1) \) and \( a = a(N, p, \alpha) \).

Remark 3.9. Instead of choosing \( k \geq 1 \), if we were to equate both the terms in the expression for \( A_n \) from (5.6), we would get the estimate in Corollary 3.7 (resp. Corollary 3.8) with \( \bigwedge 1 \) replaced with \( \bigwedge \left( \frac{\rho^2}{\theta} \right)^{\frac{1}{p-1}} \) (resp. \( \bigwedge \left( \frac{\rho^2}{\theta} \right)^{\frac{1}{p-1}} \)) and \( A \) replaced by the analogous expression that comes with this calculation. This is the version that is given in [7, Theorem 5.1 and Theorem 5.2 of Chapter VIII].

Theorem 3.10. Let \( 1 < p < \infty \) and \( u \) be a weak solution of (8.1) with (8.2) in force. Moreover, assume that \( |\nabla u| \in L^\infty_{\text{loc}} \), then given any cylinder \( Q_0 = B_{R_0} \times (-R_0^2, R_0^2) \), there exists \( \alpha = \alpha(N, p, C_0, C_1) \in (0, 1) \) such that for any \( z_0, z_1 \in Q_0 \), there holds

\[
|\nabla u(z_0) - \nabla u(z_1)| \leq C \mu_0 \left( \frac{d(z_0, z_1)}{R_0} \right)^\alpha,
\]

where \( \mu_0 = \max \{1, \sup_{Q_0} |\nabla u|\} \), \( C = C(N, p, C_0, C_1) \) and \( a = a(N, p, \alpha) \).

Remark 3.11. In the proof of \( C^{1,\alpha} \) regularity, we shall denote the exponent ‘\( \alpha' \) to be a number that depends on \( N, p, C_0, C_1 \) and the Hölder exponent \( \alpha \). This exponent ‘\( \alpha' \) appears on the right hand side of Theorem 3.10 and in the proof, by an abuse of notation, we redefine ‘\( \alpha' \) at every step to be the larger of all the occurrences of ‘\( \alpha' \). This works in our case since \( \mu_0 \geq 1 \).

4. Proof of rough Lipschitz bound - Theorem 3.1

In order to prove uniform Lipschitz estimates, we will study the prototype equation so that we can follow some of the calculations from [7, Chapter VIII] for ease of reading. It must be noted that the result can be extended to more general equations of the form (1.1) satisfying (1.2) (see [7, Section 1-(ii) of Chapter VIII] for more on this) with standard modifications. Let us recall the prototype equation

\[
\frac{\partial u}{\partial t} - \text{div}(|\nabla u|^{p-2} \nabla u) = 0. \tag{4.1}
\]

Let us first recall the well known energy estimate proved in [7, Proposition 3.2 of Chapter VIII].

Lemma 4.1. Let \( u \) be a local, weak solution of (4.1) and let \( f(\cdot) \) be a non-negative, bounded, Lipschitz function on \( \mathbb{R}^+ \). Then there exists a constant \( C = C(N, p) \) such that on the cylinder \( Q_{(p, \theta)}(z_0) = B_p(x_0) \times (t_0 - \theta, t_0) \), we have

\[
\sup_{t_0 - \theta \leq t \leq t_0} \int_{B_p(z_0)} \left( \int_0^r s f(s) \, ds \right) \zeta^2 \, dz \left. \right|_{t_0 - \theta}^t + \int_{Q_{(p, \theta)}(z_0)} v^{p-2} |\nabla u|^2 f(v) \zeta^2 \, dz + \int_{Q_{(p, \theta)}(z_0)} v^{p-1} |\nabla u|^2 f'(v) \zeta^2 \, dz + (p - 2) \int_{Q_{(p, \theta)}(z_0)} v^{p-3} (\nabla v \cdot \nabla u)^2 f'(v) \zeta^2 \, dz \tag{4.2}
\]

\[
\leq C(N, p) \int_{Q_{(p, \theta)}(z_0)} v^p f(v) |\nabla \zeta|^2 \, dz + \int_{Q_{(p, \theta)}(z_0)} \left( \int_0^r s f(s) \, ds \right) \zeta \zeta_t \, dz,
\]

where we have denoted \( v := |\nabla u| \).
Let us first fix some constants:

**Definition 4.2.** Let $\alpha, \beta, \gamma$ be positive constants satisfying the relations

(i) Choose $\gamma$ such that $p - 2 + \gamma > 0$. So since we have $p > \frac{2N}{N + 2}$, we let $\gamma = \frac{4}{N + 2}$.

(ii) Choose $\alpha$ and $\beta$ such that $\alpha \geq \gamma$ and $\beta \geq 1$.

For the proof of Theorem 3.3, we will take $\alpha = \gamma = \frac{4}{N + 2}$ and $\beta = p - 1 + \gamma$ which is admissible since $p - 2 + \gamma > 0$. For the proof of Corollary 3.7 ($p \geq 2$), we take $\alpha = \gamma = 0$ and $\beta = p - 1$ and for the proof of Corollary 3.8 ($p \leq 2$), we take $\alpha = \gamma = 2 - p$ and $\beta = 1$.

4.1. Energy type estimate

Let radii $\rho, \theta$ be given and for some fixed $\sigma \in (0, 1)$, let us define the following:

$$
\alpha_n := \sigma \frac{\rho + \rho_n + \rho_{n+1}}{2}, \quad \rho_n := \sigma \rho + \frac{(1 - \sigma)\rho}{2} \quad \text{and} \quad Q_n := Q(\rho_n, \theta_n).
$$

(4.3)

With the above choices of radii, we note that $Q_0 = Q(\rho, \theta)$ and $Q_\infty = Q(\sigma \rho, \sigma \theta)$. For a fixed $k \in (0, \infty)$ to be eventually chosen, we denote

$$
k_n := k - \frac{k}{2n},
$$

(4.4)

and consider the following cut-off functions:

$$
\begin{cases}
\zeta_n = 1 \text{ on } \tilde{Q}_n, \\
\zeta_n = 0 \text{ on } \partial_{\rho} \tilde{Q}_n
\end{cases}
\quad \text{with} \quad |\nabla \zeta_n| \leq \frac{2^{n+2}}{(1 - \sigma)\rho} \quad \text{and} \quad \left| \frac{d\zeta_n}{dt} \right| \leq \frac{2^{n+2}}{(1 - \sigma)\rho}.
$$

(4.5)

Note that since we start off with the assumption $u \in W^{1, q}_{\text{loc}}$ for any $q < \infty$, the choice of $f(v) = v^\alpha (v - k_{n+1})^{\beta}$ in (4.2) is admissible, where $\alpha, \beta, \gamma, \varepsilon$ as chosen to satisfy Definition 4.2, noting that $\beta \geq 1$ due to Item (ii) (see [7, Corollary 3.1, Chapter VIII] for more on the admissibility of the choice of $f(v)$).

Let us state the lemma that will be proved:

**Lemma 4.3.** With $\alpha, \beta, \gamma$ as in Definition 4.2 and $k, \tilde{Q}_n, \zeta_n$ as given above, we have the following energy estimate:

$$
\sup_{I_n} \int_{\tilde{B}_n} \left( (v - k_{n+1})^{\alpha + \beta + 2 - \gamma} \zeta_n \right)^2 \, dx + 2 \int_{I_n} \left( \frac{k}{2} \right)^{p - 2 + \gamma} \int_{\tilde{Q}_n} \left( \nabla \left( (v - k_{n+1})^{\alpha + \beta + 2 - \gamma} \zeta_n \right) \right)^2 \, dz
\leq (\alpha + \beta + 4 - \gamma) \int_{Q_n} v^{p + \alpha + \beta} \chi_{\{v \geq k_{n+1}\}} |\nabla \zeta_n|^2 \, dz + \int_{Q_n} v^{2 + \alpha + \beta} \chi_{\{v \geq k_{n+1}\}} |(\zeta_n)_{t}| \, dz.
$$

(4.6)

**Proof.** Recalling the choice $f(v) = v^\alpha (v - k_{n+1})^{\beta}$, we see that in order to prove the lemma, we make use of (4.2), thus we estimate each of the terms from (4.2) as follows:

**Estimate for the first term:** Since $\zeta_n$ vanishes on the parabolic boundary of $Q_n$, the sup term appearing on the left hand side of (4.2) is zero, thus we can estimate this term as follows:

$$
\sup_{t_0 - \theta \leq t \leq t_0} \int_{B_n} \left( \int_0^v ss^\alpha (s - (k_{n+1}))^{\beta} \, ds \right) \zeta_n^2 \, dx = \frac{\left( \alpha + \beta + 2 - \gamma \right)}{\alpha + \beta + 2 - \gamma} \sup_{t_0 - \theta \leq t \leq t_0} \int_{B_n} \left( (v - k_{n+1})^{\alpha + \beta + 2 - \gamma} \zeta_n \right)^2 \, dx.
$$

Note that here we required $1 + \alpha \geq \gamma$. 

8
Estimate for the second term: We estimate the second term appearing on the left hand side of (4.2) as follows:
\[
\iint_{Q_n} v^p |\nabla^2 u|^2 v^\alpha (v - k_{n+1})^\beta \zeta^2_n \, dz \geq (k_{n+1})^{p-2+\gamma} \iint_{Q_n} |\nabla^2 u|^2 (v - k_{n+1})^{\alpha+\beta-\gamma} \zeta^2_n \, dz
\]
\[
= \frac{2}{\alpha + \beta + 2 - \gamma} (k_{n+1})^{p-2+\gamma} \iint_{Q_n} \left( \nabla (v - k_{n+1})^{\alpha+\beta-\gamma} \zeta_n \right)^2 \, dz.
\]
Note that here we required \( \alpha \geq \gamma \).

Estimate for the fifth term: The first term appearing on the right hand side of (4.2) is estimated as follows:
\[
\iint_{Q_n} v^p f(v) |\nabla \zeta_n|^2 \, dz \leq \iint_{Q_n} v^{p+\alpha+\beta} \chi_{\{v \geq k_{n+1}\}} |\nabla \zeta_n|^2 \, dz.
\]

Estimate for the sixth term: Analogously, the second term on the right hand side of (4.2) is estimated as follows:
\[
\iint_{Q_n} \left( \int_{0}^{v} ss^\alpha (s - k_{n+1})^\beta \, ds \right) \zeta_n(\zeta_n) \, dz \leq \frac{1}{\alpha + \beta + 2} \iint_{Q_n} v^{2+\alpha+\beta} \chi_{\{v \geq k_{n+1}\}} |(\zeta_n)| \, dz.
\]

In (4.2), let us ignore the third and fourth term appearing on the left hand side and make use of the bounds obtained for the first, second, fifth and sixth terms above to get
\[
\frac{(k_{n+1})^{\gamma}}{\alpha + \beta + 2 - \gamma} \sup_{I_n} \left( \frac{v - k_{n+1}}{\sqrt{\alpha+\beta+2-\gamma}} \right)^2 \, dx + \frac{2(k_{n+1})^{p-2+\gamma}}{\alpha + \beta + 2 - \gamma} \iint_{Q_n} \left( \nabla (v - k_{n+1})^{\alpha+\beta+2-\gamma} \zeta_n \right)^2 \, dz
\]
\[
\leq \iint_{Q_n} v^{p+\alpha+\beta} \chi_{\{v \geq k_{n+1}\}} |\nabla \zeta_n|^2 \, dz + \frac{1}{\alpha + \beta + 2} \iint_{Q_n} v^{2+\alpha+\beta} \chi_{\{v \geq k_{n+1}\}} |(\zeta_n)| \, dz.
\]

From simple chain rule, we see that
\[
\iint_{Q_n} \left| \nabla (v - k_{n+1})^{\alpha+\beta+2-\gamma} \right|^2 \zeta_n^2 \, dz + \iint_{Q_n} (v - k_{n+1})^{\alpha+\beta+2-\gamma} |\nabla \zeta_n|^2 \, dz = \iint_{Q_n} \left| \nabla \left( (v - k_{n+1})^{\alpha+\beta+2-\gamma} \zeta_n \right) \right|^2 \, dz,
\]
thus we estimate the second term appearing on the right hand side of (4.8) as follows:
\[
\iint_{Q_n} (v - k_{n+1})^{\alpha+\beta+2-\gamma} |\nabla \zeta_n|^2 \, dz = \iint_{Q_n} \frac{v^{p+\alpha+\beta}}{v^{p-2+\gamma}} \chi_{\{v \geq k_{n+1}\}} |\nabla \zeta_n|^2 \, dz
\]
\[
\leq \frac{1}{(k_{n+1})^{p-2+\gamma}} \iint_{Q_n} v^{p+\alpha+\beta} \chi_{\{v \geq k_{n+1}\}} |\nabla \zeta_n|^2 \, dz.
\]
Adding \( \frac{2(k_{n+1})^{p-2+\gamma}}{\alpha + \beta + 2 - \gamma} \iint_{Q_n} (v - k_{n+1})^{\alpha+\beta+2-\gamma} |\nabla \zeta_n|^2 \, dz \) to both sides of (4.7) and making use of (4.8) and (4.9), we get
\[
\sup_{I_n} \frac{k_{n+1}^{\gamma}}{\alpha + \beta + 2 - \gamma} \int_{2}^{\infty} \left( (v - k_{n+1})^{\alpha+\beta+2-\gamma} \zeta_n \right)^2 \, dx + \frac{2k_{n+1}^{p-2+\gamma}}{\alpha + \beta + 2 - \gamma} \iint_{Q_n} \left( \nabla \left( (v - k_{n+1})^{\alpha+\beta+2-\gamma} \zeta_n \right) \right)^2 \, dz
\]
\[
\leq \left( 1 + \frac{\alpha + \beta + 2 - \gamma}{\alpha + \beta + 2 - \gamma} \right) \iint_{Q_n} v^{p+\alpha+\beta} \chi_{\{v \geq k_{n+1}\}} |\nabla \zeta_n|^2 \, dz + \frac{1}{\alpha + \beta + 2} \iint_{Q_n} v^{2+\alpha+\beta} \chi_{\{v \geq k_{n+1}\}} |(\zeta_n)| \, dz.
\]

Since \( \alpha + \beta + 2 - \gamma \leq \alpha + \beta + 2 \), we multiply (4.10) with \( \alpha + \beta + 2 - \gamma \) to get the desired estimate.

For any \( q \in (0, \infty) \), let us recall the well known Chebyshev's inequality:
\[
\iint_{Q_n} \chi_{\{v \geq k_{n+1}\}} \, dx \leq \frac{1}{k_{n}^{\gamma}} \iint_{Q_n} (v - k_{n})^q \, dx.
\]

Remark 4.4. Following the calculation from [7, Estimate (7.5) of Chapter V], for any \( \delta > 1 \) and some parabolic cylinder \( Q \), we have
\[
\iint_{Q} (v - k_{n})^\delta \, dz \geq \iint_{Q} (v - k_{n})^\delta \chi_{\{v \geq k_{n+1}\}} \, dz \geq \iint_{Q} (v^\delta - \frac{2n+1-2}{2n+1-1})^\delta \chi_{\{v \geq k_{n+1}\}} \, dz \geq \frac{1}{2n^\delta} \iint_{Q} v^\delta \chi_{\{v \geq k_{n+1}\}} \, dz.
\]
4.2. Proof of rough Lipschitz bound

Let us now prove that $|\nabla u| \in L^\infty$ with the estimate depending on the quantities $N, p, C_0, C_1$ and $\|\nabla u\|_{L^s}$ for some $s \in (0, \infty)$ which is finite by hypothesis.

Let us take any $\alpha, \beta, \gamma$ such that the two conditions in Definition 4.2 are satisfied and define

$$Y_n := \int_{Q_n} (v - k_n)^{\alpha + \beta + 2 - \gamma} \, dz.$$ 

Then applying Sobolev-Poincare inequality from [7, Proposition 3.1 of Chapter I], we get

$$Y_{n+1} \leq \left( \int_{Q_n} \left( \nabla (v - k_{n+1})^{\alpha + \beta + 2 - \gamma} \right)^\frac{2(n+2)}{\alpha + \beta + 2 - \gamma} \, dz \right)^\frac{\alpha + \beta + 2 - \gamma}{n+2} \left( \sup_{t_n} \int_{B_n} \left( (v - k_{n+1})^{\alpha + \beta + 2 - \gamma} \right)^2 \, dz \right)^\frac{n}{n+2},$$

where

$$\begin{align*}
&\int_{Q_n} \left( \nabla (v - k_{n+1})^{\alpha + \beta + 2 - \gamma} \right)^\frac{2(n+2)}{\alpha + \beta + 2 - \gamma} \, dz \\
\leq& \frac{2^{(n+2)}}{\rho^2} \left( \int_{Q_n} v^{\alpha + \beta + 2 - \gamma} \chi_{\{v \geq k_{n+1}\}} \, dz \right) \left( \int_{Q_n} v^{(p + \alpha + \beta)(\alpha + \beta + 2 - \gamma)} \chi_{\{v \geq k_{n+1}\}} \, dz \right)^\frac{\alpha + \beta + 2 - \gamma - A}{\alpha + \beta + 2 - \gamma},
\end{align*}$$

and

$$\begin{align*}
&\int_{Q_n} v^{2 + \alpha + \beta} \chi_{\{v \geq k_{n+1}\}} \, dz \\
\leq& \frac{2^{(n+2)}}{\theta} \left( \int_{Q_n} v^{\alpha + \beta + 2 - \gamma} \chi_{\{v \geq k_{n+1}\}} \, dz \right) \left( \int_{Q_n} v^{(2 + \alpha + \beta)(\alpha + \beta + 2 - \gamma - B)} \chi_{\{v \geq k_{n+1}\}} \, dz \right)^\frac{\alpha + \beta + 2 - \gamma - B}{\alpha + \beta + 2 - \gamma - B},
\end{align*}$$

where $A = \frac{\alpha + \beta + 2 - \gamma}{2}$ and $B = \frac{\alpha + \beta + 2 - \gamma}{2}$. Thus combining (4.14) and (4.15) into (4.13) gives

$$I \leq \frac{2^{(n+2)}}{k^\gamma} \left( \int_{Q_n} v^{\frac{p + \alpha + \beta}{2(p + \alpha + \beta)}} \frac{\partial}{\partial t} Y_n^\frac{1}{2} + \int_{Q_n} v^{\frac{2 + \alpha + \beta}{2(2 + \alpha + \beta)}} \frac{\partial}{\partial t} Y_n^\frac{1}{2} \right).$$

Estimate for $II$: Making use of (4.6) and proceeding analogous to (4.16), we get

$$II \leq \frac{2^{(n+2)}}{k^\gamma} \left( \int_{Q_n} v^{\frac{p + \alpha + \beta}{2(p + \alpha + \beta)}} \frac{\partial}{\partial t} Y_n^\frac{1}{2} + \int_{Q_n} v^{\frac{2 + \alpha + \beta}{2(2 + \alpha + \beta)}} \frac{\partial}{\partial t} Y_n^\frac{1}{2} \right).$$

Estimate for $III$: We can apply (4.11) to get

$$III \leq \frac{2^n}{k^\gamma} \frac{(\alpha + \beta + 2 - \gamma)}{\alpha + \beta + 2 - \gamma} Y_n.$$
Lemma 2.6, we see that
(4.6).
(4.6)

Theorem 3.3, we see that
(4.3), let us set
Y := \left( \frac{B_1}{k^{(N+2)}} \right)^{-(N+2)} D^{-(N+2)}

then from Lemma 2.6, we see that \{Y_n\} converges to zero as \( n \to \infty \). This concludes the proof of Theorem 3.1.

5. Proof of Uniform Lipschitz estimate - Theorem 3.3

Thanks to Theorem 3.1, we see that \(|\nabla u| \in L^\infty_{\text{loc}}\) and thus we can prove the required quantitative estimates. The first step is to iterate the energy estimates by making careful choice of the level sets.

5.1. First Iteration

Let us now take \( \alpha = \gamma \) and \( \beta = p - 1 + \gamma > 1 \) and then prove the Lipschitz regularity. The proof of the Lipschitz regularity requires two iterative steps, the first of which is proved in this subsection. With \( k_n \) as defined in (4.4), let us set

\[
Y_n := \int_{Q_n} (v - k_n)^{p+1+\gamma} \, dz.
\]

Since we want to suitably control \( Y_{n+1} \) in terms of \( Y_n \), we proceed as follows:

\[
Y_{n+1} \leq \int_{Q_n} (v - k_{n+1})^{p+1+\gamma} \, dz
\]

\[
\leq \left( \int_{Q_n} \left( (v - k_{n+1})^{p+1+\gamma} \right)^{2} \, dz \right)^{\frac{1}{2}} \left( \int_{Q_n} \chi_{\{v \geq k_{n+1}\}} \, dz \right)^{\frac{2}{N+2}}
\]

\[
\leq \left( \sup_{I_n} \int_{B_n} (v - k_{n+1})^{\alpha + \beta + 2 - \gamma} \, dz \right)^{\frac{2}{N+2}} \frac{Y_n}{k_n^{p+1+\gamma}}
\]

where to obtain (a), we enlarged the domain of integration and make use of (4.3) and (4.5), to obtain (b), we applied Hölder’s inequality and finally to obtain (c), we made use of Sobolev embedding theorem along with (4.11).

We can estimate the first two terms appearing on the right hand side of (5.1) by making use of Lemma 4.3. Let us thus estimate each of the terms appearing on the right hand side of (4.6) as follows:

**Estimate of first term:** We estimate the first term appearing on the right hand side of (4.6) as follows:

\[
\int_{Q_n} v^{p+\alpha + \beta} \chi_{\{v \geq k_{n+1}\}} |\nabla \zeta_n|^{2} \, dz
\]

\[
\leq \int_{Q_n} v^{p+\gamma + p+1+\gamma - 2} \chi_{\{v \geq k_{n+1}\}} |\nabla \zeta_n|^{2} \, dz
\]

\[
\leq \frac{4^{n+2}}{(1 - \sigma)^2 \rho^2 (k_{n+1})^2} \int_{Q_n} v^{p+1+\gamma} \chi_{\{v \geq k_{n+1}\}} \, dz
\]

**Remark 4.4**

\[
\leq \frac{4^{n+2} \rho (p+1+\gamma) (k_{n+1})^2}{(1 - \sigma)^2 \rho^2} Y_n.
\]

**Estimate of second term:** Analogously, we estimate the second term appearing on the right hand side of (4.6)
as follows:

\[
\int_{Q_n} v^{2+\alpha+\beta} \chi_{\{v \geq k_{n+1}\}}(\zeta_n) \, dz = \int_{Q_n} v^{2+\gamma+p-1+\gamma+p-p} \chi_{\{v \geq k_{n+1}\}}(\zeta_n) \, dz
\]

\[
(4.5)
\]

\[
\leq \frac{2^{n+2}}{(1-\sigma)\theta} \left( \sup_{Q_{n}} v \right)^{p+\gamma} \int_{Q_n} v^{p+1+\gamma} \chi_{\{v \geq k_{n+1}\}} \, dz
\]

\[
(5.3)
\]

Remark 4.4. Since \(v \to \infty\) as \(n \to \infty\), we have

\[
\sup_{Q_{n}} v \leq C_{1} \frac{B_{n}^{n}}{k^{pN+4+2(p+1+\gamma)}} \left( \frac{1}{1-\sigma} \right)^{2} \left( \sup_{Q_{n}} v \right)^{p+\gamma} Y_{n}^{\frac{\gamma}{p+2}+1} A.
\]

Applying Lemma 2.6, we see that \(Y_{n} \to 0\) as \(n \to \infty\) if

\[
Y_{0} = \int_{Q_{\rho,\theta}} v^{p+1+\gamma} \, dz = \frac{C_{1} A}{k^{pN+4+2(p+1+\gamma)}(1-\sigma)^{2}} \left( \sup_{Q_{\rho,\theta}} v \right)^{p+\gamma} \frac{Y_{n}^{\frac{\gamma}{p+2}+1} A}{(N+2)^{2}}.
\]

In particular, if we make the following choice of \(k\)

\[
k = \left( \frac{B_{n}^{n}}{C_{1} A} \right)^{\frac{1}{p+2}} \left( \frac{1}{1-\sigma} \right)^{2} \left( \sup_{Q_{\rho,\theta}} v \right)^{p+\gamma} \frac{Y_{n}^{\frac{\gamma}{p+2}+1} A}{(N+2)^{2}} \wedge 1.
\]

Before we proceed, let us define a few constants and make an important remark:

**Definition 5.1.** Let us define the following exponents required for the proof:

(i) \(B := 2^{p+2+\gamma - \frac{2n+4+2\gamma}{p+2}}\)

(ii) \(\Sigma := \frac{N+2}{pN+4+2(p+1+\gamma)}\)

(iii) \(A := \left( \frac{2^{\gamma}}{p^{2}} + \frac{2^{p+\gamma-2}}{\theta} \right)\)

(iv) \(X := pN+4+2(p+1+\gamma)\).

Remark 5.2. Since \(1 \geq \frac{2n}{2n} \geq \frac{1}{2}\) for any \(n \geq 1\), we have

\[
A_{n} := \left( \frac{1}{p^{2}(k_{n+1})^{\gamma}} \right) \left( \frac{1}{1-\sigma} \right)^{2} \left( \sup_{Q_{n}} v \right)^{p+\gamma} \left( \frac{1}{(k_{n+1})^{\gamma}} \right) \leq \left( \frac{1}{p^{2} k^{\gamma}} + \frac{1}{\theta k^{p+\gamma-2}} \right) = A,
\]

where \(A\) is as defined in **Definition 5.1** and \(k_{n+1}\) is as defined in (4.4) for some \(k \geq 1\).

It is easy to see that if we were to balance both the terms in \(A_{n}\), then we go back to the original estimate arising in the proofs of [7, Theorems 5.1 and 5.2, Chapter VIII], which is where the distinction between the singular and degenerate cases was needed.

Substituting (5.4) and (5.5) along with (5.6) into (5.1) and recalling the definition of constants from **Definition 5.1**, we have the following iterative estimate

\[
Y_{n+1} \leq C_{1} \frac{B_{n}^{n}}{k^{pN+4+2(p+1+\gamma)}} \left( \frac{1}{1-\sigma} \right)^{2} \left( \sup_{Q_{\rho,\theta}} v \right)^{p+\gamma} Y_{n}^{\frac{\gamma}{p+2}+1} A.
\]

Applying Lemma 2.6, we see that \(Y_{n} \to 0\) as \(n \to \infty\) if

\[
Y_{0} = \int_{Q_{\rho,\theta}} v^{p+1+\gamma} \, dz = \frac{C_{1} A}{k^{pN+4+2(p+1+\gamma)}(1-\sigma)^{2}} \left( \sup_{Q_{\rho,\theta}} v \right)^{p+\gamma} \frac{Y_{n}^{\frac{\gamma}{p+2}+1} A}{(N+2)^{2}}.
\]

In particular, if we make the following choice of \(k\)

\[
k = \left( \frac{B_{n}^{n}}{C_{1} A} \right)^{\frac{1}{p+2}} \left( \frac{1}{1-\sigma} \right)^{2} \left( \sup_{Q_{\rho,\theta}} v \right)^{p+\gamma} \frac{Y_{n}^{\frac{\gamma}{p+2}+1} A}{(N+2)^{2}} \wedge 1,
\]

12
then $Y_n \to 0$ as $n \to \infty$. From this, we see that $Y_\infty = 0$ is equivalent to the following estimate:

$$\sup_{Q_{(\sigma, \rho, \theta)}} v \leq k = \left( E^{N+2} \left( \int_{Q_{(\rho, \theta)}} v^{p+1+\gamma} \, dz \right)^{\frac{1}{p+1+\gamma}} C_1 A \frac{1}{(1-\sigma)^2} \left( \sup_{Q_{(\rho, \theta)}} v \right)^{p+\gamma} \right) \wedge 1, \quad (5.7)$$

### 5.2. Second iteration

The first iteration gives a bound of the form (5.7) which does not give the desired Lipschitz bound due the sup term appearing also on the right hand side. In order to overcome this, we now iterate (5.7) a second time in this subsection.

Let us fix some $\varepsilon_0 \in (0, 1)$ (to be eventually chosen later) and rewrite (5.7) as follows:

$$\sup_{Q_{(\rho, \sigma, \theta)}} v \leq \left( E^{N+2} \left( \int_{Q_{(\rho, \theta)}} v^{p+\varepsilon_0} \, dz \right)^{\frac{1}{p+\varepsilon_0}} C_1 A \frac{1}{(1-\sigma)^2} \left( \sup_{Q_{(\rho, \theta)}} v \right)^{p+\varepsilon_0} \right) \wedge 1. \quad (5.8)$$

With $\sigma \in (0, 1)$, consider the family of cylinders $Q_n := Q_{(\rho_n, \theta_n)}$ where

$$\rho_n := \rho + (1-\sigma) \rho \sum_{i=1}^{n} 2^{-i}$$

and

$$\theta_n := \theta + (1-\sigma) \theta \sum_{i=1}^{n} 2^{-i}.$$

Then we have $Q_0 = Q_{(\rho, \sigma, \theta)}$ and $Q_\infty = Q_{(\rho, \theta)}$. If we denote

$$M_n := \sup_{Q_n} v,$$

then (5.8) applied over $Q_{n+1}$ and $Q_n$ can be rewritten for any $\varepsilon \in (0, 1]$ as

$$M_n \leq 2^{2n+\Sigma} M_{n+1}^{\frac{2}{2p+2}} \left( E^{N+2} \left( \int_{Q_{(\rho, \theta)}} v^{p+\varepsilon} \, dz \right)^{\frac{1}{p+\varepsilon}} C_1 A \frac{1}{(1-\sigma)^2} \right) \wedge 1 \quad (5.9)$$

Recalling the constants defined in Definition 5.1, we have $X = (p+\gamma)(N+2) + 2(1+\gamma)$ and $p-2+\gamma > 0$, thus we can rewrite (5.9) as

$$M_n \leq 2^{2n+\Sigma} M_{n+1}^{1-\frac{2\varepsilon}{2p+2}} \left( E^{N+2} \left( \int_{Q_{(\rho, \theta)}} v^{p+\varepsilon} \, dz \right)^{\frac{1}{p+\varepsilon}} C_1 A \frac{1}{(1-\sigma)^2} \right) \wedge 1 \quad (5.10)$$

Thus iterating the above estimate using Lemma 2.7, we get

$$M_0 = \sup_{Q_{(\rho, \sigma, \theta)}} v \leq \left( 2^{2n+\Sigma} M_{n+1}^{1-\frac{2\varepsilon}{2p+2}} \left( E^{N+2} \left( \int_{Q_{(\rho, \theta)}} v^{p+\varepsilon} \, dz \right)^{\frac{1}{p+\varepsilon}} C_1 A \frac{1}{(1-\sigma)^2} \right) \wedge 1 \right) \wedge 1,$$

which gives the desired estimate.

### 6. Proof of Corollary 3.7

In (4.10), let us take $\alpha = \gamma = 0$ and $\beta = p-1$ which are admissible since $p \geq 2$. Then the analogue of (5.7) in the case $p \geq 2$ becomes

$$\sup_{Q_{(\rho, \sigma, \theta)}} v \leq k = \left( E^{N+2} \left( \int_{Q_{(\rho, \theta)}} v^{p+1} \, dz \right)^{\frac{1}{p+1}} C_1 A \frac{1}{(1-\sigma)^2} \left( \sup_{Q_{(\rho, \theta)}} v \right)^{p-2} \right) \wedge 1, \quad (5.11)$$

Note that the constants $E$, $C_1$ and $A$ are the analogous versions of those defined in Definition 5.1 with this specific choices of $\alpha, \beta, \gamma$, but by an abuse of notation, we still use the same symbols.
For any $\epsilon \in (0, 2]$, we get the following analogue of (5.9)

$$M_n \leq 2^{2n\Sigma} M_{n+1}^{\frac{\xi - 2\epsilon}{2}} \left( E_{p+2} \left( \int_{Q_{(p, s)}} v^{p-2+\epsilon} \, dz \right)^{\frac{2}{p+2}} \frac{C_1A}{(1 - \sigma)^2} \right) \wedge 1,$$

where $\xi = \frac{N}{p-2} + 2(p+1)$ and $\Sigma = \frac{N + 2}{N(p-2) + 2(p+1)}$. Thus iterating the above estimate using Lemma 2.7 gives the desired estimate.

7. Proof of Corollary 3.8

In (4.10), let us take $\alpha = \gamma = 2 - p$ and $\beta = p - 1 + \gamma = 1$ which are admissible since $\frac{2N}{N+2} < p \leq 2$. Then the analogue of (5.7) in this case becomes

$$\sup_{Q_{(p, s)}} v \leq k = \left( E_{p+2} \left( \int_{Q_{(p, s)}} v^3 \, dz \right)^{\frac{2}{p+2}} \frac{C_1A}{(1 - \sigma)^2} \left( \sup_{Q_{(p, s)}} v \right)^{2-p} \right)^{\frac{N+2}{2(p+1)}} \wedge 1,$$

Note that the constants $E, C_1$ and $A$ are the analogous versions of those defined in Definition 5.1 with this specific choices of $\alpha, \beta, \gamma$, but by an abuse of notation, we still use the same symbols.

For any $\epsilon \in (2-p, 3]$, we get the following analogue of (5.9)

$$M_n \leq 2^{2n\Sigma} M_{n+1}^{\frac{\xi - 2p\epsilon}{2}} \left( E_{p+2} \left( \int_{Q_{(p, s)}} v^p \, dz \right)^{\frac{2}{p+2}} \frac{C_1A}{(1 - \sigma)^2} \right) \wedge 1,$$

where $\xi := 6 + p(N+2)$ and $\Sigma = \frac{N + 2}{(2-p)N+6}$ which are defined in Definition 5.1 with $\alpha = \gamma = 2 - p$ and $\beta = 1$. We see that $\epsilon + p - 2 > 0 \iff 2 - p < \epsilon$, thus iterating the above estimate using Lemma 2.7 gives the desired estimate.

8. Proof of Uniform $C^{1, \alpha}$ estimate - Theorem 3.10

With $Q_0 = B_{R_0} \times (-R_0^2, R_0^2)$, let us consider equations of the form

$$u_t - \text{div} \mathcal{A}(\nabla u) = 0 \quad \text{on} \quad 4Q_0,$$

with $\mathcal{A}(\zeta)$ satisfying the following structural assumptions for some $s \in (0, 1]$:

$$|\mathcal{A}(\zeta)| + |\mathcal{A}'(\zeta)||\zeta|^2 + s^2 \frac{3}{2} \leq C_1(|\zeta|^2 + s^2)^{\frac{p+1}{2}},$$

$$\langle \mathcal{A}'(\zeta), \zeta \rangle \geq C_0(|\zeta|^2 + s^2)^{\frac{p+1}{2}} |\zeta|^2,$$

where we have denoted $\mathcal{A}'(\zeta) := \frac{d\mathcal{A}(\zeta)}{dz}$. Let us fix the following constant:

$$\mu_0 := \max\{1, \sup_{4Q_0} |\nabla u|\},$$

then for any $z \in Q_0$, we consider the cylinder $Q^{\mu_0}_S(z) = B_{\mu_0^{-1}S(x)} \times (t - \mu_0^{-p}S^2, t + \mu_0^{-p}S^2)$ to be the largest cylinder such that $Q^{\mu_0}_S(z) \subset 4Q_0$ and $Q^{\mu_0}_S(z) \cap (4Q_0 \setminus 2Q_0) \neq \emptyset$. Note that this fixes the radius $S$ and is independent of the point $z \in Q_0$. As a consequence, we have the following observations:

(O1): Since $Q^{\mu_0}_S(z) \subset 4Q_0$, we see that $S \leq \min\{\mu_0, \mu_0^{p/2}\}3R_0$ must hold.

(O2): Moreover, since $Q^{\mu_0}_S(z)$ has to go outside $2Q_0$, we note that $S \geq \max\{\mu_0, \mu_0^{p/2}\}R_0$. 

14
Since the proof will be independent of the point \( z \in Q_0 \), we shall ignore writing the location of the cylinder \( Q_R^\mu \).

**Proposition 8.1.** For some \( \mu \leq \mu_0 \) and \( R \leq S \), there exists numbers \( \nu \in (0, 1/2) \) and \( \kappa, \delta \in (0, 1) \) depending only on \((N, p, C_0, C_1)\) such that if

\[
|\{z \in Q_R^\mu : |\nabla u(z)| < \mu/2\}| < \nu|Q_R^\mu| \quad \text{and} \quad s + \sup_{Q_R^\mu} |\nabla u| \leq \mu,
\]

is satisfied, then the following conclusion follows:

\[
\int_{Q_{\delta i+1}^\mu} |\nabla u - (\nabla u)_{Q_{\delta i+1}^\mu}|^2 \, dz \leq \kappa \delta^{N+2} \int_{Q_{\delta i}^\mu} |\nabla u - (\nabla u)_{Q_{\delta i}^\mu}|^2 \, dz
\]

for all \( i \in \mathbb{Z} \).

**Proposition 8.2.** For some \( \mu \leq \mu_0 \) and \( R \leq S \) with \( \nu \) as fixed in Proposition 8.1, there exists numbers \( \frac{1}{2} < \sigma < \eta < 1 \) such that if

\[
|\{z \in Q_R^\mu : |\nabla u(z)| < \mu/2\}| \geq \nu|Q_R^\mu| \quad \text{and} \quad s \leq \mu,
\]

is satisfied, then the following conclusion follows:

\[
|\nabla u(z)| \leq \eta \mu \quad \text{for all} \ z \in Q_{\sigma R}^\mu.
\]

**Remark 8.3.** Let us explain a little more about the condition \( s \leq \mu \) in Proposition 8.2. Since we apply Proposition 8.2 starting from \( Q_{S}^{\mu_0} \), we automatically have \( \sup_{Q_{S}^{\mu_0}} |\nabla u| \leq \mu_0 \) and since \( s \in (0, 1) \implies s \leq \mu_0 \), we get

\[
s + \sup_{Q_{S}^{\mu_0}} |\nabla u| \leq 2\mu_0,
\]

which is the hypothesis needed in Proposition 8.12. From this, we get that \( \sup_{Q_{S}^{\mu_1}} |\nabla u| \leq \mu_1 \), but there is no bound obtainable for \( s \) in terms of \( \mu_1 \). Thus we need to assume \( s \leq \mu_1 \) in order to satisfy the hypothesis from Proposition 8.12 to be able to iterate Proposition 8.2. We proceed this way until either the measure density hypothesis fails or \( s \geq \mu_i \).

If the measure density condition fails first at step \( i_0 \), then we still have \( s \leq i_0 \) and the conclusion of Proposition 8.2 at step \( i_0 \) gives \( \sup_{Q_{R_0}^{\mu_{i_0}}} |\nabla u| \leq \mu_{i_0} \), which satisfies the hypothesis needed to apply Proposition 8.1.

On the other hand, if \( s \geq \mu_{i_0} \) happens first, then we are automatically in the uniformly parabolic situation after the rescaling as in (8.26) and we can directly apply Lemma 8.16 to get the conclusion of Proposition 8.1. This is
possible since we have

\[ \mu_{i_0} \leq s \leq s + \sup_{Q^\mu_{i_0}} |\nabla u| \leq \mu_{i_0-1} + \mu_{i_0} = \left( \frac{1 + \eta}{\eta} \right) \mu_{i_0}, \]

which immediately gives (8.43) with \( \mu \). Recall that \( \eta \) is a universal constant.

8.1. Covering argument

Let us first fix some notation needed for the proof:

**Definition 8.4.** Let \( \mu \) be as in (8.3) and \( \eta, \sigma \) be as in Proposition 8.2. Then let us denote

\[ \mu = \mu_0, \quad \mu_{n+1} := \eta \mu_n, \quad R := S, \quad R_{n+1} := c_0 R_n = c_0^{n+1} R, \]

where \( c_0 := \frac{1}{2} \sigma \min\{\eta, \eta^{p/2}\} \). Here we note that \( c_0 \in (0, 1) \) since \( \sigma, \eta \in (0, 1) \) and \( \mu \geq 1 \).

We have the following result regarding the intrinsic geometry.

**Claim 8.5.** With the notation as in Definition 8.4, we have the inclusion \( Q^\mu_{R_{i+1}} \subset Q^\mu_{R_i} \).

**Proof.** In order for this to hold, we must have the following:

\[ \begin{cases} (\eta \mu)^{-1} c_0 R & \leq \mu^{-1} \sigma R \\ (\eta \mu)^{-p} (c_0 R)^2 & \leq \mu^{-p} (\sigma R)^2 \end{cases} \iff \begin{cases} c_0 & \leq \eta \sigma \\ c_0^2 & \leq \eta^{p \sigma^2} \end{cases} \]

From the restriction \( \sigma < \eta < 1 \) and the choice \( c_0 = \frac{1}{2} \sigma \min\{\eta, \eta^{p/2}\} \), we see that the above two restrictions hold.

**Claim 8.6.** We have \( R_n = c_0^n R \) and \( \eta^n = \left( \frac{R_n}{R} \right)^{\alpha_1} \) where \( \alpha_1 = -\log_\eta c_0 \).

**Proof.** From direct computation, we see that

\[ \frac{1}{\alpha_1} = -\log c_0 \leq \frac{1}{\log \eta} = \frac{-\log c_0}{\log \eta} = \frac{\log c_0}{\log \eta} \]

Thus we get

\[ \eta^n = \left( \frac{R_n}{R} \right)^{\frac{\log c_0}{\log \eta}} \iff n \log \eta = \log c_0 \log \left( \frac{R_n}{R} \right) \iff \log c_0^n = \log \left( \frac{R_n}{R} \right) \iff R c_0^n = R_n. \]

This proves the claim.

8.1.1. Switching radius

Since we have to study the interplay between Proposition 8.1 and Proposition 8.2, we need to define what is known as the switching radius:

**Definition 8.7 (Switching Radius).** With the notation from Definition 8.4, suppose the hypothesis of Proposition 8.2 holds at level \((R_1, \mu_1), (R_2, \mu_2) \ldots (R_i, \mu_i) \), i.e.,

\[ |\{ Q^\mu_{R_i} : |\nabla u| < \mu_i/2 \} | \geq \nu |Q^\mu_{R_i}| \quad \text{and} \quad s \leq \mu_i, \] (8.4)

holds, then applying Proposition 8.2, we conclude

\[ |\nabla u| \leq \eta \mu_i = \mu_{i+1} \quad \text{on} \quad Q^\mu_{R_{i+1}} \supseteq Q^\mu_{R_{i+1}}. \]

Continuing this way, we denote \( n_0 \) (called the switching number) and radius \( R_{n_0} \) (called the switching radius) to be the first instance where one of the conditions from (8.4) fails, i.e.,

\[ |\{ Q^\mu_{R_{n_0}} : |\nabla u| < \mu_{n_0}/2 \} | < \nu |Q^\mu_{R_{n_0}}| \quad \text{or} \quad \mu_{n_0} \leq s. \] (8.5)
From Definition 8.7, we see that the following estimates hold:

- For \( n = 1, \ldots, n_0 \), due to Proposition 8.2 and Claim 8.5 there holds
  \[
  \sup_{\Omega_{\rho_{n_0}}^{R_{n_0}}} |\nabla u| \leq \mu_n = \eta^n \mu. \tag{8.6}
  \]

- We also have for any \( n = 0, 1, \ldots, n_0 \),
  \[
  \sup_{\Omega_{\rho_{n_0}}^{R_{n_0}}} |\nabla u| \leq \mu \left( \frac{R_n}{R} \right)^{\alpha_1}. \tag{8.7}
  \]

  This follows from (8.6) and Claim 8.6.

- Since the first condition from (8.5) is the hypothesis required in Proposition 8.1, there holds
  \[
  \iint_{\Omega_{\rho_{n_0}}^{R_{n_0}}} |\nabla u - (\nabla u)_i|^2 \, dz \leq \kappa^i \iint_{\Omega_{\rho_{n_0}}^{R_{n_0}}} |\nabla u - (\nabla u)_0|^2 \, dz \leq \kappa^i \mu_{n_0}^2 \quad \text{for} \quad i = 1, 2, \ldots, \tag{8.8}
  \]
  where to obtain (a), we have used the notation \((\nabla u)_i := \iint_{\Omega_{\rho_{n_0}}^{R_{n_0}}} \nabla u \, dz\) and the bound \(\sup_{\Omega_{\rho_{n_0}}^{R_{n_0}}} |\nabla u| \leq \mu_{n_0}\),
  which holds due to (8.6). Moreover, we made use of \(\iint |f - (f)|^2 \, dz = \inf_{a \in \mathbb{R}} \iint |f - a|^2 \, dz \leq \iint |f|^2\).

- If the second condition from (8.5) holds, then we have following bounds:
  \[
  \sup_{\Omega_{\rho_{n_0}}^{R_{n_0}}} |\nabla u| \leq \mu_{n_0} \quad \text{and} \quad \mu_{n_0} \leq s \leq s + \sup_{\Omega_{\rho_{n_0}}^{R_{n_0}}} |\nabla u|,
  \]
  which in particular implies that the hypothesis required to apply Lemma 8.16 is satisfied. Thus even in this case, we can directly apply Lemma 8.16 to get the conclusion given in (8.8), see more details in Remark 8.3.

**Remark 8.8.** From triangle inequality, for \( i = 1, 2, \ldots \), we also have
\[
\iint_{\Omega_{\rho_{n_0}}^{R_{n_0}}} |\nabla u - (\nabla u)_{i-1}|^2 \, dz \leq \iint_{\Omega_{\rho_{n_0}}^{R_{n_0}}} |\nabla u - (\nabla u)_i|^2 \, dz + \frac{|\Omega_{\rho_{n_0}}^{R_{n_0}}|}{|\Omega_{\rho_{n_0}}^{R_{n_0}}|} \iint_{\Omega_{\rho_{n_0}}^{R_{n_0}}} |\nabla u - (\nabla u)_{i-1}|^2 \, dz \leq \kappa^{i-1} \left( 1 + \frac{1}{\delta^{N+2}} \right). \tag{8.9}
\]

**Lemma 8.9.** Let \( \kappa, \delta, \eta \) be as given in Proposition 8.1 and Proposition 8.2, then we have the following important consequences from (8.6) and (8.8).

- **(C1):** The sequence \( \{ (\nabla u)_i \}_{i=1}^{\infty} \) from (8.8) is a Cauchy sequence and converges to \( \nabla u(z_0) \) where \( z_0 = (x_0, t_0) \) is the center of the parabolic cylinders considered in (8.6) and (8.8).

- **(C2):** The following decay estimate holds:
  \[
  |\nabla u(x_0, t_0) - (\nabla u)_i| \leq C \kappa^j \mu_{n_0}^2 \quad \text{for all} \quad i = 1, 2, \ldots.
  \]

- **(C3):** For any \( 0 < \rho < R_{n_0} \) with \( (\nabla u)_\rho := \iint_{\Omega_{\rho_{n_0}}^{R_{n_0}}} \nabla u \, dz \), the following decay estimate holds:
  \[
  |\nabla u(x_0, t_0) - (\nabla u)_\rho| \leq C(\delta) \kappa^i \mu_{n_0}^2,
  \]
  where \( i \in \mathbb{Z} \) is such that \( \delta^i R_{n_0} \leq \rho \leq \delta^{i-1} R_{n_0} \).
\((\text{C}4):\) Let us define \(\alpha_3 := \min\{\alpha_1, \alpha_2/2\}\) where \(\alpha_2 := -\log_2 \delta\) and \(\alpha_1\) is from Claim 8.6, then for any \(0 < \rho < R\), with \((\nabla u)_\rho := \iint_{Q_{\rho}^{R_{\rho}}(x_0, t_0)} \nabla u \, dz\), there holds
\[
|\nabla u(x_0, t_0) - (\nabla u)_\rho| \leq C \mu \left( \frac{\rho}{R} \right)^{\alpha_3}.
\]

\((\text{C}5):\) With \(\alpha_3\) as defined in \((\text{C}4)\), for any \(0 < \rho \leq R\) with \((\nabla u)_\rho := \iint_{Q_{\rho}^{R_{\rho}}(x_0, t_0)} \nabla u \, dz\), we also have
\[
\iint_{Q_{\rho}^{R_{\rho}}} |\nabla u - (\nabla u)_\rho|^2 \, dz \leq C \mu^2 \left( \frac{\rho}{R} \right)^{2\alpha_3}.
\]

\textbf{Proof.} From simple triangle inequality, we have
\[
|\nabla u|_{i+1} - |\nabla u|_i|^2 \leq 2|\nabla u - (\nabla u)|_{i+1}^2 + 2|\nabla u - (\nabla u)|_i^2.
\]  
(8.10)

\textbf{Proof of (C1):} From Lebesgue differentiation theorem, we see that \((\nabla u)_i \overset{i \to \infty}{\to} \nabla u(z_0)\) provided \(z_0\) is a Lebesgue point. So all that remains to show that is that the sequence \((\nabla u)_i\) is Cauchy which follows from the following sequence of estimates noting that \(\kappa \in (0, 1)\):
\[
|\nabla u|_{i+1} - |\nabla u|_i|^2 \overset{(8.10)}{=} 2 \iint_{Q_{\rho}^{R_{\rho}}} |\nabla u - (\nabla u)|_{i+1}^2 \, dz + 2 \iint_{Q_{\rho}^{R_{\rho}}} |\nabla u - (\nabla u)|_i^2 \, dz
\leq 2 \kappa^{i+1} \mu^2_{n_0} + \frac{2}{\delta^{N+2}} \kappa^i \mu^2_{n_0} = 2 \kappa^i \mu^2_{n_0} \left( \kappa + \frac{1}{\delta^{N+2}} \right). \tag{8.11}
\]

\textbf{Proof of (C2):} By adding and subtracting, for any \(j \geq 1\), we have
\[
|\nabla u|_{i+j} - |\nabla u|_i|^2 \overset{(8.11)}{\leq} 2 \kappa^i \mu^2_{n_0} \left( \kappa + \frac{1}{\delta^{N+2}} \right) (1 + \kappa + \kappa^2 + \ldots + \kappa^j)
\leq 2 \kappa^i \mu^2_{n_0} \left( \kappa + \frac{1}{\delta^{N+2}} \right) \frac{1}{1 - \kappa} = C(\kappa, \delta) \kappa^i \mu^2_{n_0}.
\]

In particular, letting \(j \to \infty\), the following holds:
\[
|\nabla u(x_0, t_0) - (\nabla u)|_i|^2 \leq C \kappa^i \mu^2_{n_0} \quad \text{for all} \quad i = 1, 2, \ldots.
\]

\textbf{Proof of (C3):} Let \(\rho < R_{n_0}\) be given, then there exists \(i \in \mathbb{Z}\) such that
\[
\delta^i R_{n_0} \leq \rho \leq \delta^{i-1} R_{n_0}, \tag{8.12}
\]
holds, where \(\delta\) is the constant from Proposition 8.1. Thus, we have the following sequence of estimates:
\[
|\nabla u|_\rho - |\nabla u|_i|^2 \leq \iint_{Q_{\rho}^{R_{\rho}}} |\nabla u - (\nabla u)_i|^2 \, dz \leq \iint_{Q_{\rho}^{R_{\rho}}} |\nabla u - (\nabla u)|_i^2 \, dz \overset{(a)}{\leq} \frac{\delta^{i-1} R_{n_0}^{N+2}}{\rho^{N+2}} C \kappa^i \mu^2_{n_0} \overset{(8.12)}{\leq} \frac{C}{\delta^{N+2}} \kappa^i \mu^2_{n_0}, \tag{8.13}
\]
where to obtain (a), we made use of (8.8) and (8.9). Thus from triangle inequality, we get the desired conclusion:
\[
|\nabla u(x_0, t_0) - (\nabla u)|_\rho|^2 \leq 2|\nabla u(x_0, t_0) - (\nabla u)|_i^2 + 2|\nabla u)_i - (\nabla u)|_\rho|^2 \leq C(\delta, \kappa) \kappa^i \mu^2_{n_0}. \tag{8.14}
\]

\textbf{Proof of (C4):} We split the proof into two cases, either \(\rho \leq R_{n_0}\) or \(R_{n_0} \leq \rho \leq R\).

\textbf{Case} \(\rho \leq R_{n_0}\): In this case, there exists \(i \in \mathbb{Z}\) such that \(8.12\) holds. Following the calculation from Claim 8.6,
we see that \( \kappa^i \leq \left( \frac{\rho}{R_{n_0}} \right)^{\alpha_2} \). Using this, we get the following sequence of estimates:

\[
|\nabla u(x_0, t_0) - (\nabla u)_\rho| \leq C \kappa^i \mu_{n_0} \leq C \left( \frac{\rho}{R_{n_0}} \right)^{\alpha_2} \mu_{n_0} = C \left( \frac{\rho}{R_{n_0}} \right)^{\alpha_2} \eta_{n_0} \mu
\]

\[\text{Claim 8.6} \]

\[
= C \left( \frac{\rho}{R_{n_0}} \right)^{\alpha_2} \left( \frac{R_{n_0}}{R} \right)^{\alpha_1} \mu
\]

\[\leq \left( \frac{\rho}{R_{n_0}} \right)^{\alpha_3} \mu
\]

\[= C \left( \frac{\rho}{R} \right)^{\alpha_3} \mu.
\]

In the above estimate, we noted that \( \rho \leq R_{n_0} \leq R \) and \( \alpha_3 = \min \{ \alpha_1, \alpha_2 / 2 \} \).

**Case** \( R_{n_0} \leq \rho \leq R \): From the definition of \( R_{n_0} \) (see Definition 8.7), we see that there exists some \( i \in \{1, 2, \ldots, n_0\} \) such that \( c_0^i R \leq \rho \leq c_0^{i-1} R \) (recall \( c_0 \in (0, 1) \) from Claim 8.5). Thus we get the following sequence of estimates:

\[
|\nabla u(x_0, t_0) - (\nabla u)_\rho| \leq 2 \sup_{Q_{\rho^{i-1}}} |\nabla u| \leq 2 \eta^{-1} \mu
\]

\[
= 2 \left( \frac{R_{n_0}}{R} \right)^{\alpha_1} \mu \quad \text{Definition 8.4} \]

\[\leq \frac{2}{c_0^i} \left( \frac{\rho}{R} \right)^{\alpha_1} \mu
\]

\[\leq C \left( \frac{\rho}{R} \right)^{\alpha_3} \mu.
\]

**Proof of (C5)**: If \( \rho \geq R_{n_0} \), then the conclusion follows directly from (8.16). In the case \( \rho \leq R_{n_0} \), there exists \( i \in \mathbb{Z} \) such that (8.12) holds. Using this, we get

\[
\int \nabla u - (\nabla u)_\rho | dz \leq 2 \int Q_{\rho^i}^n |\nabla u - (\nabla u)_i|^2 dz + 2 |(\nabla u)_i - (\nabla u)_\rho|^2
\]

\[
\leq \frac{|Q_{\rho^{i-1}}^{n_0}|}{|Q_{\rho}^n|} \int Q_{\rho^{i-1}}^{n_0} |\nabla u - (\nabla u)_{i-1}|^2 dz + |(\nabla u)_{i-1} - (\nabla u)_\rho|^2
\]

\[= C \frac{\rho^{i-1}}{\delta^{N+P}} \mu_{n_0}^2
\]

\[= C \left( \frac{\rho}{R} \right)^{2\alpha_3} \mu^2,
\]

where to obtain (a), we made use of (8.8) along with (8.13).

This completes the proof of the lemma.

8.1.2. **Proof of gradient H"older continuity in time in the cylinder** \( Q_R^u \)

Let us fix any point \( z_0 = (x_0, t_0) \in Q_0 \) and let \( \tilde{z}_1 = (x_1, t_0) \in Q_0 \) be given. In this case, we assume that the two points \( z_0, \tilde{z}_1 \in Q_R^u \) belong to the same cylinder for some cylinder \( Q_R^u \). We will prove that the gradient is Hölder continuous at \( z_0 \) and since the point \( z_0 \in Q_0 \) is arbitrary, this proves Hölder regularity in cylinders of the form \( Q_R^u \).

**Remark 8.10.** From Definition 8.7, let us denote \( \mu_{n_0} \) (respectively \( \mu_{n_1} \)) to be the switching number and \( R_{n_0} \) (respectively \( R_{n_1} \)) to be the switching radius corresponding to the point \( z_0 \) (respectively \( z_1 \)). Note that even though these two switching numbers depend on the point, all the estimates and constants in Lemma 8.9 are independent of the point.

Let \( \rho := d(z_0, \tilde{z}_1) \), then from triangle inequality, we have

\[
|\nabla u(x_0, t_0) - \nabla u(x_1, t_1)| \leq |\nabla u(x_0, t_0) - (\nabla u)_{C_0}| + |(\nabla u)_{C_1} - \nabla u(x_0, t_1)| + |(\nabla u)_{C_1} - (\nabla u)_{C_1}|,
\]

(8.17)
where we have used the notation \((\nabla u)_{C_1} := \iint_{C_1} \nabla u \, dz\) and \(C_0, C_1\) are cylinders that will eventually be chosen. In what follows, we shall use \(i \in \{0, 1\}\) to denote quantities which are related to \(z_0\) and \(z_1\) respectively.

Without loss of generality, let us assume

\[ \mu_0 \leq \mu_n. \quad (8.18) \]

**Case** \(\min\{\mu_{n_0}^{-p/2}R_{n_0}, \mu_{n_1}^{-p/2}R_{n_1}\} \geq 2\rho\): Let \(\rho = \sqrt{|t_0 - t_1|}\) and set \(S_i := \mu_{n_i}^{p/2}\sqrt{|t_0 - t_1|} = \mu_{n_i}^{p/2}\rho\). Now let us consider the cylinder \(C_0 = Q_{S_0}^\mu(z_0)\) and \(C_1 = Q_{S_1}^\mu(z_1)\), then we have the following observations:

- \(C_0, C_1 \in Q_R^\mu\). In order to see this, we need to show \(S_i\mu_{n_i}^{-1} \leq \mu R\) and \(\mu_{n_1}^{-p}S_1^2 \leq \mu^{-p}R^2\). The space inclusion holds due to the following calculations:

  \[ S_i\mu_{n_i}^{-1} \leq \mu^{-1}R \iff \rho \leq \left(\frac{n_i}{\mu n_i^{\frac{p}{2}}}\right)^{n_0} \mu_{n_1}^{\frac{p}{2}}R_{n_1} \iff \rho \leq \left(\frac{2}{\sigma}\right)^{n_0} \mu_{n_1}^{\frac{p}{2}}R_{n_1}, \]

  and the last inequality holds true since we are in the case \(2\rho \leq \mu_{n_1}^{\frac{p}{2}}R_{n_1}\) where \(\sigma \in (0, 1)\). The time inclusion is trivial since \(\mu_{n_1}^{-p}S_1^2 = |t_0 - t_1| = \rho^2 \leq \mu^{-p}R^2\) which holds true since \(z_0, z_1 \in Q_R^\mu\).

- Clearly, we see that \(C_0 \cap C_1 \neq \emptyset\) and we additionally have the following:

  \[ |C_0 \cap C_1| \approx \min\{\mu_{n_0}^{-1}, \mu_{n_1}^{-1}\}^N|t_0 - t_1|^{\frac{N-1}{2}}, \quad |C_0| \approx \mu_{n_0}^{N(\frac{\xi}{2}-1)}|t_0 - t_1|^{\frac{N-1}{2}}, \quad |C_1| \approx \mu_{n_1}^{N(\frac{\xi}{2}-1)}|t_0 - t_1|^{\frac{N-1}{2}}. \]

Thus, making use of \((8.18)\), we additionally observe the following:

\[
\begin{align*}
\frac{|C_0|}{|C_0 \cap C_1|} & = \min \left\{ \frac{N(p-1)}{\mu_{n_0} N_{n_0}} \cdot \frac{\mu_{n_0}^{N-1}}{\mu_{n_0}}, \frac{N(p-1)}{\mu_{n_1} N_{n_1}} \cdot \frac{\mu_{n_1}^{N-1}}{\mu_{n_1}} \right\} \leq \left( \frac{\mu_{n_1}}{\mu_{n_0}} \right)^{\frac{N}{2}} \quad \text{(8.18)}, \\
\frac{|C_1|}{|C_0 \cap C_1|} & = \min \left\{ \frac{N(p-1)}{\mu_{n_0} N_{n_0}} \cdot \frac{\mu_{n_0}^{N-1}}{\mu_{n_1}}, \frac{N(p-1)}{\mu_{n_1} N_{n_1}} \cdot \frac{\mu_{n_1}^{N-1}}{\mu_{n_1}} \right\} \leq \left( \frac{\mu_{n_1}}{\mu_{n_0}} \right)^{\frac{N(p-1)}{2}}. \quad (8.19)
\end{align*}
\]
Thus the first two terms on the right hand side of (8.17) is estimated as follows (recall from Remark 3.11):

\[
|\nabla u(x_0, t_i) - (\nabla u)_{C_i} | \leq C \mu \left( \frac{S_i}{R} \right)^{\alpha_3} = C \mu \left( \frac{\mu_n^{p/2} \rho}{R} \right)^{\alpha_3},
\]

(8.20)

The third term on the right hand side of (8.17) is estimated as follows:

\[
| (\nabla u)_{C_0} - (\nabla u)_{C_1} | \leq \int_{C_0 \cap C_1} |\nabla u(z) - (\nabla u)_{C_0} | dz + \int_{C_0 \cap C_1} |\nabla u(z) - (\nabla u)_{C_1} | dz
\]

(8.21)

Removing dependence on \{\mu_n\}: The estimates in (8.21) still contain \mu_n and \mu_n, which we control as follows.

Let \gamma \in (0, 1) be a constant satisfying

\[
\gamma \leq \frac{\alpha_3}{2N \max\{p - 1, 1\}}.
\]

We then consider the following two cases, either \mu_n \leq \mu_n \left( \frac{\rho}{R_0} \right)^\gamma or \mu_n \leq \mu_n \left( \frac{\rho}{R_0} \right)^\gamma. Let us first obtain the Hölder continuity of the gradient in the case

\[
\mu_n \geq \mu_n \left( \frac{\rho}{R_0} \right)^\gamma.
\]

Thus making use of (8.22) along with (8.20), we get the following estimates:

\[
|\nabla u(x_0, t_0) - \nabla u(x_0, t_1)| \leq C \mu_0 \left( \frac{\sqrt{t_0 - t_1}}{R_0} \right)^{\alpha_3 - \frac{\alpha_3}{2} \max\{p - 1, 1\}},
\]

Now we consider the case (8.22) fails, then we see that \(z_0, \tilde{z}_1\) both belong to \(C_0\), since we are in the case \(2\rho \leq \min\{\mu_n^{p/2} R_n, \mu_n^{p/2} R_n\}\). Thus, we have

\[
|\nabla u(x_0, t_0) - \nabla u(x_0, t_1)| \leq 2\mu_n \leq 2\mu_0 \left( \frac{\rho}{R_0} \right)^\gamma \leq 2\mu \left( \frac{\sqrt{t_0 - t_1}}{R_0} \right)^\gamma.
\]

Case \(\max\{\mu_n^{p/2} R_n, \mu_n^{p/2} R_n\} \leq 2\rho\): In this case, we directly proceed as follows (recall notation from Remark 3.11):

\[
|\nabla u(z_0) - \nabla u(z_1)| \leq \sup_{Q_{R_n}^{\mu_n^{p/2} R_n}(z_0)} |\nabla u| + \sup_{Q_{R_n}^{\mu_n^{p/2} R_n}(z_1)} |\nabla u| \leq \eta_n \mu + \mu^{\alpha_1} \mu
\]

Claim 8.6

\[
\leq C \mu_0 \left( \frac{R_0}{R} \right)^{\alpha_1} \leq \mu \left( \frac{\rho}{R} \right)^{\alpha_1},
\]

(8.22)

Case \(\min\{\mu_n^{p/2} R_n, \mu_n^{p/2} R_n\} \leq 2\rho \leq \max\{\mu_n^{p/2} R_n, \mu_n^{p/2} R_n\}\): Recalling Claim 8.5 and (8.18), we see that this case becomes \(\mu_n^{p/2} R_n \leq 2\rho \leq \mu_n^{p/2} R_n\). Let \(n_0 \leq n_s \leq n_1\) be a number such that \(\mu_n^{p/2} R_n \leq \mu_n^{p/2} R_n \leq \mu_n^{p/2} R_n\) holds. Recalling Claim 8.5, we then have

\[
2\rho \leq \min\{\mu_n^{p/2} R_n, \mu_n^{p/2} R_n\}.
\]

Thus we can replace \(C_0\) in the first case with \(C_0 := Q_{R_n}^{\mu_n^{p/2} R_n}(z_0)\) and make use of (8.7) followed by going through the
calculations of the first case to obtain the desired regularity. We now present a rough sketch below for the gradient Hölder regularity in time noting that analogous estimates follow for the space counterpart:

The first term on the right hand side of (8.17) is estimated as follows:

\[ |\nabla u(x_0, t_0) - (\nabla u)_{c_1}| \leq 2\mu_n \leq C\mu \left( \frac{R_n}{R} \right)^{\alpha_1} \leq C\mu^a \left( \frac{\rho^{n+1}}{R} \right)^{\alpha_1} \leq C\mu^a \left( \frac{\rho}{R} \right)^{\alpha_1}. \]  
(8.23)

The second term on the right hand side of (8.17) is estimated exactly as (8.20). In order to estimate the last term on the right hand side of (8.17), we make the following observations:

\[ \frac{|C_1|}{|C_\ast \cap C_1|} \leq \frac{\min \left\{ \frac{N(p-2)}{\mu_n^2}, \frac{N(p-2)}{\mu_n \alpha_n} \right\}}{\min \left\{ \frac{N(p-2)}{\mu_n^2}, \frac{N(p-2)}{\mu_n \alpha_n} \right\}} \leq \left( \frac{\mu_n}{\mu_n} \right) \frac{N(p-1)}{2}, \]  
(8.24)

where to obtain the last estimate, we used the fact that \( n_\ast \leq n_1 \) and proceeded as in (8.19). Now we estimate the last term on the right hand side of (8.17) as follows:

\[ |(\nabla u)_{c_1} - (\nabla u)_{c_1}| \leq \frac{\min \left\{ \frac{N(p-2)}{\mu_n^2}, \frac{N(p-2)}{\mu_n \alpha_n} \right\}}{\min \left\{ \frac{N(p-2)}{\mu_n^2}, \frac{N(p-2)}{\mu_n \alpha_n} \right\}} \leq \left( \frac{\mu_n}{\mu_n} \right) \frac{N(p-1)}{2}. \]  
(8.20)

where to obtain (a), we made use of (8.23) along with (8.24) and (C5). From here onwards, we can proceed as the first case to remove the dependence of the estimate on \( \left( \frac{\mu_n}{\mu_n} \right) \frac{N(p-1)}{2} \).

8.1.3. Proof of gradient Hölder continuity in space in the cylinder \( Q_{R}^\mu \)

Let us fix any point \( z_0 = (x_0, t_0) \in Q_0 \) and let \( \hat{z}_1 = (x_1, t_0) \in Q_0 \) be given. In this case, we assume that the two points \( z_0, \hat{z}_1 \in Q_{R}^\mu \) belong to the same cylinder for some cylinder \( Q_{R}^\mu \). We will prove that the gradient is Hölder continuous at \( z_0 \) and since the point \( z_0 \in Q_0 \) is arbitrary, this proves Hölder regularity in cylinders of the form \( Q_{R}^\mu \).

Let \( \rho := d(z_0, \hat{z}_1) \), we have the same estimate as (8.17) and without loss of generality, let us assume (8.18) holds. Since the proof is very similar to the time case, we only give a rough sketch.

**Case** \( \min \{ \mu_n^{-1} R_n, \mu_n^{-1} R_n \} \geq 2\rho \) Let us set \( \rho = |x_0 - x_1| \) and \( S_i := \mu_n |x_0 - x_1| = \mu_n \rho \) and construct the cylinders \( C_0 = Q_{S_0}^{\mu_n} (z_0) \) and \( C_1 = Q_{S_1}^{\mu_n} (\hat{z}_1) \), then we have the following observations:
• In order for $C_i \in Q^p_R$, we need $S_i \mu_n^{-1} \leq \mu^{-1} R$ and $\mu_n^{-1} S_i^2 \leq \mu^{-p} R^2$. The time inclusion holds due to the following calculations:

$$S^2 \mu_n^{-p} \leq \mu^{-p} R^2 \iff \rho \leq \left( \frac{p/2}{c_0} \right)^{n_1} \mu_n^{-1} R_n \iff \rho \leq \left( \frac{2}{\sigma} \right)^{n_0} \mu_n^{-1} R_n,$$

and the last inequality holds true since we are in the case $2 \rho \leq \mu_n^{-1} R_n$, where $\sigma \in (0, 1)$. The space inclusion is trivial since $\mu_n^{-1} S_i = |x_0 - x_1| = \rho \leq \mu^{-1} R^2$ which holds true since $z_0, z_1 \in Q^p_R$.

• The analogue of (8.19) becomes

$$\frac{|C_0|}{|C_0 \cap C_1|} = \frac{\mu_n^{-1} \mu_n}{\min \{ \mu_n^{-1} \mu_n, \mu_n^{-1} \mu_n \}} \leq \left( \frac{\mu_n}{\mu_n} \right)^{p-1},$$

and

$$\frac{|C_1|}{|C_0 \cap C_1|} = \frac{\mu_n^{-1} \mu_n}{\min \{ \mu_n^{-1} \mu_n, \mu_n^{-1} \mu_n \}} \leq \left( \frac{\mu_n}{\mu_n} \right)^{p-1}.$$

• Rest of the calculations goes through verbatim as in the time case.

Case $\max \{ \mu_n^{-1} R_n, \mu_n^{-1} R_n \} \leq 2 \rho$: This case follows exactly as the time case.

Case $\min \{ \mu_n^{-1} R_n, \mu_n^{-1} R_n \} \leq 2 \rho \leq \max \{ \mu_n^{-1} R_n, \mu_n^{-1} R_n \}$: Recalling Claim 8.5 and (8.18), this case becomes $\mu_n^{-1} R_n \leq 2 \rho \leq \mu_n^{-1} R_n$. We can now replace $C_0$ with $C_*$ where $C_* = Q^p_{R_n}(z_0)$ where $n_0 \leq n_s \leq n_1$ is a number satisfying $\mu_n^{-1} R_n \leq 2 \rho \leq \mu_n^{-1} R_n$. Rest of the calculations go through exactly as in the time case.

8.1.4. Proof of gradient Hölder continuity in $Q_0$

In the previous subsection, we proved gradient Hölder continuity at any two points provided both of them belonged to $z_0, z_1 \in Q^p_R$. In this subsection, we consider the case where both of the points does not belong to a single $Q^p_R$, i.e., if we consider the cylinder $Q^p_{R_n}(z_0)$, we are in the case $z_1 \notin Q^p_{R_n}(z_0)$. Thus we observe

$$\max \{ \mu, \mu^{p/2} \} d(z_0, z_1) = d_\mu(z_0, z_1) \geq R \geq \max \{ \mu, \mu^{p/2} \} R_0,$$

holds, using which we get

$$\left| \frac{\nabla u(z_0) - \nabla u(z_1)}{d(z_0, z_1)} \right| \leq \frac{2 \mu}{R_0}. \tag{8.25}$$

This completes the proof of gradient Hölder continuity.
8.2. Proof of first alternative - Proposition 8.1

Since \( u \) is a weak solution of (8.1) on \( Q^r \) for some \( r \in (0, R] \), let us perform the following rescaling: Define
\[
w(x, t) = \frac{w(\mu^{-1}rx, \mu^{-p}r^2t)}{\mu^{-1}r},
\]
then \( w \) solves
\[
\mu^{-1}w_t - \mu \text{div} A(\nabla w) = 0 \quad \text{on} \quad Q_1 := B_1 \times I_1.
\]

Let us first prove an energy estimate satisfied by (8.27).

**Lemma 8.11.** Let \( k \in \mathbb{R} \) and \( \phi \in C^\infty(Q_1) \) be any cut-off function with \( \phi = 0 \) on \( \partial_p(Q_1) \), then the following estimate holds:
\[
\sup_{t \in I_1} \mu^{-1} \int_{B_1} (w_{x_i} - k)^2 \phi^2 \, dx + C_0 \mu \int_{Q_1} (s^2 + |\nabla w|^2)^{\frac{p-2}{2}} |\nabla (w_{x_i} - k)|^2 \phi^2 \, dz \\
\leq 2 \mu^{-1} \int_{Q_1} (w_{x_i} - k)^2 \phi \, dz + \frac{4C^2}{C_0} \mu \int_{Q_1} (s^2 + |\nabla w|^2)^{\frac{p}{2}} |\nabla \phi|^2 \chi_{(w_{x_i} < k)} \, dz \\
+ 2C_1 \mu \int_{Q_1} (s^2 + |\nabla w|^2)^{\frac{p-1}{2}} (w_{x_i} - k)^{-1} (|\phi| |\nabla^2 \phi| + |\nabla \phi|^2) \, dz.
\]

**Proof.** Let us differentiate (8.27) with respect to \( x_i \) for some \( i \in \{1, 2, \ldots, N\} \) and then take \( (w_{x_i} - k)_- \phi^2 \) with \( \phi \in C^\infty_0 \) as a test function to get
\[
\mu^{-1} \int_{Q_1} (w_{x_i}) (w_{x_i} - k)_- \phi^2 \, dz + \mu \int_{Q_1} \langle \partial_i A(\nabla w), \nabla \phi \rangle (w_{x_i} - k)_- \, dz + \mu \int_{Q_1} \langle \partial_i A(\nabla w), \nabla (w_{x_i} - k)_- \rangle \phi^2 \, dz = 0.
\]

Recalling the notation from (8.2), we see that \( \partial_i A(\nabla w) = A'(\nabla w) \nabla w_{x_i} \). Let us estimate each of the terms as follows:

**Estimate for I:** This can be estimated as follows:
\[
\mu^{-1} \int_{Q_1} (w_{x_i}) (w_{x_i} - k)_- \phi^2 = \mu^{-1} \int_{Q_1} ((w_{x_i} - k)_- \phi)_t \, dz - \mu^{-1} \int_{Q_1} (w_{x_i} - k)_- \phi \, dz.
\]

**Estimate for II:** Integrating by parts, we get
\[
\mu \int_{Q_1} \langle \partial_i A(\nabla w), \nabla \phi \rangle \phi (w_{x_i} - k)_- \, dz = -\mu \int_{Q_1} \langle A(\nabla w), \partial_i \nabla \phi \rangle \phi (w_{x_i} - k)_- \, dz \\
- \mu \int_{Q_1} \langle A(\nabla w), \nabla \phi \rangle \partial_i (w_{x_i} - k)_- \, dz
\]
\[
\overset{(a)}{=} C_1 \mu \int_{Q_1} (s^2 + |\nabla w|^2)^{\frac{p-1}{2}} (|\nabla^2 \phi| + |\nabla \phi|^2) |w_{x_i} - k| \, dz \\
+ C_1 \mu \varepsilon \int_{Q_1} (s^2 + |\nabla w|^2)^{\frac{p}{2}} |\nabla \phi|^2 |\nabla (w_{x_i} - k)|^2 \, dz \\
+ C_1 \mu \varepsilon \int_{Q_1} (s^2 + |\nabla w|^2)^{\frac{p-1}{2}} |\nabla \phi|^2 \, dz,
\]

where to obtain (a), we made use of (8.2), the trivial bound \( |\phi| \leq 1 \) and Young’s inequality.

**Estimate for III:** We estimate this term as follows:
\[
\mu \int_{Q_1} \langle \partial_i A(\nabla w), \nabla (w_{x_i} - k)_- \rangle \phi^2 \, dz \overset{(8.2)}{\geq} C_0 \mu \int_{Q_1} (s^2 + |\nabla w|^2)^{\frac{p-1}{2}} |\nabla (w_{x_i} - k)_- |^2 \phi^2 \, dz.
\]

Combining (8.28), (8.29) and (8.30) gives the desired estimate. \( \square \)
8.2.1. DeGiorgi type iteration for $u_x$.

The main proposition we prove is the following:

**Proposition 8.12.** Let $r \in (0, R]$ and assume that

$$s + \sup_{Q_r^\mu} \|\nabla u\| \leq A\mu$$  \hspace{1cm} (8.31)

holds for some $A \geq 1$. For any $i \in \{1, 2, \ldots, N\}$, there exists universal constant $\nu \in (0, 1/2)$ such that if

$$|\{(x, t) \in Q_r^\mu : u_{x_i} < \mu/2\}| \leq \nu |Q_r^\mu|,$$

holds, then we have the following conclusion:

$$u_{x_i} \geq \frac{\mu}{4} \quad \text{on} \quad Q_{r/2}^\mu.$$

Following the rescaling from (8.26), we can restate the following equivalent version of Proposition 8.12 for $w$ as follows:

**Lemma 8.13.** Suppose

$$s + \sup_{Q_1} \|\nabla w\| \leq A\mu$$  \hspace{1cm} (8.32)

holds for some $A \geq 1$. For any $i \in \{1, 2, \ldots, N\}$, there exists universal constant $\nu \in (0, 1/2)$ such that if

$$|\{(x, t) \in Q_1 : w_{x_i} < \mu/2\}| \leq \nu |Q_1|,$$

holds, then we have the following conclusion:

$$w_{x_i} \geq \frac{\mu}{4} \quad \text{on} \quad Q_{1/2}^1.$$

**Proof.** Let us define the following constants:

$$k_m := k_0 - \frac{H}{8(1 + A)} \left(1 - \frac{1}{2^m}\right) \quad \text{where} \quad H := \sup_{Q_1} (w_{x_i} - k_0) \quad \text{and} \quad k_0 := \frac{\mu}{2}.$$

This choice of $k_m$ satisfy the following bounds:

$$k_m - k_{m+1} \geq \frac{\mu}{2^{m+6}(1 + A)}, \quad k_m \geq \frac{\mu}{4} \quad \text{and} \quad k_m \to k_\infty = k_0 - \frac{H}{8(1 + A)} > \frac{\mu}{4}. \hspace{1cm} (8.34)$$

Let us now define the following sequence of dyadic parabolic cylinders:

$$Q_m := Q_{\rho_m} \quad \text{where} \quad \rho_m := \frac{1}{2} + \frac{1}{2^{m+1}}.$$

Note that $Q_m \to Q_{1/2}$ and $Q_0 = Q_1$. Furthermore, let us consider the following cut-off function $\eta_m \in C^\infty(Q_m)$ with $\eta_m = 0$ on $\partial_p Q_m$ and $\eta_m \equiv 1$ on $Q_{m+1}$. The following bounds hold,

$$|\nabla^2 \eta_m| + |\nabla \eta_m|^2 + |(\eta_m)_t| \leq C(n)4^m.$$

We will split the proof of the lemma into several steps.

**Step 1:** In this step, we show that without loss of generality, we can assume $4H \geq \mu$. Suppose not, then we would have $4H < \mu$ which implies

$$\sup_{Q_1} (w_{x_i} - k_0)_- = \frac{\mu}{2} - \inf_{Q_1} w_{x_i} < \frac{\mu}{4}.$$

This says $w_{x_i} > \frac{\mu}{4}$ and the desired conclusion of Lemma 8.13 follows.

**Step 2:** In this step, we will apply Lemma 2.3 to obtain an estimate for the level sets. In order to do this, let us
define

\[ A_m := \{(x, t) \in Q_m : w_{x_t} < k_m \}, \]

and consider the following function

\[ \bar{w}_m := \begin{cases} 0 & \text{if } w_{x_t} > k_m, \\ k_m - w_{x_t} & \text{if } k_m \geq w_{x_t} > k_{m+1}, \\ k_m - k_{m+1} & \text{if } k_{m+1} \geq w_{x_t}. \end{cases} \]

Since \( \eta_m = 1 \) on \( Q_{m+1} \), we obtain the following sequence of estimates:

\[
\mu^{p-1}(k_m - k_{m+1})^2|A_{m+1}| = \mu^{p-1}\left\| \bar{w}_m \right\|_{L^2(A_{m+1})}^2 \\
\leq \mu^{p-1}\left\| \bar{w}_m \right\|_{L^2(Q_m)}^2 \\
\leq \mu^{p-1}\left\| \bar{w}_m \right\|_{L^2(Q_m)}^2 |A_m|^{\frac{2}{p-1}},
\]

where to obtain (a), we enlarged the domain and made use of the fact that \( \text{spt}(\eta_m) \subset Q_m \) and to obtain (b), we made use of [7, Corollary 3.1 - Page 9] noting that \( \bar{w}_m \eta_m \) is non-negative. This additionally also implies

\[ |\{Q_m : \bar{w}_m \eta_m > 0\}| \leq |\{Q_m : \bar{w}_m > 0\}| = |A_m|. \]

Then, observing that

\[ \bar{w}_m \leq (w_{x_t} - k_m) \quad \text{and} \quad |\nabla \bar{w}_m| \leq |\nabla (w_{x_t} - k_m)|\chi_{Q_1 \setminus \{w_{x_t} < k_{m+1}\}}, \]

and recalling Definition 2.2, we get

\[
\mu^{p-1}\left\| \bar{w}_m \eta_m \right\|_{L^2(Q_m)}^2 \leq \sup_{-1 < t < 0} \mu^{p-1} \int_{B_1} (w_{x_t} - k_m)^2 \eta_m^2 \, dx \\
+ \mu^{p-1} \int_{Q_1} |\nabla (w_{x_t} - k_m)|^2 \chi_{Q_1 \setminus \{w_{x_t} < k_{m+1}\}} \eta_m^2 \, dz \\
+ \mu^{p-1} \int_{Q_1} (w_{x_t} - k_m)^2 |\nabla \eta_m|^2 \, dz.
\]

**Step 3:** In this step, we shall estimate (8.36) and obtain a suitable decay of the level set \( A_m \) as follows: From (8.34), we see that

\[
\mu \leq 4k_{m+1} \leq 4w_{x_t} \leq 4|\nabla w| \quad \text{on} \quad Q_1 \setminus \{w_{x_t} < k_{m+1}\}. \]

Thus making use if (8.37) into (8.36), we get

\[
\mu^{p-1}\left\| \bar{w}_m \eta_m \right\|_{L^2(Q_m)}^2 \leq \sup_{-1 < t < 0} \mu^{p-1} \int_{B_1} (w_{x_t} - k_m)^2 \eta_m^2 \, dx \\
+ \int_{Q_1} (s^2 + |\nabla w|^2)^{\frac{p-2}{2}} |\nabla (w_{x_t} - k)|^2 \chi_{Q_1 \setminus \{w_{x_t} < k_{m+1}\}} \eta_m^2 \, dz \\
+ \mu^{p-1} \int_{Q_1} (w_{x_t} - k_m)^2 |\nabla \eta_m|^2 \, dz.
\]

From (8.32), we see that

\[ A \mu \int_{Q_1} (s^2 + |\nabla w|^2)^{\frac{p-2}{2}} |\nabla (w_{x_t} - k)|^2 \phi^2 \, dz \geq \int_{Q_1} (s^2 + |\nabla w|^2)^{\frac{p-2}{2}} |\nabla (w_{x_t} - k)|^2 \phi^2 \, dz. \]

Thus substituting (8.39) into (8.38) and making use of Lemma 8.11 to estimate each of the terms appearing on
the right hand side of (8.38) using (8.32), we get
\[
\begin{align*}
\mu^{p-1} \| \tilde{w}_m \eta_m \|_{V^2(Q_m)}^2 &
\leq \mu \iint_{Q_1} (s^2 + |\nabla w|^2) \tilde{w} |\nabla \eta_m|^2 \chi_{\{w_r \leq k_m\}} \, dz \\
& + \mu^p \iint_{Q_1} (w_{x_i} - k_m)_- (|\eta_m||\nabla^2 \eta_m| + |\nabla\eta_m|^2) \\
& + \mu^{p-1} \iint_{Q_1} (w_{x_i} - k_m)_+ |\nabla \eta_m|^2 \\
& \leq C \mu^{p+1} |A_m|.
\end{align*}
\]
(8.40)

Thus combining (8.40) with (8.35) and making use of (8.34), we get
\[
\mu^{p-1} \frac{\mu^2}{4^{m+6}(1 + A)^2} |A_{m+1}| \leq C A^m \mu^{p+1} |A_m|^{1 + \frac{m}{n+4}} \iff |A_{m+1}| \leq C A^m |A_m|^{1 + \frac{2}{m+n}}.
\]

We can now apply Lemma 2.6 to obtain the existence of a universal constant \( \nu \in (0, 1) \) such that if
\[
|A_0| = \{|Q_1 : \omega_r < \mu/2| < \nu|Q_1|,
\]
then \( |A_m| \to 0 \) as \( m \to \infty \). In particular, this says
\[
w_{x_i} \geq \frac{\mu}{4} \quad \text{on} \quad Q_{1/2}.
\]

This completes the proof of the lemma.

Following analogous calculations, we can also obtain the dual version of Proposition 8.12.

Proposition 8.14. Let \( r \in (0, R] \) and assume that
\[
s + \sup_{Q_r^\mu} \| \nabla u \| \leq A\mu
\]
holds for some \( A \geq 1 \). For any \( i \in \{1, 2, \ldots, N\} \), there exists universal constant \( \nu \in (0, 1/2) \) such that if
\[
|\{(x, t) \in Q_r^\mu : u_{x_i} > -\mu/2\}| \leq \nu|Q_r^\mu|,
\]
(8.41)
then we have the following conclusion:
\[
u_{x_i} \leq -\frac{\mu}{4} \quad \text{on} \quad Q_{r/2}^\mu.
\]

8.2.2. Proof of the decay estimate from applying the linear theory

Let us first recall the weak Harnack inequality and a decay estimate that will be needed to complete the proof of Proposition 8.1, the details can be found in [12, Lemma 3.1].

Lemma 8.15. Let \( v \in L^2(-1, 1; W^{1,2}(B_1)) \) be a weak solution to the linear parabolic equation
\[
v_t - \text{div}(B(x, t) \nabla v) = 0,
\]
where the matrix \( B(x, t) \) is bounded, measurable and satisfies
\[
C_{0} |\zeta|^2 \leq \langle B(x, t) \zeta , \zeta \rangle \quad \text{and} \quad |B(x, t)| \leq C_1,
\]
for any \( \zeta \in \mathbb{R}^N \) and \( 0 < C_0 \leq C_1 \) are fixed constants. Then there exists a constant \( C = C(N, C_0, C_1) \geq 1 \) and \( \beta = \beta(N, C_0, C_1) \in (0, 1) \) such that the following estimates are satisfied:
\[
\sup_{Q_{r/2}} |v| \leq C \left( \iint_{Q_1^\mu} |v|^q \, dz \right)^{\frac{1}{q}} \quad \text{for any} \ q \in [1, 2],
\]
\[
\left( \iint_{Q_s} |v - (v)_{Q_s}|^q \, dz \right)^{\frac{1}{q}} \leq C \delta^\beta \left( \iint_{Q_1^\mu} |v - (v)_{Q_1}|^q \, dz \right)^{\frac{1}{q}} \quad \text{whenever} \ q \in [1, 2] \text{ and} \ \delta \in (0, 1).
\]
We now have all the estimates needed to prove Proposition 8.1 which will be given the following lemma. The proof follows very closely to [12, Lemma 3.2] and hence we will only present the rough sketch of the proof.

Lemma 8.16. Assume that in the cylinder \( Q_r^p \), the following is satisfied for a fixed \( A \geq 1 \):

\[
0 < \frac{\mu}{4} \leq \| \nabla u \|_{L^\infty(Q_r^p)} \leq s + \| \nabla u \|_{L^\infty(Q_r^p)} \leq A \mu.
\]

Then there exists constants \( \beta = \beta(N, p, C_0, C_1, A) \in (0, 1) \) and \( C = C(N, p, C_0, C_1, A) \geq 1 \) such that the following holds for any \( \delta \in (0, 1) \) and \( q \geq 1 \):

\[
\left( \iint_{Q_{r/2}^p} |\nabla u - (\nabla u)_{Q_r^p}|^q \, dz \right)^{\frac{1}{q}} \leq C \delta^\beta \left( \iint_{Q_r^p} |\nabla u - (\nabla u)_{Q_r^p}|^q \, dz \right)^{\frac{1}{q}}.
\]

Proof. Without loss of generality, let us assume \( \delta \in (0, 1/2) \) noting that when \( \delta \in [1/2, 1) \), we can enlarge the domain of integration (see [12, Lemma 3.2 and Proposition 3.3] for the details) and obtain the desired conclusion, albeit with a larger constant, the details of which is given in Proposition 8.17. As in the proof of Lemma 8.13, we rescale according to (8.26), then the hypothesis becomes

\[
0 < \frac{\mu}{4} \leq \| \nabla w \|_{L^\infty(Q_1)} \leq s + \| \nabla w \|_{L^\infty(Q_1)} \leq A \mu.
\] (8.42)

Differentiating (8.27) and dividing the resulting equation by \( \mu^{p-1} \), we see that \( w_{x_i} \) solves

\[
(w_{x_i})_t - \text{div}(B(x, t)\nabla w_{x_i}) = 0 \quad \text{where} \quad B(x, t) = \mu^{2-p} \partial A(\nabla w).
\]

Furthermore, making use of (8.2) along with (8.42), we see that the matrix \( B(x, t) \) satisfies the following bounds:

\[
C|\zeta|^2 \leq \langle B(x, t)\zeta, \zeta \rangle \leq C \left( \frac{s + \mu}{\mu} \right)^{p-2} |\zeta|^2 \leq C |\zeta|^2,
\]

(8.43)

for any \( \zeta \in \mathbb{R}^N \) and \( C = C(N, p, C_0, C_1, A) \). Thus we can apply Lemma 8.15 to get the desired conclusion whenever \( q \in [1, 2] \). On the other hand, if \( q \geq 2 \), then we can follow the calculations from [12, Equations (3.33) and (3.34)] to get the desired conclusion. This completes the proof of the lemma. \( \square \)

8.2.3. Summary of the proof of Proposition 8.1

Collecting all the calculations from the previous subsections, we have the following proposition, the proof of which follows verbatim as in [12, Proposition 3.3].

Proposition 8.17. Assume that (8.31) is in force, then there exists \( \nu = \nu(N, p, C_0, C_1, A) \in (0, 1/2) \) such that if there exists \( i \in \{1, 2, \ldots, N\} \) such that either (8.33) or (8.41) holds, then there exists \( \beta = \beta(N, p, C_0, C_1, A) \in (0, 1) \) and \( C = C(N, p, C_0, C_1, A) \) such that for any \( \delta \in (0, 1) \), the following conclusions hold:

\[
\left( \iint_{Q_{r/2}^p} |\nabla u - (\nabla u)_{Q_r^p}|^q \, dz \right)^{\frac{1}{q}} \leq C \delta^\beta \left( \iint_{Q_r^p} |\nabla u - (\nabla u)_{Q_r^p}|^q \, dz \right)^{\frac{1}{q}},
\]

\[
\| \nabla u \| \geq \frac{\mu}{4} \quad \text{on} \quad Q_{r/2}^\mu.
\]

8.3. Proof of second alternative - Proposition 8.2

Having fixed \( \nu \) according to Proposition 8.1, we are now in the situation that (8.33) and (8.41) does not hold. This condition is equivalent to

\[
|\{Q_r^\mu : u_{x_i} \geq \mu/2\}| < (1 - \nu)|Q_r^\mu| \quad \text{and} \quad |\{Q_r^\mu : u_{x_i} \leq -\mu/2\}| < (1 - \nu)|Q_r^\mu|.
\]

(8.44)
From (8.3), as in (8.31), we again assume the following is always satisfied for some \( r \in (0, R) \) and \( A \geq 1 \):

\[
s + \sup_{Q_r} \|\nabla u\| \leq A\mu. \tag{8.45}
\]

Let us perform the following change of variables:

\[
w(x, t) = \frac{u(\mu^{-1}rx, \mu^{-p}r^2t)}{rA} \quad \text{for} \quad (x, t) \in Q_1. \tag{8.46}
\]

Then we have

\[
\frac{s}{\mu A} + \sup_{Q_1} \|\nabla w\| \leq 1 \quad \text{on} \quad Q_1. \tag{8.47}
\]

Moreover, (8.44) becomes

\[
\left| \left\{ Q_1 : w_{x_i} \geq \frac{1}{2A} \right\} \right| < (1 - \nu)|Q_{r_1}| \quad \text{and} \quad \left| \left\{ Q_1 : w_{x_i} \leq -\frac{1}{2A} \right\} \right| < (1 - \nu)|Q_{r_1}|. \tag{8.48}
\]

### 8.3.1. Choosing a good time slice

First let us show there exists a good time slice.

**Lemma 8.18.** There exists \( t_* \) such that \( -1 \leq t_* \leq -\frac{\nu}{2} \) and

\[
\left| \left\{ B_1 : w_{x_i}(x, t_*) \geq \frac{1}{2A} \right\} \right| \leq \left( \frac{1 - \nu}{1 - \nu/2} \right) |B_1| \tag{8.49}
\]

**Proof.** The proof is by contradiction, suppose (8.49) does not hold for any \( t \in (-1, -\frac{\nu}{2}) \), then we have

\[
(1 - \nu)|Q_1| > \left| \left\{ Q_1 : w_{x_i} \geq \frac{1}{2A} \right\} \right| = \int_{-1}^{-\frac{\nu}{2}} \left| \left\{ B_1 : w_{x_i}(x, t) \geq \frac{1}{2A} \right\} \right| dt \tag{8.49} \]

which is a contradiction. \( \square \)

### 8.3.2. Rescaling the equation

Recalling (8.46), let us define

\[
\hat{A}(\zeta) := \frac{1}{\mu^{\rho-1}A} A(\mu \zeta), \tag{8.50}
\]

then, we see that \( w \) solves

\[
w_t - \text{div} \hat{A}(\nabla w) = 0 \quad \text{in} \quad Q_1. \tag{8.51}
\]}
Lemma 8.19. Differentiating \((8.51)\) with respect to \(x_i\) for some \(i \in \{1, 2, \ldots, N\}\), we get
\[
(w_{x_i})_t - \text{div} \hat{A}(\nabla w) \nabla w_{x_i} = 0 \quad \text{in} \quad Q_1. \tag{8.52}
\]
Then \(\hat{A}(\xi)\) and \(\hat{A}(\xi)\) satisfies the following structure conditions:
\[
|\hat{A}(\xi)| + |\partial \hat{A}(\xi)| \left( |\xi|^2 + \left( \frac{s}{\lambda} \right)^2 \right)^{\frac{p-1}{2}} \leq C_1 A^{p-2} \left( |\xi|^2 + \left( \frac{s}{\lambda} \right)^2 \right)^{\frac{p-1}{2}},
\]
\[
C_0 A^{p-2} \left( |\xi|^2 + \left( \frac{s}{\lambda} \right)^2 \right)^{\frac{p-2}{2}} |\eta|^2 \leq \langle \partial \hat{A}(\xi) \eta, \eta \rangle. \tag{8.53}
\]
Proof. The proof follows directly from (8.2) and (8.50).

8.3.3. Logarithmic and Energy estimates

Lemma 8.20. Let \(0 < \eta_0 < \nu\) and \(k \geq \frac{1}{4A}\) be any two fixed numbers (recall \(\nu\) is from Proposition 8.2) and consider the function
\[
\Psi(z) := \log^+ \left( \frac{\nu}{\nu - (z - (1 - \nu) + \eta_0)} \right).
\]
Then there exists constant \(C = C(N, p, C_0, C_1, A)\) such that for all \(t_1, t_2 \in (-1, 1)\) with \(t_1 < t_2\) and all \(s \in (0, 1)\), there holds
\[
\int_{B_s \times \{t_2\}} \Psi^2((w_{x_i} - k)_{+}) \, dx \leq \int_{B_s \times \{t_1\}} \Psi^2((w_{x_i} - k)_{+}) \, dx + \frac{C}{(1-s)^2} \int_{B_s \times \{t_1, t_2\}} \Psi((w_{x_i} - k)_{+}) \, dz. \quad \text{(8.54)}
\]

Proof. Let us take \(\Psi((w_{x_i} - k)_{+})\Psi'((w_{x_i} - k)_{+})\xi^2\) as a test function in (8.51). Since \(k \geq \frac{1}{4A}\), we see that \((w_{x_i} - k)_{+} = 0\) whenever \(w_{x_i} \leq \frac{1}{4A}\). On this set, \(\Psi(0) = \log^+ \left( \frac{\nu}{\nu + \eta_0} \right) = 0\), thus we see that \(\text{spt} \Psi((w_{x_i} - k)_{+})\) is contained in the set \(\left\{ w_{x_i} \geq \frac{1}{4A} \right\}\). Thus making use of (8.47), we make the following observation,
\[
\frac{1}{A^{p-1}4^{p-1}} \left( |\nabla w|^2 + \left( \frac{\nu}{\lambda} \right)^2 \right)^{\frac{p-1}{2}} \leq 4A \quad \text{on} \quad \left\{ w_{x_i} \geq \frac{1}{4A} \right\}. \tag{8.55}
\]
Substituting (8.55) into (8.53), we see that \(\partial \hat{A}(\nabla w)\) is uniformly elliptic and we rewrite (8.52) as follows:
\[
(w_{x_i})_t - \text{div} B(x, t) \nabla w_{x_i} = 0, \quad \text{with} \quad \frac{C_0}{A4^{p-1}4^{p-1}} |\xi|^2 \leq (B(x, t) \xi, \xi) \leq 4C_1 A^{p-1} |\xi|^2. \tag{8.56}
\]
We can now follow the proof of [7, Proposition 3.2 of Chapter II] (see also [7, (12.7) of Chapter IX]) to get
\[
\int_{B_s \times \{t_2\}} \Psi^2((w_{x_i} - k)_{+}) \, dx \leq \int_{B_s \times \{t_1\}} \Psi^2((w_{x_i} - k)_{+}) \, dx + \frac{C(N, p, C_0, C_1, A)}{(1-s)^2} \int_{B_s \times \{t_1, t_2\}} \Psi((w_{x_i} - k)_{+}) \, dz. \quad \text{(8.54)}
\]

Lemma 8.21. Let \(k \geq \frac{1}{4A}\) be some fixed constant, then \(w\) solving (8.51) satisfies
\[
\sup_{t_0 < t < t_1} \int_{B_1} (w_{x_i} - k)_{+}^2 \xi^2 \, dx + C \int_{B_1 \times \{t_0, t_1\}} |\nabla (w_{x_i} - k)_{+}|^2 \xi^2 \, dz \leq \int_{B_1 \times \{t_0\}} (w_{x_i} - k)_{+}^2 \xi^2 \, dx + C \int_{B_1 \times \{t_0, t_1, t_2\}} (w_{x_i} - k)_{+}^2 |\nabla \xi|^2 \, dz
\]
\[
+ C \int_{B_1 \times \{t_0, t_2\}} (w_{x_i} - k)_{+}^2 |\xi|^2 \, dz
\]
where \(t_0, t_1 \in [-1, 1]\) with \(t_0 < t_1\) are some fixed time slices and \(\xi \in C^\infty(Q_1)\) is a cut-off function.
Proof. We take \((w_{x_i} - k_+)\) as a test function in (8.51), noting that we retain the uniformly elliptic structure of \(B(x, t)\) in (8.56) from the choice of test function. Thus proceeding according to [7, Proposition 3.1 of Chapter II] gives the desired estimate. \(\square\)

8.3.4. DeGiorgi type iteration

Let us first prove an expansion of positivity in time result:

**Lemma 8.22.** There exists \(\eta_0 = \eta_0(N, p, \nu, A) \in (0, \nu)\) such that for all \(t \in (t_*, 1)\) with \(t_*\) as in Lemma 8.18, there holds

\[
|\{x \in B_1 : w_{x_i}(x, t) > (1 - \eta_0)\}| \leq \left(1 - \frac{\nu^2}{4}\right)|B_1|.
\]

**Proof.** Let us apply (8.54) over the time interval \((t_*, t)\) and estimate each of the terms as follows:

**Estimate for the term on the left hand side of (8.54):** On the set \(\{w_{x_i} > (1 - \eta_0)\}\), we see that the log function satisfies \(\Psi(w_{x_i} - k_+) \geq \log \left(\frac{\nu}{2\eta_0}\right)\), thus we get

\[
\int_{B_1 \times \{(t_2)\}} \Psi^2((w_{x_i} - k_+^2) dx \geq \log^2 \left(\frac{\nu}{2\eta_0}\right) |\{\{B_1 : w_{x_i}(x, t) \geq 1/(4A)\}\} = \left(1 - \nu^2/2\right) \log^2 \left(\frac{\nu}{\eta_0}\right) |B_1|.
\]

**Estimate for the first term on the right hand side of (8.54):** Since the log function \(\Psi(w_{x_i} - k_+)\) vanishes on the set \(\{w_{x_i} \leq 1/(4A)\}\), we estimate this term as follows:

\[
\int_{B_1 \times \{t = t_*\}} \Psi(w_{x_i} - k_+^2) dx \leq \log^2 \left(\frac{\nu}{\eta_0}\right) |\{B_1 : w_{x_i}(x, t) \geq 1/(4A)\}| \leq \left(1 - \nu^2/2\right) \log^2 \left(\frac{\nu}{\eta_0}\right) |B_1|.
\]

**Estimate for the second term on the right hand side of (8.54):** Analogously, we estimate this term as follows:

\[
\frac{C}{(1-s)^2} \int_{B_1 \times \{(t_2)\}} \Psi((w_{x_i} - k_+) dz \leq \frac{C}{(1-s)^2} |B_1| \log \left(\frac{\nu}{\eta_0}\right).
\]

Combining (8.57), (8.58) and (8.59), we get the following sequence of estimates:

\[
|\{B_1 : w_{x_i}(x, t) > (1 - \eta_0)\}| \leq |\{B_1 : w_{x_i}(x, t) > (1 - \eta_0)\}| + (1 - s)|B_1|
\]

\[
\leq \left(1 - \frac{\nu^2}{1 - \nu^2/2}\right) \log^2 \left(\frac{\nu}{\eta_0}\right) + \frac{C}{(1-s)^2} \log^2 \left(\frac{\nu}{\eta_0}\right) + (1 - s) |B_1|.
\]

Choosing \(s\) followed by \(\eta_0\) appropriately gives the desired conclusion. \(\square\)

Given Lemma 8.22, the rest of the proof of the second alternative Proposition 8.2 is as in the linear case. We nevertheless provide details for the sake of completeness.

**Definition 8.23.** Having determined \(\eta_0\) as in Lemma 8.22, let \(j_0\) be the largest positive integer such that \(2^{-j_0} \geq \eta_0\). Then for \(j \geq j_0\) and any \(t \in (t_*, 1)\), let us define

\[
A_j(t) := \{B_1 : w_{x_i}(x, t) > (1 - 2^{-j})\} \quad \text{and} \quad A_j^s := \int_{t_*}^s A_j(t) dt,
\]

for some \(s \in (t_*, 1)\) with \(s - t_* \geq \frac{1}{8}\).
Lemma 8.24. For every $\delta \in (0,1)$, there exists $j_\delta \geq j_0$ such that for any $s \in (t_\ast,1]$ with $s - t_\ast \geq \frac{1}{8}$, the following holds:

$$A_{j_\delta}^s \leq \delta |B_1||s - t_\ast|.$$  

Note that the estimates and choice of $j_\delta$ are all independent of the choice of $s$.

Proof. From Lemma 8.22, we see that

$$|B_1 \setminus A_j(t)| \geq \left( \frac{\nu}{2} \right)^2 |B_1| \quad \text{for any} \quad t \in (t_\ast,1].$$  

Let us apply Lemma 2.5 with the choices $l = 1 - \frac{1}{2^{j+1}}$ and $k = 1 - \frac{1}{2^j}$ for some $j \geq j_0$ where $j_0$ is as defined in Definition 8.23 to get

$$\frac{1}{2^{j+1}} |A_{j+1}(t)| \leq \frac{C}{|B_1 \setminus A_j(t)|} \int_{A_j(t) \setminus A_{j+1}(t)} |\nabla w_{x_1}| \, dx \leq C_{(N,p,\nu)} \left( \int_{B_1} |\nabla (w_{x_1} - (1 - 2^{-j}))|^2 \, dx \right)^{\frac{1}{2}} |A_j(t) \setminus A_{j+1}(t)|^{\frac{1}{2}}.$$  

Integrating over $(t_\ast,s)$, we get

$$\frac{1}{2^{j+1}} |A_{j+1}(t)| \leq C_{(N,p,\nu)} \left( \int_{B_1 \times (t_\ast,s)} |\nabla (w_{x_1} - (1 - 2^{-j}))|^2 \, dz \right)^{\frac{1}{2}} |A_j(t) \setminus A_{j+1}(t)|^{\frac{1}{2}}.$$  

From (8.47) and the choice of $k = 1 - \frac{1}{2^j}$, we get

$$(w_{x_1} - k)_+ \leq \frac{1}{2^j},$$  

using which, we now estimate the gradient term on the right hand side of (8.61) using Lemma 8.21 to get

$$\int_{B_1 \times (t_\ast,s)} |\nabla (w_{x_1} - (1 - 2^{-j}))|^2 \, dz \leq C \frac{1}{s - t_\ast} |B_1||s - t_\ast|,$$  

where we used the fact that $|\zeta_t| \leq \frac{C}{s - t_\ast} \leq C$. Substituting (8.62) into (8.61) and summing over $j = j_0, j_0 + 1, \ldots, j_\delta$, we get

$$(j_\delta - j_0) |A_{j_\delta}^s|^2 \leq C |B_1||s - t_\ast| \sum_{j = j_0}^{j_\delta} |A_j^s \setminus A_{j+1}^s| \leq C \left( |B_1||s - t_\ast| \right)^2.$$  

Choosing $j_\delta$ sufficiently large followed by taking square roots gives the desired conclusion. 

8.3.5. Combining all the estimates to prove Proposition 8.2

Let us prove the following result by applying DeGiorgi iteration, using which we easily conclude Proposition 8.2.

Proposition 8.25. Suppose the first (analogously second) inequality in (8.48) holds, then there exist positive constant $\eta = \eta(N,p,C_0,C_1,A) \in (0,1)$ such that the following conclusion holds:

$$\left| \left\{ Q_{\frac{1}{2}} : w_{x_1} > (1 - \eta) \right\} \right| = 0 \quad \text{analogously} \quad \left| \left\{ Q_{\frac{1}{2}} : w_{x_1} < (1 - \eta) \right\} \right| = 0.$$  

Proof. Consider the family of nested cylinders

$$Q_n := B_{\rho_n} \times (-\rho_n^2, \rho_n^2) \quad \text{where} \quad \rho_n := \frac{1}{2} + \frac{1}{2^{n+2}} \quad \text{with} \quad n = 0, 1, 2, \ldots,$$

and increasing intervals

$$k_n := 1 - \frac{1}{2^{j_\ast+2}} - \frac{1}{2^{j_\ast+2+n}},$$  

32
where \( j_\star \) is a number to be eventually chosen. Furthermore, define

\[
Y_n := |\{Q_n : w_{x_i} > k_n\}|
\]

From Lemma 8.21 and Lemma 2.3 applied with \( \tilde{p} = 2 \), we get the following sequence of estimates (note that \( \zeta_n \) is the usual cut-off function with \( \zeta_n \equiv 1 \) on \( Q_{n+1} \) and \( spt(\zeta_n) \subset Q_n \)):

\[
\int_{Q_{n+1}} (w_{x_i} - k_n)_+^2 \, dz \overset{(a)}{\leq} \int_{Q_n} (w_{x_i} - k_n)_+^2 \zeta_n^2 \, dz \overset{(b)}{\leq} \left( \int_{Q_n} |(w_{x_i} - k_n)_+ + \zeta_n| \, dz \right)^{\frac{N}{N+2}} Y_n^{\frac{2}{N+2}} \overset{(c)}{\leq} \| (w_{x_i} - k_n)_+^2 \zeta_n \|_{V^{2,2}(Q_n)}^2 Y_n^{\frac{2}{N+2}} \overset{(d)}{\leq} C 4^{j_\star} Y_n^{1+\frac{2}{N+2}},
\]

where to obtain (a), we enlarged the domain of integration noting that \( \zeta_n \equiv 1 \) on \( Q_{n+1} \), to obtain (b), we applied H"older inequality and made use of the definition of \( Y_n \), to obtain (c), we applied Lemma 2.3 with \( \tilde{p} = 2 \) and \( \tilde{q} = \frac{2(N+2)}{N} \) and finally to obtain (d), we made use of the definition of \( k_n \) along with (8.47).

On the other hand, from (4.11), we also have

\[
Y_{n+1} \leq C 4^{n+j_\star} \int_{Q_{n+1}} (w_{x_i} - k_n)_+^2 \, dz.
\]

Combining (8.64) and (8.63), we get

\[
Y_{n+1} \leq C 4^{n} Y_n^{1+\frac{2}{N+2}},
\]

and applying Lemma 2.6, we see that if

\[
Y_0 \leq C \frac{N+2}{2} \left( \frac{N+2}{2} \right)^2 := \delta,
\]

then \( Y_n \to 0 \) as \( n \to \infty \).

Note that \( \delta = \delta(\{C_0, C_1, N, p\}) \), thus with this choice of \( \delta \), we can obtain \( j_\delta \) (and set \( j_\star = j_\delta \)) from Lemma 8.24 such that the condition for \( Y_0 \) is satisfied and thus \( Y_\infty = 0 \) which is the desired conclusion with \( \eta = 2^{-(j_\star+2)} \).

References

[1] Karthik Adimurthi, Sun-Sig Byun, and Jehan Oh. Interior and boundary higher integrability of very weak solutions for quasilinear parabolic equations with variable exponents. *Nonlinear Anal.*, 194:111370, 54, 2020.

[2] Karthik Adimurthi, Sun-Sig Byun, and Jung-Tae Park. End point gradient estimates for quasilinear parabolic equations with variable exponent growth on nonsmooth domains. *arXiv preprint arXiv:1806.00944*, 2018.

[3] Karthik Adimurthi and Sukjung Hwang. Uniform boundedness for weak solutions of quasilinear parabolic equations. *Proc. Amer. Math. Soc.*, 148(2):653–665, 2020.

[4] Hi Jun Choe. Hölder regularity for the gradient of solutions of certain singular parabolic systems. *Comm. Partial Differential Equations*, 16(11):1709–1732, 1991.
[5] Cristiana De Filippis. Gradient bounds for solutions to irregular parabolic equations with \((p,q)\)-growth. *arXiv preprint arXiv:2004.01452*, 2020.

[6] E. DiBenedetto. \(C^{1+\alpha}\) local regularity of weak solutions of degenerate elliptic equations. *Nonlinear Anal.*, 7(8):827–850, 1983.

[7] Emmanuele DiBenedetto. *Degenerate parabolic equations*. Universitext. Springer-Verlag, New York, 1993. 2, 3, 4, 6, 7, 8, 9, 10, 12, 26, 30, 31

[8] Emmanuele DiBenedetto and Avner Friedman. Regularity of solutions of nonlinear degenerate parabolic systems. *J. Reine Angew. Math.*, 349:83–128, 1984.

[9] Cyril Imbert, Tianling Jin, and Luis Silvestre. Hölder gradient estimates for a class of singular or degenerate parabolic equations. *Adv. Nonlinear Anal.*, 8(1):845–867, 2019.

[10] Juha Kinnunen and John L. Lewis. Higher integrability for parabolic systems of \(p\)-Laplacian type. *Duke Math. J.*, 102(2):253–271, 2000.

[11] Juha Kinnunen and John L. Lewis. Very weak solutions of parabolic systems of \(p\)-Laplacian type. *Ark. Mat.*, 40(1):105–132, 2002.

[12] Tuomo Kuusi and Giuseppe Mingione. The Wolff gradient bound for degenerate parabolic equations. *J. Eur. Math. Soc. (JEMS)*, 16(4):835–892, 2014.

[13] John L. Lewis. Regularity of the derivatives of solutions to certain degenerate elliptic equations. *Indiana Univ. Math. J.*, 32(6):849–858, 1983.

[14] Peter Tolksdorf. Regularity for a more general class of quasilinear elliptic equations. *J. Differential Equations*, 51(1):126–150, 1984.